Vol. 190, No. 1, 1993 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
January 15, 1993 Pages 134-139

NUCLEOTIDE SEQUENCE OF A *DICTYOSTELIUM DISCOIDEUM* GENE ENCODING A PROTEIN HOMOLOGOUS TO THE YEAST RIBOSOMAL PROTEIN S31

Ursula Hoja, Jörg Hofmann, Rolf Marschalek, and Theo Dingermann[1]

Institut für Biochemie, Med. Fak., Universität Erlangen-Nürnberg, Fahrstraße 17, D-8520 Erlangen, Germany

Received November 25, 1992

A cDNA clone has been isolated whose coding potential is significantly homologous to the yeast ribosomal protein S31. The single copy genomic gene contains a 271 bp intron immediately downstream from the ATG translation initiation codon and is flanked by cannonical exon/intron junctions. The intron carries a CAATCAAT motif which has been described as inducer element for discoidin Iγ expression and which has also been found within the intron of the rp29 gene form *D. discoideum*. The deduced protein contains 110 amino acids and is slightly basic.

Ribosomal proteins belong to a group of high abundant cellular proteins. Sufficient supply of these proteins is either acchieved through redundant genetic information or through highly efficient gene expression. Thus, several eucaryotic organisms are known to contain multiple copies of some of their ribosomal protein genes (1, 2, 3), although frequently at least some of these copies appear to be pseudogenes (4). In other eucaryotes ribosomal protein genes have been found to be single copy, including those which have been cloned from *Drosophila melanogaster* (5, 6) and a few from yeast (7, 8).

The genes for all ribosomal proteins from the cellular slime mold *Dictyostelium discoideum* so far examined are present in a single copy per haploid genome (9). Analysis of these genes has revealed several features which are characteristic for *D. discoideum* genes, and some which have been found in ribosomal protein genes of other eucaryotes, as well. Here we report the structure of a single copy gene from *D. discoideum*, designated rp12, encoding a protein homologous to the yeast ribosomal protein S31.

MATERIALS AND METHODS

Strains and Plasmids

The axenic *Dictyostelium discoideum* strain AX-2 was grown in HL-5 medium (10). Lambda gt11 phage was propagated in *E. coli* strain Y1090 (lacU 169, proA$^+$, Δlon, araD 139, strA, supF trpC 22:Tn 10 (pMC 9)) which was grown in Luria Broth supplemented with ampicillin at a

[1]Present address: Institut für Pharmazeutische Biologie, Johann Wolfgang Goethe-Universität Frankfurt/Main, Georg-Voigt-Straße 16, D-6000 Frankfurt /Main 11, Germany.

0006-291X/93 $4.00

concentration of 50µg/ml (11). Plasmid DNAs were amplified in *E. coli* strain DH5α (endA1, hsdR17(rk^- mk^+), supE44, thi-1, λ^-, recA1, gyrA96, relA1, Δ (arg, lacZYA) U196, φ80, Δ lacZ Δ M15) (12).

Isolation and Sequencing of rp12 clones

Poly A^+ messenger RNA was isolated from growing *D. discoideum* cells. Synthesis of cDNA and cloning into the *Eco*RI site of phage λgt 11 were carried out using components and the protocol of the Amersham cDNA synthesis kit. Particular cDNA clones were selected through two rounds of screening with a polyclonal antibody against yeast cytochrome b_2. From a positive phage the 0.4 kb *Eco*RI fragment was ligated into pUC18. The resulting clone was termed prp12c.

Genomic DNA was isolated from AX-2 cells and digested with the restriction endonuclease *Eco*RI. The fragments were ligated into the *Eco*RI site of the plasmid pUC19. The recombinant plasmids were screened with nick-translated prp12c fragment as a probe. A positive clone (prp12E1/5) contained an *Eco*RI fragment of 4.0 kb.

DNA sequence analysis was performed according to the dideoxy chain termination method (13) using appropriate synthetic oligonucleotides as primer for chain elongation.

Primer extension analysis

Total RNA (40µg) prepared from vegetative growing *D. discoideum* cells was hybridized in 10µl to 10 pmol labelled oligonucleotide U2 (5'-GGCGGCAGCTTTGGATG-3'; 5 x 10^5 cpm) in a buffer containing 50mM Tris/HCl pH 7.0 and 50mM KCl. The annealing reaction was carried out at 68°C for 1h. After chilling on ice, 5µl of buffer was added containing 50mM Tris/HCl pH 7.0, 50mM KCl, 20mM $MgCl_2$, 6mM DTT, 2.5mM dNTPs and 3U AMV reverse transcriptase. Assays were incubated at 42°C for 1h and cDNA synthesis was terminated by ethanol precipitation. Dry pellets were dissolved in 3µl water and 3µl loading buffer (95% formamide, 10mM EDTA, 0.1% bromophenol blue, 0.1% xylene cyanol). Samples were boiled for 5 min, chilled on ice and products were separarted on 8% polyacrylamide gels containing 50% urea (13).

