PPRESearch: Peroxisome Proliferator Activator Element Search Database

Gireedhar Venkatachalam 1, Meena Kishore Sakharkar 2,3, Alan Prem Kumar 4,5,*, Marie Veronique Clement 1,6,*

1 Dept. of Biochemistry, YLL School of Medicine, National University of Singapore, Singapore
2 Biomedical Engineering Research Centre, Nanyang Technological University, Singapore
3 Advanced Design and Modeling Lab, Nanyang Technological University, Singapore
4 National University Medical Institutes, YLL School of Medicine, National University of Singapore, Singapore
5 Dept. of Physiology, YLL School of Medicine, National University of Singapore, Singapore
6 NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore

Submitted: 11 Nov. 2009; Accepted: 8 Dec. 2009

Abstract
Peroxisome proliferator activated receptors (PPARs) are ligand-activated transcription factors involved in processes such as angiogenesis, cell cycle, apoptosis, and lipid metabolism thereby having a significant effect on cell fate and cell metabolism. After activation, PPARs heterodimerize with the 9-cis-retinoic acid receptor (RXR) and subsequently bind to DNA on specific response elements termed Peroxisome Proliferator Response Elements (PPREs), located in regulatory regions of target genes, thereby modulating their transcriptional activity. All PPREs identified initially consist of the juxtaposition of two derivatives of the canonical hexamer sequence AGGTCA spaced by one nucleotide, commonly called Direct Repeat 1 (DR1). However, recent published reports demonstrated of a DR2 (spaced by two nucleotides) as a PPAR-responsive element which creates urgency for development of a new PPRESearch database with a capacity to search for both DR1 and DR2 repeats. Our PPRESearch database is based on in vivo and in vitro experimentally derived data for PPAR binding regions. Thus, this database could determine PPREs with high accuracy and reliability. Uniquely, our PPRESearch has the capability to analyze if the putative PPREs could be bound by PPARα, PPARβ/δ, or PPARγ. If all PPAR isoforms could bind to the PPRE sites, our database is then able to predict the most preferred isoform based on binding affinity.

Keywords: PPARs, PPRE, DR1 and DR2, 5'-flanking sequences, PPRE search database.

INTRODUCTION
Peroxisome proliferator-activated receptors (PPARs) are characterized by three general functional domains: the N-terminal domain, the DNA binding domain and the ligand binding domain (Fig. 1). The N-terminal domain serves as key determinant of isotype-selective gene expression. Of all the domains present, the DNA-binding domain is the most conserved domain of nuclear receptors. A crucial role in DNA-binding specificity is played by the P-box amino acids located at the carboxyl end of the first zinc finger. On the other hand, the ligand-binding domain (LBD) of nuclear receptors plays a pivotal role in the transduction of the hormonal signal into transcriptional activation via induction of PPAR binding to specific response elements (PPREs), consisting of a direct repeat hexamer DNA recognition motif spaced by one nucleotide (DR1) (Lemberger et al., 1996).

PPAR activation is potentiated by ligand binding to the PPAR molecule. In the absence of a ligand, PPARs act as transcriptional repressors complexed with repressor proteins. Ligand binding causes a conformational change in the AF2 (activation factor 2) helix, which can then form heterodimeric complex with other transcription factors (e.g., 9-cis-retinoic acid receptor, RXR). Subsequently, various co-activators and co-repressors then fine-tune the ability of the activated PPAR:RXR complex to bind to PPRE of target genes to regulate gene transcription (Fig. 2).

*Corresponding author:
Marie Véronique Clement, Ph.D.
Department of Biochemistry,
Yong Loo Lin School of Medicine,
National University of Singapore,
8 Medical Drive, Singapore 117597
Email: bchmvc@nus.edu.sg; nmiapk@nus.edu.sg
The three PPARs identified are α, δ (also called β, NUC-1 or FAAR) and γ, which constitute a distinct subfamily of the superfamily of nuclear hormone receptors. These PPARs are activated by naturally occurring fatty acids or fatty acid derivatives. While PPARα and β stimulate fatty acid catabolism in tissues known for their high rates of fatty acid oxidation such as liver, heart, kidney, and brown adipose tissue, PPARγ favors fatty acid storage by stimulating triglyceride accumulation in adipocytes. PPARα is found predominantly in skeletal muscle, liver, heart, and kidney whereas PPARβ RNA is more expressed with maximal expression in placenta and skeletal muscle. PPARγ is expressed in the insulin-responsive tissue such as skeletal muscle, heart, and liver (Mukherjee et al., 1997).

