RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

Journal Name

ARTICLE

Density functional theory study of the mechanism in dipeptide-catalyzed intermolecular aldol reaction---the effects of steric repulsion interactions on stereoselectivity

Xiaofei Zhang[a,b], Min Pu[*a]

The mechanism of the dipeptide (S)-pro-(S)-asp catalyzed intermolecular aldol reaction with acetone as the donor and three aromatic aldehydes (benzaldehyde, p-methyl benzaldehyde and p-nitrobenzaldehyde) as the acceptors was studied by means of density functional theory (DFT) at the level of B3LYP/6-31G(d,p). The calculated results showed that there were four steps in the reaction path: i. the nucleophilic attack of amino group on carbonyl for the formation of intermediate A, which was the rate-determining step due to the largest energy barrier of 44.33 kcal/mol; ii. the dehydration process to form s-cis- or s-trans-enamine through imine-generating step; iii. the electrophilic addition of aldehyde, which decided the stereoselectivity of product because of the steric repulsion interactions between enamine and aldehyde; iv. the removal of dipeptide to the final products. According to the results analysis, it was found that the dipeptide-catalyzed aldol reaction via s-trans-enamine was more energy favorable to obtain R-product (with an ee value > 99%). The energy variations in the reaction path were verified by using CAM-B3LYP and M06-2X methods in the same basis set. The solvation effects were explored based on B3LYP/6-31G(d,p) combined with a polarizable continuum model (PCM), the substituent effects of aromatic aldehydes were also considered. The computed results provided a reference for experiment that DMSO and H_2O as the solvents could decrease the energy barriers in reaction path and the impact of substituent effects might be small. The feasibility of dipeptide provided a possibility for the protein as the catalyst which was green and nontoxic.

1. Introduction

A series of peptides and amino acids used as catalysts for the asymmetric intermolecular aldol reaction which was an important C-C bond-forming reaction in synthesis of organic products and pharmaceuticals had been studied and attracted much more attention.[1-13] L-proline-based small peptides were developed as efficient catalysts for the asymmetric direct aldol reactions of acetone with aldehydes, and high yield and stereoselectivity of up to 96% ee obtained in aqueous media.[14] The subsequently research approved that small di- to tetra-peptides with a primary amine functionality also catalyzed the asymmetric intermolecular aldol reaction with excellent enantioselectivity.[15,16] In addition, the small peptides could catalyze the aldol reaction in water and non-aqueous media with high yield asymmetric product.[17] Córdova et al. investigated the origins of the stereoselectivity for asymmetric intermolecular aldol reaction by using DFT calculations.[18,19] The results showed that amino acids with a primary amino

functionality were able to catalyze this kind of reaction with high stereoselectivity.

Heine and Wong et al.[20,21] researched the enzymatic-catalyzed asymmetric aldol reaction and concluded that the reaction involved two classes intermediates, Type I aldolases (enamine intermediates) and Type II aldolases (Zinc enolates intermediates). In Type I aldolases, the primary amino group of amino acids residue had a large effect to the reaction process, in which formed enamine intermediates with the help of proton transfer by neighboring amino acids. In Type II aldolases, chiral enolates, chiral aldehydes, or chiral auxiliaries were used to control the stereoselectivity of aldol reactions.

In 1974, Hajos et al.[22] found that amino acids were able to catalyze the asymmetric aldol reaction via an enamine intermediate path. Nevertheless, it was not rapidly developed until List et al.[23,24] demonstrated that proline and its derivatives were highly enantioselective catalysts for the intermolecular asymmetric aldol reaction between ketones and aldehydes in

2000. From then on, amino acids were used more often as highly selective catalysts for intermolecular aldol reaction.[25,26] Generally speaking, this kind of proline-catalyzed reaction involved an enamine-forming process where one proline molecule took part in the transition state as suggested by Houk and List.[27-30] A lot of work to get high efficient and stereoselective asymmetric catalysts had been carried out since that time, in which the catalysis of proline derivatives and the expansion of scope of substrates were widely researched. [31-33] The high dr (up to 18:1) and *ee* value (up to 98%) of the targeted product were obtained by using primary amine-containing amino acids as catalysts, such as L-threonine and O-tBu-L-tyrosine.[34]

Recently, small peptides were also suggested as effective catalysts for the asymmetric aldol reaction.[35] That isolated from living cells in vitro could catalyze the asymmetric aldol reaction, demonstrating that non-enzymatic enantioselective catalysis could occur in living cells and be of biological relevance.[17] The beneficial effect of small peptide catalyzed asymmetric aldol reaction was due to the faster hydrolysis of the intermediates in the enamine catalytic cycle, as well as the suppression of non-productive imidazolidinone formation.[15] However, the mechanism of small peptide catalyzed intermolecular aldol reaction was not distinct for the moment.

Proline typically gave superior results of yield and enantioselectivity in the above mentioned catalytic studies, even promoted the finding of several other catalysts. To the best of our knowledge, there were few theoretical investigations concerning the detailed variation of the dipeptide-catalyzed aldol reactions between ketones and aromatic aldehydes, although great efforts had been made to explain the stereoselectivity of reactions. Accordingly, the mechanism and stereoselectivity of intermolecular aldol reaction catalyzed by dipeptide were researched and the transition states of C-C bond formation step were found by using density functional theory methods in this paper. The calculations successfully predicted the stereoselectivity of the observed product and provided the key insights into the detailed reaction paths.

2. Computational Methods

All the calculations were performed using the gradient-corrected function B3LYP of Density Functional Theory (DFT). The traditional hybrid Becke, three-parameter, Lee-Yang-Parr (B3LYP) exchange correlation functions, the long-range-corrected coulomb-attenuating method functional (CAM-B3LYP) and hybrid meta-generalized gradient approximation functional (M06-2X) of DFT were applied. The geometries were optimized using the standard double zeta plus polarization basis set 6-31G(d,p). Meanwhile, the single point energies of species in gas were tested and verified at the level of B3LYP/6-311++G(d,p) which listed in ESI. Frequency calculations were performed at the same theory level as the optimizations to obtain zero-point energies (ZPEs) and to confirm the nature of the stationary points. To investigate the transition states to the

two desired minima of the proposed mechanism, the potential energy profile was generated by using the intrinsic reaction coordinate (IRC) calculations. The solvent effects on the mechanism were discussed based on the polarizable continuum model (PCM) of self-consistent reaction field (SCRF) theory. All of the quantum chemical calculations were applied by using Gaussian 09 programs.[36]

