Supplementary Information for:

Anisotropic Charge Distribution and Anisotropic van der Waals Radius Leading to Intriguing Anisotropic Noncovalent Interactions

Hahn Kim[1][†], Van Dung Doan[2][†], Woo Jong Cho[3], Miriyala Vijay Madhav[2], and Kwang S. Kim[2,3*]

[1]Department of Chemistry, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea

[2]Department of Chemistry, Pohang University of Science and Technology, Pohang 790-784, Korea

[3]Department of Chemistry, School of Natural Science, Ulsan National Institute of Science and Technology (UNIST), Ulsan 689-798, Korea
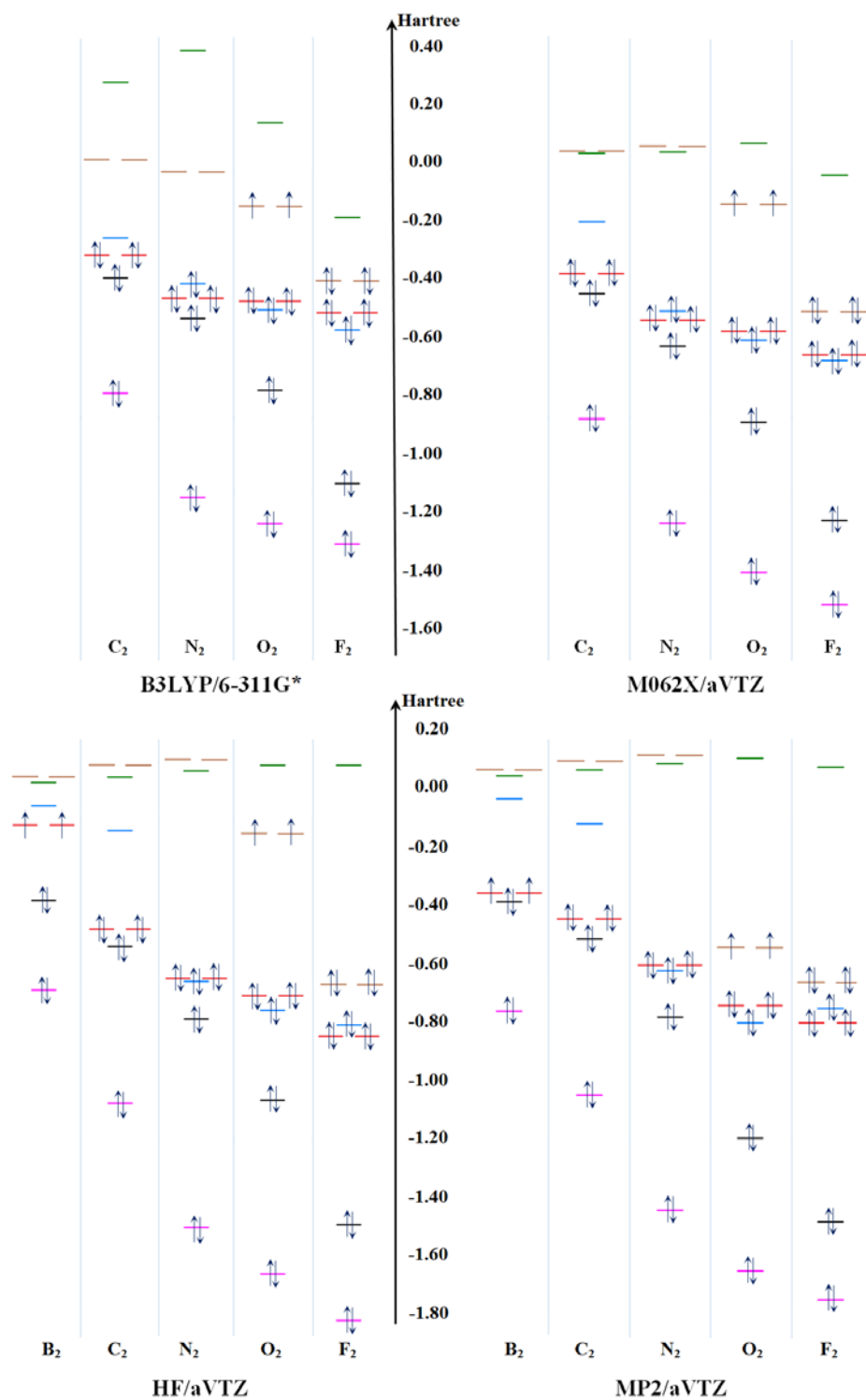
[†]These authors contributed equally.

