ELSEVIER

Catalysis Today 50 (1999) 451–477

CATALYSIS TODAY

Heterogeneous catalysis: looking forward with molecular simulation

J.W. Andzelm, A.E. Alvarado-Swaisgood, F.U. Axe, M.W. Doyle, G. Fitzgerald, C.M. Freeman, A.M. Gorman, J.-R. Hill, C.M. Kölmel, S.M. Levine, P.W. Saxe, K. Stark, L. Subramanian, M.A. van Daelen, E. Wimmer, J.M. Newsam*

Molecular Simulations Inc. 9685 Scranton Road, San Diego, CA 92121-3752, USA

Abstract

Some of the areas in which we anticipate, over the next five years, notable advances in the application of molecular simulation to problems in heterogeneous catalysis are considered, in the context of recent progress to date. The areas specifically addressed are:

- expanding access to methods,
- quantitative structure-property relationships,
- building structural models to focus or pre-screen experiments,
- confidence in predicting local and extended structure
- reaction mechanisms, barriers and kinetics, and
- data for chemical process simulations.

In each of these areas, we indicate why we consider the topic significant, provide reference to topical work and suggest opportunities for future developments.

Keywords: Heterogeneous; Molecular simulations; Quantitative structure-property relationship

1. Introduction

The value of simulation in general stems from its ability to deliver, in a reliable fashion, information about the structure, properties or performance of a system in a more rapid, more detailed or more cost-effective way than is possible by experiment. This is true whether simulation is applied to the design of a new aircraft, to optimizing the layout of an integrated circuit, or to analyzing the fate of air-borne pollutants in the environs of a major city. It is equally true when simulation is applied to the time and length scales characteristic of molecular phenomena, the province of chemistry and the domain of catalysis.

The benefits of simulation have become steadily more obvious as the price:performance ratio of hardware has become relentlessly more compelling. However, the molecular simulation of materials is, firstly, a recent progeny, barely into its teens, and, secondly, the dramatic bounds in the capabilities of molecular simulation arise at least as much from improvements in theory, in algorithms and in strategy, as from the building of hardware brawn.

The drivers for the application of molecular simulation to heterogeneous catalysis are similar; if anything they are more persuasive. It is, however, only rela-

*Corresponding author. E-mail: jmn@msi.com

0920-5861/99/$ – see front matter
PII: S0920-5861(98)00482-9

tively recently that molecular simulation has become accepted as a practical contributor. There are several reasons, although mention of but a few helps underscore why future momentum of the field is expected to be strong.

Heterogeneous catalysts tax today's simulation armory in several respects:

1. Heterogeneous catalysts are solid state materials, and the influence of the extended surface or bulk environment on the characteristics of the active site geometry or chemistry is usually significant.
2. Most heterogeneous catalysts are multiphasic, with active phases or sites that are typically dilute and hence poorly scrutinized by experiment.
3. Depending on the specific application, heterogeneous catalysts may contain first, second or third row transition metals, or lanthanides, in addition to the main block elements. This elemental diversity, the presence of d- and f-electron containing metal centers, and the complexities of simulating the behavior of the heavier elements of the second and third transition rows demands imposes special requirements on molecular or quantum method machineries.
4. The paucity of detailed experimental data relevant to the active elements of heterogeneous catalysts under real conditions imposes, concomitantly, extra demands on the reliability of simulation in predicting geometrical structure.
5. Finally, we are dealing with chemical conversions requiring, for probing by molecular simulation, an ability to track the actual process of reaction.

A heterogeneous catalyst's complexities also represent some real advantages for simulation:

- The historical Edisonian route to catalyst development and improvement is frustratingly slow and can provide no insight into the ultimate performance of a given system; simulation can potentially both focus and accelerate an R&D effort and also reveal fundamental chemical constraints.
- Maximal insight can be coaxed from analytical data when interpreted and simulated based on an atomic level model.
- Methods and parameterizations can be validated on bulk data and applied then confidently to surface or defect site models.
- Discrete questions can be asked; whereas, experimentally, observation is an inevitable convolution of many different factors, we can, with simulation, often isolate one parameter and probe selectively its influence on the whole.
- Commercial catalysis is, fundamentally, a chemical phenomenon, a macroscopic manifestation of events at the molecular level. Even in these cases where a fundamental understanding of the catalytic chemistry is lacking, we can exploit the knowledge that performance is governed by molecular level phenomena to develop rules to guide optimization of materials and processes.

2. Areas of substantial ongoing progress

A forward-looking view must be taken from a current vantage point. We therefore highlight four exemplary application areas that have witnessed substantial progress over the past five years, as noted in the other papers in this and related volumes.

2.1. Catalyst characterization

In 1990, in discussing the molecular-level simulation of heterogeneous catalysts in general, the hurdle most frequently encountered was that of insufficient understanding of structural details at the atomic-level. Considerable attention has therefore been paid to developing and applying simulation as a complement to analytical measurement for catalyst characterization [1,2]. Most types of analytical data can now be computed directly from the appropriate atomic model; such computations then facilitate atomic level data being extracted from the experimental results. Some progress has also been made with direct data inversion, the development of structural models directly from experimental data.

2.2. Zeolite structural chemistry

Literally hundreds of recent papers describe applications of modeling and simulation to diverse questions relating to the structure, properties, performance and application of zeolites and related crystalline microporous materials [3–8]. The extent, breadth and sheer number of such applications far exceeds

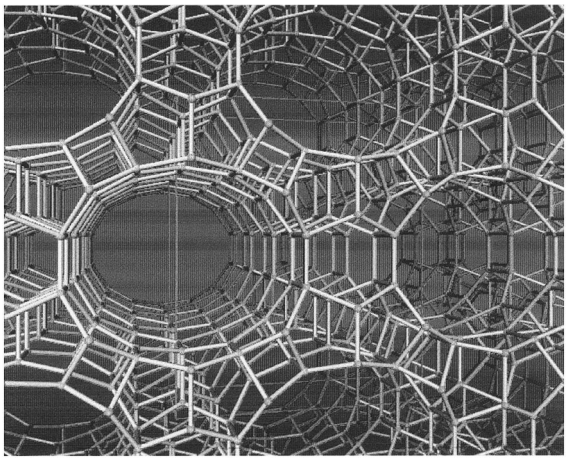

Fig. 1. The framework structure of the microporous material UTD-1 drawn as tubes connecting adjacent framework cation sites.

the scrutiny or success-rate that other catalytic systems have enjoyed. Reasons why zeolites are afforded such prominence include:

1. They are of prime industrial import, in several application areas, both catalytic and sorptive, as well as ion-exchange.
2. Zeolite macroscopic properties can be related directly to bulk, atomic-level structure (Fig. 1).
3. Zeolites can typically be prepared as pure, homogeneous, crystalline or polycrystalline samples, facilitating the application of a panoply of analytical methods.
4. Zeolites are compositionally relatively simple, containing only main group elements.

Partially because macroscopic properties relate so directly to atomic-level structure, zeolite structural characterization is a substantial field. The significant role that simulation plays is exemplified in, for example, the development of the simulated annealing approach to framework structure solution and design [9,10], the furtherance of crystallographic tools to accelerate physical model development [11], and the simulation of a slate of analytical fingerprints [1,2,12,13].

Attractive properties arise from the incorporation of aluminum into the siliceous zeolite framework. Simulation has probed local aluminum site preferences [14–19], framework aluminum distributions [12,20–24] relative to the experimental configurations measured by ^{29}Si NMR data [25], and interpreted experimental unit cell constant data in terms of disordered aluminum distributions in ZSM-5 [19]. ^{29}Si chemical shift dependencies on local environment have also been sampled [26,27].

The non-framework cations that charge-compensate for the framework Al^{3+} for Si^{4+} substitutions, using nominal valence charge language, can control or fine-tune a zeolite's macroscopic properties. Recent progress has demonstrated the feasibility of predicting non-framework cation configurations based solely on a framework structure and framework cation distribution model, not only for relatively simple systems such as zeolite Li–A(BW) [28], but also for the more technologically important cases of zeolite 4A [28], mixed Na–Ca and Na–Li zeolite A [29,30], sodium zeolite X [31], calcium LSX [30] and ETS-10 [32].

An occasional, and perhaps misguided, criticism of simulation, as of surface science, is that it applies to perfect, homogeneous samples and rarified conditions, such as in vacuo, or, in the case of structure optimizations, at 0 K. Although varying levels of complexity can be introduced into a model, a prime advantage of simulation is that it can be applied to probe one specific question or aspect of a system, independent of the many other issues that typically complicate the real material. Behavior at finite temperature is simulated by several techniques, such Monte Carlo (MC) simulations of adsorption heats and Grand Canonical Monte Carlo (GCMC) simulations of sorption uptake isotherms. Molecular dynamics (MD) is also, intrinsically, a finite temperature technique and, in addition to sampling the motion of sorbed or occluded organic species, MD has been used to explore pore dynamics in several framework topologies and pore mouth breathing [33,34]. To compute properties such as expansivities and phase transformations necessitates evaluation of structural free energies; these require calculations of the lattice dynamical contributions to the system entropy, which can be computationally demanding. Simulations of these types [35,36] have confirmed occasional experimental observations of negative thermal expansivities in certain zeolite frameworks [37].

Many technological applications of zeolites are based upon the behavior of hydrocarbons within their

micropores, a technological interest mirrored to some degree in the directions of fundamental research efforts. The established literature is substantial [5–8]. There are four common classes of simulation applied to organic sorbates within zeolites. The first seeks preferred binding sites, usually by a Monte Carlo docking procedure [38]. The second samples the set of configurations accessible at finite temperature and, by accumulating the corresponding configuration integrals, computes the heat of adsorption at that defined temperature [39–41]; these types of computations were pioneered by Kiselev [42,43], well in advance of the availability of modern computers. The third uses a Grand Canonical Monte Carlo protocol to simulate the equilibrium between the sorbed phased and an infinite vapor bath [44,45], hence permitting computation of sorption uptake isotherms or isobars for individual or competing components. The fourth uses molecular dynamics (MD) protocols [46–48] to sample configuration space, and to probe molecular motion, including rotational and translation diffusion, the latter providing a means of computing diffusion constants. Given the extent of the literature to date, we have not selected specific instances to cite; the most recent papers in the primary physical chemistry journals or zeolite conference proceedings provide swift access to the latest and earlier work.

The computational cost typically increases substantially from docking, to Metropolis Monte Carlo, to GCMC simulations. The cost of MD calculations depends on the size of the system, the form of the interatomic potentials and the time-span sampled. At least for small molecules, the required simulation mechanics are largely in place and validated. The rate-limiting constraint is generally the availability of suitable interatomic potential functions. For larger molecules, 'brute force' calculations become swiftly intractable and, for example, configuration bias Monte Carlo methods are used for probing binding sites and sorption heats [49] and transition state theory [50–52] or kinetic Monte Carlo [53] methods are applied to computing translational diffusion rates.

Many valuable applications of simulation relative to zeolite-based processes reflect not computation of absolute energies or rates, but rather the use of relative values as a guide in materials selection, screening or property rationalization. For example, molecular simulation was used to

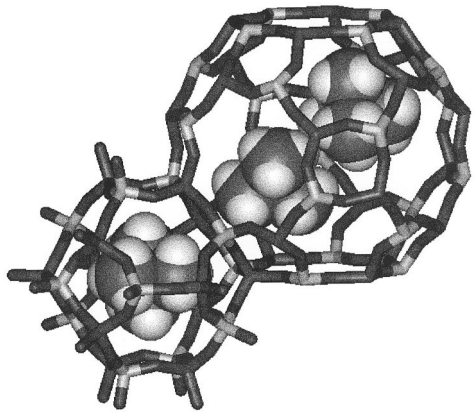

Fig. 2. Tetramethylammonium cations occluded within the sodalite or β-cage (left) and α- or supercage (right) in the LTA-framework of zeolite ZK-4 (after [76]).

1. screen a menu of zeolites for potential application in the preferential production of 2,6- vs. 2,7-dialkylnaphthalenes in the solid acid catalyzed alkylation of naphthalene with propylene [54],
2. probe the promotional role of dithiane oxide for dehydration chemistry in zeolite Y [55],
3. explore shape selectivity constraints on the preparation of neoalkanals and novel cyclic dioxepenes [56], and
4. sampling possible metal atom placements [57].

The technology initially developed for computing preferred binding sites for adsorbates within zeolites was quickly applied to the related question of the location of organic templates or structure-directing agents (SDA) occluded within zeolite structures during synthesis (Fig. 2). Close agreement between 'predicted' SDA positions and those measured by diffraction [58–60] has helped us understand the directing role that these agents play [61–64], promoted the use of simulation to assist in template selection (see, e.g., [62]) and encouraged the development of methods for de novo design of SDA's [65] that might promote the formation of known or hypothetical frameworks.

Simulation has also been applied extensively to probing the chemical reactivity of local sites within zeolites and, in certain cases, reaction mechanisms. The literature is, again, extensive [5–8]; other papers in this volume provide good examples and also con-

duits to the earlier works. These calculations almost exclusively employ quantum methods, generally in the cluster approximation, but, increasingly, using either hybrid quantum-mechanical methods to take account of the influence of the zeolite environment on the given active site model [66,67], or full periodic quantum methods using Hartree–Fock [68] or density functional theory [69]. Ab initio molecular dynamics, using plane wave basis sets within the Car–Parrinello formalism, have recently been brought to bear on Brønsted sites in offretite [70], methanol in sodalite [71,72], methanol in chabazite [73], water in zeolite A [74] and, in what is at the time of writing the cutting edge of such computations, at least in terms of system size, methanol in ZSM-5 [75]. It is, however, sobering that even expert practitioners can reach differing conclusions on the gross structural chemistry in these systems. For example, an early report suggested, perhaps counter-intuitively, that sorbed methanol was little perturbed in an acid sodalite, but protonated in an acid chabazite at a similar framework composition [69]. Subsequent dynamics calculations on the same chabazite system, again using plane wave basis sets, indicate the protonated state to be a local minimum on the complex potential energy surface, the hydrogen-bonded, neutral methanol configuration being of lower energy [73].

