Research in Pharmaceutical Sciences, February 2017; 12(1): 60-66
Received: May 2016
Accepted: October 2016

School of Pharmacy & Pharmaceutical Sciences
IsfahanUniversity of Medical Sciences

*Original Article*

# *In silico* designing of a new cysteine analogue of hirudin variant 3 for site specific PEGylation

**Seyed Mehdi Sajjadi[1], Hamzeh Rahimi[2], Saeed Mohammadi[3], Mohammad Faranoush[4], Hasan Mirzahoseini[5,*], and GholamrezaToogeh[6,*]**

[1] *Cellular and Molecular Research Center, Birjand University of Medical Sciences, Birjand, I.R. Iran.*
[2]*Molecular Medicine Group, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, I.R. Iran.*
[3]*Hematology, Oncology and Stem Cell Transplantation Research Center, Tehran University of Medical Sciences, Tehran, I.R. Iran.*
[4]*Rasoul Akram Medical Center, Iran University of Medical Sciences, Tehran, I.R. Iran.*
[5]*Biotechnology Group, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, I.R. Iran.*
[6]*Thrombosis and Homeostasis Research Center, Imam Khomeni Hospital Complex, Tehran University of Medical Sciences, Tehran, I.R. Iran.*

## *Abstract*

Hirudin is an anticoagulant agent of the salivary glands of the medicinal leech. Recombinant hirudin (r-Hir) displays certain drawbacks including bleeding and immunogenicity. To solve these problems, cysteine-specific PEGylation has been proposed as a successful technique. However, proper selection of the appropriate cysteine residue for substitution is a critical step. This study has, for the first time, used a computational approach aimed at identifying a single potential PEGylation site for replacement by cysteine residue in the hirudin variant 3 (HV3). Homology modeling (HM) was performed using MODELLER. All non-cysteine residues of the HV3 were replaced with the cysteine. The best model was selected based on the results of discrete optimized protein energy score, PROCHECK software, and Verify3D. The receptor binding was investigated using protein-protein docking by ClusPro web tool which was then visualized using LigPlot+ software and PyMOL. Finally, multiple sequence alignment (MSA) using ClustalW software and disulfide bond prediction were performed. According to the results of HM and docking, Q33C, which was located on the surface of the protein, was the best site for PEGylation. Furthermore, MSA showed that Q33 was not a conserved residue and LigPlot+ software showed that it is not involved in the hirudin-thrombin binding pocket. Moreover, prediction softwares established that it is not involved in disulfide bond formation. In this study, for the first time, the utility of the *in silico* approach for creating a cysteine analogue of HV3 was introduced. Our study demonstrated that the substitution of Q33 by cysteine probably has no effect on the biological activity of the HV3. However, experimental analyses are required to confirm the results.

***Keywords:*** *In silico*; Hirudin variant 3; PEGylation

## INTRODUCTION

Hirudin is an anticoagulant agent of the salivary glands of the medicinal leech; Hirudomedicinalis. This protein is a single polypeptide that contains 65–66 amino acids and has a molecular weight of approximately 7KDa. It also has three intra-molecular disulfide bands between the residues 6-14, 16-28, and 22-39 (1,2). Because of the pharmaceutical usefulness of hirudin and its limited availability in a natural form, the production of recombinant hirudin (r-Hir) is of great interest (3,4). Nevertheless, recombinant hirudin displays a short plasma half-life of approximately 60-100 min; therefore, several injections are needed, which makes it more expensive for patients (5). Furthermore, it has certain drawbacks, including bleeding and immunogenicity (6).

*Corresponding authors:
H. Mirzahoseini, Tel: 0098 9126933992, Fax: 0098 2188601599
Email: mirzahoseini@pasteur.ac.ir
Gh. Toogeh, Tel: 0098 9121346329, Fax: 0098 2188060717
Email: gh_toogeh@yahoo.com

**Access this article online**

**Website:** http://rps.mui.ac.ir

**DOI:** 10.4103/1735-5362.199048

To resolve these issues, PEGylation has been proposed as a successful technique to improve the half-life and immunogenicity (7,8).

