BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **219,** 249–255 (1996)
ARTICLE NO. 0213

Molecular Cloning of a Human cDNA Encoding a Trifunctional Enzyme of Carbamoyl-Phosphate Synthetase-Aspartate Transcarbamoylase-Dihydroorotase in *de Novo* Pyrimidine Synthesis

Hiroyuki Iwahana, Miwa Fujimura, Setsuko Ii, Maki Kondo, Maki Moritani, Yoko Takahashi, Takashi Yamaoka, Katsuhiko Yoshimoto, and Mitsuo Itakura

Otsuka Department of Clinical and Molecular Nutrition, The University of Tokushima, Tokushima, 770 Japan

Received December 30, 1995

A human CAD cDNA encoding a trifunctional enzyme of carbamoylphosphate synthetase-aspartate transcarbamoylase-dihydroorotase, which catalyzes the first three steps of *de novo* pyrimidine nucleotide biosynthesis, was cloned from a human fibroblast cell line of TIG-1-20 by polymerase chain reaction (PCR). The predicted open reading frame encodes a protein of 2,225 amino acids with a deduced molecular weight (Mr) of 242,913. The deduced amino acid sequence exhibits 95.3 and 76.1% identity with the CAD sequences of hamster and *Squalus acanthias.* The DNA fragment of 6,679 bp containing the full-length coding sequence was amplified by nested PCR using the first-strand cDNA of human cell lines of TIG-1-20 and COLO205 as a template. Southern blot analysis suggested that the CAD gene exists as a single copy in the human genome.

Pyrimidines as well as purines are essential components of DNA, RNA etc. and indispensable for cell proliferation. Elevated activities of *de novo* pyrimidine biosynthetic enzymes have been reported in rapidly growing cells such as ascites tumor cells, hepatoma nodules, and regenerating liver (1,2). Pyrimidine and purine derivatives, such as fluorouracil, arabinosyl cytosine, and mercaptopurine are clinically used as anticancer drugs. Inhibition of amidophosphoribosyltransferase (ATase), a presumed rate-limiting step of *de novo* purine biosynthesis, with an antisense oligodeoxynucleotide complementary to the mRNA of ATase, which we cloned (3,4), inhibited cell proliferation. Inhibition of the rate-limiting step of *de novo* pyrimidine biosynthesis is thus another target to develop anticancer drugs.

In mammalian species, the first three steps of *de novo* pyrimidine biosynthetic pathway are catalyzed by a trifunctional polypeptide called CAD, which has enzymatic activities of carbamoylphosphate synthetase II (CPSase II; EC 6.3.3.5), aspartate transcarbamoylase (ATCase; EC 2.1.3.2), and dihydroorotase (DHOase; EC 3.5.2.3) (5). CPSase II is considered a rate-limiting enzyme of this pathway and the three separate functional domains are localized in the order of NH_2-CPSase II-DHOase-ATCase-COOH (5). The gene expression of CAD and ATase increased in G_1/S phase of the cell cycle (6,7). An E-box influenced the CAD gene expression (8) but E2F1 did not activate the promoter of the CAD gene (9). There were E-boxes and E2F sites in the promoter region of the ATase gene (7). The CPSase activity of the *in vitro* phosphorylated CAD at two sites by cAMP-dependent protein kinase (PKA) was resistant to feedback control by UTP (10). ATase was also shown to be phosphorylated *in vitro* by PKA (11). It is important to understand the concordant regulation of the gene expression and activities of CAD and ATase during cell proliferation.

The full-length sequence of human CAD cDNA was not known, although those of hamster (12,13), *Squalus acanthias* (spiny dogfish) (14), *Dorsophila melanogaster* (15), and *Dictyostelium discoideum* (16) have been determined. Davidson *et al.* (17) have reported the nucleotide sequence of a human genomic DNA fragment including the 3′ end of DHOase and the complete coding region of ATCase. An EST tag (293 bp) related to the human CAD gene was registered to the GenBank (T29023).

