Supporting Information

The speciation of phosphates adsorbed on γ–alumina revealed by ^{31}P NMR, AIMD and machine learning

Adrian Hühn,[a] Tao Jiang,[a] Manuel Corral Valero,[b] Mickaël Rivallan,[b] Anne Lesage,[c] Carine Michel,[a]* and Pascal Raybaud[a,b]*

[a]   CNRS, Laboratoire de Chimie UMR 5182, ENS de Lyon, 46 allée d'Italie, Lyon F-69342, France. E-mail : carine.michel@ens-lyon.fr

[b]   IFP Énergies nouvelles, Rond-point de l'échangeur de Solaize, BP 3, 69360 Solaize, France. E-mail : pascal.raybaud@ifpen.fr

[c]   Université de Lyon, CNRS, ENS Lyon, Université Lyon 1, Centre de RMN à Hauts Champs de Lyon, UMR 5082, 5 rue de la Doua, 69100 Villeurbanne, France.

Contents

SI.1 Complementary methodological aspects

Table SI.1. Parametrization of Local Many-Body Tensor Representation (LMBTR) descriptors used in this work.

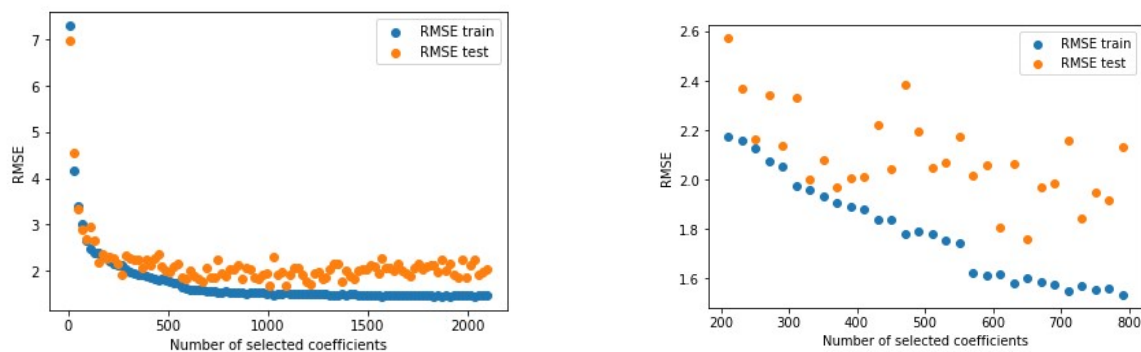|	Two body terms	Three body terms
Geometry function	Inverse distance Cut-off at 5.5 Å	Cosine of angles
Grid discretization	60 points between 0.18 and 1.2 Å^{-1}	80 points between -1 and +1
Weighting function	Exponential f(x) = e^{-sx} function with s = 0.5 Cut-off: 10^{-3}	
Normalization of results	Euclidian lengths were normalized to unity	

Figure SI.1. Root mean square error (RMSE) of the model after removing the lowest rank coefficients. Left: all coefficients considered. Right: zoom for models fitted with between 200 and 800 coefficients.

