True molecular conformation and structure determination by three-dimensional electron diffraction of PAH by-products potentially useful for electronic applications

Iryna Andrusenko,[a] Charlie L. Hall,[b] Enrico Mugnaioli,[a,c] Jason Potticary,[b] Simon R. Hall,[b] Werner Schmidt,[d] Siyu Gao,[e] Kaiji Zhao,[e] Noa Marom[e] and Mauro Gemmi[a]*

[a]Center for Material Interfaces, Electron Crystallography, Instituto Italiano di Tecnologia, Pontedera 56025, Italy, [b]School of Chemistry, University of Bristol, Bristol BS8 1TS, United Kingdom, [c]Department of Earth Sciences, University of Pisa, Pisa 56126, Italy, [d]PAH Research, Igling-Holzhausen, D-86859, Germany, and [e]Department of Materials Science and Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA. *Correspondence e-mail: mauro.gemmi@iit.it

The true molecular conformation and the crystal structure of benzo[e]dinaphtho[2,3-a;1′,2′,3′,4′-ghi]fluoranthene, 7,14-diphenylnaphtho[1,2,3,4-cde]bisanthene and 7,16-diphenylnaphtho[1,2,3,4-cde]helianthrene were determined ab initio by 3D electron diffraction. All three molecules are remarkable polycyclic aromatic hydrocarbons. The molecular conformation of two of these compounds could not be determined via classical spectroscopic methods due to the large size of the molecule and the occurrence of multiple and reciprocally connected aromatic rings. The molecular structure of the third molecule was previously considered provisional. These compounds were isolated as by-products in the synthesis of similar products and were at the same time nanocrystalline and available only in very limited amounts. 3D electron diffraction data, taken from submicrometric single crystals, allowed for direct ab initio structure solution and the unbiased determination of the internal molecular conformation. Detailed synthetic routes and spectroscopic analyses are also discussed. Based on many-body perturbation theory simulations, benzo[e]dinaphtho[2,3-a;1′,2′,3′,4′-ghi]-fluoranthene may be a promising candidate for triplet–triplet annihilation and 7,14-diphenylnaphtho[1,2,3,4-cde]bisanthene may be a promising candidate for intermolecular singlet fission in the solid state.

1. Introduction

Polycyclic aromatic hydrocarbons (PAHs) are organic molecules containing a series of conjugated carbon rings. The term 'PAH' is generally restricted to compounds consisting of only C and H atoms, and comprising two or more aromatic rings bonded in various arrangements (Lawal, 2017). Early syntheses of PAHs (Anschütz, 1886; Scholl et al., 1910; Scholl & Meyer, 1932) attracted considerable interest due to their optoelectronic properties, which led to the creation of a number of functionalized compounds based on PAH backbones (Buchlovič et al., 2013; Ko et al., 2018; Castro et al., 2019). In particular, in the mid-1900s, Clar and co-workers synthesized and identified a number of PAHs and related compounds with unique physical and chemical properties (Clar et al., 1964, 1981; Clar & Schmidt, 1975).

In most cases, the identity of the newly synthesized organic compounds can be determined from combustion analysis and by a combination of NMR, IR, UV and mass spectrometry. In particular, a comparison of the UV and photoelectron (PE) spectral data provides information about the purity and helps

1 μm

view down (100)

view down (010)

view down (001)
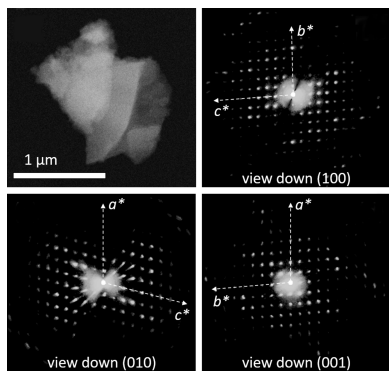
to distinguish 'isotopic PAHs', *i.e.* compounds with different shapes but identical UV band positions (Clar & Schmidt, 1976, 1979; Clar *et al.*, 1981).

It is often impossible to characterize PAHs by single-crystal X-ray diffraction (XRD) because of the difficulty in obtaining suitable sufficiently large crystals. Therefore, the final confirmation is generally attained by powder XRD analysis. However, when there are many possible molecular structures that can match a particular stoichiometry, it may be difficult to determine the exact conformation based on spectroscopic data alone. Furthermore, PXRD is intrinsically limited to one-dimensional data and hampered by the overlap of independent reflections, which may become rather severe for unit-cell parameters longer than 20 Å. As a consequence of this unsatisfactory state of affairs, the conformation of several large PAHs mentioned in the literature is uncertain or unknown, and many more are considered suspect (Clar *et al.*, 1981; Fetzer, 2000, 2007; Wilkes, 2010).

In recent years, electron diffraction (ED) has become a routine option for crystal structure determination when crystals sufficiently large for single-crystal XRD cannot be grown. The key development was the establishment of procedures for collecting and recombining 3D electron diffraction (3D ED) data (Kolb *et al.*, 2007; Wan *et al.*, 2013; Gemmi *et al.*, 2019; Nannenga & Gonen, 2019). The 3D ED method has been used for the structure determination of hundreds of previously unknown structures, including inorganic materials (Kaiukov *et al.*, 2020; Krysiak *et al.*, 2021; Kapaca *et al.*, 2021), small-molecule organics (Andrusenko *et al.*, 2019; Jones *et al.*, 2019; Brázda *et al.*, 2019; Cui *et al.*, 2020; Levine *et al.*, 2020; Bruhn *et al.*, 2021; Hall *et al.*, 2021; Andrusenko *et al.*, 2021; Papi *et al.*, 2021), peptides and macromolecules (Sawaya *et al.*, 2016; Krotee *et al.*, 2018; Lanza *et al.*, 2019; Xu *et al.*, 2019; Warmack *et al.*, 2019). It is noteworthy that *ab initio* crystal structure determination does not require information about the molecular conformation, but only a rough estimation of the atomic content of the unit cell. In this article, we report the true molecular conformation and the crystal structure determined by the 3D ED method for three PAHs belonging to two different systems.

Fluoranthenes are one of the most studied classes of PAHs (Haritash & Kaushik, 2009). They are also known as 'non-alternant PAHs' because they contain six-membered benzene rings fused with an additional five-membered ring (Gupte *et*

al., 2016). Fluoranthene-based PAHs serve as a basic unit in preparing chromophores for organic light-emitting diodes (OLEDS) and field-effect transistors (Saranya *et al.*, 2011). Benzo[*e*]dinaphtho[2,3-*a*;1′,2′,3′,4′-*ghi*]fluoranthene (I) is formed by pyrolysis together with two other hydrocarbons, namely, benzo[*a*]phenanthro[9′,10′-*c*]tetracene (II) and dibenzo[*a,k*]naphtho[1,2,3,4-*ghi*]perylene (III). Fig. 1 illustrates a stepwise model of the corresponding synthetic route described by Clar and co-workers (Clar *et al.*, 1964). The molecular conformation of II is supported by several spectroscopic data, its reactive behaviour and, most conclusively, by an alternative synthesis with Zn dust that produces II exclusively (Fig. S1). Conversely, the molecular conformations of I and III were only reported as tentative (Clar *et al.*, 1964). Spectroscopic data for II and III are presented and discussed in the supporting information.

The second system studied in this article involves both bisanthene and helianthrene compounds. Initially, 4,11-diphenylbisanthene (IV) was synthesized by Sauvage, who only determined the maximum positions of the longwave absorption bands (Sauvage, 1947*a*,*b*). Renewed interest in this compound was sparked in the 1970s, as it was shown to be one of the few molecules to fluoresce at IR wavelengths (Maulding, 1970; Rauhut *et al.*, 1975). Following the procedure of Sauvage's synthesis (Sauvage, 1947*a*,*b*), we attempted to synthesize new IR emitters with a high fluorescent quantum yield starting from helianthrone. Fig. 2 illustrates a stepwise model of the corresponding synthetic route. During this reaction, in addition to 4,11-diphenylbisanthene (IV), a number of by-products occur, among which only 7,16-diphenylhelianthrene (V) was identified previously (Arabei & Pavich, 2004). This compound appears in a greater yield when N_2 is bubbled through the system during the reaction (Fort, 2010). Two additional by-products of the reaction, VI and VII, were not identified *via* spectroscopic data, but appear to form *via* a Diels–Alder condensation reaction. The molecular structures of by-products VI and VII have hitherto remained unknown, due to the large size of the molecules and the occurrence of multiple and reciprocally connected aromatic rings.

