PCCP

ROYAL SOCIETY OF CHEMISTRY

PAPER

Published on 06 August 2026

Cite this: *Phys. Chem. Chem. Phys.*, 2026, **28**, 20872

Quantum chemical profiling of the electronic structure and hydrolytic stability of modified ribonucleosides

Volga Kojasoy and Ronald T. Raines *

Received 16th April 2026,
Accepted 5th August 2026

DOI: 10.1039/d6cp01428c

rsc.li/pccp

Chemical modifications to RNA play essential roles in regulating structure, stability, and biological function, yet a unifying physicochemical framework for understanding how these structural modifications perturb the underlying electronic landscape and influence intrinsic reactivity remains lacking. Here, we apply density functional theory to compute electronic-structure descriptors for a comprehensive set of naturally occurring modified ribonucleosides. By analyzing HOMO–LUMO gaps as measures of global electronic softness and Wiberg bond indices as local descriptors of glycosidic bond strength, we establish systematic relationships linking stereoelectronic substitution patterns and nucleobase π-conjugation to molecular reactivity and hydrolytic stability. We find that sulfur and selenium incorporation and major-groove substitutions tend to narrow HOMO–LUMO gaps and weaken glycosidic bonds, whereas C-glycosides (as in pseudouridines) confer electronic stabilization. These results reveal physical principles governing the intrinsic reactivity of modified RNA building blocks and provide a predictive framework for anticipating modification-dependent behavior relevant to RNA stability, degradation, and next-generation sequencing technologies used to characterize the epitranscriptome.

1 Introduction

RNA occupies a central position in biological information transfer and regulation.[1] Beyond its classical role as an intermediary between DNA and proteins, RNA functions as a catalyst, regulator, sensor, and structural scaffold.[2] Moreover, RNA is now known to be both the cause of (*via* RNA viruses like SARS-CoV-2) and the cure for (*via* RNA-based vaccines) some of humanity's most formidable challenges.

The chemical versatility of RNA is expanded by post-transcriptional modifications that diversify the four canonical ribonucleosides—adenosine (A), guanosine (G), cytidine (C), and uridine (U)—into a chemically rich repertoire. Since the identification of pseudouridine as the "fifth ribonucleoside" in 1951,[3] the catalog of known RNA modifications has grown to more than 170 distinct chemical entities distributed across all domains of life.[4–11]

In recent years, the biological significance of RNA modifications has become increasingly apparent. Modified nucleosides are known to regulate the folding, localization, decay, and translational fidelity of RNA and to play critical roles in development, stress responses, and disease.[4–11] Their importance is underscored by the success of nucleoside-modified mRNA vaccines, enabled by discoveries showing that base modifications suppress innate immune recognition while enhancing translational efficiency and stability.[12–15] These advances have renewed interest in RNA chemistry across biomedicine, biotechnology, agriculture, and information storage, and have inspired the "Human RNome Project".[16,17] From a physical chemistry perspective, these diverse post-transcriptional modifications constitute a vast, naturally occurring library of stereoelectronic perturbations to the canonical nucleobase architecture.

Despite their functional importance, the physicochemical principles that govern the intrinsic reactivity of modified ribonucleosides remain incompletely understood. Experimental studies have provided valuable insights into enzymatic recognition and biological outcomes, but a general quantum mechanical framework that connects chemical modifications to fundamental electronic structure—such as frontier molecular orbital (FMO) topologies and localized bond orders—is lacking. This gap is particularly evident for the glycosidic bond linking the nucleobase to the ribose sugar. Cleavage of this bond produces abasic RNA, altering molecular integrity and function and complicating analytical methods such as RNA sequencing.[18–21] Understanding how specific modifications affect glycosidic bond stability is therefore essential for both fundamental RNA chemistry and biology, as well as the rational design of sequencing technologies and reference standards.

