



HAL
open science

Causal inference machine learning leads original experimental discovery in CdSe/CdS core/shell nanoparticles

Rulin Liu, Junjie Hao, Jiagen Li, Shujie Wang, Haochen Liu, Ziming Zhou, Marie-Hélène Delville, Jiaji Cheng, Kai Wang, Xi Zhu

► **To cite this version:**

Rulin Liu, Junjie Hao, Jiagen Li, Shujie Wang, Haochen Liu, et al.. Causal inference machine learning leads original experimental discovery in CdSe/CdS core/shell nanoparticles. *Journal of Physical Chemistry Letters*, 2020, 11, pp.7232-7238. 10.1021/acs.jpcllett.0c02115 . hal-02922356

HAL Id: hal-02922356

<https://hal.science/hal-02922356v1>

Submitted on 26 Aug 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The Causal Inference Machine Learning Leads Original Experimental Discovery in CdSe/CdS Core/Shell Nanoparticles

Rulin Liu^{1, †}, Junjie Hao^{2,3 †}, Jiagen Li¹, Shujie Wang¹, Haochen Liu², Ziming Zhou²,

*Marie-Hélène Delville³, Jiaji Cheng⁴, Kai Wang^{2, *}, Xi Zhu^{1, *}*

1 Shenzhen Institute of Artificial Intelligence and Robotics for Society (AIRS),

The Chinese University of Hong Kong, Shenzhen, Guangdong 518172, China

2 Department of Electrical and Electronic Engineering, Southern University of
Science and Technology, Shenzhen, 518055, China

3 CNRS, Univ. Bordeaux, Bordeaux INP, ICMCB, UMR 5026, Pessac, F-33608,
France

4 School of Materials Science and Engineering, Hubei University, Wuhan 430062,
China

† These authors contributed equally to this study.

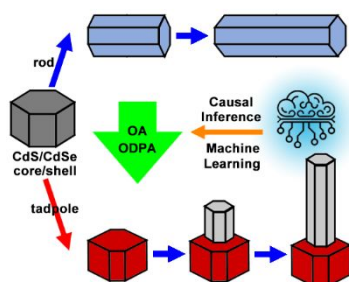
* Address correspondence to

wangk@sustc.edu.cn, zhuxi@cuhk.edu.cn

1
2
3
4 **Keywords:** Machine Learning, Causality Analysis, Entropy, Nucleation Theory,
5
6 Quantum Dots, Heterostructure
7
8
9
10

11
12
13 **ABSTRACT:** The synthesis of CdSe/CdS core/shell nanoparticles was revisited with
14 the help of a causal inference machine learning framework. The tadpole morphology
15 with 1-2 tails was experimentally discovered. The causal inference model revealed the
16 causality between the oleic acid (OA), octadecyl-phosphonic acid (ODPA) ligands and
17 the detailed tail shape of the tadpole morphology. Further, with the identified causality,
18 a neural network was provided to predict and directly lead to the original experimental
19 discovery of new tadpole-shaped structures. An entropy-driven nucleation theory was
20 developed to understand both the ligand & temperature dependent experimental data
21 and the causal inference from the machine learning framework. This work provided a
22 vivid example of how the artificial intelligence technology, including machine learning,
23 could benefit the materials science research for the discovery.
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

TOC Graphic



The application of artificial intelligent technology, including machine learning (ML) in materials science, has recently generated intense research attention.¹⁻⁶ The utilization of the ML technology could significantly reduce the workload and accelerate discoveries in both computational and experimental explorations.⁶⁻⁸ As a numerical methodology, the ML technology is advantageous in extremely complicated systems, such as nanoscale materials.

In the present work, we revisited the CdSe/CdS core/shell quantum dots (QDs), a popular type of optoelectronic materials.⁹⁻¹¹ The morphologies of the CdSe/CdS core/shell structure are as various as hexagonal shape¹², triangular shape¹³⁻¹⁴, rod-like shape^{10, 15}, etc. The causality of the morphology formation is ingenious yet intricate, where the general understanding is *via* ligand passivation¹⁶⁻¹⁷ for the preference of ligands to bind with specific facets.¹⁸⁻¹⁹ One of the routine methods to investigate the ligand-surface binding phenomena and ligand-morphology causality is the density functional theory (DFT)¹⁶. However, except for the intrinsic drawbacks,²⁰ temperature also degrades the performance of DFT. The surface energy provided by DFT is a

1
2
3
4 thermodynamic quantity at zero K, yet most of the ligand-surface binding events occur
5
6 above 500 K. Thus, at high temperature where QDs nucleate and grow, it is difficult to
7
8 trust the temperature-morphology causality derived from the zero-temperature-level
9
10 DFT methodology.
11
12

13
14 Since most of the properties of QDs are strictly dependent on the morphology,²¹⁻²² an
15
16 alternative and rational methodology is essential to understand the morphology
17
18 causality. A causal inference in ML framework is hereby provided to explore the long-
19
20 standing morphology causality problems in CdSe/CdS core-shell QDs. An ML
21
22 framework was trained, which automatically found that oleic acid (OA) is the principal
23
24 factor for tadpoles. A separate neural network was trained to predict the ODP amount
25
26 dependence of the number of tails. The predictions for the three-tail and five-tail tadpole
27
28 structures agree well with the experimental data. Besides the causal inference, the
29
30 growth mechanism of the tadpoles was also analyzed *via* entropy endorsed nucleation
31
32 theory.
33
34
35
36
37
38
39
40
41
42

43 The synthesis was done in AIR-Chem²³ and takes two steps as in **Figure 1a-b**.
44
45 Detailed descriptions can be found in **S1**. The process yielded tadpoles at 320 °C, and
46
47 triangles at 260 °C.
48
49
50
51
52
53
54
55
56
57
58
59
60

