# Supporting Information

# Revision 1

# Influence of Fe(II), Fe(III) and Al(III) Isomorphic Substitutions on Acid-Base Properties of Edge Surfaces of Cis-Vacant Montmorillonite: Insights from First-Principles Molecular Dynamics Simulations and Surface Complexation Modeling

**Pengyuan Gao[1,2,3], Xiandong Liu[4,5], Zhijun Guo[1,2], Christophe Tournassat[3,6*]**

*1. Radiochemistry Laboratory, School of Nuclear Science and Technology, Lanzhou University, Lanzhou 730000, China*
*2. Frontiers Science Center for Rare Isotopes, Lanzhou University, Lanzhou 730000, China*
*3. Institut des Sciences de la Terre d'Orléans, Université d'Orléans-CNRS-BRGM, Orléans 45071, France*
*4. State Key Laboratory for Mineral Deposits Research, School of Earth Sciences and Engineering, Nanjing University, Nanjing 210093, China*
*5. Frontiers Science Center for Critical Earth Material Cycling, Nanjing University, Nanjing 210023, China*
*6. Earth and Environmental Sciences Area, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

**\* Corresponding author: <u>christophe.tournassat@univ-orleans.fr</u>**

**This file contains 51 pages, including 6 tables and 2 figures.**

## Contents

**Supplementary: Method to calculate acidity constant**

In this method, the half-reaction scheme was applied (Costanzo et al. 2011; Cheng et al. 2012, 2014;). The proton of an acid (denoted as AH) is gradually transformed into a ghost atom that can be considered a classical particle without any charge. The free energy of the transformation is computed as

$$\Delta A = \int_0^1 \langle \Delta E \rangle_\eta \, d\eta \qquad (S1)$$

Here the $\Delta E$ is the vertical energy gap which is defined as the difference between the potential energies of the reactant and the product states. The subscript $\eta$ is a coupling parameter that increases from 0 (the protonated state) to 1 (the deprotonated state). The mean value of $\Delta E$ is derived with the restrained mapping Hamiltonian:

$$H_\eta = (1-\eta)H_R + \eta H_P + V_r \qquad (S2)$$

$H_R$ and $H_P$ refer to the reactant and product states, respectively. The restrained harmonic potential $V_r$ is used to ensure that the ghost atom stays at the same position as that in the simulation of the reactant state. The restraining potential is of the form familiar from classical force field models:

$$V_r = \sum_{i=1}^{n_d} \frac{1}{2} k_d (d-d_0)^2 + \sum_{i=1}^{n_\alpha} \frac{1}{2} k_\alpha (\alpha-\alpha_0)^2 \qquad (S3)$$

The leading term is a sum of harmonic potentials of bond distance and bond angles with the respective equilibrium values of $d_0$ and $\alpha_0$, which were derived from the prior simulations of their reactant states without constraint. The used restraining parameters of Vr are given in Tables S1-S3.

In actual calculations, the right-hand side of Eqs. S1 is estimated by the three-point

Simpson's rule which requires running the simulations of $\eta = 0.0$, $0.5$, and $1.0$, respectively. Then the deprotonation-free energy can be expressed as:

$$\Delta A = \frac{1}{6}\left(\langle\Delta E\rangle_0 + \langle\Delta E\rangle_1\right) + \frac{2}{3}\langle\Delta E\rangle_{0.5} \tag{S4}$$

With the same method, one proton of an $H_3O^+$ in the solution area of the same system gradually switched off to become a dummy, the free energy of which is denoted as $\int_0^1 d\eta \langle\Delta E_{H_3O^+}\rangle_\eta$.

Combining the results for the two steps provides the result for the $pK_a$:

$$2.30 k_B T pK_a = \int_0^1 d\eta \langle\Delta E\rangle_\eta - \int_0^1 d\eta \langle\Delta E_{H_3O^+}\rangle_\eta + k_B T \ln\left[c^0 \Lambda_{H^+}^3\right] \tag{S5}$$

Here, $c^0$ is the unit molar concentration whose value is 1.0 mol/L. $\Lambda_{H^+}$ is the thermal wavelength of the proton (Costanzo et al. 2011). The third term $k_B T \ln\left[c^0 \Lambda_{H^+}^3\right]$ represents the translational entropy caused by proton dissociation, approximated by the chemical potential of a free proton at the standard concentration (-0.19 eV or 3.2 $pK_a$ units).

**TABLE S1.** Parameters used in harmonic potentials restrain the dummy protons of the $Fe_{II}$-sub model. $H_d$ means the dummy proton. Equilibrium bond lengths ($d_0$) are in bohr and equilibrium angles ($\alpha_0$) are in radians. All the coupling constants are in a.u.

| Surface | Acids | $n_d$ | $d_0$ | $k_d$ | $n_\alpha$ | $\alpha_0$ | $k_\alpha$ |
|---|---|---|---|---|---|---|---|
| $\perp[010]$ | $\equiv SiOH$ | 1 | 1.89 | 0.1 | 1 | 1.94 (Si-O-Hd) | 0.1 |
| | $\equiv Si(OH)_2Fe$ | 1 | 1.92 | 0.1 | 2 | 1.88 (Si-O-Hd) 1.93 (Fe-O-Hd) | 0.1 |
| | $\equiv Si(O)(OH)Fe$ | 1 | 1.93 | 0.1 | 2 | 1.88 (Si-O-Hd) 1.93 (Fe-O-Hd) | 0.1 |
| $\perp[0\bar{1}0]$ | $\equiv SiOH$ | 1 | 1.89 | 0.1 | 1 | 2.02 (Si-O-Hd) | 0.1 |
| | $\equiv Al(OH_2)_2$ | 1 | 1.90 | 0.1 | 2 | 2.02 (H-O-Hd) 2.06 (Al-O-Hd) | 0.1 |
| | $\equiv Fe(OH)Al$ | 1 | 1.93 | 0.1 | 2 | 2.20 (Al-O-Hd) 1.98 (Fe-O-Hd) | 0.1 |
| $\perp[110]$ | $\equiv SiOH^U$ | 1 | 1.89 | 0.1 | 1 | 1.96 (Si-O-Hd) | 0.1 |
| | $\equiv SiOH^L$ | 1 | 1.89 | 0.1 | 1 | 2.01 (Si-O-Hd) | 0.1 |
| | $\equiv Si(OH)Al$ | 1 | 1.89 | 0.1 | 2 | 2.20 (Si-O-Hd) 2.10 (Al-O-Hd) | 0.1 |
| | $\equiv Al(OH_2)$ | 1 | 1.83 | 0.1 | 2 | 1.89 (H-O-Hd) 2.02 (Al-O-Hd) | 0.1 |
| $\perp[\bar{1}\bar{1}0]$ | $\equiv SiOH^U$ | 1 | 1.89 | 0.1 | 1 | 2.04 (Si-O-Hd) | 0.1 |
| | $\equiv SiOH^L$ | 1 | 1.92 | 0.1 | 1 | 2.09 (Si-O-Hd) | 0.1 |
| | $\equiv Fe(OH_2)_2$ | 1 | 1.89 | 0.1 | 2 | 1.89 (H-O-Hd) 2.13 (Fe-O-Hd) | 0.1 |
| | $\equiv Fe(OH)Al$ | 1 | 1.87 | 0.1 | 2 | 2.25 (Al-O-Hd) 1.94 (Fe-O-Hd) | 0.1 |
| | $H_3O^+$ | 3 | 1.89 | 1.0 | 3 | 1.94 (H-O-Hd) 1.94 (H-O-Hd) 1.94 (H-O-H) | 0.1 |

**TABLE S2.** Parameters used in harmonic potentials restrain the dummy protons of the $Fe_{III}$-sub model. $H_d$ means the dummy proton. Equilibrium bond lengths ($d_0$) are in bohr and equilibrium angles ($\alpha_0$) are in radians. All the coupling constants are in a.u.

| Surface | Acids | $n_d$ | $d_0$ | $k_d$ | $n_\alpha$ | $\alpha_0$ | $k_\alpha$ |
|---|---|---|---|---|---|---|---|
| $\perp[010]$ | $\equiv$SiOH | 1 | 1.89 | 0.1 | 1 | 1.94 (Si-O-Hd) | 0.1 |
| | $\equiv$Si(OH)$_2$Fe | 2 | 1.92<br>1.92 | 0.1 | 2 | 1.88 (Si-O-Hd)<br>1.93 (Fe-O-Hd) | 0.1 |
| | $\equiv$Si(O)(OH)Fe | 1 | 1.92 | 0.1 | 2 | 1.88 (Si-O-Hd)<br>1.93 (Fe-O-Hd) | 0.1 |
| $\perp[0\bar{1}0]$ | $\equiv$SiOH | 1 | 1.89 | 0.1 | 1 | 2.02 (Si-O-Hd) | 0.1 |
| | $\equiv$Al(OH$_2$)$_2$ | 2 | 1.88<br>1.88 | 0.1 | 2 | 1.88 (H-O-Hd)<br>2.05 (Al-O-Hd) | 0.1 |
| | $\equiv$Al(OH)(OH$_2$) | 1 | 1.88 | 0.1 | 2 | 1.95 (H-O-Hd)<br>2.16 (Al-O-Hd) | 0.1 |
| | $\equiv$Fe(OH)Al | 1 | 1.93 | 0.1 | 2 | 2.20 (Al-O-Hd)<br>1.98 (Fe-O-Hd) | 0.1 |
| $\perp[110]$ | $\equiv$SiOH$^U$ | 1 | 1.89 | 0.1 | 1 | 1.96 (Si-O-Hd) | 0.1 |
| | $\equiv$SiOH$^L$ | 1 | 1.89 | 0.1 | 1 | 2.01 (Si-O-Hd) | 0.1 |
| | $\equiv$Si(OH)Al | 1 | 1.89 | 0.1 | 2 | 2.20 (Si-O-Hd)<br>2.10 (Al-O-Hd) | 0.1 |
| | $\equiv$Al(OH$_2$) | 1 | 1.83 | 0.1 | 2 | 1.89 (H-O-Hd)<br>2.05 (Al-O-Hd) | 0.1 |
| $\perp[\bar{1}\bar{1}0]$ | $\equiv$SiOH$^U$ | 1 | 1.89 | 0.1 | 1 | 2.04 (Si-O-Hd) | 0.1 |
| | $\equiv$SiOH$^L$ | 1 | 1.92 | 0.1 | 1 | 2.09 (Si-O-Hd) | 0.1 |
| | $\equiv$Fe(OH$_2$)$_2$ | 1 | 1.89 | 0.1 | 2 | 1.89 (H-O-Hd)<br>2.10 (Fe-O-Hd) | 0.1 |
| | $\equiv$Fe(OH)(OH$_2$) | 1 | 1.89 | 0.1 | 2 | 1.89 (H-O-Hd)<br>2.13 (Fe-O-Hd) | 0.1 |
| | $\equiv$Fe(OH)Al | 1 | 1.87 | 0.1 | 2 | 2.28 (Al-O-Hd)<br>1.94 (Fe-O-Hd) | 0.1 |
| | H$_3$O$^+$ | 3 | 1.89 | 1.0 | 3 | 1.94 (H-O-Hd)<br>1.94 (H-O-Hd)<br>1.94 (H-O-H) | 0.1 |

**TABLE S3.** Parameters used in harmonic potentials restrain the dummy protons of the Al-sub model. $H_d$ means the dummy proton. Equilibrium bond lengths ($d_0$) are in bohr and equilibrium angles ($\alpha_0$) are in radians. All the coupling constants are in a.u.

