Volume 231, number 1, 237-242 FEB 05775 April 1988

Photosynthetic reaction centre of *Chloroflexus aurantiacus*

I. Primary structure of L-subunit

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

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

Received 18 February 1988

The L-subunit primary structure of the reaction centre from *Chloroflexus aurantiacus* composed of 310 amino acid residues has been determined by parallel analysis of the protein and corresponding DNA. Significant homology between this protein and L-subunits from reaction centres of purple bacteria is observed. This implies close similarity in the tertiary structure of these proteins.

Photosynthesis; Reaction center, L-subunit; Amino acid sequence; Nucleotide sequence; (*Chloroflexus aurantiacus*)

1. INTRODUCTION

The photosynthetic apparatus of green thermophilic bacteria *Chloroflexus aurantiacus* consists of the light absorbing pigment-protein complex (antenna) and photochemical reaction centre (RC) [1]. Though mostly similar in spectral characteristics to purple bacteria RC [2,3] composed of three subunits, those of *Chloroflexus* are formed by two subunits [4,5]. This is the simplest RC known at present.

Here we present the L-subunit amino acid sequence (designated by analogy with L-subunits of purple bacteria) deduced from the nucleotide sequence of a corresponding gene.

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

Abbreviations: LDAO, lauryldimethylamine *N*-oxide; RC, reaction centre

† Deceased

The nucleotide sequence presented here has been submitted to the EMBL/GenBank database under the accession number

2. MATERIALS AND METHODS

RCs were isolated according to [3] with some modifications: (i) the cells disrupted in a cooled cell of the French press (twice); (ii) the LDAO extract centrifuged for 2 h, at 30000 × *g*, the floating carotenoids removed and supernatant applied to a DE-52 cellulose column. pH of the eluted fraction containing RC was adjusted to 9.0 with 1 M Tris. This solution was diluted twice to reduce the NaCl concentration and applied to Mono Q column (FRLC-System, Pharmacia). RCs were eluted in the NaCl gradient (50 mM Tris-HCl, pH 9.0, 0.1% LDAO).

Immobilization of denatured polypeptides on activated glass (CPG-Thiol, Pierce), their cleavage with cyanogen bromide and enzymes as well as peptide separation were carried out as described [6]. The amino acid sequence was determined by the manual [7] and automated Edman degradation on a gas-phase sequenator (model 470A, Applied Biosystems). The C-terminal amino acid residues were established using carboxypeptidase A and by hydrazinolysis [8,9].

To obtain the genome library, *Chloroflexus* DNA was partially hydrolyzed by endonuclease *Sau*3AI, the fragments formed were cloned into vector pUC8 cut with endonuclease *Sal*GI as in [10]. A second library was generated by ligation of blunt fragments resulting from DNA partially cleaved with endonuclease *Alu*I and vector pUC8 cut with endonuclease *Sma*I. Oligonucleotide probes were synthesized by phosphoramidite method on an Applied Biosystems synthesizer. Nucleotide sequences were determined by the modified Maxam-Gilbert method [11,12].

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/88/$3.50

3. RESULTS AND DISCUSSION

RC preparations with A_{280}/A_{806} less than 1.6 were studied by protein chemistry methods. Depending on the RC sample preparation conditions SDS electrophoresis in PAAG shows one band, M_r = 43000 (incubation for 1 h at room temperature) or two bands M_r = 28000 and 32000 (at 2 min boiling). Interestingly the ratio of staining intensity of the two bands changed in every experiment. The N-terminal and amino acid analyses of the material electroeluted from these bands according to [13] showed no difference. Thus either RC consists of two identical subunits separated electrophoretically by some unknown reasons (see [14]), or the separated band contains the same mixture of two different polypeptides. The automated Edman degradation of the entire RC preparation revealed the main amino acid sequence S-R-A-K-A-K-D-P-R-F-P-D-F-S-F-T-V-V that agrees with [15].

Over 50 peptides were isolated by the membrane protein fragmentation method we developed earlier [6]. Complete or partial amino acid sequences were established for 30 peptides.

