Supporting Information

Stability Prediction of Gold Nanoclusters with Different Ligands and Doped Metals: Deep Learning and Experimental Tests

Yuming Gu[a, †], *Shisi Tang*[a, †], *Xu Liu*[a], *Xinyi Liang*[a], *Qin Zhu*[a], *Hongfeng Wu*[b], *Xiao Yang*[c], *Weihao Jin*[a], *Hongwei Chen*[a], *Chunyan Liu* [d, *], *Yan Zhu* [a, *], *Jing Ma* [a, *]

[a]Key Laboratory of Mesoscopic Chemistry of Ministry of Education, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing, 210023, P. R. China

[b] Medical School, Kunming University of Science and Technology, Kunming, 650500, P. R. China

[c] National Engineering Research Center for Biomaterials, Sichuan University, Chengdu, 610064, P. R. China

[d] School of Chemistry and Chemical Engineering, Hunan University, Changsha, 410082, P. R. China

[†]These authors contribute equally

*E-mail: majing@nju.edu.cn, zhuyan@nju.edu.cn, cyliu@hnu.edu.cn

List of Contents

S1 Illustration of dataset

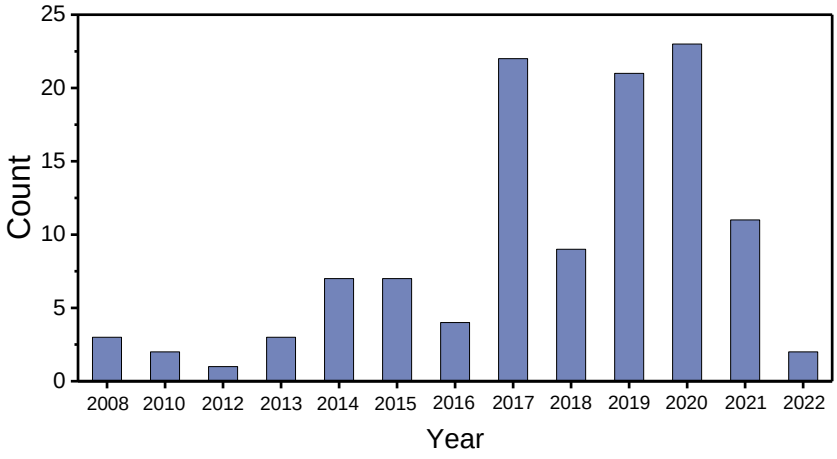

Fig. S1 The numbers of collected Au nanoclusters from literatures from the years 2008 to 2022.

The conformational sampling of ligands is obtained by using Molclus software, [S1] which sequentially invokes the xTB [S2] and Gaussian 16 [S3] programs. Initially, molecular dynamics (MD) simulations were conducted for 100 ps in the canonical ensemble (NVT) at the theoretical level of GFN0-xTB by employing the xTB software. A time step of 1 fs was employed, with output frames generated at intervals of 50 fs during the molecular dynamics simulations. The temperature was maintained at 400 K to enhance the efficiency of conformation sampling. Conformations were deemed highly similar when the energy deviation and root-mean-square deviation (RMSD) fell within the lenient thresholds of 0.5 kcal/mol and 0.4 Å, respectively. As a result, a representative conformation was chosen to characterize the entire ensemble. Subsequently, optimizations were performed for the selected low-energy conformations at the PBE/6-31G(d) level using Gaussian 16 software. Frequency calculations were conducted to validate that the optimized conformations indeed corresponded to local minima on the potential surfaces. Boltzmann distribution of conformations was evaluated at 298 K from the computed energy according to **equation S1**.

$$p_i = \frac{e^{-\Delta E_i/RT}}{\sum_{i=1}^{n} e^{-\Delta E_i/RT}} \quad \text{(S1)}$$

where ΔE_i is the energy of conformation i, R is the gas constant and T is the temperature, n is the number of the conformations selected for each molecule. The formation energy can be calculated as follows

$$E_{form} = \frac{E_{cluster} - N_{metal} * E_M - N_{ligand} * E_L}{N_{metal}} \quad \text{(S2)}$$

