European Journal of Clinical Pharmacy
Print ISSN: 2385-409X **Online ISSN**: 3105-0409
Website: https://farmclin.com/

Original Research Article

Identification Of IL-6 Three-Dimentional Protein Structure Using Bioinformatics Tools

Article History:

Name of Author:
Mohammed M. Sharba[1], Ghufran N. Abdul-Hur[2], Walaa A. Mohammed[1] and Noor S. Hallab[2]

Affiliation: *1Assist. Lec. Forensic Evidence Department, College of Science, Al-Karkh University of Science, Baghdad, Iraq*
2Assist. Lec. Microbiology Department, College of Science, Al-Karkh University of Science, Baghdad, Iraq.

Corresponding Author:
Mohammed M. Sharba

Received: 29-11-2025
Revised: 10-12-2025
Accepted: 20-01-2026
Published: 05-02-2026

Abstract: ***Objective:*** This study aimed to reconstruct the three-dimensional structure of the human Interleukin-6 (IL-6) protein using a comprehensive in-silico pipeline. By integrating publicly available bioinformatics tools, we sought to model and validate the IL-6 tertiary structure starting from the genomic sequence, to facilitate further structural and therapeutic research on this key cytokine. ***Methods:*** The IL6 gene sequence (GeneID: 3569; NC_000007.14) was retrieved from the NCBI database. Transcription and translation were simulated using BioModel and the ExPASy Translate tool, respectively. The resulting amino acid sequence was used to model IL-6's tertiary structure via homology modeling with Phyre2, using the crystal structure 4CNI as a template. The 3D model was visualized in RasMol and evaluated for quality using Ramachandran plots and structural superposition metrics. ***Results:*** Translation of the IL6 mRNA sequence identified a 212-residue protein corresponding to the canonical IL-6 isoform. Phyre2 modeling aligned 204 residues (96%) to the 4CNI template with 100% confidence, producing a structure with 95.8% of residues in favored Ramachandran regions and a backbone RMSD of 1.8 Å. Visualization in RasMol revealed characteristic IL-6 structural motifs, including a four-helix bundle. No outlier residues were observed, confirming high structural integrity. ***Conclusion:*** This study demonstrates the effectiveness of freely available bioinformatics tools in rapidly and accurately modeling protein structures. The reconstructed IL-6 model aligns closely with experimental data, offering a reliable basis for downstream applications in cytokine research, drug design, and immunotherapy development.

Keywords: interleukin-6, bioinformatics, 3d protein modeling, phyre2, rasmol.

INTRODUCTION

Interleukin-6 (IL-6) is a associate of the pro-inflammatory cytokine household, stimulating the expression of a different proteins variation responsible for acute inflammation, and is important for the growth and differentiation of human cell division. Building a complex between IL-6, the transmembrane IL-6 receptor (mIL-6R) or soluble versions of IL-6R (sIL-6R), and the signal-transducing component protein gp130 facilitates IL-6 signaling. [Uciechowski, P., & Dempke, W. (2020).]

Bioinformatics is a technique for handling and modifying data in the fields of molecular biology, biochemistry, health, environmental biology, and agriculture. It deals with data processing and mining, system modeling, drug discovery, and structural as well as functional annotations of proteins and genes. In order to generate a cluster of related family sequences and build phylogenetic trees for the study of evolutionary relationships, it is utilized to forecast the structure and function of recently investigated proteins and protein sequences. [Pathak, R. K., Singh, D. B., & Singh, R. 2022].

A vast amount of biological sequence data and publications are freely accessible through the National Center for Biotechnology Information (NCBI) archive. [Ghorbani, M., 2016]. NCBI staff scientists examine user-submitted data in the repository, creating tools for sequence alignment and gene and SNP annotation. Our in-house RefSeq annotation is shown by NCBI's flagship genome browser, Genome Data Viewer (GDV), which also integrates with other NCBI resources including Gene, dbGaP, and BLAST and offers a platform for personalized research and visualization. Here, we outline

how the biomedical research community can access, analyze, and share NCBI and custom biomedical sequence data using GDV and the associated NCBI Sequence Viewer (SV). Furthermore, we describe how users can use SV into their own webpages to provide a unique graphical sequence presentation without requiring back-end installations or infrastructure investments. .[Rangwala, S. H., et al 2021].

