THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 273, No. 38, Issue of September 18, pp. 24754–24759, 1998

Printed in U.S.A.

The Sequence, Bacterial Expression, and Functional Reconstitution of the Rat Mitochondrial Dicarboxylate Transporter Cloned via Distant Homologs in Yeast and *Caenorhabditis elegans**

(Received for publication, May 8, 1998)

Giuseppe Fiermonte‡, Luigi Palmieri‡, Vincenza Dolce‡, Francesco M. Lasorsa‡, Ferdinando Palmieri‡, Michael J. Runswick§, and John E. Walker§¶

From the ‡Department of Pharmaco-Biology, Laboratory of Biochemistry and Molecular Biology, University of Bari, Via Orabona 4, 70125 Bari, Italy, and the §Medical Research Council, Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, United Kingdom

The dicarboxylate carrier (DIC) belongs to a family of transport proteins found in the inner mitochondrial membranes. The biochemical properties of the mammalian protein have been characterized, but the protein is not abundant. It is difficult to purify and had not been sequenced. We have used the sequence of the distantly related yeast DIC to identify a related protein encoded in the genome of *Caenorhabditis elegans*. Then, related murine expressed sequence tags were identified with the worm sequence, and the murine sequence was used to isolate the cDNA for the rat homolog. The sequences of the worm and rat proteins have features characteristic of the family of mitochondrial transport proteins. Both proteins were expressed in bacteria and reconstituted into phospholipid vesicles where their transport characteristics closely resembled those of whole rat mitochondria and of the rat DIC reconstituted into vesicles. As expected from the role of the DIC in gluconeogenesis and ureogenesis, its transcripts were detected in rat liver and kidney, but unexpectedly, they were also detected in rat heart and brain tissues where the protein may fulfill other roles, possibly in supplying substrates to the Krebs cycle.

The dicarboxylate carrier (DIC)[1] is a component of the inner membranes of mitochondria that plays an important role in gluconeogenesis, urea synthesis, and sulfur metabolism, particularly in liver, by transporting dicarboxylates such as malate and succinate across the inner membrane in exchange for phosphate, sulfate, sulfite, or thiosulfate (1–4). The transport and biochemical characteristics of the rat protein have been studied extensively (1–14), but because the protein is not abundant, neither its sequence nor the coding sequence has been established until now. To solve this problem, as described below, we took advantage of a protein encoded in *Saccharomyces cerevisiae* which has been shown to be a DIC (15, 16). Its sequence has features characteristic of members of a family of membrane proteins that are responsible for the transport of metabolites involved in oxidative phosphorylation and in other important functions in mitochondria. The yeast genome contains 35 members of this family (15). In all species that have been investigated, the sequences of family members are made of three related domains repeated in tandem about 100 amino acid in length, each probably being folded into two antiparallel transmembrane α-helices linked by an extensive hydrophilic sequence. The three repeats are joined together by shorter hydrophilic sequences. The repeats in the various family members are all related to each other, and several characteristic sequence features are conserved (17–20).

Comparisons of sequences of members of the family of known biochemical function in yeast with those of their mammalian counterparts show that the homologous pairs are too distant for the yeast sequences (protein or DNA) to provide a feasible basis for cloning the mammalian proteins. Therefore, the sequence of the *S. cerevisiae* DIC has been compared with the proteins encoded in the available genomic sequences of *Caenorhabditis elegans*. Currently, it encodes 32 members of the mitochondrial transporter family, and the two most closely related proteins (named K11G12.5 and B0432.4) were identical to the yeast DIC in 37% and 30% of their residues, respectively. The K11G12.5 protein has been expressed in *Escherichia coli*, reconstituted into phospholipid vesicles, and demonstrated to be a DIC. Then the sequence of the *C. elegans* DIC was compared with protein sequences encoded in mammalian expressed sequence tags. In this way, two partial murine expressed sequence tags (AA199557 and AA041737) were found to encode a fragment of a closely related protein, and synthetic oligonucleotides based on their sequences were used to amplify a DNA segment from rat liver cDNA. Using this segment, the complete rat cDNA was then obtained by PCR experiments. The encoded protein was shown to have transport specificity and other biochemical characteristics that were identical to those of the DIC isolated from rat liver mitochondria.

EXPERIMENTAL PROCEDURES

Reagents and Biochemicals—Thermus aquaticus DNA polymerase, T4 DNA ligase, and restriction endonucleases were obtained from Boehringer Mannheim Biochemicals. Amersham Italy S.r.l. (Milan, Italy) supplied L-[1,4(2, 3)-^{14}C]malic acid and [^{33}P]. Cardiolipin and sarkosyl (N-lauroylsarcosine) were purchased from Sigma Chemical Company (Milan, Italy). All other reagents were of the highest purity commercially available.

Synthetic Oligonucleotides—Adaptor primers AP1, AP2, and CDS were purchased from CLONTECH (Palo Alto, CA). Oligonucleotides were made with a DNA synthesizer (Applied Biosystems model 320B). Some of them had additional 5′-linkers and were employed as primers in PCRs. Others were used as hybridization probes for the reaction products, as described before (21). Some unique 17-base oligonucleo-

* This work was supported in part by the Consiglio Nazionale delle Ricerche Target Project "Biotechnology" and by the Ministero dell' Universitá e della Ricerca Scientifica e Technologica. 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) AJ223355.

¶ To whom correspondence should be addressed. E-mail: walker@mrc-lmb.cam.ac.uk.

[1] The abbreviations used are: DIC, dicarboxylate carrier protein; PCR, polymerase chain reaction; kb, kilobase pair(s); PIPES, piperazine-*N*,*N*′-bis(2-ethanesulfonic acid).

tides were used as sequencing primers.

DNA Sequence Analysis—An automatic DNA sequenator (Applied Biosystems, Prism 377) was employed with dye terminator kits (Perkin-Elmer). Sequence errors from PCRs were avoided by sequencing at least three independent clones in both directions from each reaction. Ambiguities were resolved by sequencing more clones. Data bases were compiled and analyzed with the program Autoassembler (Perkin-Elmer).

