Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Mechanistic and kinetic insights into secondary phenol formation from cresol and toluene: A combined quantum chemical and machine learning study

Trang T. T. Hoang
Hanoi Architectural University
Tien V. Pham
tienxp@gmail.com

Hanoi University of Science and Technology

Research Article

Keywords: Cresol, OH radical, ML, RRKM, CCSD(T)

Posted Date: November 17th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11607815/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Mechanistic and kinetic insights into secondary phenol formation from cresol and toluene: A combined quantum chemical and machine learning study

Trang T. T. Hoang,[1] Tien V. Pham[2,*]

[1]Department of Chemistry, Hanoi Architectural University, Hanoi, Vietnam

[2]School of Chemistry and Life Sciences, Hanoi University of Science and Technology, Hanoi, Vietnam

Abstract

Context In this study, the reaction mechanisms and kinetics of four pathways leading to phenol + CH_3 formation were systematically investigated, including the reactions of *o*-, *m*-, and *p*-cresol with H atoms and the reaction of toluene with OH radicals. The results show that the addition of H atoms to the three cresol isomers requires activation energies of 7.6, 8.4, and 8.6 kcal/mol, respectively, leading to the formation of the corresponding intermediates IS1, IS2, and IS3. In the toluene + OH reaction, a pre-reactive complex is formed before the system proceeds via a submerged transition state to generate IS4. All four intermediates subsequently undergo CH_3 elimination to yield the same bimolecular product, phenol + CH_3. The transition-state energies predicted by the machine learning (ML) model are in good agreement with the corresponding quantum chemical results, with deviations of less than 1 kcal/mol. The calculated rate coefficients indicate that the *o*-cresol + H reaction is the fastest among the four pathways over the investigated temperature and pressure ranges. In addition, the rate coefficients obtained for the toluene + OH reaction agree well with the previous theoretical results of Wu et al., with a mean absolute relative deviation of 6.3%. These findings provide insights into the competing pathways for phenol formation and their temperature- and pressure-dependent kinetics, with implications for understanding aromatic oxidation and combustion chemistry.

Methods The geometries of all reactants, intermediates, transition states, and products were optimized at the B3LYP/6-311++G(3df,2p) level of theory. Single-point energies were subsequently calculated using the CCSD(T) method with the same basis set. The ML model was employed to predict transition-state energies and evaluate their agreement with the quantum chemical calculations. Temperature- and pressure-dependent rate coefficients were calculated using RRKM/master equation (ME) theory, as implemented in the MESMER software package. All quantum chemical calculations were performed using Gaussian 16.

Keywords: Cresol, OH radical, ML, RRKM, CCSD(T)

Corresponding author: Tien V. Pham, email address: tienxp@gmail.com

1. INTRODUCTION

Phenol (C_6H_5OH) is an important aromatic compound widely detected in the atmosphere and associated with both natural and anthropogenic emission sources, including biomass burning, fossil-fuel combustion, and industrial activities.[1,2] In addition to its direct emission, phenol can also be produced through the atmospheric oxidation of other aromatic volatile organic compounds (VOCs). The formation and subsequent oxidation of phenolic compounds are of particular interest because these processes can influence the chemical composition of the atmosphere and contribute to the formation of secondary organic aerosols (SOA) and other oxidation products.[3,4] Recent atmospheric observations and chemical transport modeling have further highlighted the importance of aromatic precursors, particularly toluene, in the secondary production of phenol, suggesting that direct emissions alone may not fully account for its atmospheric abundance.[3] Therefore, identifying and characterizing the chemical pathways responsible for secondary phenol formation is essential for improving our understanding of aromatic oxidation chemistry.

Among the potential aromatic precursors, cresol and toluene are of particular interest because both contain a methyl substituent attached to a benzene ring and can participate in radical-initiated reactions under atmospheric and combustion conditions. Cresol exists in three structural isomers, namely *o*-, *m*-, and *p*-cresol, whose different methyl-group positions may influence their reaction mechanisms and kinetic behavior. Toluene, on the other hand, is one of the most abundant aromatic VOCs in urban environments and undergoes atmospheric oxidation primarily through reactions with hydroxyl (OH) radicals.[4] Previous experimental studies have demonstrated that the OH-initiated oxidation of toluene can produce cresols as major ring-retaining products, while a dealkylation pathway leading directly to phenol has also been identified, with a reported phenol yield of 5.4 ± 1.2% under the investigated conditions.[5] These findings provide experimental evidence that toluene can serve as a precursor to phenol through more than one chemical pathway. In addition, reactions involving H atoms and substituted aromatic compounds have been recognized as important ipso-substitution processes in the pyrolysis and oxidation chemistry of aromatic species.[6,7] In particular,

the reaction of H atoms with cresol can lead to the replacement of the methyl substituent by hydrogen, thereby forming phenol and CH_3 radicals.

Despite these advances, a detailed understanding of the elementary reaction mechanisms and kinetic characteristics of the individual pathways leading to phenol remains important. Previous theoretical investigations have established kinetic parameters and rate rules for H-atom-induced ipso-substitution reactions involving a range of mono- and disubstituted aromatic compounds, including cresol isomers.[6] Other studies have developed kinetic models for substituted phenolic compounds and their reactions under pyrolysis and combustion conditions.[7] The OH-initiated oxidation of toluene has also been extensively investigated through experimental and theoretical approaches, with particular attention given to the formation of aromatic adducts, peroxy radicals, and phenolic products.[4,8] Nevertheless, further systematic analysis is needed to clarify the detailed addition–elimination mechanisms associated with the specific phenol + CH_3 product channel, particularly the energetic characteristics of the intermediate complexes and transition states. Moreover, a direct comparison of the three cresol isomers reacting with H atoms and toluene reacting with OH radicals within a consistent theoretical framework would help elucidate the effects of precursor structure and radical identity on phenol formation. The temperature- and pressure-dependent kinetics of these competing pathways are also important for evaluating their relative kinetic significance under different reaction conditions.

Besides, machine learning (ML)[9,10] has emerged as a promising approach for predicting molecular properties and reaction activation energies, potentially reducing the computational cost associated with high-level quantum chemical calculations. In particular, graph-based neural network models can learn relationships between molecular structures and their corresponding energetic properties from quantum chemical datasets. However, the reliability of ML predictions for transition states, which involve substantial changes in bonding configurations, must be carefully assessed against accurate quantum chemical reference data. Combining ML predictions with high-level electronic structure calculations therefore offers an opportunity to evaluate the applicability of data-

driven models to the energetic characterization of complex reaction pathways. Such an approach is especially useful when several transition states need to be examined across structurally related reaction systems.

In this work, the reaction mechanisms and kinetics of four pathways leading to phenol + CH_3 formation were systematically investigated, namely the reactions of *o*-, *m*-, and *p*-cresol with H atoms and the reaction of toluene with OH radicals. The potential energy surfaces were characterized using the B3LYP/6-311++G(3df,2p) method for geometry optimization and CCSD(T)/6-311++G(3df,2p) for single-point energy calculations. A graph neural network-based ML model was additionally employed to predict the energies of the transition states, and its performance was evaluated by comparison with the corresponding quantum chemical results. Furthermore, temperature- and pressure-dependent rate coefficients were calculated using RRKM/master equation theory over the temperature range of 300-2000 K and at pressures of 0.1-1.2 atm. By comparing the energetic profiles and kinetic behavior of these four pathways, this study aims to clarify the role of cresol isomers and toluene as potential aromatic precursors of phenol and to provide a more comprehensive understanding of the mechanisms governing its formation under gas-phase reaction conditions.