RESULTS AND DISCUSSION

The cDNA plasmid prp12c encodes a protein homologous to the yeast ribosomal protein S31

A cDNA clone coding for a putative ribosomal protein was isolated upon screening a cDNA library from *Dictyostelium discoideum* using a polyclonal antibody raised against cytochrome b_2 from yeast. This cDNA clone contained an *Eco*RI fragment of 384 bp carrying an open reading frame of 333 nucleotides. Surprisingly, the deduced protein sequence was found to lack almost any notable similarity to yeast cytochrome b_2 except for a stretch of 31 amino acids with 10 identical and 8 conservative amino acid substitutions. Instead, a comparison of the predicted amino acid sequence with protein sequences collected in the Swissprot database identified the isolated gene as a putative ribosomal protein gene whose product exhibits significant similarity to the yeast ribosomal protein S31 (14). Analysis of the yeast and *Dictyostelium* sequences showed a homology of about 80% with 40% identical amino acids. A minor gap of 5 amino acids was found at the amino-terminus of the yeast protein (Fig.1). The deduced *D. discoideum* protein has a pI of 11.4 due to 22.5% basic amino acids (lysine or arginine).

Several other *D. discoideum* ribosomal proteins have also been identified on the basis of their sequence homology to known ribosomal proteins from other species. For example, *D. discoideum* rp17 is 69% similar to rat ribosomal protein L7 (15), while the N-terminal half of the V18 protein is similar to yeast L16 (16). The first *D. discoideum* gene encoding a ribosomal protein for the small ribosomal subunit, rp29, has 78% sequence similarity to the yeast ribosomal protein S4 (17).

Dictyostelium rp12 gene contains one intron

Like all other characterized ribosomal protein genes from *D. discoideum* (9, 15, 17) also the rp12 protein is encoded by a single copy gene. Genomic Southern blot analyses probed with labeled insert DNA from prp12c identified the gene on a 4 kb *Eco*RI fragment (data not shown).

```
Y     MPPKQQLSK. ....AAKAAA ALAGGKKSKK KWSKKSMKDR AQHAVILDQE 45
      ||||   :|       |:|||| |  ::|  ::| ||||    ::  : |:: | |
Dd    MPPKAAGGKS KQIQASKAAA KGSSGGAGRK KWSKGRSREK LNNAILFDKE 50

Y     KYDRILKEVP TYRYVSVSVL VDRLKIGGSL ARIALRHLEK EGIIKPISKH 95
       | ::|||:| | : :: :|:  :|:|  ||| || |:: |    |:|| | |
Dd    TYAKLLKEMP TAKVITTAVV SERMKENGSL ARRAIKELLS KGLIKQIIKG100

Y     SKQAIYTRAT A  106
        :::||:|
Dd    HGNGVYTKAQ    110
```

Fig. 1. Comparison of the *Dictyostelium* rp12 protein sequence with the yeast S31 protein. Lines (|) indicate identities and colons (:) indicate conservative substitutions. A gap of five amino acids has been inserted into the yeast sequence for optimal alignment.

In order to investigate the genomic organisation of the *D. discoideum* rp12 gene in detail genomic DNA was digested to completion with *Eco*RI and fragments were ligated into pUC19. Recombinant clones were screened using the cDNA fragment of prp12c as a probe and a plasmid (prp12E1) was isolated containing the predicted *Eco*RI inserts of 4.0 kb. The nucleotide sequence of the gene is presented in figure 2.