SI.2 Experimental ^{31}P NMR spectra

All experimental details regarding ^{31}P NMR spectra measurements and sample preparations can be found in our previous published works.[1] Here, we just quote an excerpt from the latter reference as a reminder of the upmost relevant experimental synthesis conditions: γ-Al$_2$O$_3$ samples were prepared from commercial powdered Pural SB3 (Sasol), which were shaped as extrudates, calcined at 450°C for 4 hours and crushed to reach particle sizes between 400 and 800 µm. The support has a surface area of 209 m² g^{-1} and a pore volume of 0.88 cm3 g^{-1}. The different samples were impregnated with phosphoric acid by the aqueous incipient wetness impregnation technique. A solution of diluted phosphoric acid was prepared from a commercial solution (85 %, Alfa Aesar) in different concentrations impregnated on the support. All samples were then dried at 120 °C.
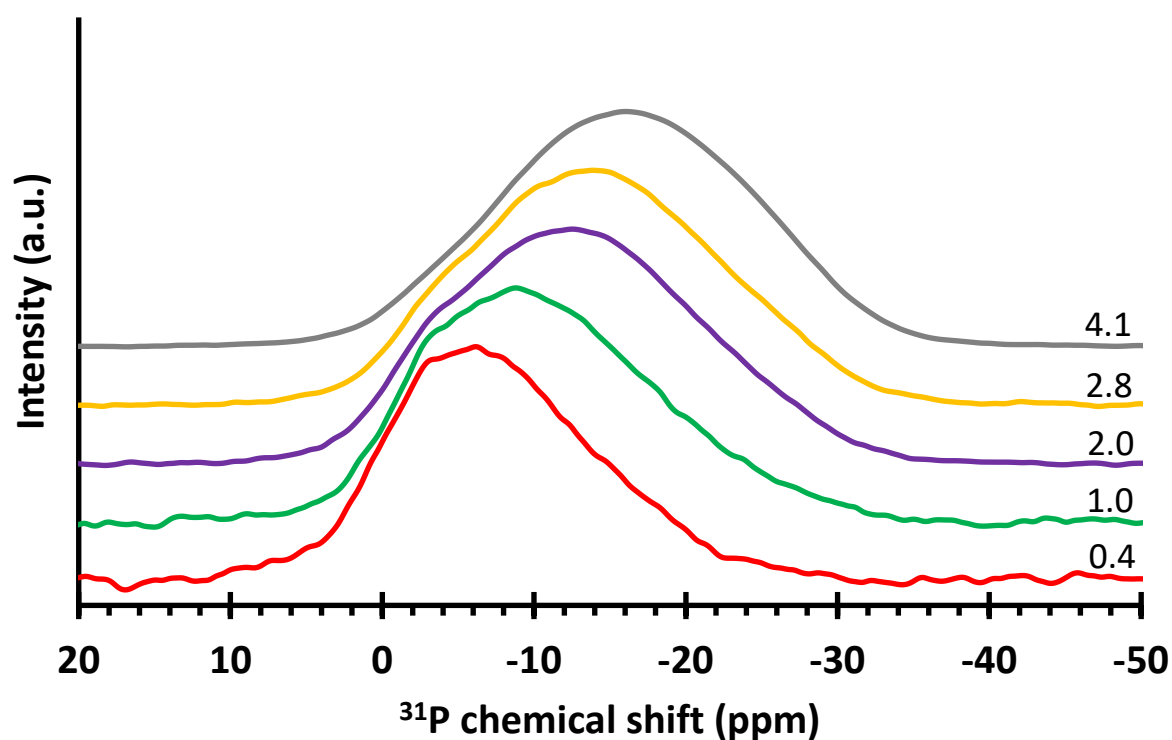

Figure SI.2. One-dimensional Experimental ^{31}P CP MAS spectra at different surface concentrations (in P nm^{-2}) of phosphate species impregnated on γ−alumina samples and dried at 120°C. The spectra were recorded at 9.4 T (400 MHz 1H resonance frequency), in a 2.5 mm zirconia rotor at a MAS rate of 30 kHz at room temperature. For more details, see reference.[1]

SI.3 Complementary results from the ML model

Histogram for Monophosphate case 100-v2

a)

Mean during AIMD = 3.7 +/- 6.0 (ppm)
At T= 0K, CS = 4.2 (ppm)

$(1\ 0\ 0)\ \nu_2$

Histogram for Monophosphate case 110-v3

b)

Mean during AIMD = -14.0 +/- 6.5 (ppm)
At T= 0K, CS = -11.0 (ppm)

$n(1\ 1\ 0)\ \nu_3$

Histogram for Monophosphate case R110-v2

c)

Mean during AIMD = -4.5 +/- 5.4 (ppm)
At T= 0K, CS = -4.4 (ppm)

$R(1\ 1\ 0)\ \nu_2$

Histogram for Monophosphate case R110-v3

d)

