Identification and partial DNA sequence of the gene for the α-subunit of the ATP synthase complex of *Chlamydomonas reinhardii* chloroplasts

Richard B. Hallick*

Department of Chemistry, Campus Box 215, University of Colorado, Boulder, CO 80309, USA

Received 18 September 1984

The partial DNA sequences of two unidentified genes flanking the gene for the large subunit of ribulose bisphosphate carboxylase of *Chlamydomonas reinhardii* have been reported [(1982) J. Mol. Biol. 162, 775–793]. Based on a comparison of the derived amino acid sequence of one of these genes with the corresponding sequences from *Nicotiana tabacum* chloroplast DNA and the *E. coli* atp (unc) operon, one *Chlamydomonas* gene is identified as coding for the α-subunit of the ATP synthase complex.

Chlamydomonas *Chloroplast DNA* *atpA* *ATP synthase* *DNA sequence*

1. INTRODUCTION

The chloroplast ATP synthase complex is composed of an intrinsic thylakoid membrane domain (CF_0), and an extrinsic domain located on the stromal surface of the thylakoids (CF_1). Like its mitochondrial and bacterial counterparts, the CF_0–CF_1 complex catalyzes the synthesis of ATP in response to a proton flux across the photosynthetic membranes. The chloroplast ATP synthase consists of at least 8 polypeptides, including 3 CF_0 components designated I–III, and 5 CF_1 components designated α–ϵ [1–2]. In spinach chloroplasts, three of the CF_1 subunits (α, β, and ϵ) and two CF_0 subunits (polypeptides I and III) are encoded in the plastid DNA and synthesized on chloroplast ribosomes. The other subunits are encoded in the nucleus, translated in the cytoplasm and, transported into the organelle for processing and assembly [3].

Recently the chloroplast genes for 4 different ATP synthase subunits have been identified in higher plant chloroplast genomes and sequenced. These include the atpBE locus (for nomenclature see [4]) for the CF_1 polypeptides β and ϵ of spinach, maize, tobacco, and barley [5–8], the atpA locus for the CF_1 α-polypeptide of tobacco [9], and the atpH locus for the DCCD-binding proteolipid (subunit III) of CF_0 from wheat and spinach [10,11]. By comparison of this published data to a sequenced, but unidentified region of the *Chlamydomonas reinhardii* chloroplast DNA [12], it has been possible to identify the amino terminus and first 112 codons of the gene for the α-subunit of the CF_1 component of the ATP synthase complex.

2. MATERIALS AND METHODS

DNA sequences were retrieved from the Los Alamos DNA sequence data base via a communications link, and manipulated with the computer programs for the IBM-PC described in [13]. Homologies were detected by visual alignment of amino acid sequences, as well as with the computer program PROTPLOT and the IBM-PC computer [14].

* Present address: Department of Biochemistry, Biological Sciences West, University of Arizona, Tucson, AZ 85721, USA

Published by Elsevier Science Publishers B.V.

3. RESULTS AND DISCUSSION

Authors in [12] reported the sequence of 4037 bp of DNA from *Eco*RI fragment R15 of *C. reinhardii* chloroplast DNA. These data included the rbcL gene which encodes the large subunit of ribulose-1,5-bisphosphate carboxylase, and two unidentified flanking genes. One of these genes, designated the x-gene, was of opposite polarity to rbcL, arranged in a 5′- to 5′-orientation in terms of the direction of transcription. The x-gene gave rise to a transcript of 2.4 kb as detected by RNA blot hybridization [12]. The second gene, termed the y-gene, was arranged 3′- to 3′- to rbcL, and had a 2.2 kb transcript. When the derived amino acid sequence of the x-gene was compared to other known chloroplast genes, it was apparent that it had a very high homology to the atpA gene of *Nicotiana tabacum* chloroplast DNA [9], and also the gene for the α-subunit of the *Escherichia coli* ATP synthase from the atp (unc) operon [15,16]. A comparison of the amino terminal 112 amino acids of these 3 genes is shown in fig.1.

The derived amino acid sequence from 112 codons of the x-gene is identical in 72 positions to the *N. tabacum* atpA locus. This is a homology of 64%. The same x-gene region has 42% homology (47 of 112 codons) to the *E. coli* gene for the α-subunit of the ATP-synthase. The 40 amino acid substitutions out of 380 possibilities that are known to occur with a higher than average frequency in homologous proteins have been tabulated [19]. If these are counted in addition to identical amino acids, then the x-gene has homologies of 86 and 63%, respectively, to the *N. tabacum* chloroplast atpA gene and the *E. coli* α-subunit gene (not shown). Since the *N. tabacum* and *E. coli* amino acid sequences for the 112 amino terminal residues are only 44% identical (49 of 112 codons) and 68% homologous if frequent amino acid substitutions are considered, and this homology was the primary basis for identification of the *N. tabacum* gene [9], the assignment of the x-gene of *C. reinhardii* chloroplast DNA as atpA based solely on a comparison of these amino acid homologies is justifiable.

