Se···O/S and S···O Chalcogen Bonds in Small Molecules and Proteins: A Combined CSD and PDB Study

Himansu S. Biswal,*[a, b] Akshay Kumar Sahu,[a, b] Bartomeu Galmés,[c] Antonio Frontera,*[c] and Deepak Chopra*[d]

Dedicated to Prof. Sanjay Wategaonkar on the occasion of his 65th birthday and superannuation and Prof T. N. Guru Row on the occasion of his 70th birthday.

The importance of selenium-centered noncovalent chalcogen bonds represented as Se···A (A = O/S) has been explored for short directional contacts in small molecules and proteins. In addition, S···O centered contacts have been analyzed. Computational analyses involving the quantitative assessment of the associated energetics, the molecular electrostatic potentials (MEP), and electron density derived topological parameters, namely, quantum theory of atom in molecules (QTAIM) analyses, and NBO (natural bond orbital) based calculations, have been performed to unequivocally establish the strength, stability, and attractive role of chalcogen bonds in the solid-state. This investigation has been performed in molecules from both the Cambridge Structural Database (CSD) and Protein Data Bank (PDB). Thus futuristic materials may be designed keeping in mind the significance of these interactions, including their relevance in biology.

Introduction

The relevance of σ-hole centered noncovalent interactions is now well recognized in the literature.[1] Recent investigations establish the importance of halogen,[2–3] chalcogen,[4] and tetrel bonding[5] in molecules. This also has implications in reactivity profiles, namely catalysis in these classes of compounds.[6] Selenium is an important main group element and is present in molecules that find application in materials,[7] in drugs,[8] and in proteins (containing selenocysteine)[9] as well. Over the last few decades, several investigations on chalcogen bonded systems have been performed. ^{17}O and ^{77}Se NMR spectroscopic measurements have been utilized to probe intramolecular Se···O contact in substituted benzeneselenyl derivatives.[10] The concept of molecular balances has been utilized to experimentally evaluate the chalcogen bonding interactions, wherein the role of solvent effects are important.[11] *In situ cryo-crystallization* technique has been utilized to probe the role of different chalcogen-centered interactions that are present in the crystal packing of Ph_2Se and Ph_2Te.[12] The directional preferences of non-covalent contacts with sulfur, which also suggested the non-spherical nature of sulfur atom, were established by studies done via a statistical analysis of the molecules from the Cambridge Structural Database.[13] Nonbonded sulfur/ selenium-oxygen interactions in thiazole and selenazole nucleosides have been evaluated from computational studies.[14] This was followed by the work of Iwaoka and co-workers who have investigated S···O contacts, and suggested that directionality plays a key role and their relevance in protein structures was also suggested.[15] Quantitative molecular orbital analysis in chalcogen bonded $D_2Ch\cdots A^-$ [Ch = chalcogen, D, A = F, Cl, Br] complexes, establish that there exists significant covalent character on account of the strong HOMO-LUMO interactions along with structural distortions arising from charge transfer from A- to D_2Ch.[16] In-depth characterization of chalcogen interactions, in terms of geometrical changes at donor and acceptor atoms, including detailed topological analysis have also been done as well.[17] Keeping in mind the features mentioned above, we have undertaken to perform a quantum mechanical study of non-covalent interactions involving Ch···Ch contacts, namely, Se···S, Se···O and S···O in both small molecules (from CSD)[18] and proteins (from PDB).[19] The contacts considered in the study involve the contact distance as the sum of the van der Waals radii of participating atoms with the directionality being near linear (≥ 160°). DFT-based calculations were performed, and the results establish the stabilizing nature of these interactions, with the strength being "weak" to "moderately strong." Furthermore, quantum theory of "atoms in molecules" (QTAIM),[20] noncovalent interaction (NCI) index,[21] and natural bond orbital (NBO) analyses[22] have been performed to

[a] Prof. Dr. H. S. Biswal, A. K. Sahu
School of Chemical Sciences
National Institute of Science Education and Research (NISER)
PO-Bhimpur-Padanpur, Via-Jatni, District-Khurda, Bhubaneswar (India)
E-mail: himansu@niser.ac.in
[b] Prof. Dr. H. S. Biswal, A. K. Sahu
Training School Complex, Homi Bhabha National Institute
Anushakti Nagar, Mumbai 400094 (India)
[c] B. Galmés, Prof. Dr. A. Frontera
Departament de Química, Universitat de les Illes Balears
Crta de Valldemossa km 7.5, 07122 Palma de Mallorca (Spain)
E-mail: toni.frontera@uib.es
[d] Prof. Dr. D. Chopra
Department of Chemistry
Indian Institute of Science Education and Research Bhopal
Bhopal Bypass Road, Bhauri, Bhopal 462066, Madhya Pradesh (India)
E-mail: dchopra@iiserb.ac.in

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cbic.202100498

characterize these noncovalent interactions' bonding features and provide quantitative insights into the role of orbital contributions towards the overall stabilization of the formed complexes.

Results and Discussion

Investigation of intermolecular interactions from the crystal structures in CSD

A search of organic X-ray structures from the CSD (updated May 2021) was initially performed, limiting the searches to those CIFs containing 3D coordinates and "no errors". Moreover, only those X-ray structures with Ch···Ch (Ch = Se, S, and O) distances $\leq \Sigma R_{vdw} + 0.2$ Å and Y–Ch···Ch (Y = any atom) angle $\geq 140^{\circ}$ were selected and further inspected manually and statistically (see ESI for further details) and the statistical plots are represented in Figure 1. From the above analyses, we found 6992, 514, and 340 hits for the S···O, Se···O and Se···S interactions, respectively. A detailed geometrical distribution for all Ch···Ch interactions is represented in Figure 1. An approximate Gaussian like distribution was observed for S/Se···O, showing a maximum population near 3.3 Å (Figure 1(a),(d)). The distance and angle distribution were simultaneously plotted in a radial density plot (Figure 1(c),(f)) evidencing that the maximum density is found around 165° for both O···S and O···Se interactions [∠O···S–C (Figure 1(b)), and ∠O···Se–C (Figure 1(e))].

For the Se···S interaction, the geometrical distribution is given in Figures 1(g)–(h). The radial distribution (Figure 1(i)) is

