WORLD JOURNAL OF PHARMACEUTICAL RESEARCH
Coden USA: WJPRAP Impact Factor 8.63
Volume 15, Issue 22, 635-652. Research Article ISSN 2277–7105

COMPUTATIONAL INVESTIGATION OF THERMODYNAMIC STABILITY AND QUANTUM CHEMICAL REACTIVITY DESCRIPTORS OF ADENINE-INITIATED RNA TRINUCLEOTIDES: AN AM1/COSMO STUDY

Dr. Bojja Rajeshwar Rao*

Senior Chemist (Retired), Chemical Division, Kakatiya Thermal Power Project (O&M), Chelpur-506 170, Telangana State, India.

Article Received on 13 Oct. 2026,
Article Revised on 03 Nov. 2026,
Article Published on 16 Nov. 2026,
https://doi.org/10.5281/zenodo.24278696

***Corresponding Author**
Dr. Bojja Rajeshwar Rao
Senior Chemist (Retired), Chemical Division, Kakatiya Thermal Power Project (O&M), Chelpur-506 170, Telangana State, India.

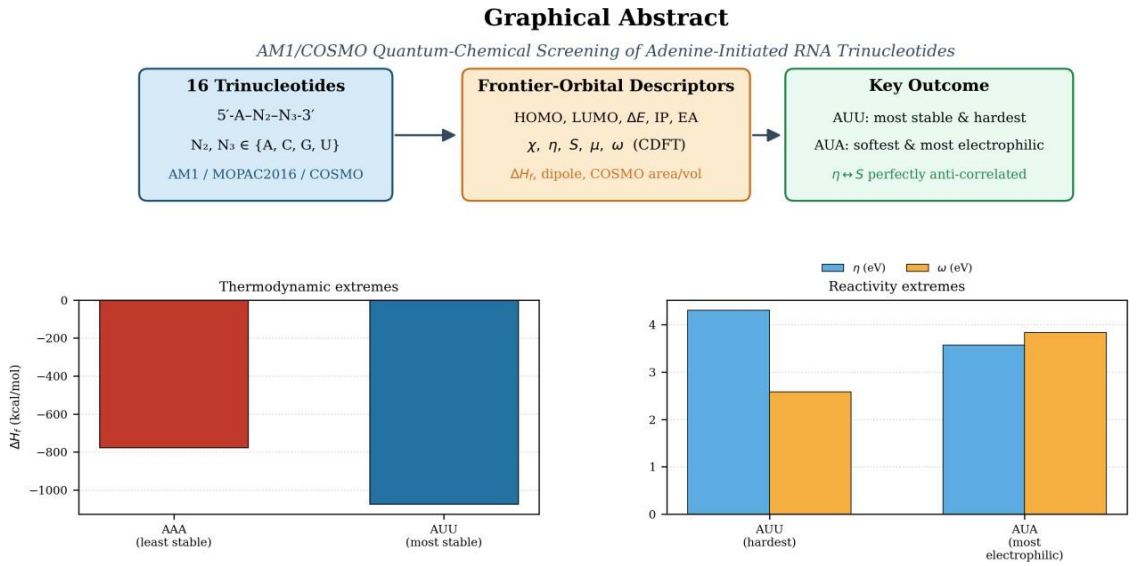
How to cite this Article: Dr. Bojja Rajeshwar Rao* (2026). Computational Investigation Of Thermodynamic Stability And Quantum Chemical Reactivity Descriptors Of Adenine-Initiated Rna Trinucleotides: An Am1/Cosmo Study. World Journal of Pharmaceutical Research, 15(22), 635–652
This work is licensed under Creative Commons Attribution 4.0 International license.

ABSTRACT

RNA trinucleotides provide minimal oligomeric models for probing the intrinsic electronic and thermodynamic effects of nucleobase sequence on nucleic-acid reactivity. In this study, the complete set of sixteen adenine-initiated RNA trinucleotides (5′-A–R_2–R_3-3′; R = A, C, G, U) was investigated using the semi-empirical Austin Model 1 (AM1) Hamiltonian as implemented in MOPAC2016 (v. 22.2.5w). Full geometry optimization was performed in an implicit aqueous environment using the Conductor-like Screening Model (COSMO). Heats of formation (ΔH_f), dipole moments, COSMO solvent-accessible surface areas and volumes, and frontier molecular orbital (FMO) energies were calculated for all sixteen sequences. These quantities were subsequently used to derive Koopmans'-theorem-based global conceptual density functional theory (CDFT) reactivity descriptors, including ionization potential (IP), electron affinity (EA), electronegativity (χ), chemical hardness (η), softness (S), chemical potential (μ), and electrophilicity index (ω). Among the investigated sequences, AUU exhibited the most negative heat of formation (ΔH_f = −1072.94 kcal mol^{-1}), the largest HOMO–LUMO gap (ΔE = 8.607 eV), and high chemical hardness (η = 4.304 eV), whereas AAA showed the least negative ΔH_f (−777.08 kcal mol^{-1}) and the largest dipole moment (15.949 D). AUA exhibited

the highest electrophilicity index (ω = 3.831 eV) and greatest softness (S = 0.140 eV^{-1}), consistent with its comparatively narrow HOMO–LUMO gap. Pearson correlation analysis revealed an inverse relationship between electrophilicity and chemical hardness (r = −0.64) and an essentially perfect inverse relationship between softness and the HOMO–LUMO gap (r = −1.00), reflecting the mathematical interdependence of these frontier-orbital-derived descriptors within the adopted CDFT framework. Overall, the results demonstrate that variation in the second and third nucleobases substantially modulates the calculated thermodynamic stability and electronic reactivity of adenine-initiated RNA trinucleotides, providing sequence-resolved computational descriptors for future investigations of structure–reactivity relationships and RNA–ligand interactions.

KEYWORDS: RNA trinucleotides, semi-empirical AM1 method, MOPAC2016, COSMO solvation, frontier molecular orbitals, HOMO–LUMO gap, conceptual density functional theory, global electrophilicity index.



Graphical Abstract: - Sixteen Adenine-initiated RNA trinucleotides were optimized at the AM1/COSMO level (MOPAC2016) and characterized by a complete set of frontier-orbital and conceptual-DFT reactivity descriptors, identifying AUU as the most stable/hardest and AUA as the softest/most electrophilic sequence.