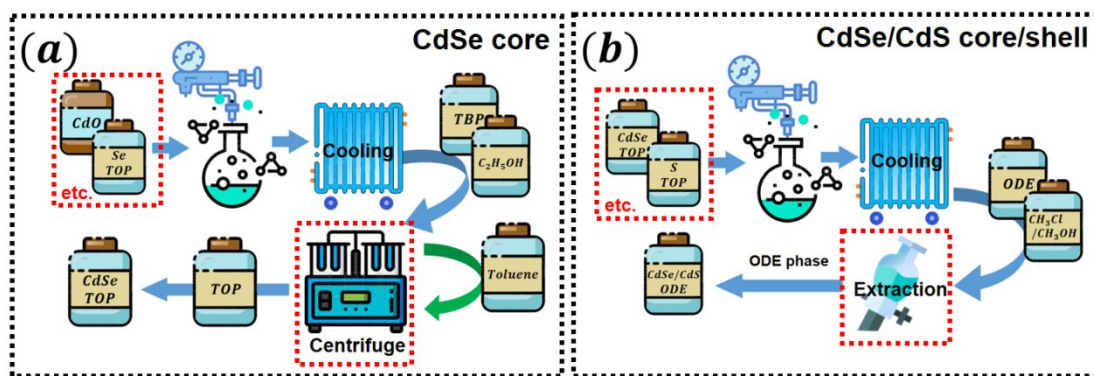


Figure 1, Experimental setup. Experimental setup of the synthesis process of the wurtzite CdSe cores, and the seeded growth to the CdSe/CdS core-shell structures. Red dotted frames in Figures 1a and 1b emphasize the difference between the two procedures. The full method is detailed in **S1**.

As an illustration, the tadpole morphology was obtained as mentioned in **S1**, and its characterizations by TEM versus time are shown in **Figure 2**. **Figures 2a-g** show the TEM images of the CdSe/CdS core-shell samples at different growth reaction times (0 s, 30 s, 1 min, 2 min, 4 min, and 8 min, respectively) in the presence of OA (1.02 ml) and ODPA (285 mg). The QDs first grew triangular, and a tail-like shape formed only after 2 minutes and kept on growing until 8 minutes, as shown in **Figures 2f** and **2g**. The tail grows along the [0002] axis of the wurtzite CdS phase²⁴, the CdS phase is coherent with the head part and the CdSe core in the center of this head. The tadpole shape indicates a complete anisotropic growth pattern due to the routine selective facet coating of ligands.¹⁶

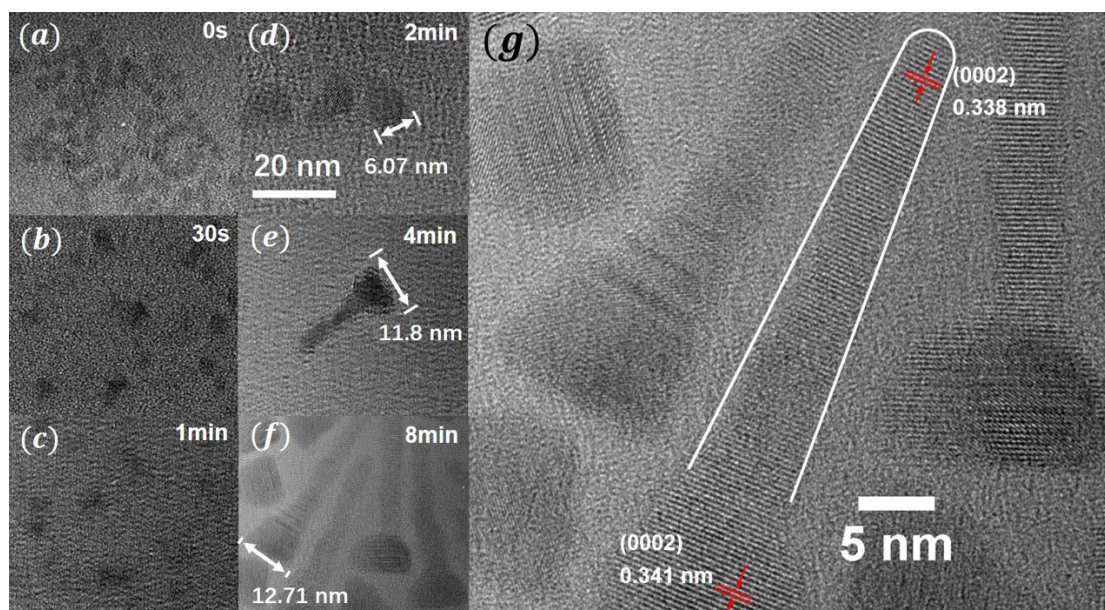


Figure 2, Time-dependent TEM images. TEM images of the tadpole morphology with respect to time: **(a)** core at 0 s, **(b)** 30 s, **(c)** 1 min, **(d)** 2 min, **(e)** 4 min, **(f)** 8min. The nanoparticles grew first into a triangular shape, and the tail started to grow between 2 min and 4 min. The images before 2 min are too blurred to recognize any pattern. **(g)** A closer look at the 8 min products. The fringes show clearly that the tail grew along the (0002) facet.

DFT is one of the popular methodologies used to study the surface science problems^{16, 25-28}, which has successfully explain numerically some ligands-facets related anisotropic growth problem.^{16, 25-30} Yet, the “DFT success” is still an issue because of its treatment of temperature.

Here, we present the causal inference (CI) in ML framework³¹⁻³³ to identify the causal relationship between the experiments parameters and the morphology of the resulting nanomaterials. The CI takes into account the treatment $a \in \{0, 1\}$, the covariates X ,

1
2
3
4 and the response y , to search for the most efficient covariates, and then uses methods
5
6 such as inverse probability weighting³⁴ to eliminate the difference resulting from the
7
8 distribution of these covariates. The causal relationship was evaluated *via* the
9
10 standardized mean difference (SMD) of the covariates by dividing the data into two
11
12 classes based on the treatment a , as the horizontal axis in **Figure 3a**. The SMD of two
13
14 groups are
15
16
17
18
19

$$SMD = \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}},$$

20
21
22
23
24 where μ is the mean of the covariates in different classes, σ is the standard
25
26 deviation, and the subscripts denote the classes of data. The two considered classes are
27
28 the two denoted in the treatment a . This parameter is effective in expressing the size
29
30 of the intervention effect of different covariates in X . The covariates with the highest
31
32 standardized mean difference are believed to be the most relevant to the variability of
33
34 the response y . In our case, X refers to the space of experimental parameters, y is
35
36 the feature space of the QDs. Since the Bayesian probability $P(\text{tadpole} | \text{ligand})$ of
37
38 OA is the largest, the addition of OA was taken as the treatment a . Since the
39
40 temperature and time are trivial as causalities for the QDs growth, we focus more on
41
42 the ligands, such as the effects of ODPA on the number of tails, and study $P(n_{\text{tail}}$
43
44 $| \text{ODPA}, T = 320^\circ\text{C}, t = 8\text{min})$. As shown in **Figure 3a**, it was found that OA and
45
46 ODPA were most relevant to the feature “tail or not”. The different synthesis routes
47
48 have been sketched in **Figure 3a**, where the coloring scheme is the same as in **Figure**
49
50
51
52
53
54
55
56
57
58
59
60

4a, denoting the different growth mechanisms that are going to be analyzed in the following sections.

