

From mono- to hexa-interstitials: Computational insights into carbon defects in diamond

Nima Ghafari Cherati^{1,2}, Arsalan Hashemi^{3,4}, and Ádám Gali^{1,2,5,*}

¹*HUN-REN Wigner Research Centre for Physics, P.O. Box 49, H-1525 Budapest, Hungary*

²*Department of Atomic Physics, Institute of Physics, Budapest University of Technology and Economics, Műegyetem rakpart 3, H-1111 Budapest, Hungary*

³*European Laboratory for Learning and Intelligent Systems (ELLIS) Institute Finland, Maarintie 8, 02150 Espoo, Finland*

⁴*Department of Technical Physics, University of Eastern Finland, P.O. Box 1627, FI-70211 Kuopio, Finland*

⁵*MTA-WFK Lendület “Momentum” Semiconductor Nanostructures Research Group, P.O. Box 49, H-1525 Budapest, Hungary*



(Received 9 December 2025; accepted 23 March 2026; published 29 April 2026)

We present a comprehensive first-principles investigation of carbon self-interstitial defects in diamond, ranging from mono- to hexa-interstitial complexes. By density functional theory, empowered by interatomic potential models, we efficiently sample the complex configurational landscape and identify both known and previously unreported defect geometries. Our results reveal a pronounced energetic driving force for aggregation: The formation energy per interstitial decreases systematically from isolated split interstitials to compact multi-interstitial clusters, with the tetra-interstitial platelet emerging as a particularly stable structural motif. Additionally, charge analysis indicates that the predominantly covalent bonding in diamond becomes more polar within the defect centers. Analysis of defect energy levels shows that only the investigated mono-, di-, penta-, and hexa-interstitial complexes introduce in-gap electronic states, whereas the tri- and tetra-interstitial clusters are electronically inert. Vibrational spectroscopies further reveal that self-interstitials generate characteristic signatures. Short carbon-carbon bonds inside the defect cores give rise to high-frequency vibrational modes between 1375 and 1925 cm^{-1} , which are strongly infrared active but exhibit weak Raman activity. Through a systematic analysis of metastable configurations, we identify the 3H defect center as a neutral di-interstitial defect. Based on this identification, we further suggest that the TR12 center may arise from a 3H-containing defect like a metastable hexa-interstitial configuration. Taken together, these findings provide a coherent picture of the structural, electronic, and vibrational characteristics of carbon self-interstitials and establish a robust framework for their experimental identification.

DOI: [10.1103/kxgc-k5lq](https://doi.org/10.1103/kxgc-k5lq)

I. INTRODUCTION

One of the long-standing challenges in materials engineering is the controlled growth of truly pristine diamond. Even the highest-purity single-crystal diamonds produced by modern high-pressure–high-temperature or chemical vapor deposition methods contain both extrinsic and intrinsic defects, including nitrogen (N), hydrogen (H), boron (B), and carbon (C) vacancies, and self-interstitials, typically at ppm levels [1,2]. Remarkably, the standard classification scheme is itself defined by defects: Type IaA, IaB, and Ib diamonds incorporate different configurations of N impurities; type IIa is essentially free of both N and B impurities, whereas type IIb contains B as the dominant impurity [3–6].

However, defects do not always deteriorate the properties of diamond; on the contrary, they can endow it with

technologically valuable functionalities [7–9]. Among the hundreds of reported centers, the prototypical nitrogen-vacancy (NV) defect stands out for its optically detected magnetic resonance (ODMR) and millisecond-scale spin coherence at room temperature, which together underpin its powerful quantum sensing capabilities [10]. In contrast to the NV center, which has been extensively studied over the past decade and measures only the projection of the magnetic field along its single symmetry axis, diamond also hosts a multiorientation center, TR12, that couples to the field along 12 symmetry-related axes. This enables true vector magnetometry with wide angular coverage from a single ensemble spectrum, allowing for the reconstruction of the full magnetic field without sample rotation and enhancing robustness to arbitrary field orientations [11]. Despite its significance, the atomistic structures of TR12 and several putative C self-interstitial complexes, including the 3H and R-family centers [12,13], remain unresolved.

In the ongoing effort to identify the atomistic structure of TR12, numerous experimental investigations have converged on the view that TR12 is an intrinsic interstitial-related defect in type IIa diamond. Following the first report of TR12 in 1956 [14], Walker showed that the center can be generated in high-purity type IIa diamond by electron irradiation,

*Contact author: gali.adam@wigner.hun-ren.hu

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

without introducing foreign impurities, implying that only lattice atoms participate in its formation [15]. Its nonlinear dose dependence further points to a multiparticle intrinsic complex rather than a simple point defect. Optical studies subsequently established TR12 as a distinct vibronic system with a zero-phonon line (ZPL) near 2.63 eV, observed in nitrogen-free diamond after damage, consistent with an aggregated intrinsic damage center rather than an impurity-related defect [16]. Annealing experiments then provided a mechanistic insight: While several self-interstitial-related centers anneal out in the temperature window below vacancy mobility (~ 900 K), the TR12 (and 5RL) bands grow in the same range, indicating that TR12 is fed by mobile carbon self-interstitials released from precursor complexes [17].

Ion-implantation studies strengthened this picture and highlighted TR12's photophysical promise, reporting optically active centers with a Debye-Waller factor of order 0.1 at cryogenic temperatures, higher than that of the NV center, suggesting improved photon emission into the ZPL [18]. Spatially resolved hyperspectral mapping further showed that the TR12 distribution along an implantation track can be treated as a *split self-interstitial map*, with TR12 produced throughout much of the profile before and after high-temperature annealing and suppressed in the most heavily damaged end-of-range region, supporting an association with split-interstitial-derived configurations [19]. Single-center ODMR measurements then established TR12 as a spin-active defect with multiorientation response and a level structure involving singlet states coupled via a long-lived metastable triplet, enabling vector magnetometry without sample rotation and with enhanced robustness to arbitrary field directions [11]. Most recently, focused carbon-ion microbeam experiments directly correlated depth-resolved TR12 formation with the local interstitial reservoir and competition with the 3H center, supporting TR12 as a self-interstitial cluster defect that can be engineered for high-performance magnetometry [20].

Despite the wealth of compositional evidence, pinpointing the exact atomistic structures of defect centers in bulk diamond remains exceptionally challenging, because nearly all available experimental probes provide only indirect signatures rather than direct real-space structural information. Techniques such as optical spectroscopy, electron paramagnetic resonance (EPR), ODMR, and infrared or Raman scattering constrain only the symmetry, spin state, charge state, and local vibrational characteristics. As a result, many distinct microscopic configurations can produce nearly identical experimental fingerprints in these observables [21,22]. Moreover, bulk diamond is both extremely stiff and dense, rendering direct real-space imaging of buried point defects at atomic resolution effectively impossible. Compounding the difficulty, many relevant centers occur at low concentrations and can exhibit multiple charge states or dynamic Jahn-Teller distortions, which further broaden or obscure their spectral signatures [11,21].

In this context, density functional theory (DFT) provides a powerful complementary route: Candidate defect geometries can be constructed, relaxed at the atomistic level, and evaluated through their formation energies, charge-transition levels, and local relaxations. Importantly, DFT can also predict directly comparable observables, including ZPL energies,

zero-field splitting (ZFS) parameters, and local vibrational modes, enabling stringent, quantitative comparisons with experiment [23–25]. Comparison between computed and measured signatures can exclude broad classes of structural models and, in favorable cases, support a unique microscopic assignment. For more complex centers, such as TR12, ongoing DFT efforts are steadily narrowing the landscape of viable carbon self-interstitial complexes consistent with their spectroscopic fingerprints [11,24].

However, early DFT studies, predominantly based on traditional semilocal functionals [26,27], showed that the tetrahedral, hexagonal, and bond-centred mono-interstitial configurations are highly unstable, whereas the $\langle 001 \rangle$ split interstitial is the only robust elementary self-interstitial, with a large formation energy and a central role in self-diffusion [23,28]. In a complementary line of work, Goss and co-workers investigated di-, tri-, and tetra-interstitial clusters, proposing these small aggregates as building blocks for larger planar interstitial structures (platelets) and assigning several configurations to specific triplet EPR centres, including R2, R1, and O3 [24,29–31]. More recent hybrid-functional studies have refined the vibrational and EPR signatures of $\langle 100 \rangle$ interstitials, their N-containing analogs and vicinal double-interstitial complexes, and have also examined selected vacancy-interstitial pairs and their recombination pathways [32–34]. Yet, these calculations have been limited to the previously reported high-symmetry configurations, typically with no more than four interstitial atoms. This gap motivates the present work: a comprehensive, high-accuracy screening of interstitial complexes as microscopic models for TR12.

In this work, we perform a systematic first-principles investigation of carbon self-interstitial defects. Motivated by suggestions that TR12 may involve aggregates of up to six carbon self-interstitials [35], we extend our analysis to clusters containing one to six self-interstitial atoms that can still be treated accurately with first-principles methods in a computationally feasible model system. We first investigate the structural properties of these configurations and then compare their relative formation energies to identify the most stable defects. We also investigate the electronic and vibrational properties of the most stable complexes, providing Raman and infrared spectroscopic fingerprints that enable the identification of electronically and optically silent configurations. Furthermore, we assess candidate carbon-interstitial-related color centers in the visible range, including the TR12 center, by comparing computed photoluminescence spectra with experiment.

II. METHODOLOGY

DFT calculations were performed using VASP [36] within the projector-augmented-wave framework [37]. The wave functions were expanded in a plane-wave basis with a kinetic-energy cutoff of 370 eV. Total energies and electronic-structure properties were evaluated using the screened hybrid Heyd-Scuseria-Ernzerhof (HSE06) functional [38]. For HSE06, all structures were fully relaxed until the total energy difference between successive ionic steps was below 10^{-5} eV and the forces on each atom were less than

10^{-2} eV/Å. For vibration assessment, the computationally cheaper Perdew-Burke-Ernzerhof (PBE) functional [26] was used. To this end, the total energies in both geometry relaxation and phonon calculations were converged within 10^{-6} eV and the forces to within 10^{-3} eV/Å. We verified that no imaginary phonon modes were present, confirming that all structures were fully optimized and dynamically stable. The optimized lattice constant is $a = 3.57$ Å (PBE) and 3.54 Å (HSE06), compared with the experimental 3.57 Å, and HSE06 reproduces a 5.4 eV band gap within 0.1 eV of experiment [39].

Defect model systems were constructed in a $4 \times 4 \times 4$ cubic supercell containing 512 carbon atoms, derived from the conventional diamond unit cell. With this supercell size, Brillouin-zone sampling at the Γ point provides converged defect energetics and electronic properties. During structural relaxations, the lattice vectors were kept fixed to isolate local distortions induced by the defects and to avoid spurious cell-wide strain effects. Within this framework, the carbon number density increases from 3.522 to 3.556 g/cm³ when going from mono- to hexa-interstitial carbon complexes.

To easily generate more realistic initial configurations for carbon self-interstitial clusters in diamond, we employed a geometric Frenkel-pair construction carried out directly in configuration space. A spherical cavity of radius R was created at the center of a 512-atom diamond supercell by removing all atoms within a distance R from the cell center. The cavity radii (1.6–3.9 Å) were selected based on the first several coordination shells of diamond, enabling a systematic modeling of voids capable of accommodating mono- to multi-interstitial complexes. Within this cavity, m additional carbon atoms ($1 \leq m \leq 6$) were inserted using rejection Monte Carlo sampling with a hard-sphere exclusion distance of 1.2 Å, ensuring physically reasonable C–C separations. Candidate positions were drawn from an isotropic, volume-uniform distribution, allowing the initial aggregates to sample a broad range of compact and extended motifs. For each cluster size and R , 200 statistically independent initial structures were generated. By deliberately sampling high-energy, distorted, and nonequilibrium configurations, we ensure adequate exploration of the defect configurational space.

These configurations provide realistic starting points for subsequent geometry optimizations using a pretrained neuroevolutionpotential (NEP) model [40] and an empirical interatomic potential [41], as implemented in GPUMD [42]. The NEP model developed by Wang *et al.* [43] has been validated for carbon systems spanning mixed sp^2/sp^3 C–C bonding carbon-based materials and a broad range of densities, from slightly perturbed diamondlike structures to high-energy amorphous phases. Although the NEP model performs well in many cases, particularly for metastable configurations, some geometries are better captured by an empirical potential, which we use to complement our dataset. These preoptimizations are performed in GPUMD under the NVT ensemble with a Berendsen thermostat [44] at 1 K for 5×10^2 steps using a 0.5 fs time step. Each structure is first rapidly preoptimized using the interatomic potential, enabling efficient identification of low-energy geometries and substantially reducing the number of costly, high-precision DFT calculations required for final screening.