The PPARs modulate the expression of numerous target genes that play a central role in regulating glucose, lipid and cholesterol metabolism where imbalances can lead to diabetes, obesity, cardiovascular disease and cancer. This has made PPARs attractive therapeutic targets and the pharmaceutical industry has invested enormous amount of its drug discovery efforts into developing PPAR modulating agents for the treatment of diseases. The fibrate class of hypolipidemic drugs including clofibrate, fenofibrate and bezafibrate exert their actions primarily through activation of PPARα. Clinically, activators of PPARα have been shown to lower serum triglycerides and increase HDL cholesterol (HDL-c) in hyperlipidemic patients (Berger et al., 2005). PPARα expression is also upregulated in human prostate adenocarcinomas (Collett et al., 2000). PPARα ligands suppress the growth of several cancer lines, including colon (Tanaka et al., 2001), endometrial (Saïdi et al., 2006), and breast (Suchanek et al., 2002) and metastatic potential of melanoma cells (Grabacka et al., 2006) in vivo or in vitro.

As the master regulator of fat-cell formation, PPARγ is required for the accumulation of adipose tissue that contributes to obesity (Lehrke et al., 2005). A role for PPARγ in limiting inflammation has also been reported (Hsueh et al., 2004). Clinically relevant ligands for PPARγ comprise TZDs, the insulin-sensitizing drugs licensed for use in patients with Type 2 diabetes mellitus (Krentz et al., 2006). PPARγ activation was reported to inhibit the proliferation of malignant cells from different lineages such as liposarcoma (Tontonoz et al., 1997), breast adenocarcinoma (Elstner et al., 1998; Mueller et al., 1998; Kumar et al., 2009), prostate carcinoma (Kubota et al., 1998), colorectal carcinoma (DuBois et al., 1998; Sarraf et al., 1998), non-small cell lung carcinoma (Chang, 2000), pancreatic carcinoma (Motomura et al., 2000), bladder cancer cells (Guan et al., 1999) and gastric carcinoma (Sato et al., 2000) both in vitro and in vivo.

In contrast, less is known about the physiological role of the PPARδ isoform, although there is some evidence supporting its involvement in embryo implantation and development (Barak et al., 2002; Lim et al., 1999), epidermal maturation and wound healing (Di-Poi et al., 2003), and regulation of fatty acid metabolism (Wang...
et al., 2003). Recently, the effect of PPARδ function on colon carcinogenesis has been reported (He et al., 1999). Recent evidences suggest that PPARδ plays a role in the regulation of fatty acid catabolism, energy metabolism and reverse cholesterol transport (Luquet et al., 2005). In addition, potent and selective PPARδ agonists have been shown to improve insulin resistance and reduce plasma glucose in animal models of type 2 diabetes (Tanaka et al., 2003).

In this report, we developed an in silico approach to search genes containing PPRE in their promoter region. Since all PPREs described previously consist of the juxtaposition of the degenerated hexamer AGGTCA sequence separated by one nucleotide (DR1) and with several recent reports demonstrating a DR2 site is also a PPAR-responsive element (Fontaine et al., 2003; Gervois et al., 1999; Kumar et al., 2004; Kumar et al., 2009; Venkatachalam et al., 2009), we incorporated into our database to search for both DR1 and DR2 and the 5'-flanking region that is required for binding of PPAR/RXR to a DR2 site (Gervois et al., 1999). Using these properties, a programming methodology was formulated and an in-built program was written to search for PPRE elements in gene promoters. This program is available for public use at http://www.classicrus.com/PPRE/.