Rat cDNA Fragments for the DIC—Touchdown PCRs (22) were conducted in a thermal cycler (Perkin-Elmer model 480). The reaction mixtures (70 μl) contained adaptor-ligated double-stranded rat liver cDNA (1 ng, CLONTECH) with primers (final concentration 0.5 μM). After an initial incubation at 94 °C for 2 min without *Taq* polymerase, the PCR cycle was denaturation at 94 °C (30 s) and annealing at 72 °C (4 min). These steps were repeated 5 times. Then the cycle was repeated for 5 and 20 times with both annealing and extension temperatures at 70 °C or 68 °C, respectively. The entire procedure was repeated once again with each of the previous final PCR mixtures (1 μl) as template, with two nested primers (final concentration 1 μM), and 2 units of *Taq* polymerase. The cDNA sequence was extended to the 5′-end with primers AP1 and AP2 (modified by substituting the *Not*I restriction site by *Eco*RI) and to the 3′-end with the primer CDS. The products were identified and recovered from agarose gels as described before (21). The *Eco*RI and *Bam*HI sites in the linkers were digested, and the products were cloned into the pUC19 vector.

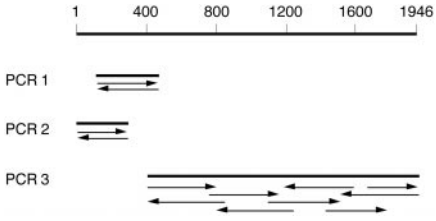

FIG. 1. **Generation by PCR and sequence analysis of clones encoding the DIC from rat liver.** The *heavy horizontal lines* represent the cDNA segments generated in PCRs 1–3. The *arrows* indicate the directions and the extents of the determined DNA sequences. The scale is in bases.

Construction of the Rat and C. elegans DIC Expression Plasmids—Full-length cDNAs for the rat and *C. elegans* DIC were amplified by 60 or 30 cycles of PCR, respectively, as described before (21), except that the annealing temperature was 60 °C. The templates were rat liver cDNA (CLONTECH) and for *C. elegans*, phage DNA (1 ng) from clone cm15f11 (23). The forward and reverse primers were nucleotides 34–55 and 869–894 of the rat cDNA (see Fig. 2), with 5′-*Nde*I and *Eco*RI restriction sites, respectively, and nucleotides 15–38 and 863–886 of clone cm15f11, with additional 5′-*Nde*I and *Bam*HI restriction sites, respectively. The amplified products, 0.9 kb in length, were purified (with Qiaquick from Qiagen), cloned into the pMW7 expression vector, and the constructs were transformed into *E. coli* DH5α cells. Transformants selected on 2xTY plates containing ampicillin (100 μg/ml) were screened by direct colony PCR and by restriction digestion of purified plasmids. The sequences of inserts were verified.

Bacterial Expression—The rat and *C. elegans* DICs were overexpressed at 37 °C in *E. coli* CO214(DE3), a mutant of *E. coli* C41(DE3) selected for its ability to overexpress the mitochondrial phosphate carrier without toxicity,[2] as described before for improvement of host strains (24). Control cultures with the empty vector were processed in parallel. Inclusion bodies were purified according to Fiermonte *et al.* (25). The pellets were washed twice with buffer containing 10 mM PIPES, pH 7.0, Triton X-114 (3%, w/v), and 1 mM EDTA and then once with TE buffer. Then the proteins were solubilized in a solution of 10 mM Tris-HCl, pH 7.0, sarkosyl (1.8%, w/v), and 0.1 mM EDTA. A small insoluble residue was removed by centrifugation at 258,000 × *g* for 1 h at 4 °C.

Samples taken at various points of growth of bacterial cultures and purified inclusion bodies were examined by SDS-polyacrylamide gel electrophoresis in 17.5% gels (26). Proteins were either stained with

