Volume 290, number 1,2, 4–8 FEBS 10152 September 1991

ADONIS 0014579391008638

Analysis of the polyadenylation consensus sequence context in the genes of nuclear encoded mitochondrial proteins

Nenad Juretić and Michael Theus

Laboratorium für Biochemie I, ETH Zentrum, CH-8092, Zürich, Switzerland

Received 27 June 1991

A compilation of the pre-mRNA ends of the genes of nuclear encoded mitochondrial proteins resulted in a consensus sequence of the type (T/A)NTTNNNNNTTTNAATAAA. Nucleotide positions +8, +13, +14, +16 and +17 downstream of the AATAAA sequence show also a predominance of nucleotide T. This consensus sequence suggests the importance of the immediate surroundings of the cannonical polyadenylation signal sequence AATAAA on the efficiency of the cleavage and polyadenylation of this specific group of pre-mRNAs.

Nuclear encoded mitochondrial protein; RNA processing

1. INTRODUCTION

Splicing of the introns and processing of the 3′-end of the pre-mRNA are important for the normal generation of almost all gene products in cells of higher eukaryotes (reviewed in [1–3]). Previous publications have stressed the importance of the three conserved sequences: AAUAAA upstream of the processing site [4], the nucleotide to which poly (A) is added (which generally is an adenosine) [5], and the region downstream of this nucleotide [6]. Experimental evidence obtained with site-directed mutagenesis has confirmed these findings [1,7], although it has not shown why some polyadenylation sequences present at other sites are not used [8,9]. In vitro polyadenylation is also stimulated by the presence of an upstream intron [10]. In this work, we report theoretical evidence for the consensus sequence upstream of the sequence AAUAAA, specific for the genes of nuclear encoded mitochondrial proteins.

2. RESULTS AND DISCUSSION

Nucleotide sequence information of 3′-ends of all known genes of nuclear-encoded mitochondrial proteins has been used for the analysis: rat cytochrome c [11], (R cyt c); bovine cytochrome P-450 (SSC) [12], (B cytP450); bovine adrenodoxin [13], (B Adrdox); rat ornithine aminotransferase [14], (R OAT); porcine mitochondrial aspartate aminotransferase [15], (P mAspAT); bovine mitochondrial ATP synthase proteolipid [16], (B P1ATPas, B P2ATPas); human aldehyde dehydrogenase 2 [17], (H ALDH2); rat carbamyl synthetase 1 [18], (R cps); human F1-ATPase (subunit) [19], (H ATPsyn); human cytochrome P-450 (SSC) [20], (H cytP450); chicken 5-aminolevulinate synthase [21], (C 5′ala); bovine ADP/ATP carrier [22], (B ADPATPc); human ornithine aminotransferase [23], (H OAT); mouse aspartate aminotransferase [24], (M mAspAT); rat serine; pyruvate aminotransferase [25], (R srpyatase); bovine ATP synthase; precursors for oligomycin sensitivity conferring protein, factor 6 and adenosinetriphosphatase inhibitor protein [26], (B OSCP, B F6, B ATPasin); rat aspartate aminotransferase [27], (R mAspAT); human medium-chain acyl-CoA dehydrogenase [28], (H acylCod); human 5-aminolevulinate synthase [29], (H 5′ala); human subunit 4 of cytochrome oxidase [30], (H coxIV); human ADP/ATP carrier [31], (H ADPATPc); rat E1α subunit of branched chain α-ketoacid dehydrogenase [32], (R E1 BCKD); porcine and human lipoamide dehydrogenase [33], (H LPdehyd, P LPdehyd); bovine adrenodoxin reductase [34], (B Adrdred); bovine steroid 11β-hydroxylase (p450 c11) [35], (B cyt45011); rat ornithine carbamoyltransferase [36], (R oct); human pyruvate dehydrogenase E1α subunit [37], (H PyrDhE1); bovine mitochondrial adenylate kinase 2 [38], (B AK2A, B AK2B); bovine mitochondrial carrier protein [39], (B mtPcP); rat mitochondrial 3-oxoacyl-CoA thiolase [40], (R 3oxCoth); human liver glutamate dehydrogenase [41], (H GDH); rat heart and liver subunit VIa of cytochrome oxidase [42], (R hcoxVIa, R lcoxVIa); mouse mitochondrial uncoupling protein [43], (M Ucp, M Ucp/min); mouse ornithine transcarbamylase [44], (M otc); mouse mitochondrial malate dehydrogenase [45], (M mMDH); human cytochrome oxidase subunit VIc [46], (H coxVIc); human aspartate amino transferase [47], (H mAspAT); rat liver mitochondrial NADH dehydrogenase 24-kDa subunit

Correspondence address: N. Juretić, Laboratorium für Biochemie I, ETH Zentrum, CH-8092 Zürich, Switzerland.

Published by Elsevier Science Publishers B.V.

Table I

Pre-mRNA ends of 110 genes of nuclear encoded mitochondrial proteins. Polyadenylation signal AATAAA is not written, only deviations of this sequence are mentioned.