Here, we present the *ab initio* structure determination of compounds I, VI and VII, obtained through 3D ED data. This method also allowed for the confirmation of the predicted molecular conformation for compound I and for the true molecular identification of compounds VI and VII. Moreover,
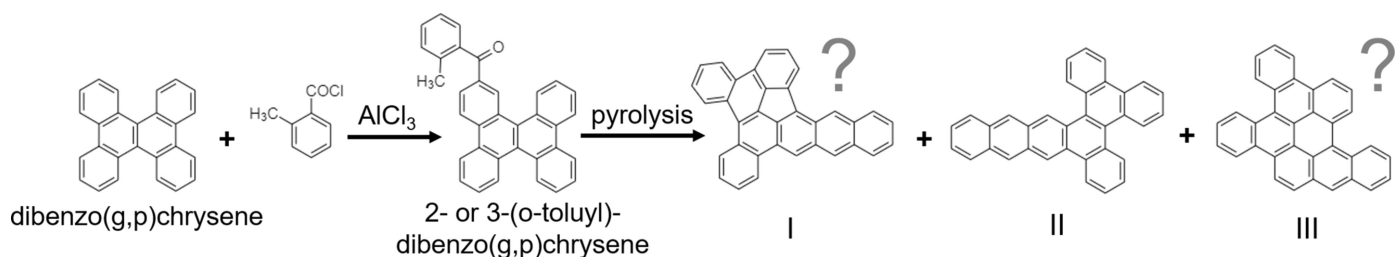
Figure 1
A stepwise model of the synthesis of benzo[*e*]dinaphtho[2,3-*a*;1′,2′,3′,4′-*ghi*]fluoranthene (I) and related products (II and III). The molecule of I was considered provisional before its structure determination by 3D ED, whereas the molecular structure of III is still speculative.
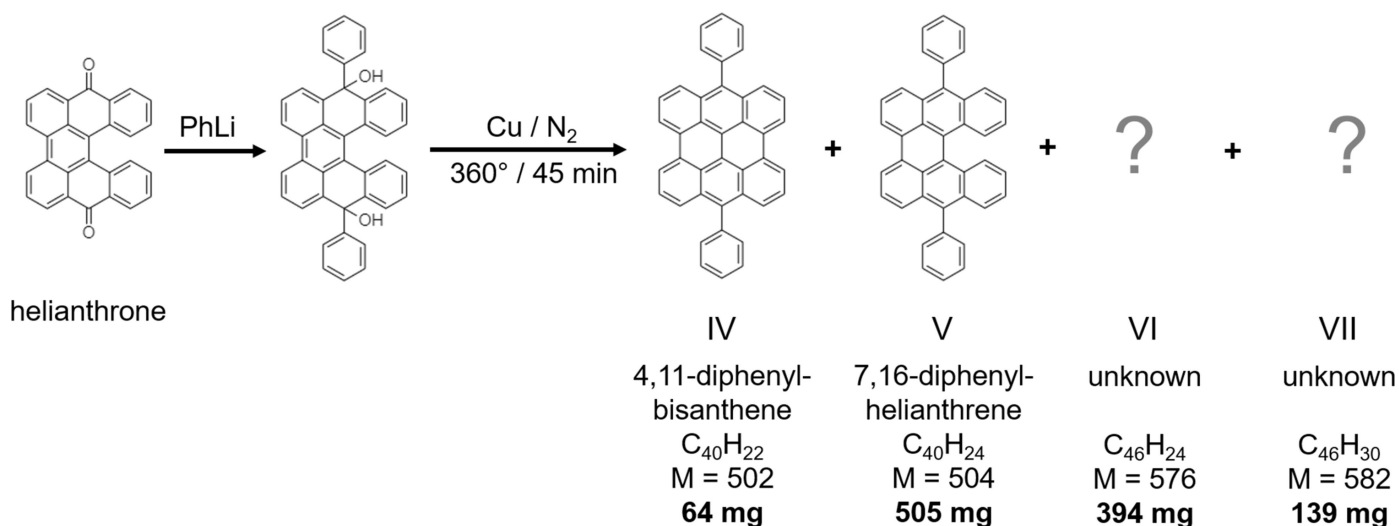
Figure 2
A stepwise model of the synthesis of 4,11-diphenylbisanthene (IV). The by-product V was identified and determined in Arabei & Pavich (2004). The molecular conformation and crystal structure of by-products VI and VII were unknown prior to their determination by the 3D ED method. For each product of the reaction, the formula, molecular weight (M) and yield (in bold) are reported.

detailed synthetic routes and spectroscopic analyses of all the synthetic products are reported. Having solved the structures of these compounds, we eventually use many-body perturbation theory calculations within the GW approximation and the Bethe–Salpeter equation (BSE) (Golze *et al.*, 2019; Blase *et al.*, 2020) to assess their potential usefulness in photovoltaics.

2. Methods

2.1. Synthesis

Samples I, II and III were obtained starting from a Friedel–Crafts reaction (Friedel & Crafts, 1877) between dibenzo[g,p]chrysene (10 g) and o-toluyl chloride, with $AlCl_3$ as catalyst. The two reactants were connected by a ketone group, either 2- or 3-(o-toluyl)dibenzo[g,p]chrysene (12 g). The 1- and 4-positions are excluded on steric grounds. After Elbs pyrolysis (15 g, 420–430 °C, 10 min under CO_2), fractional crystallization and chromatography showed that three products had formed (Fig. 1): I (0.46 g, violet–red needles, m.p. 348–350 °C), II (1.5 g, thick orange–yellow needles, m.p. 258–259 °C) and III (0.1 g, fibrous orange–yellow needles, m.p. 282°C).

Multiple syntheses of samples IV, V, VI and VII (Fig. 2 and Fig. S2) started from adding phenyl–Li to helianthrone, which can be prepared on a 100 g basis (Scholl & Mansfeld, 1910). The resulting diol was pyrolyzed with ten times the quantity of Cu powder for 45 min at 360 °C under N_2. Gradient sublimation gave an excellent yield of compound V and a smaller yield of IV, both not free of each other. There were also two major by-products: VI (m.p. >427 °C) and VII (m.p. 426–427 °C). Repeated gradient sublimation and chromatography gave four pure (>99%) products. Besides VI and VII, we eventually identified at least four other by-products, but these were not obtained free of each other.

2.2. Spectroscopy

UV spectroscopic analyses were conducted using an Agilent Cary 300 spectrophotometer. UV spectra were collected at room temperature using benzene (above 275 nm) or cyclohexane (above 200 nm) solvents (Fig. 3, Figs. S4–S7 and Tables S1–S3). Over 150 UV spectra had to be recorded to ensure the identity and homogeneity of the various fractions. Fractional sublimation was obtained using a Heraeus tube oven ROK 3/30 with a Kelvitron REK 19-20 controller and a melting-point correction. Fluorescence (FL) spectroscopic analyses were carried out with an Aminco–Bowman SPF-500 spectrofluorometer (Table S4). Compound I was insufficiently soluble in cyclohexane solvent, so for the determination of reliable extinction coefficients, only benzene was used as solvent. Compound VI is well soluble in benzene and less soluble in cyclohexane. Compound VII is well soluble in both solvents. The gas-phase PE spectrum for compound I was obtained using a PerkinElmer PS-18 spectrometer with a Helectros Developments photon source (Fig. 3(b)). Samples were also analysed by positive ion mode Matrix-assisted laser desorption/ionization on a Bruker Daltonics ultrafleXtreme II mass spectrometer using Colloidal Graphite as the matrix (Figs. S8–S10).

### 2.3. TEM microscopy and 3D electron diffraction

High-angle annular dark-field scanning transmission electron microscopy (HAADF–STEM) imaging and ED data were recorded with a Zeiss Libra 120 TEM operating at 120 kV and equipped with a LaB_6 thermionic source. 3D ED was performed in STEM mode after defocusing the beam. Therefore, the beam for the 3D ED experiments was almost parallel when crossing the sample (Köhler illumination). The TEM was set in STEM mode, but the beam, which is normally very convergent in this configuration, was purposely defocused (Benner & Probst, 1994). The actual beam convergence

