Supporting Information

Anisotropic growth of ZnO nanoparticles deciphered through 2D size plots and multivariate analysis

Zhihua Zhao,[a, b] Yinping Wang,[a, b] Céline Delmas,[c] Christophe Mingotaud, [b] Jean-Daniel Marty[b] * and Myrtil L. Kahn [a]*

a. Laboratory of Coordination Chemistry, CNRS UPR 8241, University of Toulouse, 205 route de Narbonne, 31077 Toulouse, France. E-mail: myrtil.kahn@lcc-toulouse.fr

b. Laboratoire des IMRCP, Université de Toulouse, CNRS UMR 5623, Université Paul Sabatier, 118, route de Narbonne 31062 Toulouse Cedex 9, France. E-mail: marty@chimie.ups-tlse.fr

c. MIAT, Université de Toulouse, INRA, 31326 Castanet-Tolosan (France).

A. 2D plot analysis and multivariate analysis.

B. Effect of the time of hydrolysis

C. Effect of the waiting time after mixing

D. Influence of the speeding rate of addition and amount of water

A. 2D plot analysis and multivariate analysis.

2D plot analysis. Each particle is characterized by two sizes, noted D1 and D2, measured on perpendicular axes. These two sizes generally correspond to the length and width of the nano-objects. For each particle and on a same graph, we plot D1 as a function of D2 and also D2 as a function of D1. The aspect ratio of the particle (noted AR) is defined as the ratio between the two lengths AR=D1/D2 (with D1>D2). It is related to a theta angle through the equation AR=tan θ. Higher anisotropy in the shape of a particle corresponds to higher AR value and to a theta angle closer to 90° in the proposed 2D plot. 2D plots present point clouds whose structures can be the result of different sub-populations. To identify these sub-populations, a multivariate analysis was performed with the MIXMOD software (http://mixmod.org) using R package. Apart from the Gaussian character of the probability densities, no assumption about the orientation, shape, and volume of the different sub-populations was made during calculation. The number of sub-populations that composed the point clouds was fixed by the user or chosen numerically thanks to the BIC criteria (Bayesian Information Criterion). Each sub-population was then characterized by the mean of the two studied variables (i.e., short and long axis lengths), as well as the corresponding standard deviations. Additionally, the correlation parameter ρ between both variables was calculated. The correlation is equal to zero when the two variables are totally independent and equal to 1 when they are affinely related to each other.

MIXMOD software

The main purpose of the MIXMOD software (http ://www.mixmod.org) is to discover group structures in multivariate data sets. It is an exploratory data analysis tool for solving clustering and classification problems. Mathematically speaking, for quantitative multivariate data, the MIXMOD software modelises the probability density of the data $x_i = (x_i^1, \cdots, x_i^d)^T$, $i = 1, \cdots, N$ by a mixture f of K multivariate Gaussian densities $h_1, \cdots, h_K$:

$$f(x_i, K, (\mu_k, p_k, \Sigma_k)_{k=1,\cdots,K}) = \sum_{k=1}^{K} p_k h_k(x_i, \mu_k, \Sigma_k) \qquad (1)$$

where :

$$h_k(x_i, \mu_k, \Sigma_k) = \frac{1}{\sqrt{2\pi}^d \sqrt{|\Sigma_k|}} \exp\left(-\frac{1}{2}(x_i - \mu_k)^T \Sigma_k^{-1} (x_i - \mu_k)\right)$$

h_k is caracterized by a mean vector μ_k and a variance-covariance matrix Σ_k. $|\Sigma_k|$ denotes the determinant of Σ_k. The estimation of the parameters μ_k, p_k, Σ_k for $k = 1, \cdots, K$ is done by an EM (Expectation-Maximization) algorithm. The aim of this algorithm is to find the "best", the "most likely" estimators of the parameters μ_k, p_k, Σ_k for $k = 1, \cdots, K$ that is to say the estimators that maximize the likelihood :

$$\prod_{i=1}^{N} f(x_i, K, (\mu_k, p_k, \Sigma_k)_{k=1,\cdots,K})$$

This algorithm consists in calculating iteratively until convergence :

$$\alpha_{ik}^{(n)} = \frac{p_k^{(n)} h_k(x_i, \mu_k^{(n)}, \Sigma_k^{(n)})}{\sum_{k=1}^{K} p_k^{(n)} h_k(x_i, \mu_k^{(n)}, \Sigma_k^{(n)})}$$

$$\mu_k^{(n+1)} = \left(\sum_{i=1}^{N} x_i \alpha_{ik}^{(n)}\right) / \left(\sum_{i=1}^{N} \alpha_{ik}^{(n)}\right)$$

$$\Sigma_k^{(n+1)} = \left(\sum_{i=1}^{N} \alpha_{ik}^{(n)} (x_i - \mu_k^{(n)})^T (x_i - \mu_k^{(n)})\right) / \left(\sum_{i=1}^{N} \alpha_{ik}^{(n)}\right)$$

$$p_k^{n+1} = \left(\sum_{i=1}^{N} \alpha_{ik}^{(n)}\right) / N$$

When the estimators $\hat{\mu}_k$, $\hat{p}_k$, $\hat{\Sigma}_k$ for $k = 1, \cdots, K$ are found ; each observation x_i is affected to its "most likely" component $\hat{k}_i$:

$$\hat{k}_i = \mathrm{argmax}_k \frac{\hat{p}_k h_k(x_i, \hat{\mu}_k, \hat{\Sigma}_k)}{\sum_{k=1}^{K} \hat{p}_k h_k(x_i, \hat{\mu}_k, \hat{\Sigma}_k)}$$

Thus a classification in K classes is obtained.
For each K we can calculate $\hat{\mu}_k$, $\hat{p}_k$, $\hat{\Sigma}_k$ for $k = 1, \cdots, K$. To estimate K we choose a model (1) that fits well the data but without too many parameters. For that we choose K that minimizes the BIC (Bayesian Information Criterion) criterion :

$$\hat{K} = \text{argmin}_K \text{BIC}(K) = \text{argmin}_K (-2\ln \prod_{i=1}^{N} f(x_i, K, (\hat{\mu}_k, \hat{p}_k, \hat{\Sigma}_k)_{k=1,\cdots,K}) + \nu_K \ln(N))$$

where ν_K is the number of free parameters in the mixture model with K components.

Statistical indices

The mean vector of a data set $x_i = (x_i^1, \cdots, x_i^d)^T$, $i = 1, \cdots, N$ is estimated by

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i.$$

The variance is estimated by

$$\hat{\text{Var}}(x) = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^2$$

The standard deviation is defined as the square root of the variance. The covariance between two real variables x^1 and x^2 is estimated by

$$\hat{\text{Cov}}(x^1, x^2) = \frac{1}{N-1} \sum_{i=1}^{N} (x_i^1 - \bar{x^1})(x_i^2 - \bar{x^2}).$$

The correlation between x^1 and x^2 is estimated by

$$\hat{\rho}(x^1, x^2) = \frac{\hat{\text{Cov}}(x^1, x^2)}{\sqrt{\hat{\text{Var}}(x^1)}\sqrt{\hat{\text{Var}}(x^2)}}$$

The correlation is equal to zero between two independent variables and equal to 1 between two identical variables.

