

Systembasierte Analyse von modifizierten tRNA-Basen**

Daniel Globisch, David Pearson, Antje Hienzsch, Tobias Brückl, Mirko Wagner, Ines Thoma, Peter Thumbs, Veronika Reiter, Andrea Christa Kneuttinger, Markus Müller, Stephan A. Sieber und Thomas Carell*

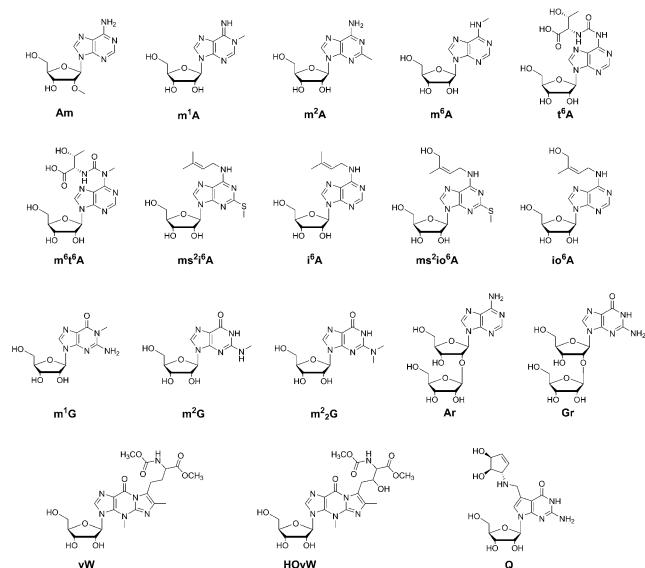
Das genetische System weist verschiedene Informationsebenen auf. Auf einer ersten Ebene ist genetische Information in Form von Sequenzen der kanonischen Basen A, C, G und T/U in DNA und RNA kodiert, die mithilfe von spezifischen Basentriplets in Aminosäuren übersetzt wird. In einer zweiten Ebene kodiert der Methylierungsstatus der Cytosinbase auf der DNA epigenetische Information und trägt zur Unterteilung der Gene in aktive und inaktive Bereiche bei. Beide Informationsebenen sind chemisch sehr gut untersucht. Weniger erforscht ist eine mögliche dritte Informationsebene, die mit der chemischen Modifikation von RNA-Nukleobasen assoziiert ist. In RNA und insbesondere tRNA sind über 100 verschiedene Modifikationen beschrieben worden,^[1] deren exakte Informationsfunktion jedoch größtenteils unbekannt ist. Für eine Reihe von Modifikationen wurde nachgewiesen, dass sie die Aufrechterhaltung des Leserahmens (reading frames) verbessern,^[2] einen Einfluss auf die RNA-Stabilität haben^[3] oder am Korrekturlesen durch tRNA-Synthetasen beteiligt sind.^[4]

Kürzlich wurde entdeckt, dass das gesamte Kollektiv von modifizierten tRNA-Nukleosiden ein regulativer Bestandteil der Stressantwort ist. Dies gibt uns den ersten Hinweis, dass Zellen möglicherweise aktiv das Modifikationsmuster als Antwort auf externe Faktoren anpassen können.^[5] Um mehr über die Funktion modifizierter Nukleobasen zu lernen, untersuchten wir die Beziehung zwischen unterschiedlichen Arten bezüglich ihres spezifischen Modifikationsgehalts. Unter Verwendung eines parallelen, systembasierten Ansatzes entdeckten wir, dass der kollektive Satz von modifizierten Basen in hohem Maße artenspezifisch und mit phylogenetischen Beziehungen assoziiert ist. Diese Daten ermöglichen

uns, einen detaillierten phylogenetischen Stammbaum zu berechnen, der mit phylogenetischen Stammbäumen übereinstimmt, die aus konventionellen Daten wie rRNA-Sequenzhomologien,^[6] konservierten orthologen Genen,^[7] Sequenzen von tRNA-Synthetasen^[8] und tRNA-abhängigen Amidotransferasen^[9] berechnet wurden. Dieses Ergebnis zeigt, dass die Gesamtheit der Basenmodifikationen nicht universell, sondern stark artenspezifisch ist. Sie wurden unter Evolutionsdruck entwickelt, um das entsprechende Basentriplett mit der übereinstimmenden kognitiven Aminosäure mit optimaler Effizienz ablesen zu können.

