Glycogenin-2, a Novel Self-glucosylating Protein Involved in Liver Glycogen Biosynthesis*

(Received for publication, April 21, 1997, and in revised form, August 5, 1997)

James Mu, Alexander V. Skurat, and Peter J. Roach‡

From the Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, Indiana 46202-5122

Glycogenin is a self-glucosylating protein involved in the initiation phase of glycogen biosynthesis. A single mammalian gene had been reported to account for glycogen biogenesis in liver and muscle, the two major repositories of glycogen. We describe the characterization of novel forms of glycogenin, designated glycogenin-2 (GN-2), encoded by a second gene that is expressed preferentially in certain tissues, including liver, heart, and pancreas. Cloning of cDNAs encoding glycogenin-2 indicated the existence of multiple species, including three liver forms (GN-2α, GN-2β, and GN-2γ) generated in part by alternative splicing. Overall, GN-2 has 40–45% identity to muscle glycogenin but is 72% identical over a 200-residue segment thought to contain the catalytic domain. GN-2 expressed in *Escherichia coli* or COS cells is active in self-glucosylation assays, and self-glucosylated GN-2 can be elongated by skeletal muscle glycogen synthase. Antibodies raised against GN-2 produced in *E. coli* recognized proteins of Mr ~66,000 present in extracts of rat liver and in cultured H4IIEC3 hepatoma cells. In H4IIEC3 cells, most of the GN-2 was present as a free protein but some was covalently associated with glycogen fractions and was only released by treatment with α-amylase. H4IIEC3 cells also expressed the muscle form of glycogenin (glycogenin-1), which was attached to a chromatographically separable glycogen fraction.

Glycogen, a branched polymer of glucose, is a metabolic energy reserve accumulated in many cell types (1). In mammals, the major glycogen deposits in absolute amount are those of skeletal muscle and liver. These glycogen reserves have somewhat different functions, but both pools contribute to blood glucose homeostasis. In eukaryotes, the pathway of glycogen biogenesis consists of an initiation step and a subsequent phase of bulk polysaccharide synthesis mediated by glycogen synthase and the branching enzyme (see, for example, Ref. 2). The existence of a specific initiator protein, postulated from the kinetic studies of glycogen synthase and the branching enzyme (see, for example, Ref. 2), was responsible for glycogen biogenesis in both liver and muscle (16). Ercan et al. (17) have also described a self-glucosylating protein of similar size present in rat liver, whereas several other reports have noted mammalian proteins, of different sizes, that had some of the properties of glycogenin (reviewed in Refs. 7 and 18). Whether or not these species are related by amino acid sequence to the known glycogenin was not established. In other organisms, however, there is clear evidence for multiple glycogenin-like proteins (see Ref. 6). Two self-glucosylating proteins, Glg1p and Glg2p, from the yeast *Saccharomyces cerevisiae* have been characterized and shown to be required for glycogen accumulation in that organism (19). In addition, searching GenBank revealed that *Caenorhabditis elegans*, *Caenorhabditis briggsae*, and *Arabidopsis thaliana* all express multiple glycogenin-like messages (6). We report here the existence of a second mammalian glycogenin gene preferentially expressed in liver, heart, and, to a lesser extent, pancreas. The protein, which we designate glycogenin-2, has all the properties expected of glycogenin; it can self-glucosylate, can act as a substrate for glycogen synthase, and is released from glycogen by α-amylase treatment. This finding makes it necessary to reconsider the enzymic components implicated in the pathway for glycogen biogenesis in these tissues.

**EXPERIMENTAL PROCEDURES**

Cloning and Sequencing of cDNAs Encoding Glycogenin-2—Two EST clones, identified by database searches, were obtained from Genome Systems Inc. (St. Louis, MO). The brain EST clone (corresponding to the sequence with accession number H04891) and the breast EST clone (corresponding to the sequence with accession number R71874) were completely sequenced on both strands. Several errors were found.

*This work was supported by National Institutes of Health Grant DK27221. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) U94357, U94358, U94362, U94363, U94364, U94369, U94361, U94359.

‡ To whom correspondence should be addressed. Tel.: 317-274-1582; Fax: 317-274-4686; E-mail: peter_roach@iupui.edu.

