Re-examination of Proline-Catalyzed Intermolecular Aldol Reactions: An Ab Initio Kinetic Modelling Study

Li-Juan Yu, Mitchell T. Blyth, Michelle L. Coote*

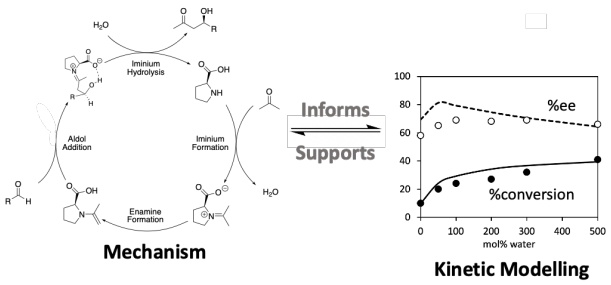
Research School of Chemistry, Australian National University, Canberra, Australian Capital Territory, 2601, Australia

Email: michelle.coote@anu.edu.au

ABSTRACT: The full catalytic cycle of the proline-catalyzed intermolecular aldol reaction of acetone and p-nitrobenzaldehyde in acetone solvent has been investigated by quantum chemistry at the G3(MP2,CC)//M062X/6-31+G(d)/SMD level of theory, and the results used to develop an ab initio kinetic model. Proline catalyzes the aldol reaction according to the enamine mechanism. The initial reaction between proline and acetone was reinvestigated, and a revised mechanism for enamine formation is proposed in which a second proline assists the process contributing to the enamine formation. Using various initial concentrations of proline while keeping the experimental ratio of water, aldehyde and acetone constant, we find that the enamine formation from the first-order to proline pathway dominates when the concentration of proline is low (< 0.005 M); while the second-order enamine formation pathways contribute and then dominate as the proline concentration is increased. The relative rates of formation of the *syn* and *anti*-enamine are not important, as these interconvert via C-N bond rotation and equilibrate faster than their subsequent reaction, which follows the standard Houk/List mechanism. While the stereochemistry can be predicted from an analysis of the alternative C-C bond formation pathways, their relative contributions to the major and minor product yields are influenced by their subsequent rates of hydrolysis. Indeed, while C–C bond formation is normally considered rate determining, our kinetic simulations show that the kinetic model is more complicated than this and under typically used concentrations, the process of initial enamine formation, C-C bond formation and the initial stages of product release all contribute to the overall reaction rate. Using our kinetic model, we predict that yield and %ee are optimal for concentrations of [proline] = 0.005 M, [acetone] = 2.25 M, [aldehyde] = 0.1 M, and [water] = 0.6 M. Using excess acetone (< 2.6 M) increases both conversion and %ee. Excess aldehyde increases %ee but decreases conversion, and excess catalyst increases the conversion but decreases %ee. Aside from the indirect effect of increasing the solubility of the proline catalyst, water increases both conversion and %ee up to a point, but at large concentrations (> 1.0 M) excess water is expected to decrease %ee. Side reactivity, including aldol condensation, acetone self-aldolization, oxazolidinone formation and azomethine and 1-oxapyrrolizidine formation were all considered in our kinetic model but shown to have a negligible effect (<2%) on the yield and %ee over the full range of reaction conditions investigated.

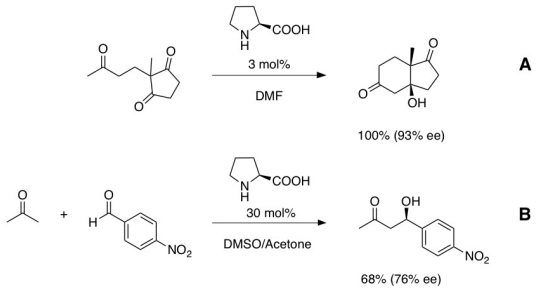
Keywords:
Proline-catalysed aldol reaction, kinetic modelling, quantum chemistry

Table of Contents Graphic



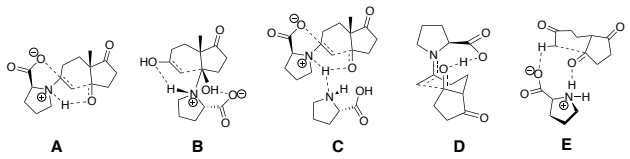
INTRODUCTION

The aldol reaction is widely regarded as one of the most important C–C bond formation reactions utilized in organic synthesis. Eder and coworkers [1] and Hajos and Parrish [2] discovered that (*S*)-proline catalyzes intramolecular aldol reactions with high enantiomeric excesses and chemical yields (Scheme 1A) in the 1970s. In 2000, List and coworkers reported intermolecular aldol reactions catalyzed by proline (Scheme 1B) [3]. Understanding the kinetic mechanism of these proline-catalyzed aldol reactions, and the origin of the catalysis and enantioselectivity, is crucial to process optimization and improved catalyst design.



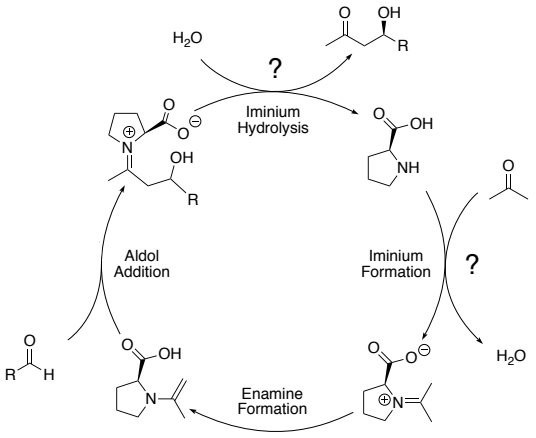
Scheme 1. Key (A) intra- and (B) intermolecular aldol reactions catalyzed by proline [1-3].

Although many studies of proline-catalyzed aldol reactions have been reported, the reaction mechanism remains controversial, and a number of alternative models have been proposed. For example, Hajos and Parrish [2] suggested two mechanisms (Scheme 2): A, which involves the formation of a protonated enamine and of an oxazolidinone ring; and B, which involves the addition of (*S*)-proline in its zwitterionic form to one of the carbonyl groups. The stereochemistry of mechanism B was subsequently questioned by Jung [4]. Additionally, Kagan and coworkers [5] proposed a side-chain enamine mechanism that involves two proline molecules in the C–C bond formation transition state, one engaged in enamine formation and the other as a proton transfer mediator (Scheme 2, C). This mechanism was challenged by List and coworkers [3,6], who proposed a one-proline enamine mechanism. Based on density functional theory (DFT) calculations, Houk and coworkers [7-12] proposed a very similar mechanism to List for the intramolecular aldol reaction (Scheme 2, D). Rzepa and co-workers also confirmed the Houk-List mechanism, while highlighting the importance of dispersion in the C-C bond forming transition states [13]. In contrast, Swaminathan and coworkers [14] preferred a heterogeneous aldolization mechanism on the surface of crystalline proline, despite that fact that many proline-catalyzed aldolizations are completely homogenous (Scheme 2, E).