MATERIALS AND METHODS

PPRE motifs database

Keywords “PPRE”, “PPAR consensus sequences” “PPAR binding sequences” were used to search for literatures in Pubmed for identifying PPRE elements. Only experimentally validated PPRE elements were incorporated into the collection. Our database contains 412 DR1 (Direct Repeat 1) and two DR2 (Direct Repeat 2) consensus patterns. The database stores the information for each PPRE on its reported consensus, isoform specificity (i.e. whether the gene is regulated by PPAR α, β, or γ), the reported in vivo and in vitro binding efficiencies and Pubmed identification numbers. Each gene input promoter sequence is restricted to 10,000kb per gene. Thus, a total of 100 genes can be searched at a time however processing speed may be comprised. Our PPRE database contains, PPRE patterns that were tested experimentally using in vitro and in vivo binding assays and its corresponding reported binding efficiency. PPRESearch reports the binding efficiency of the PPRE pattern found in the DNA sequence. This binding efficiency is the quality score of the PPREs found. PPRESearch output is in the table and can be easily copied to excel for further processing. PPRESearch usually generates its output in the tab limited text file.

PPRE motifs detection module

PPRE motif detection module uses the PPRE motifs from the database to mine for PPRE elements in the promoter region of the gene. An in-built CGI script in the PPRESearch helps the user to search their promoter sequence using the PPRE database (Fig. 3 [Supplementary data]). The unique search features incorporated in the PPRESearch are summarized below.

PPRESearch in DR1 and DR2 Motif identification

The general DR1 and DR2 consensus are AGGTCA N AGGTCA (6-N-6) and AGGTCA NN AGGTCA (6-NN-6), respectively. Both the above patterns are matched against the promoter region of the gene of interest, first by looking for the surrounding hexamers and then looking for the spacer i.e. one nucleotide or two nucleotides separation. The motifs for DR1 and DR2 are identified based on the number of ‘N’s that separate the two hexamer consensus pattern. The users could restrict the search for either DR1 or DR2 motifs in their promoter region.

PPRESearch and flanking sequence upstream of hexamer 1

It must be noted that nucleotides upstream of Hexamer 1 are involved in PPAR binding and PPAR isoforms specificity to the response elements (PPRE). This 5’flanking site is reported to be AT rich and consists of 7 nucleotide sequences with the consensus C(A/G)(A/G)AA(T)CT. This motif contributes to selective binding of PPAR/RXR hetero-dimer from other binding complexes of the nuclear receptor family such as RXR/RXR, or APR1 (apolipoprotein regulatory protein-1). This extended binding requirement ensures that PPARs occupy 5’ half site of DR1 or DR2. The presence of this flanking sequence promotes the formation of PPAR and RXR heterodimer formation. The genes with few imperfect DR1 or DR2 repeats and perfect 5’ flanking sequence from its consensus, still have PPAR/RXR heterodimer binding. Genes with imperfect 5’ flanking sequences or with few changes from its consensus would show difference in PPAR/RXR binding strength. This extended half site is a requirement for PPAR binding to the consensus. PPRESearch searches for 5’ flanking sequence of the predicted PPRE element and displays the sequence and its total number of matches in its flanking consensus. Generally PPREs having greater than 4 matches in the flanking region consensus, have better binding strength to PPARs. PPRESearch also allows the user to set flanking count threshold when predicting PPREs in the gene promoter region.
PPRESearch in finding competitive nuclear receptors

Nuclear receptors are a large family of receptors, which include subfamilies such as Thyroid hormone Receptors (TR, RAR, PPAR), Retinoid hormone receptors (HNF4, RXR, COUP), Estrogens receptors (ER) etc. These receptors bind to cis-acting sequences in target genes and regulate the expression (Mangelsdorf et al., 1995). These receptors bind to response elements which includes directed repeats of AG(G/T)TCA (Hexamers) with 0 to 5 spacing. This promiscuous cis acting element (DR1 type) is recognised by RXR (retinoid X receptor), COUP-TFI (Chicken ovalbumin upstream promoter transcription factor), ARP1 (apolipoprotein A1) and HNF 4 (Hepatocyte nuclear factor 4) as homodimer. While PPAR:RXR, RAR:RXR, COUP-TFI:RXR, ARP1:RXR also bind as heterodimers (Nakshatri et al., 1998). Hence, in addition to ligand activated PPAR/RXR binding to DR1 sequence, this sequence is bound by ARP1, HNF1, HNF4, ARP1, COUP-TFI and RAR. Nakshatri et al studied the effect of degenerate bases within repeated motifs and spacer elements in receptor binding (Nakshatri et al., 1998). Using his study as the basis, the rules for other nuclear receptors binding to the PPRE's are formulated and these rules are incorporated in our PPRESearch to mine for possible competitive receptor in PPAR binding (Table 1).