INTRODUCTION

Ribonucleic acid (RNA) is no longer regarded solely as an intermediary between genomic information and protein synthesis. In addition to its canonical roles in messenger, ribosomal, and transfer RNA, the identification of catalytically active ribozymes, riboswitches,

microRNAs, and numerous long non-coding RNAs has established RNA as a structurally and functionally diverse macromolecule. Its chemical reactivity, conformational behaviour, and molecular-recognition properties are strongly influenced by nucleobase sequence and the resulting electronic structure.[1]

Trinucleotides provide particularly useful model systems for investigating sequence-dependent molecular properties while retaining a computationally tractable molecular size. The sixteen possible RNA trinucleotides beginning with adenine (5′-A–R_2–R_3-3′; R = A, C, G, U) encompass systematic variation in the identities of the second and third nucleobases. This complete sequence space therefore provides a controlled framework for examining how nucleobase composition influences thermodynamic stability, molecular polarity, and frontier-orbital characteristics. Although such short oligomers do not reproduce the conformational and environmental complexity of folded RNA, they provide well-defined molecular models for isolating intrinsic sequence-dependent electronic effects and establishing quantitative structure–property relationships.

Semi-empirical molecular-orbital methods offer a practical approach for systematic quantum-chemical investigation of nucleic-acid fragments. Among these methods, the Austin Model 1 (AM1) Hamiltonian has been widely applied to the calculation of molecular geometries and electronic properties of organic and biological molecules. Its relatively low computational cost makes it particularly suitable for the exhaustive characterization of complete sequence libraries. When combined with the Conductor-like Screening Model (COSMO), AM1 calculations can additionally provide an approximate treatment of solvent polarization effects, allowing molecular properties to be evaluated within an implicit aqueous environment.

Frontier molecular orbital (FMO) analysis provides a useful framework for examining relationships between electronic structure and chemical reactivity. Within conceptual density functional theory (CDFT), frontier-orbital energies can be used, under the Koopmans'-theorem approximation, to estimate global reactivity descriptors such as ionization potential (IP), electron affinity (EA), electronegativity (χ), chemical potential (μ), chemical hardness (η), softness (S), and the electrophilicity index (ω) introduced by Parr, von Szentpály, and Liu.[2] These descriptors have been widely used to characterize the electronic properties and relative reactivity of nucleobases and nucleic-acid systems. Nevertheless, to the best of our knowledge, a systematic sequence-resolved characterization of the complete set of sixteen adenine-initiated RNA trinucleotides, using a common computational protocol and an

integrated set of thermodynamic, structural, and frontier-orbital-derived descriptors, has not yet been reported. The present study therefore addresses this gap by providing a comprehensive AM1/COSMO characterization of all sixteen sequences and by examining the relationships among sequence composition, thermodynamic properties, molecular polarity, frontier-orbital energies, and CDFT reactivity descriptors.

The present study investigates all sixteen adenine-initiated RNA trinucleotides using the semi-empirical AM1 Hamiltonian as implemented in MOPAC2016, with COSMO implicit solvation, in continuous of research work.[3, 4, 5, 6] The specific objectives are to.

(i) Calculate and compare the heats of formation (ΔH_f), dipole moments, and COSMO-derived solvent-accessible surface areas and volumes for all sixteen sequences;

(ii) determine the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) energies and derive the corresponding Koopmans'-theorem-based global conceptual density functional theory (CDFT) reactivity descriptors; and

(iii)Statistically evaluate relationships among nucleobase sequence, calculated thermodynamic properties, molecular polarity, frontier-orbital characteristics, and global reactivity descriptors.

The resulting sequence-resolved dataset is intended to provide a quantitative computational reference for characterizing intrinsic sequence-dependent electronic effects in adenine-initiated RNA trinucleotides and to inform future investigations of RNA structure–reactivity relationships and RNA–ligand interactions.

MATERIALS AND METHODS

Sequence generation and molecular construction *Molecular Model Construction*

All sixteen possible adenine-initiated RNA trinucleotides of the general form 5′-A–R_2–R_3-3′ (R_2, R_3 ∈ {A, C, G, U}) were constructed: AAA, AAC, AAG, AAU, ACA, ACC, ACG, ACU, AGA, AGC, AGG, AGU, AUA, AUC, AUG, and AUU as per Diagram-1. Each trinucleotide was modelled as a single-stranded RNA fragment comprising three ribonucleotide residues connected by canonical 3′–5′ phosphodiester linkages, with a 5′-phosphate group and a 3′-hydroxyl group. The canonical nucleobase tautomer corresponding to Watson–Crick base-pairing forms were used throughout. The ribose moieties were initially constructed in a conventional β-D-ribofuranose conformation compatible with RNA structure. The resulting molecular structures were subsequently subjected to full geometry optimization using the semi-empirical Austin Model 1 (AM1) method.

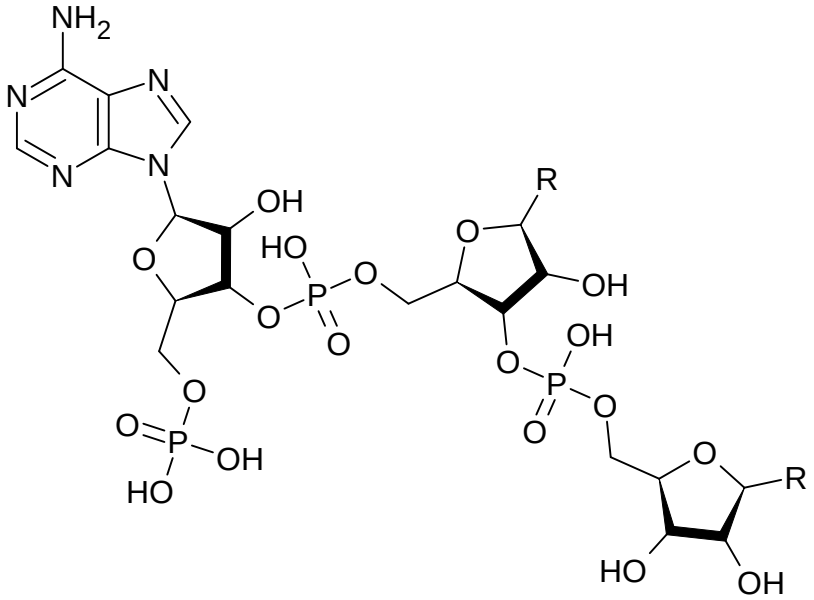


Diagram - 1: Adenine-initiated RNA-trinucleotides
R = Adenine, Guanine, Cytosine, Uracil

