Computational mechanistic investigation of the kinetic resolution of α-methyl-phenylacetaldehyde by norcoclaurine synthase

Check for updates

Shiqing Zhang[1,2], Chenghua Zhang[1,3], Aijing Guo[1,4], Baoyan Liu ®[1,2], Hao Su[1,2,4] ✉ & Xiang Sheng ®[1,2,4] ✉

Norcoclaurine synthase from *Thalictrum flavum* (*Tf*NCS) demonstrated high stereospecificity and yield in catalyzing the Pictet-Spengler reaction of dopamine with chiral aldehydes, achieving kinetic resolution of aldehydes. However, the mechanism and the factors contributing to the stereoselectivity remain unclear. Herein, by using quantum chemical calculations, the mechanisms of *Tf*NCS-catalyzed reactions of dopamine with both enantiomers of α-methyl-phenylacetaldehyde are studied. The calculations reveal a mechanism mirroring the reaction of natural substrates, for which the deprotonation of the C5−H of the cyclized intermediate is rate-limiting. The calculated overall barriers are 20.1 kcal mol^{-1} and 21.6 kcal mol^{-1} for the reactions of (*R*)- and (*S*)-α-methyl-phenylacetaldehyde, respectively. The M97 and L72 residues are proposed to be the key residues contributing to the stereospecificity. The obtained detailed information is helpful for designing new variants of *Tf*NCS with extended substrate scope, and also advancing our understanding of *Tf*NCS reactions for potential applications.

The benzylisoquinoline alkaloids (BIAs) are groups of natural products found in plants[1]. Many BIAs have been extensively studied for their pharmacological activities, with examples including berberine, glaucine, morphine, and noscapine, which display antimicrobial, anti-inflammatory, and analgesic properties, respectively[2–5]. Furthermore, compounds like (*R*)-coclaurine and (*S*)-norcoclaurine have demonstrated potent anti-HIV activities[6]. The core structure of BIAs is commonly synthesized by a well-established organic chemistry reaction, namely the Pictet-Spengler (PS) reaction, which involves the condensation of an amine with an aldehyde or ketone. The corresponding enzymes are named as Pictet-Spenglerases (PSases)[7].

Norcoclaurine synthases (NCSs) are a subset of PSases that catalyze the PS condensation between dopamine and 4-hydroxyphenylacetaldehyde (4-HPAA), leading to the formation of norcoclaurine, which contains a tetrahydroisoquinoline (THIQ) skeleton. This condensation reaction is the key step in the biosynthesis of BIAs[8–11], making NCSs have received great attention in the designed biosynthetic route of novel alkaloids[12–17]. The NCS

from *Thalictrum flavum* (*Tf*NCS) stereospecifically catalyzes the PS reaction between dopamine and 4-HPAA to generate (*S*)-norcoclaurine (Fig. 1a)[8,18,19]. This enzyme exhibits a broad substrate spectrum for aldehydes. For instance, it has high activity in the conversion of phenylacetaldehydes substituted with diverse groups[20,21]. Recently, the enzyme has been shown to accept also various ketone compounds, including phenylacetone, cyclohexanone, and some of their derivatives[22]. For the amine substrate, it has been demonstrated that in addition to the natural dopamine substrate, O-hydroxyphenylethylamine and 1-methoxy-2-hydroxy-phenylethylamine can also be recognized[23]. The broad substrate spectrum makes *Tf*NCS to be a promising biocatalyst in producing non-natural alkaloids.

The PS condensation between amine and achiral aldehydes produces compounds bearing one chiral center. While chiral aldehyde substrates are used, products possessing two chiral centers can be generated. A recent study showed that *Tf*NCS accepts chiral α-substituted aldehydes to produce THIQs with two chiral centers with high conversions[24]. As *Tf*NCS displays

[1]Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin 300308, PR China. [2]National Center of Technology Innovation for Synthetic Biology, National Engineering Research Center of Industrial Enzymes and Key Laboratory of Engineering Biology for Low-Carbon Manufacturing, Tianjin 300308, PR China. [3]School of Pharmacy, North Sichuan Medical College, Nanchong 637100, PR China. [4]University of Chinese Academy of Sciences, Beijing 100049, PR China. ✉e-mail: suhao@tib.cas.cn; shengx@tib.cas.cn

(S)-enantioselectivity at C1 position, only the (1S, 1'R)- and (1S, 1'S)-diastereoisomers can be formed when the aldehyde substrates are (R)- and (S)-enantiomers, respectively. Interestingly, the enzyme converts the (R)-enantiomer more preferentially to (S)-enantiomer, thus leading to the kinetic resolution of the chiral aldehyde racemate[24]. Taking α-methyl-phenylacetaldehyde (MPAA) as an example (Fig. 1b), (1S, 1'R)-diastereoisomer was experimentally measured to be the major product with diastereoisomeric ratio (d. r.) of 94:6[24]. To further extend the application of TfNCS in the kinetic resolution of chiral aldehydes, it is important to understand the factors controlling the stereoselectivity. To this end, the reaction mechanism should be first revealed.

Many efforts have been devoted to reveal the reaction mechanism of the natural reaction of TfNCS. On the basis of the available crystal structures, the two substrates can be proposed to bind to the active site in different sequence, giving rise to two potential binding modes: the "HPAA-first"[25,26] and "dopamine-first" binding modes[27,28]. Recent experimental and computational studies proved that the two binding modes are accessible at the stage of the enzymes-substrates complex, but only the "dopamine-first" mechanism is catalytically relevant[24,29]. The reaction mechanism of TfNCS was demonstrated to follow the general mechanism of PS reaction, involving the sequential C–N bond formation, dehydration, cyclization and the final deprotonation of C–H group of the cyclized intermediate. Nevertheless, the detailed reaction mechanisms and the factors controlling the stereospecificity toward the enantiomers of the chiral α-methyl-phenylacetaldehyde remain elusive.

In this work, the quantum chemical calculations using an active site cluster model of the enzyme (Supplementary Fig. 1) were employed to explore the reaction mechanisms of TfNCS-catalyzed condensation between dopamine and the (R)- and (S)-enantiomers of α-methyl-phenylacetaldehyde and to reveal the factors governing the stereospecificity. Detailed mechanisms are proposed, and the potential sources contributing to the kinetic resolution of α-methyl-phenylacetaldehyde are discussed.

