# STEROID REGULATION OF TERMINAL PROTEIN GLYCOSYLTRANSFERASE GENES: MOLECULAR AND FUNCTIONAL HOMOLOGIES WITHIN SIALYLTRANSFERASE AND FUCOSYLTRANSFERASE FAMILIES

Lucia Medvedova<sup>1</sup>, Jan Knopp, Robert Farkas

Institute of Experimental Endocrinology, Slovak Academy of Sciences, 833 06 Bratislava, Slovakia; <sup>1</sup>Department of Genetics, Graduate Programme, Faculty of Science, Comenius University, Bratislava, Slovakia

**Objective.** Terminal glycosylations such as fucosylation and especially sialylation are crucial in maturation of proteins for their particular function. In cells and tissues where their function is under hormonal control, often terminal processes including glycosylations are tightly regulated by hormones. Therefore, it is not surprising that two key enzyme families involved in these steps, fucosyltransferases and sialyltransferases, are directly controlled by steroid hormones.

**Methods.** Gene structure and protein sequences have been analyzed by set of Unix-based complex sequence analysis programs of the Wisconsin GCG package, and from some other resources. Protein secondary structure predictions were based on several independent programs and applications complementing each other depending on the algorithm used, all running on Unix platforms.

**Results.** Comparison of these two types of enzymes from organisms where they are known to be under direct hormonal regulation has revealed that although sialyltransferases are broad family of proteins, the only known enzymes to be under hormonal control are  $\alpha$ -2,6-sialyltransferases and in the case of fucosyl moieties transferring enzymes these are  $\alpha$ -1,2/6-fucosyltransferases. Interestingly, hormonally regulated sialyltransferases from as evolutionary distant organisms as human and fruitfly share several important features including secondary structure characteristics encompassing sialyl L and S motifs in which they differ from other sialyltransferases. Analogous organizational and structural similarities were found also in hormonally regulated fucosyltransferases.

**Conclusions.** Data indicate that hormonal control of these classes of enzymes as a regulatory mechanism could be attained relatively early during the evolution in order to obtain selective control over specific protein modification process which plays an important role during ontogenic development.

**Key words:** Steroid hormones - Sialyltransferase - Fucosyltransferase - Gene structure and expression - Protein sequence and secondary structure

In recent past it is growingly appreciated that the carbohydrate moieties of glycoproteins provide important prerequisites for various biological functions. By their potential for structural diversity they can serve as ideal candidates to carry also biological information. Therefore, it is not surprising to unravel programmed and strictly regulated changes in carbohydrate composition and sequence in glycoconjugates within the course of complex processes such as development and differentiation. Along these lines,

strong changes in presentation of cellular glycoproteins have also been detected upon malignant transformation and tumorigenesis (Schauer et al. 1988; Sharon 1991; Hakomori 1996). In addition, mevalonate-sensitive *N*-linked glycosylation of proteins is necessary for cell growth, including initiation and propagation of DNA synthesis in normal and tumortransformed cells (Carlberg and Larsson 1993; Wejde et al. 1993; Carlberg et al. 1996; Chan et al. 2001).

Glycoproteins as prominent and most variable group of glycoconjugates play a prominent part (e.g. as mucins) in the secretion. Developmental changes including tissue differentiation are accompanied by changes in glycosylation patterns. e.g. like a shift from sialylation to fucosylation or vice versa, depending on coordinate changes in glycosyltransferase activities, in sugar-nucleotide breakdown or synthesis and activity of regulatory proteins (BIOL-N'GARAGBA and Louisot 2003; Biol-N'GARAGBA et al. 2003). These activities are largely sensitive to developmental stimulations involving hormonal triggers (Sommer and Cowley 2001). Small lipophilic hormones represented mainly by steroids acting as ligands for cognate nuclear receptors are known to affect protein synthesis at translational and transcriptional levels. Terminal stages of O-glycosylation in various secretory or surface presented proteins are often associated with sialylation or fucosylation. This terminal posttranslational modification thus may play a crucial role in regulating competence of these proteins to become mature and capable to fulfill their function. As many steps following after transcriptional control up to executing the function of particular protein may not be tightly regulated or appear to be "left" under much less control, the very early as well as very last, terminal, steps seems to share strong inductive/decisive type of regulation. There is no doubt that both the sialylation and the fucosylation represents such a terminal stages of long process which require precise and coordinated control mechanisms. In the present paper we show that enzymes responsible for these terminal steps, sialyltransferases and fucosyltransferases, or their specific groups display quite interesting and common features in relation to their hormonal regulation which most probably became strongly conserved during evolution in such distant animals as mammals and invertebrates.