3. Results and Discussion

3.1 Dipeptide structure

Dipeptide, perhaps, functioned as a "micro-aldolase" that provided both the nucleophilic amino group and electrophilic carbonyl cocatalyst in the form of the carboxylate. To investigate the mechanism of dipeptide-catalyzed intermolecular aldol reaction, a model with acetone as the donor and benzaldehyde as the acceptor was chosen to study (*S*)-pro-(*S*)-asp-catalyzed aldol reaction. The catalyst (*S*)-pro-(*S*)-asp was formed by the dehydration of (*S*)-proline and (*S*)-aspartic acid. As shown in Fig.1, the four atoms of peptide bond and its attached C atoms of both ends were in the same plane. The peptide bond was stable covalent bond because the bonding of C-N was sp^2 hybrid in which the conjugation structure was subsistent.
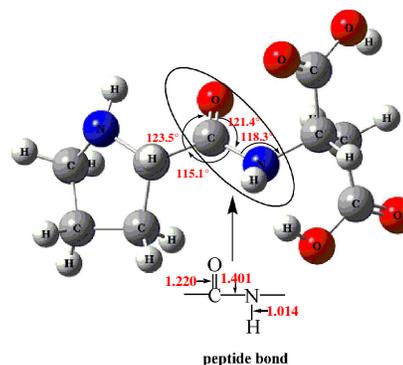

Fig.1 Optimized structure of dipeptide (*S*)-pro-(*S*)-asp

3.2 Reaction mechanism

The most possible mechanism of the dipeptide-catalyzed intermolecular aldol reaction was depicted in Fig.2. The proposed mechanism contained four steps: **i.** the nucleophilic attack on carbonyl by amino group for the formation of intermediate A; **ii.** the dehydration process to form *s-cis-* or *s-trans*-enamine; **iii.** the electrophilic addition of aldehyde; **iv.** the deprivation of catalyst-dipeptide to the final products. The discussions will be carried out in detailed step by step in the following parts. Pioneering theoretical studies had determined that the aldol reactions proceed via enamine intermediates, while the transition states for the crucial C-C bond-forming step (nucleophilic addition of the enamine intermediate to an electrophilically activate aldehyde) showed that an arrangement of the reacting atoms was stabilized by a hydrogen-bonding interaction between the proton of the amino-group moiety in

peptide bond and the oxygen atom of the electrophile.[37-39] However, the stereoselectivity of products was subjected by the spatial orientation and steric repulsion of groups in transition states.
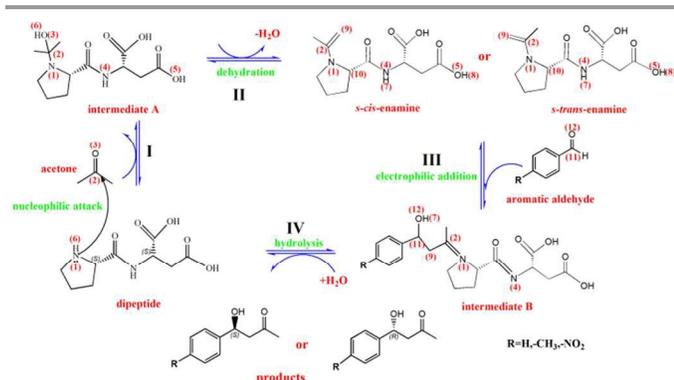

Fig.2 Reaction mechanism of the dipeptide-catalyzed aldol reaction

3.2.1 FORMATION OF INTERMEDIATE A. The mechanism of first step involved the attack of dipeptide on the C atom of carbonyl group to form intermediate A that acted as the nucleophile in the subsequent C-C bond formation with concomitant N-H···O hydrogen transfer. Initially, dipeptide reacted with acetone by nucleophilic addition via transition state TS1 (Fig.3) to form the intermediate A with the energy barrier of 44.33 kcal/mol. With the approaching of dipeptide to acetone, the electrostatic attraction between N(1) and C(2) atoms led to the formation of the intermediate A and was associated with the proton H(6) transferred from N(1) to the carbonyl O(3) atom. Due to the formation of N(1)-C(2) bond, it has also a charge transfer in the first step. The energy barrier of this step was 44.33 kcal/mol, which was not a low energy barrier at the room temperature. As discussed above, the first step was considered to be the rate-determining step, so full mixing of catalyst and reactants even operated at the elevated temperatures were essential under the experimental condition.
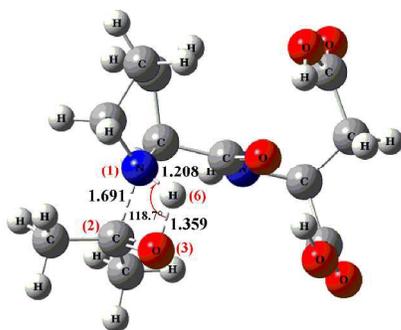

Fig.3 Optimized structure of transition state TS1

3.2.2 DIFFERENT PATHS OF ENAMINES FORMING. The second step was dehydration process of intermediate A to form *s-cis-* or *s-trans-*enamine. At the beginning, we assumed that the

dehydration was hydroxyl reacted with the hydrogen atom of either methyl to form enamine (**C-path** in Fig.4). The calculated energy barriers of this path were 65.54 (*s-cis-*enamine) and 64.06 kcal/mol (*s-trans-*enamine), respectively, which were not low energy barriers for the room temperature.
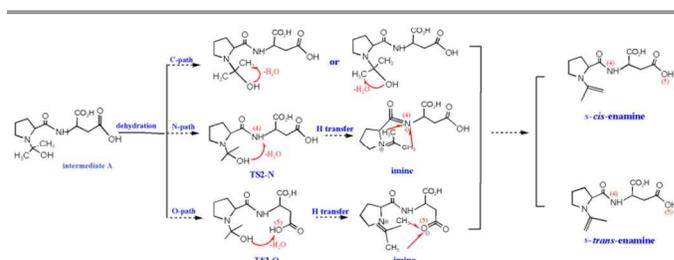