was 60 µrad. ED patterns were collected with a beam size of about 150 nm in diameter, obtained using a 5 µm C2 condenser aperture. Data were recorded by a single-electron ASI MEDIPIX detector (Nederlof *et al.*, 2013). An extremely low dose illumination, corresponding to 0.01 e$^-$ Å^{-2} s^{-1}, was used in order to avoid beam damage. The total dose during data collection depends on many experimental parameters, such as the number of frames, the exposure time and the image
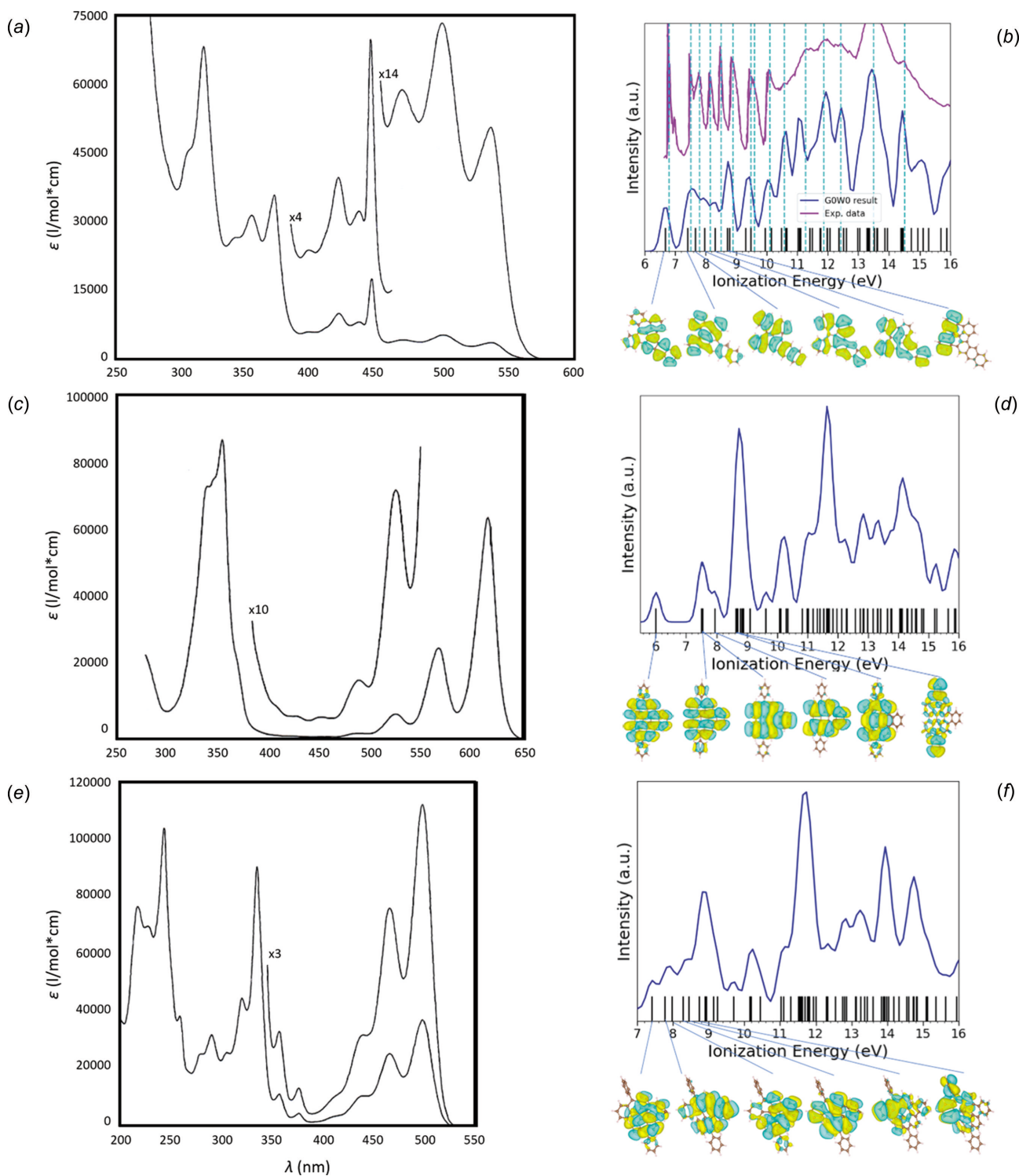
Figure 3
(*a*)/(*c*)/(*e*) UV and (*b*)/(*d*)/(*f*) simulated GW@PBE0 spectra of (*a*)/(*b*) compound I, (*c*)/(*d*) compound VI and (*e*)/(*f*) compound VII. On the UV spectra, numbers indicate zoomed areas. Comparison of the simulated GW@PBE0 spectrum (in blue) with the experimental PE spectrum (in purple) of compound I are highlighted (in (*b*)). Frontier orbitals of all three species are also visualized. Gaussian broadening of 0.2 eV was applied to the computed ionization energies to simulate the resolution of the experiment.
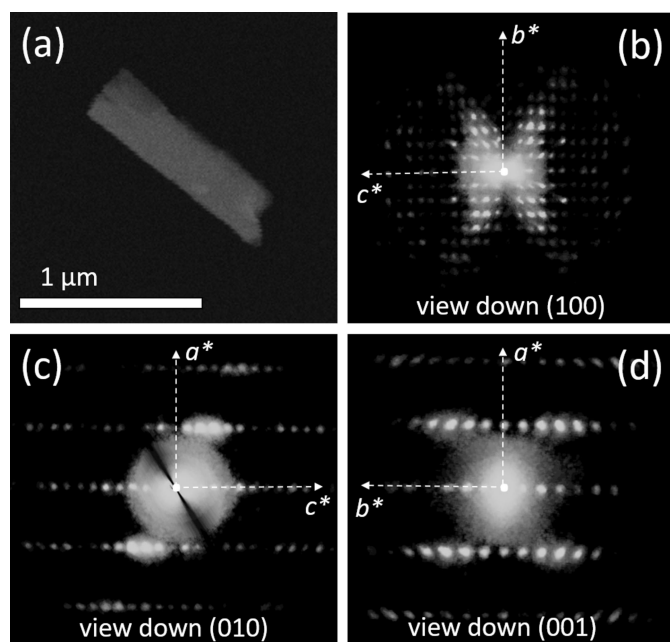
Figure 4
(a) HAADF–STEM image of a typical crystal of compound I, used for 3D ED data acquisition. (b)/(c)/(d) Reconstructed 3D ED data viewed along the main crystallographic directions. Note that these are projections of a 3D diffraction volume and not 2D oriented ED patterns, and therefore axial extinctions are not visible.

tracking mode, and is therefore different for each data collection. A rough estimation of the total dose during a stepwise data collection is in a range of about $1–5\,e^{-}\,\mathring{A}^{-2}$. When working with such low doses, the availability of a
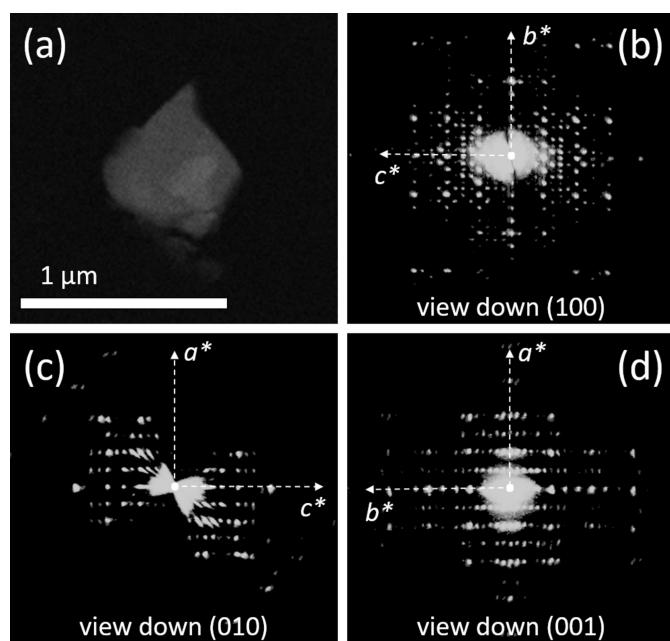
Figure 5
(a) HAADF–STEM image of a typical crystal of compound VI, used for 3D ED data acquisition. (b)/(c)/(d) Reconstructed 3D ED data viewed along the main crystallographic directions.

hybrid-pixel single-electron detector is crucial for acquiring a diffraction pattern with a satisfactory signal-to-noise ratio.

3D ED acquisitions were performed rotating the sample around the TEM goniometer axis in steps of $1°$, in total tilt ranges up to $120°$ (Figs. 4–6). The exposure time per frame was 1 s. The camera length was 180 mm, allowing a resolution in real space of up to 0.7 Å. During the experiment, the beam was precessed around the optical axis by an angle of $1°$. Precession was obtained using a Nanomegas Digistar P1000 device. After each tilt, the crystal position was tracked by STEM imaging and a diffraction pattern was acquired. All data acquisitions were performed at room temperature.

3D ED data were analysed using the software *PETS2.0* (Palatinus *et al.*, 2019) and *ADT3D* (Kolb *et al.*, 2011). Structure determination was carried out by standard direct methods (SDM) as implemented in the software *SIR2014* (Burla *et al.*, 2015), using a fully kinematical approximation (I_{hkl} proportional to $|F_{hkl}|^2$). Kinematical least-squares structure refinement was performed with the software *SHELXL* (Sheldrick, 2015) using electron atomic scattering factors. Structures were finally refined considering dynamical effects (Palatinus *et al.*, 2013, 2015) using the software *JANA2006* (Petrícek *et al.*, 2014) and assuming a simple platelet shape.

2.4. Computational details

Geometry relaxations of the crystal structures of compounds I, VI and VII, and of the single-molecule structure of compound I, were conducted using density functional theory (DFT) with the FHI-aims code (Blum *et al.*, 2009). The Perdew–Burke–Ernzerhof (PBE) (Perdew *et al.*, 1996, 1997) generalized gradient approximation was combined with the Tkatchenko–Scheffler (TS) (Tkatchenko & Scheffler, 2009) pairwise dispersion method. Tight numerical settings and tier 2 basis sets were used. Full unit-cell relaxation was performed until no force component on any atom exceeded 0.01 eV Å^{-1}.

GW+BSE calculations for the crystal structures of compounds I, VI and VII were performed using the BerkeleyGW code (Deslippe *et al.*, 2012). *Quantum ESPRESSO* (Giannozzi, 2009) was used to compute the mean-field eigenvectors and eigenvalues, and to generate the mean-field coarse-grid and fine-grid wave functions with the PBE exchange-correlation functional. For compound I, we used a coarse k-grid of $8 \times 2 \times 1$ and a fine k-grid of $8 \times 4 \times 2$. For compound VI, we used a coarse k-grid of $4 \times 1 \times 1$ and a fine k-grid of $8 \times 2 \times 2$. For compound VII, we used a coarse k-grid of $2 \times 2 \times 2$ and a fine k-grid of $4 \times 4 \times 4$. Troullier–Martins norm-conserving pseudopotentials (Troullier & Martins, 1991) were used and the kinetic energy cut-off was set at 50 Ry. About 550 unoccupied bands were included in the GW calculation. The static remainder correction was applied to accelerate the convergence with respect to the number of unoccupied states (Deslippe *et al.*, 2013). The polarizability, inverse dielectric matrix and GW self-energy operator were constructed based on the mean-field eigenvalues and eigenfunctions using the coarse k-point settings. Optical properties, including excitation energies, exciton wave functions and absorption spectra were

calculated by solving the BSE within the Tamm–Dancoff approximation (TDA). 24 valence bands and 24 conduction bands were included in the BSE calculation (Figs. S11–S16). Taking the full dielectric matrix as input to screen the attraction between the electron (e) and hole (h), the e–h interaction kernel was constructed on the coarse k-point grid. To construct the Bethe–Salpeter Hamiltonian, the GW quasiparticle energies and e–h interaction kernel calculated with coarse k-point settings were interpolated onto the fine k-point grid. The subsequent diagonalization yielded the excitation energies and wave functions. The exciton wave functions of compounds I and VI were converged using supercells of $8 \times 4 \times 2$ and $8 \times 2 \times 2$, respectively, based on the criterion proposed in Liu $et\ al.$ (2020c). The degree of singlet exciton charge-transfer character (%CT) was calculated by double-Bader analysis (DBA) (Wang $et\ al.$, 2018; Liu $et\ al.$, 2020c). The results for rubrene and pentacene are from Wang $et\ al.$ (2016), the results for quaterrylene, perylene and tetracene are from Wang $et\ al.$ (2018), the results for anthracene are from Liu $et\ al.$ (2020c), and the results for terrylene are from Hall $et\ al.$ (2021).