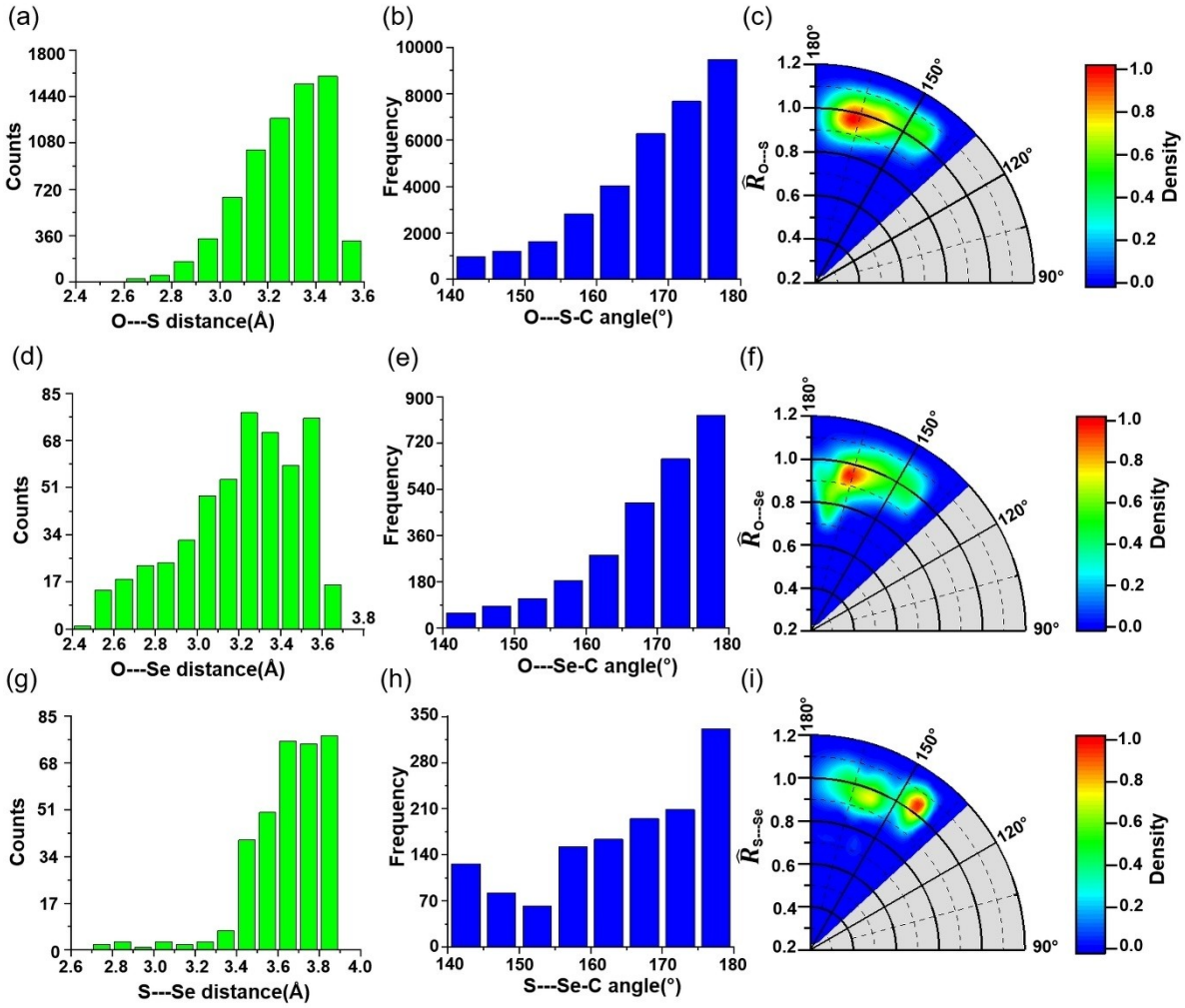
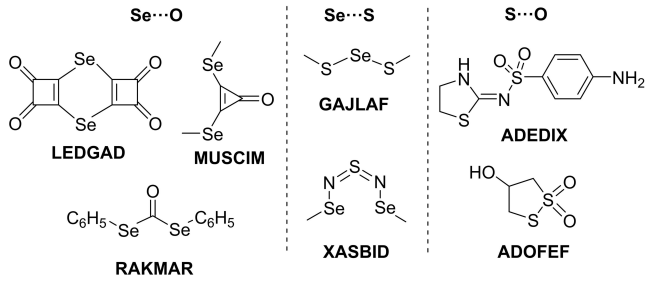


Figure 1. (a) Distribution of distance between S and O, (b) cone corrected O···S–C angle distribution where the frequency is generated by counts divided by respective sin(θ), (c) radial density distribution plot where the distance between O and S is normalized with respect to the sum of their van der Waals radii ($\widehat{R}_{S\cdots O}$), the density bar normalized to the maximum count, (d) Distribution of distances between Se and O, (e) cone corrected O···Se–C angle distribution where the frequency is generated by counts divided by respective sin(θ), (f) radial density distribution plot where the distance between O and Se is normalized with respect to their sum of the van der Waals radii ($\widehat{R}_{Se\cdots O}$), the density bar normalized to the maximum count, (g) Distribution of distances between Se and S, (h) cone corrected S···Se–C angle distribution where the frequency is generated by counts divided by respective sin(θ), (i) radial density distribution plot where the distance between S and Se is normalized with respect to their sum of the van der Waals radii ($\widehat{R}_{Se\cdots O}$), the density bar normalized to the maximum count.

slightly different, showing two different zones where the density of hits is important. One intense region is found at around 160° (∠S···Se–C) similar to S···O and Se···O interactions and another one (the maximum) at a lower angle (~140°). This different behavior is likely due to a mixing of two possible situations, one where the *Se* is the *σ-hole donor* and the *S-atom* is the *σ-lump donor* and another one that corresponds to the reverse combination, where the *S* is the *σ-hole donor* and the *Se-atom* is the *σ-lump donor*.

From the CSD searches, we have selected several hits with high directionality and short Ch···Ch distance to analyze them energetically using DFT calculations, MEP analysis, and a combination of QTAIM and NCIplot index computational tools (see Scheme 1).Tables S1–S3 (ESI) reflect the interaction energies of ChB dimers extracted from the X-ray structures, the MEP values at the Se/S ChB donors and O/S ChB acceptors, and the magnitudes of the electron density at the bond CPs connecting both the Ch atoms. From the examination of the results, several issues can be extracted. First, the interaction energies in neutral systems range from −3.2 to −14.9 kcal/mol. However, most of the Ch···Ch dimers present additional contacts, mostly H-bonds (see QTAIM analyses represented in Figures S1 to S3). Second, the MEP values at the selenium's σ-holes are quite different depending on the structure, ranging from just 1.5 kcal/mol in **DUNSAG** to 115 kcal/mol in **FIDZUG** (due to its cationic nature). As expected, the MEP values at the Se σ-holes are more positive than those at the S σ-holes in most cases. Finally, the MEP values at the O-atoms are more negative than those at the S-atoms, thus anticipating stronger Se···O interactions.

We have selected some representative hits of each CSD search based on those structures where the dimers present a minimum number of ancillary interactions (Scheme 1). Consequently, in those hits the ChB is most likely the dominant interaction. Figure 2 shows the MEP surfaces of the three selected Se-bond donor molecules that participate in Se···O ChBs (see ESI for computational details). Moreover, Figure 3



Scheme 1. Chemical diagrams of the selected hits from the CSD searches with the indication of the CSD codes.

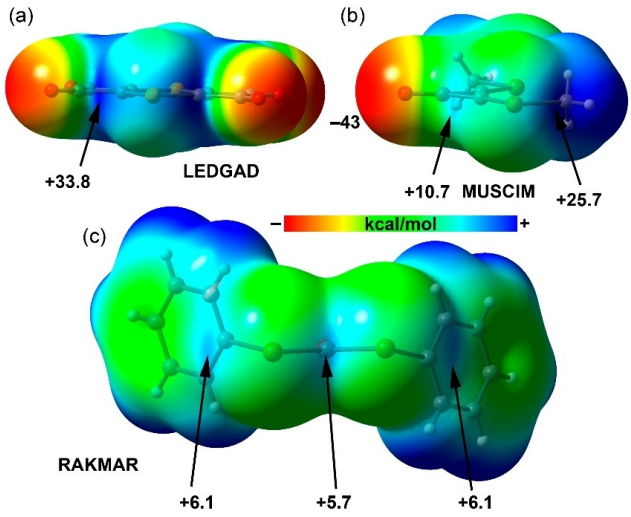


Figure 2. MEP surfaces of refcodes **LEDGAD** (a), **MUSCIM** (b), and **RAKMAR** (c). The MEP values at the σ-holes are indicated in kcal/mol. See Scheme 1 for the chemical diagrams of the compounds.

