We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

3,900
Open access books available

116,000
International authors and editors

120M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Bioinformatics as a Tool to Identify Infectious Disease Pathogen Peptide Sequences as Targets for Antibody Engineering

Lavanya Suneetha, Prasanna Marsakatla, Rachel Supriya Suneetha and Sujai Suneetha

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.71011

Abstract

Bioinformatics is an interdisciplinary field of information technology for understanding biological data from genome to protein. It includes a combination of fields of science, computer science, statistics, mathematics, and engineering to analyze, interpret and derive biological data. This chapter describes how to use Bioinformatics to identify pathogen virulence factor peptide sequence similarities in human nerve tissue proteins and for evaluation as antibody engineering target peptides.

Keywords: bioinformatics, infectious diseases, peptide, recombinant antibody

1. Introduction

Bioinformatics is the application of techniques derived from disciplines such as applied mathematics, computer science, and statistics to analyze and interpret biological data. In this chapter, you will learn how to use bioinformatic techniques to identify pathogen virulence factor (VF) peptide sequence similarities to human nerve tissue proteins and then how to identify target peptides that could form the basis for engineering recombinant antibodies. Also, wet experiments could be conducted on the identified overlapping sequences to help us to single out target antibodies to be tested for tissue culture studies [1, 2]. The most ideal targeted peptide sequences for antibody engineering are those physiologically relevant, easy to access, and comprise amino acid sequence regions which have high specificity in pathogenic steps and reduced amino acid string length.
1.1. Bioinformatics and its role in peptide discovery

The accessibility to the extensive genomic and proteomic databases and the availability of tools to compare and evaluate the information have given rise to a new interdisciplinary field that combines biology and computer science [3]. Bioinformatics conceptualizes physical and chemical biology in terms of macromolecules and then applies “informatics” techniques (derived from disciplines such as applied mathematics, computer science, and statistics) to assimilate and organize the information associated with these molecules, on a large scale [4]. Bioinformatics is an exciting and exploratory method for peptide discovery in antibody engineering and development of antimicrobial therapies and vaccination strategies [5].

There is significantly growing evidence that a number of neurodegenerative diseases are a result of the association of host cell proteins with viral and bacterial infectious agents [6]. When pathogenic microorganisms such as bacteria, viruses, parasites, or fungi cause an infectious disease, there are many molecular interactions between the host-pathogen proteins and host peptides [7] through all the stages of the disease whether incubation, prodromal illness, decline, and convalescence. There is much experimental evidence identifying the virulence factors (VF) of pathogen and host components such as receptors and tissue-specific proteins [8, 9]. Though the pathogenic pathway of the infectious agent in various host tissues is unknown, many of these processes are suspected to be attributable to the yet undiscovered role of molecular mimics identified in pathogenic microorganisms and its corresponding host tissue proteins. The sequence and structural similarities between the pathogenic VF protein and nerve peptides could impact either directly or indirectly the pathogenesis of the infectious disease [10–12]. It could contribute to molecular mimicry, steric hindrance, receptor binding, cell signaling, and autoantibody production events (involved in neuro degeneration) in the host.

Leprosy patients with peripheral nerve damage develop autoimmunity to myelin P0 (nerve protein). The above conclusion was drawn by gathering known scientific evidence that are as follows: (1) labeling and binding studies found that *Mycobacterium leprae* (bacterium causing leprosy) binds to myelin P0 [13]; (2) clinical studies confirmed the production of autoantibodies as a response of the bacterium to interact with myelin P0 [14, 15]; and (3) bioinformatics searches identified sequences and structural similarities between *M. leprae* and the immunoglobulin regions of myelin P0 [16].

Identification of molecular mimics in pathogen-host peptide sequences is one approach to identify target peptides for antibody engineering. There are about 180 extensive biological databases to retrieve information on sequence and functional aspects of biological molecules. The updated list is available in Nucleic Acids Research [17].

