Volume 14 Number 10 1986 Nucleic Acids Research

Complete nucleotide and derived amino acid sequence of cDNA encoding the mitochondrial uncoupling protein of rat brown adipose tissue: lack of a mitochondrial targeting presequence

Robert G.Ridley, Hasmukh V.Patel, Gerhard E.Gerber, Robert C.Morton and Karl B.Freeman*

Department of Biochemistry, McMaster University, 1200 Main St. West, Hamilton, Ontario, Canada L8N 3Z5

Received 14 March 1986; Accepted 24 April 1986

ABSTRACT

A cDNA clone spanning the entire amino acid sequence of the nuclear-encoded uncoupling protein of rat brown adipose tissue mitochondria has been isolated and sequenced. With the exception of the N-terminal methionine the deduced N-terminus of the newly synthesized uncoupling protein is identical to the N-terminal 30 amino acids of the native uncoupling protein as determined by protein sequencing. This proves that the protein contains no N-terminal mitochondrial targeting prepiece and that a targeting region must reside within the amino acid sequence of the mature protein.

INTRODUCTION

Uncoupling protein (UCP) is an integral inner membrane protein unique to brown adipose tissue mitochondria and is thought to be instrumental in the thermogenesis of this tissue (1-3). The protein acts as a dimer in the mitochondrial inner membrane, forming a proton channel (1, 2, 4, 5) which can lead to uncoupling of oxidative phosphorylation by dissipation of the electrochemical potential across the inner membrane. Purine nucleotides are known to bind to the uncoupling protein and inhibit proton permeability (1-5), while it is possible that free fatty acids can reverse this inhibition (2, 4, 6). Recently, the hamster uncoupling protein has been sequenced and a model for its disposition in the mitochondrial inner membrane has been proposed (7).

Our interest in the uncoupling protein has been primarily within the area of mitochondrial biogenesis. Most mitochondrial proteins are nuclear-encoded and synthesized on free polysomes in the cytosol with N-terminal extensions or prepieces (8). These prepieces serve to direct the newly synthesized protein to

© IRL Press Limited, Oxford, England.

mitochondria after which they are proteolytically cleaved by a matrix protease (9-11). The uncoupling protein does not appear to have an N-terminal prepiece as the protein synthesized in a reticulocyte lysate, primed with brown adipose tissue polysomes or mRNA, displays a mobility on sodium dodecyl sulphate-polyacrylamide gels identical to that of the mature protein (12-14). Categorical proof of this however requires sequence data for the N-terminus of both the mature protein and the newly synthesized protein.

In this report we present the isolation and sequencing of a cDNA clone spanning the entire coding region of the rat uncoupling protein mRNA. The deduced N-terminus of the newly synthesized protein, with the exception of the N-terminal methionine, is identical to the N-terminus of the mature protein as determined by protein sequencing. Since newly synthesized uncoupling protein is imported by isolated mitochondria (12) a targeting region must exist within the amino acid sequence of the mature protein.

MATERIALS AND METHODS

Isolation of cDNA clones

Construction of a rat brown adipose tissue cDNA library, enriched for uncoupling protein, within the PstI site of pBR322 has been described previously (15). A total of 1500 colonies were arrayed on nitrocellulose filters and screened by the Grunstein and Hogness procedure (16). The PstI insert of $pUCP_{rat}2$, previously isolated from this cDNA library (15) and known to correspond to the 3'-end of rat UCP mRNA was used as a probe. The insert was first isolated by electroelution from agarose gels (17) and then radiolabelled by an oligolabelling technique (18). Prehybridizations and hybridizations were carried out in a solution containing 50% (v/v) formamide, 5 x Denhardt's solution, 5 x SSPE, 0.1% (w/v) SDS and 100 µg denatured salmon sperm DNA/ml. The filters were washed in 2 x SSC, 0.1% SDS (room temperature, 4 x 15 min) followed by 1 x SSC, 0.1% SDS (68°C, 2 x 1 h). Promising clones were selected for re-screening and small scale plasmid preparations obtained (19). Plasmid samples were denatured in 0.5 M NaOH (10 µl, room

temperature, 20 min) followed by addition of 1 M NaCl, 50 mM EDTA, 50 mM Tris-HCl pH 8.0 (15 µl) and then 1 M Tris-HCl pH 6.8 (25 µl) and H_2O (50 µl). Aliquots (10 µl) were then spotted on nitrocellulose and probed as described above.