The *Dictyostelium* gene contains a single intron which starts immediately after the AUG initiation codon. This unusual location of an intron has been previously described also for the myosin heavy chain gene from *D. discoideum* (18) and for the ribosomal protein L14 from *Xenopus laevis* (19). The intron in the rp12 gene is 271 bp long and contains 77% A/T nucleotides, making it slightly longer and less A/T rich than the average *D. discoideum* introns. In this respect the rp12 intron is similar to the ribosomal protein gene rp1024 from *D. discoideum*. This gene as well shows a relatively long intron of 350 bp with an A/T content of 75% (20). Notable is a $(ATG)_7$ sequence and two CAATCAAT motifs, whose complement, ATTGATTG, has been previously shown to operate as an inducer element for discoidin Iγ expression during growth and early development (21). Such an element has also been found within the intron of the *D. discoideum* rp29 gene encoding another ribosomal protein for the small ribosomal subunit. It has been speculated that this motif possess enhancer-like qualities and may thus be able to act in a position-independent manner (17). Intronic cis-acting sequences have been found within several other genes (22, 23).

Sequences at the splice junctions /5´GTATGT.....TAG 3´/ were found to agree with both the *D. discoideum* and the general consensus (24, 25). The rp12 intron contains 23% G/C nucleotides which is clearly above average of *D. discoideum* introns with average G/C contents of 8%. (26). Also the coding region of the rp12 gene is slightly more G/C rich (40%) than average *D. discoideum* exons (34%). This is also observed in *S. cerevisiae* where ribosomal protein gene exons have a higher G/C content than exons of other genes (26).

Characterization of the transcribed region of the rp12 gene

The 5´start of transcription was mapped by primer extension. The 17mer oligonucleotide 5´-GGCGGCAGCTTTGGATG-3´ complementary to the mRNA starting 57 bases downstream of the intron junction was used as a primer and was extended by reverse transcription in the presence of deoxynucleoside triphosphates. Figure 3 shows that the longest transcript starts with an A-residue at position -30 relative to the ATG translation initiation codon. This result is consistent with capping data in *D. discoideum* where 85% of the mRNAs have a 5´terminal A residue (27). This

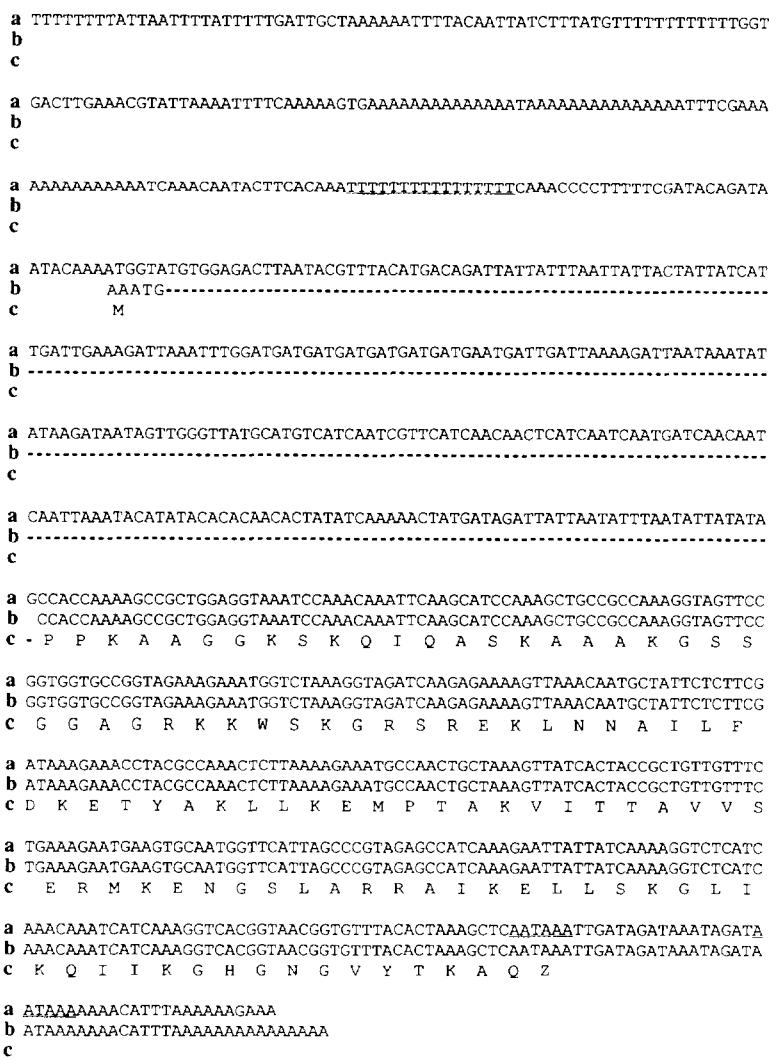

Fig. 2. DNA (**a**), cDNA (**b**), and deduced amino acid (**c**) sequence of the *D. discoideum* gene encoding the putative ribosomal protein. A oligo (dT) stretch in the upstream region and two possible polyadenylation signals are underlined. The intron is marked by a dashed line.

initiation nucleotide resides in a pyrimidine rich region which is strikingly similar to many transcription start sites of ribosomal protein genes in other higher eucaryotes (2, 19). The oligo $(dT)_{16}$ tract just upstream of the initiation site is highly characteristic for *D. discoideum* where all known transcription start sites, except one (28), are preceeded by $(dT)_n$ residues (16). Other transcription initiation sites may occur at positions -26, -25, -18, and -16 (Figure 3).