1.2. The use of bioinformatics in identifying sequence similarities

This section teaches you how to conduct a search for proteins present in a target host, how to obtain its amino acid sequence/s from the existing databases, how to compare the sequence/s of the host protein to that of the pathogen protein, and finally how to interpret the results based on existing evidential data. In our case study, we identify the virulence factor peptide
sequence similarities of a few selected infectious agents with human nerve tissue proteins for selecting peptides to engineer antipeptide antibodies which recognizes corresponding host/viral proteins.

1.2.1. Selection of nerve proteins

63 proteins were extracted from the Human Protein Atlas Database that were enriched and enhanced in the nervous tissue as observed by immunohistochemistry (Figure 1).

To conduct a search for human proteins in the nervous tissue, access the website (www.proteinatlas.org), enter the tissue of study (e.g. nervous tissue) into the search box provided and click on search.

Manual protein selection was carried out based on their tissue expression (enriched and enhanced) and also on immunohistochemistry evidence (Figure 2).

The list of selected proteins are as follows: agrin (AGRN_HUMAN, O00468), calbindin (CALB1_HUMAN, P05937), n-chimaerin (CHIN_HUMAN, P15882), secretogranin-2 (SCG2_HUMAN, P13521), neuromodulin (NEUM_HUMAN, P17677), kinesin (KIFC1_HUMAN, Q9BW19), tau (TAU_HUMAN, P10636), 2′,3′-cyclic-nucleotide 3′-phosphodiesterase (CN37_HUMAN, P09543), myelin-associated glycoprotein (MAG_HUMAN, P20916), myelin P0 (MYP0_HUMAN, P25189), myelin P2 (MYP2_HUMAN, P02689), oligodendrocyte-myelin
glycoprotein (OMGP_HUMAN, P23515), brain-derived neurotrophic factor (BDNF_HUMAN, P23560), ciliary neurotrophic factor (CNTF_HUMAN, P26441), neurotrophin-3 (NTF3_HUMAN, P20783), beta-nerve growth factor (NGF_HUMAN, P01138), nestin (NEST_HUMAN, P48681), neurofilament heavy polypeptide (NFH_HUMAN, P12036), neurogranin (NEUG_HUMAN, Q92686), voltage-dependent T-type calcium channel subunit alpha-1G (CAC1G_HUMAN, O43497), hippocalcin (HPCL1_HUMAN, P37235), neurocalcin-delta (NCALD_HUMAN, P61601), recoverin (RECO_HUMAN, P35243), bombesin receptor subtype-3 (BRS3_HUMAN, P32247), kininogen-1/bradykinin (KNG1_HUMAN, P01042), calcitonin (CALC_HUMAN, P01258), cholecystokinin (CCKN_HUMAN, P06307), galanin peptides (GALA_HUMAN, P22466), pro-neuropeptide Y (NPY_HUMAN, P01303), neuropeptide/neuromedin N (NEUT_HUMAN, P34900), protein S100-B1 (S100B_HUMAN, P04271), synapsin-1 (SYN1_HUMAN, P17600), probable tubulin polyglutamylase (TTL1_HUMAN, O95922), myelin basic protein (MBP_HUMAN, P02686), protein phosphatase 1 regulatory subunit 1B (PP1B_HUMAN, Q9UD71), Arf-GAP with GTPase, ANK repeat and PH domain-containing protein 2 (AGAP2_HUMAN, Q99490), cathepsin L2 (CATL2_HUMAN, O60911), D1 dopamine receptor (DRD1_HUMAN, P21728), BDNF/NT-3 growth factors receptor (NTRK2_HUMAN, Q16620), melanoma-associated antigen E1 (MAGE1_HUMAN, Q9HC5), microtubule-associated protein 6 (MAP6_HUMAN, Q96E9), protocadherin alpha-12 (PCDAC_HUMAN, Q9UN75), carboxypeptidase E (CBPE_HUMAN, P16870), Down syndrome cell adhesion molecule (DSCAM_HUMAN, O60469), dyslexia-associated
protein KIAA0319 (K0319_HUMAN, Q5VV43), uncharacterized protein KIAA1211-like (K121L_HUMAN, Q6NV74), microtubule-associated protein IB (MAP1B_HUMAN, P46821), neuronal calcium sensor 1 (NC51_HUMAN, P62166), neurofilament light polypeptide (NFL_HUMAN, P07196), receptor expression-enhancing protein 2 (REEP2_HUMAN, Q9BRK0), secretogranin-3 (SCG3_HUMAN, Q8WXD2), ubiquitin carboxyl-terminal hydrolase isozyme L1 (UCHL_HUMAN, P09936), galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase 1 (B3GA1_HUMAN, Q9P2W7), beta-1,4 N-acetylgalactosaminyltransferase 1 (B4GN1_HUMAN, Q00973), caprin-2 (CAPR2_HUMAN, Q6IMN6), dopamine beta-hydroxylase (DOPO_HUMAN, P09172), FAM81A (FA81A_HUMAN, Q8TB8), mitogen-activated protein kinase 10 (MK10_HUMAN, P53779), N-terminal EF-hand calcium-binding protein 1 (NECA1_HUMAN, Q8NV87), neurofilament protein C and casein kinase substrate in neurons protein 1 (PACN1_HUMAN, Q9BY11), sodium channel protein type 7 subunit alpha (SCN7A_HUMAN, Q01118), and clathrin coat assembly AP180 (AP180_HUMAN, O60641). The biological accepts of the proteins have been derived from the information presented in UniProt database for each protein [18–20].

