Journal of Cellular Biochemistry 100:1558–1572 (2007)

In Silico Determination of Intracellular Glycosylation and Phosphorylation Sites in Human Selectins: Implications for Biological Function

Ishtiaq Ahmad,[1] Daniel C. Hoessli,[2] Ramneek Gupta,[3] Evelyne Walker-Nasir,[1] Saleem M. Rafik,[1] M. Iqbal Choudhary,[4] Abdul Rauf Shakoori,[5] and Nasir-ud-Din[1,4]*

[1]Institute of Molecular Sciences and Bioinformatics, Lahore, Pakistan
[2]Department of Pathology and Immunology, Centre Médical Universitaire, Geneva, Switzerland
[3]Center for Biological Sequence Analysis Bl-208, Technical University of Denmark, DK-2800 Lyngby, Denmark
[4]HEJ Research Institute of Chemistry, University of Karachi, Karachi, Pakistan
[5]School of Biological Sciences, University of the Punjab, New Campus, Lahore 54590, Pakistan

Abstract Post-translational modifications provide the proteins with the possibility to perform functions in addition to those determined by their primary sequence. However, analysis of multifunctional protein structures in the environment of cells and body fluids is made especially difficult by the presence of other interacting proteins. Bioinformatics tools are therefore helpful to predict protein multifunctionality through the identification of serine and threonine residues wherein the hydroxyl group is likely to become modified by phosphorylation or glycosylation. Moreover, serines and threonines where both modifications are likely to occur can also be predicted (YinYang sites), to suggest further functional versatility. Structural modifications of hydroxyl groups of P-, E-, and L-selectins have been predicted and possible functions resulting from such modifications are proposed. Functional changes of the three selectins are based on the assumption that transitory and reversible protein modifications by phosphate and *O*-GlcNAc cause specific conformational changes and generate binding sites for other proteins. The computer-assisted prediction of glycosylation and phosphorylation sites in selectins should be helpful to assess the contribution of dynamic protein modifications in selectin-mediated inflammatory responses and cell–cell adhesion processes that are difficult to determine experimentally. J. Cell. Biochem. 100: 1558–1572, 2007. © 2007 Wiley-Liss, Inc.

Key words: post-translational modifications; glycosylation; phosphorylation; YinYang sites; multifunctional proteins

Multifunctional proteins are involved in diverse and often unrelated activities, and may perform different tasks in different biological environments [Jeffery, 1999]. Many proteins are physiologically multifunctional, both as extracellular components of body fluids and as cellular components. Selectins, a family of sugar-binding adhesion proteins, are multifunctional molecules involved in a variety of cellular interactions occurring in inflammatory responses (Fig. 1). Reversible phosphorylations induce changes in secondary and tertiary structures and control the modular interactions of proteins [Pawson and Gish, 1992]. In addition, phosphorylation and alternative glycosylation of the same hydroxyl group also control the functional behaviour of a growing list of proteins [Cheng and Hart, 2001]. Phosphorylation of serine/threonine or tyrosine residues has different effects resulting either in promotion of proteolysis [Elorza et al., 2003] or resistance to proteolysis, as in the case of carboxyl or sulfate groups in glycoproteins [Nasir-ud-Din et al., 2003].

Protein functional information can be deduced from their 3-D structures [Bork et al., 1998; Attwood, 2000]. However, determination of the 3-D structure in vivo is difficult, as a given configuration is constantly modified by

*Correspondence to: Nasir-ud-Din, Institute of Molecular Sciences and Bioinformatics, 28 Nisbet Road, Lahore, Pakistan.
E-mail: nasir@super.net.pk or prof_nasir@yahoo.com
Received 20 July 2006; Accepted 25 August 2006
DOI 10.1002/jcb.21156
© 2007 Wiley-Liss, Inc.

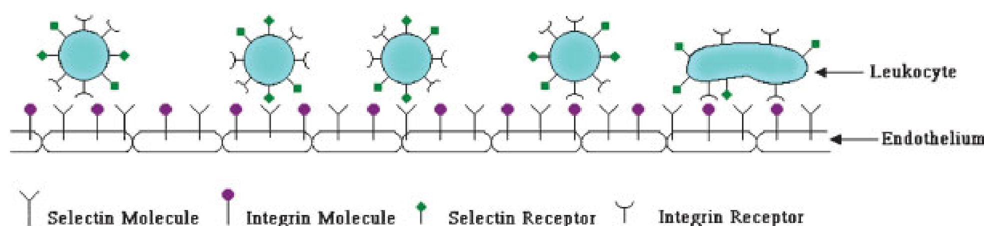


Fig. 1. Inflammatory cascade in which leukocyte rolls on endothelium mediated by selectins and adheres firmly by integrins through their respective receptors and finally transmigrates to the site of inflammation.

intra- and intermolecular interactions occurring between proteins in body fluids or in the cell. Most of the structural data available in protein databases have been determined by X-ray crystallography and NMR, but these structural determination methods provide information that is only partially relevant to the dynamic behavior of proteins in vivo.

The *O*-GlcNAc and phosphate modifications of proteins are temporary and reversible [Roquemore et al., 1996; Shafi et al., 2000; Kamemura et al., 2002] and allow proteins to sequentially perform multiple functions. However, the conformational changes caused by these modifications are still elusive and their realistic assessment through experimentation remains improbable for the time being. Theoretical assessment of modifications at hydroxyl-, amino-, or carboxy-amino acids, particularly at the serine and threonine residues would be very informative, as these modifications are known to result in functional switches in multifunctional proteins [Kamemura et al., 2002]. A list of proteins whose stability is affected by phosphorylation or *O*-β-GlcNAc glycosylation is shown in Table I.

It is pertinent to define the structure–function relationship of a protein or a modified protein after experimental verification of its three dimensional structure. However, knowing that the majority of blood proteins are multifunctional and a large number of the modifications causing functional switches are only transitory, relating structure (primary sequence) and function should be considered nearly impossible at the present time.

Computational biology and computational chemistry have considerably helped understanding how post-translational protein modifications such as phosphorylation, sulfation, and glycosylation regulate biological processes. Several programs have been developed to predict different structural and functional aspects of protein post-translational modifications with reliable accuracy [Nielsen et al., 1999]. Protein *O*-GlcNAc glycosylation of MSP-1 [Hoessli et al., 2003] and phosphorylation have been predicted and experimentally verified for the malarial vaccine candidate protein MSP-1. Prediction of modifications and their experimental verification is only possible where the half life of the protein and its availability allows such analysis. The computer-assisted methods are essential in cases where conditions of limited availability and instability of proteins prevail and transitory proteins compete for modifications to perform diverse functions.

Prediction programs to assess the potential of a protein for substituting hydroxyl or amino groups with phosphate or *O*-GlcNAc would provide significant incentives to experimentally

TABLE I. Stability/Turnover of Different Proteins Controlled by PTMs

Protein	PTM/s	Function	Stability/turnover
c-Myc	*O*-GlcNAc Phosphate	Central regulator for cell growth and proliferation	Affected by phosphorylation especially at Thr 58
Ap1	*O*-GlcNAc Phosphate	Transcription factor	Affected by phosphorylation
RNA polymerase II	*O*-GlcNAc Phosphate	Enzyme catalyzing the mRNA synthesis in nucleoplasm	Affected by phosphorylation
BimEL	Phosphate	Apoptotic cascade activating protein	Affected by phosphorylation
Rap1GAP	Phosphate	Activator of GTP hydrolytic activity of Rap1	Affected by phosphorylation
Sp1	*O*-GlcNAc	Transcription factor	Affected by *O*-GlcNAc modification
Pyruvate kinase	*O*-GlcNAc Phosphate	Glycolytic enzyme	Affected by phosphorylation

verify protein multifunctionality. Conserved residues and motifs are known [La et al., 2005; Nikolaidis et al., 2005] to be involved in performing essential functions. The identification of conserved Ser/Thr residues with marginally positive or negative modification potential (false positive and false negative) is mandatory to define real modification potential [Ahmad et al., 2006a,b].

C-type mammalian lectins, or selectins, are cell–cell adhesion molecules that play key roles in transendothelial migration of hematopoietic inflammatory cells [Vestweber and Blanks, 1999] (Fig. 1). These molecules mediate the initial steps of leukocyte adhesion to the vascular endothelium at the site of inflammation [Vestweber and Blanks, 1999]. Amongst the selectin family [Bevilacqua and Nelson, 1993], L-selectins act as primary cell adhesion molecules in lymphocyte recirculation [Gallatin et al., 1983] and in leukocyte recruitment to the site of inflammation [Rosen and Bertozzi, 1994], whereas E- and P-selectins are expressed on endothelial cells and platelets. Selectin has an *N*-terminal domain that is critical for adhesion [Tedder et al., 1995] and contains potential sites for *N*- and *O*-glycosylation.

