Original Article

In silico design and analysis of a new hyperglycosylated analog of erythropoietin to improve drug efficacy

Anvarsadat Kianmehr[1, 2], Hamid Shahbaz Mohammadi[1], Mohammad Ali Shokrgozar[3], Eskandar Omidinia[1]

[1]Biochemistry Department, Genetic and Metabolism Research Group, Pasteur Institute of Iran, Tehran [2]Department of Medical Biotechnology, School of Advanced Medical Sciences, Tabriz University of Medical Sciences, Tabriz, [3]National Cell Bank of Iran, New Technologies Research Group, Pasteur Institute of Iran, Tehran, Iran

Abstract

Background: The enhancement of glycosylation by applying glycoengineering approaches has become widely used to boost properties for protein therapeutics. The objective of this work was to engineer a new hyperglycosylated analog of erythropoietin (EPO) with appropriately targeted N-linked carbohydrates through bioinformatics tools.

Materials and Methods: The EPO protein sequence was retrieved from NCBI protein sequence database. Prediction of N-glycosylation sites for the target protein was done using the prediction server, NetNGlyc. The three-dimensional model of glycoengineered EPO (named as kypoetin) was constructed using the homology modeling program. Ramchandran plot obtained from PROCHECK server was used to check stereochemical property. Meanwhile, 3D model of kypoetin with attached N-carbohydrates was built up using the GlyProt server.

Results: In the new modified analog, three additional N-glycosylation sites at amino-acid positions 30, 34 and 86 were inserted. Ramchandran plot analysis showed 81.6% of the residues in the most favored region, 15.6% in the additional allowed, 1.4% in the generously allowed regions and 1.4% in the disallowed region. 3D structural modeling showed that attached carbohydrates were on the proper spatial position. The whole solvent accessible surface areas of kypoetin (15132.69) were higher than EPO (9938.62).

Conclusions: Totally, various model evaluation methods indicated that the glycoengineered version of EPO had considerably good geometry and acceptable profiles for clinical studies and could be considered as the effective drug.

Key Words: Erythropoietin (EPO), glycoengineering, in silico

Address for correspondence:
Dr. Eskandar Omidinia, Biochemistry Department, Genetic and Metabolism Research Group, Pasteur Institute of Iran, Tehran, Iran.
Email: eomid8@yahoo.com

Received: 23.06.2013, **Accepted:** 24.09.2013

Access this article online	
Quick Response Code:	Website: www.advbiores.net
	DOI: 10.4103/2277-9175.161548

INTRODUCTION

Glycosylation is a co-and post-translational modification that involves the selective attachment of carbohydrates to proteins. Carbohydrates are linear and branched, which are covalently linked by glycosidic bonds. They play an important role in modulating physicochemical and biological properties of proteins. Carbohydrates can affect on

How to cite this article: Kianmehr A, Mohammadi HS, Shokrgozar MA, Omidinia E. In silico design and analysis of a new hyperglycosylated analog of erythropoietin to improve drug efficacy. Adv Biomed Res 2015;4:142.

protein stability and solubility, protein function, immunogenicity and susceptibility to proteolysis.[1,2] The rational manipulation of glycosylation parameters (glycoengineering) is widely applied to obtain improved therapeutic proteins. Glycoengineering can enhance in vivo activity even in proteins that do not normally contain N-glycosylation sites.[3] The glycosylation enhancement by applying glycoengineering has become widely used to boost properties for protein therapeutics.[4] Experimental detection of glycosylation sites in proteins is an expensive and laborious process. Therefore, the use of bioinformatics tools to assist the rational design and insertion of glycosylation sites can be helpful.[5] In silico techniques are inexpensive methods that shorten the length of time spent in developing of new and efficient drugs. Successful applications for the use of in silico pharmacology (computational pharmacology) in drug discovery include HIV integrase,[6] urotensin antagonists,[7] CCR5 antagonist[8] and mesangial cell proliferation inhibitor discovery.[9] These efforts suggest that in silico-based approaches have considerable versatility and applicability to design new and efficient drugs.[10]

