Majalah Kedokteran Sriwijaya,
Th. 56 Nomor 1, January 2024

BIOINFORMATICS ANALYSIS OF INTERLEUKIN-6

Desi Oktariana[1*]
[1]Clinical Pathology Department, Biomedical Science Department, Medical Faculty, Universitas Sriwijaya, Palembang, Indonesia.

desioktariana@fk.unsri.ac.id

Abstract

Interleukin-6 (IL-6) has important role as pro-inflammatory cytokine. Study of the molecular characteristic is needed before conducting further research, thus can provide insight and understanding regarding the properties and characteristics of IL-6. Therefore, this study aims to provide comprehensive bioinformatics analysis of IL-6 at biomolecular level.Bioinformatic analysis of IL-6 was obtained from several websites that can be accessed freely on internet, such as NCBI, PSIPRED, PROTPARAM, TMHMM, PROTSCALE, PEPTIDECUTTER, PROSITE, NETNGLYC, TARGETP, SIGNALP, and Swissmodel.Human IL-6 gene is located on chromosome 7p15.3. The IL-6 protein is 212amino acids (aa) long, secondary structure is mostly in helical-shaped amino acid form. The protein is unstable, with half time 30 hours. Most of the IL-6 protein is located outside the cell membrane. It has glycosylation site at the 73 and 172 amino acid positions, with domain motif at positions 101-126. The prediction of the target location is mostly in the signal peptide pathway. Predicted cleavage site is between position 27 and 28 of amino acid.Bioinformatics analysis is very useful in determining the molecular characteristics of IL-6.

Keywords : bioinformatics, interleukin-6. biomolecular

1. INTRODUCTION

Interleukin-6 (IL-6) is a cytokine that has an important role in pro-inflammatory activity. It was first discovered to be a B-cell stimulatory factor. It was first recognized that T cells create the lymphokine that ultimately causes B cells to be mature cells that manufacture antibodies. This cytokine is involved in numerous aspects of immune system, metabolism, and nervous system, both regulation and coordination. It has a part in numerous autoimmune disorders, as well as the body's defense mechanisms against infections, numerous regeneration processes, and body weight regulation. Activated immune cells and stromal cells, such as hepatocytes, endothelial cells, monocytes/macrophages, fibroblasts, and T cells, generate IL-6.[1]

Stressful environmental conditions including infections and tissue damage cause an instantaneous and transitory expression of IL-6. A warning signal is sent out by this expression, and the host's stress-reduction systems are triggered. In addition, IL-6 stimulates the generation of effector T cells and antibodies, which contributes significantly to the acquired immune response. Additionally, numerous nonimmune cells can be stimulated to differentiate or proliferate by IL-6.[2,3]

Interleukin-6 is a key factor in inducing acute phase reactions as well as cellular and humoral immune responses to affected cells

and mucosal humoral responses targeted against reinfection in chronic diseases, which are typically exemplified by immune stressors such as tumors and chronic intracellular infections. This cytokine is in charge of promoting both the bone marrow's production of neutrophils and acute phase protein synthesis. It is hostile to regulatory T cells and promotes the development of B cells.[4,5]

Bioinformatics science is growing fast in this digital era. It can also be used for many research in medical and biomedical science.[6] Study of biomolecular characteristic is needed before conducting further research. It can provide insight, thus can assist researchers in understanding the properties and characteristics of a compound. Therefore, this study aims to provide comprehensive bioinformatics analysis of IL-6 at biomolecular level.

2. METHODS

The tools to analyze bioinformatic information of IL-6 was obtained from several websites that can be accessed freely on internet:

a. To analyze genetic and protein characteristics of IL-6, we use the features present on the National Center for Biotechnology Information's website (www.ncbi.org).
b. To analyze the prediction of protein secondary structure, we use PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/).
c. To analyze the physicochemical characteristics of a protein, we use PROTPARAM (https://web.expasy.org/protparam/).
d. To analyze the Topology of Trans-membrane Proteins, we use TMHMM (https://services.healthtech.dtu.dk/services/TMHMM-2.0/).
e. To analyze the hydrophobicity, we used PROTSCALE (https://web.expasy.org/protscale/).
f. To predict cleavage by proteases, we used PEPTIDECUTTER (https://www.expasy.org/#proteome).
g. To predict the potential of glycosylation sites, we used NETNGLYC (https://services.healthtech.dtu.dk/services/NetNGlyc-1.0/).
h. To analyze protein domain motifs, we used PROSITE (https://prosite.expasy.org/).
i. To predict the location of proteins in cells, we use TARGETP (https://services.healthtech.dtu.dk/services/TargetP-2.0/).
j. To predict bearer location code, we use SIGNALP (https://services.healthtech.dtu.dk/services/SignalP-5.0/).
k. To show the 3D shape of the IL-6 protein structure, we used Swissmodel (https://swissmodel.expasy.org/)

3. RESULTS

3.1 Interleukin-6 Gene

The Interleukin-6 (IL-6) gene was obtained from the NCBI site with the code NC_000007.14 (22727200..22731998, complement). The official name of this gene in humans is Interleukin 6 with the symbol "IL 6". This gene is also sometimes written with the following symbols: IL-6; HGF; CDF; HSF; IFN-beta-2; IFNB2; BSF-2; BSF2.[7]

The location of the IL-6 gene in humans is located on chromosome 7p15.3. This means that the gene that carries the genetic information for IL-6 is on chromosome no. 7, position of the short arm (p), 1st arm, band 5

and sub band 3. Figure 1 shows a schematic of the location of the IL-6 gene nucleotides. After analyzing the sequence with the identity code NG_011640.1, the IL-6 gene has a total length of 4,799 base pairs (bp), consisting of 5 exons and 4 introns. The IL-6 gene sequence reference can be seen in Figure 2.[8] Interleukin-6 has three isoforms, namely interleukin-6 isoform 1 precursor (NP_000591.1)[9], interleukin-6 isoform 2 (NP_001305024.1)[10], and interleukin-6 isoform 3 (NP_001358025.1)[11]. The IL-6 gene has 6 transcript variants as seen in Figure 3.[12]

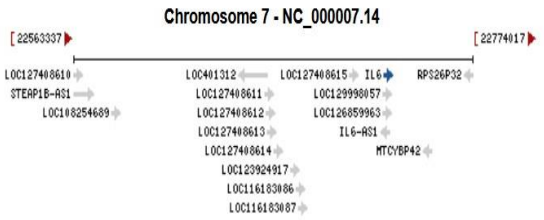

Figure 1. Gene location of IL-6.[7]

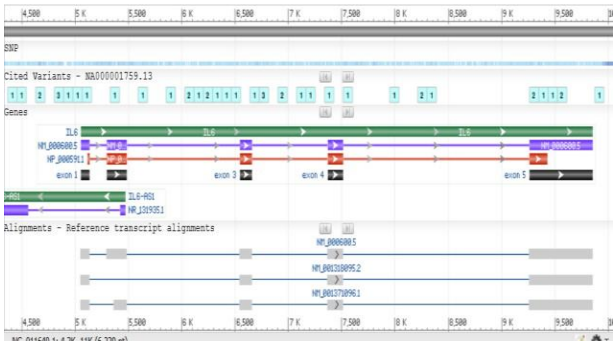

Figure 2 Ref Sequence (NG_012088).[8]

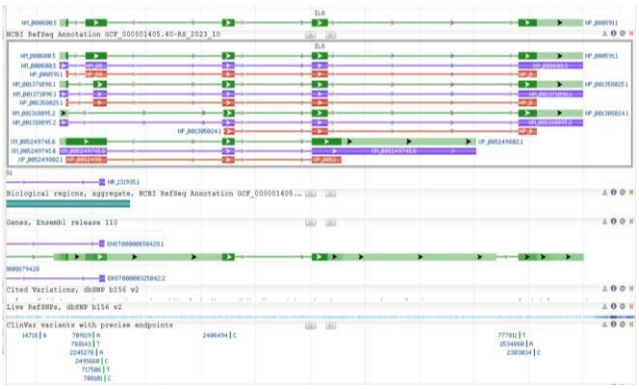

Figure 3. Gene annotation of IL-6[12]