Results and discussion
Reaction mechanism

As discussed above, previous study suggested that the "dopamine-first" binding mode is the catalytically active mode for the TfNCS reaction[29]. Consequently, in the present work, we exclusively considered this binding mode. For both the R- and S-models, we optimized the geometries of a number of enzyme–substrate complexes with different conformations of the substrates and active site residues, and evaluated the corresponding energies. The structures of the lowest-energy R- and S-models, denoted as E:S$_R$ and E:S$_S$, respectively, are presented in Fig. 2, and other structures are shown in Supplementary Figs. 2 and 3. The calculations reveal that E:S$_R$ and E:S$_S$ have very similar energies, with E:S$_S$ being only 0.1 kcal mol^{-1} lower than that of E:S$_R$. In both R- and S-models, the D141 is in the protonated state, while the E110 is deprotonated.

The reaction mechanisms of the TfNCS-catalyzed condensation of dopamine with the enantiomers of MPAA were then investigated. The calculations revealed that the reaction for R- and S-models follow the same mechanism (Fig. 3). Given (R)-enantiomer is preferential by the enzyme, only the pathway for the PS reaction between dopamine and (R)-MPAA (R-pathway) is here discussed in detail and compared to the reaction of the natural substrates. The entire catalytic cycle can be divided into four key processes: the nucleophilic attack of dopamine at the aldehyde forming the C–N bond, the dehydration step resulting in the generation of the iminium intermediate, cyclization of the iminium intermediate, and finally, the deprotonation of the cyclized intermediate leading to the generation of the THIQ product. The calculated energy profiles for the catalytic cycle are shown in Fig. 4, in which the energy of E:S$_S$ is set as the zero of the energy profiles, and the optimized structures of corresponding species are shown in

Fig. 1 | Enzymatic Pictet–Spengler reaction catalyzed by the norclaurine synthase from *Thalictrum flavum* (*Tf*NCS). The reaction schemes between the natural amine substrate dopamine and (a) the natural aldehyde substrate 4-hydroxyphenylacetaldehyde (4-HPAA) or (b) the non-natural aldehyde substrate α-methyl-phenylacetaldehyde (MPAA).
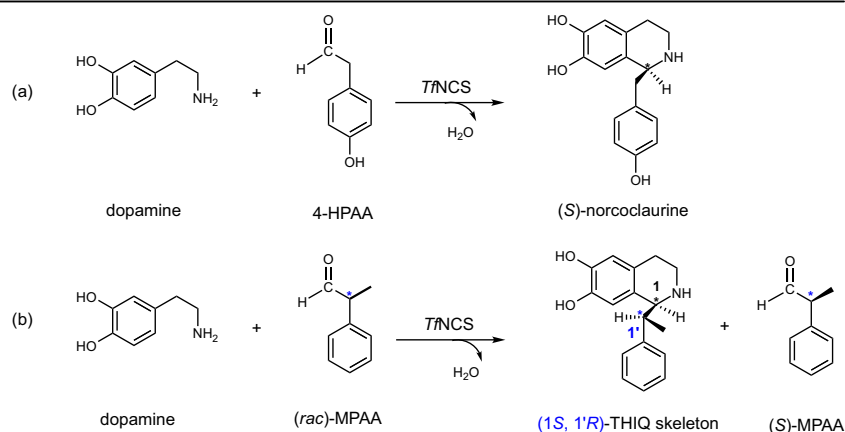
Fig. 2 | The optimized structures of the lowest-energy enzyme-substrates complexes: *R*-model (E:S$_R$) and *S*-model (E:S$_S$). The substrates are shown in green and orange stick in E:S$_R$ and E:S$_S$, respectively. The energies relative to E:S$_S$ are given in kcal mol^{-1} and selected distances are given in Å. For clarity, most of the hydrogen atoms are omitted and the constraints made on the model during geometry optimization are indicated in Supplementary Fig. 1.
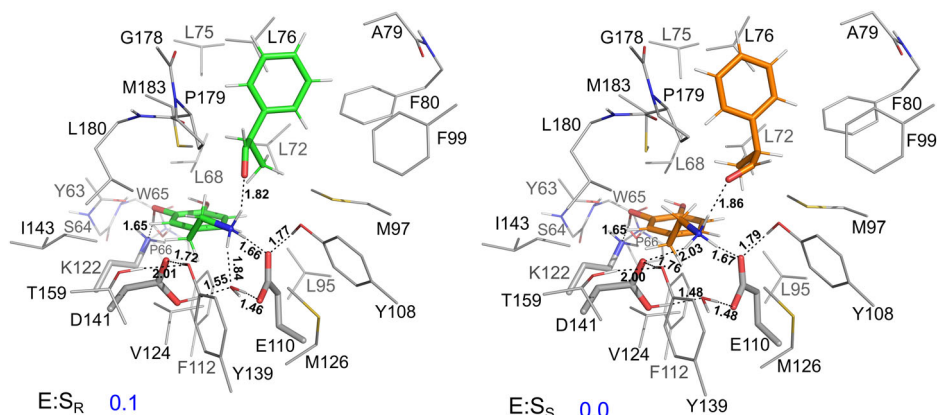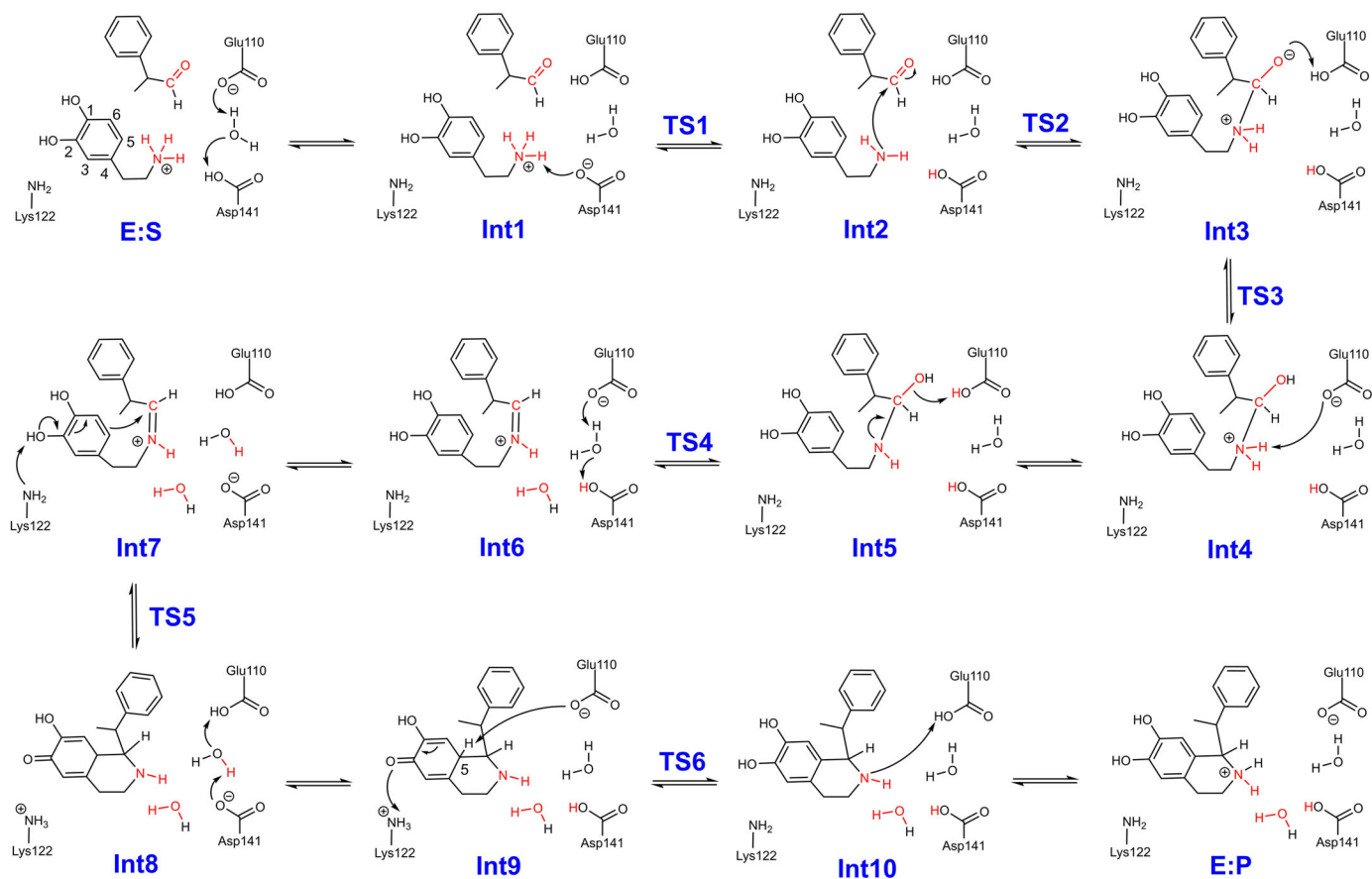
Fig. 3 | Detailed reaction mechanism suggested on the basis of the current calculations. The reaction between dopamine and α-methyl-phenylacetaldehyde is proposed to follow the "dopamine-first" mechanism, involving the sequential C–N bond formation, dehydration, cyclization and the final deprotonation of C–H group of the cyclized intermediate. A number of proton transfer events occur between the substrates and active site residues throughout the reaction process.

