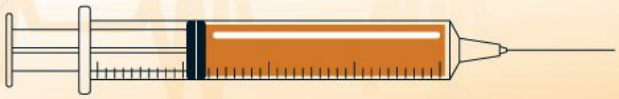
VACCIREVIEW



Applied immunoinformatic in modern vaccine design: a comprehensive review of available computational tools

Bibliography

Miles S, Mourglia-Ettlin G, Chabalgoity JA. Applied immunoinformatics in modern vaccine design: a comprehensive review of available computational tools. Vaccine. 2026;77:128392.

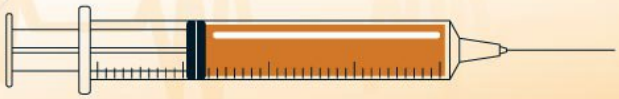
Summary

This review describes how immunoinformatics has transformed vaccine development from empirical experimentation into a structured, computationally driven workflow spanning target discovery, safety filtering, epitope prediction, structural design, and in silico immune simulation. The authors frame immunoinformatics as the integration of immunology, molecular biology, and data science, enabling rapid transition from pathogen genomes to rationally designed vaccine candidates, and emphasize the increasing role of artificial intelligence and deep learning across this pipeline.

After a concise overview of innate and adaptive immunity, including trained immunity and the central role of antigen presentation via HLA, the article situates immunoinformatics within two major conceptual advances: **reverse vaccinology and systems vaccinology**. Reverse vaccinology leverages omics data to move from genome to antigen, as exemplified by Rino Rappuoli for the *Neisseria meningitidis* serogroup B vaccine, while systems vaccinology uses multi-omics signatures to predict vaccine responsiveness and identify correlates of protection.

The core of the review is organized along the stages of a contemporary in silico vaccine design workflow. First, genomic and evolutionary analyses identify conserved, functionally essential targets and account for paralogy and redundancy using tools such as MEGA for phylogenetics and ConSurf, OrthoFinder, EggNOG-mapper, and related resources for conservation and orthology. Candidate proteins are then characterized for physicochemical properties, subcellular localization, signal peptides, and membrane topology with tools like ProtParam, PSORTb, WoLF PSORT, DeepLoc/DeepLocPro, SignalP, and DeepTMHMM, refining selections toward accessible, manufacturable antigens.

Functional importance is assessed through Gene Ontology, KEGG, InterProScan, DeepGOPlus, and interaction resources (STRING, BioGRID, SPAAN, D-SCRIPT) to prioritize virulence-related, surface-exposed, or adhesin-like proteins. Safety-oriented “negative design” is treated as a parallel stream: virulence factors (VFDB, DeepVF, VirulentPred), host–pathogen interactions, toxicity (ToxinPred2/3, ToxDL2), allergenicity (AlgPred2, AllerCatPro, AllerTOP), and autoimmunity risk via BLASTp against the human proteome are systematically screened. The Human Protein Atlas



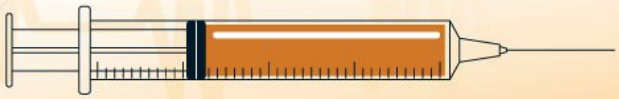
is highlighted as a key resource to avoid cross-reactivity with proteins expressed in critical human tissues.

Immunological characterization focuses on antigenicity, B-cell and T-cell epitope prediction, and population coverage. Tools such as VaxiJen, ANTIGENpro, and IApred estimate antigenicity, while NetBCE, BepiPred-3.0, and structure-based DiscoTope, EpiGraph, SEMA, and GraphBepi identify linear and conformational B-cell epitopes. For T cells, integrated pipelines model proteasomal cleavage (NetChop, PCPS), TAP transport (TAPpred, DeepTAP), MHC binding (NetMHCpan, NetMHCIIpan, MixMHCpred, MixMHC2pred, MHCflurry, BERTMHC), and even TCR recognition (ERGO-II, DeepImmuno). HLA polymorphism and global population coverage are addressed with AFND-based tools like PopCover and IEDB resources. Adaptive immune receptor repertoire analysis (GLIPH/GLIPH2, TCRdist3, immuneML) links predicted epitopes to observed TCR/BCR patterns, supports biomarker discovery, and informs neoantigen selection.

