Aspects of the Control of Phosphoenolpyruvate Carboxykinase Gene Transcription*

Published, JBC Papers in Press, July 27, 2009, DOI 10.1074/jbc.R109.040535
Jianqi Yang†, Lea Reshef†, Hanoch Cassuto§, Gabriela Alemán§, and Richard W. Hanson†

From the †Department of Biochemistry, Case Western Reserve University School of Medicine, Cleveland, Ohio 44106-4936 and the ‡Department of Biochemistry and Molecular Biology, Hebrew University-Hadassah School of Medicine, Jerusalem 91120, Israel

Overview of PEPCK-C Gene Transcription

The cytosolic form of phosphoenolpyruvate carboxykinase (GTP) (EC 4.1.1.32) (PEPCK-C) occupies an important position in energy homeostasis because it is involved in the regulation of glucose synthesis, fatty acid re-esterification, and the cataplerosis of citric acid cycle anions. PEPCK-C gene transcription can change very rapidly; cAMP administration causes a 10-fold increase in transcription in liver within 20 min (1), whereas insulin rapidly inhibits transcription (i.e. a 50% decrease in PEPCK-C mRNA levels within 30 min) (2). In contrast, the activity of the enzyme changes far less rapidly, i.e. a 2-fold increase in hepatic PEPCK-C activity was noted 3 h after Bt2cAMP administration (3). This suggests regulation of PEPCK-C mRNA processing or post-translational modification of the enzyme. A recent study using both yeast and HepG2 hepatoma cells has demonstrated that PEPCK-C is post-translationally regulated by acetylation at Lys-19 and Lys-514 of PEPCK-C; this increases PEPCK-C activity in yeast and hepatoma cells (4). However, it is generally true that the induced changes in the amount of immunoreactive PEPCK-C protein correlate well with the corresponding alterations in the activity of the enzyme (5).

PEPCK-C is a single copy gene whose transcription is regulated by the binding of transcription factors to specific sites on the gene promoter. Fig. 1 summarizes the location of transcription factor-binding sites that, together with the transcription factors and co-regulators, influence the expression of the gene in mammalian tissues. Regulatory elements in the PEPCK-C gene promoter may best be understood as “units,” each of which is composed of several different transcription factor-binding sites that function together in response to hormonal and dietary stimuli (Fig. 1B). These include the cAMP regula-

tory unit in liver (6), the acidosis regulatory unit in kidney (7), and the GRU, which is critical for the response of the gene to glucocorticoids and insulin in liver (8, 9). PEPCK-C is robustly expressed in four tissues (liver, kidney, and white and brown adipose), and several hormones differentially regulate its transcription in a tissue-specific manner. The rate of transcription of the gene for PEPCK-C is altered rapidly and differentially in these tissues in response to changes in the dietary and hormonal status of the animal. Thus, a stimulus that induces gene transcription in one tissue might repress it in another. One important caveat is that low levels of PEPCK-C are present in many mammalian tissues (10), but there is very little information concerning the factors that control transcription of the gene in tissues other than the four listed above. Finally, PEPCK-C in mammals first appears at birth (11), an event that contributes to the initiation of gluconeogenesis and glyceroneogenesis in the newborn. The regulation of transcription of the PEPCK-C gene had been reviewed in detail previously (12, 13). This minireview focuses on areas that have received less attention over the years.

Role of Chromatin Structure and Methylation in the Developmental and Tissue-specific Regulation of PEPCK-C Gene Transcription

The mechanisms involved in the tissue-specific expression of PEPCK-C and the pattern of its expression during development are critical areas for future investigation. PEPCK-C activity is present at significant levels in the perportal region of the liver, kidney cortex, and brown and white adipose tissue, all known to be involved in either gluconeogenesis or glyceroneogenesis. However, an immunohistochemical analysis of PEPCK-C in mouse by Zimmer and Magnuson (10) detected the enzyme in a wide variety of mouse tissues, most of which do not make glucose and store little triglyceride (i.e. bladder, stomach, small intestine, vagina, parotid gland, submaxillary gland, and brain). During mouse development, immunoreactive PEPCK-C was observed in the nervous system and in somites of 15-day mouse embryos (10). PEPCK-C is largely absent in mammalian liver before birth and develops to greater than adult levels in rats livers within 8 h after birth (14); this parallels the initiation of hepatic gluconeogenesis. The factors that control the tissue-specific development of PEPCK-C gene expression are poorly understood (Fig. 1C). It is likely that the developmentally regulated expression of key transcription factors in specific tissues patterns the transcription of the gene for PEPCK-C. As an example, deleting the gene for CAAT enhancer-binding protein-α in mice completely represses expression of PEPCK-C in liver at birth and leads to profound hypoglycemia in the newborn (15). This shifts the emphasis from the PEPCK-C gene itself to factors that control expression of genes for transcription factors during development. The transcriptional control of PEPCK-C gene expression becomes prominent during development (especially during the perinatal period) when the gene undergoes dramatic changes from dormancy to being actively transcribed. This involves chromatin remodeling, which has...
been studied by determining (a) the methylation status of the DNA, (b) the chromatin status, and (c) the methylation of Arg-17 in histone H3.

