Bioorg. Med. Chem. Lett. 44 (2021) 128105

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

Semi-enzymatic synthesis of pseudouridine

Andrew T. Riley [a,c], Tristan C. Sanford [c], Austin M. Woodard [b], Elliot P. Clerc, Minako Sumita [*]

Department of Chemistry, Southern Illinois University Edwardsville, Edwardsville, IL 62026, United States

ARTICLE INFO

Keywords:
RNA modification
Pseudouridine
Pseudouridine 5′-monophosphate
Pseudouridine monophosphate glycosidase
Semi-enzymatic synthesis

ABSTRACT

Modifications of RNA molecules have a significant effect on their structure and function. One of the most common modifications is the isomerization from uridine to pseudouridine. Despite its prevalence in natural RNA sequences, organic synthesis of pseudouridine has been challenging because of the stereochemistry requirement and the sensitivity of reaction steps to moisture. Herein, a semi-enzymatic synthetic route is developed for the synthesis of pseudouridine using adenosine 5′-monophosphate and uracil as the starting materials and a reverse reaction catalyzed by the pseudouridine monophosphate glycosidase. This synthetic route has only three steps and the overall yield of β-pseudouridine production was 68.4%.

More than one hundred RNA modifications have been discovered in nature and each contributes a wide range of biological functions. One of the most common and first discovered modifications is the isomerization from uridine to pseudouridine (Ψ). Pseudouridine is found in many types of RNA, such as tRNA, rRNA, and even in coding regions of mRNA.[1–6] The pseudouridine in coding regions of mRNA alters aminoacyl-tRNA selection and the speed of protein synthesis by the ribosome.[2,7] Recently, N1-methyl pseudouridine has been used to prepare the mRNA vaccines against SARS-CoV-2 that has induced the current COVID-19 pandemic. The modified nucleotide is used because it enhances the antigen expression and immune response due to inhibition of signal transduction.[8–12]

Pseudouridine is an isomer of uridine. It has a featured C5-C1′ glycosidic bond (Fig. 1).[13] In nature, the site-specific isomerization from uridine to pseudouridine is catalyzed by pseudouridine synthase that involves the cleavage of the regular N1-C1′ glycosidic bond, 180° rotation of the uracil base along the N3-C6 axis, and coupling of C5 and C1′ to form the glycosidic bond.[14] Pseudouridine can contribute to the new local RNA structure with the second N1 imino proton available for hydrogen bonding interactions. The extra imino hydrogen can form water-mediated hydrogen bonding with the 5′-side of the phosphodiester bond to stabilize the RNA structure.[15] Also, pseudouridine can increase the base-stacking with neighboring bases as compared to uridine itself.[16,17] The C—C glycosidic bond offers more rotational freedom that allows for an overall increase in structural flexibility. Although the hydrogen-bonding pattern with an adenosine residue can be the same between *syn* and *anti*-conformations of pseudouridine, it prefers *syn* conformation in nature.[18–19] However, despite the knowledge, the functions of pseudouridine within RNA remain unknown. Therefore, further investigations of the pseudouridine function are necessary to understand its unique structural and functional role in RNA conformations. In nature, most of the nucleosides are β anomers. In the β anomer, the nitrogenous base and the 5′-CH_2OH group are on the same side.

To obtain pseudouridine, extraction from RNAs is not suitable because the ratio of pseudouridine over uridine is only 0.2 ~ 0.7% in mammalian cell lines and tissues.[20] Therefore, organic chemists have been trying to develop synthetic methods to obtain large quantities of pseudouridine. The first reported synthesis of pseudouridine was in 1961 by coupling 2,4-dimethoxy-pyrimidine-5-lithium and 2,3,5-tri-O-benzoyl D-ribofuranosyl chloride. The overall yield of the synthesis was only 2%, and it was a mixture of α and β anomers.[21] Ten years after the first reported pseudouridine synthesis, the synthetic route was significantly improved which used the reaction of 2,4-di-*tert*-butoxypyrimidine-5-lithium with 2,4:3,5-di-O-benzylidene-aldehydo-D-ribose. The β-isomer was separated by column chromatography. The reaction increased the overall yield to 18%.[22] More recently, the asymmetric β-pseudouridine synthesis was improved to 47% by coupling protected ribonolactone and 5-iodo-2,4-dimethoxypyrimidine

* Corresponding author.
E-mail address: msumita@siue.edu (M. Sumita).
[a] Current Address: Department of Cell and Developmental Biology, University of Illinois Urbana-Champaign, Urbana, IL 60801, United States.
[b] Current Address: Department of Chemistry, Carl R. Woese Institute for Genomic Biology, University of Illinois Urbana-Champaign, Urbana, IL 61801, United States.
[c] Both authors contributed to this manuscript equally.

