Vol. 167, No. 1, 1990
February 28, 1990
BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
Pages 130-136

RAT PANCREATIC COLIPASE mRNA: NUCLEOTIDE SEQUENCE OF A cDNA CLONE AND NUTRITIONAL REGULATION BY A LIPIDIC DIET

Catherine Wicker * and Antoine Puigserver

Centre de Biochimie et de Biologie Moléculaire du Centre National de la Recherche Scientifique, 31 Chemin Joseph-Aiguier 13402 Marseille Cedex 9, France

Received December 6, 1989

A cDNA clone encoding rat pancreatic colipase was isolated using as a probe a synthetic deoxyoligonucleotide corresponding to a highly conserved amino acid sequence region in colipases from other species.The cloned messenger codes for a protein of 95 amino acids plus a signal peptide of 17 amino acids. The structure of the full-length cDNA was also determined and the corresponding amino acid sequence showed a high degree of homology with those of other known colipases .Quantification of the homologous mRNA in the pancreas of animals fed a high-lipid diet was consistent with a specific though moderate induction of colipase messenger by the nutritional manipulation.

Colipase is a small protein, which is synthesized and secreted by the exocrine pancreas (1).Various forms of colipase have been obtained, depending on the methods of purification used. Despite the lack of a number of amino acids in the N-terminal propeptide part and/or in the C-terminal region of the polypeptide chain as a result of limited proteolysis, all these proteins did share the same biological activity.The colipase which is secreted as procolipase in the pancreatic juice is activated by trypsin in the upper intestine (2). It is known to act as a cofactor for pancreatic lipase, by anchoring the latter protein to the substrate interface in the presence of bile salts and, consequently, enabling hydrolysis of dietary triglycerides to proceed (1).Thus,*in vivo*digestion of dietary fat depends strongly on this molecule. Conflicting results have been obtained on the possible effect of dietary lipids in the modulation of colipase activity (3,4). It seems that the nutritional adaptative process occurred in the rat only when daily protein intake was between 3.5 and 6.0 g (5).

We report in this paper the isolation and nucleotide sequence determination of a rat pancreatic colipase cDNA clone and its deduced amino acid sequence. Several important residues for the binding of colipase to lipase and bile salts were found to be present.We also report the induction of colipase mRNA in the rat pancreas as a result of the ingestion of a high-lipid diet.

MATERIALS AND METHODS

Materials: (α^{32}P)dCTP(110 TBq/mmol), (γ^{32}P)dATP (110 TBq/mmol) and (^{35}Sα)dATP (37 TBq/mmol) were purchased from Amersham Corp. (Les Ulis, France). T4 polynucleotide

*To whom correspondence should be addressed.

0006-291X/90 $1.50

kinase was from Boehringer (Mannheim, FRG). The other reagents have already been described elsewhere (6).

Synthetic oligonucleotides: All following deoxyoligonucleotides were chemically synthesized by means of an Applied Biosystems oligonucleotide synthesizer:

A. 5' TC(AG)CA(ACGT)GG(AG)CA(CT)TT(AG)TA(AG)TA 3'
B. 5' TGCTGCCAACATGACAC 3'
C. 5' TGTCCCTGTGAGCGGGGCC 3'

Construction of the 20-mer oligonucleotide set A was based on the following amino acid sequence: 58Tyr-Tyr-Lys-Cys-Pro-Cys-Glu64 which is highly conserved in porcine (7) equine (8-9) and human (10) colipases, and on codon usage tables for mammalian genes. The nucleotide sequence was complementary to the derived coding sequence. The 17- and 19-mer oligonucleotides B and C were used as primers in cDNA sequencing.

cDNA library screening and nucleotide sequencing: The Wistar rat pancreatic cDNA library was constructed in pUC 9 (6) and screened with the chemically synthesized 20-mer oligonucleotide set A which was previously radiolabeled by treatment with (γ^{32}P)ATP and T4 polynucleotide kinase (11). Nucleotide sequencing of rat colipase cDNA was carried out by the dideoxy chain termination technique of Sanger *et al.* (12) using the pUC sequencing kit from Boehringer.

