



THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à Chimie ParisTech

**Origines microscopiques de la séparation xénon/krypton
dans les matériaux nanoporeux**

Microscopic origins of the xenon/krypton separation in
nanoporous materials

Présentée par

Emmanuel REN

Soutenance prévue le
XX Septembre 2023

École doctorale n°388
**Chimie Physique et
Chimie Analytique de
Paris Centre**

Spécialité

Chimie Physique

Composition du jury :

NAME SURNAME TITLE, PLACE/University	<i>Role</i>
François-Xavier COUDERT Directeur de Recherche, Chimie ParisTech	<i>Directeur de thèse</i>

Chimie Paris



| **PSL**

REMERCIEMENTS

En premier lieu, je voudrais adresser ici mes plus vifs remerciements

TABLE OF CONTENTS

General introduction	1
1 High-throughput computational screening of nanoporous materials	5
1.1 Nanoporous materials	5
1.1.1 Introduction and concepts.	5
1.1.2 Computational databases	8
1.1.3 Exploring the chemical and structural space	10
1.2 An overview of screening methodologies	11
1.2.1 Non-adsorption properties	11
1.2.2 Transport adsorption properties	16
1.2.3 Thermodynamic adsorption properties	20
1.3 Xenon/krypton separation.	26
1.3.1 Industrial applications	26
1.3.2 Promising materials for the xenon/krypton separation	27
1.3.3 From the computer to the test tube: a rare use case.	28
1.3.4 The future of screening.	29
2 Thermodynamic exploration of xenon/krypton separation	33
2.1 Calculation of thermodynamic descriptors	33
2.1.1 Thermodynamics of adsorption separation	33
2.1.2 Grand canonical Monte Carlo	34
2.1.3 Widom insertion	34
2.2 Preliminary analyses	34
2.2.1 Structure-selectivity relationships	35
2.2.2 Thermodynamic quantities correlations	35
2.3 Selectivity drop	39
2.3.1 Thermodynamic origins	39
2.3.2 Detailed investigation	43
2.3.3 Conclusions and perspectives	46
3 Adsorption molecular simulations	49
3.1 Voronoi sampling	49
3.2 Rapid Adsorption Enthalpy Surface Sampling (RAESS)	49
3.2.1 Software description	49
3.2.2 Comparison to standard methodologies	49
3.3 Grid Adsorption Energies Descriptors (GrAED)	49

4 Adsorption properties prediction	51
4.1 Machine learning	51
4.1.1 Introduction	51
4.1.2 eXtreme Gradient Boosting	51
4.2 Ambient-pressure prediction	51
4.2.1 From infinite dilution to ambient pressure	51
4.2.2 Interpretation of the ML model.	51
5 Xenon and krypton transport properties	53
5.1 Existing methodologies	53
5.1.1 Molecular dynamics	53
5.1.2 Fast kinetic Monte Carlo	53
5.2 ML modeling	53
5.3 Molecular dynamics calculation results.	54
5.3.1 Why is it relevant?	54
5.3.2 Correlations	54
5.4 Fast diffusion calculation algorithm	54
5.4.1 Implementation in C++	54
5.4.2 Preliminary results	54
5.4.3 Visualization tool	54
5.4.4 ML model training.	54
6 Towards the next generation of screenings	57
6.1 Flexibility	57
6.1.1 Problem, literature.	57
6.1.2 Snapshot.	57
6.2 Noble gas polarisability	57
6.2.1 Problem, literature.	57
6.2.2 Perspectives.	57
General conclusions	59



List of Publications	61
Peer-reviewed papers	61
Preprint	61
Résumé en français	63
Introduction	63

GENERAL INTRODUCTION

Nanoporous materials are material

[Just a copy paste from last article]

Gas separation and purification are essential processes since they provide key reactants and inert gases for the chemical industry, as well as medical or food grade gases. Among them, we can find easily extractable or synthesizable molecules such as nitrogen, oxygen, carbon dioxide, noble gases, hydrogen, methane, or nitrous oxide. Moreover, gas separation is crucial in mitigating negative environmental impact at the end of industrial processes, such as facilities emitting green house gases (e.g. concrete or steel plants) or treating volatile radioactive wastes like ^{85}Kr . Cryogenic liquefaction or distillation is currently the mainstream technique to achieve industrial gas separation, while adsorbent beds made of nanoporous materials (activated alumina or zeolites) are mostly used as a less energy-intensive pre-purification system.[**kerry2007industrial**]

A wider use of nanoporous materials could reduce the energy consumption of current separation processes since adsorption is way less energy intensive than liquefaction.[**national2019research**] For instance, some prototypes involving beds of nanoporous materials have been developed for xenon/krypton separation to avoid employing cryogenic distillation.[**Banerjee2018**] For the process to be viable, materials need to perform even better and many studies focus on synthesizing ever more selective materials by leveraging all chemical intuitions around noble gas adsorption properties.[**Chen_2014, Li_2019, Pei_2022**] In order to speed the discovery process of novel materials with key properties, computational screening can identify factors explaining the performance and pre-select candidates for further experimental studies. As recently conceptualized by Lyu et al., a synergistic workflow combining computational discovery and experimental validation can push material discovery to the next stage.[**Lyu_2020, Jablonka_2022**] But to efficiently guide experimental discoveries, computational chemists are facing two major challenges: generating reliably more structures and evaluating them with fast and accurate models.

The number of nanoporous materials is potentially unlimited; for the metal–organic frameworks (MOFs) alone, over 90,000 structures have been synthesized [**Groom_2016**] and 500,000 computationally constructed [**Wilmer_2012, Boyd_2016, Colon_2017**]. To deal with this ever-increasing amount of structures, we need to design more efficient screening procedures as well as faster performance evaluation tools. To go beyond the time-consuming calculations over the whole dataset, computational chemists developed funnel-like screening procedures to reduce the need for expensive simulations and introduced machine learning (ML) models to

replace them with faster evaluation tools.[Ren_2022] To further improve the selectivity screening for Xe/Kr separation, we will need to design better performing structural and energy-based descriptors.

Simon et al. published one of the first articles on an ML-assisted screening approach for the separation of a Xe/Kr mixture extracted from the atmosphere.[Simon_2015] Their model's performance was highly relying on the Voronoi energy, which is basically an average of the interaction energies of a xenon atom at each Voronoi node.[Rycroft_2009] To rationalize this increase in performance, we regarded this Voronoi energy as a faster proxy for the adsorption enthalpy. By comparing it to the standard Widom insertion, we found that although it is faster, it is less accurate; and we developed a more effective alternative, the surface sampling (RAESS) using symmetry and non-accessible volumes blocking.[Ren_2023] Recently, Shi et al. used an energy grid to generate energy histograms as a descriptor for their ML model, which gives an exhaustive description of the infinitely diluted adsorption energies,[Shi_2023] but can be computationally expensive.

All the approaches described above can have good accuracy in the prediction of low-pressure adsorption (i.e., in the limit of zero loading) but are not suitable for prediction of adsorption in the high-pressure regime, when the material is near saturation uptake. While this later task is routinely performed by Grand Canonical Monte Carlo (GCMC) simulations, there is a lack of methods at lower computational cost for high-throughput screening. To better frame our challenge, in this work we are essentially trying to predict the selectivity in the nanopores of a material at high pressure, where adsorbates are interacting with each other, while only having information on the interaction at infinite dilution. The comparison between the low and high pressure cases gives key information on the origin of the differences of selectivity. For instance, we previously showed that selectivity could drop between the low and ambient pressure cases in the Xe/Kr separation application, and it was mainly attributed to the presence of different pore sizes and potential reorganizations due to adsorbate–adsorbate interactions.[Ren_2021]

[\[Xe/Kr applications in the industry\]](#)

This thesis presents my work on the



HIGH-THROUGHPUT COMPUTATIONAL SCREENING OF NANOPOROUS MATERIALS

1.1	Nanoporous materials	5
1.1.1	Introduction and concepts.	5
1.1.2	Computational databases	8
1.1.3	Exploring the chemical and structural space	10
1.2	An overview of screening methodologies	11
1.2.1	Non-adsorption properties	11
1.2.2	Transport adsorption properties	16
1.2.3	Thermodynamic adsorption properties	20
1.3	Xenon/krypton separation.	26
1.3.1	Industrial applications	26
1.3.2	Promising materials for the xenon/krypton separation	27
1.3.3	From the computer to the test tube: a rare use case.	28
1.3.4	The future of screening	29

1.1 NANOPOROUS MATERIALS

1.1.1 Introduction and concepts

ADSORPTION ISOTHERMS AND GEOMETRICAL DESCRIPTORS

Nanoporous materials are defined as materials with a nanoscale structure constituted of pores and cavities, which some are connected by a network of channels. These pores can be empty or filled with a variety of substances called adsorbates. By adhering molecules from a liquid or gas phase into the internal surfaces of the material, we can use it in various applications such as gas separation and purification,[**Li_2009, Lagorsse_2007**] energy storage and conversion, [**Morris_2008, Qiu_2020**] heterogeneous catalysis[**Bell_2003, Singh_2019, Pascanu_2019**] drug delivery,[**Della_Rocca_2011, Bernini_2014**] or sensing.[**Breslin_1976**] By designing the chemical nature, size, shape and distribution of the pores, we can tailor the physicochemical properties to the targeted application.[**Yan_2020**]

The process of adhering particles or molecules on surfaces is called adsorption. Adsorption occurs due to attractive forces between adsorbates and the adsorbent surface, such as van der Waals forces, hydrogen bonding, and electrostatic interactions. The adsorption performance depends on the chemical nature of the interface, its exposed surface area and the shape of the pores. We usually characterize adsorption properties of an adsorbate compound by measuring the amounts of adsorbed molecules as a function of its pressure at a given temperature, which is called the adsorption isotherm. These isotherms can eventually be used, among other techniques, to specify the pore size distribution, accessible surface area and pore volume. [Rouquerol_1994] Here, we will propose a quick overview on the different methodologies to define and measure these key structural descriptors, which values highly depend on the methodology chosen.

Most of the materials studied in this thesis will have pores with a size around the nanometer called “nanopores”. The International Union of Pure Applied Chemistry classifies these pores into three categories according to their size: micropores (≤ 2 nm), mesopores (2 nm–50 nm) and macropores (> 50 nm). [Sing_1985] Here, we will use a single terminology (nanopore) to designate all pores of under a few nanometers. A good characterization of the nanopores of these materials is key to fine-tuning the adsorption properties. [Yan_2020] The pore size distribution (PSD) can be computationally determined if we have resolved the structure of the nanoporous material (using X-ray diffraction on crystallized porous solids). This is the most accurate determination method of the PSD, but it relies on considering that the structure is perfectly rigid and crystalline so that only one structural data can characterize it. Other experimental methods rely on assumptions, model systems (e.g. cylindrical) or adsorption characteristics. For instance, stereological analyses based on plane sections cut through a porous material can evaluate the PSD. [Haynes_1973] The Horvath-Kawazoe (HK) method is a semi-empirical analytic model of adsorption isotherm that can extract PSD. Small angle X-ray and small angle neutron scattering methods are non-destructive methods of pore characterization. [Radlinski_2004] In this thesis, we will rely on computationally analyzing X-ray diffusion data to deduce pore sizes and other geometrical characteristics.

The pore volume consists in the measure of the volume of “closed” and “open” pores of nanoporous materials. Depending on the way of measuring it, different quantities are probed. Some pores could not be accessed by some adsorbate; depending on the probe size the volume calculated will not be the same. Methods that do not rely on adsorption like scattering or stereology will however measure the total pore volume. The porosity or void fraction would be defined as the ratio between the pore volume and the apparent framework volume. Depending on the method, we can therefore retrieve either the total porosity, the porosity opened or closed to a given probe adsorbate.

The cavities of the nanoporous material lay out an incredibly large adsorbable surface area, which is extremely useful in increasing the number of molecules in a given volume or mass of material, several thousands of square meters can be found in a gram of some nanoporous materials. [Farha_2012] The higher the surface area, the more molecules can be adsorbed for storage, separation or reaction purposes; it is therefore crucial to measure this surface area using experimental and computational methods. The most extensively used method to experimentally measure surface areas from adsorption isotherms is based on the Brunauer–Emmett–Teller (BET) theory. [Detsi_2011] Most BET areas are calculated on N₂ isotherm at its boiling tem-

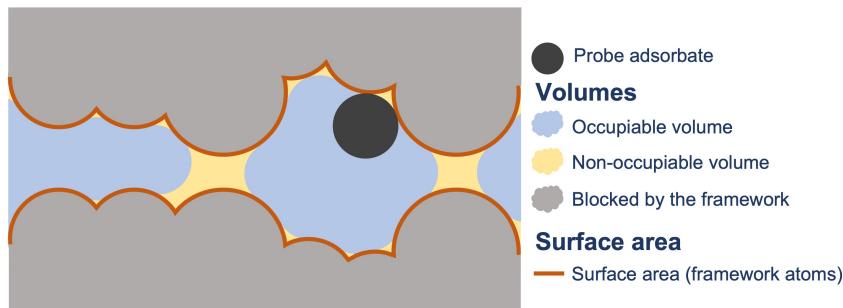


Figure 1.1: Illustration of the pore surface area and volume in a nanoporous material. As illustrated, there are different definitions of the pore volume: we can either consider the whole volume of the pores or the volume occupiable by a given probe. The surfaces also changes depending on the definition. Studies have shown that occupiable volume has a better accordance with experimental data.[vol_Ongari2017]

perature (77 K); although different probe adsorbates can be considered, they are not standard.[Tian_2017] However, the definition of the surface area depends highly on the condition of measurement but also on the fitting methodology; a dozen isotherms were given to 61 labs for BET area calculation, and the statistical experiment yielded to a high level of disparities in the calculation.[Osterrieth_2022]

Beyond the experimental techniques, some software programs like Zeo++ or PoreBlazer focus on computing pore size distributions, surface areas and void fractions using well-defined structure files.[Zeo++, PoreBlazer] The definition of these values depend also on the probe size chosen to model a given adsorbate, the size of the framework atoms and the quality of the input structure. The computational values do not rely on adsorption models or on isotherm data as in the BET area, they are now relying on more comprehensive structural data. As illustrated figure 1.1, they however very much rely on a well-designed definition of the volume, the surface and the pore size we want to evaluate. Moreover, these values also highly depend on the radii of the framework atoms and the adsorbate we consider.[Hung_2021]. In this thesis, we will rely on these computational methods to define these geometrical descriptors of nanoporous materials.

CLASSES OF NANOPOROUS MATERIALS

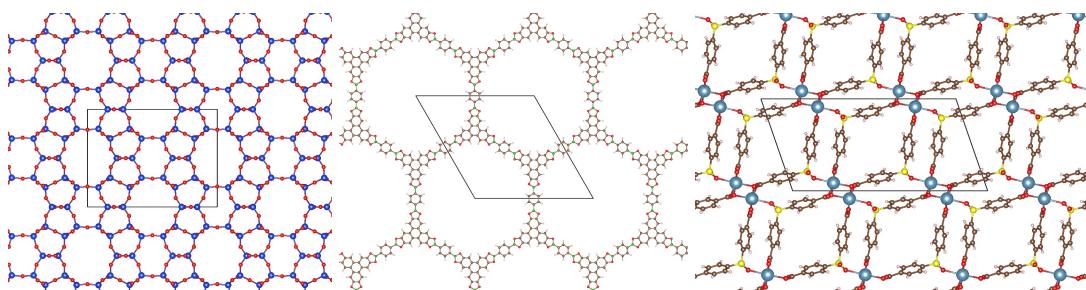


Figure 1.2: Representation of a zeolite FER[FER], a COF[Cote_2005] and a MOF[KAXQIL]. Color code: brown for C, white for H, red for O, blue for Si, cyan for Ca, yellow for S and green for B.

Nanoporous materials can have different degrees of crystallinity from perfectly crystalline to completely amorphous. Most of the computational work is focused on crystalline structures,

since the atoms are well-described within a periodic framework, which enables faster simulations. The presence of defects are also usually neglected, which could explain some discrepancies between simulations and experiments. And amorphous materials are described by thousands of atomic positions in order to grasp their intrinsic non-periodicity.[**Thyagarajan_2020**] Activated carbons, a famous class of amorphous material, are extensively used in the industry for gas purification, but cannot be rationally studied to characterize their adsorption properties. One can distinguish roughly three main classes of crystalline nanoporous materials: the inorganic materials like zeolites (aluminosilicates or aluminophosphates), the organic materials like the porous polymer networks (PPNs) or the covalent organic frameworks (COFs) and the metal–organic frameworks (MOFs).

Zeolites are naturally occurring nanoporous aluminosilicate materials that are commonly synthesized to be used in the industry as a commercial adsorbent and heterogeneous catalyst.[**Ozin_1989, Ma_2000**] It is considered as one of the most mature nanoporous material technology at our disposal. This class of material leaves also a wide room for innovation since different Al/Si ratios of a same zeolite type pan out a wide range of structures. Furthermore, zeolite materials inspired the synthesis of zeolitic frameworks harboring different atoms such as the aluminophosphates or the zeolitic imidazolate frameworks.[**Wang_2012, Chen_2014_zeo**]

Porous polymer networks (PPNs) are porous materials based on the already very developed polymer material technology.[**Lu_2010, Wang_2020, Che_2020**] However, one of the major drawback of this type of material is the formation of irreversible covalent bonds, which make the synthesis kinetically controlled leading to difficulties in crystallizing PPNs.[**Feng_2012**] In order to create crystalline porous materials, Cote et al. figured out a way of using boron-based organic compounds to form reversible bounds, which formed thermodynamically stable materials COF-1 and COF-5.[**Cote_2005**] This initiative was lead by the group of Yaghi who was at the initiative of another very promising and well-known class of materials. A decade earlier, Yaghi hydrothermally synthesized a metal–organic framework presenting large rectangular channels.[**Yaghi_1995**]

Metal–organic frameworks (MOFs) are a class of nanoporous materials formed by metallic centers connected with organic linkers to form a stable crystalline solid. Even if the first synthesis of such a material was done since the early 90s,[**Abrahams_1991**] and brought about a sparking interest in the scientific community a couple of decades later.[**Kuppler_2009, Furukawa_2013**] Because plenty of combinations of linkers and metals are imaginable, an infinite amount of MOFs could theoretically be designed. Their structure can be tuned to our needs to boost their performance in the targeted application.[**Ejsmont_2021**] This diversity of nanoporous materials offer a wide range of potential candidates that could be evaluated for any targeted applications.

1.1.2 Computational databases

All the previously described materials have been either synthesized and resolved using X-ray crystallography or computationally constructed. By combining almost all possible nanoporous materials, almost a million structures have been considered for separation or storage applications.[**Simon_2015, Simon_2015_EES, Thornton_2017**] This extended database can be broken down into the synthesized materials and hypothetical ones for all the above-mentioned classes of material.