Scheme 2. Alternative proposed mechanisms for proline-catalyzed aldol reactions

There have been a number of computational [7,10,11,15-19] and experimental [20-25] studies of the reaction mechanism. List and coworkers [20] proved the majority of 10 crystal structures were consistent with the Houk-List transition states of proline-catalyzed reactions by comparing the crystal structures with the ground and transition state structures predicted by DFT. Meanwhile, Gschwind and coworkers [21] experimentally detected and characterized the enamine intermediates in proline-catalyzed aldol reaction and showed the direct formation of enamine carboxylic acids from oxazolidinones in DMSO. Following this, Sunoj and coworkers [18] explored the competing enamine and oxazolidinone pathways using DFT and MP2 calculations, and concluded that Houk-List transition model is better able to rationalize the experimentally observed enantio- and diastereoselectivities. Indeed, we recently used DFT calculations of the C-C bond formation step of the Houk-List model to successfully predict the experimentally observed enantioselectivities in reaction of acetone with *p*-nitrobenzaldehyde, as catalyzed by proline and its derivatives Me_2bdc-Pro (bdc = 1,4-benzenedicarboxylate) and Me_2bpdc-Pro (bpdc = 4,4′-biphenyldicarboxylate) [19].

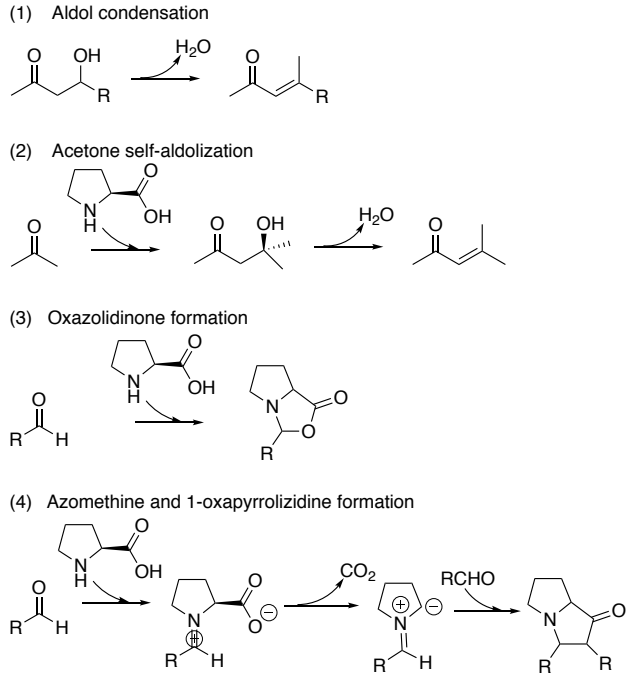


Scheme 3. Catalytic cycle of proline catalyzed aldol reactions [26].

While most experimental and computational evidence supports the C–C bond formation step proposed by the Houk-List model [18-20], the remaining details of the overall catalytic process have received little attention (Scheme 3). That is, the mechanism of enamine formation along with the release of catalyst and final aldol product are relatively poorly understood. The rate-determining step of the process also remains controversial. Houk and coworkers reported that the C–C bond formation is the rate-determining step in their early work [7,8,27]. The conclusion was then revised firstly in 2004, reporting both the enamine formation step via deprotonation by carboxylate group of the iminium and the subsequent C–C bond formation step as rate-determining steps because of their equal barrier heights [11]. By conducting a kinetic isotope effect study experimentally and computationally, Meyer, Houk and coworkers later specified that rate-limiting step precedes C–C bond formation [12]. Meanwhile, kinetic and isotope studies from Blackmond and coworkers concluded that the rate-determining step succeeds enamine formation but precedes product hydrolysis [22-25]. Moreover, proline-mediated aldol reactions can also undergo a variety of potential side-reactions (see Scheme 4), including the formation of the α, β-unsaturated ketone via aldol condensation and acetone self-aldolization catalyzed by proline.[8,27] Aromatic aldehydes (including *p*-nitrobenzaldehyde) can condense with proline to form azomethine which can undergo further 1,3-dipolar cycloaddition reaction [22,28,29]. While the products of these side reactions have been experimentally observed [22,28,29], their significance as a function of reaction conditions is yet to be established.

Herein we use high-level quantum chemical calculations to study the full catalytic cycle of proline-catalyzed intermolecular aldol reaction of acetone and *p*-nitrobenzaldehyde in acetone, including enamine formation, C–C bond formation, aldol product release, and side reactions. These results of these

calculations are then used to develop a first principles kinetic model of the entire process. That is, individual rate expressions for all known reaction pathways (including side reactions), populated with our calculated rate coefficients, are combined to form the induced mass-action kinetic differential equation that is then solved numerically for a wide range of different starting concentrations. In this way we can ascertain the relative importance of various steps in the mechanism as a function of reaction conditions and provide guidelines for optimizing the yields and product ratios.



Scheme 4. Side reactions considered in our study.

COMPUTATIONAL PROCEDURES

Quantum Chemistry. All standard ab initio calculations and density functional theory calculations were carried out using Gaussian 09 revision E.01 [30] and Molpro 2015 [31] software packages. Conformational searching was performed with the energy-directed tree search (EDTS) algorithm [32] so as to identify conformations with the lowest Gibbs free energy in solution. Solution-phase Gibbs free energies were obtained *via* a thermocycle in which Gibbs free energies in the gas phase at 298 K, as calculated via standard ideal gas partition functions, were combined with Gibbs free energies of solvation and the necessary phase change correction term [33]. The SMD solvent model [34] at the recommended [35] M062X/6-31+G(d) level of theory was used to calculate Gibbs free energies of solvation in acetone. For the gas-phase component, geometries were optimized at the M06-2X/6-31+G(d) level of theory [36]. Frequency calculations were carried out at the same level of theory. Geometries were verified either as local minima or transition states. Entropies, thermal corrections, and zero-point vibrational energies were scaled by recommended scale factors [37]. Single-point energies were calculated using the high-level composite ab initio method G3(MP2,CC) to improve the accuracy [38]. Having obtained the Gibbs free energy barriers in solution, corresponding rate coefficients were then obtained *via* the Eyring equation for use in the kinetic modelling. Benchmarking of the entire methodology is presented in the Supporting Information (Appendix S1) where we show that is capable of reproducing the experimental pKa value for proline in water (1.99 [39]) to within 0.5 pKa units (2.52).

Kinetic Modelling. Kinetic modelling was performed with the Modelica package, version 2.3.2, implemented within Mathematica 12 [40]. The kinetic modelling approach is as described in our previous

work [41]. An induced mass-action kinetic differential equation is constructed from the rate expressions of all forward and reverse reactions in the process. This is then solved numerically to yield concentrations of all species as a function of time, given a set of specified starting concentrations. By construction, the induced kinetic differential equation is mass balanced and well-defined, (i.e., the solution is unique, given unique starting conditions) yielding a concentrated-variable model (i.e. the model assumes no spatial inhomogeneities) [42]. Adaptive step sampling was conducted using the DASSL ODE solver, by default; no change in kinetic profile was observed with changes in step-tolerance (up to three orders of magnitude; minimum of 10^{-9}), or choice of ODE solver. Results were unchanged when the induced kinetic differential equation was solved manually with NDSolve. The kinetic mechanism is presented in Scheme S1 (Supporting Information) and includes a simplified model of the uncatalyzed process, several off-pathway reactions (as shown in Scheme 4), and the full proline-catalyzed intermolecular aldol reaction, including enamine formation, C-C bond formation, and product release. The full catalyzed mechanism includes all three possible pathways to enamine formation, all four possible C-C bond forming transition states with synchronous proton transfer, and the full mechanism of product release. All rate parameters are calculated from quantum chemistry with the exception of those that are diffusion limited. For these we universally assume a rate coefficient of 10^8 L mol^{-1} s^{-1}. The forward Gibbs free barriers ($\Delta G^{\ddagger}$) and reaction energies (ΔG) are listed in Table S2 of the Supporting Information, following the same nomenclature as Scheme S1. A dash in Table S2 indicates that the reaction was barrierless and thus assumed to be diffusion controlled in the forward direction. We use our model to make testable predictions of reaction outcomes over a wide range of experimental conditions, and provide the full set of reaction conditions and associated predictions in the Supporting Information. We also validate the numerical approach using the same approach for the uncatalyzed reaction and showing that the numerical and analytical solutions are identical (Appendix S3, Supporting Information).