Fig. 4 | Calculated energy profiles for the *Tf*NCS-catalyzed PS reaction between dopamine and α-methyl-phenylacetaldehyde. Reaction pathways with (*R*)- and (*S*)-MPAA, named as *R*-pathway and *S*-pathway, are shown in black and red, respectively. The energy of E:S$_S$ is set as the zero of the energy profiles.
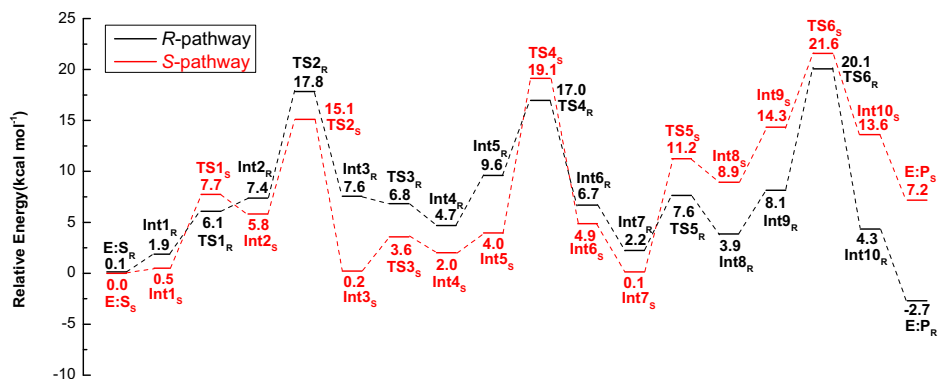

Fig. 5 and Supplementary Figs. 4 and 5. The cartesian coordinates of the optimized structures of all the species involved in the lowest-energy pathways and their corresponding energies are shown in Supplementary Data 1.

As illustrated in Fig. 3, the reaction starts with the deprotonation of D141 residue by E110 via a water bridge (E:S$_R$ → Int1$_R$), raising the energy by 1.9 kcal mol^{-1}. The deprotonated D141 then extracts a proton from the amino group of the dopamine (Int1$_R$ → [TS1$_R$] → Int2$_R$) with a barrier of 6.1 kcal mol^{-1} relative to E:S$_S$, which is very similar to the corresponding step in the natural reaction (5.0 kcal mol^{-1})[29]. The subsequent nucleophilic attack of dopamine at the carbonyl group of (*R*)-MPAA leads to the formation of a tetrahedral intermediate (Int2$_R$ → [TS2$_R$] → Int3$_R$). This step was calculated to have an energy barrier of 17.8 kcal mol^{-1} relative to E:S$_S$ (Fig. 4), significantly higher than the equivalent step in the reaction of the

natural substrates (8.0 kcal mol^{-1} from the previous study[29]). Structural comparison of the transition states (TSs) showed that the higher barrier for C–N bond formation here could be attributed to the steric hindrance between the α-methyl group of the (*R*)-MPAA substrate and P179, which is not present in the reaction of the natural substrates.