Erythropoietin (EPO) is a glycoprotein hormone that is the primary regulator of the rate of erythropoiesis. It binds to specific receptors on the cell surface of red blood cell precursors in the bone marrow, promoting their proliferation, differentiation and survival, causing an increase in the circulating red blood cell mass. The gene encoding human EPO was cloned in 1985 leading to the production of recombinant human EPO (rHuEPO). rHuEPO has been used for the treatment of anemia associated with chronic renal failure, cancer and HIV infection, and in the surgical setting to reduce allogeneic blood transfusions.[11, 12] What is more, it serves as a potent neuroprotective cytokine which hinders harm to cells in nervous system following physical and metabolic stresses.[13] EPO is synthesized in the adult kidney where the 165-amino-acid polypeptide is posttranslationally modified to contain three N-linked and one O-linked carbohydrate chains attached at asparagine residues 24, 38 and 83, and serine 126, respectively. Research on EPO indicated that the carbohydrate moieties, in particular the sialic acid residues, are necessary for in vivo biological activity.[14] There is a direct relationship between sialic acid content, serum half-life and in vivo biological activity. Generally, the molecules with the highest sialic acid content have the longest half-life and greatest in vivo biological activity.[15] There is a desire for new erythropoiesis-stimulating molecules that allow for new treatment options, including flexible or less frequent dosing. One successful strategy used to enhance the activity of erythropoiesis-stimulating protein is glycoengineering, whereby consensus sequences for N-linked carbohydrate addition are introduced into the peptide backbone of EPO. In this regard, a new glycoengineered erythropoietic analog entitled darbepoetin alfa has been created.[16] It has been engineered to contain five N-linked carbohydrate chains, two more than rHuEPO. The two new sites of N-linked glycosylation have been introduced by changing five amino acid residues in rHuEPO by site-directed mutagenesis. Darbepoetin alfa has a threefold longer circulating half-life and higher in vivo potency than rHuEPO. Due to its longer half-life and increased potency, darbepoetin alfa can be administered less frequently than rHuEPO to obtain an equivalent biological effect.[17] As the continuation of these researches, the goal of this work was to reengineer and insertion of N-glycosylation sites in EPO through bioinformatics tools whereby consensus sequences for N-linked carbohydrate addition [Asn-Xaa-Ser/Thr] were removed or added into the EPO peptide backbone.

MATERIALS AND METHODS

All experiments were performed by means of in silico technology. The methodology was as follows:

Retrieval of data set

Protein sequence was retrieved from NCBI protein sequence database. The crystal structure of EPO [PDB code: 1BUY] was also obtained from Brookhaven Protein Data Bank [PDB] database.

Engineering and insertion of N-glycosylation sites in EPO

The strategy used in glycoengineering of EPO involved the introduction of N-glycosylation sequons to increase carbohydrate content in target protein. Generally, in eukaryotes, glycan molecules are attached to the asparagine residue from sequons: Asn-x-Ser and Asn-x-Thr, or in some rare cases in Asn-x-Cys where x is not a proline residue.[18] The prediction of N-glycosylation sites for the EPO was done using the available prediction server, NetNGlyc.[19] In the case of NetNGlyc, the predicted Asn-x-Ser/Thr motifs are highlighted in red color, and a graph showing potential N-glycosylation versus amino acids position is also given.

Homology modeling of EPO glycoengineered version

Homology modeling estimates the 3D structure of a given protein sequence by using its alignment to one or more protein template of known structure. The crystal structure of EPO was selected as a template for homology modeling. Multiple sequence alignment was performed with Clustal W program. Alignment were checked for deletions and insertions in structurally conserved regions and finally fine-tuned manually

modified before 3D modeling. The 3D model of kypoetin was constructed using the homology modeling program Modeler version 9v4 [default parameters], based upon the crystal structure of EPO. The geometry of loop regions was corrected using MODELER/Refine Loop command.