Fig.4 Formation of two enamines by three paths

The another probable path (**N-path**) was the hydroxyl firstly reacted with the H(7) atom on amino group of peptide bond to remove the H_2O molecular (transition state of dehydration in **N-path** in Fig.5 (a)) and form imine simultaneously, then hydrogen atom of either methyl transferred to N(4) atom to form *s-cis-* or *s-trans-*isomer. The corresponding first energy barrier of imine forming was 13.35 kcal/mol, and the next energy barriers were 8.26 and 19.35 kcal/mol for *s-cis-* and *s-trans-*enamines, respectively, which revealed that the *s-cis-*enamine was more energy favorable and occurred easily. The calculated conversion energy barrier of two enamines was 5.06 kcal/mol indicated the *s-cis-*enamine was achieved mostly in proton transfer process, and then the *s-trans-*enamine was obtained through configuration conversion from *s-cis-*enamine in N-path. This path was a facile process and attained easily under the experimental condition. Besides, there was a potential path (**O-path**) that hydroxyl firstly reacted with the H(8) atom on terminal carboxylic acid group of dipeptide to remove the H_2O molecular (transition state of dehydration in **O-path** in Fig.5 (b)) and form imine simultaneously, then hydrogen atom of either methyl transferred to O(5) atom to form *s-cis-* or *s-trans-*isomer. The corresponding first energy barrier of imine forming was 13.36 kcal/mol, and the next energy barriers were 1.86 and 1.69 kcal/mol for *s-cis-* and *s-trans-*enamines, respectively, which demonstrated that the two isomers were almost achieved halves and performed easily at room temperature. The energy variations in three paths to form two enamines were shown in Fig.6.

The energy barriers of proton transfer forming enamine in O-path were smaller than that of N-path, it might because proton transferred from imine to the carboxylic group was easier to overcome steric repulsion and bone torsion force. In summary, the O-path was preferential to generate half-to-half *s-cis-* and *s-trans-*enamines in second step in the perspective of energy. Therefore, the subsequent discussion of mechanism, energy variation and solvent effects was carried around the favorable O-path.
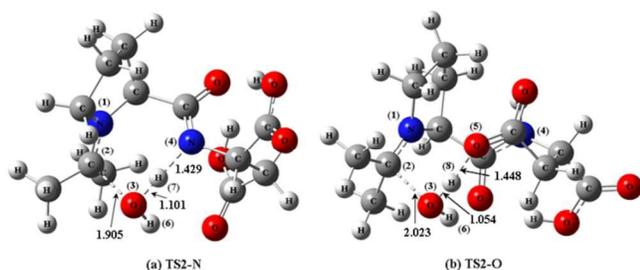
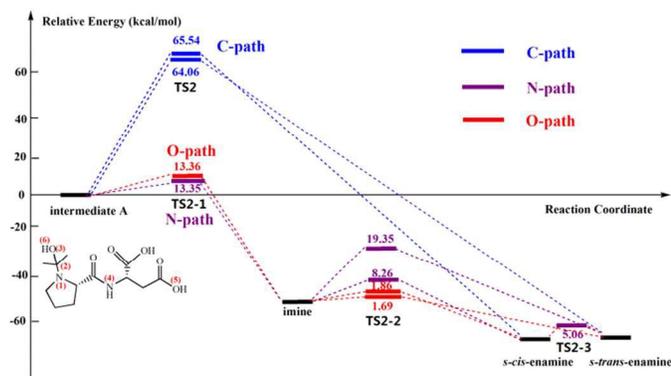Fig.5 Transition states of dehydration process in N-path and O-path

Fig.6 The energy variations in three paths to form enamines

The structures of two enamines were shown in Fig.7 (The detailed geometric parameters of enamines were listed in ESI).
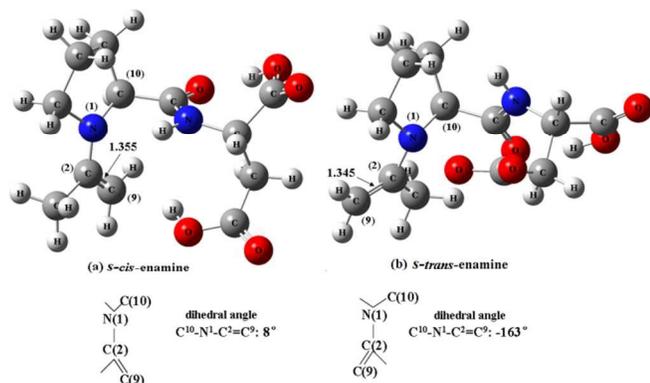

Fig.7 Structural models of two enamines

3.2.3 STEREOSELECTIVITY OF PRODUCTS. The discussion of stereoselectivity-determining step took benzaldehyde as example that the aldol reaction of acetone and benzaldehyde led to the formation of one stereoselective center, resulting in two possible stereoisomers. The enamine could be in either *s-cis-* or *s-trans*-conformation, while the enamine could attack from its different faces (*Re* or *Si*) of the benzaldehyde to form *R*- or *S*-β-hydroxyketone, giving rise to four different transition states (Fig.8). The notation used for the TS3, for example, " *t* " and " *R* " in TS3-*tR* was consistent with previous conventions, " *t* "

denoted the " *s-trans*-enamine " and " *R* " denoted the " *R*-products ".

The aldol reaction was proposed involving an enamine intermediate, and the stereo-controlling step was the formation of C-C bond between enamine intermediate and benzaldehyde. The computed energies of transition structures associated with C-C bond-formation step could reasonably explain the stereoselectivity of intermolecular aldol reaction.