In **equation S2**, N_{metal} and N_{ligand} were the numbers of metal atom and ligand, respectively. E_{cluster} was the total energy of the nanocluster, E_{M} was the energy of the metal atom, and E_{L} was the average energy of the ligands with different configurations as follows

$$E_L = \sum_i^n p_i * E_{L\_i} \quad \text{(S3)}$$

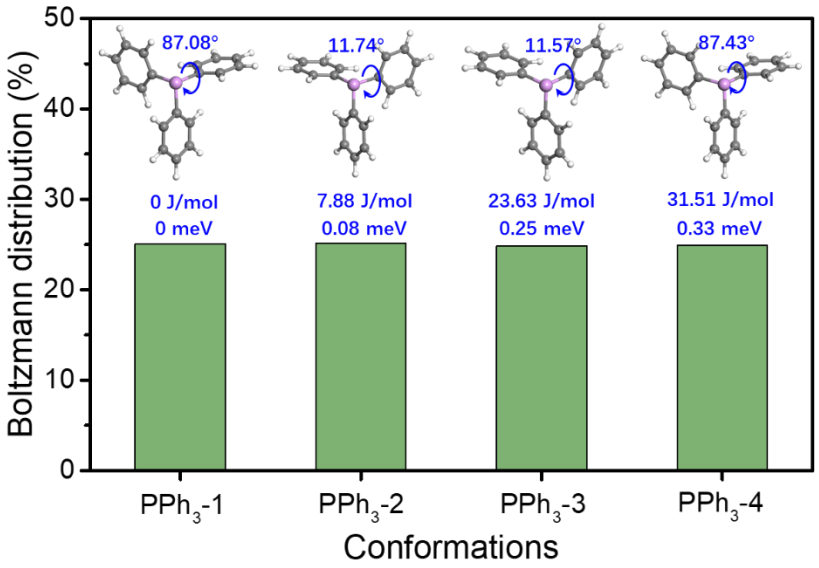

Fig. S2 Boltzmann distribution of different conformations for PPh_3 molecule, inserted structures are low-energy conformer estimated according to **equation S1**.