The N-terminus of the cyanogen bromide peptide F-A-Q-V-N-Y-... was used to synthesize the oligonucleotide probe 1a 3′-TACAAGCG$^{G}_{T}$G TTCA$^{G}_{T}$TTGAT-5′. Here and in other syntheses we utilized the probes with a lower degree of degeneracy: G or T in the third position of codon instead of G/A or T/C, respectively, and G/T instead of G/A/T/C. The probe was used for screening 30000 clones of the first library. Forty-two clones yielded positive signals of diverse intensity. These clones were hybridized with probe 1b 5′-ATGTTTGC$^{G}_{T}$CAGGT$^{G}_{T}$AATTA-3′, that corresponds to the coding sequence of the same cyanogen bromide peptide. Only two clones, pRCa9 and pRCG10, showed positive signals. Clone pRCa9 containing a larger insertion was used for the sequence analysis. It appeared to contain the major part of the gene encoding the L-subunit and 1/3 of the gene coding for the N-terminal sequence of the M-subunit. In order to find the clone with the required N-terminal part, oligonucleotide probe 2 3′-TA$^{G}_{T}$CTTCTTTA$^{G}_{T}$CTTGG-5′ corresponding to the L-subunit sequence 32 to 37 (see fig.2) was synthesized. Of 30000 recombinants of the second library, five clones showed positive signals. Only one of them was hybridized with nick-translated insertion of clone pRCa9. This clone designated pRCII-9 contained DNA fragment encoding the N-terminal part of the L-subunit gene. Fig.1 presents the restriction map of the gene and the sequencing strategy. The amino acid sequence derived from the nucleotide sequence (fig.2) following ATG codon (17–19) preceded by Shine-Dalgarno sequences [16] absolutely coincides with that of the entire RC preparation.

Carboxypeptidase A treatment of RC polypeptides immobilized on CPG-Thiol results in release of valine that accords with the L-subunit C-terminal residue deduced from DNA sequence. Two C-terminal amino acid residues, valine and proline, are identified by hydrazinolysis. C-terminal peptide A-G-I-N-Y-P-Q-G-P-T-P-P-V-(S,L,P) was isolated from the *S. aureus* protease digest of RC. Since the C-terminus of this peptide does not coincide with that of the L-subunit we suggest that proline is the C-terminal residue of the M-subunit. Thus the mature L-subunit consists of 310 amino acid residues (calculated M_r 35014). The apparent formylmethionine is removed by post-translation modification as in the case of L- and M-subunits of purple bacteria RC. Comparison of the primary structures of L-subunits of purple bacteria and *Chloroflexus* RC shows considerable homology (fig.3).

The fairly good coincidence of hydrophobicity profiles of L-subunits (not shown) and conservation of amino acid residues involved in cofactor binding indicates the similar tertiary structures of purple bacteria and *Chloroflexus* RC (fig.4). Additional data supporting this were obtained when comparing membrane parts of the L-subunits. Recently amino acid residues of purple bacteria RC buried inside the protein were shown to be more conservative than those exposed outside the molecule [17]. Comparison of α-helical regions of the L-subunit from *R. sphaeroides* with those of *Chloroflexus* shows that 37% (48/130) of amino acid residues is conservative. Conservation of amino acid residues buried inside the protein is 54% (42/78) and those exposed outside the protein only 17% (5/29). That explains a close similarity in the RC pigment arrangement of *Chloroflexus* [18] and purple bacteria [19,20]. The characteristic feature of the *Chloroflexus* RC L-subunit is the

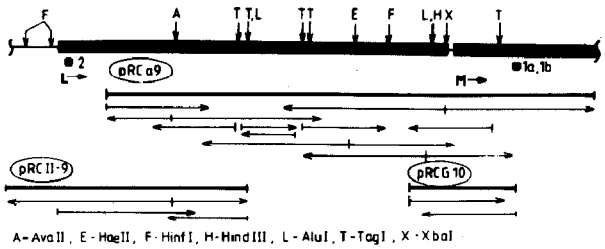

Fig.1. Location of insertions of isolated clones in restriction map of the RC L-subunit gene including part of the M-subunit gene, sequencing strategy. Black squares indicate nucleotide probes used for screening of libraries.

presence of the extended N-terminal sequence. The N-terminal sequences of the purple bacteria RC L- and M-subunits contact the H-subunit. Whether the extended N-terminal sequence of the *Chloroflexus* L-subunit exercises the function of missing H-subunit remains unclear.

