Mutation-prone Positions Within the Estrogen Receptor

Viroj Wiwanitkit*
Wiwanitkit House, Bangkhae, Bangkok, Thailand.

SUMMARY

Objective: Estrogen is an important female hormone. The estrogen receptor (ER) plays a critical role in the development of breast cancer. Mutations within the ER can have clinically significant consequences.

Materials and Methods: Identification of sites vulnerable to mutation is a new trend aimed at extending our knowledge of diseases at the genomic and proteomic levels. In this study, bioinformatics analysis was performed to determine the positions corresponding to specific peptide motifs in the amino acid sequence of the ER. A new computational tool called GlobPlot was used to identify the weak linkage positions within the ER.

Results: The results allowed the identification of mutation-resistant positions.

Conclusion: This study showed that the weak linkages within the ER could be identified, and these could provide the basis for further studies aimed at predicting possible new ER mutations. [*Taiwan J Obstet Gynecol* 2009; 48(2):130–132]

Key Words: estrogen, mutation, receptor, weak linkage

Introduction

Estrogen is an important female hormone. It has several roles in gynecology and obstetrics. Estrogen receptors (ERs), both α and β types, are members of the steroid/thyroid nuclear receptor superfamily of ligand-dependent transcription factors [1]. The impact of ER-α on breast cancer pathophysiology and disease progression is now well established [1,2], and the ER gene has been studied in many clinical conditions. Pathologic aberrations of ER structure and function are a major cause of estrogen insensitivity in affected females and a possible underlying cause of breast cancer. Indeed, estrogen plays important roles in the pathologic proliferation of cancer cells in reproductive organs, especially those processes in the pathologies of the breast and uterus [3]. Based on the observation of bilateral risks and frequent multifocality with atypical ductal hyperplasia, atypical lobular hyperplasia and lobular breast carcinoma *in situ*, it is concluded that estrogen may represent both risk factors and precursors [4]. ER-β-positivity and ER-α-negativity can characterize the highest levels of proliferative breast cancer cell activity [4].

Many naturally occurring splice variants of both ER isoforms have been detected in normal epithelium and in diseased or cancerous tissues [5]. In contrast, only a few ER point mutations have been confirmed in human patient samples from a variety of disease states, including breast cancer, endometrial cancer and some specific psychiatric diseases [5]. Depth analysis of the ER can improve our understanding of the pathogenesis of ER-related disorders. The prevalence of mutations in the ER receptor gene is unknown. Analysis of mutation-susceptible positions in the ER is the first step in furthering research into ER mutations, followed by the functional characterization of these mutations.

The prediction of protein nanostructure and function is a major focus of proteomic and genomic research. Identification of points vulnerable to mutation is a new approach that could extend our understanding of many diseases at the genomic and proteomic level [6,7]. Mutated regions in proteins usually contain specific,

*Correspondence to: Dr Viroj Wiwanitkit, Wiwanitkit House, Bangkhae, Bangkok 10160, Thailand.
E-mail: wviroj@yahoo.com
Accepted: July 20, 2008

```
MTMPLPNKTT gvtflhqiqs seletltrpp lkislerplg emyvennrtg IFNYPegtty dfaaaaapvy ssaslsyaas setFGSSSLT
GLHTLNNVPP Spvvflaklp qlspfihhhg qqvpyylese qgtfavreaa pptFYRSSSD NRRQSGRERM SSANDKGPPS MEStketryc
avcsdyasgy hygvwscegc kaffkrsiqg hNDYMCPATn qctidknrrk scqacrlrkc yevgmmkggi rkdrrggrll khkrqkeeqe
qkndvdpsei rtasiwvnps vksmklspvl sltaeqlisa lmeaeapivy sehdstkpls easmmtlltn ladrelvhmi nwakrvpgfv
dltlhdqvhl lecawleilm vgliwrsveh pgklsfapnl lldrnqgrcv eglveifdml vttatrfrmm rlrgeeficl ksiillnsgv
ytflsstles ledtdlihii ldkiidtlvh fmaksglslq qqqrrlaqll lilshirhms nkgmehlysm kcknvvplyd lllemldahr
ihtpkdkTTT QEEDSRSPPT TTVNGASPCL QpyytnteeV SLQStv
```

Lower case letters represent amino acids and upper case letters represent resistant positions.

Figure. Mutation-resistant and mutation-prone positions within the estrogen receptor amino acid sequence.

short linear peptide motifs that are crucial for protein function. Identification of these peptide motifs in the amino acid sequence can help to predict the weak linkages in a protein [6,7]. In this study, the author performed bioinformatics analysis to determine the vulnerable positions characterized by those specific peptide motifs in the amino acid sequence of the ER.

Materials and Methods

Determination of the ER sequence

The Expert Protein Analysis System (ExPASy) [8] was used to determine the amino acid sequence of the ER. ExPASy is a proteomics server copyrighted to the Swiss Institute of Bioinformatics [8]. ExPASy is designed for the analysis of protein sequences and structures, as well as for two-dimensional polyacrylamide gel electrophoresis [8]. The derived ER sequence was used as a template sequence for identifying weak linkages within the ER.

Identification of weak linkages within the ER

A computational tool called GlobPlot [9] was used to identify weak linkages in the ER using bioinformatics techniques. GlobPlot is a web-based service aimed at helping researchers to plot tendencies for order/globularity and disorder in target proteins [9]. The specific interface is simple to use; pasting a sequence or entering the SWISS-PROT/SWALL accession or entry code allows further bioinformatics analysis to be performed [9]. Functionally, the GlobPlot server retrieves the entered sequence and a description of the polypeptide from an ExPASy server using Biopython.org software. Putative globular and disorder segments are further selected using a simple peak finder algorithm known as PeakFinder [9]. Specific identified peaks are selected when the first derivative shows positive values over a continuous stretch of a minimum length [9]. Using this bioinformatics process, GlobPlot can identify inter-domain segments containing linear motifs and ordered regions with no recognized domains [9]. Theoretically, the resulting determined motifs are the weak linkages in a protein that are prone to point mutations [6,7].