Modified mixmod software

In the modified version of the MIXMOD software we assume that the first component of the mixture (1) is known. Thus the probability density of the data $x_i = (x_i^1, \cdots, x_i^d)^T$, $i = 1, \cdots, N$ is a mixture f of K multivariate Gaussian densities $h_1, \cdots, h_K$:

$$f(x_i, K, (\mu_k, p_k, \Sigma_k)_{k=1,\cdots,K}) = \sum_{k=1}^{K} p_k h_k(x_i, \mu_k, \Sigma_k)$$

where only μ_1 and Σ_1 are known. The other parameters are unknown. The points belonging to the first component are chosen in a 5% confidence interval around μ_1. Then the standard MIXMOD software is used to classify the other points and determine the number of classes.

B. Effect of the time of hydrolysis

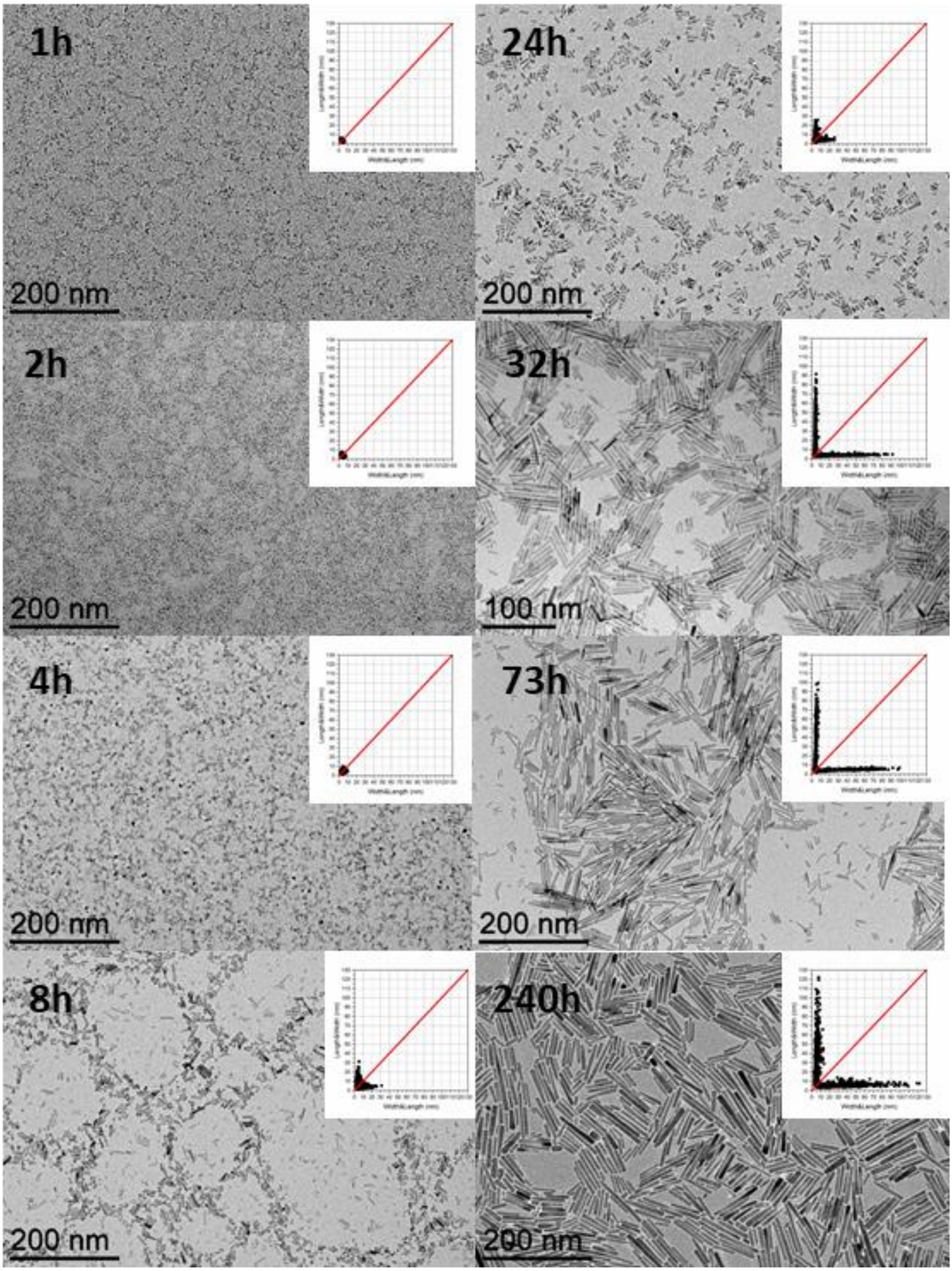


Figure S1. 2D size plots and TEM pictures of ZnO NPs versus hydrolysis time (from 1 to 240 hours). During the first hours, these pristine nanoparticles are maturing and their mean size increases.

Table S1.

Multivariate analysis of the 2D plots corresponding to the TEM images of the Figure S1 (ZnO NPs versus hydrolysis time from 1 to 240 hours) through Rmixmod program (the dispersion is given as twice the standard deviation obtained from calculations, and number of cluster was set to 1).

Hydrolysis time (hour)	Length (nm)	Width (nm)	Correlation
1	3.2 ± 1.6	2.1 ± 1.1	0.60
2	3.2 ± 1.8	2.0 ± 1.0	0.50
4	3.5 ± 2.5	2.5 ± 1.7	0.79
8	6.4 ± 7.6	3.0 ± 2.4	0.58
24	11.0 ± 8.2	4.1 ± 1.7	0.20
32	21.9 ± 36.7	3.7 ± 1.5	0.45
73	33.1 ± 46.2	4.5 ± 2.0	0.55
240	36.9 ± 41.7	5.4 ± 3.5	0.35

Table S2: Multivariate analysis of the 2D plots corresponding to the TEM images of Figure S1 (ZnO NPs versus hydrolysis time from 1 to 240 hours) through a modified Rmixmod program allowing access to the number and consequently the percent of pristine NPs, N1, and anisotropically growing NPs, N2.

Hydrolysis time (hours)	Number of N1	Number of N2	Percent of N1 (%)	Percent of N2 (%)
1	2102	46	97.9	2.1
2	4509	173	96.3	3.7
4	1184	131	90	10
8	604	604	50	50
24	65	975	6.3	93.7
32	77	682	10.1	89.9
73	19	467	3.9	96.1
240	1	596	0.2	99.8

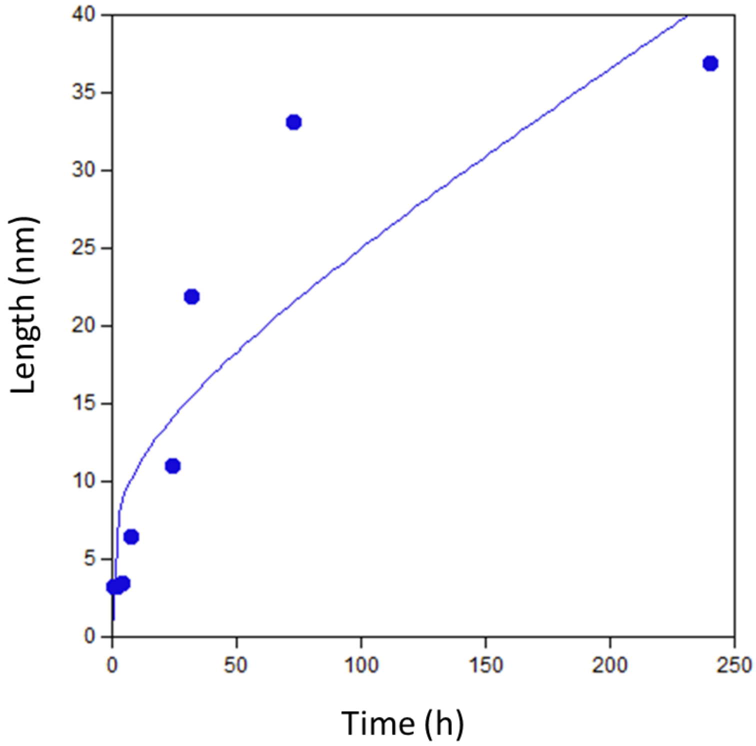


Figure S2. Fitting of the evolution of ZnO NPs average length versus hydrolysis time (from 1 to 240 hours) (see main text).