[2] G. Fiermonte and J. E. Walker, unpublished work.

```
                                                                M  A  E  A  R  T  S  R  W  Y  F  G  G  L  A  S  C  G  A  A  C  C  T  H  P  L  D  L  L   29
CCGGGGTCGGGCCAGGTCGCTGCTGCTCTGGCCATGGCTGAGGCACGCACGTCTCGCTGGTACTTTGGAGGGCTGGCTTCCTGCGGAGCTGCCTGCTGCACGCACCCTCTAGACCTGCTC                    120
-----→ AP1/AP2                                                                                                  Primer 1F ⟶

 K  V  H  L  Q  T  Q  Q  E  V  K  L  R  M  T  G  M  A  L  Q  V  V  R  T  D  G  F  L  A  L  Y  N  G  L  S  A  S  L  C  R   69
AAGGTGCATTTGCAGACCCAACAGGAGGTGAAGCTTCGAATGACTGGAATGGCACTGCAGGTGGTGCGAACCGATGGCTTCCTGGCACTCTACAATGGCCTGAGTGCCTCGCTGTGCAGG                    240
Primer 2F ⟶

 Q  M  T  Y  S  L  T  R  F  A  I  Y  E  T  M  R  D  Y  M  T  K  D  S  Q  G  P  L  P  F  Y  S  K  V  L  L  G  G  I  S  G   109
CAGATGACCTACTCCCTGACCCGGTTCGCTATCTACGAGACCATGCGGGACTACATGACCAAGGACAGCCAGGGGCCTCTCCCCTTCTACAGCAAGGTGTTGCTGGGCGGCATCAGTGGT                    360
                    ⟵ Primer4R                                                                  ⟵ Primer 3R

 L  T  G  G  F  V  G  T  P  A  D  L  V  N  V  R  M  Q  N  D  M  K  L  P  L  S  Q  R  R  N  Y  S  H  A  L  D  G  L  Y  R   149
TTAACTGGAGGCTTCGTGGGGACCCCAGCAGATTTGGTCAATGTCAGGATGCAGAATGACATGAAGCTGCCTCTGAGCCAGCGCCGCAACTACTCTCACGCCCTGGATGGTCTGTACCGT                    480
          Primer 3F ⟶                Primer 4F ⟶                                    ⟵ Primer 2R

 V  A  R  E  E  G  L  K  K  L  F  S  G  A  T  M  A  S  S  R  G  A  L  V  T  V  G  Q  L  S  C  Y  D  Q  A  K  Q  L  V  L   189
GTAGCCCGTGAAGAAGGCCTGAAGAAGCTCTTCTCTGGAGCAACTATGGCATCCAGCCGTGGGGCCCTCGTCACTGTGGGCCAGCTGTCCTGCTACGACCAGGCCAAGCAGCTGGTTCTC                    600
                              ⟵ Primer 1R

 S  T  G  Y  L  S  D  N  I  F  T  H  F  L  S  S  F  I  A  G  G  C  A  T  F  L  C  Q  P  L  D  V  L  K  T  R  L  M  N  S   229
AGTACTGGGTACCTGAGTGACAACATTTTCACCCACTTCCTCTCCAGTTTCATTGCGGGCGGATGTGCCACATTTCTCTGCCAGCCCCTGGACGTGCTGAAGACTCGCTTGATGAACTCC                    720

 K  G  E  Y  Q  G  V  F  H  C  A  V  E  T  A  K  L  G  P  Q  A  F  F  K  G  L  V  P  A  G  V  R  L  V  P  H  T  V  L  T   269
AAGGGCGAGTATCAGGGTGTTTTCCATTGTGCCGTGGAGACAGCAAAGCTTGGACCACAGGCCTTTTTCAAGGGTCTCGTTCCTGCGGGCGTCCGTCTCGTCCCCCACACCGTGCTTACC                    840

 F  M  F  L  E  Q  L  R  K  H  F  G  I  K  V  A  T  *                                                          286
TTCATGTTCCTGGAGCAGCTGCGGAAGCACTTCGGTATCAAAGTGGCAACCTGACACCGCCAGGGACACCTGGGCTGCGCTCAGTCGCTGTGCTGAGCCCCTTGGAAGAGTAGGAAGGGA                    960
GTGGGCTCCCTTCCTCCTTGGCCTGGGCCCGTGCTGATCCCCAGAAGATTCCTGTTCTTCCCACCCTTGGGCTGCTGGCTACGCCTTCCGACCCTGCCTTGGCCCCACTGAAGTGGCACC                   1080
TCTGCCCTACTGGCTCCCAGGCTCTCCCCACTGGGACGTCCCATCTTCCTACCCGATGATTCACTCAGAAGAGGTCTGGCCTGGCTGGTGTCACTGTCCCCACCTCCCTGGCTGCTACCC                   1200
TGCCCTGCCTGTCAAGCCCGGAGAAGCTGAGTACACTTCCTGCTCTTGCTGGCCCACTGAACAGCTTCCACCCAGGACTTGGGTAGTGTGGCAGGGTACAGCCCCTGGCAGCTTCTGCTT                   1320
ACCAAATGACTAGAGCACGCACGCACACACATACACACATGCACTTACGTCACAAGAGGGACCACCGTGCAGTGTTCTGGAAGGTAGTGCCTTCAGGAGAGGGAACAGGGCAGCAGCACA                   1440
GATCACCAGCAAACGCCATGACCGCGGAGTCCGGAGAAAGTACCTGGGGTTCCCTGTGCACCTCCCGTGCTCAGCCTTGGCTGCTTTAAGAGTCAGTTTTTGCGGAACCCTCCTACTCAG                   1560
CAGCCACTACCATCATCAGGAACTGGCTGTGTCCCTCATACACCTGTGCCCTCCCTTGCCTGGCTTCCCCCAGGGCCAGGGGCTCATGCTGGAGCTGGGGCTATGATGGATTCATTTTAT                   1680
GTGTCCTCCTAGGACCTGTTTTCCCGTATGGCAGGGGTCACCCCTCACCATCCCTCAGGGTTCAAAGCAGCCTGTTTTCCCTCAAATTGGGGTTGTGTGTATATCAACGTGAGGTTGGGC                   1800
CCTGTGCCCGTTTGCCTCCCCTCAGCACCAGGCCCTCACCCAGGAGGTGCGGGAGGTGGCAGTGGGCCCAGGGGCACTGCCTACATCTCTCTTCAGGATCTAATAAACCAAGTGGCCAAA                   1920
AAAAAAAAAAAAAAAAAAAAAAAAAA                                                                                                                1946
        ⟵ Primer CDS
```

FIG. 2. **Sequence of the cDNA for the DIC from rat liver mitochondria and the encoded protein sequence.** The amino acids are numbered from 1 to 286. An *asterisk* denotes the stop codon. Unique primer sequences are *boxed*. *Horizontal arrows* pointing to the *left* or *right* indicate that the primers were synthesized as either the sequence shown or its complement, respectively. The *underlined* sequence is a potential polyadenylation signal.

```
C.elegans       MAEDKTKRLGR-----WYFGGVAGAMAACCTHPLDLLKVQLQTQQQGKLTIGQLSLKIYKNDGILAFYNGVSASVLRQLTYSTTRFGIYETVKKQLPQD-Q
M. musculus     ---MAEARASR-----WYFGGLASCGAACCTHPLDLLKVHLQTQQEVKLRMTGMALQVVRTDGFLALYNGLSASLCRQMTYSLTRFAIYETMRDYMTKDSQ
Rat             ---MAEARTSR-----WYFGGLASCGAACCTHPLDLLKVHLQTQQEVKLRMTGMALQVVRTDGFLALYNGLSASLCRQMTYSLTRFAIYETMRDYMTKDSQ
S. cerevisiae   MSTNAKESAGKNIKYPWWYGGAAGIFATMVTHPLDLAKVRLQAAPMPKPTLFRMLESILANEGVVGLYSGLSAAVLRQCTYTTVRFGAYDLLKENVIPREQ
                          •     * •** *   *•  ****** **.***.    *  •  •   •  ••* •  * *.**•. ** **•  **  *• ••  •    *

C.elegans       -PLPFYQKALLAGFAGACGGMVGTPGDLVNVRMQNDSKLPLEQRRNYKHALDGLVRITREEG-FMKMFNGATMATSRAILMTIGQLSFYDQIKQTLISSG-
M. musculus     GPLPFYNKVLLGGISGLTGGFVGTPADLVNVRMQNDMKLPPSQRRNYSHALDGLYRVAREES-LRKLFSGATMASSRGALVTVGQLSCYDQAKQLVLSTG-
Rat             GPLPFYSKVLLGGISGLTGGFVGTPADLVNVRMQNDMKLPLSQRRNYSHALDGLYRVAREEG-LKKLFSGATMASSRGALVTVGQLSCYDQAKQLVLSTG-
S. cerevisiae   -LTNMAYLLPCSMFSGAIGGLAGNFADVVNIRMQNDSALEAAKRRNYKNAIDGVYKIYRYEGGLKTLFTGWKPNMVRGILMTASQVVTYDVFKNYLVTKLD
                         .*  **  *•  *.**.*****  *   .**** .*.**•. •• * *     .* *      *  *.*  *•  **  *• •••

C.elegans       VAEDNLQTHFASSISAASVATVMTQPLDVMKTRMMNAAPGEFKGILDCFMFTA-KLGPMGFFKGFIPAWARLAPHTVLTFIFFEQLRLK-FGYAPPVKA
M. musculus     YLSDNIFTHFVSSFIAGGCATFLCQPLDVLKTRLMNSK-GEYQGVFHCAMETA-KLGPQAFFKGLFPAGIRLIPHTVLTFMFLEQLRKH-FGIKVPT--
Rat             YLSDNIFTHFLSSFIAGGCATFLCQPLDVLKTRLMNSK-GEYQGVFHCAVETA-KLGPQAFFKGLVPAGVRLVPHTVLTFMFLEQLRKH-FGIKVAT--
S. cerevisiae   FDASKNYTHLTASLLAGLVATTVCSPADVMKTRIMNGS-GDHQPALKILADAVRKEGPSFMFRGWLPSFTRLGPFTMLIFFAIEQLKKHRVGMPKEDK-
                      **  .*  *   ** .. * **.***.** **   *•. •       •  * **   *.*  *•  ** * *.* *   ***. •  *
```