The following important substitutions also deserve attention: Gln-143 for Glu-104 which in purple bacteria is near bacteriopheophytin of electron-conducting pathway and probably causes its spectral red shift [21]; Asp-281 for Ala-245

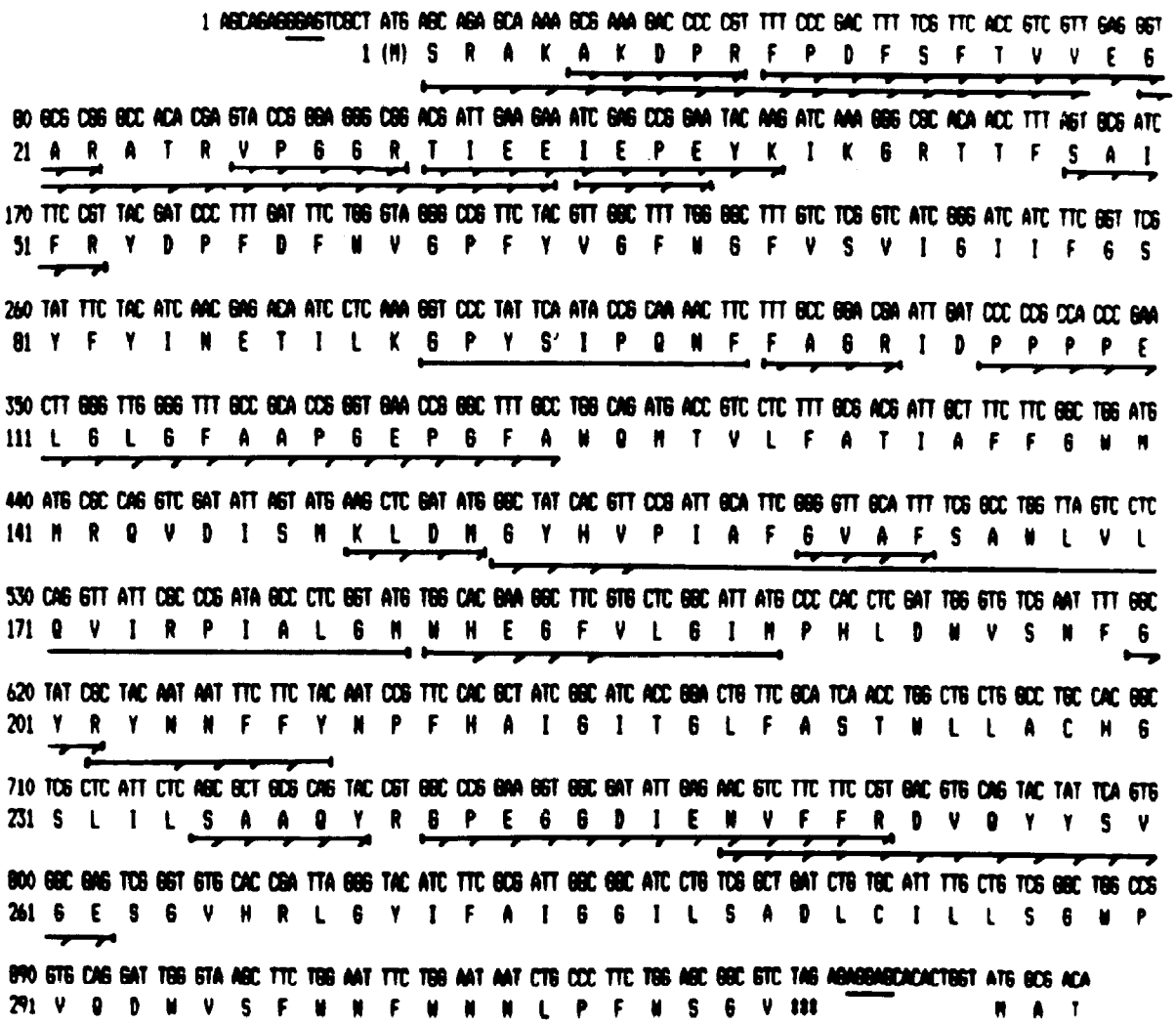

Fig.2. Nucleotide sequence of the gene coding for the L-subunit, the start of the M-subunit, and derived amino acid sequences. Amino acid sequences underlined were confirmed by protein sequencing. Shine-Dalgarno sequences preceding genes underlined [16].

```
                 1        10        20        30        40        50        60
C.AUR.     L SRAKAKDPRFPDFSFTVVEGARATRVPGGRTIEEIEPEYKIKGRTTFSAIFRYDPFDFWV
R.VIR.     L ---------------ALLSFERKYRVRGGTLIGG------------------DLFDFWV
R.SPH.     L ---------------ALLSFERKYRVPGGTLVGG------------------NLFDFWV
R.CAPS.    L ---------------ALLSFERKYRVPGGTLIGG------------------SLFDFWV

                        70        80        90       100       110       120
C.AUR.     L GPFYVGFWGFVSVIGIIFGSYFYINETILKGPYSIPQNFFAGRIDPPPPELGLGFAAPGE
R.VIR.     L GPYFVGFFGVSAIFFIFLGVSLIGY-AASQGPTWDPF---AISINPPDLKYGLGA-APLL
R.SPH.     L GPFYVGFFGVATFFFAALGIILIAWSAVLQG-TWNPQ---LISVYPPALEYGLGG-APLA
R.CAPS.    L GPFYVGFFGVTTIFFATLGFLLILWGAAMQG-TWNPQ---LISIFPPPVENGLNV-AALD

                       130       140       150       160       170       180