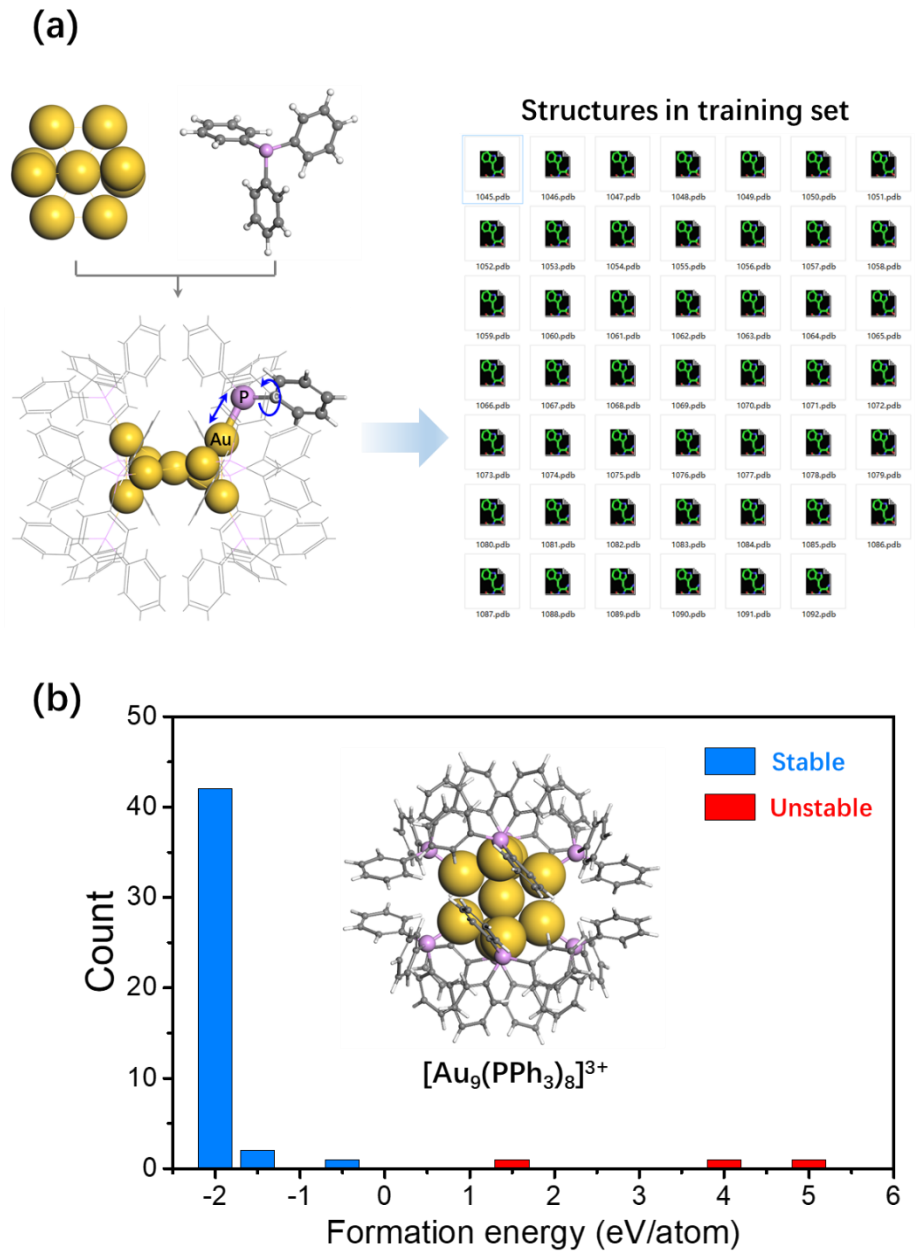

Fig. S3 (a) Illustration of structures generation in training set. (b) The formation energy distribution of $[Au_9(PPh_3)_8]^{3+}$.

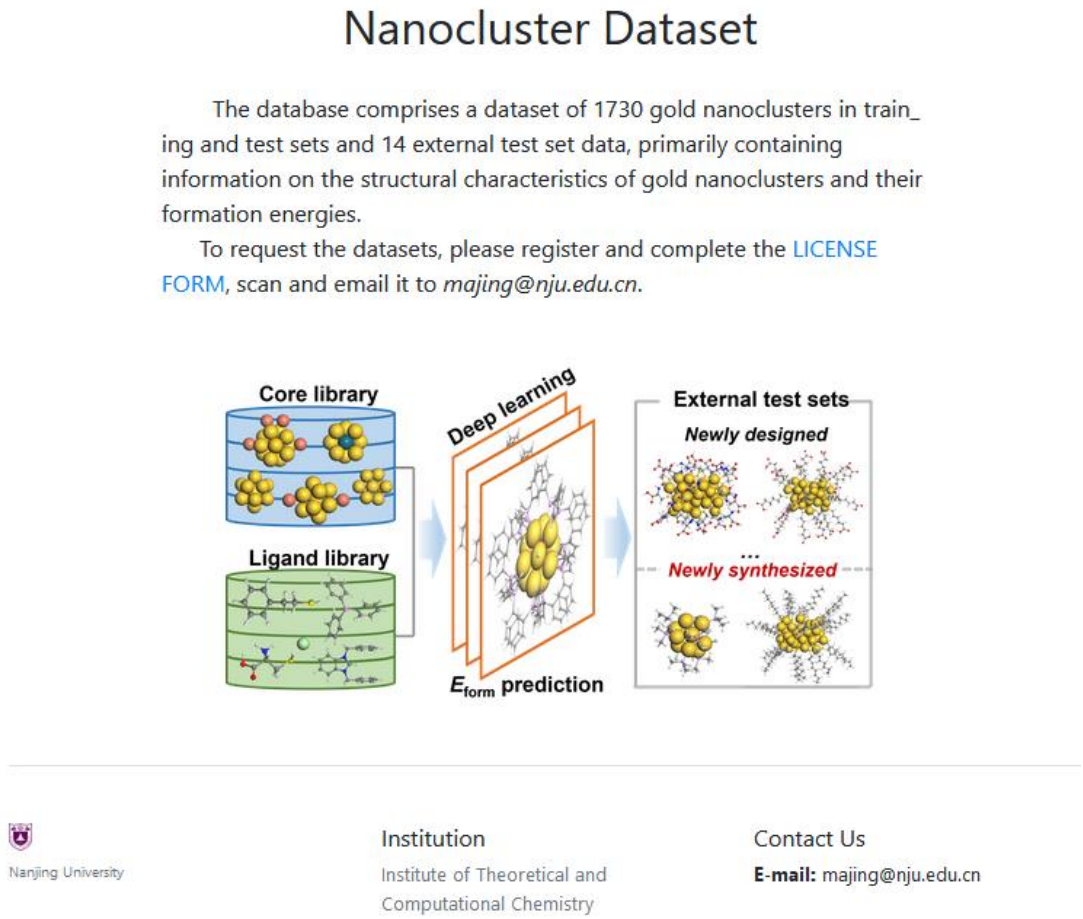

Fig. S4 Illustration of the webpage for Nanocluster dataset (http://106.15.196.160:5666/).

