

Review

# Isotope-Labeled RNA Building Blocks for NMR Structure and Dynamics Studies

Lukasz T. Oleginski <sup>1,†</sup>, Kehinde M. Taiwo <sup>1,†</sup> , Regan M. LeBlanc <sup>1,2</sup> and Theodore K. Dayie <sup>1,\*</sup>

<sup>1</sup> Center for Biomolecular Structure and Organization, Department of Chemistry and Biochemistry, University of Maryland, College Park, MD 20742, USA; lolengin@umd.edu (L.T.O.); ktaiwo@umd.edu (K.M.T.); Regan\_LeBlanc@vrtx.com (R.M.L.)

<sup>2</sup> Vertex Pharmaceuticals, 50 Northern Avenue, Boston, MA 02210, USA

\* Correspondence: dayie@umd.edu; Tel.: +1-301-405-3165

† These authors have contributed equally to this work.

**Abstract:** RNA structural research lags behind that of proteins, preventing a robust understanding of RNA functions. NMR spectroscopy is an apt technique for probing the structures and dynamics of RNA molecules in solution at atomic resolution. Still, RNA analysis by NMR suffers from spectral overlap and line broadening, both of which worsen for larger RNAs. Incorporation of stable isotope labels into RNA has provided several solutions to these challenges. In this review, we summarize the benefits and limitations of various methods used to obtain isotope-labeled RNA building blocks and how they are used to prepare isotope-labeled RNA for NMR structure and dynamics studies.

**Keywords:** stable isotope; enzymes; nucleotides; NMR spectroscopy; RNA; structure; dynamics



**Citation:** Oleginski, L.T.; Taiwo, K.M.; LeBlanc, R.M.; Dayie, T.K. Isotope-Labeled RNA Building Blocks for NMR Structure and Dynamics Studies. *Molecules* **2021**, *26*, 5581. <https://doi.org/10.3390/molecules26185581>

Academic Editor: Esteban Alonso

Received: 25 July 2021

Accepted: 7 September 2021

Published: 14 September 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

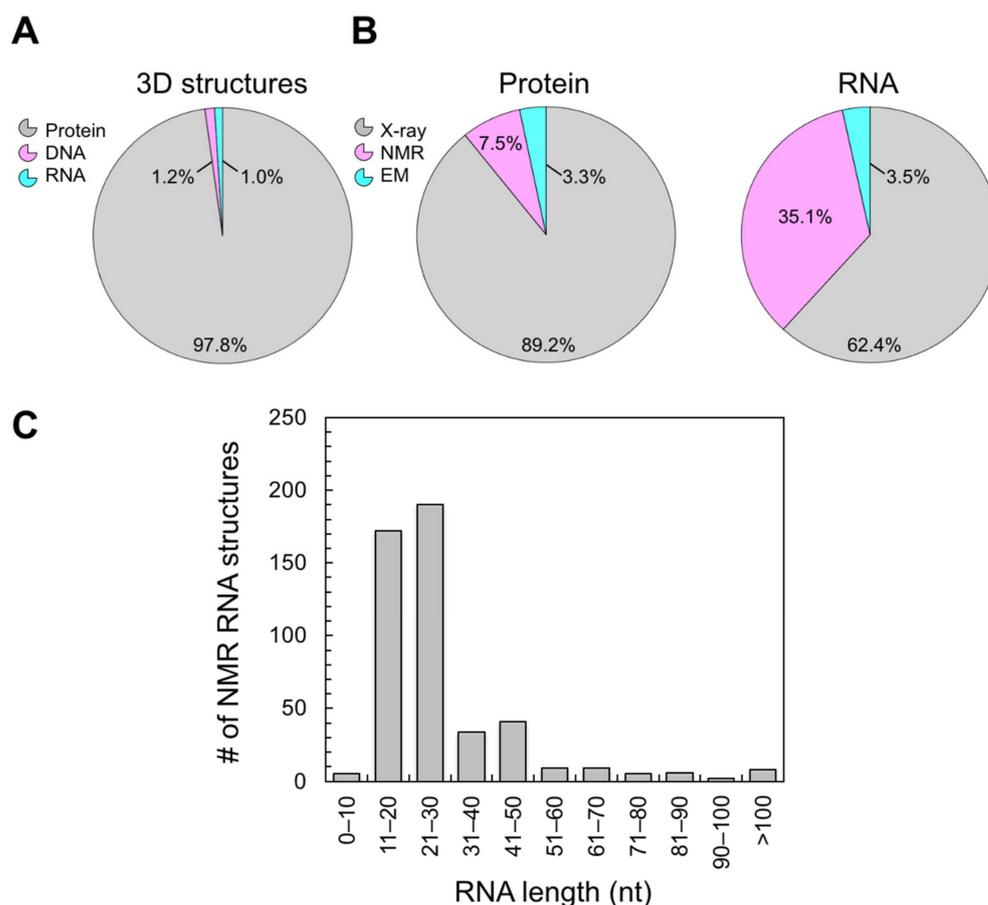


**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

RNA is a dynamic macromolecule that has many biological functions, including gene regulation [1–5], catalysis [6–8], structural organization [9,10], and viral replication [11,12]. Almost without exception, RNA's intricate three-dimensional (3D) structure and conformational plasticity are required to carry out these functions [13,14]. A robust understanding of RNA function therefore requires high-resolution structure and dynamics data. Unfortunately, there is a scarcity of RNA 3D structures as compared to those of proteins. There are only ~1500 RNA structures in the Nucleic Acid Database (NDB), whereas more than 150,000 protein structures are in the Protein Database (PDB) (Figure 1A). This observation can, in part, be explained by RNA's dynamic nature, which can impede crystallization and complicate cryo-electron microscopy (cryo-EM) data analysis. While RNA's conformational heterogeneity hinders X-ray crystallography and cryo-EM analysis, it is compatible with solution nuclear magnetic resonance (NMR) spectroscopy. Recent advances in NMR are even capable of probing low populated states that would otherwise be undetectable by traditional biophysical methods [15–17].

NMR is a competitive RNA structure determination technique that accounts for 35% of the structures in the NDB, as compared to the 8% of protein structures in the PDB (Figure 1B). However, obstacles remain. Specifically, NMR analysis of RNA suffers from spectral crowding and broad linewidths [18–24]. The former is a result of the limited chemical diversity of the RNA building blocks adenosine (A), guanosine (G), cytidine (C), and uridine (U), and the narrow chemical shift dispersion of ribose protons other than H1' (i.e., H2', H3', H4', H5', and H5''). Both conditions worsen in larger RNAs and limit our understanding of their structures and functions. In fact, only 23 RNA structures >60 nucleotides (nt) have been solved by solution NMR (some requiring additional methodologies, e.g., cryo-EM) (Figure 1C). Advancements in stable isotope (e.g., <sup>2</sup>H, <sup>13</sup>C, <sup>15</sup>N, and <sup>19</sup>F) labeling of RNA have revolutionized the analysis of RNA by NMR and opened the door to functional understanding.



**Figure 1.** Structural biology statistics. The distribution of high-resolution 3D structures deposited in the PDB/NDB separated by (A) macromolecule type; and (B) protein and RNA structures separated by structure determination technique. (C) A histogram of RNA NMR structures in the NDB sorted by RNA length. Statistics were accessed in July 2021.

The main approaches to obtaining isotope-labeled RNA are enzymatic or chemical syntheses. For the enzymatic approach, almost all methods are based on DNA template-directed T7 RNA polymerase (RNAP)-based *in vitro* transcription (IVT) using labeled ribonucleoside 5'-triphosphates (rNTPs) [22,23,25–31]. The alternative method is chemical solid-phase synthesis (SPS) using RNA phosphoramidites (amidites) [32–36]. Both approaches can use unlabeled and isotope-labeled building blocks (rNTPs and amidites) to generate versatile RNA labeling patterns. The four strategies to obtain such building blocks for enzymatic RNA synthesis are: (1) purchase commercially available isotope-labeled rNTPs; (2) use simple organisms to incorporate isotope-labeled precursors into their rNTPs; (3) complete *de novo* biosynthesis of rNTPs; (4) utilize a hybrid chemo-enzymatic approach that combines chemical syntheses of ribose and nucleobases and their enzymatic coupling to prepare rNTPs.

In this review, we detail the benefits and limitations of these four methods (Section 2). With labeled building blocks in-hand, we discuss the various ways of using them to make isotope-labeled RNA for NMR structure and dynamics studies (Section 3). Finally, we analyze isotope labeling in the context of RNA structural biology and comment on where we think the field is headed (Section 4).

## 2. Stable Isotope Labeling of RNA Building Blocks

Throughout this review, we follow the IUPAC/IUB guidelines for RNA atom numbering [37,38]. When describing RNA labeling, we use four categories: uniform, nucleotide-specific, atom-specific, and position-specific. Uniform labeling is when every atom of

a certain type (e.g.,  $^2\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ , or both of the latter) is enriched; nucleotide-specific labeling is when every nucleotide of a certain type (e.g., all uridines) is enriched; atom-specific labeling is when every atom of a certain type (e.g., uridine C6) is enriched; and position-specific labeling is when an individual nucleotide (e.g., uridine 7) is labeled. In the latter case, the type of label that is incorporated site-specifically can be uniformly or atom-specifically-labeled. Thus, these labeling categories are not mutually exclusive.

### 2.1. Commercial Isotopes Sources

The simplest approach to obtaining isotope-labeled rNTPs and amidites is to purchase the desired  $^2\text{H}$ ,  $^{13}\text{C}$ , and/or  $^{15}\text{N}$  isotope labels from a commercial source. As of July 2021, Cambridge Isotope Laboratories (CIL), Sigma-Aldrich, Cassia LLC, Silantes, and INNotope are the major suppliers of isotope-labeled rNTPs, whereas Silantes and INNotope are the only suppliers of isotope-labeled RNA amidites. Unfortunately, these products can be prohibitively expensive. Uniformly-labeled  $^2\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ,  $^2\text{H}/^{15}\text{N}$ , and  $^{13}\text{C}/^{15}\text{N}$ -labeled and atom-specifically  $^2\text{H}$ ,  $^{13}\text{C}$ ,  $^2\text{H}/^{13}\text{C}$ -labeled rNTPs are available for \$800–5600 per 100  $\mu\text{mol}$  or 50 mg (Table 1). Additionally, uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$ , and  $^{13}\text{C}/^{15}\text{N}$ -labeled and atom-specifically  $^2\text{H}$ ,  $^{13}\text{C}$ ,  $^2\text{H}/^{13}\text{C}$ , and  $^{15}\text{N}$ -labeled amidites are available for \$900–6600 per 50 mg (Table 1). For reference, IVT (20 mL) typically requires 250–1000  $\mu\text{L}$  (100  $\mu\text{mol}$  stock) per rNTP and yields 0.2–2.0 mM in 300  $\mu\text{L}$  of RNA (Section 3.2). SPS (1  $\mu\text{mol}$ ) generally requires 10–20 mg (0.1 M stock) per amidite coupling and yields 0.2–0.6 mM in 300  $\mu\text{L}$  of RNA (Section 3.1).

These considerations underscore a sobering fact: the approximate cost per NMR sample when using commercial building blocks often exceeds \$1000. Since it takes multiple samples for complete RNA resonance assignment, total costs of robust RNA analysis by NMR can easily reach \$10,000 based on isotope labeling costs alone. For example, a recently determined structure of the 43 nt SAM/SAH-binding riboswitch [39,40] required 7 uniform and nucleotide-specifically-labeled samples by IVT and 13 atom- and position-specifically-labeled samples for SPS. This financial burden partly explains the slow rate of RNA structure depositions (Figure 1A). It is therefore crucial to reduce the costs of obtaining isotope-labeled RNA and to expand the accessibility of NMR analysis of RNA.

**Table 1.** Price of commercial isotope-labeled rNTPs and RNA amidites.

Building Block	Price <sup>a</sup> (\$)	Supplier <sup>b</sup>
<b>Uniformly <math>^2\text{H}</math>-labeled rNTPs</b>		
rNTP (N = A, C, G, or U)	1300	Silantes
<b>Selectively <math>^2\text{H}</math>-labeled rNTPs</b>		
[3',4',5',5''- $^2\text{H}_4$ ]-X (X = ATP or GTP)	800	CIL
[2- $^2\text{H}$ ]-ATP	1200	CIL
[5,1',2',3',4',5',5''- $^2\text{H}_7$ ]-CTP	1800	CIL
[5,3',4',5',5''- $^2\text{H}_5$ ]-Y- (CTP or UTP)	800	CIL
[5,6- $^2\text{H}_2$ ]-CTP	1800	CIL
[5,1',2',3',4',5',5''- $^2\text{H}_7$ ]-UTP	1400	CIL
<b>Uniformly <math>^{13}\text{C}</math>-labeled rNTPs and Amidites</b>		
rNTP (N = A, C, G, or U)	1400	Silantes
N (N = A, C, G, or U Amidites)	6600	Silantes
<b>Selectively <math>^{13}\text{C}</math>-labeled rNTPs and Amidites</b>		
[8- $^{13}\text{C}$ ]-X (X = ATP or GTP)	1400	Silantes
[8- $^{13}\text{C}$ ]-A	900	Silantes
[2,8- $^{13}\text{C}_2$ ]-A	2700	Silantes
[8- $^{13}\text{C}$ ]-G	1000	INNotope

Table 1. Cont.