3.2 Interleukin-6 Protein

The IL-6 protein (AAD13886.1) has molecular weight is around 21-26 kDa. It has a length of 212 amino acids (aa)[13], with an N-terminal signal peptide of 29 amino acids and a four-helix bundle organized in an up-up-down-down topology. With lengthy loops linking the helices, each helix has 20–25 residues and is referred to historically as A through D, corresponding to the N terminus to the C terminus. The crystal structure only shows the final seven residues of this flexible tail; the remaining twenty residues in the N-terminal do not appear to adopt any secondary structure. A common characteristic of the long chain family of four-helix bundle proteins is the presence of a fifth small helix, consisting of 12 residues (amino acids 141–152), in the long loop between helices C and D.[14] The IL-6 protein sequence can be seen in figure 4.

```
ORIGIN
        1 mnsfstsafg pvafslglll vlpaafpapv ppgedskdva aphrqpltss eridkqiryi
       61 ldgisalrke tcnksnmces skealaennl nlpkmaekdg cfqsgfneet clvkiitgll
      121 efevyleylq nrfesseeqa ravqmstkvl iqflqkkakn ldaittpdpt tnaslltklq
      181 aqnqwlqdmt thlilrsfke flqsslralr qm
//
```

Figure 4. Protein Sequence of IL-6[13]

3.3 Secondary Structure Prediction (PSIPRED)

Protein IL-6 secondary structure is mostly in helical-shaped amino acid form, followed by coil-shaped amino acid form (Figure 6).[15]

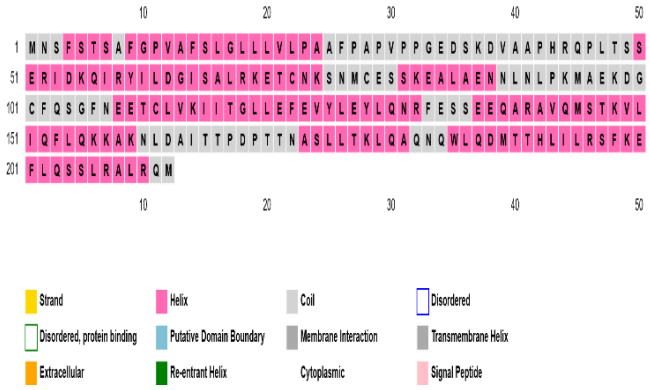

Figure 6. Secondary structures of IL-6 protein

3.4 Physical-Chemical Characteristics (PROTPARAM)

Physical-chemical characteristics of protein divided into number of amino acids, molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficients, estimated half-life, instability index, aliphatic index, and hydroplasticity (table 1).[16]

Table 1. Physical-chemical characteristics of IL-6 protein

Characteristic	**Protein IL-6 normal**
Number of amino acids	212
Molecular weight	23718.22
Theoretical pI	6.17
Amino acid composition	Ala (A) 18 8.5% Arg (R) 9 4.2% Asn (N) 11 5.2% Asp (D) 8 3.8% Cys (C) 4 1.9% Gln (Q) 14 6.6% Glu (E) 16 7.5% Gly (G) 7 3.3% His (H) 2 0.9% Ile (I) 9 4.2% Leu (L) 28 13.2% Lys (K) 14 6.6% Met (M) 6 2.8% Phe (F) 11 5.2% Pro (P) 11 5.2% Ser (S) 19 9.0% Thr (T) 13 6.1% Trp (W) 1 0.5% Tyr (Y) 3 1.4% Val (V) 8 3.8% Pyl (O) 0 0.0% Sec (U) 0 0.0% (B) 0 0.0% (Z) 0 0.0% (X) 0 0.0% **Total number of negatively charged residues (Asp + Glu):** 24 **Total number of positively charged residues (Arg + Lys):** 23
Atom composition	Carbon C 1049 Hydrogen H 1685 Nitrogen N 283 Oxygen O 321 Sulfur S 10 **Formula:** $C_{1049}H_{1685}N_{283}O_{321}S_{10}$ **Total number of atoms:** 3348
Extinction coefficients	Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water. Ext. coefficient 10220 Abs 0.1% (=1 g/l) 0.431, assuming all pairs of Cys residues form cystines Ext. coefficient 9970 Abs 0.1% (=1 g/l) 0.420, assuming all Cys residues are reduced
Estimated half-life	The N-terminal of the sequence considered is M (Met). The estimated half-life is: • 30 hours (mammalian reticulocytes, in vitro). • >20 hours (yeast, in vivo). • >10 hours (Escherichia coli, in vivo).
Instability Index	57.70 This classifies the protein as unstable.
Aliphatic Index	87.50
Grand average of hydropathicity (GRAVY):	-0.271

