Investigation of the PGC-1-alpha Co-activator Interacting with the Human Estrogen Receptor-alpha

Rachael Beach

Department of Chemistry
The University of North Carolina at Asheville
One University Heights
Asheville, North Carolina 28804 USA

Faculty Advisor: Noah W. Allen, III, Ph.D.

Abstract

The development of cancer in breast, thyroid, and uterine tissue is largely due to the uncontrolled expression of the estrogen receptor (ER). A well-founded understanding of the various factors and their interplay in the regulation of ER activity is imperative for the development of treatment and prevention strategies for these cancers. This research proposes to focus on the interactions between ER alpha and the co-activator PGC-1-alpha. The PGC-1-alpha actions are less sensitive towards the nature of the bound ligand in the ER, whereas classical co-activators interact with the ER in a highly ligand dependent fashion. Additionally, cases have been found where PGC-1-alpha interacts with ER in the absence of a ligand suggesting a binding mode, diverging from that of classical co-activators. Potential co-activator binding interfaces on the ligand binding domain (LBD), DNA binding domain (DBD), and hinge region of the ER alpha are of particular interest here. However, the tertiary structure of PGC-1-alpha is largely unknown. Homology modeling and tertiary structure prediction methods had to be applied to get reasonable and realistic tertiary structure predictions of PGC-1-alpha. Single point mutations and changes in rigidity of these interfaces as well as of the complementary interfaces on the co-activator are performed systematically. Interactions are simulated and analyzed, utilizing computational methods such as molecular dynamics, geometry optimization, and visualization. The resulting data should then provide insight into the possible binding modes of the PGC-1-alpha co-activator to the ER alpha. Furthermore, a basis will be established for future in vitro studies.

Keywords: Estrogen Receptor-alpha, DNA binding domain, Ligand binding domain, Hinge region, nuclear receptors, co-activators, structure prediction, molecular dynamics studies, PGC-1, SRC-1

1. Introduction

Nuclear receptors are proteins that are ligand and gene activated. They act as intracellular receptors for hydrophobic molecules such as steroids. These receptor-ligand complexes act as transcription factors in the nucleus. Estrogen receptors are members of this nuclear receptor family and classically bind the steroid estradiol and in turn act as transcription factors for regulating the reading of DNA and thereby the production of proteins. The binding complex between estrogen and the ER doesn’t occur solely between the two, but rather other proteins called co-activators or co-repressors are needed to help the binding between the two. The binding of estrogen to the ER exists in complex with co-activators as well. Where the problem occurs in this system of the ER, Estrogen, and co-activators is when the ER is deregulated, meaning its operation and, in turn, production of proteins is left uncontrolled, thereby causing cancer.

Co-activators or co-repressors are auxiliary proteins that, classically, are not a part of DNA binding regulation or transcriptional machinery. Co-activators do however play a vital role within the nuclear receptor-signaling pathway. The classical co-activators of ER-alpha consist of those from the p160 family of co-activators which consist of the co-activators SRC-1, SRC-2, and SRC-3. The tightness in binding of p160 co-activators in vivo depends not
exclusively on the LxxLL motif, but is apparently induced by the sequences immediately flanking the LxxLL motif. The LxxLL motif is a special sequence of amino acids (leucine, any amino acid, any amino acid, leucine, leucine) that exists on the nuclear receptor box (NR-Box) of co-activators, not the nuclear receptors. The NR-Box is the part of the co-activator that interacts with the nuclear receptor. SRC-1 consists of 1440 amino acids, and among those there are 7 LxxLL motif sites. These motifs on the co-activators recognize a hydrophobic patch on the surface of the LBD whose structure is stabilized by the binding of agonist ligands. Also, the LBD of the ER-alpha is the only part of the receptor that the SRC-1 family of co-activators is known to interact with. This SRC-1 co-activator is found in many different tissues of the body such as the endometrium, heart, brain, liver, skeletal muscle, kidney, pancreas, lung, and placenta. SRC-1 is a co-activator for nuclear receptors like ER, glucocorticoid receptor (GR), retinoid receptors (RAR’s and RXR’s), thyroid hormone receptors (TR), and peroxisome proliferated activated receptors (PPAR).