Restriction analyses and determination of DNA sequences

Restriction enzyme analysis of plasmid DNA (19) was performed under standard conditions. Suitable fragments of $pUCP_{rat}15$ were subcloned into M13mp18 and sequenced by the dideoxy chain termination method (20). Either the 17 base universal M13 sequencing primer or a synthetic oligonucleotide derived from the cDNA sequence was used for sequence determination.

Determination of protein sequences

Rat uncoupling protein was purified from rat brown adipose tissue in a Triton X-100 micellar form (21). Only one band at 32,000 Mr was observed on sodium dodecyl sulphate-polyacrylamide gels. Triton X-100 was removed from the protein by chromatography on Sephadex LH-60 in 88% formic acid/ethanol (30:70, v/v). The fractions eluting at the void volume were dried under reduced pressure and redissolved in 88% formic acid. The protein was then subjected to automated Edman degradation using a Beckman 890C sequencer and the standard Beckman 0.1 M Quadrol programme with Beckman reagents. Phenylthiohydantoin amino acids were identified by reverse phase HPLC using a Waters Picotag column.

Computer Analysis

Comparisons of protein sequences and data bank searches were performed using the Beckman MICROGENIE software package (22).

RESULTS

Isolation of UCP clones

The library screened for UCP clones had previously been prepared from a poly(A)$^+$ RNA fraction enriched for UCP message by sucrose density gradient fractionation (15). A total of 1500 colonies were screened initially by the procedure of Grunstein and Hogness (16) and possible positive clones re-screened as described in the Methods section. Twelve clones, $pUCP_{rat}6$-17 were identified by this procedure; the identification of ten of

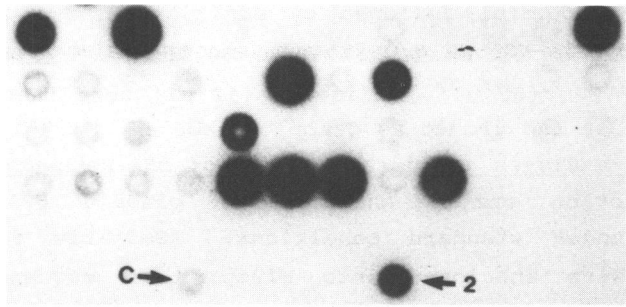

Fig. 1 Detection of clones encoding uncoupling protein. Forty five clones were selected from an initial screening and re-screened as described in the Methods section. The spot labelled '2' is a positive control consisting of a $pUCP_{rat2}$ plasmid preparation. Spot 'c' represents a negative control, pBR322.

these clones, $pUCP_{rat8-17}$, from 45 possible positives, is shown in Fig. 1. The 12 clones, $pUCP_{rat6-17}$, were all specifically shown to contain cDNA inserts corresponding to UCP mRNA by using the radiolabelled PstI insert of $pUCP_{rat2}$ (15) to probe a Southern blot (23) of PstI digests of $pUCP_{rat6-17}$ inserts. The presence of a unique RsaI site within the $pUCP_{rat2}$ insert (15) also allowed the determination of the orientation of clones $pUCP_{rat6-17}$ by a Southern hybridization of their RsaI digests using the same probe. Three clones, $pUCP_{rat14-16}$, all contained inserts of approximately 1200 bp which we calculated to be large enough to encode the entire uncoupling protein sequence. These clones appeared identical by restriction enzyme analysis. Interestingly all of these clones, as well as the only other clone with an insert greater than 1 kb, were oriented so that no formation of a ß-lactamase hybrid protein could occur, in contrast to $pUCP_{rat1-5}$ (15).

Determination of UCP cDNA and protein sequence

One clone, $pUCP_{rat15}$, was selected and sequenced as described in the Methods section. The sequencing strategy is outlined in Fig. 2. The nucleotide sequence of $pUCP_{rat15}$ and the deduced amino acid sequence of rat UCP is given in Fig. 3. This alignment allows only one open reading frame due to the presence of several 'stop' codons upstream of the ATG initiating codon. Protein sequencing of the N-terminal 30 amino acids of purified