H 5'ala	AAT CTA TGA TAA AAA CAT AGT CCT GGA	AATAAA	TCT GCT TAA TGG TG
C 5'ala	ATC ATA GTT AAA GCA CTA CGC TCT GAA		TTT CTA GAG CCC CTG
R cyt c	GCA CCG ATT TAA GTA AAA TTG ACT TGT	CATAAA	GTG GAT ATG ATC
R cyt c2	ATT TAC TGT GGT TTA TGT ATG ATA TCA		GAG TAT TTA ACA CTT CTT
R tescytc	CAG CCT AAA GTA ATT TTA ACA TGT TAA		GGT TGA TGT TTT TGA TTG
R cps	TTA CTG TGC CTT AAT TGT GAA CTT TTA		ATA CTA TTG AGT G
R oct	ATG CTT ACA ATG CAT ACC ATC TAA GTC	ATTAAA	TGT AAT CCA TGC TTA TTA
H otc	TAT CAT ACA CAT TTC CTT CCA CTA AAC	ATTAAA	CAC TTT GCT TAC AAT GTC
M mMDH	AAT TAT GTC TGG TCT GTT GAT AAT GAC	AGTAAA	GCA GGC TCT GAT TTT CTT
C mAspAT	CTG AGG TCA AGC AAT TTC TTG CCC CTT	AATAAG	AAC TTC CTC TGA C
H mAspAT	CGT ACT GCC GCC TGG TTT CTC TGT TAC		ATT ACT ATA GAC
R mAspAT	GTA GAC GTA CCT GGT TTT CTC TGT TAC		ATT ACT ACA GAC CC
H mAspAT	TTA TTG CGG CCT GGT TTT CTC TGT TAC		ATT ACT GTA GAC CC
P mAspAT	CTT ACT GCT GCC TGG TTT CTC TGT TAC		GTT ACT GTA GAT TTA G
B P1ATPas	CGG TGC TGG AGT GTG CTA ACC TTT ACC	ATTAAA	CAC AAT GTT TCT CT
B P2ATPas	AGA AGG AAG ACA AAT AAA TAC TGT ATT	AATAAG	AAA AAA AC
B cytP450	TCC CTT CTT CGC CCA TCC CAT GAA GGC		CGA GTG AAC GCT G
P cytP450	CAC TTC TTC ATC CAC TCC CAT GAA GGC		TAG CTG AAC TAT GT
H cytP450	ACT CCT CTT CAC CCA CCC CAT GGA GAC		CAG CTG AAC CAT CG
H coxIV	GTT ATG CCA AAC AGT TGT ACC ACT GCT		TGA CCA GTT TAC CTG
R coxIV	AAG CTG CTG TGT CCA ATG GTC CAT GTT		TGA CCA GTT TAC GTG G
H coxVb	AAA TGT GCT GTA AAG TTT CTT CTT TCC	AGTAAA	GAC TAG CCA TTG CAT TGG
B coxVb	CAC TCA GGA TGT GGA TTT TCT TCT TCC	AGTAAA	GAC TAG CAC ACA TGG TCT
H coxVIb	AGG ATC CTA AAT CAT GAC TTA CCT GCT		AAC TCA TTG GAA AAG TG
H coxVIc	GTG GGC TAA GAA TAG TTC CTC TTG ATA		CAA TTA AC
R MAOb	TCC ACT CAT AGA TTT AAG ACA TGC TTA	AATAAT	TAA AAA TAA AGC TG
R NADHdeh	ATT CAA CTG TAA AAA TGT CAC TGG AGA		ATA TGA A
H MnSOD	TTT GTT ATT GGG CAA GTG ATT GAA AAT	AGTAAA	TGC TTT GTG TGA TTG
MsMnSOD	GAA GCC ATT CCC ACT CTT GTC TCC ATT		ATC AGC TGA GCT TCC GAT
M Ucp	ATT GGG CTT CTA TGC TGG GAG ACC ACG		ACC AAC CAA AGA AAT C
M Ucp/min	TGA CCA CAG TTC ACA GCT AAT ATA CTC		ATA TTG CTA ATT CCA TC
B Ucp	TTA TTA CTG TTA TCA GTT AAT ATA TTC		TAT TCG TAA TAT CTC CTT
B mtPcP	AAG CAA GAG TTT CAA ACT TAA TGT GGA		CCC AAC TGT ACA TG
B OSCP	AGT GAG AAT TCT TAA AGT TGG AGC AAC		CAC TTC CTG AAC AGG GC
B F6	TCA AAT GTT CTT TTA ATT CTG ATT TCA		TTA TTT GGT GAT GTT AGG
B ATPasin	TCA ACC TGT GTG CTT CCA ACA AAT TAT		TTG TCA TTA GTG
H ADPATPc	ATT TTT TAT TTC AGT CAC TCC TGA CTA	AATAAC	AAT TTG GAG AAA TAA AAA
B ADPATPc	ATG AGA TCA AAA AGT TTG TCT AAT GTG	ATTAAA	ACT CGA TTT CAT TGG T
R ATPsyn	TCT CTC TGA AGA GTA TTT AAA GTT TTC		GTA TAT ACC CT
H ATPsyn	TCT TTC TGA ACA GTA TTT AAG GTT TCC		ATC GGA ATT C
H PyrDhE1	AAT TGC ATG CAG TTT GTA CAT TAC TGC	ATTAAA	AGA TGA ATT AAT GAG TGC
H PyrDhE1	GTT TAC TGT TAG TTT GTT TTG ATA AGG		GGA ATT TCT ATA GCT AG
H OAT	TTA TTT TCA ATA CTT CTT TAA ATT TAA		GCT TAT ATT TCA AAT GTC
R OAT	AAA ACA TTT ATT TTC AGT ATT TCT TTG		GCT TCG TTT TTC TTT
H ALDH2	CCT GCT TTG TAT TCT GGG CTA AGA TTC	ATTAAA	AAC TAG CTG CTC TT
R srpyats	GCA GGC CAC TGG ACT TCG GGA ATA TTC		GTA CTT GCC AGA CAT CT
B AK2A	CCA GCA AGA ATA TCA TTT GAT GTA TTG	ATTAAA	AAA GCA CTT GCT TCA TGT
B AK2B	TCC AAA TGC TGC CTT GTT TGG CCC TTA		GTG TGT TAA AAG TTT G
H E1 BCKD	AGG TGC GAG TGG CCA GCA GAG GTC ACG		CTG CAT CTC TGC GCC TGG
R E1 BCKD	TCA GAG AAA TAG CTG ATT GGC AGA GTA	AGTAAA	CCA CAG CTC TGC CTT TG
H E2 BCKD	TCA AAT AGT GAT CTT TTT TAG ACT AGA	ATTAAA	GGT ATG GGG TAA AAC ATT
H KADH	CAG TCA TTC AGA TAT GAT TCA AAT GTC	TATAAA	CCG AAC TGA TGT AAG T
H Metmlmt	GCA CCA GAA AAT AAA ATC CAT ATA TTA	ATTAAA	ACC TAT CTT G
R 3oxCoth	TGC CTT ACA TGG TGA AAT TAC AAA CTG		TGT TGC CTT AAC TCC
H AcylCod	ATA TGT TTG CAT TTT GGC AAA GAA CTT		ATT GTT CAG TGC TTA TTA
H LPdehyd	TAT TAC AGT GGG GAA TGA AGA TAC TGA		CGT CTT AAA TAT TC
P LPdehyd	ATT AAT GGT GGG AAA TGA AGA TAT GGA		CAT GTC TTA AAT
B ubiqred	TGC ATG TAC TAG AAA AGG AGA AAT CAC		TTA GCC ACG TCT GGC CCC
H mtCK	TGC CCC CGC ATC CCC TGC CTC CAT CCT	AGTAAA	GAC TCC TTG CTA TGC TGC
H Adrdox	GTT ATA TTA CAA AAA TGT CAA TCA AAT	ATTAAA	AAA TAG TTA GTG TCA TAG
H Adrdox2	ACT TCA AAG ATA GTT ATT GAC CTT ATA		TAT TTC AAA ATT TTG
H Adrdox3	TTG TAC TCA AGA AAA ATA ATG CTG ATT	TATAAA	TCT CTG CCT ATA ATA GAA

