Volume 232, number 2, 364–368 FEB 05888 May 1988

Photosynthetic reaction centre of *Chloroflexus aurantiacus*

Primary structure of M-subunit

Yu.A. Ovchinnikov†, N.G. Abdulaev, B.E. Shmuckler, A.A. Zargarov, M.A. Kutuzov, I.N. Telezhinskaya, N.B. Levina and A.S. Zolotarev

Shemyakin Institute of Bioorganic Chemistry, USSR Academy of Sciences, ul. Miklukho-Maklaya 16/10, 117871 GSP Moscow V-437, USSR

Received 24 March 1988

The M-subunit primary structure of the reaction centre (RC) from *Chloroflexus aurantiacus* composed of 306 amino acid residues has been determined by parallel analysis of the protein and corresponding DNA. The blocked N-terminus as well as replacement of the essential histidine liganding Mg of an accessory bacteriochlorophyll in purple bacteria by leucine distinguishes the M-subunit of *Chloroflexus* RC from that of purple bacteria.

Photosynthesis; Reaction center; Subunit structure; Amino acid sequence; Nucleotide sequence; *(Chloroflexus aurantiacus)*

1. INTRODUCTION

The reaction centre (RC) is a protein-pigment complex where the primary events of light energy conversion into chemical energy occur. The reaction centre of filamentous green bacterium *Chloroflexus aurantiacus*, the smallest RC known today, consists of two protein subunits [1]. The primary structure of one subunit has been recently determined in this laboratory by parallel analysis of the protein and corresponding DNA [2]. The homology between this subunit and L-subunits of the RC of purple bacteria indicates that the functional analogy of *Chloroflexus* and Rhodospirillaceae RC relies upon their structural similarity. That is why one may propose the second subunit of *Chloroflexus* RC to be homologous to the M-subunit of the purple bacteria. This paper presents the primary structure and tentative folding of the polypeptide chain of the *Chloroflexus* RC M-subunit.

Correspondence address: N.G. Abdulaev, Shemyakin Institute of Bioorganic Chemistry, USSR Academy of Sciences, ul. Miklukho-Maklaya 16/10, 117871 GSP Moscow V-437, USSR

† Deceased

The nucleotide sequence presented here has been submitted to the EMBL/GenBank database under the accession no. Y00972

2. MATERIALS AND METHODS

Experimental procedures applied for separation of highly purified preparation of *Chloroflexus* RC, cleavage of the proteins, separation of resulting peptides, and construction of the genomic library and DNA sequence analysis are presented in a previous paper [2] from this laboratory.

3. RESULTS

Clone pRCa9 from the *Chloroflexus* genomic library was shown [2] to contain approximately one third of the gene encoding the N-terminal sequence of the M-subunit. In order to identify the clone that includes the missing sequence of this subunit, the first genomic library of 20000 transformants was screened with oligonucleotide probe 1b [2]. As a result the only positive clone, pRCA5, was isolated (fig.1). The nick-translated insert of this clone was used to screen 30000

Published by Elsevier Science Publishers B.V. (Biomedical Division)

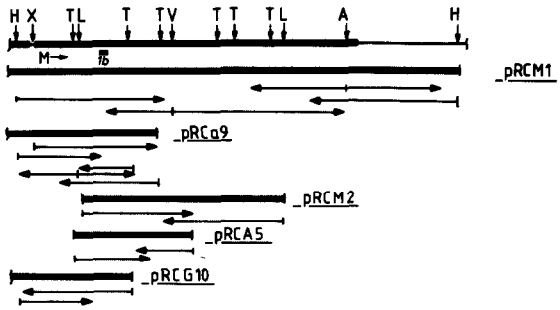

Fig.1. Location of insertions of isolated clones in restriction map of the *Chloroflexus* RC M-subunit and sequencing strategy. A, *Ava*II; H, *Hin*dIII; L, *Alu*I; T, *Taq*I; X, *Xba*I; V, *Ava*I.

transformants of the same library. Of seven positive clones, the two, pRCM1 and pRCM2, with the largest inserts were used for sequencing.