## **Materials and Methods**

**DNA and protein sequence analyses** including some secondary structure predictions were performed with GCG Wisconsin package (Accelrys Inc.) run under Irix platform (version 6.5.3) on Silicon Graphics Origin 2000 and under Solaris platform (version 5.8.2) on Sun Fire Enterprise 280 R computers. For

searching purposes, individual genomic versus cDNA clones were compared for assembly purposes using the program Compare of GCG package, and then individual responsive elements and binding motifs were found by using BestFit (SMITH et al. 1981) and Gap (Needleman and Wunch 1970) programs of GCG. Program Blast (Basic Local Alignment Search Tool) was initially used to determine most homologous partners among different organisms (ALTSCHUL et al. 1990). Prosite program which uses the sites and patterns defined in the PROSITE-database, developed and maintained by BAIROCH (1992), was used to identify protein modifications sites. SPScan predicts secretory signal peptides (SPs) using weigth matrix method of Nielsen et al. (1997). SignalSeq predicts the most likely cleavage site in a signal peptide sequence based on the rules and probability tables described in VON HAIJNE (1986).

Secondary structure predictions were made by several different methods including model recognition approach (MRA) and position specific iterated prediction (PSI-PRED) of Jones et al. (1994) and Jones (1991, 1999a,b), and according to Frishman and Argos (1992, 1995, 1996, 1997a,b) and Abagyan et al. (1994). The coordinates of the structurally conserved regions were then directly transferred from the template protein to the one to be modeled, while the structurally variable regions were built *de novo*. The resulting model was checked for the correct chirality, backbone angles, and possible bad steric contacts.

All the prediction and graphical computational work was performed on Silicon Graphics computers (Origin 2000 R10000 250 MHz or Octane 225 MHz) with graphical output at SGI O<sub>2</sub> R5000 terminal.

### **Results**

Among fucosyltransferases, the mouse fucosyltransferase FUT1 (accession No. Af214655; O09160) was unambiguously found to be under control of progesterone and estrogen (Sidhu and Kimber 1999). Search for any hormonally regulated fucosyltransferases provided evidence that all these enzymes fall into the subfamily of  $\alpha$ -1,2-fucosyltransferases (EC 2.4.1.69). As identified by bestfit and gap program analysis, the mouse and the  $\alpha$ -1,2-fucosyltransferases in other species including mouse, rat, pig, bovine