**Methylation**—Benvenisty et al. (16) analyzed the extent of methylation of 18 sites in the PEPCK-C gene and its surrounding sequences in adult liver, kidney, spleen, and heart in fetal liver using CpG site-sensitive enzymes. There was extensive undermethylation of the gene in adult liver and kidney (both actively express PEPCK-C), whereas the gene in spleen, heart, and fetal liver (tissues that do not express PEPCK-C) was heavily methylated, suggesting that the level of methylation is directly related to the expression of the gene. Injection of 5-azacytidine (which decreases methylase activity) into 19-day rat fetuses caused demethylation of the gene (after 48 h) and a marked increase in PEPCK-C mRNA in liver but not in spleen (16). Bt$_2$cAMP induced the premature transcription of the PEPCK-C gene in liver within 3 h, whereas demethylation of the gene was observed 2 days later (17). Although these studies suggest that changes in the pattern of methylation of the PEPCK-C gene locus accompany the developmental and tissue-specific expression of PEPCK-C, other activating factors, which are likely not sensitive to methylation status, are clearly involved in this process.

**Chromatin Status**—The PEPCK-C gene retains a compact structure, extremely DNase I-resistant, in fetal rat liver until the perinatal period. However, intra-uterine injection of Bt$_2$cAMP into 19-day rat fetuses for 3 h induced not only PEPCK-C gene transcription but concomitantly chromatin associated with the gene acquired sensitivity to DNase I (17). The partial DNase I sensitivity of the PEPCK-C chromatin elicited by Bt$_2$cAMP is in accord with the hepatocytes composing only 35% of the entire fetal liver cell population at this gestational age (18).

**Modifications of Histone H3/H4 by Methylation/Demethylation**—Changes in the methylation status of histone H3/H4 have been shown to modulate the transcription and chromatin occupation of transcription factors in the PEPCK-C and glucose-6-phosphatase genes. Hall et al. (19) noted that ablation of the insulin-responsive sequence failed to prevent the dominant-negative response to insulin in liver, as would be predicated if this area of the gene promoter were the only factor...
involved in the negative effect of insulin on transcription. They found a gradual but rather rapid disappearance of many key transcription factors from the chromatin of the genes for both PEPCK-C and glucose-6-phosphatase within 3 min after the addition of insulin to hepatoma cells; this correlated with a parallel decrease in methylated Arg-17 residues from histone H3.

Relationship between DNase I-hypersensitive Sites and Tissue-specific Expression of the PEPCK-C Gene—In addition to the developmental acquisition of global DNase I-sensitive chromatin, Ip et al. (20) identified five distinct DNase I-hypersensitive sites in the PEPCK-C gene in rat liver-derived hepatoma cells (Fig. 2A). Site A mapped to a 220-bp region centered at −4800 bp; Site B mapped at about −1300 bp; Site C mapped between −370 to −30 bp; Site D mapped at +4650 bp (within intron 8 of the gene); and Site E is composed of two closely spaced sites at +6100 and +6300 bp, which is within the 3′-untranslated region of the PEPCK-C gene (all sites are numbered relative to the transcription start site, i.e. +1). Several of the DNase I-hypersensitive sites have been linked to the tissue-specific transcription of the PEPCK-C gene. Williams et al. (21) noted that hypersensitive Site A (at −4800 bp) is strictly liver-specific and is absent in kidney and NIH3T3–442A preadipocytes and adipocytes. Another hypersensitive site at about −1100 bp of the PEPCK-C gene promoter (termed Site B) is present in NIH3T3–442A cells (which differentiate into adipocytes) but not in NIH3T3–C2 cells (which do not differentiate into adipocytes); this site was not detected in mouse liver. This region of the PEPCK-C gene promoter contains a PPARγ2-binding site that is critical for the adipose tissue-specific expression of the gene (Figs. 1C and 2B) (22). It is thus likely that an alteration in the chromatin conformation in this region of the PEPCK-C gene promoter is involved in the control of adipose tissue-specific expression of the gene.

Cissell and Chalkley (23) examined the tissue-specific chromatin structure of the PEPCK-C gene in kidney and liver. They detected several DNase I-hypersensitive sites in the gene in kidney. As expected, one of the sites mapped to between −370 and −30 bp relative to the start site of transcription (termed Site C); this site contains sequences required for general transcription in virtually all tissues that express PEPCK-C. A second site was at +1900 bp, within intron 4 (termed Site K); hypersensitive Site K is unique to kidney tubules and was not found in a line of cultured kidney cells that did not express the PEPCK-C gene.

