Michał Glanowski

Modelowanie mechanizmu reakcji bakteryjnych dehydrogenaz ketosteroidowych – katalizatorów do modyfikacji leków steroidowych

Praca doktorska

Prof. dr hab. Maciej Szaleniec
Instytut Katalizy i Fizykochemii Powierzchni im. Jerzego Habera
Polskiej Akademii Nauk

Prof. dr hab. Andrzej J. Bojarski
Instytut Farmakologii im. Jerzego Maja
Polskiej Akademii Nauk

KRAKÓW 2023

Rozprawa doktorska powstała w trakcie realizacji
Środowiskowych Studiów Doktoranckich
w ramach Projektu nr POWR.03.02.00-00-I013/16,
"Interdyscyplinarność dla medycyny innowacyjnej" InterDokMed,
realizowanych w ramach Programu Operacyjnego Wiedza Edukacja Rozwój 2014-2020,
współfinansowanego ze środków Europejskiego Funduszu Społecznego.

Niniejsza dysertacja opiera się na 3 artykułach naukowych opublikowanych w czasopismach z listy filadelfijskiej oraz jednego rozdziału w recenzowanej książce:

Artykuł P1:

Michał Glanowski, Sangita Kachhap, Tomasz Borowski, Maciej Szaleniec, *Model Setup and Procedures for Prediction of Enzyme Reaction Kinetics with QM-Only and QM:MM Approaches*. In: Vanhaelen, Q. (eds) *Computational Methods for Estimating the Kinetic Parameters of Biological Systems. Methods in Molecular Biology,* vol 2385. (2022) 175-236 Humana, New York, NY. DOI: 10.1007/978-1-0716-1767-0_10, ISBN: 978-1-0716-1766-3

Artykuł P2:

Patrycja Wójcik, **Michał Glanowski**, Agnieszka M. Wojtkiewicz, Ali Rohman, Maciej Szaleniec, *Universal Capability of 3-Ketosteroid Δ1-Dehydrogenases to Catalyze Δ1-Dehydrogenation of C17-Substituted Steroids*, **Microbial Cell Factories**, 2021, 20(1), article number: 119, DOI: 10.1186/s12934-021-01611-5

Artykuł P3:

Michał Glanowski, Patrycja Wójcik, Magdalena Procner, Tomasz Borowski, Dawid Lupa, Przemysław Mielczarek, Maria Oszajca, Katarzyna Świderek, Vicent Moliner, Andrzej J. Bojarski, Maciej Szaleniec, *Enzymatic Δ1-Dehydrogenation of 3-Ketosteroids-Reconciliation of Kinetic Isotope Effects with the Reaction Mechanism*, **ACS Catalysis**, 2021, 11, 8211-8225, DOI: 10.1021/acscatal.1c01479

Artykuł P4:

Patrycja Wójcik, **Michał Glanowski**, Beata Mrugała, Magdalena Procner, Olga Zastawny, Monika Flejszar, Katarzyna Kurpiewska, Ewa Niedziałkowska, Wladek Minor, Maria Oszajca, Andrzej J. Bojarski, Agnieszka M. Wojtkiewicz, and Maciej Szaleniec, *Structure, Mutagenesis, and QM:MM Modeling of 3-Ketosteroid $Δ^1$-Dehydrogenase from Sterolibacterium denitrificans─The Role of a New Putative Membrane-Associated Domain and Proton-Relay System in Catalysis,* **Biochemistry,** 2023, 62, 3, 808-823, DOI: 10.1021/acs.biochem.2c00576

Spis treści

Streszczenie pracy doktorskiej mgr inż. Michała Glanowskiego

Δ^1-Dehydrogenazy 3-ketosteroidowe (Δ^1-KSTD) to klasa enzymów katalizujących odwracalną reakcję wprowadzenia wiązania podwójnego między pozycje C1 i C2 3-ketosteroli. Są one badane już od lat 60-tych XX wieku, jednakże do chwili obecnej nie przeprowadzono modelowania katalizowanej przez nich reakcji odwodornienia.

Niniejsza rozprawa, wypełniająca tę lukę, wykorzystuje metody chemii obliczeniowej do analizy właściwości katalitycznych Δ^1-KSTD. Badania przeprowadzono w oparciu o dwie dostępne struktury krystaliczne Δ^1-KSTD, tj. pochodzące z bakterii *Rhodococcus erythropolis* (KSTD1) oraz *Sterolibacterium denitrificans* (AcmB).

Poprzez symulacje dynamiki molekularnej z różnymi substratami przeanalizowano wiązanie różnych ketosteroidów do centrum aktywnego Δ^1-KSTD. Dla wszystkich analizowanych cząsteczek (w tym również dla takich, które zawierały rozbudowane podstawniki w pozycji C17) uzyskano zbliżone wartości oszacowanej entalpii swobodnej wiązania. Te dane zasugerowały, że taki substrat jak cholest-4-en-3-on powinien ulegać dehydrogenacji z KSTD1, wbrew doniesieniom literaturowym wykazującym na brak konwersji tego substratu. Przewidywana aktywność enzymatyczna została ostatecznie potwierdzona doświadczalnie pod wpływem przeprowadzonych obliczeń.

Obliczenia QM:MM MD potwierdziły mechanizm dehydrogenacji zaproponowany w literaturze i dostarczyły dodatkowych informacji na jego temat. Zweryfikowano kwestię stanu protonacyjnego jednej z tyrozyn w centrum aktywnym i dowiedziono, że jej istnienie w postaci jonu tyrozylowego jest kluczowe dla reaktywności enzymu. Jest to szczególnie interesujące, ze względu na potwierdzoną eksperymentalnie aktywność niektórych Δ^1-KSTD w nieznacznie kwasowym pH.

Szczególnie istotnym punktem pracy było obliczenie wartości kinetycznego efektu izotopowego i ich porównanie z wynikami otrzymanymi eksperymentalnie. Pozwoliło to na wyjaśnienie dotychczasowych, pozornych sprzeczności dotyczących mechanizmu reakcji katalizowanej przez Δ^1-KSTD. W literaturze postulowano, że reakcja przebiega dwuetapowo, a drugi krok reakcji jest procesem limitującym szybkość reakcji procesu redukcji enzymu przez substrat. Jednocześnie obserwowano znaczące wartości kinetycznego efektu izotopowego, niezależnie od tego, czy związki były deuterowane w pozycji, na którą jest czuły pierwszy czy drugi etap reakcji. Wyjaśnieniem tego zjawiska okazał się oszacowany profil entalpii

swobodnej reakcji: energie stanów przejściowych są na tyle blisko siebie, aby obydwa etapy miały istotny wpływ na wypadkową szybkość całego procesu.

W przypadku AcmB dokonano analizy wpływu mutacji kluczowych aminokwasów w centrum aktywnym na profil energetyczny reakcji. Uzyskany trend jest w zgodzie z eksperymentalnymi pomiarami aktywności i ukazuje rolę poszczególnych wiązań wodorowych w otoczeniu substratu w procesie katalitycznym.

Ponadto zbadano rolę tak zwanej „pętli", czyli 50-cio aminokwasowego fragmentu występującego w AcmB oraz wielu innych KSTD, który po rozwiązaniu struktury okazał się domeną wiążącą się do membrany. Fragment ten nie tylko wiąże enzym do membrany, ale oddziałuje z różnymi substratami, szczególnie tymi podstawionymi w pozycji C17. Dzięki temu AcmB wykazuje większe powinowactwo względem substratów takich jak cholest-4-en-3-on niż takich jak progesteron. Na koniec przeprowadzono obliczenia wyjaśniające enancjoselektywność aktywacji steroidu.

Abstract of the doctoral thesis of mgr inż. Michał Glanowski

The 3-ketosteroid $\Delta 1$-dehydrogenases (Δ^1-KSTD) are a class of enzymes that catalyze the reversible reaction of dehydrogenation, introducing a double bond between the C1 and C2 positions of 3-ketosterols. Although they have been studied since the 60's of the XXth century, their catalytic mechanism has not been described by modeling approach.

This dissertation fills the gap by using computational chemistry methods to analyze the catalytic properties of Δ^1-KSTD. The research was conducted based on two available crystal structures of Δ^1-KSTD, i.e., from bacteria *Rhodococcus erythropolis* (KSTD1) and *Sterolibacterium denitrificans* (AcmB).

The binding of different ketosteroids to the active site of Δ^1-KSTD was studied by means of molecular dynamics simulations. For every analyzed molecule (including those with extended substituents in the C17 position), similar values of estimated binding free energy were obtained. These data suggested that substrates like cholest-4-ene-3-dione should be dehydrogenated by KSTD1, contrary to literature reports indicating a lack of conversion of this substrate. Prompted by modeling predictions, the enzymatic activity was finally confirmed experimentally.

QM:MM MD calculations confirmed the mechanism of dehydrogenation proposed in the literature and delivered additional mechanistic details. The protonation state of one of the tyrosines at the active site was verified. It was proved that its deprotonated state is a crucial factor for enzyme activity. It is especially interesting due to the experimentally confirmed activity of some Δ^1-KSTD from *S. denitrificans* at slightly acidic pH.

One of the most relevant points of the dissertation was the calculation of the kinetic isotopic effect and the comparison with experimentally measured values. This allowed for an explanation of apparent contradictions related to the mechanism of the reaction catalyzed by Δ^1-KSTD. According to the literature, the reaction proceeds in two steps, with the second being a rate-limiting process of the reduction of the enzyme by a substrate. Simultaneously, significant values of the kinetic isotopic effect were observed for substrates deuterated in the position for which the first step is sensitive, as well as for isotopic substitution in a position related to the second one. An explanation for this phenomenon is the estimated free energy profile of the process: the free energies of both transition states are close enough to each other for both processes to influence the overall rate of the reaction.

In the case of AcmB, the effect of mutations of the key amino acids in the active site on the reaction profile was analyzed. The obtained trend is in agreement with experimental activity measurements and demonstrates the importance of each hydrogen bond in substrate surroundings in the catalytic process.

In addition, the role of the so-called 'loop', a 50-amino-acid fragment found in AcmB and many other Δ^1-KSTDs was studied. After resolving of its structure it turned out to be a putative membrane-binding domain not a loop. This fragment binds the enzyme to the membrane and interacts with various substrates, especially those substituted at the C17 position. As a result, AcmB exhibits a higher affinity for substrates, such as cholest-4-en-3-one, than for others, such as progesterone. Finally, calculations were performed explaining the enantioselectivity of steroid activation.

Wstęp

Δ^1-Dehydrogenazy 3-ketosteroidowe

Steroidy to klasa związków chemicznych, które zawierają w swojej strukturze szkieletu węglowego charakterystyczny układ czterech sprzężonych pierścieni. Mają one duże znaczenie w systemach biologicznych jako składowe błon komórkowych oraz jako cząsteczki sygnałowe.[1]

Wiele mikroorganizmów jest zdolnych do modyfikacji lub pełnej mineralizacji steroidów.[2] Jednym z najważniejszych steroidów jest cholesterol. Proces jego degradacji do dwutlenku węgla i wody składa się z wielu etapów, katalizowanych przez różne enzymy. Niektóre z tych białek są interesujące z punktu widzenia przemysłu farmaceutycznego ze względu na zdolność do specyficznych modyfikacji steroidów. Δ^1-Dehydrogenacja jest przykładem takiego procesu, kluczowego w początkowych etapach bakteryjnej degradacji steroidów oraz modyfikacji leków steroidowych.

Pomimo, że enzymy tej klasy występują w ponad 500 organizmach, w chwili obecnej znane są tylko dwie struktury krystaliczne Δ^1-dehydrogenaz 3-ketosteroidowych (Δ^1-KSTD), KSTD1 z *Rhodococcus erythropolis* oraz AcmB ze *Sterolibacterium denitrificans*.[3] Znane sekwencje Δ^1-KSTD są względnie zróżnicowane, w pracy Rohmana i Dijkstry z 2019 roku wyróżniono 4 klasy enzymów.[3] Dwa enzymy analizowane w tej pracy, AcmB i KSTD1, cechują się 35,1% identycznością i 54,1% podobieństwem sekwencji. Porównanie struktur oraz nałożenie kluczowych reszt w centrum aktywnym omawianych enzymów przedstawiono na Rysunku 1. Oba enzymy mają wydłużony kształt i można wyróżnić w ich budowie dwie domeny: wiążącą FAD i katalityczną. Ponadto AcmB posiada w swojej strukturze dodatkową domenę, której funkcję zaproponowano m.in. na podstawie obliczeń przeprowadzonych w niniejszej rozprawie. Centra aktywne obu białek wykazują szczególnie duże podobieństwo. Trzy tyrozyny obecne w KSTD1 (Tyr487, Tyr318, Tyr119) mają swoje odpowiedniki w AcmB (Tyr536, Tyr363, Tyr118). Jednak w centrum aktywnym tego ostatniego enzymu obecne są trzy dodatkowe tyrozyny (Tyr115, Tyr467, Tyr466), którym w białku z *R. erythropolis* odpowiadają fenyloalaniny (Phe116, Phe418, Phe418).
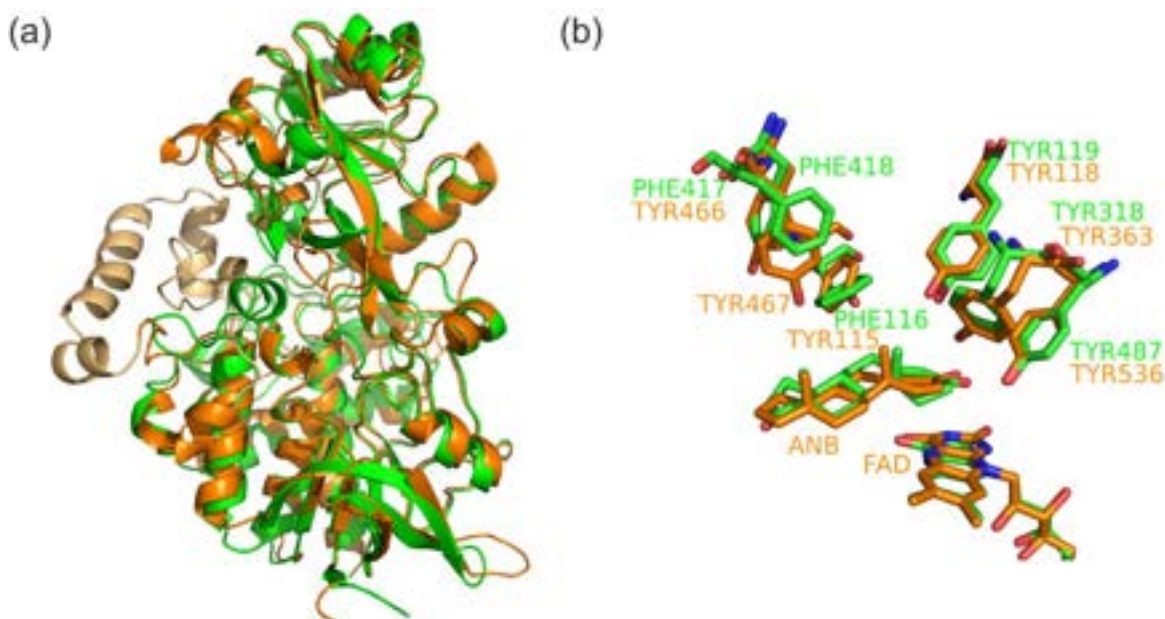
Rysunek 1. (a) Nałożenie struktur AcmB z S. denitrificans (pomarańczowy) i KSTD1 z R. erythropolis (zielony). Fragment sekwencji AcmB niemający odpowiednika w KSTD1 wyróżniono jaśniejszym kolorem. (b) Porównanie kluczowych reszt w centrach aktywnych AcmB (pomarańczowy) i KSTD1 (zielony), widoczny również produkt reakcji androstadi-1,4-en-3,17- dion (ANB).

Pomimo, iż dehydrogenazy 3-ketosteroidowe są badane od ponad 60 lat, postulowany w literaturze mechanizm 1,2-dehydrogenacji nie został dotąd zweryfikowany metodami chemii kwantowej. Obliczeniowa analiza Δ^1-KSTD stała się możliwa w 2013 roku wraz z opublikowaniem pierwszej struktury krystalicznej z bakterii *Rhodococcus erythropolis* (KSTD1, pdb 4C3X i 4C3Y).[4] Kilka lat później struktura innej Δ^1-KSTD z *Sterolibacterium denitrificans* (AcmB) została rozwiązana w naszym laboratorium.[5] Dzięki przeprowadzonym badaniom eksperymentalnym dla obu enzymów stało się możliwe dokonanie wyczerpującej analizy obliczeniowej i weryfikacji hipotezy mechanistycznej zaproponowanej w literaturze.

Δ^1-KSTD katalizują stereoselektywne odwodornienie 3-ketosteroidu między atomem C1 i C2 (por. Rysunek 2). Opierając się na strukturze krystalicznej i danych eksperymentalnych, Dijkstra i jego współpracownicy zaproponowali dwuetapowy mechanizm reakcji.[4] Hipoteza mechanistyczna zakłada istnienie w centrum aktywnym anionu tyrozylowego, który w pierwszej fazie umożliwia oderwanie protonu z pozycji C2. Następnie karboanionowy produkt przejściowy jest stabilizowany przez tautomerię keto-enolową, a później jon wodorkowy zostaje przeniesiony z pozycji C1 steroidu do atomu N5 w cząsteczce FAD.
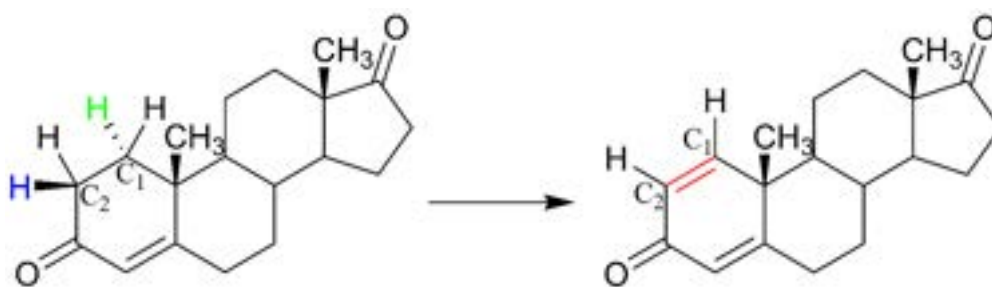
Rysunek 2. Reakcja dehydrogenacji na przykładzie cząsteczki androst-4-en-3-onu. Rysunek zaadaptowano z P3.

Aktualny stan wiedzy na temat reakcji katalizowanej przez Δ^1-KSTD

Dwa najczęściej spotykane mechanizmy związane z enzymatycznym wprowadzeniem podwójnego wiązania to dehydrogenacja i dehydratacja. Dehydrogenacja przebiega poprzez bezpośrednią eliminację atomów wodoru, natomiast dehydratacja wymaga najpierw hydroksylacji jednej z pozycji, a następnie eliminacji cząsteczki wody. Hipoteza zakładająca dehydratację 3-ketosteroidów przez Δ^1-KSTD nie znalazła potwierdzenia eksperymentalnego. Po pierwsze Δ^1-KSTD katalizują 1,2-odwodornienie w warunkach beztlenowych, podczas gdy wiele enzymów katalizujących dehydratację wymagają tlenu cząsteczkowego w celu przeprowadzenia hydroksylacji.[6] Po drugie hydroksylowane steroidy (tj. hipotetyczne produkty pośrednie) nie są konwertowane przez Δ^1-KSTD.[7] M.in. na tej podstawie przyjęto hipotezę mechanistyczną dehydrogenacji.

Wg tej hipotezy reakcję katalizowaną przez dehydrogenazy 3-ketosteroidowe można podzielić na dwie etapy. Pierwszy z nich obejmuje redukcję FAD przez substrat, który jednocześnie ulega utlenieniu (*reductive half-reaction*, RHR), natomiast podczas drugiego akceptor elektronowy (naturalnie najprawdopodobniej witamina K2)[8] utlenia układ flawinowy (*oxidative half-reaction*, OHR) (Rysunek 3). W niniejszej pracy zdecydowana większość uwagi jest poświęcona etapowi RHR.
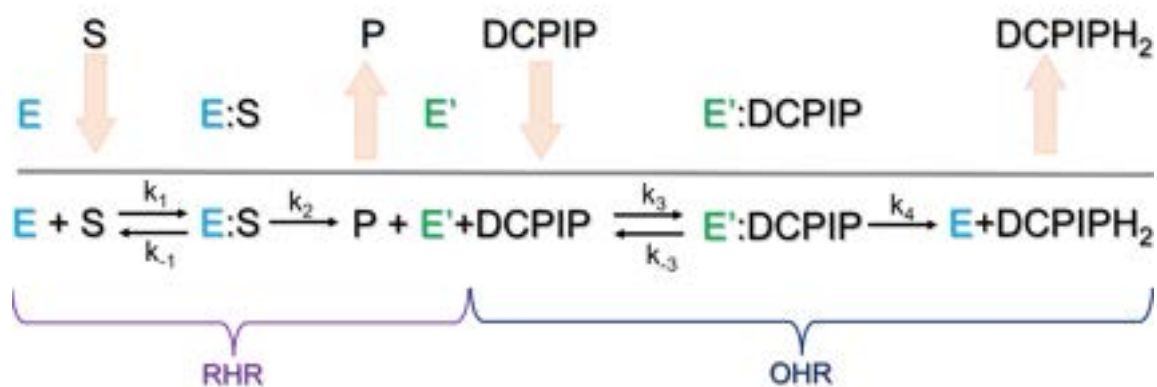
Rysunek 3. Schemat reakcji z udziałem Δ^1-KSTD z 2,6-dichloroindofenolem (DCPIP) jako sztucznym akceptorem elektronowym. Rysunek zaadaptowano z P3. E – enzym z utlenionym FAD, E' – enzym ze zredukowanym FADH$^-$/FADH$_2$.

Obecnie przyjmuje się, że w wyniku RHR następuje stereoselektywne przeniesienie atomów wodoru z pozycji C2β i C1α. Zostało to stwierdzone m.in. na podstawie eksperymentów z substratami podstawionymi izotopowo.[9] Dotychczasowe doświadczenia wykazały, że chociaż możliwa jest inicjalizacja reakcji z pozycji C2α, pozycja C2β jest silnie preferowana.[10]

Kolejność odrywania atomów wodoru określono na podstawie doświadczeń z enzymem zawierającym zredukowaną formą FADu i ciężką wodą (D_2O). Okazało się, że w takich warunkach mimo, że nie zachodzi proces dehydrogenacji, następuje wprowadzenie atomu deuteru do substratu w pozycji C2, tj. następuje wymiana izotopowa H/D.[11] Eksperyment ten dowodzi, że w pierwszym etapie RHR dochodzi do oderwania protonu przez zasadę tyrozylową i wtedy może dojść do wymiany protu na deuter dzięki interakcji tyrozyny z cząsteczkami rozpuszczalnika. Ponieważ proces przeniesienia jonu hydroniowego nie może zajść (FAD jest zredukowany), tyrozyna oddaje atom deuteru do aktywowanego substratu uwalniając znakowany steroid.

Te przesłanki doprowadziły jeszcze w latach 60-tych do sformułowania hipotezy, zgodnie z którą dehydrogenacja przeprowadzana przez enzymy Δ^1-KSTD przebiega w dwóch etapach. W pierwszym z nich, proton zostaje oderwany z pozycji C2β. Powstały karboanion jest stabilizowany za pomocą rezonansu keto-enolowego oraz wiązań wodorowych z atomem tlenu przy C3. W drugim etapie, jon wodorkowy jest przenoszony z pozycji C1α na atom N5 cząsteczki FAD. Nieco więcej problemów nastręczała analiza wyników eksperymentów kinetycznych. Postulowano, że pierwszy etap jest względnie szybki w stosunku do drugiego,

kontrolującego obserwowaną szybkość reakcji.[12] Więcej informacji na ten temat dostarczyły pomiary Kinetycznego Efektu Izotopowego (KIE). Wartości otrzymane dla enzymu z organizmu *Bacillus sphaericus* wykazały porównywalne wartości KIE dla substratów podstawionych w pozycjach C1 i C2. To z kolei sugerowało, że obydwa etapy reakcji przebiegają z porównywalną szybkością, lub etapem limitującym szybkość procesu jest jednoczesne przeniesienie obu atomów wodoru.

Publikacja pierwszej struktury krystalicznej enzymu z *R. erythropolis* dostarczyła kolejnych szczegółów dotyczących enzymów Δ^1-KSTD (kody PDB 4C3X i 4C3Y). Bazując na strukturze enzymu z cząsteczką FAD i produktem reakcji w centrum aktywnym autorzy określili kluczowe dla reakcji reszty. Tyr487 i Gly491 znajdowały się w pozycji dogodnej do tworzenia wiązań wodorowych z grupą ketonową steroidu. Tyr318 okazała się być w odpowiedniej pozycji do oderwania atomu wodoru z pozycji C2 substratu. Ponadto wiązanie wodorowe między Tyr119 i Tyr318 mogło wpływać na własności kwasowo-zasadowe Tyr318 (patrz Rysunek 1).