The members of the PGC-1 family are the PGC-1alpha co-activators and the PGC-1 beta co-activators. Homology comparisons show sequence identity in several distinct domains like 35% at the central regulatory domain, 40% at a N-terminal activation domain and 48% at a C-terminal RNA binding domain. Also, flexibility of the co-activator, PGC-1 in this case, is essential for selectivity, but certain rigidity is needed for interaction stability. PGC-1 is a very versatile co-activator in that it is involved in many other reactions and mechanisms. The mRNA of this co-activator is found in many different tissues of the body such as the skeletal muscle, brown fat, heart, kidney, brain, and liver. These tissues are also rich in mitochondria and are areas that have high energy demands. The implication is that PGC-1 induces the expression of genes that facilitate cells to adjust to metabolic demands. Other than ER, PGC-1 can co-activate in the peroxisome proliferator activated receptor (PPAR-gamma), glucocorticoid receptor (GR), thyroid hormone receptor (TR), Peroxisome proliferator activated receptor-alpha, and the mineralocorticoid receptor (MR). The PGC-1 family of co-activators displays some interesting characteristics with the ER. When compared to the function of SRC-1 co-activators, the amino acids flanking the LxxLL motif on PGC-1-alpha don’t appear to be necessary for recognition. Also, PGC-1-alpha has only three LxxLL motifs whereas SRC-1 has seven. Second, there is interaction with the AF-2 activation region of the ligand binding domain (LBD) of the ER along with the DNA binding domain (DBD) and the hinge region. This is interesting because PGC-1-alpha is known to interact with the LBD, suggesting that the DBD and hinge region of the ER are used to help stabilize the binding of PGC-1-alpha to the LBD of the ER. Lastly, interaction between PGC-1-alpha and ER is possible in a ligand independent manner.

Furthermore, studies using molecular modeling of the interaction between the nuclear receptor (NR) box 2 and 3 peptides with the ER alpha LBD indicate that the charge clamp found in the SRC-1 and ER complexes may not be implicated in the case of PGC-1alpha. This may possibly also account for the fact that compared to the affinity between SRC-1 and ER alpha LBD, the affinity between PGC-1-alpha and ER alpha LBD is lower. In this research, the co-activator PGC-1-alpha is being focused on because cases have been found in which PGC-1-alpha interacts with ER in the absence of a ligand suggesting a binding mode, diverging from that of classical co-activators. Therefore, a strong understanding of the various factors and their interplay in the regulation of ER is imperative for the development of treatment and prevention strategies for cancers such as uterine, thyroid and breast, which are the leading cause of death for women between the ages of 35 and 54.

2. Materials and Methods

Before calculations could be run and decisions where point mutations are going to be made, an extensive primary literature research was conducted to find the primary sequence of PGC-1-alpha so that secondary and then tertiary structures could be found, and in this case, predicted. Typically, tertiary structures of molecules are obtained though crystallography or NMR (nuclear magnetic resonance). These known structures are usually deposited into a data base, such as the RCSB Protein Data Bank online (http://www.rcsb.org). A tertiary structure of PGC-1-alpha via crystallography or NMR does not exist except for 12 residues from amino acids 205-216 (RPASELLKYLTT). This is unfortunate because it is not enough to run any conclusive calculations. Therefore, necessary measures were to be taken to try and predict a realistic and reasonable tertiary structure. The primary structure is known, and was taken from the Human Protein Reference Database (HPRD).

Single point mutations and changes in rigidity of the LBD, DBD, and hinge region interfaces as well as of the complementary interfaces on the co-activator are to be performed systematically. Interactions are to be simulated and analyzed utilizing computational methods such as molecular dynamics, geometry optimization, and visualization. The molecular dynamics simulations software used is AMBER.
2.1. secondary structure prediction

The secondary structure for PGC-1-alpha isn’t known exactly, therefore different secondary structure prediction analysis programs were used and the results evaluated. Although the secondary structure is not necessary for running calculations, it is still good to obtain one in order to get a better idea of the base of the structure and to have something to refer back to if there is a problem in the tertiary structure prediction. The programs used were obtained via links from the ExPASy Proteinomics tools website. The following programs used and compared were Porter, PSLpred, and SOPMA (Figure 1). Porter, a server for secondary structure protein prediction, is an advancement of the SSpro server that is based on 45 bidirectional recurrent neural networks\(^2\). PSLpred is a secondary structure protein prediction method integrating two feed-forward neural networks which perform analysis on output obtained from Position Specific Iterated BLAST\(^3\). Lastly, SOPMA (Self Optimized Prediction Method from Alignment), is another secondary structure protein prediction program by using consensus prediction from multiple alignments\(^4\). These programs were chosen based on their scientific quality and broad acceptance within the scientific community.