The ATG initiator codons of the *C. reinhardii* rbcL and atpA genes are separated by a spacer region of 854 bp [12]. The 50 bp preceding the start codon are very AT-rich (92%), and contain no obvious ribosome binding site. It is noteworthy that in many higher plant chloroplast genomes, the rbcL gene is organized in a similar 5′- to 5′-fashion, but with the gene for the β subunit of the ATP synthase, atpB [5–7].

Since only 112 amino terminal codons of atpA are present on the *Eco*RI fragment R15, unless introns are present, the remainder of the gene must be on the adjacent fragment, R7 [12]. There have been two reports citing this fragment (but not R15) as the probable locus for atpA. It was found [17] that R7 yielded a polypeptide which could be immunoprecipitated with a mixed antibody to α- and β-subunits following in vitro transcription and translation of cloned R7. Authors in [18] detected heterologous hybridization with a spinach atpA derived DNA sequence to this same *Eco*RI fragment. Therefore these two observations have been confirmed and extended. It is now possible to identify R15 as the coding site for the atpA amino terminus, and to describe the exact location, polarity, and partial amino acid sequence of the atpA gene.

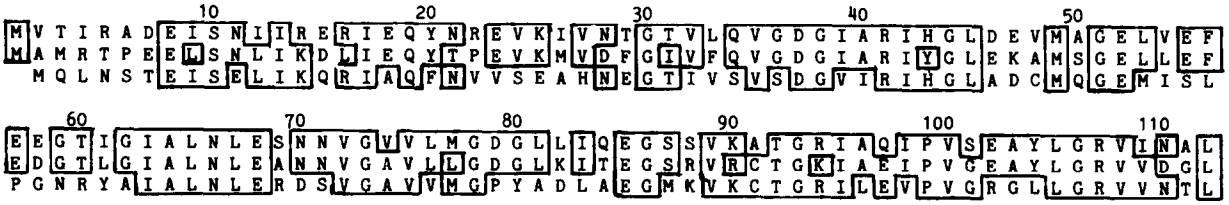

Fig.1. Comparison of the amino terminal sequences of the α subunit of the ATP synthase of *N. tabacum* (top), *C. reinhardii* (middle), and *E. coli* (bottom). The *N. tabacum* sequence is from [9], and the *E. coli* sequence from [15,16]. The *C. reinhardii* atpA amino acid sequence is from the x-gene reported by authors in [12], with 26 additional codons added at the amino terminus from their published DNA sequence. The boxed areas contain amino acids shared in common by 2 or 3 of the sequences. Of 112 amino acids compared, 37 are shared in common by all 3 polypeptides.

ACKNOWLEDGEMENT

This work was supported by NIH Grant GM21351.

REFERENCES

[1] Pick, U. and Racker, E. (1979) J. Biol. Chem. 254, 2793–2799.
[2] Nelson, N., Nelson, H. and Schatz, G. (1980) Proc. Natl. Acad. Sci. USA 77, 1361–1364.
[3] Westhoff, P., Nelson, N., Buneman, H. and Herrmann, R.G. (1981) Curr. Genet. 4, 109–120.
[4] Hallick, R.B. and Bottomley, W. (1983) Plant Mol. Biol. Reporter 1, 38–43.
[5] Zurawski, G., Bottomley, W. and Whitfeld, P.R. (1982) Proc. Natl. Acad. Sci. USA 79, 6260–6264.
[6] Krebbers, E.T., Larrinua, I.M., McIntosh, L. and Bogorad, L. (1982) Nucleic Acids Res. 10, 4985–5002.
[7] Shinozaki, K., Deno, H., Kato, A. and Sugiura, M. (1983) Gene 24, 147–155.
[8] Zurawski, G. and Clegg, M.T. (1984) Nucleic Acids Res. 12, 2549–2559.
[9] Deno, H., Shinozaki, K. and Sugiura, M. (1983) Nucleic Acids Res. 11, 2185–2191.
[10] Howe, C.J., Auffret, A.D., Doherty, A., Bowman, C.M., Dyer, T.A. and Gray, J.C. (1982) Proc. Natl. Acad. Sci. USA 79, 6903–6907.
[11] Alt, J., Winter, P., Sebald, W., Moser, J.G., Schedel, R., Westhoff, P. and Herrmann, R.G. (1983) Curr. Genet. 7, 129–138.
[12] Dron, M., Rahire, M. and Rochaix, J.D. (1982) J. Mol. Biol. 162, 775–793.
[13] Conrad, B. and Mount, D.W. (1982) Nucleic Acids Res. 10, 31–38.
[14] Mount, D.W. and Conrad, B. (1984) Nucleic Acids Res. 12, 811–817.
[15] Gay, N.J. and Walker, J.E. (1981) Nucleic Acids Res. 9, 2187–2194.
[16] Kanazawa, H., Kayano, T., Mabuci, K. and Futai, M. (1981) Biochem. Biophys. Res. Commun. 103, 604–612.
[17] Rochaix, J.D. (1981) Experientia 37, 323–332.
[18] Woessner, J.P., Masson, A., Harris, E.H., Bennoun, P., Gillham, N.W. and Boyton, J.E. (1984) Plant Mol. Biol. 3, 177–190.
[19] McLachlan, A.D. (1971) J. Mol. Biol. 61, 409–424.