Supplementary Material: Intrinsic Point Defect Tolerance in Selenium for Indoor and Tandem Photovoltaics

S1 Crystal Structure of *t*-Se

Table S1 Calculated lattice parameters of trigonal selenium, using various DFT functionals with and without dispersion corrections. Errors with respect to room temperature experimental values[44–46] are given as percentages

Functional	a (Å)	Δa (%)	c (Å)	Δc (%)	Δ Volume (%)
PBE	4.51	+3.1	5.05	+1.9	+8.5
PBE+D3	4.19	−4.1	5.13	+3.4	+4.8
PBEsol	4.03	−7.7	5.14	+3.5	−11.7
PBEsol+D3	3.96	−9.5	5.13	+3.4	−15.1
HSE06	4.86	+11.2	4.87	−1.8	+21.9
HSE06+D3	4.35	−0.5	4.96	0.0	−1.0
PBE0+D3	4.34	−0.6	4.95	−0.2	−1.3

From Table S1, we see that hybrid DFT functionals with van der Waals dispersion corrections (HSE06/PBE0+D3) give good agreement with experimental lattice parameters, while semi-local functionals give quite poor structural predictions here. The resultant calculated band gap is also decreased by ∼0.5 eV when using the HSE06+D3-relaxed cell with the HSE06+SOC functional, rather than the expanded HSE06-relaxed cell. Further analysis of the functional dependence of calculated structural parameters for *t*-Se is given in Stoliaroff *et al.*[47]

S2 Selenium Allotropes

Table S2 Lattice parameters and relative energies of various Se allotropes, calculated using hybrid DFT with dispersion corrections (HSE06+D3), neglecting finite temperature effects. Errors with respect to room temperature[a] experimental values are given as percentages

Allotrope	a (Å)	Δa (%)	b (Å)	Δb (%)	c (Å)	Δc (%)	Δ Volume (%)	ΔE (meV/atom)
Trigonal ($P3_121$)	4.35	−0.5	—	—	4.96	0.0	−1.0	6
(Experiment)[44–46]	4.37	—	—	—	4.95	—	—	—
Rhombohedral ($R\bar{3}$)	11.34	−0.2	—	—	4.41	+0.5	−0.9	20
(Experiment)[145]	11.36	—	—	—	4.43	—	—	—
β-monoclinic ($P2_1/c$)	9.42	+1.2	8.17	+1.2	13.02	+1.3	+3.8	12
(Experiment)[146]	9.31	—	8.07	—	12.85	—	—	—
γ-monoclinic ($P2_1/c$)	15.14	+0.8	14.68	−0.2	8.81	+0.2	+0.8	0
(Experiment)[147]	15.02	—	14.71	—	8.79	—	—	—
δ-monoclinic ($P2_1/c$)	9.28	+1.0	9.1	+1.4	14.76	+1.7	+6.0	8
(Experiment)[148]	9.04	—	8.97	—	14.52	—	—	—

[a] For δ-monoclinic Se the only available experimental measurement is at 150 K[148]

The favoured phase of Se depends on temperature, pressure and synthesis procedures (e.g. epitaxial growth on lattice-matched substrates). To assess the relative stability of the various crystalline structures, we compute the athermal formation energies of each low-energy ($\Delta E_{\text{hull}} \leq 0.2\,\text{eV/atom}$) experimentally-observed Se allotrope on the Materials Project database, using hybrid DFT including dispersion corrections (HSE06+D3) — as shown in Table S2.

As expected, van der Waals (vdW) dispersion interactions are found to impact the structural parameters and relative energies of these low-dimensional pure-covalent compounds, compressing unit cell volumes by $\sim 25\,\%$ on average (Table S3†).[47] The γ-monoclinic ($P2_1/c$ space group) polymorph is predicted to be the lowest energy phase, with t-Se ($P3_121$) being slightly higher in energy ($\Delta E = 6\,\mathrm{meV/atom}$) — neglecting temperature and pressure effects. Finite-temperature effects are expected to favour t-Se (having the highest symmetry of all Se allotropes), though we find that free energy contributions from *harmonic* phonons (Fig. S1 and Table S4†) do not result in a thermodynamic preference for t-Se over (monoclinic) γ-Se at elevated temperatures. Anharmonic contributions[149] or residual errors in the underlying DFT functional / dispersion corrections could also affect the predicted relative energies, and there could be a kinetic preference for t-Se chain growth during crystallisation — typically achieved via low-temperature annealing of amorphous Se.* The energy differences listed in Table S2 are small, all being less than thermal energy at room temperature ($k_BT \sim 25\,\mathrm{meV/particle}$). This indicates a possible difficulty in obtaining phase-pure samples of t-Se, with the potential for low-energy metastable Se chain rearrangements or ring-like formations in the bulk of the material — a point revisited in the discussion.

Table S3 Lattice parameters and relative energies of various allotropes of Selenium, calculated using hybrid DFT without dispersion corrections (HSE06), neglecting finite temperature effects. Errors with respect to room temperature[a] experimental values are given as percentages

Allotrope	a (Å)	Δa (%)	b (Å)	Δb (%)	c (Å)	Δc (%)	Δ Volume (%)	ΔE (meV/atom)
Trigonal ($P3_121$)	4.86	+11.2	—	—	4.87	−1.8	+21.9	21
(Experiment)[44–46]	4.37	—	—	—	4.96	—	—	—
Rhombohedral ($R\bar{3}$)	11.32	−0.4	—	—	5.51	+24.4	+23.5	35
(Experiment)[145]	11.36	—	—	—	4.43	—	—	—
β-monoclinic ($P2_1/c$)	10.49	+12.7	8.53	+5.7	13.73	+6.9	+26.7	9
(Experiment)[146]	9.31	—	8.07	—	12.85	—	—	—
γ-monoclinic ($P2_1/c$)	15.71	+4.6	15.17	+3.1	9.80	+11.5	+20.2	0
(Experiment)[147]	15.02	—	14.71	—	8.79	—	—	—
δ-monoclinic ($P2_1/c$)	10.41	+15.2	9.42	+5.0	15.32	+5.5	+28.9	10
(Experiment)[148]	9.04	—	8.97	—	14.52	—	—	—

[a] For δ-monoclinic Se the only available experimental measurement is at 150 K[148]

Phonon calculations were performed using `phonopy`[143] and `ThermoParser`[144] with the HSE06+D3 functional and a $4\times4\times4$ supercell of the primitive t-Se unit cell. No imaginary phonon modes were obtained with this $4\times4\times4$ supercell, while some small negative frequencies were obtained with a $3\times3\times3$ supercell.

Table S4 Vibrational contributions to the internal energy (E), entropy (S) and free energy (F) as computed using the harmonic phonons dispersions shown in Fig. S1, calculated using HSE06+D3, for t-Se ($P3_121$) and γ-monoclinic Se ($P2_1/c$), at temperatures of 0 K, 300 K and 1000 K

	t-Se ($P3_121$)			γ-monoclinic Se ($P2_1/c$)		
Temperature (K)	E (meV/atom)	S (meV/atom.K)	F (meV/atom)	E (meV/atom)	S (meV/atom.K)	F (meV/atom)
0 (Zero-Point Energy)	25.50	0	25.50	25.84	0	25.84
300	81.37	0.44	-49.74	81.69	0.45	-52.44
1000K	259.68	0.74	-482.87	259.78	0.75	-492.39

Similar results for the vibrational contributions to free energies of t-Se and monoclinic ring-like Se were reported by Stoliaroff *et al.*[47] using semi-local DFT, with monoclinic Se having lower-energy acoustic modes and thus more negative vibrational free energy contributions with increasing temperature.

* For instance, improved crystallisation quality and adhesion is obtained for t-Se when deposited on isomorphic t-Te.[17,18,25]

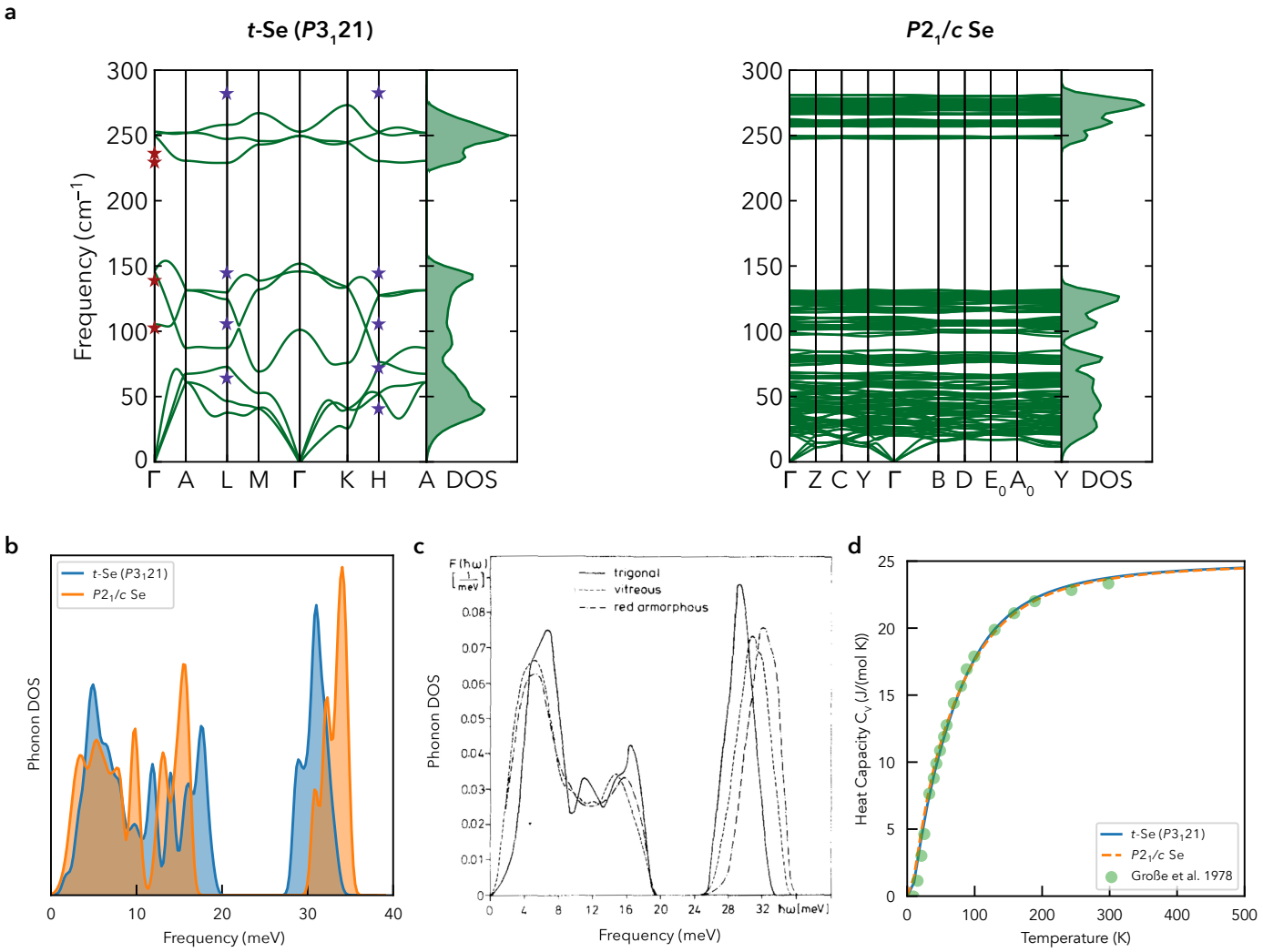

Fig. S1 Vibrational properties of trigonal ($P3_121$) and $P2_1/c$ Se. (a) Harmonic phonon dispersions for t-Se ($P3_121$) and γ-monoclinic Se ($P2_1/c$), calculated using HSE06+D3, alongside the corresponding phonon densities of states (DOS)(which have an additional Gaussian broadening of 0.1 THz) and experimental datapoints from Geick *et al.*[150]. The lower DOS peak for the optical modes corresponds to the Raman peak at $236\,\mathrm{cm}^{-1}$ measured by Lu *et al.*[24], while the prediction of higher energy optical modes for ring-like ($P2_1/c$) Se matches their measurements of higher energy Raman peaks for 'disordered' Se. (b) Phonon DOS of trigonal (t-Se) and monoclinic ($P2_1/c$) Se, calculated using HSE06+D3. (c) Phonon DOS of trigonal (t-Se), vitreous and red amorphous selenium as measured by inelastic neutron scattering. Reproduced with permission from Gompf[151]. (d) Calculated heat capacity as a function of temperature for trigonal ($P3_121$) and $P2_1/c$ Se, along with experimental datapoints for t-Se from Grosse *et al.*[152].

S3 Electronic Structure of *t*-Se

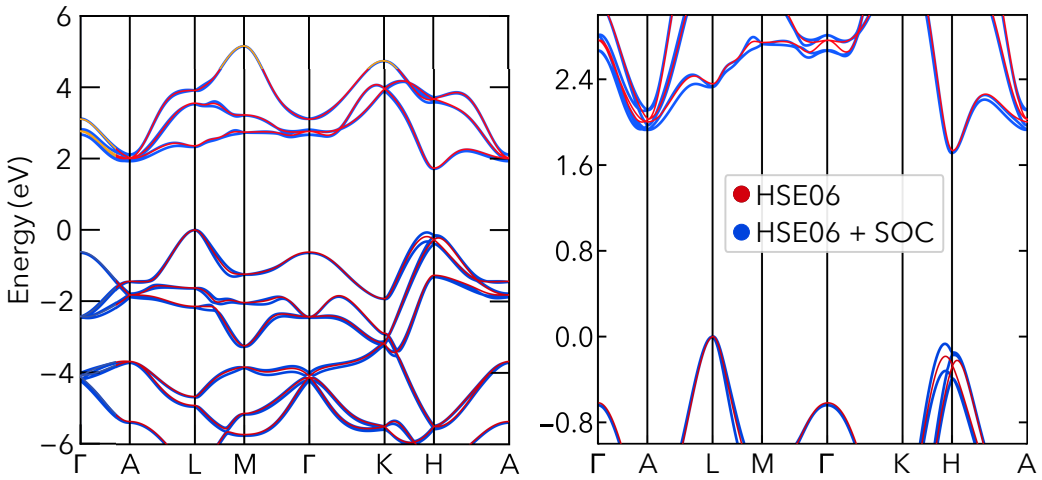

Fig. S2 Effect of spin-orbit coupling (SOC) on the electronic band structure of *t*-Se. The band structure calculated using hybrid DFT including SOC (HSE06+SOC) is shown in blue, while that calculated without SOC effects (HSE06) is overlaid in red. The VBM is set to 0 eV in both cases. SOC is found to have a small but significant effect on the carrier effective masses, decreasing the light carrier masses for both electrons and holes.

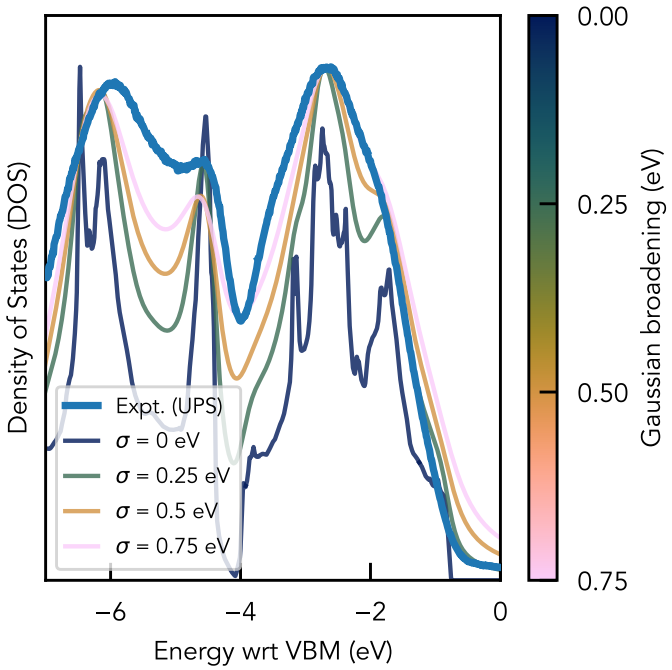

Fig. S3 Electronic density of states (DOS) of the upper valence band of *t*-Se, calculated using HSE06+D3+SOC, alongside experimental data from UPS measurements – shifted so that the leading edges align. The calculated DOS is shown after applying various levels of Lorentzian broadening (σ), normalised to the maximum peak height, and using the UPS cross-sectional weights from `Galore`.[141]

In Fig. S3, we see that the best agreement between the calculated *t*-Se density of states (DOS) and experiment is found for high levels of Gaussian energy smearing, around $\sigma \sim 0.5$ eV. Once again, this indicates strong thermal fluctuations in the electronic bands of *t*-Se at room temperature, resulting in these high levels of thermal broadening.

