Translation termination factor aRF1 from the archaeon *Methanococcus jannaschii* is active with eukaryotic ribosomes

Maria Dontsova[a], Ludmila Frolova[b], Julia Vassilieva[a,c], Wolfgang Piendl[c,*], Lev Kisselev[b], Maria Garber[a]

[a] *Institute of Protein Research, Russian Academy of Sciences, 142292 Pushchino, Moscow Region, Russia*
[b] *Engelhardt Institute of Molecular Biology, Vavilov str. 32, 117984 Moscow, Russia*
[c] *Institute of Medical Chemistry and Biochemistry, University of Innsbruck, Fritz-Pregl-Str. 3, A-6020 Innsbruck, Austria*

Received 12 March 2000

Edited by Horst Feldmann

Abstract Class-1 translation termination factors (release factors (RFs)) from Eukarya (eRF1) and Archaea (aRF1) exhibit a high degree of amino acid sequence homology and share many common motifs. In contrast to eRF1, function(s) of aRF1 have not yet been studied in vitro. Here, we describe for the first time the cloning and expression in *Escherichia coli* of the gene encoding the peptide chain RF from the hyperthermophilic archaeon *Methanococcus jannaschii* (MjaRF1). In an in vitro assay with mammalian ribosomes, MjaRF1, which was overproduced in *E. coli*, was active as a RF with all three termination codon-containing tetraplets, demonstrating the functional resemblance of aRF1 and eRF1. This observation confirms the earlier prediction that eRF1 and aRF1 form a common structural–functional eRF1/aRF1 protein family, originating from a common ancient ancestor.

Key words: Eukaryote; Archaeon; Translation termination; eRF1; aRF1; Polypeptide release factor; *Methanococcus jannaschii*

1. Introduction

Termination of protein synthesis is carried out on the ribosome with the help of two key components, the termination codon in the mRNA encountering the A site and the class-1 polypeptide chain release factor (RF) inducing peptidyl-tRNA hydrolysis (for reviews, see [1–4]). In bacteria, RF1 decodes the stop codons UAG and UAA and RF2 decodes UGA and UAA. Mitochondrial RFs from Eukarya belong to the bacterial type both structurally and functionally. In Eukarya, only one class-1 factor, eRF1, was found which decodes all three termination codons [5,6]. In all Archaea, where the complete sequence of the genome is available, only one gene encoding a peptide chain RF, aRF1, was identified so far [7–10]. The aRF1 family is highly homologous to the eRF1 family, indicating a common origin and ancestor molecule [3,4,11–13]. For this reason, eRF1 and aRF1 sequences were joined into a single protein family, entirely different from the bacterial/mitochondrial family RF1/2/mit [6,12].

The close structural similarity between eukaryal and archaeal RFs led to the assumption that these proteins could also be functionally related [14]. In particular, aRF1 may recognize all three stop codons and function with eukaryotic ribosomes. Up to now, this hypothesis had not been tested, as, to the best of our knowledge, no aRF1 had been isolated from any archaeal species or cloned and expressed as a recombinant protein.

The aim of this work was to clone and to express the aRF1 gene from the hyperthermophilic archaeon *Methanococcus jannaschii*. The purified factor was shown to possess an in vitro release activity with mammalian ribosomes toward all three termination codons, proving for the first time functional compatibility for eukaryal and archaeal translation termination machineries.

2. Materials and methods

2.1. Cloning and overexpression of the aRF gene from M. jannaschii

In order to clone the gene encoding MjaRF1 into the regulated expression vector pET11a, the corresponding DNA fragments were generated by PCR. Synthetic oligonucleotides were obtained from MWG-Biotech (Ebersburg, Germany): 5′-CTCATGGTGATACATATGGCATCAACTGATTC-3′ (corresponding to the 5′ end of the aRF gene), 5′-GCATAAATTTGAAGGGATCCAGGAGTTATTGAT-3′ (corresponding to the 3′ end). The primer complementary to the 5′ end of the gene contains the initiator ATG codon. The *Nde*I site and the *Bam*HI site created for inserting the MjaRF1 gene into pET11a [15] are underlined. Plasmid AMJHW52 (available from ATCC) carrying the MjaRF1 gene was used as template for PCR. For the PCR reactions, the Advantage cDNA Polymerase Mix (Clontech, Palo Alto, CA, USA), which provides an antibody-mediated 'hot start' and a proofreading activity, was used. The reaction mixture contained 10 ng of template DNA, 50 pmol of each primer, 200 μM of each of the four deoxynucleotide triphosphates, 1 μl of the polymerase mix, and the PCR buffer supplied by the manufacturer, in a total volume of 50 μl. Amplification was allowed to proceed through 30 cycles at 94°C (30 s), 50°C (30 s) and 72°C (1 min). The amplified DNA was digested with *Nde*I and *Bam*HI and subcloned into pET11a. The correct sequence of the resulting plasmid, pMjaRF1.4, was confirmed by double-stranded DNA sequencing.

