Immunoinformatics and its role in microbes and vaccines

Chate P.B.1, Kayarkar N.A.2, Durgude S.G.2, Maurya B.D.2, Pawar S.V.2 and Gomase V.S.3

1Bhavan’s College, Andheri (west), Mumbai, 400 058, India
2V.E.S. College of Arts, Science and Commerce, Chenab, Mumbai, 400 071, India
3* School of Technology, S.R.T.M. University, Sub-Centre, Latur, 413512, India

Abstract- Immunoinformatics utilizes bioinformatics applications such as immune related databases with integration of mathematics, information science, computer engineering, genomics, proteomics which bridge immunology and informatics. Plants lack adaptive immune system and rely on innate immunity which consists of many protein interactions, which protect them from infections. The plant innate immunity consists of many molecular switches that help in activation of plant innate immunity. The plant innate immunity also consists of many elicitors and suppressors that elicit and suppress the plant innate immunity respectively. Fragment-based drug discovery is a new approach that builds drugs from small chemical structures.

Keywords- Immunoinformatics, Plant innate Immunity, Protein interactions, Molecular switches, Elicitors, Suppressors, Fragment Based Drug design

Introduction

With the burgeoning immunological data in the scientific literature, scientists increasingly rely on bioinformatics applications well developed for some immunological areas, to inform and enhance their work [1, 2]. There is an agreeable synergy between the growing collections in immune-related databases such as GenBank/GenPept, EMBL/TrEMBL, DDBJ/DAD, PIR, SWISS-PROT, PDB, PROSITE, etc. among which the IMGT database contains high quality annotations of DNA and protein sequence of Ig, TCR and MHC. These computational tools contribute to improved understanding of immune responses, and evolution of pathogens under immune pressure. For development of immunoinformatics tools we need the integration of immunological database with generic interfaces and also the integration of system level mathematical models with molecular level models leading to application in fields such as development of novel therapeutic regimens, vaccine designing and disease management [1, 2]. A number of computational methods have been developed to identify MHC-binding peptides and their subset of T-cell epitopes that helps improve our understanding of specificity of immune responses which is important for discovery of vaccines and immunotherapies [6, 7]. These computational methods consist of a variety of statistical and machine learning approaches making computational prescreening of antigens for CTL epitopes a standard approach in epitope-mapping studies [7]. Selection of antigen sequences as essential T-cell epitopes of supertype human leukocyte antigen (HLA) alleles lead to production of T-cell epitope based vaccines [3]. A web server, PEPPVAC (Promiscuous EPitope-based VACCine), was used for formulation of multi-epitope vaccines with broad population coverage [4]. In Dengue viruses (DENV) study, sequence fragments that were conserved across the majority of available DENV sequences evaluated their relevance as candidate vaccine targets, using various bioinformatics-based methods (NCBI Entrez protein database) and immune assay [5]. Plants lack mobile defender cells and a somatic adaptive immune system. They rely on the innate immunity of each cell and on systemic signals emanating from infection sites [8]. The plant innate immunity consist of PTI (PAM triggered immunity) and effector-triggered immunity (ETI) which involves interactions of proteins [9]. There are many molecular switches, which regulate the plant innate immunity such as NB-ARC, HSP90, SGT1, RAR1 etc [17, 18]. The plant innate immunity also consists of elicitors such as oomycete-derived Nep1, Avr9 [25, 29] and suppressor such as Cyclic beta-(1, 2)-Glucan, Xanthan [35, 40] that help to induce or suppress the plant innate immunity. Basic concept of fragment-based drug discovery was developed about 25 years go by William Jencks and it includes building of drugs from small molecular pieces and it has a great advantage of finding new drugs [41, 42].

Protein Interactions to Regulate Plant Immunity

Plants lack mobile defender cells and a somatic adaptive immune system. They rely on the innate immunity of each cell and on systemic signals emanating from infection sites [8]. Plants consists of trans-membrane receptors at their cell surface which recognize microbe- or pathogen associated molecular patterns (MAMPs or PAMPs) such as cell wall fragments, chitin or peptide motifs in bacterial flagella which induces the primary or basal defense responses, referred to as PTI (PAM triggered immunity). Many plant pathogens produce and deliver effector proteins in the host. To recognize these effector proteins plants evolved secondary defense referred to as effector-triggered immunity (ETI) and is mediated by resistance (R) proteins [9]. PTI is also called as primary driving force of plant–microbe interactions as it is the first facet of active plant defense [10]. Most plant species recognize a
highly conserved 22-amino-acid epitope, flg22, present in the flagellin N-terminus, as best-characterized PAMP in plants [10]. Leucine-rich repeat receptor-like kinase (LRRRLK) FLAGELLIN-SENSING 2 (FLS2) pattern recognition receptor is responsible for flagellin recognition in the plant model Arabidopsis thaliana [10]. Mutation in FLS2 makes the plant more susceptible to pathogenic bacterium Pseudomonas syringae pv. tomato DC3000 (Pto DC3000) [11]. Another recognized PAMP in Arabidopsis and other members of the family Brassicaceae is a most abundant bacterial protein Elongation factor Tu (EF-Tu) [10].

