Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supplementary Information

Theoretical search for characteristic atoms in supported gold nanoparticles: A large-scale DFT study

Shengzhou Li,[a,b] Tsuyoshi Miyazaki[b] and Ayako Nakata [*,a,b,c]

[a] Department of Computer Science, University of Tsukuba, Tsukuba, Ibaraki 305-8573, Japan.

[b] Research Center for Materials Nanoarchitectonics (MANA), National Institute for Materials Science (NIMS), Tsukuba, Ibaraki 305-0044, Japan

[c] Precursory Research for Embryonic Science and Technology (PRESTO), Japan Science and Technology Agency (JST), Kawaguchi, Saitama 332-0012, Japan

*NAKATA.Ayako@nims.go.jp

S1. PAO basis set dependence of PCA

We calculated the local density of states (LDOS) by Eq. (5) in the manuscript. The result of Eq. (5) depends on the basis set, similar to Mulliken's population analysis [S1]. Therefore, we investigated the basis set dependence of principal component analysis (PCA). The examined pseudo atomic orbital (PAO) sets are summarised in Table S1. The PCA results for different PAOs are provided in Figure S4, in which the LDOS calculated with the PAO sets were projected with the PCA vectors obtained using double-ζ polarisation (DZP) functions. As shown in Figure S4, we confirm that the distribution of the PCA points does not significantly depend on the basis set in the present study.

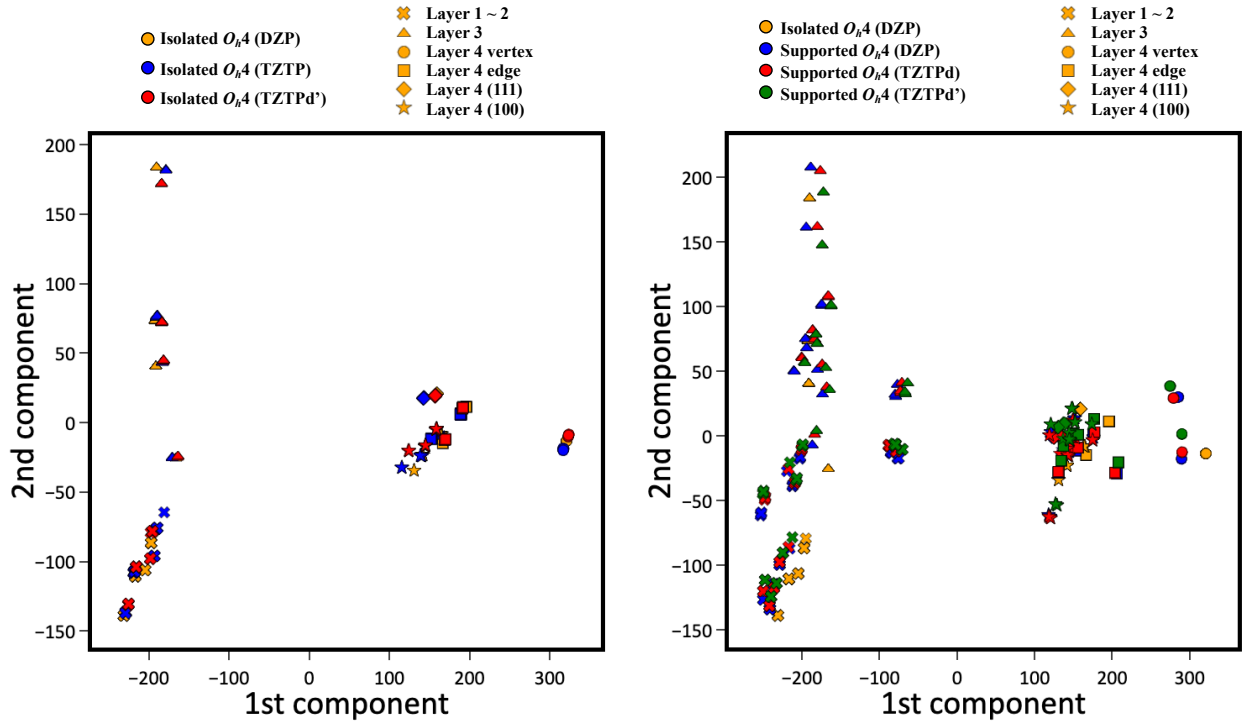

Figure S1. Distribution of the LDOS for atoms in gold nanoparticles consisting of 923 atoms (O_h6), calculated with the PAOs in Table S1 in two-dimensional space by PCA.