Computational method

Quantum-Chemical Calculations and Solvation Model

Full geometry optimization of all sixteen adenine-initiated RNA trinucleotides was performed using the semi-empirical Austin Model 1 (AM1) Hamiltonian[7], as implemented in MOPAC2016 (Version 22.2.5w).[8] AM1 was selected because its relatively low computational cost permits systematic characterization of the complete trinucleotide sequence library while providing a practical semi-empirical description of the electronic structure and thermochemical properties of nucleic-acid fragments. Previous applications of AM1 to nucleobases and nucleotide systems further demonstrate its utility for investigating the molecular and electronic properties of nucleic-acid components.[9]

Solvent effects were incorporated implicitly using the Conductor-like Screening Model (COSMO)[10], as implemented in MOPAC2016. The dielectric constant was set to that of bulk water (ε = 78.4). The COSMO calculations employed the MOPAC-defined solvent-accessible cavity with the surface discretization parameters NSPA = 92 and NPPA = 1082. Geometry optimizations were performed without symmetry constraints using the MOPAC eigenvector-following (EF) optimization procedure, with convergence monitored using the GNORM criterion. Optimization was continued until the specified convergence criteria were satisfied, and the final molecular geometries were inspected in the corresponding MOPAC output to confirm successful convergence.

For each optimized trinucleotide, the heat of formation (ΔH_f, kcal mol^{-1}), total dipole moment (D), COSMO solvent-accessible surface area ($Å^2$), COSMO cavity volume ($Å^3$), molecular weight (g mol^{-1}), and frontier molecular orbital energies namely, the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) were extracted from the MOPAC output. These calculated quantities constituted the primary molecular-property dataset and were subsequently used to derive the global reactivity descriptors and perform the statistical analyses.

Derivation of global reactivity descriptors

Within Koopmans' approximation, the vertical ionization potential (IP) and electron affinity (EA) of each trinucleotide were estimated directly from the frontier orbital energies as IP ≈ −E(HOMO) and EA ≈ −E(LUMO). The remaining conceptual density functional theory (CDFT) descriptors were then computed using the standard finite-difference expressions of Parr and co-workers.[2, 11]

Electronegativity: $\chi = (IP + EA) / 2$

Chemical potential: $\mu = -\chi = -(IP + EA) / 2$

Global hardness: $\eta = (IP - EA) / 2$

Global softness: $S = 1 / (2\eta)$

Electrophilicity index: $\omega = \mu^2 / (2\eta)$

The electrophilicity index ω, first proposed by Parr, von Szentpály, and Liu[2] and subsequently applied to nucleic acid bases[12], provides a single quantitative measure of the propensity of a molecule to accept electron density from the surrounding environment, and was used here as the principal descriptor for comparing the relative electrophilic character of the sixteen trinucleotide sequences.

Statistical and comparative analysis

The complete descriptor dataset, comprising heat of formation (ΔH_f), dipole moment, COSMO solvent-accessible surface area and cavity volume, molecular weight, HOMO and LUMO energies, HOMO–LUMO energy gap (ΔE), ionization potential (IP), electron affinity (EA), electronegativity (χ), chemical hardness (η), softness (S), chemical potential (μ), and electrophilicity index (ω), was compiled into a sequence-indexed data matrix containing all sixteen adenine-initiated RNA trinucleotides. Pairwise Pearson correlation coefficients (r) were calculated across the sixteen sequences to quantify linear associations among the molecular descriptors. Particular attention was given to relationships between the primary

quantum-chemical observables (ΔHf, dipole moment, COSMO surface area and volume, HOMO, and LUMO) and the derived global reactivity descriptors, thereby enabling sequence-dependent relationships among thermodynamic, structural, and electronic properties to be assessed.

Correlations among descriptors mathematically derived from the same frontier-orbital quantities were interpreted primarily as measures of their expected internal relationships rather than as independent evidence of chemical causation. Statistical analyses and figure generation were performed using standard scientific Python libraries, including pandas, NumPy, and Matplotlib.

RESULTS AND DISCUSSION

Thermodynamic Stability

The calculated heats of formation (ΔH_f) of the sixteen adenine-initiated RNA trinucleotides span nearly 300 kcal mol^{-1}, ranging from −1072.94 kcal mol^{-1} for AUU to −777.08 kcal mol^{-1} for AAA (Table - 1; Figure - 1A). This substantial range indicates pronounced sequence dependence in the calculated thermochemical properties of the trinucleotide series. Thus, variation in the identities of the second and third nucleobases produces differences in the calculated ΔH_f values that cannot be attributed solely to molecular composition or size, indicating an important contribution from sequence-dependent electronic and structural effects.

The uracil-terminated sequences AAU, ACU, AGU, and AUU generally exhibit more negative heats of formation than several corresponding sequences terminating in adenine, cytosine, or guanine. This pattern suggests that incorporation of uracil at the 3′-terminal position is associated with more favorable calculated thermochemical energetics within the optimized AM1/COSMO structures. However, the origin of this trend cannot be assigned uniquely to steric effects, hydrogen bonding, base-stacking interactions, or solvation on the basis of ΔH_f alone. Rather, the observed differences may reflect the combined influence of sequence-dependent molecular geometry, intramolecular interactions, electronic structure, and implicit solvation.

AAA represents the least thermochemical favorable sequence in the series, with a calculated heat of formation (ΔH_f) of −777.08 kcal mol^{-1}, and also exhibits the largest dipole moment (15.949 D; Figure - 1B). The simultaneous occurrence of the least negative ΔH_f and the

largest molecular dipole moment distinguishes the all-adenine sequence from the mixed-base trinucleotides in terms of its calculated thermochemical and electronic properties. The large dipole moment indicates a pronounced net polarity arising from the charge distribution of the optimized molecular structure. However, the origin of this enhanced polarity cannot be assigned solely from the present descriptors; structural and energetic analyses would be required to determine whether it is associated with particular stacking geometries, conformational effects, intramolecular interactions, or other sequence-dependent features.