Use of Transgenic Mice to Correlate the Function of the DNase I-hypersensitive Site in the PEPCK-C Gene Promoter—Support for the importance of the kidney-specific hypersensitive sites for renal PEPCK-C gene expression can be found in studies with transgenic mice in which the tissue-specific regulation of the gene promoter was studied. Transgenes that included segments of the PEPCK-C gene promoter ranging from −2000 to +73 or −109 to +73 bp linked to a marker gene were expressed in the appropriate mouse tissues (i.e. liver and adipose tissue) (24). However, expression of the transgene in kidney was only partially induced compared with a transgene containing the entire PEPCK-C gene (25). These findings suggest that the kidney-specific DNase I-hypersensitive site at +1900 bp (Site K in intron 4) is required for the full physiological level of expression.
MINIREVIEW: Control of PEPCK Gene Transcription

of the gene in kidney cortex. The role of the sequences in Site K in the regulation of renal PEPCK-C gene transcription is not clear at this time.

An analysis of the PEPCK-C gene promoter in transgenic mice indicated that a region between −540 and −362 bp is required for the glucocorticoid induction of PEPCK-C gene transcription in kidney (25); this is consistent with the position of the GR, which maps at −410 bp. Site C also contains a duplicated HNF-1-binding site (Fig. 2B), which is required for the renal specific transcription of the PEPCK-C gene and for its response to metabolic acidosis (26). Patel et al. (24) noted that the mutation of this site in the PEPCK-C gene promoter in transgenic mice resulted in complete loss of transcription from the promoter in kidney. As mentioned above, the region in hypersensitive Site C is not restricted to the control of renal gene transcription; it is a key region for the expression of the PEPCK-C gene in all tissues studied to date. We assume that proteins binding at Sites K and C act cooperatively to ensure appropriate expression of the PEPCK-C gene in kidney cortex.

Is There a LCR Associated with the Hepatic PEPCK-C Gene?—DNase I-hypersensitive Site A, which maps at −4800 bp upstream of the PEPCK-C gene, can function to increase both the copy-dependent enhancement and frequency of transcription of a reporter gene that has been introduced into transgenic mice (27). In this regard, it shows properties of a LCR and resembles the hypersensitive sites flanking the coding region of the β-globin gene (28, 29) and the T-cell-specific CD2 gene (30). Site A contains a sequence that resembles a doublet of the extended GRU (discussed above).3 As discussed in detail below, this region is highly conserved in the mouse, human, rat, and dog genomes.

Regulatory Elements in the PEPCK-C Gene Promoter Are Conserved during Evolution

DNase I-hypersensitive Sites—It is noteworthy that significant blocks of nucleotide sequence in the PEPCK-C gene and surrounding regions are highly conserved in several vertebrate genomes. We assume that the high level of sequence identity noted in areas that are outside of the coding region of the PEPCK-C structural gene provides strong support for the importance of these regions of DNA in the control of PEPCK-C gene expression. In Fig. 2A, we present the rat PEPCK-C gene and compare it with the sequence of the same gene from eight other vertebrate species (human, mouse, dog, cow, opossum, chicken, frog, and fish). The rat PEPCK-C gene has 10 exons and 9 introns; the intron/exon alignment is virtually identical in the five mammalian species studied and in chicken, but the first exon is not found in the frog and fish PEPCK-C genes. There is a highly conserved region in intron 6 of the PEPCK-C gene that is found in the five mammalian species and in frog; the potential regulatory significance of this region of the PEPCK-C gene is not clear.

The relative positions of the DNase I-hypersensitive sites discussed above are annotated at the top of Fig. 2A. Sites C, K, and D are conserved in the vertebrate genome. DNA sequences that correspond in location to Site A (at about −4800 bp), Site B (−1100 bp), Site −3.1, and Site E (at about +6200 bp) are not conserved in their exact locations, as originally aligned. However, Site A is relatively close to a highly conserved sequence located at −5700 bp. This site is present within the same region (relative to the start site of PEPCK-C gene transcription) in all five of the mammalian species analyzed. The positions of the original DNase I-hypersensitive sites were based on measurements of the migration of DNA fragments in agarose gels, which is subject to some variation, so it is possible that Site A corresponds to the conserved site at −5700 bp noted in the analysis in Fig. 2A. If so, it is a strong verification of the importance of this site as a LCR that is involved in potentially controlling the development of PEPCK-C gene transcription, as suggested by Ip et al. (20).