RESULTS AND DISCUSSION

Part 1. Quantum-chemical calculations.

High-level ab initio molecular orbital calculations were first used to evaluate the kinetics and thermodynamics of every known mechanistic pathway for the proline catalyzed process, including all of the possible side reactions in Scheme 4, so as to provide the rate and equilibrium constants for incorporation into a kinetic model.

Enamine Formation. The first stage in the aldol catalytic cycle is the formation of the enamine intermediate (Figure 1). Noteworthy aspects of the process are the initial inversion of stereochemistry required to activate the proline, which proceeds through a low barrier (22.9 kJ mol^{-1}) [17], followed by a relatively high barrier (~80 kJ mol^{-1} on an absolute scale) to form the zwitterionic intermediate **E′**. Proton transfer involving some external source then occurs to form **F**, which is justified by the aldol reaction proceeding under acid or base conditions [43]. C–O bond cleavage can then occur *via* another barrier (75.4 kJ mol^{-1} on an absolute scale) to generate the iminium complex **G**. In the lowest energy pathway (TS4 in Figure 1), another proline molecule assists the transformation of the iminium complex to the nucleophilic enamine complex **H′**. This reaction occurs via TS4 rather than its charge separated form (TS4" in Figure 1) which is 17 kJ mol^{-1} higher in energy (76.9 kJ mol^{-1} versus 93.9 kJ mol^{-1}), though the latter is 10.2 kJ mol^{-1} more thermodynamically favorable. The corresponding intramolecular pathway for the formation of enamine (T4' in Figure 1), as proposed earlier by Houk and coworkers [11], has a slightly higher barrier (82.7 kJ mol^{-1} on absolute scale) than the lowest bimolecular pathway TS4 for the proton transfer step. Nonetheless, due to concentration effects, this intramolecular pathway actually dominates kinetically, except when the concentration of proline is high (*vide infra*).

Interestingly, the intramolecular (TS4') and zwitterionic bimolecular (TS4") pathways both lead to the *syn*-enamine, whereas the competing non-zwitterionic bimolecular pathway (TS4) leads to the *anti*. However, the relative importance of these pathways does not affect the *syn*- and *anti*-enamine concentrations, as these can interconvert through rapid N–C bond rotation ($\Delta G^{\ddagger}$ = 54.9 kJ mol^{-1} on an absolute scale). The reaction itself is essentially thermoneutral (ΔG = 1.0 kJ mol^{-1} for *syn* to *anti*), and considerably faster than the subsequent C-C bond formation steps (*vide infra*). As a result, and in accordance with the Curtin-Hammett principle [44], the *syn*- and *anti*-enamine concentrations are able to equilibrate and equalize prior to the C-C bond formation step, which is why the enantioselectivity has been found to be governed these subsequent steps rather than the enamine formation rates [18,19].

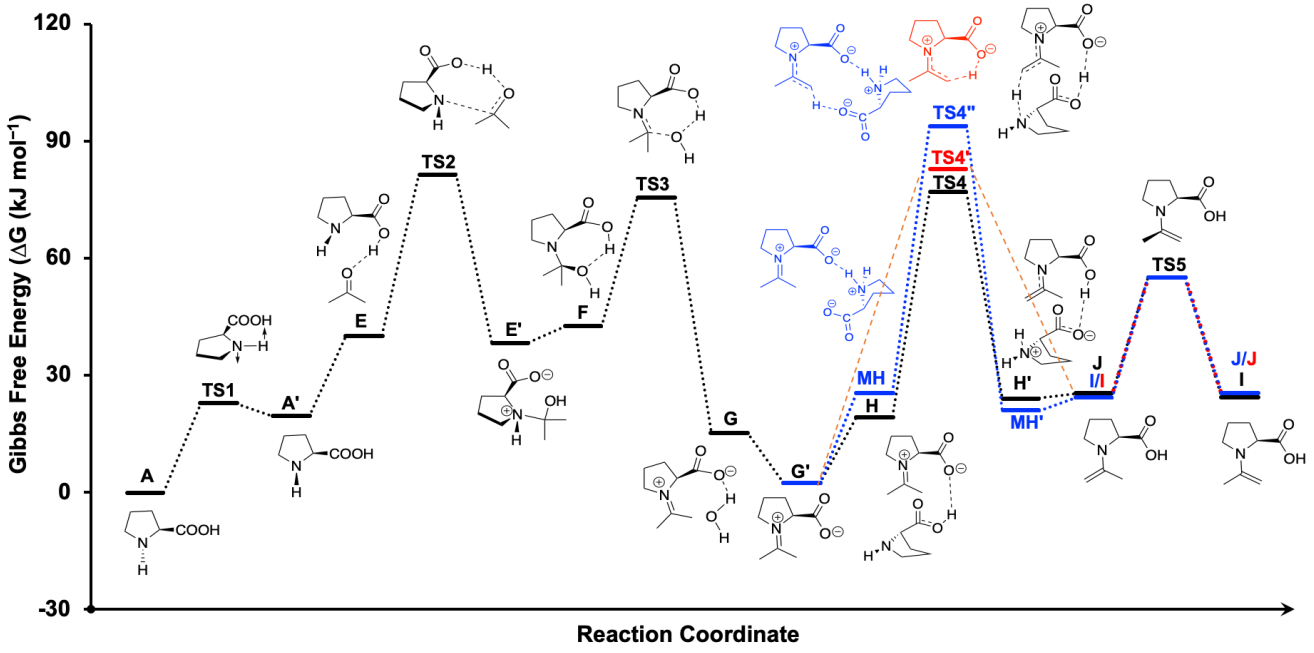
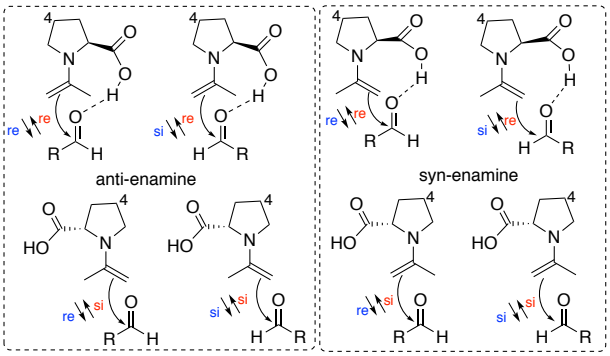


Figure 1. Gibbs free energy surface (25°C, kJ mol^{-1}) at G3(MP2,CC)//M062X/6-31+G(d) level of theory for enamine formation between acetone and proline.

