In Silico Pharmacology (2024) 12:84
https://doi.org/10.1007/s40203-024-00252-x

ORIGINAL RESEARCH

Investigation of therapeutic potential of the IL-24-p20 fusion protein against breast cancer: an in-silico approach

Shahnila Qureshi[1] · Nadeem Ahmed[4] · Hafiz Muhammad Rehman[1,2] · Muhammad Imran Amirzada[3] · Fiza Saleem[1,2] · Kainat Waheed[1,2] · Afeefa Chaudhry[5] · Iram Kafait[6] · Muhammad Akram[1] · Hamid Bashir[1]

Received: 20 June 2024 / Accepted: 13 August 2024
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

Abstract

Targeted delivery of therapeutic anticancer chimeric molecules enhances drug efficacy. Numerous studies have focused on developing novel treatments by employing cytokines, particularly interleukins, to inhibit the growth of cancer cells. In the present study, we fused interleukin 24 with the tumor-targeting peptide P20 through a rigid linker to selectively target cancer cells. The secondary structure, tertiary structure, and physicochemical characteristics of the constructed chimeric IL-24-P20 protein were predicted by using bioinformatics tools. In-silico analysis revealed that the fusion construct has a basic nature with 175 amino acids and a molecular weight of 20 kDa. By using the Rampage and ERRAT2 servers, the validity and quality of the fusion protein were evaluated. The results indicated that 93% of the chimeric proteins contained 90.1% of the residues in the favoured region, resulting in a reliable structure. Finally, docking and simulation studies were conducted via ClusPro and Desmond Schrödinger, respectively. Our results indicate that the constructed fusion protein exhibits excellent quality, interaction capabilities, validity, and stability. These findings suggest that the fusion protein is a promising candidate for targeted cancer therapy.

Keywords Breast cancer · IL-24 · P20 · Chimeric protein · Anticancer peptides · Bioinformatics · Cancer therapy

Introduction

Breast cancer (BC) is a prevalent illness that significantly impacts young women and is the leading cause of cancer-related deaths among females, accounting for 22.9% of all cancer types in women. In 2019, more than 3.8 million women with BC histories lived in the United States. Over 268,600 new cases of invasive BC were diagnosed among women living in the United States (DeSantis et al., 2019). From 2020 to 2030, the age-standardized incidence rate of breast cancer in women is expected to rise significantly, especially for those aged 45–49. In high-socio-demographic Index (SDI) regions, however, the incidence rate for women aged 40–49 is projected to decline. Death and disability-adjusted life year rates due to breast cancer are also anticipated to increase in high-SDI and high–middle-SDI regions, particularly for those aged 40–49 and 45–49 (Zhang et al., 2024). The incidence of breast cancer generally increases with age. Most cases are diagnosed in women aged 50 and older. Women aged 40–59 have the highest incidence rates of breast cancer. The risk continues to increase with age, peaking at around 70–74 years. In women under 40, breast

✉ Hamid Bashir
hamid.camb@pu.edu.pk

[1] Centre for Applied Molecular Biology, University of the Punjab, 87-West canal, Bank Road, Lahore 53700, Pakistan

[2] University Institute of Medical Laboratory Technology, Faculty of Allied Health Science, University of Lahore, Lahore 54590, Pakistan

[3] Department of Pharmacy, COMSATS University, Islamabad Abbottabad, Pakistan

[4] Centre of Excellence in Molecular Biology (CEMB), University of the Punjab, 87-West canal, Bank Road, Lahore 53700, Pakistan

[5] Department of Biology, Lahore Garrison University, Avenue 4, sector phase 6 DHA, Lahore, Pakistan

[6] Institute of Molecular Biotechnology, Graz University of Technology, Graz, Austria

Published online: 18 September 2024

cancer is less common but often more aggressive and diagnosed at a later stage. (DeSantis et al., 2019; Sung et al., 2021). Treatment options for breast cancer include chemotherapy, surgery, hormone therapy, and radiotherapy. Chemotherapy remains the most effective method of treatment for advanced or metastatic cancer (Thun et al., 2010). Most chemotherapy drugs, however, affect normal cells in addition to cancer cells, resulting in significant side effects and remarkable resistance. New treatment strategies are required to selectively target cancer cells and reduce the chances of resistance developing in cancer cells.

Cytokines have anticancer properties, regulate immune function, and play an essential role in disease prevention. Multiple cytokines including interleukin-2, interleukin-15, and interleukin-24, have been well studied as promising anticancer drugs. These cytokines have a limited role in cancer therapy since substantial dosages are needed to increase immune activity against cancer cells. It is therefore critical that novel anticancer drugs with increased specificity and selectivity be developed (Dong and Mumper 2010; Chakraborty and Rahman 2012). Human interleukin-24, also known as melanoma differentiation-associated gene-7/interleukin-24, is a new tumor suppressor gene that was previously classified as a member of the IL-10 cytokine gene family. Interleukin-24 is a cytokine that stimulates the immune system and modulates the immune system, without adversely affecting normal cells (Menezes et al., 2015; Jiang et al., 1995). IL-24 has intriguing properties that model it as an effective and perfect tool for cancer therapy, such as inhibiting angiogenesis (Nishikawa et al., 2004), modulating the immune system(Gao et al., 2008), serving as a monoclonal antibody, and serving as a chemotherapy agent (Emdad et al., 2006), and it has significant anticancer activity. IL-24 induces cancer cell death by binding to its receptor and activating Janus kinase (JAK)/signal transducer and activator of transcription (STAT) pathway. This interaction triggers a series of events leading to cancer cell death (Sauane et al., 2008). Previously, intracellular IL-24 was shown to induce stress in the endoplasmic reticulum (ER), which leads to mitochondrial apoptosis through the suppression of protein translation (Dent et al., 2010).

Conventional chemotherapy has substantial hazardous side effects due to its low selectivity towards cancer cells. An efficient alternative method of selectively delivering high doses of chemotherapeutic drugs to tumor areas while sparing healthy tissues involves the use of tumor-targeting peptides (Bidwell 2012). Peptides are promising candidates for drug design due to their low toxicity, good solubility, high selectivity, ability to penetrate tumors, versatility in being used alone or in combination with other chemotherapeutics, and capacity to target various types of cancers. (Thundimadathil, Craik et al., 2013). The construction of chimeric proteins by fusing cytokines with tumor-targeting peptides has the potential to increase onsite delivery with the least amount of side effects. A notable bioactive peptide, P20 [-CSSRTMHHC-], identified through phage display, has demonstrated antitumour activity both in vivo and in vitro. It inhibited tumor growth in 60% of mice implanted with subcutaneous melanoma cells, demonstrating its ability to function as a therapeutic agent. Furthermore, this peptide significantly inhibited tumor cell invasion in vitro and in vivo, as well as cell-to-cell adhesion (Matsuo et al., 2010).

In the present study, a novel targeted anticancer fusion protein was constructed by fusing IL24 with P20 through a rigid linker to take advantage of targeted delivery and cytotoxicity against breast cancer cells. Short amino acid sequences known as linkers are used to join two different moieties into a single polypeptide chain. We utilized a rigid type of linker in our design, which offers several advantages. Rigid linkers maintain a fixed distance between the functional domains, preventing them from interfering with each other and preserving their activities. This stability enhances the overall structural integrity and functionality of the fusion protein, leading to improved targeting and therapeutic efficacy. Current literature suggests that polyalanine linkers are known for their stability and resistance to proteolytic degradation. Studies have shown that rigid linkers are less prone to cleavage by proteases, ensuring the integrity of the fusion protein in a biological environment. However, it is important to note that in vivo conditions can vary, and additional empirical studies could provide more definitive insights into the stability of the linker in a physiological context. (Muhammad Rehman et al., 2023; Ehsasatvatan and Baghban Kohnehrouz 2023; Huang et al., 2021)

First, a Bioinformatics approach was used to assess the structural and functional qualities of potential drug candidates. In addition to understanding the macromolecular structure, molecular dynamics simulations are used to determine functional relationships (Muhammad Rehman et al., 2023). An MD simulation enables us to investigate intriguing aspects, such as physical processes, molecular interactions, and precise movements of atoms or molecules within a specific timeframe, as well as minimal protein geometries for drug-binding free energies (Aslam et al., 2023). Furthermore, this method can be used to assess and design fusion proteins by evaluating residue interactions, placement, wrapping, distance, and potential participation of harmful or desired moieties. (Karplus and Kuriyan 2005; Mahnam et al., 2014; Wadood et al., 2013). The primary aim of this study was to design a recombinant protein encoding interleukin-24 and the antitumour inhibitory peptide P20 via a rigid linker, to analyse the chimeric proteins via a bioinformatics approach and to evaluate the potential of this protein for treating breast cancer.

