

Estimation of Chemical Reactivity Parameters and Physical Properties of Organic Molecules Using SPARC

S. H. Hilal^a, L. A. Carreira^b and S. W. Karickhoff^a

^aEnvironmental Research Laboratory, U.S. Environmental Protection Agency,
Athens, GA 30605, U.S.A.

^bDepartment of Chemistry, University of Georgia, Athens, GA 30602, U.S.A.

1. GENERAL INTRODUCTION

The major differences between behavior profiles of molecules in the environment are attributable to their physicochemical properties. The need for physical and chemical constants of chemical compounds has greatly accelerated both in industry and government as assessments are made of pollutant exposure and risk. Although considerable progress has been made in process elucidation and modeling for chemical and physical processes, values for the fundamental thermodynamic and physicochemical properties (i.e., rate/equilibrium constants, Henry's law constant, distribution coefficients between immiscible solvents, solubility in water, etc.) have been achieved for only a small number of molecular structures.

For most chemicals, only fragmentary knowledge exists about those properties that determine each compound's fate in the environment. Chemical-by-chemical measurements of the required properties is not practical because of expense and because trained technicians and adequate facilities are not available for measurement efforts involving thousands of chemicals. In fact physical and chemical properties have actually been measured for, perhaps, 1 percent of the approximately 70,000 industrial chemicals listed by the U.S. Environmental Protection Agency's Office of Toxic Substances (OTS) [1].

Fortunately, estimation techniques that employ the judgment of expert chemists are available to provide the required data in a cost-effective manner. These techniques include the application of linear free energy relationships (LFER) [2,3], structure activity relationships (SAR) [4,5], and other estimation methods particularly in the drug and environmental fields. Even so, methodologies and values often are not available for the parameters needed in the sophisticated mathematical models used for environmental exposure assessment.

Recently [6-10], we described our approach for predicting numerous physical properties and chemical reactivity parameters of organic compounds strictly from molecular structure using the new prototype computer program SPARC (SPARC Performs Automated Reasoning in Chemistry). The goal of this new computer program is to apply the reasoning process that an organic chemist might undertake in reactivity analysis. The approach primarily involves deductive reasoning and is theory/mechanism oriented. The computational approach is based on existing mathematical models of chemistry.

Our new computer program costs the user only a few minutes of computer time and provides greater accuracy and a broader scope than is possible with conventional estimation techniques. The user needs to know only the molecular structure for the compound to predict a property of interest. The user provides the program with the molecular structure either by direct entry as SMILES (Simplified Molecular Input Line Entry System) notation or via the molecular editor that will generate the structure and translate it to SMILES notation. SPARC is programmed with the ALS (Applied Logic Systems) version of Prolog (PROgramming in LOGic). It is executable on machines with the UNIX operating system or 386/486 MS DOS machines with 16 MB of extended memory.

SPARC presently predicts ionization pK_a , electron affinity, and numerous physical properties such as vapor pressure (at any temperature), boiling point (at any pressure), activity coefficient/distribution coefficient (for any solvent), retention times for gas and liquid chromatography, etc. In this chapter we report the calculation of ionization equilibrium constants, electron affinity, and gas chromatography retention indices for a squalane liquid phase strictly from molecular structure for a large number of organic compounds using SPARC.

1.1. SPARC Computational Approach

SPARC does not do "first principles" computation; rather, it analyzes chemical structure relative to a specific reactivity query much as an expert chemist might. SPARC utilizes directly the extensive knowledge base of organic chemistry. For physical properties, intermolecular interactions are expressed as a summation over all the interaction forces between molecules (i.e., dispersion, induction, dipole and H-bonding). Each of these interaction forces is expressed in terms of a limited set of molecular-level descriptors (density-based volume, molecular polarizability, molecular dipole, and H-bonding parameters) that, in turn, are calculated from molecular structure. For chemical reactivity, molecular structure is broken into functional units. Reaction centers with known intrinsic reactivity are identified and the impact on reactivity of appended molecular structure is quantified using mechanistic perturbation models.

A "toolbox" of mechanistic perturbation models has been developed that can be implemented where needed for a specific reactivity query. Resonance models were developed and calibrated on light absorption spectra [6]; electrostatic models were developed on ionization equilibrium constants [7,8]. Solvation models (i.e.,

dispersion, induction, H-bonding, dipole, etc.) have been developed on physical properties (i.e., vapor pressure, solubilities, distribution coefficient, gas chromatographic retention times, etc.) [9,10]. Ultimately these mechanistic components will be fully implemented for the aforementioned chemical and physical property models and will be extended to additional properties such as hydrolytic and redox processes.

The computational approaches in SPARC are a blend of conventional LFER, SAR, and Perturbed Molecular Orbital (PMO) methods [11,12]. In general, SPARC utilizes LFER to compute thermodynamic or thermal properties and PMO theory to describe quantum effects such as delocalization energies or polarizabilities of π electrons. In reality, every chemical property involves both quantum and thermal contributions and necessarily requires the use of both perturbation methods for prediction.

Any predictive method should be understood in terms of the purpose for which it is conceived and should be structured by appropriate operational constraints. SPARC's predictive methods can be characterized as engineering applications in environmental assessments. More specifically these methods provide :

(a) an *a priori* estimate of physicochemical parameters for physical and chemical process models when measured data are not available, (b) guidelines for ranking a large number of chemical parameters and processes in terms of relevance to the question at hand, thus establishing priorities for measurements or study, (c) an evaluation or screening mechanism for existing data based on "expected" behavior, and (d) guidelines for interpreting or understanding existing data and observed phenomena.