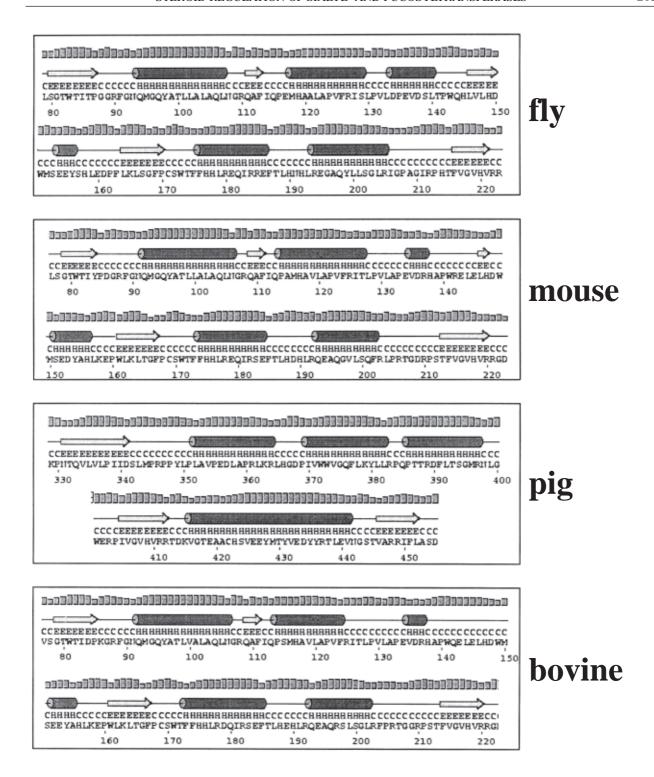


Fig 1 High degree of similarity in fucosyl motifs between *Drosophila*, mouse, pig and bovine  $\alpha$ -1,2/6-fucosyltransferases as compared by prediction of their secondary structures. Cylinders represent  $\alpha$ -helices, right-oriented arrows represent  $\beta$ -sheets and lines between them coil structures (according to the Psi-Pred method). Note conservation in sequence/length of  $\alpha$ -helices and  $\beta$ -sheets in particular regions.

and highly homologous Drosophila α-1,6-fucosyltransferase (accession No. Af441264; Q9vyv5; EC 2.4.1.68) contain hormone responsive elements with well conserved (95-100% homology) nucleotide sequences within the promoter regions, mostly -3500 to -500 bp upstream of the transcription start site, thus conferring direct hormonal responsiveness via nuclear steroid receptors. In contrast to other  $\alpha$ -1,2fucosyltransferases (so called FUT2), α-1,3-fucosyltransferases (so called FUT4, FUT6 and FUT7) or o-fucosyltransferases (OFU, EC 2.4.1.221), hormonally regulated vertebrate  $\alpha$ -1,2-fucosyltransferases of FUT1 subfamily and hormonally regulated Drosophila α-1,6-fucosyltransferase share significant homology in their N-terminally located fucosyl motifs systemically organized into 3 successive α-helices from which first two are always equally long and in some species interrupted with extremely short  $\beta$ sheet (Figure 1, upper row of each of the 4 panels). Second fucosyl motif is regularly composed of seven amino acid long β-sheet followed by two hydrophobic  $\alpha$ -helices or one apparently longer  $\alpha$ -helix which may represent evolutionary older version of originally unsplit helical structure (Figure 1, lower row of each of the 4 panels).

Another important group of O-glycosylation terminal transferases is represented by broad family of sialyltransferases from which only subgroup of  $\alpha$ -2,6-sialyltransferases (EC 2.4.99.1) displays hormone responsiveness. For example, rat  $\alpha$ -2,6-sialyltransferase (accession No. M18769; P13721) is controlled by glucocorticoids and in Drosophila by ecdysone (accession No. Af218237; Medvedova et al. in prep-

aration). In spite of their great evolutionary distance, hormonally regulated  $\alpha$ -2,6-sialyltransferases display several common features. Among various vertebrate species strongly conserved glucocorticoid response elements are found not only in promoters within -600 to -100 bp, but also within first intron (Figure 2). Analysis of the sialyl motifs revealed that major and first sialyl motif called "L" motif is in hormonally regulated α-2,6-sialyltransferases always split into 2 almost equally sized halfs by first intron which contains mentioned hormone responsive elements. The model for steroid hormone/receptor transcriptional cascade seems to be at the present best elaborated in Drosophila where we do know several other factors among early immediate steroid inducible genes the products of which, in addition to ecdysone receptor (EcR), participate in transducing and modulating hormonal signal. One of them is Broad-Complex (BR-C), a BTB/POZ Zn-finger transcription factor, encoded by ecdysone-inducible complex locus BR-C. Above mentioned intron which lies between two first exons of *Drosophila* α-2,6-sialyltransferase gene contains at least 2 BR-C binding sites, close to ecdysone responsive elements (EcREs). Furthermore, secondary structure prediction revealed that L motifs of vertebrate α-2,6-sialyltransferases are regularly composed of b-sheet followed by one or two  $\alpha$ -helices and another one or two  $\beta$ -sheets interconnected by 7 to 15 amino acids long coiled coils. As shown in Figure 3 (upper row of each of the 4 panels), in *Droso*phila the L motif of ecdysone-controlled  $\alpha$ -2,6-sialyltransferase appears to be most similar to human α-2,6-sialyltransferase (accession No. P15907). Even