Materials and methods

IL24-P20 chimeric protein construction

The Universal Protein Knowledge database (UniProtKB) (http://www.Uniprot.org) provided the amino acid sequence (Saif et al., 2022) for human interleukin-24 (NM_006850.3). The 3D crystal structure of the soluble heterodimer receptor for IL-24 (interleukin-22RA and interleukin-20RB, protein Database ID: 6DF3) was retrieved from the protein databank (RCSB PDB: 6DF3) shown in Table 1. A reported P20 peptide (Matsuo et al., 2010) was fused with IL24 through a rigid linker (PAPAPAPAP) (Ghavimi et al., 2020).

Prediction of the secondary structure of the chimeric protein

The GOR-IV online server was used to examine the secondary structure of the chimeric protein. (https://npsa-prabi) (McGuffin et al., 2000). This tool predicts several functional properties, including the locations of disulfide bridges, low-complexity regions, coiled-coil domains, solvent-accessible surfaces, and segments lacking regular structures (Javaid et al., 2021).

Tertiary structure analysis of the chimeric protein

The I-TASSER (version 5.2) online program was used to construct the three-dimensional structure of the fusion proteins (https://zhanggroup.org/I-TASSER/) (Zhang 2008). I-TASSER uses iterative template fragment assembly simulations and multiple-threading alignments to model protein structures. The quality and structure validation of the modeled chimeric protein was assessed using the online SAVES v6.1 toolkit. This comprehensive toolkit offers access to various online programs for detailed quality assessment and validation. Specifically, ERRAT2 and Verify 3D were employed for quality assessment, ensuring the reliability and accuracy of the protein model. Additionally, the Rampage server was used to generate the Ramachandran plot, providing structural validation of the modeled chimeric protein (Hafiz Muhammad et al., 2024). The overall quality factor was examined using the ERRAT2 server (version 2.0) (https://saves.mbi.ucla.edu/results?job=1033540&p=errat), and the compliance of the amino acid sequence with the three-dimensional atomic model was assessed using the Verify3D server (http://services.mbi.ucla.edu/Verify 3D) (Bowie et al., 1991). The percentage of amino acids in the most favorable regions of the chimeric protein was predicted using the Rampage server (version 1.0) (http://mordred.bioc.cam.ac.ukrapper/rampage.php) (Pourhadi et al., 2019).

Refinement of the 3D structure of the chimeric protein

To refine the complete protein model, the most appropriate model created through I-TASSER was eventually submitted to the Galaxy Refine software (version 2.0) (https://galaxy.seoklab.org/cgibin/submit.cgi?type=REFINE) (Heo et al., 2013). This method minimizes the energy of the modelled protein to improve its structural geometry and resolves major steric conflicts that occur during protein modelling (Alshehri et al., 2024).

Prediction of the physical and chemical characteristics of the chimeric protein

Several physical and chemical properties of chimeric proteins were investigated using the online program ProtParam (https://web.expasy.org/program/) (Gasteiger et al., 2005) The parameters investigated included in vivo or in vitro half-life, instability index, amino acid composition, theoretical isoelectric point, GRAVY, hydropathicity, extinction coefficient, aliphatic index, and molecular weight. The solubilities of the chimeric proteins were determined with the Soluprot online program (https://protein-sol.manchester.ac.uk/) (Hebditch et al., 2017).

Evaluation of the allergenicity and Antigenicity of the Fusion protein

The AllerTOP 2.0 server was used to evaluate the allergenicity of the fusion protein (http://www.pharmfac.net/allertop/).AllerTOP 2.0 provides in-silico allergen evaluation by analysing the physical and chemical properties of proteins (Dimitrov et al., 2013). To assess the antigenicity of the proteins, the VaxiJen v2.0 server was used. (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html). This server classifies antigens based on their physio-chemical properties without requiring sequence alignment. A threshold value of 0.5 was used to differentiate antigenic proteins from non-antigenic proteins (Rehman et al., 2023).

Molecular docking and interaction studies

IL-24 triggers apoptosis upon binding to its receptor. Using the ClusPro server (https://cluspro.bu.edu/login.php), docking studies of the fusion protein and the IL22RA-IL20RB receptor were carried out to investigate the potential of these interactions. Analysis of the interactions and binding patterns between the modelled protein and its particular receptor was performed. Based on their docking scores, the top 10 docked complexes were selected, and PyMOL (https://

Table 1 Primary sequence of peptides used for constructing fusion protein

Protein	Nucleotide sequence	Amino acid sequence	No of Amino acid	References
Interleukin-24	CAAGAATT CCACTTTGG GCCCTGCC AAGTGAAG GGGGTTGT TCCCCAGA AACTGTGG GAAGCCTT CTGGGCTG TGAAAGAC ACTATGCAA GCTCAGGA TAACATCAC GAGTGCCC GGCTGCTG CAGCAGGA GGTTCTGC AGAACGTC TCGGATGCT GAGAGCTG TTACCTTGT CCACACCC TGCTGGAG TTCTACTTG AAAACTGT TTTCAAAA ACTACCAC AATAGAAC AGTTGAAG TCAGGACT CTGAAGTC ATTCTCTAC TCTGGCCA ACAACTTT GTTCTCATC GTGTCACA ACTGCAAC CCAGTCAA GAAAATGA GATGTTTTC CATCAGAG ACAGTGCA CACAGGCG GTTTCTGCT ATTCCGGA GAGCATTC AAACAGTT GGACGTAG AAGCAGCT CTGACCAA AGCCCTTG GGGAAGTG GACATTCTT CTGACCTG GATGCAGA AATTCTACA AGCTC	QEFHFGPCQVKGVVPQKLWEAFWAVKDTMQAQDNITSARLLQQEVLQNVS-DAESCYLVHTLLEFYLKTVFKNYHNRTVEVRTLKSFSTLANNFVLIVSQLQP-SQENEMFSIRDSAHRRFLLFRRAFKQLDVEAALTKALGEVDILLTWMQKFYKL	155	accession Q13007 (UniProtKB)

Table 1 (continued)

Protein	Nucleotide sequence	Amino acid sequence	No of Amino acid	References
Linker	CCGGCGCC GGCGCCG GCGCCGGC GCCGGCG CCG	PAPAPAPAPAP	11	[24]
P20	TGCAGCAG CCGCACCA TGCATCAT TGC	CSSRTMHHC	09	[18]

Fig. 1 Diagrammatic illustration of the interleukin 24-P20 recombinant protein

pymol.org/2/), was used to further evaluate the interactions of the complex with the lowest energy (DeLano 2002).

Molecular dynamic simulation

The protein-antigen complex was subjected to MD (Molecular dynamic) Simulation using VMD and NAMD (Phillips et al., 2020). The input files were generated using Ambertools 21 (Case et al., 2021), which also added the missing hydrogens using the LeaP program (Case et al., 2005). The complex was solvated using a 10 Å solvation box containing TIP3P water molecules (Jorgensen et al., 1983). To neutralize the system, Na + and Cl- counterions were also added. The system was minimized using the f14SB forcefield (Duan et al., 2003) to prevent any collisions. The solvation systems underwent three more temperature equilibrations at 200, 250, and 300 K to further stabilize them. Following that, the complex was simulated for 100 ns, saving MD trajectories every 2 ps. Lastly, the R program was used to analyze the trajectory (Grant et al., 2021).

Results

Construction of the Interleukin24-P20 chimeric protein

Amino acid sequences of P20 and interleukin-24 were retrieved from NCBI and fused with rigid linkers to create a single peptide of 175 amino acids (Fig. 1). The separation of functional peptides is assured by rigid linkers, which also prevent the formation of disulfide bonds. A restriction site was added to the C and N termini of the fusion construct to clone it within an appropriate vector for invitro expression analysis.