S3.1 *t*-Se: Electronic Band Alignment

We note that Shen *et al.*[42] measured the electronic band alignment of *n*-type Sb_2Se_3 with an ultrathin 2 nm *t*-Se passivation layer. Similar to our prediction and measurement in Fig. 2b, they find a type I band alignment, with the wider band gap of *t*-Se straddling the

smaller gap of Sb_2Se_3. However, the natural valence band offset (VBO) of Sb_2Se_3 and *t*-Se that we derive from both our predictions and measurements (when used in combination with experimental values for the Sb_2Se_3 band positions)[50–52] is ∼0.7 eV, compared to a value of 0.1 eV reported by Shen *et al.*[42] using XPS measurements. However, the ultraviolet photoelectron spectroscopy (UPS) results from Shen *et al.*[42] seem to match our values, giving VBM positions of $0.56+4.55=5.11$ eV and $0.70+5.18=5.88$ eV relative to vacuum for Sb_2Se_3 and *t*-Se respectively, compared to our values of 5.1 eV (from the experimental literature)[50–52] and 5.8 eV (calculated and experimentally measured). Moreover, the small thickness of the *t*-Se layers in Shen *et al.*[42] may result in significant residual band bending effects for these measurements, especially given the large depletion regions expected for *t*-Se.

Similar electron affinity and ionisation potential values for *t*-Se as measured here have been reported in Refs. 7,66,153.

S3.2 *t*-Se: Optical Anisotropy

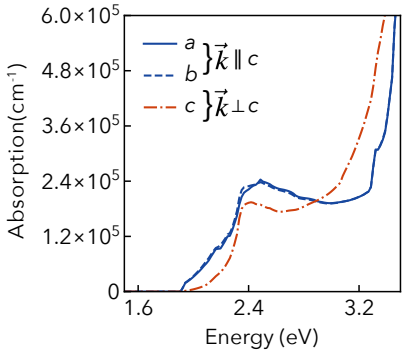

Fig. S4 Optical absorption of single-crystal *t*-Se calculated using hybrid DFT including SOC (HSE06+SOC), as a function of polarization direction. The corresponding propagation directions ($\vec{k}$) as perpendicular ($\perp$) or parallel ($\parallel$) to the Se chain direction (c) are also indicated. No smearing or spectral broadening has been applied, to highlight the spectral differences.

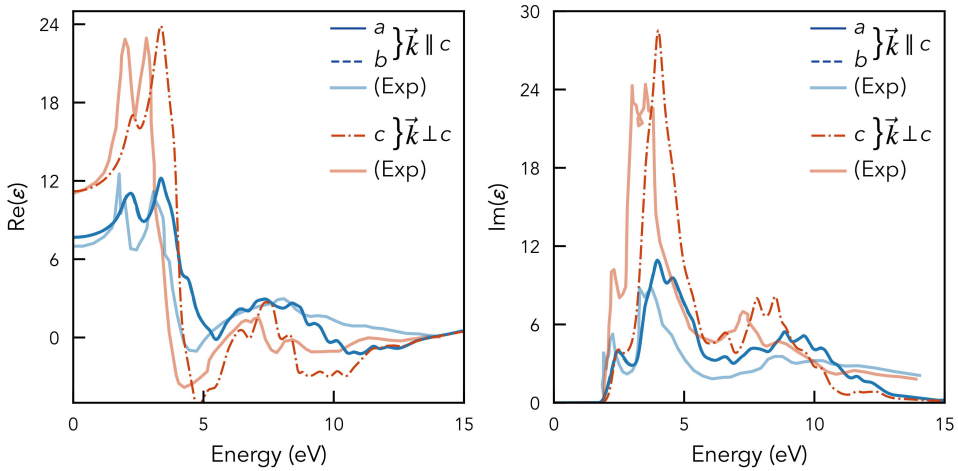

Fig. S5 Real and imaginary components of the dielectric function of single-crystal *t*-Se calculated using hybrid DFT including SOC (HSE06+SOC), as a function of polarisation direction. The corresponding propagation directions ($\vec{k}$) as perpendicular ($\perp$) or parallel ($\parallel$) to the Se chain direction (c) are also indicated. Gaussian broadening of 0.1 eV is applied to the calculated spectra. Alongside are plotted the results from reflectivity measurements by Tutihasi and Chen[154] and Leiga[155].

As shown in Figs. S4 and S5, the optical absorption of *t*-Se exhibits significant anisotropy, particularly near the onset energy where light propagation directions along c (which should dominate for epitaxial grown films along the Se chain direction) show greater light absorption. In combination with the fact that *t*-Se films oriented along the c-axis (chain direction) are expected to favour benign grain boundaries[63] — similar to the case of Sb_2Se_3 solar cells,[48], this indicates that *t*-Se films primarily oriented along the c direction (by epitaxial growth or otherwise) should exhibit greater photovoltaic efficiencies due to increases in both J_{sc} and V_{oc}.

We note that Ref. Moustafa *et al.*[38] also computed the anisotropic imaginary dielectric spectrum for *t*-Se, finding stronger absorption for light polarised *along* the Se chains (i.e. $\vec{k} \perp c$) for the 0 - 10 eV range they plotted over. Our results agree with this conclusion, but crucially for photovoltaic applications it is the strength of absorption at energies just above the band gap which are most important for

performance, for which we find light polarised *perpendicular* to the Se chains (i.e. with the direction of propagation along the c-axis; $\vec{k} \parallel c$) to give the strongest absorption (Fig. S4).

S3.3 t-Se: Direct Band Gap

Given the slight underestimation ($\Delta E_g = -0.12\,\text{eV}$) of the direct band gap of t-Se using the range-separated hybrid DFT functional HSE06, including spin-orbit coupling (SOC) and using the HSE06+D3-relaxed crystal structure (i.e. including van der Waals dispersion corrections), some alternative DFT functional choices were trialled for predicting the host band gap.

Firstly, using a dielectric-dependent hybrid functional (DDH@HSE) approach, where the fraction of exact Hartree-Fock exchange α_{exx} in the HSE hybrid DFT functional is set to $\frac{1}{\varepsilon_\infty} = \frac{1}{7.59} = 13.2\,\%$ (using the harmonic mean of ε_∞ as calculated using HSE06+SOC@HSE06+D3), we obtain far larger errors in both the structural parameters and band gap values for t-Se. Relaxing the structure with HSE(13.2 %)+D3 gives lattice parameters $a = 4.25\,\text{Å}, c = 5.05\,\text{Å}$, corresponding to errors of $-2.8\,\%$, $1.7\,\%$ and $-3.9\,\%$ for a, c and the cell volume compared to experimental values at room temperature.[44,46] The predicted direct band gap with this dielectric-dependent hybrid functional now underestimates the experimental value of 1.95 eV by $\Delta E_{g,\,\text{DDH}} \simeq -0.6\,\text{eV}$. This behaviour can be related to the unusually large high-frequency dielectric constant of t-Se for a system with a band gap close to 2 eV (as discussed in the main text), and thus low exchange fraction α within the DDH approach. As mentioned in the main text, the calculated high-frequency dielectric constants (using our HSE06+D3+SOC functional) match very well with the experimental values, and so the value of the dielectric is not expected to be a source of error. Alternative DDH approaches, which instead modify the range-separated exchange screening parameters, may perform better in this case.

The PBE0 hybrid DFT functional was also trialled. Unlike HSE06, PBE0 does not screen long-range exchange interactions and often better predicts band gaps for wide-gap semiconductors and insulators, though this is mostly due to the weaker screening in wider-gap compounds, which is not the case for t-Se as mentioned above. With PBE0+D3+SOC, a direct band gap of 2.46 eV is obtained, which now overestimates the experimental value of ∼1.95 eV to a greater extent than the underestimation (∼0.1 eV) of HSE06+D3+SOC.

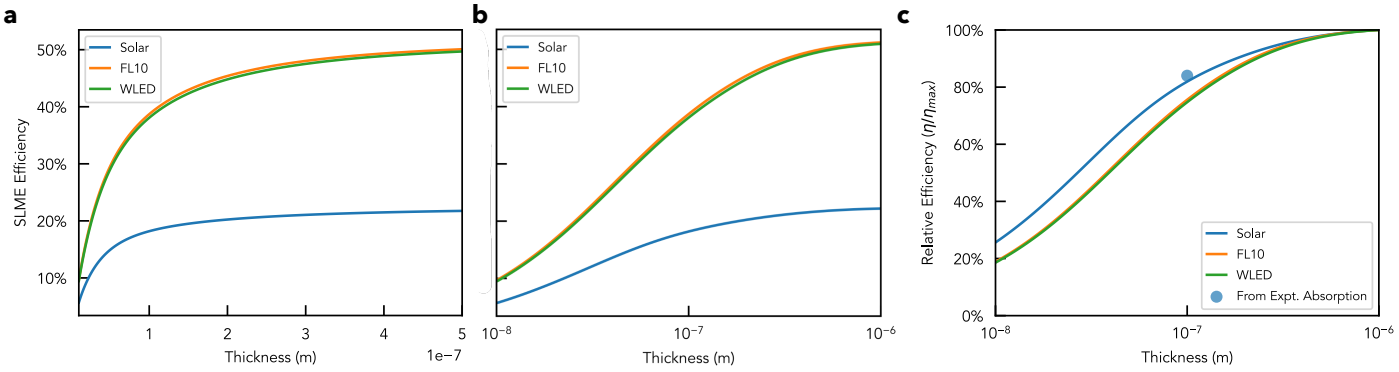

Fig. S6 Spectroscopic-Limited Maximal Efficiency (SLME) for single-junction t-Se photovoltaic (PV) cells as a function of absorber thickness, with linear (a) and log (b) scale x-axes. Subfigure (c) plots the relative efficiencies as a function of thickness, along with the expected value based on the reported transmittance of 16 % for a FTO + 100 nm Se layer by Todorov *et al.*[25] The calculated (i-)SLME values are shown for both solar irradiation and typical indoor light sources; fluorescent lamp (FL10) and white light-emitting diodes (WLED) at 1000 lux.[136] 'i-SLME' refers to the SLME metric adapted for indoor light sources. The calculated absorption curve of t-Se from HSE06+SOC@HSE06+D3 is slightly upshifted to give a direct band gap of 1.95 eV and indirect band gap of 1.85 eV, matching the experimental values.

S3.4 t-Se: Deformation Potential Analysis

We calculate a surface energy of $0.177\,\text{J/m}^2$ for t-Se along the [100] lattice direction (i.e. inter-chain direction; ⊥), in excellent agreement with the measured lower bound of $0.175\,\text{J/m}^2$ reported by Lee[54] in 1971. This low surface energy (c.f. $0.25\,\text{J/m}^2$ for layered SnO)[156] indicates that weak vdW dispersion forces dominate the inter-chain Se interactions.

In Fig. S7, we plot the calculated internal energy per atom of t-Se as a function of the unit cell volume, where the cell shape and internal atomic coordinates are either kept fixed or allowed to vary. We see that the energy variation with this volume deformation is extremely small for the physically-relevant case where cell shape and internal coordinates are relaxed at each volume, varying by just over 2 meV/atom as the volume is varied by ±5 % from the equilibrium value. Fitting this energy-volume curve using the Birch-Murnaghan equation of state[157] gives quite a low bulk modulus of 15.0 GPa, which along with the low phonon frequencies shown in Fig. S1 indicates the facile compressibility and weak inter-chain binding of this low-dimensional system. However, despite the small energy variation, both the indirect and direct band gap vary quite dramatically with this volume deformation, over a 0.34 eV range for the indirect band gap and a 0.46 eV range for the direct band gap. Combined, this suggests that thermal fluctuations of the band gap are likely to be significant in t-Se.

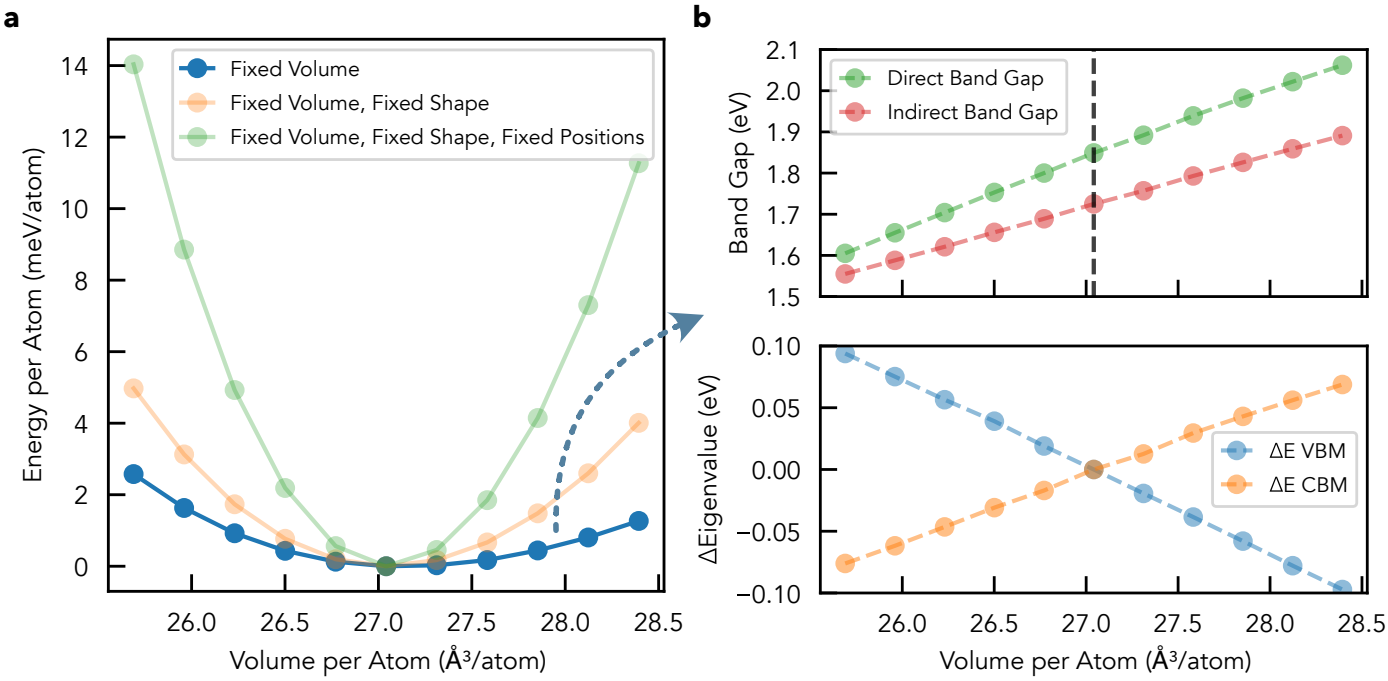

Fig. S7 Effects of volume deformation on energetic and electronic properties of t-Se, calculated using HSE06+D3+SOC. (a) Energy per atom as a function of volume per atom in t-Se. Results are also shown for fixed volumes and fixed shapes (i.e. c/a lattice parameter ratio), and for fixed internal atomic positions within the unit cell. (b) Direct and indirect band gap (upper) and band edge eigenvalue changes (lower) as a function of volume per atom, where cell shape and internal positions have been allowed to vary (corresponding to 'fixed volume' in (a). Atomic core level potentials were used to align the band edge eigenvalue changes.

We note that Fu *et al.*[59] also recently observed significant differences in energy-volume (deformation) curves in the low-dimensional $CuSbSe_2$ system when allowing the internal coordinates to relax as opposed to keeping them fixed, which is typically not the case for densely-packed 3D solids such as II-VI semiconductors. Allowing cell shape and internal coordinate relaxation is crucial when performing volume deformation analysis of low-dimensional/compressible solids.

S4 Point Defect Thermodynamics

Both equilibrium defect concentrations and the positions of charge transition energy levels are determined by the formation energies of defects in their various charge states. The formation energy $E^f_{X,q}(E_F,\mu)$ of a defect X in charge state q is given by:[69,70,158]

$$E^f_{X,q}(E_F,\mu) = [E_{X,q} - E_H] - \sum_i n_i\mu_i + qE_F + E_{Corr}(q) \quad \text{(S1)}$$

$E_{X,q}$ is the calculated total energy of the finite-sized *defect-containing* supercell, while E_H is the energy of the *pristine* bulk supercell. The summation $\sum_i n_i\mu_i$ incorporates the thermodynamic cost of exchanging n_i atoms with their reservoir chemical potential(s) μ_i (upon forming the defect species X_q from the ideal host material). Analogously, qE_F accounts for the energetic cost of addition or removal of charge q to the defect site, from the electronic reservoir of the system (the Fermi energy E_F). Finally, $E_{Corr}(q)$ is an *a posteriori* correction to remove any spurious interactions introduced by the use of a finite-size supercell — in this case, the Kumagai-Oba[122] (eFNV) charge correction scheme was used, as automated in `doped`[71] via the `pydefect`[123] API.

In calculating the defect and charge-carrier concentrations, and thus self-consistent Fermi level positions, the 'frozen defect approximation' was used.[121] Within this model, total defect concentrations are computed at the given synthesis/annealing temperature and then fixed to these values, while the relative concentrations of different charge states of the same defect are then allowed to re-equilibrate at room temperature under this total concentration constraint.

For the point defects computed in t-Se, charge state ranges of +3 to -3 were tested for V_{Se}, Se_i, As_{Se}, As_i, Sb_{Se} and Sb_i; +5 to -3 for P_{Se} and P_i; +1 to -3 for N_{Se} and N_i; +1 to -1 for H_{Se}, H_i, X_{Se} and X_i (X = halogen, except F_{Se} and Cl_{Se} for which +/-2 charge states were also trialled); 0 to -2 for O_{Se} & O_i; and +2 to -2 for S_{Se}, S_i, Te_{Se}, Te_i.