0006-291X/96 $18.00

In this paper, we report the human CAD cDNA sequence encoding all three enzymes of CPSase II, DHOase, and ATCase with the PCR strategy.

MATERIALS AND METHODS

Cell Culture

The human fibroblast cell line TIG-1-20 (18) and the human colon adenocarcinoma cell line COLO205 (19) were provided by the Japanese Cancer Research Resources Bank (Tokyo, Japan) and ATCC (Rockville, MD), respectively. Cells of each cell line were grown in 90-mm dishes in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) in an atmosphere of 5% CO_2-95% air at 37°C.

RNA Isolation and First-Strand cDNA Synthesis

Total RNA was prepared from each cultured cell line of TIG-1-20 and COLO205 using ISOGEN (Nippon Gene, Tokyo, Japan). The first-strand cDNA was synthesized from 10 μg total RNA using a First-strand cDNA Synthesis Kit (Life Sciences, Petersburg, FL) in a total volume of 25 μl. The synthesized first-strand cDNA was stored at −20°C until use.

PCR Amplification

Primers used for this experiment are listed in Table 1. They were synthesized by an Applied Biosystems 392 DNA/RNA synthesizer (Foster City, CA). Nine degenerate primers for the conserved regions were designed according to the amino acid alignment of hamster CAD, rat mitochondrial CPSase I, *D. melanogaster* CAD, *D. discoideum pyr*1-3 gene product, *S.*

TABLE 1
Primers Used to Clone Human CAD cDNA

Name	Sequence	Position	AA sequence
oRB1377	5′-AGTGT(AGCT)CA(AG)TT(CT)CA(CT)CC(AGCT)GA-3′	1020 → 1039	SVQFHPE
oRB1378	5′-AGA(AGCT)CCTGA(GC)TA(AG)TC(AG)AA(CT)TC-3′	1277 → 1257	EFDYSGS
oRB1386	5′-AT(ACT)TG(CT)GC(AGCT)TTGA(CT)TG(CT)GG-3′	558 → 577	ICALDCG
oRB1408	5′-AGCTGTGGCCTCTTTCACAG-3′	1112 → 1093	
oRB1434	5′-CTGAA(AG)GT(AGCT)AT(ACT)GA(AG)TG(CT)AA-3′	3654 → 3673	LKVIECN
oRB1435	5′AC(CT)TC(AGCT)CC(AGCT)GTACT(AGCT)GTCAT(CT)TC-3′	3889 → 3867	EMTSTGEV
oRB1442	5′-CTGAAGGGCCAGCTTGGTGC-3′	1059 → 1040	
orb1464	5′-ATCGTGGCAGAGACATCTGATCTG-3′	614 → 591	
oRB1465	5′-CTATTGGCCTTAGCCATTGG-3′	795 → 814	
oRB1477	5′-GCTCTGGT(AGCT)TTGGA(AG)GA(CT)GG-3′	33 → 52	ALVLEDG
oRB1479	5′-AAAAGGTGTCGACTGGGCCTTGGAG-3′	5348 → 5324	
oRB1484	5′-GG(AGCT)ATGGT(AGCT)GG(AGCT)TA(CT)CC(AGCT)GA-3′	120 → 139	GMVGYPE
Orb1485	5′TACTGGATCCACACCCAGTGTCTTG-3′	37131 → 3707	
oRB1486	5′-TTCTGGAGTCGACGGAGTAAAGGTG-3′	3791 → 3815	
oRB1487	5′-GCTTTGGTCAGTTCATCCTC-3′	6508 → 6527	
oRB1488	5′-CTAAAGAGTCGACGCTGAGGAAGCC-3′	6735 → 6711	
oRB1502	5′-GACATCTAGAGAGCGGCTGACTGAG-3′	1130 → 1154	
oRB1503	5′-GTCTTGTCGACGAAGGGGAAGGAGC-3′	3712 → 3688	
oRB1507	5′-GCCTGAATTCTCCTTCTCCCGCTTG-3′	3815 → 3839	
oRB1511	5′-CCGGTTTGAAACACCACTTC-3′	121 → 102	
oRB1516	5′-TCAGAGGATAGTTGAGCACT-3′	195 → 176	
oRB1517	5′-CAT(AGCT)AC(CT)TC(AGCT)CC(AGCT)TCACT(CT)TTCAT-3′	2324 → 2301	MKSVGEVM
oRB1518	5′-TCCAGTGT(AGCT)GA(AG)TT(CT)GA(CT)TGGTG-3′	2865 → 2887	SSVEFDWC
oRB1536	5′-TCTGGGTTATAGTTCACCATGATGG-3′	2956 → 2932	
oRB1543	5′-CCCGCCTCTGAGCTCCCTTC-3′	5 → 24	
oRB1745	5′-CCCCGCCTCTGAGCTCCCTTCCCAT-3′	4 → 28	
oRB1747	5′-CATGGCGGCCCTAGTGTTGGAGGAC-3′	26 → 50	
oRB1751	5′CTAGAAACGGCCCAGCACGGTGGCTAA-3′	6704 → 6678	
oRB205	5′-GACTCGAGTCGACATCGATTTTTTTTTTTTTTTT-3′		
oRB254	5′-ATCGACTCGAGTCGACATCGATTT-3′	Primers for RACE	
oRB1348	5′-CCGACTCGAGTCGACATCG-3′		