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts, 02139, USA. E-mail: rtraines@mit.edu

Here, we address this challenge by using quantum chemical calculations to systematically evaluate electronic descriptors across a broad set of naturally occurring modified ribonucleosides. We focus on two complementary quantities: the HOMO–LUMO gap, which reflects the ease of electron redistribution and orbital mixing, and the Wiberg bond index (WBI) of the glycosidic bond, which provides a quantitative measure of local bond strength. By correlating these descriptors with the modification type, nucleobase identity, and substitution topology, we identify the physical principles governing modification-dependent reactivity. These insights establish a predictive framework for understanding RNA stability at the molecular level and lay the groundwork for future experimental and technological advances in epitranscriptomics.

2 Methods

We focused on 147 naturally occurring modified ribonucleosides (Fig. 1 and 2).[9] Each ribonucleoside was modeled as a 3′,5′-bis(methylphosphate) (Me-pNp-Me) to mimic the electronic influence of terminal phosphates while avoiding artificial charge delocalization associated with extended polyanionic backbones. This capping strategy provides a chemically realistic yet computationally tractable representation of RNA building blocks for comparative electronic analysis. We also assigned charges to amino, amidino, guanidino, carboxyl, phosphoryl, and sulfonyl groups expected at near-neutral pH.

The geometry of each modified nucleoside was optimized using density functional theory (DFT) with the M06-2X exchange–correlation functional and the 6-31+G(d,p) basis set, incorporating aqueous solvation *via* the SMD implicit solvent model (SMD–H_2O).[22,23] This level of theory provides a reliable treatment of heterocyclic systems, noncovalent interactions, and the heavier main-group elements (*e.g.*, sulfur and selenium) relevant to modified nucleobases. Full geometry optimizations were performed without symmetry constraints.

Within conceptual DFT, the maximum hardness principle states that molecules naturally evolve their electronic structures to achieve maximum stability, which corresponds to maximum absolute hardness, or equivalently, minimum softness.[24–29] Because chemical softness is inversely proportional to the HOMO–LUMO energy gap, a larger gap corresponds to lower chemical reactivity. Thus, the HOMO–LUMO gap serves as a global electronic descriptor of molecular softness, indicating the ease of electron redistribution and orbital mixing, and thereby providing a comparative measure of intrinsic chemical reactivity.[30–32] HOMO–LUMO gaps were obtained directly from the converged Kohn–Sham orbital energies. Although Kohn–Sham eigenvalue differences do not correspond rigorously to excitation energies, they provide reliable relative trends across structurally homologous series when computed at a consistent level of theory.

The WBI is a measure of electron density shared between atoms A and B in an A–B bond, determined as the sum of the off-diagonal square of the density matrix D, where p and q are atomic orbitals localized on atoms A and B, respectively, as in eqn (1).[33,34]

$$\mathrm{WBI}_{\mathrm{AB}} = \sum_{p\mathrm{A}} \sum_{q\mathrm{B}} D_{pq}^{2} \tag{1}$$

WBIs were calculated for the bond between C1′ of ribose and the nitrogen (or carbon) of the nucleobase using Natural Bond Orbital (NBO) analysis (version 7.0) as implemented in Gaussian 16.[33,35,36] NBO analysis transforms delocalized Kohn–Sham orbitals into localized, Lewis-like bonding patterns, providing a well-defined quantum mechanical analogue to classical bond order.

3 Results and discussion

3.1 Electronic structure of modified ribonucleosides

RNA modifications encompass a wide range of chemical transformations, including methylation, thiolation, oxidation, reduction, ring closure, and conjugation to bulky substituents (Fig. 1 and 2). These modifications perturb the electronic structure of the nucleobase and, in turn, influence its reactivity. Although HOMO–LUMO gaps do not constitute quantitative predictors of absolute reaction rates, they provide a consistent and internally comparable measure of intrinsic electronic softness across a chemically diverse yet structurally homologous set of nucleosides.

Across all four nucleobases, larger and more electronically diverse substituents generally reduce the HOMO–LUMO gap, indicating increased electronic softness and reactivity (Table S1 and Fig. 3A). This effect is particularly pronounced for modifications that disrupt base planarity or introduce low-lying acceptor orbitals, such as the incorporation of highly polarizable, diffuse sulfur or selenium atoms. In contrast, simple methylations at many positions tend to increase the HOMO–LUMO gap relative to the unmodified nucleoside, consistent with the inductive electron-donating character of methyl groups, which raise occupied orbital energies without introducing low-lying vacant orbitals. Dihydrouridines exhibit the largest HOMO–LUMO gaps in the entire dataset (Fig. 3A), as saturation of the C5–C6 bond eliminates π-conjugation within the pyrimidine ring, confining electron density to σ-bonding frameworks and markedly decreasing electronic softness.

