Comprehending B-Cell Epitope Prediction to Develop Vaccines and Immunodiagnostics

Salvador Eugenio C. Caoili*

Biomedical Innovations Research for Translational Health Science (BIRTHS) Laboratory, Department of Biochemistry and Molecular Biology, College of Medicine, University of the Philippines Manila, Manila, Philippines

Keywords: B-cell epitope prediction, B-cell epitopes, surface accessibility, conformational disorder, immunodominance, peptide-based vaccines, antipeptide antibodies, protective immunity

1 INTRODUCTION

B-cell epitope prediction (BCEP) is the original subject of immunoinformatics (i.e., bioinformatics applied to immunology). This began with protein sequence analysis to identify hydrophilic peptide fragments bound by antibodies that recognize whole proteins (1), to enable the earlier proposed development of synthetic peptide-based vaccines for inducing protective antibody-mediated immunity (2). BCEP was thus initially "prediction of protein antigenic determinants," with each antigenic determinant being a B-cell epitope (BCE): a structural feature (e.g., sequence segment) recognized by a paratope (i.e., antigen-binding site of an immunoglobulin such as an antibody) (3, 4). Now understood as the computational identification of putative BCEs, BCEP has since grown to comprise much more sophisticated methods for analyzing both sequence (5–9) and higher-order structure (10, 11) on ever larger scales (e.g., applying genomics and proteomics for vaccine design (12, 13)). However, the full potential of BCEP for peptide-based vaccine design remains to be realized, for which reason the utility of BCEP as such has been called into question (14–17). Nevertheless, BCEP can support the development of vaccines and immunodiagnostics provided that its limitations are adequately comprehended and addressed.

2 ACCESSIBLE DISORDER (AD) AND IMMUNODOMINANCE

Accessible disorder (AD) is the state of a BCE that is simultaneously both paratope-accessible and disordered (i.e., conformationally unconstrained), such that BCE-paratope binding can occur via induced fit and/or conformational selection (8). Clearly, BCEs must be accessible to paratopes for physical contact to occur between them. In the context of antibody-mediated protective (e.g., antipathogen) immunity, this is most readily feasible for BCEs that are on outwardly protruding solvent-accessible molecular surfaces at extracellular sites (e.g., among secreted biomolecules). Hence, vaccine-design initiatives tend to selectively apply BCEP with a focus on surface-exposed sites among biomolecular targets of the pertinent (e.g., pathogen) secretome (i.e., totality of secreted biomolecules) and surfome (i.e., surface proteome) (18, 19). Additionally, surface complementarity must be attained between BCEs and their corresponding paratopes upon physical contact if stable BCE-paratope binding is to occur, for which the BCEs must adopt suitable conformations. This is favored where the BCEs are disordered prior to binding by the paratopes (20, 21). Vaccines can thus
be produced from peptides comprising disordered BCEs of selected target proteins (e.g., pathogen virulence factors), to elicit production of antipeptide antibodies (i.e., peptide antibodies (22)) that can neutralize the biological activity of the proteins (4, 23–36). Such an approach is viable where the BCEs are disordered in both the peptides and the proteins; but if the BCEs are conformationally constrained (e.g., folded) in the proteins, their binding by the antipeptide antibodies may fail to occur, as is thought to be the case among unsuccessful attempts at peptide-based vaccine development (16).

AD among BCEs thus facilitates BCE-paratope binding; but BCE-specific antibody production is also subject to the phenomenon of immunodominance (i.e., bias of immune responses toward subsets of BCEs encountered in the course of immunization), as depicted in Figure 1. Driven by Darwinian competition among B-cell clones, immunodominance tends to be favored by greater numbers of functional BCE-recognizing precursor B cells as well as stronger binding of BCEs by B cells in terms of both affinity (i.e., strength of binding per individual BCE-paratope pairwise interaction) and avidity (i.e., overall strength of cooperative binding among paratopes that simultaneously bind two or more BCEs on a single antigen, as is possible with engagement of one or more bivalent immunoglobulin molecules) (37). Consequently, individual host life history of antigenic exposure (e.g., via infection and immunization) influences immunodominance. Immunodominance may thus be precluded by immune tolerance (i.e., selective inability to mount immune responses to particular BCEs, due to functional deletion or inactivation of their corresponding B cells), which is often induced by BCEs of host self antigens (i.e., autoantigens) and of other antigens (e.g., in food) to which the host has been exposed in a natural physiologic setting (rather than in the course of infectious disease or vaccination) (38–40). Alternatively, immunodominance may be heightened by the immunological memory of prior immunization (e.g., via infection or vaccination), as occurs in the phenomenon of original antigenic sin (i.e., antigenic imprinting) whereby memory B-cell clones generated by past immunization continue to dominate antibody responses to more recent immunizations, possibly even compromising the ability to mount protective immune responses against newly encountered pathogen variants (41, 42).

