Pakistan J. Zool., vol. 38(2), pp. 85-92, 2006.

Mammalian Epidermal Growth Factor Receptor: Role of Phosphorylation and Glycosylation

TAYYABA WAJAHAT, AFSHAN KALEEM, ISHTIAQ AHMAD, DANIEL C. HOESSLI, EVELYNE WALKER-NASIR, WAJAHAT M. QAZI, MUHAMMAD SALEEM, ABDUL RAUF SHAKOORI AND NASIR-UD-DIN*

Institute of Molecular Sciences and Bioinformatics, Lahore, Pakistan (TW, AK, IA, EWN, WQ, MS & NUD); Department of Pathology and Immunology, Centre Médical Universitaire, Geneva, Switzerland (DCH); School of Biological Sciences, University of the Punjab Lahore, Pakistan (ARS) and HEJ Research Institute of Chemistry, University of Karachi, Karachi, Pakistan (NUD)

Abstract.- Large number of proteins *in vivo* perform multiple functions. The post-translational modifications induce conformational, more often transitory, changes in most of the cytoplasmic and nuclear proteins, which result in the regulation of diverse functions including signal transduction cascade. The *O*-β-GlcNAc modification on Ser/Thr has been described as dynamic as phosphorylation and interplay between the two modifications at the same or neighboring residue(s) can result in diverse functional regulation of proteins. Assessing temporary functional changes due to transitory protein structural changes *in vivo* is difficult. Computer assisted predictions of modification potential of proteins direct possible functional changes and help design experiments to verify these possible multiple functionality of proteins. In this study using computer-assisted techniques functional changes of epidermal growth factor receptor (EGFR) were investigated due to *O*-GlcNAc and phosphorylation modifications of EGFR-protein.

Key words: Post transcriptional modifications, signal transduction, neural network based predictions, 3D structure of proteins, Computer-assisted modifications.

INTRODUCTION

Epidermal growth factor receptor (EGFR) is a phosphoprotein, which is present on the cell membrane as a transmembrane protein and plays significant role in regulation of cellular growth, differentiation, proliferation, survival, apoptosis and cell migration (Schlessinger, 2000). EGFR is activated by binding with its specific ligand, including epidermal growth factor (EGF) and transformation growth factor alpha (TGF-∝) besides others, thereby stimulating an intrinsic protein receptor tyrosine kinase (RTK) activity in the cytoplasmic domain. Once EGFR binds to a ligand it interacts with its neighboring monomer (EGFR, Erb2, Erb3 or Erb4) and dimerizes (either as a homo- or hetrodimer). After dimerization of EGFR, its tyrosine (Tyr) kinase domain auto-phosphorylates in its cytoplasmic tail on specific Tyr residues (992, 1068, 1086, 1148 and 1173) (Downward *et al*., 1984). Auto-phosphorylation is an intramolecular process and is more efficient than exogenous substrate phosphorylation (Bertics and Gill, 1985). Auto-phosphorylation elicits down-stream activation and signaling events of other proteins. Once EGFR is phosphorylated it becomes docking site for Src homology 2 (*SH2)* containing signal proteins (like Gbr-2 and Shc) or phospho-tyrosine-binding (PTB) proteins (Schlessinger, 2000), and provide a crucial connection between the external stimuli and internal signal transduction pathways which is considered as an important step in signal transduction cascade.

Several kinases also phosphorylate EGFR independently such as Ca^{2+}/calmodulin dependent kinase II (CK-II), mitogen activated protein kinase (MAPK) and protein kinase A and C (PKA and PKC). MAPK phosphorylates EGFR on Thr 669 *in vivo* (Morrison *et al*., 1993) and indirectly regulates receptor tyrosine kinase (RTK) by activating a tyrosine phosphatase SHP-I. The SHP-I dephosphorylates EGFR on Tyr 1173 and presents a major mechanism of negative regulation of tyrosine kinase receptor signaling (Tomic *et al*., 1995).

Signal transduction by EGFR is regulated by change in its phosphorylation status. Phosphorylated EGFR deactivates its protein tyrosin kinase activity.