Table I *(continued)*

Pre-mRNA ends of 110 genes of nuclear encoded mitochondrial proteins. Polyadenylation signal AATAAA is not written, only deviations of this sequence are mentioned.

B Adrdox	TTT TTA TAA TTA TTA TTT CTT AAT GTA	ATTAAA	TGA GAA CAT GGA TGA ATG
R mtFumra	AAA GTG ACT TAA ATA AAT TTA TAT TAA		CAT GTA TGA AAT TT
H Adrdred	GGG TCA CCA GGT TGG GAA CAT GCT GGA		ACA GCT GCA CCC
B Adrdred	GGG TCC CCA GGT CAG GAA TAT GCT GGA		GCA CCT GCC ACC TAG
B cyt45011	AAC ACA ACA TAG GAA GCC CAC CCA GTT		TAC CAT TAA CAA T
H Ubiqbpr	TAT TTG GTC TGA AGA TGT TTT ACT TTA		TGT CTA TTG TAA TGG CTG
R GDH	ATG ATT AAC TTG GTG ATA AAA GCA GTT	ATTAAA	AGT CTA CCT TTT CC
H GDH	TGG CTT AAC CTG GTG ATA AAA GCA GTT	ATTAAA	AGT CTA CGT TTT CC
R sucCosy	TTC ACT TAC TGT GTA ACA GAG ACA GAT		TCT ATC ATT TGA TTT G
H AD/TPt1	ATT GAT AAT AAC TGA ATG TGA AAC ATC		GAC CAC TTA ATG CAC
B AD/TPt1	ATC ATG AAT AAC TCA ATG TGA AAT GTC		GAC CAC TTA ATG C
H AD/TPt2	TCT ATG TTG GGC CTC CTG CTG CAA AAC		CAG AGG ACG CAG
B AD/TPt2	TCT CCG TCG GGC ATT TCC GCT GCG GAC		ATC AGG ACA CAG
H P1	TAT CCA ATT ATG TGA CAA CCT TTG TGT		AAT TTG TTT AAA GTT
R hcoxVIa	GAG GTT CCA GTG GAC AGT TCC AAG CGC		GGT GTG AAA CTT T
R lcoxVIa	CCT TTG CTC GTA AGA GGA GAT GGC TTA		TAC GGA ATT C
M NADhchd	TGC AGT TGG GTT TTT TGT TTG TTT TTA		GCT CAT CTG TCG GGT GTG
H coxVIIa	ATT TAT TGA CTT GTA GTA ACT GCC ACC		GCA GTC TTT ACC ATG
M coxVa	AAA CCA TGT AAT AGT AAC TTG GAC TTT		GGG AAA TGA GTT TGA ACT
R coxVa	AGT ATC AAA CCA TGT AAC TTG GAC TTT		GGG AAA TGA GTT TGA CCC
H coxVIa	CCT TTG CTT GTG GCA GGA GAT GGC TTA		TAA CTT AAA CTT
R coxVb	TCA GAA TGT AAA GGA AAA CGT CTC TCT		GAC TAG CCA ACG CAC TGG
B cyt45011	AGC AAT TAT CCT CCG ATT AAA		TAC ATT AAA CAA T
R SCCoade	TCT GTG GCT GTA GAG CTG CCT GTG GTC		GCT CAC CTG TGT CTT G
R HMGCoAs	TTT TGT ATC TAA ACT TTT AAT ATG GCG	ATTAAA	AGG AGA GAA GG
H Rcytbc1	AGG TAC AAT GTC TTT TAG CTA ATT CTA	ATTAAA	AAT TAC AGA CTG GTG TAC AAG
DrmtGS1	TCC ATT GAG ATA CAT AAA TAC ATT TTG	AGTAAA	AGC AGA ACT GCG TAC AAA GCA
R MAOa	TTC TCT TTC CCA TCT GAA AAA TTG TCA	ATTAAA	AAA ACT TTT CAA AGG TTT CAA
H coxVIIc	TCA TAT GCC ATA CTA GAT ATG TTT GTC		CTT ATG ACG TG
B 75kDasb	GTG TTG AAT GTA AAT GTT AAA TAG TTT		CAT TAT ATA GAG G
B AK3	TAT TAA ACT GCT ATA ATA TTC AAC AGC		TTT TAA AAG ATG TT
H MTHFdeh	CTG ATT AAA GGG TTT TCT TTC TTT TTT		ACA CTC TGT CTG GTG TGG
R F6	AGT TGT ACA ATT AAT CTA AAA AAT TCA		CAT TCA TTC ACA GTT
H PDHE1	ACC TCA GGT CAA AGA CAT CAT ATT TGC		GAA AAC ATT AAA TAT TTA
R CTPII	CTC AGC TGT ATT TAT TGT TTA AAT AGA		ACC TGG TGT TGC TTA G
B Rieske	GTG CAG TGT CTT TGA TAG TTA ATT CTA	ATTAAA	AAT TGT AGA TGA GTG TAT AAA
R IVD	ACC ACA CAG GCT CTG CTG CCA TTT TGT	ATTAAA	CAT ACT GCA AAG T
B ATPsynt	AAC ACT TCA AAA TAT TAA CAT ACT TGT	ATTAAA	ATT TCT TAG TAA GAA GCT
H Sarc.CK	AAA ATT GTA GAT CCT GCC TAT CTT TAC		CTC TCC TTA ATA T
H MAOb	AGG ACA GTA CAC ACT TGG GTA ATT AAA		AGT TGA TTG ACC AT
H cytbc1	GTA ATT ATT CCC AGC TGA CCT AAA GTC		ACA TTC TGT TT
M coxVb	TCA GAA TGT AAA GAA AAA CTT CTC TCT		GAC TAG CCA TTG CAC CTG CTC
B coxVIII	CAC ACC TGG CAC AGA CAG GAA GAT CAG	ATTAAA	GCC CTC CTC TTT CTC CTT T
B coxVIIIi	CCT GCT CAG TAG GGC CTC CCT TGC AAC		GTC TAT TTA AAG CGT G
H coxVIII	CCT GCT TGG TGG GGT CCC CCT TGT AAC		TCT ATT TAA ACT TT
C CK	GTA TAT CTC TGT ACA AAT AGC TTG TGT		GCA CTG CCT GAT ACT TG
H cyt45011	GCC CTG GAA TAG GGT CCT GCA GGG TAG		AAG GCC CCT GTG GTC CCT G
R coxVIIc	ATT TAT TGA TGT GTA TTA ACT GCC ACC		GCA GTC CTT AAC CAT T

