AIChE

Rational Design of Shape Selective Separations and Catalysis: Lattice Relaxation and Effective Aperture Size

Chrysanthos E. Gounaris, James Wei, and Christodoulos A. Floudas
Dept. of Chemical Engineering, Princeton University, Princeton, NJ 08544

Rajiv Ranjan and Michael Tsapatsis
Dept. of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, MN 55455

DOI 10.1002/aic.12016
Published online September 29, 2009 in Wiley InterScience (www.interscience.wiley.com).

Gounaris et al. presented a computational method that can be used for the quick screening of zeolite structures and provide predictions regarding which of them have the potential to exhibit high selectivity among a set of molecules of interest. This article builds upon this earlier work and furthers our understanding of diffusion processes in zeolites and other microporous metal oxides. We first present an augmented formulation to account for the flexibility of the zeolitic portal and conduct an analysis to assess the effect of varying the parameters of the associated quadratic potential. We then introduce a methodology to map the energetic landscape, identify all locally optimal conformations, and probabilistically account for the multiplicity of conformers. Finally, we conduct sensitivity analysis on the effective size of the aperture, and show how the methodology can be fine–tuned through experimental observations. A comprehensive database of 290 molecules of industrial interest and a total of 123 different zeolite structures were used in this study. © 2009 American Institute of Chemical Engineers
AIChE J, 56: 611–632, 2010
Keywords: catalysis, molecular sieves, selectivity, separations, supramolecular chemistry, zeolites

Introduction

Selective separation is based on the sizes and shapes of the sorbate molecules in comparison with the windows and channels of the sorbent hosts. (Chen et al., 1996; Kresge and Dhingra, 2004; Smit and Maesen, 2008). An example of suitable hosts is the family of zeolites, which are used routinely in industry as molecular sieves, adsorbents, and catalyts. Their success lies in the fact that their structure features nanoporous windows, of comparable size with that of small molecules, offering the potential for shape–selectivity. However, despite the great variety of existing structures, only a handful of them (e.g., LTA, MFI, FAU) are being utilized in practice and no significant effort is being made to find potential application for the remaining ones; it is not unlikely that some may have openings that are highly suited to be selective in a process of interest. To bridge this gap, we developed an integrated framework (Gounaris et al., 2006a,b) that allows us to screen large host structure databases and provide us with educated predictions on which structures are more likely to exhibit selectivity between a given pair of molecules of interest.

Through means of global optimization search, the method was able to identify the guest conformation that would penetrate the host portal in the energetically most favourable way (i.e., with the globally minimum activation energy required).

Correspondence concerning this article should be addressed to C. A. Floudas at floudas@titan.princeton.edu

The energetic interactions were considered to be due to the strain exerted when the guest molecule attempted to penetrate through the confined space of the window, as this is quantified by a Lennard–Jones potential. Given an objective to selectively distinguish between two molecules A and B, and a set of host structures for which we have obtained results, a screening process can be performed. This is based on finding a zeolite that has a much lower activation energy for molecule A, which passes readily, in comparison with a much higher activation energy for molecule B, which has difficulty in passing. The proposed approach can discover surprising candidate structures. It can also limit considerably the search space for potential research, directing the efficient allocation of funds and resources. The key assumptions in the approaches by (Gounaris et al., 2006a,b) are that (a) the zeolite portal is rigid, and that (b) translation and rotation are the only degrees of freedom of the guest molecule.

This article compiles a number of extensions to the original approach. In ''Flexibility of Portal'' Section, we present an augmented formulation that takes into account the flexibility of the zeolite portals, through use of an appropriate quadratic potential. This extra layer of complexity offers a better physical description of the interactions, improving the accuracy of the predictions. In ''Entropic Contributions of Locally Optimal Conformers'' Section, we demonstrate how one can make use of local optimization search techniques so as to account for the entropy contribution from multiplicity of conformers that may be explored during penetration of a guest moplecule into the host. On the basis of exhaustive initialization from a fine grid of initial conformations, we map the energetic landscape of the interactions and identify the totality of locally optimal conformations. Then, depending on an assumption regarding the distribution of initial states, we derive a probability measure for the occurence of each local solution. The calculated probabilities allow us to weigh appropriately the various solutions and derive expectations of the various quantities of interest. Finally, in ''Effective Radii of Portal Oxygen Atoms'' Section, we address the uncertainty involved in some of the model parameters. In particular, we perturb the values for the effective radii of the portal defining oxygen atoms and perform a sensitivity analysis on select results. By comparing the results with experimental observation, we are able to deduce the appropriate values for these parameters, thus fine tuning the framework.

Synopsis of Rigid Portal Formulation

In this section, we discuss briefly the mathematical model used in this study. A more detailed discussion is available in Gounaris et al. (2006b), where the model was presented originally.

We start from a simple molecular model where every atom $i = 1,2,\ldots,M$ corresponds to a sphere with center (x_i^0, y_i^0, z_i^0) and radius r_i. This initial fixed three-dimensional conformation can be obtained from quantum or molecular mechanics calculations, while some ''effective'' radius (e.g., Van der Waals) can be used for the value of the parameters r_i. We also consider a two-dimensional portal that lies on the xy-plane and consists of atoms $j = 1,2,\ldots,N$ that are represented by circles of center (xh_j, yh_j) and radius r_j.

Parameterizing with the three Euler angles, ϕ, θ, ψ, we rotate suitably the molecule in three-dimensional space, translate it by (xt,yt), and then project it onto the two-dimensional plane where the portal lies. This projection results into a set of circles, each corresponding to a particular molecule atom, and having center (x_i,y_i) and radius r_i. Our objective is to determine a three-dimensional orientation and positioning that would allow the molecule to penetrate the portal; that is, its projection to not overlap with the ring of the opening.

As the molecule may be too large to fit the opening completely (while no orientation or translation can prevent this), and according to the analysis presented in the earlier publications (Gounaris et al., 2006a,b), we allow reduction of the atom effective radii, and we penalize this distortion according to a Lennard–Jones type of potential. This penalization provides a measure of the energetic interactions between guest molecule and host portal; in particular, it takes into account the dominant nonbonded interactions due to overlapping electron orbitals.

We use variable δ to denote the amount of distortion on a single atom as

$$\delta = \frac{r_s}{r_0}, \tag{1}$$

where r_s is the required—possibly squeezed—radius, and r_0 is the original—Van der Waals— radius.

The strain experienced by atom i is defined as

$$S_i = \left(\frac{1}{\delta_i^{12}} - \frac{1}{\delta_i^6}\right). \tag{2}$$

The total strain required for penetration of the molecule through the portal quantifies the extent of the necessary distortion, and provides a rigorous characterization of the geometric affinity between the molecule and the portal. The objective is to minimize this total strain, S, required for penetration.

$$\min_{\substack{\phi,\theta,\psi \\ xt,yt \\ \delta_i,\delta_j}} S = \min_{\substack{\phi,\theta,\psi \\ xt,yt \\ \delta_i,\delta_j}} \left\{ \sum_i \left(\frac{1}{\delta_i^{12}} - \frac{1}{\delta_i^6}\right) + \sum_j j \in J_O \left(\frac{1}{\delta_j^{12}} - \frac{1}{\delta_j^6}\right) \right\}, \tag{3}$$

where J_O is the set of O–atoms on the portal side.

The reasoning behind accounting only for the distortion of the O–atoms of the host is that the T–atoms are positioned at the exterior of the portal (interior to the framework) and thus interact only weakly with the penetrating molecule. Furthermore, they are cations, which means that they have been stripped off a portion of their original electron clouds, making their interactions even more negligible. This assumption is being used routinely in the literature and traces its origin at the work of Kiselev and Du (1981) and Kiselev et al. (1985).

For every atom of the guest molecule, the position of its projected center should correspond to a valid rotation and translation of the original conformation provided. According

to the rotation and projection analysis (Gounaris et al., 2006b), we have

$$\begin{aligned} x_i &= \cos\theta\cos\phi x_i^0 + \cos\theta\sin\phi y_i^0 - \sin\theta z_i^0 + xt \quad \forall i \\ y_i &= (\sin\psi\sin\theta\cos\phi - \cos\psi\sin\phi)x_i^0 \\ &\quad + (\sin\psi\sin\theta\sin\phi + \cos\psi\cos\phi)y_i^0 \\ &\quad + \cos\theta\sin\psi z_i^0 + yt \quad \forall i. \end{aligned} \tag{4}$$

For every pair of host–guest atoms, it must be imposed that their effective circles cannot intersect with each other. This implies that the circles have to be distorted (squeezed) so as the projection of the guest to fit the opening of the host.

$$\begin{aligned} &\delta_{ij} \geq \delta_i r_i + \delta_j r_j \\ &\qquad\qquad \forall (i,j) \\ &\delta_{ij}^2 = \left(x_i - xh_j\right)^2 + \left(y_i - yh_j\right)^2 \end{aligned} \tag{5}$$

To avoid obtaining (otherwise valid) solutions where the guest molecule lies completely outside the portal area, we include in the formulation a set of constraints indexed over k. This set, abstractly defined here as $H_k\ (x_i,y_i,\delta_i) \leq 0,\ \forall\ (i,k)$ imposes that the molecule penetrates through the interior side of the opening. Typically, a set of linear constraints corresponding to the convex hull of the portal is used.

$$a_k x_i + b_k y_i \leq c_k - \delta_i r_i \quad \forall (i,k), \tag{6}$$

where parameters a_k,b_k,c_k can be easily calculated from the host atom coordinates $\begin{bmatrix} xh_j \\ yh_j \end{bmatrix}$.

Finally, the following bounds have to be applied to the decision variables of the problem:

$$-\pi < \phi, \theta, \psi \leq +\pi \tag{7}$$

$$\begin{aligned} &\delta^L \leq \delta_i \leq 1 \quad \forall i \\ &\delta^L \leq \delta_j \leq 1 \quad \forall j \end{aligned} \tag{8}$$

The bounds on the Euler angles are imposed so as not to obtain periodic solutions, while a cutoff lower bound $\delta^L = 0.1$ is imposed on all the δ–variables so as to avoid numerical difficulties (divisions by zero). Note that no bounds are imposed on the translation variables xt and yt, which are allowed to vary freely.

The requirement to distort a molecule for penetration corresponds to some activation energy, E_S, which leads to a decrease of the equilibrium concentration in the pores of the adsorbent, C, according to the Boltzmann equation:

$$\frac{C}{C_0} = \exp\left(-\frac{E_S}{RT}\right) = \exp\left(-\frac{4\varepsilon S^*}{RT}\right), \tag{9}$$

where C_0 is the concentration of the species at the exterior of the adsorbent, and ε is an averaged L-J parameter that serves as a hardness coefficient*.

*A geometric mean of atom–specific parameters may be used.

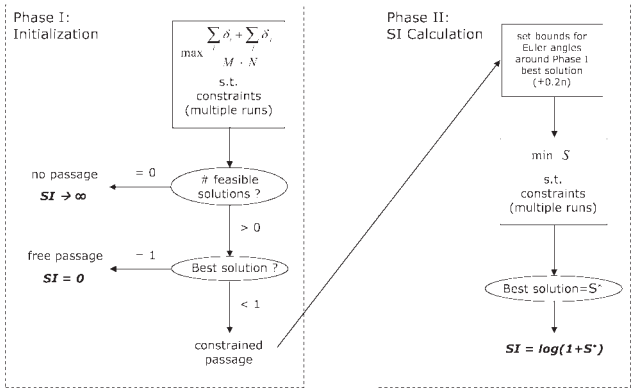


Figure 1. Schematic of strain index calculation algorithm.

Given that two molecules A and B require different amounts of distortion, selectivity is produced. The latter can be defined as the absolute difference between their reduced equilibrium concentrations.

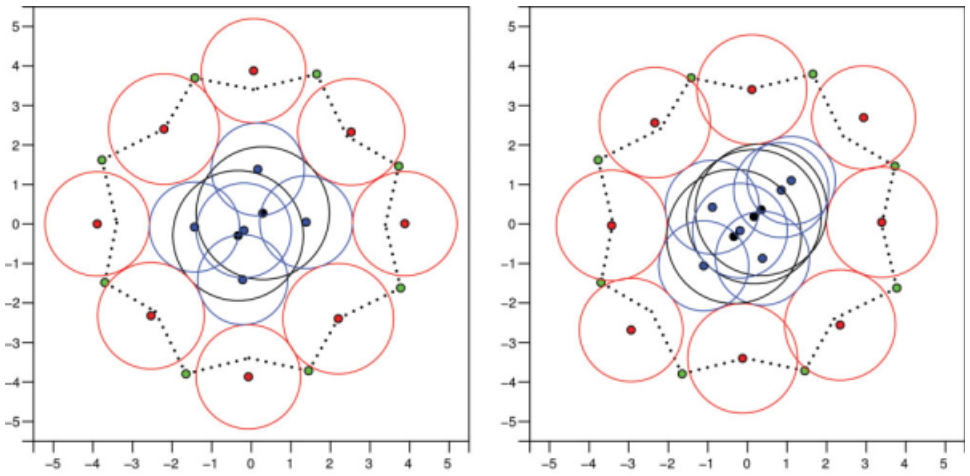
$$\gamma_{AB} = \left| \left(\frac{C}{C_0}\right)^A - \left(\frac{C}{C_0}\right)^B \right| \tag{10}$$

Note that the concentration reductions are calculated at equilibrium which inherently assumes that concentrations C_0^A and C_0^B are constant in the time scale of interest. If the concentration fields outside the adsorbent are subject to variations due to non steady state conditions (e.g., ongoing isomerization reactions $A \rightarrow B$), then this should be explicitly taken into account.

The minimization of the objective function (Eq. 3), subject to the constraints of Eqs. 4–8, constitutes a nonlinear programming problem (NLP) that involves continuous variables. The problem is highly non convex, rendering it very difficult to address it with reasonable amount of computational effort. To be able to address such a difficult optimization problem, we have chosen to instead employ local optimization methods augmented with an insightful initialization scheme that assists in avoiding convergence to nonglobal solutions. The details of this two—phase algorithm can be found in Gounaris et al. (2006b). A schematic is depicted in Figure 1. Note that *SI* denotes the "Stain Index" which was introduced as a suitable logarithmic scale.

Flexibility of Portal

The possibility that the zeolite window is flexible and does not maintain its original "nominal" structure when being penetrated by a guest molecule is an important issue that deserves attention. Indeed, there is a lot of experimental evidence that this might happen in many cases. For example, Karsli et al. (1992) showed that freshly regenerated aluminum–deficient ZSM–5 (silicalite–1) exhibits a different performance on *o*–xylene intake than a sample that has firstly been treated with *p*–xylene and then outgassed. Tezel and Ruthven (1990) observed an unusual hysteresis behavior for benzene adsorption on NaX, implying that the sorbend



(a) Propane (b) Propylene

Figure 2. Zeolite lattice relaxation under sorbate loading; dotted line represents nominal portal configuration.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

adapted its structure after being exposed to the sorbate for some time. Finally, Boulicaut et al. (1998) showed that high sorbate loading, such as the case in a liquid hexane stream, slightly expands the silicalite channels.