The energetically most accessible transition state, TS3-*tR*, had several features that made it to have lower energy than all other ones in gas phase. The configuration of dipeptide was optimal to form a hydrogen bond (N-H···O) between dipeptide amide and carbonyl that was developing on benzaldehyde. The amide moiety was also in an optimum position to donate its proton and prevent steric repulsion from framework. In addition, the developing anion of the benzaldehyde interacted favorably with the C=C moiety of the *s-trans*-enamine. These interactions stabilized the TS3-*tR* to a very high degree, making the concerted C-C bond formation and proton transfer from amide group very synchronous. Furthermore, the aspartic acid substituent of dipeptide was pointing away from benzaldehyde and at the same time avoiding steric repulsion from dipeptide backbone.
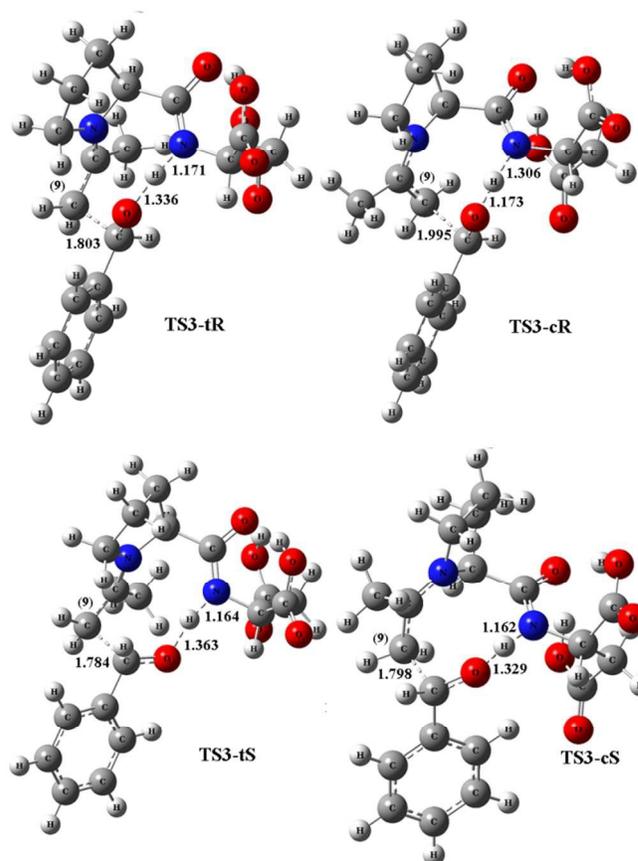

Fig.8 Transition states in third step

It should be worth noting that there were two possible stereoselective channels including the *cis*-isomer and *trans*-isomer reaction channels. When the enamine was in the *s-cis*-conformation, the transition states leading to the products were all existed the distortion of *s-cis*-enamine framework, resulting in higher energies. In these transition states, other effects that contributed to the higher energy were seen, such as steric repulsion between methyl group of *s-cis*-enamine and benzaldehyde, distortion of the aspartic acid substituent of *s-cis*-enamine. In TS3-*cR*, the favorable interactions of the aspartic acid substituent with benzaldehyde, the hydrogen bond of NH group with carbonyl were very similar to the one found for TS3-*tR*. However, to accommodate these interactions in an optimal way, the peptide backbone was rotated such that the methyl substitute of enamine experiences steric repulsion from both five-membered ring and peptide bond. In the *s-cis*-enamine version of this transition state, it was possible to avoid this steric repulsion, but this came at the price of losing the interaction to the enamine proton, which resulted in a very similar energy. These effects led to energies that were considerably higher than those transition states obtained via *s-trans*-enamine reaction channel. This caused the higher energy and was thus the source of enantioselectivity.

For the final *S*-β-hydroxyketone product, two transition states were found to have energies both lying at 6.65 and 7.58 kcal/mol higher than TS3-*tR*. The lowest transition state leading to the *S*-β-hydroxyketone enantiomer of product TS3-*tS* had an energy of 6.65 kcal/mol higher than TS3-*tR*. However, changing the face of the benzaldehyde caused a ring-ring steric repulsion that raised the energy. Finally, in the highest transition state leading to *S*-β-hydroxyketone-product TS3-*cS*, two effects were responsible for raising the energy by 7.58 kcal/mol compared to TS3-*tR*. Previous theoretical studies showed that proton transfer from the amide group of dipeptide to the forming of intermediate was essential for charge stabilization and to facilitate C-C bond formation in the transition states.[40]

The relative energies of transition states in stereoselectivity-determining step leading to the two different stereoisomers were 0, 4.02, 6.65, and 7.58 kcal/mol for TS3-*tR,* TS3-*cR,* TS3-*tS* and TS3-*cS*, respectively. Theoretically, the *R*-product was found to be the major product, with an enantiomeric excess > 99%, which translated approximately to an energy difference of 6.65 kcal/mol between the TS3-*tR* and TS3-*tS* transition states.[41] This was due to the steric repulsion interaction between the carbonyl of benzyl substituent and enamine terminus in the *S*-face attack. This also explained why the *R*-face transition state was preferred over the *S*-face for electrophilic attack.

In summary, C-C bond and hydrogen bonding interactions made TS3-*tR* the optimal transition state, having a lower energy barrier than the other transition states leading to the final product. Taken together, the features of steric repulsion interactions in other transition states discussed above jointly made the TS3-*tR* the most stable transition state. The stabilization of transition states provided by hydrogen bonding interactions was likely one of the main reasons that enable

dipeptide to catalyze intermolecular aldol reaction with high enantioselectivity.

As discussed above, the main source of stereoselectivity was found to be the steric repulsion interactions of enamines with benzaldehyde. The origin of enantioselectivity was arised from the distortion of attack face of benzaldehyde and the distortion of enamine backbone in transition state geometries. B3LYP/6-31G(d,p) calculations matched the general experimental trends and provided useful insights into the origins of the variations in stereoselectivity.

3.2.4 PRODUCTS STRUCTURE. The last step of reaction was the deprivation of catalyst-dipeptide to obtain the final chiral products in which the transition states-TS4 was formed by the intermediate B and water. At this moment, the discussion of this step was seldom because the stereoselectivity of products was determined in stereoselectivity determining step (TS3) even the energy barrier was smaller than that of rate-determining step (TS1). The structural models of final products were shown in Fig.9, the *R*-product was found to be the major product and the *ee* value calculated from the energy barriers difference between *R*-product and *S*-product was > 99%. Theoretically, the dipeptide (*S*)-pro-(*S*)-asp was excellent catalyst for the intermolecular aldol reaction.
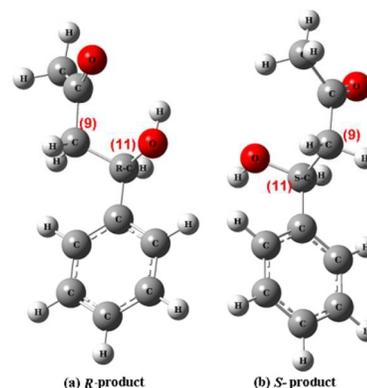
Fig.9 Structural models of final products

