FULLY IN SILICO CALIBRATION OF EMPIRICAL PREDICTIVE MODELS FOR ENVIRONMENTAL FATE PROPERTIES OF NOVEL MUNITIONS COMPOUNDS

PAUL G. TRATNYEK (1), ALEXANDRA J. SALTER-BLANC (1), ERIC J. BYLASKA (2), AND KURT R. GLAESEMANN (2)

PREDICTING THE ENVIRONMENTAL FATE OF NOVEL MUNITIONS COMPOUNDS REQUIRES DATA FOR THE KEY PROPERTIES THAT CONTROL CONTAMINANT FATE, AND THE MOST COMMON WAY TO PREDICT THESE PROPERTIES IS WITH EMPirical QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIPS (QSARs). HOWEVER, THE TRADITIONAL APPROACH TO QSAR DEVELOPMENT (CALIBRATION WITH EXPERIMENTAL DATA) IS CHALLENGING FOR NEW EXPLOSIVES COMPOUNDS BECAUSE OF LIMITATIONS TO THE AVAILABILITY OF THESE MATERIALS. SOME OF THE NECESSARY ENVIRONMENTAL FATE PROPERTIES CAN BE CALCULATED DIRECTLY FROM MOLECULAR STRUCTURE THEORY, BUT RELIABLE CALCULATIONS OF THIS TYPE REQUIRE CONSIDERABLE THEORETICAL EXPERTISE AND COMPUTATIONAL EFFORT. TO OVERCOME THIS COMBINATION OF CHALLENGES, WE USED A HYBRID, PARTIALLY IN SILICO APPROACH TO QSAR DEVELOPMENT, WHERE SOME OF THE CALIBRATION DATA (BOTH THE TARGET VARIABLE AND THE DESCRIPTOR VARIABLES) WERE CALCULATED FROM MOLECULAR STRUCTURE THEORY.
# TABLE OF CONTENTS

Table of Contents .............................................................................................................. ii
List of Tables .................................................................................................................... iv
List of Figures..................................................................................................................... v
List of Acronyms .............................................................................................................. viii
Keywords .......................................................................................................................... ix
Acknowledgements .......................................................................................................... ix
Abstract ............................................................................................................................... 1
Objectives............................................................................................................................ 3
Background ........................................................................................................................ 4

## Results and Discussion

1. **Objective 1—Hydrolysis** ...................................................................................... 7
   1.1 Introduction........................................................................................................ 7
   1.2 Methods........................................................................................................... 8
   1.3 Results.............................................................................................................. 9
      1.3.1 Formulation of hydrolysis reaction pathways/mechanisms and mining previously-published data on the reactants and products (Task 1.1) .............................................................. 9
      1.3.2 Calculation of target/response variable data for training set compounds with high-level theory (Task 1.2) ................................................. 16
      1.3.3 Calculate descriptor variable data for training set compounds using lower level theory (Task 1.3) ........................................................... 19
      1.3.4 Correlation Analysis and fitting QSARs (Task 1.4) .................................. 20
      1.3.5 Verification of QSARs with experimental values (Task 1.5) ............... 21
   1.4 Conclusions..................................................................................................... 22

2. **Objectives 2 and 3—Nitro Reduction** .............................................................. 22
   2.1 Introduction...................................................................................................... 23
   2.2 Methods.......................................................................................................... 24
   2.3 Results............................................................................................................ 24
      2.3.1 Formulation of reduction reaction pathways/mechanisms and mining previously-published data on the reactants and products (Task 2/3.1) ......................................................... 24
      2.3.2 Calculation of target/response variable data for training set compounds with high-level theory (Task 2/3.2) ......... 25
      2.3.3 Calculate descriptor variable data for training set compounds using lower level theory (Task 2/3.3) ............................................. 26
      2.3.4 Task 2/3.4. Correlation analysis and fitting QSARs (Task 2/3.4)........ 28
      2.3.5 Verification of QSARs with experimental values (Task 2/3.5) ........ 28
   2.4 Conclusions..................................................................................................... 31

3. **Objective 4—Amine Oxidation** ................................................................. 31
   3.1 Introduction...................................................................................................... 31
   3.2 Methods.......................................................................................................... 32
   3.3 Results............................................................................................................ 33
3.3.1 Formulation of amine oxidation reaction pathways/mechanisms and mining previously-published data on the reactants and products (Task 4.1) .................................................................33
3.3.2 Calculation of target/response variable data for training set compounds with high-level theory (Task 4.2) .................................................................34
3.3.3 Calculation of descriptor variable data with high-level and lower level theory (Task 4.3) .........................................................................................34
3.3.4 Correlation analysis and fitting QSARs (Task 4.4) .....................................36
3.3.5 Verification of QSARs with experimental values (Task 4.5) .....................38
3.4 Conclusions .................................................................................................40

Conclusions and Implications ..............................................................................40

References ..........................................................................................................42

Appendices ...........................................................................................................53

4. Supporting Data ..............................................................................................53

5. List of Scientific/Technical Publications (Chronological) ..............................53

6. Other Supporting Materials .............................................................................53
LIST OF TABLES

Table 1. Free energies of reaction in kcal/mol for substitution reactions involving nitroaromatic compounds calculated at PBE, B3LYP, and PBE0 theories with COSMO solvation model. ................................................................................................................17

Table 2. Free energies of reaction in kcal/mol for Meisenheimer addition reactions involving nitroaromatic compounds calculated at PBE, B3LYP, and PBE0 theories with COSMO solvation model. ........................................................................17

Table 3. List of priority compounds for QSAR development...............................................21

Table 4. Calculated one-electron reduction potentials ($E^{1}_{NAC}$) and comparison to a set of measured values. Calculated values were determined with density functional theory (DFT) using the the 6-311++G(2d,2p) basis set with the specified exchange correlation functionals. Solvation was accounted for using the COSMO model. ................................................................................................................27

Table 5. QSAR verification and validation data. $k_{Fe(II)P}$ were determined from experimental data. ........................................................................................................................................30

Table 6. Normalized rate constants ($k_{rel}$) for aromatic amine oxidation by MnO$_2$. ...........34
LIST OF FIGURES

Figure 1. Structures of key members of three families of modern/emerging energetic substances, where each family represents a series of compounds that are sufficiently closely related to form the basis for a QSAR model of an environmental property. .................................................................5

Figure 2. A selection of energetic NACs included or considered for use in munitions formulations. ................................................................................................................8

Figure 3. (A) Alternative initial steps in the mechanism of reaction between TNT and OH\(^-\). (B) Alternative initial steps in the mechanism of reaction between DNAN and OH\(^-\). Meisenheimer complex formation at C1, which has been hypothesized based on experimental observations [52], is shown in bold. ..........11

Figure 4. RDX. (A) Arrhenius plot displaying reported and calculated \(k_2\) data from designated sources [39, 42-44]. Data was fit to the displayed linear model to obtain the collective Arrhenius parameters. The data point from Hwang, et al. [43] was not included in the regression analysis. (B) Dependence of \(k_{obs}\) on temperature and pH. Isothermal trends for differing temperatures calculated using the Arrhenius parameters determined in part A of this figure are shown with diagonal lines. Reported \(k_{obs}\) values from designated sources are also displayed [39, 42-45]. ................................................................................................12

Figure 5. HMX. (A) Arrhenius plot displaying reported \(k_2\) data from designated sources [42, 44]. Data was fit to the displayed linear model to obtain the collective Arrhenius parameters. (B) Dependence of \(k_{obs}\) on temperature and pH. Isothermal trends for differing temperatures calculated using the Arrhenius parameters determined in part A of this figure are shown with diagonal lines. Reported \(k_{obs}\) values from designated sources are also displayed [42, 44] ..................13

Figure 6. CL-20. (A) Arrhenius plot displaying reported and calculated \(k_2\) data from designated sources [48, 49]. Data was fit to the displayed linear model to obtain the collective Arrhenius parameters. (B) Dependence of \(k_{obs}\) on temperature and pH. Isothermal trends for differing temperatures calculated using the Arrhenius parameters determined in part A of this figure are shown with diagonal lines. Reported \(k_{obs}\) values from designated sources are also displayed [45, 48, 49] ............13

Figure 7. TNT. Arrhenius plot displaying reported and calculated \(k_2\) data from designated sources [16, 19]. Data was fit to the displayed linear model to obtain the collective Arrhenius parameters.........................................................................................14

Figure 8. Screen shot of a spreadsheet-like view of the database, showing basic information about each compound........................................................................................................15
Figure 9. “Form view” page for an individual compound (RDX) with data linked from the database component shown in Figure 8 and from another component containing kinetic data mined from the literature. ...............................................................15

Figure 10. Results of a query for a structural moiety involved in the mechanism described above for the hydrolysis of RDX and related compounds. It is likely compounds in this set hydrolyze by a similar mechanism, so it will be logical to group these compounds in the same QSAR.........................................................16

Figure 11. (a) Two dimensional potential energy surface for the reaction, \( \text{C}_6\text{H}_5\text{NO}_2\text{(aq)} + \text{OH}^-\text{(aq)} \rightarrow \text{C}_6\text{H}_5\text{OH}\text{(aq)} + \text{NO}_2^-\text{(aq)} \), calculated at the B3LYP/6-311++G(3p,3d) //COSMO level. The reaction barrier for this potential energy surface is \( E_{\text{act}} = +31 \) kcal/mol. (b) Illustration of an AIMD/MM simulation for the \( \text{C}_6\text{H}_5\text{NO}_2\text{(aq)} + \text{OH}^-\text{(aq)} \rightarrow \text{C}_6\text{H}_5\text{OH}\text{(aq)} + \text{NO}_2^-\text{(aq)} \) reaction and the two dimensional free energy surface of the reactant well obtained from the AIMD/MM metadynamics simulation. The reaction barrier for this potential energy surface is \( E_{\text{act}} = +30 \) kcal/mol. ........................................................................18

Figure 12. Comparison of DFT calculations of \( E_{\text{LUMO}} \) from structures optimized using HF theory and DFT. .................................................................................................20

Figure 13. (A) Arrhenius plot and (B) Eyring plot showing second order rate constants for alkaline hydrolysis of TNT. Both new data and data reported in the literature are shown. Arrhenius parameters as well as the enthalpy and entropy of the transition state (\( \Delta H^\ddagger \) and \( \Delta S^\ddagger \)) were obtained from fitting the data to the appropriate model. The free energy for the transition state (\( \Delta G^\ddagger \)) at 25 °C was calculated from \( \Delta H^\ddagger \) and \( \Delta S^\ddagger \). .....................................................................................22

Figure 14. \( \log(k_{\text{red}}) \) for various NACs vs. (A) \( E_{1\text{NAC}}^{\ddagger} / 0.059 \) V and (B) \( \Delta G^\ddagger \). Reductants include the electron carrier/donor pairs juglone/hydrogen sulfide \([11, 12]\), lawsone/hydrogen sulfide \([11]\), Fe(II) porphyrin/cysteine \([11]\), electrochemically reduced AQDS \([70]\), and the Fe(II)-ligand complexes Fe(II)-tiron \([76]\) and Fe(II)-DFOB \([77]\). Closed symbols represent data for which the descriptor variable was determined from measured \( E_{1\text{red}}^{\ddagger} \). Open symbols represent data for which the descriptor variable was determined from \( E_{1\text{red}}^{\ddagger} \) values that were estimated from LFERs \([11, 12]\). Only data for which \( E_{1\text{red}}^{\ddagger} \) was measured were included in the fits. The correlated data were fit to (A) Equation 6 or (B) the Marcus equation. In (B), an extrapolation of the fit is shown (as a dashed line) for cases where the data is well fit to the Marcus model.........................26

Figure 15. QSARs for NAC reduction by Fe(II)P calibrated using rate constants \( \log(k_{\text{Fe(II)P}}) \) obtained from the literature \([11]\) and \( E_{1\text{NAC}}^{\ddagger} \) calculated at (A) the B3LYP/6-311++G(2d,2p) level or (B) the M06-2X/6-311++G(2d,2p) level,
both with an applied linear transformation based on correlations to a measured dataset (denoted as $E^{1}_{\text{NAC, B3LYP}}$ and $E^{1}_{\text{NAC, M06-2X}}$, respectively).

**Figure 16.** Comparison of newly collected log($k_{\text{Fe(II)P}}$) data to the calibration dataset and QSAR. Data for non-energetic NACs were used to verify agreement of the new data with the calibration data (method verification data). Data for energetic NACs were used to test the validity of the QSAR to energetic compounds (validation data).

**Figure 17.** Correlation of log $k_{\text{rel}}$ for oxidation of aromatic amines to $E_{\text{HOMO}}$. $E_{\text{HOMO}}$ was calculated using Gaussian09 with B3LYP, 6-31G*, and SMD. Fitting coefficients and statistics for the regression line given in the Supporting Information of [121]. Labels correspond to numbers in Table 6.

**Figure 18.** log $k_{\text{rel}}$ for oxidation of aromatic amines to $E_{\text{ox}}$. Calculated with NWChem, M06-2X, 6-311++G(2d,2p), and COSMO. Fitting coefficients and statistics for the regression line given in the Supporting Information of [121]. Labels correspond to numbers in Table 6.

**Figure 19.** Comparison of correlations for normalized oxidation rates constants ($k_{\text{rel}}$) for aromatic amines versus MnO2 (this study and prior literature) and previously published data for other oxidants. The data and fit for MnO2 (gray) are from Figure 17. The data for other oxidants are compiled in in the Supporting Information of [121], and the fit of these data is described in the text.