CI is capable of explaining the role of ligands in the formation of tadpole in a qualitative way. With more data acquired,³⁵⁻⁴⁴ the CI could extend its capability to the synthesis of other morphologies of other materials. Data collected from other publications are listed as Supplementary file S4 “Experimental data” and some related features analyzed in S5.

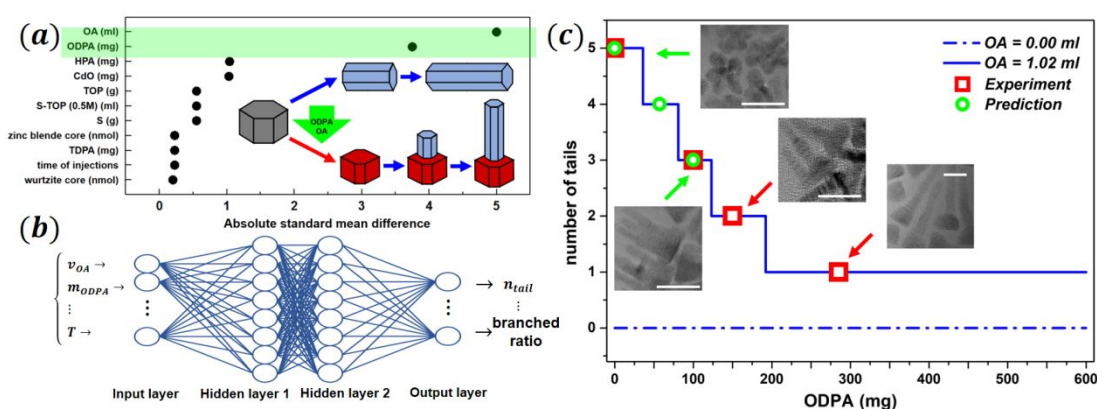


Figure 3, Causal Inference Analysis. (a), Causal relationship for the transition from quantum rod to tadpole. The SMD for OA and ODPa are the highest. The figure inserted shows that the growth route can be altered by the addition of OA and ODPa. The coloring scheme of the structures is the same as the ones shown in Figure 4a. (b) The detailed structure of the neural network. The *Hidden layer 1* and the *Output layer* both take a *tanh* activation function and a linear operation afterward, while the others are purely linear. The dimension of the neural network is $n \times 8 \times 8 \times m$, where n is the number of controls, and m is the number of QD features. In the experiment present, $n = 4$ refers to the amount of OA, ODPa, HPA, temperature, and $r = 1$

1
2
3
4 refers to the number of tails, respectively. (c), Neural network predicted causal
5
6 relationship between the amount of ODPA and the number of tails at 320 °C, trained
7
8 with the experimental data. The solid red rectangles refer to the experimental data with
9
10 tail numbers 1 and 2 and their corresponding TEM images; the green circles refer to the
11
12 predicted structures made by the neural network with 3, 4 and 5 tails. The scale bars in
13
14 the figures denote represent 20 nm. Full-sized TEM image is provided in **S6**.
15
16
17
18
19
20
21
22

23 The causal inference provided a qualitative understanding, while the neural network
24
25 is quantitative. The neural network is shown in **Figure 3b**, whose structure provided
26
27 was chosen because of its monotonicity and proper numerical range. The dimension of
28
29 the neural network is $n \times 8 \times 8 \times m$, where n is the number of controls, and m is
30
31 the number of QD features. In the present experiment, $n = 4, m = 1$, refer to OA,
32
33 ODPA, HPA, temperature (the highest causality with the tadpole morphology in **Figure**
34
35 **3a**) as four inputs for n , and the number of tails as the output m . The loss function was
36
37 MSE, and an adaptive backpropagation was conducted for training.
38
39
40
41
42
43
44

45 Since ODPA was found to be relevant to the number of tails, the prediction between
46
47 the number of tails and the amount of ODPA was first examined using the trained
48
49 network, in which OA was fixed at 0, 1.02 mL, and HPA at 0 g (HPA was found to be
50
51 relevant to the formation of rods, and it did not appear in the formal experiments where
52
53 tadpole was first discovered, so it was set to 0 g), and ODPA ranging from 0 to 600 mg.
54
55 The result is plotted in **Figure 3c**, where the prediction agrees well with the
56
57
58
59
60

1
2
3
4 experimental data for 0, 1, 2 tails, and the prediction of new morphologies with 3, 4, 5
5
6 tails is also given.
7

8
9 According to the prediction in **Figure 3c**, 100 mg, 50 mg, 5 mg (picked from around
10 the middle of the temperature range shown by the prediction curve) of ODPA was
11 picked for validation. By applying these parameters, we successfully synthesized
12 tadpole with 3 and 5 tails with 100mg and 5mg ODPA, the corresponding TEM images
13 of the samples were shown beside the prediction data points in **Figure 3c**. The 4-tail
14 samples were not obtained maybe because of the narrow tolerance window.
15
16
17
18
19
20
21
22
23

24 As shown in **Figure 3a**, OA was the main reason accounting for the formation of
25 tadpoles. OA restricted the growth of the elongated part along the radial directions ($[01$
26 $\bar{1}0]$ and $[10\bar{1}0]$) of CdSe/CdS because of the OA/facets binding.⁴⁵ Thus, the vertical
27 elongation carried on, while the horizontal does not, which resulted in the thinner tail.
28
29 Another proof is that the more the number of tails of a tadpole QD, the larger its head
30 (as shown in the TEM images provided in **Figure 3c**), since more tails provides larger
31 surface area, which dissipates the OA/facet binding. ODPA is relevant to the growth
32 along the vertical direction $\langle 0001 \rangle$, and thus the number of tails in a discrete manner,
33 as shown by the blue lines of CI and neural network results in **Figure 3c**. Upon reaching
34 one threshold by increasing the amount of ODPA, the number of the tails decreases by
35 one. Quantitative calculations of the ligands can be found in **S7**.
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