* To whom correspondence be addressed. E-mail: prof_nasir@yahoo.com, nasir@super.net.pk

0030-9923/2006/0002-0085 $ 8.00/0

Induction of post-translational modifications (PTMs) in EGFR is an important mechanism to regulate its function. One of the most common PTMs in EGFR is phosphorylation that regulates its activity and in its response this phosphorylated EGFR by exogenous kinases deactivates its protein tyrosine kinase activity, resultantly the signal transduction cascade is stopped. This phosphorylation and dephosphorylation is a common mechanism for regulation of protein function. Like phosphorylation, glycosylation is also crucial in regulating protein functions, for example phosphorylation on serine and threonine can be inhibited by the addition of *O*-GlcNAc on these residues (Kearse and Hart, 1991). Interplay between *O*-GlcNAc modification and phosphorylation on the same amino acid residues has been observed in several nucleo- and cytoplasmic proteins (Kamemura *et al*., 2002).

These PTMs are dynamic and result in temporary conformational changes and regulate many functions of the protein. This observation led to the formulation of the Yin Yang hypothesis, according to which phosphorylation and *O*-glycosylation compete for the same site or on the same region of a polypeptide (Zachara and Hart, 2002). These sites can be predicted and analyzed, using various computer-assisted neural network-based programs, which can help to determine protein functions by assessing their modification potential. Such computational methods have become intrinsic to modern biological research.

In this report, we describe Ser/Thr of EGFR as potential sites for *O*-linked glycosylation and phosphorylation and their possible interplay. These sites have been predicted and analyzed by using different prediction methods. The conserved phylogenetic residues are known to act as key functional sites (La *et al*., 2005). PTMs at these conserved residues may regulate certain functions of protein so these sites are expected to act as important functional sites.

MATERIALS AND METHODS

Sequence data

The *Mus musculus* was selected as representative of mammals. The sequence of EGFR of mouse has been described by many workers (Paria *et al*., 1993; Avivi *et al.*, 1992). The sequence data used for predicting phosphorylation and *O*-linked glycosylation sites for EGFR of *Mus musculus* was retrieved from SWISS-PROT (Luetteke *et al*., 1994) sequence data base with entry name EGFR_MOUSE and primary accession number Q01279. SWISS-PROT result showed that first 24 amino acids are signal peptide. BLAST search was made using the NCBI data base of non-redundant sequences (Altschul *et al*., 1997). The BLAST search was made for all organisms' sequences with expect value set to 10 using blosum62 matrix and low complexity filter selecting nr database. A total of 1539 hits were obtained. Of these 1539 blast hits the first 170 were with highest bit scores and zero expect value. Among these170 hit scores 12 hits with highest bit score and zero expect value were selected. The 11 selected sequences were from different vertebrates groups: six from mammals (*Mus musculus, Rattus norvegicus, Canis familiaris, Sus scrofa, Bos taurus* and *Felis catus* with primary accession numbers; CAA55587.1, NP_113695.1, XP_533073.2, NP_ 999172.1, XP_592211.2, and AAW23986.1, respectively), three from pisces (*Danio rerio, Xiphophorus xiphidium, Tetraodon nigroviridis* with their primary accession numbers; NP_919405.1, AAP55673.1, and CAG07098.1, respectively) and one from aves (*Gallus gallus,* with the accession numbers NP_001025536.1, Isoform of selected sequences were neglected.

To find out conserved residues in *Mus musculus* EGFR, all selected sequences were multiple aligned using CLUSTALW (Wilbur and Lipman, 1983). CLUSTALW is a general purpose multiple sequence alignment program for DNA or proteins. It produces biologically meaningful multiple sequence alignment of divergent sequences. It calculates the best match for the selected sequences, and lines them up so that the identities, similarities and differences can be seen. The sequence of EGFR was multiple aligned using CLUSTALW in three steps. In the first step the sequence of *Mus musculus* was aligned with all other mammalian species described above. In the second step all the six mammalian sequences were aligned with those of pisces, and in the third and

final step all the selected 12 sequences including that of *Gallus gallus,* an avian, were aligned collectively.