2.3. Metal surfaces and reactions on metal surfaces

Given the importance of metal surface chemistry and the broad industrial use of supported metal catalysts, simulations of metals, metal surfaces and reactions on metal surfaces comprise a substantial literature. Many metallic systems, including metal clusters, metal surfaces, and metal surfaces or clusters with surface adsorbates have been simulated, with some ingenuity and computational effort, using first principles calculations based on density functional theory [77,78].

Recent density functional theory calculations using plane wave basis sets [79] include those of O_2 on Ag(1 1 0) [80,81] and Pt(1 0 0) [82], H_2 dissociation over Cu(1 1 1) [83] and W(1 0 0) [84] and the interaction of carbon monoxide with metal surfaces [85] such as palladium [86,87]. Applications of density functional theory to cluster models, using localized basis sets, include, from our own past work, studies of small Pd_nH_m clusters [88], simulation of the geometrical structures of small $(MgO)_n$ clusters [89], analysis of the influence of chlorine on cluster models of oxided silver, of interest in ethylene epoxidation catalysis [90], quantitative studies of the dissociative chemisorption of diatomic molecules such as NO on cluster models of different copper surfaces [91,92] (Fig. 3) and, using the Harris functional [93], the sulfiding of copper surfaces [94,95].

To study the dynamics of large systems, that demand large simulation cells and potentially protracted simulation times, more approximate simulation schemes have proven useful and continue to be used extensively.

Empirical or semi-empirical approaches to the simulation of metal surfaces and their behavior, such as defect formation, boundary interactions and surface diffusion, exploit potential functions that can be rapidly evaluated; the embedded atom method (EAM) potential developed by Daw and Baskes [96], equivalent crystal theory (ECT) [97] and effective medium theory [98]. In addition to such approaches, where very large systems (thousands of atoms) may be handled, tight binding simulation methods have a substantial history of application to metallic systems, [99,100] offering a treatment analogous to the semi-empirical approach widely employed in organic chemistry. The explicit treatment of small molecule adsorption on surfaces, using empirical energy functions, has been demonstrated at a practical level by Sellers and Shustorovich [101,102] using simple potentials and the premise of system conserved bond order. The resulting formalism makes the treatment of chemisorption and surface reaction phenomena possible without excessive computational effort.

A complementary approach applies dynamic Monte Carlo simulations to surface rate processes [103], to which experiment can provide substantial parameter input [104].

2.4. Homogeneous catalysts

Organometallic complexes promote a diverse range of chemistries. As such complexes can often be characterized accurately by single crystal X-ray diffraction, perhaps combined with probes of structure in solution, such as NMR, they are attractive targets for

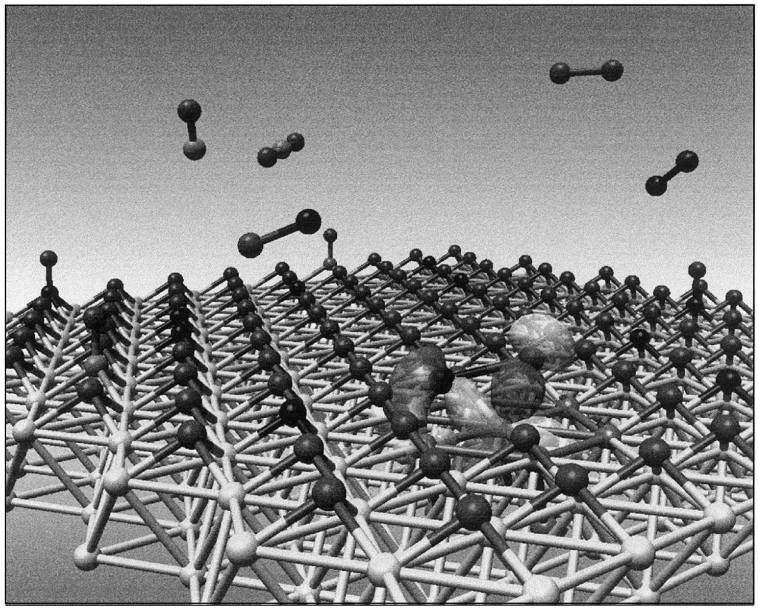

Fig. 3. Illustration of discrete 10-atom cluster model of the Cu(1 1 1) surface, and the computed transition state for NO dissociation on this cluster, together with the HOMO in the transition state [91,92].

simulation. However, current force field technology for organometallic systems has, until fairly recently, been less robust for metal complexes and the size of practical organometallic catalyst centers has usually placed them on, or beyond, the cutting edge of first principles methods.

One family of the materials to whose development molecular simulation has contributed are metallocene polyolefin catalysts, which promise a next generation of plastics for industrial applications [105,106]. Force–field technology has been a helpful tool in the rationalization and even design of the ligand environment around the titanium or zirconium metal center, the steric environment defined by the ligand substituents being known to control the tacticity of polyolefins that are produced. As such single site, or constrained geometry catalysts have ostensibly well defined, discrete molecular structures, they also avail themselves to study by first principles methods. Several studies describe significant aspects of the geometrical and electronic structures of relatively simple metallocene complexes or models thereof (e.g., [107,108] and, as examples from our own work, [109,110]), and of stages along the reaction coordinate defined by the accepted Cossée mechanism of olefin insertion. The role of the methylalumoxane used as co-catalyst, and the effect of solvent has, as yet, been little explored. Recent first principles dynamics calculations, initially on the archetypal ferrocene [111] and subsequently on models of metallocene polyolefin catalysts [112,113] demonstrate the tractability now of these types of calculations, albeit still at substantial time or computational cost.

These studies also reinforce that raw molecular dynamics are not a preferred technique for probing activated phenomena, such as essentially all catalyzed chemical conversions or, as noted above, translational diffusion of even modestly sized physisorbed molecules in a microporous host. In the first principles molecular dynamics study of the model $(CpSiH_2NH)TiR^+$ center, for example, it was necessary to first determine the transition states and reaction coordinate by 'static' methods and then perform dynamics for only degrees of freedom essentially orthogonal to this defined reaction coordinate [113]. Raw first principles molecular dynamics is not a technique for exploring, de novo, possible reaction chemistries that involve significant barriers.

3. Expanding access to methods

The widespread implementation of small computers for instrument control, analysis, information and data access, and 'office productivity' provides an environment well primed to accommodate broader deployment of molecular simulation. From a hardware perspective, this is fueled by steadily improving price:performance and by the 'digitalization' of experiments. From a software perspective, growth is gated by the degree to which simulation adds practical value to the experimental laboratory environment.

3.1. Necessary ingredients

To allow non-simulations specialists to take real benefit from molecular simulation requires three ingredients. First, the methods need adequate validation; to pre-package a specific application it must be clear for which problems a particular method is appropriate, what force–field is apt, or which quantum mechanical framework is suitable. Preferably, suitable boundaries can be defined within which, even, error bars on the computational results can be provided.

Second, access to the set of capabilities must be structured as what we term a 'computational instrument', analogous to an experimental instrument, in which a specific, well-validated calculation is pre-parameterized, precalibrated and provided in a well-structured form. That is, it should permit the 'sample' to be specified, a small number of intuitive and understandable controls to be set, and then deliver the desired simulated data. The analog to an X-ray powder diffractometer, would be a powder diffraction computational instrument; the sample (the structure) is specified, the type of instrument or instrumental parameters are set, and the 'computational instrument' delivers the simulated powder X-ray diffraction profile. Similarly, for other X-ray, neutron or electron diffraction or scattering data, infra-red, UV-vis spectra, EXAFS, high resolution transmission electron micrographs (HRTEM) etc. (Table 1).

Thirdly, such computational instruments will become useful broadly only when they are routinely available, when provided on the experimentalist's workbench or desktop; they need to be accessible alongside the other productivity tools. Even the industrial simulations specialist [114], whose work profile is dominated by applications of simulation, can benefit from closer integration between the desktop environment and their simulations workbench.

Table 1
Some experimental sources of structural information when experiment is combined with appropriate data simulations

Meso–micro scale
Light scattering
Microtomography, NMR imaging
Long range structure
Powder diffraction (X-ray, neutron, electron)
Single crystal diffraction
Fiber, film diffraction
Wide-angle X-ray, neutron scattering
High resolution lattice images (electron microscopy)
Scanning electron microscopy (SEM) micrographs
Surface structures
Scanning tunneling microscopy (STM), Atomic force microscopy (AFM)
Grazing incidence X-ray diffraction
LEED, RHEED
Local structure
Extended X-ray absorption fine structure (EXAFS)
Solid state NMR
Electron spin resonance (ESR)
Electron spin echo spectrometry
Electronic structure
UV-vis spectra
X-ray photoemission
Structure dynamics
Infra-red, Raman spectroscopy
Inelastic, quasielastic neutron scattering

3.2. Satisfying these requirements

Validation is cumulative; it is contributed by the global simulations community, at least to the extent that results are disseminated in an open and complete fashion. Validation enables the design and software engineering of the 'computational instrument'. To permit access from the desktop is ostensibly also a software engineering issue, but, in practice, it has been made possible more by the step-change in our approach to global computation created by the Intra- and Internet, by the global adoption of the Web paradigm and of Web technology.

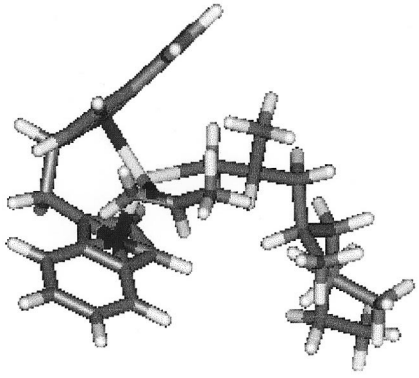

Fig. 4. Titanocene olefin polymerization catalyst center, with a growing polythene chain. This image is 'cut-and-paste' from the WebLab Viewer and, in the Microsoft Word document from which the manuscript was printed, is incorporated dynamically via the OLE2 protocol.

What might be listed as a first requirement is a molecular graphics capability that runs natively. Past protocols that allow a terminal to access the graphics created on a networked machine, such as X-windows, have been helpful, but comparatively rudimentary in performance terms. Fig. 4 is an illustration of a metallocene catalyst center to which is bonded a growing polyolefin chain. In this hard-copy printing the image is static. However, on the PC on which this text was crafted, this image is embedded via the object linking and embedding standard protocol (OLE2)[1]. The WebLab Viewer used in this particular case, one amongst several similar types of freely-available molecular graphics applications, runs natively, hence avoiding the communication shortfalls of early client–server configurations for graphics and, at the same time, benefiting from the impressive throughput of, for example, the Intel Pentium Pro processor to provide acceptable molecular modeling performance. For displays of simple molecules there are many options currently, mostly at no charge; PC-based display systems for most heterogeneous catalysts are, at present, lacking.

We suggest, as the second requirement, access to an appropriate suite of validated server methods. Had Web technology not become accepted as a global standard, this might have required implementation of various methods locally on a PC with what is, today, still significant limitations on performance, or a likely esoteric client–server mode. The WWW has introduced an expectation of truly seamless client–server architecture. It does not matter 'where' the site providing the data or connection is located. The URL could be local to the machine running the Browser, within the same office or building, within the same campus, within an internal company network protected by a firewall or geographically remote in the richness of the Internet. This, we believe, is a boon for molecular simulation, as the 'client', the front end, need provide only the necessary display, interface and access capabilities, and the CPU demanding computations can then be farmed seamlessly to a suitable machine, perhaps within the security of the corporate intranet. Many of the key simulation methods are optimized to run on such a server machine.

An integral component of this network computational environment recognizes that computer experiments are not intended to stand independently, but rather to complement the other data gathered in the lab. Computer networking has facilitated the communication of data within laboratories; over the past year, the technology that created the world wide web has demonstrated its value as both a communication device and as an infrastructure for data management. This same technology can be leveraged to provide an access to the wide range of complementary computational methods currently accessible to dedicated applications on scientific workstations.

Computing trends will clearly continue to lower the access barrier to molecular simulation methods, and innovative software design is emerging to best leverage this progress. However, computer experiments still require a knowledge base and much further validation of methodologies is needed. The following sections describe many of the areas where computer methods are having an increasing impact on catalyst research.

4. Quantitative structure-property relationships

4.1. Introduction

Frequently, we are well aware that macroscopic properties or performance are determined by molecu-

[1]The WebLab Viewer for Macintosh and PC can be downloaded from http://www.msi.com.

lar level behavior, but we neither know the details of this inter-relationship, nor are we able to simulate the macroscopic properties directly. In such cases correlative methods can be useful [115–118]. Molecular attributes or 'descriptors' are computed for a 'training' set of molecules or materials for which properties have been measured; correlations or patterns are then sought, by a variety of analytical or statistical techniques, between the determining molecular descriptors and the macroscopic properties. Validation of this correlative model by prediction of the properties of a subset of systems measured, but not included within the training set, helps define the bounds of the response surface of the system [119,120], within which the model can be used to predict the properties of other candidate systems. These correlative methods are termed quantitative structure property relationships (QSPR) [121] or quantitative structure activity relationships (QSAR). They have long been used in drug discovery and lead optimization [115–118,122,123]. They also have potentially broad applicability in the materials sciences (Fig. 5). As a general rule, these methods are applicable within the compass defined by the training data set; they are substantially less reliable when applied in any way that is extrapolative.