Building Block	Price <sup>a</sup> (\$)	Supplier <sup>b</sup>
<b>Selectively <sup>2</sup>H/<sup>13</sup>C-labeled rNTPs and Amidites</b>		
[6- <sup>13</sup> C-5- <sup>2</sup> H]- Y (Y = CTP or UTP)	1600	Silantes
[6- <sup>13</sup> C-5- <sup>2</sup> H]-C or U	1000	INNotope
<b>Uniformly <sup>15</sup>N-labeled rNTPs and Amidites</b>		
rNTP (N = A, C, G, or U)	900	CIL
A, or G, or C, or U	1400	Silantes
<b>Selectively <sup>15</sup>N-labeled Amidites</b>		
[1- <sup>15</sup> N]-A	1000	INNotope
[1- <sup>15</sup> N]-G	1100	INNotope
[3- <sup>15</sup> N]-C- or U	1000	Silantes
[1,3- <sup>15</sup> N <sub>2</sub> ]-C	1200	Silantes
[1,3,4- <sup>15</sup> N <sub>3</sub> ]-C	1000	INNotope
[1,3- <sup>15</sup> N <sub>2</sub> ]-U	1000	INNotope
<b>Uniformly <sup>2</sup>H/<sup>15</sup>N-labeled rNTPs</b>		
rNTP (N = A, C, G, or U)	5600	Silantes
<b>Uniformly <sup>13</sup>C/<sup>15</sup>N-labeled rNTPs and Amidites</b>		
rNTP (N = A, C, G, or U)	1100	CIL
N (N = A, C, G, or U Amidites)	5300	Silantes

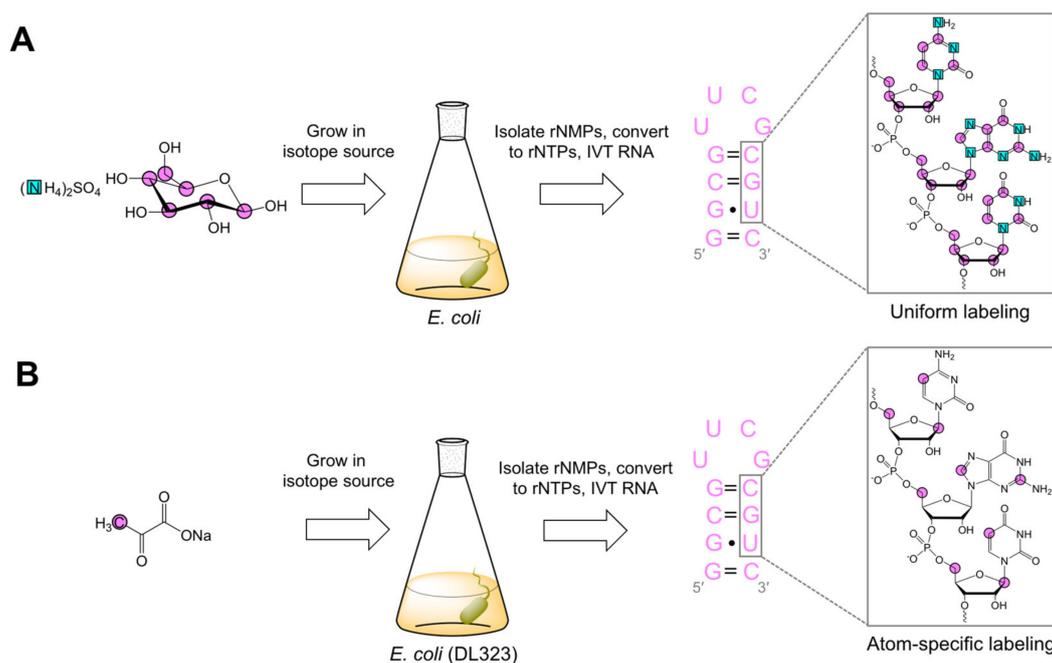
<sup>a</sup> Price of rNTPs (CIL, per 100 μmol or Silantes, per 50 mg) and amidites (Silantes or INNotope, per 50 mg) were accessed from their respective websites in July 2021. <sup>b</sup> While some of the labeled material is available from multiple suppliers, all prices reflect the cheapest available and are rounded up to the nearest \$100.

## 2.2. Biomass Labeling

Biomass labeling incorporates isotope-labeled building blocks into simple organisms' RNA. This labeling approach was established by the Pardi [22] and Williamson [23] research groups. In general, their method includes growing organisms on <sup>13</sup>C and/or <sup>15</sup>N source(s), harvesting cells and extracting RNA, hydrolyzing RNA to rNMPs, and converting those to rNTPs. Triphosphate conversion can be achieved by chemical [41] or enzymatic [42] means, depending on the expertise and resources available. The latter method is usually superior to the chemical approach, yielding rNTPs of >95% purity [22,23,41–43]. Although biomass methods permit new and commercially unavailable rNTP labeling patterns, the overall cost advantage is minimal, and the purification steps are laborious. Nevertheless, many research groups have utilized biomass labeling to prepare RNA for NMR analysis [22,23,41–45].

### 2.2.1. Biomass Uniform Labeling

Uniform <sup>13</sup>C labeling was first achieved by growing *E. coli* [22,23], *M. methylotrophus* [22], or *M. extorquens* [23] with either <sup>13</sup>C-glucose or <sup>13</sup>C-methanol. For uniform <sup>13</sup>C/<sup>15</sup>N labeling, *E. coli* was grown with <sup>13</sup>C-glucose and <sup>15</sup>N-ammonium sulfate (Figure 2A) [22,23]. The use of *M. methylotrophus* and *M. extorquens* gained popularity due to their compatibility with the more cost-effective <sup>13</sup>C-methanol [22]. Nevertheless, these organisms have significantly lower rNTP contents and more difficult growth conditions as compared to *E. coli* [22]. The rNTPs obtained from this method were used in IVT to make uniformly and nucleotide-specifically-labeled RNAs for multi-dimensional NMR experiments [22,23,43,46]. These new experiments greatly simplified resonance assignment strategies and the structure determination of small (<30 nts) RNAs. However, alternative labeling strategies were needed to overcome spectral crowding in larger RNAs.



**Figure 2.** Schematic representations of biomass labeling adapted from the Pardi [22] and Williamson [23] research groups, and our previous work [44]. **(A)** *E. coli* grown in  $^{13}\text{C}$ -glucose and  $^{15}\text{N}$ -ammonium sulfate yields uniformly  $^{13}\text{C}/^{15}\text{N}$ -labeled rNMPs, which can be phosphorylated, forming rNTPs, and then used in IVT to make uniformly (or nucleotide-specifically) labeled RNA. **(B)** *E. coli* (DL323) grown in  $[3\text{-}^{13}\text{C}]$ -pyruvate affords rNTPs, which can then make atom-specifically-labeled RNA. Magenta circles and cyan squares represent  $^{13}\text{C}$  and  $^{15}\text{N}$ , respectively. Additional details can be found in the original works [22,23,44].

### 2.2.2. Biomass Atom-Specific Labeling

Hoffman and Holland modified the biomass method to make atom-specifically-labeled rNTPs [47]. In their approach, *E. coli* were grown with different  $^{13}\text{C}$ -sodium acetate sources to make various  $^{13}\text{C}$ -labeled rNTPs. For example,  $[2\text{-}^{13}\text{C}]$ -sodium acetate labeled purine (A and G) C2, C5, and C8 (>95%); pyrimidine (U and C) C5 and C6 (>90%); and ribose C1', C4', and C5' (~90%). In addition, they made  $[1\text{-}^{13}\text{C}]$ -sodium acetate labeled purine C4 and C6 (>90%), pyrimidine C2 and C4 (>95%), and ribose C3' (~75%) [47]. Among others, Hoogstraten and co-workers employed a similar methodology wherein *E. coli* strains deficient in enzymes involved in the tricarboxylic acid cycle (DL323) [45] or the oxidative pentose phosphate pathway (K10-1516) [48] were grown with various  $^{13}\text{C}$ -glycerol sources [49]. While most labeling patterns had yields <50%,  $[2',4'\text{-}^{13}\text{C}_2]$ -AMP was created with an 80% yield in K10-1516 cells grown with  $[2\text{-}^{13}\text{C}]$ -glycerol. Importantly, Hoogstraten and co-workers measured longitudinal ( $R_1$ ) and rotating-frame ( $R_{1\rho}$ ) relaxation rates in  $[2',4'\text{-}^{13}\text{C}_2]$ -AMP and uniformly  $^{13}\text{C}$ -labeled AMP to demonstrate that  $^{13}\text{C}$ - $^{13}\text{C}$  dipolar couplings lead to over-estimated relaxation rates [49].

Our research group used a similar method but with a different isotope source [44]. Purine C2 and C8 (~95%), pyrimidine C5 (~98%), and ribose C1' (42%) and C5' (95%) were labeled in DL323 cells fed with  $[3\text{-}^{13}\text{C}]$ -pyruvate (Figure 2B). To demonstrate the utility of these labeling patterns, we used our atom-specifically-labeled rNTPs to make a 27 nt RNA via IVT for NMR analysis. In agreement with previous work,  $R_1$  rate measurements showed a discrepancy between uniformly and atom-specifically-labeled samples for pyrimidine C5 and ribose C1' and C5' [44,49].

### 2.3. Ribonucleotide De Novo Biosynthesis

A ribonucleotide de novo biosynthesis uses enzymes from the pentose phosphate and nucleotide salvage biosynthetic pathways (Table 2), various cofactor regeneration systems, and isotope-labeled precursor compounds to synthesize purine [27] and pyrimidine [28]

rNTPs in a one-pot enzymatic reaction. The benefits of this route include reduced reaction time and increased product yield and specificity, compared with traditional chemical synthesis [50–54]. Moreover, this approach produces cost-effective uniformly  $^{13}\text{C}/^{15}\text{N}$ -labeled and atom-specifically-labeled rNTPs. On the other hand, laborious protein cloning is involved, but the pay offs in preparation and purification more than make up for this initial outlay. Still, de novo labeling to prepare RNA for NMR studies has been used with some success [27,28].

### 2.3.1. Purine De Novo Biosynthesis

Williamson and co-workers were the first to describe the de novo biosynthesis of isotope-labeled purine rNTPs [27]. Their approach used enzymes from the pentose phosphate pathway to convert glucose to 5-phospho-D-ribose-1-pyrophosphate (PRPP) and enter a linear cascade of reactions to assemble the purine ring and produce inosine monophosphate (IMP), a precursor for both ATP and GTP. The de novo biosynthesis of ATP and GTP also required NAD(P)H and rNTP regeneration systems, folate, aspartate, and glutamine. Isotope-labeled precursor compounds  $^{13}\text{C}$ -D-glucose,  $^{13}\text{C}$ -sodium bicarbonate ( $\text{NaH}^{13}\text{CO}_3$ ),  $^{15}\text{N}$ -ammonium chloride ( $^{15}\text{NH}_4\text{Cl}$ ), and  $^{13}\text{C}/^{15}\text{N}$ -L-serine enabled atom-specific labeling. As shown in Figure 3A, purine labeling is provided by these precursor compounds. Specifically, N1, N2, N3, N6, and N9 are derived from  $^{15}\text{NH}_4\text{Cl}$ . Similarly,  $^{13}\text{C}/^{15}\text{N}$ -L-serine labels C2, C4, C5, C8, and N7, and  $\text{NaH}^{13}\text{CO}$  labels C6. Finally,  $^{13}\text{C}$ -D-glucose provides the label for the ribose carbons (C1', C2', C3', C4', and C5') and C6.

While this methodology is very powerful, it also comes with a number of drawbacks. The use of labeled glucose can limit the potential labeling patterns if both ribose and nucleobase labels are desired. For example, production of  $\text{CO}_2$  via decarboxylation of 6-phosphogluconate to ribulose-5-phosphate during PRPP production links the isotope label of the C1 of glucose to C6 in the purine nucleobase. If ribose labeling is not required, and only purine C6 labeling is needed, then PRPP must be made directly from unlabeled ribose. Similarly, care must be taken to prevent isotopic dilution from atmospheric  $\text{CO}_2$ , if both  $^{13}\text{C}$ -ribose and  $^{13}\text{C}$ -purine labeling are required. Additionally, the C2 and C8 positions are labeled together or not at all. Finally, C6 and ribose labeling are limited by commercial sources of D-glucose.

In summary, a total of 28 biosynthetic enzymes (Table 2) were used for the efficient one-pot synthesis of ATP and GTP over two days, producing yields of up to 66% [27]. Specifically, 23 enzymes were used to synthesize  $[2,8-^{13}\text{C}_2]$ -ATP in a 57% yield and 26, 24, and 27 enzymes helped synthesize uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}/^{15}\text{N}$ -labeled GTP in 66, 24, and 42% yields, respectively [27]. Their atom-specific rNTPs were used in IVT to make a 30 nt RNA for NMR analysis. Their labeling patterns helped identify specific nucleobase interactions and greatly reduced spectral crowding.