[48], (R NADHdeh); rat liver monoamineoxidase B [49], (R MAOb); human manganese-containing superoxide dismutase [50], (H MnSOD); maize manganese-containing superoxide dismutase [51], (MsMnSOD); human monoamine oxidase B [52], (H MAOb); human dihy-

drolipoamide dehydrogenase component of α-ketoacid dehydrogenase complexes [53], (K KADH); rat mitochondrial ATP synthase subunit β [54], (R ATPsyn); human and bovine cytochrome oxidase subunit Vb [55], (H coxVb, B coxVb); chicken creatine kinase [56], (C CK); rat glutamate dehydrogenase [57], (R GDH); human mitochondrial ubiquinone-binding protein of complex 3 [58], (H Ubqbpr); human adrenodoxin reductase [59], (H Adrdred); rat testis specific cytochrome c [60], (R tescytc); human adrenodoxin [61], (H Adrdox3); human pyruvate dehydrogenase subunit β [62], (H PyrDhE1); rat liver succinyl-CoA synthetase [63], (R sucCoAsyn); bovine cytochrome P-450 (11β) [64], (B cytp45011); human branched chain acyltransferase subunit E2 [65], (H E2BCKD); bovine NADH: ubiquinone reductase 75 kDa subunit [66], (B 75kDasubc); human cytochrome c oxidase subunit VIII [67], (H coxVIII); murine NAD-dependent methylenetetrahydrofolate dehydrogenase-methylenyltetrahydrofolate cyclohydrolase [68], (M MTHFdeh); bovine GTP:AMP phosphotransferase [69], (B AK3); bovine NADH-ubiquinone reductase [70], (B ubiqred); human methylmalonyl-CoA mutase [71], (H Metmlmt); human mitochondrial protein homologous to the bacterial and plant chaperonins [72], (H P1); porcine testis cytochrome P-450 (SCC) [73], (P cytp450); human cytochrome oxidase subunit VIb [74], (H coxVIb); human creatine kinase [75], (H CK); rat fumarase [76], (R mt-Fumra); human branched chain α-keto acid dehydrogenase complex subunit E1 α [77], (H E1 BCKD); rat brain and liver cytochrome oxidase subunit IV [78], (R coxIV); bovine uncoupling protein [79], (B Ucp); rat cytochrome c oxidase subunit VIa (Vb) [80], (R coxVb); human cytochrome c oxidase subunit VIa [81], (H coxVIa); human NAD-dependent methylene tetrahydrofolate dehydrogenase-cyclohydrolase [82], (H MTHF-deh); rat short chain acyl-CoA dehydrogenase and isovaleryl-CoA dehydrogenase [83], (R SCCode, R IVD); mouse cytochrome oxidase subunit Va [84], (M coxVa); human cytochrome bc1 complex core protein 2 [85]; human cytochrome c oxidase subunit VIIa [81], (H coxVIIa); human ADP/ATP translocase [87], (H AD/TPt1, H AD/TPt2); bovine ADP/ATP translocase [88], (B AD/TPt1, B AD/TPt2); rat cytochrome c oxidase subunit Va [89], (R coxVa); bovine ATP synthase subunit γ [90], (B ATPsynt); rat 3-hydroxy-3-methyl-glutaryl-CoA synthase [91], (R HMGCoAs); human sarcomere-specific creatine kinase [92], (H sarc.CK); human cytochrome c oxidase subunit VIIc [93], (H coxVIIc); rat monoamine oxidase A [94], (R MAOa); bovine Rieske iron–sulfur protein of ubiquinol-cytochrome c reductase [95], (B Rieske); human Rieske iron–sulfur protein of cytochrome bc1 complex [96], (H cytbc1); mouse cytochrome c oxidase subunit Vb [97], (M coxVb); rat liver carnitine palmitoyltransferase 2 [98], (R CPTII); human pyruvate dehydrogenase E1β subunit [99], (H PDHE1); chicken aspartate aminotransferase [100], (C mAspAT); rat ATP synthase coupling factor 6 [101], (R F6); *Drosophila* glutamine synthetase [102], (DmtGS1); human cytochrome P-450 (11) [103], (H cyt45011) and bovine cytochrome oxidase subunit VIII [104], (B coxVIII, BcoxVIIIi) and rat cytochrome oxidase subunit VIIa [105], (R coxVIIa). Abbreviations in the brackets are used in Table I for every gene.