From an evolutionary standpoint, pathogens may co-evolve with their hosts to evade immune destruction in part by altering their BCE repertoires to limit the expression of immunodominant pathogen BCEs on key virulence factors (e.g., via molecular mimicry, with pathogen BCEs tending to resemble host self BCEs) while possibly also expressing immunodominant pathogen BCEs that serve as antigenic decoys to detract from protective host immune responses (43). Furthermore, immune tolerance may be broken under certain circumstances (e.g., infection by a pathogen employing molecular mimicry), which may result in antibody-mediated (e.g., autoimmune) disease (44). These various scenarios highlight the potential complexity of vaccine development with the diversity of BCEs and possible immune responses thereto. Peptide-based vaccine design thus provides opportunities to systematically restrict the repertoire of vaccine BCEs and thereby selectively target key biomolecules (e.g., critical virulence factors) while avoiding harmful or otherwise counterproductive antibody responses (e.g., to BCEs of autoantigens and antigenic decoys).

**FIGURE 1** | Identification of plausible candidate vaccine peptide BCEs. Accessible disorder (AD) is realized for BCEs that are simultaneously both paratope-accessible and disordered (i.e., conformationally unconstrained) in both peptide-based immunogens and cognate native antigenic targets in situ (e.g., extracellular pathogen virulence factors). Immunodominant BCEs are identified empirically as they occur in peptide-based immunogens (e.g., using immunogenic carrier molecules and immunologic adjuvants) versus other contexts (e.g., in native antigenic targets comprising antigenic decoys).
3 TOWARD NEW VACCINES AND IMMUNODIAGNOSTICS

In essence, BCEP consists of two steps: structural partitioning of a prospective target (e.g., protein) into plausible candidate BCEs (e.g., peptidic sequences) and evaluation of these to assign them numerical scores that can inform subsequent decisions (e.g., on selecting components for inclusion in vaccines). Ideally, the scores would directly quantify functional impact (e.g., degree of antibody-mediated host protection against a protein toxin). In practice, functional impact can be estimated from BCE-paratope binding affinity in conjunction with a limited set of other key parameters (e.g., concentrations of antibody and its target), with said affinity itself being estimated as the BCE-paratope standard free-energy change of binding ($\Delta G_{b\circ}$) based on target structure (8). This can be simplified by considering only candidate BCEs of paratope-accessible disordered target regions. Viewed more comprehensively, BCEP can be performed largely by excluding target regions that are at least partially inaccessible to paratopes. Such regions may be inaccessible due to biomolecular folding, complex formation (e.g., oligomerization) and anatomic compartmentalization (e.g., due to biomembranes). Accordingly, target regions may be identified for exclusion using appropriate computational tools to predict folding (vis-à-vis disorder, e.g., using AlphaFold (45, 46)) and higher-order structural organization (e.g., among supramolecular assemblies such as biomembranes (47)). For convenience, inaccessibility may be generalized to also include forms of posttranslational modification resulting from covalent linkage of nonprotein moieties (e.g., glycosylation), which can be computationally predicted to mark candidate BCEs for exclusion as well. Likewise, generalization of inaccessibility can also be extended to regions featuring disulfide-bond formation between cysteine residues, as this may impede BCE-paratope binding. After exclusion of implausible or otherwise potentially problematic target regions based on anticipated inaccessibility, $\Delta G_{b\circ}$ and in turn functional impact may be estimated for the remaining candidate BCEs.

