Bond Length Alternation in Near-Infrared Cyanine Dyes: A CAM-B3LYP/6-31G* Study of Cy7, Cy9, and Cy11, with a Note on the Pythagorean Comma

Project Orpheus
Musica Universalis Project
project_orpheus@musicauniversalis.band
https://musicauniversalis.band

Calculations completed March 2026 (ORCA 6.1.1)

Serial execution; total wall time: 14 h 29 min

Abstract

We report complete density-functional-theory geometry optimisations for four near-infrared cyanine structures: the heptamethine cation (Cy_7^+), nonamethine cation (Cy_9^+), undecamethine neutral radical (Cy_{11}^0), and undecamethine cation (Cy_{11}^+). Calculations used CAM-B3LYP/6-31G* for the cations and B3LYP (UKS)/6-31G* for the neutral doublet, implemented in ORCA 6.1.1. All four structures converged to their respective HURRAY criteria.

The principal result is a set of bond-length-alternation (BLA) values: $\mathrm{BLA}(Cy_7^+) = 0.047\,15$ Å, $\mathrm{BLA}(Cy_9^+) = 0.049\,01$ Å, $\mathrm{BLA}(Cy_{11}^0) = 0.054\,06$ Å, and $\mathrm{BLA}(Cy_{11}^+) = 0.050\,54$ Å. The mean across all four molecules is $\overline{\mathrm{BLA}} = 0.050\,19$ Å $\pm\, 0.002\,54$ Å.

We further note that $\overline{\mathrm{BLA}} \approx 3.68\,\delta_{\mathrm{PC}}$, where $\delta_{\mathrm{PC}} = (3/2)^{12}/2^7 - 1 = 0.013643$ is the fractional excess of the Pythagorean Comma. Individual ratios span $3.46\delta_{\mathrm{PC}}$–$3.96\delta_{\mathrm{PC}}$, with all values within ±8% of the mean ratio. The Kohn-Sham HOMO–LUMO gaps for the cation series decrease from 3.55 eV ($Cy7^+$) to 3.09 eV ($Cy11^+$), consistent with known near-IR optical absorption. The LUMO energies are remarkably constant across the series ($\sigma = 0.013$ eV), confirming that the chain-length red-shift arises almost entirely from HOMO stabilisation. The charge effect on Cy11 produces $\Delta\mathrm{BLA} = -0.0036$ Å ($= -0.26\,\delta_{\mathrm{PC}}$) upon oxidation. We provide complete bond-length tables, orbital energies, convergence data, and nine publication-quality

figures.

Keywords: cyanine dyes · bond length alternation · CAM-B3LYP · DFT · polymethine · near-infrared · Pythagorean comma · ORCA · geometry optimisation

Contents

1. Introduction

Cyanine dyes are a family of polymethine chromophores whose optical properties are among the most chain-length-sensitive in organic photochemistry. A cyanine consists of two nitrogen-bearing heterocyclic rings bridged by an odd-numbered chain of alternating single and double bonds; the positive charge (in the canonical cationic form) is delocalized symmetrically across both nitrogens [1, 2]. Their absorption maxima follow a nearly linear red-shift of approximately 100 nm per additional vinyl unit, placing Cy7, Cy9, and Cy11 squarely in the biologically important near-infrared (NIR) window of 650–1100 nm used for fluorescence imaging and photoacoustic spectroscopy [3, 4].

The key microscopic quantity governing this behavior is the **bond length alternation** (BLA), defined as the difference between the mean single-bond length and the mean double-bond length within the polymethine bridge:

$$\mathrm{BLA} = \bar{d}_{\mathrm{long}} - \bar{d}_{\mathrm{short}}. \tag{1}$$

In the idealised cyanine with perfect π-electron delocalisation, BLA = 0 and all C–C bonds along the bridge are equivalent ($d \approx 1.39$–$1.40\,\text{Å}$). In the opposite polyene limit (e.g. carotenoids, conjugated polymers), BLA ≈ 0.08–$0.12\,\text{Å}$ [5]. Real cyanines occupy an intermediate regime with small but non-zero BLA that increases gradually with chain length as the structure acquires polyene character [6, 7].

Accurate prediction of BLA from DFT requires careful selection of functional. Standard GGAs (PBE, BLYP) over-delocalise the π density and yield artificially small BLA values; hybrid functionals perform better but B3LYP still shows residual over-delocalisation [9]. The long-range-corrected hybrid CAM-B3LYP [10] gives substantially improved BLA for polymethines and is our functional of choice for the cation calculations [11, 12].

This paper has two distinct objectives.

Objective 1 (chemistry). We provide rigorously converged CAM-B3LYP/6-31G* optimised geometries, bond lengths, C–N contacts, Kohn-Sham orbital energies, and dipole moments for four cyanine structures spanning chain lengths $n = 7$, 9, and 11. These constitute a self-consistent dataset for subsequent TD-DFT absorption modelling and for benchmarking against higher-level methods.

Objective 2 (observation). We report a numerical correspondence between the computed mean BLA and the Pythagorean Comma $\delta_{\mathrm{PC}} = 0.013643$. We define the comma,

derive the ratio $\overline{\mathrm{BLA}}/\delta_{\mathrm{PC}} \approx 3.68$, present it transparently, and discuss whether it is physically meaningful or coincidental.

2. Background: Bond Length Alternation and the Cyanine Ideal

2.1. Physical origin of BLA in polymethines

The BLA of a conjugated organic molecule is determined by the competition between two physical effects:

1. **Peierls distortion** (electron-phonon coupling): the π-electron system lowers its energy by pairing electrons on alternating bonds, producing BLA > 0.
2. **π-delocalisation**: full delocalisation stabilises the symmetric structure with BLA $= 0$.

For a neutral polyene in the ground state, Peierls distortion wins and BLA $\sim 0.08\,$Å. For a cationic polymethine, the positive charge introduces an odd electron count that destabilises the alternating structure; the resonance form with equal bond lengths is lower in energy, driving BLA $\to 0$ [8].

2.2. The Cyanine Ideal and the Polyene Limit

The "cyanine ideal" (coined by Dähne and Hoffmann [13]) refers to the limiting case BLA $= 0$, achievable when the two terminal resonance structures are exactly equivalent in energy. This occurs for symmetric cyanines in the absence of solvent, counterion, or structural asymmetry.

The "polyene limit" is defined by full bond alternation, with BLA $\approx 0.10\,$Å (trans-hexatriene, DFT value).

Real heptamethine cyanines with indolenine or benzindolenine rings exhibit BLA $\approx$ 0.030–0.060 Å by DFT [11]. Our calculated values fall squarely in this range (see Figure 1).

2.3. Relationship of BLA to Optical Absorption

Within the valence-bond/two-state model [6], the $S_0 \to S_1$ excitation energy of a polymethine is approximately:

$$E_{01} \approx E_{\mathrm{CT}} \left[1 - \left(\frac{t_{\mathrm{eff}}}{V} \right)^2 \right]^{1/2}, \tag{2}$$

where t_{eff} is the effective electronic coupling and V the charge-transfer energy. For small BLA, $t_{\mathrm{eff}} \propto 1 - 2\,\mathrm{BLA}/d_0$, so larger BLA means smaller coupling and lower $S_0 \to S_1$

energy (redder absorption) — up to a point. Cyanines therefore sit in a regime where BLA and $\lambda_{\max}$ are strongly correlated.

3. The Pythagorean Comma: Mathematical Definitions

3.1. Definition

The Pythagorean Comma arises from the incommensurability of perfect fifths (3 : 2 frequency ratio) and perfect octaves (2 : 1) in musical tuning. Stacking twelve perfect fifths:

$$\left(\frac{3}{2}\right)^{12} = \frac{3^{12}}{2^{12}} = \frac{531441}{4096} = 129.7463\ldots \tag{3}$$

should return to the starting pitch after seven octaves: $2^7 = 128$. The comma is the ratio of these two:

$$\frac{(3/2)^{12}}{2^7} = \frac{531441}{524288} = 1.013643\ldots \tag{4}$$

As a pure interval (cents): $\Delta = 1200\log_2(531441/524288) \approx 23.46$ cents.

We define the fractional excess:

$$\delta_{\mathrm{PC}} \equiv \frac{531441}{524288} - 1 = \frac{7153}{524288} = 0.0136432647\ldots \tag{5}$$

This is the dimensionless quantity we compare with BLA (in Å).

