Vol. 185, No. 1, 1992 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
May 29, 1992 Pages 341-349

DESIGN AND FUNCTIONAL EXPRESSION IN *Escherichia coli* OF A SYNTHETIC GENE ENCODING *Clostridium pasteurianum* 2[4Fe-4S] FERREDOXIN*

Valérie Davasse and Jean-Marc Moulis

Département de Biologie Moléculaire et Structurale, Laboratoire Métalloprotéines, C.E.N.G. 85X, 38041 Grenoble Cedex, France

Received April 14, 1992

ABSTRACT : A gene encoding the exact sequence of *Clostridium pasteurianum* 2[4Fe-4S] ferredoxin and containing 11 unique restriction endonuclease cleavage sites has been synthesized and cloned in *Escherichia coli*. The synthetic gene is efficiently expressed in *E.coli* and its product has been purified and characterized. The N-terminal sequence is identical to that of the protein isolated from *C.pasteurianum* and the recombinant ferredoxin contains the exact amount of [4Fe-4S] clusters (2 per monomer) expected for homogeneous holoferredoxin. It displays reduction potential and kinetic parameters as electron donor to *C.pasteurianum* hydrogenase I identical to those determined for the native ferredoxin. All of these properties demonstrate that the 2[4Fe-4S] ferredoxin expressed in *E.coli* is identical to the parent clostridial protein.

The 2[4Fe-4S] ferredoxin from *Clostridium pasteurianum* has been one of the very first iron-sulfur proteins to be purified (1). Since then, a wealth of structural and functional information has been gathered on this small (55 aminoacids) protein or on other homologous clostridial ferredoxins (2). The main reactions in which *C.pasteurianum* ferredoxin participates *in vivo* have been listed (3) and the physiological importance of this protein is illustrated by its prominent role as an electronic shuttle between the catabolic and anabolic pathways of *C.pasteurianum* (3) and of most Clostridia. More recently, the gene encoding *C.pasteurianum* ferredoxin has been cloned and sequenced (4) and its transcription studied (5). Despite all of these data, some questions about the detailed mechanism by which clostridial ferredoxins fulfill their function remain unanswered. These problems would largely benefit from the availability of molecular variants.

*Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession No. M83832.

0006-291X/92 $4.00

A few ferredoxin genes have been cloned and expressed in *Escherichia coli* (6-11). Most of them encode relatively long (more than *ca.* 80 aminoacids) proteins of the [2Fe-2S] type and, among those encoding short [4Fe-4S] ferredoxins, no report on site-directed mutants of this class of ferredoxin has yet appeared. A complementary approach can be carried out with these small proteins, where a synthetic gene may avantageously replace the natural gene. It is shown here that *C.pasteurianum* 2[4Fe-4S] ferredoxin can be expressed in *E.coli* from a synthetic gene and the full characterization of the recombinant protein indicates it to be indistinguishable from the parent protein purified from *C.pasteurianum*.

MATERIALS AND METHODS

All common methods used to manipulate DNA molecules were as described (12).

STRAINS AND MATERIALS. Restriction enzymes, T4 DNA ligase and Isopropyl-1-thio-β-D-galactoside (IPTG) were purchased from Boehringer Mannheim, plasmids pUC18 and pKK223-3 as well as *E.coli* strain JM109 were obtained from Pharmacia, T4 polynucleotide kinase and competent *E.coli* DH5α were from Gibco-BRL. *E.coli* strain SG20043 (13) was obtained from Dr. P. Brodin, University of Umeå, Sweden.