Für diese Studie verwendeten wir unsere kürzlich entwickelte LC-MS-basierte Quantifizierungsmethode modifizierter Nukleoside mit isotopenmarkierten Standards.^[10] Für die parallele Quantifizierung synthetisierten wir 18 tRNA-Modifikationen sowohl in der natürlich vorkommenden, als auch in isotopenmarkierter Form (Schema 1). Die meisten dieser Nukleoside kommen in 3'-Position zum Anticodon in Position 37 vor, während andere über die ganze tRNA verteilt sind. Die untersuchten Modifikationen sind an einer Reihe von biologischen Prozessen beteiligt, z.B. der Strukturstabilität, der Stabilisierung der Codonbindung und der Translationsinitiation.^[2a,3,10b,11]

Mit diesen tRNA-Nukleosiden analysierten wir das tRNA-Modifikationsmuster von 16 Arten, um unterschiedli-



Schema 1. Modifizierte Nukleoside, die in natürlich vorkommender und isotopenmarkierter Form synthetisiert und für die parallelen Quantifizierungsstudien als Standards verwendet wurden.

[*] Dr. D. Globisch, Dr. D. Pearson, Dipl.-Chem. A. Hienzsch, Dr. T. Brückl, Dipl.-Chem. M. Wagner, M. Sc. I. Thoma, Dipl.-Chem. P. Thumbs, Dipl.-Chem. V. Reiter, M. Sc. A. C. Kneuttinger, Dr. M. Müller, Prof. S. A. Sieber, Prof. T. Carell
Center for Integrated Protein Science, Department of Chemistry, Ludwig-Maximilians-Universität München
Butenandtstraße 5–13, 81377 München (Deutschland)
E-Mail: thomas.carell@cup.uni-muenchen.de
Homepage: <http://www.carellgroup.de>

[**] Wir danken dem Exzellenzcluster CiPSM, der Deutschen Forschungsgemeinschaft (CA-275 8/4) und dem SFB749 für finanzielle Unterstützung. Wir danken Prof. Wolfgang Steglich für die Bereitstellung des Pilzmaterials und dessen taxonomische Einordnung. Wir danken außerdem Kerstin Kurz und Dilek Özden für ihre technische Unterstützung.

Hintergrundinformationen zu diesem Beitrag sind im WWW unter <http://dx.doi.org/10.1002/ange.201103229> zu finden.

a)

	m^1A	m^2A	m^6A	Am	t^6A	m^6t^6A	t^6A	ms^2t^6A	io^6A	ms^2io^6A	m^1G	m^2G	m^2G	Q	yW	$HoYW$	Ar	Gr
Proteobacteria																		
<i>E. coli</i>	0	280	66	0	71	31	6	111	0	0	232	0	0	179	0	0	0	0
<i>P. putida</i>	59	303	37	9	60	30	85	0	79	0	163	0	0	240	0	0	0	0
<i>B. thailandensis</i>	0	93	5	5	89	0	20	43	0	35	77	0	0	120	0	0	0	0
Firmicutes																		
<i>B. subtilis</i>	244	38	261	8	182	0	43	34	0	0	219	38	0	184	0	0	0	0
<i>L. welshimeri</i>	180	29	94	0	207	0	71	0	0	0	144	0	0	0	0	0	0	0
<i>S. aureus</i> NCTC	184	2	86	0	118	0	69	27	0	0	128	0	0	154	0	0	0	0
<i>D. radiodurans</i>	73	5	5	3	237	0	58	12	0	2	542	0	0	0	0	0	0	0
Eukaryoten																		
<i>S. cerevisiae</i>	400	0	30	37	250	0	102	0	0	0	888	617	486	0	22	0	13	0
<i>C. albicans</i>	400	0	30	84	169	0	100	0	0	0	869	430	315	0	16	0	0	16
<i>C. nebularis</i>	382	1	4	45	68	0	3	0	68	0	703	338	474	59	0	0	0	24
<i>F. fomentarius</i>	474	0	3	104	34	0	8	0	14	0	604	168	290	5	0	0	0	13
<i>S. scrofa</i> (Herz/Mitochondrien)	549	0	52	16	145	3	12	80	0	0	249	51	413	36	0	1	0	0
<i>S. scrofa</i> (Herz/Cytoplasma)	1158	0	46	31	171	15	33	6	0	0	656	723	450	25	0	17	0	0

b)

