Diamond & Related Materials 158 (2025) 112654

Contents lists available at ScienceDirect

Diamond & Related Materials

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

The first-formed hydrogen and nitrogen defects in natural diamond

Maxwell C. Day [a,*], Etienne Balan [b], Razvan Caracas [c,d], Michael C. Jollands [e], Francesca Innocenzi [a], Davide Novella [a], Fabrizio Nestola [a], Martha G. Pamato [a]

[a] Dipartimento di Geoscienze, Università degli Studi di Padova, Via Gradenigo 6, I-35131 Padova, Italy
[b] Sorbonne Université, CNRS, Muséum National d'Histoire Naturelle, Institut de Minéralogie, Physique des Matériaux et Cosmochimie (IMPMC), Paris, France
[c] Institut de Physique du Globe de Paris, Université de Paris Cité, CNRS, Paris, France
[d] The Research Institute of the University of Bucharest (ICUB), University of Bucharest, Bucharest, Romania
[e] Gemological Institute of America (GIA), 50W, 47th St., New York, NY 10036, USA

ARTICLE INFO

Keywords:
Diamond
FTIR spectroscopy
Hydrogen defects
Nitrogen aggregation
Vibrational modes
Ab-initio simulations

ABSTRACT

Nitrogen (N) and hydrogen (H) defects in diamond provide crucial information about diamond growth. The electromagnetic properties that N/H-defects impart on synthetic diamond are exploited in novel technological applications including quantum information storage and sensing. Access to the *Gemological Institute of America* database of infrared spectra recorded from rare *Type Ib* + *IaA* natural diamonds has allowed identification of the first-formed N/H-defects, NH^0 and N_2H^0, assigned to IR peaks at 3394 and 3343 cm^{-1}, respectively. Density functional perturbation theory and ab initio molecular dynamic simulations of NH^0 and N_2H^0, including the first machine-learning molecular dynamics simulation performed for N/H-defects in diamond, show excellent agreement between observed and calculated C—H stretching frequencies (18 to 66 cm^{-1} difference). New aggregation sequences are proposed for VN_3H^0, the commonly observed *end-state* of N/H-aggregation, that differ from the currently accepted sequence. This suggests that NH^0 and N_2H^0 form during, or shortly after, diamond formation inhibiting diffusive loss of H during mantle residence and thus the H (isotopic) composition of diamond may reflect parental fluids from primordial and/or recycled mantle H_2O reservoirs. These findings also have important technological implications for N/V-defects, such as NV^-, with useful electromagnetic properties that may be passivated by H.

1. Introduction

Nitrogen (N) and hydrogen (H) are the most abundant impurities in natural diamonds forming point defects involving substitutional N, carbon (C) vacancies (*V*) and H atoms (or ions) that passivate dangling bonds on carbon (C) [1,2]. Several different *V*/N/H-related defects have been proposed and assigned to signals (peaks) observed in spectroscopic experiments [3,4] based on quantum chemical ab initio simulations [5–7]. However, most observed peaks have not been assigned to specific defects, especially those attributed to H observed in the C—H stretching region (3100–3500 cm^{-1}) of the Fourier Transform Infrared (FTIR) spectra of natural diamonds [8]. Nevertheless, *V*/N/H-defects have received considerable attention due to their useful electromagnetic properties exploited for semiconductor [9] and quantum computation and sensing applications [10,11]. From a geologic perspective, *V*/N/H-defects provide important information about the composition and sources of fluids/melts associated with diamond formation and growth [12,13]. The mechanisms and rates of N-defect aggregation and migration (e.g. vacancy-assisted migration [14]) are well understood. In *Type Ib* diamond, N occurs predominately as single-substitutional N_S^0 (***C***-centers) and with continued annealing, ***C***-centers aggregate to form N_2^0 (***A***-centers) in *Type IaA* diamonds and eventually VN_4 (***B***-centers) in *Type IaB* diamonds [1]. Deconvoluting the composite signal of these defects in the mid-infrared (MIR) allows calculation of the *N-aggregation state* (i.e., %IaA or %IaB) [15]. As the time/temperature dependence of the C → A → B aggregation sequence is known, %IaA (or %IaB) and N contents can be used to derive the time/temperature of diamond residence in the mantle post-formation and pre-eruption [16–18]. The C → A → B aggregation sequence in reality is more complex, and involves several intermediate N-defects (e.g., ***Y***-centers [19] and ***X***-centers (N_S^+) [20]) and interstitials [21,22] (e.g., N_i and C_i).

Unfortunately, the behaviour of H-related defects in natural diamond is much less understood. The consensus is that N/*V*-defects act as sinks for interstitial H (H_i) where C → N(*V*) substitution produces dangling C-

* Corresponding author.
E-mail address: maxwell.day@unipd.it (M.C. Day).

https://doi.org/10.1016/j.diamond.2025.112654
Received 30 June 2025; Received in revised form 17 July 2025; Accepted 21 July 2025
Available online 22 July 2025
0925-9635/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

bonds that act to trap H_i [2,5,23]. Only a few H-related IR peaks have been assigned to specific defects [8] including the 3107 cm^{-1} and 3236 cm^{-1} peaks assigned to the VN_3H^0 and VN_4H^0 defects, respectively [5,24]. These defects represent an *end-state* of N/H aggregation [8] and are thus ubiquitous in *Type Ia* diamonds which represent the latter- or end-stages of mantle residence (annealing). As >95 % of natural diamonds are *Type Ia* [25], the first-formed *V*/N/H-defects are often not preserved and information about the initial stages of *V*/N/H-defect formation and aggregation is lost. Understanding the complete behaviour of H in diamond requires analysis of H-bearing *Type Ib* diamonds that have experienced minimal annealing and N-aggregation. This is a difficult task given the rarity of H-bearing *Type Ib* samples, which account for <0.1 % of natural diamonds [25].

In this study, 178 FTIR spectra of natural *Type Ib + IaA* diamonds (covering almost the entire range of %IaA) were selected from the *Gemological Institute of America* (GIA) database. In addition, FTIR and photoluminescence (PL) spectra were recorded from H-bearing and H-free zones in one *Type Ib* diamond from the Zimmi alluvial locality (West Africa) (Fig. 1a, b). Of particular interest are three peaks located at 3343, 3373, and 3394 cm^{-1} that are the only C—H stretching peaks observed in relatively pure *Type Ib* (low %IaA) diamond and thus are attributed to the first H-defects that form in natural diamond. The identity of the H-related defects associated with the 3343 and 3394 cm^{-1} peaks is proposed here based on density functional perturbation theory (DFPT) and ab-initio and machine-learning molecular dynamics (AIMD and MLMD) simulations, observed C—H bending, stretching, and combination modes, and peak intensity correlations. Our findings show that H_i in N-bearing diamonds is trapped much earlier than previously thought mitigating diffusive loss of H and thus H-isotopic re-equilibration in diamond during mantle residence.

2. Materials and methods

2.1. Fourier transform infrared (FTIR) spectroscopy

A set of 178 FTIR spectra recorded from *Type Ib + IaA* diamonds were selected from the GIA database. Only spectra recorded from natural diamonds (as opposed to CVD or HPHT synthetic diamond) were included to ensure geological relevance. FTIR spectra were recorded at the GIA using a Thermo Fisher Nicolet iS50 unit equipped with a KBr beam splitter and a LN-cooled MCT detector. The spectra were recorded at an operating resolution of 1 cm^{-1} with 32 scans, over the range 6000–650 cm^{-1}. Ten IR spectra were recorded from sample Z-16 (six from the colorless zone and four from the yellow zone, Fig. 1a) from the Zimmi alluvial locality (West Africa). These spectra were recorded at the University of Padova (Italy) using a Thermo Fisher Nicolet iN10 InfraRed microscope equipped with a KBr beam splitter and a LN-cooled MCT detector, spectra were recorded at an operating resolution of 2 cm^{-1} with 64 scans over the range 6000–750 cm^{-1}.

The IR spectra from the GIA were previously processed by Day et al. [26] who showed that poor quality fits of the N-region of *Type Ib + IaA* diamonds are due to a previously unaccounted for defect, the ***Y***-center [19]. ***Y***-centers are common in natural *Type Ib + IaA* diamonds and Day et al. [26] developed a new deconvolution routine (i.e., *Caxbd_Inherit_2024-Ib* Excel spreadsheet) in which the spectrum of ***Y***-centers is including during least-squares fitting of the N-region. This routine was used here to determine the N-defect contents, N_{tot} and %IaA (%IaA = $N_A/(N_A + N_C + N_X + N_Y) \times 100$) of all samples. The normalized linear intensity (absorption coefficient, cm^{-1}) of peaks between 1353 and 1387 cm^{-1} and H-related peaks at 3107, 3343, 3373, and 3394 cm^{-1} was determined following the procedure for the 3107 cm^{-1} peak in the DiaMap Excel spreadsheet [15]. Exact peak centers were determined by fitting peaks with Voigt functions using the Fityk V0.9.8 spectra software [27]. To verify processing consistency, normalized H-related peak intensities were also determined manually for each spectrum using the OMNIC and Fityk V0.9.8 spectra software [27].

The statistical significance of peak intensity correlations (see Section 3) is determined by calculating the Pearson correlation coefficient (*r*) and conducting standard two-tailed *t*-tests at $\alpha = 0.05$ (i.e., 97.5 % confidence). The R^2 values for these correlations are also reported.

2.2. Photoluminescence (PL) spectroscopy

Photoluminescence spectra were recorded from sample Z-16 at the GIA under liquid N_2, using a Renishaw inVia Raman microscope equipped with 457 nm (Ar ion), 633 nm (Ne ion) and 830 nm (diode) lasers, using a 5× objective. Spectra were recorded from both the yellow and colorless zones of the polished plate (Fig. 1a). As spectra were recorded from a single polished plate, peak intensities from spectra collected using the same excitation wavelength were considered comparable and the spectra were not Raman normalized. The raw data for all PL spectra are provided in Appendix E.