*Correspondence and requests for materials should be addressed to K.S.K. (email: kimks@unist.ac.kr)
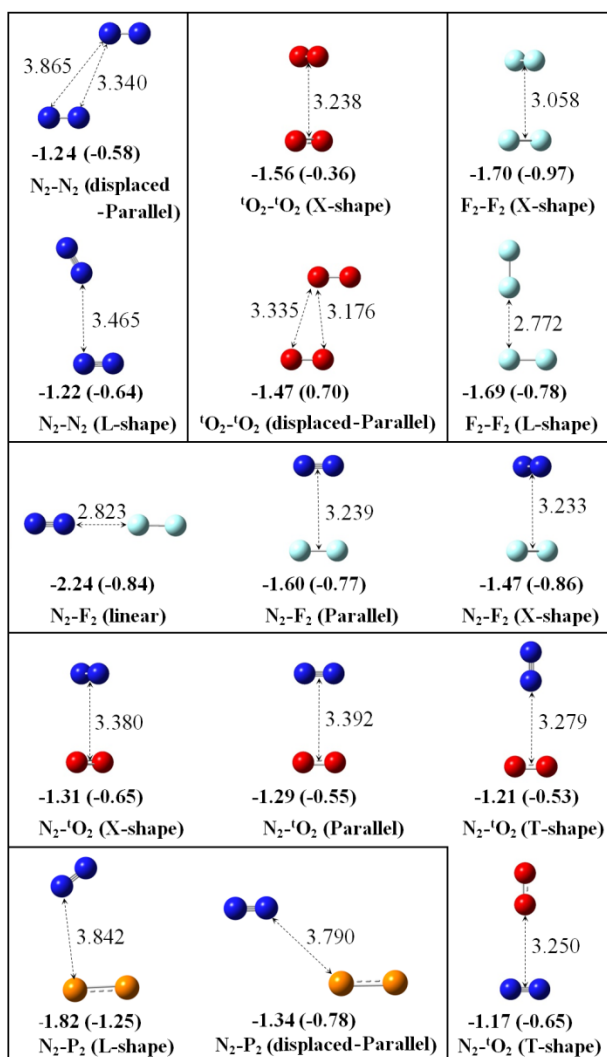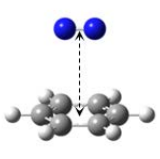
Supplementary Figures
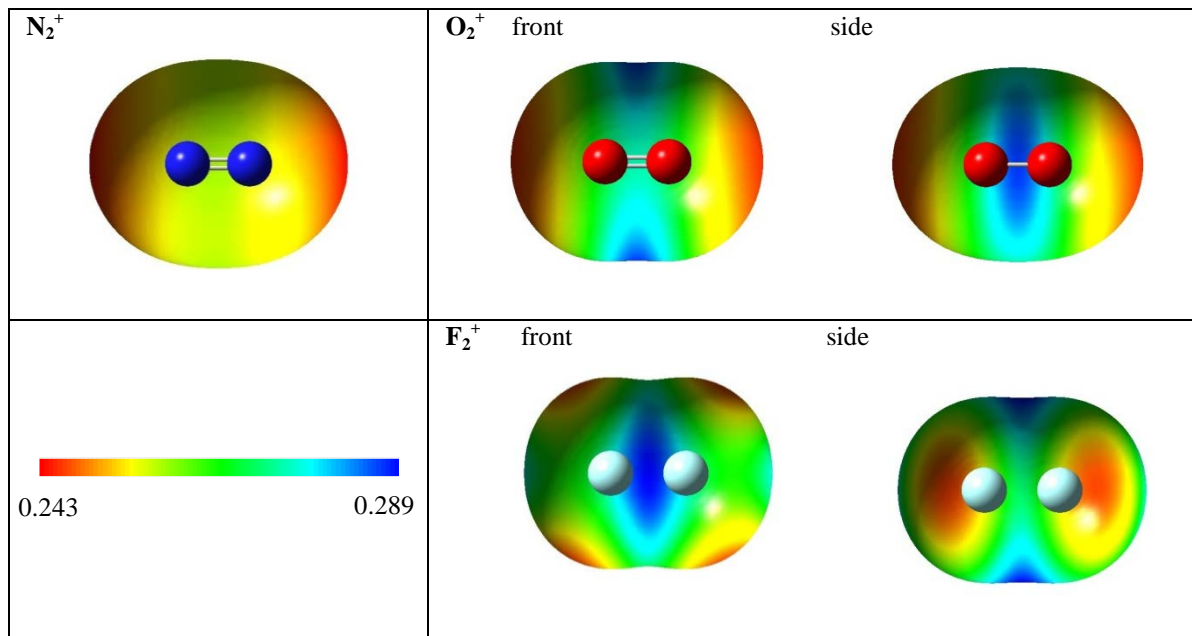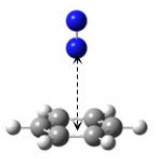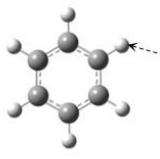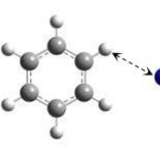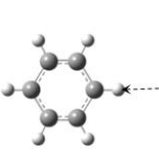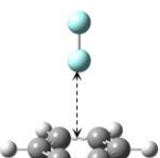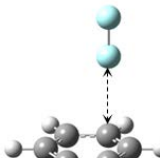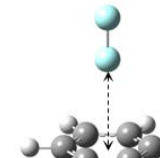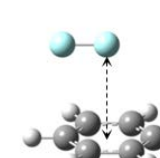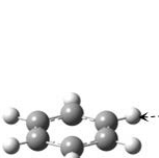
Supplementary Figure S1 | ESP map for N_2^+, O_2^+ and F_2^+ (MP2/aVTZ).

