ACS chemical biology

Point of VIEW

# The Purine Machine Scores a Base Hit

**T. Joseph Kappock***

Department of Biochemistry, Purdue University, 175 S. University Street, West Lafayette, Indiana 47907

**ABSTRACT** A new synthetic method allows incorporation of $^{13}C$ or $^{15}N$ into selected positions within purine nucleotide bases, starting from simple labeled precursors. The procedure harnesses diverse enzymes to support biosynthesis by the pentose phosphate and *de novo* purine pathways. Selective isotope incorporation should expand the range of RNAs that are amenable to NMR analysis.

*Corresponding author, kappock@purdue.edu.

Published online August 15, 2008
10.1021/cb8001842 CCC: $40.75



Synthesis is a defining skill for those working at the chemistry–biology interface. The ability to prepare small or large molecules (better yet, both) frees an investigator to explore the world beyond the Sigma catalogue. However, synthesis on the scale and purity needed for incisive biomolecular experiments is often daunting. Quick, easy preparative procedures spread rapidly and arguably reach perfection when the key catalyst is a Web site password.

On page 499 of this issue (*1*), Schultheisz *et al*. report another milestone on their way to a more perfect synthesis of purine nucleotides (ATP and GTP) containing stable isotope labels (reviewed in ref 2), which they incorporate into RNA for NMR studies. Purines are biosynthesized by either the salvage pathway, an intact-base installation, or the *de novo* pathway, which is exploited here. The latter accretes the base on a foundation of ribose 5-phosphate, consuming a handful of ATPs and small (amino) acids in at least 10 reactions from phosphoribosylpyrophosphate to inosine monophosphate (*3*). The new synthesis supports ribose and base assembly from comparatively cheap, off-the-shelf labeled precursors. While the required investment in materials and synthesis time appears to be lower than ever, a whopping 28 enzymes are employed. More on that later.

Why bother? NMR studies open a window into the solution structure and dynamics of RNA, but they get tougher as the molecules get larger (*4, 5*). Proton ($^{1}H$) NMR spectra of most RNAs that are large enough to be both structured and interesting have thickets of overlapping peaks, due to the awful chemical shift dispersion in ribose, and other problems. Assigning $^{1}H$, $^{13}C$, $^{15}N$, and $^{31}P$ resonances within individual residues is done by any means necessary, including multiple multidimensional spectral methods and multiple samples with differing isotope compositions. Parallel challenges for the NMR analysis of larger proteins have been successfully addressed by a convergence of new spectral methods (*5*) and selective *in vivo* amino acid deuteration strategies (*6, 7*). The acme of selective isotope labeling of proteins, a method involving chemical-enzymatic synthesis of each residue and protein synthesis by *in vitro* translation (*8*), is not for everyone. In contrast, the *in vivo* deuteration methods developed by the Kay laboratory are compatible with most bacterial overexpression protocols and are thus broadly implemented. All aim to reduce the clutter that is characteristic of large-molecule NMR spectra.

RNA NMR samples are made by *in vitro* transcription of NTPs. Pure labeled NTPs are prepared from degraded rRNAs, which generally gives uniformly labeled products, or synthesized with targeted isotope incorporation using organic or enzymatic reactions (cited in ref 1). Uniformly labeled RNAs do not usually solve the spectral overlap problem. Base synthesis is more laborious but allows incorporation of $^{13}C$ or $^{15}N$ into the right places, yielding less complex spectra and facilitating assignments.