The reason for choosing selectins as a model for this study was precisely their multifunctionality in adhesion and their recent designation as targets for drug delivery and anti-inflammatory drug development as well as in anti-cancer therapy [Ehrhardt et al., 2004]. Utilizing computer-assisted studies, we demonstrate that selectins have potential for both *O*-β-GlcNAc modification and phosphorylation. Phosphorylation is an important reversible protein modification involved in transmembrane signaling, growth control, cell cycle progression, membrane trafficking, platelet activation, and protein turnover [Corvello et al., 1993; Fujimoto and McEver, 1993; Smeets et al., 1993; Tonks and Neel, 1996; Yoshida et al., 1998] (Table I). Rapid selectin phosphorylation (15–30 s) and selective dephosphorylation after platelet activation is quite likely to be involved in signal transduction [Corvello et al., 1993]. A similarity between phosphorylation and glycosylation and dynamic interplay between *O*-glycosylation and *O*-phosphorylation have been described [Roquemore et al., 1996; Shafi et al., 2000; Kamemura et al., 2002]. The potential modification sites, *O*-glycosylation and *O*-phosphorylation, have been found on conserved residues. This prompted us to investigate the role of transitory modifications by *O*-GlcNAc and phosphate in the cytoplasmic domains of transmembrane selectins, similar to what has been described for cytoplasmic and nucleoplasmic proteins [Whelan and Hart, 2003].

In this study, we have investigated the primary amino acid sequence of human selectins for the presence of signals for phosphorylation, glycosylation and alternative phosphorylation/glycosylation in phylogenetically conserved sites that are functionally relevant. False negatively predicted Yin Yang sites have also been identified. The computer-assisted modifiable sites identified in this study suggest functional diversity generated by substitutions.

MATERIALS AND METHODS

The Sequence Data

The sequence data of all three selectins of *Homo sapiens* used for predicting phosphorylation and glycosylation sites was taken from Entrez protein database, the NCBI sequence viewer. The L-selectin sequence with Entrez database No A34015 and SWISSPROT entry name LEM1_HUMAN [Ord et al., 1990], the E-selectin sequence with Entrez database No. A35046 and SWISSPROT entry name LEM2_HUMAN [Collins et al., 1991], and the sequence of P-selectin with Entrez database No. P16109 and SWISSPROT entry name LEM3_HUMAN [Johnston et al., 1989] were retrieved. All three sequences of human selectins were precursor sequences having the signal peptide at C-terminus (L-selectin 1–28, E-selectin 1–21, and P-selectin 1–41). The same sequences with signal peptide regions were used for further analysis just to retain uniformity in results for the analysis and prediction. The numbers of residues cited in this article include the signal peptide, thus for finding the exact position of any amino acid in sequence of each functional mature selectin molecule, the signal peptide region may be removed. Blast search was made against all three selectins using NCBI-protein program [Altschul et al., 1997]. Default parameters were used for the search. Total 1,691 hits for L-selectin, 2,711 hits for E-selectin, and 3205 hits for P-selectins were obtained. From these

hits, 10 for L-selectin (human, chimpanzee, orangutan, rhesus monkey, baboon, rabbit, dog, bovine, rat, and mouse), 10 for E-selectin (human, horse, dog, mouse, rat, rabbit, sheep, mule deer, bovine, and pig), and 9 for P-selectin (human, horse, dog, mouse, rat, sheep, bovine, pig, and rabbit) were selected from their respective blast search results to find out conserved regions in cytoplasmic part of different mammalian species (Table II). Conserved sites were identified using CLUSTALW [Thompson et al., 1994].

Glycosylation Prediction Methods

The methods used for defining potential glycosylation sites involved prediction of both *O*- and *N*-linked sites. The three methods for predicting *O*-linked glycosylation sites include *NetOGlyc* 2.0 and 3.1 (http://www.cbs.dtu.dk/services/NetOGlyc/) [Hansen et al., 1998; Julenius et al., 2005], that predicts *O*-glycosylation sites in mucin-type proteins (i.e., for *O*-GalNAc sites), *DictyOGlyc* 1.1 [Gupta et al., 1999] (http://www.cbs.dtu.dk/services/DictyOGlyc/) that predicts *O*-α-GlcNAc sites in eukaryotic proteins and YinOYang 1.2 (http://www.cbs.dtu.dk/services/YinOYang/) (unpublished), that predicts *O*-β-GlcNAc sites in eukaryotic proteins. *NetNGlyc* 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc/) (unpublished) was used for predicting *N*-glycosylation sites. The above-mentioned four methods for predicting glycosylation sites are neural network based.

Phosphorylation Prediction Method

For prediction of phosphorylation sites in selectins, *NetPhos* 2.0 [Blom et al., 1999] (http://www.cbs.dtu.dk/services/NetPhos/) was utilized. *NetPhos* 2.0 is also a neural network-based program designed by training the neural networks through protein phosphorylation data from phosphoBase 2.0. NetPhosK 1.0 (http://www.cbs.dtu.dk/services/NetPhosK/) [Blom et al., 2004] was also used to find the potential of three selectin cytoplasmic domains for their specificity to be phosphorylated by different kinases at Ser/Thr/Tyr.

Prediction methods described above are neural network-based and are designed by

TABLE II. Sequences Selected From Blast Search

S. no.	Species	Database	Accession no.	Blast search results		
				E-value	Bits score	% Identity
L-selectin						
1	*Homo sapiens*	Swissprot	P14151	0.00	764	100
2	*Pan troglodytes*	RefSeq	NP_001009074.1	0.00	762	99
3	*Pongo pygmaeus*	Swissprot	Q95235	0.00	755	98
4	*Macaca mulata*	Swissprot	Q95198	0.00	730	94
5	*Papio hamadryas*	Swissprot	Q28768	0.00	728	94
6	*Oryctalagus cuniculus*	GenBank	AAA67896.1	0.00	656	85
7	*Bos taurus*	RefSeq	NP_776607.1	6e-178	625	78
8	*Canis familaris*	RefSeq	XP_537201.2	0.00	643	81
9	*Rattus norvegicus*	Swissprot	P30836	e-168	594	77
10	*Mus musculus*	RefSeq	NP_035476.1	2e-168	594	76
E-Selectin						
1	*Homo sapiens*	Swissprot	P16581	0.00	1275	100
2	*Equus caballus*	GenBank	AAK48712	0.00	1056	82
3	*Canis familaris*	RefSeq	NP_001003310	0.00	993	76
4	*Mus musculus*	Swissprot	Q00690	0.00	944	73
5	*Rattus norvegicus*	Swissprot	P98105	0.00	758	70
6	*Oryctalagus cuniculus*	Swissprot	P27113	0.00	728	65
7	*Ovis aries*	GenBank	AAK48711	0.00	633	67
8	*Odocoileus hemionus*	GenBank	AAK48710	2e-180	632	67
9	*Bos taurus*	Swissprot	P98107	8e-178	624	66
10	*Sus scrofa*	Swissprot	P98110	e-156	553	59
P-Selectin						
1	*Homo sapiens*	Swissprot	P16109	0.00	1802	100
2	*Equus caballus*	GenBank	AAS79432	0.00	1451	79
3	*Canis familaris*	RefSeq	XP_537202.2	0.00	1163	65
4	*Mus musculus*	Swissprot	Q01102	0.00	1134	63
5	*Rattus norvegicus*	Swissprot	P98106	0.00	1130	64
6	*Ovis aries*	Swissprot	P98109	0.00	1130	64
7	*Bos taurus*	Swissprot	P42201	0.00	770	60
8	*Sus scrofa*	RefSeq	NP_999243.1	0.00	761	59
9	*Oryctalagus cuniculus*	GenBank	AAA81385.1	0.00	734	59

memorizing the known sequence environment data of glycosylated/phosphorylated serine/threonine and non-glycosylated/non-phosphorylated serine/threonine. A jury of networks is used to evaluate the performance of neural networks. The results obtained from all the networks are sigmoidally arranged and averaged to obtain a value between zero and one by these prediction methods. Usually a threshold of 0.5 is used for prediction, which means that a site with an output of more than 0.5 is assigned as having a potential to be glycosylated or phosphorylated. These methods use back propagation algorithm to learn the examples and adjusting saliency of weights [Rumelhart et al., 1996]. Cross validation data has also been used by these methods to check the performance and sensitivity of the neural networks using Matthews correlation coefficients [Matthews, 1975]. Among them, the YinOYang 1.2 method (*unpublished*) is efficient in a cross validation test as it correctly identifies 72.5% of the glycosylated sites and 79.5% of the non-glycosylated sites in the test set, verifying the Matthews correlation coefficient of 0.22 on the original data, and 0.84 on the augmented data set. The method has the capability to predict the Yin Yang sites that can be glycosylated and alternatively phosphorylated. NetPhos 2.0 predicts phosphorylation on the OH- function of serine, threonine or tyrosine residues with a sensitivity range of 69–96% [Blom et al., 1999].