3.5 Trans-membrane Protein Topology (TMHMM)

Most of the IL-6 protein is located outside the cell membrane. The transmembrane position is at amino acid positions approximately 30-202, a small portion is inside the cell (Figure 7).[17]

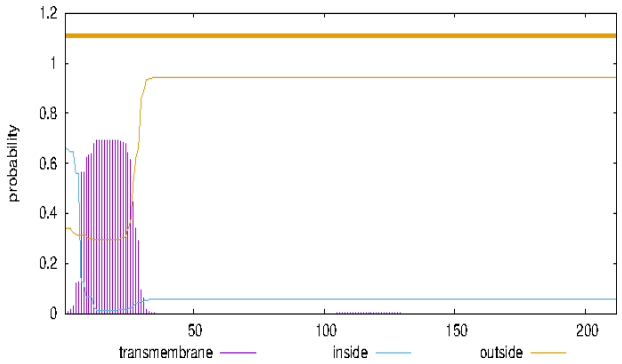

Figure 7. TMHMM of IL-6 protein[17]

3.6 Hydrophobicity of IL-10 (PROTSCALE)

Hydrophobicity analysis using the the scale Hydropath. / Kyte & Doolittle by PROTSCALE method, and converted to scatter form with the help of Microsoft Excel, showed hydrophobicity score each position (Figure 8).[18]

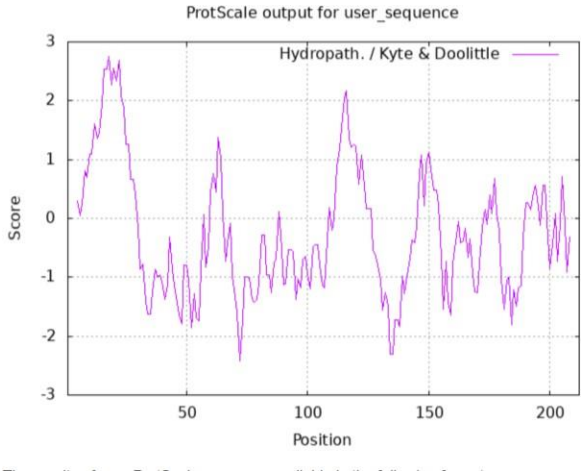

Figure 8. PROTSCALE of IL-6 protein [18]

3.7 Prediction of Cleavage by Protease (PEPTIDECUTTER)

The normal IL-6 proteins have glycosylation site at the 73 and 172 amino acid positions (Figure 9). [19]

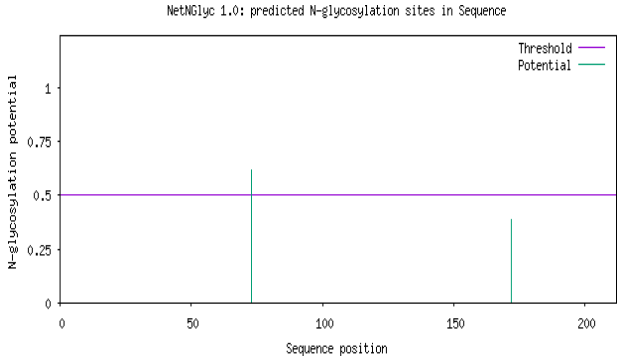

Figure 9. Position of glycosylation sites of IL-6 protein[19]

3.8 Protein Domain Motif (PROSITE)

The normal IL-6 protein domain motif is: CfqsgfneetClvkiitGLleFevyL at positions 101-126.[20]

3.9 Protein Location Prediction in Cells (TARGETP)

The prediction of the target location of normal IL-6 proteins in cells mostly in the signal peptide pathway (0.971), a little in the other (0.0029), and none in mitochondrial transfer peptide pathway (0). This is related to the function of the IL-10 protein as a signal peptide that must be on the cell surface (figure 10).[21]

Figure 10. Target location of normal IL-6 protein[21]

