Supplementary Information: Acquiring hybrid molecular crystal materials through high-throughput screening

Yu-Han Yang[1†], Yao-Long Kang[1†], Zi-Ye Li[1], Jia-Bin Liu[1], Fu-Sheng Zhang[1], Chen-Dong Jin[1], Rui-Ning Wang[1], Xing-Qiang Shi[1] and Jiang-Long Wang[1], Peng-Lai Gong[1,*]

[1]*Key Laboratory of Optic-Electronic Information and Materials of Hebei Province, Hebei Research Center of the Basic Discipline for Computational Physics, College of Physics Science and Technology, Hebei University, Baoding 071002, P. R. China*

Contents

Table S1. Number of molecular crystals with different space groups and configurations found through RG^2 search. Sum-X represents the total number of crystals found for each space group, while Sum-Y represents the total number of crystals found for each configuration. The percentages following the numerical values in SUM-X and SUM-Y represent the proportion of each corresponding space group or configuration, respectively, out of all crystals found. The cell at the intersection of the Sum-X and Sum-Y rows/columns, in the bottom right corner of the table, shows the total number of crystals found by RG^2 search (346).

Space Group	r6x	r68	r68x	r6	r8x	Sum-X
195	23	4	21	0	0	48(13.87%)
196	0	4	0	0	0	4(1.16%)
198	0	5	0	0	0	5(1.45%)
200	60	0	45	0	0	105(30.34%)
208	0	15	5	0	0	20(5.78%)
213	0	3	0	0	0	3(0.87%)
215	54	2	29	0	0	85(24.57%)
216	0	0	0	5	0	5(1.45%)
218	4	0	0	0	0	4(1.16%)
221	2	0	13	0	52	67(19.36%)
Sum-Y	143 (41.33%)	33 (9.54%)	113 (32.66%)	5 (1.45%)	52 (15.03%)	346

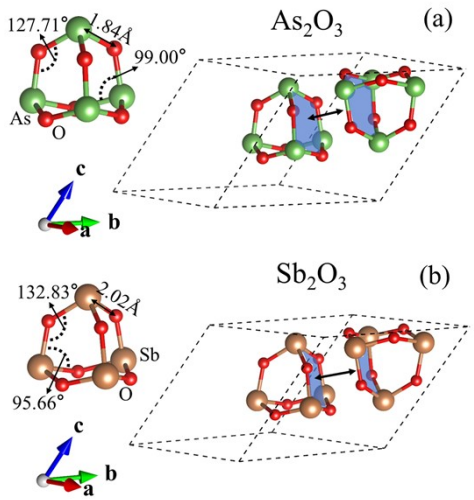

Figure S1. The primitive cell diagrams of As_2O_3 (a) and Sb_2O_3 (b). The blue-filled surfaces and black arrows indicate the nearest-neighbor contacts between adjacent molecular cages.

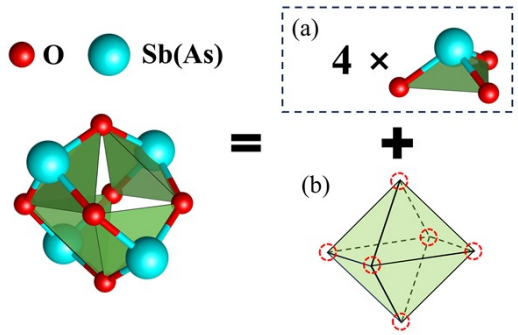

Figure S2. Decomposition diagram of molecular cage As_4O_6 or Sb_4O_6. Each molecular cage can be viewed as being composed of four triangular pyramids, with Sb (or As) atoms at the apex and O atoms forming the base (a). Connecting the positions of the O atoms (indicated by the red dashed circle lines) form a regular octahedron (b), and the center of which corresponds to the center of the molecular cage.

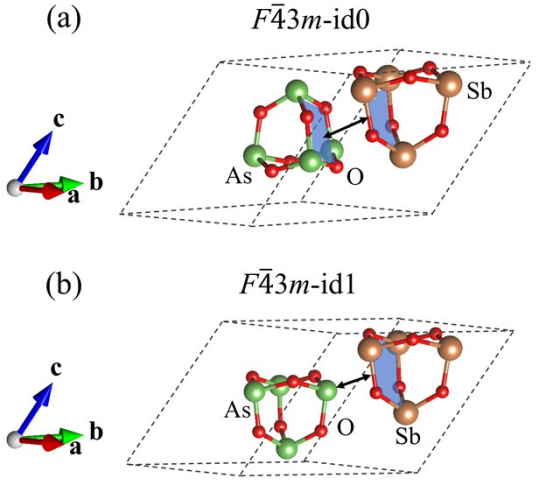

Figure S3. The primitive cell diagrams of $F\bar{4}3m$-id0 (a) and $F\bar{4}3m$-id1 (b). The blue-filled surfaces and black arrows indicate the nearest-neighbor contacts between adjacent molecular cages.