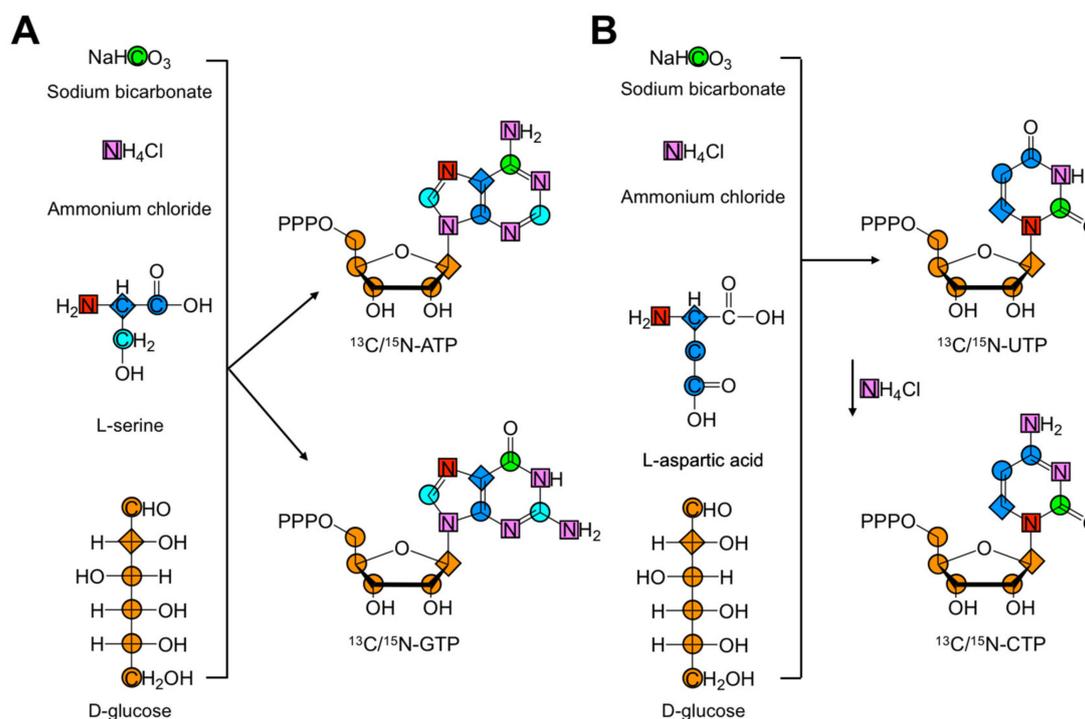
**Table 2.** Enzymes of the purine de novo biosynthetic pathway.

Enzyme <sup>a</sup>	Gene	EC Number	PDB ID	Organism <sup>b</sup>
Hexokinase	hvk1/2	2.7.1.1	1HKG	<i>Saccharomyces cerevisiae</i>
Glucokinase	glk	2.7.1.2	1Q18	<i>Escherichia coli</i>
Glucose-6-phosphate dehydrogenase	zwf1	1.1.1.49	2BHL	<i>Homo sapiens</i>
Phosphogluconate dehydrogenase	gndA	1.1.1.44	2ZYA	<i>Escherichia coli</i> K-12
Ribose-5-phosphate isomerase	rpiA	5.3.1.6	1O8B	<i>Escherichia coli</i>
Ribose-phosphate diphosphate kinase	prsA	2.7.6.1	3Q89	<i>Staphylococcus aureus</i>
Amido phosphoribosyl-transferase	purF	2.4.2.14	IECF	<i>Escherichia coli</i>
Phosphoribosylamine-glycine ligase	purD	6.3.4.13	5VEV	<i>Neisseria gonorrhoea</i>
Phosphoribosylglycinamide formyltransferase	purN	2.1.2.2	3P9X	<i>Bacillus halodurans</i>
Phosphoribosylformylglycinamide synthase	purL	6.3.5.3	1VK3	<i>Thermotoga maritima</i>
Phosphoribosylformylglycinamide cyclo-ligase	purM	6.3.3.1	5VK4	<i>Neisseria gonorrhoea</i>
Phosphoribosylamino-imidazole carboxylase (catalytic subunit)	purE	4.1.1.21	4GRD	<i>Burkholderia cenocepacia</i>

Table 2. Cont.

Enzyme <sup>a</sup>	Gene	EC Number	PDB ID	Organism <sup>b</sup>
Phosphoribosylamino-imidazole carboxylase (ATPase subunit)	purK	4.1.1.21	2Z04	<i>Aquifex aeolicus</i>
Phosphoribosylamino-imidazole-succinocarboxamide synthase	purC	6.3.2.6	3NUA	<i>Clostridium perfringens</i>
Adenylosuccinate lyase	purB	4.3.2.2	3GZH	<i>Escherichia coli</i>
Phosphoribosylamino-imidazole-carboxamide formyltransferase	purH	2.1.2.3	1ZCZ	<i>Thermotoga maritima</i>
Inosine-monophosphate cyclohydrolase	purH	3.5.4.10	2IU0	<i>Gallus gallus</i>
Adenylosuccinate synthase	purA	6.3.4.4	2J91	<i>Homo sapiens</i>
Inosine-monophosphate dehydrogenase	guaB	1.1.1.205	1B30	<i>Homo sapiens</i>
Guanosine-monophosphate synthase	guaA	6.3.5.2	2YWB	<i>Thermus thermophilus</i>
Adenylate kinase	plsA	2.7.4.3	1E4Y	<i>Escherichia coli</i>
Creatine phosphokinase	ckmT	2.7.3.2	2CRK	<i>Oryctolagus cuniculus</i>
Guanylate kinase	spoR	2.7.4.8	2ANC	<i>Escherichia coli</i>
Glycine hydroxymethyltransferase	glyA	2.1.2.1	5VMB	<i>Acinetobacter baumannii</i>
Methylene-tetrahydrofolate dehydrogenase	folD	1.5.1.5	1B0A	<i>Escherichia coli</i> K-12
Methenyl-tetrahydrofolate cyclohydrolase	folD	3.5.4.9	5TCA	<i>Homo sapiens</i>
Aspartate ammonia-lyase	aspA	4.3.1.1	1JSW	<i>Escherichia coli</i>
Glutamate dehydrogenase (NAD(P)+)	glud1/ghA	1.4.1.3	4fcc	<i>Escherichia coli</i>
Glutamine synthetase	glnA	6.3.1.2	4IS4	<i>Medicago truncatula</i>
Inorganic diphosphatase	ppa	3.6.1.1	1IPW	<i>Escherichia coli</i>

<sup>a</sup> All enzymes are commercially available except phosphoribosylformylglycinamide cyclo-ligase, phosphoribosylamino-imidazole carboxylase, phosphoribosylamino-imidazole-carboxamide formyltransferase, inosine-monophosphate cyclohydrolase, and methenyl-tetrahydrofolate cyclohydrolase. <sup>b</sup> Organisms are the sources for the enzymes whose structures are represented by the PDB IDs.



**Figure 3.** De novo biosynthesis of rNTPs adapted from Williamson and co-workers [27,28]. The four isotope-labeled precursors needed for the de novo biosynthesis of uniformly <sup>13</sup>C/<sup>15</sup>N-labeled (A) purine and (B) pyrimidine rNTPs are shown, and the heteroatoms are color coded to match their corresponding labeling positions. As shown in (B), UTP is converted to CTP in one additional step. For consecutive carbon atoms that are incorporated concomitantly, the first position is diamond shaped to indicate the direction of incorporation. All colored shapes represent either a <sup>13</sup>C or a <sup>15</sup>N atom. The isotope label of the C1 of glucose can complicate the label of C6 in the purine nucleobase. Care is usually taken to prevent dilution. Additional details can be found in the original works [27,28].

### 2.3.2. Pyrimidine De Novo Biosynthesis

Extending previous work, Williamson and co-workers developed the first de novo biosynthesis of isotopically labeled pyrimidines [28]. In contrast to purine synthesis, where the nucleobase was constructed step-by-step on the ribose, the nucleobase was directly and enzymatically coupled to the ribose to synthesize pyrimidines. Rather than directly coupling uracil, orotidine 5'-monophosphate (OMP) was produced and then converted to UTP. In a final step, UTP was converted to CTP with CTP synthetase (CTPS) (EC 6.3.4.2, PDB ID: 6NUI) and  $\text{NH}_4\text{Cl}$  [55]. This method still relied on enzymes from the pentose phosphate and nucleotide salvage biosynthetic pathways, albeit with two enzymes cloned from species other than *E. coli*: the carbamoyl-phosphate synthase-like carbamate kinase enzyme (cpkA) was cloned from the thermophile *Pyrococcus furiosus* (EC 6.3.5.5, PDB ID: 1E19), and dihydro-orotate dehydrogenase A (pydA) from *Lactococcus lactis* (EC 1.3.5.2, PDB ID: 2DOR). The de novo biosynthesis of UTP and CTP also required ATP and NADPH regeneration systems and isotope-labeled precursor compounds  $^{13}\text{C}$ -D-glucose,  $\text{NaH}^{13}\text{CO}_3$ ,  $^{15}\text{NH}_4\text{Cl}$ , and  $^{13}\text{C}/^{15}\text{N}$ -L-aspartate. Figure 3B shows the sources of pyrimidine rNTP labels. Specifically, C4, C5, C6, and N1 were derived from  $^{13}\text{C}/^{15}\text{N}$ -L-aspartate; N3 and CTP N4 were delivered by  $^{15}\text{NH}_4\text{Cl}$ ; and C2 came from  $\text{NaH}^{13}\text{CO}_3$ . All ribose carbons were provided by  $^{13}\text{C}$ -D-glucose. Again, this labeling methodology has a number of drawbacks. The use of labeled glucose can limit the potential labeling patterns if both ribose and nucleobase labels are desired. For example, production of  $\text{CO}_2$  via decarboxylation of 6-phosphogluconate to ribulose-5-phosphate during PRPP production, and of OMP to UMP, links the isotope label of the C1 of glucose and the C1 of aspartate to C2 in the pyrimidine nucleobase. If ribose labeling is not needed and only pyrimidine C2 labeling is wanted, then PRPP must again be made directly from unlabeled ribose. Similarly, care must be taken to prevent isotopic dilution from solvent and atmospheric  $\text{CO}_2$ , if both  $^{13}\text{C}$ -ribose and  $^{13}\text{C}$ -pyrimidine labeling are required. Alternatively, commercially available  $^{13}\text{C}$ -1-aspartate can be used to label C2 without C4, C5, and C6 labeling.

To summarize, a total of 16 biosynthetic enzymes were used for the efficient one-pot synthesis of UTP (and CTP) over 3–4 days with yields of up to 45% [28]. Specifically, 15 enzymes were used to synthesize atom-specific  $[1',6\text{-}^{13}\text{C}_2]$ -UTP with a 25% yield; and batches of 15, 16, and 16 enzymes helped synthesize uniformly  $^{13}\text{C}$ ,  $^{13}\text{C}/^{15}\text{N}$ ,  $^{15}\text{N}$ - $[5,3',4',5',5''\text{-}^2\text{H}_5]$ -labeled UTP with 40, 45, and 30% yields, respectively [28]. Additionally, CTPS converted uniformly  $^{15}\text{N}$ - $[5,3',4',5',5''\text{-}^2\text{H}_5]$ -UTP to its CTP counterpart with a 48% yield. The utility of their UTP and CTP labels was demonstrated for the same 30 nt RNA model system. This labeling scheme reduced spectral crowding even more than before, owing to ribose deuteration.

### 2.4. Chemo-Enzymatic Labeling

Chemo-enzymatic labeling is a hybrid approach that we developed, taking inspiration from the Tinoco [29] and Williamson [27,28] research groups. In brief, this method uses enzymes from the nucleotide salvage biosynthetic pathways and cofactor regeneration systems to couple a nucleobase and ribose, followed by subsequent phosphorylation to the rNTP in a one-pot enzymatic reaction [30,31,56]. Moreover, the nucleobase and ribose building blocks can be unlabeled, isotope-labeled, chemically synthesized, or commercially available, permitting a diverse set of labeling patterns. In collaboration with the Kreutz research group, we have prepared rNTPs with a variety of commercially unavailable labeled nucleobases at reduced costs [30,31]. This approach has many advantages over previously reported de novo [27,28] and chemical [50–54] synthesis methods, including fewer enzymes, fewer synthetic steps, and greater yields.

We described the first chemo-enzymatic synthesis of  $[1',5',6\text{-}^{13}\text{C}_3\text{-}1,3\text{-}^{15}\text{N}_2]$ -UTP and -CTP in 90 and 95% yields, respectively, by coupling  $[6\text{-}^{13}\text{C}_2\text{-}1,3\text{-}^{15}\text{N}_2]$ -uracil to  $[1,5\text{-}^{13}\text{C}_2]$ -ribose [30]. These atom-specifically-labeled rNTPs were then used in IVT to make a variety of RNAs ranging in size from 27 to 155 nts. This RNA labeling scheme significantly reduced spectral crowding, increased signal-to-noise ratios, facilitated direct carbon detection

experiments, and eliminated  $^{13}\text{C}$ - $^{13}\text{C}$  scalar and dipolar couplings. The latter benefit alleviated systematic overestimations in  $\text{C1}'$   $R_1$  rates that were observed in uniformly  $^{13}\text{C}/^{15}\text{N}$ -labeled RNA, in agreement with previous reports [30,44,49]. Extending this work, we also detailed the first chemo-enzymatic synthesis of various atom-specifically-labeled ATPs and GTPs by combining  $[8\text{-}^{13}\text{C}]$ -adenine or  $[8\text{-}^{13}\text{C}]$ -guanine with either  $[1\text{-}^{13}\text{C}]$ , or  $[2\text{-}^{13}\text{C}]$ , or  $[1,5\text{-}^{13}\text{C}_2]$ -ribose to yield the desired rNTPs in 70–95% yields [31]. These atom-specifically-labeled ATPs and GTPs were then used (along with the previously described CTP and UTP) in IVT to make a variety of RNAs ranging in size from 27 to 59 nts. This labeling pattern offered substantial sensitivity improvements and was used to develop a novel resonance assignment strategy [57]. Additional synthetic details, including a list of enzymes used, can be found in the original works [30,31,56], and the method has also been used in recent studies by our group [58,59]. This topic will also be the focus of an upcoming review.