The International Zeolite Association (IZA) gave a standardized set of 244 zeolites (in their idealized all-silica form) that can be used for screening purposes. To generate a dataset of structures, existing experimental database like the Cambridge Structural Database can be exploited. However, the raw structures determined experimentally by X-ray cannot be used directly as is. To obtain a computation-ready dataset, Chung et al. used algorithmic cleaning procedures to build the publicly available Computation-Ready Experimental MOF (CoRE MOF) database.[**Chung_2014, Chung_2019**] CoRE MOF 2019 contains about 14,000 MOF structures, which is the biggest experimental database. Similar approach applied to organic frameworks led to the construction of a set of 187 COFs with disorder-free and solvent-free structures.[**Tong_2017, Ongari_2019**]

These experiment-based databases can already be used in computational screenings to retrieve valuable information, but unknown structures that are yet to be discovered are not represented. To overcome the limits and biases of experimental synthesis, artificial ways of generating nanoporous material datasets can be used, which proved to be extremely efficient. The first *in silico* generated database of about 130,000 MOFs used a recursion-based assembly (or tinkertoy-like) algorithm to combine 102 building blocks.[**Wilmer_2012**] Martin and Haranczyk then proposed a topology-specific structure assembly algorithm that leverage the topological information of the structures.[**Martin_2014**] Inspired by this algorithm, topology-based databases emerged a few years later with the set of 13,000 MOF structures generated using the Topologically Based Crystal Constructor (ToBaCCo) algorithm developed by Colon, Gómez-Gualdrón and Snurr.[**Colon_2017**] Later, Boyd and Woo proposed another topology-based algorithm using a graph theoretical approach and generated a 300,000 structures database (BW-DB) based on 46 different network topologies.[**Boyd_2016**] Similar approaches are used for other classes of materials, Deem and coworkers proposed a dataset of nearly 2.6 million hypothetical zeolite structures.[**Earl_2006, Deem_2009, Popale_2011**] However, one could wonder if these hypothetical structures are synthesizable and can remain stable under operational conditions (e.g. thermal, mechanical, radioactive constraints). To discuss their synthetic likelihood, Anderson and Gómez-Gualdrón computed the free energies of 8,500 hypothetical structures and compared them to experimentally observed MOF structures.[**Anderson_2020**] Later, Nandy et al. performed a meta-analysis of thousands of articles associated to the CoRE MOF 2019 database to extract their experimental solvent-removal stability and thermal decomposition temperature.[**Nandy_2021**] These data are then leveraged in the training of multiple ML models to predict stability properties. These predictions can be very useful to gauge the relative stability of each material and to only consider stable structures. Other types of materials have been explored, Turcani et al. published 60,000 organic cage structures and used machine learning to predict their stability based on the shape persistence metric.[**Turcani_2018**]

The Materials Genome Initiative, a 100 million dollar effort from the White House that aims to “discover, develop, and deploy new materials twice as fast”, led to the creation of the “Materials Project”, a centralized database containing all the above-mentioned structures.[**kalil2011national, Matgenome, Jain_2013**] The fast development of this nanoporous materials genome motivated Boyd et al. to write a comprehensive review on all the initiatives on generating new data for computational analysis.[**Boyd_2017**]

Yet, the sole increase in size of the databases is not enough. One needs to add diversity to have more general knowledge on the maximum performance and the explanatory features of such per-

formance. Moreover, the diversity of structures ensure the quality of the predicted best materials for a given application. To qualitatively or quantitatively assess the diversity of a database, inventive methodologies have been developed. For instance, Martin, Smit and Haranczyk proposed a Voronoi hologram representation as a way of measuring similarities between structures to generate geometrically diverse subsets of a database.[[Martin_2011](#)] Moosavi et al. made a comparative study of the diversity of three well-known databases CoRE MOF 2019,[[Chung_2019](#)] BW-DB[[Boyd_2016](#)] and ToBaCCo[[Gomez_Gualdrón_2016](#), [Colon_2017](#)] using geometrical and chemical descriptors to design a theoretical strategy for generating the most diverse set of materials.[[Moosavi_2020](#)] Another approach consists in searching for similarities instead of differences in the materials by studying topological patterns in the data.[[Lee_2017](#)] These investigations on the data structures give a solid ground to develop novel materials by objectively defining similarity, diversity and novelty. From the analysis gathered so far, one would need to radically change the approach by proposing materials with new chemistry, topology or mechanism (e.g. flexibility) in order to significantly improve the diversity of the current databases.

1.1.3 Exploring the chemical and structural space

With the development of ever-increasing nanoporous material databases, computational chemists proposed more and more inventive methods to evaluate or screen thousands of structures. Other challenges arose, such as the design of more efficient methods than the brute force screening or the analysis of big data. Two research groups in Northwestern University led by R. Snurr and J. Hupp began to address those questions, they used a “funnel-like” approach to efficiently screen about 130,000 hypothetical MOF structures.[[Wilmer_2012](#)] To do so, they performed a first screening involving fewer steps of simulation on the whole dataset, then they extracted a subset of top performing structures to perform a second round with more simulation steps. This procedure is repeated until a few materials are selected by a final round of simulations with reasonable accuracy. Similar “funnel-like” procedures have then been used in other field of applications as described in the Figure 1.3. This type of screening saves precious computation time by balancing the complexity of the calculation with the amount of data to be screened. The most demanding simulations or experiments are only applied to the few most promising structures. This method can rather efficiently identify top candidates, but it can't draw quantitative structure-property relationships (QSPR), beside facing scalability issues above a critical dataset size.

To overcome these new challenges, people are looking increasingly towards transferable models trained by a machine learning (ML) algorithm on a diverse and size-limited sub-sample. Ideally, such a model is transferable to potentially millions of structures and can provide valuable QSPR. For instance, Fernandez et al.[[Fernandez_2013](#)] used multiple linear regression analysis, decision tree regression, and nonlinear support-vector machine models to extract QSPR and establish rules of designing well-performing MOFs for methane storage, while identifying promising structures. In this first work they only used geometrical descriptors to describe methane storage,[[Fernandez_2013](#)] but realizing the importance of chemical descriptors, they proposed the atomic property weighted radial distribution function as a powerful descriptor to predict CO₂ uptakes.[[Fernandez_2013_rdf](#)] More importantly, they proved that ML can be used as a pre-screening tool to avoid running time-costly simulations by correctly identifying around 95 % of the top 1000 best performing materials. Recently, the same group used similar

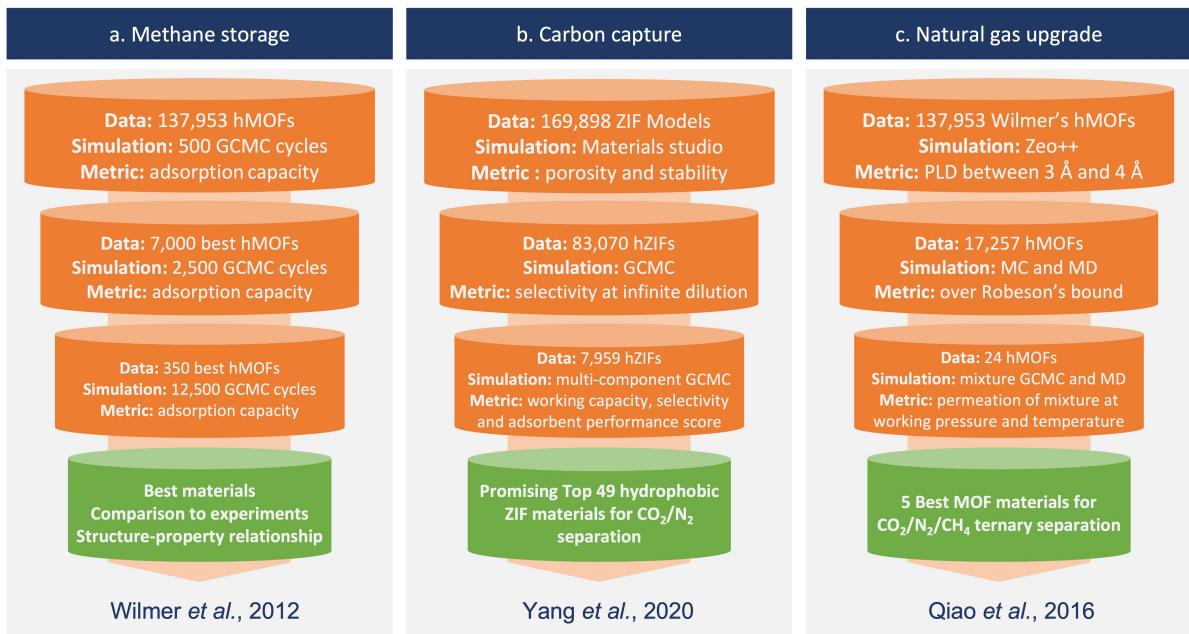


Figure 1.3: Simplified representation of typical funnel-type screening procedures, exemplified on three different applications from the published literature. (a) Wilmer et al.[Wilmer_2012] used a series of bi-component Grand Canonical Monte Carlo (GCMC) calculations at different levels of complexity to screen a large dataset of hypothetical MOFs for methane storage application. (b) Yang et al.[Yang_2020] used simulations at infinite dilution to pre-screen the dataset before using computationally demanding simulations and multiple metrics to find the most promising ZIFs for carbon capture. (c) In Qiao et al.[Qiao_2016], transport properties were screened along standard adsorption properties to find the best materials for the targeted CO₂/N₂/CH₄ ternary separation; similarly, cheaper calculations at infinite dilution were carried out in a first step, before using more expensive calculations at working pressure and temperature.

techniques to predict CO₂ working capacity as well as CO₂/H₂ selectivity in MOFs for pre-combustion carbon capture.[Dureckova_2019]

1.2 AN OVERVIEW OF SCREENING METHODOLOGIES

1.2.1 Non-adsorption properties

Due to their high internal surface area, adsorption applications were a natural outlet for nanoporous materials. However, these materials can be used in many other applications. This section is dedicated to the physical and chemical properties not directly related to the adsorption process inside nanoporous materials such as catalytic activity,[Singh_2015, Greeley_2006, Back_2020] mechanical properties,[Chibani_2019, Gaillac_2020] or thermal properties.[Toher_2014, Sarikurt_2020, I...] These properties require a more refined description of the atomic interactions within the material. DFT simulations are usually performed to accurately retrieve these properties. However, the computational cost required is multiplied by several orders of magnitude compared to classical simulations. The size of the datasets screened are therefore much smaller (a few hundreds maximum), and the use of ML can potentially speed up the whole process. ML is based on lower cost descriptors,[Evans_2017, Ducamp_2022] or it can be used in ML potentials for molecular simulations[Eckhoff_2019, Friederich_2021].

CATALYTIC ACTIVITY

Beyond adsorption properties, screening procedures have been applied to chemical properties such as catalytic activities. Heterogeneous catalysis is generally performed using metallic nonporous structures, the use of nanoporous materials can increase dramatically the active surface area and the catalytic activity. Consequently, MOFs have been demonstrated to show catalytic properties for several chemical reactions. Just to cite a few, one can think of hydrogenation, hydrolysis, oxidation, among others explicitly covered by McCarver et al. in their review.[[McCarver_2021](#)] Considering the sheer amount of possible materials, computational studies are potentially more effective than experimental ones. Therefore, computational screenings evolved in the last decade aiming at studying larger datasets.

Although the vast majority of computational screenings have been done on small series, there are a few systematic screenings of larger datasets. The scarcity of the latter can be explained by the high level of computational cost required. Here, we show some examples of such attempts by focusing on the example of C–H bond activation for the conversion of alkanes into alcohols in the presence of nitrous oxide.

Inspired by enzymatic catalysis of the reaction of small alkanes with N₂O into alcohols, Vogiatzis et al. identified 7 iron containing MOF structures out of 5,000 structures from the CoRE MOF database.[[Vogiatzis_2016](#)] They found two descriptors that govern the catalytic activity: 1) the N–O dissociation energy of N₂O on the adsorption site and 2) the energy difference between two spin states of the intermediate. Using a screening on these descriptors, three structures were identified as promising for further experimental studies. The best one has been computationally demonstrated to catalytically and selectively oxidize ethane to ethanol in presence of N₂O. Moreover, the authors found that defects played a major role in the observed catalytic activity.

Later, Rosen et al. enlarged the scope of materials screened to other metals.[[Rosen_2019](#)] From an 838 DFT-optimized MOFs subset of CoRE MOF 2014, the authors selected 168 MOFs that were likely to have open metal sites and pore-limiting diameters that allows the diffusion of the reactants. They then used a fully automated workflow to place the reactants in the adsorption site and relaxed the system using periodic DFT calculations. As shown in Figure 1.4, using the bond activation energy E_{a,C–H} and the metal-oxo formation energy ΔE_O as key parameters, they classified the materials according to their relative stability and reactivity to find the best materials for the application. These energies were then analyzed using physicochemical descriptors such as the spin density on the oxygen and the metal–oxygen distance.

This type of brute force screening can be quickly cumbersome, as a result many researchers in the field are trying to find key structure-activity relationships to accelerate future computational screenings. Several descriptors have been developed for high-throughput screenings: Butler et al. used electron removal energies to explain photocatalytic behaviors of MOFs;[[Butler_2014](#)] Rosen et al. showed that the energy required to form the metal-oxide intermediate was a key descriptor of the thermal catalysis of alkane oxidation by N₂O;[[Rosen_HTPDFT_2019](#)] and Fumanal et al. show a screening protocol based on two energy-based descriptors to predict photocatalytic properties of MOFs.[[Fumanal_descriptor_2020](#)] Lately, Rosen et al. screened thousands of MOF structures to compare different DFT functionals and leveraged

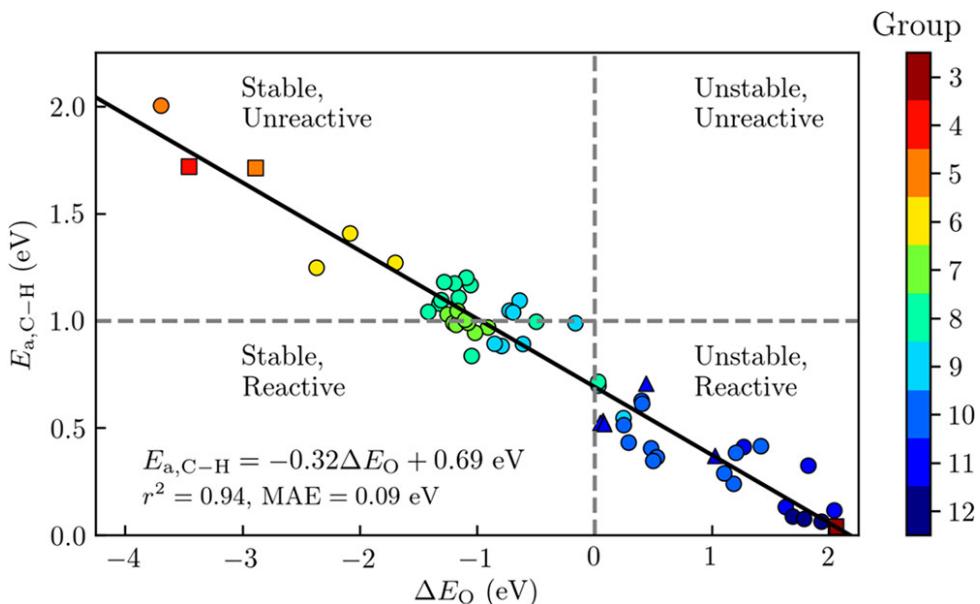


Figure 1.4: Analysis of a diverse set of experimentally derived metal–organic frameworks (MOFs) with accessible metal sites for the oxidative activation of methane. The graph shows the predicted barrier for the C–H bond activation of methane, E_a , as a function of the metal-oxo formation energy, ΔE_O . For each material, the symbol color refers to the group number of the metal in the periodic table. The best-fit line has is plotted in black, and has a mean absolute error (MAE) of 0.09 eV. MOFs with $E_a < 1$ eV are classified as being reactive toward C–H bond activation and MOFs with $\Delta E_O < 0$ as having thermodynamically favored active sites when using O_2 as the reference state. Reprinted with permission from Ref. Rosen_2019. Copyright 2019 American Chemical Society.

the data calculated to train machine learning models that can rapidly predict MOF band gaps.[Rosen_2022_high]

The development of ML methods are also critical in the field,[Rosen_2021] but the lack of centralized database with high precision descriptors is a challenge for the future of these methods. The influence of defects, the different ways of modelling MOFs as periodic structures or clusters, the diversity of structures and the stability of such structures remain open problems. Yet, it does not threaten the major role of high-throughput screenings in the early design process of any nanoporous materials for catalysis. To conclude this brief overview, we point the readers to a more exhaustive presentation of the matter.[Rosen_2022]

MECHANICAL PROPERTIES

In the past decade, there has been a growing interest in the systematic study of physical properties of various classes of materials, including inorganic materials and framework materials. Among these physical properties, mechanical properties have been a topic of particular interest, as they are crucial for many applications, and at the same time can be computed by relatively standard methodologies. In particular, is it possible to calculate linear elastic constants (the second-order elastic tensor) in the zero-Kelvin limit by strain/stress or strain/energy approaches, performing a series of DFT calculations of strained structures and calculating the elastic constants. From these constants, all other mechanical properties can be evaluated by tensorial analysis,[Marmier_2010] including the bulk modulus, Young's modulus, shear

modulus, Poisson's ratio, etc. This type of calculation can be coupled with any available quantum chemistry code,[**Golesorkhtabar_2013**] and is even integrated in some packages, like CRYSTAL17.[**Dovesi_2018**]

One of the first studies that investigated systematically the elastic properties of a family of materials was a 2013 study of all-silica zeolites,[**Coudert_2013**] i.e., crystalline and porous SiO₂ polymorphs. While this dealt with only 121 zeolitic frameworks out of 244 known structures, it showed that systematic studies at the DFT level were computationally tractable, and that they provided physical insight into the link between microscopic structure and macroscopic physical properties. This study demonstrated, among other things, that a few zeolites presented large negative linear compressibility (NLC), which could be linked to the wine-rack motif of their frameworks.

Outside the specific case of zeolites, other groups have applied DFT calculations of elastic constants in a high-throughput manner. de Jong et al. leveraged the structures of the Materials Project[**Matgenome, Jain_2013**], trying to chart the diversity of elastic properties across the whole space of inorganic crystalline compounds.[**deJong_2015**] As shown in the Figure 1.5, they provided a database containing the full elastic information of 1,181 inorganic compounds initially, and has grown steadily since then, containing more almost 14,000 records to date. [**MaterialsProject**] This dataset has been used in two different ways by researchers in the field.