	Pathogene/resistente Bakterien	<i>P. aeruginosa</i>	<i>B. cenocepacia</i>	<i>L. monocytogenes</i>	<i>S. aureus</i> Mu50
0	314	37	8	52	34
0	92	6	6	92	0
259	31	117	0	279	0
247	2	111	7	177	0

Abbildung 1. tRNA-Modifikationsmuster der hier untersuchten Arten. Die Mengen an tRNA-Modifikationen sind als modifizierte Nukleoside pro 1000 tRNAs wiedergegeben (mittlere Standardabweichung 7%). Dunkle Felder repräsentieren höhere Modifikationsmengen. a) Eukaryotische und nichtpathogene/resistente Bakterienarten, gruppiert nach ihrer Taxonomie. b) Pathogene/resistente Bakterien. Die Verbindungslien rechts markieren die Verknüpfung des jeweiligen Bakterienstamms mit den entsprechenden nichtpathogenen/nichtresistenten Stämmen.

che Zweige des phylogenetischen Stammbaums abzudecken. Neben fünf Eukaryoten, fünf Gram-negativen Proteobakterien und fünf Gram-positiven Bakterien der Firmicutes untersuchten wir das Bakterium *Deinococcus radiodurans*, das Merkmale beider Bakterienklassen aufweist. *D. radiodurans* wird normalerweise als ein Gram-positives Bakterium betrachtet, hat aber zusätzliche Zellwände, die denen der Gram-negativen Bakterien ähneln.^[12] Die gesamte tRNA wurde entweder aus frischem Gewebe oder aus Bakterien- und Hefekulturen extrahiert, die in komplexen Medien unter Optimalbedingungen gezüchtet wurden. Anschließend wurde die tRNA enzymatisch verdaut und die Nukleosidmischung mit unserer quantitativen LC-MS-Methode analysiert.^[10a]

Die Ergebnisse für alle untersuchten Arten sind in Abbildung 1 dargestellt. Das erste Schlüsselergebnis ist, dass die Modifikationsmengen drastisch zwischen unterschiedlichen Arten variieren. Zum Beispiel unterscheidet sich die Menge an m^2A von 1 (1 Modifikation pro 1000 tRNA-Molekülen) im Pilz *Clitocybe nebularis* bis 314 im Proteobakterium *Pseudomonas aeruginosa*, und die von m^2G von 38 im Gram-positiven Bakterium *Bacillus subtilis* bis 723 im Säugetier *Sus scrofa*. Hohe Mengen an bestimmten Modifikationen sind charakteristisch für spezifische Gruppen von Organismen. Zum Beispiel ist der Wert von m^2A hauptsächlich in Gram-negativen Bakterien hoch, während die Modifikationen m^1A , m^1G , m^2G und m^2G in großen Mengen in Eukaryoten vorkommen. Diese ersten Beobachtungen zeigen deutlich, dass sich evolutionäre Beziehungen nicht nur im artenspezifischen Vorkommen, sondern auch in der Quantität von tRNA-Modifikationen widerspiegeln.

Vier N^6 -Isopentenyladenosin-Derivate (t^6A , ms^2t^6A , io^6A und ms^2io^6A) wurden parallel quantifiziert. Dadurch konnten wir die Beziehung zwischen diesen Modifikationen im Detail