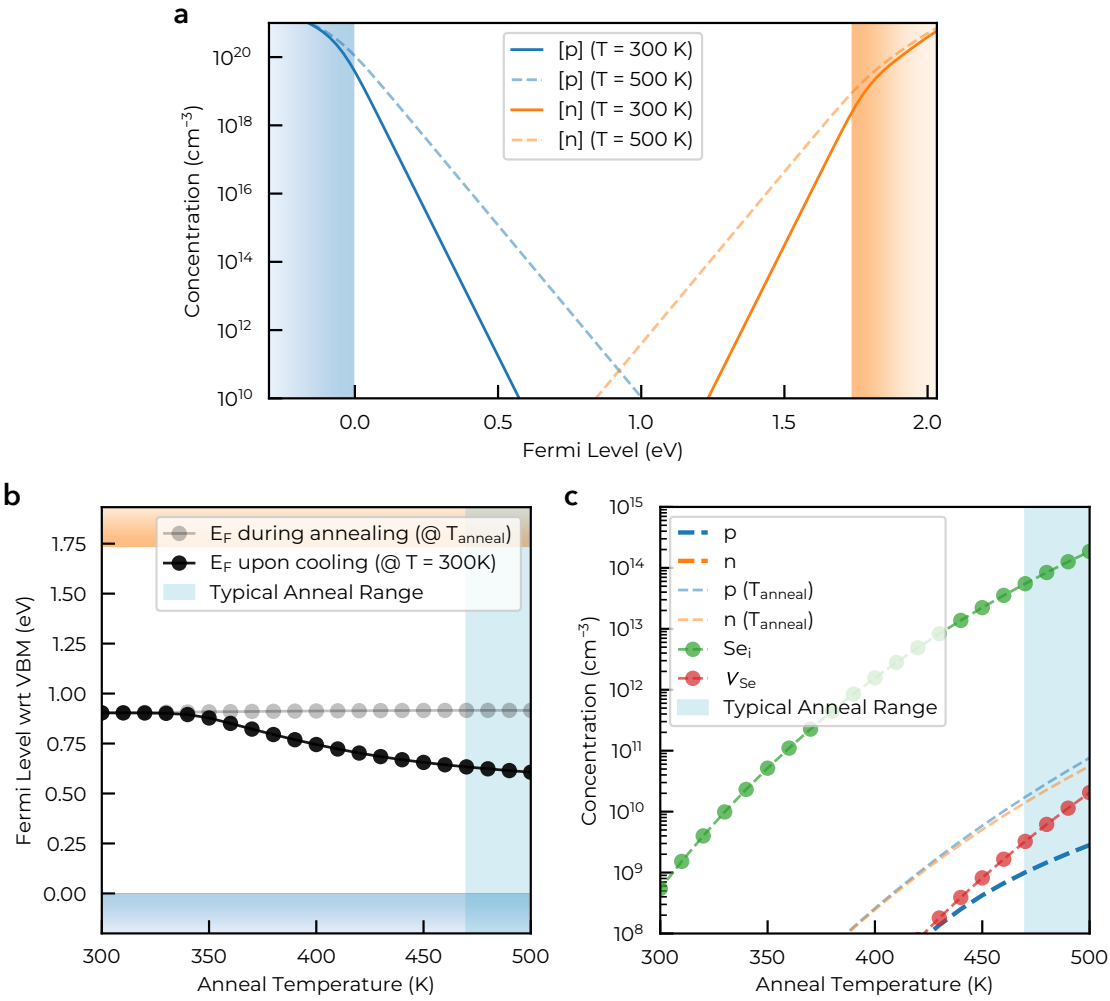

Fig. S8 Defect and charge carrier concentrations in undoped t-Se, as output by the doped defect simulation python package using the HSE06+D3+SOC defect formation energies and electronic density of states. (a) Charge carrier concentrations in t-Se as a function of Fermi level position, at elevated annealing temperatures ($T \sim 500\,\mathrm{K}$) and when cooled to room temperature ($T \sim 300\,\mathrm{K}$). (b) Self-consistent Fermi level position in intrinsic, undoped t-Se during annealing (grey) and upon cooling to room temperature ($T = 300\,\mathrm{K}$; black), as a function of annealing temperature. Valence and conduction band edges shown in blue and orange respectively. Typical anneal temperatures for t-Se crystallisation are indicated by the light blue region.[18,24,25] (c) Calculated defect and carrier concentrations as a function of anneal temperature in intrinsic, undoped t-Se. Carrier concentrations under annealing are also shown in faded blue/orange.

S5 Further Defect Analysis

S5.1 Valence Alternation Pair (VAP) Model

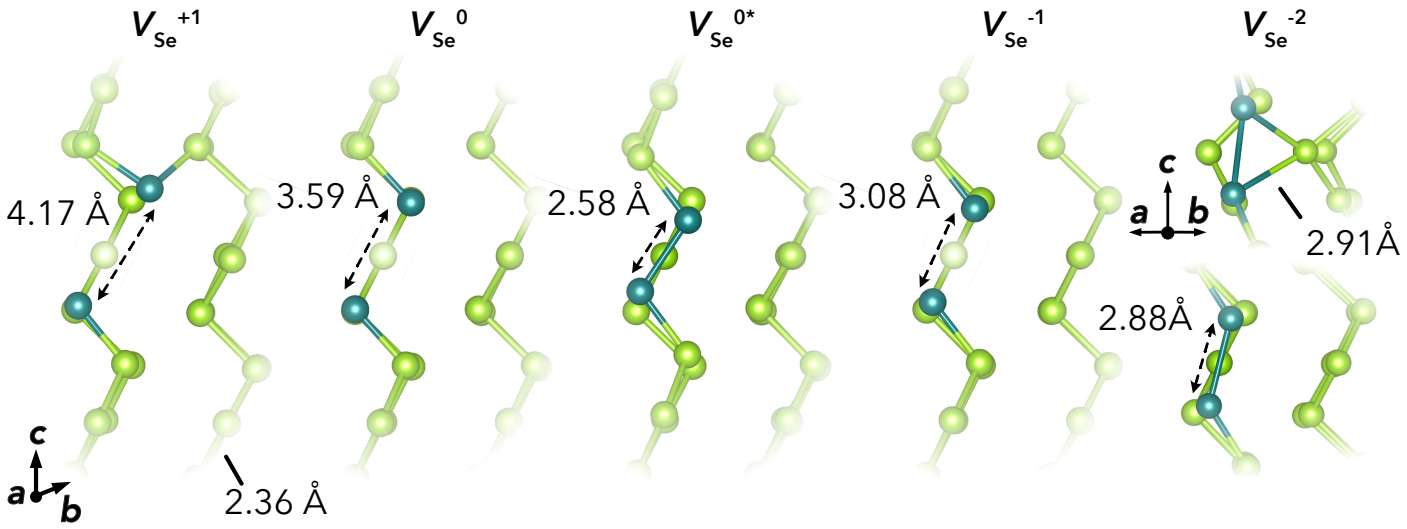

Fig. S9 Geometries of the intrinsic vacancy (V_{Se}) in *t*-Se, under various charge states, looking down the [100] lattice direction. The two neighbouring Se atoms to the vacancy site are shown in teal as a guide to the eye. V_{Se}^{0*} refers to the metastable neutral vacancy, and V_{Se}^{-2} is additionally shown along the [110] direction to make the bridging bonds between Se chains visible.

In 1975, Street and Mott[159] proposed negative-U centre behaviour for dangling bonds in glassy chalcogenide compounds such as As_2Se_3, where two neutral singly-coordinated ("C_1^0") dangling bonds would disproportionate to form a positive triply-coordinated atom ("C_3^+") and a negatively-charged dangling bond ("C_1^-").[160] This work was built upon by Kastner *et al.*[94], who termed such species as 'valence-alternation pairs' (VAPs). This model aligns with our findings for the positively-charged vacancy V_{Se}^{+1}, where we obtain the expected "$C_3^+ + C_1^0$" configuration (with the three-fold coordinated Se atom at the inter-chain bridging bond) being the positively-charged species as suggested by the VAP model. As expected, for V_{Se}^{+2}, we also find it to adopt the expected "$C_3^+ + C_3^+$" configuration where the two dangling chain ends bond with the neighbouring chains (Fig. S10), rather than the "$C_1^+ + C_1^+$" state which is predicted to be higher energy by the VAP model. For V_{Se}^{0} however, we note that our findings differ to that expected from the VAP model, where the two neutral dangling bonds ("$2C_1^0$") for V_{Se}^{0} are found to be the ground-state arrangement, rather than "$C_3^+ + C_1^-$" as predicted by the VAP model, with the 'self-healed' metastable state V_{Se}^{0*} also being only slightly higher in energy. We note that structure-searching (`ShakeNBreak`),[80,81] with hybrid DFT including spin polarisation and no symmetry constraints, was used to sample a range of potential geometries here — including manual testing of the "$C_3 + C_1$" structure of V_{Se}^{+1} in the 0 and -1 charge states, both of which relaxed back to the previously-identified ground-state structures. We postulate that this may be the result of a high cost of charge localisation ("$C_3^+ + C_1^-$" vs "$2C_1^0$") in this pure-covalent, low-dielectric compound, as well as a possible correlation effect between the two hole states of the bipolaron-like V_{Se}^{0} state.

S5.2 Further Vacancy Analysis

The 'self-healing' type rearrangement of metastable V_{Se}^{0*} is similar to that of iso-electronic, recombination-active V_P^{+1} in the structurally-analogous NaP compound.[3] However, unlike V_P^{+1} in NaP, this geometry results in bond lengths and angles which deviate significantly from the ideal bulk values; 2.58 Å at the vacancy position and 2.44 Å for the two next-nearest Se-Se bonds in the chain, and bond angle deviations of up to 10°. This strain cost for V_{Se}^{0*}, and the charge-localisation cost for the bipolaron-type V_{Se}^{0}, result in significantly greater energy costs of formation for neutral vacancies compared to self-interstitials in *t*-Se.

Given the small energy difference between spin states, the effect of spin-orbit coupling (SOC) is significant for the relative energies of the singlet and triplet states for the V_{Se}^{0} bipolaron. The ground-state spin arrangement is a singlet ($S = 0$), but the triplet ($S = 1$) state is calculated to be only 10 meV higher in energy, or 46 meV when SOC effects are excluded.

Notably, the similarity of geometries for the intrinsic point defects in *t*-Se and NaP[3] — in particular V_{Se}^{0*} and V_P^{+1} — suggests that a bipolaron-like state with two terminated chains is likely to be metastable for V_P^{+1} (which has not yet been reported). This is also supported by the electronic structure of V_P^{0} and V_P^{-1} as shown in Figure 6 of Kumagai *et al.*[3], exhibiting similar in-gap polaron levels, though now (partially) occupied in these charge states. From Figure 6 of Kumagai *et al.*[3] we see that the 'self-healed' motif for V_P^{+1}

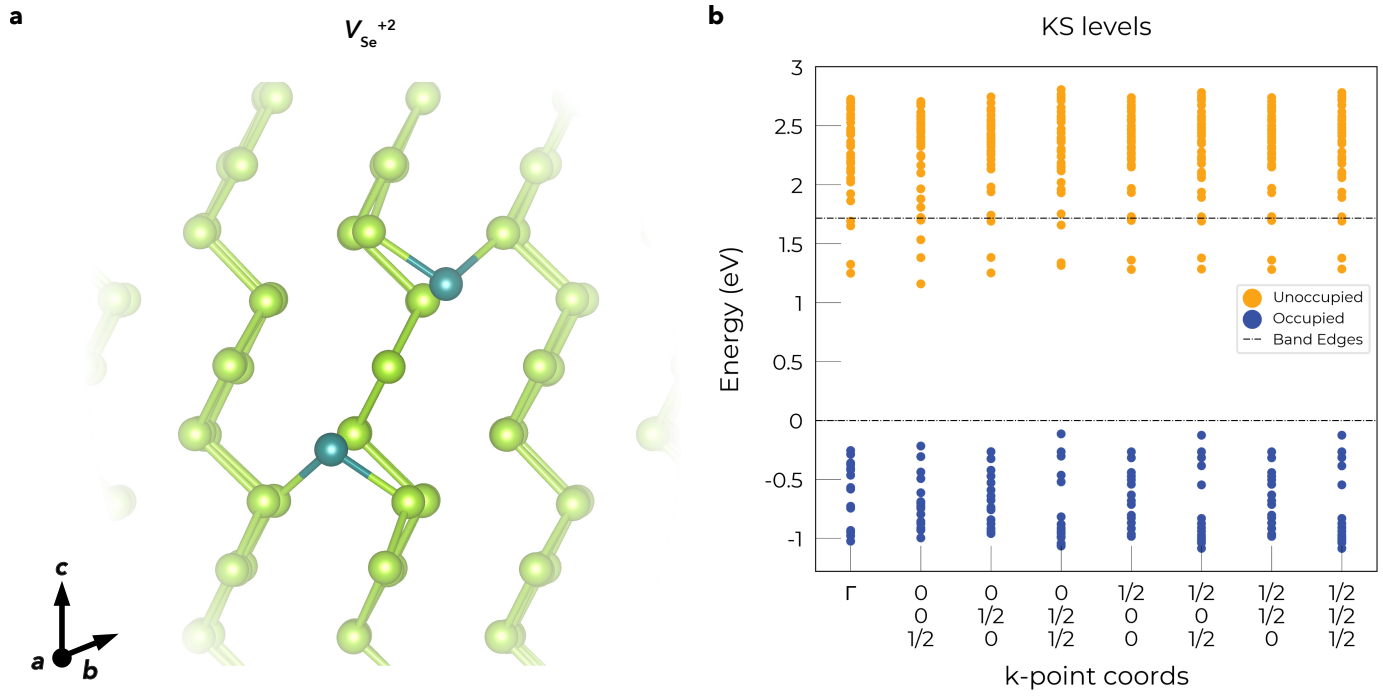

Fig. S10 (a) Geometry of the doubly-positive intrinsic vacancy (V_{Se}^{+2}) in *t*-Se, looking down the [100] lattice direction. The two neighbouring Se atoms to the vacancy site are shown in teal as a guide to the eye. (b) Single-particle eigenvalues from HSE06+SOC@HSE06+D3 DFT supercell calculations of V_{Se}^{+2} in *t*-Se, as output by doped[71] using the `pydefect`[123] parsing functions, showing two unoccupied localised states just below the CBM. "KS levels" = Kohn-Sham levels.

involves much less deviation from bulk bond lengths than for V_{Se}^{0*} in *t*-Se — likely a consequence of the fourfold P chain helix rather than threefold for *t*-Se, indicating a much lower strain cost and explaining why it is the observed ground-state structure for V_P^{+1} in NaP. The $V_P^{+1} \leftrightarrow V_P^{0}$ transition is already predicted to be the dominant recombination mechanism for intrinsic defects in NaP,[3] and so the potential presence of a low-energy V_P^{+1*} state with similar geometry to V_P^{0} could act to further catalyse this recombination behaviour.

The calculated charge-capture coefficients and cross-sections for V_{Se} in *t*-Se, as obtained from fitting of the anharmonic potential energy surfaces shown in Fig. 6, are shown in Fig. S11.

As discussed in the main text, the fastest charge capture rate is given by σ_h^0 (hole capture at the $\varepsilon(+1/0)$ transition level) as expected, which exhibits a relatively small classical capture barrier of 135 meV and a large capture cross-section $\sigma_h^0 = 3.2 \times 10^{-17}\,\mathrm{cm}^{-2}$ @ $T = 300\,\mathrm{K}$ — of similar magnitude to the Se atomic cross-section $\sigma_{Se} = \pi(a_{Se})^2 \simeq 4.5 \times 10^{-16}\,\mathrm{cm}^{-2}$. All other capture cross-sections for V_{Se}, however, are > 5 orders of magnitude smaller at room temperature, corresponding to weak traps. We note that this differs from V_P in the structurally-analogous NaP compound for which fast hole *and* electron capture was predicted,[3] primarily due to the softer PES and thus 'inverted Marcus regions'[†] for the slow transitions here. The slow charge capture rates for V_{Se} here can mostly be ascribed to the soft and anharmonic PES along the paths between equilibrium geometries of the various charge states (Fig. 6), which itself is related to the relatively soft bonding and small structural differences between different charge states (Fig. 5).

In each case[‡] the $0 \leftrightarrow \pm 1$ transition is faster for the *ground-state* neutral vacancy V_{Se}^{0} rather than V_{Se}^{0*}, due to its closer structural similarity to the V_{Se}^{+1} & V_{Se}^{-1} states (each involving a ruptured Se chain at the vacancy position; Fig. 5) as indicated by the smaller mass-weighted displacements (ΔQ) between equilibrium geometries in Fig. 6. Charge carrier capture rates are typically expected to be roughly inverse-exponentially dependent on the energetic distance of the defect level to the corresponding band-edge,[162–164] based on the strength of coupling between defect and band-edge electronic states. This trend is partly witnessed here, with hole capture rates decreasing for V_{Se} levels further from the VBM; $\sigma_h^0 > \sigma_h^{-1} > \sigma_h^{-2}$ — though not inverse-exponentially. This is not a general rule however, particularly in the case of lower symmetry or anharmonic bonding,[48,72,79,83,165] and we see that the electron capture cross-sections do not follow this trend; $\sigma_e^{+1} > \sigma_e^{-1} > \sigma_e^{0}$ at room temperature.