**Figure 20.** Correlation between the rates of nitro reduction for various energetics compounds ($k$), relative to TNT, versus the descriptor variable that provides the best overall QSAR (E$^{1}$). The right axis shows rate constants converted to approximate half-lives to give. The blue symbols are experimental data, obtained with IMC related amino compounds.
## LIST OF ACRONYMS

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Arrhenius pre-exponential factor</td>
</tr>
<tr>
<td>AT</td>
<td>Atom transfer reaction</td>
</tr>
<tr>
<td>B3LYP</td>
<td>Becke Lee-Yang-Parr 3-parameter fit</td>
</tr>
<tr>
<td>CCSD(T)</td>
<td>Coupled Cluster with Singles, Doubles, and perturbative Triples</td>
</tr>
<tr>
<td>CL-20</td>
<td>2,4,6,8,10,12-hexanitro-2,4,6,8,10,12-hexaazaisowurtzitane</td>
</tr>
<tr>
<td>Cmpnds.</td>
<td>Compounds</td>
</tr>
<tr>
<td>COSMO</td>
<td>COnductor-like Screening MOdel</td>
</tr>
<tr>
<td>DFT</td>
<td>Density Functional Theory</td>
</tr>
<tr>
<td>$E_a$</td>
<td>Activation energy</td>
</tr>
<tr>
<td>$E_{\text{GAP}}$</td>
<td>HOMO-LUMO energy gap</td>
</tr>
<tr>
<td>$E_{\text{HOMO}}$</td>
<td>Energy of the highest occupied molecular orbital</td>
</tr>
<tr>
<td>$E_{\text{LUMO}}$</td>
<td>Energy of the lowest unoccupied molecular orbital</td>
</tr>
<tr>
<td>ESP</td>
<td>Environmental Simulation Program</td>
</tr>
<tr>
<td>$E_T$</td>
<td>Total Energy</td>
</tr>
<tr>
<td>ET</td>
<td>Electron Transfer</td>
</tr>
<tr>
<td>GAMESS</td>
<td>General Atomic and Molecular Electronic Structure System</td>
</tr>
<tr>
<td>HF</td>
<td>Hartree-Fock</td>
</tr>
<tr>
<td>HMX</td>
<td>1,3,5,7-tetranitro-1,3,5,7-tetrazocane</td>
</tr>
<tr>
<td>HOMO</td>
<td>Highest Occupied Molecular Orbital</td>
</tr>
<tr>
<td>HPLC</td>
<td>High performance liquid chromatography</td>
</tr>
<tr>
<td>$k_2$</td>
<td>Second-order rate constant</td>
</tr>
<tr>
<td>$k_t$</td>
<td>Rate constant for the forward portion of a reversible reaction</td>
</tr>
<tr>
<td>$k_{t,\text{obs}}$</td>
<td>Observed rate constant for the forward portion of a reversible reaction</td>
</tr>
<tr>
<td>$k_{\text{obs}}$</td>
<td>Observed rate constant</td>
</tr>
<tr>
<td>$k_r$</td>
<td>Rate constant for the reverse portion of a reversible reaction</td>
</tr>
<tr>
<td>LLNL</td>
<td>Lawrence Livermore National Laboratory</td>
</tr>
<tr>
<td>LUMO</td>
<td>Lowest Unoccupied Molecular Orbital</td>
</tr>
<tr>
<td>MM</td>
<td>Molecular Mechanics</td>
</tr>
<tr>
<td>MOPAC</td>
<td>Molecular Orbital PACkage</td>
</tr>
<tr>
<td>MP2</td>
<td>Møller-Pleset perturbation theory through second order</td>
</tr>
<tr>
<td>NEB</td>
<td>Nudged Elastic Band</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>NOM</td>
<td>Natural Organic Matter</td>
</tr>
<tr>
<td>OHSU</td>
<td>Oregon Health &amp; Science University</td>
</tr>
<tr>
<td>PCM</td>
<td>Polarizable Continuum Model</td>
</tr>
<tr>
<td>PIMC</td>
<td>Path Integral Monte Carlo</td>
</tr>
<tr>
<td>PMF</td>
<td>Potential of Mean Force</td>
</tr>
<tr>
<td>PNNL</td>
<td>Pacific Northwest National Laboratory</td>
</tr>
<tr>
<td>PT</td>
<td>Proton Transfer</td>
</tr>
<tr>
<td>QCISD(T)</td>
<td>Quadratic Configuration Interaction with Singles, Doubles, and perturbative Triples</td>
</tr>
<tr>
<td>QM</td>
<td>Quantum Mechanics</td>
</tr>
<tr>
<td>QM/MM</td>
<td>Quantum Mechanics/Molecular Mechanics</td>
</tr>
<tr>
<td>QSAR</td>
<td>Quantitative Structure-Activity Relationship</td>
</tr>
<tr>
<td>RDS</td>
<td>Rate Determining Step</td>
</tr>
<tr>
<td>RDX</td>
<td>1,3,5-trinitro-1,3,5-triazine</td>
</tr>
<tr>
<td>SIC</td>
<td>Self-interaction correction</td>
</tr>
</tbody>
</table>
KEYWORDS

Munitions Compounds, Nitro aromatic compounds, Aromatic amines, TNT, DNAN, Environmental Fate Properties, Hydrolysis, Reduction, Oxidation, Kinetics, Rate Constants, Ab Initio Quantum Mechanical Calculations, Quantitative structure activity relationships (QSARs)

ACKNOWLEDGEMENTS

Calculations performed at PNNL were performed using the Molecular Science Computing Capability in the William R. Wiley Environmental Molecular Science Laboratory (EMSL), a national scientific user facility sponsored by the U.S. Department of Energy's Office of Biological and Environmental Research and located at the Pacific Northwest National Laboratory, operated for the Department of Energy by Battelle. In addition, this work makes use of the open source quantum chemistry package NWChem maintained by the EMSL user facility. We would also like to thank the PNNL Institutional Computing (PIC) program at Pacific Northwest National Laboratory (PNNL) for a small block of computing time.
ABSTRACT

Objectives: Predicting the environmental fate of novel munitions compounds requires data for the key properties that control contaminant fate, and the most common way to predict these properties is with empirical quantitative structure-activity relationships (e.g., QSARs). However, the traditional approach to QSAR development (calibration with experimental data) is challenging for new explosives compounds because of limitations to the availability of these materials. Some of the necessary environmental fate properties can be calculated directly from molecular structure theory, but reliable calculations of this type require considerable theoretical expertise and computational effort. To overcome this combination of challenges, we used a hybrid, partially in silico approach to QSAR development, where some of the calibration data (both the target variable and the descriptor variables) were calculated from molecular structure theory.

Technical Approach: The technical approach of the project we divided into three objectives, corresponding to the three reaction pathways that contribute most to the overall fate of explosives in the environment. These include: (i) hydrolysis and associated elimination reactions, which are ubiquitous in water; (ii) nitro reduction by outer-sphere electron transfer, which are dominant in soils in sediments; and (iii) oxidation of the amino products soil minerals like manganese dioxide.

Results: After extensive analysis of prior work, it was concluded that mechanism(s) of hydrolysis of the nitro aromatic energetic compounds TNT and DNAN was not sufficiently advanced to support conventional QSAR development. The possible difference in mechanisms means that predicting the reaction mechanism for one based on the other may lead to unreliable predictions of environmental fate. This, along with uncertainties in the consistency of the calculated results with experimental values, presents a challenge for developing QSARs calibrated “in silico” that predict the hydrolysis behavior of the diverse range of energetic NACs. However, new experimental and computational results reported here provide insight into these mechanisms.

In contrast to the hydrolysis pathway, the effect of complexity in the nitro reduction pathway was effectively captured by contrasting two approaches: theory-based correlations to $E_1$ and empirical correlations to $E_{LUMO}$. The former provides a variety of insights into the fundamental processes controlling nitro reduction and the latter provides simple QSARs that should be useful for estimating relative reduction rates of nitro aromatic compounds. The models were calibrated and/or tested with a variety of model and actual energetic compounds, including TNT and DNAN. An overall conclusion of this aspect of the study is that most alternative energetic compounds will react by nitro reduction more slowly than TNT.

For oxidation of anilines a variety of QSARs were developed that represent a range of convenience and accuracy. The QSARs based on the most convenient descriptors (e.g., pKa) were less accurate and models based on the more specialized descriptors (e.g., $E_{ox}^1$) were the most accurate. Comparison of the model for MnO$_2$ to other environmentally-relevant oxidants showed that the former was more sensitive to substituent effects than the latter.

Benefits. The QSARs provided by this project can be used to predictive the relative rates of major transformation processes that contribute to the environmental fate of energetic
compounds. The models are flexible enough to handle some of the diverse molecular structures that are found in candidate compounds for new explosives. They are being integrated into models developed by others to predicted the overall fate of energetic compounds.
OBJECTIVES

Predicting the environmental fate of novel munitions compounds requires data for the key properties that control contaminant fate, and the most common way to predict these properties is with empirical quantitative structure-activity relationships (e.g., QSARs). However, the traditional approach to QSAR development (calibration with experimental data) is impractical for new explosives compounds because of limitations to the availability of these materials. Some of the necessary environmental fate properties can be calculated directly from molecular structure theory, but reliable calculations of this type require considerable theoretical expertise and computational effort.

To overcome this combination of challenges, we proposed a novel, *in silico* approach to QSAR development, where most of the calibration data (both the target variable and the descriptor variables) are calculated from molecular structure theory. Rates of the most likely breakdown pathways (target variables) were to be calculated with the highest level of theoretical accuracy and these data were correlated to molecular properties (descriptor variables) that were obtained with computational methods that are more available and feasible for most chemists. As QSARs were obtained by this approach, we attempted to validate them by predicting data for (safe and available) model compounds and comparing to measured experimental values.

The technical approach of the project was originally divided into four objectives (later streamlined to three), corresponding directly to the three reaction pathways that contribute most to the overall fate of explosives in the environment. These four pathways are:

- Hydrolysis and associated elimination reactions, which are ubiquitous in water.
- Homogeneous nitro reduction by outer-sphere electron transfer with dissolved electron donors or at the surfaces of reducing minerals.
- Polymerization of the amino products (with themselves and with the phenolic moieties associated with natural organic matter) of reduction, by oxidative coupling or nucleophilic condensation.

Each one of these objectives was pursued via five tasks: (i) reaction formulation and mining of existing data, (ii) calculation of target variable data with high level theory, (iii) calculation of descriptor variable data with lower level theory, (iv) correlation analysis and fitting QSARs, and (v) validation of data predicted with the QSAR.
BACKGROUND

Predicting chemical properties is a long-standing challenge that has received extensive study for many applications (chemical engineering, green chemistry, environmental chemistry, toxicology, pharmacology, etc.). For chemical properties that determine the fate and effects of environmental contaminants, the field is mature enough to have already engendered several generations of compilations of predictive models. However, a realistic assessment of the state-of-the-art in this field reveals some pervasive limitations. In particular, most existing property prediction models are based on empirical correlations between the target property and descriptor variables (e.g., quantitative structure-activity relationships, QSARs), which must be calibrated with measured data, and are of dubious validity for extrapolation outside the training set data-space. Expanding or validating the scope of application of such QSARs requires more experimental data, which often ends up being the same data that the QSAR was supposed to serve to estimate. To overcome this dilemma, novel approaches to chemical property prediction are needed.

With respect to chemical property prediction to support development of new, environmentally benign chemicals for explosives, the challenge noted above is particularly severe for two reasons. First, safety, regulatory, and security considerations make it impractical to do the experiments necessary to calibrate and validate QSARs. Second, while most current explosives and candidates for future explosives are characterized by >N-NO₂ or Ar-NO₂ moieties, they do not fit into simple homologous series of congeners that are described by traditional QSARs. This can be seen in Figure 1, where we have shown the structures of some current and future explosives arranged to illustrate groups that have a high enough degree of homology and number of potential members to make meaningful QSARs feasible. The groups are defined by a variety of analogous backbone types containing variable numbers of a single moiety that is subject to a common pathway of reaction. The properties of such groups of chemicals can be described by QSARs, but they must be based on suitable descriptor variables. Most descriptor variables that are suitable for this purpose are molecular properties, which are generally measured, which—as noted above for target properties—is impractical for most explosives.

Alternatively, most properties that directly impact environmental fate can—in principle—be calculated directly from molecular structure theory. However, performing accurate calculations of this type requires specialized theoretical expertise and considerable computational resources, so the fully “ab initio” approach to chemical property estimation is also not a practical solution for most of those who are responsible for assessing the potential environmental fate and effects of novel explosives.

To overcome this unique combination of challenges, we proposed take a novel hybrid approach that provides QSAR-like models calibrated with data obtained partially from

---

1 Most prominent among these is the Handbook of Chemical Property Estimation Methods compiled by Lyman et al. [2], which became so highly regarded that a subsequent volume by Mackay and Boethling [3] was promoted as the 2nd Edition of Lyman et al. and its publication was consecrated with very well attended symposium, with Lyman as keynote speaker, at the 20th Annual meeting of the Society for Environmental Toxicology and Chemistry, 14-18 November 1999, Philadelphia, PA.
theoretical calculations based on molecular structure theory. This “fully in silico”
approach to QSAR development uses descriptor variable data that can be calculated using
methods that are feasible for non-specialists and target variable data calculated (instead of
measured) for complete characterization of reaction coordinates using state-of-the-art
computational methods of highest accuracy. For example, consider the formation of
hydride-Meisenheimer complexes, which frequently controls rates of biodegradation for
2,4,6-trinitrotoluene (TNT) and other nitro aromatics [4]. There are no reliable kinetic
data for this reaction and they would be challenging to obtain experimentally, but rates
constants for this reaction could be calculated from theory since the reduction mechanism
is well–defined. Correlating the calculated rate constants (target variable) with an easily
calculated molecular property (descriptor variable) like charge density on key atoms
should produce a QSAR that would be easy to apply to other nitro aromatics.

<table>
<thead>
<tr>
<th>Nitramines</th>
<th>Nitrofurans</th>
<th>Furoxans</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNAZ (etc.)</td>
<td>O₂N N₂O₂</td>
<td>Cl-14 (ADNBF, etc.)</td>
</tr>
<tr>
<td>RDX (HMX, etc.)</td>
<td>N₂N N₂O₂</td>
<td>DDF (etc.)</td>
</tr>
<tr>
<td>CL-20 (TEX, etc.)</td>
<td>O₂N N₂O₂</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 1.** Structures of key members of three families of modern/emerging energetic
substances, where each family represents a series of compounds that are sufficiently
closely related to form the basis for a QSAR model of an environmental property.

In principle, this approach could be applied to any process affecting the fate of any
type of contaminant, but the focus of this study on explosive compounds allowed us to
narrow the range of processes to be considered. By necessity, explosive compounds are
generally oxidized, polar, and unstable compounds. Therefore, in the environment, they
show relatively little tendency to adsorb or bioaccumulate, facile reduction but relatively
little susceptibility to oxidation, and significant hydrolysis initiated by “spontaneous” elimination reactions. This means that we can cover the most important processes controlling the environmental persistence of explosives by considering just hydrolysis and reduction reactions. Hydrolysis can be adequately described as a reaction in homogeneous solvent, but modeling reduction is more complex. Since the overall rate of most contaminant reductions is controlled by the rate of the initial electron attachment, we developed a generic model for reduction by single electron transfer from a solution-phase donor.