1 The abbreviations used are: EST, expressed sequence tag; PCR, polymerase chain reaction; PAGE, polyacrylamide gel electrophoresis; kb, kilobase pair(s).
in the original sequences in the data base, and complete corrected sequences for the two cDNA clones have been deposited in GenBank with accession numbers U94357 (brain) and U94358 (breast). Sequences from the other ends of the breast clone (accession number R7875) and brain clone (accession number H04798) were already presented in our previous communication. A small overlap in the presence of 10 m bases identical sequence was found in the brain and breast cDNA clones. These cDNAs were used as probes to screen human brain and human liver cDNA libraries (5'-stretch plus cDNA libraries, in Agt10, from CLONTECH). A 320-base pair NcoI-EcoRI fragment (probe 1) of the brain clone was used to screen the brain library, and a 0.3-kb EcoRI-NcoI fragment from the 5'-end of a breast clone 9-1 (as "Results") was used for liver library screening. Totals of 4 x 10^6 plaques were screened for each probe. Probes were randomly labeled by [32P]dCTP using a New England Biolabs NEBlot kit. Approximately 2 x 10^6 cpm/ml probe was used for hybridization which was carried out in 5 x SSPE (5 x SSPE: 0.15 mM NaCl, 0.01 mM Na_2HPO_4, 1 mM EDTA, pH 7.4), 10 x Denhardt's solution (1 x Denhardt's solution: 0.2 μg/μl Ficoll, 0.2 μg/μl polyvinylpyrrolidone, 0.2 μg/μl bovine serum albumin), 100 μg/ml single-stranded DNA, 0.2% SDS, 50% formamide, 0.05% pyrophosphate at 42 °C for 18 h. Filters were washed twice in 2 x SSC (1 x SSC: 0.15 M NaCl, 0.015 mM sodium citrate, pH 7.0) and 0.5% SDS at room temperature for 20 min each and then washed twice for 20 min in 1 x SSC and 0.1% SDS at 65 °C. Positive clones were picked, replated, and taken to plaque purity. λ phage DNA was purified using polyethylene glycol/NaCl precipitation of the plate lysate, followed by phenol/chloroform (1:1) DNA extraction. Purified λ DNA preparations were digested with EcoRI and subcloned into the EcoRI site of the pGEM7 vector (Promega) for sequencing. DNA was sequenced by the dyeox analysis method of Sanger (20), either manually or with an Applied Biosystems automated sequencer. All DNA sequences were reported in this work result from sequencing the corresponding cDNAs on both strands. For cDNA clones with internal EcoRI sites, the order of the EcoRI fragments was confirmed by polymerase chain reaction (PCR) using uncut λDNA as templates and primers that straddle the EcoRI sites.

**Northern Blot Analysis**—A human multiple tissue Northern blot was obtained from CLONTECH. The blot was hybridized with probes from the brain clone (H04798/H04881) using either probe 1 or a 1.2-kb PstI-NolI fragment (probe 2). The filter was also probed with a 0.9-kb NolI-Sall fragment (probe 3) from the breast clone (R7874/R7875). Probes were randomly labeled by [32P]dCTP, as described above. Normally 2 x 10^6 cpm/microgram hybridization solution was used. Hybridization was performed in 5 x SSPE, 10 x Denhardt's solution, 100 μg/ml single-stranded DNA, 0.1% SDS, 0.05% pyrophosphate at 42 °C for 18 h. The filter was washed for 20 min in 2 x SSC and 0.1% SDS at room temperature and then washed for 20 min twice in 0.1 x SSC and 0.1% SDS at 60 °C. The filter was exposed to x-ray film at -80 °C. Before each re-use, the filter was stripped in boiling 0.5% SDS solution for 10 min, followed by a wash with 2 x SSC for 10 min according to the manufacturer's recommendations, and then exposure of the film to air at -80 °C to keep the filter dry and thus functioning at a much lower cost.

**Expression, Purification, and Assay of Recombinant Glycogenin-2**—The coding sequences from the liver cDNAs encoding glycogenin-2a and glycogenin-2y (clones 7' and 1', respectively) were cloned into the E. coli expression vector pET-28a so as to express NH2-terminally His6-tagged glycogenin-2 purified by Ni^{2+}-agarose chromatography was used for antibody production in guinea pigs (Cocalico Biologicals, Inc.). The specificity of the antisera was tested by Western hybridization using extracts from COS-1 cells expressing glycogenin-2 (see "Results").

**Immunoprecipitation—** Recombinant His6-tagged glycogenin-2a purified by Ni^{2+}-agarose chromatography was used for antibody production in guinea pigs (Cocalico Biologicals, Inc.). The specificity of the antisera was tested by Western hybridization using extracts from COS-1 cells expressing glycogenin-2 (see "Results"). The Western blot procedure was essentially as described previously (21). Signals were visualized by using either the Enhanced Chemiluminescence (ECL) system (Amersham) or 125I-protein A (5 μg/ml hybridization). Normally, the antisera was diluted 1:2000 for Western analysis. Rat hepatoma H4IEC3 cells were grown in Dulbecco's modified Eagle's medium containing 10% fetal bovine serum. Confuent monolayers were harvested in 1 ml of buffer A (50 mM Tris-HCl, pH 7.8, 10 mM EDTA, 2 mM EGTA, 100 mM NaF, 0.1 mM N-tertosyl-L-lysine chloromethyl ketone, 1 mM phenylmethylsulfonyl fluoride, 1 mM dithiothreitol) and frozen in liquid nitrogen. After thawing, cells were homogenized by repeated passage through a 22-gauge needle. The supernatant (1.2 mg of protein/ml) at 1:200 dilution; for the anti-sa amylase for 30 min at 30 °C. Antisera to glycogenin-2 was added to 1 ml of 10% rat liver supernatant (1.2 mg of protein/ml) at 1:200 dilution; for the antisyndactylase after 30 min of self-glucosylation, purified glycogen synthase (22) that are effective for immunoprecipitation but not for Western analysis. These were available as a purified IgG fraction from rabbits or polyclonal antisera to muscle glycogenin (23), also produced in this laboratory, are effective for both immunoprecipitation and Western hybridization. Serum containing polyclonal antibodies to rabbit skeletal muscle glycogen synthase produced in chicken was kindly provided by John C. Lawrence Jr., University of Virginia. It was used for Western analysis of glycogen synthase from 1,400 angiotensin for 30 min at 30 °C. The supernatant from this low speed spin was centrifuged on 5000 x g for 10 min. Protein taken from the pellet fraction was used for Western analysis.