Cytoplasmic content of colipase mRNA in the pancreas of lipid-fed rats: Male Wistar rats, weighing 180-200 g were shifted from a control diet containing 3% sunflower oil to an isocaloric-isoproteic lipidic diet containing 25% sunflower oil.Daily food ingestion was about 20g (protein intake: 4.4 g/day). The composition of the diets were the same as those previously described (6), except for bicarbonate content, which was replaced by cellulose. Colipase activity was measured in pancreatic tissue homogenates by a titrimetric method (13). Cytoplasmic RNA was purified according to Chirgwin *et al.* (14) and total RNA was dot-blotted onto nitrocellulose membrane (15). Hybridization of immobilized RNA to the ^{32}P-radiolabeled cDNA probe and washing conditions were as previously detailed (6). Hybridized radioactivity was determined by Cerenkov counting in a Beckman LS 3800 automatic counting equipment.

RESULTS AND DISCUSSION

Library screening: Despite a minor change in the amino acid sequence extending between Tyr 58 and Glu 64 in the rat, as compared to the conserved one in other species (lysine 60 being replaced by an arginine, thus giving rise to a one base-pair mismatch), the ^{32}P-labeled 20-mer oligonucleotide set A was efficiently used for screening the cDNA library. One out of the few recombinant plasmids, a 525 base-pair fragment which was long enough to encode a nearly full-length copy of colipase mRNA, was selected for sequence analysis.

Nucleotide sequence: The nucleotide sequence of procolipase mRNA and the deduced amino acid sequence are shown in Fig. 1. The entire cDNA clone contains 525 nucleotides, including a poly(A) tail of 80 nucleotides.The initiator codon for methionine was found following a 14-base long 5'-untranslated region. The 5'-noncoding region did not show any significant complementarity with the 3' end of 18 S rRNA (5'UGCGGAAGGAU3') although the nucleotides flanking the initiator codon are not randomly distributated. Colipase mRNA has, as 80% of eukaryotic messengers, an adenine in position -3, and, as more than 50% of mRNAs, cytidines in positions -1 and -4 (16). The occurrence of purines in positions -3 (adenine) and

```
CAGCCGTACCAGTCACC ATG AAG GTC CTT GTT GTT CTG CTT
                  Met Lys Val Leu Val Val Leu Leu
                                -15                -10

GTA ACC CTC GTT GCC GTG GCC TAT GCA GTT CCT GGA
Val Thr Leu Val Ala Val Ala Tyr Ala Val Pro Gly
                -5                -1  1

CCC CGG GGT CTT TTT ATC AAC CTG GAG GAC GGT GAG
Pro Arg Gly Leu Phe Ile Asn Leu Glu Asp Gly Glu
    5                       10                  15

ATC TGC GTA AAC AGT ATG CAG TGT AAG AGC AGA TGC
Ile Cys Val Asn Ser Met Gln Cys Lys Ser Arg Cys
                20                  25

TGC CAA CAT GAC ACC ATC CTG GGC ATC GCC CGG TGC
Cys Gln His Asp Thr Ile Leu Gly Ile Ala Arg Cys
        30                  35

ACA CAC AAG GCC ATG GAG AAC AGC GAG TGC TCC CCA
Thr His Lys Ala Met Glu Asn Ser Glu Cys Ser Pro
40                  45                  50

AAG ACC CTC TAT GGG ATC TAC TAC AGG TGT CCC TGT
Lys Thr Leu Tyr Gly Ile Tyr Tyr Arg Cys Pro Cys
            55                  60

GAG CGG GGC CTG ACC TGT GAG GGG GAC AGG AGC ATC
Glu Arg Gly Leu Thr Cys Glu Gly Asp Arg Ser Ile
    65                  70                  75

ATT GGC GCC ATC ACC AAC ACC AAC TAC GGC GTC TGC
Ile Gly Ala Ile Thr Asn Thr Asn Tyr Gly Val Cys
                80                  85

CTC GAC TCC ACC CGC TCC AAG CAG TGA GATCG TGCAG
Leu Asp Ser Thr Arg Ser Lys Gln TGA STOP
        90                  95

TGAGC TGGGC CACCT CTCCC TTTCC CTTCA CTCGC CCCAT CTGAG

TCACC CATTG GC AATTAAA GCCCA TTGCA ACCTTG AAAA.....
```

Figure 1:
Rat pancreatic nucleotide sequence and the predicted aminoacid sequence. The subscript numbers refer to the amino acid positions. The sequence from -17 to -1 is the putative hydrophobic leader sequence. The polyadenylation signal (AATTAAA) is underlined.