Reduction of nitro groups produces amino-substituted products and the environmental fate of these amines is controlled mainly by oxidative coupling to soil/sediment organic matter [5, 6]. Like reduction, this is a complex and potentially system-specific process, but we should be able to model relative rates of the process by assuming a rate-limiting oxidation to produce radical intermediates that couple with phenolic moieties of natural organic matter (NOM) model compounds. In this case, disappearance of the original contaminant is not the most challenging issue; instead, the greatest concern is with assessing the potential for reversing this reaction (decoupling) and release of the amino residues, which are potentially problematic contaminants. With the theoretical approach used here, it should eventually be possible to consider both coupling and decoupling (i.e., the reversibility of this process).

While all of the model development and some of the model validation used in this study were done in silico, we tested each model against experimental data in several ways. First, where experimental data were available for the target or descriptor variables describing hydrolysis, reduction, or oxidative coupling of explosives, we compared these directly with our calculated values. However, since there were many data gaps, our approach had the fundamental advantage that we can calculate properties of other model compounds for which data are available and use these in validation of our QSARs. Finally, we used the combination of our predictive models and the validation data we compiled from other work, to prioritize what additional experimental data is necessary to validate our models. We made some of these measurements as part of this project, and mined experimental data from other recent, SERDP-funded work on the environmental fate of insensitive munitions compounds.
RESULTS AND DISCUSSION

The main results of this project are presented in three sections that correspond to the three major objectives of the proposal. Each of these sections is derived from one or more peer-reviewed publications, which are specified at the beginning of each section.

1. **Objective 1—Hydrolysis**


**Abstract:** The environmental impacts of energetic compounds can be minimized through the design and selection of new energetic materials with favorable fate properties. Building predictive models to inform this process, however, is difficult because of uncertainties and complexities in some major fate-determining transformation reactions such as the alkaline hydrolysis of energetic nitroaromatic compounds (NACs). Prior work on the mechanisms of the reaction between NACs and OH\(^{-}\) has yielded inconsistent results. In this study, the alkaline hydrolysis of 2,4,6-trinitrotoluene (TNT) and 2,4-dinitroanisole (DNAN) was investigated with coordinated experimental kinetic measurements and molecular modeling calculations. For TNT, the results suggest reversible formation of an initial product, which is likely either a Meisenheimer complex or a TNT anion formed by abstraction of a methyl proton by OH\(^{-}\). For DNAN, the results suggest that a Meisenheimer complex is an intermediate in the formation of 2,4-dinitrophenolate. Despite these advances, the remaining uncertainties in the mechanisms of these reactions—and potential variability between the hydrolysis mechanisms for different NACs—mean that it is not yet possible to generalize the results into predictive models (e.g., quantitative structure–activity relationships, QSARs) for hydrolysis of other NACs.

1.1 **Introduction**

Many energetic compounds are nitroaromatic compounds (NACs). The most familiar member of this group is 2,4,6-trinitrotoluene (TNT), which has long been used in a variety of energetic materials. Also included is a shock-insensitive alternative to TNT, 2,4-dinitroanisole (DNAN) [7], which is used in modern munitions formulations such as PAX-21 [8]. While TNT and DNAN comprise a subfamily of substituted nitrobenzenes, other NACs used in emerging munitions formulations include substituted heterocycles, such as 5-nitro-1,2,4-triazol-3-one (NTO)—another insensitive replacement for TNT that is used in emerging munitions formulations [7]—and polyaromatic compounds. A representative sample of energetic NACs is shown in **Figure 2**; these and many more have been described previously [1, 9].
Figure 2. A selection of energetic NACs included or considered for use in munitions formulations.

The main abiotic transformation processes influencing the environmental fate of NACs in groundwater are reduction and alkaline hydrolysis [10]. While the pathways and kinetics associated with the reduction of NACs have been studied extensively (e.g., [11-14]), the pathways associated with alkaline hydrolysis are not well characterized, despite considerable proof-of-concept testing of this process for engineered remediation of TNT-contaminated wastewaters [15-23]. In the case of TNT, there are significant inconsistencies in the experimental data, which likely arise from the sensitivity of the reaction to the nature of the solvent [24, 25], difficulty observing products in water due to their poor solubility [19], and the influence of the concentration ratio between TNT and base on product distribution [24, 25], difficulty observing products in water due to their poor solubility [19]. Previous computational studies of TNT reaction energetics have also failed to definitively determine the pathways of its reaction with OH$^-$. This incomplete understanding of the mechanisms of TNT hydrolysis in natural and engineered systems inhibits the prediction of degradation properties of future energetic NACs through the use of correlations calibrated with data determined fully in silico. Our goal under this objective was to clarify the mechanisms of TNT degradation by OH$^-$ in water by taking a combined approach that emphasizes reconciliation of both experimental and computational data. An analysis of the kinetics of TNT disappearance in the presence of OH$^-$ is presented and used to determine the experimental activation free energies ($\Delta G^*$) and reaction free energies ($\Delta G_{rxn}$) for the processes. Also presented are $\Delta G^*$ and $\Delta G_{rxn}$ values for possible mechanisms of TNT degradation determined using molecular modeling. The measured and modeled results are then compared to assess which mechanisms predominate. Also reported is an analogous analysis for DNAN, which is structurally similar to TNT.

1.2 Methods

Method details are more fully documented in Salter-Blanc et al. [29].

**Batch experiments.** Batch experiments were carried out in 20-mL amber VOA vials capped with Teflon-lined silicon septa (SUN Sri; Rockwood, TN). Each vial initially
contained 20 mL phosphate buffer (50 mM) at pH 11.0, 11.7, or 12.0. Vials were
temperature equilibrated in water baths or in a cold room to 11.0, 25.0, 40.0, 55.0, or 65.0
°C. After temperature equilibration, TNT or DNAN was introduced to the reaction vial
by injecting 200 μL of a 10-g/L stock solution prepared in acetonitrile (reaction
initiation). Following introduction of TNT or DNAN, the vials were shaken by hand for
~1 min (in the water bath) to insure proper mixing. 1-mL aliquots were removed at
specified times and quenched by mixing with an equal volume of acidified acetonitrile.
Quenched aliquots were analyzed by high-pressure liquid chromatography (HPLC).

**Molecular Modeling.** In this study, the solution phase $\Delta G_{\text{rxn}}$ and $\Delta G^\ddagger$ were directly
calculated from gas-phase reaction energy, entropy, electronic structure calculations,
continuum solvation models, and gas-phase entropy estimates. All calculations were
performed using the NWChem program suite [30]. The electronic structure calculations
were either performed using density functional theory (DFT) [31] with the B3LYP [32,
33] exchange correlation potential, or with second-order Møller-Plesset perturbation
theory (MP2) [34]. In all cases, the 6-311++G(2d,2p) basis set was used (obtained from
the Extensible Computational Chemistry Environmental Basis Set Database [35]). The
solvation energies were estimated using the self-consistent reaction field theory of Klamt
and Schüürmann (COSMO) [36], with the cavity defined by a set of overlapping atomic
spheres with radii suggested by Stefanovich and Truong [37]. The dielectric constant of
water used for all of the solvation calculations was 78.4. The solvent cavity discretization
was generated from the surfaces of non-overlapping spheres that were discretized by an
iterative refinement of triangles starting from a regular octahedron.

The geometries and harmonic frequencies for the reactants, products and transition
states were consistently optimized using the B3LYP calculation with COSMO.
Transitions states were determined for the Meisenheimer and proton abstraction reactions
by performing constrained optimizations at a series of C-OH and CH-OH bond distances.
The virtual entropies for each compound were estimated using formulas derived from
statistical mechanics that are broken into translational, rotational, and frequency terms. In
addition to the COSMO correction, cavitation, and dispersion contributions to the
solvation energy were added a posteriori using empirically-derived expressions that
depend only on the solvent accessible surface area. In this study, we used the
parameterized formula given by Sitkoff et al. [38].

### 1.3 Results

1.3.1 Formulation of hydrolysis reaction pathways/mechanisms and mining
previously-published data on the reactants and products (Task 1.1)

#### 1.3.1.1 Mechanism Determination

**Cyclic and cage-cyclic nitramines.** The most highly-studied cyclic nitramine, with
respect to hydrolysis reaction mechanisms, is RDX. Aqueous alkaline hydrolysis of RDX
occurs through a concerted two-step process involving proton abstraction by hydroxide
and loss of a nitrite ion (NO$_2^-$) [39] (E2 elimination reaction). Initial attack by hydroxide
is rate-limiting [40, 41] and the reaction is second order with respect to hydroxide
concentration [39, 41-44]. Denitration is followed by ring cleavage and spontaneous
decomposition to N$_2$O, HCHO, and 4-NDAB [45]. Identical products are observed following hydrolysis of HMX and MNX, which—like RDX—are observed to produce one equivalent of NO$_2^-$ in the reaction process [45]. This supports earlier suggestions that cyclic nitramines other than RDX also react via an E2 elimination reaction [46]. One theoretical study suggests that an S$_N$1 hydroxide ion substitution of RDX is more energetically favorable than the E2 elimination reaction [47], but this suggestion is in disagreement with experimentally observed products [45].

Structurally similar to the cyclic nitramines is CL-20, a cage-cyclic (or polycyclic) nitramine. Kinetic studies of CL-20 show second-order behavior with respect to CL-20 and hydroxide concentrations [45, 48, 49]. One study suggests the reaction proceeds by initial denitration—similarly to the monocyclic nitramines—however, the reaction produces two equivalents of NO$_2^-$ and different final products than were observed in the monocyclic experiments. It has also been suggested that this reaction proceeds differently under high or low concentrations of hydroxide, producing products that have different susceptibilities to photodegradation [47]. Comparison of experimental FTIR spectra for CL-20 hydrolysis and DFT theorized spectra for 1,5- and 1,7- dihydrodiimidazo[4,5-b:4',5'-e]pyrazine suggest CL-20 hydrolysis produces a mixture of these tautomers through cleavage of the C-C bond attaching the two 5-membered rings [50].

**Nitroaromatic compounds.** Most relevant environmental literature on the alkaline hydrolysis of nitroaromatic compounds focuses on TNT. The reaction of TNT with hydroxide is second order with respect to TNT and hydroxide concentrations [16, 19]. The pathways of TNT hydrolysis are complex—intermediate and product formation appear to be affected by pH [23], the TNT/hydroxide concentration ratio [24], and the reaction media [19]. Prior studies led to three main hypotheses regarding the first step of TNT alkaline hydrolysis; these are: (i) formation of a Meisenheimer complex [17, 19, 51], (ii) direct nitro substitution [20], and (iii) abstraction of a methyl proton [24, 25]. More details are given in the publication resulting from this objective [29]. Alternative initial steps in the reaction between TNT and OH$^-$ are shown in Figure 3A. A limited number of studies have addressed the alkaline hydrolysis of 2,4-dinitroanisole (DNAN). The reaction is suggested to occur by substitution of the methoxy group to form 2,4-nitrophenol, which in turn may react with OH$^-$ to from 2,4-dinitrophenolate (Figure 3B).
Figure 3. (A) Alternative initial steps in the mechanism of reaction between TNT and OH\(^{-}\). (B) Alternative initial steps in the mechanism of reaction between DNAN and OH\(^{-}\). Meisenheimer complex formation at C1, which has been hypothesized based on experimental observations [52], is shown in bold.

1.3.1.2 Determination of “Best Values” of kinetic data from Literature

Rate constants for the base-catalyzed hydrolysis of various nitramines were collected from the literature. Reported values for \(k_2\), the second-order rate constant, were compared in Arrhenius plots. In some cases, \(k_2\) was not reported directly, but was calculated from reported values of \(k_{obs}\), the pseudo-first-order rate constant, measured at various pH’s.

This set of data was used to determine the temperature dependence of the reaction as defined by the Arrhenius equation (Equation 1), where \(A\) is the pre-exponential factor and \(E_a\) is the activation energy.

\[
k_2 = Ae^{-\frac{E_a}{RT}}
\]  

(1)

The natural logarithm of Equation 1 is shown in Equation 2:

\[
\ln(k_2) = -\frac{E_a}{R} \frac{1}{T} + \ln(A)
\]  

(2)

Equation 2 was used to fit \(k_2\) vs. 1/T data for RDX (Figure 4A), HMX (Figure 5A) and CL-20 (Figure 6A) to obtain the Arrhenius parameters \(A\) and \(E_a\). In all cases, the fit shows excellent agreement between the \(k_2\) values obtained from different sources. The only data point that appears to deviate is the value for RDX calculated from data reported by Hwang, et al. [43] (Figure 4A). This data point was excluded from the final fit. Also
shown is a plot for TNT (Figure 7A). Agreement between the values in this figure is not as good as seen with the other compounds.

The Arrhenius parameters obtained from fitting the data in Figure 4A, Figure 5A, and Figure 6A were used to predict the dependence of the observed rate constants \(k_{\text{obs}}\) on pH and temperature by using the relationship \(k_{\text{obs}} = k_2 [\text{OH}^-]\) and Equation 5. The results are shown in Figure 4B, Figure 5B, and Figure 6B with diagonal lines. Also on these plots are the \(k_{\text{obs}}\) data used by Hoffstetter, et al. [39], Heilmann, et al. [42], Hwang, et al. [43], Monteil-Rivera, et al. [44], Karakaya, et al. [48], and Santiago, et al. [49] to obtain the \(k_2\) values shown in Figure 4A, Figure 5A, and Figure 6A. Additional \(k_{\text{obs}}\) data from Balakrishnan, et al. [45] are also shown, although these data were collected under heterogeneous conditions, whereas the other data were collected under homogeneous conditions. With the exception of the Balakrishnan, et al. [45] data, the figures show excellent agreement between the values reported by the different sources and the calculated Arrhenius parameters. This agreement supports the use of these parameters to calculate “best values” for the hydrolysis rate constants to be used in correlation analysis.