The nucleotide sequence of the gene encoding the M-subunit and the deduced amino acid sequence are shown in fig.2. The sequences of peptides derived from analysis at the protein and gene levels match perfectly.

Only one subunit of the RC preparation can be identified by stepwise Edman degradation [2,4]. The sequence thus obtained corresponds to that of the L-subunit. This indicates that the N-terminal amino group of the M-subunit is posttranslationally modified. Treatment of the RC preparation with *Staphylococcus aureus* protease and trypsin results in the two peptides with blocked N-termini. According to amino acid analyses these peptides correspond to the sequences X-A-T-I-N-M-T-P-G-D-L-E and X-A-T-I-N-M-P-G-D-L-E-L-G-R deduced from the DNA analysis. The difference in 60 Da was obtained when comparing the results of FAB mass spectrometry and amino acid analysis of both peptides. Chemical characterisation of the blocking group is a subject for further studies.

The amino acid sequence preceding the termination codon of the M-subunit coincides with that of the C-terminal peptide of this subunit (see [2]). Thus the mature M-subunit of *Chloroflexus* RC contains 306 amino acid residues and has a molecular mass of 34982 Da (including the blocking group).

```
1    ATG GCG ACA ATC AAT ATG ACG CCG GGC GAT CTG GAA CTG GGT CGT GAT CGC GGT CGG ATC GGA AAA CCA ATT GAG ATT CCG CTC CTG GAA
1    (M) A   T   I   N   M   T   P   G   D   L   E   L   G   R   D   R   G   R   I   G   K   P   I   E   I   P   L   L   E

91   AAC TTC GGC TTC GAT AGC CAG CTC GGC CCG TTC TAT CTC GGC TTT TGG AAC GCA GTG GCT TAC ATC ACC GGT GGT ATC TTT ACC TTT ATC
30   N   F   G   F   D   S   Q   L   G   P   F   Y   L   G   F   W   N   A   V   A   Y   I   T   G   G   I   F   T   F   I

181  TGG CTG ATG GTG ATG TTT GCT CAG GTG AAT TAC AAT CCG GTG GCC TTT GCC AAG TAC TTT GTG GTC TTG CAA ATC GAT CCG CCG TCA TCA
60   W   L   M   V   M   F   A   Q   V   N   Y   N   P   V   A   F   A   K   Y   F   V   V   L   Q   I   D   P   P   S   S

271  CGC TAT GGC CTG AGC TTT CCG CCG CTG AAT GAG GGT GGA TGG TGG CTC ATC GCG ACC TTC TTC CTG ACG GTC TCG ATC TTC GCG TGG TAT
90   R   Y   G   L   S   F   P   P   L   N   E   G   G   W   W   L   I   A   T   F   F   L   T   V   S   I   F   A   W   Y

361  ATG CAC ATC TAT ACC CGC GCC AAG GCG CTC GGG ATT AAG CCC TAC CTG GCT TAC GGC TTC ACC GGT GCG ATT GCG CTC TAC CTG GTG ATC
120  M   H   I   Y   T   R   A   K   A   L   G   I   K   P   Y   L   A   Y   G   F   T   G   A   I   A   L   Y   L   V   I

451  TAT ATC ATT CGT CCG GTG TGG ATG GGT GAT TGG TCG GAA GCG CCG GCC CAC GGC ATC AAA GCG CTG CTC GAC TGG ACG AAT AAT GTG AGC
150  Y   I   I   R   P   V   W   M   G   D   W   S   E   A   P   A   H   G   I   K   A   L   L   D   W   T   N   N   V   S

541  GTG CGT TAC GGC AAT TTC TAC TAC AAT CCG TTC CAC ATG CTC TCG ATC TTC TTC TTG CTC GGA TCA ACG TTG TTG CTG GCG ATG CAC GCC
180  V   R   Y   G   N   F   Y   Y   N   P   F   H   M   L   S   I   F   F   L   L   G   S   T   L   L   L   A   M   H   A

631  GGT ACC ATT TGG GCA CTC GAA AAG TAC GCT GCC CAC GAG GAA TGG AAT GAG ATT CAG GCT CCC GGT ACC GGT ACC GAG CGT GCT CAG CTC
210  G   T   I   W   A   L   E   K   Y   A   A   H   E   E   W   N   E   I   Q   A   P   G   T   G   T   E   R   A   Q   L

721  TTC TGG CGC TGG TGC ATG GGC TTT AAC GCG AAT GCC TAC TCG ATC CAC CTG TGG GCA TTC TGG TTC GCC TGG TTG TGC GGT ATT ACC GGT
240  F   W   R   W   C   M   G   F   N   A   N   A   Y   S   I   H   L   W   A   F   W   F   A   W   L   C   G   I   T   G

811  GCA TTG GGT GTC TTC TTC TCA ATG CCC GAT TTT GTC AAT AAC TGG TTC CAA TGG GGT ATT GAA GCC GGC ATC AAC TAT CCG CAA GGT CCA
270  A   L   G   V   F   F   S   M   P   D   F   V   N   N   W   F   Q   W   G   I   E   A   G   I   N   Y   P   Q   G   P

901  ACG CCA CCG GTT TCA CTG CCG TAG CGTTCGTGACACTTCTCTACCTTCCTCCCGGCCTCTCGCCCGTCATCAGACGGGCGAGAGGTTGTTTATTGCACGGGGAGAGGTGAC
300  T   P   P   V   S   L   P   Stop

1112 TATCTGAAGACTGGTGTATCAAATCTGGCTTCATAACAGTCATTGCACGCCGTTTTCAGGTGTGCCACCGGCGCGTCGGTGGCTTGTTGTCGGTGGGCCGGTAGGGCACGGCATGCCGT

1231 GCCCCTACGCCTCCTGTCGTACCAGTAGAGACGCATGCGTCGCCCCTACCAGCGTCCAGCGGTGAATTGTGTAGGTGCTGCATCAGGGGGATC
```