Resistance proteins

The NB-LLR is the core of R proteins, in this NB refers to nucleotide binding domain and LRR refers to leucine rich repeat domain which is fused to nucleotide binding domain. This R protein is equipped with variable amino- and sometimes also carboxy-terminal domains. On the basis of presence or absence of an amino-terminal Toll/interleukin-1 receptor-like domain the NB-LLR consists of two major subfamilies. The non- TIR NB-LRR proteins contain predicted coiled coil (CC) motifs and this family is referred as CC-NB-LRRs [9]. Pathogenic strains of Pseudomonas syringae delivers type III effector protein encoded by a virulence gene B (AvrB) and localizes to plasma membrane and induces immunity by the Arabidopsis coiled-coil (CC)-nucleotide binding (NB)-leucine-rich repeat (LRR) disease resistance protein RPM1 [13]. Several NB-LRR proteins recognize type III effectors indirectly, by detecting products produced by their action on host targets, consistent with the ‘guard hypothesis’. Arabidopsis RRS1-R at its CC domain of the potato (Solanum tuberosum) as molecular switch

Coiled-coil (CC) domain, Pto kinase, Prf and HDA19 [20].

R proteins as molecular switch

For two tomato (Lycopersicon esculentum) R proteins, I-2 and Mi-1, NB-ARC domain functions as a molecular switch whose state (on/off) depends on the nucleotide bound (ATP/ADP). Specific mutations were introduced in conserved motifs of the NB-ARC domain to investigate the role of nucleotide binding and hydrolysis for the function of I-2 in planta, and it was found that the...
ATP- rather than the ADP-bound state of I-2 is the active form that triggers defense signaling [23, 52].

SGT1 and Pti4, Pti5, and Pti6 proteins as molecular switch
SGT1 is a positive regulator of disease resistance which is conferred by many Resistance (R) proteins. AtSGT1a and AtSGT1b are two SGT1 proteins in Arabidopsis which are induced in leaves upon infection [24]. SGT1 may be involved in the proper folding of the Bs2 protein [63]. Pti4, Pti5, and Pti6 proteins from tomato activate the expression of GCC box–containing pathogenesis-related (PR) genes and play important in plant defense [53].

Elicitation of plant innate immunity
In addition to PAMP or AVR effector-mediated nonself recognition, breakdown products of the plant cell wall serve as endogenous danger signals that monitor distress of host structures and elicit plant immune responses. Such plant-derived elicitors are probably released by glucohydrolytic activities from attacking microbes [25].

Elicitors that induce immunity
Oomycete-derived Nep1 as elicitor
In Arabidopsis thaliana oomycete-derived Nep1 (for necrosis and ethylene-inducing peptide1)–like proteins (NLPS) trigger an extensive reprogramming of transcriptome, which was revealed by transcript profiling [25]. flg22 region of Xcc flagellin region and chitin as elicitor. In Arabidopsis it was found that the flg22 region of Xanthomonas campestris pv campestris (Xcc) flagellin was the only region responsible for detectable elicitation of Arabidopsis defense responses [26]. A Lysin motif (LysM) receptor-like protein (LysM RLK1) in Arabidopsis is required for chitin (a polymer of N-acetyl-D-glucosamine, found in fungal cell walls) signaling [54]. The LysM motif is a ubiquitous protein [55].

Bacterial induced stomatal closure
Bacterium-induced stomatal closure, which requires PAMP signaling and SA and ABA homeostasis, appears to be part of the plant innate immune system and can be activated by bacterial PAMPs such as the flagellin peptide flg22 [27]. Mitogen-Activated Protein Kinase3 (MPK3) in Arabidopsis is required for stomatal immune response [56].

Pepper pectin methylesterase inhibitor protein CaPMEI1 as elicitor
In pepper leaves infection with bacterial pathogens and treatment with plant hormones such as SA, ethylene, MeJA and ABA induces CaPMEI1 expression suggesting that this gene may be involved in the early stages of the active defense responses [28].

Avr9 as elicitor
In Nicotiana benthamiana Cff-9 and Cff-4 dependent hypersensitive response (HR) was elicited by three Avr9/Cff-9 Rapidly Elicited (ACRE) genes [29].