**Immunoprecipitation—** Supernatants from extracts of H4IEC3 cells were obtained as described above and incubated with 80 μg/ml anti-amylase for 30 min at 30 °C. Antisa to glycogenin-2 was added to 1 ml of supernatant (1.2 mg of protein/ml) at 1:200 dilution; for the anti-glycogen synthase IgG, 2 μg of antibody was added. Cell extracts were incubated on ice for 1 h before addition of 20 μl of protein A-agarose (Life Technologies, Inc.). The samples were then incubated for another 1 h on a nutator at 4 °C, following by two washes with 1 ml of cold buffer containing 1 ml of 0.1% Tween-20 and 0.1% Triton X-100 and 0.3 M NaCl and then with 1 ml of 0.25 M sucrose, adjusted to pH 7.4 with Tris base) and centrifuged at 700 x g for 10 min. The supernatant from this low speed spin was centrifuged at 5000 x g for 10 min. Protein taken from the pellet fraction was used for Western analysis.
Novel Mammalian Glycogenin Gene

Despite their multiplicity, all of the cDNAs appear to be derived from a single gene since there were common segments of absolute sequence identity. We propose to name the corresponding protein glycogenin-2 because, as described below, it has properties consistent with its serving a similar function to the known glycogenin. Muscle glycogenin can be termed glycogenin-1 where needed for clarity. After skeletal muscle, liver is the major repository of glycogenin in the body, and so most effort was applied to analyzing clones from this source. The three independent cDNAs analyzed each had long open reading frames and differed in regions immediately following the start codon; they are named glycogenin-2α (accession number U94362), glycogenin-2β, (accession number U94363), and glycogenin-2γ (accession number U94364). Glycogenin-2α (501 amino acids) has an insertion of 40 residues compared with glycogenin-2γ, whereas glycogenin-2γ has an insert of 9 residues that aligns with the extreme NH₂-terminus of glycogenin-1. This 9-amino acid insert in glycogenin-2β is also present in glycogenin-2α. It seems most likely that the three different forms are generated by alternative splicing. Where they overlap, the three liver cDNAs have identical sequences except in three locations. First, the cDNA encoding glycogenin-2γ has a region of 5'-sequence that does not match glycogenin-2α or glycogenin-2β. The region of sequence dissimilarity ends with an EcoRI site and is probably due to an artifact during library construction. Second, glycogenin-2α and γ share a codon insertion with respect to glycogenin-2β. A possible explanation for this difference is the existence of different alleles. Finally, glycogenin-2α has a G → T substitution in the 3' noncoding sequence. The liver glycogenin-2 isoforms α, β, and γ have predicted Mr of 55,150, 51,841, and 50,995 with predicted pI values of 4.75, 4.72, and 4.76, respectively. Overall, the proteins would have 42–45% amino acid identity to muscle glycogenin (Fig. 2). However, there is 72% identity over an NH₂-terminal region of about 200 amino acids thought to be involved in catalysis (6).

The other partial cDNAs potentially define as many as four more isoforms of glycogenin-2. H04789/H04891 and brain clone 9–1 (accession number U94360) have identical sequences where they overlap and would predict a protein with an internal deletion of 71 COOH-terminal residues with respect to liver glycogenin-2α. We designate this isoform glycogenin-2δ. Brain clones 4–1 and 9–2 are identical (accession number U94361) and similarly code for a protein with a deletion in the COOH terminus with respect to glycogenin-2α, but in this case the deletion is only 36 residues. This form is termed glycogenin-2ε. Brain clone 2–1 (accession number U94359) has two internal EcoRI sites, and the two EcoRI fragments at the 5'-end are most likely due to artifacts during library construction. The 3' EcoRI fragment has a region of sequence matching the 3' regions of the liver glycogenin-2 messages except for a 13-base insertion in the noncoding sequence. However, preceding this segment of identity is a sequence not present in any of the other clones and which does not have an open reading frame. We are therefore doubtful as to the legitimacy of this clone. Finally, the breast clone would encode a partial protein sequence consistent with that, for example, predicted by the liver clones except that the sequence is truncated by a stop codon introduced by a single base change. This form is designated glycogenin-2τ. The full significance of the various partial glycogenin-2 cDNAs will require further study.