Istotność poszczególnych aminokwasów została zweryfikowana za pomocą ukierunkowanej mutagenezy KSTD1 i testów aktywności katalitycznej zmutowanych wariantów enzymu. Kolejne tyrozyny były zastępowane przez fenyloalaninę co wprowadzało duża zmianę funkcjonalną przy minimalnym zaburzeniu struktury. Mutacja Y318F doprowadziła do kompletnej utraty aktywności. Tymczasem, wariant Y487F cechował się aktywnością na poziomie 2,6% w odniesieniu do natywnego enzymu. Co ciekawe, mutacja Y119F okazała się jeszcze bardziej znacząca, skutkując spadkiem aktywności do około 0,5%.[4]

Te fakty doświadczalne pozwoliły na zaproponowanie roli każdej spośród wymienionych tyrozyn.[4] Tyr487 miałaby stabilizować ketosteroid w trakcie reakcji. Tyr318 byłaby zasadą, która jest niezbędna do oderwania protonu 2Hβ w pierwszym etapie reakcji. Postulowano, że aminokwas ten występuje w centrum aktywnym w postaci zdeprotonowanej - anionu tyrozylowego. Tyr119 z kolei miałaby oddziaływać z Tyr318 i być odpowiedzialna za zmianę mikrootoczenia Tyr318 co skutkowałoby obniżeniem pK_a grupy fenolowej, w taki sposób, aby jony tyrozylowe istniały w dostatecznej ilości w lekko zasadowym i neutralnym pH (dla jakiego obserwowano optimum reakcji dehydrogenacji).[7,13–17]

Podany w literaturze mechanizm zaprezentowano na rysunku 4.
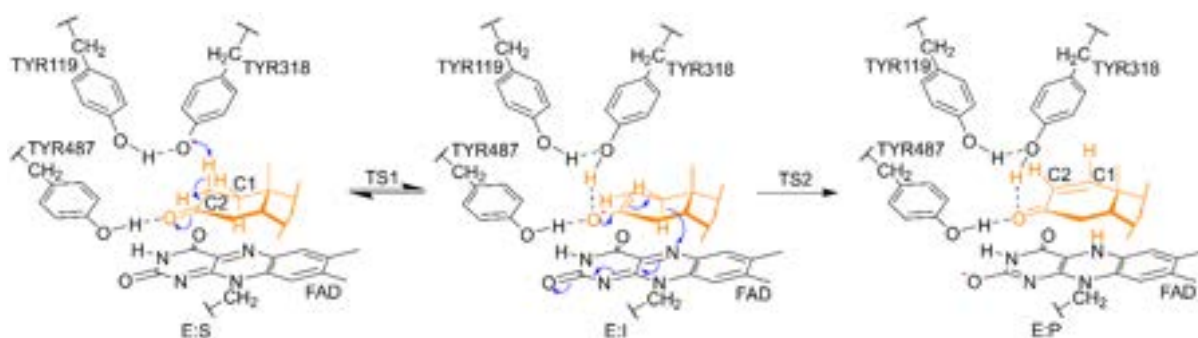
Rysunek 4. Mechanizm reakcji dehydrogenacji. Rysunek zaadaptowano z P3.

Ważnym parametrem różniącym poszczególne Δ^1-KSTD jest specyficzność substratowa. Niektóre enzymy katalizują odwodornienie tylko tych steroidów, które mają wiązanie podwójne między atomami C4 i C5 (tzw. Δ^4-steroidy),[14,18,19] istnieją jednak takie, które wykazują większą aktywność wobec substratów bez tego wiązania. Fragmentem struktury niezbędnym do zajścia reakcji jest natomiast grupa ketonowa w pozycji C3. Eksperymenty z innymi grupami funkcyjnymi w tej pozycji wykazały brak aktywności związków w reakcji 1,2-odwodornienia.[13,19,20] Szczególnie ciekawa jest kwestia podstawników w pozycji C17. Większość Δ^1-KSTD jest aktywna wobec substratów z niewielkimi podstawnikami w tej pozycji, z kolei te z bardziej rozbudowanymi (np. cholest-4-en-3-on) są substratami dla niewielkiej grupy enzymów (enzymy z *R. equi* oraz *N. simplex*),[19,21] do której zalicza się również AcmB z *Sterolibacterium denitrificans*.[20] **W dotychczasowych publikacjach postulowano brak aktywności KSTD1 wobec cholest-4-en-3-onu, co jest również analizowane w niniejszej dysertacji.**[3]

Jedną z najistotniejszych różnic strukturalnych między KSTD1 i AcmB jest 50 aminokwasowy fragment sekwencji występujący m.in. w dehydrogenazie z *S. denitrificans*. Rola biologiczna i struktura drugorzędowa tego fragmentu nie była oczywista, dopóki nasz zespół nie rozwiązał struktury krystalicznej AcmB.[5] Modelowanie homologiczne nie było pomocne w tym zakresie ze względu na brak odpowiednich struktur szablonowych dla tego fragmentu sekwencji. Kolejną niewiadomą AcmB była kwestia optimum pH reakcji odwodornienia. Większość opisanych w literaturze KSTD wykazuje optimum reakcji w pH w zakresie od 8-10.[7,13–17] Tymczasem enzym z *S. denitrificans,* w zależności od zastosowanego reutleniacza, może wykazywać więcej niż jedno optimum aktywności w funkcji pH,[22] przy czym gdy stosowany jest jeden ze standardowych utleniaczy KSTD, dichloroindofenol (DCPIP), optimum to znajduje się przy pH 6,5. Jest to szczególnie intrygujący aspekt w kontekście wymogu istnienia w centrum aktywnym zdeprotonowanej tyrozyny potrzebnej

do aktywacji substratu. Przedstawiona w dysertacji publikacja P4 podnosi między innymi ten temat.

Kinetyczny efekt izotopowy (KIE)

Szybkość reakcji chemicznej w większości przypadków zależy od mas atomów przenoszonych w reakcji, lub znajdujących się w ich bliskim sąsiedztwie. Dysponując dwoma stałymi szybkości dla danej reakcji chemicznej: jedną odpowiadającą układowi, w którym dokonano podstawienia izotopowego wybranych atomów (k^H), i drugą, który nie zawiera takiego podstawienia (k^L), można zdefiniować kinetyczny efekt izotopowy (*Kinetic Isotopic Effect*, KIE):

$$KIE = \frac{k^L}{k^H}$$

Stała k^H może odpowiadać reagentom, gdzie np. wybrane atomy wodoru zostały wymienione na deuter. W większości przypadków k^H jest mniejsze od k^L co oznacza, że wartość KIE jest większa od jedności. W tej sytuacji mowa o normalnym kinetycznym efekcie izotopowym (reakcje z reagentami podstawionymi cięższym izotopem przebiegają wolniej). Możliwe jest jednak, aby KIE uzyskało wartości mniejsze od jedności i tego rodzaju przypadki określa się jako odwrotny KIE.

Ponadto istnieje podział KIE ze względu na to, które atomy zostały podstawione. Jeśli są to atomy uczestniczące bezpośrednio w tworzonych lub zrywanych wiązaniach, mówimy o pierwszorzędowym KIE, w przeciwnym przypadku – o drugorzędowym.

Podstawienie izotopowe jest niezwykle użytecznym narzędziem w badaniu mechanizmów reakcji. Pozwala na identyfikację, które atomy są przenoszone w reakcjach (efekt znakowania izotopowego) oraz które etapy elementarne są dostatecznie wolne by mieć wpływ na szybkość badanej reakcji.

Aby oszacować KIE metodami chemii kwantowej potrzebne jest obliczenie wysokości bariery entalpii swobodnej dla badanej reakcji. Szczegóły metody obliczania KIE zostały opisane w streszczeniu artykułu P1 oraz w samej pracy P1.

Izotopowy Efekt Wiązania

Analogicznie jak dla stosunku stałych szybkości zdefiniowany jest KIE, tak dla stosunku stałych równowagi definiuje się równowagowy efekt izotopowy (EIE). Jeśli stałe te opisują tworzenie kompleksu enzym-substrat, to uzyskana wartość nosi nazwę izotopowego efektu wiązania (*Binding Isotopic Effect,* BIE).

$$BIE = \frac{K_{E:S}^L}{K_{E:S}^H}$$

W powyższym równaniu $K_{E:S}^L$ to stała równowagi tworzenia kompleksu E:S w układzie bez podstawienia izotopowego, a $K_{E:S}^H$ to analogiczna wielkość związana z układem, gdzie jeden lub kilka z atomów zamieniono na inny izotop.

Kwestia obliczeniowego oszacowania BIE jest analogiczna jak w przypadku KIE, z tym, że zamiast operować na różnicy entalpii swobodnej aktywacji, operuje się na różnicy entalpi swobodnej wiązania substratu do enzymu.

$$BIE = exp\left(\frac{-(\Delta G_{E:S}^L - \Delta G_{E:S}^H)}{RT}\right)$$

W powyższym równaniu $\Delta G_{E:S}^L$ i $\Delta G_{E:S}^H$ to entalpie swobodne wiązania substratu, odpowiednio, dla układu niepodstawionego i podstawionego izotopowo.

Warto zauważyć, że do wyznaczenia różnic wartości $\Delta G_{E:S}^L$ i $\Delta G_{E:S}^H$ nie ma potrzeby, aby wyznaczać je osobno. Do tego celu wystarczy znajomość poprawek wibracyjnych dla obydwu układów.[23] Obliczeniowo są one uzyskiwane na podstawie zoptymalizowanych struktur substratu w rozpuszczalniku oraz kompleksu enzymu z substratem.

Cel pracy

Celem niniejszej dysertacji była analiza reakcji 1,2-dehydrogenacji katalizowanej przez Δ^1-dehydrogenazy 3-ketosteroidowe z *R. erythropolis* i *S. denitrificans* metodami chemii obliczeniowej, oraz weryfikacja dostępnych i dostarczenie nowych informacji na temat mechanizmu katalizowanej przez nie reakcji odwodornienia.

Najważniejszym zadaniem była weryfikacja hipotezy mechanistycznej zaproponowanej przez Dijkstrę i współpracowników.[24] Postulowany w literaturze proces został porównany z alternatywnym mechanizmem, w którym odwodornienie rozpoczyna się od oderwania atomu wodoru przy atomie C1. Innym potencjalnym wariantem był jednoczesny transfer obydwu atomów wodoru z substratu z pozycji 2β i 1α. Kolejnym aspektem mechanistycznym możliwym do zbadania była stereoselektywność reakcji, co zostało osiągnięte przez obliczenie profilu energetycznego dla dehydrogenacji, w której pierwszy przenoszony atom wodoru pochodzi z pozycji C2α a nie C2β.

Następnym pytaniem badawczym był stan protonacyjny kluczowego z punktu widzenia reakcji aminokwasu Tyr318 w KSTD1 oraz Tyr363 w AcmB. Zakłada się, że ten aminokwas jest zdeprotonowany na początku RHR.[24] Rozważenie tej tyrozyny w jej standardowym stanie protonacyjnym (tj. formie fenolowej), doprowadziło do kolejnej alternatywy mechanistycznej. Jest to szczególnie istotna kwestia zważywszy na szeroki zakres aktywności AcmB w zakresie pH od 6–9.

Kolejnym celem pracy, była konfrontacja przewidywanych przez modelowanie wartości kinetycznego efektu izotopowego z wynikami eksperymentalnymi. Metody teoretyczne pozwalają na analizę wpływu podstawienia izotopowego poszczególnych atomów na profil reakcji. Istotnym celem był również opis ilościowy wpływu mutacji poszczególnych aminokwasów w centrum aktywnym na profil reakcji i powiązanie tych wyników z eksperymentalnie wyznaczonymi szybkościami reakcji. Istotne również było skonfrontowanie zdolności predykcyjnych teoretycznego mechanizmu 1,2-dehydrogenacji do opisu reakcji różnych 3-ketosteroidów i porównanie przewidywań modelu z pomiarami kinetycznymi.

Istotnym aspektem było również wyjaśnienie roli unikatowego fragmentu „pętli" AcmB, nieobecnego w KSTD1. W tym przypadku szczególnie pomocne były symulacje dynamiki molekularnej przeprowadzone dla obydwu enzymów z różnymi ketosteroidami

związanymi w centrum aktywnym dzięki czemu możliwe było zaproponowanie roli „pętli"
w katalizie

Podsumowując, praca zakładała dogłębną charakteryzację Δ^1-KSTD ze szczególnym nastawieniem na aspekty mechanistyczne i kinetyczne, za pomocą współcześnie dostępnych metod chemii obliczeniowej.

Metody obliczeniowe

Obliczenia kwantowo-mechaniczne

Metody obliczeniowe w chemii służą rozwiązaniu równania Schrödingera i uzyskaniu wartości energii będącej wartością własną hamiltonianu, celem dostarczenia informacji chemicznych na temat badanego układu. Większość popularnych metod obliczeniowych jest opartych na przybliżeniu Borna-Oppenheirmera. Jego podstawą jest stwierdzenie, że ze względu na różnice mas między jądrami atomowymi i elektronami, można przyjąć, że stan elektronowy jest zawsze dopasowany do aktualnych położeń jąder. Tzn. nie ma sprzężenia pomiędzy ruchem elektronów i jąder, a więc funkcja falowa $\Psi(R,r)$ może być przedstawiona jako iloczyn funkcji elektronowej $\chi(R,r)$, zależnej od współrzędnych elektronowych i w sposób parametryczny od współrzędnych jąder atomowych, i funkcji jądrowej $\phi(R)$.

$$\Psi(R,r) = \chi(R,r)\phi(R)$$

Prekursorem wielu nowoczesnych metod chemii kwantowej jest metoda Hartree-Focka (HF).[25] Energia elektronowa w tej metodzie dana jest następującym równaniem:

$$E_{HF} = \Sigma_{i=1}^{n} h_i + \Sigma_{i>j=1}^{n}\left(J_{ij} - K_{ij}\right)$$

W powyższym równaniu sumowanie odbywa się po obsadzonych orbitalach, gdzie h_i to całka jednoelektronowa, na którą składa się energia oddziaływania kulombowskiego elektron-jądra atomowe oraz energia kinetyczna elektronu, J_{ij} to dwuelektronowa całka kulombowska reprezentująca oddziaływanie elektron-elektron, a K_{ij} to z kolei całka wymienna.

Problem z dokładnością tej metody wynika z nieuwzględnienia korelacji ruchu elektronów. Wiele różnych podejść zostało wypracowanych, aby rozwiązać ten problem. Powstał cały szereg metod obliczeniowych, gdzie problem korelacji próbowano rozwiązać poprzez modyfikacje używanej funkcji falowej. W teorii, jest możliwe, aby obliczyć dokładną energię z uwzględnieniem korelacji elektronowej poprzez rozwinięcie funkcji falowej w liniową kombinację wszystkich możliwych wyznaczników Slatera (pełne CI). W praktyce jednak jest to możliwe tylko dla wyjątkowo małych układów ze względu na koszty obliczeniowe pełnej metody CI. Pewnym kompromisem jest wybranie tylko tych wyznaczników, które zawierają największą część poprawki – ten sposób realizuje m.in. metoda

sprzężonych klasterów (CC). Alternatywą dla tych metod jest teoria funkcjonałów gęstości (DFT), gdzie zamiast na funkcji falowej operuje się na gęstości elektronowej. Z pierwszego twierdzenia Hohenberga-Kohna wiadomo, że gęstość elektronowa zawiera dość informacji, by znaleźć dokładną energię układu (uwzględniając energie korelacji). Z kolei drugie twierdzenie Hohenberga-Kohna stwierdza, że minimum energii uzyskuje się dla dokładnej gęstości elektronowej.[26] Dzięki tym twierdzeniom możliwe było powstanie całej gamy różnych funkcjonałów gęstości. Udowodniono bowiem, że z samej gęstości elektronowej można rozwiązać problem korelacji, jednakże nie pozwala to na jednoznaczne określenie jakie operacje są potrzebne by uzyskać poprawne rezultaty. Energia elektronowa w metodach DFT jest obliczana z równania Kohna-Shama:

$$E_{KS}[\rho] = T[\rho] + V_{ne}[\rho] + J[\rho] + E_{xc}[\rho]$$

Pierwszy człon tego równania ($T[\rho]$) jest związany z energią kinetyczną elektronów, drugi ($V_{ne}[\rho]$) z oddziaływaniem elektronów z potencjałem wytworzonym przez ładunki jąder atomowych. Z kolei $J[\rho]$ to energia oddziaływania kulombowskiego między elektronami (analogiczna do odpowiedniego członu z równania na energie E_{HF}). Ostatni element równania to energia korelacyjno-wymienna, obliczana z funkcjonału korelacyjno-wymiennego, którego dokładna postać jest nieznana.

Nie istnieje analityczny sposób wyprowadzenia matematycznej formy E_{xc}. Znane są tylko pewne więzy fizyczne, które muszą być spełnione przez dokładny funkcjonał.

Istnieją funkcjonały hybrydowe, w których wykorzystuje się pewien ułamek energii wymiany (obliczany jak w metodzie HF). Tego rodzaju podejście doprowadziło do powstania jednych z najpopularniejszych funkcjonałów, m.in. B3LYP (Becke, 3-parameter, Lee-Yang-Parr).[27] Wynika to z względnej zgodności wyników obliczeniowych z eksperymentalnie wyznaczonymi.

Metody QM:MM

W metodach QM:MM (*quantum mechanicss: molecular mechanic*) część układu jest opisywana kwantowo-mechanicznie (warstwa QM o współrzędnych r$_{QM}$), a część za pomocą mechaniki molekularnej (warstwa MM o współrzędnych r$_{MM}$). Pozwala to na analizowanie bardzo dużych układów, niemożliwych do obliczeń za pomocą „czystego" DFT. Metoda ONIOM[28] realizuje subtraktywny wariant QM:MM co oznacza, że aby uzyskać końcową

energię układu ($E^{ONIOM}(r_{QM}, r_{MM})$) należy obliczyć energię całego systemu w MM ($E^{MM}(r_{QM}, r_{MM})$), odjąć energię warstwy QM obliczonej również na poziomie MM ($E^{MM}(r_{QM})$), a następnie dodać energię warstwy QM obliczonej na poziomie QM ($E^{QM}(r_{QM})$):

$$E^{ONIOM}(r_{QM}, r_{MM}) = E^{MM}(r_{QM}, r_{MM}) - E^{MM}(r_{QM}) + E^{QM}(r_{QM})$$

Konsekwencją tej metodologii jest konieczność ustalenia parametrów MM dla wszystkich warstw. W pracach **P3** i **P4** zastosowano inny rodzaj obliczeń QM:MM, tj. addytywny wariant tej metody. W tym przypadku całkowita energia układu ($E^{QMMM\ add}(r_{QM}, r_{MM})$) to suma energii warstwy QM obliczonej na poziomie kwantowo-mechanicznym oraz energii warstwy MM obliczonej na poziomie mechaniki molekularnej. Ten schemat jest zaimplementowany między innymi w bibliotece fDYNAMO.[29]

$$E^{QMMM\ add}(r_{QM}, r_{MM}) = E^{MM}(r_{MM}) + E^{QM}(r_{QM})$$

Metody półempiryczne

Metoda HF i DFT skalują się nominalnie jak O(n^4) (w notacji dużego O), gdzie n to liczba funkcji bazy. Jest to spowodowane liczbą całek dwuelektronowych potrzebnych do obliczenia energii układu. Dzięki pewnym technikom optymalizacyjnym (m.in. pominięciu całek, które są zbyt małe, by mieć faktyczny wkład w energię), rzeczywista złożoność tych metod to około O(n^3). Metody półempiryczne idą znacznie dalej w stronę redukcji kosztu obliczeniowego, jednakże kosztem dokładności.

Pierwsze istotne uproszczenie stosowane w metodach półempirycznych to baza funkcyjna. Zwykle tylko orbitale walencyjne są opisane za pomocą funkcji orbitalnych, w dodatku z wykorzystaniem tzw. bazy minimalnej. Zakłada się, że powłoki niewalencyjne nie ulegają zmianie i mogę być sparametryzowane. Te uproszczenia ograniczają m.in. możliwość polaryzacji cząsteczki.[30]

Różne podejścia zostały opracowane celem redukcji liczby obliczanych całek dwuelektronowych. W przybliżeniu ZDO (*Zero-Differential Overlap*) obliczane są tylko całki postaci (aa|bb). Powoduje to zmniejszenie złożoności obliczeniowej z O(n^4) do O(n^2). Z kolei w metodzie INDO (*Intermediate Neglect of Differentia Overlap*) stosowane jest podobne podejście, ale bez pominięcia całek zawierających funkcje scentrowane na jednym atomie. Kolejnym przybliżeniem jest NDDO (*Neglect of Diatomic Differentia Overlap*), gdzie oblicza

się także całki postaci (ab|cd), gdy orbitale a,b są scentrowane na jednym atomie, a orbitale c,d na drugim. Jedną z popularniejszych metod empirycznych realizujących to przybliżenie jest AM1 (*Austin Model 1*).[31] Powstała jako udoskonalenie starszej metody MNDO (*Modified of Neglect Diatomic Overlap*), również opartą o przybliżenie NDDO. AM1 zostało pomyślane, aby poprawić opis wiązań wodorowych i oddziaływań van der Waalsa przez MNDO (przeszacowane siły odpychające). Zostało to osiągnięte poprzez modyfikacje oddziaływania między rdzeniami atomów, tj. dodanie odpowiednio sparametryzowanych funkcji gaussowskich. Poza tym członem równania AM1 są analogiczne jak MNDO.

Metody półempiryczne często są skalibrowane dla specyficznych układów, z ograniczeniem użycia do określonego podzbioru pierwiastków.

Potencjał średniej siły

Pojęcie potencjału średniej siły (PMF – *Potential of Mean Force*) zostało wprowadzone przez Kirkwooda w 1935 roku.[32] Aby wyjaśnić to pojęcie warto zacząć od relacji pomiędzy potencjałem (energią, U), a siłą (F).

$$F(\xi; r) = \nabla U(\xi; r)$$

Zarówno siła jak i energia potencjalna są funkcjami nie tylko współrzędnej reakcji ξ ale też współrzędnych innych atomów r. Bardziej jednak niż profil energii dla pewnej kombinacji r, interesująca jest jego uśredniona wartość dla różnych wartości r, tak aby był on funkcją tylko współrzędnej reakcji ξ. Aby ta wielkość miała sens fizyczny, powinna być to średnia ważona, gdzie wagą jest gęstość prawdopodobieństwa tego, że układ znajduje się w stanie $U(\xi; r)$ (co można uzyskać z rozkładu Boltzmanna). Tak więc średnia siła w funkcji ξ może być opisana równaniem:

$$<F(\xi; r)> = \frac{\int exp\left(-\beta U(\xi; r)\right)\nabla U(\xi; r)dr}{\int exp\left(-\beta U(\xi; r)\right)dr}$$

W powyższym równaniu β oznacza odwrotność iloczynu stałej Boltzmanna i temperatury. Uzyskaną wartość średniej siły można interpretować jako gradient pewnego potencjału ($w(\xi)$), który określa się jako potencjał średniej siły (*Potential of Mean Force*).

$$\nabla w(\xi) = <F(\xi; r)>$$

W kontekście niniejszej pracy, PMF pozwala odnaleźć taki potencjał wzdłuż zadanej współrzędnej (reakcji), który reprodukuje zaobserwowany rozkład tej współrzędnej w analizowanym systemie.

Obliczanie PMF bezpośrednio z definicji jest najczęściej niepraktyczne. Wymagałoby to symulacji bez żadnych dodatkowych więzów, podczas której system zmienia swój stan, np. zachodzi proces chemiczny. Taka symulacja dostarczyłaby dużą ilość danych dotyczących minimów lokalnych, ale bardzo mało na temat stanów przejściowych (o ile w ogóle byłyby one osiągnięte, szczególnie gdy badalibyśmy mniej prawdopodobne ścieżki reakcji). Dlatego często przeprowadza się symulacje, podczas których wymusza się na danej współrzędnej, aby przyjmowała zadany zakres wartości. Jedną z technik umożliwiających realizację takiego schematu obliczeniowego jest *umbrella sampling*.[33] Do tego celu konieczne jest wygenerowanie geometrii początkowych, które reprezentują różne stadia analizowanego procesu. Każda z takich geometrii jest punktem wyjścia do symulacji dynamiki z dodatkowym potencjałem (najczęściej parabolicznym), który zapobiega odejściu zbyt daleko od stanu początkowego:

$$w_i(\xi) = \frac{K_i}{2}(\xi - \xi_i^c)^2$$

W powyższym równaniu $w_i(\xi)$ to dodatkowy potencjał nałożony na i-te okno symulacji, który jest funkcją współrzędnej reakcji. K_i to wartość stałej siłowej dla tego okna, a ξ_i^c to wartość współrzędnej reakcji, dla której potencjał osiąga minimum.

Ostatni krok powyższej procedury zakłada obliczenie profilu PMF na podstawie danych z poszczególnych okien symulacji. Do tego celu można użyć metody WHAM (*Weighted Histogram Analysis Method*).[34] W tej metodzie, na podstawie zaobserwowanego rozkładu współrzędnej reakcji oblicza się jej niezaburzoną (odpowiadającą sytuacji gdzie nie narzucono żadnych dodatkowych potencjałów) gęstość prawdopodobieństwa ($P(\xi)$). Dane wejściowe dla tej metody składają się tylko i wyłącznie z listy zaobserwowanych wartości ξ oraz danych definiujących dodatkowy potencjał (czyli wartości K_i i ξ_i^c) w danym oknie symulacji.

Do bezpośredniego uzyskania $P(\xi)$ potrzebne są wartości stałych f_j, które również zależą od $P(\xi)$, dlatego też równania WHAM są rozwiązywane iteracyjnie. Stosowane równania samouzgodnione (dla przypadku, gdy każde okno ma identyczną czasową funkcję autokorelacji) mają następującą postać:

$$P(\xi) = \frac{\Sigma_{i=1}^{N_w} h_i(\xi)}{\Sigma_{j=1}^{N_w} n_j \, exp\left[-\beta\left(w_j(\xi) - f_j\right)\right]}$$

$$exp(-\beta f_i) = \int d\xi \, exp\left(-\beta w_i(\xi)\right) P(\xi)$$

W powyższym równaniu β oznacza odwrotność iloczynu stałej Boltzmanna i temperatury, n_j to liczba wszystkich uzyskanych wartości ξ w j-tym oknie symulacji, a $h_i(\xi)$ reprezentuje rozkład prawdopodobieństwa wystąpienia współrzędnej reakcji ξ w tymże oknie, z kolei N_w to liczba okien.

Ostatni krok procedury to obliczenie potencjału średniej siły $w(\xi)$ na podstawie $P(\xi)$.

$$w(\xi) = -\beta^{-1} \ln\left(\frac{P(\xi)}{P(\xi_0)}\right)$$

Wartość ξ_0 to wybrana wartość współrzędnej reakcji, dla której przyjmujemy PMF za równy zeru, zwykle odpowiada ona minimum potencjału związanemu z substratem.

Poprawki interpolacyjne *Spline*

Ograniczenia związane z wyznaczeniem profili PMF wynikają z konieczności obliczenia tysiące razy energii i gradientu energii. Dlatego też tego rodzaju obliczenia często są możliwe jedynie z zastosowaniem metod półempirycznych w warstwie QM.