Substitutions oriented toward the major groove (*e.g.*, at the 5 position of a pyrimidine nucleoside) often lead to smaller gaps than analogous minor-groove substitutions, consistent with stereoelectronic interactions with the ribosyl and phosphoryl groups in the backbone. Purine nucleosides exhibit systematically smaller HOMO–LUMO gaps than pyrimidines (Fig. 3B), reflecting the greater polarizability and extended π-conjugation of the fused bicyclic system. These trends suggest that global electronic descriptors can serve as useful predictors of modification-dependent reactivity in RNA.

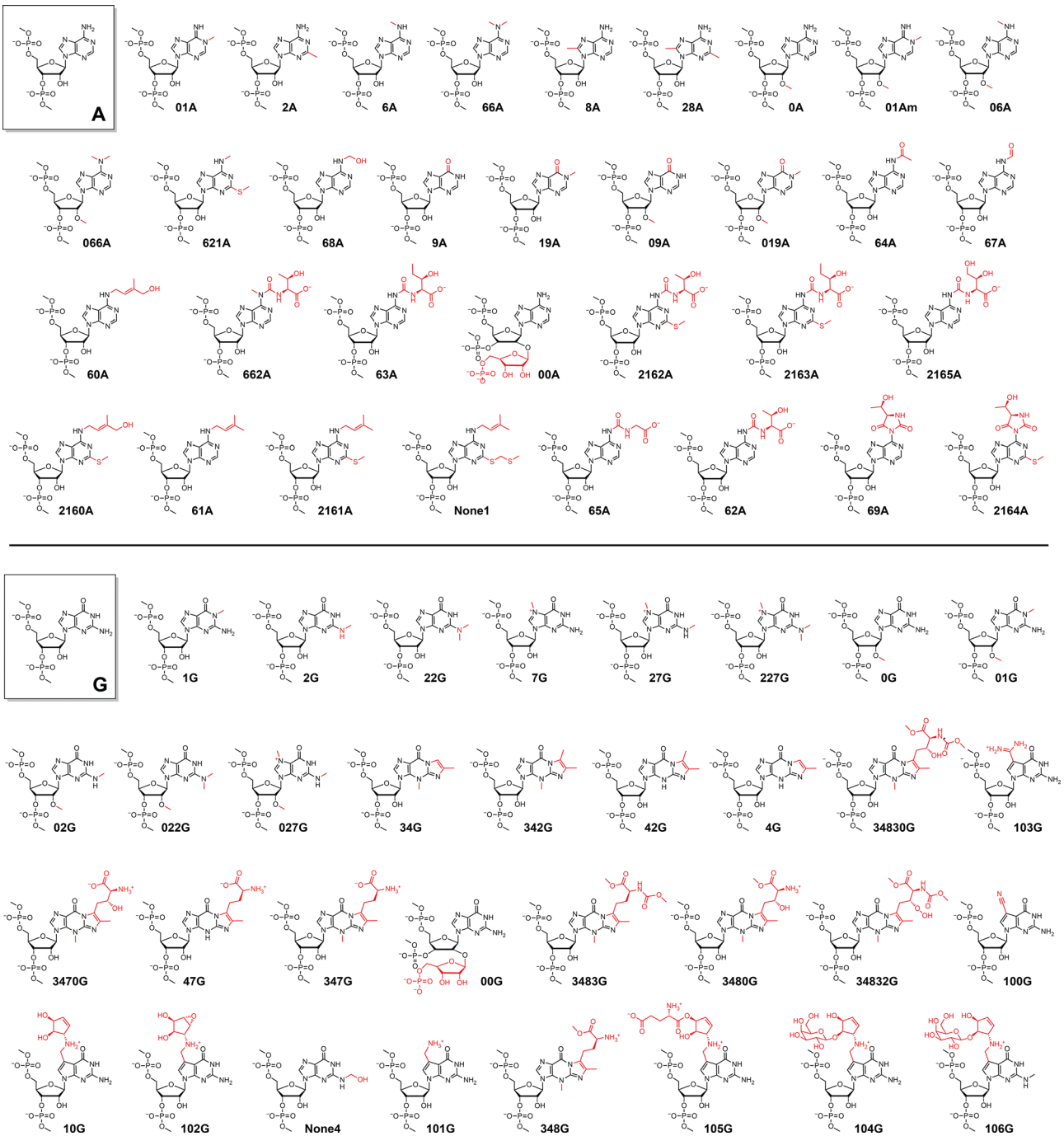

Fig. 1 Chemical structures of purine nucleosides in natural RNA.[9] The nucleosides are depicted with 5′-and 3′-methylphosphates and in the charge state used for calculations. Post-transcriptional modifications are indicated in red.

3.2 Glycosidic bond stability

RNA can undergo hydrolysis of the glycosidic bond linking the ribose sugar to the nitrogenous base, producing abasic sites that alter RNA stability, processing, and biological activity.[18–21] Modified bases can be especially labile: wybutosine (**3483G**) at position 37 of phenylalanine tRNA is readily lost, leaving an abasic site adjacent to the anticodon,[18] and 7-methylguanosine (**7G**) can be removed reductively with sodium borohydride.[19] The reactivity of **7G** ($E_{HOMO–LUMO}$ = 7.31 eV; WBI = 0.8764) was exploited in early RNA sequencing methods,[20] illustrating how glycosidic bond instability can be both a natural consequence of RNA modification and a useful experimental tool. Purine glycosidic bonds are generally more labile than those of pyrimidines, because the bicyclic purine scaffold enhances charge delocalization, stabilizing the cationic intermediate formed upon bond cleavage.[37] Abasic sites in RNA can also arise from

