Supporting Information

Mechanistic characterization of zeolite-catalyzed aromatic electrophilic substitution at realistic operating conditions

Massimo Bocus¹, Louis Vanduyfhuys¹, Frank De Proft², Bert M. Weckhuysen³ and Veronique Van Speybroeck¹,*

¹Center for Molecular Modeling, Ghent University, Technologiepark 46, 9052 Zwijnaarde (Belgium)
²Eenheid Algemene Chemie (ALGC), Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel (Belgium)
³Inorganic Chemistry and Catalysis Group, Debye Institute for Nanomaterials Science, Utrecht University, Universiteitsweg 99, 3584 CG Utrecht (The Netherlands)

*Corresponding author: Veronique.Vanspeybroeck@Ugent.be
Content
S1. Catalyst model ........................................................................................................3
S2. Advance sampling techniques ..................................................................................4
  S2.1. Collective variables definition ..............................................................................4
  S2.2. Umbrella sampling simulations details .................................................................5
  S2.3. 2-D expansion of the US FESs. ............................................................................8
  S2.4. Walls applied during the simulations ...................................................................10
  S2.5. Error estimation ..................................................................................................12
S3. Benzene alkylation mechanism ...................................................................................14
  S3.1. 2D US FESs .........................................................................................................14
  S3.2. Mobility analysis ..................................................................................................18
S4. Comparison with previous literature reports ..............................................................20
S5. Insights into the role of water in the protonation reaction ..........................................21
  S5.1. Cluster size involved in the reaction ....................................................................21
  S5.2. Expansion of the reaction profile in two dimensions ............................................21
S6. References ...............................................................................................................23
S1. Catalyst model

Figure S1. Reference H-ZSM-5 unit cell, used as starting point to build the initial structures for the molecular dynamics simulations, as seen along the three cell vectors. The approximate unit cell lengths are shown. The Al tetrahedron and the Brønsted acid site location (O_{zeo,1}) are highlighted with larger spheres.
S2. Advance sampling techniques

S2.1. Collective variables definition

Figure S2. Schematic picture showing the passages from a 2-dimensional to a 1-dimensional free energy surface when one of the coordination numbers increases of one unit while the other decreases of one unit. The initial 2D profile (a) can be projected on the difference between the two coordination numbers (red line) to obtain a final 1D free energy profile (b). The values of the CNs in the figure have been chosen to mimic the ones of TS3 but will change based on the number of atoms involved in the coordination number definition for different reactions.

Figure S3. Schematic picture showing the passages from a 2-dimensional to a 1-dimensional free energy surface when one of the partial collective variables (CN1-CN2 in this case) varies twice as the other (CN3). Firstly, the initial 2D profile (a) is squeezed along the horizontal axis by multiplying by 0.5. In this way, both the partial CVs vary of one unit (b). Then, the profile is projected on the sum of the two partial CVs (red line in scheme b) giving the final 1D profile (c). The values of the (partial) CVs in the figure have been chosen to mimic the ones actually encountered in our simulations, but in principle any 2D CV space in which one CV varies twice as the other can be projected in 1D with a similar procedure.
S2.2. Umbrella sampling simulations details

Table S1. Summary of the collective variables adopted for the simulation of all the transition states of benzene ethylation in the US simulations (see also Figure S4). The CNs used to define the CV are also reported, with their relative $r_0$ value (Å). The approximate values that the CV and the CNs are supposed to assume in the reactant (R) and product (P) state of each reaction are also shown (according to the reaction arrows of Figure 1 in the main manuscript). Each CN can include the oxygen atoms in the first-coordination sphere of the Al defect in the zeolite framework (O$_{zeo}$), the aliphatic α-carbon of ethanol and SES or both the aliphatic carbon atoms of ethene (C$_a$), the oxygen of the hydroxyl group of ethanol (O$_{al}$) the six aromatic carbons of benzene (C$_{ar}$) or their respective hydrogens (H$_{ar}$) and the other non-aromatic hydrogen atoms (H$_{al}$).

| TS    | CV                                      | CV(R) | CV(P) | CN                                      | CN(R) | CN(P) | $r_0$ |
|-------|-----------------------------------------|-------|-------|-----------------------------------------|-------|-------|-------|
| TS0$^{OH}$ | CN(O$_{zeo}$;C$_a$)-CN(O$_{al}$;C$_a$) | 1     | -1    | CN(O$_{zeo}$;C$_{ar}$)                  | 0     | 1     | 2.0   |
|       |                                         |       |       | CN(O$_{ar}$;C$_{ar}$)                   | 1     | 0     | 2.0   |
|       | CN(C$_{ar}$;O$_{zeo}$)+0.5(CN(C$_{ar}$;H$_{ar}$)-CN(O$_{zeo}$;H$_{al}$)) | 1.5   | 3.5   | CN(C$_{ar}$;O$_{zeo}$)                  | 0     | 1     | 2.2   |
|       |                                         |       |       | CN(C$_{ar}$;H$_{ar}$)                   | 4     | 5     | 1.5   |
|       |                                         |       |       | CN(O$_{zeo}$;H$_{al}$)                  | 1     | 0     | 1.5   |
| TS0$^*$ | CN(C$_{ar}$;O$_{zeo}$)-0.5(CN(C$_{ar}$;H$_{ar}$)-CN(O$_{zeo}$;H$_{al}$)) | -1.5  | -1.5  | CN(C$_{ar}$;O$_{zeo}$)                  | 0     | 1     | 2.2   |
|       |                                         |       |       | CN(C$_{ar}$;H$_{ar}$)                   | 4     | 5     | 1.5   |
|       |                                         |       |       | CN(O$_{zeo}$;H$_{al}$)                  | 1     | 0     | 1.5   |
| TS1   | CN(C$_{ar}$;C$_{al}$)-CN(O$_{zeo}$;C$_{al}$) | -1    | 1     | CN(C$_{ar}$;C$_{al}$)                   | 0     | 1     | 1.9   |
|       |                                         |       |       | CN(O$_{al}$;C$_{al}$)                   | 1     | 0     | 1.9   |
| TS2$^{OH}$ | CN(C$_{ar}$;C$_{al}$)-CN(O$_{al}$;C$_{al}$) | -1    | 1     | CN(C$_{ar}$;C$_{al}$)                   | 0     | 1     | 1.9   |
|       |                                         |       |       | CN(O$_{al}$;C$_{al}$)                   | 1     | 0     | 1.9   |
| TS2$^*$ | CN(C$_{ar}$;C$_{al}$)+0.5(CN(C$_{ar}$;H$_{al}$)-CN(O$_{zeo}$;H$_{al}$)) | 1.5   | 3.5   | CN(C$_{ar}$;C$_{al}$)                   | 0     | 1     | 1.8   |
|       |                                         |       |       | CN(C$_{ar}$;H$_{al}$)                   | 4     | 5     | 1.5   |
|       |                                         |       |       | CN(O$_{zeo}$;H$_{al}$)                  | 1     | 0     | 1.5   |
| TS3$^i$ | CN(C$_{ar}$;H$_{al}$)-CN(O$_{zeo}$;H$_{al}$) | 1     | -1    | CN(C$_{ar}$;H$_{al}$)                   | 1     | 0     | 1.5   |
|       |                                         |       |       | CN(O$_{zeo}$;H$_{al}$)                  | 0     | 1     | 1.5   |
| TS3$^p$ | CN(C$_{ar}$;H$_{al}$)-CN(O$_{zeo}$;H$_{al}$)$^b$ | 0     | 2     | CN(C$_{ar}$;H$_{al}$)                   | 1     | 2     | 1.5   |
|       |                                         |       |       | CN(O$_{zeo}$;H$_{al}$)                  | 1     | 0     | 1.5   |
| TS3$^{p,r,ω=0}$ | CN(C$_{ar}$;H$_{al}$)$^b$ | 1     | 2     | -                                         | -     | -     | 1.5   |