Post-translational modification prediction methods

Method used for predicting potential phosphorylation sites in *Mus musculus* EGFR was NetPhos 2.0 (Blom *et al*., 1999) (http:// www.cbs.dtu.dk/services/Netphos/). NetPhos 2.0 is a neural network-based method for predicting potential phosphorylation sites at Ser, Thr, and Tyr in protein sequences, with sensitivity in the range from 69% to 96%. NetphosK 1.0 (Blom *et al*., 2004) (http:// www.cbs.dtu.dk/services/NetPhosK/) was used to predict Kinase specific phosphorylation sites in mouse EGFR. The NetPhosK 1.0 predicts the kinase specific acceptor substrate sites in a protein.

To predict potential *O*-β-GlcNAc modification sites YinOYang 1.2 (http://www. cbs.dtu.dk/services/YinOYang/) was used. This method is also capable of predicting the potential phosphorylation sites as well and hence predicting the Yin Yang sites. The potential Yin Yang sites were also predicted using this method.

Neural networks-based prediction methods

Artificial neural networks based methods have been extensively used in biological sequence analysis and predicting the potentials of proteins for modifications (Baldi and Brunak, 2002). The methods developed using machine learning approach includes memorizing the neural networks with the sequence environment windows of phosphorylated / glycosylated and non-phosphorylated / non-glycosylated sites. The input data in form of sequence windows is encoded by a binary encoding system. A threshold value in forms of bits is set for positive hit and zero for negative hits. The learning process and performance is checked with the data reserved for cross validation using statistical equations. During learning the error is computed and weights given to each neuron are set to get the maximum correct predictions, reducing the error and hence increasing the false positive and false negative prediction sites.

RESULTS

The possible phosphorylation sites predicted by Net Phos 2.0 in *Mus musculus* have shown a high potential for phosphorylation. A total of 52 residues were predicted to be phosphorylated (Fig. 1). These include 20 Ser (at 203, 205, 222, 262, 356, 428, 433, 506, 730, 935, 969, 1006, 1015, 1020, 1033, 1046, 1047, 1080, 1142, 1181), 13 Thr (at 15, 235, 373, 548, 656, 671, 703, 729, 870, 980, 1007) and 19 Tyr (at 45, 64, 89, 93, 261, 291, 447, 561, 602, 742, 779, 847, 994, 1068, 1086, 1101, 1114, 1148, 1173) (Fig.1). Similarly, prediction results of EGFR for *O*-β-GlcNAc modification sites showed that the protein had the potential for *O*-β-GlcNAc modification. The prediction results for *O*-GlcNAc modification showed that there were 12 potential sites highly predicted to be modified by *O*-GlcNAc, 9 on Ser residues, at positions 185, 186, 366, 525, 923, 1032, 1130, 1204, and 1205 and 3 on Thr residues at positions 302, 1031,and 1034 (Fig. 2). There were 2 Yin Yang sites according to the prediction results (at Thr 1007, and Ser 1180) (Fig.3).

Apart from these, there are many other Ser and Thr residues, which were predicted to be non-glycosylated but having potential for phosphorylation, either very close to the threshold or higher than the threshold value. Interestingly, these residues are also conserved in mammals. Such residues appear to be as false-negative sites. These may act as possible Yin Yang sites other than those which have been predicted by the YinOYang 1.2 method. The false negative sites (FN), along with their conservation status and modification potential, are shown in Table 1. As different kinases are involved in the phosphorylation of mammalian EGFR so kinase specific phosphorylation substrate potential was accessed by using NetPhosK 1.0. The results obtained by NetPhosK 1.0 had shown the involvement of different kinases in phosphorylation of predicted FN-YinYang sites of mammalian EGFR. PKA can phosphorylate Ser 222, 1046, 1142 and Thr 654, similarly PKC can phosphorylate Ser (428, 935) and Thr (235, 238), while Thr 671 can be phosphorylated by p38MAPK. Calmodulin dependent kinase I and II (CKI and CKII) can also phosphorylate Ser 433 and 1142, respectively.