3.10 Location Code Carrier Prediction (SIGNALP)

Predicted cleavage site is between position 27 and 28 of amino acid (AFP-AP), with max cleavage score: 0.257 (position 28). Peptide signals are at position 1-27 of amino acid (mean S; 0.887), with the S peak at position 19 of amino acid (0.963). With a high mean-S and D-score, it indicates that the IL-6 protein and the mutant are included in the "signal peptide" group (figure 11).[22]

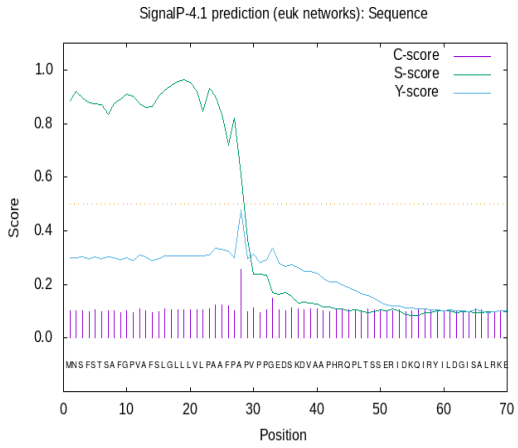

Figure 11. Signal Prediction of IL-6[22]

3.11 Image of Protein Structure

The 3-dimensional shape of the IL-6 protein structure is referenced from the Swiss model. The protein structure template refers to AlphaFold DB model of IL6_HUMAN (figure 12).[23]

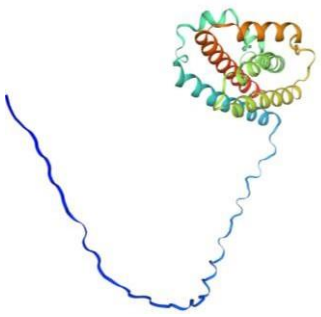

Figure 12. The 3-dimensional shape of IL-6[23]

4. DISCUSSION

Human IL-6 gene is located on chromosome 7p15.3. The IL-6 protein is 212amino acids (aa) long, secondary structure is mostly in helical-shaped amino acid form. The protein is unstable, with half time 30 hours. Most of the IL-6 protein is located outside the cell membrane. It has glycosylation site at the 73 and 172 amino acid positions, with domain motif at positions 101-126. The prediction of the target location is mostly in the signal peptide pathway. Predicted cleavage site is between position 27 and 28 of amino acid. The bioinformatics analysis is very useful in determining the biomolecular characteristics of IL-6.

5. CONCLUSION

The author would like to personally thank Prof. Dr. dr. Irfanuddin, Sp.KO, M.Pd.Ked (professor at medical faculty, Universitas Sriwijaya) who introduced the author to various tools in the field of molecular bioinformatics.

REFERENCES

1. Aliyu M, Zohora FT, Anka AU, Ali K, Maleknia S, Saffarioun M, et al. Interleukin-6 cytokine: An overview of the immune regulation, immune dysregulation, and therapeutic approach. Int Immunopharmacol. 2022;111:109130.
2. Trovato M, Sciacchitano S, FAccIOLÀ A, Valenti A, Visalli G, Di Pietro A. Interleukin-6 signalling as a valuable cornerstone for molecular medicine. Int J Mol Med. 2021;47(6):1–9.
3. Velazquez-Salinas L, Verdugo-Rodriguez A, Rodriguez LL, Borca M V. The role of interleukin 6 during

viral infections. Front Microbiol. 2019;10:1057.