Results

The ER sequence (P81559) was derived by ExPASy searching and was then used for further study. Several motifs were identified (as shown in the Figure). Positions 1–10, 51–55, 84–101, 144–173, 212–219, 548–571 and 580–584 were identified as resistant positions.

Discussion

The temporal and tissue-specific actions of estrogen are physiologically mediated by ERs α and β [10]. ERs are steroid hormone receptors that modulate the transcription of target genes when specifically bound to their ligand [10]. There is a large and increasing body of experimental and clinical data supporting the existence of variant forms of ER proteins in both normal and neoplastic estrogen target tissues, including human breast tissue [11]. The physiologic and pathologic functions of variant ER proteins remain unclear, though a role for some ER variants in breast tumorigenesis and breast cancer progression would be consistent with the accumulated published data [11].

ER mutations are involved in some rare syndromes characterized by estrogen disorders, while the molecular pathophysiology of other ER disorders, including breast cancer, are associated with post-receptor modifications. Identification of the mutation points within the ER can be useful for furthering research into the pathogenesis of these estrogen disorders. However, identification of ER mutation points using classical *in vitro* experiments is time consuming and has a low probability of identifying positive cases. Developments in bioinformatics, however, can help to solve this problem [12]. In this study, the author used a computational bioinformatics algorithm to identify the positions in the ER amino acid sequence that were susceptible to

mutation. The standard protein database tool ExPASy was used for the initial data-searching step. This computational database tool has been widely used and can produce highly reliable results [13]. GlobPlot was used for further identification of mutation-prone sites within the ER. This new technique has been used to identify mutations in many previously published studies [14–18].

In the current study, the tool successfully identified several positions. Some of these positions were previously-documented mutation points, whereas others were identified for the first time in this study. Based on these data, the weak linkages in the ER can be successfully identified and can provide useful data for further studies aimed at predicting possible new ER mutations associated with overt ER abnormalities. Further translational research is needed to relate phenotypic expression to the mutation-prone positions in the ER. These results will also be useful for further functional characterization predictions, which can be performed using new gene ontology technology techniques [19,20]. The functional significance of the identified mutation-resistant locations within the ER should also be further assessed and functionally characterized in future studies.

References

1. Leclercq G, Lacroix M, Laios I, Laurent G. Estrogen receptor alpha: impact of ligands on intracellular shuttling and turnover rate in breast cancer cells. *Curr Cancer Drug Targets* 2006;6:39–64.
2. Cordera F, Jordan VC. Steroid receptors and their role in the biology and control of breast cancer growth. *Semin Oncol* 2006;33:631–41.
3. Ikeda K, Inoue S. Estrogen receptors and their downstream targets in cancer. *Arch Histol Cytol* 2004;67:435–42.
4. Schneider HP, Bocker W. Hormones and progeny of breast tumor cells. *Climacteric* 2006;9:88–107.
5. Herynk MH, Fuqua SA. Estrogen receptor mutations in human disease. *Endocr Rev* 2004;25:869–98.
6. Lee C, Wang Q. Bioinformatics analysis of alternative splicing. *Brief Bioinform* 2005;6:23–33.
7. Levin JM, Penland RC, Stamps AT, Cho CR. Using in silico biology to facilitate drug development. *Novartis Found Symp* 2002;247:222–38.
8. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* 2003;31:3784–8.
9. Linding R, Russell RB, Neduva V, Gibson TJ. GlobPlot: exploring protein sequences for globularity and disorder. *Nucleic Acids Res* 2003;31:3701–8.
10. Pearce ST, Jordan VC. The biological role of estrogen receptors alpha and beta in cancer. *Crit Rev Oncol Hematol* 2004; 50:3–22.
11. Murphy LC, Dotzlaw H, Leygue E, Douglas D, Coutts A, Watson PH. Estrogen receptor variants and mutations. *J Steroid Biochem Mol Biol* 1997;62:363–72.
12. Gogarten JP, Olendzenski L. Orthologs, paralogs and genome comparisons. *Curr Opin Genet Dev* 1999;9:630–6.
13. Golaz O, Wilkins MR, Sanchez JC, Appel RD, Hochstrasser DF, Williams KL. Identification of proteins by their amino acid composition: an evaluation of the method. *Electrophoresis* 1996;17:573–9.
14. Wiwanitkit V. Where is the weak linkage in the globin chain? *Int J Nanomedicine* 2006;1:109–10.
15. Wiwanitkit V. Mutation-prone points in thrombin receptor. *J Thromb Thrombolysis* 2008;25:190–2.
16. Wiwanitkit V. Weak linkage in androgen receptor: identification of mutation-prone points. *Fertil Steril* 2009;91:e1–3.
17. Wiwanitkit V. Identification of weak points prone for mutation in ferredoxin of *Trichomonas vaginalis*. *Indian J Med Microbiol* 2008;26:158–9.
18. Wiwanitkit V. Identification of mutation-prone points in bile salt export pump. *HPB (Oxford)* 2007;9:444–6.
19. Azuaje F, Al-Shahrour F, Dopazo J. Ontology-driven approaches to analyzing data in functional genomics. *Methods Mol Biol* 2006;316:67–86.
20. Khan S, Situ G, Decker K, Schmidt CJ. GoFigure: automated Gene Ontology annotation. *Bioinformatics* 2003;19:2484–5.