CONSTRUCTION OF THE GENE. Oligonucleotides were synthesized by phosphoramidite chemistry on a 381A synthesizer from Applied Biosystems. They were obtained with the protecting dimethoxytrityl group at the 5' end and purified on an Aquapore RP-300 column (Brownlee Labs). After deprotection, the purity was assessed by Urea-PAGE and, when necessary, the deprotected oligonucleotides were further purified by another passage on the same column. All oligonucleotides constituting the gene, except the two protruding ones at the 5' ends, were phosphorylated for 1 h at 37°C and heated to 95°C. The two remaining oligonucleotides were then added and the mixture was annealed by decreasing the temperature to 30°C over a 2 h-period. Overnight ligation with T4 DNA ligase at 15°C afforded a series of DNA fragments which were resolved on 8% PAGE-SDS in Tris-borate-EDTA buffer. The band migrating at the expected apparent size was sliced and extracted with ammonium acetate 0.5 M, pH 8.0, EDTA 1 mM. This material was cloned into *Eco*RI-*Hin*dIII-cleaved pUC18 (14) and its DNA sequence was determined by the dideoxy method (15).

PURIFICATION OF RECOMBINANT FERREDOXIN. *E.coli* JM109/pKCFd2 (see below) was grown on Luria Broth supplemented with 100 µl/ml ampicillin and induced at late log phase with 1 mM IPTG for 4 hours. The cell pellet was obtained by centrifugation of the culture for 20 min (4°C, 4500 ***g***), resuspended in 100 mL Tris-Cl 20 mM, pH 7.4 containing 0.5 mM phenyl-methyl-sulfonylfluoride (buffer A) and stored at -80°C. The thawed suspension was sonicated and centrifuged for 30 min (4°C, 45000 ***g***). All further purification steps were carried out under strictly anaerobic conditions. The soluble extract was loaded on a DEAE-cellulose column (DE-52, Whatman, U.K.) equilibrated with buffer A. After washing the column with 2 volumes each of NaCl 0.15 and 0.25 M in buffer A, the brown ferredoxin fraction was eluted with the same buffer containing 0.4 M NaCl. The protein solution was treated with 10 µg/ml each of deoxyribonuclease I (Boehringer Mannheim) and ribonuclease A (Sigma), both

from bovine pancreas, at 37°C for 1 h, and filtrated through Sephadex G-50 (Pharmacia) equilibrated with buffer A. The ferredoxin-containing fraction was easily followed by its color and concentrated after elution on a Amicon YM5 membrane. The recombinant ferredoxin was further purified by passage through an anion exchange h.p.l.c. column (PL-SAX from Polymer Laboratories Ltd., U.K.) equilibrated with Tris-Cl 20 mM, pH 8.0, NaCl 0.1 M and developed with a gradient of increasing NaCl concentration. Pure ferredoxin (see Results) was eluted at a NaCl concentration of 0.4 M.

OTHER METHODS. Protein sequence analysis was carried out on a model 477A sequencer coupled to on-line detection model 120A (Applied Biosystems). Redox potential measurements were performed in Tris-Cl 20 mM, NaCl 40 mM pH 8.0 using the method of mixtures with methyl viologen (16). A value of -449 mV *vs* the Normal Hydrogen Electrode (NHE) for the mid-point potential of methyl viologen was assumed (16). Native *C.pasteurianum* ferredoxin and hydrogenase I (EC 1.18.3.1) were purified and assayed in Tris-Cl 0.1 M, pH 7.0, NaCl 0.1 M as previously described (17).

RESULTS AND DISCUSSION

Due to the design of the synthetic *C.pasteurianum* Fd gene (Figure 1), all oligonucleotides building the complete sequence could be assembled in a single reaction. Characterization by restriction endonuclease analysis and sequencing of the insert in *Eco*RI-*Hin*dIII-digested pUC18 (plasmid pUCFd2) have indicated that the construct displayed the exact sequence shown in Figure 1. 11 restriction endonucleases recognition sequences of six or more base pairs have been distributed along the gene to ease its handling and future modifications. An ATG initiation codon has been included to provide a translation start signal. Codons found in highly expressed *E.coli* genes (18) have generally been introduced, with two main exceptions, ATA for I23 and ACA for T36, which serve to build the recognition sequences for *Eco*RV and *Nsi*I, respectively.