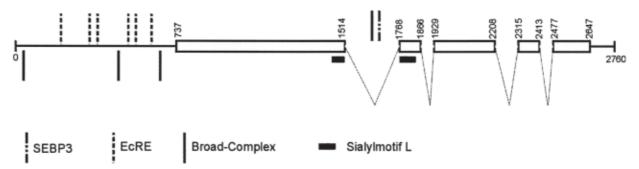


Fig 2 Genomic structure and organization of the  $\alpha$ -2,6-sialyltransferase gene from *Drosophila* with indication promoter sequences, introns and exons. Binding sites and responsive elements which are analogously distributed also in mammalian orthologues, and accounting for glucocorticoid response elements. Dark thick lines overlapping first intron are L sialyl motif encoded within the end of first exon and the beginning of second exon. Legends within transcription unit: unbroken vertical bars = Broad-Complex binding sites; dotted vertical bars = EcREs; two-dotted interrupted vertical bars = SEBP3 binding sites.

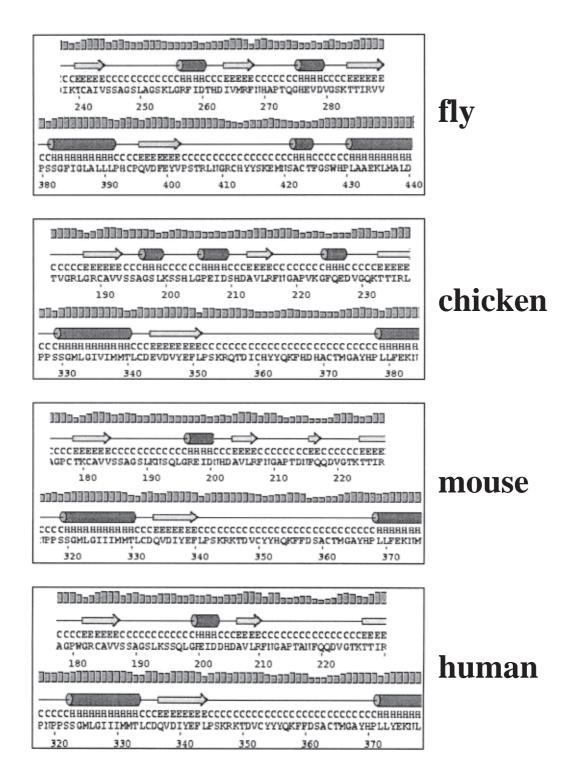


Fig 3 High degree of similarity in sialyl motifs between *Drosophila*, chicken, mouse and human  $\alpha$ -2,6-sialyltransferases as compared by prediction of their secondary structures using Psi-Pred method. Cylinders represent  $\alpha$ -helices, right-oriented arrows represent  $\beta$ -sheets and lines between them coil structures. Note that overall distribution (sequence and/or length) of  $\alpha$ -helices and  $\beta$ -sheets in particular regions is strongly conserved among evolutionary distant species.

higher degree of structural homology has been found between "S" and "VS" motifs of invertebrate and vertebrate hormonally regulated  $\alpha$ -2,6-sialyltransferases composed of ten mostly hydrophobic amino acids long  $\alpha$ -helix and seven largely acidic or aromatic amino acids long  $\beta$ -sheet followed by longer coil and second hydrophobic  $\alpha$ -helix, altogether spanning over 50 amino acid residues (Figure 3, lower row of each of the 4 panels).