1.2.2. Retrieving FASTA formats

FASTA formats for each of the above proteins were retrieved from NCBI PubMed. The FASTA format is a text-based format obtained from the PubMed search and represents either nucleotide sequences or peptide sequences (Figure 3).

Upon accessing the website, select the database in which the search is to be conducted (e.g. Protein). Type the name of the protein and its species in brackets into the search text box provided (e.g. Agrin (Homo sapiens)) and click on the search button.

The protein with the highest number of amino acids is chosen. Click on the hyperlinked protein to access its gene bank. Upon reaching the gene bank of the selected protein, click on the hyperlinked FASTA (Figures 4, 5 and 6).

Obtain the FASTA format by copying all the information (Starting from the > symbol).

1.2.3. Arranging the FASTA formats

All the FASTA formats of the human proteins are saved in a sequence on Microsoft Notepad (Figure 7).

1.2.4. Running the BLAST

Pathogen-protein mimics, nerve protein sequences were BLAST (Basic Local Alignment Search Tool; Version 2.7.1; e-value ≤0.01) [21] against a pathogen genome (Figure 8).

Access the BLAST website at https://blast.ncbi.nlm.nih.gov/Blast.cgi and click on Protein (Protein–Protein) BLAST. The FASTA formats of 63 nerve proteins were copied and pasted from the notepad into the text box provided. Enter the species of the organism against which the blast has to be performed/sequence comparison has to be carried out specifying its Tax ID (Figure 8).
Figure 3. Conducting a search on the PubMed database.

Figure 4. List of available sequenced protein information.
Figure 5. Gene information of agrin.

Figure 6. FASTA format of agrin.
Figure 7. FASTA formats of the 63 proteins in sequence.

Figure 8. BLAST home page.
The pathogen genome sequences that were compared with the human nerve proteins are as follows: HIV (Tax ID: 11,676), Polio (Tax ID: 138,950), Japanese Encephalitis (Tax ID: 64,320), M. leprae (Tax ID: 1769), Human herpes virus 1 (Tax ID: 10,298), Human herpes virus 2 (Tax ID: 10,310), Rabies virus (Tax ID: 11,292), Zika virus (Tax ID: 64,320), Corona virus (Tax ID: 11,118), Varicella zoster virus (Tax ID: 10,335).