**Figure 4.** RDX. (A) Arrhenius plot displaying reported and calculated \(k_2\) data from designated sources [39, 42-44]. Data was fit to the displayed linear model to obtain the collective Arrhenius parameters. The data point from Hwang, et al. [43] was not included in the regression analysis. (B) Dependence of \(k_{\text{obs}}\) on temperature and pH. Isothermal trends for differing temperatures calculated using the Arrhenius parameters determined in part A of this figure are shown with diagonal lines. Reported \(k_{\text{obs}}\) values from designated sources are also displayed [39, 42-45].
Figure 5. HMX. (A) Arrhenius plot displaying reported $k_2$ data from designated sources [42, 44]. Data was fit to the displayed linear model to obtain the collective Arrhenius parameters. (B) Dependence of $k_{obs}$ on temperature and pH. Isothermal trends for differing temperatures calculated using the Arrhenius parameters determined in part A of this figure are shown with diagonal lines. Reported $k_{obs}$ values from designated sources are also displayed [42, 44].

Figure 6. CL-20. (A) Arrhenius plot displaying reported and calculated $k_2$ data from designated sources [48, 49]. Data was fit to the displayed linear model to obtain the collective Arrhenius parameters. (B) Dependence of $k_{obs}$ on temperature and pH. Isothermal trends for differing temperatures calculated using the Arrhenius parameters determined in part A of this figure are shown with diagonal lines. Reported $k_{obs}$ values from designated sources are also displayed [45, 48, 49].
Figure 7. TNT. Arrhenius plot displaying reported and calculated $k_2$ data from designated sources [16, 19]. Data was fit to the displayed linear model to obtain the collective Arrhenius parameters.

1.3.1.3 Organization of Data in a Chemically-Intelligent Database

In addition to mining mechanism and rate constant data from the literature, another aspect of Task 1.1 was the development of a chemoinformatic database for storing this and other data. After inspecting chemically intelligent database packages, we decided to use Instant JChem by ChemAxon (Budapest, Hungary). This software allows for storage and query of numerical, textual, and structural data and for the development of relational databases. This will be very useful for storing complex data and meta data and for sorting by mechanistically-relevant structural moieties. The ChemAxon tools also include certain property predictors such as a pKa calculator, which may be used in determining descriptors for QSAR development. Screen shots of the database are shown in Figure 8 to Figure 10.

The decision to use Instant JChem was partly motivated by the adoption of ChemAxon tools by Dr. Eric Weber’s group of the EPA, Athens (for his SERDP project, but also for broader goals of EPA’s “CompTox” programs. Dr. Weber’s group has been working on developing an environmental fate simulator for novel and proposed munitions compounds. The co-adoption of ChemAxon tools by Dr. Weber’s group and our group facilitated collaboration and the integration of our results (predicted using QSARs calibrated with fully in silico methods) into the proposed fate simulator.
Figure 8. Screen shot of a spreadsheet-like view of the database, showing basic information about each compound.

Figure 9. “Form view” page for an individual compound (RDX) with data linked from the database component shown in Figure 8 and from another component containing kinetic data mined from the literature.
Figure 10. Results of a query for a structural moiety involved in the mechanism described above for the hydrolysis of RDX and related compounds. It is likely compounds in this set hydrolyze by a similar mechanism, so it will be logical to group these compounds in the same QSAR.

1.3.2 Calculation of target/response variable data for training set compounds with high-level theory (Task 1.2).

As reviewed for Tasks 1.1, there are several hydrolysis mechanisms for explosive compounds. The three most likely are substitution,

\[ R-\text{NO}_2 + \text{OH}^- \rightarrow R-\text{OH} + \text{NO}_2^- \] (3)

Meisenheimer addition where the hydroxyl attaches to a neighboring carbon atom in an aromatic ring,

\[ \text{Ar-NO}_2 + \text{OH}^- \rightarrow (\text{OH})\text{-Ar-NO}_2^- \] (4)

and the highly favorable $\beta$-elimination reaction for nitramine explosives.

\[ R-\text{CH-N-NO}_2 + \text{OH}^- \rightarrow R-\text{C=N} + \text{H}_2\text{O} + \text{NO}_2^- \] (5)

Examples of our calculated reaction free energies for a few of the nitroaromatics are shown in **Table 1** and **Table 2**. In general for the nitroaromatic compounds it was found that the overall thermodynamics of the substitution reactions was ~20 kcal/mol more favorable than the Meisenheimer reactions for the same compound. It was also found that Meisenheimer complexes were only thermodynamically stable when the nitroaromatic compounds contained 3 nitro groups (e.g. TNT, TNB, tetryl).
In our initial calculations, we attempted to determine the activation barriers of the substitution reactions of nitroaromatics by OH\(^-\) using a two-step process, where we first calculated the transition-state and activation barrier for the hydrolysis reaction in the gas-phase. Followed by calculation of solvent effects the activation barriers using continuum solvation models and QM/MM models. However, it turns out that the hydrolysis reaction.

Table 1. Free energies of reaction in kcal/mol for substitution reactions involving nitroaromatic compounds calculated at PBE, B3LYP, and PBE0 theories with COSMO solvation model.

<table>
<thead>
<tr>
<th>Substitution Reactions</th>
<th>PBE</th>
<th>B3LYP</th>
<th>PBE0</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNT + OH(^-) → TNT-2-OH + NO(_2^+)</td>
<td>-30.9</td>
<td>-33.5</td>
<td>-33.3</td>
</tr>
<tr>
<td>TNT + OH(^-) → TNT-4-OH + NO(_2^+)</td>
<td>-28.4</td>
<td>-30.4</td>
<td>-30.9</td>
</tr>
<tr>
<td>TNT + OH(^-) → TNT-1-OH-2-CH3 + NO(_2^+)</td>
<td>-28.8</td>
<td>-32.1</td>
<td>-31.1</td>
</tr>
<tr>
<td>TNT + OH(^-) → TNT-3-OH-4-H + NO(_2^+)</td>
<td>-28.7</td>
<td>-30.0</td>
<td>-30.1</td>
</tr>
<tr>
<td>TNT + OH(^-) → TNT-3-OH-2-H + NO(_2^+)</td>
<td>-29.5</td>
<td>-34.1</td>
<td>-33.7</td>
</tr>
<tr>
<td>NB + OH(^-) → NB-1-OH + NO(_2^+)</td>
<td>-21.4</td>
<td>-23.7</td>
<td>-23.8</td>
</tr>
<tr>
<td>TNB + OH(^-) → TNB-1-OH-2,3-NO(_2^+)</td>
<td>-24.6</td>
<td>-25.7</td>
<td>-25.1</td>
</tr>
<tr>
<td>TNB + OH(^-) → TNB-1-OH-2,4-NO(_2^+)</td>
<td>-34.6</td>
<td>-36.7</td>
<td>-35.9</td>
</tr>
<tr>
<td>TNB + OH(^-) → TNB-1-OH-2,5-NO(_2^+)</td>
<td>-31.4</td>
<td>-33.0</td>
<td>-32.5</td>
</tr>
<tr>
<td>TNB + OH(^-) → TNB-1-OH-2,6-NO(_2^+)</td>
<td>-27.4</td>
<td>-29.1</td>
<td>-28.5</td>
</tr>
<tr>
<td>TNB + OH(^-) → TNB-1-OH-3,4-NO(_2^+)</td>
<td>-27.8</td>
<td>-30.0</td>
<td>-30.1</td>
</tr>
<tr>
<td>TNB + OH(^-) → TNB-1-OH-3,5-NO(_2^+)</td>
<td>-31.1</td>
<td>-34.5</td>
<td>-33.6</td>
</tr>
<tr>
<td>DNA + OH(^-) → DNA-2-OH + NO(_2^+)</td>
<td>-25.1</td>
<td>-28.2</td>
<td>-28.1</td>
</tr>
<tr>
<td>TATB + OH(^-) → TATB-2-OH + NO(_2^+)</td>
<td>-24.8</td>
<td>-28.9</td>
<td>-28.1</td>
</tr>
<tr>
<td>Tetryl + OH(^-) → tetryl-2-OH + NO(_2^+)</td>
<td>-33.2</td>
<td>-36.3</td>
<td>-36.4</td>
</tr>
<tr>
<td>Tetryl + OH(^-) → tetryl-4-OH + NO(_2^+)</td>
<td>-34.2</td>
<td>-35.4</td>
<td>-34.8</td>
</tr>
<tr>
<td>DNAN + OH(^-) → DNAN-2-OH + NO(_2^+)</td>
<td>-24.4</td>
<td>-27.8</td>
<td>-26.6</td>
</tr>
<tr>
<td>DNAN + OH(^-) → DNAN-4-OH + NO(_2^+)</td>
<td>-14.1</td>
<td>-17.6</td>
<td>-16.0</td>
</tr>
</tbody>
</table>

Table 2. Free energies of reaction in kcal/mol for Meisenheimer addition reactions involving nitroaromatic compounds calculated at PBE, B3LYP, and PBE0 theories with COSMO solvation model.

<table>
<thead>
<tr>
<th>Meisenheimer Reactions</th>
<th>PBE</th>
<th>B3LYP</th>
<th>PBE0</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNT + OH(^-) → TNT-1-OH(^+)</td>
<td>-9.6</td>
<td>-3.8</td>
<td>-10.1</td>
</tr>
<tr>
<td>TNT + OH(^-) → TNT-3-OH(^+)</td>
<td>-7.9</td>
<td>-1.7</td>
<td>-8.5</td>
</tr>
<tr>
<td>NB + OH(^-) → NB-2-OH(^+)</td>
<td>13.7</td>
<td>19.7</td>
<td>13.5</td>
</tr>
<tr>
<td>NB + OH(^-) → NB-3-OH(^+)</td>
<td>14.2</td>
<td>14.9</td>
<td>14.3</td>
</tr>
<tr>
<td>NB + OH(^-) → NB-4-OH(^+)</td>
<td>15.0</td>
<td>14.2</td>
<td>14.0</td>
</tr>
<tr>
<td>TNB + OH(^-) → TNB-2-OH(^+)</td>
<td>-15.8</td>
<td>-11.2</td>
<td>-16.4</td>
</tr>
<tr>
<td>DNA + OH(^-) → DNA-1-OH(^+)</td>
<td>-0.5</td>
<td>5.1</td>
<td>-0.9</td>
</tr>
<tr>
<td>DNA + OH(^-) → DNA-3-OH(^+)</td>
<td>2.9</td>
<td>9.1</td>
<td>3.9</td>
</tr>
<tr>
<td>Tetryl + OH(^-) → tetryl-1-OH(^+)</td>
<td>-17.1</td>
<td>-10.7</td>
<td>-19.2</td>
</tr>
<tr>
<td>Tetryl + OH(^-) → tetryl-3-OH(^+)</td>
<td>-9.0</td>
<td>-7.8</td>
<td>-3.6</td>
</tr>
<tr>
<td>Tetryl + OH(^-) → tetryl-4-OH(^+)</td>
<td>-15.9</td>
<td>-12.7</td>
<td>-14.2</td>
</tr>
<tr>
<td>DNAN + OH(^-) → DNAN-1-OH(^+)</td>
<td>0.7</td>
<td>5.3</td>
<td>0.9</td>
</tr>
<tr>
<td>DNAN + OH(^-) → DNAN-3-OH(^+)</td>
<td>19.1</td>
<td>16.6</td>
<td></td>
</tr>
<tr>
<td>DNAN + OH(^-) → DNAN-5-OH(^+)</td>
<td>19.0</td>
<td>16.2</td>
<td>19.9</td>
</tr>
</tbody>
</table>
does not have an activation barrier in the gas-phase, because its overall reaction energy is extremely favorable (e.g. \( \text{C}_6\text{H}_5\text{NO}_2^-\text{(g)} + \text{OH}^-\text{(g)} \rightarrow \text{C}_6\text{H}_5\text{OH}\text{(g)} + \text{NO}_2^-\text{(g)} \), \( E_{\text{rxn}}(\text{B3LYP}) = -50 \text{ kcal/mol} \)).

Given difficulties with this approach, we had to use computationally more expensive strategies to determine these activation barriers. The approach that worked best was to map out the two-dimensional reaction energy surface in solution (i.e. with a continuum model) as a function of the R-NO\(_2\) and R-OH distances (Figure 11). This approach was considerably more expensive than our initial approach as it requires \( \sim 500 \) constrained geometry optimizations to compute each reaction energy surface. For example, at the DFT level, it takes \( \sim 150 \text{K CPU hours (0.73 CPU years) to calculate a surface for the reaction of TNT + OH}^-\). Using this approach, we have determined the reaction activation barriers for TNT, NB, 1,4-DNB, and 1,3-DNB (Table 4).

![Figure 11](image)

**Figure 11.** (a) Two dimensional potential energy surface for the reaction, \( \text{C}_6\text{H}_5\text{NO}_2\text{(aq)} + \text{OH}^-\text{(aq)} \rightarrow \text{C}_6\text{H}_5\text{OH}\text{(aq)} + \text{NO}_2^-\text{(aq)} \), calculated at the B3LYP/6-311++G(3p,3d) /COSMO level. The reaction barrier for this potential energy surface is \( E_{\text{act}} = +31 \text{ kcal/mol} \). (b) Illustration of an AIMD/MM simulation for the \( \text{C}_6\text{H}_5\text{NO}_2\text{(aq)} + \text{OH}^-\text{(aq)} \rightarrow \text{C}_6\text{H}_5\text{OH}\text{(aq)} + \text{NO}_2^-\text{(aq)} \) reaction and the two dimensional free energy surface of the reactant well obtained from the AIMD/MM metadynamics simulation. The reaction barrier for this potential energy surface is \( E_{\text{act}} = +30 \text{ kcal/mol} \).

**Table 4.** Activation barriers in kcal/mol for TNT, NB, 1,4-DNB, and 1,3-DNB from two dimensional potential energy surfaces calculated at the B3LYP/6-311++G(3p,3d)//COSMO level.

<table>
<thead>
<tr>
<th>Compound</th>
<th>( E_{\text{act}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNT</td>
<td>20 kcal/mol</td>
</tr>
<tr>
<td>NB</td>
<td>31 kcal/mol</td>
</tr>
<tr>
<td>1,4-DNB</td>
<td>15 kcal/mol</td>
</tr>
<tr>
<td>1,3-DNB</td>
<td>23 kcal/mol</td>
</tr>
</tbody>
</table>
We have also been using AIMD/MM (33) free energy approaches to calculate potential energy surfaces for OH− substitution reactions. In these types of simulations, it is sufficient to compute ΔG and ΔG‡ of reactions as just the Helmholtz free energy, ΔA, and activation barrier, ΔA‡, because the additional terms PΔV and PΔV‡ in ΔG and ΔG‡ respectively are usually very small. We used the metadynamics free energy sampling method with an AIMD/MM model (Figure 11) to determine the ΔA and ΔA‡ for the various OH−/NO2− exchange processes in nitro containing compounds. Metadynamics is a powerful, non-equilibrium molecular dynamics method that accelerates the sampling of the multidimensional free energy surfaces of chemical reactions. For the OH− substitution the reactions, these simulations can be performed on a modest workstation in about a week.