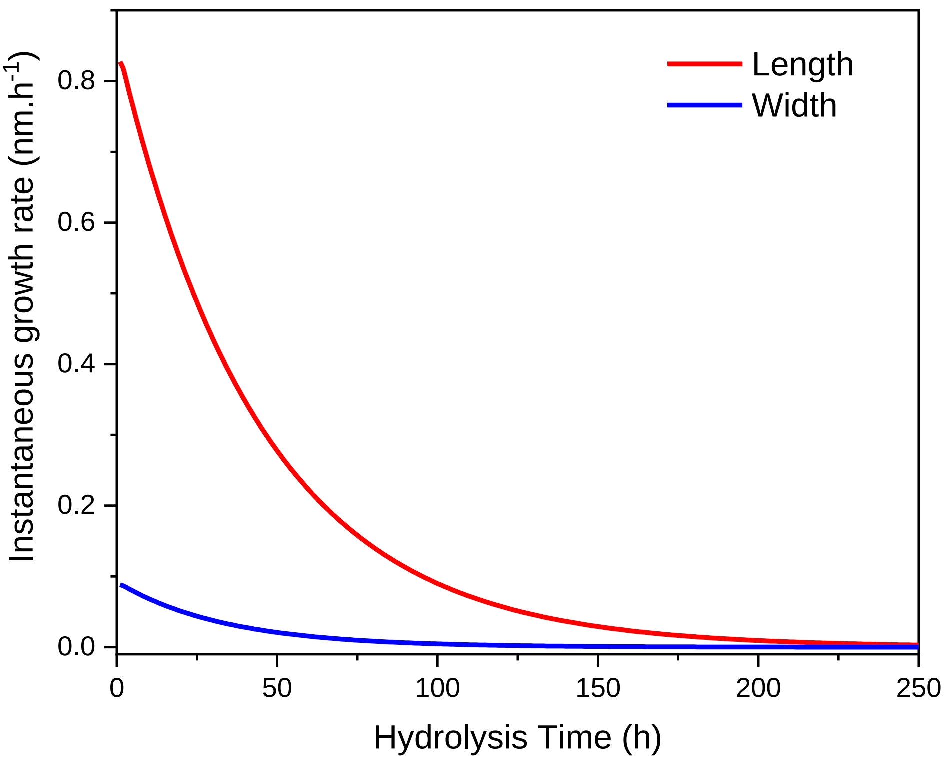


Figure S3. Instantaneous growth rate (nm.h^{-1}) as a function of time of ZnO NPs (from 1 to 240 hours, first derivative of Figure 1 –top right–).

C. Effect of the waiting time after mixing

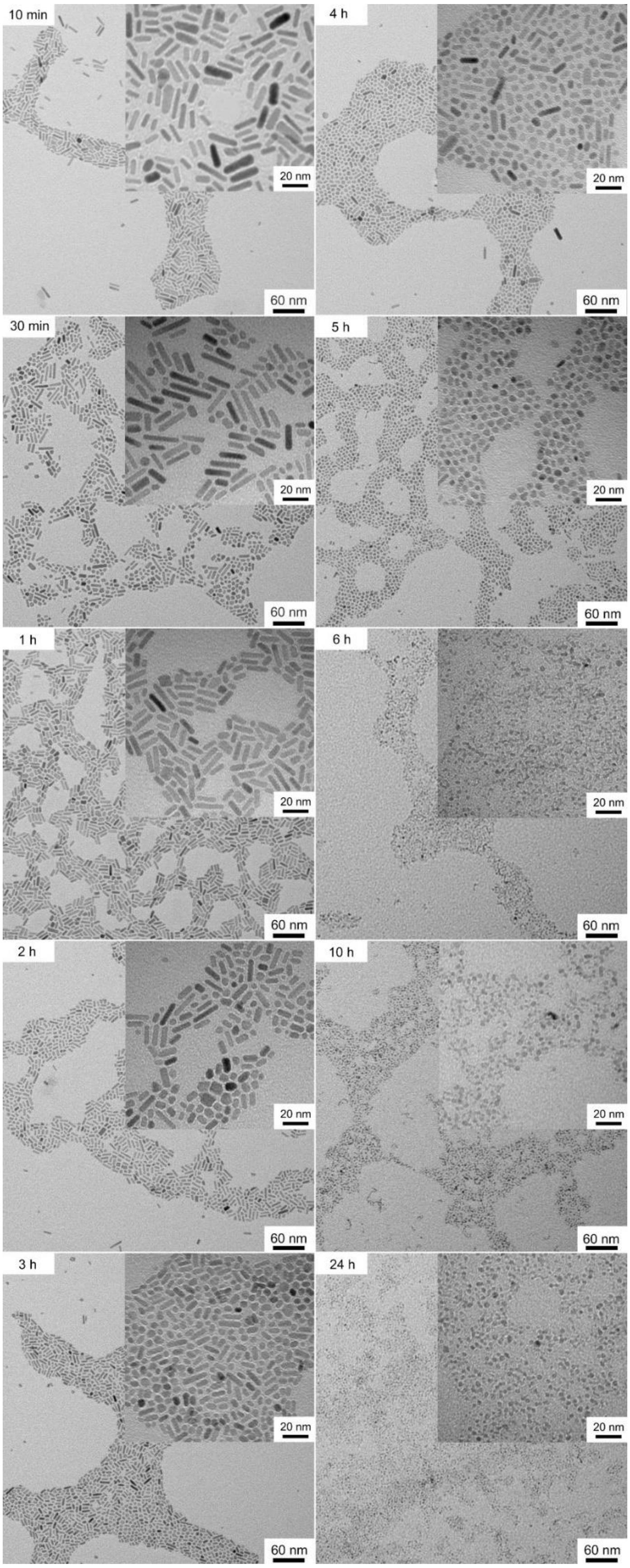


Figure S4. TEM images of the ZnO/2DDA NPs *versus* waiting time after mixing (or aging time before hydrolysis).