Prediction of the secondary structure of the fusion protein

GOR IV predicts the secondary structure, which is defined as the protein's alpha helix or beta sheets. The chimeric protein contained 50.29% alpha helices (88 residues), 9.71% beta sheets (17 residues), and 40.00% random coils (70 residues) (Fig. 2). The larger the fraction of **α**-helices is, the greater the structural integrity of the molecule (Muhammad Rehman et al., 2023). The coil-shaped stiff linker was expected to separate interleukin 24 and P20 molecules and ensure functional viability.

Estimating the tertiary structure of the fusion protein IL24-P20

The tertiary structure of the chimeric protein was generated through homology modelling via the I-TASSER server, which produced the five top models (Fig. 3). The C-score served as a guide for selecting an appropriate model based on quality and validation tests (Zheng et al., 2021). The confidence score typically ranges from 2 to 5, with higher values indicating greater confidence (Atapour et al., 2019). With a C-score of -0.11, a predicted RMSD (Root Mean Square Deviation) of 5.33, and a predictable TM-score of 0.700, the selected I-TASSER model was considered acceptable for further refinement (Table 2).

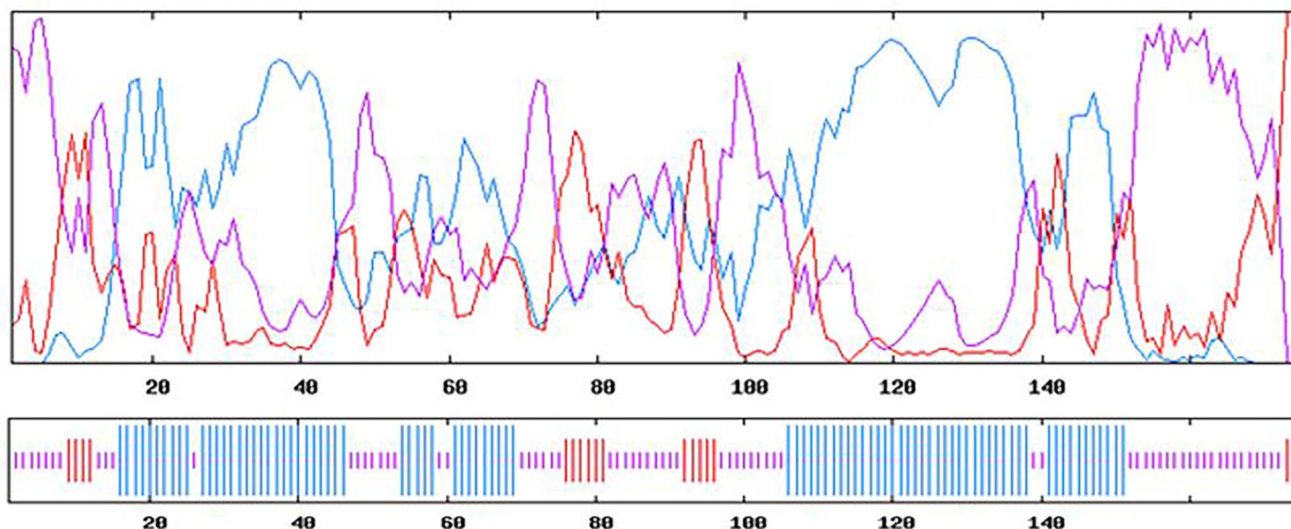
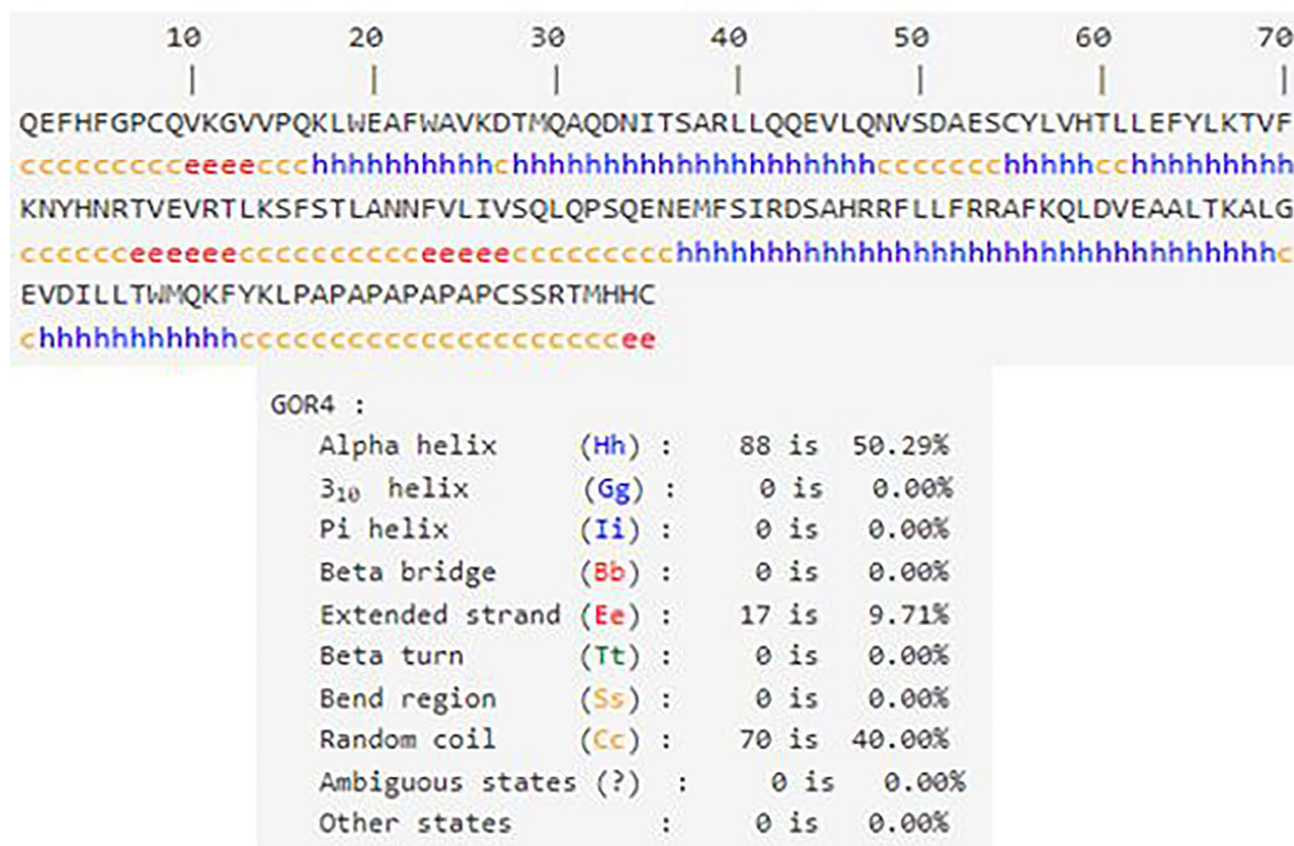


Fig. 2 The chimeric molecule's primary sequence, which is distinguished by a particular color, highlights the expected secondary structure



```
        10        20        30        40        50        60        70
         |         |         |         |         |         |         |
QEFHFGPCQVKGVVPQKLWEAFWAVKDTMQAQDNITSARLLQQEVLQNVSDAESCYLVHTLLEFYLKTVF
cccccccccceeeeccchhhhhhhhhhchhhhhhhhhhhhhhhhhhhhccccccchhhhhcchhhhhhhhhh
KNYHNRTVEVRTLKSFSTLANNFVLIVSQLQPSQENEMFSIRDSAHRRFLLFRRAFKQLDVEAALTKALG
cccccceeeeeecccccccccceeeeecccccccccchhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhc
EVDILLTWMQKFYKLPAPAPAPAPAPCSSRTMHHC
chhhhhhhhhhhccccccccccccccccccccceee

GOR4 :
   Alpha helix     (Hh) :    88 is  50.29%
   3₁₀ helix       (Gg) :     0 is   0.00%
   Pi helix        (Ii) :     0 is   0.00%
   Beta bridge     (Bb) :     0 is   0.00%
   Extended strand (Ee) :    17 is   9.71%
   Beta turn       (Tt) :     0 is   0.00%
   Bend region     (Ss) :     0 is   0.00%
   Random coil     (Cc) :    70 is  40.00%
   Ambiguous states (?)  :    0 is   0.00%
   Other states         :     0 is   0.00%
```

Fig. 3 A schematic representation of primary or secondary structure estimation via the GOR IV

Table 2 Top five final models predicted by I-TASSER

Models	Model 1	Model 2	Model 3	Model 4	Model 5
C-score	-0.11	-1.09	-1.29	-2.54	-3.57

Validation analysis, quality assurance, and selection of three-dimensional structures

The Galaxy Refine server (version 2.0) was used to refine the top 3D model (Table 3).

