

Experimental Versus Predicted Heats of Reaction for Some Common Reaction Types in Pharmaceutical Industry

Katherine L. Young, Joseph Humes, Caitlin Justice, and Ray Mentzer*

Cite This: https://do	oi.org/10.1021/acs.oprd.2c00263	Read	Online	
ACCESS	III Metrics & More		E Article Recommendations	

ABSTRACT: Common pharmaceutical reactions were modeled using the commercially available CHETAH program and the novel TCIT program to obtain heats of reaction values. The programs were used to predict thermochemical data for isolated molecules based on their functional groups. Both programs use group theory, CHETAH, commercially available through ASTM and TCIT using a more modern approach and under development at Purdue. Reactions such as amide coupling (via both uronium reagent and T3P), debenzylation, bromination, *tert*-butyloxycarbonyl deprotection, Suzuki couplings, and halogenation were analyzed. These values were then compared to experimental values to assess the accuracy of each program's predictions. Experimental values were provided by several pharmaceutical companies, who participated and advised in this endeavor, as listed in Acknowledgements. Deviations were on average within $\pm 20\%$ of experimental values for both programs, but outlier behavior with respect to certain functional groups was distinct. CHETAH calculations were at times hampered by missing groups, which was not the case for TCIT, which is designed to be arbitrarily extensible. This project resulted in a broader molecular library for TCIT, exposure of shortcomings in CHETAH, and strategies to improve accuracy moving forward.

KEYWORDS: heats of reaction, CHETAH, TCIT, process safety, reaction hazards

INTRODUCTION

Knowledge of the heat a chemical reaction produces is vital to any industry for process scale-up.¹ If a reaction makes a small amount of heat at lab scale, then it will make significantly more heat when scaled up by an order of magnitude or more. Additionally, heat of reaction prediction can be used as a screening tool to direct resources for hazard investigations or toward the replacements of particularly hazardous chemistry. Thus, some industries, especially the pharmaceutical industry due to their constantly evolving product line, rely on heat of reaction data when planning the manufacture of a product. When a new reaction is discovered or synthetic pathway is achieved, it is advantageous to model how much heat is given off during the reaction without resorting to experimental measurements. Experimental methods can be problematic in early stages of development due to limited availability of materials for testing. Limiting or eliminating the need for them would also be advantageous in the interest of cost and time required before production can begin.

The heat of reaction is commonly used in the thermodynamic analysis of a reaction.^{2,3} Essentially, it measures how much heat is given off or taken in during a reaction. Heat of reaction (ΔH_r) values are mathematically simple to calculate: the heat of formation of the reactants are subtracted from the heats of formation of the products. Thus, to quantify ΔH_r one needs the heat of formation of the products and reactants.

Two programs were used to model the heat of reaction for comparison with measured values: CHETAH and TCIT.⁴ CHETAH (Chemical Thermodynamic and Energy Release) is a commercially available program marketed by ASTM that has been used for the last two decades to analyze chemical reactions.⁵ TCIT is a novel program by Dr. Savoie's group at Purdue University. Both programs predict a variety of thermochemical properties with different methodologies employed by each program.

LITERATURE REVIEW

The merits of predicting heats of reaction are plentiful in the literature. Weisenburger et al. note that reaction heat can be





used to determine how much the temperature will rise in a reaction chamber.⁶ The resulting heat produced must be removed and managed safely. This is especially important during scale-up calculations, where a bench-top reaction is scaled to plant operation size. There are several well-

Received: August 19, 2022



Figure 2. With TCIT, the heat of formation is calculated by breaking apart the molecule and analyzing the adjacent atoms, as well as the atoms two positions away.⁷

documented cases of reactor vessel's contents getting too hot resulting in a runaway reaction, without remediation this risks the viability of surrounding equipment and safety of human lives from a potential explosion.⁶ Predicting heats of reaction is not only a possible time and money saver but also a valuable procedure to assess the likelihood of safety concerns of the operation.

One of the ways to predict heats of reaction is to use a group theory. The basis of such a theory is relatively simple: separate the molecule(s) of interest into component parts, in which each part has a known heat of formation. The heats of formation of the various parts are then added together to get an estimate of the heat of formation of the molecule.^{6,7} A diagram explaining the breakdown of a molecule can be seen in Figure 1.

Considering the molecule above, the way in which groups would be related to neighboring groups is determined by the group theory in use for a particular calculation. CHETAH uses Benson group increment theory, which draws on programmed groups experimentally determined. However, Benson group theory looks at a specific atom, as well as the atom(s) or group(s) directly next to it, or at one adjacent position. Each group has an associated value, namely group additivity values, which can be added together to determine the heat of formation of the molecule.⁸ Problematically, if the group has not been quantified and/or programmed into CHETAH, the system indicates a substitution to an existing group must be made for the calculation to be complete.