GW+BSE calculations for the single molecule of compound I were conducted using the FHI-aims code, using augmented tier-2 basis sets (Liu $et\ al.$, 2020a). PBE was used as the DFT starting point. A detailed account of the GW implementation in FHI-aims is provided in Ren $et\ al.$ (2012) and Caruso $et\ al.$ (2013). Briefly, the self-energy is first calculated on the imaginary frequency axis and then analytically continued to the real frequency axis. A 16-parameter Pade approximation

was used in the analytical continuation (Ren $et\ al.$, 2012; Van Setten $et\ al.$, 2015). Using the GW quasiparticle energies, BSE calculations were performed to obtain the singlet and triplet excitation energies (Liu $et\ al.$, 2020c). Results for the known TTA chromophores, which we compare to compound I here, are from Wang $et\ al.$ (2020), in which the same methodology was used. For GW calculations of the gas-phase PE spectra of I, VI and VII, we used the PBE-based hybrid functional (PBE0) (Adamo & Barone, 1999) as the mean-field starting point, with tight numerical settings and tier 4 basis sets. This method was benchmarked in Marom $et\ al.$ (2012) and Knight $et\ al.$ (2016), and shown to yield spectra in good agreement with experiments for a variety of organic compounds. For compounds VI and VII, only calculated GW@PBE0 spectra are presented, in the absence of experimental data.

3. Results and discussion

3.1. Benzo[e]dinaphtho[2,3-a;1′,2′,3′,4′-ghi]fluoranthene

The molecular conformation of compound I was reported as provisional (Clar $et\ al.$, 1964). New absorption measurements (Fig. S3) match well with those reported by Clar $et\ al.$ (1964), apart from the spurious long-wavelength band at 584 nm. The alleged composition from mass spectrometry was $C_{34}H_{18}$ (Fig. S8), also consistent with the related UV spectrum (Fig. 3(a) and Table S1). The occurrence of cyclodehydrogenation, with the formation of a five-membered ring, is accompanied by a red shift of the first UV band by 74.5 nm. In particular, the sharp band at about 440 nm is often found in fluoranthene-type PAHs, and never in alternant $cata$- or $peri$-condensed PAHs. The strongest argument in favour of Clar's assignment comes from the PE spectrum, which fits the calculated one (Clar $et\ al.$, 1981). Also, our GW@PBE0 calculation, based on the structure determination by 3D ED data, is in excellent agreement with experimental PE data (Fig. 3(b)). We note that the computed ionization energies are vertical values. We do not consider the relaxation of the atomic positions in response to the electronic excitation, vibrational effects, scattering cross-section effects, detector resolution and other sources of noise in the experiment, which may affect the peak breadth and intensity. The comparison to experiment is focused on peak positions, with the dashed lines at the experimental peak maxima serving as a guide to the eye. Additionally, this spectrum shows an impurity band at 6.95 eV. This peak is not related to the presence of III, which should instead give a band at about 6.67 eV (Clar $et\ al.$, 1981).

3D ED data were recorded from three flat crystal fragments of I, with sizes less than 1 µm (Fig. 4(a)). All data sets were consistent with a primitive orthorhombic unit cell, with parameters $a = 5.1\,(1)$, $b = 17.7\,(4)$ and $c = 23.2\,(5)$ Å (Figs. 4(b)–(d)). Such a cell would likely contain four molecules. Upon inspection of the reciprocal space reconstruction, the extinction rules 0k0: $k = 2n$ and 00l: $l = 2n$ were observed. All 3D ED acquisitions miss information about h00, because this direction is always parallel to the main surface of the platelet and, standing vertical, cannot be sampled inside the
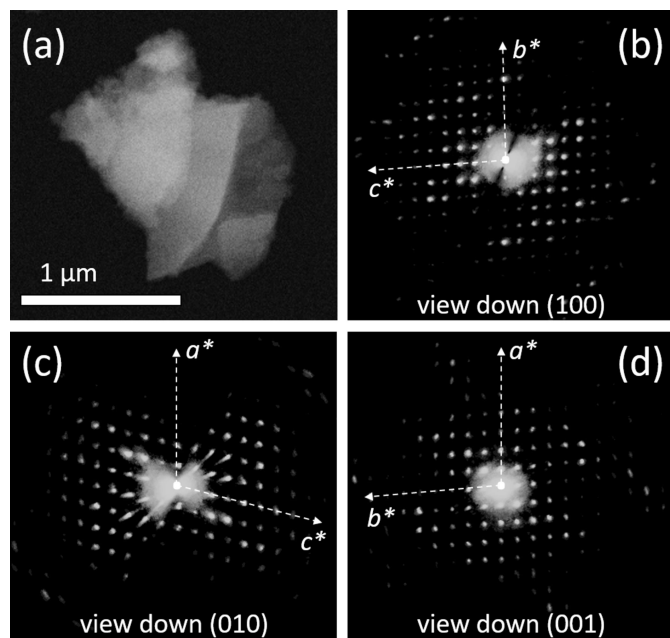
Figure 6
(a) HAADF–STEM image of a typical crystal of compound VII, used for 3D ED data acquisition. (b)/(c)/(d) Reconstructed 3D ED data viewed along the main crystallographic directions. Note, in the case of compound VII, the HAADF image shows more than one single crystal, while the 3D ED data were collected from a single-crystal area (upper right) that appear as thin as the other two compounds.

TEM goniometer range. The possible space groups for compound I were therefore $P22_12_1$ (No. 18) or $P2_12_12_1$ (No. 19), both with multiplicity 4.

Structure solution was performed by SDM using the 3D ED data set with the highest angular range and the greatest resolution. The solution eventually converged in the space group $P2_12_12_1$. All but one of the 34 non-H atoms of the molecule were located *ab initio*. The missing C and all H atoms were generated during least-squares refinement against 3D ED data imposing constraints on the aromatic rings. More details about structure determination and refinement are reported in Table S5.

The structure determination of compound I allowed for unequivocal confirmation of the molecular model inferred from spectroscopic data, which was previously considered provisional (Clar *et al.*, 1964). Molecules of I are flat and arranged in four columns per unit cell, which extend along *a*. Inside each column, molecules form a herringbone stack, parallel to one of the four planes of the {3,10,6} family (Figs. 7(*a*) and 7(*b*)). The intermolecular distance is about 3.5 Å, similar to that observed in terrylene, another PAH recently solved by the 3D ED method and potentially promising for electronic applications (Hall *et al.*, 2021).

3.2. 7,14-Diphenylnaphtho[1,2,3,4-*cde*]bisanthene and 7,16-diphenylnaphtho[1,2,3,4-*cde*]helianthrene

The two side products VI and VII obtained during the synthesis of IV were separated by repeated gradient sublimation. Compound VI appears as dark-green crystals that do not melt up to 427 °C and shows a deep-red colour in concentrated benzene solution or a carmine colour in dilute solution. While exposed to UV light, it appears violet–blue and has an intense orange–red fluorescence in daylight. Based on mass spectrometry (Fig. S9) and combustion data, VI has the composition $C_{46}H_{24}$. The molecular conformation could not be deduced by spectroscopic data alone and no crystal produced was sufficiently large for single-crystal XRD analysis. In the absence of experimental PE data, we present here the computed GW@PBE0 spectra of compounds VI and VII (Figs. 3(*d*) and 3(*f*)).

3D ED data collected on three VI micrometric crystals (Fig. 5(*a*)) consistently indicated a primitive orthorhombic cell with parameters $a = 9.9$ (2), $b = 26.5$ (5) and $c = 20.7$ (4) Å (Figs. 5(*b*)–(*d*)). Such a cell would likely contain eight VI molecules. Upon inspection of the reciprocal space reconstruction, the extinction rules uniquely indicated the space group *Pbca* (No. 61), with multiplicity 8. Despite the remarkable complexity of the molecule, its structure solution was obtained *ab initio* by SDM on the basis of 3D ED data. All 46 non-H atoms of the molecule were located automatically in the first Fourier map, eventually revealing the molecular conformation that could not be obtained by spectroscopic data (Fig. 7(*c*)). More details about the structure determination and refinement are reported in Table S5.

Molecules of VI are flat and pack as a sandwich herringbone structure parallel to the (102) and (10$\bar{2}$) planes

(Fig. 7(*d*)). The phenyl groups act as nonplanar satellites of the molecule, forming with it an angle of about 66°. The intermolecular distance is about 3.5 Å.

Apparently, during the synthesis, phenyl–Li acts as a masked benzyne and connects to the reactive diene positions. A literature search supports the 3D ED result. In fact, the dimesityl analogue of VI was synthesized in an unequivocal manner by benzyne addition to 7,14-bis(mesityl)bisanthene (Konishi *et al.*, 2013), when benzyne was liberated *in situ* from *o*-anthranilic acid and isoamyl nitrite. Due to severe twisting about the formal single bond, the phenyl and mesityl groups show little difference; hence, the UV spectra reported by Konishi *et al.* (2013) agrees with that obtained for compound VI (Fig. 3(*c*)). Compared to IV, lateral annelation shifts the first UV band by 79 nm to the blue (from 684.5 to 613 nm), due to the gain of one Clar sextet. There is a good mirror relationship between the absorption and fluorescence emission, and the Stokes shift is 6 nm.