C.AUR.     L PGFAWQMTVLFATIAFFGWMMRQVDISMKLDMGYHVPIAFGVAFSAWLVLQVIRPIALGM
R.VIR.     L EGGFWQAITVCALGAFISWMLREVEISRKLGIGWHVPLAFCVPIFMFCVLQVFRPLLLGS
R.SPH.     L KGGLWQIITICATGAFVSWALREVEICRKLGIGYHIPFAFAFAILAYLTLVLFRPVMMGA
R.CAPS.    L KGGLWQVITVCATGAFCSWALREVEICRKLGIGFHIPVAFSMAIFAYLTLVVIRPMMMGS

                       190       200       210       220       230       239
C.AUR.     L WHEGFVLGIMPHLDWVSNFGYRYNNFFYNPFHAIGITGLFASTWLLACHGSLILSAAQY-
R.VIR.     L WGHAFPYGILSHLDWVNNFGYQYLNWHYNPGHMSSVSFLFVNAMALGLHGGLILSVANPQ
R.SPH.     L WGYAFPYGIWTHLDWVSNTGYTYGNFHYNPAHMIAISFFFTNALALALHGALVLSAANPE
R.CAPS.    L WGYAFPYGIWTHLDWVSNTGYTYGNFHYNPFHMLGISLFFTTAWALAMHGALVLSAANPV

                       250       260       270       280       290
C.AUR.     L RGPE--GGDIENVFFRDVQYYSVGESGVHRLGYIFAIGGILSADLCILLSGWPVQD-WVS
R.VIR.     L DGDKVKTAEHENQYFRDVVGYSIGALSIHRLGLFLASNIFLTGAFGTIASGPFWTRGWPE
R.SPH.     L KGKEMRTPDHEDTFFRDLVGYSIGTLGIHRLGLLLSLSAVFFSALCMIITGTIWFDQWVD
R.CAPS.    L KGKTMRTPDHEDTYFRDLMGYSVGTLGIHRLGLLLALNAVFWSACCMLVSGTIYFDLWSD

                  300       310
C.AUR.     L FWNFWNNLPFWSGV*
R.VIR.     L WWGWWLDIPFWS*
R.SPH.     L WWQWWVKLPWWANIPGGING*
R.CAPS.    L WWYWWVNMPFWADMAGGING*
```

Fig.3. Alignment of amino acid sequences of the RC L-subunits from *Chloroflexus* (C. AUR.), *R. viridis* (R. VIR.), *R. sphaeroides* (R. SPH.), and *R. capsulatus* (R. CAPS.). Conserved residues are indicated by −.