2. COMPUTATIONAL METHODS

The potential energy surface (PES) for phenol formation through the reactions of cresol isomers with atomic hydrogen and toluene with the OH radical were investigated using the CCSD(T)//B3LYP/6-311++G(3df,2p) approach.[12-15] All stationary-point geometries were optimized at the B3LYP/6-311++G(3df,2p) level of theory. Harmonic frequency calculations were then performed to distinguish minima from transition states and to determine the zero-point vibrational energy (ZPVE) corrections. Minima were characterized by the absence of imaginary frequencies, while transition states were confirmed by the presence of a single imaginary frequency. The calculated ZPVEs were scaled by a factor of 0.985, following previous recommendations.[16-21] Intrinsic reaction coordinate (IRC) calculations[22] were also carried out to verify the connections between each transition state and the corresponding adjacent minima. The electronic energies were

subsequently refined by single-point CCSD(T) calculations using the 6-311++G(3df,2p) basis set. The final relative energies were obtained by combining these electronic energies with the scaled ZPVE corrections, and the resulting PESs were used for the mechanistic and kinetic analyses. All calculations were performed with the Gaussian 16 program.[23]

The reliability of the single-reference approach was evaluated by calculating the T1 diagnostic for all stationary points at the CCSD(T)/6-311++G(3df,2p) level. As reported in Table S1 of the ESI, the resulting values fall within the range of 0.011-0.037; in which, T1 diagnostic values of closed-shell species are all lower than 0.02, and those of open-shell species are all lower than 0.045.[24] This suggests that insignificant multireference character was involved in this study, supporting the use of CCSD(T) for the present reaction systems. In addition, spin contamination in the open-shell species was evaluated by examining the expectation value of the spin-squared operator, $\langle S^2 \rangle$. The values before spin annihilation ranged from 0.75 to 0.78 and approached the ideal doublet value of 0.75 after annihilation. The small deviations from the theoretical value indicate that spin contamination is limited and is not expected to substantially affect the calculated properties or relative energies.

In the present study, a graph neural network (GNN)[9,10] was also employed to predict the energies of eight transition states (T1P-T4P and T0/1-T0/4). The model was trained on a dataset of 10,000 molecular and transition-state structures with reference energies calculated at the CCSD(T) level. The training dataset included a substantial number of transition states involving bond formation and bond dissociation processes similar to those investigated in this work. All eight target transition states were excluded from the training set. The GNN-predicted energies were subsequently compared with the corresponding quantum-chemical results at the B3LYP, CBS-QB3, and CCSD(T) levels to assess the accuracy of the ML predictions. The resulting deviations from the CCSD(T) reference energies are reported in Table 1.

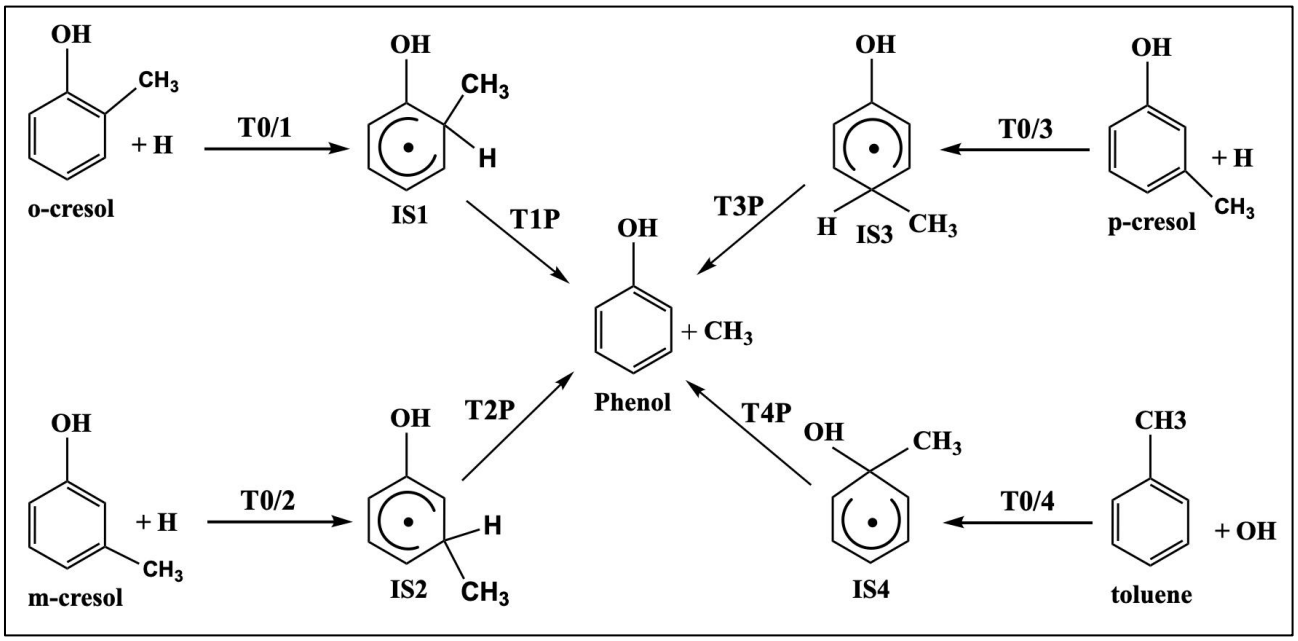
The temperature- and pressure-dependent kinetics of the investigated reactions were analyzed using Rice–Ramsperger–Kassel–Marcus (RRKM) theory coupled with master-equation (ME) calculations.[25,26] The energies, rotational constants, and vibrational frequencies required for these

calculations were taken from the quantum-chemical results presented in the preceding section. To improve the treatment of the low-frequency motions of the CH_3 and OH groups, the corresponding modes were modeled as hindered internal rotations instead of being described by the harmonic oscillator approximation. Tunneling effects were accounted for using a two-dimensional asymmetric Eckart potential.[27] For the chemically activated intermediates generated by H-atom addition to cresols and OH-radical addition to toluene, microcanonical unimolecular rate coefficients were calculated using RRKM theory. The pressure dependence of the kinetics was then obtained by solving the master equation with an energy grain size of 100 cm^{-1}. Nitrogen was selected as the bath gas, with Lennard–Jones parameters of ε/k_B = 48 K and σ = 3.9 Å.[28] In the absence of specific collision parameters for the chemically activated intermediates, their Lennard-Jones parameters were estimated from the reported values for 2-methylphenol (ε/k_B = 561.62 K and σ = 6.36 Å).[29] Collisional energy transfer was represented by the single-exponential-down model, using an average downward energy transfer parameter, $\langle \Delta E_{down} \rangle$, of 400 cm^{-1}.[30] All RRKM/ME calculations were carried out with the MESMER code.[31] Rate coefficients and product branching fractions were determined over the temperature range of 300-2000 K and the pressure range of 0.1-1.2 atm, covering conditions relevant to both atmospheric chemistry and high-temperature combustion processes.