Compound VII results in ochre crystals with a melting point of 426–427 °C and shows a permanganate colour in solution with benzene or cyclohexane. Under UV light it emits a green fluorescence and is moderately stable in daylight. This feature contrasts with the extreme sensitivity of helianthrene, whose fluorescence fades in daylight within seconds (Seip & Brauer,
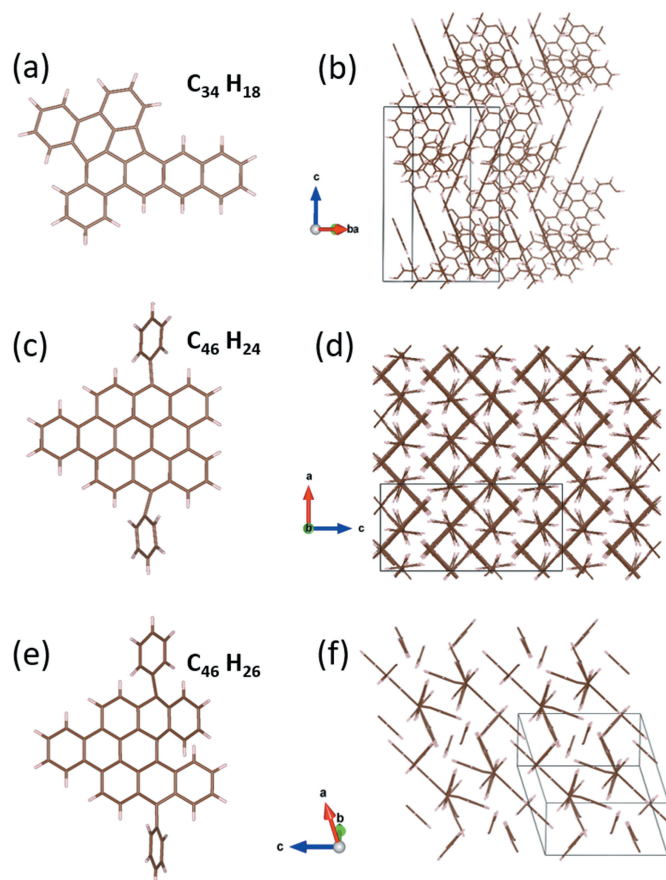
Figure 7
The molecular configuration and crystalline structure of (*a*)/(*b*) compound I, (*c*)/(*d*) compound VI and (*e*)/(*f*) compound VII. C atoms are shown in brown and H atoms are shown in light grey.

1992). The combustion and mass spectrometry (Fig. S10) analysis pointed to a composition of $C_{46}H_{26}$. As for compound VI, the molecular conformation could not be deduced by spectroscopic data and no crystal was sufficiently large for single-crystal XRD analysis.

3D ED data were recorded from six crystal fragments with a size of less than 1 μm (Fig. 6(a)). All of them delivered a triclinic unit cell with parameters $a = 10.5$ (2), $b = 11.6$ (3), $c = 12.8$ (3) Å, $\alpha = 85.3$ (5), $\beta = 76.1$ (5) and $\gamma = 84.9$ (5)° (Figs. 6(b)–(d)). No possibility of cell centring was envisaged, and no extinction rule pointing to higher symmetry was recognised. Such a cell would conveniently host only two VII molecules. Structure solution of this compound was indeed obtained *ab initio* by SDM in the space group $P\bar{1}$ (No. 2). All 46 non-H atoms of the molecule were automatically spotted in the first Fourier map, allowing also, in this case, the molecular conformation to be established, which could not be obtained by spectroscopic data (Fig. 7(e)). More details about the structure determination and refinement are reported in Table S5.

The two terminal phenyl rings in VII act again as nonplanar satellites, forming an angle of about 65° with the main part of the molecule. Their presence partially breaks the aromaticity of the helianthrene group. The nonplanarity of VII allows for both C—H⋯π and π–π interactions, leading to the reduction of symmetry down to triclinic. No herringbone motif is present (Fig. 7(f)).

In the corresponding UV spectrum, a first sharp band was observed at 501 nm (Fig. 3(e)). Consistent with the 3D ED results, in going from V to VII, there is a blue shift by 79 nm due to the gain of one Clar sextet in the molecule. The Stokes shift is 16.5 nm (Fig. S7, Tables S2 and S3).

3.3. Electronic properties

Having definitively resolved the molecular and crystal structures of compounds I, VI and VII by 3D ED, we are now able to use computer simulations to assess their potential usefulness for photovoltaic applications. In particular, we are interested in discovering new materials capable of undergoing singlet fission (SF), the down-conversion of one singlet exciton into two triplet excitons (Smith & Michl, 2010, 2013; Monahan & Zhu, 2015; Michl, 2019), and its reverse process, triplet–triplet annihilation (TTA), the up-conversion of two triplet excitons into one singlet exciton (Baldo *et al.*, 2000; Schmidt & Castellano, 2014; Schulze & Schmidt, 2015). Both of these processes can be harnessed to reduce losses and thus increase the conversion efficiency of solar cells. In a typical solar cell, the absorption of one photon produces one charge carrier. The absorption of a high-energy photon generates a highly excited singlet exciton, which subsequently thermalizes to the lowest excited state, such that the excess photon energy is lost to heat. SF can be potentially utilized to convert that excess photon energy into an additional charge carrier by generating two triplet excitons from one high-energy singlet exciton. Another source of losses in solar cells is that photons with energies below the gap of the absorber cannot be absorbed

and their energy is lost. TTA can be utilized to convert two low-energy photons into an additional charge carrier. Hence, supplementing the traditional absorber with SF and TTA materials can significantly enhance the efficiency of solar cells by reducing losses at both the high end and the low end of the solar photon energy spectrum. SF and TTA materials are rare because few chromophores meet the stringent requirements for the excited-state energetics, as detailed below. Most of the known and predicted SF and TTA chromophores are PAHs, mainly acene and perylene derivatives. Therefore, we evaluate the PAHs studied here for these purposes.

To assess the potential of molecular crystals to undergo SF in the solid state, we consider a two-dimensional descriptor (Wang *et al.*, 2018; Liu *et al.*, 2020b,c), as shown in Fig. 8. The primary descriptor, displayed on the *x* axis, is the thermodynamic driving force for SF, which is the difference between the singlet exciton energy and twice the triplet exciton energy ($E_s − 2E_t$). A high driving force indicates that a material is likely to undergo SF with a high rate. However, an overly high driving force would lead to losses in solar energy conversion. Therefore, it has been suggested that materials with $E_s \simeq 2E_t$ may be preferable (Wu *et al.*, 2014). Owing to the approximations used in GW+BSE calculations, the values of $E_s − 2E_t$ are systematically underestimated. Hence, we restrict the discussion to qualitative comparisons between materials.

The secondary descriptor, displayed on the *y* axis, is the degree of charge-transfer character (%CT) of the singlet exciton wave function. Exciton wave functions have two spatial variables corresponding to the electron and hole probability distributions. If the hole is located on one molecule, the degree of charge-transfer character corresponds to the probability of finding the electron on other molecules. This descriptor is motivated by the growing body of experimental
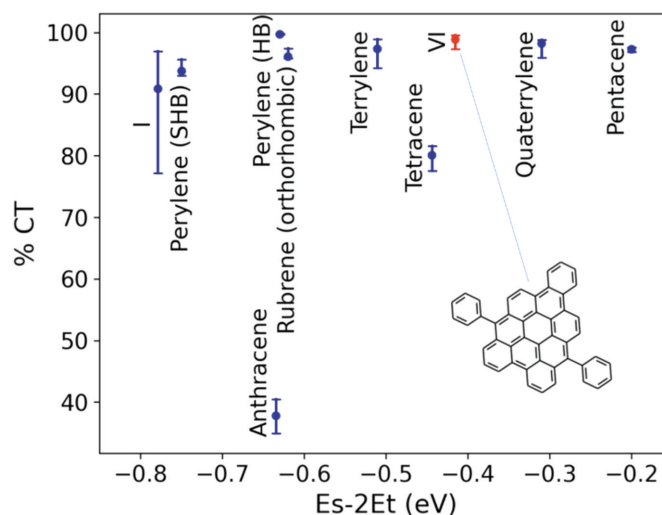
Figure 8
A two-dimensional descriptor for assessing SF candidates. The thermodynamic driving force for SF ($E_s − 2E_t$) is displayed on the *x* axis and the singlet exciton charge-transfer character (%CT) is displayed on the *y* axis. The error bars correspond to the range of %CT values obtained by using different hole positions when performing the double-Bader analysis. Compounds I and VI (the latter highlighted in red) are compared to some acene and rylene phases.

evidence for the involvement of an intermediate charge-transfer state in the SF process (Monahan & Zhu, 2015; Chan *et al.*, 2013; Kim *et al.*, 2019; Miyata *et al.*, 2019; Margulies *et al.*, 2017). A singlet exciton with a high degree of charge-transfer character is thought to be favourable for SF (Monahan & Zhu, 2015; Sharifzadeh *et al.*, 2013, 2015; Broch *et al.*, 2018; Hart *et al.*, 2018). To evaluate %CT, we use double-Bader analysis (DBA), an extension of the Bader charge-partitioning scheme, to exciton wave functions with two spatial variables. The %CT is calculated by performing nested sums over the electron distributions obtained for different hole positions within a molecule. The error bars correspond to the range of %CT values obtained for different hole positions within the double-Bader analysis (Wang *et al.*, 2018; Liu *et al.*, 2020c).