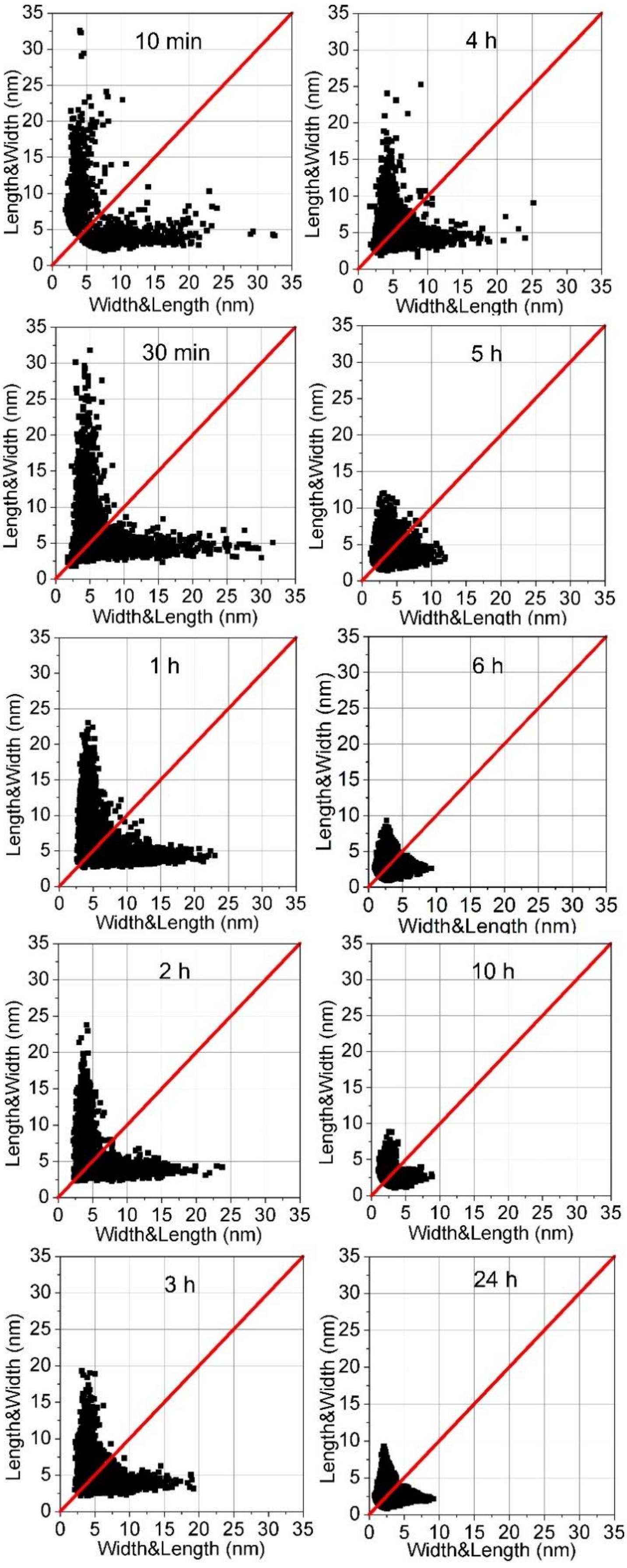


Figure S5. 2D size plots corresponding to the TEM images of Figure S4 (ZnO/2DDA NPs *versus* waiting time after mixing).

Table S3. Multivariate analysis of the 2D plots corresponding to the TEM images of the Figure S4 (ZnO/2DDA NPs *versus* waiting time after mixing) through Rmixmod program (the dispersion is given as twice the standard deviation obtained from calculations, and number of cluster was set to 1).

Aging time	Length (nm)	Width (nm)	correlation
10 minutes	9.3 ± 7.5	3.7 ± 1.8	0.40
30 minutes	8.6 ± 9.0	4.4 ± 1.9	0.13
1 hour	10.9 ± 7.3	4.4 ± 1.8	-0.02
2 hours	8.3 ± 6.7	3.8 ± 1.4	0.10
3 hours	7.7 ± 5.9	4.1 ± 1.6	0.02
4 hours	6.6 ± 5.4	4.2 ± 1.7	0.32
5 hours	5.1 ± 2.8	3.6 ± 1.7	0.41
6 hours	3.4 ± 2.6	2.1 ± 1.3	0.61
10 hours	3.4 ± 2.1	2.0 ± 1.0	0.40
24 hours	3.4 ± 2.7	2.1 ± 1.2	0.54

Table S4: Multivariate analysis of the 2D plots corresponding to the TEM images of Figure S4 (ZnO/2DDA NPs *versus* waiting time after mixing) through a modified Rmixmod program.

Aging time	Number of N1	Number of N2	Percent of N1 (%)	Percent of N2 (%)
10 min	39	975	3.8	96.2
30 min	71	1499	4.5	95.5
1 hour	14	1056	1.3	98.7
2 hours	76	1166	6.1	93.9
3 hours	50	904	5.2	94.8
4 hours	105	1375	7.1	92.9
5 hours	664	1956	25.3	74.7
6 hours	1691	244	87.4	12.6
10 hours	978	94	91.2	8.8
24 hours	1309	203	86.6	13.4

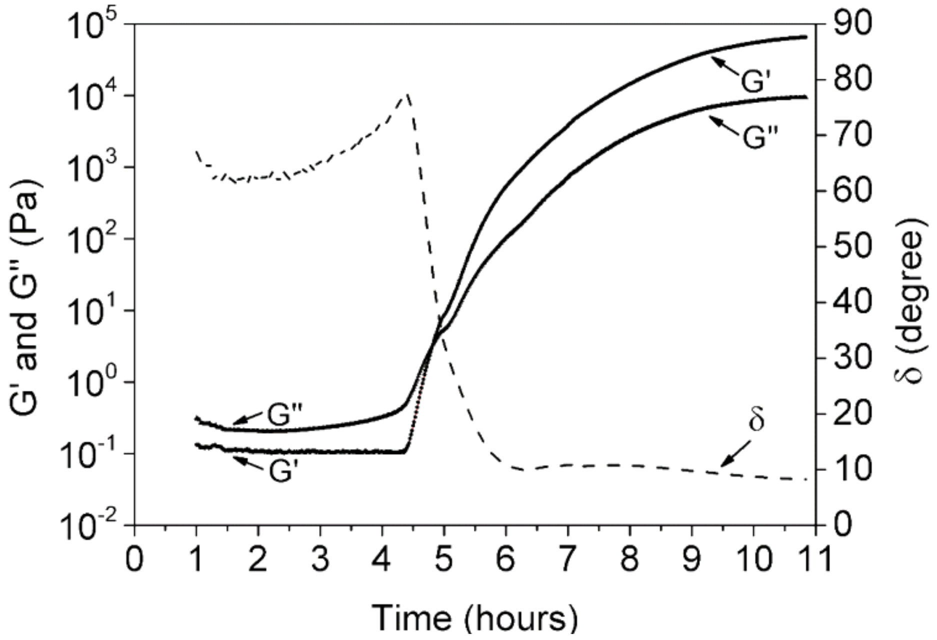
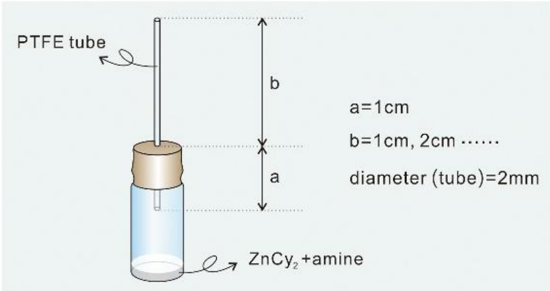


Figure S6. Time dependence of the elastic modulus G’, the viscous modulus G'' and the phase angle, δ, between the stress and strain measured under sinusoidal stress (frequency 1 Hz) for the [$ZnCy_2$]/2DDA mixture under nitrogen. The origin of the time scale corresponds to the mixing process. (extracted from ref 1)

D. Influence of the speeding rate of addition and amount of water.

Water feeding rate was controlled by using PTFE tubes of different lengths as depicted schematically in Scheme S1.