The lattice relaxation of zeolites has, thus, become a topic of rigorous research (Conner et al., 1990; Van Santen et al., 1991). Kopelevich and Chang (2001) icorporated lattice flexibility into transition state theory and their results came to much better agreement with experimental diffusion rate data than their studies with rigid structures. They reported considerable reductions in the activation barrier when the zeolite window is allowed to be stretched. Using molecular dynamics, Demontis et al. (1992) came to similar conclusions when studying the temperature dependence of methane diffusion in silicalite. Yashonath and Santikary (1993, 1994), Santikary and Yashonath (1994), and Bandyopadhyay and Yashonath (1995) conducted a series of molecular dynamics simulations of NaA, NaCaA, and VPI–5 zeolites, incorporating framework flexibility, in an attempt to verify the existence of anomalous diffusion. Their results suggest that the window expands during intercage migration only if the sorbate size is comparable to the window diameter and that the anomaly is independent of the type of host substance. Sartbaeva et al. (2006) have, in fact, conjectured that structural flexibility is a necessary feature that enables zeolite synthesis.

Figure 2 presents two schematics that illustrate how a zeolite portal may get distorted when being penetrated by a molecule. We observe that the rectangular projection of the propane molecule penetrates more comfortably when the top, left most, bottom, and right most O–atoms of the host are displaced, while the other four O–atoms need not depart or deviate signifcantly from their nominal positions. Meanwhile, the more eccentric propylene molecule benefits from displacement of the O–atoms that lie on its projection's major axis (top right towards bottom left). Note that the depicted lattice distortion has been artificially enhanced so as to better illustrate the concept.

Had the portal been rigid, as we have considered it in previous work (Gounaris et al., 2006a,b), there would have been much more strain exerted on the two systems and the equilibrium concentrations achieved would have been significantly lower. However, the portal's flexibility offers potential for sorption even for molecules that are of comparable size with the opening.

Formulation enhancements

Modeling flexibility of the portal requires that the host atom coordinates are no longer fixed parameters $\begin{bmatrix} xh_j \\ yh_j \end{bmatrix}$, but variables $\begin{bmatrix} xhv_j \\ yhv_j \end{bmatrix}$. This results into the replacement of Eq. 5 by the following:

$$\begin{aligned} &\delta_{ij} \geq \delta_i r_i + \delta_j r_j \\ &\qquad\qquad\qquad\qquad \forall (i,j) \\ &\delta_{ij}^2 = \left(x_i - xhv_j\right)^2 + \left(y_i - yhv_j\right)^2 \end{aligned} \tag{11}$$

Furthermore, additional nonlinear constraints have to be included in the optimization formulation, so as to describe the connectivity between atoms as well as the overall coherence of the structure. It is also our concern, for computational efficiency and robustness reasons, that any attempt to include window flexibility into the model shall be done in a manner that maintains the additional complexity to a minimum.

The optimization should still aim at minimizing the activation energy required for penetration, but the latter would now consist of two components: (a) the strain incurred by squeezing the atom spheres (repulsive forces due to overlapping electron orbitals) and (b) the energy required for distortion of the zeolite window.

$$\min_{\substack{\phi,\theta,\psi \\ xt,yt \\ \delta_i,\delta_j \\ xhv_j,yhv_j}} E = \min_{\substack{\phi,\theta,\psi \\ xt,yt \\ \delta_i,\delta_j \\ xhv_j,yhv_j}} E_S + E_Z \tag{12}$$

The first component, E_S, was addressed briefly in "Synopsis of Rigid Portal Formulation" Section and involved a Lennard–Jones type potential.

$$E_S = 4\varepsilon S = 4\varepsilon\left\{\sum_i\left(\frac{1}{\delta_i^{12}} - \frac{1}{\delta_i^{6}}\right) + \sum_{j\in J_O}\left(\frac{1}{\delta_j^{12}} - \frac{1}{\delta_j^{6}}\right)\right\} \quad (13)$$

Note that in the case where individualized L-J coefficients are available for every atom type involved, one could effectively take them into account by instead using Eq. 14.

$$E_S = \sum_i 4\varepsilon_i\left(\frac{1}{\delta_i^{12}} - \frac{1}{\delta_i^{6}}\right) + \sum_{j\in J_O} 4\varepsilon_j\left(\frac{1}{\delta_j^{12}} - \frac{1}{\delta_j^{6}}\right) \quad (14)$$

For the quantification of the second component, E_Z, a harmonic (quadratic) potential that penalizes deviation from the nominal T—O and O—(T)—O lengths can be used.

$$E_Z = \sum_j \frac{1}{2}k_{TO}\left(d_j^{TO} - d_j^{TO,nom}\right)^2 + \sum_{j\in J_O}\frac{1}{2}k_{OO}\left(d_j^{OO} - d_j^{OO,nom}\right)^2, \quad (15)$$

where d^{TO} is the Euclidean distance between two adjacent atoms of the portal (a T–atom and a O–atom) and d^{OO} is the distance between two successive O–atoms of the portal (that are both bonded to an intermediate T–atom). $d^{TO,nom}$ and $d^{OO,nom}$ are the respective nominal distances, as calculated from the structure database (Baerlocher et al., 2001; Baerlocher and McCusker web) using Eqs. 17, while k_{TO} and k_{OO} are suitable parameters (see "Computational Results" Section for exact values).

This potential reproduces satisfactorily the main features of aluminosilicate frameworks, as it has been reported in various studies (Demontis et al., 1988; Kopelevich and Chang, 2001, Santikary and Yashonath, 1994). For specific host–guest interaction simulations where a greater degree of accuracy is required, more sophisticated models (Deem et al., 1992, Schrimpf et al., 1992) can be used, provided reliable values for their parameters are available.

The definitions of distances d^{TO} and d^{OO} are also included in the formulation:

$$\begin{aligned} d_j^{TO} &= \sqrt{(xhv_{j+1} - xhv_j)^2 + (yhv_{j+1} - yhv_j)^2} \quad \forall j \\ d_j^{OO} &= \sqrt{(xhv_{j+2} - xhv_j)^2 + (yhv_{j+2} - yhv_j)^2} \quad \forall j \in J_O \end{aligned} \quad (16)$$

In a similar fashion, the nominal distances can be calculated as

$$\begin{aligned} d_j^{TO,nom} &= \sqrt{(xh_{j+1} - xh_j)^2 + (yh_{j+1} - yh_j)^2} \quad \forall j \\ d_j^{OO,nom} &= \sqrt{(xh_{j+2} - xh_j)^2 + (yh_{j+2} - yh_j)^2} \quad \forall j \in J_O. \end{aligned} \quad (17)$$

For periodicity reasons, we also need to define that

$$\begin{aligned} xhv_{(N+1)} &= xhv_1 \quad xhv_{(N+2)} = xhv_2 \\ yhv_{(N+1)} &= yhv_1 \quad yhv_{(N+2)} = yhv_2 \end{aligned} \quad (18)$$

and similarly for the respective parameters xh_j, yh_j, when $j = N + 1$ or $N + 2$.

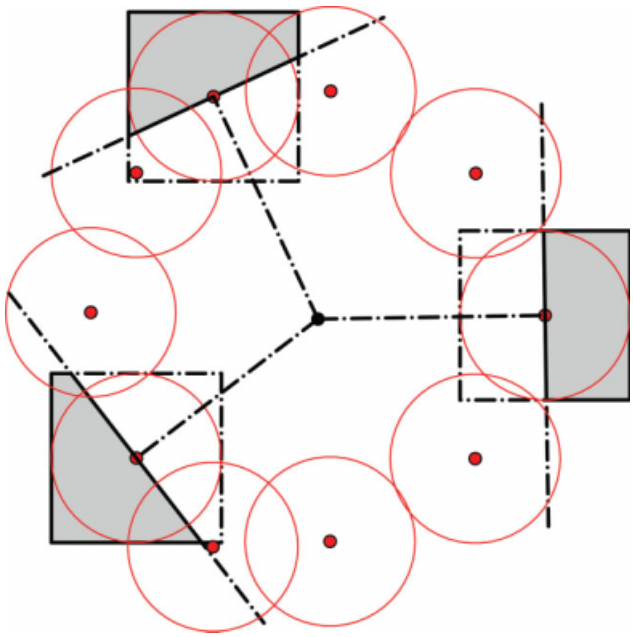


Figure 3. Feasible locations for some of the portal's O–atoms in the case of flexible template.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

We further impose a set of linear constraints to somewhat restrict the positioning of the host O–atoms. These constraints impose effectively that the portal can only be pushed outwards and not pulled inwards.

$$f_j xhv_j + g_j yhv_j \geq h_j \quad \forall j \in J_O, \quad (19)$$

where parameters f_j, g_j, h_j can be easily calculated from the host atom nominal positions $\begin{bmatrix} xh_j \\ yh_j \end{bmatrix}$. Note that these lines are defined with repect to some portal center. A suitable point could be the center of mass or the arithmetic average of all the coordinates.

Finally, the following bounds are applied to the positions of the portal's O–atoms.

$$\begin{aligned} xh_j - r_j \leq xhv_j \leq xh_j + r_j \\ \forall j \in J_O \\ yh_j - r_j \leq yhv_j \leq yh_j + r_j \end{aligned} \quad (20)$$

Note that no bound constraints apply to the positions of the portal's T–atoms, as these are not allowed to vary and essentially are kept fixed in their nominal positions.

$$\begin{aligned} xhv_j = xh_j \\ \forall j \notin J_O \\ yhv_j = yh_j \end{aligned} \quad (21)$$

The shaded areas depicted in Figure 3 correspond to the feasible areas for the positions of some of the portal's O–atoms. They constitute the intersections of the half-planes defined by Eq. 19 and the bounds of Eqs. 20.

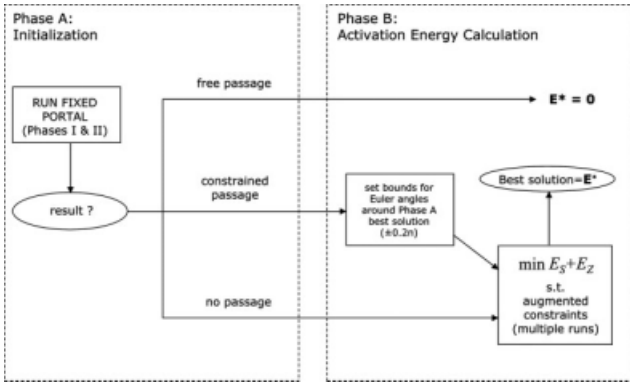


Figure 4. Algorithm for the calculation of activation energy with flexible portals.

Algorithmic enhancements

Given the additional nonconvexities, some algorithmic enhancements were required in order to obtain robust and reliable optimization solutions. The implementation consists of two phases; an overall schematic is depicted in Figure 4.

Phase A — Rigid portal. For a given molecule/portal pair, we apply the rigid portal algorithm that is synopsized in Figure 1. If the algorithm returns a "free passage" result, then we adopt this result and terminate, as there will be free passage irrespectively of whether the portal has the capability to stretch/expand or not. If the algorithm returns a "constrained passage" result, then the exact solution (ϕ^{*A}, θ^{*A}, ψ^{*A}, xt^{*A}, yt^{*A}) is stored and the program moves on to Phase B where the new (flexible portal) methodology will be used for the calculation of the exact activation energy. If the algorithm returns a "no passage" result, we still need to move on to Phase B, since relaxation of the portal might actually allow the guest molecule to be admitted, and we need to investigate whether this will occur or not. In this case, (ϕ^{*A}, θ^{*A}, ψ^{*A}, xt^{*A}, yt^{*A}) can be set to a random conformation.

Phase B — Activation Energy Calculation. This phase utilizes the nonlinear objective function of Eq. 12, that accounts for the combined effect of strain and portal flexibility.

As we expect that the optimal solution is in the vicinity of the Phase A solution, we set bounds in the Euler angles around (ϕ^{*A},θ^{*A},ψ^{*A}):

$$\begin{aligned} \phi^{*A} - \xi 2\pi \le \phi \le \phi^{*A} + \xi 2\pi \\ \theta^{*A} - \xi 2\pi \le \theta \le \theta^{*A} + \xi 2\pi \\ \psi^{*A} - \xi 2\pi \le \psi \le \psi^{*A} + \xi 2\pi, \end{aligned} \tag{22}$$

where ξ is a parameter that we set to 0.1. Note that, in the case where Phase A resulted in a "no passage" result, there is no special reason why the actual global solution is in the vicinity of a randomly generated instance. Therefore, we are required to re-examine the complete sorbate rotation space, and relax the above constraints by setting $\xi = 0.5$ (so as Eqs. 22 to be equivalent with Eqs. 7).

The minimization of activation energy (Eq. 12) shall be subject to a set of equality and inequality constraints that come from the analysis of "Synopsis of Rigid Portal Formulation" and "Formulation Enhancements" Section. In particular, the formulation includes the constraints of Eqs. 4,6,8,11,13,15–22.

We solve this optimization problem locally starting from 5^3 different initial points, selected from from the elements of the set

$$\Omega = \left\{ \phi - \phi^{*A}, \theta - \theta^{*A}, \psi - \psi^{*A} \in (-\xi 2\pi, -\xi\pi, 0, +\xi\pi, +\xi 2\pi) \right\} \tag{23}$$

The translation variables are initialized at their Phase A optimal values, the host's atom coordinates are initialized at their nominal values, and the δ–variables are initialized at their lower bounds.

The best solution out of all runs, E^*, corresponds to the final result. This activation energy ought to be less than or equal to the strain energy that corresponds to the rigid portal (Phase A) solution. This holds true because the updated model that takes into account portal flexibility is a relaxation (from a mathematical programming point of view) of the original model. The rigid portal formulation corresponds to a particular instance of the updated formulation when $xhv_j = xh_j, yhv_j = yh_j$, $\forall j$. Also, note that when Phase A returns a "constrained passage" result, then Phase B cannot be infeasible since it is guarranteed to admit at least one feasible point (the solution of Phase A, while host atom positions are in their nominal values).

Computational results

The algorithm of Figure 4 was implemented in GAMS (Brooke et al., 2005). The results of Phase A were readily available to us from our work in Gounaris et al. (2006a,b). The NLP optimization solver used for Phase B was CONOPT3 (Drud 1996). Initialization of the δ–variables at their lower bounds increases the chances of starting the optimization procedure from a feasible point and, therefore, assists the solver in locating the solution faster.

The methodology was applied on the same data sets that were used in Gounaris et al. (2006a,b). These consist of a collection of 38 molecules (Table A1, entries typed in boldface) and a zeolite structure database that consists of 123 different structures that correspond to 217 different portals (Table B1).