$^a$H$_{al}$ includes, in this case, the para H of ethylbenzene and the BAS.
$^b$H includes the BAS, all water protons and the proton initially bonded to the ethylbenzene para carbon.

Figure S4. CVs used in the umbrella sampling simulations for the formation of a surface ethoxide species from ethene (TS0$^*$) or ethanol (TS0$^{OH}$), for the direct alkylation of benzene with ethanol (TS2$^{OH}$), for the protonation of ethylbenzene on a specific position (TS3$^i$, ipso and para carbon are considered in this work; in the para case, also the proton already linked to the carbon is considered in the CN as it becomes indistinguishable from the BAS once the protonation has occurred) and for the para protonation of ethylbenzene in the presence of water (TS3$^{p,r}$ with ω≠0, 3 water molecules are shown in the figure as an example). TS0$^*$ required some extra 2-dimensional umbrellas to improve the sampling of the transition state region, therefore two separate CVs are reported (see Section S2.3 for more information). For the nomenclature of the group of atoms see also the caption of Table S1.
Each umbrella sampling simulation consists of multiple umbrella windows, distributed between reactants and products along the collective variable. In each window, a quadratic bias potential in the form $V_i(CV) = \frac{\kappa_i}{2} (CV - CV_0)^2$ is applied, being $CV_0$ the center of the umbrella and $\kappa$ its force constant. These two parameters are reported for each of the investigated transition states in the following tables. In general, low $\kappa$ values were chosen to sample flat regions of the FES (the minima, in particular) while higher $\kappa$ values were chosen for region with a large or unknown gradient of the FES.

- **TS0**

  **Table S2.** Parameters for the umbrella sampling simulation of TS0. Some of the umbrellas (14-19) are two-dimensional, as better described in the following section (Figure S5).

| $U_{num}$ | $CV_0$ | $CV_0'$ | $\kappa$ (kJ·mol$^{-1}$) | $\kappa'$ (kJ·mol$^{-1}$) |
|-----------|--------|--------|------------------------|---------------------|
| 1         | 1.5    | -      | 500                    | 0                   |
| 2         | 1.7    | -      | 1000                   | 0                   |
| 3         | 1.9    | -      | 1000                   | 0                   |
| 4         | 2.1    | -      | 1000                   | 0                   |
| 5         | 2.3    | -      | 1000                   | 0                   |
| 6         | 2.5    | -      | 1000                   | 0                   |
| 7         | 2.7    | -      | 1000                   | 0                   |
| 8         | 2.8    | -      | 1500                   | 0                   |
| 9         | 2.9    | -      | 1000                   | 0                   |
| 10        | 3.1    | -      | 1000                   | 0                   |
| 11        | 3.3    | -      | 1000                   | 0                   |
| 12        | 3.5    | -      | 1000                   | 0                   |
| 13        | 3.7    | -      | 500                    | 0                   |
| 14        | 2.6    | -1.7   | 1000                   | 1000                |
| 15        | 2.6    | -1.9   | 1000                   | 1000                |
| 16        | 2.7    | -2.2   | 1000                   | 1000                |
| 17        | 2.75   | -1.65  | 1000                   | 1000                |
| 18        | 2.75   | -1.8   | 1000                   | 1000                |
| 19        | 2.85   | -1.4   | 2000                   | 1000                |

- **TS0$^{\text{OH}}$**

  **Table S3.** Parameters for the umbrella sampling simulation of TS0$^{\text{OH}}$.

| $U_{num}$ | $CV_0$ | $\kappa$ (kJ·mol$^{-1}$) |
|-----------|--------|------------------------|
| 1         | -0.8   | 500                    |
| 2         | -0.6   | 1000                   |
| 3         | -0.4   | 1000                   |
| 4         | -0.2   | 1000                   |
| 5         | 0.0    | 1000                   |
| 6         | 0.2    | 1000                   |
| 7         | 0.4    | 1000                   |
| 8         | 0.6    | 1000                   |
| 9         | 0.8    | 500                    |
| 10        | 1.0    | 500                    |

- **TS1**

  **Table S4.** Parameters for the umbrella sampling simulation of TS1.

| $U_{num}$ | $CV_0$ | $\kappa$ (kJ·mol$^{-1}$) |
|-----------|--------|------------------------|
| 1         | -0.8   | 500                    |
| 2         | -0.6   | 1000                   |
| 3         | -0.4   | 1000                   |
| 4         | -0.2   | 1000                   |
| 5         | 0.0    | 1000                   |
| 6         | 0.2    | 1000                   |
| 7         | 0.4    | 1000                   |
| 8         | 0.6    | 1000                   |
| 9         | 0.8    | 1000                   |
| 10        | 1.0    | 500                    |
**TS2**

Table S5. Parameters for the umbrella sampling simulation of TS2.

| U<sub>num</sub> | CV<sub>0</sub> | κ (kJ·mol<sup>-1</sup>) |
|-----------------|----------------|--------------------------|
| 1               | 1.4            | 500                      |
| 2               | 1.6            | 1000                     |
| 3               | 1.8            | 1000                     |
| 4               | 2.0            | 1000                     |
| 5               | 2.2            | 1000                     |
| 6               | 2.4            | 1000                     |
| 7               | 2.5            | 1500                     |
| 8               | 2.6            | 1000                     |
| 9               | 2.8            | 1000                     |
| 10              | 3.0            | 1000                     |
| 11              | 3.2            | 1000                     |
| 12              | 3.4            | 500                      |