The synthetic gene cloned in pKK223-3 (19) gave plasmid pKCFd2, which produced in *E.coli* JM109 significantly larger amounts of ferredoxin than pUCFd2. The main characteristics of the protein purified as described in the Materials and Methods section from this new source are described below.

The electronic spectra of Figure 2 display the broad absorption bands centered at 388 and 302 nm expected from $[4Fe\text{-}4S]^{2+}$ clusters (20, 21). The contribution of the protein around 280 nm in these spectra is small and the measured A_{388}/A_{280} ratio of 0.83 is the best available indicator of the purity of the holoferredoxin (17), both for the protein purified from *C.pasteurianum* or the one obtained from *E.coli* JM109/pKCFd2. These data indicate a significant improvement in the purity of the preparation, or at least in its holoferredoxin content, over other strategies used to produce the same protein outside its natural host (10, 22, 23). The EPR spectra of reduced clostridial ferredoxins are highly characteristic since they arise from the magnetic interaction between two

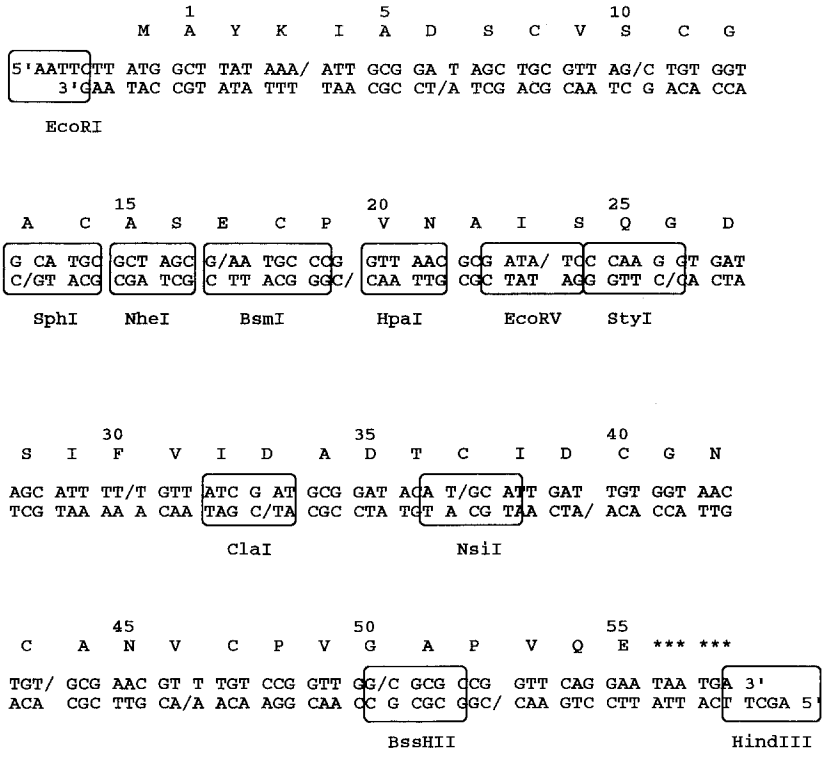

FIGURE 1. Design and sequence of the synthetic gene encoding *C.pasteurianum* ferredoxin. The sequences of the coding and complementary DNA strands resulting from the concatenation of the synthetic fragments are shown. The corresponding protein sequence is indicated above. Unique restriction endonucleases recognition sequences are boxed and slashes separate the oligonucleotides used for gene assembly. The oligonucleotides used to assemble both strands of the synthetic gene encoding *C.pasteurianum* Fd were designed mainly to enable modular mutagenesis of the gene. The aminoacid sequence of *C.pasteurianum* Fd (4, 34) was back-translated and screened for potential restriction endonucleases recognition sequences of six or more nucleotides. Outside the selected restriction sites, the codon usage deduced from the subdatabase of highly expressed *E.coli* genes (18) was applied. The synthetic gene was divided into oligonucleotides of reasonable length for efficient purification (around 20 nucleotides), while managing to have each oligonucleotide on one strand overlapping with at least 7 nucleotides of the hybridizing oligonucleotides of the complementary strand. These overlapping sequences were checked not to resemble each other in order to avoid misassembly of the gene. The other guidelines followed to design the synthetic gene were, when applicable, those detailed previously (35).