The initial three-dimensional conformations, (x_i^0,y_i^0,z_i^0), of the molecules were computed with use of the HyperChem software (Hypercube, 1996). The semi-empirical setup with the AM-1 method was utilized. The geometry refinement used the Polak-Ribiere (conjugate gradient) method down to an RMS gradient of 0.01 kcal/molÅ. This optimized configuration has the lowest possible energy and corresponds to the state at absolute zero temperature. As effective radii r_i,r_j, we used the van der Waals radii. The exact values used are presented in Table 1. The nominal distances between the various host atom centers, $d^{\mathrm{TO,nom}}$ and $d^{\mathrm{OO,nom}}$, were easily computed from the coordinates $\begin{bmatrix} xh_j \\ yh_j \end{bmatrix}$ that are part of the structure database. The values used for the parameters of the

Table 1. Van der Waals Radii Used (in Å)

Atom	H	C	N	O	F	S	U
r_{vdw}	1.20	1.70	1.50	1.40	1.35	1.85	1.86

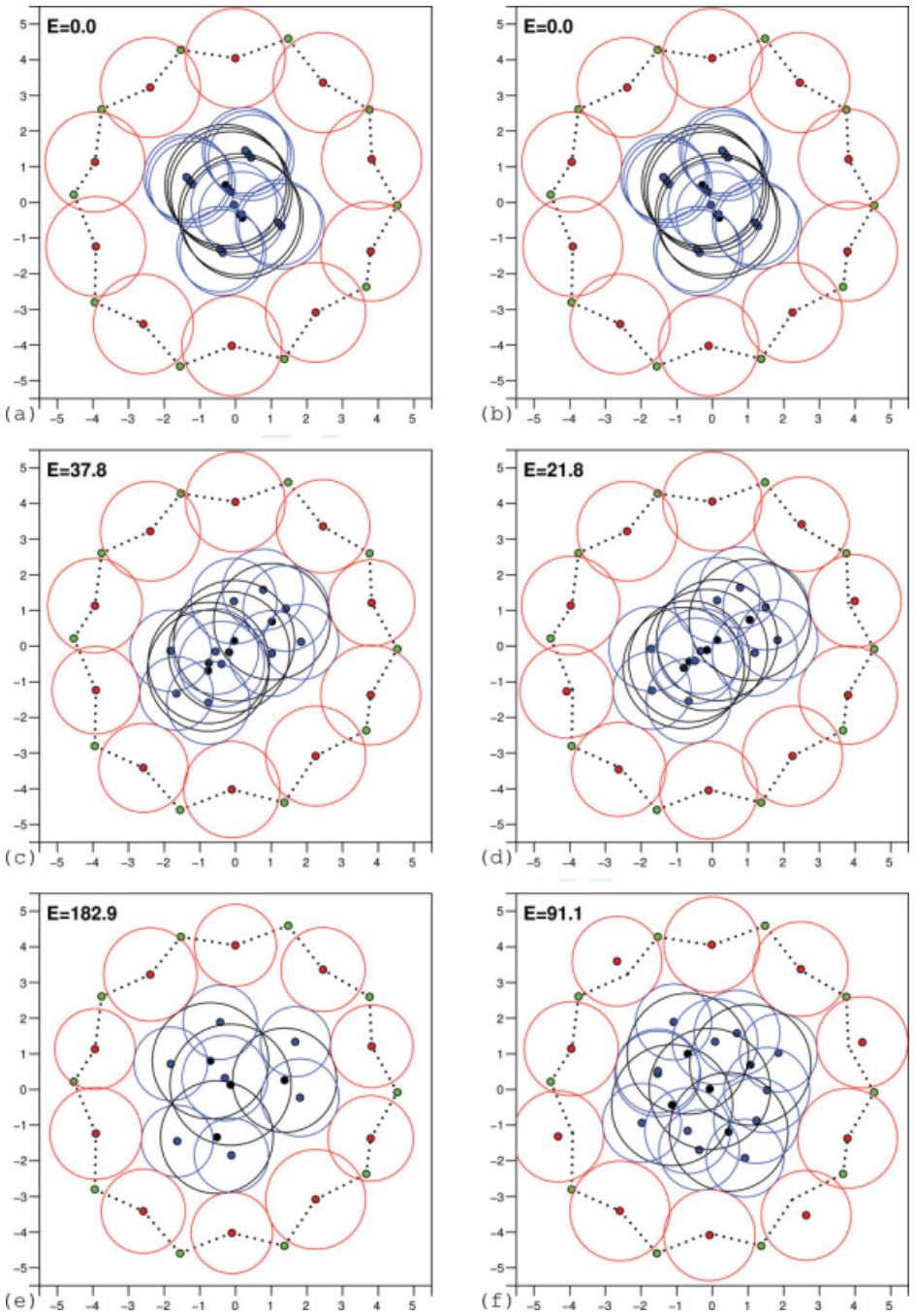


Figure 5. Pentane isomers vs. MFI2 portal: (a,b) *n*-C_5; (c,d) iso-C_5; (e,f) neo-C_5; (a,c,e) rigid; (b,d,e) flexible.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

lattice relaxation potential were k_{TO} = 2,092 kJ/mol$Å^2$ and k_{OO} = 430 kJ/mol$Å^2$, as have been reported by Santikary and Yashonath (1994)[†] Note that these parameter values have been specifically proposed for a natrolite framework. However, given the scarcity of relevant experimental data, we are using these values on a universal basis. A suitable sensitivity analysis shall, therefore, be conducted to gain better understanding of the robustness of these results.

We present here the example of the separation of pentanes using ZSM–5 (MFI). Figure 5 depicts the relevant illustrations for one of the rings, MFI2[‡]. Note that these depictions include also the positions of the Si cations that serve as T–atoms on the ZSM–5 framework and link the O–atoms of the windows. No radii are drawn around these cations, as these are relatively small and effectively do not contribute to the interactions. It can be seen that the n–pentane/MFI2

[†]For a typical silicalite having a T—O bond length of ℓ_0 = 1.610 Å and an O—T—O bond angle of θ_0 = 109.47 Å (Sartbaeva et al., 2006), this k_{OO} value translates to a k_θ value of 0.113 kJ/mol.

[‡]The illustrations for the other ring, MFI1, are very similar and their depiction is omitted since it does not provide additional insight.

Figure 6. Activation energy map for the case of rigid portals (energy levels in kJ/mol).

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

system corresponded to a "free passage" result ($E = 0$ kJ/mol) in the case of a fixed portal. This result is maintained in the case where the portal is allowed to relax (as a matter of fact, no Phase B run is necessary to infer this). In the case of iso–pentane, a rigid portal system would exhibit an activation energy of 37.8 kJ/mol. Under flexible portal conditions, though, some oxygen atoms on the zeolite side have been pushed back to avoid further squeezing of the soft–spherical atoms. This has occured in an energetically favorable way, so as the overall activation energy to drop down to 21.8 kJ/mol when flexibility is allowed. This latter number reflects the combined contributions of the two objective function components. Roughly 40% of it is due to window flexibility, leaving the remaining 60% to be due to distortion of the atoms involved. The case of neo–pentane is similar. In fact, the effect of portal relaxation is more severe and can be more easily observed in the illustrations. The energy drops from 182.9 kJ/mol to 91.1 kJ/mol, with the split between strain and window flexibility to be about 35/65(%). Note that the passage of iso– and neo–C_5 is highly "constrained." The activation energies are relatively high, even for the case of flexible portals, and practically no meaningful equilibrium concentration can be achieved within the pores of ZSM–5. Therefore, ZSM–5 is a structure that we would expect to be selective in the separation of n–C_5 from its two isomers, a result that is in good agreement with commercial experience.

The complete activation energy map of results is presented in Figure 7. For comparison purposes, we present first Figure 6 that depicts the equivalent map in the case of rigid portals. This map was calculated from the results of Gounaris et al. (2006a,b) and corresponds to the case of no lattice flexibility ($k_{TO} = k_{OO} \to \infty$). Finally, as a matter of sensitivity analysis, we also present Figure 8. This corresponds to the case of very flexible portals, as it is reflected by a tenfold decrease in the strength of the harmonic potential ($k_{TO} = 209.2$ kJ/molÅ^2 and $k_{OO} = 43.0$ kJ/molÅ^2).

Given the uncertainty on the actual degree of flexibility of the zeolitic frameworks and the possibility that this differs from structure to structure, one cannot expect that the results of Figure 7 are highly accurate. However, one can assume with confidence that the actual numbers should lie somewhere between those presented in Figure 7 and 8. We have effectively bounded from above and below the actual values, and provided a reliable range for quantities such as activation energy or equilibrium concentration.

In an effort to illustrate the differences between these figures, we present Table 2. Its entries correspond to average activation energy reductions[§] from the rigid portal case for a number of different molecule and window sets. We observe that portal flexibility becomes crucial when small windows are involved, as there is a greater potential for framework relaxation. In contrast, the case of 14–ring windows remains practically unaffected, as the majority of the systems correspond to "free passage" results. The reduction in activation energy translates into an increase in the achieved equilibrium concentration. The average increases[¶] are listed in Table 3 (for ambient temperature). We note that the effect on actual equilibrium concentration is moderate since, in many cases, the reduction of activation energy is not sufficient to facilitate a considerable increase in the level of penetration.

[§]Defined as $\Delta_r E = (E^{rig} - E^{flex})/E^{rig}$.

[¶]Defined as $\Delta \frac{C}{C_0} = \left(\frac{C}{C_0}\right)^{flex} - \left(\frac{C}{C_0}\right)^{rig}$.

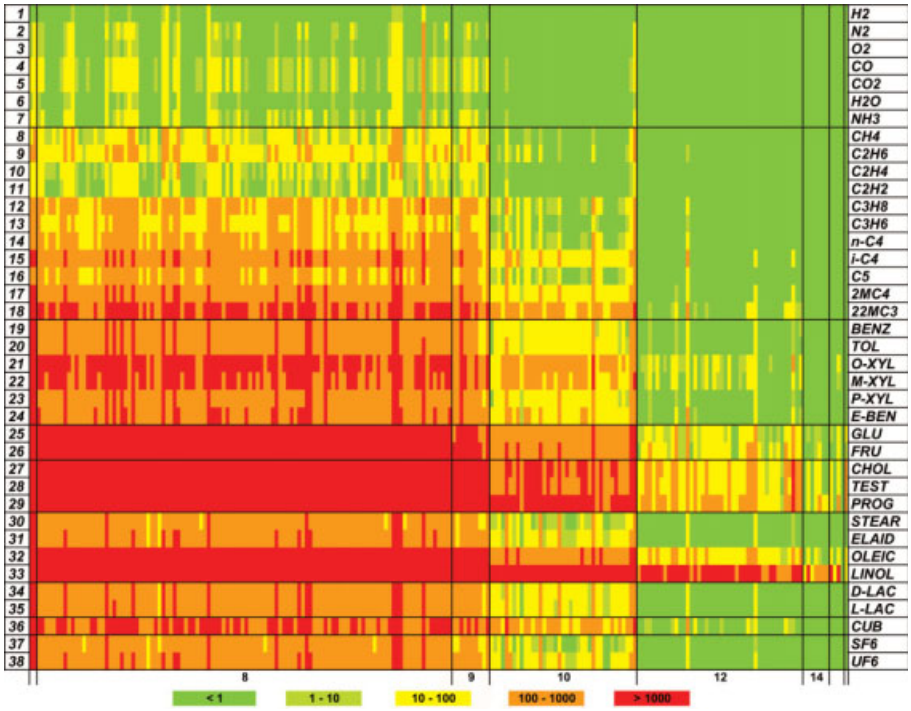


Figure 7. Activation energy map for the case of flexible portals (energy levels in kJ/mol).

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Finally, we present Figure 9 that depicts the extent of activation energy reductions that apply when portal flexibility is taken into account. Energies higher than the cut–off value of $E = 20$ kJ/mol are not depicted, as they typically lead to insignificant equilibrium concentrations. We also present Figure 10 that depicts the cases of propane and propylene over a wider range of activation energies. We observe that when portal flexibility is taken into account, the actual activation

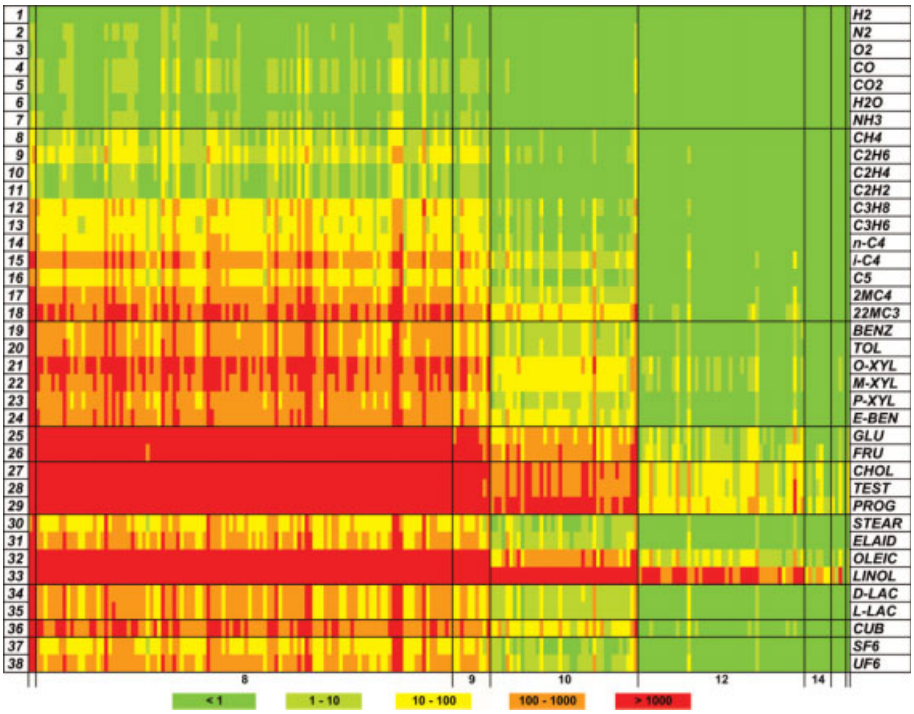


Figure 8. Activation energy map for the case of very flexible portals (energy levels in kJ/mol).

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Table 2. Average Activation Energy Reduction As a Function of Window Size

Molecule Set		8	9	10	12	14
Light gases	1–7	17% (37%)	27% (45%)	1% (2%)	0% (0%)	0% (0%)
Chain H/C	8–18	66% (87%)	71% (90%)	26% (48%)	2% (5%)	0% (1%)
Aromatic H/C	19–24	90% (94%)	89% (96%)	59% (91%)	14% (28%)	0% (0%)

Case of very flexible portals in parentheses.

Table 3. Equilibrium Concentration Increase As a Function of Window Size

Molecule Set		8	9	10	12	14
Light gases	1–7	0.019 ± 0.048	0.078 ± 0.200	0.001 ± 0.008	0.000 ± 0.000	0.000 ± 0.000
Chain H/C	8–18	0.009 ± 0.043	0.012 ± 0.051	0.012 ± 0.036	0.002 ± 0.014	0.001 ± 0.006
Aromatic H/C	19–24	0.000 ± 0.000	0.000 ± 0.000	0.006 ± 0.025	0.021 ± 0.088	0.000 ± 0.002

Entries correspond to $\mu \pm 1\sigma$. Results at $T = 300$ K.

energies required for penetration are only one-third of what would have been predicted by the rigid portal approach (one-tenth if portals are considered to be very flexible).

Entropic Contributions of Locally Optimal Conformers

In the earlier approaches (either with or without flexibility of portal), the main aim was to extract the globally optimal solution of the applicable optimization model. This solution was subsequently used for the calculation of quantities of interest such as the equilibrium concentration at a given temperature, or the selectivity between two molecules. The whole approach inherently relies on the assumption that every individual molecule that is eventually adsorbed into the pores of the guest structure will penetrate the pore opening(s) having attained a particular three dimensional conformation that corresponds to the globally minimum activation energy. However, there can very well be individual molecules that eventually get admitted into the pores through a different transition conformation. This conformation would be one that corresponds to some energetic local minimum that is different than the global one and, thus, corresponds to a higher value of activation energy.