Table II

Nucleotide distribution at 27 positions upstream and 18 positions downstream of the polyadenylation signal.

Position	-27	-26	-25	-24	-23	-22	-21	-20	-19	-18	-17	-16	-15	-14
%G	16	17	19	17	14	21	17	25	19	24	25	21	22	23
%A	35	22	26	29	24	23	28	26	30	30	31	33	28	28
%T	34	34	36	33	34	36	35	28	33	28	31	29	30	32
%C	16	28	19	21	28	20	19	20	18	18	13	17	20	17

Position	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	A
%G	11	21	20	15	15	12	17	21	17	11	25	19	10	0
%A	42	26	23	26	26	37	29	35	27	24	23	24	31	97
%T	39	35	42	42	28	31	32	30	34	47	42	41	25	2
%C	8	18	15	16	31	20	22	15	22	18	10	16	35	1

Position	A	T	A	A	A	+1	+2	+3	+4	+5	+6	+7	+8	+9
%G	6	0	0	0	2	29	18	12	16	15	21	15	15	18
%A	75	1	100	100	96	32	36	27	25	31	25	27	21	39
%T	19	99	0	0	1	22	24	35	36	35	32	35	42	32
%C	0	0	0	0	1	17	22	25	23	20	23	24	22	11

Position	+10	+11	+12	+13	+14	+15	+16	+17	+18
%G	20	18	20	20	21	27	26	28	31
%A	31	36	27	12	22	21	13	16	17
%T	32	28	31	43	41	32	43	40	36
%C	16	18	22	25	16	20	19	16	17

This bank contains the presently known 110 pre-mRNA ends and is most likely representative for all genes of nuclear encoded mitochondrial proteins.

As can be seen from Table II, in the upstream part of the pre-mRNA ends there is a strong preference for the nucleotides T and A, resulting in the consensus sequence: (T/A)NTTNNNNNTTTNAATAAA. Interestingly, nucleotides at positions +8, +13, +14, +16 and +17 show predominance of nucleotide T. This finding is in good accordance with the mutagenesis experiments, reporting the importance of the region downstream of the AATAAA sequence [7]. This consensus may reflect the existence of one or more processing factors specific for the pre-mRNA ends of this group of nuclear encoded proteins, as already reported for the possible splicing factors [106].

Acknowledgements: We are grateful to PD Dr. C. Richter and Prof. Dr. K.H. Winterhalter for critically reading the manuscript, and Dr. K. Vukelić for help in the calculations.

REFERENCES

[1] McLaughlan, J., Gafney, D., Whitton, J.L. and Clements, J.B. (1985) Nucleic Acids Res. 13, 1347–1368.
[2] Christophori, G. and Keller, W. (1988) Cell 54, 875–889.
[3] Utans, U. and Krämer, A. (1990) EMBO J. 9, 4119–4126.
[4] Birnstiel, M.L., Busslinger M and Strub, K. (1985) Cell 41, 349–359.