**TS2<sup>OH</sup>**

Table S6. Parameters for the umbrella sampling simulation of TS2<sup>OH</sup>.

| U<sub>num</sub> | CV<sub>0</sub> | κ (kJ·mol<sup>-1</sup>) |
|-----------------|----------------|--------------------------|
| 1               | -0.8           | 500                      |
| 2               | -0.6           | 1000                     |
| 3               | -0.4           | 1000                     |
| 4               | -0.2           | 1000                     |
| 5               | 0.0            | 1000                     |
| 6               | 0.2            | 1000                     |
| 7               | 0.4            | 1000                     |
| 8               | 0.6            | 1000                     |
| 9               | 0.8            | 500                      |
| 10              | 1.0            | 500                      |

**TS3<sub>i</sub>**

Table S7. Parameters for the umbrella sampling simulation of TS3.<sub>i</sub>.

| U<sub>num</sub> | CV<sub>0</sub> | κ (kJ·mol<sup>-1</sup>) |
|-----------------|----------------|--------------------------|
| 1               | -0.9           | 500                      |
| 2               | -0.7           | 1000                     |
| 3               | -0.5           | 1000                     |
| 4               | -0.3           | 1000                     |
| 5               | -0.2           | 500                      |
| 6               | -0.1           | 1000                     |
| 7               | 0.0            | 500                      |
| 8               | 0.1            | 1000                     |
| 9               | 0.2            | 500                      |
| 10              | 0.3            | 1000                     |
| 11              | 0.4            | 500                      |
| 12              | 0.5            | 1000                     |
| 13              | 0.7            | 1000                     |
| 14              | 0.9            | 500                      |

**TS3<sub>p<sup>nw</sup></sub>**

Table S8. Parameters for the umbrella sampling simulation of TS3<sub>p<sup>nw</sup></sub> (n=1,3,6).

| U<sub>num</sub> | CV<sub>0</sub> | κ (kJ·mol<sup>-1</sup>) |
|-----------------|----------------|--------------------------|
| 1               | 0.9            | 1000                     |
| 2               | 1.0            | 1000                     |
| 3               | 1.1            | 1000                     |
| 4               | 1.2            | 1000                     |
| 5               | 1.3            | 1000                     |
| 6               | 1.4            | 1000                     |
| 7               | 1.5            | 1000                     |
| 8               | 1.55           | 1500                     |
| 9               | 1.6            | 1000                     |
| 10              | 1.7            | 1000                     |
2-D expansion of the US FESs.

While the linear combinations of coordination numbers used as collective variables allows to drastically reduce the computational time, making the reaction profile one-dimensional, the risk of poorly exploring important regions of the higher-dimensional FES – function of the separate CNs – increases with the number of degrees of freedom that are projected on the linear combination. Luckily, by applying some simple statistical analysis, it is possible to expand a one-dimensional profile as a function of 2 or more new collective variables. In formula:

\[
F(CV_1, CV_2) = -k_B T \ln \left( \int_{-\infty}^{+\infty} p(CV_1, CV_2|CV) \exp \left( -\frac{F(CV)}{k_B T} \right) dCV \right)
\]  

(S1)

Where \(k_B\) is Boltzmann’s constant, \(T\) the temperature and \(F\) the free energy. This formula allows to expand a 1-dimensional FES function of an initial collective variable (\(F(CV)\)) in a new 2-dimensional FES function of two new collective variables (\(F(CV_1, CV_2)\)). To do so, the conditional probability \(p(CV_1, CV_2|CV)\) is used, which encodes the likelihood of observing the values \(CV_1\) and \(CV_2\) when, during the simulation, the value \(CV\) is visited. In practice, the 2-dimensional free energy \(F(CV_1, CV_2)\) will be low when \(CV_1\) and \(CV_2\) are very likely to occur (high \(p(CV_1, CV_2|CV)\)) in correspondence of a \(CV\) value of low free energy (high \(\exp(-F(CV)/k_B T)\)) and vice versa. This formula was used to expand all the US FESs in two dimensions and check that the path connecting reactant and products was well-sampled (see Section S3.1) but also to convert the FES of TS3_Nw as a function of two new collective variables to track the proton position (see Section S5.2).

While in most of the simulations the whole reaction path resulted to be well-sampled, TS0° presented obvious regions around the transition states that were insufficiently visited by the 1D umbrellas, when the free energy was expanded in function of \(x=0.5[CN(C_\text{al};H_\text{al})-CN(O_\text{zeo};H_\text{al})]\) and \(y=CN(Cl_\text{al};O_\text{zeo})\). Such two dimensional FES would correspond to the schematic one shown in Figure S3b. Projecting the free energy on the sum of the two partial collective variables (\(x+y\), as done in Figure S3c) can also be seen as an initial 45° rotation of the FES together with some stretching, creating a new 2D space defined by the horizontal axis \(x+y\) and its orthogonal direction \(x-y\) (Figure S5a), followed by a projection on the horizontal axis. In this intermediate 2-dimensional space, the 1D umbrellas acting on \(x+y\) can be seen as 2D, with a force constant of zero in the orthogonal \(x-y\) direction (Figure S5b). This allows to introduce proper 2-dimensional umbrellas, with force constants different from zero in both the \(x+y\) and \(x-y\) directions, to sample difficult regions of the FES and subsequently combining them with the original 1-dimensional umbrellas with 2-dimensional WHAM analysis. The obtained 2-dimensional FES can then be projected on the horizontal axis to retrieve a final 1-dimensional free energy profile, using the simple formula:

\[
F(CV_1) = -k_B T \ln \left( \int_{-\infty}^{+\infty} \exp \left( -\frac{F(CV_1, CV_2)}{k_B T} \right) dCV_2 \right)
\]  

(S2)
Figure S5. When a 2-dimensional free energy surface is projected on the sum of the two collective variables of which it is function (see also Figure S3), the process can also be visualized as a 45° degrees rotation of the space (together with some stretching, as this is not a rigid rotation) followed by a projection on the new horizontal axis, x+y (a). In this new space, the vertical direction will consist of the difference between the two CVs instead of their sum (x-y). 1-dimensional umbrellas acting on the x+y direction can also be seen as 2-dimensional in this new space, with a force constant of 0 in the x-y direction (b) and can then be combined with ‘regular’ 2D umbrellas (force constant ≠0 in both directions) using a 2-dimensional WHAM analysis.
S2.4. Walls applied during the simulations