cerevisiae pyrimidine-specific CPSase, *S. cerevisiae* arginine-specific CPSase, *E. coli car*A gene product, *E. coli car*B gene product, and *S. typhimurium car*A gene product, and the amino acid sequence of the amino terminus of hamster CAD (12,13). Primers of oRB1479, oRB1487, oRB1488, and oRB1751 were designed according to the sequence reported by Davidson *et al.* (17).

The PCR amplification was performed using 1–2 μl of the synthesized first-strand cDNA as a template in a total volume of 20–50 μl which consisted of 1 μM each of a pair of primers, 200 μM each of four deoxynucleotides, 0.025 units/μl of AmpliTaq (Perkin Elmer, Norwalk, CT) or TaKaRa Ex Taq (Takara, Kyoto, Japan) in a buffer recommended by each manufacturer. Thirty cycles of the reaction at 94, 40–60, and 72°C for 0.5-1, 0.5-1, and 2–3 min, respectively, were carried out in the Program Temp Control System PC700 (Astec, Fukuoka, Japan). The first denaturation at 94°C was extended to 2.5–3 min, and the last polymerization at 72°C was extended to 10 min to complete the elongation reaction.

The 5′- and 3′-end of CAD cDNA were amplified by the method of RACE (20). The first-strand cDNA corresponding to the 5′-end was tailed with dATP using terminal deoxynucleotidyl transferase (Toyobo, Osaka, Japan), and the nested PCR was performed.

A DNA fragment of 6,679 bp containing the full-length coding region was amplified from the first-strand cDNA of TIG-1-20 and COLO205 by nested PCR. The first PCR was performed using a pair of primers of oRB1745 and oRB1488, and the second PCR was performed using a pair of primers of oRB1747 and oRB1751. TaKaRa Ex Taq was used for the polymerization reaction. The elongation reaction at 72°C was extended to 10 min.

Nucleotide Sequencing

PCR-amplified DNA fragments were gel-purified and subcloned into pCR&otrade;II (Invitrogen, Leek, Netherlands) or pBluescript II SK(+) (Stratagene, La Jolla, CA). DNA sequencing was carried out using *Taq* Dye Primer Cycle Sequencing Core Kit (401112) in the Applied Biosystems 373A or 377 DNA sequencer. Cloned pCR™II or pBluescript was used as a template.