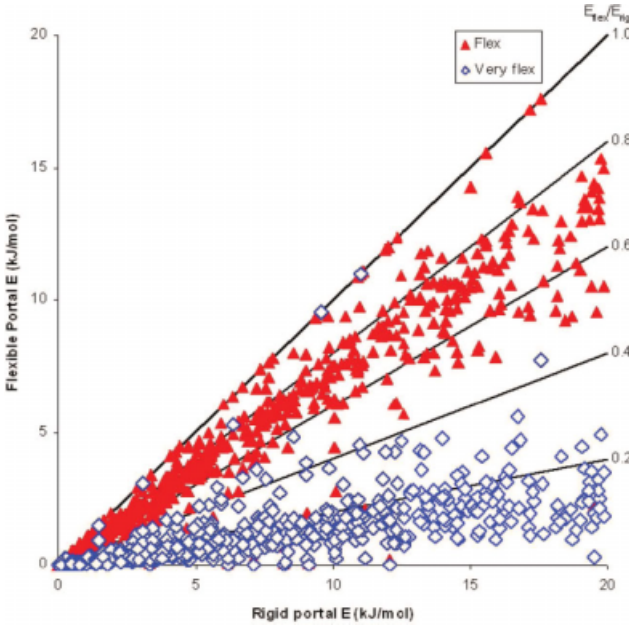


Figure 9. Rigid vs. flexible portals: Effect on activation energy for all the molecules used in this study.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Given the complexity of the systems involved, it is not uncommon for the optimization model to exhibit multiplicity of solutions. An appropriate correction needs to be employed that accounts for the possibility that a molecule will penetrate a guest portal with an energetically local optimum conformation. To illustrate this, we present Figure 11 that depicts four different locally optimal conformations for the case of methane approaching the opening of the deca–dodecasil 3R (DDR) zeolite. We observe that even for the case of a relatively simple and symmetrically shaped molecule such as methane, there occur transition conformations that are energetically stable, but correspond to energy barriers larger than the smallest possible value corresponding to the global optimum. If we only took into account the global solution, the strain index** for this system would have been SI = 0.81, and the ambient T equilibrium concentration $C/C_0 \simeq 0.003$ (according to Eq. 9). Assuming that each of the four locally optimal conformations had equal probability of occuring, the observed equilibrium concentration would be only $C/C_0 \simeq 0.001$, as each of the four conformations would have contributed one fourth of the final result.

Having enumerated the various possible local solutions, we need to take all of them into account by weighting them appropriately. To do this, we need to know their frequencies of occurence; that is, how often do molecules penetrate the opening in a particular locally optimal conformation. The probability of penetration by each of these possible conformations can be postulated, estimated, or computed with various methodologies. We will address this issue in the next Section, where two different probability measures will be introduced. Once an appropriate probability measure has been adopted, the respective probabilities can serve suitably as weights in the calculation of expectations of quantities of interest, such as the equilibrium concentration:

$$\left\langle \frac{C}{C_0} \right\rangle = \sum_{\ell=1}^{L} p_\ell \left(\frac{C}{C_0} \right)_\ell = \sum_{\ell=1}^{L} p_\ell \exp\left(-\frac{E_\ell}{RT} \right), \qquad (24)$$

where L is the number of distinct optimal conformations, p_ℓ is the probability that a molecule will penetrate with conformation ℓ,

**The strain index is defined as SI = $\log(1 + \langle S^* \rangle)$ (Gounaris et al., (2006a).

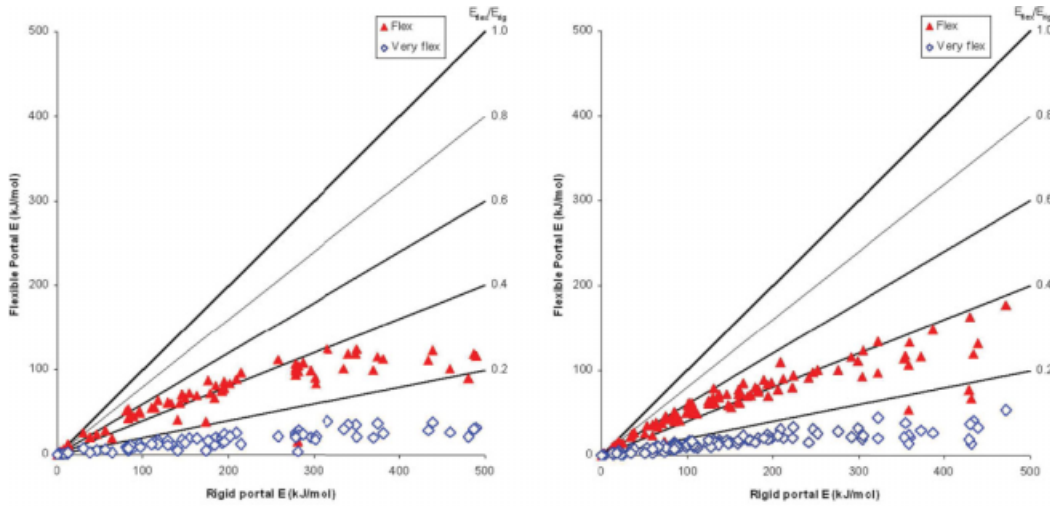


Figure 10. Rigid vs. flexible portals: Effect on activation energy for the C_3/C_3 = system.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

E_ℓ is the associated energetic barrier that has to be overcome, and $\left(\frac{C}{C_0}\right)_\ell$ is the corresponding equilibrium concentration.

The reason behind performing the expectation operation on $\frac{C}{C_0}$, instead of doing so directly on E, lies in the result by Jarzynski (1997), which dictates that the actual equilibrium free energy difference (and, thus, the macroscopically observed equilibrium concentration) can be computed by this weighted average.

Probability of occurrence of transition conformers/local optima

Each locally optimal solution (distinguished from the other ones in terms of activation energy E) corresponds to a set of transition conformers that are all characterized by this activation energy. The purpose of this analysis is to derive the conditional probability that a molecule will penetrate the opening as a transition conformer that belongs to a particular set, given that it will indeed penetrate the opening. This probability is appropriate for use as a factor in a weighted sum of contributions by all possible locally optimal solutions, since it is proportional to the frequency of occurence of the corresponding transitions.

Let $\mathcal{R}$ be the full conformational space (ϕ,θ,ψ,xt,yt) for the guest molecule, and let $\mathcal{R}_F \subseteq \mathcal{R}$ be the subspace of initial points (states) that lead to feasible optimization runs (local energy minimization used here in the context of representing a physical path towards a transition). Let also index $\ell = 1,2,\cdots,L$ denote the various different local optima (sets of transition conformers), and $\mathcal{R}(\ell)$ be the corresponding regions of attraction. Obviously,

$$\mathcal{R}(\ell) \subseteq \mathcal{R}_F \quad \forall \ell, \qquad \mathcal{R}(\ell_1) \cap \mathcal{R}(\ell_2) \equiv \varnothing \quad \forall \ell_2 \neq \ell_1, \quad \text{and} \quad \bigcup_{\ell=1}^{L} \mathcal{R}(\ell) \equiv \mathcal{R}_F.$$

The probability calculation makes use of the Bayes formula and conditioning on the initial state. The rigorous derivation has as follows:

$$\begin{aligned}
p_k &= \mathbb{P}\left\{ \begin{array}{c} \text{“molecule penetrates as} \\ \text{a conformer of set } k\text{”} \end{array} \middle| \begin{array}{c} \text{“molecule does} \\ \text{penetrate”} \end{array} \right\} \\
&= \frac{\mathbb{P}\left\{ \begin{array}{c} \text{“molecule penetrates as} \\ \text{a conformer of set } k\text{”} \end{array} \cap \begin{array}{c} \text{“molecule does} \\ \text{penetrate”} \end{array} \right\}}{\mathbb{P}\left\{ \begin{array}{c} \text{“molecule does} \\ \text{penetrate} \end{array} \right\}} = \frac{\mathbb{P}\left\{ \begin{array}{c} \text{“molecule penetrates as} \\ \text{a conformer of set } k\text{”} \end{array} \right\}}{\mathbb{P}\left\{ \begin{array}{c} \text{“molecule does} \\ \text{penetrate”} \end{array} \right\}} \\
&= \frac{\int_{i\in\mathcal{R}} \mathbb{P}\left\{ \begin{array}{c} \text{“molecule penetrates as} \\ \text{a conformer of set } k\text{”} \end{array} \middle| \begin{array}{c} \text{“initial state} \\ \text{is } i\text{”} \end{array} \right\} \mathbb{P}\left\{ \begin{array}{c} \text{“initial state} \\ \text{is } i\text{”} \end{array} \right\}}{\int_{i\in\mathcal{R}} \mathbb{P}\left\{ \begin{array}{c} \text{“molecule does} \\ \text{penetrate”} \end{array} \middle| \begin{array}{c} \text{“initial state} \\ \text{is } i\text{”} \end{array} \right\} \mathbb{P}\left\{ \begin{array}{c} \text{“initial state} \\ \text{is } i\text{”} \end{array} \right\}} \\
&= \frac{\int_{i\in\mathcal{R}} \mathbf{1}\left\{ \begin{array}{c} \text{“local minimization run starting from} \\ \text{point } i \text{ leads to local optimum } k\text{”} \end{array} \right\} \mathbb{P}\left\{ \begin{array}{c} \text{“initial state} \\ \text{is } i\text{”} \end{array} \right\}}{\int_{i\in\mathcal{R}} \mathbf{1}\left\{ \begin{array}{c} \text{“local minimization run starting from} \\ \text{point } i \text{ is feasible”} \end{array} \right\} \mathbb{P}\left\{ \begin{array}{c} \text{“initial state} \\ \text{is } i\text{”} \end{array} \right\}} \\
&= \frac{\int_{i\in\mathcal{R}(k)} \mathbb{P}\left\{ \begin{array}{c} \text{“initial state} \\ \text{is } i\text{”} \end{array} \right\}}{\int_{i\in\mathcal{R}_F} \mathbb{P}\left\{ \begin{array}{c} \text{“initial state} \\ \text{is } i\text{”} \end{array} \right\}} = \frac{\int_{i\in\mathcal{R}(k)} \pi_i}{\int_{i\in\mathcal{R}_F} \pi_i} \approx \frac{\sum_{i\in\mathcal{R}(k)} \pi_i}{\sum_{i\in\mathcal{R}_F} \pi_i}
\end{aligned}$$

where π_i is the probability that the molecule's initial state is state i.

The final step provides a discretization approximation that is necessary in an actual implementation, where sampling is finite. The accuracy of this approximation is improved with finer sampling resolution. Note that this sampling is required to be uniform, in order not to include spatial bias in the calculations.

We have therefore derived an expression for p_k, the overall probability that an adsorbed molecule penetrated the opening as a conformer of the set k; that is, after having confronted a particular energetic barrier E_k. This probability depends on π_i, the probability that the process initiated from one of the associated initial states i.

The first way to calculate such probabilities is through the assumption that every initial state has equal probability of occuring. This leads trivially to a probability that is proportional to the number of initial (sampled) points that resulted into the particular local solution. Let f_k be the frequency of occurence of the kth local optimum, and N_k be the number of feasible local minimization runs that resulted into this particular solution. The following uniform distribution applies:

$$p_k = f_k = \frac{N_k}{\sum_{\ell=1}^{L} N_\ell} \tag{25}$$

The second way for the calculation of these probabilities is to assume that higher energy initial states have a smaller probability of occuring than lower energy ones. In this case, we can define a partition function Z_k for each state k, and the probability is given by the following Gibbs–Boltzmann distribution:

$$p_k = \frac{Z_k}{\sum_{\ell=1}^{L} Z_\ell} = \frac{\sum_{n=1}^{N_k} \exp\left(-\frac{E_n^0}{RT}\right)}{\sum_{n=1}^{N_F} \exp\left(-\frac{E_n^0}{RT}\right)}, \tag{26}$$

where N_F is the total number of initial points that led to a feasible optimization run, n is an index that spans over the appropriate subset of initial points, and E_n^0 is some energy level associated with the corresponding initial point.

Calculating the effect of multiplicity of local solutions

A suitable methodology has been developed to incorporate into the overall framework the multiplicity of transition conformations that a guest molecule can explore during the process of adsorbing into the pores of a host structure. The efficient algorithms developed for the cases of rigid (Gounaris et al., 2006b) and flexible ("Algorithmic enhancements" Section) portal, in conjunction with access to powerful distributed computing machinery, has allowed us to follow an approach that relies on exhaustive sampling of the molecule's conformational space, in terms of both orientation and translation, and using these points to initialize a series of local energy minimization runs. The associated probability of the initial conformation is allotted to the corresponding solution, until all runs have been completed. The results can then be post–processed for the calculation of the expected equilibrium concentration $\left\langle \frac{C}{C_0} \right\rangle$. The core of the algorithm consists of two phases, which are described below in detail. The overall flow diagram, including the post–processing phase, is depicted in Figure 13.

Phase 1 — Selection of Initial Point and Feasibility Check. The selection of an initial point is done in a manner consistent with an overall initialization strategy that aims at finely sampling the complete conformational space of a guest molecule which is about to enter the portal. Although the strategy shall not exclude any initial points that can potentially lead to feasible solutions, a practical implementation shall exclude those points that have no reasonable chance of leading to penetration, like points very far away from the portal or orientations completely misaligned with the axis of the porous channel. As an unambiguous rule, we consider all initial points that correspond to a "no passage" result (that is, a projection that lies outside the portal, even for minimum values of δ) as points that shall be excluded from consideration. The conformational space of the molecule includes the three Euler angles (ϕ,θ,ψ), that define its orientation, and the two coordinates (xt,yt), that define its position with repect to the position of the window[††]. Therefore, an initial point n corresponds to the vector $(\phi_n^0,\theta_n^0,\psi_n^0,xt_n^0,yt_n^0)$.

For each of the three angles, one out of N_{or} values from their full range $(-\pi,\pi]$ is selected. As the sampling is required to be spatially uniform, so as not to include a relevant bias in the calculations, these values are equally far apart and selected from the elements of the set

$$\Omega_{or} = \left\{ \phi_n^0, \theta_n^0, \psi_n^0 \in \left(-\pi + \frac{i}{N_{or}} 2\pi \ , \ i = 1, 2, \cdots N_{or} \right) \right\}. \tag{27}$$

For xt_n^0 (and similarly for yt_n^0), one out of N_{tr} values in the range (xh_{min},xh_{max}) is selected from the elements of the set

$$\Omega_{tr} = \left\{ xt_n^0 \in \left(-xh_{min} + \frac{i}{N_{tr}+1}(xh_{max} - xh_{min}), \right. \right. \\ \left. \left. i = 1, 2, \cdots N_{tr} \right) \right\}. \tag{28}$$

The values xh_{min} and xh_{max} correspond, to the minimum and maximum value of the host O–atom coordinates, respectively, that is, $xh_{min} = \min_{j \in J_O} xh_j$ and $xh_{max} = \max_{j \in J_O} xh_j$. Clearly, for any molecule at any possible orientation, (xt,yt) values outside these bounds would result to initial points that would have been excluded from further consideration.

Once an initial point has been selected, a feasibility check needs to be performed so as to assess whether the selected conformation corresponds or not to a "no passage" result. Fixing the model's variables according to

$$\phi = \phi_n^0 \ , \ \theta = \theta_n^0 \ , \ \psi = \psi_n^0 \tag{29}$$

$$xt = xt_n^0 \ , \ yt = yt_n^0 \tag{30}$$

$$\delta_i = \delta^L \ \forall i \ , \ \delta_j = \delta^L \ \forall j, \tag{31}$$

[††]For consistency purposes, the atom coordinate databases have been shifted so that both the molecules' centers of mass and the windows' centers lie at point (0,0).