Model evaluation and validation

Ramchandran plot obtained from PROCHECK server was used to check stereochemical property.[20] It provides the facility to check the stereochemical quality of a protein sequence by analyzing residue-by-residue geometry. Model constructed from MODELLER was finally chosen for subsequent analysis as they possessed good geometry and energy profile. Visualization and analysis of the model were done with PyMOL program, which was also used for illustrations.

Prediction of kypoetin 3D structure with attached N-glycans

Tertiary glycoprotein structure having attached N-glycans for kypoetin was modeled using the GlyProt server. This facilitates the identification of spatially unfavorable N-glycosylation sites.[21] The 3D structure model was built using the structure of the native target protein as a template. The sequence used as input to build the 3D model has to contain the inserted N-glycan sequins.

RESULTS

The process flowchart for glycoengineering is depicted in Figure 1. Firstly, the EPO protein sequence was retrieved from NCBI database. Glycosylation pattern was predicted using NetNGlyc [Figure 2]. The NetNglyc server predicts N-Glycosylation locations in proteins using artificial neural networks that examine the sequence context of Asn-Xaa-Ser/Thr sequons. As shown in Figure 2a, the predicted N-glycan sites in the modified target protein were six positions. All these sites had required potential for N-glycosylation [Figure 2b]. To analyze the structural model, modified sequence was applied to perform homology modeling using Modeller program [Figure 3b]. The alignment of the EPO new variant with its native sequence is shown in Figure 3a. 1000 models were constructed using MODELLER. The best model was selected for subsequent analysis by PROCHECK program.[20] In this model, Ramchandran plot analysis showed 81.6% of the residues in the most favored region, 15.6% in the additional allowed, 1.4% in the generously allowed regions and 1.4% in the disallowed region [Figure 4a]. 3D glycoprotein structures of the engineered EPO (named as kypoetin) and also darbepoetin alfa with attached N-glycans was predicted using the GlyProt server [Figure 5]. Based on the physico-chemical features [Table 1], structural modeling of new analog showed that the attached carbohydrates were on the proper spatial position. For example, total solvent accessible surface areas of kypoetin [15132.69] were higher than EPO [9938.62].

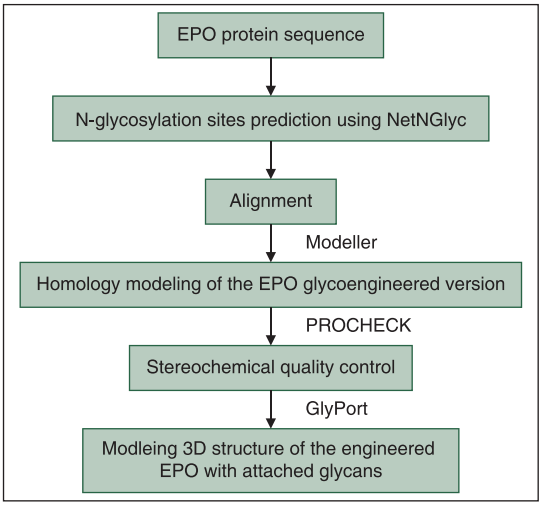

Figure 1: In silico glycoengineering of EPO protein. The scheme illustrates the steps involved in this process

DISUCSSION

Increasingly over the last decade in silico pharmacology has been developed and applied to pharmacology hypothesis development and testing. Computational methods are helping us to make decision and simulate virtually every facet of drug discovery, moving the pharmaceutical industry closer to engineering-based disciplines.[10] In this research, we aimed to engineer and evaluate new functional N-glycosylation sites