Aby poprawić uzyskane w ten sposób wyniki, można zastosować poprawki interpolacyjne *spline* (interpolacja funkcjami sklejanymi).[35] Podejście to zostało oparte o prace zespołu Thrulara.[36] W tej metodologii seria struktur uzyskanych na poziomie półempirycznym reprezentująca analizowany proces jest wykorzystana do obliczeń energii metodą QM:MM w wariancie wysokiego poziomu teorii ($E_{QMMM}^{HL}(\xi; r)$, gdzie r to pozycje atomów), oraz niskiego ($E_{QMMM}^{LL}(\xi, r)$), tj. takiego samego jak wykorzystanego w symulacjach dynamik w obliczeniach PMF. Następnie obliczana jest między nimi różnica energii $\Delta E^U(\xi)$:

$$\Delta E^U(\xi) = E_{QMMM}^{HL}(\xi; r) - E_{QMMM}^{LL}(\xi; r)$$

która zostaje wykorzystana do udokładnienia profilu PMF ($E^{LL}(\xi)$).

$$E_{corr}(\xi) = E^{LL}(\xi) + spline[\Delta E^U(\xi)]$$

Interpolacja *spline* wykorzystuje wielomiany niewielkiego stopnia (w przypadku wyników opisanych w dysertacji były to wielomiany stopnia trzeciego) do interpolacji poszczególnych fragmentów energii w funkcji współrzędnej reakcji. Uzyskana funkcja jest ciągła i zawiera wszystkie węzły interpolacji.

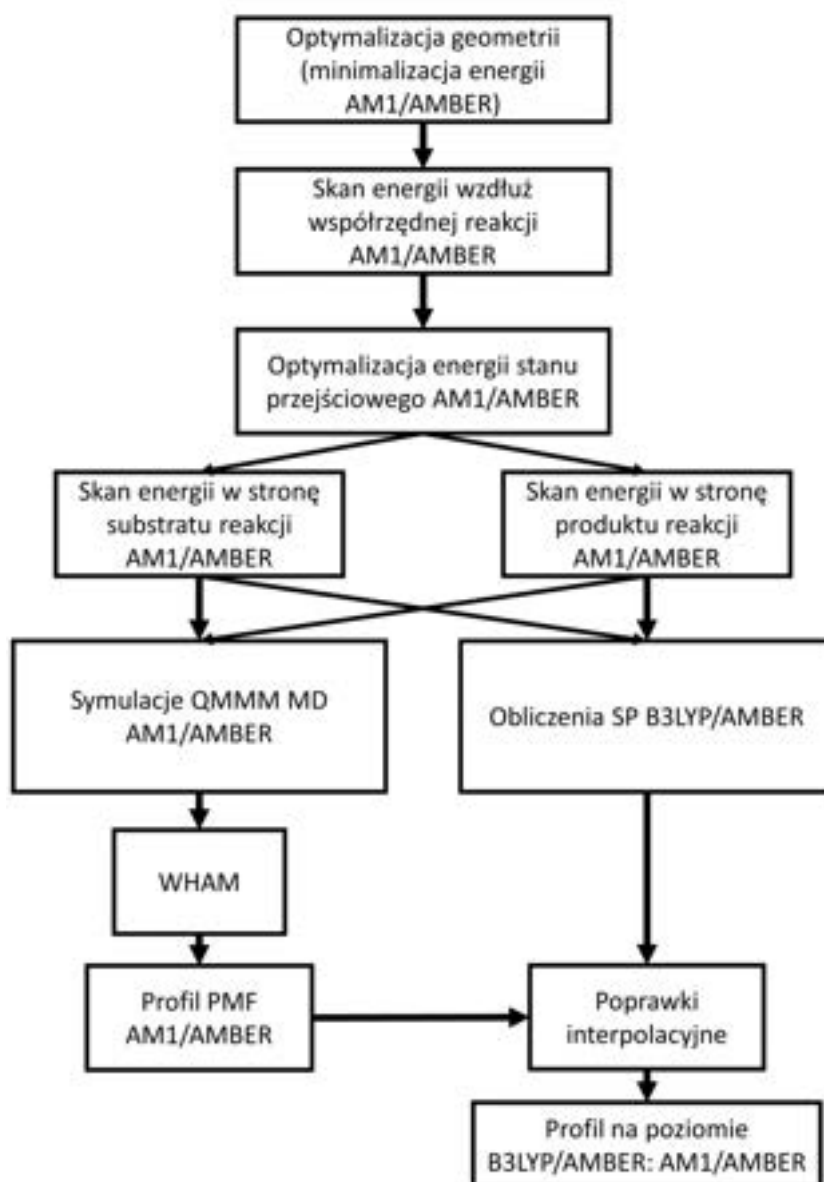## Schemat obliczeń zastosowany do modelowania enzymów Δ^1-KSTD

Poniższa sekcja prezentuje szczegóły dotyczące sposobu uzyskania profilów PMF i poprawek *spline* jaki został wykorzystany w pracach załączonych do niniejszej dysertacji (patrz Rysunek 5). Ten etap obliczeń był przeprowadzany po uzyskaniu dwuwymiarowych map energii potencjalnej i weryfikacji tego, że analizowany mechanizm może być opisany na poziomie teorii AM1.

Wszystkie prezentowane wyniki związane z obliczeniami QM:MM zostały uzyskane z wykorzystaniem biblioteki fDYNAMO.[37] Mając strukturę wyjściową do modelowania, która pochodzi z symulacji MD lub wcześniejszego modelowania QM:MM, pierwszym krokiem była optymalizacja jej geometrii na poziomie AM1/AMBER. Następnie na tym samym poziomie teorii przeprowadzano jednowymiarowy skan wzdłuż wybranej współrzędnej reakcji. Geometria odpowiadająca maksimum energii stanowiła punkt wyjścia do optymalizacji stanu przejściowego za pomocą algorytmu Bakera.[38] Ten krok obliczeń był weryfikowany przez liczbę urojonych wartości częstości drgań uzyskanych z obliczeń hesjanu. Poprawna lokalizacja stanu przejściowego była kluczowa z punktu widzenia dalszych obliczeń, ponieważ pozwalała na ustawienie jednego z okien symulacji QM:MM MD w taki sposób, aby minimum dodanego potencjału parabolicznego pokrywało się ze współrzędną bliską maksimum energii.

Z tak zlokalizowanej geometrii stanu stacjonarnego przeprowadzano następnie dwa jednowymiarowe skany energii: jeden w kierunku substratów, a drugi w kierunku produktów analizowanego procesu. Kolejnym krokiem walidacyjnym było potwierdzenie, że każdy z tych skanów zawiera lokalne minimum energii i ponadto zawiera kilka punktów za nim. Dzięki temu minima lokalne były również odpowiednio pokryte oknami symulacji QM:MM MD.

Geometrie z tych skanów oraz zlokalizowanego stanu przejściowego stanowiły struktury początkowe symulacji QM:MM MD. Wartość stałej siłowej użytej do kontroli współrzędnej reakcji wynosiła 2500 $\frac{kJ}{mol \cdot A^2}$. Każde okno obliczeń składało się z 10 ps relaksacji układu i 20 ps właściwej symulacji. Jedną z metod walidacji tych obliczeń, było sprawdzenie

czy zakresy wartości współrzędnej reakcji uzyskane dla kolejnych okien posiadają część wspólną, aby mieć pewność, że analizowany proces jest odpowiednio pokryty przez symulacje. Te same struktury, które stanowiły geometrie początkowe symulacji QM:MM MD, były następnie wykorzystane do obliczeń energii (SP) na poziomie B3LYP/6-311++G(2d,2p)/AMBER. Uzyskane wartości energii w połączeniu z profilem PMF na poziomie AMBER/AM1 pozwalały obliczyć poprawki *spline*, co skutkowało uzyskanemu PMF na poziomie B3LYP/6-311++G(2d,2p)/AMBER:AM1/AMBER.

Rysunek 5. Graficzna reprezentacja schematu obliczeniowego zastosowanego do modelowania enzymów Δ^1-KSTD.

Artykuł P1 – M. Glanowski, S. Kachhap, T. Borowski, M. Szaleniec, Model setup and procedures for prediction of enzyme reaction kinetics with QM–only and QM:MM approaches. In: Vanhaelen, Q. (eds) Computational methods for estimating the kinetic parameters of biological systems. Methods in molecular biology, vol 2385, 175–236, Humana, New York, NY 2022, ISBN 978-1-0716-1767-0

Pierwszy artykuł jest praktycznym przeglądem różnych technik obliczeniowych użytecznych w modelowaniu układów biologicznych. Nie jest on prezentacją konkretnych wyników, raczej poradnikiem, w jaki sposób prowadzić modelowanie molekularne enzymów. Zaprezentowane zostały fragmenty skryptów i komend, które należy uruchomić, aby przygotować modele do obliczeń lub przeprowadzić omawiane analizy. W ramach niniejszej dysertacji omówienie oraz sam artykuł **P1** stanowi opis metod stosowanych w pracy.

Wkład autora tej dysertacji w pracę **P1** stanowią sekcje 2.1 (*Parametrization of an Organic Molecule*), 2.2 (*Parametrization of Nonstandard Amino Acid*), 7 (*QM:MM MD Simulation*) i 8 (*Prediction of Kinetic Isotope Effects*).

Pierwszym omawianym tematem (sekcja 2) jest kwestia parametryzacji niestandardowych (niedostępnych w typowych polach siłowych) cząsteczek na potrzeby obliczeń metodami mechaniki molekularnej (MM) i symulacji dynamiki molekularnej (MD). Ten temat jest istotny z kilku powodów. Po pierwsze często się zdarza, że model jest tworzony na podstawie geometrii, która nie musi ściśle odpowiadać białku w roztworze, np. dla struktury krystalicznej. Poprawnie przeprowadzona symulacja MD powinna zbliżyć stan modelu do stanu enzymu w środowisku, w którym faktycznie katalizuje on reakcje (np. roztworu wodnego). Po drugie, w przypadku, gdy wyjściowa struktura pochodzi z dyfrakcji rentgenowskiej nie zawiera ona atomów wodoru oraz najczęściej zawiera tylko nieliczne cząsteczki wody (tzw. wodę krystalizacyjną). Posiadając parametry MM dla wszystkich cząsteczek obecnych w enzymie jesteśmy w stanie przeprowadzić optymalizację geometrii białka, która powinna m.in. odtworzyć wiązania wodorowe występujące w rzeczywistości zarówno pomiędzy resztami aminokwasowymi jak i cząsteczkami rozpuszczalników. Obliczenia prowadzone metodami mechaniki molekularnej, poza tym, że same mogą dostarczyć interesujących informacji, pozwalają przygotować modelu do dalszych analiz, które są znacznie bardziej kosztowne obliczeniowo. W rozdziale przedstawiono w jaki sposób

obliczyć cząstkowe ładunki atomów, określić ich rodzaj (tzw. typ zdefiniowany w polu siłowym) oraz nałożyć odpowiednie więzy związane z symetrią cząsteczki. Ponadto zaprezentowano jak umieścić ligand w centrum aktywnym opierając się na współrzędnych innej cząsteczki obecnej w strukturze krystalicznej. Opisywana procedura jest zaprezentowana na przykładzie dihydrotestosteronu (DHT, por. Rysunek 6, sekcja 2.1). Ta cząsteczka nie cechuje się żadną symetrią, co stanowi pewne ułatwienie w procedurze fitowania ładunków cząstkowych. Nieco bardziej skomplikowany przykład stanowi parametryzacja niestandardowego aminokwasu, w tym przypadku zdeprotonowanej tyrozyny (por. Rysunek 7, sekcja 2.2). Parametry dla tej reszty nie są obecne w standardowym polu siłowym AMBER, a jest to molekuła kluczowa dla modelowania reakcji dehydrogenaz 3-ketosteroidowych. Uzyskane parametry były wykorzystane w artykułach P2 i P3. Pierwszym krokiem była optymalizacja geometrii za pomocą DFT, następnie obliczany był potencjał elektrostatyczny wokół cząsteczki, na podstawie którego wyznaczone zostały ładunki cząstkowe Merza-Kollmana.[39]
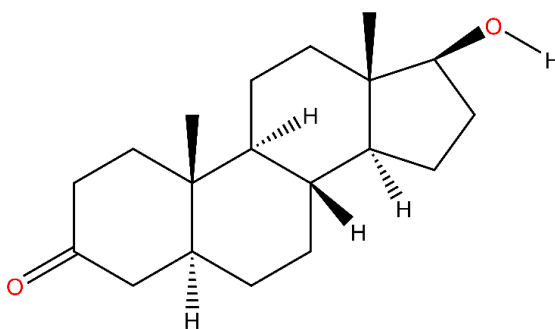
Rysunek 6. Dihydrotestosteron DHT.
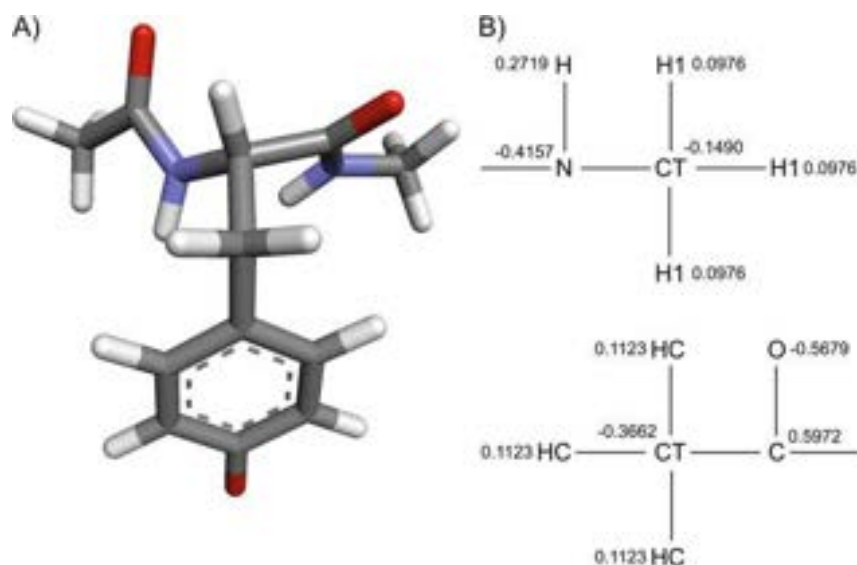
Rysunek 7. (A) Struktura wykorzystana w obliczeniach parametrów zdeprotonowanej tyrozyny.
(B) Ładunki na grupach acylowej i metyloaminowej tworzących wiązanie amidowe z tyrozyną.
*Rysunek zaadaptowano z **P1**.*

Sekcja siódma rozdziału jest poświęcona metodyce modelowania QM:MM MD. Zaprezentowano w niej jak otrzymać profil PMF na poziomie teorii AM1/AMBER. Przykładowe obliczenia zostały przeprowadzone dla pierwszego etapu dehydrogenacji DHT przez KSTD1.

Ostatnia część rozdziału opisuje podstawy obliczeń kinetycznego efektu izotopowego na podstawie różnicy entalpi swobodnej aktywacji (a właściwie różnicy poprawek wibracyjnych, ponieważ sama energia elektronowa nie zależy od mas atomów) dla cząsteczki podstawionej i niepodstawionej izotopowo.

$$KIE(\Delta G_{akt}) = e^{\frac{\Delta G_{akt}^L - \Delta G_{akt}^H}{RT}}$$

W powyższym równaniu $KIE(\Delta G_{akt})$ jest kinetycznym efektem izotopowym związanym ze zmianą w entalpii swobodnej aktywacji ze względu na podstawienie izotopowe, a ΔG_{akt}^L oraz ΔG_{akt}^H to entalpie swobodne aktywacji, odpowiednio układu niepodstawionego i podstawionego cięższym izotopem.

Pokazano również oszacowanie wpływu tunelowania na KIE z wykorzystaniem równania Wignera.[40]

$$\kappa(T) = 1 + \frac{1}{24}\left(h\frac{Im(\nu)}{k_B T}\right)^2$$

gdzie h to stała Plancka, k_B – stała Boltzmanna, T – temperatura, a ν to częstotliwość urojona pochodząca z zoptymalizowanej struktury stanu przejściowego. Wówczas KIE uwzględniający tunelowanie można zapisać jako:

$$KIE = KIE(\Delta G_{akt})\frac{\kappa^L(T)}{\kappa^H(T)} = KIE(\Delta G_{akt})KIE(\kappa)$$

gdzie $\kappa^L(T)$ i $\kappa^H(T)$ to współczynniki Wignera obliczone dla odpowiednio niepodstawionego i podstawionego izotopowo układu. Są to najbardziej podstawowe równania związane z kinetycznym efektem izotopowym, aczkolwiek w pewnych przypadkach wystarczające do zweryfikowania danego mechanizmu.

Pozostałe sekcje, których autorami są pozostali autorzy rozdziału, dotyczą obliczeń MD, klastrowego modelowania QM, modelowania QM:MM i porównaniu wartości eksperymentalnych stałych szybkości z wynikami obliczeniowymi. W sekcjach tych przedstawiono protokół związany z przygotowaniem, uruchomieniem i analizą symulacji MD. Wszystkie obliczenia dynamiki molekularnej opisane w pozostałych artykułach tworzących dysertację, zostały przeprowadzone zgodnie z tą metodologią. Przygotowanie modelu oznacza minimalizację jego energii i stopniowe ogrzewanie do temperatury zadanej w symulacji MD. W tej sekcji znajduje się również opis metod użytych do wybrania reprezentatywnej geometrii z całej symulacji. Ten krok jest nietrywialny i wymaga analiz zmian parametrów geometrycznych w czasie symulacji oraz zastosowania analizy skupień do agregacji podobnych geometrii. W kolejnym podrozdziale opisano metodę modelowania QM w oparciu o modele klastrowe QM. Polega ona na wybraniu odpowiedniego fragmentu enzymu, zamrożeniu pozycji wyselekcjonowanych atomów na skrajach klastra, a następnie potraktowaniu całości jedną metodą kwantowo-mechaniczną, np. DFT. Wady tej metody to problem z uwzględnieniem otoczenia reakcji. Pewne poprawki są osiągalne dzięki wykorzystaniu modeli ciągłych polaryzowalnych ośrodków. Ponadto ze względu na możliwości współczesnych komputerów i koszt obliczeniowy metod kwantowo-chemicznych rozmiar studiowanych systemów jest ograniczony. Zaprezentowano również w jaki sposób porównywać zmierzony z eksperymentu k_{cat} z obliczoną wartością bariery energetycznej dla różnych substratów. Skomentowano także kiedy taka analiza ma sens i zamieszczono praktyczne obserwacje.

Wyniki

Artykuł P2 – P. Wójcik, M. Glanowski, A. M. Wojtkiewicz, A. Rohman, M. Szaleniec, Universal Capability of 3-Ketosteroid Δ1-Dehydrogenases to Catalyze Δ1-Dehydrogenation of C17-Substituted Steroids, *Microbial Cell Factories*, 2021, 20(1), article number: 119

Wkładem autora dysertacji do pracy **P2** było udowodnienie na drodze symulacji komputerowych, iż KSTD1 może katalizować reakcje odwodornienia 3-ketosteroidów o rozbudowanych podstawnikach, przeprowadzenie symulacji ilościowo potwierdzających tę hipotezę, analiza wyników z nich uzyskanych oraz dyskusja rezultatów badań w odniesieniu do wyników eksperymentalnych. Wzory strukturalne wszystkich steroidów, które były przedmiotem modelowania w całej pracy doktorskiej, są widoczne na Rysunku 8.
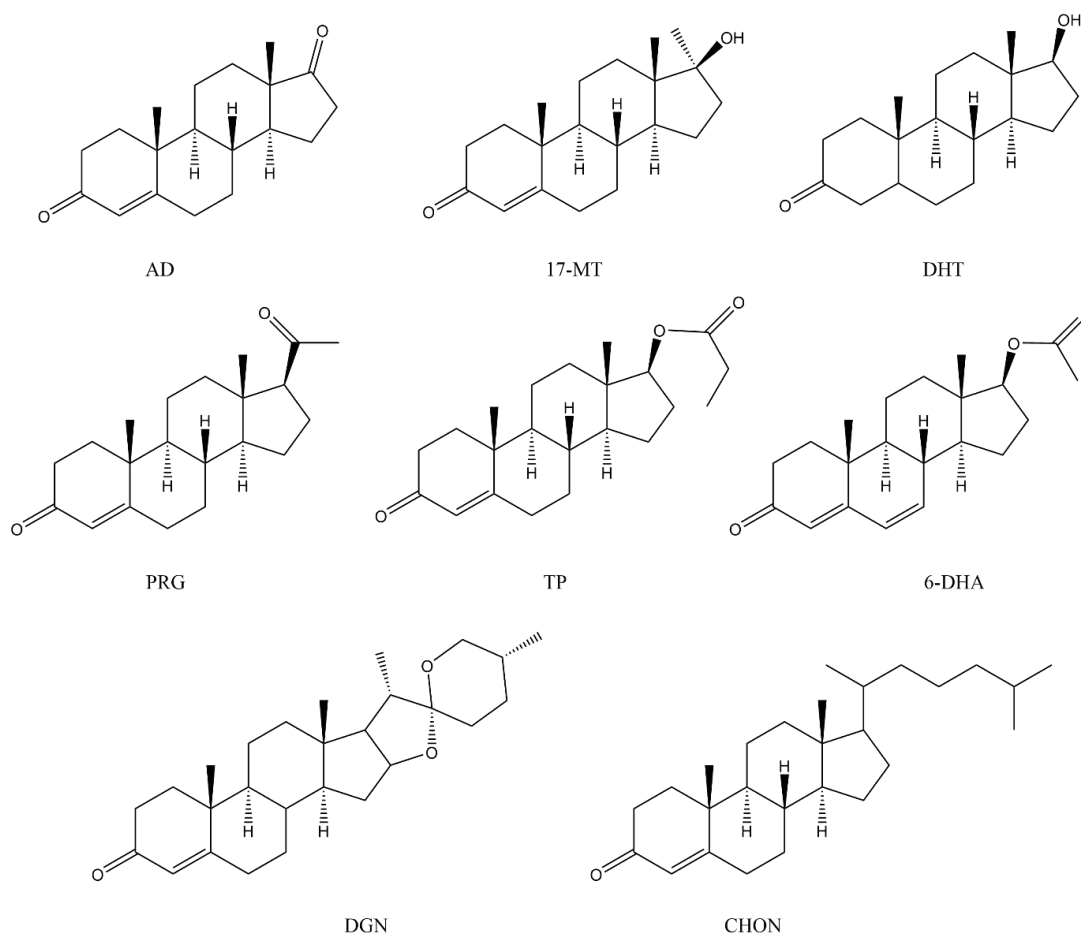
Rysunek 8. Struktury steroidów wykorzystane w symulacjach opisanych w dysertacji: androst-4-en-3-on (AD), 17-metylotestosteron (17-MT), androstanolon (DHT), progesteron (PRG),

propanian testosteronu (TP), octan 6-dehydrotestosteronu (6-DHA), diosgenon (DGN), cholest-4-en-3-on (CHON).

Jednym z pierwszych pytań badawczych, które pojawiły się przy podejściu do analizy Δ^1-KSTD z *R. erythropolis*, była kwestia specyficzności substratowej KSTD, tj. wiązania różnych substratów do centrum aktywnego oraz szybkości ich odwodornienia przez enzym (tzw. aktywności). W literaturze postulowano, że AcmB jest unikatowym enzymem, jeśli chodzi o spektrum substratowe, ze względu na jego zdolność do katalizowania odwodornienia substratów o rozbudowanym podstawniku przy atomie C17 (np. cholest-4-en-3-on, CHON).[41] Unikatowość AcmB zdawały się potwierdzać wyniki eksperymentalne wykazujące brak aktywności enzymu KSTD1 z *R. erythropolis* z takimi substratami oraz liczne publikacje wykazujące aktywność enzymów z klasy Δ^1-KSTD z 3-ketosteroidami pozbawionymi dużych podstawników przy C17.[7,13–15,18–20,42,43]

Wstępne badania teoretyczne przeprowadzone za pomocą dokowania substratów 3-ketosteroidowych w centrum aktywnym KSTD1 wykazały, że zasadniczo nie ma przeszkód uniemożliwiających katalitycznie aktywne wiązanie CHON. Rozbudowany podstawnik wystawał bowiem z centrum aktywnego do rozpuszczalnika, zaś w centrum aktywnym wiązany był przede wszystkim układ pierścieni steroidowych. Zasadniczo więc, fakt braku reakcji KSTD1 z CHON wydawał się nieuzasadniony.

Problem ten był eksplorowany za pomocą symulacji MD oraz obliczeń MMPBSA dla kompleksów E:S KSTD1 z różnymi 3-ketosteroidami o zróżnicowanym stopniu rozbudowania podstawnika przy C17.[44] Do tych ostatnich obliczeń wykorzystano 5 ns z każdej trajektorii, które cechowały się największą stabilnością parametrów geometrycznych opisujących związanie substratu w centrum aktywnym. Uzyskane entalpie swobodne wiązania (ΔG_b) różnych ketosteroidów (Rysunek 8) zostały zebrane w Tabeli 1. Okazało się, że podstawniki w pozycji C17 nie wpływają znacząco na entalpię swobodą wiązania, nawet w przypadku dużych grup takich jak podstawnik izooktylowy w CHON czy dodatkowe dwa pierścienie w diosgenonie (DGN)). Wszystkie obliczone ΔG_b zawierają się w przedziale $-6{,}17\pm0{,}15$ do $-10{,}97\pm0{,}15$ kcal/mol. Oszacowana ΔG_b CHON wyniosła $-8{,}40\pm0{,}19$ kcal/mol, dla DGN $-6{,}17\pm0{,}15$ kcal/mol, podczas gdy dla AD $-8{,}02\pm0{,}13$ kcal/mol. Analiza trajektorii z symulacji MD wyjaśnia te obserwacje - podstawniki przy pozycji C17 nie znajdują się w bezpośrednim otoczeniu aminokwasów enzymu, lecz wystają z centrum aktywnego do rozpuszczalnika (por.