By biasing along the two CVs chosen to sample the FES of TS 1 and TS 2 with MTD, the system could also evolve towards an undesired (meta)stable state while approaching the high energies of the transition state. Moreover, the diffusion of reactant/products also has to be limited in order to make barrier recrossings more likely to occur and improve the FES convergence. For this reason, a series of quadratic walls was applied to specific collective variables in the MTD simulations. A quadratic wall has the form $V(CV) = 0.5 \cdot \kappa (CV - CV_0)^2$ if $CV > CV_0$ and 0 elsewhere. A list of the walls applied in the MTD simulations, together with their parameters and purposes, is reported in Table S10.

| CV            | TS     | upper value $(CV_0)$ | $\kappa$ (kJ·mol$^{-1}$) | Purpose                                         |
|---------------|--------|----------------------|---------------------------|------------------------------------------------|
| dist(C$_{al}$;H$_a$)$^a$ | 1      | 1.3Å                 | 3000                      | Prevent unwanted proton transfers involving the ethyl fragment. |
| CN(C$_{al}$;H)$^b$     | 1.2$^a$| 1.4 $(r_0=1.5\text{Å})$ | 2500                    | Prevent undesired protonation of the aromatic ring.               |
| min[dist(C$_{ar}$;C$_a$)]$^c$ | 2$^a$  | 6.0Å                 | 1000                     | Improve barrier recrossing by limiting ethene diffusion.         |
| CN(C$_{al}$;O$_{zeo}$) | 2$^a$  | 0.25 $(r_0=2.0\text{Å}$, NN=10) | 2500 | Prevent the formation of a surface ethoxide species.             |

$^a$5 separate walls, one for each C$_{al}$-H$_a$ bond.
$^b$six separate walls, one for each C$_{al}$ with all the H of the system.
$^c$To make the function continuous, the minimal distance is defined by PLUMED as $d_{\text{min}} = \beta(\log \sum_i \exp ((\beta/d_i)))^{-1}$ with $\beta=500$.

The same applies for the umbrella sampling simulations. A complete list of all upper walls (i.e. acting when CV $> CV_0$) adopted in the US simulations, together with their parameters and purposes, is reported in Table S11. For TS1, also a lower wall (i.e. acting when CV $< CV_0$) was needed to prevent unwanted processes to occur. For the sake of clarity, it is reported separately in Table S12.

| TS     | CV            | upper value $(CV_0)$ | $\kappa$ (kJ·mol$^{-1}$) | Purpose                                         |
|--------|---------------|----------------------|---------------------------|------------------------------------------------|
| TS0$^*$| min[dist(O$_{zeo}$;C$_{al}$)]$^a$ | 5.0Å                 | 1000                      | Improve barrier recrossing by limiting ethene diffusion.         |
| TS0$^{0,1}$| min[dist(O$_{zeo}$;C$_{al}$)]$^a$ | 4.0Å                 | 1000                      | Improve barrier recrossing by limiting ethanol diffusion.        |
|        | dist(O$_{al}$;C$_{al}$) | 4.0Å                 | 1000                      | Improve barrier recrossing by limiting water diffusion.          |
|        | dist(C$_{al}$;H$_a$)$^b$ | 1.19Å                | 5000                      | Prevent unwanted proton transfers involving the ethyl fragment.  |
| TS1    | dist(C$_{al}$;H$_a$)$^c$ | 1.2Å                 | 6000                      | Prevent unwanted proton transfers involving the ethyl fragment.  |
|        | CN(C$_{ar}$;H$_a$)$^c$ | 1.5 $(r_0=1.5\text{Å})$ | 2500 | Prevent undesired protonation of the aromatic ring.             |
| TS2$^*$| min[dist(C$_{ar}$;C$_{al}$)]$^a$ | 6.0Å                 | 1000                      | Improve barrier recrossing by limiting ethene diffusion.         |
|        | CN(C$_{al}$;O$_{zeo}$) +0.5[CN(C$_{al}$;H$_a$) -CN(O$_{zeo}$;H$_a$)] | 2.8 (r0 same as CV of TS0$^*$, see Table S1) | 3000 | Prevent the formation of a surface ethoxide species.           |
|        | dist(C$_{ar}$;H$_a$)$^d$ | 1.4Å                 | 2500                      | Prevent undesired deprotonation of the aromatic ring.            |

Table S10. Summary of the quadratic walls adopted during the multiple-walkers well-tempered metadynamics simulations, together with their parameters and their purpose. The labelling of the atom species is the same as in Table S1.
Few comments can be made on the walls selection; firstly, they can be divided into walls used to improve barrier recrossing and walls used to prevent unwanted side reactions. Starting from the formers, their presence will have an impact on the reaction FES. Indeed, they will drastically reduce the available configurational space for the adsorbed molecules, possibly causing a significant entropic penalty to the free energy (in their absence, the diffusing molecule could in principle explore the whole available volume in the zeolite unit cell). Being more interested in an estimate of the intrinsic reaction kinetics than in a full analysis of the transport properties in the material, we decided to set similar cutoff distances for similar reactions which should allow for a sufficient mobility of the species around the active site while allowing a fast convergence of the FES estimate.

The second type of walls, used to prevent side reactions, should have a minimal impact on the FES. The type of wall and the related parameters were mostly chosen based on chemical intuition and on a trial-and-error procedure. For instance, TS3, required to limit the interactions between the reacting proton and all the other aromatic carbons, while for TS3p only walling the meta carbon atoms proved to be enough to avoid protonation of the spectating aromatic carbons (being the para protonation inherently easier than the ipso). The bias due to the walls presence was carefully monitored in all simulations, to ensure that they were hit only seldomly whenever the system was trying to go through an undesired reaction path (because of the random statistical motion of the atoms) and not in a continuous fashion, symptom of an improperly placed wall that would significantly impact the final free energy estimate.