53 Besides the ligands, the reaction temperature and time also matter. For the same
54 experimental conditions as in **Figure 2**, but with temperature changed from 320 °C to
55 260 °C, the morphology would remain small and triangular even after 8 min (the time
56
57
58
59
60

dependent size is as depicted in the 260 °C curve in **Figure 4b**, and the morphology is similar to the one in **Figure 2d**). Since the tadpole occurs only at a higher temperature (320 °C), and the morphology is different from the Platonic solids described with enthalpy in classical theories,⁴⁶⁻⁴⁸ and since the tadpole morphology is an hybrid of sphere and rod, the mechanism could be an entropy-driven two-step nucleation. The entropy considered in the present work is assumed to be proportional to the surface area of the QDs by facet dependent coefficients defined as $\chi(Jm^{-2}K^{-1})$.⁴⁹ The expression of the free energy from the surfaces is thus:

$$\Delta G = \Delta H - T\Delta S = \sum_i (\gamma_i - T\chi_i)\Delta A_i \#(1)$$

where the subscript i refers to different facets, γ is the surface energy, ΔA is the change of the surface area. Thus, the morphology favoured under different temperature ranges can be specified by taking the minimum of the free energy ΔG .

Assuming the volume of the hexagonal prism QDs is V , and the additional volume from growth is ΔV ($\Delta V \ll V$). If ΔV grows to keep the hexagonal prism,

$$\Delta A_{hex} = \left(\frac{9\sqrt{3}}{16} + 3\right) \frac{16}{9\sqrt{3}} V^{\frac{2}{3}} \frac{\Delta V}{V} \#(2)$$

if ΔV induces a triangular morphology,

$$\Delta A_{tri} = \left(\frac{8}{3\sqrt{3}}\right)^{\frac{2}{3}} \left(\frac{3}{2} + \frac{3\sqrt{3}}{4}\right) V^{\frac{2}{3}} \frac{\Delta V}{V} \#(3)$$

if ΔV gives a rod-like morphology along the (0001) facet,

$$\Delta A_{rod} = \frac{16}{\sqrt{3}} \left(\frac{3\sqrt{3}}{8}\right)^{\frac{1}{3}} V^{\frac{2}{3}} \frac{\Delta V}{V} \#(4)$$

Eq (2) - (4) was plugged into eq (1) and ΔG minimized to determine the temperature-dependent favored facet. Note that ΔG depends on V and ΔV , so it is only upon

reaching a critical volume that the growth turns to favor another morphology. One can thus define the most favored morphology as a function of the temperature, together with the size where the transition starts. **Figure 4a** shows the ΔG plot with $V = 158.47n \text{ m}^3$, $\Delta V = 1 \text{ nm}^3$. The gray, red and blue colors refer to different growth mechanisms analyzed above, and the directions and lengths of the green arrows denote the growth directions and rates. An explicit temperature dependence for the morphologies can be observed: below 140°C , hexagonal is favored; while a triangle phase emerges in the range $140^\circ\text{C} \sim 320^\circ\text{C}$ ⁵⁰⁻⁵²; the rod-like morphology emerges at 320°C . This corresponds to the transition from triangular to rod-like morphology as shown by the TEM images in **Figure 2**.

Next, the kinetic mechanism of the tadpole morphology was investigated. The growth rates of the head and tail parts of the tadpole morphology are:

$$dr_{head} = dr, \quad (5)$$

$$dl_{tail} = \frac{\Theta(\Delta H_{diff} - T\Delta S_{diff})}{\sum_{n_{tail}=1}^5 n_{tail} (\Theta(c_{ODPA} - c_{tail}^{n_{tail}}) - \Theta(c_{ODPA} - c_{tail}^{n_{tail}-1}))} dh$$

$$= \begin{cases} 0, & \text{if } \Delta H_{diff} < T\Delta S_{diff} \\ \frac{dh}{n_{tail}}, & \text{if } \Delta H_{diff} > T\Delta S_{diff} \end{cases} \quad (6)$$

where

$$\frac{dr}{dt} = \frac{v_m c_\infty D_{head} k_{head}}{D_{head} + k_{head} r} \left(\frac{2\gamma_{head} v_m}{RT r} \right) \left(r \left(\frac{2\gamma_{head} v_m}{RT} \frac{c_\infty}{c - c_\infty} \right)^{-1} - 1 \right)$$

$$\approx \frac{2\gamma_{head} v_m^2 c_\infty D_{head}}{RT r^2} \left(r \left(\frac{2\gamma_{head} v_m}{RT} \frac{c_\infty}{c - c_\infty} \right)^{-1} - 1 \right), \quad (7)$$

$$\frac{dh}{dt} = \frac{v_m c_\infty D_{tail} k_{tail}}{D_{tail} + k_{tail} (\lambda - h)} \left(\frac{2\gamma_{tail} v_m}{RT r_{tail}} \right) \left(r_{tail} \left(\frac{2\gamma_{tail} v_m}{RT} \frac{c_\infty}{c - c_\infty} \right)^{-1} - 1 \right). \quad (8)$$

r_{head} in eq 5 is the radius of the tadpole's head, l_{tail} in eq 6 is the length of the tail; r and h in eqs 5 and 6 refer to the radius and height of the sphere and cylinder (detailed in **S8**) that form a simplified spherical morphology with a cylindrical tail, into which the tadpole morphology has been simplified. v_m is the molar volume of the material; c_∞ is the bulk concentration of the monomers, c is the concentration of the monomers in solution; D, k are the diffusion and reaction coefficients, where the subscripts denote the part they refer to; γ is the surface energy, with subscripts referring to the different parts; λ is a considerable enough distance from a QD, where the concentration of the monomers c does not deviate from c_∞ (detailed in **S8**).

The subscript *diff* of ΔH and ΔS in the numerator of eq 6 refers to the difference in them between the triangular and tadpole-like morphologies. Note that ΔH_{diff} refers to the chemical bonding energy which is almost temperature independent, thus, $T\Delta S_{diff}$ becomes more decisive at high temperatures. As can be observed from this model, an increasing monomer concentration would result in a longer and pointier tail, since that the growth route according to **Figure 4a** depends on only the temperature and ligands, and that the parameters D, k, c will increase with the monomer concentration.

The growth rate of the triangle morphology is

$$\frac{dx}{dt} = \frac{kc_\infty\gamma v_m^2}{RT} \left(\left(\frac{\gamma v_m c_\infty}{RT c - c_\infty} \right)^{-1} - \frac{1}{h} + \frac{1}{x - 2l} \right), \quad (9)$$

$$\frac{dl}{dt} = \frac{Dk v_m c_\infty \gamma v_m}{D \ln(\delta/l) + kl RT} \left(\left(\frac{\gamma v_m c_\infty}{RT c - c_\infty} \right)^{-1} - \frac{1}{h} + \frac{1}{l - 2x} \right). \quad (10)$$

where x is the distance from the center of the triangular plane to the edge of the facet that grows faster, and l is the distance to the slower one, δ is defined the same as λ above, h here is different from above and denotes the thickness of the triangle. The models are detailed in S9.