Table S1. Part of data (the most stable configuration) in dataset of deep learning to predict the formation energy

Nanoclusters [a]	E_{form}^{DFT} **(eV)**	E_{form}^{DL} **(eV)**
$[Au_4Pd_2(PET)_8]^0$	-3.36	-3.36
$[Au_4Pt_2(PET)_8]^0$	-4.16	-3.36
$[Au_4Ru_2(PET)_8(PPh_3)_2]^0$	-5.41	-5.41
$[Au_5Ru_2(PPh_3)_3Cl_2((CH_3)_5C_5)_2]^0$	-4.62	-4.61
$[Au_9(PPh_3)_8]^{3+}$	-1.79	-1.78
$[Au_8Pd(PPh_3)_8]^{2+}$	-2.46	-2.45
$[Au_9Pd(FPh_3P)_7Br_3]$	-1.21	-1.33
$[Au_{11}(PPh_3)_7Cl_3]^0$	-3.37	-3.35
$[Au_{11}(PPh_3)_8Cl_2]^+$	-3.19	-3.22
$[Au_{13}(dppe)_5Cl_2]^{3+}$	-2.08	-2.08
$[Au_{13}(NHC)_9Cl_3]^{2+}$	-3.64	-3.64
$[Au_{12}Ir(dppe)_5Cl_2]^+$	-3.54	-3.55
$[Au_{12}Ir(dppe)_5Cl_2]^{2+}$	-3.04	-3.06
$[Au_{12}Ir(dppm)_6]^{3+}$	-2.73	-2.73
$[Au_{12}Ru(dppm)_6]^{2+}$	-3.21	-3.21
$[Au_{12}Ru(dppm)_6]^{2+}$	-3.31	-3.32
$[Au_{14}Cd(SAdm)_{12}]^0$	-3.09	-3.08
$[Au_{14}Cd(t\text{-}BuS)_{12}]^0$	-17.46	-17.46
$[Au_{13}Cu_2(PET)_6(PPh_3)_6]^{1+}$	-3.37	-3.20
$[Au_{13}Cu_4(PPh_3)_4(PyS)_8]^{1+}$	-3.08	-3.04
$[Au_{15}Ag_3(CHT)_{14}]^0$	-3.17	-3.16
$[Au_{17}Cd_2(CHT)_{12}(dppp)_2]^+$	-3.05	-3.88
$[Au_{25}(PET)_{18}]^-$	-3.33	-3.62
$[Au_{25}(PET)_5(PPh_3)_{10}Cl_2]^{2+}$	-3.16	-3.14
$[Au_{38}(PET)_{24}]^0$	-3.25	-3.25
$[Au_8(PPh_3)_7]^{2+}$	-1.46	-2.11
$Au_{10}(PPh_3)_7Cl_3$	-2.55	-2.03
$Au_{18}(CHT)_{14}$	-3.22	-3.68
$Au_{25}(cis\text{-}R\text{-}AB)_{18}$	-	-6.67
$Au_{25}(cis\text{-}F\text{-}AB)_{18}$	-	-8.53
$Au_{25}(trans\text{-}F\text{-}AB)_{18}$	-	-8.43
$Au_{28}(CHT)_{20}$	-3.45	-3.83
$Au_{36}(TBBT)_{24}$	-	-8.43
$Au_{38}Cys_{24}$	-	-4.09
$Au_{38}MHA_{24}$	-	-6.89
$Au_{38}OT_{24}$	-12.86	-14.07
$Au_{38}OT_{18}MUS_6$	-	-18.40
$Au_{38}MUS_{24}$	-	-36.82
$Au_{44}(TBBT)_{28}$	-	-26.40

[a] Geometry coordination will be provided when you e-mail to majing@nju.edu.cn, if you are interested