2.3. Static ab initio simulations of N/H-related point defects

The DFPT modeling was performed within the density functional

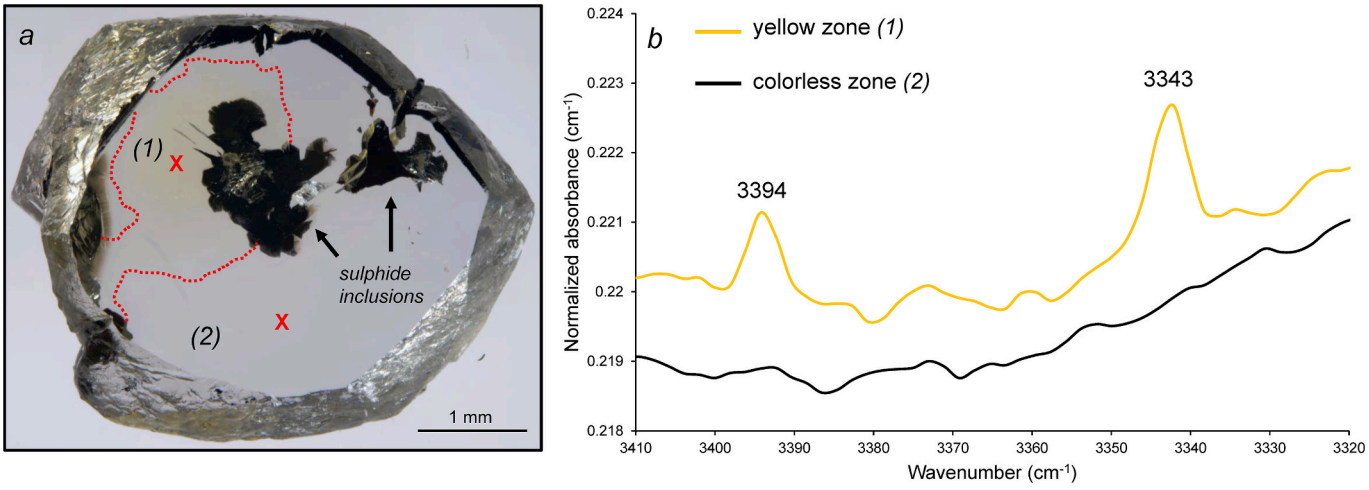


Fig. 1. Diamond sample Z-16. (a) Transmitted light image of sample Z-16 showing the colorless zone and the circular, H-bearing yellow zone. (b) FTIR spectra in the C—H stretching region recorded from the colorless zone and the yellow zone which shows H-related peaks at 3343 and 3394 cm^{-1}. Red X mark the points in the colorless and yellow zones where PL spectra were recorded.

theory (DFT) framework using the generalized gradient approximation (GGA) to the exchange-correlation functional as proposed by Perdew et al. [28]. The modeling scheme used periodic boundary conditions, and a plane-wave basis set as implemented in the PWscf code of the Quantum Espresso package [29]. Ionic cores were described using optimized norm-conserving Vanderbilt (ONCV) pseudopotentials [30,31]. The wavefunction and electron density were expanded in plane waves with cutoffs of 120 Ry and 720 Ry, respectively.

Structural properties of defects were determined using 2 × 2 × 2 diamond supercells. The Brillouin zone sampling for electronic integration was performed using a 2 × 2 × 2 Monkhorst-Pack mesh of k-points and the relaxation of atomic internal coordinates was performed until the residual forces were $<10^{-4}$ Ry per atomic unit. The unit-cell parameter of pure diamond was optimized at zero pressure (a = 3.5675 Å) and used without further relaxation to produce the defect-bearing supercells. Spin-polarized calculations have been used to account for the paramagnetic character (S = 1/2) of the N_2H^0, NH^+ and NH^- defects. The vibrational properties at the Brillouin zone center (Γ point) of the relaxed structures were calculated using the density-functional perturbation theory [32], as implemented in the PHonon code [29].

2.4. Molecular dynamics simulations of N/H-related point defects

Ab initio molecular dynamics (AIMD) simulations were performed using the *Vienna* Ab initio *Software Package* (VASP) [33,34] with the density functional theory (DFT) and projector-augmented wave (PAW) method [35]. The exchange correlation energy was computed within the Perdew-Burke-Ernzerhof formulation of the generalized gradient approximation [36]. A cut-off energy of 550 eV was used for the plane-wave-basis set and 880 eV for augmentation charges. A 2 × 2 × 2 high-symmetry grid of k-points [37] was used to calculate the electronic density. Simulations were conducted in the *NVT* ensembles with fixed composition (*N*), volume (*V*), and temperature (*T*). The temperature was controlled using the Nosé-Hoover thermostat [38,39]. A timestep of 1 femtosecond was used and the simulations lasted for 35 ps after thermalization. The simulations were performed in 2 × 2 × 2 (64C atom) supercells, in which one to three C atoms were replaced to produce the defect.

Further modeling of the NH^0 defect was performed with more realistic defect concentrations using MLMD simulations. Several AIMD simulations were performed with different NH^0 defect concentrations, an on-the-fly ML procedure built in VASP was used to construct the ML interatomic potentials [40]. Potentials were first built and tested in the 2 × 2 × 2 supercell, then refined in 2 × 2 × 4 and then in 2 × 4 × 4 supercells. The refinement was done iteratively until the MLMD simulations reproduced the AIMD forces and energies. The MLMD simulations were then performed using 8 × 8 × 8 (4096C atom) supercells. One C atom was replaced with the NH^0 defect. The MLMD simulation lasted for up to 200 ps. For AIMD and MLMD simulations, the vibrational spectra were computed from the self-correlation of the atomic velocities using the UDM package [41]. The resulting H vibrational spectra are shown in Fig. D.1a–D.1c (Appendix D).

3. Results and discussion

3.1. Spectroscopic analysis of N- and H-related defects

Selected data (N-defect contents, N_{tot}, %IaA, normalized peak intensities) for all IR spectra from the GIA database and from the Zimmi diamond (sample Z-16) are provided in Tables A.1 and A.2 (Appendix A), respectively. Sample Z-16 has a yellow zone surrounded by a colorless zone (Fig. 1a) as described for other *Type Ib* diamonds from the same locality [42]. IR spectra recorded from the yellow zone show a higher average N content (35.4 at.ppm) and a lower average N-aggregation state (22.9 %IaA) compared to the colorless zone (12.4 at.ppm N and 25.2 %IaA) (Table 1). IR peaks at 3343 and 3394 cm^{-1}, in addition to a relatively weak 3107 cm^{-1} peak, are only observed in the yellow zone (Fig. 1b).

Analyses of all IR spectra revealed >20 peaks in the C—H stretching region (Fig. 2) that have either not been observed or are simply noted in previous studies [43–45]. In addition to signal due to the ***C***-, ***X***-, ***Y***-, and ***A***-center absorption systems, weak peaks at 1353, 1358, 1363, 1374, and 1387 cm^{-1}, attributed to C_i/N_i-related (self-)interstitial complexes [26], are observed. A large number of peaks in the C—H stretching region, most relatively weak, were observed and are listed in Table 2. In addition to the peak at 3107 cm^{-1}, relatively intense peaks at 3343, 3373, and 3394 cm^{-1} were observed in most spectra, these peaks (Fig. 2a) have not been described in detail until now. When relatively intense, the 3343 and 3394 cm^{-1} peaks show shoulders at 3335 and 3388 cm^{-1} (Fig. 2b). They are accompanied by two broad peaks at 4764 and 4797 cm^{-1} (Fig. 2c) and a relatively weak peak, at 3373 cm^{-1}. Apart from peaks at 1350–1390 cm^{-1}, the lower frequency C—H bending region contains peaks at 1405 (associated with 3107 cm^{-1}), 1408, 1415, 1419, 1424, and1428 cm^{-1} (Figs. 2d and B.1., Appendix B).

At ~≤30 % IaA, the 3343, 3373, and 3394 cm^{-1} peaks are relatively intense and often observed alone or alongside only a few other relatively weak peaks (e.g., 3107 cm^{-1}) (Fig. 3a). With increasing %IaA, more peaks (e.g., 3144, 3182–85, 3236–37, and 3310–12 cm^{-1}) are observed in the C—H stretching region and the relative intensity of the 3107 cm^{-1} peak increases (Fig. 3b). Continued N-aggregation is accompanied by a significant decrease in the intensity of the 3343, 3373, and 3394 cm^{-1} peaks and a further increase in the intensity of the 3107 cm^{-1} peak (Fig. 3c). The 3144 cm^{-1} peak increases with 3107 cm^{-1} and with increasing %IaA (see Fig. B.2, Appendix B). Some spectra contain an extremely large number of peaks (>40) in the C—H stretching region (e.g., Fig. 3d), such samples typically show intermediate N-aggregation states (~20–40 % IaA) and likely represent *transitional* aggregation states (see Section 3.3).

As shown in Fig. 4a, the normalized intensities of the 3343 and 3394 cm^{-1} peaks are positively correlated. In Fig. 4b, intensities are re-normalized to N_{tot} and the positive correlation becomes stronger ($r(178) = 0.98$ compared to 0.95) suggesting that the corresponding defects contain N. This is further evidenced by the apparent effect of %IaA on the intensity of the 3343 and 3394 cm^{-1} peaks as described above. The intensity of the 3343 and 3394 cm^{-1} peaks is approximately inversely proportional to the 3107 cm^{-1} peak intensity (Fig. 4c and d) suggesting the 3343 and 3394 cm^{-1} defects are involved in aggregation

Table 1
Selected data from infrared spectra recorded from sample Z-16.

	Colorless zone (6 points)	Yellow zone (4 points)
N-aggregation state (%IaA)	25.2	22.9
N_C (at.ppm)	2.1	9.4
N_A (at.ppm)	3.2	8.1
N_X (at.ppm)	0.1	0.7
N_Y (at.ppm)	7.0	17.3
N_{tot} (at.ppm)	12.4	35.4
3107 (cm^{-1})	0	0.0013
3343 (cm^{-1})	0	0.0022
3373 (cm^{-1})	0	0.0004
3394 (cm^{-1})	0	0.0013
1353 (cm^{-1})	0.0197	0.0233
1358 (cm^{-1})	0.0359	0.0707
1363 (cm^{-1})	0	0
1374 (cm^{-1})	0	0.0150
1387 (cm^{-1})	0	0.0195

FTIR data was calculated by averaging the deconvolution results of six spectra recorded from the colorless zone and four spectra recorded from the yellow zone. FTIR data processed from all 10 points is given in Table A.2 (Appendix A). The ***C***-, ***A***-, ***X***-, and ***Y***-center content was calculated using the excel spreadsheet *Caxbd_Inherit_2024_Ib* from [26]. The absorption coefficients (cm^{-1}) for H-related peaks were calculated using the automated baseline and normalization routine from the DiaMap Excel spreadsheet [15].