Select program PSI BLAST as the BLAST algorithm for a more position-sensitive search. It looks deeper into the database to best match to your query. Click on the BLAST button and wait for the results. Take screen shots of your result and also download the provided excel format (Figure 9).

The output of the BLAST identified the significant peptide sequence similarities between the human protein and its pathogenic counterpart Figure 10. These peptide sequence similarities are identified by amino acid positions, in which amino acids exist in single-letter codes.

The BLAST provides us with the number of sequence similarities between the pathogenic genomic sequence and its host proteins. It also identifies viral counterpart peptides and the region of similarity on the host proteins.

Further interpretations of the results can be made by referring to the Uniprot database to obtain the biological and functional aspects of the host and the pathogen proteins (Figures 11 and 12).

1.2.5. Ascribing a biological role and application

The results show a number of sequence similarities existing between host proteins and various pathogen proteins. The maximum number of peptide sequence similarities were found between host protein caprin-2 which had 495 similarities with polio; neurogranin had 230
Figure 10. BLAST results of nerve proteins showing similarity to pathogen proteins.

Figure 11. UniProtKB screenshot showing the biological and functional data of the human protein.
similarities with HHV2; secretogranin-3 had 221 similarities with Japanese encephalitis; agrin had 212 similarities with varicella; caprin-2 had 198 similarities with rabies virus; galanin peptides had 87 similarities with Zika virus; kinesin had 54 similarities with HIV; neurofilament heavy polypeptide had 46 similarities with corona virus; neurogranin had 39 similarities with HHV1; and 2′,3′-cyclic-nucleotide 3′-phosphodiesterase had 21 similarities with M. leprae.

This method identifies significant virulent factors which have sequence similarities to human nerve tissue proteins. The nerve proteins that exhibited sequence similarities with four or more pathogenic virulent factors are displayed in Table 1. All 63 proteins are found to have sequence similarities with M. leprae proteins.

Agrin is a heparin sulphate basal lamina glycoprotein with a molecular mass of 217,232 Da. It plays a central role in the formation and maintenance of the neuromuscular junction. It is known to direct events in postsynaptic differentiation. Agrin also induces the phosphorylation and activation of muscle-specific kinase (MUSK), the clustering of Acetyl choline esterase receptor (AChR) in the postsynaptic membrane, regulates calcium ion homeostasis in neurons, and is involved in regulation of neuritis outgrowth [22, 23].