2.2. Overproduction and purification of aRF1

In the aRF1 gene of *M. jannaschii*, the codons mainly used for arginine are AGG and AGA (18 codons) and that for isoleucine is AUA (13 codons). These are the less frequently used codons in *Escherichia coli*. To obtain maximal yields and to avoid a potential high-level misincorporation of amino acids (e.g. lysine instead of arginine [16]) when MjaRF1 is overexpressed in large amounts from pMjaRF1.4, we used the *E. coli* expression strain BL21-CodonPlus(DE3)-RIL which carries extra copies of the tRNA genes recognizing AGA/AGG and AUA [17]. After induction with isopropyl-β-thiogalactopyranoside (IPTG) (1 mM final concentration), cells were grown for an additional 3 h, harvested by centrifugation and resus-

*Corresponding author. Fax: (43)-512-507 2872.
E-mail: wolfgang.piendl@uibk.ac.at

Abbreviations: DTT, 1,4-dithiothreitol; IPTG, isopropyl-β-thiogalactopyranoside; PMSF, phenylmethylsulfonylfluoride

0014-5793/00/$20.00
PII: S0014-5793(00)01466-6

pended in a buffer containing 50 mM Tris–HCl, pH 7.6, 0.8 M NaCl, 150 mM $MgCl_2$, 5 mM β-mercaptoethanol, 0.1 mM phenylmethylsulfonylfluoride (PMSF). After a double passage through a chilled French pressure cell at 10 000 lb/in^2, *E. coli* cell debris was removed by centrifugation (11 000 × *g* for 30 min). The supernatant was then heated at 90°C for 20 min and the denatured *E. coli* proteins removed by centrifugation at 10 000 × *g* for 1 h at 4°C. MjaRF1 was retained in the soluble fraction.

The supernatant containing MjaRF1 was diluted 16-fold to a final concentration of 50 mM NaCl and loaded onto a SP-Sepharose Fast Flow (Pharmacia) column. The column was equilibrated with 50 mM Tris–HCl, pH 7.6, 0.05 M NaCl, 1 mM 1,4-dithiothreitol (DTT), 0.1 mM PMSF. After washing with the equilibration buffer, the protein was eluted at 4°C with a linear gradient of 0.05–1.2 M NaCl. This step allows separation of the protein from RNA contamination. Fractions containing MjaRF1 were collected, pooled, diluted with 50 mM NaAc, pH 5.5, 1 M $(NH_4)_2SO_4$, 1 mM DTT, 0.1 mM PMSF, and applied to a Butyl-Sepharose Fast Flow (Pharmacia) column. The protein was eluted with a linear gradient of 1–0 M $(NH_4)_2SO_4$. After this step, the protein was approximately 80% pure. Fractions containing MjaRF1 were collected, pooled and the protein precipitated with ammonium sulfate (55% saturation) at 4°C. The protein pellet was dissolved in a buffer containing 50 mM Tris–HCl, pH 7.6, 25 mM NaCl, 1 mM DTT, 0.1 mM PMSF and loaded onto a DEAE-Sepharose (Pharmacia) column. Protein was eluted with a linear gradient of 0.025–1 M NaCl. The purity of the resulting MjaRF1 preparation was greater than 95% as estimated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE). The concentration of the protein was determined with the Bio-Rad protein assay, using bovine serum albumin as a standard.