In Fig. 8, compounds I and VI are compared to perylenes and acenes with respect to these two descriptors. Pentacene has the highest SF driving force of the materials shown here and is known to undergo fast SF with a high triplet yield (Chan *et al.*, 2011; Wilson *et al.*, 2011). SF in tetracene is experimentally known to be slightly endoergic (Tomkiewicz *et al.*, 1971; Grumstrup *et al.*, 2010; Burdett *et al.*, 2010; Chan *et al.*, 2012; Burdett & Bardeen, 2012, 2013; Arias *et al.*, 2016). Compound VI has a somewhat higher SF driving force, as well as a higher singlet exciton CT character than tetracene. Based on this, it may be a promising candidate for intermolecular SF in the solid state. Compound I has a lower SF driving force

than anthracene and perylene, whose derivatives are known to undergo TTA (Renaud *et al.*, 2013; Jiang *et al.*, 2013; Eaton *et al.*, 2013; Mirjani *et al.*, 2014; Renaud & Grozema, 2015; Le *et al.*, 2016; Würthner *et al.*, 2016). Therefore, it may be able to undergo TTA. The computed singlet and triplet excitation energies of compound VII are 2.90 and 2.00 eV, respectively. It is thus not a likely candidate for either SF or TTA; therefore, it is not shown in Figs. 8 and 9.

To further assess the potential of compound I to undergo TTA, we compare its excited-state energetics as an isolated molecule to several chromophores experimentally known to undergo TTA (Fig. 9). The comparison is based on the criteria proposed in Wang *et al.* (2020). The primary criterion for a chromophore to undergo TTA is that the energy release in the singlet pathway, $E_{\mathrm{TTA,S}}^{\mathrm{loss}} = 2T_1 - S_1$, plotted on the x axis, should be positive to provide thermodynamic driving force, but not overly large to avoid significant energy losses. As discussed in detail in Wang *et al.* (2020), GW+BSE@PBE systematically underestimates $E_{\mathrm{TTA,S}}^{\mathrm{loss}}$. Therefore, we assess new chromophores based on a comparison with known TTA chromophores, similar to our procedure for assessing SF candidates. In rubrene, TTA is experimentally known to be approximately isoergic (Cheng *et al.*, 2010). Therefore, we consider the $E_{\mathrm{TTA,S}}^{\mathrm{loss}}$ of rubrene as the lower limit for the singlet pathway to be open. Molecules with a smaller $E_{\mathrm{TTA,S}}^{\mathrm{loss}}$ than rubrene may be more likely to undergo SF than TTA. This is indicated by the left dashed vertical line in Fig. 9. Of the experimentally known TTA chromophores calculated in Wang *et al.* (2020), pyrene has the highest $E_{\mathrm{TTA,S}}^{\mathrm{loss}}$. We then consider pyrene as the upper limit, above which the energy loss is too high for efficient TTA. This is indicated by the right vertical dashed line in Fig. 9.

The main process competing with TTA is the combination of two triplet excitons into a higher-energy triplet state, T_2, which decays nonradiatively. Hence, a secondary criterion for the quantum yield of TTA to be as high as possible is for the energy release in the triplet pathway, $E_{\mathrm{TTA,T}}^{\mathrm{loss}} = 2T_1 - T_2$, to be as small as possible compared to $E_{\mathrm{TTA,S}}^{\mathrm{loss}}$. In Wang *et al.* (2020), we have found that most of the experimentally known TTA chromophores fall into the range between $E_{\mathrm{TTA,T}}^{\mathrm{loss}} = E_{\mathrm{TTA,S}}^{\mathrm{loss}}$ and $E_{\mathrm{TTA,T}}^{\mathrm{loss}} = E_{\mathrm{TTA,S}}^{\mathrm{loss}} + 0.32$ eV. This is indicated by the two diagonal lines in Fig. 9. Compound I (highlighted in red) is well within this range, between perylene and an anthracene derivative. Therefore, based on energetic considerations, compound I may be a promising candidate for TTA. We note, however, that favourable energetics are a necessary but not sufficient condition for good TTA chromophores (Wang *et al.*, 2020).
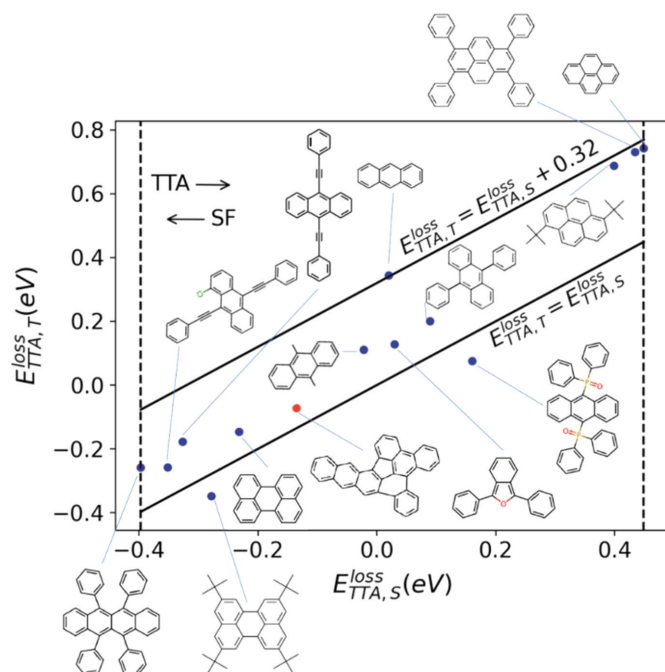
Figure 9
Assessment of compound I (highlighted in red) as a TTA candidate, compared to several known TTA chromophores. The energy release in the singlet pathway, $E_{\mathrm{TTA,S}}^{\mathrm{loss}} = 2T_1 - S_1$, is plotted on the x axis and the energy release in the competing triplet pathway, $E_{\mathrm{TTA,T}}^{\mathrm{loss}} = 2T_1 - T_2$, is plotted on the y axis. The dashed line on the left correspond to the $E_{\mathrm{TTA,S}}^{\mathrm{loss}}$ of rubrene and the dashed line on the right corresponds to the $E_{\mathrm{TTA,S}}^{\mathrm{loss}}$ of pyrene. The region delineated by the two diagonal lines is where most experimentally known TTA chromophores are found in Wang *et al.* (2020).

4. Conclusions

Organic syntheses typically yield several side products, which are generally discarded even if their quantity amounts to 20–50% of the total. Such side products are often unwanted. Moreover, the assessment of their potential use is complicated by the difficulty of their characterization. Spectroscopic data may be insufficient to unambiguously determine the molecular

conformation, especially when molecules are complex or the presence of more than one species is suspected. Powder X-ray diffraction requires a sufficient amount of material and may suffer from peak overlap, in particular when the phase of interest cannot be properly purified, its unit-cell parameters are long and/or small crystal size causes severe peak broadening. In these cases, 3D electron diffraction emerges as a promising technique for conclusive determination of the molecular conformation and crystal structure, as we have demonstrated here for benzo[*e*]dinaphtho[2,3-*a*;1′,2′,3′,4′-*ghi*]-fluoranthene (compound I), 7,14-diphenyl-naphtho[1,2,3,4-*cde*]bisanthene (compound VI) and 7,16-diphenylnaphtho-[1,2,3,4-*cde*]helianthrene (compound VII).

Once the structure is determined, computer simulations may be used to predict the electronic and optical properties of side products, in particular when the amount of material initially produced is insufficient for detailed experimental characterization. This can help to assess the potential usefulness of side products for various applications. In some cases, unintended side products may turn out to be as useful as the main reaction products. For example, the second system analysed here comprises four related molecules. Of these, IV exhibits exceptional optical properties (Gorelenko *et al.*, 1977) and V found practical application as a pigment for visible-light-sensitive actinometers (Brauer *et al.*, 1983). Based on the many-body perturbation theory simulations presented here, compounds I and VI may be useful for photovoltaic applications. Compound I may exhibit triplet–triplet annihilation, which may enable harvesting photons with energies below the gap of the absorber in a solar cell. Compound VI may exhibit intermolecular singlet fission in the solid state, which may enable the harvesting of two charge carriers from one high-energy photon in a solar cell. Compound VII is a wide-gap insulator.

We hope that the results reported in this article will inspire others to pursue further characterization of organic side products by 3D ED and computer simulations. This may lead to the discovery of potentially useful compounds for various applications. Thus, 3D ED is a promising new avenue for enriching our knowledge of organic synthetic routes and exploiting side products that would otherwise be discarded.

References

Adamo, C. & Barone, V. (1999). *J. Chem. Phys.* **110**, 6158–6170.

Andrusenko, I., Hamilton, V., Lanza, A. E., Hall, C. L., Mugnaioli, E., Potticary, J., Buanz, A., Gaisford, S., Piras, A. M., Zambito, Y., Hall, S. R. & Gemmi, M. (2021). *Int. J. Pharm.* **608**, 121067.

Andrusenko, I., Hamilton, V., Mugnaioli, E., Lanza, A., Hall, C., Potticary, J., Hall, S. R. & Gemmi, M. (2019). *Angew. Chem.* **131**, 11035–11038.

Anschütz, R. (1886). *Justus Liebigs Ann. Chem.* **235**, 299–341.

Arabei, S. M. & Pavich, T. A. (2004). *J. Appl. Spectrosc.* **71**, 187–193.

Arias, D. H., Ryerson, J. L., Cook, J. D., Damrauer, N. H. & Johnson, J. C. (2016). *Chem. Sci.* **7**, 1185–1191.

Baldo, M. A., Adachi, C. & Forrest, S. R. (2000). *Phys. Rev. B*, **62**, 10967–10977.

Benner, G. & Probst, W. (1994). *J. Microsc.* **174**, 133–142.

Blase, X., Duchemin, I., Jacquemin, D. & Loos, P. F. (2020). *J. Phys. Chem. Lett.* **11**, 7371–7382.

Blum, V., Gehrke, R., Hanke, F., Havu, P., Havu, V., Ren, X., Reuter, K. & Scheffler, M. (2009). *Comput. Phys. Commun.* **180**, 2175–2196.