| Surface | Acids | $n_d$ | $d_0$ | $k_d$ | $n_\alpha$ | $\alpha_0$ | $k_\alpha$ |
|---|---|---|---|---|---|---|---|
| $\perp[010]$ | $\equiv Al^T(OH_2)$ | 1 | 1.88 | 0.1 | 2 | 1.91 (H-O-Hd) | 0.1 |
| | | | | | | 2.11 ($Al^T$-O-Hd) | |
| | $\equiv Al^T(OH)$ | 1 | 1.88 | 0.1 | 1 | 2.03 ($Al^T$-O-Hd) | 0.1 |
| | $\equiv Al^T(OH)_2Al$ | 1 | 1.82 | 0.1 | 2 | 1.80 ($Al^T$-O-Hd) | 0.1 |
| | | | | | | 2.07 (Al-O-Hd) | |
| | $\equiv Al^T(O)(OH)Al$ | 1 | 1.82 | 0.1 | 2 | 1.80 ($Al^T$-O-Hd) | 0.1 |
| | | | | | | 2.07 (Al-O-Hd) | |
| $\perp[0\overline{1}0]$ | $\equiv Al^T(OH_2)_2$ | 1 | 1.89 | 0.1 | 2 | 1.88 (H-O-Hd) | 0.1 |
| | | | | | | 2.13 ($Al^T$-O-Hd) | |
| | $\equiv Al^T(OH)$ | 1 | 1.89 | 0.1 | 1 | 2.02 ($Al^T$-O-Hd) | 0.1 |
| | $\equiv Al(OH_2)_2$ | 1 | 1.89 | 0.1 | 2 | 2.01 (H-O-Hd) | 0.1 |
| | | | | | | 2.21 (Al-O-Hd) | |
| | $\equiv Al(OH)(OH_2)$ | 1 | 1.89 | 0.1 | 2 | 1.95 (H-O-Hd) | 0.1 |
| | | | | | | 2.01 (Al-O-Hd) | |
| $\perp[110]$ | $\equiv Al^T(OH_2)_2$ | 1 | 1.86 | 0.1 | 2 | 1.86 (H-O-Hd) | 0.1 |
| | | | | | | 2.12 ($Al^T$-O-Hd) | |
| | $\equiv Al^T(OH)$ | 1 | 1.86 | 0.1 | 1 | 2.02 ($Al^T$-O-Hd) | 0.1 |
| | $\equiv Al^T(OH)Al$ | 1 | 1.88 | 0.1 | 2 | 2.01 ($Al^T$-O-Hd) | 0.1 |
| | | | | | | 1.92 (Al-O-Hd) | |
| | $\equiv Al(OH_2)$ | 1 | 1.87 | 0.1 | 2 | 1.82 (H-O-Hd) | 0.1 |
| | | | | | | 2.12 (Al-O-Hd) | |
| $\perp[\overline{11}0]$ | $\equiv Al^T(OH_2)_2$ | 1 | 1.89 | 0.1 | 2 | 1.81 (H-O-Hd) | 0.1 |
| | | | | | | 2.03 ($Al^T$-O-Hd) | |
| | $\equiv Al^T(OH)$ | 1 | 1.89 | 0.1 | 1 | 2.01 ($Al^T$-O-Hd) | 0.1 |
| | $\equiv Al(OH_2)_2$ | 1 | 1.92 | 0.1 | 2 | 1.91 (H-O-Hd) | 0.1 |
| | | | | | | 2.19 (Al-O-Hd) | |
| | $\equiv Al(OH)(OH_2)$ | 1 | 1.92 | 0.1 | 2 | 1.91 (H-O-Hd) | 0.1 |
| | | | | | | 2.19 (Al-O-Hd) | |
| | $H_3O^+$ | 3 | 1.89 | 1.0 | 3 | 1.94 (H-O-Hd) | 0.1 |
| | | | | | | 1.94 (H-O-Hd) | |
| | | | | | | 1.94 (H-O-H) | |

**TABLE S4.** Free energies (in eV) and pKa values of surface groups of the $Fe_{II}$-sub model.

| Models | Sites | $\eta = 1.0$ | $\eta = 0.5$ | $\eta = 0.0$ | $\Delta_{dp}A/eV$ | pKas |
|---|---|---|---|---|---|---|
| $\perp[010]$ | $\equiv SiOH$ | 12.86 $\pm$ 0.03 | 18.31 $\pm$ 0.03 | 20.25 $\pm$ 0.05 | 17.73 $\pm$ 0.03 | 12.2 $\pm$ 1.0 |
| | $\equiv Si(OH)_2Fe$ | 14.31 $\pm$ 0.05 | 17.96 $\pm$ 0.04 | 20.57 $\pm$ 0.03 | 17.79 $\pm$ 0.04 | 13.2 $\pm$ 1.1 |
| | $\equiv Si(O)(OH)Fe$ | 13.69 $\pm$ 0.03 | 18.52 $\pm$ 0.02 | 20.38 $\pm$ 0.03 | 18.03 $\pm$ 0.02 | 17.2 $\pm$ 0.8 |
| | $H_3O^+$ | 13.04 $\pm$ 0.03 | 17.18 $\pm$ 0.02 | 19.08 $\pm$ 0.04 | 16.81 $\pm$ 0.03 | — |
| $\perp[0\bar{1}0]$ | $\equiv SiOH$ | 14.42 $\pm$ 0.05 | 17.95 $\pm$ 0.04 | 20.00 $\pm$ 0.04 | 17.70 $\pm$ 0.04 | 13.5 $\pm$ 1.1 |
| | $\equiv Al(OH_2)_2$ | 13.46 $\pm$ 0.08 | 17.91 $\pm$ 0.05 | 19.93 $\pm$ 0.05 | 17.51 $\pm$ 0.03 | 10.2 $\pm$ 0.9 |
| | $\equiv Fe(OH)Al$ | 14.14 $\pm$ 0.04 | 18.39 $\pm$ 0.06 | 20.88 $\pm$ 0.03 | 18.10 $\pm$ 0.03 | 20.1 $\pm$ 1.2 |
| | $H_3O^+$ | 13.45 $\pm$ 0.03 | 16.83 $\pm$ 0.02 | 19.50 $\pm$ 0.02 | 16.71 $\pm$ 0.02 | — |
| $\perp[110]$ | $\equiv SiOH^U$ | 13.94 $\pm$ 0.04 | 18.25 $\pm$ 0.03 | 20.06 $\pm$ 0.02 | 17.83 $\pm$ 0.03 | 13.8 $\pm$ 0.9 |
| | $\equiv SiOH^L$ | 13.65 $\pm$ 0.06 | 18.23 $\pm$ 0.04 | 20.03 $\pm$ 0.05 | 17.77 $\pm$ 0.05 | 12.7 $\pm$ 1.2 |
| | $\equiv Si(OH)Al$ | 13.38 $\pm$ 0.03 | 16.78 $\pm$ 0.05 | 18.83 $\pm$ 0.02 | 16.56 $\pm$ 0.04 | -7.7 $\pm$ 1.1 |
| | $\equiv Al(OH_2)$ | 14.40 $\pm$ 0.02 | 17.88 $\pm$ 0.01 | 19.94 $\pm$ 0.02 | 17.64 $\pm$ 0.01 | 10.7 $\pm$ 0.6 |
| | $H_3O^+$ | 12.34 $\pm$ 0.04 | 17.21 $\pm$ 0.02 | 19.76 $\pm$ 0.02 | 16.82 $\pm$ 0.02 | — |

| | | | | | | |
|---|---|---|---|---|---|---|
| | $\equiv SiOH^U$ | 14.26 ± 0.04 | 18.25 ± 0.03 | 20.10 ± 0.03 | 17.89 ± 0.03 | 12.3 ± 1.0 |
| | $\equiv SiOH^L$ | 13.75 ± 0.03 | 18.32 ± 0.08 | 20.00 ± 0.02 | 17.84 ± 0.06 | 11.4 ± 1.5 |
| $\perp[\overline{11}0]$ | $\equiv Fe(OH_2)_2$ | 13.48 ± 0.03 | 18.25 ± 0.06 | 20.36 ± 0.02 | 17.81 ± 0.05 | 10.9 ± 1.3 |
| | $\equiv Fe(OH)Al$ | 15.61 ± 0.05 | 18.36 ± 0.02 | 20.98 ± 0.01 | 18.34 ± 0.02 | 19.8 ± 0.9 |
| | $H_3O^+$ | 12.55 ± 0.02 | 17.46 ± 0.03 | 19.40 ± 0.03 | 16.97 ± 0.03 | — |

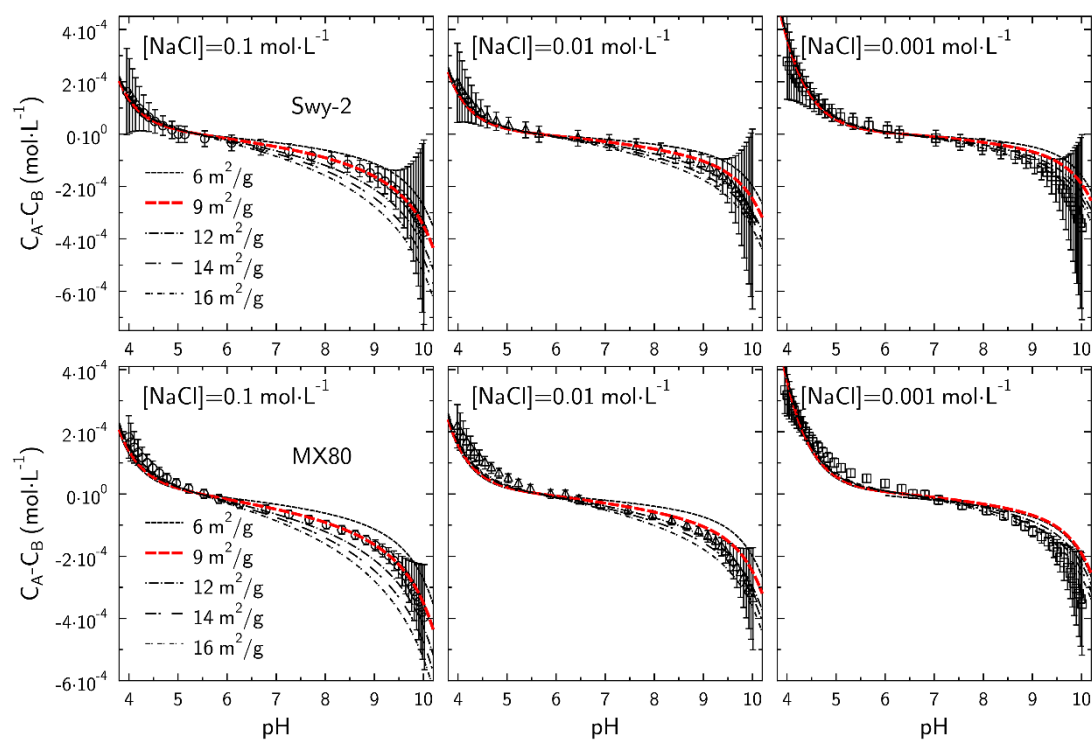**TABLE S5.** Free energies (in eV) and pKa values of surface groups of the $Fe_{III}$-sub model.