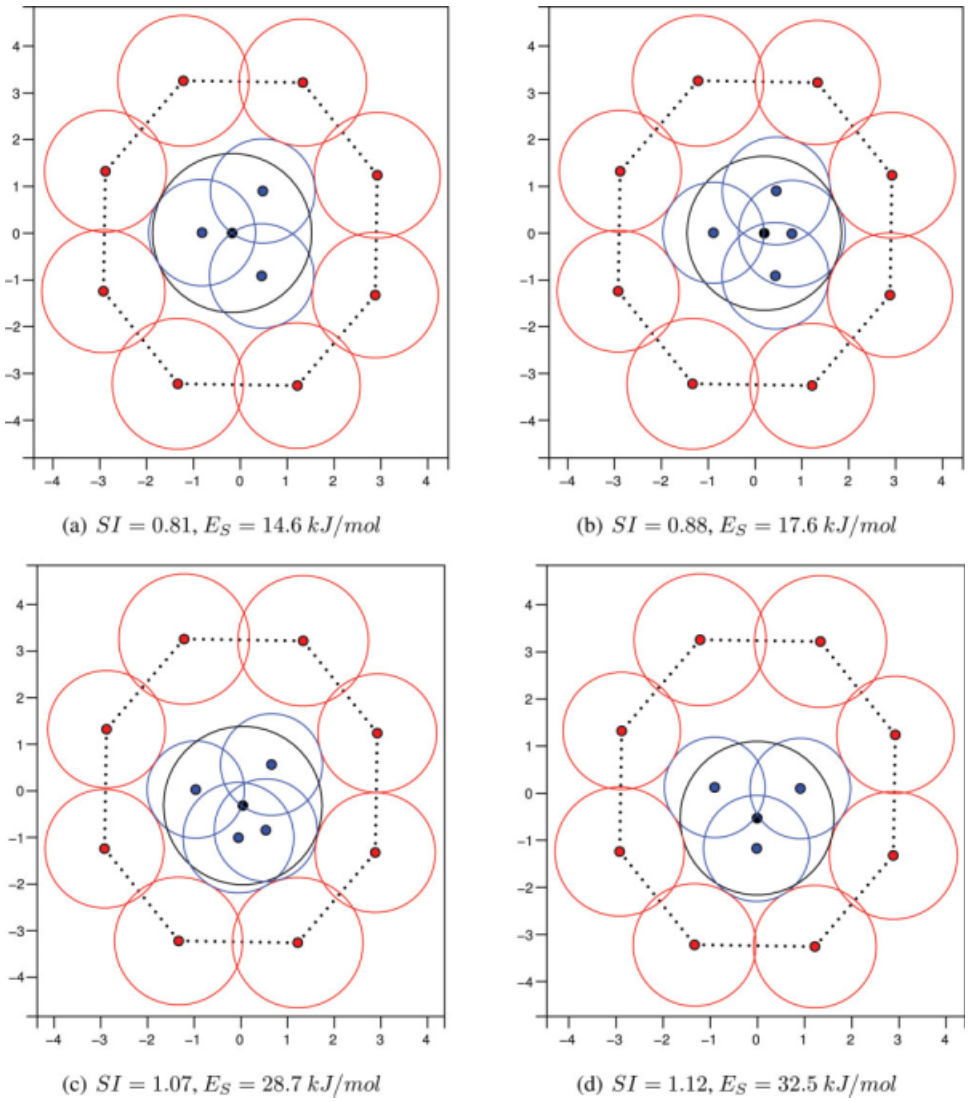


Figure 11. Four distinct locally optimum orientations when CH_4 is approaching DDR.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

the feasibility test corresponds to checking whether the constraints of Eqs. 4–6 are satisfied. Note that the δ–variables are set at their lower bounds, since this instance is the most likely to be feasible (if it is not, then no other would). In the case of flexible portals, the window may be fixed to its nominal shape, or some other relaxed instance that satisfies Eqs. 19–21.

Figure 12 depicts an (xt,yt)–grid, as it is defined by the set Ω_{tr}. As Eq. 6 is part of the feasibility check, each grid point that lies outside the window is automatically discarded. Note that each dot corresponds to a set of N_{or}^3 initial points that share the same translation but have different orientation, and it may happen that only a subset of those (or even none) is actually valid.

If the feasibility check is successful, the initial point is deemed valid, and the algorithm moves on to Phase 2. At this point, one may also calculate an initial energy E_n^0 that is associated with this initial point. This energy could be used in determining the point's contribution to the respective probability p_k through a partition function Z_k. A plausible selection would be to use $E_n^0 = \min_{\delta_i,\delta_j} 4\varepsilon S$, subject to the constraints of Eqs. 4–6, 8, 29, 30. Note that this constitutes a convex optimization formulation that is guarranteed to be feasible.

If the feasibility check fails, the initial point is discarded, the next one is selected, and the algorithm repeats Phase 1. Obviously, if all initial points have been discarded, the algorithm concludes with a "no passage" result and an equilibrium concentration of $\langle \frac{C}{C_0} \rangle = 0$.

Phase 2 — Calculation of E_n and Identification of Solution. The initialization scheme described in Phase 1 provides a feasible initial point (defined by Eqs. 29–31) for the local optimization run that constitutes this phase. The second phase of either the rigid (Gounaris et al.2006b) or the flexible portal ("Algorithmic enhancements" Section) algorithm may be used. Since the optimization problem has at least one feasible point (the initial one), the run oughts to be feasible and terminate at a local optimum. The optimal

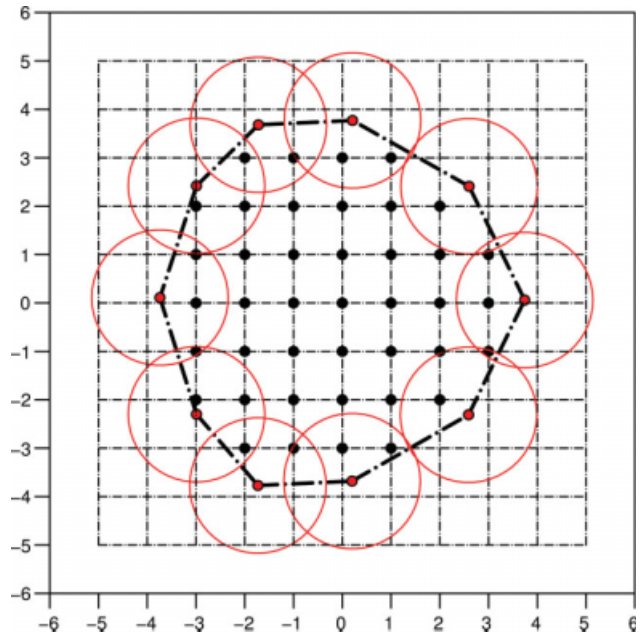


Figure 12. Exhaustive gridding and potentially valid initial points.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

solution is the required activation energy E_n (or, alternatively, $E_n = 4\varepsilon S^*$ if it is the total strain that we minimize).

As we go through multiple runs of Phase 2, starting from different feasible initial points selected by Phase 1, we keep a list of the distinct (in terms of energy) solutions that we encounter. If the solution has already been observed in a previous run, then the associated bin is updated (i.e., increase the corresponding counter, allocate the run's contribution to the partition function, etc.). Otherwise, a new bin is generated and the new solution becomes its first member.

At the end of the process, we have compiled a list of L different local solutions, that correspond to different activation energies E_ℓ. Each of these solutions has been observed N_ℓ times, and has an associated partition function Z_ℓ. Depending on the probability measure we have opted to use (Eq. 25 or Eq. 26), the respective probabilities p_ℓ can now be calculated and used as weights in the expectation calculation of the equilibrium concentration.

One question that arises in an implementation is how finely to perform the sampling process; that is, what values to select for parameters N_{or} and N_{tr}. A plausible strategy would be to perform a sequence of calculations, each time taking into account an increased number of initial points (finer sampling). The process should stop when two or more consecutive iterations provided practically the same results. This is an indication that the effect of discretization has been eliminated, and the sampling process constitutes a good approximation of the infinite case.

Computational results

The algorithm of Figure 13 was implemented in GAMS (Brooke et al., 2005). The NLP optimization solver used for the calculation of E_n^0 in Phase 1 was MINOS (Murtagh and Saunders, 1987), while the calculation of E_n in Phase 2 was performed by CONOPT3 (Drud, 1996). The initialization used a grid of size $(N_{or}, N_{tr}) = (6,11)$. This was deemed fine enough, as the calculations performed with a grid of size (5,9) had produced similar results (only 2% average difference).

The methodology was applied on the same zeolite structure database as before (Table B1), in conjunction with an extended guest molecule database that consists the 290 entries listed in Table A1. The latter is a very comprehensive database that includes a wide range of molecules: from light gases such as H_2 and *He* to large molecules such as the buckminsterfullerene ball, the drug paclitaxel, and the protein insulin. Note that we consider a total of 290 × 217 =

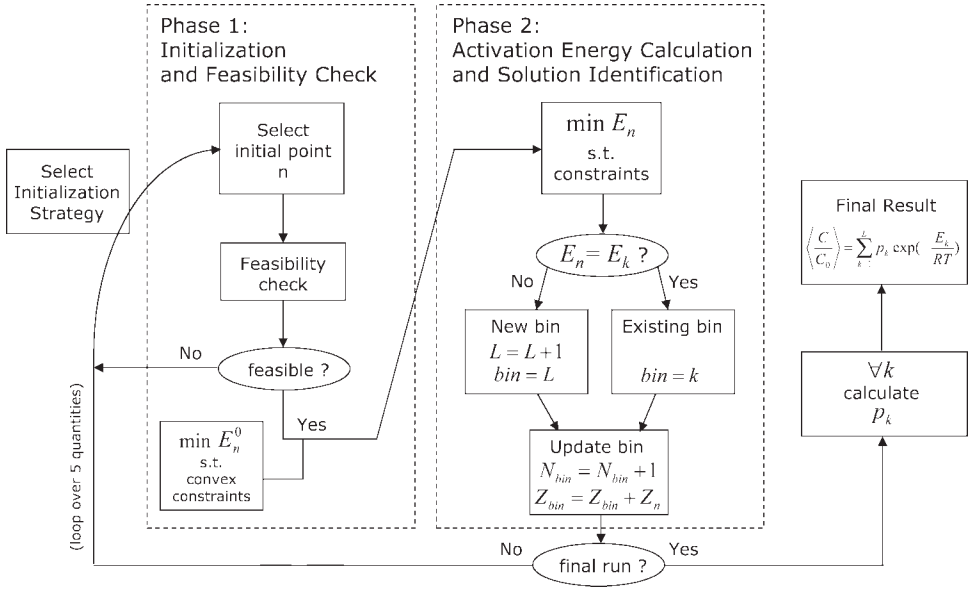


Figure 13. Overall algorithm for the calculation of expected equilibrium concentration $\left\langle \frac{C}{C_0} \right\rangle$.

Table 4. Expectation Calculation for the Equilibrium Concentration of Ethane in AEL at 300 K

ℓ	SI_ℓ	E_ℓ	$\left(\frac{C}{C_0}\right)_\ell$	N_ℓ	$p_\ell^{(\alpha)}$	$\left(p^{(\alpha)}\frac{C}{C_0}\right)_\ell$	Z_ℓ	$p_\ell^{(\beta)}$	$\left(p^{(\beta)}\frac{C}{C_0}\right)_\ell$
1	0.00	0.00	1.000	5,851	0.591	0.591	0.329	0.078	0.078
2	0.01	0.06	0.977	18	0.002	0.002	0.000	0.000	0.000
3	0.03	0.18	0.931	18	0.002	0.002	0.000	0.000	0.000
4	0.04	0.24	0.908	22	0.002	0.002	0.000	0.000	0.000
5	0.07	0.44	0.839	161	0.016	0.014	3.856	0.920	0.772
6	0.08	0.51	0.817	425	0.043	0.035	0.000	0.000	0.000
7	0.09	0.58	0.794	421	0.043	0.034	0.000	0.000	0.000
8	0.10	0.65	0.772	528	0.053	0.041	0.000	0.000	0.000
9	0.11	0.72	0.749	1,417	0.143	0.107	0.008	0.002	0.001
10	1.33	50.89	0.000	398	0.040	0.000	0.000	0.000	0.000
11	1.35	53.41	0.000	399	0.040	0.000	0.000	0.000	0.000
12	1.43	64.72	0.000	125	0.013	0.000	0.000	0.000	0.000
13	1.44	66.28	0.000	111	0.011	0.000	0.000	0.000	0.000
$L = 13$			Sums:	9,894	1.000	0.828	4.192	1.000	0.852

α, based on uniform distribution; β, based on Gibbs–Boltzmann distribution.
Numbers calculated to higher precision than presented. E_ℓ in kJ/mol.

62,930 host/guest pairs and, given the selected initialization scheme, we conducted roughly 527 million local optimization runs.

Let us focus on the illustrative example of ethane approaching AlPO–11 (AEL). According to the results of the approach based on the globally optimal conformer, this particular host/guest system would correspond to a ''free passage'' result, which means that it would exhibit a —globally minimum— strain index of SI = 0.00, an associated activation energy of $E = 0$ and, therefore, a $T = 300$ K equilibrium concentration of $\frac{C}{C_0} = 1.000$. However, the system exhibits 13 distinct local minima and, according to the enhanced methodology, the globally minimum activation energy would only contribute a portion of the final answer. All of the locally optimal conformations have a chance of being explored by a penetrating molecule and, therefore, all should be weighted appropriately. This calculation is performed in Table 4, for both probability measures introduced in ''Probability of occurence of transition conformers/local optima'' Section. The two final results are $\langle\frac{C}{C_0}\rangle^{(\alpha)} = 0.828$ and $\langle\frac{C}{C_0}\rangle^{(\beta)} = 0.852$, and can be found in the last line where all contributions are added together. We observe that the actual ethane equilibrium concentration achieved within the pores of AlPO–11 would not be $C = C_0$, rather than a concentration smaller by about 16%.

Figure 14 presents the map of expected equilibrium concentrations at $T = 300$ K, as they are calculated using the probability measure based on uniform distribution (Eq. 25), while Figure 15 presents the corresponding results based on the Gibbs–Boltzmann distribution (Eq. 26). Both figures include only the results for a reduced database of 38 molecules so that they can be compared to the results presented in Figure 16. The latter is borrowed from Gounaris et al. (2006b) and corresponds to the globally optimal solutions.

We observe that not taking into account higher activation energy solutions via an expectation calculation results in overestimation of the achieved equilibrium concentration. The average differences[#] over the complete database are 24.6×10^{-3} and 9.6×10^{-3}, respectively for the first and second probability measure. Tables 5 and 6 present average deviations from the global solution for select molecule and window sets. We observe that the deviations resulting from use of the uniform distribution are slightly elevated with respect to their Gibbs–Boltzmann distribution counterparts, however they are all relatively small. The only entries that report some considerable average deviation correspond to the aromatic hydrocarbons over the larger windows.

[#]Defined as $\Delta\frac{C}{C_0} = \left(\frac{C}{C_0}\right)^* - \langle\frac{C}{C_0}\rangle$.

Effective Radii of Portal Oxygen Atoms

The methodology that has been developed attempts to estimate the ability of porous structures to perform as molecular sieves, by effectively comparing the shapes of the apertures with that of the projections of the sorbate molecules. It should be apparent, though, that the effective size of an opening depends critically on the assumed values for the radii of the framework atoms.

Cook and Conner (1999) considered the question of what are the appropriate atomic dimensions for the silicon and oxygen atoms which make up the crystalline lattice. They suggested that the universally employed ionic or van der Waals radii are inappropriate for estimating pore diameters in adsorption studies, and that the effective pore dimensions are considerably larger than conventional analyses would imply. Their suggestion was based on the common experimental observation where molecules diffuse into pores of significantly smaller diameter than expected; such is the case, for example, of cyclohexanes that pass through MFI pores (Cavalcante and Ruthven, 1995; Chon and Park, 1988; Conner et al., 1990; Magalhaes et al., 1998). Although some of these results can be attributed to the flexibility and relaxation of the lattice, this cannot do justice to the large discrepancies observed in some of the cases. The authors concluded that using the so-called Norman radius (Norman, 1976) as the effective one would generally provide more accurate results.