It is worth noting that there are other approaches to enzymatically coupling nucleobase and ribose sources. Indeed, the synthetic method of Serianni and co-workers begins with the chemical synthesis of inosine, followed by its phosphorylation by purine nucleoside phosphorylase (PNPase) (EC 2.4.2.1) to give  $\alpha$ -D-ribofuranosyl 1-phosphate sodium salt ( $\alpha\text{R1P}$ ) [60]. The  $\alpha\text{R1P}$  intermediate is then glycosylated enzymatically to the ribonucleoside which can easily be converted to the desired rNTP.

### 2.5. RNA Phosphoramidite Labeling

Thus far, all discussions of isotope-labeled building blocks (Sections 2.2–2.4) have focused on rNTPs for use in IVT of RNA. However, RNA can also be prepared by chemical SPS with unlabeled or isotope-labeled amidites. Nearly all amidite synthetic schemes have nucleoside intermediates. Thus, any method to incorporate isotope labels into nucleobases, nucleosides, or rNTPs (Sections 2.2–2.4) can be converted into amidites. Indeed, in collaboration with the Kreutz group, we used the chemo-enzymatic method to make  $[1',8\text{-}^{13}\text{C}_2]$ -AMP which was then dephosphorylated to adenosine and used as the entry point for amidite synthesis [61]. Several groups have also developed strategies to obtain  $^{13}\text{C}/^{15}\text{N}$ -labeled amidites. Initial efforts were developed by the Pitsch [62] and Jones [63–66] research groups. More recently, the Micura [67,68] and Kreutz [17,61,69–71] research groups have dramatically improved the efficiency and scalability of amidite synthesis for NMR analysis. These methods were recently reviewed [36].

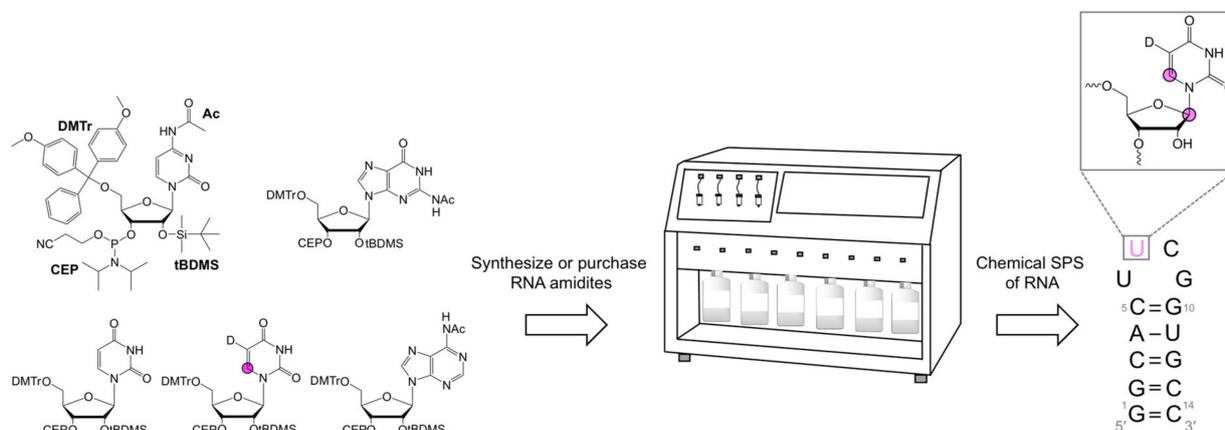
## 3. RNA Preparation Methods

With isotope-labeled building blocks in-hand, we can now discuss how they are used to prepare isotope-labeled RNA for NMR analysis. In general, amidites are used in chemical SPS to make small-to-medium sized RNAs (20–40 nts) with atom- and position-specific labeling. However, recent work from the Kreutz research group extended this length to ~80 nts [69]. All other in vitro methods of RNA preparation require T7 RNAP based IVT and use rNTPs. IVT is the most widely used method to prepare medium sized RNAs (~20–100 nts) but has major NMR limitations due to spectral overlap. Nevertheless, large RNAs can still be profitably analyzed if made from atom-specifically-labeled rNTPs [30,31,61,72–75]. The remaining RNA preparation methods rely on segmental labeling of large RNAs (>100 nts) from smaller fragments or enzymatic incorporation of position-specific isotope labels into the RNA.

### 3.1. Chemical Solid-Phase RNA Synthesis

Originally developed for DNA synthesis by Beaucage and Caruthers [76], the amidite method has since been adapted to RNA [32–36]. This approach has been further adapted to incorporate isotope-labeled amidites [36]. SPS is carried out in an automated synthesizer, requires amidite building blocks, and occurs in four steps (Figure 4). First, the 4,4'-dimethoxytrityl (DMTr) protecting group at the 5'-hydroxyl (OH) of the solid-support bound 3'-nucleoside is removed. Second, the deprotected 5'-OH then attacks the activated

amidite to couple the two nucleosides. Synthesis continues to grow the RNA molecule 3'-to-5' by repeating the first two steps following oxidation of the phosphite-triester to the phosphotriester and subsequent deprotection. Cleavage from the solid-support terminates the cycle.



**Figure 4.** Chemical solid-phase synthesis of position-specifically-labeled RNA. Amidite building blocks (unlabeled and isotope-labeled) are used in an automated synthesizer to make pre-programmed RNA sequences with position-specific labels. Magenta nucleotides harbor stable isotope labels. Magenta circles and D represent  $^{13}\text{C}$  and  $^2\text{H}$  atoms, respectively. This example shows a [6- $^{13}\text{C}$ -5- $^2\text{H}$ ]-UTP-labeled RNA at nucleotide U7.

SPS efficiency depends on the protecting group (PG) choice. RNA amidites are 5'-O-DMTr-protected, and the nucleobase exocyclic amino groups are protected with acetyl (Ac), phenoxyacetyl (Pac), benzoyl (Bz), or isobutyryl (iBu) groups. The choice of the various 2'-OH PGs requires careful deliberation. These PGs can be classified as acid- [77], photo- [78], and fluoride-labile [33,79,80]. While RNAs have been synthesized with a variety of 2'-OH PGs, only [(triisopropylsilyl)oxy]methyl (TOM) [79], *tert*-butyldimethylsilyl (tBDMS) [33], and 2'-cyanoethoxymethyl (CEM) [80] amidites have been widely used in NMR studies. Even though the CEM is the only PG that enables synthesis of RNAs >60 nts, it is commercially unavailable, which requires researchers to synthesize unlabeled and labeled CEM amidites in-house [69]. On the other hand, unlabeled TOM and tBDMS amidites and isotope-labeled tBDMS amidites are commercially available (Table 1) and therefore more accessible to a larger group of researchers. In practice, chemical SPS is rarely employed to make RNAs >60 nt for NMR analysis, and therefore TOM and tBDMS amidites are mainly used. For example, the Kreutz research group synthesized [6- $^{13}\text{C}$ ]- and [6- $^{13}\text{C}$ -5- $^2\text{H}$ ]-uridine and -cytidine, and [2,8- $^{13}\text{C}_2$ ]- and [8- $^{13}\text{C}$ ]-adenosine and -guanosine TOM and tBDMS amidites for use in  $^1\text{H}$ - [70] and  $^{13}\text{C}$ -relaxation dispersion [17,71] NMR experiments to study conformational dynamics of RNAs up to 53 nts.

### 3.2. T7 RNA Polymerase-Based *In Vitro* Transcription

IVT with DNA-dependent RNA polymerases from bacteriophage SP6, T3, or T7 (EC 2.7.7.6) is a widely used enzymatic method for RNA synthesis (Table 3) [25,26,81–83]. T7 RNAP IVT is undoubtedly the standard approach for making RNAs for NMR analysis. In practice, IVT is performed with chemically synthesized single-stranded or double-stranded DNA templates comprising one of two T7 RNAP promoter sequences (class II  $\phi$ 2.5 or class III  $\phi$ 6.5) [26,83,84]. While this approach overcomes the size restrictions of chemical SPS, it has limitations of its own. First, the widely used class III promoter is GTP-initiated and requires 5'-GG for efficient initiation [83]. Second, repeated failed transcription initiation results in 5'-end heterogeneity [85,86]. Third, T7 RNAP often adds additional non-templated rNTPs to the 3'-end of the nascent RNA [26,87]. Lastly, T7 RNAP is not immediately amenable to position-specific labeling of RNA, though some exceptions may arise. Fortunately, the 3'- and 5'-end heterogeneities are dramatically reduced by incorporat-

ing ribozyme sequences in the template in cis and trans [88,89], by chemically incorporating 2'-O-CH<sub>3</sub> rNTPs at the 3'-end [90], or by judicious choice of 5' sequences that minimize 5'-end heterogeneity (Table 4). In addition to template modification, the efficiency of T7 RNAP IVT with nucleotides bearing 2'-O-F, 2'-O-NH<sub>2</sub>, or 2'-O-CH<sub>3</sub> modified ribose is enhanced by introducing Y639F and H784A mutations [91–94]. Although the addition of non-templated rNTPs remains a challenge to T7 RNAP IVT, Roy and co-workers found no detectable 3'-end products when transcribing RNA of various sizes at higher temperature [95].

**Table 3.** Enzymes involved in IVT and DNA and RNA ligation.

Enzyme <sup>a</sup>	Gene	EC Number	PDB ID	Availability	Organism <sup>a</sup>
T7 RNA polymerase	1	2.7.7.6	1ARO	Thermo Fisher Scientific	<i>Escherichia coli</i>
T4 DNA ligase	1.3	6.5.1.1	5WFY	Sigma-Aldrich	<i>Escherichia coli</i>
T4 Polynucleotide kinase	pseT	2.7.1.78	1LY1	Thermo Fisher Scientific	<i>Escherichia coli</i>
T4 RNA ligase	63	6.5.1.3	5TT6	Thermo Fisher Scientific	<i>Escherichia coli</i>
Alkaline phosphatase	ALPL	3.1.3.1	1K7H	New England Biolabs	<i>Pandalus borealis</i>

<sup>a</sup> Organisms are the sources for enzymes whose structures are represented by the PDB IDs.

Despite its limitations, T7 RNAP IVT is an extremely versatile method and is compatible with unlabeled and isotope-labeled rNTPs. The latter can either be purchased (Section 2.1) or prepared in-house by biomass, de novo biosynthesis, or chemo-enzymatic methods (Sections 2.2–2.4). As such, this method can yield uniformly, nucleotide-specifically, and atom-specifically-labeled RNA. In the latter case, the preparation of deuterated RNA greatly simplifies NMR spectra and reduces dipolar relaxation. Deuteration of the H1', H2', H3', H4', and H5'/5'' positions in the ribose and H5 position in C/U nucleobases can be achieved with commercially available rNTPs (Table 1). On the other hand, atom-specific deuteration of purine H8 can be accomplished by heating the purine nucleobase to 55 °C in D<sub>2</sub>O [21,96] and helps distinguish adenine H8 from guanine H8 and adenine H2 [74]. These methods have been extended by the Summers research group [72–75] to study very large RNAs. Similarly, uracil H5 is easily deuterated using D<sub>2</sub>O and triethylamine [70] and then readily converted to [5-<sup>2</sup>H]-UTP and [5-<sup>2</sup>H]-CTP enzymatically [57]. This specific deuteration eliminates competing relaxation pathways among dipolar coupled protons such as H5–H6, leading to sharp linewidths, increased signal-to-noise, and reduced chemical shift overlap in NMR spectra [19]. Using these strategies, perdeuterated, fully protonated, and atom-specifically deuterated samples were prepared by IVT (and sometimes segmental labeling schemes, see Section 3.3.2) to overcome the ambiguous assignment problems of large RNAs and determine their structures [72–75].

**Table 4.** Sequences requirements for maximum RNA 5'-end homogeneity in T7 RNA polymerase in vitro transcription.

5'-End Sequences <sup>a</sup>	5'-End Heterogeneity (<1%)
GGG	No
GAG	No
GCG	No
GUG	Yes
GGA	Yes
GAA	Yes
GCA <sup>b</sup>	Yes
GUA	Yes
GGC	No
GAC	Yes
GCC	No
GUC	Yes
GGU	No

Table 4. Cont.

5'-End Sequences <sup>a</sup>	5'-End Heterogeneity (<1%)
GAU	Yes
GCU	Yes
GUU	Yes
AGG	No
AAG	Yes
ACG	Yes
AUG <sup>b</sup>	Yes
AGA	No
AAA	No
ACA	No
AUA	No
AGC	Yes
AAC	No
ACC	No
AUC	No
AGU <sup>b</sup>	Yes
AAU	No
ACU	Yes
AUU	Yes

<sup>a</sup> Sequences synthesized from either wild type or mutant T7 (p266L) RNA polymerase [97]. <sup>b</sup> Sequences synthesized from mutant p266L T7 only. Sequences starting with a G were synthesized with the T7 RNA polymerase bearing a class III promoter. T7 RNA polymerase with a class III  $\phi$ 2.5 promoter was used for sequences starting with an A. All but three of the sequences, as reported by Legault and co-workers, were synthesized with either the wild type or mutant T7 [98].