The tissue distribution of glycogenin-2 expression was analyzed by Northern blotting of poly(A)⁺ RNA isolated from several different tissues. Using three different probes derived from the original brain (H04789/H04891) and breast (R71874/R71875) clones (see “Experimental Procedures”), we obtained
Novel Mammalian Glycogenin Gene

Fig. 2. Alignment of glycogenin-2a with other glycogenins or glycogenin-like proteins. The alignment is of all current full-length sequences for glycogenins or glycogenin-like proteins. Shown, however, are only the regions of similarity: the first 281 residues of glycogenin-2a and the extreme COOH terminus of the C. elegans sequence CEG-14 has accession number Z72554 and CEG-11 is CEG-14664 from the St. Louis C. elegans sequencing project. The alignment was made with the MACAW program, followed by manual adjustments.

Uppercase letters indicate the domains selected by the computer. Several conserved domains, as defined by Roach and Skurat (6) are indicated. Note that here the alignment of domain IV includes mammalian and C. elegans sequences. Residues conform in five or more of the six sequences have an asterisk beneath. The bullets mark the Tyr residues involved in self-glucosylation in glycogenin-1, Glg1p, and Glg2p.

essentially the same results. A single hybridizing signal corresponding to a size of 3.6 kb was detected, with the strongest signal for liver RNA (Fig. 3). The band was broad enough, however, that we would not have distinguished the small differences in size predicted by the cDNAs for glycogenin-2a, b, and g. A species of similar size was clearly visible with heart and, more weakly, with pancreas. The longest liver cDNA was 3.2 kb, suggesting that we lack little of the mRNA sequence. Indeed, if the 3' of the brain H04789/H04891 clone, which does extend to a poly(A) tail, is shared by the glycogenin-2a clone, we can account for 3.4 kb of the message.

Biochemical Properties of Glycogenin-2—To test whether glycogenin-2 has properties expected of a self-glucosylating initiator protein, we expressed it in bacterial and mammalian cells. The glycogenin-2a coding sequence was inserted into a pET vector so that it would be produced in E. coli as a fusion protein with an NH2-terminal polyhistidine tag. The protein was expressed in E. coli and purification by Ni2+ -agarose chromatography yielded a major species of apparent Mr 66,000, compared with a predicted Mr of 57,329 for the His6-tagged protein (Fig. 4A). The reason for this aberrant electrophoretic mobility is not known, but it is probably not due to glucosylation since it was not significantly affected by α-amylase. Some smaller protein species were present and these probably represent proteolytic degradation products. After incubation with UDP-[U-14C]glucose, glycogenin-2a incorporated radioactive label indicating that it was capable of self-glucosylation. Note that several lower Mr, labeled species accumulated suggesting that any proteolysis did not eliminate self-glucosylation. For comparison, muscle glycogenin-1 was similarly analyzed (Fig. 4B) using 7-fold less protein. After the self-glucosylation reaction, purified glycogen synthase was added to the glycogenin together with an excess of unlabeled UDP-glucose; any high molecular weight radioactivity would then track the glycogenin protein attached to high Mr carbohydrate. For both glycogenin-1-2a and -2, a significant amount of label was moved to higher Mr, including material too large to enter the gel, indicating that glycogenin-2a was able to serve as a substrate for elongation by glycogen synthase. Similar efforts to analyze glycogenin-2g resulted in expression of protein that was not active under the

FIG. 3. Northern analysis of glycogenin-2 distribution. A human multiple tissue Northern blot was hybridized with probes from both brain and breast EST clones (see “Experimental Procedures”). The same pattern was obtained with either probe. The numbers to the right indicate the molecular weights (in kb) of standards.
conditions of the self-glucosylation assay (data not shown). The three liver glycogenin-2 coding regions were also inserted into mammalian expression vectors, as described under “Experimental Procedures,” and transiently expressed in COS-1 cells (Fig. 5). In all three cases, antibodies raised to glycogenin-2 recognized species of M₃, ~66,000. No signal was observed from control COS cell extracts, suggesting that endogenous COS cell glycogenin either was not recognized by the antibodies or else was present at very low levels. Lysates from the expressing cells were subjected to self-glucosylation assays. Both glycogenin-2α and β were active, whereas little or no label became associated with glycogenin-2γ consistent with the results of expression in E. coli.

Several enzyme kinetic properties of glycogenin-1 and glycogenin-2 were compared. Self-glucosylation of glycogenin-2, like glycogenin-1, was dependent on the presence of divalent metal ions of which Mn²⁺ was the most effective (data not shown). The reaction was not significantly dependent on pH from pH 6 to pH 9 (data not shown). In absolute terms, self-glucosylation by glycogenin-2 was slower than that by recombinant glycogenin-1, ~1 nmol/min/mg compared with ~11 nmol/min/mg, while the apparent \( K_m \) for UDP-glucose was significantly higher than that of glycogenin-1 (Table I). Like glycogenin-1, glycogenin-2 was also capable of transferring glucose to small molecule acceptors such as n-dodecyl β-D-maltoside. Under standard conditions (21), glycogenin-1 catalyzed transglucosylation at a rate more than 50 times greater than self-glucosylation, whereas, for glycogenin-2α, self-glucosylation and transglucosylation occurred at similar rates (data not shown).