JA as elicitor
The herbivore susceptibility in plants is associated with the reduced levels of jasmonic acid–isoleucine (JA-Ile), but when Ile or JA-Ile is applied to the wounds of Threonine deaminase (TD)-silenced plants; it restores herbivore resistance [30].

(AvrPtoB1) as elicitor
The physical interaction of either sequence-dissimilar type III effector proteins AvrPto or AvrPtoB (HopAB2) from Pseudomonas syringae pv. Tomato with the host Ptokinase leads to elicitation of Pto/Prf-dependent immunity against Pseudomonas syringae pv. Tomato [31]. AvrPtoB homologs from diverse P. syringae pathovars have conserved avirulence and virulence activities similar to AvrPtoB activity and also elicit the Pto/Prf-dependent immunity [57].

Lipopolysaccharides (LPSs) and lipooligosaccharides (LOSs) as elicitor
Lipopolysaccharides (LPSs) and lipooligosaccharides (LOSs) are major components of the cell surface that are present in Gram-negative bacteria and have diverse roles in bacterial pathogenesis of animals and plants that include elicitation of host defenses [32].

Suppression of basal innate immunity
Some strains of vascular wilt fungus Fusarium oxysporum f. sp. lycopersici (Fol) secrete a small protein Avr2 that suppresses the activity of two disease resistance genes of tomato [33, 58].

Sinorhizobium meliloti (LPS) as suppressor
A specific concentration of S. meliloti LPS results in suppression of invertase induced oxidative burst in M. truncatula [34].

Bacterial Cyclic beta-(1, 2)-Glucan as suppressor
The black rot pathogen Xanthomonas campestris pv campestris (Xcc) consist of nodule development B (ndvB) gene which synthesizes cyclic beta-(1,2)-glucan which causes virulence. This was studied by introducing mutation to ndvB gene and so did not produce virulence but when beta-(1, 2)-glucan was supplied it produced virulence [35].

AvrPtoB and E3 ubiquitin ligase activity as suppressor
AvrPtoB type III effector protein of tomato pathogen *Pseudomonas syringae* suppresses programmed cell death (PCD) associated with plant immunity. It also exhibits E3 Ub ligase activity. The C terminus of AvrPtoB alone is sufficient for both anti-PCD and E3 Ub ligase activities and this suggest that the two functions are associated [36]. AvrPtoB a single bacterial effector elevate ABA levels, enhance bacterial growth, and suppress PAMP-responsive genes [39].

**Suppression of microRNA pathway and suppression by HopAO1 or HopF2**

*Arabidopsis* mutants deficient in microRNAs (miRNAs) partly restore growth of a type-three secretion-defective mutant of *Pseudomonas syringae* and also sustained growth of non-pathogenic *Pseudomonas fluorescens* and *Escherichia* coli strains which implies miRNAs is a key component in plant basal defense [37]. *Arabidopsis thaliana* that express either of two HopAO1 or HopF2, type III effector protein suppressed the HopA1-induced hypersensitive response (HR) [59].

**EIN3 and EIL1 as suppressor**

*Arabidopsis thaliana* over accumulating transcription factors ETHYLENE INSENSITIVE3 (EIN3) exhibit enhanced disease susceptibility to *Pseudomonas syringae* and is compromised in PAMP defenses. ETHYLENE INSENSITIVE3-LIKE1 (EIL1) also controls negatively PAMP response genes [38].

**Xanthan as suppressor**

The xanthan minus mutant (strain 8397) and the mutant strain 8396 fail to cause disease in both *Nicotiana benthamiana* and *Arabidopsis* (Arabidopsis thaliana) plants but when this strains are treated with xanthan, enhances the susceptibility of both *N. benthamiana* and *Arabidopsis* plants to both the mutant strains [40].

**Conserved effector loci (CEL) as suppressor**

Salicylic acid (SA) present in *Arabidopsis* plants induce resistance against *Pseudomonas syringae* mutated in conserved effector loci (CEL) but plants that were mutated in salicylic acid (SA) production did not provide resistance against the mutated CEL. This showed that salicylic acid (SA) is important for resistance [60].

**Fragment based drug design**

Fragment-based drug discovery was developed about 25 years ago by William Jencks. Fragment-based drug discovery builds drugs from small chemical structures (fragments) that may only exhibit weak binding affinity. Strategies are then applied to increase affinity. Thus, it attempts to build a ligand piece-by-piece, in a modular fashion [41]. Larger potential chemical diversity can be sampled with fewer compounds. This is its main advantage; which is particularly important for new target classes [42]. There are two key components of FBD: the detection technology and the compound library [62].