Then, two successive low-barrier proton transfer processes take place, from E110 to the alkoxide (Int3$_R$ → [TS3$_R$] → Int4$_R$) and from ammonium ion to the E110 (Int4$_R$ → Int5$_R$). Owing to the presence of the steric effect between the P179 residue and the moiety of (*R*)-MPAA, the energies of the intermediates Int4$_R$ and Int5$_R$ are higher than E:S$_S$. This contrasts with the reaction of natural substrates, where the energies of the two corresponding intermediates are lower than their enzyme-substrate complexes. The following dehydration step (Int5$_R$ → [TS4$_R$] → Int6$_R$) was calculated to be
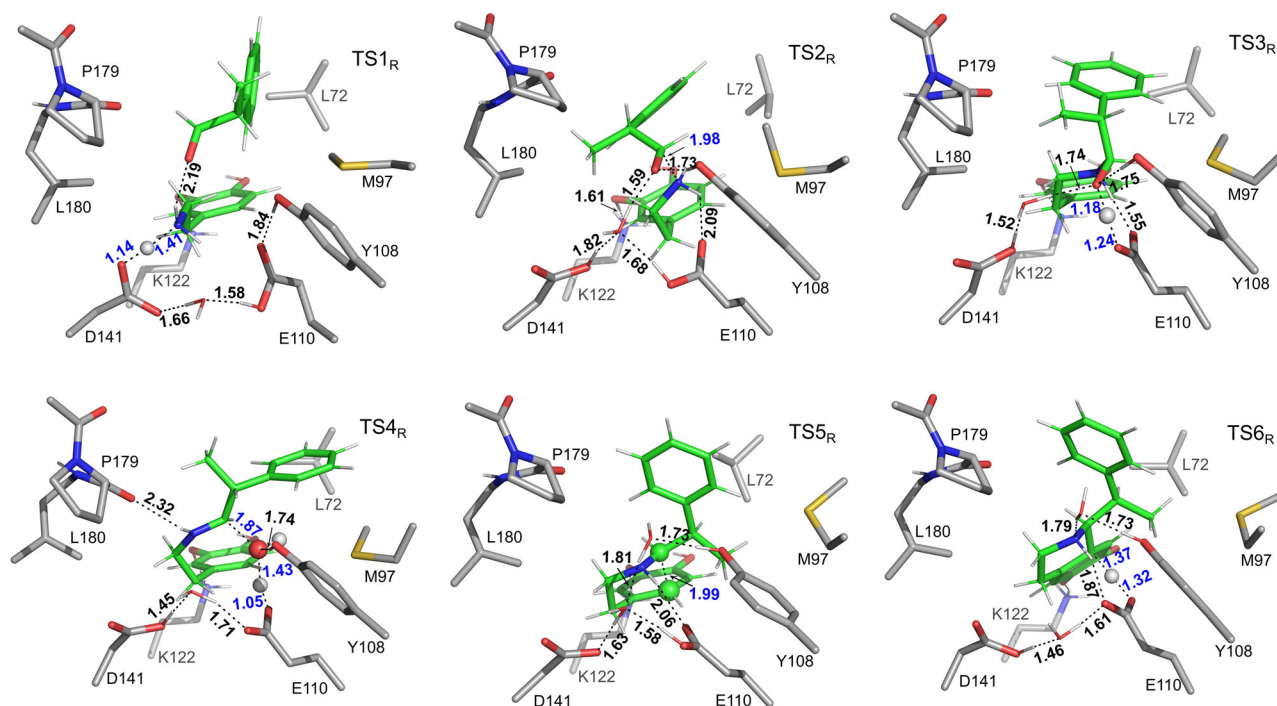
Fig. 5 | The optimized structures of the transition states involved in the *Tf*NCS-catalyzed PS reaction between dopamine and (*R*)-MPAA. Substrates are depicted as green sticks. Selected distances are given in Å. For clarity, only a small part of the model is shown. Structures with the full model are given in Supplementary Fig. 4.

17.0 kcal mol^{-1} relative to E:S$_S$, and the resulting iminium intermediate (Int6$_R$) is 6.7 kcal mol^{-1} higher than E:S$_S$. The overlap of Int6$_R$ with the co-crystallized structure of *Tf*NCS with an (*R*)-iminium mimic (PDB code: 6RP3) displays very high similarity (Supplementary Fig. 6), which indicates that our model could give a good representation to the real enzymatic process.

Following the formation of Int6$_R$, a proton transfer from D141 to E110 takes place, reducing the energy of the system by 4.5 kcal mol^{-1} (Int6$_R$ → Int7$_R$). Next, proton abstraction by K122 initiates the cyclization of the iminium intermediate (TS5$_R$) with an energy barrier of 5.4 kcal mol^{-1} relative to Int7$_R$, and the calculated energy of the cyclized intermediate Int8$_R$ is 1.7 kcal mol^{-1} higher than Int7$_R$. To form the THIQ skeleton, the C5–H must be deprotonated by a general base group. The most plausible residue is E110 in the vicinity. However, in Int7$_R$, this residue is in the protonated form. Therefore, a proton transfer from E110 to D141 takes place (Int8$_R$ → Int9$_R$) prior to the deprotonation of C5–H group. This proton transfer was calculated to be endothermic by 4.2 kcal mol^{-1}. After the proton transfer, the deprotonation of the C5–H of the cyclized intermediate by E110 occurs with a calculated barrier of 20.1 kcal mol^{-1} relative to E:S$_S$ (Int9$_R$ → [TS6$_R$] → Int10$_R$). Finally, the proton transfer from E110 to the amine group of the newly-formed six-membered ring generates the final product and restores the enzyme for the next catalytic cycle (Int10$_R$ → E:P$_R$).