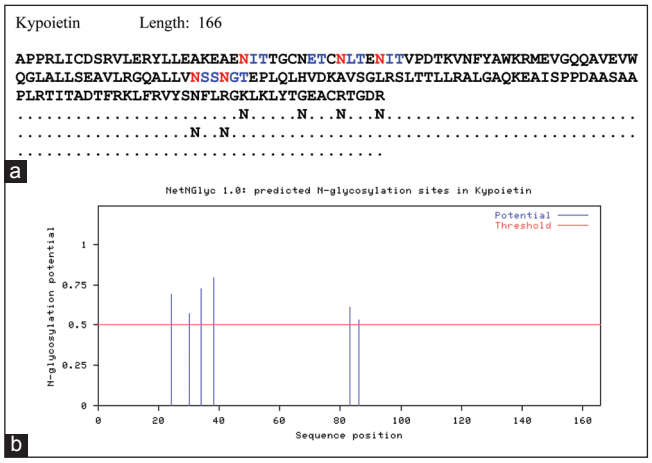

Figure 2: Predicted N-glycosylation pattern in the EPO new variant. (a) The predicted Asn-x-ser/Thr motifs. Asn-Xaa-Ser/Thr sequons in the sequence output are highlighted in blue. Asparagines predicted to be N-glycosylated are highlighted in red. N represents a predicted N-glycosylation site. (b) A graph showing potential N-glycosylation versus amino acids position is also given

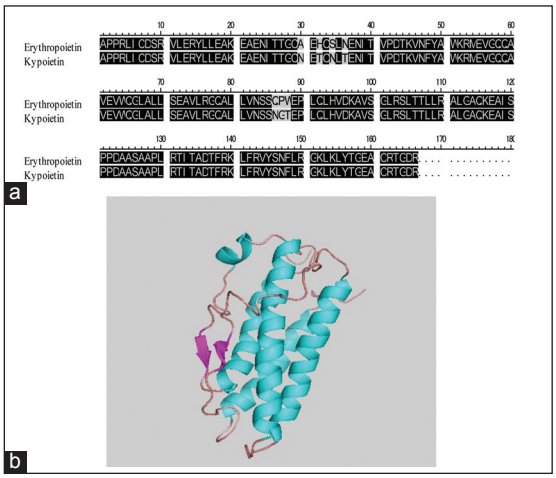

Figure 3: The alignment (a) and the structural model (b) of the kypoetin are shown as a representative homology-based structural model. The identical residues between the query and template are colored black. The modified residues in Kypoetin are in grey. The structural model retrieved from Modeller is rendered as a cartoon using PyMol [http://www.pymol.org]

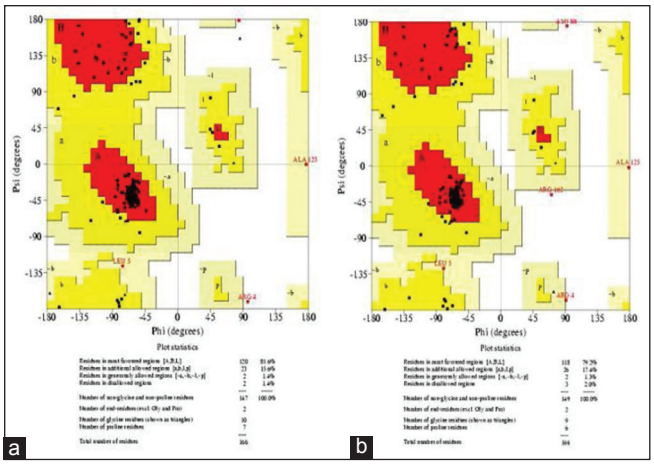

Figure 4: Ramchandran plot analysis of kypoetin (a) and darbepoetin alfa (b)

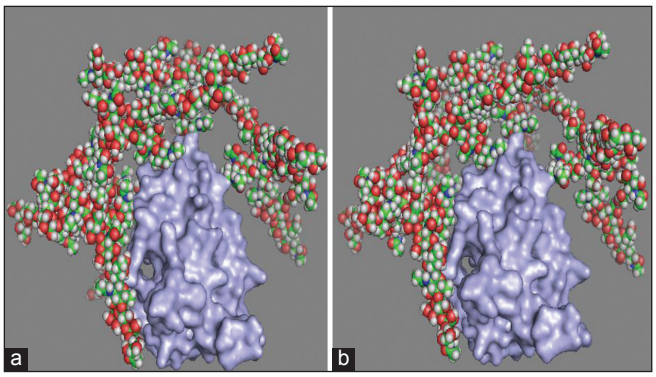

Figure 5: The three-dimensional structures of glycosylated kypoetin (a) and darbepoetin alfa (b). The modified target protein is given as a surface representation. The attached N-glycan molecules are represented as a space fill model