Table 3 Refined 3D structures predicted by glaxy refine tool

Models	GDT-HA	RMSD	MolProbity	Clash score	Poor rotamers	Rama favored
Initial	1.0000	0.000	2.691	3.8	10.3	78.6
MODEL 1	0.9586	0.421	2.023	11.7	1.9	96.5
MODEL 2	0.9586	0.417	1.729	11.3	0.6	97.1
MODEL 3	0.9557	0.426	1.672	12.4	0.6	97.7
MODEL 4	0.9543	0.425	1.614	10.7	0.6	97.7
MODEL 5	0.9586	0.423	1.944	13.4	1.3	96.5

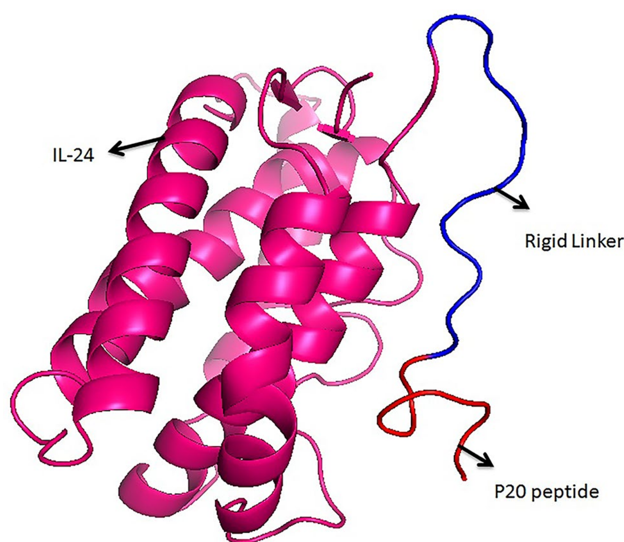
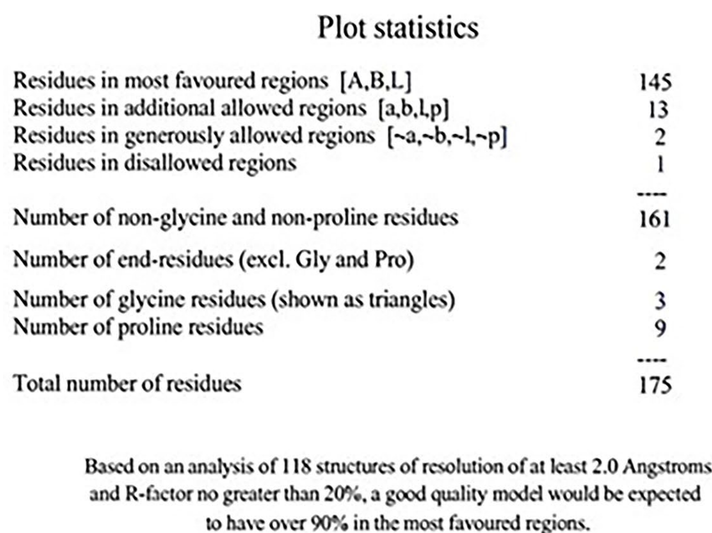


Fig. 4 The tertiary structure of the chimeric protein

The SAVES server was used to validate the best candidate model among the modified tertiary structures. The Ramachandran plot illustrates the distribution of phi and psi for each amino acid residue, demonstrating the structural correctness, accuracy, and dependability of the method. A higher proportion of amino acids in the favored zone and a lower percentage in the additionally allowed region indicate a good model (Wang et al., 2019). Based on the Ramachandran plot, the selected fusion construct contained 8.1% of its residues in additional allowed regions and 90.1% in favored regions. (Fig. 4).

The Verify 3D program was utilized for additional validation. A three-dimensional layout was generated by evaluating the compatibility of the modelled protein with its original amino acid sequence.(Vangone and Bonvin 2015). At least 80% of the amino acid residues with an average 3D-1D score greater than 0.2 were considered suitable for further investigation. According to the improved chimeric protein model, 90.86% of the residues had a 3D-1D score greater than 0.2 (Fig. 5).

The ERRAT2 algorithm was used to assess the overall quality of the fusion protein structure. The results showed that the predicted error for the chimeric protein was 93.4132% or lower, indicating an acceptable high-resolution structure (Fig. 6).



Plot statistics

Residues in most favoured regions [A,B,L]	145	90.1%
Residues in additional allowed regions [a,b,l,p]	13	8.1%
Residues in generously allowed regions [~a,~b,~l,~p]	2	1.2%
Residues in disallowed regions	1	0.6%
	----	------
Number of non-glycine and non-proline residues	161	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	3	
Number of proline residues	9	

Total number of residues	175	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

