Original Paper



Ann Neurosci 2018;25:1–10 DOI: 10.1159/000481809 Received: January 27, 2017 Accepted: June 28, 2017 Published online: October 23, 2017

Estrogen Receptor α- and β-Interacting Proteins Contain Consensus Secondary Structures: An Insilico Study

Vijay Paramanik^b Harini Krishnan^a Mahendra Kumar Thakur^a

^aDepartment of Zoology, Biochemistry and Molecular Biology Laboratory, Banaras Hindu University, Varanasi, India; ^bDepartment of Zoology, Cellular and Molecular Neurobiology and Drug Targeting Laboratory, Indira Gandhi National Tribal University, Amarkantak, India

Keywords

Estrogen receptor $\alpha \cdot$ Estrogen receptor $\beta \cdot$ Estrogen receptor coregulators \cdot Secondary structure \cdot Homology \cdot Insilico study

Abstract

Background: Estrogen receptor (ER)a and ERB are ligandactivated transcription factors that regulate gene expression by binding to estrogen-responsive elements and interacting with several coregulators through protein-protein interactions. Usually, these coregulators bind to the various conserved and functional domains of the receptor through a consensus LXXLL sequence, although variations can be found. The interaction of receptor domains and the consensus motif can be a possible target for nuclear receptor (NR) pharmacology, since modifications in these are responsible for possible pathogenesis of various diseases. **Purpose:** The present study focuses on the secondary structure and conserved domains of the ERa and ERB interacting proteins, using bioinformatics tools and their relation to the function of the coregulators. Methods: Bioinformatics-based prediction tools like STRING, PSIPRED, PROTPARAM and Conserved Domain Database (CDD) were used. The prediction tools utilized in this study basically determines the characteristics of a possible coregulator by using an already existing protein as a template and determines the presence of any conserved

KARGER

© 2017 S.Karger AG, Basel

E-Mail karger@karger.com www.karger.com/aon consensus sequence. Coregulators have been enlisted with the help of NCBI, STRING and iHOP. The secondary structures were analyzed using PSIPRED and conserved domains were determined using CDD. **Results:** The analysis of the structure has shown the presence of conserved domains and homology between the various coregulators. Each interacting protein contains conserved domains like the nuclear coactivators' domain, the helix-loop-helix domain and the SRC domain. **Conclusion:** Such studies give the characteristic features of ERα and ERβ interacting proteins and maybe useful to determine their family and uses in NR pharmacology in health and diseases. © 2017 S.Karger AG, Basel

Introduction

Estrogens are steroid molecules and primarily secreted by the female gonads. Apart from playing an important role in the regulation of the female reproductive systems, they have important functions in bone development, in lipid metabolism and in maintaining the cardiovascular system and the neuronal systems [1, 2]. Estrogens are known to exert their functions through ER, which are ligand-activated transcription factors belonging to the nuclear receptor (NR) superfamily. Like the other members of the NR superfamily, ER contains 6 functional domains

Vijay Paramanik

Department of Zoology, Cellular and Molecular Neurobiology and Drug Targeting Laboratory, Indira Gandhi National Tribal University Amarkantak 484887, Madhya Pradesh (India) E-Mail vijay.paramanik@igntu.ac.in; vijayparamanik@gmail.com

[3, 4]. The DNA-binding domain shares the highest homology with the other NR superfamily members and is the most conserved domain. The ligand binding domain (LBD) is relatively less conserved in nature and contains the ligand-dependent transcription activation function activation function (AF)-2. The most variable portion of the ER is the A/B domain that also contains a ligand independent, constitutively active transcription activation factor AF-1.

Biochemical analysis has shown the presence of 2 isoforms of the ER, namely, ER α and ER β . Although ER α an ER β share the DBD, they differ to some extent in their LBD and thus their ligand-binding affinity and expression patterns. ER β modulated the function of ER α when they are co-expressed and supplements the activity of ER α . Most effects of estrogen are cell-specific, which can be attributed to the differential expression of the ER in different tissues. Studies have shown the presence of ER α in the uterus, liver, kidney and heart, whereas ER β is shown to be present in the gastrointestinal tract, ovary, prostate, lung, bladder and the central nervous system. The co-expression of the 2 isoforms is seen in the mammary gland, epididymis, adrenal gland and regions of brain like the hypothalamus and the amygdala [5].

Usually, the ERs are bound to the heat shock proteins 90 present in the cytoplasm [6]. On interaction with the ligand, phosphorylation takes place at the serine and the threonine residues, which causes a conformational change in the receptor. This releases the receptor from the heat shock proteins complex and interaction with the respective response elements (ERE) is facilitated. This acts as a scaffold for the recruitment for a number of other factors termed as coregulators. The coregulators facilitate estrogen-mediated chromatin reorganization. By altering the chromatin structure, the coregulators provide interaction with receptor and the transcription machinery, thereby controlling the transcription.

Coregulators exist as multi-protein complexes and can be broadly classified into 2 categories. The first class contains the interacting proteins that have histone-acetylating or -deacetylating activities. The most widely studied coregulators family under this category is the p160 family of coregulators. As the name suggests, the family contains molecules that are 160kD in size and basically comprise SRC1, GRIP1/TIF2 and AIB1/ACTR/pCIP [7–9]. They function by binding histone acetyl transferases such as p300/CBP and pCAF, which in turn regulated the chromatin remodeling. The second class contains the interacting proteins that facilitate chromatin remodeling through ATP hydrolysis. This class includes the SWI/ SNF family of coregulators. They have homology similar to the helicases [10, 11]. The coregulators can be coactivators like SRC1, ERAP140 [12], GRIP1, TRAP220 and so on or can be corepressors like NCoR and SMRT (Table 1). The variation in the expression and interaction of ER coregulators and their dysregulation is shown to have severe consequences for cellular homeostatsis and often contributes to pathogenesis.