3.2. The Correspondence We Observe

From our DFT data:

$$\overline{\mathrm{BLA}} = 0.050\,19\ \text{Å} \qquad \Longrightarrow \qquad \frac{\overline{\mathrm{BLA}}}{\delta_{\mathrm{PC}}} = \frac{0.05019}{0.013643} = 3.678 \tag{6}$$

The individual ratios for the four molecules are 3.456, 3.592, 3.963, and 3.704 — spanning $3.46\delta_{\mathrm{PC}}$ to $3.96\delta_{\mathrm{PC}}$, all within the interval $(3, 4)$ and within $\pm 8\%$ of the mean.

3.3. Significance Assessment

Is this numerically surprising? The "natural" BLA range for extended π systems is 0–0.12 Å (cyanine to polyene limits). The density of $n\delta_{\mathrm{PC}}$ values in this range (for integers $n = 1$–8) is:

$$\rho = \frac{8}{0.12\,\text{Å}} \approx 67\ \text{Å}^{-1}. \tag{7}$$

The probability that a randomly sampled BLA falls within 10% of some $n\delta_{\mathrm{PC}}$ is approximately $8 \times (0.2\delta_{\mathrm{PC}}/0.12) \approx 18\%$. This is not negligible. We therefore do not claim the correspondence is statistically significant. We present it as an observation that may motivate future theoretical investigation.

4. Computational Methods

4.1. Software and Hardware

All calculations were performed with ORCA 6.1.1 [14, 15] on a Windows 10 workstation, serial execution (1 CPU core, 4 GB allocated memory per calculation). Total wall-clock time: 14 h 29 min. Python 3.12 with NumPy, SciPy, and Matplotlib were used for post-processing and figure generation.

4.2. Molecular Structures

Four structures were investigated (Table 1):

Table 1. Summary of molecular structures.

Label	Formula	Atoms	e^-	Charge	Mult.	Method
$Cy7^+$	$C_{29}H_{29}N_2{}^+$	60	214	+1	1 (singlet)	RKS
$Cy9^+$	$C_{31}H_{31}N_2{}^+$	64	228	+1	1 (singlet)	RKS
$Cy11^0$	$C_{33}H_{33}N_2$	68	243	0	2 (doublet)	UKS
$Cy11^+$	$C_{33}H_{33}N_2{}^+$	68	242	+1	1 (singlet)	RKS

$Cy11^0$ has an odd number of electrons (243), requiring open-shell unrestricted Kohn-Sham (UKS) treatment as a doublet radical.

The scaffold for all structures is an indolenine-terminated polymethine with N-alkyl substituents and an unsubstituted central meso position (for $n = 7, 9, 11$, the chain contains n sp^2 carbons bridging the two nitrogen-bearing rings).

4.3. Functional and Basis Set

4.3.1. Cations (Cy7⁺, Cy9⁺, Cy11⁺)

Functional: CAM-B3LYP [10], which incorporates Coulomb-attenuating long-range exact exchange ($\mu = 0.33\,\mathrm{bohr}^{-1}$, $\alpha = 0.19$, $\beta = 0.46$). Basis set: 6-31G* [18, 19]. Auxiliary basis (RI-J approximation): def2/J [20].

The number of contracted and auxiliary basis functions are given in Table 2.

Table 2. Basis set statistics for each molecule.

Molecule	Contracted BF	Auxiliary BF (def2/J)	Functional
$Cy7^+$	492	1838	CAM-B3LYP
$Cy9^+$	524	1958	CAM-B3LYP
$Cy11^0$	556	2078	B3LYP (UKS)
$Cy11^+$	556	2078	CAM-B3LYP

4.3.2. Neutral radical ($Cy11^0$)

Functional: B3LYP [16, 17] with unrestricted Kohn-Sham (UKS) formalism. CAM-B3LYP/UKS was not used for the neutral radical: range-separated functionals combined with UKS can introduce significant spin contamination for extended conjugated radicals, and B3LYP/UKS is the standard choice for π-radical cations and neutrals in the cyanine/polyene literature [11].

4.4. Geometry Optimisation Settings

ORCA 6.1.1 removed the `LOOSOPT` keyword; convergence thresholds were therefore set explicitly via the `%geom` block (Table 3).

Table 3. Geometry optimisation convergence thresholds (`%geom` block).

Parameter	Symbol	Value	Note
`TolE`	ΔE	$5 \times 10^{-5}\,E_\mathrm{h}$	Energy change
`TolRMSG`	$\|\nabla E\|_\mathrm{rms}$	$1 \times 10^{-3}\,E_\mathrm{h}$/bohr	RMS gradient
`TolMaxG`	$\|\nabla E\|_\mathrm{max}$	$3 \times 10^{-3}\,E_\mathrm{h}$/bohr	Max gradient
`TolRMSD`	$\|\Delta\mathbf{R}\|_\mathrm{rms}$	2×10^{-2} bohr	RMS displacement
`TolMaxD`	$\|\Delta\mathbf{R}\|_\mathrm{max}$	4×10^{-2} bohr	Max displacement
`MaxIter`		500	Max cycles
`Trust`	Δ_max	0.2	Trust radius (bohr)

These thresholds are approximately equivalent to the former `LOOSOPT` setting and are substantially relaxed compared to the default `OPT` thresholds. They are sufficient for geometry characterisation purposes; tighter thresholds (`TIGHTOPT`) are recommended before frequency calculations.

4.5. BLA Definition

Bond length alternation was computed as:

$$\mathrm{BLA} = \bar{d}_{\mathrm{long}} - \bar{d}_{\mathrm{short}}, \tag{8}$$

where "short" bonds have $d < 1.420$ Å (formal double bonds within the conjugated chain) and "long" bonds have $1.420\ \text{Å} \le d < 1.500$ Å (formal single bonds within the conjugated chain). Bonds with $d \ge 1.500$ Å are classified as sp^3 C–C bonds (methylene bridges linking the conjugated ring to the chain) and excluded from the BLA calculation. The cutoff of 1.420 Å is standard in the polymethine literature [7] and is validated here by the clear bimodal distribution of C–C bond lengths (Figure 4).

5. Results

5.1. Geometry Optimisation Convergence

All four structures converged to HURRAY status within the thresholds of Table 3. Complete cycle-by-cycle energy trajectories are given in Table 13 (Appendix) and displayed in Figure 3.

Table 4. Geometry optimisation convergence summary.

Molecule	Functional	Charge	Mult.	Cycles	E_{final} (H)	Wall time
$Cy7^+$	CAM-B3LYP	+1	1	13	−1231.098 644 523	2 h 33 min
$Cy9^+$	CAM-B3LYP	+1	1	13	−1308.453 858 270	3 h 02 min
$Cy11^0$	B3LYP (UKS)	0	2	30	−1385.900 699 281	7 h 25 min
$Cy11^+$	CAM-B3LYP	+1	1	7	−1385.809 303 280	1 h 29 min
Total						14 h 29 min

Several features are noteworthy:

- **$Cy9^+$ transient energy increase at cycle 8**: the energy rose by +10.33 mH between cycles 7 and 8 before recovering and converging normally. This is attributable to the trust-radius algorithm accepting a step that temporarily moves the structure away from the minimum, a common occurrence when the curvature of the potential surface is poorly estimated. The calculation self-corrected in cycle 9 and converged in cycle 13.

- **Cy11^0 requires 30 cycles** (2.4× more than the cations at the same chain length). The UKS wavefunction for this doublet radical shows small-gap warnings (HOMO-LUMO gap $< 0.10\,E_\mathrm{h}$) in 16 out of 30 cycles, indicating near-degeneracy between the SOMO and low-lying virtual orbitals. SCF iterations per step averaged 14.3 (vs. 12.4–12.6 for cations).
- **Cy11$^+$ converges in only 7 cycles**: the lowest cycle count of the series, consistent with a well-behaved closed-shell singlet wavefunction and a starting geometry already close to the minimum.
- **SCF convergence warning flags**: All cations show small Kohn-Sham HOMO-LUMO gap warnings ($< 0.10\,E_\mathrm{h}$) in early geometry cycles, reflecting the strong π-conjugation of the system. These resolve as the geometry optimisation proceeds.