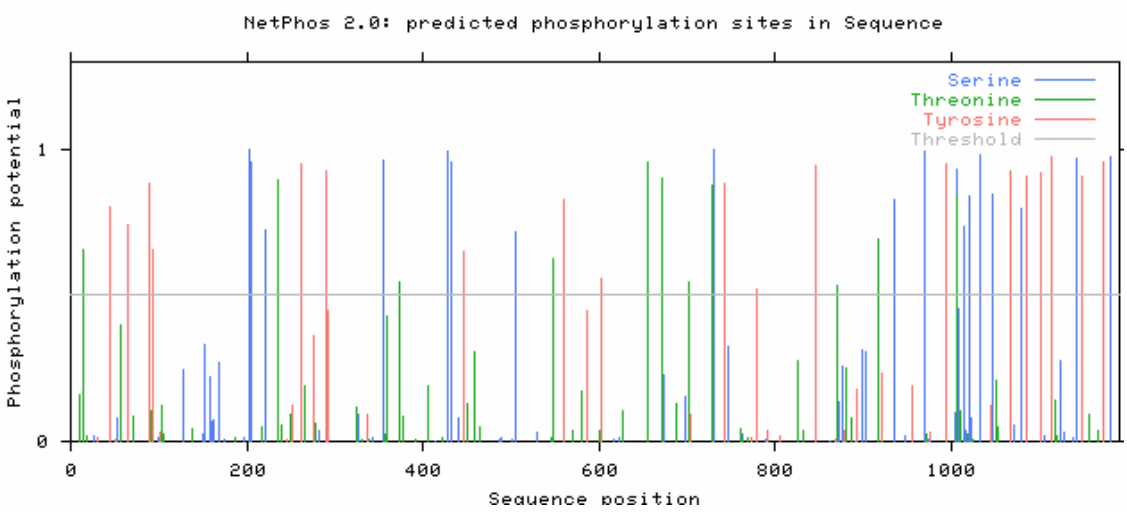

Fig. 1. Graphic presentation of potential for phosphate modification at serine and threonine residues in the mammalian EGFR. Green vertical lines show the potential of Serine / Thr residue for phosphate modification and light blue horizontal wavy line shows threshold for modification potential.

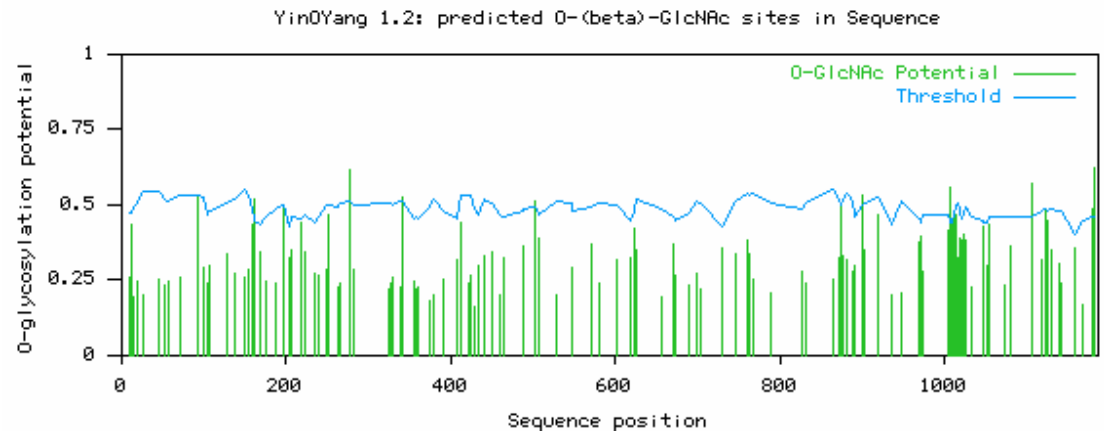

Fig. 2. Graphic representation of potential for O-GlcNAc modification for serine and threonine residues in mammalian EGFR. Here green vertical lines show the potential of Serine / Thr residue for O-GlcNAc modification and light blue horizontal wavy line shows threshold for modification potential.