Firstly, the exploration of the database of elastic properties by tensorial analysis has allowed studying quantitatively the occurrence of certain "anomalous" or rare mechanical behavior, including negative linear compressibility, very high anisotropy, or negative Poisson's ratio (also called *auxeticity*). Indeed, such properties are considered rare and usually sought after — the materials exhibiting these anomalous behaviors are mechanical metamaterials.[**Coudert_2019**] In addition to their fundamental interest, such materials have applications in materials engineering: for example in energy dissipation (as shock absorbers and for bulletproofing), energy storage, as well as acoustics.[**Surjadi_2018**] However, it was not possible until now to quantify exactly "how rare" they are. Chibani et al. showed through a systematic exploration of available mechanical properties of crystalline materials that general mechanical trends, which hold for isotropic (noncrystalline) materials at the macroscopic scale, also apply on average for crystals. Moreover, they could quantify the presence of materials with rare anomalous mechanical properties: 3% of the crystals were found to feature negative linear compressibility, and only 0.3% to exhibit complete auxeticity (negative Poisson's ratio in all directions of space).

Secondly, the datasets of mechanical properties were used as a basis to accelerate the discovery of novel materials with targeted behavior. Dagdelen et al. used search algorithms to identify 38 candidate materials exhibiting features correlating with auxetic behavior, from more than 67,000 materials in the Materials Project database.[**Dagdelen_2017**] Performing DFT calculations on these 38 structures, they could identify 7 new auxetic compounds. In a more complex setup, Gaillac et al. [**Gaillac_2020**] have used a multiscale modelling strategy for the fast exploration and identification of novel auxetic materials. They combined classical force fields MD simulations with DFT calculations on candidate materials, and then used this reference DFT data to train an ML algorithm. They found that the accuracy of this multiscale method exceeds the current low-computational-cost approaches for screening. In a similar work, Moghadam et

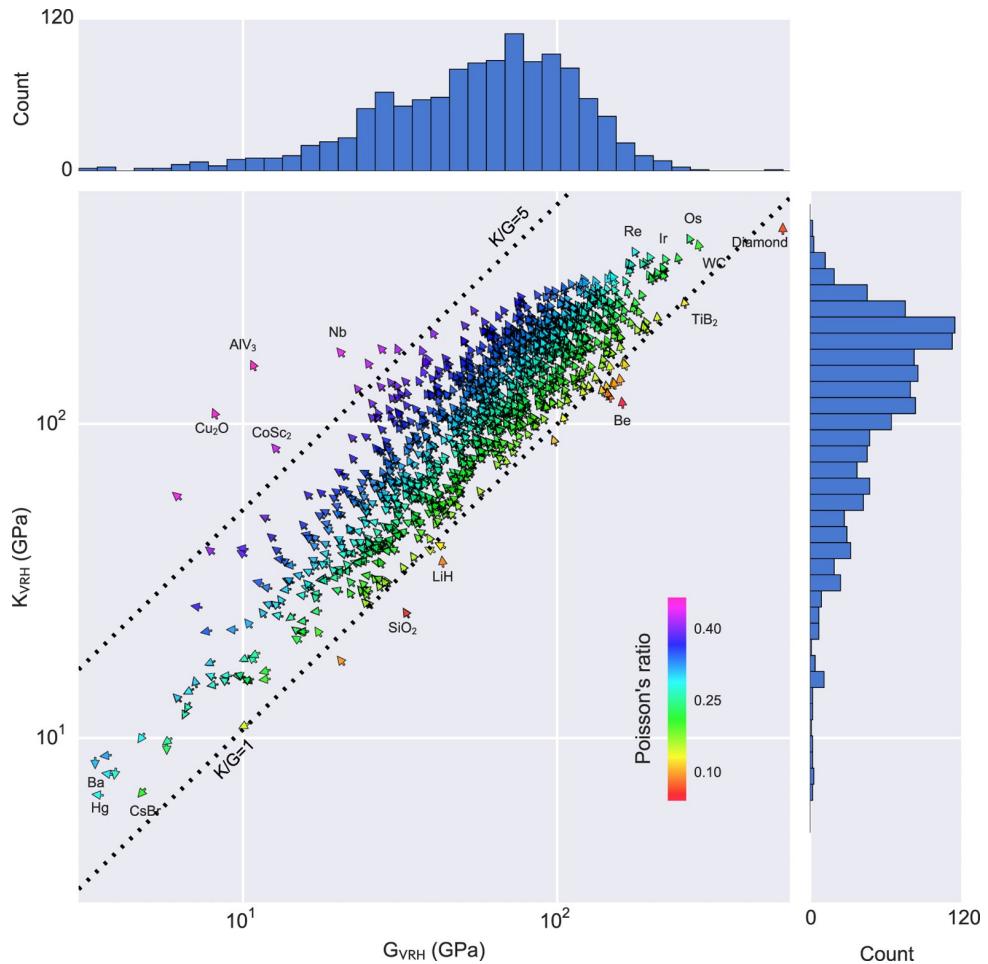


Figure 1.5: Statistical analysis of the calculated volume per atom, Poisson's ratio, bulk modulus K_{VRH} and shear modulus G_{VRH} of 1,181 compounds in the Materials Project database. In the vector field-plot, arrows pointing at 12 o'clock correspond to minimum volume-per-atom and move anti-clockwise in the direction of maximum volume-per-atom, which is located at 6 o'clock. Reprinted from Ref. deJong_2015 under CC-BY license. Copyright 2015 de Jong et al.

al. used molecular simulation to train an artificial neural network (ANN) for the prediction of the bulk modulus of metal–organic frameworks.[**Moghadam_2019**] This shows the potential of such methodologies to treat very different (chemically as well as structurally) classes of materials.

Thermal Properties

While mechanical properties (in the elastic regime) have been by far the most studied physical property in nanoporous materials, others have also been occasionally screened. We can cite, in particular, the systematic study of piezoelectric tensors by de Jong et al., on almost a thousand crystalline compounds, by first-principles calculations based on density functional perturbation theory.[**deJong2015_piezo**] We can also cite efforts to calculate thermal properties in a high-throughput setup, using the quasi-harmonic approximation (QHA).[**Togo_2010**] This method requires the calculation of each structure's phonon modes at various volumes, and can be coupled to any electronic structure program.[**Togo_2015**] It is, however, quite computationally intensive, and sensitive to the parameters of the QHA methodology (range of volume, range of temperature, precision of the frequency calculation, etc.). Therefore, it has been limited

so far to modest numbers of structures: a dataset of 75 inorganic structures by Toher et al.,[Toher_2014] and more recently a dataset of 134 pure SiO₂ zeolites by Ducamp et al. [Ducamp_2021] Very recent work in our group on the prediction of thermal properties through machine learning based on structural features alone indicates that thermal behavior is more difficult than mechanical behavior to predict, and might require the use of a wider set of structural descriptors, or more advanced ML models.[Ducamp_2022]

1.2.2 Transport adsorption properties

In the previous section, the thermodynamic properties only described the state of equilibrium of the adsorption process. But sometimes the transient state can last long before reaching the equilibrium, which makes the process more time-consuming. Thus, the transport properties complete the thermodynamic description of the adsorption process inside a nanoporous material. For example, a low diffusion rate would mean for storage applications more time and energy needed to fill-up the tanks, or for separation applications a less selective process than expected. In more extreme cases of molecular sieves for fluid separation, the transport properties become predominant to assess the performance. One can leverage the difference of the molecules' diffusion coefficients to selectively filter gas mixtures through a nanoporous membrane.[Miandoab_2021] Here, the main subject becomes the transient state and not the equilibrium. This section is thus dedicated to the kinetics of the adsorption process to better model the time required to reach the equilibrium or to study out-of-equilibrium processes such as molecular sieving by nanoporous membranes.

KINETIC PROPERTIES

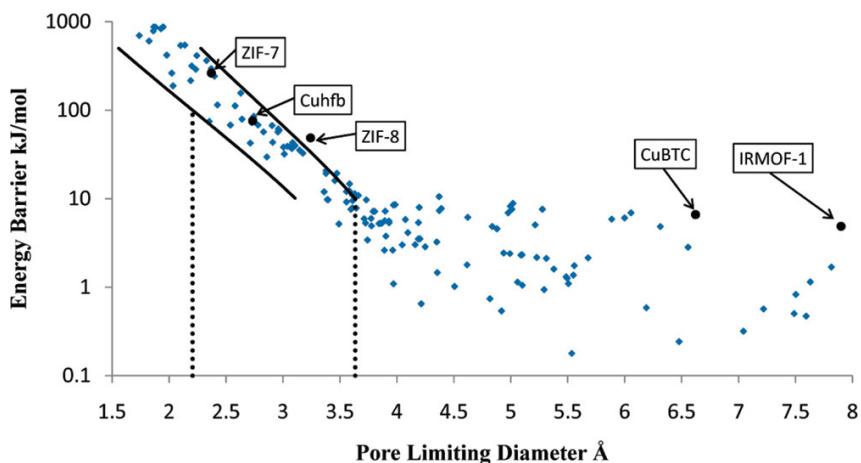
In most computational screenings, the diffusion coefficient considered is the self-diffusion coefficient that describes an infinite-dilution case. Other multi-component diffusion coefficients could be considered, but for simplicity and clarity they won't be mentioned in this review. The calculation of the self-diffusion coefficient gives a first estimation of the kinetics in a storage or a separation process in the limit of low adsorption loading.

There are two approaches to estimate the diffusion inside a porous material: the first one relies on molecular dynamics (MD) and the second one on transition state theories. In the first approach, one can analyze the mean squared displacement of the adsorbed molecule moving in the material. In the second, one identifies minimum energy path along the material to identify transition states (TS) to calculate diffusion energy barriers. The MD-based method requires fewer assumptions and is therefore more reliable than the TS-based method, but the latter is computationally more efficient in the case of low diffusion rate (diffusivity lower than 10⁻¹¹ m² s⁻¹).

State-of-the-art MD simulations could calculate rather accurate diffusion coefficients, but the computational cost scales quickly with the number of structures. To use this method on a large dataset without spending too much computation time, Watanabe and Sholl pre-screened the pore sizes of 1,163 MOFs to select only the structures within a certain range of PLD (pore limiting diameters).[Watanabe_2012] A restricted list of 359 MOFs was then used to carry out MD simulations to calculate diffusion coefficients. The results of this final screening are then used to extract the most promising structures for further experimental or computational investigation. Similarly, Qiao et al. used a multi-stage screening to find the best membrane-material within about 130,000 hypothetical MOFs for a CO₂/N₂/CH₄ separation.[Qiao_2016]

They started to select materials based on pore geometry analysis; then they calculated Henry's coefficient and diffusion coefficients at infinite dilution; finally they compared the binary permselectivities to extract 24 promising MOFs for ternary adsorption and diffusion calculation at the desired pressure and temperature conditions.

Another approach replaces MD simulations with more computationally efficient TS-based methods to determine diffusion coefficients. Haldoupis et al. developed an algorithm to identify diffusion paths by exploiting an energy grid with a clustering algorithm. The diffusion paths are then analyzed to identify the pores and the channels, and to calculate key geometric (PLD, the largest cavity diameter) and energetic (Henry's constant, diffusion activation energy) features.[**Haldoupis_2010**] As represented in see Figure 1.6, they found a clear dependence of the diffusion energy barrier to the PLD. As one of the first TS-based screenings, it is still subject to many development perspectives. For instance, the approach is limited to spherical adsorbates and rigid frameworks. Moreover, the diffusion coefficients are approximated using a simplistic hopping model for a qualitative analysis. This method is highly efficient, but the accumulation of approximations makes a quantitative systematic analysis of diffusion coefficients out of reach.



*Figure 1.6: Calculated energy barrier for the diffusion of CH₄ in 216 metal–organic frameworks (MOFs), shown as a function of the pore-limiting diameter. The solid lines represent statistical upper and lower bounds on the energy barrier, in a transition state theory approach. Reprinted with permission from Ref. **Haldoupis_2010**. Copyright 2010 American Chemical Society.*

Later, Kim et al. introduced a flood fill algorithm to obtain all the points within a given energy.[**Kim_2013**] These points are then identified as channels or blocked regions. Along the channels, local minimums of energy are defined as lattice sites and transition states are defined perpendicular to the diffusion direction. A random walk is then computed along the lattice sites with hop-rates defined according to the activation energy. A diffusion coefficient is then calculated in each three directions of the space and an average diffusion coefficients is finally determined. A comparison with the MD method on the IZA zeolite structures shows good agreement, but there are still some discrepancies explained by correlated hops in the case of rapid diffusion or by the presence of complicated channel profiles. Inspired by this work, Mace et al. developed a similar method that progressively fill the energy grid to detect transition states, hence removing the previous restriction to orthogonal cells only.[**Mace_2019**] The diffusion coefficient is now computed using a kinetic Monte Carlo simulation allowing the

adsorbate to jump freely in all directions instead of restricting it in a single dimension. This new method, called TuTraSt, handles very complex diffusion paths (like in the AEI zeolite). This new approach seems to be promising as it is in good agreement with MD simulations, while being 2-3 orders of magnitude faster. However, the time performance could improve tremendously by translating it from Matlab to C++ and by implementing parallelization procedures.

Very recently a massively parallel GPU-accelerated string method has been implemented and shared publicly to compute very efficiently diffusion coefficients based on the transition state theory.[**Zhou_2021**] The recent developments in the prediction of diffusion coefficients in nanoporous materials point towards a promising future for the screening of transport properties applied to even larger databases. Going further, Bukowski et al. reviewed thoroughly diffusion in nanoporous solids as an attempt to connect theory to experiments.[**Bukowski_2021**]

MEMBRANE MATERIALS

In separation application, the study of the transport properties can evaluate the feasibility of the thermodynamic equilibrium, crucial for any bed separation process. If this separation is not feasible, kinetic separation or partial molecular sieving are to be considered. Some notable examples are: air separation in zeolites using pressure swing adsorption,[**ruthven1990air**] N₂/O₂ separation in carbon molecular sieves,[**Reid_1999**] or N₂ removal from natural gas. [**Wang_2019**] In kinetic separation, the valuable metric is not the selectivity anymore, but the permselectivity, i.e. the product of the selectivity and the permeability (ratio of diffusion coefficients). Therefore, the screening of diffusion coefficients gives complementary information to the thermodynamic selectivity screenings. Here, we give some examples of such screening and the main descriptors that partially explains the computed figures of merit.

To give an overview on the potential of computational screenings to predict transport properties, we are now going to focus on the membrane separation applied to natural gas upgrading. The separation of CH₄ from N₂ and CO₂ is a crucial step of this upgrading process. In 2016, a large scale high-throughput screening (see Figure 1.3 for the approach) of hypothetical MOF membranes for upgrading natural gas has been performed using MD simulations.[**Qiao_2016**] Qiao et al. confirmed the existence of MOF materials beyond the upper bound for N₂/CH₄ and CO₂/CH₄ separations determined by Robeson on a large set of polymeric membranes. [**robeson1991correlation**] This Robeson's upper bound is systematically crossed by MOF materials in computational screenings, see as an example the Figure 1.7. This can be explained by the fact that MOFs perform better than polymeric frameworks and the simulations at this level of theory They also identified 24 MOFs suitable for the ternary CO₂/N₂/CH₄ separation using a multi-stage screening described in the previous section.

Two years later, Qiao et al. used the same approach to study this ternary separation on a database of synthesized structures.[**Qiao_2018**] Applying machine learning techniques to their data, they performed a QSPR analysis. Using a principal component analysis, they notably found that the permeability is higher when materials have high PLD and void fraction coupled with low density and percentage of pores within a characteristic range. The opposite was found to be true for high membrane selectivity for the CO₂/CH₄ separation. Using decision tree algorithms, they gave objective procedures of selecting the best separation membranes based on some key descriptors. Finally, they studied in detail some top performing materials found by a support vector machine algorithm.

Altintas and Keskin later performed a screening on the same database for CO_2/CH_4 membrane separation to identify the best performing materials and perform more computationally demanding simulations.[[Altintas_2018](#)] The simulations in rigid structures at infinite dilution show numerous structures above the Robeson's upper bound as shown in figure 1.7, this crossing of the upper bound can be explained by either a better performance of MOF membranes compared to the polymeric membranes used by Robeson, or an overestimation due to oversimplified assumptions (infinite dilution, rigidity). But when higher pressures and flexibility are considered, the selectivity values are dropping down closer to the upper boundary, hence confirming the overestimation of the performance in screenings based on rigid approximations at infinite dilution. But the best performing materials are still above the Robeson's upper bound and can therefore be used in mixed matrix membranes with polymeric membranes. Budhathoki et al. developed a screening methodology for MOFs in mixed matrix membranes for carbon capture applications by estimating permeation values in these composite materials using a Maxwell model.[[Budhathoki_2019](#)] The authors even proposed a pricing for each material compared to their relative performance. Similar studies have been carried out on different materials, Yan et al. showed the influence of decorating COFs with different chemical compounds on the membrane selectivity.[[Yan_2018](#)]

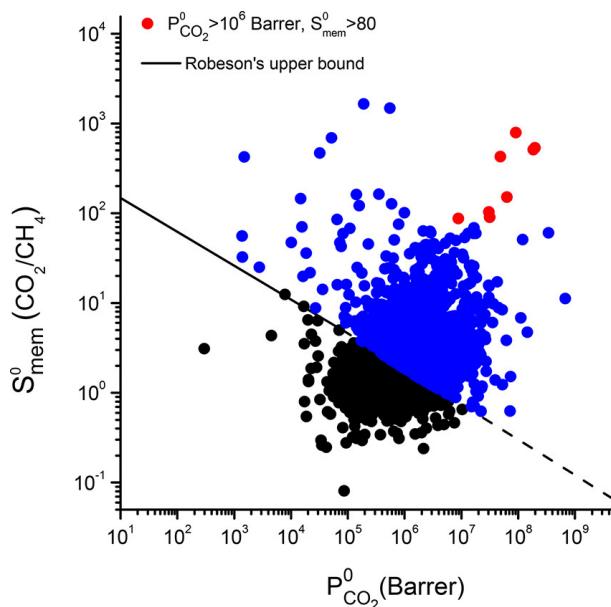


Figure 1.7: Selectivity and permeability of metal–organic framework (MOF) membranes for CO_2/CH_4 separation, computed at infinite dilution by combining Grand Canonical Monte Carlo and molecular dynamics simulations.[[Altintas_2018](#)] The black solid line represents the Robeson's upper bound.[[robeson1991correlation](#), [Robeson_2008](#)] MOFs that can exceed the bound are shown in blue, and the 8 top-performing MOF membranes are shown with red symbols. Reprinted with permission from Ref. [Altintas_2018](#). Copyright 2018 American Chemical Society.