### Table S12. Summary of the lower quadratic walls adopted during the umbrella sampling simulations, together with their parameters and their purpose.

| TS   | CV       | lower value (CV$_0$) | $\kappa$ (kJ·mol$^{-1}$) | Purpose                                      |
|------|----------|----------------------|---------------------------|----------------------------------------------|
| TS2$^{OH}$ | dist(C$_{al}$;H$_{al}$) | 1.2Å | 3000 | Prevent unwanted proton transfers involving the ethyl fragment. |
|      | CN(C$_{al}$;O$_{zeo}$) | 0.25 (r$_0$=2.0Å) | 2500 | Prevent the formation of a surface ethoxide species. |
|      | min[dist(C$_{ar}$,C$_{al}$)] | 4.0Å | 1000 | Improve barrier recrossing by limiting ethanol diffusion. |
|      | dist(O$_{al}$;C$_{al}$) | 4.0Å | 1000 | Improve barrier recrossing by limiting water diffusion. |
|      | dist(C$_{al}$;H$_{al}$) | 1.4Å | 3000 | Prevent undesired deprotonation of the aromatic ring. |
| TS3$_i$ | CN(C$_{ar,ips}$;H$_{al}$) | 0.5 (r$_0$=1.5Å) | 2500 | Prevent undesired protonation of the aromatic ring. |
| TS3$_p$ | CN(C$_{ar,meta}$;H$_{al}$) | 0.7 (r$_0$=1.5Å) | 2500 | Prevent undesired protonation of the aromatic ring. |

$^a$To make the function continuous, the minimal distance is defined by $d_{min} = \beta(\log \Sigma \exp (\beta/d_i))^{-1}$ with $\beta$=500.

$^b$5 separate walls, one for each C$_{ar}$-H$_{al}$ bond.

$^c$6 separate walls, one for each C$_{ar}$ with all the H$_{ar}$ of the system.

$^d$6 separate walls, one for each C$_{ar}$-H$_{al}$ bond.

$^e$5 separate walls, one for each C$_{ar}$ not involved in the reaction with the H being transferred.

$^f$2 separate walls, one for each meta C$_{ar}$ with the H being transferred.

$^g$Only 6 separate walls, one for each C$_{ar}$ with all the H$_{ar}$ of the system.
Error estimation

As stated in the main manuscript, the kinetic constant of a reaction can be estimated from an umbrella sampling simulation using classical transition state theory with the formula:

$$\kappa_{TST} = \sqrt{\frac{1}{2\pi\beta} \langle \nabla_x F(V) \rangle_{CV^*} \int_{CV^*}^\infty e^{-\beta F(V)} dCV}$$

(S3)

To obtain a final estimate for the confidence interval associated with the kinetic constant, the error associated with the FES estimate ($F(CV)$) and with the rate prefactor ($A = \sqrt{1/2\pi\beta} \langle \nabla_x F(CV) \rangle_{CV^*}$) were evaluated separately.

An accurate estimate of the reaction free energy surface and the relative confidence interval can be obtained only by fully decorrelating the CV time series obtained from the US simulations. To achieve this, the Pymbar library$^6,7$ was used, with which we also automatically identified and discarded the simulation equilibration time. For the 2-dimensional umbrellas of TS0, the highest statistical inefficiency among the two computed separately for each CV was used to subsample the trajectories. In some umbrellas, very large statistical inefficiencies were computed which led to discard more than 95% of the data. While this approach necessarily increases the uncertainty of the free energy surface, it is also the only way to obtain a reliable and unbiased estimate of it.

In order to adequately estimate error bars on the free energy surface obtained from WHAM analysis, we start by interpreting the WHAM equations as a maximum likelihood estimate (MLE).$^8$ Using the asymptotic normality of the MLE, one can construct the Fisher information matrix and use it to derive the covariance matrix associated with the MLE.$^9$ However, since the WHAM equations represent an MLE associated with various simulations, we start by computing the Fisher matrix $\overline{F}_i$ for each simulation $i$ (consisting of $N_i$ simulation steps) separately and combine them afterwards into the global covariance matrix $\overline{\sigma}$:

$$\overline{\sigma} = \left( \sum_i N_i \overline{F}_i \right)^{-1}$$

(S4)

Note that this formula can be interpreted as simply adding up all available information. The square root of the diagonal elements of the covariance matrix then allows to compute the standard deviation on the probability (and hence free energy) of each bin in the free energy profile. By taking an error bar as two times the standard deviation, we compute a 2-sigma error bar corresponding to a 95% confidence interval (assuming normal distribution of the estimated free energy values, which is indeed the case for large samples sizes according to the asymptotic normality property of the MLE).

Since the rate prefactor is evaluated only on a small interval around the transition state value (0.02 in CV units), the system will cross it only seldomly during the simulation, even in the umbrella on top of the TS. For this reason, the data obtained from the trajectories are much less correlated. The confidence interval on the rate prefactor was then estimate using the block averaging method.$^{10}$ Both the MLE and block averaging methods are implemented in the ThermoLIB library, which was used for the data analysis of all reaction steps.$^{11}$

To combine the two error estimates together, $F(CV)$ and $A$ were assumed to be normally distributed around their average value, with standard deviation computed from the MLE or block averaging methods, respectively. $\kappa_{TST}$ was then recomputed 100,000 times using a Monte Carlo (MC) procedure to pick random values of both $F(CV)$ and $A$ from the respective normal distribution. It must be pointed out that, by selecting random values for $F(CV)$, the position of the transition state (intended as the point with highest free energy between reagents and products) can also change. However, computing the rate prefactor for every possible TS value that can occur in the MC simulation would be extremely time consuming and practically unfeasible. We then decided to fix the average value and confidence interval of $A$ to the ones obtained from the original FES. To validate this approach, TS0$^{OH}$ was chosen as a case study (the relative FES can be seen in Figure S8). Firstly, we analyzed in how far the transition state position (intended as the point on the FES with higher free energy between reagents and products) can
change during the MC run, since each point in the FES is randomly shifted up or down based on the confidence interval (bar plot in Figure S6, left). All values counted for more than $10^3$ times lie within 0.1 CV units from the original TS and the maximal and minimal observed variation in the TS position are contained within 0.4 CV units. We then computed the rate prefactor on various equally spaced points within this interval (purple dots in Figure S6, left). While at the original TS a value of $\sim 3 \cdot 10^{12}$s$^{-1}$ is computed, at the extremes of the intervals a decrease to $\sim 2.5 \cdot 10^{12}$s$^{-1}$ can be seen. Such difference would correspond to a decrease in the kinetic constant of about $k_{TST}^{(2)} / k_{TST}^{(1)} = A^{(1)} / A^{(2)} \approx 0.8$, thereby abundantly remaining in the same order of magnitude. Moreover, the difference in the computed phenomenological barrier would be of only $\Delta F_f^{(1)} - \Delta F_f^{(2)} = \ln(A^{(1)} / A^{(2)})^{-1 / \beta} \approx 0.9$kJ·mol$^{-1}$. Once the rate constant and the related phenomenological barrier (both forward and backward) were computed for each MC step, we considered the average value of the barrier as final estimate with an associated 95% confidence interval given by twice the standard deviation. While this is possible since the values of the phenomenological barrier are normally distributed, the rate constant presents a log-normal distribution (Figure S6, right). For this reason, we computed the confidence interval as $\exp(\langle \log(k) \rangle) / \exp(2 \cdot \sigma(\log(k)))$, $\exp(\langle \log(k) \rangle) \cdot \exp(2 \cdot \sigma(\log(k)))$.