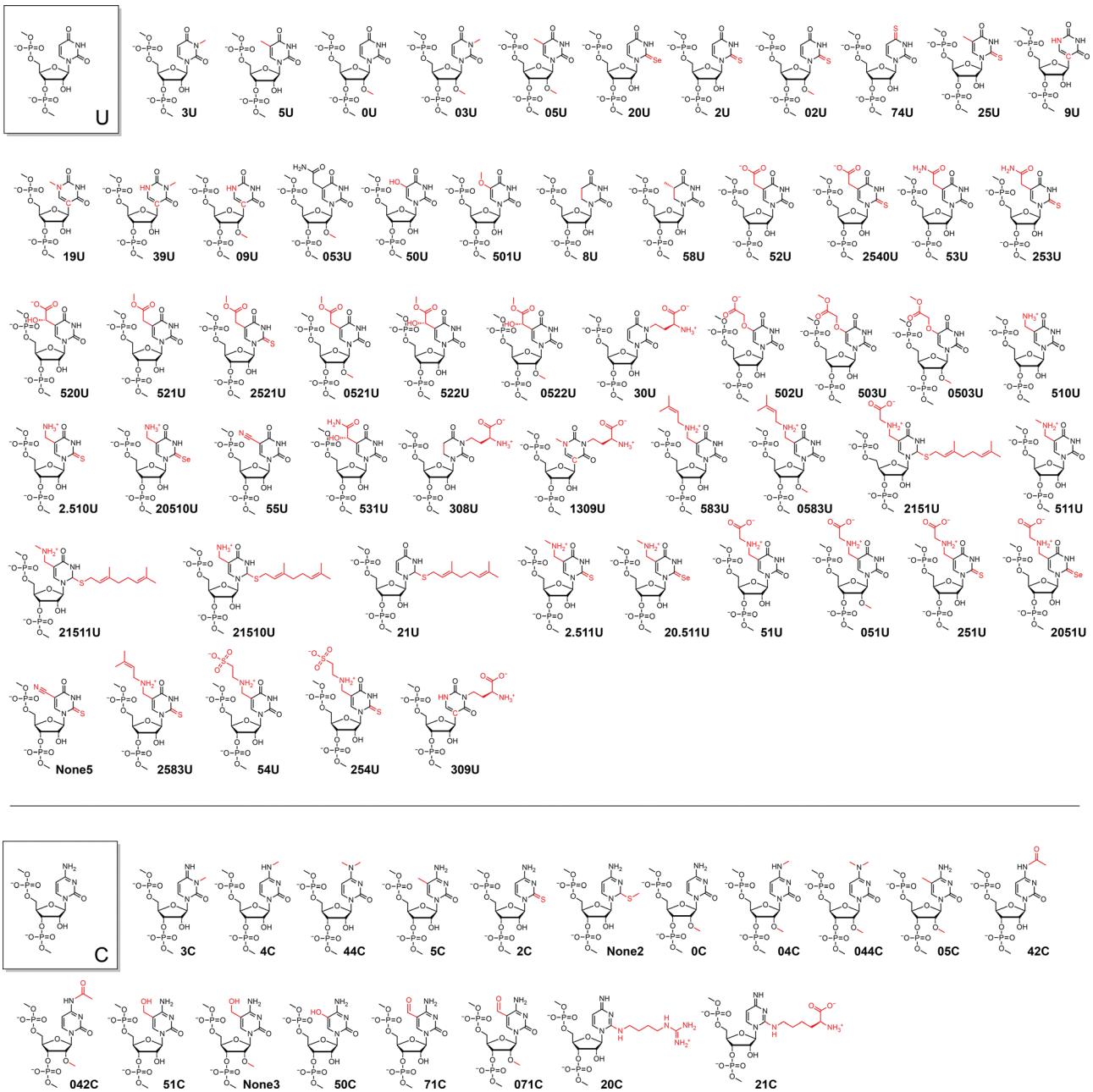

Fig. 2 Chemical structures of pyrimidine nucleosides in natural RNA.[9] The nucleosides are depicted with 5′- and 3′-methylphosphates and in the charge state used for calculations. Post-transcriptional modifications are indicated in red.

methylation-induced hydrolysis, radical damage, or enzymatic depurination and are emerging as key intermediates in cellular signaling and transcriptional regulation.[21,38–41]

To assess how chemical modifications influence glycosidic bond strength, we calculated WBIs for the C1′–nucleobase linkage in each modified ribonucleoside (Table S2). Because glycosidic bond cleavage necessarily involves redistribution of electron density from this linkage, the WBI provides a direct electronic measure of local bond integrity that complements global reactivity descriptors. C-Glycosidic bonds, characteristic of pseudouridines (*e.g.*, **39U**, **9U**, **09U**, **309U**, **19U**, and **1309U**), display markedly higher WBIs than do N-glycosidic bonds (Fig. 3A), consistent with their enhanced resistance to hydrolysis and the greater covalent character of the C–C linkage relative to the C–N linkage. The C-glycosidic linkage in pseudouridines confers stability not only to the glycosidic bond itself but also to the adjacent phosphodiester bond. We previously showed that the *C*-for-*N* substitution at the glycosidic position diminishes the inductive effect that acidifies the 2′-hydroxy group, thereby attenuating nucleophilic attack on the P–O$^{5''}$ bond in both spontaneous and ribonuclease-catalyzed pathways.[42] That finding was corroborated independently by Kool and coworkers using complementary experimental approaches.[43] Thus, the electronic consequences of