1.3.3 Calculate descriptor variable data for training set compounds using lower level theory (Task 1.3)

Energy-related descriptor variables such as ET, EHOMO, ELUMO, and EGAP were calculated using relatively-accessible methods. Calculations were performed on a desktop computer (MacMini, 2.66 GHz Intel Core 2 Dual processor) using Gaussian09 with the GaussView5 interface. These calculations were performed using density functional theory (DFT) with the B3LYP functional and 6-31G(d) basis set. These calculations were successful, but the efficient optimization of larger structures (e.g., CL-20) required greater computational resources than provided by the desktop system. In order to provide a less computationally expensive alternative, a recommended “research-level” model chemistry was also employed. In this method, B3LYP/6-31G energy calculations were performed on structures optimized using Hartree-Fock (HF) theory (HF/3-21G(d)).

A comparison of ELUMO values calculated from DFT and HF optimized structures is shown in Figure 12. The overall correlation (regression line) shows a near 1:1 relationship, with a slope of 0.996, intercept of 0.024, and scatter about the regression line. The significance of these variations (especially scatter) with respect to QSAR formulation and is discussed below.
1.3.4 Correlation Analysis and fitting QSARs (Task 1.4)

Two groups of compounds were selected for QSAR prediction of hydrolysis rate constants based on structural similarities likely to lead to similar mechanisms. These groupings are “nitramine-like” compounds (i.e., with a C-N(-NO$_2$) moiety) and nitroaromatic compounds, with our main focus being on the nitroaromatic compounds. The reaction of these compounds with hydroxide in water is well understood [51] and there are a variety of existing and emerging munitions compounds that fit this description, as well as a number of model compounds.

Although we had a large number of nitroaromatic compounds in our database (>60), we focused on compounds with a nitrobenzene backbone. We designated a handful of these compounds as “top-priority”. The selection of these compounds was based on factors including novelty, designation as a SEDP priority compound, structural similarity to novel or priority compounds, current usage, previous publication history, or availability for use in bench-top experiments (munitions and model compounds). We also further organized the list into “Priority I” and “Priority II” compounds, where the Priority I group contains munitions compounds and important/available model compounds, and the Priority II group contains a number substituted nitrobenzenes that have been well studied by the environmental chemistry community [11, 12, 53-55].

Our current list is shown in Table 3. We also plan to include 2,6-diamino-3,5-dinitropyrazine-1-oxide (LLM-105) in our study, although it is not a substituted nitrobenzene, because of its importance as an emerging munitions compound.

**Figure 12.** Comparison of DFT calculations of $E_{\text{LUMO}}$ from structures optimized using HF theory and DFT.
Table 3. List of priority compounds for QSAR development.

<table>
<thead>
<tr>
<th>Priority I</th>
<th>Priority II</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,4,6-Trinitroluene (TNT)</td>
<td>1,4-Dinitrobenzene (1-4-DNB)</td>
</tr>
<tr>
<td>2,4-Dinitroanisole (DNAN)</td>
<td>2,4-Diamino-6-nitrotoluene (2,4-DANT)</td>
</tr>
<tr>
<td>Nitrobenzene (NB)</td>
<td>2,6-Diamino-4-nitrotoluene (2,6-DANT)</td>
</tr>
<tr>
<td>1,3-Dinitrobenzene (1,3-DNB)</td>
<td>2-Methylnitrobenzene (2-CH3-NB)</td>
</tr>
<tr>
<td>2-Aminodinitrobenzene (2-ADNT)</td>
<td>2-Acetynitrobenzene (2-COCH3-NB)</td>
</tr>
<tr>
<td>3,5-Dinitroaniline (DNA)</td>
<td>2-Nitroaniline (2-NH2-NB)</td>
</tr>
<tr>
<td>1,3,5-Triamino-2,4,6-trinitrobenzene (TATB)</td>
<td>3-Methylnitrobenzene (3-CH3-NB)</td>
</tr>
<tr>
<td>Trinitrobenzene (TNB)</td>
<td>3-Acetynitrobenzene (3-COCH3-NB)</td>
</tr>
<tr>
<td>Tetryl</td>
<td>3-Nitroaniline (3-NH2-NB)</td>
</tr>
<tr>
<td>Musk ketone</td>
<td>4-COCH3-NB</td>
</tr>
<tr>
<td>Musk xylene</td>
<td></td>
</tr>
<tr>
<td>4-Amino-2,6-dinitrotoluene (4-ADNT)</td>
<td></td>
</tr>
</tbody>
</table>

1.3.5 Verification of QSARs with experimental values (Task 1.5)

We measured hydrolysis rates for verifying the predictability of potential QSARs. As a first step, we determined rate constants for the hydrolysis of TNT. This has allowed us to compare our values to those reported in the literature, and thereby check the accuracy of our methods. We will also determine rate constants for the hydrolysis of other nitroaromatic compounds with a nitrobenzene backbone (both energetic materials and model compounds). The results for the TNT experiments are shown in Figure 13 along with data reported in the literature [16, 19]. There is some scatter in the data. Part of the scatter in the data is likely due to the fact that Mills et al. [19] fit their concentration vs. time data to a reversible first-order model (which is appropriate for this reaction), whereas Emmrich [16] neglected to take into account the reversibility of the reaction. We have taken into account the reversibility of the reaction, yet our data seems to fall in between the data gathered by Mills et al. and Emmrich. The reason for this is not clear at this time.
Figure 13. (A) Arrhenius plot and (B) Eyring plot showing second order rate constants for alkaline hydrolysis of TNT. Both new data and data reported in the literature are shown. Arrhenius parameters as well as the enthalpy and entropy of the transition state ($\Delta H^\ddagger$ and $\Delta S^\ddagger$) were obtained from fitting the data to the appropriate model. The free energy for the transition state ($\Delta G^\ddagger$) at 25 °C was calculated from $\Delta H^\ddagger$ and $\Delta S^\ddagger$.

1.4 Conclusions

Previous efforts have failed to unambiguously define mechanisms for the reactions of TNT and DNAN with OH$^-$, especially in the case of TNT where product characterization has been particularly challenging. The experimental and computational observations reported here provide insight into these mechanisms, although some ambiguity remains, especially in the case of TNT. Because of this ambiguity, it is uncertain whether TNT and DNAN react by the same mechanism. The possible difference in mechanisms means that predicting the reaction mechanism for one based on the other may lead to unreliable predictions of environmental fate. This, along with uncertainties in the consistency of the calculated results with experimental values, presents a challenge for developing QSARs calibrated “in silico” that predict the hydrolysis behavior of the diverse range of energetic NACs.

2. Objectives 2 and 3—Nitro Reduction

The major results of Objective 2/3 were published in two manuscripts:


Abstract. The evaluation of new energetic nitroaromatic compounds (NACs) for use in green munitions formulations requires models that can predict their environmental fate. Previously invoked linear free energy relationships (LFER) relating the log of the rate
constant for this reaction (log(k)) and one-electron reduction potentials for the NAC ($E^1_{NAC}$) normalized to 0.059 V have been re-evaluated and compared to a new analysis using a (non-linear) free-energy relationship (FER) based on the Marcus theory of outer-sphere electron transfer. For most reductants, the results are inconsistent with simple rate limitation by an initial, outer-sphere electron transfer, suggesting that the strong linear correlation between log(k) and $E^1_{NAC}$ is best regarded as an empirical model. This correlation was used to calibrate a new quantitative structure-activity relationship (QSAR) using previously reported values of log(k) for non-energetic NAC reduction by Fe(II) porphyrin and newly reported values of $E^1_{NAC}$ determined using density functional theory at the M06-2X/6-311++G(2d,2p) level with the COSMO solvation model. The QSAR was then validated for energetic NACs using newly measured kinetic data for 2,4,6-trinitrotoluene (TNT), 2,4-dinitrotoluene (2,4-DNT), and 2,4-dinitroanisole (DNAN). The data show close agreement with the QSAR, supporting its applicability to other energetic NACs.


Abstract: Many redox reactions of importance in aquatic chemistry involve elementary steps that occur by single-electron transfer (SET). This step is often the first and rate limiting step in redox reactions of environmental contaminants, so there has been a great deal of interest in the corresponding one-electron reduction potentials ($E^1$). Although $E^1$ can be obtained by experimental methods, calculation from first-principles chemical structure theory is becoming an increasingly attractive alternative. Sufficient data are now available to perform a critical assessment of these methods—and their results—for two types of contaminant degradation reactions: dehalogenation of chlorinated aliphatic compounds (CACs) and reduction of nitro aromatic compounds (NACs). Early datasets containing $E^1$’s for dehalogenation of CACs by dissociative SET contained a variety of errors and inconsistencies, but the preferred datasets show good agreement between values calculated from thermodynamic data and quantum mechanical models. All of the datasets with $E^1$’s for reduction of NACs by SET are relatively new, were calculated with similar methods, and yet yield a variety of systematic differences. Further analysis of these differences is likely to yield computational methods for $E^1$’s of NAC nitro reduction that are similar in reliability to those for CAC dechlorination. However, comparison of the $E^1$ data compiled here with those calculated with a more universal predictive model (like SPARC) highlight a number of challenges with implementation of models for predicting properties over a wide range of chemical structures.

2.1 Introduction

This application of green chemistry principals [56] to the design and selection of new energetic compounds requires data on their chemical fate properties, often in the early stages of evaluation (including prior to synthesis) where experimentation with the candidate compounds is impractical [29, 57]. To overcome this obstacle, robust models are needed to predict key environmental fate properties, including rate constants for nitro reduction, which is a major determinant of their fate [58, 59]. The reliance of these
models on experimental input data must be limited, due to the challenges associated with performing experiments with novel and/or energetic compounds. For this reason, models such as linear free-energy relationships (LFERs) and quantitative structure-activity relationships (QSARs) utilizing calculated descriptor variables (e.g., determined using molecular modeling) are desirable.

2.2 Methods

Batch Experiments. Batch experiments were performed in 60-mL clear glass reaction vials capped with butyl rubber septa (Fisher Scientific). Each vial was prepared in an anoxic glove box (5% H₂ in N₂) and initially contained 49 mL phosphate buffer, 1 mL aqueous cysteine solution, and an aliquot (25–400 μL) of aqueous porphyrin stock solution. Vials were equilibrated to 25.0 °C in a water bath. After temperature equilibration, the specified NAC was introduced to the reaction vial by injecting 50 μL of a 0.1 M stock solution. 1 mL aliquots were removed over time and quenched by mixing with an equal volume of oxygen-containing (i.e., not deoxygenated) HPLC grade methanol. Quenched aliquots were analyzed by high performance liquid chromatography (HPLC) using a Varian ProStar 210 solvent delivery module, 410 autosampler, and 330 photodiode array detector, with a Platinum C18 5μ 250-mm x 4.6-mm column (Grace). For most analyses, the mobile phase consisted of 1:1 DI water and HPLC-grade methanol, the flow rate was 1 mL min⁻¹, and the detection wavelength was 254 nm.

Computational Methods. Values of \( E^1_{\text{NAC}} \) were determined from the free energy difference, \( \Delta G_{\text{rxn}} \), for the one-electron half reaction. These energy differences were calculated using electronic structure calculations with the COSMO continuum solvation model [36] using the NWChem program suite [30]. The electronic structure calculations were performed using density functional theory (DFT) [60] with the 6-311++G(2d,2p) basis set [61, 62] and the LDA [63], PBE96 [64], B3LYP [32, 33], PBE0 [65], and M06-2X [66] exchange correlation functions. Additional methods details are given in Bylaska et al. [67] and Salter-Blanc et al. [68].

2.3 Results

2.3.1 Formulation of reduction reaction pathways/mechanisms and mining previously-published data on the reactants and products (Task 2/3.1)

In order to develop models for predicting rates of NAC reduction, it is useful to consider the mechanism of the reduction reaction. In a substantial amount of previous work (e.g., [11, 12, 58, 69]), LFERs relating NAC reduction rate constants in the presence of a variety of reductants (\( k_{\text{red}} \)) to one-electron reduction potentials (\( E^1_{\text{NAC}} \)) according to the following model

\[
\log k_{\text{red}} = a \frac{E^1_{\text{NAC}}}{2.3RT/F} + b = a \frac{E^1_{\text{NAC}}}{0.059} + b
\]

were interpreted as evidence that the rate of NAC reduction is controlled by the rate of the initial electron transfer in an outer-sphere reaction (Equation 7).
More recently, evidence from compound-specific nitrogen isotope analysis of NAC reduction reactions has suggested that the rate limiting step in NAC reduction is either (i) the dehydration of $Ar\text{-N(OH)}_{2}$ to form $Ar\text{-NO}$, or (ii) proton or electron transfer to the nitroaromatic radical anion ($Ar\text{-NO}^\cdot$), depending on the reductant and solution chemistry [13, 70, 71]. These conclusions are broadly consistent with electrochemical studies of NAC reduction, which suggest that either dehydration or electron transfer to $Ar\text{-NO}^\cdot$ is rate limiting, depending on experimental conditions [72, 73]. In order to provide some reconciliation between the interpretation of LFERs involving NAC reduction and the more direct evidence from isotope fractionation and electrochemical studies, and to provide a more robust basis for defining predictive models, we revisited the interpretation of the correlation between $\log(k_{\text{red}})$ and $E_{1\text{NAC}}$ as evidence that initial outer-sphere electron transfer is rate limiting. This was done using both the model described in Equation 6 and a (non-linear) free energy relationship (FER) based on the Marcus theory of outer-sphere electron transfer [74], described for application to organic reactions by Eberson [75].