A number of potential effectors of glycogenin-2 activity were also tested for their effect on self-glucosylation (Table II). In general, the most effective inhibitors of glycogenin-1 were similarly potent for glycogenin-2a. For example, CDP and UDP-xylose were strong inhibitors of both glycogenins, as reported previously for glycogenin-1 (27, 28). The same was true of some polyphosphate compounds. The most striking difference was the effect of transglucosylation acceptors, like maltose and n-dodecyl β-D-maltoside, which presumably inhibit self-glucosylation by competing as substrates. Consistent with the finding that glycogenin-2 was less active in transglucosylation, neither of these compounds was an effective inhibitor of glycogenin-2 self-glucosylation. Finally, ATP was reproducibly a stronger inhibitor of glycogenin-2 than glycogenin-1.

### FIG. 5. Expression and self-glucosylating activity of glycogenin-2 in COS-1 cells.

As described under “Experimental Procedures,” glycogenin-2α (lane 3), glycogenin-2β (lane 4), and glycogenin-2γ (lane 2) were transiently expressed in COS-1 cells. Equal amounts of cell lysates (20 μg of protein) were used for self-glucosylation (A) and Western blot analysis (B). For self-glucosylation, samples were incubated with 77 μM UDP-[U-14C]glucose at 30 °C for 30 min and separated by SDS-PAGE. The dried gel was exposed to x-ray film for 5 days. The signal for Western blotting was visualized by interaction with 125I-protein A, followed by autoradiography. The numbers to the left indicate the molecular mass values (in kDa) of standards.

### TABLE I

| Protein            | \( K_m \) (μM) | \( V_{max} \) (nmol/min/mg) |
|--------------------|----------------|-----------------------------|
| Glycogenin-1       | 12.9 ± 1.1     | 11.2 ± 0.2                  |
| Glycogenin-2α      | 70.6 ± 10.6    | 0.95 ± 0.04                 |

Analysis of Glycogenin-2 in Cells—Extracts from H4IIEC3 hepatoma cells or from the liver of fed rats were analyzed by Western hybridization using anti-glycogenin-2 antibodies. A single immunoreactive species of M₃, ~66,000 was detected in rat liver. The fraction giving the strongest signal was an 8500 ∗ g pellet (Fig. 6C), whereas only a weak signal was seen in the the 8500 ∗ g supernatant (data not shown). We also analyzed extracts from H4IIEC3 cells and likewise observed a single immunoreactive species of similar size (Fig. 6A). When the H4IIEC3 cell extract was first treated with α-amylase to degrade glycogen and release free glycogenin, there was only a small change in the signal strength for glycogenin-2, suggesting that most of this protein was not attached to carbohydrate.
TABLE II

**Effectors of glycogenin self-glucosylation**

The self-glucosylation of glycogenin-1 (55 μg/ml) and glycogenin-2 (120 μg/ml) was measured as described under “Experimental Procedures” with the indicated additions. The UDP-glucose concentrations were 19 μM and 77 μM for glycogenin-1 and glycogenin-2, respectively. Assays were for 5 min or 30 min for glycogenin-1 or glycogenin-2, respectively.

| Compound               | Glycogenin-2a | Glycogenin-1 |
|------------------------|--------------|-------------|
| Control                | 100          | 100         |
| Glucose                | 94           | 95          |
| Maltose                | 28           |             |
| Maltotriose            | 109          |             |
| Maltotetraose          | 101          |             |
| Maltopentose           | 110          |             |
| UDP-glucose-6-P        | 102          |             |
| UDP-glucuronic acid    | 102          |             |
| NAD                    | 66           |             |
| NADP                   | 73           |             |
| AMP                    | 71           |             |
| ADP                    | 48           |             |
| ATP                    | 60           |             |
| ADP-glucose            | 86           |             |
| Glucose-6-P            | 102          |             |
| Glucose-1-P            | 96           |             |
| Phosphate              | 72           |             |
| Pyrophosphate          | 10           |             |
| Tripolyphosphate       | 12           |             |
| Trimetaphosphate       | 89           |             |
| Tetrapolyphosphate     | 9             |             |

*% of control based on comparison of the signal intensities with purified recombinant glycogen as the standard, most glycogenin-2 is at least 10 times more abundant than glycogenin-1 in H4IIEC3 cells. The H4IIEC3 cell extract was also analyzed by chromatography on concanavalin A-Sepharose which binds carbohydrates such as glycogen (Fig. 7). The glycogen content of H4IIEC3 cells grown under these conditions is 19 μg of glycogen/mg of protein. Most glycogenin-2 did not bind to the column and was present in the flow-through fraction as a species of M_r 66,000, as judged by SDS-PAGE (data not shown). The column was eluted with a glucose gradient and the fractions were subjected to Western analysis using antibodies against glycogenin-1 or glycogenin-2. No signal was seen for fractions not treated with α-amylase (data not shown). However, after treatment with α-amylase, both glycogenin-1 (Fig. 7C) and glycogenin-2 (Fig. 7D) were present.