[5] Humphries, T. and Proudfoot, N.J. (1988) Trends Genet. 4, 243–245.
[6] Manley, J.L. (1988) Biochim. Biophys. Acta 950, 1–12.
[7] Sheets, M.D., Ogg, S.C. and Wickens, M.P. (1990) Nucleic Acids Res. 18, 5799–5805.
[8] Adami, G. and Nevins, J.R. (1988) EMBO J. 7, 2107–2116.
[9] Levitt, N., Briggs, D., Gil, A. and Proudfoot, N.J. (1989) Genes Dev. 1, 532–543.
[10] Niwa, M., Rose, D.S. and Berget, S.M. (1990) Genes Dev. 4, 1552–1559.
[11] Scarpulla, R.C. and Wu, R. (1983) Cell 32, 473–482.
[12] Morohashi, K., Fujii-Kuriyama, Y., Okada, Y., Sogawa, K., Hirose, T., Inayama and Omura (1984) Proc. Natl. Acad. Sci. USA 81, 4647–4651.
[13] Okamura, T., John, M.E., Zuber, M.X., Simpson, E.R. and Waterman, M.R. (1985) Proc. Natl. Acad. Sci. USA 82, 5705–5709.
[14] Mueckler, M.M. and Pitot, H.C. (1985) J. Biol. Chem. 260, 12993–12997.
[15] Joh, T., Nomiyama, H., Maeda, S., Shimada, K. and Morino, Y. (1985) Proc. Natl. Acad. Sci. USA 82, 6065–6069.
[16] Gay, N.J. and Walker, J.E. (1985) EMBO J. 4, 3519–3524.
[17] Hsu, L.C., Tani, K., Fujiyoshi, T., Kurachi, K. and Yoshida, A. (1985) Proc. Natl. Acad. Sci. USA 82, 3771–3775.
[18] Nyunoya, H., Broglie, R.E., Widgren, E.E. and Lusty, C.J. (1985) J. Biol. Chem. 260, 9346–9356.
[19] Ohta, S. and Kagawa, Y. (1986) J. Biochem. 99, 135–141.
[20] Chung, B., Matteson, K.J., Voutilainen, R., Mohandas, T.K. and Miller, W.L. (1986) Proc. Natl. Acad. Sci. USA 83, 8962–8966.
[21] Maguire, D.J., Day, A.R., Borthwick, I.A., Srivastava, G., Wigley, P.L., May, B.K. and Elliot, W.H. (1986) Nucleic Acids Res. 14, 1379–1390.
[22] Rasmussen, U.B. and Wohlrab, H. (1986) Biochem. Biophys. Res. Commun. 138, 850–857.
[23] Inana, G.,Totsuka, S., Redmond, M., Dougherty, T., Nagle, J., Shiono, T., Ohura, T., Kominami, E. and Katunuma, N. (1986) Proc. Natl. Acad. Sci. USA 83, 1203–1207.
[24] Obaru, K., Nomiyama, H., Shimada, K., Nagashima, F. and Morino, Y. (1986) J. Biol. Chem. 261, 16976–16983.
[25] Oda, T., Miyajima, H., Suzuki, Y. and Ichiyama, A. (1987) Eur. J. Biochem. 168, 537–542.
[26] Walker, J.E., Gay, N.J., Powell, S.J., Kostina, M. and Dyer, M.R. (1987) Biochemistry 26, 8613–8619.
[27] Mattingly, J.R., Rodriguez-Berrocal, F.J., Gordon, J., Iriarte, A. and Martinez-Carrion, M. (1987) Biochem. Biophys. Res. Commun. 149, 859–865.
[28] Kelly, D.P., Kim, J., Billadello, J., Hainline, B.E., Chu, T.W. and Strauss, A.W. (1987) Proc. Natl. Acad. Sci. USA 84, 4068–4072.
[29] Bawden, M.J., Borthwick, I.A., Healy, H.M., Morris, C.P., May, B.K. and Elliot, W.H. (1987) Nucleic Acids Res. 15, 8563.
[30] Zeviani, M., Nakagawa, M., Herbert, J., Lomax, M.I., Grossman, L.I., Sherbany, A.A., Miranda, A.F., DiMauro, S. and Schon, E.A. (1987) Gene 55, 205–217.
[31] Battini, R., Ferrari, S., Kaczmarek, L., Calabretta, B., Chen, S. and Baserga, R. (1987) J. Biol. Chem. 262, 4355–4359.
[32] Zhang, B., Kuntz, M.J., Goodwin, G.W., Harris, R.A. and Crabb, D.W. (1987) J. Biol. Chem. 262, 15220–15224.
[33] Otulakowski, G. and Robinson, B.H. (1987) J. Biol. Chem. 262, 17313–17318.
[34] Sagara, Y., Takata, Y., Miyata, T., Hara, T. and Horiuchi, T. (1987) J. Biochem. 102, 1333–1336.
[35] Chua, S.C., Szabo, P., Vitek, A., Grzeschik, K., John, M. and White, P.C. (1987) Proc. Natl. Acad. Sci. USA 84, 7193–7197.
[36] Takiguchi, M., Murakami, T., Miura, S. and Mori, M. (1987) Proc. Natl. Acad. Sci. USA 84, 6163–6140.
[37] Dahl, H.M., Hunt, S.M., Hutchison, W.M. and Brown, G.K. (1987) J. Biol. Chem. 262, 7398–7403.
[38] Kishi, F., Tanizawa, Y. and Nakazawa, A. (1987) J. Biol. Chem. 262, 11785–11789.