† 'Inverted Marcus region'[161] refers to the case where the excited state minimum is located *within* the basin of the ground-state energy surface (i.e. $\varepsilon(0/-1)$ & $\varepsilon(-1/-2)$), whereas the $V_{Se}^{0(*)} + h^+ \rightarrow V_{Se}^{+1}$ transitions correspond to the normal Marcus region.

‡ $0 \leftrightarrow \pm 1$ transitions are faster for V_{Se}^{0} than V_{Se}^{0*} except for electron capture at $\varepsilon(0/-1)$ (σ_e^0) which is extremely slow in both cases.

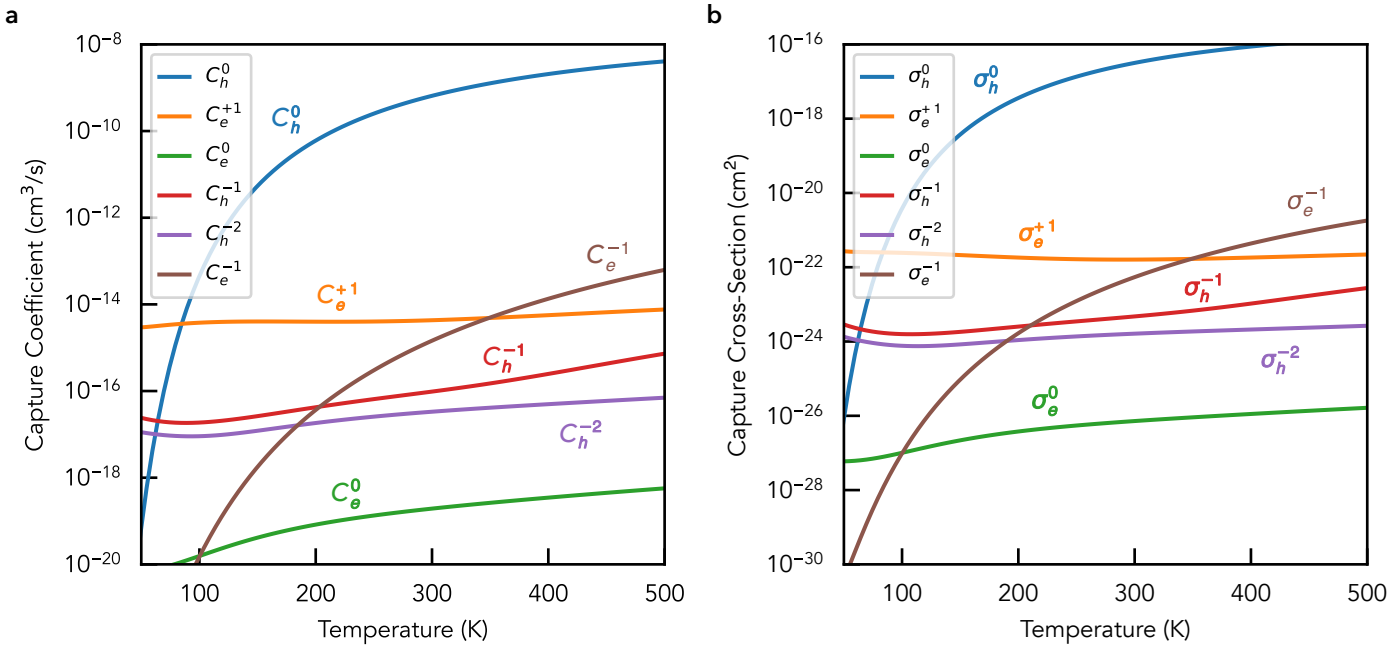

Fig. S11 Charge capture coefficients (a) and cross-sections (b) of vacancy defects in t-Se, as a function of temperature. The charge of the defect is given by the superscript, and the charge carrier is denoted by the subscript; e/h for electron/hole; $\sigma^q_{e/h}$.

S5.3 Further Interstitial Analysis

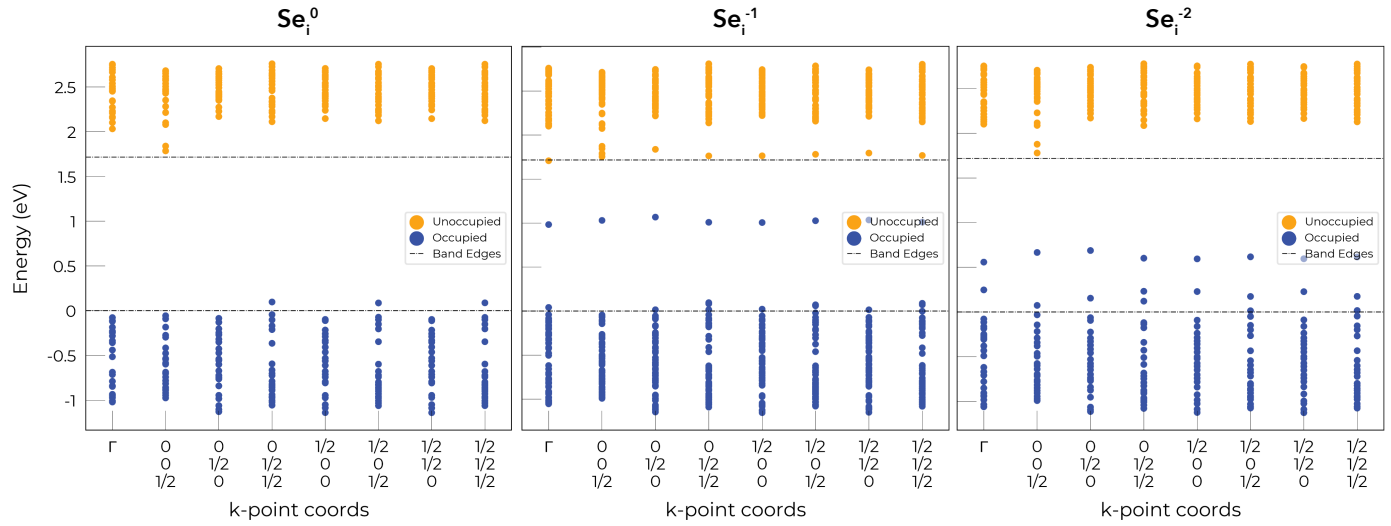

Fig. S12 Single-particle eigenvalues from HSE06+SOC@HSE06+D3 DFT supercell calculations of Se_i^0 and Se_i^{-2} in t-Se, as output by doped[71] using the `pydefect`[123] parsing functions.

As shown in Fig. S13, for Se_i^{-2} the Se chain is ruptured and two bridging bonds are formed by the terminated ends with neighbouring chains to stabilise the charge, while for Se_i^{-1} the split-interstitial geometry is retained, but with a lengthening of the $(Se\text{-}Se)_{i,\,split}$ bond to 2.64 Å and the formation of a bridging bond to a neighbouring chain.

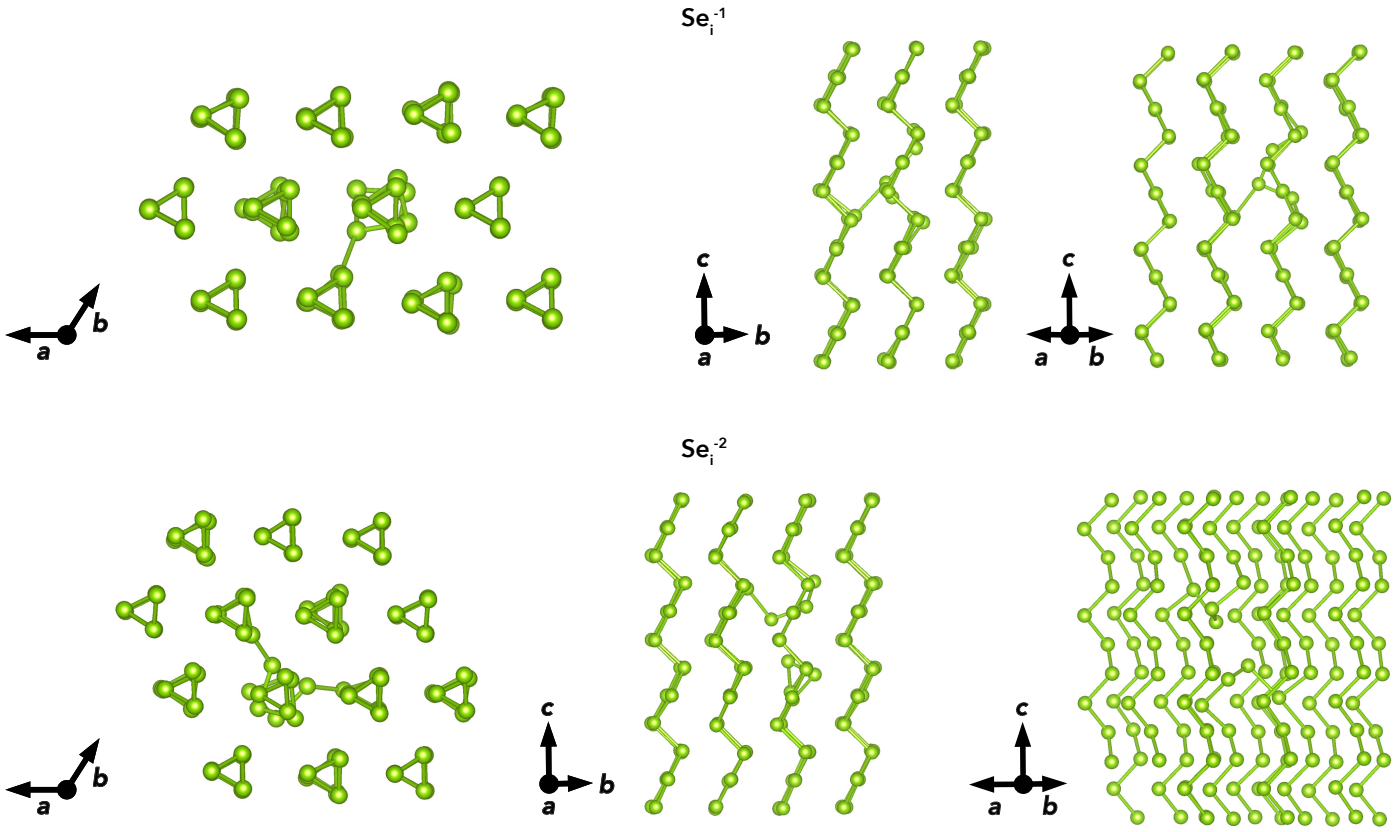

Fig. S13 Relaxed geometries of Se_i^{-1} and Se_i^{-2} defects in *t*-Se, along different crystal directions.

S5.4 High Concentration Impurity Levels

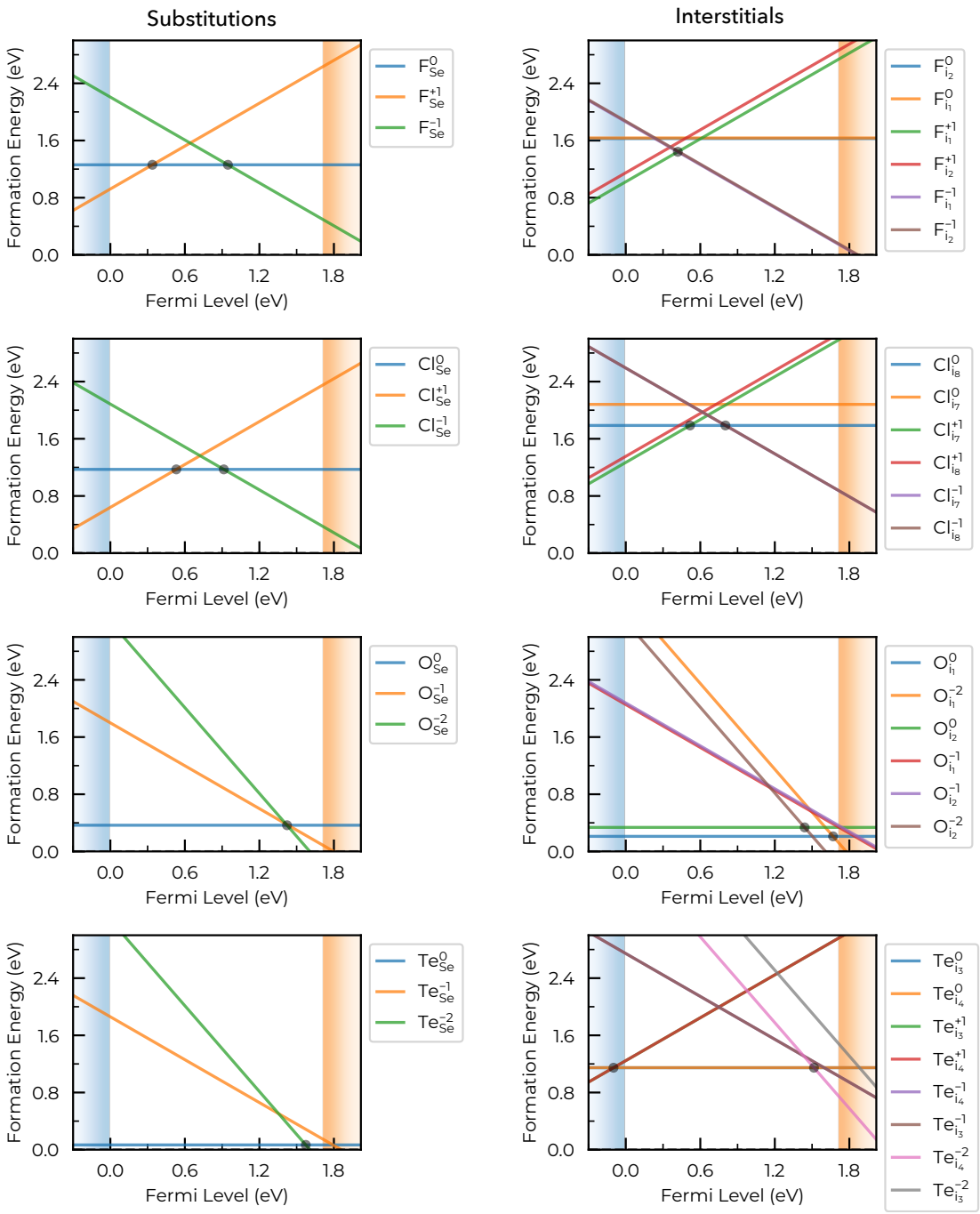

Fig. S14 Formation energy diagrams for F, Cl, O and Te impurities in *t*-Se, separated into substitutional and interstitials defects. All identified (metastable) states are included, and the charge transition levels are also tabulated in Section S5.6.

t-Se exhibits two photoluminescence (PL) peaks at energies 1.67 eV and 1.79 eV, which are thought to be associated with defects.[18] These optical transitions could possibly be caused by high-concentration impurities such as F, Cl, Te or O, in addition to the multiple charge transition levels exhibited by V_{Se} (Fig. 3). The calculated charge transition levels for these defects, including all identified (metastable) states, are plotted in Fig. S14 and tabulated in Section S5.6 below.

S5.5 ShakeNBreak

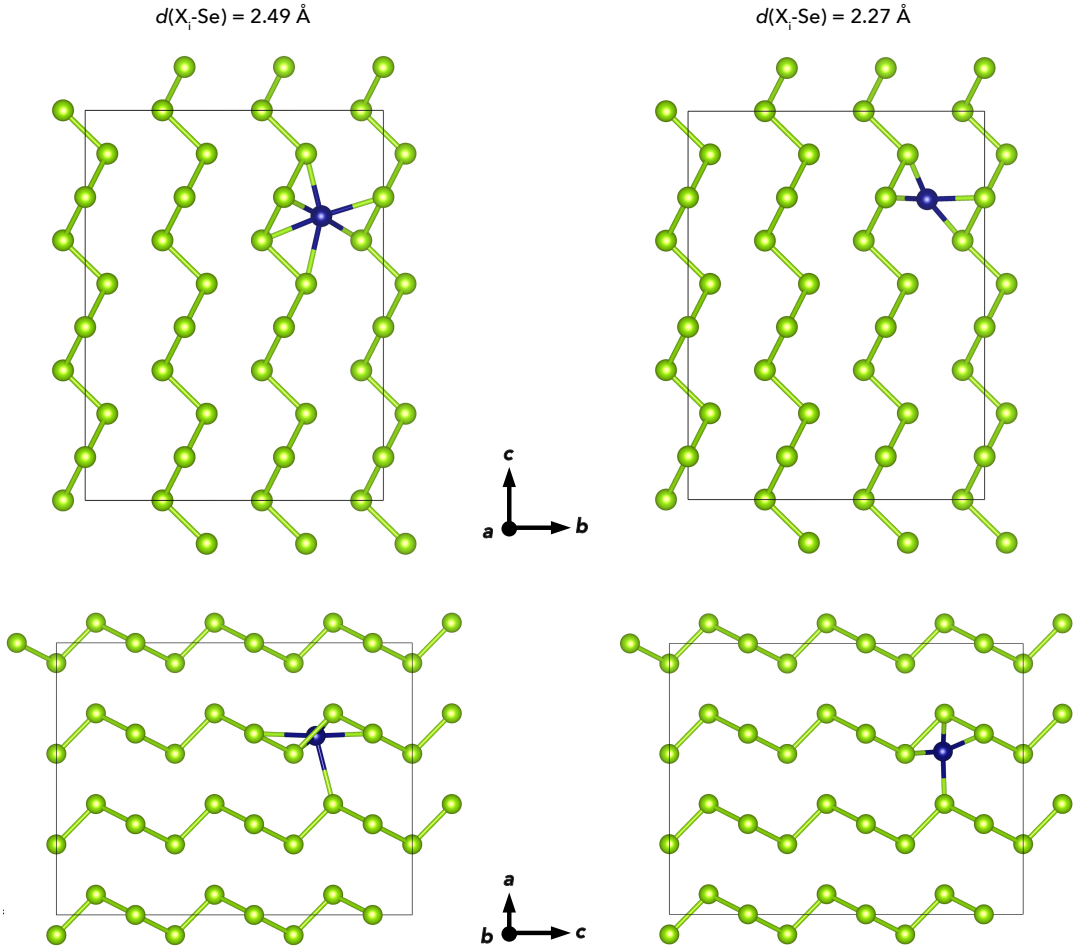

Fig. S15 Interstitial positions trialled for each interstitial defect species (X_i) in *t*-Se, as viewed along the crystal *a* (top) and *b* (bottom) axes. X_i shown in midnight blue, Se in green. The distances of the two interstitial positions to the nearest Se atoms are given as titles.