The stability of various defect forms was evaluated based on the calculated formation energies. Specifically, we computed the formation energy E_f^q for defects in a charge state q as follows [45]:

$$E_f^q = E_{\text{tot}}^q - E_{\text{bulk}} - \sum_i n_c \mu_c + qE_{\text{Fermi}} + E_{\text{corr}}^q, \quad (1)$$

where E_{tot}^q and E_{bulk} are the total energies of defective and pristine diamond, respectively, μ_c is the chemical potential of carbon, n_c is a the number of added carbons as interstitials to the bulk system, E_{Fermi} is the Fermi energy, referenced to the valence band maximum (VBM) and E_{corr}^q is a correction term for the total energies of charged supercells [46]. The chemical potential of carbon was calculated from the diamond reference as the total energy per carbon atom.

The LOBSTER package [47] was used to assess defect bonding strength via the integrated crystal orbital Hamilton population (ICOHP). Within this framework, more negative ICOHP values correspond to stronger, more bonding interactions between the two carbon atoms, whereas less negative (or positive) values indicate weaker or antibonding interactions.

The Raman activity of a phonon mode m with frequency ω_m and eigenvector $\mathbf{v}_m$ is determined by the derivative of the polarizability tensor χ with respect to $\mathbf{v}_m$. Within DFT calculations, this quantity can be equivalently evaluated from the derivative of the macroscopic dielectric tensor ϵ_{mac} as

$$R_m \propto \frac{\partial \chi}{\partial \mathbf{v}_m} = \frac{\partial \epsilon_{\text{mac}}}{\partial \mathbf{v}_m}. \quad (2)$$

The nonresonant first-order Raman intensity is expressed as

$$I_m^R \propto |\mathbf{e}_s \cdot R_m \cdot \mathbf{e}_i|^2, \quad (3)$$

where $\mathbf{e}_i$ and $\mathbf{e}_s$ denote the polarization vectors of the incident and scattered light, respectively. Spatial averaging over all possible scattering geometries yields the unpolarized Raman intensity used in this work.

The first-order IR intensity is obtained from the derivative of the total dipole moment μ with respect to the mode eigenvector,

$$I_m^{\text{IR}} \propto \left| \frac{\partial \mu}{\partial \mathbf{v}_m} \right|^2. \quad (4)$$

To evaluate the Raman and IR activities according to each vibrational mode, we compute Eqs. (2) and (3) by displacing the atoms along both the positive and negative directions defined by the phonon eigenvectors. This procedure requires $6N - 6$ single-point calculations, where N is the number of atoms in the supercell. Phonon modes were obtained using the finite-difference method as implemented in VASP. The resulting eigenvectors were mass-scaled to yield normalized displacement vectors. The displacement amplitude was set to 0.015 Å. To generate continuous spectra, all discrete mode intensities were convoluted using Gaussian functions with a broadening of 5 cm^{-1} .

Photoluminescence (PL) line shape of the defects was evaluated within the Huang-Rhys formalism [48,49]. This approach requires (1) the adiabatic potential energy surfaces of the ground and excited electronic states and (2) the vibrational normal modes that couple to the optical transition. The excited-state geometry was obtained using the

constrained-occupation Δ SCF method as was explained for nitrogen-vacancy center in diamond [50]. In this method, the electronic configuration of the target excited state is enforced by constraining the occupations of the Kohn-Sham (KS) orbitals, and the ionic positions are then relaxed in exactly the same way as for the ground state. This yields the minima of the potential-energy surfaces in both the ground state and the excited state. Denoting the corresponding total energies by $E_g^{\min}$ and $E_e^{\min}$, the ZPL energy is obtained as $E_{\text{ZPL}} = E_e^{\min} - E_g^{\min}$. The relaxed geometries were then used as reference structures for the analysis of phonon modes that participate in the optical transition. The electron-phonon interactions can be formalized using the partial Huang-Rhys factor (HRF) as described below:

$$S_\lambda = \omega_\lambda Q_\lambda^2 / 2\hbar, \quad (5)$$

where $\hbar$ is the reduced Planck constant and the configurational coordinate Q_λ for an optical process is defined as

$$Q_\lambda = \sum_\alpha \sqrt{m_\alpha} \langle (\mathbf{R}_{e,\alpha} - \mathbf{R}_{g,\alpha}) | \mathbf{u}_{\alpha,\lambda} \rangle. \quad (6)$$

Here, $\mathbf{R}_g$ and $\mathbf{R}_e$ are ground- and excited-state atomic coordinates, while $\mathbf{u}_{\alpha,\lambda}$ indicates the normalized displacement vector corresponding to mode λ with frequency ω_λ and m_α is mass of atom α . The total HRF and Q^2 are defined as $\text{HRF} = \sum S_\lambda$ and $Q^2 = \sum Q_\lambda^2$, respectively. They provide measures for the average number of phonons involved in the emission and the difference between the initial and final state geometries. From the HRF, the DWF can be defined as $\text{DWF} = \exp(-\text{HRF})$. The spectral function that underlies the computation of the PSB is

$$S(\omega) = \sum_\lambda S_\lambda \delta(\omega - \omega_\lambda), \quad (7)$$

where the δ function is replaced by Gaussian with a broadening of 3 meV. Once the electron-phonon spectral function is computed, the spectral distribution function can be determined as

$$A(E_{\text{ZPL}}/\hbar - \omega) = \int dt \exp(S(0) - S(t)) \exp(i\omega t - \gamma|t|), \quad (8)$$

where $S(t)$ is the Fourier transform of $S(\omega)$ and γ determines the lifetime broadening of individual contributions. The PL intensity is proportional to $\omega^3 A(\omega)$.

III. RESULTS AND DISCUSSION

In the remainder of this paper, defect complexes are labeled according to the number of interstitial carbon atoms they contain. Mono- to hexa-interstitial clusters are denoted as C_{1i}^n , C_{2i}^n , C_{3i}^n , C_{4i}^n , C_{5i}^n , and C_{6i}^n , where the subscript n labels distinct isomers within each cluster size.

A. Structural properties

During annealing, defects with low formation energies are expected to be stable, although metastable configurations can also persist when the energy barriers to their annihilation are sufficiently high. Therefore, we applied an energy threshold, retaining as metastable candidates only those structures whose DFT-calculated relative formation energy lies within 4 eV

TABLE I. Relative energies (eV) of the $I_1^{(001)}$ configuration for different symmetries and spin states obtained with LDA, PBE, and HSE06. The lowest-energy configuration for each functional is taken as the zero reference. The value reported in parentheses is calculated within the von Barth procedure.

Sym.	Spin	LDA	PBE	HSE06
D_2	Singlet	0.02 (0.04)	0.33 (0.67)	0.27 (0.55)
D_2	Triplet			
D_{2d}	Singlet	0.49 (0.97)	0.79 (1.59)	0.00
D_{2d}	Triplet	0.00	0.00	0.02 (0.04)

of the lowest-energy configuration. Structures exceeding this threshold are unlikely to form or survive under the annealing conditions considered. The resulting optimized geometries are shown in Fig. 1, ordered by increasing energy.

For the C_{1i} defects, two distinct configurations were identified: (001)-oriented split interstitial ($I_1^{(001)}$) and bond-centered defect (cf. Fig. 1, C_{1i}^a and C_{1i}^b , respectively).

Our HSE06 calculations indicate that the neutral $I_1^{(001)}$ in diamond has a singlet ($S = 0$) ground state, with a triplet ($S = 1$) state lying approximately 20 meV higher in energy. In fact, this energy difference equals 40 meV when the von Barth procedure is employed [51], which avoids spin-state mixing. These findings agree with the previous theoretical studies [29–31]. In addition, experimental measurements have identified a defect, known as the R2 center, exhibiting EPR activity, which has recently been shown through ZFS calculations to correspond to the $I_1^{(001)}$ configuration [33]. Its temperature-dependent EPR intensities place the triplet roughly 50 meV above the singlet ground state, in close agreement with our result [52].

We find that the structural symmetry of $I_1^{(001)}$ exhibits a pronounced dependence on both the xc-functional and the spin multiplicity, as summarized in Table I. Using semilocal functionals, namely, the local-density approximation (LDA) [27] and PBE, we obtain the triplet D_{2d} configuration as the ground state. In contrast, HSE06 stabilizes a singlet state and consequently predicts the triplet configuration to be metastable. When switching from the stable to the metastable configuration, the symmetry changes to D_2 in PBE and LDA calculations, whereas in HSE06 the structure remains in the D_{2d} symmetry. In the former cases, this behavior can be interpreted as a Jahn-Teller distortion: The singlet state with D_2 symmetry is about 0.45 eV lower in energy than its D_{2d} counterpart, in good agreement with the value reported by Goss *et al.* [31]. It is noteworthy that in HSE06 calculations the D_{2d} geometry is 0.27 eV lower in energy than the D_2 configuration.

Our results also indicate that the bond-centered defect is 2.37 eV higher in energy than the more stable $I_1^{(001)}$ configuration. This defect with a D_{3d} symmetry stabilizes in a triplet spin state while having a singlet state at 0.85 eV higher energy, in agreement with the previous reports [23,28,53].

Three $2C_i$ configurations, known as π -bonded, Humble, and di-(001) split interstitial ($I_2^{(001)}$), shown in Fig. 1 as C_{2i}^a - C_{2i}^c , comprise our dataset, listed in order of increasing energy. Further structural details are also provided in

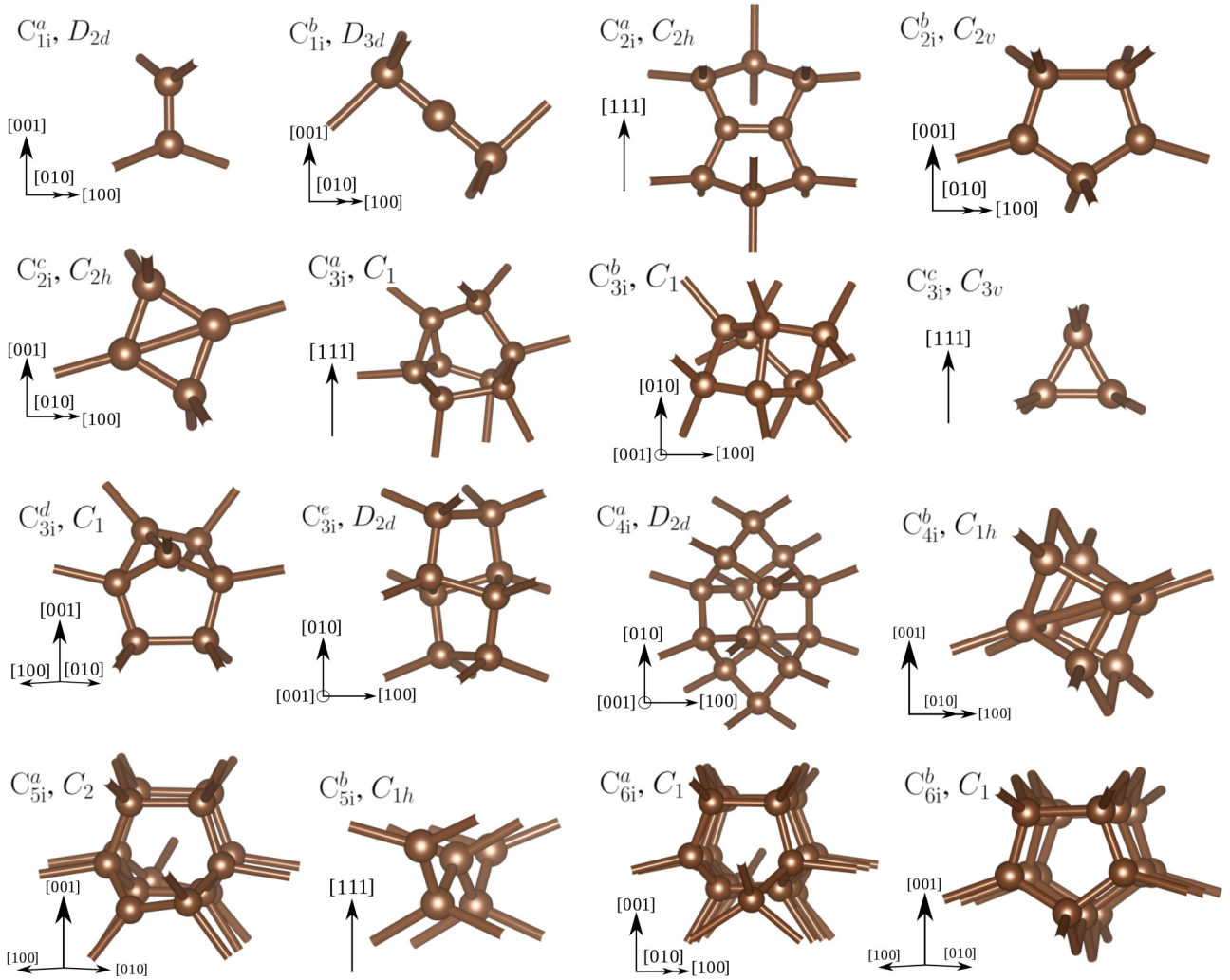


FIG. 1. Optimized geometries of carbon self-interstitial configurations, ordered by increasing energy, together with their corresponding structural symmetries. Panels: C_{1i}^a and C_{1i}^b mono carbon interstitials; C_{2i}^a – C_{2i}^c di-carbon; C_{3i}^a – C_{3i}^e tri carbon; C_{4i}^a and C_{4i}^b tetra carbon; C_{5i}^a and C_{5i}^b penta carbon; and C_{6i}^a and C_{6i}^b hexa carbon.