FIG. 3. **Alignment of *C. elegans*, *M. musculus*, rat, and *S. cerevisiae* DIC proteins.** *Asterisks* and *dots* indicate residues in all four sequences which are identical and conserved, respectively.

Coomassie Blue dye or transferred to polyvinylidene difluoride membranes, stained with Coomassie Blue dye, and their NH_2 termini sequenced (Applied Biosystems, model 477A). The yield of purified protein was estimated by laser densitometry (LKB 2202 Ultroscan) of SDS-polyacrylamide gels stained with Coomassie Blue dye, with carbonic anhydrase as standard.

Reconstitution into Liposomes—Sarkosyl-soluble proteins were diluted 10-fold with a buffer containing 2 mM PIPES, pH 7.0, 0.6% (w/v) Triton X-114, and 0.2 mM EDTA and reconstituted into liposomes in the presence of appropriate substrates, as described before (25), except that the reconstitution mixture contained cardiolipin (1 mg/ml). The extent of incorporation of the protein into liposomes was estimated as follows. The proteoliposomes were passed through a Sephadex G-75 column, centrifuged at 300,000 × *g* for 30 min, precipitated with cold acetone, and delipidated with organic solvents (27). The protein was dissolved in 0.1% SDS, and after SDS-polyacrylamide gel electrophoresis, the amount of the DIC was measured by laser densitometry of stained samples, as described above.

Activity Assays—External substrates were removed from proteoliposomes on Sephadex G-75 (28). Transport at 25 °C was started by adding [^{14}C]malate or ^{33}P to the eluted proteoliposomes and, unless otherwise indicated, terminated after 20 s (*C. elegans* DIC) or 40 s (rat DIC) by the addition of either 30 mM butylmalonate or a mixture containing 30 mM pyridoxal 5′-phosphate and 10 mM bathophenanthroline, (this is the "inhibitor-stop" method (28)). In controls, the inhibitor was added with the labeled substrate. Extraliposomal labeled substrate was removed from quenched samples on Dowex AG1-X8, and eluted radioactivity was measured (28). Transport activities were calculated from the experimental values minus the controls. Various other transport activities were also assayed by the inhibitor-stop method.

Northern and Southern Blotting—Total RNA was extracted from rat liver, kidney, heart, and brain according to Chirgwin *et al*. (29). The hybridization experiments were performed as described before (30), except that the probe was nucleotides 34–894 of the rat cDNA sequence (see Fig. 2). Samples were autoradiographed at −70 °C for 15 days with an intensifying screen. Hybridization signals were normalized with an 850-base pair probe encoding part of rat glyceraldehyde-3-phosphate dehydrogenase (31) employed under the same conditions as described above.

Samples (10 μg) of rat liver genomic DNA (32) were digested for 6 h at 37 °C with one of the restriction enzymes *Nco*I, *Eco*RI, or *Bam*HI. DNA fragments were separated on 0.8% agarose gels, transferred to a Hybond-N membrane (33), and hybridized with a probe from the cDNA for the rat DIC (nucleotides 91–536; see Fig. 2) using the same procedures described for Northern blotting.

Homology Searches—Homologs of the yeast DIC in *C. elegans* were identified by screening the nematode (the Sanger Center, Hinxton, U. K.) and the National Center for Biotechnology Information data bases with the sequence of the *S. cerevisiae* protein (15) using the BLASTP program. The EMBL nonredundant expressed sequence tag data base was probed with a protein sequence encoded in the nematode clone K11G12.5 using the TBLASTN program. The same program was used to search all available sequence data bases for murine fragments coding for protein sequences related to the putative rat DIC protein sequence. Amino acid sequences were aligned with ClustalW (version 1.7).

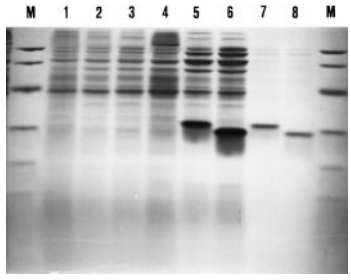

FIG. 4. **Expression of *C. elegans* and rat DIC proteins in *E. coli*.** Proteins were separated by SDS-polyacrylamide gel electrophoresis and stained with Coomassie Blue dye. *Lanes M*, markers (rabbit muscle phosphorylase *b*, bovine serum albumin, ovalbumin, bovine carbonic anhydrase, soybean trypsin inhibitor, and from hen egg lysozyme); *lanes 1–6*, *E. coli* C0214(DE3) containing the expression vector, without (*lanes 1* and *4*), with the coding sequence of the *C. elegans* DIC (*lanes 2* and *5*), and with the coding sequence of the rat DIC carrier (*lanes 3* and *6*). Samples were taken immediately before induction (*lanes 1–3*) and 5 h later (*lanes 4–6*). The same number of bacteria was analyzed in each sample. *Lane 7*, purified *C. elegans* DIC (8 μg); *lane 8*, purified recombinant rat DIC (6.5 μg) obtained from the bacteria analyzed in *lanes 5* and *6*, respectively.