1.2.6. HHV3 peptide similarity to human protein agrin

Agrin UniProtKB-O00468 (AGRIN_HUMAN) (AA position 1269–1326) (Figure 13) has a similarity to membrane glycoprotein C (Sequence ID: AEW88711.1 AA Position 43–122) of the varicella zoster virus UniProtKB-Q9J3M8 (GE_VZVO) which by its similarity has the
<table>
<thead>
<tr>
<th>S. No</th>
<th>Query No.</th>
<th>Proteins</th>
<th>HIV</th>
<th>Polio</th>
<th>JE</th>
<th>HHV 1</th>
<th>HHV 2</th>
<th>M. lepra</th>
<th>Corona</th>
<th>Zika</th>
<th>Rabies</th>
<th>Vericella</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Q00468</td>
<td>Agrin</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>212</td>
</tr>
<tr>
<td>2</td>
<td>P17677</td>
<td>Neurromodulin</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>28</td>
<td>1</td>
<td>0</td>
<td>75</td>
</tr>
<tr>
<td>3</td>
<td>Q9BW19</td>
<td>Kinesin</td>
<td>54</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>P10836</td>
<td>Tau protein</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>5</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>19</td>
</tr>
<tr>
<td>5</td>
<td>P25189</td>
<td>Myelin protein P0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>22</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>P23515</td>
<td>Oligodendrocyte-myelin glycoprotein</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>23</td>
</tr>
<tr>
<td>7</td>
<td>P48681</td>
<td>Nestin</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>8</td>
<td>2</td>
<td>30</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>8</td>
<td>P17600</td>
<td>Synapsin-1</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>11</td>
<td>2</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>P02686</td>
<td>Myelin basic protein</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>9</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>Q16620</td>
<td>BDNF/NT-3 growth factors receptor</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>23</td>
<td>8</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>11</td>
<td>Q5V43</td>
<td>Dyslexia-associated protein KIAA0319</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>37</td>
<td>21</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>P07196</td>
<td>Neurofilament light polypeptide</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>77</td>
</tr>
<tr>
<td>13</td>
<td>Q8WXD2</td>
<td>Secretogranin-3</td>
<td>3</td>
<td>5</td>
<td>221</td>
<td>0</td>
<td>10</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>Q00973</td>
<td>Beta-1,4 N-acetylgalactosaminyltransferase 1</td>
<td>1</td>
<td>29</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 1. Sequence similarities of human nerve tissue proteins with human virulent factors. Multiple alignments obtained in a single BLAST search could result in identities of the amino acids or substitutions of the amino acids in the same peptide region.
potential to bind to the tissue cell receptor. Experimental evidence in epithelial cells shows that the hetero demonization of viral receptors could spread the virus by sorting nascent virion to nerve tissue cell junctions. The virus particles can spread to adjacent cells through interactions with cellular receptors at these cell junctions. The virus at cell junctions spreads extremely rapidly into the tissues [24, 25]. Sequence mimics of agrin to the varicella membrane glycoprotein could have an effect on either viral entry into host cell, evasion or on tolerance of host immune response to the virus and virion attachment to the host cell. These similarities in peptide regions warrant further exploration to understand pathogenesis and to identify target peptides for antibody engineering [26].

1.2.7. Poliovirus and rabies virus peptide similarities to human protein caprin-2

Caprin-2UniProtKB-Q6IMN6 (CAPR2_HUMAN) is a protein of molecular mass 68,429 Da. The structure of caprin-2 was found to be similar to the polio and rabies viruses. Caprin-2 (AA position: 136–176) has a similarity to the polyprotein of polio virus UniProtKB–E0WCG5 (E0WCG5_9ENTO) (polyprotein sequence ID: ACZ05040.1 AA position: 1994–2070) (Figures 14 and 15). Caprin-2 (AA position: 13–54) also has a similarity to the phosphoprotein of rabies virus UniProtKB-Q80IL8 (Q80IL8_9RHAB) (phosphoprotein sequence ID: AAO60615.1 AA position 76–110) (Figure 15). Caprin-2 has a significant role in influencing phosphorylation of the Wnt-signaling pathways (PubMed:18,762,581) [27]. Caprin-2 also facilitates LRP6 phosphorylation by CDK14/CCNY during G2/M stage of the cell cycle, which may potentiate cells for transport or translation of mRNAs, modulate the expression of neuronal proteins involved in synaptic plasticity [28], while simultaneously influencing cell cycle signaling and regulation of viral transcription and replication [29, 30].
1.2.8. *Mycobacterium leprae* peptide similarity to 2′, 3′-cyclic-nucleotide 3′-phosphodiesterase

2′, 3′-cyclic-nucleotide 3′-phosphodiesterase UniProtKB-P09543 (CN37_HUMAN) is a protein of molecular mass 47,579 Da. 2′, 3′-cyclic-nucleotide 3′-phosphodiesterase (sequence ID: WP_010908292.1 AA position 191–261) has a similarity to thiamin pyrophosphokinase of *M. leprae* UniProtKB A0A197SEI9 (A0A197SEI9_MYCLR) (AA position: 170–2166) (Figure 16). 2′, 3′-cyclic-nucleotide 3′-phosphodiesterase is involved in RNA metabolism of the myelinating cell, CN37 (2′, 3′-cyclic-nucleotide 3′-phosphodiesterase) is the one of the most abundant myelin protein in nervous system. The sequence similarities identified could impact cell signaling and also regulate energy metabolism [31].