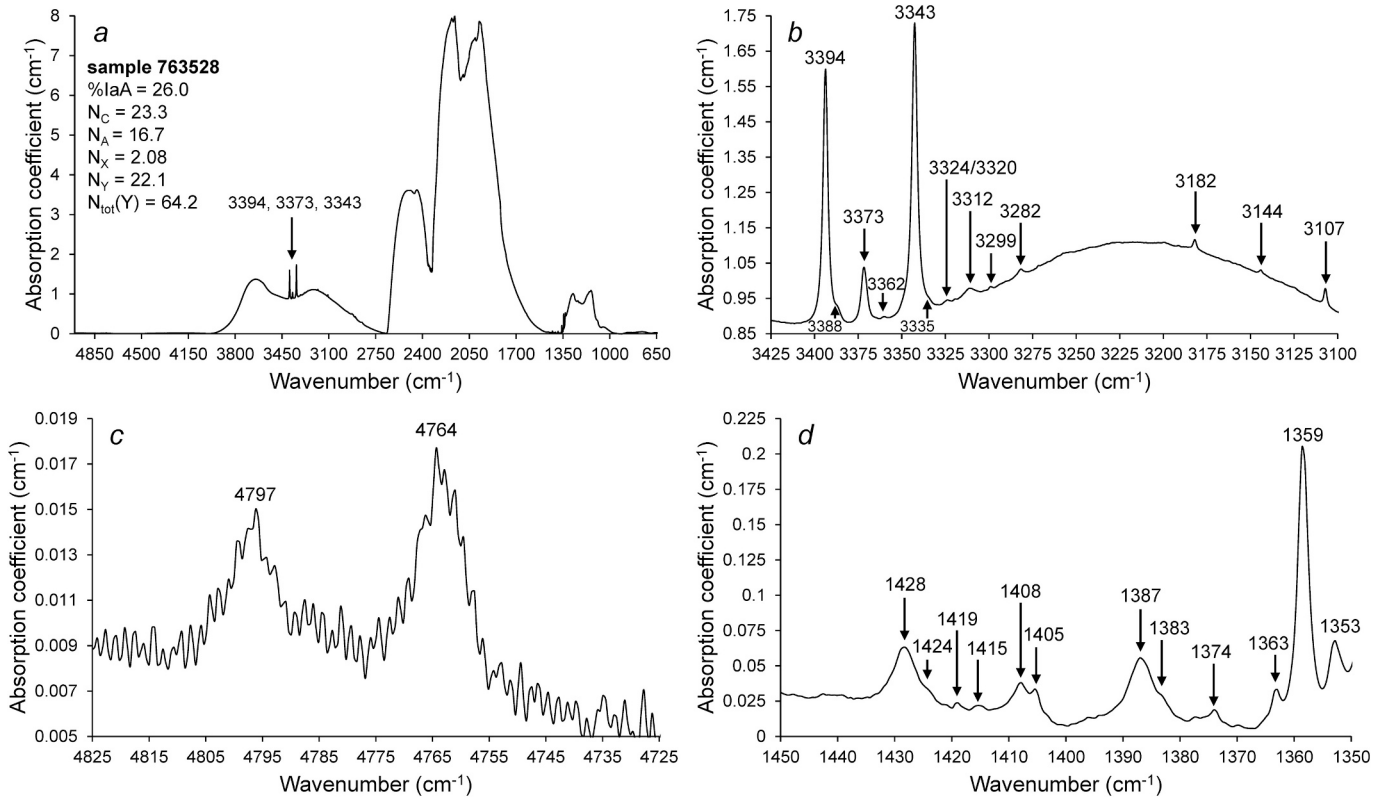


Fig. 2. FTIR spectrum of a *Type Ib + IaA* diamond (sample 763528). (a) FTIR spectrum in the 5000–650 cm^{-1} range showing intense H-related peaks at 3343, 3373, and 3394 cm^{-1}. (b) FTIR spectrum collected in the 3425–3100 cm^{-1} range showing peaks at 3343, 3373, and 3394 cm^{-1} in additional to several other relatively weak H-related peaks. (c) FTIR spectrum in the 4825–4725 cm^{-1} range showing two weak peaks at 4764 and 4797 cm^{-1}. (d) FTIR spectrum in the 1450–1350 cm^{-1} range. N-defect contents and the N-aggregation state (%IaA) of sample 763528 were calculated using the deconvolution routine of [26] and are indicated in (a).

Table 2
Positions (cm^{-1}) of observed peaks in the C—H stretching region of *Type Ib + IaA* diamonds.

3050	3082	**3093**	**3098**	**3107***s*	3117	**3124**	3137	**3144***s*	3149
3151	**3154-56**	3163	3167	3170-73	**3181-85***s*	**3190-91**	3196	3200	3213
3218	3224	3228-29*b*	3234	**3236-37**	3240	3245	3250	3254-57	3262-63*b*
3272	3276	**3282-87***s*	3290*s*	3300	3305	**3310-13***s/b*	3320-24	3328	3331
3343*s*	3359	3362	**3371-76***s/b*	**3394***s*	3401	3407	3410		

Commonly observed peaks in *Type Ib + IaA* spectra are shown in bold, those that are observed in transitional spectra (10 spectra, 20–40 %IaA, see Section 3.3) are shown in normal font.
Peak centers were determined by fitting peaks with Voigt functions, relatively broad (non-Voigt) peaks are indicated with *b*, peaks that contain one or more shoulders are indicated with *s*, the position of shoulders is not indicated.

sequences that produce VN_3H^0 and thus must also contain H in addition to N. Similar trends are observed for the 3373 cm^{-1} peak (see Table A.1, Appendix A) and thus it may also correspond to an N/H-related defect. However, this peak shows variable degrees of broadening and positions ranging from 3371 to 3376 cm^{-1}. Therefore, the origin of the 3373 cm^{-1} peak is likely more complicated and may be related to an extended defect [1] or a combination (or overtone) of other N/H-related point-defects. For these reasons, a defect identity is not proposed for the 3373 cm^{-1} peak.

3.2. *Ab initio simulations and identification of H-related point defects: NH^0 and N_2H^0*

Based on the descriptions given above, it is reasonable to assume that the 3343 and 3394 cm^{-1} peaks correspond to defects that contain N and H. Although no correlation is observed between N_{tot} and the 3343 and 3394 cm^{-1} peak intensities, this is expected in diamonds where H, as opposed to N, is likely the limiting reagent in the formation of N/H-defects. Deviation from the 1:1 line in Fig. 4a and b indicates that the 3343 and 3394 cm^{-1} peaks correspond to different defects. In *Type Ib + IaA* diamonds, the highest concentration of substitutional N-defects capable of trapping H_i are ***C***- (N^0) and ***A***-centers (N_2^0), and thus the NH^0 and N_2H^0 defects are logical candidates for the 3343 and 3394 cm^{-1} peaks.

To test this hypothesis, static DFPT and ab initio molecular dynamics (AIMD) simulations of the vibrational properties of NH and N_2H defects were conducted and the results are given in Table 3. Several NH and N_2H configurations are possible for which different structures and charge-states are assumed. The validity of each simulation method was verified by simulating the fundamental C—H stretching ($\boldsymbol{\nu_S}$ in cm^{-1}) and bending ($\boldsymbol{\nu_B}$ in cm^{-1}) frequency of the VN_3H^0 defect that corresponds to known peaks at 3107 and 1405 cm^{-1}, respectively [5]. Here, both DFPT and AIMD $\boldsymbol{\nu_S}$ and $\boldsymbol{\nu_B}$ frequencies provide consistent results and are in excellent agreement with observation (1.22 and 0.48 % difference for the DFPT and AIMD stretching frequencies, respectively). A large-scale molecular dynamics machine-learning (MLMD) simulation with machine-learning interatomic potentials of the NH^0 defect in configuration **(1)** in an 8 × 8 × 8 (4096C atom) supercell produce $\boldsymbol{\nu_S}$ (3376

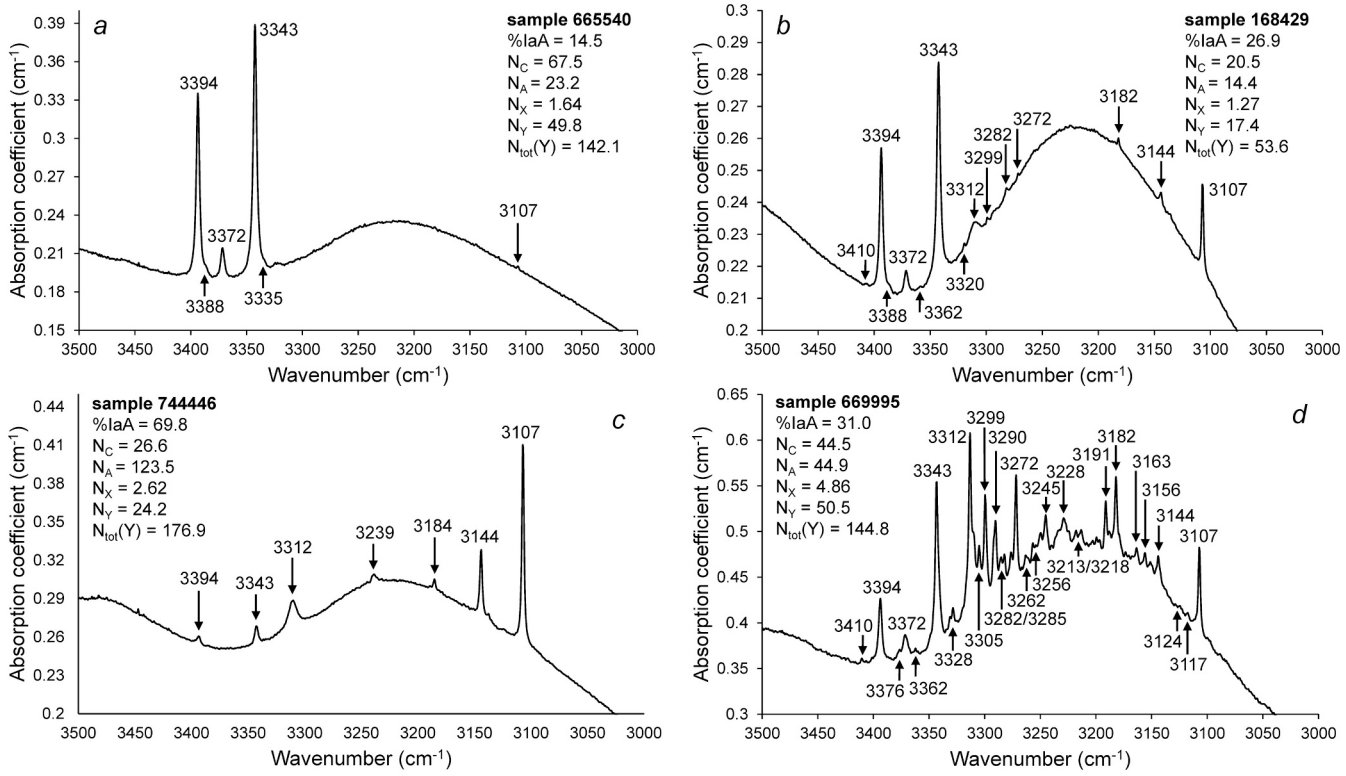


Fig. 3. Trends in the occurrence and intensity of H-related peaks with %IaA. (a) FTIR spectrum of sample 665540 (%IaA = 14.5) showing intense 3343, 3373 and 3394 cm^{-1} peaks and a very weak 3107 cm^{-1} peak. (b) FTIR spectrum of sample 168429 (%IaA = 26.9) showing relatively less intense 3343, 3373 and 3394 cm^{-1} peaks and a more intense 3107 cm^{-1} peak in addition to several other weak H-related peaks. (c) FTIR spectrum of sample 744446 (%IaA = 69.8) showing very weak 3343 and 3394 cm^{-1} peaks and a relatively intense 3107 cm^{-1} peak. (d) FTIR spectrum of sample 669995 (%IaA = 31.0) showing an extremely large number of different peaks including those at 3343, 3373, 3394 and 3107 cm^{-1}. N-defect contents and N-aggregation states (%IaA) of each sample were calculated using the deconvolution routine of [26] and are indicated in the respective figures.

cm^{-1}) close to the AIMD ν_S (3349 cm^{-1}), indicating that the DFPT and AIMD simulations of smaller supercells still provide relevant results to interpret the observed diamond spectrum with significantly lower defect contents. The ν_S computed for VN_3H^0 and NH^0 in previous studies [7] (Table C.1, Appendix C) are also consistent with the ν_S values reported here. The vibrational properties of N_2H^0 have not been previously reported in the literature.

In Fig. 5, the calculated ν_S for NH, N_2H and VN_3H^0 are plotted against the corresponding (observed) peak positions. Overall, ν_S calculated for the neutral NH^0 defect in configuration **(1)** and for the neutral N_2H^0 defect in configuration **(3)** are in best agreement with the observed 3394 and 3343 cm^{-1} peaks, respectively (differences range from 0.5 to 2.3 %, Table 3). The calculated ν_B for NH^0 and N_2H^0 defects range from 1401 to 1463 cm^{-1}, consistent with the observation of several peaks between 1405 and 1428 cm^{-1} (Figs. 2d and B.1, Appendix B). Although previous simulations of $VN_2H_2^0$ (Table C.1, Appendix C) also produce ν_S close to 3394 and 3343 cm^{-1}, two C—H stretching peaks with proportional intensity would be expected. The calculated ν_S for VH and VH_2 in previous studies (Table C.1, Appendix C) are also close to 3394 and 3343 cm^{-1} but they do not contain N and thus are unlikely candidates for these two peaks. The VH_3, VNH_3 and VH_4 defects were not considered as the calculated ν_S for such defects is typically >3500 cm^{-1} [8]. Moreover, it is unlikely that isolated vacancies (as opposed to N) serve as the dominant means by which H_i is trapped (and thus observed in experiment) as isolated vacancies are annealed out of diamond at relatively low temperatures (800 °C) [46].