### 3.3. Enzymatic Ligation

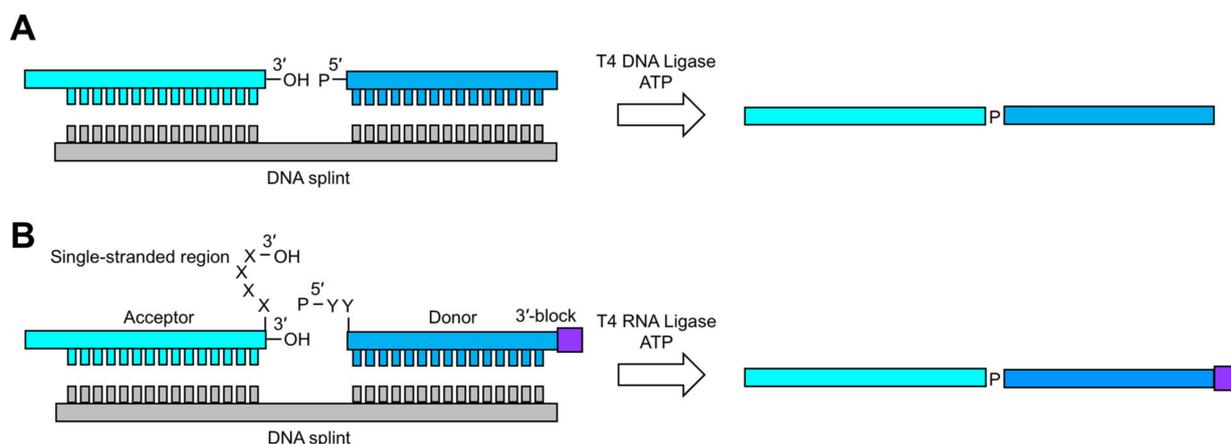
One approach that enables position-specific and segmental isotope labeling is the ligation of two RNA molecules by T4 DNA (EC 6.5.1.1) or RNA ligase (EC 6.5.1.3) (Table 3). These ligating enzymes have also been combined with self-cleaving ribozymes to segmentally label RNA. In these methods, multiple fragments of RNA are ligated to produce a larger isotope-labeled RNA that can be studied by NMR. Depending on the RNA sequence under investigation, researchers can devise unique labeling patterns to incorporate position-specific labels and greatly reduce spectral overlap and NMR analysis.

#### 3.3.1. T4 DNA and RNA Ligation

The standard method for RNA ligation uses T4 DNA ligase (EC 6.5.1.1) [99]. In the presence of ATP, this enzyme recognizes a nicked double-stranded substrate and joins a 5'-monophosphate (P) RNA (donor) with a 3'-OH RNA (acceptor) (Figure 5A). The donor and acceptor RNA fragments can either be prepared by chemical SPS or IVT. In the former case, the 5'-P can be added during or after donor RNA synthesis using T4 polynucleotide kinase (T4 PNK) (EC 2.7.1.78) and ATP. In the latter case, a donor RNA can be initiated with GMP or a 5'-XpG-3' dinucleotide. Alternatively, transcribed RNAs can be dephosphorylated with recombinant shrimp alkaline phosphatase (rSAP) (EC 3.1.3.1) and then phosphorylated with T4 PNK [100]. The 3'- and 5'-end heterogeneities dramatically reduce ligation efficiency, and therefore great care must be taken to purify the RNA of interest [99]. The two main advantages of T4 DNA ligation are that undesired side products (e.g., circularization and oligomerization) are minimized and enzymatic activity is independent of ligation junction sequence. A major disadvantage of the methodology is that T4 DNA ligase requires large quantities of RNA and is relatively inefficient at joining RNA strands [99].

An alternative method for RNA ligation is using T4 RNA ligase (EC 6.5.1.3). Like its DNA counterpart, RNA ligase requires a 5'-P donor, a 3'-OH acceptor, and ATP (Figure 5B) [101]. However, RNA ligase requires single-stranded ligation junctions, complicating the use of cDNA as a template. To overcome this limitation, Bain and Switzer

designed a DNA splint that positioned the donor and acceptor in close proximity [102]. This designed single-stranded region is compatible with T4 RNA ligase and resulted in ligation efficiencies of 53%. Building on this work, Rader and co-workers optimized ligation efficiency to near completion in less than an hour [103]. To achieve this, they (1) protected the donor 3'-OH with a 5'-silyl-2'-acetoxy-ethyl orthoester (2'-ACE) group to minimize side products, (2) chemically incorporated the 5'-P to minimize 5'-end heterogeneities, and (3) designed an optimized linker at the ligation junction.

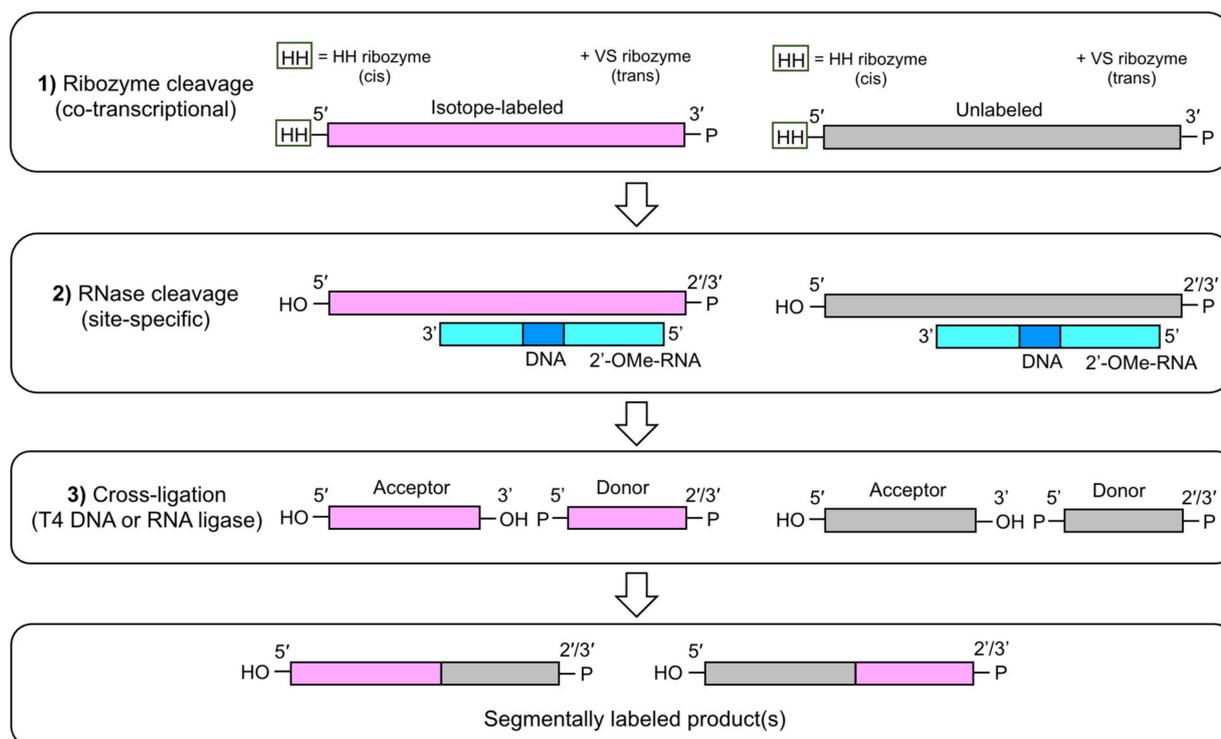


**Figure 5.** Enzymatic RNA ligation [19,99]. DNA splinted ligation schemes are shown using (A) T4 DNA ligase and (B) RNA ligase. Shown in B are the optimized 3'-end linker and 3'-end block of the donor and acceptor fragments, respectively. Sequence requirements for T4 RNA Ligase are also shown. Additional details can be found in earlier works [19,99].

### 3.3.2. Segmental RNA Labeling

Another unique ligation strategy employs RNase H and hammerhead (HH), Varkud satellite (VS), and hepatitis delta virus (HDV) self-cleaving ribozymes. This approach dramatically reduces 3'- and 5'-end heterogeneities and has therefore been embraced as a popular method to segmentally label large RNAs for NMR studies. Two such examples have come from the Puglisi [104] and Lukavsky [105] research groups. Their method was streamlined by the construction of a plasmid containing a T7 RNAP promoter, the 3'-fragment and its 3'-HH ribozyme in cis confirmation, and the RNA of interest. Despite the attractive design, the protocol took 12–14 days and only yielded 20–22 nmol RNA [104,105]. Nevertheless, this approach enabled simplified NMR structural analysis of 74 nt [105] and 77 nt [104] RNAs.

Building on this work, Wijmenga and co-workers developed an efficient two-step ligation method to selectively label central positions of large RNAs [106]. The utility of these labeling patterns to simplify resonance assignment was showcased with a 61 nt viral RNA. Still, this method only yielded 15–30 nmol RNA and required 9–11 days [106]. Finally, Allain and co-workers developed an alternative approach for segmental labeling of RNA based on IVT of two full-length RNAs with identical sequence: one labeled and one unlabeled [107]. The RNAs were flanked at the 5'- and 3'-end by the HH and VS ribozymes, respectively. After ribozyme and RNase H cleavage steps, the acceptor and donor fragments were cross-ligated using T4 DNA or RNA ligase (Figure 6). The power of this method was demonstrated in a 72 nt non-coding RNA containing four stem-loops. Four NMR samples were made: each with only one of the four stem-loops isotope-labeled. This approach also provided ~10-fold better yield (90–260 nmol RNA) and required less time (5–7 days) than did previous methods [107].



**Figure 6.** Segmental RNA labeling strategy adapted from Allain and co-workers [107]. Labeling is achieved via IVT of identical unlabeled and isotope-labeled and fragments along with HH (cis) and VS (trans) ribozymes which are cleaved co-transcriptionally (step 1). Then, site-specific RNase H cleavage is facilitated by a DNA/RNA chimera (step 2) following cross ligation reactions with T4 DNA or RNA ligases to yield segmentally labeled RNA (step 3). Additional details can be found in the original work [107].

### 3.4. Enzymatic Position-Specific RNA Labeling

The final two RNA preparation methods rely on enzymatic incorporation of position-specific isotope labels within RNA. These approaches hold promise for enabling site-specific NMR measurements of large RNAs, and therefore combine the benefits of both enzymatic and chemical SPS preparation methods.

#### 3.4.1. Position-Selective Labeling of RNA (PLOR)

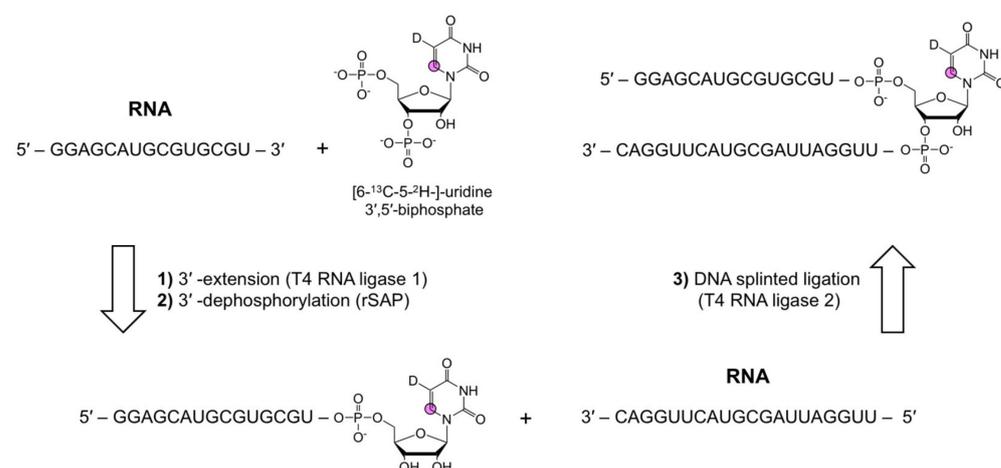
Wang and co-workers developed a powerful but laborious method to obtain position-specific isotope-labeled RNA [108–110]. This hybrid solid-liquid phase transcription technique uses an automated robotic platform known as position-selective labeling of RNA (PLOR) to prepare isotope-labeled RNAs in a three-step process that includes initiation, elongation, and termination. In PLOR, the DNA template is attached to beads and IVT is initiated by the addition of T7 RNAP and a mixture missing one of the four rNTPs, stalling RNA elongation. The beads are then washed to remove unincorporated rNTPs and elongation by T7 RNAP is resumed by the addition of a new mixture containing the previously omitted rNTP. Repetition of the elongation, stalling, and reinitiation steps enables synthesis of position-selective labeled RNA. The main drawbacks of this method are the inaccessibility of the automated synthesizer and the need for stoichiometric amounts of T7 RNAP and DNA template. Additionally, a stretch of identical nucleotides (e.g., UUU) cannot be individually labeled. Wang and co-workers showcased the utility of PLOR to aid NMR structural studies of a 71 nt RNA [108]. The power of PLOR for synthesizing site-specific and fluorescently labeled 140 nt viral RNA has also been reported [111]. Segmental labeling of the 71 nt RNA dramatically improved signal overlap and aided the unambiguous detection of multiple conformations of a single nucleotide in a position-specifically-labeled

sample. While this method holds promise, the requisite equipment needed for segmental labeling with PLOR is not publicly available.

To facilitate the use of this technique for the synthesis of larger RNAs, Liu and co-workers recently characterized the sequence requirement for the initiation step which is crucial for an efficient synthesis of RNA at optimal yields [112]. From the study of 16 RNAs derived from the adenine (riboA) and thiamine pyrophosphate (riboTPP) riboswitches, it was observed that the first six nucleotides at the 5'-end were more crucial for the PLOR-based synthesis of RNA transcripts than the nucleotides following. In addition, and consistent with IVT where G nucleotides at the first and second position of the sequence are preferred by T7 RNA polymerase, the presence of a 5'-GGG sequence was optimal for all of the sequences studied [85,113–115]. Finally, initiation lengths between 11 and 19 nt were also reported to be crucial for the synthesis of RNA transcripts at optimal yield. Following this new discovery, the future of studying large (>60 nt), biologically relevant RNAs with PLOR looks promising, and addressing the hardware limitation of this method may help to make it more broadly adopted.