3.3 The energy variation in reaction path

It should be worth noting that there were two possible stereoselective channels including the form of *cis*- and *trans*-isomerization in aldol reaction. The energy variations via O-path including the *s-cis*- and *s-trans*-enamine channels were shown in Fig.10 and Fig.11, respectively. The energies of acetone and dipeptide were set as 0.00 kcal/mol as reference in the energy profile. According to the energy variation in reaction path, the energy barrier of first step was the maximum (44.33 kcal/mol), so the first step was the rate-determining step. In the second step, there were three possible channels (C-path, N-path, O-path) that O-path was the preferential channel in aldol reaction from previous analysis. The corresponding former

energy barrier of imine forming (dehydration) was 13.36 kcal/mol, and the next energy barriers of proton transfer were 1.86 and 1.69 kcal/mol for *s-cis-* and *s-trans*-enamines respectively. The energy barrier of proton transfer of imine in O-path were 1.86 (*s-cis*-enamine) and 1.69 (*s-trans*-enamine) kcal/mol (Table 1). It was too small to neglected in the course of energy variation in Fig.10 and Fig.11. Subsequently, the activation energy barrier in TS3-*tR* path was 6.65 kcal/mol lower than that in TS3-*tS* path and the energy barrier in TS3-*cR* path was 3.56 kcal/mol lower than that in TS3-*cS* path of two channels which demonstrated *R*-product was the major product. Thus, the high *ee* value (> 99%) could be explained by the larger energy difference (6.65 kcal/mol) between TS3-*tS* and TS3-*tR* transition states. Therefore, the O-path was thermodynamically more favorable, supporting the preference of reaction through the conformation of *s-trans*-enamine path leading to *R*-product.
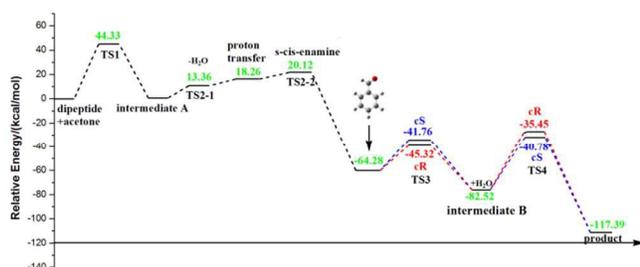

Fig.10 The energy variation via *s-cis*-enamine channel in O-path

More specifically, we focused our attention on the nucleophilic attack of the chiral enamine at the carbonyl group of benzaldehyde, since the stereochemistry of the product controlled in the step. In the equilibrium between the benzaldehyde and the enamine, an intermediate B was formed. However, to achieve this N-H···O hydrogen transfer, there was a significant loss in conjugation between the nitrogen lone pair and the C=C double bond. Mechanism involved attack of an enamine intermediate accompanied by proton transfer from amide moiety to the developing benzaldehyde.
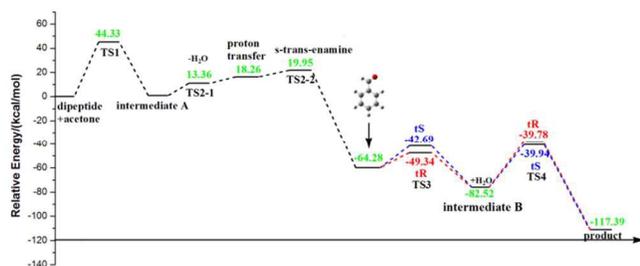

Fig.11 The energy variation via *s-trans*-enamine channel in O-path

Furthermore, two other methods (CAM-B3LYP and M06-2X) were used to calculate the energy variations in reaction path at the same basis set level of 6-31G(d,p). CAM-B3LYP combined the hybrid qualities of B3LYP and the long-range correction, while performed well for charge transfer in dipeptide model.[42] M06-2X, a hybrid meta-GGA functional, had excellent performance for main group thermochemistry and noncovalent interactions, predicted accurate valence energies. A benchmark paper by Truhlar showed the good performances of the M06-2X functional for thermodynamic and kinetic properties,[43] which agreed with their evaluation showing that M06-2X outperformed the B3LYP functional.[44-46] M06-2X was chosen for the increased accuracy with main group atom energies and reaction kinetics, as well as for its ability to describe these medium-range electron correlations and Van der Waals interactions much more effectively than B3LYP.[47] Previous studies indicated that the M06-2X functional could properly describe the thermodynamic stability within a series of transition states, hence, this was the functional chosen for this paper.[48-51]

The geometries of dipeptide, reactants, transition states, intermediates and products were calculated with M06-2X and CAM-B3LYP methods via O-path. The energy barriers in O-path calculated at CAM-B3LYP/6-31G(d,p) and M06-2X/6-31G(d,p) levels in gas phase were listed in Table 1. The M06-2X method aimed to improve the thermodynamic data for the aldol reaction using a functional that was expected to overcome known deficiencies of the B3LYP functional when considering medium-range electron correlation effects. The past researches showed that CAM-B3LYP and M06-2X calculations were in closer agreement with the available experimental data than the B3LYP calculations.[52] We considered that due to the prevalence of hydrogen in chemical systems, the good reproduction of the total atomic energies, hydrogen in particular, was more important than the slight loss in quality of the atomization and ionization energies.

As shown in Table 1, the energy variation tendency was practically analogical by using such methods. For the energy variation in reaction path by using two correction functionals, the energy barrier of first step was also the maximum (45.23 and 41.65 kcal/mol, respectively), so the first step was the rate-determining step. In second step, O-path was the preferential channel in aldol reaction from previous analysis. The energy barrier of dehydration of intermediate A calculated by CAM-B3LYP was increased, while it decreased by M06-2X method. The disparities of energy barriers to form two enamines were enlarged used by two correction functionals which proved the *s-trans*-enamine was energy favorable. Additionally, the activation energy of TS3-*tR* was lower than the other three transition states which demonstrated *R*-product was the major product calculated by three different functionals. Accordingly, the conformation of *s-trans*-enamine was thermodynamically more favorable, supporting the preference of reaction leading to *R*-product.

Table 1 The energy barriers in O-path calculated by using different methods in gas phase

	$\Delta E(B3LYP)/$ (kcal/mol)	$\Delta E(CAM-B3LYP)/$ (kcal/mol)	$\Delta E(M06-2X)/$ (kcal/mol)
TS1	44.33	45.23	41.65
TS2-1	13.36	16.09	17.20
TS2-2 (*cis-*)	1.86	3.11	4.07
TS2-2 (*trans-*)	1.69	2.53	2.60
TS3-*tR*	14.94	14.94	11.83
TS3-*cR*	18.96	18.66	15.24
TS3-*tS*	21.59	21.90	21.37
TS3-*cS*	22.52	22.83	19.86
TS4-*tR*	42.74	39.97	38.46
TS4-*cR*	42.07	39.04	36.38
TS4-*tS*	42.58	39.54	36.90
TS4-*cS*	41.74	38.72	34.66

3.4 Solvation effects

According to the mechanism, this reaction occurred under solvent-free condition at room temperature. Generally, the aldol reaction was operated in solvent, and therefore, the solvent effects considered being crucial. In order to investigate the possible solvent effects, DMSO (ϵ_{DMSO} = 46.8) and H_2O (ϵ_{H2O} = 78.4) were chosen as solvents using PCM model in our calculations.