Previous assignments of H-related defects to specific peaks are based largely on their occurrence in the C—H stretching region [43,44] and, with the exception of 3107 cm^{-1} (VN_3H^0), no rigorous peak-defect assignments have been made for *Type Ib + IaA* diamonds until now. Woods and Collins [44] were the first to observe peaks at 3343, 3373, and 3394 cm^{-1} in *Type Ib + IaA* diamonds although they were accompanied by other relatively intense H-related peaks such as 3107 cm^{-1}. Gentile et al. [7] tentatively assigned the 3394 cm^{-1} peak to the NH defect based on agreement with data from ab initio simulations but no attempts to assign a specific defect to the 3343 cm^{-1} peak has been made until now.

Based on the agreement with DFPT, AIMD and MLMD simulations and peak intensity correlations, we assign the 3394 and 3343 cm^{-1} peaks to ν_S of the NH^0 and N_2H^0 defects in configurations **(1)** and **(3)**, respectively (Table 3). This peak-defect assignment is supported by observation of shoulders at 3335 and 3388 cm^{-1} (Fig. 2b) which are consistent with substitution of ^{12}C by ^{13}C which is known to have a C—H peak-shifting effect of ~9 cm^{-1} [5] (e.g., 3343–3335 = 8 cm^{-1}), as is observed at ~3098 cm^{-1} for the 3107 cm^{-1} peak [44]. The NH^0 defect in configuration **(1)** is shown in Fig. 6a, this defect has point symmetry C_{3v} where the C—H bond points towards the N atom (Fig. 6b) (no symmetry constraint was applied during modeling, leading to a very slight departure of the model from C_{3v} symmetry). The N_2H^0 defect in configuration **(3)** is shown in Fig. 6c, this defect has point symmetry C_1 where one N atom bonds directly to the H-bonded C atom and the C—H bond points towards the other N atom (Fig. 6d). Vacancies are not involved in either defect.

For NH^0 and N_2H^0, one (ν_B) and two C—H bending modes (ν_{B1} and ν_{B2}) are expected, respectively (Table 3). A rigorous assignment of the ν_B of the NH^0 and N_2H^0 defects to observed peaks between 1405 and 1428 cm^{-1} (Figs. 2d and B.1, Appendix B) is difficult due to (*i*) overlap at 1405 cm^{-1} due to the ν_B of the VN_3H^0 defect [5], (*ii*) peak overlap due to (self-)interstitial complexes [26] and (*iii*) noise due to atmospheric H_2O. However, two observed peaks at 4797 and 4764 cm^{-1} (Fig. 2c) likely correspond to $\nu_S + \nu_B$ combination bands, akin to the observed 4496

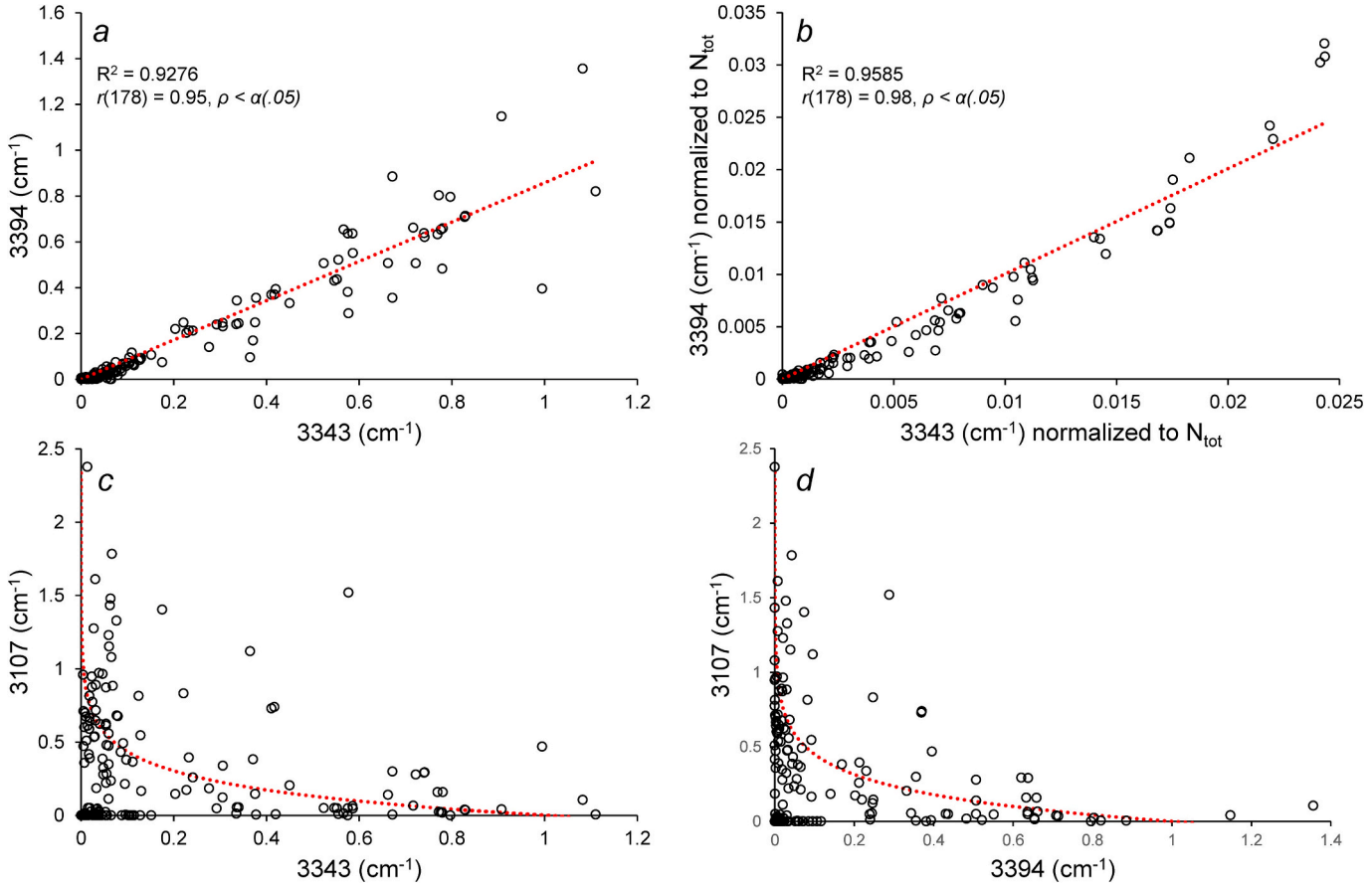


Fig. 4. Intensity correlations between the 3343, 3394 and 3107 peaks. (a) The 3394 cm^{-1} peak intensity plotted as a function of the 3343 cm^{-1} peak intensity showing a positive linear correlation. (b) The 3394 cm^{-1} peak intensity plotted as a function of the 3343 cm^{-1} peak intensity both normalized to the total N content (N_{tot}) showing a stronger linear correlation compared to (a). (c) The 3107 cm^{-1} peak intensity plotted as a function of the 3343 cm^{-1} peak intensity showing a negative exponential correlation. (d) The 3107 cm^{-1} peak intensity plotted as a function of the 3394 cm^{-1} peak intensity showing a negative exponential correlation. The statistical significance of the correlation in (a) and (b) is determined by calculating the Pearson correlation coefficient (r) and conducting standard two-tailed *t*-tests at $\alpha = 0.05$ (i.e., 97.5 % confidence). The R^2 values are also reported in (a) and (b).

cm^{-1} peak assigned to a (3107 + 1405 cm^{-1}) combination band of the VN_3H^0 defect [8]. Assuming that each peak is related to different defects, the best agreement with observed peaks in the C—H bending region leads to the assignment of the 4797 and 4764 cm^{-1} bands to NH^0 and N_2H^0, respectively. Here, expected C—H bending modes at ~1403 cm^{-1} (4797–3394 = 1403 cm^{-1}) for NH^0 and ~ 1421 cm^{-1} (4764–3343 = 1421 cm^{-1}) for N_2H^0 may correspond to observed peaks at 1408 (or 1415) and 1428 cm^{-1}, respectively (Fig. 2d). Two peaks at 2895 and 2904 cm^{-1} (Fig. B3, Appendix B) are observed when the 3343 and 3394 cm^{-1} peaks are relatively intense. These peaks may correspond to combination or overtone bands of the NH^0 and N_2H^0 defects, akin to the 2786 cm^{-1} peak assigned to a (2 × 1405 cm^{-1}) overtone band of the VN_3H^0 defect [8].

3.3. Formation and aggregation of the first-formed H-related defects

The 3343 and 3394 cm^{-1} peaks are often the only C—H stretching peaks observed in poorly aggregated diamonds (≤30 %IaA). Therefore, NH^0 and N_2H^0 are likely the first H-bearing point defects that form in *Type I* (N-bearing) natural diamond. Inspection of Fig. 7 shows that annealing to ~15 %IaA is required to generate significant NH^0 and N_2H^0 contents which decrease at ~35–40 %IaA. During the initial mantle residence (annealing) period (up to ~40 %IaA), NH^0 and N_2H^0 may form as follows;

[1] $N_S^0 + H_i = NH^0$

[2] $NH^0 + N_S^0 = N_2H^0$

[3] $N_2^0 + H_i = N_2H^0$

If one assumes [2] is the dominant sequence by which N_2H^0 forms, one should expect an inverse correlation between the 3343 and 3394 cm^{-1} peak intensities which is not observed (Fig. 7). Thus, N_2H^0 likely forms according to sequence [3], which acts to decrease ***A***-center (N_2^0) content and thus %IaA. This is evidenced in sample Z-16, where the yellow zone that contains NH^0 and N_2H^0, has a slightly lower %IaA than the colorless zone (Table 1). At ~30 %IaA, the average VN_3H^0 content increases coincident with a decrease in NH^0 and N_2H^0 content suggesting they are consumed with continued aggregation to produce VN_3H^0. This observation is in contrast with the commonly proposed formation mechanism for VN_3H^0 where an N3-center (VN_3^0) traps H_i [5]. The IR spectrum of the yellow zone of sample Z-16 shows a weak peak at 3107 cm^{-1} but no peaks in the PL spectrum (Table A.3, Appendix A) attributed to VN_3^0 [3] suggesting VN_3H^0 formed via alternative sequences.

Our results suggest the majority of H_i is trapped by N_S^0 and N_2^0 (sequences [1,3] above) much earlier during C → A aggregation that previously thought. Assuming *vacancy-assisted* aggregation, the following sequences for VN_3H involving NH^0 and N_2H^0 are possible;

[4] $NH^0 + NV^0 = VN_2H^0$

$VN_2H^0 + N_S^0 = VN_3H^0$

Table 3
Calculated C—H vibrational frequencies (cm^{-1}) and C—H bond lengths from DFPT, AIMD and MLMD simulations of different *V*/N/H-related defects in diamond.