4. Rose-John S. Interleukin-6 family cytokines. Cold Spring Harb Perspect Biol. 2018;10(2):a028415.
5. Choy E, Rose-John S. Interleukin-6 as a multifunctional regulator: inflammation, immune response, and fibrosis. J Scleroderma Relat Disord. 2017;2(2_suppl):S1–5.
6. Irfannuddin. Karakteristik Biomolekuler Mutasi Amyloid Precursor Protein Sebagai Penyebab Penyakit Alzheimer yang Dianalisis Melalui Media Bioinformatika [Internet]. Vol. 1, Sriwijaya Journal of Medicine. 2018. Available from: www.ncbi.org
7. National Center for Biotechnology Information (NCBI). IL6 interleukin 6 [Homo sapiens (human) [Internet]. 2023 [cited 2023 Dec 3]. Available from: https://www.ncbi.nlm.nih.gov/gene/3569
8. National Center for Biotechnology Information (NCBI). Homo sapiens interleukin 6 (IL6), RefSeqGene on chromosome 7 [Internet]. 2023 [cited 2023 Dec 3]. Available from: https://www.ncbi.nlm.nih.gov/gene/3569#reference-sequences
9. National Center for Biotechnology Information (NCBI). Interleukin-6 isoform 1 precursor [Homo sapiens] [Internet]. 2023 [cited 2023 Dec 3]. Available from: https://www.ncbi.nlm.nih.gov/protein/NP_000591.1
10. National Center for Biotechnology Information (NCBI). Interleukin-6 isoform 2 [Homo sapiens] [Internet]. 2023 [cited 2023 Dec 3]. Available from: https://www.ncbi.nlm.nih.gov/protein/NP_001305024.1
11. National Center for Biotechnology Information (NCBI). Interleukin-6 isoform 3 [Homo sapiens] [Internet]. 2023 [cited 2023 Dec 3]. Available from: https://www.ncbi.nlm.nih.gov/protein/NP_001358025.1
12. National Center for Biotechnology Information (NCBI). Genome Data Viewer [Internet]. 2023 [cited 2023 Dec 3]. Available from: https://www.ncbi.nlm.nih.gov/genome/gdv/browser/gene/?id=3569
13. National Center for Biotechnology Information (NCBI). Interleukin-6 [Homo sapiens] [Internet]. 2023 [cited 2023 Dec 3]. Available from: https://www.ncbi.nlm.nih.gov/protein/AAD13886.1
14. Gelinas AD, Davies DR, Edwards TE, Rohloff JC, Carter JD, Zhang C, et al. Crystal structure of interleukin-6 in complex with a modified nucleic acid ligand. Journal of Biological Chemistry. 2014;289(12):8720–34.
15. UCL Department of Computer Science. Secondary Structure of IL-6 [Internet]. 2023 [cited 2023 Dec 3]. Available from: http://bioinf.cs.ucl.ac.uk/psipred/&uuid=f9eadaa8-5b80-11ec-bcf4-00163e100d53

16. Expasy. ProtParam Analysis of IL-6 [Internet]. 2023 [cited 2023 Dec 3]. Available from: https://web.expasy.org/cgi-bin/protparam/protparam

17. DTU Health Tech. TMHMM Result of IL-6 [Internet]. 2023 [cited 2023 Dec 3]. Available from: https://services.healthtech.dtu.dk/cgi-bin/webface2.cgi?jobid=656C6EA0000073A37466CE2E&wait=20

18. Expasy. ProtScale Analysis of IL-6 [Internet]. 2023 [cited 2023 Dec 3]. Available from: https://web.expasy.org/cgi-bin/protscale/protscale.pl?1

19. DTU Health Tech. Prediction of Cleavage by Protease of IL-6 [Internet]. 2023 [cited 2023 Dec 4]. Available from: https://services.healthtech.dtu.dk/cgi-bin/webface2.cgi?jobid=656DED51000007A324181DB7&wait=20

20. Expasy. Protein Domain Motif of IL-6 [Internet]. 2023 [cited 2023 Dec 4]. Available from: https://prosite.expasy.org/cgi-bin/prosite/scanprosite/ScanView.cgi?scanfile=560417955488.scan.gz

21. DTU Health Tech. Target location of normal IL-6 protein [Internet]. 2023 [cited 2023 Dec 4]. Available from: https://services.healthtech.dtu.dk/cgi-bin/webface2.cgi?jobid=656DF50B000058558CD5465D&wait=20

22. DTU Health Tech. Signal Prediction of IL-6 [Internet]. 2023 [cited 2023 Dec 4]. Available from: https://services.healthtech.dtu.dk/cgi-bin/webface2.cgi?jobid=656DF9E7000009CD80FF2733&wait=20

23. Biozentrum. The 3-dimensional shape of IL-6 [Internet]. 2023 [cited 2023 Dec 4]. Available from: https://swissmodel.expasy.org/interactive/qq8UR9/templates/