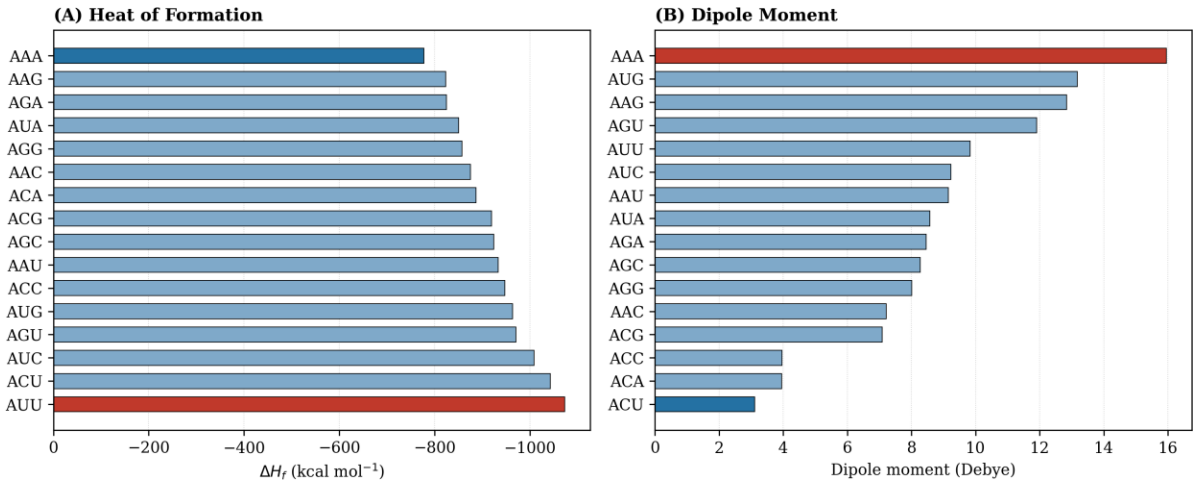


Figure – 1: (A) Computed heats of formation (ΔH_f, kcal mol^{-1}) and (B) total dipole moments (Debye) for the sixteen Adenine-initiated RNA trinucleotides at the AM1/COSMO level. In each panel the most extreme value is highlighted in red (least favorable / highest) and blue (most favorable / lowest).

Molecular polarity and COSMO solvation descriptors

The calculated dipole moments range from 3.109 D for ACU to 15.949 D for AAA, representing an approximately five-fold variation across the sixteen sequences. This substantial variation highlights the pronounced sequence dependence of the molecular charge distribution and overall polarity of the optimized trinucleotides. Notably, the three sequences with the lowest dipole moments of ACU, ACC, and ACA, all contain cytosine at the second position, whereas AAA, AGU, and AUG exhibit comparatively large dipole moments. These observations suggest that substitutions at the second and third nucleobase positions can substantially alter the spatial distribution of polar functional groups and, consequently, the resultant molecular dipole moment.

The COSMO-derived solvent-accessible surface areas range from 627 to 758 Å^2, while the corresponding cavity volumes span 750–1052 Å^3. Both descriptors exhibit moderate positive

correlations with the calculated heats of formation (Pearson's $r = 0.67$ for surface area and $r = 0.66$ for cavity volume; Figure 4A). Thus, within the present dataset, trinucleotides with larger COSMO-defined molecular dimensions tend to exhibit less negative heats of formation. These correlations indicate an association between molecular geometry and thermochemical properties but do not establish a causal relationship. The observed trends may reflect sequence-dependent differences in optimized molecular conformations, solvent-accessible geometry, and intramolecular interactions. Nevertheless, COSMO surface area and cavity volume are geometric descriptors and cannot independently establish the extent or energetic contribution of base stacking or hydrogen bonding. Further structural analysis of the optimized geometries would therefore be necessary to determine whether variations in these interactions contribute to the observed thermochemical trends.

Frontier molecular orbital analysis

The HOMO and LUMO energies of the sixteen adenine-initiated RNA trinucleotides are summarized in Table - 2 and compared in the frontier-orbital energy-level diagram (Figure - 2). The HOMO energies span a relatively narrow range, from −8.210 to −9.168 eV, whereas the LUMO energies exhibit a substantially broader sequence-dependent variation, ranging from −0.306 eV for AUG to −1.661 eV for AUA. These results indicate that, within the present AM1/COSMO dataset, substitutions at the second and third nucleobase positions influence the calculated LUMO energies more strongly than the HOMO energies. This behavior suggests greater sequence sensitivity of the unoccupied frontier-orbital energies, while the occupied frontier-orbital energies remain comparatively constrained. The observation is broadly consistent with the established sensitivity of nucleobase frontier-orbital energies to molecular structure and substitution, although direct comparison with previous studies requires consideration of the computational methods and molecular environments employed.[13]

The resulting HOMO–LUMO energy gaps (ΔE) range from 7.132 eV for AUA to 8.607 eV for AUU, corresponding to a variation of 1.475 eV across the sequence set (Figure - 3). This substantial variation demonstrates that nucleobase substitutions at the second and third positions modulate the calculated frontier-orbital separation of the adenine-initiated RNA trinucleotides. Within the conceptual framework of frontier-orbital theory and conceptual density functional theory, a smaller HOMO–LUMO gap is generally associated with lower chemical hardness and greater global softness, whereas a larger gap indicates greater hardness and reduced softness. Accordingly, AUA, with the smallest calculated gap, exhibits

the greatest softness expected from the adopted gap-based relationship, while AUU, with the largest gap, exhibits the greatest hardness. These trends provide an electronic-structure basis for interpreting the sequence-dependent variations in the global reactivity descriptors discussed in the following section. However, the calculated gap should be interpreted as a qualitative indicator of electronic structure and potential reactivity rather than as a direct measure of experimental excitation energies or kinetic reactivity.