DISCUSSION

The post-translational modifications such as phosphorylation and glycosylation may regulate proper folding and maintain the 3D structure of a protein. (Hounsell *et al.*, 1996). To understand the biological functions of a protein, knowledge of its 3D structure is essential (Siew and Fischer, 2001). However, determination of the 3D structure *in vivo* is difficult, because given conformations is constantly being modified by intra- and intermolecular interactions between the proteins present in body fluid or in the cell. Computer-assisted methods are useful in predicting the molecular interactions between multifunctional proteins *in vivo*. Such techniques access the modification potential of the proteins involved. Fast output results and easy accessibility have made these theoretical prediction methods vital for the study and elucidation of the biological functions of

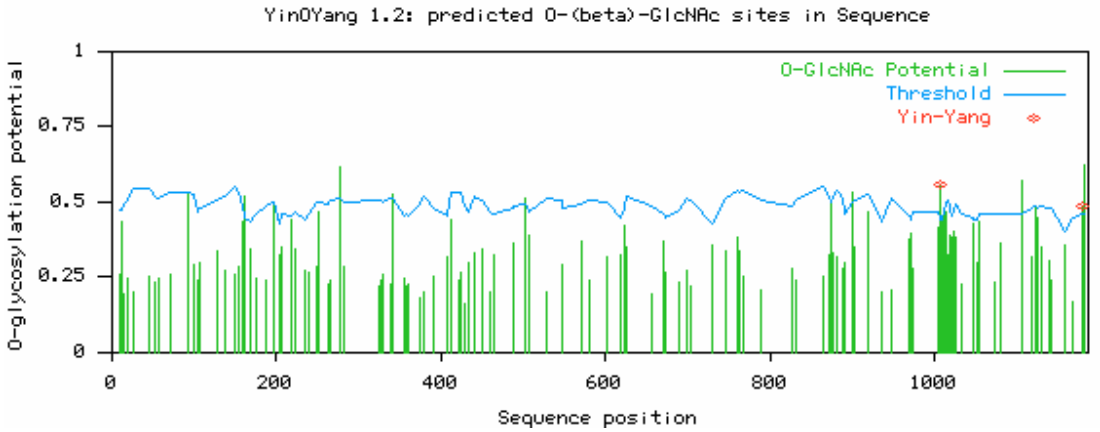

Fig. 3. Graphic representation of sites with potential for both *O*-GlcNAc and phosphate modifications, the Yin Yang sites of mammalian EGFR. Here green vertical lines show the potential of Serine / Thr residue for O-GlcNAc modification, while green lines with red asterisk at top representing the Yin Yang sites, and light blue horizontal wavy line shows threshold for modification potential.

proteins. Predictions performed by these methods have been shown to be accurate (Nielsen *et al*., 1999) and it also provides important hints and directions for practical verification. Computer-assisted techniques were, therefore, used to elucidate the PTMs of EGFR.

Signal transduction initiated by EGFR is regulated by the status of activated second messenger pathway in cell. The molecular detail of how the integration of signaling occurs is incompletely understood. However, the functional control of EGFR is considered to be regulated by change in phosphorylation status of EGFR. Upon binding of EGFR to its ligand, EGFR undergoes dimerization and shows its tyrosine kinase activity. This leads to the auto-phosphorylation of EGFR (Schlessinger, 2000). The five auto-phosphorylation sites have already been described in the cytoplasmic domain of mammalian EGFR at tyrosine residues including 992, 1068, 1086, 1148, and 1173 (Downward *et al*., 1984; Walton *et al*., 1990). These are all conserved sites in mammals. According to prediction results of NetPhos 2.0 all the above described auto-phosphorylation sites have potential for phosphorylation. These phosphorylated tyrosine residue serve as the binding domain for growth factor receptor receptor bound protein Grb2/Son of seven less protein homolog (SOS) complex and activate the MAPK signaling cascade which influence the cell proliferation, migration, and differentiation (Oksvold *et al*., 2002).