Fig.2. Nucleotide sequence of the gene coding for the M-subunit and derived amino acid sequence. The positions of isolated peptides are indicated by |——|; residues identified by Edman degradation are designated by ——. In the 3′-untranslated region the sequence with features typical of Rho-independent terminator is underlined [3].

```
C. AUR.          X-ATINMTPGDLELGRDRGRIGKPIEIPLLENFGFDSQLGPFYLGFWNAVAY   50
R. VIR.   ADYQTIYTQIQARGPHITVSGEWGDNDRVGKPFYSYWL-GKIGDAQIGPIYLGASGIAAF   59
R. SPH.   AEYQNIFSQVQVRGPADLGMTEDVNLANRSGVGPFSTLLGWFGNAQLGPIYLGSLGVLSL   60
R. CAPS.  AEYQNFFNQVQVAGAPEMGLKEDVDTFERTPAGMFNIL-GWMGNAQIGPIYLGIAGTVSL   59

C. AUR.   ITGGIFTFIWLMVMFAQVNYNPVAFAKYFVVLQIDPPSSRYGLSF-PPLNEGGWWLIATF  109
R. VIR.   AFGSTAILIILFNMAAEVHFDPLQFFRQFFWLGLYPPKAQYGMGI-PPLHDGGWWLMAGL  118
R. SPH.   FSGLMWFFTIGIWFWYQAGTNPAVFLRDLFFFSLEPPAPEYGLSFAAPLKEGGLWLIASF  120
R. CAPS.  AFGAAWFFTIGVWYWYQAGFDPFIFMRDLFFFSLEPPPAEYGLAI-APLKQGGVWQIASL  118

C. AUR.   FLTVSIFAWYMHIYTRAKALGIKPYLAYGFTGAIALYLVIYIIRPVWMGDWSEAPAHGIK  169
R. VIR.   FMTLSLGSWWIRVYSRARALGLGTHIAWNFAAAIFFVLCIGCIHPTLVGSWSEGVPFGIW  178
R. SPH.   FMFVAVWSWWGRTYLRAQALGMGKHTAWAFLSAIWLWMVLGFIRPILMGSWSEAVPYGIF  180
R. CAPS.  FMAISVIAWWVRVYTRADQLGMGKHMAWAFLSAIWLWSVLGFWRPILMGSWSVAPPYGIF  178

C. AUR.   ALLDWTNNVSVRYGNFYYNPFHMLSIFFLLGSTLLLAMHAGTIWALEKYAAHEEWNEIQA  229
R. VIR.   PHIDWLTAFSIRYGNFYYCPWHGFSIGFAYGCGLLFAAHGATILAVARFGGDREIEQITD  238
R. SPH.   SHLDWTNNFSLVHGNLFYNPFHGLSIAFLYGSALLFAMHGATILAVSRFGGERELEQIAD  240
R. CAPS.  SHLDWTNQFSLDHGNLFYNPFHGLSIAALYGSALLFAMHGATILAVTRPGGERELEQIVD  238

C. AUR.   PGTGTERAQLFWRWCMGFNANAYSIHLWAFWFAWLCGITGALGVFFSMPDFVNNWFQWGI  290
R. VIR.   RGTAVERAALFWRWTIGFNATIESVHRWGWFFSLMVMVSASVGILLTGT-FVDNWYLWCV  297
R. SPH.   RGTAAERAALFWRWTMGFNATMEGIHRWAIWMAVLVTLTGGIGILLSGT-VVDNWYVWGQ  299
R. CAPS.  RGTASERAALFWRWTMGFNATMEGIHRWAIWMAVMVTLTGGIGILLSGT-VVDNWYVWAQ  297

C. AUR.   EAGINYPQGPTPPVSLP*
R. VIR.   KHGAAPDYPAYLPATPDPASLPGAPK*
R. SPH.   NHGMAPLN*
R. CAPS.  VHGYAPVTP*
```