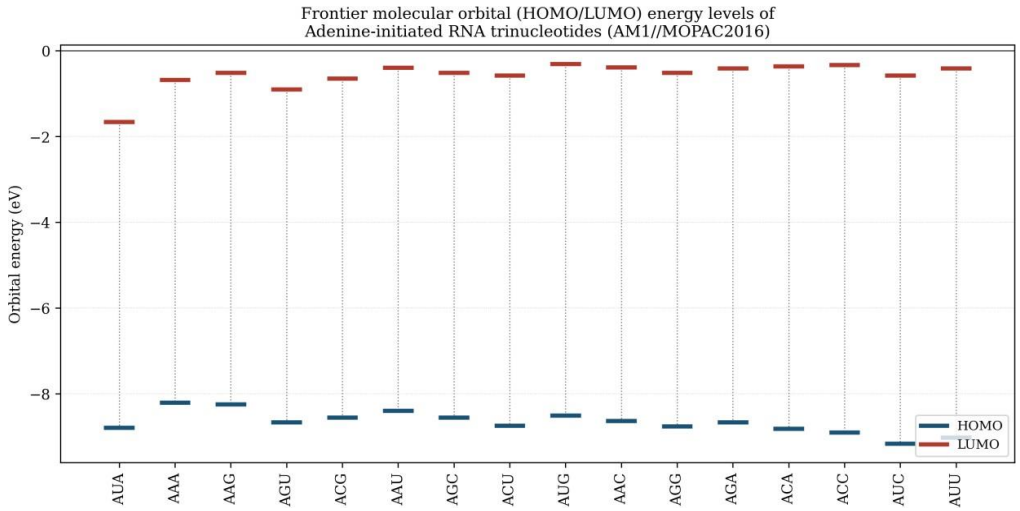


Figure – 2: Frontier molecular orbital (HOMO, blue; LUMO, red) energy-level diagram for the sixteen Adenine-initiated RNA trinucleotides, ordered by increasing HOMO–LUMO gap from left to right.

Global reactivity descriptors

Table - 2 summarizes the complete set of global reactivity descriptors derived using the Koopmans'-theorem approximation. AUU, which exhibits the largest HOMO–LUMO energy gap (Δ E = 8.607 eV), also possesses the highest chemical hardness in the series (η = 4.304 eV) and a comparatively low electrophilicity index (ω = 2.583 eV). These results are consistent with the general conceptual-DFT expectation that greater chemical hardness is associated with lower global softness and reduced susceptibility to electronic redistribution. This interpretation is broadly aligned with the maximum hardness principle [11, 14], although the principle should not be regarded as a universal predictor of molecular stability or reactivity under all conditions.

In contrast, AUA exhibits the smallest HOMO–LUMO gap (Δ E = 7.132 eV), the lowest chemical hardness (η = 3.566 eV), and the highest global softness ((S = 0.140 eV^{-1}) among the sixteen sequences. AUA also has the highest electrophilicity index (ω = 3.831 eV), suggesting

a comparatively greater capacity for electronic stabilization through charge acceptance within the adopted conceptual-DFT framework. The concurrent occurrence of relatively low hardness, high softness, and high electrophilicity identifies AUA as an electronically distinctive sequence in the present dataset. These descriptors provide complementary perspectives on its potential charge-transfer behavior and are broadly compatible with electrophilicity-based descriptions of nucleobase interactions.[12] Nevertheless, they should be interpreted as qualitative indicators of electronic properties rather than direct evidence of enhanced chemical reactivity in an RNA strand. Establishing sequence-specific reactivity would require further investigation using explicit intermolecular interaction energies, charge-transfer analyses, or reaction-path calculations.

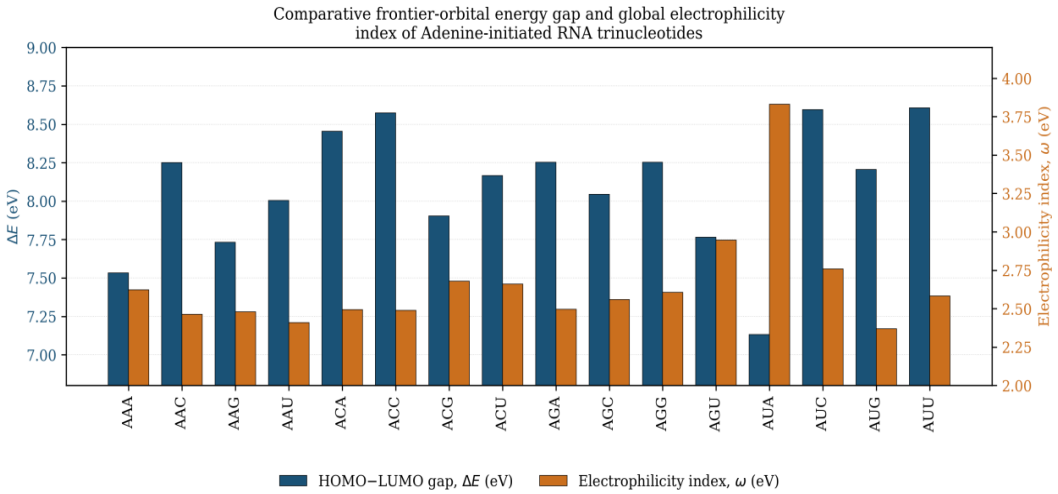


Figure – 3: Comparative bar chart of the HOMO–LUMO energy gap (ΔE, left axis, blue) and the global electrophilicity index (ω, right axis, orange) for each of the sixteen Adenine-initiated RNA trinucleotides.

Correlation analysis and structure–reactivity relationships

Pearson correlation analysis of the complete descriptor matrix (Figure - 4A) reveals the expected mathematical relationships among the conceptual density functional theory (CDFT) descriptors derived from the HOMO and LUMO energies. Chemical hardness and global softness are perfectly inversely correlated ((r=-1.00)), consistent with their reciprocal dependence on the HOMO–LUMO energy gap under the adopted definitions. Electrophilicity exhibits a strong positive correlation with electronegativity ((r=0.90)) and a strong negative correlation with chemical potential and electron affinity (r=-0.90) in each case. These associations largely reflect the mathematical interdependence of the derived descriptors and should therefore be interpreted primarily as indicators of internal consistency rather than as

independent chemical relationships.

In contrast, the correlations of electrophilicity with heat of formation ((r=0.08)) and dipole moment ((r=0.02)) are negligible in the present dataset. These findings indicate that the calculated electrophilicity index has little linear association with either the heat of formation or the overall molecular dipole moment across the sixteen adenine-initiated RNA trinucleotides. Thus, thermochemical properties and frontier-orbital-derived reactivity descriptors capture distinct aspects of sequence-dependent molecular behavior. A sequence with a relatively favorable calculated heat of formation may nevertheless exhibit a comparatively high electrophilicity index, and vice versa, depending on its electronic structure. Similarly, the magnitude of the molecular dipole moment does not, by itself, predict the electrophilicity index.