However, this would call for employing an effective radius of just 1.0Å for the zeolite O–atoms, a significant change from the original radius of 1.4Å that has been used routinely. Furthermore, by definition, the Norman radius depends on the overall neighboring structure and, thus, is not the same for all applications. Instead of adopting and using universally a different value for the radii of the portal's O–

Figure 14. Equilibrium concentration database at $T = 300$ K (uniform distribution of initial states).

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

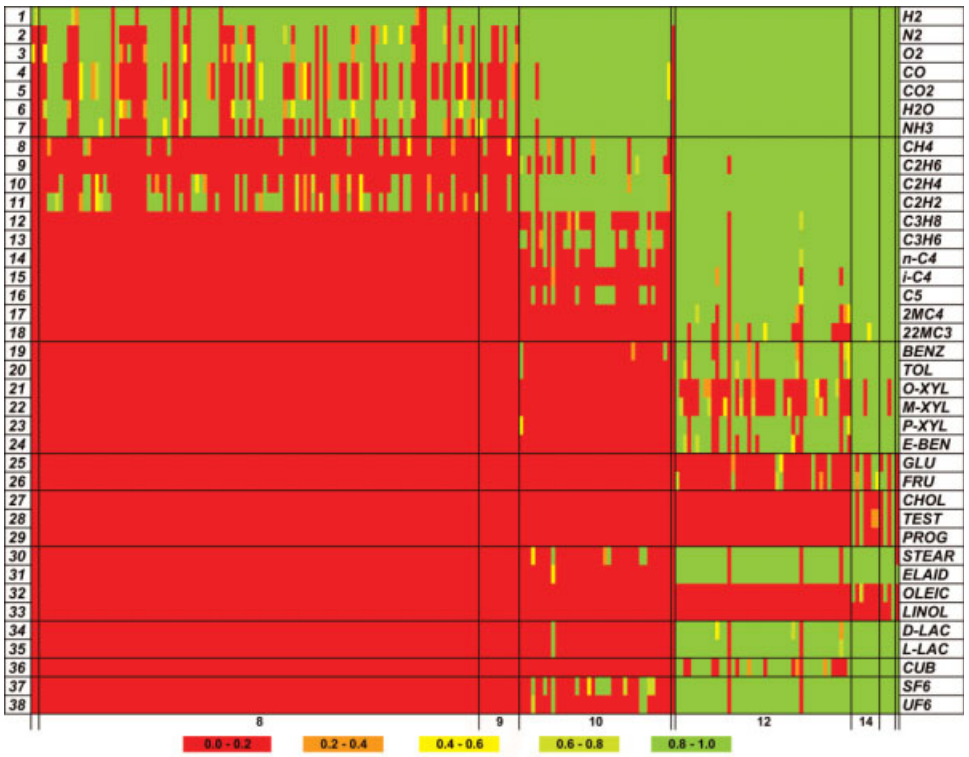


Figure 15. Equilibrium concentration database at $T = 300$ K (Gibbs–Boltzmann distribution of initial states).

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

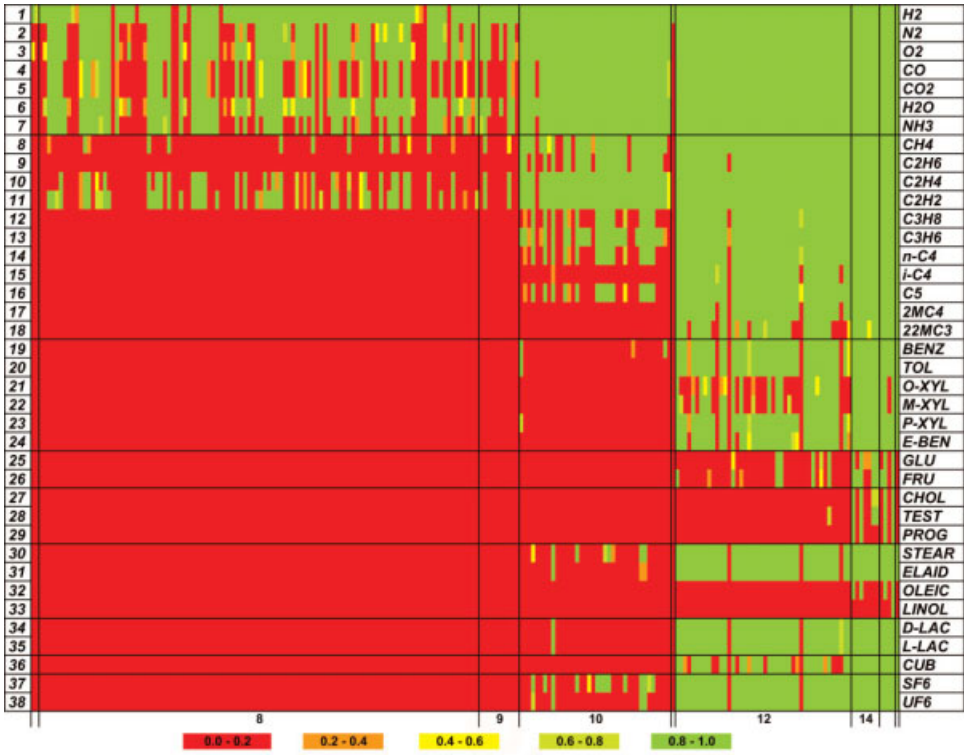


Figure 16. Equilibrium concentration database at $T = 300$ K (globally optimal conformer).

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

atoms, our approach in this section will be to perform a sensitivity analysis so as to quantify how important the dependency of our predictions on the oxygen radius really is.

The case of the ZSM–5 framework

As it was mentioned earlier, the framework of ZSM–5 (MFI) exhibits some discrepancies between actual experimental data, and expected observations. This has been mainly attributed to a potentially smaller O–atom radius on the MFI rings than the commonly used value of 1.40 Å. A sensitivity analysis on this value was performed for the system of the three xylene molecules. The flexible framework methodology was used, and in Figure 17 we present the results in terms of activation energy (global solution) and its associated equilibrium concentration at $T = 600$ K.

We notice that with the O–atoms considered to have an effective radius of 1.40 Å, the pores are relatively small and none of the isomers manages to penetrate to any significant extent; thus, no selectivity. Our method with $r_{\text{eff}} = 1.40$ Å essentially cannot predict the good potential of ZSM–5 for xylene separation, a fact that has been demonstrated extensively on both a laboratory and an industrial scale (Gu et al., 2006). However, the plot suggests that had the actual effective radii been somewhat reduced, the opening would be of the appropriate size to admit *p*–xylene, but not the other two isomers, exhibiting selectivity and potential for separation. We conclude that the appropriate value to use for the O–atom effective radii in the framework of the ZSM–5 zeolite should lie in the range between 0.90 and 1.10 Å.

The case of a titanosilicate framework

Engelhard titanosilicates are mixed oxide microporous molecular sieves that contain titanate structural units, in addition to the tetrahedral silicate units found in conventional zeolitic materials. The presence of titania chains, in which the titanium atom is octahedrally coordinated, provides wide–gap semiconductor properties to these materials (Borello et al., 1997). Catalytic activity has also been observed, with ETS–10 reportedly catalyzing the dehydration of butanol, the isomerization of aromatic compounds, as well as hexane reforming reactions (Das et al., 1996; Philippou et al., 1998a,b).

Table 5. Equilibrium Concentration Difference Between Global Solution and Expectation (Uniform Distribution)

Molecule Set		8	9	10	12	14
Light gases	1–7	0.039 ± 0.062	0.031 ± 0.061	0.020 ± 0.081	0.006 ± 0.037	0.002 ± 0.004
Chain H/C	8–18	0.016 ± 0.079	0.004 ± 0.022	0.058 ± 0.131	0.030 ± 0.086	0.015 ± 0.061
Aromatic H/C	19–24	0.000 ± 0.000	0.000 ± 0.000	0.005 ± 0.028	0.180 ± 0.215	0.123 ± 0.163

Note: Entries correspond to $\mu \pm 1\sigma$. Results at $T = 300$ K.

Table 6. Equilibrium Concentration Difference Between Global Solution and Expectation (Gibbs–Boltzmann distribution)

Molecule Set		8	9	10	12	14
Light gases	1–7	0.008 ± 0.029	0.005 ± 0.022	0.002 ± 0.011	0.000 ± 0.000	0.000 ± 0.000
Chain H/C	8–18	0.007 ± 0.070	0.002 ± 0.017	0.034 ± 0.144	0.009 ± 0.058	0.000 ± 0.000
Aromatic H/C	19–24	0.000 ± 0.000	0.000 ± 0.000	0.002 ± 0.022	0.099 ± 0.246	0.048 ± 0.213

Note: Entries correspond to $\mu \pm 1\sigma$. Results at $T = 300$ K.

We are interested here in Sr–ETS–4 (Braunbarth et al.2000; Nair et al., 2001a), which is a different type of titanosilicate. This material is especially interesting, since its framework is known to contract after thermal treatment (Kuznicki et al., 2001), and this phenomenon has been rigorously studied by neutron diffraction and Raman spectroscopy (Nair et al., 2001b). Of particular interest are Sr–ETS–4 structures that are formed after heat treatment in the temperature range between 200 and 300 °C, which are known to be selective for methane over nitrogen (Maple and Williams, 2008; Delgado et al., 2008). A related structure, Na–ETS–4, is also known to exhibit this type of selectivity and, furthermore, known to selectively admit oxygen over argon (Pillai et al., 2008). Based on the structure of the Sr–ETS–4 pores formed after heat treatment at $T = 573$ K (Nair et al., 2001b), Figure 18 presents the equilibrium concentration curves, as a function of portal atomic radii, for various light gases.

We observe that, if the effective radii are considered to be in the range between 1.30 and 1.40 Å, the predictions of our method would come in excellent agreement with the experimental observations (Delgado et al., 2008; Maple and Williams, 2008; Pillai et al., 2008), since the hydrogen and oxygen molecules would penetrate freely while argon and methane would experience great difficulty. This probable range of values is different than the corresponding range obtained from the ZSM–5 analysis in ''The case of the ZSM–5 framework'' Section. However, this is not surprising as it is reasonable that these values are in general structure–dependent. Moreover, access to the pore openings of ETS–4 may be affected by the presence of extra-framework cations.

The separation of furfural, HMF, and vanillin

Zeolite pretreatment of hydrolyzates can improve the yield of lignocellulosic bio-ethanol production by removing fermentation inhibitors such as furfural, hydroxymethyl–furfural (HMF), and vanillin (Palmqvist and Hahn–Hägerdal, 2000) from the fermentable sugars stream. Furthermore, zeolites can be used to isolate these commercially valuable compounds and, thus, enhance the financial viability of the overall process.

Applying the method that takes into account the multiplicity of local solutions (uniform distribution of initial states) on this molecular system, we present Table 7 that lists the most promising structures to (a) exclude only vanillin, and (b) admit only furfural. These zeolites would be good candidates to serve as adsorbents in a two-stage separation process where the mixture of the three inhibitor compounds is sharply divided into its three constituents.

After examining the entries of Table 7, Ranjan et al. (2009) conducted relevant experimental studies using Ferrierite (FER) and ZSM–5 (MFI). These two structures were selected because they are commonly available, as well as because they can be synthesized in pure silica form. This latter feature makes them hydrophobic, something desirable in the context of sugar fermentation that takes place in an aqueous solution (the hydrophobicity typically assists the separation by keeping the water molecules outside of the pores).

The equilibrium adsorption isotherms that resulted from the single component adsorption experiments of Ranjan et al. (2009) are in very good agreement with our predictions. The authors reported that both FER and MFI adsorb more furfural than HMF, while FER exhibits higher furfural/

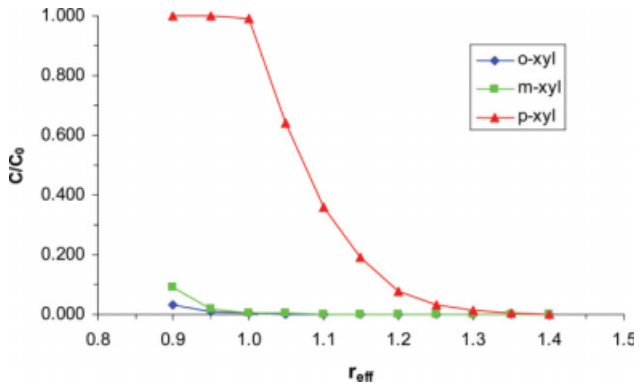


Figure 17. Equilibrium concentration of the xylene isomers in MFI (T = 600 K) as a function of effective atomic radii (in Å).

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

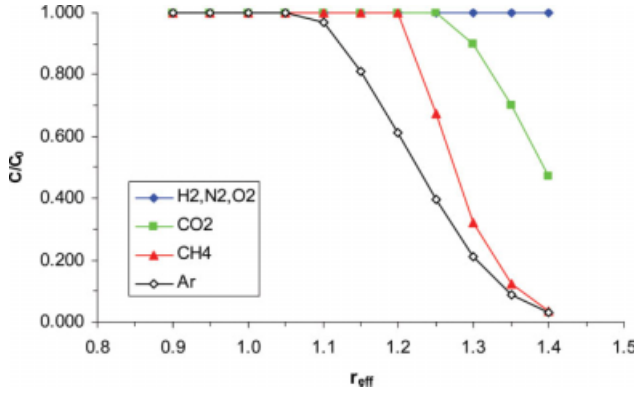


Figure 18. Equilibrium concentration of light gases in Sr–ETS–4 (T = 573 K) as a function of effective atomic radii (in Å).

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

HMF selectivity than MFI. Vanillin is clearly excluded by both structures.

Effect on predictions: furfural/HMF selectivity

In "The case of the ZSM–5 framework" and "The case of a titanosilicate framework" Section, we focused on particular structures and interpreted our results through facts known from experimental observations. This allowed us to deduce potential values for the effective O–atom radii in the portals that we studied. In the previous section, we assumed a universal value for the O–atom radii of zeolites, $r_{\text{eff}} = 1.10$ Å, and studied the molecular system of furfural, HMF and vanillin, offering some predictions that were later validated by experimental research. In this section, we are interested in finding out how robust these predictions would have been with respect to the actual value used for r_{eff}.

We focus on the separation of furfural from HMF, which notably is a rather difficult separation since the two molecules enjoy a great degree of resemblance. The selectivity computation was performed using the method that takes into account the multiplicity of local solutions (probability measure based on uniform distribution). Figure 19 plots the $T = 300$ K selectivity results for seven structures of interest (FER, MEL, MFI, MTT, RRO, STT, TON), as a function of effective radii. These structures were selected because they exhibit high selectivity ($\gamma \geq 0.60$ for at least some r_{eff} value in the depicted range), as well as because they can be available in a pure silica form and, thus, be highly hydrophobic.

Table 7. Candidate Zeolite Structures for a Two-Stage Separation of Fermentation Inhibitors ($T = 300$ K)

	C/C_0		
Zeol	Furf	HMF	Van
(a) Structures Excluding Only Vanillin			
AEL	0.860	0.843	0.000
AFR	0.988	1.000	0.047
AFS	0.984	1.000	0.012
BPH	0.875	1.000	0.000
CAN	0.814	0.888	0.000
GME	0.991	1.000	0.019
GON	0.801	0.801	0.000
ISV	0.921	1.000	0.000
IWW	0.855	1.000	0.000
MEI	0.995	1.000	0.013
MTW	0.811	0.913	0.000
OFF	0.911	1.000	0.000
VET	0.803	0.969	0.000
(b) Structures admitting only furfural			
AHT	0.624	0.058	0.000
DAC	0.674	0.000	0.000
EUO	0.610	0.025	0.000
FER	**0.724**	**0.001**	**0.000**
HEU	0.732	0.000	0.000
LAU	0.671	0.000	0.000
MEL	0.623	0.001	0.000
MFI	**0.717**	**0.085**	**0.000**
MTT	0.664	0.006	0.000
MWW	0.599	0.000	0.000
RON	0.687	0.000	0.000
STI	0.549	0.023	0.000
STT	0.700	0.000	0.000

Bold typeface indicates experimental validation by Ranjan et al. (2009).
The effective oxygen radius is taken as $r_{\text{eff}} = 1.10$ Å (other values may have to be used for certain structures).