C-C Bond Formation. Formation of the C–C bond in the proline-catalyzed aldol reaction has been extensively studied, and while most experimental evidence supports the Houk/List mechanism of C–C bond formation [20,21], several key mechanistic questions remain. Namely, does C–C bond formation occur simultaneously with proton transfer? Additionally, does the enamine react in the *syn* or *anti* form **(Scheme 5)**, and is the barrier to isomerization (**Figure 1, TS5**) relevant for the stereochemistry of the final product. To address these questions, **Figure 2** shows the four possible transition states describing C–C bond formation in which proton transfer is concerted for either the *syn* or *anti* enamine starting material, while Figure S2 of the Supporting Information shows those same transition states without concerted proton transfer. It should be noted that only the lowest energy conformations of the pyrrolidine ring (i.e., with the C4 "up" with respect to the carboxylic acid) are considered here as we have previously shown that consideration of the "down" conformers are significantly less stable (by up to 9.4 kJ/mol) and neglecting them does not affect the predicted enantiomeric excess [19]. In all cases the barrier heights for C-C bond formation with concerted proton transfer are up to 50 kJ mol^{-1} lower than those without, and thus we conclude that concerted proton transfer is necessary, consistent with the previous computational work of Houk and coworkers [7,8]. Houk and coworkers showed that proton transfer from the carboxylic acid of proline to the forming alkoxide is essential for charge stabilization and facilitates C–C bond formation in the transition state. Additionally, the lowest-energy transition states in **Figure 2** involve the *anti*-conformation of the enamine, which is why the rapid equilibration of the *anti*- and *syn*-enamine concentrations by N–C bond rotation is important. The TS_{arr} and TS_{srr} barriers to the C–C bond

formation are approximately 10 kJ mol^{-1} lower than that of TS_{ars} and TS_{srs}, respectively. Of these four transition states, TS_{arr} and TS_{srr} lead to the (*R*)-major aldol product, and TS_{ars} and TS_{srs} lead to the (*S*)-minor product, in agreement with experiment and as predicted by List and Houk and coworkers [10].



Scheme 5. Stereoselectivities in the C–C bond formation transition states. Red color means the enamine is attacked and blue color means the aldehyde is attacked. The uppermost structures feature simultaneous C–C bond formation and proton-transfer, while the lower structures feature asynchronous proton transfer.

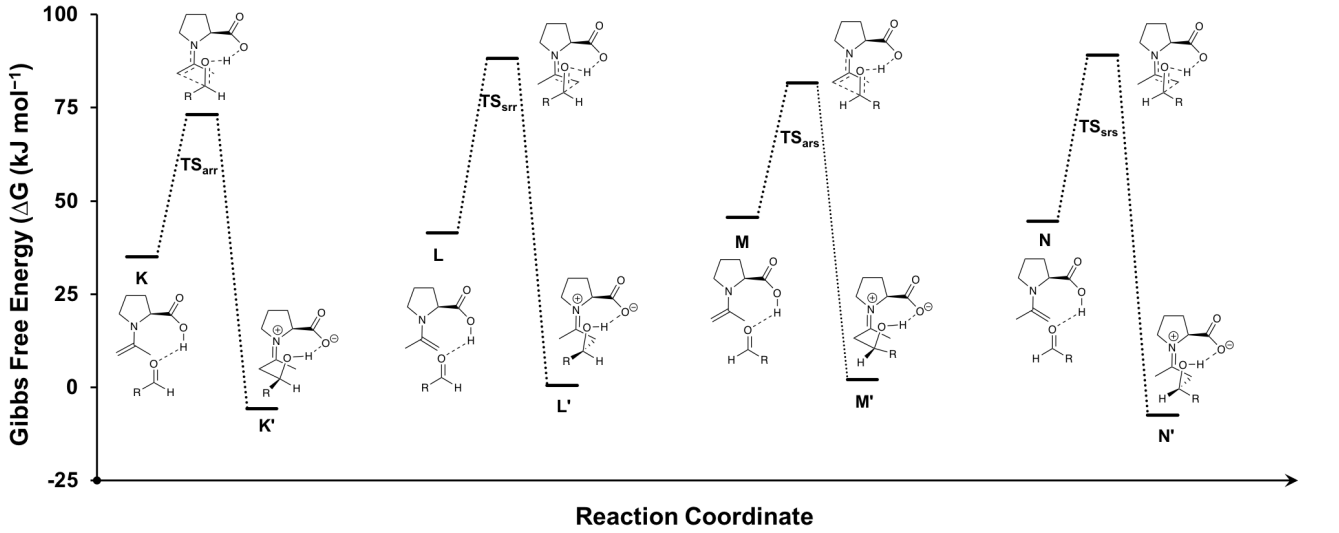


Figure 2. Gibbs free energy surface (25°C, kJ mol^{-1}) for C–C bond formation coupled with proton transfer at G3(MP2, CC)//M062X/6-31+G(d) level of theory. Here, R is *p*-nitrobenzaldehyde. The zero of energy is the original isolated reactants and catalyst.

Product Release. Taking the lowest-energy transition state for C–C bond formation (TS_{arr}), we next examined the mechanism of product release (Figure 3). The corresponding Gibbs free energy surfaces for product release after the other transition states, TS_{srr}, TS_{ars} and TS_{srs} are shown in the Supporting Information (Figure S3, Figure S4, and Figure S5 respectively). Interestingly, and contrary to previous reports stating that C–C bond formation was rate determining [7,8,15,16,22-25,27,45,46] we find here that product release proceeds through a higher barrier transition state. Once again, however, kinetic information is concentration dependent, and kinetic modelling shows that both C–C bond formation and product release contribute to the rate of product formation (*vide infra*).

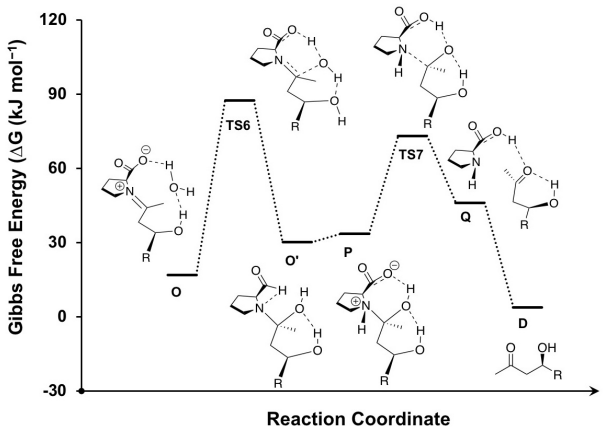


Figure 3. Gibbs free energy surface (25°C, kJ mol^{-1}) for product release following C–C bond formation *via* TS_{arr} at G3(MP2, CC)//M062X/6-31+G(d) level of theory. The zero of energy is the original isolated reactants and catalyst.