Supplementary Figure S2 | Frontier orbital energy levels for homonuclear diatomic molecules of second period elements. The energy ordering depends sensitively on the methods applied (B3LYP/6-311G*, M062X/aVTZ, HF/aVTZ, MP2/aVTZ). Orbitals are denoted as [$1\sigma_g$:purple, $2\sigma_g^*$:black, $1\pi_u$:red, $3\sigma_u$:blue, $2\pi_u^*$:brown, $4\sigma_u^*$:green]; occupied electrons, as thin dark blue arrows. Only the DFT orbital energy levels match those illustrated in many textbooks of general chemistry and physical chemistry.

Supplementary Figure S3 | Structures and interaction energies of homonuclear homodimers (N_2−N_2, $^tO_2$−$^tO_2$, F_2−F_2, O_2−O_2) and homonuclear heterodimers (N_2−F_2, N_2−$^tO_2$, $^tO_2$−F_2, N_2−P_2). Distances (in Å) are given at the CCSD(T)/aVTZ level, and interaction energies E_e (and E_o in parentheses with the MP2/aVTZ ZPE correction) in kJ/mol are given at the CCSD(T)/CBS level. In the case of short intermolecular distances (<< 3 Å) as in N_2−F_2 (linear) and N_2−F_2, (L-shape), the CCSD(T)/CBS geometry optimization could be required to avoid the sharp decrease of energy towards to the hard wall region. For CCSD(T)/CBS geometry optimization, in the case of N_2−F_2 the nearest N…F distance is 2.894 Å and the interaction energy E_e (E_o) is -2.28 (-0.88) kJ/mol; in the case of F_2−F_2, the nearest F…F distance is 2.828 Å and the interaction energy E_e (E_o) is -1.71 (-0.80) kJ/mol.