The transport properties screening is based on the calculation of diffusion coefficients at infinite dilution and in rigid molecules. There are different methods to calculate them (mainly MD and TS-based methods). Flexibility and pressure dependence are very hard to incorporate directly in the screening procedures. Researchers usually consider these factors at the end of the screening on the most promising structures because of the computational complexity of the corresponding simulations. To take account of pressure dependence, we need an MD

simulation of several adsorbates that takes much more time than running single component simulations,[**Keskin_2007, Keskin_2009**] which makes it harder to include in a high-throughput screening. Flexibility could be taken account by calculating snapshots and running multiple MD simulations, or by using flexible force fields, which means in both cases an increase in computational run-time. Some faster methods of quantitatively predicting the impact of flexibility on diffusion are being investigated in ZIFs and could give an interesting alternative to these expensive methodologies.[**Han_2020**]

1.2.3 Thermodynamic adsorption properties

In its early development, computational screening was mainly used to predict thermodynamic properties in adsorption processes. Three main applications have been identified in the associated literature: gas storage (for energy or medical applications), gas separation (noble gas, hydrocarbons, carbon dioxide, etc.) and post-combustion CO₂ capture. These applications are closely linked to urgent environmental and energy issues that are yet to be solved. Screening can guide the development of better performing materials by shedding light upon unknown structure-property relationship, probes possible theoretical limitations (unreachable targets) and identifies potential candidates that need to be experimentally tested.

GAS STORAGE

One can leverage the high surface density of the nanoporous materials, especially the MOFs, to stock in very low-density gas. In the field of energy storage or transportation, natural gas (mainly methane) or hydrogen are considered plausible alternative fuels to replace conventional ones for transport. The US Department of Energy (US DOE) recently financed research programs and set target for methane and hydrogen storage. Nanoporous materials could reduce energy, infrastructure and security cost due to the required compression and cooling. In this section, we are focusing on high-throughput screening for methane storage in nanoporous materials, before broadening the scope hydrogen and other perspectives.

One of the pioneering works in computational screening was published in 2012 by Wilmer et al.[**Wilmer_2012**]. They performed a large-scale screening of 137,953 hypothetical MOF structures to estimate the methane storage capacity of each MOF at 35 bar and 298 K based on the US DOE standards. Back then, the US DOE set a target methane capacity value of 180 vol^{STP}vol⁻¹ (which has since been achieved by several materials reported in the literature). In their large-scale analysis, Wilmer et al. found over 300 hypothetical MOFs that meet the targeted requirements and the best one can store up to 267 vol^{STP}vol⁻¹, surpassing the state-of-the-art of the time. From their large dataset, a preliminary structure-property relationship analysis revealed that void fraction values of approximately 0.8 and gravimetric surface areas in a range 2500-3000 m²g⁻¹ resulted in the highest methane capacities. Optimal pore size are also shown to be around the size of one or two methane molecule(s). Maximization of gravimetric surface area was a common strategy in the MOF design for storage applications, but this study showed the existence of an optimal range of surface area values. Computational screenings can draw clear relationships between structural descriptors and performance. Later, a more quantitative relationship was drawn by Fernandez et al. using ML models as illustrated on Figure 1.8. Beware not to over-interpret the relation given by the response surface, since the identified maxima do not always have a physical reality, especially where there is no training data in the area pointed by the red arrows. However, it highlights promising unexplored feature space and shows potential research directions.

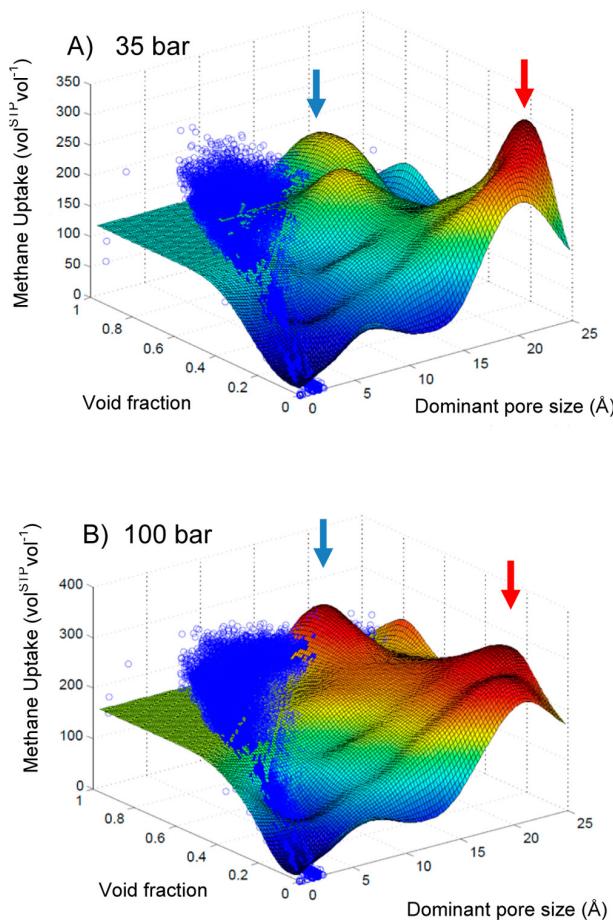


Figure 1.8: Two-dimensional response surfaces of the support vector machine (SVM) models trained by Fernandez et al. for methane storage at (A) 35 bar and (B) 100 bar using void fraction and dominant pore size. The blue dots represent the GCMC simulated uptake values. The color of the surface represents the methane storage value, from blue (the lowest values) to red (the highest values). Blue and red arrows indicate maxima on the response surface. Reprinted with permission from Ref. Fernandez_2013. Copyright 2013 American Chemical Society.

Since then new materials above the target have been found and the US DOE decided to set a higher target of $315 \text{ vol}^{\text{STP}} \text{ vol}^{-1}$. Until now, this new target is not yet reached. This is why the recent developments have focused on assessing the feasibility of such a target by accelerating the screening methods so that more data can be screened, and by interpreting the QSPR models to extract important knowledge for the design of novel materials. For instance, Gómez-Gualdrón et al. showed that even by artificially quadrupling the Lennard-Jones interaction factor ϵ and by increasing the delivery temperature by 100 K, the newly set target is only reached by a handful of MOFs. [Gomez_Gualdrón_2014] This study suggests the impossibility to reach the DOE target using a preconceived (experimentally or theoretically) material to store methane. However, this theoretical limitation can be overcome by increasing the surface density of sites with high affinity with methane and by increasing the delivery temperature.

Later, a larger-scale screening on methane storage was carried out by Simon et al. on 650,000 experimental and hypothetical structures of zeolites, MOFs, and PPNs. This study confirmed that the classes of materials currently being investigated were unlikely to meet the new target. The authors suggested that it wasn't surprising since the target was based on economical

arguments, while the screening is based on thermodynamic arguments.[**Simon_2015_EES**] This example illustrates the power of large scale screening to settle questions of physical feasibility (if simulations are accurate) and hence avoiding experimental efforts spent on impossible tasks.

More recently, a dataset containing trillions of hypothetical MOFs have been screened for methane storage.[**Lee_2021**] Lee et al. developed a methodology using machine learning combined with genetic algorithm to perform the largest screening until now. In addition to confirming most of the results (theoretical limits and QSPR) found by previous screenings, 96 MOFs were found to outperform the current world record. This study shows the scaling potential of ML-assisted screenings in handling “Big data”.

Similarly, computational high-throughput screenings have been applied to other storage applications such as hydrogen storage. Computational screenings showed that cryogenic storage of hydrogen can meet the DOE target of 50 g L^{-1} .[**Gomez_Gualdrón_2016, Bobbitt_2016, Thornton_2017**] Anderson et al. performed a large scale screening based on neural networks to test out multiple pressure/temperature swing conditions to find that the maximal deliverable capacity cannot exceed 62 g L^{-1} .[**Anderson_2018**] Compared to the density of liquid hydrogen (72 g L^{-1}), this upper limit seems reasonable since the adsorbent material takes at least 10-20% of the tank. Here, we only showed some flagship results of the field. For a more detailed meta-analysis, Bobbitt and Snurr wrote a very complete review on computational high-throughput screening of MOFs for hydrogen storage.[**Bobbitt_2019**]

XENON/KRYPTON SEPARATION

As a representative example of what could be done in the field of gas separation, we are going to focus on Xe/Kr separation. Nanoporous materials can be used as a safer, cheaper and less energy-intensive option for this gas separation. However, experimental design of top performing materials can be cumbersome. Computational screenings is an ideal tool to kick-start the development of this new technology by identifying rapidly the best candidates.

SMALL-SCALE SCREENINGS

Metal–organic frameworks, and later other supramolecular porous materials like covalent organic frameworks (COFs), have been proposed for applications in separation of noble gases for a decade. With no aim of being exhaustive, we highlight some milestones in that area, both from experimental and computational point of view.

In 2012, Liu et al.[**Liu_2012**] published an experimental study of two MOFs, HKUST-1 and Ni/DOBDC, for adsorption of Xe and Kr at ppm (part-per-million) levels in air. The target application was the removal of Xe and Kr from nuclear fuel reprocessing plants. The same group later proposed a two-column method for the separation of Kr and Xe from process off-gases[**Liu_2014**], based on MOF materials. At about the same time, Bae et al.[**Bae_2013**] combined a computational Grand Canonical Monte Carlo (GCMC) study with experimental breakthrough measurements of the separation of a Xe/Kr mixture on MOF-505 and HKUST-1.

Parkes et al.[**Parkes_2013**] studied sixteen different MOF materials for Kr, Ar, and N_2 adsorption and separation, through GCMC simulations. They concluded on the potential of MOFs for separation, and a general correlation between the Henry’s constant and the isosteric heat of adsorption for the three gases studied. A year later, in 2014, Chen et al.[**Chen_2014**]

demonstrated, again through a combined computational and experimental study, the potential of porous organic cages for selective binding of xenon over krypton.

Later experimental work expanded these early separation studies to different types of MOF materials. Xiong et al.[**Xiong_2015**] studied a flexible zinc tetrazolate framework for xenon selective adsorption over krypton, argon and nitrogen. Thermodynamic analysis of the adsorption isotherms at various temperatures confirmed the occurrence of a “breathing” structural transition upon Xe uptake, contributing to a high working capacity for a pressure swing adsorption (PSA) cycle. Lee et al.[**Lee_2016**] compared the selective adsorption properties for Xe/Kr mixtures on three highly-studied MOFs, namely UiO-66(Zr), MIL-100(Fe) and MIL-101(Cr), and confirmed a high potential of UiO-66(Zr) for separations under dynamic flow conditions. These authors also assessed the hydrothermal and radioactive stability of the material, a test seldom performed in the existing literature, and found it to be good. In a further study,[**Lee_2018**] they demonstrated that Xe/Kr selectivity could be further improved by ligand substitution.

In parallel, computational studies were published to provide insight at the microscopic level into the mechanisms behind good (and bad) separation properties. Wang et al.[**Wang_2014**] studied 6 MOFs and COFs for adsorption of Xe and Xe/N₂ separation, through GCMC simulations looking at the impact of pressure (and therefore pore filling) on selectivity. Anderson et al.[**Anderson_2017**] combined GCMC and biased MD simulations to elucidate the nature of adsorption- and diffusion-based Kr/Xe separation mechanisms in four archetypal nanoporous materials: SAPO-34, ZIF-8, UiO-66, and IRMOF-1. These authors draw a couple of general conclusions, including the fact that diffusion selectivity for krypton dominates membrane separation selectivity, and large pore cages and stiff pore windows are desirable – however the scope of these conclusions is inherently limited by the small number of materials actually studied.

In a different family of materials, Tong et al.[**Tong_2017**] have surveyed the structure–property relationships of covalent organic frameworks (COFs) for noble gas separation, by GCMC simulations of 187 different materials for Kr/Ar, Xe/Kr and Rn/Xe separations. These authors included in their calculations some adsorption figures of merit (AFM), representative of the conditions of industrial vacuum (VSA) and pressure swing adsorption (PSA) processes.

One area that has been particularly explored is the tuning and improvement of separation properties through the presence and nature of coordinatively unsaturated sites (or open metal sites) in MOFs. In 2016, Vazhappilly et al.[**Vazhappilly_2016**] used density functional theory (DFT) calculations of host–guest binding energies to probe the impact of the metal atoms in a specific framework (MOF-74) on Xe and Kr adsorption. Later, Zarabadi-Poor et al.[**ZarabadiPoor_2018**] investigated – again through computational methods – a series of metal–BTC MOFs for recovering xenon from exhaled anesthetic gas, i.e., mixtures of CO₂, O₂, and N₂.

LARGE-SCALE COMPUTATIONAL SCREENING

In its early stage, computational screening has been used on small series of nanoporous materials to generate specific knowledge on some subclasses of materials. These small-scale screenings combined with experiments helped faster identification of good performing candidates, but they failed to establish general rules of design or to explore the unknown. Larger-scale screenings

overcame these limitations by trying to exhaustively cover the whole spectrum of nanoporous materials.

The first large-scale computational screening on Xe/Kr adsorption-based was performed by Sikora et al. based on the same approach previously developed for methane storage by their group at the Northwestern University.[**Sikora_2012**] This study was based on the same 137,000 structures of hypothetical MOFs.[**Wilmer_2012**] They calculated the Xe/Kr selectivity using Monte Carlo molecular simulations on the whole database by iteratively increasing the number of steps and selecting the best materials similar to the approach on Figure 1.3. By analyzing the relationships between pore sizes and selectivity, they confirmed a hypothesis from a smaller scale study that the pores should be between the size of 1 to 2 xenon molecules.[**Ryan_2010**] Tube-like channel were also found to favor better selectivity. Moreover, they found that top performing materials could have selectivities around 500; but we can only conclude on the order of magnitude of the theoretical limitation of the Xe/Kr selectivity, considering the statistical uncertainty of the simulation.

Seizing the opportunity of a formidable expansion of the nanoporous materials database triggered by the Materials Genome Initiative, Simon et al. screened 670,000 experimental and hypothetical nanoporous material structures for Xe/Kr separation.[**Simon_2015**] It is one of the largest-scale screening performed in this area. Inspired by the work of Fernandez and co-workers,[**Fernandez_2013**] they used ML algorithms to train a model on a diverse subset of 15,000 structures. This method allowed them to run time-consuming molecular simulations only on this training set, before applying the ML model to predict the selectivity values on the larger set of structures. On top of analyzing the links between pore descriptors and selectivity, they rationalized it using theoretical pore models of spherical and cylindrical geometries to confirm the findings of Snurr and co-workers.[**Ryan_2010, Sikora_2012**] By comparing the structural descriptors of good-performing and bad-performing structures, they concluded that geometrical descriptors wasn't enough to explain the performance (see Figure 1.9). The analysis of a few top candidates suggests that different chemical insights could explain their good performance. For SBMOF-1 or KAXQIL,[**KAXQIL**] an experimental MOF, its higher performance was explained by the tube-like 1D channel with a very favorable binding site formed by carbon aromatic rings. This nanoporous material was later tested using breakthrough experiments and proved to be one of the most promising candidates.[**Banerjee_2016**] This close collaboration between computation and experimentation is a testimony of the potential of computational screenings to find nanoporous materials for any targeted application.

The experimental work on Xe/Kr separation on SBMOF-1 revealed discrepancies between the selectivity values obtained experimentally and computationally.[**Banerjee_2016**] The assumption of rigid crystal structures in the molecular simulations could partially explain the difference observed. Witman et al. proposed that the flexibility of the materials, that weren't considered in the screening of Simon et al., could explain the lower selectivity observed experimentally.[**Witman_2017**] In this study, they screened the Henry regime separation of about 4,000 MOF structures of the CoRE MOF 2014 database[**Chung_2014**], and found that intrinsic flexibility, i.e. the thermal vibration of the material, can make the pore size derive from the ideal value for the separation and hence lower the selectivity. This study further confirms the importance of the pore size by highlighting the effect of its evolution over time.

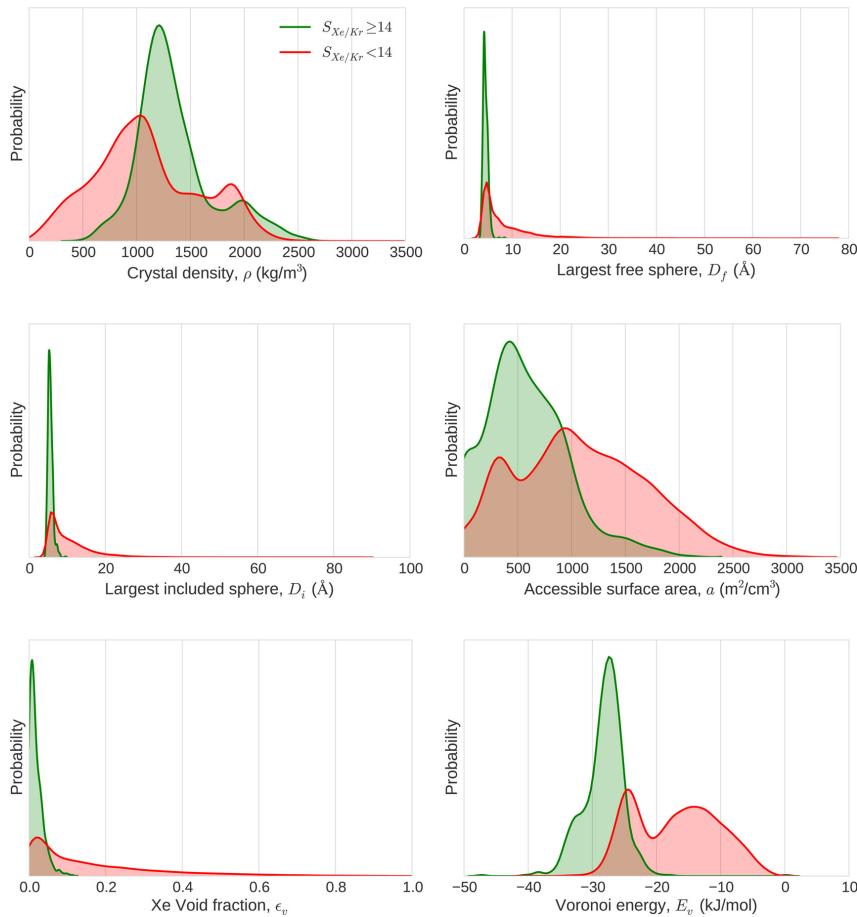


Figure 1.9: Statistical analysis of the adsorption separation of xenon/krypton mixtures by nanoporous materials. The graphs represent the distributions of structural descriptors explored by highly selective (green) and poorly selective (red) materials separately. Reprinted with permission from Ref. Simon_2015. Copyright 2015 American Chemical Society.

In 2019, Chung et al. screened the most extensive simulation-ready and experimentally synthesized MOF structures for Xe/Kr separation.[**Chung_2019**] This study pointed out the potential of coordinated solvent molecules to fine-tune the selectivity for any separation application, since their presence can enhance selectivity in some cases. The results of their screening confirms the potential of structures such as SBMOF-1 found by Simon et al., but they also described a few structures with similar selectivity but with better xenon uptake. The authors emphasize the importance of considering other figures of merit such as the adsorption capacity. Other factors should be taken into account to find the best trade-off between all the relevant figures of merit; we could think of the kinetics of such a separation, the effect of flexibility on the performance, the stability of the materials (especially in radioactive environment), the financial aspects, and more.

After this quick overview of the different screening studies in the field of xenon/krypton separation, we are now going to detail its industrial context, the foreseen top materials that could fulfill the industrial separation and what further studies are needed to better understand the process while discovering new materials.