Side Reactions. Finally, several side reactions have also been considered in the catalytic cycle (see Scheme 4), including aldol product condensation to form the α,β-unsaturated ketone, acetone self-aldolization catalyzed by proline, as well as the experimentally observed oxazolidinone, azomethine and 1-oxapyrrolizidine formed by condensation of proline and *p*-nitrobenzaldehyde [22,28,29]. Starting with aldol product condensation, two pathways have been investigated (see Figure S6, Supporting Information). For acetone self-aldolization, the enamine reacts with a second acetone molecule via either *re* or *si* attack, and two different pathways were predicted (Figure S7 and Figure S8, Supporting Information). The calculated Gibbs free energy surfaces for the condensation reactions between proline and *p*-nitrobenzaldehyde are depicted in Figure S9 and Figure S10 (Supporting Information). All of the processes in Figures S6-S10 proceed through at least one very high barrier (> 160 kJ mol^{-1}) and thus are not expected to contribute significantly, but all are nonetheless included the kinetic model for the sake of completeness.

Part 2 Kinetic Modelling.

Having predicted energetics and hence rate and equilibrium coefficients for all known reaction pathways, we then used them to construct a complete kinetic model of the process, which was in turn used to study the mechanism of the process and identify reaction conditions for optimizing the enantiomeric excess (%ee) and the aldol product yields.

Model Validation. To help validate the model, we first simulated the conversion and %ee for the published experimental reaction conditions of Pikho [47]. In that work, the reactants, temperature and reported concentrations were same as in the simulations, but their reactions were performed in DMF, and data reported after 3 days instead of the 30 minutes used in the simulations (Figure 4). The effect of water on the conversion and %ee was consistent, but clearly the simulated reactions were much faster. This may in part reflect the low solubility, and hence lower than reported effective concentration, of proline in DMF. Considering the poor solubility of proline in DMF, we screened 66 conditions by varying the initial proline concentration (from 0.1 M to 0.002 M) while keeping the experimental conditions of acetone, aldehyde and water constant (see Table S5). By comparing the screening results with experiment, we found that if the proline concentration is decreased from 0.1 M to 0.02 M, while maintaining the other concentrations at their experimentally reported values, we can better reproduce the experimental conversions except when the concentration of water is nil (see Figure S16), which is likely due to the even poorer solubility of proline in DMF without water. Although the substrate is different (m-Cl-benzaldehyde), we also compared the effect of water on the reaction rate for the reaction conditions of Blackmond [22] and again obtained

qualitatively similar results, although in this case reaction rates are faster for Blackmond's substrate than ours (see Figure S11, Supporting Information).

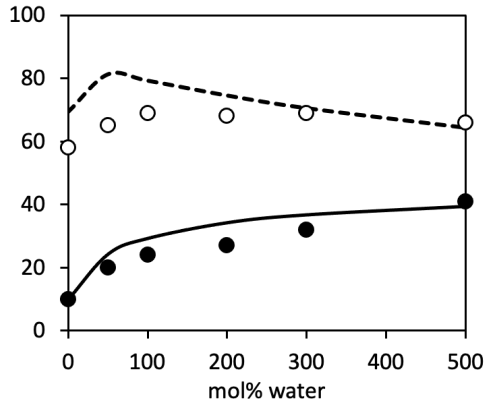


Figure 4. %Conversion (solid) and %ee (dashed) for the proline (10 mol%) catalysed aldol reaction of p-nitrobenzaldehyde (1M) and acetone (1M) as a function of water concentration. Solid lines are the kinetic model predictions after 30 minutes; points are the experimental data taken from Ref [47].

Process Optimization. Having validated the kinetic model, we can use it to identify optimal reaction conditions to, for example, maximise the %ee or the conversion to the major or minor product. To this end, we systematically screened 186 reaction condition sets in which the initial catalyst, acetone, and water concentrations were varied. The temperature was kept constant at 298 K, which corresponds to the typically used conditions in the literature. From these screening studies, the best conditions (with respect to conversion of the major and minor product, respectively) were identified (Figure S12). Regions of high product yield and %ee (Figure S12) were also identified, with good coverage of possible outcomes, such that one could in principle select a set of conditions to achieve expected %ee or product conversion. The full results of screening are provided in Table S3. We find that the high concentration of the aldehyde leads to very low aldol product conversion (< 10%) when the concentration of proline is very low (i.e. 0.002 M). If everything else is constant, the aldol product conversion increases with the increase of any of proline, acetone and water, but decreases when the aldehyde is increased. With the ratio of proline, acetone and aldehyde fixed, increasing the concentration of water would increase the aldol product conversion, which is consistent with the experimental results of Pihko and co-workers for the same reaction [47]. Interestingly, the enantiomeric excess was largely unaffected for all the screening conditions except for those conditions with extremely high "acetone-excess" (i.e. 2.6 M) and high concentration of proline (i.e. 0.03 M). The latter result might seem counterintuitive but is because at very high proline concentrations, the hydrolysis step and its role in determining enantioselectivity (vide infra) becomes relatively more important. The conditions which are predicted to maximize product conversion (i.e. 89.4%) are similar to those of List and co-workers, except for the initial concentration of proline, which likely differs due to its poor solubility. Finally, we chose condition 140 (Table S3) as our optimal model, that is [proline]=0.005M, [acetone]=2.25M, [aldehyde]=0.1M, and [water]=0.6M, resulting in 76.6% ee and 89.4% yield. Figure 5 shows that the predicted yield and enantiomeric excess conducted by kinetic modelling using this optimal model.

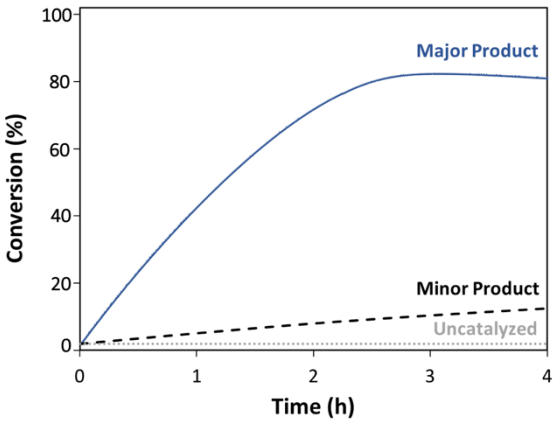


Figure 5. Predicted major and minor aldol product using our kinetic modelling with the optimal reaction condition ([proline]=0.005M, [acetone]=2.25M, [aldehyde]=0.1M, and [water]=0.6M). The predicted conversion and %ee are 89.4% and 68 % respectively, compared with experimental conversion of 68% and %ee of 76%.

Order in Proline. With our optimal kinetic model in hand, we can use this to assess the relative importance of different reactions and side reactions in the overall reaction scheme. Under optimal conditions, the predicted mechanism of enamine formation is mainly first-order with respect to (*S*)-proline and occurs through an intramolecular proton transfer to yield the *syn*-enamine (55.0%, TS4' in Figure 1), though the lower barrier pathway that is second-order with respect (*S*)-proline contributes 44.8% (TS4 in Figure 1), and the other higher-energy zwitterionic second order pathway examined contributes less than 0.3% (TS4" in Figure 1). This is notable because the first-order pathway has a higher barrier to enamine formation (82.7 kJ mol^{-1} on an absolute scale) than the lowest second-order pathway (76.9 kJ mol^{-1}), which illustrates the importance of kinetic modelling when discriminating between plausible pathways that exhibit similar barriers, but have different stoichiometric requirements. In this regard, it is also worth noting that the order of the reaction in proline is dependent on the proline concentration. Figure 6 shows the % contribution of the three pathways as a function of proline concentration, while keeping the concentrations of [water] = 0.6 M, [aldehyde] = 0.1 M, and [acetone] = 2.25 M constant. Increasing the initial concentration of proline from 0.001 to 2.25 M, we find that the proportion of enamine formation from the first order pathway (TS4') decreases from 80.9% to 32.2%, while that of the second order pathway (TS4) increases until the concentration of proline reaches 0.01 M but then also decreases in favour of the zwitterionic bimolecular pathway TS4" which ultimately increases to 39.6%. This zwitterionic pathway has highest barrier of all three on an absolute scale (93.9 kJ mol^{-1}) but is 10.2 kJ mol^{-1} more thermodynamically favourable than the lower barrier bimolecular pathway (TS4), which is why it becomes significant at very high proline concentrations. Summing up when the initial concentration of proline is low (< 0.005 M) the first-order enamine formation pathway dominates; while the second-order enamine pathways become significant at higher catalyst concentrations. This dependence on catalyst concentration, coupled with the low solubility of proline under typical experimental conditions, explains the conflicting conclusions in the literature with respect to the order of the reaction in proline [3-12,22-25,48].