Refs. [30,31,54]. The π -bonded di-interstitial, formed by two bond-centered mono-interstitials on opposite sides of the same hexagonal ring, exhibits C_{2h} symmetry and a singlet ground state. The Humble configuration consists of two $I_1^{(001)}$ units occupying *next-nearest-neighbor* sites. This exhibits C_{2v} symmetry and a singlet ground state; the triplet state lies 0.13 eV higher in energy. Notably, the positively charged Humble defect has been proposed to be related to the 3H optical center, with a ZPL energy of 2.46 eV [31]. However, our findings indicate that such a conclusion about the charge state is incorrect because that charge state is not photostable under blue emission (see Sec. III C). We rather identify 3H optical center with the *neutral* Humble (C_{2i}^b) defect in Sec. III F. Finally, the $I_2^{(001)}$ consists of two $I_1^{(001)}$ defects occupying *nearest-neighbor* sites, exhibits C_{2h} symmetry, and has historically been attributed to the triplet-state EPR center R1 [33,54]. In agreement with prior works, we recover the same stability ordering: the π -bonded configuration is lowest in energy; the Humble one is 0.84 eV higher; and the $I_2^{(001)}$ is 1.40 eV higher, all relative to the π -bonded configuration.

We identify five different tri-interstitial complexes (Fig. 1, C_{3i}^a – C_{3i}^e); to our knowledge, the structures C_{3i}^a – C_{3i}^d have not been reported previously. Configuration C_{3i}^a , with symmetry of C_1 , consists of a triangular moiety with one bond to a Humble-like di-interstitial [cf. Fig. 1 C_{2i}^b], and the third vertex is capped by a C–C bond that links to the opposite side of the Humble unit. Note that an analogous triangular tri-interstitial motif has been discussed in silicon, so-called Coomer [55]. Complex C_{3i}^b can be viewed as comprising four split-interstitial units (cf. C_{1i}^a), three of which are nearly aligned along the $\langle 010 \rangle$ orientation, while the fourth is tilted. Alternatively, it can be described as two dihedral motifs with internal angles of 116.44° and 118.94° , related by a rotation about the $[010]$ axis. This configuration has a C_1 symmetry and lies 1.46 eV higher in energy than the most stable tri-interstitial structure. Defect C_{3i}^c also adopts a Coomer-like geometry, forming a highly compact complex that induces substantial local distortion in the crystal lattice [55,56]. This configuration exhibits C_{3v} symmetry and is energetically close to C_{3i}^b , lying 1.47 eV above the most stable tri-interstitial

structure. Configuration C_{3i}^d consists of a Humble-like di-interstitial [cf. C_{2i}^b] adjacent to a Coomer-like motif [cf. C_{3i}^c], bridged by a single additional carbon atom. It has a C_1 symmetry. Energetically, configuration C_{3i}^d lies 2.90 eV above the most stable tri-interstitial. It is structurally analogous to C_{3i}^a , but the capping carbon dimer is replaced by a single carbon atom. The final tri-interstitial configuration considered, C_{3i}^e , has D_{2d} symmetry and lies 3.48 eV above the most stable tri-interstitial structure. We note that in Ref. [30] this configuration was reported as the lowest-energy tri-interstitial. In our work, configuration C_{3i}^e is recovered: placing four split-interstitials on sites adjacent to a vacancy reconstructs into this complex. All tri-interstitial configurations are found to be nonmagnetic, a point that will be further discussed in a later section. It is worth noting that the structure reported in Ref. [30,31] was also identified in our searches; however, it lies outside the considered energy range of this study, being more than 4 eV higher in energy than the most stable configuration.

Our calculations reveal two distinct tetra-interstitial complexes. Configuration C_{4i}^a , the well-known platelet defect [57], has D_{2d} symmetry and consists of four $I_1^{(010)}$ split interstitials that fully reconstruct, eliminating all dangling bonds. This structure has been proposed as the fundamental building block of one-dimensional platelets in type I diamond, where nitrogen is present as an impurity [24,29]. Platelets themselves are extended, planar defects oriented along a single crystallographic direction. Configuration C_{4i}^b consists of two parallel di-interstitial C_{2i}^c units and lies 1.77 eV higher in energy than the platelet configuration C_{4i}^a . Both tetra-interstitial complexes are nonmagnetic.

For the penta-interstitial, we identify two novel configurations. Configuration C_{5i}^a consists of five aggregated split-interstitial units and has crystal symmetry C_2 and a triplet ground state. The singlet spin state is nearly degenerate with the triplet state, lying 5 meV lower in energy. Configuration C_{5i}^b lies 3.91 eV higher in energy than C_{5i}^a ; it is oriented perpendicular to the (111) plane and is formed by a zigzag stacking of five carbon interstitials occupying hexagonal and tetrahedral interstitial sites. This defect has a C_{1h} symmetry.

Finally, we identify two novel hexa-interstitial configurations, both formed through the aggregation of split-interstitial units. Configuration C_{6i}^a induces a pronounced lattice distortion and exhibits C_1 symmetry. This configuration is very similar to structure C_{5i}^a ; however, the presence of an additional split-interstitial, arranged so as to break one of the symmetry planes, lowers the symmetry. Here, the singlet and triplet spin states are nearly degenerate, with the singlet lying within 1 meV below the triplet. Configuration C_{6i}^b , which at first glance resembles three parallel di-interstitials of type C_{2i}^b , in fact consists of a combination of one C_{2i}^b di-interstitial and one C_{4i}^a tetra-interstitial unit. Two dangling bonds are present in the C_{2i}^b -like Humble segment of this complex. This configuration lies 3.18 eV higher in energy than the most stable hexa-interstitial structure. The results discussed above are summarized in Table II. However, singlet-triplet splittings below 10 meV remain inconclusive for identifying the true ground state within the numerical and methodological uncertainty of our approach.

TABLE II. The spin multiplicities of defects in the neutral charge state are listed according to the labels used in this work and their conventional names, found in the literature. The triplet-singlet energy difference (ΔE_{TS}) is reported in meV, with a negative value indicating a triplet ground state.

Defect	Labels	Spin ground state	ΔE_{ST}
C_{1i}^a	$\langle 001 \rangle$ split	singlet	40
C_{1i}^b	Bond-centered	triplet	−850
C_{2i}^a	π -bond	singlet	−
C_{2i}^b	Humble	singlet	130
C_{2i}^c	di- $\langle 001 \rangle$ split	triplet	−60
C_{5i}^a	C_{5i}^a	singlet	5
C_{5i}^b	C_{5i}^b	singlet	140
C_{6i}^a	C_{6i}^a	singlet	1
C_{6i}^b	C_{6i}^b	singlet	60

B. Bond analysis

Figure 2 shows the dependence of the ICOHP on the C–C bond length for the lowest-energy mono- to hexa-interstitial configurations, C_{1i}^a – C_{6i}^a .

For pristine diamond, the HSE06 calculations yield an ICOHP value of −11.4 eV, whereas PBE gives a value of −9.64 eV, consistent with the prior report [58]. Notably, even in an ideal crystal, bonds with nearly identical geometric lengths can display a small spread in ICOHP, reflecting the fact that ICOHP is governed not only by interatomic distance but also by the local electronic environment and the underlying orbital interactions. Unless stated otherwise, all defect configurations are analyzed at the HSE06 level.

For the C_{1i}^a , most C–C bonds lie within a narrow range of 1.50–1.55 Å. A distinct double-bond-like feature appears at a shorter bond length of 1.29 Å with an ICOHP of −14.97 eV; this bond is associated with the C atoms forming the $I_1^{(001)}$ configuration. The next isolated data points correspond to the second most negative ICOHP value of −12.75 eV and, in fact, represent four equivalent bonds that connect the defect center to the diamond lattice.

For configuration C_{2i}^a , the C–C bond lengths are predominantly distributed in the range 1.49–1.56 Å. In this structure, only the central π bond, with a length of 1.28 Å, and the four ancillary bonds linking the neighboring C atoms to the two π -bond-forming carbons, with lengths of 1.33 Å, exhibit pronounced shortening. Together, these two groups of bonds, characterized by ICOHP values of −16.8 and −13.7 eV, define the defect cluster. Embedding this cluster in the diamond lattice induces both compression and stretching of hosting bonds, giving rise to shorter bonds down to 1.40 Å and elongated bonds exceeding 1.60 Å.

In configuration C_{3i}^a , the C–C bond lengths are narrowly distributed within the 1.50–1.56 Å range, indicating strongly localized covalent bonding. Within the defect core, however, two significantly shortened bonds emerge at 1.35 and 1.37 Å. Notably, the more negative ICOHP value corresponds not to the shortest bond, but to the second-shortest one (1.37 Å). The resulting contraction of the defect core forces the surrounding lattice to relax outward, producing neighboring C–C bonds that are elongated relative to the pristine lattice and appear in the upper (less negative ICOHP) region of the plot.

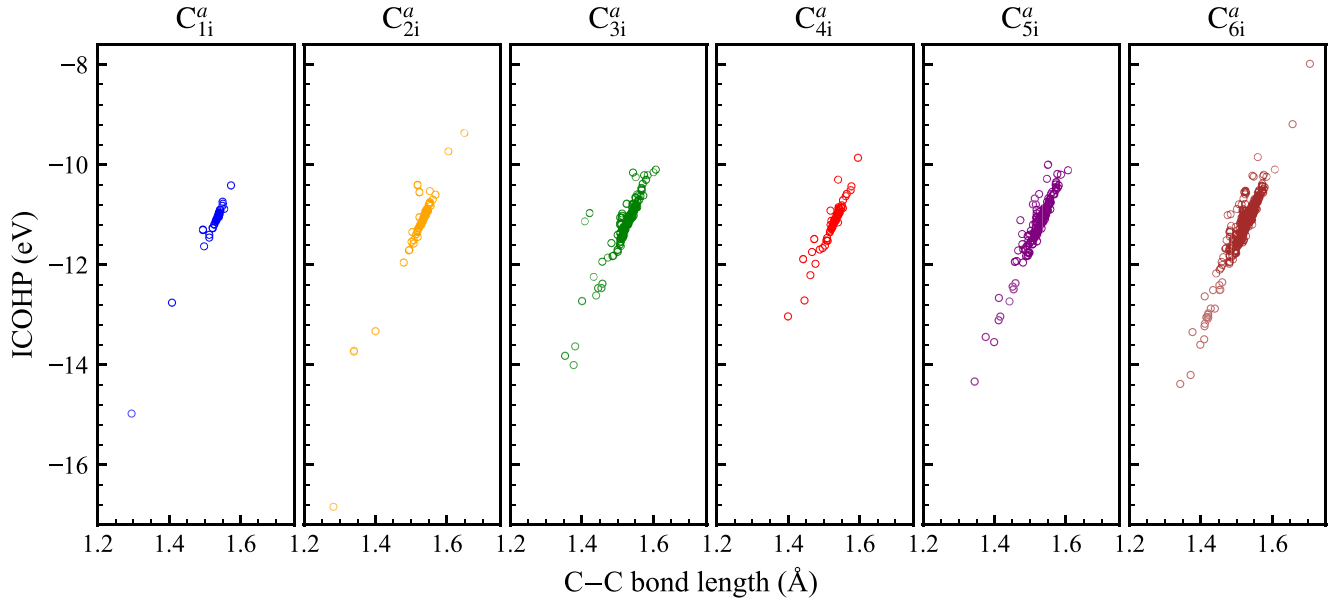


FIG. 2. The ICOHP values for the most stable of each interstitial defect vs their bond lengths. Each data point represents one C–C bond.