The hardness–softness relationship presented in Figure - 4B further illustrates the systematic variation in the frontier-orbital-derived descriptors across the sequence series. The four AU-terminated sequences AUA, AUC, AUG, and AUU share the same first two nucleobases but differ at the 3′-terminal position. Their distribution across the hardness–softness plot demonstrates how substitution of the terminal nucleobase can alter the calculated electronic descriptors while the first two positions remain unchanged. This comparison highlights the contribution of the 3′-terminal base to sequence-dependent frontier-orbital characteristics within the present AM1/COSMO computational model.

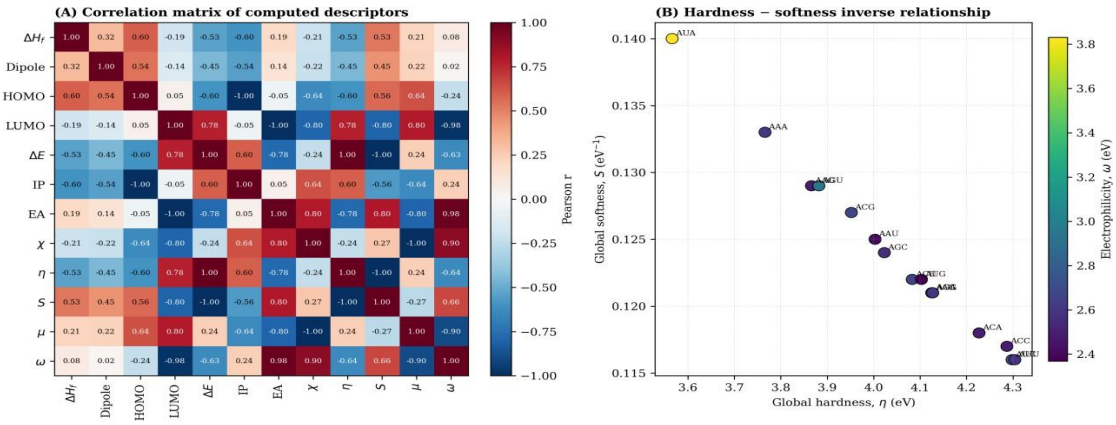


Figure – 4: (A) Pearson correlation heat-map among all computed thermodynamic, geometric, and CDFT reactivity descriptors. (B) Global hardness (η) versus global softness (S) for the sixteen trinucleotides, coloured by electrophilicity index (ω).

Methodological workflow and structural interpretation

The overall computational workflow adopted in this study is summarized schematically in Figure - 5. The procedure encompasses enumeration of the complete sequence set, AM1/COSMO geometry optimization, extraction of frontier molecular orbital energies, calculation of Koopmans'-theorem-based global reactivity descriptors, and subsequent statistical correlation and structure–property analyses.

Figure - 6 presents a qualitative schematic representation of the sugar–phosphate backbone of a generic adenine-initiated RNA trinucleotide and illustrates the conceptual interpretation of frontier-orbital-derived reactivity. In this representation, regions of high highest occupied molecular orbital (HOMO) density are associated with electron-rich character and potential nucleophilic behavior, whereas regions of high lowest unoccupied molecular orbital (LUMO) density indicate locations that may be susceptible to electron acceptance and nucleophilic attack. These interpretations provide a conceptual framework for understanding the potential electronic reactivity of the trinucleotides.

However, the calculated HOMO and LUMO energies provide information about the relative energetic positions of the frontier orbitals but do not, by themselves, establish their spatial localization within the molecular structure. Consequently, any proposed preferential localization of the HOMO near the 5′-terminal adenine residue or of the LUMO near the 3′ terminus should be regarded as a hypothesis rather than a demonstrated finding. Establishing sequence-specific orbital localization would require explicit visualization of the HOMO and LUMO distributions for the optimized geometries, supported where appropriate by quantitative analyses of orbital contributions from individual nucleobases and sugar–phosphate regions. Such analyses would help determine whether the proposed terminal localization is supported by the calculated electronic structures.

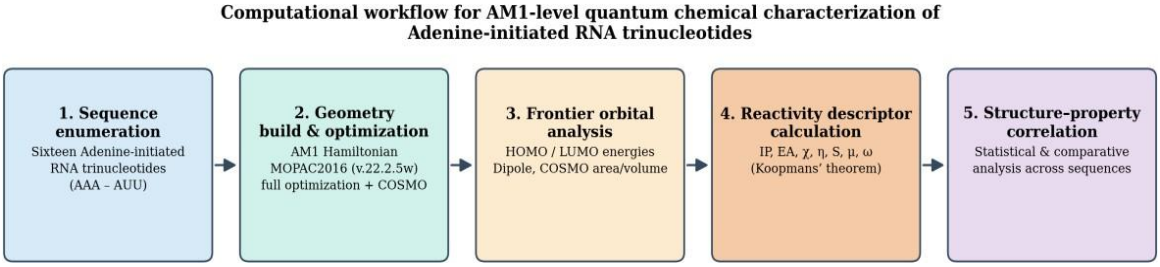


Figure – 5: Schematic computational workflow used for the sequence-resolved AM1/COSMO characterization of the sixteen Adenine-initiated RNA trinucleotides, from sequence enumeration to structure–reactivity correlation.