RESULTS

The prediction results obtained for *O*-β-GlcNAc sites by *YinOYang* 1.2 showed that potential *O*-β-GlcNAc sites were the most frequent among all three types of *O*-glycosylation (Fig. 2a, Table III). L-selectin was predicted to have altogether four potential sites for *O*-β-GlcNAc modifications, at Ser 177, 286, 319, and Thr 170, and no YinYang site. E-selectin was predicted to have a total of 12 potential sites for *O*-β-GlcNAc modifications, at Ser 4, 19, 147, 227, 288, 314, 477, 592, 607 and at Thr 140, 146, 303, with one YinYang site at Ser 227; whereas P-selectin was predicted to have a total of 13 potential sites for *O*-β-GlcNAc modifications, at Ser 87, 287, 371, 462, 529, 530, 619, 829 and at Thr 48, 511, 618, 628, 696, with two YinYang sites, at Ser 287 and 619 (Fig. 2a). Utilizing *NetPhos* 2.0, *O*-linked phosphorylation at serine, threonine, and tyrosine in the different selectins were predicted (Fig. 2b, Table III), showing that in L-selectin, there were nine Ser (at 106, 226, 249, 263, 313, 322, 324, 364, and 367); six Thr (at 10, 99, 104, 239, 240, and 301), and five Tyr (at 61, 75, 132, 331, and 372). The E-selectin prediction results showed that 20 Ser (at 37, 131, 149, 185, 207, 209, 218, 226, 227, 318, 320, 383, 387, 392, 397, 407, 446, 460, 505, and 596), 8 Thr (at 26, 32, 86, 266, 298, 326, 452, and 514), and 6 Tyr (at 58, 69, 377, 453, 603, and 608) had potential for phosphorylation. Potential phosphorylation sites predicted in P-selectin included 17 Ser at 52, 227, 278, 287, 291, 385, 472, 531, 534, 557, 582, 586, 619, 623, 685, 689, and 813, 10 Thr at 106, 251, 334, 522, 561, 568, 647, 746, 756, and 767, and 10 Tyr at 64, 78, 135, 196, 234, 358, 408, 532, 544, and 726. The percentage of phosphorylatable Ser is highest of the three amino acids for all three selectins.

The NetPhosK 1.0 prediction results showed that the cytoplasmic domains of three selectins had utilization potential for different kinases at different Ser/Thr and Tyr residues. The predictions results suggested that the cytoplasmic domain of L-selectin have phosphorylation potential by PKC (at Ser 364), PKA (at 364 and 367), PKG (at Ser 367), and INSR (at Tyr 372) acceptor substrate sites. The potential sites for different kinases activity predicted in cytoplasmic domain of E-selectin were all at Ser on positions 596 by PKC, 599 by CKI, 602 by DNAPK, PKC, and cdc2, and 607 by PKA. P-selectin cytoplasmic domain contained only two potential acceptor substrate sites for different kinases, which include Ser 829 by GSK3 and cdk5; and Thr 817 by PKC.

Besides these predicted, false negative Yin Yang sites in cytoplasmic domains of three selectins were identified. In L-selectins, the cytoplasmic Ser 364, and Ser 367 showed positive potential for phosphorylation but negative (but close to threshold) potential for *O*-GlcNAc, and both of these residues are conserved in most mammalian species (Fig. 3a). If a kinase can access these residues, OGT is also likely to access them [Kamemura et al., 2002]. Thus Ser 364 and 367 seem to be false negative Yin Yang sites, as both phosphorylation and *O*-β-GlcNAc addition could occur at these sites. Similarly for E-selectin, cytoplasmic Ser 592 and Ser 596 appear to be false negative Yin Yang site (Figs. 2a and 3b), and for P-selectin, Tyr 818 is conserved in mammals

(Fig. 3c) is false negative prediction site for phosphorylation (Fig. 2b). However, Ser 829, which is almost conserved in mammals (Fig. 3c), is a positive prediction site for *O*-GlcNAc modification but negative for phosphorylation and should be considered a false negative Yin Yang site.

DISCUSSION

Definition of the functional structures of proteins in biological contexts has to take into account the simultaneous presence of a high number of interacting proteins. For instance in human blood, nearly 1,444 proteins interact [Chang et al., 2004] and definition of an individual protein structure is not feasible with the tools available. Under such circumstances, one could begin to approach the multifunctional nature of proteins by establishing the potential for transitory conformational changes. This could be achieved by mapping the functional groups and sequence signals likely to become post-translationally modified. For instance, the anionic group modification may change the protein conformation [Berlot et al., 2002] and so contribute to the formation of active sites capable of specific interactions [Zaia et al., 2001]. Similarly, the negative charge of phosphate groups [Varki and Kornfeld, 1980] may alter the balance of non-covalent interactions

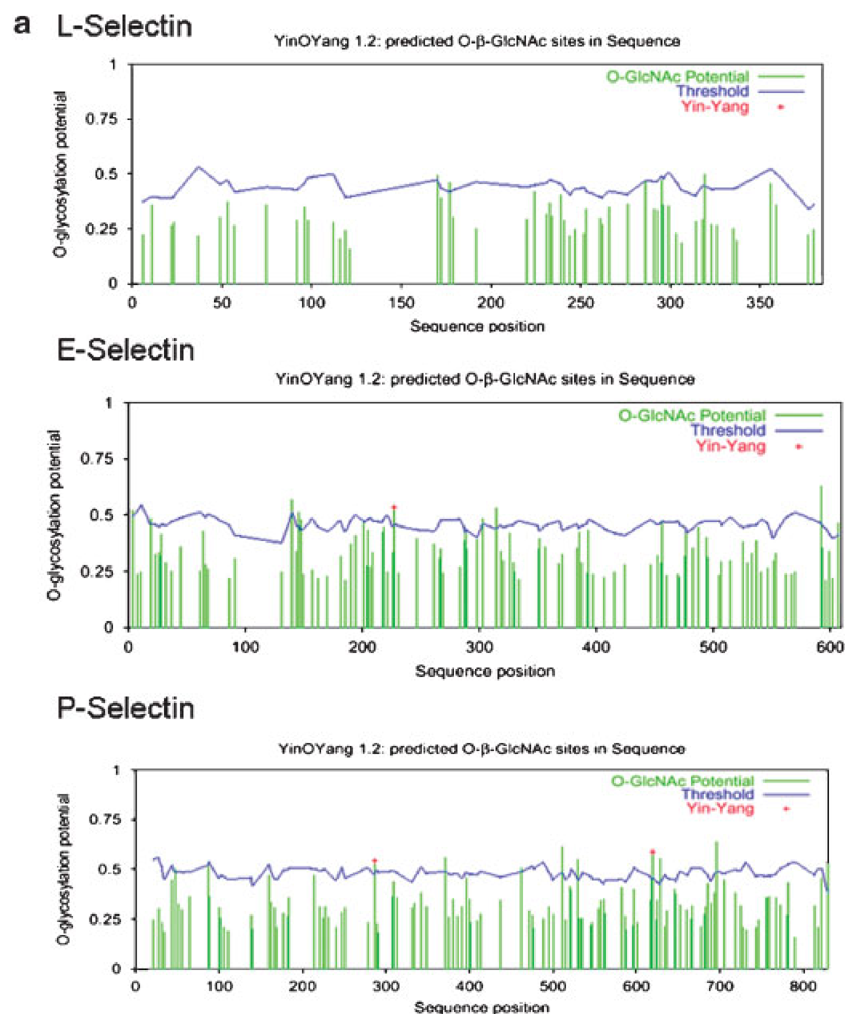
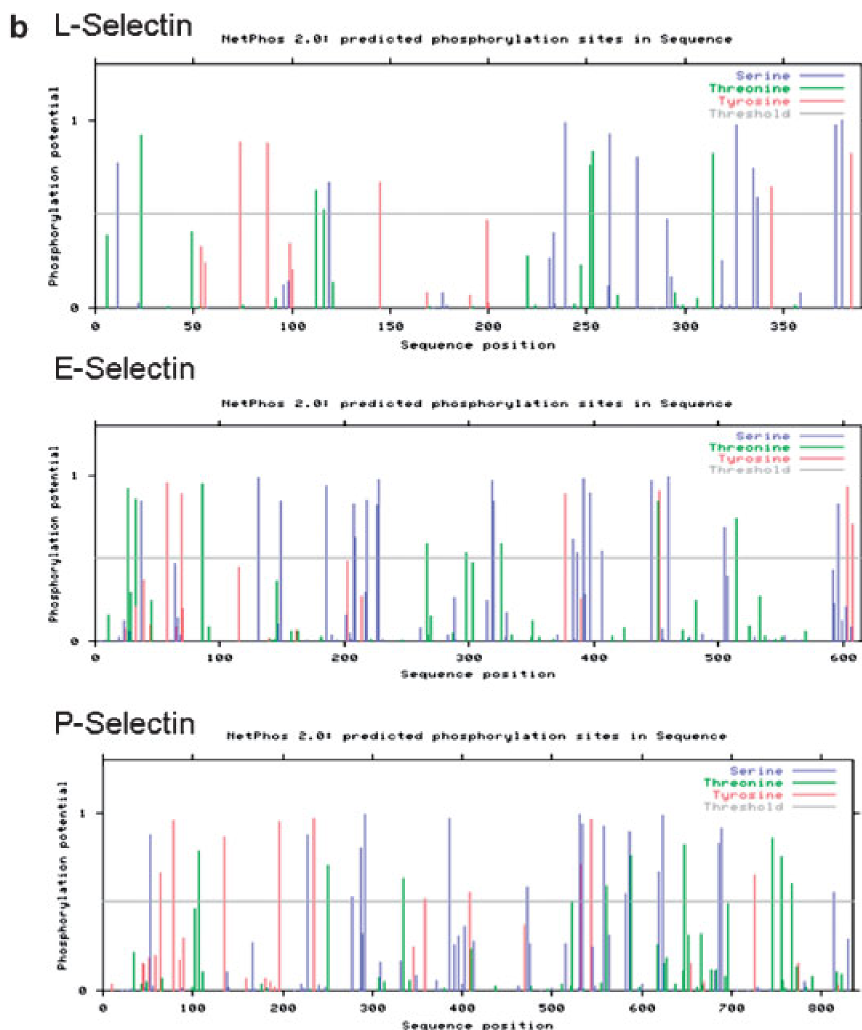