analysieren. Während die einfachste Modifikation dieser Gruppe, t^6A , in allen untersuchten Arten vorkommt, sind die komplexeren Modifikationen unterschiedlich verteilt. Die Modifikation ms^2t^6A ist in allen Bakterien außer *Listeria* und *P. putida* vorhanden und kommt bei Eukaryoten nur in *S. scrofa*, aber nicht in den untersuchten Pilzen vor. In Übereinstimmung mit der Literatur kommen die Hydroxyderivate io^6A und ms^2io^6A hauptsächlich in γ -Proteobakterien wie *Pseudomonas* vor.^[13] Überraschenderweise konnten wir auch große Mengen dieser Verbindungen im β -Proteobakterium *Burkholderia* finden. Spuren von ms^2io^6A konnten wir auch in *D. radiodurans* nachweisen, was sehr überraschend ist, da die Sequenz des modifizierenden Enzyms *MiaE* nicht in Genanalysen gefunden werden konnte.^[14] Dieses Ergebnis weist darauf hin, dass ein Enzym mit einer signifikant anderen Sequenz für diese Modifikation verantwortlich sein könnte. Die beiden komplexen Pilze *C. nebularis* und *Fomes fomentarius* verwenden hauptsächlich io^6A , was sie von den anderen untersuchten Eukaryoten unterscheidet.

Unsere Daten belegen, dass die Modifikationen m^1G und t^6A in großen Mengen in allen untersuchten Arten vorkommen. Insbesondere kommen diese Modifikationen in extrem hoher Zahl in dem ungewöhnlichen Bakterium *D. radiodurans* vor, wo sie den Hauptanteil aller quantifizierten Modifikationen ausmachen. Diese beiden modifizierten Nukleoside sind die einzigen tRNA-Nukleoside, die in allen drei Reichen des Lebens und auch in Mitochondrien und Chloroplasten vorkommen.^[15] Deshalb folgern wir, dass diese Modifikationen zu den ältesten RNA-Modifikationen gehören, die möglicherweise zur frühen Entwicklung des Lebens beitrugen.^[16] Die Entdeckung, dass diese Nukleoside sehr häufig in *D. radiodurans* vorkommen, deutet darauf hin, dass sich diese Bakterienart aus einer sehr alten evolutionären

phylogenetischen Abzweigung entwickelt hat, was in Übereinstimmung mit evolutionären Schlussfolgerungen aus anderen Daten ist.^[17]

Um eine statistische Analyse der unterschiedlichen Modifikationsmengen zwischen den unterschiedlichen Spezies durchzuführen, haben wir einen hierarchischen Clustering-Algorithmus auf unsere Daten mit den Programmen „Cluster“ und „Treeview“ angewendet.^[18] Mit diesen Berechnungen werden Spezies mit ähnlichen Modifikationsmengen zusammen gruppiert. Wir verwendeten hierfür den euklidischen Abstandsalgorithmus ohne Normalisierung der Daten, da der Algorithmus die absoluten Wertunterschiede ins Verhältnis setzt. Somit sollten alle tRNA-Funktionalitäten repräsentiert und geringe Variationen der Modifikationsmengen zwischen den Spezies einbezogen sein. Interessanterweise ergeben die Clustering-Ergebnisse unserer quantitativer Daten (Abbildung 2a) einen hochgradig genauen phylogenetischen Stammbaum. Das erhaltene Bakterien-Clustering wird durch Bootstrap-Rechnungen stark gestützt. Das Clustering der eukaryotischen Arten ist nur geringfügig schlechter. Zum möglichen Vergleich mit etablierten Methoden berechneten wir einen konventionellen phylogenetischen Stammbaum,

der auf genetische Unterschieden basiert (Abbildung 2b). Dieser Stammbaum wurde basierend auf einer publizierten Methode aufgebaut,^[7a] wobei genetische Daten für alle möglichen Arten verwendet wurden, die wir in dieser Studie untersucht haben.

Überraschenderweise zeigen beide Clustering-Experimente sehr ähnliche Gruppierungsmuster. Dies bestätigt, dass die Mengen an modifizierten Nukleosiden mit genetischen Unterschieden zwischen Arten eng verknüpft sind. Eukaryoten sind deutlich getrennt von Bakterien, und es gibt eine klare Unterteilung zwischen Gram-negativen und Gram-positiven Bakterien. Ein genauerer Blick ermöglicht die exakte Unterscheidung der Gram-positiven Firmicutes-Bakterien (*Bacillus*, *Listeria* und *Staphylococcus*) und demonstriert die Auflösung und Genauigkeit von unserer Analyse. Ähnlich clustern die Eukaryoten in exzellenter Korrelation mit dem phylogenetischen Stammbaum und einer deutlichen Abtrennung zwischen Säugetieren und Pilzen, sowie zwischen komplexen Pilzen und Hefen untereinander. Auch zeigt der Cluster der Gram-negativen Bakterien eine gute Übereinstimmung mit der phylogenetischen Einstufung. Eine Ausnahme bildet hier lediglich *Pseudomonas aeruginosa*, welches