Topology automated force-field interactions (TAFFI) component increment theory (TCIT) relies on component theory. The difference from group theory is in which atoms are analyzed and how the fundamental "building blocks" are calculated.⁸ In Figure 1, the atoms analyzed are two positions away from the atom of interest. This has several major benefits. First, contributions to a single atom from a greater portion of the molecule is considered, increasing the accuracy. Second, there are few molecules in which TCIT cannot characterize.⁷ This means that any user error inherent to picking Benson groups is eliminated in TCIT. A visual depiction of the method of calculation can be seen in Figure 2.

Figure 2 indicates the use of a Gaussian-4 (G4) database, which is the training data for deriving group values in TCIT. G4 is a composite quantum chemistry computational model that relies on post-Hartree–Fock methods and basis set extrapolation to calculate the heat of formation of a molecule.^{7,9} In TCIT, G4 calculations are only performed for small molecules during the group derivation.⁷ For large molecules, the component groups are added together to produce a total value for the heat of formation of the molecule of interest.

However, TCIT automatically includes ring corrections when executing calculations.^{7,10} In CHETAH, a ring correction is not included and must be chosen from a list of available Benson groups.^{1,8} This introduces an element of error in CHETAH calculations, as the results are heavily dependent on the Benson groups chosen in the model. TCIT and CHETAH both have limitations on the molecules they can be applied to (e.g., charged and free-radical species as well as fused rings cannot be inputted).

METHODS

It was known through past work that both programs could not manage some specific types of compounds for analysis. This

Tal	ble	1.	Samp	le	Calcu	lation	for	Sc	heme	1
-----	-----	----	------	----	-------	--------	-----	----	------	---

Molecule	Program	Heat of Formation (-kJ/mol)
S_	CHETAH	-45.76
	TCIT	-42.318
	СНЕТАН	406.82
H_2O_2	TCIT	330.448
0 <u></u>	СНЕТАН	322.367
S.	TCIT	301.493
	СНЕТАН	571.66
H ₂ O	TCIT	514.936
Products - Reactants	CHETAH	532.967
	TCIT	528.299





included ionic compounds (generally salts), charged species, radicals, fused rings, and various functional groups (such as nitro groups).^{1,7,8} To analyze these compounds, simplifications had to be made to the molecule in question, such that the molecule could be analyzed. If the molecule was charged, for example, a possible simplification would be to input the protonated species into the program, such that the charge was no longer present. Similarly, if a fused ring were present, a simplification would be to eliminate one atom of a ring, such that there were two alkyl chains. Or if that were not enough simplification for the program, one ring could be eliminated entirely. These simplifications add a layer of inaccuracy to the calculations.

The steps for analyzing each compound were relatively simple. The molecule in question would be drawn in ChemDraw and would be converted to a SMILES string (simplified molecular-input-line-entry).^{11,12} Once the molecules were inputted into the respective programs, the analysis would begin. In most reactions, reagents were in the liquid phase, which meant a correction would be needed. TCIT will





automatically calculate the heat of vaporization, which would be subtracted from the gas-phase heat of formation to yield the liquid-phase heat of formation.7 CHETAH was more complicated, as oftentimes there was no liquid-phase enthalpy data available and manual corrections were performed.^{1,8,} Frurip, a codeveloper of CHETAH, published these guidelines to do so: for nonstrongly interacting species were 0.42 kJ/mol/ Da, hydrogen bonding species were 0.63-0.84 kJ/mol/Da, and halogenated species were 0.275 kJ/mol/Da (Da is dalton or atomic mass unit).^{7,8,13} Thus, the correction was simply the factor (determined by the structure of the molecule), multiplied by the molecular weight, then subtracted from the gas-phase value.¹³ Another common correction made which will be denoted as "acid/base" indicates that an acid base reaction occurred after the formation of the products and needed to be added to the overall heat output. A maximum heat of formation value of -56 kJ/mol for the reaction of a strong acid and base was used as the correction factor. This value was used as it is commonly found in literature for the reaction of a strong acid and base.¹⁴ It is theoretically possible to quantify the individual acid-base reactions occurring or approximate them to something more reasonable, but with the intent of the work being the validity of the programs for safety purposes, a maximum value was used on every reaction.

One sample calculation can be found in Table 1 using Scheme 1, a sulfide to sulfone reaction.