| Models | Sites | $\eta = 1.0$ | $\eta = 0.5$ | $\eta = 0.0$ | $\Delta_{dp}A/eV$ | pKas |
|---|---|---|---|---|---|---|
| $\perp[010]$ | $\equiv SiOH$ | 14.23 ± 0.01 | 17.97 ± 0.03 | 19.50 ± 0.01 | 17.60 ± 0.02 | 8.8 ± 0.6 |
| | $\equiv Si(OH)_2Fe$ | 14.15 ± 0.03 | 17.28 ± 0.03 | 19.05 ± 0.02 | 17.05 ± 0.03 | -0.4 ± 0.7 |
| | $\equiv Si(O)(OH)Fe$ | 14.95 ± 0.02 | 17.48 ± 0.01 | 19.63 ± 0.03 | 17.42 ± 0.02 | 5.7 ± 0.5 |
| | $H_3O^+$ | 12.81 ± 0.02 | 17.42 ± 0.01 | 18.84 ± 0.02 | 16.89 ± 0.02 | — |
| $\perp[0\bar{1}0]$ | $\equiv SiOH$ | 14.12 ± 0.02 | 17.79 ± 0.02 | 19.64 ± 0.04 | 17.49 ± 0.02 | 9.0 ± 0.7 |
| | $\equiv Al(OH_2)_2$ | 13.05 ± 0.01 | 17.29 ± 0.03 | 19.45 ± 0.03 | 16.94 ± 0.03 | -0.1 ± 0.8 |
| | $\equiv Al(OH)(OH_2)$ | 13.06 ± 0.03 | 17.78 ± 0.01 | 19.71 ± 0.02 | 17.32 ± 0.02 | 6.1 ± 0.6 |
| | $\equiv Fe(OH)Al$ | 13.71 ± 0.01 | 17.90 ± 0.03 | 19.55 ± 0.04 | 17.48 ± 0.03 | 8.9 ± 0.8 |
| | $H_3O^+$ | 12.27 ± 0.02 | 17.32 ± 0.02 | 18.99 ± 0.01 | 16.76 ± 0.02 | — |
| $\perp[110]$ | $\equiv SiOH^U$ | 14.22 ± 0.01 | 18.31 ± 0.02 | 19.77 ± 0.03 | 17.87 ± 0.02 | 8.4 ± 0.7 |
| | $\equiv SiOH^L$ | 14.17 ± 0.03 | 18.29 ± 0.04 | 19.42 ± 0.05 | 17.79 ± 0.04 | 7.1 ± 1.0 |
| | $\equiv Si(OH)Al$ | 13.35 ± 0.02 | 16.80 ± 0.07 | 18.47 ± 0.02 | 16.50 ± 0.05 | -14.6 ± 1.2 |
| | $\equiv Al(OH_2)$ | 14.83 ± 0.04 | 17.81 ± 0.02 | 19.12 ± 0.02 | 17.53 ± 0.02 | 2.7 ± 0.7 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | $H_3O^+$ | 13.26 $\pm$ 0.01 | 17.69 $\pm$ 0.02 | 19.04 $\pm$ 0.03 | 17.18 $\pm$ 0.02 | — |
| $\perp[\overline{1}\overline{1}0]$ | $\equiv SiOH^U$ | 14.31 $\pm$ 0.02 | 17.95 $\pm$ 0.06 | 20.08 $\pm$ 0.03 | 17.70 $\pm$ 0.05 | 7.9 $\pm$ 1.3 |
| | $\equiv SiOH^L$ | 13.85 $\pm$ 0.03 | 18.08 $\pm$ 0.03 | 20.09 $\pm$ 0.01 | 17.71 $\pm$ 0.03 | 8.1 $\pm$ 0.9 |
| | $\equiv Fe(OH_2)_2$ | 14.61 $\pm$ 0.07 | 16.82 $\pm$ 0.01 | 18.76 $\pm$ 0.06 | 16.78 $\pm$ 0.03 | -7.7 $\pm$ 0.9 |
| | $\equiv Fe(OH)(OH_2)$ | 13.53 $\pm$ 0.03 | 17.51 $\pm$ 0.08 | 20.30 $\pm$ 0.04 | 17.37 $\pm$ 0.07 | 2.3 $\pm$ 1.5 |
| | $\equiv Fe(OH)Al$ | 13.69 $\pm$ 0.01 | 18.20 $\pm$ 0.01 | 19.74 $\pm$ 0.03 | 17.71 $\pm$ 0.01 | 8.0 $\pm$ 0.7 |
| | $H_3O^+$ | 12.80 $\pm$ 0.03 | 17.66 $\pm$ 0.03 | 18.77 $\pm$ 0.01 | 17.04 $\pm$ 0.03 | — |

**TABLE S6.** Free energies (in eV) and pKa values of surface groups of the Al-sub model.

| Models | Sites | $\eta = 1.0$ | $\eta = 0.5$ | $\eta = 0.0$ | $\Delta_{dp}A$/eV | pKas |
|---|---|---|---|---|---|---|
| $\perp[010]$ | $\equiv Al^T(OH_2)$ | 13.88 $\pm$ 0.08 | 17.68 $\pm$ 0.03 | 20.13 $\pm$ 0.04 | 17.46 $\pm$ 0.04 | 3.1 $\pm$ 1.3 |
| | $\equiv Al^T(OH)$ | 13.57 $\pm$ 0.03 | 18.48 $\pm$ 0.04 | 21.11 $\pm$ 0.03 | 18.10 $\pm$ 0.04 | 14.0 $\pm$ 1.3 |
| | $\equiv Al^T(OH)_2Al$ | 14.34 $\pm$ 0.03 | 17.91 $\pm$ 0.05 | 20.17 $\pm$ 0.05 | 17.69 $\pm$ 0.05 | 7.1 $\pm$ 1.4 |
| | $\equiv Al^T(OH)(O)Al$ | 13.27 $\pm$ 0.01 | 18.88 $\pm$ 0.03 | 20.58 $\pm$ 0.02 | 18.23 $\pm$ 0.03 | 16.1 $\pm$ 1.0 |
| | $H_3O^+$ | 12.76 $\pm$ 0.02 | 17.43 $\pm$ 0.04 | 19.99 $\pm$ 0.04 | 17.08 $\pm$ 0.04 | — |
| $\perp[0\bar{1}0]$ | $\equiv Al^T(OH_2)_2$ | 13.48 $\pm$ 0.03 | 17.21 $\pm$ 0.01 | 20.64 $\pm$ 0.03 | 17.16 $\pm$ 0.02 | 2.9 $\pm$ 0.7 |
| | $\equiv Al^T(OH)$ | 14.67 $\pm$ 0.05 | 17.81 $\pm$ 0.03 | 20.84 $\pm$ 0.04 | 17.79 $\pm$ 0.04 | 14.2 $\pm$ 1.0 |
| | $\equiv Al(OH_2)_2$ | 12.90 $\pm$ 0.05 | 17.85 $\pm$ 0.04 | 20.44 $\pm$ 0.03 | 17.46 $\pm$ 0.04 | 7.8 $\pm$ 1.2 |
| | $\equiv Al(OH)(OH_2)$ | 12.49 $\pm$ 0.04 | 18.31 $\pm$ 0.03 | 20.14 $\pm$ 0.04 | 17.65 $\pm$ 0.03 | 11.0 $\pm$ 1.0 |
| | $H_3O^+$ | 12.55 $\pm$ 0.02 | 17.05 $\pm$ 0.03 | 19.82 $\pm$ 0.02 | 16.76 $\pm$ 0.03 | — |
| $\perp[110]$ | $\equiv Al^T(OH_2)_2$ | 14.16 $\pm$ 0.06 | 17.96 $\pm$ 0.04 | 20.08 $\pm$ 0.02 | 17.68 $\pm$ 0.04 | 3.0 $\pm$ 1.1 |
| | $\equiv Al^T(OH)$ | 14.19 $\pm$ 0.03 | 18.75 $\pm$ 0.03 | 21.00 $\pm$ 0.05 | 18.37 $\pm$ 0.03 | 14.5 $\pm$ 1.2 |
| | $\equiv Al^T(OH)Al$ | 14.23 $\pm$ 0.03 | 18.18 $\pm$ 0.02 | 20.51 $\pm$ 0.01 | 17.91 $\pm$ 0.02 | 6.9 $\pm$ 0.9 |

| | | | | | |
|---|---|---|---|---|---|
| | $\equiv Al(OH_2)$ | $15.24 \pm 0.05$ | $18.23 \pm 0.02$ | $20.54 \pm 0.02$ | $18.12 \pm 0.03$ | $10.4 \pm 1.0$ |
| | $H_3O^+$ | $13.08 \pm 0.04$ | $17.72 \pm 0.03$ | $19.91 \pm 0.03$ | $17.31 \pm 0.03$ | — |
| $\perp[\overline{1}\overline{1}0]$ | $\equiv Al^T(OH_2)_2$ | $13.84 \pm 0.02$ | $17.48 \pm 0.03$ | $20.74 \pm 0.06$ | $17.42 \pm 0.03$ | $1.5 \pm 1.0$ |
| | $\equiv Al^T(OH)$ | $14.91 \pm 0.03$ | $18.34 \pm 0.06$ | $20.86 \pm 0.03$ | $18.07 \pm 0.05$ | $14.4 \pm 1.3$ |
| | $\equiv Al(OH_2)_2$ | $14.38 \pm 0.07$ | $17.87 \pm 0.02$ | $20.03 \pm 0.04$ | $17.69 \pm 0.03$ | $6.0 \pm 1.0$ |
| | $\equiv Al(OH)(OH_2)$ | $14.32 \pm 0.03$ | $18.16 \pm 0.05$ | $20.49 \pm 0.03$ | $17.91 \pm 0.04$ | $9.7 \pm 1.2$ |
| | $H_3O^+$ | $13.37 \pm 0.03$ | $17.44 \pm 0.03$ | $19.60 \pm 0.02$ | $17.14 \pm 0.03$ | — |

**FIGURE S1.** Comparison of model predictions based on the cis-vacant structure at different edge-specific surface areas (lines) and potentiometric titration data (symbols) for MX80 montmorillonite (bottom) and SWy-2 montmorillonite (top).
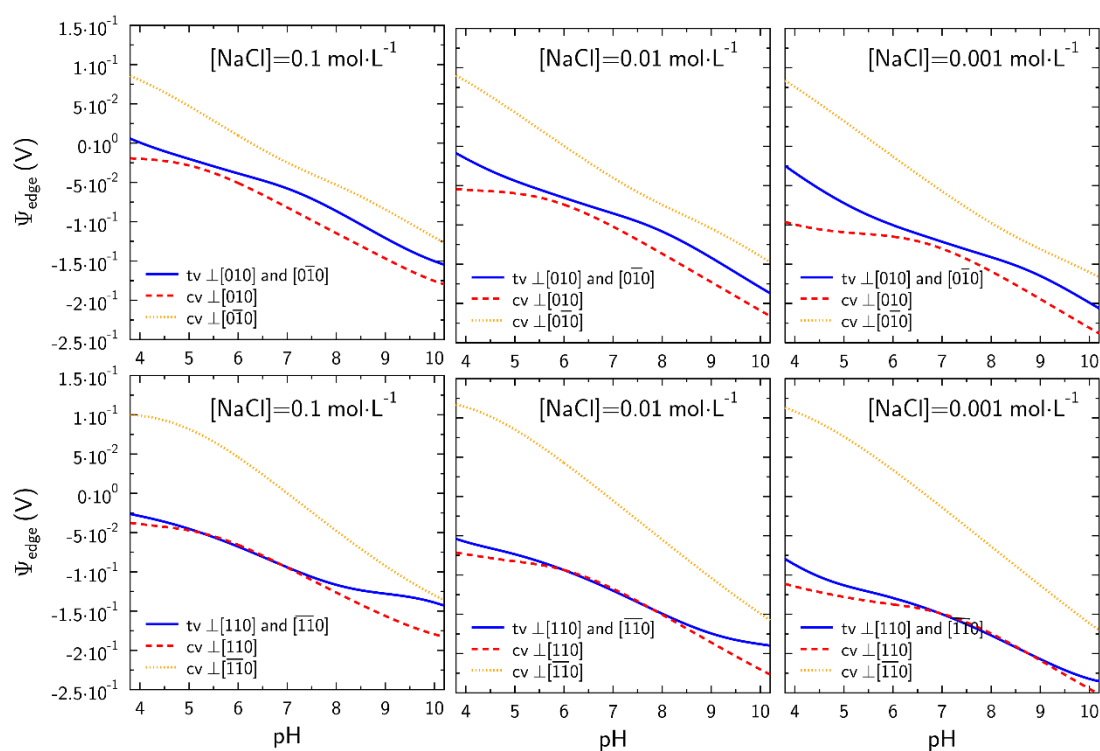
**FIGURE S2.** Swy-2 clay edge surface potentials as a function of surface directions, pH, and ionic strength. Comparison of our cis-vacant model and the trans-vacant model by Tournassat et al. (Tournassat et al. 2016).

**References**

Cheng, J., and Sprik, M. (2012) Alignment of Electronic Energy Levels at Electrochemical Interfaces. Physical Chemistry Chemical Physics, 14, 11245-11267.

Cheng, J., Liu, X., VandeVondele, J., Sulpizi, M., and Sprik, M. (2014) Redox Potentials and Acidity Constants from Density Functional Theory Based Molecular Dynamics. Accounts of Chemical Research, 47, 3522–3529.

Costanzo, F., Sulpizi, M., Valle, R. G. D., and Sprik, M. (2011) The Oxidation of Tyrosine and Tryptophan Studied by a Molecular Dynamics Normal Hydrogen Electrode. The Journal of Chemical Physics, 134, 244508.

Tournassat, C., Davis, J.A., Chiaberge, C., Grangeon, S., and Bourg, I.C. (2016) Modeling the Acid–Base Properties of Montmorillonite Edge Surfaces. Environmental Science & Technology, 50, 13436–13445.