2.3.2 Calculation of target/response variable data for training set compounds with high-level theory (Task 2/3.2) 

Assessment of free-energy relationships (FER). In order to reevaluate and assess the use of FERs to describe NAC reduction rates, and to evaluate the consistency of these FERs with evidence regarding the rate-limiting step of the reactions, we compiled the major datasets for NAC reduction by homogeneous reductants (consisting of second order rate constants, $k_{\text{red}}$) and plotted the data vs. $E_{1\text{NAC}} / 0.059$ V according to Equation 6 (in the convention of previous studies) [11, 12, 58, 69] and vs. $\Delta G^\circ$ (in a Marcus-Eberson plot). The results are shown in Figure 14 A and B, respectively. The compiled data include $k_{\text{red}}$ for NAC reduction by the electron carrier/donor pairs juglone/hydrogen sulfide [11, 12], lawsone/hydrogen sulfide [11], and Fe(II) porphyrin/cysteine [11], as well as electrochemically-reduced 9,10-anthraquinone-2,6-disulfonate (AQDS) and Fe(II) complexed by the organic ligands tiron [76] and desferrioxamine B (DFOB) [77]. Throughout this work, we refer to these redox systems simply as juglone, lawsone, AQDS, Fe(II) porphyrin, Fe(II)-tiron, and Fe(II)-DFOB.
Figure 14. log($k_{\text{red}}$) for various NACs vs. (A) $E_{1\text{NAC}}^{1}/0.059$ V and (B) $\Delta G^\circ'$. Reductants include the electron carrier/donor pairs juglone/hydrogen sulfide [11, 12], lawsone/hydrogen sulfide [11], Fe(II) porphyrin/cysteine [11], electrochemically reduced AQDS [70], and the Fe(II)-ligand complexes Fe(II)-tiron [76] and Fe(II)-DFOB [77]. Closed symbols represent data for which the descriptor variable was determined from measured $E_{1\text{red}}^{1}$. Open symbols represent data for which the descriptor variable was determined from $E_{1\text{red}}^{1}$ values that were estimated from LFERs [11, 12]. Only data for which $E_{1\text{red}}^{1}$ was measured were included in the fits. The correlated data were fit to (A) Equation 6 or (B) the Marcus equation. In (B), an extrapolation of the fit is shown (as a dashed line) for cases where the data is well fit to the Marcus model.

The $E_{1\text{NAC}}^{1}$ data used in the preparation of Figure 14 were those reported in conjunction with the original $k_{\text{red}}$ datasets [11, 12, 76, 77]. Note that in some cases, the values of $E_{1\text{NAC}}^{1}$ were not measured, but instead were estimated based on LFERs for juglone and lawsone [11, 12]. We included these data in Figure 14 but excluded them from subsequent fitting. For Figure 14B, $\Delta G^\circ'$ was calculated for all NAC/reductant pairs. For this, it was determined that the electrostatic term differentiating $\Delta G^\circ'$ from $\Delta G^\circ$ (described in detail elsewhere) [75, 78, 79] was equal to zero in cases where the reductant carries a charge of -1 [75], and negligible in all other cases [78, 79]. Therefore, $\Delta G^\circ'$ was taken to be equal to $\Delta G^\circ$. $\Delta G^\circ$ was calculated from $E_{1\text{NAC}}^{1}$ and the one-electron reduction potential for the reductant ($E_{1\text{red}}^{1}$).

2.3.3 Calculate descriptor variable data for training set compounds using lower level theory (Task 2/3.3)

Measured data for $E_{1\text{NAC}}^{1}$ are often lacking or difficult to obtain, particularly for novel energetic NACs (e.g., due to lack of synthesized material or experimental hazards). To overcome this, it is desirable to use calculated values, such as those obtained using molecular modeling. Multiple efforts have been made to determine $E_{1\text{NAC}}^{1}$ using various computational techniques [67, 80-82]. We compiled and compared the results of these studies previously, in addition to presenting preliminary results from our own calculations of $E_{1\text{NAC}}^{1}$ [67]. In this work, we expanded on our previous results to include calculations using other exchange correlation potentials and applied these methods to an expanded set of compounds.

$E_{1\text{NAC}}^{1}$ were determined at five levels of theory, using the 6-311++G(2d,2p) basis set with the LDA, PBE, B3LYP, PBE0, and M06-2X exchange correlation functionals with solvation determined using the COSMO solvation model. In order to determine which model chemistry gives results that are most consistent with experimental data, a subset of the calculated $E_{1\text{NAC}}^{1}$ values were compared to a previously compiled set of measured $E_{1\text{NAC}}^{1}$ [83]. The measured validation dataset is given in Supporting Information of Salter-Blanc et al. [68], along with the corresponding computed values determined at the levels of theory used in this study. Deviations between the calculated and measured datasets were characterized by calculating the mean absolute deviation (MAD), root mean square deviation (RMSD), and the largest positive and negative deviation between the datasets.
Table 4. Calculated one-electron reduction potentials ($E_{\text{NAC}}^\dagger$) and comparison to a set of measured values. Calculated values were determined with density functional theory (DFT) using the 6-311++G(2d,2p) basis set with the specified exchange correlation functionals. Solvation was accounted for using the COSMO model.

<table>
<thead>
<tr>
<th>Cmpd</th>
<th>Measured</th>
<th>Calculated $E_{\text{NAC}}^\dagger$ (V)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Suatoni</td>
<td>LDA</td>
</tr>
<tr>
<td>NB</td>
<td>-0.486</td>
<td>-0.512</td>
</tr>
<tr>
<td>2-CH$_3$-NB</td>
<td>-0.590</td>
<td>-0.595</td>
</tr>
<tr>
<td>3-CH$_3$-NB</td>
<td>-0.475</td>
<td>-0.561</td>
</tr>
<tr>
<td>4-CH$_3$-NB</td>
<td>-0.500</td>
<td>-0.594</td>
</tr>
<tr>
<td>3-Cl-NB</td>
<td>-0.405</td>
<td>-0.272</td>
</tr>
<tr>
<td>4-Cl-NB</td>
<td>-0.450</td>
<td>-0.475</td>
</tr>
<tr>
<td>4-NH$_2$-NB</td>
<td>-0.568</td>
<td>-0.964</td>
</tr>
<tr>
<td>3-COCH$_3$-NB</td>
<td>-0.437</td>
<td>-0.439</td>
</tr>
<tr>
<td>4-COCH$_3$-NB</td>
<td>-0.356</td>
<td>-0.174</td>
</tr>
<tr>
<td>1,2-DNB</td>
<td>-0.287</td>
<td>0.033</td>
</tr>
<tr>
<td>1,3-DNB</td>
<td>-0.345</td>
<td>-0.339</td>
</tr>
<tr>
<td>1,4-DNB</td>
<td>-0.257</td>
<td>0.197</td>
</tr>
<tr>
<td>2,4-DNT</td>
<td>-0.397</td>
<td>-0.542</td>
</tr>
<tr>
<td>2,6-DNT</td>
<td>-0.402</td>
<td>-0.318</td>
</tr>
<tr>
<td>TNT</td>
<td>-0.253</td>
<td>-0.232</td>
</tr>
<tr>
<td>2-CHO</td>
<td>-0.355</td>
<td>-0.151</td>
</tr>
<tr>
<td>4-CHO</td>
<td>-0.322</td>
<td>-0.065</td>
</tr>
<tr>
<td>4-CH$_2$OH</td>
<td>-0.478</td>
<td>-0.554</td>
</tr>
<tr>
<td>2-ADNT</td>
<td>-0.417</td>
<td>-0.586</td>
</tr>
<tr>
<td>4-ADNT</td>
<td>-0.449</td>
<td>-0.707</td>
</tr>
<tr>
<td>2,4-DANT</td>
<td>-0.502</td>
<td>-0.870</td>
</tr>
<tr>
<td>2-Cl-NB</td>
<td>---</td>
<td>-0.458</td>
</tr>
<tr>
<td>2-COCH$_3$-NB</td>
<td>---</td>
<td>-0.424</td>
</tr>
<tr>
<td>DNAN</td>
<td>---</td>
<td>-0.517</td>
</tr>
<tr>
<td>$\text{MAD}^c$</td>
<td>0.161</td>
<td>0.207</td>
</tr>
<tr>
<td>$\text{RMSD}^d$</td>
<td>0.043</td>
<td>0.058</td>
</tr>
</tbody>
</table>

**Largest Positive Deviation**
- 0.454$^e$ 0.268$^e$ 0.316$^e$ 0.182$^e$ 0.140$^e$ 0.099$^e$ 0.078$^e$

**Largest Negative Deviation**
- -0.396$^f$ -0.493$^g$ -0.308$^g$ -0.374$^g$ -0.273$^g$ -0.104$^g$ -0.112$^g$

*a) Abbreviations: nitrobenzene (NB), 2-methylnitrobenzene (2-CH$_3$-NB), 3-methylnitrobenzene (3-CH$_3$-NB), 4-methylnitrobenzene (4-CH$_3$-NB), 3-chloronitrobenzene (3-Cl-NB), 4-chloronitrobenzene (4-Cl-NB), 4-nitroaniline (4-NH$_2$-NB), 3-acetylnitrobenzene (3-COCH$_3$-NB), 4-acetylnitrobenzene (4-COCH$_3$-NB), 1,2-dinitrobenzene (1,2-DNB), 1,3-dinitrobenzene (1,3-DNB), 1,4-dinitrobenzene (1,4-DNB), 2,4-dinitrotoluene (2,4-DNT), 2,6-dinitrotoluene (2,6-DNT), 2,4,6-trinitrotoluene (TNT), 2-nitrobenzaldehyde (2-CHO), 4-nitrobenzaldehyde (4-CHO), 4-nitrobenzyl alcohol (4-CH$_3$OH), 2-amino-4,6-dinitrotoluene (2-ADNT), 4-amino-2,5-dinitrotoluene (4-ADNT), 2,4-diamino-6-nitrotoluene (2,4-DANT). b) Dataset previously compiled by Phillips et al. [83]. c) Mean absolute deviation. d) Root mean square deviation. e) Deviant value for 1,4-DNB. f) Deviant value for 4-NH$_2$-NB. g) Deviant value for 2,4-DANT.
2.3.4 Task 2/3.4. Correlation analysis and fitting QSARs (Task 2/3.4)

In order to test the use of $E^{1}_{\text{NAC}}$ values calculated using the B3LYP* and M06-2X* methods as descriptor variables for QSARs, previously reported values of $k_{\text{Fe(II)P}}$ (also shown in Figure 14A and B) were correlated against the newly calculated $E^{1}_{\text{NAC}}$ values ($E^{1}_{\text{NAC, B3LYP*}}$ and $E^{1}_{\text{NAC, M06-2X*}}$). The resulting QSARs are shown in Figure 15A and B along with linear fits. The strongest correlation is seen when using the $E^{1}_{\text{NAC, B3LYP*}}$ values ($r^2 = 0.74$ vs. $0.55$ for $E^{1}_{\text{NAC, M06-2X*}}$). This is not surprising, since $E^{1}_{\text{NAC, B3LYP*}}$ most closely approximates the measured values of $E^{1}_{\text{NAC}}$ for the subset of compounds used in QSAR calibration (MAD = 0.037 V and RMSD = 0.002 V for $E^{1}_{\text{NAC, B3LYP*}}$ vs. MAD = 0.055 V and RMSD = 0.004 V for $E^{1}_{\text{NAC, M06-2X*}}$). The data show uniform and modest residuals, suggesting that the QSAR calibrated with the $E^{1}_{\text{NAC, B3LYP*}}$ values is an adequate representation of the data. The QSAR resulting from the linear fit of the data is given in Equation 7 (with standard deviations of the fitting coefficients in parentheses).

$$\log(k_{\text{Fe(II)P}}) = 12.4(\pm 2.6) \cdot E^{1}_{\text{NAC,B3LYP*}} + 5.8(\pm 1.1)$$

(7)

Figure 15. QSARs for NAC reduction by Fe(II)P calibrated using rate constants ($\log(k_{\text{Fe(II)P}})$) obtained from the literature [11] and $E^{1}_{\text{NAC}}$ calculated at (A) the B3LYP/6-311++G(2d,2p) level or (B) the M06-2X/6-311++G(2d,2p) level, both with an applied linear transformation based on correlations to a measured dataset (denoted as $E^{1}_{\text{NAC, B3LYP*}}$ and $E^{1}_{\text{NAC, M06-2X*}}$, respectively).

2.3.5 Verification of QSARs with experimental values (Task 2/3.5)

In order to validate the QSAR given in Equation 7 for predicting rate constants for energetic compounds, we measured rate constants for reduction of three energetic NACs (TNT, 2,4-DNT, and DNAN) by Fe(II)P/cysteine. Rate constants were also measured for
three non-energetic NACs (NB, 4-Cl-NB, and 1,3-DNB) in order to check the agreement between our results and the previously published $k_{\text{Fe(II)P}}$ data we used to calibrate the FER given in Equation 7.

In the absence of a comprehensive kinetic model for the cases that show deviation from first-order behavior, the initial portion of the data alone were fit to first-order kinetics (similar to the method of initial rates [84]). This treatment resulted in fitting one to three half-lives of NAC disappearance data. It is likely that the kinetics at the beginning of the experiment were dominated by simple reduction of the NACs, making this simplification a satisfactory approximation of the rate constant for that process. Details regarding data fitting for all the NACs evaluated are given in the Supporting Information of Salter-Blanc et al. [68]. The resulting log($k_{\text{Fe(II)P}}$) values are reported in Table 5.
Table 5. QSAR verification and validation data. $k_{Fe(II)P}$ were determined from experimental data.

<table>
<thead>
<tr>
<th>NAC</th>
<th>$k_{Fe(II)P}$ (L M$^{-1}$ s$^{-1}$)</th>
<th>log($k_{FeP}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NB</td>
<td>1.06 ± 0.16</td>
<td>0.026</td>
</tr>
<tr>
<td>4-Cl-NB</td>
<td>1.946 ± 0.083</td>
<td>0.286</td>
</tr>
<tr>
<td>1,3-DNB</td>
<td>16.6 ± 0.77</td>
<td>1.22</td>
</tr>
<tr>
<td>2,4-DNT</td>
<td>14.34 ± 0.76</td>
<td>1.156</td>
</tr>
<tr>
<td>TNT</td>
<td>112.1 ± 3.1</td>
<td>2.050</td>
</tr>
<tr>
<td>DNAN</td>
<td>17.50 ± 0.55</td>
<td>1.243</td>
</tr>
</tbody>
</table>

In order to verify that our methods of measuring $k_{Fe(II)P}$ produced values similar to the calibration dataset, log($k_{Fe(II)P}$) values for a few non-energetic compounds (NB, 4-Cl-NB, and 1,3-DNB) were appended to a plot of the calibration data from Figure 15 (shown in Figure 16) by plotting the data vs. the associated $E_{1NAC, B3LYP^*}$ value. The data agree well with the QSAR, thereby supporting the general agreement of our experimental data with the calibration data. In addition, the data for compounds overlapping with the calibration dataset (NB and 4-Cl-NB) show very good agreement with the corresponding individual data points. This suggests that the measured values of $k_{Fe(II)P}$ are generally quite accurate and that scatter observed in Figure 16 comes mostly from $E_{1NAC, B3LYP^*}$ or limitations of the model.