RESULTS

The data collected must first be qualified. In many instances, the reaction schemes provided by the pharmaceutical

Table 3. Amide	Coupling	Reactions	via	Uronium	Reagents
Data Measured	and Calcu	lated			

rxn #	measured $\Delta H_{ m rxn} \ (-kJ/mol)$	$\begin{array}{c} \text{CHETAH} \\ \Delta H_{\text{rxn}} \\ (-\text{kJ/mol}) \end{array}$	TCIT $\Delta H_{\rm rxn}$ (-kJ/mol)	% diff CHETAH	% diff TCIT
1	159	126	131	21	18
2	174	149	155	14	11
3	159	161	139	1	13
4	127	137	141	7	11
5	123	125	110	2	11
6	136	173	150	27	10
6	174	173	150	1	14
7	207	173	161	16	22

companies included an "R" group. This is a section of the molecule that does not participate in the reaction and had to be hidden for intellectual property purposes. The group removed could contain any number of groups and atoms, but typically, a methyl or ethyl group was inserted unless otherwise directed by the company representative. Despite the fact that those parts of the molecule do not participate in the reaction,

Table 4. Debenzylation Reactions



Table 5. Debenzylation Data Measured and Calculated

rxn #	measured $\Delta H_{ m rxn} \ (-{ m kJ/mol})$	$\begin{array}{c} \text{CHETAH} \\ \Delta H_{\text{rxn}} \\ (-\text{kJ/mol}) \end{array}$	$\begin{array}{c} \text{TCIT } \Delta H_{\text{rxn}} \\ \left(-\text{kJ/mol}\right) \end{array}$	% diff CHETAH	% diff TCIT
1 ²	116	119	103	3	11
2	114	91	89	20	22
3	349	245	267	30	24
4	210	198	189	6	10

they still contribute to the heats of formation, but are likely similar in reactants and products, so should not have an appreciable net effect. Additionally, the heat of reaction values were determined using various types of equipment (e.g., EasyMax or RC-1) and methodology; thus, there could be error introduced in the data collection methods as well. Because of the accumulating layers of error, the industry representatives concluded that a calculation falling within a 20% margin of error would be considered successful for risk assessment purposes.

The first series of reactions studied were amide couplings using either the HATU, HBTU, or TBTU (uronium coupling agents) reactants to facilitate the formation of amides. Reactions 1–3 are HATU; 4 and 5 are HBTU, and 6–7 are TBTU. Nearly all reactions required acid/base corrections. Additionally, CHETAH results were converted from gas values to liquid using the methodology mentioned previously.¹³ All measured heats of reaction are in the liquid phase, except for reaction 3 in Table 2 which is a gas phase CHETAH value calculated by a company expert. By recalculating the value in CHETAH, the method employed for CHETAH calculations throughout this work can be validated relative to work completed by an expert. Reaction 6 was measured by more than one company which employed different methods to measure the heat of reaction which demonstrates the innate variability in the experimental data.

Per Tables 2 and 3 for various amide-coupling classes, several notable trends arise. For the first two HATU amidecoupling reactions, TCIT performed better, but both programs were reasonably close to the experimental values. For the third reaction, which was not experimentally measured but calculated with CHETAH, the CHETAH value found was nearly equal to the value found by a company. This was a good indication of the viability of values calculated in CHETAH. For the TBTU amide coupling reactions 6 and 7, CHETAH performed the best, except for one of the experimental values for reaction 6.

The next set of reactions analyzed were debenzylations. The mechanism and molecular structures were relatively simple, and there were no acid/base corrections to account for. The reactions and calculations are shown in Tables 4 and 5.

Reactions 1 and 2 had reasonable agreement relative to each other and to the experimental value. Reactions 3 and 4 had repeated units; so, it was assumed for simplicity reasons that the molecules only had one of the repeated units. No attempt was made to make the predicted results better match the experimental results by altering the number of units, although the reaction 4 predictions are quite good for both programs. These reactions only served as a benchmark to learn some of the intricacies of both programs, but an investigation into polymer chain effects on heat of formation/reaction could be completed using these programs.

The bromination reaction results are shown in Tables 6 and 7. There were several structures that one or both programs were incapable of analyzing.

This set of reactions had several difficulties. First, Nbromosuccinimide required a Benson group substitution for

Table 6. Bromination Reactions



Table 7. Bromination Data Measured and Calculated

rxn #	measured $\Delta H_{\rm rxn} \ (-{\rm kJ/mol})$	CHETAH ΔH _{rxn} (–kJ/mol)	$\begin{array}{c} \text{TCIT } \Delta H_{\text{rxn}} \\ (-\text{kJ/mol}) \end{array}$	% diff CHETAH	% diff TCIT
1	79	93	73	18	7
2	106	81	128	24	21
3	79	88	74	12	7
4	136	205	122	51	10

chlorine in CHETAH because of the extremely limited number of halogen containing groups available. Next, the fused ring in reaction 2 had to be simplified to one ring (altered reaction shown below original), as neither program could analyze fused rings, resulting in ~20% difference from measured values with both programs. The nitro group in reaction 4 was converted to the reduced version (an amine group, NH₂), as nitro groups have charged oxygens in them. An amine drastically changes the electronics of the ring, given that an amine is an electron donor, and a nitro group is an electron-withdrawing group, resulting in the introduction of more error into the calculation. Despite the corrections made, the TCIT results were all reasonable compared to the experimental values; however, the CHETAH results suffered more significantly from the substitutions made.