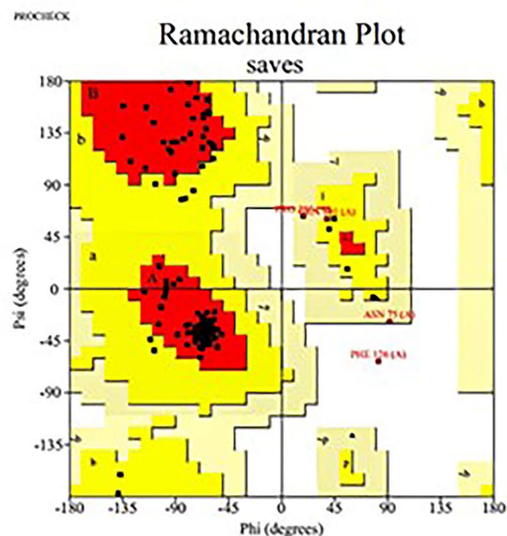


Fig. 5 A Ramachandran plot of the IL24-p20 fusion protein

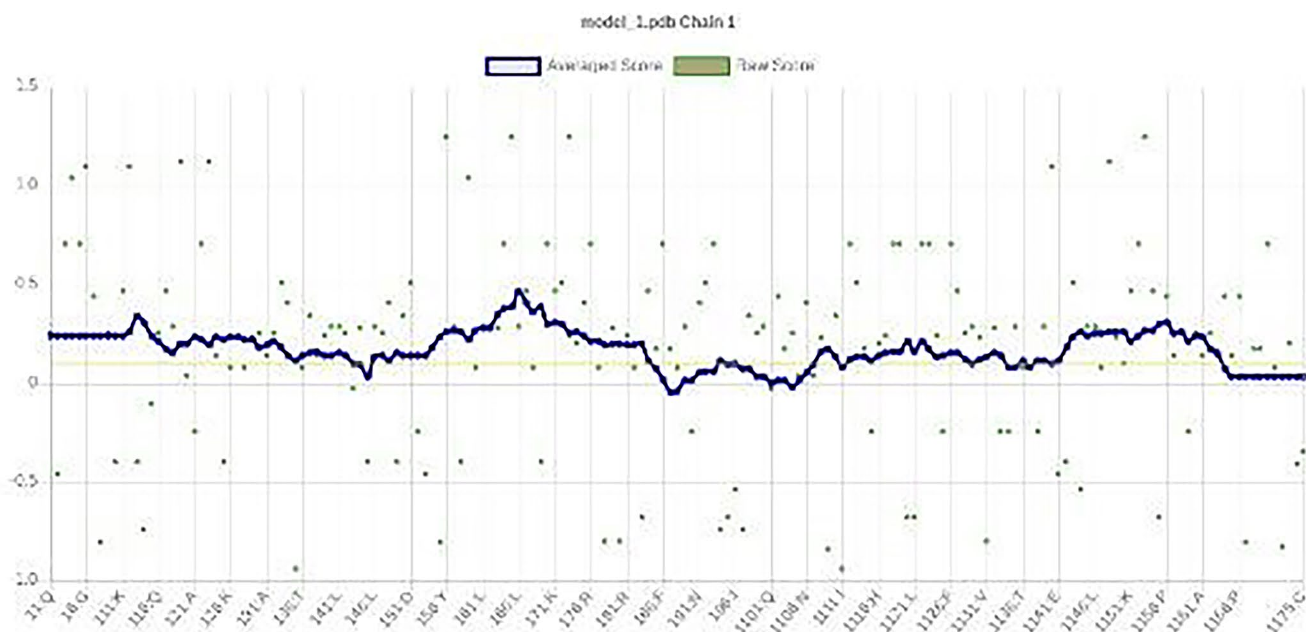


Fig. 6 Verification of the 3D structure of the chimeric protein

Table 4 Different physicochemical properties of chimeric protein

Physiochemical properties	Analytical values
Amino acids number	175
Theoretical pI	8.72
Molecular weight	20144.2
Total number of -ve charged residues	(Glu + Asp): 16
Total number of + ve charged residues	(Lys + Arg): 19
Number of atoms overall	2835
Estimated half life	10 h (Escherichia coli, in-vivo). 0.8 h (mammalian reticulocytes, in-vitro). 10 min (yeast, in-vivo).
Aliphatic index	85.83
GRAVY	-0.178

Physicochemical parameter analysis of chimeric proteins

As indicated in Table 4, the ProtParam web service was used to determine the physicochemical characteristics of the chimeric proteins. With more basic amino acid residues (Arg + Lys = 19) compared to acidic residues (Asp + Glu = 16), the analysis showed that the chimeric protein is actually basic. The 1.094 extinction coefficient of the modelled proteins indicated that these proteins may be involved in protein-protein interactions. The molecular weight was estimated to be 20 kDa, and the theoretical isoelectric point was 8.72. The estimated half-lives are 10 min in yeast (in vivo), 0.8 h in mammalian reticulocytes (in vitro), and 10 h in *Escherichia coli* (in vivo). Strategies such as PEGylation, fusion to albumin or Fc fragments, and encapsulation in nanoparticles could be considered to extend the half-life and reduce dosing frequency. With an aliphatic index of 85.83 and GRAVY values of -0.178, the fusion protein is hydrophilic and stable over a broad temperature range. Additionally, ProtParam identified Leu as the most abundant amino acid, constituting 11.4% of the protein.

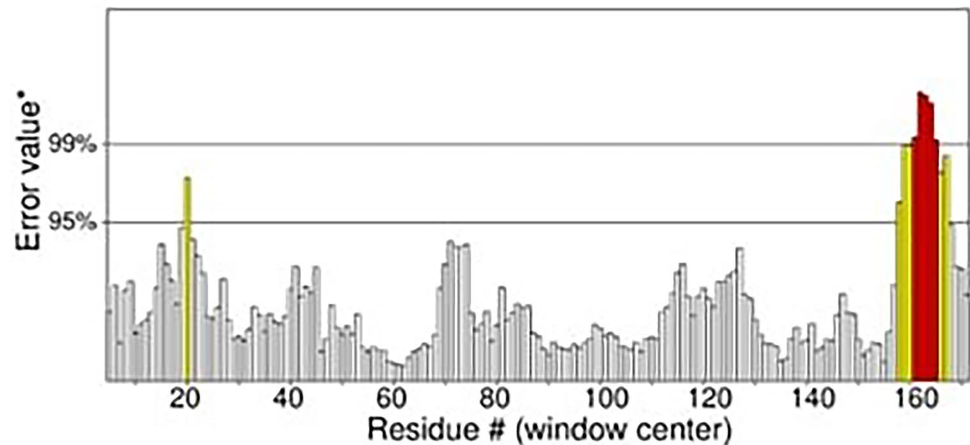
The online program Soluprot was utilized to predict the solubility of a protein in *E. coli* by analysing a single amino acid FASTA sequence. An expression is considered soluble if the solubility score is greater than 0.5, and insoluble if it is less than 0.5. The fusion protein had received a solubility index of 0.591, indicating that it is likely soluble in *Escherichia coli.*

Immunoinformatics assessment

Evaluation of antigenicity and allergenicity

The chimeric protein was estimated to be non-antigenic at a 0.5 set threshold using Vaxijen, with a bacterial model antigenicity score of 0.444. The allergenicity of the chimeric construct was assessed using AllerTOP 1.0, and the findings indicated that the protein is typically non-allergic (Atapour et al., 2019).

Fig. 7 Prediction of the quality error for each amino acid in the Interleukin 24-P20 fusion protein via the ERRAT2 program



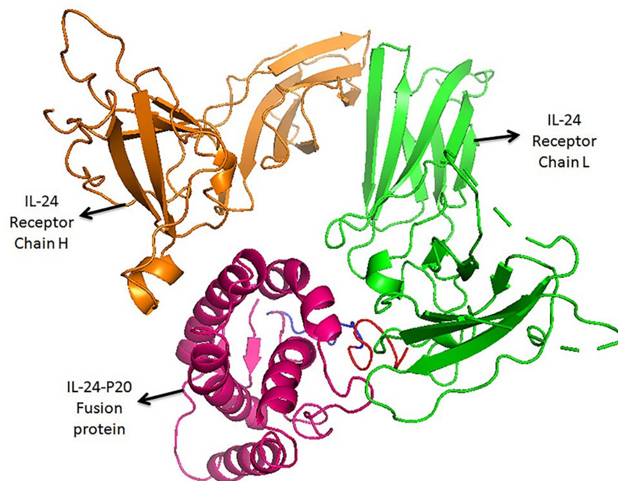


Fig. 8 ClusPro-generated docked model complex among the ligand (IL24-P20) and receptor

Fusion protein docking analysis

The ClusPro server was used to conduct docking studies to investigate the mechanics of the interaction between the chimeric protein and its receptor. The top docking models were downloaded in PDB format for further evaluation. The interaction between the IL24-P20 fusion protein and its receptor was strong, as evidenced by the extremely negative energy score (-910.8) for the molecular docking data from ClusPro 2.0 (Fig. 7).

Interactions of the IL24-p20 fusion protein with the IL22RA-IL20RB receptor

Using the PDBSum web service, the interactions between the receptor and the chimeric protein were examined. In this analysis, the IL24-p20 fusion protein is represented by chain A, and the IL24 receptor chains are denoted by chains L and H. Ten hydrogen bonds, 158 nonbonded contacts, and one salt bridge between chains A and L were found in the results. Additionally, there were 1 salt bridge, 6 hydrogen bonds, and 37 nonbonded contacts among chains H and L. Four hydrogen bonds, forty-five nonbonded contacts, and two salt bridges were found between chains A and H (refer to Fig. 8). Figure 9 shows the interactions between residues in chains A-L, H-L, and A-H.

MD simulation

Several analyses were carried out to assess the stability and dynamics of the fusion protein complex throughout the 100 ns molecular dynamics simulation. The stability of the complex over time was evaluated by calculating the root mean square deviation (RMSD) of backbone atoms. The complex was found to have reached equilibrium with an RMSD value of approximately 3 Å at 15 ns. Following equilibration, the RMSD increased somewhat toward the conclusion of the simulation but otherwise was largely constant between 2 and 3 Å (see Fig. 10). It appears from the constant RMSD trajectory that fusion protein remained stable during simulation.