## Supplementary: PHREEQC Files for cis-vacant structure

## Surface_database.phr

```
PHASES
Fix_pHa
H+ = H+
log_k 0

Fix_pHb
H+ = H+
log_k 0

SURFACE_MASTER_SPECIES
### Cation exchange
[Swy]_ex  [Swy]_ex-

### (0100) ###
# (0100) no substitution
[Swy-0100]_wa [Swy-0100]_waOH4+
# (0100) Mg substitution
[Swy-0100]_wb [Swy-0100]_wbOH4
# (0100) Al substitution
[Swy-0100]_wc [Swy-0100]_wcOH4
# (0100) FeII substitution
[Swy-0100]_wd [Swy-0100]_wdOH4
# (0100) FeIII substitution
[Swy-0100]_we [Swy-0100]_weOH4+

### (0101) ###
# (0101) no substitution
[Swy-0101]_wa [Swy-0101]_waOH4+
# (0101) Mg substitution
[Swy-0101]_wb [Swy-0101]_wbOH4
# (0101) Al substitution
[Swy-0101]_wc [Swy-0101]_wcOH4
# (0101) FeII substitution
[Swy-0101]_wd [Swy-0101]_wdOH4
# (0101) FeIII substitution
[Swy-0101]_we [Swy-0101]_weOH5+

### (1100) ###
# (1100) no substitution
[Swy-1100]_wa [Swy-1100]_waOH4+
# (1100) Mg substitution
```

```
[Swy-1100]_wb [Swy-1100]_wbOH4
# (1100) Al substitution
[Swy-1100]_wc [Swy-1100]_wcOH5
# (1100) FeII substitution
[Swy-1100]_wd [Swy-1100]_wdOH4
# (1100) FeIII substitution
[Swy-1100]_we [Swy-1100]_weOH4+

### (1101) ###
# (1101) no substitution
[Swy-1101]_wa [Swy-1101]_waOH4+
# (1101) Mg substitution
[Swy-1101]_wb [Swy-1101]_wbOH4
# (1101) Al substitution
[Swy-1101]_wc [Swy-1101]_wcOH5
# (1101) FeII substitution
[Swy-1101]_wd [Swy-1101]_wdOH4
# (1101) FeIII substitution
[Swy-1101]_we [Swy-1101]_weOH5+

SURFACE_SPECIES
### Cation exchange
[Swy]_ex- = [Swy]_ex-
log_k 0

[Swy]_ex- + Na+ = [Swy]_exNa
log_k 0

[Swy]_ex- + H+ = [Swy]_exH
log_k 0.5

2 [Swy]_ex- + Ca+2 = [Swy]_ex2Ca
log_k 0.5

2 [Swy]_ex- + Mg+2 = [Swy]_ex2Mg
log_k 0.5

3 [Swy]_ex- + Al+3 = [Swy]_ex3Al
log_k 0.5

#### (0100) #######

# (0100) no substitution
[Swy-0100]_waOH4+ = [Swy-0100]_waOH4+
log_k 0
```

```
[Swy-0100]_waOH4+ = [Swy-0100]_waOH3 + H+
log_k 0.8
[Swy-0100]_waOH3 = [Swy-0100]_waOH2- + H+
log_k -5.9
[Swy-0100]_waOH2- = [Swy-0100]_waOH-2 + H+
log_k -6.8
[Swy-0100]_waOH-2 = [Swy-0100]_waO-3 + H+
log_k -6.8


#  (0100) Mg substitution
[Swy-0100]_wbOH4 = [Swy-0100]_wbOH4
log_k 0
[Swy-0100]_wbOH4 = [Swy-0100]_wbOH3-1 + H+
log_k -5.3
[Swy-0100]_wbOH3-1 = [Swy-0100]_wbOH2-2 + H+
log_k -8.8
[Swy-0100]_wbOH2-2 = [Swy-0100]_wbOH-3 + H+
log_k -9.0
[Swy-0100]_wbOH-3 = [Swy-0100]_wbO-4 + H+
log_k -9.0


#  (0100) Al  substitution
[Swy-0100]_wcOH4 = [Swy-0100]_wcOH4
log_k 0
[Swy-0100]_wcOH4 = [Swy-0100]_wcOH3-1 + H+
log_k -2.4
[Swy-0100]_wcOH3-1 = [Swy-0100]_wcOH2-2 + H+
log_k -6.4
[Swy-0100]_wcOH2-2 = [Swy-0100]_wcOH-3 + H+
log_k -6.8
[Swy-0100]_wcOH-3 = [Swy-0100]_wcO-4 + H+
log_k -11.6


#  (0100) FeII substitution
[Swy-0100]_wdOH4 = [Swy-0100]_wdOH4
log_k 0
[Swy-0100]_wdOH4 = [Swy-0100]_wdOH3-1 + H+
log_k -12.2
[Swy-0100]_wdOH3-1 = [Swy-0100]_wdOH2-2 + H+
log_k -12.2
[Swy-0100]_wdOH2-2 = [Swy-0100]_wdOH-3 + H+
log_k -13.2
[Swy-0100]_wdOH-3 = [Swy-0100]_wdO-4 + H+
log_k -17.2
```

```
#  (0100) FeIII substitution
[Swy-0100]_weOH4+ = [Swy-0100]_weOH4+
log_k 0
[Swy-0100]_weOH4+ = [Swy-0100]_weOH3 + H+
log_k 0.4
[Swy-0100]_weOH3 = [Swy-0100]_weOH2- + H+
log_k -5.7
[Swy-0100]_weOH2- = [Swy-0100]_weOH-2 + H+
log_k -8.8
[Swy-0100]_weOH-2 = [Swy-0100]_weO-3 + H+
log_k -8.8


#### (0101) #######


#  (0101) no substitution
[Swy-0101]_waOH4+ = [Swy-0101]_waOH4+
log_k 0
[Swy-0101]_waOH4+ = [Swy-0101]_waOH3 + H+
log_k -5.7
[Swy-0101]_waOH3 = [Swy-0101]_waOH2- + H+
log_k -8.1
[Swy-0101]_waOH2- = [Swy-0101]_waOH-2 + H+
log_k -8.1
[Swy-0101]_waOH-2 = [Swy-0101]_waO-3 + H+
log_k -13.2


#  (0101) Mg substitution
[Swy-0101]_wbOH4 = [Swy-0101]_wbOH4
log_k 0
[Swy-0101]_wbOH4 = [Swy-0101]_wbOH3-1 + H+
log_k -5.9
[Swy-0101]_wbOH3-1 = [Swy-0101]_wbOH2-2 + H+
log_k -10.1
[Swy-0101]_wbOH2-2 = [Swy-0101]_wbOH-3 + H+
log_k -11.0
[Swy-0101]_wbOH-3 = [Swy-0101]_wbO-4 + H+
log_k -11.0


#  (0101) Al substitution
[Swy-0101]_wcOH4 = [Swy-0101]_wcOH4
log_k 0
[Swy-0101]_wcOH4 = [Swy-0101]_wcOH3-1 + H+
log_k -2.9
[Swy-0101]_wcOH3-1 = [Swy-0101]_wcOH2-2 + H+
log_k -7.2
```

```
[Swy-0101]_wcOH2-2 = [Swy-0101]_wcOH-3 + H+
log_k -10.3
[Swy-0101]_wcOH-3 = [Swy-0101]_wcO-4 + H+
log_k -11.0


#  (0101) FeII substitution
[Swy-0101]_wdOH4 = [Swy-0101]_wdOH4
log_k 0
[Swy-0101]_wdOH4 = [Swy-0101]_wdOH3-1 + H+
log_k -10.2
[Swy-0101]_wdOH3-1 = [Swy-0101]_wdOH2-2 + H+
log_k -13.5
[Swy-0101]_wdOH2-2 = [Swy-0101]_wdOH-3 + H+
log_k -13.5
[Swy-0101]_wdOH-3 = [Swy-0101]_wdO-4 + H+
log_k -20.1


#  (0101) FeIII substitution
[Swy-0101]_weOH5+ = [Swy-0101]_weOH5+
log_k 0
[Swy-0101]_weOH5+ = [Swy-0101]_weOH4 + H+
log_k 0.1
[Swy-0101]_weOH4 = [Swy-0101]_weOH3- + H+
log_k -6.1
[Swy-0101]_weOH3- = [Swy-0101]_weOH2-2 + H+
log_k -8.9
[Swy-0101]_weOH2-2 = [Swy-0101]_weOH-3 + H+
log_k -9.0
[Swy-0101]_weOH-3 = [Swy-0101]_weO-4 + H+
log_k -9.0



 #### (1100) #######

#  (1100) no substitution
[Swy-1100]_waOH4+ = [Swy-1100]_waOH4+
log_k 0
[Swy-1100]_waOH4+ = [Swy-1100]_waOH3 + H+
log_k 11.7
[Swy-1100]_waOH3 = [Swy-1100]_waOH2- + H+
log_k -5.6
[Swy-1100]_waOH2- = [Swy-1100]_waOH-2 + H+
log_k -6.3
[Swy-1100]_waOH-2 = [Swy-1100]_waO-3 + H+
log_k -6.8
```

```
#  (1100) Mg substitution
[Swy-1100]_wbOH4 = [Swy-1100]_wbOH4
log_k 0
[Swy-1100]_wbOH4 = [Swy-1100]_wbOH3-1 + H+
log_k 7.1
[Swy-1100]_wbOH3-1 = [Swy-1100]_wbOH2-2 + H+
log_k -8.5
[Swy-1100]_wbOH2-2 = [Swy-1100]_wbOH-3 + H+
log_k -9.1
[Swy-1100]_wbOH-3 = [Swy-1100]_wbO-4 + H+
log_k -10.4

#  (1100) Al substitution
[Swy-1100]_wcOH5 = [Swy-1100]_wcOH5
log_k 0
[Swy-1100]_wcOH5 = [Swy-1100]_wcOH4-1 + H+
log_k -3.0
[Swy-1100]_wcOH4-1 = [Swy-1100]_wcOH3-2 + H+
log_k -6.8
[Swy-1100]_wcOH3-2 = [Swy-1100]_wcOH2-3 + H+
log_k -6.9
[Swy-1100]_wcOH2-3 = [Swy-1100]_wcOH-4 + H+
log_k -10.4
[Swy-1100]_wcOH-4 = [Swy-1100]_wcO-5 + H+
log_k -13.9

#  (1100) FeII substitution
[Swy-1100]_wdOH4 = [Swy-1100]_wdOH4
log_k 0
[Swy-1100]_wdOH4 = [Swy-1100]_wdOH3-1 + H+
log_k 7.7
[Swy-1100]_wdOH3-1 = [Swy-1100]_wdOH2-2 + H+
log_k -10.7
[Swy-1100]_wdOH2-2 = [Swy-1100]_wdOH-3 + H+
log_k -12.7
[Swy-1100]_wdOH-3 = [Swy-1100]_wdO-4 + H+
log_k -13.8

#  (1100)   FeIII substitution
[Swy-1100]_weOH4+ = [Swy-1100]_weOH4+
log_k 0
[Swy-1100]_weOH4+ = [Swy-1100]_weOH3 + H+
log_k 14.6
[Swy-1100]_weOH3 = [Swy-1100]_weOH2- + H+
```

```
log_k -2.7
[Swy-1100]_weOH2-1 = [Swy-1100]_weOH-2 + H+
log_k -7.1
[Swy-1100]_weOH-2 = [Swy-1100]_weO-3 + H+
log_k -8.4


  #### (1101) #######


#  (1101) no substitution
[Swy-1101]_waOH4+ = [Swy-1101]_waOH4+
log_k 0
[Swy-1101]_waOH4+ = [Swy-1101]_waOH3 + H+
log_k -7.0
[Swy-1101]_waOH3 = [Swy-1101]_waOH2- + H+
log_k -7.4
[Swy-1101]_waOH2- = [Swy-1101]_waOH-2 + H+
log_k -8.1



#  (1101) Mg substitution
[Swy-1101]_wbOH4 = [Swy-1101]_wbOH4
log_k 0
[Swy-1101]_wbOH4 = [Swy-1101]_wbOH3-1 + H+
log_k -8.9
[Swy-1101]_wbOH3-1 = [Swy-1101]_wbOH2-2 + H+
log_k -9.2
[Swy-1101]_wbOH2-2 = [Swy-1101]_wbOH-3 + H+
log_k -15.1

#  (1101) Al substitution
[Swy-1101]_wcOH5 = [Swy-1101]_wcOH5
log_k 0
[Swy-1101]_wcOH5 = [Swy-1101]_wcOH4-1 + H+
log_k -1.5
[Swy-1101]_wcOH4-1 = [Swy-1101]_wcOH3-2 + H+
log_k -6.0
[Swy-1101]_wcOH3-2 = [Swy-1101]_wcOH2-3 + H+
log_k -8.1
[Swy-1101]_wcOH2-3 = [Swy-1101]_wcOH-4 + H+
log_k -8.6
[Swy-1101]_wcOH-4 = [Swy-1101]_wcO-5 + H+
log_k -12.4

#  (1101) FeII substitution
[Swy-1101]_wdOH4 = [Swy-1101]_wdOH4
```

```
log_k 0
[Swy-1101]_wdOH4 = [Swy-1101]_wdOH3-1 + H+
log_k -10.9
[Swy-1101]_wdOH3-1 = [Swy-1101]_wdOH2-2 + H+
log_k -11.4
[Swy-1101]_wdOH2-2 = [Swy-1101]_wdOH-3 + H+
log_k -12.3
[Swy-1101]_wdOH-3 = [Swy-1101]_wdO-4 + H+
log_k -19.8


#  (1101) FeIII substitution
[Swy-1101]_weOH5+ = [Swy-1101]_weOH5+
log_k 0
[Swy-1101]_weOH5+ = [Swy-1101]_weOH4 + H+
log_k 7.7
[Swy-1101]_weOH4 = [Swy-1101]_weOH3-1 + H+
log_k -2.3
[Swy-1101]_weOH3-1 = [Swy-1101]_weOH2-2 + H+
log_k -7.9
[Swy-1101]_weOH2-2 = [Swy-1101]_weOH-3 + H+
log_k -8.1
[Swy-1101]_weOH-3 = [Swy-1101]_weO-4 + H+
log_k -8.7
```

