Volume 308, number 1, 54–58 FEBS 11414 August 1992

Ubiquitin of *Entamoeba histolytica* deviates in six amino acid residues from the consensus of all other known ubiquitins

Claudia Wöstmann[a], Egbert Tannich[b] and Tilly Bakker-Grunwald[a]

[a]*Universität Osnabrück, Abteilung Mikrobiologie, Barbarastrasse 11, D-4500 Osnabrück, Germany* and [b]*Bernhard-Nocht-Institut, D-2000 Hamburg, Germany*

Received 1 June 1992; revised version received 30 June 1992

The amino acid sequence of ubiquitin from *Entamoeba histolytica*, as deduced from a cDNA nucleotide sequence, deviated at six positions from the consensus of all other known ubiquitins (ranging from *Trypanosoma cruzi* to *Homo sapiens*). The corresponding residues were scattered over the primary sequence, but came close together on the surface of the folded protein structure. We conclude that (i) *E. histolytica* branched off very early from the main eukaryotic line, and (ii) this organism may yield clues as to the evolutionary development of the ubiquitin system.

Eukaryote evolution; Ubiquitin; *Entamoeba*

1. INTRODUCTION

Entamoeba histolytica, the parasite causing amebiasis in man, is a structurally simple eukaryote lacking mitochondria, a well-developed ER/Golgi apparatus and typical lysosomes [1]. In fact, the only typically eukaryotic organelle in this cell is the nucleus, but even this may deviate appreciably from its counterpart in the more complex eukaryotes; for instance, the basic DNA-binding proteins of *E. histolytica* are different from the usual histones [2].

Against this background we thought it interesting to investigate the amebal ubiquitin. Ubiquitin is a 76-amino acid protein found in all eukaryotic cells where people have looked for it; it has not been found in any prokaryote yet. The amino acid sequence of ubiquitin has been highly conserved throughout evolution, which may be a necessary consequence of its astonishing multifunctionality. The best-characterized biological role of ubiquitin is that of a covalently bound recognition signal for non-lysosomal proteolysis; ubiquitin is also found in linear ubiquitin–protein fusions. Putative functions of ubiquitination include DNA repair, cell cycle control, stress response and ribosome biogenesis. In a recent review [3] it has been estimated that over forty gene products participate in ubiquitin trafficking; we refer to this and to another review [4] for more information and for literature citations.

Below we present the coding sequence for ubiquitin derived from an amebal cDNA clone. The amino acid sequence deduced from this nucleotide sequence deviated substantially from that of all other ubiquitins analyzed so far; in particular, six of the variant positions were unique for the *E. histolytica* ubiquitin. We discuss the implications of this finding for the evolutionary history of both *E. histolytica* and the ubiquitin system.

Abbreviation: SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

The sequence of ubiquitin from *Entamoeba histolytica* reported here has been deposited in the EMBL database under accession no. X65525 EHUBIQUIT.

Correspondence address: T. Bakker-Grunwald, University of Osnabrück, Department of Microbiology, Barbarastrasse 11, D-4500 Osnabrück, Germany. Fax: (49) (541) 969 2870.

2. EXPERIMENTAL

2.1. *Cells*

E. histolytica HM1:IMSS trophozoites were grown axenically at 36°C in TYI-S-33 medium [5] with 15% serum, supplemented with penicillin (100 $\mu g \cdot ml^{-1}$) and streptomycin sulfate (100 $mg \cdot ml^{-1}$). The amebae were harvested in late-logarithmic growth by chilling on ice and a 10-min centrifugation at 400 × *g*, and washed twice in phosphate-buffered saline.

2.2. *Enrichment of ubiquitin in amebal protein extract*

The cells were suspended in phosphate-buffered saline with iodoacetamide (2 mM) to suppress amebal protease activity [6], and disrupted with a Branson sonifyer (25 pulses of 0.5 s at 40–50 W). The procedure for the enrichment of ubiquitin followed that described in [7] up to and including the ammonium sulfate fractionations. The ubiquitin-containing material precipitated by 80% (w/v) ammonium sulfate was dissolved in 50 mM Tris-HCl, pH 7.4, and subjected to SDS-PAGE and immunoblot analysis.