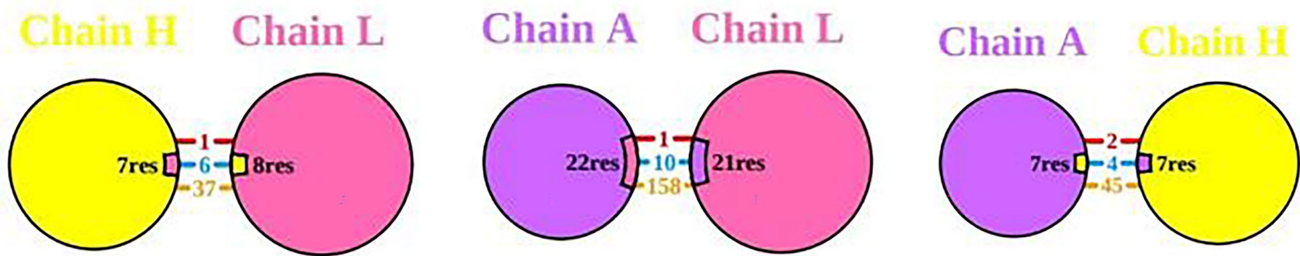


Fig. 9 Coloured lines link interacting chains. H-bonds (blue lines), salt bridges (red lines), and nonbonded contacts (orange tick marks) are all shown in the key

Fig. 10 Types of interactions between residues

By determining the chimeric protein's radius of gyration (Rg), which provides information about the protein's overall size and structural integrity, the compactness of the complex was examined. The Rg was approximately 29.2 Å.

at equilibrium, increased to a maximum of 29.4 Å at 25 ns, and then continued to increase to an average of 29 Å (see Fig. 10). With an overall Rg fluctuation of less than 0.6 Å, the complex was shown to remain compact during the simulation.

Protein residue flexibility was evaluated using the computation of root mean square fluctuations (RMSFs). Greater flexibility is indicated by higher RMSF values, while stiffness is suggested by lower values, especially in secondary structures. With a maximum RMSF value of 7 Å, RMSF analysis showed that residues 390 to 400 were more flexible.

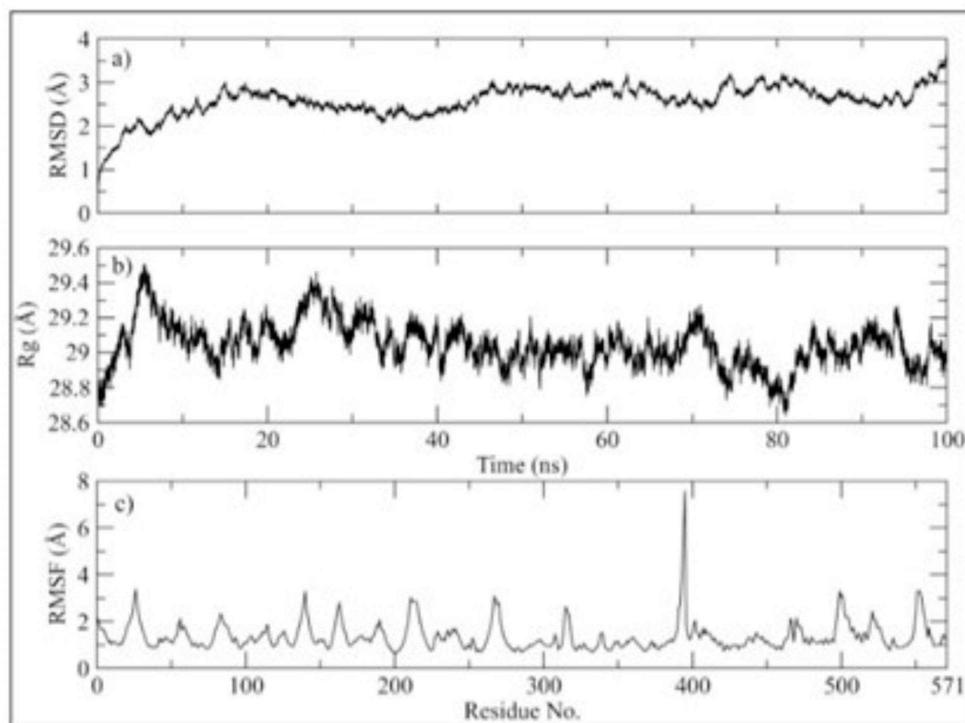


Fig. 11 (**a**) RMSD of the backbone atoms of the chimeric protein during the 100-ns simulation. (**b**) The system compactness was evaluated using a chimeric protein Rg analysis. (**c**) Analysis of the fluctuation of the residual fluctuation of chimeric protein

This is probably because this region of the protein contains many loop areas (see Fig. 10). Nevertheless, the majority of the protein residues were rather stiff, with RMSF values not exceeding 2 Å. Overall, RMSF analysis showed that throughout the simulation, the fusion protein maintained stable conformations.

All analyses showed that during the 100 ns molecular dynamics simulation, the chimeric protein complex stayed compact, stable, and structurally stiff and had only slight fluctuations noticed within some protein areas Fig. 11.

Discussion

Cancer pharmacotherapy aims to target and eliminate cancer cells while minimizing damage to healthy tissues. The monotherapeutic treatment selection regarding cancer includes surgery, hormonal therapy, radiotherapy, and chemotherapy. These monotherapies' success is largely limited due to their adverse side effects and remarkable increase in resistance. (Zahreddine and Borden 2013). Conventional therapies often face issues such as non-specific targeting, severe side effects, and the development of resistance. Additionally, they have the potential to cause secondary malignancies and are unable to eradicate cancer cells that are slow-growing and drug-resistant (Chabner and Roberts Jr 2005). Our fusion protein approach is designed to address these problems by targeting cancer cell receptors more precisely, which can reduce off-target effects and minimize systemic toxicity while improving therapeutic efficacy. Fusion proteins are designed by combining different protein domains or peptides to create novel molecules capable of specifically targeting and destroying cancer cells. This approach holds the potential to create innovative treatments for breast cancer and other malignancies, offering a pathway to enhance therapeutic efficacy while reducing harmful side effects on healthy tissues. Previous successes in interleukin-peptide fusion proteins initially analyzed through computational analysis and subsequently validated through in-vitro expression and purification, demonstrate a streamlined and effective approach to protein engineering. Notable achievements include the evaluation of interleukin 24-p28 peptide against breast cancer (Jahanian-Najafabadi et al., 2020), interleukin 24-BR2 peptide against MCF-7 cells (Pourhadi et al., 2019), interleukin 24-RGD peptide against MCF-7 cancer cells (Xiao et al., 2009), interleukin 24-p28 peptide against breast cancer (Jahanian-Najafabadi et al., 2020), TAT-interleukin 24-KDEL peptide in inducing tumor apoptosis, and interleukin 24-NRC peptide against breast cancer (Soleimani et al., 2016). A fusion protein consisting of IL-24 and P20 kills cancerous cells by acting as a toxic moiety while P20 facilitates IL-24's penetration within the membrane of cancer cells. *In silico* techniques are used to assess a novel drug candidate's structural and functional properties. In the current study, the IL-24 protein was computationally fused with the P20 peptide to design a fusion protein and assess its anti-cancer potential using an *in silico* approach. A rigid linker was inserted between