## Surface_init.phr

```
SELECTED_OUTPUT
 -reset false
 -file surface_model.phr
# The surface properties of the montmorillonite can be
defined here
USER_PUNCH
 -start
  # Define here your montmorillonite structure
 2  m_arg=4.3             # Solid concentration in g/L
 6  CEC=0.9               # Cation exchange capacity in mol
kg-1
 8   FeIIsub=0.0015            # Fe(II) substitutions in
octahedra (value from 0 to 1)
 9   FeIIIsub=0.056           # Fe(III) substitutions in
octahedra (value from 0 to 1)
 10  Alsub=0.0325                 # Al substitutions in
tetrahedra (value from 0 to 1)
```

```
12 Mgsub=0.0625                      # Mg substitutions in
octahedra (value from 0 to 1): The sum of substitutions
values must NOT exceed 1
13 ratio0100=0.15                 # ratio of edge surfaces
perpendicular to [0100] (value from 0 to 1)
14 ratio1100=0.35                 # ratio of edge surfaces
perpendicular to [1100] (value from 0 to 1)
16 SSA=6              # specific surface area


# Calculation of parameters from information given above.
DO NOT CHANGE
20  ratio0101=0.15              # ratio of edge surfaces
perpendicular to [0101] (value from 0 to 1)
21 ratio1101=0.35   # ratio of edge surfaces perpendicular
to [0101] (value from 0 to 1)
22 Nav=6.022e23 # mol-1
23 SSA_0100=SSA*ratio0100
24 SSA_0101=SSA*ratio0101
25 SSA_1100=SSA*ratio1100
26 SSA_1101=SSA*ratio1101
28 wa=2.06*(1-Mgsub-Alsub-FeIIsub-FeIIIsub) # site nm-2
30 wb=2.06*Mgsub # site nm-2
32 wc=2.06*Alsub # site nm-2
34 wd=2.06*FeIIsub # site nm-2
36 we=2.06*FeIIIsub # site nm-2


# Writing of the input file for calculation
100 sc$ = chr$(59)
150   punch eol$ + 'SOLUTION 1'
155   punch eol$ + 'pH 7'
160   punch eol$ + ' Na 0.3'
170   punch eol$ + ' Cl 0.3 charge'
175   punch eol$ + ' Mg 0.0000005 '
176   punch eol$ + ' Si 0.07 '
180   punch eol$ + 'SURFACE 1'
200   punch eol$ + '-equilibrate 1 '
201   punch eol$ + '-bourg'
202   punch eol$ + '[Swy]_ex ' + str$(CEC*m_arg/1000) + '
750 ' + str$(m_arg)
205   if SSA_0100=0 then goto 237
230        punch    eol$   +   '[Swy-0100]_wa   '   +
str$(wa*1e18/Nav*SSA_0100*m_arg) + ' ' + str$(SSA_0100) +
' ' + str$(m_arg)
232        punch    eol$   +   '[Swy-0100]_wb   '   +
str$(wb*1e18/Nav*SSA_0100*m_arg)
```

```
234        punch   eol$   +    '[Swy-0100]_wc    '    +
str$(wc*1e18/Nav*SSA_0100*m_arg)
235        punch   eol$   +    '[Swy-0100]_wd    '    +
str$(wd*1e18/Nav*SSA_0100*m_arg)
236        punch   eol$   +    '[Swy-0100]_we    '    +
str$(we*1e18/Nav*SSA_0100*m_arg)
237   if SSA_0101=0 then goto 245
238        punch   eol$   +    '[Swy-0101]_wa    '    +
str$(wa*1e18/Nav*SSA_0101*m_arg)  + ' ' + str$(SSA_0101) +
' ' + str$(m_arg)
239        punch   eol$   +    '[Swy-0101]_wb    '    +
str$(wb*1e18/Nav*SSA_0101*m_arg)
240        punch   eol$   +    '[Swy-0101]_wc    '    +
str$(wc*1e18/Nav*SSA_0101*m_arg)
241        punch   eol$   +    '[Swy-0101]_wd    '    +
str$(wd*1e18/Nav*SSA_0101*m_arg)
242        punch   eol$   +    '[Swy-0101]_we    '    +
str$(we*1e18/Nav*SSA_0101*m_arg)
245   if SSA_1100=0 then goto 251
246        punch   eol$   +    '[Swy-1100]_wa    '    +
str$(wa*1e18/Nav*SSA_1100*m_arg)  + ' ' + str$(SSA_1100) +
' ' + str$(m_arg)
247        punch   eol$   +    '[Swy-1100]_wb    '    +
str$(wb*1e18/Nav*SSA_1100*m_arg)
248        punch   eol$   +    '[Swy-1100]_wc    '    +
str$(wc*1e18/Nav*SSA_1100*m_arg)
249        punch   eol$   +    '[Swy-1100]_wd    '    +
str$(wd*1e18/Nav*SSA_1100*m_arg)
250        punch   eol$   +    '[Swy-1100]_we    '    +
str$(we*1e18/Nav*SSA_1100*m_arg)
251   if SSA_1101=0 then goto 300
252        punch   eol$   +    '[Swy-1101]_wa    '    +
str$(wa*1e18/Nav*SSA_1101*m_arg)  + ' ' + str$(SSA_1101) +
' ' + str$(m_arg)
253        punch   eol$   +    '[Swy-1101]_wb    '    +
str$(wb*1e18/Nav*SSA_1101*m_arg)
254        punch   eol$   +    '[Swy-1101]_wc    '    +
str$(wc*1e18/Nav*SSA_1101*m_arg)
255        punch   eol$   +    '[Swy-1101]_wd    '    +
str$(wd*1e18/Nav*SSA_1101*m_arg)
256        punch   eol$   +    '[Swy-1101]_we    '    +
str$(we*1e18/Nav*SSA_1101*m_arg)
300   punch eol$ + ' '
340   punch eol$ + 'SAVE SURFACE 1'
350   punch eol$ + 'END'
```

```
  -end
```

**Input.phr**

```
INCLUDE$ Surface_database.phr
SOLUTION 0
END


INCLUDE$ Surface_init.phr
SOLUTION 0
END



SELECTED_OUTPUT
 -reset false
 -file titration.loop1
USER_PUNCH
 -start
20 Ionic_str=0.1     # in NaCl
30  m_arg=4.3        # g/L
40   FI_stock=0.0003     #  Ionic_str  in  stock  suspension
0.0003 mol/L NaCl in Duc et al. 2005
100
DebyeL=1/((2*96485*96485*1000*FI_stock/(78.3*8.85419e-
12*8.314*298))^0.5)
200 Volume_DL=DebyeL*2*750*1000*m_arg
300 volume_Sol=(1-Volume_DL)
400 dilution_f=1/(1-Volume_DL)



1000 for j = 3 to 6 step 0.1
1020  punch eol$ + 'USE SURFACE 1'
1150  punch eol$ + 'SOLUTION 1'
1155  punch eol$ + '-water ' + str$(volume_Sol)
1160  punch eol$ + ' Na ' + str$(Ionic_str*1000*dilution_f)
1170  punch eol$ + ' Cl ' + str$(Ionic_str*1000*dilution_f)
+ ' charge'
1200  punch eol$ + ' EQUILIBRIUM_PHASES 1'
1210  punch eol$ + ' Fix_pHa ' + str$(-j) + ' HCl 10'
1390  punch eol$ + 'END'
1400 next j
 -end
SOLUTION 0
END
```

```
SELECTED_OUTPUT
 -reset false
 -file titration.loop2
USER_PUNCH
 -start
20 Ionic_str=0.1      # in NaCl
30  m_arg=4.3         # g/L
40   FI_stock=0.0003      # Ionic_str in stock suspension
0.0003 mol/L NaCl in Duc et al. 2005
100
DebyeL=1/((2*96485*96485*1000*FI_stock/(78.3*8.85419e-
12*8.314*298))^0.5)
200 Volume_DL=DebyeL*2*750*1000*m_arg
300 volume_Sol=(1-Volume_DL)
400 dilution_f=1/(1-Volume_DL)

2000 for j = 6 to 11 step 0.1
2020   punch eol$ + 'USE SURFACE 1'
2150   punch eol$ + 'SOLUTION 1'
2155   punch eol$ + '-water ' + str$(volume_Sol)
2160   punch eol$ + ' Na ' + str$(Ionic_str*1000*dilution_f)
+ ' charge'
2170   punch eol$ + ' Cl ' + str$(Ionic_str*1000*dilution_f)
2200   punch eol$ + ' EQUILIBRIUM_PHASES 1'
2210   punch eol$ + ' Fix_pHb ' + str$(-j) + ' NaOH 10'
2390   punch eol$ + 'END'
2400 next j
 -end
SOLUTION 0
END

PRINT ; -reset false ; -selected_out false
SELECTED_OUTPUT
 -reset false
 -high_precision true
 -file titration_swy2_cis_a.prn
 -user_punch true
 -pH true
 -totals Na Cl
 USER_PUNCH
-headings   proton_charge   Potential_0100   DeltaH   Mgconc
Potential_0100   charge0100   Potential_0101   charge0101
Potential_1100 charge1100 Potential_1101 charge1101
-start
```

```
10        PUNCH        SURF("H","[Swy-0100]")+SURF("H","[Swy-
0101]")+SURF("H","[Swy-1100]")+SURF("H","[Swy-
1101]")+SURF("H","[Swy]")
20 PUNCH EDL("Psi","[Swy-0100]")
30 PUNCH -EQUI_DELTA("Fix_pHa")
40 PUNCH TOT("Mg")
50 PUNCH EDL("Psi","[Swy-0100]")
60 PUNCH EDL("Sigma","[Swy-0100]")
70 PUNCH EDL("Psi","[Swy-0101]")
80 PUNCH EDL("Sigma","[Swy-0101]")
90 PUNCH EDL("Psi","[Swy-1100]")
100 PUNCH EDL("Sigma","[Swy-1100]")
110 PUNCH EDL("Psi","[Swy-1101]")
120 PUNCH EDL("Sigma","[Swy-1101]")

-end
PRINT
-reset false
-high_precision true
-selected_out true
INCLUDE$ surface_model.phr
INCLUDE$ titration.loop1


PRINT ; -reset false ; -selected_out false
SELECTED_OUTPUT
 -reset false
 -high_precision true
 -file titration_swy2_cis_b.prn
 -user_punch true
 -pH true
 -totals Na Cl
USER_PUNCH
-headings  proton_charge  Potential_0100  DeltaH  Mgconc
Potential_0100  charge0100  Potential_0101  charge0101
Potential_1100 charge1100 Potential_1101 charge1101
-start
10        PUNCH        SURF("H","[Swy-0100]")+SURF("H","[Swy-
0101]")+SURF("H","[Swy-1100]")+SURF("H","[Swy-
1101]")+SURF("H","[Swy]")
20 PUNCH EDL("Psi","[Swy-0100]")
30 PUNCH EQUI_DELTA("Fix_pHb")
40 PUNCH TOT("Mg")
50 PUNCH EDL("Psi","[Swy-0100]")
60 PUNCH EDL("Sigma","[Swy-0100]")
```

```
70  PUNCH EDL("Psi","[Swy-0101]")
80  PUNCH EDL("Sigma","[Swy-0101]")
90  PUNCH EDL("Psi","[Swy-1100]")
100 PUNCH EDL("Sigma","[Swy-1100]")
110 PUNCH EDL("Psi","[Swy-1101]")
120 PUNCH EDL("Sigma","[Swy-1101]")

-end
PRINT
-reset true
-high_precision true
-selected_out true
INCLUDE$ surface_model.phr
INCLUDE$ titration.loop2


END
```