**Figure S6.** (left) Distribution of the highest free energy point by randomly perturbing the FES of TS$^{\text{OH}}$ in the 100,000 Monte Carlo steps. The value of the rate prefactor ($A$), computed at equally spaced values along the distribution, is also shown with the associated confidence interval. The vertical dotted line indicates the position of the transition state in the original free energy surface. (right) Distribution of the forward and backward kinetic constant and phenomenological barrier as derived from the Monte Carlo simulation.
S3. Benzene alkylation mechanism

S3.1. 2D US FESs

As previously mentioned (Section S2.3), the US 1-dimensional FESs were expanded in 2 dimensions representing the CNs part of the linear combination defining the 1D CV. For CVs with format \( CV_A = \text{CN}(\alpha; \beta) - \text{CN}(\beta; \gamma) \), the 2D FES is reported as function of \( \text{CN}(\alpha; \beta) \) and \( \text{CN}(\beta; \gamma) \) while for CVs with format \( CV_B = \text{CN}(\alpha; \beta) + \frac{1}{2}[\text{CN}(\beta; \gamma) - \text{CN}(\gamma; \delta)] \) as function of \( \text{CN}(\alpha; \beta) \) and \( \text{CN}(\beta; \gamma) - \text{CN}(\gamma; \delta) \), but also as function of \( \text{CN}(\beta; \gamma) \) and \( \text{CN}(\gamma; \delta) \) to indeed show that the sampling is adequate along every separate coordination number. In the following figures the final 1-dimensional free energy profile with associated error bars and the expanded two-dimensional free energy surface(s) for the various steps of the benzene ethylation reaction are shown. In the case of TS0\(^{\pm}\), because of the additional 2-dimensional umbrellas, the final 2D FES is shown together with its projection along the horizontal axis.

![Diagram showing 2D US FESs](image)

Figure S7. a) 2-dimensional free energy profile as obtained from US. Black crosses highlight the center of the additional two-dimensional umbrellas used to improve the sampling of the transition state region. b) 1-dimensional free energy profile obtained by projecting the 2-dimensional FES on the horizontal axis. The grey shadow represents the error bars as obtained from the MLE method. c) 2-dimensional free energy surface obtained by expansion of the 1D profile, as function of the CNs sampling the protonation of ethene.
Figure S8. a) 1-dimensional free energy profile as obtained from US. The grey shadow represents the error bars as obtained from the MLE method. b) 2-dimensional free energy surface obtained by expansion of the original 1D profile, as function of the CNs constituting the final CV.

Figure S9. a) 1-dimensional free energy profile as obtained from US. The grey shadow represents the error bars as obtained from the MLE method. b) 2-dimensional free energy surface obtained by expansion of the original 1D profile, as function of the CNs constituting the final CV.
**TS2**

![Figure S10](image)

**Figure S10.** (a) 1-dimensional free energy profile as obtained from US. (b) 2-dimensional free energy surface obtained by expansion of the original 1D profile, as function of the CNs constituting the final CV. (c) 2-dimensional free energy surface obtained by a further expansion, useful to focus on the proton transfer process.

**TS2OH**

![Figure S11](image)

**Figure S11.** (a) 1-dimensional free energy profile as obtained from US. The grey shadow represents the error bars as obtained from the MLE method. (b) 2-dimensional free energy surface obtained by expansion of the original 1D profile, as function of the CNs constituting the final CV.
Figure S12. a) 1-dimensional free energy profile as obtained from US. b) 2-dimensional free energy surface obtained by expansion of the original 1D profile, as function of the CNs constituting the final CV.

Figure S13. a) 1-dimensional free energy profile as obtained from US. b) 2-dimensional free energy surface obtained by expansion of the original 1D profile, as function of the CNs constituting the final CV.
S3.2. Mobility analysis

Figure S14. Time evolution of the distances between the hydroxyl group oxygen and the BAS (H1) and the hydroxyl hydrogen (H2, see structure in the inset) during the MD simulation of the umbrella in the reactant state of TS0^{OH} (CV_0=-0.8). As visible two proton exchange events occur during the simulation and a strong hydrogen bond persist for its whole duration.

Figure S15. Time evolution of the minimum distances between the oxygens in the first two coordination spheres around the Al defect and the BAS (H1) and the hydroxyl hydrogen (H2) during the MD simulation of the umbrella in the reactant state of TS0^{OH} (CV_0=-0.8). It can be seen how the distance of the non-bonded hydrogen is, on average, around 2-3Å, thus indicating a very weak hydrogen bond.

Figure S16. Distribution of the minimum distances between the alkyl carbon atoms and the zeolite oxygens (in blue) and the alkyl carbon atoms and the BAS (in orange, only for ethene) taken from the umbrella in the reactant state of ethene (a) and ethanol (b). As visible, the latter has a narrow distribution that spikes between 3 and 4Å, in line with the formation of a strong hydrogen bond between its hydroxyl group and the BAS. For ethene, on the other hand, the molecule can go relatively far from the BAS and would certainly diffuse away without the restraining wall.
Figure S17. 2-dimensional expansion of the free energy profile of TS0^{OH} (right) and TS2^{OH} (left). The vertical CV encodes the proton transfer from the zeolite to the hydroxyl group, with CN(O_{al};H) going from ~1 to ~2 and CN(O_{zeo};H) from ~1 to ~0. At the transition state of TS0^{OH} the proton transfer is complete (CN(O_{al};H)-CN(O_{zeo};H)~1.8), while in TS2^{OH} the proton transfer takes gradually place while the alkylation proceeds (as indicated by the wide span of CN(O_{al};H)-CN(O_{zeo};H) values in the transition state region).