### 3.4.2. Chemo-Enzymatic Position-Specific Labeling

In an attempt to make position-specifically-labeled RNA more accessible, Schwalbe and co-workers developed a chemo-enzymatic synthetic approach. Therein, a single modified nucleoside-3',5'-bisphosphate is incorporated to the 3'-end of an RNA fragment followed by DNA-splinted ligation to complete the desired internally labeled RNA sequence (Figure 7) [116]. This method was used to introduce photocaged, photoswitchable, and isotope-labeled ribonucleosides into RNAs of up to 392 nts. Furthermore, this method uses standard laboratory equipment and the commercially available enzymes T4 RNA ligase 1, rSAP, and T4 RNA ligase 2, making it readily accessible to most research groups. However, the relatively low yields of bis-phosphorylation (6–22%) and ligation (9–49%) reactions are a major drawback of this approach [116]. Improvements in these reaction steps would encourage the widespread adoption of this methodology.



**Figure 7.** A position-specific RNA labeling strategy adapted from Schwalbe and co-workers [116]. Labeling is achieved via 3'-end extension with a modified 3',5'-biphosphate nucleoside, followed by 3'-end dephosphorylation of the nucleoside by T4 RNA ligase 1 and rSAP. Lastly, DNA splinted ligation with T4 RNA ligase 1 combines the modified RNA with a 5'-phosphorylated donor RNA. Additional details can be found in the original work [116].

## 4. Conclusions

We have presented a detailed overview of the various methods for obtaining isotope-labeled rNTPs (Section 2) and how they are used to prepare isotope-labeled RNA for NMR structure and dynamics studies (Section 3). Despite these advances, there is still a tremendous bottleneck for complete RNA resonance assignment, which is a prerequisite for RNA

structure determination. On the one hand, there are dramatic costs associated with NMR sample preparation. Many samples are needed for complete RNA resonance assignment, and even more are needed for RNA structure determination. In addition, extreme time investments are required to characterize RNAs. The work needed to determine the structure of a medium-sized RNA often spans an entire PhD or post-doc, if not more. Our example from Section 2.1 illustrates this point. The solution NMR-based structure determination of a 43 nt RNA required 20 NMR samples and 10 contributing authors [39,40]. The costs of materials for NMR sample preparation and the labor are prohibitively expensive for most research groups. Moreover, the aforementioned RNA is only of modest size; studying larger RNAs will involve greater financial and time commitments.

However, RNA structural biology is moving toward larger and larger RNAs, especially as cryo-EM gains in resolution and popularity [117,118]. Nevertheless, solution NMR studies, unlike X-ray crystallography and cryo-EM, attempt to replicate the appropriate physiological environments and temperatures, and are therefore more apt for investigating the structural dynamics (from picoseconds-to-seconds) of macromolecules. Certainly, the technical advances and explosion of new data from X-ray and/or cryo-EM structures portend exciting times for the use of solution NMR in integrative structural biology and drug discovery projects [73,118–121]. Still, the challenges associated with NMR studies of large systems must be met head-on. Attention must center on either (1) developing approaches that are capable of efficient position-specific or segmental labeling of large RNAs, (2) strategic “divide-and-conquer” designs of atom-specifically-labeled RNAs, or (3) reducing the cost of selective deuteration. In the first case, improving ligation efficiency would open the door for larger RNAs to be constructed from any variations of atom-specifically-labeled RNAs from IVT and/or position-specifically-labeled RNAs from SPS. Secondly, careful design of multiple small, functional, and folded core RNAs that represent larger RNAs would be powerful if used in combination with atom-specific labeling. Indeed, this “divide-and-conquer” strategy has been used successfully to study RNAs as large as 155 nts [73,74,122]. Lastly, NMR experiments with selectively deuterated rNTPs have been used with great success by the Summers research group to determine the structures of large RNAs [72–75]. However, the extreme prices of these rNTPs prevent their widespread use (Table 1). Taken together, analysis of large RNAs by solution NMR spectroscopy will always be a challenge. Such studies will involve trade-offs between costs and efficient labeling methods. Fortunately, the methodological developments described herein demonstrate a research community that has adapted to previous challenges and will continue to do so.

**Author Contributions:** Conceptualization, writing, editing, and reviewing: L.T.O., K.M.T., R.M.L. and T.K.D. Funding: T.K.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Science Foundation (NSF 1808705).

**Acknowledgments:** We are grateful for ongoing collaboration with the research group of Christoph Kreutz (University of Innsbruck).

**Conflicts of Interest:** None declared.

## References

1. Wightman, B.; Ha, I.; Ruvkun, G. Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* **1993**, *75*, 855–862. [[CrossRef](#)]
2. Lee, R.C.; Feinbaum, R.L.; Ambros, V. The *C. elegans* Heterochronic Gene *lin-4* Encodes Small RNAs with Antisense Complementarity to *lin-14*. *Cell* **1993**, *75*, 843–854. [[CrossRef](#)]
3. Mironov, A.S.; Gusarov, I.; Rafikov, R.; Lopez, L.E.; Shatalin, K.; Kreneva, R.A.; Perumov, D.A.; Nudler, E. Sensing small molecules by nascent RNA: A mechanism to control transcription in bacteria. *Cell* **2002**, *111*, 747–756. [[CrossRef](#)]
4. Winkler, W.; Nahvi, A.; Breaker, R.R. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **2002**, *419*, 952–956. [[CrossRef](#)]
5. Grundy, F.J.; Henkin, T.M. tRNA as a positive regulator of transcription antitermination in *B. subtilis*. *Cell* **1993**, *74*, 475–482. [[CrossRef](#)]

6. Steitz, T.A.; Steitz, J.A. A general two-metal-ion mechanism for catalytic RNA. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 6498–6502. [[CrossRef](#)] [[PubMed](#)]
7. Kruger, K.; Grabowski, P.J.; Zaug, A.J.; Sands, J.; Gottschling, D.E.; Cech, T.R. Self-Splicing RNA: Autoexcision and Autocyclization of the Ribosomal RNA Intervening Sequence of Tetrahymena. *Cell* **1982**, *31*, 147–157. [[CrossRef](#)]
8. Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **1983**, *35*, 849–857. [[CrossRef](#)]
9. Zappulla, D.C.; Cech, T.R. RNA as a flexible scaffold for proteins: Yeast telomerase and beyond. *Cold Spring Harb. Symp. Quant. Biol.* **2006**, *71*, 217–224. [[CrossRef](#)]
10. Yik, J.H.N.; Chen, R.; Nishimura, R.; Jennings, J.L.; Link, A.J.; Zhou, Q. Inhibition of P-TEFb (CDK9/Cyclin T) kinase and RNA polymerase II transcription by the coordinated actions of HEXIM1 and 7SK snRNA. *Mol. Cell* **2003**, *12*, 971–982. [[CrossRef](#)]
11. D'Souza, V.; Summers, M.F. How retroviruses select their genomes. *Nat. Rev. Microbiol.* **2005**, *3*, 643–655. [[CrossRef](#)]
12. Tycowski, K.T.; Guo, Y.E.; Lee, N.; Moss, W.N.; Vallery, T.K.; Xie, M.; Steitz, J.A. Viral noncoding RNAs: More surprises. *Genes Dev.* **2015**, *29*, 567–584. [[CrossRef](#)]
13. Ganser, L.R.; Kelly, M.L.; Herschlag, D.; Al-Hashimi, H.M. The roles of structural dynamics in the cellular functions of RNAs. *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 25–27. [[CrossRef](#)]
14. Marušič, M.; Schlagnitweit, J.; Petzold, K. RNA Dynamics by NMR Spectroscopy. *ChemBioChem* **2019**, *20*, 1–27. [[CrossRef](#)]
15. Zhao, B.; Zhang, Q. Characterizing excited conformational states of RNA by NMR spectroscopy. *Curr. Opin. Struct. Biol.* **2015**, *30*, 134–146. [[CrossRef](#)] [[PubMed](#)]
16. LeBlanc, R.M.; Longhini, A.P.; Tugarinov, V.; Dayie, T.K. NMR probing of invisible excited states using selectively labeled RNAs. *J. Biomol. NMR* **2018**, *71*, 165–172. [[CrossRef](#)] [[PubMed](#)]
17. Strebitzer, E.; Nußbaumer, F.; Kremser, J.; Tollinger, M.; Kreutz, C. Studying sparsely populated conformational states in RNA combining chemical synthesis and solution NMR spectroscopy. *Methods* **2018**, *148*, 39–47. [[CrossRef](#)] [[PubMed](#)]
18. Wijmenga, S.S.; Buuren, B.N.M. Van The use of NMR methods for conformational studies of nucleic acids. *Prog. Nucl. Magn. Reson. Spectrosc.* **1998**, *32*, 287–387. [[CrossRef](#)]
19. Dayie, K.T. Key labeling technologies to tackle sizeable problems in RNA structural biology. *Int. J. Mol. Sci.* **2008**, *9*, 1214–1240. [[CrossRef](#)]
20. Tolbert, T.J.; Williamson, J.R. Preparation of specifically deuterated and <sup>13</sup>C-labeled RNA for NMR studies using enzymatic synthesis. *J. Am. Chem. Soc.* **1997**, *119*, 12100–12108. [[CrossRef](#)]
21. Lu, K.; Miyazaki, Y.; Summers, M.F. Isotope labeling strategies for NMR studies of RNA. *J. Biomol. NMR* **2010**, *46*, 113–125. [[CrossRef](#)]
22. Nikonowicz, E.P.; Sirr, A.; Legault, P.; Jucker, F.M.; Baer, L.M.; Pardi, A. Preparation of <sup>13</sup>C and <sup>15</sup>N labelled RNAs for heteronuclear multi-dimensional NMR studies. *Nucleic Acids Res.* **1992**, *20*, 4507–4513. [[CrossRef](#)]
23. Batey, R.T.; Inada, M.; Kujawinski, E.; Puglisi, J.D.; Williamson, J.R. Preparation of isotopically labeled ribonucleotides for multidimensional NMR spectroscopy of RNA. *Nucleic Acids Res.* **1992**, *20*, 4515–4523. [[CrossRef](#)]
24. Barnwal, R.P.; Yang, F.; Varani, G. Applications of NMR to structure determination of RNAs large and small. *Arch. Biochem. Biophys.* **2017**, *628*, 42–56. [[CrossRef](#)] [[PubMed](#)]
25. Milligan, J.; Groebe, D.; Whherell, G.; Uhlenbeck, O. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.* **1987**, *15*, 8783–8798. [[CrossRef](#)] [[PubMed](#)]
26. Milligan, J.F.; Uhlenbeck, O.C. Synthesis of Small RNAs Using T7 RNA Polymerase. *Methods Enzymol.* **1989**, *180*, 51–62. [[PubMed](#)]
27. Schultheisz, H.L.; Szymczyzna, B.R.; Scott, L.G.; Williamson, J.R. Pathway engineered enzymatic de novo purine nucleotide synthesis. *ACS Chem. Biol.* **2008**, *3*, 499–511. [[CrossRef](#)] [[PubMed](#)]
28. Schultheisz, H.L.; Szymczyzna, B.R.; Scott, L.G.; Williamson, J.R. Enzymatic de novo pyrimidine nucleotide synthesis. *J. Am. Chem. Soc.* **2011**, *133*, 297–304. [[CrossRef](#)] [[PubMed](#)]
29. SantaLucia, J.; Shen, L.X.; Cai, Z.; Lewis, H.; Tinoco, I. Synthesis and NMR of RNA with selective isotopic enrichment in the bases. *Nucleic Acids Res.* **1995**, *23*, 4913–4921. [[CrossRef](#)] [[PubMed](#)]
30. Alvarado, L.J.; LeBlanc, R.M.; Longhini, A.P.; Keane, S.C.; Jain, N.; Yildiz, Z.F.; Tolbert, B.S.; D'Souza, V.M.; Summers, M.F.; Kreutz, C.; et al. Regio-Selective Chemical-Enzymatic Synthesis of Pyrimidine Nucleotides Facilitates RNA Structure and Dynamics Studies. *ChemBioChem* **2014**, *15*, 1573–1577. [[CrossRef](#)]
31. Longhini, A.P.; Leblanc, R.M.; Becette, O.; Salguero, C.; Wunderlich, C.H.; Johnson, B.A.; D'souza, V.M.; Kreutz, C.; Dayie, T.K. Chemo-enzymatic synthesis of site-specific isotopically labeled nucleotides for use in NMR resonance assignment, dynamics and structural characterizations. *Nucleic Acids Res.* **2015**, *44*, 52. [[CrossRef](#)]
32. Ogilvie, K.K.; Theriault, N.; Sadana, K.L. Synthesis of Oligoribonucleotides. *J. Am. Chem. Soc.* **1977**, *99*, 7741–7743. [[CrossRef](#)] [[PubMed](#)]
33. Ogilvie, K.K.; Sadana, K.L.; Thompson, E.A.; Quilliam, M.A.; Westmore, J.B. The use of silyl groups in protecting the hydroxyl functions of ribonucleosides. *Tetrahedron Lett.* **1974**, *15*, 2861–2863. [[CrossRef](#)]
34. Reese, C.B. The Chemical Synthesis of Oligo- and Poly-ribonucleotides. In *Nucleic Acids and Molecular Biology*; Springer: Berlin/Heidelberg, Germany, 1989; pp. 164–181. [[CrossRef](#)]
35. Beaucage, S.L.; Reese, C.B. Recent advances in the chemical synthesis of RNA. In *Current Protocols in Nucleic Acid Chemistry*; John Wiley & Sons: Hoboken, NJ, USA, 2009; Volume 38, pp. 2.16.1–2.16.31. [[CrossRef](#)]