**Supplementary: PHREEQC Files for trans-vacant structure**

**Surface_database.phr**

```
PHASES
Fix_pHa
H+ = H+
log_k 0

Fix_pHb
H+ = H+
log_k 0

SURFACE_MASTER_SPECIES
### Cation exchange
[Swy]_ex  [Swy]_ex-

### (010) ###
# (010) no substitution
[Swy-010]_wa [Swy-010]_waOH4+
# (010) Mg substitution
[Swy-010]_wb [Swy-010]_wbOH4
# (010) Al substitution
[Swy-010]_wc [Swy-010]_wcOH4
# (010) FeII substitution
[Swy-010]_wd [Swy-010]_wdOH4
# (010) FeIII substitution
[Swy-010]_we [Swy-010]_weOH4+

### (110) ###
# (110) no substitution
[Swy-110]_wa [Swy-110]_waOH4+
# (110) Mg substitution
[Swy-110]_wb [Swy-110]_wbOH4
# (110) Al substitution
[Swy-110]_wc [Swy-110]_wcOH4


SURFACE_SPECIES
### Cation exchange
[Swy]_ex- = [Swy]_ex-
log_k 0

[Swy]_ex- + Na+ = [Swy]_exNa
log_k 0
```

```
[Swy]_ex- + H+ = [Swy]_exH
log_k 0.5


2 [Swy]_ex- + Ca+2 = [Swy]_ex2Ca
log_k 0.5


2 [Swy]_ex- + Mg+2 = [Swy]_ex2Mg
log_k 0.5


3 [Swy]_ex- + Al+3 = [Swy]_ex3Al
log_k 0.5


#### (010) #######

#  (010) no substitution
[Swy-010]_waOH4+ = [Swy-010]_waOH4+
log_k 0
[Swy-010]_waOH4+ = [Swy-010]_waOH3 + H+
log_k -3.1
[Swy-010]_waOH3 = [Swy-010]_waOH2- + H+
log_k -7
[Swy-010]_waOH2- = [Swy-010]_waOH-2 + H+
log_k -7
[Swy-010]_waOH-2 = [Swy-010]_waO-3 + H+
log_k -8.3


#  (010) Mg substitution
[Swy-010]_wbOH4 = [Swy-010]_wbOH4
log_k 0
[Swy-010]_wbOH4 = [Swy-010]_wbOH3-1 + H+
log_k -10.8
[Swy-010]_wbOH3-1 = [Swy-010]_wbOH2-2 + H+
log_k -10.8
[Swy-010]_wbOH2-2 = [Swy-010]_wbOH-3 + H+
log_k -13.2


#  (010) Al substitution
[Swy-010]_wcOH4 = [Swy-010]_wcOH4
log_k 0
[Swy-010]_wcOH4 = [Swy-010]_wcOH3-1 + H+
log_k -4.9
[Swy-010]_wcOH3-1 = [Swy-010]_wcOH2-2 + H+
log_k -7
[Swy-010]_wcOH2-2 = [Swy-010]_wcOH-3 + H+
```

```
log_k -8.5
[Swy-010]_wcOH-3 = [Swy-010]_wcO-4 + H+
log_k -15.1


#  (010) FeIII substitution
[Swy-010]_weOH4+ = [Swy-010]_weOH4+
log_k 0
[Swy-010]_weOH4+ = [Swy-010]_weOH3 + H+
log_k -1.2
[Swy-010]_weOH3 = [Swy-010]_weOH2- + H+
log_k -5.1
[Swy-010]_weOH2- = [Swy-010]_weOH-2 + H+
log_k -8.6
[Swy-010]_weOH-2 = [Swy-010]_weO-3 + H+
log_k -8.6


#  (010) FeII substitution
[Swy-010]_wdOH4 = [Swy-010]_wdOH4
log_k 0
[Swy-010]_wdOH4 = [Swy-010]_wdOH3-1 + H+
log_k -6.6
[Swy-010]_wdOH3-1 = [Swy-010]_wdOH2-2 + H+
log_k -10.2
[Swy-010]_wdOH2-2 = [Swy-010]_wdOH-3 + H+
log_k -11.2
[Swy-010]_wdOH-3 = [Swy-010]_wdO-4 + H+
log_k -11.2


 #### (110) #######

#  (110) no substitution
[Swy-110]_waOH4+ = [Swy-110]_waOH4+
log_k 0
[Swy-110]_waOH4+ = [Swy-110]_waOH3 + H+
log_k -1.7
[Swy-110]_waOH3 = [Swy-110]_waOH2- + H+
log_k -5.5
[Swy-110]_waOH2- = [Swy-110]_waOH-2 + H+
log_k -8.3
[Swy-110]_waOH-2 = [Swy-110]_waO-3 + H+
log_k -8.3


#  (110) Mg substitution
[Swy-110]_wbOH4 = [Swy-110]_wbOH4
log_k 0
```

```
[Swy-110]_wbOH4 = [Swy-110]_wbOH3-1 + H+
log_k -4.2
[Swy-110]_wbOH3-1 = [Swy-110]_wbOH2-2 + H+
log_k -11
[Swy-110]_wbOH2-2 = [Swy-110]_wbOH-3 + H+
log_k -11


#  (110) Al substitution
[Swy-110]_wcOH4 = [Swy-110]_wcOH4
log_k 0
[Swy-110]_wcOH4 = [Swy-110]_wcOH3-1 + H+
log_k 2.4
[Swy-110]_wcOH3-1 = [Swy-110]_wcOH2-2 + H+
log_k -10.2
[Swy-110]_wcOH2-2 = [Swy-110]_wcOH-3 + H+
log_k -12.7
[Swy-110]_wcOH-3 = [Swy-110]_wcO-4 + H+
log_k -17.5
```

**Surface_init.phr**

```
SELECTED_OUTPUT
 -reset false
 -file surface_model.phr
USER_PUNCH
 -start
   # Define here your montmorillonite structure
 2  m_arg=4.3            # Solid concentration in g/L
 6  CEC=0.9              # Cation exchange capacity in mol
kg-1
 8   FeIIsub=0.003            # Fe(II) substitutions in
octahedra (value from 0 to 1)
 9   FeIIIsub=0.112           # Fe(III) substitutions in
octahedra (value from 0 to 1)
 10 Alsub=0.065          # Al substitutions in tetrahedra
(value from 0 to 1)
 12 Mgsub=0.125          # Mg substitutions in octahedra
(value from 0 to 1): The sum of substitutions values must
NOT exceed 1
 13  ratio010=0.3                 # ratio of edge surfaces
perpendicular to [010] (value from 0 to 1)
 16 SSA=16               # specific surface area


 # Calculation of parameters from information given above.
DO NOT CHANGE
```

```
20 ratio110=1-ratio010
22  Nav=6.022e23 # mol-1
24 SSA_010=SSA*ratio010
26 SSA_110=SSA*ratio110
28 wa=2.06*(1-Mgsub-Alsub-FeIIsub-FeIIIsub) # site nm-2
30 wb=2.06*Mgsub # site nm-2
32 wc=2.06*Alsub # site nm-2
34 wd=2.06*FeIIsub # site nm-2
36 we=2.06*FeIIIsub # site nm-2


100 sc$ = chr$(59)
150   punch eol$ + 'SOLUTION 1'
155   punch eol$ + 'pH 7'
160   punch eol$ + ' Na 0.3'
170   punch eol$ + ' Cl 0.3 charge'
175   punch eol$ + ' Mg 0.0000005 '
176   punch eol$ + ' Si 0.07 '
180   punch eol$ + 'SURFACE 1'
200   punch eol$ + '-equilibrate 1 '
201   punch eol$ + '-bourg'
202   punch eol$ + '[Swy]_ex ' + str$(CEC*m_arg/1000) + '
750 ' + str$(m_arg)
205   if SSA_010=0 then goto 239
230           punch   eol$   +   '[Swy-010]_wa   '   +
str$(wa*1e18/Nav*SSA_010*m_arg)  + ' ' + str$(SSA_010) + '
' + str$(m_arg)
232           punch   eol$   +   '[Swy-010]_wb   '   +
str$(wb*1e18/Nav*SSA_010*m_arg)
234           punch   eol$   +   '[Swy-010]_wc   '   +
str$(wc*1e18/Nav*SSA_010*m_arg)
235           punch   eol$   +   '[Swy-010]_wd   '   +
str$(wd*1e18/Nav*SSA_010*m_arg)
236           punch   eol$   +   '[Swy-010]_we   '   +
str$(we*1e18/Nav*SSA_010*m_arg)
239   if SSA_110=0 then goto 300
244           punch   eol$   +   '[Swy-110]_wa   '   +
str$((wa+we)*1e18/Nav*SSA_110*m_arg) + ' ' + str$(SSA_110)
+ ' ' + str$(m_arg)
246           punch   eol$   +   '[Swy-110]_wb   '   +
str$((wb+wd)*1e18/Nav*SSA_110*m_arg)
248           punch   eol$   +   '[Swy-110]_wc   '   +
str$((wc)*1e18/Nav*SSA_110*m_arg)
300   punch eol$ + ' '
340   punch eol$ + 'SAVE SURFACE 1'
350   punch eol$ + 'END'
```

```
  -end
```