https://doi.org/10.1016/j.bmcl.2021.128105
Received 5 March 2021; Received in revised form 5 May 2021; Accepted 9 May 2021
Available online 13 May 2021
0960-894X/

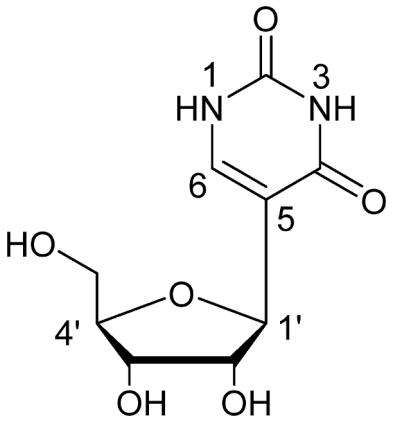

Fig. 1. The structure of pseudouridine is shown.

along with ring opening and closing with Zn^{2+} chelation.[18,23] Nonetheless, the organic synthesis of pseudouridine has been time-consuming, requires multiple steps, and has a low overall yield. In this work, we report a semi-enzymatic synthesis of pseudouridine using pseudouridine 5′-monophosphate glycosidase (ΨMP glycosidase) and alkali phosphatase (Fig. 2). It combines organic and enzymatic syntheses which requires fewer steps and has higher overall yield than the reported pure organic syntheses. One of the starting materials, ribose 5′-monophosphate, is chemically synthesized by depurination reaction of adenosine 5′-monophosphate (AMP). The coupling reaction between uracil and ribose 5′-monophosphate and the dephosphorylation are the enzymatic reactions. This semi-enzymatic reaction does not need to deal with any stereoselectivity required and moisture sensitivity encountered in pure organic synthesis which has greatly challenged the organic synthetic chemists to synthesize β-pseudouridine.

The ΨMP glycosidase is a metabolic enzyme used in the nucleotide catabolic pathway in prokaryotes.[24] Some eukaryotes have a similar enzyme, but it functions as both a kinase and ΨMP glycosidase.[25] However, higher organisms, including humans, do not have ΨMP glycosidase or similar enzymes.[26] The bacterial ΨMP glycosidase selectively cleaves the C—C glycosidic bond in pseudouridine 5′-monophosphate to form uracil and ribose 5′-monophosphate. The enzyme can also catalyze the reverse reaction.[27] The structure of ΨMP glycosidase has been solved by x-ray crystallography.[25] Based on the crystal structure, ΨMP glycosidase is a homotrimer. The active site of the ΨMP glycosidase contains one Mn^{2+} ion per subunit.[25] The metal ion interacts with the phosphate group of the substrate via water-mediated interaction.[25] Since ΨMP glycosidase is not commercially available, the psuG gene (yeiN is a synonym) that encodes ΨMP glycosidase was cloned and overexpressed with standard procedure in this work.[25] The *E. coli* K12 psuG gene was PCR amplified, then double digested with *Xho*I and *Nde*I enzymes in CutSmart Buffer (New England Biolabs, Inc, Ipswich, MA). After cloning the genomic DNA from *E. coli*, the DNA length was verified via agarose gel by comparing with the DNA ladder. The sequence of the genomic DNA was further confirmed by sequencing results from the Roy J. Carver Biotechnology Center at the University of Illinois (Urbana-Champaign, IL). The sequence results were run through a nucleotide to protein BLAST program (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and compared to the sequence of ΨMP glycosidase (Protein Data Bank 4GIJ).[25] While there was some ambiguity on the 5′ and 3′ ends of the sequencing products, the data that was recovered had a 99% match to the published sequence. The 1% difference of sequence was further compared with the published ΨMP glycosidase with ribose 5′-monophosphate adduct (Protein Data Bank 4GIK), and ΨMP covalent adduct (Protein Data Bank 4GIL), but none of the differences were observed at the reaction center (data not shown). The digested psuG was ligated into a pET-28a plasmid (MilliporeSigma, St. Louis, MO) then transformed into *E. coli* BL21 for overexpressing ΨMP glycosidase. The harvested ΨMP glycosidase was purified by a Ni-NTA column. The purified enzyme solution was analyzed by SDS-PAGE and quantified by Bradford assay. For 1 L culture of transformed *E. coli* BL21, 100–150 mg of ΨMP glycosidase was produced.