These values are based on comparison of the signal intensities with purified recombinant standards of human glycogenin-2α and rabbit glycogenin-1. Over the 160 residues for which the sequences are known, rat glycogenin-1 is 92% identical to rabbit glycogenin-1 at the amino acid level. There are no non-human sequences for glycogenin-2 to make a similar comparison. The antibodies are polyclonal and were raised against whole recombinant protein.

![Fig. 6. Western analysis of glycogenin-2 in H4IIEC3 hepatoma cells and in rat liver.](image)

![Fig. 7. Concanavalin A chromatography of a soluble extract from H4IIEC3 cells.](image)
antibodies. Tracts after treatment with genin-2. Shown in Fig. 8 are immunoprecipitates from cell extracts. Immunoprecipitates contained material cross-reacting with an- judged by Western blotting. Likewise, anti-glycogen synthase g glycogen synthase, as able that the liver isoform is present. Further evidence for an (data not shown). Therefore, we cannot be certain which glyco- cogen molecules that are chromatographically separable. The fractions from concanavalin A chromatography were also ana- lysed by Western blotting with anti-glycogen synthase antibod- ies. An immunoreactive species of approximately 82,000 was also detected in a wide range of fractions. Previously, we had observed glyco- cogen synthase in H4IIEC3 cells as species with molecular weights of 86,000 or 90,000 daltons. The difference may be due to the use of different molecular weight standards or due to proteolysis, since the liver isoform of glycogen synthase is known to be particularly susceptible to degradation. The strongest signal was in coincidence with glycogenin-1. However, with longer exposures, signal was also clearly present in earlier fractions where glycogenin-2 was dominant. We found that these chicken antibodies, although raised against muscle glycogen synthase, cross-reacted weakly with purified liver glycogen synthase (data not shown). Therefore, we cannot be certain which glyco- gen synthase isoform is being detected, although it seems prob- able that the liver isoform is present. Further evidence for an association between glycogen synthase and glycogenin-2 came from immunoprecipitation experiments (Fig. 8). Anti-glyco- genin-2 immunoprecipitates contained glycogen synthase, as judged by Western blotting. Likewise, anti-glycogen synthase immunoprecipitates contained material cross-reacting with anti-glycogenin-2 antibodies and of appropriate size to be glyco- genin-2. Shown in Fig. 8 are immunoprecipitates from cell ex- tracts after treatment with anti-glycogen synthase. However, we obtained essentially the same results if the anti-glycogen synthase digestion was omitted, presumably because most of the glycogenin-2 was free anyway. This result provides evidence for a physical asso- ciation between glycogen synthase and glycogenin-2 in H4IIEC3 cells.

**DISCUSSION**

The most significant outcome of the present study is to rec-ognize the existence of a novel glycogenin gene expressed in tissues that are important for glycogen metabolism. Liver is a major site of glycogen synthesis, and, in the heart, glycogen may have a special significance for the function of cardiac muscle. Expression of glycogenin-2 in the pancreas, although weaker, is also intriguing because of the connection between glycogen metabolism and blood glucose homeostasis. Although not all glycogenin-like species previously reported matched glycogenin-1 in every respect, there has been relatively little discussion of the presence of a separate gene. In liver, glycoge- nin-1 was found to be covalently linked to liver glycogen and its low abundance compared with skeletal muscle was rationalized by the fact that liver glycogen molecules are much larger. Ercan et al. (17) had likewise identified a self-glucosylating species of the correct size to be glycogenin-1 but had questioned whether glycogenin-1 was involved in the major pathway of liver glycogen synthesis. Interestingly, Calder and Geddes (31) did describe a protein of 60,000 daltons that was released from rat liver glycogen after digestion. If the amino acid composition analysis of this protein was correct, however, it is unlikely to be glycogenin-2 (or glycogenin-1) since it was extremely rich in Ser and Gly, which accounted for 43% of the residues. Glycoge- nin-2 has a Ser + Gly content of 16%. Thus, the significance of the Calder and Geddes protein is unknown. A substantial fraction of glycogenin-2 is not covalently attached to glycogen and would not have been purified by methods aimed at isolat- ing covalently linked proteins. On the other hand, one might have expected to detect glycogenin-2 as a self-glucosylating protein in experiments such as those of Ercan et al. (17). The assay of Ercan et al. (17) used 5 μM UDP-glucose, commensu- rate with the low Km of glycogenin-1 for the nucleoside diphos- phate sugar, so that conditions would not have been optimal for the higher Km glycogenin-2. However, we have no certain explanation of why Ercan et al. (17) failed to detect glycogenin-2. Krisman and colleagues (32, 33) have described multiple glyco- sylated heart proteins, including one of approximately 60,000 daltons, whose formation is inhibited by Mn2+ and which they associate with an activity termed “glycogen initiator 2.” It is not clear, how- ever, whether this species was considered capable of support- ing heart glycogen synthesis independent of other self-gluco- sylated species of 38,000 and 42,000 daltons, of similar size to glycogenin-1. Detection of a second glycogenin gene has been slow to emerge in the sequence data bases since the glyco- genin-2 message is simply not well represented in the libraries that are the subject of mass sequencing projects. For example, the number of human and mouse glycogenin-1 ESTs exceeds 100 and is growing steadily, whereas we are aware of only about 20 ESTs (representing 12 clones) of glycogenin-2, most of which contain only the 3′-noncoding sequence.