Nuclear receptor binding rules incorporated in PPRESearch

HNF4 binding rules
1. T at position 1 instead of a purine reduces the affinity for ARP1, RAR:RXR and PPAR:RXR when compared to HNF4. Hence the Response element (RE) of HNF4 will have T at position 1. (TGGGCA A AGGTCA) could be a classical RE for HNF4.
2. Additional base preferences include bases at position 4 and 7. G instead of T at base 4 reduces the binding strength of PPAR:RXR and RAR:RXR
3. The DR1 spacer A instead of G at position 7 is beneficial for HNF4 binding.
4. It is noted that among the naturally occurring HNF 4 response elements, 4 elements contain pyrimidine at position 1 and A at position 4.

ARP1 binding rules
1. The bases at position 1, 4 and 7 neither increase nor decrease the ARP1 affinity to DNA.
2. G instead of A at position 7 decrease affinity for HNF4 and RAR:RXR without affecting ARP1 affinity.
3. In general DR1 spacer having C instead of A, G or T at position 7 bind preferentially to ARP1 as other receptors have poor binding.
4. G at position 2 when replaced with A decreases the affinity for HNF4,RAR:RXR and PPAR:RXR much more when compared to ARP1

RAR:RXR binding rules
1. In view of RAR:RXR binding, none of the bases in DR1 repeats provide selective advantage for RAR.
2. DR1 element functions as RARE only in the cells which has little PPAR, HNF4 and ARP1. Embryonic carcinoma cells P19 and F9 which lack HNF4, PPAR, ARP1 and COUP TFI, DR1 elements act as RARE (Ben-Shushan et al., 1995).

Firstly the PPRE elements are searched using the PPREs collected in the database and then the above rules were applied on each PPREs to check for the competitive nuclear receptors.

PPRESearch in PPAR isoform specificity search

PPRE database is a comprehensive collection of all PPRE motifs published so far. It includes the PPRE elements that are specifically recognized by PPAR α, β and γ. It also contains binding efficiencies of PPARs to these PPRE elements. PPRESearch uses these collections to search PPAR isoform specificity for the PPRE element in the gene promoter. It reports isoforms that bind to PPRE in the promoter with its binding efficiencies. The information on PPRE isoform specificity and its corresponding binding strength help the user understand the PPAR isoform that binds to PPRE element in gene promoter and preferences of PPARs.

Other special features in PPRESearch

PPRESearch allows the user to set the threshold binding efficiency when searching for PPRE elements in a gene promoter region. In addition, user can search for PPREs with few mismatches from the PPRE consensus in the database. For better user experience and understanding PPRESearch displays the predicted PPRE element in the gene promoter region with the mismatch nucleotides from its consensus in bold letters and outputs the start position of the gene promoter region.

DISCUSSION

In this work we have constructed an improved in silico approach to finding genes containing PPRE in their promoter region. PPRESearch database contain all PPREs that are found to date. Each sequence in the
database is matched with the genomic sequences and the binding strength of each sequences are known. All the retrieved outputs correspond to the binding strength of known PPREs to which they are matched. Hence, there is a possibility of knowing original binding strengths as stated in the literature for a particular PPRE pattern found in the promoter region. Recent studies evidenced that the PPAR:RXR heterodimer can recognise DR2 repeats. Till now none of the online transcription factor binding programs are able to search for DR2 repeats with respect to PPAR binding. A detailed study was done on PPAR isoform-specific binding properties, 5’flanking sequence properties, DR1 and DR2 properties and competitive nuclear receptor properties that determine the PPAR binding. PPRESearch searches for both DR1 and DR2 binding sites. PPRESearch has the capability to analyse if the particular site is PPAR, PPARβ, PPARγ, any of the two isoforms or all the three isoforms. If all the 3 isoforms bind to the PPRE site, it can predict which isoform is most preferred based on binding affinities. This helps to understand the binding affinity of each isoforms on a PPRE site present in a gene and their control. Ongoing pharmaceutical research continues to pursue PPAR ligands with enhanced therapeutic efficacy and superior margins of safety. Given the effort on discovering the mechanisms driving the undesired effects of PPAR agonists, our PPRESearch database should help better identify candidate genes for favorable clinical activities.

Acknowledgement

This work was supported by grant R-183-000-204-213 from the National Medical Research Council, Singapore to MVC and APK.

References