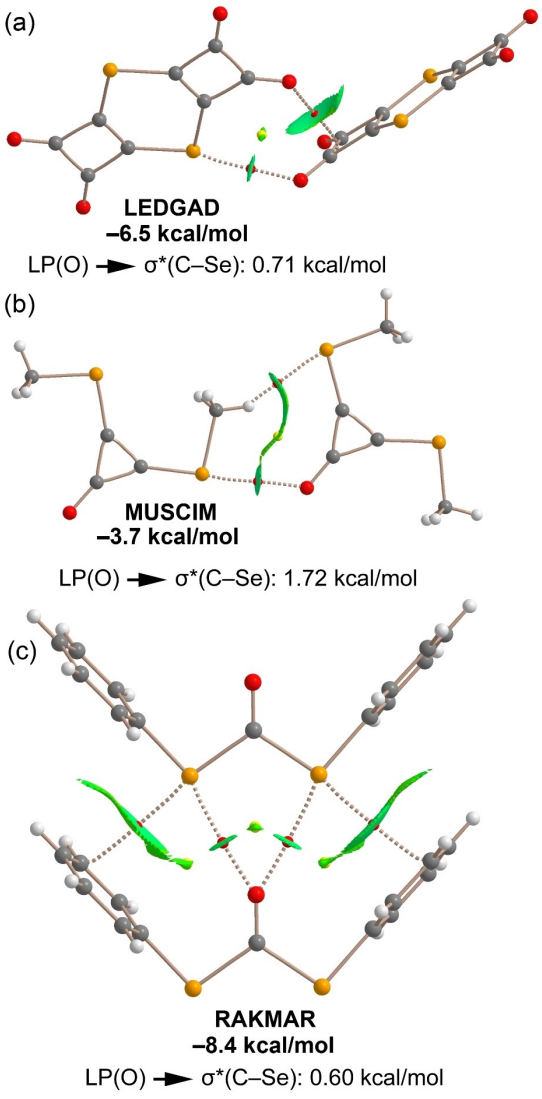


Figure 3. Combined QTAIM (bond and red critical points represented as red and ring spheres, respectively) and NCIPlot (ρ cut-off 0.04 a.u., $|RGD| = 0.5$, color range $-0.04 < (\text{sign}\lambda_2)\,\rho < 0.04$ a.u. of refcodes **LEDGAD** (a), **MUSCIM** (b) and **RAKMAR** (c). Only intermolecular interactions are represented for clarity. The orbital stabilization energies obtained using the second-order perturbation analysis are also indicated along with the orbital involved. LP stands for lone pair and σ* for antibonding sigma orbital.

shows the QTAIM/NCIplot index analyses of the dimers extracted from their solid-state structures. The CSD reference code **LEDGAD** presents a six-membered diselenacycle in the structure, where each Se-atom exhibits two symmetric σ-holes (+33.8 kcal/mol, see Figure 2a). In the solid-state, it forms dimers stabilized by a combination of Se···O and O(lp)···π interactions, both characterized by the corresponding bond CPs and green NCI plot isosurfaces (see Figure 3a). The dimerization energy is moderately strong (−6.5 kcal/mol), confirming the importance of the ChB contact in the crystal packing. The O(lp)···π interacts between the electrostatically more negative oxygen with the electrostatically positive four-membered ring indicating the attractive nature of the interaction in the NCI plot.

In the case of bis(methylseleno)cyclopropenone (**MUSCIM**), the Se-atom exhibits two different σ-holes, the one opposite to the cyclopropenone ring being the most positive (+25.7 kcal/mol see Figure 2b). Interestingly, the Se···O interactions observed in the solid-state of **MUSCIM** corresponds to the most positive σ-hole that interacts with the O-atom of the adjacent molecule (see Figure 2b). The dimerization energy of **MUSCIM** is −3.7 kcal/mol, which is smaller than that in LEDGAD and is likely due to the strongest contribution of the O···π interaction in the latter. Finally, the MEP surface of Se,Se′-diphenyl carbonodiselenoate (**RAKMAR**, see Figure 2c) structure presents *three σ-holes* (MESP in the range of 5.7–6.1 kcal/mol, two σ-holes merged into one) that are similar in energy. All of them participate in the dimer formation, establishing the bifurcated nature of Se···O and two Se···π ChBs (see Figure 2c). These interactions are characterized by bond CPs (red spheres) and bond paths connecting the Se-atoms to the O-atom of the adjacent molecule and two C-atoms of the phenyl rings. The green NCI isosurfaces establish the attractive nature of the Se···O and Se···π interaction. The interaction energy is large (−8.4 kcal/mol) due to the cooperative formation of these three concurrent ChB interactions. For the sake of comparison, we have computed at the same level of theory, two prototypical H-bonded complexes, that are the water dimer and the formamide dimer. For the water dimer (only one H-bond, see Figure S4a) the interaction energy is −5.7 kcal/mol and for the self-assembled dimer of formamide (two H-bonds, see Figure S4b), the interaction energy is −14.4 kcal/mol. Therefore, the ChBs energies reported for the Se···O contacts are more or less comparable to those of the H-bonds.

In order to further evidence the existence of chalcogen bonding interaction and the importance of orbital effects, we have carried out the natural bond orbital analysis of the homodimers represented in Figure 3. We have focused the study on the second-order perturbation analysis since it is useful to account for orbital donor-acceptor interactions. The results ($E^{(2)}$ energies) are indicated in Figure 3, showing in all three cases an electron donation from the lone pair (LP) at the O-atom to the antibonding (Se–C) orbital with concomitant stabilization energies that range from 0.60 kcal/mol in **RAKMAR** to 1.72 kcal/mol in MUSCIM. The existence of such electron donation is typical in σ-hole interactions. For comparison purposes, we have also computed the $E^{(2)}$ energies in the prototypical H-bonded dimers commented above (see Figure S4), which are significantly larger than those obtained for the ChBs reported herein. For instance, for the formamide dimer, each N–H···O H-bonds exhibits an electron donation from the lone pair (LP) at the O-atom to the antibonding (N–H) orbital with a consequent stabilization of the system in $E^{(2)} = -12.3$ kcal/mol.

Figure 4 shows the MEP surfaces of the two Ch-bond donors selected to highlight the Se···S interactions. For bis(methylmercapto)selenane (**GAJLAF**, Figure 4a), two symmetric σ-holes are present in the Se atom.