For the larger clusters, C_{4i}^a – C_{6i}^a , the dispersion in ICOHP broadens with increasing cluster size. This enhanced spread indicates the coexistence of comparatively strong and weak C–C interactions within a single defect complex, reflecting the increasing structural and bonding diversity as additional interstitial carbon atoms are incorporated. Across all panels, shorter C–C bonds are generally associated with more negative ICOHP values, indicating stronger bonding, although a modest scatter is present and occasional deviations occur depending on the local coordination and electronic environment. Overall, the figure supports ICOHP as a reliable descriptor of C–C bond strength across the full range of interstitial cluster sizes.

Bader charge analysis [59] shows that, across all configurations from C_{1i}^a to C_{6i}^a , a charge transfer of roughly 1.5 electrons occurs among the defect carbon atoms, with similar redistributions observed in the remaining systems. This systematic charge imbalance introduces a polar component into the core C–C bonds, partially shifting their character away from purely covalent. The resulting polarization enhances the local electrostatic attraction between the involved atoms, thereby strengthening these bonds relative to the surrounding, more covalent lattice bonds. Consequently, the bonds experiencing the largest charge transfer consistently correspond to the most negative ICOHP values within each defect structure, reflecting their increased bond strength.

C. Formation energies

We calculated the formation energy per interstitial of the optimized defects (cf. Fig. 1) to show their (relative) stability in different charge states (see Fig. 3).

For the mono-interstitial, both configurations can exist in positive and negative charge states. In the neutral charge state, configurations C_{1i}^a and C_{1i}^b have formation energies of 11.7 and 14.1 eV, respectively. Although *p*-type or *n*-type doping shifts the Fermi level and thus can modify the preferred charge

state, we find that the neutral defect remains thermodynamically stable over a wide range, indicating that it is not easily ionized under typical conditions. Generally, defect charging lowers the formation energy when electron exchange with the Fermi reservoir allows the defect to relax into a lower-energy electronic and structural configuration.

For the di-interstitial complexes, configuration C_{2i}^a is stable only in neutral and positive charge states, indicating that it has no empty defect level in the band gap that could accommodate an additional electron. In the neutral state, configurations C_{2i}^a , C_{2i}^b , and C_{2i}^c have formation energies of 8.2, 8.6, and 8.9 eV per interstitial, respectively. In contrast, configurations C_{2i}^b and C_{2i}^c are stable in positive, neutral, and negative charge states. At low and high Fermi-level energies, the positively and negatively charged states of the C_{2i}^b defect, respectively, are the most stable among the di-interstitial configurations considered, consistent with previous findings [31]. We note that these defects are photostable under visible illumination only in the neutral charge state: Both the positive and negative charge states are photoconverted to neutral via single-photon absorption for photon energies above 1.4 eV. Consequently, the assignment of the positively charged C_{2i}^b defect to the 3H optical center proposed in Ref. [31], which is excited in the blue, is no longer supported.

All tri-interstitial defects are found to be stable only in the neutral charge state. The calculated formation energies per interstitial are 5.96, 6.44, 6.47, 6.94, and 7.12 eV for configurations C_{3i}^a , C_{3i}^b , C_{3i}^c , C_{3i}^d , and C_{3i}^e , respectively. The absence of stable charged states indicates that these complexes do not introduce partially filled defect levels within the band gap; i.e., no dangling-bond-derived states are present. Consequently, no charge-transition levels appear in the gap. If all defect-related electronic states lie deep below the valence band maximum or high above the conduction band minimum, they remain inactive within the accessible Fermi level range. This implies that the tri-interstitial defects cannot be ionized and thus remain electronically neutral irrespective of *p*-type or *n*-type doping.

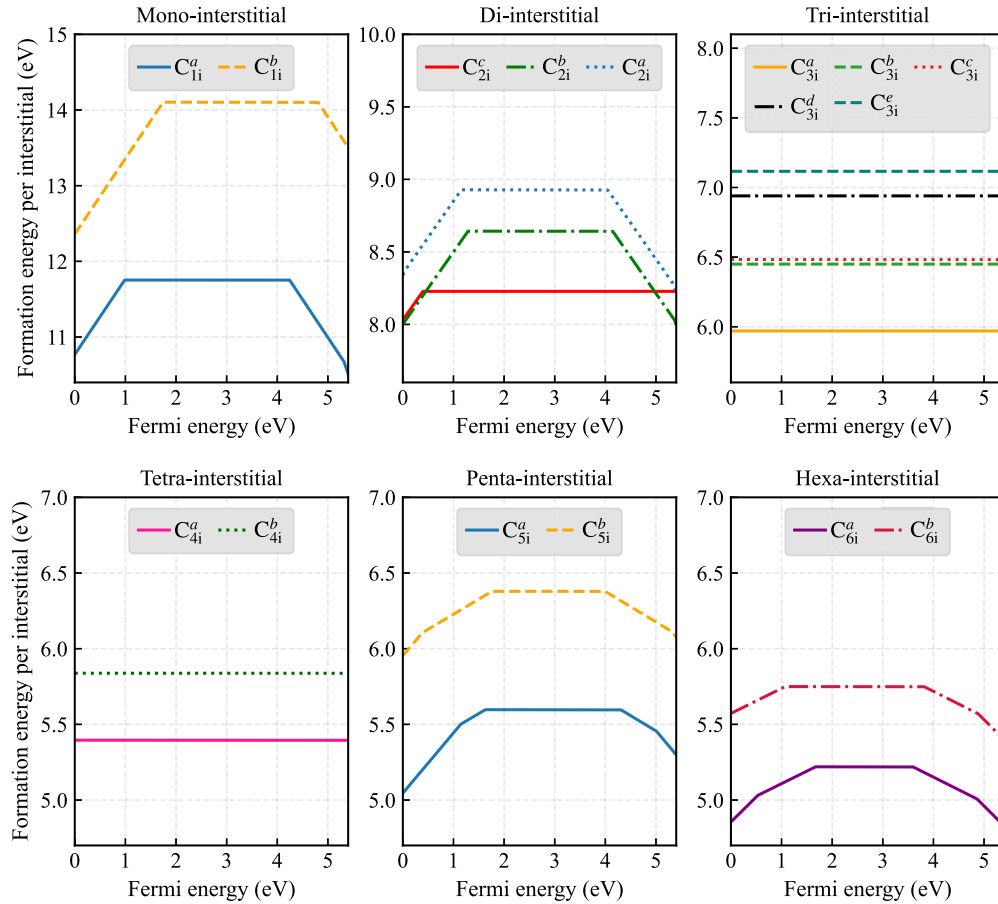


FIG. 3. Formation energy of the interstitial defects depicted in Fig. 1 as a function of the position of the Fermi level. The valence level energy is aligned to zero for the sake of simplicity. The conduction level energy is set at 5.4 eV.

The tetra-interstitial defects exhibit the same behavior as tri-interstitials: They are stable only in the neutral charge state. The calculated formation energies per interstitial are 5.40 eV for configuration C_{4i}^a and 5.84 eV for configuration C_{4i}^b .

For the penta-interstitials, configuration C_{5i}^a is stable in charge states ranging from doubly positive to doubly negative. The presence of dangling bonds in this complex gives rise to several occupied and empty defect levels within the band gap. Electrons can be added to the empty levels or removed from the occupied ones, which leads to pronounced multicharge state behavior in the formation energies as a function of the Fermi level. A similar situation occurs for configuration C_{5i}^b , which is stable in the doubly positive, singly positive, neutral, and singly negative charge states. In the neutral charge state, the formation energies per interstitial are 5.60 eV for C_{5i}^a and 6.38 eV for C_{5i}^b .

For the hexa-interstitial complexes, both configurations exhibit multiple stable charge states, similar to the penta-interstitials. In the neutral state, configurations C_{6i}^a and C_{6i}^b have formation energies per interstitial of 5.22 and 5.75 eV, respectively. Among all defects studied, C_{6i}^a shows the narrowest Fermi level range over which the neutral charge state is thermodynamically stable, implying that even modest Fermi level shifts (e.g., due to doping) will drive it into a charged state. Thus, C_{6i}^a is the most easily ionized complex in our set.

We predict an overall trend in which the formation energy per interstitial atom decreases as the number of interstitials in the complex increases, with a pronounced drop from the single and di-interstitials to the tri- and tetra-interstitial complexes. This behavior indicates a thermodynamic driving force toward aggregation, consistent with earlier reports on the relative stability of extended self-interstitial complexes in diamond [24,29]. Although this does not imply spontaneous clustering under all conditions, it shows that, once several interstitials are present, compact configurations such as the tetra-interstitial platelet C_{4i}^a are energetically more favorable than an equivalent number of isolated split interstitials.

D. Electronic structure

The ground-state electronic structures of all the most stable interstitial defects (C_{1i}^a – C_{6i}^a) are summarized in Fig. 4(a). The corresponding orbitals are visualized in Fig. 4(b).

The mono-interstitial adopts a singlet ground state with a wave function of D_{2d} symmetry. Its a_1 state is fully occupied and lies within the valence band. In this case, the presence of two dangling bonds introduced by the p orbital of the two sp^2 coordinated neighboring carbon atoms gives rise to a pair of degenerate e levels in each spin channel within the band gap.

We illustrated the orbital localization of the defect, in which the occupied e_x state is highly localized on the upper

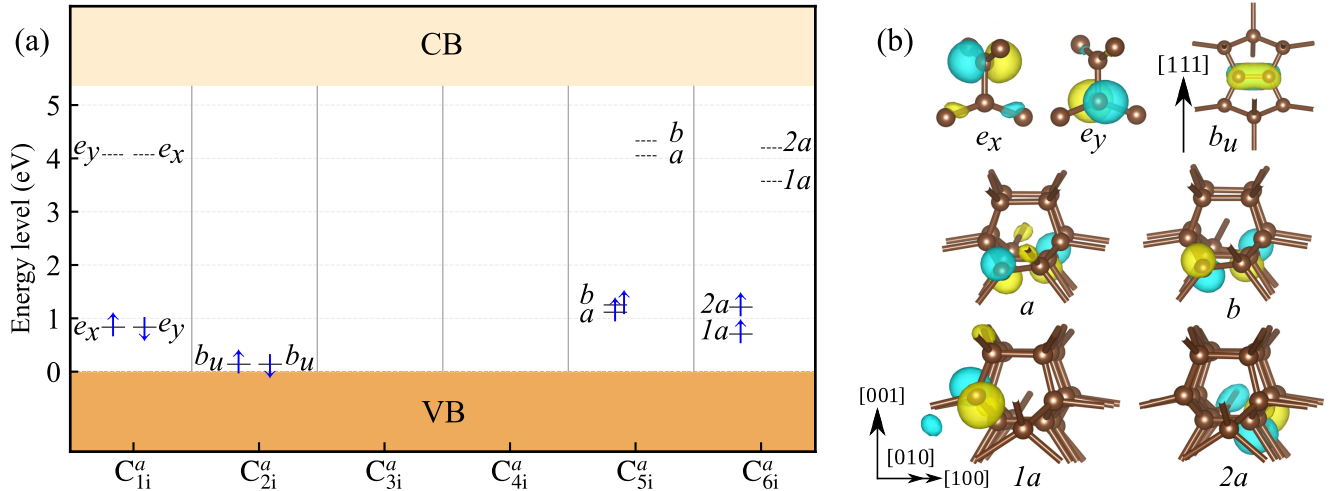


FIG. 4. (a) Electronic structure for the ground states of the neutral defects using the HSE06 functional. KS energy levels are represented by spin-up ($\uparrow$) and spin-down ($\downarrow$). The valence band (VB) and conduction band (CB) are depicted in brown and cream, respectively. (b) Representation of the KS orbitals of the relevant states. The light cyan and yellow lobes exhibit negative and positive isovalues. The isosurface absolute value is set to $7 \times 10^{-7} \text{ \AA}^{-3}$.

atom of the split interstitial with p -like character. We observe a small contribution from the neighboring atoms. This confirms that the e_x state is a defect localized pure p character, with nearly equal projection onto the p_x and p_z orbitals in the global frame and negligible s or d admixture. In addition, the unoccupied e_y level shows an analogous localization pattern on the other atom of the defect. Here again, most of the weight resides on this atom, with almost equal p_x and p_z contributions and only small s/p tails on the adjacent neighbors.