Mean during AIMD = -5.3 +/- 5.2 (ppm)
At T= 0K, CS = -1.7 (ppm)
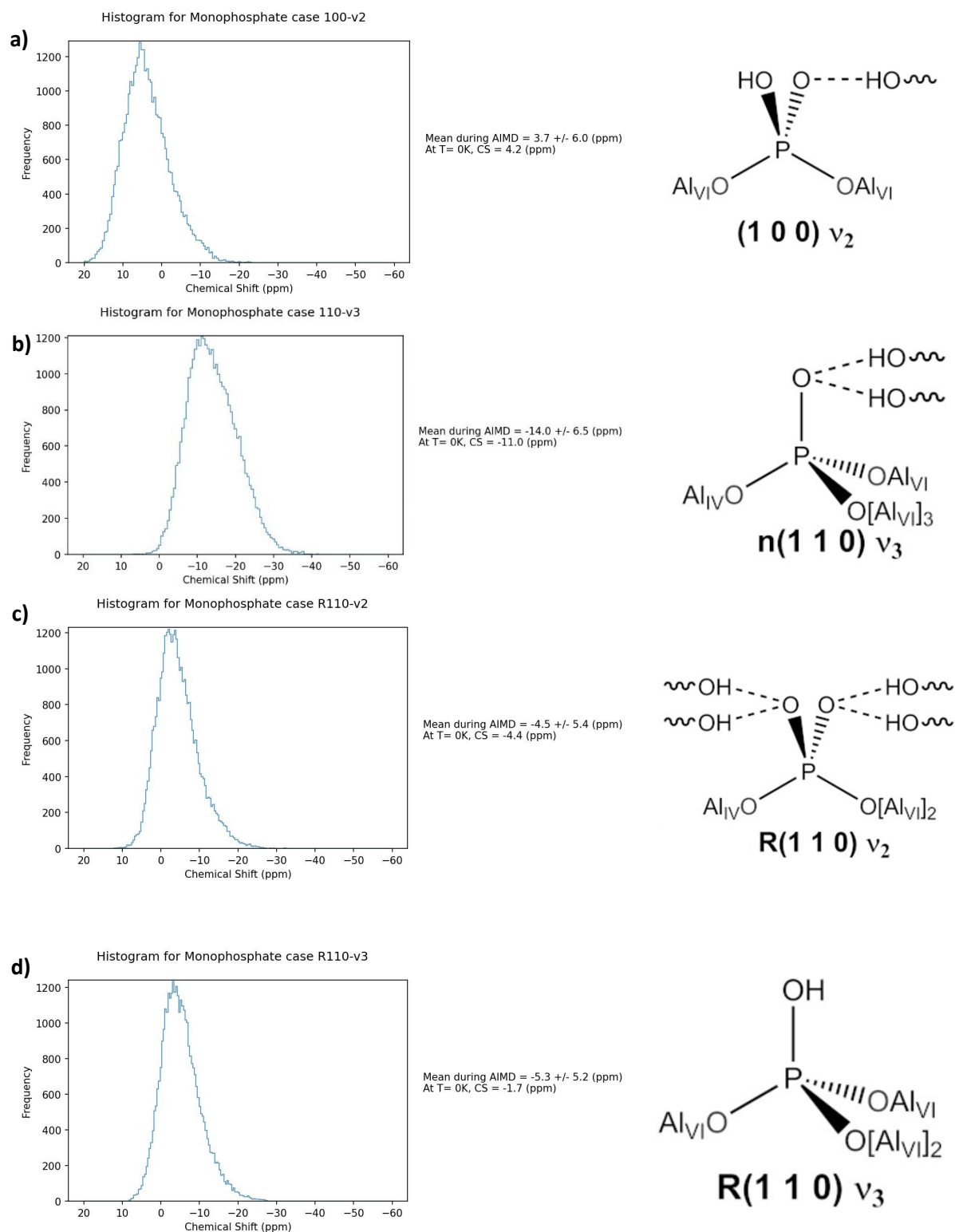
$R(1\ 1\ 0)\ \nu_3$