2.3. Overproduction and purification of human eRF1 protein

Plasmid pQE-30 carrying the full-length cDNA encoding human eRF1 was used to produce eRF1 in *E. coli* M15(pREP4) [13]. Protein expression was induced by adding IPTG to a final concentration of 0.4 mM. The bacterial pellet was resuspended in lysis buffer containing 50 mM Tris–HCl, pH 7.5, 0.2 M KCl, 10 mM imidazole, 1% Nonidet P40, 10% glycerol, 2 mM β-mercaptoethanol, 1 mM PMSF and 0.2 mg/ml of lysozyme, incubated 15 min at room temperature (RT), sonicated and centrifuged at 10 000 × *g* for 30 min. Then, 150 ml of Talon™ metal affinity resin (Clontech) equilibrated with buffer A (20 mM Tris–HCl, pH 8.0, 0.2 M KCl, 10% glycerol) containing 10 mM imidazole was added to the supernatants. After 30 min incubation at RT, the Talon™ suspension was transferred onto a column and the resin washed with 10 volumes of buffer A containing 10 mM imidazole. His-tagged protein was eluted with buffer A containing 0.1 M imidazole; 0.2 ml fractions were collected and analyzed by 10% SDS–PAGE. Protein-containing fractions were pooled and dialyzed against 50 mM Tris–HCl, pH 7.5, 50 mM KCl, 10% glycerol, 0.1 mM EDTA and 1 mM DTT. Protein concentration was determined as described above.

2.4. Assays for eRF1 and aRF1 activities in vitro

The eRF1 and aRF1 release activities were measured in an in vitro assay, as hydrolysis of f[^{35}S]Met-tRNA$_f^{Met}$ associated with 80S ribosomes programmed with one of three mini-mRNAs, UUCAUGUAAA, UUCAUGUAGA or UUCAUGUGAA (synthesized by A. Veniaminova and M. Ryabkova, Novosibirsk State University), containing the initiation (AUG) and termination (underlined) codons [18]. Incubation mixtures (25 μl) contained 20 mM Tris–HCl, pH 7.5, 10 mM $MgCl_2$, 8 mM NH_4Cl and 1.5 pmol of f[^{35}S]Met-tRNA$_f^{Met}$·mini-mRNA·ribosome complex. The concentration of the human eRF1 or aRF1 was 0.3 mM. The amount of f[^{35}S]Met released at zero time (1000–1200 cpm) was subtracted from all values. Isolation of rabbit reticulocyte ribosomal subunits was described previously [19,20].

3. Results and discussion

On the basis of the nucleotide sequence, aRF1 from *M. jannaschii* was predicted to be a protein of 425 amino acids [7]. Inspection of the DNA sequence located upstream of the assigned AUG initiation codon revealed that the proposed start codon is not preceded by any Shine–Dalgarno-like sequence. The second in-frame AUG triplet, which is located four codons downstream, is preceded at the appropriate distance of four nucleotides by a ribosome binding sequence (GGTGAT) complementary to the 3′ end of 16S rRNA. Therefore, we assume that the second in-frame AUG codon is used as a start codon. This notion is supported by the alignment of RF1 factors. The N-terminus of the shorter, 421 amino acid long protein is in perfect agreement with that of the related Archaea *Methanobacterium thermoautotrophicum* [10] and *Pyrococcus horikoshii* [8].

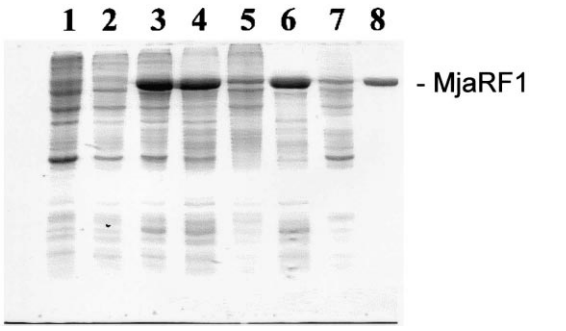

Fig. 1. Overproduction and purification of RF aRF1 from *M. jannaschii*. A 15% SDS–polyacrylamide gel stained with Coomassie blue is presented. Lane 1, total cellular protein from *E. coli* BL21(DE3); lane 2, total cellular protein from *E. coli* BL21(DE3) transformed with pMjaRF1.4, before induction with IPTG; lane 3, overproduction of aRF1 in *E. coli* BL21(DE3) carrying pMjaRF1.4 (total cellular protein 6 h after induction with IPTG); lane 4, supernatant after cell destruction; lane 5, corresponding cell debris pellet; lane 6, supernatant after heat step; lane 7, pellet after heat step (denatured *E. coli* proteins); lane 8, purified MjaRF1.