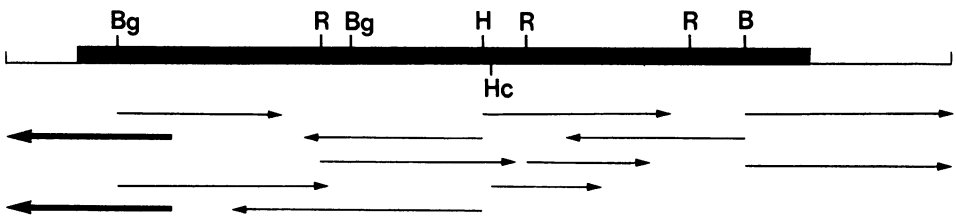

Fig. 2 Partial restriction map and sequencing strategy for clone $pUCP_{rat}15$. Sequencing was carried out as described in the Methods section. Only those restriction sites used for sequencing are presented: B, BamHI; Bg, BglII; H, HindIII; Hc, HincII and R, RsaI. The map does not contain poly(dG.dC) tails but is otherwise drawn to scale. The thick line corresponds to the coding region. Thin arrows indicate sequences determined using the universal primer. The thick arrows indicate a sequence determined using a synthetic oligonucleotide complementary to a subcloned insert as a primer.

rat UCP was performed as described in the Methods section. The results, with the exception of the N-terminal methionine, were in perfect agreement with the amino acid sequence predicted from the cDNA sequence (Fig. 3).

DISCUSSION

Lack of N-terminal prepiece

Two species of rat UCP mRNA have been detected (15, 24), which are probably due to the presence of two polyadenylation sites within the pre-mRNA (25, 26). We have determined the entire coding sequence, the entire 3'-untranslated sequence and part of the 5'-untranslated sequence of one of these species of rat UCP mRNA (Fig. 3).

As mentioned in the Introduction, most mitochondrial proteins are made as precursors with N-terminal presequences (8). The data presented here proves that, with the exception of the N-terminal methionine, newly synthesized UCP has no N-terminal prepiece. The proposed N-terminal 30 amino acids deduced from the cDNA sequence of $pUCP_{rat}15$ were in complete agreement with the N-terminal 30 amino acids of the mature protein determined by protein sequencing (Fig. 3). The proposed initiating ATG codon is the only possible choice from the data obtained. No ATG codons (in or out of frame) are observed upstream of the proposed