Two interstitial sites were trialled for each interstitial defect type in Se, which are shown in Fig. S15. For Se_i^{-1} and Se_i^{-2}, standard gradient relaxations starting from the $X_{i,\,2.27\,Å}$ position are lower energy by 0.5 eV and 0.2 eV, though `ShakeNBreak`[80,81,95] structure-searching starting from the $X_{i,\,2.49\,Å}$ site additionally identifies these lower energy configurations.

For the neutral vacancy V_{Se}^{0}, a Γ-point only calculation in the 81-atom supercell predicts the 'self-healed' V_{Se}^{0*} structure to be 25 meV lower energy than the 'bipolaron' split-chain V_{Se}^{0} geometry, while calculations with a $2\times2\times$ *k*-point mesh flips this energy ordering to predict the split-chain V_{Se}^{0} to be 26 meV lower energy than the 'self-healed' V_{Se}^{0*}. When distinct geometries with small energy differences are identified with coarse relaxations during structure searching, it is important to continue the energy calculations with tight fully-converged calculation parameters to accurately predict the relative energies.

For the extrinsic defect species, lower energy structures were identified for all X_{Se}^{-1} where *X* is a halogen (F, Cl, Br, I) with $\Delta E \simeq 0.7$ eV. As shown in Fig. S16, the defect geometry returned by a standard gradient relaxation ('unperturbed') has the halogen placed on the Se site, with a repulsion of the neighbouring Se atoms – e.g. giving a Cl_{Se}-Se bond length of 2.83 Å, compared to the bulk Se-Se bond length of 2.36 Å. For the ground-state defect geometry identified by `ShakeNBreak` ($\Delta E \simeq 0.7$ eV), the halogen atom instead attracts one of the neighbouring Se atoms (e.g. Cl_{Se}-Se bond length of 2.21 Å), while the other neighbouring Se bonds with a neighbouring chain. Note that for the `ShakeNBreak` distortion plots provided here, only defect species which showed a significant energy lowering relative to 'unperturbed' relaxations are shown.

For interstitials, in all cases ShakeNBreak identified the lower energy of the two trial interstitial sites (Fig. S15), using the targeted bond distortions and atom rattling in this method, when starting with the interstitial atom initially placed at the higher energy site. Moreover, several lower-energy structures were identified for halogen interstitials which again are missed by standard gradient relaxations from either of these trial interstitial sites as shown in Fig. S17, with ΔE of ∼0.04 eV, 0.11 eV, 0.15 eV, 0.25 eV for X_i^{+1}; X = F,

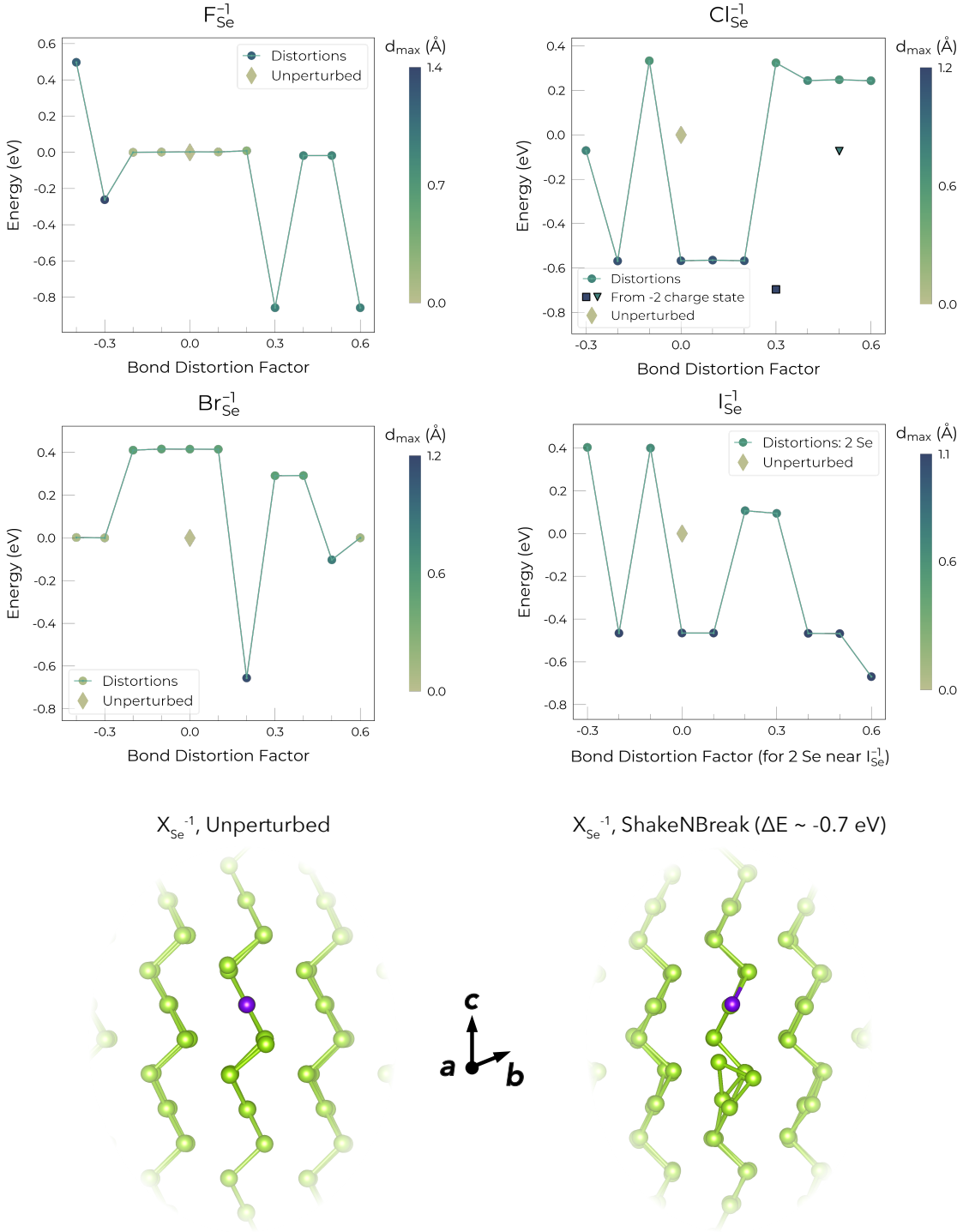

Fig. S16 (Upper) `ShakeNBreak` results for halogen substitutions in *t*-Se, as output directly by the `snb-plot` utility. Above are plotted the final energies obtained from geometry relaxations of trial structures generated for the given defect species by `ShakeNBreak`, as a function of the bond distortion factor (corresponding to the fractional distortion of nearest neighbour bond lengths) applied to the initial geometry. The number of nearest-neighbour displacements is dictated by the change in valence electrons for the defect species (a.k.a. excess charge), which for X_{Se}^{-1} (X = halogen) is 2 Se neighbours. 'Rattling' of all other atomic positions in the supercell is also applied, using Monte-Carlo sampling from a Gaussian distribution, normalised by the bulk bond length of the host. Further details are provided in Mosquera-Lois *et al.*[80] The 'Unperturbed' datapoint, shown with a diamond symbol in faded gold, corresponds to the result obtained with a standard unperturbed supercell relaxation. Datapoints are coloured according to the maximum displacement of any atom relative to the final 'unperturbed' geometry, as an indicator of structural disparity. (Lower) Final geometries of X_{Se}^{-1} (X = halogen) defects in *t*-Se, as obtained from unperturbed and `ShakeNBreak` relaxations.

Cl, Br, I; ∼0.05 eV,[§] 0.3 eV, 0.3 eV, 0.6 eV for X_i^0. For Br_i^{-1}, a ∼0.3 eV lower-energy ground-state structure was also identified. We note

§ 0.1 eV energy difference relative to the higher-energy unperturbed interstitial site, from which the structure search that identified th

that the F_i^{+1} geometry has the same ruptured chain with two bridging bonds as shown in Fig. S10a, now with the F_i^{-1} bonded to one of these bridging Se atoms. The combination of high non-equilibrium impurity concentrations, with kinetic inhibition of the structural transformation from the F_i^{-1} geometry to that of F_i^{+1} upon cooling from the annealing temperature, could hypothetically lead to significant (but metastable) hole concentrations on the order of 10^{14}-10^{17} cm^{-3} in *t*-Se, which would give time and temperature-dependent carrier concentrations. No such metastability/transience has been reported for *t*-Se to the knowledge of the authors however.

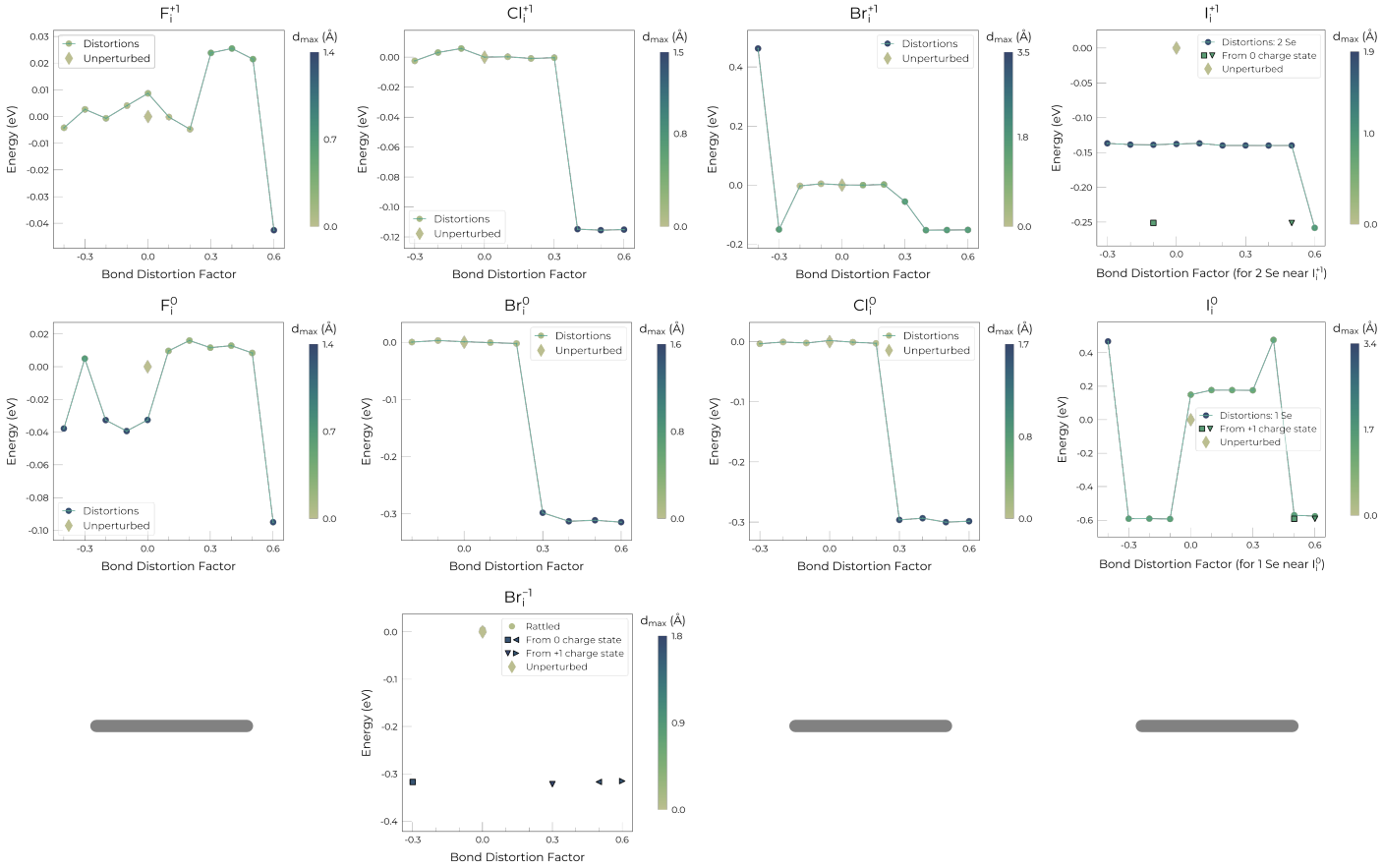

Fig. S17 `ShakeNBreak` results for halogen interstitials in *t*-Se, as output directly by the `snb-plot` utility. Above are plotted the final energies obtained from geometry relaxations of trial structures generated for the given defect species by `ShakeNBreak`, as a function of the bond distortion factor (corresponding to the fractional distortion of nearest neighbour bond lengths) applied to the initial geometry. The number of nearest-neighbour displacements is dictated by the change in valence electrons for the defect species (a.k.a. excess charge), which for X_i^q (X = halogen) is $q+1$ Se neighbours. 'Rattling' of all other atomic positions in the supercell is also applied, using Monte-Carlo sampling from a Gaussian distribution, normalised by the bulk bond length of the host. Further details are provided in Mosquera-Lois *et al.*[80] The 'Unperturbed' datapoint, shown with a diamond symbol in faded gold, corresponds to the result obtained with a standard unperturbed supercell relaxation. Datapoints are coloured according to the maximum displacement of any atom relative to the final 'unperturbed' geometry, as an indication of structural disparity.

For hydrogen and chalcogen point defects in *t*-Se, energy-lowering distortions in the range of 0.15 - 0.55 eV were identified by `ShakeNBreak` for some singly-negative charge states, as shown Fig. S18. For pnictogen point defects in *t*-Se, many energy-lowering distortions (missed by standard energy gradient relaxations) are found by `ShakeNBreak`, as shown in the distortion plots in Figs. S19 and S20. For As_{Se}^{+1} a higher energy structure was originally found, which made As_{Se} a slight outlier compared to P_{Se} and Sb_{Se} in terms of defect level positions. Starting a gradient relaxation from the P_{Se}^{+1} geometry gave a lower energy structure ($\Delta E = 0.3$ eV), making the behaviour of As_{Se} consistent with P_{Se} and Sb_{Se} as expected.

For pnictogen impurities, only P_i and As_i are found to be stable in -3 charge states, with transition levels just below the CBM. For these defects, the geometries correspond to that of the P/As_i^{-1} states, but with an additional rupturing of the Se chain on which the substitution is located (at the first Se-Se bond neighbouring the substitution site), in the same configuration as witnessed for the -2 charge self-interstitial (Fig. S13 (the Se chain is ruptured and two new bridging bonds are formed, with an electron localising on an atom involved in each). We also note that while the formation energies of N_{Se} and N_i are very similar (Fig. 8), in both neutral and

in Fig. S17), but only 0.05 eV lower energy than the unperturbed relaxation starting from the other, lower-energy interstitial site.