While PEGylation can increase the size, drug solubility, and bioavailability of r-Hir, it also decreases the dose frequency and immunogenicity (9).

Among the different PEGylation methods, lysine PEGylation is most commonly used; however, due to the high prevalence of this amino acid in the molecule and the low chemical selectivity of the amine group, positional isomers occur (5,7).

A potential result of non-specific PEGylation is a decrease in biologic activity (10). Thus, site specific PEGylation is an effective method to obtain a mono-PEGylated form. Because the number of free cysteine residues on the protein's surface is lower than that of other conjugable groups (i.e. lysine residues), free cysteine PEGylation is considered to be an efficient tool for the enhancement of the pharmacokinetic properties of therapeutic proteins (11).

The three-dimensional structure of protein must not be altered by the insertion of cysteine residue (12); however, increasing the potential for disulfide bond formation and protein dimerization are drawbacks of the addition of free cysteines via genetic engineering (13). Therefore, the proper selection of the appropriate sites for substitution with cysteine residue is a critical step in the process (10,12). The *in silico* design of new variants with the proper site for PEGylation is currently considered a useful method, and it is therefore frequently used in studies of folding, kinetics, and different interactions of proteins.

The results of computational studies have revealed insights into the mechanism of protein dynamics. These studies have also helped identify a variety of structural and dynamical characteristics of protein interactions (14-16).

For the first time, a computational approach was used in this study that aimed to identify a single potential PEGylation site for replacement by cysteine residue in the hirudin variant 3 (HV3; hirPA).

## MATERIALS AND METHODS

### *Homology modeling and design mutant library*

To the best of the authors' knowledge, a three-dimensional structure for HV3 has not been identified; therefore, HM was performed using MODELLER version 9.11 (http://www.salilab.org) (17). Sequence-structure alignment against the protein data bank (PDB) database was carried out using the HHpred toolkit (18) to determine the most appropriate template. Ten thousand models were generated using MODELLER version 9.11 (17), and a crystal structure of hirudin variant 1 (HV1) (PDB code: 1HRT, chain I) from the protein data bank (http://www.rcsb.org) was used as a template. To identify the best location for cysteine-specific PEGylation, all non-cysteine residues of the HV3 (60 amino acids) were replaced with cysteine. Among the 10,000 generated molecules of the cysteine analogs of HV3, the five best models of each mutant were selected based on their discrete optimized protein energy (DOPE) score (19). The quality of the models was then checked by a torsion angles analysis of the protein backbone with PROCHECK software (20) and Verify3D (21).

### *Receptor binding*

To investigate the receptor (thrombin) binding of new mutants, protein-protein docking was applied by using the ClusPro web tool (22), which functions based on the surfaces' complementarities. Subsequently, the resulting models were sorted by clustering based on their docking score, as described by Comeau, *et al*. (23). A two-dimensional interacting residues diagram was visualized using LigPlot+ software (24) and the 3D structure was drawn by means of PyMOL (25).

### *Multiple sequence alignment*

The FASTA format of hirudin variants from the universal protein resource (http://www.uniprot.org) (26) was used in multiple sequence alignment (MSA) using ClustalW software with the blocks substitution matrix (BLOSUM62) scoring matrix and default parameter settings. For the alignment visualization and the secondary structure

elements' predictions, JalView and Jpred softwares were applied, respectively (27,28).

***Disulfide bond prediction***

To predict the bonding state of cysteines in hirudin, online software that used the different algorithms, including Dipro (29),CYSPRED (30), Disulfind (31) and Metadetector (32), were employed.

## RESULTS

For this study, the computational methods were employed to determine the best site for cysteine specific PEGylation in HV3. HM was applied to predict the 3D structure of HV3 and its binding to the receptor.

The best template for HM was obtained by sequence-structure alignment using the HHpred toolkit. The HV1 (PDB code: 1HRT), which had an 85% identity and 1.518 similarity, was selected as the best template for HM (Table1). Among 10,000 generated models, the best five models were selected based on their DOPE score (Table 2). The best model (Fig. 1A) was then subjected to further analyses, including stereochemical conformation checking (Ramachandran plot) and Verify 3D, to identify the association between the 3D structure and the sequence. ProsaII was then applied for structural quality validation. It was found that the covalent bonds and the related angles were acceptable (Fig. 1B and 1C).