It has been described earlier that the binding of EGFR to its ligand causes activation of its tyrosine kinase domain (Schlessinger, 2000). The desensitization is important physiological function that controls the receptor signaling. This process of desensitization is associated with an inhibition of tyrosin kinase activity of EGFR (Countaway *et al*., 1992). The phosphorylation of EGFR at Thr 654 by protein kinase C decreases the intrinsic tyrosine kinase activity of EGFR (Davis, 1988). In addition to this phosphorylation of serine 1002 by Cell division control protein 2 (cdc2) and Alpha- and gamma-adaptin-binding protein (p34) (Kuppuswamy *et al*., 1993) and of Ser (1046/1047) by CKII (Countaway *et al*., 1992) are also associated with inhibition of tyrosine kinase. In addition to these sites the mammalin EGFR can also be phosphorylated on Thr 669 by MAPK and on tyr 845, 891, 820 (Northwood *et al*., 1991), but these Tyr residues are not auto-phosphorylated. These phoshotyrosine can also provide docking side for SH2 domain containing protein. These above described residues are all conserved among mammals. According to prediction results of NetPhos 2.0, among these known phosphorylated sites Ser 1046, 1047 and Thr 669 have the potential for phosphorylation.

As for *O*-GlcNAc modification of EGFR is concerned direct evidences are not available. As

Table I: Conserved Ser /Thr residue with post-translational modifications of human EGFR.

No. of observations	Residues	Conservation status			Modification potential		Yin Yang sites
		Mammals	Pisces	Aves	Phosphate	O-GlcNAc	
1	Thr 1007	CS	NCS	CS	VHP	VHP	Pr
2	Ser 1180	NC	CS	NCS	HP	VHP	Pr
3	Ser 222	CR	NCS	CS	VHP	VCT	FN
4	Thr 235	NC	NCS	NCS	HP	VCT	FN
5	Ser 356	NCS	NCS	NCS	HP	VCT	FN
6	Ser 428	CR	CR	CR	HP	VCT	FN
7	Ser 433	NCS	CR	CR	VHP	VCT	FN
8	Ser 506	NC	CR	CR	VHP	VCT	FN
9	Thr 656	CR	CR	CR	VHP	VCT	FN
10	Thr 671	CR	CR	CR	VHP	VCT	FN
11	Ser 935	NCS	NCS	NCS	HP	VCT	FN
12	Ser 1006	CS	CS	CS	VHP	VCT	FN
13	Ser 1046	CR	NC	NC	VHP	VCT	FN
14	Ser 1047	NCS	NC	NC	VHP	VCT	FN
15	Ser 1142	NC	CR	CR	HP	VCT	FN
16	Ser 1008	CR	CS	NCS	VCT	VHP	FN

Abbreviations used: CR, conserved residue**;** CS, conserved substitution residue; FN, False negative; HP, High potential; N, Negative; NC, non-conserved; NCS, Non conserved substitution residue; Pr, Predicted Ying Yang site; VCT, Very close to threshold; VHP, Very high potential.

reviewed by Zachara and Hart (2002) the site mapping of *O*-GlcNAc modified protein had shown that *O*-GlcNAc occur at sites on proteins that are similar to those modified by protein kinases. Furthermore, *O*-GlcNAc can compete for phosphorylation at the same or neighboring sites on protein. The prediction results of NetPhosK 1.0 had also shown the potential of phosphorylation by specific kinases in EGFR. So *O*-GlcNAc modification may occur on these specific sites where kinases are involved in phosphorylation. Further more as phosphorylation and glycosylation compete for the same site on proteins so it is possible that interplay of phosphorylation and glycosylation may occur on these sites. Moreover YinOYang 1.2 prediction results had showed, that EGFR has the potential for *O*-linked glycosylation (Fig. 3). On the bases of conserved nature of Ser/Thr residues and their modification potential for phosphorylation and *O*-β-GlcNAc modification the possible FN-YinYang sites of murine EGFR has been identified that are Ser222, 356, 428, 433, 506, 1006, 1046, 1047, 1142 and Thr 235, 654 and 659 (Table I).

Similarly, Ser 1008 has very high potential for *O*-β-GlcNAc modification, while its potential for phosphorylation is very close to threshold value, and this site is also conserved in mammals, aves, and pices so this is also predicted as a FN-YinYang site (Table I). The interplay of phosphorylation and glycosylation on these predicted YinYang sites of EGFR may possibly modulate the function of Tyr kinase activity of EGFR.