3. RESULTS AND DISCUSSION

In the present study, the B3LYP/6-311++G(3df,2p) geometries of all species are shown in Figure S1 of the ESI. The energies of the stationary points at different levels of theory are given in Table S2. In addition, Tables S3-S5 of the ESI file present vibrational frequencies and Cartesian coordinates along with thermodynamic properties at 298 K such as standard enthalpies, Gibbs free energies, and entropies. The obtained results provide comprehensive thermodynamic insights into the reaction systems, allowing the energetic viability of the individual pathways to be evaluated, together with their exothermic or endothermic characteristics and the relative stability of the resulting intermediates and products.

To construct the potential energy landscape for the reactions of cresol isomers with H atoms, a reaction scheme connecting the reactants to the intermediates and, ultimately, to the phenol + CH_3 products was first established, as shown in Scheme 1. The structures of all relevant species were then constructed and optimized, followed by single-point energy calculations. The relative energies of these species were subsequently determined with respect to the *o*-cresol + H reactants, which were taken as the reference. The resulting energy data are shown in Figure 1.



Scheme 1. The formation of phenol from the reactions of cresols + H and toluene + OH

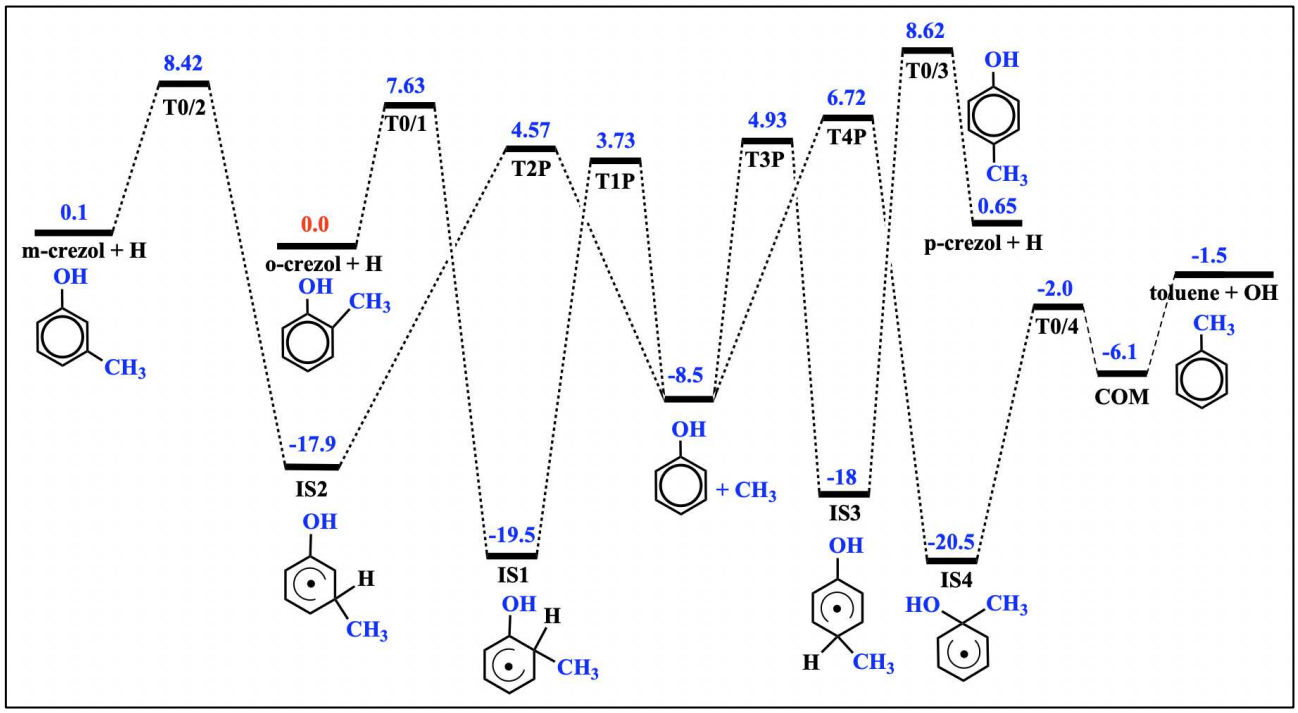


Figure 1. The potential energy profile of the reactions of cresols + H and toluene + OH. Energies (kcal/mol) were calculated at the CCSD(T)/6-311++G(3df,2p) level, including ZPVE corrections.

The PES reveals four reaction pathways leading to phenol formation, originating from *o*-cresol + H, *m*-cresol + H, *p*-cresol + H, and toluene + OH. The first three pathways proceed through a sequential H-addition and methyl-elimination mechanism, whereas the fourth involves the reaction of OH with toluene. Each pathway contains two transition states separated by a relatively stable intermediate, suggesting that product formation involves initial radical addition followed by cleavage of the C-CH_3 bond. The first pathway, starting from *o*-cresol + H, proceeds through T0/1 and the intermediate IS1 before reaching phenol via T1P. T0/1 lies 7.63 kcal/mol above the reactants and corresponds to H-atom addition to the ortho carbon bearing the methyl substituent. The H···C distance of 1.815 Å and the imaginary frequency of 731.8i cm^{-1} are consistent with the formation of a new C-H bond. The resulting adduct IS1 is strongly stabilized, lying 19.50 kcal/mol below the reactants. Subsequent cleavage of the C-CH_3 bond occurs through T1P, located 3.73 kcal/mol above the reactant asymptote. In this transition state, the C-CH_3 distance has increased to 2.133 Å, accompanied by an imaginary frequency of 510i cm^{-1}, consistent with substantial elongation of the bond being broken. Although T1P is lower in energy than T0/1 when both are referenced to the reactants, it lies 23.23 kcal/mol above IS1. Thus, methyl elimination requires the system to climb out of the deep IS1 potential well, and the energy of T1P relative to IS1 is more relevant to assessing the barrier for this elementary step.

The second pathway, originating from *m*-cresol + H, follows a similar mechanism through T0/2, IS2, and T2P. The entrance transition state T0/2 lies 8.42 kcal/mol above the reactants and is associated with H addition at the meta carbon bearing the methyl group. The H-C distance is 1.821 Å, and the corresponding imaginary frequency is 836i cm^{-1}. The resulting intermediate IS2 lies 17.90 kcal/mol below the point of *o*-cresol + H. The pathway then proceeds through T2P, at 4.57 kcal/mol relative to the reactants, where the C-CH_3 bond is elongated to 2.179 Å; its imaginary frequency is 508i cm^{-1}. Compared with the *o*-cresol pathway, both transition states are slightly higher in energy, while IS2 is less stabilized than IS1. These differences suggest a modest energetic preference for

phenol formation from *o*-cresol. However, the energy differences alone do not establish the origin of this positional effect, which may involve differences in local bonding and the electronic structures of the transition states and intermediates.

The third pathway, beginning with *p*-cresol + H, also involves H addition followed by methyl-group elimination through T0/3, IS3, and T3P. The relative energies of T0/3 and T3P are 8.62 and 4.93 kcal/mol, respectively, while IS3 lies 18 kcal/mol below the reactants. These values are close to those calculated for the *m*-cresol pathway: the entrance transition state is only 0.2 kcal/mol higher, the elimination transition state is 0.36 kcal/mol higher, and the intermediate is stabilized by 0.1 kcal/mol more. Such small differences indicate that the two pathways have very similar potential energy profiles. Consequently, phenol formation from *m*-cresol and *p*-cresol may be competitive under comparable conditions. This similarity may reflect the comparable energetic requirements for H addition and C-CH_3 bond cleavage at the *meta* and *para* positions. Nevertheless, a definitive comparison of their reaction rates requires kinetic calculations, because differences in activation entropy, energy transfer, and the subsequent dynamics of the energized intermediates may also influence the product-forming rates.