For the most part, BCEP can thus be regarded as prediction of AD. This varies with the envisioned practical application (e.g., vaccines versus immunodiagnostics). For vaccines, AD is defined by what can be achieved in vivo based on endogenous antibody production, with intracellular targets tending to be inaccessible under physiologic conditions, though antibodies are sometimes internalized by host cells in either free or pathogen-bound form to mediate immunity within certain intracellular compartments (48, 49). Such limitation may be overcome via immunotherapeutics using exogenously supplied antibodies and derivatives thereof (e.g., antibody fragments), notably with artificially produced cell-penetrating antibodies that can cross plasma membranes to bind intracellular targets (50, 51). For immunodiagnostics, the potential extent of AD is even greater, as constraints in vivo (e.g., on temperature and chemical composition) can be transcended in vitro (e.g., within a diagnostic test kit). For instance, membrane permeabilization (e.g., using detergents and/or organic solvents (52)) can extend paratope access into virtually all intracellular compartments; and treatment with chaotropic agents (e.g., urea (53) or ammonium thiocyanate (54)) can disrupt both intermolecular and intramolecular interactions, thereby increasing AD via order-to-disorder transitions, as in biomolecular disassembly and unfolding. BCEP itself is expanding in scope to support other applications, notably the design of novel antibodies and related constructs. BCEs can thus be identified initially as targets for binding by complementary peptides, which may be computationally designed and subsequently grafted on an antibody scaffold to produce novel antibodies that recognize the BCEs (55). The complementary peptide sequences might themselves be disordered, possibly remaining so on the antibody scaffold and even after binding their targets; but such persistent disorder could pose challenges for biotechnological antibody production, particularly with disordered regions that are recognized to initiate proteasome-mediated degradation (56). Nevertheless, said degradation could be circumvented via cell-free antibody production, which can also be used to produce cell-penetrating antibodies (57). Additionally, preexisting (e.g., antihapten) antibodies could be used with a diverse repertoire of peptide-based adaptors (e.g., hapten-labeled complementary peptides) for chemically programmable immunity that entails redirection of the antibodies to particular targets according to the specific choice of adaptors (58). Such repurposing of preexisting antibodies could enable more rapid responses to emerging threats (e.g., novel pathogens) than would be possible via development of novel antibodies or similar constructs.

Beyond predicting AD, BCEP also encompasses the more computationally challenging task of predicting conformational BCEs (i.e., BCEs that are to at least some extent conformationally constrained). Among folded proteins, conformational BCEs are surface-exposed regions that each constitute a paratope footprint and are thus typically discontinuous BCEs in the sense of comprising paratope-contacting residues that are noncontiguous along the protein sequence, sometimes even on separate polypeptide chains where proteins form supramolecular complexes such as viral capsids (3). Whereas prediction of AD can be cast as sequence profiling that is reminiscent of the earliest BCEP methods, BCEP for conformational BCEs entails the nontrivial cascading problems of delineating discontinuous candidate BCEs, predicting their hierarchy of immunodominance and estimating $\Delta G_{b\circ}$ for cross-reactive binding of their corresponding paratopes to disordered candidate BCEs (e.g., peptidic sequences forming parts of discontinuous candidate BCEs) (59). This is further complicated by possible protein unfolding in vivo and consequent uncertainty regarding relevant candidate BCEs, which confounds interpretation of data on binding of antiprotein antibodies to peptide fragments of the cognate protein antigens (60). In light of these considerations, the apparently poor performance of BCEP methods benchmarked against said data (15, 61) is unsurprising and likely reflects the unmitigated complexity of factors underlying BCEP to identify short peptide sequences that are recognized by antiprotein antibodies. Yet, such challenges pose barriers to development...
of peptide-based constructs as immunodiagnostics for antibody responses to vaccination, and as means for generating protective immunity (e.g., via active or passive immunization), may serve as immunodiagnostics reagents for antigen detection.

4 CONCLUSION

BCEP can be framed mainly as computational identification of putative paratope-accessible disordered peptidic sequences in an appropriate translational context (e.g., in vivo versus in vitro), thereby transcending limitations of earlier approaches to BCEP. This enables development of vaccines and immunodiagnostics, most notably by selecting BCEs for inclusion among peptide-based constructs that elicit production of antipeptide antibodies. Such antibodies can mediate protective immunity (e.g., by neutralizing pathogen virulence factors in vivo) and/or be used for antigen detection (e.g., with the aid of surfactants and other chaotropic agents to increase the extent of AD among pathogen-derived proteins in vitro). BCEP thus supports the design of peptide-based constructs that are potentially useful as vaccine components, as companion immunodiagnostics for monitoring antipeptide antibody responses to vaccination, and as means for generating antipeptide antibodies that, apart from mediating protective immunity (e.g., via active or passive immunization), may serve as immunodiagnostics reagents for antigen detection.