Purines and purine biosynthetic intermediates can also be prepared using enzymes (cited in refs 1 and 9), a route with advantages in selectivity and specificity, but one

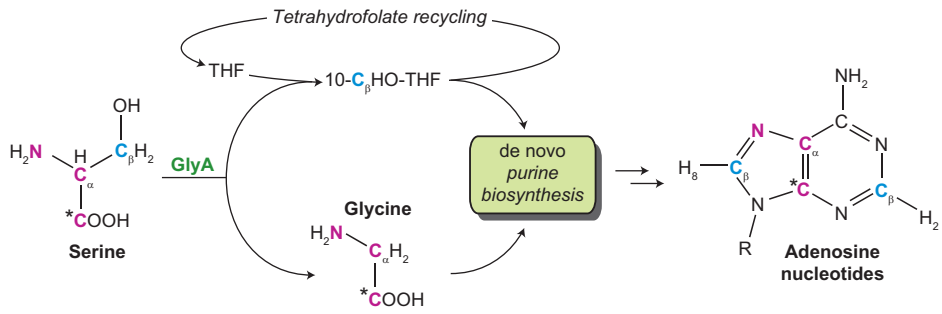


**Figure 1. Atom-economical $^{13}C/^{15}N$ transfer from serine to adenosine nucleotides in an enzyme-dependent synthesis of purines (*1*). Serine hydroxymethyltransferase (GlyA) excises the β-carbon, which is shuttled by 10-CHO-THF to the *de novo* purine biosynthesis pathway, which installs this atom at both the adenine C8 and C2 positions (indicated by the β atom labels). Meanwhile, the remaining non-hydrogen atoms from the glycine byproduct are incorporated as a unit into the imidazole ring. Dual $^{13}C2,^{13}C8$ labeling can be used to assign H2 and H8 resonances (indicated by subscripts). Different labeling patterns are accessible by switching the labeled precursor(s). R represents ribose 5′-phosphate.**

that requires separation of the nucleotide product(s) from the nucleotide co-substrates for the enzymes used. The biosynthesis of labeled ATP has the additional challenge that isotope dilution from co-substrates must be avoided. It would certainly seem easier to engineer some useful bug to do all this work for us, as with partial amino acid deuteration. However, living cells actively maintain (deoxy)nucleotide pools of the appropriate size and composition, so it is difficult to envision how NMP hyperproduction in an engineered strain could be achieved. On the other hand, it is tricky and costly to use and regenerate enzyme cofactors *in vitro*.

A compromise might be to mix chemical and enzymatic synthesis to complete the purine base from a synthetic compound incorporating the desired labels. However, this requires synthesis of a *de novo* purine metabolite. Synthetic challenges for these compounds include the acid lability of imidazole, purine nucleotides, and the β-anomeric configuration; the chemical lability of the bases of several intermediate metabolites; and most vexing, the requirement for 5′-phosphorylation. Some metabolites are only accessible by enzyme-mediated syntheses, like the final acyclic metabolite in the *de novo* pathway, formylglycinamidine ribonucleotide.

The Williamson group has developed chemical-enzymatic routes to RNAs with selected $^{13}C$ labels or partial deuteration that give improved NMR spectra (*2, 10*). Their new route to purine bases with defined $^{13}C/^{15}N$ incorporation should further improve matters. As an example, purines contain nonexchangeable protons at the opposite ends of each base: H8 (adenine and guanine) and H2 (adenine). Correlating these signals is an important step in assigning adenine spin systems (*11*). Starting from a single labeled compound, Schultheisz *et al.* synthesize adenine with $^{13}C$ attached only to H8 and H2. Isotope-edited spectra of RNAs containing these compounds should allow a better view of interactions made by these critical protons.