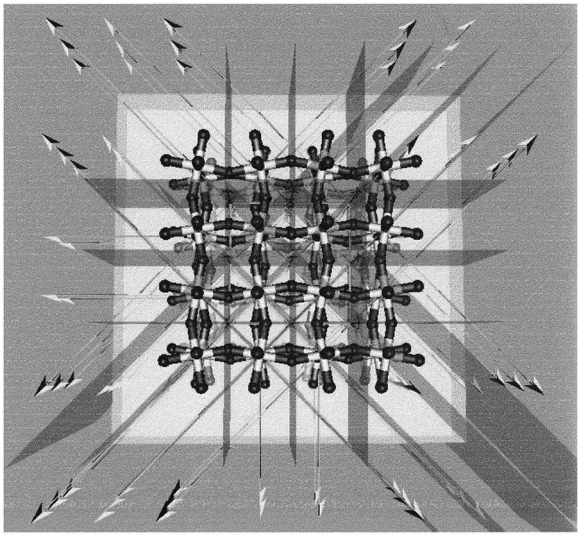

Fig. 5. Configuration of tilted, corner-shared octahedra in a model perovskite with the corresponding symmetry elements overlaid. The rich compositional diversity of materials with the perovskite structure suggests opportunities for correlative methods.

4.2. QSAR, neural nets and expert systems

Traditional or stochastic methods of QSAR analysis such as forward regression, stepwise regression, principal component or factor analyses [115–118] require that

1. the number of features measured within the experiment is large compared to the number of parameters of the model,
2. the descriptor functions are not complex or non-linear in their response, and
3. the descriptor functions are not strongly interrelated.

Recent developments have helped broaden, significantly, the applicability of QSAR approaches.

A computational neural net (CNN) [123–125] represents a system containing many highly interlinked nodes that captures the complex interrelationships between an input set of data, the molecular descriptors in this case, and the output, the corresponding properties. A neural net is initially 'trained' on a set of 'experiences', input and output responses. Neural nets have been applied to a broad range of optimization problems, such as of refinery operation, smart weapons targeting and traffic patterns (see, e.g., [126,127]. They have also been used to relate mixed metal oxide compositions and characteristics to catalytic performance [128]. Two key disadvantages, though, of CNNs are opacity and irreproducibility – even a net that has predictive reliability often yields little insight into the fundamental physics of the problem, and each training run will typically yield a different network, even based on the same set of input data.

Non-linear or non-stochastic models have been used in a range of expert systems for the prediction of performance and process variables, such as in steel manufacturing, automotive design, and chemical processing, in navigation and global positioning, pattern recognition, such as for satellite image interpretation, optical character recognition (OCR) in fax machines and palm top computers, and in financial modeling for guiding stock trading and equity flow (see, e.g., [129–131]).

Genetic algorithms (GAs) mutate a discrete set of parameter values for a given function, propagating those mutations which represent a 'fitter' or lower cost parameter set, and hence converging by 'natural selection' over a large number of 'populations' to the optimal parameter values [122]. Genetic function algorithms (GFAs) [125] mutate not only the parameters within a defined analytical expression, but also the functions that comprise the analytical expression itself and can therefore accommodate non-linear, complex and even noisy data [119]. They output, however, sets of simple descriptive functions that can often be interpreted on a physical basis and highlight the nature of any non-linearities. The effectiveness of these simple functions can then often be easily tested.

4.3. QSAR for property prediction

QSAR is a proven vehicle for predicting a broad range of macroscopic properties, including solution phase behavior, partition functions, and diverse polymer properties [132] such as glass transition temperature [120], Young's modulus and thermal expansion coefficient. Non-linear modeling methods such as GAs and CNNs have been applied to ceramics, semi-conductors, metals and alloys, and matrix composites. CNNs and GAs have also been used in calculations of potential energy surfaces for chemical reactions, by interpolating the complex energy hypersurface, sampled at a small number of discrete points by quantum or experimental methods, using splines and non-linear expressions. Blank et al. [133], for example, studied the reaction surface for diffusion of chemisorbed CO on Ni(1 1 1). Neural nets have been applied to reactions that include Markovnikoff and Saytzeff eliminations, and cyclo-, and Diels–Alder additions. Recent advances in GAs and ring counting analyses using Kier and Hall indices have also allowed QSAR methods to be applied to metallocene polyolefin catalysts.

4.4. QSAR in analytical studies

Non-linear modeling methods such as CNNs, GAs and clustering methods can assist in data classification and pattern matching, applicable in areas such as spectroscopy and chemometrics [134]. Munk and coworkers [135] identified, for example, infra-red (IR) peaks characteristic of particular side groups and then used their occurrence in spectra to identify the species present. Other spectroscopic applications [136] include NMR [137], atomic absorption, IR and UV spectroscopies. Spectral analysis tasks include instrument geometry optimization, feature and difference feature identification, curve fitting and peak decomposition. Non-linear fitting has, for example, been used to interpret the laser spectroscopy of alkanes, the near infra-red spectra of polysaccharides and the Mössbauer spectra of iron containing minerals [138]. In chromatography, various QSAR methods have been used to relate the retention characteristics on a particular column phase to the molecular characteristics of the eluent [139,140]. Similarly, QSAR models have been used to correlate the fragmentation and charging patterns reflected in mass spectra with molecular structural features. GAs have also been exploited in structure solutions and predictions [141–144].

4.5. QSAR summary

Linear and non-linear correlative methods, in common with other function optimization procedures, find diverse applications; in the present context, this is, indeed, one of their attractions. Some topical areas include the use of CNNs and GAs for 'experimental design' [145]; it is conceptually interesting to consider applying statistical methods to the library of solutions that a genetic function algorithm (GFA) yields, to explore parameters such as diversity or consensus, and to base choices for subsequent experiments on the degree to which they are, based on a correlative analysis, most likely to extend our knowledge of the system.

Process design and optimization, as well as on-line process control, require optimization of overall performance by altering layout or operating conditions and, possibly, feedstock composition, subject to a complicated set of constraints. Such optimization problems are well primed for the deployment of correlative methods and exemplary application areas encompass semi-conductor processing, steel rolling, resin composite formulation, and catalytic reactor optimization.

Correlative methods, as embodied in QSAR techniques, have a substantial history of successful application in the rational drug design field. This well-validated framework, extended by developments in CNNs, GAs, GFAs and expert systems tools, offers a fresh approach to many materials research and development problems. Some early materials examples validate this potential and set the stage for what is expected to be an expanding number of practical applications.

5. Building structural models to focus experiment

5.1. Structure characterization

An atomic-level structural model is a usual starting point, not only for further simulation work, but also as the basis for any fundamental understanding or rationalization of catalytic behavior. Further, structural characterization at the atomic level will likely be required more and more frequently, in the process of describing or defining a material or materials class, such as is required for a composition of matter patent application. Crystal structure determinations are already generally required for drug molecule patent applications, and the deeper insights into analytical or performance characteristics that stem from an atomic-level structural understanding can be an aid in the patenting of inorganic materials also (e.g., [146]).

As noted above, simulation has been applied broadly to structural characterization, which, for many classes of catalytic materials can be difficult. For example, planar faulting, to which several crystalline microporous solids are prone, complicated the analyses of ETS-10 [147] and zeolite beta [148,149].

Cumulative structural data provide a rich foundation for crystal structure model development, analysis and design. Several crystallographic databases are readily available, including the Cambridge Structural Database (CSD) [150], the inorganic crystal structure database and more specialized compilations such as the metals and alloys databases, the database of surface structures and the on-line atlas of zeolite structure types (access to all of these data is possible via the Internet; a suitably phrased search will yield the URL's for these sites and several others). The data from each of these databases can be imported directly into crystallographic modeling and simulation packages, providing starting points for edits and other model design approaches.

Structural edits may range from the adjustment of atom types to account for atomic substitutions, to the construction of supercells or derivative structures to accommodate defined adsorbates or critical elements of a particular system. As many analytical fingerprints, such as powder diffraction, can be computed dynamically from the developing model (Table 1), the degree to which such edits increase the match of the model with the characteristics of a measured system can be gauged.

Crystallographic modeling tools have become steadily more accessible and more useful [151], allowing application to accelerated crystal structure design and development [11]. Additionally, automated, Monte Carlo based approaches to model development or structure solution are quite promising. Notably, simulated annealing techniques have been employed in the determination of zeolite structures [9,10,152,153] and in analogous applications [154–158] (Fig. 6). Recent progress in using MC, MD and structure optimization techniques for the location of non-framework cations or anions within inorganic framework structures was noted above.

For more disordered structures or amorphous structures, such as silicate glasses, microporous carbons and transition aluminas, structural models are less easily validated. A variety of methods, such as molecular dynamics (MD), Monte Carlo (MC) or reverse Monte Carlo (RMC) can be applied to general trial structures. Geometric information, readily obtained from the computer-generated models, such as bond angle distributions (BAD) and radial distribution functions, can then be compared with interpretations of NMR, wide-angle X-ray or neutron scattering, or spectroscopic data [159].

Validated in application to models of the crystallographic bulk, classical and quantum mechanical methods can be applied to local defect and surface structures. Quantum mechanical treatments are discussed further below. Force field technology has been applied frequently to calculate the relative energetics involved in the creation of a variety of different surfaces, the quantitative bases for predicting details

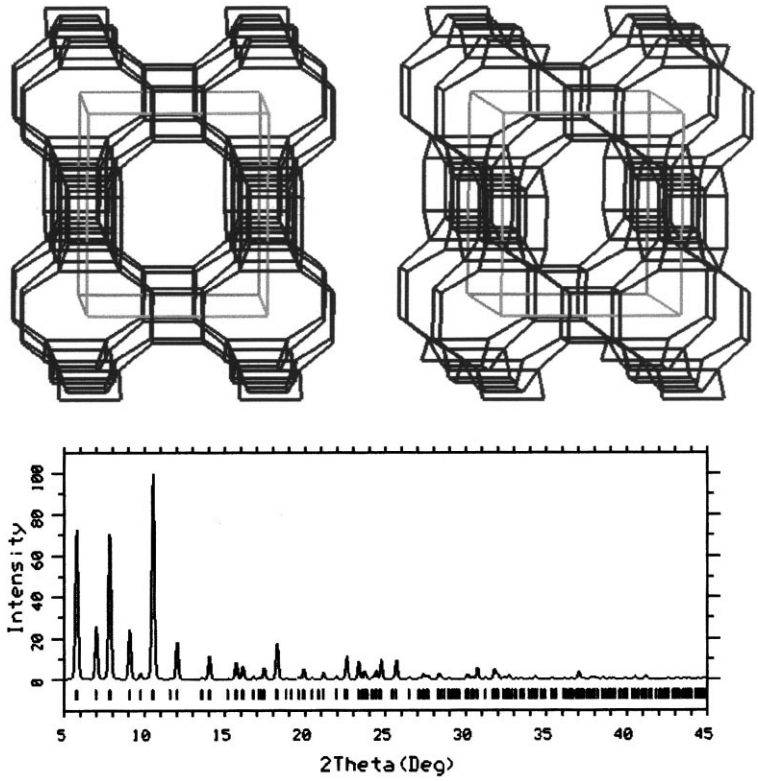

Fig. 6. One of the many viable framework models developed as models for a new zeolite material (upper; stereoview, framework atoms drawn as connections with adjacent nodes) and the corresponding simulated powder X-ray diffraction profile (lower: computed using apical oxygen atoms introduced, automatically, at the mid-points of the T–T vectors).

of crystal habit and morphology [13,160]. The effect of defect and sorbent interactions on such surface stabilities may then be investigated [161].

5.2. *Hypothetical models*

A prime advantage of simulation, and, in terms of identifying ultimate materials properties, what may be the most impactful advantage, is that it can be applied as readily to hypothetical models as to those validated for real materials. The reader can explore computationally which geometrical configuration will maximize the HOMO–LUMO difference, or yield the maximal difference in interaction energies between physisorbed molecules, or prevent attainment of a transition state along an undesired reaction pathway. These calculations may be performed whether or not a synthetic route to instantiate the configuration is available; the results of the analysis should indicate whether, indeed, such a synthetic pursuit would be warranted.

There are few examples of this kind as yet in the literature, largely because of a global emphasis on validating molecular simulations directly against various direct experimental data. Santilli et al. [162] used simulation to probe the influence of pore dimension on the product selectivities obtained in hydrocracking of long-chain alkanes in Pd-loaded zeolites with one-dimensional channels. In the ratios of computed sorption heats for *n*-hexane and dimethylbutane, and in the corresponding production of these isomers in hydrocracking, they observed a preference towards the branched isomer at intermediate channel sizes, a phenomenon they termed inverse shape selectivity. Certain pore diameters, however, were not represented in the slate of materials available. They therefore performed the sorption heat calculations on hypothetical models with the ‘missing’ dimensions, helping to

identify whether a synthetic program targeted to produce materials with such pore dimensions would be worthwhile.