An incorrect assignation of the initiation codon has been demonstrated for several genes encoding archaeal ribosomal proteins [21]. Dennis [14] suggested that many of the assigned start codons lacking a Shine–Dalgarno sequence may be incorrect, as methanogenic Archaea, in general, use Shine–Dalgarno sequences to identify AUG, GUG or UUG codons as initiation codons.

Fig. 1 shows the result of overexpression of the MjaRF1 gene in *E. coli* and purification of the overproduced protein. The size of the protein induced in the recombinant strain (lane 3) corresponds to the molecular mass (47 500) calculated from the deduced amino acid sequence of MjaRF1 [7]. The advantage of *M. jannaschii* aRF1 purification over purification of mesophilic recombinant proteins stems from the hyperthermophilic nature of this organism. A protein denaturing step removing most of the *E. coli* proteins could be included (Fig. 1, lanes 6 and 7). Apart from its simplicity, it also inactivates the host proteins which may interfere with further protein purification (e.g. proteases). The combination of ion-exchange and hydrophobic chromatography steps resulted in purification of MjaRF1 of greater than 95% (Fig. 1, lane 8), which was sufficient to run further functional tests.

Purified human and *Xenopus laevis* eRF1 are active in vitro with each of the three stop codon-containing tetraplets and rabbit ribosomes [6]. Under the same experimental conditions, pure aRF1 from *M. jannaschii* was fully active (Table 1). The ability of aRF1 to respond to all three termination codons clearly distinguishes aRF1 from the bacterial/mitochondrial RF1/2/mit family, where any factor is unable to decode more than two termination codons.