Fig. 2. The potential for *O*-β-GlcNAc modification on Ser and Thr residues and phosphorylation at Ser, Thr, and Tyr residues taking sequence position horizontally and the potential vertically. **a**: Vertical green lines show the potential of each Ser and Thr. The blue wavy line shows the threshold potential for *O*-β-GlcNAc addition. The lines crossing, touching, or very close to the threshold designate residues with potential to be modified. Also shown are the Yin Yang sites by red + mark. **b**: Vertical green lines represent Thr, blue represent Ser, and red lines represent Tyr. The gray horizontal line shows the fixed threshold (0.5) for phosphorylation.

Fig. 2. (*Continued*)



modify the secondary, tertiary, or even quaternary structures. Such alterations may induce conformational changes in a protein and modify its biological function or multimeric state [Yuan et al., 2003]. Thus, the predictions based on amino acid sequence signals should become useful to define modifications and guide further exploration of protein function.

The dynamic and regulatory *O*-GlcNAc modification was found to occur at sites where phosphorylation also occurred [Kelly et al., 1993; Felin et al., 1994; Chou et al., 1995; Medina et al., 1998; Cheng and Hart, 2001]. An interplay between glycosylation and phosphorylation was noted in proteins such as RNA polymerase II [Kelly et al., 1993], estrogen receptor-β [Cheng and Hart, 2001], SV-40 large T-antigen [Medina et al., 1998], and c-Myc proto-oncogene product [Chou et al., 1995], which led to the formulation of the Yin-Yang hypothesis. According to this hypothesis, *O*-GlcNAc modifications and phosphorylations compete for the same site or region on a protein backbone. For instance Thr58 in c-Myc is alternatively glycosylated (by *O*-GlcNAc) and phosphorylated, as evidenced by the use of Thr-58 site-specific antibodies [Kamemura et al., 2002]. It was further observed by mutagenic replacement of Ser62 with Ala that phosphorylation of Thr58 is dependent upon prior phosphorylation of Ser62.

The role of *O*-GlcNAc modification of nuclear and cytoplasmic proteins exhibiting lectin-like activity [Felin et al., 1994] remains to be determined. Further to the specific roles of *O*-GlcNAc modification in signaling, nuclear transport, regulation of protein multimer formation, modulation of the half-life, and

TABLE III. Prediction Results of Glycosylation and Phosphorylation in Human L-, E-, and P-Selectins

	L-selectin	E-selectin	P-selectin
Modification type and residues information			
Total number of amino acids	372	610	830
Total number of Ser residues and (percentage)	30 (8.06%)	66 (10.81%)	71 (8.55%)
Total number of Thr residues and (percentage)	23 (6.18%)	40 (6.557%)	55 (6.62%)
Total number of Tyr residues and (percentage)	13 (3.49%)	18 (2.95%)	29 (3.49%)
Total number of Asn residues and (percentage)	21 (5.64%)	30 (4.91%)	47 (5.66%)
Number and percentage of *O*-α-GalNAc sites			
Number and percentage of potential Ser residues	1 (3.33%)	2 (3.03%)	4 (5.63%)
Number and percentage of potential Thr residues	0 (0.00%)	1 (1.51%)	3 (4.22%)
Total modification sites and percentage	1 (1.88%)	3 (2.83%)	7 (5.55%)
Number and percentage of *O*-α-GlcNAc sites			
Number and percentage of potential Ser residues	2 (6.66%)	01, 1.51%	01, 1.41%
Number and percentage of potential Thr residues	0 (0.00%)	00, 0.00%	00, 0.00%
Total modification sites and percentage	2 (3.77%)	01, 0.94%	01, 0.79%
Number and percentage of *O*-β-GlcNAc sites			
Number and percentage of potential Ser residues	03, 10.00%	08, 12.12%	08, 11.26%
Number and percentage of potential Thr residues	01, 03.33%	03, 04.54%	05, 07.04%
Total modification sites and percentage	04, 7.55%	11, 10.38%	13, 10.32%
Number and percentage of *N*-glycosylation sites			
Number of potential Asn residues within sequon N-Xaa-S/T and percentage	06, 28.57%	11, 36.66%	12, 25.53%
Number and percentage of other Asn with +ve score that don't lie within sequon N-Xaa-S/T	10, 47.62%	11, 36.66%	25, 53.19%
Number and percentage of Asn with −ve score	05, 23.81%	08, 26.66%	10, 21.28%
Number and percentage of phosphorylation sites			
Number and percentage of potential Ser residues	09, 30.00%	20, 30.30%	17, 23.94%
Number and percentage of potential Thr residues	06, 26.08%	08, 20.00%	10, 18.18%
Number and percentage of potential Tyr residues	05, 38.46%	06, 33.33%	10, 34.48%

regulation of protein activity [Kelly et al., 1993; Felin et al., 1994; Chou et al., 1995], other possible roles of *O*-GlcNAc modification have been proposed [O'Donnell, 2002]. These include blocking of phosphorylation by occupying serine and threonine residues at the same position or region, and regulating protein turnover by blocking phosphorylation-dependent proteolysis and protein–protein interactions.

Different domains (C-type lectin, EGF-like, Sushi, and cytoplasmic domains) of all three selectins share different degrees of sequence homology. The sequence homology is higher in the C-type lectin domain than in the cytoplasmic domain (Table IV). Both, pair-wise and multiple alignments of three selectins reveal the same trend in different domains. The cytoplasmic domains of the three selectins share little homology, although they are related structurally and phylogenetically [Spertini et al., 1991; Green et al., 1994].

From studies on calmodulin binding, a Ca^{2+}-dependent process [Matala et al., 2001], and cleavage of L-selectin, it was inferred that changes in intracellular Ca^{2+} may be responsible for the dissociation of calmodulin from L-selectin before proteolysis [Matala et al., 2001]. Cleavage of L-selectin occurs within a membrane-proximal region resulting in the shedding of an extra-cellular fragment and retention of a 6-kDa transmembrane fragment [Kahn et al., 1994]. The membrane-proximal cleavage site shows relaxed sequence specificity, but requires spatial and possibly conformational changes [Migaki et al., 1995]. The relaxed sequence specificity for the action of proteases has also been supported by prediction results of NetPhop *2.0* (http://www.cbs.dtu.dk/services/NetChop/) (prediction results not shown). In addition, the cytoplasmic domain of L-selectin is phosphorylated on tyrosine [Migaki et al., 1995] and serine [Haribabu et al., 1997] residues following cell activation. The important point is that a highly conserved serine residue in the NH_2-terminal region of the cytoplasmic domain of L-selectin is phosphorylated before shedding [Haribabu et al., 1997; Kilian et al., 2004]. It was also shown that the recognition site(s) for protein kinase C (PKC) is exclusively contained within the ectodomain [Fors et al., 2001] and that a metalloprotease is involved in the extracellular cleavage [Bennett et al., 1996], but the mechanism for L-selectin recognition by the surface metalloprotease is not clearly understood. It is well possible that intracellular signals might be responsible for the metalloprotease action on L-selectin, either by activating the enzyme or by inducing a conformational change in the extracellular domain of L-selectin that exposes the cleavage site [Werb and Yan,