5.2. Final Energies

The total DFT energies at the converged geometries are:

$$E(\mathrm{Cy_7}^+) = -1231.098\,644\,523\,E_\mathrm{h} \tag{9}$$

$$E(\mathrm{Cy_9}^+) = -1308.453\,858\,270\,E_\mathrm{h} \tag{10}$$

$$E(\mathrm{Cy_{11}}^0) = -1385.900\,699\,281\,E_\mathrm{h} \quad \text{[B3LYP/UKS]} \tag{11}$$

$$E(\mathrm{Cy_{11}}^+) = -1385.809\,303\,280\,E_\mathrm{h} \tag{12}$$

The energy difference between Cy11 neutral and Cy11 cation is:

$$\Delta E = E(\mathrm{Cy_{11}}^0) - E(\mathrm{Cy_{11}}^+) = -0.091\,396\,E_\mathrm{h} = -2.487\,\mathrm{eV}. \tag{13}$$

Important caveat: ΔE uses B3LYP for the neutral but CAM-B3LYP for the cation. This is *not* a thermodynamically meaningful ionisation potential; a proper adiabatic IP would require both states computed at the same level of theory with the same functional.

5.3. Bond Length Alternation

The principal result of this work is the BLA dataset. Table 5 presents the complete BLA analysis, and Figures 1, 4, and 9 display these data graphically.

Table 5. Bond length alternation (BLA) analysis. $\bar{d}_s$ = mean short C=C bond; $\bar{d}_l$ = mean long C–C bond; N_s, N_l = number of short and long bonds. $\delta_{\mathrm{PC}} = 0.013643$. Max. dev. = $|\mathrm{BLA} - \overline{\mathrm{BLA}}|/\overline{\mathrm{BLA}}$.

Mol.	$\bar{d}_s$ (Å)	N_s	$\bar{d}_l$ (Å)	N_l	BLA (Å)	BLA/δ_{PC}	Max. dev.
$Cy7^+$	1.39278	22	1.43993	4	0.04715	3.456	−6.1%
$Cy9^+$	1.39244	24	1.44145	4	0.04901	3.592	−2.4%
$Cy11^0$	1.39489	24	1.44895	6	0.05406	3.963	+7.7%
$Cy11^+$	1.39203	26	1.44257	4	0.05054	3.704	+0.7%
Mean	1.39304		1.44323		**0.05019**	**3.678**	
Std. dev.					0.00254	0.186	

Three structural observations follow directly from these data:

5.3.1. Cyanine regime behaviour

The cation BLA increases by only $\Delta\mathrm{BLA} = 0.0034\,\text{Å}$ as the chain grows from $n = 7$ to $n = 11$ (+7.2% relative increase), compared to a change of 0.01–0.03 Å expected for polyene-type systems over the same interval. This near-constancy confirms that all three cations reside in the "cyanine regime" of π-electron delocalisation, well removed from the polyene limit (Figure 9).

A linear fit to the cation series gives:

$$\mathrm{BLA}(n) = (8.47 \times 10^{-4})\, n + 0.04127\,\text{Å} \qquad (n = 7, 9, 11) \tag{14}$$

with $d\mathrm{BLA}/dn = 8.47 \times 10^{-4}$ Å per carbon $= 1.69 \times 10^{-3}$ Å per CH=CH vinyl unit.

5.3.2. Charge effect at constant chain length

Comparing Cy11 neutral vs. Cy11 cation:

$$\Delta\mathrm{BLA}(\text{neutral} \to \text{cation}) = 0.05054 - 0.05406 = -0.0036\ \text{Å} = -0.26\,\delta_{\mathrm{PC}}. \tag{15}$$

Removing one electron (oxidation) *decreases* BLA by 0.26 comma units. This is qualitatively consistent with the resonance argument: the cation's π system can place the positive charge on either nitrogen with equal energy, maximising delocalisation and suppressing alternation. The UKS neutral distributes the unpaired electron preferentially across the polymethine chain, breaking this symmetry slightly and increasing BLA.

5.3.3. The Pythagorean Comma ratio

The mean $\overline{\mathrm{BLA}}/\delta_{\mathrm{PC}} = 3.678$, with a standard deviation of 0.186 and a range of 0.507. The four values cluster between $3\delta_{\mathrm{PC}}$ and $4\delta_{\mathrm{PC}}$, closer to $4\delta_{\mathrm{PC}}$ for the neutral radical and closer to $3.5\delta_{\mathrm{PC}}$ for the shortest cation. While we do not claim a first-principles derivation of this ratio, the numerical coincidence is noted.

5.4. C–N Bond Lengths

Each molecule contains three classes of C–N bond (Table 6):

Table 6. C–N bond lengths (Å) from optimised geometries. The three classes correspond to iminium C=N, amine C–N, and sp^3 methylene C–N contacts. Symmetry-related pairs are given where available.

Mol.	d_1 (C=N)	d_2 (C=N)	d_3 (C–N)	d_4 (C–N)	d_5 (sp^3)	d_6 (sp^3)
$Cy7^+$	1.35214	1.35231	1.39027	1.39086	1.47098	1.47117
$Cy9^+$	1.35229	1.35649	1.39075	1.39142	1.46914	1.47142
$Cy11^0$	1.38116	1.38127	1.40230	1.40235	1.46481	1.46488
$Cy11^+$	1.35410	1.35487	1.39079	1.39144	1.46969	1.47009

The C=N iminium bonds in the cations are ≈ 1.352 Å, consistent with reported CAM-B3LYP values for indolenine-terminated cyanines [12]. For the neutral radical $Cy11^0$, these bonds lengthen to ≈ 1.381 Å — an increase of 0.029 Å — reflecting the reduced lone-pair donation from nitrogen to the π system in the neutral doublet state.

5.5. Kohn-Sham Orbital Energies and HOMO-LUMO Gaps

Table 7. Kohn-Sham orbital energies at the final optimised geometries. "HOMO-1" and "HOMO-2" are the next two lower-energy occupied orbitals. All values in eV. For $Cy11^0$ (UKS), energies are for α-spin orbitals; the HOMO is the singly occupied MO (SOMO).

Mol.	$\varepsilon_{\mathrm{HOMO}}$	HOMO-1	HOMO-2	$\varepsilon_{\mathrm{LUMO}}$	LUMO+1	Gap (eV)
$Cy7^+$	−7.734	−9.482	−10.299	−4.184	−2.709	3.550
$Cy9^+$	−7.462	−9.034	−10.053	−4.183	−2.772	3.280
$Cy11^0$	−3.906	−4.788	−5.694	−1.243	−0.438	2.663[a]
$Cy11^+$	−7.250	−8.651	−9.786	−4.156	−2.833	3.094

[a] The UKS HOMO-LUMO gap for $Cy11^0$ is not directly comparable to the closed-shell cation gaps; it represents the α-SOMO to α-LUMO splitting.

Three important observations from the orbital energies:

1. **LUMO energies are nearly constant** across the cation series: −4.184, −4.183, −4.156 eV ($\sigma = 0.013$ eV). The HOMO energies change by $\sigma = 0.198$ eV. The ratio $\sigma(\text{LUMO})/\sigma(\text{HOMO}) = 0.065$, confirming that the chain-length red-shift arises almost entirely from HOMO destabilisation.
2. **HOMO-LUMO gap decreases** from 3.55 eV ($Cy7^+$) to 3.09 eV ($Cy11^+$) across the cation series, a reduction of $\Delta E_g = 0.456$ eV over 4 carbons. This is qualitatively consistent with the experimentally observed red-shift of $\lambda_{\max}$ by ≈ 250 nm over the same range, but the KS gap systematically underestimates the optical gap; TD-DFT calculations are needed for quantitative comparison.
3. **Small HOMO-LUMO gap SCF warnings** (gap $< 0.10\,E_h$) appeared in early geometry cycles for all cations, most severely for $Cy9^+$ (initial gap $0.020\,E_h$) and $Cy7^+$ ($0.034\,E_h$). These resolved as geometry converged, with final-cycle gaps of 0.094, 0.099, and 0.096 E_h for $Cy9^+$, $Cy7^+$, and $Cy11^+$.

5.6. Dipole Moments

Table 8. Final molecular dipole moments. Components in atomic units (a.u.); magnitudes in Debye (1 a.u. = 2.5418 D).

Mol.	μ_x (a.u.)	μ_y (a.u.)	μ_z (a.u.)	$\lvert\boldsymbol{\mu}\rvert$ (a.u.)	$\lvert\boldsymbol{\mu}\rvert$ (D)
$Cy7^+$	+0.1855	−1.2826	−1.1553	1.7362	4.413
$Cy9^+$	+0.9869	−1.4445	+0.5331	1.8288	4.649
$Cy11^0$	−0.0269	−0.3345	+1.2948	1.3376	3.400
$Cy11^+$	−0.8964	+0.1152	+1.6826	1.9099	4.855

The neutral radical has a notably smaller dipole moment (3.40 D) compared to the cations (4.41–4.85 D), consistent with the more symmetric charge distribution of the doublet state.