**Input.phr**

```
INCLUDE$ Surface_database.phr
SOLUTION 0
END


INCLUDE$ Surface_init.phr
SOLUTION 0
END



SELECTED_OUTPUT
 -reset false
 -file titration.loop1
USER_PUNCH
 -start
20 Ionic_str=0.1 # in NaCl
30  m_arg=4.3        # g/L
40   FI_stock=0.0003     #  Ionic_str  in  stock  suspension
0.0003 mol/L NaCl in Duc et al. 2005
100
DebyeL=1/((2*96485*96485*1000*FI_stock/(78.3*8.85419e-
12*8.314*298))^0.5)
200 Volume_DL=DebyeL*2*750*1000*m_arg
300 volume_Sol=(1-Volume_DL)
400 dilution_f=1/(1-Volume_DL)



1000 for j = 3 to 6 step 0.1
1020   punch eol$ + 'USE SURFACE 1'
1150   punch eol$ + 'SOLUTION 1'
1155   punch eol$ + '-water ' + str$(volume_Sol)
1160   punch eol$ + ' Na ' + str$(Ionic_str*1000*dilution_f)
1170   punch eol$ + ' Cl ' + str$(Ionic_str*1000*dilution_f)
+ ' charge'
1200   punch eol$ + ' EQUILIBRIUM_PHASES 1'
1210   punch eol$ + ' Fix_pHa ' + str$(-j) + ' HCl 10'
1390   punch eol$ + 'END'
1400 next j
 -end
SOLUTION 0
END
```

```
SELECTED_OUTPUT
 -reset false
 -file titration.loop2
USER_PUNCH
 -start
20 Ionic_str=0.1 # in NaCl
30  m_arg=4.3        # g/L
40   FI_stock=0.0003     #  Ionic_str  in  stock  suspension
0.0003 mol/L NaCl in Duc et al. 2005
100
DebyeL=1/((2*96485*96485*1000*FI_stock/(78.3*8.85419e-
12*8.314*298))^0.5)
200 Volume_DL=DebyeL*2*750*1000*m_arg
300 volume_Sol=(1-Volume_DL)
400 dilution_f=1/(1-Volume_DL)

2000 for j = 6 to 11 step 0.1
2020   punch eol$ + 'USE SURFACE 1'
2150   punch eol$ + 'SOLUTION 1'
2155   punch eol$ + '-water ' + str$(volume_Sol)
2160   punch eol$ + ' Na ' + str$(Ionic_str*1000*dilution_f)
+ ' charge'
2170   punch eol$ + ' Cl ' + str$(Ionic_str*1000*dilution_f)
2200   punch eol$ + ' EQUILIBRIUM_PHASES 1'
2210   punch eol$ + ' Fix_pHb ' + str$(-j) + ' NaOH 10'
2390   punch eol$ + 'END'
2400 next j
 -end
SOLUTION 0
END

PRINT ; -reset false ; -selected_out false
SELECTED_OUTPUT
 -reset false
 -high_precision true
 -file titration_swy2_trans_a.prn
 -user_punch true
 -pH true
 -totals Na Cl
 USER_PUNCH
-headings  proton_charge  Edge_potential_010  DeltaH  Mgconc
charge010 charge110 Edge_potential_110
-start
```

```
10        PUNCH        SURF("H","[Swy-010]")+SURF("H","[Swy-
110]")+SURF("H","[Swy]")
20 PUNCH EDL("Psi","[Swy-010]")
30 PUNCH -EQUI_DELTA("Fix_pHa")
40 PUNCH TOT("Mg")
50 PUNCH EDL("Sigma","[Swy-010]")
60 PUNCH EDL("Sigma","[Swy-110]")
70 PUNCH EDL("Psi","[Swy-110]")
-end
PRINT
-reset false
-high_precision true
-selected_out true
INCLUDE$ surface_model.phr
INCLUDE$ titration.loop1


PRINT ; -reset false ; -selected_out false
SELECTED_OUTPUT
 -reset false
 -high_precision true
 -file titration_swy2_trans_b.prn
 -user_punch true
 -pH true
 -totals Na Cl
USER_PUNCH
-headings proton_charge Edge_potential_010 DeltaH Mgconc
charge010 charge110 Edge_potential_110
-start
10        PUNCH        SURF("H","[Swy-010]")+SURF("H","[Swy-
110]")+SURF("H","[Swy]")
20 PUNCH EDL("Psi","[Swy-010]")
30 PUNCH EQUI_DELTA("Fix_pHb")
40 PUNCH TOT("Mg")
50 PUNCH EDL("Sigma","[Swy-010]")
60 PUNCH EDL("Sigma","[Swy-110]")
70 PUNCH EDL("Psi","[Swy-110]")
-end
PRINT
-reset true
-high_precision true
-selected_out true
INCLUDE$ surface_model.phr
INCLUDE$ titration.loop2
END
```

**Supplementary: Example of relevant input sections for CP2K**

```
&MOTION
&CONSTRAINT
     &COLLECTIVE
      TARGET 1.89
      INTERMOLECULAR
      COLVAR 1
      &RESTRAINT
       k 0.1
      &END RESTRAINT
    &END COLLECTIVE
    &COLLECTIVE
      TARGET 2.20
      INTERMOLECULAR
      COLVAR 2
      &RESTRAINT
       k 0.1
      &END RESTRAINT
    &END COLLECTIVE
    &COLLECTIVE
      TARGET 2.10
      INTERMOLECULAR
      COLVAR 3
      &RESTRAINT
       k 0.1
      &END RESTRAINT
    &END COLLECTIVE
  &END CONSTRAINT
  &MD
    COMVEL_TOL 2.0E-6
    &THERMOSTAT
         TYPE NOSE
         REGION GLOBAL
      &NOSE
        LENGTH 3
        YOSHIDA 3
        TIMECON 1000
        MTS 2
      &END NOSE
    &END THERMOSTAT
    ENSEMBLE NVT
    STEPS 30000
    TIMESTEP 0.5
    TEMPERATURE 300.0
```

```
       TEMP_TOL 50.0
   &END MD
&END MOTION
&MIXED
   MIXING_TYPE GENMIX
   &GENERIC
     MIXING_FUNCTION k*X+(1-k)*Y
     VARIABLES X Y
     PARAMETERS k
     VALUES ${LAMBDA}
   &END
&END MIXED
&DFT
   &MGRID
     CUTOFF 400
     NGRIDS 4
   &END MGRID
   &QS
     EPS_DEFAULT 1.0E-12
     EXTRAPOLATION ASPC
     EXTRAPOLATION_ORDER 2
   &END QS
   &SCF
     SCF_GUESS RESTART
     EPS_SCF 1.0E-6
     MAX_SCF 50
     &OUTER_SCF
        EPS_SCF 1.0E-6
        MAX_SCF 5
     &END
     &OT
        MINIMIZER DIIS
     &END
     &PRINT
       &RESTART
         FILENAME =RESTART-${a}.wfn
         BACKUP_COPIES 1
       &END RESTART
     &END PRINT
   &END SCF
   &XC
     &XC_FUNCTIONAL PBE
     &END XC_FUNCTIONAL
     &VDW_POTENTIAL
        DISPERSION_FUNCTIONAL PAIR_POTENTIAL
```

```
            &PAIR_POTENTIAL
               TYPE DFTD3
               PARAMETER_FILE_NAME ./dftd3.dat
               REFERENCE_FUNCTIONAL PBE
                        R_CUTOFF [angstrom] 15
            &END PAIR_POTENTIAL
         &END VDW_POTENTIAL
      &END XC
      PLUS_U_METHOD MULLIKEN
   &END DFT
   &SUBSYS
      &CELL
         ABC [angstrom] 12.4472 10.4405 33.5611
      &END CELL
        &COLVAR
        &DISTANCE
         ATOMS 424 103
        &END DISTANCE
       &END COLVAR
       &COLVAR
        &ANGLE
         ATOMS 424 103 24
        &END ANGLE
       &END COLVAR
          &COLVAR
        &ANGLE
         ATOMS 424 103 165
        &END ANGLE
       &END COLVAR
      &COORD
      atomic coordinates
      &END COORD
      &KIND Si
            BASIS_SET DZVP-MOLOPT-SR-GTH-q4
          POTENTIAL GTH-PBE-q4
         &END KIND
        &KIND O
          BASIS_SET DZVP-MOLOPT-SR-GTH-q6
        POTENTIAL GTH-PBE-q6
          &END KIND
        &KIND Al
         BASIS_SET DZVP-MOLOPT-SR-GTH-q3
          POTENTIAL GTH-PBE-q3
         &END KIND
         &KIND H
```

```
     BASIS_SET DZVP-MOLOPT-SR-GTH-q1
      POTENTIAL GTH-PBE-q1
    &END KIND
    &KIND H1
   BASIS_SET NONE
   POTENTIAL GTH-PBE-q1
  &END KIND
     &KIND Fe
      BASIS_SET DZVP-MOLOPT-SR-GTH-q16
      POTENTIAL GTH-PBE-q16
         MAGNETIZATION 4.0
     &DFT_PLUS_U
         EPS_U_RAMPING 0.001
                 U_RAMPING 0.1
                 INIT_U_RAMPING_EACH_SCF F
                 L 2
                 U_MINUS_J [ev] 3
            &END
           &END KIND
&END SUBSYS
```

## Supplementary: XYZ files for cis-vacant edge surface model

163

[010]&[0$\bar{1}$0]

```
Si   10.076755    9.669243   11.604888
Si   10.287237    9.176820    2.554471
Si   10.181187    9.369385    8.617950
Si   10.225623    6.648329    7.075216
Si   10.133016    7.090444   13.046448
Si   10.242178    6.679640    4.096874
Si    4.603970    8.351872   11.541737
Si    4.772505    8.576896    2.493965
Si    4.701432    8.610399    8.566701
Si    4.710643    0.829958    7.015041
Si    4.432448    0.400472   12.988918
Si    4.762797    0.701582    4.045625
 O   10.669241    9.042704   10.168949
 O   10.837528    0.209055    8.067446
 O   10.462992    0.715463   12.008921
 O   10.640738    0.199662    3.046783
 O    8.284358   10.021077   14.294919
 O    8.511271    9.784155    5.216728
 O    8.524091    9.805969    8.725074
 O    8.407959    9.469581   11.531487
 O    8.623420    9.257499    2.472474
 O   11.064690    6.657304   14.386503
 O   10.864912    1.097983    5.578109
 O   10.707455    8.606850   12.722758
 O   10.898519    7.956619    3.547362
 O   10.580120    8.107435    7.568851
 O    8.380326    7.360034    9.687217
 O    8.487372    7.191997   13.068270
 O    8.585207    7.115911    4.204329
 O    8.561433    6.729425    6.993549
 O    4.098241    8.858281    9.959328
 O    4.052186    7.314215    7.988755
 O    4.411039    6.722432   11.698788
 O    4.399214    7.141795    2.985947
 O    6.521088    7.913618   14.273097
 O    6.462710    8.071219    5.195567
 O    6.353175    8.170266    8.705897
 O    6.241399    8.577105   11.518262
 O    6.432464    8.500810    2.451974
 O    3.723207    6.188679   14.258879
```

```
 O     4.101228     1.110925     5.512644
 O     3.752457     9.346319    12.529669
 O     4.113551     9.842499     3.467348
 O     4.338098     9.831540     7.506030
 O     6.467565     0.262806     9.746399
 O     6.049713     0.160314    13.263013
 O     6.414789     0.261449     4.184822
 O     6.370850     0.754185     6.973381
 H     8.705198     6.844964    10.516775
 H     8.738183     9.716106    15.126791
 H     8.999077     9.293299     5.944451
 H     6.313967     0.871498    10.503218
 H     5.849810     8.315342    14.889396
 H     5.990464     8.592552     5.930234
Al     7.463217     6.261530    14.543554
Al     7.487737     1.105602     5.478920
Al     7.447782     6.375503     8.488341
Al     7.388284     9.058882     9.992722
Al     7.247755     8.945221    13.027272
Al     7.508898     8.903681     3.967265
 O     4.162593     8.608521     1.212431
 O    11.109748     9.327219     1.337741
 O     6.499803     5.897275    16.140182
 H     3.210008     8.592656     0.776988
 H     8.128769     9.279209     1.478921
 H    11.872015     9.356235     0.531327
 H     2.789417     5.854587    14.145141
 H     5.590535     5.780171    15.999662
 H    11.104894     5.670790    14.758466
 H     6.817456     8.566668     1.413456
Si    10.130651     4.399551    11.587695
Si    10.287237     3.958270     2.554471
Si    10.181187     4.150835     8.617950
Si    10.225623     1.429778     7.075216
Si     9.988505     1.831085    13.135057
Si    10.242178     1.460985     4.096874
Si     4.594759     3.049721    11.499747
Si     4.772505     3.358346     2.493965
Si     4.701432     3.391745     8.566701
Si     4.710643     6.048613     7.015041
Si     4.611065     5.592929    12.968419
Si     4.762797     5.920132     4.045625
 O    10.742556     3.822171    10.105798
 O    10.837528     5.427709     8.067446
 O    10.644721     5.922220    11.879975
```