2. CHEMICAL REACTIVITY PARAMETERS: ESTIMATION OF IONIZATION pK_a

INTRODUCTION

A knowledge of the acid-base ionization properties of organic molecules is essential to describing chemical transport, transformation or potential environmental effects. For ionizable compounds, solubility, partitioning phenomena, and chemical reactivity are all highly dependent upon the state of ionization in the solution phase. The ionization pK_a of an organic compound is vital to environmental exposure assessment because it can be used to define the degree of ionization and the propensity for sorption to soil and sediment by cation exchange. These, in turn, can determine mobility, reaction kinetics, bioavailability, complexation etc. In addition to being highly significant in evaluating environmental fate and effects, acid-base ionization equilibria provide an excellent development arena for electrostatic effects models. Because the gain or loss of protons results in a change in molecular charge, these processes are extremely sensitive to electric field effects within the molecule.

Unfortunately, up to now no reliable method has been available for predicting pK_a over a wide range of molecular structures either for simple compounds or for complicated molecules such as dyes. The object of this study was to demonstrate the application of SPARC to the prediction of pK_a for a wide range of molecular structures.

2.1. SPARC's Chemical Modeling

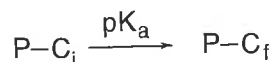
Chemical properties describe molecules in transition, that is, the conversion of a reactant molecule to a different state or structure. For a given chemical property, the transition of interest may involve electron redistribution within a single molecule or bimolecular union to form a transition state or distinct product. The behavior of chemicals depends on the differences in electronic properties of the initial state of the system and the state of interest. For example, a light absorption spectrum reflects the differences in energy between the ground and excited electronic states of a given molecule. Chemical equilibrium constants depend on the energy differences between the reactants and products. Electron affinity depends on the energy differences between the LUMO (Lowest Unoccupied Molecular Orbital) state and the HOMO (Highest Unoccupied Molecular Orbital) state.

For any chemical property addressed in SPARC, the energy differences between the initial state and the final state are small compared to the total binding energy of the reactant involved. Calculating these small energy differences by *ab initio* computational methods is difficult, if not impossible. On the other hand, perturbation methods provide these energy differences with more accuracy and with more computational simplicity and flexibility than *ab initio* methods. These methods treat the final state as a perturbed initial state and the energy differences between these two energy states are determined by quantifying the perturbation. For pK_a , the perturbation of the initial state, assumed to be the protonated form, versus the unprotonated final form, is factored into the mechanistic contributions of resonance and electrostatic effects and any other additional perturbations such as hydrogen bonding, steric contributions or solvation.

2.2. pK_a Computational Procedure

Molecular structures are broken into functional units called the reaction center and the perturber. The reaction center, C, is the smallest subunit that has the potential to ionize and lose a proton to a solvent. The perturber, P, is the molecular structure appended to the reaction center, C. The perturber structure is assumed to be unchanged in the reaction. The pK_a of the reaction center is known either from direct measurement or inferred indirectly from pK_a measurements. The pK_a of the reaction center is adjusted for the molecule in question using the mechanistic perturbation models described below.

Like all chemical reactivity parameters addressed in SPARC, pK_a is analyzed in terms of some critical equilibrium component:



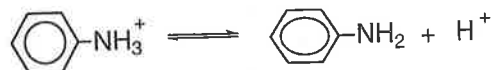
where C_i denotes the initial protonated state. C_f is the final unprotonated state of the reaction center, C. P is the "perturber". The pK_a for a molecule of interest is expressed in terms of the contributions of both P and C.

$$pK_a = (pK_a)_c + \delta_p(pK_a)_c \quad (1)$$

where $(pK_a)_c$ describes the ionization behavior of the reaction center, and $\delta_p(pK_a)_c$ is the change in ionization behavior brought about by the perturber structure. SPARC computes reactivity perturbations, $\delta_p(pK_a)_c$, that are then used to "correct" the ionization behavior of the reaction center for the compound in question in terms of potential "mechanisms" for interaction of P and C as

$$\delta_p(pK_a)_c = \delta_{ele} pK_a + \delta_{res} pK_a + \delta_{sol} pK_a + \dots \quad (2)$$

where $\delta_{res} pK_a$, $\delta_{ele} pK_a$ and $\delta_{sol} pK_a$ describe the differential resonance, electrostatic and solvation effects of P with the protonated and unprotonated states of C, respectively. Electrostatic interactions are derived from local dipoles or charges in P interacting with charges or dipoles in C. $\delta_{ele} pK_a$ represents the difference in the electrostatic interactions of the P with the two states. $\delta_{res} pK_a$ describes the change in the delocalization of π electrons of the two states due to P. This delocalization of π electrons is assumed to be into or out of the reaction center. Additional perturbations include direct interactions of the structural elements of P that are contiguous to the reaction center such as H-bonding or steric blockage of solvent access to C. For example, in the ionization of aniline, $-NR_2$ is the reaction center (denoted C) and the phenyl group is the perturber (denoted P).



The ionization equilibrium constant can be expressed as

$$pK_a = (pK_a)_c + \delta_{res} pK_a \quad (3)$$

where $(pK_a)_c$ is the pK_a for the reaction center NR_2 and is equal to 8.93, and $\delta_{res} pK_a$ is the resonance contribution to pK_a .

2.2.1. pK_a Modeling Approach.

The modeling of the perturber effects for chemical reactivity relates to the structural representation $S-R_j-C$, where $S-R_j$ is the perturber structure, P,

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts



Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research



With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips



Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.