P: **P**arallel	**T**: **T**-shape	**Sp**: parallel-**S**ided	**S2**: bi-**S**ided	**S1**: uni-**S**ided
N_2 −6.35 d_{cc}=3.35	N_2 −3.24 d_{cn}=3.22	N_2 −2.93 d_{hn}=2.80	N_2 −2.81 d_{hn}=2.88	N_2 −2.00 d_{hn}=2.73
$^tO_2$ −5.52 d_{cc}=3.27	$^tO_2$ −4.16 d_{co}=3.11	$^tO_2$ −2.99 d_{ho}=2.70	$^tO_2$ −2.65 d_{ho}=2.73	$^tO_2$ −1.45 d_{ho}=2.61
O_2 −11.32 d_{cc}=2.91	O_2 −4.13 d_{co}=3.09	O_2 −3.90 d_{ho}=2.71	O_2 −3.26 d_{ho}=2.71	O_2 −1.69 d_{ho}=2.57
Tb: **T**-shape-on-bond	**Ta**: **T**-shape-on-atom	**T**: **T**-shape	**Pd**: displaced-**P**arallel	**Sc**: cross-**S**ided
F_2 −6.04 d_v=2.89	F_2 −5.85 d_v=2.89	F_2 −5.64 d_{cf}=3.02	F_2 −4.27 d_v=3.22	F_2 −3.56 d_{hf}=2.75
Cl_2 −12.73 d_v=3.10	Cl_2 −12.37 d_v=3.09	Cl_2 −11.25 d_{ccl}=3.28	Cl_2 −8.79 d_v=3.46	Cl_2 −6.48 d_{hcl}=3.09

Supplementary Figure S4 | CCSD(T)/CBS energies for benzene-X_2 (X = N, ^{t}O, O, F, Cl). The geometries were optimized first at the BSSE-corrected MP2/aVTZ level, and then the interaction distance in each complex was optimized at the BSSE uncorrected CCSD(T)/aVTZ level. Interaction energies (E_e) are in kJ/mol and the distances denoted by dotted lines (d) are in Å. In d_{ca}, subscript "c" denotes center; subscript "a" designates an atom (a; N/O/F/Cl/H). d_v is the vertical distance shown in an arrow.

Supplementary Tables

Supplementary Table S1 | Hydridization character (in %) of σ-lone pairs of N_2, O_2, F_2, O_2^{2+}, C_2^{2-} calculated at the QCISD/aug-cc-pVTZ//MP2/aug-cc-pVTZ level.

species	s-character	p-character
N_2	63	37
N_2(1.4Å)	80	20
$^tO_2$	83/60	16/20 (up/down spin)
F_2	95	5
F_2(1.1Å)	84	16
P_2	73	27
$^tS_2$	87/85	13/15 (up/down spin)
Cl_2	93	7
C_2^{2-}	55	45
O_2^{2+}	82	18

Supplementary Table S2 | SAPT-DFT energy decomposition [kJ/mol] with the asymptotically corrected (AC) PBE0 functional and the aug-cc-pVTZ basis set on the MP2/aVTZ optimized geometries for molecular complex of each homonuclear dimer interacting with either a homonuclear dimer or a benzene (Bz) molecule.[a]