Glycogenin-2 has strong sequence similarity to glycoge- nin-1. Recombinant protein expressed in heterologous sys- tems is active as a self-glucosylating protein, and it can serve as a substrate for purified glycogen synthase. The protein can be detected in rat hepatoma cells and in rat liver. Fur- thermore, a portion of the glycogenin-2 can only be detected after treatment with α-amylase. Glycogenin-2 and glycogen syn- thase co-immunoprecipitate from H4IIEC3 cells. Therefore, glycogenin-2 has all of the properties that would be expected for it to function as an initiator of glycogen synthesis, and we propose that it functions in this capacity in vivo. This pro- posal would force consideration of two new concepts regard- ing glycogen synthesis. First, at least in hepatoma cells, both glycogenin-1 and glycogenin-2 appear to be simultaneously involved in glycogen synthesis since both are released from glycogen by α-amylase. Furthermore, the corresponding glyco- cogen molecules were separable by chromatography on concanavalin A-Sepharose. One interpretation would be the the existence of two populations of glycogen associated with the different glycogenin isoforms. Further work is needed to in-
vestigate this suggestion. In muscle, glycogen synthase and glycogenin-1 are found in approximately equimolar amounts (34), whereas, in liver, this ratio is very much in favor of glycogen synthase. Our work suggests that a major proportion of the glycogenin bound to liver glycogen synthase is likely to be glycogenin-2. Second, unlike glycogenin-1 in muscle extracts (35), most glycogenin-2 was present as a low molecular weight species even when glycogen was present. Free glycogenin-1 in fasted rat liver had been reported previously by Ercan et al. (17). The presence of significant amounts of free glycogenin-2 could be explained if only a small proportion of the molecules at any given time are recruited to glycogen biosynthesis. Future work should address whether the amount of free glycogenin varies under different physiological conditions. Alternatively, there could be "recycling" whereby the glycogenin is removed from the glycogen at some stage in the biosynthesis (discussed in Refs. 6 and 18) and the glycogenin bound to glycogen would represent the small fraction not yet released from the polysaccharide. This hypothesis would imply that most of the glycogen molecules have no attached glycogenin-2. Interestingly, Stetten and Katzen (36) many years ago reported the presence of reducing groups in liver glycogen, and it is likely that, unlike in muscle, most liver glycogen molecules are not covalently linked to protein (see discussion by Gannon and Nuttall (18)). An interesting consequence of such a model would be to predict the occurrence of a cleavage reaction to release the glycogenin. It will be of interest to test whether such an enzymatic activity exists in liver.

The glycogenin-2 sequence is 72% identical to glycogenin-1 over the NH2-terminal 200 residues. Like glycogenin-1 and the enzymic activity exists in liver. glycogenin. It will be of interest to test whether such an isoform-specific, and the size differences among isoforms are quite small, with glycogenin-2a no more than 4 kDa larger than the other forms. For the most part, the different potential splice variants would provide alternative NH2 and COOH termini attached to a common 335-residue segment containing the region with homology to other self-glucosylating proteins and implicated in catalysis. Whether or not all would self-glucosylate is difficult to predict. We have shown that recombinant glycogenin-2y was inactive, but it is difficult to interpret such a failure since minor changes in sequence can sometimes affect whether a functional protein is produced in bacteria. Although a role in some form of glycosyl transfer reaction is the most logical expectation for the various glycogenin-2 isoforms, it is possible that not all isoforms are involved in glycogen metabolism.

With the original identification of muscle glycogenin, it was attractive to propose regulation at the initiation step of glycogen biosynthesis, as happens for the synthesis of many other biopolymers. To date, there is relatively little evidence for such regulation. In particular, the idea that nutritional cues might set the glycogen storage capacity of cells by regulating the glycogenin level has not been substantiated. Now, the discovery of glycogenin-2 requires that some aspects of glycogen biosynthesis and its control be reevaluated, especially in liver and heart. Of specific interest would be any hormonal controls such as by insulin or glucagon. Nutritional control of gene expression is common in liver, and we can ask again whether longer term controls of glycogen storage could be exerted through the levels of glycogenin-2 expression. Finally, the metabolism of glycogen is so closely linked to blood glucose homeostasis that any gene acting positively in the pathway of glycogen deposition is a candidate to be impaired in non-insulin-dependent diabetes mellitus. Thus, glycogenin-2 will be examined in this context and cloning of the glycogenin-2 gene is in progress.

Acknowledgments—We thank Drs. Robert A. Harris and Pengfei Wu for helpful discussions and for help in preparing rat liver samples.