This molecule in the solid-state forms self-assembled dimers (see Figure 5a) stabilized by two Se···S contacts characterized by the corresponding bond CP and green NCI plot index isosurface. The dimerization energy is −3.9 kcal/mol, thus revealing that each ChB is modest (around −2 kcal/mol), in line with the moderate σ-hole MEP value. The second selected example corresponds to(*Z*,*Z*)-1,5-diphenyl-3-thia-1,5-diselena-2,4-diaza-2,3-pentadiene (refcode **XASBID**)that exhibits one large and positive σ-hole opposite to the Se–N bond (+15 kcal/mol, see Figure 4b). This compound forms infinite 1D supramolecular chains in the solid-state. One dimer extracted from this chain is represented in Figure 5b. It forms two symmetrically equivalent ChBs, and the dimerization energy is −7.6 kcal/mol, which is stronger than **GAJLAF**. This is in accordance with the more intense σ-hole observed in the Se-atom of **XASBID**. The presence of green isosurfaces in both the molecules **GAJLAF** and **XASBID** establishes the attractive nature of Se···S contacts in the solid-state. In contrast, the orbital contribution [LP-(S)→σ*(Se–Y), Y=N,S] is larger in **GAJLAF** (3.40 kcal/mol) than in **XASBID** (0.64 kcal/mol), likely due to the better directionality of the interaction in GAJLAF, allowing a better orbital overlap. In case of GAJLAF dimer, the interaction energy using the highly

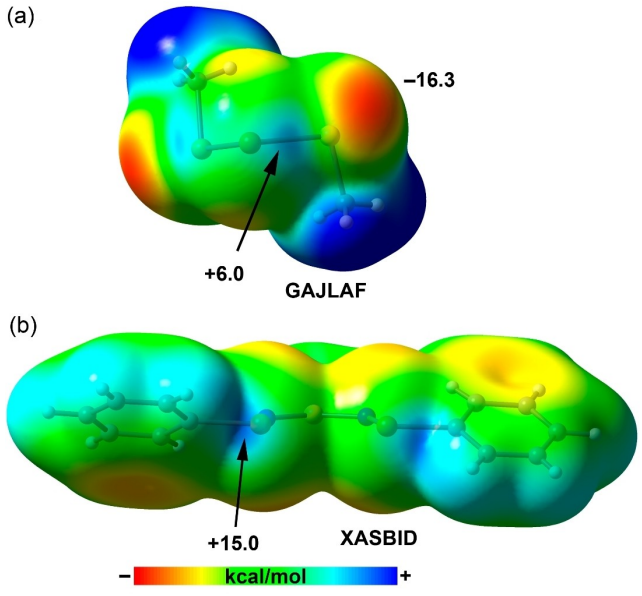


Figure 4. MEP surfaces of refcodes **GAJLAF** (a) and **XASBID** (b). The MEP values at the σ-holes are indicated in kcal/mol. See Scheme 1 for the chemical diagrams of the compounds

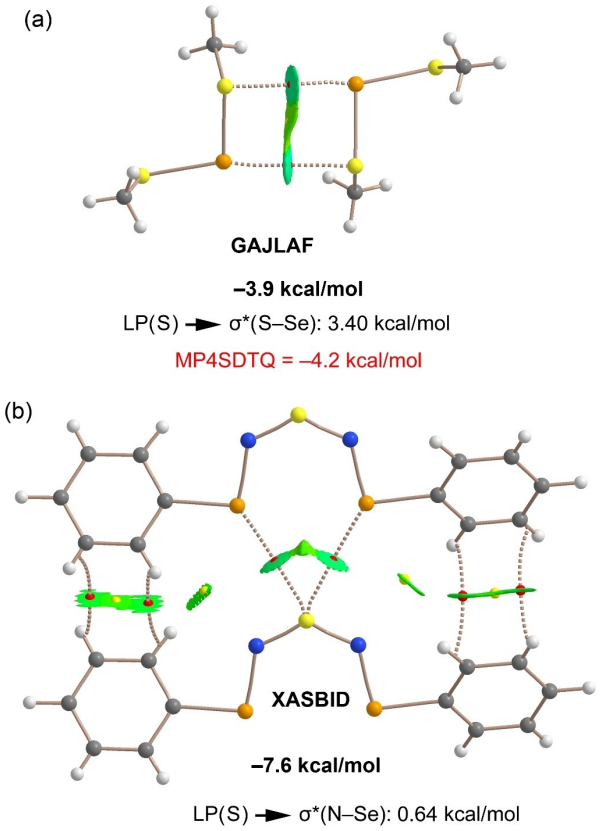


Figure 5. Combined QTAIM (bond and red critical points represented as red and ring spheres, respectively) and NCIPlot (ρ cut-off 0.04 a.u., $|RGD| = 0.5$, color range $-0.04 < (sign\lambda_2)\,\rho < 0.04$ a.u. of refcodes **GAJLAF** (a) and **XASBID** (b). Only intermolecular interactions are represented for clarity. The orbital stabilization energies obtained using the second-order perturbation analysis are also indicated along with the orbital involved. LP stands for lone pair and σ* for antibonding sigma orbital.

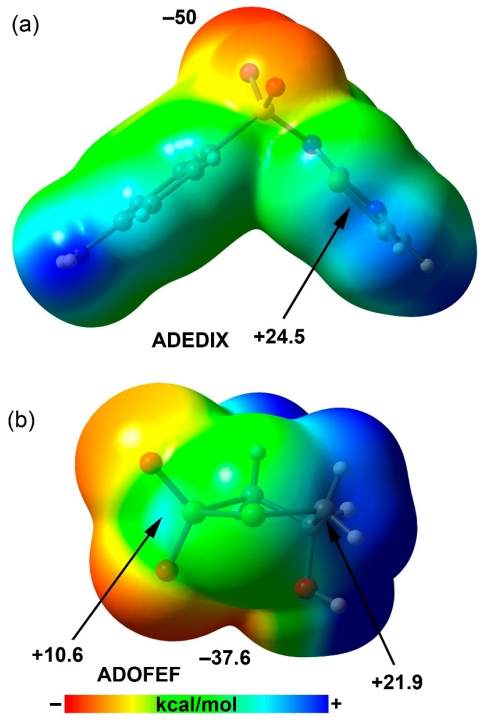


Figure 6. MEP surfaces of refcodes **ADEDIX** (a) and **ADOFEF** (b). The MEP values at the σ-holes are indicated in kcal/mol. See Scheme 1 for the chemical diagrams of the compounds.

correlated MP4(SDTQ) method has been also calculated in order to give reliability to the DFT method. It can be observed that the difference between both methods is very small (0.3 kcal/mol) thus supporting the PBE0 energies included in the Figures and Tables S1 to S4 (ESI).

Figure 6 shows the MEP surfaces of the two Ch-bond donors used to highlight the S···O interactions. For sulfathiazole (**ADEDIX**, Figure 6a). The S-atom exhibits a σ-hole that is significantly positive (+24.5 kcal/mol). This molecule in the solid-state forms self-assembled dimers (see Figure 7a) stabilized by two S···O interactions and one S···S contact that is expected to be very weak due to its poor directionality. Each interaction is characterized by the corresponding bond CP and green NCI plot index isosurface that establishes the attractive nature of the S···O interaction. The dimerization energy is large −12.3 kcal/mol, in accordance with the large MEP value at the σ-hole. The second selected example corresponds to 4-hydroxy-1,2-dithiolane 1,1-dioxide (refcode **ADOFEF**), exhibiting two quite different σ-holes (+21.9 and +10.6 kcal/mol, see Figure 6b), where the most positive one is the opposite to the -SO_2 electron-withdrawing group. This compound also forms self-assembled dimers in the solid-state (see Figure 7b) where two