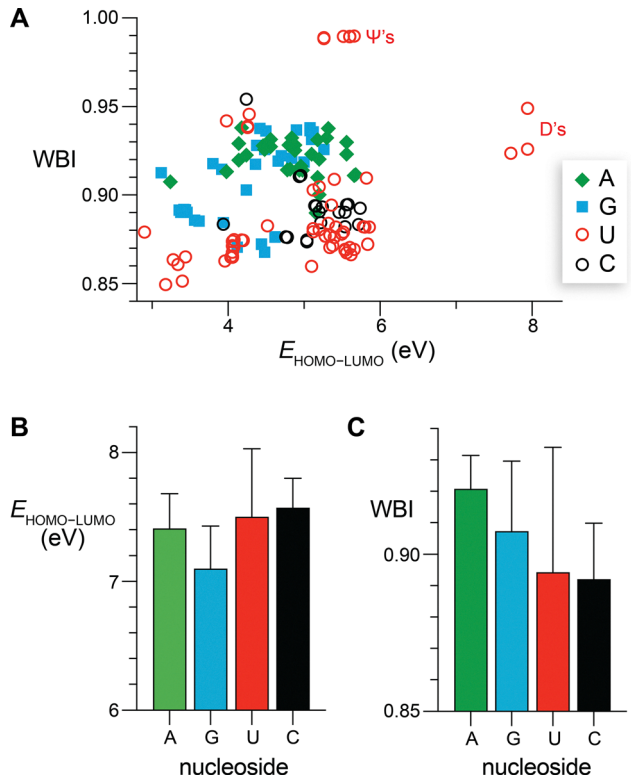

Fig. 3 Compiled data from quantum chemical calculations of Me-pNp-Me ribonucleosides. (A) Plot of $E_{HOMO-LUMO}$ *versus* WBI values for 147 modified ribonucleosides. Data are listed in Tables S1 and S2. Ψ, pseudouridine; D, dihydrouridine. (B) $E_{HOMO-LUMO}$ values (mean ± SD) for A, G, U, and C nucleosides from panel A. Data are listed in Table S1. (C) WBI values (mean ± SD) for A, G, U, and C nucleosides from panel A. Data are listed in Table S2.

C-glycosidic connectivity manifest in the stability of two distinct bonds: the glycosidic bond (this work) and the phosphodiester bond.[42,43] Pyrimidine N-glycosidic bonds tend to exhibit lower WBIs than their purine counterparts (Fig. 3C). Modifications that introduce sulfur or selenium or that are directed toward the major groove tend to weaken the glycosidic bond (Table S2).

3.3 Coupling between global reactivity and local bond strength

Although HOMO–LUMO gaps and glycosidic bond indices probe distinct aspects of molecular stability, they reveal complementary facets of modification-dependent reactivity. In several chemically important classes of RNA modifications, particularly sulfur-containing purines and heavily substituted nucleosides, reduced HOMO–LUMO gaps are accompanied by lower glycosidic WBIs, consistent with enhanced electronic reactivity occurring alongside local glycosidic bond weakening (Tables S1 and S2). Across the broader landscape of naturally occurring RNA modifications, however, these descriptors capture distinct contributions to stability rather than exhibiting a universal one-to-one relationship. Together, they provide a molecular framework for interpreting experimental observations of modification-dependent depurination and RNA lability.

4 Conclusion

Quantum chemical analysis reveals clear physical principles governing the intrinsic reactivity and hydrolytic stability of modified ribonucleosides. By combining global electronic descriptors with local measures of glycosidic bond strength, we establish a unified framework that rationalizes how nucleobase identity, substitution topology, and functional group chemistry dictate modification-dependent behavior. Inherently predictive, these results provide a foundation for future experimental validation and offer practical guidance for developing next-generation RNA sequencing technologies and reference standards. Moreover, the computational design of hydrolytically stable RNA molecules could enhance the utility of RNA-based therapeutic agents. In that context, the electronic consequences of C-glycosidic connectivity in pseudouridines, which are the most abundant modified nucleosides in natural RNA, extend to both glycosidic bond integrity (this work) and phosphodiester bond stability.[42,43] We speculate that this electronic protection helps to preserve natural RNA structures, particularly those vulnerable to spontaneous cleavage, and may constitute one biological rationale for the evolutionary prevalence of modified RNAs. Taken together, these results demonstrate how simple quantum chemical descriptors can anticipate modification-dependent vulnerabilities in RNA without recourse to experiment. More broadly, this work provides a physical framework for understanding structure–reactivity relationships in functionalized heteroaromatic biopolymers.

Conflicts of interest

There are no conflicts to declare.

Data availability

The data supporting the findings of this study, including the calculated HOMO–LUMO energy gaps and Wiberg bond indices (WBIs) of the glycosidic bonds, are available in the supplementary information (SI) accompanying this article. See DOI: https://doi.org/10.1039/d6cp01428c.