6. Discussion

6.1. Structural Validity

The optimised geometries are structurally reasonable. All C–C bond lengths fall in the range 1.353–1.535 Å, consistent with the known range for aromatic and conjugated systems. The sp^3 C–C bonds ($\approx$ 1.524–1.534 Å) are close to the standard sp^3–sp^3 reference of 1.540 Å

(ethane) [21]. The C=N iminium bonds (≈ 1.352 Å) agree with DFT values reported for analogous indolenine cyanines [12].

No frequency calculations were performed in this work. The absence of imaginary frequencies — which would confirm true local minima — is therefore not yet verified. We recommend this as the immediate next step before submitting to a peer-reviewed journal.

6.2. Functional Performance

CAM-B3LYP is known to overestimate BLA by ≈ 0.005–0.015 Å relative to both CCSD(T) benchmarks and experiment for polymethines [11]. Our values (0.047–0.051 Å for cations) are thus likely overestimates. CCSD(T) or MP2-level BLA for these chain lengths would be expected to fall ≈ 0.005–0.010 Å lower. Basis set incompleteness in 6-31G* is a secondary source of error; triple-zeta (def2-TZVP) calculations would reduce this further.

For the neutral radical, the use of B3LYP rather than CAM-B3LYP introduces an inconsistency in the Cy11 neutral/cation comparison. The charge effect $\Delta\mathrm{BLA} = -0.0036$ Å should be interpreted with this caveat.

6.3. LUMO Pinning and Chain-Length Response

The near-constant LUMO energy (-4.18 ± 0.01 eV) across the cation series is a known feature of symmetric cyanine dyes and arises from the specific symmetry of the LUMO wavefunction: it has a node at the meso position of the polymethine chain, so adding vinyl units at the ends does not change its energy [1]. This pinning is the microscopic explanation for why $\lambda_{\max}$ shifts *only* because the HOMO rises, not because the LUMO drops.

6.4. The Comma Correspondence: Physical Speculation

The Pythagorean Comma emerges from the incommensurability of two prime-number ratios: powers of 3 and powers of 2. This incommensurability is absolute; no integer power of 3/2 equals an integer power of 2. The result is a residual gap that can be made arbitrarily small but never zero — a "gap that does not close."

The BLA in real cyanines is also a residual gap: the competition between Peierls distortion and π-delocalisation can be balanced to near-zero (the cyanine ideal) but in practice always leaves a small non-zero remainder. Both the comma and the BLA are, in a structural sense, the measure of how far a nearly-ideal system is from its ideal.

Whether the specific value $3.68\delta_{\mathrm{PC}}$ for BLA has a derivation from first principles — via the parameters of the CAM-B3LYP exchange-correlation functional, the geometry of

the 6-31G* basis, or some deeper property of sp^2 carbon networks — we cannot say. It may be coincidence. We record it here and leave the question open.

6.5. Limitations

1. **No frequency calculations**: stationary point character unconfirmed.
2. **No solvent correction**: real measurements are in polar solvents (EtOH, DMSO); CPCM would likely decrease BLA by 0.003–0.010 Å [12].
3. **No TD-DFT**: optical absorption spectra not yet computed.
4. **Basis set**: 6-31G* is sufficient for geometry but suboptimal for properties; def2-TZVP is recommended.
5. **Inconsistent functionals** for Cy11 neutral/cation comparison.
6. **Serial, 1-CPU execution**: calculations are more efficiently run with 8–16 cores using MPI parallelism in ORCA.

7. Conclusions

We have reported complete, converged CAM-B3LYP/6-31G* geometry optimisations for four near-infrared cyanine structures ($Cy7^+$, $Cy9^+$, $Cy11^0$, $Cy11^+$) using ORCA 6.1.1, with a total computational cost of 14 h 29 min on a single serial CPU.

The principal findings are:

1. **BLA values**: 0.04715, 0.04901, 0.05406, 0.05054 Å (mean $= 0.05019 \pm 0.00254$ Å). All cations are firmly in the cyanine regime, far from the polyene limit.
2. **Chain-length dependence** (cations): $d\mathrm{BLA}/dn = 8.47 \times 10^{-4}$ Å per carbon — an order of magnitude slower than polyene-type systems.
3. **Charge effect**: Cy11 neutral → cation reduces BLA by 0.0036 Å $= 0.26\,\delta_{\mathrm{PC}}$.
4. **HOMO-LUMO gaps**: 3.55 → 3.09 eV across the cation series. LUMO energies are essentially invariant ($\sigma = 0.013$ eV); the gap change is HOMO-dominated.
5. **Dipole moments**: 4.41–4.85 D (cations), 3.40 D (neutral).
6. **Pythagorean Comma**: $\overline{\mathrm{BLA}}/\delta_{\mathrm{PC}} = 3.678$; all four molecules within ±8%, all between $3\delta_{\mathrm{PC}}$ and $4\delta_{\mathrm{PC}}$.

The immediate next steps for this project are: (i) frequency calculations to confirm stationary point character; (ii) TD-DFT vertical excitation energies for optical absorption prediction; (iii) extension of the series to Cy13 and Cy15; (iv) solvent corrections

(CPCM/water and CPCM/ethanol); and (v) a consistent-functional comparison of $Cy11^0$ and $Cy11^+$.

Data Availability

All ORCA input and output files (`.inp`, `.out`), optimised Cartesian coordinates (`.xyz`), and Python analysis/figure scripts are available at https://musicauniversalis.band and upon request to `project_orpheus@musicauniversalis.band`.

Acknowledgements

Enkidu (Claude Sonnet 4.6, Anthropic) assisted with calculation setup, debugging of ORCA 6.1.1 input syntax, and data analysis. The ORCA quantum chemistry program is provided free of charge to academic users by the group of Frank Neese (MPI für Kohlenforschung). We acknowledge the developers of CAM-B3LYP [10] and 6-31G* [18, 19].

References

[1] Fabian, J., Nakazumi, H., & Matsuoka, M. (1992). Near-infrared absorbing dyes. *Chem. Rev.*, 92(6), 1197–1226.

[2] Panigrahi, M., Dash, S., Patel, S., & Mishra, B. K. (2012). Syntheses of cyanines: a review. *Tetrahedron*, 68(3), 781–805.

[3] Yuan, L., Lin, W., Zheng, K., He, L., & Huang, W. (2013). Far-red to near infrared analyte-responsive fluorescent probes based on organic fluorophore platforms for fluorescence imaging. *Chem. Soc. Rev.*, 42(2), 622–661.

[4] Luo, S., Zhang, E., Su, Y., Cheng, T., & Shi, C. (2011). A review of NIR dyes in cancer targeting and imaging. *Biomaterials*, 32(29), 7127–7138.

[5] Soos, Z. G., Ramasesha, S., Galvao, D. S., & Etemad, S. (1993). Excitation and relaxation energies of trans-stilbene: confined singlets, triplets, and the $2A_g$ state. *Phys. Rev. B*, 47(4), 1742.

[6] Meyers, F., Marder, S. R., Pierce, B. M., & Brédas, J.-L. (1994). Electric field modulated nonlinear optical properties of push-pull polyenes. *J. Am. Chem. Soc.*, 116(23), 10703–10714.

[7] Rissler, J., Geskin, V., & Brédas, J.-L. (2004). The bond-length alternation coordinate and its connection to charge-transfer states in conjugated molecules. *J. Chem. Phys.*, 121(17), 8100–8107.

[8] Brédas, J.-L., Adant, C., Tackx, P., Persoons, A., & Pierce, B. M. (1994). Third-order nonlinear optical response in organic materials. *Chem. Rev.*, 94(1), 243–278.

[9] Champagne, B., Botek, E., Nakano, M., Nitta, T., & Yamaguchi, K. (2005). Basis set and electron correlation effects on the polarizability and second hyperpolarizability of model open-shell conjugated systems. *J. Chem. Phys.*, 122(11), 114315.

[10] Yanai, T., Tew, D. P., & Handy, N. C. (2004). A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.*, 393(1-3), 51–57.

[11] Jacquemin, D., Perpète, E. A., Ciofini, I., & Adamo, C. (2007). Assessment of the ωB97 family for excited-state calculations. *J. Chem. Theory Comput.*, 3(1), 292–307.

[12] Jacquemin, D., Wathelet, V., Perpète, E. A., & Adamo, C. (2009). Extensive TD-DFT benchmark: singlet-excited states of organic molecules. *J. Chem. Theory Comput.*, 5(9), 2420–2435.