In addition, these coregulators have been shown to interact with the receptor through a consensus LXXLL sequence, also called NR boxes [13], although variation is reported. These structural motifs are found as helices. The leucines create a hydrophobic surface that interacts with the hydrophobic groove of the receptor [14]. The interaction between the receptor motifs and the NR boxes is used as a possible target for NR pharmacology studies [15]. The structure and properties of the interacting proteins per se coregulators play an important role in determining their interaction with ER α and ER β and function. There exists a great deal of homology among various coregulators due to the presence of conserved domains and responsible for specific functions. The present study is focused to determine the secondary structure of ERa and ER^β interacting proteins and their conserved domains using various bioinformatics tools. Such a study may be useful to determine the family of the coregulators, newly identified, and helpful to understand the etiology of different diseases and treatment by NR pharmacology.

Methods

Enlisting ERα and ERβ Interacting Proteins

ER α and ER β interacting proteins were enlisted using the NCBI, STRING and iHOP databases. STRING is a database that maintains a list of known or predicted protein interactions that are derived from 4 basic sources: genomic analysis, high throughput experiments, conserved coexpression and previously known literature. iHOP provides information about the proteins and their interacting members in a concise manner with access to related literature. NCBI is a database maintained by the National Institutes of Health, USA. Amino acid sequences of the protein can be obtained in the FASTA format from the NCBI database.

Prediction of Secondary Structure of ER α and ER β Interacting Proteins

The secondary structure was predicted using prediction tools such as PSIPRED. PSIPRED is a tool that predicts the secondary structure by using the information gained from PSI-BLAST. The amino acid sequence of the interacting domain was input in the FASTA format, derived from NCBI database and the secondary structure is determined in the form of helices denoted by cylindrical figures and strands denoted by arrow marks.

S. No.	Name of the protein	Tissue specificity	Disease correlation
1	SRC-1	Brain (hypothalamus, hippocampus, cerebellum, PVN)	Breast cancer, endometrial cancer
2	TIF2	Pancreas, skeletal muscle, liver, lung, placenta, brain, heart	Colorectal cancer, acute mylemonocytic leukemia
3	GRIP1	Muscle, neural tissue, mammary gland	Breast cancer
4	AIB1	Heart, skeletal muscle, pancreas and placenta	Polycystic ovarian syndrome, ovarian cancer, breast cancer
5	ARA70	Adipose tissue, testis	Polycystic ovarian syndrome, prostate cancer
6	NCoA3	Heart, skeletal muscle, pancreas, and placenta	Breast cancer
7	СВР	Brain, skeletal muscles	Rubinstein-Taybi syndrome, Huntington's disease
8	pCAF	Intestines	Colorectal cancer
9	TRAP220 (MED1)	Ubiquitously present	Lung adenocarcinoma, crohn's disease
10	SRA	Mammary gland, ovaries, placenta, adipose tissue	Breast cancer, prostate cancer, ovarian cancer
11	p68	Mammary gland, adipose tissue	Breast cancer
12	ASC-2	Mammary gland, spleen, liver	Breast cancer
13	PNRC	Liver, lung, adipose tissue, NK/T cells	Hepatitis B, breast Cancer
14	SPBP (TCF20)	Brain, heart, testis, kidney, liver, lungs	Arrhythmia, cardiac problems
15	DHX9	Neuronal cells, pituitary, ovary	Werner syndrome
16	STAU1	Brain, pancreas, heart, skeletal muscles, liver, lung, kidney	Cardiac diseases, lung carcinomas
17	EIF4G	Heart, skeletal muscles	Encephalomyocarditis virus infection
18	PELP1	Breast cancer cells	Endometrial cancer, breast cancer
19	PGC1	Heart, brain, skeletal muscles	Obesity, type II diabetes
20	NRIP1	Placenta, ovary, brain, lung, stomach, kidney	Post-menopausal osteoporosis, infertility
21	N-COR	Lung, spleen, brain, mammary gland	Breast cancer, myeloid leukemia
22	SMRT (NCOR2)	Ubiquitous, high levels in embryo	B-cell lymphoma, myeloid leukemia
23	RIP140	Placenta, ovary, liver	Post-menopausal osteoporosis, infertility
24	CARM1	High expression in prostate, epithelial cells	Prostate cancer, vascular diseases
25	SHP	Liver, heart, pancreas, placenta	Obesity, endometrial cancer, insulin resistance.
26	DAX	Adrenal gland, pituitary	Hypogonadism, congenital adrenal hyperplasia
27	HET-SAFB	Mammary gland, stomach, liver	Breast cancer, duodenal ulcer

$\textbf{Table 1.} ER \alpha \ interacting \ proteins$

Prediction of the Physical Parameters of the ER α and ER β Interacting Proteins

The physical parameters were determined using ExPasy tool and PROTPARAM. The protein sequence was input in the FASTA format from the NCBI database. PROTPARAM gives information about molecular weight, stability index, extinction coefficients, half-life and the amino acid composition of the protein.

 $ER\alpha\mathchar`-$ and $\beta\mathchar`-$ Interacting Proteins Contain Consensus Secondary Structures

Prediction of Conserved Sequences among ER α and ER β Interacting Proteins

Homology modeling was done using the bioinformatics tool CD Search, which is a part of the Conserved Domain Database (CDD). The amino acid sequence of the target protein is input in the FASTA format or as the PDB ID. The software performs alignment with the help of BLAST and CLUSTAL and provides the ho-

Table 2	ERβ	interacting	proteins
---------	-----	-------------	----------

S. No.	Name of the protein	Tissue specificity	Disease correlation
1	NCoA3	Heart, skeletal muscle, pancreas, and placenta	Breast cancer
2	PELP1	Breast cancer cells	Endometrial cancer, breast cancer
3	TRIM24	Thyroid, mammary gland	Breast cancer, myeloid leukemia
4	ERAP140	Brain, mammary gland, ovary, uterus, stomach, bladder	Breast cancer, ovarian cancer
5	AIB-1	Heart, skeletal muscle, pancreas, and placenta	Polycystic ovarian syndrome, ovarian cancer, breast cancer
6	TRKA	Neuronal cells	Congenital insensitivity to pain with anhidrosis
7	CREB	Brain	Spina bifida, neuroblastoma
8	GRIP1	Muscle, neural tissue, mammary gland	Breast cancer
9	MED1	Ubiquitously present	Lung adenocarcinoma, crohn's disease

mology model based on the entries. It displays a schematic structure of the conserved domains and gives the gi number of the proteins that share homology with these domains.