Table 2 displayed the energy barriers in O-path calculated at B3LYP/6-31G(d,p) level in gas phase and in H_2O and DMSO using PCM model. Inclusion of the solvent DMSO and H_2O reduced the activation energy barriers for stereoselective intermediates formation significantly, and yielded a greater exothermic reaction and stable *R*-product. However, the activation energy barriers of two conformations of enamines via proton transfer process were increased. In order to illustrate this phenomenon, the proton transfer process of forming *s-trans*-enamine in gas and DMSO via O-path was presented in Fig.12 for instance.

Table 2 The energy barriers in O-path calculated at B3LYP/6-31G(d,p) level in gas phase and in water, DMSO using PCM model

	$\Delta E(gas)/(kcal/mol)$	$\Delta E(H_2O)/(kcal/mol)$	$\Delta E(DMSO)/(kcal/mol)$
TS1	44.33	37.61	40.60
TS2-1	13.36	10.77	8.15
TS2-2 (*cis-*)	1.86	5.16	7.84
TS2-2 (*trans-*)	1.69	5.93	9.19
TS3-*tR*	14.94	11.98	12.04
TS3-*cR*	18.96	15.32	15.55
TS3-*tS*	21.59	15.14	17.14
TS3-*cS*	22.52	15.43	17.06
TS4-*tR*	42.74	40.46	5.07
TS4-*cR*	42.07	40.00	11.42
TS4-*tS*	42.58	40.10	5.63
TS4-*cS*	41.74	41.46	8.86

As shown in Fig.12, proton transferred to O(5) atom forming the *s-trans*-enamine, in the meantime, the hydrogen bond between dipeptide amide N(4) and carbonyl existed. In gas phase, the hydrogen bond (N-H⋯O) distance was 1.992 Å, while it was 2.416 Å in DMSO which almost dissociated. The higher activation energy barrier was due to the hinder of hydrogen bond in solvent. The close similarity of the barriers for the dipeptide-catalyzed aldol reaction in H_2O and DMSO in contrast to the catalytic effect highlights once more the importance of explicit solute-solvent interactions.
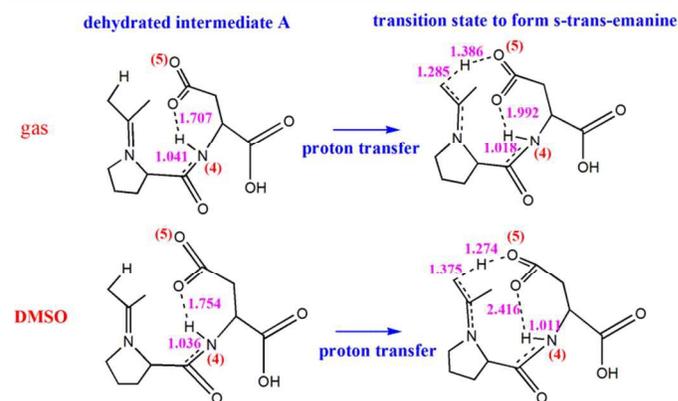
Fig.12 Proton transfer to form *s-trans*-enamine in gas and DMSO via O-path

The geometries of transition state TS3-*tR* in gas and DMSO were given in Fig.13. In DMSO, the activation energy (12.04 kcal/mol) needed for the formation of transition state TS3-*tR* was smaller than that (14.94 kcal/mol) calculated in gas phase. The effect of solvent (DMSO) was predicted to further favor *R*-product formation, and thus a stereoselective reaction, even more.
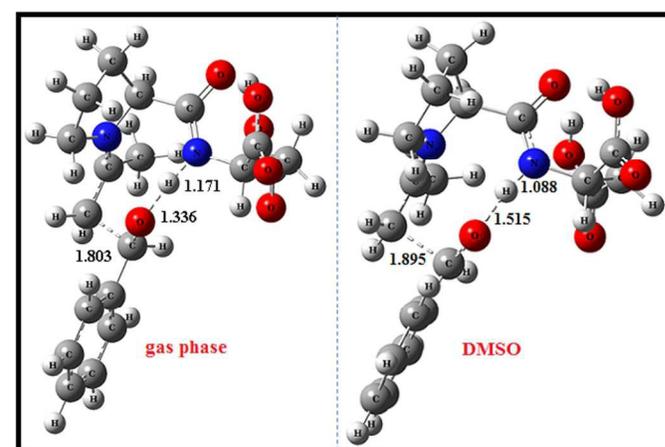
Fig.13 Structural model of TS3-*tR* in gas and DMSO

Overall, whether in the gas phase or in DMSO and H_2O, there was obviously excellent agreement between the predicted and observed stereoselectivities for the dipeptide-catalyzed intermolecular aldol reaction. The relative energies of TS4 were decreased acutely in DMSO which demonstrated DMSO was

an advantageous solvent for the dipeptide catalyzed asymmetric aldol reaction in experiment.

3.5 Substituent effects of aromatic aldehydes

As shown in Table 3, the energy variations in reaction path of aldol reaction between acetone and three aromatic aldehydes calculated at B3LYP/6-31G(d,p) level in gas. All calculations of frequency with these transition states geometries yielded one and only one imaginary frequency and all real for the local minima. The data in Table 3 showed that the change of substituent from H atom to methyl-(electron donating group) and nitro-(electron withdrawing group) only had slight influences on formation of the transition states. Similar to benzaldehyde, the lowest energy transition state was predicted to be TS3-*tR* in stereoselectivity determining step, corresponding to generate the final product, *R*-β-hydroxyketone, which was in agreement with the experimental observations. The formation of dominating *R*-product was nearly unaffected by the change of substituent of aldehydes due to the faint inductive effects of *p*-position in benzene ring and hence similar values for enantiomeric access (>99%) were observed for the aldol reactions. These results suggested that the introduction of substituent at the *p*-position of benzaldehyde effectively provided the desired stereoselectivity for asymmetric intermolecular aldol reaction.