The toluene + OH pathway also provides a potentially competitive route to phenol. The reported entrance transition state T0/4 lies at -2.0 kcal/mol, following formation of a pre-reactive complex at -6.1 kcal/mol. The intermediate IS4 is located at -20.5 kcal/mol, and the product-forming transition state T4P lies at 6.72 kcal/mol on the reported energy scale. Taking the stated energy of toluene + OH (-1.5 kcal/mol) as the reference for this pathway, the corresponding energy barriers are 8.22 kcal/mol for T4P and 0.5 kcal/mol below the reactant level for T0/4. The T4P barrier is therefore comparable to, though slightly higher than, the entrance barrier of the *o*-cresol + H pathway (7.63 kcal/mol). This comparison suggests that the toluene + OH route may also be energetically accessible, although its kinetic contribution cannot be established from the PES alone.

Overall, the calculated PES identify the initial H-addition step as the highest electronic-energy threshold for each of the three cresol + H pathways. The *o*-cresol route has the lowest entrance barrier, while the *m*-cresol and *p*-cresol routes exhibit very similar energetics. The toluene + OH pathway may provide an additional route to phenol with a comparable energetic threshold. These findings establish a basis for comparing the four mechanisms; their relative kinetic importance, however, should ultimately be assessed using rate calculations that account for the relevant thermodynamic and dynamical effects.

Table 1. Comparison of activation energies (kcal/mol) obtained from quantum chemical calculations and machine learning (ML) predictions

Transition States	B3LYP	CBS-QB3	CCSD(T)	ML prediction
T1P	3.9	3.09	3.73	3.25
T2P	4.42	4.28	4.57	4.28
T3P	5.71	4.31	4.93	4.37
T4P	6.23	6.12	6.72	6.98
T0/1	7.35	7.15	7.63	7.26
T0/2	9.27	8.56	8.42	8.34
T0/3	10.4	8.01	8.62	8.53
T0/4	0.5	-1.8	-2.0	-1.95

In this investigation, a machine learning model, namely GNN, was also used to predict the relative energies for all transition states appearing on the PES. The reliability of the GNN model was evaluated by comparing its predictions for eight transition states (T1P–T4P and T0/1–T0/4) with the corresponding CCSD(T) energies (Table 1). The model was trained on 10,000 molecular and transition-state structures with CCSD(T)-level reference energies. Notably, the training set contained a substantial number of transition states involving bond formation and bond dissociation processes similar to those encountered in the cresol + H reaction, whereas all eight transition states considered here were excluded from the training set. The predicted energies show close agreement with the CCSD(T) results, with an MAE of 0.27 kcal/mol, an RMSE of 0.32 kcal/mol, and a maximum absolute deviation of 0.56 kcal/mol. This agreement across both entrance-channel and product-forming transition states suggests that the GNN model captures relevant structural and energetic

features of the reaction pathways examined. The results therefore support the use of the trained GNN as a complementary approach for estimating transition-state energies at substantially lower computational cost than direct high-level calculations. Although the present comparison is limited to eight target transition states, the chemical relevance of the training data and the exclusion of these structures from training provide encouraging evidence of the model's predictive capability for the reaction system studied and the similar reaction systems.

4. RATE CONSTANT CALCULATIONS

For the convenience of kinetic calculations, four reaction paths on the PES are denoted as follows:

o-cresol + H → IS1 → phenol + CH_3 (k_1)

m-cresol + H → IS2 → phenol + CH_3 (k_2)

p-cresol + H → IS3 → phenol + CH_3 (k_3)

toluene + OH → IS4 → phenol + CH_3 (k_4)

Rate constants, k_1 - k_4, were calculated using the Mesmer program within the RRKM/ME framework over the temperature range of 300–2000 K and at pressures of 0.1, 0.6, 1.0, and 1.2 atm to examine the effects of both temperature and pressure on the reaction kinetics. The calculated rate constants for the individual pathways are presented in Tables S6-S9 of the ESI and the temperature- and pressure-dependent variations in the calculated rate coefficients are illustrated in Figures 2-5.

The calculated results show that all four reactions exhibit an increase in their rate coefficients with increasing temperature, indicating the strong thermal activation associated with the formation of phenol through these reaction pathways. Nevertheless, the magnitude of the rate coefficients and their responses to temperature and pressure differ considerably among the four systems. Among the three cresol isomers, the reaction of o-cresol with H (k_1) consistently exhibits the highest rate coefficient, followed by the reactions of m-cresol (k_2) and *p*-cresol (k_3). This ordering is maintained throughout the investigated temperature and pressure ranges. For instance, at 0.1 atm, k_1 rises from 7.83×10^{-19} cm^3 $molecule^{-1}$ s^{-1} at 300 K to 5.62×10^{-13} cm^3 $molecule^{-1}$ s^{-1} at 2000 K. The corresponding values for

k_2 and k_3 increase from 7.49×10^{-20} and 5.06×10^{-20} cm^3 molecule^{-1} s^{-1} to 1.22×10^{-13} and 6.40×10^{-14} cm^3 molecule^{-1} s^{-1}, respectively. The consistently higher rate coefficient of the *o*-cresol reaction suggests that the *ortho*-substituted structure provides a more favorable kinetic pathway for phenol formation than the *meta*- and *para*-substituted isomers. The *T*-dependence of these three reactions is particularly pronounced at low temperatures. At 0.1 atm, the rate coefficients of k_1, k_2, and k_3 increase by approximately 5-6 orders of magnitude between 300 and 2000 K. However, their relative differences do not remain constant throughout this temperature interval. For example, the ratio k_1/k_2 reduces from approximately 10.5 at 300 K to 4.6 at 2000 K, indicating that the kinetic difference between the *ortho*- and *meta*-isomers becomes less pronounced at elevated temperatures. In contrast, the relative contribution of the *para*-isomer remains comparatively limited, as k_3 always exhibits the lowest rate coefficient among the three cresol reactions.