the IL24-P20 (as rigid linkers being the most efficient for constructing fusion proteins) and I-TASSER was utilized to achieve a more reliable and accurate structure. Following that, quality and validity of currently generated models were assessed using various software tools. ERRAT2 server was used to examine the overall quality factor (Awais et al., 2023). To evaluate the compatibility of amino acid sequence with a three-dimensional atomic model Verify3D server was employed. To predict amino acid percentage in the most favorable region of chimeric protein Rampage server was used. The selected model underwent structural refinement to enhance its quality and underwent reassessments for validation. The best model generated by I-TASSER was uploaded on Galaxy Refine Software. I-TASSER generated five top models & best model was chosen for quality and validation testing based on C-score. For each amino acid residue, the distribution of phi and psi is provided by the Ramachandran plot, which imply structural validity, accuracy, and reliability. Higher % of amino acids are in preferred zone, and a lower percentage are in the additionally allowed region. (Wang et al., 2019). Based on chemical validation and quality score, a trustworthy hit from predicted model was chosen. The chosen modeled fusion construct has 90.1% of its residues in favored regions & 8.1% in additionally allowed region of the Ramachandran plot. The results underscore the efficacy of structural refinement in achieving a validated protein structure before in-depth analyses. Figure 11 represents the predicted secondary structure composition of the fusion protein. Notably, a significant portion of amino acids adopt a stable and well-structured alpha helix conformation, comprising 50.29% of the secondary structure. This dominance of the alpha helix suggests the formation of rigid, coiled structures, which play crucial roles in protein functions, such as facilitating protein-protein interactions (Xie et al., 2023). Furthermore, the substantial presence of alpha helices in the protein's secondary structure indicates potential stability and resistance to conformational changes. 9.71% of the secondary structure consists of extended strands, or beta sheets, implying regions with sheet-like structures that may contribute to structural stability or binding interactions with receptors. The fusion protein also exhibited a notable percentage (40.00%) of random coils, representing flexible and disordered regions lacking a specific secondary structure. These regions may undergo conformational changes or contribute to the protein's intrinsic flexibility, adding a dynamic aspect to its structural characteristics. ClusPro 2.0 server was used for docking that explain mechanism of fusion protein and its receptor interaction. The stability of a multichain complex is influenced by amino acid interactions and the interface area. The top-ranked "zero number" model was selected for detailed molecular interaction analysis using PDBsum. The examination revealed a total of 10 hydrogen bonds, 158 unbonded contacts, and 1 salt bridge. between chain A and L there were, while 1 salt bridge, 6 hydrogen bonds 37 un bonded contacts between chain H and L, and between chain A and H there were 4 hydrogen bonds, 45 unbonded contacts, and 2 salt bridge (Fig. 8). The stability, compactness, and flexibility of a chimeric protein complex were evaluated over a 100 ns molecular dynamics simulation. The root mean square deviation (RMSD) of backbone atoms showed the complex reached equilibrium around 15 ns with an RMSD of approximately 3 Å, maintaining stability between 2 and 3 Å until the end. The radius of gyration (Rg) analysis indicated the complex remained compact, with an Rg value around 29 Å. Root mean square fluctuations (RMSF) revealed that most residues were rigid, except for residues 390–400, which were more flexible due to large loop regions.

Despite many advantages of fusion peptides, there are inherent limitations in relying solely on *in silico* studies, including the lack of empirical validation and the possibility of discrepancies between computational predictions and actual biological behavior. To address these limitations, future work will involve comprehensive in vitro studies, such as protein arrays, X-ray crystallography, and NMR spectroscopy, to validate the computational findings and assess the practical efficacy and stability of the fusion protein. These experimental approaches will provide crucial insights into the protein's functionality and support its development as a viable therapeutic option.

Conclusion

A computational approach to designing antitumour fusion proteins can considerably reduce both the time and cost, as well as eliminate the ethical issues of in vivo experiments. In this paper, we present a novel interleukin 24-P20 fusion protein construct that exhibits increased activity against breast cancer. Several parameters including, solubility, interaction studies, physical and chemical properties, and molecular dynamics simulations were used to determine the quality, validation, accuracy, and validity of the chimeric protein. The chimeric protein has shown a specific binding affinity for the IL22RA/IL20RB receptor on breast cancer cells. Due to the nonallergenic nature of the chimeric proteins, they could be used as stable antitumour drugs, making them a promising candidate for targeted cancer therapy. Our in vitro studies in the future will have the opportunity to reveal improved insights into the anticancer activity of the theoretical model presented here.

Acknowledgements This work was supported by grants from the HEC Pakistan under NRPU (**Project No. 16935)** 2021 and the University of the Punjab, Lahore, Pakistan.

Author contributions Methodology, formal analysis, validation, S.Q, N.A, H.B, investigation resources, M.I.A, F.S, I.k, data curation H.M.R, K.W, M.A, and original draft preparation S.Q, A.C, H.B. All authors have read & agreed to published version of manuscript.

Data availability No datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare no competing interests.

Conflict of interest All authors declare no conflicts of interest.
The data will be provided by Dr. Hamid Bashir, (email: hamid.camb@pu.edu.pk) upon request.

References

ALSHEHRI AA, ALMUTAIRI AM, SHAFIE A, ALSHEHRI NA, ALMUTAIRI SM, ANJUM F (2024) Identification of potential inhibitors targeting DNA adenine methyltransferase of Klebsiella pneumoniae for antimicrobial resistance management: a structure-based molecular docking study. Advancements Life Sci 10:604–608

ASLAM S, REHMAN HM, SARWAR MZ, AHMAD A, AHMED N, REHMAN AMIRZADAMI, YASMIN HM, NADEEM H, T., BASHIR H (2023) Computational modeling, high-level soluble expression and in vitro cytotoxicity assessment of recombinant Pseudomonas aeruginosa azurin: a promising anti-cancer therapeutic candidate. Pharmaceutics 15:1825

ATAPOUR A, MOKARRAM P, MOSTAFAVIPOUR Z, HOSSEINI SY, GHASEMI Y, MOHAMMADI S, NEZAFAT N (2019) Designing a fusion protein vaccine against HCV: an in silico approach. Int J Pept Res Ther 25:861–872

AWAIS H, AFZAAL ZAHIDA, MANNAN A, T., HABIB H (2023) Coding Genome Sequence and Protein Sequence Analysis of Dengue Strains: in Silico correlation. Advancements Life Sci 10:48–53

BIDWELL GL (2012) Peptides for cancer therapy: a drug-development opportunity and a drug-delivery challenge. Therapeutic Delivery 3:609–621

BOWIE JU, LÜTHY, R., EISENBERG D (1991) A method to identify protein sequences that fold into a known three-dimensional structure. Science 253:164–170

CASE DA, III CHEATHAM, DARDEN TE, GOHLKE T, MERZ HLUOR, JR KM, ONUFRIEV A, SIMMERLING C, WANG B, WOODS RJ (2005) The Amber biomolecular simulation programs. J Comput Chem 26:1668–1688

CASE DA, AKTULGA HM, BELFON K, BEN-SHALOM I, CERUTTI BROZELLSR (2021) Amber 2021. T. E., CRUZEIRO, V. W. D., DARDEN, T. A. & DUKE, R. E., University of California, San Francisco, III

CHABNER BA, ROBERTS TG JR (2005) Chemotherapy and the war on cancer. Nat Rev Cancer 5:65–72

CHAKRABORTY S, RAHMAN T (2012) The difficulties in cancer treatment. Ecancermedicalscience, 6

CRAIK DJ, FAIRLIE DP, LIRAS S, PRICE D (2013) The future of peptide-based drugs. Chem Biol Drug Des 81:136–147

DELANO WL (2002) Pymol: an open-source molecular graphics tool. CCP4 Newsl Protein Crystallogr 40:82–92

DENT P, YACOUB A, PARK HAMEDHA, SARKAR MADASHRB-HUTIASK, GUPTA DWANGX-Y, P., EMDAD L (2010) The development of MDA-7/IL-24 as a cancer therapeutic. Pharmacol Ther 128:375–384

DESANTIS CE, GAUDET MAJ, GODING MMNEWMAN-LAMILLERKD, SAUER A, JEMAL, A., SIEGEL RL (2019) Breast cancer statistics, 2019. *CA: a cancer journal for clinicians*, 69, 438–451

DIMITROV I, FLOWER, D. R., DOYTCHINOVA I (2013) AllerTOP-a server for in silico prediction of allergens. BMC bioinformatics, Springer, 1–9

DONG X, MUMPER RJ (2010) Nanomedicinal strategies to treat multidrug-resistant tumors: current progress. Nanomedicine 5:597–615

DUAN Y, WU C, CHOWDHURY S, LEE MC, XIONG G, ZHANG W, YANG R, LUO CIEPLAKP, R., LEE T (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 24:1999–2012

EHSASATVATAN M, BAGHBAN KOHNEHROUZ B (2023) Designing and computational analyzing of chimeric long-lasting GLP-1 receptor agonists for type 2 diabetes. Sci Rep 13:17778

EMDAD L, SARKAR D, LEBEDEVA IV, SU ZZ, GUPTA P, MAHASRESHTI PJ, DENT P, CURIEL DT, FISHER PB (2006) Ionizing radiation enhances adenoviral vector expressing mda-7/IL-24-mediated apoptosis in human ovarian cancer. J Cell Physiol 208:298–306

GAO P, CHEN SUNX, WANG X, SUBJECK YFOSTERBA, FISHER J, P. B., WANG X-Y (2008) Secretable chaperone Grp170 enhances therapeutic activity of a novel tumor suppressor, mda-7/IL-24. Cancer Res 68:3890–3898