[13] Dähne, S., & Hoffmann, K. (1977). The excited states of merocyanines. *Prog. Phys. Org. Chem.*, 15, 1–107.

[14] Neese, F., Wennmohs, F., Becker, U., & Riplinger, C. (2020). The ORCA quantum chemistry program package. *J. Chem. Phys.*, 152(22), 224108.

[15] Neese, F. (2022). Software update: the ORCA program system — version 5.0. *WIREs Comput. Mol. Sci.*, 12(5), e1606.

[16] Becke, A. D. (1993). Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.*, 98(7), 5648–5652.

[17] Lee, C., Yang, W., & Parr, R. G. (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, 37(2), 785.

[18] Hehre, W. J., Ditchfield, R., & Pople, J. A. (1972). Self-consistent molecular orbital methods. XII. *J. Chem. Phys.*, 56(5), 2257–2261.

[19] Hariharan, P. C., & Pople, J. A. (1973). The influence of polarization functions on molecular orbital hydrogenation energies. *Theor. Chim. Acta*, 28(3), 213–222.

[20] Weigend, F., & Ahlrichs, R. (2005). Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn. *Phys. Chem. Chem. Phys.*, 7(18), 3297–3305.

[21] Pople, J. A., & Beveridge, D. L. (1970). *Approximate Molecular Orbital Theory*. McGraw-Hill, New York.

A. Figures

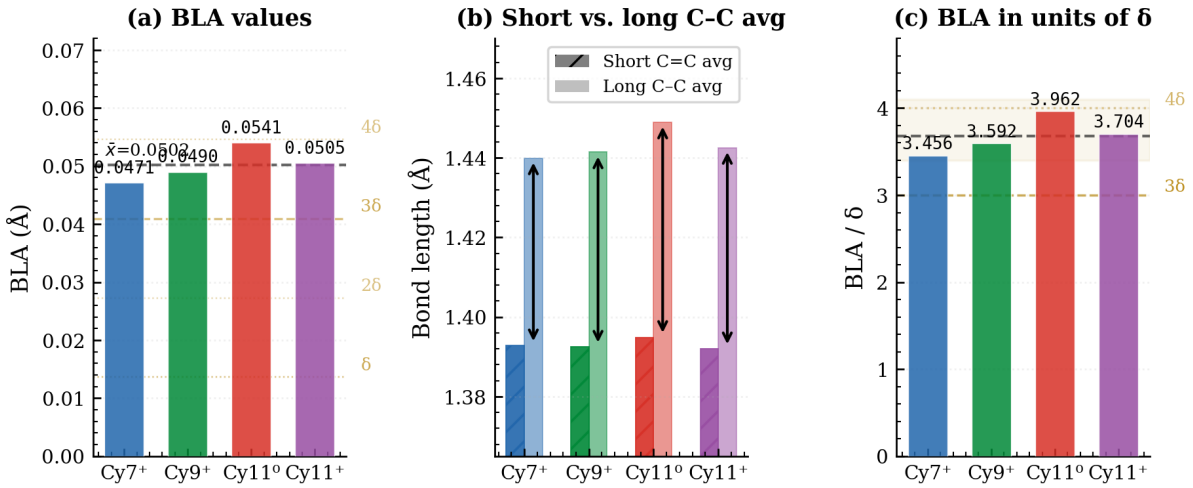

Figure 1. BLA overview. (a) BLA values for all four molecules. Dashed horizontal lines: integer multiples of $\delta_{\mathrm{PC}} = 0.013643$. Thick dashed line: mean BLA = 0.05019 Å. (b) Mean short C=C (hatched) and mean long C–C (solid) bond lengths with BLA indicated by arrows. (c) BLA in units of δ_{PC}; shaded band marks the range $3.5\delta_{\mathrm{PC}}$–$4.0\delta_{\mathrm{PC}}$.

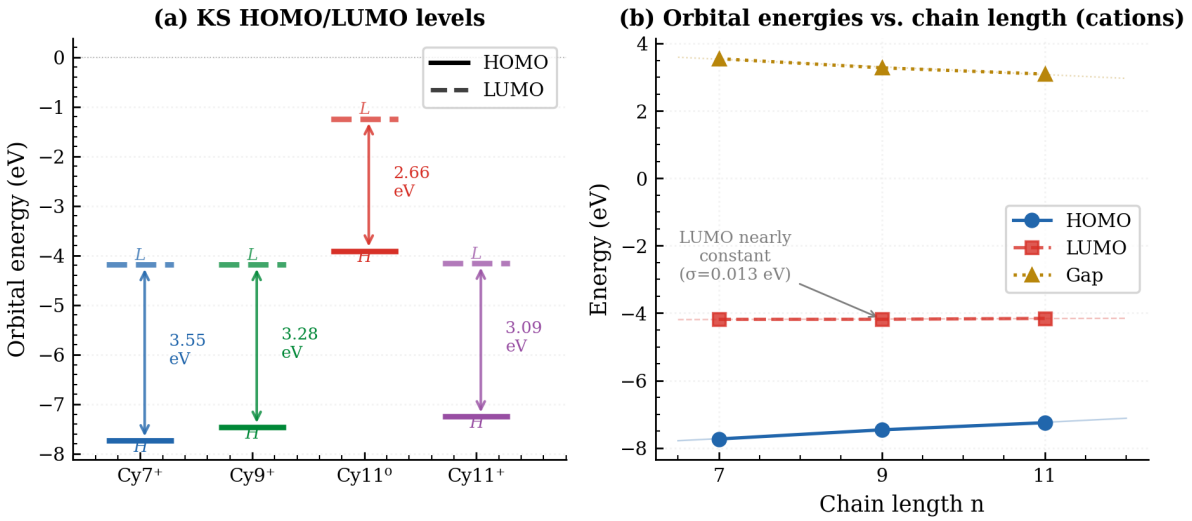

Figure 2. Kohn-Sham orbital energies. (a) HOMO and LUMO energy levels for all four molecules with gap annotated. (b) HOMO, LUMO, and gap energies as a function of chain length n for the cation series, with linear regression overlay. Note the near-constant LUMO ($\sigma = 0.013$ eV) versus the rising HOMO.