Surprisingly, the activity of *M. jannaschii* aRF1 with rabbit

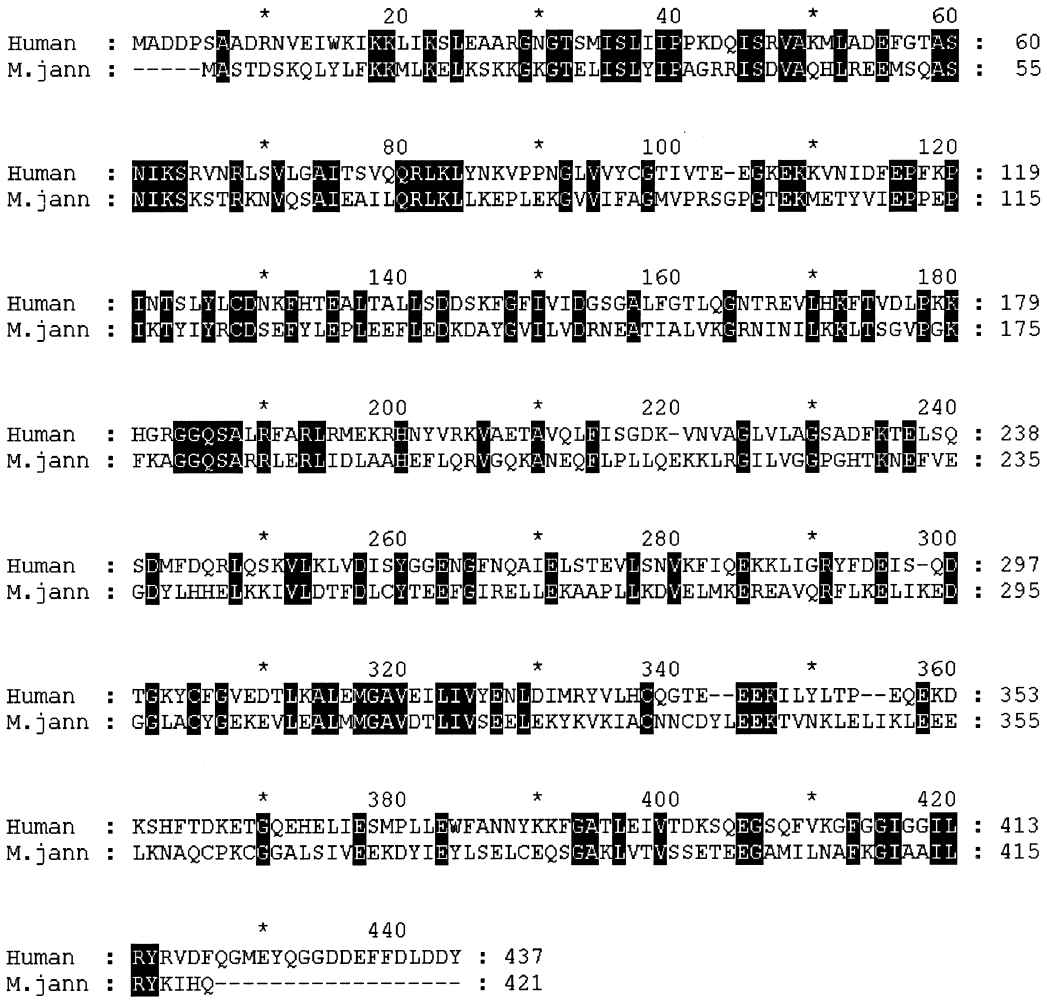

Fig. 2. Amino acid sequence alignment of human eRF1 (SWISS-PROT Acc. P46055) and MjaRF1 (SWISS-PROT Acc. Q58239). Strictly conserved residues are boxed in black.

ribosomes was virtually identical to that of eRF1 for all three stop codons, though the optimal growth temperature of the hyperthermophilic archaeon *M. jannaschii* and that of mammals differs by about 50°C. Despite that, aRF1 is active even at 37°C. Since recombinant *E. coli* was grown at 37°C, but the extract was then heated to 90°C (see Section 2), it remains unknown whether the biologically active folding of *M. jannaschii* aRF1 proceeded at normal or elevated temperatures.

One of the puzzling problems of protein structures is the identification of amino acids which confer thermostability to the protein molecule. The aRF1 family includes not only thermophilic but also mesophilic proteins. Therefore, the eRF1/aRF1 family may become a very useful model system for delineation of a structural basis of thermosensitivity/stability.

Table 1
Release activity of purified human eRF1 and MjaRF1 with rabbit ribosomes programmed with stop codon-containing tetranucleotides

Termination factor	f[^{35}S]Met released (cpm) Mini-mRNA containing stop codon		
	UAA	UAG	UGA
Human eRF1	3000	4060	2350
M. jannaschii aRF1	2600	3700	2530

For experimental details, see Section 2.

Studying hybrid translational systems provides insight into the functional and structural similarity of ribosomes and translational factors from different phylogenetic kingdoms. The functional compatibility of mammalian ribosomes with archaeal RF1, described here, suggests that the architecture of the binding site for class-1 RFs is identical or very similar in archaeal and eukaryal ribosomes, and that the overall shape of eRF1 and aRF1 molecules should also be very similar.

An amino acid sequence alignment of human eRF1 and aRF1 from *M. jannaschii* shows that the N-terminus of MjaRF1 is five amino acids shorter than human eRF1 and that the N-termini (up to amino acid residue 18) show very little identity (Fig. 2). Nevertheless, both factors are active with mammalian ribosomes demonstrating that the N-terminal variable region in the eRF1/aRF1 family is not essential for function. This result is fully consistent with the earlier observation that deletion of the first 20 N-terminal amino acid residues in human eRF1 causes no change in its functional activity [18].

At the C-terminus, aRF from *M. jannaschii* is 18 amino acids shorter than human eRF1 (Fig. 2), but this natural 'truncation' does not abolish RF activity, demonstrating that the C-terminal amino acids are not involved in mainte-

nance of RF activity. Our data on the release activity of MjaRF1 contradict the assumption that, in yeast eRF1/Sup45p, the homolog of human eRF1, the C-terminal acidic amino acids are essential for its function [22,23]. Our observation is consistent with earlier data showing that deletion of the 22 C-terminal amino acids in human eRF1 did not interfere with eRF1 binding to eRF3 [24] and with an in vitro RF activity [18]. Within the human eRF1 electron density map, determined recently [25], these residues are disordered and are assumed to be mobile regions of the protein.

Deletion of residues 411–415, which are highly conserved in Eukarya and Archaea [12,24], abolishes the ability of eRF1 to interact with eRF3. No genes homologous to eRF3 have yet been found in archaeal genomes [7–10]. However, the high degree of homology with the eRF3 binding region of eukaryal eRF1 suggests that aRF1 could form a complex with eRF3.