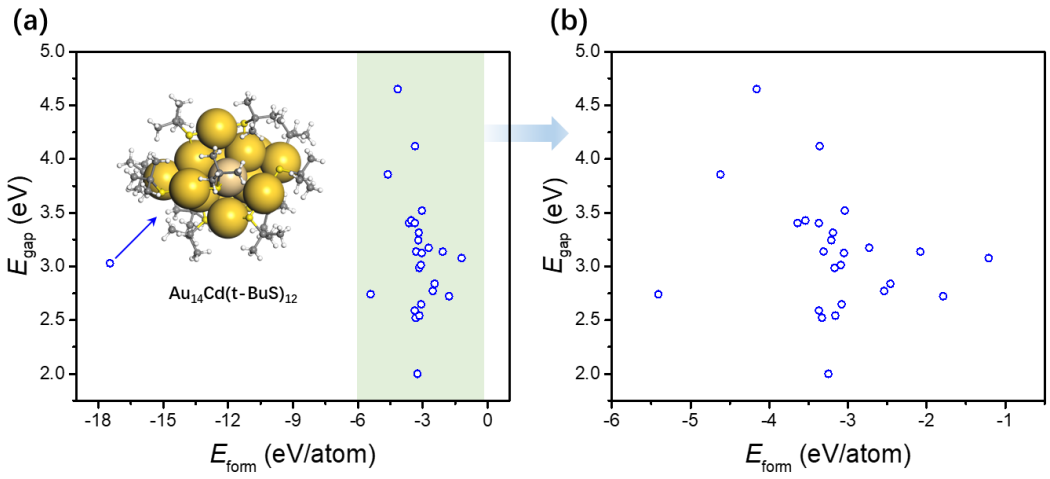

Fig. S5. The correlation between the formation energy (E_{form}) and HOMO-LUMO gap (E_{gap})

S2 Deep learning

DeepMoleNet, utilizing atom and atom pair representations as node and edge inputs, adeptly captures molecular information to autonomously predict the Au nanoclusters, as shown in **Fig. S6**. These inputs consist of real number vectors, employing one-hot encoding to encapsulate atom and atom pair details. The atom node inputs encompass atomic type, atomic number, van der Waals radius, and atom node degree. Edge features are characterized by bond type and Gaussian expanded distance. The hyper-parameters of DeepMoleNet are listed in **Table S2**, and Atom-Centered Symmetry Functions (ACSFs) are computed using the Dscribe package.

The Au nanocluster structures including 1730 data were employed in the graph convolution neural network, where 1384 data were chosen as the training set, 173 for validation set, and the rest for the test set.

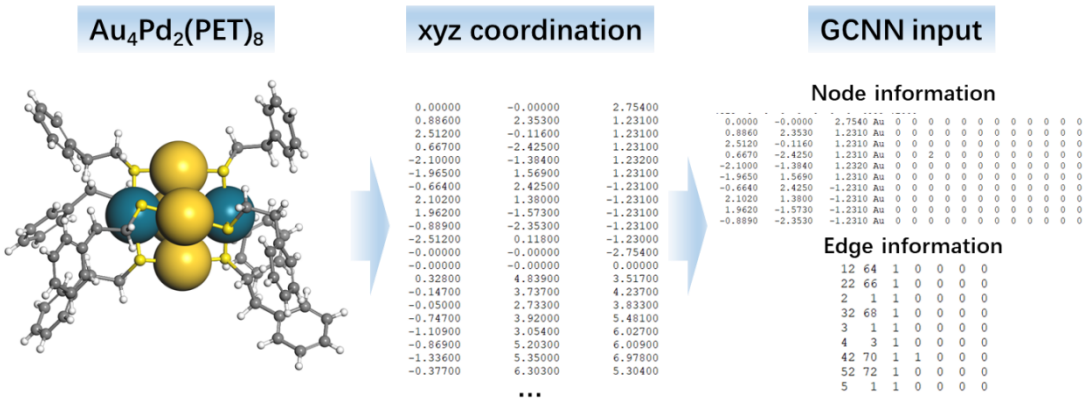

Fig. S6 The flow chart from the geometry of Au nanoclusters to the GCNN inputs.

Table S2. Hyper-parameters of graph convolutional neural network

Category	Hyper-parameter	Value
Message Passing Function	Message passing steps T	6
	Edge network layers	3
	Edge network hidden dim	128
Update Function	Node hidden units dim	64
	Attention layer	2
Readout Function	Atom-wise NN	2
	Output NN hidden layer	3
	Output NN hidden units	128
Auxiliary Target/ACSFs	Radial functions	[1, 1], [1, 2], [1, 3]
	Angular functions	[1, 1, 1], [1, 2, 1], [1, 1, -1], [1, 2, -1]
Training	Initial learning rate	2×10^{-3}
	Scheduler	Cosine Annealing LR
	Optimizer	Adam
	Batch size	64
	Training epochs	3000

Table S3. Performance with different learning rates of deep learning model via DeepMoleNet

Learning_rate	MAE (eV/atom)	R^2
0.9×10^{-3}	0.1282	0.9917
1.0×10^{-3}	0.1278	0.9836
2.0×10^{-3}	0.1222	0.9981
3.0×10^{-3}	0.1288	0.9962
4.0×10^{-3}	0.2926	0.8391