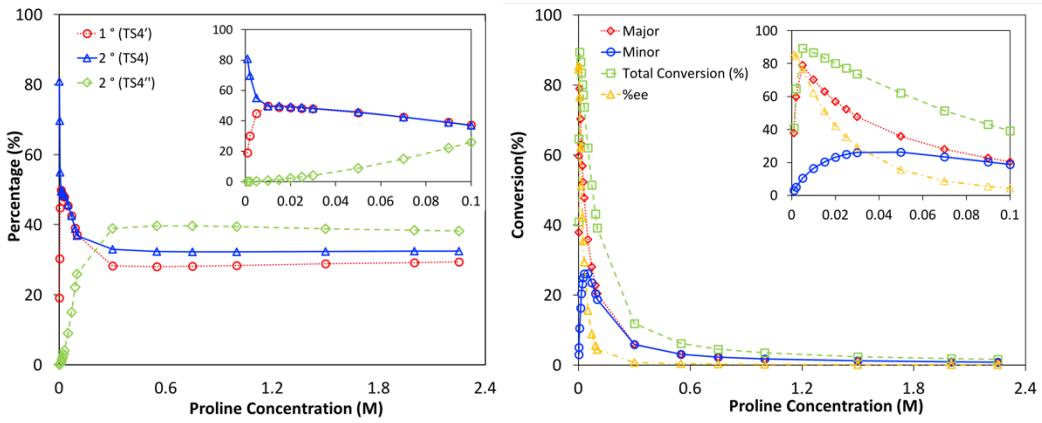


Figure 6. The enamine formation proportions from the three enamine formation pathways (left) and the conversion of aldol products (including the major and minor aldol products) and enantiomeric excess (right) with increasing the concentration of proline from 0.001M to 2.25M while keeping [acetone] = 2.25 M, [aldehyde] =0.1 M and [water] = 0.6 M.

Stereochemistry. As anticipated, under the optimal kinetic model, the stereochemistry is governed entirely by the barrier heights of the alternative C-C bond formation transition states. This is because the concentrations of the *anti*- and *syn*-enamine equilibrate and equalize via a rapid thermoneutral N-C bond rotation step. When this ability to interconvert is artificially removed (i.e., when the barrier to rotation is made infinitely high), we observe that the predicted enantiomeric excess decreases from 76.6 % to 40.4% under our optimal conditions, and the conversion decreases only slightly, from 89.4% to 87.7%. The major (R)-aldol product is formed in roughly equal proportion *via re*-facial attack of the *p*-nitrobenzaldehyde upon the *re*-face of the *anti*-enamine and *syn*-enamine (50.4 and 49.6% respectively), for which we predict the barriers to C-C bond with concerted proton transfer to be 38.1 (TS_{arr}) and 46.7 (TS_{srr}) kJ/mol, respectively. The reason that both contribute more or less equally, is their barriers for the corresponding hydrolysis are 70.6 and 46.3 kJ/mol, respectively. This highlights the key role of hydrolysis, and the reason why the %ee depends on water and other reaction conditions. That the %ee was successfully predicted in previous studies based on the C-C bond formation alone [18,19], is due to the fact that both of these pathways lead to the same product. The minor (S)-aldol product is formed almost exclusively *via si*-facial attack of the *p*-nitrobenzaldehyde upon the *re*-face of the *syn*-enamine (70.2%) and via *si*-facial attack of the *p*-nitrobenzaldehyde upon the *re*-face of the *anti*-enamine (29.8%) (Figure 7).

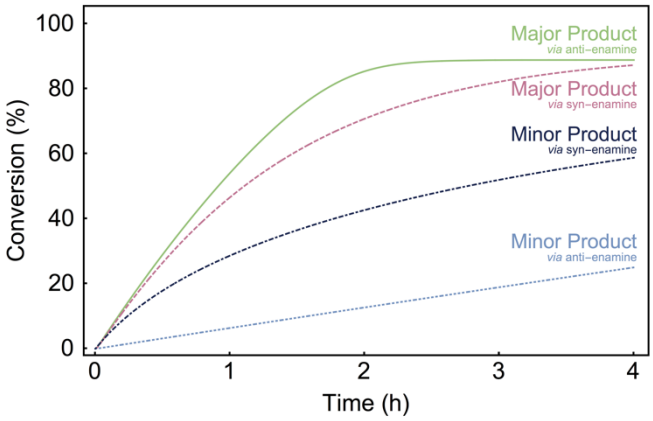


Figure 7. Product Formation Pathways. Selectively inhibiting pathways in our kinetic model to construct model subsets yields mechanistic information that is otherwise difficult to access experimentally. We find that the major (R)-aldol product is formed in roughly equal proportion via re-facial attack of the p-nitrobenzaldehyde upon the re-face of the anti-enamine and syn-

enamine (50.4 and 49.6% respectively), which we predict the barriers to C-C bond formation (38.1 and 46.7 kJ/mol, respectively) and the following hydrolysis process (70.6 and 46.3 kJ/mol, respectively).

Rate Determining Steps. As noted in the Introduction, there has been significant debate as to which step(s) in the process are rate determining [12,15,22-24]. We note in passing that concept of a "rate-determining step" has itself been debated in the literature (see for example [49] for an excellent discussion of this problem). Here we use the IUPAC definition of a "rate controlling step", that is "an elementary reaction the rate constant for which exerts a strong effect—stronger than that of any other rate constant— on the overall rate" [44]. To assess which step(s) of the model are rate determining, we used our optimal concentrations and artificially modified our rate coefficients in turn by ±10% to assess which rates had the largest impact on the conversion (Figures S14 and S15, Supporting Information). Our rate sensitivity analysis suggests that changes to initial stages of enamine formation (i.e. k_3 to k_7), the rate of enamine isomerisation (i.e. k_{14}), the rate of *re*-facial attack of the *p*-nitrobenzaldehyde upon the *re*-face of the *anti*-enamine (i.e. k_{15} to k_{16}), and the rate of water complexation with the aldol adduct (which initiates product release, i.e. k_{17} to k_{19}) have the greatest effect upon the rate of conversion to the major (*R*)-aldol product (Figure S14, Supporting Information). Likewise, the rate of conversion to the minor (S)-aldol product is most sensitive to changes in the rate of enamine isomerisation (i.e. k_{14}), the rate of *si*-facial attack of the *p*-nitrobenzaldehyde upon the *re*-face and *si*-face of the *syn*-enamine and the rate of water complexation with the aldol adduct (i.e. k_{22} to k_{25} and k_{29} to k_{31}) (Figure S15, Supporting Information). The rate of conversion to the minor product is also moderately sensitive to changes in the rate of major product formation *via* the preferred mechanism detailed above; this merely reiterates that the rate of minor product conversion is adversely affected by consumption of the aldehyde substrate to form the major (*R*)-aldol product. To sum up, under our optimal reaction conditions, the process of initial enamine formation, C-C bond formation and the initial stages of product release all contribute to the overall reaction rate, in agreement with our analysis of the mechanism of C-C bond formation previously. The observation that all of these steps contribute to the rate under these conditions, and that their relative contributions is dependent on the concentrations used, goes some way to explaining the conflicting conclusions in the literature as to which step is rate determining [12,15,22-24] — in reality all are under the right conditions