Figure SI.3. Left, histograms of modelled chemical shifts (CS) obtained from AIMD trajectories computed with monophosphates. Right: Lewis structure representation of the site at T= 0K.
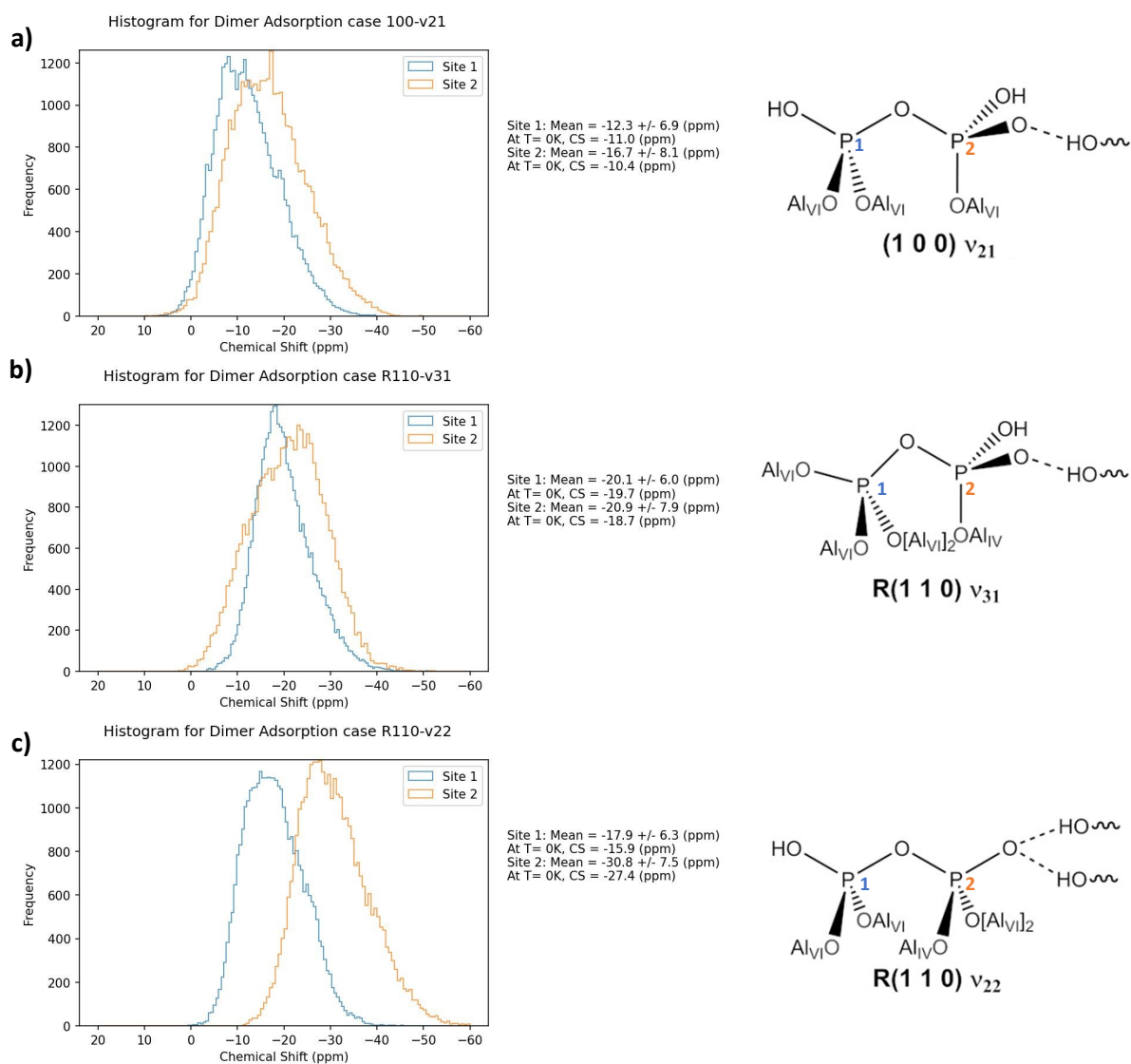
a) Histogram for Dimer Adsorption case 100-v21

Site 1: Mean = -12.3 +/- 6.9 (ppm)
At T= 0K, CS = -11.0 (ppm)
Site 2: Mean = -16.7 +/- 8.1 (ppm)
At T= 0K, CS = -10.4 (ppm)

(1 0 0) ν_{21}

b) Histogram for Dimer Adsorption case R110-v31

Site 1: Mean = -20.1 +/- 6.0 (ppm)
At T= 0K, CS = -19.7 (ppm)
Site 2: Mean = -20.9 +/- 7.9 (ppm)
At T= 0K, CS = -18.7 (ppm)

R(1 1 0) ν_{31}

c) Histogram for Dimer Adsorption case R110-v22

Site 1: Mean = -17.9 +/- 6.3 (ppm)
At T= 0K, CS = -15.9 (ppm)
Site 2: Mean = -30.8 +/- 7.5 (ppm)
At T= 0K, CS = -27.4 (ppm)