Figure S18. Superposition of multiple structures taken every 100 simulation steps from the transition state umbrella of (from left to right) TS1, TS2^- and TS2^{OH}. Benzene and ethene are shown in yellow and blue stick representation, respectively, while the hydroxyl group in TS2^{OH} as purple spheres. The H-ZSM-5 framework is in grey and shown along the straight channel, the Al defect (in pink) and the oxygen in its first coordination sphere (in red) are highlighted with larger colored spheres.
S4. Comparison with previous literature reports

Table S13. Activation energies (kJ mol\(^{-1}\)) for the benzene ethylation reaction in H-ZSM-5 derived from static calculations, as extracted from the available literature. The data refer to simple electronic energies (\(\Delta E_{el}\)), zero-point corrected electronic energies (\(\Delta E_{el}+ZPE\)) or free energies (\(\Delta G\)). The Level Of Theory (LOT) at which the calculation has been performed is also reported. If available, both the forward (f) and backwards (b) barriers for every reaction step are reported.

| Data          | LOT                                      | TS0\(^f\) | TS1  | TS2\(^b\) | TS3          | TS0\(^{OH}\) | TS2\(^{OH}\) | Ref  |
|---------------|------------------------------------------|-----------|------|-----------|-------------|--------------|---------------|------|
| \(\Delta E_{el}\) | B3LYP/6-31G* (4T cluster)               | 132       | 230\(^a\) |           |             |              |               | 12   |
| \(\Delta E_{el}+ZPE\) | B3LYP/6-31G* (4T cluster)               | 131       | 226\(^a\) |           |             |              |               | 13   |
| \(\Delta E_{el}+ZPE\) | B3LYP/DZSV(P) (33T cluster)            | 67        | 107   | 131       | 84          | 96           | 115           | 1–36\(^b\) 70–116\(^b\) | 14   |
| \(\Delta E_{el}+ZPE\) | PBE/PW + higher LOT corrections (periodic, corrections on cluster) | 130       | 143   | 103       | 190\(^a\)   | 130          | 221\(^a\)     | 15   |
| \(\Delta G(573K)\) | PBE-D3/PW (periodic)                     | ~95        | 76–83 | ~85       | ~117        | ~126         |               | 16   |
| \(\Delta G(573K)\) | oB97XD/6-31G(d,p):AM1 (72T cluster)     | 84        | 106   | 107       | 266\(^a\)   | 107          | 158\(^a\)     | 17   |
| \(\Delta G(573K)\) | PBE-D3/TZVP (periodic, dynamic)          | 101±7      | 92±7  | 80±5      | 64±5        | 102±3        | 55±3          | 18±1 96±1 115±3 80±3 108±8 44±8 | This work |

\(^{a}\)Referred to neutral ethylbenzene, no Wheland complex considered/found as minimum.

\(^{b}\)Value dependent on the considered adsorbate conformation.
S5. Insights into the role of water in the protonation reaction

S5.1. Cluster size involved in the reaction

When simulating the para protonation of ethylbenzene in the presence of water, the n (1,3 or 6) water molecules were manually placed in proximity of the active site at the beginning of the simulations. No restraints were imposed on their positions and therefore they were able to freely diffuse in the unit cell. To understand whether water remains close to the active site, possibly playing a role in the protonation reaction, the size of the water cluster in proximity of the Al defect was monitored in the simulations. For each umbrella, 1 frame every 5 fs was considered and, for each of such frames, a cutoff radius of 5 Å around the 4 O atoms in the first coordination sphere of the Al defect (O\textsubscript{zeo}) was set. Then, the presence of at least one water oxygen (O\textsubscript{w}) within this cutoff radius was assessed. If no O\textsubscript{w}s were found, the cluster size was set to 0 and the following frame was considered. Otherwise, the cluster size was set to 1. From the O\textsubscript{w} that is closest to O\textsubscript{zeo}, a second cutoff radius of 4 Å was defined and all other O\textsubscript{w}s in such cutoff counted and their number added to the cluster size. Then, the newly counted O\textsubscript{w}s became the center of a new cutoff radius to look for not yet included O\textsubscript{w}s and so on, until no further O\textsubscript{w}s were found (Figure S19).

For every water loading, the results of all umbrellas were joined and the percentage of occurrence of the various cluster sizes as function of the reaction CV reported (Figure 5d-f in the main manuscript).

Figure S19. Schematic depiction of the H-ZSM-5 channel intersection, as seen along the sinusoidal channel. The approximate location of ethylbenzene during the simulations of TS3nw is also shown. Three water molecules are randomly drawn in the intersection to illustrate the definition of water cluster size in proximity of the active site. Firstly, water 1 is found to be the closest to O\textsubscript{zeo} and within the 5 Å cutoff radius (blue region in the figure). Then, it is found that water oxygen 2 lies within 4 Å from water oxygen 1. Finally, no further O\textsubscript{w}s are found within 4 Å from water 1 or 2 (green circles in the figure). A water cluster size of 2 would then be associated with a simulation snapshot resembling this schematic picture (color code: H white, C grey, O red, Al pink and Si yellow).

S5.2. Expansion of the reaction profile in two dimensions

To quantitatively describe the energetic effect of the water-mediated proton hopping in the protonation reaction at various loadings, the free energy profile was expanded in term of two new collective variables, defining the proton-O\textsubscript{zeo} and proton-C\textsubscript{para} distances. To define such fictitious proton atom, a combination of the collective variables recently proposed by Pérez de Alba Ortíz et al. and Grifoni et al. was used. Firstly, each proton in the system that could be involved in the protonation reaction (the BAS, all the water protons and the proton bounded to the ethylbenzene para carbon) is attributed to its closer proton acceptor (the water oxygens, O\textsubscript{w}, the ethylbenzene para carbon, C\textsubscript{para}, and the oxygens in the first coordination sphere of the Al, O\textsubscript{zeo}) using the formula:
\[ W_i = \sum_{j \in H} e^{-\lambda |r_i - r_j|} \sum_{m \in A} e^{-\lambda |r_m - r_j|} \]  

(S5)

In which \( i \) and \( m \) index the proton acceptors (A) and \( j \) runs over the protons. \( r_i \) is the position of atom \( i \) and \( \lambda \) is a parameter chosen to be 6 in this case. Thanks to this formulation, \( W_i \) roughly equals the number of protons that are closer to atom \( i \) than to any other proton acceptor in the system. Indeed, the argument of the main summation tends to 1 if proton \( j \) is closer to atom \( i \) than to any other proton acceptor and quickly drops to zero otherwise. If a proton is exactly in between two proton acceptors, it will contribute to \( W_i \) by 0.5 for both. For each \( i \)th proton acceptor an offset value \( W_{0,i} \) can be defined, from which the fictitious proton position can be computed:

\[ r_p = \frac{\sum_i e^{\gamma(W_i - W_{0,i})} r_i}{\sum_i e^{\gamma(W_i - W_{0,i})}} \]  

(S6)

By using \( \gamma = 14 \), this weighted average will ensure that the position of the fictitious proton atom \( r_p \) corresponds with the position of the proton acceptor whose number of closer protons (\( W_i \)) is greater than its offset value \( W_{0,i} \), i.e. the atom that carries the extra proton. Here, we selected an offset value of 0 for the four \( O_{zeo} \), of 2 for all \( O_w \) of 1 for \( C_{para} \). Then, as final collective variables, we considered the minimum distance between the proton and \( O_{zeo} \) (\( CV_{1,p} = \min(|r_p - r_{O_{zeo}}|) \)) and the distance between the proton and \( C_{para} \) (\( CV_{2,p} = |r_p - r_{C_{para}}| \)). The former will be equal to zero when the proton is attached to the zeolite framework and greater than zero otherwise, while the latter will be equal to zero when the ethylbenzene has been protonated and greater than zero otherwise (Figure S20).