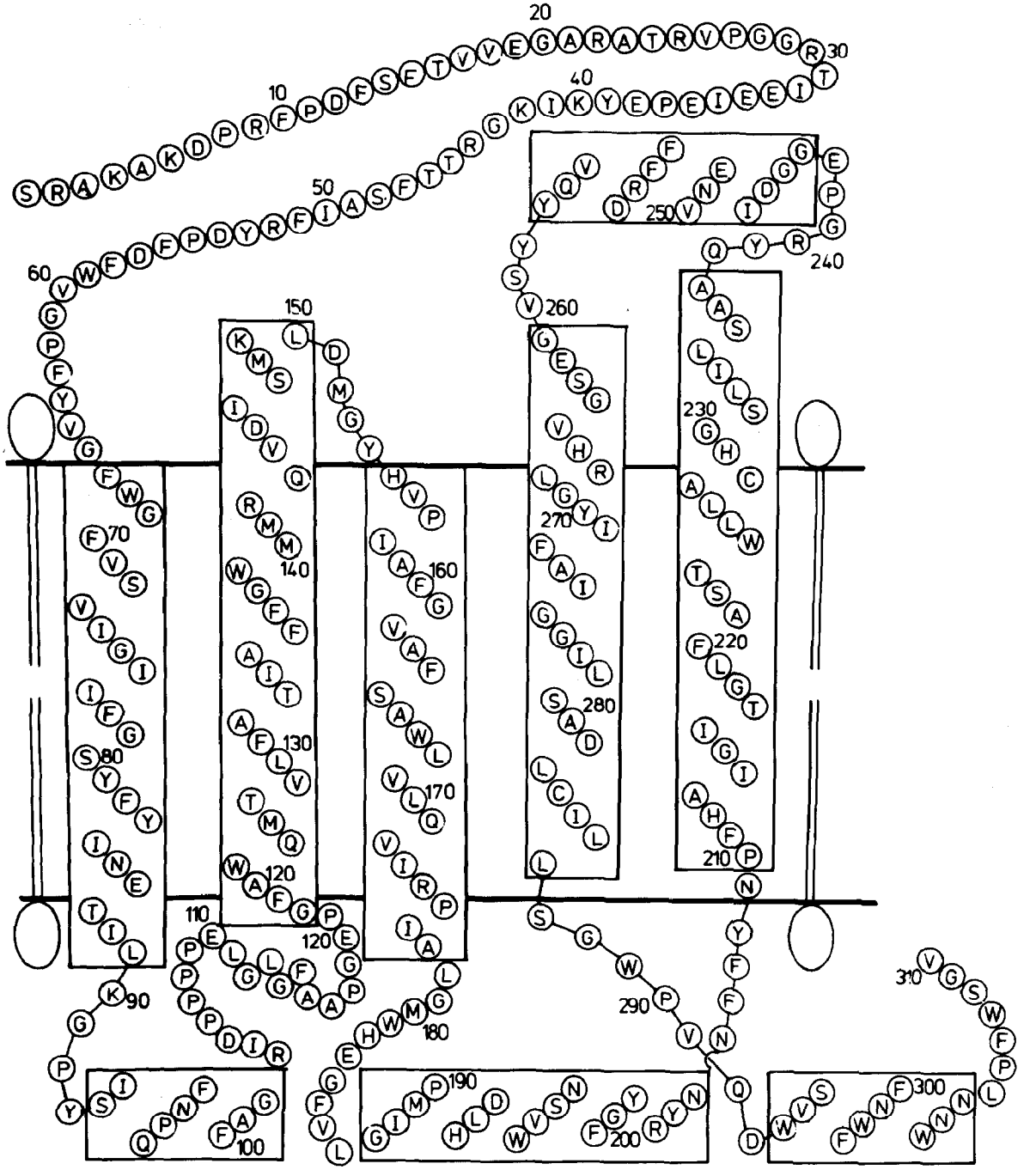

Fig.4. Predicted folding of the amino acid sequence of the RC L-subunit from *Chloroflexus*. Model is based on: (i) homology between the L-subunit of *Chloroflexus* and those of purple bacteria; (ii) similarity of their hydropathy index plots, and (iii) the X-ray structure of *R. viridis* and *R. sphaeroides* reaction centres.

located in the middle of the lipid bilayer. As known oxidation-reduction potential of the special pair of bacteriochlorophylls of *Chloroflexus* (+386 mV) is more negative comparing to those of *R. sphaeroides* (+454 mV) [18] and *R. viridis* (+515 mV) [22]. It can be explained by the

presence of nucleophilic Asp-281 in the vicinity of the special pair which could provide easy electron distraction from the special pair.