The next series of reactions analyzed were from a published paper by Weisenburger et al.⁶ There were 49 reactions presented, along with experimental heat of reaction values. The reactions analyzed were chosen based on available groups in TCIT and CHETAH. For example, a number of reactions that produced ionic species were avoided, as both programs were incapable of analyzing those groups. Furthermore, fused rings were generally avoided, except for reaction 2 (49 in the publication), which was simplified by opening one of the rings (such that the product is a bi-substituted ring). The results from the reactions can be seen below in Table 8.

Both programs were challenged with these reactions, as shown in Table 9. From these results, there are several important conclusions. An epoxide ring correction was made for reaction 1 external to the software based on direction from the experts on the significant heat it produces. Another point to note is in reactions 6 and 7, in which both the TCIT and CHETAH values significantly differ from the experimental values. One possible explanation is that the experimental values could be improperly indicative of the heat of reaction. The thought is that intermolecular forces (mainly hydrogen

Table 8. Pfizer Reactions⁶



Table 9. Pfizer Data Measured and Calculated⁶

rxn #	measured $\Delta H_{ m rxn} \ (-kJ/mol)$	$\begin{array}{c} \text{CHETAH} \\ \Delta H_{\text{rxn}} \\ (-\text{kJ/mol}) \end{array}$	$\begin{array}{c} \text{TCIT } \Delta H_{\text{rxn}} \\ \left(-\text{kJ/mol}\right) \end{array}$	% diff CHETAH	% diff TCIT
1	88	144	123	64	40
2	149	162	139	8	7
3	7	16	7	68	6
4	566	584	567	3	0.1
5	122	93	88	24	28
6	23	73	82	221	260
7	38	77	67	101	76
8	310	209	285	32	8

bonding by the nitrogen and keto groups present) inhibit the molecule, resulting in a lower (more endothermic) heat of reaction than would otherwise be observed. This is difficult to prove, as this theory is largely dependent on the solvent system used. Obviously, a polar solvent such as water, alcohol, acetonitrile, or acetone would amplify the hydrogen bonding effects more than a nonpolar solvent (hexane, chloroform, pentane, etc.).

The next set of reactions analyzed were BOC deprotection (deBoc) reactions. These reactions are useful in the pharmaceutical industry, especially on an important functional group. The proposed mechanism can be seen in Scheme 2.

pubs.acs.org/OPRD

From the reaction scheme, there are many products that form from the deprotection. The amine is generated, as well as carbon dioxide. It is important to note that one of the products not mentioned is isobutene, which comes off as a gas with carbon dioxide.¹⁶ It is also important to know that the alcohol formed depends on the source of hydrogens present. If there is just water, t-BuOH is formed. However, if the hydrogen source is an alcohol, a tertbutyl-ether could form. Thus, the BOCdeprotection reaction products are heavily dependent on the solvent system. The reaction schemes and resulting notes for each reaction can be seen in Table 10, whereas the results from the analysis can be seen in Table 11.

From Table 10, the only reaction in which solvent effects must be considered is in reaction 1. That reaction was run in a mixture of ethanol and MTBE. The ethanol reacts with the tert-butyl group to form the ether. However, the rest of the reactions were run in a nonnucleophilic solvent (DCM, EtOAc, MTBE/PhCN), negating the ether product.

There are several notes to be made in the deBoc reactions. First, in every deBoc reaction, an acid is used to catalyze the reaction; thus, the TFA, MSA, or HCl is assumed to not have an impact on the heat of reaction. The last major facet of deBoc reactions is the gas product. The initial predicted values were significantly different from the experimental values because the carbon dioxide and isobutene gas leaving the system removes heat from the reaction (acting as a pseudoheat exchanger). Thus, a method of correcting for the gaseous product was required. It was assumed that all gaseous species were initially formed as a liquid. The heat of vaporization was added to the heat capacity (at an arbitrary temperature) for each gas, then combined to form the correction factor.¹⁹⁻²¹ In this reaction set, the TCIT values agreed with the experimental results relatively well, except for reaction four. However, the CHETAH results are drastically different. After collaboration with CHETAH SMEs, the exact cause of this difference remains uncertain. The only accepted thought being that there could have been side products generated.