PyMOL software was employed to visualize the selected model. As a result, a folded globular anti-parallel beta sheet was shown at the N-terminal domain, and an unstructured tail with a 3-10 helix was found at the C-terminal region (Fig. 1A).

**Table 1.**HHpred toolkit analysis

| PDB code | E-value | Score | Aligned columns | Identities | Similarity |
|---|---|---|---|---|---|
| IHRT | 7.1e-51 | 255.21 | 65 | 85% | 1.518 |

**Table 2.**The best models of HV3 based on the DOPE score and the related analyses

| No | Model | DOPE | Verify 3D | Prosa score | Most allowed (%) | Allowed (%) | Generously allowed (%) | Disallowed (%) |
|---|---|---|---|---|---|---|---|---|
| 1 | hirPA.B99991375 | -4032.8562 | 0.302 | 0.641 | 86.3 | 13.7 | 0.0 | 0.0 |
| 2 | hirPA.B99991258 | -4007.28369 | 0.354 | 0.662 | 82.4 | 15.7 | 0.0 | 0.0 |
| 3 | hirPA.B99997408 | -3986.17334 | 0.373 | 0.777 | 76.5 | 21.5 | 2.0 | 0.0 |
| 4 | hirPA.B99996774 | -3985.3147 | 0.293 | 0.789 | 82.4 | 17.6 | 0.0 | 0.0 |
| 5 | hirPA.B99993446 | -3983.53003 | 0.355 | 0.752 | 78.4 | 19.6 | 2.0 | 0.0 |

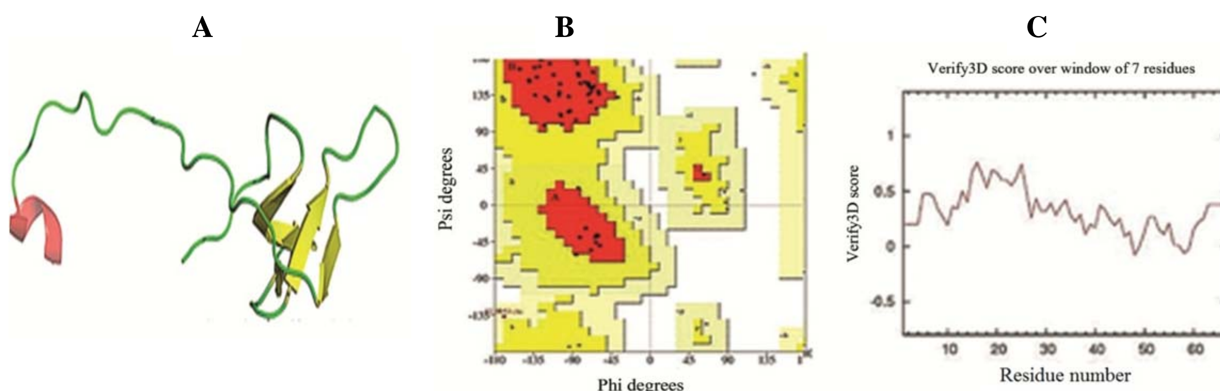


**Fig. 1.**(A) 3D structure of the best model of HV3. The structure contains a 3-10 helix in the C-terminal, an anti-parallel beta sheet, and an unstructured region in the C-terminal. No alpha helix is observed. (B) Stereochemical analysis of the model HV3-normal is drawn from PROCHECK. The plot confirms the stereochemical quality with 90.2% of residues fall in the most favored region and 7.8% in allowed region and 2% in generously allowed region. No residue was located in the disallowed region. (C) Verify 3D analysis of HV3. As the scores were higher than zero, the model correctness was confirmed.