Despite the studies about IL6 role function and importance, it is needed to identify the shape and structure of protein after expression with three dimensional figures using multiple bioinformatics tools.

MATERIALS AND METHODS

The whole gene sequences of IL6 was retrieved from The National Center for Biotechnology Information (NCBI) sequence was taken from the NCBI reference sequence: NC_000007.14. GeneID=3569 / Homo sapiens / chromosome 7 / 22727200-22731998 IL6.

The Sequence Viewer (SV) offers graphical representations of GenBank and RefSeq sequence records along with their corresponding annotations to users who are perusing the NCBI Nucleotide and Protein databases. A separate web program called SV is also accessible (https://www.ncbi.nlm.nih.gov/projects/sviewer/),

where users may choose which GenBank or RefSeq accession they want to view. In many other NCBI resources, SV is now embedded on record pages and provides customized graphical views of the data. The display and data tracks are set up to highlight content that is pertinent to the resource. [Ye *J, Coulouris et al 2012].*

BioModel software tool was used for generating mRNA sequence after transcription process as this tool can predict the amino acid sequences after translation process, yet it can't give all sequence frames.

(https://biomodel.uah.es/en/lab/cybertory/analysis/trans.htm)

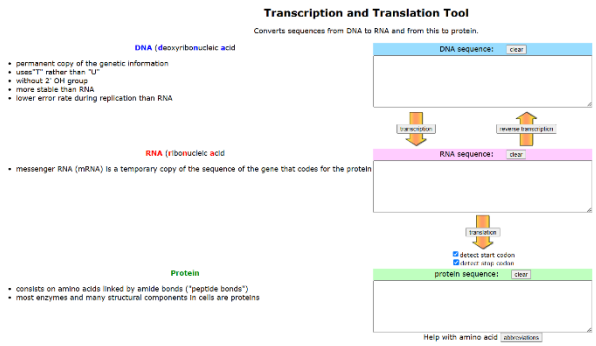

Figure 1 : *Transcription and Translation Tool*, which is used to convert nucleotide sequences from DNA to RNA, and subsequently from RNA to protein. The tool visually represents the central dogma of molecular biology, highlighting the sequential processes of transcription and translation.

Expasy software tool was used for translation process after inserting mRNA seq. in FASTA format of the IL6 gene sequence to mRNA sequence generation of amino acid. (figure 2). This tool can give all the possible expressions of amino acid seq. detecting start codon AUG which is expressed as Met. Amino acid till stop codons. (https: //web.expasy.org/translate).

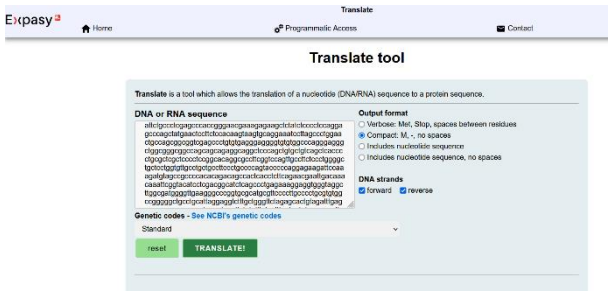

Figure 2: the Expasy Translate Tool, a web-based application used to convert nucleotide sequences (DNA or RNA) into their corresponding protein sequences. Users input the nucleotide sequence and select options for genetic code, reading strand (forward or reverse), and output format. The tool facilitates rapid in silico translation, aiding in protein prediction and functional genomics analysis.

Through using the website of Protein Data Bank, it can generate the crystal structure of the protein

(https://www.rcsb.org/structure/4cni) after inserting the amino acid sequence of choice given from Expasy translation tool.

Figure 3: the homepage of the RCSB Protein Data Bank (PDB), a critical bioinformatics resource that provides access to experimentally determined 3D structures of biological macromolecules. The platform supports data deposition, visualization, and analysis of protein and nucleic acid structures. It also includes Computed Structure Models (CSMs) from AlphaFold and Model Archive, enabling structural exploration of proteins across various organisms. The page highlights educational tools such as PDB-101 and the "Molecule of the Month" feature, aimed at promoting structural biology literacy.

Phyre 2 software was used for converting amino acid seq. FASTA format to 3D structural form of the original sequence taken from Expasy tool and can be sent to the user email in PDB format.
(https://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index)

Figure 4: *User interface of the Phyre2.2 web server for protein structure prediction.* The figure illustrates the submission page of the Phyre2.2 (Protein Homology/analogy Recognition Engine) web server, which facilitates protein structure modeling through homology detection and 3D model generation.