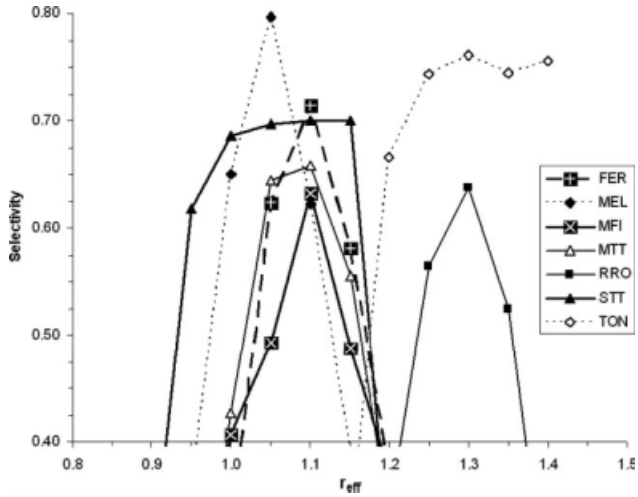


Figure 19. Selectivity between hydroxymethyl–furfural and furfural ($T = 300$ K) as a function of effective atomic radii (in Å) for seven hydrophobic structures.

We observe that the results for SSZ–23 (STT) and Theta–1 (TON) are more robust, as these structures maintain high selectivity for a considerable portion of the r_{eff}–range considered. There is greater confidence that they would indeed be selective, irrespectively of how large the radii of the O–atoms actually are. Moreover, depending on the actual r_{eff} value, other structures may emerge as good candidates for exhibiting selectivity. For example, RUB–41 (RRO) would be selective only if higher values of r_{eff} are applicable, while an opposite trend holds for Ferrierite (FER), ZSM–11 (MEL), ZSM–5 (MFI) and ZSM–23 (MTT) which would be selective if $r_{\text{eff}} \in [1.00,1.15]$.

Conclusions

The penetration by a guest molecule of the confined space of a host portal is a complex process that depends on the joint influence of many competing forces. In an effort to develop a simple and computationally efficient framework to study this process, we describe the interactions in terms of three flexibilities. The first corresponds to reducing the radii of the molecule and portal atoms, and this was studied in two earlier publications Gounaris et al. (2006a,b). This article provides an enhancement to that original approach by incorporating portal flexibility through variable portal atom locations. We observed that the activation energies required for penetration may be significantly smaller than those that were computed before. The differences are more prominent when activation energies are large, however the actual effect on equilibrium concentrations remains moderate. Given the uncertainty regarding the extent of flexibility of zeolite structures, we also conducted a suitable sensitivity analysis that effectively bounds the real case between the cases of rigid and very flexible portals. A further type of flexibility, that of allowing the guest molecule's bond lenghts and bond angles to vary, presents another level of complexity and is left to be addressed in the future.

In addition to studying the effect of flexible zeolite portals, in this article we also developed an improved methodology in order to account for the multiplicity of possible transition conformers. We showed that even relatively simple systems exhibit such multiplicity and, through extensive use of local optimization techniques, were able to quantify the overestimation of results that are obtained if this is not taken into account.

Finally, we subjected the oxygen atom radii values into sensitivity analysis. It was shown that the results are indeed sensitive to these values and that their precise knowledge would lead to more accurate predictions. The analysis serves the purpose of providing reliable bounds on sorption performance, as well as dictate that some of the predictions can be drawn with greater confidence. Furthermore, for given structures for which extensive experimental observation is available, such an analysis can help us fine tune the relevant parameters into more appropriate values. Note that the effective atomic radii can be envisioned also as a surrogate for many other types of possible uncertainties such as the presence and exact location of charge–balancing cations in the portal structure.

The enhancements presented in this paper are crucial in improving the accuracy of our method's predictions, particularly when constrained passage cases are involved. This is due to the fact that, in these cases, the sensitivity of the optimization solutions is higher and even minor adjustments can lead to considerable changes in equilibrium concentration. On the other hand, the extreme cases of free passage and no passage are typically not affected significantly.

Acknowledgments

The authors gratefully acknowledge support from the National Science Foundation.

Literature Cited

Baerlocher C, McCusker LB. *Database of Zeolite Structures*. Available at: http://www.iza-structure.org/databases/.

Baerlocher C, Meier WM, Olson DH. *Atlas of Zeolite Framework Types: Fifth Revised Edition*. Amsterdam, The Netherlands: Elsevier, 2001.

Bandyopadhyay S, Yashonath S. Diffusion anomaly in silicalite and VPI–5 from molecular dynamics simulations. *J Phys Chem.* 1995;99: 4286–4292.

Borello E, Lamberti C, Bordiga S, Zecchina A. Quantum–size effects in the titanosilicate molecular sieve. *App Phys Lett*. 1997;71:2319–2321.

Boulicaut L, Brandani S, Ruthven DM. Liquid phase sorption and diffusion of branched and cyclic hydrocarbons in silicalite. *Microporous Mesoporous Mater*. 1998;95:81–93.

Braunbarth C, Hillhouse HW, Tsapatsis M, Burton A, Lobo RF, Jacubinas RM, Kuznicki SM. Structure of strontium ion-exchanged ETS–4 microporous molecular sieves. *Chem Mater*. 2000;12:1857–1865.

Brooke A, Kendrick D, Meeraus A, Raman R. *GAMS: A Users Guide*. Washington, DC: GAMS Development Corporation, 2005.

Cavalcante CL, Ruthven DM. Adsorption of branched and cyclic paraffins in silicalite. 1. Equilibrium. *Indus Eng Chem Res.* 1995;34:177–184.

Chen NY, Garwood WE, Dwyer FG. *Shape Selective Catalysis in Industrial Applications*. New York: Marcel Dekker, 1996.

Chon H, Park DH. Diffusion of cyclohexanes in ZSM–5 zeolites. *J Catal*. 1988;114:1–7.

Conner WC, Vincent R, Man P, Fraissard J. Flexibility in zeolites: ^{29}Si NMR studies of ZSM–5 frame transitions. *Catal Lett*. 1990;4: 75–84.

Cook M, Conner WC. *How big are the pores of zeolites?* In Treacy M, editor. *XIIth Congress on Zeolites*. 1999;409–414. Warrendale, PA: M.R.S.

Das TK, Chandwadkar AJ, Sivasanker SJ. Studies on the synthesis, characterization and catalytic properties of the large pore titanosilicate, ETS–10. *J Mol Catal A*. 1996;107:199–205.

Deem MW, Newsam JM, Creighton JA. Fluctuations in zeolite aperture dimensions simulated by crystal dynamics. *J Am Chem Soc*. 1992; 114:7198–7207.

Delgado JA, Uguina MA, Agueda VI, Garcia-Sanz A. Adsorption and diffusion parameters of methane and nitrogen on microwave-synthesized ETS–4. *Langmuir*. 2008;24:6107–6115.

Demontis P, Suffritti GB, Fois ES, Quartieri S. Molecular dynamics studies on zeolites. 6. Temperature dependence of diffusion of methane in silicalite. *J Phys Chem*. 1992;96:1482–1490.

Demontis P, Suffritti GB, Quartieri S, Fois ES, Gamba A. Molecular dynamics studies on zeolites. 3. Dehydrated zeolite A. *J Phys Chem*. 1998;92:867–871.

Drud AS. *CONOPT: A system for large scale nonlinear optimization*, In *Reference Manual for CONOPT Subroutine Library*. Bagsvaerd, Denmark, ARKI Consulting and Development A/S, 1996.

Gu X, Dong J, Nenoff TM, Ozokwelu DE. Separation of p-xylene from multicomponent vapor mixtures using tubular MFI zeolite membranes. *J Membr Sci*. 2006;280:624–633.

Gounaris CE, Floudas CA, Wei J. Rational design of shape selective separation and catalysis. I. Concepts and analysis. *Chem Eng Sci*. 2006a;61:7933–7948.

Gounaris CE, Wei J, Floudas CA. Rational design of shape selective separation and catalysis. II. Mathematical model and computational studies. *Chem Eng Sci*. 2006b;61:7949–7962.

Hypercube, Inc. *HyperChem Getting Started Release 5.0 for Windows*. Gainesville, FL, 1996.

Jarzynski C. Nonequilibrium equality for free energy differences. *Phys Rev Lett*. 1997;79:2690–2693.

Karsli H, Culfaz A, Yuecel H. Sorption properties of silicalite-1 of pure silica form: the influence of sorption kinetics of critically sized molecules. *Zeolites*. 1992;12:728–732.

Kiselev AV, Du PQ. Molecular statistical calculation of the thermodynamic adsorption characteristics of zeolites using the atom-atom approximation. 2. Adsorption of non-polar and polar inorganic molecules by zeolites of type-X and type-Y. *J Chem Soc–Faraday Trans II*. 1981;77:1–15.

Kiselev AV, Lopatkin AA, Shulga AA. Molecular statistical calculation of gas-adsorption by silicalite. *Zeolites*. 1985;5:261–267.

Kopelevich DI, Chang HS. Diffusion of inert gases in silica sodalite: importance of lattice flexibility. *J Chem Phys*. 2001;115:9519–9527.

Kresge CT, Dhingra SS. *Molecular sieves*. In *Kirk-Othmer Encyclopedia of Chemical Technology*, Wiley, 2004

Kuznicki SM, Bell VA, Nair S, Hillhouse HW, Jacubinas RM, Braunbarth CM, Toby BH, Tsapatsis M. A titanosilicate molecular sieve with adjustable pores for size-selective adsorption of molecules. *Nature*. 2001;412:720–724.

Magalhaes FD, Laurence RL, Conner WC. Diffusion of cyclohexane and alkylohexanes in silicalite. *J Phys Chem B*. 1998;102:2317–2324.

Maple MJ, Williams CD. Separating nitrogen/methane on zeolite-like molecular sieves. *Microporous Mesoporous Mater*. 2008;111:627–631.

Murtagh BA, Saunders MA. MINOS 5.1 User's Guide. Report SOL 83–20R, 1987.

Nair S, Jeong HK, Chandrasekaran A, Braunbarth CM, Tsapatsis M, Kuznicki SM. Synthesis and structure determination of ETS–4 single crystals. *Chem Mater*. 2001a;13:4247–4254.

Nair S, Tsapatsis M, Toby BH, Kuznicki SM. A study of heat-treatment induced framework contraction in strontium–ETS–4 by powder neutron diffraction and vibrational spectroscopy. *J Am Chem Soc*. 2001b; 123:12781–12790.

Norman JG.. Non–empirical versus empirical choices for overlapping–sphere radii ratios in SCF–XALPHA–SW calculations on CLO–4(-) and SO–2. *Mole Phys*. 1976;31:1191–1198.

Palmqvist E, Hahn–Hägerdal B. Fermentation of lignocellulosic hydrolysates II: inhibitors and mechanisms of inhibition. *Bioresource Technol*. 2000;74:25–33.

Philippou A, Naderi M, Pervaiz N, Rocha J, Anderson MW. N-hexane reforming reactions over basic Pt–ETS–10 and Pt–ETAS–10. *J Catal*. 1998a;178:174–181.

Philippou A, Naderi M, Rocha J, Anderson MW. Dehydration of t-butanol over basic ETS–10, ETAS–10 and AM–6 catalysts. *Catal Lett*. 1998b;53:221–224.

Pillai RS, Peter SA, Jasra RV. Adsorption of carbon dioxide, methane, nitrogen, oxygen and argon in NaETS-4. *Microporous Mesoporous Mater*. 2008;113:268–276.

Ranjan R, Thust S, Gounaris CE, Woo M, Floudas CA, Von Keitz M, Valentas KJ, Wei J, Tsapatsis M. Adsorption of fermentation inhibitors from lignocellulosic biomass hydrolyzates for improved ethanol yield and value–added product recovery. *Microporous Mesoporous Mater*. 2009; 122:143–148.

Santikary P, Yashonath S. Dynamics of zeolite cage and its effect on the diffusion properties of sorbate: Persistence of diffusion anomaly in NaA zeolite. *J Phys Chem*. 1994;98:9252–9259.

Sartbaeva A, Wells SA, Treacy MMJ, Thorpe MF. The flexibility window in zeolites. *Nature Mater*. 2006;5:962–965.

Schrimpf G, Schlenkrich M, Brickmann J, Bopp P. Molecular dynamics simulation of zeolite NaY. A study of structure, dynamics, and thermalization of sorbates. *J Phys Chem*. 1992;96:7404–7410.

Smit B, Maesen TLM. Towards a molecular understanding of shape selectivity. *Nature*. 2008;451:671–678.

Tezel OH, Ruthven DM. Sorption of benzene in NaX zeolite: An unusual hysteresis effect. *J Colloid Interfar Sci*. 1990;139:581–583.

Van Santen RA, De Man AJM, Jacobs WPJH, Teunissen EH, Kramer GJ. Lattice relaxation of zeolites. *Catal Lett*. 1991;9:273–286.

Yashonath S, Santikary P. Sorbate properties and cage-to-cage diffusion of argon in NaCaA: A molecular dynamics study. *J Phys Chem*. 1993;97:13778–13787.

Yashonath S, Santikary P. Diffusion in zeolites: Amomalous dependence on sorbate diameter. *J Chem Phy*. 1994;100:4013–4016.