2.3. *SDS-PAGE and immunoblot analysis*

Proteins were separated by SDS-PAGE (16% acrylamide, 6 M urea) according to Schägger and von Jagow [8], and transferred to nitrocellulose or Immobilon-P membrane by electroblotting. Ubiquitin was detected with polyclonal antibodies against SDS-denatured bovine

 Published by Elsevier Science Publishers B.V.

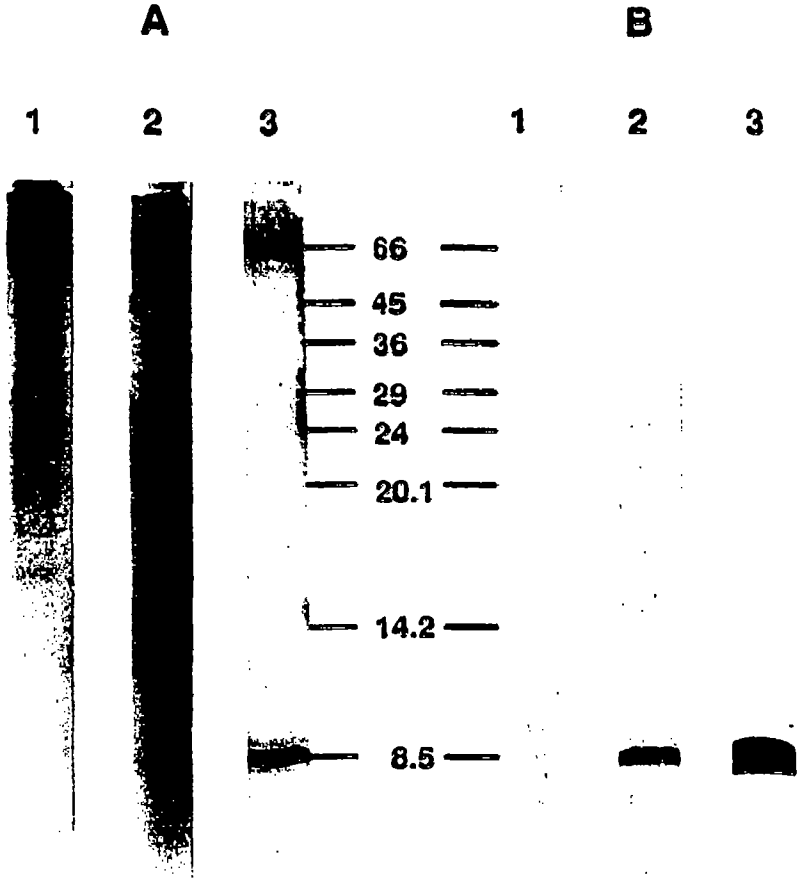

Fig. 1. Immunoblot analysis of *E. histolytica* proteins with polyclonal antibodies against SDS-denatured bovine ubiquitin. (A) Coomassie-stained gels. (B) Immunoblots. Lanes 1, amebal homogenate (20 μg); lanes 2, fraction enriched for ubiquitin (20 μg); lanes 3, bovine ubiquitin (1 μg). For further details, see section 2.

ubiquitin (prepared according to [7]); these were a kind gift from Dr. A.L. Haas, Medical College of Wisconsin.

2.4. *Amino acid sequencing*

Ubiquitin bands were cut out from Immobilon-P membrane and subjected to automated protein sequencing with an Applied Biosystems 477A gas-phase sequencer equipped with an on-line 120 A PTH amino acid analysis system.

2.5. *Isolation of genomic DNA and amplification procedure*

Genomic DNA was obtained from isolated nuclei as described in [9]. Based on the available ubiquitin sequences (Fig. 3) we choose two highly conserved regions to design two oligonucleotide primers, Eh-UB-S20 (ATG CAA ATR TTT GTR AAA AC) and Eh-UB-AS20 (TC TTT TTG RAT ATT ATA ATC) (R = A or T); in view of the high AT content of amebal coding sequences [10] we reduced the degeneracy of the primers by incorporating adenosine and thymidine at variable positions. The polymerase chain reaction was performed as described [11]. The thermal cycler (Thermal Reactor; Hybaid) was programmed for 30 cycles as follows: melting at 94°C, 1 min; annealing at 37°C, 2 min; and polymerization at 72°C, 3 min. In the first ten cycles an additional step (1 min at 55°C) was inserted after annealing. After the last cycle, polymerisation was extended for 20 min. An aliquot (5 μl) of the reaction product was separated by electrophoresis in 2% (w/v) agarose. After staining with ethidium bromide, a single fragment of about 190 bp was observed. The fragment was cut out from the gel and the DNA was extracted with phenol/chloroform, 1:1 (v/v) and with chloroform and precipitated with ethanol. We determined the nucleotide sequence of this fragment and established that the deduced amino acid sequence showed substantial homology to known ubiquitins (see Results).