Complex	E_{es}	E_{ind}^{*}	E_{disp}^{*}	E_{exch}^{*}	δH_{HF}	E_{tot}	$E_{CCSD(T)/CBS}$
N_2−N_2 (Pd)	-0.74	-0.04	-2.47	2.21	-0.13	-1.16	-1.24
N_2−N_2 (L)	-0.76	-0.03	-2.18	1.90	-0.06	-1.12	-1.22
N_2−N_2 (T)	-0.87	-0.03	-2.34	2.26	-0.07	-1.05	-1.17
F_2−F_2 (X)	-0.57	0.00	-2.87	2.11	-0.03	-1.35	-1.70
F_2−F_2 (L)	-1.13	-0.05	-3.17	3.35	-0.23	-1.24	-1.69
N_2−F_2 (l)	-2.58	-0.09	-3.63	5.28	-0.61	-1.63	-2.24
N_2−F_2 (P)	-1.05	0.00	-3.37	3.19	-0.13	-1.36	-1.60
N_2−F_2 (X)	-0.89	0.00	-3.14	2.88	-0.10	-1.24	-1.47
N_2−P_2 (L)	-1.20	-0.02	-3.50	3.11	-0.33	-1.93	-1.82
N_2−P_2 (D)	-0.82	-0.03	-2.88	2.60	-0.27	-1.41	-1.34
Bz−N_2 (P)	-4.55	-0.25	-11.67	11.42	-1.28	-6.32	-6.35
Bz−N_2 (T)	-0.95	-0.20	-8.41	6.96	-0.79	-3.40	-3.24
Bz−N_2 (Sp)	-1.96	-0.20	-6.99	7.17	-0.52	-2.50	-2.93
Bz−N_2 (S2)	-1.82	-0.12	-5.28	4.97	-0.28	-2.26	-2.81
Bz−N_2 (S1)	-1.90	-0.14	-3.44	3.94	-0.22	-1.74	-2.00
Bz−O_2 (P)	-13.01	0.18	-20.84	43.19	-19.80	-10.30	-11.32
Bz−O_2 (T)	-3.35	-0.14	-10.54	12.21	-1.81	-3.63	-4.13
Bz−O_2 (Sp)	-3.80	-0.17	-6.54	7.41	-0.32	-3.41	-3.90
Bz−O_2 (S2)	-2.75	-0.10	-5.70	6.26	-0.94	-3.24	-3.26
Bz−O_2 (S1)	-1.24	-0.06	-3.54	3.78	-0.26	-1.33	-1.69
Bz−F_2 (Tb)	-5.87	0.11	-9.24	13.62	-3.46	-4.84	-6.04
Bz−F_2 (Ta)	-5.72	0.13	-8.90	13.29	-3.50	-4.69	-5.85
Bz−F_2 (T)	-3.22	-0.11	-8.11	8.10	-0.98	-4.32	-5.64
Bz−F_2 (Pd)	-3.03	0.11	-9.32	9.70	-0.77	-3.30	-4.27
Bz−F_2 (Sc)	-2.45	0.00	-7.06	7.59	-0.30	-2.22	-3.56
Bz−Cl_2 (Tb)	-13.10	-1.76	-18.52	26.69	-5.52	-12.21	-12.73
Bz−Cl_2 (Ta)	-12.90	-1.77	-18.04	26.45	-5.64	-11.90	-12.37
Bz−Cl_2 (T)	-7.89	-0.98	-15.53	15.72	-1.84	-10.53	-11.25

[a] L: L-shape, T: T-shape, X: X-shape, l: linear, Pd: displaced-Parallel, P: Parallel, Sp: parallel-Sided, S2: bi-Sided, S1: uni-Sided, Sc: cross-Sided. The geometries of complexes were optimized at the MP2/aVTZ level. The triplet states were not calculated because the SAPT can handle only the singlet states.

Supplementary Notes

Supplementary Note 1 | Anisotropic charge distribution in charged states of N_2, O_2 and F_2.

N_2^+, O_2^+, and F_2^+ also show anisotropic ESP (Supplementary Fig. 1). For N_2^+, the cylindrical-bond-surface is more positive, and the bond-ends are only slightly positive. For F_2^+, the cylindrical-bond-surface is only slightly positive, and the bond-ends are much more positive. This phenomenon reflects the ESP at their corresponding neutral diatomic species. To analyze the electronic configuration of N_2, O_2 and F_2 upon ionization, the well-known MO energy level diagram in typical textbooks can be a useful guide. However, in this case, the MO energy levels for such physico-chemical interpretation are better presented by Kohn-Sham orbitals, but not by Hartree-Fock orbitals, as has been recently addressed (See Supplementary Fig. 2 for the MO energy levels of B_2, C_2, N_2, $^tO_2$ and F_2, while Be_2 is not bound). Furthermore, when N_2 is doubly ionized as N_2^{++}, the triplet state is more stable than the singlet. The triplet state has an up spin in both σ and π orbitals. Since the electron correlation due to the double electron excitations should be taken into account, this case cannot be explained simply by normal orbital diagram representation in typical textbooks of general chemistry and physical chemistry. Thus, in this case the MO energy level diagrams should not be overstressed.

Supplementary Note 2 | Detailed analysis of interactions in the homo-dimers and hetero-dimers of homonuclear diatomic molecules.