It is concluded, that phosphorylation of EGFR provides an important mechanism to regulate its function proteins in addition to its interplay with *O*-GlcNAc modification at different sites in cytoplasmic part. The phosphorylation of EGFR by PKC, CKII and MAPK inhibit its tyrosine kinase activity, and thereby terminate the signal transduction, which results in down regulation of EGFR, while *O*-GlcNAc modification inhibits phosphorylation on specific sites, which may result in sustaining the signal transduction. We propose here that alternative phosphorylation and *O*-GlcNAc modification at Thr 1007, and Ser 1180 in the catalytic domain of EGFR are potential Yin Yang sites and are involved in modulating the function of EGFR. Moreover Ser 222, 356, 428, 433, 506 of extra cellular domain,Thr 656 and 671 of tyrosine kinase domain and Ser 1006, 1008, 1046, 1047, 1142, and Thr 1007of catalytic domain are proposed as FN-Yin Yang sites involved in PTMs of mammalian EGFR. These FN-sites can also act as

substrates for OGT due to similarity of kinase and OGT recognition and modification site. We propose the involvement of *O*-GlcNAc modification in blocking the phosphorylation on these predicted FN sites. Thus an addition of *O*-GlcNAc is proposed to be involved in desensitization of Tyr kinase activity, involved in maintaining a normal cellular growth until removed by *O*-GlcNAcase and phosphorylation can occur on the same site(s) leading to down regulation of EGFR's function followed by degradation.

ACKNOWLEDGEMENTS

NUD thanks HEC Pakistan for DNP grant and PAS for partial support to this work.

REFERENCES

ALTSCHUL, S.F., MADDEN, T.L., SCHÄFFER, A.A., ZHANG, J., ZHANG, Z., MILLER, W. AND LIPMAN, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res*., **25**: 3389-3402.

AVIVI, A., SKORECKI K., YAYON A. AND GIVOL, D., 1992. Son of seven less protein homolog 1. Promoter region of the murine fibroblast growth factor receptor 2 (bek/KGFR) genes. *Oncogene*, **7**: 1957-1962.

BALDI, P. AND BRUNAK, S., 2002. *Bioinformatics, The machine learning approach*. 2nd edition, MIT Press, Cambridge., MA.

BERTICS, P. J.AND GILL, G. N. 1985. Self-phosphorylation enhances the protein-tyrosine kinase activity of the epidermal growth factor receptor. *J. biol. Chem,* **260**: 14642-14647.

BLOM, N., GAMMELTOFT, S. AND BRUNAK, S., 1999. Sequence- and structure-based prediction of eukaryotic protein phosphorylation sites. *J. mol. Biol.,* **294**: 1351–1362.

BLOM, N., SICHERITZ-PONTEN, T., GUPTA, R., GAMMELTOFT, S. AND BRUNAK, S., 2004. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics,* **4**: 1633-1649.

COUNTAWAY, J. L., NAIRNG, A. C. AND DAVIS, R., J. 1992. Mechanism of desensitization of the epidermal growth factor receptor protein-tyrosine kinase. *J. biol. Chem,* **267:**1129-1140.

DOWNWARD, J., PARKER, P. AND WATERFIELD, M.D., 1984. Autophosphorylation sites on the epidermal growth factor receptor. *Nature,* **311**: 483-485.

DAVIS, R.J., 1988. Independent mechanisms account for the regulation by protein kinase C of the epidermal growth factor receptor affinity and tyrosine-protein kinase activity. *J. biol. Chem,* **263:** 9462-9469.

HOUNSELL, E.F., DAVIES, M.J. AND RENOUF, D.V., 1996. O-linked protein glycosylation structure and function. *Glycoconj. J.,* **13**:19-26.