2.6. *Construction and screening of the cDNA library*

The λZAP cDNA library from HM1:IMSS has been described in [9]; it contains 10^7 independent recombinant phages. The 190-bp amplified cDNA fragment was labeled with digoxigenin-11 dUTP (Non-radioactive DNA labeling and Detection Kit, Boehringer-Mannheim) and used to screen the cDNA library according to the instructions of the manufacturer. Hybridization was carried out with 5× SSC at 55°C and the filters were washed with 1× SSC (1× SSC = 0.15 M NaCl, 0.015 M sodium citrate, pH 7). Hybridizing phages were isolated and the plasmids were released according to the instructions of Stratagene. → ((Information on genomic DNA library))

2.7. *Nucleotide sequencing*

The nucleotide sequence of the amplified fragment and the inserts of the hybridizing CDNA and genomic DNA clones was determined. Sequence analysis was performed by the dideoxy chain-termination method [12]. A T7-Sequencing Kit (Pharmacia) was employed according to the instructions of the manufacturer.

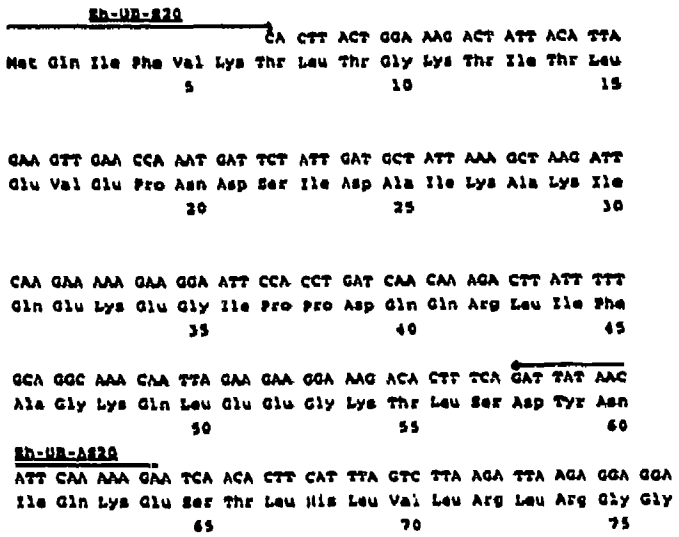

Fig. 2. Ubiquitin from *E. histolytica*. Shown is the nucleotide sequence of the cDNA clone; it codes for amino acids 8–76. The deduced amino acid sequence is given underneath. Amino acids 1–7 were determined by N-terminal sequencing of the protein. The oligonucleotide primers used for genomic amplification (Eh-UB-A20 and Eh-UB-AS20; see section 2.5) hybridize with the regions marked by arrows.

3. RESULTS

3.1. *E.* histolytica *contains ubiquitin*

E. histolytica contained a protein crossreacting with antibodies against bovine ubiquitin; this protein was appreciably enriched by a procedure designed to concentrate ubiquitin, and it had an apparent molecular weight identical to that of bovine ubiquitin (Fig. 1). Also, its seven N-terminal amino acids corresponded to those of other ubiquitins (see Fig. 2). Based on these observations we concluded that *E. histolytica*, similar to all other eukaryotes investigated so far, contains ubiquitin.

3.2. *Isolation and characterization of an amebal cDNA clone for ubiquitin*

Approximately 200,000 recombinant phages from an *E. histolytica* cDNA library were screened with a 190-bp amplified fragment covering the major part of an amebal ubiquitin gene (see section 2). One hybridizing phage was purified and the nucleotide sequence of its cDNA insert was determined. The sequence contained a single open reading frame (Fig. 2) terminated by a stop codon and followed by a 3′ non-coding region plus poly(A) tail (not shown). The amino acid sequence deduced from the cDNA corresponded with a ubiquitin molecule, from position 8 (Leu) up to position 76 (Gly). The region flanked by the primers for the polymerase chain reaction (indicated by arrows in Fig. 2) and the amplified fragment predicted identical amino acids (not shown).