The structures of the homo-dimers and hetero-dimers for homonuclear diatomic molecules are optimized at the CCSD(T)/aVTZ level. The zero-point-energy (ZPE) uncorrected interaction energies (E_e) in kJ/mol are given at the CCSD(T)/CBS level, and the ZPE-corrected CCSD(T)/CBS energies (E_o) are provided with the ZPE obtained from MP2/aVTZ (Supplementary Fig. 3). For the N_2 dimer (N_2-N_2), the displaced-Parallel (Pd) shape (E_e=-1.24 kJ/mol; E_o=-0.58 kJ/mol at a distance of the nearest N atoms d_{NN}= 3.340 Å) and the L-shape (E_e= -1.22 kJ/mol; E_o=-0.64 kJ/mol at the distance between two N atoms d_{NN}= 3.465 Å) are the most stable. In both Pd and L shapes, the electrostatic energy (E_{es}=-0.74 and -0.76 kJ/mol, respectively) plays an important role. The next most stable structure is the T-shape (E_e=-1.17 kJ/mol; E_o=-0.45 kJ/mol at the distance between N and the mid-point of N_2 d_{Nm} = 3.459 Å) for which an electrostatically negative bond-end interacts with the mid-point of the electrostatically positive cylindrical-bond-surface. In the case of the ${}^t O_2$ dimer (${}^t O_2-{}^t O_2$), an X-shaped structure (E_e=-1.56 kJ/mol; E_o=-0.36 kJ/mol at the distance between the two mid-points of both ${}^t O_2$ molecules d_{mm} = 3.238 Å) has the lowest energy. In this structure the van der Waals interaction becomes significant due to the weakened electrostatic potential. A parallel displaced structure with a significant interaction energy (E_e=-1.47 kJ/mol; E_o=0.70 kJ/mol), where the θ=60° direction is important, is not stable due to significant zero point energy (ZPE). The F_2 dimer (F_2-F_2) has two nearly isoenergetic structures (in E_e): an X-shape (E_e=-1.70 kJ/mol; E_o=-0.97 kJ/mol at the distance d_{mm}=3.058 Å) where the van der Waals interaction becomes significant due to the increased number of electrons, and an L-shape (E_e=-1.69 kJ/mol; E_o=-0.78 kJ/mol at the distance between two F atoms d_{FF}= 2.772 Å) for which an electrostatically positive F atom's end interacts with the F in another molecule at an electrostatically negative cylindrical-bond-surface edge. The X-shape is determined to be slightly more stable in terms of E_o because the L-shape has a significant ZPE due to two short-ranged F atoms' interaction between the two F_2 molecules.

In the case of the hetero-molecular interaction between N_2 (which has electrostatically negative bond-ends) and F_2 (which has electrostatically positive bond-ends), a linear structure is the most stable in the potential energy surface (E_e=-2.24 kJ/mol; E_o=-0.84 kJ/mol which becomes -0.88 kJ/mol for the full CCSD(T)/CBS geometry optimization), followed by a parallel structure (E_e=-1.60 kJ/mol; E_o=-0.77 kJ/mol) and an X-shape structure (E_e=-1.47 kJ/mol; E_o=-0.86 kJ/mol). However, when the ZPE is considered, the strong N…F binding between N_2 and F_2 results in high ZPE in the linear structure, and so in terms of E_o, the linear structure (having a large ZPE) and X-shape structure (having a minimal ZPE energy due to van der Waals driven interaction) turn out to be nearly isoenergetic. In the case of the interaction between N_2 and ${}^t O_2$ (both of which have electrostatically negative bond-ends and electrostatically positive cylindrical-bond-surfaces), the X-shape structure (E_e=-1.31 kJ/mol; E_o=-0.65 kJ/mol) is the most stable, followed by a parallel structure (E_e=-1.29 kJ/mol; E_o=-0.55 kJ/mol), a T-shape structure with the N_2 end interacting with the O_2 bond (E_e=-1.21 kJ/mol; E_o=-0.53 kJ/mol), and another T-shape structure with the O atom's end interacting with the N_2 bond (E_e=-1.17 kJ/mol; E_o=-0.65 kJ/mol). Since the latter T-shape structure has a small ZPE, it is as stable as the X-shape structure after ZPE correction. In the case of the interaction between N_2 (which has electrostatically negative bond-ends) and P_2 (which has electrostatically positive bond-ends), they preferentially form an L-shape structure (E_e=-1.82 kJ/mol; E_o=-1.25 kJ/mol), followed by a displaced structure (E_e=-1.34 kJ/mol; E_o=-0.78 kJ/mol).