The ribose 5′-monophosphate was chemically synthesized by depurination reaction of AMP in acidic conditions. 500.0 mg of AMP (1.4 mmol) was stirred in 30 mL of 1 M aqueous HCl at 100 ℃ for an hour. Then, the reaction mixture was cooled to room temperature and adjusted to pH 8.2 with saturated aqueous KOH solution. The mixture was desalted by isopropanol precipitation three times and crude ribose 5′-monophosphate was obtained as pellets. The crude product was purified with Dowex 1x8 anion exchange column using a step-gradient of monochloroacetic acid solution from 0 to 1.0 M (0, 0.25, 0.5, 0.75, and 1.0 M). The purified ribose 5′- monophosphate was obtained by lyophilization and analyzed by 400 MHz Bruker NMR (D_2O as solvent) with water-gate solvent suppression (Supplemental Information Fig. S2). The yield was 300.0 mg (95%).

Before the enzymatic synthesis of pseudouridine 5′-monophosphate, the reaction conditions were optimized including temperature, pH, and metal concentration. All reactions were at a total volume of 100 μL and monitored via thin layered chromatography (TLC) (eluent was 7:2:1 isopropanol: double-deionized H_2O: ammonium hydroxide). The reaction temperature was investigated at 25, 30, and 35 ℃. At 35 ℃, less production of pseudouridine 5′-monophosphate was observed. The reduced formation may be caused by the equilibrium shift toward glycosylation at higher temperatures. Thus, the 30 ℃ reaction temperature was decided as the optimal condition. The reaction pH was investigated at pH 6.5, 7.1, and 7.5. Upon observations of the TLC, a pH

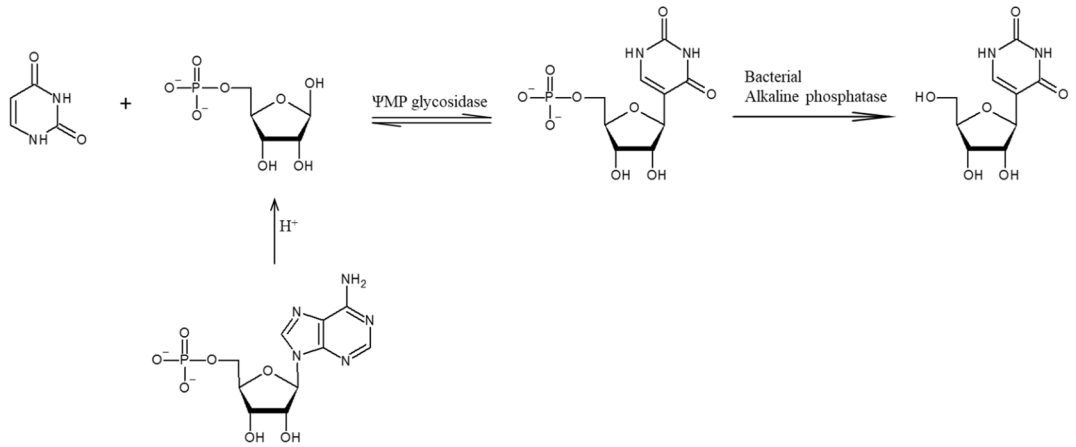

Fig. 2. The semi-enzymatic reaction scheme of pseudouridine is shown.

of 6.5 visibly produced less product compared to pH 7.1 and 7.5. The difference can be due to His-137 within the active site. His-137 plays a role in creating the salt bridge by water-mediated hydrogen bonding. Therefore, it is possible to alter the structure of the binding pocket due to the presence of His-137 at a lower pH. When comparing pH of 7.1 and 7.5, there was no visible difference in the TLC intensity of the product. However, instead of going with a more physiological pH, pH 7.1 was decided as the optimal pH to avoid possible glycosidation.[27]

The metal ion plays a major role in the formation of pseudouridine 5′-monophosphate for assisting the correct orientation of ribose 5′-phosphate within the active site and stabilizing the negative charge on the phosphate group. Manganese concentrations of 0.5 and 1 mM were tested. Using a TLC Analyzer software, the intensities of uracil and pseudouridine 5′-monophosphate were monitored.[28] Since ribose 5′-phosphate is not shown under UV, uracil and pseudouridine 5′-monophosphate were used for reaction analysis. The pseudouridine 5′-monophosphate synthesis took about 72 h with 0.5 mM Mn^{2+}, but it only took 24 h with 1 mM Mn^{2+} (Supplemental Information Fig. S1). Therefore, 1 mM Mn^{2+} was decided as the optimal condition.