Southern Blot Analysis

Aliquots of DNA isolated from human peripheral white blood cells in 5 μg were digested with indicated restriction enzymes. Digested DNA samples were subjected to 1.0% agarose gel electrophoresis and transferred to a nylon membrane (Hybond-N, Amersham, Buckinghamshire, UK) by the method of Southern (21). A 236 bp DNA fragment of hCAD1434-1435 was amplified with a pair of primers of oRB1434 and oRB1435 using a plasmid with a known sequence as a template and gel-purified. This fragment was radiolabeled with [α^{-32}P]dCTP (3000 Ci/mmol; Amersham) using a Multiprime DNA labelling system (Amersham) and used as a probe.

RESULTS

Nucleotide and Amino Acid Sequences

The 6,989 bp nucleotide sequence of the cDNA encoding human CAD and the deduced amino acid sequence are shown in Fig. 1. The nucleotide sequence was confirmed by analyzing both strands of at least 3 clones. A part of sequence of 1,480 bp from 5,280 to 6,759 was identical with sequences of exons determined by Davidson *et al.* (17). There is a single open reading frame starting at the first AUG, encompassing positions 27–29, and ending with a UAG at nucleotides 6,702–6,704. The polyadenylation signal AAUAAA is located at 6,949–6,954, which is 19 residues upstream of the poly(A) addition site. The deduced amino acid sequence encoded by this open reading frame is a peptide chain of 2,225 amino acids with the calculated Mr of 242,913.

Amplification of the Full-Length Coding Region

As shown in Fig. 2, a 6,679 bp DNA fragment was amplified by nested PCR with primers of oRB1745 and oRB1488 as an outer primer set and oRB1747 and oRB1751 as an inner primer set using the first-strand cDNA synthesized from total RNA of COLO205 as a template. A 6,679 bp DNA fragment was also amplified from the first-strand cDNA of TIG-1-20 (data not shown). Computer analysis showed that this 6,679 bp DNA fragment could be digested with restriction enzymes of *Fba*I, *Hind*III, *Kpn*I, and *Pst*I. All expected DNA fragments created by the digestion with these four enzymes were detected (Fig. 2).