R(1 1 0) ν_{22}

Figure SI.4. Left, histograms of modelled chemical shifts (CS) obtained from AIMD trajectories computed with dimers. Right: Lewis structure representation of the site at T= 0 K.
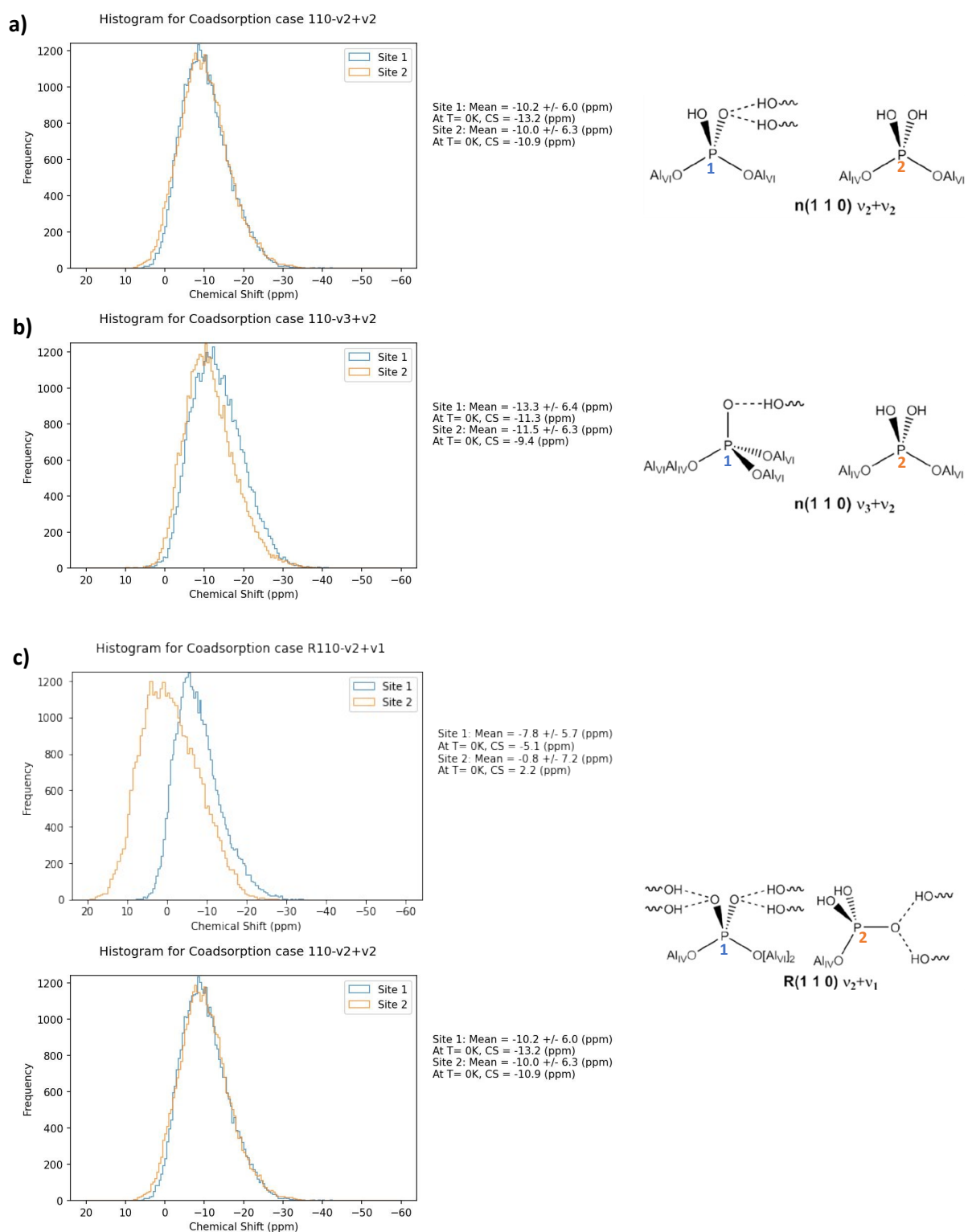
a)

Histogram for Coadsorption case 110-v2+v2

Site 1: Mean = -10.2 +/- 6.0 (ppm)
At T= 0K, CS = -13.2 (ppm)
Site 2: Mean = -10.0 +/- 6.3 (ppm)
At T= 0K, CS = -10.9 (ppm)

n(1 1 0) $\nu_2+\nu_2$

b)

Histogram for Coadsorption case 110-v3+v2

Site 1: Mean = -13.3 +/- 6.4 (ppm)
At T= 0K, CS = -11.3 (ppm)
Site 2: Mean = -11.5 +/- 6.3 (ppm)
At T= 0K, CS = -9.4 (ppm)

n(1 1 0) $\nu_3+\nu_2$

c)

Histogram for Coadsorption case R110-v2+v1

Site 1: Mean = -7.8 +/- 5.7 (ppm)
At T= 0K, CS = -5.1 (ppm)
Site 2: Mean = -0.8 +/- 7.2 (ppm)
At T= 0K, CS = 2.2 (ppm)

R(1 1 0) $\nu_2+\nu_1$

Histogram for Coadsorption case 110-v2+v2

Site 1: Mean = -10.2 +/- 6.0 (ppm)
At T= 0K, CS = -13.2 (ppm)
Site 2: Mean = -10.0 +/- 6.3 (ppm)
At T= 0K, CS = -10.9 (ppm)

Figure SI.5. Left, histograms of modelled chemical shifts (CS) obtained from AIMD trajectories computed with co-adsorbates. Right: Lewis structure representation of the site at T= 0K.

Table SI.2. Calculated adsorption energies (ΔG_{ads} in kJ mol^{-1}) corresponding to experimental drying conditions (T=393 K) and calculated ^{31}P chemical shift (CS in ppm) at DFT (T=0 K) level, with ML model (T=0 K) and after AIMD (T=300 K as in analytical conditions) for phosphate species on γ-alumina surfaces considered in this work.