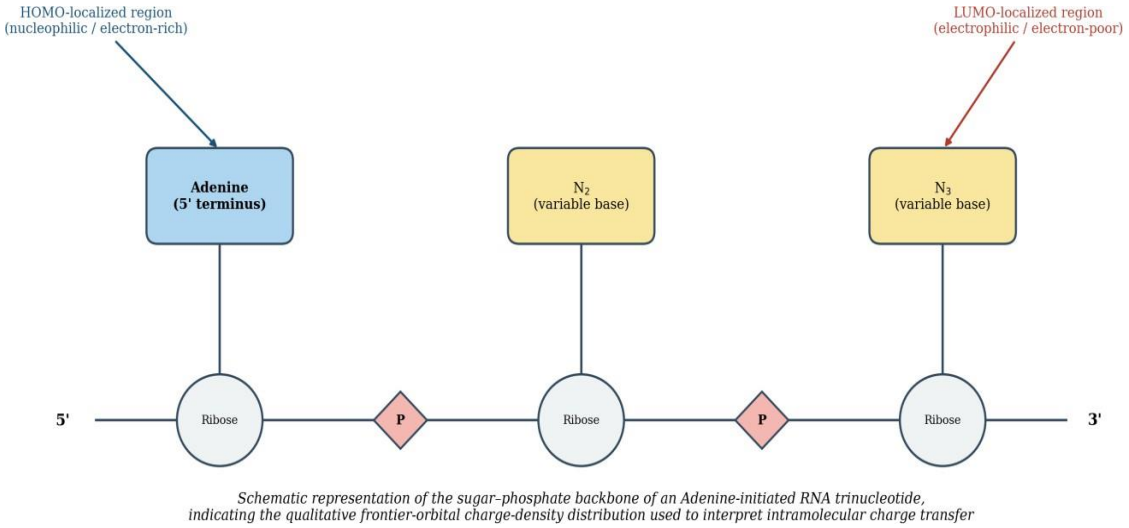


Figure – 6: Schematic representation of the RNA trinucleotide sugar–phosphate backbone, indicating the qualitative frontier-orbital charge-density distribution (HOMO-versus LUMO-localized regions) used to interpret the intramolecular charge-transfer trends discussed.

Tables

Table – 1: Computational study on thermodynamic properties of Adenine-initiated RNA trinucleotides by the Austin Model 1 (AM1) method using MOPAC2016 (Version 22.2.5w).

Adenine-initiated RNA trinucleotides	ΔHf (kcal/mol)	Dipole (Debye)	HOMO (eV)	LUMO (eV)	COSMO Area (Å²)	COSMO Volume (Å³)	Mol. Wt.
AAA	-777.0797	15.949	-8.210	-0.678	717.42	1032.48	1005.6406
AAC	-874.7115	7.221	-8.632	-0.382	739.90	970.57	981.6156
AAG	-823.3932	12.843	-8.244	-0.513	726.74	1051.10	1021.6400
AAU	-932.8388	9.154	-8.395	-0.390	696.97	792.87	982.6004
ACA	-886.5045	3.956	-8.818	-0.364	738.47	972.50	981.6156
ACC	-947.0159	3.969	-8.907	-0.333	709.27	945.70	957.5906
ACG	-919.0751	7.083	-8.552	-0.649	707.95	998.62	997.6150
ACU	-1042.8099	3.109	-8.744	-0.579	627.47	946.51	958.5754
AGA	-824.4023	8.459	-8.666	-0.413	758.26	1009.12	1021.6400
AGC	-923.8897	8.281	-8.559	-0.514	637.53	965.46	997.6150
AGG	-856.8803	8.016	-8.764	-0.511	757.91	1012.72	1037.6394
AGU	-970.2586	11.901	-8.665	-0.901	667.20	986.76	998.5998
AUA	-850.6102	8.579	-8.793	-1.661	683.48	980.15	982.6004
AUC	-1008.4418	9.238	-9.168	-0.572	701.13	938.11	958.5754
AUG	-963.1318	13.170	-8.511	-0.306	700.53	984.60	998.5998
AUU	-1072.9379	9.831	-9.019	-0.412	662.18	749.79	959.5602

Table – 2: Global quantum chemical reactivity descriptors derived from the frontier molecular orbital energies of Adenine-initiated RNA trinucleotides.

Adenine-initiated RNA trinucleotides	HOMO (eV)	LUMO (eV)	ΔE (eV)	IP (eV)	EA (eV)	EN (χ) (eV)	η (eV)	S (eV^{-1})	μ (eV)	ω (eV)
AAA	-8.210	-0.678	7.532	8.210	0.678	4.444	3.766	0.133	-4.444	2.622
AAC	-8.632	-0.382	8.250	8.632	0.382	4.507	4.125	0.121	-4.507	2.462
AAG	-8.244	-0.513	7.731	8.244	0.513	4.379	3.866	0.129	-4.379	2.480
AAU	-8.395	-0.390	8.005	8.395	0.390	4.393	4.003	0.125	-4.393	2.410
ACA	-8.818	-0.364	8.454	8.818	0.364	4.591	4.227	0.118	-4.591	2.493
ACC	-8.907	-0.333	8.574	8.907	0.333	4.620	4.287	0.117	-4.620	2.489
ACG	-8.552	-0.649	7.903	8.552	0.649	4.601	3.952	0.127	-4.601	2.678
ACU	-8.744	-0.579	8.165	8.744	0.579	4.662	4.083	0.122	-4.662	2.661
AGA	-8.666	-0.413	8.253	8.666	0.413	4.540	4.127	0.121	-4.540	2.497
AGC	-8.559	-0.514	8.045	8.559	0.514	4.537	4.023	0.124	-4.537	2.558
AGG	-8.764	-0.511	8.253	8.764	0.511	4.638	4.127	0.121	-4.638	2.606
AGU	-8.665	-0.901	7.764	8.665	0.901	4.783	3.882	0.129	-4.783	2.947
AUA	-8.793	-1.661	7.132	8.793	1.661	5.227	3.566	0.140	-5.227	3.831
AUC	-9.168	-0.572	8.596	9.168	0.572	4.870	4.298	0.116	-4.870	2.759
AUG	-8.511	-0.306	8.205	8.511	0.306	4.409	4.103	0.122	-4.409	2.369
AUU	-9.019	-0.412	8.607	9.019	0.412	4.716	4.304	0.116	-4.716	2.583

ΔE = HOMO–LUMO gap; IP = ionization potential; EA = electron affinity; EN (χ) = electronegativity; η = global hardness; S = global softness; μ = chemical potential; ω = global electrophilicity index. All quantities derived from Koopmans' theorem.

CONCLUSIONS

A systematic AM1/COSMO quantum-chemical investigation of the complete set of sixteen adenine-initiated RNA trinucleotides has established a sequence-resolved dataset comprising heats of formation, dipole moments, COSMO-derived molecular surface areas and cavity volumes, frontier molecular orbital energies, and Koopmans'-theorem-based conceptual density functional theory (CDFT) global reactivity descriptors. The principal findings are as follows.