Brauer, H.-D., Schmidt, R., Gauglitz, G. & Hubig, S. (1983). *Photochem. Photobiol.* **37**, 595–598.

Brázda, P., Palatinus, L. & Babor, M. (2019). *Science*, **364**, 667–669.

Broch, K., Dieterle, J., Branchi, F., Hestand, N. J., Olivier, Y., Tamura, H., Cruz, C., Nichols, V. M., Hinderhofer, A., Beljonne, D., Spano, F. C., Cerullo, G., Bardeen, C. J. & Schreiber, F. (2018). *Nat. Commun.* **9**, 954.

Bruhn, J. F., Scapin, G., Cheng, A., Mercado, B. Q., Waterman, D. G., Ganesh, T., Dallakyan, S., Read, B. N., Nieusma, T., Lucier, K. W., Mayer, M. L., Chiang, N. J., Poweleit, N., McGilvray, P. T., Wilson, T. S., Mashore, M., Hennessy, C., Thomson, S., Wang, B., Potter, C. S. & Carragher, B. (2021). *Front. Mol. Biosci.* **8**, 648603.

Buchlovič, M., Kříž, Z., Hofr, C. & Potáček, M. (2013). *Bioorg. Med. Chem.* **21**, 1078–1081.

Burdett, J. J. & Bardeen, C. J. (2012). *J. Am. Chem. Soc.* **134**, 8597–8607.

Burdett, J. J. & Bardeen, C. J. (2013). *Acc. Chem. Res.* **46**, 1312–1320.

Burdett, J. J., Müller, A. M., Gosztola, D. & Bardeen, C. J. (2010). *J. Chem. Phys.* **133**, 144506.

Burla, M. C., Caliandro, R., Carrozzini, B., Cascarano, G. L., Cuocci, C., Giacovazzo, C., Mallamo, M., Mazzone, A. & Polidori, G. (2015). *J. Appl. Cryst.* **48**, 306–309.

Caruso, F., Rinke, P., Ren, X., Rubio, A. & Scheffler, M. (2013). *Phys. Rev. B*, **88**, 075105.

Castro, K. P., Bukovsky, E. V., Kuvychko, I. V., DeWeerd, N. J., Chen, Y.-S., Deng, S. H. M., Wang, X.-B., Popov, A. A., Strauss, S. H. & Boltalina, O. V. (2019). *Chem. Eur. J.* **25**, 13547–13565.

Chan, W.-L., Berkelbach, T. C., Provorse, M. R., Monahan, N. R., Tritsch, J. R., Hybertsen, M. S., Reichman, D. R., Gao, J. & Zhu, X.-Y. (2013). *Acc. Chem. Res.* **46**, 1321–1329.

Chan, W.-L., Ligges, M., Jailaubekov, A., Kaake, L., Miaja-Avila, L. & Zhu, X.-Y. (2011). *Science*, **334**, 1541–1545.

Chan, W.-L., Ligges, M. & Zhu, X.-Y. (2012). *Nat. Chem.* **4**, 840–845.

Cheng, Y. Y., Fückel, B., Khoury, T., Clady, R. G. C. R., Tayebjee, M. J. Y., Ekins-Daukes, N. J., Crossley, M. J. & Schmidt, T. W. (2010). *J. Phys. Chem. Lett.* **1**, 1795–1799.

Clar, E., Guye-Vuillème, J. F. & Stephen, J. F. (1964). *Tetrahedron*, **20**, 2107–2117.

Clar, E., Robertson, J. M., Schloegl, R. & Schmidt, W. (1981). *J. Am. Chem. Soc.* **103**, 1320–1328.

Clar, E. & Schmidt, W. (1975). *Tetrahedron*, **31**, 2263–2271.

Clar, E. & Schmidt, W. (1976). *Tetrahedron*, **32**, 2563–2566.

Clar, E. & Schmidt, W. (1979). *Tetrahedron*, **35**, 2673–2680.

Cui, P., Svensson Grape, E., Spackman, P. R., Wu, Y., Clowes, R., Day, G. M., Inge, A. K., Little, M. A. & Cooper, A. I. (2020). *J. Am. Chem. Soc.* **142**, 12743–12750.

Deslippe, J., Samsonidze, G., Jain, M., Cohen, M. L. & Louie, S. G. (2013). *Phys. Rev. B*, **87**, 1–6.

Deslippe, J., Samsonidze, G., Strubbe, D. A., Jain, M., Cohen, M. L. & Louie, S. G. (2012). *Comput. Phys. Commun.* **183**, 1269–1289.

Eaton, S. W., Shoer, L. E., Karlen, S. D., Dyar, S. M., Margulies, E. A., Veldkamp, B. S., Ramanan, C., Hartzler, D. A., Savikhin, S., Marks, T. J. & Wasielewski, M. R. (2013). *J. Am. Chem. Soc.* **135**, 14701–14712.

Fetzer, J. C. (2000). *Chemical Analysis, A Series of Monographs on Analytical Chemistry and Its Applications*, Vol. 158, *Large (C> = 24) Polycyclic Aromatic Hydrocarbons: Chemistry and Analysis*, series edited by J. D. Winefordner. New York: Wiley Interscience.

Fetzer, J. C. (2007). *Polycycl. Aromat.. Compd.* **27**, 143–162.

Fort, E. H. (2010). Doctoral dissertation, Boston College University, Newton, USA.

Friedel, C. & Crafts, J. M. (1877). *Comptes Rendus*, **84**, 1392–1395.

Gemmi, M., Mugnaioli, E., Gorelik, T. E., Kolb, U., Palatinus, L., Boullay, P., Hovmöller, S. & Abrahams, J. P. (2019). *ACS Cent. Sci.* **5**, 1315–1329.

Giannozzi, P., Baroni, S., Bonini, N., Calandra, M., Car, R., Cavazzoni, C., Ceresoli, D., Chiarotti, G. L., Cococcioni, M., Dabo, I., Dal Corso, A., de Gironcoli, S., Fabris, S., Fratesi, G., Gebauer, R., Gerstmann, U., Gougoussis, C., Kokalj, A., Lazzeri, M., Martin-Samos, L., Marzari, N., Mauri, F., Mazzarello, R., Paolini, S., Pasquarello, A., Paulatto, L., Sbraccia, C., Scandolo, S., Sclauzero, G., Seitsonen, A. P., Smogunov, A., Umari, P. & Wentzcovitch, R. M. (2009). *J. Phys. Condens. Matter*, **21**, 395502.

Golze, D., Dvorak, M. & Rinke, P. (2019). *Front. Chem.* **7**, 377.

Gorelenko, A. Ya., Tolkachev, V. A. & Khalimanovich, D. M. (1977). *J. Appl. Spectrosc.* **26**, 710–712.

Grumstrup, E. M., Johnson, J. C. & Damrauer, N. H. (2010). *Phys. Rev. Lett.* **105**, 257403.

Gupte, A., Tripathi, A., Patel, H., Rudakiya, D. & Gupte, S. (2016). *Open Biotechnol. J.* **10**, 363–378.

Hall, C. L., Andrusenko, I., Potticary, J., Gao, S., Liu, X., Schmidt, W., Marom, N., Mugnaioli, E., Gemmi, M. & Hall, S. R. (2021). *ChemPhysChem*, **22**, 1631–1637.

Haritash, A. K. & Kaushik, C. P. (2009). *J. Hazard. Mater.* **169**, 1–15.

Hart, S. M., Silva, W. R. & Frontiera, R. R. (2018). *Chem. Sci.* **9**, 1242–1250.

Jiang, H., Zhang, K. K., Ye, J., Wei, F., Hu, P., Guo, J., Liang, C., Chen, X., Zhao, Y., McNeil, L. E., Hu, W. & Kloc, C. (2013). *Small*, **9**, 990–995.

Jones, C. G., Asay, M., Kim, L. J., Kleinsasser, J. F., Saha, A., Fulton, T. J., Berkley, K. R., Cascio, D., Malyutin, A. G., Conley, M. P., Stoltz, B. M., Lavallo, V., Rodríguez, J. A. & Nelson, H. M. (2019). *ACS Cent. Sci.* **5**, 1507–1513.

Kaiukov, R., Almeida, G., Marras, S., Dang, Z., Baranov, D., Petralanda, U., Infante, I., Mugnaioli, E., Griesi, A., De Trizio, L., Gemmi, M. & Manna, L. (2020). *Inorg. Chem.* **59**, 548–554.

Kapaca, E., Jiang, J., Cho, J., Jordá, J. L., Díaz-Cabañas, M. J., Zou, X., Corma, A. & Willhammar, T. (2021). *J. Am. Chem. Soc.* **143**, 8713–8719.

Kim, V. O., Broch, K., Belova, V., Chen, Y. S., Gerlach, A., Schreiber, F., Tamura, H., Della Valle, R. G., D'Avino, G., Salzmann, I., Beljonne, D., Rao, A. & Friend, R. (2019). *J. Chem. Phys.* **151**, 164706.

Knight, J. W., Wang, X., Gallandi, L., Dolgounitcheva, O., Ren, X., Ortiz, J. V., Rinke, P., Körzdörfer, T. & Marom, N. (2016). *J. Chem. Theory Comput.* **12**, 615–626.

Ko, S. H., Lee, T., Park, H., Ahn, D.-S., Kim, K., Kwon, Y., Cho, S. J. & Ryoo, R. (2018). *J. Am. Chem. Soc.* **140**, 7101–7107.

Kolb, U., Gorelik, T., Kübel, C., Otten, M. T. & Hubert, D. (2007). *Ultramicroscopy*, **107**, 507–513.

Kolb, U., Mugnaioli, E. & Gorelik, T. E. (2011). *Cryst. Res. Technol.* **46**, 542–554.