RESULTS AND DISCUSSION

Homologs of the S. cerevisiae DIC—The nematode gene product K11G12.5 is identical to the yeast DIC sequence (15, 16) and to the bovine oxoglutarate-malate carrier proteins (34) in 37 and 35% of its residues, respectively. However, the bovine oxoglutarate-malate carrier protein is 64% identical to the *C. elegans* protein B0432.4, which is therefore more likely to be the worm oxoglutarate-malate carrier protein than K11G12.5. Two overlapping cDNA clones (AA199557 and AA041737) from *Mus musculus* contain 622 nucleotides of unique sequence, and the encoded protein sequence is 47% identical to part of the *C. elegans* K11G12.5 protein.

Cloning and Sequencing of the cDNA for the Rat DIC—Four nested oligonucleotides (1F, 2F, 1R, 2R) based on the sequence of the AA199557 and AA041737 cDNA sequences from *M. musculus* were used to amplify a segment from a rat liver cDNA (see PCR1 in Figs. 1 and 2). The product gave a single band of about 350 base pairs. Its nucleotide sequence is 94% identical to the *M. musculus* partial sequence. The rat cDNA sequence was completed by experiments PCR2 and PCR3 (see Fig. 1).

The final rat cDNA sequence of 1946 nucleotides (see Fig. 2) encodes a protein with a molecular mass of 31,455. The assignment of the proposed start codon is supported by the closely related murine sequence where the corresponding codon is preceded by an in-frame stop codon. The fragment of murine

TABLE I

Dependence on internal substrate of the transport properties of proteoliposomes containing bacterially expressed C. elegans K11G12.5 protein or its rat homolog

Proteoliposomes reconstituted with the *C. elegans* K11G12.5 protein or its rat homolog were preloaded internally with various substrates (concentration 20 mM). Transport was started by the external addition of 0.5 mM [^{14}C]malate or 1.5 mM ^{33}P. Transport was measured over 20 s or 40 s in proteoliposomes reconstituted with the *C. elegans* or rat protein, respectively. Similar results were obtained in four different experiments. Values shown are mmol of substrate transported/min/g of protein.

Internal	*C. elegans* K11G12.5		Rat homolog	
	[^{14}C]Malate	^{33}P	[^{14}C]Malate	^{33}P
	mmol/min/g protein			
None (Cl^- present)	1.5	2.4	0.9	0.5
Malate	66.4	78.0	17.8	20.4
Phosphate	85.2	93.2	18.7	24.9
Malonate	65.6	76.4	16.7	18.6
Succinate	53.5	57.1	14.5	13.3
Sulfate	51.4	53.4	13.3	12.1
Thiosulfate	50.1	45.1	11.6	12.6
Oxoglutarate	3.7	4.1	1.0	0.9
Fumarate	1.6	0.8	1.4	0.2
Aspartate	1.1	2.7	1.4	0.5
Glutamate	0.7	2.3	0.3	0.6
Citrate	0.8	3.5	0.4	0.5
ATP	1.5	3.2	1.2	0.8
Pyruvate	0.9	1.5	0.4	0.5
Carnitine	1.1	1.9	0.6	0.4
Ornithine	1.8	2.1	0.8	0.5

protein sequence (132 amino acids) is 96% identical to the corresponding part of the rat sequence. The 5′- and 3′-untranslated regions are 33 and 1055 base pairs, respectively, and the sequence is terminated by a run of A residues separated by a 13-nucleotide sequence from the preceding AATAAA sequence, a typical signal for polyadenylation of mRNA (32).

Recent additions to sequence data bases have allowed us to assemble the entire coding sequence of the murine protein. In the alignment of the four homologous protein sequences (Fig. 3), the putative rat DIC is 96, 59, and 37% identical to the murine, *C. elegans*, and *S. cerevisiae* proteins, respectively.

Bacterial Expression of the Putative DIC Proteins from C. elegans and Rat—The overexpression of the *C. elegans* K11G12.5 protein and its rat homolog in *E. coli* C0214(DE3) (see Fig. 4, *lanes 5* and *6*) yielded about 60 and 45 mg of purified protein per liter of culture, respectively. Both purified proteins gave single bands on SDS-polyacrylamide gels (Fig. 4, *lanes 7* and *8*) with apparent molecular masses of 32 and 29.5 kDa, respectively (the corresponding calculated molecular masses including the initiator methionine are 32.23 and 31.55). Neither protein was detected in bacteria harvested immediately before induction of expression (*lanes 1–3*) nor in cells harvested after induction but lacking the coding sequence in the expression vector (*lane 4*). The determined NH_2-terminal sequences (AEDKTKRLGR and AEARTSRWYF, respectively) of residues 1–10 of the purified 32- and 29.5-Da proteins are identical to those predicted for residues 2–11 of the *C. elegans* K11G12.5 protein and its rat homolog, respectively.

Transport Properties of the C. elegans K11G12.5 Protein and Its Rat Homolog—The proportion of rat and *C. elegans* proteins incorporated into liposomes in various experiments was usually in the range of 11–18% of the protein used in the experiment, and they were incorporated to approximately the same extent. Both proteins catalyzed a [^{14}C]malate/phosphate exchange that was inhibited by butylmalonate (see Table I), as does the DIC in rat mitochondria (1–8). No such activity was found by reconstitution of sarkosyl-solubilized material from bacterial cells either lacking the expression vector, or harvested immediately before induction of overexpression. Neither reconstituted protein catalyzed the homoexchange of pyruvate, citrate, oxoglutarate, glutamate, aspartate, glutamine, carnitine, ornithine, ADP or ATP, all being substrates for various mitochondrial transporters.

TABLE II

Effect of inhibitors and externally added substrates on the [^{14}C]malate/phosphate exchange by proteoliposomes containing the C. elegans K11G12.5 protein or its rat homolog

Proteoliposomes were preloaded with 20 mM phosphate, and transport was started by adding 0.5 mM [^{14}C]malate. The incubation time was 20 s and 40 s for the reconstituted *C. elegans* and rat proteins, respectively. Thiol reagents, pyridoxal 5′-phosphate, carboxyatractyloside, and α-cyanocinnamate were added 2 min before the labeled substrate; the other inhibitors and external substrates were added together with [^{14}C]malate. All inhibitors and substrates were used at a concentration of 10 mM except for organic mercurials, carboxyatractyloside and α-cyanocinnamate (0.1 mM) and *N*-ethylmaleimide (0.1 or 1 mM). Similar results were obtained in three independent experiments.