+4 (adenine) was shown to enhance the binding efficiency to ribosome during formation of initiation complexes (16). The 5'-noncoding region is followed by an open reading frame encoding a presecretory protein of 112 amino acids. The 3' region following the TGA stop codon consists of a 89-nucleotide stretch and a poly(A) tail of 80 nucleotides. The polyadenylation recognition site (AATTAAA) is slightly altered, as that in amylase mRNA (17), and located 16 nucleotides upstream the poly(A) tract of colipase mRNA.

Amino acid sequence: The protein encoded by the open reading frame is built of 112 amino acids, and has calculated molecular weight of 12246 for the processed molecule. The amino acid sequence deduced from the cDNA begins with a 17-amino acid hydrophobic segment corresponding to the signal peptide. The prepeptide is comparable to those from other pancreatic secretory enzymes, as far as its length and hydrophobic character are concerned. Based on the homology existing between the activation peptides from pig (18), horse(9) and rat (this study) procolipases, as well as on the presence of a Val residue at the NH_2-terminus

of rat procolipase (19), the cleavage of the signal peptide should occur between Ala -1 and Val +1. The propeptide segment consists of 5 residues, four of which are strictly identical with those from pig and horse (see Figure 2). The presence of a glycine residue in position 3, instead of the highly conserved aspartic acid, should be mentioned. Tryptic hydrolysis of the Arg 5-Gly 6 bond gives rise to the mature protein. Both procolipase and colipase forms can bind either to lipase (20) or to a triacylglycerol interface covered by bile salts (23), and consequently participate together with lipase to the hydrolysis of emulsified triacylglycerol substrates (24). However,the lagtime for hydrolysis of intralipid is much shorter for the tryptic form of the protein (24).

Figure 2 shows a comparison of the amino acid sequence of human, pig, horse, chicken and rat colipases. The respective positions of the 10 cystein residues are conserved in the rat colipase. The protein has a phenylalanine residue (position 8), two methionines (positions 21 and 44), two histidines (positions 30 and 41) and four tyrosines (positions 55, 58, 59 and 84).Comparison of the six sequences indicates substantial stretches of identity between the proteins. Rat protein shares 63,60,61,69 and 51 % identical residues with pig, horse form A,horse form B, human and chicken proteins, respectively. If chemically similar amino acids are taken into account (25), then the homology increases to 80, 76, 77, 83 and 72%, respectively.

The deduced amino acid sequence of rat colipase contains all of the residues which contribute to the cofactor activity. On the one hand, the negatively charged residues in position12 (Glu) and 15 (Glu), known to be important for the anchoring of lipase (26) are conserved. On the other hand, the hydrophobic region Leu 54-Tyr 59 including the 3 tyrosine residues constituting the micelle binding site is identical to the corresponding human sequence. NMR experiments suggested the existence in porcine and equine proteins of a specific interaction between Tyr 58 or 59 and a histidine residue, in the presence of bile salts

```
                      10                  20                  30
Rat     VPGPRGLFINLEDGEICVNSMQCKSRCCQHDTILG
Pig     VPDPRGIIINLDEGELCLNSAQCKSNCCQHDTILS
Horse A VPDPRGVIINLEAGEICLNSAZCKSECCHQEESLS
Horse B VPDPRGVIINLEAGEICMNSAQCKSECCHRESSLS
Human       GIIINLENGELCMNSAQCKSNCCQHSSALG
Chicken     GLIFNLDTGELXLQSAQCKSECCQEDSGLS

            40                  50                  60                  70
Rat     IARCTHKAMENSECSPKTLYGIYYRCPCERGLTCE
Pig     LLRCALKARENSECSAFTLYGVYYKCPCERGLTCE
Horse A LARCAAKASENSECSAWTLYGVYYKCPCERGLTCQ
Horse B LARCAAKASENSECSAWTLYGVYYKCPCERGLTCQ
Human   LARCTSMASENSECSVKTLYGIYYKCPCERGLTCE
Chicken LAXC

                      80                  90                  100
Rat     GDRSIIGAITNTNYGVCLDSTRSKQ
Pig     GDKSLVGSITNTNFGICHNVGRSDS(EDKT)ARLS
Horse A VDKTLVGSIMNTNFGICFNAAKER
Horse B VDKTLVGSIMNTNFGICFDAARSEZR
Human   GDKTIVGSITNTNFGICHDAG
```

Figure 2:
Comparative analysis of the predicted aminoacid sequence of rat colipase to those of pig(residues 6-91 (7), 1-5 (18), and 96-103 (20), horse forms A (residues 1-55 (21), 1-95 (9) and B (8), man (10) and chicken (22). Equine colipase A residue 22 was E (9) or Q (21).