Adsorption configuration	ΔG_{ads}*	DFT CS*	Static ML-CS	Mean AIMD ML-CS	SD**	DFT CS*	Static ML-CS	Mean AIMD ML-CS	SD**
	T= 393 K	T= 0 K	T= 0 K	T= 300 K		T= 0 K	T= 0 K	T= 300 K	
Monomers									
(100) v_2	-156	4.2	3.6	3.7	6.0				
n(110) v_3	-150	-11.0	-15.5	-14.0	6.5				
R(110) v_2	-160	-4.4	-4.7	-4.5	5.4				
R(110) v_3	-168	-1.7	-0.8	-5.3	5.2				
Dimers		**First P atom** (1) in Figure SI.4				**Second P atom** (2) in Figure SI.4			
(100) v_{21}	-106	-11.0	-10.9	-12.3	6.9	-10.4	-12.7	-16.7	8.1
R(110) v_{31}	-117	-19.7	-18.9	-20.1	6.0	-18.7	-21.7	-20.9	7.9
R(110) v_{22}	-116	-15.9	-16.7	-17.9	6.3	-27.4	-29.3	-30.8	7.6
Monomers co-adsorbed		**First P atom** (1) in Figure SI.4				**Second P atom** (2) in Figure SI.4			
n(110) v_2+v_2	-127	-13.2	-9.9	-10.2	6.0	-10.9	-11.9	-10.0	3.3
n(110) v_3+v_2	-128	-11.3	-12.0	-13.3	6.4	-9.4	-9.4	-11.5	6.3
R(110) v_2+v_1	-132	-5.1	-4.2	-7.8	5.7	2.2	2.1	-0.8	7.2

*Thermochemistry values and CS at T= 0 K are taken from work in[1].

**SD: standard deviation of the mean (in ppm).

SI.4 MCR-ALS analysis for the spectra decomposition

The set of theoretical data was used to fit the experimental ^{31}P CP MAS spectra by using multivariate curve resolution alternating least squares method (MCR-ALS) as implemented in PEAXACT chemometrics software (PEAXACT is developed by S-PACT GmbH. For more information, see www.s-pact.de/en/about_us).

The analysis is based on ^{31}P CP spectra, which are inherently not quantitative due to potential variations in CP efficiency between the different sites. This variation can arise from differences in the number and dynamics of surrounding protons. However, since all phosphorus sites are surrounded by multiple hydroxyl groups in the present conditions of mild drying (hydroxylation of alumina surfaces remains high), we believe that the ^{31}P CP signals provide reliable estimates of the relative contributions of each species. In other words, we expect that the efficiency of CP transfer is consistent for all surface species.

Multivariate curve resolution alternating least squares method (MCR-ALS) was used to decompose the experimental NMR spectra denoted as D (representing the recorded NMR spectra as function of P content on the alumina support), using pure components represented by S corresponding to theoretical NMR spectra of individual P containing species and their corresponding concentrations denoted by C. Mathematically, this decomposition can be expressed by the matrix equation $D = CS^t + E$, with E as error matrix.

1. Decomposition based on the AIMD-ML simulated histograms

Pure components of dimers and co-adsorbed species have been used after addition of theoretical contributions of both P sites. The RMSE values corresponding to the decomposed spectra illustrated in **Figure SI.6** are reported in **Table SI.3**. Any attempt to fit the spectra with species by using Gaussian functions and without any theoretical inputs lead to inconsistent solutions.

Table SI.3. RMSE values corresponding to the MCR-ALS analysis using AIMD-ML histograms.

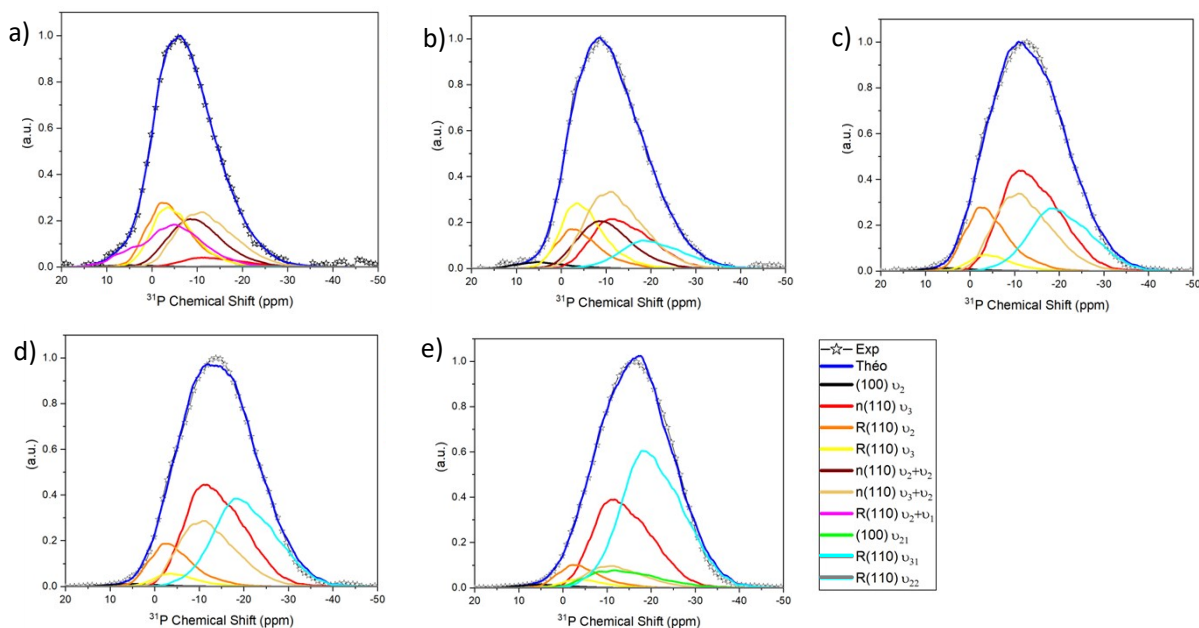| Sample P concentration (P nm^{-2})	RMSE
0.4	0.015
1.0	0.014
2.0	0.017
2.8	0.017
4.1	0.019