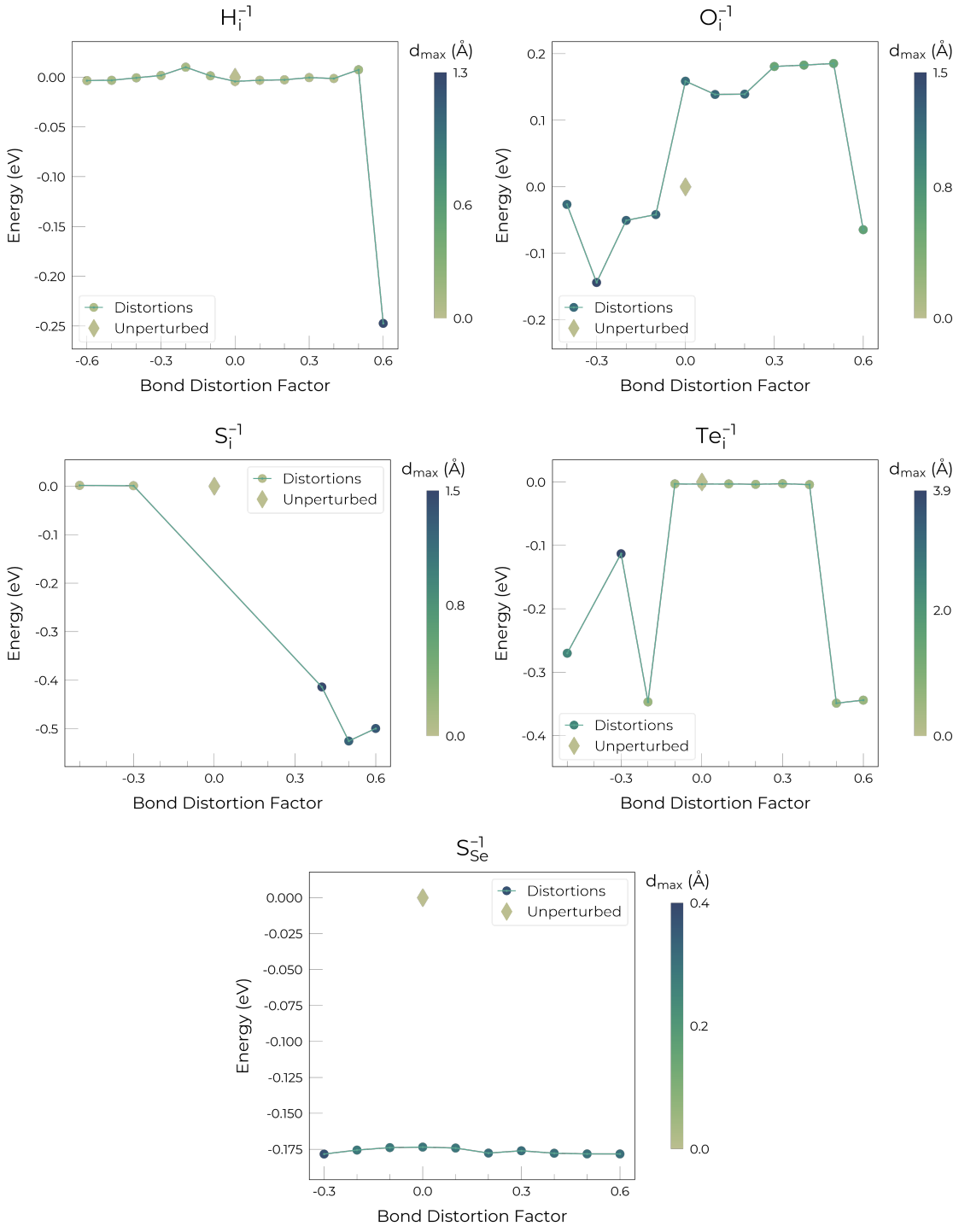

Fig. S18 `ShakeNBreak` results for hydrogen and chalcogen defects in *t*-Se, as output directly by the `snb-plot` utility. Above are plotted the final energies obtained from geometry relaxations of trial structures generated for the given defect species by `ShakeNBreak`, as a function of the bond distortion factor (corresponding to the fractional distortion of nearest neighbour bond lengths) applied to the initial geometry. The number of nearest-neighbour displacements is dictated by the change in valence electrons for the defect species (a.k.a. excess charge). 'Rattling' of all other atomic positions in the supercell is also applied, using Monte-Carlo sampling from a Gaussian distribution, normalised by the bulk bond length of the host. Further details are provided in Mosquera-Lois *et al.*[80] The 'Unperturbed' datapoint, shown with a diamond symbol in faded gold, corresponds to the result obtained with a standard unperturbed supercell relaxation. Datapoints are coloured according to the maximum displacement of any atom relative to the final 'unperturbed' geometry, as an indication of structural disparity.

-1 charge states, which can sometimes be an indicator of spontaneous complex formation (e.g. $N_{Se} \rightarrow N_i$ & N_{Se}), we find that this is not the case here and they are distinct substitutional and interstitial defects. The eigenvalue analysis and plotting tools implemented in `doped`, via the `pydefect` API were highly useful for analysing and identifying shallow states ('perturbed host states') for unfavourable

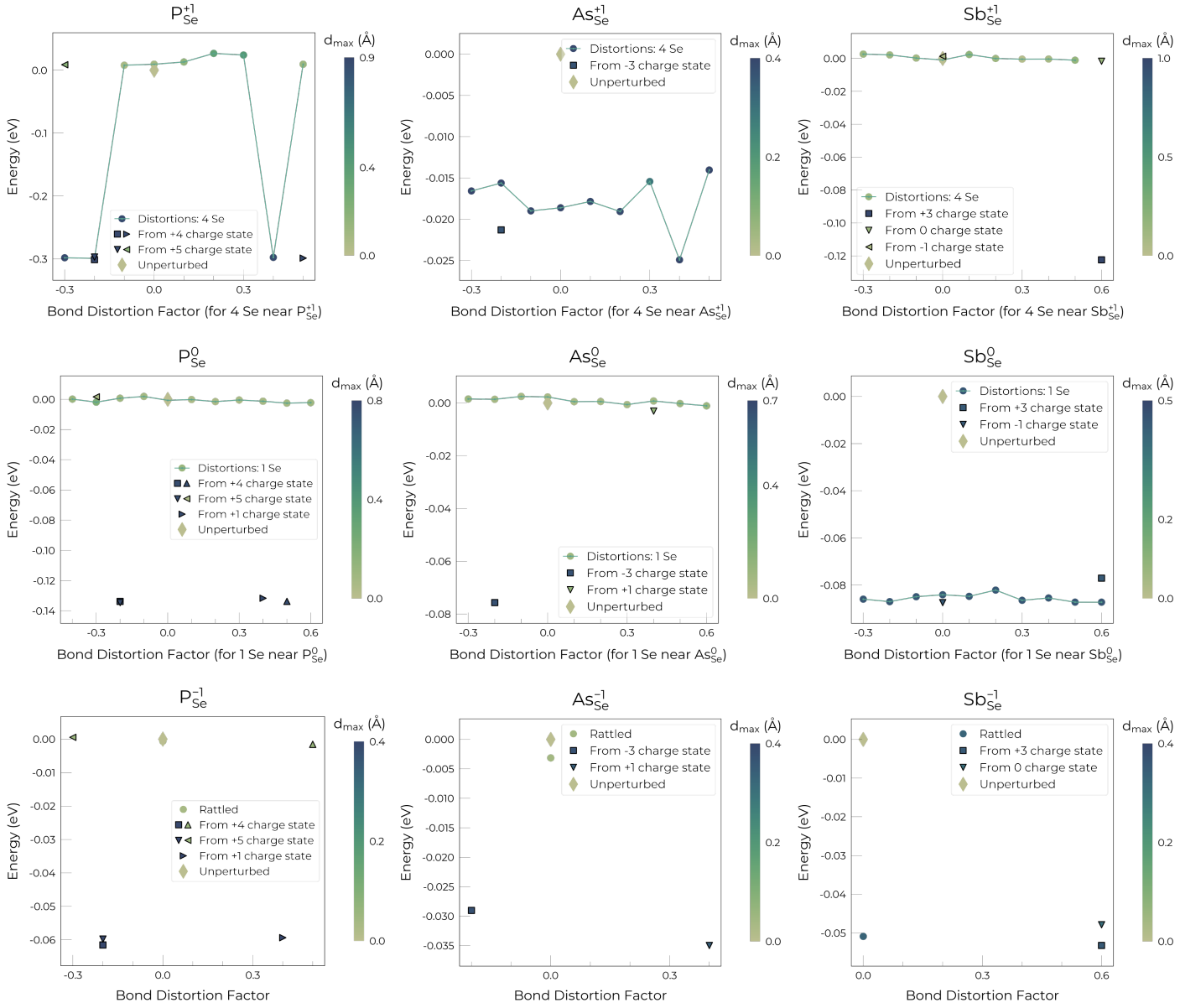

Fig. S19 `ShakeNBreak` results for pnictogen substitutions in *t*-Se, as output directly by the `snb-plot` utility. Above are plotted the final energies obtained from geometry relaxations of trial structures generated for the given defect species by `ShakeNBreak`, as a function of the bond distortion factor (corresponding to the fractional distortion of nearest neighbour bond lengths) applied to the initial geometry. The number of nearest-neighbour displacements is dictated by the change in valence electrons for the defect species (a.k.a. excess charge). 'Rattling' of all other atomic positions in the supercell is also applied, using Monte-Carlo sampling from a Gaussian distribution, normalised by the bulk bond length of the host. Further details are provided in Mosquera-Lois *et al.*[80] The 'Unperturbed' datapoint, shown with a diamond symbol in faded gold, corresponds to the result obtained with a standard unperturbed supercell relaxation. Datapoints are coloured according to the maximum displacement of any atom relative to the final 'unperturbed' geometry, as an indication of structural disparity.

defect charge states.

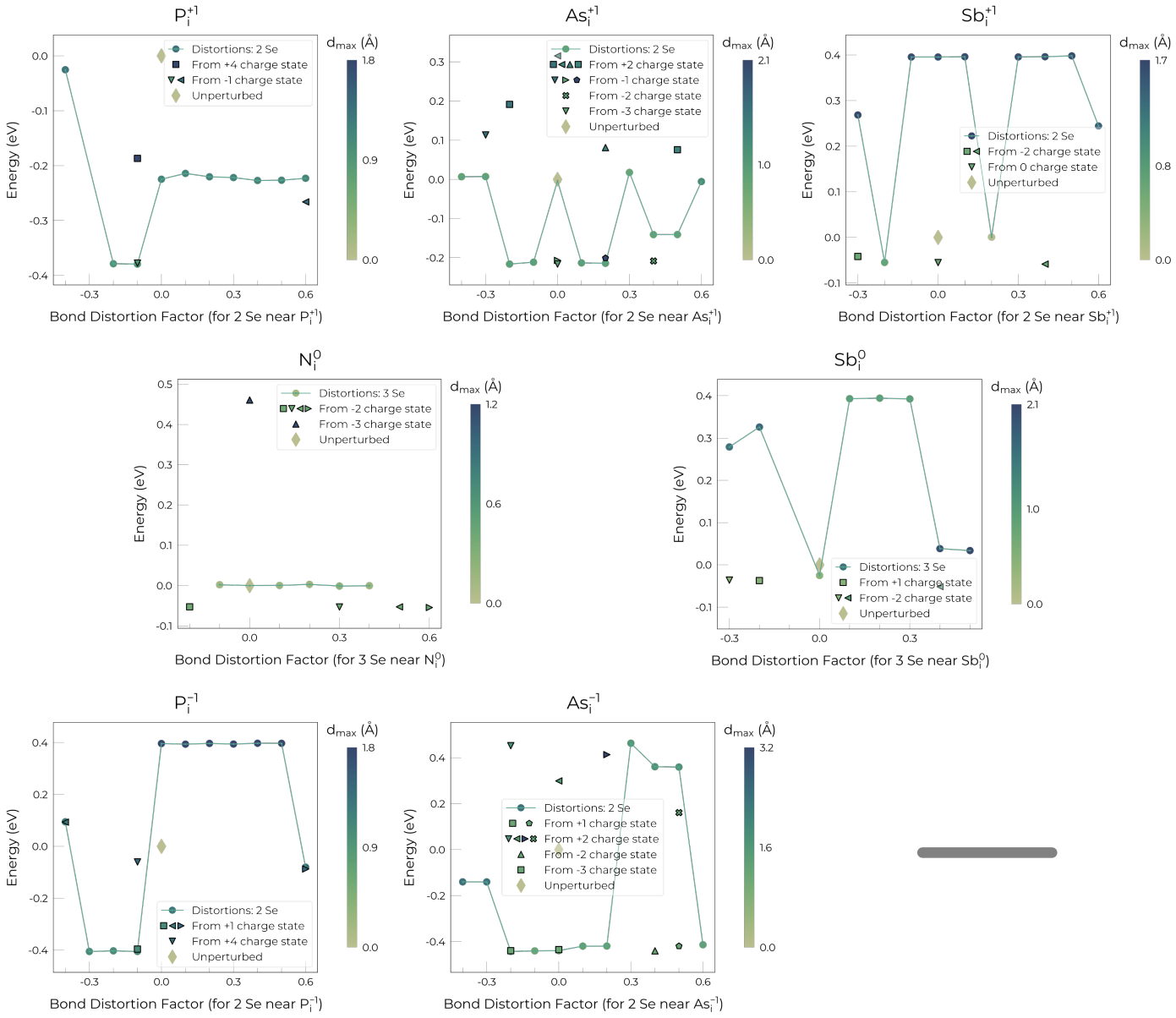

Fig. S20 `ShakeNBreak` results for pnictogen interstitials in *t*-Se, as output directly by the `snb-plot` utility. Above are plotted the final energies obtained from geometry relaxations of trial structures generated for the given defect species by `ShakeNBreak`, as a function of the bond distortion factor (corresponding to the fractional distortion of nearest neighbour bond lengths) applied to the initial geometry. The number of nearest-neighbour displacements is dictated by the change in valence electrons for the defect species (a.k.a. excess charge). 'Rattling' of all other atomic positions in the supercell is also applied, using Monte-Carlo sampling from a Gaussian distribution, normalised by the bulk bond length of the host. Further details are provided in Mosquera-Lois *et al.*[80] The 'Unperturbed' datapoint, shown with a diamond symbol in faded gold, corresponds to the result obtained with a standard unperturbed supercell relaxation. Datapoints are coloured according to the maximum displacement of any atom relative to the final 'unperturbed' geometry, as an indication of structural disparity.

S5.6 Defect Formation Energy Terms

The formation energies of all calculated defects in *t*-Se, including metastable states and showing all terms in the formation energy equation (Eq. (S1)), are provided below. The point group symmetries of the defect site in the pristine bulk crystal, and relaxed defect, corresponding orientational, spin and total degeneracy factors, and site multiplicities in the 3-atom primitive cell of *t*-Se are also provided. Lastly, the resultant calculated charge transition levels, including metastable states, are also tabulated. This data is directly output by the `DefectThermodynamics` functions in `doped`,[71] which are intended to encourage transparency, reproducibility and ready comparison to other calculations, *in addition to* providing calculation data on openly-available online repositories.

Table S5 Formation energies (E^f) of all native point defects in *t*-Se, including metastable states (i.e. $V_{Se}{}^{0*}$), showing all contributions to the formation energy term (Eq. (S1)), as directly output by the `DefectThermodynamics.get_formation_energies` function in `doped`.[71] $\Delta E^f(E_g/2)$ and $\Delta E^f(E_{VBM})$ refer to the formation energies for Fermi levels positioned at the mid-gap ($E_F = 0.86\,eV$) and VBM positions respectively. ΔE_{corr} is the auto-determined estimate of the charge correction error from variance in the electrostatic potentials in the sampling region.[71] The automatically-determined point group symmetries of the defect site *in the pristine bulk crystal* ($Site_{Symm}$) and relaxed defect ($Defect_{Symm}$), corresponding orientational (g_{Orient}), spin (g_{Spin}) and total (g_{Tot}) degeneracy factors, and site multiplicities ("Mult") in the 3-atom primitive cell of *t*-Se are also shown, as directly output by the `DefectThermodynamics.get_symmetries_and_degeneracies` function in `doped`.[71] All energy terms are in units of eV, and ΔE^{raw} refers to the raw DFT energy difference between defect and bulk supercells ($[E_{X,q} - E_H]$ in Eq. (S1))

Defect	q	ΔE^{raw}	qE_{VBM}	$qE_F(E_g/2)$	E_{corr}	ΔE_{corr}	$\Delta E^f(E_g/2)$	$\Delta E^f(E_{VBM})$	$Site_{Symm}$	$Defect_{Symm}$	g_{Orient}	g_{Spin}	g_{Tot}	Mult
V_{Se}	+2	1.122	4.566	1.717	0.278	0.032	3.009	1.292	C_2	C_2	1.0	1	1.0	3.0
	+1	3.677	2.283	0.859	0.011	0.013	2.156	1.297	C_2	C_1	2.0	2	4.0	3.0
	0	6.171	0	0	0	0	1.498	1.498	C_2	C_2	1.0	1[a]	1.0[a]	3.0
	0*	6.198	0	0	0	0	1.525	1.525	C_2	C_2	1.0	1	1.0	3.0
	-1	8.874	-2.283	-0.859	0.244	0.004	1.303	2.162	C_2	C_2	1.0	2	2.0	3.0
	-2	11.861	-4.566	-1.717	0.76	0.016	1.665	3.382	C_2	C_2	1.0	1	1.0	3.0
Se_i	0	-3.85	0	0	0	0	0.823	0.823	C_1	C_2	0.5	1	0.5	6.0
	-1	0.14	-2.283	-0.859	0.046	0.011	1.717	2.576	C_1	C_1	1.0	2	2.0	6.0
	-2	3.543	-4.566	-1.717	0.203	0.115	2.138	3.855	C_1	C_2	0.5	1	0.5	6.0

[a] For the bipolaronic $V_{Se}{}^0$ ground-state, the presence of the triplet state 10 meV higher in energy than the ground-state singlet means that the effective spin (and total) degeneracy is actually $1 + 3\exp(-10\,meV/k_BT)$; see discussion in Mosquera-Lois *et al.*[82] for further details.