Defect	Configuration	Sim. type	ν_S (cm^{-1})	ν_B (cm^{-1})	ovt. bands (cm^{-1})	comb. bands (cm^{-1})	$\nu_{S(obs)}/\nu_{S(calc)}$ (%)	<*d*C-H> (Å)
NH^0	**(1)** N ← H-C	DFPT	**3326**	1411*	–	4737 ($\nu_S + \nu_B$)	2.02	1.040
NH^0	**(2)** N → H-C	DFPT	3511	1871*	–	5382 ($\nu_S + \nu_B$)	3.36	1.030
NH^0	**(1)** N ← H-C	AIMD	**3349**	1401	2795 ($2x\nu_B$)	4693 ($\nu_S + \nu_B$)	1.34	
					4166 ($3x\nu_B$)	1948		
					6674 ($2x\nu_S$)	2535		
						2715		
NH^0	**(1)** N ← H-C	MLMD	**3376**	1451	2895 ($2x\nu_B$)	4833 ($\nu_S + \nu_B$)	0.53	
						4667		
NH^+	**(1)** N ← H-C	DFPT	2601	1337	–	3938 ($\nu_S + \nu_B$)	26.45	1.178
NH^-	**(1)** N ← H-C	DFPT	2995	1472	–	4467 ($\nu_S + \nu_B$)	12.49	1.061
N_2H^0	**(3)** trig. N bonded to CH	DFPT	**3277**	1408 (ν_{B1})	–	4685 ($\nu_S + \nu_{B1}$)	1.99	1.044
				1463 (ν_{B2})		4740 ($\nu_S + \nu_{B2}$)		
N_2H^0	**(3)** trig. N bonded to CH	AIMD.	**3268**	1394 (ν_{B1})	2792 ($2x\nu_{B1}$)	4667 ($\nu_S + \nu_{B1}$)	2.27	
				1451 (ν_{B2})	2902 ($2x\nu_{B2}$)	4720 ($\nu_S + \nu_{B2}$)		
					4113 ($3x\nu_{B1}$)	3786		
					4446 ($3x\nu_{B2}$)	4013		
					6541 ($2x\nu_S$)	6174		
N_2H^+	**(4)** tetra. N bonded to CH	DFPT	3149	1429 (ν_{B1})	–	4578 ($\nu_S + \nu_{B1}$)	5.98	1.053
				1483 (ν_{B2})		4642 ($\nu_S + \nu_{B2}$)		
N_2H^+	**(5)** CH pointing to one N	DFPT	3484	1385 (ν_{B1})	–	4869 ($\nu_S + \nu_{B1}$)	4.13	1.031
				1397 (ν_{B2})		4881 ($\nu_S + \nu_{B2}$)		
N_2H^+	**(6)** H between both N	DFPT	2339	1611*	–	3950 ($\nu_S + \nu_B$)	35.37	
VN_3H^0	**(7)** trigonal	DFPT	**3145**	1376*	–	4521 ($\nu_S + \nu_B$)	1.22	1.076
VN_3H^0	**(7)** trigonal	AIMD	**3092**	1371	2739 ($2x\nu_B$)	4466 ($\nu_S + \nu_B$)	0.48	
					4113 ($3x\nu_B$)	3572		
					6201 ($2x\nu_S$)	3919		
$VN_2H_2^0$	**(8)**	DFPT	3350	1287	–	–	1.30 ($\boldsymbol{\nu_{S\text{-}3394}}$)	1.053
			3524	1294			0.21 ($\boldsymbol{\nu_{S\text{-}3343}}$)	
				1339			3.76 ($\boldsymbol{\nu_{S\text{-}3394}}$)	
				1640			5.27 ($\boldsymbol{\nu_{S\text{-}3343}}$)	

Density functional perturbation theory (DFPT) simulations performed at 0 K using a GGA functional and 64-atom cell.
Ab initio molecular dynamics (AIMD) simulations performed at 300 K (room temperature) using a GGA functional and 64-atom cell and machine-learning molecular dynamics (MLDM) simulations performed at 300 K using a GGA functional and 4096-atom cell. * twicely degenerated modes.
Harmonic and anharmonic vibrational frequencies (cm^{-1}) are reported for DFPT and dynamical simulations, respectively.
Configurations **(1)** (N ← HC) (Fig. 6a/b) and **(2)** (N → HC) indicate an NH^0 defect where the C—H bond points towards and away from N, respectively. For N_2H^0 in configuration **(3)** (Fig. 6c/d), $\boldsymbol{\nu_{B1}}$ and $\boldsymbol{\nu_{B2}}$ have calculated IR-activities of 0.0396 and 3.5194 $(D/Å)^2$/amu, respectively.
The origin of calculated overtone and combination bands in dynamical simulations that are not assigned a vibrational mode (e.g., ($\nu_S + \nu_B$)) are unknown.
The difference between calculated ($\boldsymbol{\nu_{S(calc)}}$) and observed ($\boldsymbol{\nu_{S(obs)}}$) C—H stretching frequencies ($\boldsymbol{\nu_{S(obs)}/\nu_{S(calc)}}$, expressed as % difference) is determined by assuming NH^0 and N_2H^0 correspond to the observed 3394 and 3343 cm^{-1} peaks, respectively. For VN_3H^0, $\boldsymbol{\nu_{S(calc)}}$ is compared to the known C—H stretching mode observed at 3107 cm^{-1} and for $VN_2H_2^0$, $\boldsymbol{\nu_{S(calc)}}$ is compared to both the 3394 and 3343 cm^{-1} peaks (e.g., $\boldsymbol{\nu_{S\text{-}3394}}$ and $\boldsymbol{\nu_{S\text{-}3343}}$).

[5] $NH^0 + V^0 = VNH^0$
$VNH^0 + N_2^0 = VN_3H^0$
[6] $NH^0 + VN_2^0 = VN_3H^0$
[7] $N_2H^0 + V^0 = VN_2H^0$
$VN_2H^0 + N_S^0 = VN_3H^0$
[8] $N_2H^0 + NV^0 = VN_3H^0$

Alternatively, such sequences may involve charged defects, for example;

[9] $NH^0 + NV^- = VN_2H^-$
$VN_2H^- + N_S^+ = VN_3H^0$

Evidence for the reagents (i.e., $NV^{0/-}$, N_S^+, VN_2^0, V^0) in each sequence are observed in the PL spectra of sample Z-16 where several peaks have been assigned, although tentatively, to various N/*V*-related defects and different nickel (Ni-) and (self-)interstitial (C_i/N_i)-related defects (Table A.3, Appendix A). Moreover, intermediate defects such as VN_2H^0 may correspond to additional H-related peaks (e.g., 3144 and 3182–85 cm^{-1}, Fig. 3) that appear at ~30 %IaA and become relatively more intense with increasing %IaA (e.g., Fig. B.2b, Appendix B). In fact, previous studies have shown an intensity correlation between the VN_2H^0 electron paramagnetic resonance (EPR) signal and IR peaks at 1375 and 1378 cm^{-1}, both of which are observed here [4,47,48] (Figs. 2d and B.1, Appendix B). The other intermediate H-related defect, VNH^0, corresponds to a known peak at 3123–3124 cm^{-1} that is rarely observed (or is relatively weak) in the spectra studied here, in accord with observations of this peak in almost exclusively in CVD diamonds [49,50]. Therefore, sequence [5] may be less operative especially in diamonds with high N contents which may drive aggregation of NH^0 according to sequences [4] and/or [6]. A relatively weak peak at 3236–37 cm^{-1} is observed in spectra with high %IaA suggesting formation of VN_4H^0 [24] requires higher degrees of annealing.

Ten spectra show an unusually large number of peaks between 3100 and 3400 cm^{-1} (e.g., Fig. 3d). As eight of these spectra have N-aggregation states of ~20–40 %IaA, it is tempting to attribute these peaks to a *transitional behaviour* of H-defects where NH^0 and N_2H^0 combine with other V/N-related defects producing many H-related defects and thus many distinct C—H modes (peaks). Similarly complex signal in the C—H stretching region has been observed *Type IbXY* orange-brown diamonds [45] and *Type Ib + IaA* orange-yellow *480 nm diamonds* [51] which have been shown to contain different Ni- and oxygen(O)-related defects [51,52]. Therefore, it is possible that some of the C—H peaks are associated with H-defects containing atoms other than N, such as Ni and/or O.

3.4. Incorporation of N and H during diamond growth

During growth, N is incorporated as a structural (covalently bonded) atom, rather than in an interstitial form (e.g., N_i *or* N_{2i}). Diamond synthesis experiments show that N is readily incorporated as N_S^0 during growth with preference for octahedral (111) sectors compared to cuboid (100) sectors [53,54]. However, this does not preclude formation of N_i or N_{2i} as substitution for N_S^0 by C_i ($N_S^0 + C_i = N_i + C$) or N_2^0 by C_i ($N_2^0 + C_i$

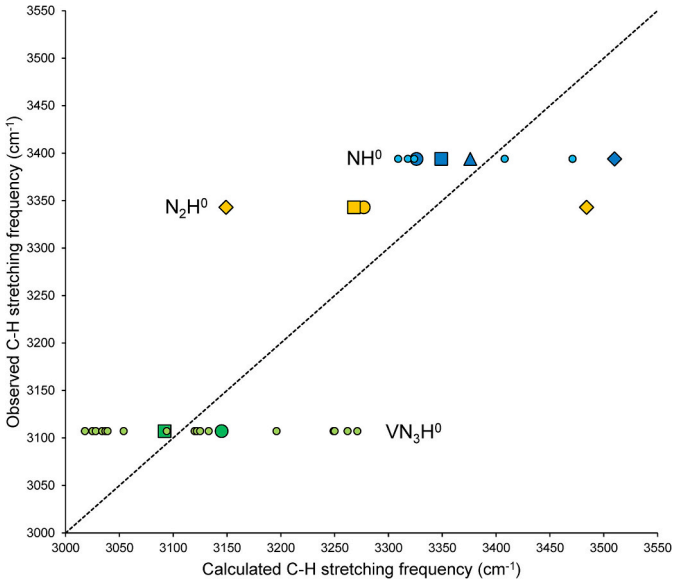


Fig. 5. Comparison of the observed and calculated C—H stretching frequencies for the NH^0, N_2H^0 and VN_3H^0 defects. The C—H stretching frequencies (IR peak positions) for the NH^0, N_2H^0 and VN_3H^0 defects are shown in blue, yellow and green, respectively. Small circles indicate C—H stretching frequencies computed in previous studies (see Table C.1, Appendix C). For NH^0 in configuration **(1)**, large circles indicate C—H stretching frequencies from DFPT simulations (at 0 K using a GGA functional and 64-atom cell). For NH^0 in configuration **(2)**, the diamond indicates the C—H stretching frequency from DFPT simulations (at 0 K using a GGA functional and 64-atom cell). For N_2H^0 in configuration **(3)**, the large circle indicates the C—H stretching frequency from DFPT simulations (at 0 K using a GGA functional and 64-atom cell). For NH^0 in configuration **(4)** and **(5)**, diamonds indicate the C—H stretching frequencies from DFPT simulations (at 0 K using a GGA functional and 64-atom cells). For VN_3H^0 in configuration **(7)**, large circles indicate C—H stretching frequencies from DFPT simulations (at 0 K using a GGA functional and 64-atom cell). Squares indicate C—H stretching frequencies from AIMD simulations (at 300 K using a GGA functional and 64-atom cell) and triangles from MLMD simulations (at 300 K using a GGA functional and 4096-atom cell). Data from all simulations conducted in this study are given in Table 3.

