

The RNA modification database—1998

James A. McCloskey^{1,2} and Pamela F. Crain^{1,*}

¹Department of Medicinal Chemistry and ²Department of Biochemistry, 311A Skaggs Hall, University of Utah, Salt Lake City, UT 84112, USA

Received October 6, 1997; Accepted October 8, 1997

ABSTRACT

The RNA modification database provides a comprehensive listing of posttranscriptionally modified nucleosides from RNA, and is maintained as an updated version of the initial printed report [Limbach, P.A., Crain, P.F. and McCloskey, J.A. (1994) *Nucleic Acids Res.*, 22, 2183–2196]. Information provided for each nucleoside includes: the type of RNA in which it occurs and phylogenetic distribution; common chemical name and symbol; Chemical Abstracts registry number and index name; chemical structure; initial literature citations for structural characterization or occurrence, and for chemical synthesis. The data are available through the World Wide Web at: <http://www-medlib.med.utah/RNAMods/RNAMods.html>

INTRODUCTION

Both the structural diversity and extent of posttranscriptional modification in RNA are remarkable, with more than 95 different nucleosides presently known in all types of RNA. The discovery of new modified nucleosides as well as increasing knowledge of the array of functional roles of modification, based largely on extensive studies of tRNA (1–3), mandate the need for a comprehensive database of RNA nucleosides. The RNA Modification Database is maintained as an extension of the initial version published in mid-1994 (4), and consists of all RNA-derived ribonucleosides of known structure, including those from established sequence positions (5) as well as those detected or characterized in mixtures from hydrolysates of RNA. The information provided permits access to the modified nucleoside literature both through provision of computer-searchable Chemical Abstracts registry numbers, as well as key literature citations.

The reader is referred to the earlier publication (4) or to the text portion of the present database for comments or discussion of the following topics: usage of modified nucleoside symbols and notations; accuracy of assignments in the early literature; problems associated with modification assignments in rRNA; and exclusion from the database collection of nucleoside degradation products and of erroneous structures or distributions.

During the past year, a listing has been included of the presently known modified nucleosides in *Escherichia coli* 16S and 23S rRNAs and their sequence locations. Also, a table has been added to provide a convenient summary of the phylogenetic distribution

(archaea, bacteria, eukarya) of each of the modified nucleosides found in tRNA. The database has been updated to include the presence of pseudouridine in 23S rRNA from *Haloferax volcanii* (6), the first report of nucleoside modification in archaeal LSU rRNA.

The authors invite comments concerning new entries, errors or omissions, and on the format presently used for electronic access to the database. The E-mail address for this purpose is: rnmods@ecclab.med.utah.edu.

DESCRIPTION OF THE DATABASE

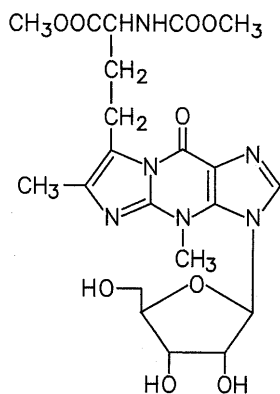
In general, each nucleoside file consists of the following information:

- Type(s) of RNA in which the nucleoside occurs: tRNA (presently consisting of 80 nucleosides), rRNA (28), mRNA (12), snRNA (11), chromosomal RNA (two) and other RNAs (one).
- Phylogenetic occurrence of the RNA in which found: archaea (archaeobacteria), (eu)bacteria, eukarya, and the corresponding literature citations for each. Ribosomal RNA entries are further distinguished by RNA subunit, e.g., 16S, 28S.
- Common chemical names and symbols.
- Chemical Abstracts registry numbers for the ribonucleoside, and corresponding base if assigned.
- Chemical Abstracts index name, which in some cases includes stereochemical information not shown in the database chemical structure.
- Chemical structure of the nucleoside.
- Literature citation to structure assignment of the nucleoside.
- Literature citation to the first reported chemical synthesis of the nucleoside, or in limited cases of the base. Subsequent reports of synthesis, which often include refinements, can be accessed effectively by computer through the Chemical Abstracts registry numbers.
- Comments on any of the above entries, including additional literature citations or alternate nomenclature.