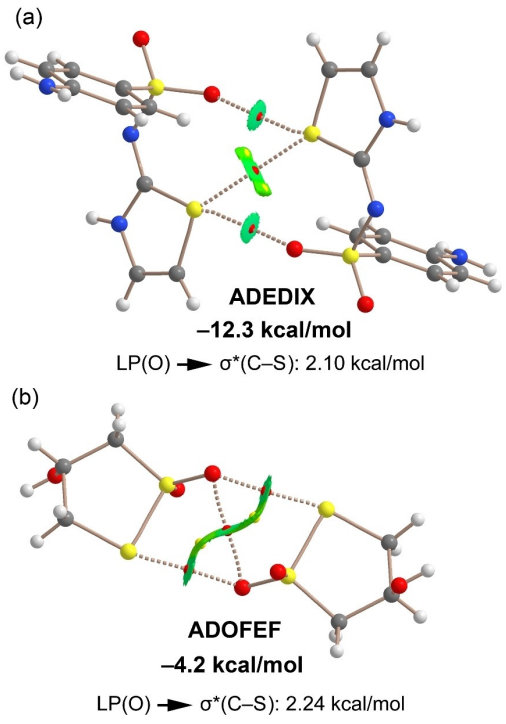


Figure 7. Combined QTAIM (bond and red critical points represented as red and ring spheres, respectively) and NCIPlot (ρ cut-off 0.04 a.u., $|RGD| = 0.5$, color range $-0.04 < (sign\lambda_2)\,\rho < 0.04$ a.u. of refcodes **ADEDIX** (a) and **ADOFEF** (b). Only intermolecular interactions are represented for clarity. The orbital stabilization energies obtained by using the second-order perturbation analysis are also indicated along with the orbital involved. LP stands for lone pair and σ* for sigma antibonding orbital.

S···O and one O···O contacts are present. In this case, the dimerization energy is modest (−4.2 kcal/mol) because the relatively less positive σ-hole at the S-atom is involved in the interaction. However, the orbital contribution is similar in both compounds that exhibit electron donation from the LP orbital at the O-atom to the antibonding σ*(S–C) orbital with concomitant stabilization energies of 2.10 and 2.25 kcal/mol for **ADEDIX** and **ADOFEF**, respectively.

In order to further assess the role of the ChBs in the crystal packing, we have fully optimized two structures, which are represented in Figure 8 along with the dimerization energies. The optimized dimer of refcode **GAJLAF** shows a very similar geometry compared to the one retrieved from the X-ray structure (see Figure 5a). The Se···S distance is slightly longer in the optimized structure (3.69 Å) than the experimental one (3.50 Å) likely due to packing effects. The dimerization energy (−4.2 kcal/mol) is also very similar to the one obtained using the X-ray coordinates (−3.9 kcal/mol). Similar results have been found for MUSCIM dimer. In this case the fully optimized geometry exhibits a shorter H···O distance and longer Se···O distance compared to the X-ray dimer (values in red).

The dimerization energy is identical to the one obtained for the X-ray structure, thus suggesting a compensating effect between the ChB and HB in the theoretical geometry with respect to the experimental one. The fact that the assemblies and ChBs are similar in the fully optimized and X-ray geometries supports the relevance and structure guiding role of the ChBs in the solid state.

Investigation of intermolecular interactions from the crystal structures in PDB

The home-written python code was used to analyze the 13178 PDB files downloaded from the RCSB website[11] (October 2020 release) [Resolution ≤ 1.5 Å]. The criteria used to look into the interactions in PDB files were as follows: (a) the distances between S and O: 2 Å < $d_{S\cdots O}$ ≤ 3.32 Å, Se and O: 2 Å < $d_{Se\cdots O}$ ≤ 3.42 Å, Se and S: 2.5 Å < $d_{Se\cdots S}$ ≤ 4.07 Å; (b) the angle was set to 130° ≤ ∠C–Se···O/S, ∠C–S···O ≤ 180°.

The lower limit of the distance was kept at 2 Å for S···O, Se···O and 2.5 Å for Se···S, which is greater than their respective sum of their covalent radius. The covalent radius for O, S, and Se was considered to be 0.66 Å, 1.05 Å, and 1.20 Å, respectively. The upper limits of the distances for S···O and Se···O were less than their respective sum of the van der Waals radius,[23] whereas the distance for Se···S is kept less than 4.07 Å, which is 10% higher than their sum of the van der Waals radii. The secondary structure analysis of the interacting residues was performed using the software STRIDE.[24] The PDB files with the angles ∠C–S···O, ∠S–Se···O/S ≥ 160° were chosen for model systems. We considered i−1, i, i+1 residues from the selected PDB files where i is the interacting residue. The C_α position of the C-terminus (i−1) and N-terminus (i+1) residue was replaced with a methyl group.

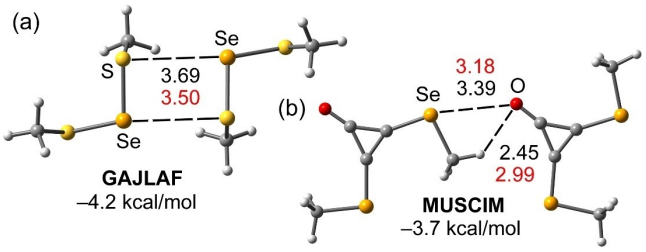


Figure 8. Fully optimized geometries of **GAJLAF** (a) and **MUSCIM** (b) dimers at the PBE0-D3/def2-TZVP level of theory. Distances in Å, in black the optimized ones and in red the experimental ones.

A detailed geometrical distribution for S···O and Se···O interactions are represented in Figure 9. A Gaussian like distribution was observed for S···O, which shows that the maximum population is present near 3.3 Å (Figure 9a). The distance and angle distribution were simultaneously plotted in a radial density plot (Figure 9b). The presence of maximum density around 165° for ∠O···S–C confirms the directional nature of the interaction. From analysis of the secondary structures, it was observed that these interactions are present in all secondary structures. However, the O atoms preferred the turn (25.37%) secondary structure followed by strand (24.32%), coil (22.41%), α-helix (19.63%), 3_{10}-helix (6.54%), and bridge (1.73%) (Figure 9c) structures, respectively.

From the above analysis, we found 1689, 50, and 10 hits for the S···O, Se···O and Se···S interactions, respectively.

Similarly, the Se···O distance is maximum near 3.3 Å (Figure 8d). Nevertheless, the maximum density around 175° for ∠O···Se–C in radial distribution plot showed the highly directional nature of the interaction (Figure 9e). The O atom of O···Se interactions is present in all secondary structures as well but preferred the coil (30.6%) secondary structure followed by strand (22.4%), α-helix (18.4%), 3_{10}-helix (14.3%), turn (10.2%) and bridge (4.1%) (Figure 9f) structures respectively.

Tables S4 and S5 depict the interaction energies of the chalcogen bonding adducts, the MEP values at the Se/S and O/S in the monomers, and the electron density values at the bond CPs of the examples extracted from the PDB searches. The inspection of the results included in Tables S4 and S5 reveals that (i) the interaction energies in neutral systems range from −0.8 to −7.8 kcal/mol, where most of the dimers present additional contacts (see Figures S5 and S6); (ii) the MEP values at the selenium's σ-holes are modest ranging from 5.3 kcal/mol in **4IRX** to 13.8 kcal/mol in **3CHV**. This is reasonable taking into consideration that the Se is not bonded to electron-withdrawing groups; (iii) the MEP values at the S-atoms acting as electron acceptors (σ-hole donors) in the S···O contacts are in most cases negative at the extension of the C–S bonds, indicating the absence of positive σ-hole in most of the hits. Therefore, the favorable interaction energies are likely due to the additional contacts and are not due to the ChB itself, and (iv) the MEP values at the O-atoms are more negative than those at the S-atoms (acting as electron donors), apart from 3CHV, because the electron donor is cationic.