Appendix Table A1. Complete List of Supported Molecules

#	Category	Molecules
1–6	Noble gases	*He, Ne, Ar, Kr, Xe, Rn*
7–12	Light gases	**H_2, N_2, O_2, CO, CO_2, NH_3**
13–16	H_2x	H_2O, *H_2Se*, *H_2S*, *H_2Te*
17–24	Inorganic acids	*H_2CO_3*, *HNO_3*, *H_3PO_4*, *H_2SO_4*, *HF, HCl, HBr, HI*
25–49	Chain hydrocarbons	**CH_4, C_2H_6, C_2H_4, C_2H_2, C_3H_8, C_3H_6**, Propadiene, nC_4H_{10}, iC_4H_{10}, iso–Butylene, alpha–Butylene, cis–beta–Butylene, trans–beta–Butylene, Butadiene, Methylacetylene, **nC_5H_{12}, iC_5H_{12}, $neoC_5H_{12}$**, Isoprene, Spiropentane, Diethyl–Pentane, *$C_{10}H_{22}$*,*$C_{15}H_{32}$*,*$C_{20}H_{42}$*, TetraDecaneMethane (MOBIL1)
50–69	Aromatic hydrocarbons	**Benzene**, **Toluene**, **o–Xylene**, **m–Xylene**, **p–Xylene**, Mesitylene, **Ethyl–Benzene**, Benzocyclobutene, Biphenyl, Napthalene, Anthracene, Phenanthrene, Napthacene, Triphenylene, Pyrene, Fluoranthene, Tetraphene, Chrysene, Benzopyrene, Coronene
70–79	Alcohols	*CH_3OH*,*C_2H_5OH*,*C_3H_7OH*, Isopropyl–Alcohol, Isobutyl–Alcohol, Tertamyl–Alcohol, Phenol, Ethylene–Glycol, Propylene–Glycol, Glycerol
80–87	Aldehydes/Ketones	*HCHO*,*CH_3CHO*, Acetone, *$CH_3COCOOH$*,*C_6H_5COH*, Benzophenone, Methyl–Urea, Urea
88–89	Ethers	MethylTertButyl–Ether, TertAmylMethyl–Ether
90–105	Carboxylic acids and anhydrides	*HCOOH*,*CH_3COOH*,*C_2H_5COOH*,*$CH_2CHCOOH$*,*$CH_3CH_2CH_2COOH$*,*C_6H_5COOH*, Oxalic–Acid, Malonic–Acid, Succinic–Acid, Maleic–Acid, Glutaric–Acid, Adipic–Acid, Salicylic–Acid, Citric–Acid, Maleic–Anhydride, Phthalic–Anhydride
106–110	Esters	Propyl–Methanoate, Ethyl–Ethanoate, Propyl–Ethanoate, Ethyl–Propanoate, Methyl–Butanoate
111–117	Biomolecules	**Cholesterol**, **Progesterone**, **Testosterone**, **L–Lactic–Acid**, **D–Lactic–Acid**, Estrone, Insulin
118–126	Fatty acids	Lauric–Acid, **Stearic–Acid**, **Oleic–Acid**, **Elaidic–Acid**, **Linoleic–Acid**, alpha–Linolenic–Acid, gamma–Linolenic–Acid, Arachidonic–Acid, Erucic–Acid
127–130	Vitamins	Retinol (Vitamin–A), Riboflavin (Vitamin–B2), Niacin (Vitamin–B3), Ascorbic–Acid (Vitamin–C)
131–136	Sugars and sweeteners	**Fructose**, **Glucose**, Sucrose, Aspartame, Cyclamate, Saccharin
137–156	Amines	*CH_3NH_2*,*$C_2H_5NH_2$*, Dimethyl–Amine, Trimethyl–Amine, Monoethanol–Amine, Methylethanol–Amine, Dimethylethanol–Amine, Diethanol–Amine, Methyldiethanol–Amine, Triethanol–Amine, Tris, Bis–Tris, Tricine, Aniline, Diphenyl–Amine, Triphenyl–Aniline, Amphetamine, Ethylene–Diamine, Diethylene–Triamine, Triethylene–Tetramine
157–168	Cyclic	Cyclohexane, Methyl–Cyclohexane, Nitrobenzene, Chlorobenzene, Benzyl–Chloride, Menthol, Nitroaniline, Nicotine, Caffeine, Uric–Acid, Alizarin, Indigo
169–190	Heterocyclic	Aziridine, Azetidine, Pyrrolidine, Piperidine, Pyrrole, Indole, Pyridine, Quinoline, Isoquinoline, Acridine, Pyrimidine, Purine, Furan, Benzofuran, Furfural, Hydroxymethyl–Furfural, Vanillin, Thiophene, Benzothiophene, Dibenzothiophene, MDBT, DMDBT
191–211	Drugs	Acetyl–Salicylic–Acid (ASPIRIN), Amlodipine (NORVASC), Ampicillin, Atorvastatin (LIPITOR), Barbituric–Acid, Clopidogrel (PLAVIX), Cocaine, Codeine, Ephedrine, Heroin, Methadone, Methamphetamine, Morphine, Naloxone, Olanzapine (ZYPREXA), Paclitaxel (TAXOL), Paracetamol (TYLENOL), Penicillin–G, Phenacetin, Phenylephrine, Simvastatin (ZOCOR)
212–218	Explosives and propellants	*H_2O_2*, Hydrazine, Monomethyl–Hydrazine, Unsymmetrical–Dimethyl–Hydrazine, Dinitrogen–Tetroxide, Picric–Acid, Trinitrotoluene
219–251	Haloalkanes	*CCl_4*, *CCl_3F*, *CCl_2F_2*, *$CClF_3$*, *CF_4*, *CF_3I*, *$CBrF_3$*, *$CBrClF_2$*, *CHF_3*, *$CHCl_3$*, *$CHBr_3$*, *$CHClF_2$*, *CH_2F_2*, *CH_2Cl_2*, *CH_2Br_2*, *CH_3F*, *CH_3Cl*, *CH_3Br*, *C_2Cl_6*, *Cl_3CCH_3*, *$Cl_2FCCClF_2$*, *Cl_2FCCH_3*, *Cl_2HCCH_3*, *ClH_2CCH_2Cl*, *ClF_2CCClF_2*, *Cl_3CCF_3*, *ClF_2CCF_3*, *ClF_2CCH_3*, *$ClFHCCF_3$*, *F_2HCCF_3*, *FH_2CCF_3*, *F_2HCCHF_2*, *F_2HCCH_3*
252–271	Amino acids	Alanine, Arginine, Asparagine, Aspartic–Acid, Cysteine, Glutamic–Acid, Glutamine, Glycine, Histidine, Isoleucine, Leucine, Lysine, Methionine, Phenylalanine, Proline, Serine, Threonine, Tryptophan, Tyrosine, Valine
272–276	Nucleobases	Adenine, Cytosine, Guanine, Thymine, Uracil
277–280	Fluorocompounds	*SF_4*, **SF_6, UF_6**, *S_2F_{10}*
281–285	Geometrically shaped	Adamantane, **Cubane**, Octanitrocubane, *C_{20}*,*C_{60}*
286–290	Various	*HCN*, Camphor, Dimethyl–Sulfoxide, Perfluorooctane–Sulfonate, Perfluorooctanoic–Acid

Original database entries typed in boldface.

Appendix Table B1. List of Zeolite Windows Used

ABW	[8]	3.8 × 3.4	DAC1	[10]	5.3 × 3.4	JBW	[8]	4.8 × 3.7	RTH2	[8]	5.6 × 2.5
ACO1	[8]	3.5 × 2.8	DAC2	[8]	4.8 × 3.7	KFI	[8]	3.9 × 3.9	RWR	[8]	5.0 × 2.8
ACO2	[8]	3.5 × 3.5	DDR	[8]	4.4 × 3.6	LAU	[10]	5.3 × 4.0	SAO1	[12]	7.2 × 6.5
AEI1	[8]	3.8 × 3.8	DFO1	[12]	7.3 × 7.3	LEV	[8]	4.8 × 3.6	SAO2	[12]	7.0 × 7.0
AEI2	[8]	3.8 × 3.8	DFO2	[8]	5.6 × 3.4	LOV1	[9]	4.5 × 3.2	SAS	[8]	4.2 × 4.2
AEI3	[8]	3.8 × 3.8	DFO3	[12]	6.2 × 6.2	LOV2	[9]	4.2 × 3.0	SAT	[8]	5.5 × 3.0
AEL	[10]	6.5 × 4.0	DFO4	[10]	6.4 × 5.4	LOV3	[8]	3.7 × 3.6	SAV1	[8]	3.8 × 3.8
AEN1	[8]	4.3 × 3.1	DFT1	[8]	4.1 × 4.1	LTA	[8]	4.1 × 4.1	SAV2	[8]	3.9 × 3.9
AEN2	[8]	5.0 × 2.7	DFT2	[8]	4.7 × 1.8	LTL	[12]	7.1 × 7.1	SBE1	[12]	7.4 × 7.2
AET	[14]	8.7 × 7.9	DFT3	[8]	4.7 × 1.8	MAZ1	[12]	7.4 × 7.4	SBE2	[8]	4.0 × 4.0
AFI	[12]	7.3 × 7.3	DON	[14]	8.2 × 8.1	MAZ2	[8]	3.1 × 3.1	SBE3	[8]	3.1 × 3.0
AFN1	[8]	4.6 × 1.9	EAB	[8]	5.1 × 3.7	MEI1	[12]	6.9 × 6.9	SBS1	[12]	6.8 × 6.8
AFN2	[8]	4.9 × 2.1	EDI1	[8]	3.8 × 2.8	MEI2	[7]	3.5 × 3.2	SBS2	[12]	7.0 × 6.9
AFN3	[8]	4.0 × 3.3	EDI2	[8]	3.1 × 2.0	MEL	[10]	5.4 × 5.3	SBS3	[8]	5.2 × 3.5
AFO	[10]	7.0 × 4.3	EMT1	[18]	7.3 × 7.3	MER1	[8]	3.5 × 3.1	SBT1	[12]	7.4 × 6.4
AFR1	[12]	6.9 × 6.7	EMT2	[14]	7.5 × 6.5	MER2	[8]	3.6 × 2.7	SBT2	[12]	7.8 × 7.3
AFR2	[8]	3.7 × 3.7	EPI1	[10]	5.6 × 3.4	MER3	[8]	5.1 × 3.4	SBT3	[8]	4.4 × 3.5
AFS1	[12]	7.0 × 7.0	EPI2	[10]	6.2 × 2.6	MER4	[8]	3.3 × 3.3	SFE	[12]	7.6 × 5.4
AFS2	[8]	4.0 × 4.0	EPI3	[8]	4.5 × 3.7	MFI1	[10]	5.5 × 5.1	SFF	[10]	5.7 × 5.4
AFT1	[8]	3.8 × 3.2	ERI	[8]	5.1 × 3.6	MFI2	[10]	5.6 × 5.3	SFH	[14]	8.7 × 6.4
AFT2	[8]	3.8 × 3.6	ESV	[8]	4.7 × 3.5	MFS1	[10]	5.4 × 5.1	SFN	[14]	8.5 × 6.2
AFX	[8]	3.6 × 3.4	ETR1	[18]	10.1 × 10.1	MFS2	[8]	4.8 × 3.3	SSY	[12]	7.6 × 5.0
AFY1	[12]	6.1 × 6.1	EUO	[10]	5.4 × 4.1	MON1	[8]	4.4 × 3.2	STF	[10]	5.7 × 5.4
AFY2	[8]	4.3 × 4.0	FAU	[12]	7.4 × 7.4	MON2	[8]	3.6 × 3.6	STI1	[10]	5.0 × 4.7
AHT	[10]	6.8 × 3.3	FER1	[10]	5.4 × 4.2	MOR1	[12]	7.0 × 6.5	STI2	[8]	5.6 × 2.7
AWW	[8]	3.9 × 3.9	FER2	[8]	4.8 × 3.5	MOR2	[8]	5.7 × 2.6	STT1	[9]	5.3 × 3.7
BCT	[8]	2.4 × 2.4	GIS1	[8]	4.5 × 3.1	MTF	[8]	3.9 × 3.6	STT2	[7]	3.5 × 2.4
BEA1	[12]	6.7 × 6.6	GIS2	[8]	4.5 × 3.1	MTT	[10]	5.2 × 4.5	TER1	[10]	5.0 × 5.0
BEA2	[12]	5.6 × 5.6	GME1	[12]	7.0 × 7.0	MTW	[12]	6.0 × 5.6	TER2	[10]	7.0 × 4.1
BEC1	[12]	7.5 × 6.3	GME2	[8]	3.9 × 3.6	MWW1	[10]	5.5 × 4.0	THO1	[8]	3.9 × 2.3
BEC2	[12]	6.9 × 6.0	GON1	[12]	6.8 × 5.4	MWW2	[10]	5.1 × 4.1	THO2	[8]	4.0 × 2.2
BIK	[8]	3.7 × 2.8	GON2	[8]	4.3 × 1.3	NAT1	[8]	3.9 × 2.6	THO3	[8]	3.0 × 2.2
BOG1	[12]	7.0 × 7.0	GOO1	[8]	4.0 × 2.8	NAT2	[9]	4.1 × 2.5	TON	[10]	5.7 × 4.6
BOG2	[10]	5.8 × 5.5	GOO2	[8]	4.1 × 2.7	NES	[10]	5.7 × 4.8	TSC1	[8]	4.2 × 4.2
BPH1	[12]	6.3 × 6.3	GOO3	[8]	4.7 × 2.9	OFF1	[18]	6.8 × 6.7	TSC2	[8]	5.6 × 3.1
BPH2	[8]	3.5 × 2.7	HEU1	[10]	7.5 × 3.1	OFF2	[10]	4.9 × 3.6	UOZ	[12]	4.7 × 4.6
BRE1	[8]	5.0 × 2.3	HEU2	[8]	4.6 × 3.6	OSI	[12]	6.0 × 5.2	USI1	[12]	6.2 × 6.1
BRE2	[8]	4.1 × 2.8	HEU3	[8]	4.7 × 2.8	OSO1	[14]	7.3 × 5.4	USI2	[10]	6.4 × 3.9
CAN	[12]	5.9 × 5.9	IFR	[12]	7.2 × 6.2	OSO2	[8]	3.3 × 2.8	VET	[12]	5.9 × 5.9
CAS	[8]	4.7 × 2.4	ISV1	[12]	6.5 × 6.1	PAR	[10]	6.9 × 3.5	VFI	[18]	12.7 × 12.7
CFI	[14]	7.5 × 7.2	ISV2	[12]	6.6 × 5.9	PAU	[8]	3.6 × 3.6	VNI1	[8]	4.0 × 3.1
CGF1	[10]	9.2 × 2.5	ITE1	[8]	4.3 × 3.8	PHI1	[8]	3.8 × 3.8	VNI2	[8]	3.6 × 3.5
CGF2	[8]	6.7 × 2.1	ITE2	[8]	5.8 × 2.7	PHI2	[8]	4.3 × 3.0	VNI3	[8]	4.1 × 3.6
CGF3	[8]	4.8 × 2.4	ITH1	[10]	5.3 × 4.8	PHI3	[8]	3.3 × 3.2	VSV1	[9]	4.3 × 3.3
CGS1	[10]	8.1 × 3.5	ITH2	[10]	5.1 × 4.8	RHO	[8]	3.6 × 3.6	VSV2	[9]	4.2 × 2.9
CGS2	[8]	4.6 × 2.5	ITH3	[9]	4.8 × 4.0	RON	[12]	4.3 × 4.3	VSV3	[8]	2.7 × 2.1
CHA	[8]	3.8 × 3.8	ITW1	[8]	5.4 × 2.4	RRO1	[10]	6.5 × 4.0	VSV4	[8]	4.0 × 3.8
CHI	[9]	4.3 × 3.9	ITW2	[8]	4.2 × 3.9	RRO2	[8]	5.0 × 2.7	WEI1	[10]	5.4 × 3.1
CLO1	[20]	13.2 × 6.0	IWR1	[12]	6.8 × 5.8	RRO3	[8]	4.6 × 2.8	WEI2	[8]	5.0 × 3.3
CLO2	[8]	3.8 × 3.8	IWR2	[10]	5.3 × 4.6	RSN1	[9]	4.4 × 3.3	WEN1	[10]	4.8 × 2.5
CON1	[12]	7.0 × 6.4	IWR3	[12]	7.1 × 5.7	RSN2	[9]	4.3 × 3.1	WEN2	[8]	2.7 × 2.3
CON2	[12]	7.0 × 5.9	IWW1	[12]	6.4 × 6.0	RSN3	[8]	4.1 × 3.4	YUG1	[8]	3.6 × 2.8
CON3	[10]	5.1 × 4.5	IWW2	[10]	4.9 × 4.9	RTE	[8]	4.4 × 3.7	YUG2	[8]	5.0 × 3.1
CZP	[12]	7.2 × 3.8	IWW3	[8]	4.6 × 3.3	RTH1	[8]	4.1 × 3.8	ZON1	[8]	5.1 × 2.5
									ZON2	[8]	4.4 × 3.7

Number of O-atoms shown in brackets. Portal's nominal diameters are in Å.

Manuscript received Mar. 7, 2009, and revision received Jun. 21, 2009.