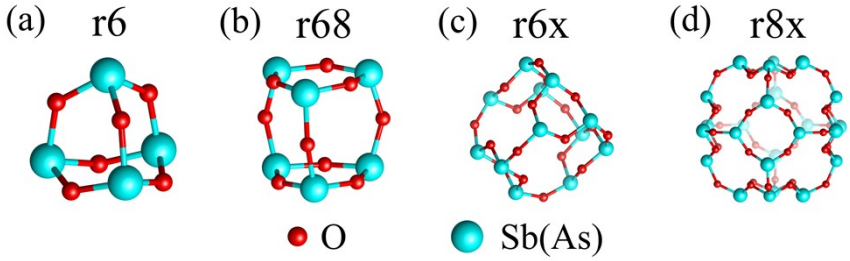

Figure S4. Example of different molecular cage structures.

Table S2. Descriptions of the molecular cages in various conformations of molecular crystals, as searched by RG[2].

Conformation	Crystal includes cages	Cages example
r6	Containing six-membered rings only	Figure S4a
r68	Containing six- and eight-membered rings	Figure S4b
r6x	Containing six-membered rings and possibly eight- or twelve-membered rings	Figure S4c. Other molecular cage configurations may be present in the unit cell, e.g., Figure S4a or Figure S4b.
r8x	Containing eight-membered rings and possibly six- or twelve-membered rings	Figure S4d. Other molecular cage configurations may be present in the unit cell, e.g., Figure S4a or Figure S4b.
r68x	Containing six- and eight-membered rings and possibly twelve-membered rings	The unit cell contains multiple molecular cages, e.g., Figure S4b and S4c.

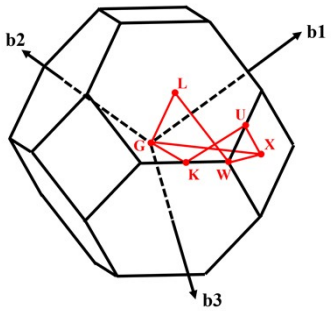

Figure S5. The first Brillouin zone of $F\bar{4}3m$-id0, $F\bar{4}3m$-id1, and $F\bar{4}3m$-id2.

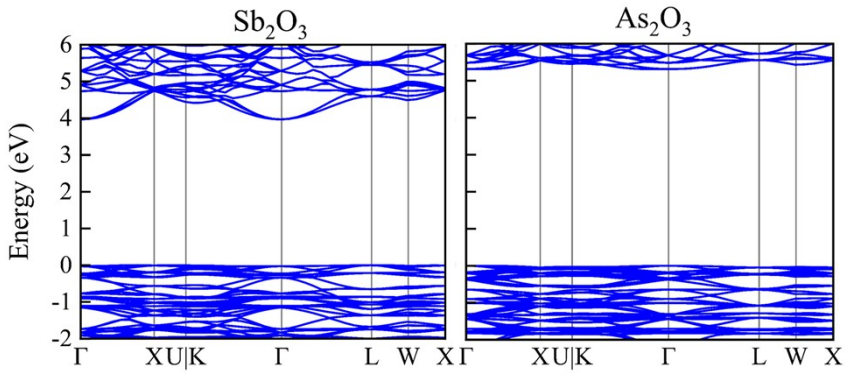

Figure S6. The HSE06 band structures of bulk Sb_2O_3 and As_2O_3, with the band edges aligned relative to the valence band maximum (VBM).

Table S3. POSCAR of stable hybrid molecular crystal $F\bar{4}3m$-id0's primitive cell.

$F\bar{4}3m$-id0 primitive cell		
1.000000		
0	5.5144952748721501	5.5144952748721501
5.5144952748721501	0	5.5144952748721501
5.5144952748721501	5.5144952748721501	0
Sb	O	As
4	12	4
Direct		
0.3670133671905453	0.3670133671905453	0.3670133671905450
0.3670133671905453	0.3670133671905453	0.8989598984283642
0.3670133671905453	0.8989598984283642	0.3670133671905453
0.8989598984283642	0.3670133671905452	0.3670133671905451
0.1785114303508379	0.1785114303508379	0.8214885696491621
0.8214885696491621	0.8214885696491621	0.1785114303508379
0.8214885696491621	0.1785114303508379	0.1785114303508379
0.1785114303508379	0.8214885696491621	0.8214885696491621
0.1785114303508379	0.8214885696491621	0.1785114303508379
0.8214885696491621	0.1785114303508379	0.8214885696491621
0.0584172611418377	0.4415827388581623	0.0584172611418377
0.4415827388581623	0.0584172611418377	0.4415827388581623
0.0584172611418377	0.0584172611418377	0.4415827388581623
0.4415827388581623	0.4415827388581623	0.0584172611418377
0.4415827388581623	0.0584172611418377	0.0584172611418377
0.0584172611418377	0.4415827388581623	0.4415827388581623
0.8933415994694556	0.3199752015916333	0.8933415994694556
0.3199752015916333	0.8933415994694556	0.8933415994694556
0.8933415994694556	0.8933415994694556	0.8933415994694556
0.8933415994694556	0.8933415994694556	0.3199752015916333