Scheme S1: Schematic illustration for the preparation of ZnO NPs depending on different length of tubes for hydrolysis (vial dimension: height: 3 cm, diameter: 1.5 cm). Water is coming from the top of the tube

Indeed, following the Fick's law, the gaseous water flow, J, from the atmosphere to the vial containing the reactants should be:

$$J = - D\, dC/dx,$$

where D is the diffusion coefficient of gaseous water in air (ca 0.25 cm^2/s),[2] C the concentration of gaseous water along the position x in the tube. The outside concentration of gaseous water, C_{out}, can be roughly estimated by the ideal gas law:

$$C_{out} = RH_{mean}.\ P_s/RT$$

where RH_{mean} is the mean value of the relative humidity of the atmosphere (ca 70%)[3] and P_s the saturated pressure of water at the average temperature T (20°C), leading to C_{out} close to 7.10^{-4} $mol.L^{-1}$. The inner concentration C_{in} should be close to zero because of the (fast) hydrolysis reaction. Therefore, the flow could be estimated to:

$$J = D.C_{out} / L,$$

where L is the total length of the PTEFE tube (i.e. a +b in Figure 2A) and the rate of water addition is J.S where S in the section of the PTFE tube, i.e. $2.2.10^{-8}$ $mol.s^{-1}$ /L(in cm) for a tube diameter of $\Phi = 2$ mm. Increasing the tube length from 2 cm to 20 cm should therefore divide the water flow rate by a factor of 10 to ca 10^{-9} $mol.s^{-1}$. With such a rudimentary theoretical rate, 70 hours (*i.e.* roughly 3 days) are needed to hydrolyze all the zinc precursor.

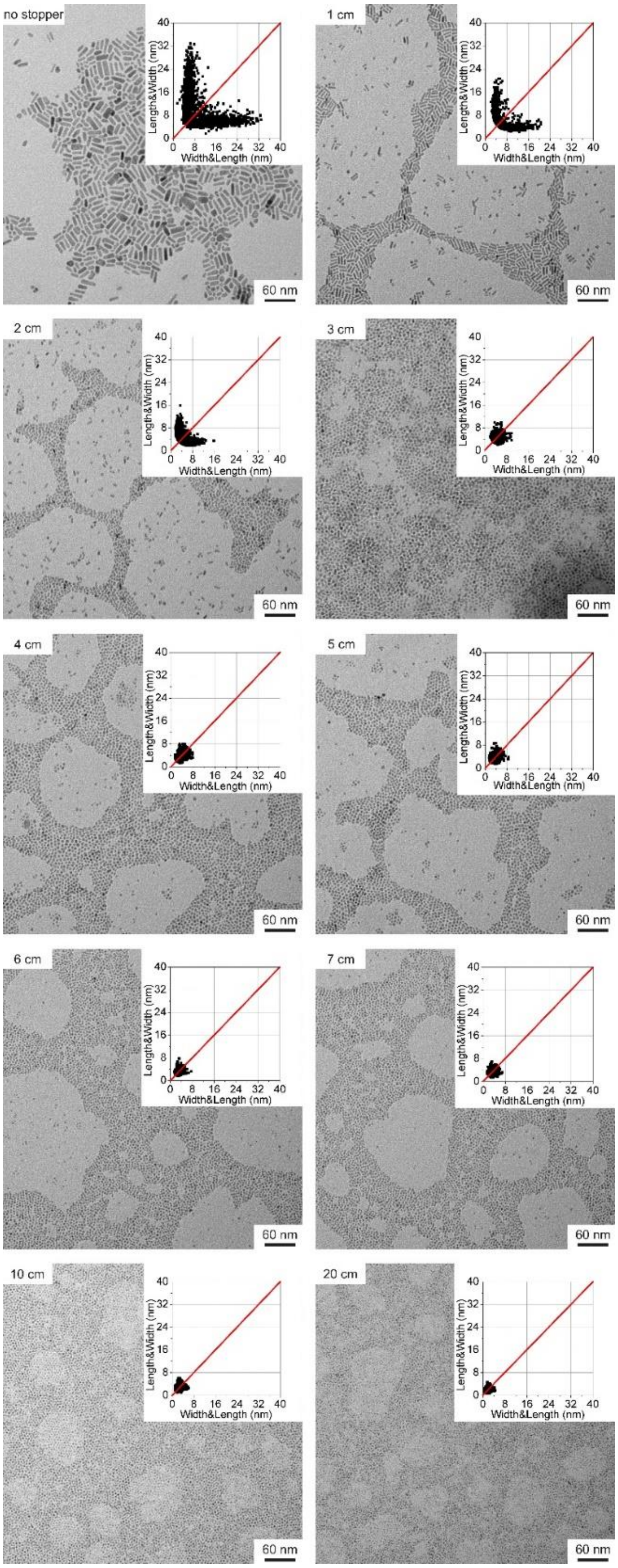


Figure S7. TEM images of synthesized ZnO NPs for various PTFE external tube length (L in cm).

Table S5. Multivariate analysis of the 2D plots corresponding to the TEM images of the Figure S7 (ZnO NPs for various PTFE external tube length) through Rmixmod program (the dispersion is given as twice the standard deviation obtained from calculations, and number of cluster was set to 1).

Outer length of the tube (cm)	Length (nm)	Width (nm)	correlation
no stopper	16.5 ± 11.6	6.1 ± 2.8	0.13
1	10.5 ± 6.9	4.1 ± 1.8	-0.09
2	5.8 ± 3.5	3.1 ± 1.3	-0.05
3	4.9 ± 1.9	3.7 ± 1.4	0.58
4	4.6 ± 1.9	3.5 ± 1.5	0.69
5	4.3 ± 1.8	3.4 ± 1.4	0.70
6	3.5 ± 1.4	2.8 ± 1.1	0.78
7	3.4 ± 1.4	2.7 ± 1.1	0.73
10	3.0 ± 1.6	2.2 ± 1.1	0.74
20	1.7 ± 1.1	1.1 ± 0.8	0.75

Table S6: Multivariate analysis of the 2D plots corresponding to the TEM images of Figure S7 (ZnO NPs for various PTFE external tube length) through a modified Rmixmod program allowing access to the number and consequently the percent of pristine NPs, N1, and anisotropically growing NPs, N2.

Outer length of the tube (cm)	Number of N1	Number of N2	Percent of N1 (%)	Percent of N2 (%)
no stopper	0	774	0.0	100.0
1	2	405	0.5	99.5
2	535	626	46.1	53.9
3	539	347	60.8	39.2
4	608	279	68.5	31.5
5	653	151	81.2	18.8
6	1043	23	97.8	2.2
7	1367	36	97.4	2.6
10	1055	8	99.2	0.8
20	1352	0	100.0	0.0