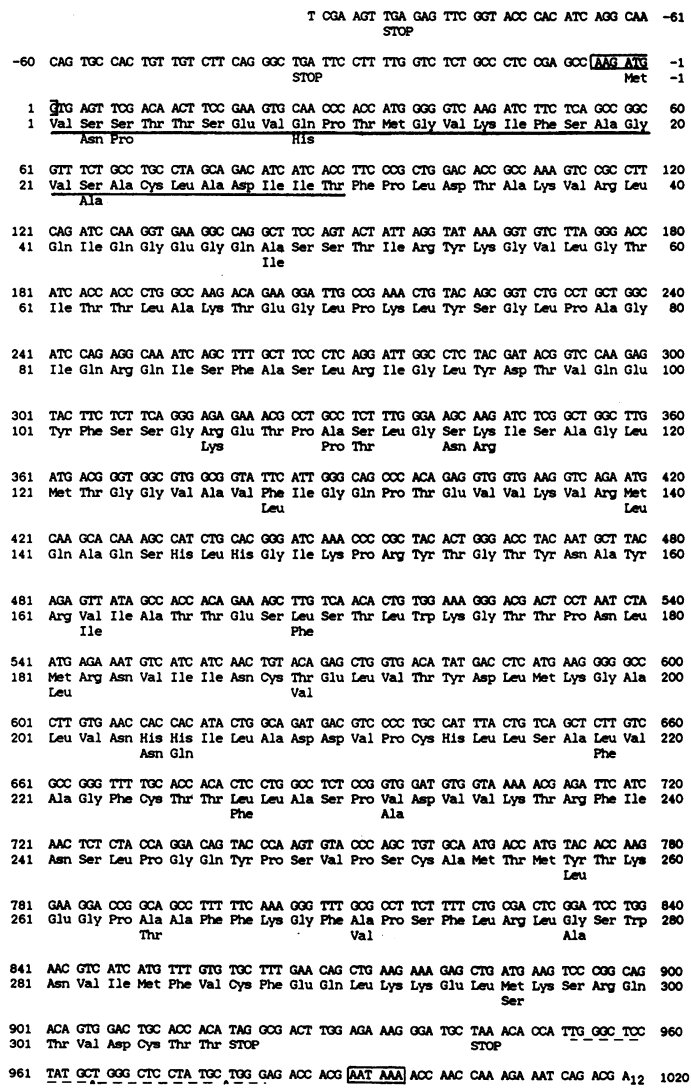

Fig. 3 Complete nucleotide sequence of the cDNA insert of pUCP$_{rat15}$ (row 1) with the derived amino acid sequence of rat UCP (row 2) and a comparison with the sequence for hamster UCP (7) (row 3). Only amino acids of the hamster sequence differing from rat are shown. Numbering indicates both the amino acid and nucleotide and begins at the N-terminal residue of the mature protein. The heavily underlined region corresponds to protein sequence data confirming the derived amino acid sequence of rat UCP. The polyadenylation consensus sequence and the initiation consensus sequence are boxed. A tandem repeat sequence in the 3' non-coding region is lightly underlined (--^--).

initiation codon while several in-frame 'stop' codons are observed. Also the initiation codon lies within the consensus (CC)AAGATGG in which the underlined purine nucleotides A at position -3, with respect to the first nucleotide of the ATG initiation codon, and G at +4 are known to be important factors affecting initiation (27). In addition the amino acid immediately C-terminal to the initiating methionine residue, and which becomes the N-terminal amino acid of the mature protein is valine. This is consistent with the known specificity for cleavage of N-terminal methionine from newly synthesized proteins (28).

Comparison of the rat and hamster UCP sequences

The predicted protein sequence of rat UCP shows a high 91.5% homology with the hamster UCP sequence (7) (Fig. 3). Both proteins contain 306 amino acid residues and rat UCP has a 33,084 Mr as opposed to a 33,185 Mr for hamster UCP. The charge distribution over the two proteins is almost identical. All but two charged residues are conserved and even in these cases the changes are conservative, Arg_{106} and Lys_{115} of rat UCP substituting for Lys_{106} and Arg_{115} in hamster. Of the 26 mismatched residues 22 can be explained by single base codon changes. The other 4, Ala_{48}, Thr_{189}, Tyr_{258} and Met_{296} of rat UCP require two base changes. Despite the high degree of similarity between the two proteins however, it is known that hamster UCP runs about 500 Mr behind rat UCP on a sodium dodecyl sulphate-polyacrylamide gel (14).

Protein sequence comparison of UCP with the ADP/ATP carrier

Upon searching a protein data bank for homologous sequences to both rat and hamster UCP, using a cut-off homology level of 25%, only one sequence showed up in both cases. This was fragment Lys_{262} - Arg_{279} of the bovine ADP/ATP carrier (AAC) (29) which corresponds to residues Lys_{260} - Arg_{276} of UCP in both rat (68% homology) and hamster (61% homology) (Fig. 4). In addition to this high regional homology, the rat and hamster UCP each displayed an overall homology of about 20% with bovine AAC. Others have documented in detail the sequence and structural homologies of hamster UCP and bovine AAC (7) and have also noted that the region mentioned above is by far the most highly

```
                                (+) (-)          (+)              (+)
rat UCP (260-276)          K _ E G P A A F F K G F A P S F _ L R
                           *   * * *   * * * * *   *   *     * *

bovine AAC (262-279)       K D E G P K A F F K G _ A W S N V L R
                           (+)(-)(-)   (+)     (+)              (+)
```

Fig. 4 Alignment of the most highly conserved region found between rat UCP (Lys_{260} - Arg_{276}) and bovine AAC (Lys_{262}-Arg_{279}). The single letter amino acid code is used. Matches are marked by * and the charge distribution of both sequences is also given.

conserved between the two proteins. They also stressed that both proteins have a similar triplicate structure consisting basically of three 100 residue repeats (7).

A major functional similarity between both UCP and ACC is that they bind adenine nucleotides. It has been proposed by homology with other adenine nucleotide binding proteins that in AAC the hydrophobic fragment Ser_{275} - Val_{297} is involved in this binding (30, 31). This corresponds to fragment Ser_{273}-Val_{295} of UCP, which is also highly hydrophobic and fits the nucleotide binding consensus (Fig. 5). Some degree of variation is expected as AAC specifically binds adenine nucleotides whereas UCP binds both adenine and guanine nucleotides (21). In the membrane disposition model of UCP proposed by Aquila et al. (7), this region is situated within the mitochondrial inner membrane facing the cytosol, which is consistent with its role as a nucleotide binding site. In addition the hamster UCP and bovine AAC show great structural similarity in this region as measured by

```
rat UCP (273-295)     S _ F L _ [R] _ _ L [G] S W N [V I M F] V C F (E) Q L (K)(K) E L
bovine AAC (275-297)  S N V L _ [R] G M G [G] A F V [L V L Y] (D) _ _ (E) I _ (K)(K) F V
Consensus             _ _ _ _ _   R   _ _ _   G   _ _ _   H H H H   _ _ _ _ _ _ _ _ _ _
```