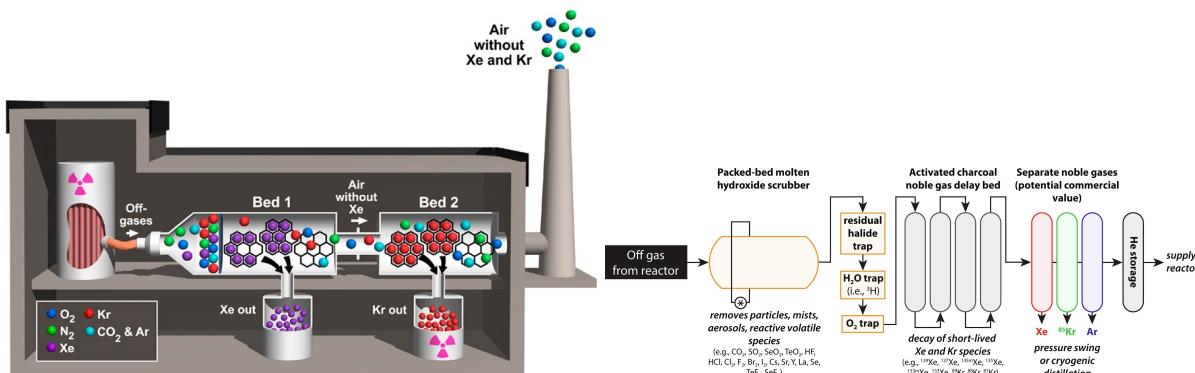
1.3 XENON/KRYPTON SEPARATION

In this section, we will try to see how we can apply the above-mentioned screening methodologies to help us understand the origins of the Xe/Kr separation and identify promising materials for industrial applications.

1.3.1 Industrial applications

The industrial interest for noble gases lies first and foremost in the many applications attached to them. For instance, xenon has multiple applications in the medical (e.g. anesthesia, painkiller, imaging), [cullen1951anesthetic, holstrater2011intranasal, Mammarappallil_2019] aeronautical[Patterson_2002, Coxhill_2005], lithographic[Abramov_2018], microelectronic[Chang_1995] or lighting sectors,[Jarman_1974, Tanaka_2019] just to cite a few. To meet the demand for these noble gases, one should consider all available sources, the most obvious one being the air we breathe. Xenon and krypton have both very low atmospheric concentrations; out of a thousand liters of air we would extract at most one-tenth of a milliliter of xenon and one of krypton.[kerry2007industrial] Nevertheless, direct extraction from the air remains the main production mean for xenon and krypton along with chemical plant off-gases that contains a higher concentration of inert gas (e.g. ammonia purge gas). In these cases, the industry more commonly uses cryogenic distillation to extract xenon and krypton, which requires a compression and cooling of the gas mixture at very low temperatures. The separation process can be broken down into three steps: first the condensation of all gas with a boiling point higher than the oxygen, then the purification of oxygen resulting in a 20-80 xenon/krypton mixture, and finally the separation of xenon from krypton. In 1997, several cases of explosion of separation units were caused by the reaction of non-filtered dangerous hydrocarbons with purified liquid oxygen produced in the second step of this long process.[distill_accident, distill_accident2] The extreme chemical and physical conditions required for cryogenic distillation support the need for less energy-intensive and safer alternatives.

The role of a dispatchable source of low-carbon energy can only be fulfilled by batteries charged by renewable energies (wind or solar) or by nuclear plants. However, one of the major criticism on this source of energy concerns the management of the radioactive waste. As promising technologies in gas separation emerge, there is an increasing need for a solution for the release of very small amount of radioactive off-gases like Kr₈₅ from nuclear spent fuels.[Blomeke_1969] Furthermore, stable xenon isotopes are also produced in these spent nuclear fuels, which can be used in all the above-mentioned applications. In the context of a regained interest in nuclear energy, the fourth generation nuclear plants are projected to be built on other technologies such as the light water or the molten salt technologies. [LeBlanc_2010] Molten salt reactors would continuously produce xenon and radioactive krypton in the electricity generation process.[Riley_2019] The development of gas separation units in these facilities would represent a promising source for xenon production. Yet, we can laboriously imagine deploying standard cryogenic distillation units in a nuclear facility for obvious security reasons. Consequently, nanoporous materials are considered as the alternative technology for xenon/krypton separation. Zeolites are already used as a pre-purification system, [kerry2007industrial] and they are now projected to be used as a standalone separation system.



*Figure 1.10: Representation of xenon/krypton separation process using porous materials in a nuclear fuel reprocessing plant and in a molten salt reactor. Reprinted with permission from reference **Banerjee_2014** copyright ©(2014) American Chemical Society and reference **Riley_2019** copyright ©(2019) Elsevier.*

Banerjee et al. proposed a two-bed system with a first bed filled with MOFs designed for xenon separation and then a second one for radioactive krypton capture.[**Banerjee_2014**] The authors proposed some examples of material that could be used for this separation unit; more research is needed to find out what are the best materials for these separations. In the following section, we will review the most promising materials for this separation and the structural explaining their high performance.

1.3.2 Promising materials for the xenon/krypton separation

Several experimental reports used the strategy outlined by computational screenings to improve separation properties, as well as tuning the chemical nature of the organic linkers. The main criteria outlined by the different studies on xenon/krypton separation call for pore size tailor-made for xenon and also for maximized interactions with the framework atoms obtained either through the chemical nature or the shape of the cavities.

In the early phase of the experimental design of materials for the xenon/krypton separation, Wang et al. synthesized a cobalt MOF $\text{Co}_3(\text{HCOO})_6$ with a selectivity of 12 that present rather narrow pores (around 5 Å) connected by zig-zag segments.[**Wang_2014**] Later, Chen et al. synthesized a selective porous cage material by not only focusing on the pore size but more importantly on the shape of the cavity, the selectivity of around 20 was considered record-high at that time. For instance, the cage windows are open for small noble gases such as krypton, whereas they close around the xenon hence maximizing the interaction.[**Chen_2014**] Mohamed et al. also designed a material with a similar selectivity, CROFOUR-1-Ni, however the performance was now explained by the chemical nature of the chromium oxide ligands that interact more strongly with the more polarizable xenon than the krypton molecules.[**Mohamed_2016**] Finally, Banerjee et al. tested a previously synthesized[**KAXQIL**] MOF after it was identified through high-throughput screening[**Simon_2015**] for its outstanding theoretical selectivity around 70. However, experimental measurements showed that its selectivity was not exceeding the one of the previous top materials. Similar emphases were made on the ideal pore size coupled with highly attractive framework atoms.[**Banerjee_2016**]

More recently, Li et al. proposed a rigid square-based MOF with “perfect pore size” (comparable with the kinetic diameter of Xe), and an internal pore surface decorated with very

polar hydroxyl groups. This material experimentally demonstrated record-high Xe/Kr selectivity of 60.6 at low pressure (0.2 bar) and ambient temperature.[**Li_2019**] Later, Pei et al. discovered even better performing materials with Xe/Kr selectivities of 74.1 and 103.4 in the same conditions 0.2 bar and 298 K. In addition to the perfectly tailored pore size, the structure features two oppositely adjacent open metal sites that strongly clamps the adsorbed xenon molecule.[**Pei_2022**] These studies clearly show the potential of polar sites that preferentially interact with the more polarizable xenon over the krypton, hence explaining these record-breaking separation performances.

1.3.3 From the computer to the test tube: a rare use case

To connect back our study to computational screenings, we are now going to present one of the rare case of direct contribution of high-throughput screenings to the lab testing of a material. In 2015, Simon et al.[**Simon_2015**] analyzed the Nanoporous Materials Genome, [**Simon_2015_EES**, **Boyd_2017**] a database of about 670,000 experimental and hypothetical porous material structures, including MOFs, zeolites, PPNs, ZIFs, and COFs, for candidate adsorbents for xenon/krypton separations. This study led to the re-discovery of SBMOF-1, a promising nanoporous material that was presented one year later.[**Banerjee_2016**]

It is possibly the largest-scale study performed in this area, both by the sheer number of frameworks involved and by the diversity of their nature. Because such a set is too big for brute-force screening with GCMC simulations, they proposed a multiscale modelling strategy combining machine learning algorithms (trained on a diverse subset of 15,000 materials) with molecular simulations (used both to generate the ML training data, and to refine the separation properties for the top performers obtained by the ML predictor). Without going into details (see Fig. 1.11 for more details), the ML model they trained was mainly based on geometric structural descriptors, with the addition of a single energy-based descriptor: the Voronoi energy (i.e. the average energy of a xenon atom at the accessible nodes in the Voronoi partition of space). In addition to identifying and describing some top performing materials, the authors also analyzed the correlations between high Xe/Kr selectivity and the geometric properties of the frameworks, in order to “rationalize the strong link between pore size and selectivity”. In particular, by developing theoretical pore models of spherical and cylindrical geometries, they could highlight the general geometrical trends observed, but also the fact that there is a wide diversity of performance beyond the geometrical features of the frameworks, which suggests the key role of the chemical nature of the cavities.

The chemical nature of the cavities were best described using the Voronoi energy descriptor they developed. This descriptor give an idea of the xenon adsorption isosteric heat of the material. Given these results, more studies should focus on describing the adsorption thermodynamic quantities such as the asorption enthalpy but also the henry constants. This study finally lead to the synthesis and testing of one of the top performing material in the field. However, we cannot stop but wondering why there is a discrepancy between the theoretical selectivity of around 70 of SBMOF-1 and its actual experimental selectivity of 16. In the final perpective chapter of this thesis we will try to give an explanation for this. In the future, such close collaboration between experimental and computational teams are crucial even if they are still too rare. A recent paper suggests that these collaborations are rare accross all nanoporous fields and a lot of improvements are needed to foster cooperations betwwen the labs.[**Li_2022**]

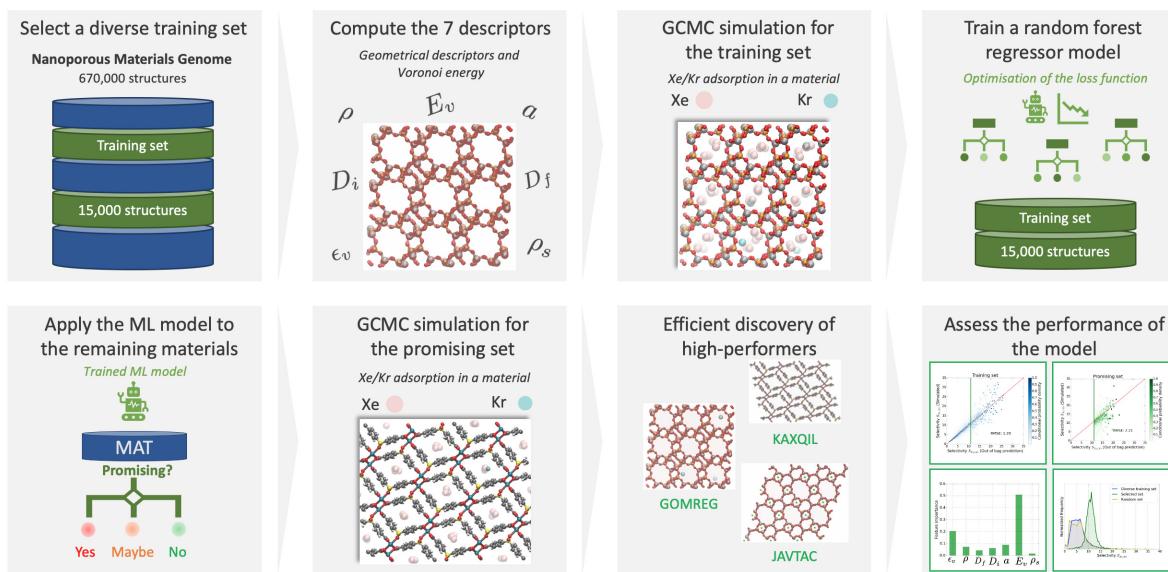


Figure 1.11: Schematic representation of large-scale screening of nanoporous materials for Xe/Kr adsorption-based separation by Simon et al.,[Simon_2015] based on a combination of Grand Canonical Monte Carlo simulations and machine learning algorithm (Random Forest Regressor). The main goal of this screening is to find high-performing materials in a large dataset of both experimental and hypothetical materials. Adapted with permission from Ref. Simon_2015. Copyright 2015 American Chemical Society.

1.3.4 The future of screening

Despite the progress made, important drawbacks of the current methodologies remain. High-throughput screenings rely too much on oversimplified assumptions such as the rigidity of the framework, the absence of defects, the use of Lennard-Jones potentials and inaccurate charges. For instance, the rigidity of the framework only takes into account one conformation of the framework. Yet, thermal agitation induces a “breathing” movement of the framework with an amplitude dependent on its intrinsic flexibility. The pores of the framework can change depending on the number of adsorbates to interact more optimally with them, which can be induced by a change in pressure. The issue of flexibility is rarely tackled, and when considered, it is only on the few most selective structures given by an inaccurate screening based on the rigid crystal approximation. One can wonder about the results obtained if it is applied to larger sets of structures. Witman et al. found that flexibility applied to top performing materials can decrease the selectivity, because the pore does not have an optimal size anymore.[Witman_2017] In some cases, the selectivity of a well performing material can even increase to become a top performing one. Computational screenings can be closer to predict experimental values of selectivity, diffusivity, and other key performance metrics.

Many open problems remain for the design of efficient high-throughput computational screenings. The connection between different properties for a given application is not systematically integrated in the screening procedures. For example, in methane storage, the working capacity of the material is the main property to optimize, but the kinetics of the adsorption/desorption or the mechanical resistance to compaction among others also need to be considered. Designing a nanoporous material is in fact a multivariate optimization problem with tacit constraints,

for example the synthesizability. For instance, the diffusion coefficients of adsorbates in a xenon/krypton separation problem can help us better understand the breakthrough simulations and eventually the whole separation process in pressure or temperature swing adsorption beds. For this reason, studying transport properties along with uptake capacities and thermodynamic selectivity of the xenon/krypton separation can give a more complete picture of the industrial process we ultimately want to model.

Moreover, the transferability of the methodology to a broad range of materials is often achieved at the expense of accuracy in specific cases. And one can rightly question the universality of depending on faster but less elaborated models, which boils down to a trade-off problem between prediction accuracy and computational cost (or complexity). For instance, classical force-fields are broadly used in rigid materials for adsorption properties, but the switch to more costly *ab initio* methods or the addition of flexibility can result in a more accurate description at the expense of computational resources. The addition of polarization could be very promising since several top performing materials harbor open metal sites and highly polar sites that explain the acute affinity to xenon adsorbates.

The development of ML-assisted screenings is paired with the advances in data science techniques and algorithms. Recent advances in deep learning have enabled the development of transformer-based (the technology at the base of ChatGPT for example) machine learning models to predict adsorption properties.[**Kang_2023, Cao_2023**] More importantly, the construction of descriptors tailored to the many possible applications is also an on going work. This construction work cannot be dissociated to the physical and chemical intuition of the scientists. Topological, chemical, electronic and other descriptors have been developed on top of the more common geometrical and thermodynamic descriptors, which displays the importance of strong physical chemistry knowledge. Recently Shi et al. highlighted the key role of energy histograms in predicting adsorption properties.[**Shi_2023**] The discovery of novel relevant descriptors remains the main lever for increased performance of the ML models and is closely related to a rigorous theoretical work. For these reasons, we worked during this thesis on more accurate and faster ways of calculating these interaction energies to extract valuable energy/thermodynamic descriptors.

The development of databases is another key aspect in the promotion of data science in the field of materials science in general, and nanoporous materials chemistry in particular. The diversity of materials, the inclusion of experimental data (successful or failed), the addition of under studied classes of materials (e.g. amorphous) are all key aspects to upgrade the existing database. Even if existing attempts to create a centralized database have been initiated by the materials project,[**MaterialsProject**] this database does not contain all the existing information on each material. Furthermore, this high amount of data will need to be efficiently explored, and non supervise deep learning algorithms have been developed to do so.[**Park_2023**] Coupled with synthesis robot, these method can navigate through the unexplored databases to find the few most interesting candidates for a given targeted application.

In the future, computational high-throughput screening could be integrated more tightly into the design process of nanoporous materials, hence further improving its efficiency. The computational pre-screening can be coupled with automated screenings of the most promising materials to finally identify candidates for further studies. This automated design process is

described by Lyu et al. in their paper on “Digital Reticular Chemistry” and set out promising perspectives for computational screenings in the field.[**Lyu_2020**] Some studies are already pioneering this new research area by combining high-throughput characterizations, active learning algorithms and robotic synthesis.[**Greenaway_2018, Moosavi_2019**] Another step towards faster industrialization would integrate process modelling to enrich the purely atomistic approach.

2

THERMODYNAMIC EXPLORATION OF XENON/KRYPTON SEPARATION

2.1	Calculation of thermodynamic descriptors	33
2.1.1	Thermodynamics of adsorption separation	33
2.1.2	Grand canonical Monte Carlo	34
2.1.3	Widom insertion	34
2.2	Preliminary analyses	34
2.2.1	Structure-selectivity relationships	35
2.2.2	Thermodynamic quantities correlations	35
2.3	Selectivity drop	39
2.3.1	Thermodynamic origins	39
2.3.2	Detailed investigation	43
2.3.3	Conclusions and perspectives	46



2.1 CALCULATION OF THERMODYNAMIC DESCRIPTORS

dqsd

2.1.1 Thermodynamics of adsorption separation

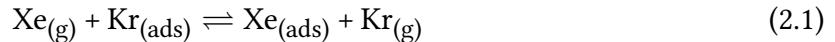
[selectivity, entropy enthalpy, Henry coefficients]

From microcanonical to Gibbs ensemble

[Exchange equilibrium]

In order to analyze thoroughly the thermodynamics of competitive xenon and krypton adsorption using the results of both pure-component and multi-component GCMC simulations, we

use in this work some thermodynamic properties (enthalpy and entropy in various conditions) associated with the following fictitious “exchange equilibrium”:



The equilibrium constant associated to the Equation (2.1) at a pressure of 1 atm for a given composition is simply the selectivity s_1 , defined above as a function of y^{Xe} , y^{Kr} , q^{Xe} and q^{Kr} , respectively the mole fractions in the gas phase and the loadings in the material of xenon and krypton in the GCMC calculation. The enthalpy difference associated with this “exchange” equilibrium, $\Delta_{\text{exc}}H_1 = \Delta_{\text{ads}}H_1^{\text{Xe}} - \Delta_{\text{ads}}H_1^{\text{Kr}}$, is defined as the difference of enthalpy between xenon and krypton within the mixture. Similarly, an entropy difference $\Delta_{\text{exc}}S_1$ can then be calculated as in Equation 2.2 (with $i = 1$):

$$\Delta_{\text{exc}}S_i = R \ln(s_i) + \frac{1}{T} \Delta_{\text{exc}}H_i \quad \text{where } i \in \{0, 1\} \quad (2.2)$$

Furthermore, this “exchange equilibrium” can be generalized to the zero loading limit. A comparison of the thermodynamic quantities between the low pressure and ambient pressure cases is made possible by considering this equilibrium equation. The equilibrium constant at low pressure similarly corresponds to s_0 , the enthalpy to $\Delta_{\text{exc}}H_0 = \Delta_{\text{ads}}H_0^{\text{Xe}} - \Delta_{\text{ads}}H_0^{\text{Kr}}$ and the entropy to $\Delta_{\text{exc}}S_0$ as written in Equation 2.2 (with $i = 0$).