Two additional features of hormonally regulated  $\alpha$ -1,2/6-fucosyltransferases and  $\alpha$ -2,6-sialyltransferases came to our attention. Both, α-1,2-fucosyltransferases and *Drosophila* α-1,6-fucosyltransferase share several common and conserved protein kinase C (PKC) phosphorylation and myristoylation sites indicating their regulation by PKC system which has been shown to be also under steroid hormone control (ARNOLD et al. 1994; FUJIMOTO and KATZENELLEN-BOGEN 1994; QIU et al. 2003), this way potentially ramificating hormonal signaling cascade to various cellular circuits, whereas myristoylation sites support independent Prosite analysis predicting discussed α-1,2-fucosyltransferases and *Drosophila* α-1,6-fucosyltransferase as microsomal membrane or endoplasmic reticulum-resident enzymes. However, detailed sequence analysis of all  $\alpha$ -2,6-sialyltransferases under study surprisingly revealed that besides being protein kinase A (PKA) and unspecified tyrosine kinase targets, they have greatest potential to be extracellular or secretory enzymes.

## **Discussion**

Broad spectrum of fucosyltransferases which are hormonally regulated represent  $\alpha$ -1,2-fucosyltransferase in mammals (Barreaud et al. 2000; Domino et al. 2001) and  $\alpha$ -1,6-fucosyltransferase in *Drosophila* which despite difference in substrate specificity/preference ( $\alpha$ -1,2 *versus*  $\alpha$ -1,6) show significant conservation in overall primary and secondary structure including relative size and organizational sequence of  $\alpha$ -helices and  $\beta$ -sheets. It appears that developmentally-linked switches in presentation or disappearance of fucosylated determinants on proteins had great importance and therefore became strongly conserved during evolution at various levels such as hormone responsiveness, location and orientation of hormone responsive elements within transcription

units, location of motifs and domains within protein molecule. Relatively high degree of conservation in the location and the alternating of specific amino acid motifs forming  $\alpha$ -helices or  $\beta$ -sheets suggest that interactions taking place at the level of tertiary protein structure within e.g. hydrophilic clefts or hydrophobic surfaces are similar if not identical between these fucosyltransferase and reflect their participation in closely related biological processes, possibly having highly homologous substrates in distant animal species.

As mentioned above, major sialyl motif "L" in hormonally regulated α-2,6-sialyltransferases has been found to be encoded by 2 separate exons being separated by an intron which bears hormone responsive elements, additional to those found in promoter region. This has not been observed in all other  $\alpha$ -2,6sialyltransferases which are hormone irresponsive. Thus, hormonally regulated sialyltransferases appear to constitute specific group of proteins the transcriptional regulation and gene structure organization of which was under similar if not conserved evolutionary pressure even in as distant species as insects and mammals separated by hundreds millions years of independent development. Therefore, there is a good reason to believe that under these circumstances development could be so tight that also some other factors and/or players of this machinery were kept conserved as well which may include some components of glucocorticoid early immediate response orthologous to those found in Drosophila EcR and BR-C cascade.

Most sialyltransferases including  $\alpha$ -2,6-sialyltransferase family members along with mannosidases I and II are considered to be membrane-bound Golgi proteins involved in terminal modification of secretory proteins with extracellular function (Dong et al. 2000). Our data suggest that amino acid sequence of these enzymes bears features allocating them with high probability to extracellular space as well. This surprising outcome is gaining experimental support in observations of so-called SGS secretory protein sialylation in *Drosophila* salivary glands where tissue-specific hormonal control of  $\alpha$ -2,6-sialyltransferase expression has been documented (MED-VEDOVA et al. in preparation). Under this system, the process of granule maturation by sialylation of SGS proteins is continued upon and probably after exocytosis of the secretory material to extracellular space. Presence of signal peptide sequence and anchoring motif near the N-terminal end of at least  $Drosophila\,\alpha\text{-}2,6\text{-sialyltransferase}$  indicates that even in the case of being secreted, this group of proteins may remain inserted in the membraneous structure. This finding extends function of  $\alpha\text{-}2,6\text{-sialyltransferases}$  behind their regular subcellular regions, and provides reasonable link between predictible analysis based on protein primary structure and its true function-related activity.