36. Becette, O.; Oleginski, L.T.; Dayie, T.K. Solid-phase chemical synthesis of stable isotope-labeled RNA to aid structure and dynamics studies by NMR spectroscopy. *Molecules* **2019**, *24*, 3476. [CrossRef]
37. Abbreviations and Symbols for the Description of Conformations of Polynucleotide Chains: Recommendations 1982. *Eur. J. Biochem.* **1983**, *131*, 9–15. [CrossRef]
38. Markley, J.L.; Bax, A.; Arata, Y.; Hilbers, C.W.; Kaptein, R.; Sykes, B.D.; Wright, P.E.; Wüthrich, K. Recommendations for the presentation of NMR structures of proteins and nucleic acids. *J. Mol. Biol.* **1998**, *280*, 933–952. [CrossRef]
39. Weickmann, A.K.; Keller, H.; Duchardt-Ferner, E.; Strebiter, E.; Juen, M.A.; Kremser, J.; Wurm, J.P.; Kreutz, C.; Wöhnert, J. NMR resonance assignments for the SAM/SAH-binding riboswitch RNA bound to S-adenosylhomocysteine. *Biomol. NMR Assign.* **2018**, *12*, 329–334. [CrossRef]
40. Weickmann, A.K.; Keller, H.; Wurm, J.P.; Strebiter, E.; Juen, M.A.; Kremser, J.; Weinberg, Z.; Kreutz, C.; Duchardt-Ferner, E.; Wöhnert, J. The structure of the SAM/SAH-binding riboswitch. *Nucleic Acids Res.* **2019**, *47*, 2654–2665. [CrossRef]
41. Hoard, D.E.; Ott, D.G. Conversion of Mono- and Oligodeoxyribonucleotides to 5'-Triphosphates. *J. Am. Chem. Soc.* **1965**, *87*, 1785–1788. [CrossRef] [PubMed]
42. Simon, E.S.; Grabowski, S.; Whitesides, G.M. Convenient Syntheses of Cytidine 5'-Triphosphate, Guanosine 5'-Triphosphate, and Uridine 5'-Triphosphate and Their Use in the Preparation of UDP-glucose, UDP-glucuronic Acid, and GDP-mannose. *J. Org. Chem.* **1990**, *55*, 1834–1841. [CrossRef]
43. Michnicka, M.J.; King, G.C.; Harper, J.W. Selective Isotopic Enrichment of Synthetic RNA: Application to the HIV-1 TAR Element. *Biochemistry* **1993**, *32*, 395–400. [CrossRef]
44. Thakur, C.S.; Dayie, T.K. Asymmetry of  $^{13}\text{C}$  labeled 3-pyruvate affords improved site specific labeling of RNA for NMR spectroscopy. *J. Biomol. NMR* **2012**, *52*, 65–77. [CrossRef] [PubMed]
45. LeMaster, D.M.; Kushlan, D.M. Dynamical mapping of E. coli thioredoxin via  $^{13}\text{C}$  NMR relaxation analysis. *J. Am. Chem. Soc.* **1996**, *118*, 9255–9264. [CrossRef]
46. Hines, J.V.; Landry, S.M.; Varani, G.; Tinoco, I. Carbon-Proton Scalar Couplings in RNA: 3D Heteronuclear and 2D Isotope-Edited NMR of a  $^{13}\text{C}$ -Labeled Extra-Stable Hairpin. *J. Am. Chem. Soc.* **1994**, *116*, 5823–5831. [CrossRef]
47. Hoffman, D.W.; Holland, J.A. Preparation of carbon-13 labeled ribonucleotides using acetate as an isotope source. *Nucleic Acids Res.* **1995**, *23*, 3361–3362. [CrossRef]
48. Fraenkel, D.G. Selection of Escherichia coli mutants lacking glucose-6-phosphate dehydrogenase or gluconate-6-phosphate dehydrogenase. *J. Bacteriol.* **1968**, *95*, 1267–1271. [CrossRef]
49. Johnson, J.E.; Julien, K.R.; Hoogstraten, C.G. Alternate-site isotopic labeling of ribonucleotides for NMR studies of ribose conformational dynamics in RNA. *J. Biomol. NMR* **2006**, *35*, 261–274. [CrossRef]
50. Vorbrüggen, H.; Krolkiewicz, K.; Benua, B. Nucleoside syntheses, XXIII) Nucleoside synthesis with trimethylsilyl triflate and perchlorate as catalysts. *Chem. Ber.* **1981**, *114*, 1234–1255. [CrossRef]
51. Földesi, A.; Nilson, F.P.R.; Glemarec, C.; Gioeli, C.; Chattopadhyaya, J. Synthesis of  $1',2',3',4',5',5''\text{-}^2\text{H}_6\text{-}\beta\text{-D-ribo}$ nucleosides and  $1',2',2'',3',4',5',5''\text{-}^2\text{H}_7\text{-}\beta\text{-D-}2'$ -deoxyribonucleosides for selective suppression of proton resonances in partially-deuterated oligo-DNA, oligo-RNA and in 2,5A core (1H-NMR window). *Tetrahedron* **1992**, *48*, 9033–9072. [CrossRef]
52. Toyama, A.; Takino, Y.; Takeuchi, H.; Harada, I. Ultraviolet Resonance Raman Spectra of Ribosyl C(1')-Deuterated Purine Nucleosides: Evidence of Vibrational Coupling between Purine and Ribose Rings. *J. Am. Chem. Soc.* **1993**, *115*, 11092–11098. [CrossRef]
53. Cook, G.P.; Greenberg, M.M. A General Synthesis of C2'-Deuterated Ribonucleosides. *J. Org. Chem.* **1994**, *59*, 4704–4706. [CrossRef]
54. Kline, P.C.; Serianni, A.S.  $^{13}\text{C}$ -Enriched Ribonucleosides: Synthesis and Application of  $^{13}\text{C}$ - $^1\text{H}$  and  $^{13}\text{C}$ - $^{13}\text{C}$  Spin-Coupling Constants To Assess Furanose and A-Glycoside Bond Conformations. *J. Am. Chem. Soc.* **1990**, *112*, 7373–7381. [CrossRef]
55. Lunn, F.A.; MacDonnell, J.E.; Bearne, S.L. Structural requirements for the activation of Escherichia coli CTP synthase by the allosteric effector GTP are stringent, but requirements for inhibition are lax. *J. Biol. Chem.* **2008**, *283*, 2010–2020. [CrossRef]
56. Arthur, P.K.; Alvarado, L.J.; Dayie, T.K. Expression, purification and analysis of the activity of enzymes from the pentose phosphate pathway. *Protein Expr. Purif.* **2011**, *76*, 229–237. [CrossRef]
57. LeBlanc, R.M.; Longhini, A.P.; Le Grice, S.F.J.; Johnson, B.A.; Dayie, T.K. Combining asymmetric  $^{13}\text{C}$ -labeling and isotopic filter/edit NOESY: A novel strategy for rapid and logical RNA resonance assignment. *Nucleic Acids Res.* **2017**, *45*, e146. [CrossRef] [PubMed]
58. Oleginski, L.T.; Dayie, T.K. Chemo-enzymatic synthesis of  $[2\text{-}^{13}\text{C}, 7\text{-}^{15}\text{N}]\text{-ATP}$  for facile NMR analysis of RNA. *Mon. für Chem.* **2020**, *151*, 1467–1473. [CrossRef]
59. Taiwo, K.M.; Becette, O.B.; Zong, G.; Chen, B.; Zavalij, P.Y.; Dayie, T.K. Chemo-enzymatic synthesis of  $^{13}\text{C}$ - and  $^{19}\text{F}$ -labeled uridine-5'-triphosphate for RNA NMR probing. *Mon. für Chem.-Chem. Mon.* **2021**, *1*, 3. [CrossRef]
60. Zhang, W.; Turney, T.; Surjancev, I.; Serianni, A.S. Enzymatic synthesis of ribo- and 2'-deoxyribonucleosides from glycofuranosyl phosphates: An approach to facilitate isotopic labeling. *Carbohydr. Res.* **2017**, *449*, 125–133. [CrossRef] [PubMed]
61. Chen, B.; Longhini, A.; Nußbaumer, F.; Kreutz, C.; Dinman, J.; Dayie, T.K. CCR5 RNA Pseudoknots: Residue and Site-Specific Labeling correlate Internal Motions with microRNA Binding. *Chem.-A Eur. J.* **2018**, *24*, 5462–5468. [CrossRef]
62. Wenter, P.; Pitsch, S. Synthesis of Selectively  $^{15}\text{N}$ -Labeled 2'-O-[[Triisopropylsilyloxy]methyl](=tom)-Protected Ribonucleoside Phosphoramidites and Their Incorporation into a Bistable 32Mer RNA Sequence. *Helv. Chim. Acta* **2003**, *86*, 3955–3974. [CrossRef]

63. Zhang, X.; Gaffney, B.L.; Jones, R.A.  $^{15}\text{N}$  NMR of a Specifically Labeled RNA Fragment Containing Intrahelical GU Wobble Pairs. *J. Am. Chem. Soc.* **1997**, *119*, 6432–6433. [[CrossRef](#)]
64. Zhang, X.; Gaffney, B.L.; Jones, R.A.  $^{15}\text{N}$  NMR of RNA fragments containing specifically labeled GU and GC pairs. *J. Am. Chem. Soc.* **1998**, *120*, 615–618. [[CrossRef](#)]
65. Shallop, A.J.; Gaffney, B.L.; Jones, R.A. Use of Both Direct and Indirect  $^{13}\text{C}$  Tags for Probing Nitrogen Interactions in Hairpin Ribozyme Models by  $^{15}\text{N}$  NMRI. *Nucleosides Nucleotides Nucleic Acids* **2004**, *23*, 273–280. [[CrossRef](#)] [[PubMed](#)]
66. Gaffney, B.L.; Jones, R.A. Regioselective 2'-Silylation of Purine Ribonucleosides for Phosphoramidite RNA Synthesis. *Curr. Protoc. Nucleic Acid Chem.* **2001**, *6*, 2.8.1–2.8.13. [[CrossRef](#)] [[PubMed](#)]
67. Neuner, S.; Kreutz, C.; Micura, R. The synthesis of  $^{15}\text{N}$ (7)-Hoogsteen face-labeled adenosine phosphoramidite for solid-phase RNA synthesis. *Mon. für Chem.* **2017**, *148*, 149–155. [[CrossRef](#)]
68. Neuner, S.; Santner, T.; Kreutz, C.; Micura, R. The “speedy” Synthesis of Atom-Specific  $^{15}\text{N}$  Imino/Amido-Labeled RNA. *Chem.-A Eur. J.* **2015**, *21*, 11634–11643. [[CrossRef](#)] [[PubMed](#)]
69. Kremser, J.; Strebitzer, E.; Plangger, R.; Juen, M.A.; Nußbaumer, F.; Glasner, H.; Breuker, K.; Kreutz, C. Chemical synthesis and NMR spectroscopy of long stable isotope labelled RNA. *Chem. Commun.* **2017**, *53*, 12938–12941. [[CrossRef](#)] [[PubMed](#)]
70. Juen, M.A.; Wunderlich, C.H.; Nußbaumer, F.; Tollinger, M.; Kontaxis, G.; Konrat, R.; Hansen, D.F.; Kreutz, C. Excited States of Nucleic Acids Probed by Proton Relaxation Dispersion NMR Spectroscopy. *Angew. Chem. Int. Ed.* **2016**, *55*, 12008–12012. [[CrossRef](#)] [[PubMed](#)]
71. Wunderlich, C.H.; Spitzer, R.; Santner, T.; Fauster, K.; Tollinger, M.; Kreutz, C. Synthesis of (6- $^{13}\text{C}$ )pyrimidine nucleotides as spin-labels for RNA dynamics. *J. Am. Chem. Soc.* **2012**, *134*, 7558–7569. [[CrossRef](#)]
72. D'Souza, V.; Dey, A.; Habib, D.; Summers, M.F. NMR structure of the 101-nucleotide core encapsidation signal of the moloney murine leukemia virus. *J. Mol. Biol.* **2004**, *337*, 427–442. [[CrossRef](#)]
73. Zhang, K.; Keane, S.C.; Su, Z.; Irobalieva, R.N.; Chen, M.; Van, V.; Sciandra, C.A.; Marchant, J.; Heng, X.; Schmid, M.F.; et al. Structure of the 30 kDa HIV-1 RNA Dimerization Signal by a Hybrid Cryo-EM, NMR, and Molecular Dynamics Approach. *Structure* **2018**, *26*, 490–498.e3. [[CrossRef](#)] [[PubMed](#)]
74. Keane, S.C.; Heng, X.; Lu, K.; Kharytonchik, S.; Ramakrishnan, V.; Carter, G.; Barton, S.; Hosic, A.; Florwick, A.; Santos, J.; et al. Structure of the HIV-1 RNA packaging signal. *Science* **2015**, *348*, 917–921. [[CrossRef](#)]
75. Marchant, J.; Bax, A.; Summers, M.F. Accurate Measurement of Residual Dipolar Couplings in Large RNAs by Variable Flip Angle NMR. *J. Am. Chem. Soc.* **2018**, *140*, 6978–6983. [[CrossRef](#)] [[PubMed](#)]
76. Beaucage, S.L.; Caruthers, M.H. Deoxynucleoside phosphoramidites—A new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Lett.* **1981**, *22*, 1859–1862. [[CrossRef](#)]
77. Scaringe, S.A.; Kitchen, D.; Kaiser, R.J.; Marshall, W.S. Preparation of 5'-silyl-2'-orthoester ribonucleosides for use in oligoribonucleotide synthesis. *Curr. Protoc. Nucleic Acid Chem.* **2004**, *16*, 2–10. [[CrossRef](#)] [[PubMed](#)]
78. Schwartz, M.E.; Breaker, R.R.; Asteriadis, G.T.; deBear, J.S.; Gough, G.R. Rapid synthesis of oligoribonucleotides using 2'-O-(o-nitrobenzyloxymethyl)-protected monomers. *Bioorg. Med. Chem. Lett.* **1992**, *2*, 1019–1024. [[CrossRef](#)]
79. Pitsch, S.; Weiss, P.; Jenny, L.; Stutz, A.; Wu, X. Reliable Chemical Synthesis of Oligoribonucleotides (RNA) with 2'-O-[(Triisopropylsilyl)oxy]methyl(2'-O-tom)-Protected Phosphoramidites. *Helv. Chim. Acta* **2001**, *84*, 3773–3795. [[CrossRef](#)]
80. Shiba, Y.; Masuda, H.; Watanabe, N.; Ego, T.; Takagaki, K.; Ishiyama, K.; Ohgi, T.; Yano, J. Chemical synthesis of a very long oligoribonucleotide with 2-cyanoethoxymethyl (CEM) as the 2'-O-protecting group: Structural identification and biological activity of a synthetic 110mer precursor-microRNA candidate. *Nucleic Acids Res.* **2007**, *35*, 3287–3296. [[CrossRef](#)]
81. Krieg, P.A.; Melton, D.A. In Vitro RNA Synthesis with SP6 RNA Polymerase. *Methods Enzymol.* **1987**, *155*, 397–415. [[CrossRef](#)]
82. Pokrovskaya, I.D.; Gurevich, V.V. In Vitro transcription: Preparative RNA yields in analytical scale reactions. *Anal. Biochem.* **1994**, *220*, 420–423. [[CrossRef](#)] [[PubMed](#)]
83. William Studier, F.; Rosenberg, A.H.; Dunn, J.J.; Dubendorff, J.W. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **1990**, *185*, 60–89. [[CrossRef](#)]
84. Coleman, T.M.; Wang, G.; Huang, F. Superior 5' homogeneity of RNA from ATP-initiated transcription under the T7 phi 2.5 promoter. *Nucleic Acids Res.* **2004**, *32*, e14. [[CrossRef](#)] [[PubMed](#)]
85. Pleiss, J.A.; Derrick, M.L.; Uhlenbeck, O.C. T7 RNA polymerase produces 5' end heterogeneity during in vitro transcription from certain templates. *RNA* **1998**, *4*, 1313–1317. [[CrossRef](#)] [[PubMed](#)]
86. Helm, M.; Brule, H.; Giege, R.; Florentz, C. More mistakes by T7 RNA polymerase at the 5' ends of in vitro-transcribed RNAs. *RNA* **1999**, *5*, 618–621. [[CrossRef](#)] [[PubMed](#)]
87. Krupp, G. RNA synthesis: Strategies for the use of bacteriophage RNA polymerases. *Gene* **1988**, *72*, 75–89. [[CrossRef](#)]
88. Ferré-D'Amaré, A.R.; Doudna, J.A. Use of cis- and trans-ribozymes to remove 5' and 3' heterogeneities from milligrams of in vitro transcribed RNA. *Nucleic Acids Res.* **1996**, *24*, 977–978. [[CrossRef](#)] [[PubMed](#)]
89. Grosshans, C.A.; Cech, T.R. A hammerhead ribozyme allows synthesis of a new form of the Tetrahymena ribozyme homogeneous in length with a 3' end blocked for transesterification. *Nucleic Acids Res.* **1991**, *19*, 3875–3880. [[CrossRef](#)]
90. Kao, C.; Zheng, M.; Rüdiger, S. A simple and efficient method to reduce nontemplated nucleotide addition at the 3' terminus of RNAs transcribed by T7 RNA polymerase. *RNA* **1999**, *5*, 1268–1272. [[CrossRef](#)]
91. Brieba, L.G.; Sousa, R. Roles of Histidine 784 and Tyrosine 639 in Ribose Discrimination by T7 RNA Polymerase. *Biochemistry* **2000**, *39*, 919–923. [[CrossRef](#)]