**Fragment based approach and detection of fragments**

Fragment-based lead discovery involves identifying from very much smaller compound libraries low molecular weight (<250) chemical fragments (also known as scaffolds or templates) and combining or optimizing them to produce a new compound. The fragments that are selected should consist of molecular weight of less than 300, CLogP equal to 3, and not more than 3 hydrogen bond donors and three acceptors. The fragments that are selected are detected by X-Ray Crystallography which provides detailed profile of fragment-binding [43]. The other method used for detection is NMR Screening which is a versatile technique for various aspects of hit identification, validation and optimization [45]. Fragments are generally less potent than hits obtained via HTS, and because of this they are subjected to various processes to convert them into potential drug leads. The strategies available to do this are the following:

- a. In Fragment Evolution the initial fragments that are identified by direct binding techniques are built up into larger, more complex molecules that target additional interactions in the active site of the protein.
- b. In Fragment Linking Two fragments that are identified bind in separate sites but which are close enough together to be chemically linked resulting in a larger, higher-affinity molecule.
- c. In Fragment Self-Assembly fragments undergo self-assembly in the presence of a template.
- d. In Targeted Libraries fragment used as the core template can efficiently map the features of the receptor allowing rapid generation of SAR [43]. To interrogate much larger compound libraries the method used is molecular docking [50].

Vernalis approach called SeeDs (Structural exploitation of experimental Drug startpoints) are used in fragment based drug discovery. The process is used to discover compounds against the oncology targets Hsp90 and PDK1 [61]. The use of differentiated fragment collections containing new, diverse scaffold sets may be used to more efficiently navigate chemical space towards areas that are currently unexplored and which are safe [44]. Computational chemistry can play an important role in producing a target focused fragment library prior to a fragment screen, and also in evolution of a drug-like molecule from a fragment hit, both with and
without the available fragment-target co-complex structure post-screening [46]. A 3D-pharmacophore that fit the active site of edema factor (EF) of Bacillus anthracis was constructed from fragments in a structure-based method to identify non-nucleotide inhibitors of EF [47]. SILCS: Site identification by ligand competitive saturation method is a method used to solve the problem of detecting and characterizing fragment binding. This method is applied to the BCL-6 protein, which is implicated in a variety of cancers [48]. Fragment-based drug discovery methods are capable of identifying minimal bonding determinants of active-site side-chain rearrangements and the mechanistic origins of spectroscopic shifts this result was found by amide ligands that bind weakly but specifically to the ricin active site, and produce significant shifts in positions of the critical active site residues Arg180 and Tyr80 [49].

Current research scenario

Hydropobicity analysis
Hydrophobicity is the physical property of the molecule such as amino acid which is related to its transfer free energy from a polar medium to an apolar medium [65, 66]. Hydrophobic residue sequences are used for revealing patterns related to protein tertiary structure [67]. Effect of peptide hydrophobicity on the action of antimicrobial peptide can also be studied [68].

Toxicogenomics
In toxicogenomics the adverse biological effects of exogenous agents on genes with the help of omics-based techniques such as genomics, transcriptomics, proteomics, metabolomics etc. are studied [69, 70]. Toxicogenomics has been applied in drug development and biomarker discovery [71].

Transgenomics
Transgenomic’s SURVEYOR Nuclease was used to screen PKD1 and PKD2 variants in diagnosis and prognosis of autosomal dominant polycystic kidney disease (ADPKD) [72]. Transgenomic’s WAVE System is also used for diagnosis of (ADPKD) and for early detection of drug resistance mutations in chronic myeloid leukemia [72, 73].

Cheminformatics
Cheminformatics is very useful in determining the drug-like characteristics of a compound [75]. ChemReader is a cheminformatic tool used for extracting chemical structure diagrams in research articles and the analog-to-digital conversion is done thus it has the basic application of storing informations that are related to compounds [74, 76].

Pharmacoinformatics
Pharmacoinformatics consist of various new immigring information technologies that lead to drug discovery, it consists of internet, cheminformatics, immunoinformatics, etc to solve drug related problems and provide improved patient safety [77, 78]. Multiple model (MM) is used to achieve therapeutic goals [79].

Pharmacophore modeling
It is a method to identify new potential drugs for the targets whose 3D structure are not known, it consist of ligand based approach [80, 81]. It is important computational tool in rational drug design [82].

New lead discovery
This consists of fragment based lead discovery, in this low molecular weight fragments or compounds are used to obtain new drugs [83]. Structural biology along with bioinformatics has contributed in target identification and lead discovery [84]. Thus the computer aided technologies are important for new drug discovery [85].

Plant pathological condition and assay
Two important techniques used in immunosorbent assay in plant pathology are immunosorbent electron microscopy (ISEM) and ELISA [86]. Laboratory assay were performed to study the effect of P. infestans on leaves that were kept under different conditions [87]. To find markers common to all isolates of Fusarium poae that infect the wheat, PCR was carried out [88].

References