mit *E. coli* und nicht mit dem enger verwandten *P. putida* gruppiert. Die eng verwandten Bakterienklassen der γ -Proteobakterien (*Escherichia* und *Pseudomonas*) und der β -Proteobakterien (*Burkholderia*) sind jedoch klar zu unterscheiden. Interessanterweise lässt sich *D. radiodurans* keiner der zwei Bakterienklassen zuordnen, was dessen mehrdeutigen Charakter unterstreicht.

Als weiteres Ergebnis wurde gefunden, dass die Modifikationsmengen in mitochondrialer tRNA aus dem Herzgewebe von *S. scrofa* mit den Bakteriendaten clustern, während die zytosolischen Modifikationsmengen erwartungsgemäß mit den anderen Eukaryoten clusterten. Diese Resultate belegen, dass mitochondriale tRNA in Bezug auf ihre Modifikationsmengen prokaryotischen Charakter beibehält, was in Übereinstimmung mit qualitativen Analysen von tRNA-Sequenzen und tRNA-Modifikationen ist.^[19] Unser Ergebnis deutet zusätzlich darauf hin, dass Mitochondrien möglicherweise prokaryotische Systeme aufrecht erhalten, um die Regulation der Modifikationsmengen zu ermöglichen.

Die hohe Auflösung der Analyse zeigt, dass die Quantifizierung von tRNA-Modifikationen als Methode zur Unterscheidung zwischen