**Figure 16.** Comparison of newly collected log($k_{Fe(II)P}$) data to the calibration dataset and QSAR. Data for non-energetic NACs were used to verify agreement of the new data with the calibration data (method verification data). Data for energetic NACs were used to test the validity of the QSAR to energetic compounds (validation data).

Following method verification, log($k_{Fe(II)P}$) data for the energetic compounds tested (2,4-DNT, TNT, and DNAN) were appended to Figure 16 to test the validity of the QSAR to energetic compounds. The data point for TNT plots in line with the QSAR and the data.
points for 2,4-DNT and DNAN fall slightly outside the range of scatter, but is still within one log value. The good agreement of the data with the QSAR suggests that it is an accurate model for predicting reduction rates of energetic NACs by Fe(II)P.

### 2.4 Conclusions

In contrast to the hydrolysis pathway—where complexity in the mechanism and pathway prevented us from obtaining any broadly applicable QSARs for predicting degradation rates—the effect of complexity in the nitro reduction pathway was effectively captured by contrasting two approaches: theory-based correlations to $E^1$ and empirical correlations to $E_{LUMO}$. The former provides a variety of insights into the fundamental processes controlling nitro reduction and the latter provides simple QSARs that should be useful for estimating relative reduction rates of nitro aromatic compounds. The models were calibrated and/or tested with a variety of model and actual energetic compounds, including TNT and DNAN.

### 3. Objective 4—Amine Oxidation

The major results of Objective 4 were published in:


**Abstract:** New energetic compounds are designed to minimize their potential environmental impacts, which includes their transformation and the fate/effects of their transformation products. The nitro groups of energetic compounds are readily reduced to amines and the resulting aromatic amines are subject to oxidation and coupling reactions. Manganese dioxide (MnO$_2$) is a common environmental oxidant and model system for kinetic studies of aromatic amine oxidation. In this study, a training set of new and previously reported kinetic data for oxidation of model and energetic-derived aromatic amines was assembled and subjected to correlation analysis against descriptor variables that ranged from general purpose [Hammett sigma constants ($\sigma^-$), pKa’s of the amines, and energies of the highest occupied molecular orbital ($E_{HOMO}$)] to specific for the likely rate limiting step [one-electron oxidation potentials ($E_{ox}$)]. The selection of calculated descriptors (pKa, $E_{HOMO}$, and $E_{ox}$) was based on validation with experimental data. All of the correlations gave satisfactory quantitative structure-activity relationships (QSARs), but they improved with the specificity of the descriptor. The scope of correlation analysis was extended beyond MnO$_2$ to include literature data on aromatic amine oxidation by other environmentally relevant oxidants (ozone, chlorine dioxide, phosphate and carbonate radicals) by correlating relative rate constants (normalized to 4-chloroaniline) to $E_{HOMO}$ (calculated with a modest level of theory).

### 3.1 Introduction

Many explosives, pesticides, and pharmaceuticals are nitro aromatic compounds, and these compounds are readily reduced to aromatic amines under environmental conditions
Any source of aromatic amino compounds in the environment is a significant concern because this class of compounds is characterized by high carcinogenicity. Therefore, the green design of new materials involving nitro aromatic compounds must take into account the potential environmental effects of any aromatic amines that might arise from their nitro reduction. This is challenging for the development of new energetic materials, which should be designed for minimal environmental impact, but usually are not amenable to experimental measurement of environmental fate determining properties (because of restrictions related to security and safety).

The environmental fate of simple aromatic amines has been studied extensively from multiple points of view: biodegradation, sorption to soil, wastewater treatment, etc. In most of this work, the major transformation process is oxidation, some of which leads to polymerization and oxidative coupling to soil/sediment organic matter. Detailed studies of the oxidation of aromatic amines are available for manganese dioxide, chlorine dioxide, ozone, alkaline-activated persulphate, iron-activated hydrogen peroxide, acidified manganese dioxide, hydrogen peroxide catalyzed by poloxometalate, triplet excited-state organic sensitizers, and electrochemical oxidation. In this study, we used MnO₂ as the primary oxidant system because protocols are well established for its use in studies of aromatic amine oxidation and MnO₂ is relevant to the “natural attenuation” of aromatic amines in soils and groundwater, as well as in systems engineered for remediation. A comparison of results obtained with MnO₂ to other oxidants of aromatic amines (ClO₂, O₃, carbonate radical, phosphate radical, and triplet excited-state natural organic matter) is provided in the last part of this section on “cross-correlation” analysis.

3.2 Methods

Rate Constants. Rate constants for aromatic amine oxidation were measured in batch experiments using MnO₂ that was freshly synthesized using an adaptation of the methods described by Murray and Villalobos. The expected product from this procedure is mainly birnessite, which was confirmed by X-ray diffraction. The resulting precipitate was washed three times with deionized (DI) water and then suspended in pH 6, 10 mM carbonate buffer, with 0.1 M ionic strength (NaCl). For batch experiments, the MnO₂ solutions were used as prepared or diluted to the desired concentrations with additional carbonate buffer, placed in amber vials, and sealed. Vials were then injected with 50 μL of 10 mM aromatic amine stock solutions, shaken, and then stirred throughout the experiment. Periodically, 1 mL aliquots were removed and the reaction was quenched by removing MnO₂, either by filtration or reaction with ascorbic acid in a methanolic solution. Analysis of the samples was performed by HPLC with C-18 reverse-phase column and photodiode array detection. Quantification was performed from absorbance peaks at 234 nm.

Descriptor Data. Several types of descriptor data were calculated using methods that can be efficiently applied to most aromatic amines. Values of pKa were calculated using the pKa calculators available from ACD/Labs (as part of Percepta Predictors and SciFinder), ChemAxon (as part of InstantJChem and chemicalize.org) and ARChem (as part of SPARC). Values of EHOMO for the aromatic amines were calculated using Hartree-Fock.
(HF) or density functional theory (DFT) using Gaussian 09 (G09) [119] or NWChem [120] software. Free energies of oxidation were calculated for one-electron oxidation ($E_{ox}$) using NWChem [120]. For both sets of quantum chemical calculations, a representative range of functionals and solvation models were used and the results were compared.

Additional details regarding the experimental and computational methods used in this work are described in Salter-Blanc et al. [121]

3.3 Results

3.3.1 Formulation of amine oxidation reaction pathways/mechanisms and mining previously-published data on the reactants and products (Task 4.1)

**Rate Constants for Oxidation of Aromatic Amines.** To provide a training set of kinetic data for correlation analysis and QSAR development, two types of aromatic amines were selected: (i) model compounds (usually anilines with one non-reactive substituent) for which kinetic data are available from previous work on oxidation by MnO₂ [100, 101] and (ii) possible daughter products from nitro reduction of important energetic compounds [122, 123]. Compounds of the latter type are possible products of nitro reduction at individual nitro group substituents on the energetic nitroaromatic compounds 2,4-dinitroanisole (DNAN), 2,4-dinitrotoluene (2,4-DNT), 2,4,6-trinitrotoluene (TNT), and 3-nitroaniline. Where applicable, the products of nitro reduction at both favorable and unfavorable sites (based on computational predictions of regioselectivity) [122] have been included. Identifying information on all of the selected training set compounds is given in Table S1 in the Supporting Information of [121].

In the study of model aromatic amine oxidation by Klausen et al., [101] kinetic data were reported as relative rate constants ($k_{rel}$) using 4-chloroaniline as the reference compound for normalization. This convention is analogous to normalizing nitro reduction rate constants to 4-chloronitrobenzene, which has been done in many studies by Schwarzenbach et al. [12, 55, 69, 124] The main rationale for this type of normalization is that it should allow comparison of kinetic data obtained in different ways and/or under different experimental conditions (e.g., MnO₂ mass concentration or molarity, MnO₂ size or surface area, and pH), thereby enabling correlation analyses across more diverse data sets. For the training set of rate constants used in this study, we adopted the convention of normalizing to the 4-chloro substituted congener, and defined $k_{rel}$ according to Equation 8,

$$\log(k_{rel}) = \log \frac{k_{obs}}{k_{4Cl}}$$

where $k_{obs}$ is the observed rate constant for the compound of interest, and $k_{4Cl}$ is the observed rate constant for 4-chloronitrobenzene. In Klausen et al., [101] only $k_{rel}$ was reported and their data were used here without further adjustment. Laha and Luthy [100] reported initial disappearance rates, which we divided by the initial aniline concentration.

---

33
to obtain estimated values of $k_{obs}$, so their data could be used in Equation 8. The values of $k_{rel}$ for both studies are summarized in Table 6.

### Table 6. Normalized rate constants ($k_{rel}$) for aromatic amine oxidation by MnO$_2$.

<table>
<thead>
<tr>
<th>No.</th>
<th>IUPAC Name</th>
<th>$\log k_{rel}$</th>
<th>$\log k_{rel}$</th>
<th>$\log k_{rel}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Laha and Luthy$^a$</td>
<td>Klausen et al.$^b$</td>
<td>This Study$^c$</td>
</tr>
<tr>
<td>1</td>
<td>aniline</td>
<td>-0.93</td>
<td>0.48</td>
<td>-0.100</td>
</tr>
<tr>
<td>2</td>
<td>3-chloroaniline</td>
<td></td>
<td>-0.96</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4-chloroaniline</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>4</td>
<td>2-methylaniline</td>
<td>0.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3-methylaniline</td>
<td>0.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>4-methylaniline</td>
<td>0.34</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>2-methoxyaniline</td>
<td></td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>3-methoxyaniline</td>
<td>0.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>4-methoxyaniline</td>
<td>2.0</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>3-nitroaniline</td>
<td></td>
<td>-1.34</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>4-nitroaniline</td>
<td>-4.0</td>
<td>~ -4.11$^d$</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>2-methyl-5-nitroaniline</td>
<td></td>
<td>-1.40</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>4-methyl-3-nitroaniline</td>
<td></td>
<td>-1.20</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>2-methoxy-5-nitroaniline</td>
<td></td>
<td>-0.279</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>4-amino benzoic acid</td>
<td>0.45</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$ Calculated from $k_{exp}$ data reported in Laha and Luthy [100]

$^b$ Calculated from concentration vs. time data in Figure 8 of Klausen et al. [101]

$^c$ Details in Supporting Information of Salter-Blanc et al. [121].

$^d$ Approximate value because reaction was slow.

### 3.3.2 Calculation of target/response variable data for training set compounds with high-level theory (Task 4.2)

Only experiment data (from Task 4.1) were used.

### 3.3.3 Calculation of descriptor variable data with high-level and lower level theory (Task 4.3)

**Correlations Based on $E_{HOMO}$.** The highest occupied molecular orbital energies ($E_{HOMO}$) of compounds often work well as descriptors for correlation analysis of oxidation rates because oxidation general involves loss of electrons from the HOMO (to the acceptor/oxidant). $E_{HOMO}$ is a molecular descriptor that is usually calculated from chemical structure theory using commonly available software, but $E_{HOMO}$ is approximately equal to the negative of the ionization potential (by Koopman’s Theorem [125]), and the ionization potential (IP) can be measured by photoemission spectroscopy and related methods.

For this part of the study, we were able to compile experimental values of the adiabatic or vertical IP (VIP) for most of the aromatic amines in the training set, and the values are tabulated in Table S6 in the Supporting Information of [121]. Values of $E_{HOMO}$ were calculated for all of the training set compounds using Gaussian09 and NWChem.
software; HF, B3LYP or M06-2X functionals; 6-31G* or 6-311++G(2d,2p) basis sets; and COSMO, CPCM, PCM, or SMD solvation models. The quality of the \( E_{\text{HOMO}} \) data were evaluated by (i) plotting a matrix of correlations of all pairs of \( E_{\text{HOMO}} \) data sets (not shown because all of the correlations were nearly 1:1) and (ii) plotting the difference between calculated \( E_{\text{HOMO}} \) and measured VIP, as in Figure S2 in the Supporting Information of [121]. Six representative sets of \( E_{\text{HOMO}} \) data were chosen for inclusion in the figure, to reflect the full range of calculation conditions and results. The most significant dissimilarity among the results shown is that the differences are relatively small for \( E_{\text{HOMO}} \) calculated with HF (<0.5 eV), but large (1–2.5 eV) for \( E_{\text{HOMO}} \)'s calculated with DFT (M06-2X and B3LYP). This pattern has been reported previously [126], and is the reason that we included HF calculations in this study. Within the results for each DFT method, the variability among the differences are small (based on MAD, RMSD, and median), but they follow consistent and expected trends: e.g., the largest basis set (6-311++G(2d,2p)) gives the smallest differences.

**Correlations Based on \( E_{\text{o,ox}} \).** The putative rate limiting step for aromatic amine oxidation is loss of one-electron to give the corresponding aryl amino radical [102, 103, 127], so the free energy or potential of this half-reaction (\( \text{ArNH}_2 \rightarrow \text{ArNH}_2^+ + e^- \)) is a promising candidate descriptor variable for mechanistically-based QSARs. Data for these descriptors can be obtained experimentally—e.g., half-wave potentials (\( E_{1/2} \)) from electrochemical measurements—or by theoretical calculations. Measured values of \( E_{1/2} \) from Suatoni et al. [128] have been used as descriptors in other QSAR studies on the oxidation kinetics of substituted anilines [99-101] (and phenols [129-131]), but their data only cover a portion of the compounds included in the training set of this work. Therefore, we used the measured values of \( E_{1/2} \) from Suatoni et al. to verify our calculated values of one-electron oxidation potentials (\( E_{\text{o,ox}} \)) and then used \( E_{\text{o,ox}} \) for correlation analysis and QSARs. To obtain \( E_{\text{o,ox}} \), we first calculated the free energies of one-electron oxidation (\( \Delta G^0_{\text{o,ox}} \)) with NWChem and 10 combinations of functionals, basis sets, and solvation models. From the resulting values of \( \Delta G^0_{\text{o,ox}} \), six representative sets were selected and these were then used to calculate the final values of \( E_{\text{o,ox}} \) (in Volts vs. SHE), which are summarized in Table S8 in the Supporting Information of [121] along with the half-wave potentials (\( E_{1/2} \)) from Suatoni et al. [128].