[39] Runswick, M.J., Powell, S.J., Nyren, P. and Walker, J.E. (1987) EMBO J. 6, 1367–1373.
[40] Arakawa, H., Takiguchi, M., Amaya, Y., Nagata, S., Hayashi, H. and Mori, M. (1987) EMBO J. 6, 1361–1366.
[41] Amuro, N., Ooki, K., Ito, A., Goto, Y. and Okazaki, T. (1989) Nucleic Acids Res. 17, 2356.
[42] Schlerf, A., Droste, M., Winter, M. and Kadenbach, B. (1988) EMBO J. 7, 2387–2391.
[43] Kozak, L.P., Britton, J.H., Kozak, U.C. and Wells, J.M. (1988) J. Biol. Chem. 263, 12274–12277.
[44] Scherer, E.S., Veres, G. and Caskey, T.C. (1988) Nucleic Acids. Res. 16, 1593–1601.
[45] Takeshima, H., Joh, T., Tsuzuki, T., Shimada, K. and Matsukado, Y. (1988) J. Mol. Biol. 200, 1–11.
[46] Otsuka, M., Mizuno, Y., Yoshida, M., Kagawa, Y. and Ohta, S. (1988) Nucleic Acids Res. 16, 10916.
[47] Pol, S., Bousquet-Lemercier, B., Pave-Preux, M., Pawlak, A., Nalpas, B., Berthelot, P., Hanoune, J. and Barouki, R. (1988) Biochem. Biophys. Res. Commun. 157, 1309–1315.
[48] Nishikimi, M., Hosokawa, Y., Toda, H., Suzuki, H. and Ozawa, T. (1988) Biochem. Biophys. Res. Commun. 157, 914–920.
[49] Ito, A., Kuwahara, T., Inadome, S. and Sagara, Y. (1988) Biochem. Biophys. Res. Commun. 157, 970–976.
[50] Ho, Y. and Crapo, J.D. (1988) FEBS Lett. 229, 256–260.
[51] White, J.A. and Scandalios, J.G. (1988) Biochim. Biophys. Acta 951, 61–70.
[52] Bach, A.W.J., Lan, N.C., Johnson, D.L., Abell, C.W., Bembenek, M.E., Kwan, S., Seeburg, P.H. and Shih, J.C. (1988) Proc. Natl. Acad. Sci. USA 85, 4934–4938.
[53] Pons, G., Raefsky-Estrin, C., Carothers, D.J., Pepin, R.A., Javed, A.A., Jesse, B.W., Ganapathi, M.K., Samols, D. and Patel, M.S. (1988) Proc. Natl. Acad. Sci. USA 85, 1422–1426.
[54] Garboczi, D.N., Fox, A.H., Gerring, S.L. and Pedersen, P.L. (1988) Biochemistry 27, 553–560.
[55] Zeviani, M., Sakoda, S., Sherbany, A.A., Nakase, H., Rizzuto, R., Samitt, C.E., DiMauro, S. and Schon, E.A. (1988) Gene 65, 1–11.
[56] Hossle, J.P., Schlegel, J., Wegmann, G., Wyss, M., Böhlen, P., Eppenberger, H.M., Wallimann, T. and Perriard, J. (1988) Biochem. Biophys. Res. Commun. 151, 408–416.
[57] Amuro, N., Ooki, K., Ito, A., Goto, Y. and Okazaki, T. (1989) Nucleic Acids. Res. 17, 2356.
[58] Suzuki, H., Hosokawa, Y., Toda, H., Nishikimi, M. and Ozawa, T. (1988) Biochem. Biophys. Res. Commun. 156, 987–994.
[59] Solish, S.B., Picado-Leonard, J., Morel, Y., Kuhn, R.W., Mohandas, T.K., Hanukoglu, I. and Miller, W.L. (1988) Proc. Natl. Acad. Sci. USA 85, 7104–7108.
[60] Virbasius, J.V. and Scarpulla, R.C. (1988) J. Biol. Chem. 263, 6791–6796.
[61] Picado-Leonard, J., Voutilainen, R., Kao, L., Chung, B., Strauss III, J.F. and Miller, W.L. (1988) J. Biol. Chem. 263, 3240–3244.
[62] Koike, K., Ohta, S., Urata, Y., Kagawa, Y. and Koike, M. (1988) Proc. Natl. Acad. Sci. USA 85, 41–45.
[63] Henning, W.D., Upton, C., McFadden, G., Majumdar, R. and Bridger, W.A. (1988) Proc. Natl. Acad. Sci. USA 85, 1432–1436.
[64] Kirita, S., Morohashi, K., Hashimoto, T., Yoshioka, H., Fujii-Kuriyama, Y. and Omura, T. (1988) J. Biochem. 104, 683–686.
[65] Hummel, K.B., Litwer, S., Bradford, A.P., Aitken, A., Danner, D.J. and Yeaman, S.J. (1988) J. Biol. Chem. 263, 6165–6168.
[66] Runswick, M.J., Gennis, R.B., Fearnley, I.M. and Walker, J.E. (1989) Biochemistry 28, 9452–9459.
[67] Rizzuto, R., Nakase, H., Darras, B., Francke, U., Fabrizi, G.M., Mengel, T., Walsh, F., Kadenbach, B., DiMauro, S. and Schon, E.A. (1989) J. Biol. Chem. 264 (18) 10595–10600.
[68] Belanger, C. and Mackenzie, R.E. (1989) J. Biol. Chem. 264 (9) 4837–4839.
[69] Yamada, M., Shahjahan, M., Tanabe, T., Kishi, F. and Nakazawa, A. (1989) J. Biol. Chem. 264 (15) 19192–19199.