(i) Calculated heat of formation varies by nearly 300 kcal mol^{-1} across the sequence set, with AUU exhibiting the most negative value and AAA the least negative value;

(ii) LUMO energies exhibit greater sequence-dependent variation than HOMO energies, resulting in substantial differences in the HOMO–LUMO energy gap;

(iii) AUU combines the most negative calculated heat of formation with the largest HOMO–LUMO gap and highest chemical hardness, whereas AUA exhibits the smallest gap, lowest hardness, greatest global softness, and highest electrophilicity index; and

(iv) Electrophilicity exhibits negligible linear correlations with heat of formation and dipole moment, indicating that thermochemical properties, molecular polarity, and frontier-orbital-

derived reactivity capture distinct aspects of sequence-dependent molecular behavior.

Collectively, these findings demonstrate that substitution of the second and third nucleobases can substantially modulate the calculated thermochemical and electronic properties of RNA trinucleotides while preserving the same 5′-adenine residue. The resulting sequence-resolved descriptors provide a computational reference for future investigations of RNA–ligand interactions, charge-transfer phenomena, and sequence-dependent electronic reactivity, including potential susceptibility to oxidative or electrophilic modification. However, such applications require explicit modelling of the relevant interacting species, solvent effects, and, where appropriate, reaction pathways. The present calculations characterize isolated trinucleotides within the adopted AM1/COSMO framework and do not independently establish experimental RNA stability, biological reactivity, or modification susceptibility.

Future work should extend this protocol to guanine-, cytosine-, and uracil-initiated RNA trinucleotide families and benchmark representative structures against higher-level electronic-structure methods and available experimental redox or reactivity data. Explicit frontier-orbital visualization, charge-distribution analysis, and calculations of intermolecular interactions would further clarify the structural origins of the observed sequence-dependent trends. Such extensions will help establish the transferability, predictive utility, and limitations of AM1/COSMO-derived descriptors for the computational characterization of RNA sequence-dependent properties.

Declaration

The authors declare no conflict of interest.

Funding

The author received no financial support for the research work, and publication of this article.

Author’s contribution

The author has participated in conception, design, analysis, interpretation of the data, drafting the article, and approval of the final version.

ACKNOWLEDGEMENT

The author is grateful to the assistance of **Claude AI** and **ChatGPT** in improving the presentation of the manuscript. The author takes full responsibility for the scientific content, data analysis, interpretations, and conclusions presented in this article.

REFERENCES

1. Cech, T. R.; Steitz, J. A. The Noncoding RNA Revolution—Trashing Old Rules to Forge New Ones. *Cell*, 2014; 157(1): 77–94. https://doi.org/10.1016/j.cell.2014.03.008
2. Parr, R. G.; von Szentpály, L.; Liu, S. Electrophilicity Index. *J. Am. Chem. Soc*, 1999; 121(9): 1922–1924. https://doi.org/10.1021/ja983494x
3. Rajeshwar Rao, B., Sequence-dependent thermodynamic and quantum chemical descriptors of adenine-initiated DNA trinucleotides: a semi-empirical AM1/MOPAC2016 investigation, *World J Pharmacy and Pharmaceutical Research*, 2026; 15(17): 1821-1834. https://doi.org/10.5281/zenodo.22267154 www.wjpr.net
4. Rajeshwar Rao, B., Computational study of thermodynamic stability and frontier-orbital-based reactivity descriptors of guanine-initiated DNA trinucleotides: an AM1/MOPAC2016 investigation, *World J Pharmacy and Pharmaceutical Sciences*, 2026; 15(9): 843-857. https://doi.org/10.5281/zenodo.22252151
5. Rajeshwar Rao, B., Computational insights into the thermodynamic stability and quantum-chemical reactivity of cytosine-initiated DNA trinucleotides: an AM1 semi-empirical study, *World J Pharmacy and Pharmaceutical Sciences*, 2026; 15(9): 1094-1107. https://doi.org/10.5281/zenodo.22688003
6. Rajeshwar Rao, B., Thermodynamic stability and frontier molecular orbital-derived reactivity descriptors of thymine-initiated DNA trinucleotides: a semi-empirical AM1 (MOPAC2016) study, *World J Pharmacy and Pharmaceutical Research*, 2026; 15(18): 915-932. https://doi.org/10.5281/zenodo.22768734 www.wjpr.net
7. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. Development and Use of Quantum Mechanical Molecular Models. 76. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.*, 1985; 107(13): 3902–3909. https://doi.org/10.1021/ja00299a024
8. Stewart, J. J. P. *MOPAC2016*, Version 22.2.5w; Stewart Computational Chemistry: Colorado Springs, CO, USA, 2016. http://OpenMOPAC.net
9. Ünal Civcir, P. A Theoretical Study of Tautomerism of Cytosine, Thymine, Uracil and Their 1-Methyl Analogues in the Gas and Aqueous Phases Using AM1 and PM3. *J. Mol. Struct. (THEOCHEM),* 2000; 532: 157–169. https://doi.org/10.1016/S0166-1280(00)00556-X
10. Klamt, A.; Schüürmann, G. COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *J. Chem. Soc., Perkin Trans*, 1993; *2:* 799–805. https://doi.org/10.1039/P29930000799

11. Parr, R. G.; Pearson, R. G. Absolute Hardness: Companion Parameter to Absolute Electronegativity. *J. Am. Chem. Soc*, 1983; 105(26): 7512–7516. https://doi.org/10.1021/ja00364a005
12. Padmanabhan, J.; Parthasarathi, R.; Subramanian, V.; Chattaraj, P. K. Electrophilicity-Based Charge Transfer Descriptor. *J. Phys. Chem. A.*, 2007; 111(7): 1358–1361. https://doi.org/10.1021/jp0649549
13. Foster, M. E.; Wong, B. M. Nonempirically Tuned Range-Separated DFT Accurately Predicts Both Fundamental and Excitation Gaps in DNA and RNA Nucleobases. *J. Chem. Theory Comput*, 2012; 8(8): 2682–2687. https://doi.org/10.1021/ct300420f
14. Pearson, R. G. Absolute Electronegativity and Hardness Correlated with Molecular Orbital Theory. *Proc. Natl. Acad. Sci. U.S.A.*, 1986; 83(22): 8440–8441. https://doi.org/10.1073/pnas.83.22.8440