Results

List of ER α and ER β Interacting Proteins Using NCBI, STRING and iHOP

A list of the interacting proteins was generated (Tables 1, 2; Fig. 1, 2), based on the literature of the predicted as well as experimentally derived proteins. NCBI, STRING and iHOP maintain a pool of information on the proteins found in the primary protein databases such as the PDB and SWISS PROT and so on.

Secondary Structure of the Interacting Proteins Contain Alpha Helices

The secondary structures predicted using PSIPRED showed the presence of alpha helices in the AF1- and AF2-binding regions of the protein. These alpha helices were found to be rich in leucine residues. The leucine residues are known to form a hydrophobic region that fits into the hydrophobic groove of the receptor (Fig. 3, 4). Biochemical and structural studies clearly showed that the LBDs of the NRs recognize a short alpha helix present in the co-activators. Apart from alpha helices, random coils were also predicted. In addition, the secondary structures of AREG and TRKA showed the presence of beta turns in their interacting regions (Fig. 3). TRKA is widely reported in neuronal cell cultures and have been shown to be downregulated in case of anhidrosis and congenital insensitivity to pain [16]. AREG, on the other hand, has been found to be an essential regulator of ERa in the mammary gland and its dysregulation has been known to cause breast cancer [17].

Homology Prediction Using CDD

The conserved domains were determined using CD SEARCH. The conserved domain of the NCoR-1 and PELP1 molecule was shown in Figure 5. The NCoR-1 conserved domain structure acts as a template for a wide variety of coregulators. PELP1 was chosen as the other template, since it contains the RIX1 superfamily conserved domain, which is absent in the SRC-1 domain structure but is a part of many co-regulators like SCAF. The characteristic feature of any co-activator viz SRC domain (basically contains the NR box), NR coactivators domain (forms an alpha helical structure) and the helixloop-helix domain that are specifically found in transcription factors were present. Further, proteins were found, which share a similar homology pattern with the NCOR1. Some of the proteins that were determined were NCoA3 (gi 82112094), SCAF14542 (gi 82267094), novel protein similar to vertebrate NCoA1 (gi 82196443) and the others (Fig. 6, 7). Proteins that were found to be homologous with PELP1 were modulators of non-genomic activity of ER (gi 82072956), At1g30240 (gi 122236349), Acetyltransferase catalyses diacyl glycerol esterification (gi 254569024; Fig. 7). There exists a great amount of homology between the interacting proteins. Some of the proteins having homology are enlisted in Table 3.







Fig. 2. Interacting proteins of $ER\beta$ by STRING tool.

Discussion

Usually, bioinformatics prediction tools, such as the CDD, PSIPRED and STRING, help in utilizing information from different proteins as templates and derive the structure and function of the newly discovered protein. In the present study, the classical example of the coregu-

 $ER\alpha\mathchar`-$ and $\beta\mathchar`-$ Interacting Proteins Contain Consensus Secondary Structures

lators, that is, the members of the p160 family of coregulators were used as template to compare with predicted coregulators enlisted using NCBI, STRING and iHOP. The secondary structure of the enlisted proteins share structural homology with the p160 family of proteins. The AF1 and AF2 interacting domains of the coregulators show the presence of alpha helices, which are con-



RIP140 (78-303 aa)

SMRT (first 289 aa)



Fig. 3. Secondary structures of ERa interacting proteins; PNRC (a), RIP140 (b), SMRT (c), AREG (d) and TRKA (e).





Fig. 4. Secondary structures of ER β interacting proteins; ERAP140 (a), SIN3A (b), PELP1 (c) and GIOT-4 (d).

 $ER\alpha\mathchar`-$ and $\beta\mathchar`-$ Interacting Proteins Contain Consensus Secondary Structures

Conserved domains on Igil6754800/refNP 0350	011 11		View concise res	ult 🛛 🖸
nuclear receptor coactivator 1 [Mus musculus]	,			
Graphical summary show options »				?
Query seq	750	1000	1250 14	15
Specific hits HLH PAS Non-specific HLH PAS Superfanilies HLH su PAS superf Multi-domains PAS	SRC-1 SRC-1 su	Nuc_	DUF15 DUF	
• • • • • • • • • • • • • • • • • • •				÷.
	NCoR-1			
Conserved domains on [gi]21426922[gb]AAC17	708.2]		View full result	?
Graphical summary show options »				?
Query seq. Superfamilies (RIX1 superfamily)	00 750 FUC202	<pre></pre>	1284	
<				
	PELP1			

Fig. 5. Conserved domains of NCoR-1 and PELP1.