Fig.3. Alignment of amino acid sequences of the RC M-subunits from *Chloroflexus* (C. AUR.), *R. viridis* (R. VIR.), *R. sphaeroides* (R. SPH.), and *R. capsulata* (R. CAPS.) [2–6]. Stars above the sequence indicate a match. X = 60 Da.

```
M                          X-ATINMTPGDLELGRDRGRIGKPIEIPLLENFGFDSQL   37
L   SRAKAKDPRFPDFSFTVVEGARATRVPGGRTIEEIEPEYKIKGRTTFSAIFRYDPFDFWV   60

M   GPFYLGFWNAVAYITGGIFTFIWLMVMFAQV-NYNPVAFAKYFVVLQIDPPSSRYGLSF-   95
L   GPFYVGFWGFVSVIGIIFGSYFYINETILKGPYSIPQNF----FAGRIDPPPPELGLGFA  116

M   PPLNEGGWWLIATFFLTVSIFAWYMHIYTRAKALGIKPYLAYGFTGAIALYLVIYIIRPV  155
L   APGEPGFAWQMTVLFATIAFFGWMMRQVDISMKLDMGYHVPIAFGVAFSAWLVLQVIRPI  176

M   WMGDWSEAPAHGIKALLDWTNNVSVRYGNFYYNPFHMLSIFFLLGSTLLLAMHAGTIWAL  215
L   ALGMWHEGFVLGIMPHLDWVSNFGYRYNNFFYNPFHAIGITGLFASTWLLACHGSLILSA  236

M   EKYAAHEEWNEIQAPGTGTERAQLFWRWCMGFNANAYSIHLWAFWFAWLCGITGALGVFF  275
L   AQY----------RGPEGGDIENVFFRDVQYYSVGESGVHRLGYIFAIGGILSADLCILL  296

M   SMPDFVNNWFQWGIEAGINYPQGPTPPVSLP*
L   SGWPVQDWVSFWNFWNNLPFWSGV*
```

Fig.4. Comparison of amino acid sequences of the L- and M-subunits from *Chloroflexus*. Positions of the putative transmembrane helices are indicated by lines with letters indicating the corresponding helix.