6. Confidence in predicting local and extended structure by density functional methods

6.1. Density functional theory background

Progress in first-principles quantum mechanical methods has culminated in an unprecedented capability to predict atomistic structures of molecules, solids, and surfaces. Compared with experiments, present theoretical and computational approaches allow the prediction of interatomic distances with an accuracy of some 0.04 Å and bond angles within some 4° for a wide variety of compounds. This confidence is based on a large number of calculations for molecules and ordered crystalline compounds, where accurate experimental data are available. It is reasonable to assume that the same level of accuracy is achieved for defects, surfaces, and clusters or for systems at high pressures, where experimental data are less accessible.

This current status is the result of systematic and actually quite predictable developments over the past decades. Three major milestones characterize this progress, namely

1. the accurate calculation of the electronic structure for molecules, solids, and surfaces containing any element in the periodic table,
2. the evaluation of total energies and forces, and
3. the combination of molecular dynamics with ab initio electronic structure methods.

Density functional theory (DFT) provides the necessary framework to meet all of these requirements. This theory was originally intended for extended systems such as bulk solids and surfaces, but has proven to be successful also for isolated molecules (e.g., [163]). It is this unifying aspect of DFT which makes this approach particularly suited for problems in heterogeneous catalysis.

Practical DFT implementations use either an all-electron approach as, for example in methods such as Dmol [164,165], Dgauss [166], deMon [167], Turbomole-DFT [168] and Gaussian [170], and ESOCS, that employs the atomic sphere approximation and augmented spherical wave basis functions [170], and the full-potential linearized augmented plane wave (FLAPW) method [171] or pseudopotentials as implemented in Turbomole [172] and plane wave programs such as CASTEP [79]. For an overview of the various methods, the reader is referred to articles such as [78,79]. A notable feature of these various programs is their availability, either as academic codes or with full commercial support.

By now, these approaches are well understood and tested. Density functional theory and the local density approximation were formulated over thirty years ago; the use of pseudopotentials dates back to the early days of band theory; plane waves have gone hand-in-hand with pseudopotentials since then. The first density functional total energy calculations on bulk solids and surfaces were performed around 1980. However, present calculations are distinguished in at least three major aspects:

1. the level of confidence,
2. the computational efficiency, and
3. the inclusion of dynamical aspects.

The *accuracy* is mainly governed by the functional used (local, non-local or hybrid), the type and quality of the basis set(s) and pseudopotentials, if employed, and the choice of other computational parameters such as numerical integration grid, the density fit or the high energy cut-off in the plane wave case. Both local (LSD) and non-local (NLSD) versions of the method allow for an efficient computer implementation. For a system of N atoms (or, strictly, basis functions), the computational time is expected to scale as some N^3, whereas the Hartree–Fock (HF) ab initio method and the hybrid DFT–HF adiabatic connection method (ACM) scale as N^4. In the limit of large N, both methods scale as N^2; however, for densely packed systems representing heterogeneous catalysts, such as surfaces or clusters, the LSD or NLSD approach can be implemented as much faster algorithms than is possible for the HF or ACM methods. Semi-empirical programs are orders of magnitude faster than DFT algorithms, but they are restricted in application, like force fields, to the types of system for which they were parameterized. On the other hand, DFT methods can be applied to molecules, clusters and surfaces including all the elements from periodic table. No adjustment of DFT parameters is needed and an already

large and growing applications literature validates the useful accuracy of the DFT approach.

For molecules and clusters, the relatively new hybrid DFT–HF adiabatic connection method (ACM) appears to represent a significant advance [173]. As implemented in Turbomole-DFT [168] or its equivalent called B3P or B3LYP in Gaussian [169], this hybrid scheme combines elements of Hartree–Fock and Kohn–Sham methods. It was initially proposed by Becke [174] as a semi-empirical combination of various density functionals together with a term representing the exact exchange. The coefficients of this formula were determined by the best fit to the experimental data from Pople's G1 set [175]. The significant theoretical effort which followed provided rigorous justification of the ACM approach within exact DFT theory (see references in [173]).

The *confidence* in the DFT results is based on a large body of calculations on bulk solids, surfaces, and molecules. Accurate LSD calculations give equilibrium bond distances typically within about 0.02 Å compared with experiment, except for weakly bonded systems such as those involving hydrogen bonds or van der Waals interactions, where computed bond lengths tend to be too short. Relative energies, for example between two similar adsorption sites of an atom on a surface, are significant to within several kcal mol^{-1}, or even better, on the LSD level. Current gradient corrected density functionals give absolute binding energies typically within several kcal mol^{-1} and result often in better geometries for weakly bonded systems.

The *computational efficiency* is the result of the development of methods such as the pseudopotential plane wave approach using conjugent–gradient techniques for obtaining the LSD eigenstates. Additional approximations, as introduced by the pseudopotentials and truncations in basis sets, have been validated by comparison with accurate all-electron approaches such as the full-potential linearized augmented plane wave (FLAPW) [171] and are thus well controlled.

6.2. Geometries – chemisorption on metal surfaces

Although the nature of the model to be simulated guides the first choice of method, decisions about the level of approximation to employ impact significantly the computational cost and, frequently, the accuracy of the results. Calculations on the simple, but technologically important process of hydrogen adsorption on the Pd(1 0 0) surface, are illustrative of such issues in probing chemisorption on metal surfaces (analogous calculations on optimized cluster models have been reported previously [88]). Several clusters were employed (Fig. 7) to simulate H atom adsorption over the 4-fold site of Pd(1 0 0), (C-position); 2-fold site (B-position) and 1-fold site (A-position). The clusters contain 5, 8 and 9 palladium atoms, for C, B and A-positions, respectively (see Fig. 7). In each case all the first neighbors in the surface surround the H adsorption site. Certainly, the size of the cluster and its mode of termination may significantly influence results, particularly for binding energies, which are known to converge slowly with the size of the clusters [88]. However, these clusters are a reasonable test case on which to compare relative adsorption sites as calculated by using various DFT-approaches.

Calculations were performed by using the Turbomole-DFT program [168]. Local (Vosko–Wilk–Nusair (VWN)), non-local (Becke–Perdew (BP)) and adia-

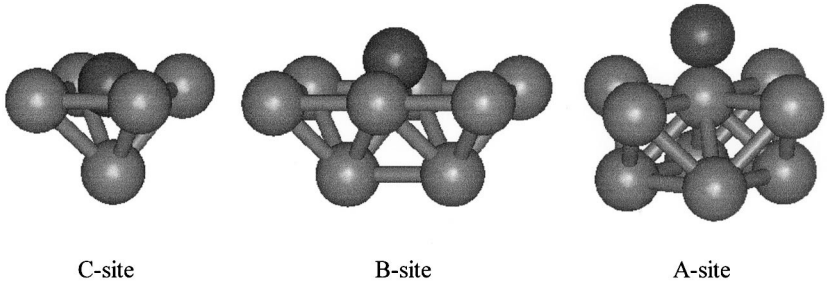

Fig. 7. Cluster models of the immediately local environments of H atom adsorption on the prime sites on a Pd(1 0 0) surface; the C (Pd_5H: C4v), B (Pd_8H: C2v) and A (Pd_9H: C4v) positions reflect 4-, 2- and 1-fold sites, respectively, for the H atom adsorption.

batic-connection (ACM) spin-polarized calculations were performed for clusters with and without hydrogen. The position of the H atom above the surface was optimized in each case; the Pd atoms were fixed at the (1 0 0) unreconstructed surface geometry. A split valence+polarization basis set for hydrogen and relativistic pseudopotential for Pd were used. The 4s, 4p, 4d, were active in the valence space, giving a total of 18 electrons per Pd atom. A standard numerical integration grid and default convergence parameters were used. The energies of clusters without hydrogen were calculated using the spin-restricted DFT approach and the same basis sets and other parameters as in the case of presence of hydrogen. The results are presented in Table 2.

The results from Table 2 clearly indicate that the geometry of H in each adsorption site studied depends little on the level of theory employed in the DFT-calculations. The maximum differences are within 0.01, 0.02 and 0.03 Å for C, B and A-positions, respectively. Clearly, for many practical applications this level of agreement is more than adequate and indicates that the simplest DFT approach, the local spin density method, may be used successfully to study geometries of adsorbed species on palladium clusters.

All the DFT-methods find the same 4-fold position (C-position) to be the most stable adsorption site. This is in full agreement with the experimental results which indicate that the 4-fold sites are the preferred hydrogen chemisorption sites on Pd(1 0 0) surfaces. The B-position is less preferable by about 12 kcal mol^{-1} according to VWN calculations and 14 kcal mol^{-1} from the ACM study. Again the spread of results of some 2 kcal mol^{-1} is probably not significant. The A-position is clearly the least stable and here the results from the different theory levels vary somewhat more, by as much as 8 kcal mol^{-1}. Further examination of the results for A-site absorption reveals that for the BP case, the lowest electronic state for the clusters was probably not yet attained, as the Fermi statistics are not precisely maintained in BP calculations (although they are in the ACM case). The 16a1 orbital of spin beta for Pd_9H (C4v) should probably be occupied instead of the 6a2 orbital. This results in a somewhat different electronic structure for the ACM and BP clusters, which may contribute somewhat to the binding energy difference between the BP and ACM calculations. The lower binding energy of BP corresponds to the longer bond distances of H–Pd. Further optimizations of the electronic structures of these clusters using BP functionals are in progress, but the qualitative conclusions of this work are not expected to change.

Table 2
Optimized geometry of H in C, B, and A positions above the Pd cluster and the relative binding energies in kcal mol^{-1} as a function of the DFT-approach used

DFT method	Adsorption site		
	C	B	A
Distance H-nearest Pd atom (Å)			
VWN	1.956	1.638	1.530
BP	1.962	1.649	1.543
ACM	1.957	1.629	1.513
Relative binding energies (kcal mol^{-1})			
VWN	0	12	33
BP	0	11	29
ACM	0	14	37

The present calculations indicate that the geometry of the simple chemisorbed species, such as H, on metal clusters is likely to be described well with the simplest DFT-approach – the local spin density approximation. Since all the methods confirm the 4-fold position to be the most stable for H-chemisorption, we can have confidence in using the more economical DFT calculations, such as in Dmol or FastStructure, to answer basic questions on the geometry of chemisorbed species or to locate preferred adsorption sites. Most likely, use of a finite cluster model of the periodic surface, the influence of which can be assessed by using larger clusters or slab models, can impact results more than that the improvement in the DFT Hamiltonian. These conclusions may not hold if more complicated molecules chemisorb on the surface and do, in general, apply only to geometries, not energetics, particularly when high accuracy is needed.

6.3. *Structures via first principles molecular dynamics*

The inclusion of dynamical aspects is made possible by the accurate calculation of the forces acting on each atom. This is particularly efficient in the framework of pseudopotentials and plane waves, but it is

also possible for all-electron localized basis set methods. This capability has opened the field of ab initio molecular dynamics.

Since the original paper by Car and Parrinello [176], progress on this frontier has been steady, but three obstacles have now become clear:

1. current ab initio calculations are limited to relatively small systems containing of the order of 100 atoms,
2. only short time intervals, of the order of several pico-seconds, can be simulated by ab initio MD within a reasonable computational effort, and
3. the accuracy in energies still requires improvements, especially in the calculation of reaction barriers.

The size limit is not a hard boundary and systems of over 1000 atoms have been tackled. Ab initio dynamics has even been carried out for unit cells of zeolites with methanol containing over 250 atoms [75], albeit with a considerable investment in supercomputer time. Progress in computer hardware and linearly scaling methods will extend this frontier continuously. The limitation of very short time intervals is perhaps more serious, and is often overlooked. The pico-second range of ab initio molecular dynamics methods compares with actual reaction events which are infrequent and which can be 10 orders of magnitude slower. Few chemical conversions are barrierless, and all processes that entail a significant activation barrier are poorly sampled by molecular dynamics. For this reason, our own applications of brute force molecular dynamics focus more on determining structure and sampling low energy conformations than on attempting, in an unconstrained manner, to simulate chemistry.

In order to gain information about possible structures and conformations, ab initio calculations are, and will remain, computationally demanding and time consuming, especially as one needs to explore the many degrees of freedom of a complex system such as a catalytically active surface. While quasi-classical methods based on empirical potentials work fairly well for organic molecular systems and for ionic compounds, it is hard to anticipate the possible errors of these methods, especially for systems where little experimental information is available. For these reasons, significant efforts have been dedicated to the development of methods which are rooted in rigorous quantum mechanics, but introduce a sequence of well-controlled approximations which allow one to find the optimal balance between computational speed and accuracy. A promising method is based on the Harris functional [177].

As an illustration of this approach, we present a simulated annealing study of the reconstruction of a C-terminated β-SiC(0 0 1) surface [178]. Silicon carbide offers an intriguing combination of mechanical, electronic, thermal and chemical properties. The compound is well known for its great hardness, its high thermal stability and melting point, and its chemical stability. SiC exists in a cubic α-form as well as a hexagonal β-form with over 170 polytypes with significantly varying electronic properties. For these reasons, SiC and its surfaces represent a fascinating topic for experimental and theoretical studies. Despite significant effort, many of the surface reconstructions of SiC are not known or understood.

The β-SiC(0 0 1) surface is modeled by a nine-layer slab which is two-dimensionally periodic in the form of a (2×2) supercell. The slab is repeated in the direction perpendicular to the surface thus forming a super-slab geometry.