As to the second subunit (in preparation) its amino acid sequence is homologous to the M-subunits of purple bacteria.

Acknowledgements: The authors thank N.S. Bystrov for synthesis of oligonucleotide probes, K.A. Ishchenko and A.A. Bucharov for sequencing certain DNA fragments and Dr A.V. Shuvalov for helpful discussions.

REFERENCES

[1] Betti, J.A., Blankenship, R.E., Natarajan, L.V., Dickinson, L.D. and Fuller, R.C. (1982) Biochim. Biophys. Acta 680, 194–201.
[2] Bruce, B.D., Fuller, R.C. and Blankenship, R.E. (1982) Proc. Natl. Acad. Sci. USA 79, 6532–6536.
[3] Pierson, B.K. and Thornber, J.P. (1983) Proc. Natl. Acad. Sci. USA 80, 80–84.
[4] Pierson, B.K., Thornber, J.P. and Seftor, R.E.B. (1983) Biochim. Biophys. Acta 723, 322–326.
[5] Blankenship, R.E., Feick, R., Bruce, B.D., Kirmaier, C., Holten, D. and Fuller, R.C. (1983) J. Cell. Biochem. 22, 251–261.
[6] Ovchinnikov, Yu.A., Abdulaev, N.G., Bogachuk, A.S. and Morris, C.A. (1987) in: Methods in Protein Sequence Analysis, 1986 (Walsh, K.A. ed.) pp.189–209, Humana Press, Clifton, NJ.
[7] Chen, R. (1976) Hoppe-Seyler's Z. Physiol. Chem. 357, 873–886.
[8] Ambler, R.P. (1972) Methods Enzymol. 25, 262–272.
[9] Mesrob, B. and Heleysovsky, V. (1967) Collect. Czech. Chem. Commun. 32, 1976–1982.
[10] Zabarovsky, E.R. and Allikmets, R.L. (1985) Bioorg. Khim. 11, 849–852.
[11] Maxam, A. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560–564.
[12] Chuvpilo, S.A. and Kravchenko, V.V. (1984) FEBS Lett. 179, 34–36.
[13] Hunkapiller, M.W., Lijan, E., Ostrander, F. and Hood, L.E. (1983) Methods Enzymol. 91, 227–236.
[14] Hegemann, P., Blanck, A., Vogelsang-Wenke, H., Lottspeich, F. and Oesterhelt, D. (1987) EMBO J. 6, 259–264.
[15] Shiozawa, J.A., Lottspeich, F. and Feick, R. (1987) Eur. J. Biochem. 167, 595–600.
[16] Shine, J. and Dalgarno, L. (1974) Proc. Natl. Acad. Sci. USA 71, 1342–1346.
[17] Yeates, T.O., Komiya, H., Rees, D.C., Allen, J.P. and Feher, G. (1987) Proc. Natl. Acad. Sci. USA 84, 6438–6442.
[18] Shuvalov, V.A., Shkuropatov, Yu., Kulakova, S.M., Ismailov, M.A. and Shkuropatova, S.M. (1986) Biochim. Biophys. Acta 849, 337–346.
[19] Deisenhofer, J., Epp, O., Miki, K., Huber, R. and Michel, H. (1984) J. Mol. Biol. 180, 385–398.
[20] Allen, J.P., Feher, G., Yeates, T.O., Komiya, H. and Rees, D.C. (1987) Proc. Natl. Acad. Sci. USA 84, 5730–5734.
[21] Michel, H., Epp, O. and Deisenhofer, J. (1986) EMBO J. 5, 2445–2451.
[22] Klimov, V.V., Shuvalov, V.A., Krakhmaleva, I.N., Klevanik, A.V. and Krasnovsky, A.A. (1977) Biokhimiya 42, 519–530.