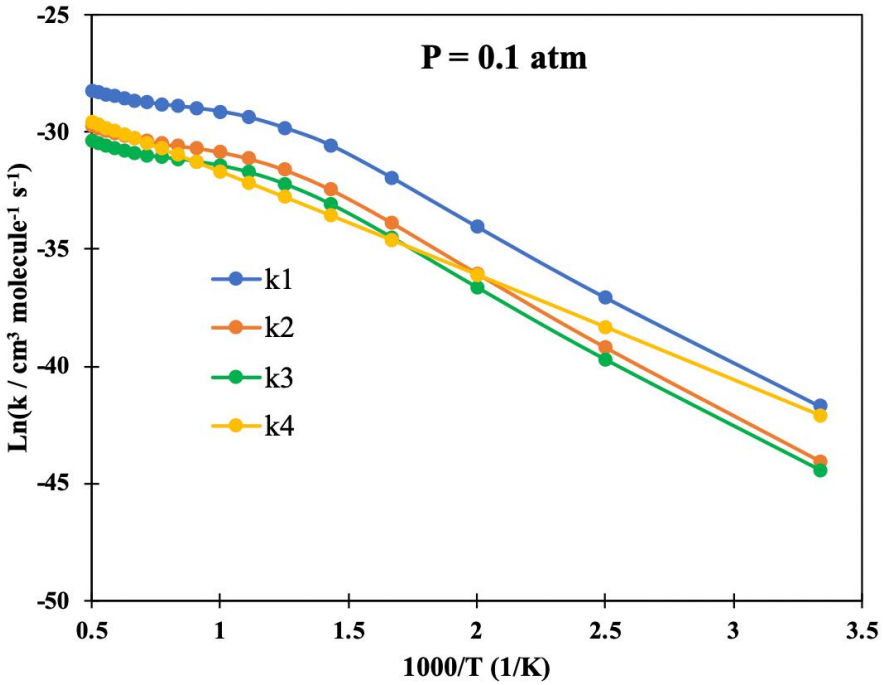


Figure 2. Rate coefficients of the four channels forming the phenol + CH_3 product over a wide temperature range of 300 - 2000 K and pressure of 0.1 atm.

The reaction of toluene with OH (k_4) exhibits a distinct temperature-dependent behavior. At 0.1 atm, its rate coefficient increases from 5.24×10^{-19} cm^3 molecule^{-1} s^{-1} at 300 K to 1.42×10^{-13} cm^3 molecule^{-1} s^{-1} at 2000 K. Although k_4 is comparable to k_1 at low temperatures, its relative kinetic importance changes considerably as the temperature increases. At 0.1 atm, k_4 is initially higher than

k_2 and k_3 at 300-400 K, but becomes lower than both pathways over most of the intermediate-temperature region. Upon further heating, k_4 gradually regains its competitive position, exceeding k_3 at approximately 1100-1200 K and k_2 at approximately 1400-1500 K. Consequently, at temperatures above approximately 1500 K, the relative ordering becomes $k_1 > k_4 > k_2 > k_3$. This behavior demonstrates that the kinetic importance of the toluene + OH pathway is strongly dependent on temperature, particularly under low-pressure conditions.

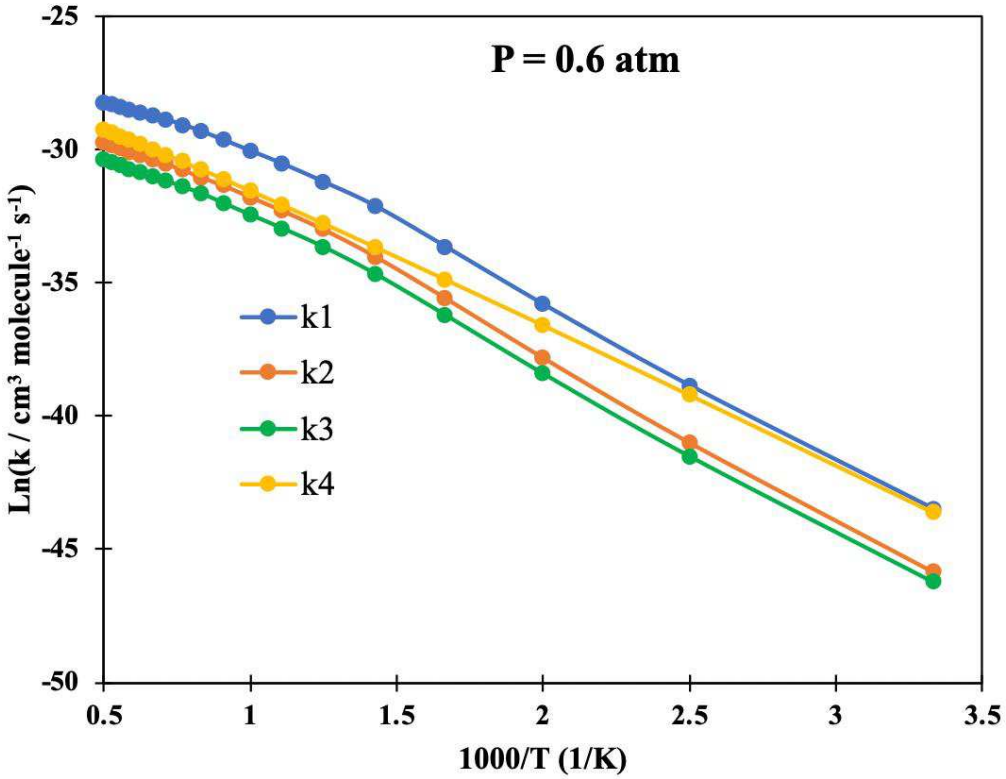


Figure 3. Rate coefficients of the four channels forming the phenol + CH_3 product over a wide temperature range of 300 - 2000 K and pressure of 0.6 atm.

A particularly interesting feature is the pressure dependence of the rate coefficients. For the three cresol + H reactions, increasing pressure leads to a substantial decrease in the rate coefficients at low temperatures. For example, at 300 K, increasing the pressure from 0.1 to 1.2 atm decreases k_1, k_2, and k_3 by approximately a factor of 12. However, this pressure effect becomes progressively weaker as the temperature increases. At 2000 K, the corresponding rate coefficients decrease by only approximately 2.4% when the pressure is increased from 0.1 to 1.2 atm. This behavior suggests that the three cresol reactions approach a relatively pressure-insensitive kinetic regime at elevated

temperatures. In contrast, the toluene + OH reaction displays a more complex pressure dependence. At low temperatures, increasing the pressure decreases k_4, following a trend similar to those observed for the cresol reactions. At 300 K, for example, k_4 decreases from 5.24×10^{-19} cm^3 molecule^{-1} s^{-1} at 0.1 atm to 5.74×10^{-20} cm^3 molecule^{-1} s^{-1} at 1.2 atm. Nevertheless, the pressure dependence becomes progressively weaker with increasing temperature, and the rate coefficients calculated at the four pressures approach one another around 900-1000 K. Above this temperature region, the trend is reversed: increasing pressure results in an increase in k_4. At 2000 K, k_4 increases from 1.42×10^{-13} cm^3 molecule^{-1} s^{-1} at 0.1 atm to 2.86×10^{-13} cm^3 molecule^{-1} s^{-1} at 1.2 atm, corresponding to an approximately twofold enhancement. This reversal in pressure dependence distinguishes the toluene + OH pathway from the three cresol + H reactions and highlights the importance of pressure-dependent unimolecular kinetics in determining its overall rate.