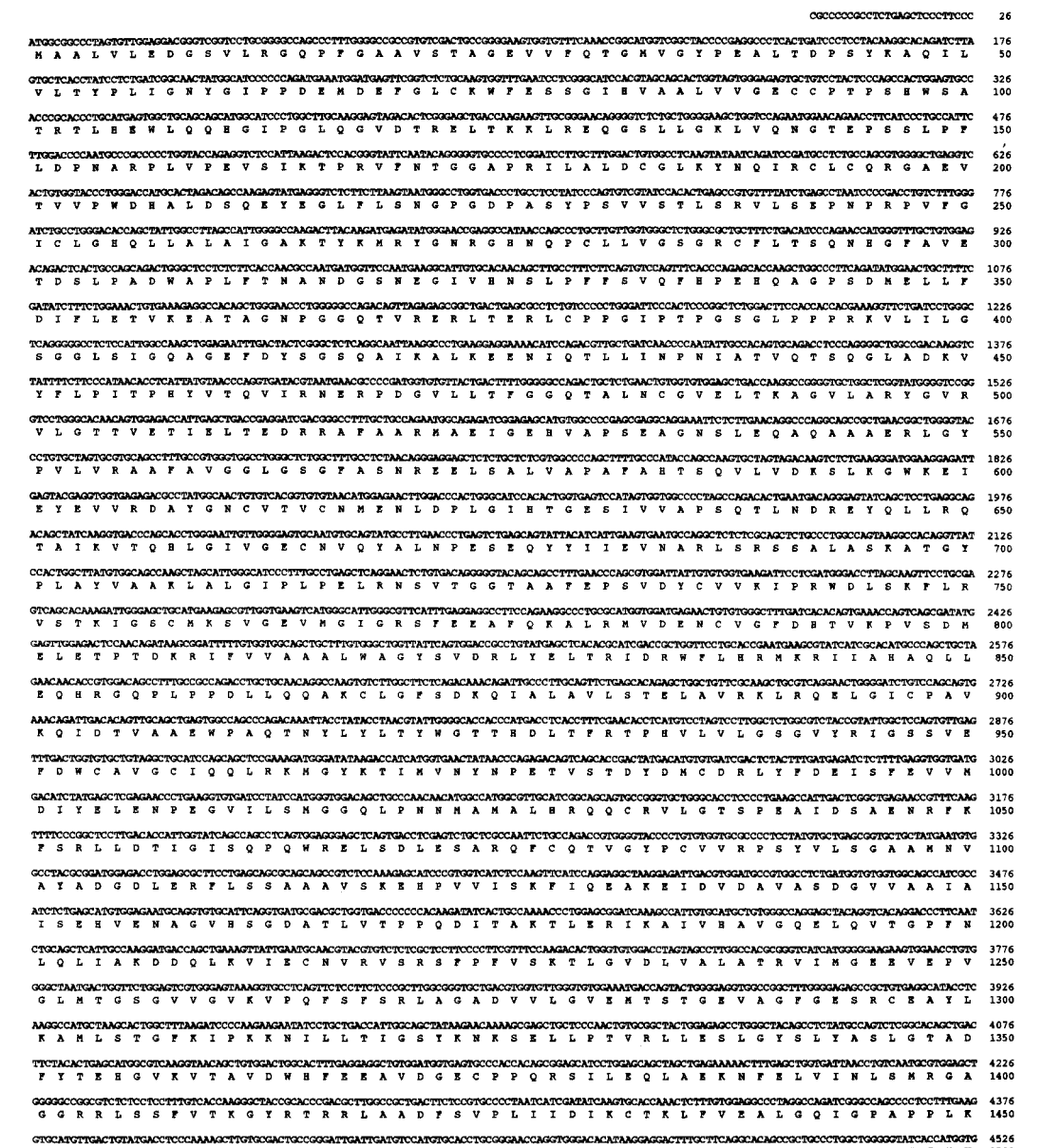

FIG. 1. Nucleotide sequence of the human CAD cDNA and deduced amino acid sequence. The nucleotide sequence is arbitrarily numbered from the nucleotide at the 5′ end. The amino acid sequence shown by one letter is numbered from the amino terminus. DDBJ, EMBL, and GenBank accession number is D78586.

Southern Blot Analysis

Genomic DNA isolated from human peripheral white blood cells was digested, respectively, by one of 6 restriction enzymes with 6 base recognition sequences: *Eco*RI, *Hind*III, *Kpn*I, *Pst*I, *Sac*I, and *Xba*I. The genomic Southern blot prepared from these digests was probed with hCAD1434–