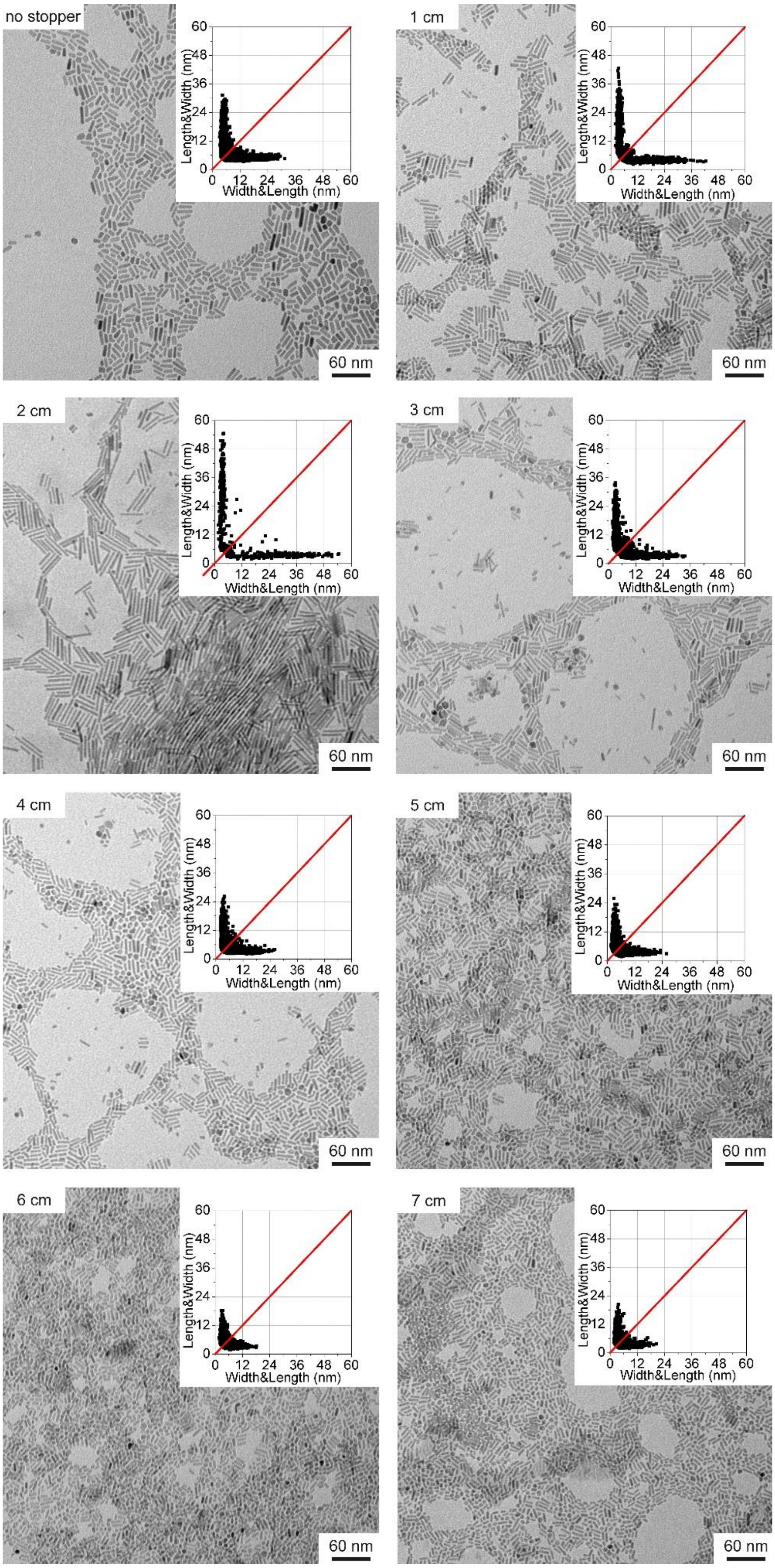


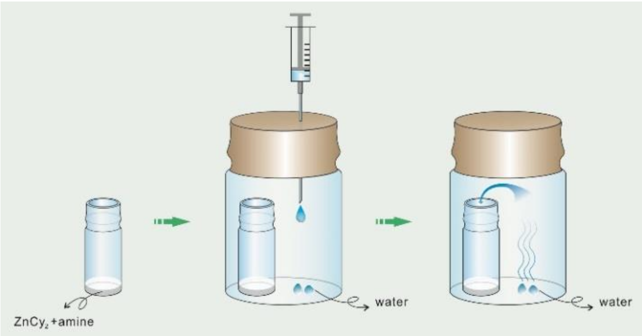
Figure S8. ZnO NPs prepared as the process shown in Figure S7, but using 4 mm PTFE tube.

Table S7. Multivariate analysis of the 2D plots corresponding to the TEM images of the Figure S8 (ZnO NPs for various 4 mm PTFE external tube length) through Rmixmod program (the dispersion is given as twice the standard deviation obtained from calculations, and number of cluster was set to 1).

Outer length of the tube (cm)	Length (nm)	Width (nm)	Correlation
no stopper	13.9 ± 10.0	5.1 ± 1.8	-0.11
1 cm	15.8 ± 15.4	4.2 ± 2.0	-0.34
2 cm	25.9 ± 24.0	3.5 ± 2.1	-0.07
3 cm	15.0 ± 12.0	3.6 ± 2.3	-0.25
4 cm	11.2 ± 8.7	3.7 ± 2.1	-0.19
5 cm	9.3 ± 7.4	3.6 ± 1.9	-0.09
6 cm	7.4 ± 5.3	3.7 ± 1.6	-0.18
7 cm	7.6 ± 5.7	3.4 ± 1.7	-0.09

Table S8: Multivariate analysis of the 2D plots corresponding to the TEM images of Figure S8 (ZnO NPs for various 4 mm PTFE external tube length) through a modified Rmixmod program allowing access to the number and consequently the percent of pristine NPs, N1, and anisotropicaly growing NPs, N2.

Outer length of the tube (cm)	Number of N1	Number of N2	Percent of N1 (%)	Percent of N2 (%)
no stopper	0	1014	0.0	100.0
1	0	705	0.0	100.0
2	0	288	0.0	100.0
3	3	585	0.5	99.5
4	39	612	6.0	94.0
5	93	968	8.8	91.2
6	202	888	18.5	81.5
7	290	1253	18.8	81.2



Scheme S2. Schematic illustration for the preparation of ZnO NPs depending on different equivalent of water.