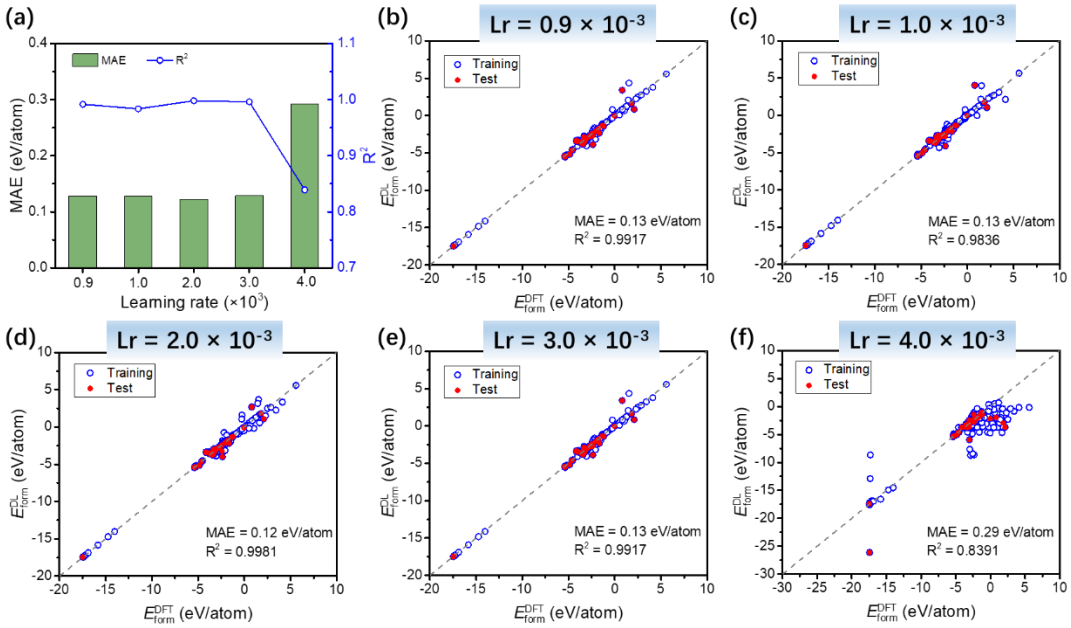

Fig. S7 (a) Comparison of the performance of different learning rates according to the mean absolute error (MAE) and coefficient of determination (R^2); (b-f) prediction of formation energy with different learning rates.

S3 Cell viability experiments of Insoluble Au nanoclusters

S3.1 Material preparation of $Au_9(PPh_3)_8$, $Au_{11}(PPh_3)_8Cl_2$, $Au_{18}(CHT)_{14}$, $Au_{25}(PPh_3)_{10}(PET)_5Cl_2$, $Au_{28}(CHT)_{20}$, $Au_{36}(TBBT)_{24}$, $Au_{44}(TBBT)_{28}$

Before inoculating cells, seven types of Au nanocluster materials (3 parallel samples for each material) were placed into a 24-well plate. 1 mL of 75% ethanol was added to each well, and the plate was placed in a biosafety cabinet. The 75% ethanol was allowed to air dry before use.

S3.2 Cellular cultivation

To investigate the antitumor properties of Au nanoclusters, rat osteosarcoma cells (UMR106) were selected. The cells were cultured in DMEM complete medium containing 10% fetal bovine serum and 1% penicillin–streptomycin solution, and maintained in a 37 ℃ incubator with 5% CO_2.

S3.3 Cell proliferation analysis

After digesting and resuspending UMR106 cells in the logarithmic growth phase, they were seeded into a 24-well plate at a density of 1.5×10^4 cells per well. The cells were incubated for 24 hours in a CO_2 incubator, after which the supernatant was removed. Each well was then supplemented with 1 mL of DMEM complete medium containing a specific concentration of Au nanoclusters for the experimental group, 1 mL of DMEM complete medium for the control group. For the blank group, only DMEM complete medium was added without cells and Au nanoclusters. Three replicate wells were prepared for each sample group. After 3 days of incubation, remove the culture medium, and replace it with fresh medium containing 10% CCK-8 reagent. Then the plate was placed in a light-protected incubator for 2 hours. Subsequently, 100 μL of the supernatant from each well was transferred to a 96-well plate (with 4-5 replicate wells per well), and the absorbance was measured at a wavelength of 450 nm using a microplate reader. Cell viability was obtained by the equation (S2):