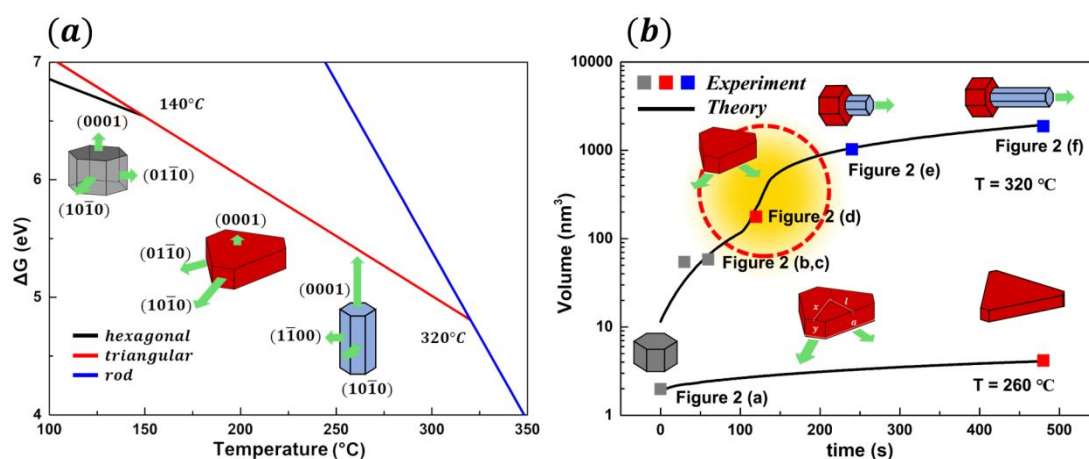


Figure 4, The entropy-driven nucleation theory. (a), ΔG of growth into different morphologies, plotted with respect to the temperature. Different colors in the figure refer to different growth mechanisms; the direction of the green arrows refer to the growth directions, and the magnitude of them refer to the respective growth rate. Since this plot was calculated with $V = 158.47\text{nm}^3$, as measured from the TEM in **Figure 2d**, it is only when the nanoparticle grows to this size that it begins to grow a tail, while below which the growth will be triangular, accounting for the triangular shape that occurred in the TEM image at 2 min. The surface entropy⁴⁹ $\chi_{\text{uniform}} = 10 \times 10^{-4}\text{Jm}^{-2}$, $\chi_{\text{triangle}} = 15 \times 10^{-4}\text{Jm}^{-2}$, $\chi_{\text{tail}} = 20 \times 10^{-4}\text{Jm}^{-2}$. The surface energies γ were taken as the same in the growth models. **(b),** The total volume of the tadpoles at $T = 320$ $^{\circ}\text{C}$, together with the triangle at $T = 260$ $^{\circ}\text{C}$. Both the theoretical and

1
2
3
4 experimental results are included. The data point labeled as **Figure 2a** shows up in both
5
6 theoretical curves since it refers to the seed. The golden-colored red-dotted-line-circled
7
8 emphasized area in the 320 °C curve denotes the emergence of the tail, as the second
9
10 stage of the growth.
11
12
13
14
15
16

17 **Figure 4b** shows the numerical plot of eq 5 and 6, the time-dependent growth at
18
19 320 °C, and eq 9 and 10 at 260 °C. The colors in **Figure 4b** refer to the same growth
20
21 mechanisms in Figure 4a. The experimental data shown in **Figure 2a** to **2f** are also
22
23 labelled in Figure 4b, which agrees well with our theoretical model. By eq 6, the
24
25 entropic competition between ΔH and $T\Delta S$ items in the θ function is the critical
26
27 factor for the system to choose the triangle or tadpole channel.
28
29
30
31

32 The emergence of the tadpole morphology requires a more aggressive volume
33
34 increase, indicating a three-stage growth mechanism induced by the entropy. The three
35
36 stages are: (1) the growth of the head to reach a critical volume at a given temperature,
37
38 which will be slow and diffusion-controlled in later time; (2) the fast reaction controlled
39
40 emergence of the tail; and (3) the slowing down of the tail's growth, where both the
41
42 head and the tail end up in diffusion-controlled growth. This mechanism cannot be
43
44 established without the help of entropy, for enthalpy analysis considers the competition
45
46 amongst existing facets, rather than the appearance of a new one. The enthalpy analysis
47
48 based on surface energy and chemical potential⁴⁷ can describe the morphology
49
50 transition of nanoparticles, yet it is more like a ripening phenomenon of existing
51
52 surfaces. The endorsement of entropy enables us to find another alternative growth
53
54
55
56
57
58
59
60

1
2
3
4 route for the nanoparticles besides the morphology transition from enthalpy
5
6 competition.
7
8
9

10
11 To conclude, in this work, the causal inference has successfully learned from the
12 existing experimental data of ours of 0, 1, 2-tail tadpoles, and managed to provide
13
14 predictions for 3, 4, 5-tail tadpoles. Experiments were then conducted, and 3, 5-tail
15
16 tadpoles were synthesized as expected. An entropy-driven three-stages growth
17
18 mechanism was developed under the guidance of causal inference. This work provides
19
20 a vivid example of how ML can help researchers in recognizing the pattern of the
21
22 experimental data, and even in developing new knowledge.
23
24
25
26
27
28
29
30
31
32

33 **ASSOCIATED CONTENT**

34 **Supporting Information:**

35
36
37 Additional data and figures are included in the Supplementary Information. S1:
38
39 Experimental methodology; S2: Characterization of tadpoles, S3: Drawbacks of DFT;
40
41 S4: Experimental data (The experimental data from literatures, which were used for the
42
43 training of CI, were formatted in an Excel-formatted PDF file and attached in an “S4.
44
45 Experimental Data.pdf”), S5: CI for other morphologies, S6: Full TEM images of
46
47 tadpoles, S7: Attachment of ligand molecules, S8: Derivation of growth in reduced
48
49 dimensions, S9: Derivation and numerical calculation of triangular and tadpole
50
51
52
53
54
55
56
57
58
59
60

morphology. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Authors:

Kai Wang – *Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen, 518055, China, orcid.org/0000-0003-0443-6955 Email: wangk@sustc.edu.cn*

Xi Zhu – *Shenzhen Institute of Artificial Intelligence and Robotics for Society (AIRS), The Chinese University of Hong Kong, Shenzhen (CUHK-Shenzhen), Shenzhen, Guangdong 518172, China, Email: zhuxi@cuhk.edu.cn*

Authors

Rulin Liu – *Shenzhen Institute of Artificial Intelligence and Robotics for Society (AIRS), The Chinese University of Hong Kong, Shenzhen, Guangdong 518172, China*

Junjie Hao – *Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen, 518055, China and CNRS, Univ. Bordeaux, Bordeaux INP, ICMCB, UMR 5026, Pessac, F-33608, France*

Jiagen Li – *Shenzhen Institute of Artificial Intelligence and Robotics for Society (AIRS), The Chinese University of Hong Kong, Shenzhen, Guangdong 518172, China*

1
2
3
4 Shujie Wang – *Shenzhen Institute of Artificial Intelligence and Robotics for Society*
5
6 *(AIRS), The Chinese University of Hong Kong, Shenzhen, Guangdong 518172, China*
7
8

9
10 Haocheng Liu – *Department of Electrical and Electronic Engineering, Southern*
11
12 *University of Science and Technology, Shenzhen, 518055, China*
13
14

15
16 Ziming Zhou – *Department of Electrical and Electronic Engineering, Southern*
17
18 *University of Science and Technology, Shenzhen, 518055, China*
19
20

21
22 Marie-Hélène Delville – *CNRS, Univ. Bordeaux, Bordeaux INP, ICMCB, UMR*
23
24 *5026, Pessac, F-33608, France*
25
26

27
28 Jiaji Cheng – *School of Materials Science and Engineering, Hubei University, Wuhan*
29
30 *430062, China, orcid.org/0000-0002-2663-7881. Email: jiajicheng@hubu.edu.cn*
31
32
33

34 35 36 37 **Author Contributions**

38
39
40 R. Liu and J. Hao contributed equally to this work.
41
42
43
44

45 46 **ACKNOWLEDGMENTS**

47
48 This work is supported by the Shenzhen Fundamental Research Foundation
49
50 (JCYJ20170818103918295, JCYJ20180508162801893), and the National Natural
51
52 Science Foundation of China (grant no. 21805234 and 61875082). Also supported by
53
54 funding (2019-INT018,2020-IND002) from Shenzhen Institute of Artificial
55
56 Intelligence and Robotics for Society (AIRS). TEM data were obtained using
57
58
59
60

1
2
3
4 equipment maintained by Southern University of Science and Technology Core
5
6 Research Facilities, and the authors would like to acknowledge the technical support
7
8
9 from Dongsheng He and Yang Qiu in SUSTech CRF.
10

11 12 13 14 REFERENCE

- 15
16
17 1. Meredig, B. Five High-Impact Research Areas in Machine Learning for Materials
18
19 Science. *Chem. Mater.* **2019**, *31* (23), 9579-9581.
20
21
- 22 2. Sun, W.; Zheng, Y.; Yang, K.; Zhang, Q.; Shah, A. A.; Wu, Z.; Sun, Y.; Feng, L.;
23
24 Chen, D.; Xiao, Z.; Lu, S.; Li, Y.; Sun, K. Machine learning–assisted molecular design
25
26 and efficiency prediction for high-performance organic photovoltaic materials. *Sci. Adv.*
27
28 **2019**, *5* (11), eaay4275.
29
30
- 31 3. Granda, J. M.; Donina, L.; Dragone, V.; Long, D.-L.; Cronin, L. Controlling an
32
33 organic synthesis robot with machine learning to search for new reactivity. *Nature*
34
35 **2018**, *559* (7714), 377-381.
36
37
- 38 4. Hase, F.; Roch, L. M.; Kreisbeck, C.; Aspuru-Guzik, A. Phoenix: a Bayesian
39
40 optimizer for chemistry. *ACS Cent. Sci.* **2018**, *4* (9), 1134-1145.
41
42
- 43 5. Bai, Y.; Wilbraham, L.; Slater, B. J.; Zwijnenburg, M. A.; Sprick, R. S.; Cooper,
44
45 A. I. Accelerated discovery of organic polymer photocatalysts for hydrogen evolution
46
47 from water through the integration of experiment and theory. *J. Am. Chem. Soc.* **2019**,
48
49 *141* (22), 9063-9071.
50
51
- 52 6. Butler, K. T.; Davies, D. W.; Cartwright, H.; Isayev, O.; Walsh, A. Machine
53
54 learning for molecular and materials science. *Nature* **2018**, *559* (7715), 547-555.
55
56
57
58
59
60

- 1
2
3
4 7. Kitchin, J. R. Machine learning in catalysis. *Nat. Catal.* **2018**, *1* (4), 230-232.
- 5
6
7 8. Li, F.; Thevenon, A.; Rosas-Hernández, A.; Wang, Z.; Li, Y.; Gabardo, C. M.;
8
9 Ozden, A.; Dinh, C. T.; Li, J.; Wang, Y.; Edwards, J. P.; Xu, Y.; McCallum, C.; Tao,
10
11 L.; Liang, Z.-Q.; Luo, M.; Wang, X.; Li, H.; O'Brien, C. P.; Tan, C.-S.; Nam, D.-H.;
12
13 Quintero-Bermudez, R.; Zhuang, T.-T.; Li, Y. C.; Han, Z.; Britt, R. D.; Sinton, D.;
14
15 Agapie, T.; Peters, J. C.; Sargent, E. H. Molecular tuning of CO₂-to-ethylene
16
17 conversion. *Nature* **2020**, *577* (7791), 509-513.
- 18
19
20
21
22 9. Alivisatos, A. P. Semiconductor clusters, nanocrystals, and quantum dots. *Science*
23
24
25 **1996**, *271* (5251), 933-937.
- 26
27
28 10. Zavelani-Rossi, M.; Lupo, M. G.; Krahne, R.; Manna, L.; Lanzani, G. Lasing in
29
30 self-assembled microcavities of CdSe/CdS core/shell colloidal quantum rods.
31
32 *Nanoscale* **2010**, *2* (6), 931-935.
- 33
34
35 11. Nan, W.; Niu, Y.; Qin, H.; Cui, F.; Yang, Y.; Lai, R.; Lin, W.; Peng, X. Crystal
36
37 structure control of zinc-blende CdSe/CdS core/shell nanocrystals: synthesis and
38
39 structure-dependent optical properties. *J. Am. Chem. Soc.* **2012**, *134* (48), 19685-19693.
- 40
41
42
43 12. Cassette, E.; Mahler, B.; Guigner, J.-M.; Patriarche, G.; Dubertret, B.; Pons, T.
44
45 Colloidal CdSe/CdS dot-in-plate nanocrystals with 2D-polarized emission. *ACS Nano*
46
47 **2012**, *6* (8), 6741-6750.
- 48
49
50
51 13. Sugunan, A.; Zhao, Y.; Mitra, S.; Dong, L.; Li, S.; Popov, S.; Marcinkevicius, S.;
52
53 Toprak, M. S.; Muhammed, M. Synthesis of tetrahedral quasi-type-II CdSe–CdS core–
54
55 shell quantum dots. *Nanotechnology* **2011**, *22* (42), 425202.
- 56
57
58
59
60