In an ideal symmetric singlet, one electron occupies each spin channel, distributed equally between the two components of the e manifold. This behavior is notably reproduced by the PBE functional: The two KS levels remain strictly degenerate, and each carries an occupation of ≈ 0.5 per spin channel, corresponding to a half-filled e doublet with the correct singlet occupation. In contrast, the HSE06 hybrid functional lifts this degeneracy. Although the two states still transform as an E representation, one component (e_x in spin-up and e_y in spin-down) is lowered in energy and becomes fully occupied, while the complementary component (e_y in spin-up and e_x in spin-down) is shifted upward by roughly 3 eV and remains empty. Thus, HSE06 drives the system from the symmetric half-half occupation characteristic of PBE to an integer-occupation solution in which one member of the e pair is filled and the other is depopulated. We emphasize that the KS level splitting is not a consequence of a Jahn-Teller distortion nor of any underlying lattice-symmetry breaking; rather, it reflects the hybrid-functional's intrinsic tendency to penalize fractional occupations, together with the increased level of separation introduced by Fock exchange. This electron configuration instead corresponds to one of the Slater determinants of the 1B_1 multiplet [31]. Promoting a spin-down electron from e_y to e_x yields a mixture of the 1A_1 and 1B_2 multiplet states, which are optically inactive in the high-symmetry configuration but may become dynamically activated through coupling to symmetry-distorting phonons [31]. Possible correlations of this defect with near-infrared and ultraviolet optical centers are discussed in Ref. [31]. Since the main focus of this work is to understand

the properties of carbon self-interstitial clusters, we continue our investigation with the di-interstitial complexes.

In the di-interstitial, with a singlet ground state, the defect wave function exhibits C_{2h} symmetry. Deep defect levels appear in the valence band, including b_g and b_u states located about 2.5 eV below the VBM energy. Within the band gap, we find only a single fully occupied b_u level, with no corresponding empty state available for electronic promotion. The associated KS orbital shows that this in-gap b_u state is strongly localized on the two interstitial atoms forming the short bond. The b_u state is a localized p - p bonding orbital of π symmetry between the two interstitial atoms. In other words, the in-gap level corresponds to a defect-localized π bond constructed predominantly from the p orbitals of the di-interstitial.

For the tri-interstitial defects, neither the most stable configuration nor the metastable ones introduce electronic states in the band gap. These configurations are therefore *silent* (i.e., electrically inert) defects: They do not generate in-gap states and thus exhibit no electrical or optical activity. However, the stable defect with C_1 symmetry possesses a deep, localized a state in the valence band, while no corresponding resonant state is found in the conduction band.

For the tetra-interstitial defects, in both stable and metastable configurations, we do not observe any states in the band gap, as all bonds are saturated. In the platelet defect, consisting of an array of I_4 units, all the carbon atoms are fourfold coordinated so no p orbitals appear due to threefold coordinated, sp^2 carbon atoms. This defect exhibits two levels, a_1 and b_2 , both of which lie deep within the valence band. Consequently, this defect is optically and electronically inactive, i.e., a silent defect. Goss *et al.* [31] previously reported only the a_1 state of this defect, but because they could not accurately determine the band edges by semilocal DFT functionals, they assumed that this level lies in the band gap. In contrast, our results show that this state, together with an additional defect state, lies deep within the valence band.

The most stable penta-interstitial defect has a triplet ground state, and the wave functions of the localized states exhibit C_2

symmetry. The spin-up channel contains fully occupied a and b states, whereas the corresponding states in the spin-down channel are empty. The orbital localization for both the a and b states shows characteristic localization on the p orbitals of the sp^2 coordinated C atoms, similar to that of a split-interstitial carbon defect. However, they are spatially separated on opposite sides of the complex.

The hexa-interstitial configuration has a triplet ground state, with two fully occupied a states in the spin-up channel and two empty a states in the spin-down channel. The corresponding defect orbitals exhibit a localization pattern similar to that of the penta-interstitial configuration.

We note that, in the interstitial defects, gap states arise whenever there are sp^2 -hybridized carbon atoms with unsaturated p_z orbitals. In contrast, the tri- and tetra-interstitial configurations contain only sp^3 -hybridized carbon atoms and therefore lack p_z orbitals that could generate localized states in the band gap. If the surrounding sp^3 bonds were significantly weakened, their associated defect levels would be expected to fall into the gap. However, as shown in Sec. III B, the tri- and tetra-interstitials actually exhibit stronger sp^3 C–C bonds than pristine diamond. Consequently, the corresponding defect states remain resonant with the bulk bands rather than forming isolated gap levels. In contrast, in the mono-, di-, penta-, and hexa-interstitials, the presence of sp^2 carbon leads to p_z -derived defect states within the diamond band gap.

All the most stable clusters that introduce midgap states and undergo electron excitation to the conduction band lead to bound-exciton excitations. These transitions typically occur in the ultraviolet regime; therefore, such defects are not considered in the present study. Based on our KS level analysis, none of the lowest-energy carbon self-interstitial cluster configurations exhibits the electronic structure required to account for the TR12 center. The TR12 center has an experimental ZPL of ~ 2.63 eV, implying that any viable defect candidate must possess two well-isolated defect states whose KS energy separation exceeds this value. No such pair of in-gap states is found for the neutral self-interstitial cluster configurations considered here. Although their formation energies suggest that these lowest-energy defects can exist in multiple charge states, which is interesting in its own right, we do not pursue them further here.

E. Vibrational properties

For pristine diamond, the nonresonant first-order Raman spectrum shows a single, sharp peak at 1300 cm^{-1} , which is slightly lower than the experimental value of 1332 cm^{-1} , corresponding to the triply degenerate zone-center optical phonon characteristic of the sp^3 -bonded lattice [60]. In contrast, no features appear in the infrared spectrum because all zone-center optical modes are IR inactive (silent) in diamond. Note that the small deviation between the experiment and theory can be attributed to the choice of functional used in our vibrational calculations [61].

Self-interstitials introduce local lattice disorder and thereby shorten phonon lifetimes. In vibrational spectroscopy, this is presented as broadening, frequency shifts, and the appearance of additional peaks in the line shapes [62]. Overall, the breakdown of translational symmetry relaxes the

momentum selection rule, enabling phonons with finite wave vectors to contribute to features that are nominally associated with Γ -point active modes. These effects give rise to defect-induced signatures in both Raman and IR spectra, which provide a sensitive fingerprint of the presence and nature of lattice imperfections.

The interstitial defects exhibit two distinct structural behaviors: In some configurations (e.g., structures C_{1i}^a , C_{2i}^a , C_{5i}^a , and C_{6i}^a in Fig. 1), the introduction of additional carbon atoms gives rise to dangling bonds, whereas in others (e.g., structures C_{3i}^a and C_{4i}^a) the surrounding lattice reconstructs such that no explicit dangling bonds are present. In both classes, our results show that the defects can nonetheless induce the formation of local sp^2 -like bonding within the diamond lattice. It is noteworthy that a change in the chemistry of the bonds can modify the IR activity of a mode, as the latter depends on changes in the dipole moment. In particular, sp^2 -hybridized carbon has a higher effective electronegativity than sp^3 -hybridized carbon owing to its larger s character, which alters the local charge distribution and, consequently, the associated dipole moment.

As shown in Fig. 2, self-interstitials give rise to bond lengths of $1.28\text{--}1.40\text{ \AA}$, up to $\sim 17\%$ shorter than in pristine diamond ($1.54\text{ \AA}$), indicating locally stronger bonds and correspondingly higher-frequency vibrational modes. These bond lengths are comparable to those in molecular hydrocarbons, spanning a single C–C bond in ethane ($1.54\text{ \AA}$), a double C=C bond in ethene ($1.34\text{ \AA}$), and a triple C $\equiv$ C bond in ethyne ($1.20\text{ \AA}$). More specifically, the C=C and C $\equiv$ C stretching modes occur at 1623 and 1974 cm^{-1} , respectively [63]. This indicates that the additional high-energy features arise directly from interstitial-induced bonding, rather than phonon folding.

Alongside the Raman and IR spectra shown in Fig. 5, we also evaluated the contribution of the defect centers to the activity of each vibrational mode. The contribution of the defect-center atoms to a given vibrational mode is quantified as

$$\rho = \sum_n \sqrt{|v_{n,x}|^2 + |v_{n,y}|^2 + |v_{n,z}|^2}, \quad (9)$$

where $v_{n,\alpha}$ denotes the eigenvector component of the defect-center atom n along the direction α , which is a real number at the Γ point. Partial intensities are then obtained by scaling the total mode activities by the corresponding ρ values. Defect-center atoms are defined as those that are strongly displaced from their ideal crystallographic sites; their numbers are 6, 6, 8, 12, 14, and 18 atoms for the C_{1i}^a – C_{6i}^a complexes, respectively.

Our results show that the Raman spectra in the lattice phonon region are broadly similar across all structures, whereas interstitial defects introduce several additional peaks. A systematic blueshift is also observed, which we attribute to internal strain in the lattice. In our calculations, the cell vectors are always fixed to those of the pristine crystal. At the highest defect density, an average strain of about 1.5% is expected; depending on the defect orientation, this strain can be uniaxial, biaxial, or triaxial. From our calculations on compressively strained lattices, uniaxial, biaxial, and triaxial strains of 1.5% shift the vibrational frequencies by approximately 31 , 48 , and 66 cm^{-1} , respectively.

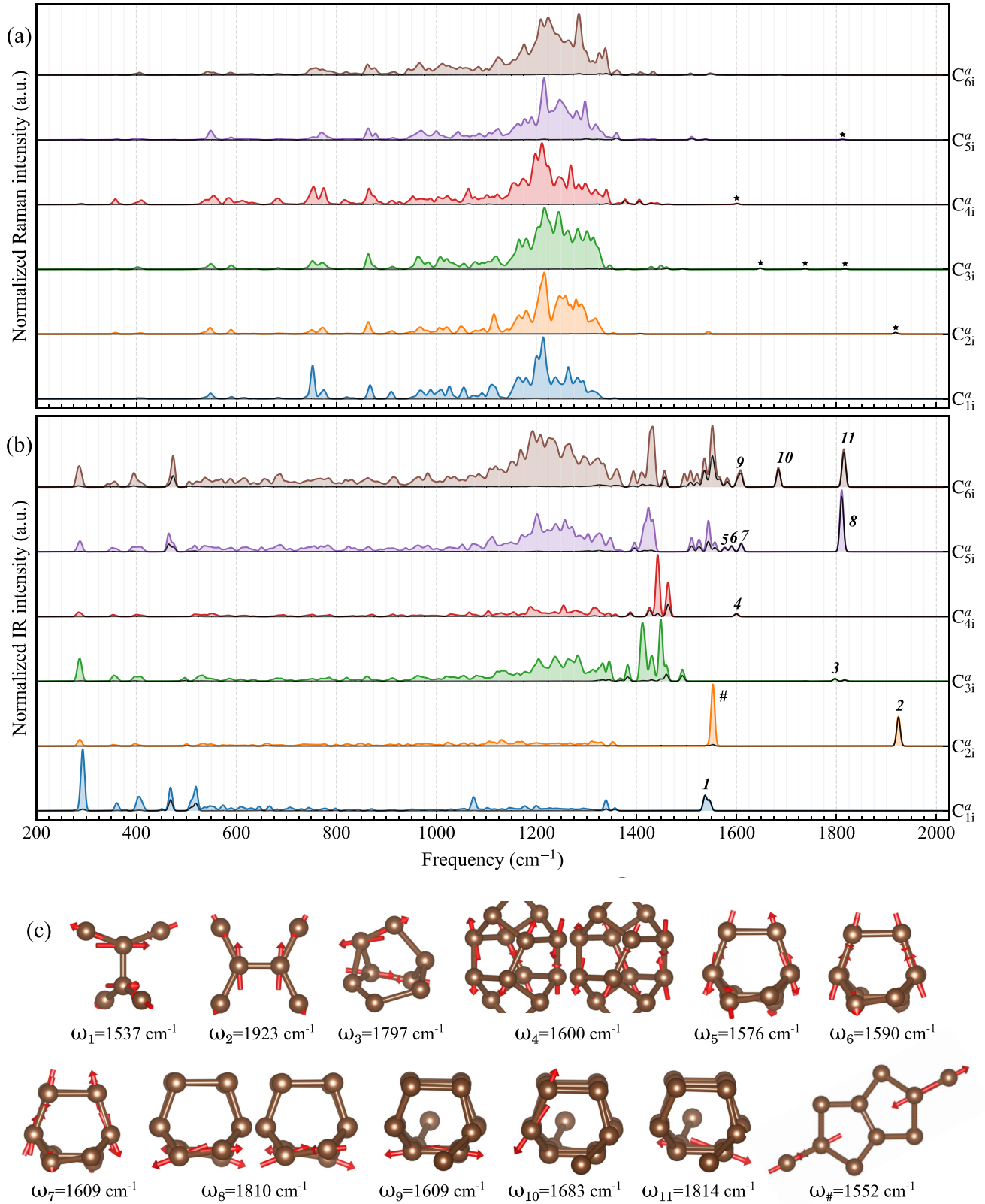


FIG. 5. (a) First-order Raman spectra for defect complexes C_{1i}^a to C_{6i}^a . Pentagram symbols mark Raman-active modes that are predominantly localized on the defect centers. Each spectrum is normalized to its own maximum (arbitrary units). The black line shows the contribution of the defect atoms to the total spectrum, which are eligible in Raman cases. (b) IR spectra of the corresponding defects, where pronounced defect-localized modes are numbered and visualized in panel (c). (c) Atomic displacement patterns of the selected IR-active modes, where only the defect atoms are active. Red arrows indicate the direction of atomic motion, and their lengths are scaled by the displacement magnitude.