92. Padilla, R.; Sousa, R. A Y639F/H784A T7 RNA polymerase double mutant displays superior properties for synthesizing RNAs with non-canonical NTPs. *Nucleic Acids Res.* **2002**, *30*, e138. [[CrossRef](#)]
93. Sousa, R.; Padilla, R. A mutant T7 RNA polymerase as a DNA polymerase. *EMBO J.* **1995**, *14*, 4609–4621. [[CrossRef](#)] [[PubMed](#)]
94. Kostyuk, D.A.; Dragan, S.M.; Lyakhov, D.L.; Rechinsky, V.O.; Tunitskaya, V.L.; Chernov, B.K.; Kochetkov, S.N. Mutants of T7 RNA polymerase that are able to synthesize both RNA and DNA. *FEBS Lett.* **1995**, *369*, 165–168. [[CrossRef](#)]
95. Wu, M.Z.; Asahara, H.; Tzertzinis, G.; Roy, B. Synthesis of low immunogenicity RNA with high-temperature in vitro transcription. *RNA* **2020**, *26*, 345–360. [[CrossRef](#)] [[PubMed](#)]
96. Huang, X.; Yu, P.; Leproust, E.; Gao, X. An efficient and economic site-specific deuteration strategy for NMR studies of homologous oligonucleotide repeat sequences. *Nucleic Acids Res.* **1997**, *25*, 4758–4763. [[CrossRef](#)] [[PubMed](#)]
97. Guilleres, J.; Lopez, P.J.; Proux, F.; Launay, H.; Dreyfus, M. A mutation in T7 RNA polymerase that facilitates promoter clearance. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 5958–5963. [[CrossRef](#)]
98. Salvail-Lacoste, A.; Di Tomasso, G.; Piette, B.L.; Legault, P. Affinity purification of T7 RNA transcripts with homogeneous ends using ARiBo and CRISPR tags. *RNA* **2013**, *19*, 1003–1014. [[CrossRef](#)]
99. Moore, M.J.; Query, C.C. Joining of RNAs by splinted ligation. *Methods Enzymol.* **2000**, *317*, 109–123. [[CrossRef](#)]
100. Porecha, R.; Herschlag, D. RNA radiolabeling. In *Methods in Enzymology*; Academic Press Inc.: Cambridge, MA, USA, 2013; Volume 530, pp. 255–279, ISBN 9780124200371.
101. Romaniuk, P.J.; Uhlenbeck, O.C. Joining of RNA molecules with RNA ligase. *Methods Enzymol.* **1983**, *100*, 52–59. [[CrossRef](#)]
102. Bain, J.D.; Switzer, C. Regioselective ligation of oligoribonucleotides using DNA splints. *Nucleic Acids Res.* **1992**, *20*, 4372. [[CrossRef](#)]
103. Stark, M.R.; Pleiss, J.A.; Deras, M.; Scaringe, S.A.; Rader, S.D. An RNA ligase-mediated method for the efficient creation of large, synthetic RNAs. *RNA* **2006**, *12*, 2014–2019. [[CrossRef](#)]
104. Kim, I.; Lukavsky, P.J.; Puglisi, J.D. NMR study of 100 kDa HCV IRES RNA, using segmental isotope labeling. *J. Am. Chem. Soc.* **2002**, *124*, 9338–9339. [[CrossRef](#)] [[PubMed](#)]
105. Tzakos, A.G.; Easton, L.E.; Lukavsky, P.J. Complementary segmental labeling of large RNAs: Economic preparation and simplified NMR spectra for measurement of more RDCs. *J. Am. Chem. Soc.* **2006**, *128*, 13344–13345. [[CrossRef](#)]
106. Nelissen, F.H.T.; van Gammeren, A.J.; Tessari, M.; Girard, F.C.; Heus, H.A.; Wijmenga, S.S. Multiple segmental and selective isotope labeling of large RNA for NMR structural studies. *Nucleic Acids Res.* **2008**, *36*, e89. [[CrossRef](#)] [[PubMed](#)]
107. Duss, O.; Maris, C.; Von Schroetter, C.; Dé, F.; Allain, H.-T. A fast, efficient and sequence-independent method for flexible multiple segmental isotope labeling of RNA using ribozyme and RNase H cleavage. *Nucleic Acids Res.* **2010**, *38*, e188. [[CrossRef](#)]
108. Liu, Y.; Holmstrom, E.; Zhang, J.; Yu, P.; Wang, J.; Dyba, M.A.; Chen, D.; Ying, J.; Lockett, S.; Nesbitt, D.J.; et al. Synthesis and applications of RNAs with position-selective labelling and mosaic composition. *Nature* **2015**, *522*, 368–372. [[CrossRef](#)]
109. Liu, Y.; Yu, P.; Dyba, M.; Sousa, R.; Stagno, J.R.; Wang, Y.X. Applications of PLOR in labeling large RNAs at specific sites. *Methods* **2016**, *103*, 4–10. [[CrossRef](#)]
110. Stagno, J.R.; Yu, P.; Dyba, M.A.; Wang, Y.X.; Liu, Y. Heavy-atom labeling of RNA by PLOR for de novo crystallographic phasing. *PLoS ONE* **2019**, *14*, e0215555. [[CrossRef](#)]
111. Liu, Y.; Holmstrom, E.; Yu, P.; Tan, K.; Zuo, X.; Nesbitt, D.J.; Sousa, R.; Stagno, J.R.; Wang, Y.X. Incorporation of isotopic, fluorescent, and heavy-atom-modified nucleotides into RNAs by position-selective labeling of RNA. *Nat. Protoc.* **2018**, *13*, 987–1005. [[CrossRef](#)] [[PubMed](#)]
112. Zhang, X.; Li, M.; Liu, Y. Optimization and characterization of position-selective labelling of RNA (PLOR) for diverse RNA and DNA sequences. *RNA Biol.* **2020**, *17*, 1009–1017. [[CrossRef](#)]
113. Martin, C.T.; Muller, D.K.; Coleman, J.E. Processivity in Early Stages of Transcription by T7 RNA Polymerase. *Biochemistry* **1988**, *27*, 3966–3974. [[CrossRef](#)] [[PubMed](#)]
114. Imburgio, D.; Rong, M.; Ma, K.; McAllister, W.T. Studies of promoter recognition and start site selection by T7 RNA polymerase using a comprehensive collection of promoter variants. *Biochemistry* **2000**, *39*, 10419–10430. [[CrossRef](#)] [[PubMed](#)]
115. Cunningham, P.R.; Weitzmann, C.J.; Ofengand, J. SP6 RNA polymerase stutters when initiating from an AAA... sequence. *Nucleic Acids Res.* **1991**, *19*, 4669–4673. [[CrossRef](#)] [[PubMed](#)]
116. Keyhani, S.; Goldau, T.; Blümmler, A.; Heckel, A.; Schwalbe, H. Chemo-Enzymatic Synthesis of Position-Specifically Modified RNA for Biophysical Studies including Light Control and NMR Spectroscopy. *Angew. Chem. Int. Ed.* **2018**, *57*, 12017–12021. [[CrossRef](#)] [[PubMed](#)]
117. Kappel, K.; Zhang, K.; Su, Z.; Watkins, A.M.; Kladwang, W.; Li, S.; Pintilie, G.; Topkar, V.V.; Rangan, R.; Zheludev, I.N.; et al. Accelerated cryo-EM-guided determination of three-dimensional RNA-only structures. *Nat. Methods* **2020**, *17*, 699–707. [[CrossRef](#)] [[PubMed](#)]
118. Su, Z.; Zhang, K.; Kappel, K.; Li, S.; Palo, M.Z.; Pintilie, G.D.; Rangan, R.; Luo, B.; Wei, Y.; Das, R.; et al. Cryo-EM structures of full-length Tetrahymena ribozyme at 3.1 Å resolution. *Nature* **2021**, *596*, 603. [[CrossRef](#)] [[PubMed](#)]
119. Seffernick, J.T.; Lindert, S. Hybrid methods for combined experimental and computational determination of protein structure. *J. Chem. Phys.* **2020**, *153*, 240901. [[CrossRef](#)]
120. Shimada, I.; Ueda, T.; Kofuku, Y.; Eddy, M.T.; Wüthrich, K. GPCR drug discovery: Integrating solution NMR data with crystal and cryo-EM structures. *Nat. Rev. Drug Discov.* **2018**, *18*, 59–82. [[CrossRef](#)]

121. Gauto, D.F.; Estrozi, L.F.; Schwieters, C.D.; Effantin, G.; Macek, P.; Sounier, R.; Sivertsen, A.C.; Schmidt, E.; Kerfah, R.; Mas, G.; et al. Integrated NMR and cryo-EM atomic-resolution structure determination of a half-megadalton enzyme complex. *Nat. Commun.* **2019**, *10*, 2697. [[CrossRef](#)]
122. Pratap Barnwal, R.; Loh, E.; Godin, K.S.; Yip, J.; Lavender, H.; Tang, C.M.; Varani, G. Structure and mechanism of a molecular rheostat, an RNA thermometer that modulates immune evasion by *Neisseria meningitidis*. *Nucleic Acids Res.* **2016**, *44*, 9426–9437. [[CrossRef](#)]