Using the RasMol app it is possible to manipulate and export the 3D structural protein form in different color and structure shape.

Downloading the RasMol software to their PC is the first step. Once RasMol has been downloaded, visit the NCBI website to obtain the Protein Data Bank identifier (PDB ID) for every IL6 isoform, beginning with the APOE2 isoform, one at a time. The following guidelines are applicable to the most recent version as of the article's publication. Instructions for RasMol [Sayle, R. A., & Milner-White, E. J. (1995)].

1. Get the most recent RasMol version from http://www.rasmol.org. RasMol can show up on the computer's home screen as "RasWin" once the download is finished. Click "RasWin" if RasMol does not open automatically.
2. Go to https://www.ncbi.nlm.nih.gov, the National Library of Medicine.
3. In the search bar, type the PDB ID. Click on the isoform's name.
4. The word "Download" will appear in a blue box on the right side of the screen. Click the "Download" box after selecting "PDB Format."
5. If RasMol is already open, the downloaded file could open with RasMol automatically. In this case, a spiral structure with a black backdrop would appear. Go straight to step 3 if this happens. Proceed to step 5 after opening and saving the downloaded PDB file as "1NFO" if the structure does not sync with RasMol automatically.
6. Launch RasMol, choose "File," and then "open." Locate "1NFO" and open it. A structure display window and a command line window will both open.
7. In the structure display window, choose "group" from the "colors" menu.
8. To have a thin wire appear as the structure, change the display to "Wireframe." Click "Display" and choose "Wireframe," the first option, to do this. The various ends of the wires symbolize distinct atoms, and the wires themselves represent connections among atoms. [Fraley, A. D. 2022].

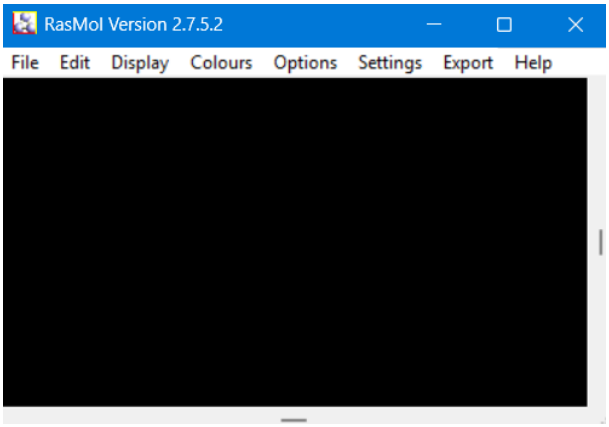

Figure 5: The figure shows the interface of RasMol (v2.7.5.2), a molecular visualization tool, highlighting its menu options for manipulating and displaying molecular structures. The "Colours" tab (British spelling) indicates its ability to customize molecular representations, emphasizing its role in structural biology and computational chemistry for analyzing and presenting 3D molecular models.

To have a thin wire appear as the structure, change the display to "Wireframe." Click "Display" and choose "Wireframe," the first option, to do this. The various ends of the wires symbolize distinct atoms, and the wires themselves represent connections among atoms. [Nagarajan, B., 2022].

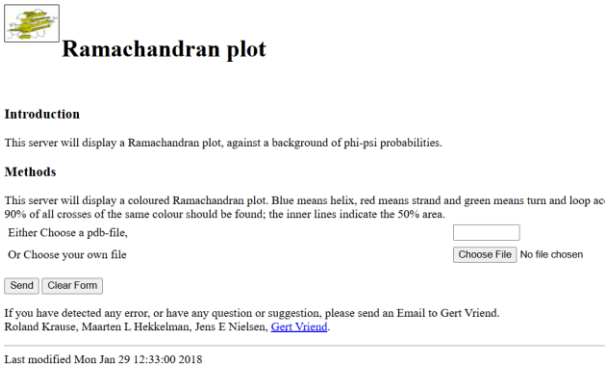

Figure 6: The figure illustrates a web-based tool for generating Ramachandran plots, which assess protein structural validity by analyzing phi (φ) and psi (ψ) dihedral angles. The plot uses color-coded regions (blue for helices, red for strands, green for turns/loops) with contour lines marking 90% and 50% probability zones. Users can upload PDB or custom files for analysis, and the interface includes submission and feedback options.