$= N_{2i} + C$) have relatively low activation energies (eV = 1.6) [55]. Here, C_i may form via structural deformation and/or irradiation or by C → Ni substitution [56,57]. PL spectra of sample Z-16 shows several peaks attributed to N_i/C_i-related defects (Table A.3, Appendix A). Moreover, some of the IR spectra show amber centers (broad bands at ~4064 and ~4109 cm^{-1}) indicating plastic deformation [58,59] which may generate vacancies and in turn, C_i and N_i.

Theoretical work and synthesis experiments show that H may be incorporated as hydrogenated sp^2 carbon and/or in a molecular (H_2) or atomic/ionic form occupying dislocations or interstitial sites [2,60]. The rate at which H_i is trapped by N_S^0 is a function of the concentration and spatial distribution of N_S^0 and H_i and the H_i diffusion rate. For sample Z-16, the yellow zone shows a higher average N_{tot} than the colorless zone (Table 1) and in-turn serves as a larger sink for H_i. The lower N-aggregation state of the yellow zone reflects how trapping of H_i by N_S^0 and N_2^0 may supress C → A aggregation. Interestingly, samples with %IaA < ~10 show only relatively weak peaks at 3343 and 3394 cm^{-1}. This is expected for 3343 cm^{-1} as some degree of annealing is required to drive sequence [3] and in-turn, produce N_2^0 (via $N_S^0 + N_S^0$) such that it can trap H_i to form N_2H^0. However, for 3394 cm^{-1}, this observation implies that initial trapping of H_i by N_S^0 occurs gradually from ~0–10 %IaA. During this initial annealing period, incorporated H_i diffuses extremely quickly [61,62] increasing the probability that it encounters N_S^0 or N_2^0 and is trapped.

These results also have important implications for the controlled production of point-defects in synthetic diamond with useful technological applications, for example the NV^- defect which has an electron spin state conducive to quantum sensing and information storage applications [10,11]. The properties of NV^- are altered when passivated by H^+ (forming VNH^0 defects) representing a major problem in the synthesis of NV^--bearing diamond via chemical vapor deposition from gas mixtures containing H_2 [62]. PL peaks at 575.0 and 637.1 nm attributed to NV^- and NV^0, respectively (Table A.3, Appendix A) are significantly more intense in the yellow zone of sample Z-16 indicating that the majority of H_i is trapped by N_S^0 and N_2^0 (forming NH^0 and N_2H^0) and not $NV^{-/0}$ thus mitigating production of $VNH^{-/0}$ [62]. However, although relatively weak, the 3124 cm^{-1} peak (due to VNH^0) is observed in some spectra suggesting that during primary growth (or post-growth heat/irradiation treatment), VNH^0 may form by $NH^0 + V^0$ also reducing the yield of NV^-. These early N/H-aggregation sequences should also be considered during the production of semiconducting diamond where H may complex with substitutional N-defects (and B-defects) modifying their function as electron donors or acceptors [63].

3.5. Implications for diffusion and re-equilibration of H during mantle residence

The appearance of NH^0 and N_2H^0 at low N-aggregation states (Fig. 7) indicates H_i is trapped by N during, or shortly after, diamond formation. As the N content is typically much greater than the H-content of most *Type I* diamonds, most H_i is likely trapped by N-defects. This H is retained during mantle residence as disaggregation of most N/H-related defects in experiments occurs at temperatures (>1600 °C) [23,64,65] that exceed temperatures of normal mantle residence (1160 ± 100 °C) [66] and kimberlite magmatism (~1300–1400 °C) [67,68]. It follows that post-growth re-equilibration of H with the ambient mantle via diffusive-loss of H_i is unlikely. However, re-equilibration via diffusive-gain of H may be possible but the degree of H-diffusion into the interior of the diamond would be impeded by N, especially in diamonds with high N contents. Moreover, secondary episodic growth may result in zones (or coats) with highly variable N and H contents [69,70] but as NH^0 and N_2H^0 form shortly after diamond crystallization it is likely there is minimal diffusion of H between growth zones.

Consequently, the content and isotopic composition of H in *Type I* diamond may reflect that of the diamond forming media (fluid or melt) and thus the parental reservoir from which such media is sourced. *Type I* diamonds may be the only means by which otherwise inaccessible primordial (or recycled) H (H_2O) reservoirs in the mantle can be studied. However, care must be taken when selecting such samples, high N content is favourable as it increases the probability that all initial H_i is preserved (trapped). On the other hand, samples with evidence of secondary growth or resorption may reflect different sources (compositions) of H and therefore may not reflect the original H-bearing fluid source and instead a potentially mixed recycled and primordial fluid.

4. Conclusions

By filtering the GIA database, a suite of 178 FTIR spectra recorded from extremely rare H-bearing *Type Ib + IaA* diamonds (<0.1 % of natural diamonds) [25] were selected, affording us a unique opportunity to examine the initial stages of H-defect formation and aggregation in diamond. Although numerous peaks are observed in the C—H stretching region, three peaks at 3343, 3373, and 3394 cm^{-1} are relatively intense and are often the only peaks in the C—H stretching region of poorly aggregated diamonds (<30 %IaA). It follows that these peaks may represent the first-formed H-related defect in diamond. Based on different peak-intensity correlations and excellent agreement with DFPT, AIMD and MLMD simulations, the 3343 and 3394 cm^{-1} peaks are assigned to the C—H stretching modes of the N_2H^0 and NH^0 defects, respectively. This assignment is further supported by the observed peaks at 1408, 1415 and 1428 cm^{-1} which are in agreement with calculated

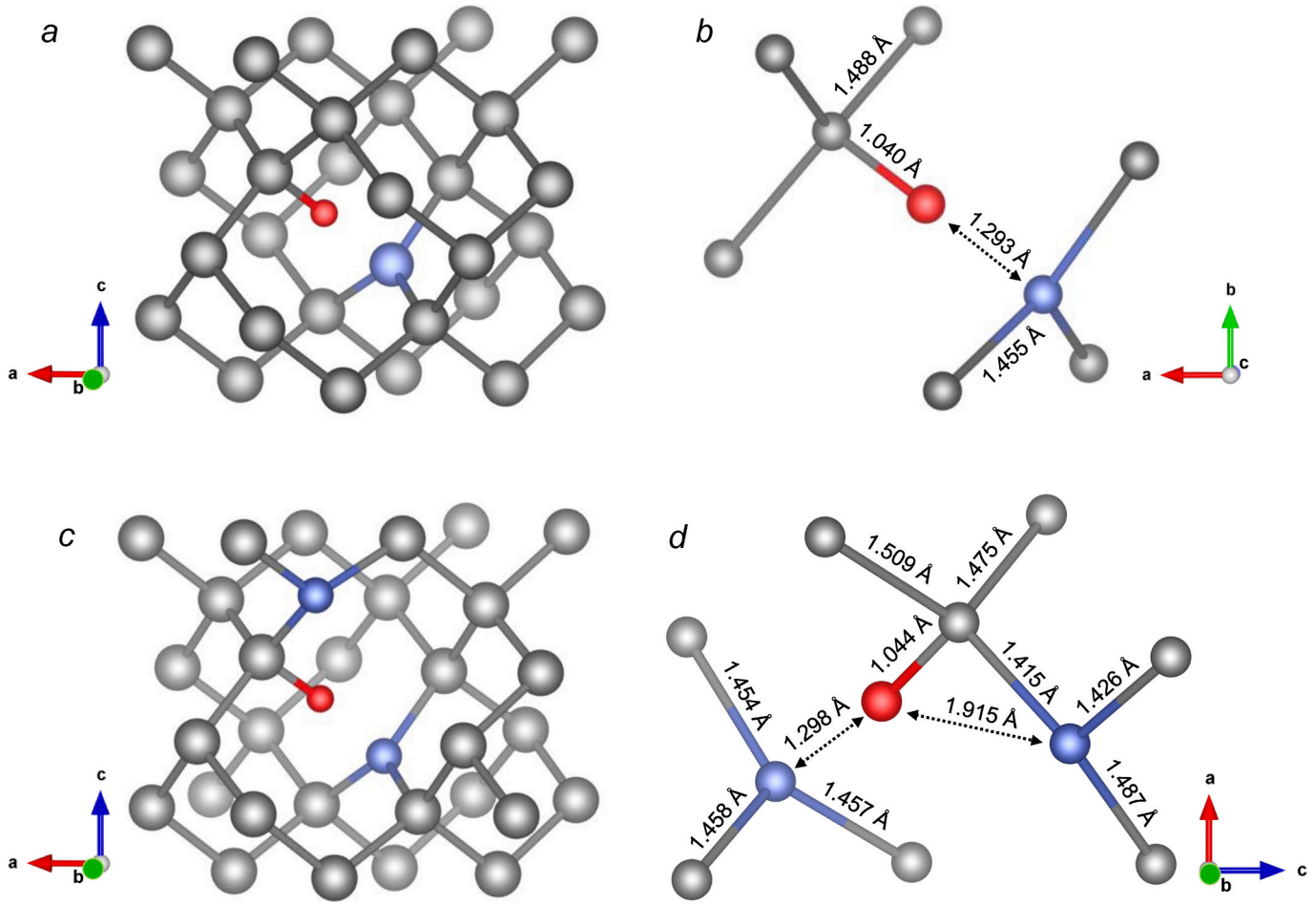


Fig. 6. The structure of the NH^0 and N_2H^0 defects. (a, b) The structure of the NH^0 defect in configuration **(1)** (see Table 3). (c, d) The structure of the N_2H^0 defect in configuration **(3)**. Carbon, nitrogen and hydrogen atoms are shown as grey, blue and red spheres, respectively. All bond distances and interatomic spacings are given in angstroms (Å). Structures were generated using the VESTA V3.4.8 software.

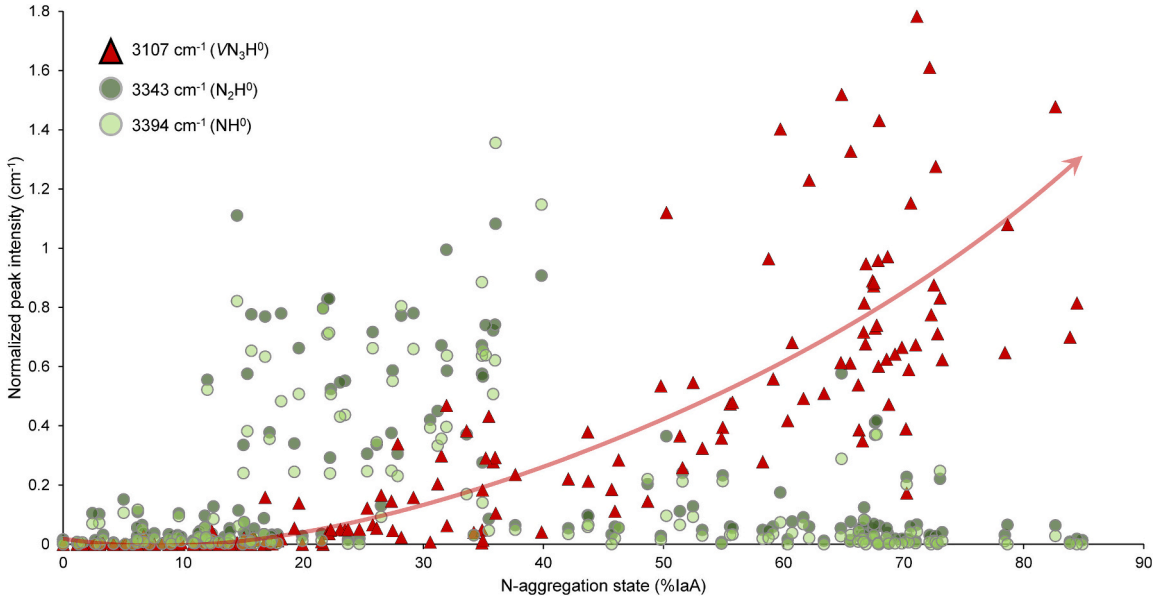


Fig. 7. Behaviour of the NH^0, N_2H^0 and VN_3H^0 defects with increasing N-aggregation. The normalized intensity of the 3343, 3394 and 3107 cm^{-1} peaks plotted as a function of N-aggregation state (%IaA). The normalized 3343 cm^{-1} (N_2H^0) and 3394 cm^{-1} (NH^0) peak intensities are shown as dark green and light green circles, respectively. The normalized 3107 cm^{-1} (VN_3H^0) peak intensities are shown as red triangles. The red arrow shows how an increase in VN_3H^0 content (3107 cm^{-1} peak intensity) occurs coincident with a decrease in the NH^0 and N_2H^0 content (3394 and 3343 cm^{-1} peak intensities, respectively) at ~35–40 %IaA.