query	8 SH	KLVOLLTTTAEOO	[10] KDVT. [12]	1.SSGGS	CPSSHS	SLTERHKILHRLLOEG	S 76
	634 SH	KLVOLLTTTAEOO	[10] . KDVL. [6]	1.SSASS. [6	L.CPSSHS	SLTERHKILHRLLOEG	5 702
gi 82266828	670 HT	KLLOLLTRKPEPL.	[12].KDOP.[5]	I.AGHNG	GNAHAT	SLKEKHKILHOLLONN.	[11.5 734
gi 82112094	627 HT	KLLOLLTTKTEPT.	[15].KDGG.[6]	L.GSOGG	NGSHAT	SLKEKHKILHBLLONS	[1].5 695
gi 82267094	586 HT	KLLOLLTTKADPS.	[15].KDOL.[5]	I.VDGAG	HNNNPT	SLKEKHKILHRLLOTS.	[1].5 653
gi 13626594	638 OT	KLLOLLTTKSDOM.	[13].KDST.[1]	1.SLPGS	GSTHGT	SLKEKHKILHRLLODS.	[1].5 699
gi 13124365	624 00	KLLKLLTTKSDOM.	[14].KDSL.[4]	I.SNSMS	ASAHGT	SLKEKHKILHRLLODS.	[1].5 689
gi 82196443	540 NO. [4]	1.KLSOLLDGASAPE	ODSO.[3]	I.SPTSO	CPASHS	SLTERHKILHRLLODN	S 593
gi 158518533	627 SQ. [3]	.KLVOLLTTTAEOO.	[10].KDVL.[12]	1.SSGGS	CPSSHS	SLTERHKILHRLLOEG	S 698
gi 82194874	629 SQ. [4]	.KLVQLLASTAEQQ.	[10].KDSL.[13]].PPSST	CPSSHS	SLTERHKILHRLLQEG	S 702
gi 23396770	616 HK	KLLQLLTCFTEER.	[12].KDSS.[9]	1.SSTSI	GVSSTS.[5	.MLQEKHRILHKLLONG.	[1].5 689
gi 82268011	151 HK	KLLQLLTSPTDEL.	[17].KDGP.[13]].SSTGG	NHSTVS.[11	.SLQEKHKILHKLLQNG.	[1].T 239
2000 0		T (2) F 86					
ZERG_C	702 DEDITI	[J].E 00					
di esseesse	735 PUDIA	E.[3] E 744					
gi 82112004	FOE DUDIN	X.[3].E 744					
gi 82267094	654 DVELAR	K.[3] F 663					
gi 12626504	700 PUDIAL	X.[3].E 003					
gi 13020394	600 PVDLAP	X.[3].E 709					
gi 92106443	594 PADAGE	X.[3].E 635					
gi 150510522	600 DEDITI	T (2) F 709					
gi 130310333	703 PSDITI	T (3) F 712					
gi 23396770	690 PAEVAI	(3) F 699					
gi 82268011	240 PDEVAL	R [3] F 249					
g1 01200011							
			jues proteins of	the SRC-1 c	onserved dom	ain	
2C52 B	5 EGRNDEK	ALLEQLVSFLSGKD. [2].ELAELDRALGI	DKLV Q	.[8]. 51		
2C52_B query	5 EGRNDEK	ALLEQLVSFLSGKD.[2 ALLEQLVSFLSGKD.[2].ELAELDRALGI	DKLV Q	.[8]. 51 .[8]. 976		
2C52_B query gi 82266828 1	5 EGRNDEK 930 EGRNDEK 1093 ANGVDEG	ALLEQLVSFLSGKD.[2 ALLEQLVSFLSGKD.[2 ALMSOLYTALKDFN	<pre>?].ELAELDRALGI ?].ELAELDRALGI GMEELDRALGI</pre>	DKLV Q DKLV Q PALV.[1].0	.[8]. 51 .[8]. 976 .[8]. 113 8		
2C52_B query gi 82266828 1 gi 82112094 1	5 EGRNDEK 930 EGRNDEK 093 ANGVDEG	ALLEQLVSFLSGKD. [2 ALLEQLVSFLSGKD. [2 ALMSQLYTALKDFN ALLSOLYSALKDFD	2].ELAELDRALGI 2].ELAELDRALGI GMEELDRALGI GLEELDRALGI	DKLV Q DKLV Q PALV.[1].Q	.[8]. 51 .[8]. 976 .[8]. 1138		
2C52_B query gi 82266828 1 gi 82112094 1 gi 13626584 1	5 EGRNDEK 930 EGRNDEK 1093 ANGVDEG 1061 EGPADEG	ALLEQLVSFLSGKD. [1 ALLEQLVSFLSGKD. [1 ALMSQLYTALKDFN ALLSQLYSALKDFD	2].ELAELDRALGI 3].ELAELDRALGI GMEELDRALGI GLEEIDRALGI GLEEIDRALGI	DKLV Q DKLV Q PALV.[1].Q PALV.[1].Q	.[8]. 51 .[8]. 976 .[8]. 1138 .[8]. 1106		
2C52_B query gi 82266828 1 gi 82112094 1 gi 13626594 1	5 EGRNDEKJ 930 EGRNDEKJ 1093 ANGVDEGJ 1061 EGPADEGJ 1071 ESPSDEGJ	ALLEQLVSFLSGKD. [1 ALLEQLVSFLSGKD. [1 ALMSQLYTALKDFN ALLSQLYSALKDFD ALLDQLYLALRNFD	<pre>2].ELAELDRALGI 2].ELAELDRALGI GMEELDRALGI GLEEIDRALGI GLEEIDRALGI CLEEIDR</pre>	DKLV Q DKLV Q PALV.[1].Q PALV.[1].Q PELV.[1].Q	.[8]. 51 .[8]. 976 .[8]. 1138 .[8]. 1106 .[8]. 1116		