Figure SI.6. Decomposition of the experimental NMR spectra by using the theoretical histograms of each species obtained with the AIMD-ML approach for the five P concentrations considered in this work: a) 0.4, b) 1, c) 2, d) 2.8 and e) 4.1 P nm^{-2}. For improving the fit with experimental spectra and enhancing the clarity of obtained results, the raw simulated histograms in SI.3. have been smoothed.

2. Decomposition using Gaussian signals centered on 0 K DFT calculated chemical shifts

In this section, we assume that the spectrum of each species is determined by chemical shift obtained by DFT calculations at 0 K and a Gaussian signal centered on 0 K DFT calculated chemical shift with 6 ppm of full width at mid height (FWMH). Note that double gaussians must also be used for dimers and coadsorbed species.

First of all, as shown in the following table, the corresponding results exhibit RMSE values that are larger with this approach than with the one proposed in the current manuscript. At a first glance, the evolution of the species proportion exhibits some similar chemical trends: such as the evolutions of R(110) ν_2 and R(110) ν_{31} species. However, we also notice some significant changes in these evolutions: in that case, the n(110) ν_3 contribution is not evidenced anymore, whereas n(110) $\nu_2+\nu_2$ and R(110) ν_{22} become present, while the proportion of n(110) $\nu_3+\nu_2$ keeps increasing up to the highest P concentration. These differences are due, for one part, to the strong overlapping of the symmetrical gaussian signals centered at ca. -11 ppm where feature of n(110) $\nu_3+\nu_2$ takes the lead. Using asymmetrical components calculated from AIMD gives more constraints for fitting the experimental signal and allows revealing different species. For the second part, one should not overlook that there exists a small but non negligible impact of the temperature on the mean CS values (Table SI.3) which may change the identified species.

Table SI.4. RMSE values corresponding to the MCR-ALS analysis using Gaussian signals.

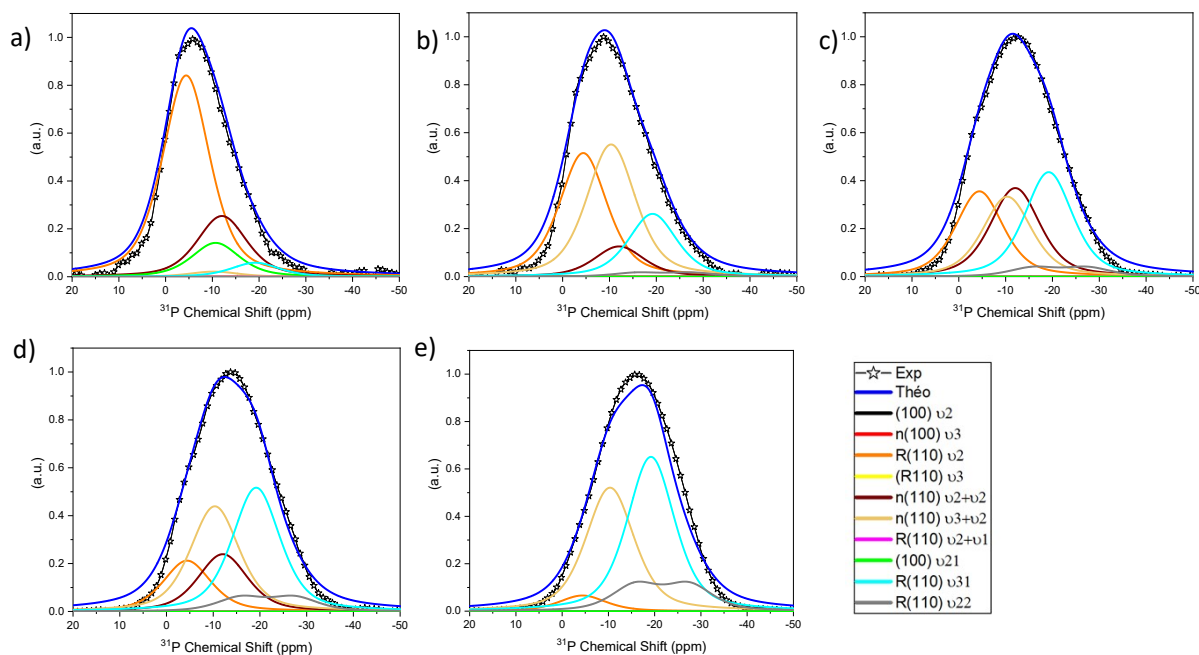| Sample P concentration (P nm^{-2})	RMSE
0.4	0.020
1.0	0.023
2.0	0.019
2.8	0.019
4.1	0.020