Side Reactivity. Taking our optimal reaction conditions, we also artificially modified our model so as to "switch off" side reactions, with a view to elucidating the various factors controlling off-target reactivity. It is well-understood that additional water suppresses side-reactivity by suppressing spectator species [22] but the effects upon off-target reactivity of simultaneously changing the concentrations of acetone, (*S*)-proline, water, and also the aldehyde is less well-understood, and is difficult to discern experimentally. To that end, we screened the same 186 reaction conditions sets that we used to optimise the model and compared their results to the same set of parameters with the barriers to side-product formation made unfeasibly high. The full results of screening are provided in Table S4 and Figure S13. As expected, for over 50% all tested cases, removing the possibility of side-reactivity increased the predicted yield, while less than 50% of the tested cases decreased the predicted yield. However, in all cases the predicted effect of removing side-reactivity was very small, with a maximal conversion difference of ~2.0% relative to the optimized kinetic model. Side-reactivity is predicted to be minimized (< 0.3%) at high "acetone-excess" (>2.0 M), and where the "acetone-excess" smaller than 1.5 M results in the most significant side-reactivity (~2.0%). Meanwhile, we find the side-reactivity is sensitive to the concentrations of proline and aldehyde when the concentration of acetone is fixed.

CONCLUSIONS

Herein, using the reaction of *p*-nitrobenzaldehyde and acetone as a case study, we propose a revised mechanism for intermolecular proline catalyzed aldol reactions, and use it to derive optimal reaction conditions for the process. Using a combination of G3(MP2,CC)//M062X/6-31+G(d)/SMD quantum-chemical calculations and ab initio kinetic modelling we show that, contrary to previous reports, the enamine formation benefits from participation of a second proline moiety except when the proline concentration is very low (< 0.005M). However, once the enamine is formed, C–C bond formation occurs via the established List/Houk mechanism. We show that the relative rates of formation of the *syn* and *anti*-enamine are not important, as these interconvert via C-N bond rotation and equilibrate faster than their subsequent reaction. However, although the stereochemistry has been previously successfully predicted from the C-C bond formation barriers alone [18,19], our kinetic analysis shows that the subsequent rates of hydrolysis also influence the relative contributions of the different C-C bond formation pathways though not in a manner that changes the %ee in this instance. Moreover, while C–C bond formation is normally considered rate determining [3,4,22-25,48], the kinetic model is more complicated than this and under typically used concentrations, the process of initial enamine formation, C-C bond formation and the initial stages of product release all contribute to the overall reaction rate. Using our kinetic model, we predict that yield and %ee are optimal for concentrations of [proline] = 0.005 M, [acetone] = 2.25 M, [aldehyde] = 0.1 M, and [water] = 0.6 M. Using excess acetone (< 2.6 M) increases both conversion and %ee, excess aldehyde increases %ee but decreases conversion, and excess catalyst increases the conversion but decreases %ee. Aside from the indirect effect of increasing the solubility of the proline catalyst, water increases both conversion and %ee up to a point, but at large concentrations (> 1.0 M) excess water is expected to decrease %ee. Side reactions were shown to have a negligible effect (< 2%) on conversion and %ee under the full range of conditions explored.

ASSOCIATED CONTENT

Supporting Information

Complete computational and kinetic modeling details.

Corresponding Author

* Email: michelle.coote@anu.edu.au

ACKNOWLEDGMENT

The authors acknowledge financial support from the Australian Research Council (ARC) Centre of Excellence for Electromaterials Science, an ARC Laureate Fellowship (to M.L.C.), and generous supercomputing time from the National Computational Infrastructure. The authors acknowledge helpful discussions with Dr. Stephen Dale.

Funding Sources

Australian Research Council (FL170100041, CE140100012)

REFERENCES

1. Eder U, Sauer G, Wiechert R (1971) New type of asymmetric cyclization to optically active steroid CD partial structures. Angew Chem Int Ed 10:496-497.
2. Hajos ZG, Parrish DR (1974) Asymmetric synthesis of bicyclic intermediates of natural product chemistry. J Org Chem 39:1615-1621.
3. List B, Lerner RA, Barbas CF (2000) Proline-Catalyzed Direct Asymmetric Aldol Reactions. J Am Chem Soc 122:2395-2396.
4. Jung ME (1976) A review of annulation. Tetrahedron 32:3-31.