C—H bending modes. Moreover, combination bands at 4764 and 4797 cm^{-1} and overtone bands at 2895 and 2904 cm^{-1} are in agreement with calculated ($\nu_S + \nu_B$) and (2 x ν_B) modes for N_2H^0 and NH^0, respectively and shoulders at 3335 and 3388 cm^{-1} (on the 3343 and 3394 cm^{-1} peaks, respectively) are in agreement with calculated C—H shifting effect due to $^{12}C \rightarrow {}^{13}C$ (~9 cm^{-1}) [5].

Our data shows that the NH^0 and N_2H^0 defects form when ***C***- and ***A***-centers trap H_i, respectively suggesting that H_i is trapped during, or shortly after, the formation of *Type I* diamonds. This implies that ubiquitous H-defects like VN_3H^0 (3107 cm^{-1}) [5] and VN_4H^0 (3236 cm^{-1}) [24] may not form by trapping H_i (e.g., $VN_3^0 + H_i = VN_3H^0$), as is commonly proposed [5], but instead by progressive aggregation of earlier formed defects like NH^0 and N_2H^0. This is evidenced by a clear inverse correlation between the 3343/3394 cm^{-1} and 3107 cm^{-1} peak intensities where a decrease in the NH^0 and N_2H^0 contents (at ~35–40 % IaA) is coincident with an increase in VN_3H^0 contents. Although the 3343 and 3394 cm^{-1} peaks have not been observed in synthetic (CVD) diamond, the new N/H-related aggregation sequences proposed here may be particularly important when attempting to increase the yield of technologically useful defects (e.g., NV^- used for quantum computing and information storage) whose electromagnetic properties are passivated by H_i. In fact, the prevalence of the $VNH^{0/-}$ defect (3124 cm^{-1} peak) in CVD diamond suggests that H_i may act to impede the production of $NV^{0/-}$ by $NV^{0/-} + H_i = VNH^{0/-}$ where H is sourced from the H_2 gas. However, the prevalence of $NV^{0/-}$ in sample Z-16 suggest that N and N_2 preferentially trap H_i and thus may offer a solution for enhancing production of $NV^{0/-}$ defects in synthetic diamond.

In natural diamonds with sufficiently high N-contents, the majority of the H_i incorporated during diamond formation is trapped, thereby minimizing re-equilibration of H with the surrounding ambient mantle regardless of the high diffusivities of H in diamond [61]. It follows that the H isotopic signature of *Type I* diamonds may reflect that of the diamond-forming fluid (or melt) and the primordial and/or recycled mantle reservoirs from which parental fluids are sourced.

CRediT authorship contribution statement

Maxwell C. Day: Writing – original draft, Investigation, Formal analysis, Data curation, Conceptualization. **Etienne Balan:** Writing – review & editing, Software, Investigation, Formal analysis. **Razvan Caracas:** Writing – review & editing, Software, Investigation, Formal analysis. **Michael C. Jollands:** Writing – review & editing, Resources, Investigation, Conceptualization. **Francesca Innocenzi:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Davide Novella:** Writing – review & editing, Validation. **Fabrizio Nestola:** Writing – review & editing, Visualization. **Martha G. Pamato:** Writing – review & editing, Supervision, Investigation.

Declaration of competing interest

The authors have no financial or non-financial conflicts of interest to disclose. The authors confirm that this work is original and has not been published elsewhere, nor is it currently under consideration for publication elsewhere. No artificial intelligence tools or services were used to write, construct or organize any part of this manuscript.

Acknowledgements

This work was funded by the European Union (ERC, INHERIT, Starting Grant No.101041620). Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. The authors express their gratitude to A'Dhi Lall and Taryn Linzmeyer for helping with data collection.

Appendices A-E. Supplementary data

The online version contains supplementary material (Appendices A, B, C, D and E) available at https://researchdata.cab.unipd.it/id/eprint/1601. Supplementary material can also be accessed from the University of Padova repository at https://researchdata.cab.unipd.it/id/eprint/1601. All data is also available upon request to Maxwell C. Day.

Data availability

Data will be made available on request.

References

[1] M.C. Day, M.G. Pamato, D. Novella, F. Nestola, Imperfections in natural diamond: the key to understanding diamond genesis and the mantle, Riv. Nuovo Cim. 46 (2023) 381–471.
[2] J.P. Goss, R. Jones, M.I. Heggie, C.P. Ewels, P.R. Briddon, S. Öberg, Theory of hydrogen in diamond, Phys. Rev. B 65 (2002) 115207, https://doi.org/10.1103/PhysRevB.65.115207.
[3] B. Dischler, Handbook of Spectral Lines in Diamond, 2012, https://doi.org/10.1007/978-3-642-22215-3.
[4] M.N.R. Ashfold, J.P. Goss, B.L. Green, P.W. May, M.E. Newton, C.V. Peaker, Nitrogen in Diamond, Chem. Rev. 120 (2020) 5745–5794, https://doi.org/10.1021/acs.chemrev.9b00518.
[5] J.P. Goss, P.R. Briddon, V. Hill, R. Jones, M.J. Rayson, Identification of the structure of the 3107 cm^{-1} H-related defect in diamond, Phys. Condens. Matter 26 (2014) 145801, https://doi.org/10.1088/0953-8984/26/14/145801.
[6] S. Salustro, F.S. Gentile, A. Erba, P. Carbonniére, K.E. El-Kelany, R. Dovesi, The characterization of the VNxHy defects in diamond through the infrared vibrational spectrum. A quantum mechanical investigation, Carbon N Y 132 (2018), https://doi.org/10.1016/j.carbon.2018.02.045.
[7] F.S. Gentile, S. Salustro, G. Di Palma, M. Causà, P. D'Arco, R. Dovesi, Hydrogen, boron and nitrogen atoms in diamond: a quantum mechanical vibrational analysis, Theor. Chem. Accounts 137 (2018) 154, https://doi.org/10.1007/s00214-018-2375-0.
[8] M.C. Day, M.C. Jollands, D. Novella, F. Nestola, R. Dovesi, M.G. Pamato, Hydrogen-related defects in diamond: a comparison between observed and calculated FTIR spectra, Diam. Relat. Mater. 143 (2024) 110866, https://doi.org/10.1016/j.diamond.2024.110866.
[9] C.J.H. Wort, R.S. Balmer, Diamond as an electronic material, Mater. Today 11 (2008) 22–28, https://doi.org/10.1016/S1369-7021(07)70349-8.
[10] J.J. Hamlin, B.B. Zhou, Extreme diamond-based quantum sensors, Science (1979) 366 (2019) 1312–1313, https://doi.org/10.1126/science.aaz4982.
[11] G. Thiering, A. Gali, Color Centers in Diamond for Quantum Applications, 2020, pp. 1–36, https://doi.org/10.1016/bs.semsem.2020.03.001.
[12] M. Palot, P. Cartigny, J.W. Harris, F.V. Kaminsky, T. Stachel, Evidence for deep mantle convection and primordial heterogeneity from nitrogen and carbon stable isotopes in diamond, Earth Planet. Sci. Lett. 357 (2012), https://doi.org/10.1016/j.epsl.2012.09.015.
[13] T. Stachel, P. Cartigny, T. Chacko, D.G. Pearson, Carbon and nitrogen in mantle-derived diamonds, in: Diamond: Genesis, Mineral Geochem, 2023, https://doi.org/10.2138/rmg.2022.88.15.
[14] A. Mainwood, Nitrogen and nitrogen-vacancy complexes and their formation in diamond, Phys. Rev. B 49 (1994), https://doi.org/10.1103/PhysRevB.49.7934.
[15] D. Howell, C.J. O'Neill, K.J. Grant, W.L. Griffin, N.J. Pearson, S.Y. O'Reilly, M-FTIR mapping: distribution of impurities in different types of diamond growth, Diam. Relat. Mater. 29 (2012), https://doi.org/10.1016/j.diamond.2012.06.003.
[16] W.R. Taylor, D. Canil, H. Judith Milledge, Kinetics of Ib to IaA nitrogen aggregation in diamond, Geochim. Cosmochim. Acta 60 (1996) 4725–4733, https://doi.org/10.1016/S0016-7037(96)00302-X.
[17] W.R. Taylor, A.L. Jaques, M. Ridd, Nitrogen-defect aggregation characteristics of some Australasian diamonds: time-temperature constraints on the source regions of pipe and alluvial diamonds, Am. Min. 75 (1990).
[18] I. Kiflawi, H. Kanda, D. Fisher, S.C. Lawson, The aggregation of nitrogen and the formation of a centres in diamonds, Diam. Relat. Mater. 6 (1997) 1643–1649, https://doi.org/10.1016/S0925-9635(97)00207-0.
[19] T. Hainschwang, E. Fritsch, F. Notari, B. Rondeau, A new defect center in type Ib diamond inducing one phonon infrared absorption: the Y center, Diam. Relat. Mater. 21 (2012) 120–126, https://doi.org/10.1016/j.diamond.2011.11.002.
[20] S.C. Lawson, D. Fisher, D.C. Hunt, M.E. Newton, On the existence of positively charged single-substitutional nitrogen in diamond, Phys. Condens. Matter 10 (1998) 6171–6180, https://doi.org/10.1088/0953-8984/10/27/016.
[21] I. Kiflawi, A. Mainwood, H. Kanda, D. Fisher, Nitrogen interstitials in diamond, Phys. Rev. B 54 (1996) 16719–16726, https://doi.org/10.1103/PhysRevB.54.16719.
[22] J.P. Goss, P.R. Briddon, S. Papagiannidis, R. Jones, Interstitial nitrogen and its complexes in diamond, Phys. Rev. B 70 (2004) 235208, https://doi.org/10.1103/PhysRevB.70.235208.
[23] F. De Weerdt, A.T. Collins, Optical study of the annealing behaviour of the 3107 cm^{-1} defect in natural diamonds, Diam. Relat. Mater. 15 (2006) 593–596, https://doi.org/10.1016/j.diamond.2006.01.005.