Table 3 The energy barriers of three aromatic aldehydes calculated at B3LYP/6-31G(d,p) level in gas

	ΔE(benzaldehyde)/ (kcal/mol)	ΔE(p-methyl benzaldehyde)/ (kcal/mol)	ΔE(p-nitrobenzaldehyde)/ (kcal/mol)
TS3-*tR*	14.94	15.59	12.38
TS3-*cR*	18.96	19.42	16.66
TS3-*tS*	21.59	22.14	18.80
TS3-*cS*	22.52	23.08	19.12
TS4-*tR*	42.74	42.32	42.58
TS4-*cR*	42.07	42.16	41.59
TS4-*tS*	42.58	42.56	42.61
TS4-*cS*	41.74	41.61	42.23

It was useful to confirm the HOMO and the LUMO of transition states in stereoselectivity-determined step. The relative ordering of the occupied and virtual orbital provided a reasonable qualitative indication of the excitation properties. The HOMO and LUMO energies of these transition states in the stereoselectivity determining step were listed in Table 4. Table 4 displayed that the order of transition states for HOMO orbital energies used three aromatic aldehydes was $E_{(p\text{-methyl benzaldehyde})} > E_{(benzaldehyde)} > E_{(p\text{-nitrobenzaldehyde})}$, indicating that the introduction of methyl- in *p*-position of benzaldehyde enhanced the HOMO orbital energy, while the nitro- substituent reduced the HOMO orbital energy. The faint inductive effects in benzene ring due to the *p*-position substituent that methyl-group could enhanced the electronic density as the electron donating group, on the contrary, nitro- group could reduced the electronic density as the electron withdrawing group. The

results proved the different substituents in the *p*-position of benzene ring initiated the inductive effects; however *p*-position was relatively distant to have slight influence in the energy barriers and no influence in stereoselectivity of major product. From the comparison of energy barriers, the value of the same transition state of methyl-substituent increased and that of nitro-substituent decreased as the result of introduction of substituent in *p*-position of benzaldehyde. The electronic density contours of the frontier orbitals (HOMO and LUMO orbitals) of transition states in stereoselectivity determining step were plotted in ESI.

Table 4 HOMO and LUMO orbital energies and HOMO-LUMO gap for transition states of three aromatic aldehydes calculated at B3LYP/6-31G(d,p) level in gas

		HOMO(eV)	LUMO(eV)	Gap(eV)
benzaldehyde	TS3-*tR*	-0.229	-0.076	0.153
	TS3-*cR*	-0.235	-0.075	0.160
	TS3-*tS*	-0.228	-0.080	0.148
	TS3-*cS*	-0.223	-0.079	0.144
p-methyl benzaldehyde	TS3-*tR*	-0.225	-0.074	0.151
	TS3-*cR*	-0.231	-0.074	0.157
	TS3-*tS*	-0.225	-0.078	0.147
	TS3-*cS*	-0.219	-0.077	0.142
p-nitro benzaldehyde	TS3-*tR*	-0.241	-0.096	0.145
	TS3-*cR*	-0.246	-0.098	0.148
	TS3-*tS*	-0.238	-0.098	0.140
	TS3-*cS*	-0.233	-0.096	0.137

4. Conclusions

The mechanism of dipeptide-catalyzed intermolecular aldol reaction was investigated by using different DFT methods. There were four steps in the reaction path that the first step was the rate-determining step due to the largest energy barrier of 44.33kcal/mol, but the third step decided the stereoselectivity of the product because steric repulsions of functional groups governed the stereoselectivity of products severely. The effects of leading *R*-product through *s-trans*-isomer path were the steric repulsion between aspartic acid substitute of enamine and aldehyde; and the steric repulsion between attack face of aldehyde and enamine. It was found that the dipeptide-catalyzed aldol reaction via *s-trans*-enamine showed more energy favorable to obtain *R*-product (with an *ee* value > 99%). Additionally, the computed results provided reference for experiment that full mixing of catalyst and reactants or operated at the elevated temperatures was essential and DMSO or H_2O as the solvent could decrease the energy barriers in reaction path. We displayed that the CAM-B3LYP and M06-2X density functionals performed as well as and better than the B3LYP functional for charge transfer transition states with steric effects

but had better performance for bond energies, noncovalent interactions, and chemical reaction barrier heights for representative systems. The introduction of substituent at the *p*-position of benzaldehyde had slight influence in the energy barriers and no influence in stereoselectivity of major product for asymmetric intermolecular aldol reaction. Owing to few researches were concerned with the detailed variations in entire path, we concluded that dipeptide should be preferred for studies requiring the exploration of potential green synthesis as well as nontoxic catalysts, provided that the dipeptide of this sort even protein included.

Acknowledgements

ᵃ State Key Laboratory of Chemical Resource Engineering, Beijing University of Chemical Technology, Beijing, 100029.
Email: pumin@mail.buct.edu.cn
ᵇ Inner Mongolia Key Laboratory of Photoelectric Functional Materials, Chifeng University, Chifeng, 024000.

† Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/b000000x/

Notes and references

1   J. S. Johnson, D. A. Evans, *Acc. Chem. Res.* 2000, **31**, 325-335.

2   Z. Tang, F. Jiang, L. T. Yu, X. Cui, L. Z. Gong, A. Q. Mi, Y. Z. Jiang, Y. D. Wu, *J. Am. Chem. Soc.* 2003, **125**, 5262-5263.

3   Y. Hayashi; H. Gotoh, T. Hayashi, M. Shoji, *Angew. Chem. Int. Ed.* 2005, **44**, 4212-4215.

4   D. Enders, M. R. M. Hüttl, C. Grondal, G. Raabe, *Nature.* 2006, **441**, 861-863.

5   S. Gandhi, V. K. Singh, *J. Org. Chem.* 2008, **73**, 9411-9416.

6   X. H. Liu, L. L. Lin, X. M. Feng, *Chem. Commun.* 2009, **41**, 6145-6158.

7   C. D. Maycock, M. R. Ventura, *Tetrahedron: Asymmetry.* 2012, **23**, 1262-1271.

8   A. K. Pandey, D. Naduthambi, K. M. Thomas, N. J. Zondlo, *J. Am. Chem. Soc.* 2013, **135**, 4333-4363.

9   T. D. Machajewski, C.-H. Wong, *Angew. Chem. Int. Ed.* 2000, **39**, 1352-1375.

10   G. Luppi, P. G. Cozzi, M. Monari, B. Kaptein, Q. B. Broxterman, C. Tomasini, *J. Org. Chem.* 2005, **70**, 7418-7421.