Structural biology and design receive extensive coverage. Secondary structure predictors (PSIPRED, SPIDER3, Porter5, NetSurfP-3.0) and modern 3D structure tools (AlphaFold2/3, RoseTTAFold, ColabFold, ESMFold, AlphaFold DB) are positioned as enablers of structure-based epitope mapping, docking (HADDOCK, ZDOCK, ClusPro, AlphaFold-Multimer, RoseTTAFold-All-Atom, AF2Complex), and molecular dynamics (GROMACS, NAMD, AmberTools, OpenMM). A major focus is generative protein design (“Reverse Vaccinology 3.0”) using RFdiffusion for backbone generation and ProteinMPNN and CAPE-Beam for sequence design and de-immunization, plus Rosetta MotifGrafting for epitope scaffolding and VLP-like displays.

The authors then discuss computational immunology and immune simulation, including C-ImmSim, ENISI-MSM, Serosim, Immunaut, MiStImm, DeepImmuno, VaxRank, pVACview, ImmuneApp, and cytokine-prediction tools (IFNepitope2, IL4pred2, IL-6-Pred, MultiFeatVotPIP, ProIn-Fuse), as well as resources for adjuvant selection such as the Vaccine Adjuvant Compendium. A final vaccine design section addresses stability (RaSP, DeepSTABp, iStable2, ProTstab2, DeepTM), hydrophobicity and hydration, intrinsic disorder (IUPred3, DEPICTER2, IDP-LM, DeepIDP-2L), solubility (PML_Sol, SKADE, EPSOL, ProtSolM, ProSol-multi), aggregation (AGGRESCAN, Tango, PASTA2, AgMata, AggreProt, AGGRESCAN4D), half-life (PLTNUM), and specific optimization of mRNA vaccines with LinearDesign and RNAdegformer.

Several success stories illustrate real-world impact: reverse vaccinology for meningococcal B vaccines, structure-based prefusion F design for RSV vaccines, and a multi-epitope Group B *Streptococcus* construct that conferred full protection in mice



and 100% survival after lethal challenge, including protection via passive antibody transfer. The authors close by stressing that, despite major advances, experimental validation remains indispensable and many in silico candidates never reach in vivo testing.

Comment

This review offers an impressively comprehensive and technically detailed map of the immunoinformatics landscape, spanning more than 250 tools and multiple conceptual layers from genomics to structural design and immune simulation. Its main strengths are the methodical, pipeline-oriented organization and the clear distinction between functional “positive” design (target and epitope choice) and “negative” design (safety filtering), which mirrors real-world vaccine R&D workflows. The authors also succeed in highlighting genuine paradigm shifts, such as generative AI for protein and antibody design and systems vaccinology for baseline-signature–driven prediction of vaccine responsiveness.

The breadth of coverage comes at the cost of depth in some areas. For many tools, performance is described qualitatively rather than with quantitative, head-to-head benchmarking data, leaving the reader without a clear sense of how to prioritize among overlapping methods. While the limitations section rightly stresses imperfect accuracy, allele- and context-dependence, and the translation gap between in silico predictions and in vivo outcomes, the discussion remains largely conceptual and could benefit from more concrete failure examples or guidance on minimal experimental validation packages. The authors note that many pipelines ignore baseline immunity, age, co-morbidities, microbiome, and geographic variation, but they stop short of proposing standardized strategies to integrate such data into routine analyses.

Practical usability is a recurrent theme, yet issues such as software maintenance, version drift, web-server stability, licensing, and required computational resources are only briefly touched upon in the tables. For translational groups, more explicit recommendations on a “minimal robust stack” for bacterial versus viral pathogens, or for subunit versus mRNA vaccines, would be valuable. Finally, although success stories are well chosen, they understandably emphasize positive outcomes; a more balanced appraisal of why most informatics-designed candidates fail to advance would sharpen the conclusions. Overall, this is a highly useful state-of-the-art map for specialists, but turning it into a practical handbook for diverse vaccine developers will require additional, context-specific implementation guidance.

*Brought to you by Chief Editor **Joe Schmitt** – Supported by AI*