S=1/2 spin systems, each contributed by one cluster, *ca.* 1 nm apart (24). The spectrum of the ferredoxin expressed in *E.coli* is shown in Figure 3 and clearly demonstrates that the two clusters are properly assembled and oriented.

Aminoacid sequencing of the carboxymethylated ferredoxin expressed in *E.coli* gives the expected N-terminal sequence AYKIA, without significant amounts of other aminoacids. The sequencing yields of the N-terminal alanine

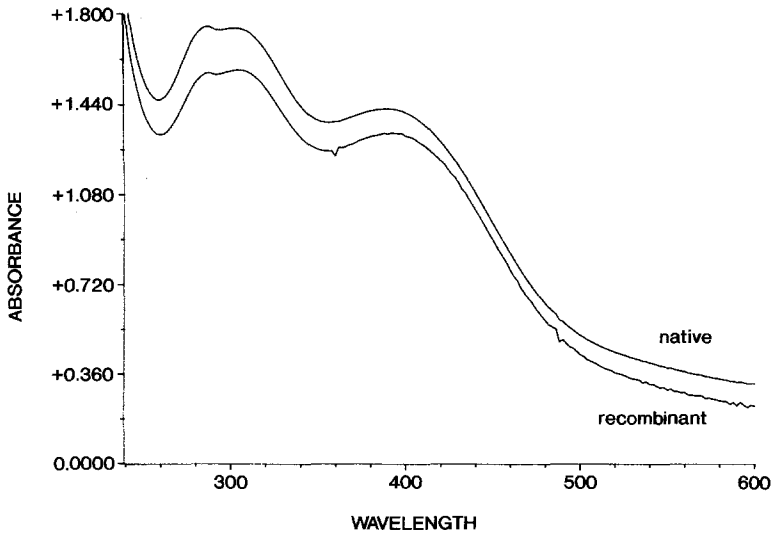

FIGURE 2. Electronic absorption spectra of *C.pasteurianum* ferredoxin. The native and recombinant proteins were in Tris-Cl 20 mM, NaCl 0.4 M, pH 8.0. For clarity, an offset of +0.1 absorbance units has been applied to the spectrum of native ferredoxin. The spectra were recorded with a Hewlett-Packard 8452 spectrophotometer.

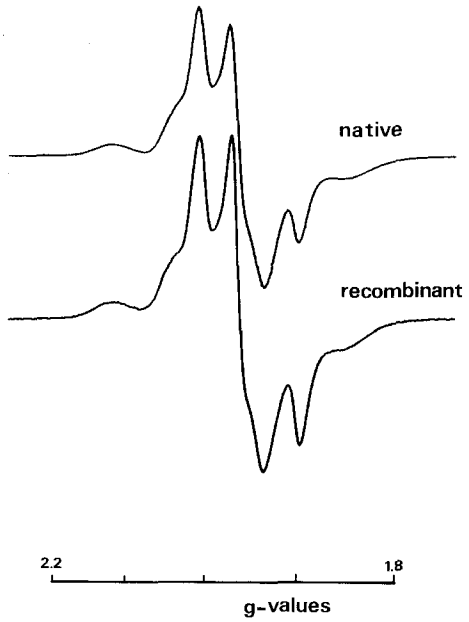

FIGURE 3. X-band EPR spectra of *C.pasteurianum* ferredoxin. The spectra were recorded with a X-band Varian E-109 spectrometer equipped with a liquid helium transfer system (Oxford Instruments ESR 900) at a klystron frequency of 9.225 GHz and a temperature of 10 K with a microwave power of 0.2 mW.