The obtained mechanism and the calculated energy profile for the *Tf*NCS-catalyzed condensation of dopamine and (*R*)-MPAA are generally consistent with the reaction of the natural substrates in the previous study but obvious differences are observed[29]. In particular, the deprotonation of the C5–H group of the cyclized intermediate is rate-limiting in both cases. However, the barrier for the overall pathway of the reaction of (*R*)-MPAA (20.1 kcal mol^{-1}) is 1.3 kcal mol^{-1} higher than that of the natural reaction. This is a result of the steric hindrance between the methyl substituent of the substrate and the M97 residue in the reaction of (*R*)-MPAA. In the reaction of 4-HPAA, this effect does not exist, allowing the substrate to adjust the conformation more freely in the active site.

The kinetic data is not available for the considered reaction here but accessible for the reactions of structurally similar aldehyde substrates[30]. The measured k_{cat} values for two other non-natural substrates, benzaldehyde and 4-biphenylaldehyde, are 0.237 s^{-1} and 0.001 s^{-1}, respectively, which can be converted to barriers of ~19 kcal mol^{-1} and 23 kcal mol^{-1}. The calculated barriers for the reactions of (*R*)-MPAA (20.1 kcal mol^{-1}) and (*S*)-MPAA (21.6 kcal mol^{-1}, see next section) are thus reasonable, as they are in good agreement with the experimentally derived barriers of the two comparable substrates.

Enantioselectivity

The *Tf*NCS-catalyzed condensation of dopamine and MPAA racemate exhibits high enantiospecificity in favor of the conversion of (*R*)-enantiomer, leading to the formation of (1*S*, 1'*R*)-product as the major product. To decipher the origins of the enantiospecificity, the pathway for the reaction of (*S*)-enantiomer, leading to the formation of (1*S*, 1'*S*)-product, was also studied (*S*-pathway). The calculated profile for the lowest-energy *S*-pathway is shown in Fig. 4, and the optimized structures of corresponding species are shown in Fig. 6 and Supplementary Figs. 7 and 8.

As shown in Fig. 4, in *S*-pathway, the proton transfer from the amino group of dopamine to the D141 is also easy with an energy barrier of 7.7 kcal mol^{-1} (TS1$_S$ relative to E:S$_S$). For the formation of C–N bond, the calculated energy barrier in the *S*-pathway (TS2$_S$) is 2.7 kcal mol^{-1} lower than that of TS2$_R$. By scrutinizing the optimized structures of TS2$_R$ and TS2$_S$ (Figs. 5 and 6), it was found that this difference can be ascribed to the fact that the P179 residue introduces steric interaction with the methyl group of the (*R*)-MPAA in TS2$_R$ but not in TS2$_S$. Due to this steric effect, the calculated energies of the species from Int3 to Int5 in the *R*-pathway (Int3$_R$ to Int5$_R$) are consistently higher than those in the *S*-pathway (Int3$_S$ to Int5$_S$).

The following dehydration step of *S*-pathway has a barrier of 19.1 kcal mol^{-1} (TS4$_S$ relative to E:S$_S$, Fig. 4), which 2.1 kcal mol^{-1} higher than the corresponding TS of *R*-pathway (TS4$_R$). By analyzing the optimized structures, it can be identified that in TS4$_R$ the –OH moiety of the substrate adjusts its hydrogen bonding interactions with the surrounding environment to accommodate the dehydration process and to form the *E*-configuration iminium (Supplementary Fig. 4). Additionally, the methyl group of the (*R*)-MPAA rotates away from P179 thus eliminating the previously mentioned steric interaction in TS4$_R$. This trend remains in the
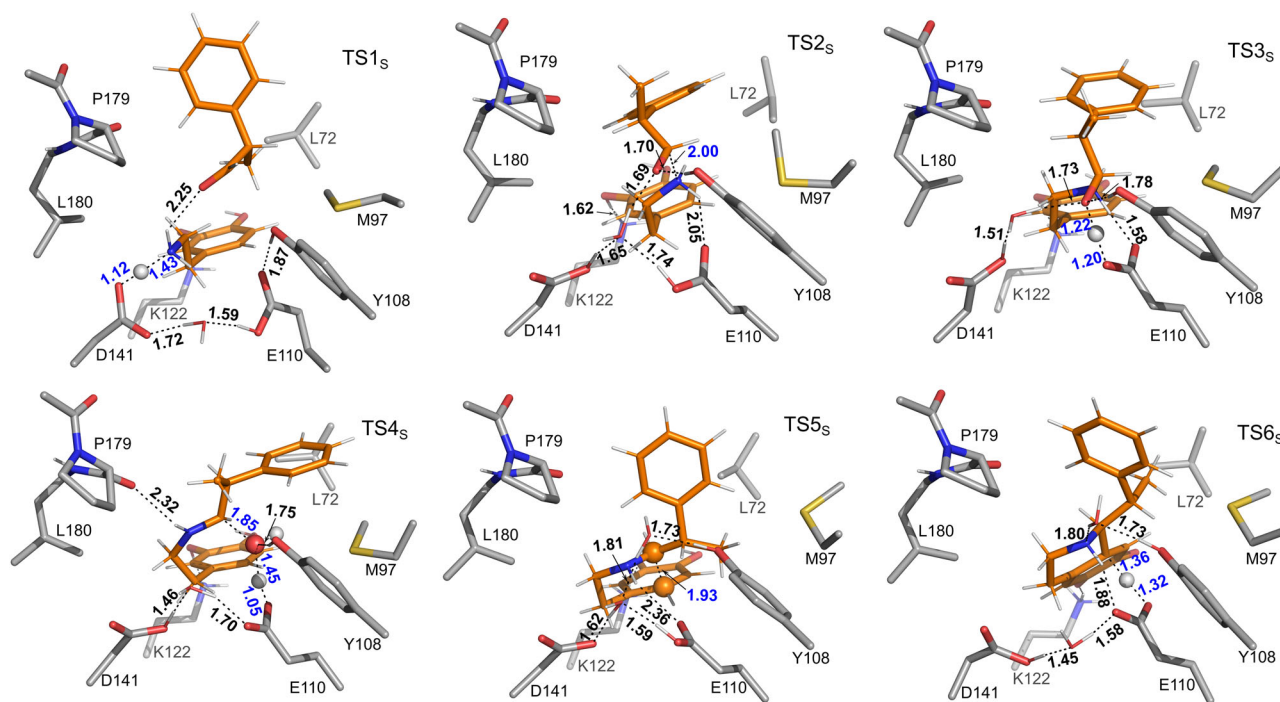
Fig. 6 | The optimized structures of the transition states involved in the _Tf_NCS-catalyzed PS reaction between dopamine and (_S_)-MPAA. Substrates are depicted as orange sticks. Selected distances are given in Å. For clarity, only a small part of the model is shown. Structures with the full model are given in Supplementary Fig. 7.