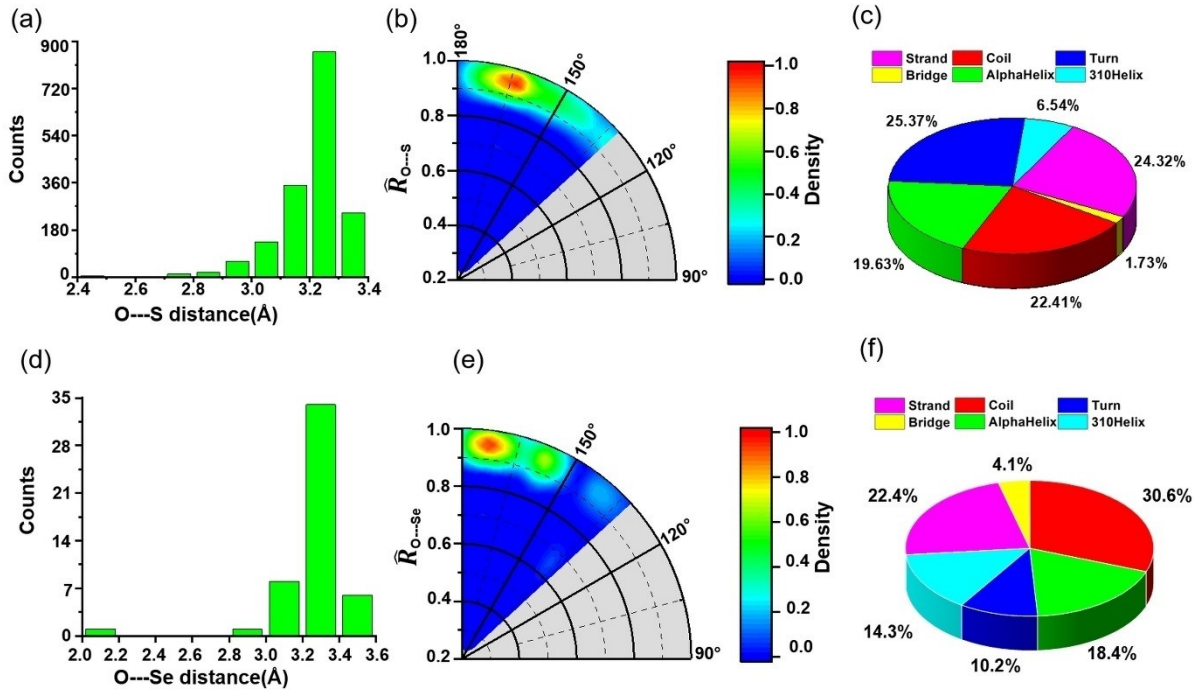


Figure 9. (a) Distribution of distance between S and O, (b) radial density distribution plot where the distance between O and S is normalized with respect to the sum of their van der Waals radii ($\widehat{R}_{S\cdots O}$), the density bar normalized to the maximum count, (c) the distribution of secondary structures of O atom in O···S interaction, (d) distribution of distances between Se and O, (e) radial density distribution plot where the distance between O and Se is normalized with respect to their sum of the van der Waals radii ($\widehat{R}_{Se\cdots O}$), the density bar normalized to the maximum count (f) the distribution of secondary structures of O atom in Se···O interaction.

We have also selected some representative hits of both PDB searches based on those structures where the model dimers presented the lowest ancillary interactions. Consequently, in those hits the ChB interaction is dominant. Figure 10 shows the MEP surfaces of two selected Se-bond donor molecules that participate in Se···S ChBs, and Figure 10 shows the QTAIM/NCI plot index analyses of the dimers. The model of the PDB code 3S9X presents selenomethionine in the structure, where the Se-atom exhibits a modest σ-hole (+6.3 kcal/mol, see Figure 9a). In the X-ray structure, this Se atom is in short contact with the S-atom of a nearby cysteine amino acid, forming a ChB that is characterized by the corresponding bond CPs and green NCI plot isosurfaces (see Figure 11a). The dimerization energy is very small (−0.8 kcal/mol) and is likely due to the poor directionality of this interaction since the lone pair at the electron donor S-atom does not point toward the selenium's σ-hole. In the case of **3DSB**, the σ-hole that participates in the interaction is the opposite of the methyl group, which shows a slightly weaker σ-hole (+4.7 kcal/mol, see Figure 10b). This amino acid also forms a ChB with a nearby cysteine and the computed interaction energy −3.8 kcal/mol. It is greater (in absolute value) than that of the model of **3S9X** because the lone pair at the S-atoms points to the σ-hole with better directionality. This explanation agrees well with the NBO analysis results that show a larger orbital contribution for **3DSB** (0.95 kcal/mol) compared to **3S9X** (0.53 kcal/mol).

Figure 12a shows the MEP surface of the unique Ch-bond donors that presents a positive MEP value at the S-atom (see Table S5), which is a cysteine residue of PDB code **1GT9**. The S-atom exhibits a σ-hole that is very small but positive (+1.7 kcal/mol) partially due to the influence of the adjacent S–H bond. This residue interacts with an O-atom of the protein backbone in the X-ray structure of **1GT9**. The interaction is characterized by the corresponding bond CP and green NCI plot index isosurface connecting the S to the O-atom (see Figure 12b). The interaction energy is small −3.2 kcal/mol, which is in accordance with the small MEP value at the σ-hole. The NBO analysis reveals an orbital contribution from the LP at the O-atom to the antibonding σ*(C–S) that is 1.01 kcal/mol, further confirming the existence of the ChB interaction in the model of **1GT9**.

Conclusions

The results obtained from a systematic analysis of the molecules in the CSD and PDB bring out the relevance of chalcogen-

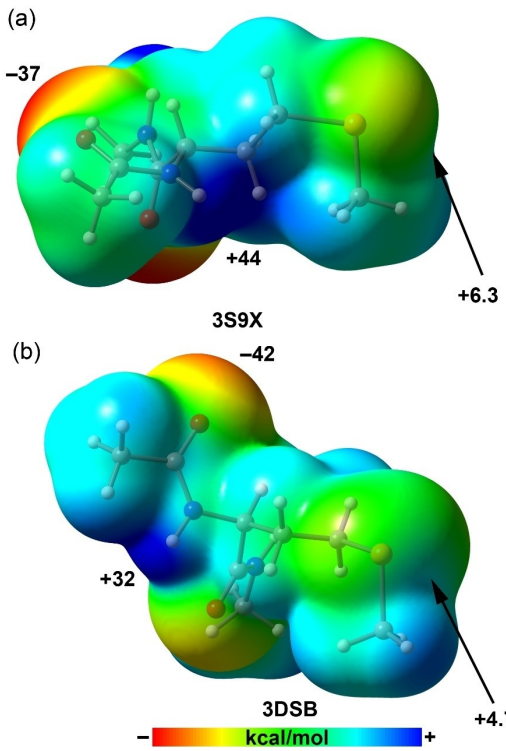


Figure 10. MEP surfaces of the models of the PDB codes **3S9X** (a) and **3DSB** (b). The MEP values at the σ-holes are indicated in kcal/mol.