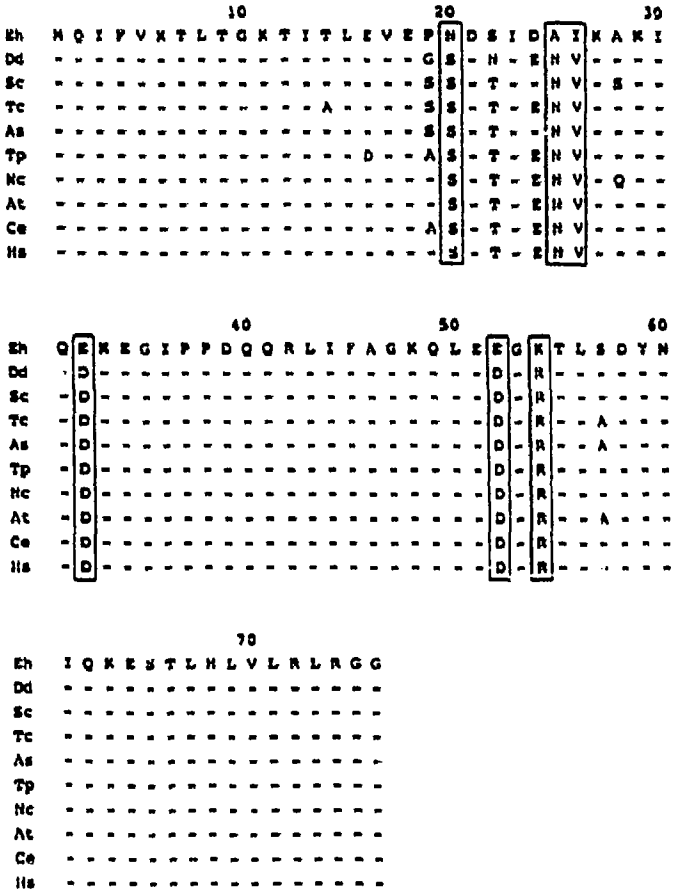

Fig. 3. Comparison of the amino acid sequence of ubiquitin from *E. histolytica* with that from other eukaryotes. Matching residues are indicated by dashes; residues that are uniquely variant in *E. histolytica* are boxed. Dd, *Dictyostelium discoideum*; Sc, *Saccharomyces cerevisiae*; Tc, *Trypanosoma cruzi*; As, *Avena sativa*; Tp, *Tetrahymena pyriformis*; Nc, *Neurospora crassa*; At, *Arabidopsis thaliana*; Ce, *Caenorhabditis elegans*; Hs, *Homo sapiens*. The sequence of *H. sapiens* is identical to those of *Xenopus laevis*, chicken, *Drosophila melanogaster*, *Manduca sexta* and *Coprinus congregatus*; that of *A. thaliana* is identical to those of *Helianthus annus*, *Hordeum vulgare*, *Glycine max*, *Linus polyphyllus* and *Chlamydomonas reinhardtii*. For references, see [4].

3.3. *Comparison of the amino acid sequence of ubiquitin from E.* histolytica *with that from other eukaryotes*

Fig. 2 shows the complete amino acid sequence of the amebal ubiquitin, based on N-terminal amino acid sequencing (residues 1–7) and the data from the cDNA (deduced residues 8–76). Meanwhile we have confirmed the full sequence with two genomic DNA clones (data not shown).

In Fig. 3 we compare the amino acid sequence of the amebal ubiquitin with the sequences published for other organisms. So far, seven residues (at positions 14, 16, 19, 22, 24, 28 and 57) were known to be variable. At these positions the amebal ubiquitin contained, with one exception (Ser 22), amino acids found in other ubiquitins as well. On top of this, the *E. histolytica* sequence deviated from the consensus for all other sequences at