(27,28). His 88 has been proposed as the interacting residue in the pig protein whereas interaction in the horse protein was assumed to be assigned to histidine 29 which is the only histidine residue of the molecule (9, 27, 29).In the rat, residue 88 is a leucine and consequently one out of the two histidines in positions 30 and 41 probably participates in a similar way to the micelle binding site.An additional residue, tryptophan 52, has been shown to be involved in the binding site of colipase for micelle (9) in the horse protein, but has no counterpart in colipases from other species including the rat.The N-terminal part of the protein (residues 7-11) which is highly hydrophobic, was also assumed to be involved in the binding of bile salts. Differences appeared mainly in the C-terminal part of the rat protein, which is shorter than the porcine one.The fact that limited tryptic activation was shown to generate degraded forms of the porcine protein in its C-terminal part without any decrease in the activity (20), suggests that this region is probably not of prime importance for its activity.

Colipase mRNA response to a high-lipid diet: Once the probe was characterized, colipase mRNA response to a high-lipid diet was investigated. The results are shown in Table 1. Between 0 and 5 days of diet ingestion, an about 1.5-fold increase in the concentration of colipase mRNA was observed, a plateau being reached on the 3rd day of diet intake. Although the enhancement of colipase mRNA was rather small, it was nevertheless significant. This suggested that colipase synthesis is controlled at the level of transcription of the corresponding gene or turnover of the mRNA during the nutritional manipulation. However, the increase in the concentration of colipase mRNA did not parallel the change in colipase specific activity. Indeed, as already pointed out (30), colipase activity did not adapt to the lipid substrate during the 5 first days of diet intake under the same nutritional conditions. This was mainly due to a high protein level in the pancreas of control rats. However, from day 2 to day 5, specific activity increased slightly (about 1.2-fold). Thus, the discrepancy between colipase activity

Table 1:Effect of lipid ingestion on the specific activity of rat pancreatic colipase and the relative level of its mRNA

Days on diet	0	1	2	3	5	10
Specific activity (U/mg) (from Ref. 30)	234±16	177±42	176±22	197±42	220±46	ND
mRNA level	1	1.14±0.1	1.16±0.2	1.46±0.1	1.68±0.2	1.55±0.1.

Colipase specific activities are taken from Ref. 30 and represent the mean value of five determinations. Values for colipase mRNA content are the average of 3 determinations effected on total RNA from 5 animals. Total denaturated RNA was dotted on nitrocellulose filters and hybridized with the above-characterized cDNA probe, which was previously labeled by nick translation using (^{32}P)dCTP . The specific activity of the probe was (0.5-1)x10^9 cpm/µg. Radioactivity in individual RNA dots was assayed by Cerenkov counting, lines (RNA concentration *versus* cpm hybridized) were obtained by linear regression and the pancreatic level of colipase mRNA was normalized to that of control rats.

and messenger level suggests that, in addition to the transcriptional control,alteration in secretion of the cofactor and/or in translation of the messenger may occur. It is worthstressing here that, under the same nutritional conditions, lipase synthesis was enhanced by 1.6-fold whereas lipase messenger level was 1.8-fold increased, as a result of a modulation of the transcription of the corresponding gene (31).Thus, the non parallelism existing between lipase and colipase messenger concentrations explain the observed differences in the corresponding activity levels.

CONCLUSION

This study is the first reported nucleotide sequence of a pancreatic colipase. If the existence of the procolipase in rat pancreatic juice has been reported (19), its amino acid sequence had not been yet elucidated, as far as we know. The reported nucleotide sequence of a full-length cDNA clone encoding this protein enabled us to give the deduced primary structure of the protein. All the residues that are known to be important for the activity of the cofactor (the binding sites for lipid substrates and lipase) were also found to be present in rat colipase. In addition, we demonstrated that the content in cytoplasmic messenger for rat colipase was able to undergo adaptative variations as a result of dietary manipulations.