Site B, which is located at −1100 bp in rat, most likely corresponds to a highly conserved site at −1000 bp that is present in all mammalian species annotated in Fig. 2A. As discussed above, this region contains a PPARγ2-binding domain that is required for the adipose tissue transcription of the mouse PEPCK-C gene. Finally, the 3′-end of the PEPCK-C gene contains a region at +5700 bp that is highly conserved in the five mammalian species studied. This site most likely corresponds to Site E, which Ip et al. (20) mapped at +6300 bp.

Conserved Regulatory Regions of the PEPCK-C Gene Promoter—We have carried out a detailed comparison of the 5′-flanking region of PEPCK-C genes from a number of vertebrate species, ranging from humans to birds (Fig. 2B). Within the first 1000 bp of sequence 5′ to the start site of gene transcription, there are a number of highly conserved (>95% sequence identity between mammalian species) regions that correspond directly to known transcription factor-binding sites within the rat PEPCK-C gene promoter. The chicken PEPCK-C gene promoter shares with other mammals (Fig. 2B) a 400-bp region of sequence identity immediately 5′ to the transcription start site, but with little identity beyond. Several new insights are forthcoming from this comparison. For example, the cAMP response element of the PEPCK-C gene promoter is adjacent to an Sp1-binding site that lies between a ubiquitous NF1-binding site and the cAMP response element; this Sp1 site is present in the human, chimp, dog, mouse, rat, and chicken genomes. A previously unidentified Sp1 site was also noted at approximately −70 bp of the gene promoter in these species. We have shown that mutations in both of these Sp1 sites markedly alters basal and protein kinase A-stimulated transcription from the PEPCK-C gene promoter.4 The HNF-1-binding site in the promoter is repeated twice within what we previously considered to be a single regulatory element (31, 32); this arrangement is conserved in all species we have analyzed (Fig. 2B). Finally, a hitherto unrecognized HNF-4α-binding site at −272 to −252 bp was identified by phylogenetic sequence comparison. We demonstrated in subsequent studies that HNF-4α binding to this site is responsive to SIRT1-catalyzed deacetylation induced by resveratrol (37).

In contrast to the highly conserved sites, there is considerable variability among species regarding the occurrence and position of the SREBP-1c-binding site in the PEPCK-C gene promoter; this transcription factor is involved in the response of


4 J. Yang, G. Aleman, and R. W. Hanson, unpublished data.
the PEPCK-C gene to insulin (33, 34). Along with HNF-4α, FOXO, and PGC-1α, SREBP-1c regulates the repressive response of the hepatic PEPCK-C gene to insulin. This response is critical for the regulation of hepatic glucose output and for glyceroneogenesis in liver. It is thus of particular interest that the human PEPCK-C gene promoter has several SREBP-1c regulatory elements with near-perfect agreement with the SREBP-1c-binding site in the low density lipoprotein gene promoter (34). These elements were shown by electrophoretic mobility shift assay to bind tightly to SREBP-1c. In contrast, the rat PEPCK-C gene promoter requires the nuclear high mobility group protein to bind SREBP-1c effectively (34). On the other hand, the mouse PEPCK-C gene promoter does not have a detectable SREBP-1c-binding site based on the consensus sequence of these sites.

As mentioned in detail above, there is a highly conserved region at about −1000 bp in the PEPCK-C gene promoter that is found in the genome of all mammalian species we analyzed (i.e. human, rat, mouse, cow, and dog) but is not present in opossum, fish, chicken, and frog (Fig. 2A). This region contains a PPARγ2-binding site that controls transcription of the PEPCK-C gene in adipose tissue (22) and the distal component of the extended GRU in liver (9). Thus, in mammals, glucocorticoids coordinates the circulation of fatty acids between liver and adipose tissue. This involves a reciprocal regulation of PEPCK-C gene transcription in the two tissues via the same conserved regions in the gene promoter in adipose tissue and the extended GRU in liver. The absence of this conserved region in birds, fish, and frogs also suggests a different mode of transcriptional control in adipose tissue gene transcription from these species.

Concluding Remarks

Despite advances in our understanding of specific transcription factors that bind to regions of the PEPCK-C gene promoter in proximity to areas of hypomethylation, there is little information on the mechanisms involved in controlling the methylation status of the gene. Likewise, factors that modify nucleosomes associated with the gene, which recruit regulatory proteins complexes to the promoter, remain elusive. Chromatin modifications such as changes in acetylation, ubiquitination, sumoylation, and methylation are also likely to play a role in the control of PEPCK-C gene transcription. The research summarized in this minireview provides an excellent starting point from which to probe the mechanisms by which chromatin modification controls PEPCK-C gene transcription.

Acknowledgment—We are grateful to Dr. David Samols for critically reading this article while in manuscript form.

REFERENCES