To design the cysteine variants, single amino acid replacement was performed using HM. Then, the best models were introduced to Cluspro to investigate the binding affinity of HV3 to the receptor. Finally, the best mutant for cysteine PEGylation was selected based on structural quality and binding energy in docking. The HM and docking showed that the Q33C was the best site for PEGylation. Because the Q33 is located on the surface of the protein, it would be an appropriate site for the PEGylation (Fig. 2A). The surface accessibility for Q33 was defined using the visualization with PyMOL. After further analyses using Ramachandran plot and Verify3D, the accuracy of the model was confirmed; the quality and binding pattern of this variant were similar to the normal pattern (Fig. 2B and 2C).

The 2D visualization of hirudin-thrombin using LigPlot+ software showed that the binding pocket had not changed. In other words, Q33C was not involved in ligand-receptor binding.

For further confirmation, MSA was carried out using ClustalW software. Different variants of hirudin were subjected to MSA. As can be found from the Fig. 3, the Q33 residue (code: 09944) was not conserved; therefore, it was not likely to be essential to the protein function and could be changed into another amino acid.

The disulfide bond formation prediction was performed using different online software. The results showed that the hirudin did not use the Q33C for disulfide bond formation. Hence, the native disulfide bonds of hirudin did not change after the cysteine insertion.

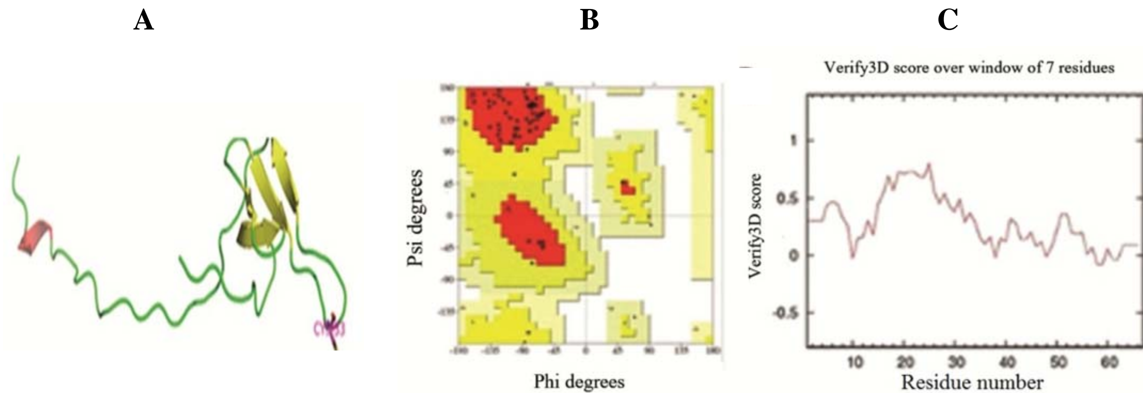


**Fig. 2.** (A) 3D structure of the HV3-Q33C. Cys33 is located on the surface of the protein so, it would be accessible for the PEGylation. (B) Stereochemical property analysis of the model HV3-Q33C is drawn from PROCHECK. The plot confirms the stereochemical quality of the model, with 92.2% of residues fall in the most favored region and 7.8% in allowed region. No residue is located in the disallowed region. (C) Verify 3D analysis of HV3-Q33C. As the scores were higher than zero, the model correctness was confirmed.