Acknowledgements: We are very much obliged to P. Kaiser for critical reading of the manuscript. This work was supported by the Russian Academy of Sciences, the Russian Foundation for Basic Research (Grant 99-04-48217) and the Russian Program for Scientific School Support. M.D. was supported by a Summer Short-term Fellowship from FEBS, L.K. was supported by the Human Frontier Science Program (Grant RG0032/1997-M), W.P. was supported by the Austrian Science Foundation (Grant P12070-MOB). M.G. was supported in part by an International Research Scholar's award from the Howard Hughes Medical Institute.

References

[1] Buckingham, R.H., Grentzmann, G. and Kisselev, L. (1997) Mol. Microbiol. 24, 449–456.
[2] Nakamura, Y. and Ito, K. (1998) Genes Cells 3, 265–278.
[3] Kisselev, L. and Frolova, L. (1999) Biochemistry (Moscow) 64, 14–24.
[4] Kisselev, L. (1999) Mol. Biol. (Moscow) 33, 1054–1062.
[5] Konecki, D.S., Aune, K.S., Tate, W.P. and Caskey, C.T. (1977) J. Biol. Chem. 252, 4514–4520.
[6] Frolova, L., Le Goff, X., Rasmussen, H.H., Cheperegin, S., Drugeon, G., Kress, M., Arman, I., Haenni, A.-L., Celis, J.E., Philippe, M., Justesen, J. and Kisselev, L. (1994) Nature 372, 701–703.
[7] Bult, C.J., White, O., Olsen, G.J., Zhou, L. and Fleischmann, R.D. et al. (1996) Science 273, 1058–1073.
[8] Kawarabayashi, Y., Sawada, M., Horikawa, H., Haikawa, Y. and Hino, Y. et al. (1998) DNA Res. 5, 147–155.
[9] Klenk, H.-P., Clayton, R.A., Tomb, J.-F., White, O. and Nelson, K.E. et al. (1997) Nature 390, 364–370.
[10] Smith, D.R., Doucette-Stamm, L.A., Deloughery, C., Lee, H. and Dubois, J. et al. (1997) J. Bacteriol. 179, 7135–7155.
[11] Karamyshev, A.L., Ito, K. and Nakamura, Y. (1999) FEBS Lett. 457, 483–488.
[12] Kisselev, L., Oparina, N. and Frolova, L.Y. (2000) Mol. Biol. (Moscow) 34 (in press).
[13] Frolova, L.Y., Tsivkovskii, R., Sivolobova, G., Oparina, N., Serpinsky, O., Blinov, V., Tatkov, S. and Kisselev, L. (1999) RNA 5, 1014–1020.
[14] Dennis, P.P. (1997) Cell 89, 1007–1010.
[15] Dubendorff, J.W. and Studier, F.W. (1991) J. Mol. Biol. 219, 61–68.
[16] Calderone, T.L., Stevens, R.D. and Oas, T.G. (1996) J. Mol. Biol. 262, 407–412.
[17] Carstens, C.P. and Waesche, A. (1999) Stratagies 12, 49–51.
[18] Frolova, L.Yu, Merkulova, T.I. and Kisselev, L.L. (2000) RNA 6 (in press).
[19] Frolova, L., Le Goff, X., Zhouravleva, G., Davydova, E., Philippe, M. and Kisselev, L. (1996) RNA 2, 334–341.
[20] Frolova, L.Yu., Simonsen, J.L., Merkulova, T.I., Litvinov, D.Y., Martensen, P.M., Rechinsky, V.O., Camonis, J., Kisselev, L.L. and Justesen, J. (1998) Eur. J. Biochem. 256, 36–44.
[21] Kraft, A., Lutz, C., Lingenhel, A., Gröbner, P. and Piendl, W. (1999) Genetics 152, 1363–1372.
[22] Ito, K., Ebihara, K. and Nakamura, Y. (1998) RNA 4, 958–972.
[23] Ebihara, K. and Nakamura, Y. (1999) RNA 5, 739–750.
[24] Merkulova, T.I., Frolova, L.Y., Lazar, M., Camonis, J. and Kisselev, L. (1999) FEBS Lett. 443, 41–47.
[25] Song, H., Mugner, P., Das, A.K., Webb, H.M., Evans, D.R., Tuite, M.F., Hemmings, B.A. and Barford, D. (2000) Cell 100, 311–321.