Access to individual nucleoside files within the database can be made through any of the following four entry points: common name; abbreviation symbol; RNA type; and chemical structure (displayed graphically via WWW format). An example of the information contained in a single nucleoside file is shown in Figure 1.

The database currently (October, 1997) contains 95 modified ribonucleoside entries, distributed by RNA type and phylogenetic

* To whom correspondence should be addressed. Tel: +1 801 581 5581; Fax: +1 801 581 7457; Email: pamela.crain.rna.pharm.utah.edu



Symbol: yW
 Common name: wybutosine
 RNA type: tRNA
 Phylogenetic source: eukarya
 Structure: refs. 1, 2
 Synthesis: ref. 2
 CA registry numbers: ribonucleoside 55196-46-8
 base 35693-91-5
 CA index name: 3H-Imidazo[1,2-a]purine-7-butanoic acid, 4,9-dihydro-
 [(methoxycarbonyl)amino]-4,6-dimethyl-9-oxo-3-β-D-
 ribofuranosyl-, methyl ester, (S)-
 Comment: Nomenclature for the yW family of nucleosides is discussed in
 refs. 3 and 4. Errata for structure determination of yW (ref. 1)
 are reported in ref. 2. Alternate commonly used name is "Y
 nucleoside". Commonly used symbol for the corresponding base is
 Y. Structural characterization was made from the acid-liberated
 base. A Chemical Abstracts registry number was then assigned to
 the corresponding ribonucleoside.

References

1. Nakanishi, K., Furutachi, N., Funamizu, M., Grunberger, D. and Weinstein, I.B. (1970) *J. Am. Chem. Soc.*, 92, 7617-7619.
2. Funamizu, M., Terahara, A., Feinberg, A.M. and Nakanishi, K. (1971) *J. Am. Chem. Soc.*, 93, 6706-6708.
3. Blobstein, S.H., Grunberger, D., Weinstein, I.B. and Nakanishi, K. (1973) *Biochemistry*, 12, 188-193.
4. Kasai, H., Goto, M., Ikeda, K., Zama, M., Mizuno, Y., Takemura, S., Matsuura, S., Sugimoto, T. and Goto, T. (1976) *Biochemistry*, 15, 898-904.

Figure 1. Database record for the modified nucleoside wybutosine ('Y nucleoside').

source as shown in Table 1. Users of this database are requested to cite the present paper as the source of information.

Table 1. Phylogenetic distribution of modified ribonucleosides reported in RNA

RNA	Source		
	Archaea	Bacteria	Eukarya
tRNA	36	42	47
rRNA		1 ^a	4 ^a
SSU	3	8	18
LSU	1	14	12
5S	3		1
5.8S			5
mRNA			12
snRNA			11
Chromosomal RNA			2
Other small RNA			1

^aSubunit origin not shown, or is reported in a mixture of SSU and LSU rRNAs.

AVAILABILITY

The nucleoside modification database is currently accessible via the World Wide Web at:
<http://www-medlib.med.utah.edu/RNAmods/RNAmods.html>

ACKNOWLEDGMENTS

We are grateful to D.He, S.C.Pomerantz and N.Lombardo for assistance in implementing the electronic version of the database. Maintenance of the database is supported by NIH grant GM29812.

REFERENCES

1. Björk, G.R., Ericksson, J.U., Gustafsson, C.E.D., Hagervall, T.G., Jönsson, Y.H. and Wikström, P.M. (1987) *Ann. Rev. Biochem.* 56, 263-287.
2. Björk, G.R. (1992) In Hatfield, D.L., Lee, B.J. and Pirtle, R.M. (eds), *Transfer RNA in Protein Synthesis*. CRC Press, Boca Raton, FL, pp. 23-85.
3. Yokoyama, S. and Nishimura, S. (1995) In Söll, D. and RajBhandary, U.L. (eds), *tRNA Structure, Biosynthesis, and Function*. American Society for Microbiology, Washington, DC, pp. 207-223.
4. Limbach, P.A., Crain, P.F. and McCloskey, J.A. (1994) *Nucleic Acids Res.* 22, 2183-2196.
5. Horn, C., Brown, M., Steinberg, S. and Sprinzl, M. (1998) *Nucleic Acids Res.* 26, 148-153.
6. Ofengand, J. and Bakin, A. (1997) *J. Mol. Biol.* 266, 246-268.