Rysunek 9). Wyniki te potwierdziły wstępną hipotezę, że wiązanie substratu do centrum aktywnego nie powinno być przeszkodą dla KSTD1 w odwodornieniu cholest-4-en-3-onu.

W celu przetestowania postawionej przeze mnie hipotezy wynikającej z obliczeń, przeprowadzone zostały testy katalityczne. Jak dowiodły wyniki eksperymentalne przedstawione w pracy **P2**, KSTD1 z *R. erythropolis* faktycznie jest aktywny w reakcji odwodornienia CHON, co jest bardzo ważnym odkryciem mającym konsekwencje dla wszystkich enzymów podobnych do KSTD1. Okazało się, że sposób prowadzenia eksperymentów kinetycznych przez odkrywców KSTD1 (rozpuszczanie steroidów w izopropanolu) uniemożliwiał wykrycie aktywności z CHON ze względu na jego bardzo niską rozpuszczalność w wodzie. Dopiero zastosowanie solubilizatora (tj. izopropylowanej β-cyklodekstryny, HBC) umożliwiło eksperymentalne wykrycie takiej aktywności.

Tabela 1. Obliczone energie wiązania dla poszczególnych ketosteroidów

Ligand	$\Delta G_{\text{wiązania}}$ [kcal/mol]	Błąd standardowy $\Delta G_{\text{wiązania}}$ [kcal/mol]
Androst-4-en-3-on	−8,02	0,13
Cholest-4-en-3-on	−8,40	0,19
Diosgenon	−6,17	0,15
Androstanolon	−10,97	0,15
Progesteron	−8,23	0,17
Propanian testosteronu	−7,85	0,15
Octan 6-dehydrotestosteronu	−8,53	0,15
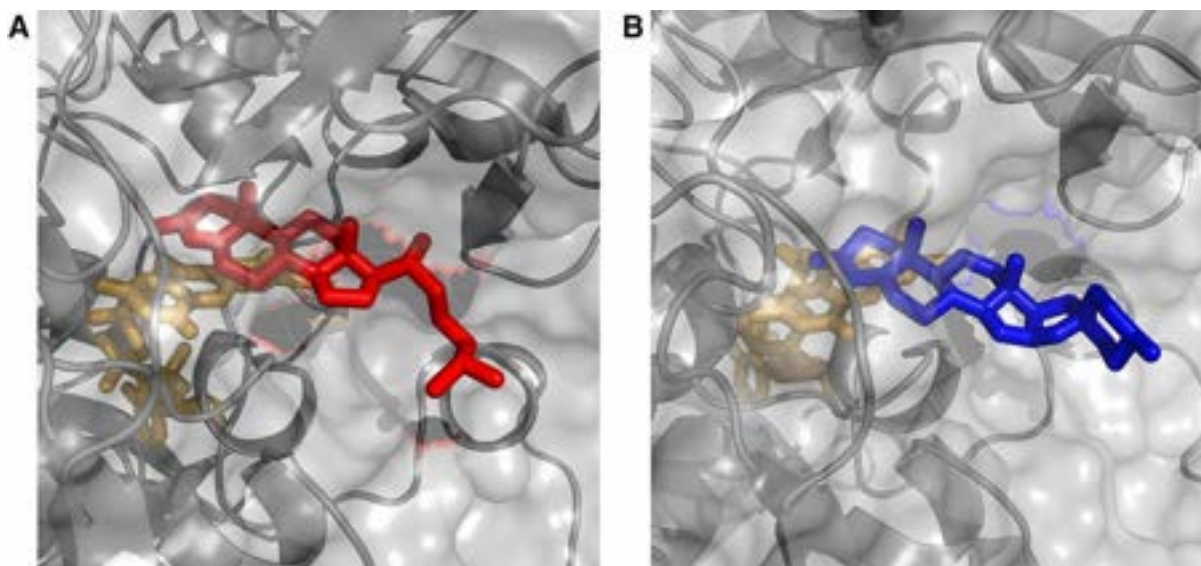
Rysunek 9. Przykładowe klatki z symulacji MD prezentujące (A) cholest-4-en-3-on (czerwony) i (B) diosgenon (niebieski) związany w centrum aktywnym KSTD1. Na pomarańczowo zaznaczono FAD. Rysunek zaadaptowano z P2.

Artykuł **P2** przeprowadza dyskusję w jaki sposób KSTD1 i AcmB są zdolne do dehydrogenacji substratów podstawionych w pozycji C17. Z punktu widzenia eksperymentalnego, wyznaczenie aktywności enzymu dla CHON jest nietrywialne ze względu na jego niewielką rozpuszczalność w wodzie. Aby poradzić sobie z tym problemem, do roztworów dodawano HBC jako solubilizatora. Dzięki temu możliwe było przeprowadzenie pomiarów, ale jednocześnie analiza wyników stała się bardziej złożona. W takiej mieszaninie reakcyjnej tworzą się kompleksy HBC:steroid w stosunku 1:1 i 1:2. Artykuł **P2** opisuje procedurę pozwalającą uzyskać dwie stałe równowagi związane z tworzeniem kompleksów, oraz stężenie wolnego substratu. Z tą wiedzą możliwe było obliczenie stężenia kompleksów z HBC. Okazało się, że AD występuje w roztworze głównie (99%) jako kompleks 1:1, natomiast w przypadku CHON udział procentowy kompleksów 1:1 i 1:2 jest znaczący (odpowiednio 53,52% oraz 43,55%). Oznacza to, że kompleks enzymu z substratem powstaje w rzeczywistości na skutek interakcji ze steroidem związanym z HBC a nie z wolnym steroidem w roztworze co również wpływa na $\Delta G_{wiązania}$ a przez to na stałą K_m.

Innym istotnym aspektem artykułu **P2** jest kwestia natywnego substratu AcmB. Głównym argumentem użytym w dyskusji jest wartość k_{cat}/K_m otrzymana dla AD i CHON. Co ciekawe, w przypadku AcmB, ta wartość jest wyższa dla drugiego z tych ketosteroidów, co sugeruje, że to CHON może być substratem natywnym. Analogiczne dane dla KSTD1 pokazują preferencje dla AD.

Najważniejsze wnioski z pracy **P2**:

- KSTD1 jest aktywne w reakcji 1,2-odwodornienia CHON oraz DGN,
- Oszacowane ΔG wiązania różnych ketosteroidów do centrum aktywnego KSTD1 mieszczą się w wąskim zakresie; różne podstawniki w pozycji C17 tylko w niewielkim stopniu wpływają na wartość ΔG.

Artykuł P3 – M. Glanowski, P. Wójcik, M. Procner, T. Borowski, D. Lupa, P. Mielczarek, M. Oszajca, K. Świderek, V. Moliner, A. J. Bojarski, M. Szaleniec, Enzymatic Δ1-dehydrogenation of 3-ketosteroids – reconciliation of kinetic isotope effects with the reaction mechanism, *ACS Catal.*, **2021, 11, 8211−8225**

Publikacja **P3** stanowi najważniejszą pracę przedstawionego cyklu gdyż prezentuje wyniki dotyczące głównego celu badań, tj. weryfikację przedstawionej w literaturze hipotezy mechanistycznej 1,2-odwodornienia katalizowanego przez Δ^1-dehydrogenazy 3-ketosteroidowe. W artykule **P3** zaprezentowano szczegółową dyskusję mechanizmu reakcji dehydrogenacji 3-ketosteroidów przez KSTD1. Łączy on modelowanie teoretyczne oraz wyczerpującą interpretację danych eksperymentalnych. Wkład autora dysertacji polegał na przeprowadzeniu wszystkich symulacji komputerowych i ich analizy, interpretacji wyników KIE i ich dyskusji z wynikami eksperymentalnymi. Przedstawione interpretacje KIE w oparciu o analizę wyników teoretycznych umożliwiły pogodzenie pozornie sprzecznych wyników eksperymentalnych sprzed 50 lat z mechanizmem reakcji opartym o dane strukturalne i przeprowadzone przeze mnie obliczenia teoretyczne.

W pracy **P3** doświadczalnie dowiedziono, że reakcja katalizowana przez enzym KSTD1 zachodzi zgodnie z niesekwencyjnym mechanizmem Ping-Pong bi-bi. Inne rozważane możliwości to mechanizm sekwencyjny uporządkowany oraz sekwencyjny przypadkowy, który zakłada, że oba reagenty muszą związać się w centrum aktywnym by zaszła reakcja, zostały wykluczone jako mniej prawdopodobne. Wynik ten, choć spodziewany dzięki pracom Itagakiego z połowy lat '90[11] oraz możliwości przeprowadzenia reakcji połówkowych, był podstawą do dalszych rozważań kinetycznych.

Wyniki eksperymentów kinetycznych w stanie przedstacjonarnym dla RHR zostały wykorzystane do globalnego, numerycznego dopasowania wartości stałych kinetycznych. Zastosowany model składał się z dwóch, odwracalnych etapów (wiązania cząsteczki przez enzym i procesu katalitycznego odwodornienia). Dopasowanie bardziej złożonego opisu reakcji, np. rozważenie dwuetapowego kroku katalitycznego okazało się niemożliwe ze względu na brak wystarczającej informacji w danych doświadczalnych. Możliwe było numeryczne dopasowanie wartości stałych szybkości tak aby zminimalizować błąd średniokwadratowy, ale nie wszystkie uzyskane w ten sposób parametry były wyznaczone

jednoznacznie. Swiadczy to o tym, że model zawiera zbyt dużo parametrów w stosunku do dostępnych danych. W takiej sytuacji nie było możliwe, aby zdeterminować wartości wszystkich stałych.

Model do obliczeń mechanizmu reakcji metodami QM:MM MD został otrzymany ze struktury krystalicznej KSTD1 z *R. erythropolis* (PDB 4c3y).[4] W pierwszym etapie enzym został przygotowany do modelowania poprzez relaksację geometrii za pomocą symulacji MD, z tej występującej w krysztale do tej bliższej białku znajdującemu się w roztworze w optymalnym pH dla reakcji (8). Struktura do dalszych obliczeń została wybrana na podstawie analizy skupień stabilnej trajektorii z dynamiki molekularnej. Jako współrzędną użyto pierwiastka z odchylenia średniokwadratowego pozycji ciężkich atomów kluczowych reszt w centrum aktywnym (RMSD). Ostatecznie wybrano strukturę, która leżała najbliżej centroidu najliczniejszego klastra. Następnie przeprowadzono symulacje dynamiki QM:MM MD. Warstwa QM składała się z ketosteroidu, reszty Tyr318 i fragmentu FAD (Rysunek 10).
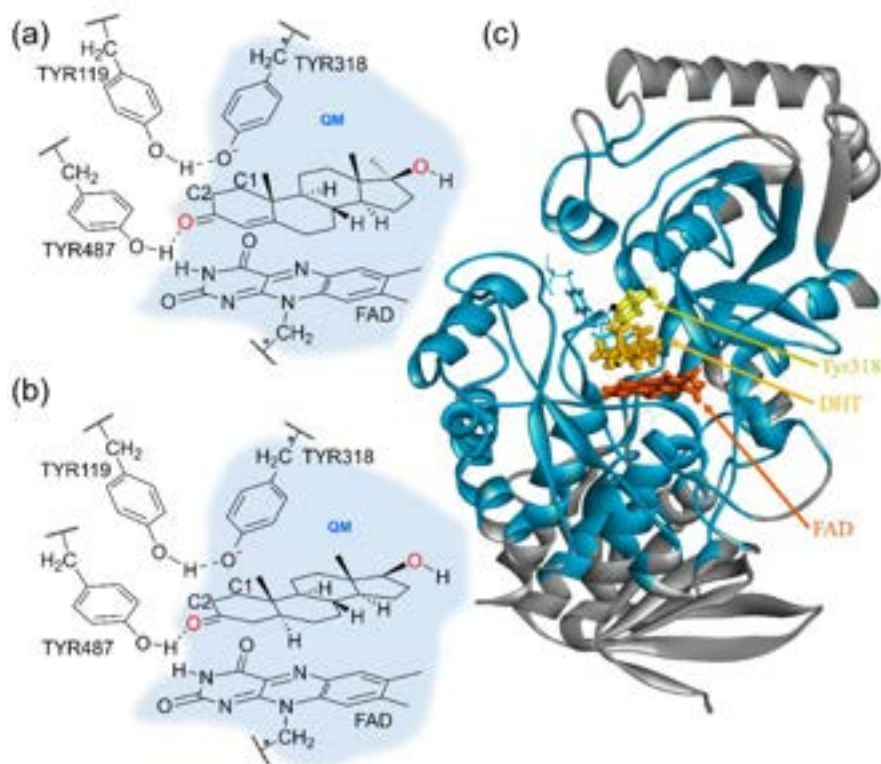
*Rysunek 10. Schematyczna reprezentacja centrum aktywnego z substratami: 17MT (a) oraz DHT (b). Warstwa QM została zaznaczona jasnoniebieskim obszarem. (c) Poglądowy rysunek modelu z zadokowanym DHT. Szare fragmenty białka były nieruchome w czasie symulacji. Rysunek zaadaptowano z **P3**.*

Na potrzeby analizy mechanizmu zdefiniowano dwie współrzędne reakcji ξ_1 i ξ_2, jako różnice dystansów pomiędzy ciężkimi atomami a przenoszonym atomem wodoru, (Rysunek 11) W rezultacie współrzędne ξ_1 i ξ_2 odpowiadały transferowi atomów wodoru, odpowiednio, z C2β na Tyr318 i z C1α na atom N5 w FAD.

$$\xi_1 = D_1 - D_2$$
$$\xi_2 = D_3 - D_4$$
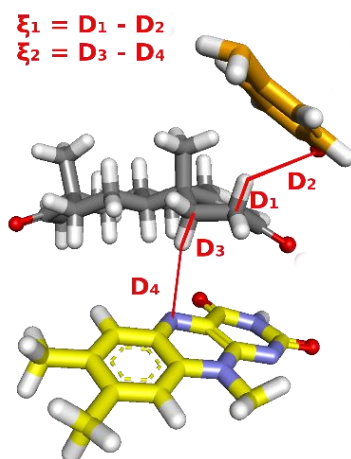
Rysunek 11. Współrzędne reakcji opisujące Δ¹-dehydrogenację 3-ketosteroidów.

Możliwe mechanizmy dla tak zadanego układu zbadano dwuwymiarowym skanem energii na poziomie AM1/AMBER. Dla zoptymalizowanych w ten sposób struktur przeprowadzono następnie obliczenia Single Point Energy (SP) na poziomie B3LYP/6-31G(d,p)/AMBER co pozwoliło na otrzymanie mapy widocznej na Rysunku 12.
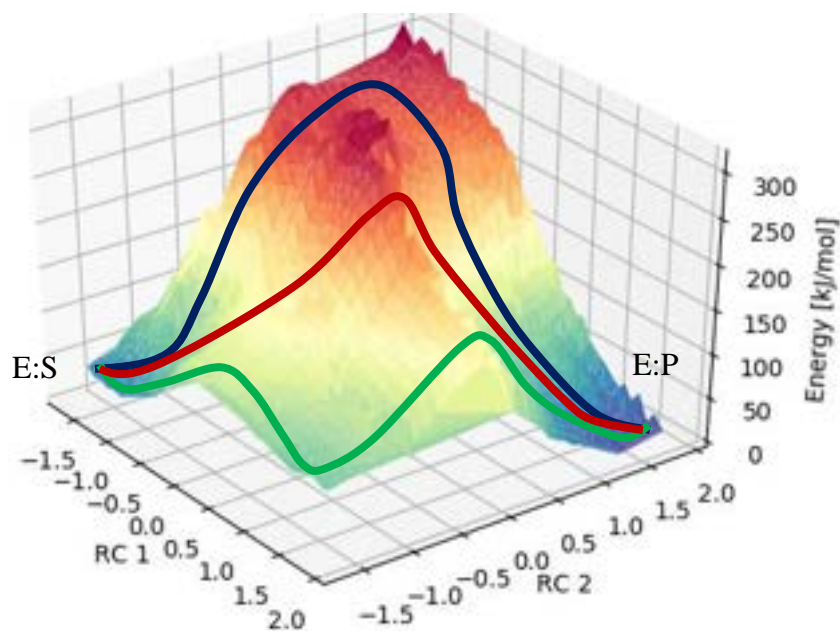
Rysunek 12. Wynik dwuwymiarowego skanu wzdłuż zdefiniowanych współrzędnych reakcji (RC1 – ξ_1 i RC2 – ξ_2). Geometrie zoptymalizowane na poziomie AM1/AMBER, energie

pochodzą z obliczeń Single Point Energy (SP) na poziomie teorii B3LYP/6-31G(d,p)/AMBER. Obszar odpowiadający kompleksowi enzym:substrat oznaczono jako E:S, a kompleksowi enzym:produkt jako E:P.Krzywe zielona, czerwona i granatowa reprezentują odpowiednio mechanizmy: proponowany w literaturze, jednoetapowy oraz zaczynający się z pozycji C1α.

Otrzymane rezultaty jasno sugerują, że nie istnieje struktura stanu przejściowego odpowiadająca równoczesnemu transferowi atomów wodoru z pozycji C2β i C1α (brak punktu siodłowego wzdłuż czerwonej krzywej na Rysunku 12). Z kolei transfer jonu hydroniowego w pierwszej kolejności z pozycji C1α, a dopiero później protonu z C2β (krzywa granatowa) cechuje się znacznie wyższą barierą niż mechanizm, w którym przeniesienie atomów wodoru następuje w odwrotnej kolejności (krzywa zielona, najpierw proton 2β, później jon hydroniowy 1α). W rezultacie ten klasyczny mechanizm został poddany dalszemu modelowaniu. Seria dynamik QM:MM MD pozwoliła na obliczenie profilu PMF na poziomie AM1/AMBER. Wykorzystanie metody AM1 w warstwie QM było uzasadnione ze względu na szybkość obliczeń, konieczną, aby przeprowadzić symulacje QM:MM MD o odpowiedniej długości w dostępnym czasie. Ponadto wcześniej zweryfikowano czy w metodzie AM1 zachowany jest charaktery chemiczny mechanizmu reakcji jaki był obserwowany w obliczeniach prowadzonych metodą DFT. Ostatecznie profil PMF został udokładniony poprawkami interpolacyjnymi z wykorzystaniem obliczeń SP na poziomie B3LYP/6-311++G(2d,2p)/AMBER (Rysunek 13).
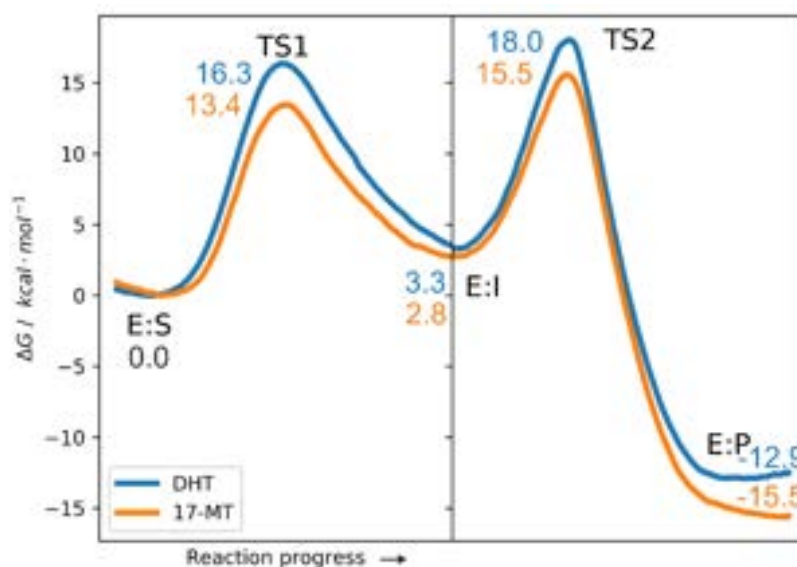
*Rysunek 13. Profil PMF reakcji odwodornienia udokładniony poprawkami spline dla 17MT (linia pomarańczowa) oraz DHT (linia niebieska). Wyniki uzyskane na poziomie teorii B3LYP/6-311++G(2d,2p):AM1/AMBER. Rysunek zaadaptowano z **P3**.*

Uzyskane profile są podobne dla obydwu analizowanych substratów. W przypadku 17-metylotestosteronu (17MT) energie swobodne TS1 i TS2 wynoszą odpowiednio: 13,4 oraz 15,5 kcal/mol, natomiast dla DHT wartości te są równe 16,3 oraz 18,0 kcal/mol. Energia swobodna drugiego stanu przejściowego jest tylko nieco wyższa od pierwszego (1,7–2,1 kcal/mol), co ma istotne konsekwencje dla kinetycznego efektu izotopowego. Wybrane struktury z trajektorii QM:MM MD zostały zoptymalizowane na poziomie teorii B3LYP/6-31G(d,p)/AMBER, tak aby zlokalizować wszystkie stany stacjonarne związane z badanym mechanizmem. Reprezentatywne struktury dla DHT są widoczne na Rysunku 14. Liczba drgań urojonych służyła potwierdzeniu czy dana geometria jest stanem przejściowym czy minimum energii.
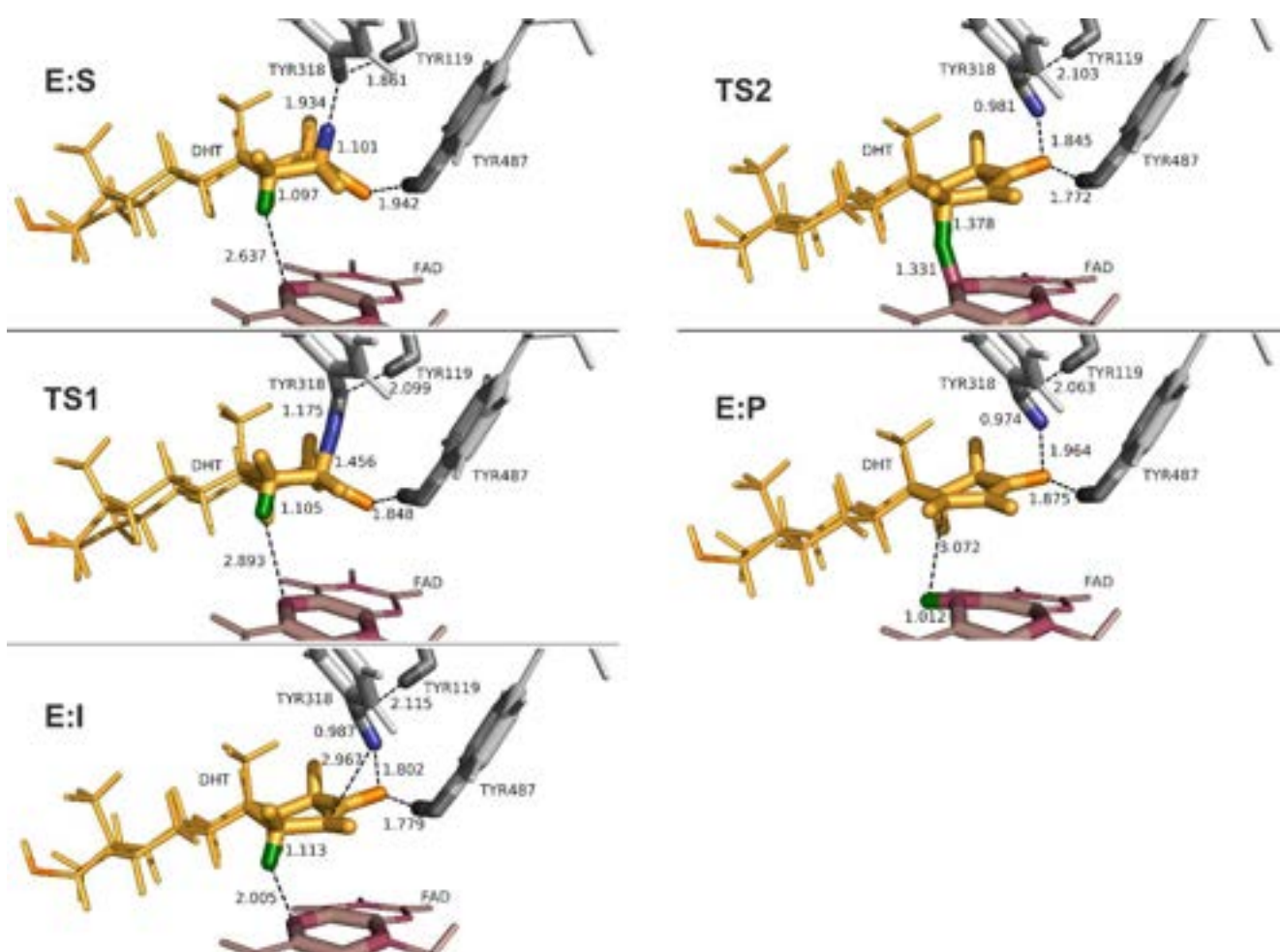
Rysunek 14 Reprezentatywne struktury stanów stacjonarnych dla reakcji odwodornienia otrzymanych dla DHT. Przenoszone atomy zaznaczono kolorem zielonym i niebieskim. Rysunek zaadaptowano z P3.

Proces RHR może zostać przedstawiony za pomocą następującego równania:

$$E + S \underset{k_{-1}}{\overset{k_1}{\leftrightarrow}} E{:}S \overset{k_2}{\rightarrow} E{:}P$$
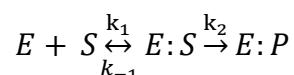
Przeprowadzono pomiary kinetycznego lepkościowego efektu rozpuszczalnikowego (KSVE) z różnymi stężeniami PEG (wpływającym na makroplepkość) oraz gliceryny (modyfikującej mikrolepkość). Wykazano silną zależność zarówno stałej kinetycznej k_2 jak i k_2/K_D (gdzie K_D jest stałą równowagi rozpadu kompleksu E:S) w funkcji mikrolepkości i minimalną zależność tych wielkości od makrolepkości. Można stąd wnioskować, że zarówno wiązanie substratu jak i proces katalityczny (k_2) wewnątrz centrum aktywnego odbywa się bez istotnych zmian konformacyjnych białka, co jest zgodne z tym co zaobserwowano w symulacjach MD i QM:MM MD. Zależność od mikrolepkości sugeruje z kolei, że dyfuzja reagentów w roztworze może limitować wiązanie substratu do centrum aktywnego.