The next set of reactions that were analyzed were Suzuki-Miyaura coupling. This reaction is especially useful in synthetic chemistry, in which the goal is to make a carbon-carbon bond. The reaction mechanism for a Suzuki coupling is shown in Figure 3.

The Suzuki mechanism in Figure 3 makes note of several products that form. There is the acid formed from the alkyl halide (1), the coupled alkene (2), and organoborane product (3). However, this scheme does not demonstrate the solvent effects. If the reaction is run under nonprotic conditions, this mechanism is likely to occur. However, if the reaction is run in a protic system, then the boron product that forms will be boric acid. Of the Suzuki reactions that were analyzed, only reaction 1 used a nonprotic solvent. The rest had boric acid as a product, in addition to the generated acid and coupled carbons. The results from the Suzuki coupling reactions can be seen in Tables 12 and 13.

Scheme 2. Mechanism for a BOC Deprotection Reaction^a

^aNote that R can be any carbon chain.¹⁵

Table 10. DeBoc Reactions^{17,18}

Rxn #	Reaction	Comments
1	$\wedge OH + \bigwedge_{H} \stackrel{O}{\longrightarrow} \stackrel{HCI}{\longrightarrow} \stackrel{HCI}{\longrightarrow} + \bigwedge_{NH_2} + \bigwedge_{O} \stackrel{HCO_2}{\longleftarrow} + CO_2$	Assumed that the ether is produced by reaction with ethanol solvent, R_1 -ethyl group, and added 2 acid/base corrections to account for dihydrochloride salt formation. Used heat of formation of CO ₂ and isobutylene in gas phase ^{17,18}
2	$CF_3 \longrightarrow OH$ $CF_3 \longrightarrow OH$ $TFA \longrightarrow CF_3 \longrightarrow CF_3 \longrightarrow R_1 + CO_2 + M_3N^{-R_1} + CO_2 + M_3N^{-R_1}$	R_1 -methyl group. One acid/base correction, and used heat of formation of CO_2 and isobutylene in gas phase ^{17,18}
3	$(H_3 \longrightarrow OH) (H_3 \longrightarrow OH$	R_1 -methyl group. One acid/base correction, and used heat of formation of CO_2 and isobutylene in gas phase ^{17,18}
4	$ \underbrace{\overset{O}{\overset{H}{\overset{H}{\overset{H}{\overset{H}{\overset{H}{\overset{H}{\overset{H}$	One acid/base correction, and used heat of formation of CO_2 and isobutylene in gas phase ^{17,18}

Table 11. DeBoc Data Measured and Calculated

rxn #	$\begin{array}{c} \text{measured} \\ \Delta H_{\text{rxn}} \\ (-\text{kJ/mol}) \end{array}$	CHETAH ΔH _{rxn} (–kJ/mol)	TCIT $\Delta H_{\rm rxn}$ (-kJ/mol)	% diff CHETAH	% diff TCIT
1	138	-234	141	270	2
2	52	-501	56	198	7
3	73	-40	54	155	26
4	31	-55	-61	277	299

There are several things to note within these sets of reactions. Both TCIT and CHETAH agreed reasonably well with reactions 1, 2, and 4, but poorly with reaction 3. There are two possible reasons for the latter discrepancy. The first is that there are unexplained side reactions occurring, which would make the measured value not indicative of the heat of reaction for the proposed reaction. The other explanation is that the heat of dosing (Q_{dos}) data has not been factored into the predicted value. The heat of dosing value is a critical consideration when reagents are added at different temperatures and bring a certain amount of additional energy to a reaction, such that the measured heat of reaction can be affected. In a sense, a calorimeter could be measuring the heat of reaction *in addition* to the energy introduced by the reagents



Figure 3. Reaction mechanism for a Suzuki coupling. Make note of the palladium catalyst, as well as the cyclical nature of the mechanism. The reaction products are labeled 1-3.²²

added at a temperature different than the reaction temperature. Thus, to get the true heat of reaction value, Q_{dos} must be subtracted from the measured value. Correcting this is an ongoing effort and may be part of the future direction of related work.

Table 12. Suzuki Reactions



Table 13. Suzuki Data Measured and Calculated

rxn #	$\begin{array}{c} {\rm measured} \\ \Delta H_{\rm rxn} \\ (-{\rm kJ/mol}) \end{array}$	$\begin{array}{c} \text{CHETAH} \\ \Delta H_{\text{rxn}} \\ (-\text{kJ/mol}) \end{array}$	TCIT $\Delta H_{\rm rxn}$ (-kJ/mol)	% diff CHETAH	% diff TCIT
1	252	206	234	18	7
2	193	174	224	10	16
3	120	151	210	23	75
4	356	267	309	25	13

The next set of reactions analyzed all contained propanephosphonic acid anhydride (T3P) reagent for peptide formation.