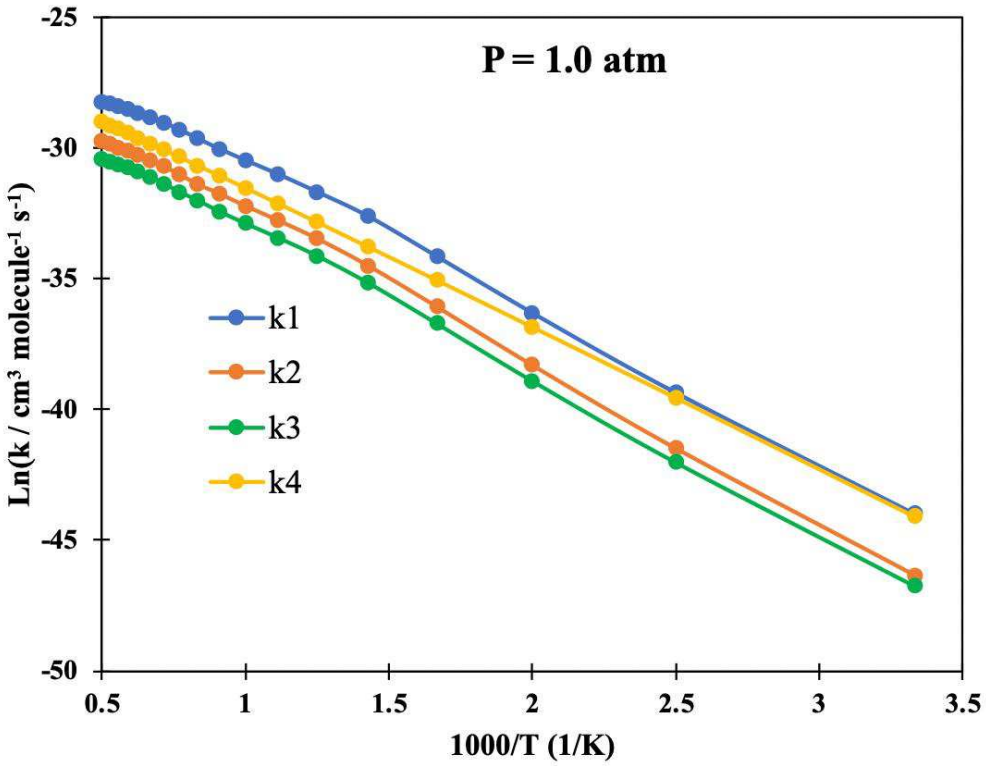


Figure 4. Rate coefficients of the four channels forming the phenol + CH_3 product over a wide temperature range of 300 - 2000 K and pressure of 1.0 atm.

To further assess the reliability of the present kinetic calculations, the rate coefficients obtained for the toluene + OH reaction were compared with the previously reported theoretical results of Wu et al.[8] over the common temperature range of 300-1000 K and at pressures of 0.1, 0.6, 1.0, and 1.2

atm, as presented in Table 2 and illustrated in Figure 6. Overall, very good agreement is obtained between the two studies. The mean absolute relative deviation across the 32 corresponding data points is approximately 6.34%, while the maximum deviation is approximately 17.74%. The mean absolute deviations at 0.1, 0.6, 1.0, and 1.2 atm are 7.44%, 3.98%, 6.26%, and 7.68%, respectively. The agreement is particularly good at 0.6 atm, where the mean absolute deviation is below 4%. At 0.1 atm, the present rate coefficients are consistently slightly lower than those reported by Wu et al., with deviations ranging from approximately 5.3% to 10.7%. At higher pressures, some deviations become more pronounced at 400 K; however, the overall agreement remains satisfactory, particularly at temperatures of 500-1000 K. Notably, at 900 K, the calculated values are almost identical to those reported in the previous study, with deviations of 0.0%, 1.7%, and 3.4% at 0.6, 1.0, and 1.2 atm, respectively (see Table 2).

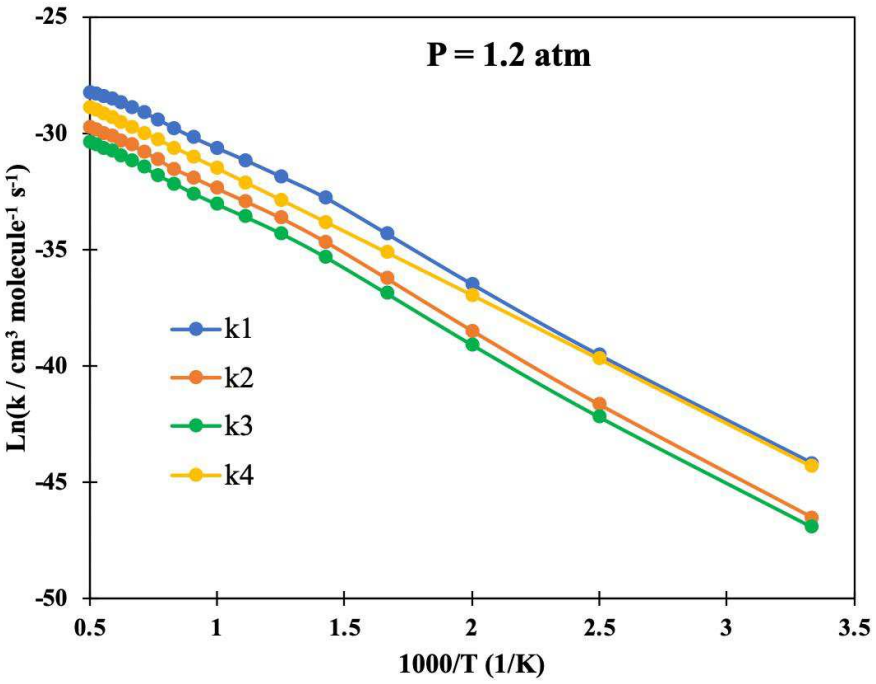


Figure 5. Rate coefficients of the four channels forming the phenol + CH_3 product over a wide temperature range of 300 - 2000 K and pressure of 1.2 atm.

In addition to the agreement in the absolute rate coefficients, the two studies reproduce similar pressure-dependent trends. At 300 K, increasing the pressure from 0.1 to 1.2 atm decreases the rate coefficient by approximately a factor of 9.1 in the present study, compared with a factor of 8.8 in the work of Wu et al. At 600 K, the corresponding decreases are approximately 1.7 and 1.6 times,

respectively. At 1000 K, the present calculations indicate a slight increase in the rate coefficient with increasing pressure, whereas the values reported by Wu et al. are nearly pressure-independent. These observations demonstrate that the present calculations satisfactorily reproduce not only the magnitude but also the general temperature- and pressure-dependent behavior of the toluene + OH reaction.

Table 2. Comparison of the present rate constants with the previously calculated data of Wu et al. for the (k_4) channel in the 300 - 2000 K temperature range and the 0.1 - 1.2 pressures.

T (K)	k_4 at 0.1 atm		k_4 at 0.6 atm		k_4 at 1.0 atm		k_4 at 1.2 atm	
	This work	Wu et al.	This work	Wu et al.	This work	Wu et al.	This work	Wu et al.
300	5.24E-19	5.57E-19	1.16E-19	1.22E-19	7.22E-20	7.57E-20	5.74E-20	6.36E-20
400	2.27E-17	2.44E-17	9.52E-18	8.51E-18	6.77E-18	5.75E-18	5.73E-18	4.96E-18
500	2.14E-16	2.37E-16	1.30E-16	1.31E-16	1.01E-16	1.00E-16	8.96E-17	9.03E-17
600	9.43E-16	1.02E-15	7.21E-16	7.71E-16	6.05E-16	6.63E-16	5.55E-16	6.22E-16
700	2.70E-15	2.85E-15	2.42E-15	2.54E-15	2.15E-15	2.35E-15	2.03E-15	2.28E-15
800	5.92E-15	6.28E-15	5.91E-15	5.94E-15	5.53E-15	5.74E-15	5.35E-15	5.64E-15
900	1.08E-14	1.17E-14	1.17E-14	1.17E-14	1.15E-14	1.17E-14	1.13E-14	1.17E-14
1000	1.75E-14	1.96E-14	2.01E-14	1.96E-14	2.04E-14	1.96E-14	2.06E-14	1.96E-14

The consistency with the previous theoretical results provides an important validation of the kinetic treatment employed in this work. It also supports the reliability of the present approach for investigating the three cresol + H reactions, for which a consistent comparison with the toluene + OH pathway is essential to establish the relative kinetic importance of the four phenol-forming channels.