As we did for the predicted/calculated values of \( pK_a \) and \( E_{\text{HOMO}} \), the quality of the calculated \( E_{\text{o,ox}} \) data were evaluated by (i) plotting a matrix of correlations of all pairs of \( E_{\text{o,ox}} \) data sets (again, not shown because all of the correlations were nearly 1:1) and (ii) plotting the deviation between \( E_{\text{o,ox}} \) and the measure values of \( E_{1/2} \) (Figure S3 in the Supporting Information of [121]). The main conclusion from the figure is that the trends among aromatic amines are consistent between the different levels of theory, and therefore all should give similar results when used as descriptor variables in QSARs. The \( E_{\text{o,ox}} \) reported here also correlate well (\( r^2 > 0.9 \), slopes ≈ 1.0, intercepts < 0.5 Volts; details not shown) with the one-electron oxidation potentials for aromatic amines reported previously by Winget et al. [132], Arnold [110], and Erickson et al. [109]. Careful inspection of the differences (individual, median, MAD, and RMSD) suggests that the absolute accuracy of \( E_{\text{o,ox}} \) calculated with PBE and B3LYP are slightly better than for M062X, and that the differences between COSMO and COSMO-SMD were negligible.
3.3.4 Correlation analysis and fitting QSARs (Task 4.4)

All 6 sets of $E_{\text{HOMO}}$ correlate well with log $k_{\text{rel}}$ and the resulting QSARs are summarized in Table S9 in the Supporting Information of [121]. Consideration of the fitting statistics and outlier distribution in all cases suggests that the 6 sets of $E_{\text{HOMO}}$’s are equally satisfactory as descriptor variables. Therefore, in the discussion that follows, we chose to feature a correlation to $E_{\text{HOMO}}$ obtained using methods (B3LYP/6-31G*/SMD) that are easily performed with common hardware and software (e.g., a personal computer running Gaussian09). The result, shown in Figure 17, is better than any of the correlations to $\sigma$ constants or pKa’s, and most of the unexplained variance (MAD, RSMD) is due to just 2-3 aromatic amines with the greatest residuals. One of these outliers is 4-amino benzoic acid (#15), which is easily rationalized because the anionic carboxylate group is likely to influence the affinity of this compound to the MnO$_2$ surface. Another outlier is 4-nitroaniline (#11), which has the greatest uncertainty of the values of $k_{\text{rel}}$ measured in this study, due to its slow reaction rate with MnO$_2$.

![Figure 17](image1.png)

**Figure 17.** Correlation of log $k_{\text{rel}}$ for oxidation of aromatic amines to $E_{\text{HOMO}}$. $E_{\text{HOMO}}$ was calculated using Gaussian09 with B3LYP, 6-31G*, and SMD. Fitting coefficients and statistics for the regression line given in the Supporting Information of [121]. Labels correspond to numbers in Table 6.

The relative success of the correlation shown in Figure 17 is unsurprising, given the abundance of QSARs based on $E_{\text{HOMO}}$ for oxidation of substituted phenols (e.g., [105, 133, 134]) and the strong analogy between substituted phenols and substituted anilines as model systems for structure-activity relationships. There are relatively few reported QSARs for substituted anilines based on $E_{\text{HOMO}}$, however, two are environmentally relevant and are notable because they bracket the result reported here with respect to the trade-offs between rigor and scope: Lee et al. [105] obtained more precise correlations using rate constants for a homogeneous model system (aqueous O$_3$) and optimal selection
of molecular orbital energies, whereas Colón et al. [99] reported a worse correlation using rate constants measured in aerobic sediment suspensions and $E_{\text{HOMO}}$’s calculated using semiempirical methods. These correlations to $E_{\text{HOMO}}$ are further evaluated in the final section on meta-analysis.

All 6 sets of $E_{\text{ox}}$ correlate well with $\log k_{\text{rel}}$ and the resulting QSARs are summarized in Table S9 in the Supporting Information of [121]. After consideration of the fitting statistics and outlier distribution in all cases (not shown), we concluded that the best correlation to $E_{\text{ox}}$ was obtained using the descriptor dataset calculated using M06-2X and 6-311++G(2d,2p). This correlation is shown in Figure 18, along with the regression results that define the QSAR. Overall, this correlation is better than those given above for $\sigma$ constants, pKa’s, or $E_{\text{HOMO}}$’s. In this case, most of the variance (MAD, RSMD) is due to just 2 aromatic amines with the greatest residuals, the overall confidence and prediction intervals are narrower, and the only clear outlier is 4-aminobenzoic acid. As noted above, the anomalous behavior of this compound is unsurprising given that its carboxylate group is likely to alter the interactions of this compound with the MnO$_2$ surface.

![Figure 18](image)

**Figure 18.** $\log k_{\text{rel}}$ for oxidation of aromatic amines to $E_{\text{ox}}$. Calculated with NWChem, M06-2X, 6-311++G(2d,2p), and COSMO. Fitting coefficients and statistics for the regression line given in the Supporting Information of [121]. Labels correspond to numbers in Table 6.

The correlation between $k_{\text{rel}}$ and $E_{\text{ox}}$ in Figure 18 can be compared to two previous studies that used $E_{\text{ox}}$ as the descriptor for aromatic amine oxidation. Arnold [110] obtained correlations with similar statistical quality, but a significantly larger slope, using rate constants for oxidation of a wider range of resonance stabilized N-containing organics by carbonate radical. A rationale for this difference in sensitivity (slope) is introduced in the next section. Colon et al. [99] obtained statistically less satisfactory...
correlations for aromatic amine oxidation using kinetic data obtained with a complex, natural material as the oxidant (suspensions of aerobic sediment). Using MnO2 as the oxidant, correlations that are qualitatively similar to Figure 18 were reported in the two previous studies and $E_{1/2}$ as the descriptor variable $[100, 101]$, and these correlations were interpreted as evidence that one-electron oxidation was the rate determining step in this reaction. Further interpretation of the slopes of these correlations in terms of electron-transfer theory would require converting $E_{ox}$ to free energies of aromatic amine oxidation by MnO2 $[68]$, but we were not able to take this step because the necessary one-electron reduction potential for MnO2 is not available.

3.3.5 Verification of QSARs with experimental values (Task 4.5)

Cross-Correlation and Meta-Analysis. The approach to correlation analysis represented by Figure 17 to Figure 18 is conventional in that it describes the reactivity of a homologous series of reactants (aromatic amines with various substitutions) by one reaction (oxidation by MnO2) under similar or comparable conditions (neutral pH, etc.). To determine how these results relate to other reactions or reaction conditions, it can be informative to perform “cross-correlation” analysis between the rate constants for one reaction (or set of reaction conditions) and the rate constants of another $[69, 134]$. This approach is less common than conventional correlation analysis, but it can be useful for estimation of missing rate constants and for diagnosis of reaction pathways or mechanisms. The best and most relevant example of this is the cross-correlation analysis of rate constants for oxidation of substituted phenols by environmentally significant oxidants ($^{1}$O$_2$, O$_3$, ClO$_2$, and HOCl), which gives strong linear relationships for most pairs of oxidants $[134, 135]$.

To extend this study to include cross- and meta-correlation analysis of aromatic amine oxidation, we found suitable published kinetic data for four oxidants: O$_3$, ClO$_2$, carbonate radical, and phosphate radical. Literature data for triplet excited state methylene blue $[109]$ and natural organic matter ($^{3}$NOM$^{+}$) $[110]$ were not included due to complicating factors or insufficient number of anilines, respectively. The data we compiled, which are summarized in Table S10 in the Supporting Information of $[121]$, were subjected to cross-correlation analysis, but the results (not shown) were of limited value because not enough compounds were common among the data sets. To overcome this, we performed a meta-analysis of the data sets by calculating $k_{rel}$ (with Equation 8, which requires $k_{4Cl}$, but this value was available in all the major datasets), and comparing their correlations to the best descriptor variable that can readily be calculated for any substituted aromatic amine ($E_{HOMO}$ obtained with the same conditions chosen for Figure 17). The results are shown in Figure 19, with the results from Figure 17 included, for comparison.
Figure 19. Comparison of correlations for normalized oxidation rates constants ($k_{rel}$) for aromatic amines versus MnO$_2$ (this study and prior literature) and previously published data for other oxidants. The data and fit for MnO$_2$ (gray) are from Figure 17. The data for other oxidants are compiled in in the Supporting Information of [121], and the fit of these data is described in the text.

The meta analysis represented by Figure 19 suggests three types of trends in the reactivity of substituted aromatic amines. First, the data for chlorine dioxide [136, 137] and phosphate radical [138, 139] exhibit no correlation to $E_{HOMO}$, which is surprising, but the number of amines are few, especially for the phosphate radical data, and the original reports of these data provide only limited structure-reactivity analysis. Second, the data for ozone [105, 140] and carbonate radical [141-143] show a very strong correlation to $E_{HOMO}$ for most amines, with a small number of outliers that are similar to those that tend to deviate from other correlations (e.g., nitro substituents, etc.). Fitting the ozone and carbonate radical data (pooled, without outliers) gives the new regression line (slope 1.54 ± 0.11, intercept = 8.57 ± 0.64, $r^2 = 0.85$) and narrow confidence intervals shown in Figure 19. For comparison, data for MnO$_2$ (compiled from this and previous studies) conform satisfactorily to a correlation based on $E_{HOMO}$, with a slope (6.07 ± 0.79) that suggests a significantly (roughly 5-fold) greater sensitivity to substituent effects than for ozone and carbonate radical. One reason for this difference in slope is likely to be that MnO$_2$ is the only heterogeneous oxidant included in the comparison, and interfacial steric or electronic effects on surface complexation might compound the effect of substituents on reactivity. Another possibility is that the difference in slope is the effect of differences in potential of the oxidants that is predicted by electron transfer theory, which would be analogous to the interpretation we gave to data on nitro reduction kinetics in the predecessor to this study [68].

The concentration of data near the fitted line for ozone and carbonate radical data in Figure 19 undoubtedly indicates these data are of high quality, but probably also reflects the relatively simple substitutions of the anilines included in these data sets (mostly
single methyl, methoxy, or halogen substituents) and homogeneity of the oxidant solutions. The training set used in our study includes anilines derived from energetic compounds— which often include multiple nitro groups, and therefore are more challenging for correlation analysis—and heterogeneous oxidants. Heterogeneous systems, in particular, tend to produce complex kinetics, which is true for our results and the two previous studies using suspensions of MnO₂ as oxidant. Another study done in even more heterogeneous systems (aerobic sediment suspensions) [99] also gave complex kinetics, and a correlation to \( E_{\text{HOMO}} \) with a relatively high degree of scatter. For these reasons, and because there were few substituted anilines in common with the other studies included in this meta-analysis, we have not included them in Figure 19. However, qualitative analysis of the data is consistent with relatively shallow slope, similar to the correlation for ozone and carbonate radical, and significantly less than the slope of the correlation for MnO₂.

### 3.4 Conclusions

The ultimate goal of this meta-analysis is to determine if there is sufficient generality to QSARs derived for individual, model oxidants that useful predictions can be made about the reactivity of nitro reduction products of new energetic materials under environmental conditions. Clearly, predicting absolute rates of oxidation is not within the scope this work, but relative rates (e.g., as represented by \( k_{\text{rel}} \)) vary with aniline substitution in a qualitatively similar way for most oxidants. These trends predict that oxidation (and therefore oxidative coupling and immobilization) of putative reduction products of energetic compounds will be relatively slow (one-order magnitude or more) compared with aniline and most common model substituted anilines.

### CONCLUSIONS AND IMPLICATIONS

An overarching objective of this project was to demonstrate that much of the QSAR calibration process could be done with data obtained from computational chemistry; hence the project title “fully in silico calibration”. This could be significant because some emerging energetics (e.g., insensitive munition compounds (IMCs)) are not readily available for experimental determination of environmental fate determining properties, and therefore calibration data derived in silico can be the only option.

With respect to hydrolysis, the first step in the QSAR development process (defining the reaction that determines the rate constant) turned out to be a greater challenge than expected. All prior experimental and theoretical work on degradation of nitro aromatic compounds by hydrolysis was compiled and shown to be inconclusive with respect to the reaction pathways and products. New experimental and theoretical work performed as part of this project clarified the most favorable hydrolysis pathways, but there still was too much complexity to this reaction for QSARs to reliably describe the reaction rates over a significant variety of IMCs.

The reduction of nitro groups is one of the most widely studied environmental contaminant transformation pathways and the most important to energetic compounds, including IMCs. Kinetic data on nitro reduction were compiled and analyzed using a theoretically based QSAR model, which allows one model to describe rates of nitro
reduction by a variety of environmental reductants. Using new experimental data, and newly calculated reduction potentials of IMCs and related nitro compounds, a QSAR was developed that can be used to predict nitro reduction rates of other IMCs.

Using the QSAR for nitro reduction, relative rate constants were estimated for a large number of energetic nitro compounds, and then representative IMCs were superimposed, for comparison. Figure 1 shows that many IMCs will have relatively slow rates of nitro reduction, in some cases more than an order of magnitude slower than TNT. This comparison illustrates how QSAR models could be used to guide the design of IMCs for minimal environmental impact.

Figure 20. Correlation between the rates of nitro reduction for various energetics compounds ($k$), relative to TNT, versus the descriptor variable that provides the best overall QSAR ($E^1$). The right axis shows rate constants converted to approximate half-lives to give. The blue symbols are experimental data, obtained with IMC related amino compounds.

With respect to oxidation of amines, previously reported kinetic data were compiled and combined with new experimentally measured rate constants using several IMC related aromatic amines. To develop QSARs with these data, many possible descriptors were evaluated, ranging from simple to advanced. The simple descriptors gave adequate correlations, but the advanced descriptors gave the best QSARs. The best QSAR was used in predictive mode to compare the relative sensitivity of different aromatic amines to oxidation. In this analysis, most of the IMCs relate amines were relatively more susceptible to oxidation, which might mean that these reduction products will bind more readily to soils, and therefore be less mobile.
REFERENCES


49


APPENDICES

4. Supporting Data
No Supporting data provided.

5. List of Scientific/Technical Publications (Chronological)


6. Other Supporting Materials
No other supporting materials provided.