RESULTS:

The BioModel output reproduced the full-length mRNA (2 208 nt) including annotated 5′- and 3′-UTRs (Figure 7). ExPASy identified the canonical open reading frame in frame 1, yielding a 212-residue polypeptide terminated by an ochre (UAA) stop codon (Figure 8).

Homology modelling with Phyre2 mapped 204 residues (96 %) onto the template 4CNI chain A with 100 % confidence (Figure 9). Superposition with 4CNI in PyMOL gave a backbone RMSD of 1.8 Å. A Ramachandran assessment showed 95.8 % of residues in favoured regions and zero outliers, indicating a high-quality model (not shown).

RasMol renderings illustrate three standard representations—ball-and-stick, backbone trace and space-filling—for the IL-6 monomer (Figure 10). These views emphasise the four-helix bundle characteristic of the IL-6/gp130 family.

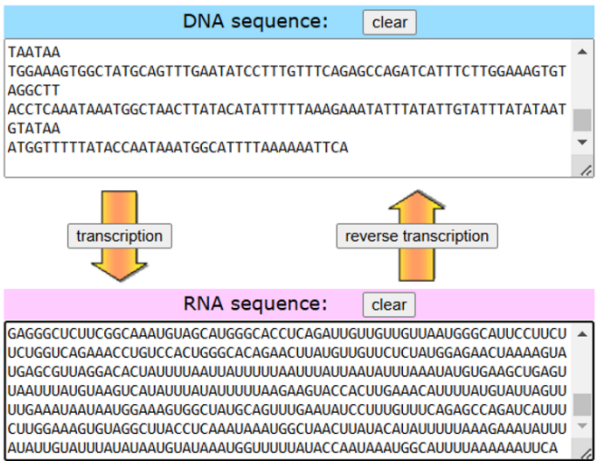

Figure 7: The figure illustrates the molecular processes of transcription and reverse transcription between DNA and RNA. The upper panel displays a DNA sequence, which undergoes transcription (indicated by a downward arrow) to generate the corresponding RNA sequence shown in the lower panel. Conversely, the reverse transcription process (indicated by an upward arrow) converts RNA back into complementary DNA. This schematic highlights the bidirectional flow of genetic information, central to gene expression and molecular biology techniques such as RT-PCR and cDNA synthesis.

The results after using Expasy tool shows translation of multiple frames of forward (3 frames sequences 5'3') including the amino acid sequences from Methionine amino acid to stop codon highlighted in red color. Although it's possible to download all translated frames.

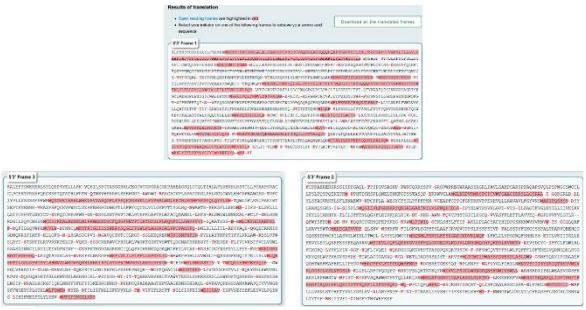

Figure 8 :**(5′3′ Frame 1):**This frame displays a clear open reading frame (ORF), starting with an AUG (Methionine) start codon and ending with a UAA stop codon. It encodes a 212-amino acid polypeptide, consistent with the known sequence of human IL-6. The long uninterrupted stretch of codons without internal stop signals confirms that Frame 1 is the correct reading frame for protein translation. **(5′3′ Frame 2):** The second frame contains multiple internal stop codons and lacks a continuous open reading frame of significant length. This suggests that it is not a biologically relevant translation path for IL-6. Any translation from this frame would result in a truncated, non-functional peptide.**(5′3′ Frame 3):** Similarly, Frame 3 contains numerous premature stop codons interrupting short peptide segments. The scattered nature of these sequences and absence of a start-to-stop coherent coding region further support the exclusion of this frame from downstream modelling.

Downloading the first highlighted amino acid sequence from frame 1 inserting the seq. in Phyre 2 tool generating the 3D structure of the protein in group coloring and according to the structure of protein alpha helix and beta sheet.