Konishi, A., Hirao, Y., Matsumoto, K., Kurata, H. & Kubo, T. (2013). *Chem. Lett.* **42**, 592–594.

Krotee, P., Griner, S. L., Sawaya, M. R., Cascio, D., Rodriguez, J. A., Shi, D., Philipp, S., Murray, K., Saelices, L., Lee, J., Seidler, P., Glabe, C. G., Jiang, L., Gonen, T. & Eisenberg, D. S. (2018). *J. Biol. Chem.* **293**, 2888–2902.

Krysiak, Y., Maslyk, M., Silva, B. N., Plana-Ruiz, S., Moura, H. M., Munsignatti, E. O., Vaiss, V. S., Kolb, U., Tremel, W., Palatinus, L., Leitão, A. A., Marler, B. & Pastore, H. O. (2021). *Chem. Mater.* **33**, 3207–3219.

Lanza, A., Margheritis, E., Mugnaioli, E., Cappello, V., Garau, G. & Gemmi, M. (2019). *IUCrJ*, **6**, 178–188.

Lawal, A. T. (2017). Cogent *Environ. Sci.* **3**, 1339841.

Le, A. K., Bender, J. A. & Roberts, S. T. (2016). *J. Phys. Chem. Lett.* **7**, 4922–4928.

Levine, A. M., Bu, G., Biswas, S., Tsai, E. H. R., Braunschweig, A. B. & Nannenga, B. L. (2020). *Chem. Commun.* **56**, 4204–4207.

Liu, C., Kloppenburg, J., Yao, Y., Ren, X., Appel, H., Kanai, Y. & Blum, V. (2020*a*). *J. Chem. Phys.* **152**, 044105.

Liu, X., Tom, R., Gao, S. & Marom, N. (2020*b*). *J. Phys. Chem. C*, **124**, 26134–26143.

Liu, X., Tom, R., Wang, X., Cook, C., Schatschneider, B. & Marom, N. (2020*c*). *J. Phys. Condens. Matter*, **32**, 184001.

Margulies, E. A., Logsdon, J. L., Miller, C. E., Ma, L., Simonoff, E., Young, R. M., Schatz, G. C. & Wasielewski, M. R. (2017). *J. Am. Chem. Soc.* **139**, 663–671.

Marom, N., Caruso, F., Ren, X., Hofmann, O. T., Körzdörfer, T., Chelikowsky, J. R., Rubio, A., Scheffler, M. & Rinke, P. (2012). *Phys. Rev. B*, **86**, 245127.

Maulding, D. (1970). *J. Org. Chem.* **35**, 1221–1223.

Michl, J. (2019). *Mol. Front. J.* **03**, 84–91.

Mirjani, F., Renaud, N., Gorczak, N. & Grozema, F. C. (2014). *J. Phys. Chem. C*, **118**, 14192–14199.

Miyata, K., Conrad-Burton, F. S., Geyer, F. L. & Zhu, X.-Y. (2019). *Chem. Rev.* **119**, 4261–4292.

Monahan, N. & Zhu, X.-Y. (2015). *Annu. Rev. Phys. Chem.* **66**, 601–618.

Nannenga, B. L. & Gonen, T. (2019). *Nat. Methods*, **16**, 369–379.

Nederlof, I., van Genderen, E., Li, Y.-W. & Abrahams, J. P. (2013). *Acta Cryst.* D**69**, 1223–1230.

Palatinus, L., Brázda, P., Jelínek, M., Hrdá, J., Steciuk, G. & Klementová, M. (2019). *Acta Cryst.* B**75**, 512–522.

Palatinus, L., Jacob, D., Cuvillier, P., Klementová, M., Sinkler, W. & Marks, L. D. (2013). *Acta Cryst.* A**69**, 171–188.

Palatinus, L., Petříček, V. & Corrêa, C. A. (2015). *Acta Cryst.* A**71**, 235–244.

Papi, F., Potticary, J., Lanza, A. E., Hall, S. R. & Gemmi, M. (2021). *Cryst. Growth Des.* **21**, 6341–6348.

Perdew, J. P., Burke, K. & Ernzerhof, M. (1996). *Phys. Rev. Lett.* **77**, 3865–3868.

Perdew, J. P., Burke, K. & Ernzerhof, M. (1997). *Phys. Rev. Lett.* **78**, 1396.

Petříček, V., Dušek, M. & Palatinus, L. (2014). *Z. Kristallogr. Cryst. Mater.* **229**, 345–352.

Rauhut, M. M., Roberts, B. G., Maulding, D. R., Bergmark, W. & Coleman, R. (1975). *J. Org. Chem.* **40**, 330–335.

Ren, X., Rinke, P., Blum, V., Wieferink, J., Tkatchenko, A., Sanfilippo, A., Reuter, K. & Scheffler, M. (2012). *New J. Phys.* **14**, 053020.

Renaud, N. & Grozema, C. (2015). *J. Phys. Chem. Lett.* **6**, 360–365.

Renaud, N., Sherratt, P. A. & Ratner, M. A. (2013). *J. Phys. Chem. Lett.* **4**, 1065–1069.

Saranya, G., Kolandaivel, P. & Senthilkumar, K. (2011). *J. Phys. Chem. A*, **115**, 14647–14656.

Sauvage, G. (1947*a*). *Ann. Chim.* **2**, 844–873.

Sauvage, G. (1947*b*). *Comptes Rendus*, **225**, 247–249.

Sawaya, M. R., Rodriquez, J., Cascio, D., Collazo, M. J., Shi, D., Reyes, F. E., Hattne, J., Gonen, T. & Eisenberg, D. S. (2016). *PNAS*, **113**, 11232–11236.

Schmidt, T. W. & Castellano, F. N. (2014). *J. Phys. Chem. Lett.* **5**, 4062–4072.

Scholl, R. & Mansfeld, J. (1910). *EurJIC*, **43**, 1734–1746.

Scholl, R. & Meyer, K. (1932). *Ber. Dtsch Chem. Ges. A/B*, **65**, 902–915.

Scholl, R., Seer, C. & Weitzenböck, R. (1910). *Ber. Dtsch Chem. Ges.* **43**, 2202–2209.

Schulze, T. F. & Schmidt, T. W. (2015). *Energy Environ. Sci.* **8**, 103–125.

Seip, M. & Brauer, H.-D. (1992). *J. Am. Chem. Soc.* **114**, 4486–4490.

Setten, M. J. van, Caruso, F., Sharifzadeh, S., Ren, X., Scheffler, M., Liu, F., Lischner, J., Lin, L., Deslippe, J. R., Louie, S. G., Yang, C., Weigend, F., Neaton, J. B., Evers, F. & Rinke, P. (2015). *J. Chem. Theory Comput.* **11**, 5665–5687.

Sharifzadeh, S., Darancet, P., Kronik, L. & Neaton, J. B. (2013). *J. Phys. Chem. Lett.* **4**, 2197–2201.

Sharifzadeh, S., Wong, C. Y., Wu, H., Cotts, B. L., Kronik, L., Ginsberg, N. S. & Neaton, J. B. (2015). *Adv. Funct. Mater.* **25**, 2038–2046.

Sheldrick, G. M. (2015). *Acta Cryst.* C**71**, 3–8.

Smith, M. B. & Michl, J. (2010). *Chem. Rev.* **110**, 6891–6936.

Smith, M. B. & Michl, J. (2013). *Annu. Rev. Phys. Chem.* **64**, 361–386.

Tkatchenko, A. & Scheffler, M. (2009). *Phys. Rev. Lett.* **102**, 073005.

Tomkiewicz, Y., Groff, R. P. & Avakian, P. (1971). *J. Chem. Phys.* **54**, 4504–4507.

Troullier, N. & Martins, J. L. (1991). *Phys. Rev. B*, **43**, 1993–2006.

Wan, W., Sun, J., Su, J., Hovmöller, S. & Zou, X. (2013). *J. Appl. Cryst.* **46**, 1863–1873.

Wang, X., Garcia, T., Monaco, S., Schatschneider, B. & Marom, N. (2016). *CrystEngComm*, **18**, 7353–7362.

Wang, X., Liu, X., Cook, C., Schatschneider, B. & Marom, N. (2018). *J. Chem. Phys.* **148**, 184101.

Wang, X., Tom, R., Liu, X., Congreve, D. N. & Marom, N. (2020). *J. Mater. Chem. C*, **8**, 10816–10824.

Warmack, R. A., Boyer, D. R., Zee, C.-T., Richards, L. S., Sawaya, M. R., Cascio, D., Gonen, T., Eisenberg, D. S. & Clarke, S. G. (2019). *Nat. Commun.* **10**, 3357.

Wilkes, H. (2010). *Methods of Hydrocarbon Analysis*, in *Handbook of Hydrocarbon and Lipid Microbiology*, edited by K. N. Timmis. Berlin, Heidelberg: Springer.

Wilson, M. W. B., Rao, A., Clark, J., Kumar, R. S. S., Brida, D., Cerullo, G. & Friend, R. H. (2011). *J. Am. Chem. Soc.* **133**, 11830–11833.

Wu, T. C., Thompson, N. J., Congreve, D. N., Hontz, E., Yost, S. R., Van Voorhis, T. & Baldo, M. A. (2014). *Appl. Phys. Lett.* **104**, 193901.

Würthner, F., Saha-Möller, C. R., Fimmel, B., Ogi, S., Leowanawat, P. & Schmidt, D. (2016). *Chem. Rev.* **116**, 962–1052.

Xu, H., Lebrette, H., Clabbers, M. T. B., Zhao, J., Griese, J. J., Zou, X. & Högbom, M. (2019). *Sci. Adv.* **5**, eaax4621.