- 1
2
3
4 14. Song, X.; Vladislav, K.; Liu, C.; Xu, W. The preparation of anisotropic hybrid
5
6 nanostructures based on CdSe and CdS by the ligand combination method. *Nanoscale*
7
8
9 **2017**, *9* (11), 3934-3940.
- 10
11
12 15. Qin, J.; Hao, J.; Chen, W.; Deng, J.; Wang, D.; Xu, B.; Wu, D.; Wang, K.; Li, S.;
13
14 Wen, Z. In *34-2: A Rapid, Highly Emissive Procedure Synthesize of Giant Pure Red*
15
16 *Coreshell Quantum Rods by Using Modified Tributylphosphine-assisted Method*, SID
17
18 Symposium Digest of Technical Papers, Wiley Online Library: 2016; pp 428-431.
- 19
20
21
22 16. Puzder, A.; Williamson, A. J.; Zaitseva, N.; Galli, G.; Manna, L.; Alivisatos, A. P.
23
24 The effect of organic ligand binding on the growth of CdSe nanoparticles probed by ab
25
26 initio calculations. *Nano Lett.* **2004**, *4* (12), 2361-2365.
- 27
28
29
30 17. Manna, L.; Wang, L. W.; Cingolani, R.; Alivisatos, A. P. First-principles modeling
31
32 of unpassivated and surfactant-passivated bulk facets of wurtzite CdSe: A model
33
34 system for studying the anisotropic growth of CdSe nanocrystals. *J. Phys. Chem. B*
35
36 **2005**, *109* (13), 6183-6192.
- 37
38
39
40 18. Peng, Z. A.; Peng, X. Mechanisms of the shape evolution of CdSe nanocrystals. *J.*
41
42 *Am. Chem. Soc.* **2001**, *123* (7), 1389-1395.
- 43
44
45
46 19. Yin, Y.; Erdonmez, C.; Aloni, S.; Alivisatos, A. P. Faceting of nanocrystals during
47
48 chemical transformation: From solid silver spheres to hollow gold octahedra. *J. Am.*
49
50 *Chem. Soc.* **2006**, *128* (39), 12671-12673.
- 51
52
53
54 20. Lejaeghere, K.; Van Speybroeck, V.; Cottenier, S.; Van Oost, G. The ground state
55
56 elemental crystals as a benchmark set for solid state DFT: intrinsic accuracy and code
57
58 comparison. *Crit. Rev. Solid State Mater. Sci.* **2012**, *39*, 1.
- 59
60

- 1
2
3
4 21. Brus, L. Electronic wave functions in semiconductor clusters: experiment and
5
6 theory. *J. Phys. Chem.* **1986**, *90* (12), 2555-2560.
7
8
9 22. Markovich, G.; Collier, C. P.; Henrichs, S. E.; Remacle, F.; Levine, R. D.; Heath,
10
11 J. R. Architectonic quantum dot solids. *Acc. Chem. Res.* **1999**, *32* (5), 415-423.
12
13
14 23. Li, J.; Lu, Y.; Xu, Y.; Liu, C.; Tu, Y.; Ye, S.; Liu, H.; Xie, Y.; Qian, H.; Zhu, X.
15
16 AIR-Chem: Authentic Intelligent Robotics for Chemistry. *J. Phys. Chem. A* **2018**, *122*
17
18 (46), 9142-9148.
19
20
21 24. Wang, X.; Chen, S.; Thota, S.; Wang, Y.; Tan, H.; Tang, M.; Quan, Z.; Zhao, J.
22
23 Anisotropic Arm Growth in Unconventional Semiconductor CdSe/CdS Nanotetrapod
24
25 Synthesis Using Core/Shell CdSe/CdS as Seeds. *J. Phys. Chem. C* **2019**, *123* (31),
26
27 19238-19245.
28
29
30 25. Yamamoto, M.; Chan, C. T.; Ho, K.; Kurahashi, M.; Naito, S. First-principles
31
32 calculation of the longitudinal phonon in the surface-normal direction of a zirconium
33
34 (0001) slab: Localization mode at the subsurface. *Phys. Rev. B* **1996**, *53* (20), 13772.
35
36
37 26. Rempel, J. Y.; Trout, B. L.; Bawendi, M. G.; Jensen, K. F. Density functional
38
39 theory study of ligand binding on CdSe (0001),(0001), and (1120) single crystal relaxed
40
41 and reconstructed surfaces: implications for nanocrystalline growth. *J. Phys. Chem. B*
42
43 **2006**, *110* (36), 18007-18016.
44
45
46 27. Aruguete, D. M.; Marcus, M. A.; Li, L.-s.; Williamson, A.; Fakra, S.; Gygi, F.;
47
48 Galli, G. A.; Alivisatos, A. P. Surface structure of CdSe nanorods revealed by combined
49
50 X-ray absorption fine structure measurements and ab initio calculations. *J. Phys. Chem.*
51
52 *C* **2007**, *111* (1), 75-79.
53
54
55
56
57
58
59
60