2.1.2 Grand canonical Monte Carlo

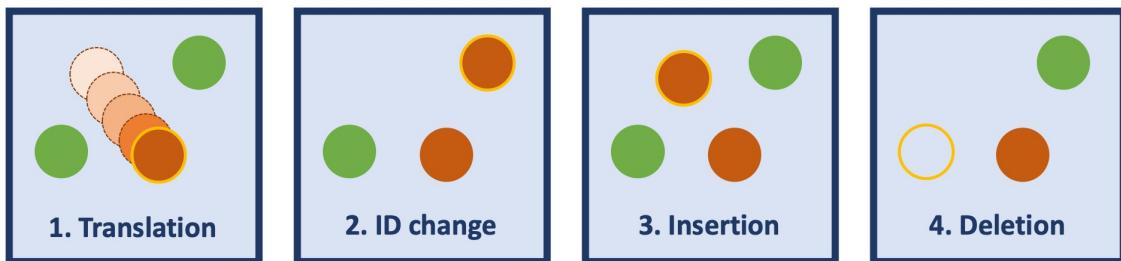


Figure 2.1: MC moves for monoatomic atoms with multiple components

2.1.3 Widom insertion

2.2 PRELIMINARY ANALYSES

As we have seen above in the existing literature, the computational screening of the nanoporous materials – both existing frameworks and hypothetical structures – for targeted adsorption properties has been the object of many studies, and several of those high-throughput screening studies have focused on noble gas separation, and Xe/Kr separation, in particular. For large-scale studies we have found that, in addition to the testing and validation of methodological developments, the screening aimed in most cases at one of three objectives: (i) to identify top performing materials for synthesis and/or characterization; (ii) to better understand the limits of possible performance, and the relationships and trade-offs between various metrics of performance (selectivity, uptake, etc); (iii) identify structure–property relationships, correlating separation performance with structural properties of the materials that can be more easily determined (i.e., at low computational cost).

2.2.1 Structure–selectivity relationships

[How SA VF and LCD are calculated]

[Put analyses of selectivity relationship with other descriptors: uptake, void fraction, surface area, etc.]

[Trade-offs etc.]

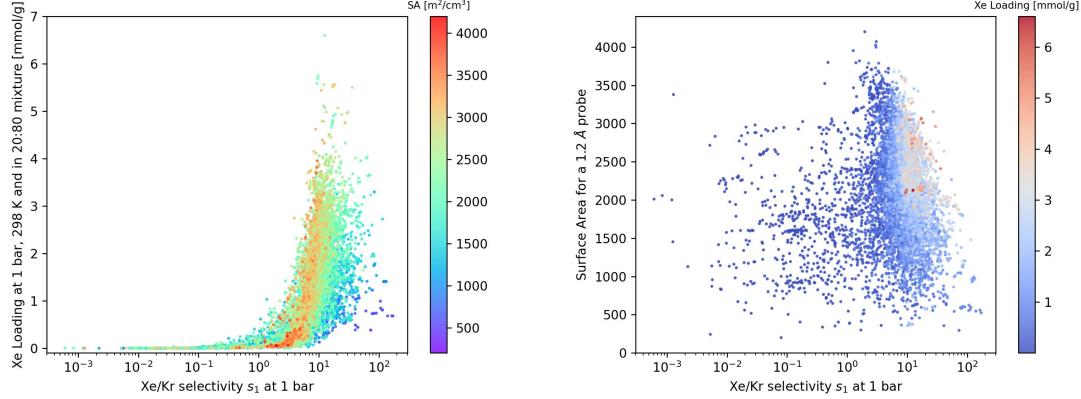


Figure 2.2: Selectivity / Xe Uptake / Surface area

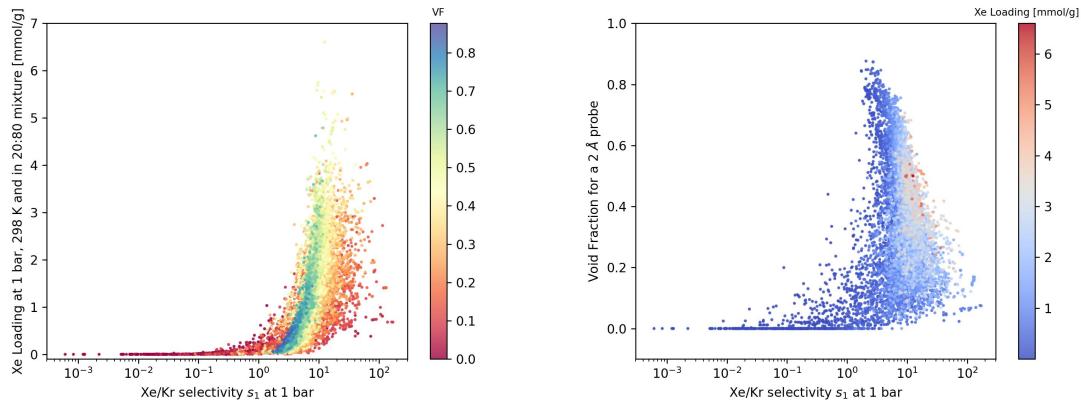


Figure 2.3: Selectivity / Xe Uptake / Void Fraction

2.2.2 Thermodynamic quantities correlations

In the present work, our goal is not directly to address the structure–property relationships, but rather to map out the details of the thermodynamic features of Xe/Kr adsorption and separation in nanoporous materials. We used the high-throughput screening methodology as a way to map out the space of thermodynamic properties, going beyond the usual quantities of selectivity and uptake, to focus more specifically on the role of adsorption enthalpy and entropy, the differences between Xe and Kr adsorption thermodynamics, and the differences between selectivity at low and high pressure.

To evaluate the performance of a given nanoporous material for separation in the low loading (or low pressure) limit, Henry’s constants are often calculated from linear fits of low-pressure adsorption isotherm data — both experimentally and computationally. In this section, we investigate the thermodynamics of Xe and Kr adsorption at low pressure. Here, We have

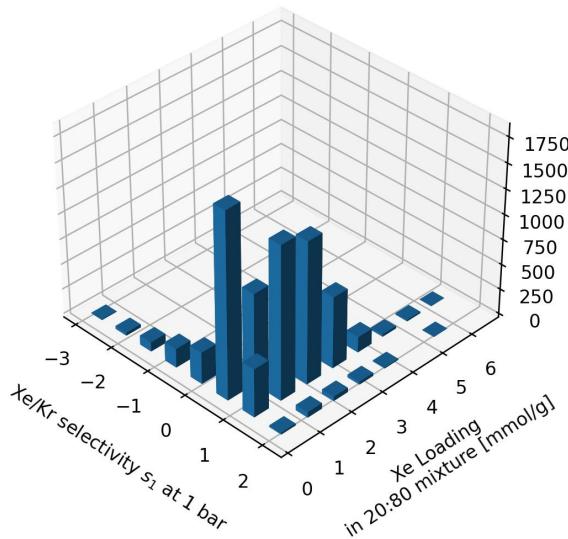


Figure 2.4: Distribution selec / uptake

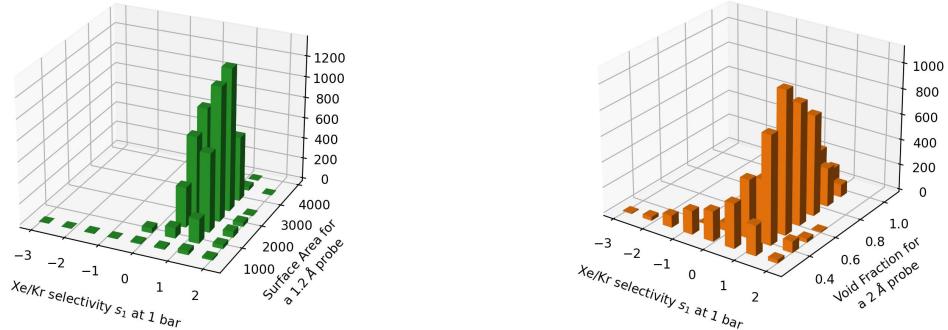


Figure 2.5: Distributions, selec vs sa vol

calculated the low-pressure adsorption properties by using the Widom insertion method [**Widom1963, frenkel2001widom**] on 9,668 structures from the dataset selected. It has higher accuracy than the fitting of isotherms, where it can be difficult to know what the extent of the linear adsorption regime is. With these simulations, we could obtain for each material the Henry's constant K and the adsorption enthalpy $\Delta_{\text{ads}}H_0$ (at the zero loading limit) for both xenon and krypton. The Xe/Kr thermodynamic selectivity s_0 in the low-pressure limit is then determined by the ratio $s_0 = K^{\text{Xe}}/K^{\text{Kr}}$ of the Henry's constants for the two gases. In the following, we look at the statistical relationships between the thermodynamic quantities at low pressure: s_0 , K^{Xe} , K^{Kr} , $\Delta_{\text{ads}}H_0^{\text{Xe}}$, $\Delta_{\text{ads}}H_0^{\text{Kr}}$ and $\Delta_{\text{exc}}H_0$ (which is defined in Section 2.1.1).

We display the distribution of thermodynamic properties of materials with favorable thermodynamic Xe/Kr selectivity ($s_0 > 1$) in Figure 2.6 — we restrict these plots to selectivity above 1, because those are the materials of interest for separation, and doing so removes several outliers with specific geometries or binding sites (but does not change the overall conclusions). We can first see that although the logarithm of the Xe Henry's constant K^{Xe} is weakly correlated to

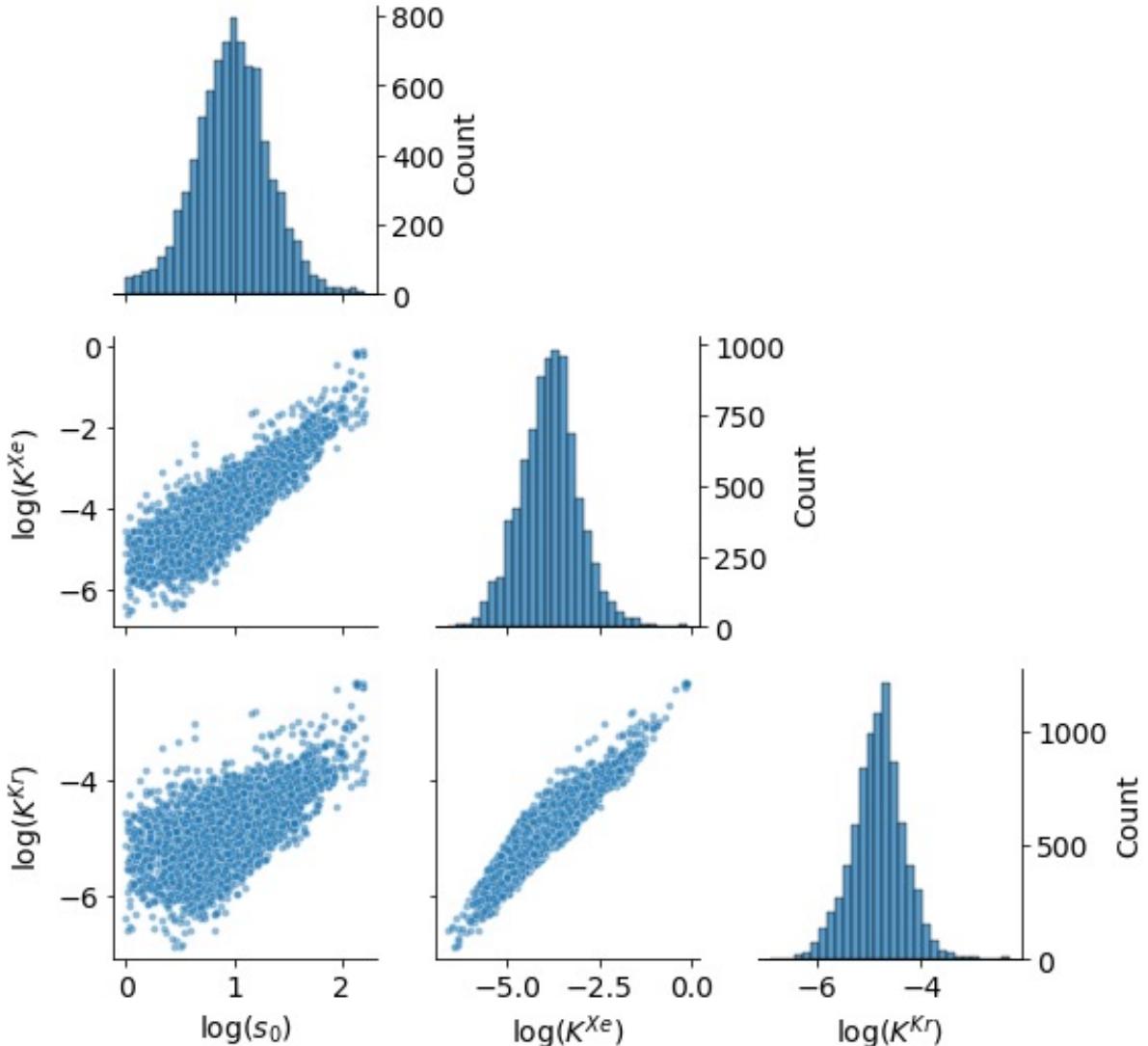


Figure 2.6: For 8,401 MOFs with favorable thermodynamic Xe/Kr selectivity ($s_0 > 1$), pair-plots of $\log_{10}(s_0)$, $\log_{10}(K^{Xe})$ and $\log_{10}(K^{Kr})$ (the Henry's constants are in $\text{mmol g}^{-1} \text{Pa}^{-1}$) in the off-diagonal subplots (note that the y-axis is displayed on the right side) and the distribution of each quantity are on the diagonal (note that the y-axis displayed on the right side corresponds to the count and the x-axis is correctly labeled below each subplot).

the logarithm of the selectivity s_0 , this correlation is stronger for highly selective materials. Therefore, in a multistep screening study to identify the most selective materials, it could be possible to use as a “first filter” criterion based purely on Xe adsorption, discarding materials below a certain threshold (e.g., the materials with $s_0 \geq 30$ are contained in the subset with $K^{Xe} \geq 2.7 \cdot 10^{-1} \text{ mmol g}^{-1} \text{ Pa}^{-1}$). The correlation between K^{Kr} and s_0 , on the other hand, is weaker.

With regards to Henry's constants, we see a large range of behaviour, with K^{Xe} ranging from $2.6 \cdot 10^{-7} \text{ mmol g}^{-1} \text{ Pa}^{-1}$ to $7.9 \cdot 10^{-1} \text{ mmol g}^{-1} \text{ Pa}^{-1}$, and K^{Kr} ranging from $1.3 \cdot 10^{-7} \text{ mmol g}^{-1} \text{ Pa}^{-1}$ to $5.1 \cdot 10^{-3} \text{ mmol g}^{-1} \text{ Pa}^{-1}$. We also see that statistically, a high affinity for xenon usually translates into a high (relative) affinity for krypton, which is a general trend for noble gases where the adsorption sites are not strongly specific. In order to look more in detail into the

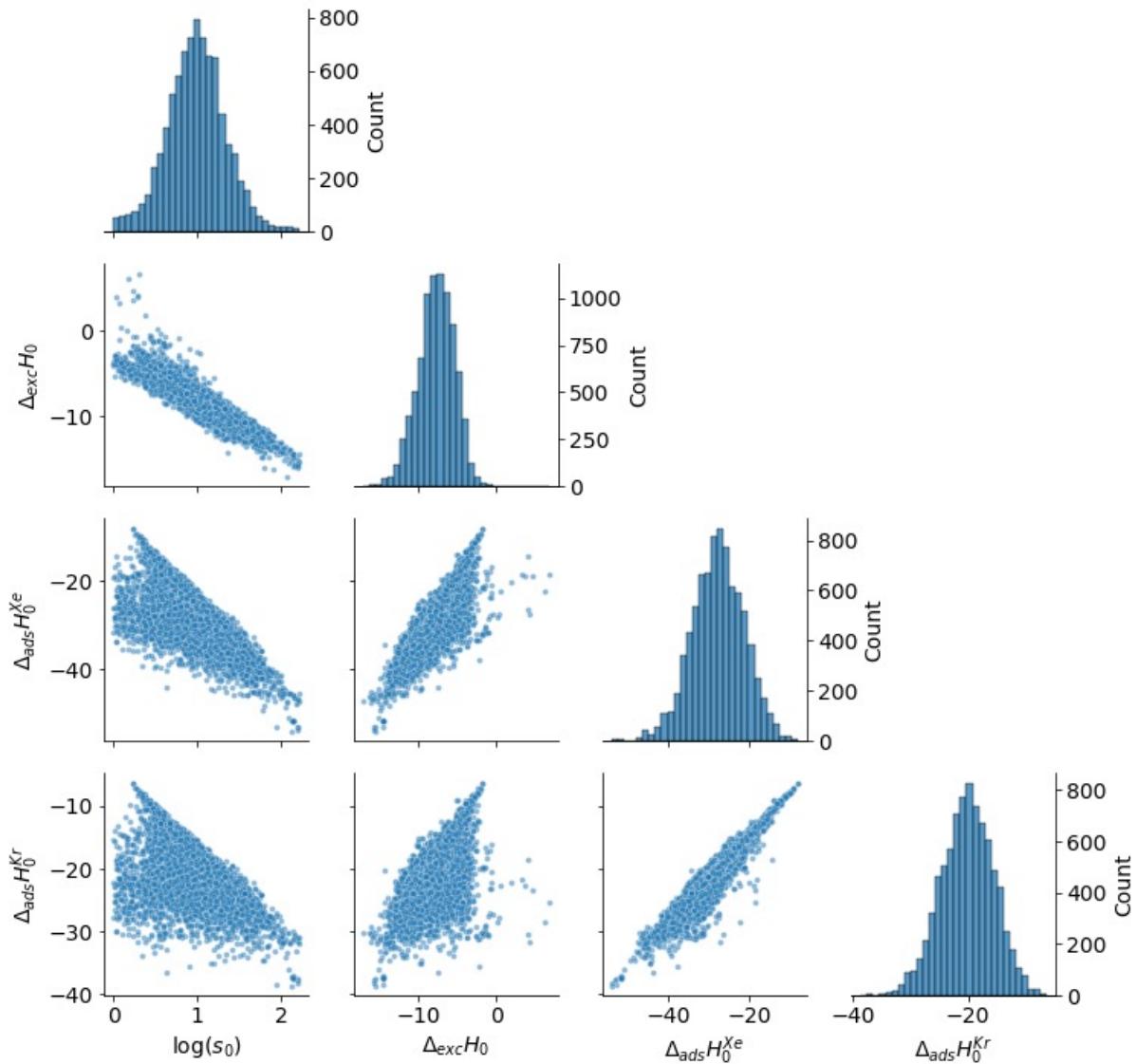


Figure 2.7: For 8,401 MOFs with favorable thermodynamic Xe/Kr selectivity ($s_0 > 1$), pair-plots of $\log(s_0)$, $\Delta_{\text{exc}}H_0$, $\Delta_{\text{ads}}H_0^{\text{Xe}}$ and $\Delta_{\text{ads}}H_0^{\text{Kr}}$ (the enthalpies are in kJ mol^{-1}) in the off-diagonal subplots and the distribution of each quantity are on the diagonal.

thermodynamics behind this large diversity in behaviour, we plot in Figure 2.7 the enthalpies involved.