Reagents	Malate transported	
	C. elegans K11G12.5	Rat homolog
	mmol/min/g protein	
Experiment 1		
None	67.3	17.0
Butylmalonate	7.7	2.1
Benzylmalonate	11.6	2.8
Phenylsuccinate	23.3	6.1
Mersalyl	22.7	0.3
p-Chloromercuribenzenesulfonate	29.7	0.8
N-Ethylmaleimide (0.1 mM)	64.5	15.1
N-Ethylmaleimide (1.0 mM)	60.3	9.7
Bathophenanthroline	0.6	0
Pyridoxal 5′-phosphate	2.7	0.2
Carboxyatractyloside	59.3	14.6
Phthalonate	63.1	15.3
α-Cyanocinnamate	63.5	15.2
Experiment 2		
None	63.9	19.3
Malate	2.1	1.6
Phosphate	8.1	2.7
Malonate	2.3	1.1
Succinate	5.8	2.0
Sulfate	7.0	2.3
Thiosulfate	6.9	2.2
Oxoglutarate	52.7	16.7
Fumarate	57.7	17.7
Citrate	53.7	16.6
ADP	54.8	17.0
Aspartate	57.9	17.6
Glutamate	58.1	17.5
Glutamine	56.6	17.2

The substrate specificity of the *C. elegans* and rat recombinant proteins was examined in greater detail by measuring the uptake of [^{14}C]malate and ^{33}P into proteoliposomes that had been preloaded with a variety of substrates (Table I). In the absence of internal substrate, uptake of labeled malate or phosphate was not observed. The highest activities were detected in the presence of internal L-malate, phosphate, and malonate. To a somewhat lower extent, succinate, sulfate, and thiosulfate were exchanged for both external malate and phosphate. Virtually no exchange was observed with internal oxoglutarate, fumarate, glutamate, aspartate, citrate, ATP, pyruvate, carnitine, or ornithine. The residual activity in the presence of these substrates was approximately the same as the activity observed in the presence of sodium chloride. Therefore, proteoliposomes reconstituted with the *C. elegans* and rat recombinant proteins import external malate and phosphate only in exchange for certain dicarboxylates, inorganic sulfur-containing anions, and phosphate, similar to the rat DIC in liver mitochondria (1–8).

The [^{14}C]malate/phosphate exchange reaction of both proteins was inhibited by the substrate analogs butylmalonate,

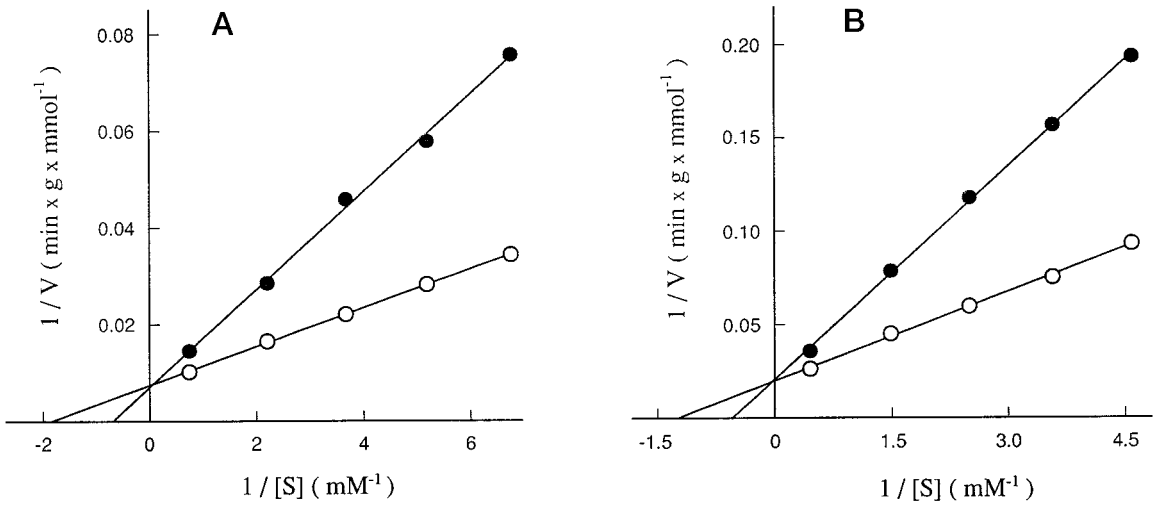

FIG. 5. **Lineweaver-Burk plots of the [^{14}C]malate/malate and ^{32}P/malate exchanges in proteoliposomes reconstituted with the *C. elegans* and the rat DIC proteins.** Radioactive malate (○) or phosphate (●) was added to proteoliposomes containing 10 mM malate reconstituted with the *C. elegans* protein (*panel A*) and the rat protein (*panel B*).

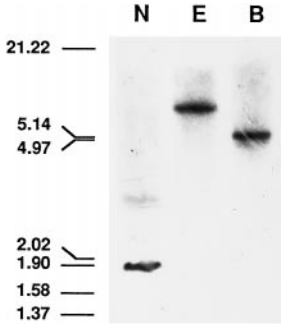

FIG. 6. **Hybridization of restriction digests of rat genomic DNA with a probe derived from the cDNA for the rat DIC.** The DNA was digested with the restriction enzymes *Nco*I, *Eco*RI, and *Bam*HI, denoted by *N*, *E*, and *B*, respectively. The probe was nucleotides 91–536 (see Fig. 2 and "Experimental Procedures"). The positions of the DNA markers (in kb) are shown at the *left*.

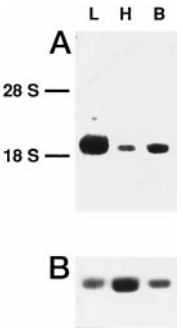

FIG. 7. **Expression of the DIC in rat tissues.** The following hybridization probes were used in Northern blot analyses: a 0.85-kb probe from the rat DIC cDNA (*panel A*) and a 0.85-kb probe from a rat cDNA encoding glyceraldehyde 3-phosphate dehydrogenase (*panel B*). The RNA samples were from liver (*L*), heart (*H*), and brain (*B*). About 20 μg of total RNA was loaded in each slot. After 15 days of autoradiography at −70 °C, the blot in *A* was stripped by boiling in 0.1% (w/v) SDS for 3 min and rehybridized with the glyceraldehyde 3-phosphate dehydrogenase probe. On the *left* side of *panel A* are the positions of 28 S and 18 S ribosomal RNAs (4,800 and 1,900 nucleotides in length, respectively).

benzylmalonate, and phenylsuccinate (Table II). Butylmalonate was more effective than phenylsuccinate, as with the DIC in mitochondria (3, 6), whereas the reverse is true for the oxoglutarate-malate carrier protein (35). The activities of both recombinant proteins were also inhibited by organic mercurials, pyridoxal 5′-phosphate and bathophenanthroline, all of them inhibitors of the rat DIC, both in mitochondria (3, 5, 8, 36) and after purification and reconstitution into liposomes (9, 10, 13, 14). The yeast DIC also has similar properties (15, 16, 37). Inhibitors of other characterized mitochondrial carriers, such as carboxyatractyloside, phthalonate, and α-cyanocinnamate, had little or no effect on the activities of the *C. elegans* and the rat reconstituted proteins.