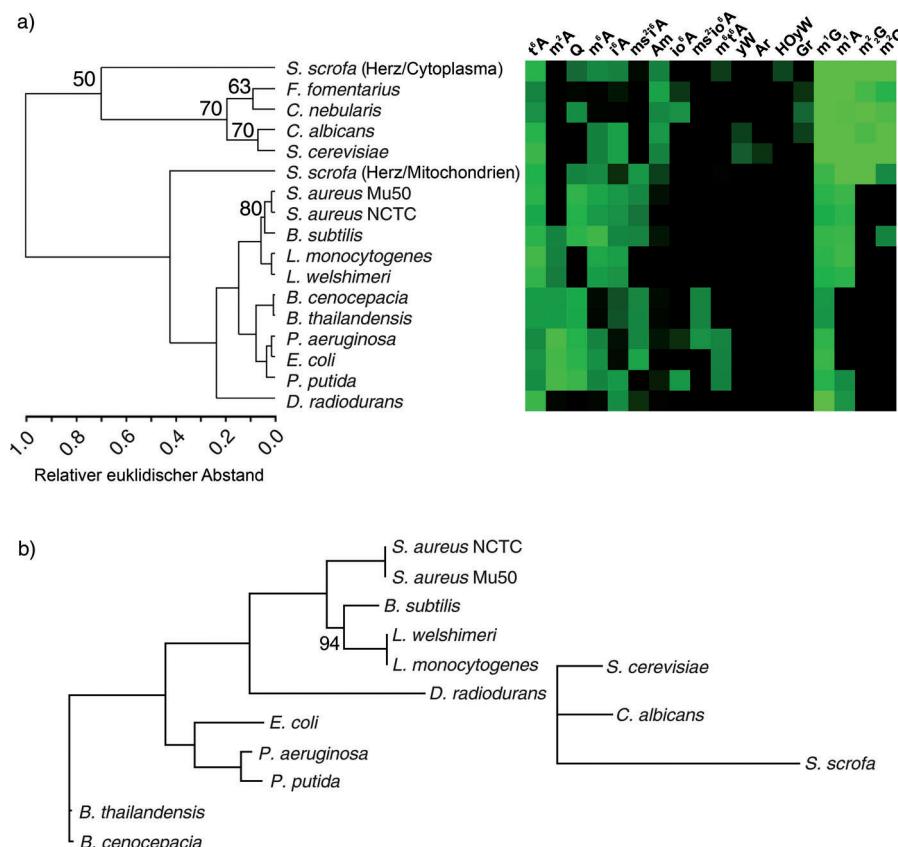


Abbildung 2. Cluster-Artenanalyse. a) Artenclustering basierend auf unseren quantitativen tRNA-Modifikationsdaten. Die Werte an den Verzweigungspunkten sind Bootstrap-Werte der schwächeren Cluster; nichtmarkierte Verzweigungspunkte weisen sehr hohe Bootstrap-Werte (> 95) auf. Die horizontale Skala zeigt die relativen euklidischen Abstände zwischen Clustern an jeder Verzweigung. b) Phylogenetischer Stammbaum basierend auf der Analyse einer Gruppe orthologer Gene. Die Bootstrap-Werte wurden durch wiederholte Probenahme bestimmt; nichtmarkierte Verzweigungspunkte weisen sehr hohe Bootstrap-Werte (100) auf. Die Längen der horizontalen Linien repräsentieren die relative Häufigkeit von Aminosäuresubstitutionen zwischen Arten/Verzweigungspunkten.

pathogenen Bakterien und deren nichtpathogenen Spezies angewendet werden könnte. Um das Potential unseres Ansatzes einzuschätzen zu können, verglichen wir drei Paare von pathogenen und nichtpathogenen Bakterien der Genera *Pseudomonas*, *Burkholderia* und *Listeria* sowie ein Paar Methicillin-resistenter und nichtresistenter *S.-aureus*-Stämme (Abbildung 1). Diese Bakterien gehören zu den bedrohlichsten klinischen Pathogenen und sind für viele Todesfälle verantwortlich.

Tatsächlich sind pathogene und nichtpathogene Bakterien deutlich zu unterscheiden. Die beiden *Pseudomonas*-Spezies weisen sogar unterschiedliche tRNA-Modifikationsarten auf. Während die pathogene Spezies *P. aeruginosa* alle vier **i⁶A**-Derivate aufweist (**i⁶A**, **ms²i⁶A**, **io⁶A** und **ms²io⁶A**), enthält die nichtpathogene Spezies *P. putida* nur **i⁶A** und **io⁶A** ohne die Methylthiogruppe. Zusätzlich ist das Nukleosid **m¹A** nur in der nichtpathogenen Spezies vorhanden. Im Fall der *Listerien* und *Staphylokokken* ist der Modifikationsgehalt in der pathogenen und resistenten Spezies jeweils deutlich höher. Dies ist ein Hinweis auf einen veränderten Translationsprozess mit einem erhöhten Bedarf an Modifikationen. Nur die *Burkholderia*-Spezies waren nicht deutlich durch ihr Modifikationsmuster zu unterscheiden. Generell bietet unsere Methode aber eine neue Möglichkeit zur Unterscheidung zwischen pathogenen und nichtpathogenen Bakterien.

Zusammengefasst haben wir 18 tRNA-Modifikationen in 16 Spezies quantitativ bestimmt. Die Ergebnisse ermöglichen einen tieferen Einblick in die Evolution von tRNA-Modifikationen, und es zeigt sich, dass diese Modifikationen unterschiedliche Spezies sehr fein charakterisieren und mit phylogenetischen Unterschieden einhergehen. Die beschriebenen Ergebnisse können zudem zur Unterscheidung von Arten und sogar von pathogenen und nichtpathogenen Bakterien verwendet werden.

Eingegangen am 11. Mai 2011

Online veröffentlicht am 31. August 2011

Stichwörter: Bioanalytische Methoden · Isotopenmarkierung · Molekulare Evolution · Nukleoside · tRNA

- [1] a) H. Grosjean, R. Benne, *Modification and Editing of RNA, Vol. 1*, ASM, 1998; b) F. Jühling, M. Mörl, K. Hartmann, M. Sprinzl, F. Stadler, J. Pütz, *Nucleic Acids Res.* **2009**, 37, D159–D162.