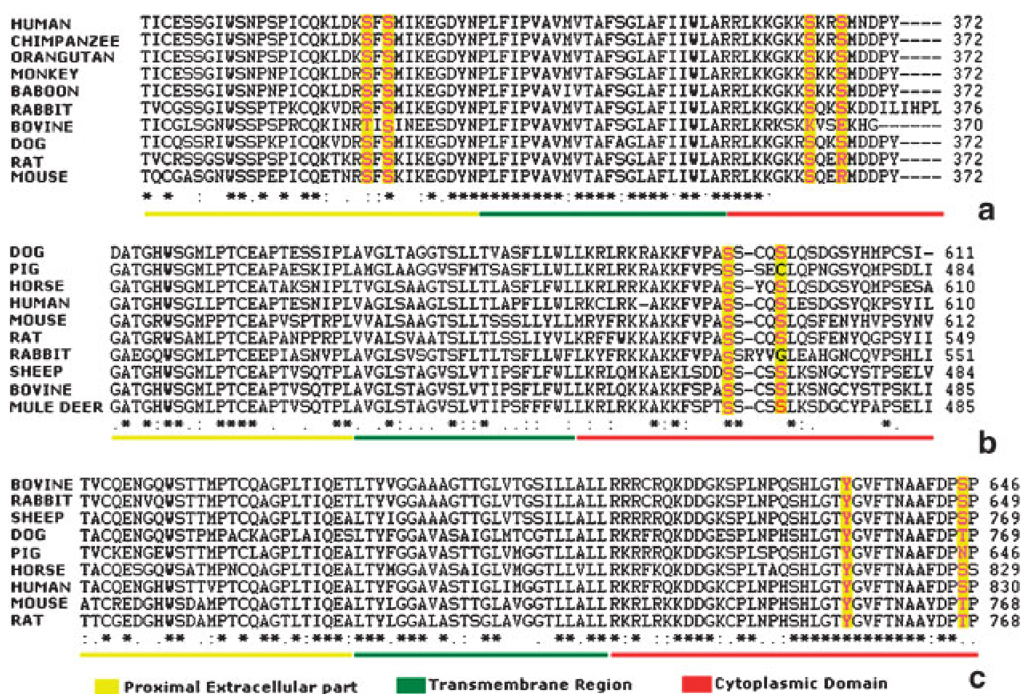


Fig. 3. Multiple alignment of L-selectin (**a**), E-selectin (**b**), and P-selectin (**c**) in different animal groups. Important conserved Ser/Thr residues with modification potential are highlighted in yellow.

1998]. Shedding of the extracellular part of L-selectin may result in both conformational changes in the residual extracellular part of L-selectin and activate a trans-membrane metalloprotease (Figs. 4a and 5b). It has been described that metalloprotease inhibitors downregulate L-selectin shedding [Feehan et al., 1996; Walcheck et al., 1996], whereas, phorbol esters induce collective shedding of a wide variety of cell-surface molecules including L-selectin. Moreover, phorbol myristate acetate (PMA) stimulates PKC [Kaldjian et al., 1988] (Fig. 5b,c) which in turn may activate metalloprotease through phosphorylation signaling to bind and cleave the extracellular part of transmembrane proteins such as L-selectin [Arribas et al., 1996; Mullberg et al., 1997].

On the basis of available data and according to our prediction results, we propose that phosphorylation at Ser364 or 367 (Fig. 4a) within the cytoplasmic domain may be responsible for metalloprotease binding following conformational changes in the membrane proximal, extracellular part of L-selectin. PMA activates PKC [Kaldjian et al., 1988] which in turn activates the metalloprotease (Fig. 5b,c). It is also expected that at the same time, PKC also triggers inside-out signaling through phosphorylation of the L-selectin cytoplasmic domain at Ser 364 which is highly conserved in different mammalian species (Fig. 3a) and bears modification potential as predicted by both NetPhos 2.0 and NetPhosK 1.0. Phosphorylation of cytoplasmic domain of L-selectin by PKC has

TABLE IV. Homology Among Different Domains of Three (L-, E-, and P) Selectins of *Homo Sapiens*

	% Identity				% Similarity			
Selectin Types	CTLD	EGF	Sushi D	Ct. D	CTLD	EGF	Sushi D	Ct. D
E & L	66%	66%	30–37%	N. S	75%	77%	44–51%	N. S
E & P	68%	63%	37–39%	N. S	75%	74%	55–58%	N. S
L & P	68%	64%	37–45%	N. S	77%	77%	48–62%	N. S
E, L & P	58%	53%	—	N. S	64%	58%	—	N. S

CTLD, C-type lectin domain; EGF, epidermal growth factor; Sushi D, Sushi domain; Ct. D, cytoplasmic domain; N.S, not significant.

also been evidenced [Kilian et al., 2004]. Similarly, inactivation and decreased membrane attachment of PKC and PKA by increased levels of *O*-GlcNAc incorporation and vice versa has been shown [Griffith and Schmitz, 1999; Matthews et al., 2005]. Thus, we propose that the false negative Yin Yang sites at Ser 364 or 367, when modified by *O*-β-GlcNAct may block or delay phosphorylation and consequently increase the half life of L-selectin. This complex interplay of phosphorylation and *O*-β-GlcNAc modification on Ser 364 and 367 seems to regulate the signaling events of L-selectin in leukocyte during inflammatory response. Inside-out signals may cause phosphorylation of Ser 322 and 324 and induce conformational changes that favor metalloprotease binding. In summary, we propose here that L-selectin shedding is regulated in both the cytoplasmic and extracellular domains through specific modifications, inducing conformational changes in the metalloprotease binding region of L-selectin (Figs. 4a and 5b).

Rapid endocytosis of E-selectin was reported to depend upon tyrosine-containing motifs [Chuang et al., 1997]. There are, however, endocytosis motifs that do not involve tyrosine, such as the di-leucine [Spertini et al., 1991] and di-lysine [Itin et al., 1995] sequences. Depending on the particular protein, phosphorylation of specific cytoplasmic serine residues, membrane-proximal to the di-leucine motifs, may or may not be required for efficient endocytosis [Spertini et al., 1991]. The signal for the internalization of E-selectin seems to be complex and might involve a novel route. Chuang et al. [1997] showed that Ser581, along with additional upstream residue(s), is necessary for internalization. They suggested that Ser581 may be phosphorylated and serve as a constitutive endocytosis signal. E-selectin is indeed phosphorylated at cytoplasmic serine but not at tyrosine residues [Smeets et al., 1993]. Specific cytoplasmic serine phosphorylation of other surface receptors was shown to be associated with ubiquitination of neighboring

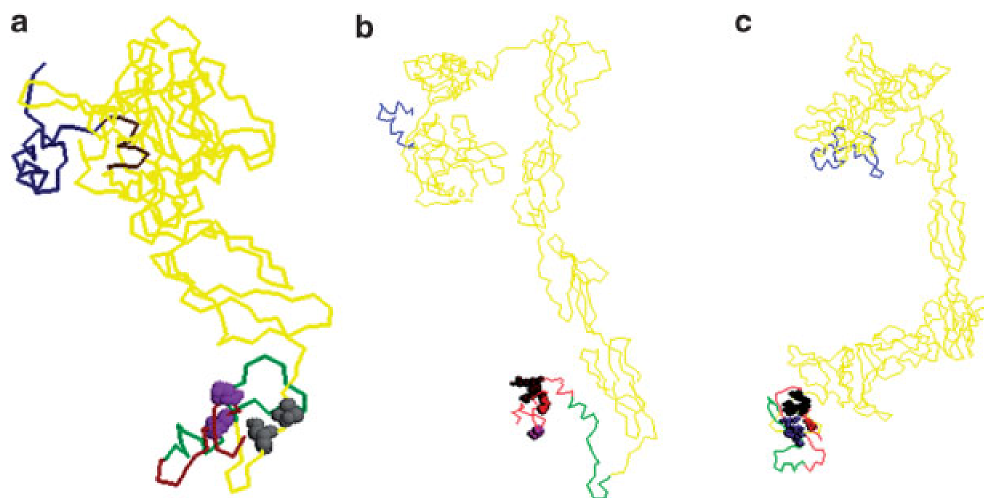


Fig. 4. Different domains of L-, E-, and P-selectins and the residues involved in regulating their function. The figure was prepared by generating theoretical models of three selectins (L-Selectin with 372, E-Selectin with 610, and P-selectin with 830 amino acids) with the Robetta Structure Prediction Server. RasMol was used to view and annotate the structural models. **a**: *L-selectin*. Cytoplasmic domain (356–372, C-terminal region) in red. Transmembrane domain (a.a. 333–355) in green. Extracellular domain in yellow (39–332), terminal signal peptide (1–28) in blue, and pro-peptide (29–38) in brown. Ser 364 and 367 in the cytoplasmic domain are shown in purple (space fill style). Phosphorylation of these residues may result in signals for conformational changes in the membrane proximal extracellular domain at Ser 322 and 324 (in gray). **b**: *E-Selectin*. Cytoplasmic domain (579–610, C-terminal region) in red. Transmembrane domain (557–578) in green. Extracellular domain in yellow (22–556), terminal signal peptide (1–21) in blue. Ser592 in red has potential for *O*-GlcNAc modification and may block phosphorylation at Ser596, in violet. The neighboring Lys586 and 587 in brown are considered to be ubiquitinated and to signal for endocytosis of E-Selectin. **c**: *P-Selectin*. Cytoplasmic domain (796–830, C-terminal region), in red. Transmembrane domain (772–795) in green. Extracellular part in yellow (a.a. 42–771), and terminal signal peptide (1–41) in blue. Positive signal sequence KCPL (Lys806-Leu809), in purple and negative signal sequences YGVF (Tyr818-Phe821) in gray and the DPSP (Asp827-Pro830) with Ser829 in red. The penultimate Ser829 in the DPSP sequence has been proposed as a site for interplay of GlcNAc and phosphate in regulating negative signals.