measured on samples of ferredoxin purified either from *C.pasteurianum* or from *E.coli* JM109/pKCFd2 are similar, between 80 and 90% and they do not increase upon treatment with cyanogen bromide. Therefore, the N-terminal methionine encoded by the start codon of the synthetic gene is cleaved by *E.coli* in much the same way as it is from the native protein in *C.pasteurianum*. The removal of the initial methionine from plasmid-encoded proteins in *E.coli* appears to strongly depend on the nature of the second translated amino acid (25), the cleavage efficiency being inversely correlated to the radius of gyration of the latter residue (26). The present finding that the N-terminus of *C.pasteurianum* ferredoxin expressed in *E.coli* is exclusively alanine agrees with such trends.

The reduction potentials of samples of 2[4Fe-4S] ferredoxin isolated from *C.pasteurianum* and from *E.coli* transformed with pKCFd2 have been found to be -418 mV and -415 mV *vs* NHE, respectively. Considering the intrinsic uncertainties associated with these measurements, it can safely be concluded that the ferredoxins from both origins display the same reduction potential. Among the various physiological reactions in which 2[4Fe-4S] clostridial ferredoxins serve as electron transfer agents (3), *C.pasteurianum* ferredoxin can supply electrons to hydrogenase I of the same organism. This reaction has been studied *in vitro* and, as can be seen from Figure 4, the activity of the ferredoxin expressed from pKCFd2 is indistinguishable from that of the same protein isolated from *C.pasteurianum*. In the conditions of the experiment (Figure 4), the kinetic constants are K_m=28 μM and k_{cat}=200 s^{-1}.

The present report describes the first successful production of a fully active recombinant iron-sulfur protein from a synthetic gene. *C.pasteurianum* ferredoxin is unusual as it contains 8 cysteines out of 55 aminoacids; it has been documented that a high cysteine content in a strongly expressed protein may impede either its solubility or its correct folding (27). The synchronous assembly of the inorganic core with the folding of the nascent polypeptide is most probably responsible for the shielding of the cysteine residues and for the synthesis of stable and soluble holoferredoxin molecules. The same reason might explain the previously reported failure to satisfactorily express a ferredoxin similar to that of *C.pasteurianum* in *E.coli* (28); in that case, the addition of a signal sequence aimed at excreting the product has most probably hampered the insertion of the clusters during or right after translation and has led to untractable forms of the molecule. The accumulation of *C. pasteurianum* ferredoxin in *E.coli* cannot be considered as a trivial task, since small heterologous polypeptides are generally easily recognized in the enterobacterium and rapidly degraded through a cascade of proteolytic reactions (29). The ATP-dependant protease La encoded by *lon* (30) catalyzes one of the first steps in this process. We have observed here that the steady-