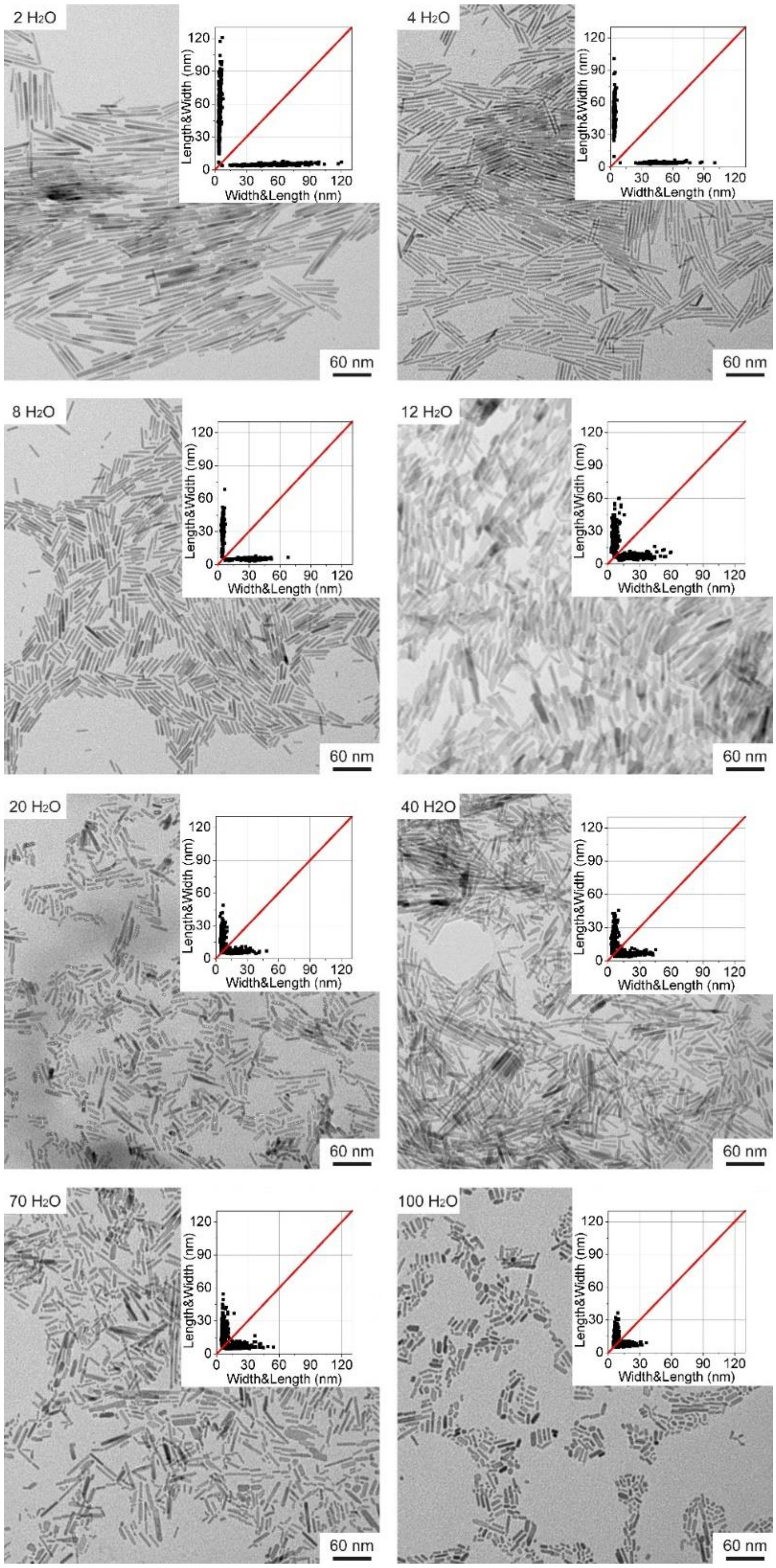


Figure S9. TEM picture and 2D size plots of ZnO NPs prepared in the presence of 2 eq. DDA for different amounts of water.

Table S9. Multivariate analysis of the 2D plots corresponding to the TEM images of the Figure S9 (ZnO NPs prepared in the presence of 2 eq. DDA for different amounts of water) through Rmixmod program (the dispersion is given as twice the standard deviation obtained from calculations, and number of cluster was set to 1).

Equivalent of H_2O	Length (nm)	Width (nm)	Correlation
2	53.8 ± 41.2	4.8 ± 1.5	0.51
4	48.1 ± 22.3	4.0 ± 1.1	0.21
8	31.9 ± 16.7	4.8 ± 1.2	0.21
12	24.4 ± 19.4	6.4 ± 3.5	0.35
20	19.1 ± 13.5	6.4 ± 2.2	0.15
40	19.3 ± 15.7	5.9 ± 2.6	0.36
70	19.5 ± 16.5	6.6 ± 3.3	0.18
100	14.9 ± 9.9	7.3 ± 2.3	0.31

Table S10: Multivariate analysis of the 2D plots corresponding to the TEM images of Figure S9 (ZnO NPs prepared in the presence of 2 eq. DDA for different amounts of water) through a modified Rmixmod program allowing access to the number and consequently the percent of pristine NPs, N1, and anisotropicaly growing NPs, N2.

Equivalent of H_2O	Number of N1	Number of N2	Percent of N1 (%)	Percent of N2 (%)
2	0	300	0	100
4	0	301	0	100
8	2	298	0.6	99.4
12	2	302	0.6	99.4
20	14	286	4.7	95.3
40	19	281	6.3	93.7
70	17	284	5.6	94.4
100	32	268	10.6	89.4

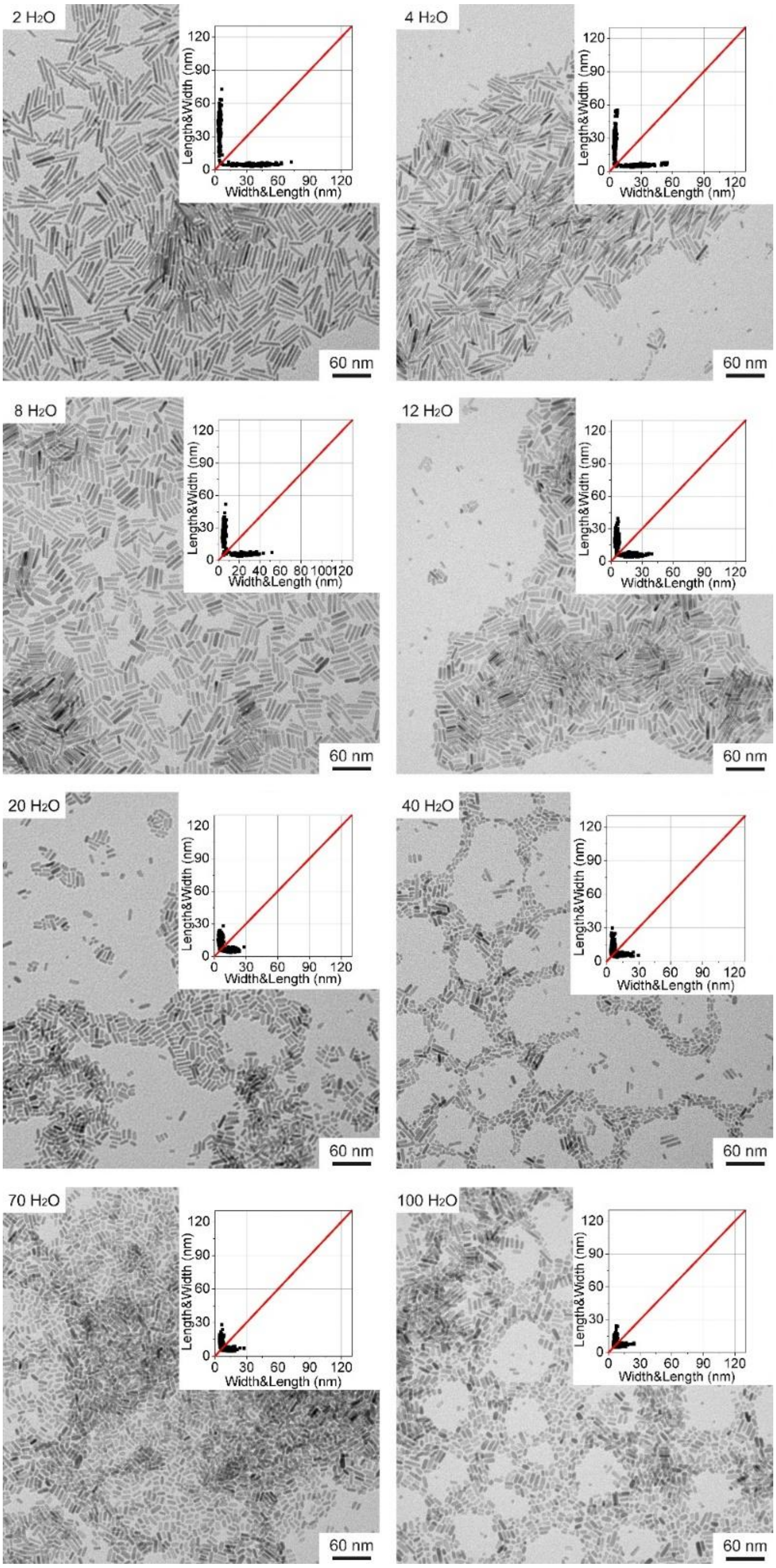


Figure S10. TEM picture and 2D size plots of ZnO NPs prepared in the presence of 2 eq. OA for different amount of water.