Table S1. PAO ranges in DZP functions, triple-ζ with triple-polarisation functions including d functions of Mg (TZTP(d)), and elongated TZTP(d) (TZTP(d)') [Bohr].

PAO			Au		
	5s	5p	5d	6s	polarization p
DZP	2.81	3.36	7.12, 3.74	7.12, 3.74	7.12
TZTP(d)	2.81	3.36	7.12, 5.41, 3.74	7.12, 5.41, 3.74	7.12, 5.41, 3.74
TZTP(d)'	2.81	3.36	9.14, 6.39, 3.74	9.14, 6.39, 3.74	9.14, 6.39, 3.74
			Mg		
	2s	2p	3s	polarization p	polarization d
DZP	2.77	3.16	9.03, 4.69	9.03	
TZTP(d)	2.77	3.16	9.03, 6.86, 4.69	9.03, 6.86, 4.69	9.03, 6.86, 4.69
TZTP(d)'	2.77	3.16	11.06, 7.83, 4.96	11.06, 7.83, 4.96	11.06, 7.83, 4.96
			O		
	2s	2p	polarization d		
DZP	4.91, 2.58	4.91, 2.58	4.91		
TZTP(d)	4.91, 3.74, 2.58	4.91, 3.74, 2.58	4.91, 3.74, 2.58		
TZTP(d)'	6.95, 4.74, 2.58	6.95, 4.74, 2.58	6.95, 4.74, 2.58		

S2. K-means clustering and decision tree analyses

To analyse the LDOS differences, we also performed k-means clustering and used decision tree (DT) methods. For k-means clustering (Figure S3), we changed the number of clusters from five to nine to determine how the LDOS is divided stepwise. The clusters divided in earlier steps have larger differences so that we can identify which atoms have larger differences in the LDOS. The LDOS of the surface atoms at the vertex and (100) face are divided in the first step and never divided into smaller clusters after that, which means that the vertex and (100) face atoms have characteristic electronic structures. Edge and (100) face atoms in the second layer (which is the subsurface of O_h3) are divided at the last step, which means that the difference in the LDOS among them is the smallest.