The presence of highly conserved PKC phosphorylation motifs in  $\alpha\text{--}1,2/6\text{-fucosyltrans}$  ferases and PKA motifs in  $\alpha\text{--}2,6\text{-sialyltrans}$  ferases indicates that besides hormone responsiveness at transcriptional level the control of the particular enzymatic function is under more complex regulation which offers possibility of its posttranslational control by several signaling pathways converging at this point, and still may include

original steroid hormone activated cascade as reflected in steroid receptor modification by PKC and PKA activity (Arnold et al. 1994; Kuiper and Brinkmann 1994; Kato et al. 1995; Bunone et al. 1996; El-Tanani and Green 1997; Joel et al. 1998).

# Acknowledgements

We would like to express our gratitude to Computing Centre of the Slovak Academy of Sciences for services and backup provided during computational and graphical analyses presented in this paper. This work was supported, in part, by the Research Grant 2/999533 from GAV and by Grants No. 2/7194/20 and 2/3025/23 from VEGA Slovakia, and US-Slovak collaborative grant 029/2001 from APVT to R.F. International collaborative grants from the NATO (CRG-972173 and LST.CLG-977559) are also acknowledged.

## REFERENCES

- ABAGYAN R, FRISHMAN D, ARGOS P: Recognition of distantly related proteins through energy calculations. Proteins 19: 132-140, 1994
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: Basic alignment search tool. J Mol Biol **215**: 403-410, 1990 Arnold SF, Obourn JD, Jaffe H, Notides AC: Serine 167 is the major estradiol-induced phosphorylation site on the human estrogen receptor. Mol Endocrinol **8**: 208-1214, 1994
- BAIROCH A: PROSITE: a dictionary of sites and patterns in proteins. Nucl. Acid. Res. 20s: 2013-2018, 1992
- Barreaud JP, Saunier K, Souchaire J, Delourme D et al.: Three bovine alpha-2-fucosyltransferase genes encode enzymes that preferentially transfer fucose on Gal-beta1-3GalNAc acceptor substrates. Glycobiology **10**: 611-621, 2000
- BIOL-N'GARAGBA M-C, LOUISOT P: Regulation of the intestinal glycoprotein glycosylation during postnatal development: role of hormonal and nutritional factors. Biochimie **85**: 331-352, 2003
- BIOL-N'GARAGBA M-C, NIEPCERON E, MATHIAN B, LOUISOT P: Glucocorticoid-induced maturation of glycoprotein galactosylation and fucosylation processes in the rat small intestine. J Steroid Biochem Mol Biol 84: 411-422, 2003
- Bunone G, Briand PA, Miksicek RJ, Picard D: Activation of the unliganded estrogen receptor by EGF involves the MAP kinase pathway and direct phosphorylation. EMBO J 15: 2174-2183, 1996
- Carlberg M, Dricu A, Blegen H, Wang M et al.: Mevalonic acid is limiting for N-linked glycosylation and translocation of the insulin-like growth factor-1 receptor to the cell surface. Evidence for a new link between 3-hydroxy-3-methylglutaryl-coenzyme a reductase and cell growth. J Biol Chem **271**: 17453-17462, 1996
- CARLBERG M, LARSSON O: Role of *N*-linked glycosylation in cell-cycle progression and initiation of DNA synthesis in tumor-transformed human fibroblasts. Anticancer Res **13**: 167-171, 1993
- Chan FL, Choi HL, Ho SM: Analysis of glycoconjugate patterns of normal and hormone-induced dysplastic noble rat prostates, and an androgen-independent noble rat prostate tumor, by lectin histochemistry and protein blotting. Prostate **46**: 21-32, 2001
- Domino SE, Zhang L, Lowe JB: Molecular cloning, genomic mapping, and expression of two secretor blood group alpha-(1,2)-fucosyltransferase genes differentially regulated in mouse uterine epithelium and gastrointestinal tract. J Biol Chem **276**: 23748-23756, 2001
- EL-Tanani MK, Green CD: Two separate mechanisms for ligand-independent activation of the estrogen receptor. Mol Endocrinol 11: 928-937, 1997