The optimized structures reported herein are available in the ioChem-BD repository,[44] see: https://doi.org/10.19061/iochem-bd-6-647. https://iochembd.bsc.es/browse/review-collection/100/492844/51ac834875bdc0bf6f0fcc55.

Acknowledgements

We are grateful to Prof. Vivian G. Cheung (Brown University) for helpful discussions. This work was supported by grants R01 CA073808 and R35 GM148220 (NIH) and used computational resources through allocation BIO230178 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program, which was supported by grants 2137603, 2138259, 2138286, 2138296, and 2138307 (NSF). RTR is a member of the Human RNome Project.

Notes and references

1 F. H. C. Crick, in *Symposia of the Society for Experimental Biology*, Cambridge University Press, Cambridge, UK, 1958, vol. 12, pp. 138–163.
2 T. R. Cech, *The Catalyst: RNA and the Quest to Unlock Life's Deepest Secrets*, W. W. Norton & Company, New York, NY, 2024.
3 W. E. Cohn and E. Volkin, *Nature*, 1951, **167**, 483–484.
4 C. S. Chow, T. N. Larnichhane and S. K. Mahto, *ACS Chem. Biol.*, 2007, **2**, 610–619.
5 M. Frye, S. R. Jaffrey, T. Pan, G. Rechavi and T. Suzuki, *Nat. Rev. Genet.*, 2016, **17**, 365–372.
6 P. Boccaletto, M. A. Machnicka, E. Purta, P. Piatkowski, B. Gaginski, T. K. Wirecki, V. de Crécy-Lagard, R. Ross, P. A. Limbach, A. Kotter, M. Helm and J. M. Bujnicki, *Nucleic Acid Res.*, 2017, **46**, D303–D307.
7 R. J. Ontiveros, J. Stoute and K. F. Liu, *Biochem. J.*, 2019, **476**, 1227–1245.
8 T. Suzuki, *FASEB J.*, 2020, **34**, 1.
9 P. J. McCown, A. Ruszkowska, C. N. Kunkler, K. Breger, J. P. Hulewicz, M. C. Wang, N. A. Springer and J. A. Brown, *Wiley Interdiscip. Rev.: RNA*, 2020, **11**, e1595.
10 C. Xue, Z. Chu, Q. Zheng, S. Jiang, Z. Bao, Y. Su, J. Lu and L. Li, *Signal Transduction Targeted Ther.*, 2022, **7**, 142.
11 B. Sharma, W. Prall, G. Bhatia and B. D. Gregory, *Annu. Rev. Plant Biol.*, 2023, **74**, 53–85.
12 K. Kariko, M. Buckstein, H. Ni and D. Weissman, *Immunity*, 2005, **23**, 165–175.
13 K. Kariko, H. Muramatsu, F. A. Welsh, J. Ludwig, H. Kato, S. Akira and D. Weissman, *Mol. Ther.*, 2008, **16**, 1833–1840.
14 B. R. Anderson, H. Muramatsu, S. R. Nallagatla, P. C. Bevilacqua, L. H. Sansing, D. Weissman and K. Kariko, *Nucleic Acids Res.*, 2010, **38**, 5884–5892.
15 F. Krammer and P. Palese, *Proc. Natl. Acad. Sci. U.S.A.*, 2024, **121**, e2400423121.
16 *Charting a Future for Sequencing RNA and Its Modifications: A New Era for Biology and Medicine*, The National Academies Press, Washington, DC, 2024, https://www.nationalacademies.org/projects/DELS-BLS-22-08/publication/27165.
17 E. Pennisi, *Science*, 2024, **383**, 1398.
18 R. Thiebe and H. G. Zachau, *Eur. J. Biochem.*, 1968, **5**, 546–555.
19 W. Wintermeyer and H. G. Zachau, *FEBS Lett.*, 1970, **11**, 160–164.
20 D. A. Peattie, *Proc. Natl. Acad. Sci. U.S.A.*, 1979, **76**, 1760–1764.