2C52_B query gi 82266828 1 gi 82112094 1 gi 13626594 1 gi 82267094 1	5 EGRNDEK 930 EGRNDEK 1093 ANGVDEG 1061 EGPADEG 1071 ESPSDEG 1009 ESQPDEG	ALLEQLVSFLSGKD. [1 ALLEQLVSFLSGKD. [2 ALMSQLYTALKDFN ALLSQLYSALKDFD ALLDQLYLALRNFD ALLDQLCSVLKDYE	2].ELAELDRALGI GMEELDRALGI GLEEIDRALGI GLEEIDRALGI GLEEIDRALGI GLEEIDRALGI	DKLV Q DKLV Q PALV. [1].Q PALV. [1].Q PELV. [1].Q PTLA. [1].Q	.[8]. 51 .[8]. 976 .[8]. 1138 .[8]. 1106 .[8]. 1116 .[8]. 1054		
2C52_B query gi 82266828 1 gi 82112094 1 gi 13626594 1 gi 82267094 1 gi 82078550	5 EGRNDEK 930 EGRNDEK 1093 ANGVDEG 1061 EGPADEG 1071 ESPSDEG 109 ESQPDEG 18 EGRNDEK	ALLEQLVSFLSGKD. [/ ALLEQLVSFLSGKD. [/ ALLSQLYTALKDFN ALLSQLYAALKDFD ALLDQLYLALRNFD ALLDQLCSVLKDYE ALLEQLVTFLSHTD. [/	<pre>2].ELAELDRALGI ?].ELAELDRALGI GMEELDRALGI GLEEIDRALGI GLEEIDRALGI GLEEIDKMLGI ?].ELAELDRALGI</pre>	DKLV Q DKLV Q PALV.[1].Q P PALV.[1].Q P PLV.[1].Q P PLV.[1].Q Q DKLV Q	.[8]. 51 .[8]. 976 .[8]. 1138 .[8]. 1106 .[8]. 1116 .[8]. 1054 .[5]. 61		
2C52_B query gi 82266828 1 gi 82112094 1 gi 13626594 1 gi 82267094 1 gi 82078550 gi 82251966	5 EGRNDEK 930 EGRNDEK 1093 ANGVDEG 1061 EGPADEG 1071 ESPSDEG 1009 ESQPDEG 18 EGRNDEK 735 EARNDEK	ALLEQLVSFLSGKD. [ALLEQLVSFLSGKD. [ALMSQLYTALKDFN ALLSQLYSALKDFD ALLQQLYLALRNFD ALLNQLCSVLKDYE ALLEQLVTFLSHTD. [ALLEQLVTFLSGTD. [ALLEQLVTFLSGTD. [2].ELAELDRALGI GMEELDRALGI GLEEIDRALGI GLEEIDRALGI GLEEIDRALGI SLEEIDKMLGI !.ELAELDRALGI !.ELAELDRALGI	DKLV Q DKLV Q PALV.[1].Q P PALV.[1].Q P PELV.[1].Q P PTLA.[1].Q D DKIV Q DKIV Q	.[8]. 51 .[8]. 976 .[8]. 1138 .[8]. 1106 .[8]. 1116 .[8]. 1054 .[5]. 61 .[8]. 781		
2C52_B query gi 82266828 1 gi 82112094 1 gi 82267094 1 gi 82078550 gi 82251966 gi 54036169	5 EGRNDEK 930 EGRNDEK 1093 ANGVDEG 1061 EGPADEG 1071 ESPSDEG 1009 ESQPDEG 18 EGRNDEK 735 EARNDEK 930 EGRNDEK	ALLEQLVSFLSGKD. [ALLEQLVSFLSGKD. [ALLSQLYTALKDFN ALLSQLYSALKDFD ALLDQLYLALRNFD ALLNQLCSVLKDYE ALLEQLVTFLSHTD. [ALLEQLVTFLSGTD. [ALLEQLVSFLSGKD. [<pre>2].ELAELDRALGI 2].ELAELDRALGI GLEEIDRALGI GLEEIDRALGI GLEEIDRALGI 3].ELAELDRALGI 2].ELAELDRALGI 2].ELAELDRALGI 3].ELAELDRALGI</pre>	DKLV Q DKLV Q PALV.[1].Q PALV.[1].Q PELV.[1].Q PTLA.[1].Q DKLV Q DKLV Q	.[8]. 51 .[8]. 976 .[8]. 1138 .[8]. 1106 .[8]. 1116 .[8]. 1054 .[5]. 61 .[8]. 781 .[8]. 976		
2C52_B query gi 82266828 1 gi 82112094 1 gi 82267094 1 gi 82267094 1 gi 82078550 gi 82251966 gi 54036169 gi 82268011	5 EGRNDEK 930 EGRNDEK 1093 ANGVDEG 1061 ESPADEG 1071 ESPSDEG 109 ESQPDEG 18 EGRNDEK 735 EARNDEK 930 EGRNDEK 610 EGQSDER	ALLEQLVSFLSGKD. [/ ALLEQLVSFLSGKD. [/ ALLSQLYTALKDFN ALLSQLYLALKDFD ALLDQLYLALKNFD ALLQLVIFLSHTD. [/ ALLEQLVTFLSHTD. [/ ALLEQLVTFLSGTD. [/ ALLEQLVSFLSGKD. [/ ALLDQLDSLLNTAD. [/	2] .ELAELDRALGI 2] .ELAELDRALGI GMEELDRALGI GLEEIDRALGI GLEEIDRALGI 2] .ELAELDRALGI 2] .ELAELDRALGI 2] .ELAELDRALGI 2] .ELAELDRALGI 2] .ALQEIDRALGI	DKLV Q DKLV Q PALV.[1].Q PALV.[1].Q PELV.[1].Q PTLA.[1].Q DKLV Q DKLV Q DKLV Q DKLV Q PEIV.[1].Q PEIV.[1].Q	.[8]. 51 .[8]. 976 .[8]. 1138 .[8]. 1106 .[8]. 1116 .[8]. 1054 .[5]. 61 .[8]. 781 .[8]. 976 .[8]. 657		