Examining the findings in Tables 14 and 15 closely, several conclusions can be drawn. First, CHETAH and TCIT were both within the 20% margin for all reactions. However, CHETAH was not able to model the T3P reagent, so literature values were used to complete the calculations.^{23,24} Additionally, CHETAH was not able to complete the heat of formation calculation for reagents containing a nitrogen close to a benzene ring which required substitutions in reactions 1 and 3.

Another important note is that agreement between CHETAH and TCIT was excellent for this set of reactions except in the third reaction which had significant CHETAH substitutions, as previously mentioned.

The final set of reactions analyzed were halogenation reactions which can be found in Tables 16 and 17.

The reaction complexities with these reactions were the heat of dissolution of the SO_2 because it would remain dissolved in the solvent. Literature values were used to approximate those interactions to the exact solvent or as similar as possible.²⁷ TCIT modeled the three reactions well. However, CHETAH was unable to work with the nitrogen substituted ring in reaction 1, so a 5 membered ring with one unsaturation was used instead. Finally, CHETAH struggled with the last reaction because it does not have bromine in the system; thus, CHETAH is only suitable for chlorine-based halogen reactions.

DISCUSSION AND CONCLUSIONS

In terms of user experience, there are significant differences between CHETAH and TCIT. CHETAH requires more user

pubs.acs.org/OPRD

Table 14. Amide Coupling Reactions via T3P Reagent



Table 15. Amide Coupling Reactions via T3P Reagent Data Measured and Calculated

rxn #	measured ΔH _{rxn} (-kJ/mol)	CHETAH ΔH _{rxn} (–kJ/mol)	TCIT ΔH_{rxn} (-kJ/mol)	% diff CHETAH	% diff TCIT
1	178	150	209	16	18
2	189	188	186	0.5	2
3	150	152	177	2	18
4	198	186	185	6	7

expertise in the event of required substitutions, whereas TCIT requires no in-program corrections to be made. However, TCIT requires some level of programming abilities, whereas

CHETAH is more user friendly. Therefore, it is much quicker to input and receive values from CHETAH.

Compiling all % errors from the 39 reactions studied, TCIT was able to come within the desired 20% margin of error on 74% of them. On the other hand, CHETAH was within 20% on 54% of reactions. To better understand if the programs were consistently under or overpredicting heats of reaction, parity plots were developed, as shown in Figure 4.

Looking at the parity plot for CHETAH, it tends to underestimate the heat of reaction for the reactions with higher heats of reaction, TCIT to a lesser extent. Considering the process safety implications, this would be an important thing to

Article

Table 16. Halogenation Reactions^{25,26}

Rxn #	Reaction	Comments
1	$\begin{array}{c c} & R & \\ HO & N & \\ & N & N & \\ & N & N & \\ & H & \\ & H & \\ & H & \\ & H & \\ & & N & \\ & & H & \\ & & H & & H & \\ & & & H & \\ & & & H & \\ & & H & & H & \\ & $	CHETAH: Literature values used for SOCl ₂ , HCl and SO ₂ . ²⁵ Reactant was substituted for 5 member ring with 1 unsaturation. Both: acid/base and heat of dissolution corrections.
2	$N \longrightarrow OH + SOCI_2 \longrightarrow N \longrightarrow CI + HCI + SO_2$	CHETAH: Literature values used for SOCl ₂ , HCl and SO ₂ . ²⁵ Both: acid/base and heat of dissolution corrections.
3	\sim_{N} \rightarrow_{HBr} \rightarrow_{Br} \rightarrow_{HBr}	CHETAH: Literature values used for SOBr ₂ , HCl and SO ₂ . ²⁶ Both: acid/base and heat of dissolution corrections.

Table 17. Halogenation Data Measured and Calculated

rxn #	measured ΔH _{rxn} (-kJ/mol)	$\begin{array}{c} \text{CHETAH} \\ \Delta H_{\text{rxn}} \\ (-\text{kJ/mol}) \end{array}$	TCIT ΔH_{rxn} (-kJ/mol)	% diff CHETAH	% diff TCIT
1	162	145	165	10	2
2	196	203	161	4	18
3	126	84	131	33	4

note. Generally, TCIT tends to slightly underpredict the heats of reaction, but not as significantly as is seen with CHETAH.