[70] Fearnley, I.M., Runswick, M.J. and Walker, J. (1989) EMBO J. 8, 665–672.
[71] Jansen, R., Kalousek, F., Fenton, W.A., Rosenberg, L.E. and Ledley, F.D. (1989) Genomics 4, 198–205.
[72] Jindal, S., Dudani, A.K., Singh, B., Harley, C.B. and Gupta, R.S. (1989) Mol. Cell. Biol. 9, 2279–2283.
[73] Mulheron, G.W., Stone, R.T., Miller, W.L. and Wise, T. (1989) Nucleic Acids Res. 17, 1773.
[74] Taanman, J., Schrage, C., Ponne, N., Bolhuis, P., De Vries, H. and Agsteribbe, E. (1989) Nucleic Acids. Res. 17, 1766.
[75] Haas, R.C., Korenfeld, C., Zhang, Z., Perryman, B., Roman, D. and Strauss, A.W. (1989) J. Biol. Chem. 264, 2890–2897.
[76] Suzuki, T., Sato, M., Yoshida, T. and Tuboi, S. (1989) 2581–2586.
[77] Fischer, C.W., Chuang, J.L., Griffin, T.A., Lau, K.S., Cox, R.P. and Chuang, D.T. (1989) J. Biol. Chem. 264, 3448–3453.
[78] Goto, Y., Amuro, N. and Okazaki, T. (1989) Nucleic Acids Res. 17, 2851.
[79] Casteilla, L., Bouillaud, F., Forest, C. and Ricquier, D. (1989) Nucleic Acids Res. 17, 2131.
[80] Goto, Y., Amuro, N. and Okazaki, T. (1989) Nucleic Acids Res. 17, 6388.
[81] Fabrizi, G.M., Rizzuto, R., Nakase, H., Mita, S., Kadenbach, B. and Schon, E.A. (1989) Nucleic Acids Res. 17, 6409.
[82] Peri, K.G., Belanger, C. and Maskenzie, R.E. (1989) Nucleic Acids Res. 17, 8853.
[83] Matsubara, Y., Indo, Y., Naito, E., Ozasa, H., Glassberg, R., Vockley, J., Ikeda, Y., Kraus, J. and Tanaka, K. (1989) J. Biol. Chem. 264, 16321–16331.
[84] Nielsen, P.J., Ayane, M. and K hler, G. (1989) Nucleic Acids Res. 17, 6723.
[85] Hosokawa, Y., Suzuki, H., Toda, H., Nishikimi, M., Ozawa, T. (1989) J. Biol. Chem. 264, 13483–13488.
[86] Fabrizi, G.M., Rizzuto, R., Nakase, H., Mita, S., Lomax, M.I., Grossman, L.I. and Schon, E.A. (1989) Nucleic Acids Res. 17, 7107.
[87] Cozens, A.L., Runswick, M.J. and Walker, E. (1989) J. Mol. Biol. 206, 261–280.
[88] Powell, S.J., Medd, S.M., Runswick, M.J. and Walker, J.E. (1989) Biochemistry 28, 866–873.
[89] Droste, M., Schon, E. and Kadenbach, B. (1989) Nucleic Acids Res. 17, 4375.
[90] Dyer, M.R., Gay, N.J., Powell, S.J. and Walker, J.E. (1989) Biochemistry 28, 3670–3680.
[91] Ayte, J., Gil-Gomez, G., Haro, D., Marrero, P.F. and Hegardt, F.G. (1990) Proc. Natl. Acad. Sci. USA 87, 3874–3878.
[92] Haas, R.C. and Strauss, A.W. (1990) J. Biol. Chem. 265, 6921–6927.
[93] Koga, Y., Fabrizi, G.M., Mita, S., Arnaudo, E., Lomax, M.I., Aqua, M.S., Grossman, L.I. and Schon, E.A. (1990) Nucleic Acids Res. 18, 684.
[94] Kuwahara, T., Takamoto, S. and Ito, A. (1990) Agric. Biol. Chem. 54, 253–257.
[95] Usui, S., Yu, L. and Yu, C. (1990) Biochem. Biophys. Res. Commun. 167, 575–579.
[96] Nishikimi, M., Hosokawa, Y., Toda, H., Suzuki, H. and Ozawa, T. (1990) Biochem. Internat. 20, 155–180.
[97] Basu, A. and Avadhani, N.G. (1990) Biochim. Biophys. Acta 1087, 98–100.
[98] Woeltje, K.F., Esser, V., Weis, B.C., Sen, A., Cox, W.F., McPhaul, M.J., Slaughter, C.A., Foster, D.W. and McGarry, J.D. (1990) J. Biol. Chem. 265, 10720–10725.
[99] Huh, T., Cassazza, J.P., Huh, J., Chi, Y. and Song, B.J. (1990) J. Biol. Chem. 265, 13320–13326.
[100] Juretić, N., Mattes, U., Ziak, M., Christen, P. and Jaussi, R. (1990) Eur. J. Biochem. 192, 119–126.
[101] Higuti, T., Osaka, F., Yoshihara, Y., Tsurumi, C., Kawamura, Y., Tani, I., Toda, H., Kakuno, T., Sakiyama, F., Tanaka, K. and Ichihara, A. (1990) Biochem. Biophys. Res. Commun. 171, 1079–1086.
[102] Caizzi, R., Bozzetti, M.P., Caggese, C. and Ritossa, F. (1990) J. Mol. Biol. 212, 17–26.
[103] Kawamoto, T., Mitsuuchi, Y., Toda, K., Miyahara, K., Yokoyama, Y., Nakao, K., Hosoda, K., Yamamoto, Y., Imura, H. and Shizutu, Y. (1990) FEBS Lett. 269, 345–349.
[104] Lightowlers, R., Ewart, G., Aggeler, R., Zhang, Y., Calavetta, A. and Capaldi, R.A. (1990) J. Biol. Chem. 265, 2677–2681.
[105] Enders, C., Schlerf, A., Mell, O., Grossman, L.I. and Kadenbach, B. (1990) Nucleic Acids Res. 18, 7143.
[106] Juretić, N., Naussi, R., Mattes, U. and Christen, P. (1987) Nucleic Acids Res. 15, 10083–10087.