The new methods produce the desired output with the minimum input of small molecules and are perhaps more interesting and instructive than the products obtained. Familiar linear reaction sequences are framed in generally applicable concepts: for example, consumable compounds are categorized as fuels, which drive accessory cofactor-regenerating "loops", or substrates, which are incorporated into the products. These loops provide nearly all needed cosubstrates for pentose phosphate and *de novo* purine biosynthesis enzymes. Some are straightforward, like the folate loop (Figure 1), which supports *in situ* generation of $^{13}C$-labeled $N^{10}$-formyl-tetrahydrofolate (10-CHO-THF), a substrate for the PurN and PurH formyltransferases that is not commercially available. When used with GlyA/FolD, a remarkably efficient conversion of serine to labeled purines is achieved (Figure 1). Other loops are more networked. By its conversion to glutamate, the fuel α-ketoglutarate supports the recycling of reduced nicotinamide coenzymes produced during production of GMP and ribulose 5-phosphate and by the folate-recycling loop. Meanwhile, the glutamate product is converted into glutamine, incorporating $^{15}N$ label that is delivered by three amidotransferase reactions, in a process that consumes another fuel, creatine phosphate. There are also other pleasing touches, like the use of dATP as an enzyme cofactor. dATP is easily separated from labeled product NTPs using a diol-binding boronate column, thereby avoiding the isotope dilution problem. Here, Schultheisz *et al.* orchestrate small innovations and previously reported procedures in a masterly way.

Despite the elegance of its biochemical logic, the new method is at its heart a hammer-and-tongs way to get material. Other routes are just harder to implement, particularly in laboratories that are not equipped with research-grade chemical synthesis and analysis facilities. As with partial protein deuteration, ease of implementation is a key virtue. There is only one catch, and it is a doozy: acquiring 28 enzymes, only a few of which are commercially available. Schultheisz *et al.* use His-tagged constructs and a generic purification method to obtain

most of the rest, including the entire *de novo* purine biosynthetic pathway. Since purine metabolites are a headache to synthesize as substrates for enzyme assays, arbitrary specific activities are assigned to most *de novo* purine enzymes, saving a great deal of time but injecting ambiguity into the troubleshooting of reaction conditions. Little optimization of reaction conditions is reported. A need for some retooling is clear from the eventual loss of product formation as one or more of the reaction cocktail components fails over time. If the many new gene constructs described here are readily disseminated, communal tweaking and improvement of the reaction conditions should naturally follow. In any case, one hopes students everywhere will be spared the soul-killing task of rebuilding all of those clones!

The new methodology is open and adaptable to the synthesis of other labeled purines or purine precursors for NMR studies or other tasks like studies of purine, thiamine, or cofactor biosynthesis. A simple example is a synthesis of purines labeled at C6, which derives uniquely from bicarbonate. One could also exploit the detours in the "classical" 10-step *de novo* pathway. The discovery of the role of bacterial PurK and the differentiation of animal and bacterial PurE functions defined the first fork in the *de novo* pathway (*12, 13*). Other pathway variants are associated with the enzymes PurT, PurO, and PurP (*14–16*). In the current method, label for C2 and C8 comes from the same source (Figure 1). Since the precursor of GTP C8 can be also be installed using the "variant" formate-dependent PurT GAR formyltransferase instead of the "classical" folate-dependent PurN GAR formyltransferase, one can place different labels in C8 and C2 by selection of appropriate formate and serine precursors. Labeled DNA precursors, which one might want to study DNA damage, could be made with ribonucleotide reductase.

One suspects most RNA NMR specialists will be content to make the labeled purines described in this paper, although some may tweak the recipe a bit. By their bold and successful use of a couple dozen enzymes, Williamson and co-workers have encouraged other natural product cooks to consider new ways to deploy enzymes in their recipes.