REFERENCES

1. Hopp TP, Woods KR. Prediction of Protein Antigenic Determinants From Amino Acid Sequences. Proc Natl Acad Sci USA (1981) 78:3824–8. doi: 10.1073/pnas.78.6.3824
2. Arnon R. Synthetic Vaccines—a Dream or Reality. Adv Exp Med Biol (1972) 31:209–22. doi: 10.1007/978-1-4684-3225-1_17
3. Van Regenmortel MH. What is a B-Cell Epitope? Methods Mol Biol (2009) 524:3–20. doi: 10.1007/978-1-59745-450-6_1
4. Caoili Sec. Beyond B Cell Epitopes: Curating Positive Data on Antipeptide Paratope Binding to Support Peptide-Based Vaccine Design. Protein Pept Lett (2021) 28:953–62. doi: 10.2174/0929866528666210218125624
5. El-Manzalawy Y, Dobbs D, Honavar VG. In Silico Prediction of Linear B-Cell Epitopes on Proteins. Methods Mol Biol (2017) 1484:255–64. doi: 10.1007/978-1-4939-6406-2_17
6. Ras-Carmona A, Pelaez-Prestel HF, Lafuente EM, Reche PA. BCEPS: A Web Server to Predict Linear B Cell Epitopes With Enhanced Immunogenicity and Cross- Reactivity. Cells (2021) 10:2744. doi: 10.3390/cells10102744
7. Zhu J, Gouru A, Wu F, Berzofsky JA, Xie Y, Wang T. BepiTBR: T-B Reciprocity Enhances B Cell Epitope Prediction. iScience (2022) 25:103764. doi: 10.1016/j.isci.2022.103764
8. Caoili SEC. Prediction of Variable-Length B-Cell Epitopes for Antipeptide Paratopes Using the Program HAPPTC. Protein Pept Lett (2022) 39:328–39. doi: 10.2174/092986652966620203101808
9. Alghamdi W, Attique M, Alzahrani E, Ullah MZ, Khan YD. LBCEPred: A Machine Learning Model to Predict Linear B-Cell Epitopes. Brief Bioinform (2022) 23:bbb035. doi: 10.1093/bib/bbb035
10. Jespersen MC, Peters B, Nielsen M, Marcattil P. BepiPred-2.0: Improving Sequence-Based B-Cell Epitope Prediction Using Conformational Epitopes. Nucleic Acids Res (2017) 45:W24–9. doi: 10.1093/nar/gkx346
11. da Silva BM, Myung Y, Ascher DB, Fries DEV. Epitope3D: A Machine Learning Method for Conformational B-Cell Epitope Prediction. Brief Bioinform (2022) 23:bbb0263. doi: 10.1093/bib/bbb0263
12. Rawal K, Sinha R, Abbasi BA, Chaudhary A, Nath SK, Kumari P, et al. Identification of Vaccine Targets in Pathogens and Design of a Vaccine Using Computational Approaches. Sci Rep (2021) 11:17626. doi: 10.1038/s41598-021-96863-x
13. Kesavwani V, Gupta R, Vetukuri RR, Kushwaha SK, Gandhi S. Identification of Unique Peptides for SARS-CoV-2 Diagnostics and Vaccine Development by an In Silico Proteomics Approach. Front Immunol (2021) 12:725240. doi: 10.3389/fimmu.2021.725240
14. Van Regenmortel MH, Pellequer JL. Predicting Antigenic Determinants in Proteins: Looking for Unidimensional Solutions to a Three-Dimensional Problem? Pept Res (1994) 7:224–8.
15. Blythe MJ, Flower DR. Benchmarking B Cell Epitope Prediction: Underperformance of Existing Methods. Protein Sci (2005) 14:246–8. doi: 10.1100/ps.041039505
16. Chen SW, Van Regenmortel MH, Pellequer JL. Structure-Activity Relationships in Peptide-Antibody Complexes: Implications for Epitope Prediction and Development of Synthetic Peptide Vaccines. Curr Med Chem (2009) 16:953–64. doi: 10.2174/092986709787581914
17. Van Regenmortel MHV. HIV/AIDS: Immunochemistry, Reductionism and Vaccine Design: A Review of 20 Years of Research. Cham, Switzerland: Springer Nature Switzerland (2019).
18. Naz A, Awan FM, Obaid A, Muhammad SA, Paracha RZ, Ahmad J, et al. Identification of Putative Vaccine Candidates Against Helicobacter pylori: Exploiting Exoproteome and Secretome: A Reverse Vaccinology Based Approach. Infect Genet Evol (2015) 32:280–91. doi: 10.1016/j.meegid.2015.03.027
19. Dwivedi P, Alam SI, Tomar RS. Secretome, Surfome and Immunome: Emerging Approaches for the Discovery of New Vaccine Candidates Against Bacterial Infections. World J Microbiol Biotechnol (2016) 32:155. doi: 10.1007/s11274-016-2107-3
20. MacrAld CA, Richards JS, Anders RF, Norton RS. Antibody Recognition of Disordered Antigens. Structure (2016) 24:148–57. doi: 10.1016/j.str.2015.10.028
21. Uversky VN, Van Regenmortel MHV. Mobility and Disorder in Antibody and Antigen Binding Sites Do Not Prevent Immunoochemical Recognition. Curr Rev Biochem Mol Biol (2021) 56:149–56. doi: 10.1080/10492380.2020.1869683
22. Trier N, Hansen P, Houen G. Peptides, Antibodies, Peptide Antibodies and More. Int J Mol Sci (2019) 20. doi: 10.3390/ijms20246289