KAMEMURA, K., HAYES, B.K., COMER, F.I. AND HART, G.W., 2002. Dynamic interplay between *O*-glycosylation and *O*-phosphorylation of nucleocytoplasmic proteins: alternative glycosylation/ phosphorylation of THR-58, a known mutational hot spot of c-Myc in lymphomas, is regulated by mitogens. *J. biol. Chem*, **277**: 19229–19235.

KEARSE, K.P. AND HART, G.W., 1991. Lymphocyte activation induces rapid changes in nuclear and cytoplasmic glycoproteins. *Proc. natl. Acad. Sci*., **88**: 1701-1705.

KUPPUSWAMY, D., DALTON, M. AND PIKE, L.J., 1993. Serine 1002 is a site of *in vivo* and *in vitro* phosphorylation of the epidermal growth factor receptor. *J. biol. Chem,* **268**: 19134-19242.

LA, D., SUTCH, B. AND LIVESAY, D.R. 2005. Predicting protein functional sites with phylogenetic motifs. *Proteins***, 58**: 309–320.

LUETTEKE N.C., PHILLIPS, H.K., QIU, T.H., COPELAND, N.G., EARP, H.S., JENKINS, N.A. AND LEE D.C., 1994.The mouse waved-2 phenotype results from a point mutation in the EGF receptor tyrosine kinase. *Genes Devel*, **8**:399-413.

MORRISON, P., TAKISHIMA, K. AND ROSNER, M. R.1993. Role of threonine residues in regulation of epidermal growth factor receptor by protein kinase C and mitogen-activated protein kinase. *J. biol. Chem,* **268**: 15536-15543.

NIELSEN, H., BRUNAK, S. AND VONHEIJNE, G.,1999. Machine learning approach for prediction of signal peptide and other protein signals. *Protein Eng.,* **12**: 33-39.

NORTHWOOD, I. C., GONZALEZ, F. A., WARTMANN, M., RADEN, D. L. AND DAVIS, R.J., 1991. Isolation and characterization of two growth factor-stimulated protein kinases that phosphorylate the epidermal growth factor receptor at threonine 669. *J. biol. Chem.,* **266**: 15266-15276.

OKSVOLD, M.P., HUITFELDT, H.S., OSTVOLD, A.C. AND SKARPEN, E., 2002. UV induces tyrosine kinase-independent internalization and endosome arrest of th EGF receptor. *J. Cell Sci*., **115**: 793-803.

PARIA, B.C., DAS, S.K., ANDREWS, G.K. AND DEY, S.K., 1993. Expression of the epidermal growth factor receptor gene is regulated in mouse blastocysts during delayed implantation. *Proc. natl. Acad. Sci. U.S.A*, **90**: 55-59.

SCHLESSINGER, J., 2000. Cell signaling by receptor tyrosine kinases. *Cell,* **103:** 211–225.

SIEW, N. AND FISCHER, D., 2001. Convergent evolution of protein structure prediction and computer chess tournaments: CASP, Kasparov, and CAFASP. *IBM Syst. J.,* **40**: 410-450.

TOMIC, S., GREISER, U.**,** LAMMERS, R., KHARITONENKOV, A., IMYANITOV, E., ULLRICH, A. AND BÖHMER, F.D., 1995. Association of SH2 domain protein tyrosine phosphatases with the epidermal growth factor receptor in human tumor cells. phosphatidic acid activates receptor dephosphorylation by ptp1c. *J. biol. Chem.*, **270**: 21277-21284.

WALTON, G.M., CHEN, W.S., ROSENFELD, M.G. AND GILL, G.N., 1990. Analysis of deletions of the carboxyl terminus of the epidermal growth factor receptor reveals self-phosphorylation at tyrosine 992 and enhanced *in vivo* tyrosine phosphorylation of cell substrates. *J. boil. Chem*., **265**: 1750-1754.

WILBUR, W. J. AND LIPMAN, D.J., 1983. Rapid similarity searches of nucleic acid and protein data banks. *Proc. Natl. Acad. Sci. USA*, **80**: 726-730.

ZACHARA, N.E. AND HART, G.W., 2002.The emerging significance of O-GlcNAc in cellular regulation. *Chem. Rev*, **102**: 431-438.

(Received 11 January 2006, revised 13 April 2006)