Overall, TCIT more consistently met the 20% error criteria with no substitutions needed. Any major errors can generally be accounted for due to an incomplete understanding of the reactions. The most consistent correction needed was the acid base. Alternatively, CHETAH was much less capable of handling the complex reactions studied, but with more experience with the program, and more reasonable substitutions can perhaps be made. This program can give a reasonable first approximation if the molecules can be reduced to the simplest form of the components taking part in the reaction.

Throughout this study, there have been a variety of observed effects on the measured heat of reaction. For example, in the deBoc reactions, gases leaving the system can take energy away, causing the heat of reaction to appear lower than the true value. Furthermore, there can be side reactions (i.e., acid/base reactions, side-couplings) that can change the measured value. In some reactions of the Pfizer paper, intermolecular forces are thought to entrap the molecules, thereby impacting the measured heat of reaction. In addition, in the Suzuki reactions, there was heat of dosing affects that had to be considered. Overall, there is often more to an experimental heat of reaction than just the difference in the heats of formation of the products and reactants. There are many complex interactions that impact the experimental value. By continuing to expand the understanding of these interactions, the models can be better utilized to provide a more realistic depiction of the heat of reactions.

As for the future of this work for TCIT, we want to continue to find limitations in its database so it is able to reliably calculate heats of formation for virtually any molecule. Adding an ionic and free-radical data base would be a major accomplishment on that front. Lastly, CHETAH can continue to be studied to find the most reasonable group substitutions. If there are additions to the Benson groups currently available, this work could be quickly reevaluated, but with the



Figure 4. Parity plots for CHETAH and TCIT, respectively.

methodology employed it would likely not be as effective as

AUTHOR INFORMATION

Corresponding Author

Ray Mentzer – Purdue Process Safety and Assurance Center, Davidson School of Chemical Engineering, Purdue University, West Lafayette, Indiana 47906, United States; Orcid.org/ 0000-0002-7108-9175; Email: rmentzer@purdue.edu

Authors

TCIT.

- Katherine L. Young Purdue Process Safety and Assurance Center, Davidson School of Chemical Engineering, Purdue University, West Lafayette, Indiana 47906, United States; orcid.org/0000-0001-9735-5366
- Joseph Humes Purdue Process Safety and Assurance Center, Davidson School of Chemical Engineering, Purdue University, West Lafayette, Indiana 47906, United States
- Caitlin Justice Purdue Process Safety and Assurance Center, Davidson School of Chemical Engineering, Purdue University, West Lafayette, Indiana 47906, United States

Complete contact information is available at:

https://pubs.acs.org/10.1021/acs.oprd.2c00263

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors would like to acknowledge and profoundly thank the group of pharmaceutical representatives that enabled this work to be possible not only with their data but also with their insights and expertise. They are Roy Flanagan (GSK), Jeffrey Sperry (Vertex), Thomas Vickery (MSD), Min Sheng (Corteva), Derek Brown (Amgen), Ayman Allian (Eli Lilly) Han Xia (Eli Lilly), Li-Gen Ping (JM), and Sai Sathanapally (JM). Additionally, the authors would like to thank Professor Brett Savoie and Qiyuan Zhao for their support in using TCIT.

REFERENCES

(1) Grewer, T.; Frurip, D. J.; Keith Harrison, B. Prediction of Thermal Hazards of Chemical Reactions. *J. Loss Prev. Process Ind.* **1999**, *12*, 391–398.

(2) Grewer, T.; Klusacek, H.; Löffler, U.; Rogers, R. L.; Steinbach, J. Determination and assessment of the characteristic values for the evaluation of the thermal safety of chemical processes. *J. Loss Prev. Process Ind.* **1989**, *2*, 215–223.

(3) Stoessel, F. Thermal Safety of Chemical Processes Risk Assessment and Process Design; Wiley-VCH: Weinheim, 2020.

(4) Downey, J. R.; Frurip, D. J.; LaBarge, M. S.; Syverud, A. N.; Grant, N. K.; Marks, M. D.; Harrison, B. K.; Seaton, W. H.; Treweek, D. N.; Selover, T. B. ASTM Data Series DS 51B, Chetah 11.0. The ASTM Computer Program for Chemical Thermodynamics and Energy Release Evaluation (NIST Special Database 16), Philadelphia, PA, 1994.

(5) The ASTM Computer Program for Chemical Thermodynamic and Energy Release Evaluation—Chetah Version 11.0. https://www. a s t m . o r g / d s 5 1 h o l - e b . h t m l # : ~ : t e x t =

The%20CHETAH%C2%AE%20(Chemical%20Thermodynamic (accessed Oct 20, 2022).