**Figure S20.** Schematic depiction of the H-ZSM-5 channel intersection, as seen along the sinusoidal channel. The adsorbed ethylbenzene and some water molecules (3, in this example) are also included. The variation of the two collective variables used to monitor the proton position when different atoms are protonated is shown. 

- **a.** When the extra proton is located on the zeolite framework, \( CV_{1,p} = 0 \) (the location of the fictitious proton atom corresponds with the \( O_{zeo} \) on which the BAS is located) and \( CV_{2,p} > 0 \). When the proton is solvated by the water forming an hydronium ion, both \( CV_{1,p} \) and \( CV_{2,p} \) are greater than 0.
- **b.** When the ethylbenzene has been protonated, \( CV_{1,p} > 0 \) and \( CV_{2,p} = 0 \) (the location of the fictitious proton atom corresponds with \( C_{para} \)). Color code: H white, C grey, O red, Al pink and Si yellow.
S6. References

(1) Van Den Broeck, E.; Verbraeken, B.; Dedecker, K.; Cnudde, P.; Vanduyfhuys, L.; Verstraelen, T.; Van Hecke, K.; Jerca, V. V.; Catal, S.; Hoogenboom, R.; et al. Cation-Π Interactions Accelerate the Living Cationic Ring-Opening Polymerization of Unsaturated 2-Alkyl-2-Oxazolines. *Macromolecules* **2020**, *53*(10), 3832–3846.

(2) De Wispelaere, K.; Bailleul, S.; Van Speybroeck, V. Towards Molecular Control of Elementary Reactions in Zeolite Catalysis by Advanced Molecular Simulations Mimicking Operating Conditions. *Catal. Sci. Technol.* **2016**, *6*(8), 2686–2705.

(3) Bailleul, S.; Dedecker, K.; Cnudde, P.; Vanduyfhuys, L.; Waroquier, M.; Van Speybroeck, V. Ab Initio Enhanced Sampling Kinetic Study on MTO Ethene Methylation Reaction. *J. Catal.* **2020**, *388*, 38–51.

(4) Rey, J.; Bignaud, C.; Raybaud, P.; Bučko, T.; Chizallet, C. Dynamic Features of Transition States for β-Scission Reactions of Alkenes over Acid Zeolites Revealed by AIMD Simulations. *Angew. Chem. Int. Ed.* **2020**, *59*(43), 18938–18942.

(5) Cnudde, P.; Wispelaere, K. De; Vanduyfhuys, L.; Demuynck, R.; Mynsbrugge, J. Van Der; Waroquier, M.; Speybroeck, V. Van. How Chain Length and Branching Influence the Alkene Cracking Reactivity on H-ZSM-5. *ACS Catal.* **2018**, *8*, 9579–9595.

(6) Chodera, J. D.; Swope, W. C.; Pitera, J. W.; Seok, C.; Dill, K. A. Use of the Weighted Histogram Analysis Method for the Analysis of Simulated and Parallel Tempering Simulations. *J. Chem. Theory Comput.* **2007**, *3*(1), 26–41.

(7) Shirts, M. R.; Chodera, J. D. Statistically Optimal Analysis of Samples from Multiple Equilibrium States. *J. Chem. Phys.* **2008**, *129*, 124105.

(8) Zhu, F.; Hummer, G. Convergence and Error Estimation in Free Energy Calculations Using the Weighted Histogram Analysis Method. *J. Comput. Chem.* **2012**, *33*(4), 453–465.

(9) Härdle, W. K.; Simar, L. *Applied Multivariate Statistical Analysis, Fourth Edition*; Springer Berlin Heidelberg, 2015.

(10) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids: Second Edition*; Oxford University Press, 2017.

(11) ThermoLIB | Center for Molecular Modeling https://molmod.ugent.be/software/thermolib (accessed Nov 19, 2021).

(12) Vos, A. M.; Schoonheydt, R. A.; De Proft, F.; Geerlings, P. Reactivity Descriptors and Rate Constants for Acid Zeolite Catalyzed Ethylation and Isopropylation of Benzene. *J. Phys. Chem. B* **2003**, *107*(9), 2001–2008.

(13) Arstad, B.; Kolboe, S.; Swang, O. Theoretical Investigation of Arene Alkylation by Ethene and Propene over Acidic Zeolites. *J. Phys. Chem. B* **2004**, *108*(7), 2300–2308.

(14) Hansen, N.; Brüggemann, T.; Keil, F. J.; Bell, A. T. Theoretical Investigation of Benzene Alkylation with Ethene over H-ZSM-5. *J. Phys. Chem. C* **2008**, *112*, 15402–15411.

(15) Hansen, N.; Kerber, T.; Sauer, J.; Bell, A. T.; Keil, F. J. Quantum Chemical Modeling of Benzene Ethylation over H-ZSM-5 Approaching Chemical Accuracy: A Hybrid MP2:DFT Study. *J. Am. Chem. Soc.* **2010**, *132*(33), 11525–11538.

(16) Wang, D.; Wang, C. M.; Yang, G.; Du, Y. J.; Yang, W. M. First-Principles Kinetic Study on Benzene Alkylation with Ethanol vs. Ethylene in H-ZSM-5. *J. Catal.* **2019**, *374*, 1–11.

(17) Acharya, D.; Chen, W.; Yuan, J.; Liu, Z.; Yi, X.; Xiao, Y.; Zheng, A. Stepwise or Concerted Mechanisms of Benzene Ethylation Catalyzed by Zeolites? Theoretical Analysis of Reaction Pathways. *Catal. Letters* **2021**, *151*, 3048–3056.

(18) Pérez De Alba Ortiz, A.; Tiwari, A.; Puthenkalathil, R. C.; Ensing, B. Advances in Enhanced Sampling along Adaptive Paths of Collective Variables. *J. Chem. Phys.* **2018**, *149*(7).

(19) Grifoni, E.; Piccini, G. M.; Parrinello, M. Microscopic Description of Acid–Base Equilibrium. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*(10), 4054–4057.