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This work was funded by the University of the Philippines System, via One UP professorial chair grant 2019-100965.
33. Long Q, Yang Y, Yang M, Bai H, Sun W, Yang X, et al. Recombinant VLPs
32. Du X, Xue J, Jiang M, Lin S, Huang Y, Deng K, et al. A Multiepitope Peptide,
31. Wang C, Li Y, Wang S, Yan X, Xiao J, Chen Y, et al. Evaluation of a Tandem
30. Freidl R, Gstöttner A, Baranyi U, Swoboda I, Stolz F, Focke-Tejkl M, et al.
29. Caoili Comprehending B-Cell Epitope Prediction
28. Verma S, Singh K, Bansal A. Multi-Epitope DnaK Peptide Vaccine Accords
27. Abbott RK, Crotty S. Factors in B Cell Competition and Immunodominance.
26. Petrás M, Kralová Lesná I. SARS-CoV-2 Vaccination in the Context of Original Antigenic Sin.}
25. Wang C, Li Y, Wang S, Yan X, Xiao J, Chen Y, et al. Evaluation of a Tandem
24. Achtenes E, Reyeh, Randoval A. Protective Efficacy of Peptides From
23. MacRaid CA, Seow J, Das SC, Norton RS. Disordered Epitopes as Peptide
22. Allergy
21. doi: 10.1016/j.jaut.2017.04.008
20. doi: 10.1002/jmr.768
19. doi: 10.4155/fsoa-2017-0101
18. doi: 10.1016/s0022-1759(00)00239-8
17. doi: 10.1038/s41586-021-03819-2
16. doi: 10.1016/j.coi.2018.07.011
15. doi: 10.1038/jci.insight.127475
14. doi: 10.1172/jci.insight.122551
13. doi: 10.1002/bit.25985
12. doi: 10.1038/ncomms12861
11. doi: 10.1172/jci.insight.125067
10. doi: 10.1038/nature.2021.488571
9. doi: 10.1038/s41597-021-00661-x
8. doi: 10.1093/jb/mfs415
7. doi: 10.1073/pnas.1422401112
6. doi: 10.1002/ibm.201807127
5. doi: 10.1002/jbc.201804153
4. doi: 10.1038/s41557-021-00661-x
3. doi: 10.1038/s41586-021-03816-9
2. doi: 10.1038/ncomms12860
1. doi: 10.1002/jbc.201804152