The reaction termination time was also optimized. The reaction should be quenched after the equilibrium is established between production and degradation of pseudouridine 5′-monophosphate. In addition, the pseudouridine 5′-monophosphate reaction was quenched with heat due to easier purification. ΨMP glycosidase was precipitated out upon heating. Therefore, simple filtration was used to remove the enzyme after the completion of the reaction.

Based on the optimized conditions, the reaction was scaled to 30 mL where uracil (33.6 mg, 300 mmol), ribose 5′-monophosphate (82.8 mg, 300 mmol), and manganese chloride hexahydrate (6.1 mg, 30 μmol) were added to 25 mM HEPES buffer (pH 7.1). The reaction mixture was then warmed up to 30 ℃. Before addition of the enzyme, the reaction mixture was allowed to equilibrate at the optimal temperature for 10–15 min. 200 U of ΨMP glycosidase was added and the reaction mixture was swirled until the pseudouridine 5′-monophosphate production was maximized by TLC. After ~24 h, the reaction was terminated by heating the reaction mixture to 80 °C and the supernatant was freeze dried. The crude pseudouridine 5′-monophosphate was purified by column chromatography using an eluent of 7:2:1 isopropanol: double-deionized H_2O: ammonium hydroxide mixture. To convert from pseudouridine 5′-monophosphate to its nucleoside, the phosphate group was removed with Quick CIP phosphatase (New England Biolabs). The product was purified by HPLC on a SUPELCOSIL™ LC-18-S HPLC column (5 μm, 25 cm × 4.6 mm, SUPELCO, PA) in which the eluent was 1 M ammonium phosphate buffer, pH 5.2, with a 0 to 5% linear gradient of acetonitrile over 10 min at a flow rate of 1.0 mL/min. The pseudouridine peak was eluted at 7.0 min. Then, the HPLC purified product was desalted by Sephadex G-10 column (Cytiva, Marlborough, MA). The final pseudouridine product was analyzed by Bruker 400 MHz NMR (D_2O as solvent) with water-gate solvent suppression (Supplemental Information Fig. S3) and electrospray ionization mass spectrometry (ESI-MS) (Supplemental Information Fig. S5). The overall yield of this two-step enzymatic reaction was 69.7 mg (72%).

Pseudouridine was successfully synthesized by this three-step semi-enzymatic reaction. The overall yield was 68.4% from the chemical synthesis of ribose 5′-monophosphate and enzymatic pseudouridine synthesis of the final product. The synthesized pseudouridine can be further converted to phosphoramidite for the purpose of solid-phase synthesis of RNA. The intermediate, pseudouridine 5′-monophosphate, could be easily converted to pseudouridine 5′-triphosphate chemically or enzymatically to be incorporated into RNA by *in vitro* transcription. Since the starting materials are uracil and AMP, this semi-enzymatic synthesis would contribute to the synthesis of isotope-labeled pseudouridine in the RNA structure and dynamics studies using NMR techniques.

The synthesized ribose 5′-monophosphate was a mixture of α and β anomers. The NMR spectrum without water suppression showed two peaks at 5.4 ppm (α-H1′) and 5.2 (β-H1′), and the ratio of α and β anomers are 1:2. The proposed mechanism of ΨMP glycosidase-catalyzed reaction is Lys166 residue bound to the C1 position of the linear form of ribose 5′-monophosphate with dehydration before coupling with uracil. Therefore, the stereochemistry of ribose 5′-monophosphate does not affect the production of the desired β-isomer of pseudouridine 5′-monophosphate. When the enzymatic synthesis of pseudouridine 5′-monophosphate was optimized with a small-scale reaction (100 μL), the product was analyzed by Bruker 400 MHz NMR (D_2O as a solvent) with water-gate solvent suppression. The NMR spectrum shows that the ΨMP glycosidase reaction did not produce any side-product and the produced pseudouridine 5′-monophosphate was the β-isomer (Supplemental Information Fig. S2).