11   H. J. Martin, B. List, *Synlett.* 2003, **12**, 1901-1902.

12   F. Kolundzic, M. N. Noshi, M. Tjandra, M. Movassaghi, S. J. Miller, *J. Am. Chem. Soc.* 2011, **133**, 9104-9111.

13   S. Bayat, B. A. Tejo, A. B. Salleh, E. Adbmalek, Y. M. Normi, M. B. A. Rahman, *Chirality.* 2013, **25**, 726-734.

14   T. Zhuo, Z. H. Yang, L. F. Cun, L. Z. Gong, A. Q. Mi, Y. Z. Jiang, *Org. Lett.* 2004, **6**, 2285-2287.

15   W. Zou, I. Ibrahem, P. Dziedzic, H. Sundén, A. Córdova, *Chem. Commun.* 2005, **39**, 4946-4948.

16   E. Machuca, E. Juaristi, *Tetrahedron. Lett.* 2015, **56**, 1144-1148.

17   P. Dziedzic, W. Zou, J. Hafren, A. Córdova, *Org. Biol. Chem.* 2006, **4**, 38-40.

18   A. Bassan, W. Zou, E. Reyes, F. Himo, A. Córdova, *Angew. Chem. Int. Ed.* 2005, **44**, 7028-7032.

19   A. Córdova, W. Zou, I. Ibrahem, E. Reyes, M. Engqvist, W. W. Liao, *Chem. Commun.* 2005, **36**, 3586-3588.

20   A. Heine, G. Desantis, J. G. Luz, M. Mitchell, C.-H. Wong, I. A. Wilson, *Science*, 2001, **294**, 369-374.

21   T. D. Machajewski, C. -H. Wong, *Angew. Chem.* 2000, **112**, 1406-1430.

22   Z. G. Hajos, D. R. Parrish, *J. Org. Chem.* 1974, **39**, 1615-1621.

23   B. List, R. A. Lerner, C. F. III. Barbas, *J. Am. Chem. Soc.* 2000, **122**, 2395-2396.

24   W. Notz, B. List, *J. Am. Chem. Soc.* 2000, **122**, 7386-7387.

25   M. Amedjkouh, *Tetrahedron: Asymmetry.* 2005, **16**, 1411-1414.

26   M. Limbach, *Tetrahedron Lett.* 2006, **47**, 3843-3847.

27   S. Bahmanyar, K. N. Houk, *J. Am. Chem. Soc.* 2001, **123**, 12911-12912.

28   L. Hoang, S. Bahmanyar, K. N. Houk, B. List, *J. Am. Chem. Soc.* 2003, **125**, 16-17.

29   B. List, P. Pojarliev, W. T. Biller, H. J. Martin, *J. Am. Chem. Soc.* 2002, **124**, 827-833.

30   C. Allemann, J. M. Um, K. N. Houk, *J. Mol. Catal. A: Chem.* 2010, **324**, 31-38.

31   F. R. Clemente, K. N. Houk, *J. Am. Chem. Soc.* 2005, **127**, 11294-11302.

32   Y. K. Zhang, J. Zhu, N. Yu, H. Yu, *Chin. J. Chem.* 2015, **33**, 169-174.

33   K. Liu, G. Zhang, *Tetrahedron. Lett.* 2015, **56**, 243-246.

34   S. S. V. Ramasastry, H. Zhang, F. Tanaka, C. F. III. Barbas, *J. Am. Chem. Soc.* 2007, **129**, 288-289.

35   S. Pizzarello, A. L. Weber, *Science*, 2004, **303**, 1151-1151.

36   Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gau-ssian 09 Revision A.1 Gaussian, Inc. Wallingford, CT

37   A. Fu, H. Li, F. Tian, S. Yuan, H. Si, Y. Duan, *Tetrahedron: Asymmetry.* 2008, **19**, 1288-1296.

38   B. List, L. Hoang, H. J. Martin, *Proc. Natl. Acad. Sci. U.S.A.* 2004, **101**, 5839-5842.

39   G. Yang, Z. Yang, L. Zhou, R. Zhu, C. Liu, *J. Mol. Catal. A: Chem.* 2010, **316**, 112-117.

40   P. Hammar, A. Córdova, F. Himo, *Tetrahedron: Asymmetry.* 2008, **19**, 1617-1621.

41 O. B. Gadzhiev, L. A. G. D. L. Rosa, F. J. M. Bustamante, C. A. D. Parrodi, H. H. Abdallah, A. I. Petrov, T. Scior, *J. Phys. Org. Chem.* 2012, **25**, 971–978.

42 T. Yanai, D. P. Tew, N. C. Handy. *Chem. Phys. Lett.* 2004, **393**, 51-57.

43 Y. Zhao, D. G. Truhlar, *Acc. Chem. Res.* 2008, **41**, 157-167.

44 A. D. Becke, *J. Chem. Phys.* 1993, **98**, 1372-1377.

45 C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B: Condens. Matter.* 1988, **37**, 785-789.

46 A. D. Bose, J. M. L. Martin, *J. Chem. Phys.* 2004, **121**, 3405-3416.

47 S. Rayne, K. Forest, *Theochem.* 2010, **948**, 102–107.

48 D. H. Ess, S. Liu, F. D. Proft, *J. Phys. Chem. A.* 2010, **114**, 12952-12957.

49 Y. Zhao, D.G. Truhlar, *Org. Lett.* 2006, **8**, 5753–5755.

50 Y. Zhao, D.G. Truhlar, *Theory Chem. Acc.* 2008, **120**, 215–241.

51 R. Huenerbein, B. Schirmer, J. Moellmann, S. Grimme, *Phys. Chem. Chem. Phys.* 2010, **12**, 6940–6948.

52 R. Li, J. Zheng, D. G. Truhlar, *Phys. Chem. Chem. Phys.* 2010, **12**, 12697-12701.

Graphical Abstract

The mechanism of dipeptide-catalyzed intermolecular aldol reaction was investigated by using different calculated methods of DFT that few researches were concerned with the detailed variations in entire path.