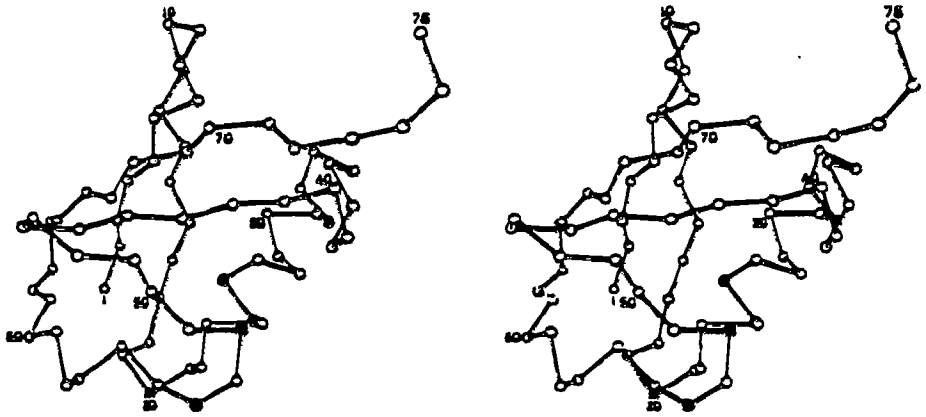

Fig. 4. Stereo drawing of the α-carbon backbone of ubiquitin [19]. The six positions (20, 25, 26, 32, 52, 54) at which *E. histolytica* side chains deviate from the consensus of the other ubiquitins have been filled out in black. Note that they lie close together on the 3D structure, and that all six residues except Ile-26 are exposed on the surface.

six positions (boxed in Fig. 2). These corresponded with 20 S>N, 25N>A, 26 V>I, 32 D>E, 52 D>E and 54 R>K; note that all substitutions are conservative. Overall, the amebal sequence deviated at 8–11 positions from each of the other sequences, whereas the latter deviate pairwise at, maximally, 3 positions. This unambiguous cleft between the pairwise distances immediately designates *E. histolytica* as an outgroup to all other organisms of which the ubiquitin has been analyzed. We confirmed this with two methods for making evolutionary trees, the unweighted pair group method with arithmetic mean [13] and the computer program PROTPARS [14]. The latter yielded 24 equivalent trees with, in each of these, *E. histolytica* branching off first; none of these trees looked particularly convincing with respect to the other organisms (data not shown).

4. DISCUSSION

4.1. *Phylogenetic status of* E. histolytica

Based, among others, on its simple structural organization, Cavalier-Smith [15] has classified *Entamoeba* (together with other primitive protozoa such as *Giardia* and the microsporidia) as Archezoa; these organisms presumably branched off very early from the main eukaryote line. For *Giardia* [16] and the microsporidia [17] this classification has been confirmed by a phylogenetic tree inferred from 16 S-like rRNA; however, *E. histolytica* branches off this tree much later (between *Euglena gracilis* and *Dictyostelium discoideum* [18]), suggesting that its primitive structure may be due to regression rather than representing an archaic trait. As shown above, the deviations in the inferred amino acid sequence of its ubiquitin designate *E. histolytica* as an unambiguous outgroup to all other organisms of which ubiquitin has been characterized; these organisms include *T. cruzi*, which is thought to be one of the oldest mitochondrion-containing cells [18]. We interpret this as very strong evidence in favor of the early branching of *Entamoeba*.

Why would we have more confidence in ubiquitin than in 16 S-like rRNA as an evolutionary marker for *Entamoeba*? Of course, 16 S-like rRNA has proven to be an invaluable generic probe for evolutionary relationships; conversely, we do not claim that ubiquitin would make a good generic probe for evolution (judging from the PROTPARS output, it does not). The point we want to make here, though, is that clearcut molecular evidence for an evolutionary cleft should outweigh more indirect evidence based on a large number of gradual transitions. Thus, whereas the amebal ubiquitin data allow for an immediate and straightforward interpretation, the calculation of 16 S-like rRNA trees is based on several assumptions; for instance, the sequences have to be aligned [18]. An even more compelling argument is that, of all variant positions in the amebal ubiquitin, six were unique for this organism. As discussed below, this finding suggests that *E. histolytica* is separated from the other organisms by a jump in the evolutionary development of the ubiquitin system.

4.2. *Evolution of the ubiquitin system*

The six uniquely variant residues in the amebal ubiquitin were spread over nearly half of the primary sequence (Fig. 3). Strikingly, though, they came close together on the tertiary structure of the protein (Fig. 4; [19]). Of the six residues, one (Asn-20) was located in a reverse turn, three (Ala-25, Ile-26, Glu-32) on the single α-helix, and the remaining two (Glu-52 and Lys-54) in the large loop; all of them except Ile-26 faced outwardly. We explain this by the following scenario.