A TATA box with the consensus sequence TATAAA(T_A)A (25) is normally located 30-40 bp upstream from the transcription initiation site. Apparently, this box is missing in front of the rp12 gene. This is also the case for the rp1024 gene encoding another *D. discoideum* ribosomal protein (20) as well as for several ribosomal protein genes from other organisms (4, 19, 29).

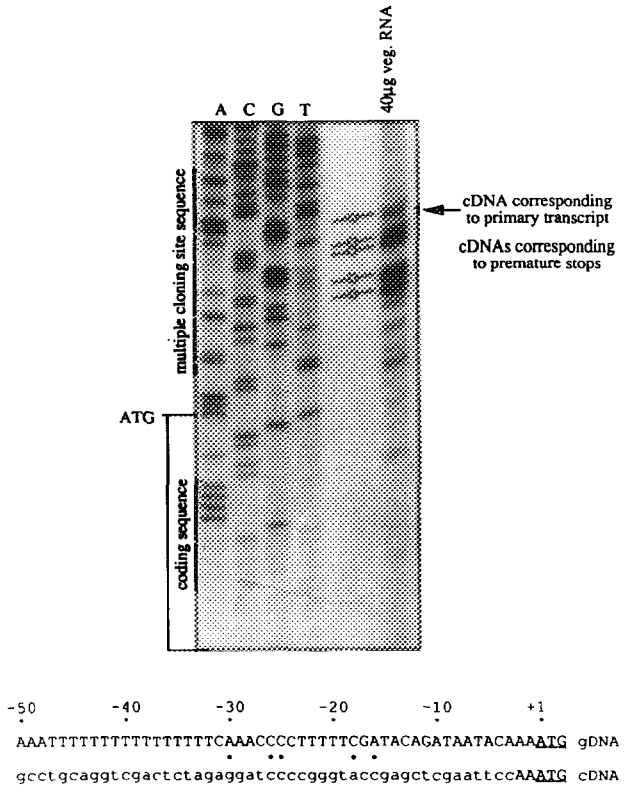

Fig. 3. Primer extension analysis of the 5'start of transcription.
The ^{32}P-end-labelled oligonucleotide U2 binds specificaly to the second exon of the analyzed *D. discoideum* gene. The right lane shows the products of the primer extension reaction. Five putative initiations sites are marked. The sequence of the isolated cDNA clone served as size marker (left four lanes). Sequence analysis was initiated with the phosphorylated oligonucleotide U2 oligonucleotide. The sequence from the isolated genomic clone as well as from the cDNA clone is shown below. Cloning vector sequences are indicated by lower case letters. Dots between both sequences mark the putative initiation sites. Some of the smaller cDNAs might not reflect true initiation sites rather than premature stops of reverse transcriptase during cDNA synthesis.

The sequence immediately surrounding the AUG codon of the rp12 mRNA shows homology to the optimal sequence for translation initiation by eucaryotic ribosomes (30). An invariant A at position -1 is present as well as the consensus A at position -3 relative to the AUG codon. Furthermore, like in rp1024 mRNA (20) also the rp12 mRNA contains a C at -4 which further confirms to the consensus derived from other eucaryotic translation start sites (31, 32, 33).

For rp12 polyadenylation signals (AAUAAA) occur precisely at the end of coding region and 19 nt downstream from the UAA stop codon. A putative polyadenylation signal as part of the translational stop codon has also been described for the *D. discoideum* D11 gene (34).

ACKNOWLEGDMENTS

We thank Walter and Helga Kersten for support and M. Bach for excellent technical assistance. This work was supported by grants from the Deutsche Forschungsgemeinschaft and Fonds der Chemischen Industrie. R.M. was a recipient of a postdoctoral fellowship from the Deutsche Forschungsgemeinschaft.