REFERENCES

1. Preiss, J., and Walsh, D. A. (1981) in Biology of Carbohydrates (Ginsburg, V., and Robbins, P., eds) Vol. 1, pp. 199–314, John Wiley & Sons, New York
2. Skurat, A. V., and Roach, P. J. (1995) in Diabetics Mellitus: A Fundamental and Clinical Text (Leboit, D., Olefsky, J. E., and Taylor, S., eds) pp. 193–222, J. B. Lippincott Co., Philadelphia
3. Krisman, C. R., and Barrengo, R. (1975) Eur. J. Biochem. 52, 117–123
4. Poster, J., Smythe, C., and Cohen, P. (1988) Eur. J. Biochem. 176, 391–395
5. Lomako, J., Lomako, W. M., and Whelan, W. J. (1988) FASEB J. 2, 3097–3103
6. Roach, P. J., and Skurat, P. J. (1997) Prog. Nucleic Acids Res. Mol. Biol. 57, 289–316
7. Alonso, M. D., Lomako, J., Lomako, W. M., and Whelan, W. J. (1995) FASEB J. 9, 1126–1137
8. Smythe, C., and Cohen, P. (1991) Eur. J. Biochem. 200, 625–631
9. Cao, Y., Steinrauf, L. K., and Roach, P. J. (1995) Arch. Biochem. Biophys. 319, 293–298
10. Roach, P. J. (1996) FASEB J. 4, 2961–2968
11. Newgard, C. B., Hwang, P. K., and Fletterick, R. J. (1989) CRC Crit. Rev. Biochem. Mol. Biol. 24, 69–99
12. Thom, V. J., Khalil, M., and Cannon, J. F. (1993) J. Biol. Chem. 268, 7509–7513
13. Yang, B. Z., Ding, J. H., Inghild, J. J., Yao, Y., and Chen, Y. T. (1992) J. Biol. Chem. 267, 9259–9266
14. Campbell, D. G., and Cohen, P. (1989) Eur. J. Biochem. 185, 119–125
15. Viskupic, E., Cao, Y., Zhang, W., Cheng, C. DePaluRo-Agosh, A., and Roach, P. J. (1997) J. Biol. Chem. 272, 25759–25763
16. Smythe, C., Villar-Palasi, C., and Cohen, P. (1989) Eur. J. Biochem. 183, 205–209
17. Ercan, N., Gannon, M. C., and Nuttall, F. Q. (1994) J. Biol. Chem. 269, 22328–22333
18. Gannon, M. C., and Nuttall, F. Q. (1996) Trends Glycoconjug. Glycobiol. 6, 163–184
19. Cheng, C., Mu, J., Farkas, I., Huang, D., Goebel, M. G., and Roach, P. J. (1995) Mol. Cell. Biol. 15, 6632–6640
20. Sanger, F., Nicken, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467

---

3 H. Zong, J. Mu, and P. J. Roach, unpublished data.
21. Mu, J., Cheng, C., and Roach, P. J. (1996) J. Biol. Chem. 271, 26554–26560
22. Lawrence, J. C., Jr., Hiken, J. F., DePaoli-Roach, A. A., and Roach, P. J. (1983) J. Biol. Chem. 258, 10710–10719
23. Skurat, A. V., Lim, S.-S., and Roach, P. J. (1997) Eur. J. Biochem. 245, 147–155
24. Bradford, M. M. (1976) Anal. Biochem. 72, 248–254
25. Skurat, A. V., Wang, Y., and Roach, P. J. (1994) J. Biol. Chem. 269, 25534–25542
26. Laemmli, U. K. (1970) Nature 227, 680–685
27. Manzella, S., Ananth, S., Oegema, T. R., Roden, L., Rosenberg, L. C., and Meezan, E. (1995) Arch. Biochem. Biophys. 320, 361–368
28. Alonso, M. D., Lagzdins, E. J., Lomako, J., Lomako, W. M., and Whelan, W. J. (1995) FEBS Lett. 359, 110–112
29. Wang, Y., Camici, M., Lee, F. T., Ahmad, Z., DePaoli-Roach, A. A., and Roach, P. J. (1986) Biochim. Biophys. Acta 888, 225–236
30. Camici, M., DePaoli-Roach, A. A., and Roach, P. J. (1982) J. Biol. Chem. 257, 9898–9901
31. Calder, P. C., and Geddes, R. (1988) Biochem. Int. 17, 711–717
32. Blumenfeld, M. L., and Krisman, C. R. (1986) Eur. J. Biochem. 156, 163–169
33. Telmasky, D. S., and Krisman, C. R. (1996) Cell. Mol. Biol. 42, 589–598
34. Pitcher, J., Smythe, C., Campbell, D. G., and Cohen, P. (1987) Eur. J. Biochem. 169, 497–502
35. Smythe, C., Watt, P., and Cohen, P. (1990) Eur. J. Biochem. 189, 199–204
36. Stetten, M. R., and Katzen, H. M. (1961) J. Am. Chem. Soc. 83, 2912–2918
37. Cao, Y., Mahrenholz, A. M., DePaoli-Roach, A. A., and Roach, P. J. (1993) J. Biol. Chem. 268, 14687–14693