5. Puchot C, Samuel O, Dunach E, Zhao S, Agami C, Kagan H (1986) Nonlinear effects in asymmetric synthesis. Examples in asymmetric oxidations and aldolization reactions. J Am Chem Soc 108:2353-2357.
6. List B, Hoang L, Martin HJ (2004) New mechanistic studies on the proline-catalyzed aldol reaction. Proc Natl Acad Sci USA 101:5839-5842.
7. Bahmanyar S, Houk KN (2001) Transition states of amine-catalyzed aldol reactions involving enamine intermediates: theoretical studies of mechanism, reactivity, and stereoselectivity. J Am Chem Soc 123:11273-11283.
8. Bahmanyar S, Houk KN (2001) The origin of stereoselectivity in proline-catalyzed intramolecular aldol reactions. J Am Chem Soc 123:12911-12912.
9. Bahmanyar S, Houk KN (2003) Origins of opposite absolute stereoselectivities in proline-catalyzed direct Mannich and aldol reactions. Org Lett 5:1249-1251.
10. Bahmanyar S, Houk KN, Martin HJ, List B (2003) Quantum mechanical predictions of the stereoselectivities of proline-catalyzed asymmetric intermolecular aldol reactions. J Am Chem Soc 125:2475-2479.
11. Clemente FR, Houk KN (2004) Computational evidence for the enamine mechanism of intramolecular aldol reactions catalyzed by proline. Angew Chem 116:5890-5892.
12. Zhu H, Clemente FR, Houk KN, Meyer MP (2009) Rate Limiting Step Precedes C− C Bond Formation in the Archetypical Proline-Catalyzed Intramolecular Aldol Reaction. J Am Chem Soc 131:1632-1633.
13. Armstrong A, Boto RA, Dingwall P, Contreras-García J, Harvey MJ, Mason NJ, Rzepa HS (2014) The Houk–List transition states for organocatalytic mechanisms revisited. Chem Sci 5:2057-2071.
14. Rajagopal D, Moni M, Subramanian S, Swaminathan S (1999) Proline mediated asymmetric ketol cyclization: a template reaction. Tetrahedron: Asymmetry 10:1631-1634.
15. Rankin KN, Gauld JW, Boyd RJ (2002) Density functional study of the proline-catalyzed direct aldol reaction. J Phys Chem A 106:5155-5159.
16. Arnó M, Domingo LR (2002) Density functional theory study of the mechanism of the proline-catalyzed intermolecular aldol reaction. Theor Chem Acc 108:232-239.
17. Ajitha MJ, Suresh CH (2011) A higher energy conformer of (S)-proline is the active catalyst in intermolecular aldol reaction: evidence from DFT calculations. J Mol Catal A Chem 345:37-43.
18. Sharma AK, Sunoj RB (2010) Enamine versus Oxazolidinone: What Controls Stereoselectivity in Proline‐Catalyzed Asymmetric Aldol Reactions? Angew Chem Int Ed 122:6517-6521.
19. Yu L-J, Coote ML (2021) Electrostatic Switching of Stereoselectivity in Aldol Reactions. J Org Chem 86:9076–9083.
20. Bock DA, Lehmann CW, List B (2010) Crystal structures of proline-derived enamines. Proc Natl Acad Sci USA 107:20636-20641.
21. Schmid MB, Zeitler K, Gschwind RM (2010) The elusive enamine intermediate in proline‐catalyzed aldol reactions: NMR detection, formation pathway, and stabilization trends. Angew Chem 49:4997-5003.
22. Zotova N, Franzke A, Armstrong A, Blackmond DG (2007) Clarification of the role of water in proline-mediated aldol reactions. J Am Chem Soc 129:15100-15101.
23. Zotova N, Broadbelt LJ, Armstrong A, Blackmond DG (2009) Kinetic and mechanistic studies of proline-mediated direct intermolecular aldol reactions. Bioorg Med Chem Lett 19:3934-3937.
24. Zotova N, Moran A, Armstrong A, Blackmond DG (2009) A Coherent Mechanistic Rationale for Additive Effects and Autoinductive Behaviour in Proline-Mediated Reactions. Adv Synth Catal 351:2765-2769.
25. Mathew SP, Klussmann M, Iwamura H, Wells JDH, Armstrong A, Blackmond DG (2006) A mechanistic rationalization of unusual kinetic behavior in proline-mediated C–O and C–N bond-forming reactions. Chem Commun:4291-4293.
26. Cheong PH-Y, Legault CY, Um JM, Çelebi-Ölçüm N, Houk KN (2011) Quantum mechanical investigations of organocatalysis: mechanisms, reactivities, and selectivities. Chem Rev 111:5042-5137.
27. Hoang L, Bahmanyar S, Houk KN, List B (2003) Kinetic and stereochemical evidence for the involvement of only one proline molecule in the transition states of proline-catalyzed intra-and intermolecular aldol reactions. J Am Chem Soc 125:16-17.
28. Orsini F, Pelizzoni F, Forte M, Sisti M, Bombieri G, Benetollo F (1989) Behaviour of aminoacids and aliphatic aldehydes in dipolar aprotic solvents: Formation of oxazolidinones—behaviour of aminoacids and aliphatic aldehydes in dipolar aprotic solvents. J Heterocycl Chem 26:837-841.
29. Seebach D, Beck AK, Badine DM, Limbach M, Eschenmoser A, Treasurywala AM, Hobi R, Prikoszovich W, Linder B (2007) Are oxazolidinones really unproductive, parasitic species in proline catalysis?–Thoughts and experiments pointing to an alternative view. Helv Chim Acta 90:425-471.
30. Frisch M, Trucks G, Schlegel H, Scuseria G, Robb M, Cheeseman J, Scalmani G, Barone V, Petersson G, Nakatsuji H, al. e (2016) Gaussian 09, Revision E. 01, Gaussian. Inc, Wallingford CT.
31. Werner H, Knowles P, Knizia G, Manby F, Schütz M, Celani P, Györffy W, Kats D, Korona T, Lindh R, al. e (2015) MOLPRO, version 2015.1, a package of ab initio programs.
32. Izgorodina EI, Lin CY, Coote ML (2007) Energy-directed tree search: an efficient systematic algorithm for finding the lowest energy conformation of molecules. Phys Chem Chem Phys 9: 2507-2516.

33. Ho JM, Klamt A, Coote ML (2010) Comment on the Correct Use of Continuum Solvent Models. J Phys Chem A 114:13442-13444.
34. Marenich AV, Cramer CJ, Truhlar DG (2009) Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. J Phys Chem B 113:6378-6396.
35. Xu L, Coote ML (2019) Methods To Improve the Calculations of Solvation Model Density Solvation Free Energies and Associated Aqueous pKa Values: Comparison between Choosing an Optimal Theoretical Level, Solute Cavity Scaling, and Using Explicit Solvent Molecules. J Phys Chem A 123:7430-7438.
36. Zhao Y, Truhlar DG (2008) The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. Theor Chem Acc 120:215-241.
37. Alecu I, Zheng J, Zhao Y, Truhlar DG (2010) Computational thermochemistry: scale factor databases and scale factors for vibrational frequencies obtained from electronic model chemistries. J Chem Theory Comput 6:2872-2887.
38. Curtiss LA, Raghavachari K, Redfern PC, Baboul AG, Pople JA (1999) Gaussian-3 theory using coupled cluster energies. Chem Phys Lett 314:101–107.
39. O'Neil MJ (ed) (2018) The Merck Index - An Encyclopedia of Chemicals, Drugs, and Biologicals. Royal Society of Chemistry, Cambridge, UK.
40. Wolfram Research I (2019) Mathematica. Version 12.0 edn. Wolfram Research, Inc., Champaign, Illinois.
41. Blyth MT, Coote ML (2019) A pH-Switchable Electrostatic Catalyst for the Diels–Alder Reaction: Progress toward Synthetically Viable Electrostatic Catalysis. J Org Chem 84:1517-1522.
42. Tóth J, Nagy AL, Papp D (2018) The Induced Kinetic Differential Equation. Reaction Kinetics: Exercises, Programs, and Theorems. Springer.
43. Zhang X, Houk KN (2005) Acid/base catalysis by pure water: the aldol reaction. J Org Chem 70:9712-9716.
44. IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson. Blackwell Scientific Publications, Oxford (1997). Online version (2019-) created by S. J. Chalk. ISBN 0-9678550-9-8. https://doi.org/10.1351/goldbook.
45. Reymond J-L, Chen Y (1995) Catalytic, enantioselective aldol reaction with an artificial aldolase assembled from a primary amine and an antibody. J Org Chem 60:6970-6979.
46. Reymond J-L (1998) Stereoselectivity of aldolase catalytic antibodies. J Mol Catal B Enzym 5:331-337.
47. Nyberg AI, Usano A, Pihko PM (2004) Proline-Catalyzed Ketone-Aldehyde Aldol Reactions are Accelerated by Water. Synlett 2004:1891-1896.
48. Klussmann M, Mathew SP, Iwamura H, Wells Jr. DH, Armstrong A, Blackmond DG (2006) Kinetic Rationalization of Nonlinear Effects in Asymmetric Catalysis Based on Phase Behavior. Angew Chem Int Ed 45:7989-7992.
49. Kozuch S, Martin JML (2011) The Rate-Determining Step is Dead. Long Live the Rate-Determining State! ChemPhysChem 12:1413-1418.