- [2] a) P. F. Agris, F. A. Vendeix, W. D. Graham, *J. Mol. Biol.* **2007**, 366, 1–13; b) A. Ambrogelly, S. Paloura, D. Söll, *Nat. Chem. Biol.* **2007**, 3, 29–35; c) J. F. Atkins, G. R. Björk, *Microbiol. Mol. Biol. Rev.* **2009**, 73, 178–210.
[3] Y. Motorin, M. Helm, *Biochemistry* **2010**, 49, 4934–4944.
[4] L. A. Sylvers, K. C. Rogers, M. Shimizu, E. Ohtsuka, D. Söll, *Biochemistry* **1993**, 32, 3836–3841.
[5] C. T. Chan, M. Dyavaiah, M. S. DeMott, K. Taghizadeh, P. C. Dedon, T. J. Begley, *PLoS Genet.* **2010**, 6, e1001247.
[6] a) G. E. Fox, E. Stackebrandt, R. B. Hespell, J. Gibson, J. Maniloff, T. A. Dyer, R. S. Wolfe, W. E. Balch, R. S. Tanner, L. J. Magrum, L. B. Zablen, R. Blakemore, R. Gupta, L. Bonen, B. J. Lewis, D. A. Stahl, K. R. Luehrs, K. N. Chen, C. R. Woese, *Science* **1980**, 209, 457–463; b) H. J. Hori, *Mol. Evol.* **1975**, 7, 75–86.
[7] a) F. D. Ciccarelli, T. Doerks, C. von Mering, C. J. Creevey, B. Snel, P. Bork, *Science* **2006**, 311, 1283–1287; b) M. Wu, J. A. Eisen, *Genome Biol.* **2008**, 9, R151–R151.
[8] K. Shiba, H. Motegi, P. Schimmel, *Trends Biochem. Sci.* **1997**, 22, 453–457.
[9] K. Sheppard, D. Söll, *J. Mol. Biol.* **2008**, 377, 831–844.
[10] a) T. Brückl, D. Globisch, M. Wagner, M. Müller, T. Carell, *Angew. Chem.* **2009**, 121, 8074–8077; *Angew. Chem. Int. Ed.* **2009**, 48, 7932–7934; b) D. Pearson, A. Hienzsch, M. Wagner, D. Globisch, V. Reiter, D. Özden, T. Carell, *Chem. Commun.* **2011**, 47, 5196–5198; c) M. Müntzel, D. Globisch, T. Brückl, M. Wagner, V. Welzmiller, S. Michalakis, M. Müller, M. Biel, T. Carell, *Angew. Chem.* **2010**, 122, 5503–5505; *Angew. Chem. Int. Ed.* **2010**, 49, 5375–5377.
[11] C. Förster, K. Chakraburty, M. Sprinzl, *Nucleic Acids Res.* **1993**, 21, 5679–5683.
[12] J. R. Battista, *Annu. Rev. Microbiol.* **1997**, 51, 203–224.
[13] J. J. Janzer, J. P. Raney, B. D. McLennan, *Nucleic Acids Res.* **1982**, 10, 5663–5672.
[14] K. H. Kaminska, U. Baraniak, M. Boniecki, K. Nowaczyk, A. Czerwoniec, J. M. Bujnicki, *Proteins* **2008**, 70, 1–18.
[15] A. Morin, S. Auxilien, B. Senger, R. Tewari, H. Grosjean, *RNA* **1998**, 4, 24–37.
[16] G. R. Björk, K. Jacobsson, K. Nilsson, M. J. O. Johansson, A. S. Byström, O. P. Persson, *EMBO J.* **2001**, 20, 231–239.
[17] a) R. S. Gupta, *Microbiol. Mol. Biol. Rev.* **1998**, 62, 1435–1491; b) K. S. Makarova, L. Aravind, Y. I. Wolf, R. L. Tatusov, K. W. Minton, E. V. Koonin, M. J. Daly, *Microbiol. Mol. Biol. Rev.* **2001**, 65, 44–79.
[18] a) M. B. Eisen, P. T. Spellman, P. O. Brown, D. Botstein, *Proc. Natl. Acad. Sci. USA* **1998**, 95, 14863–14868; b) N. Jessani, Y. Liu, M. Humphrey, B. F. Cravatt, *Proc. Natl. Acad. Sci. USA* **2002**, 99, 10335–10340.
[19] a) H. Grosjean, V. de Crécy-Lagard, C. Marck, *FEBS Lett.* **2010**, 584, 252–264; b) K. Watanabe, *Proc. Jpn. Acad. Ser. B* **2010**, 86, 11–39.