At the time the ancestral *E. histolytica* branched off, the six residues that are now variant still largely conformed to the consensus of the other organisms. Some time thereafter, but before the next organism (presuma-

bly *T. cruzi*) branched off, ubiquitin in the main-line cells became engaged in one or more additional functional interaction that fixed part of the surface (lower right in Fig. 3) at its status quo. However, because the amebal ubiquitin had escaped this functional fixation, its corresponding residues were relatively free to drift. (As noted above, all of these exchanges are conservative; this is probably dictated by the physical properties of the protein [20].)

We thus interpret our data to mean that one or more of the functions of ubiquitin developed after *E. histolytica* branched off. In agreement with this interpretation, the pattern of hybridization of amebal mRNA with either ubiquitin cDNA or the amplified fragment looks much simpler than the corresponding patterns [21–23] from other eukaryotes (1 vs. ≥3 bands; experiments in progress).

Acknowledgements: We thank Dr. A.L. Haas for his generous gift of antibodies, Dr. H. Lill for running the PROTPARS program, Dr. J. Hoppe for N-terminal sequencing, Dr. S. Jentsch for helpful information and Dr. W.J. Cook for his kind permission to reproduce the stereo drawing of ubiquitin. This work was supported by the Deutsche Forschungsgemeinschaft (SFB171 Project C2) and by the Fonds der Chemischen Industrie.

REFERENCES

[1] Martinez-Palomo, A. (1982) The Biology of *Entamoeba histolytica*, Research Studies Press, Chichester.
[2] Torres-Guerrero, H., Peattie, D.A. and Meza, I. (1991) Mol. Biochem. Parasitol. 45, 121–130.
[3] Finley, D. and Chau, V. (1991) Annu. Rev. Cell Biol. 7, 25–69.
[4] Jentsch, S., Seufert, W. and Hauser, H.P. (1991) Biochim. Biophys. Acta 1089, 127–139.
[5] Diamond, L.S., Harlow, D.R. and Cunnick, C.C. (1978) Trans. R. Soc. Trop. Med. Hyg. 72, 431–432.
[6] Scholze, H. and Schulte, W. (1988) Biomed. Biochim. Acta 47, 115–123.
[7] Haas, A.L. and Wilkinson, K.D. (1985) Prep. Biochem. 15, 49–60.
[8] Schägger, H. and von Jagow, G. (1987) Anal. Biochem. 166, 368–379.
[9] Tannich, E., Bruchhaus, I., Walter, R.D. and Horstmann, R.D. (1991) Mol. Biochem. Parasitol. 49, 61–72.
[10] Tannich, E. and Horstmann, R.D. (1992) J. Mol. Evol. 34, 272–273.
[11] Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) Science 239, 487–491.
[12] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[13] Li, W.-H. and Graur, D. (1991) Fundamentals of Molecular Evolution, Sinauer Associates Inc., Sunderland, MA.
[14] Felsenstein, J. (1989) PHYLIP 3.2 Manual, University of California Herbarium, Berkeley.
[15] Cavalier-Smith, T. (1987) Nature 326, 332–333.
[16] Sogin, M.L., Gunderson, J.H., Elwood, H.J., Alonso, R.A. and Peattie, D.A. (1989) Science 243, 75–77.
[17] Vossbrinck, C.R., Maddox, J.V., Friedman, S., Debrunner-Vossbrinck, B.A. and Woese, C.R. (1987) Nature 326, 411–414.
[18] Sogin, M.L. (1989) Am. Zool. 29, 487–499.
[19] Vijay-Kumar, S., Bugg, C.E., Wilkinson, K.D. and Cook, W.J. (1985) Proc. Natl. Acad. Sci. USA 82, 3582–3585.
[20] Ecker, D.J., Butt, T.R., Marsh, J., Sternberg, E.J., Margolis, N., Monia, B.P., Jonnalagadda, S., Khan, M.I., Weber, P.L., Mueller, L. and Crooke, S.T. (1987) J. Biol. Chem. 262, 14213–14221.
[21] Swindle, J., Ajioka, J., Eisen, H., Sunwal, B., Jacquemot, C., Browder, Z. and Buck, G. (1988) EMBO J. 7, 1121–1127.
[22] Callis, J., Raasch, J.A. and Vierstra, R.D. (1990) J. Biol. Chem. 265, 12486–12493.
[23] Müller-Taubenberger, A., Westphal, M., Jaeger, E., Noegel, A. and Gerisch, G. (1988) FEBS Lett. 229, 273–278.