The simulation is commenced from an ideal, unreconstructed surface which is created by truncating the bulk crystal (Fig. 8). Without bias, the simulated annealing procedure leads to a dimerization of the surface C atoms thereby creating a (2×1) structure (Fig. 8). In this calculation, only superposed spherical atomic densities are used to construct the effective potential and the charge density is not iterated to self-consistency [93,177]. Given the mixed covalent and ionic character of SiC, this is a rather crude approximation. It is surprising and reassuring to see that this model already captures the major features of the bonding mechanism, although the details of the geometric parameters deviate from self-consistent calculations. For example, the C–C bond distance is found to be about 1.20 Å compared with the value of 1.36 Å reported for accurate self-consistent pseudopotential calculations using Gaussian-type basis sets [179]. The Si–C back-bond between the surface and sub-surface atoms is obtained to be 1.94 Å compared with 1.86 Å from the self-consistent calculations. In the interior of the slab, the Harris functional gives a Si–C distance of 1.84 Å compared with the experimental value of

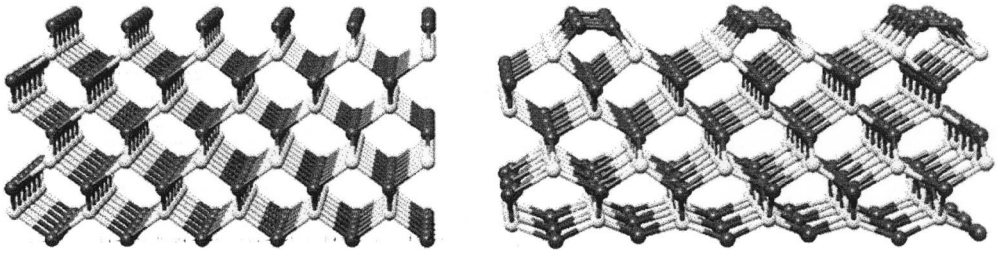

Fig. 8. Determination of the reconstruction of the C-terminated β-SiC(0 0 1) surface by the Harris functional methods implemented in the FastStructure program. The left panel shows the unreconstructed surface and the right panel displays a snapshot of the simulated annealing close to the final geometry.

1.883 Å. The non-self-consistent Harris functional approach gives the correct topology and deviations between 0.04–0.10 Å in bond distances compared with fully self-consistent density functional calculations. This demonstrates that simplified, but parameter-free quantum mechanical approaches based on the Harris functional can be used to identify the topology of low-energy structures. Further geometry optimizations using more accurate (and more time consuming) methods can be applied subsequently to refine the details of a structure and to obtain reliable energies.

7. Probing reactivity, reaction mechanisms, barriers and kinetics by density functional methods

Post Hartree–Fock methods that employ high levels of electron correlation yield the most accurate results for relative energies, for the reaction profiles and kinetics. They scale, however, as N^n, $n \geq 5$, and can be applied only to tiny models. Density functional methods offer useful accuracy within practical timeframes, and have been used to study a number of gas phase and surface reactions that are prototypical of reactions on heterogeneous catalysts. Development of analytic gradients and force constants accelerates the calculation of potential energy surfaces (PES) [164,166,180].

The adiabatic connection method can improve the accuracy of reaction barrier predictions significantly. The $OH+H_2 \rightarrow H_2O+H$ reaction, for example, exhibits a barrier only at the ACM level of calculations; other DFT approaches are clearly wrong in this specific case. The caliber of the ACM results depends on the basis set used, the use of very extensive basis sets lowering the ACM barrier for this specific reaction to about 1 kcal mol^{-1} which is now only in a qualitative agreement with experiment.

7.1. Using DFT to compute molecular reactivity

DFT is frequently used to predict qualitative properties of reactive species. Various properties, such as volumetric displays of electron densities, electrostatic potentials and Fukui functions, simulated optical spectra, etc. help in the characterization of reactive sites and species. For example, the evaluation and visualization of the electrostatic potentials in ketones [181] or Fukui functions in amines [182] predict the sites in the molecules susceptible to electrophilic attack.

A quantitative understanding of reaction mechanism requires calculation of the potential energy surface (PES) from which, by transition state theory (TST), rate constants can be calculated. Potentially, therefore, the efficiency of a catalytic process might be predicted, although applications to date have been for quite simple reactions, without, as far as we are aware, any direct computation of catalytic effect. Transition state theory permits the calculation of the PES to be focussed on the reactant, product and transition states, although it may be necessary to calculate the reaction path connecting those points, and properties of the PES in its vicinity. Fortunately, the absolute accuracies in the predicted structures and energies of reactant, product and transition states are less important than their relative positions on the PES.

The simplest, local functional (LSD) of DFT theory, is capable of reproducing, for example, rotational

barriers in hydrocarbons [166] or of identifying the correct ordering of CO adsorption sites on a Pd surface [183]. However, deficiencies of the LSD approximation are numerous, particularly when significantly different points on a PES are compared, or when weak intermolecular interactions are important. The problem seems to be more acute for chemistry, when systems with very different electron densities are compared, than for studies of the solid state, in which cases the electron density is more uniform and complies better with the prime LSD assumption of a homogeneous electron gas. The LSD model assumes that the exchange and correlation effects are independent of inhomogeneities in the electron density, and that they are short range; long-range dispersion forces are absent in DFT calculations [184].

A significant improvement in accuracy of DFT method to calculate PESS was achieved with the advent of gradient corrected (known as NLSD or GGA) DFT functionals [185–188]. The NLSD functional includes gradients of the density in computing the exchange and correlation terms, in addition to the density alone. The most popular versions of NLSD, such as BLYP [187,188] or BP [187,189] have contributed to many successes of the NLSD approach. For example, reactions involving 1,3-dipolar cycloadditions have unphysical, negative barrier heights at the LSD level, whereas NLSD results lead to barriers within 2 kcal mol^{-1} of accurate, yet incomparably more expensive MP4 calculations [190]. This and other work [191,192] indicates that the NLSD barriers are uniformly lower than those obtained from high level ab initio calculations, such as MP4, potentially reflecting an intrinsic inaccuracy of the DFT approach.

The energies of the optimized structures of the isomers and the transition states for various transformations of diazene [191,192] are compared in Table 3. Clearly, all the barriers for internal rotation, isomerization *trans–cis* and dissociation to H_2 and N_2 are much lower than the MP4 results. Calculations by Smith [193] using the G2 method, which appeared about the same time, confirmed the initial NLSD predictions. The G2 method [194] is considered the most accurate method for small molecules and it is capable of predicting the reaction energetics with an accuracy of ±2 kcal mol^{-1}. This is a case where the NLSD method performs well, in agreement with the best ab initio calculations. LSD results predict quite accurate relative energies of isomers, although the transition states are somewhat less accurate.

Table 3
Relative energies (in kcal mol^{-1}) for isomers and transition states (ts) of diazene

Molecule	VWN	BP	ACM	MP4	G2
H_2N=N	21	22	22	29	24
Isomerization ts	66	73	75	79	70
trans-HN=NH	0	0	0	0	0
Internal rotation ts	46	50	50	56	50
cis-HN=NH	5	5	5	6	5
Dissociation ts	–	72	79	82	–

DFT calculations performed using Dmol (VWN, BP) and Turbomole-DFT (ACM); MP4 and G2 calculations with Gaussian 92.

Studies on several transition metal complexes have confirmed the reliability of the ACM method in calculating structures, energies and transition states. Such systems can be considered prototypical of organometallic homogeneous catalysts and heterogeneous catalysts. Some of the systems investigated were: M^+, MH^+, MCH_3^+, MCH_2^+, $M(CO)_x$, where M spans all of the first-row transition metals (see, the references in [173]).

Although NLSD has been successfully overall in predicting the reactivity of many organic and inorganic processes, a growing number of failures have been reported over the last few years, indicating a need for more reliable functionals. For example, Raghavachari et al. [195] reported a complete reversal of the energy ordering (in excess of 100 kcal mol^{-1}) for the isomers of C_{20} between MP2 and NLSD. NLSD predicts barriers which are typically too low, and disappear completely for reactions with low barriers involving radicals. Neither LSD nor NLSD [196] were able to predict the existence of a barrier in the radical reaction $OH+H_2 \rightarrow H_2O+H$, for which a barrier of 3.0–4.6 kcal mol^{-1} is observed experimentally. LSD (VWN) yields a completely attractive curve, of −12 kcal mol^{-1} in the vicinity of the transition state calculated ab initio. NLSD results are not much better, the BLYP functional yielding a PES that is entirely attractive, whereas BP at the 6-31G* basis set level leads to a transition state with an insignificant barrier of about 0.1 kcal mol^{-1}. A calculation with a better basis set removes even this tiny barrier.

7.2. *Relative energies of different points on the potential energy surface*

A recent validation project has established the limits of the ACM method in calculating various points on a PES, including the relative energies of isomers, binding energies and transition states [168,173]. One hundred and eight small molecules comprised of H, first- and second-row atoms were calculated using several DFT functionals and various basis sets. From the energies of the molecules and corresponding atoms, the heats of more than 300 chemical reactions were computed. The reactions include: atomization, bond dissociation, hydrogenation, oxygenation, isomerization and isodesmic processes. Overall, the ACM method as implemented in Turbomole-DFT [168,172] is found to be superior for geometries and energies. The results are summarized in Table 4.

The Hartree–Fock method fails when bond breaking occurs or when different types of bonds occur in reactants and products. Only for the hydrogenation reactions are the results of an accuracy comparable with other methods. The VWN (LSD) method is inferior for all but isomerization and oxygenation processes, but LSD is notably better in describing the oxygenation reactions. Combustion processes cause notorious problems for quantum chemistry methods, including even the sophisticated G2 method [194]. Analysis of the reaction components, namely, O_2, CO_2 and H_2O, indicates that the atomization energy errors for these systems are both small and of comparable magnitude, 1–3 kcal mol^{-1}. However, the signs of the errors are such that they are additive and combustion reactions may exhibit a large error of some 25 kcal mol^{-1}. This is also the case for the following combustion reaction calculated at the ACM level.

$$CH_3{-}CH{=}CH_2 + \tfrac{9}{2}O_2 \rightarrow 3CO_2 + 3H_2O$$

These examples demonstrate that a uniform error in the PES energies is key to achieving accurate reaction energies. The MP2 method is less accurate than the NLSD and ACM methods for reactions involving bond breaking. The two NLSD approaches are of similar quality and, on average, exhibit larger errors than the ACM method. However, the ACM method is not an improvement over the NSLD methods for certain classes of reactions such as oxygenation, isomerization or isodesmic conversions.

Table 4
Mean absolute theory vs. experiment errors of heats of reactions obtained using a large uncontracted Dunning basis set (kcal mol^{-1})

Method	AT	BD	HY	OX	ISM	ISD
HF[a]	119.2	58.8	8.5	44.5	6.9	2.5
MP2[a]	22.0	8.8	7.0	11.2	5.1	3.0
VWN	56.4	23.7	8.5	11.7	5.5	4.7
BLYP	7.1	5.2	7.3	15.0	4.5	2.9
BP91	7.0	5.4	5.3	15.9	3.9	3.3
ACM	4.1	4.4	3.7	14.6	4.1	3.0

Calculations were performed with the Turbomole-DFT program.
[a] These calculations were performed using smaller basis set, 6-31G**.
AT: atomization; BD: bond dissociation; HY: hydrogenation; OX: oxygenation; ISM: isomerization; ISD: isodesmic.

8. Data for process models

Designing, selecting and optimizing a chemical process is a complicated endeavor involving a balance between the laws of nature, manufacturing costs, and the quality of the product produced. In order to screen the possible engineering and operational space, process simulation is employed [197]. Process simulation programs allow the chemical engineer to construct a virtual prototype of a chemical plant, assembling the engineering for the underlying chemical processes by connecting a sequence of unit operations that correspond, e.g., to distillation columns, extraction units, reactors, or mixers. In order to simulate the behavior of any one of these operations, thermodynamic, kinetic and mechanistic information is usually a necessary input. Such data has historically come from direct experimental measurements, from empirical correlations, or by estimation.

Such data is calculable via atomistic models. Given the cost of obtaining accurate experimental data, the limitations of correlative methods, and the ever improving price:performance ratio of computation, the use of molecular models to predict thermochemical and thermophysical properties is becoming steadily more attractive.

8.1. Thermochemistry

Thermochemistry is a quantitative measure of chemical reactivity. The thermochemical change associated with a chemical reaction is an important indicator of whether a reaction is favorable, a compound is hazardous, how much heat is released in a reaction, or the stability of a material. Due to the rapid development of novel chemical compounds and the experimental challenges associated with measuring their thermochemical data, theoretical estimates are becoming the first source of this important information for new materials. Typical thermochemical quantities such as ΔH and ΔG of reactions and formation can be estimated by both quantum mechanical and group contribution methods.

Semi-empirical, density functional and traditional ab initio levels of quantum mechanical theory have been used to predict thermochemical data. MOPAC, using the MNDO type Hamiltonians developed by Dewar and coworkers [198,199], is the most widely applied semi-empirical package. MOPAC is extremely fast at calculating heats for formation, ΔH_f, and can handle molecules comprised primarily of first and second row elements. While fast, its accuracy is variable and MOPAC is most useful for obtaining rough estimates quickly and for qualitative trends. For greater accuracy in thermochemical estimates, first principles quantum methods [175,194] must be employed.

As noted above, DFT has become a valuable route to thermochemical data (Fig. 9), with an accumulating applications portfolio. The ACM has been found [168] to be the most accurate DFT functional for predicting the enthalpy changes for a series of chemical reactions comprised of molecules containing first- and second-row elements (Table 4). The mean absolute deviations from experiment for 56 atomization reactions using the best basis set was 3.0 kcal mol^{-1} (Table 4), which is quite reasonable agreement considering the cost of the calculations. The standard Becke–Perdew (BP) functional using the same basis set on the same set of atomization reactions gave results with an absolute deviation of 5.2 kcal mol^{-1}.