| | | | |
|---|---|---|---|
| O | 10.640738 | 5.418316 | 3.046783 |
| O | 8.391653 | 4.777687 | 14.210276 |
| O | 8.511271 | 4.565605 | 5.216728 |
| O | 8.524091 | 4.587419 | 8.725074 |
| O | 8.443682 | 4.499852 | 11.480240 |
| O | 8.623420 | 4.038845 | 2.472474 |
| O | 10.865285 | 1.440737 | 14.445687 |
| O | 10.864912 | 6.316533 | 5.578109 |
| O | 10.572527 | 3.332044 | 12.758797 |
| O | 10.898519 | 2.737965 | 3.547362 |
| O | 10.580120 | 2.888885 | 7.568851 |
| O | 8.457374 | 2.188764 | 9.813850 |
| O | 8.423643 | 2.103911 | 13.275246 |
| O | 8.585207 | 1.897465 | 4.204329 |
| O | 8.561433 | 1.510875 | 6.993549 |
| O | 4.032644 | 3.590884 | 10.011899 |
| O | 4.052186 | 2.095561 | 7.988755 |
| O | 4.348803 | 1.512649 | 11.816493 |
| O | 4.399214 | 1.923244 | 2.985947 |
| O | 6.272642 | 2.719595 | 14.387826 |
| O | 6.462710 | 2.852773 | 5.195567 |
| O | 6.353175 | 2.951716 | 8.705897 |
| O | 6.308614 | 3.185925 | 11.552979 |
| O | 6.432464 | 3.282259 | 2.451974 |
| O | 3.534258 | 1.054460 | 14.185479 |
| O | 4.101228 | 6.329475 | 5.512644 |
| O | 3.833862 | 4.086647 | 12.566369 |
| O | 4.113551 | 4.623949 | 3.467348 |
| O | 4.338098 | 4.612990 | 7.506030 |
| O | 6.353673 | 5.375107 | 9.685563 |
| O | 6.206049 | 5.346092 | 13.194241 |
| O | 6.414789 | 5.479895 | 4.184822 |
| O | 6.370850 | 5.972735 | 6.973381 |
| H | 8.777641 | 1.605226 | 10.510492 |
| H | 9.053844 | 4.363543 | 14.816988 |
| H | 8.999077 | 4.074748 | 5.944451 |
| H | 5.818941 | 5.742075 | 10.447011 |
| H | 5.461707 | 2.846615 | 14.900968 |
| H | 5.990464 | 3.373897 | 5.930234 |
| Al | 7.194979 | 1.047154 | 14.570005 |
| Al | 7.487737 | 6.324152 | 5.478920 |
| Al | 7.447782 | 1.156848 | 8.488341 |
| Al | 7.517237 | 3.829372 | 10.032067 |
| Al | 7.300158 | 3.751929 | 13.070584 |
| Al | 7.508898 | 3.685131 | 3.967265 |

```
O     4.162593    3.389970    1.212431
O    11.109748    4.108773    1.337741
O     8.479033    1.568696   16.060829
H     3.210008    3.374106    0.776988
H     8.128769    4.060658    1.478921
H    11.872015    4.137789    0.531327
H     2.820411    1.705318   13.988752
H     9.419668    1.613263   15.851208
H     8.233574    2.422347   16.527352
H     6.817456    3.348117    1.413456
O     8.910950    7.006634   15.798637
H     9.824077    6.939524   15.342364
H     9.006172    7.866651   16.275741
O     5.971917    0.514549   15.929568
H    11.284507    0.504321   14.581246
H     5.072981    0.862313   15.842281
H     6.357283    0.523107   16.970401
H     6.759825    5.905520   17.219367
H     3.788305    1.057069   15.265987
O    10.462992   11.152564   12.008921
O    10.640738   10.636762    3.046783
O    10.837528   10.646156    8.067446
O     4.338098   -0.605561    7.506030
O     3.752457   -1.090781   12.529669
O     4.113551   -0.594602    3.467348
O     8.511271   -0.652945    5.216728
O     6.467565   10.699907    9.746399
O     6.049713   10.597414   13.263013
O     8.284358   -0.416023   14.294919
O     8.524091   -0.631131    8.725074
H     8.999077   -1.143802    5.944451
H     6.313967   11.308598   10.503218
H     8.738183   -0.720995   15.126791


160

[110]&[1̄1̄0]

Si    7.508275    4.554668   13.835200
Si    7.571756    1.495080    9.242362
Si    7.734815    9.512653    5.122911
Si    7.449525    7.193296   12.599940
Si    7.601878    4.176513    7.956151
Si    7.647560    1.788144    3.312026
Si    7.371108    9.965249   13.765357
```

| Si | 7.630756 | 6.697894 | 9.503856 |
|----|----------|----------|----------|
| Si | 7.676561 | 4.309525 | 4.860406 |
| Si | 7.482883 | 2.107937 | 12.304029 |
| Si | 7.659260 | 9.380163 | 8.218657 |
| Si | 7.705190 | 6.991794 | 3.574870 |
| Si | 2.027649 | 3.953087 | 12.470374 |
| Si | 2.060136 | 1.843270 | 7.812752 |
| Si | 2.223194 | 9.861053 | 3.693301 |
| Si | 2.019309 | 6.557156 | 10.987452 |
| Si | 2.126604 | 4.439927 | 6.394276 |
| Si | 2.172410 | 2.051558 | 1.750151 |
| Si | 1.998771 | 9.205076 | 12.385683 |
| Si | 2.119011 | 7.045876 | 8.074583 |
| Si | 2.164817 | 4.657612 | 3.430458 |
| Si | 2.040221 | 1.315399 | 10.764761 |
| Si | 2.184608 | 9.643577 | 6.656782 |
| Si | 2.230414 | 7.255103 | 2.012994 |
| O | 8.141340 | 6.029702 | 13.523771 |
| O | 8.221625 | 2.618582 | 8.386688 |
| O | 8.184905 | 0.238357 | 4.095495 |
| O | 8.016121 | 8.666555 | 12.989986 |
| O | 8.249630 | 5.186841 | 8.952526 |
| O | 8.295312 | 2.798367 | 4.308401 |
| O | 8.045621 | 4.171711 | 15.351525 |
| O | 7.981145 | 1.786683 | 10.797490 |
| O | 8.073628 | 9.702983 | 6.695584 |
| O | 5.929224 | 2.099376 | 15.085983 |
| O | 5.792802 | 10.354157 | 10.909509 |
| O | 5.941671 | 7.587112 | 6.559607 |
| O | 5.987476 | 5.198743 | 1.915820 |
| O | 5.825289 | 7.214386 | 12.665397 |
| O | 5.939679 | 4.450994 | 8.130930 |
| O | 5.985485 | 2.062625 | 3.486805 |
| O | 5.894620 | 4.579726 | 13.717107 |
| O | 5.906819 | 1.776973 | 9.015622 |
| O | 6.069752 | 9.794547 | 4.896171 |
| O | 8.102380 | 0.846098 | 13.252155 |
| O | 8.279504 | 7.822127 | 8.648181 |
| O | 8.325310 | 5.433654 | 4.004394 |
| O | 8.188141 | 3.435342 | 12.879653 |
| O | 8.107359 | 10.432252 | 9.044301 |
| O | 8.352942 | 8.001286 | 4.571919 |
| O | 7.756472 | 9.760406 | 15.371096 |
| O | 8.042011 | 6.947109 | 11.038401 |
| O | 8.016246 | 4.498290 | 6.432403 |

| | | | |
|---|---|---|---|
| O | 5.768530 | 7.631066 | 15.313398 |
| O | 5.665218 | 4.913821 | 10.993863 |
| O | 5.883542 | 2.383566 | 6.297439 |
| O | 5.846948 | 0.003445 | 2.006246 |
| O | 5.811100 | 2.252329 | 12.328661 |
| O | 5.997434 | 9.654435 | 8.393436 |
| O | 6.043240 | 7.266066 | 3.749648 |
| O | 5.788197 | 9.852387 | 13.670881 |
| O | 5.965445 | 6.979683 | 9.277790 |
| O | 6.011251 | 4.591314 | 4.634003 |
| O | 1.486569 | 5.046416 | 11.357255 |
| O | 1.454331 | 3.381260 | 7.302586 |
| O | 1.500137 | 0.992892 | 2.658461 |
| O | 1.354504 | 9.571537 | 13.881088 |
| O | 1.444995 | 6.819944 | 9.527137 |
| O | 1.490801 | 4.431679 | 4.883012 |
| O | 1.489432 | 2.545707 | 11.642704 |
| O | 1.697300 | 0.783873 | 6.694234 |
| O | 1.860359 | 8.801446 | 2.574783 |
| O | 3.702544 | 1.397670 | 13.446503 |
| O | 3.861619 | 9.421925 | 9.346960 |
| O | 3.907301 | 7.033661 | 4.702834 |
| O | 3.555418 | 9.163314 | 15.182820 |
| O | 3.668936 | 6.400444 | 10.952362 |
| O | 3.782331 | 3.943064 | 6.324769 |
| O | 3.828136 | 1.554695 | 1.680982 |
| O | 3.676156 | 4.073361 | 12.647514 |
| O | 3.721588 | 1.674552 | 7.988206 |
| O | 3.884647 | 9.692230 | 3.869092 |
| O | 1.571583 | 10.300284 | 11.228363 |
| O | 1.513082 | 8.583970 | 7.564755 |
| O | 1.558763 | 6.195602 | 2.920967 |
| O | 1.496402 | 4.124624 | 13.992096 |
| O | 1.385747 | 1.617651 | 9.264968 |
| O | 1.548805 | 9.635433 | 5.145518 |
| O | 1.550921 | 7.752176 | 11.967968 |
| O | 1.756798 | 5.986583 | 6.955390 |
| O | 1.802479 | 3.598214 | 2.311603 |
| O | 3.753702 | 6.682338 | 13.790662 |
| O | 3.803491 | 4.218589 | 9.084117 |
| O | 3.849297 | 1.830324 | 4.440329 |
| O | 3.705531 | 3.611787 | 15.209813 |
| O | 3.687732 | 1.283764 | 10.715836 |
| O | 3.840584 | 9.146818 | 6.586937 |
| O | 3.886265 | 6.758449 | 1.943150 |

| | | | |
|----|----------|----------|-----------|
| O  | 3.644291 | 9.259888 | 12.548315 |
| O  | 3.780712 | 6.877262 | 8.250374  |
| O  | 3.826394 | 4.488998 | 3.606586  |
| O  | 8.529568 | 1.796706 | 1.889164  |
| O  | 8.468079 | 7.154144 | 2.092286  |
| O  | 1.666805 | 2.196890 | 0.393421  |
| O  | 1.672904 | 7.671679 | 0.467314  |
| H  | 5.317568 | 7.125641 | 16.063125 |
| H  | 5.706294 | 3.964780 | 11.119717 |
| H  | 6.374958 | 1.457181 | 6.258637  |
| H  | 6.682279 | 2.581414 | 15.430817 |
| H  | 5.913665 | 9.453768 | 10.977330 |
| H  | 6.430970 | 6.659891 | 6.521142  |
| H  | 6.476776 | 4.271522 | 1.877355  |
| H  | 3.256934 | 6.312117 | 14.551861 |
| H  | 3.326390 | 3.845863 | 9.799765  |
| H  | 3.372195 | 1.457494 | 5.155978  |
| H  | 3.432440 | 0.801622 | 14.181047 |
| H  | 3.381406 | 9.050556 | 10.060922 |
| H  | 3.427212 | 6.662187 | 5.416796  |
| H  | 8.611346 | 9.225644 | 15.578604 |
| H  | 7.847960 | 4.697599 | 16.133307 |
| H  | 0.334581 | 9.873476 | 13.916179 |
| H  | 0.467143 | 4.232996 | 14.248192 |
| H  | 3.857636 | 9.301650 | 16.122510 |
| H  | 2.807964 | 3.933354 | 14.967890 |
| H  | 9.354569 | 1.694911 | 0.547618  |
| H  | 9.217774 | 7.062477 | 0.697766  |
| H  | 0.569957 | 2.464271 | 0.146436  |
| H  | 0.690571 | 8.286207 | 0.147786  |
| H  | 6.117550 | 8.437595 | 15.718291 |
| H  | 6.249366 | 1.191470 | 14.985772 |
| H  | 5.658621 | 5.381243 | 0.861410  |
| H  | 5.420133 | 0.210272 | 0.987265  |
| Al | 4.748358 | 0.530377 | 12.178850 |
| Al | 4.901085 | 8.262821 | 8.078631  |
| Al | 4.946891 | 5.874452 | 3.435181  |
| Al | 4.590403 | 8.304583 | 13.822380 |
| Al | 4.848309 | 5.592663 | 9.447508  |
| Al | 4.893990 | 3.204189 | 4.803720  |
| Al | 4.639694 | 5.662092 | 12.482858 |
| Al | 4.843205 | 3.059693 | 7.816126  |
| Al | 4.888887 | 0.671324 | 3.172338  |
| Al | 4.864241 | 2.896090 | 13.894923 |
| Al | 4.788687 | 0.389639 | 9.186014  |

```
Al   4.951745    8.407317    5.066564
 H   2.621505    9.341847   15.012090
 H   3.764656    3.824982   16.223058
 H   4.586918    1.561272    0.870858
 H   4.813332    6.765549    1.332774
 H   7.003541    9.765209   16.186618
 O   8.107359   -0.008248    9.044301
 O   8.184905   10.678857    4.095495
 O   8.102380   11.286598   13.252155
 O   1.571583   -0.140216   11.228363
 O   1.697300   11.224373    6.694234
 O   5.792802   -0.086343   10.909509
 O   3.644291   -1.180612   12.548315
 O   3.884647   -0.748271    3.869092
 O   5.997434   -0.786066    8.393436
 O   3.861619   -1.018575    9.346960
 H   5.913665   -0.986732   10.977330
 H   3.381406   -1.389944   10.060922
```