Table 1: Comparison of some physico-chemical properties of the EPO and kypoetin

Content	EPO	Kypoetin
Number of non-HoH, non-H atoms	1284	1589
Total ASA	9938.62	15132.69
Polar	4712.89	7257.99
Non-polar	5225.62	7874.70
+charge	1360.46	1325.96
−charge	778.52	749.52
Structure contains cavities	12	12

ASA: Solvent accessible surface areas

in EPO protein using bioinformatics approaches to improve its pharmacokinetics efficacy.[22] Our designed locations were three more than rHuEPO and one more than darbepoetin alfa.[16] The new hyperglycosylated analog, designed in this work, was named as kypoetin. These N-linked carbohydrates were introduced at amino acids positions 24, 30, 34, 38, 83 and 86. The positions 24, 30, 38 and 83 are the same as those on EPO and darbepoetin alfa, however, the positions 34 and 86 were newly inserted. According to this result, it can be anticipated that more carbohydrates attach to kypoetin. The quality of protein model verifies the informatics can be mined from it. Hence, evaluation of the accuracy of protein models is essential for their interpretation. For this purpose, stereo-chemical properties of the model were evaluated by PROCHECK program.[20] Prominently, these features were nearly better than the obtained results by darbepoetin alfa [Figure 4b]. In comparison to darbepoetin alfa [79.2%], 81.6% of kypoetin residues were in the most favored region. It ensured that most residues were in consistent phi-psi distribution and were reliable for further analysis. Finally, 3D structures with attached N-glycans for the kypoetin and darbepoetin alfa were built and compared. As can be found [Figure 5], the glycosylated 3D models of kypoetin and darbepoetin alfa were similar. Therefore, it can be expected that the physico-chemical properties of these two glycoproteins were nearly identical; although, the biological activities may be different because the N-glycosylation locations were not the same. Our data indicated the positive influence of attached N-glycans on the conformational changes of EPO new variant. Similar studies on the glycoengineering of other proteins such as leptin and MPI ligand has also been preformed.[3] Overwhelmingly, our interpretation of this project was that the prediction of protein glycosylation helps us to find out important analogs of glycoproteins. Successful examples of in silico approach for the development of novel therapeutics have been reported.[10] For examples, in silico has been applied for the development of interventions in the neuroprotection and neurotherapy of Alzheimer's disease.[23] Here, we propose an idea to reengineer the analysis of glysoylated analogs using computational

methods before starting the experimental phases. Homology modeling and molecular dynamic are the most important techniques to fulfill the desired results. In summary, various model evaluation methods indicated that the EPO glycoengineered version with three additional N-linked carbohydrates had considerably good geometry and acceptable profiles for clinical studies and can be considered as the effective drug. Our proposed structure suggested the possible application of kypoetin for production and clinical studies. All research was performed using in silico techniques. Further experiments for cloning, recombinant production and the clinical evaluation of kypoetin are in progress. This is the first report on the in silico glycoengineering of EPO.

ACKNOWLEDGMENTS

This project was financially supported by the Pasteur Institute of Iran. The authors wish to express their deep gratitude to all who provided support, especially enzyme technology lab. of Metabolic and Genetic Research Group, Pasteur Institute of Iran.