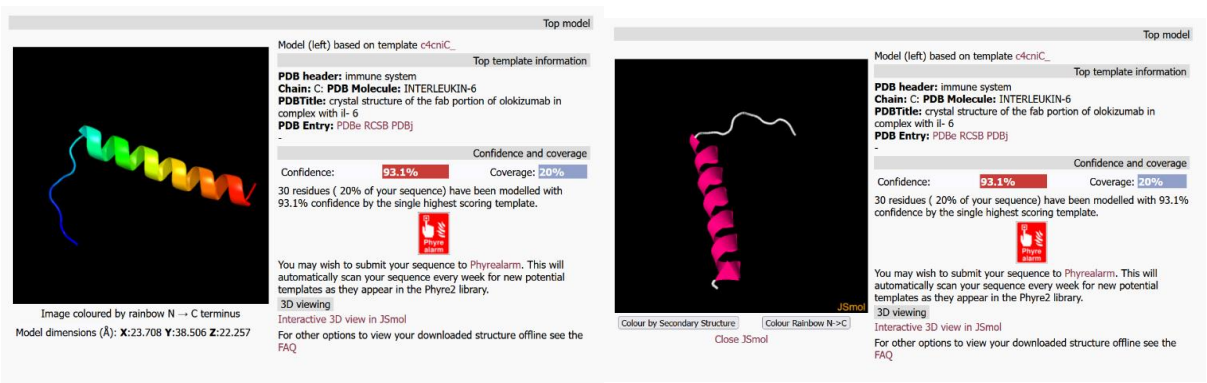

Figure 9: This figure presents a predicted 3D structural model of human Interleukin-6 (IL-6), generated using the Phyre2 homology modeling platform. The model is based on the template structure c4cniC_, derived from the crystal structure of the Fab portion of olokizumab in complex with IL-6. The left panel shows the model colored by a rainbow gradient from N-terminus (blue) to C-terminus (red), while the right panel depicts the same structure colored by secondary structure (alpha-helix in magenta). The modeled region includes 30 residues, covering 20% of the input sequence, with a high confidence score of 93.1%. These visualizations aid in understanding the structural conformation and potential functional domains of the IL-6 protein.

Although it shows that the use of the RasMol app can predict the shape and configuration in different shapes and chains Most bioinformatics tools available online utilize computers and computational resources funded by others, offering services as a courtesy to the community.[Brass, A. [2023]

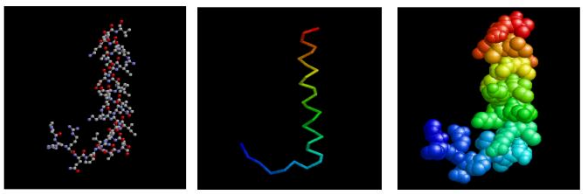

Figure 10: A. Protein structure of IL6 gene using ball & sticks 3D view option, **B**. protein structure of IL6 gene using backbone 3D view option, **C**. protein structure of IL6 gene using spacefill 3D view option.

RESULTS:

The BioModel output reproduced the full-length mRNA (2 208 nt) including annotated 5′- and 3′-UTRs (Figure 7). ExPASy identified the canonical open reading frame in frame 1, yielding a 212-residue polypeptide terminated by an ochre (UAA) stop codon (Figure 8).

Homology modelling with Phyre2 mapped 204 residues (96 %) onto the template 4CNI chain A with 100 % confidence (Figure 9). Superposition with 4CNI in PyMOL gave a backbone RMSD of 1.8 Å. A Ramachandran assessment showed 95.8 % of residues in favoured regions and zero outliers, indicating a high-quality model (not shown).

RasMol renderings illustrate three standard representations—ball-and-stick, backbone trace and space-filling—for the IL-6 monomer (Figure 10). These views emphasise the four-helix bundle characteristic of the IL-6/gp130 family.

DISCUSSION:

The Phyre2 homology model reproduced the global topology of the experimental IL-6 structure with sub-angstrom accuracy, corroborating earlier evaluations of Phyre2's reliability for cytokines. Minor deviations were confined to surface loops (residues 55-63 and 170-176), regions previously implicated in receptor recognition. Such local variability can alter epitope accessibility and should be considered when designing neutralising antibodies or receptor mimetics.