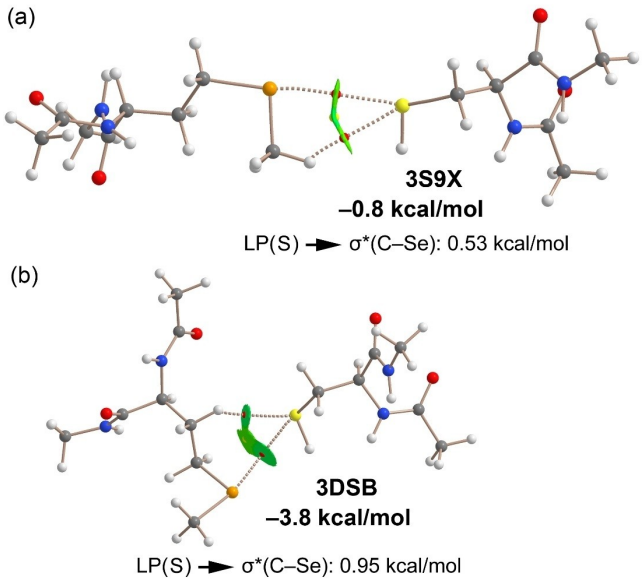


Figure 11. Combined QTAIM (bond and red critical points represented as red and ring spheres, respectively) and NCIPlot (ρ cut-off 0.04 a.u., $|RGD| = 0.5$, color range $-0.04 < (\text{sign}\lambda_2)\,\rho < 0.04$ a.u. of the models of the PDB codes **3S9X** (a) and **3DSB** (b). Only intermolecular interactions are represented for clarity. The orbital stabilization energies obtained by using the second-order perturbation analysis are also indicated along with the orbital involved. LP stands for lone pair and σ* for sigma antibonding orbital.

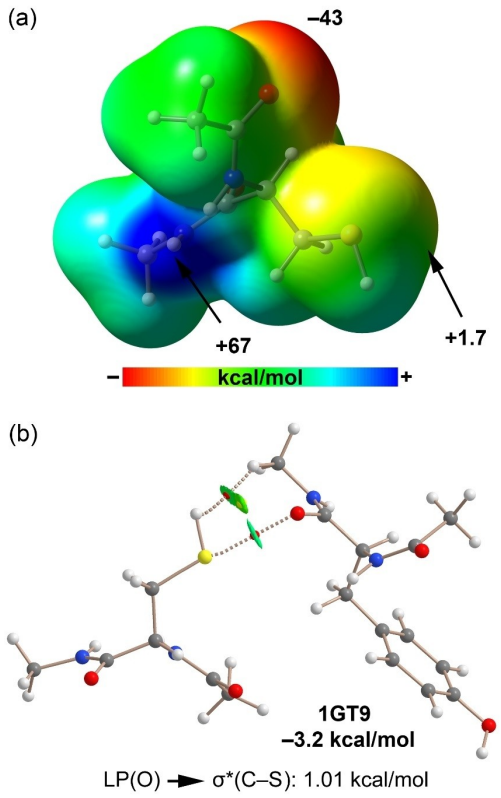


Figure 12. (a) MEP surfaces of the model of PDB code **1GT9**. The MEP values at the σ-hole are indicated in kcal/mol. (b) Combined QTAIM (bond and red critical points represented as red and ring spheres, respectively) and NCIPlot (ρ cut-off 0.04 a.u., $|RGD| = 0.5$, color range $-0.04 < (\text{sign}\lambda_2)\,\rho < 0.04$ a.u. of the model of PDB code **1GT9**. Only intermolecular interactions are represented for clarity. The orbital stabilization energies obtained by using the second-order perturbation analysis are also indicated along with the orbital involved. LP stands for lone pair and σ* for sigma antibonding orbital.

centered Se···O, Se···S, and S···O interactions in the solid-state. The stabilization energies for the dimers are in the range of −2.3 to −7.8 kcal/mol, −0.8 to −3.8 kcal/mol, and −0.7 to −14.1 kcal/mol for Se···O, Se···S and S···O chalcogen bonds in PDB. The corresponding magnitudes, i.e. the stabilization energies for molecules in CSD are −3.7 to −78.0 kcal/mol, −3.8 to −13.2 kcal/mol and −3.6 to −14.9 kcal/mol for the above-mentioned interactions in CSD. The second-order interaction energies are in the range of 0.6 to 3.4 kcal/mol for lone pair donation from O/S to σ* antibonding orbital of C–Se bond, and 2.10 to 2.24 kcal/mol for lone pair donation from O to the σ* antibonding orbital of the C–S bond in case of small molecules from CSD. The corresponding values are in the range of 0.53 to 0.95 kcal/mol for lone pair donation from S to σ* antibonding orbital of C–Se bond and 1.01 kcal/mol for lone pair donation from O to σ* antibonding orbital of the C–S bond in case of molecules in PDB. The MEP analysis clearly depicts the presence of the σ-holes on the chalcogen atoms. The weak attractive nature of these chalcogen bonds was established from the NCI plot visualization index. Such studies establish the significance of these interactions unequivocally and are relevant in the

design of materials related to applications in energy and medicine.

Experimental Section

Computational methods

The energies of all assemblies included in this study were computed at the PBE0-D3/def2-TZVP level of theory by using the program TURBOMOLE v7.0.[25] For the calculations we have used the Weigend def2-TZVP[26,27] basis set and the PBE0[28] DFT functional. The terms interaction energy and binding energy have been used indistinctly in this manuscript and calculated as the energy difference between the complex and the sum of the energies of the monomers. The molecular electrostatic potential surfaces have been plotted using the 0.001 a.u. isosurface using Gaussian-16 software at the PBE1PBE-D3/def2-TZVP level of theory. The PBE1PBE is the notation used by Gaussian-16 for the hybrid functional PBE0. The MP4SDQT/def2-TZVP ab initio calculations were performed using Gaussian-16.[29]

The QTAIM method has been used to obtain the critical points and bond paths via analysis of the topology of the electron density,[30,31] using the PBE1PBE-D3/def2-TZVP wavefuntion and the AIMALL program.[32] The NCIPLOT is a computational index for visualization and identification of NCIs in real space. The noncovalent contacts are identified with the peaks that emerge in the reduced density gradient (RDG) at low densities.[33] These are plotted by mapping an isosurface of s ($s = |\nabla\rho|/\rho^{4/3}$) for a low value of RDG. When a supramolecular dimer is formed, the RDG changes at the CPs in between the monomers due to the annihilation of the density gradient at these CPs. Therefore, this index visualizes the extent to which NCIs stabilize a supramolecular assembly qualitatively and reveals which molecular regions interact. The color code is red-yellow-green-blue where the red and yellow colors are used for strong and weak repulsive (ρ_{cut}^{+}), respectively and the blue and green colors for strong and weak attractive (ρ_{cut}^{-}) forces, respectively. The Natural Bond Orbital analysis of orbital donor-acceptor interactions has been performed at the same level using the NBO 3.1 program.[22]

Cambridge Structural Database (CSD) study (small molecules)

The Cambridge Structure Database (CSD) version 5.42 with updates until May 2021 was inspected using ConQuest version 2021.1.0 (build 319319). All queries were limited to crystal structures that contained 3D coordinates and had no errors. Only those .cif files with Ch···Ch (Ch = Se, S and O) distances $\leq$ than the sum of van der Waals radii + 0.2 Å and C–Ch···Ch angle $\geq 140^{\circ}$ were selected and further inspected manually and statistically. To avoid H-bonded hits, no H-atoms were allowed to be bonded to O and S atoms. Finally, "only organics" option was chosen to eliminate metal bridged hits and related structures.