- 1
2
3
4 28. Kilin, D. S.; Prezhdo, O. V.; Xia, Y. Shape-controlled synthesis of silver
5
6 nanoparticles: Ab initio study of preferential surface coordination with citric acid.
7
8 *Chem. Phys. Lett.* **2008**, *458* (1-3), 113-116.
9
10
11 29. Bealing, C. R.; Baumgardner, W. J.; Choi, J. J.; Hanrath, T.; Hennig, R. G.
12
13 Predicting nanocrystal shape through consideration of surface-ligand interactions. *ACS*
14
15 *Nano* **2012**, *6* (3), 2118-2127.
16
17
18 30. Hu, H.; Reven, L.; Rey, A. First-principles density functional theory (DFT) study
19
20 of gold nanorod and its interaction with alkanethiol ligands. *J. Phys. Chem. B* **2013**,
21
22 *117* (41), 12625-12631.
23
24
25 31. Verma, T.; Pearl, J. *Equivalence and synthesis of causal models*. UCLA, Computer
26
27 Science Department: 1991.
28
29
30 32. Spirtes, P.; Glymour, C. N.; Scheines, R.; Heckerman, D. *Causation, prediction,*
31
32 *and search*. MIT press: 2000.
33
34
35 33. Pearl, J. Causal inference in statistics: An overview. *Stat. Surv.* **2009**, *3*, 96-146.
36
37
38 34. Mansournia, M. A.; Altman, D. G. STATISTICS NOTES Inverse probability
39
40 weighting. *BMJ-BRIT MED J* **2016**, *352*.
41
42
43 35. Carbone, L.; Nobile, C.; De Giorgi, M.; Sala, F. D.; Morello, G.; Pompa, P.; Hytch,
44
45 M.; Snoeck, E.; Fiore, A.; Franchini, I. R. Synthesis and micrometer-scale assembly of
46
47 colloidal CdSe/CdS nanorods prepared by a seeded growth approach. *Nano letters*
48
49 **2007**, *7* (10), 2942-2950.
50
51
52
53 36. Fiore, A.; Mastria, R.; Lupo, M. G.; Lanzani, G.; Giannini, C.; Carlino, E.; Morello,
54
55 G.; De Giorgi, M.; Li, Y.; Cingolani, R. Tetrapod-shaped colloidal nanocrystals of II-
56
57
58
59
60

- 1
2
3
4 VI semiconductors prepared by seeded growth. *J. Am. Chem. Soc.* **2009**, *131* (6), 2274-
5
6 2282.
7
8
9 37. Gafni, T.; Cohen, K. Learning in restless multi-armed bandits via adaptive arm
10
11 sequencing rules. *arXiv preprint arXiv:1906.08120* **2019**.
12
13
14 38. Kunneman, L. T.; Zanella, M.; Manna, L.; Siebbeles, L. D.; Schins, J. M. Mobility
15
16 and spatial distribution of photoexcited electrons in cdse/cds nanorods. *The Journal of*
17
18 *Physical Chemistry C* **2013**, *117* (6), 3146-3151.
19
20
21
22 39. Manna, L.; Scher, E. C.; Li, L.-S.; Alivisatos, A. P. Epitaxial growth and
23
24 photochemical annealing of graded CdS/ZnS shells on colloidal CdSe nanorods. *J. Am.*
25
26 *Chem. Soc.* **2002**, *124* (24), 7136-7145.
27
28
29
30 40. Peng, X. Mechanisms for the shape - control and shape - evolution of colloidal
31
32 semiconductor nanocrystals. *Adv. Mater.* **2003**, *15* (5), 459-463.
33
34
35 41. Peng, Z. A.; Peng, X. Nearly monodisperse and shape-controlled CdSe
36
37 nanocrystals via alternative routes: nucleation and growth. *J. Am. Chem. Soc.* **2002**, *124*
38
39 (13), 3343-3353.
40
41
42
43 42. Talapin, D. V.; Nelson, J. H.; Shevchenko, E. V.; Aloni, S.; Sadtler, B.; Alivisatos,
44
45 A. P. Seeded growth of highly luminescent CdSe/CdS nanoheterostructures with rod
46
47 and tetrapod morphologies. *Nano Lett.* **2007**, *7* (10), 2951-2959.
48
49
50
51 43. Wang, W.; Banerjee, S.; Jia, S.; Steigerwald, M. L.; Herman, I. P. Ligand control
52
53 of growth, morphology, and capping structure of colloidal CdSe nanorods. *Chem.*
54
55 *Mater.* **2007**, *19* (10), 2573-2580.
56
57
58
59
60

- 1
2
3
4 44. Wang, X.; Feng, Z.; Fan, D.; Fan, F.; Li, C. Shape-controlled synthesis of CdS
5 nanostructures via a solvothermal method. *Crystal growth & design* **2010**, *10* (12),
6 5312-5318.
7
8
9
10
11 45. Shanavas, K.; Sharma, S. M.; Dasgupta, I.; Nag, A.; Hazarika, A.; Sarma, D. First-
12 principles study of the effect of organic ligands on the crystal structure of CdS
13 nanoparticles. *The Journal of Physical Chemistry C* **2012**, *116* (11), 6507-6511.
14
15
16
17
18 46. Lifshitz, I. M.; Slyozov, V. V. The kinetics of precipitation from supersaturated
19 solid solutions. *J. Phys. Chem. Solids* **1961**, *19* (1-2), 35-50.
20
21
22
23
24 47. Sugimoto, T. Preparation of monodispersed colloidal particles. *Adv. Colloid*
25 *Interface Sci.* **1987**, *28*, 65-108.
26
27
28
29
30 48. Wagner, C. Theorie der alterung von niederschlägen durch umlösen (Ostwald-
31 reifung). *Z. Elektrochem.* **1961**, *65* (7-8), 581-591.
32
33
34
35 49. Molleman, B.; Hiemstra, T. Size and shape dependency of the surface energy of
36 metallic nanoparticles: unifying the atomic and thermodynamic approaches. *PCCP*
37 **2018**, *20* (31), 20575-20587.
38
39
40
41
42 50. Chen, W.; Chen, K.; Peng, Q.; Li, Y. Triangular CdS nanocrystals: rational
43 solvothermal synthesis and optical studies. *Small* **2009**, *5* (6), 681-684.
44
45
46
47
48 51. Poulouse, A. C.; Veerananarayanan, S.; Yoshida, Y.; Maekawa, T.; Kumar, D. S.
49 Rapid synthesis of triangular CdS nanocrystals without any trap emission. *J. Nanopart.*
50 *Res.* **2012**, *14* (4), 789.
51
52
53
54
55 52. Warner, J. H.; Tilley, R. D. Synthesis and Self-Assembly of Triangular and
56 Hexagonal CdS Nanocrystals. *Adv. Mater.* **2005**, *17* (24), 2997-3001.
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60