Defects introduce both stiffer and softer bonds into the system. Softer bonds could, in principle, contribute to Raman activity at low frequencies; however, our results show no localized defect-related contribution in this region. In contrast, several high-frequency peaks appear between 1375 and 1925 cm^{-1} , which can be attributed to vibrations of the defect centers or to modes involving stiffer bonds between the defect centers and the host lattice. Nevertheless, the corresponding nonresonant first-order Raman activities are negligible.

The IR spectra clearly exhibit many distinct peaks, indicating that IR spectroscopy is highly sensitive to self-interstitials and can reliably discriminate between different defect types. Whereas a pristine diamond is IR inactive at the zone center because inversion symmetry forbids dipole-active phonons, interstitial defects, especially those introducing local sp^2 bonding, break this symmetry and activate vibrational modes with nonzero dipole moments. This symmetry breaking enables a range of IR-active features. Many of the stiffer, defect-induced bonds give rise to high-frequency IR modes with well-separated and clearly distinguishable peaks. There are only four peaks at low frequency, localized on the defect; two at 470 and 518 cm^{-1} in C_{1i}^a , and one in both C_{5i}^a and C_{6i}^a at 465 and 474 cm^{-1} , respectively. All high-energy IR-active modes are fully localized on the bonds associated with the defect centers, as illustrated in Fig. 5(c).

F. Identification of 3H center and its relation with TR12 center

The experimental detection of the R1 EPR center, associated with the di-interstitial C_{2i}^c , indicates that metastable configurations can likewise appear as local minima on the potential energy surface [54,64]. To this end, we also consider configurations that can form and persist under kinetic control. Among these, the C_{2i}^b defect (the Humble di-interstitial) is particularly compelling because it introduces both filled and empty levels in the fundamental band gap of diamond.

As mentioned, Goss *et al.* proposed that the 3H optical center corresponds to C_{2i}^b in the positive charge state based on the computed local vibration modes that were compared to experiments [31,65]. Our formation-energy diagram showed that under optical excitation, the positively charged defect would ionize to the neutral charge state, so it would not remain stable at the relevant excitation energy. Therefore, this attribution appears to be incorrect. On the other hand, the neutral charge state cannot be disregarded as a viable candidate that is photostable under blue illumination.

The neutral C_{2i}^b defect forms a two-state system, as shown in Fig. 6(a). It exhibits a broken-symmetry singlet ground state, as seen from orbitals' visualization in Fig. 6(b). In order to test this assertion, we performed symmetric closed-shell (CS) calculations [Fig. 6(c)], which show that its energy is 0.56 eV higher than that obtained from our open-shell (OS) singlet calculations. In a broken-symmetry system, the wave functions spontaneously break the space-group symmetry, and the KS orbitals are not necessarily eigenstates of the symmetry operations of the defective supercell [66]. Such a dilemma can be alleviated by adopting a multiplet-state representation [67,68]. Considering the a_2 and b_1 defect orbitals in C_{2v} symmetry, four states

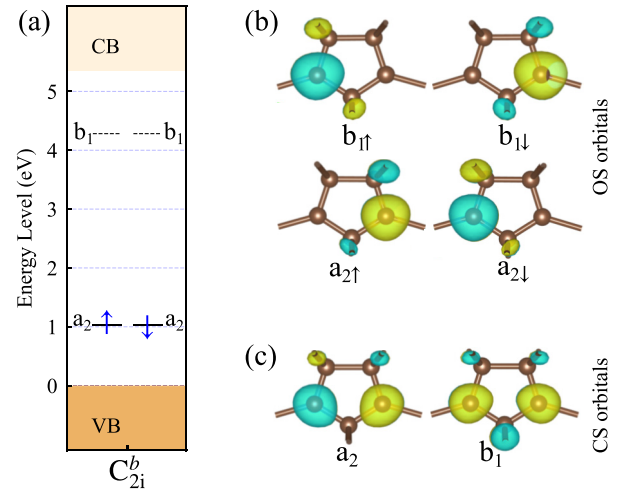


FIG. 6. (a) The Kohn-Sham levels of the open-shell singlet state of the neutral C_{2i}^b defect, represented by spin-up ($\uparrow$) and spin-down ($\downarrow$) channels. Representation of the defect-state orbitals for (b) an OS singlet broken-symmetry wave function and (c) a CS symmetric wave function treatment. The light cyan and yellow lobes exhibit negative and positive isovalues, with an absolute value of $5 \times 10^{-7} \text{ \AA}^{-3}$.

are obtained:

$$\Psi(^1A_1) = [a_2(1)a_2(2)][\uparrow\downarrow - \downarrow\uparrow], \quad (10a)$$

$$\Psi(^1A_1) = [b_1(1)b_1(2)][\uparrow\downarrow - \downarrow\uparrow], \quad (10b)$$

$$\Psi(^1B_2) = [a_2(1)b_1(2) + b_1(1)a_2(2)][\uparrow\downarrow - \downarrow\uparrow], \quad (10c)$$

$$\Psi(^3B_2) = [a_2(1)b_1(2) - b_1(1)a_2(2)] \begin{Bmatrix} [\uparrow\uparrow] \\ [\uparrow\downarrow + \downarrow\uparrow] \\ [\downarrow\downarrow] \end{Bmatrix}. \quad (10d)$$

Equation (10a) represents the CS singlet state. Equation (10b) corresponds to a double-electron excitation configuration, which lies 2.58 eV higher in energy than the ground state. However, the relevant OS singlet ground state arises when one electron occupies the spin-up channel of the a_2 orbital and the other occupies the spin-down channel of the b_1 orbital (or vice versa), corresponding to the $a_2 \otimes b_1 = ^1B_2$ configuration, as given in Eq. (10c). This state is described by a broken-symmetry wave function that reproduces the ground-state density of the multiplet state. The corresponding triplet state, 3B_2 , given by Eq. (10d), lies 0.13 eV higher in energy. The relevant excitation of the 3H center transforms the ground state [Eq. (10c)] into the excited state [Eq. (10b)] via a double-excitation process, yielding a ZPL of 2.58 eV, in good agreement with the experimental value of 2.46 eV [65]. A single-electron excitation cannot describe this transition, as it does not recover the multideterminant, spin-adapted singlet configuration required for an accurate many-body description. In principle, the two CS 1A_1 states should be treated as linear combinations (i.e., mixed). However, because the calculated energy separation between the configurations in Eqs. (10c) and (10b) is ≈ 2.0 eV, we expect the mixing to be weak and therefore regard our approach as a reasonable approximation.

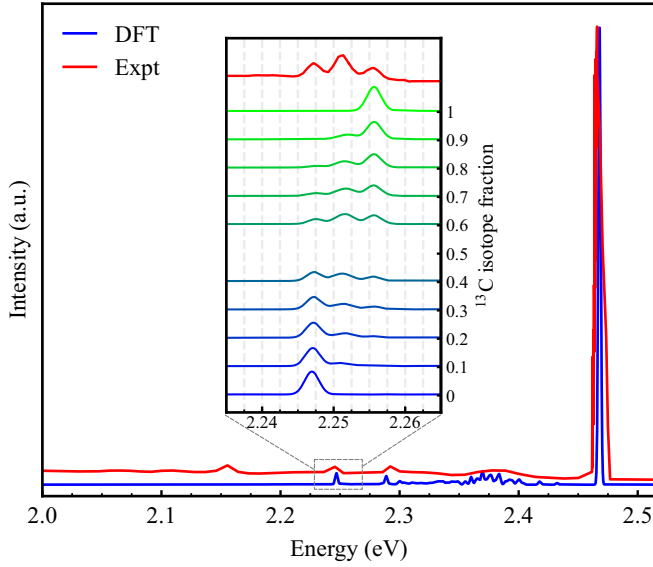


FIG. 7. The PL spectrum of the neutral C_{2i}^b defect compared with the experimental data from Ref. [65]. A rigid shift of 0.12 eV is applied to the calculated spectrum to align the zero-phonon line with experiment. The inset highlights the evolution of the PL spectrum as a function of isotope mass. The experiments were carried out for a diamond sample with a 50%–50% mixture of ^{13}C and ^{12}C isotopes (red spectrum).

To compute the vibrational properties, we adopt the optimized excited-state geometry [Eq. (10b)], which is well described as a single Slater-determinant CS configuration and can therefore be treated within the ΔSCF framework using both hybrid and semilocal density functionals. This treatment is consistent with the Huang-Rhys approximation employed here to generate the optical spectra of the defects. The calculated PL spectrum is shown in Fig. 7. After applying a rigid shift of 0.12 eV to align the calculated spectrum with experiment, we reproduce the characteristic vibrational features of the 3H center. We obtain a total Huang-Rhys factor of $S = 0.36$, corresponding to a Debye-Waller factor of 0.69 (69%).

The PL spectrum of the 3H optical center was also measured in a specially prepared diamond sample containing a 50%–50% mixture of ^{13}C and ^{12}C isotopes [65]. This isotopic disorder gives rise to additional features in the PL spectrum; most notably, a pronounced triple peak emerges in the phonon sideband (PSB) around 2.25 eV, which has been associated with local vibrational modes of ~ 230 meV. Isotope-induced shifts and splittings in the PSB provide a sensitive means to validate microscopic models of optical centers. In addition to the energies of the local modes, their relative intensities in the PL spectrum are also informative, and are accessible within the Huang-Rhys framework used here. Isotope enrichment systematically renormalizes vibrational frequencies: Increasing the average atomic mass lowers phonon frequencies and compresses the PSB toward the ZPL. To reproduce the experimentally observed triple-peak structure in the 2.24–2.26 eV range, we generated 50 random isotope configurations and report the ensemble-averaged line shape. Throughout, we assume that isotope substitution leaves the interatomic force

constants essentially unchanged; accordingly, only the atomic masses were modified when constructing the dynamical matrix (see, e.g., Ref. [69]). The dynamical matrix was then solved separately for each configuration.

Mass substitution results in an approximately 10 meV shift of the PSB. The characteristic triple-peak structure is most pronounced at a 50% ^{13}C concentration, yielding excellent agreement with experiment. This agreement provides additional support for assigning the 3H center to the neutral C_{2i}^b defect. Finally, the experimentally observed peak at 2.15 eV, which is absent from our calculations, likely originates from residual laser light or a different color center rather than a 3H-related emission feature.

We note that the 3H center is particularly interesting because it features a singlet ground state and an experimentally confirmed photoactive singlet excited state [65], together with a low-lying triplet manifold. This level structure could be exploited for qubit operation, where the metastable triplet state may serve as an optically addressable qubit. However, the large energy separation between the optically excited singlet and the metastable triplet is expected to suppress intersystem crossing (ISC), potentially limiting spin polarization and readout. One possible route to enhance ISC is a two-color excitation scheme, in which a second laser resonantly drives vibronic transitions to access spin-orbit-coupled channels more efficiently, thereby enabling qubit operation. By establishing this identification and directly overlaying the experimentally measured PL spectra of the 3H and TR12 centers, several notable similarities become apparent. Beyond the proximity of their ZPLs—2.46 eV for 3H and 2.63 eV for TR12—their PSBs exhibit multiple common spectral features [18,65]. Furthermore, annealing studies show opposing trends in their intensities: The TR12 signal generally increases with annealing, whereas the 3H signal decreases [12,70]. This anticorrelation suggests that TR12 may form as a secondary product associated with the transformation of the 3H center. Motivated by this possibility, we systematically screened our defect dataset for more complex configurations that intrinsically incorporate the Humble structural motif (i.e., C_{2i}^b). This search ultimately highlighted the C_{6i}^b defect as a prime candidate.