In contrast to the recombinant *C. elegans* protein, the rat protein is more sensitive to organic mercurials and is also inhibited by 1 mM *N*-ethylmaleimide to some extent. The rat DIC, both in mitochondria and after purification, and the yeast DIC, are insensitive to *N*-ethylmaleimide (2–5, 8–10, 14–16, 37), but this point requires further clarification. There are at least two possible explanations for this partial inhibition of the rat recombinant protein by *N*-ethylmaleimide. First, the expressed rat protein might differ from the native DIC by the absence of disulfide bridges (although there is no evidence for their presence in the native rat protein). Second, both rat liver mitochondria and the DIC purified from mitochondria might contain more than one isoform of DIC, one being relatively insensitive to *N*-ethylmaleimide.

The ability of nonradioactive potential substrates to inhibit the reconstituted [^{14}C]malate/phosphate exchange of both the *C. elegans* and rat recombinant proteins was also examined. The exchange was prevented by external addition of any one of the known substrates of the DIC (Table I), and it was unaffected by substrates of other mitochondrial carriers, such as fumarate, oxoglutarate, ADP, or citrate.

The dependence of the [^{14}C]malate/malate and ^{33}P/malate exchange rates on substrate concentration was investigated by changing the concentration of externally added [^{14}C]malate or ^{33}P at a constant internal concentration of 10 mM malate (see Fig. 5). In liposomes reconstituted with either the *C. elegans* or the rat protein (Fig. 5, *A* and *B*) linear functions were obtained which intersected the ordinate close to a common point. The different slopes indicate that the half-saturation constant (K_m) for phosphate is about three times higher than for malate (in liposomes reconstituted with the *C. elegans* protein, the mean values from four experiments are 0.54 ± 0.10 and 1.52 ± 0.36 mM for malate and phosphate, respectively; for the rat protein the corresponding values are 0.78 ± 0.14 and 1.77 ± 0.27). Thus, both the *C. elegans* and rat recombinant proteins are also similar in these respects to the rat liver DIC (K_m values for malate and phosphate are 0.26 mM and 1.5 mM, respectively, in

mitochondria (2) and 0.49 mM and 1.41 mM, respectively, after purification (11)). The K_m values of the recombinant yeast DIC for malate and phosphate are 0.59 and 1.65 mM, respectively (15). In four experiments, the V_{max} value, corrected for small differences in efficiency of reconstitution, was 149.3 ± 38.6 mmol/min/g of *C. elegans* protein and 44.7 ± 12.7 mmol/min/g of rat protein. These values are virtually the same for both the malate/malate and phosphate/malate exchanges and are independent of the type of substrate, as observed before (2, 11, 15). Furthermore, the V_{max} values of the recombinant rat and *C. elegans* proteins are higher than those determined previously for the purified rat liver and recombinant yeast DIC (11, 15). This difference may arise because the efficiency of reconstitution was not quantified in the earlier studies.

Repetitive Sequences within the DIC and Hydrophobicity Analysis—The sequences of the DICs consist of three homologous tandem repeats that are distantly related in sequence to repeats in other mitochondrial carrier proteins. Each repetitive element appears to be made of two antiparallel transmembrane hydrophobic α-helices linked by an extensive polar region. Therefore, they have the tripartite structure that is characteristic of other mitochondrial transporters (17–20, 38). Each of the three domains contains the sequence motif characteristic of the mitochondrial carriers, namely P-h-D/E-*X*-*h*-K/R-*X*-R/K-(20–30 residues)-D/E-G-(4 residues)-*a*-K/R-G (where *h* and *a* represent hydrophobic and aromatic amino acids, respectively, and *X* is unknown).

Number of Genes for the DIC in Rat—Some mitochondrial carriers are unique proteins encoded by single genes, whereas others have closely related isoforms either encoded by multiple genes (39–48) or generated by alternative splicing from a single gene (21). Because only single hybridizing bands were detected by Southern hybridization (see Fig. 6), it seems likely that the rat DIC has a single gene, although this experiment would not detect isoforms arising by alternative splicing or isoforms with distantly related coding sequences.

Expression of the DIC in Rat—Among the tissues that were examined, the rat DIC was expressed most strongly in liver. Somewhat lower levels were detected in brain, and much less was found in heart (see Fig. 7). It is notable that the DIC transcript, estimated to be about 2.1 kb, is larger than the cDNA sequence that has been determined, indicating that the transcript may have an extensive 5′-noncoding region. High levels of mRNA for the DIC in liver and kidney (not shown) are probably related to gluconeogenesis and ureogenesis, which occur mainly in these tissues. At present, there is no explanation for the relatively high levels of the DIC mRNA in brain and for its presence in heart where gluconeogenesis and ureogenesis do not operate. So the DIC may have other functions in these tissues, possibly exclusively catalyzing the supply of substrates to the Krebs cycle. It remains to be demonstrated that the DIC mRNA is translated in heart and brain.