First principles approaches beyond Hartree–Fock theory that include electron correlation effects (like MP2, CI, MCSCF, GVB and CC) are necessary for accurate energy calculations. These achieve higher accuracy in predicting thermochemical information, but at substantial computational expense relative to DFT methods. It is becoming popular to take specific

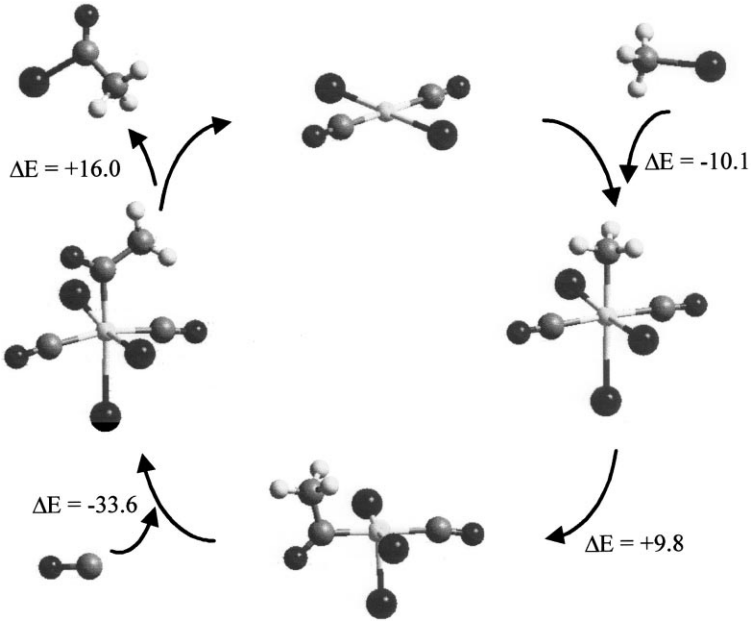

Fig. 9. Key steps involved in the catalytic conversion of CH_3I to $CH_3(CO)I$ by *trans*-$Rh(CO)_2I_2$, which is used in the commercial production of acetic acid from methanol (all activation energies given in kcal mol^{-1}).

combinations of these high level methods for geometry optimization, energy evaluation and frequency calculation. The G1 [175] and G2 [194] methods predict absolute errors for the same set of atomization reaction of 1.6 and 1.2 kcal mol^{-1}, respectively. However, the computational cost of these approaches limits them to small systems.

Empirical schemes have been used for many years to estimate thermochemical data. Benson's group additivity method assumes that ΔH_f can be estimated by summing the contributions from individual groups in a molecule. A group is defined as an atom or collection of atoms and the types of atoms bonded to it. This method is extremely fast to calculate and, in general, yields good results when suitable group data are available. Leveraging the value of high level, first principles calculations, several studies have reported Benson group parameters derived from G2 calculations rather than from experiment [200].

8.2. Kinetics

To estimate the rate of a reaction from first principles remains challenging. First, to estimate an overall rate constant, the mechanism and the full set of corresponding elementary reaction steps must be determined. Next, for the each significant elementary reaction step the transition state structure and energy must be determined, preferably using quantum mechanics. At this point there are essentially three levels of sophistication for predicting an actual rate constant. In basic transition state theory (TST) the rate constant, k, is defined by the following equation:

$$k = Q_{ts}/Q_{rs}\exp(-E_a/RT), \quad (1)$$

where E_a is the activation energy for the elementary step and Q_{ts} and Q_{rs} are the vibrational partition functions of the transition and reactant state species, respectively. This relatively simple theory provides a first order estimate of k; in many instances it is the only practical method. Even at this most basic level of theory it is evident that the rate constant is acutely sensitive to the value of E_a, which must therefore be estimated as accurately as possible. This level of transition state theory was used, for example, to estimate the rates of NO dissociation and recombination on cluster models of a copper surface (Fig. 3) [91,92].

More sophisticated extensions, like RRKM theory, include the effect of collisions between molecules in the gas phase. This method was used in a recent study of the reaction pathways for the decomposition of GeH_4 [201]. Typically, for reactions in which the potential energy surface is relatively flat near the transition state (reactions with long lived transition states) simple TST and RKKM theories begin to break down. Information about the nature of the transition state beyond its activation energy and vibrational frequencies must be included. Variational transition state theory accounts for higher order effects although, unfortunately, it is limited to fairly small molecules.

8.3. Thermophysical properties

Several approaches to calculating thermophysical properties are available, including group contribution methods, molecular based equation of state estimates, and molecular mechanics methods.

Force fields provide an empirically parameterized representation of the molecular potential energy surface of isolated molecules and molecules in condensed phases. In particular, Class II force fields have been developed with the intent of representing molecules well in all environments including the condensed phase, and of supporting computations of finite temperature behavior, such as equations of state, that are useful quantitatively [202,203]. Molecular dynamics calculations have been successful in calculating properties like solubility parameters, densities, and equation of state for liquids and polymers [204]. Monte Carlo methods have been used to directly calculate vapor–liquid and liquid–liquid phase equilibria using the Gibbs ensemble method, a technique with promise for being able to predict phase equilibria for a wide variety of systems, especially those for which experimental data are absent or which are difficult to access experimentally.

9. Other opportunity areas

9.1. A systems view

A working catalyst exemplifies well two parallel and general principles.

First, a significantly beneficial research result represents only one link in a long chain of steps that must be

taken towards ultimate commercial implementation in the field Fig. 9 (see, e.g., [205,206]).

Secondly, it illustrates the importance of a systems view. The performance of a given catalytic system represents the combination of a set of properties. A given property is dependent on the geometrical structure of the total system, and of its components, both at the atomic level (the ground state electronic structure being defined by the geometrical structure) and at longer length scales such as the mesoscopic. Structure is controlled by the conditions of synthesis and processing.

This systems view highlights the desirability, not only of detailing each of these separate facets of the system, but, particularly, of understanding the inter-facet links, how the synthesis and processing conditions affect the various structural attributes, how the individual properties are controlled by the structural details, how the combination of a defined set of properties governs the overall catalyst performance.

9.2. The challenge of synthesis

The reader can today design catalysts on the computer. As in de novo drug design, software that is readily available allows a user to construct, for example, viable new zeolite frameworks with specific compositions and pore architectures [10,11], to design particular non-framework cation configurations, and, in what is still today a small number of cases, predict the likely chemical properties of a defined structure. Molecular simulation can help identify the key discrete features necessary for activity and responsible for selectivity, and allow the effect of changes in these features to be assessed.

Although catalyst design can be approached by simulation, we are rarely able to implement such a design and a design, no matter how elegant, is, without a recipe for producing a material with those attributes, purely hypothetical. Our knowledge and level of control of synthesis is, in general, insufficient to permit this translation into practice. Bulk synthesis routes are diverse and include traditional 'heat-and-beat' solid-state chemistry, post synthesis modification by redox, by ion exchange, by adsorption or by loading with other promotional centers such as dispersed noble metals. Perhaps most opaque, at least at the molecular level, however, is the hydrothermal synthesis of zeolites and other mixed metal oxides [207]. Molecular level simulation, combined optimally with the latest experimental insights, is expected to be a major contributor to a growing understanding of the molecular mechanisms of hydrothermal synthesis.

9.3. Process improvement and troubleshooting

Often, the goal in a catalyst improvement study that is benefiting from molecular simulation is not complete replacement of a current system, but rather reduced cost, improved activity or selectivity, greater catalyst longevity, or better environmental compatibility. Mechanistic insight into deactivation yielded by molecular simulation might reasonably be expected to provide pointers to routes to control or circumvent catalyst deactivation by particular, identified mechanisms.

9.4. Accelerated experimental screening

The revolution in the pharmaceutical industry is fueled not only by the emergence of genomics and by the coming of age of molecular simulation, but also by new, parallel and high-throughput approaches to organic synthesis and screening. Combinatorial chemistry makes tractable the synthesis of large libraries of different compounds, by the introduction of many different possible substituents at each of several substitution points on a molecular scaffold.

Certain of these synthesis approaches and, particularly, high throughput parallel screening methods will have direct or indirect value in heterogeneous catalysis research. As in the organic domain, the value of high band-width experimental routes is enhanced dramatically by complementary computational techniques for library design and diversity assessment, for collating and interpreting, on a sound atomic-level foundation, synthesis and property evaluation results, and for identifying attractive target areas for a shotgun experimental chemistry approach.

9.5. Catalyst fabrication

Although the chemistry of heterogeneous catalysis is fundamentally a molecular level phenomenon, practical implementation typically requires that, for exam-

ple, catalyst formulation, support and matrix issues be considered. As yet little considered from a molecular simulation perspective, this area presents interesting questions and opportunities.

9.6. *A structured approach to bringing molecular simulation to bear*

Molecular simulation has been an already substantial contributor to several areas of heterogeneous catalyst research and development. To obtain best value from molecular simulation as a technology, however, demands a substantial level of expertise. Specifically, to apply molecular simulation to helping solve industrial problems requires that the appropriate methods be brought to bear on the appropriate issues, in the appropriate way. A general framework for this process has been described [114], based both on direct and anecdotal data. What is often key, is the initial analysis of the clearly stated research problem or business target [114]. The more complete this initial analysis, and the better phrased the questions and the corresponding follow-up items, the greater the likelihood of success in the overall project.

10. Conclusion

Combined with the other papers in this volume, even this quite cursory overview highlights that heterogeneous catalysts have already been the topic of a large number of molecular simulation studies. In one sense, however, we are still at an early stage of application. We are still accumulating fundamental information about active sites from various new and established experimental probes; many details are still sketchy. We are only now beginning to have the simulation tools needed to probe true catalytic chemistry; even today this technology is restricted to relatively simple, carefully chosen models and they are probably best applied by experienced practitioners.

Without any question, developments in methodology and application will continue to be rapid, steadily extending the complexity of systems that are tractable. The several areas considered here, amongst others, are well positioned to benefit from such developments. Notably, however, the molecular simulation technology that is already available can contribute substantially to a catalyst R&D program, if it is applied suitably to the appropriately phrased questions and combined with the relevant experimental approaches.

When, five years hence, we look back over the preceding period we will probably be at least as impressed by the many innovative and practical applications of current capabilities as we are by the development of new hardware and new software methods. Given the typical time delay that separates an industrial catalyst innovation from its discussion in the open literature, it is, indeed, quite likely that several such applications of molecular simulation are already underway or completed.

Acknowledgements

The MSI Catalysis Project is supported by a consortium of industrial, governmental and academic groups; we thank the various member representatives, including those of companion MSI consortia, which have furthered the simulation technologies mentioned here, and the scientific advisors for their ongoing input, ideas, guidance and advice.

References

[1] J.M. Newsam, C.M. Freeman, S.M. Levine, M.W. Deem, in: M. Doyama, J. Kihara, M. Tanaka, R. Yamamoto (Eds.), Computer Aided Innovation of New Materials II, Elsevier, Amsterdam, 1993, pp. 1077–1081.
[2] J.M. Newsam, J. Comput. Aided Mater. Design, (1999) in preparation.
[3] K.T. No, H. Chon, T. Ree, M.S. Jhon, J. Phys. Chem. 85 (1981) 2065–2070.
[4] K.T. No, B.H. Seo, J.M. Park, M.S. Jhon, J. Phys. Chem. 92 (1988) 6783–6789.
[5] C.R. A. Catlow (Ed.), Modeling of Structure and Reactivity in Zeolites, Academic Press, London, 1992.
[6] J. Sauer, in: J. Weitkamp, H.G. Karge, H. Pfeifer, W. Hölderich (Eds.), Zeolites and related microporous materials: state of the Art 1994, Stud. Surf. Sci. Catal. 84 (Part C) (1994) 2039–2057.
[7] J. Sauer, P. Ugliengo, E. Garrone, V.R. Saunders, Chem. Rev. 94 (1994) 2095–2160.
[8] H. Chon, S.-K. Ihm, Y.S. Uh (Eds.), Progress in zeolite and microporous materials, Stud. Surf. Sci. Catal. 105 (Parts A, B and C) (1997).
[9] M.W. Deem, J.M. Newsam, Nature 342 (1989) 260–262.
[10] M.W. Deem, J.M. Newsam, J. Am. Chem. Soc. 114 (1992) 7189–7198.