Acknowledgements

D.C. thanks IISER Bhopal for research facilities and infrastructure. H.S.B. and A.K.S. acknowledge financial support from the Department of Atomic Energy, Department of Science and Technology (Project File No: CRG/20l8/000892), Govt. of India. A.F. and B.G. thank MCIU/AEI of Spain (project PID2020-115637GB-I00, FEDER, UE funds) for financial support. A.F. thanks J. Arbona and R. Frontera (CTI, UIB) for computational facilities.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: chalcogen bonding · CSD analysis · density functional calculations · PDB analysis · sigma-hole interactions

[1] a) M. H. Kolář, P. Hobza, *Chem. Rev.* **2016**, *116*, 5155–5187; b) S. Scheiner, M. Michalczyk, W. Zierkiewicz, *Coord. Chem. Rev.* **2020**, *405*, 213136; c) A. Frontera, A. Bauzá, *Crystals* **2021**, *11*, 1205; d) A. Frontera, A. Bauzá, *Org. Biomol. Chem.* **2021**, *19*, 7816–7821; e) A. Frontera, A. Bauzá, *Org. Biomol. Chem.* **2021**, *19*, 6858–6864.
[2] P. Politzer, J. S. Murray, T. Clark, *Phys. Chem. Chem. Phys.* **2013**, *15*, 11178–11189.
[3] L. C. Gilday, S. W. Robinson, T. A. Barendt, M. J. Langton, B. R. Mullaney, P. D. Beer, *Chem. Rev.* **2015**, *115*, 7118–7195.
[4] a) P. Scilabra, G. Terraneo, G. Resnati, *Acc. Chem. Res.* **2019**, *52*, 1313–1324; b) A. Chand, D. K. Sahoo, A. Rana, S. Jena, H. S. Biswal, *Acc. Chem. Res.* **2020**, *53*, 1580–1592; c) R. M. Gomila, A. Frontera, *J. Organomet. Chem.* **2021**, *945–955*, 122092; d) M. N. Piña, A. Frontera, A. Bauzá, *ACS Chem. Biol.* **2021**, *16*, 1701–1708; e) S. Thakur, R. M. Gomila, A. Frontera, S. Chattopadhyay, *CrystEngComm* **2021**, *23*, 5087–5096; f) J. A. Fernandez-Riveras, A. Frontera, A. Bauzá, *Phys. Chem. Chem. Phys.* **2021**, *23*, 17656–17662.
[5] a) V. Rao Mundlapati, D. K. Sahoo, S. Bhaumik, S. Jena, A. Chandrakar, H. S. Biswal, *Angew. Chem. Int. Ed.* **2018**, *57*, 16496–16500; *Angew. Chem.* **2018**, *130*, 16734–16738; b) S. Scheiner, *Phys. Chem. Chem. Phys.* **2021**, *23*, 5702–5717; c) M. N. Piña, A. Frontera, T. J. Mooibroek, A. Bauzá, *ChemPhysChem* **2021**, *22*, https://doi.org/10.1002/cphc.202100613.
[6] A. M. Maharramov, K. T. Mahmudov, M. N. Kopylovich, M. F. C. Guedes da Silva, A. J. L. Pombeiro, *Noncovalent Interactions in the Synthesis and Design of New Compounds*, Wiley, Hoboken, **2016**.
[7] C. Huang, H. Yu, *ACS Appl. Mater. Interfaces* **2020**, *12*, 19643–19654.
[8] K. Satheeshkumar, G. Mugesh, *Chem. Eur. J.* **2011**, *17*, 4849–57.
[9] V. N. Gladyshev, D. L. Hatfield, *J. Biomed. Sci.* **1999**, *6*, 151–160.
[10] a) H. Komatsu, M. Iwaoka, S. Tomoda, *Chem. Commun.* **1999**, 205–206; b) M. Iwaoka, H. Komatsu, T. Katsuda, S. Tomoda, *J. Am. Chem. Soc.* **2004**, *126*, 5309–5317.
[11] D. J. Pascoe, K. B. Ling, S. L. Cockroft, *J. Am. Chem. Soc.* **2017**, *139*, 15160–15167.
[12] S. Bhandary, A. Sirohiwal, R. Kadu, S. Kumar, D. Chopra, *Cryst. Growth Des.* **2018**, *18*, 3734–3739.
[13] a) R. E. Rosenfield Junior, R. Parthasarathy, J. D. Dunitz, *J. Am. Chem. Soc.* **1977**, *99*, 4860–4862; b) T. N. Guru Row, R. Parthasarathy, *J. Am. Chem. Soc.* **1981**, *103*, 477–479.
[14] F. T. Burling, B. M. Goldstein, *J. Am. Chem. Soc.* **1992**, *114*, 2313–2320.
[15] M. Iwaoka, S. Takemoto, S. Tomoda, *J. Am. Chem. Soc.* **2002**, *124*, 10613–10620.
[16] L. de Azevedo Santos, S. C. C. van der Lubbe, T. A. Hamlin, T. C. Ramalho, F. Mattias Bickelhaupt, *ChemistryOpen* **2021**, *10*, 391–401.
[17] F. De Vleeschouwer, M. Denayer, B. Pinter, P. Geerlings, F. De Proft, *J. Comput. Chem.* **2018**, *39*, 557–572.
[18] C. R. Groom, I. J. Bruno, M. P. Lightfoot, S. C. Ward, *Acta Crystallogr. Sect. B* **2016**, *72*, 171–179.
[19] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res.* **2000**, *28*, 235–42.
[20] R. F. W. Bader, Atoms in Molecules, A Quantum Theory, Clarendon, Oxford, **1990**.
[21] E. R. Johnson, S. Keinan, P. Mori-Sanchez, J. Contreras-Garcia, A. J. Cohen, W. Yang, *J. Am. Chem. Soc.* **2010**, *132*, 6498.
[22] F. Weinhold, C. R. Landis, *Valency and Bonding: A Natural Bond Orbital Donor-Acceptor Perspective*, Cambridge University Press, Cambridge, **2005**.
[23] A. Bondi, *J. Phys. Chem.* **1964**, *68*, 441–451.
[24] M. Heinig, D. Frishman, *Nucleic Acids Res.* **2004**, *32*, W500–W502.

[25] R. Ahlrichs, M. Bär, M. Hacer, H. Horn, C. Kömel, *Chem. Phys. Lett.* **1989**, *162*, 165–169.
[26] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
[27] F. Weigend, *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.
[28] C. Adamo, V. Barone, *J. Chem. Phys.* **1999**, *110*, 6158–6169.
[29] Gaussian 16, Revision C.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, Gaussian, Inc., Wallingford CT, **2016**.
[30] R. F. W. Bader, M. T. Carroll, J. R. Cheeseman, C. Chang, *J. Am. Chem. Soc.* **1987**, *109*, 7968–7979.
[31] R. F. W. Bader, *Atoms in Molecules: A Quantum Theory*, Clarendon, Oxford, **1990**.
[32] AIMAll (Version 17.11.14), T. A. Keith, TK Gristmill Software, Overland Park KS, USA, (aim.tkgristmill.com), **2013**.
[33] E. R. Johnson, S. Keinan, P. Mori-Sanchez, J. Contreras-Garcia, A. J. Cohen, W. Yang, *J. Am. Chem. Soc.* **2010**, *132*, 6498.

Manuscript received: September 20, 2021
Revised manuscript received: October 23, 2021
Accepted manuscript online: October 25, 2021
Version of record online: November 5, 2021