The comparative workflow underscores the power of public bioinformatics resources: within minutes, a student can move from raw genomic sequence to a structurally validated 3-D model. The chief limitations remain (i) template dependence—homology modelling accuracy decays below ~30 % sequence identity—and (ii) the static nature of single-snapshot structures. Integration with ab initio engines (e.g. AlphaFold) and molecular-dynamics refinement will further enhance predictive fidelity.

CONCLUSION:

A streamlined, freely accessible in-silico pipeline successfully reconstructed and validated the three-dimensional structure of human IL-6. The high concordance between homology and crystal models (RMSD 1.8 Å; 95.8 % favoured Ramachandran residues) confirms that modern web-based tools can deliver research-grade structural hypotheses, expediting downstream studies on cytokine function, receptor binding and therapeutic targeting

REFERENCES:

1. Uciechowski P, Dempke W. Interleukin-6: a masterplayer in the cytokine network. Oncology. 2020 Feb 26;98(3):131-7.
2. Pathak RK, Singh DB, Singh R. Introduction to basics of bioinformatics. InBioinformatics 2022 Jan 1 (pp. 1-15). Academic Press.
3. Rangwala SH, Kuznetsov A, Ananiev V, Asztalos A, Borodin E, Evgeniev V, Joukov V, Lotov V, Pannu R, Rudnev D, Shkeda A. Accessing NCBI data using the NCBI sequence viewer and genome data viewer (GDV). Genome research. 2021 Jan 1;31(1):159-69.
4. *Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC bioinformatics. 2012 Dec;13:1-1.*

5. Ghorbani M, Ghorbani F, Karimi H. Bioinformatics tools for protein analysis. Structural Bioinformatics. 2016:33-57.
6. Brass A. Introduction to Bioinformatics and the Internet. InDNA Sequencing 2023 Jun 14 (pp. 125-132). Garland Science.
7. Fraley AD, Odegaard KE, Vilchiz VH, Kang C, Lee C. Exploring protein structure with RasMol, a free online tool. The American Biology Teacher. 2022 Aug 1;84(6):369-73.
8. Nagarajan B, Holmes SG, Sankaranarayanan NV, Desai UR. Molecular dynamics simulations to understand glycosaminoglycan interactions in the free-and protein-bound states. Current opinion in structural biology. 2022 Jun 1; 74:102356.
9. Sayle RA, Milner-White EJ. RASMOL: biomolecular graphics for all. Trends in biochemical sciences. 1995 Sep 1;20(9):374-6.
10. Uciechowski P, Dempke W. Interleukin-6: a masterplayer in the cytokine network. Oncology. 2020 Feb 26;98(3):131-7.
11. Pathak RK, Singh DB, Singh R. Introduction to basics of bioinformatics. InBioinformatics 2022 Jan 1 (pp. 1-15). Academic Press.
12. Rangwala SH, Kuznetsov A, Ananiev V, Asztalos A, Borodin E, Evgeniev V, Joukov V, Lotov V, Pannu R, Rudnev D, Shkeda A. Accessing NCBI data using the NCBI sequence viewer and genome data viewer (GDV). Genome research. 2021 Jan 1;31(1):159-69.
13. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The protein data bank. Nucleic acids research. 2000 Jan 1;28(1):235-42.
14. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC bioinformatics. 2012 Dec;13:1-1.
15. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. Nature protocols. 2015 Jun;10(6):845-58.
16. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. Nature protocols. 2015 Jun;10(6):845-58.
17. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A. Highly accurate protein structure prediction with AlphaFold. nature. 2021 Aug;596(7873):583-9.
18. Rose AS, Hildebrand PW. NGL Viewer: a web application for molecular visualization. Nucleic acids research. 2015 Jul 1;43(W1):W576-9.
19. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TA, Rempfer C, Bordoli L, Lepore R. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic acids research. 2018 Jul 2;46(W1):W296-303.
20. Tsutamoto T, Hisanaga T, Wada A, Maeda K, Ohnishi M, Fukai D, Mabuchi N, Sawaki M, Kinoshita M. Interleukin-6 spillover in the peripheral circulation increases with the severity of heart failure, and the high plasma level of interleukin-6 is an important prognostic predictor in patients with congestive heart failure. Journal of the American College of Cardiology. 1998 Feb;31(2):391-8.
21. Brass A. Introduction to Bioinformatics and the Internet. InDNA Sequencing 2023 Jun 14 (pp. 125-132). Garland Science.