2C52_B query gi 82266828 1 gi 82112094 1 gi 13626594 1 gi 82078550 gi 82267094 1 gi 82078550 gi 82251966 gi 54036169 gi 82260825 1 gi 82260825 1	5 EGRNDEK 930 EGRNDEK 1093 ANGVDEG 1061 EGPADEG 1009 ESQPDEG 18 EGRNDEK 930 EGRNDEK 610 EGQSDER 216 EGQSDER	ALLEQLVSFLSGKD. [ALLEQLVSFLSGKD. [ALMSQLYTALKDFN ALLSQLYSALKDFD ALLOQLYLALRNFD ALLQLCSVLKDYE ALLEQLVTFLSGTD. [ALLEQLVTFLSGTD. [ALLEQLVSFLSGKD. [ALLEQLDSLLNTAD. [ALLDQLDSLLNNTD. [2) .ELAELDRALGI 2) .ELAELDRALGI GMEELDRALGI GLEEIDRALGI GLEEIDRALGI 2) .ELAELDRALGI 2) .ELAELDRALGI 2) .ELAELDRALGI 2) .ALQEIDRALGI 2) .ALQEIDRALGI	DKLV Q DKLV Q PALV.[1].Q PALV.[1].Q PELV.[1].Q DKLV DKLV Q PEIV.[1].Q PEIV.[1].Q	.[8], 51 .[8], 976 .[8], 1138 .[8], 1106 .[8], 1116 .[8], 1054 .[5], 61 .[8], 781 .[8], 976 .[8], 657 .[8], 1263		
2C52_B query g1 82266828 1 g1 82112094 1 g1 13626594 1 g1 82078550 g1 82078550 g1 8227856 g1 54036169 g1 82268011 g1 82268011 g1 82268025 1 g1 23396768 1	5 EGRNDEK 930 EGRNDEK 1093 ANGVDEG 1061 EGPADEG 1071 ESPSDEG 18 EGRNDEK 930 EGRNDEK 610 EGQSDER 216 EGQSDER 52 EGQSDER	ALLEQLVSFLSGKD. [ALLEQLVSFLSGKD. [ALLSQLYSALKDFD ALLSQLYSALKDFD ALLDQLYALRNFD ALLQLCSVLKDYE ALLEQLVTFLSHTD. [ALLEQLVTFLSGTD. [ALLEQLVSFLSGKD. [ALLEQLSLLNTAD. [ALLDQLDSLLNTD. [ALLDQLDSLLNTD. [2	2) ELAELDRALGI GMEELDRALGI GLEEIDRALGI GLEEIDRALGI GLEEIDRALGI 2) ELAELDRALGI 2) ELAELDRALGI 2) ELAELDRALGI 2) ALQEIDRALGI 2) ALEEIDRALGI 2) ALEEIDRALGI	DKLV Q DKLV Q PALV.[1].Q PALV.[1].Q PELV.[1].Q PTLA.[1].Q DKLV Q DELV.[1].Q PELV.[1].Q	.[8]. 51 .[8]. 976 .[8]. 1138 .[8]. 1106 .[8]. 1116 .[8]. 1054 .[5]. 61 .[8]. 781 .[8]. 976 .[8]. 657 .[8]. 1263 .[8]. 1099		
2C52_B query g1 82266828 1 g1 82112094 1 g1 82267094 1 g1 82267094 1 g1 82267094 1 g1 82267094 1 g1 82268011 g1 82268011 g1 82260825 1 g1 23396768 1	5 EGRNDEK 930 EGRNDEK 1093 ANGVDEG 1061 EGPADEG 1071 ESPSDEG 1071 ESPSDEG 108 EGRNDEK 930 EGRNDEK 610 EGQSDER 216 EGQSDER 052 EGQSDER	ALLEQLVSFLSGKD. [ALLEQLVSFLSGKD. [ALLSQLYTALKDFN ALLSQLYTALKDFD ALLDQLYLALRNFD ALLDQLYLALRNFD ALLEQLVTFLSHTD. [ALLEQLVFFLSGKD. [ALLEQLVSFLSGKD. [ALLEQLDSLLNTAD. [ALLDQLDSLLNNTD. [ALLDQLDSLLNNTD. [ALLDQLHFFLSNTD. [2] ELAELDRALGI 2] ELAELDRALGI GLEEIDRALGI GLEEIDRALGI GLEEIDRALGI 2] ELAELDRALGI 2] ELAELDRALGI 2] ELAELDRALGI 2] ELAELDRALGI 2] ALQEIDRALGI 2] ALEEIDRALGI 2] GLEEIDRALGI	DKLV Q DKLV Q PALV.[1].Q PALV.[1].Q PELV.[1].Q PELV.[1].Q DKLV Q DKLV Q DKLV Q DKLV Q DKLV Q DKLV Q PELV.[1].Q PELV.[1].Q	.[8]. 51 .[8]. 976 .[8]. 1138 .[8]. 1106 .[8]. 1106 .[8]. 1054 .[5]. 61 .[8]. 781 .[8]. 781 .[8]. 976 .[8]. 657 .[8]. 1263 .[8]. 1099		