REFERENCES

1. Monk, R.J., Meyuhas, O., and Perry, R.P. (1981) Cell 24, 301-306.
2. Dudov, K.P., and Perry, R.P. (1984) Cell 37, 457-468.
3. Leer, R.J., van Raamsdonk-Duin, M.C., Molenaar, C.M.T., Witsenboer, H.M.A., Mager, W.H., and Planta, R.J. (1985) Nucl. Acids Res. 13, 5027-5039.
4. Wagner, M., and Perry, R.P. (1985) Mol. Cell. Biol. 5, 3560-3576.
5. Vaslet, C.A., O'Connell, P., Izquierdo, M., and Rosbash, M. (1980) Nature 285, 674-676.
6. Fabijanski, S., and Pellegrini, M. (1982) Gene 18, 267-276.
7. Woolford, J.L., Jr., Hereford, L.M., and Rosbash, M. (1979) Cell 18, 1247-1259.
8. Fried, H.M., Pearson, N.J., Kim, C.H., and Warner, J.R. (1981) J. Biol. Chem. 256, 10176-10183.
9. Steel, L.F., and Jacobson, A. (1986) Gene 41, 165-172.
10. Watts, D.J., and Ashworth, J.M. (1970) Biochem. J. 119, 171-174.
11. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Labatory Press, New York.
12. Hanahan, D. (1983) J. Mol. Biol. 166, 557-580.
13. Sanger, F., Nicklen, S. and Coulsen, A. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463-5467.
14. Nieuwint, R.T.M., Molenaar, C.M.Th., van Bommel, J.H., van Raamsdonk-Duin, M.M.C., Mager, W.H., and Planta, R.J. (1985) Curr. Genet. 10, 1-5.
15. Szymkowski, D.E., and Deering, R.A. (1990) Nucl. Acids Res. 18, 4695-4701.
16. Singleton, C.K., Manning, S.S., and Ken, R. (1989) Nucl. Acids Res. 17, 9679-9692.
17. Proffitt, J.A., Jagger, P.S., Wilson, G.A., Donovan, J.T.J., Widdowson, D.C.C., and Hames, B.D. (1991) Nucl. Acids Res. 19, 3867-3873.
18. Jung, G., Saxe III, C.L., Kimmel, A.R., and Hammer III, J.A. (1989) Proc. Natl. Acad. Sci. U.S.A. 86, 6186-6190.
19. Beccari, E., Mazzetti, P., Mileo, A.M., Bozzoni, I., Pierandrei-Amaldi, P., and Amaldi, F. (1986) Nucl. Acids Res. 14, 7633-7646.
20. Steel, L.F., Smyth, A., Jacobson, A. (1987) Nucleic Acids Res. 15, 10285-10298.
21. Vauti, F., Morandini, P., Blusch, J., Sachse, A., and Nellen, W. (1990) Mol. Cell. Biol. 10, 4080-4088.
22. Sternberg, E.A., Spizz, G., Perry, W.M., Vizard D., Weil, T., and Olson, E.N. (1988) Mol. Cell. Biol. 8, 2896-2909.
23. Ng, S.Y., Gunning, P., Liu, S.H., Leavitt, J., and Kedes, L. (1989) Nucl. Acids Res. 17, 601-615.
24. Breathnach, R., and Chambon, P. (1981) Ann. Rev. Biochem. 50, 349-384.
25. Kimmel, A.R., and Firtel, R.A. (1982) In Loomis, W.F. (ed.), The Development of *Dictyostelium discoideum*, Academic Press, New York, 233-324.
26. Csank, C., Taylor, F.M., and Martindale, D.W. (1990) Nucl. Acids Res. 18, 5133-5141.
27. Kimmel, A.R., and Firtel, R.A. (1983) Nucl. Acids Res. 11, 541-552.
28. Ragheb, J.A., and Dottin, R.P. (1987) Nucl. Acids Res. 15, 3891-3906.
29. Wiedemann, L.M., and Perry,R.P. (1984) Mol. Cell. Biol. 4, 2518-2528.
30. Kozak, M. (1986) Cell 44, 283-292.
31. Hamilton, R., Watanabe, C.K., and de Boer, H.A. (1987) Nucl. Acids Res. 15, 3581-3593.
32. Cavener, D.R. (1987) Nucl. Acids Res. 15, 1353-1361.
33. Lütcke, H.A., Chow, K.C., Mickel, F.S., Moss, K.A., Kern, H.F., and Scheele, G.A. (1987) EMBO J. 6, 43-48.
34. Barklis, E., Pontius, B., and Lodish, H.F. (1985) Mol. Cell. Biol. 5, 1473-1479.