21 T. Prashar, F. De La Selle and K. A. Hudak, *RNA Biol.*, 2023, **20**, 348–358.
22 Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2007, **120**, 215–241.
23 Y. Zhao and D. G. Truhlar, *Acc. Chem. Res.*, 2008, **41**, 157–167.
24 R. G. Parr and R. G. Pearson, *J. Am. Chem. Soc.*, 1983, **105**, 7512–7516.
25 R. G. Parr and P. K. Chattaraj, *J. Am. Chem. Soc.*, 1991, **113**, 1854–1855.
26 P. W. Ayers and R. G. Parr, *J. Am. Chem. Soc.*, 2000, **122**, 2010–2018.
27 R. A. Miranda-Quintana, F. Heidar-Zadeh, S. Fias, A. E. A. Chapman, S. Liu, C. Morell, T. Gmez, C. Crdenas and P. W. Ayers, *Front. Chem.*, 2022, **10**, 906674.
28 A. R. Al Nabulsi, F. Heidar-Zadeh and P. W. Ayers, *J. Phys. Chem. A*, 2026, **130**, 3221–3232.
29 A. Bajaj, F. Heidar-Zadeh, T. Stuyver, S. Habershon, P. W. Ayers and F. D. Proft, *J. Comput. Chem.*, 2026, **47**, e70365.
30 R. G. Parr and W. Yang, *Density–Functional Theory of Atoms and Molecules*, Oxford University Press, Oxford, UK, 1989.
31 P. Geerlings, F. De Proft and W. Langenaeker, *Chem. Rev.*, 2003, **103**, 1793–1874.
32 F. De Proft, P. Geerlings, F. Heidar-Zadeh and P. W. Ayers, in *Comprehensive Computational Chemistry*, eds. M. Yáñez and R. J. Boyd, Elsevier, Oxford, UK, 2024, pp. 306–321.
33 K. B. Wiberg, *Tetrahedron*, 1968, **24**, 1083–1096.
34 G. Frenking and S. Shaik, eds., *The Chemical Bond: Fundamental Aspects of Chemical Bonding*, Wiley–VCH, Weinheim, Germany, 2014.
35 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian 16, Revision C.01*, Gaussian, Inc., Wallingford, CT, 2016.
36 E. D. Glendening, J. K. Badenhoop, A. E. Reed, J. E. Carpenter, J. A. Bohmann, C. M. Morales, P. Karafiloglou, C. R. Landis and F. Weinhold, *NBO 7.0*, Theoretical Chemistry Institute, University of Wisconsin—Madison, Madison, WI, 2018.
37 J. R. Fresco and O. Amosova, *Annu. Rev. Biochem.*, 2017, **86**, 461–484.
38 F. Drabløs, E. Feyzi, P. A. Aas, C. B. Vaagbø, B. Kavli, M. S. Bratlie, J. Peña-Diaz, M. Otterlei, G. Slupphaug and H. E. Krokan, *DNA Repair*, 2004, **3**, 1389–1407.
39 K. S. Gates, *Chem. Res. Toxicol.*, 2009, **22**, 1747–1760.
40 Y. Liu, Y. Rodriguez, R. L. Ross, R. Zhao, J. A. Watts, C. Grunseich, A. Bruzel, D. Li, J. T. Burdick, R. Prasad, R. J. Crouch, P. A. Limbach, S. H. Wilson and V. G. Cheung, *Proc. Natl. Acad. Sci. U.S.A.*, 2020, **117**, 20689–20695.
41 J. A. Watts, C. Grunseich, Y. Rodriguez, Y. Liu, D. Li, J. T. Burdick, A. Bruzel, R. J. Crouch, R. W. Mahley,

S. H. Wilson and V. G. Cheung, *Nucleic Acids Res.*, 2022, **50**, 12497–12514.
42 C. S. Gutierrez, B. Silkenath, V. Kojasoy, J. A. Pich, D. C. Lim and R. T. Raines, *RNA*, 2025, **31**, 1542–1556.
43 D. Banerjee, L. Xiao, P. S. Thacker, J. Kundu, M. Manoharan and E. T. Kool, *ACS Cent. Sci.*, 2025, **11**, 2400–2409.
44 M. Álvarez-Moreno, C. de Graaf, N. López, F. Maseras, J. M. Poblet and C. Bo, *J. Chem. Inf. Model.*, 2015, **55**, 95–103.