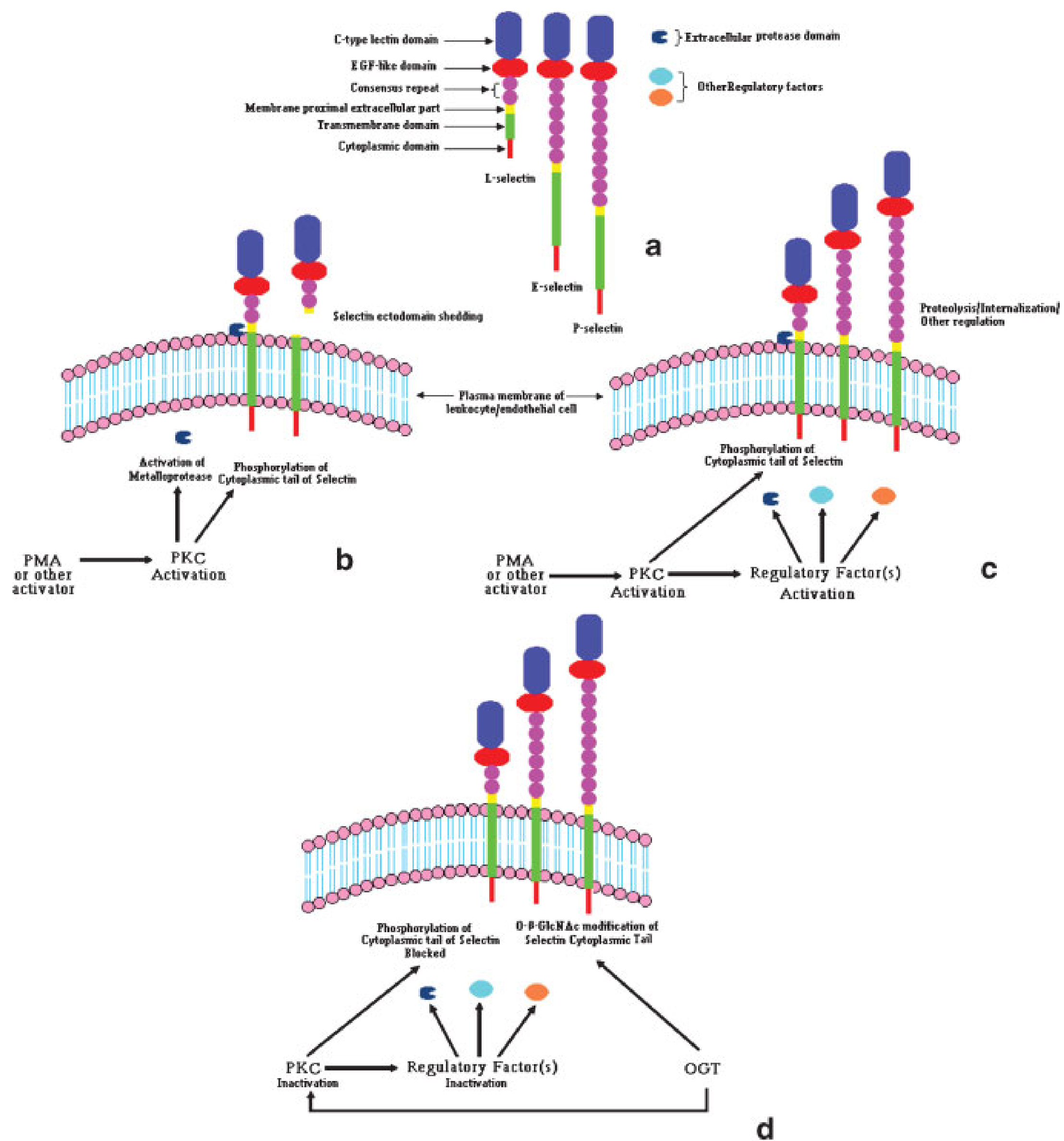


Fig. 5. The possible mechanism of PKC mediated phosphorylation of different selectins and other regulatory factors including transmembrane metalloproteases and role of *O*-GlcNAc modification in the cytoplasmic domain of three selectins and PKC (**a**). Domain structure of three selectins and their regulatory factors. **b**: Activated PKC phosphorylates directly the metalloprotease or cytoplasmic domain of three selectins resulting in binding of the metalloprotease to the membrane proximal extracellular part of L-selectin and shedding of extracellular part. **c**: Activated PKC phosphorylates cytoplasmic domains of selectins or other regulatory factors which in turn control selectin shedding and/or internalization. **d**: The activated OGT inactivates PKC by addition of *O*-GlcNAc, thus blocking all functions attributed to PKC. In addition, the OGT adds O-GlcNAc to the cytoplasmic domain of three selectins and occupies the PKC phosphorylation sites. PKC, protein kinase C; PMA, phorbol myristate acetate; OGT, *O*-GlcNAc transferase.

lysine residues that in turn signal for endocytosis [Kolling and Hollenberg, 1994].

We expect E-selectin expression to be down-regulated through a cytoplasmic serine phosphorylated signal resulting in the association of a protease with an extracellular cleavage similar to L-selectin (Figs. 4c and 5) or alternatively the activated PKC may activate other regulatory factors which are necessary for internalization of E-selectin. On the basis of our prediction results, we suggest that in the human E-selectin, Ser592 (potential for *O*-GlcNAc modification, see Fig. 2a) may block phosphorylation at Ser596 (predicted as potential phosphorylation site, see Fig. 2b). This may result in the delay of E-selectin proteolysis

(by preventing extracellular protease association) or endocytosis, prolong its surface expression, and consequently sustain the inflammatory reaction (Figs. 4b and 5c). The highly conserved Ser592 (Fig. 3b), shows potential for *O*-GlcNAc modification and the conserved Ser 596 (Fig. 3b), shows positive potential for phosphorylation with PKC kinase specificity. These Ser592 and 596 residues of E-selectin behave similarly to those of L-selectin (Ser364 and 367) as false negative Yin Yang site(s), thus providing important signals in regulating the stability and endocytosis of E-selectin (Fig. 5c).