In the case of N_2-N_2, according to a simple intuition based on the electrostatic interactions due to charge anisotropy, one might think that the T-shape could be slightly more stable than the Pd and L-shapes (Pd/L vs. T in E_{es}: -0.74/-0.76 vs. -0.86 kJ/mol). However, the Pd/L-shape is slightly more stable due to larger interaction energy gain from the sum of E_{disp}^* and E_{exch}^* than the T-shape (Pd/L vs. T in $E_{disp*}+E_{exch*}$: -0.26/-0.28 vs. -0.08 kJ/mol). The F_2-F_2 interaction shows two isoenergetic structures: X-shape and L-shape. However, in terms of E_{es} the L-shape is far more stabilized than the X-shape (L vs. X in E_{es}: -1.13 vs. -0.57 kJ/mol) because of the strong electrostatic interaction between the electrostatically positive atom's end and the electrostatically negative cylindrical-bond-surface in the L-shape. In terms of E_{disp}^* the L-shape is also more stabilized than the X-shape (E_{disp}^* for L and X shapes are -3.17 and -2.87 kJ/mol, respectively) because of the larger van der Waals interaction arising from the short contact length between the two interacting F atoms which occurs between the two F_2 molecules in the L-shape. This can also be understood by considering the van der Waals hard wall radii in Fig. 2 [$r_w(90^o)$ vs. $r_w(central)$], as the X-shape has less van der Waals interaction energy due to the larger hard wall distance r_w arising from the strong repulsion between two mid bond regions having a denser ρ_e. On the other hand, in terms of E_{exch*} the L-shape is far less stabilized than the X-shape (E_{exch*} for L and X shapes are 3.34 vs. 2.11 kJ/mol, respectively) because of the shorter contact length between two interacting F atoms in the L-shape. For this reason, both X-shape and L-shape are isoenergetic with respect to E_e,

For the interaction between N_2 and F_2, one observes that in terms of E_{es} the linear structure ($E_{es} = -2.58$ kJ/mol) is much more stabilized than the parallel ($E_{es} = -0.16$ kJ/mol) and X-shape ($E_{es} = -0.27$ kJ/mol) structures because of the strong electrostatic interaction between the electrostatically negative atom's end of N_2 and the electrostatically positive atom's end of F_2. For the N_2-P_2 interaction, the electrostatic interaction of P_2 behaves like that of F_2 which is opposite to that of N_2, but is much less anisotropic. Thus, both electrostatic and dispersion interactions are important, and as such the partly stacked L-shape is the most stable.

Supplementary Note 3 | Detailed analysis of interactions of the homonuclear diatomic molecules with benzene.

Supplementary Fig. 4 shows the low-lying energy structures of homonuclear diatomic molecules interacting with benzene (Bz). These structures reflect the charge distribution of Bz which has negative electron clouds both on and down the Bz plane and electrostatically positive radial direction of the Bz plane due to the presence of positively charged H atoms. In the case of N_2, the parallel (P) structure is the most stable (E_e=-6.3 kJ/mol), while the T-shape (T) structure (E_e=-3.2 kJ/mol) and the side structures are less stable (SS: doubly sided with two N…H bridges, E_e=-2.9 kJ/mol; S2: bi-sided with two N…H bridges involving the same N atom, E_e=-2.9 kJ/mol; S1: single-sided with one N…H bridge, E_e=-2.7 kJ/mol). In the case of $^tO_2$, the parallel structure is the most stable (E_e=-5.5 kJ/mol), while the T-shape structure (E_e=-4.2 kJ/mol) and the SS structure (E_e=-3.0 kJ/mol) are less stable. In the same way, for the singlet O_2, the parallel structure is also the most stable (E_e=-11.3 kJ/mol) since the electrostatically negative surface of the Bz plane is able to favorably interact with the electrostatically positive cylindrical-bond-surface of O_2. (top figure for O_2 in Fig. 1), but not with the electrostatically negative cylindrical-bond-surface of O_2. (bottom front-view figure for O_2 in Fig. 1). The T-shape (E_e=-4.2 kJ/mol) and SS (E_e=-3.0 kJ/mol) are far less stable.

In contrast to Bz-N_2 and Bz-O_2, for the interaction between Bz and F_2 the T-shape-on-Bond (Tb) structure is the most stable (E_e=-6.0 kJ/mol), as the most electrostatically negative CC aromatic bond of the Bz plane favorably interacts with the electrostatically positive atom's end of F_2. The displaced-Parallel (Pd) (E_e=-4.3 kJ/mol) and cross-sided (Sc) (E_e=-3.6 kJ/mol) structures are less stable. Likewise, Cl_2 behaves like F_2, but with stronger interactions. In this case, the Tb structure is the most stable (E_e=-12.7 kJ/mol), and the Pd (E_e=-8.8 kJ/mol) and Sc (E_e=-6.5 kJ/mol) structures are less stable. Based on this information, we can understand the interactions of these diatomic molecules with Bz and further with graphene.