following cyclization and deprotonation steps. The same as the _R_-pathway, the deprotonation of C5–H of the cyclized intermediate (TS6$_S$) is calculated to be rate-limiting of the overall reaction for the _S_-pathway. The calculated barrier is 21.6 kcal mol^{-1} relative to E:S$_S$, which is 1.5 kcal mol^{-1} higher than that of the _R_-pathway. According to the calculated energy profiles for the two pathways, this step is also the selectivity-determining step. The calculated difference between the barriers of _S_- and _R_-pathways is in good agreement with the experimental reports, 1.6 kcal mol^{-1} converted from the d. r. with a value of 94:6[24].

By scrutinizing the geometries of the transition states of the rate-limiting steps of the two different pathways, the reasons for the obtained energy difference can be identified and the kinetic resolution of MPAA racemate can be rationalized. As shown in Supplementary Fig. 9, the methyl substituent of the substrate engages in steric interaction with M97 and L72 in TS6$_R$ and TS6$_S$, respectively, thus raising the corresponding energy barrier for the C5–H deprotonation step compared to the corresponding step in the natural reaction. Obviously, the hindrance with L72 in TS6$_S$ is stronger than that with M97 in TS6$_R$, as _R_-pathway corresponds to a lower barrier by 1.5 kcal mol^{-1}. This trend is more obvious in the following steps, which are 9.3 and 9.9 kcal mol^{-1} in favor of the Int10$_R$ and E:P$_R$ compared to Int10$_S$ and E:P$_S$, respectively. This is in fact consistent with the experimental mutagenesis analysis, for which the replacement of M97 by a smaller valine residue resulted in higher (_R_)-enantiospecificity[24].

Conclusions
In the present study, the detailed mechanisms of the _Tf_NCS-catalyzed PS reactions of dopamine with both enantiomers of α-methyl-phenylacetaldehyde (MPAA) and the factors contributing to the kinetic resolution of MPAA racemate are studied by using quantum chemical calculations. The calculations reveal that the reaction mechanisms are consistent with that of the natural substrates proposed in the previous study. The deprotonation of C5–H of the cyclized intermediates is the rate-limiting step for the reaction of both enantiomers of MPAA, with calculated barriers are 20.1 and 21.6 kcal mol^{-1} for (_R_)- and (_S_)-MPAA, respectively.

The current calculations perfectly reproduce the experimentally measured diastereoisomeric ratio (d. r.) value of the _Tf_NCS-catalyzed

kinetic resolution of MPAA racemate (1.5 kcal mol^{-1} from the calculations vs. 1.6 kcal mol^{-1} converted from the experimental data). By scrutinizing the optimized geometries of selectivity-determining transition states, it is revealed that the steric effects between the α-methyl group of the aldehyde substrate and the L72 and M97 residues are the key factors contributing to the kinetic resolution of MPAA racemate.

We believe that the detailed information about the mechanisms and the factors contributing to the enantiospecificity of MPAA racemate will be helpful to rationally design variants of _Tf_NCS with improved reactivity and selectivity for a wider scope of substrates to generate new tetra-hydroisoquinoline skeletons. Indeed, in a recent Quantum Mechanics/Molecular Mechanics (QM/MM) study on the NCS-catalyzed reaction of the natural substrates, mutants with enhanced rate were identified through the rational transition-state macrodipole stabilization method[31]. The insights provided by the current study can also be extended to other NCS enzymes and other members of the Pictet–Spenglerase family.

From the technical point of view, the present study demonstrates again that the quantum chemical cluster approach is powerful in examining enzyme properties with small energy differences. The high accuracy in energy calculations using this methodology can be attributed, in part, to the effective cancellation of systematic errors between the transition states of very high similarity.

Method
Technical details
The quantum chemical cluster approach was chosen for investigating the reaction mechanism and exploring the factors controlling the stereo-specificity of _Tf_NCS toward the non-natural aldehyde substrate α-methyl-phenylacetaldehyde. This method has been proven to be powerful in studying the mechanism of enzymatic reactions[11,32–35], and elucidating the origins of selectivities[29,36–38]. All calculations in this work were conducted using Gaussian16 C.01 package with the B3LYP-D3(BJ) density functional method[39–43]. Geometry optimizations were performed with 6-31G (d,p) basis set for all atoms. Based on the optimized structures, at the same level of theory of geometry optimization, the single-point solvation energies using solvation model based on density (SMD) with $\varepsilon = 4.0$[44] and frequency

calculations were performed to estimate the effects of enzyme environment and to obtain zero-point energies (ZPE), respectively. The transition states were validated by the imaginary normal modes of vibrations. For more accurate electronic energies, the single-point energy calculations were performed with the larger basis set 6-311 + G(2d,2p). The presented energies in this work are the electronic energies at B3LYP-D3(BJ)/6-311 + G(2d,2p) level corrected for the solvation effect and the zero-point energy correction. The absolute energies and energy corrections of the species involved in the reaction pathways are summarized in Supplementary Table 1.