Ostateczną weryfikacją poprawności proponowanego mechanizmu miała być konfrontacja wartości teoretycznych z eksperymentalnymi wynikami pomiarów kinetycznych uzyskanych dla deuterowanych substratów. Wartości KIE zostały obliczone dla deuterowanych cząsteczek 17MT i DHT. Analogicznie obliczono wartości BIE, korzystając z zoptymalizowanych struktur E:S i ketosteroidów w wodzie. Do tych ostatnich obliczeń warstwa QM została ograniczona do cząsteczki steroidu.

Wartości KIE zostały obliczone dla każdego rozważanego etapu elementarnego. Dla 17MT (deuterowanego m.in. w pozycji C2β i C2α) przewidziano względnie wysoki normalny, pierwszorzędowy efekt dla pierwszego etapu reakcji (5,39±0,04) i odwrotny, drugorzędowy KIE dla etapu drugiego (0,92±0,02). W przypadku DHT (podstawionego deuterem m.in. w pozycji C1α) sytuacja była podobna, dla pierwszego etapu RHR obliczony KIE był równy 1,12±0,01 (efekt drugorzędowy), a dla drugiego 4,95±0,05 (efekt pierwszorzędowy).

W przypadku bezpośrednich pomiarów KIE w stanie przedstacjonarnym dla RHR pomiary dostarczyły wartości KIE odpowiednio 1,5±0,04 i 3,5±0,04 dla 2,2,4,6,6-d_5-17-MT i 1,16,16,17-d_4-DHT. Pozornie te wyniki stoją w sprzeczności z wartościami teoretycznymi. Gdyby szybkość reakcji była całkowicie kontrolowana przez przeniesienie jonu wodorkowego na FAD, powinno zaobserwować się mierzalne KIE tylko dla DHT, ale nie dla 17MT. Gdyby z kolei to abstrakcja protonu w pierwszym etapie byłaby limitująca, wówczas istotne wartości KIE powinny być obserwowalne tylko dla 17MT. Tymczasem uzyskano wartości KIE istotnie

różne od jedności dla podstawienia zarówno w pozycji C2β jak i C1α. Wyniki te były podobne do tych, które zmierzono w stanie stacjonarnym już w latach 60'.[9]

Należy zwrócić uwagę, że wartość KIE dla całego procesu dąży do wartości KIE jednego etapu elementarnego tylko wtedy, gdy jest on dostatecznie wolny. Na potrzeby dalszych rozważań warto rozłożyć proces opisywany przez stała kinetyczną k_2 na etapy elementarne zgodne z mechanizmem reakcji:

$$E{:}S \underset{k_{I-S}}{\overset{k_{S-I}}{\longleftrightarrow}} E{:}I \overset{k_{I-P}}{\longrightarrow} E{:}P$$
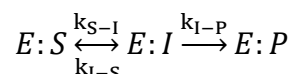
Aby ilościowo odpowiedzieć na pytanie, jakie różnice w konkretnych wartościach stałych szybkości są dostateczne duże, tak aby tylko jeden etap mógł kontrolować cały proces, można przeanalizować równanie 10 z **P3**:

$$k_2 = \frac{k_{S-I}k_{I-P}}{k_{S-I} + k_{I-S} + k_{I-P}}$$

To wyrażenie łączy wartości stałych szybkości procesów elementarnych kroku katalitycznego zachodzącego w KSTD1 z wypadkową stałą szybkości, mierzalna w eksperymencie stopped-flow. Stąd można wyprowadzić zależność pomiędzy wypadkowym KIE (odnoszącym się do k_2) a stałymi szybkości procesów elementarnych i KIE procesów elementarnych (równanie 12 z **P3**):

$$KIE_2 = KIE_{S-I}KIE_{I-P} \frac{\dfrac{1}{KIE_{S-I}} + \dfrac{1}{KIE_{I-S}}\dfrac{k_{I-S}^H}{k_{S-I}^H} + \dfrac{1}{KIE_{I-P}}\dfrac{k_{I-P}^H}{k_{S-I}^H}}{1 + \dfrac{k_{I-S}^H}{k_{S-I}^H} + \dfrac{k_{I-P}^H}{k_{S-I}^H}}$$

gdzie KIE_i oznacza $\frac{k_i^H}{k_i^D}$, czyli stosunek stałych szybkości odpowiadających niedeuterowanym (k_i^H) i deuterowanym reagentom (k_i^D), indeks i może przyjmować wartości S-I, I-S i I-P.

Niestety oszacowanie wartości stałych kinetycznych na podstawie barier wiąże się z dość dużym błędem. Różnica jednej kcal/mol powoduje około pięciokrotne przeszacowanie lub niedoszacowanie danej stałej. Dlatego w tej analizie należy mieć na uwadze dokładność dostępnych metod obliczeniowych. Z tego powodu do oceny uzyskanych wyników zostały wykreślone mapy widoczne na Rysunku 15. Zakładając, że obliczone wartości KIE dla procesów elementarnych są dokładne, można obliczyć wypadkowe KIE dla różnych wartości

barier poszczególnych etapów. Wartości eksperymentalne zostały zaznaczone czarnymi liniami. Środki kół na wykresach odpowiadają wartości wypadkowego KIE obliczonego z wykorzystaniem uzyskanych barier. Przyjmując dokładność metody równą 1 kcal/mol otrzymuje się obszar dopuszczalnych wartości KIE mieszczących się w obrębie koła. Zarówno dla 17MT jak i DHT koła są przecięte przez krzywą wartości eksperymentalnych. Wynik ten wykazuje, że obliczone wartości KIE są zgodne z wartościami zmierzonymi w eksperymencie z dokładnością do metody obliczeniowej.
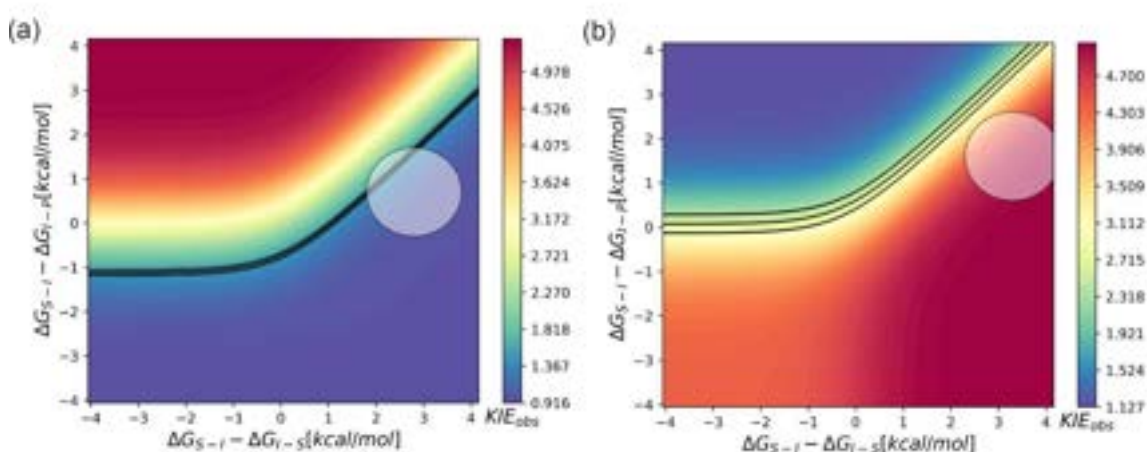
*Rysunek 15 Mapy wartości KIE dla 17MT (a) i DHT(b). Czarne linie odpowiadają wartościom eksperymentalnym. Zaadaptowano z **P3***.

Obliczone wartości BIE wyniosły 0,97±0,06 i 0,76±0,05 odpowiednio dla DHT i 17MT. Te wyniki sugerują, że szczególnie 17MT wiąże się lepiej z białkiem w wariancie zdeuterowanym. Wyniki z kompetycyjnego pomiaru KIE, które interpretujemy jako kinetyczny efekt izotopowy dla stosunku V/K, wyniosły 2,2±0,02 (17MT) i 2,4±0,07 (DHT). W porównaniu do pomiarów bezpośrednich uzyskano więc wyższą wartość dla 17MT i niższą dla DHT, co oznacza, że wartości obliczone są w jakościowej zgodności tylko dla DHT.

Należy zauważyć, że eksperymentalne wartości zostały uzyskane w roztworach zawierających dodatek rozpuszczalników organicznych, bez których osiągnięcie wysycającego enzym stężenie ketosteroidów było nieosiągalne. W symulacjach natomiast, 17MT i DHT znajdowały się w centrum aktywnym białka lub czystej wodzie. Jest to istotny czynnik, który szczególnie utrudnia liczbowe porównanie wartości BIE. Ponadto w przypadku 17MT wszystkie podstawienia izotopowe (5 pozycji) mają miejsce w pierścieniach A i B steroidu, które są otoczone przez białko w kompleksie E:S. Natomiast w cząsteczce DHT, tylko jeden atom deuteru znajduje się w pierścieniu A, a pozostałe trzy w pierścieniu D, który w dużej

mierze wystaje z centrum aktywnego i w związku z tym, jest otoczony przez cząsteczki wody w kompleksie E:S. To tłumaczy dlaczego, obliczone wartości BIE odbiegają od jedności bardziej dla 17MT niż DHT.

Rozważono również wariant ścieżki reakcji, w której Tyr318 nie jest zdysocjowana na początku reakcji. Mechanizm ten był szczególnie istotny wobec eksperymentalnie potwierdzonej aktywności AcmB w reakcji odwodornienia w pH 6,0. Schemat potencjalnego alternatywnego mechanizmu przedstawiony jest na Rysunku 16.
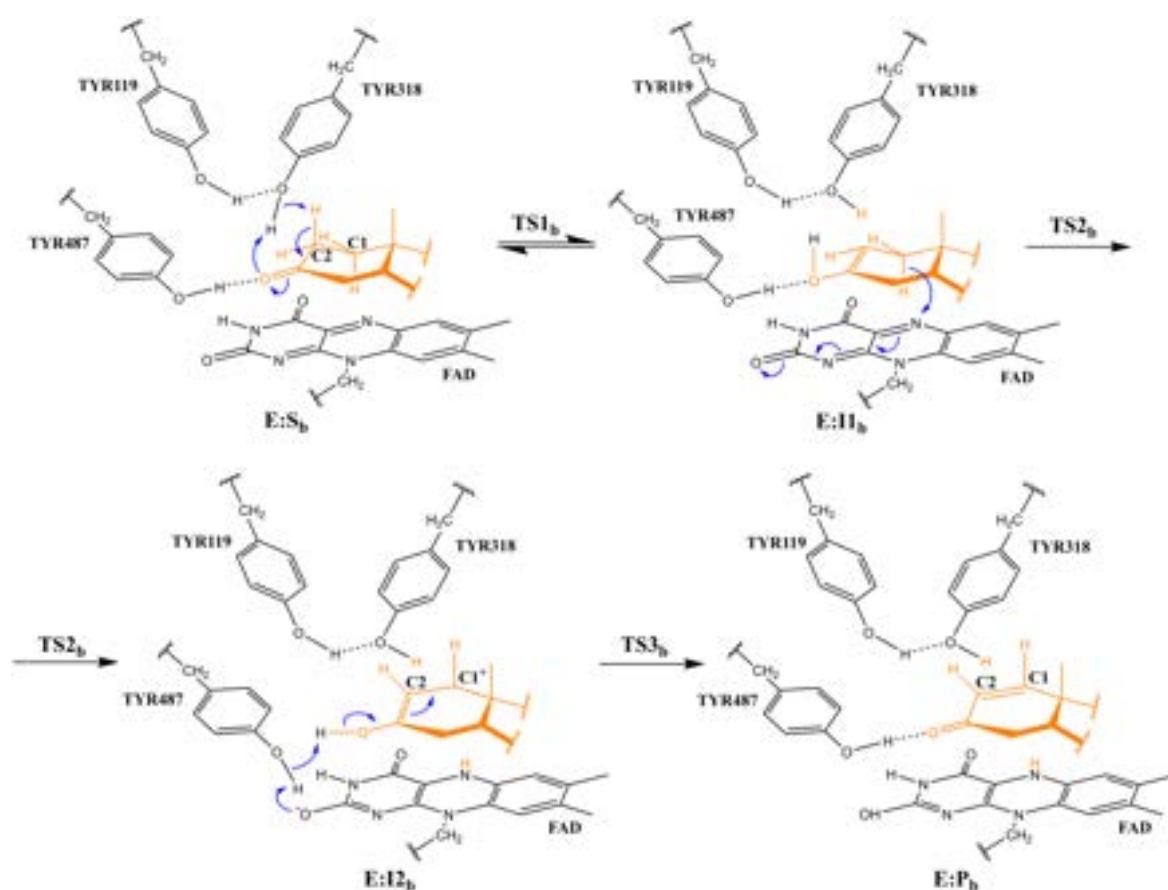
*Rysunek 16 Schemat mechanizmu alternatywnego, niezakładającego istnienia zdeprotonowanej Tyr318 w centrum aktywnym. Zaadoptowano z **P3**. W pierwszym etapie atom z pozycji C2β steroidu jest przeniesiony na atom tlenu Tyr318, równocześnie atom wodoru z grupy hydroksylowej Tyr318 jest transferowany na grupę ketonową steroidu. W drugim etapie jon wodorkowy jest odrywany z pozycji C1α na FAD. W ostatnim etapie atom wodoru grupy enolowej przechodzi na atom tlenu Tyr487, a atom wodoru związany wcześniej z tą tyrozyną jest transferowany na FAD.*

Alternatywny proces katalitycznego 1,2-odwodornienia składa się z trzech etapów. W pierwszym następuje równoczesne przeniesienie dwóch protonów: jednego z pozycji C2β

na atom tlenu Tyr318 i drugiego z grupy hydroksylowej Tyr318 na grupę ketonową steroidu. Drugi etap jest analogiczny do wcześniej omawianego mechanizmu, jon wodorkowy jest transferowany z pozycji C1α na atom N5 w cząsteczce FAD. W ostatnim etapie ponownie następuje jednoczesny transfer dwóch protonów: z grupy hydroksylowej enolu na atom tlenu grupy OH Tyr487 i z Tyr487 na FAD. Obliczenia wykazały, że mechanizm ten jest skrajnie nieprawdopodobny. Oszacowana bariera pierwszego etapu (aktywacji steroidu i przeniesienia protonu na grupę C3=O) wyniosła 44,8 kcal/mol. Energia pierwszego produktu przejściowego również okazała się zaskakująco wysoka, rzędu 27,9 kcal/mol. Zoptymalizowane struktury DHT dla tego mechanizmu są widoczne na Rysunku 17.

Najważniejsze wnioski z pracy **P3**:

- Obecność Tyr318 w stanie zdeprotonowanym jest kluczowa dla aktywności KSTD1,

- Reakcja odwodornienia przebiega zgodnie z mechanizmem zaproponowanym w literaturze; ten mechanizm jest znacznie bardziej prawdopodobny niż jednoczesny transfer atomów wodoru z pozycji C1 i C2 lub proces rozpoczynający się od transferu z pozycji C1,

- Uzyskany profil entalpi swobodnej reakcji wyjaśnia obserwowanie wartości KIE istotnie wyższych od jedności, zarówno dla ketosteroidów deuterowanych w pozycji C2β, jak i C1α,

- Szybkości pierwszego i drugiego etapu reakcji katalizowanych przez KSTD1 są na tyle zbliżone, by oba etapy miały istotny wpływ na szybkość procesu RHR,

- Tyr318 pełni rolę nie tylko zasady, która uczestniczy w transferze protonu w pierwszym etapie reakcji, ale również stabilizuje powstały produkt pośredni poprzez wiązanie wodorowe wraz z Tyr487.
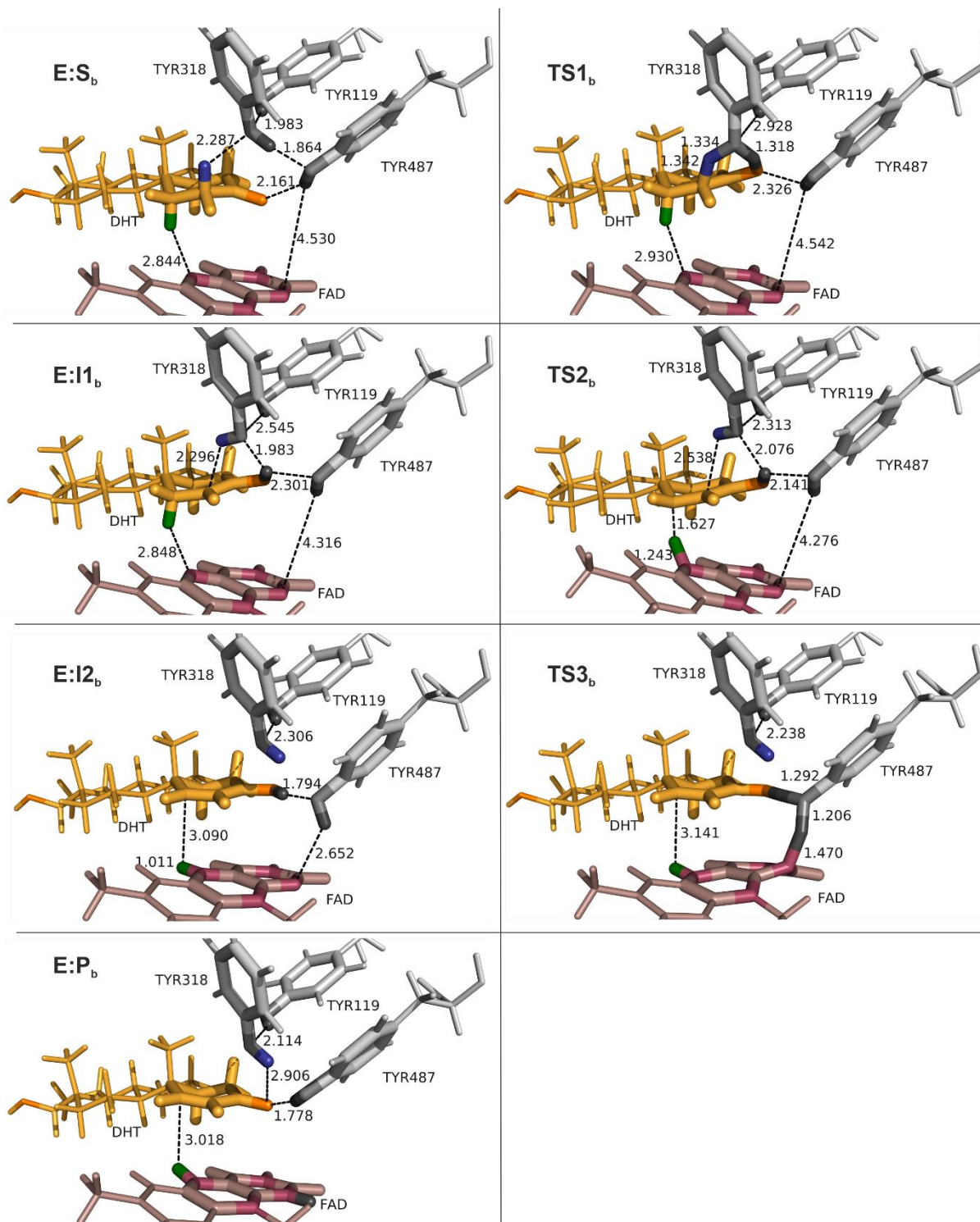
Rysunek 17 Reprezentatywne struktury stanów stacjonarnych dla DHT według mechanizmu alternatywnego. Zaadoptowano z SI P3.

Artykuł P4 – P. Wójcik, M. Glanowski, B. Mrugała, M. Procner, O. Zastawny, M. Flejszar, K. Kurpiewska, E. Niedziałkowska, W. Minor, M. Oszajca, A. J. Bojarski, A. M. Wojtkiewicz, M. Szaleniec, Structure, mutagenesis and QM:MM modelling of 3-ketosteroid Δ^1-dehydrogenase from Sterolibacterium denitrificans ‑ the role of new putative membrane-associated domain and proton-relay system in catalysis, Biochemistry 2023, 62, 3, 808-823

Praca **P4** poszerza dotychczasowe analizy o kolejny enzym z grupy Δ^1-KSTD, pochodzący z organizmu *S. denitrificans (*Δ^1-dehydrogenazę CHON, AcmB).[41] Podobnie jak w przypadku publikacji **P2** oraz **P3**, w pracy zaprezentowano zarówno wyniki eksperymentalne jak i obliczeniowe, a wkład badawczy autora dysertacji dotyczy obliczeniowej części badań. W szczególności, w pracy przedstawiono symulacje MD dla kompleksów E:S, obliczenia profili PMF dla szeregu substratów, w tym dla natywnego substratu CHON oraz dla szeregu zmutowanych wariantów enzymu. Przeprowadzono również obliczenia kinetycznego efektu izotopowego. Wykorzystano również symulację MD do zaproponowania roli jaką domena zasocjowana w membraną może pełnić w procesie katalitycznym.

W publikacji **P4** dokonano opisu nieznanej wcześniej struktury krystalicznej AcmB, przeprowadzono analizę bioinformatyczną sekwencji tego białka w kontekście ogółu znanych sekwencji KSTD. Praca zawiera również wyniki eksperymentów kinetycznych, w tym wyznaczenie KIE.

Analizy sekwencji i struktur wykazały, że AcmB wykazuje duże podobieństwo do KSTD1, z wyjątkiem jednego fragmentu (Tyr153–Arg204), który jest charakterystyczny dla wielu dehydrogenaz 3-ketosteroidowych i występuje również w enzymie z *S. denitrificans*. Niewykluczone, że ta część enzymu jest związana z asocjacją dehydrogenazy do błony komórkowej i odpowiada za pozycjonowanie wejścia do centrum aktywnego w kierunku bogatej w steroidy błony komórkowej bakterii.

Ze strony obliczeniowej, przeprowadzono symulacje MD, które pozwoliły oszacować energię swobodną wiązania szeregu ketosteroidów do białka (za pomocą metody MMPBSA). Ze względu na dużą mobilność substratów w centrum aktywnym, wyniki te są opisane

w dwóch wariantach. W pierwszym z nich do obliczeń MMPBSA wzięto cały czas symulacji dynamiki (60 ns). Uzyskane w ten sposób wartości, opisujące powinowactwo enzymu do substratu bez ingerencji eksperymentatora w wybór fragmentów symulacji, mieszczą się w przedziale −10,0 do −3,4 kcal/mol. W drugim wariancie, do analizy wybrano 5 ns cechujące się największą stabilnością parametrów geometrycznych, wskazujących na skuteczne związanie steroidu w centrum aktywnym (w szczególności trwałego wiązania grupy C3=O z resztą Tyr536 lub Gly540). Wówczas wartości oszacowanej energii wiązania zawarły się zakresie −12,4 do −5,7 kcal/mol. W obydwu przypadkach największą stabilizacją cechował się kompleks enzym:CHON. Porównano również energie interakcji poszczególnych substratów z całym białkiem oraz unikalnym fragmentem sekwencji AcmB. Okazało się, że substraty o rozbudowanych podstawnikach w pozycji C17 charakteryzują się istotnie większą energią oddziaływania z resztami 153–204. Podsumowanie wyników znajduje się w Tabeli 2.

Tabela 2. Podsumowanie obliczonych wartości energii wiązania (ΔG$_b$) oraz energii oddziaływania substratu z całym białkiem i unikalnym fragmentem AcmB.