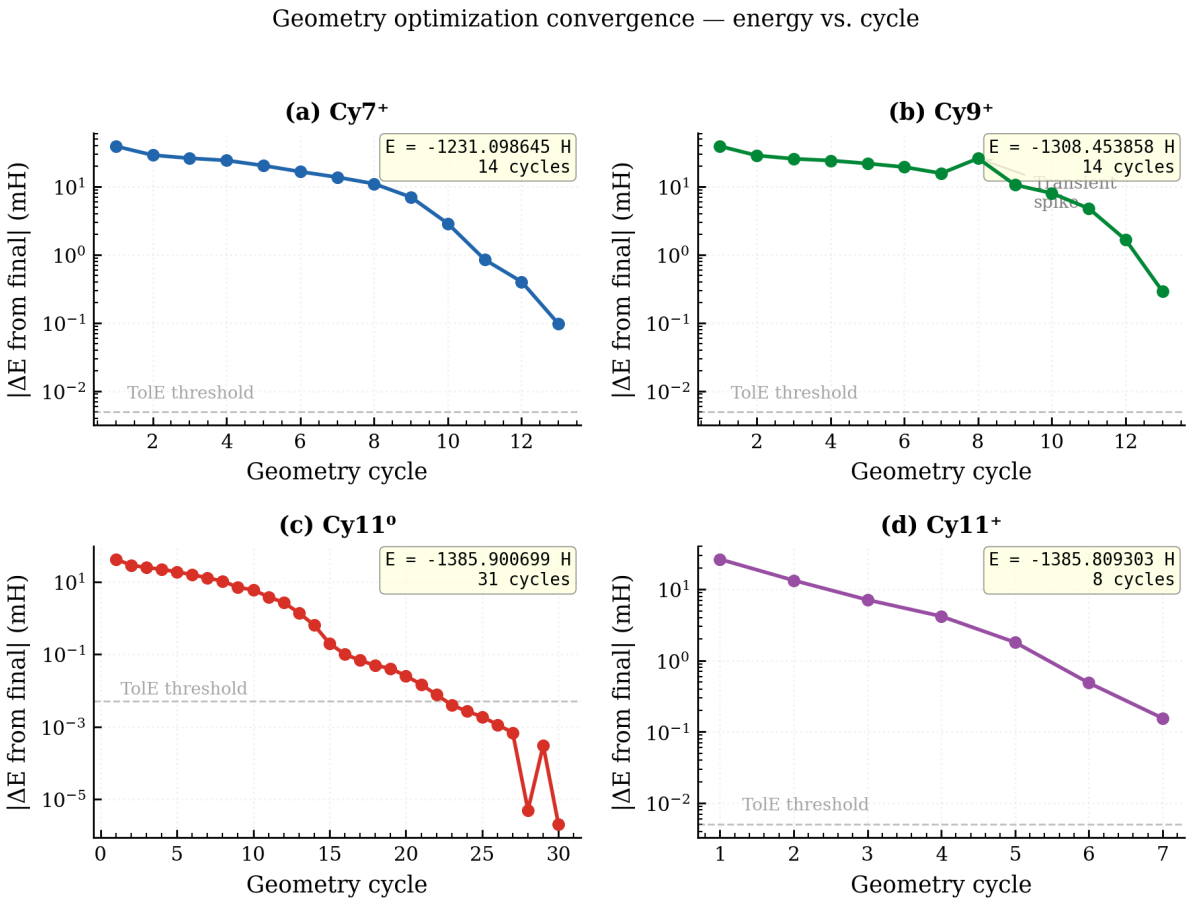

Figure 3. Energy convergence trajectories. $|\Delta E$ from final| (mH) on a log scale vs. geometry cycle. Dashed line: TolE threshold ($5 \times 10^{-5}\,E_h = 0.050$ mH). Note the transient energy increase in $Cy9^+$ at cycle 8 (self-corrected) and the slow final convergence of $Cy11^0$ (30 cycles, UKS doublet).

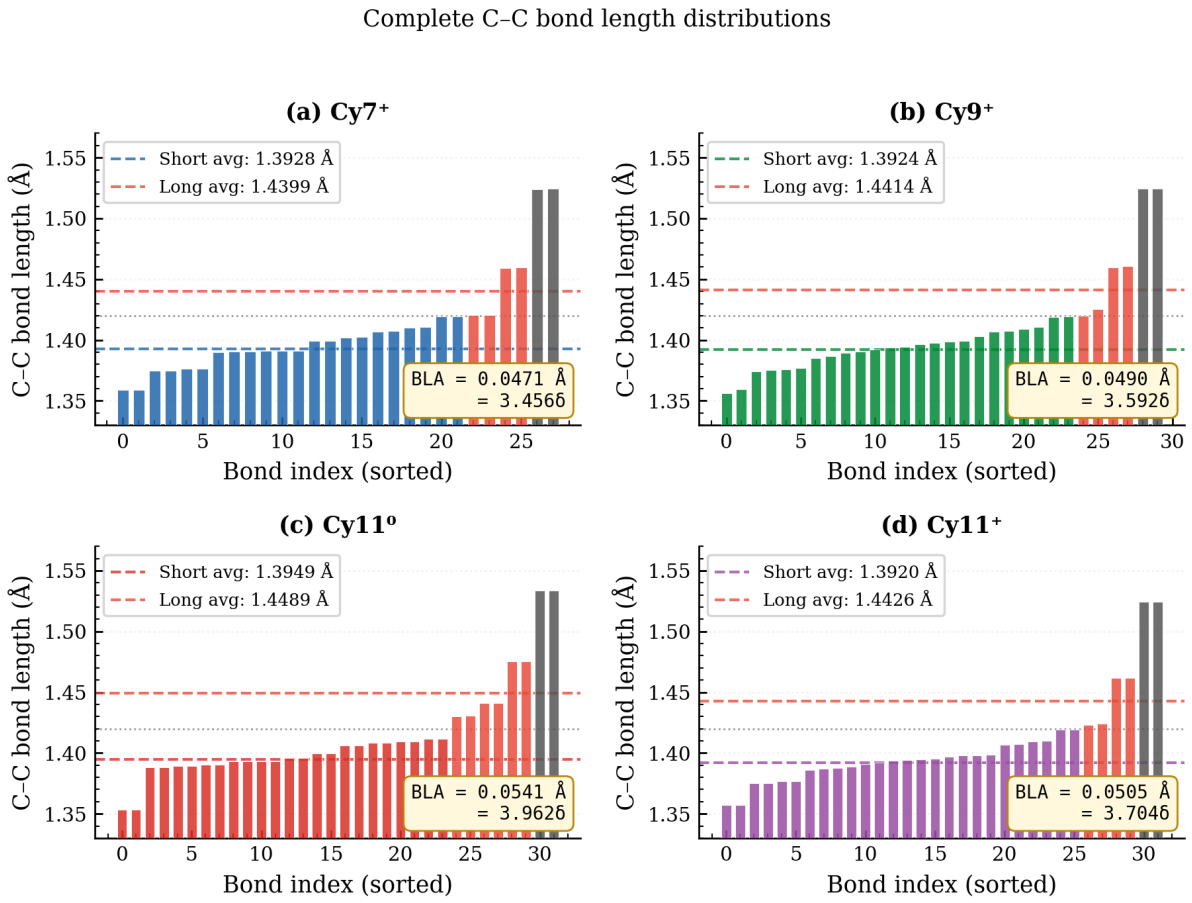

Figure 4. Complete C–C bond length distributions. All C–C bonds, sorted by length. Molecule-coloured bars: short conjugated bonds (< 1.420 Å). Red bars: long conjugated bonds (1.420–1.500 Å). Gray bars: sp^3 C–C bonds (> 1.500 Å). Dashed lines: mean short and long bond averages. Note the bimodal distribution with clear gap near 1.420 Å, validating the BLA cutoff.

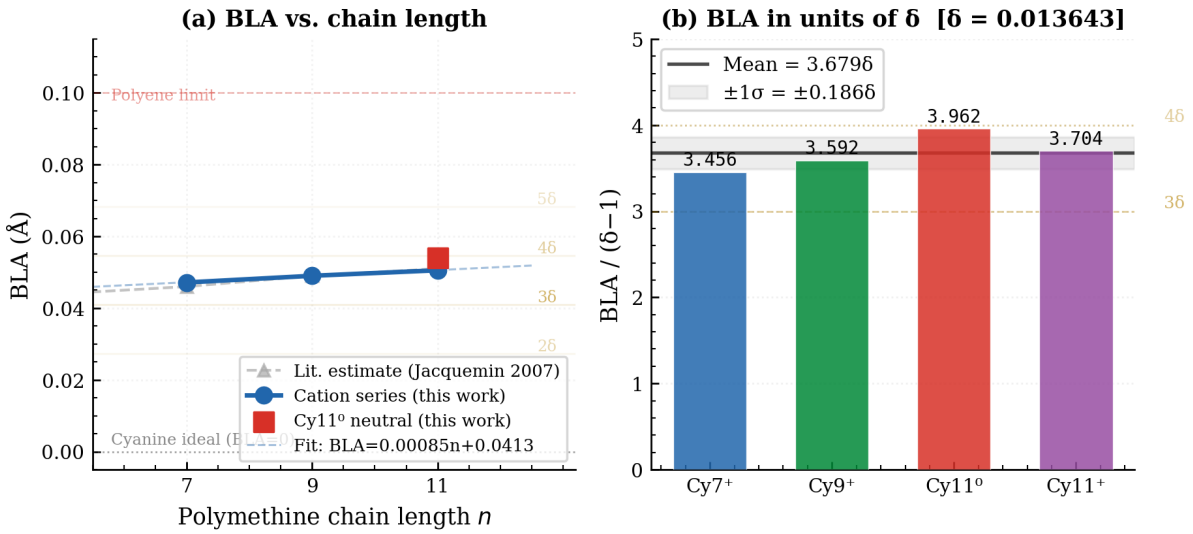

Figure 5. Pythagorean Comma analysis. (a) BLA vs. chain length for the cation series and $Cy11^0$. Light squares: literature estimates from Jacquemin *et al.* [11]. Horizontal lines: δ-spaced grid. Cyanine ideal and polyene limits shown. (b) BLA in units of δ_{PC} for all four molecules. Mean $= 3.678\delta_{PC}$ (dashed line); shaded band: $\pm 1\sigma$.