REFERENCES

1. Sola RJ, Griebenow K. Glycosylation of therapeutic proteins: An effective strategy to optimize efficacy. Bio Drugs 2010;24:9-21.
2. Vidal D, Garcia-Serna R, Mestres J. Ligand-based approaches to in silico pharmacology. Methods Mol Biol 2011;672:489-502.
3. Elliott S, Lorenzini T, Asher S, Aoki K, Brankow D, Buck L, *et al*. Enhancement of therapeutic protein in vivo activities through glycoengineering. Nat Biotechnol 2003;21:414-21.
4. Frank M, Schloissnig S. Bioinformatics and molecular modeling in glycobiology. Cell Mol Life Sci 2010;67:2749-72.
5. Ortega SS, Cara LC, Salvador MK. In silico pharmacology for a multidisciplinary drug discovery process. Drug Metab Drug Interact 2012; 27:199-207.
6. Carlson HA, Masukawa KM, Rubins K, Bushman FD, Jorgenson WL, Lins RD, *et al*. Developing a dynamic pharmacophore model for HIV-1 integrase. J Med Chem 2000;43:2100-14.
7. Flohr S, Kurz M, Kostenis E, Brkovich A, Fournier A, Klabunde T. Identification of nonpeptidic urotensin II receptor antagonists by virtual screening based on a pharmacophore model derived from structure – activity relationships and nuclear magnetic resonance studies on urotensin II. J Med Chem 2002;45:1799-805.
8. Debnath AK. Generation of predictive pharmacophore models for CCR5 anatagonists: Study with piperidine- and piperazine-based compounds as a new class of HIV-1 entry inhibitors. J Med Chem 2003;46:4501-15.
9. Kurogi Y, Miyata K, Okamura T, Hashimoto K, Tsutsumi K, Nasu M. Discovery of novel mesangial cell proliferation inhibitors using a three-dimensional database searching method. J Med Chem 2001;44:2304-7.
10. Ekins S, Mestres J, Testa B. In silico pharmacology for drug discovery: Methods for virtual ligand screening and profiling. Br J Pharmacol 2007;152:9-20.
11. Elliott S, Chang D, Delorme E, Eris T, Lorenzini T. Structural requirement for additional N- linked carbohydrate on recombinant human erythropoietin. J Biol Chem 2004;279:16854- 62.
12. Elliott S, Pham E, Macdougali IC. Erythropoietins: A common mechanism of action. Exp Heamatol 2008;36:1573-84.
13. Bartesaghi S, Marinovich M, Corsini E, Galli CL, Viviani B. Erythropoietin: A novel neuroprotective cytokine. Toxicology 2005;26:923-8.
14. Elliott S, Egrie J, Browne J, Lorenzini T, Busse L, Rogers N, *et al*. Control of rHuEPO biological activity: The role of carbohydrate. Exp Heamatol 2004; 32:1146-55.
15. Egrie JC, Dwyer E, Browne JK, Hitz A, Lykos MA. Darbepoetin alfa has a longer circulating half-life and greater in vivo potency than recombinant human erythropoietin. Exp Heamatol 2003;31:290-9.
16. Egrie JC, Browne JK. Development and characterization of darbepoetin alfa. Oncology 2002;16:1-11.
17. Shakin-Eshleman SH, Spitalnik SL, Kasturi L. The amino acid at the X position of an Asn-X-Ser sequon is an important determinant of N-linked core-glycosylation efficiency. J Biol Chem 1996;271:6363-6.
18. Julenius K. NetCGlyc 1.0: Prediction of mammalian C-mannosylation sites. Glycobiology 2007;17:868-76.
19. Laskowski R, Macarthur M, Moss D, Thornton J. PROCHECK: A program to check the stereochemical quality of protein structures. G Appl Crys 1993;26:283-91.
20. Bohne-Lang A, von der Lieth CW. GlyProt: In silico glycosylation of proteins. Nucleic Acids Res 2005;33:214-9.
21. Der Lieth CW, Bohne-Lang A, Lohmann KK, Frank M. Bioinformatics for glycomics: Status, methods, requirements and perspectives. Brief Bioinform 2004;5:164-78.
22. Mahal LK. Glycomics: Towards bioinformatic approaches to understanding glycosylation. Anticancer Agents Med Chem 2008;8:37-51.
23. Kumar P, Pillay V, Choonara YE, Modi G, Naidoo D, Toit LC. In Silico theoretical molecular modeling for Alzheimer's disease: The nicotine-curcumin paradigm in neuroprotection and neurotherapy. Int J Mol Sci 2011;12:694-724.

Source of Support: Pasteur Institute of Iran, **Conflict of Interest:** The authors declare that there is no conflict of interests regarding the publication of this article.