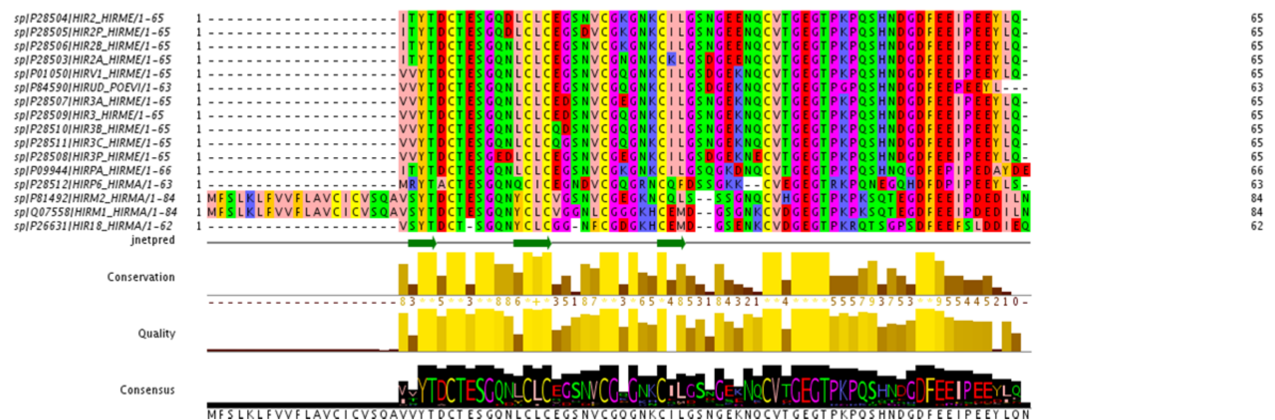


**Fig. 3.** Multiple sequence alignment using ClustalW software. The Q33 residue is shown not to be a conserve residue (code: 09944).

## DISCUSSION

*In silico* designing has been used for different proteins (33-35). In this regard, we applied this approach for site specific PEGylation, a well-known method used to prolong the circulating half-lives of protein therapeutics and to improve both their stability and efficiency (36-38). The primary challenge is determining the best conformational location for PEGylation. Cysteine-specific PEGylation includes introducing a 'free' cysteine amino acid into a target protein that has not participated in a disulfide bond formation. The PEGylation reaction is carried out with a cysteine-specific PEG reagent. A major technical problem with this method is the identification of a proper site in the protein where the substituted or inserted cysteine residue can be easily PEGylated and thus have no effect on the biological activity of the protein (39). To perform enzymatic PEGylation by microbial transglutaminase in G-CSF, two different mutations, P132Q and Q134N, were selected using bioinformatic approach. It was shown that an *in silico* study is a useful approach for identifying specific sites for PEGylation (15). Cysteine specific PEGylation is performed for different proteins, such as interferon α-2 (39) and GM-CSF (40). Hirudin, a recently FDA-approved drug, has been PEGylated with different PEGylation reagents, including PEG-NHS ester disulfide (41), SC-mPEG (42,43), and methoxyl-polyethylene glycol (mPEG) (44). For the first time, the utility of the *in silico* approach for creating a cysteine analogue of hirudin variant 3 has been introduced in this study. The results of this study showed that Q33 would be the best residue for substitution with cysteine. As can be observed from the Ramachandrn plot, the mutant model's quality is quite similar to the normal variant. Furthermore, because Q33 is located on the surface of hirudin (Fig. 2A), it would be easily accessible for the cysteine specific PEGylation reagent. In addition, based on the MSA analysis, Q33 is not a conserved residue of the hirudin molecule; therefore, it probably has no significant effect on the biological activity of the protein. Further confirmation of the appropriateness of selecting the Q33 residue was provided by the results of the online software used including DIpro (29), CYSPRED (30), Disulfind (31), and Metadetector (32), which showed the binding state of cysteines in the proteins. These software programs employed different algorithms. In this regard, Q33 was not involved in disulfide bond formation. Moreover, there has been another study in which the Asp33 in the hirudin variant 1(HV1) was changed to lysine, and PEGylation was successfully completed (8). The LigPlot+ also revealed that the Q33 was not involved in the hirudin-thrombin binding pocket. This is in accordance with previously reported findings (45).

## CONCLUSION

This was the first study in which the *in silico* approach was applied for creating a cysteine analogue of HV3. Our study demonstrated that the substitution of Q33 by cysteine probably has no effect on the biological activity of the HV3. However, experimental analyses are required to confirm the results. In conclusion, an *in silico* study can be a fast and useful approach for predicting a proper site for the cysteine-specific PEGylation of recombinant hirudin.