Fig. 6. Homologous proteins of NCoR-1.

Color version available online

27 .[2].SLAASITECGELLSA.[12].H gi 74597564 S.[3].HKLVTRISSLLQ.[3].FEGRWAAVVLVKALVEA.[3].EI 97 66 .[2].GLMCLLRLHGSVGGA.[1].N L. [3].GALVSLSNARLS. [1].IKTRFEGLCLLSLLVGE. [3].EL 123 query gi 74583175 31 .[2].QIADILVHQKCIERL.[5].A L. [3].KNWCTSLTKMLQ. [3].FRIRWSAIILIHCTISQ. [2].DC 93 gi 122236349 41 .[2].KVVSTISTHKLLSES.[12].S K. [3]. DDWVARLSALIS. [3]. PDKSWVGICLIGVTCQE. [3]. DR 111 18 .[2].EAIRGLREHGAFRGE.[1].L P.[3].SGLLSSCNSRLT.[1].ASSRIEGLSLLALAVEE.[3].DV 75 gi 82072956 gi 296809173 26 .[2].YIATTISECSVILSA.[1].T.[11].A.[3].QKLKARITSLLQ.[3].VEGRWTGVVLVKSMVES.[3].EI 96 gi 238881351 17 .[2].PLLSILHNDKQILST.[1].T K.[3].NHLISRTLNLVR.[3].VYNKWCGINLIRVIVEE.[1].SI 74 gi 260945661 19 .[2].PILSSLHNQRHSLST.[1].S K. [3].KHLTSRTLNLCR. [3].PYNVWCGVNLIYVIIDN. [1].LI 76 gi 254569024 25 .[2].EILPILADENAVNEA S K. [3].THVTSRAANYLR. [3].SEVRWFGTKLVHVLCLH. [1].RI 81 gi 260796465 18 .[2].WAIEVANEHQLLQSE.[1].N QDWVSHINTSLG.[1].AKTRLEGLCLLGTVVQQ.[3].GT 72 А gi 74597564 98 I RGSEPFVRGLMSILSK SD PASTKTMAVITLTRIFHLTYQYPT. [1].VREITTPS. [1].PGF 153 124 F.[1].QHCVSWLRSIQQVLQT PPATMELAVAVLRDLLRYAAQLPA. [1].FRDISMNH OD T.PG 179 guery gi 74583175 TPKTLEIAMITVSKMFSSTVGRPA.[1].TRELVTPN.[1].PTF 150 94 L.[1].EHGATWAKLLIALLNR PE gi 122236349 112 F.[1].KSYSVWFNSLLSHLKN.[1].AS SRIVRVASCISISDULTRUSRESN TKKDAVSH.[1].SKL 168 gi 82072956 76 F.[1].QHCVSWLRSLLQIIQS QD PPRVVSLAVFVLRSLLAHSSALPE.[1].SREISTNH.[1].PGL 132 gi 296809173 97 L RGCETWARSILAILGK PD PLSLKRLSILCLTRIFQLTHPYPT. [1].TREITTPT. [1].PQF 152 gi 238881351 75 L.[1].SEGNNLMNSLLQVLEA.[16].TI DLKILTSCIETINYMGDLIRGKPT. [1] .TREILTPK. [1] .NSI 147 gi 260945661 77 L.[1].SEGSSFFSQLLKVLES PR. [2] .DPRVFKSVVDCINKLCKNIRGKPT. [1] .TREVLTPN. [1] .SSV 135 QMTCLESAISTLDFIMDKIRHKVA. [1].TREILTPR. [1].PSI 147 82 I.[1].SDAGNYISALTKIIES.[9].VR gi 254569024 gi 260796465 73 F. [1]. QHGTTWIRMLTQVLQA YD SPLTLOMASHVLGSVVQQAAQYPE.[1].AREVATTH.[1].PTL 129 PELP1 homologous proteins of the RIX1 conserved domain gi 82072956 514 PYISSDCR. [2]. LYRLLLCLTLT. [10].C. [7].GTTEESLQVSR. [1].STEALAI. [2]. ILIHPRVPS L 583 SGEALLT 568 PYSTVRTK VYAILELWVQV GASAGMLQGGA HLLSDISPP.[15].S 630 query С gi 115502554 569 PYTSSRCR. [2].LYCLLLALLLA. [10].C. [7].GQREDSLEVSS. [1].CSEALVT. [2].ALTHPRVPP 638 С gi 82072956 378 PYSSLRCS VYRVLETWVTT GISSGVLQGPM. [1].HSDILLA NLLSDITPP.[8].F 434 gi 74705525 392 PYSTVRTK VYAILELWVQV С GASAGMLQGGA SGEALLT HLLSDISPP.[15].S 454 gi 82072956 584 QRPL 587 query 631 LQTG 634 gi 115502554 639 QPMG 642 gi 82072956 435 VQLG 438 gi 74705525 455 LQTG 458 PELP1 homologous proteins of the nuc202 conserved domain.

Fig. 7. Homologous proteins of PELP1.

Table 3.	ERa and ER	β interacting	proteins b	v STRING	(low confidence)	level)
14010 01	Dica and Dic	p miter acting	proteino o	,	(10 m connactice	10,01)

ERα						
SRC	NRIP 1	ERBB2	SMAD4	HDAC4	CREBBT	UBC
NCoA1	PELP1	MTA1	ARNT	TP53	CDH1	TRIM24
BRCA1	AKT 1	TFF1	STAT54	JUN	FOX03	
EP300	NROB2	XBP1	CEBPB	MED1	FOXA1	
NCOA3	MYC	IGF1	HDAC9	ESR2	NCOA6	
SP1	HSP90AA1	BCAR1	SAFB	CTSD	SMARCA4	
NCOA2	AHR	EGFR	NR2F1	NCOR2	PRDM2	
IGF1R	CCND1	PIK3R1	GRIP1	CAV1	KAT2B	
ERβ						
SRC	MAPK11	THRB	RARA	TCF19	RORC	
ESR1	AREG	NR1D1	NR4A2	HNF4G	NR2F1	
NCOR2	AHR	VDR	PPARA	RDRB	PPARD	
NCOA3	IGF1	PELP1	NR2 C2	GNA14	NRID2	
AKT1	NCOA1	ESRRB	RARG	NR2E1	GNA15	
NROB2	THRA	PPARG	RARB	GNG2		
ERBB2	OXT	ESRRA	RXRG	GNAL		
EGF	NR5A1	ESRRG	NR2C1	NR2C2AP		

 $ER\alpha\mathchar`-$ and $\beta\mathchar`-$ Interacting Proteins Contain Consensus Secondary Structures

Ann Neurosci 2018;25:1–10 DOI: 10.1159/000481809 Color version available online

served domains. Some of the interacting proteins, such as TRKA and AREG, possessing beta turn in their interacting domains suggesting that, apart from alpha helices, other secondary structures play an important role in gene regulation.