For Bz-N_2, the most stable parallel structure (P) shows strong electrostatic energy (E_{es}: -4.55 kJ/mol), while the E_{disp*} and E_{exch*} tend to nearly cancel ($E_{disp*}+E_{exch*}$ = -0.25 kJ/mol). On the other hand, the T-shape structure (T) shows weak E_{es} (-0.95 kJ/mol) but large $E_{disp*}+E_{exch*}$ (-1.45 kJ/mol), thus, forming a dispersion driven structure. Structure SS/S2/S1 is electrostatic driven (E_{es}: -1.96/-1.82/-1.90/ kJ/mol), because the $E_{disp*}+E_{exch*}$ (0.18/-0.31/0.50 kJ/mol) is very small or positive. The current SAPT calculation is not available for the triplet state, so the calculations were performed only for the excited singlet state O_2 with benzene. For Bz-O_2 (singlet), the most stable parallel (P) structure shows strong electrostatic energy (E_{es}: -13.01 kJ/mol), while $E_{disp*}+E_{exch*}+E_{HF}$ (2.55 kJ/mol) is slightly positive; thus, it forms an electrostatically driven structure. In the T/S structures the electrostatic energy (E_{es}: -3.35/-3.80 kJ/mol) plays a significant role in the total energy ($E_{tot}^{SAPT-DFT}$: -3.63/-3.41 kJ/mol, $E_{CCSD(T)/CBS}$: -4.75/-4.66 kJ/mol), but the dispersion interaction does not play a significant role because the value of $E_{disp*}+E_{exch*}$ (1.67/0.87 kJ/mol) is positive despite the large magnitude of dispersion interaction (E_{disp*}: -10.54/-6.54 kJ/mol).

For Bz-F_2, the most stable Tb structure exhibits a large electrostatic energy (E_{es}: -5.87 kJ/mol), while the $E_{disp*}+E_{exch*}$ (4.38 kJ/mol) is positive; as such it forms an electrostatically driven structure. The Pd/Sc structures give a significant contribution of E_{es} (-3.03/-2.45 kJ/mol) to $E_{tot}^{SAPT-DFT}$ (-3.30/-2.22 kJ/mol) but an insignificant van der Waals interaction contribution of $E_{disp*}+E_{exch*}$ (0.38/0.53 kJ/mol); for these reasons an electrostatically driven structure forms. A similar behavior is noted in Bz-Cl_2, but

with stronger interaction energy. The most stable Tb structure shows strong electrostatic energy (E_{es}: -13.10 kJ/mol), while the $E_{disp*}+E_{exch*}$ (8.17 kJ/mol) is positive, forming an electrostatically driven structure. The Pd/Sc structure gives a significant contribution of E_{es} (-6.21/-3.83 kJ/mol) to $E_{tot}^{SAPT-DFT}$: -8.22/-5.59 kJ/mol) but an insignificant van der Waals interaction contribution of $E_{disp*}+E_{exch*}$ (-0.44/-0.82 kJ/mol), therefore, an electrostatically driven structure is formed. Therefore, anisotropic charge distribution effect plays an important role in molecular recognition and assembly. When the Bz-F_2 and Bz-Cl_2 are compared, Bz-Cl_2 has a much more electrostatic contribution, while Bz-F_2 has a somewhat smaller electrostatic contribution. In this regard, Bz-Cl_2 is governed by dispersion interaction followed by electrostatic interaction, while Bz-F_2 is governed mostly by dispersion interaction, as can be noted from the simple magnitude of the dispersion and electrostatic contribution in the Supplementary Table 2. This issue was already addressed by Hobza and coworkers (ref. 44 in the main text). Indeed, it is true from the table. However, a deeper analysis indicates that the electrostatic contribution (even though its magnitude is smaller than the dispersion) can play a more significant role in conformational change for molecular complexation (refs. 13,14,42,43 in the main text).