Table S6 Charge transition level positions of all native point defects in *t*-Se, including metastable states (i.e. $V_{Se}{}^{0*}$), as directly output by the `DefectThermodynamics.get_transition_levels` function in `doped`.[71] Charge states which are not the ground-state for any Fermi level (i.e. $Se_i{}^{-1}$) are denoted with the $\times$ symbol

Defect	Charges	eV from VBM	In Band Gap?
V_{Se}	(+2/+1)	0.005	True
	(+1/0)	0.200	True
	(0/-1)	0.664	True
	(0*/-1)	0.690	True
	(-1/-2)	1.221	True
Se_i	(0/-2)	1.516	True
	(0/-1$^\times$)	1.751	False
	(-1$^\times$/-2)	1.280	True

As shown below, we note that distinct metastable configurations were identified for a number of defects in *t*-Se, as indicated by the asterisks in the defect names (e.g. $O_i{}^{-2*}$).

As noted in the main text, Refs. 38,47,166 also performed calculations of some point defects in *t*-Se using semi-local DFT, however we obtain qualitatively different results for most defect properties. These major discrepancies are mostly attributed to the use of semi-local DFT functionals, strained supercells, and neglect of van der Waals interactions and structure-searching (Section S5.5),[80,95,96] giving shallower defect levels, unphysically low capture coefficients and negative formation energies. In particular, the self-interstitial is predicted to reside in the inter-chain void, rather than forcing itself within the Se chain as we find here, giving much higher formation

Table S7 Formation energies (E^f) of all extrinsic substitutional defects investigated in t-Se, showing all contributions to the formation energy term (Eq. (S1)), as directly output by the `DefectThermodynamics.get_formation_energies` function in `doped`.[71] $\Delta E^{\mathrm{f}}(E_g/2)$ and $\Delta E^{\mathrm{f}}(E_{\mathrm{VBM}})$ refer to the formation energies for Fermi levels positioned at the mid-gap ($E_F = 0.86\,\mathrm{eV}$) and VBM positions respectively. ΔE_{corr} is the auto-determined estimate of the charge correction error from variance in the electrostatic potentials in the sampling region.[71] $\sum\mu_{\mathrm{ref}}$ and $\sum\mu_{\mathrm{formal}}$ are the summed elemental reference energies and (formal) chemical potentials, which combined give the chemical potential term in Eq. (S1). All energy terms are in units of eV, and ΔE^{raw} refers to the raw DFT energy difference between defect and bulk supercells ($[E_{X,q} - E_H]$ in Eq. (S1))

Defect	q	ΔE^{raw}	qE_{VBM}	$qE_F(E_g/2)$	$\sum\mu_{\mathrm{ref}}$	$\sum\mu_{\mathrm{formal}}$	E_{corr}	ΔE_{corr}	$\Delta E^{\mathrm{f}}(E_g/2)$	$\Delta E^{\mathrm{f}}(E_{\mathrm{VBM}})$
H_{Se}	+1	-0.939	2.283	0.859	-0.764	0	0.042	0.014	1.48	0.622
	0	1.623	0	0	-0.764	0	0	0	0.859	0.859
	-1	4.508	-2.283	-0.859	-0.764	0	0.235	0.007	0.836	1.695
N_{Se}	0	-2.776	0	0	5.504	0	0	0	2.728	2.728
	-1	-0.071	-2.283	-0.859	5.504	0	0.204	0.005	2.495	3.354
P_{Se}	+1	-3.722	2.283	0.859	1.976	0.163	-0.008	0.092	1.552	0.693
	0	-1.115	0	0	1.976	0.163	0	0	1.025	1.025
	-1	1.702	-2.283	-0.859	1.976	0.163	0.124	0.007	0.824	1.683
As_{Se}	+1	-3.22	2.283	0.859	1.17	0.399	0.004	0.088	1.495	0.636
	0	-0.612	0	0	1.17	0.399	0	0	0.957	0.957
	-1	2.195	-2.283	-0.859	1.17	0.399	0.118	0.008	0.741	1.599
Sb_{Se}	+1	-3.106	2.283	0.859	0.78	0.664	0.07	0.087	1.55	0.691
	0	-0.415	0	0	0.78	0.664	0	0	1.029	1.029
	-1	2.544	-2.283	-0.859	0.78	0.664	0.041	0.01	0.887	1.745
O_{Se}	0	-2.507	0	0	2.391	0.483	0	0	0.367	0.367
	-1	0.993	-2.283	-0.859	2.391	0.483	0.214	0.007	0.94	1.799
	-2	4.341	-4.566	-1.717	2.391	0.483	0.563	0.107	1.494	3.212
S_{Se}	0	-0.537	0	0	0.714	0	0	0	0.177	0.177
	-1	3.533	-2.283	-0.859	0.714	0	0.146	0.009	1.252	2.11
	-2	6.956	-4.566	-1.717	0.714	0	0.552	0.03	1.939	3.657
Te_{Se}	0	0.328	0	0	-0.293	0.032	0	0	0.067	0.067
	-1	4.335	-2.283	-0.859	-0.293	0.032	0.068	0.008	1.0	1.859
	-2	7.802	-4.566	-1.717	-0.293	0.032	0.238	0.128	1.497	3.214
F_{Se}	+1	-2.415	2.283	0.859	-1.334	2.352	0.037	0.014	1.78	0.922
	0	0.242	0	0	-1.334	2.352	0	0	1.26	1.26
	-1	3.256	-2.283	-0.859	-1.334	2.352	0.216	0.01	1.348	2.207
Cl_{Se}	+1	-0.443	2.283	0.859	-2.032	0.74	0.093	0.014	1.499	0.641
	0	2.464	0	0	-2.032	0.74	0	0	1.172	1.172
	-1	5.589	-2.283	-0.859	-2.032	0.74	0.071	0.066	1.227	2.085
Br_{Se}	+1	-0.02	2.283	0.859	-2.19	0.205	0.1	0.013	1.237	0.378
	0	2.961	0	0	-2.19	0.205	0	0	0.977	0.977
	-1	6.063	-2.283	-0.859	-2.19	0.205	0.056	0.069	0.993	1.852
I_{Se}	+1	0.159	2.283	0.859	-2.246	0	0.187	0.007	1.241	0.382
	0	3.322	0	0	-2.246	0	0	0	1.076	1.076
	-1	6.46	-2.283	-0.859	-2.246	0	-0.008	0.063	1.064	1.923

energies (and a stronger acceptor character). This is possibly related to issues with interstitial placement algorithms, for which we typically find the Voronoi tessellation scheme (as implemented in `doped`[71] and `PyCDT`[167]) coupled with a structure-searching strategy such as `ShakeNBreak`[81] to be most reliable. The presence of high levels of strain in the supercell, from fixing the lattice vectors to match experiment but using semi-local DFT functionals which poorly reproduce these structural parameters (Table S1), likely also plays a role here — especially given the high sensitivity of the band gap to strain (Fig. S7) and expected sensitivity of defect thermodynamics on strain in this system, as mentioned in the main text. More detailed comparisons with the results of those papers are hindered by the fact that the computational data was not made openly-available.

Table S8 Formation energies (E^f) of all extrinsic interstitial defects investigated in t-Se, showing all contributions to the formation energy term (Eq. (S1)), as directly output by the `DefectThermodynamics.get_formation_energies` function in doped.[71] $\Delta E^{\mathrm{f}}(E_g/2)$ and $\Delta E^{\mathrm{f}}(E_{\mathrm{VBM}})$ refer to the formation energies for Fermi levels positioned at the mid-gap ($E_F = 0.86\,\mathrm{eV}$) and VBM positions respectively. ΔE_{corr} is the auto-determined estimate of the charge correction error from variance in the electrostatic potentials in the sampling region.[71] $\sum\mu_{\mathrm{ref}}$ and $\sum\mu_{\mathrm{formal}}$ are the summed elemental reference energies and (formal) chemical potentials, which combined give the chemical potential term in Eq. (S1). All energy terms are in units of eV, and ΔE^{raw} refers to the raw DFT energy difference between defect and bulk supercells ($[E_{X,q} - E_H]$ in Eq. (S1))

Defect	q	ΔE^{raw}	qE_{VBM}	$qE_F(E_g/2)$	$\sum\mu_{\mathrm{ref}}$	$\sum\mu_{\mathrm{formal}}$	E_{corr}	ΔE_{corr}	$\Delta E^{\mathrm{f}}(E_g/2)$	$\Delta E^{\mathrm{f}}(E_{\mathrm{VBM}})$
H_i	+1	-6.036	2.283	0.859	3.909	0	0.153	0.007	1.168	0.309
	0	-2.942	0	0	3.909	0	0	0	0.967	0.967
	-1	-0.013	-2.283	-0.859	3.909	0	0.123	0.013	0.878	1.737
	-1*	0.11	-2.283	-0.859	3.909	0	0.067	0.036	0.945	1.804
N_i	0	-7.414	0	0	10.177	0	0	0	2.763	2.763
	-1	-4.598	-2.283	-0.859	10.177	0	0.111	0.012	2.549	3.408
	-2	-0.490	-4.566	-1.717	10.177	0.000	0.354	0.119	3.758	5.476
	-3	2.917	-6.848	-2.576	10.177	0.000	1.075	0.077	4.744	7.320
P_i	+3	-12.819	6.848	2.576	6.65	0.163	1.346	0.081	4.765	2.189
	+1	-8.158	2.283	0.859	6.65	0.163	0.234	0.007	2.03	1.172
	0	-5.237	0	0	6.65	0.163	0	0	1.577	1.577
	-1	-2.316	-2.283	-0.859	6.65	0.163	0.013	0.017	1.368	2.227
	-3	4.970	-6.848	-2.576	6.650	0.163	0.606	0.241	2.965	5.541
As_i	+3	-12.244	6.848	2.576	5.844	0.399	1.244	0.204	4.667	2.091
	+1	-7.552	2.283	0.859	5.844	0.399	0.237	0.007	2.069	1.211
	0	-4.608	0	0	5.844	0.399	0	0	1.635	1.635
	-1	-1.657	-2.283	-0.859	5.844	0.399	0.011	0.017	1.456	2.314
	-2	2.285	-4.566	-1.717	5.844	0.399	0.234	0.063	2.479	4.196
	-3	5.661	-6.848	-2.576	5.844	0.399	0.588	0.243	3.067	5.643
Sb_i	+3	-12.24	6.848	2.576	5.454	0.664	1.693	0.045	4.994	2.418
	+1	-7.316	2.283	0.859	5.454	0.664	0.302	0.008	2.245	1.386
	0	-4.186	0	0	5.454	0.664	0	0	1.932	1.932
	-1	-1.19	-2.283	-0.859	5.454	0.664	-0.055	0.018	1.731	2.59
	-2	2.885	-4.566	-1.717	5.454	0.664	0.175	0.032	2.894	4.611
	-3	6.419	-6.848	-2.576	5.454	0.664	0.665	0.057	3.776	6.352
O_i	0	-7.336	0	0	7.064	0.483	0	0	0.211	0.211
	-1	-3.334	-2.283	-0.859	7.064	0.483	0.124	0.008	1.196	2.054
	-2	-0.116	-4.566	-1.717	7.064	0.483	0.351	0.119	1.499	3.217
	0*	-7.211	0	0	7.064	0.483	0	0	0.336	0.336
	-1*	-3.31	-2.283	-0.859	7.064	0.483	0.123	0.009	1.218	2.077
	-2*	0.123	-4.566	-1.717	7.064	0.483	0.445	0.039	1.832	3.549
S_i	0	-4.599	0	0	5.388	0	0	0	0.789	0.789
	-1	-0.589	-2.283	-0.859	5.388	0	0.041	0.012	1.699	2.557
	-2	3.123	-4.566	-1.717	5.388	0	0.181	0.146	2.409	4.126
Te_i	+1	-5.768	2.283	0.859	4.381	0.032	0.318	0.007	2.105	1.246
	0	-3.266	0	0	4.381	0.032	0	0	1.147	1.147
	-1	0.643	-2.283	-0.859	4.381	0.032	-0.028	0.01	1.886	2.744
	-2	4.215	-4.566	-1.717	4.381	0.032	0.115	0.091	2.46	4.177
	-2*	4.831	-4.566	-1.717	4.381	0.032	0.234	0.028	3.194	4.911
F_i	+1	-7.089	2.283	0.859	3.34	2.352	0.138	0.012	1.882	1.023
	0	-4.056	0	0	3.34	2.352	0	0	1.636	1.636
	-1	-1.662	-2.283	-0.859	3.34	2.352	0.114	0.009	1.002	1.861
	+1*	-6.99	2.283	0.859	3.34	2.352	0.165	0.006	2.008	1.149
Cl_i	+1	-4.609	2.283	0.859	2.642	0.74	0.214	0.017	2.128	1.269
	0	-1.595	0	0	2.642	0.74	0	0	1.787	1.787
	-1	1.427	-2.283	-0.859	2.642	0.74	0.065	0.014	1.732	2.59
	+1*	-4.524	2.283	0.859	2.642	0.74	0.214	0.014	2.213	1.354
	0*	-1.301	0	0	2.642	0.74	0	0	2.081	2.081
Br_i	+1	-4.058	2.283	0.859	2.484	0.205	0.224	0.017	1.996	1.137
	0	-0.988	0	0	2.484	0.205	0	0	1.701	1.701
	-1	2.084	-2.283	-0.859	2.484	0.205	0.058	0.012	1.689	2.548
	+1*	-4.007	2.283	0.859	2.484	0.205	0.226	0.013	2.05	1.191
	-1*	2.292	-2.283	-0.859	2.484	0.205	0.047	0.016	1.886	2.744
I_i	+1	-3.638	2.283	0.859	2.427	0	0.305	0.013	2.235	1.377
	0	-0.454	0	0	2.427	0	0	0	1.973	1.973
	-1	2.514	-2.283	-0.859	2.427	0	-0.012	0.012	1.788	2.647

Table S9 Point group symmetries of the defect site *in the pristine bulk crystal* ($Site_{Symm}$) and relaxed defect ($Defect_{Symm}$), corresponding orientational (g_{Orient}), spin (g_{Spin}) and total (g_{Tot}) degeneracy factors, and site multiplicities ("Mult") in the 3-atom primitive cell for extrinsic substitutional defects in *t*-Se, as directly output by the `DefectThermodynamics.get_symmetries_and_degeneracies` function in `doped`.[71]

Defect	q	$Site_{Symm}$	$Defect_{Symm}$	g_{Orient}	g_{Spin}	g_{Tot}	Mult
H_{Se}	+1	C_2	C_1	2.0	1	2.0	3.0
	0	C_2	C_1	2.0	2	4.0	3.0
	-1	C_2	C_1	2.0	1	2.0	3.0
N_{Se}	0	C_2	C_1	2.0	2	4.0	3.0
	-1	C_2	C_2	1.0	1	1.0	3.0
P_{Se}	+1	C_2	C_1	2.0	1	2.0	3.0
	0	C_2	C_1	2.0	2	4.0	3.0
	-1	C_2	C_1	2.0	1	2.0	3.0
As_{Se}	+1	C_2	C_1	2.0	1	2.0	3.0
	0	C_2	C_1	2.0	2	4.0	3.0
	-1	C_2	C_1	2.0	1	2.0	3.0
Sb_{Se}	+1	C_2	C_1	2.0	1	2.0	3.0
	0	C_2	C_1	2.0	2	4.0	3.0
	-1	C_2	C_1	2.0	1	2.0	3.0
O_{Se}	0	C_2	C_2	1.0	1	1.0	3.0
	-1	C_2	C_1	2.0	2	4.0	3.0
	-2	C_2	C_1	2.0	1	2.0	3.0
S_{Se}	0	C_2	C_2	1.0	1	1.0	3.0
	-1	C_2	C_1	2.0	2	4.0	3.0
	-2	C_2	C_1	2.0	1	2.0	3.0
Te_{Se}	0	C_2	C_2	1.0	1	1.0	3.0
	-1	C_2	C_1	2.0	2	4.0	3.0
	-2	C_2	C_1	2.0	1	2.0	3.0
F_{Se}	+1	C_2	C_1	2.0	1	2.0	3.0
	0	C_2	C_1	2.0	2	4.0	3.0
	-1	C_2	C_1	2.0	1	2.0	3.0
Cl_{Se}	+1	C_2	C_1	2.0	1	2.0	3.0
	0	C_2	C_1	2.0	2	4.0	3.0
	-1	C_2	C_1	2.0	1	2.0	3.0
Br_{Se}	+1	C_2	C_1	2.0	1	2.0	3.0
	0	C_2	C_1	2.0	2	4.0	3.0
	-1	C_2	C_1	2.0	1	2.0	3.0
I_{Se}	+1	C_2	C_2	1.0	1	1.0	3.0
	0	C_2	C_1	2.0	2	4.0	3.0
	-1	C_2	C_1	2.0	1	2.0	3.0