REFERENCES

1. Maylié,M.F., Charles,M., Gache,C. and Desnuelle,P. (1971) Biochim. Biophys. Acta 229,286-289.
2. Borgström,B., Wieloch,T. and Erlanson-Albertson,C. (1979) FEBS Lett. 108,407-410.
3. Girard-Globa,A., Bourdel,G. and Lardeux,B. (1980) J. Nutr. 110,1380-1390.
4. Vandermeers-Piret,M.C., Vandermeers,A., Wijns,W.,Rathé,J. and Christophe,J. (1977) Am. J. Physiol. 232, 131-135.
5. Saraux,B., Girard-Globa,A., Ouagued,M. and Vacher,D. (1982) Am. J. Physiol. 243, G10-G15.
6. Wicker,C., Scheele,G. A. and Puigserver,A. (1988) Biochimie 70, 1277-1283.
7. Charles,M., Erlanson,C., Bianchetta,J., Joffre,J., Guidoni,A. and Rovery,M. (1974) Biochim. Biophys. Acta 359, 186-197.
8. Bonicel,J.,Couchoud,P.,Foglizzo,E.,Desnuelle,P. and Chapus,C. (1981) Biochim. Biophys. Acta 669,39-45.
9. Pierrot,M., Astier,J-P, Astier,M.,Charles,M. and Drenth,J. (1982) Eur. J. Biochem. 123, 347-354.
10. Sternby,B., Engström,A., Hellman,U.,Vihert,A.M.,Sternby,N. and Borgström,B. (1984) Biochim. Biophys.Acta 784,75-80.
11. Maxam,A.M. and Gilbert,W. (1980) Methods Enzymol. 65,499-560.
12. Sanger,F.,Nicklen,S. and Coulson,A.R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74,5463-5467.
13.Rovery,M., Boudouard,M. and Bianchetta,J. (1978) Biochim. Biophys. Acta 525,373-379.
14. Chirgwin,J.M., Pryzbyla,A.E., MacDonald,R.J. and Rutter,W.J. (1979) Biochemistry 18, 5294-5299.
15. Thomas,P.S. (1980) Proc. Natl. Acad. Sci. USA 77, 5201-5205.
16. Kozak,M. (1981) Nucl. Ac. Research 9, 5233-5252.
17.MacDonald,R.J., Crerar,M.M., Swain,W.F., Pictet,R.L., Thomas,G. and Rutter,W.J. (1980) Nature,287,117-122.
18.Erlanson,C., Fernlund,P. and Borgström,B. (1973) Biochim. Biophys. Acta 310, 437-445.
19. Erlanson-Albertsson,C. (1981) Biochim. Biophys. Acta 666, 299-300.
20. Larsson,A. and Erlanson-Albertsson,C. (1981) Biochim. Biophys. Acta 664, 538-548.

21. Julien,R., Bechis,G., Gregoire,J., Rathelot,J., Rochat,H. and Sarda,L. (1980) Biochem. Biophys. Res. Commun. 95, 1245-1252.
22. Bosc-Bierne,I., Rathelot,J., Canioni,P., Julien,R., Bechis,G., Gregoire,J., Rochat,H. and Sarda,L. (1981) Biochim. Biophys. Acta 667, 225-232.
23. Borgström,B. (1976) FEBS Lett. 71, 201-204
24. Borgström,B., Wieloch,T. and Erlanson-Albertsson,C. (1979) FEBS Lett. 108, 407-410.
25. Shotton,D.M. and Hartley,B.S. (1970) Nature 225, 802-806.
26. Erlanson,C., Barrowman,J.A. and Borgström,B. (1977) Biochim. Biophys. Acta 489, 150-162.
27. Cozzone,P.J. (1976) FEBS Lett. 69, 153-156.
28. Cozzone,P.J., Canioni,P., Sarda, L. and Kaptein,R. (1981) Eur. J. Biochem. 114, 119-126.
29. Sari,H., Granon,S. and Sémériva,M. (1978) FEBS Lett. 95, 229-234.
30. Wicker,C. and Puigserver,A. (1989) Eur. J. Biochem. 180, 563-567.
31. Wicker,C. and Puigserver,A. Biochem. Biophys. Res. Commun., in press.