[11] J.M. Newsam, C.M. Freeman, C.M. Kölmel, in: W.I.F. David (Ed.), Structure Determination from Powder Diffraction Data, Oxford University Press, Oxford, 1999, in preparation.
[12] J.M. Newsam, in: R. von Ballmoos, J.B. Higgins, M.M.J. Treacy (Eds.), Proceedings from the Ninth International Zeolite Conference, vol. I, Butterworth-Heinemann, Stoneham, MA, 1993, pp. 127–141.
[13] S.D. Loades, S.W. Carr, D. Gay, A. Rohl, J. Chem. Soc., Chem. Commun. (1994) 1369–1370.
[14] J.G. Fripiat, F. Berger-Andre, J.-M. Andre, E.G. Derouane, Zeolites 3 (1983) 306.
[15] K.-P. Schröder, J. Sauer, M. Leslie, C.R.A. Catlow, Zeolites 12 (1992) 20.
[16] E. de Vos Burchart, H. van Bekkum, B. van de Graaf, Collect. Czech. Chem. Commun. 57 (1992) 681.
[17] G.J. Kramer, R.A. van Santen, J. Am. Chem. Soc. 115 (1993) 2887.
[18] A. Redondo, P.J. Hay, J. Phys. Chem. 97 (1993) 11754.
[19] G. Ricchiardi, J.M. Newsam, J. Phys. Chem. 101 (1997) 9943–9950.
[20] C.M. Soukoulis, J. Phys. Chem. 88 (1984) 4898–4901.
[21] C.P. Herrero, J. Phys. Chem. 95 (1991) 3282–3288.
[22] C.P. Herrero, Chem. Phys. Lett. 183 (1991) 199–203.
[23] C.P. Herrero, J. Phys. Chem. 97 (1993) 3338–3343.
[24] M.C. Gordillo, C.P. Herrero, Chem. Phys. 211 (1996) 81–90.
[25] M.T. Melchior, D.E.W. Vaughan, K.G. Strohmaier, J. Phys. Chem. 99 (1995) 6128–6144.
[26] V. Moravetski, J.-R. Hill, U. Eichler, A.K. Cheetham, J. Sauer, J. Am. Chem. Soc. 118 (1996) 13015.
[27] B. Bussemer, K.-P. Schröder, J. Sauer, Solid State NMR (1997), in press.
[28] J.M. Newsam, C.M. Freeman, A.M. Gorman, B. Vessal, J. Chem. Soc., Chem. Commun. (1996) 1945–1946.
[29] J. Lignières, J.M. Newsam, Microporous and Mesoporous Mater. (1999), in print.
[30] A.M. Gorman, C.M. Freeman, C.M. Kölmel, J.M. Newsam, Faraday Discus. 106 (1997) 489–494.
[31] G. Vitale, C.F. Mellot, L.M. Bull, A.K. Cheetham, J. Phys. Chem. 101 (1997) 4559–4564.
[32] M.E. Grillo, J. Carrazza, J. Phys. Chem. 100 (1996) 12261–12265.
[33] M.W. Deem, J.M. Newsam, J.A. Creighton, J. Am. Chem. Soc. 114 (1992) 7198–7207.
[34] K.S. Smirnov, D. Bougeard, Zeolites 14 (1994) 203–207.
[35] P. Tschaufeser, S.C. Parker, J. Phys. Chem. 99 (1995) 10609–10615.
[36] G.W. Watson, P. Tschaufeser, A. Wall, R.A. Jackson, S.C. Parker, in: C.R.A. Catlow (Ed.), Computer Modeling in Inorganic Crystallography, Academic Press, London, 1997, pp. 55–81.
[37] J.M. Newsam, J. Phys. Chem. 93 (1989) 7689–7694.
[38] C.M. Freeman, C.R.A. Catlow, J.M. Thomas, S. Brode, Chem. Phys. Lett. 186 (1991) 137–142.
[39] J.L. Soto, A.L. Myers, Mol. Phys. 42 (1981) 971–983.
[40] K. Fiedler, B. Grauert, Adsorpt. Sci. Technol. 3 (1986) 181–187.
[41] H. Kono, A. Takasaka, J. Phys. Chem. 91 (1987) 4044–4055.
[42] P. Broier, A.V. Kiselev, E.A. Lesnik, A.A. Lopatkin, Zh. Fiz. Khim. 42 (1968) 2556–2562.
[43] A.V. Kiselev, J. Chem. Technol. Biotechnol. 29 (1979) 673–685.
[44] G.B. Woods, A.Z. Panagiotopoulos, J.S. Rowlinson, Mol. Phys. 63 (1988) 49–63.
[45] D.M. Razmus, C.K. Hall, AIChE J. 37 (1990) 769–779.
[46] P. Demontis, G.B. Suffritti, A. Alberti, S. Quartieri, E.S. Fois, A. Gamba, Gazz. Chim. Ital. 116 (1986) 459–466.
[47] E. Cohen de Lara, R. Kahn, A.M. Goulay, J. Chem. Phys. 90 (1989) 7482–7491.
[48] S. Yashonath, P. Demontis, M.L. Klein, Chem. Phys. Lett. 153 (1988) 551–556.
[49] B. Smit, T.L.M. Maesen, Nature 374 (1995) 42–44.
[50] R.L. June, A.T. Bell, D.N. Theodorou, J. Phys. Chem. 95 (1991) 8866–8878.
[51] R.Q. Snurr, A.T. Bell, D.N. Theodorou, J. Phys. Chem. 98 (1994) 11948–11961.
[52] E.J. Maginn, A.T. Bell, D.N. Theodorou, J. Phys. Chem. 100 (1997) 7155–7173.
[53] S.M. Auerbach, H.I. Metiu, J. Chem. Phys. 105 (1996) 3753–3760.
[54] J.A. Horsley, J.D. Fellmann, E.G. Derouane, C.M. Freeman, J. Catal. 147 (1994) 231–240.
[55] S. Feast, D. Bethell, G.J. Hutchings, P.C.B. Page, S.P. Saberi, F. King, C.H. Rochester, in: J. Weitkamp, H.G. Karge, H. Pfeifer, W. Hölderich (Eds.), Zeolites and Related Microporous Materials: State of the Art 1994, Stud. Surf. Sci. Catal. 84 (Part C) (1994) 1611–1616.
[56] W.F. Hölderich, M.E. Paczkowski, in: Proceedings of the Second International Conference on Microporous Crystals, Elsevier, Amsterdam, 1993.
[57] F. Blanco, G. Urbina-Villalba, M.M. Ramírez de Argudelo, in: J. Weitkamp, H.G. Karge, H. Pfeifer, W. Hölderich (Eds.), Zeolites and Related Microporous Materials: State of the Art 1994, Stud. Surf. Sci. Catal. 84 (Part C) (1994) 2155–2162.
[58] M. Hochgräfe, H. Gies, Extended Abstracts, Proceedings of the German Crystallographic Association Meeting, Bochum, Germany, 1992.
[59] R.G. Bell, D.W. Lewis, P. Voigt, C.M. Freeman, J.M. Thomas, C.R.A. Catlow, in: J. Weitkamp, H.G. Karge, H. Pfeifer, W. Hölderich (Eds.), Zeolites and Related Microporous Materials: State of the Art 1994, Stud. Surf. Sci. Catal. 84 (Part C) (1994) 2075–2082.
[60] D.W. Lewis, C.M. Freeman, C.R.A. Catlow, J. Phys. Chem. 99 (1995) 11201.
[61] C.M. Kölmel, Y.S. Li, C.M. Freeman, S.M. Levine, M.-J. Hwang, J. Maple, J.M. Newsam, J. Phys. Chem. 98 (1994) 12911–12918.
[62] T.V. Harris, S.I. Zones, in: J. Weitkamp, H.G. Karge, H. Pfeifer, W. Hölderich (Eds.), Zeolites and Related Microporous Materials: State of the Art 1994, Stud. Surf. Sci. Catal. 84 (Part A) (1994) 29–36.
[63] A.P. Stevens, A.M. Gorman, C.M. Freeman, P.A. Cox, J. Chem. Soc. Faraday Trans. 92 (1996) 2065–2073.

[64] D.W. Lewis, C.R.A. Catlow, J.M. Thomas, Faraday Discus. 106 (1996), in press.
[65] D.W. Lewis, D.J. Willock, C.R.A. Catlow, J.M. Thomas, G.J. Hutchings, Nature 382 (1996) 604–606.
[66] U. Eichler, C.M. Kölmel, J. Sauer, J. Comput. Chem. 18 (1997) 463–477.
[67] U. Eichler, M. Brändle, J. Sauer, J. Phys. Chem. B 101 (1997) 10035–10050.
[68] J.B. Nicholas, A.C. Hess, J. Am. Chem. Soc. 116 (1994) 5428.
[69] R. Shah, M.C. Payne, M.-H. Lee, J.D. Gale, Science 271 (1996) 1395–1397.
[70] L. Campana, A. Selloni, J. Weber, A. Pasquarello, I. Papai, A. Goursot, Chem. Phys. Lett. 226 (1994) 245–250.
[71] E. Nusterer, P. Blöchl, K. Schwarz, Angewandte Chemie Int. Ed. 35 (1996) 175.
[72] E. Nusterer, P.E. Blöchl, K. Schwarz, Chem. Phys. Lett. 253 (1996) 448.
[73] F. Haase, J. Sauer, J. Hutter, Chem. Phys. Lett. 266 (1997) 397.
[74] K. Schwarz, E. Nusterer, P.E. Blöchl, Chem. Phys. Lett. 253 (1996) 448.
[75] I. Stich, J.D. Gale, K. Terakura, M.C. Payne, (1997), submitted.
[76] J.M. Newsam, Science 231 (1986) 1093–1099.
[77] E. Wimmer, Science 269 (1995) 1397–1398.
[78] E.J. Wimmer, J. Comput. Aided Mater. Design 1 (1993) 215.
[79] M.C. Payne, M.P. Teter, D.C. Allan, T.A. Arias, J.D. Joannopoulos, Rev. Modern Phys. 64 (1992) 1045.
[80] P.A. Gravil, D.M. Bird, J.A. White, Phys. Rev. Lett. 77 (1996) 3933–3936.
[81] D.M. Bird, P.A. Gravil, Surf. Sci. 377 378 379 (1997) 555–562.
[82] Q. Ge, P. Hu, D.A. King, M.-H. Lee, J.A. White, M.C. Payne, J. Chem. Phys. 106 (1997) 1210–1215.
[83] B. Hammer, M. Scheffler, K.W. Jacobsen, J.K. Norskov, Phys. Rev. Lett. 73 (1994) 1400.
[84] J.A. White, D.M. Bird, M.C. Payne, Phys. Rev. B 53 (1996) 1667–1674.
[85] B. Hammer, Y. Morikawa, J.K. Norskov, Phys. Rev. Lett. 76 (1996) 2141.
[86] P. Hu, D.A. King, S. Crampin, M.H. Lee, M.C. Payne, Chem. Phys. Lett. 230 (1994) 501–506.
[87] P. Hu, D.A. King, M.H. Lee, M.C. Payne, Chem. Phys. Lett. 246 (1995) 73–78.
[88] Y.S. Li, J.M. Newsam, Materials theory and modelling, in: J. Broughton, P. Bristowe, J. Newsam (Eds.), Proceedings of the Materials Research Symposium, vol. 291, Materials Research Society, Pittsburgh, PA, 1993, pp. 573–578.
[89] S. Veliah, R. Pandey, Y.S. Li, J.M. Newsam, B. Vessal, Chem. Phys. Lett. 235 (1995) 53–57.
[90] Y.S. Li, J.M. Newsam, in: M. Doyama, J. Kihara, M. Tanaka, R. Yamamoto (Eds.), Computer Aided Innovation of New Materials II, Elsevier, Amsterdam, 1993, pp. 1043–1046.
[91] M.A. van Daelen, Y.S. Li, J.M. Newsam, R.A. van Santen, Chem. Phys. Lett. 226 (1994) 100–105.
[92] M.A. van Daelen, Y.S. Li, J.M. Newsam, R.A. van Santen, J. Phys. Chem. 100 (1996) 2279–2289.
[93] X.-P. Li, J. Andzelm, J. Harris, A.M. Chaka, in: B.B. Laird, R.B. Ross, T. Ziegler (Eds.), Chemical Applications of Density-Functional Theory, ACS Symp. Ser. 629, American Chemical Society, Washington, DC, 1996, pp. 388–401.
[94] A.M. Chaka, J. Harris, X.-P. Li, Rev. L'Institut Franc. Petrole 51 (1996) 171–181.
[95] A.M. Chaka, J. Harris, X.-P. Li, in: B.B. Laird, R.B. Ross, T. Ziegler (Eds.), Chemical Applications of Density-Functional Theory, ACS Symp. Ser. 629, American Chemical Society, Washington, DC, 1996, pp. 368–387.
[96] M.S. Daw, M.I. Baskes, Phys. Rev. B 29 (1984) 6443.
[97] J.R. Smith, T. Perry, A. Banerjea, J. Ferrante, G. Bozzolo, Phys. Rev. B 44 (1991) 6444.
[98] K.W. Jacobsen, P. Stoltze, J.K. Norskov, Surf. Sci. 366 (1996) 394.
[99] A.T. Paxton, A.P. Sutton, Acta Metall. 37 (1989) 1693–1715.
[100] F. Ruette, A.J. Hernandez, in: S. Fraga (Ed.), Computational Chemistry: Structure, Interactions and Reactivity, Stud. Phys. Theo. Chem. 77B (1991) 637–661.
[101] H. Sellers, Surf. Sci. 310 (1994) 281–291.
[102] H. Sellers, E. Shustorovich, Surf. Sci. 346 (1996) 322–336.
[103] H.C. Kang, W.H. Weinberg, Acc. Chem. Res. 25 (1992) 253–259.
[104] W.H. Weinberg, Langmuir 9 (1993) 655–662.
[105] E. Kirschner, Chem. Eng. News 11 (1995) 13–20.
[106] G. McWilliams, Business Week 23 (1996) 82.
[107] R. Fusco, L. Longo, F. Masi, F. Garbassi, Macromolecules 30 (1997) 7673–7685.
[108] R. Fusco, L. Longo, F. Masi, F. Garbassi, Macromolecules Rapid Commun. 18 (1997) 433–441.
[109] F.U. Axe, J.M. Coffin, J. Phys. Chem. 98 (1994) 2567–2570.
[110] H. Munakata, Y. Ebisawa, Y. Takashima, M.C. Wrinn, A.C. Scheiner, J.M. Newsam, Catal. Today 23 (1995) 403–408.
[111] P. Margl, K. Schwarz, P.E. Blöchl, J. Chem. Phys. 100 (1994) 8194.
[112] S. Iarlori, F. Buda, R.J. Meier, G.H. Van-DoreMaele, J. Mol. Phys. 87 (1996) 801.
[113] T.K. Woo, P.M. Margl, J.C.W. Lohrenz, P.E. Blöchl, T. Ziegler, J. Am. Chem. Soc. 118 (1996) 13021–13030.
[114] J.M. Newsam, (1999), in preparation.
[115] C. Hansch, A. Leo, D. Hoekman, Exploring QSAR – Hydrophobic, Electronic, and Steric Constants, American Chemical Society, Washington, DC, 1995.
[116] D. Livingstone, Data Analysis for Chemists. Applications to QSAR and Chemical Product Design, Oxford University Press, Oxford, 1995.
[117] K.B. Lipkowitz, D.B. Boyd (Eds.), Reviews in Computational Chemistry, vol. 9, VCH Publishers, New York, 1996.
[118] K.B. Lipkowitz, D.B. Boyd (Eds.), Reviews in Computational Chemistry, vol. 10, VCH Publishers, New York, 1996.
[119] D. Rogers, Advances in Neural Processing Systems vol. 4, Morgan Kaufmann, San Mateo, 1992.