Fig. 5 Alignment of a region of rat UCP (Ser_{273} - Leu_{295}) with a region of bovine AAC (Ser_{275} - Val_{297}) thought to correspond to a purine nucleotide binding consensus region (boxed residues) (30, 31). The single letter amino acid code is used. The consensus itself is also given, in which H indicates a hydrophobic residue. The circled residues are not part of the defined consensus (30) but acidic residues, followed by basic residues are often found soon after the hydrophobic consensus region in adenine nucleotide binding proteins (30, 31).

hydropathy and amphiphilic hydropathy plots (7). It is interesting to note that although the most highly conserved sequence between the two proteins (Fig. 4) overlaps slightly with the proposed purine nucleotide binding site (Fig. 5), it has not itself been implicated in purine nucleotide binding (30, 31).

Location of a mitochondrial targeting sequence(s) within UCP

The targeting information for most newly synthesized mitochondrial proteins resides within an N-terminal presequence (9-11). It has been shown that import of UCP into mitochondria is competitively inhibited by a synthetic peptide corresponding to the targeting presequence of a mitochondrial matrix protein, ornithine carbamyltransferase (32). This peptide also blocks the import of ornithine carbamyltransferase and malate dehydrogenase into mitochondria (32). It was hypothesized that import of all three proteins is facilitated by the same targeting recognition apparatus. If this is the case, then one would expect some homology between the targeting presequences of ornithine carbamyltransferase and malate dehydrogenase and an internal sequence(s) within UCP. Mori et al. (33) have reached similar conclusions concerning 3-ketoacyl-CoA thiolase, another mitochondrial protein which apparently lacks an N-terminal presequence. This protein, in its native form, inhibits the uptake of a variety of mitochondrial precursor proteins including ornithine carbamyltransferase.

As targeting presequences are N-terminal, an obvious place to look for a targeting sequence within UCP is the N-terminus. However, the N-terminus of UCP does not appear to share the major characteristics of targeting presequences; it has very few basic residues and also contains some acidic residues. If the N-terminus is involved in targeting then it must be due to more subtle homologies with targeting presequences, e.g. its high serine and threonine content (34) or its three dimensional shape.

An alternative possibility is that targeting sequence(s) lie at other points within the UCP protein sequence. Several basic regions exist within UCP; Arg_{53}-Lys_{66}, Lys_{72}-Arg_{91}, Lys_{137}-Arg_{161}, Lys_{174}-Arg_{182}, Lys_{268}-Arg_{276}. In the inner membrane disposition model of Aquila et al. (7), all of these regions are thought to lie in or toward the matrix side, which is where the

N-terminal targeting presequences of other mitochondrial proteins are directed and cleaved (8).

Upon completion of this work Bouilleaud et al. (35) published a complete cDNA sequence for a rat UCP cDNA clone they had previously isolated (24). Their clone contained some cloning artifacts upstream of the cDNA sequence published, but nevertheless it is in almost complete agreement with the sequence we present here. The only differences lie in the non-coding regions. At position -44 they report A whereas we detect G. We report the region around the polyadenylation consensus AAT AAA ACC AAC CAA; they report it as AAT AAA AACC AAC __A. Also, their clone, pUCP 36, lacks the final three nucleotides ACG exhibited by $pUCP_{rat15}$ prior to the poly(A) tail. We have found other clones which lack the final five nucleotides AGACG prior to the poly(A) tail (15). Whether this sequence heterogeneity prior to the poly(A) tail represents an mRNA processing phenomenon or a cloning artifact remains to be resolved.

ACKNOWLEDGEMENTS

This work was supported by the Medical Research Council of Canada.

*To whom reprint requests should be sent.