Substrat	ΔG$_b$ dla pełnej symulacji MD [kcal/mol]	ΔG$_b$ dla wybranych 5ns symulacji MD [kcal/mol]	Energia oddziaływania substrat-białko [kcal/mol]	Energia oddziaływania substrat-reszty 153–204 [kcal/mol]
Androst-4-en-3,17-dion	−4,1±0,2	−6,1 ± 0,1	−40,1±0,2	−4,93±0,05
Cholest-4-en-3-on	**−10,0 ± 0,3**	**−12,4± 0,1**	**−61,4±0,2**	**−15,68±0,08**
Dihydrotestosteron	−6,6±0,1	−7,0±0,1	−41,9±0,1	−10,63±0,06
Progesteron	−6,3±0,2	−7,3±0,1	−42,0±0,2	−7,95±0,04
17-Metylotestosteron	−7,6±0,1	−8,7±0,1	−46,3±0,2	−4,53±0,04
Propionian testosteronu	−4,9±0,3	−7,1±0,1	−49,2±0,2	−10,02±0,07
Octan 6-Dehydrotestosteronu	−3,8±0,5	−7,2±0,1	−53,4±0,2	−11,10±0,07
Diosgenon	−3,4±0,3	−5,7±0,2	**−60,0±0,2**	**−13,48±0,07**

Metodologia modelowania QM:MM wyglądała analogicznie jak w przypadku pracy **P3**. Ponownie warstwa QM składała się z trzech części: ketosteroidu, fragmentu FAD oraz fragmentu Tyr363 (analogicznie do Tyr318 w KSTD1). Współrzędne reakcji zdefiniowano analogicznie jak w przypadku modelowania enzymu z *R. erythropolis*. Dwuwymiarowe mapy PES wykazały, że mechanizm reakcji jest niemal identyczny z tym opisanym dla KSTD1 pomimo tego, że enzym został przygotowany dla pH 6,5. Ponownie transfer protonu musi najpierw zostać przeprowadzony z pozycji C2β, a następnie jonu hydroniowego z C2α (Rysunek 18). Profile FES (*Free Energy Surface*) dla natywnego wariantu enzymu zostały

obliczone dla AD, DHT, 17-MT i CHON. Ponadto dla AD obliczono również profile dla zmutowanych wariantów białka: Y115F, Y118F, Y53F, Y467F. Pozwoliło to na przebadanie wpływ mutacji kolejnych tyrozyn w bezpośrednim otoczeniu substratu na profil reakcji oraz rozkład ładunków reagentów. Okazało się, że największy wpływ na profil reakcji ma mutacja Y536F. W tym przypadku nastąpiło drastyczne podwyższenie energii E:I (z 2,7 na 9,4 kcal/mol) oraz TS2 (z 15,5 na 28,7 kcal/mol). Aktywność tego wariantu enzymu wynosiła 0,6% w stosunku do natywnej wersji AcmB. Oznacza to, że wiązania wodorowe Tyr536 z grupą ketonową przy pozycji C3 jest kluczowe dla stabilizacji zarówno produktu pośredniego, jak i stanu przejściowego związanego z transferem jonu wodorkowego na FAD. Zmiana pozostałych tyrozyn miała mniejszy efekt na energie stanów stacjonarnych, który malał wraz z narastającą odległością aminokwasu od substratu w szeregu Y118F (21,6 kcal/mol), Y115F (18,9 kcal/mol), Y467F (16,7 kcal/mol). Zamiana Tyr118 na fenyloalaninę powoduje, że produkt przejściowy i drugi stan przejściowy były stabilizowane przez jedno wiązanie wodorowe mniej. Tak więc niekorzystny wpływ tej mutacji na drugą barierę wydaje się zrozumiały. Z kolei mutacje Y115F i Y467F zaburzają istniejący łańcuch stabilizujący system transferujący protony z centrum katalitycznego do rozpuszczalnika i zmieniają polarność centrum aktywnego, co wydaje się mieć wpływ na stabilizację karboanionowego produktu przejściowego i TS2.
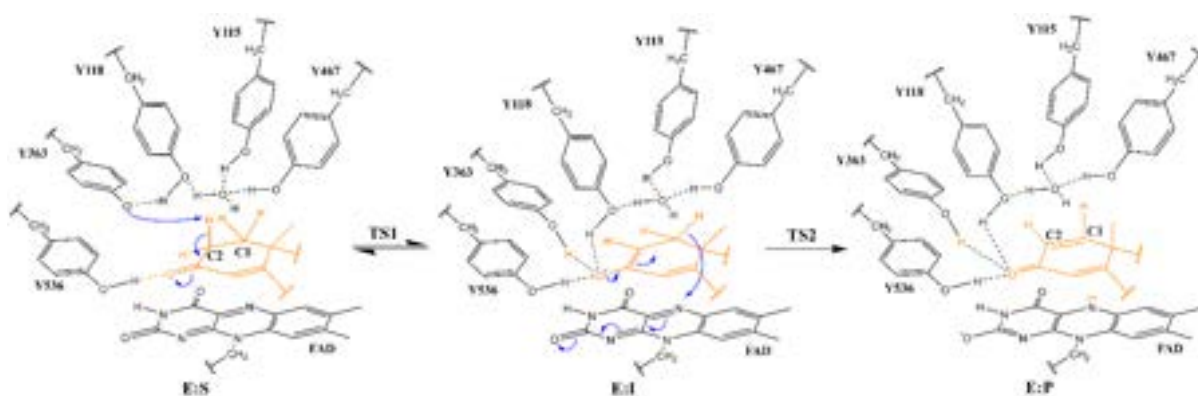
Rysunek 18 Schemat reakcji odwodornienia katalizowanej przez AcmB.

Ciekawy jest też wpływ mutacji na pierwszą barierę związaną z aktywacją steroidu. O ile w przypadku Y115F i Y467F są to wartości porównywalne (odpowiednio 12,2 kcal/mol i 14,0 kcal/mol) z tą uzyskaną dla natywnej wersji enzymu (13,2 kcal/mol), to już mutacja Y118F doprowadziła do obniżenie bariery dla TS1 do 10,8 kcal/mol. Można to wyjaśnić bardziej zasadowym charakterem Tyr363 w sytuacji bez wiązania wodorowego z Tyr118. Znajduje to odzwierciedlenie w ładunkach cząstkowych atomu tlenu grupy fenolowej (−0,9

i −0,46 odpowiednio dla Y118F i WT dla ładunków zanalizy populacyjnej APT). Jeszcze niższą pierwszą barierę przewidziano dla Y536F (9,9 kcal/mol), jednakże wówczas dochodzi także do zmiany pozycji ketosteroidu w centrum aktywnym, co doprowadza do gorszej stabilizacji w dalszych etapach reakcji.

Wykorzystując AD jako substrat rozważono również sytuacje, gdy transfer protonu z ketosteroidu w pierwszym etapie reakcji następuje z pozycji C2α, a nie C2β. Rozpoczęcie reakcji od tej alternatywnej pozycji podniosło energię TS1 z 13,2 na 16,3 kcal/mol. Wynik ten potwierdza możliwość nie enancjoselektywnej aktywacji steroidu co obserwowano już eksperymentalnie.[9,12]

Uzyskano profile energetyczne także dla kilku innych ketosteroidów. Zestawienie uzyskanych obliczeń z wynikami pomiarów kinetycznych (względnej aktywności właściwej wyznaczonej w warunkach stacjonarnych) pozwala na jakościowe potwierdzenie zdolności predykcyjnej metod obliczeniowych. Oczywiście aby prawidłowo zweryfikować obliczenia metodami kinetycznymi należałoby dysponować stałymi kinetycznymi wyznaczonymi w stanie przedstacjonarnym, co niestety okazało się niemożliwe dla AcmB ze względu na zbyt szybki proces RHR. Uzyskane wyniki dodatkowo potwierdzają eksperymentalne dowody w kwestii natywnego substratu AcmB. Najniższą barierę uzyskano dla cholest-4-en-3-onu (14,1 kcal/mol) co wskazuje, że dla tego substratu enzym będzie wykazywać się najszybszym etapem 1,2-odwodornienia.

Tabela 3. Uzyskane wartości profilu reakcji odwodornienia dla różnych wariantów AcmB razem z eksperymentalną aktywnością.

Wariant AcmB	Substrat	TS1	I	TS2	P	rSA [%]
WT 2βH	AD	13,2	2,7	15,5	−9,2	100
WT 2αH	AD	16,3	2,7	15,5	−9,2	-
WT	17-MT	13,6	2,8	15,6	−9,2	71
WT	DHT	16,5	5,9	19,0	−10,4	19
WT	CHON	12,8	2,8	14,1	−9,5	n.d.*
Y118F	AD	10,8	4,7	21,6	−1,4	śladowa
Y115F	AD	12,2	5,2	18,9	−0,9	32
Y467F	AD	14,0	5,3	16,7	−6,1	58

n.d.* - wyniki eksperymentalne nieporównywalne z pozostałym ze względu na użycie solubilizatora

Podobnie jak w przypadku pracy **P3**, eksperymentalne wyniki KIE zostały porównane z obliczonymi. Niestety tym razem nie były dostępne KIE zmierzone bezpośrednio dla procesu RHR, ze względu na zbyt dużą szybkość reakcji enzymu. Możliwy był jednie pomiar

w warunkach stacjonarnych. Zmierzone wartości KIE okazały się ponownie porównywalne dla 17MT i DHT, wynosząc odpowiednio 1,05–1,33 i 1,06–1,17, w zależności od pH dla metody bezpośredniej. Zakładając, że proces OHR jest całkowicie niewrażliwy na podstawienie izotopowe ketosteroidu, aby oszacować eksperymentalne KIE RHR, niezbędna jest znajomość stosunku stałych szybkości opisujące te procesy. Niestety ta wielkość również nie została wyznaczona, wiadomo jednak, że proces OHR jest wolniejszy od procesu RHR.

Eksperymentalne KIE z metody konkurencyjnej (interpretowane jako efekt izotopowy dla stosunku V/K) dostarczyło nieznacznie wyższych wartości dla obydwu substratów: 1,23–1,50 i 1,21-1,28, odpowiednio dla DHT i 17-MT.

Sumaryczny obliczony KIE dla RHR zawiera się w granicach 1,07–2,14 dla 17MT oraz 4,42–4,96 dla DHT, uwzględniając precyzję obliczonych energii na poziomie 1 kcal/mol. Zakładając, że rzeczywisty KIE dla procesu RHR zawiera się między skrajnymi obliczonymi wartościami, oraz przyjmując różne stosunki szybkości procesu RHR do OHR udało się pokazać, że opublikowany model jest ogólnie zgodny z eksperymentem.

Kluczowe wyniki pracy **P4**:

- Wykazano rolę domeny związanej z błoną w stabilizacji kompleksu E:S dla substratów o rozbudowanych podstawnikach przy C17; obecność tej domeny skutkuje niższymi wartościami ΔG wiązania dla CHON niż te obliczone dla KSTD1, co tłumaczy obserwowane eksperymentalnie różnice w powinowactwie tego substratu do obu enzymów,
- Potwierdzono, że mechanizm reakcji katalizowanej przez AcmB jak analogiczny jak KSTD1,
- Przeanalizowano wpływ mutacji poszczególnych tyrozyn w centrum aktywnym na profil reakcji i uzyskano pod tym względem zgodność z eksperymentalnymi pomiarami aktywności enzymu,
- Zweryfikowano możliwość oderwania protonu w pierwszym etapie reakcji z pozycji C2α; uzyskana wysokość bariery (16,3 kcal/mol) sugeruje, że jest to jak najbardziej możliwy mechanizm w przypadku, gdy transfer z pozycji C2β jest utrudniony,
- Porównanie barier uzyskanych dla różnych substratów potwierdza tezę, że to CHON jest substratem natywnym dla AcmB.

Podsumowanie

Praca przedstawia analizę reakcji katalizowanej przez enzymy Δ^1-KSTD, przeprowadzoną za pomocą metod obliczeniowych. Modelowanie oparto o struktury dwóch enzymów z tej grupy: KSTD1 i AcmB.

Modelowanie QM:MM i QM:MM MD pozwoliło na porównanie możliwych hipotez mechanistycznych. Uzyskane wyniki potwierdziły hipotezę zaproponowaną w literaturze, tzn. dwuetapowe przeniesienie atomów wodoru, najpierw z pozycji C2β, a następnie C1α. Transfer tych samych atomów w odwrotnej kolejności okazał się znacznie mniej prawdopodobny. Ponadto wykazano, że nie istnieje punkt siodłowy na powierzchni energii potencjalnej, który odpowiadałby równoczesnemu przeniesieniu dwóch atomów wodoru.

Rozważono również kwestie konieczności istnienia jonu tyrozylowego w centrum aktywnym enzymu. Obliczenia dla mechanizmu związanego z białkiem, w którym wszystkie tyrozyny są w natywnym stanie protonacyjnym, wykazały, że obecność tego anionu jest kluczowa dla reaktywności Δ^1-KSTD.

Obliczenia KIE i ich interpretacja pozwoliły na walidację modelu i wyjaśnienie niejasności dotyczących kinetyki Δ^1-KSTD. Stanowią one jedno z głównych osiągnięć pracy. Analiza uzyskanych profili FES wraz z KIE dla reakcji elementarnych doprowadziła do spójnego opisu mechanizmu reakcji Δ^1-KSTD.

W przypadku AcmB dokonano wyczerpującej, obliczeniowej analizy wpływu wiązań wodorowych w otoczeniu substratu poprzez uzyskanie profili energetycznych reakcji dla zmutowanych wariantów enzymu. Dzięki temu oraz analizom populacyjnym, udało się wyjaśnić rolę wszystkich kluczowych aminokwasów w centrum aktywnym. Uzyskane bariery dla poszczególnych wariantów enzymu układają się w szereg zgodny z eksperymentalną aktywnością.

Dzięki obliczeniom MMPBSA uzyskano oszacowanie energii wiązania różnych ketosteroidów do centrum aktywnego, to z kolei przyczyniło się do potwierdzenia aktywności KSTD1 wobec CHON, która nie była postulowana wcześniej.

Bibliografia

(1) Sedlaczek, L.; Smith, L. L. Biotransformations of Steroids. *Crit. Rev. Biotechnol.* **1988**, *7* (3), 187–236. https://doi.org/10.3109/07388558809146602.

(2) Donova, M. V.; Egorova, O. V. Microbial Steroid Transformations: Current State and Prospects. *Appl. Microbiol. Biotechnol.* **2012**, *94* (6), 1423–1447. https://doi.org/10.1007/s00253-012-4078-0.

(3) Rohman, A.; Dijkstra, B. W. The Role and Mechanism of Microbial 3-Ketosteroid Δ1-Dehydrogenases in Steroid Breakdown. *J. Steroid Biochem. Mol. Biol.* **2019**, *191*, 105366. https://doi.org/10.1016/j.jsbmb.2019.04.015.

(4) Rohman, A.; Van Oosterwijk, N.; Thunnissen, A. M. W. H.; Dijkstra, B. W. Crystal Structure and Site-Directed Mutagenesis of 3-Ketosteroid Δ1-Dehydrogenase from Rhodococcus Erythropolis SQ1 Explain Its Catalytic Mechanism. *J. Biol. Chem.* **2013**, *288* (49), 35559–35568. https://doi.org/10.1074/jbc.M113.522771.

(5) Wójcik, P.; Glanowski, M.; Mrugała, B.; Procner, M.; Zastawny, O.; Flejszar, M.; Kurpiewska, K.; Niedziałkowska, E.; Minor, W.; Oszajca, M.; Bojarski, A. J.; Wojtkiewicz, A. M.; Szaleniec, M. Structure, Mutagenesis, and QM:MM Modeling of 3-Ketosteroid Δ 1 -Dehydrogenase from Sterolibacterium Denitrificans ─The Role of a New Putative Membrane-Associated Domain and Proton-Relay System in Catalysis. *Biochemistry* **2023**, *62* (3), 808–823. https://doi.org/10.1021/acs.biochem.2c00576.

(6) Ullrich, V. Enzymatic Hydroxylations with Molecular Oxygen. *Angew. Chemie Int. Ed. English* **1972**, *11* (8), 701–712. https://doi.org/10.1002/anie.197207011.

(7) Sih, C. J.; Bennett, R. E. Steroid I-Dehydrogenase of Nocardia Restrictus. *Biochim. Biophys. Acta* **1962**, *56*, 584–592. https://doi.org/10.1016/0006-3002(62)90610-8.

(8) Abul-Hajj, Y. J. Isolation of Vitamin K2(35) from Nocardia Restrictus and Corynebacterium Simplex. A Natural Electron Acceptor in Microbial Steroid Ring A Dehydrogenations. *J. Biol. Chem.* **1978**, *253* (7), 2356–2360. https://doi.org/10.1016/S0021-9258(17)38081-X.

(9) Jerussi, R.; Ringold, H. J. The Mechanism of the Bacterial C-1,2 Dehydrogenation of Steroids. III. Kinetics and Isotope Effects *. *Biochemistry* **1965**, *4* (10), 2113–2126. https://doi.org/10.1021/bi00886a028.

(10) Hayano, M.; Ringold, H. J.; Stefanovic, V.; Gut, M.; Dorfman, R. I. The Stereochemical Course of Enzymatic Steroid 1,2-Dehydrogenation. *Biochem. Biophys. Res. Commun.* **1961**, *4* (6), 454–459. https://doi.org/10.1016/0006-291X(61)90307-2.

(11) Itagaki, E.; Matushita, H.; Hatta, T. Steroid Transhydrogenase Activity of 3-Ketosteroid-δ-Dehydrogenase from Nocardia Corallina. *J. Biochem.* **1990**, *108* (1emi), 122–127. https://doi.org/10.1093/oxfordjournals.jbchem.a123150.

(12) Ringold, H. J.; Hayano, M.; Stefanovic, V. Concerning the Stereochemistry and Mechanism of the Bacterial C-1,2 Dehydrogenation of Steroids. *J. Biol. Chem.* **1963**, *238* (6), 1960–1965. https://doi.org/10.1016/S0021-9258(18)67926-8.

(13) Itagaki, E.; Wakabayashi, T.; Hatta, T. Purification and Characterization of 3-Ketosteroid-Δ1-Dehydrogenase from Nocardia Corallina. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **1990**, *1038* (1), 60–67. https://doi.org/10.1016/0167-4838(90)90010-D.

(14) Choi, K.-P.; Molnár, I.; Yamashita, M.; Murooka, Y. Purification and Characterization of the 3-Ketosteroid-Δ1-Dehydrogenase of Arthrobacter Simplex Produced in Streptomyces Lividans1. *J. Biochem.* **1995**, *117* (5), 1043–1049. https://doi.org/10.1093/oxfordjournals.jbchem.a124804.

(15) Kondo, E. Steroid 1-Dehyclrogenation by a Crude Enzyme Preparation from Arthrobacter Simplex. *Agric. Biol. Chem.* **1963**, *27* (1), 69–70.

https://doi.org/10.1080/00021369.1963.10858065.

(16)   Aries, V. C.; Goddard, P.; Hill, M. J. Degradation of Steroids by Intestinal Bacteria. III. 3-OXO-5β-Steroid Δ1-Dehydrogenase and 3-OXO-5β-Steroid Δ4-Dehydrogenase. *Biochim. Biophys. Acta - Lipids Lipid Metab.* **1971**, *248* (3), 482–488. https://doi.org/10.1016/0005-2760(71)90238-4.

(17)   Koepsell, H. J. 1-Dehydrogenation of Steroids by Septomyxa Affinis. *Biotechnol. Bioeng.* **1962**, *4* (1), 57–63. https://doi.org/10.1002/bit.260040108.

(18)   Penasse, L.; Nomine, G. Le Centre Actif de La 3-Oxosteroide Delta1-Deshydrogenase d'Arthrobacter Simplex. *Eur. J. Biochem.* **1974**, *47* (3), 555–559. https://doi.org/10.1111/j.1432-1033.1974.tb03725.x.

(19)   Penasse, L.; Peyre, M. Studies of 3-Oxo Steroid Δ1-Oxydo Reductase of Arthrobacter Simplex. *Steroids* **1968**, *12* (4), 525–544. https://doi.org/10.1016/S0039-128X(68)80116-3.

(20)   Chiang, Y.-R.; Ismail, W.; Gallien, S.; Heintz, D.; Van Dorsselaer, A.; Fuchs, G. Cholest-4-En-3-One-Δ1-Dehydrogenase, a Flavoprotein Catalyzing the Second Step in Anoxic Cholesterol Metabolism. *Appl. Environ. Microbiol.* **2008**, *74* (1), 107–113. https://doi.org/10.1128/AEM.01968-07.

(21)   Sih, C. J.; Bennett, R. E. Enzymic Dehydrogenation of Steroid A Ring. *Biochim. Biophys. Acta* **1960**, *38*, 378–379. https://doi.org/10.1016/0006-3002(60)91268-3.

(22)   Wojtkiewicz, A. M.; Wójcik, P.; Procner, M.; Flejszar, M.; Oszajca, M.; Hochołowski, M.; Tataruch, M.; Mrugała, B.; Janeczko, T.; Szaleniec, M. The Efficient Δ1-Dehydrogenation of a Wide Spectrum of 3-Ketosteroids in a Broad PH Range by 3-Ketosteroid Dehydrogenase from Sterolibacterium Denitrificans. *J. Steroid Biochem. Mol. Biol.* **2020**, *202*, 105731. https://doi.org/10.1016/j.jsbmb.2020.105731.

(23)   Świderek, K.; Paneth, P. Binding Isotope Effects. *Chem. Rev.* **2013**, *113* (10), 7851–7879. https://doi.org/10.1021/cr300515x.

(24)   Dijkstra, B. W.; van Oosterwijk, N.; Rohman, A. Structure and Catalytic Mechanism of 3-Ketosteroid Dehydrogenases. *Procedia Chem.* **2016**, *18* (Mcls 2015), 3–11. https://doi.org/10.1016/j.proche.2016.01.006.

(25)   Self-Consistent Field, with Exchange, for Beryllium. *Proc. R. Soc. London. Ser. A - Math. Phys. Sci.* **1935**, *150* (869), 9–33. https://doi.org/10.1098/rspa.1935.0085.

(26)   Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136* (3B), B864–B871. https://doi.org/10.1103/PhysRev.136.B864.

(27)   Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648–5652. https://doi.org/10.1063/1.464913.

(28)   Dapprich, S.; Komáromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. A New ONIOM Implementation in Gaussian98. Part I. The Calculation of Energies, Gradients, Vibrational Frequencies and Electric Field Derivatives. *J. Mol. Struct. THEOCHEM* **1999**, *461–462*, 1–21. https://doi.org/10.1016/S0166-1280(98)00475-8.

(29)   Field, M. J.; Albe, M.; Bret, C.; Proust-De Martin, F.; Thomas, A. The Dynamo Library for Molecular Simulations Using Hybrid Quantum Mechanical and Molecular Mechanical Potentials. *J. Comput. Chem.* **2000**, *21* (12), 1088–1100. https://doi.org/10.1002/1096-987X(200009)21:12<1088::AID-JCC5>3.0.CO;2-8.

(30)   Christensen, A. S.; Kubař, T.; Cui, Q.; Elstner, M. Semiempirical Quantum Mechanical Methods for Noncovalent Interactions for Chemical and Biochemical Applications. *Chem. Rev.* **2016**, *116* (9), 5301–5337. https://doi.org/10.1021/acs.chemrev.5b00584.

(31)   J. S. Dewar, M.; G. Zoebisch, E.; F. Healy, E.; J. P. Stewart, J. Development and Use of Quantum Mechanical Molecular Models. 76. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **2002**, *107* (13), 3902–3909. https://doi.org/10.1021/ja00299a024.

(32) Kirkwood, J. G. Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.* **1935**, *3* (5), 300–313. https://doi.org/10.1063/1.1749657.

(33) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23* (2), 187–199. https://doi.org/10.1016/0021-9991(77)90121-8.

(34) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. THE Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13* (8), 1011–1021. https://doi.org/10.1002/jcc.540130812.

(35) Ruiz-Pernía, J. J.; Silla, E.; Tuñón, I.; Martí, S.; Moliner, V. Hybrid QM/MM Potentials of Mean Force with Interpolated Corrections. *J. Phys. Chem. B* **2004**, *108* (24), 8427–8433. https://doi.org/10.1021/jp049633g.

(36) Corchado, J. C.; Coitiño, E. L.; Chuang, Y.-Y.; Fast, P. L.; Truhlar, D. G. Interpolated Variational Transition-State Theory by Mapping. *J. Phys. Chem. A* **1998**, *102* (14), 2424–2438. https://doi.org/10.1021/jp9801267.

(37) Martin, A. C. R. A Practical Introduction to the Simulation of Molecular Systems, Martin J. Field, Cambridge University Press, 1999; ISBN 0-521-58129-X. 325 Pp. £50 (HB). *Comput. Chem.* **2000**, *24* (2), 239–240. https://doi.org/10.1016/S0097-8485(99)00063-7.

(38) Baker, J.; Kessi, A.; Delley, B. The Generation and Use of Delocalized Internal Coordinates in Geometry Optimization. *J. Chem. Phys.* **1996**, *105* (1), 192–212. https://doi.org/10.1063/1.471864.

(39) Singh, U. C.; Kollman, P. A. An Approach to Computing Electrostatic Charges for Molecules. *J. Comput. Chem.* **1984**, *5* (2), 129–145. https://doi.org/10.1002/jcc.540050204.

(40) Wigner, E. Crossing of Potential Thresholds in Chemical Reactions. *Zeitschrift für Phys. Chemie* **1932**, *B19*, 203–216.

(41) Chiang, Y.-R.; Ismail, W.; Heintz, D.; Schaeffer, C.; Van Dorsselaer, A.; Fuchs, G. Study of Anoxic and Oxic Cholesterol Metabolism by Sterolibacterium Denitrificans. *J. Bacteriol.* **2008**, *190* (3), 905–914. https://doi.org/10.1128/JB.01525-07.

(42) Knol, J.; Bodewits, K.; Hessels, G. I.; Dijkhuizen, L.; van der Geize, R. 3-Keto-5α-Steroid Δ1-Dehydrogenase from Rhodococcus Erythropolis SQ1 and Its Orthologue in Mycobacterium Tuberculosis H37Rv Are Highly Specific Enzymes That Function in Cholesterol Catabolism. *Biochem. J.* **2008**, *410* (2), 339–346. https://doi.org/10.1042/BJ20071130.

(43) Wovcha, M. G.; Brooks, K. E.; Kominek, L. A. Evidence for Two Steroid 1,2-Dehydrogenase Activities in Mycobacterium Fortuitum. *Biochim. Biophys. Acta - Lipids Lipid Metab.* **1979**, *574* (3), 471–479. https://doi.org/10.1016/0005-2760(79)90243-1.

(44) Miller, B. R.; McGee, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. MMPBSA.Py : An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **2012**, *8* (9), 3314–3321. https://doi.org/10.1021/ct300418h.