Active site model
Based on the recently solved crystal structure of *Tf*NCS in complex with an intermediate analog (PDB code: 5NON)[45], the cluster model of the active site was designed to investigate the reaction mechanism and the stereospecificity of *Tf*NCS toward the (*R/S*)-α-methyl-phenylacetaldehyde. We constructed two models for the enzyme-substrates complexes, which consist of dopamine and (*R*)-α-methyl-phenylacetaldehyde (called *R*-model) or (*S*)-α-methyl-phenylacetaldehyde (called *S*-model). The intermediate analog in the crystal structure was manually replaced by the corresponding substrate. These models include residues proposed to be general acid/base groups in the reaction (E110, K122, and D141), key residues controlling selectivity based on experimental findings (L76, A79, F80, M97, Y108), and other residues constituting the active site (Y63, S64, W65, P66, L68, L72, L75, L95, F99, F112, V124, M126, Y139, I143, T159, G178, P179, L180, M183), along with two crystallographic waters. Amino acids were truncated, and hydrogen atoms were added to maintain saturation of the truncated carbon atoms. To avoid unrealistic movements, the truncated carbons were held fixed in their crystallographic positions during the geometry optimizations (see Supplementary Fig. 1 for the fixed atoms). The coordinate locking procedure gave rise to several small imaginary frequencies, typically within $50i$ cm^{-1}. These frequencies do not contribute significantly to the ZPE and can thus be safely disregarded. The final active sites of both models consist of 396 atoms and the overall charge is 0.

The protonation states of the ionizable residues within the active site model are determined based on the previous experimental and computational studies on the NCS-catalyzed reactions[28,29]. Specifically, the established role of K122 in deprotonating the substrate's phenolic hydroxyl group emphasizes the necessity for this residue to be in its neutral form[28,29]. K122 is thus modeled in its neutral form in the present study. The previous theoretical study showed that the pathway in which both E110 and D141 are ionized has very high energies[29]. Therefore, one of them is neutral and the other one is ionized in the reaction. The current calculation results on the non-natural substrates showed that the enzyme-substrates complex with D141 in the protonated state while E110 in the deprotonated state has the lowest energy (Fig. 2). Furthermore, as reported in the previous study[28], the predicted pK_a values of K122 is much lower than the typical value, while the pK_a values of D141 and E110 are rather higher than usual, indicating that the pK_a of these residues are greatly influenced by the active site microenvironment and the energetic penalties associated with switching the protonation states of these residues are relatively modest.

Data availability
The authors declare that the coordinates of the optimized structures of all the species involved in the lowest-energy pathways and their corresponding energies are available within the paper and its Supplementary Data 1. The coordinates and energies of other considered structures are available from the corresponding authors upon reasonable request. The following calculation results are provided in the Supplementary Materials: optimized structures for the enzyme-substrates complexes; optimized structures of the intermediates and transition states; superposition of (*R*)-iminium ion intermediate (Int6$_R$) with the co-crystallized structure; schematic illustration of the selectivity-determining transition states; absolute energies and energy corrections.

References
1. Bentley, K. W. β-Phenylethylamines and the isoquinoline alkaloids. *Nat. Prod. Rep.* **23**, 444–463 (2006).
2. Roddan, R., Ward, J. M., Keep, N. H. & Hailes, H. C. Pictet–Spenglerases in alkaloid biosynthesis: future applications in biocatalysis. *Curr. Opin. Chem. Biol.* **55**, 69–76 (2020).
3. Scott, J. D. & Williams, R. M. Chemistry and biology of the tetrahydroisoquinoline antitumor antibiotics. *Chem. Rev.* **102**, 1669–1730 (2002).
4. Stermitz, F. R., Lorenz, P., Tawara, J. N., Zenewicz, L. A. & Lewis, K. Synergy in a medicinal plant: antimicrobial action of berberine potentiated by 5′-methoxyhydnocarpin, a multidrug pump inhibitor. *Proc. Natl Acad. Sci.* **97**, 1433–1437 (2000).
5. Cortijo, J. et al. Bronchodilator and anti-inflammatory activities of glaucine: in vitro studies in human airway smooth muscle and polymorphonuclear leukocytes. *Br. J. Pharmacol.* **127**, 1641–1651 (1999).
6. Kashiwada, Y. et al. Anti-HIV benzylisoquinoline alkaloids and flavonoids from the leaves of Nelumbo nucifera, and structure–activity correlations with related alkaloids. *Bioorg. Med. Chem.* **13**, 443–448 (2005).
7. Rueffer, M. & Zenk, M. H. Distant precursors of benzylisoquinoline alkaloids and their enzymatic formation. *Z. Naturforsch.* **42c**, 319–332 (1987).
8. Samanani, N., Liscombe, D. K. & Facchini, P. J. Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzylisoquinoline alkaloid biosynthesis. *Plant J.* **40**, 302–313 (2004).
9. Samanani, N. & Facchini, P. J. Isolation and partial characterization of norcoclaurine synthase, the first committed step in benzylisoquinoline alkaloid biosynthesis, from opium poppy. *Planta* **213**, 898–906 (2001).
10. Lechner, H. et al. Library of norcoclaurine synthases and their immobilization for biocatalytic transformations. *Biotechnol. J.* **13**, 1700542 (2018).
11. Zhang, L. et al. Discovery, structure, and mechanism of the (*R*, *S*)-norcoclaurine synthase for the chiral synthesis of benzylisoquinoline alkaloids. *ACS Catal.* **13**, 15164–15174 (2023).
12. Stöckigt, J., Antonchick, A. P., Wu, F. & Waldmann, H. The Pictet–Spengler reaction in nature and in organic chemistry. *Angew. Chem. Int. Ed.* **50**, 8538–8564 (2011).
13. Patil, M. D., Grogan, G. & Yun, H. Biocatalyzed C–C bond formation for the production of alkaloids. *Chem. Cat. Chem.* **10**, 4783–4804 (2018).
14. Schmidt, N. G., Eger, E. & Kroutil, W. Building bridges: biocatalytic C–C bond formation toward multifunctional products. *ACS Catal.* **6**, 4286–4311 (2016).
15. Lechner, H., Pressnitz, D. & Kroutil, W. Biocatalysts for the formation of three- to six-membered carbo- and heterocycles. *Biotechnol. Adv.* **33**, 457–480 (2015).
16. Schrittwieser, J. H. & Resch, V. The role of biocatalysis in the asymmetric synthesis of alkaloids. *RSC Adv.* **3**, 17602–17632 (2013).
17. Resch, V., Schrittwieser, J. H., Siirola, E. & Kroutil, W. Novel carbon–carbon bond formations for biocatalysis. *Curr. Opin. Biotechnol.* **22**, 793–799 (2011).
18. Rueffer, M., El-Shagi, H., Nagakura, N. & Zenk, M. H. (*S*)-norlaudanosoline synthase: the first enzyme in the benzylisoquinoline biosynthetic pathway. *FEBS Lett.* **1**, 5–9 (1981).
19. Minami, H., Dubouzet, E., Iwasa, K. & Sato, F. Functional analysis of norcoclaurine synthase in *Coptis japonica*. *J. Biol. Chem.* **282**, 6274–6282 (2007).
20. Ruff, B. M., Bräse, S. & O'Connor, S. E. Biocatalytic production of tetrahydroisoquinolines. *Tetrahedron Lett.* **53**, 1071–1074 (2012).