REFERENCES

1. Chappel, J. B., and Haarhoff, K. N. (1967) in *Biochemistry of Mitochondria* (Kaniuga, Z., Slater, E. C., and Wojtczak, L., eds) pp. 75–91, Academic Press, London
2. Palmieri, F., Prezioso, G., Quagliariello, E., and Klingenberg, M. (1971) *Eur. J. Biochem.* **22,** 66–74
3. Crompton, M., Palmieri, F., Capano, M., and Quagliariello, E. (1974) *Biochem. J.* **142,** 127–137
4. Crompton, M., Palmieri, F., Capano, M., and Quagliariello, E. (1974) *FEBS Lett.* **46,** 247–250
5. Meijer, A. J., Groot, G. S. P., and Tager, J. M. (1970) *FEBS Lett.* **8,** 41–44
6. Robinson, B. H., Williams, G. R., Halperin, M. L., and Leznoff, C. C. (1972) *J. Membr. Biol.* **7,** 391–401
7. Johnson, R. N., and Chappell, J. B. (1973) *Biochem. J.* **134,** 769–774
8. Palmieri, F., Passarella, S., Stipani, I., and Quagliariello, E. (1974) *Biochim. Biophys. Acta* **333,** 195–208
9. Kaplan, R. S., and Pedersen, P. L. (1985) *J. Biol. Chem.* **260,** 10293–10298
10. Bisaccia, F., Indiveri, C., and Palmieri, F. (1988) *Biochim. Biophys. Acta* **933,** 229–240
11. Indiveri, C., Capobianco, L., Krämer, R., and Palmieri, F. (1989) *Biochim. Biophys. Acta* **977,** 187–193
12. Indiveri, C., Dierks, T., Krämer, R., and Palmieri, F. (1989) *Biochim. Biophys. Acta* **977,** 194–199
13. Indiveri, C., Prezioso, G., Dierks, T., Krämer, R., and Palmieri, F. (1993) *Biochim. Biophys. Acta* **1143,** 310–318
14. Lancar-Benba, J., Foucher, B., and Saint-Macary, M. (1994) *Biochim. Biophys. Acta* **1190,** 213–216
15. Palmieri, L., Palmieri, F., Runswick, M. J., and Walker, J. E. (1996) *FEBS Lett.* **399,** 299–302
16. Kakhniashvili, D., Mayor, J. A., Gremse, D. A., Xu, Y., and Kaplan, R. S. (1997) *J. Biol. Chem.* **272,** 4516–4521
17. Walker, J. E., and Runswick, M. J. (1993) *J. Bioenerg. Biomembr.* **25,** 435–446
18. Kuan, J., and Saier, M. H. (1993) *Crit. Rev. Biochem. Mol. Biol.* **28,** 209–233
19. Palmieri, F. (1994) *FEBS Lett.* **346,** 48–54
20. Palmieri, F., and van Ommen, B. (1998) in *Frontiers of Cellular Bioenergetics: Molecular Biology, Biochemistry, and Physiopathology* (Papa, S., Guerieri, F., and Quagliariello, E. eds), Plenum Press, New York, in press
21. Dolce, V., Iacobazzi, V., Palmieri, F., and Walker, J. E. (1994) *J. Biol. Chem.* **269,** 10451–10460
22. Don, R. H., Cox, P. T., Wainwright, B. J., Baker, K., and Mattick, J. S. (1991) *Nucleic Acids Res.* **19,** 4008
23. Runswick, M. J., Philippides, A., Lauria, G., and Walker, J. E. (1994) *DNA Sequence* **4,** 281–291
24. Miroux, B., and Walker, J. E. (1996) *J. Mol. Biol.* **260,** 289–298
25. Fiermonte, G., Walker, J. E., and Palmieri, F. (1993) *Biochem. J.* **294,** 293–299
26. Laemmli, U. K. (1970) *Nature* **227,** 680–685
27. Capobianco, L., Bisaccia, F., Mazzeo, M., and Palmieri, F. (1996) *Biochemistry* **35,** 8974–8980
28. Palmieri, F., Indiveri, C., Bisaccia, F., and Iacobazzi, V. (1995) *Methods Enzymol.* **260,** 349–369
29. Chirgwin, J. M., Przbyla, A. F., MacDonald, A. J ., and Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299
30. Dolce, V., Fiermonte, G., and Palmieri, F. (1996) *FEBS Lett.* **399,** 95–98
31. Ness, G. C., and Pendleton, L. C. (1991) *FEBS Lett.* **288,** 21–22
32. Proudfoot, N. J., and Brownlee, G. G. (1976) *Nature* **263,** 211–214
33. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517
34. Runswick, M. J., Walker, J. E., Bisaccia, F., Iacobazzi, V., and Palmieri, F. (1990) *Biochemistry* **29,** 11033–11040
35. Palmieri, F., Quagliariello, E., and Klingenberg, M. (1972) *Eur. J. Biochem.* **29,** 408–416
36. Passarella, S., Palmieri, F., and Quagliariello, E. (1973) *FEBS Lett.* **38,** 91–95
37. Lancar-Benba, J., Foucher, B., and Saint-Macary, M. (1996) *Biochimie* (*Paris*) **78,** 195–200
38. Saraste, M., and Walker, J. E. (1982) *FEBS Lett.* **144,** 250–254
39. Walker, J. E., Cozens, A. L., Dyer, M. R., Fearnley, I. M., Powell, S. J., and Runswick, M. J. (1987) *Chem. Scripta* **27B,** 97–105
40. Powell, S. J., Medd, S. M., Runswick, M. J., and Walker, J. E. (1989) *Biochemistry* **28,** 866–873
41. Cozens, A. L., Runswick, M. J., and Walker, J. E. (1989) *J. Mol. Biol.* **206,** 261–280
42. Lawson, J. E., and Douglas, M. G. (1988) *J. Biol. Chem.* **263,** 14812–14818
43. Kolarov, J., Kolarova, N., and Nelson, N. (1990) *J. Biol. Chem.* **265,** 12711–12716
44. Bathgate, B., Baker, A., and Leaver, C. J. (1989) *Eur. J. Biochem.* **183,** 303–310
45. Iacobazzi, V., Poli, A., Blanco, A., and Palmieri, F. (1996) *Plant Physiol.* **110,** 1435–1436
46. Fleury, C., Neverova, M., Collins, S., Raimbault, S., Champigny, O., Levi-Meyrueis, C., Bouillaud, F., Seldin, M. F., Surwit, R. S., Ricquier, D., and Warden, C. H. (1997) *Nat. Genet.* **15,** 269–272
47. Boss, O., Samec, S., Paoloni-Giacobino, A., Rossier, C., Dulloo, A., Seydoux, J., Muzzin, P., and Giacobino, J.-P. (1997) *FEBS Lett.* **408,** 39–42
48. Solanes, G., Vidal-Puig, A., Grujic, D., Flier, J. S., and Lowell, B. B. (1997) *J. Biol. Chem.* **272,** 25433–25436