Dorobek naukowy kandydata

Artykuły

Związane z tematem dysertacji:

1. Agnieszka M. Wojtkiewicz, **Michał Glanowski**, Piotr Waligórski, Tomasz Janeczko, Maciej Szaleniec, *1,2-Hydrogenation and Transhydrogenation Catalyzed by 3-Ketosteroid Δ^1-Dehydrogenase from Sterolibacterium denitrificans—Kinetics, Isotope Labelling and QM:MM Modelling Studies*, **International Journal of Molecular Sciences** 2022, 23, 14660, DOI: 10.3390/ijms232314660

2. Patrycja Wójcik, **Michał Glanowski**, Beata Mrugała, Magdalena Procner, Olga Zastawny, Monika Flejszar, Katarzyna Kurpiewska, Ewa Niedziałkowska, Wladek Minor, Maria Oszajca, Andrzej J. Bojarski, Agnieszka M. Wojtkiewicz, and Maciej Szaleniec, *Structure, Mutagenesis, and QM:MM Modeling of 3-Ketosteroid Δ^1-Dehydrogenase from Sterolibacterium denitrificans─The Role of a New Putative Membrane-Associated Domain and Proton-Relay System in Catalysis*, **Biochemistry**, 2023, 62, 3, 808-823, DOI: 10.1021/acs.biochem.2c00576

3. **Michał Glanowski**, Sangita Kachhap, Tomasz Borowski, Maciej Szaleniec, *Model Setup and Procedures for Prediction of Enzyme Reaction Kinetics with QM-Only and QM:MM Approaches*. In: Vanhaelen, Q. (eds) *Computational Methods for Estimating the Kinetic Parameters of Biological Systems. Methods in Molecular Biology,* vol 2385. (2022) 175-236 Humana, New York, NY. DOI: 10.1007/978-1-0716-1767-0_10, ISBN: 978-1-0716-1766-3

4. **Michał Glanowski**, Patrycja Wójcik, Magdalena Procner, Tomasz Borowski, Dawid Lupa, Przemysław Mielczarek, Maria Oszajca, Katarzyna Świderek, Vicent Moliner, Andrzej J. Bojarski, Maciej Szaleniec, *Enzymatic Δ1-Dehydrogenation of 3-Ketosteroids-Reconciliation of Kinetic Isotope Effects with the Reaction Mechanism*, **ACS Catalysis**, 2021, 11, 8211-8225, DOI: 10.1021/acscatal.1c01479

5. Patrycja Wójcik, **Michał Glanowski**, Agnieszka M. Wojtkiewicz, Ali Rohman, Maciej Szaleniec, *Universal Capability of 3-Ketosteroid Δ1-Dehydrogenases to Catalyze Δ1-Dehydrogenation of C17-Substituted Steroids*, **Microbial Cell Factories**, 2021, 20(1), article number: 119, DOI: 10.1186/s12934-021-01611-5

Niezwiązane z tematem dysertacji:

6. Emilia Kuzniak-Glanowska, **Michał Glanowski**, Rafał Kurczab, Andrzej J. Bojarski, Robert Podgajny, *Mining Anion–Aromatic Interactions in the Protein Data Bank*, ***Chemical Science***, 2022, 13(14), 3984–3998, DOI: 10.1039/D2SC00763K, DOI: 10.1039/D2SC00763K

7. Grzegorz Mazur, Marcin Makowski, Roman Łazarski, Radosław Włodarczyk, Ewa Czajkowska, **Michał Glanowski**, *Automatic Code Generation for Quantum Chemistry Applications,* ***International Journal of Quantum Chemistry,*** 2016, 116(18), 1370 – 1381, DOI: 10.1002/qua.25187

Staże naukowe:

1. 03.02.2020 – 11.03.2020, Universitat Jaume I, Castellon de la Plana, Hiszpania, staż naukowy finansowany ze środków programu NCBR POWR.03.02.00-00-I013/16

2. 27.08.2019 – 30.08.2019, University of Warwick, Coventry, UK, staż naukowy finansowany ze środków programu NAWA PROM PPI/PRO/2018/1/00006/U/001

3. 25.03.2019 – 25.05.2019, Universitat Jaume I, Castellon de la Plana, Hiszpania, staż naukowy finansowany ze środków programu Erasmus+

Konferencje, wystąpienia ustne:

1. 20-22.09.2021, XIV Kopernikańskie Seminarium Doktoranckie, współautor komunikatu, *Analiza bazy Protein Data Bank oraz znaczenie oddziaływań anion- w cząsteczkach biologicznych,* Emilia Kuźniak-Glanowska, Michał Glanowski, Rafał Kurczab, Andrzej J. Bojarski, Robert Podgajny

2. 16-21.12.2021, 2021 International Chemical Congress of Pacific Basin Societes, Online, Komunikat, *Insight into enzymatic Δ1-dehydrogenation of 3-ketosteroids: over 50 years of research verified with QM/MM,* Michał Glanowski, Patrycja Wójcik, Magdalena Procner, Maria Oszajca, Katarzyna Świderek, Vicent Moliner, Andrzej J. Bojarski, Maciej Szaleniec

3. 13-17.09.2021, 63. Zjazd Naukowy Polskiego Towarzystwa Chemicznego, Online, komunikat, *Dehydrogenazy 3-ketosteroidowe – modelowanie QMMM skonfrontowane z pomiarami kinetycznego efektu izotopowego,* Michał Glanowski, Patrycja Wójcik, Magdalena Procner, Maria Oszajc, Katarzyna Świderek, Vicent Moliner, Andrzej J. Bojarski, Maciej Szaleniec

4. 25.27.11.2020 - 27.11.2020, 52. Ogólnopolskie Kolokwium Katalityczne, Online, flash oral, *Modeling of oxidative 1-dehydrogenetation catalyzed by 3-ketosteroid Δ1-dehydrogenases*, <u>M. Glanowski</u>, P. Wójcik, M. Procner, M. Szaleniec, A. Bojarski, K. Świderek, V. Moliner

5. 25.11.2020 - 27.11.2020, 52. Ogólnopolskie Kolokwium Katalityczne, Online, współautor wystąpienia ustnego, *Investigation of the structure and the reaction mechanism of the 3-ketosteroid dehydrogenase from Stereolibacterium denitrificansfor further application to the biosynthesis of pharmaceutica*, <u>P. Wójcik</u>, B. Mrugała, M. Glanowski, O. Zastawny, A.M. Wojtkiewicz, K. Kurpiewska, E. Niedziałkowska, W. Minor, M. Szaleniec

6. 28.10.2019 - 29.10.2019, Trends in Enzyme Catalysis, Merging theory and experiment, Benicassim, Spain, współautor wystąpienia ustnego, *Combining kinetic studies with theoretical modeling – case study for ketosteroid dehydrogenases*, <u>M. Szaleniec</u>, P. Wójcik, M. Glanowski, A. M. Wojtkiewicz, M. Procner, B. Mrugała, O. Zastawny, M. Guzik, E. Niedziałkowska, W. Minor, K. Świderek, V. Moliner

7. 01.09.2019 - 06.09.2019, Summer School of Chemical Biology, Hirshegg, Austria, współautor wystąpienia ustnego, *Combining kinetic studies with theoretical modeling – case study for ketesteroid dehydrogenases,* <u>M. Szaleniec</u>, P. Wójcik, M. Glanowski, A. Wojtkiewicz, M. Procner, B. Mrugała, O. Zastawny, M. Guzik, E. Niedziałkowska, W. Minor, K. Świderek, V.Moliner

8. 20.03.2019 - 22.03.2019, 51 Ogólnopolskie Kolokwium Katalityczne, Kraków, Polska, flash oral, *Modeling of 3-ketosteroid Δ1-dehydrogenases*, <u>M. Glanowski</u>, M. Szaleniec, A. Bojarski

Konferencje, postery:

1. 16-21.12.2021, 2021 International Chemical Congress of Pacific Basin Societes, Online, współautor posteru, *Mining anion-aromatic interactions in the Protein Data Bank,* <u>Emilia Kuźniak-Glanowska</u>, Michał Glanowski, Rafał Kurczab, Andrzej J. Bojarski, Robert Podgajny

2. 18-21.10.2021, Advances and challenges in biomolecular simulations, EMBO workshop, Online, Poster, *QMMM and Kinetic Isotope Effect,* <u>Michał Glanowski</u>, Patrycja Wójcik, Magdalena Procner, Maria Oszajca, Katarzyna Świderek, Vicent Moliner, Andrzej J. Bojarski, MaciejSzaleniec

3. 19-22.07.21, Biotrans 2021, Online, współautor posteru, *Δ1-Dehydrogenation of 3-ketosteroids by 3-ketosteroid dehydrogenase from Sterolibacterium denitrificans – structural and mechanistic studies,* Patrycja Wójcik, Beata Mrugała, Michał Glanowski, Magdalena Procner, Katarzyna Kurpiewska, Olga Zastawny, Maciej Guzik, Maria Oszajca, Maciej Szaleniec

4. 13, 22, 30.06.2021, International Conference on Theorethical Aspects of Catalysis, Online, poster, *QM/MM modeling of enzymatic Δ1-dehydrogenation of 3-ketosteroids and validation by kinetic isotope effects,* Michał Glanowski, Patrycja Wójcik, Magdalena Procner, Maria Oszajca, Katarzyna Świderek, Vicent Moliner, Andrzej J. Bojarski, Maciej Szaleniec

5. 02.09.2019 - 06.09.2019, 1st International Conference on Noncovalent Interactions, Lizbona, Portugalia, współautor posteru, *PDB and the hidden secrets – do we know everything about anion-π interactions in macromolecules?,* E. Kuźniak, M. Glanowski, R. Kurczab, A. Bojarski, R. Podgajny

6. 09.07.2019 - 16.07.2019, Workshop on data collection and structure solving in macromolecular X-ray diffraction, Kraków, Polska, współautor posteru, Optimization of membrane flavoprotein purification and crystallization conditions, P. Wójcik, B. Mrugała, M. Glanowski, E. Niedziałkowska, A. Wojtkiewicz, W. Minor, M. Szaleniec

7. 07.07.2019 - 11.07.2019, 14th International Symposium on Biocatalysis and Biotransformations (BioTrans 2019), Groningen, Holandia, współautor posteru, Biocatalytic dehydrogenation of steroids by cholest-4-en-3-one Δ1-dehydrogenase - mechanistic study and examples of applicatio, M. Szaleniec, P. Wójcik, M. Glanowski, A. Wojtkiewicz, M. Flejszar, M. Hochołowski, O. Zastawny, M. Guzik

8. 17.03.2019 - 20.03.2019, 23rd West Coast Structural Biology Workshop, Pacific Grove, CA, USA, współautor posteru, *Optimization of aggregating flavoprotein expression and purification*, P. Wójcik, B. Mrugała, M. Glanowski, C. Martin, G. Bailleul, A. Wojtkiewicz, M. Fraaije, M. Szaleniec

9. 17.09.2018 - 21.09.2018, 61 Zjazd Naukowy Polskiego Towarzystwa Chemicznego, Kraków, Polska, poster, *Modelowanie mechanizmów reakcji Δ^1-dehydrogenaz-3-ketosteroidowych,* M. Glanowski, A. Wojtkiewicz, S. Mordalski, M. Szaleniec, A. Bojarski

10. 26.08.2018 - 30.08.2018, 9th International Congress on Biocatalysis, Hamburg, Niemcy, współautor posteru, *Cholest-4-en-3-one Δ^1-dehydrogenase – a biocatalyst for*

Publikacja P1

<div align="right">

Chapter 10

</div>

Model Setup and Procedures for Prediction of Enzyme Reaction Kinetics with QM-Only and QM:MM Approaches

Michał Glanowski, Sangita Kachhap, Tomasz Borowski, and Maciej Szaleniec

Abstract

The enzyme-catalyzed reactions are traditionally studied with experimental kinetic assays. The modern theoretical modeling techniques provide a complementary way to investigate these catalytic reactions. Experimental assay frequently does not allow an unequivocal answer to the factors controlling the reaction mechanism. On the other hand, the theoretical experiments provide a precise understanding of the molecular-level steps involved in catalytic reactions. However, modeling requires at least structural data on the enzyme and reactant, and the complexity of the enzyme systems can still be a challenge.

In this chapter, we are going to describe how to apply theoretical modeling methods, such as MD simulation, QM-only cluster models of enzyme active site, or QM:MM multiscale modeling to study enzyme kinetics and even to predict kinetic isotope effect (KIE). We present a full protocol that starts from the PDB structure of the enzyme, through MD simulation of enzyme: substrate complex and statistical analysis of MD trajectory, selection of a model of the active site, and study of reaction pathways. We show how theoretical predictions basing on QM-only cluster models, QM:MM model, or multiple QM:MM models derived from QM:MM:MD simulations can be correlated with experimental kinetic results. Finally, we show how one can calculate intrinsic KIE associated with an individual molecular step.

Key words Enzyme-catalyzed reactions, Quantum mechanics, Molecular mechanics, Molecular dynamics, Gibbs free energy

1 Introduction

In this chapter, we provide detailed procedures with examples enabling modeling of the enzymatic reaction. The workflow (Fig. 1) starts with the structure of the enzyme in PDB format which is then prepared for MD simulation (methods 2 and 3.1) and subjected to molecular dynamic (MD) simulation (Subheading 3). The analysis of trajectory from MD simulation (Subheading 3.4)

<div align="center">

175

</div>
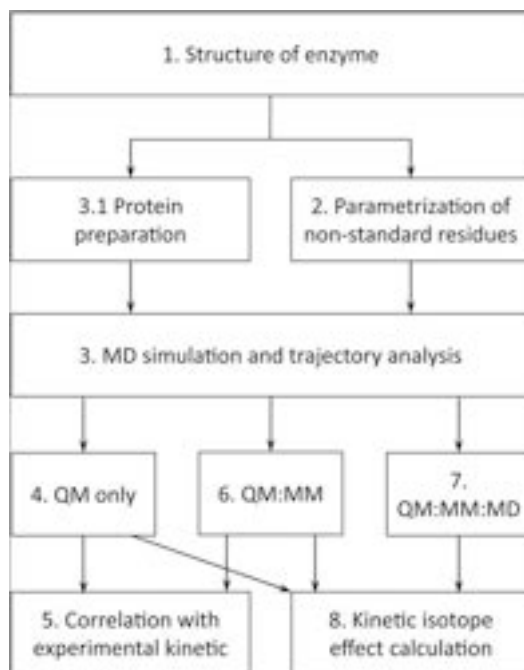
Fig. 1 Schematic workflow of modeling and organization of the chapter

allows the construction of different models like quantum mechanics-only (QM-only) cluster models of the active site (Subheading 4), quantum mechanics:molecular mechanics (QM:MM) models (Subheading 6), or QM:MM models that will be used for QM:MM:MD calculations (Subheading 7). The obtained results from such calculations can be used for correlation with the results of the kinetic experiment (Subheading 5) and prediction of the kinetic isotope effect (Subheading 8).

2 Parametrization of Nonstandard Residues

2.1 Parametrization of an Organic Molecule

In order to establish molecular mechanics (MM) charges for nonstandard molecule, we have two options:

- Find these parameters in an external database (like RESP ESP charge DdataBase R.E.DD.B, https://upjv.q4md-forcefieldtools.org/REDDB/index.php).

- Use *Restrained Electrostatic Potential* (RESP) method to calculate point charges on our own.

In this section, we explain how to apply the RESP method using a simple example: dihydrotestosterone (also known as androstanolone, DHT), which we would like to model into the active site of 3-ketosteroid dehydrogenase from *Rhodococcus*
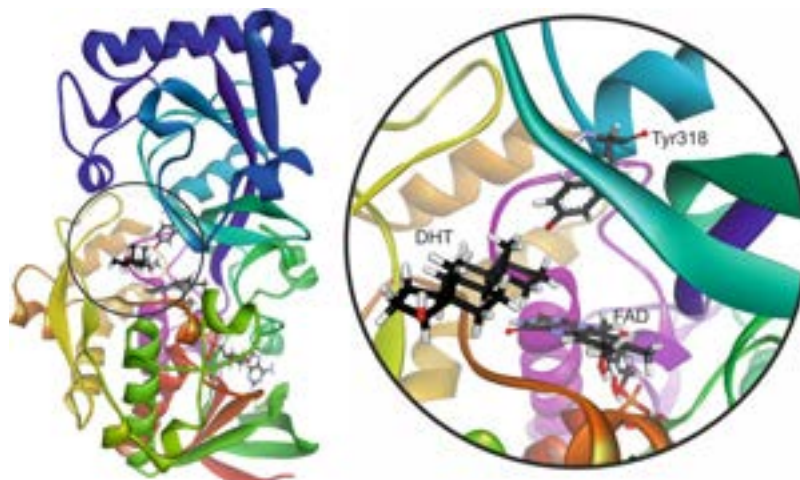
Fig. 2 Structure of 3-ketosteroid dehydrogenase from *R. erythropolis* [1] and magnification of its active site with docked dihydrotestosterone (DHT)
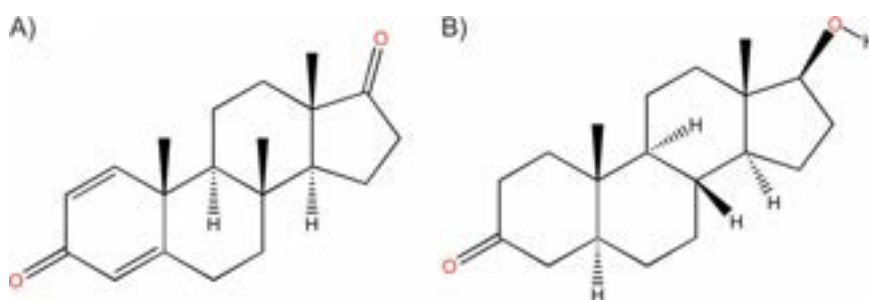
Fig. 3 (**a**) Androsta-1,4-diene-3,17-dione, (**b**) dihydrotestosterone (DHT)

erythropolis (PDB code 4c3y, Fig. 2) [1]. This molecule is not chemically bonded to the rest of the model, it contains only light atoms, and its conformational flexibility is limited. Another factor that simplifies the procedure is the fact that for DHT, we do not expect any pair of atoms to have exactly the same charge due to molecular symmetries. On the other hand, DHT differs a little from the ligand present in the active site in the crystal structure of 4c3y (androsta-1,4-diene-3,17-dione)—*see* Fig. 3. We have to bear this in mind when preparing the model.

The RESP procedure is as follows:

1. Perform geometry optimization of DHT at B3LYP/6-31G(d, p) level of theory. Make sure that the stationary point is found. It is also worth checking whether the conformation adopted by the molecule is obtainable in the active site of the enzyme. Examples of input and output files are located in 2.1. DHT_parametrization/optimization directory in the Electronic Supplementary Material (SI_Revised) on link.springer. com.

2. Using optimized geometry, calculate the electrostatic potential (**auto_gesp.inp**). To do that we have to add to the Gaussian route section iop(6/33 = 2)—to save points and potentials to the external file. Another option we use is iop(6/42 = 6) which specifies the number of the points to be calculated in a single layer around each atom.

3. Use Antechamber from Assisted Model Building with Energy Refinement (AMBER) package to obtain mol2 (all input and output files are in SI in 2.1.DHT_parametrization/antechamber directory): antechamber -i keto.gesp -fi gesp -o keto.mol2 -fo mol2 -c resp.

4. After this step, it is important to validate the results in the **keto.mol2** file. Two points that should be checked are the atomic charges (if nonphysical values are being predicted, we may suspect some error in gesp calculation) and atom types. In this example, we expect that all carbon atoms are of "c3" type except for the ketone group atom, which is of type "c." The oxygen atom from the ketone group is of "o" type, and the one from the hydroxyl group—of type "oh." Almost all hydrogens are of "hc" type except two: H27 which is of type "ho" (hydroxyl group) and H26 which is of type "h1"—this type is a special gaff type for hydrogen atoms bonded to aliphatic C with 1 electron-withdraw group. So in our example, everything is fine.

5. The next step is to generate the **keto.prepi** and **keto.frcmod** files:

```
antechamber -i keto.mol2 -fi mol2 -o keto.prepi -fo prepi
parmchk2 -f prepi -i keto.prepi -o keto.frcmod
```

6. Now we should open and check the keto.frcmod file. If this molecule requires any parameter that is not defined in the GAFF [2] force field, it should be listed there. In keto.frcmod file, there is only one nonempty section:

```
IMPROPER
c3-c3-c -o 1.1 180.0 2.0 Using the default value
```

As we see, Antechamber suggested the default value for the missing improper.

7. Use the Kabsch method to find the best coordinates of the molecule. To use this method, we have to select dihydrosterone atoms that correspond to atoms from crystal structure—in our example, these are all carbon atoms and the oxygen atom of one of the ketone groups. Kabsch algorithm can be used to transform coordinates of dihydrotestosterone in such a way as to minimize root mean square deviation (RMSD) between the coordinates of selected atoms. To do this, we execute the fit.tcl

script in Visual Molecular Dynamics (VMD) (all input and output files are located in SI 2.1.DHT_parametrization/fit).

```
set mol2id [ mol new "keto.mol2" type "mol2" first 0 last -1
step 1 waitfor 1 ]
set templateId [ mol new "anb.xyz" type "xyz" first 0 last -1
step 1 waitfor 1 ]

set templateSel [ atomselect $templateId "all" ]
set indexes "0 1 4 2 3 9 5 6 7 8 18 10 11 12 13 14 15 16 17 19"
set mol2sel [ atomselect $mol2id "index $indexes" ]

set transformation_matrix [measure fit $mol2sel $templateSel]
set mol2fullSel [ atomselect $mol2id "all" ]
$mol2fullSel move $transformation_matrix

animate write mol2 "keto_fitted.mol2" beg 0 end 0 skip 1 $mo-
l2id
exit
```

The script can be executed by typing the following commands into the terminal: vmd -dispdev text -e fit.tcl. The first lines of the script load files into vmd (**keto.mol2** and **anb.xyz**—this file contains only a subset of atoms from the original molecule). After reading the files, the next instructions create two selections—one is the template molecule (anb.xyz) and the second one describes the atoms from keto.mol2 that can be used to calculate RMSD. After fitting, the final coordinates are written in **keto_fitted.mol2**.

2.2   Parametrization of Nonstandard Amino Acid

In this section, we discuss the preparation of the nonstandard amino acids. The procedure is explained using tyrosyl anion as an example. In this type of situation, we once again start with geometry optimization at the density functional theory (DFT) level. To reproduce the electronic characteristic of the amino acid with good quality, we have to prepare a capped version of tyrosyl anion as shown in Fig. 4a. We have to remember that after deprotonation the charge of the molecule is −1 (tyrosyl anion).

After geometry minimization, we generate gesp file in the same way as in Subheading 2.1 (all files are located in SI in 2.2.TYM_-parametrization folder). The next step is to fit charges of atoms to the calculated potential with the RESP program. To do that, we have to take care of two aspects:

1. A sum of charges of the non-cap atoms must be equal to −1. To satisfy this requirement, the charges of the capped atoms must be forced to specific values which add up to zero (Fig. 4b).
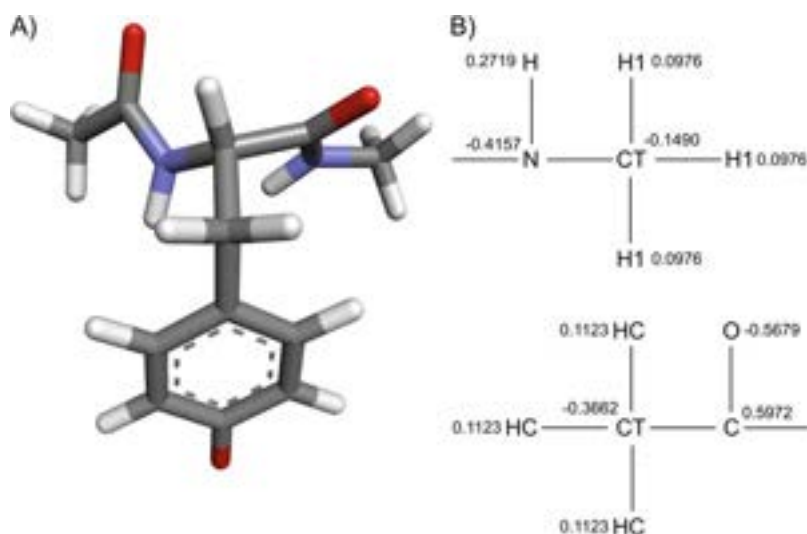
Fig. 4 (**a**) The capped version of tyrosyl anion and (**b**) the point charges of the caps used in the RESP procedure

2. Charges of equivalent atoms (e.g., at H atoms at the CB or equivalent H atoms in the atomic ring) must be forced by resp procedure to be identical.

The simplest way to ensure that those constraints are respected is to generate initial input files for resp using Antechamber (as in Subheading 2.1, **point 3**) and then modify it (examples of Antechamber outputs are located in 2.2.TYM_parametrization/antechamber). We use two of the files generated by antechamber: **ANTECHAMBER.ESP** (electrostatic potentials formatted for resp program) and **ANTECHAMBER_RESP1.IN** (resp input file). In the cntrl section of the file **ANTECHAMBER_RESP1. IN** we add iqopt = 2 option to let resp read the initial charges from an external file. The next step is to freeze the charges of cap atoms. The next section of resp input file looks as follows:

```
 1.0
Resp charges for organic molecule
-1 32
6 -1
6 -1
8 -1
1 -1
1 -1
1 -1
7 0
6 0
6 0
8 0
```

```
6 0
6 0
6 0
6 13
6 0
6 15
6 0
8 0
1 0
1 0
1 0
1 21
1 0
1 23
1 0
1 25
7 -1
6 -1
1 -1
1 -1
1 -1
1 -1
```

The first number (1.0) is the weighting factor for multiconformational RESP (in our example, there is only one conformation). The next line is a comment. Then there is a pair of numbers: molecule charge and number of atoms (−1 32). In the next lines are listed atomic numbers of each atom in the model and integer flag parameters. These parameters determine what resp does with the charge of a specific atom. 0 means normal fitting procedure, −1 means that the charge is not changed and a positive N value means that the charge of this atom must be equal to the charge of the *n*th atom of this molecule. In our example, there are five pairs of such equivalent atoms (four of them were already detected by antechamber): 13–14, 15–16, 21–22, 23–24, and 25–26. The modified resp input is located in 2.2.TYM_parametrization/resp./**resp.IN**.

The external file (**tym.qin**) with initial charges of the atoms used for capping the tyrosyl ions (Fig. 4b) looks as follows:

```
-0.366200 0.597200 -0.567900 0.112300 0.112300 0.112300 0.000000 0.000000
0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000
0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000
0.000000 0.000000 -0.415700 -0.149000 0.271900 0.097600 0.097600 0.097600
```

The charges are provided in the same order as the atoms in our model. The 0.000000 charges are replaced by the RESP procedure with new partial charges fitted to the model electrostatic potential
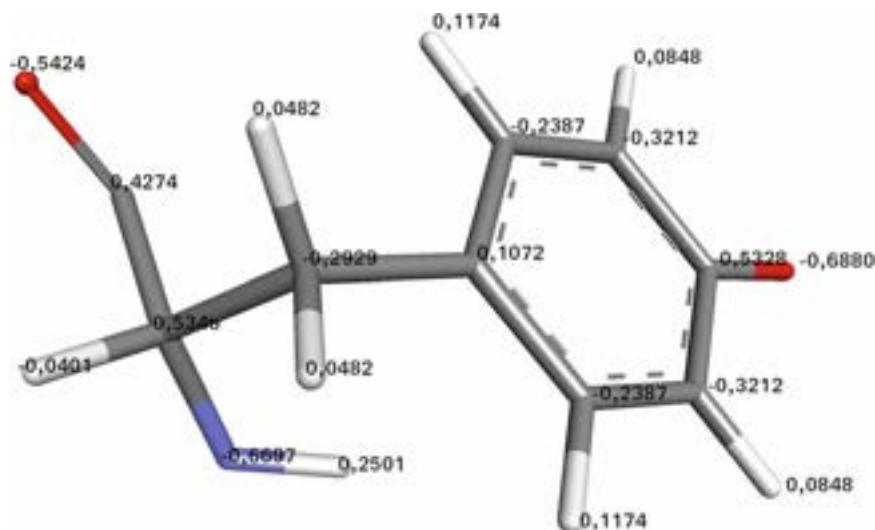
Fig. 5 The MM partial charges fitted to the tyrosyl ion residue by RESP procedure

described in the gesp file, while the charges already provided are fixed by −1 flag (*see* above). Unfortunately, RESP program is format sensitive, and you should always make sure that the input data has been correctly loaded.