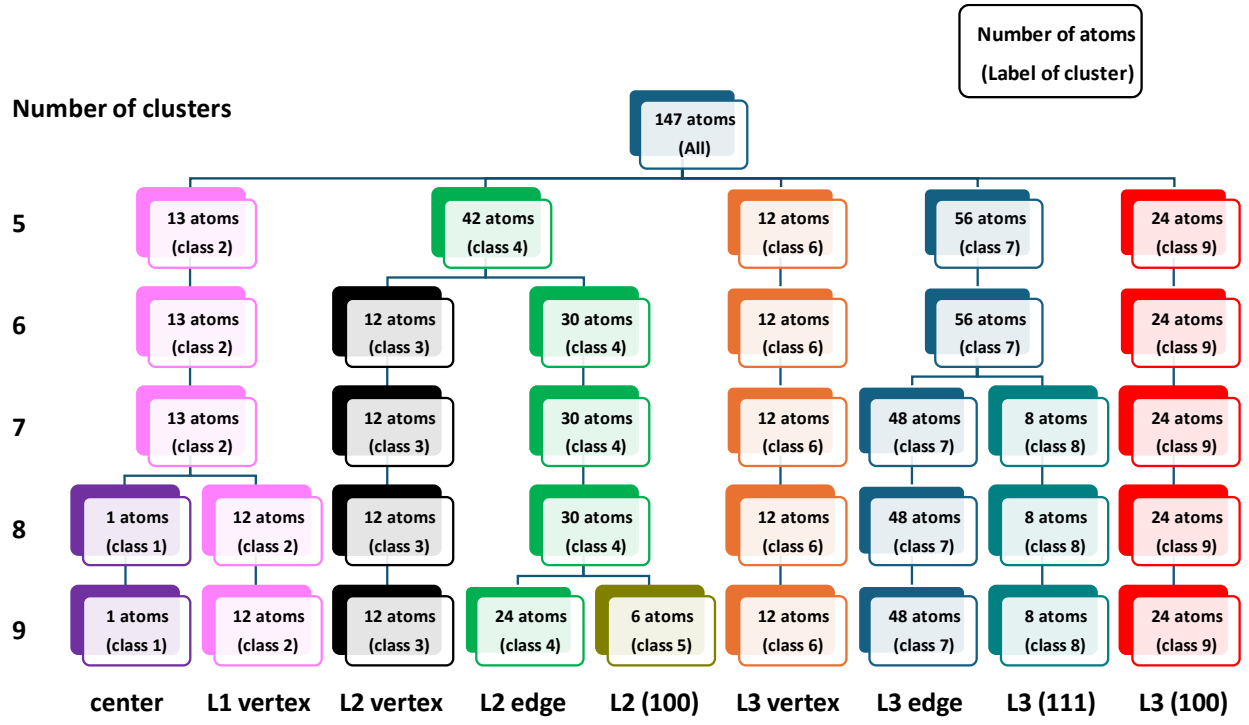

Figure S3. Cluster decomposition of the LDOS for the atoms in the nanoparticle consisting of 147 atoms (O_h3), according to the number of the clusters in k-means clustering. L1, L2 and L3 correspond to the first, second and third (surface) layers, respectively.

In the DT analysis, we determined which energy regions were used to classify the LDOS. We divided the occupied energy range into 15 energy regions in steps of 0.03 eV and integrated the LDOS in each energy region. The integrated values were used as the features of DT. Then, we found which feature (i.e., energy region) had been used to divide which atoms. The highlighted energy regions shown in Figure S4 correspond to the features that were used in the DT. The PCA components are the linear combination of the high dimensions in the raw LDOS data, which is slightly complex, but the features in the DT clearly show which energy region has a large difference.

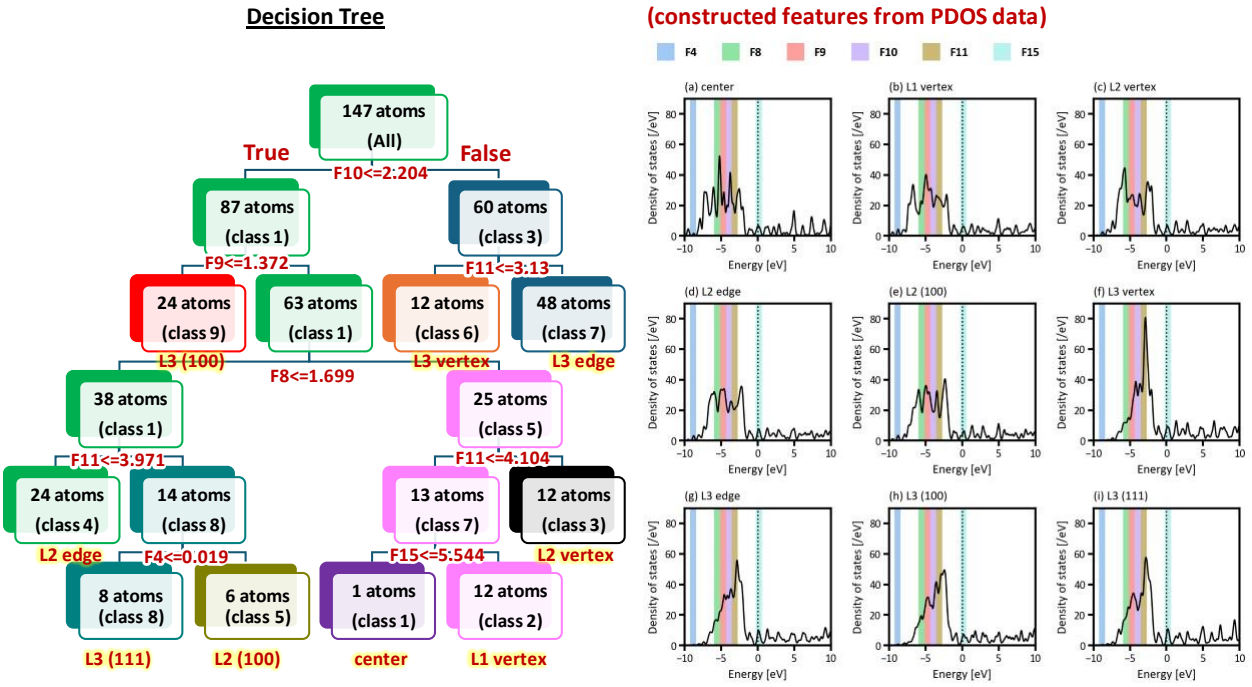

Figure S4. Decision tree analysis of LDOS of the atoms in the nanoparticle consisting of 147 atoms (O_h3). L1, L2 and L3 correspond to the first, second and third (surface) layers, respectively.

[S1] R. S. Mulliken, *J. Chem. Phys*. 1955, **23**, 1833–1840.