ERs recruit a host of coregulators for regulating estrogen-dependent gene expression [2]. The level of estradiol decreases during aging in both genders and a sharp decline is reported in females after their menopause [18]. Thus, the optimum level of estrogen, receptors in the normal form and binding of ERs with coregulators are vital factors to decrease risk factors of neurodegenerative diseases.

Here, our study shows that majority of the interacting proteins contain α helices. Their presence may induce the aggregation properties of proteins as shown in earlier studies. The study shows that the intrinsic stabilization of α helix proteins can trigger the formation of α helical aggregates [19]. Further, higher concentration helix-promoting trifluoroethanol induces stabilization of tau protein aggregates [20], a known factor in case of Alzheimer disease. In addition, the microtubule-binding domain of tau, which constitutes the core of paired helical filaments, has significant α helix propensity [21]. Taking together,

References

- Mauvais-Jarvis F, Clegg DJ, Hevener AL: The role of estrogens in control of energy balance and glucose homeostasis. Endocr Rev 2013; 34:309–338.
- 2 Thakur MK, Paramanik V: Role of steroid hormone co-regulators in health and disease. Horm Res 2009;71:194–200.
- 3 Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schutz G., Umesono K, Blumberg B, Kastner P, Mark M, Chambon P, Evans RM: The nuclear receptor superfamily: the second decade. Cell 1995;83:835–839.
- 4 Huang P, Chandra V, Rastinejad F: Structural overview of the nuclear receptor superfamily: insights into physiology and therapeutics. Annu Rev Physiol 2010;72:247–272.
- 5 Couse JF, Lindzey J, Grandien K, Gustafsson JA, Korach KS: Tissue distribution and quantitative analysis of estrogen receptor-alpha (ERalpha) and estrogen receptor-beta (ERbeta) messenger ribonucleic acid in the wild-type and ERalpha-knockout mouse. Endocrinology 1997;138:4613–4621.
- 6 Beato M: Gene regulation by steroid hormones. Cell 1989;56:335-344.
- 7 Meijer OC, Steenbergen PJ, De Kloet ER: Differential expression and regional distribution of steroid receptor coactivators SRC-1 and SRC-2 in brain and pituitary. Endocrinology 2000;141:2192–2199.

- 8 Tetel MJ: Nuclear receptor coactivators in neuroendocrine function. J Neuroendocrinol 2000;12:927–932.
- 9 Paramanik V, Thakur MK: AIB1 shows variation in interaction with $ER\beta TAD$ and expression as a function of age in mouse brain. Biogerontology 2011;12:321–328.
- 10 Kadam S, Everson BM: Transcriptional specificity of human SWI/SNF BRG1 and BRM chromatin remodeling complexes. Mol Cell 2003;11:377–389.
- 11 Ghosh S, Thakur MK: Interaction of estrogen receptor-alpha transactivation domain with nuclear proteins of mouse brain: p68 RNA helicase shows age- and sex-specific change. J Neurosci Res 2009;87:1323–1328.
- 12 Paramanik V, Thakur MK: Interaction of estrogen receptor associated protein (ERAP) 140 with ER beta decreases but its expression increases in aging mouse cerebral cortex. Cell Mol Neurobiol 2010;6:961– 966.
- 13 Lionder K, Söderström M: Functional analyses of an LXXLL motif in nuclear receptor corepressor (N-CoR). J Steroid Biochem Mol Biol 2004;91:191–196.
- 14 Heery DM, Kalkhoven E, Hoare S, Parker MG: A signature motif in transcriptional coactivators mediates binding to nuclear receptors. Nature 1997;387:733–736.

the present work may provide an insight into the structural abnormalities leading to functional dysregulation and can be an excellent target for NR pharmacology studies.

Acknowledgements

M.K.T. acknowledged the Department of Biotechnology, Government of India for financial support. H.K. acknowledged the Indian Academy of Sciences for providing Summer Research Fellowship.

Disclosure Statement

The source of funding is Department of Biotechnology, Government of India to MKT. There are no conflicts of interest among authors.

Authors Contribution

V.P. designed, performed the work and analyzed the data. M.K.T. designed the work and analyzed the data. H.K. performed the work.

- 15 Hall JM, Mcdonnell DP: Coregulators in nuclear estrogen receptor action: from concept to therapeutic targeting. Molecular Interv 2005;5:343–357.
- 16 Greco A, Villa R, Romano L, Penso D, Pierotti MA: A novel NTRK1 mutation associated with congenital insensitivity to pain with anhidrosis. Am J Hum Genet 1999;64:1207–1210.
- 17 Ciarloni L, Mallepell S, Brisken C: Amphiregulin is an essential mediator of estrogen receptor alpha function in mammary gland development. Proc Natl Acad Sci U S A 2007; 104:5455–5460.
- 18 Sharma PK, Thakur MK: Expression of estrogen receptor (ER) alpha and beta in mouse cerebral cortex: effect of age, sex and gonadal steroids. Neurobiol Aging 2006;27:880–887.
- 19 Kunjithapatham R, Oliva FY, Doshi U, Perez M, Avila J, Munoz V: Role for the alpha-helix in aberrant protein aggregation. Biochemistry 2005;44:149–156.
- 20 Luo PZ, Baldwin RL: Mechanism of helix induction by trifluoroethanol: a framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. Biochemistry 1997;36:8413–8421.
- 21 Mandelkow EV, Mandelkow E: Biochemistry and cell biology of tau protein in neurofibrillary degeneration. Cold Spring Harb Perspect Med 2012;2:a006247.