We first observe that the low-loading adsorption enthalpy of xenon ($\Delta_{\text{ads}}H_0^{\text{Xe}}$) is strongly correlated to that of krypton ($\Delta_{\text{ads}}H_0^{\text{Kr}}$). Echoing the similar correlation seen between respective Henry's constants, it suggests a rather generic physisorption mechanism is at play in the majority of materials, and that host–adsorbate affinities are mainly determined by the enthalpy. The main driver of Xe/Kr selectivity is neither the xenon or krypton adsorption enthalpy alone (both are weakly correlated to the selectivity), but as expected their difference, $\Delta_{\text{exc}}H_0$, which is strongly correlated to $\log(s_0)$. This is further confirmed by the lack of correlation between selectivity and adsorption entropies (c.f. supplementary information, Figure S2): the separation is mostly enthalpic in nature, and the entropy causes the dispersion in the correlation between selectivity $\log(s_0)$ and $\Delta_{\text{exc}}H_0$.

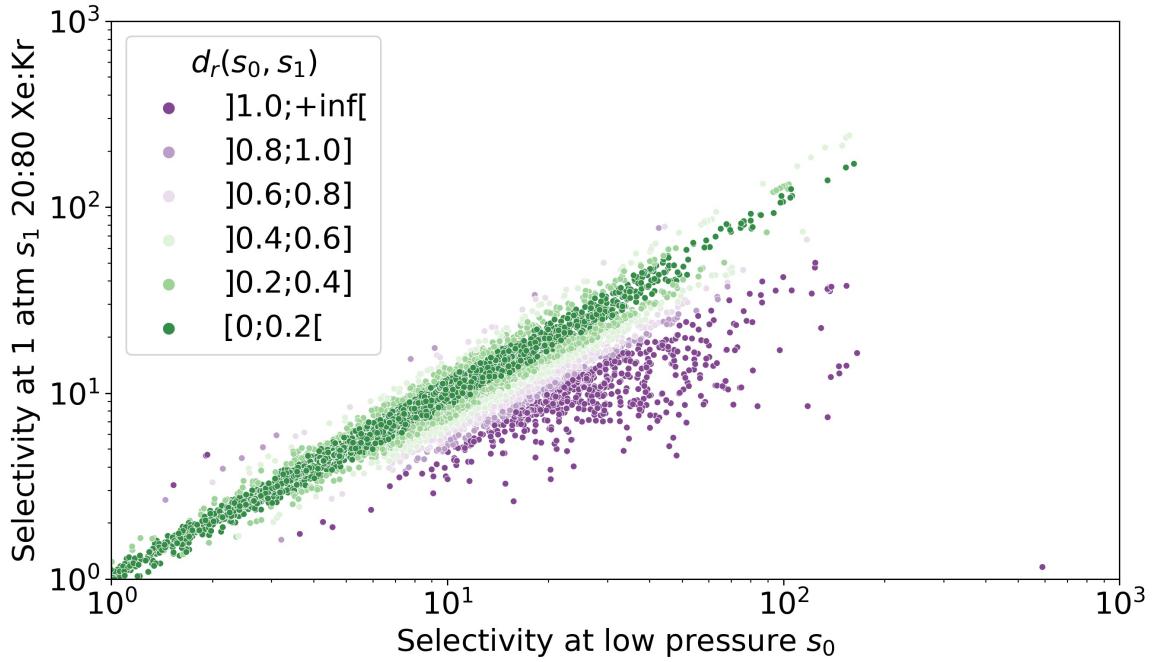


Figure 2.8: Difference of selectivity between low pressure and at a 1013 hPa pressure for a 20:80 xenon krypton composition. The relative difference between the low-pressure selectivity and the ambient-pressure is particularly high for the points labeled in purple.

2.3 SELECTIVITY DROP

2.3.1 Thermodynamic origins

In this section we focus on the impact of a change of working pressure on the adsorption selectivity, and analyze its thermodynamic origins. This is key to accurately assess the thermodynamics of adsorption in different working conditions for specific industrial processes, and any insight into the impact of pressure on selectivity may allow for faster screening limited at selected thermodynamic conditions.

We calculated the selectivity s_1 at pressure 1 atm and ambient temperature using GCMC calculations on the entire dataset, with Xe/Kr mixture composition of 20:80 (found in a byproduct stream from air separation [**kerry2007industrial**]) and 90:10 (found in the off-gas streams from nuclear waste [**auerbach2003handbook**]). For high-selectivity materials, we find that the impact of composition appears rather marginal (c.f. supplementary information, Figure S5). In the following, we discuss the selectivity for the 20:80 mixture, which is the most commonly studied one in the literature. To measure the difference in selectivity between low and ambient pressures, we consider a relative difference $d_r(s_0, s_1)$ defined as follows:

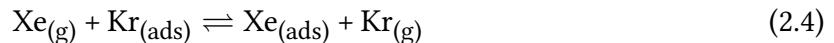
$$d_r(s_0, s_1) = \frac{|s_0 - s_1|}{\min(s_0, s_1)} \quad (2.3)$$

In Figure 2.8, the selectivity at ambient pressure s_1 is plotted against its low-pressure counterpart s_0 (for materials where $s_0 > 1$, as before). The points are color-coded according to the value of

$d_r(s_0, s_1)$, in 6 discrete categories for the sake of clarity. There is some broad level of correlation, see near the diagonal with 61.5% of materials where the difference is below 20% (near the $s_0 = s_1$ line). We also see clearly that there are many more points (74.3% among the materials with $d_r(s_0, s_1) \geq 0.2$) below the first bisector ($s_1 < s_0$) than above: for these materials the selectivity s_1 at 1 atm is significantly lower than the one at low pressure s_0 .

This drop in selectivity mainly concerns the materials with a relatively high selectivity $s_0 > 10$ (see Figure 2.8), and forewarns that considering solely pure-component Henry's constant (i.e. zero-pressure selectivity) for materials screening could be misleading in some cases. Although it is simpler and faster to calculate, those low-pressure results that can overestimate selectivity by more than 100% in a significant number of materials (646 out of 9,668 in our dataset). By using a thermodynamic approach, we now try to explain the reasons behind these shifts in selectivity.

To evaluate quantitatively the thermodynamic effects at play in the competitive adsorption in different regimes, we consider thermodynamic properties of the following "exchange equilibrium":



as described in Section 2.1.1. We plot in Figure 2.9 the exchange entropy at low pressure (plotted as $T\Delta_{\text{exc}}S_0$) against the exchange enthalpy $\Delta_{\text{exc}}H_0$. In this scatter plot, the points are color-coded according to the selectivity s_0 (with discrete categories for the sake of clarity), which is related to the enthalpy and entropy through Equation 2.2 – meaning iso-selectivity lines are parallel straight lines in this scatter plot.

In the supplementary information Figure S8, we display the distributions of the exchange enthalpy and entropy at low pressure. For the 630 most selective materials ($s_0 > 30$), the distribution of the exchange enthalpy $\Delta_{\text{exc}}H_0$ is centered on $-12.0 \text{ kJ mol}^{-1}$ with a standard deviation of 1.3 kJ mol^{-1} , whereas the distribution of the exchange entropy (plotted as $T\Delta_{\text{exc}}S_0$) is centered on -2.5 kJ mol^{-1} with a standard deviation of 0.7 kJ mol^{-1} . These figures, along with the overall distribution plotted in Figure 2.9, further confirms the moderate role of entropy in the low-pressure selectivity: it is equivalent in average to about 20% of the exchange enthalpy at low pressure.

Figure 2.10 represents a scatter plot of the exchange entropy at $P = 1 \text{ atm}$ $\Delta_{\text{exc}}S_1$ against the exchange enthalpy at ambient pressure $\Delta_{\text{exc}}H_1$. To compare it to the Fig. 2.9, the points are color-coded according to the low-pressure selectivity s_0 . Compared to the iso-selectivity s_1 straight parallel lines (c.f. supplementary information Figure S7), we can see that many materials with high s_0 have lower s_1 – seen as a migration of points to the right of the plot, compared to Fig. 2.9. This shift is therefore mainly due to a higher (less favorable) exchange enthalpy, hinting at an important role of enthalpy to determine higher pressure selectivity.

To quantify this change, we consider the distributions of the exchange enthalpy $\Delta_{\text{exc}}H_1$ and the energetic equivalent of the exchange entropy $T\Delta_{\text{exc}}S_1$ at ambient pressure (Figures S9). The enthalpy $\Delta_{\text{exc}}H_1$ is now centred on $-11.1 \text{ kJ mol}^{-1}$ with a standard deviation of 1.9 kJ mol^{-1} . Compared to the zero-pressure values, the enthalpy distribution is more dispersed, showing that there are important changes in individual values, and is higher in average – majority of materials have lower ambient pressure selectivity due to enthalpic effects. This can be explained by the very general increase of adsorption enthalpy upon loading in the gas phase,

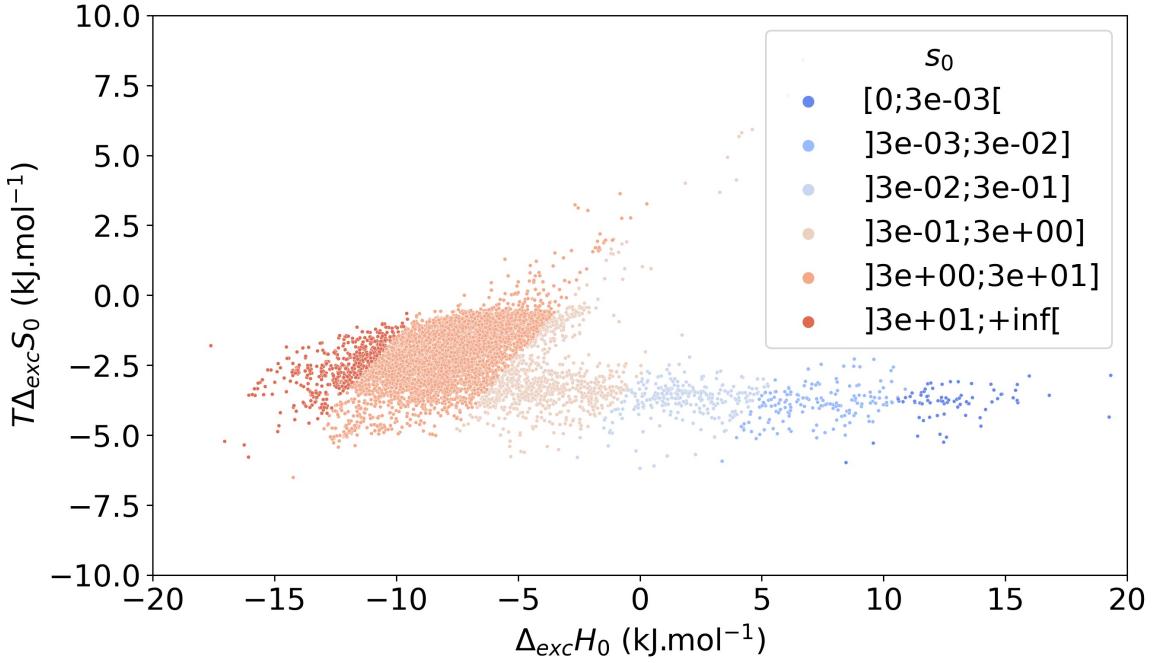


Figure 2.9: The energetic equivalent of exchange entropy $T\Delta_{\text{exc}}S_0$ and enthalpy $\Delta_{\text{exc}}H_0$ at low pressure labeled using the selectivity s_0 at low pressure. The limits between labels follows a affine function of slope $1/T$ and of intercept $-R \ln(s_0^{\lim})$ where s_0^{\lim} is the limit selectivity value (cf. Equation (2.2)). In other words, the iso-selectivity lines are all parallel lines of equation $y = f(x)$ where f is the affine function described previously.

which is linked to the presence of more adsorbed molecules. In fact, the correlations (Figure 2.6) suggest that highly selective materials have high affinity in xenon, therefore they feature significant uptake at 1 atm and the large Xe loading means the most favorable adsorption sites can be saturated, and further adsorption involves weaker host–guest interactions and therefore increases the average adsorption enthalpy at nonzero loading.

The entropic term $T\Delta_{\text{exc}}S_1$ is now centered on -2.9 kJ mol^{-1} , with a standard deviation of 0.8 kJ mol^{-1} (almost unchanged from low-pressure). The entropy is on average lower, which means an overall less favorable separation due to entropic effects: this evolution of the entropic term hints at the potential of reorganization of the adsorbed molecules inside each material. The difference in distribution of enthalpy has, overall, more impact on the high-pressure selectivity than that of entropy. This suggests that the overall contribution of enthalpy remains more decisive than the role of entropy in the selectivity change, even at ambient pressure. This is an interesting conclusion for screening studies, because evaluation of adsorption enthalpy can be computationally faster than that of the adsorption free energy (or entropy).

To further investigate the thermodynamics of the selectivity change, we quantify in this section the contributions of enthalpy and entropy. The ratio s_1/s_0 is equal to the product $k_H \times k_S$

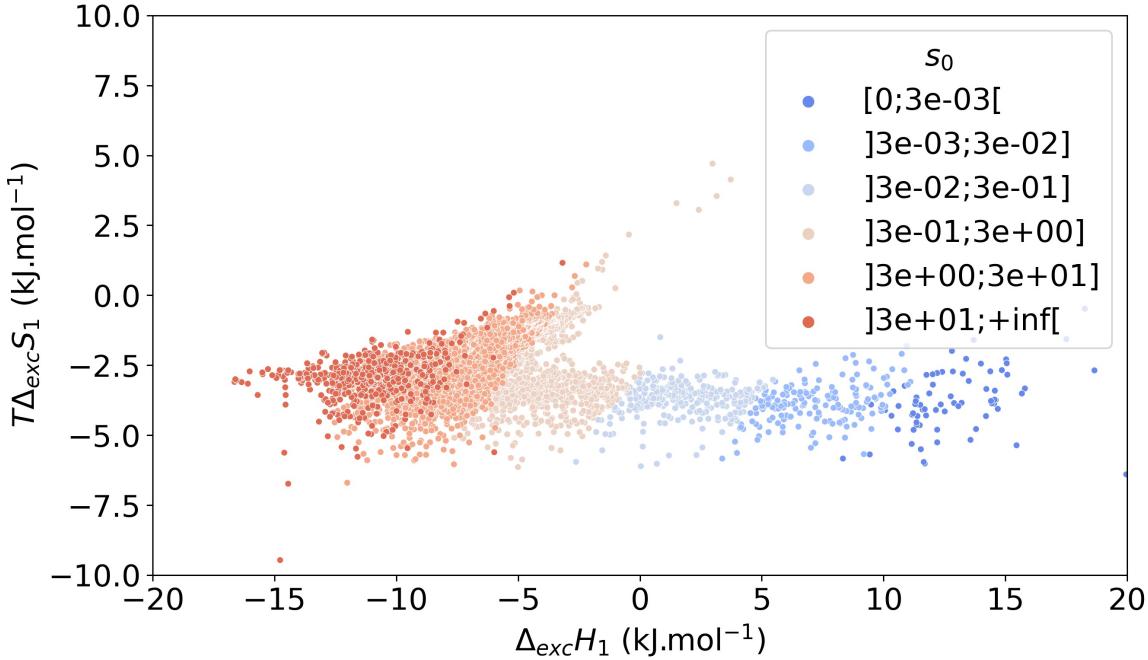


Figure 2.10: The energetic equivalent of exchange entropy $T\Delta_{\text{exc}}S_1$ and enthalpy $\Delta_{\text{exc}}H_1$ at ambient pressure labeled using the selectivity s_0 at low pressure. The points are layered so that the points with higher s_0 are always above. To see a split version of this plot, please refer to the Figure S6 in the supplementary information.

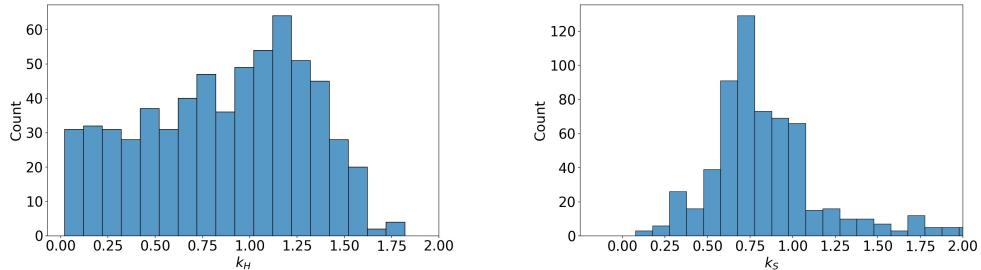


Figure 2.11: Distribution of the enthalpic k_H and entropic k_S contributions to the change of selectivity from low to ambient pressure for the 630 materials with $s_0 > 30$. k_H has a rather uniform distribution, whereas k_S has a bell-like distribution.

where k_H and k_S are the enthalpic and entropic contributions to the selectivity change defined as:

$$\begin{aligned} k_H &= \exp \left(-\frac{\Delta_{\text{exc}}H_1 - \Delta_{\text{exc}}H_0}{RT} \right) \\ k_S &= \exp \left(\frac{\Delta_{\text{exc}}S_1 - \Delta_{\text{exc}}S_0}{R} \right) \end{aligned} \quad (2.5)$$

As we can see in Figure 2.11, the entropic contribution k_S has a bell-like distribution, with a mean of 0.9 and a standard deviation of 0.6. This confirms that k_S is close to 1, and has therefore only a marginal effect on the selectivity change. On the other hand the enthalpic contribution k_H has a more uniform distribution ranging from 0.1 to 1.5, which means that

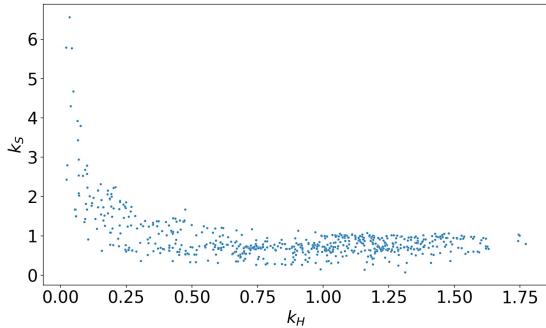


Figure 2.12: Scatter plot of the enthalpic contribution k_H and entropic contribution k_S for the 630 materials with $s_0 > 30$. The entropic compensation occurs when the enthalpic contribution is around 0.1, else its value is around 1 and has little effect on the selectivity change.

enthalpy has a crucial role in the selectivity change we observe. There are a significant number of materials with a k_H close to zero, they correspond to the same materials highlighted in Section 2.3.1.