[24] T.J. Keat, D.J.L. Coxon, R.J. Cruddace, V.G. Stavros, M.E. Newton, J. Lloyd-Hughes, The 3237 cm^{-1} diamond defect: ultrafast vibrational dynamics, concentration calibration, and relationship to the N3VH0 defect, Diam. Relat. Mater. 141 (2024) 110661, https://doi.org/10.1016/j.diamond.2023.110661.
[25] G.E. Harlow, The Nature of Diamonds, 1st ed., Cambridge University Press, Cambridge, UK, 1998.
[26] M.C. Day, M.C. Jollands, F. Innocenzi, D. Novella, F. Nestola, M.G. Pamato, Infrared spectroscopy of natural type Ib + IaA diamond: insights into the formation of Y-centers and the early aggregation of nitrogen, Am. Min. (2025), https://doi.org/10.2138/am-2024-9722 (accepted, in-press).
[27] M. Wojdyr, Fityk: a general-purpose peak fitting program, J. Appl. Crystallogr. 43 (2010) 1126–1128, https://doi.org/10.1107/S0021889810030499.
[28] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. 77 (1996) 3865–3868, https://doi.org/10.1103/PhysRevLett.77.3865.
[29] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G.L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A.P. Seitsonen, A. Smogunov, P. Umari, R.M. Wentzcovitch, QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials, Phys. Condens. Matter 21 (2009) 395502, https://doi.org/10.1088/0953-8984/21/39/395502.
[30] D.R. Hamann, Optimized norm-conserving Vanderbilt pseudopotentials, Phys. Rev. B 88 (2013) 085117, https://doi.org/10.1103/PhysRevB.88.085117.
[31] M. Schlipf, F. Gygi, Optimization algorithm for the generation of ONCV pseudopotentials, Comput. Phys. Commun. 196 (2015) 36–44, https://doi.org/10.1016/j.cpc.2015.05.011.
[32] S. Baroni, S. de Gironcoli, A. Dal Corso, P. Giannozzi, Phonons and related crystal properties from density-functional perturbation theory, Rev. Mod. Phys. 73 (2001) 515–562, https://doi.org/10.1103/RevModPhys.73.515.
[33] G. Kresse, J. Hafner, Ab initio molecular dynamics for liquid metals, Phys. Rev. B 47 (1993) 558–561, https://doi.org/10.1103/PhysRevB.47.558.
[34] G. Kresse, D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, Phys. Rev. B 59 (1999) 1758–1775, https://doi.org/10.1103/PhysRevB.59.1758.
[35] P.E. Blöchl, Projector augmented-wave method, Phys. Rev. B 50 (1994) 17953–17979, https://doi.org/10.1103/PhysRevB.50.17953.
[36] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. 77 (1996) 3865–3868, https://doi.org/10.1103/PhysRevLett.77.3865.
[37] H.J. Monkhorst, J.D. Pack, Special points for Brillouin-zone integrations, Phys. Rev. B 13 (1976) 5188–5192, https://doi.org/10.1103/PhysRevB.13.5188.
[38] W.G. Hoover, Canonical dynamics: equilibrium phase-space distributions, Phys. Rev. A (Coll Park) 31 (1985) 1695–1697, https://doi.org/10.1103/PhysRevA.31.1695.
[39] S. Nosé, A unified formulation of the constant temperature molecular dynamics methods, J. Chem. Phys. 81 (1984) 511–519, https://doi.org/10.1063/1.447334.
[40] R. Jinnouchi, J. Lahnsteiner, F. Karsai, G. Kresse, M. Bokdam, Phase transitions of hybrid perovskites simulated by machine-learning force fields trained on the fly with Bayesian inference, Phys. Rev. Lett. 122 (2019) 225701, https://doi.org/10.1103/PhysRevLett.122.225701.
[41] R. Caracas, A. Kobsch, N.V. Solomatova, Z. Li, F. Soubiran, J.-A. Hernandez, Analyzing melts and fluids from ab initio molecular dynamics simulations with the UMD package, Visualized Exp. (2021), https://doi.org/10.3791/61534.
[42] K.V. Smit, U.F.S. D'Haenens-Johansson, D. Howell, L.C. Loudin, W. Wang, Deformation-related spectroscopic features in natural type Ib-IaA diamonds from Zimmi (west African craton), Mineral. Petrol. 112 (2018) 243–257, https://doi.org/10.1007/s00710-018-0587-6.
[43] E. Fritsch, T. Hainschwang, L. Massi, B. Rondeau, Hydrogen-related optical centers in natural diamond: an update, New Diamond Front. Carbon Technol. 17 (2007).
[44] G.S. Woods, A.T. Collins, Infrared absorption spectra of hydrogen complexes in type I diamonds, J. Phys. Chem. Solids 44 (1983), https://doi.org/10.1016/0022-3697(83)90078-1.
[45] T. Hainschwang, F. Notari, G. Pamies, A defect study and classification of Brown diamonds with non-deformation-related color, Minerals 10 (2020) 914, https://doi.org/10.3390/min10100914.
[46] S.H. Connell, J.P.F. Sellschop, J.E. Butler, R.D. Maclear, B.P. Doyle, I.Z. Machi, A study of the mobility and trapping of minor hydrogen concentrations in diamond in three dimensions using quantitative ERDA microscopy, Diam. Relat. Mater. 7 (1998) 1714–1718, https://doi.org/10.1016/S0925-9635(98)00266-0.
[47] C.V. Peaker, First Principles Study of Point Defects in Diamond (PhD), Newcastle University, 2018.
[48] C.V. Peaker, J.P. Goss, P.R. Briddon, A.B. Horsfall, M.J. Rayson, Di-nitrogen-vacancy-hydrogen defects in diamond: a computational study, Phys. Status Solidi A 212 (2015) 2616–2620, https://doi.org/10.1002/pssa.201532216.
[49] S. Liggins, Identification of Point Defects in Treated Single Crystal Diamond (PhD Thesis), University of Warwick, 2010.
[50] J.P. Goss, P.R. Briddon, R. Jones, S. Sque, The vacancy–nitrogen–hydrogen complex in diamond: a potential deep centre in chemical vapour deposited material, J. Phys. Condens. Matter 15 (2003) S2903–S2911, https://doi.org/10.1088/0953-8984/15/39/014.
[51] M.Y. Lai, M.F. Hardman, S. Eaton-Magaña, C.M. Breeding, V.A. Schwartz, A. T. Collins, Spectroscopic characterization of diamonds colored by the 480 nm absorption band, Diam. Relat. Mater. 142 (2024) 110825, https://doi.org/10.1016/j.diamond.2024.110825.
[52] T. Hainschwang, Diamants de type Ib: Relations entre les propriétés physiques et gemmologiques des diamants contenant de l'azote isolé (PhD Thesis), Université de Nantes, 2014.
[53] R. Samlenski, C. Haug, R. Brenn, C. Wild, R. Locher, P. Koidl, Incorporation of nitrogen in chemical vapor deposition diamond, Appl. Phys. Lett. 67 (1995) 2798–2800, https://doi.org/10.1063/1.114788.
[54] R.C. Burns, V. Cvetkovic, C.N. Dodge, D.J.F. Evans, M.-L.T. Rooney, P.M. Spear, C. M. Welbourn, Growth-sector dependence of optical features in large synthetic diamonds, J. Cryst. Growth 104 (1990) 257–279, https://doi.org/10.1016/0022-0248(90)90126-6.
[55] J.P. Goss, B.J. Coomer, R. Jones, C.J. Fall, P.R. Briddon, S. Öberg, Extended defects in diamond: the interstitial platelet, Phys. Rev. B Condens. Matter Mater. Phys. 67 (2003), https://doi.org/10.1103/PhysRevB.67.165208.
[56] A.T. Collins, Colour centres in diamond, Gemmol 18 (1982) 37–75, https://doi.org/10.15506/JoG.1982.18.1.37.
[57] D. Fisher, S.J. Sibley, C.J. Kelly, Brown colour in natural diamond and interaction between the brown related and other colour-inducing defects, Phys. Condens. Matter 21 (2009) 364213, https://doi.org/10.1088/0953-8984/21/36/364213.
[58] L. Massi, E. Fritsch, A.T. Collins, T. Hainschwang, F. Notari, The "amber centres" and their relation to the brown colour in diamond, Diam. Relat. Mater. 14 (2005) 1623–1629, https://doi.org/10.1016/j.diamond.2005.05.003.
[59] T. Hainschwang, F. Notari, G. Pamies, A defect study and classification of Brown diamonds with deformation-related color, Minerals 10 (2020) 903, https://doi.org/10.3390/min10100903.
[60] D. Ballutaud, T. Kociniewski, J. Vigneron, N. Simon, H. Girard, Hydrogen incorporation, bonding and stability in nanocrystalline diamond films, Diam. Relat. Mater. 17 (2008) 1127–1131, https://doi.org/10.1016/j.diamond.2008.01.067.
[61] D.J. Cherniak, E.B. Watson, V. Meunier, N. Kharche, Diffusion of helium, hydrogen and deuterium in diamond: experiment, theory and geochemical applications, Geochim. Cosmochim. Acta 232 (2018) 206–224, https://doi.org/10.1016/j.gca.2018.04.029.
[62] A. Stacey, T.J. Karle, L.P. McGuinness, B.C. Gibson, K. Ganesan, S. Tomljenovic-Hanic, A.D. Greentree, A. Hoffman, R.G. Beausoleil, S. Prawer, Depletion of nitrogen-vacancy color centers in diamond via hydrogen passivation, Appl. Phys. Lett. 100 (2012), https://doi.org/10.1063/1.3684612.
[63] J.P. Goss, P.R. Briddon, R. Jones, S. Sque, Donor and acceptor states in diamond, Diam. Relat. Mater. 13 (2004) 684–690, https://doi.org/10.1016/j.diamond.2003.08.028.
[64] I. Kiflawi, D. Fisher, H. Kanda, G. Sittas, The creation of the 3107 cm^{-1} hydrogen absorption peak in synthetic diamond single crystals, Diam. Relat. Mater. 5 (1996) 1516–1518, https://doi.org/10.1016/S0925-9635(96)00568-7.
[65] Z. Song, J. Su, W. Zhu, T. Lu, Y. Wang, S. He, Spectroscopic study of the 3107 cm^{-1} and 3143 cm^{-1} H-related defects in type Ib diamonds, Crystals (Basel) 12 (2022) 1352, https://doi.org/10.3390/cryst12101352.
[66] T. Stachel, R.W. Luth, Diamond formation - where, when and how? Lithos 220 (2015) https://doi.org/10.1016/j.lithos.2015.01.028.
[67] L. Wilson, J.W. Head III, An integrated model of kimberlite ascent and eruption, Nature 447 (2007) 53–57, https://doi.org/10.1038/nature05692.
[68] R.S.J. Sparks, L. Baker, R.J. Brown, M. Field, J. Schumacher, G. Stripp, A. Walters, Dynamical constraints on kimberlite volcanism, Volcanol. Geotherm. Res. 155 (2006) 18–48, https://doi.org/10.1016/j.jvolgeores.2006.02.010.
[69] E. Vasilev, N. Gubanov, D. Zedgenizov, Point defects in coated diamonds, Diam. Relat. Mater. 148 (2024) 111519, https://doi.org/10.1016/j.diamond.2024.111519.
[70] E.A. Vasilev, D.A. Zedgenizov, I.V. Klepikov, The enigma of cuboid diamonds: the causes of inverse distribution of optical centers within the growth zones, Austrian J. Geosci. (2020) 59–70, https://doi.org/10.3190/jgeosci.301.