Figure SI.7. Decomposition of the experimental NMR spectra by using the 0 K DFT values of chemical shifts and a Gaussian feature with a fixed FWMH of 6 ppm for each species for the five P concentrations considered in this work: a) 0.4, b) 1, c) 2, d) 2.8 and e) 4.1 P nm^{-2}. For improving the fit with experimental spectra and enhancing the clarity of obtained results, the raw simulated histograms in SI.3. have been smoothed.
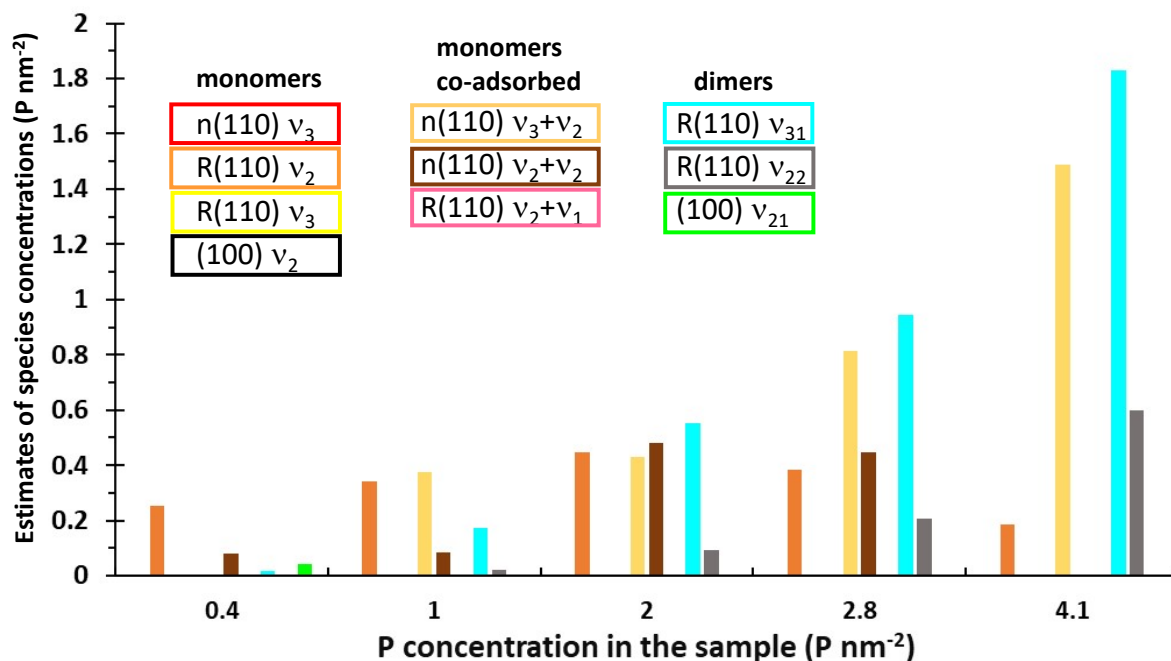

Figure SI.8. Quantitative estimates of the surface concentration for each phosphate species adsorbed on the three γ-alumina facets (in P nm^{-2}) and for each total P surface concentration. Their relative contributions were calculated from the deconvolution of the ^{31}P NMR lines as shown in **SI.7** and scaled from the knowledge of the total P surface concentration (in P nm^{-2}).

REFERENCES

1 A. Hühn, D. Wisser, M. Corral Valero, T. Roy, M. Rivallan, L. Catita, A. Lesage, C. Michel and P. Raybaud, *ACS Catal.*, 2021, **11**, 11278.