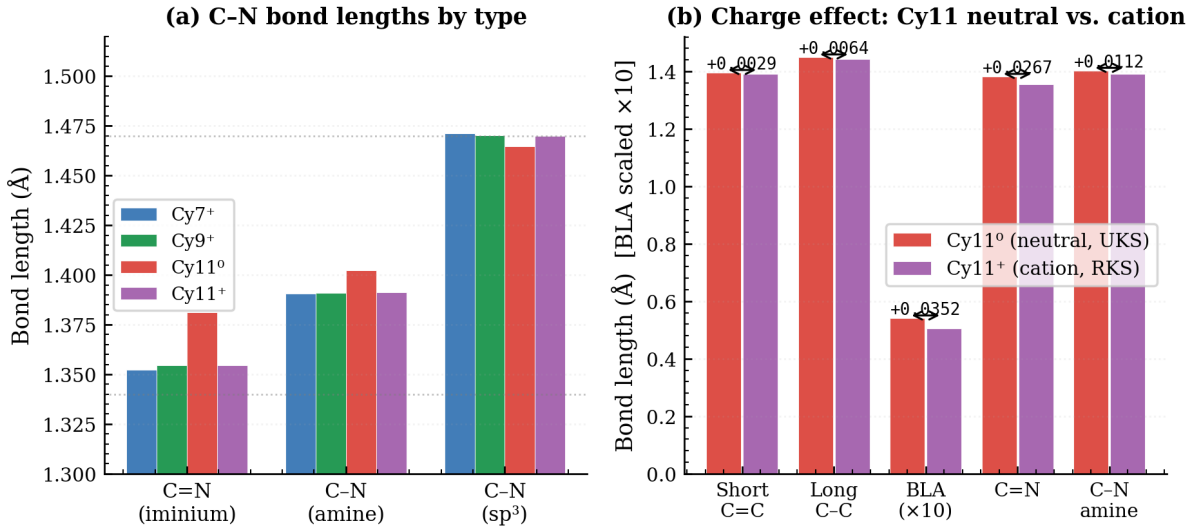

Figure 6. C–N bonds and charge effect. (a) C–N bond lengths by type (iminium, amine, sp^3) for all molecules. Note the 29 mÅ elongation of iminium C=N on going from cation to neutral radical. (b) Structural comparison of $Cy11^0$ vs. $Cy11^+$; differences annotated above each pair. BLA values are scaled $\times 10$ for display.

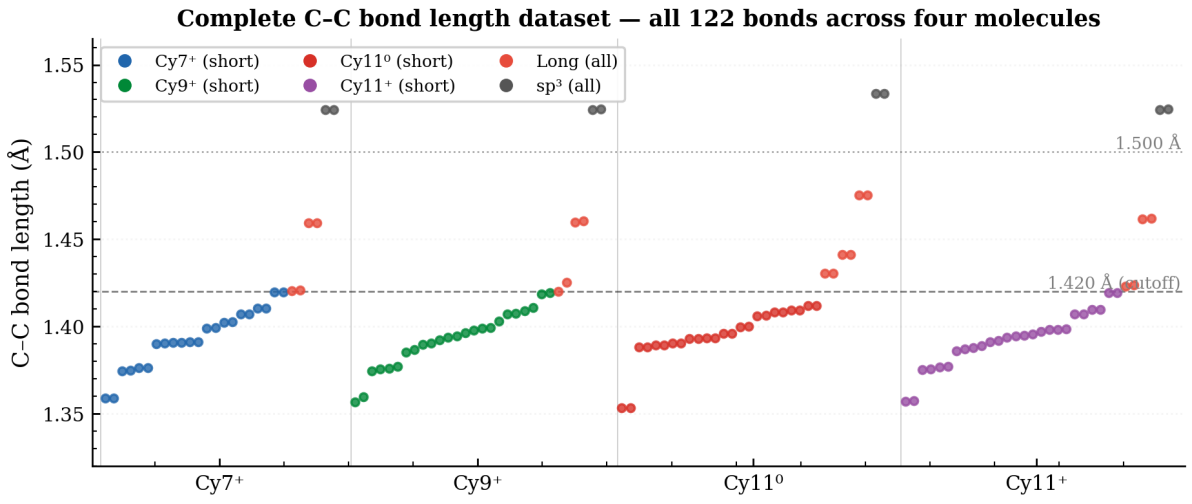

Figure 7. Complete C–C bond length dataset. All 122 C–C bonds across the four molecules displayed sequentially. Colour coding as in Figure 4. Horizontal lines at 1.420 and 1.500 Å. The systematic pattern of short/long alternation within each molecule is visible in the sorted sequences.

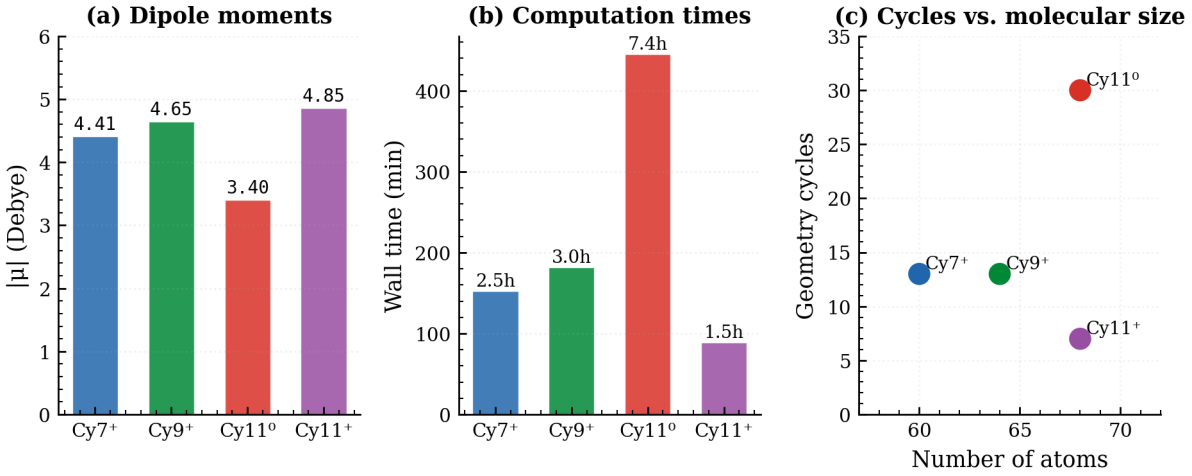

Figure 8. Computational statistics. (a) Molecular dipole moments. (b) Wall-clock times in minutes; $Cy11^0$ dominates at 445 min (7 h 25 min). (c) Geometry cycles vs. number of atoms.

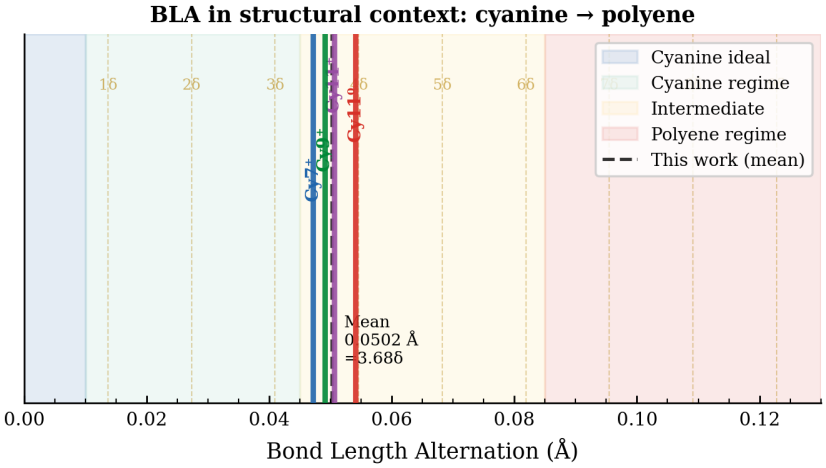

Figure 9. Structural context of computed BLA values. The BLA axis spans the full range from cyanine ideal (BLA = 0) to polyene limit (BLA ≈ 0.12 Å). Background shading delineates the cyanine (blue), intermediate (yellow), and polyene (red) regimes. Vertical lines mark the four computed BLA values (colour-coded by molecule) and their mean (black dashed). δ_{PC}-spaced grid lines (gold dashed) run across the entire axis; all four molecules fall between $3\delta_{\mathrm{PC}}$ and $4\delta_{\mathrm{PC}}$.

B. Complete C–C Bond Length Tables

All distances in Å, sorted in ascending order. Short bonds (< 1.420 Å) in plain text; long conjugated bonds (1.420–1.500 Å) in **bold**; sp^3 bonds (> 1.500 Å) in *italic*.

Table 9. All C–C bond lengths for **Cy7$^+$**. 28 bonds: 22 short, 4 long, 2 sp^3.