The importance of this semi-enzymatic pseudouridine synthesis is the use of the reverse reaction of the ΨMP glycosidase. In nature, the ΨMP glycosidase is involved in a nucleotide catabolic pathway to degrade pseudouridine 5′-monophosphate to uridine and ribose 5′-monophosphate. Pseudouridine residues are not randomly found in RNA sequence and both prokaryotes and eukaryotes have site-specific pseudouridylation enzymes, referred to as pseudouridine synthases.[29] Therefore, ΨMP glycosidase does not need to synthesize pseudouridine 5′-monophosphate in nature. To use the ΨMP glycosidase to synthesize pseudouridine 5′-monophosphate, the reverse reaction should be favored. With our overexpressed ΨMP glycosidase, the equilibrium between production and degradation of pseudouridine 5′-monophosphate was monitored by thin-layered chromatography. Since ribose 5′-monophosphate is not UV active, the analysis was based on the disappearance of the uracil and pseudouridine 5′-monophosphate on thin-layered chromatography. Based on the TLC results, both synthesis and degradation reaction equilibriums are reached in about 24 h, and the equilibrium will be kept at least for 48 h (data not shown). This suggests that the degraded uracil and ribose 5′-monophosphate did not degrade further in the presence of the ΨMP glycosidase. Also, our overall high yield indicates that our reported reaction condition favored pseudouridine 5′-monophosphate production over glycosylation. Therefore, the ΨMP glycosylation reaction by ΨMP glycosidase in nature may be favored by the simple Le Châtelier's principle that uracil and ribose 5′-monophosphate degrades further in the catabolic pathway.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors are grateful to Dr. James K. Bashkin at the University of Missouri at St. Louis for mass spectrometry analysis. This work was supported by faculty startup funds and graduate student research grants from Southern Illinois University Edwardsville.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bmcl.2021.128105.

References

1 Charette M, Gray MW. *IUBMB Life*. 2000;49(5):341–351.
2 Eyler DE, Franco MK, Batool Z, et al. *Proc Natl Acad Sci USA*. 2019;116(46):23068–23074.
3 Schwartz S, Bernstein DA, Mumbach MR, et al. *Cell*. 2014;159(1):148–162.
4 Carlile TM, Rojas-Duran MF, Zinshteyn B, Shin H, Bartoli KM, Gilbert WV. *Nature*. 2014;515(7525):143–146.
5 Lovejoy AF, Riordan DP, Brown PO. *PLoS One*. 2014;9(10), e110799.
6 Li X, Zhu P, Ma S, et al. *Nat Chem Biol*. 2015;11(8):592–597.
7 Parisien M, Yi C, Pan T. *RNA*. 2012;18(3):355–367.

8 Anderson BR, Muramatsu H, Nallagatla SR, Bevilacqua PC, Sansing LH, Weissman D, Kariko K. *Nucleic Acids Res*. 2010.
9 Andries O, Mc Cafferty S, De Smedt SC, Weiss R, Sanders NN, Kitada T. *J Control Release*. 2015;217:337–344.
10 Pardi N, Hogan MJ, Naradikian MS, et al. *J Exp Med*. 2018;215(6):1571–1588.
11 Jackson NAC, Kester KE, Casimiro D, Gurunathan S, DeRosa F. *NPJ Vaccines*. 2020;5:11.
12 Corbett KS, Edwards DK, Leist SR, et al. *Nature*. 2020;586(7830):567–571.
13 Spenkuch F, Motorin Y, Helm M. *RNA Biol*. 2014;11(12):1540–1554.
14 Yu YT, Meier UT. *RNA Biol*. 2014;11(12):1483–1494.
15 Newby MI, Greenbaum NL. *PNAS*. 2002;99(20):12697–12702.
16 Davis DR. *Nucleic Acids Res*. 1995;23(24):5020–5026.
17 Yarian CS, Basti MM, Cain RJ, et al. *Nucleic Acids Res*. 1999;27(17):3543–3549.
18 Chang YC, Herath J, Wang TH, Chow CS. *Bioorg Med Chem*. 2008;16(5):2676–2686.
19 Zhou J, Lv C, Liang B, Chen M, Yang W, Li H. *J Mol Biol*. 2010;401(5):690–695.
20 Roundtree IA, Evans ME, Pan T, He C. *Cell*. 2017;169(7):1187–1200.
21 Shapiro R, Chambers RW. *J Am Chem Soc*. 1961;83(18):3920–3921.
22 Lerch U, Burdon MG, Moffatt JG. *J Org Chem*. 1971;36(11):1507–1513.
23 Hanessian S, Marcotte S, Machaalani R, Huang G. *Org Lett*. 2003;5(23):4277–4280.
24 Thapa K, Oja T, Metsa-Ketela M. *FEBS J*. 2014;281(19):4439–4449.
25 Huang S, Mahanta N, Begley TP, Ealick SE. *Biochemistry*. 2012;51(45):9245–9255.
26 Feng B, Zheng MH, Zheng YF, et al. *J Gastroenterol Hepatol*. 2005;20(12):1913–1919.
27 Preumont A, Snoussi K, Stroobant V, Collet JF, Van Schaftingen E. *J Biol Chem*. 2008;283(37):25238–25246.
28 Hess AVI. *J Chem Educ*. 2007;84(5):842–847.
29 Ofengand J. *FEBS Lett*. 2002;514(1):17–25.