(6) Weisenburger, G. A.; Barnhart, R. W.; Clark, J. D.; Dale, D. J.; Hawksworth, M.; Higginson, P. D.; Kang, Y.; Knoechel, D. J.; Moon, B. S.; Shaw, S. M.; Taber, G. P.; Tickner, D. L. Determination of Reaction Heat: A Comparison of Measurement and Estimation Techniques. Org. Process Res. Dev. **2007**, *11*, 1112–1125. (7) Zhao, Q.; Savoie, B. M. Self-Consistent Component Increment Theory for Predicting Enthalpy of Formation. J. Chem. Inf. Model. 2020, 60, 2199–2207.

(8) Benson, S. W.; Buss, J. H. Additivity Rules for the Estimation of Molecular Properties. Thermodynamic Properties. *J. Chem. Phys.* **1958**, *29*, 546–572.

(9) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Gaussian-4 Theory. J. Chem. Phys. 2007, 126, 084108.

(10) Zhao, Q.; Iovanac, N. C.; Savoie, B. M. Transferable Ring Corrections for Predicting Enthalpy of Formation of Cyclic Compounds. J. Chem. Inf. Model. **2021**, *61*, 2798–2805.

(11) O'Boyle, N. M. Towards a Universal SMILES Representation - a Standard Method to Generate Canonical SMILES Based on the InChI. J. Cheminf. 2012, 4, 22.

(12) Evans, D.; Rubenstein, S. ChemDraw; PerkinElmer, 1985.

(13) Curtiss, L. A.; Redfern, P. C.; Frurip, D. J. Theoretical Methods for Computing Enthalpies of Formation of Gaseous Compounds. *Rev. Comput. Chem.* **2007**, *15*, 147–211.

(14) Roux, A.; Perron, G.; Picker, P.; Desnoyers, J. E. Enthalpies of Reaction and Reaction Rates by Flow Microcalorimetry: Ester Hydrolysis in Basic Medium. *J. Solution Chem.* **1980**, *9*, 59–73.

(15) Wang, J.; Liang, Y.-L.; Qu, J. Boiling Water-Catalyzed Neutral and Selective N-Boc Deprotection. *Chem. Commun.* 2009, 34, 5144.

(16) Li, B.; Li, R.; Dorff, P.; McWilliams, J. C.; Guinn, R. M.; Guinness, S. M.; Han, L.; Wang, K.; Yu, S. Deprotection of N-Boc Groups under Continuous-Flow High-Temperature Conditions. *J. Org. Chem.* **2019**, *84*, 4846–4855.

(17) Argonne National Laboratory. Isobutene Active Thermochemical Tables. https://atct.anl.gov/Thermochemical%20Data/ version%201.118/species/?species_number=862.

(18) Cox, J. D.; Wagman, D. D.; Medvedev, V. A. CODATA Key Values for Thermodynamics; Hemisphere Publishing Corporation: New York, 1989.

(19) Stephenson, R. M.; Malanowski, S. Handbook of the Thermodynamics of Organic Compounds; Springer Netherlands: Dordrecht, 1987.

(20) Chase, M. W., Jr. NIST-JANAF Thermochemical Tables, Fourth Edition. J. Phys. Chem. Ref. Data, Monogr. 1998, 9, 1–1951.

(21) Thermodynamics Research Center. Selected Values of Properties of Chemical Compounds; Thermodynamics Research Center, Texas A&M University: College Station, TX, 1997.

(22) Suzuki, A. Organoborane Coupling Reactions (Suzuki Coupling). Proc. Jpn. Acad., Ser. B 2004, 80, 359-371.

(23) PPACAlC9H21O6P3lChemSpider. http://www.chemspider. com/Chemical-Structure.100355.html (accessed Oct 20, 2022).

(24) 1,5-Dihydroxy-1,3,5-tripropyltriphosphoxane 1,3,5-trioxidel C9H23O7P3lChemSpider. http://www.chemspider.com/Chemical-Structure.74087828.html?rid=2df05ad2-09a4-4389-baad-989c1b4d80d7 (accessed Oct 20, 2022).

(25) Claydon, A. P.; Mortimer, C. T. 619. Heats of Formation and Bond Energies. Part VIII. Diethylaminotrimethylsilane, NN'-Dithiodiethylamine, NN'-Thionylbisdiethylamine, and NN'-Sulphurylbisdiethylamine. J. Chem. Soc. **1962**, 3212–3216.

(26) Finch, A.; Gardner, P. J.; Radcliffe, K. Thermodynamic Functions of Thionyl Bromide. *J. Chem. Eng. Data* **1968**, *13*, 176–177.

(27) Benoit, R. L.; Milanova, E. Vapour Pressure and Calorimetric Data for the Solution of Sulfur Dioxide in Aprotic Solvents. *Can. J. Chem.* **1979**, *57*, 1319–1323.