[120] D. Rogers, A.J. Hopfinger, J. Chem. Inf. Comput. Sci. 34 (1994) 854.
[121] J. Biscerano, Prediction of Polymer Properties, Marcel Decker, New York, 1993.
[122] J. Devillers (Ed.), Genetic Algorithms in Molecular Modeling (Principles of QSAR and Drug Design), Academic Press, London, 1996.
[123] J. Devillers (Ed.), Neural Networks in QSAR and Drug Design (Principles of QSAR and Drug Design), Academic Press, London, 1996.
[124] K. Hornik, M. Stinchcombe, H. White, Neural Networks 2 (1989) 359.
[125] A.J. Hopfinger, B.J. Burke, in: M.A. Johnson, G.A. Maggiora (Eds.), Concepts and Applications of Molecular Similarity, Wiley, New York, 1990, p. 173.
[126] M.I. Koch, D.C. Chin, R.H. Smith, in: Proceedings of the Seventh National Conference on Light Rail Transit, vol. 2. National Academy of Sciences, 1997, pp. 126–131.
[127] J.C. Spall, J.A. Cristion, IEEE Trans. Systems Man and Cybernet. B 27 (1997) 369–375.
[128] T. Hattori, S. Kito, Catal. Today 23 (1995) 347–355.
[129] H.M.G. Smets, W.F.L. Bogaerts, Mater. Sel. Des. (1992) 64.
[130] L. Xi, Miner. Met. Mater. Soc. 511 (1992) 511.
[131] B. Joseph, F.W. Hanratty, J. Kardos, J. Compos. Mater. 29 (1995) 100.
[132] M.G. Koehler, A.J. Hopfinger, Polymer 30 (1989) 116.
[133] T.B. Blank, S.D. Brown, A.W. Calhoun, D.J. Doren, J. Chem. Phys. (1997), in press.
[134] H. van de Waterbeemd (Ed.), Methods and Principles in Medicinal Chemistry, vol 2: Chemometric Methods in Molecular Design, VCH Publishers, Weinheim, 1995.
[135] E. Robb, M. Munk, Mikrochim. Acta 1 (1990) 131.
[136] Z. Hippe, R. Hippe, Appl. Spectrosc. Rev. 16 (1980) 135.
[137] R. Lindsay, B.G. Buchanan, E.A. Fiegenbaum, J. Lederberg, Dendral, New York, 1980.
[138] E.O.T. Sales, P.A.D. Junior, V.K. Geag, Radioanal. Nucl. Chem. 190 (1995) 439.
[139] J. Bojarski, L. Ekiert, Anal. Chem. Symp. Ser. 16 (1984) 35.
[140] A. Sabljic, J. Chromatogr. 314 (1984) 1.
[141] J. Chang, M. Lewis, Acta Crystallogr. Sect. D 50 (1994) 667.
[142] Y. Xiao, D.E. Williams, Comput. Chem. 18 (1994) 199.
[143] T. Dandekar, P. Argos, Protein Eng. 5 (1992) 637.
[144] G. Schneider, P. Wrede, Biophys. J. 66 (1994) 335.
[145] D. Rogers, Molecular Simulations Inc., QSAR Workshop, Trevose PA, and private communication (1997).
[146] D.E.W. Vaughan, K.G. Strohmaier, M.M.J. Treacy, J.M. Newsam, US Patent no. 5 116 590 (1992).
[147] M.W. Anderson, O. Terasaki, T. Ohsuna, A. Philippou, S.P. Mackay, A. Ferriera, A. Rocha, S. Lidin, Nature 367 (1994) 347–351.
[148] J.M. Newsam, M.M.J. Treacy, W.T. Koetsier, C.B. deGruyter, Proc. R. Soc. London A 420 (1988) 375–405.
[149] J.B. Higgins, R.B. LaPierre, J.L. Schlenker, A.C. Rohrman, J.D. Wood, G.T. Kerr, W.J. Rohrbaugh, Zeolites 8 (1988) 446–452.
[150] F.H. Allen, O. Kennard, Chem. Design Automat. News 8 (1993) 31–37.
[151] Cerius2 and InsightII crystallographic modeling modules, User Manuals (see data provided at http://www.msi.com), 1997.
[152] D.E. Akporiaye, H. Fjellvåg, E.N. Halvorsen, J. Hustveit, A. Karlsson, K.P. Lillerud, J. Chem. Soc., Chem. Commun. (1996) 601–602.
[153] D.E. Akporiaye, H. Fjellvåg, E.N. Halvorsen, T. Haug, A. Karlsson, K.P. Lillerud, J. Chem. Soc., Chem. Commun. (1996) 1553–1554.
[154] G.D. Price, I.G. Wood, D.E. Akporiaye, in: C.R.A. Catlow (Ed.), Modelling of Structure and Reactivity in Zeolites, Academic Press, London, 1992, pp. 19–42.
[155] D.A. Keen, R.L. McGreevy, Nature 344 (1990) 423–425.
[156] M.B. Boisen, G.V. Gibbs, M.S.T. Mukowinski, Phys. Chem. Miner. 21 (1994) 269.
[157] C.M. Freeman, J.M. Newsam, S.M. Levine, C.R.A. Catlow, J. Mater. Chem. 3 (1993) 531–535.
[158] C.M. Freeman, C.R.A. Catlow, J. Chem. Soc., Chem. Commun. (1991) 89–91.
[159] B. Vessal, G.N. Greaves, P.T. Marten, A.V. Chadwick, R. Mole, S. Houde-Walter, Nature 356 (1992) 504–506.
[160] S.C. Parker, P.J. Lawrence, C.M. Freeman, S.M. Levine, J.M. Newsam, Catal. Lett. 15 (1992) 123–131.
[161] S.M. Black, J. Chem. Soc., Faraday Trans. II 87 (1991) 3409.
[162] D.S. Santilli, T.V. Harris, S.I. Zones, Microporous Mater. 1 (1993) 329–341.
[163] E. Wimmer, A.J. Freeman, C.-L. Fu, P.-L. Cao, S.-H. Chou, B. Delley, in: K.F. Jensen, D.G. Truhlar (Eds.), ACS Symp. Ser. 353 (1987) 49–68.
[164] B. Delley, J. Chem. Phys. 92 (1990) 508.
[165] DMol3 User Manual, Molecular Simulations Inc., San Diego, CA, 1997.
[166] J. Andzelm, E. Wimmer, J. Chem. Phys. 96 (1992) 1280.
[167] D.R. Salahub, R. Fournier, P. Mlynarski, I. Papai, A. St. Amant, J. Ushio, in: J. Labanowski, J. Andzelm (Eds.), Density Functional Methods in Chemistry, Springer, New York, 1991, p. 77.
[168] A.C. Scheiner, J. Baker, J.W. Andzelm, J. Comput. Chem. 18 (1997) 775.
[169] M.J. Frisch, G.W. Trucks, H.B. Schlegel, P.M.W. Gill, B.G. Johnson, M.W. Wong, J.B. Foresman, M.A. Robb, M. Head-Gordon, E.S. Replogle, R. Gomperts, J.L. Andres, K. Raghavachari, J.S. Binkley, C. Gonzalez, R.L. Martin, D.J. Fox, D.J. Defrees, J. Baker, J.J.P. Stewart, J.A. Pople, Gaussian 92/DFT Gaussian Inc., Pittsburgh, PA, 1990.
[170] A.R. Williams, J. Kubler, J.R. Gelatt, Phys. Rev. B 19 (1979) 609.
[171] E. Wimmer, H. Krakauer, A.J. Freeman, in: P.N. Hawkes (Ed.), Advances in Electronics and Electron Physics, vol. 65, Academic Press, Orlando, FL, 1985, 357 pp.
[172] R. Ahlrichs, M. Bär, E. Ehrig, M. Häser, H. Horn, C. Kölmel, Turbomole User Manual, Molecular Simulations Inc., San Diego, CA, 1995.
[173] J. Baker, M. Muir, J. Andzelm, A. Scheiner, in: B.B. Laird, R.B. Ross, T. Ziegler (Eds.), Chemical Applications of

Density Functional Theory, ACS Symp. Ser. 629 (1996) 342.
[174] A.D. Becke, J. Chem. Phys. 98 (1993) 5648.
[175] J.A. Pople, M. Head-Gordon, D. Fox, K. Raghavachari, L.A. Curtiss, J. Chem. Phys. 90 (1989) 5622.
[176] R. Car, M. Parrinello, Phys. Rev. Lett. 55 (1985) 2471.
[177] J. Harris, Phys. Rev. B 31 (1985) 1770.
[178] L. Douillard, P. Soukiassian, E. Wimmer, 1997, in preparation.
[179] M. Sabisch, P. Krüger, A. Mazur, M. Rohlfing, J. Pollmann, Phys. Rev. B 53 (1996) 13121.
[180] A. Komornicki, G. Fitzgerald, J. Chem. Phys. 98 (1995) 1398.
[181] A.T. Pudzianowski, J.C. Barish, S.H. Spergel, Tetrahedron Lett. 33 (1992) 293.
[182] K. Flurchik, L. Bartolotti, J. Mol. Graphics 13 (1995) 10.
[183] J. Andzelm, D.R. Salahub, Int. J. Quant. Chem. 29 (1986) 1091.
[184] S. Kristyan, P. Pulay, Chem. Phys. Lett. 229 (1994) 175.
[185] T. Ziegler, Chem. Rev. 91 (1991) 651.
[186] J. Labanowski, J. Andzelm (Eds.), Density Functional Methods in Chemistry, Springer, New York, 1991.
[187] A.D. Becke, Phys. Rev. A 38 (1988) 3098.
[188] C. Lee, W. Yang, R.G. Parr, Phys. Rev. B 37 (1988) 785.
[189] J.P. Perdew, Phys. Rev. B 33 (1986) 8822.
[190] C. Sosa, J. Andzelm, C. Lee, J.F. Blake, B.L. Chenard, T.W. Butler, Int. J. Quant. Chem. 49 (1994) 511.
[191] J. Andzelm, C. Sosa, R.A. Eades, J. Phys. Chem. 97 (1993) 4664.
[192] J. Andzelm, J. Baker, A. Scheiner, M. Wrinn, Int. J. Quant. Chem. 56 (1995) 733.
[193] B. Smith, J. Phys. Chem. 97 (1993) 10513.
[194] L.A. Curtiss, C. Jones, G.W. Trucks, K. Raghavachari, J.A. Pople, J. Chem. Phys. 94 (1991) 7221.
[195] K. Raghavachari, D.L. Strout, G.K. Odom, G.E. Scuseria, J.A. Pople, B.G. Johnson, P.M.W. Gill, Chem. Phys. Lett. 214 (1993) 357.
[196] J. Baker, J. Andzelm, M. Muir, P.R. Taylor, Chem. Phys. Lett. 237 (1995) 53.
[197] B.C. Gates, Catalytic Chemistry, Wiley, New York, 1992.
[198] J.J.P. Stewart, Mopac Manual, FJSRL-TR-88-007, Frank J. Seiler Research Laboratory, US Air Force Academy, 1988.
[199] J.J.P. Stewart, J. Comp. Aided Mol. Design 4 (1990) 1.
[200] B.T. Colegrove, T.B. Thompson, J. Chem. Phys. 106 (1997) 1480.
[201] M. Hierlemann, H. Simka, M. Utz, K.F. Jensen, 1997, submitted.
[202] B.E. Eichinger, D.R. Rigby, M.H. Muir, Comput. Polym. Sci. 5 (1995) 147.
[203] H. Sun, D.R. Rigby, Spectrochim. Acta (1997), in press.
[204] M.P. Allen, D.J. Tildesley, Computer Simulation of Liquids, Oxford University Press, Oxford, 1987.
[205] J.M. Newsam, Y.S. Li, Catal. Today 23 (1995) 325–332.
[206] G.S. Yablonskii, in: A. Sydow (Ed.), Proceedings of the 15th IMACS World Congress on Scientific Computation, Modelling and Applied Mathematics. Vol. 6: Application in Modelling and Simulation, Verlag, Berlin, 1997, pp. 305–310.
[207] H.S. Horowitz, A.J. Jacobson, J.M. Newsam, J.T. Lewandowski, M.E. Leonowicz, Solid State Ionics 32 33 (1989) 678–690.