The C_{6i}^b defect is composed of C_{2i}^b and C_{4i}^a . While the isolated C_{4i}^a defect does not introduce localized states within the band gap and instead produces a deep valence-band state, it plays a crucial role when incorporated into the C_{6i}^b configuration. Here, we label the character of the electronic states as those for the isolated Humble C_{2i}^b defect with C_{2v} symmetry for the sake of simplicity; the symmetry of the composed defect is rather C_{1h} . As shown in Fig. 8(a), the C_{4i}^a -related state, labeled a_1 , shifts toward the valence-band maximum and thereby may participate in the electronic excitation process. When considering the a_2 and b_1 states, this configuration exhibits an electronic structure similar to that of the C_{2i}^b defect, where the singlet ground state is described by a broken-symmetry solution. The singlet-triplet energy difference is 60 meV (see Table II).

For two electrons occupying a four-state system, the many-body states can be constructed in the same manner as for the 3H configuration discussed above. The broken-symmetry singlet ground state, $a_2 \otimes b_1 = B_2$, is followed by the

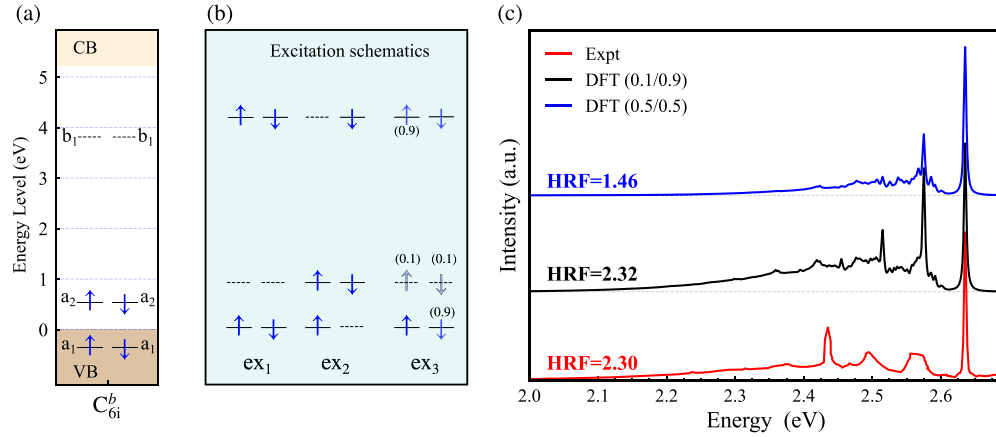


FIG. 8. (a) The Kohn-Sham energy levels of the open-shell singlet state of the neutral C_{6i}^b defect, shown separately for the spin-up ($\uparrow$) and spin-down ($\downarrow$) channels. (b) Schematic illustration of the electron excitation mechanisms investigated in the search for the TR12 PL signal. (c) The PL spectra of the neutral C_{6i}^b defect compared with the experimental data taken from Ref. [18]. For better comparison with the experiment, rigid energy shifts of 0.10 and -0.42 eV were applied to the calculated spectra to align the ZPL with the experimental value. Our calculated ZPL for the occupation numbers of 0.1 and 0.5 are 2.53 and 3.06 eV, respectively.

closed-shell $a_2 \otimes a_2 = A_1$ state, which lies 0.73 eV higher in energy. The 3B_2 triplet state is found at 0.06 eV above the broken-symmetry singlet state. The $b_1 \otimes b_1 = A_1$ configuration arises from a double-electron excitation and lies 2.46 eV above the ground state. However, incorporating the two extra states originating from C_{4i}^a leads to a six-state system with 15 many-body configurations (six singlets and three triplets with nine components), reflecting more complexity. To identify the relevant excitation process, we use the experimental ZPL (2.63 eV) and HRF (2.30) of the TR12 center as reference criteria. As illustrated schematically (Fig. 8), one of the lowest optical transitions, denoted ex_1 , corresponds to a double-electron excitation from the a_2 to the b_1 level, yielding a computed ZPL of 2.46 eV and an HRF of 3.22. Another excitation, ex_2 , involves a single-electron promotion from the a_1 to the b_1 state, for which we obtain a ZPL of 3.52 eV and an HRF of 0.84. Comparison of these two candidate excitations suggests that the system behaves as a strongly interacting two-electron two-hole manifold, in which the actual optical transition may arise from a mixture of the ex_1 and ex_2 configurations rather than from a pure single excitation. To probe this possibility, we introduced fractional occupations of the two excited configurations as a proxy for a mixed many-body state [see ex_3 in Fig. 8(b)]. Specifically, we considered fractional occupations of (0.1/0.9) and (0.5/0.5) for the two excitation components. Figure 8(c) shows the PL line shapes. For the (0.5/0.5) occupation, we obtain a ZPL of 3.06 eV and an HRF of 1.46. The (0.1/0.9) occupation yields a ZPL of 2.53 eV with an HRF of 2.32, which is closer to the experimental values. However, the corresponding PSB structure does not fully reproduce the experimental spectral features. We attribute this discrepancy either to limitations of the computational methodology employed in this work or to finite-size effects arising from the relatively small supercell used for the six-interstitial defect.

We propose that C_{6i}^b may represent a potential candidate for the TR12 center; however, a definitive assignment requires higher-level excited-state methods capable of simultaneously capturing electron-electron correlation and vibronic coupling

in the excited state. Such an investigation using multireference many-body approaches is beyond the scope of the present work. Nevertheless, based on the energetic ordering, ZPL alignment, and qualitative agreement of the PL spectral features, we conclude that the C_{6i}^b defect remains a viable candidate for the TR12 center.

IV. CONCLUSIONS

We have carried out a systematic investigation of carbon self-interstitial complexes in diamond, spanning from mono- to hexa-interstitial defects. To overcome the severe configurational complexity that emerges for larger clusters, we employed interatomic potential-assisted structure searches, followed by full relaxation and energetics within DFT. Accurate hybrid-functional calculations were used to ensure reliable defect stability ordering and defect-level alignments. This multiscale protocol enabled us to produce all established low-energy configurations for small interstitials and to identify several previously unreported tri-, penta-, and hexa-interstitial geometries.

Our results confirm that the neutral $\langle 001 \rangle$ -oriented split interstitial is lowest-energy mono-interstitial configuration, while the commonly reported lower-symmetry alternative arises from the limitations of semilocal functionals. For di-interstitials, we recover the known stability hierarchy, with the π -bonded complex as the ground state. For higher-order complexes, we find a clear thermodynamic drive toward aggregation: The formation energy per interstitial decreases with cluster size, showing a pronounced drop from isolated and di-interstitials to compact tri- and tetra-interstitials. The tetra-interstitial platelet is predicted to be a particularly stable motif, consistent with its proposed role as a building block for extended $\{100\}$ planar platelets. As additional carbon atoms are incorporated, the bond-length window remains narrow, yet the spread in bond strengths increases, reflecting the coexistence of very stiff defect-core bonds and weaker defect-lattice linkages.

The electronic structures reveal two distinct classes of self-interstitial complexes. Mono-, di-, penta-, and hexa-interstitials can host localized in-gap states and multiple charge states, associated with dangling bonds or partly reconstructed motifs. In contrast, all low-energy tri- and tetra-interstitials are electronically silent; they introduce no gap states and remain stable only in the neutral charge state, highlighting their relevance as optically and electrically silent defects in irradiated and annealed diamond. Crucially, none of the *lowest-energy* neutral self-interstitial complexes exhibits the pair of well-isolated in-gap states required to account for the TR12 center with its ~ 2.63 eV ZPL. This observation rules out the most stable self-interstitial aggregates as microscopic models of TR12. Motivated by this discrepancy, we extended our analysis to metastable configurations that may form and persist under kinetic control. Within this broader configurational landscape, our investigation of metastable configurations identifies the 3H defect center as a neutral di-interstitial defect. Building on this identification, we further propose that the TR12 center may originate from a 3H-containing defect linked to a metastable hexa-interstitial configuration.

Vibrational spectroscopy provides complementary fingerprints of defects and might be the only experimental method for exploring electronically inactive defects. For the most stable complexes, we analyze the Raman and infrared signatures associated with these defects. In Raman spectra, the lattice-phonon region remains broadly similar across defects, aside from systematic blueshifts originating from internal strain induced by fixed pristine lattice vectors. Defect-center vibrations introduce additional Raman peaks at high frequency, but their nonresonant first-order activities are negligible. In contrast, IR spectra show rich, defect-specific signatures.

Local symmetry breaking, often accompanied by sp^2 -like rehybridization, activates dipole-allowed modes that are absent in pristine diamond. Short, stiff C–C bonds within the defect cores yield well-separated high-frequency IR peaks (1375 – 1925 cm^{-1}), and mode-projection analysis confirms that these features are strongly localized on defect-center bonds. These results establish IR spectroscopy as a particularly powerful probe for discriminating among self-interstitial complexes.

Overall, this work delivers a unified and functionally robust picture of the stability, electronic activity, and vibrational signatures of carbon self-interstitial aggregates in diamond, and establishes a consistent framework for linking first-principles defect models with experimental identification.

ACKNOWLEDGMENTS

A.G. acknowledges support from the Quantum Information National Laboratory of Hungary, funded by the National Research, Development, and Innovation Office (NKFIH) under Grant No. 2022-2.1.1-NL-2022-00004, as well as the EU QuantERA II MAESTRO project (NKFIH Grant No. 2019-2.1.7-ERA-NET-2022-00045). Á.G. further acknowledges access to high-performance computational resources provided by KIFÜ (Governmental Agency for IT Development, Hungary), and funding from the European Commission through the QuSPARC (Grant No. 101186889) and SPINUS (Grant No. 101135699) projects.

DATA AVAILABILITY

The data that support the findings of this study are available in Ref. [71].

-
- [1] R. Kalish, Diamond as a unique high-tech electronic material: Difficulties and prospects, *J. Phys. D: Appl. Phys.* **40**, 6467 (2007).
 - [2] R. J. Nemanich, J. A. Carlisle, A. Hirata, and K. Haenen, CVD diamond—Research, applications, and challenges, *MRS Bull.* **39**, 490 (2014).
 - [3] J. Zhang, J. Wang, G. Zhang, Z. Huo, Z. Huang, and L. Wu, A review of diamond synthesis, modification technology, and cutting tool application in ultra-precision machining, *Mater. Des.* **237**, 112577 (2024).
 - [4] D. S. Misra, Nitrogen centers in single crystal diamond grown by chemical vapor deposition, *J. Vac. Sci. Technol. A* **41**, 042703 (2023).
 - [5] A. Zaitsev, N. Kazuchits, V. Kazuchits, K. Moe, M. Rusetsky, O. Korolik, K. Kitajima, J. Butler, and W. Wang, Nitrogen-doped CVD diamond: Nitrogen concentration, color and internal stress, *Diam. Relat. Mater.* **105**, 107794 (2020).
 - [6] C. M. Breeding and J. E. Shigley, The “type” classification system of diamonds and its importance in gemology, *Gems Gemol.* **45**, 96 (2009).
 - [7] J. Rovny, Z. Yuan, M. Fitzpatrick, A. I. Abdalla, L. Futamura, C. Fox, M. C. Cambria, S. Kolkowitz, and N. P. de Leon, Nanoscale covariance magnetometry with diamond quantum sensors, *Science* **378**, 1301 (2022).
 - [8] G. Wolfowicz, F. J. Heremans, C. P. Anderson, S. Kanai, H. Seo, A. Gali, G. Galli, and D. D. Awschalom, Quantum guidelines for solid-state spin defects, *Nat. Rev. Mater.* **6**, 906 (2021).
 - [9] G. Zhang, Y. Cheng, J.-P. Chou, and A. Gali, Material platforms for defect qubits and single-photon emitters, *Appl. Phys. Rev.* **7**, 031308 (2020).
 - [10] G. Balasubramanian, I. Y. Chan, R. Kolesov, M. Al-Hmoud, J. Tisler, C. Shin, C. Kim, A. Wojcik, P. R. Hemmer, A. Krueger, T. Hanke, A. Leitenstorfer, R. Bratschitsch, F. Jelezko, and J. Wrachtrup, Nanoscale imaging magnetometry with diamond spins under ambient conditions, *Nature (London)* **455**, 648 (2008).
 - [11] J. Foglászinger, A. Denisenko, T. Kornher, M. Schreck, W. Knolle, B. Yavkin, R. Kolesov, and J. Wrachtrup, TR12 centers in diamond as a room temperature atomic scale vector magnetometer, *npj Quantum Inf.* **8**, 65 (2022).
 - [12] K. Iakubovskii and G. Adriaenssens, Optical study of some interstitial-related centres in CVD diamond, *Phys. Status Solidi A* **181**, 59 (2000).
 - [13] J. H. N. Loubser and J. A. van Wyk, Electron spin resonance in the study of diamond, *Rep. Prog. Phys.* **41**, 1201 (1978).