4. DISCUSSION

Comparison of the M-subunit primary structures from *Chloroflexus* and purple bacteria demonstrates the high degree of homology. As shown in fig.3, 77 amino acid residues are conserved in four species. It should be noted that considerable homology (52 conserved amino acid residues) is also observed when comparing the M-subunit of *Chloroflexus* and L-subunits of purple bacteria (not shown). Designating *Chloroflexus* RC subunits L and M similarly to those of purple bacteria, we implied their similar functions, such as chromophore binding and electron transport. Correctness of the designation of subunits is confirmed by their comparison (fig.4). Unlike M-subunits, L-subunits of purple bacteria contain two deletions: the first between membrane-spanning helices A and B and the second between D and E [7]. Analogous deletions, although of different size (the first is 3 amino acid residues shorter; the second 3 amino acid residues longer) are conserved in the *Chloroflexus* RC L-subunit.

In fact, the M-subunit of *Chloroflexus* RC is several amino acid residues shorter than the L-subunit. This small difference in molecular masses

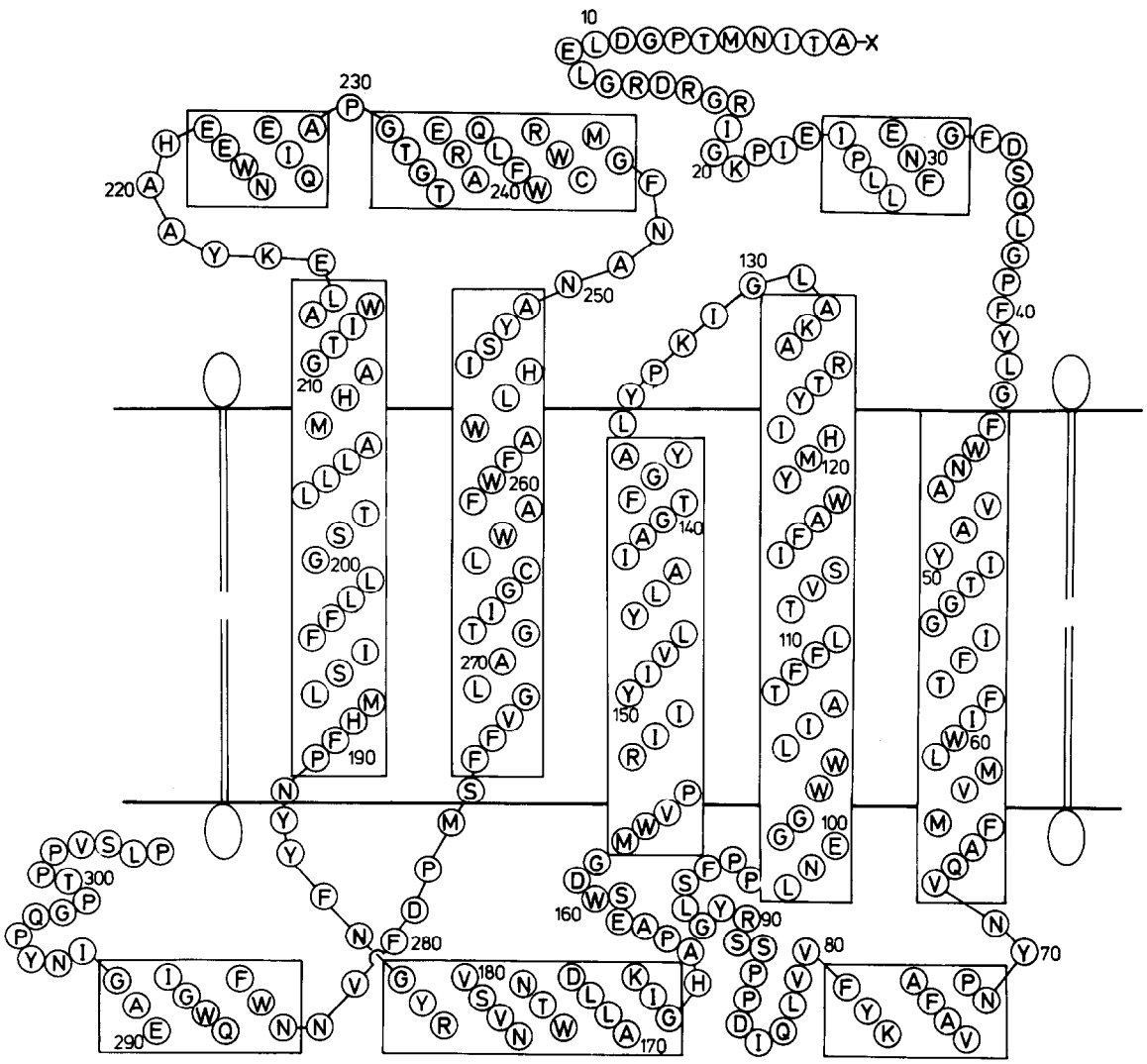

Fig.5. Predicted folding of the amino acid sequence of the RC M-subunit from *Chloroflexus*.

may explain the observed comigration of these subunits in PAGE [8].