## REFERENCES

1. Lazar JB, Winant RC, Johnson PH. Hirudin: amino-terminal residues play a major role in the interaction with thrombin. J Biol Chem. 1991;266(2):685-688.
2. Markwardt F. Hirudin: the promising antithrombotic. Cardiovasc Drug Rev. 1992;10(2):211-232.
3. Liu CC, Schultz PG. Recombinant expression of selectively sulfated proteins in Escherichia coli. Nat Biotechnol. 2006;24(11):1436-1440.
4. Greinacher A, Lubenow N. Recombinant hirudinin clinical practice focus on lepirudin. Circulation. 2001;103(10):1479-1484.
5. Hou B, LiS, Li X, Xiu Z. Design; preparation and in vitro bioactivity of mono-PEGylated recombinant hirudin. Chin J Chem Eng. 2007;15(6):775-780.
6. Warkentin TE. Bivalent direct thrombin inhibitors: hirudin and bivalirudin. Best Pract Res ClinHaematol. 2004;17(1):105-125.
7. Wang X, Hu J, Pan D, Teng H, Xiu Z. PEGylation kinetics of recombinant hirudin and its application for the production of PEGylated HV2 species. BiochemEng J. 2014;85:38-48.

8. Avgerinos GC, Turner BG, Gorelick KJ, Papendieck A, Weydemann U, Gellissen G. Production and clinical development of a Hansenulapolymorpha-derived PEGylated hirudin. Semin Thromb Hemost. 2001;27(4):357-372.
9. Pasut G, Veronese FM. PEGylation for improving the effectiveness of therapeutic biomolecules. Drugs Today (Barc). 2009;45(9):687-695.
10. Fee CJ, Van Alstine JM. PEG-proteins: reaction engineering and separation issues. Chem Eng Sci. 2006;61(3):924-939.
11. Colonna C, Conti B, Perugini P, Pavanetto F, Modena T, Dorati R, *et al.* Site-directed PEGylation as successful approach to improve the enzyme replacement in the case of prolidase. Int J Pharm. 2008;358(1-2):230-237.
12. Gaberc-Porekar V, Zore I, Podobnik B, Menart V. Obstacles and pitfalls in the PEGylation of therapeutic proteins. Curr Opin Drug Discov Devel. 2008;11(2):242-250.
13. Roberts MJ, Bentley MD, Harris JM. Chemistry for peptide and protein PEGylation. Adv Drug Deliv Rev. 2002;54(4):459-476.
14. Cohan RA, Madadkar-sobhani A, Khanahmad H, Roohvand F, Aghasadeghi MR, Hedayati MH, *et al.* Design, modeling, expression and chemoselective PEGylation of a new nanosize cysteine analog of erythropoietin. Int J Nanomedicine. 2011;6:1217-1227.
15. Maullu C, Raimondo D, Caboi F, Giorgetti A, Sergi M, Valentini M, *et al.* Site-directed enzymatic PEGylation of the human granulocyte colony-stimulating factor. FEBS J. 2009;276(22):6741-6750.
16. Jadhav AN, Dash RC, Hirwani RR. Relative stability of thrombin-hirudin complex is illustrated using molecular dynamics. J Comput Meth Mol Des. 2014;4(4):54-62.
17. Šali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol. 1993;234(3):779-815.
18. Söding J, Biegert A, Lupas AN. The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res.2005;33(Web Server issue):W244-248.
19. Eramian D, Shen My, Devos D, Melo F, Sali A, Marti-Renom MA. A composite score for predicting errors in protein structure models. Protein Sci. 2006;15(7):1653-1666.
20. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst. 1993;26(2):283-291.
21. Eisenberg D, Lüthy R, Bowie JU. VERIFY3D: assessment of protein models with three-dimensional profiles. Methods Enzymol. 1997;277:396-404.
22. Comeau SR, Gatchell DW, Vajda S, Camacho CJ. ClusPro: a fully automated algorithm for protein-protein docking. Nucleic Acids Res. 2004;32(Web Server issue):W96-99.
23. Comeau SR, Gatchell DW, Vajda S, Camacho CJ. ClusPro: an automated docking and discrimination method for the prediction of protein complexes. Bioinformatics. 2004;20(1):45-50.
24. Laskowski RA, Swindells MB. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. J ChemInf Model. 2011;51(10):2778-2786.
25. DeLano W. The PyMOL molecular graphics system. 2002, DeLano Scientific LLC. San Carlos,CA,USA. http://pymol.sourceforge.net.
26. UniProt Consortium. The Universal Protein resource (UniProt). Nucleic Acids Res. 2007;35(Database issue):D193-197.
27. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics. 2009;25(9):1189-1191.
28. Cole C, Barber JD, Barton GJ. The Jpred 3 secondary structure prediction server. Nucleic Acids Res. 2008;36(Web Server issue):W197-201.
29. Cheng J, Saigo H, Baldi P. Large-Scale Prediction of disulphide bridges using kernel methods, two-dimensional recursive neural networks, and weighted graph matching. Proteins. 2006;62(3):617-629.
30. Fariselli P, Riccobelli P, Casadio R. Role of evolutionary information in predicting the disulfide-bonding state of cysteine in proteins. Proteins. 1999;36(3):340-346.
31. Vullo A, Frasconi P. Disulfide connectivity prediction using recursive neural networks and evolutionary information. Bioinformatics. 2004;20(5): 653-659.
32. Shi W, Punta M, Bohon J, Sauder JM, D'Mello R, Sullivan M, *et al.* Characterization of metalloproteins by high-throughput X-ray absorption spectroscopy. Genome Res. 2011;21(6):898-907.
33. Maullu C, Raimondo D, Caboi F, Giorgetti A, Sergi M, Valentini M, *et al.* Site-directed enzymatic PEGylation of the human granulocyte colony-stimulating factor. FEBS J. 2009;276(22):6741–6750.
34. Golshani M, Rafati S, Jahanian-Najafabadi A, Nejati-Moheimani M, Siadat SD, Shahcheraghi F, *et al*. *In silico* design, cloning and high level expression of L7/L12-TOmp31 fusion protein of Brucella antigens.Res Pharm Sci. 2015;10(5):436–445.
35. Muhammad SA, Ali A, Ismail T, Zafar R, Ilyas U, Ahmad J. *In silico* study ofanti-carcinogenic lysyloxidase-like2 inhibitors. Comput Biol Chem. 2014;(51):71–82.
36. Fontana A, Spolaore B, Mero A, Veronese FM. Site-specific modification and PEGylation of pharmaceutical proteins mediated by transglutaminase. Adv Drug Deliv Rev. 2008;60(1):13-28.
37. Cong Y, Pawlisz E, Bryant P, Balan S, Laurine E, Tommasi R, *et al*. Site-specific PEGylation at histidine tags. Bioconjug Chem. 2012;23(2):248-263.
38. Shaunak S, Godwin A, Choi JW, Balan S, Pedone E, Vijayarangam D, *et al.* Site-specific PEGylation of native disulfide bonds in therapeutic proteins. Nat Chem Biol. 2006;2(6):312-313.