Figure 14. BLAST output of polyprotein of poliovirus showing similarity to human protein caprin-2.

Figure 15. BLAST output of phosphoprotein of rabies virus showing similarity to human protein caprin-2.
1.2.9. Zika virus peptide similarity to human protein galanin

Galanin peptide UniProtKB-P22466 (GALA_HUMAN) is a protein of molecular mass 13,302 Da. Galanin (AA position 53–99 position) has a similarity to polyprotein envelope protein E of Zika virus UniProtKB-Q73880 (Q73880_9HIV1) sequence ID: ARB07952.1 (AA position: 729–765) (Figure 17). Galanin is involved in the smooth muscle contraction of the gastrointestinal and genitourinary tract, regulation of growth hormone release, modulation of insulin release, and might also be involved in the control of adrenal secretion [32]. The envelope protein E of the Zika virus is responsible for binding to host cell surface receptors and mediating fusion between viral and cellular membranes. It is synthesized in the endoplasmic

Figure 16. BLAST output of thiamin pyrophosphate of Mycobacterium leprae showing similarity to human protein 2', 3'-cyclic-nucleotide 3'-phosphodiesterase.

Figure 17. BLAST output of polyprotein of Zika virus showing similarity to human Galanin peptide.
reticulum with protein prM and forms a heterodimer. Galanin’s similarity with the ZIKA polypeptide could subsequently affect neural regulation of muscle function and play a role in immune evasion pathogenesis and viral replication [33].

1.2.10. HIV 1 peptide similarity to human kinesin-like protein

Kinesin-like protein KIFC1 UniProtKB-Q9BW19 (KIFC1_HUMAN) is a protein of molecular mass 73,748 Da. Kinesin-like protein (AA position: 411–470) has a similarity to HIV virus envelope glycoprotein UniProtKB-D6QPK9 (D6QPK9_9HIV1) sequence ID: ADG63850.1 (AA position: 270–387) (Figure 18). KIFC1 along with microtubules contributes to movement of endocytic vesicles. These similarities could affect viral attachment to the host cell, membrane fusion, and entry into the cell and the nucleus [34, 35].

1.2.11. Corona virus peptide similarity to human neurofilament heavy polypeptide

Neurofilament heavy polypeptide UniProtKB-P12036 (NFH_HUMAN) is a protein of molecular mass 112,479 Da. Neurofilament heavy polypeptide (AA position: 819–872) has a similarity to ORF1a UniProtKB-A0A0F6SKM6 (A0A0F6SKM6_9GAMC) of Corona virus sequence ID: AKF17723.1 (AA positions: 890–1031) (Figure 19) neurofilament of the nerve tissue usually contain three intermediate filament proteins: L, M, and H (NFH-human) which is involved in the maintenance of neuronal caliber. NFH-H has an important function in axon maturation. These similarities could affect viral replication, protein processing, and could generate autoantibody production [36, 37].

1.2.12. HHV 1 and HHV 2 peptide similarity to human protein neurogranin

Neurogranin UniProtKB-Q92686 (NEUG_HUMAN) is a protein of molecular mass 7618 Da. The structure of neurogranin at identical regions has a similarity to envelope glycoprotein M of

Figure 18. BLAST output of envelope glycoprotein of HIV 1 showing similarity to human kinesin-like protein.
HHV1 and envelope glycoprotein M of HHV2 at partially overlapping positions. Neurogranin (AA position: 38–63) has a similarity to the envelope glycoprotein M of HHV1 (UniProtKB-A0A181ZHE7 (A0A181ZHE7_HHV11) (sequence ID: SBO07578.1 AA position: 347–376) (Figure 20). Neurogranin (AA position: 38–64) also has a similarity to the envelope glycoprotein M of HHV2 (UniProtKB-A0A0Y0R357 (A0A0Y0R357_HHV2)) (sequence ID: AMB66044.1 AA position 389–416) (Figure 21). Neurogranin functions as a signaling messenger, a substrate for protein kinase C and has affinity to calmodulin in the absence of calcium. These similarities

Figure 19. BLAST output of ORF1 of corona virus showing similarity to human neurofilament heavy polypeptide.