## REFERENCES


1. Schultheisz, H. L., Szymczyna, B. R., Scott, L. G., and Williamson, J. R. (2008) Pathway engineered enzymatic *de novo* purine nucleotide synthesis, *ACS Chem. Biol. 3*, 499–511.
2. Scott, L. G., Tolbert, T. J., and Williamson, J. R. (2000) Preparation of specifically $^{2}$H- and $^{13}$C-labeled ribonucleotides, *Methods Enzymol. 317*, 18–38.
3. Kappock, T. J., Ealick, S. E., and Stubbe, J. (2000) Modular evolution of the purine biosynthetic pathway, *Curr. Opin. Chem. Biol. 4*, 567–572.
4. Latham, M. P., Brown, D. J., McCallum, S. A., and Pardi, A. (2005) NMR methods for studying the structure and dynamics of RNA, *ChemBioChem 6*, 1492–1505.
5. Tzakos, A. G., Grace, C. R., Lukavsky, P. J., and Riek, R. (2006) NMR techniques for very large proteins and RNAs in solution, *Annu. Rev. Biophys. Biomol. Struct. 35*, 319–342.
6. Gardner, K. H., and Kay, L. E. (1998) The use of $^{2}$H, $^{13}$C, $^{15}$N multidimensional NMR to study the structure and dynamics of proteins, *Annu. Rev. Biophys. Biomol. Struct. 27*, 357–406.
7. Goto, N. K., Gardner, K. H., Mueller, G. A., Willis, R. C., and Kay, L. E. (1999) A robust and cost-effective method for the production of Val, Leu, Ile (δ1) methyl-protonated $^{15}$N-, $^{13}$C-, $^{2}$H-labeled proteins, *J. Biomol. NMR 13*, 369–374.
8. Kainosho, M., Torizawa, T., Iwashita, Y., Terauchi, T., Mei Ono, A., and Güntert, P. (2006) Optimal isotope labelling for NMR protein structure determinations, *Nature 440*, 52–57.
9. Meyer, E., Kappock, T. J., Osuji, C., and Stubbe, J. (1999) Evidence for the direct transfer of the carboxylate of $N^5$-carboxyaminoimidazole ribonucleotide ($N^5$-CAIR) to generate 4-carboxy-5-aminoimidazole ribonucleotide catalyzed by *Escherichia coli* PurE, an $N^5$-CAIR mutase, *Biochemistry 38*, 3012–3018.
10. Vallurupalli, P., Scott, L., Hennig, M., Williamson, J. R., and Kay, L. E. (2006) New RNA labeling methods offer dramatic sensitivity enhancements in $^{2}$H NMR relaxation spectra, *J. Am. Chem. Soc. 128*, 9346–9347.
11. Marino, J. P., Prestegard, J. H., and Crothers, D. M. (1994) Correlation of adenine H2/H8 resonances in uniformly $^{13}$C labeled RNAs by 2D HCCH-TOCSY: a new tool for $^{1}$H assignment, *J. Am. Chem. Soc. 116*, 2005–2006.
12. Mueller, E. J., Meyer, E., Rudolph, J., Davisson, V. J., and Stubbe, J. (1994) $N^5$-carboxyaminoimidazole ribonucleotide: evidence for a new intermediate and two new enzymatic activities in the de novo purine biosynthetic pathway of *Escherichia coli*, *Biochemistry 33*, 2269–2278.
13. Firestine, S. M., Poon, S. W., Mueller, E. J., Stubbe, J., and Davisson, V. J. (1994) Reactions catalyzed by 5-aminoimidazole ribonucleotide carboxylases from *Escherichia coli* and *Gallus gallus*: a case for divergent catalytic mechanisms, *Biochemistry 33*, 11927–11934.
14. Marolewski, A., Smith, J. M., and Benkovic, S. J. (1994) Cloning and characterization of a new purine biosynthetic enzyme: a non-folate glycinamide ribonucleotide transformylase from *E. coli*, *Biochemistry 33*, 2531–2537.
15. Graupner, M., Xu, H., and White, R. H. (2002) New class of IMP cyclohydrolases in *Methanococcus jannaschii*, *J. Bacteriol. 184*, 1471–1473.
16. Ownby, K., Xu, H., and White, R. H. (2005) A *Methanocaldococcus jannaschii* archaeal signature gene encodes for a 5-formaminoimidazole-4-carboxamide-1-β-D-ribofuranosyl 5′-monophosphate synthetase. A new enzyme in purine biosynthesis, *J. Biol. Chem. 280*, 10881–10887.