21. Nishihachijo, M. et al. Asymmetric synthesis of tetrahydroisoquinolines by enzymatic Pictet–Spengler reaction. *Biosci. Biotechnol. Biochem.* **78**, 701–707 (2014).

22. Lichman, B. R., Zhao, J., Hailes, H. C. & Ward, J. M. Enzyme catalysed Pictet-Spengler formation of chiral 1, 1'-disubstituted-and spiro-tetrahydroisoquinolines. *Nat. Commun.* **8**, 14883 (2017).

23. Luk, L. Y. P., Bunn, S., Liscombe, D. K., Facchini, P. J. & Tanner, M. E. Mechanistic studies on norcoclaurine synthase of benzylisoquinoline alkaloid biosynthesis: an enzymatic Pictet–Spengler reaction. *Biochemistry* **46**, 10153–10161 (2007).

24. Roddan, R. et al. Acceptance and kinetic resolution of α-methyl-substituted aldehydes by norcoclaurine synthases. *ACS Catal.* **9**, 9640–9649 (2019).

25. Samanani, N. & Facchini, P. J. Purification and characterization of norcoclaurine synthase: the first committed enzyme in benzylisoquinoline alkaloid biosynthesis in plants. *J. Biol. Chem.* **277**, 33878–33883 (2002).

26. Ilari, A. et al. Structural basis of enzymatic (*S*)-norcoclaurine biosynthesis. *J. Biol. Chem.* **284**, 897–904 (2009).

27. Pesnot, T., Gershater, M. C., Ward, J. M. & Hailes, H. C. The catalytic potential of coptis japonica NCS2 revealed–development and utilisation of a fluorescamine-based assay. *Adv. Synth. Catal.* **354**, 2997–3008 (2012).

28. Lichman, B. R. et al. Dopamine-first'mechanism enables the rational engineering of the norcoclaurine synthase aldehyde activity profile. *FEBS J.* **282**, 1137–1151 (2015).

29. Sheng, X. & Himo, F. Enzymatic Pictet–Spengler reaction: computational study of the mechanism and enantioselectivity of norcoclaurine synthase. *J. Am. Chem. Soc.* **141**, 11230–11238 (2019).

30. Zhang, M. et al. Engineering a norcoclaurine synthase for one-step synthesis of (*S*)-1-aryl-tetrahydroisoquinolines. *Bioresour. Bioprocess.* **10**, 15 (2023).

31. De Sousa, J. P. M., Oliveira, N. C. S. A. & Fernandes, P. A. Rational engineering of (*S*)-Norcoclaurine synthase for efficient benzylisoquinoline alkaloids biosynthesis. *Molecules* **28**, 4265 (2023).

32. Siegbahn, P. E. M. A quantum chemical approach for the mechanisms of redox-active metalloenzymes. *RSC Adv.* **11**, 3495–3508 (2021).

33. Salahub, D. R. Multiscale molecular modelling: from electronic structure to dynamics of nanosystems and beyond. *Phys. Chem. Chem. Phys.* **24**, 9051–9081 (2022).

34. Himo, F. & de Visser, S. P. Status report on the quantum chemical cluster approach for modeling enzyme reactions. *Commun. Chem.* **5**, 29 (2022).

35. Sheng, X. & Himo, F. The quantum chemical cluster approach in biocatalysis. *Acc. Chem. Res.* **56**, 938–947 (2023).

36. Sheng, X., Kazemi, M., Planas, F. & Himo, F. Modeling enzymatic enantioselectivity using quantum chemical methodology. *ACS Catal.* **10**, 6430–6449 (2020).

37. Dong, Y. et al. Biocatalytic Baeyer–Villiger reactions: uncovering the source of regioselectivity at each evolutionary stage of a mutant with scrutiny of fleeting chiral intermediates. *ACS Catal.* **12**, 3669–3680 (2022).

38. Sheng, X. & Himo, F. Computational study of pictet–spenglerase strictosidine synthase: reaction mechanism and origins of enantioselectivity of natural and non-natural substrates. *ACS Catal.* **10**, 13630–13640 (2020).

39. Frisch, M. J. et al. *Gaussian 16, Revision C.01* (Gaussian, Inc., 2016).

40. Lee, C., Yang, W. & Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **37**, 785 (1988).

41. Beeke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98**, 5648–5652 (1993).

42. Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **132**, 154104 (2010).

43. Grimme, S., Ehrlich, S. & Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **32**, 1456–1465 (2011).

44. Marenich, A. V., Cramer, C. J. & Truhlar, D. G. 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 (2009).

45. Lichman, B. R. et al. Structural evidence for the dopamine-first mechanism of norcoclaurine synthase. *Biochemistry* **56**, 5274–5277 (2017).

Acknowledgements

Author contributions

X.S. designed the project; X.S. and H.S. co-supervised the project; S.Z. performed research and analyzed data; C.Z., A.G., and B.L. contributed also to the calculations and data analysis; S.Z., H.S. and X.S. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42004-024-01146-x.

Correspondence and requests for materials should be addressed to Hao Su or Xiang Sheng.

Peer review information *Communications Chemistry* thanks Miquel Estevez, Pedro A Fernandes, Tiziana Marino and the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.