Cell viability (%) = ($[OD]_{test}$-$[OD]_{blank}$) / ($[OD]_{control}$ -$[OD]_{blank}$) (S4)

Table S4. The cell viability of insoluble Au nanoclusters

	Au_9	Au_{11}	Au_{18}	Au_{25}	Au_{28}	Au_{36}	Au_{44}
1 group	17.47%	14.48%	41.53%	16.87%	15.32%	15.20%	21.30%
2 group	17.95%	15.56%	42.25%	16.99%	15.32%	15.56%	24.17%
3 group	18.43%	15.80%	40.21%	16.87%	15.32%	15.32%	23.10%
4 group	17.83%	15.32%	96.10%	16.99%	15.08%	14.24%	17.83%
5 group	19.15%	15.08%	95.14%	16.75%	15.92%	15.68%	16.99%
6 group	19.51%	15.92%	104.12%	16.87%	15.32%	14.84%	16.28%
7 group	18.43%	15.20%	96.70%	16.87%	15.32%	15.08%	16.63%
8 group	17.47%	14.36%	96.82%	16.52%	15.68%	14.24%	17.23%
9 group	18.43%	14.36%	101.84%	16.52%	15.56%	15.80%	19.03%
10 group	18.31%	15.32%	100.53%	16.63%	15.08%	15.44%	18.67%
11 group	18.30%	15.14%	81.52%	16.79%	15.39%	15.14%	19.12%

$Au_9 = Au_9(PPh_3)_8$, $Au_{11} = Au_{11}(PPh_3)_8Cl_2$, $Au_{18} = Au_{18}(CHT)_{14}$,

$Au_{25} = Au_{25}(PPh_3)_{10}(PET)_5Cl_2$, $Au_{28} = Au_{28}(CHT)_{20}$, $Au_{36} = Au_{36}(TBBT)_{24}$,

$Au_{44} = Au_{44}(TBBT)_{28}$

The experimental results indicate that due to some loss of the seven Au nanocluster materials during the sterilization process with 75% ethanol, the material concentrations in the experiment were lower than the theoretical concentrations. Nevertheless, six Au nanocluster materials demonstrate excellent anti-tumor effects with cell inhibition rates exceeding 80% except for $Au_{18}(CHT)_{14}$.

S4 Cell viability experiments of Soluble Au nanocluster

S4.1 Material preparation of $Au_{25}(PET)_8$

Prior to cell inoculation, the $Au_{25}(PET)_8$ material was placed in a biosafety cabinet, and subjected to 1 hour of ultraviolet irradiation for preparation.

S4.2 Cellular cultivation

Rat osteosarcoma cells (UMR106) were selected to investigate the anti-tumor properties of $Au_{25}(PET)_{18}$, while mouse pre-osteoblast cells (MC3T3-E1) were chosen as control cells to study the applicability of $Au_{25}(PET)_{18}$ in antitumor ability. UMR106 and MC3T3-E1 cells were separately cultured in DMEM complete medium and α-MEM complete medium, and incubated in a 37 ℃ incubator with 5% CO_2.

S4.3 Cell proliferation analysis

After digesting and resuspending UMR106 and MC3T3-E1 cells in the logarithmic growth phase, they were each seeded at a density of 1.5×10^4 cells per well in a 24-well plate. The cells were incubated for 24 hours in a CO_2 incubator, after which the supernatant was removed. Each well was then supplemented with 1 mL of complete medium containing a specific concentration of $Au_{25}(PET)_{18}$ material for the experimental group, 1 mL of complete medium for the control group, and for the blank group, only complete medium was added without cells and $Au_{25}(PET)_{18}$. Three replicate wells were prepared for each sample group. After 3 days of incubation, remove the culture medium, and replace it with fresh medium containing 10% CCK-8 reagent. Then the plate was placed in a light-protected incubator for 2 hours. Subsequently, 100 μL of the supernatant from each well was transferred to a 96-well plate (with 4-5 replicate wells per well), and the absorbance was measured at a wavelength of 450 nm using a microplate reader.

The experimental results show that $Au_{25}(PET)_{18}$ exhibits a high inhibition rate of 88.89% against UMR106 cells, demonstrating good anti-tumor effects. However,

$Au_{25}(PET)_{18}$ also inhibits the viability of MC3T3-E1 cells, suggesting certain biocompatibility concerns.