<table>
<thead>
<tr>
<th></th>
<th>Alpha helix (Hh) :</th>
<th>3_10 helix (Gg) :</th>
<th>Beta bridge (Bb) :</th>
<th>Extended strand (Ee) :</th>
<th>Beta turn (Tt) :</th>
<th>Bend region (Bz) :</th>
<th>Random coil (Cc) :</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>236 is 29.57%</td>
<td>0 is 0.00%</td>
<td>0 is 0.00%</td>
<td>99 is 12.41%</td>
<td>35 is 4.39%</td>
<td>0 is 0.00%</td>
<td>428 is 53.63%</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1 SOPMA results

2.2. tertiary structure prediction

Three different methods of predicting tertiary structures were used to try and elucidate a realistic PGC-1-alpha tertiary structure. The first was through homology modeling, which is the most accurate approach. This method of predicting three-dimensional structures is based on the coordinates of known homologies found in the Protein Data Base (PDB). The homology program used was ESyPred3D\(^5\). The second way of predicting tertiary protein structures used was through threading, or fold recognition. This method recognizes common fold in proteins having...
basically no homology in sequence to any protein in the PDB. The threading program used was PHYRE (Protein Homology/analogy Recognition Engine).

2.3. protein structure evaluation

The structure predictions are to be evaluated using some different evaluation programs such as Dictionary of Secondary Structure for Proteins (DSSP)\textsuperscript{16}, PROCHECK\textsuperscript{17}, Volume, area, dihedral angle reporter (VADAR)\textsuperscript{18}, and Verify3D\textsuperscript{19}. DSSP identifies and labels seven different kind of secondary structures (alpha helices, \(3/10\) helices, \(\pi\) helices, beta bridges, extended beta strands, hydrogen bonded turns, and bends) by following a very stringent method of identifying hydrogen bonds and hydrogen bonding patterns. To calculate accessible surface area, PROCHECK uses DSSP to identify a secondary structure, and it also calculates torsion angles interatomic distances, bond angles, and other properties. To help evaluate the protein structures, VADAR measures dihedral angles, identifies hydrogen bonds, and measures interatomic. What makes it different is that it uses a more comprehensive approach to identifying secondary structures. It identifies and classifies beta turns, side chain hydrogen bonds (or salt bridges), calculates packing volume, determines exact accessible surface areas, performs packing “defect” checks, performs buried charge evaluation, calculates threading and surface free energies, and determines residue disposition. It then compares several of these values with those that would be anticipated among high quality structures. Lastly, to identify protein structure quality, Verify3D uses a form of three-dimensional threading. It uses a matrix scoring method where the secondary structure and solvent exposure tendency of each of the 20 amino acids is determined statistically from high quality PDB structures\textsuperscript{20}.

3. Conclusion

The results from the secondary protein predictions gave results in the format of each amino acid in the primary sequence of PGC-1-alpha accompanied with a letter that has specific meaning depending on the program used to define the predicted structure. Even though the secondary structure predictions are not directly used for running calculations, they are useful for reference, to construct a more complete structure of the protein.

The tertiary structures predicted from PHYRE and ESyPred3D are pictured below (Figure 2 and Figure 3). These structures are to be compared evaluated using the programs mentioned in the “protein structure evaluation” section, and then used for running calculations.

The specific areas being highly studied are the LxxLL motifs on the PGC-1-alpha co-activator in conjunction with the whole estrogen receptor. Once a tertiary structure of PGC-1-alpha can be completely predicted and verified, some point mutations and changes in rigidity around the LBD, DBD, and hinge region of the estrogen receptor are going to be made, and then molecular dynamic studies via AMBER are to be conducted to study the interaction between the two. The resulting data should provide insight into the possible binding modes of the PGC-1-alpha co-activator to the ER alpha. Furthermore, a basis will be established for future in vitro studies.
4. Acknowledgements

The author would like to thank Dr. Allen, principal investigator, and Gert Kiss, who collaborated throughout the project. The author would also like to thank all the other members of the Allen Laboratory for all their helpful input, guidance, and support on this project. Supportive friends and family of the author, especially James Hollifield, were also invaluable throughout the course of the project and proceedings.

5. References


433