Note S1: The volume calculation of molecular cages

As illustrated in Figure S2, the volume of each fully optimized molecular cage can be decomposed into two components: (1) the volume of a regular octahedron (V_8) with

O atoms at its vertices (Figure S2a) and (2) the volume of four triangular pyramids (V_3) (Figure S2b). Therefore, the total volume (V) of the molecular cage is calculated using the following formula:

$$V = V_8 + 4 \times V_3 \tag{S1}$$

$$V_8 = \frac{a^3}{6\sqrt{2}} \tag{S2}$$

$$V_3 = \frac{1}{3} \times S \times h \tag{S3}$$

In equation S2, a denotes the distance between two nearest-neighbor O atoms in the Sb_4O_6 (As_4O_6) molecular cages of various molecular crystals (see Figure S2 for details; the nearest-neighbor O atom distance a, for Sb_4O_6 (As_4O_6) are summarized as a_1(a_2) in Table S4). In equation S3, S represents the base area of the triangular pyramid and h its height. The formulas for calculating S and h are as follows:

$$S = \frac{\sqrt{3}}{4}a^2 \tag{S4}$$

$$h = \sqrt{b^2 - (\frac{1}{3}a^2)} \tag{S5}$$

where b represents the Sb(As)-O bond length (see Figure S2 for details; where the Sb(As)-O bond lengths (b) for different molecular cages are summarized as b_1(b_2) in Table S4). Combining equations S1-S5, we obtain the final formula for calculating the volume of the molecular cage:

$$V = \frac{a^3}{6\sqrt{2}} + \sqrt{\frac{b^2 - \left(\frac{1}{3}a^2\right)}{3}}\, a^2 \tag{S6}$$

$$= \frac{a^2}{3}\left(\frac{\sqrt{2}}{4} + \sqrt{3b^2 - a^2}\right)$$

By substituting the values of a and b for different molecular cages into the equation S6, we obtained the corresponding cage volumes. The volumes of the Sb_4O_6 (As_4O_6) cages are summarized as V_1 (V_2), in Table S5. Considering that the primitive cell of the hybrid molecular crystal contains one Sb_4O_6 cage and one As_4O_6 cage, for ease of comparison, we determine the total volume by summing the volumes of the Sb_4O_6 cage (V_1) and the As_4O_6 cage (V_2) in different molecular crystals and we designate this sum as V_p:

$$V_p = V_1 + V_2 \tag{S7}$$

Table S4. The distance between the neighboring O atoms (a) and the Sb (As)-O bond lengths (b) in the molecular cage Sb_4O_6 (As_4O_6). To better distinguish between the *a* and *b* parameters in the Sb_4O_6 and As_4O_6 molecular cages, we define the nearest-neighbor O atom distance and the Sb-O bond length in Sb_4O_6 as a_1 and b_1, respectively, and the corresponding parameters in As_4O_6 as a_2 and b_2.

Crystals	Sb_4O_6 (Å)		As_4O_6 (Å)	
	O···O (a_1)	Sb-O (b_1)	O···O (a_2)	As-O (b_2)
$F\bar{4}3m$-id0	2.988	2.001	2.784	1.843
$F\bar{4}3m$-id1	2.992	1.990	2.813	1.830
$F\bar{4}3m$-id2	2.987	1.981	2.795	1.822

Table S5. The volumes of individual molecular cages (V) within various molecular crystals, as well as the overall volume of the primitive cell (V_p). For the hybrid molecular crystals, V_p represents the sum of the volumes of the Sb_4O_6 and As_4O_6 cages within their respective primitive cells.

Crystals	V (Å^3)		V_p (Å^3)
	Sb_4O_6 (V_1)	As_4O_6 (V_2)	
$F\bar{4}3m$-id0	6.278	4.948	11.226

$F\bar{4}3m$-id1	6.161	4.785	10.946
$F\bar{4}3m$-id2	6.073	4.736	10.809

Note S2: The distances between neighboring molecular cages

For the analysis of distances between nearest-neighbor molecular cages, we define the distance as that between the centroids of two adjacent cages. This corresponds to the distance (d) between the centroids of the two central octahedra of these neighboring cages (see Figure S2b). Since each octahedron has six vertices that correspond to the six O atoms of the molecular cage (highlighted by the red dashed circles in Figure S2b), we can determine the centroid coordinates of each cage based on the three-dimensional (3D) coordinates of these O atoms. Finally, a pair of nearest-neighbor cages is randomly selected, and the d between their centroids is calculated. The statistical results are presented in Table S6.

Table S6. The distance (d) of different molecular cages.

Crystals	$F\bar{4}3m$-id0	$F\bar{4}3m$-id1	$F\bar{4}3m$-id2
d (Å)	4.78	5.99	7.15