GASTEIGER E, HOOGLAND C, GATTIKER A, DUVAUD SE, WILKINS MR, APPEL, R. D., BAIROCH A (2005) Protein identification and analysis tools on the ExPASy server. Springer

GHAVIMI R, MOHAMMADI E, AKBARI V, SHAFIEE F, JAHANIAN-NAJAFABADI A (2020) In silico design of two novel fusion proteins, p28-IL-24 and p28-M4, targeted to breast cancer cells. Res Pharm Sci 15:200–208

GRANT BJ, SKJÆRVEN L, YAO XQ (2021) The Bio3D packages for structural bioinformatics. Protein Sci 30:20–30

HAFIZ MUHAMMAD R, WARDAH S, MUHAMMAD NAVEED K, NUMAN Y, FAREEHA B, QURBAN HAMIDB, A., SHIMING H (2024) A Comprehensive in Silico Study of the NDB-IL-24 Fusion protein for Tumor Targeting: a promising anti-cancer therapeutic candidate. J Biol Regul Homeost Agents 38:3449–3461

HEBDITCH M, CARBALLO-AMADOR MA, CHARONIS S, CURTIS R, WARWICKER J (2017) Protein–Sol: a web tool for predicting protein solubility from sequence. Bioinformatics 33:3098–3100

HEO L, PARK H, SEOK C (2013) GalaxyRefine: protein structure refinement driven by side-chain repacking. Nucleic Acids Res 41:W384–W388

HUANG Z, ZHANG C, XING X-H (2021) Design and construction of chimeric linker library with controllable flexibilities for precision protein engineering. Methods in enzymology. Elsevier

JAHANIAN-NAJAFABADI A, GHAVIMI R, AKBARI V (2020) 10P in vitro and in vivo cytolethal and antitumor effects of a novel fusion protein targeting IL-24 toward breast cancer cells. Ann Oncol 31:S2

JAVAID MS, KAUL H, FAZAL N, YAQUB F, NASEER N, HANIF M, LATIEF N (2021) In silico analysis to reveal underlying trans differentiation mechanism of mesenchymal stem cells into Osteocytes. Advancements Life Sci 8:412–418

JIANG H, SU LINJJ, GOLDSTEIN Z, N., FISHER P (1995) Subtraction hybridization identifies a novel melanoma differentiation associated gene, mda-7, modulated during human melanoma differentiation, growth and progression. Oncogene 11:2477–2486
JORGENSEN WL, CHANDRASEKHAR, J., MADURA, J. D., IMPEY, R. W., KLEIN ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79:926–935
KARPLUS M, KURIYAN J (2005) Molecular dynamics and protein function. Proc Natl Acad Sci 102:6679–6685
MAHNAM K, SAFFAR B, MOBINI-DEHKORDI M, FASSIHI A, MOHAMMADI A (2014) Design of a novel metal binding peptide by molecular dynamics simulation to sequester Cu and Zn ions. Res Pharm Sci 9:69–82
MATSUO AL, JULIANO TANAKAAS, M. A., RODRIGUES, E. G., TRAVASSOS LR (2010) A novel melanoma-targeting peptide screened by phage display exhibits antitumor activity. J Mol Med 88:1255–1264
MCGUFFIN LJ, BRYSON K, JONES DT (2000) The PSIPRED protein structure prediction server. Bioinformatics 16:404–405
MENEZES ME, SHEN X-N, EMDAD DASSK, YUAN LGUOC, LI F, ZACKSENHAUS Y-JARCHERMC, E., WINDLE JJ (2015) MDA-7/IL-24 functions as a tumor suppressor gene in vivo in transgenic mouse models of breast cancer. Oncotarget 6:36928
MUHAMMAD REHMAN H, REHMAN HM, NAVEED M, KHAN MT, ASLAM SHABBIRMA, S., BASHIR H (2023) In silico investigation of a chimeric IL24-LK6 fusion protein as a potent candidate against breast cancer. Bioinform Biol Insights 17:11779322231182560
NISHIKAWA T, MUNSHI RAMESHR, CHADA A, S., MEYN RE (2004) Adenovirus-mediated mda-7 (IL24) gene therapy suppresses angiogenesis and sensitizes NSCLC xenograft tumors to radiation. Mol Ther 9:818–828
PHILLIPS JC, HARDY DJ, RIBEIRO MAIAJDSTONEJE, BUCH JVBERNARDIRC, FIORIN R, HÉNIN G, J., JIANG W (2020) Scalable molecular dynamics on CPU and GPU architectures with NAMD. J Chem Phys, 153
POURHADI M, JAHANIAN-NAJAFABADI JAMALZADEF, A., SHAFIEE F (2019) Expression, purification, and cytotoxic evaluation of IL24-BR2 fusion protein. Res Pharm Sci 14:320–328
REHMAN HM, REHMAN HM, AHMED N, IMRAN M (2023) In silico design and evaluation of Novel Cell Targeting Melittin-Interleukin-24 Fusion protein: a potential drug candidate against breast Cancer. Sains Malaysiana 52:3223–3237
SAIF R, ASHFAQ K, ALI G, IFTEKHAR A, ZIA S, YOUSAF MZ (2022) Computational prediction of Cassia angustifolia compounds as a potential drug agents against main protease of SARS-nCov2. Advancements Life Sci 9:36–40
SAUANE M, SU Z-Z, GUPTA P, LEBEDEVA IV, DENT P, SARKAR D, FISHER PB (2008) Autocrine regulation of mda-7/IL-24 mediates cancer-specific apoptosis. Proc Natl Acad Sci 105:9763–9768
SOLEIMANI M, MIRMOHAMMAD-SADEGHI MAHNAMK, SADEGHI-ALIABADI H, H., JAHANIAN-NAJAFABADI A (2016) Theoretical design of a new chimeric protein for the treatment of breast cancer. Res Pharm Sci 11:187–199
SUNG H, SIEGEL FERLAYJ, LAVERSANNE RL, JEMAL MSOERJOMATARAMI, A., BRAY F (2021) Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. Cancer J Clin 71:209–249
THUN MJ, CENTER DELANCEYJO, JEMAL MM, A., WARD EM (2010) The global burden of cancer: priorities for prevention. Carcinogenesis 31:100–110
THUNDIMADATHIL J (2012) Cancer treatment using peptides: current therapies and future prospects. J Amino Acids. ; 2012: 967347. Epub 2013/01/15. https://doi.org/10.1155/2012/967347 PMID: 23316341
VANGONE A, BONVIN AM (2015) Contacts-based prediction of binding affinity in protein–protein complexes. Elife 4:e07454
WADOOD A, AHMED N, AHMAD SHAHL, HASSAN A, H., SHAMS S (2013) In-silico drug design: an approach which revolutionarised the drug discovery process. OA Drug Des Deliv 1:3
WANG S-Q, SHI M, XU FANGL, WANG S-M, C., YU Z (2019) Design of dual inhibitors of human TNF-α and IL-6 with potentials for the treatment of rheumatoid arthritis. Trop J Pharm Res 18:2305–2312
XIAO B, LI W, YANG J, GUO G, MAO X-H, ZOU Q-M (2009) RGD-IL-24, a novel tumor-targeted fusion cytokine: expression, purification and functional evaluation. Mol Biotechnol 41:138–144
XIE X, VALIENTE, P. A., KIM PM (2023) HelixGAN a deep-learning methodology for conditional de novo design of α-helix structures. Bioinformatics 39:btad036
ZAHREDDINE H, BORDEN KL (2013) Mechanisms and insights into drug resistance in cancer. Front Pharmacol 4:28
ZHANG Y (2008) I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 9:1–8
ZHANG S, JIN Z, BAO L, SHU P (2024) The global burden of breast cancer in women from 1990 to 2030: assessment and projection based on the global burden of disease study 2019. Front Oncol, 14
ZHENG W, ZHANG C, LI Y, PEARCE R, BELL EW, ZHANG Y (2021) Folding non-homologous proteins by coupling deep-learning contact maps with I-TASSER assembly simulations. Cell Rep Methods, 1

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.