39. Rosendahl MS, Doherty DH, Smith DJ, Carlson SJ, Chlipala EA, Cox GN. A long-acting; highly potent interferon α-2 conjugate created using site-specific PEGylation. Bioconjug Chem. 2005;16(1):200-207.
40. Doherty DH, Rosendahl MS, Smith DJ, Hughes JM, Chlipala EA, Cox GN. Site-specific PEGylation of engineered cysteine analogues of recombinant human granulocyte-macrophage colony-stimulating factor. Bioconjug Chem. 2005;16(5):1291-1298.
41. Alibeik S, Zhu S, Brash JL. Surface modification with PEG and hirudin for protein resistance and thrombin neutralization in blood contact. Colloids Surf B Biointerfaces. 2010;81(2):389-396.
42. Li X, Xiu Z, Zhao J, Li S, Li X, Su Z. An integrated process of PEGylation and separation of hirudin on an anion exchange column. J Biotechnol. 2008;136:S501.
43. Wang XD, Teng H, Hu JJ, Xiu ZL. PEGylation of recombinant hirudin in mixed aqueous-organic solutions. Procbiochem. 2015;50(3):367–377.
44. Qin H, Xiu Z, Zhang D, BaoY, LiX, Han G. PEGylation of hirudin and analysis of its antithrombin activity *in vitro*. Chin J Chem Eng. 2007;15(4):586-590.
45. Otto A, Seckler R. Characterization, stability and refolding of recombinant hirudin. Eur J Biochem. 1991;202(1):67-73.