Table S10 Point group symmetries of the defect site *in the pristine bulk crystal* ($Site_{Symm}$) and relaxed defect ($Defect_{Symm}$), corresponding orientational (g_{Orient}), spin (g_{Spin}) and total (g_{Tot}) degeneracy factors, and site multiplicities ("Mult") in the 3-atom primitive cell for extrinsic interstitial defects in *t*-Se, as directly output by the `DefectThermodynamics.get_symmetries_and_degeneracies` function in `doped`.[71]

Defect	q	$Site_{Symm}$	$Defect_{Symm}$	g_{Orient}	g_{Spin}	g_{Tot}	Mult
H_i	+1	C_1	C_1	1.0	1	1.0	6.0
	0	C_1	C_1	1.0	2	2.0	6.0
	-1	C_1	C_1	1.0	1	1.0	6.0
	+1*	C_1	C_1	1.0	1	1.0	6.0
N_i	0	C_1	C_1	1.0	2	2.0	6.0
	-1	C_1	C_1	1.0	1	1.0	6.0
	-2	C_1	C_1	1.0	2	2.0	6.0
	-3	C_1	C_1	1.0	1	1.0	6.0
P_i	+1	C_1	C_1	1.0	1	1.0	6.0
	0	C_1	C_1	1.0	2	2.0	6.0
	-1	C_1	C_1	1.0	1	1.0	6.0
	-3	C_1	C_1	1.0	1	1.0	6.0
As_i	+1	C_1	C_1	1.0	1	1.0	6.0
	0	C_1	C_1	1.0	2	2.0	6.0
	-1	C_1	C_1	1.0	1	1.0	6.0
	-2	C_1	C_1	1.0	2	2.0	6.0
	-3	C_1	C_1	1.0	1	1.0	6.0
Sb_i	+1	C_1	C_1	1.0	1	1.0	6.0
	0	C_1	C_1	1.0	2	2.0	6.0
	-1	C_1	C_1	1.0	1	1.0	6.0
	-2	C_1	C_1	1.0	2	2.0	6.0
	-3	C_1	C_1	1.0	1	1.0	6.0
O_i	0	C_1	C_1	1.0	1	1.0	6.0
	-1	C_1	C_1	1.0	2	2.0	6.0
	-2	C_1	C_1	1.0	1	1.0	6.0
	0*	C_1	C_1	1.0	1	1.0	6.0
	-1*	C_1	C_1	1.0	2	2.0	6.0
	-2*	C_1	C_1	1.0	1	1.0	6.0
S_i	0	C_1	C_2	0.5	1	0.5	6.0
	-1	C_1	C_1	1.0	2	2.0	6.0
	-2	C_1	C_1	1.0	1	1.0	6.0
Te_i	+1	C_1	C_1	1.0	2	2.0	6.0
	0	C_1	C_1	1.0	1	1.0	6.0
	-1	C_1	C_1	1.0	2	2.0	6.0
	-2	C_1	C_1	1.0	1	1.0	6.0
	-2*	C_1	C_1	1.0	1	1.0	6.0
F_i	+1	C_1	C_1	1.0	1	1.0	6.0
	0	C_1	C_1	1.0	2	2.0	6.0
	-1	C_1	C_1	1.0	1	1.0	6.0
	+1*	C_1	C_1	1.0	1	1.0	6.0
Cl_i	+1	C_1	C_1	1.0	1	1.0	6.0
	0	C_1	C_1	1.0	2	2.0	6.0
	-1	C_1	C_1	1.0	1	1.0	6.0
	+1*	C_1	C_1	1.0	1	1.0	6.0
	0*	C_1	C_1	1.0	2	2.0	6.0
Br_i	+1	C_1	C_1	1.0	1	1.0	6.0
	0	C_1	C_1	1.0	2	2.0	6.0
	-1	C_1	C_2	0.5	1	0.5	6.0
	+1*	C_1	C_1	1.0	1	1.0	6.0
	-1*	C_1	C_1	1.0	1	1.0	6.0
I_i	+1	C_1	C_1	1.0	1	1.0	6.0
	0	C_1	C_1	1.0	2	2.0	6.0
	-1	C_2	C_2	1.0	1	1.0	3.0

Table S11 Charge transition level positions for extrinsic substitutional defects in *t*-Se, including metastable states, as directly output by the `DefectThermodynamics.get_transition_levels` function in `doped`.[71] Charge states which are not the ground-state for any Fermi level are denoted with the × symbol

Defect	Charges	eV from VBM	In Band Gap?
H_{Se}	(+1/0)	0.237	True
	(0/-1)	0.836	True
N_{Se}	(0/-1)	0.626	True
P_{Se}	(+1/0)	0.332	True
	(0/-1)	0.658	True
As_{Se}	(+1/0)	0.321	True
	(0/-1)	0.642	True
Sb_{Se}	(+1/0)	0.338	True
	(0/-1)	0.716	True
O_{Se}	(0/-1$^{\times}$)	1.432	True
	(-1$^{\times}$/-2)	1.413	True
S_{Se}	(0/-1$^{\times}$)	1.933	False
	(-1$^{\times}$/-2)	1.546	True
Te_{Se}	(0/-1$^{\times}$)	1.792	False
	(-1$^{\times}$/-2)	1.355	True
F_{Se}	(+1/0)	0.339	True
	(0/-1)	0.947	True
Cl_{Se}	(+1/0)	0.531	True
	(0/-1)	0.914	True
Br_{Se}	(+1/0)	0.599	True
	(0/-1)	0.875	True
I_{Se}	(+1/0)	0.693	True
	(0/-1)	0.847	True

Table S12 Charge transition level positions for extrinsic interstitial defects in *t*-Se, including metastable states, as directly output by the `DefectThermodynamics.get_transition_levels` function in doped.[71] Charge states which are not the ground-state for any Fermi level are denoted with the × symbol

Defect	Charges	eV from VBM	In Band Gap?
H_i	(+1/0)	0.658	True
	(0/-1)	0.770	True
	(0/-1*)	0.837	True
N_i	(0/-1)	0.644	True
	(-1/-2×)	2.068	False
	(-2×/-3)	1.845	False
P_i	(+1/0)	0.405	True
	(0/-1)	0.650	True
As_i	(+1/0)	0.424	True
	(0/-1)	0.680	True
	(-1/-2×)	1.882	False
	(-2×/-3)	1.447	True
Sb_i	(+1/0)	0.545	True
	(0/-1)	0.658	True
	(-1/-2×)	2.022	False
	(-2×/-3)	1.741	False
O_i	(0/-1×)	1.843	False
	(0/-1×,*)	1.866	False
	(0*/-1×)	1.718	False
	(0*/-1×,*)	1.741	False
	(-1×/-2)	1.140	True
	(-1×,*/-2)	1.163	True
	(-1×,*/-2*)	1.472	True
	(-1×/-2*)	1.495	True
S_i	(0/-1×)	1.715	True
	(-1×/-2)	1.569	True
Te_i	(+1/0)	-0.099	False
	(0/-1×)	1.597	True
	(-1×/-2)	1.433	True
	(-1×/-2×,*)	2.167	False
F_i	(+1/0×)	0.604	True
	(+1*/0×)	0.478	True
	(0×/-1)	0.225	True
Cl_i	(+1/0)	0.517	True
	(+1*/0)	0.432	True
	(+1/0*)	0.812	True
	(+1*/0*)	0.727	True
	(0/-1)	0.802	True
	(0*/-1)	0.508	True
Br_i	(+1/0)	0.564	True
	(+1*/0)	0.510	True
	(0/-1)	0.847	True
	(0/-1*)	1.043	True
I_i	(+1/0)	0.596	True
	(0/-1)	0.674	True

S6 Urbach Energy Analysis

The Urbach tail is the assumption that in a certain energy range the absorption coefficient is

$$\alpha \simeq \alpha_0 \exp\left(\frac{E-E_1}{E_0}\right)$$

where E_0 is the Urbach energy, and α_0 and E_1 are constants while E is the photon energy. In an absorber where the diffusion length is long compared to the un-depleted absorber thickness the quantum efficiency is

$$\mathrm{EQE} = T\left(1-\exp\left(-\alpha L\right)\right)$$

where T is the transmittance of the window layers. Then we easily arrive at

$$\alpha L = -\ln\left(1-\frac{\mathrm{EQE}}{T}\right)$$

and thus in a certain energy range

$$L\alpha_0 \exp\left(\frac{E-E_1}{E_0}\right) \simeq -\ln\left(1-\frac{\mathrm{EQE}}{T}\right) \implies$$

$$\frac{E-E_1}{E_0} \simeq \ln\left(-\ln\left(1-\frac{\mathrm{EQE}}{T}\right)\right) - \ln\left(L\alpha_0\right)$$

and thus the Urbach energy is obtained as the inverse slope of a plot of $\ln\left(-\ln\left(1-\frac{\mathrm{EQE}}{T}\right)\right)$ as a function of energy. From the EQE-spectrum of the ZnMgO-based selenium solar cell in Figure S21 (not corrected for T, which is probably ~ 0.75), we arrive at $E_0 = \frac{1\,\mathrm{eV}}{22.81789} \simeq 44\,\mathrm{meV}$.

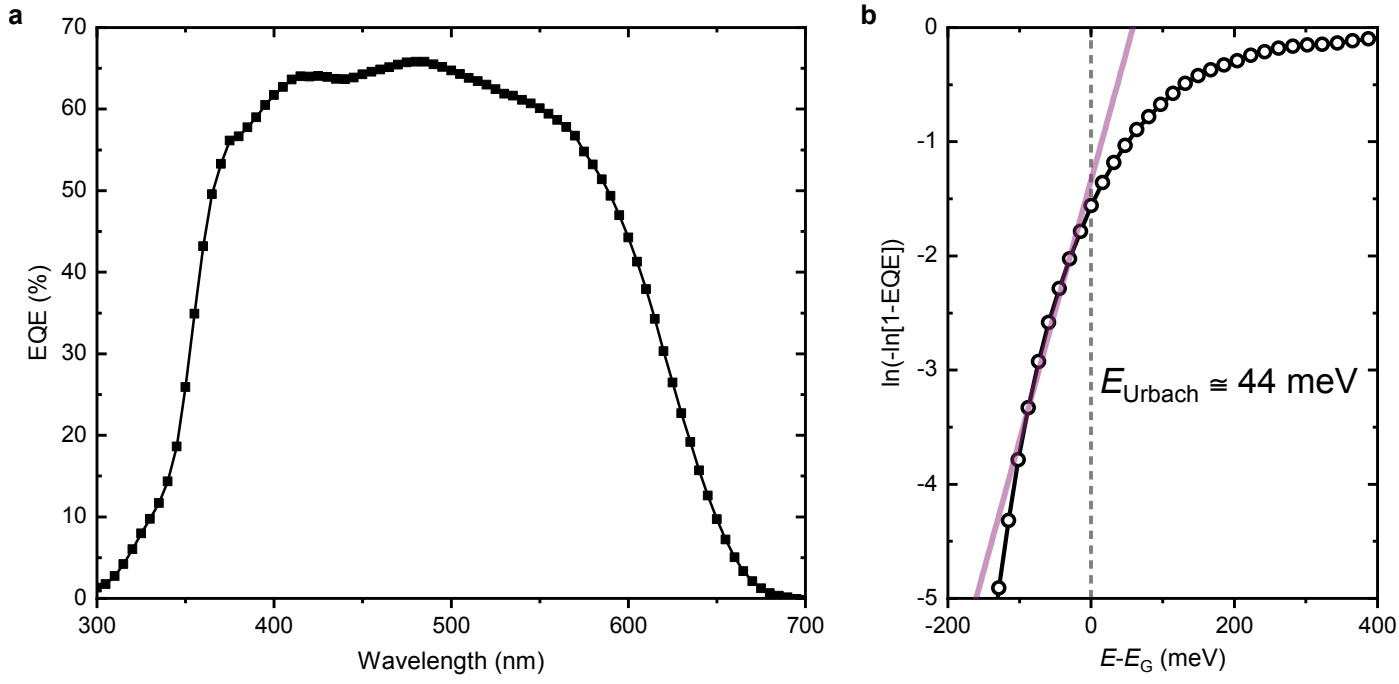

Fig. S21 Calculation of the Urbach energy in selenium. (A) EQE-spectrum of a ZnMgO-based selenium solar cell. (B) Linear fit of the tail of the EQE-spectrum, from which an Urbach energy of $E_{\mathrm{Urbach}} \simeq 44\,\mathrm{meV}$ is calculated. This value is not corrected for the transmittance of the window layers.

S7 Time-of-Flight Secondary Ion Mass Spectrometry (ToF-SIMS)

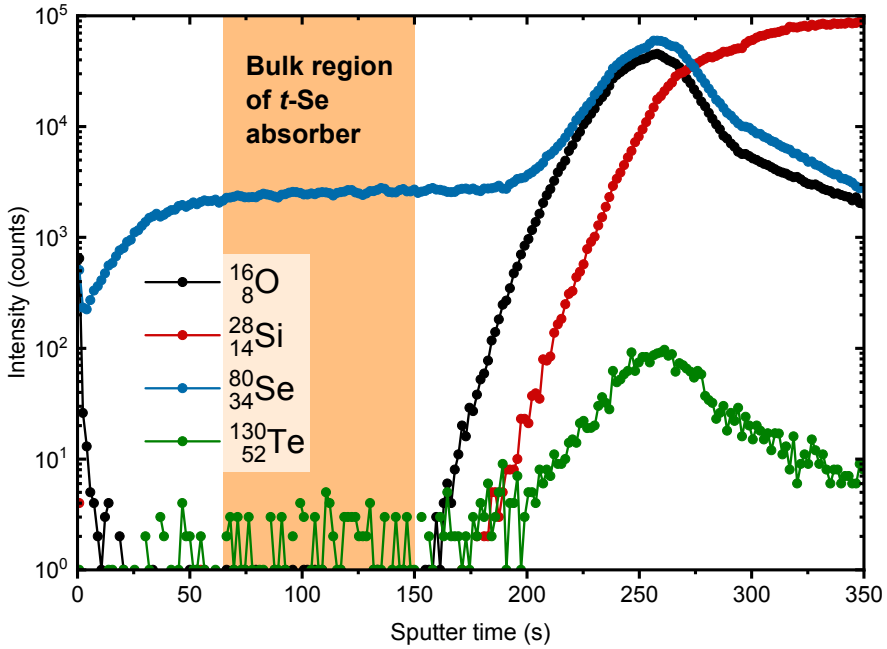

Fig. S22 ToF-SIMS analysis of a 300 nm thick crystalline selenium thin-film synthesized on a silicon substrate. The sputter depth profile shows the integrated signal of $^{16}_{8}$O, $^{28}_{14}$Si, $^{80}_{34}$Se, and $^{130}_{52}$Te. The bulk region of the selenium thin-film (highlighted in orange) has been analysed for impurities in Fig. 7e.

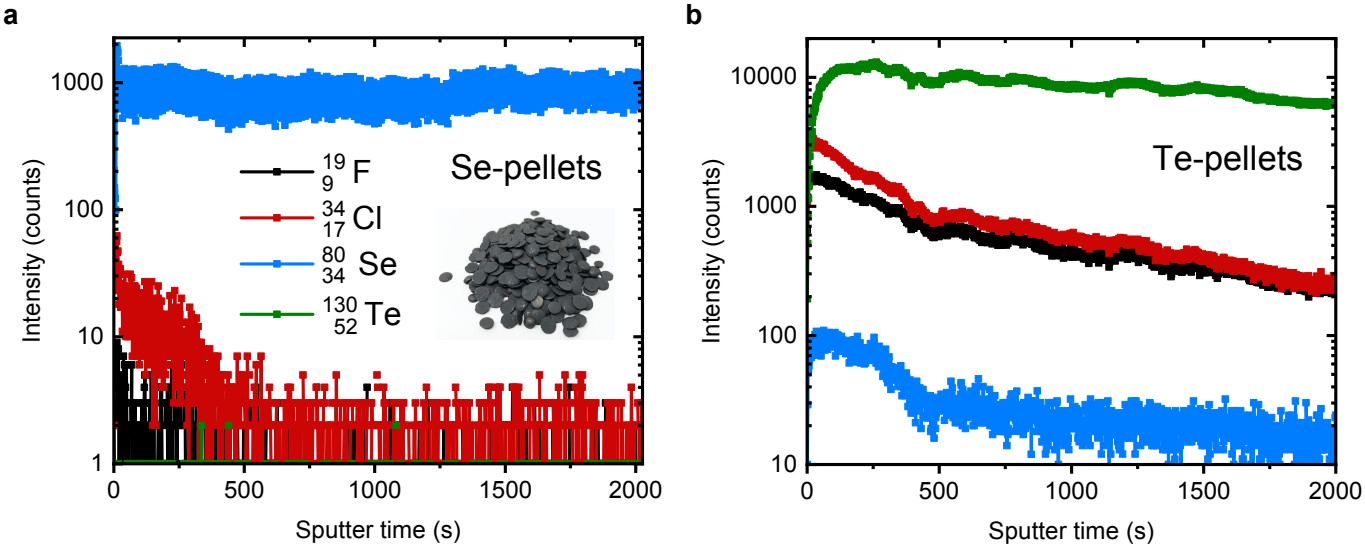

Fig. S23 ToF-SIMS analysis of the (a) selenium, and (b) tellurium pellet source materials. The halogen impurities observed in the bulk of the selenium thin-film are only detected in the tellurium source material.