Previously, we have suggested that the folding of the L- and M-subunits of purple bacteria and *Chloroflexus* should be similar (fig.5, see also fig.4 in [2]). If this is the case, a number of characteristic features distinguishing M- from L-subunits should also exist in *Chloroflexus*. In fact, extension of the peptide loop between membrane-spanning helices D and E is sufficient to form an additional α-helical segment de′ (amino acid residues 222–229) which does not exist in the L-subunit [9,10]. Furthermore, three amino acid residues of the M-subunit participate in binding of the iron atom (only two amino acid residues of the L-subunit are ligands of non-haem iron). These three amino acid residues (His_{208}, Ala_{223} and His_{255}) are conserved in the M-subunit of the *Chloroflexus* RC. It is known from X-ray data that Mg atoms of accessory bacteriochlorophylls are liganded by histidine residues. In the L-subunit of *Chloroflexus* RC the corresponding histidine residue is conserved [2]. Moreover, in the *Chloroflexus* RC one of the accessory bacteriochlorophylls is replaced by bacteriopheophytin [11]. Moreover, analysis of spectroscopic data on *Chloroflexus* RC indicates that this bacteriopheophytin should be in the subunit analogous to the M-subunit of purple bacteria [12]. If this is true, the histidine residue liganding Mg of accessory bacteriochlorophyll in the M-subunit of purple bacteria could well be replaced by other amino acid residues in *Chloroflexus*. Indeed, Leu_{171} occupies the appropriate position in the *Chloroflexus* M-subunit.

REFERENCES

[1] Pierson, B.K., Thornber, J.P. and Seftor, R.E.B. (1983) Biochim. Biophys. Acta 723, 322–326.
[2] Ovchinnikov, Yu.A., Abdulaev, N.G., Zolotarev, A.S., Shmukler, B.E., Zargarov, A.A., Kutuzov, M.A., Telezhinskaya, I.N. and Levina, N.B. (1988) FEBS Lett. 231, 237–242.
[3] Rosenberg, M. and Court, D. (1979) Annu. Rev. Genet. 13, 319–353.
[4] Shiozawa, J.A., Lottspeich, F. and Feick, R. (1987) Eur. J. Biochem. 167, 595–600.
[5] Youvan, D.C., Bylina, E.J., Alberti, M., Begusch, H. and Hearst, J.E. (1984) Cell 37, 949–957.
[6] Williams, J.C., Steiner, L.A., Ogden, R.C., Simon, M.I. and Feher, G. (1983) Proc. Natl. Acad. Sci. USA 80, 6505–6509.
[7] Michel, H., Wever, K.A., Gruenberg, H., Dunger, I., Oesterhelt, D. and Lottspeich, F. (1986) EMBO J. 5, 1149–1158.
[8] Blankenship, R.E., Feick, R., Bruce, B.D., Kirmaier, C., Holten, D. and Fuller, R.C. (1983) J. Cell. Biochem. 22, 251–261.
[9] Deisenhofer, J., Epp, O., Miki, K., Huber, R. and Michel, H. (1985) Nature 318, 618–623.
[10] Allen, J.P., Feher, G., Yeates, T.O., Komiya, H. and Rees, D.C. (1987) Proc. Natl. Acad. Sci. USA 84, 6162–6166.
[11] Blankenship, R.E. and Fuller, R.C. (1986) Encycl. Plant Physiol. 19, 390–399.
[12] Vasmel, H., Amesz, J. and Hoff, A.J. (1986) Biochim. Biophys. Acta 852, 159–168.