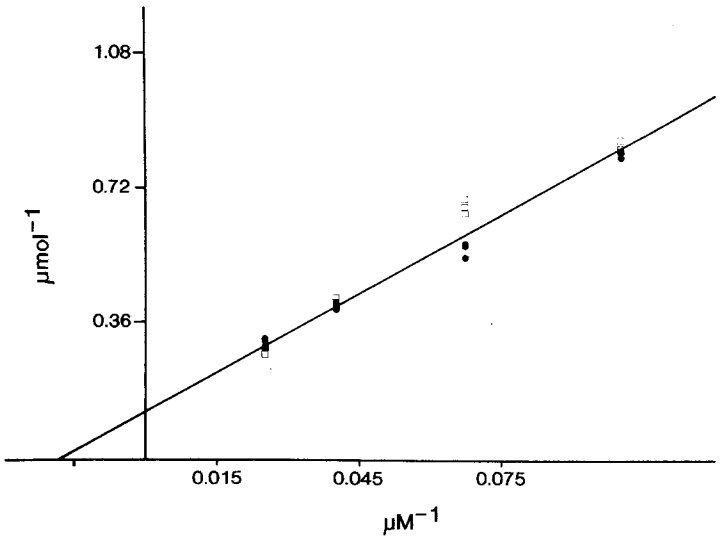

FIGURE 4. Double-reciprocal plot of the activity of *C.pasteurianum* ferredoxin. The hydrogen production from protons catalyzed by 5 µg of *C.pasteurianum* hydrogenase I was measured for 5 min as a function of the ferredoxin concentration in Tris-Cl 0.1 M, NaCl 0.1 M, pH 7.0. Open symbols, ferredoxin isolated from *C.pasteurianum*; Closed symbols, ferredoxin isolated from *Escherichia coli* JM109/pKCFd2. The least-square fit of the data is drawn and gives the values reported in the text.

state concentration of *C.pasteurianum* ferredoxin does not increase in the *E.coli lon*⁻ mutant SG20043 (13). This result minimizes the importance of degradation by this pathway in the present case; either protease La takes no part in the rate-limiting steps of heterologous ferredoxin degradation or, more plausibly, the newly synthesized ferredoxin is resistant to proteolytic degradation due to the rapid incorporation of the Fe-S centers. Also, considering the expression of synthetic genes encoding iron-containing proteins (31, 32, 33, this work), it is worth noticing that no special element on the 5' or 3' untranslated part of these genes appears to be required to trigger or adjust the incorporation of iron into the nascent polypeptide.

The correct folding and efficient incorporation of the [4Fe-4S] clusters of *C.pasteurianum* ferredoxin in *E.coli* are borne out by the observed properties of the recombinant protein. An extensive set of criteria has been used here to demonstrate that *C.pasteurianum* ferredoxin expressed in *E.coli* is identical to its native counterpart. Beside extending the dataset available for metalloproteins expressed in *E.coli*, the present work provides a valuable addition to the studies aiming at producing molecular variants of ferredoxins, especially those containing [4Fe-4S] clusters. The preparation of an almost unlimited range of modified forms can now be considered due to the presence of many restriction sites along the synthetic version of the gene. This approach

constitutes an alternative to strategies using *de novo* chemical synthesis of the polypeptide chain (22,23) in providing a more versatile way of getting molecular variants.

ACKNOWLEDGMENTS

This work was supported in part by funds from the Laboratoire d'Ingénierie des Protéines, Commissariat à l'Energie Atomique. We are grateful to P.Richaud for oligonucleotide synthesis, J.Gaillard for EPR measurements and J.Gagnon and N.Scherrer for aminoacid sequencing. We thank Dr Peter Brodin, University of Umeå, for advice on the use of strain SG20043. The comments of I.Mathieu on the manuscript and the careful reading of the manuscript, the suggestions and the continuous support of J.Meyer are gratefully acknowledged.