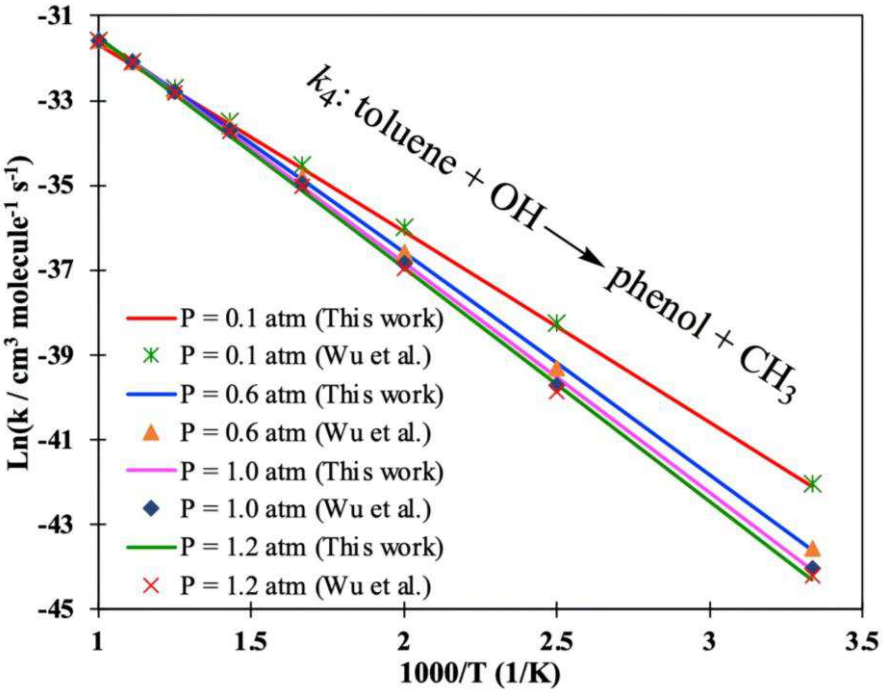


Figure 6. Rate constants (k_4) of the toluene + OH → phenol + CH_3 reaction in this work compared with those of Wu et al. in the 300 - 2000 K range at various pressures.

To provide analytical expressions for the calculated rate constants which can be readily used in future kinetic modeling studies, the rate constants of the investigated reaction pathways were fitted to the modified Arrhenius expression at P = 1 atm over the temperature range T = 300-2000 K, as summarized in Table 3.

Table 3. The modified Arrhenius expression ($k(T) = AT^n\exp(-E_a/RT)$) of the four reaction paths leading to the formation of toluene + CH_3.

Rate coefficients	A (cm^3 $molecule^{-1}$ s^{-1})	n	E_a (kcal/mol)
k_1	1.13×10^{-6}	-1.47	13.19 ± 0.28
k_2	8.74×10^{-8}	-1.32	13.54 ± 0.27
k_3	6.96×10^{-9}	-1.09	13.07 ± 0.27
k_4	8.64×10^{-11}	-0.39	11.13 ± 0.52

5. CONCLUSION

In this study, the reaction mechanisms and kinetics of phenol formation through the reactions of the three cresol isomers (*o*-, *m*-, and *p*-cresol) with H atoms and toluene with OH radicals were systematically investigated using a combination of high-level quantum chemical calculations, machine learning (ML), and RRKM/ME calculations. The PES was characterized at the CCSD(T)//B3LYP/6-311++G(3df,2p) level of theory, while the ML model was employed to predict transition-state energies and assess its performance in describing the energetics of the investigated reaction pathways. The calculated rate coefficients were obtained over the temperature range of 300-2000 K and at pressures of 0.1-1.2 atm.

The results indicate that all four reactions proceed through multistep pathways involving the formation of intermediate species (IS1–IS4), followed by CH_3-elimination processes to yield the same bimolecular product, phenol + CH_3. From an energetic perspective, the pathway involving *o*-cresol + H is more favorable than those involving *m*-cresol + H and *p*-cresol + H. Notably, the toluene + OH reaction also represents a competitive pathway for phenol formation, with its energetic profile being comparable to that of the *o*-cresol + H reaction. In addition, the activation energies predicted by ML

show good agreement with the corresponding quantum chemical results, with deviations of less than 1 kcal/mol. These results demonstrate that ML can serve as a useful complementary approach for the rapid estimation of activation energies and screening of related gas-phase radical reactions, while high-level quantum chemical calculations remain essential for detailed mechanistic characterization.

The calculated rate coefficients reveal a clear kinetic preference among the three cresol isomers, following the order ($k_1 > k_2 > k_3$) across the entire temperature and pressure ranges investigated, with the *o*-cresol + H reaction being the fastest among all four pathways. The relative contribution of the toluene + OH reaction is strongly dependent on temperature and pressure: it is generally the second-fastest pathway at 0.6-1.2 atm, while its relative kinetic importance varies with temperature at 0.1 atm. Furthermore, increasing pressure significantly suppresses the rate coefficients of the cresol + H reactions at low temperatures, although this effect diminishes at higher temperatures. In contrast, the toluene + OH reaction exhibits an opposite pressure dependence above approximately 1000 K, with its rate coefficient increasing as pressure rises.

Finally, the calculated rate coefficients for toluene + OH show good overall agreement with the previously reported results of Wu et al., with a mean absolute relative deviation of approximately 6.3% over the compared temperature and pressure conditions. This agreement provides additional support for the reliability of the present kinetic calculations. Overall, this work provides a consistent theoretical description of four chemically related pathways leading to phenol formation and establishes their temperature- and pressure-dependent kinetic characteristics. The results may contribute to a better understanding of phenol formation in aromatic oxidation and combustion environments and provide useful kinetic information for the development of more comprehensive chemical kinetic mechanisms.

ACKNOWLEDGEMENTS

This research is funded by the Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 104.06-2025.118. We would like to thank the National Center for High-Performing Computers in Taiwan for the use of its facilities.