Furthermore, the scatter plot of k_H and k_S (shown in Figure 2.12) confirms a rather moderate effect of entropy. For most of the materials with $0.25 \leq k_H \leq 1.75$, we see that k_S is close to 1. The most significant entropic contributions are found for materials where k_H is close to zero (typically below 0.25). If we look in more detail at the 29 materials with $k_S > 2$, the entropic contribution k_S moderately compensate the enthalpic contribution as the average ratio s_1/s_0 is around 0.25. In such cases, the entropy is non-negligible and it can partially compensate the enthalpic contribution to the selectivity change, but the general trend is still given by enthalpy, since the overall selectivity is decreasing as a result.

2.3.2 Detailed investigation

In this section, we go over some of the most selective materials, as identified at low pressure and listed in Table 2.1, and we provide a detailed investigation of the thermodynamic effects behind their behavior. We can split them into three main categories: materials with a slight increase in selectivity or little change in selectivity ($s_0/s_1 > 0.8$), materials with a slight decrease in selectivity ($0.5 \leq s_0/s_1 \leq 0.8$) and materials with a significant decrease in selectivity ($s_0/s_1 < 0.5$). In this section, we investigate the origins of these different behaviours: all materials are referenced by their CSD refcode.

We first study a few examples of the category of materials where ambient-pressure selectivity is close to (or even higher than) the low-pressure value. For VOKJIQ, the selectivity is multiplied by 1.5 between low and ambient pressure. We see that the adsorption enthalpy of xenon $\Delta_{\text{ads}}H^{\text{Xe}}$ decreases from $-53.9 \text{ kJ mol}^{-1}$ to $-61.1 \text{ kJ mol}^{-1}$, whereas for krypton $\Delta_{\text{ads}}H^{\text{Kr}}$ decreases from $-38.2 \text{ kJ mol}^{-1}$ to $-44.5 \text{ kJ mol}^{-1}$ (c.f. supplementary information Table S1). This increased stability of the adsorption sites upon loading is not common in nanoporous materials for rare gas adsorption, and can be linked to a cooperative effect between the adsorbed molecules. The stabilization favors the xenon molecules over the krypton molecules, due to an interatomic distance inside the pores that is a closer match to the energy well for favorable Lennard-Jones potential for xenon-xenon interactions than for krypton-krypton interactions (which is the case for a distance higher than 4.2 Å; see Figure S10).

Table 2.1: Enthalpic and entropic contributions to the selectivity change for some archetypal structures selected for their high s_0 selectivity. Every structure is identified using a CSD Refcode and the first article that mentions it

CSD Refcode	Ref.	s_0	s_1	s_1/s_0	k_H	k_S
VOKJIQ	VOKJIQ	157.17	242.73	1.54	1.46	1.06
KAXQIL	KAXQIL	103.78	132.57	1.28	1.32	0.96
JUFBIX	JUFBIX	106.11	114.83	1.08	1.08	1.00
FALQOA	FALQOA	162.20	171.10	1.05	1.09	0.96
GOMREG	GOMREG_GOMRAC	114.14	73.83	0.65	1.01	0.64
JAVTAC	JAVTAC	117.38	66.93	0.57	0.77	0.74
GOMRAC	GOMREG_GOMRAC	124.11	47.34	0.38	0.58	0.66
MISQIQ	MISQIQ	138.94	37.32	0.27	0.51	0.53
BAEDTA01	BAEDTA01	154.10	37.74	0.24	0.12	1.97
VIWMOF	VIWMOF	81.13	13.24	0.16	0.04	4.30
LUDLAZ	LUDLAZ	165.68	16.42	0.10	0.16	0.63
WOJJOV	WOJJOV	146.32	13.94	0.10	0.06	1.68
VAPBIZ	VAPBIZ	146.73	12.76	0.09	0.06	1.50

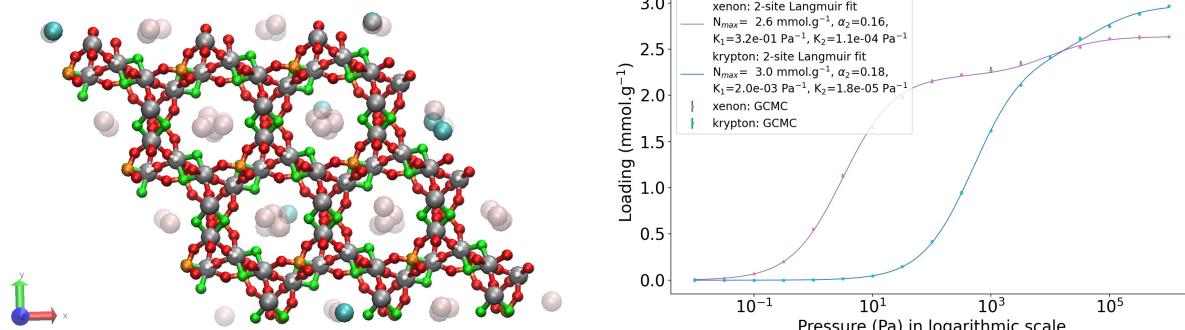


Figure 2.13: Representation of a chiral open-framework fluoroaluminophosphate $[C_4N_3H_{16}]\cdot[Al_6P_3O_{12}F_6(OH)_6]$ denoted AlPO-JU89 (referenced MISQIQ in the Cambridge structural database), which has been loaded with xenon and krypton in a GCMC simulation, on the left side. [MISQIQ] Color code: Al in silver, P in orange, O in red, H in white and F in green for the framework; and Xe in transparent pink and Kr in cyan for the adsorbates. The pure-component isotherms fitted with a 1-site Langmuir model for both xenon and krypton at 298 K on the right side (for more information on the Langmuir parameters see supplementary information Equation (S5)).

In the case of KAXQIL, the channels are one-dimensional tubes (see Figure S12) and the distance between two adsorption sites is approximately the unit cell parameter along the direction of the tube (5.6 Å). There the selectivity increases with pore filling, for enthalpic reasons, which we can explain by relatively simple reasoning. The Lennard-Jones potentials V_{LJ} can be estimated for all species at 5.6 Å: $V_{Xe-Xe} = -1.0 \text{ kJ mol}^{-1}$, $V_{Kr-Kr} = -0.3 \text{ kJ mol}^{-1}$ and $V_{Xe-Kr} = -0.5 \text{ kJ mol}^{-1}$. In a simplistic model where all adsorbed molecules are 5.6 Å apart, the cooperative effect is higher between two xenon molecules, which explains the increased selectivity at high uptake. If we look further at the adsorption enthalpy of both xenon and krypton (c.f. supplementary information Table S1), they both increase: the guest molecules move from the “ideal” adsorption

sites, and the guest–guest interactions do not fully compensate. The selectivity change in this material is therefore a consequence of the guest–guest interactions that rearranges the position of the adsorbates inside the nanopores.

To further corroborate the role of the guest–guest interactions, we look at another material with one-dimensional tube-like channels: JUFBIX, a cobalt(II) coordination polymer based on carboxylic acid linkers (see Figure S13). [JUFBIX] The periodicity along the direction of the tubes is much higher at 7.2 Å. The pair interaction energies corresponding to the LJ potentials at this distance are $V_{\text{Xe-Xe}} = -0.24 \text{ kJ mol}^{-1}$, $V_{\text{Kr-Kr}} = -0.06 \text{ kJ mol}^{-1}$ and $V_{\text{Xe-Kr}} = -0.13 \text{ kJ mol}^{-1}$. By looking at the adsorption enthalpies (Table 2.1), these values are too small to affect the position of the adsorbed molecules. At high loading, the distance between adsorbed molecules is high, and every adsorption site is independent of the others. The ambient-pressure selectivity s_1 is therefore the same as the low-pressure selectivity s_0 , since every guest–guest interactions are negligible. It confirms the crucial role of cooperative effects between guest molecules, when considering a saturated material.

GOMREG and JAVTAC are frameworks that belong to the second category of materials, with a moderate decrease in selectivity from low to ambient pressure. In GOMREG, the channels are composed of one-dimensional tubes larger than the ones found in KAXQIL or JUFBIX (see Figure S15 and Table S1). The adsorption sites are alternating from left to right inside the channel, and the adsorbed molecules organize in a “zigzag” pattern. Looking at the adsorption enthalpies, we see that both xenon and krypton have lower enthalpies by a similar margin, suggesting an equivalent stabilization for both atoms, hence the enthalpic contribution to the selectivity change is close to 1. Since krypton is smaller and less strongly tied on its adsorption site than xenon, it has more available space inside the pore space. This gives an entropic advantage to the Kr, seen in the entropic contribution k_S of 0.64 in Table 2.1. This indicates that even if enthalpic considerations mainly explain the observed changes at a statistical level, as discussed in the previous sections, for individual cases entropic considerations can be a strong factor in pressure-dependent selectivity.

The remaining materials discussed here form a third category, with a strong decrease in selectivity from low to ambient pressure. We look at several phenomena that can be at the root of this decrease, which is important for screening studies as it can limit the working performance of a material that appears to be a “top performer” based on zero-pressure screening.

For example, GOMRAC has a similar structure compared to GOMREG (see Figure S17), except for the fact that the pores and channels are smaller (see the values of D_f the diameter of the largest free sphere, and D_i the diameter of the largest included sphere, in Table S1). The distances between the adsorbed molecules – in their ideal sites – are then consequently smaller. At such distances, we can assume that the interactions between adsorbates become more stabilising for krypton than for xenon molecules in GOMRAC (see LJ potentials at distance lower than 4.2 Å in the Figure S10), which translates into an enthalpic contribution k_H of 0.58. Moreover, this is compatible with the equivalent guest–guest interactions in GOMREG, as previously discussed. It explains why difference between the adsorption enthalpies become smaller for GOMRAC, whereas it stays the same for GOMREG (between low and ambient pressure). This further validates the crucial role of the interactions between adsorbed molecules, and their relationship with the guest–guest distances when considering a high loading condition.

If we look at the case of MISQIQ, we see that the pure-component Xe isotherm in Figure 2.13 cannot be fitted by a single-site Langmuir isotherm, but is well-fitted by a two-site Langmuir model (see Figure S18). Visual inspection of the adsorbed density at various loadings shows that this is not a second, separate adsorption site that is populated at high loading: instead, the second step in the isotherm (representing about 20% of the uptake at full loading) is associated with a reorganization of the adsorbate molecules occurs at high loading, accompanying a contraction of the interatomic distances. In this case, the potential for reorganization of the adsorbate in the material's nanopores leads to the change in selectivity. This reorganization can be detected on the basis of the xenon isotherm alone, and has a major role in the selectivity at ambient pressure. This repacking of the adsorbed phase is linked to a strong entropic effect, and also impacts the enthalpic contribution to selectivity.

More extreme cases of selectivity drop can occur when more than one site is available, as is the case for materials BAEDTA01, VIWMOF, LUDLAZ, WOJJOV, and VAPBIZ. The pure-component isotherms and the representation of the materials loaded in xenon and krypton molecules (presented in supporting information Figures S19-23) confirm the existence of at least two distinct adsorption sites in each material. The most selective sites (i.e., the most favorable for Xe) are filled in priority at low loading, and the less selective sites will then be populated when the pressure increases, leading to a net selectivity drop at ambient pressure for these materials. The different types of adsorption sites, and therefore the potential for a drop in Xe/Kr selectivity (at non-zero pressure) is a factor that could be explicitly included in screening of pure-component isotherms, without the need for explicit multi-component GCMC simulations.

2.3.3 Conclusions and perspectives

In the current state of the art on Xe/Kr separation by adsorption in nanoporous materials, many studies have focused on the determination of structure/property relationships, the description of theoretical limits of performance, and the identification of top-performing materials, whether for existing experimental structures or for novel hypothetical structures yet to be synthesized. Here, we provide a study based on a high-throughput screening of the adsorption of Xe, Kr, and Xe/Kr mixtures in 12,020 experimental open-framework materials, in order to provide a better comprehension of the thermodynamics behind Xe/Kr separation in nanoporous materials and the microscopic origins of Xe/Kr selectivity at both low and ambient pressure.

The statistical correlation found between Henry's constant for Xe and Xe/Kr selectivity showed that the most selective materials are those with the highest affinity for xenon. To some degree of accuracy, we conclude that directly screening for Kr adsorption or for Xe free energy may not be necessary for a coarse-grained evaluation of a nanoporous framework selectivity. This could help build more efficient screening methodologies, for example with multistep studies with a first rough selection on Henry's constant at a low computational cost, followed by more expensive GCMC simulations on the selected materials (a gain that can be between 5 and 10-fold in our setup). Furthermore, inspection of the correlations between enthalpy and entropy contributions at low pressure showed that the adsorption-based separation process in the open-frameworks studied is mainly enthalpic in nature. We intend to extend the study in the future to other classes of nanoporous materials beyond MOFs, including covalent organic

frameworks, porous aromatic frameworks, purely inorganic porous frameworks such as zeolites, but also amorphous porous materials such as porous polymer membranes.

In order to use nanoporous materials to separate xenon from krypton, pressure swing adsorption (PSA) processes have been widely proposed: pressure is therefore a crucial thermodynamic variable in the separation cycle. Here, we studied the difference of selectivity between a system under very low pressure (at the zero loading limit, which is calculated at relatively low computational cost) and a system at ambient pressure (closer to working conditions, but obtained at higher simulation cost). We demonstrated that the selectivity could be highly dependent on the pressure, with high low-pressure selectivity that could be maintained in some materials at ambient-pressure selectivity, while in others there would be a large drop in selectivity: a high ambient-pressure selectivity requires high low-pressure selectivity, but the reverse does not hold.

Using a thermodynamic approach to describe the separation selectivity, we showed that the differences in selectivity between the different pressures (and therefore different loading regimes of the frameworks) are mainly explained by the evolution of the adsorption enthalpies for Xe and Kr. By focusing on specific examples, we uncovered the microscopic origins of these selectivity changes, and related them to the relative roles of host–guest and guest–guest interactions. Population of different adsorption sites, or repacking of the adsorbed phase at higher loading, can lead to drastic changes in the overall selectivity. The mechanisms behind selectivity at high pressure are complex and unique to each framework, requiring a good understanding of the interactions between guest molecules constrained in the nanopores. Nevertheless, our classification of the interactions at play can help in the future to design more efficient high-throughput screening procedures.

[transition to raess]

3

ADSORPTION MOLECULAR SIMULATIONS

3.1	Voronoi sampling	49
3.2	Rapid Adsorption Enthalpy Surface Sampling (RAESS)	49
3.2.1	Software description	49
3.2.2	Comparison to standard methodologies	49
3.3	Grid Adsorption Energies Descriptors (GrAED)	49

3.1 VORONOI SAMPLING

3.2 RAPID ADSORPTION ENTHALPY SURFACE SAMPLING (RAESS)

3.2.1 Software description

3.2.2 Comparison to standard methodologies

3.3 GRID ADSORPTION ENERGIES DESCRIPTORS (GRAED)

[transition into ML models]

4

ADSORPTION PROPERTIES PREDICTION

4.1	Machine learning	51
4.1.1	Introduction	51
4.1.2	eXtreme Gradient Boosting	51
4.2	Ambient-pressure prediction	51
4.2.1	From infinite dilution to ambient pressure	51
4.2.2	Interpretation of the ML model.	51

4.1 MACHINE LEARNING

4.1.1 Introduction

4.1.2 eXtreme Gradient Boosting

4.2 AMBIENT-PRESSURE PREDICTION

4.2.1 From infinite dilution to ambient pressure

4.2.2 Interpretation of the ML model

Origins of the selectivity drop

5

XENON AND KRYPTON TRANSPORT PROPERTIES

5.1	Existing methodologies	53
5.1.1	Molecular dynamics	53
5.1.2	Fast kinetic Monte Carlo	53
5.2	ML modeling	53
5.3	Molecular dynamics calculation results.	54
5.3.1	Why is it relevant?	54
5.3.2	Correlations	54
5.4	Fast diffusion calculation algorithm	54
5.4.1	Implementation in C++	54
5.4.2	Preliminary results	54
5.4.3	Visualization tool	54
5.4.4	ML model training.	54



5.1 EXISTING METHODOLOGIES

Experiment?

5.1.1 Molecular dynamics

5.1.2 Fast kinetic Monte Carlo

utrast ctutrast[Mace_2019] ML descriptors next steps

5.2 ML MODELING

Results

5.3 MOLECULAR DYNAMICS CALCULATION RESULTS

5.3.1 Why is it relevant?

5.3.2 Correlations

5.4 FAST DIFFUSION CALCULATION ALGORITHM

5.4.1 Implementation in C++

5.4.2 Preliminary results

5.4.3 Visualization tool

5.4.4 ML model training

6

TOWARDS THE NEXT GENERATION OF SCREENINGS

6.1	Flexibility	57
6.1.1	Problem, literature.	57
6.1.2	Snapshot.	57
6.2	Noble gas polarisability	57
6.2.1	Problem, literature.	57
6.2.2	Perpectives.	57

6.1 FLEXIBILITY

Final screening step, easy integration into the workflow of current screenings

6.1.1 Problem, literature

6.1.2 Snapshot

6.2 NOBLE GAS POLARISABILITY

Xe/Kr difference of polarisability Open Metal Sites/polar groups [\[20220421_pres\]](#)

[Not the best material, but interesting discussion on open metal site effect] Tao et al.[[Tao_2020](#)] looked at tuning (and improving) the selective adsorption of Xe over Kr by MOF open metal sites in the UTSA-74 framework structure.

6.2.1 Problem, literature

6.2.2 Perpectives

GENERAL CONCLUSIONS

The work presented in this thesis is



This work opens perspectives for



LIST OF PUBLICATIONS

PEER-REVIEWED PAPERS

PREPRINT

RÉSUMÉ EN FRANÇAIS

Introduction	63
------------------------	----



INTRODUCTION

[5 à 10 pages]

Les matériaux poreux sont des matériaux



RÉSUMÉ

Durant ma thèse, j'ai

MOTS CLÉS

simulation moléculaire, matériaux nanoporeux, criblage haut-débit, adsorption, apprentissage statistique

ABSTRACT

During my PhD, I

KEYWORDS

molecular simulation, nanoporous materials, high-throughput screening, adsorption, machine learning