REFERENCES

1. Mortenson, L. E., Valentine, R. C. and Carnahan, J. E. (1962) Biochem. Biophys. Res. Commun. **7**, 448-452.
2. Sweeney, W. V. and Rabinowitz, J. C. (1980) Ann. Rev. Biochem. **49**, 139-161.
3. Mortenson, L. E. and Nakos, G. (1973) in Iron-Sulfur Proteins (W. Lovenberg, Ed) vol.I, pp. 37-64, Academic Press, New York.
4. Graves, M. C., Mullenbach, G. T. and Rabinowitz, J. C. (1985) Proc. Natl. Acad. Sci. USA **82**, 1653-1657.
5. Graves, M. C. and Rabinowitz, J. C. (1986) J. Biol. Chem. **261**, 11409-11415.
6. Coughlan, V. M. and Vickery, L. E. (1989) Proc. Natl. Acad. Sci. USA **86**, 835-839.
7. Böhme, H. and Haselkorn, R. (1989) Plant Molec. Biol. **12**, 667-672.
8. Koga, H., Yamaguchi, E., Matsunaga, K., Aramaki, H. and Horiuchi, T. (1989) J. Biochem. **106**, 831-836.
9. Peterson, J. A., Lorence, M. C. and Amarneh, B. (1990) J. Biol. Chem. **265**, 6066-6073.
10. Baur, J. R., Graves, M. C., Feinberg, B. A. and Ragsdale, S. W. (1990) BioFactors **2**, 197-203.
11. Grabau, C., Schatt, E., Jouanneau, Y. and Vignais, P. M. (1991) J. Biol. Chem. **266**, 3294-3299.
12. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (1988) Current Protocols in Molecular Biology, Wiley-Interscience, New York.
13. Trisler, P. and Gottesman, S. (1984) J. Bacteriol. **160**, 184-191.
14. Yanisch-Perron, C., Viera, J. and Messing, J. (1985) Gene **33**, 103-119.
15. Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA **74**, 5463-5467.
16. Wilson, G. S. (1978) Methods Enzymol. **54**, 396-410.
17. Moulis, J.-M. and Meyer, J. (1982) Biochemistry **21**, 4762-4771.
18. Grosjean, H. and Fiers, W. (1982) Gene **18**, 199-209.
19. Brosius, J. and Holy, A. (1984) Proc. Natl. Acad. Sci. USA **81**, 6929-6933.
20. Mortenson, L. E. (1964) Biochim. Biophys. Acta **81**, 71-77.
21. Meyer, J. and Moulis, J.-M. (1981) Biochem. Biophys. Res. Commun. **119**, 828-835.

22. Smith, E. T., Feinberg, B. A., Richards, J. H. and Tomich, J. M. (1991) J. Am. Chem. Soc. **113**, 688-689.
23. Smith, E. T., Tomich, J. M., Iwamoto, T., Richards, J. H., Mao, Y. and Feinberg, B. A. (1991) Biochemistry **30**, 11669-11676.
24. Mathews, R., Charlton, S., Sands, R. H. and Palmer, G. (1974) J. Biol. Chem. **249**, 4326-4328.
25. Hirel, P.-H., Schmitter, J.-M., Dessen, P., Fayat, G. and Blanquet, S. (1989) Proc. Natl. Acad. Sci. USA **86**, 8247-8251.
26. Dalbøge, H., Bayne, S. and Pedersen, J. (1990) FEBS Lett. **266**, 1-3.
27. Schoemaker, J. M., Brasnett, A. H. and Marston, F. A. O. (1985) EMBO J. **4**, 775-780.
28. Bourdineaud, J. P., Howard, S. P., Pages, J. M., Bernadac, A., Leroy, G., Bruschi, M. and Lazdunski, C. (1990) Biochimie **72**, 407-415.
29. Gottesman, S. (1990) Methods Enzymol. **185**, 119-128.
30. Goldberg, A. L., Swamy, K. H. S., Chung, C. H. and Larimore, F. S. (1981) Methods Enzymol. **80**, 680-702.
31. Alexander, H., Alexander, S., Heffron, F., Fieser, T. M., Hay, B. N., Getzoff, E. D., Tainer, J. A. and Lerner, R. A. (1991) Gene **99**, 151-156.
32. Springer, B. A. and Sligar, S. G. (1987) Proc. Natl. Acad. Sci. USA **84**, 8961-8965.
33. Beck von Bodman, S., Schuler, M. A., Jollie, D. R. and Sligar, S. G. (1986) Proc. Natl. Acad. Sci. USA **83**, 9443-9447.
34. Tanaka, M., Nakashima, T., Benson, A., Mower, H. and Yasunobu, K. T. (1966) Biochemistry **5**, 1666-1681.
35. Ferretti, L., Karnik, S. S., Khorana, H. G., Nassal, M. and Oprian, D. D. (1986) Proc. Natl. Acad. Sci. USA **83**, 599-603.