REFERENCES

1. Nicholls, D. (1979) Biochim. Biophys. Acta 549, 1-29.
2. Nicholls, D. and Locke, R.M. (1984) Physiol. Rev. 64, 1-64.
3. Cannon, B. and Nedegaard, J. (1984) Essays Biochem. 20, 110-164.
4. Strieleman, P.J., Schalinske, K.L. and Shrago, E. (1985) J. Biol. Chem. 260, 13402-13405.
5. Klingenberg, M. and Winkler, E. (1985) EMBO J. 4, 3087-3092.
6. Locke, R.M., Rial, E., Scott, I.D. and Nicholls, D.G. (1982) Eur. J. Biochem. 129, 373-380.
7. Aquila, H., Link, T.A. and Klingenberg, M. (1985) EMBO J. 4, 2369-2376.
8. Hay, R., Böhni, P. and Gasser, S. (1984) Biochim. Biophys. Acta 779, 65-87.
9. Hurt, E.C., Pesold-Hurt, B. and Schatz, G. (1984) FEBS Lett. 178, 306-310.
10. Horwich, A.L. Kalousek, F., Mellman, I. and Rosenberg, L.E. (1985) EMBO J. 4, 1129-1135.
11. Nguyen, M., Argan, C., Lusty, C.J. and Shore, G.C. (1986) J. Biol. Chem. 261, 800-805.

12. Freeman, K.B., Chien, S.-M., Litchfield, D. and Patel, H.V. (1983) FEBS Lett. 158, 325-330.
13. Ricquier, D., Thibault, J., Bouilleaud, F. and Kuster, Y. (1983) J. Biol. Chem. 258, 6675-6677.
14. Freeman, K.B., Meyrick, K., Patel, H.V. and Ridley, R.G. (1985) Can. J. Biochem. Cell Biol. 63, 988-991.
15. Ridley, R.G., Patel, H.V., Parfett, C.L.J., Olynyk, K.A., Reichling, S. and Freeman, K.B. (1986) Biosci. Rep. (in press).
16. Grunstein, M. and Hogness, D. (1975) Proc. Natl. Acad. Sci. U.S.A. 72, 3961-3965.
17. Yang, R.C.-A., Lis, J. and Wu, R. (1979) Methods Enzymol. 68, 176-182.
18. Feinberg, A.P. and Vogelstein, B. (1983) Anal. Biochem. 132, 6-13.
19. Birnboim, H.C. (1983) Methods Enzymol. 100, 243-255.
20. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463-5467.
21. Lin, C.S. and Klingenberg, M. (1982) Biochemistry 21, 2950-2956.
22. Queen, C. and Korn, L.J. (1984) Nucl. Acids Res. 12, 581-599.
23. Maniatis, T., Fritsch, E.F. and Sambrooke, J. (1982) Molecular Cloning, A Laboratory Manual. Cold Spring Harbor Laboratory.
24. Bouilleaud, F., Ricquier, D., Thibault, J. and Weissenbach, J. (1985) Proc. Natl. Acad. Sci. U.S.A. 82, 445-448.
25. Joh, T., Nomiyama, H., Maeda, S., Shimada, K. and Morino, Y. (1985) Proc. Natl. Acad. Sci. U.S.A. 82, 6065-6069.
26. Heden, L.-O., Höög, J.-O., Larsson, K., Lake, M., Lagerholm, E., Holmgren, A., Vallee, B.L., Jörnvall, H. and von Bahr-Lindström, H. (1986) FEBS Lett. 194, 327-332.
27. Kozak, M. (1986) Cell 44, 283-292.27.
28. Teunasawa, S., Steward, J.W. and Sherman, F. (1985) J. Biol. Chem. 260, 5382-5391.
29. Aquila, H., Misra, D., Eulitz, M. and Klingenberg, M. (1982) Hoppe-Seyler's Z. Physiol. Chem. 363, 345-349.
30. Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J. (1982) EMBO J. 1, 945-951.
31. Higgins, C.F., Hiles, I.D., Whalley, K. and Jamieson, D.J. (1985) EMBO J. 4, 1033-1040.
32. Gillespie, L.L., Argan, C., Taneja, A.T., Hodges, R.S., Freeman, K.B. and Shore, G.C. (1985) J. Biol. Chem. 260, 16045-16048.
33. Mori, M., Matsue, H., Miura, S., Tatibana, M. and Hashimoto, T. (1985) Eur. J. Biochem. 149, 181-186.
34. Hurt, E.C.. Muller, U.and Schatz, G. (1985) EMBO J. 4, 3509-3518.
35. Bouilleaud, F., Weissenbach, J. and Riquier, D. (1986) J. Biol. Chem. 261, 1487-1490.