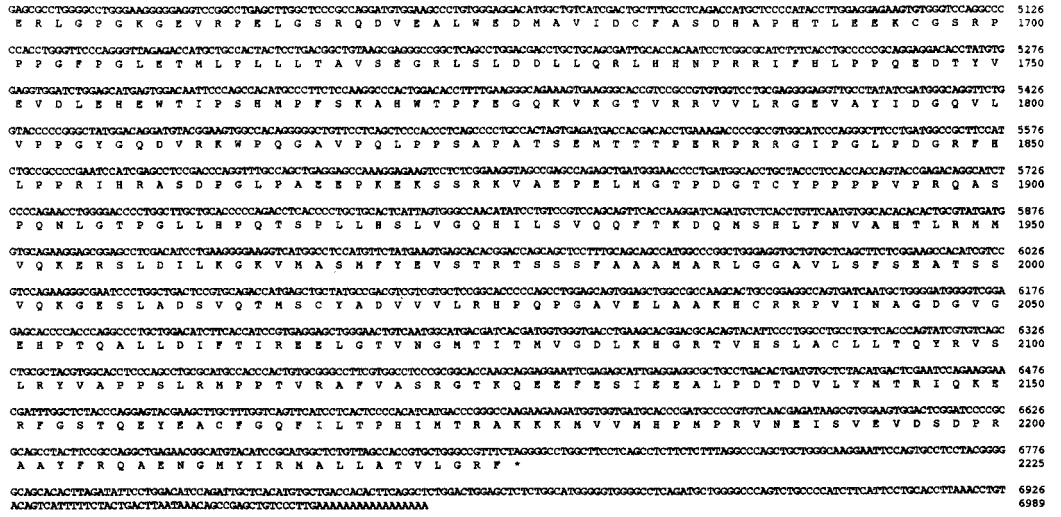

FIG. 1 —*Continued.*

1435. As shown in Fig. 3, a single band was observed in all lanes, suggesting that the CAD gene exists as a single copy in the human genome.

DISCUSSION

The determined sequence contains a single open reading frame of 6,679 bp encoding a protein of 2,225 amino acids with a deduced Mr of 242,913. The entire deduced amino acid sequence of

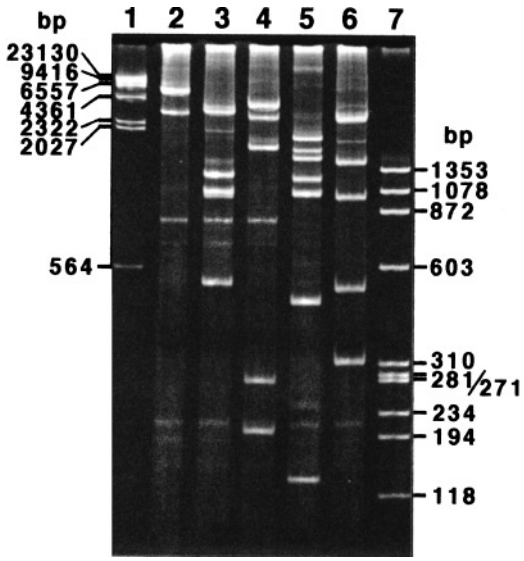

FIG. 2. PCR-amplification of the entire coding region of human CAD. A 6,679 bp DNA fragment was amplified by nested PCR with the outer primer pair of oRB1745 and oRB1488 and the inner primer pair of oRB1747 and oRB1751 using the first-strand cDNA synthesized from COLO205 total RNA as a template. Lane 1 is *Hind*III-digested λDNA. Lane 2 is the nested PCR-amplified DNA fragment. Lanes 3–6 are DNA fragments, which were digested with *Fba*I, *Hind*III, *Kpn*I, and *Pst*I, respectively. Lane 7 is *Hae*III-digested $\phi \times 174$. *Fba*I-digestion produced DNA fragments of 3,755, 1,296, 1,075, and 553 bp. *Hind*III-digestion produced DNA fragments of 4,379, 1,810, 288, and 202 bp. *Kpn*I-digestion produced DNA fragments of 2,155, 1,599, 1,274, 1,040, 479, and 132 bp. *Pst*I-digestion produced DNA fragments of 3,279, 1,489, 1,051, 533, and 327 bp.

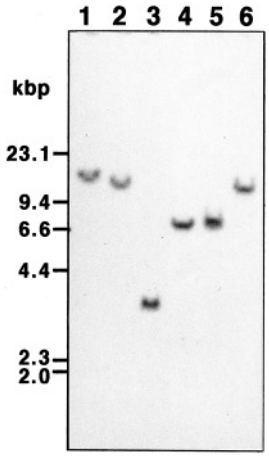

FIG. 3. Southern blot analysis. Human genomic DNA (5 μg) was digested, respectively, with *Eco*RI, *Hind*III, *Kpn*I, *Pst*I, *Sac*I, and *Xba*I and subjected to electrophoresis on a 1% agarose gel. After Southern blotting, the nylon membrane was hybridized with a probe hCAD1434–1435, which contained a 236 bp DNA fragment from base pair 3,654 to 3,889 of the human CAD cDNA.

human CAD exhibited 95.3 and 76.1% identity with those of hamster and *Squalus acanthias*, respectively (12–14).