By way of mutagenesis and inhibition studies, Fujimoto and McEver [1993] provided evidence for phosphorylation of serines and threonines in the cytoplasmic domain of P-selectin and suggested that phosphorylation is involved in membrane trafficking and other functions of this protein. Mutational studies suggested that the 35-residue cytoplasmic domain was involved in the internalization of P-selectin [Setiadi et al., 1995]. Lysosomal targeting was suggested by the existence of a balance of positive (KCPL sequence within the C-1 domain in the membrane proximal region) and negative (sequences YGVF, in gray and DPSP in red within the C-2 domain, cf. Fig. 3c) signals within the cytoplasmic tail of P-selectin [Blagoveshchenskaya et al., 1998]. It was then demonstrated that in the cytoplasmic tail, cysteine provides a major positive signal along with proline, whereas lysine and leucine play no role in lysosomal targeting of P-selectin. Both in lysosomal targeting and intracellular proteolysis, the results of mutational analysis within the C2 domain (by substitutions in the YGVF and DPSP sequences) suggested the existence of lysosomal avoidance signals. As shown in Figure 4c, a possible role for phosphorylated Tyr 818 within the YGVF sequence (gray) and for a *O*-β-GlcNAc-modified serine within the DPSP sequence (red) may be envisaged in generating lysosomal avoidance signals. However, our prediction results with NetPhos 2.0 rather predict that Tyr 818 within the YGVF sequence (Tyr818-Phe821) has low potential (i.e., lower than the threshold 0.5) for phosphorylation. But this residue is highly conserved in all mammalian species (Fig. 3c) and its phosphorylation was experimentally shown by Fujimoto and McEver [1993]. These authors suggested that phosphorylation in cytoplasmic domain of P-selectin plays an important role in lysosomal targeting of P-selectin. Therefore, it appears to be a false negative phosphorylation site. The same group showed that Ser829 is the main phosphorylation site in cytoplasmic domain of P-selectin [Fujimoto and McEver, 1993], but it was predicted negatively by NetPhos 2.0 and positively by NetPhosK1.0 as substrate for GSK3 and cdk5 kinases and YinOYang1.2 for GlcNAc (Fig. 2a). Consequently, the Ser829 within the DPSP sequence (Asp827-Pro830) is potentially a Yin Yang site. We suggest that interplay between phosphorylation and *O*-β-GlcNAc modification at the same Ser829 in the DPSP sequence is important in regulating inhibition lysosomal signals. Interestingly, this serine is also preceded by proline. Proline was shown to be preferred in close vicinity of Ser/Thr acting as an interplay site [16]. Similarly, Thr817 adjacent to Tyr818 (of YGVF sequence) is highly conserved in all mammalian species. This Thr817 showed positive potential by NetPhosK 1.0 for PKC, but is negative with both NetPhos 2.0 and YinOYang 1.2 (very high just reaching the threshold line (Fig. 2a). With experimental evidence [Fujimoto and McEver, 1993] for its phosphorylation and positive prediction by NetPhosK1.0, Thr 817 is a false negative phosphorylation as well as false negative Yin Yang site. Threonine 817 may also be another possible interplay site for phosphate and/or *O*-β-GlcNAc modification involving PKC for phosphorylation as shown earlier for L- and P-selectin. Thus, Ser817 and Ser829 could be involved in a similar mechanism of complex interplay as proposed for L- and E-selectin.

From the above discussion, one common mechanism emerges for inside out signaling regulation of three selectins from their cytoplasmic domains is the PKC regulated phosphorylation of Ser/Thr. Additionally, an upstream Ser residue is proposed as source of interplay for phosphate and *O*-β-GlcNAc modification in regulating the functions and stability of three selectins. Although the three selectins perform different functions, a common mechanism may control their half lives and multifunctional behavior (Fig. 5a).

The addition of *O*-β-GlcNAc to serine and/or threonine and interplay with phosphorylation has not been reported until now for selectins. Our prediction results suggest that addition of *O*-β-GlcNAc or alternatively of phosphate in the cytoplasmic domain of L-, E-, and P-selectins

may regulate proteolytic shedding or endocytosis of those selectins (Fig. 5b). This should provide useful information to test the potential role of O-GlcNAc and phosphate modifications in controlling the half-lives of selectins. Many proteins have the potential to be modified transitorily in vivo by various substituents, and this property could be explored more extensively utilizing computer-assisted prediction tools for designing anti-inflammatory compounds.

ACKNOWLEDGMENTS

Nasir-ud-Din acknowledges partial support from HEC, Pakistan, and Pakistan Academy of Sciences for this research effort.

REFERENCES

Ahmad I, Hoessli DC, Walker-Nasir E, Rafik SM, Shakoori AR, Nasir-ud-Din. 2006. Oct-2 DNA binding transcription factor: Role of phosphorylation glycosylation and their interplay in its different domains. Nucl Acid Res 34:175–184.

Ahmad I, Hoessli DC, Walker-Nasir E, Shakoori AR, Nasir-ud-Din. 2006. Interplay between phosphorylation and glycosylation involving protein modifications at hydroxy amino acids: Prediction of signaling functions of the human β_3-integrin family. J Cell Biochem 99:706–718.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucl Acids Res 25:3389–3402.

Arribas J, Coodly L, Vollme P, Kishimoto TK, Rose-John S, Massagué J. 1996. Diverse cell surface protein ectodomains are shed by a system sensitive to metalloprotease inhibitors. J Biol Chem 271:11376–11382.

Attwood T. 2000. The quest to deduce protein function from sequence: The role of pattern databases. Int J Biochem Cell Biol 32:139–155.

Bennett TA, Lynam EB, Sklar LA, Rogelj S. 1996. Hydroxamate-based metalloprotease inhibitor blocks shedding of L-selectin adhesion molecule from leukocytes: Functional consequences for neutrophil aggregation. J Immunol 156:3093–3097.

Berlot S, Aissaoui Z, Pavon-Djavid G, Belleney J, Jozefowicz M, Helary G, Migonney V. 2002. Biomimetic poly (methyl methacrylate)-based terpolymers: Modulation of bacterial adhesion effect. Biomacromolecules 3:63–68.

Bevilacqua MP, Nelson RM. 1993. Selectins. J Clin Invest 91:379–387.

Blagoveshchenskaya DA, Hewitt EW, Cutler DF. 1998. A balance of opposing signals within the cytoplasmic tail controls the lysosomal targeting of P-selectin. J Biol Chem 273:27896–27903.

Blom N, Gammeltoft S, 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, Brunak S. 2004. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. Proteomics 4:1633–1649.

Bork P, Dansekar T, Diaz-Lazcoz Y, Eisenhaber F, Huynen M, Yuan Y. 1998. Predicting function: From genes to genome and back. J Mol Biol 283:707–725.

Chang KC, Lucas DA, Hise D, Scaefer CF, Xio Z, Janini GM, Buetow KH, Issaq HJ, Veenstra TD, Conrads TP. 2004. Analysis of human serum proteome. Clin Proteomics 1:101–226.

Cheng X, Hart GW. 2001. Alternative O-glycosylation/O-phosphorylation of serine-16 in murine estrogen receptor beta: Post-translational regulation of turnover and transactivation activity. J Biol Chem 276:10570–10575.

Chou TY, Hart GW, Dang CV. 1995. c-Myc is glycosylated at threonine 58, a known phosphorylation site and a mutational hot spot in lymphomas. J Biol Chem 270: 18961–18965.

Chuang PI, Young BA, Thiagarajan RR, Cornejo C, Winn RK, Harlan JM. 1997. Cytoplasmic domain of E-selectin contains a non-tyrosine endocytosis signal. J Biol Chem 272:24813–24818.

Collins T, Williams A, Johnston GI, Kim J, Eddy R, Shows T, Gimbrone MA Jr, Bevilacqua MP. 1991. Structure and chromosomal location of the gene for endothelial-leukocyte adhesion molecule 1. J Biol Chem 266:2466–2473.

Corvello CS, Furie BC, Furie B. 1993. Rapid phosphorylation and selective dephosphorylation of P-selectin accompanies platelet activation. J Biol Chem 268:14590–14593.

Ehrhardt C, Kneuer C, Bakowsky U. 2004. Selectins-an emerging target for drug delivery. Adv Drug Deliv Rev 56:527–549.

Elorza A, Penela P, Sarnago S, Mayor F Jr. 2003. MAPK-dependent degradation of G protein-coupled receptor kinase 2. J Biol Chem 278:29164–29173.

Feehan C, Darlak K, Kahn J, Walcheck B, Spatola AF, Kishimoto TK. 1996. Shedding of the lymphocyte L-selectin adhesion molecule is inhibited by a hydroxamic acid-based protease inhibitor. Identification with an L-selectin-alkaline phosphatase reporter. J Biol Chem 271:7019–7024.

Felin M, Doyannette-Moyne MA, Hadj-Sahraoui Y, Aubery M, Hubert J, Seve AP. 1994. Identification of two nuclear N-acetylglucosamine-binding proteins. J Cell Biochem 56:527–535.

Fors BP, Goodarzi K, von Andrian UH. 2001. L-Selectin shedding is independent of its subsurface structures and topographic distribution. J Immunol 167:3642–3651.

Fujimoto T, McEver RP. 1993. The cytoplasmic domain of P-selectin is phosphorylated on serine and threonine residues. Blood 82:1758–1766.

Gallatin MW, Weissman IL, Butcher EC. 1983. A cell-surface molecule involved in organ-specific homing of lymphocytes. Nature 304:30–34.

Green SA, Setiadi H, McEver RP, Kelly RB. 1994. The cytoplasmic domain of P-selectin contains a sorting determinant that mediates rapid degradation in lysosomes. J Cell Biol 124:435–448.

Griffith LS, Schmitz B. 1999. O-linked N-acetylglucosamine levels in cerebellar neurons respond reciprocally to pertubations of phosphorylation. Eur J Biochem 262: 824–831.

Gupta R, Jung E, Gooley AA, Williams KL, Brunak S, Hansen J. 1999. Scanning the available *Dictyostelium discoideum* proteome for *O*-linked GlcNAc glycosylation sites using neural networks. Glycobiology 9:1009–1022.

Hansen JE, Lund O, Tolstrup N, Gooley AA, Williams KL, Brunak S. 1998. NetOGlyc: Prediction of mucin type *O*-glycosylation sites based on sequence context and surface accessibility. Glycoconj J 15:115–130.