Table S11. Multivariate analysis of the 2D plots corresponding to the TEM images of the Figure S10 (ZnO NPs prepared in the presence of 2 eq. OA for different amount of water) through Rmixmod program (the dispersion is given as twice the standard deviation obtained from calculations, and number of cluster was set to 1).

Equivalent of H_2O	Length (nm)	Width (nm)	Correlation
2	36.6 ± 10.8	4.5 ± 1.3	0.25
4	25.8 ± 15.5	5.0 ± 1.3	0.35
8	24.3 ± 12.9	5.3 ± 1.6	0.26
12	18.9 ± 9.7	5.6 ± 1.7	0.05
20	15.0 ± 6.7	5.8 ± 1.7	0.07
40	11.4 ± 9.1	5.7 ± 1.5	0.16
70	12.0 ± 5.5	6.0 ± 1.4	0.18
100	11.0 ± 5.8	6.3 ± 1.6	0.46

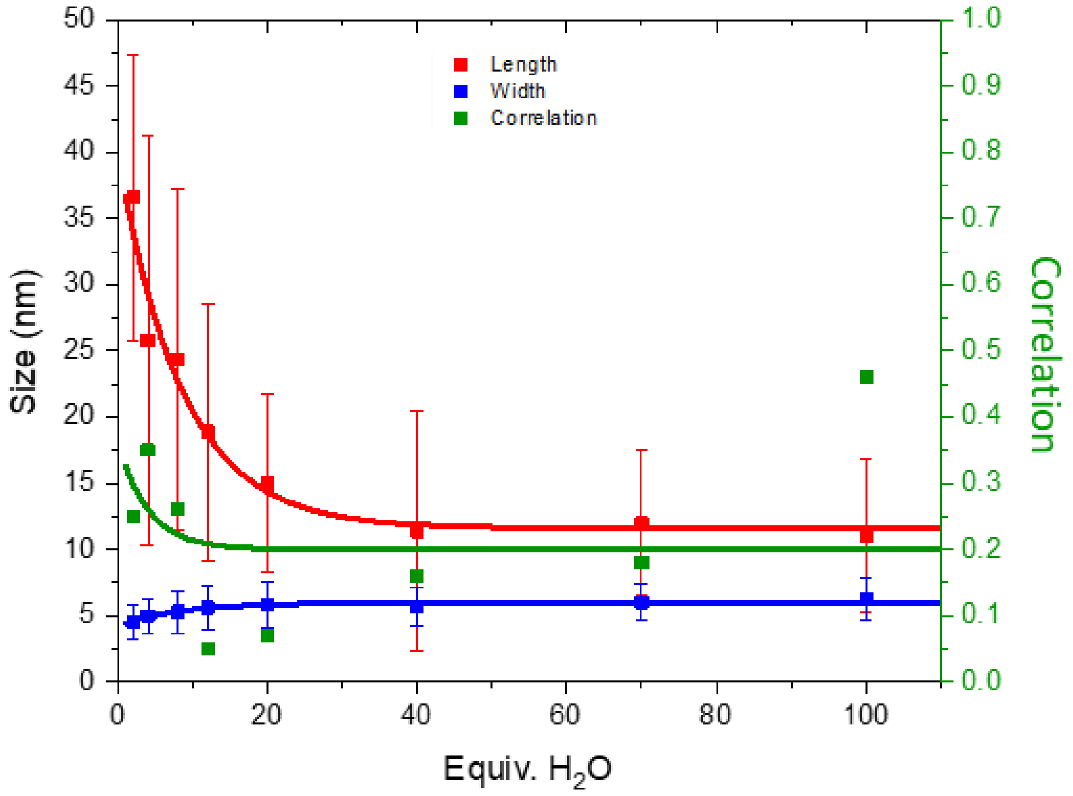


Figure S11: Mean width (red), mean length (blue) and corresponding correlation (green) issued from a multivariate analysis (with a single Gaussian) of data corresponding to ZnO NPs versus equivalent of water. The lines are just guides for the eyes. See also Figure S10 and Table S11.

Table S12: Multivariate analysis of the 2D plots corresponding to the TEM images of Figure S10 (ZnO NPs prepared in the presence of 2 eq. OA for different amount of water) through a modified Rmixmod program.

Water amount (equivalents)	Number of N1	Number of N2	Population of N1 (%)	Population of N2 (%)
2	2	298	0.7	99.3
4	9	293	3.0	97.0
8	10	291	3.3	96.7
12	45	255	15.0	85.0
20	125	175	41.7	58.3
40	198	102	66.0	34.0
70	248	52	82.7	17.3
100	266	34	88.7	11.3

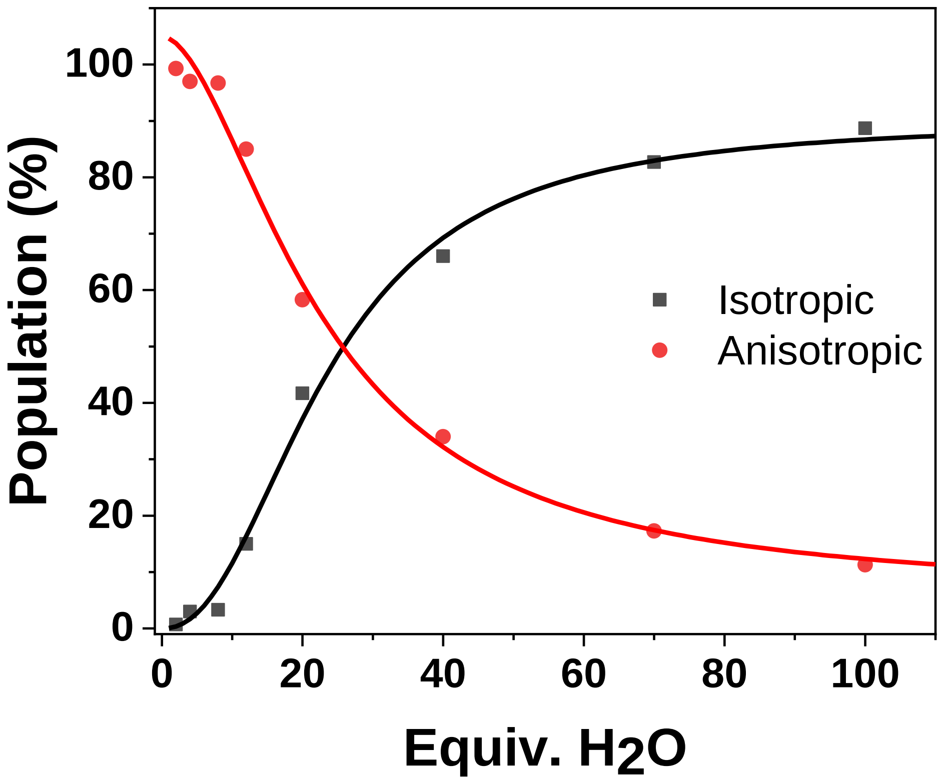


Figure S12. Percent of N1 and N2 depending on amount of water. See also the Figure S10 and Table S11.

1 Z. Zhao, Y. Coppel, J. Fitremann, P. Fau, C. Roux, C. Lepetit, P. Lecante, J.-D. Marty, C. Mingotaud, M. L. Kahn, *Chem. Mater.*, 2018, **30**, 8959-8967.
2 J. Rumble, *CRC Handbook of Chemistry and Physics, 100th Edition*, 2019.
3 https://www.wofrance.fr/weather/maps/city).