We searched protein motifs in the deduced amino acids sequence using a computer program of GENETYX-MAC Ver 7.3 (SDC, Tokyo, Japan) and a database of PROSITE. There were a glutamine amidotransferases class-I active site of PVFGICLGHQLL from 247 to 258, two sites of carbamoyl-phosphate synthase subdomain signature 1 of YPVLVRAAFAVGGLG from 550 to 564 and YPCVVRPSYVLSGAA from 1,083 to 1,097, two sites of carbamoyl-phosphate synthase signature 2 of IIEVNARL from 680 to 687 and VIECNVRV from 1,212 to 1,219, a dihydroorotase signature 1 of DVHVHLREP from 1,469 to 1,477,a dihydroorotase signature 2 of ASDHAPH-TLEEK from 1,684 to 1,695, and an aspartate and ornithine carbamoyltransferases signature of FYEVSTRT from 1,969 to 1,976 (Fig. 1). This strongly suggested that we cloned the human CAD cDNA.

From the similarity to other CPSases (22), the CPSase II domain extends from M^{1} to C^{1455}. It can be further divided into a glutaminase domain (M^{1}-P^{365}), the amino half of the CPSase domain (K^{395}-T^{932}), and the carboxyl half of the CPSase domain (P^{933}-C^{1455}). In the glutaminase domain, C^{252} is conserved as the residue required for formation of the γ-glutamyl thioester intermediate (23). Two PKA phosphorylation sites have been identified in hamster CAD (24). One of which located at the C-terminal end of the CPSase synthetase domain is R^{1403}-R^{1404}-L^{1405}-S^{1406} and conserved as well as *Squalus acanthias* CAD (14).

The DHOase domain extends from M^{1456} to R^{1789}. Five histidine residues of H^{1471}, H^{1473}, H^{1614}, H^{1642}, and H^{1690} have been considered to be candidates for the zinc-binding residues in the hamster CAD (25). These histidine residues are conserved in the human CAD DHOase domain.

The ATCase domain begins at L^{1918} and ends at F^{2225}. Between the DHOase and ATCase domains, there is a large linker region, from G^{1790} to L^{1917}. The second PKA phosphorylation site in hamster CAD has been identified as a HRAS sequence present within this interdomain linker region (26). This sequence was conserved in human CAD as H^{1856}-R^{1857}-A^{1858}-S^{1859}.

Although DNA fragments of 3,400 and 780 bp were amplified, a major 6,679 bp DNA fragment could be amplified by the nested PCR. The restriction maps of *Fba*I, *Hind*III, *Kpn*I, and *Pst*I were compatible with those made by computer analysis. Therefore, the human CAD mRNA serially encodes CPSase II, DHOase, and ATCase. The 6,679 bp fragment migrated faster than the 6,557 bp band of DNA size marker of *Hind*III-digested λDNA on an 8% polyacrylamide gel (Fig. 2, lanes

1 and 2). This was probably due to the difference of salt concentrations of solutions dissolving DNA fragments. The human CAD gene presumably exists as a single copy in the genome, because all samples in the Southern blot analysis using 6 restriction enzymes with 6 base recognition sites, showed a single band.

We cloned and sequenced the human CAD cDNA, which contained all three functional domains of CPSase II, DHOase, and ATCase. This will help us to develop anticancer drugs targeting at pyrimidine biosynthesis.

ACKNOWLEDGMENT

This study was supported in part by a grant from Otsuka Pharmaceutical Factory Inc. for Otsuka Department of Clinical and Molecular Nutrition, School of Medicine, The University of Tokushima.