- [14] C. D. Clark, R. W. Ditchburn, and H. B. Dyer, The absorption spectra of natural and irradiated diamonds, *Proc. R. Soc. Lond. A* **234**, 363 (1956).
- [15] J. Walker, An optical study of the TR12 and 3H defects in irradiated diamond, *J. Phys. C* **10**, 3031 (1977).
- [16] G. Davies, C. Foy, and K. O'Donnell, The TR12 vibronic band in diamond, *J. Phys. C* **14**, 4153 (1981).
- [17] L. Allers, A. T. Collins, and J. Hiscock, The annealing of interstitial-related optical centres in type II natural and CVD diamond, *Diam. Relat. Mater.* **7**, 228 (1998).
- [18] B. Naydenov, R. Kolesov, A. Batalov, J. Meijer, S. Pezzagna, D. Rogalla, F. Jelezko, and J. Wrachtrup, Engineering single photon emitters by ion implantation in diamond, *Appl. Phys. Lett.* **95**, 181109 (2009).
- [19] S. P. Turaga, H. Jin, E. J. Teo, and A. A. Bettiol, Cross-sectional hyperspectral imaging of proton implanted diamond, *Appl. Phys. Lett.* **115**, 021904 (2019).
- [20] X. Chai, H. Liang, C. Yang, V. X. Ho, E. J. Teo, Z. Pastuovic, and A. Bettiol, High sensitivity spin defects in carbon implanted diamond, *Adv. Opt. Mater.* **13**, 2500397 (2025).
- [21] G. Davies, *Properties and Growth of Diamond*, EMIS Datareviews Series (INSPEC, the Institution of Electrical Engineers, London, 1994).
- [22] M. Newton, B. Campbell, D. Twitchen, J. Baker, and T. Anthony, Recombination-enhanced diffusion of self-interstitial atoms and vacancy–interstitial recombination in diamond, *Diam. Relat. Mater.* **11**, 618 (2002).
- [23] S. J. Breuer and P. R. Briddon, *Ab initio* investigation of the native defects in diamond and self-diffusion, *Phys. Rev. B* **51**, 6984 (1995).
- [24] J. P. Goss, B. J. Coomer, R. Jones, C. J. Fall, P. R. Briddon, and S. Öberg, Extended defects in diamond: The interstitial platelet, *Phys. Rev. B* **67**, 165208 (2003).
- [25] A. Gali, *Ab initio* theory of the nitrogen-vacancy center in diamond, *Nanophotonics* **8**, 1907 (2018).
- [26] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized gradient approximation made simple, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [27] J. P. Perdew and A. Zunger, Self-interaction correction to density-functional approximations for many-electron systems, *Phys. Rev. B* **23**, 5048 (1981).
- [28] J. Bernholc, A. Antonelli, T. M. Del Sole, Y. Bar-Yam, and S. T. Pantelides, Mechanism of self-diffusion in diamond, *Phys. Rev. Lett.* **61**, 2689 (1988).
- [29] J. P. Goss, B. J. Coomer, R. Jones, C. J. Fall, C. D. Latham, P. R. Briddon, and S. Öberg, Small aggregates of interstitials and models for platelets in diamond, *J. Phys.: Condens. Matter* **12**, 10257 (2000).
- [30] J. Goss, B. Coomer, R. Jones, T. Shaw, P. Briddon, and S. Öberg, Interstitial aggregates in diamond, *Diam. Relat. Mater.* **10**, 434 (2001).
- [31] J. P. Goss, B. J. Coomer, R. Jones, T. D. Shaw, P. R. Briddon, M. Rayson, and S. Öberg, Self-interstitial aggregation in diamond, *Phys. Rev. B* **63**, 195208 (2001).
- [32] F. Colasuonno, F. S. Gentile, W. Mackrodt, A. M. Ferrari, A. Platonenko, and R. Dovesi, Interstitial defects in diamond: A quantum mechanical simulation of their EPR constants and vibrational spectra, *J. Chem. Phys.* **153**, 024119 (2020).
- [33] A. Komarovskikh, DFT investigation of 100-split interstitial paramagnetic centers in diamond, *Diam. Relat. Mater.* **146**, 111235 (2024).
- [34] S. Salustro, Y. Noël, C. M. Zicovich-Wilson, P. Olivero, and R. Dovesi, The V + I defects in diamond: An *ab initio* investigation of the electronic structure, of the Raman and IR spectra, and of their possible recombination, *J. Chem. Phys.* **145**, 184701 (2016).
- [35] A. Mainwood, A. T. Collins, and P. Woad, Isotope dependence of the frequency of localized vibrational modes in diamond, in *Defects in Semiconductors 17*, Materials Science Forum (Trans Tech Publications Ltd., Aedermannsdorf, Switzerland, 1993), Vol. 143, pp. 29–34.
- [36] G. Kresse and J. Hafner, *Ab initio* molecular dynamics for liquid metals, *Phys. Rev. B* **47**, 558 (1993).
- [37] P. E. Blöchl, Projector augmented-wave method, *Phys. Rev. B* **50**, 17953 (1994).
- [38] J. Paier, M. Marsman, K. Hummer, G. Kresse, I. C. Gerber, and J. G. Ángyán, Screened hybrid density functionals applied to solids, *J. Chem. Phys.* **124**, 154709 (2006).
- [39] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, New York, 1976).
- [40] Z. Fan, Z. Zeng, C. Zhang, Y. Wang, K. Song, H. Dong, Y. Chen, and T. Ala-Nissila, Neuroevolution machine learning potentials: Combining high accuracy and low cost in atomistic simulations and application to heat transport, *Phys. Rev. B* **104**, 104309 (2021).
- [41] P. Erhart and K. Albe, Analytical potential for atomistic simulations of silicon, carbon, and silicon carbide, *Phys. Rev. B* **71**, 035211 (2005).
- [42] Z. Fan, Y. Wang, P. Ying, K. Song, J. Wang, Y. Wang, Z. Zeng, K. Xu, E. Lindgren, J. M. Rahm, A. J. Gabourie, J. Liu, H. Dong, J. Wu, Y. Chen, Z. Zhong, J. Sun, P. Erhart, Y. Su, and T. Ala-Nissila, GPUMD: A package for constructing accurate machine-learned potentials and performing highly efficient atomistic simulations, *J. Chem. Phys.* **157**, 114801 (2022).
- [43] Y. Wang, Z. Fan, P. Qian, M. A. Caro, and T. Ala-Nissila, Density dependence of thermal conductivity in nanoporous and amorphous carbon with machine-learned molecular dynamics, *Phys. Rev. B* **111**, 094205 (2025).
- [44] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, Molecular dynamics with coupling to an external bath, *J. Chem. Phys.* **81**, 3684 (1984).
- [45] J. E. Northrup, R. Di Felice, and J. Neugebauer, Energetics of H and NH₂ on GaN(10 $\bar{1}$ 0) and implications for the origin of nanopipe defects, *Phys. Rev. B* **56**, R4325 (1997).
- [46] C. Freysoldt and J. Neugebauer, First-principles calculations for charged defects at surfaces, interfaces, and two-dimensional materials in the presence of electric fields, *Phys. Rev. B* **97**, 205425 (2018).
- [47] R. Nelson, C. Ertural, J. George, V. L. Deringer, G. Hautier, and R. Drnaskowski, LOBSTER: Local orbital projections, atomic charges, and chemical-bonding analysis from projector-augmented-wave-based density-functional theory, *J. Comput. Chem.* **41**, 1931 (2020).
- [48] K. Huang and A. Rhys, Theory of light absorption and non-radiative transitions in F-centres, *Proc. R. Soc. Lond. A* **204**, 406 (1950).
- [49] A. Alkauskas, B. B. Buckley, D. D. Awschalom, and C. G. V. de Walle, First-principles theory of the luminescence lineshape for the triplet transition in diamond NV centres, *New J. Phys.* **16**, 073026 (2014).

- [50] A. Gali, E. Janzén, P. Deák, G. Kresse, and E. Kaxiras, Theory of spin-conserving excitation of the $N - V^-$ center in diamond, *Phys. Rev. Lett.* **103**, 186404 (2009).
- [51] U. von Barth and L. Hedin, A local exchange-correlation potential for the spin polarized case. I, *J. Phys. C* **5**, 1629 (1972).
- [52] D. C. Hunt, D. J. Twitchen, M. E. Newton, J. M. Baker, T. R. Anthony, W. F. Banholzer, and S. S. Vagarali, Identification of the neutral carbon $\langle 100 \rangle$ -split interstitial in diamond, *Phys. Rev. B* **61**, 3863 (2000).
- [53] C. Weigel, D. Peak, J. W. Corbett, G. D. Watkins, and R. P. Messmer, Carbon interstitial in the diamond lattice, *Phys. Rev. B* **8**, 2906 (1973).
- [54] D. J. Twitchen, M. E. Newton, J. M. Baker, O. D. Tucker, T. R. Anthony, and W. F. Banholzer, Electron-paramagnetic-resonance measurements on the di- $\langle 001 \rangle$ -split interstitial center (R1) in diamond, *Phys. Rev. B* **54**, 6988 (1996).
- [55] B. Coomer, J. Goss, R. Jones, S. Öberg, and P. Briddon, Interstitial aggregates and a new model for the I_1/W optical centre in silicon, *Physica B* **273–274**, 505 (1999).
- [56] R. Jones, T. Eberlein, N. Pinho, B. Coomer, J. Goss, P. Briddon, and S. Öberg, Self-interstitial clusters in silicon, *Nucl. Instrum. Methods Phys. Res. Sect. B* **186**, 10 (2002).
- [57] P. J. Fallon, L. M. Brown, J. C. Barry, and J. Bruley, Nitrogen determination and characterization in natural diamond platelets, *Philos. Mag. A* **72**, 21 (1995).
- [58] A. L. Görne and R. Dronskowski, Covalent bonding versus total energy: On the attainability of certain predicted low-energy carbon allotropes, *Carbon* **148**, 151 (2019).
- [59] G. Henkelman, A. Arnaldsson, and H. Jónsson, A fast and robust algorithm for Bader decomposition of charge density, *Comput. Mater. Sci.* **36**, 354 (2006).
- [60] S. Praver, K. Nugent, and D. Jamieson, The Raman spectrum of amorphous diamond, *Diam. Relat. Mater.* **7**, 106 (1998).
- [61] A. A. El-Azhary and H. U. Suter, Comparison between optimized geometries and vibrational frequencies calculated by the DFT methods, *J. Phys. Chem.* **100**, 15056 (1996).
- [62] G. P. Srivastava, *The Physics of Phonons*, 2nd ed. (CRC Press, Boca Raton, FL, 2022).
- [63] T. Shimanouchi *et al.*, *Tables of Molecular Vibrational Frequencies* (U.S. Government Printing Office, Washington, DC, 1973).
- [64] D. Twitchen, M. Newton, J. Baker, W. Banholzer, and T. Anthony, Optical spin polarization in the di- $\langle 001 \rangle$ -split interstitial (R1) centre in diamond, *Diam. Relat. Mater.* **8**, 1101 (1999).
- [65] J. Steeds, T. Davis, S. Charles, J. Hayes, and J. Butler, 3H luminescence in electron-irradiated diamond samples and its relationship to self-interstitials, *Diam. Relat. Mater.* **8**, 1847 (1999).
- [66] J. Gräfenstein, E. Kraka, M. Filatov, and D. Cremer, Can unrestricted density-functional theory describe open shell singlet biradicals? *Int. J. Mol. Sci.* **3**, 360 (2002).
- [67] G. Thiering and A. Gali, Complexes of silicon, vacancy, and hydrogen in diamond: A density functional study, *Phys. Rev. B* **92**, 165203 (2015).
- [68] A. V. Ivanov, Y. L. A. Schmerwitz, G. Levi, and H. Jónsson, Electronic excitations of the charged nitrogen-vacancy center in diamond obtained using time-independent variational density functional calculations, *SciPost Phys.* **15**, 009 (2023).
- [69] N. Ghafari Cherati, I. Abdolhosseini Sarsari, A. Hashemi, and T. Ala-Nissila, Boron isotope effects on Raman scattering in bulk BN, BP, and BAs: A density functional theory study, *Phys. Rev. B* **111**, 205204 (2025).
- [70] I. Vlasov, V. Ralchenko, and E. Goovaerts, Laser-induced transformation of 3H defects in diamond, *Phys. Status Solidi A* **193**, 489 (2002).
- [71] N. Ghafari Cherati, *From mono- to hexa-interstitials: Computational insights into carbon defects in diamond* (ARP, 2026).