REFERENCES

1. Hossen MAA, Hasan Shakil M, Ehasan MF, Mohammad Tareq AR, Salam A (2025) Composition and formation mechanism of brown carbon: identification and quantification of phenolic precursors. Atmos Chem Phys 25:14629–14642. https://doi.org/10.5194/acp-25-14629-2025
2. El-Sayed SA (2025) Chemical products yielded from different pyrolysis processes of rice waste residues: a comprehensive review. Biomass Convers Biorefin 15:20615–20655. https://doi.org/10.1007/s13399-025-06624-3
3. MacFarlane SM, Fisher JA, Xu L, Wennberg PO, Crounse JD, Ball K, et al. (2025) Sources, sinks, and oxidation pathways of phenolic compounds in South Korea constrained using KORUS-AQ airborne observations. J Geophys Res Atmos 130:e2024JD043110. https://doi.org/10.1029/2024JD043110
4. Ji Y, Zhao J, Terazono H, Misawa K, Levitt NP, Li Y, et al. (2017) Reassessing the atmospheric oxidation mechanism of toluene. Proc Natl Acad Sci USA 114:8169–8174. https://doi.org/10.1073/pnas.1705463114
5. Noda J, Volkamer R, Molina MJ (2009) Dealkylation of alkylbenzenes: a significant pathway in the toluene, o-, m-, p-xylene + OH reaction. J Phys Chem A 113:9658–9666. https://doi.org/10.1021/jp901529k
6. 6. Pratali Maffei L, Faravelli T, Cavallotti C, Pelucchi M (2020) Electronic structure-based rate rules for H ipso addition-elimination reactions on mono-aromatic hydrocarbons with single and double $OH/CH_3/OCH_3/CHO/C_2H_5$ substituents: a systematic theoretical investigation. Phys Chem Chem Phys 22:20368-20387. https://doi.org/10.1039/D0CP03099F
7. Pelucchi M, Cavallotti C, Cuoci A, Faravelli T, Frassoldati A, Ranzi E (2019) Detailed kinetics of substituted phenolic species in pyrolysis bio-oils. React Chem Eng 4:490-506. https://doi.org/10.1039/C8RE00198G
8. Wu X, Huang C, Niu S, Zhang F (2020) New theoretical insights into the reaction kinetics of toluene and hydroxyl radicals. Phys Chem Chem Phys 22:22279-22288. https://doi.org/10.1039/D0CP02984J
9. Lewis-Atwell T, Townsend PA, Grayson MN (2022) Machine learning activation energies of chemical reactions. WIREs Comput Mol Sci 12:e1593. https://doi.org/10.1002/wcms.1593
10. Choi S, Kim Y, Kim JW, Kim Z, Kim WY (2018) Feasibility of activation energy prediction of gas-phase reactions by machine learning. Chem Eur J 24:12354-12358. https://doi.org/10.1002/chem.201800345
11. Friederich P, dos Passos Gomes G, De Bin R, Aspuru-Guzik A, Balcells D (2020) Machine learning dihydrogen activation in the chemical space surrounding Vaska's complex. Chem Sci 11:4584-4601. https://doi.org/10.1039/D0SC00445F
12. Becke AD (1993) Density-functional thermochemistry. III. The role of exact exchange. J Chem Phys 98:5648-5652. https://doi.org/10.1063/1.464913
13. Becke AD (1988) Density-functional exchange-energy approximation with correct asymptotic behavior. Phys Rev A 38:3098-3100. https://doi.org/10.1103/PhysRevA.38.3098
14. Stanton JF (1997) Why CCSD(T) works: a different perspective. Chem Phys Lett 281:130-134. https://doi.org/10.1016/S0009-2614(97)01144-5
15. Kowalski K, Piecuch P (2000) The method of moments of coupled-cluster equations and the renormalized CCSD[T], CCSD(T), CCSD(TQ), and CCSDT(Q) approaches. J Chem Phys 113:18–35. https://doi.org/10.1063/1.481769
16. Pham TV (2025) Computational study on the mechanisms and kinetics of the $H_2 + H_2CCCH$ reaction in the gas phase. Chem Phys Lett 879:142392. https://doi.org/10.1016/j.cplett.2025.142392

17. Pham TV, Tran AV, Nguyen NT (2026) A theoretical study on mechanism and kinetics of o-toluidine with NH_2 radicals. Comput Theor Chem 1257:115686. https://doi.org/10.1016/j.comptc.2026.115686
18. Trang HT, Nguyen TN, Nguyen NT, Pham TV (2026) Mechanistic and kinetic insights into the gas-phase reaction of p-toluidine with amino radicals: a high-level theoretical study. J Mol Model 32:71. https://doi.org/10.1007/s00894-026-06629-3
19. Pham TV (2025) A theoretical investigation of the kinetics of hydrogen abstraction and amino radical addition to 3-methylaniline in the gas phase. Theor Chem Acc 144:58. https://doi.org/10.1007/s00214-025-03217-1
20. Pham TV (2025) Mechanism and kinetics of the m-toluidine reaction with NH_2 radicals: a computational study. Chem Phys Lett 870:142094. https://doi.org/10.1016/j.cplett.2025.142094
21. Pham TV, Le TP, Nguyen NT, Nam PC (2026) Insights into mechanism and kinetics of toluene reactions with CH_3 radicals: a theoretical study. ACS Omega 11:40389-40401. https://doi.org/10.1021/acsomega.6c03091
22. Gonzalez C, Schlegel HB (1989) An improved algorithm for reaction path following. J Chem Phys 90:2154-2161. https://doi.org/10.1063/1.456010
23. 2. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Petersson GA, Nakatsuji H, Li X, Caricato M, Marenich A, Bloino J, Janesko BG, Gomperts R, Mennucci B, Hratchian HP, Ortiz JV, Izmaylov AF, Sonnenberg JL, Williams-Young D, Ding F, Lipparini F, Egidi F, Goings J, Peng B, Petrone A, Henderson T, Ranasinghe D, Zakrzewski VG, Gao J, Rega N, Zheng G, Liang W, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Throssell K, Montgomery JA, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Keith T, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Millam JM, Klene M, Adamo C, Cammi R, Ochterski JW, Martin RL, Morokuma K, Farkas O, Foresman JB, Fox DJ (2016) Gaussian 16, Revision C.01. Gaussian, Inc., Wallingford, CT, USA.
24. Alecu IM, Truhlar DG (2011) Computational study of the reactions of methanol with the hydroperoxyl and methyl radicals. 1. Accurate thermochemistry and barrier heights. J Phys Chem A 115:2811-2829. https://doi.org/10.1021/jp110024e
25. Klippenstein SJ (1992) Variational optimizations in the Rice-Ramsperger-Kassel-Marcus theory calculations for unimolecular dissociations with no reverse barrier. J Chem Phys 96:367-371. https://doi.org/10.1063/1.462472
26. Wardlaw DM, Marcus RA (1984) RRKM reaction rate theory for transition states of any looseness. Chem Phys Lett 110:230-234. https://doi.org/10.1016/0009-2614(84)85219-7
27. Eckart C (1930) The penetration of a potential barrier by electrons. Phys Rev 35:1303-1309. https://doi.org/10.1103/PhysRev.35.1303
28. Hippler H, Troe J, Wendelken HJ (1983) Collisional deactivation of vibrationally highly excited polyatomic molecules. II. Direct observations for excited toluene. J Chem Phys 78:6709-6717. https://doi.org/10.1063/1.444670
29. Mourits FM, Rummens FHN (1977) A critical evaluation of Lennard-Jones and Stockmayer potential parameters and of some correlation methods. Can J Chem 55:3007–3020. https://doi.org/10.1139/v77-418
30. Tardy DC, Rabinovitch BS (1966) Collisional energy transfer. Thermal unimolecular systems in the low-pressure region. J Chem Phys 45:3720-3730. https://doi.org/10.1063/1.1727392
31. Glowacki DR, Liang CH, Morley C, Pilling MJ, Robertson SH (2012) MESMER: An open-source master equation solver for multi-energy well reactions. J Phys Chem A 116:9545-9560. https://doi.org/10.1021/jp3051033

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupportingInformation.doc