REFERENCES

1. Calva, E., and Cohen, P. P. (1959) *Cancer Res.* **19,** 679–683.
2. Calva, E., Lowenstein, J. M., and Cohen, P. P. (1959) *Cancer Res.* **19,** 101–104.
3. Iwahana, H., Yamaoka, T., Mizutani, M., Mizusawa, N., Ii, S., Yoshimoto, K., and Itakura, M. (1993) *J. Biol. Chem.* **268,** 7225–7237.
4. Iwahana, H., Oka, J., Mizusawa, N., Kudo, E., Ii, S., Yoshimoto, K., Holmes, E. W., and Itakura, M. (1993) *Biochem. Biophys. Res. Commun.* **190,** 192–200.
5. Webster, D. R., Becroft, D. M. O, and Suttle, D. P. (1995) *in* The Metabolic and Molecular Basis of Inherited Disease (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., Eds.), Vol. II, 7th ed., pp. 1799–1837, MacGraw-Hill, New York.
6. Rao, G. N., and Church, R. L. (1988) *Exp. Cell Res.* **178,** 449–456.
7. Iwahana, H., Honda, S., Tsujisawa, T., Takahashi, Y., Adzuma, K., Katashima, R., Yamaoka, T., Moritani, M., Yoshimoto, K., and Itakura, M. (1995) *Biochim. Biophys. Acta* **1261,** 369–380.
8. Miltenberger, R. J., Sukow, K. A., and Farnham, P. J. (1995) *Mol. Cell. Biol.* **15,** 2527–2535.
9. Li, Y., Slansky, J. E., Myers, D. J., Drinkwater, N. R., Kaelin, W. G., and Farnham, P. J. (1994) *Mol. Cell. Biol.* **14,** 1861–1869.
10. Carrey, E. A., Campbell, D. G., and Hardie, D. (1985) *EMBO J.* **4,** 3735–3742.
11. Itakura, M., Yamaoka, T., Yoshikawa, H., and Yamashita, K. (1988) *Biol. Res.* **9,** Suppl. 3, 111–116.
12. Simmer, J. P., Kelly, R. E., Rinker, A. G., Jr., Scully, J. L., and Evans, D. R. (1990) *J. Biol. Chem.* **265,** 10395–10402.
13. Bein, K., Simmer, J. P., and Evans, D. R. (1991) *J. Biol. Chem.* **266,** 3791–3799.
14. Hong, J., Salo, W. L., and Anderson, P. M. (1995) *J. Biol. Chem.* **270,** 14130–14139.
15. Freund, J. N., and Jarry, B. P. (1987) *J. Mol. Biol.* **193,** 1–13.
16. Faure, M., Camonis, J. H., and Jacquet, M. (1989) *Eur. J. Biochem.* **179,** 345–358.
17. Davidson, J. N., Rao, G. N., Niswander, L., Andreano, C., Tamer, C., and Chen, K. C. (1990) *DNA Cell Biol.* **9,** 667–676.
18. Ohashi, M., Aizawa, S., Ooka, H., Ohsawa, T., Kaji, K., Kondo, H., Kobayashi, T., Noumura, T., Matsuo, M., Mitsui, Y., Murota, S., Yamamoto, K., Ito, H., Shimada, H., and Utakoji, T. (1980) *Exp. Gerontol.* **15,** 121–133.
19. Semple, T. U., Quinn, L. A., Woods, L. K., and Moore, G. E. (1978) *Cancer Res.* **38,** 1345–1355.
20. Frohman, M. A. (1994) *PCR Methods and Applic.* **4,** S40–S58.
21. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
22. Evans, D. R., Bein, K., Guy, H. I., Liu, X., Molina, J. A., and Zimmermann, B. H. (1993) *Biochem. Soc. Trans.* **21,** 186–191.
23. Zalkin, H. (1993) *Adv. Enzymol. Relat. Areas Mol. Biol.* **66,** 203–309.
24. Carrey, E. A., and Hardie, D. G. (1988) *Eur. J. Biochem.* **171,** 583–588.
25. Quinn, C. L., Stephenson, B. T., and Switzer, R. L. (1991) *J. Biol. Chem.* **266,** 9113–9127.
26. Carrey, E. A. (1992) *Biochem. J.* **287,** 791–795.