Haribabu B, Steeber DA, Ali H, Richardson RM, Snyderman R, Tedder TF. 1997. Chemoattractant receptor-induced phosphorylation of L-selectin. J Biol Chem 272: 13961–13965.

Hoessli DC, Poincelet M, Gupta R, Ilangumaran S, Nasir-ud-Din. 2003. Glycosylation and localization to low density, detergent resistant membranes in the parasitized erythrocyte. Eur J Biochem 270:366–375.

Itin C, Kappeler F, Linstedt AD, Hauri HP. 1995. A novel endocytosis signal related to the KKXX ER-retrieval signal. EMBO J 14:2250–2256.

Jeffery CJ. 1999. Moonlighting proteins. Trends Biochem Sci 24:8–11.

Johnston GI, Cook RG, McEver RP. 1989. Cloning of GMP-140, a granule membrane protein of platelets and endothelium: Sequence similarity to proteins involved in cell adhesion and inflammation. Cell 56:1033–1044.

Julenius K, Mølgaard A, Gupta R, Brunak S. 2005. Prediction, conservation analysis and structural characterization of mammalian mucin-type *O*-glycosylation sites. Glycobiology 15:153–164.

Kahn J, Ingraham RH, Shirley F, Migaki GI, Kishimoto TK. 1994. Membrane proximal cleavage of L-selectin: Identification of the cleavage site and Membrane proximal cleavage a 6-kD transmembrane peptide fragment of L-selectin. J Cell Biol 125:461–470.

Kaldjian E, McCarthy SA, Sharrow SO, Litrman DR, Klausner RD, Singer A. 1988. Nonequivalent effects of PKC activation by PMA on murine CD4 and CD8 cell surface expression. FASEB J 2:2801–2806.

Kamemura K, Hayes BK, Comer FI, Hart GW. 2002. Dynamic interplay between *O*-Glycosylation and *O*-phosphorylation of nucleoplasmic proteins. J Biol Chem 277:19229–19235.

Kelly WG, Dahmus ME, Hart GW. 1993. RNA polymerase II is a glycoprotein. Modification of the COOH-terminal domain by *O*-GlcNAc. J Biol Chem 268:10416–10426.

Kilian K, Dernedde J, Mueller E-C, Bahr I, Tauber R. 2004. The Interaction of protein kinase C isozymes α, ι, and θ with the cytoplasmic domain of L-selectin is Modulated by phosphorylation of the receptor. J Biol Chem 279: 34472–34480.

Kolling R, Hollenberg CP. 1994. The ABC-transporter Ste6 accumulates in the plasma membrane in a ubiquitinated form in endocytosis mutants. EMBO J 13:3261–3271.

La D, Sutch B, Livesay DR. 2005. Predicting protein functional sites with phylogenetic motifs. Proteins: Struct Funct Bioinform 58:309–320.

Matala E, Alexander SR, Kishimoto TK, Walcheck B. 2001. The cytoplasmic domain of L-selectin participates in regulating L-Selectin endoproteolysis. J Imunol 167: 1617–1623.

Matthews BW. 1975. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochim Biophys Acta 405:442–451.

Matthews JA, Acevedo-Duncan M, Potter RL. 2005. Selective decrease of membrane-associated PKC-α and PKC-ε in response to elevated intracellular O-GlcNAc levels in transformed human glial cells. Biochim Biophys Acta 1743:305–315.

Medina L, Grove K, Haltiwanger RS. 1998. SV40 large T antigen is modified with *O*-linked *N*-acetylglucosamine but not with other forms of glycosylation. Glycobiology 8:383–391.

Migaki GI, Kahn J, Kishimoto TK. 1995. Mutational analysis of the membrane-proximal cleavage site of L-selectin: Relaxed sequence specificity surrounding the cleavage site. J Exp Med 182:549–557.

Mullberg J, Rauch CT, Wolfson MF, Castner B, Fitzner JN, Otten-Evans C, Mohler KM, Cosman D, Black RA. 1997. Further evidence for a common mechanism for shedding of cell surface proteins. FEBS Lett 401:235–238.

Nasir-ud-Din, Hoessli DC, Rungger-Braendle E, Hussain SA, Walker-Nasir E. 2003. Role of sialic acid and sulfate groups in cervical mucus physiological functions: Study of *Macaca radiata* glycoproteins. Biochim Biophys Acta 1623:53–61.

Nielsen H, Brunak S, von Heijne G. 1999. Machine learning approaches for prediction of signal peptides and other protein signals. Protein Eng 12:33–39.

Nikolaidis N, Makalowska I, Chalkia D, Makalowski W, Klein J, Nei M. 2005. Origin and evolution of the chicken leukocyte receptor complex. Proc Natl Acad Sci USA 102:4057–4062.

O'Donnell N. 2002. Intracellular glycosylation and development. Biochim Biophys Acta 1573:336–345.

Ord DC, Ernst TJ, Zhou LJ, Rambaldi A, Spertini O, Griffin J, Tedder TF. 1990. Structure of the gene encoding the human leukocyte adhesion molecule-1 (TQ1, Leu-8) of lymphocytes and neutrophils. J Biol Chem 265:7760–7767.

Pawson T, Gish GD. 1992. SH2 and SH3 domains: From structure to function. Cell 71:359–362.

Roquemore EP, Chevrier MR, Cotter RJ, Hart GW. 1996. Dynamic *O*-GlcNAcylation of the small heat shock protein alpha B-crystallin. Biochemistry 3:2578–2586.

Rosen SD, Bertozzi CR. 1994. The selectins and their ligands. Curr Opinion Cell Biol 6:663–673.

Rumelhart DE, Hinton GE, Williams RJ. 1996. Learning internal representation by error propagation. In: Rumelhart D, McClelland J. and the PDP Research Group editors. Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 1: Cambridge, M.A.: Foundations, MIT Press, pp. 318–362.

Setiadi H, Disdier M, Green SA, Canfield WM, McEver RP. 1995. Residues throughout the cytoplasmic domain affect the internalization efficiency of P-selectin. J Biol Chem 270:26818–26826.

Shafi R, Iyer SP, Ellies LG, O'Donnel N, Marek KW, Chui D, Hart GW, Marth JD. 2000. The *O*-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. Proc Natl Acad Sci USA 97:5735–5739.

Smeets EF, de Vries T, Leeuwenberg JF, van den Eijnden DH, Buurman WA, Neefjes JJ. 1993. Phosphorylation of surface E-selectin and the effect of soluble ligand (sialyl Lewis X) on the half-life of E-selectin. Eur J Immunol 23:147–151.

Spertini O, Freedman AS, Belvin MP, Penta AC, Griffin JD, Tedder TF. 1991. Regulation of leukocyte adhesion molecule-1 (TQ1, Leu-8) expression and shedding by normal and malignant cells. Leukemia 5:300–308.

Tedder TF, Steeber DA, Chen A, Engel P. 1995. The selectins: Vascular adhesion molecules. FASEB J 9:866–873.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTALW: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl Acids Res 22:4673–4680.

Tonks NK, Neel BG. 1996. From form to function: Signaling by protein tyrosine phosphatases. Cell 87:365–368.

Varki A, Kornfeld S. 1980. Structural studies of phosphorylated high mannose-type oligosaccharides. J Biol Chem 255:10847–10858.

Vestweber D, Blanks JE. 1999. Mechanisms that regulate the function of the selectins and their ligands. Physiol Rev 79:181–213.

Walcheck B, Kahn J, Fisher JM, Wang BB, Fisk RS, Payan DG, Feehan C, Betageri R, Darlak K, Spatola AF, Kishimoto TK. 1996. Neutrophil rolling altered by inhibition of selectin shedding *in vitro*. Nature 380:720–723.

Werb Z, Yan Y. 1998. A cellular striptease act. Science 282:1279–1280.

Whelan SA, Hart GW. 2003. Proteomic approaches to analyze the dynamic relationships between nucleocytoplasmic protein glycosylation and phosphorylation. Circ Res 93:1047–1058.

Yoshida M, Szente BE, Kiely JM, Rosenzweig A, Gimbrone MA Jr. 1998. Phosphorylation of the cytoplasmic domain of E-selectin is regulated during leukocyte-endothelial adhesion. J Immunol 161:933–941.

Yuan ZQ, Feldman RI, Sussman GE, Coppola D, Nicosia SV, Cheng JQ. 2003. AKT2 inhibition of cisplatin-induced JNK/p38 and Bax activation by phosphorylation of ASK1: Implication of AKT2 in chemoresistance. J Biol Chem 278:23432–23440.

Zaia J, Boynton R, Heinegard D, Barry F. 2001. Post-translational modifications to human bone sialoprotein determined by mass spectrometry. Biochemistry 40: 12983–12991.