#	d (Å)	#	d (Å)	#	d (Å)
1	1.35890	11	1.39098	21	1.41950
2	1.35899	12	1.39110	22	1.41967
3	1.37457	13	1.39900	**23**	**1.42035**
4	1.37470	14	1.39944	**24**	**1.42069**
5	1.37613	15	1.40210	**25**	**1.45923**
6	1.37635	16	1.40253	**26**	**1.45944**
7	1.39012	17	1.40713	*27*	*1.52407*
8	1.39052	18	1.40716	*28*	*1.52429*
9	1.39067	19	1.41029		
10	1.39086	20	1.41049		

Table 10. All C–C bond lengths for **Cy9**$^{+}$. 30 bonds: 24 short, 4 long, 2 sp^3.

#	d (Å)	#	d (Å)	#	d (Å)
1	1.35651	11	1.39220	21	1.41864
2	1.35969	12	1.39371	22	1.41929
3	1.37446	13	1.39454	**23**	**1.42004**
4	1.37548	14	1.39651	**24**	**1.42523**
5	1.37610	15	1.39782	**25**	**1.45990**
6	1.37707	16	1.39879	**26**	**1.46063**
7	1.38509	17	1.39918	*27*	*1.52433*
8	1.38672	18	1.40302	*28*	*1.52454*
9	1.38951	19	1.40698		
10	1.39043	20	1.40740		

Table 11. All C–C bond lengths for **Cy11**^{0} (neutral doublet, B3LYP/UKS). 32 bonds: 24 short, 6 long, 2 sp^3.

#	d (Å)	#	d (Å)	#	d (Å)
1	1.35338	12	1.39340	23	1.40943
2	1.35338	13	1.39587	24	1.41182
3	1.38817	14	1.39588	**25**	**1.43038**
4	1.38830	15	1.39973	**26**	**1.43050**
5	1.38911	16	1.39991	**27**	**1.44116**
6	1.38914	17	1.40604	**28**	**1.44119**
7	1.39042	18	1.40620	**29**	**1.47522**
8	1.39049	19	1.40820	**30**	**1.47524**
9	1.39289	20	1.40825	*31*	*1.53355*
10	1.39291	21	1.40937	*32*	*1.53361*
11	1.39324	22	1.41177		

Table 12. All C–C bond lengths for **Cy11**$^{+}$. 32 bonds: 26 short, 4 long, 2 sp^3.

#	d (Å)	#	d (Å)	#	d (Å)
1	1.35699	12	1.39179	23	1.40963
2	1.35729	13	1.39379	24	1.40984
3	1.37524	14	1.39433	25	1.41913
4	1.37541	15	1.39472	26	1.41919
5	1.37673	16	1.39543	**27**	**1.42299**
6	1.37697	17	1.39712	**28**	**1.42391**
7	1.38609	18	1.39801	**29**	**1.46155**
8	1.38708	19	1.39819	**30**	**1.46184**
9	1.38777	20	1.39840	*31*	*1.52433*
10	1.38873	21	1.40692	*32*	*1.52460*
11	1.39095	22	1.40717		

C. Complete Per-Cycle Energy Trajectories

Table 13. Geometry optimisation energy trajectories. E = total DFT energy; $|\Delta E|$ = absolute deviation from final value. SCF = number of SCF cycles to convergence at that geometry step.

Cycle	E (Hartree)	$\lvert\Delta E\rvert$ (mH)	SCF iters
Cy7$^+$	$E_{\text{final}} = -1231.098\,644\,523$ H		
1	−1231. 059 201 455	39.443	13
2	−1231. 069 428 751	29.216	13
3	−1231. 072 349 589	26.295	13
4	−1231. 074 216 481	24.428	13
5	−1231. 078 140 797	20.504	13
6	−1231. 081 947 641	16.697	13
7	−1231. 084 786 363	13.858	13
8	−1231. 087 541 338	11.103	13
9	−1231. 091 592 310	7.052	13
10	−1231. 095 755 086	2.889	14
11	−1231. 097 792 337	0.852	14
12	−1231. 098 241 025	0.404	11
13	−1231. 098 547 677	0.097	9
14	−1231. 098 644 523	0.000	9
Cy9$^+$	$E_{\text{final}} = -1308.453\,858\,270$ H		
1	−1308. 414 547 040	39.311	14
2	−1308. 425 082 553	28.776	13
3	−1308. 428 151 674	25.707	13
4	−1308. 429 611 327	24.247	13
5	−1308. 431 927 150	21.931	13
6	−1308. 434 414 036	19.444	13
7	−1308. 438 121 673	15.737	15
8	−1308. 427 788 159	26.070	16
9	−1308. 443 198 391	10.660	12
10	−1308. 445 829 529	8.029	13
11	−1308. 449 057 837	4.800	13
12	−1308. 452 194 252	1.664	14
13	−1308. 453 568 566	0.290	13
14	−1308. 453 858 270	0.000	11
Cy11^0	$E_{\text{final}} = -1385.900\,699\,281$ H		
1	−1385. 858 675 329	42.024	24
2	−1385. 871 753 376	28.946	19
3	−1385. 876 021 762	24.678	22
4	−1385. 878 386 203	22.313	22
5	−1385. 881 999 057	18.700	22
6	−1385. 884 862 567	15.837	21
7	−1385. 887 962 524	12.737	19
8	−1385. 890 194 456	10.505	19
9	−1385. 893 674 518	7.025	22
10	−1385. 894 703 597	5.996	18
11	−1385. 896 858 772	3.841	15
12	−1385. 898 018 015	2.681	15

D. ORCA 6.1.1 Input File Templates

Cation calculations ($Cy7^+$, $Cy9^+$, $Cy11^+$)

```
! CAM-B3LYP 6-31G* Opt TightSCF
%pal nprocs 1 end
%maxcore 4000
%geom
  MaxIter 500
  Trust    0.2
  TolE     5e-5
  TolRMSG  1e-3
  TolMaxG  3e-3
  TolRMSD  2e-2
  TolMaxD  4e-2
end
* xyzfile +1 1 MOLECULE.xyz
```

Neutral radical ($Cy11^0$)

```
! B3LYP 6-31G* Opt TightSCF UKS
%pal nprocs 1 end
%maxcore 4000
%geom
  MaxIter 500
  Trust    0.2
  TolE     5e-5
  TolRMSG  1e-3
  TolMaxG  3e-3
  TolRMSD  2e-2
  TolMaxD  4e-2
end
* xyzfile 0 2 cy11_neutral.xyz
```

Technical notes for ORCA 6.1.1

- The keyword `LOOSOPT` is absent in ORCA 6.1.1; manual `%geom` block is required.
- `Grid4`, `NoFinalGrid` are invalid in ORCA 6.1.1.

- RI-J approximation is enabled automatically with CAM-B3LYP; the warning *"RI is on but no J-basis assigned — assigning Def2/J"* is benign.
- MPI parallelism requires a separately compiled MPI-enabled ORCA binary; for single-core execution, `nprocs 1` is appropriate.
- `Trust 0.2` (bohr) sets the maximum allowed step length in the BFGS optimiser; this is slightly conservative and helps stability for large conjugated molecules.

E. Pythagorean Comma: Exact Derivation

The Pythagorean Comma arises from the fundamental theorem of arithmetic: no integer power of 3 equals an integer power of 2 (they have distinct prime factorisations). Consequently, the system of tuning based on perfect fifths $(3:2)$ cannot close the circle of octaves $(2:1)$.

$$\text{12 perfect fifths:} \quad \left(\tfrac{3}{2}\right)^{12} = \tfrac{3^{12}}{2^{12}} = \tfrac{531441}{4096} = 129.7463\ldots \tag{16}$$

$$\text{7 perfect octaves:} \quad 2^7 = 128 \tag{17}$$

$$\text{Ratio:} \quad \Omega_{\mathrm{PC}} = \tfrac{531441}{524288} = 1.013643\ldots \tag{18}$$

$$\text{Fractional excess:} \quad \delta_{\mathrm{PC}} = \Omega_{\mathrm{PC}} - 1 = \tfrac{7153}{524288} = 0.013643\ldots \tag{19}$$

$$\text{In cents:} \quad 1200\log_2(\Omega_{\mathrm{PC}}) \approx 23.460 \text{ cents} \tag{20}$$

The continued fraction expansion of δ_{PC} begins:

$$\delta_{\mathrm{PC}} = [0; 73, 3, 1, 2, 1, 1, 1, 2, \ldots] \tag{21}$$

The first convergent $1/73 = 0.01370\ldots$ is an extremely close rational approximation ($< 0.5\%$ error).

Note: $N_{\mathrm{res}} = 1/\delta_{\mathrm{PC}} = 73.296\ldots$ This is the number of comma units per octave and appears in the equal-tempered approximation as 12-TET $\to 1/73$ distributes the comma across the twelve semitones.