Table S5. The optical density of UMR106 and MC3T3-E1

	UMR106		MC3T3-E1	
	control	$Au_{25}(PET)_{18}$	control	$Au_{25}(PET)_{18}$
1 group	2.1585	0.2413	1.4104	0.2357
2 group	2.2546	0.2050	1.3863	0.2006
3 group	2.2300	0.2001	1.5053	0.1999
4 group	2.2010	0.1966	1.4857	0.1925
5 group	2.3781	0.2047	1.4910	0.1977
6 group	2.3509	0.2028	1.4974	0.1997
7 group	2.3241	0.1997	1.4935	0.1980
8 group	2.3656	0.1968	1.4760	0.1931
9 group	2.1805	0.2017	1.4789	0.1987
10 group	2.1303	0.2002	1.4907	0.2007
11 group	2.1488	0.1997	1.4648	0.1949
12 group	2.2450	0.1925	1.5218	0.1922

S5 Synthesis experiments of $Au_{10}(PPh_3)_7Cl_3$ and $Au_{38}(OT)_{24}$

S5.1 Chemicals.

All chemicals and reagents are commercially available and used as received. 2-phenylethanethiol (PET), p-toluenethiol, tetra-octylammonium bromide (TOAB) Tetrachloroauric (III) acid ($HAuCl_4 \cdot 4H_2O$), and sodium borohydride ($NaBH_4$) were purchased from Aladdin. Methanol (CH_3OH), dichloromethane (CH_2Cl_2), acetonitrile (CH_3CN), toluene and petroleum ether were obtained from Sinopharm Chemical Reagent. Co. Ltd. The water used in all experiments was ultrapure with the resistivity of 18.2 MΩ·cm produced by a Milli-Q NANO pure water system.

S5.2 Synthesis of $Au_{10}(PPh_3)_7Cl_3$ nanocluster.

$Au_{10}(PPh_3)_7Cl_3$ was synthesized by the co-reduction of the $AuPPh_3Cl$ and $Pt(cod)Cl_2$ with sodium borohydride ($NaBH_4$). After the two metal precursors were fully dissolved in the dichloromethane/ethanol mixed solution, the reducing agent was added under stirring and the reaction was carried out for 3 hours. Next, the reaction solution was concentrated and aged at room temperature for a week. Finally, the target product was separated and purified by thin layer chromatography.

S5.3 Synthesis of $Au_{38}(OT)_{24}$ nanocluster.

Firstly, 10 mg of $Au_{38}(PET)_{24}$ nanoclusters were dissolved in 1 mL of toluene. Secondly, 500 μL of 1-octanethiol was added into the solution, the reaction then was

carried out under 80 ℃ for 4 hours in an oil bath. The resultant was precipitated with the addition of methanol and washed twice with excess methanol. Subsequently, the $Au_{38}(OT)_{24}$ nanoclusters were further separated by preparation thin layer chromatography (PTLC) with the mixture of dichloromethane and petroleum ether (v/v =1/4) as developing solvent.

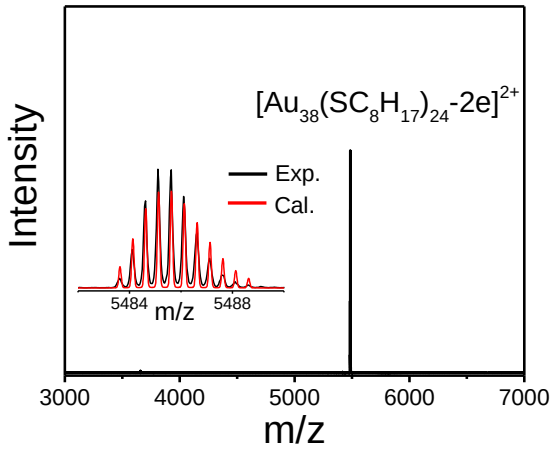

Fig. S8 ESI-MS data of $Au_{38}(OT)_{24}$ nanoclusters. The inset graph shows the comparison between the experimental and simulated isotopic pattern.

Reference

S1 T. Lu, molclus program, Version 1.9.9.5. http://www.keinsci.com/research/molclus.html (accessed 2020-10-1).
S2 C. Bannwarth, E. Caldeweyher, S. Ehlert, A. Hansen, P. Pracht, J. Seibert, S. Spicher and S. Grimme, *WIREs Comput. Mol. Sci*. 2021, **11**, e1493.
S3 M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai and T. Vreven, Gaussian 16, Revision A. 03. Gaussian, Inc., Wallingford, CT, 2016. Gaussian.