Figure 20. BLAST output of envelope glycoprotein of HHV1 showing similarity to human protein neurogranin.
of HHV1 & 2 with neurogranin could have an interaction with viral transport into the host cell Golgi network and subsequently to the host nucleus [38].

1.2.13. JE 2 peptide similarity to human protein secretogranin-3

Secretogranin-3 UniProtKB-Q8WXD2 (SCG3_HUMAN) is a protein of molecular mass 53,005 Da. Secretogranin-3 (AA position: 139–190) has a similarity to the polyprotein of Japanese encephalitis virus (UniProtKB-G3LHD8 (G3LHD8_9FLAV) (sequence ID: SBO07578.1 AA position: 2744 to (Figure 22). Secretogranin-3 is a member of the chromogranin/secretogranin

Figure 21. BLAST output of envelope glycoprotein of human alpha herpes virus 2 showing similarity to human protein neurogranin.

Figure 22. BLAST output of polyprotein of JE 2 showing similarity to human protein secretogranin-3.
family of neuroendocrine secretory proteins comprising a number of significant cellular functions. In an experimental mouse model, autoimmunity with secretogranin was associated with encephalitis [39]. These similarities identified in the host-pathogen could affect neuroendocrine secretory protein release and autoimmunity.

2. Creating a schematic model

The sequence similarities in agrin, caprin-2, 2', 3'-cyclic-nucleotide 3'-phosphodiesterase, galanin peptide, kinesin-like protein, neurofilament heavy polypeptide, neurogranin and secretogranin-3 with its corresponding pathogenic peptide/s could have a number of cellular-level implications which include alternations in receptor binding, signaling/synaptic transmission, metabolic alteration, inflammation, resulting in autoimmunity and consequently neuropathy (Figure 23) [11, 40].

![Figure 23. A model for the modes of host-pathogen interaction and possible intracellular regulation of metabolic activities.](http://dx.doi.org/10.5772/intechopen.71011)

3. Conclusion

In conclusion, it is important to conduct bioinformatic searches and design wet experiments with the objective of identifying a vast number of functionally significant peptides for further comparison and study. Bioinformatic search tools and various available databases are to be extensively explored to rapidly develop possible neuroprotective or pathogenic peptide sequences. These peptides can be further explored as targets to generate recombinant antibodies. This exercise can also be used to develop an efficacious and safe vaccine against pathogens that demonstrate no autoimmune cross-reactions. It can also contribute to design peptide/drug molecules to neutralize the effects of neurotoxins. Bioinformatics is the key to open the door of understanding medical and biological processes in the future.
Acknowledgements

We acknowledge short-term project works of Do Eon Lee of York University, 700 Keele St, Toronto, ON M3 J 1P3, Canada and Logeshwaran Vasudevan of Bharathidasan University, Palkalaiperur, Tiruchirappalli, Tamil Nadu 620024, and Dr Sharon Bushi of Morristown Med CtrIntlMedcn, 100 Madison Ave, Morristown, NJ 07960 on the preliminary work of nerve protein pathogen similarity searches.

Author details

Lavanya Suneetha*, Prasanna Marsakatla, Rachel Supriya Suneetha and Sujai Suneetha

*Address all correspondence to: drlavanyasuneetha@gmail.com

Infectious Disease Research Laboratory, Nireekshana ACET, Hyderabad, India

References