- Frishman D, Argos P: Recognition of distantly related protein sequences using conserved motifs and neural networks. J Mol Biol 228: 951-962, 1992
- FRISHMAN D, Argos P: Knowledge-based protein secondary structure assignment. Proteins 23: 566-579, 1995
- Frishman D, Argos P: Incorporation of non-local interactions in protein secondary structure prediction from the amino acid sequence. Protein Eng 9: 133-142, 1996
- Frishman D, Argos P: The future of protein secondary structure prediction accuracy. Fold Des 2: 159-162, 1997a
- Frishman D, Argos P: Seventy-five percent accuracy in protein secondary structure prediction. Proteins 27: 329-335, 1997b
- Fuлмото N, Katzenellenbogen BS: Alteration in the agonist/antagonist balance of antiestrogens by activation of protein kinase A signaling pathways in breast cancer cells: antiestrogen selectivity and promoter dependence. Mol Endocrinol 8: 296-304, 1994
- HAKOMORI S: Tumor malignancy defined by aberrant glycosylation and sphingo (glyco) lipid metabolism. Cancer Res **56**: 5309-5318, 1996
- Joel PB, Smith J, Sturgill TW, Fisher TL, Blenis J, Lannigan DA: pp90rsk1 regulates estrogen receptor-mediated transcription through phosphorylation of Ser-167. Mol Cell Biol 18: 1978-1984, 1998
- Jones DT: The application of fractal clustering to efficient molecular ray tracing on low-cost computers. J Mol Graph 9: 249-253, 1991
- Jones DT: GenTHREADER: an efficient and reliable protein fold recognition method for genomic sequences. J Mol Biol **287**: 797-815, 1999a
- Jones DT: Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol **292**: 195-202, 1999b
- Jones DT, Taylor WR, Thornton JM: A model recognition approach to the predication of all-helical membrane protein structure and topology. Biochemistry **33**: 3038-3049, 1994
- Kato S, Endoh H, Masuhiro Y, Kitamoto T et al.: Activation of the estrogen receptor through phosphorylation by mitogen-activated protein kinase. Science **270**: 1491-1494, 1995
- Kuiper GG, Brinkmann AO: Steroid hormone receptor phosphorylation: is there a physiological role? Mol Cell Endocrinol **100**: 103-107, 1994
- Needleman SB, Wunsch CD: A general method applicable to the search for similarities in the amino acid sequence of two proteins. Joliol **48**: 443-453, 1970
- NIELSEN H, ENGELBRECHT J, BRUNAK S, VON HEIJNE G: Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Engin 10: 1-6, 1997
- QIU J, BOSCH MA, TOBIAS SC, GRANDY DK et al.: Rapid signaling of estrogen in hypothalamic neurons involves a novel G-protein-coupled estrogen receptor that activates protein kinase C. J Neurosci 23: 9529-9540, 2003
- Schauer R, Fischer C, Lee H, Ruch B, Kelm S: Sialic acids as regulators of molecular and cellular interactions. In: Lectins and Glycoconjugates in Oncology (edited by Gabius H.J. and Nagel G.A.); pp. 5-23. Springer-Verlag, Berlin 1988
- Sharon N: Molecular basis of lectin-carbohydrate interactions. In: Lectins and Cancer (edited by Gabius H.J. and Gabius S); pp. 1-12. Springer-Verlag, Berlin 1991
- SMITH TF, WAREMAN HS, FITCH WM: Comparative biosequence metrics. J Mol Evol **18**: 38-46, 1981 SOMMER P, COWLEY HM: Effect o steroid hormone on membrane profiles of HeLa S3 cells. Cell Biol Int **25**: 103-111, 2001
- VON HAIJNE G: A new method for predicting signal sequence cleavage sites. Nucl Acid Res 14: 4683-4690, 1986
- Wejde J, Carlberg M, Hjertman M, Larsson O: Isoprenoid regulation of cell growth: identification of mevalonate-labelled compounds inducing DNA synthesis in human breast cancer cells depleted of serum and mevalonate. J Cell Physiol **155**: 539-548, 1993

Corresponding author: Robert Farkaš

Institute of Experimental Endocrinology

Slovak Academy of Sciences Vlárska 3, 833 06 Bratislava

Slovakia

phone: (+4212) 5477-3800, ext. 216

fax: (+4212) 5477-4247

E-mail: ueenfark@savba.savba.sk