Now we can perform resp fitting using command:

```
resp -O -i resp.IN -o tym_resp.out -p tym.punch -t tym_resp.chg
-q tym.qin -e tym.ESP -s tym_resp.esout
```

Fitted charges are written in 2.2.TYM_parametrization/resp./**tym_resp.chg** and are visualized on Fig. 5.

Finally to obtain the mol2 file of the amino acid, we need to replace the last column (charges) of the file 2.2.TYM_parametrization/antechamber/**keto.mol2** generated by Antechamber with new charges generated by RESP (2.2.TYM_parametrization/resp/**keto_new_charge.mol2**).

2.3 Parameterization of the Metal Center

The metal ions have an important role in protein structure and function. These metal ions are usually coordinated with amino acid residues. To study the protein structure and the functional relationship of these metalloproteins using MD simulation, one needs the force field parameters describing the interactions of the metal with other atoms of the system. But the force field parameters are available only for standard amino acid residues and standard organic molecules. Coordination of metal ions with amino acid residues varies from protein to protein, and these coordinated amino acid residues are considered as nonstandard residues during MD simulation if a so-called bonded model is used. The force fields for these metal centers need to be generated to perform an MD simulation. To illustrate such parameterization, here we use thebaine 6-O-

demethylase (T6ODM), a non-heme 2-oxoglutarate/Fe(II)-dependent dioxygenase that catalyzes regiospecific 6-O-demethylation of morphine precursors thebaine (i.e., (5α)-3,6-dimethoxy-17-methyl-6,7,8,14-tetradehydro-4,5-epoxymorphinan) and oripavine ((5α)-6-methoxy-17-methyl-6,7,8,14-tetradehydro-4,5-epoxymorphinan-3-ol) in morphine biosynthesis pathway. The active site of T6ODM (PDB ID: 5O7Y) hosts a Fe ion coordinated with HIS238, ASP240, HIS295, succinate, and an oxo ligand. Although there are many methods for modeling the metal center, here we are discussing the bonded model with the Seminario method to extract force constant from the Hessian matrix (matrix of second-order partial derivatives of the energy with respect to the atomic coordinates) [3].

2.3.1 Protocol in Brief to Parametrize the Metal Center by the Seminario Method

1. Generate the metal center small model (Fig. 6) from the PDB file.
2. Optimize the geometry of the small model (equilibrium bond lengths and valence angles made by the metal are used as reference values in the force field).
3. Calculate harmonic frequencies for the optimized small model to get the Hessian matrix.
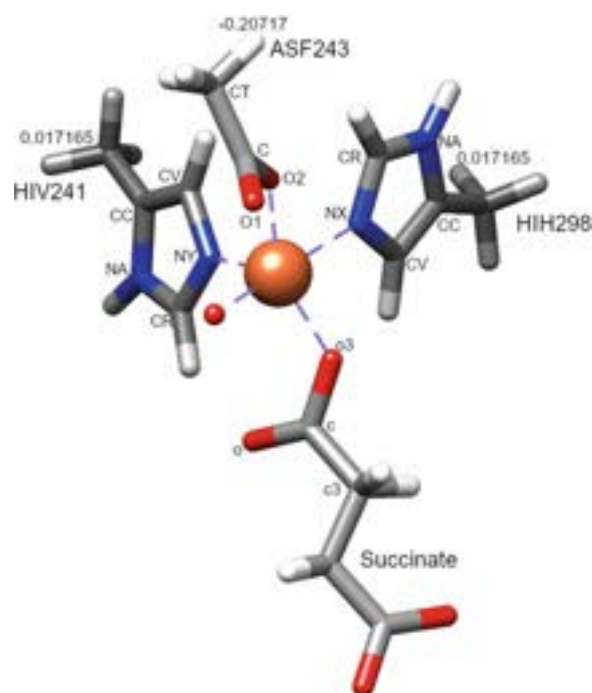
Fig. 6 Optimized T6ODM metal center with labeled point charges on H replacing CA atoms and atom types for atoms considered to make AS. frcmod parameter file

4. From the formatted checkpoint file containing the Hessian extract the missing force constants needed to prepare the **AS. frcmod** parameter file.

5. Calculate the electrostatic potential (ESP) for the optimized small model to get a **sm.gesp** file containing the ESP values computed for an array of points around the model (*see* 1.3. Fe_parametrization/gesp).

6. Fit the atomic partial charges of atoms with the RESP procedure of AMBER using the **sm.gesp** as an input file.

7. Generate prep files (containing topology information, atom types and atomic charges) for nonstandard residues.

2.3.2 Protocol in Detail

1. Prepare the **sm.xyz** coordinate file for Fe center (small model):
   In our PDB (PDBID: 5O7Y) file, there were missing residues from 38 to 40. They were regenerated by the loop modeling procedure available in MODELLER. The oxygen atom of 1,2-ethanediol coordinated with the metal site was replaced by the oxo ligand, the rest of the atoms of 1,2-ethanediol were removed, and Ni(II) was replaced with Fe(IV).

   Add hydrogen atoms to the PDB file using the H++ web server (http://biophysics.cs.vt.edu/).

   Then load the modified PDB file *t6odm.pdb* in VMD, extract the Cartesian coordinates of the metal center (*see* **Note 1**).

   VMD: Graphics→Graphical Representation→Selected Atoms

   ```
   resid 241 243 298 368 369 370 and (sidechain or name CA)
   ```

   and save it in xyz format (*sm.xyz*). CA atoms of amino acid residues are replaced with hydrogen atoms using XYZViewer, in a similar manner as shown for the construction of the cluster model (*see* below).

   (a) Select the C atom which is needed to be replaced with H
       XYZViewer: File→Load structure→select *sm.xyz* file and load it→select the C by clicking on it, also note down its atom number, e.g., 20.

   (b) Replace C atom with H atom
       XYZViewer: Edit→Show matrix→"Matrix Editor" opens→scroll down to atom number 20 and click on C atom→"Periodic Table" opens, click on H atom→selected the C atom that is replaced with H→close the Matrix Editor.

   (c) Edit bond length between replaced H and C atom to which it is connected

XYZViewer: Edit→Edit in internal coords→a single atom→select the atom to edit, and three more atoms to define its internal coordinates→"Set coordinate window" opens→set the "Bond length" 1 Å→click on OK button.

2. Geometry optimization of a small model

For geometry optimization, we need a Gaussian input file.

```
#sm-opt.com

%chk=sm-opt.chk
%mem=6GB
# ub3lyp/lanl2dz 5d scf=(xqc,maxcycle=350) nosymm
EmpiricalDispersion=GD3BJ Opt=modredundant
----a blank line----
sm optimization
-----a blank line-----
0 5
XYZ coordinates
-----a blank line-----
X 1 F
-----two blank lines-----
```

%chk=sm-opt.chk: This line is to specify the name of the checkpoint file (*sm-opt.chk*), the results of the calculation are written in this file during calculation and it can be further used as a starting point of the second calculation.

%mem=6GB: Maximum amount of memory to be used for calculation.

 # ub3lyp/lanl2dz: We are using the *uB3LYP/lanl2dz* [4] level of theory. A different level of theory can be used, but the same level of theory has to be used for geometry optimization and force constant (Hessian) calculations. *5d* is for using five d-type orbitals (as opposed to six Cartesian d functions) *SCF= (xqc,maxcycle=350)*, for SCF procedure with a maximum of 350 iterations, *nosymm* to prevent Gaussian from trying to use symmetry during the calculation. *EmpiricalDispersion=GD3BJ* [5], to enable empirical dispersion (D3) with the Becke–Johnson dumping. *Opt=modredundant*, to freeze the Cartesian coordinates of hydrogen atoms that replace CA atoms of amino acid residues to prevent any dramatic changes in the overall geometry of the small model during optimization. Keywords can be written in more than one line. This section is terminated by a blank line.

sm optimization: Title section—a brief description of calculation, this is also terminated by a blank line.

0 5 and XYZ coordinates: Molecular specification section to specify charge (0) and spin multiplicity (quintet) of the model followed by xyz Cartesian coordinates.

X 1 F: Here we can specify the atom(s) with frozen coordinates by the *modredundant* keyword.

Command to run Gaussian:

```
g16 sm_opt.com
```

After geometry optimization, a small model should be checked for the integrity of the complex (i.e., if any of the coordination bonds did not get broken and the geometry of the model is resembling that in the crystal structure). Sometimes the accuracy of a metal site geometry in a protein is poor, e.g., due to ambiguity of electron density or lack of reference structures, so some differences between the experimental and computational structures may occur. Also one has to remember that the complex is optimized without explicit constraints imposed by the protein. Sometimes the enzymes strain the coordination of the metal active site to influence its electronic structure [6].

3. Frequency calculation for optimized small model

Frequency calculation Gaussian input file (1.3.Fe_parametrization/optimization/**sm-opt.fq.com**):

```
#sm-opt.fq.com

%oldchk=sm-opt.chk
%chk=sm-opt.fq.chk
%mem=6GB
# ub3lyp/lanl2dz 5d scf=(xqc,maxcycle=350) nosymm Guess=Read
Geom=AllCheckpoint
EmpiricalDispersion=GD3BJ Freq
```

We use keyword *Guess=Read* to take orbitals from the previous optimization checkpoint file, *sm-opt.chk,* as an initial guess, and *Geom=AllCheckpoint* to retrieve from it also molecular geometry, charge, and multiplicity of the model. *Freq* is used to tell Gaussian that this is input for frequency calculation. After finishing the frequency calculation, we format the checkpoint file *sm-opt.fq.chk*:

```
formchk sm-opt.fq.chk sm-opt.fq.fchk
```

which is suitable for various visualizing programs, e.g., XYZViewer. From this formatted checkpoint, bond and angle values as well as force constant are extracted to construct the *frcmod* parameter file.

4. Make *frcmod* parameter file
   (a) Extract bond force constant

XYZViewer: File→Load structure→select *sm-opt.fq. fchk* and load it→Calculate→Forces from hessian (beta)→Forces→click the two atoms for which bond length and force constant has to be measured→"Force calculation" window opens displaying values of bond length and its force constant for that particular bond.

As an alternative to the XYZViewer program, to extract force constants, one can use Hess2FF software developed by Ulf Ryde [7].

Bonds to be measured for a metal center (Fig. 6):

- Between metal and atoms coordinated to a metal.
- (Optionally) between coordinated atom and atom connected with it.

(b) Extract angle force constant

XYZViewer: File→Load structure→select the formatted checkpoint file *sm-opt.fq.fchk* and load it→Calculate→Forces from hessian(beta)→Angle force→click the three atoms for which angle force constant has to be measured→"Force calculation" window opens displaying values of angle and its force constant for that particular angle.

Angles to be measured for the metal center:

- Between three atoms one must be metal.
- (Optionally) between three atoms, one of them must be an atom coordinated with metal.

Atom types that need to be provided in the *AS.frcmod* file are labeled in Fig. 6. Now make two columns in the Excel sheet: r for bond length and k for force constant. After measuring values for all the bonds, divide column k by 2 (this is due to the difference in assumed formula, i.e., it is $K(r-r_{eq})^2$ for Amber and $1/2\ k(r-r_{eq})^2$ for XYZViewer).

Construct the **AS.frcmod** file for the metal center which looks like below:

```
MASS
NX 14.01 0.530
FE 55.85

BOND
FE-ox 365.0 1.642
NY-FE 61.0 2.110

ANGL
NY-FE-o3 41.0 90.0
FE-NX-CV 107.0 125.0
```

```
DIHE
ox-FE-NY-CV 3 0.00 0.0 3.0
O2-FE-o3-c 3 0.00 0.0 3.0


NONB
NX 1.8240 0.1700
FE 1.409 0.0130
```

We are not measuring dihedral parameters, rather when we use our *AS.frcmod* file (without dihedral parameter) to generate topology and initial coordinate files by the leap program. The output of the first leap run contains a list of the missing dihedral angle parameters. So, from there we take those dihedral angles specifications and add them to our frcmod file while keeping the values of their barrier heights set to zero, as in the above *AS.frcmod* file.

5. ESP (electrostatic potential) calculation using Gaussian (1.3. Fe_parametrization/gesp/*sm-opt.resp.com*)

```
#sm-opt.resp.com


%oldchk=sm-opt.chk
%chk=sm-opt.resp.chk
%mem=6GB
# ub3lyp/lanl2dz 5d scf=(xqc,maxcycle=350) nosymm Guess=Read
EmpiricalDispersion=GD3BJ Pop=(MK,ReadRadii)  IOp(6/33=2) iop
(6/41=10) iop(6/50=1)
----molecular specification, and a blank line-----
Fe 1.409
-----a blank line-----
sm.gesp
-----a blank line-----
sm.gesp
-----two blank line-----
```

pop=(MK, ReadRadii): To calculate atomic charges according to Merz-Singh-Kollman scheme, ReadRadii to read atomic radius for Fe from the input file.

iop(6/33=2): To write coordinates and ESP potential values.

iop(6/41=10): Specify the number of layers at which the ESP potential is calculated, i.e., 10.

iop(6/50=1): To write Gaussian output file *sm.gesp*, which is further used for charge fitting by Antechamber.

After *xyz* Cartesian coordinates, we specified van der Waals (vdW) radii of Fe and the name of the Gaussian *sm.gesp* output file to write ESP data.

6. RESP charge fitting using Antechamber (*see* /1.3.Fe_parame-trization/resp for files)

    Run the Antechamber [8] using *gesp* file and generate input file for RESP fitting.

```
antechamber -i sm.gesp -fi gesp -o sm.mol2 -fo mol2 -c resp -j 5
```

This step also generates the *quot* (point charges for all the atoms) file. *quot* file is further modified to the *qin* file by making charges of all the atoms zero except for H atoms replacing CA. Since here we are generating a prep file for AMBER MD simulation, we need to make sure that the total charge of the system is an integer. For this purpose, we sum the backbone atoms' (N, H, CA, HA, C, O) charges of respective amino acid residues (in the metal center), which are available in the AMBER prep file for standard residues, and ascribe these total charges to H (Fig. 6) atoms replacing the CA. This *qin* file is further used by RESP in the next step of charge fitting.

A two-step charge fitting procedure by RESP is the next step.

Step 1: All charges of atoms are allowed to vary except those for hydrogen atoms replacing the CA atoms. Charges of these atoms (Fig. 6) are read from the qin file.

```
resp -O -i sm_RESP1.IN -o sm_RESP1.out -e ANTECHAMBER.ESP -p
sm_RESP1.punch -q qin -t sm_RESP1.qout
```

Step 2: Final fit with constraints on charges of the degenerate atoms. For example, one of the hydrogen atom of the methyl group in an amino acid model is a hydrogen atom replacing alpha carbon and the other two hydrogen atoms, which are bound to Cβ, are chemically equivalent (degenerate), and thus, these two have the same charge.

```
resp -O -i sm_RESP2.IN -o sm_RESP2.out -e ANTECHAMBER.ESP -p
sm_RESP2.punch -q sm_RESP1.qout -t sm_RESP2.qout
```

7. Generate prep files (1.3.Fe_parametrization/paramerization)

    From 1.3.Fe_parametrization/resp.**/sm_RESP2.qout**, point charges of Fe, oxo, and succinate and side chain atoms of HIS241, ASP243, HIS298 are extracted to construct their respective *prep* files. First, generate prep files for respective nonstandard amino acids from AMBER prep file *all_a-mino03.in* and then point charges (last column) of side chain atoms are replaced with point charges obtained after RESP fitting. Below is a prep file for ASP243:

```
#ASF.in
0 0 2
```

```
ASF
asf.db4
 ASF INT 1
 CORR OMIT DU BEG
 0.00000
 1 DUMM DU M 0 -1 -2 0.000 0.000 0.000 0.00000
 2 DUMM DU M 1 0 -1 1.449 0.000 0.000 0.00000
 3 DUMM DU M 2 1 0 1.522 111.100 0.000 0.00000
 4 N N M 3 2 1 1.335 116.600 180.000 -0.558201
 5 H H E 4 3 2 1.010 119.800 0.000 0.319676
 6 CA CT M 4 3 2 1.449 121.900 180.000 0.007225
 7 HA H1 E 6 4 3 1.090 109.500 300.000 0.082375
 8 CB CT 3 6 4 3 1.525 111.100 60.000 0.718039
 9 HB2 HC E 8 6 4 1.090 109.500 300.000 -0.140144
 10 HB3 HC E 8 6 4 1.090 109.500 60.000 -0.140144
 11 CG C B 8 6 4 1.527 109.470 180.000 0.281945
 12 OD1 O2 E 11 8 6 1.260 117.200 90.000 -0.451834
 13 OD2 O2 E 11 8 6 1.260 117.200 270.000 -0.664639
 14 C C M 6 4 3 1.522 111.100 180.000 0.443199
 15 O O E 14 6 4 1.229 120.500 0.000 -0.501445

IMPROPER
 -M CA N H
 CA +M C O
 CB OD1 CG OD2

DONE
STOP
```

For Fe, oxo, and succinate, we can use antechamber to generate mol2 files then prep files from mol2 and further replace the point charges with the above-calculated charges. Here we need to construct the two separate prep files for HIS241 and HIS298 because the coordination orientation of both residues with Fe is unique, HIS241 is in an axial, whereas HIS298 is in an equatorial position, so there is a need to distinguish them as their nitrogen atoms coordinating Fe make valence angles of very different values, e.g., NX-Fe-oxo—178.7° and NY-Fe-oxo—93.8°.

3   MD Simulation

Before QM:MM calculation, we usually need to do a MD simulation for the enzyme-substrate complex using the AMBER MD simulation package.

Protocol in brief for performing an MD simulation using AMBER:

1. Generate a prmtop file (description of molecular topology and force field parameters) and an inpcrd file (atom coordinate, velocity information, and periodic box dimensions).
2. Optimize the geometry of the model at 0 K (energy minimization).
3. Heat the system and equilibrate its density.
4. Perform production run.

3.1 Protein Preparation

Generate the prmtop and inpcrd files

Editing in the PDB file

Calculate the pK_a of amino acid residues using PDB2PQR (http://server.poissonboltzmann.org/pdb2pqr), then change the name of HIS residues to HID (delta protonation), HIE (epsilon protonation), HIP (positively charged HIS) in the PDB file accordingly. Since HIS241, ASP243, HIS298 coordinate metal center and are nonstandard residues, thus, we rename them to HIV241, ASF243, and HIH298, respectively. This PDB file has also previously docked two molecules of substrate thebaine. Since we already generated prep and parameter files for the metal center, we generate these files for thebaine using Antechamber, as is usually done for organic molecules (*see* Subheading 2.1).

Run leap program to generate *prmtop* and *inpcrd*

tleap -f leap.in (files in 3.MD-simulation/3.1 leap)

```
#leap.in

source leaprc.protein.ff03.r1
source leaprc.water.tip3p
loadamberparams AS.frcmod
loadAmberPrep HIV.in
loadAmberPrep ASF.in
loadAmberPrep HIH.in
loadAmberPrep Fe.in
loadAmberPrep oxo.in
loadAmberPrep sin.in
loadAmberPrep thb.in
mol = loadpdb t6odm.pdb
bond mol.residue1.atom1 mol.residue2.atom2
solvateBox mol TIP3PBOX 10.0
addions mol Na+ 0
charge mol
saveAmberParm mol t6odm.prmtop t6odm.inpcrd
charge mol
quit
```

To create the *t6odm.prmtop* and *t6odm.inpcrd* files, we use ff03 force field [9, 10] for protein (other protein force fields can also be used), and above-generated force field parameters (*AS.frcmod*) and prep (HIV.*in*, *ASF.in*, *HIH.in*, *Fe.in*, *oxo.in*, *sin.in*, *thb.in*) files for nonstandard atom types and residues. Since we perform explicit water MD simulation, we add TIP3P water to solvate our system and also add counter ions to neutralize the net charge. Because we adopted the bonded model for the metal center, we form bonds between Fe and coordinated atoms of respective residues using *bond* command.

Periodic boundary: When we add solvent to solute (protein in our case), we make a finite box size in which some of the solvent molecules are present at the boundary between solute and solvent, some in the bulk solvent, and some at the edge of the box. During MD simulation, water molecules at the edge may move out of the box. To avoid such effect, due to finite box size, we make the system infinite by applying the periodic boundary condition, i.e., the unit cell is replicated in all three directions to fill the space with its images. As a result, if a molecule leaves the unit cell through one wall, it reappears on the opposite side. Thus, through periodic boundary conditions, we construct an infinite system consisting of images of a unit cell.

3.2 Geometry Minimization

There is a high chance that the PDB file we are using for MD simulation has some close contacts, nonoptimal residue conformation, the sub-optimal position of added protons. Also, the addition of an explicit solvent can create close contacts between the solute and solvent. Therefore, the geometry minimization is needed to deal with those problems—otherwise, MD simulation will be unstable due to van der Waals clashes. The minimization procedure we adopt consists of three steps (*see* 3.MD-simulation/3.2. MM for files):

#Step1: Minimization of water and Na^+ ions

```
#min1.in

minimization of water and Na+ ions
&cntrl
 imin=1,
 ntpr=50,
 ntb=1, cut=10.0,
 ntr=1,
 maxcyc=5000, ncyc=500,
 restraint_wt=500.0, restraintmask=':1-372'
 /
```

imin=1: To do minimization.

ntpr=50: Energy information is printed to AMBER output file at every ntpr steps.

ntb=1: To apply constant volume periodic boundary condition.

cut=10.0: Cutoff distance for non-bonded calculation (should not be <8).

ntr=1: Turn on positional restraints on mask atoms.

maxcyc=5000: Maximum number of minimization cycles.

ncyc=500: For the first 500 steps of maxcyc, perform for steepest descent minimization, then switch to the conjugate gradient.

restraint_wt=500.0: Weight (harmonic force constant) for positional restraint.

restraintmask=':1-372: Mask atom for applying restraint.

There are many algorithms for minimization. The commonly used are the steepest descent which quickly removes larger strains but converges slowly and the second method is the conjugate gradient, which converges faster than the steepest descent. Thus, a combination of both is used to remove larger strains in the initial phase (500 cycles) and then faster convergence in the second phase (next 4500 cycles).

```
pmemd.MPI -O -i min1.in -o min1.out -p t6odm.prmtop -c t6odm.
inpcrd -ref t6odm.inpcrd -r min1.rst
```

Since we are using positional restraint on mask atoms so we are giving *t6odm.inpcrd* as a reference structure (-ref) to which we can impose restraint.

#Step2: Minimization of water and Na$^+$ ions with decreased restraint weight using restart file *min1.rst* from #Step1

#min2.in

minimization of water and Na+ ions with decreased restraint weight

```
&cntrl
 imin=1,
 ntpr=50,
 ntb=1, cut=10.0,
 ntr=1,
 maxcyc=5000, ncyc=500,
 restraint_wt=10.0, restraintmask=':1-372'
/
pmemd.MPI -O -i min2.in -o min2.out -p t6odm.prmtop -c min1.rst
-ref min1.rst -r min2.rst
```

#Step3: Minimization of the whole system using restart file *min2.rst* from #Step2

#min3.in

minimization of whole system

```
&cntrl
 imin=1,
 ntpr=50,
 ntb=1, cut=10.0,
 ntr=1,
 maxcyc=10000, ncyc=500,
/
pmemd.MPI -O -i min3.in -o min3.out -p t6odm.prmtop -c min2.rst
-r min3.rst
```

3.3 Heating, Density Equilibration and Production

3.3.1 After Minimization, Heat the System from 0 to 300 K (3. MD-Simulation/2.3. MD)

```
#NVT.in
Heating for 100 ps
&cntrl
 imin=0
 ntx=1, irest=0,
 ntpr=500, ntwr=500, ntwx=5000, iwrap=1,
 ntc=2, ntf=2, ntb=1, cut=10.0,
 ntr=1,
 nstlim=50000, dt=0.002, nscm=1000,
 tempi=0.0, temp0=300.0, ntt=3, gamma_ln=0.05,
&end
restrain backbone of the solute
1.0
FIND
* * M *
SEARCH
RES 1 372
END
END
```

imin=0: To perform MD simulation.

ntx=1: Read coordinate from restart coordinate file.

irest=0: It is not a restart of the MD simulation, velocity in the input coordinate file are ignored.

ntpr=50: Energy information is printed to mdout file at every ntpr steps.

ntwr: At an every ntwr step, restart file is written.

ntwx: At an every ntwx step, coordinates are written to mdcrd file.

iwrap=1: Wrap the coordinate to primary box while writing the restart and mdcrd file.

ntc=2: Use SHAKE to constrain bonds involving hydrogen atoms for increased efficiency of MD simulation.

ntf=2: Force evaluation, should be used with ntc=2.

ntb=1: To apply constant volume periodic boundary condition.

cut=10.0: Distance cutoff [Å] for non-bonded interactions.

nstlim=50000: Number of MD steps to perform.

dt=0.002: Time step [ps].

nscm=1000: Remove the translational and rotational motion of the whole system (centre of mass) at every nscm steps.

tempi=0.0: Initial temperature of the system.

temp0=300.0: Target temperature [K] of the system after heating.

ntt=3: Turn on Langevin dynamics.

gamma_ln=0.05: Collision frequency $[\text{ps}^{-1}]$ when ntt=3.

The Berendsen temperature coupling (ntt=1) is a weak coupling; it cannot ensure that temperature is even all over the system, which can lead to the hot solvent cold solute. Meanwhile, the Langevin thermostat (ntt=3) with collision frequency is more efficient for equilibrating the temperature and is thus recommended. There are different thermodynamic ensembles used during MD simulation. The isobaric-isothermal NPT ensemble is closer to the laboratory condition, but the heating from the low temperature at NPT for temperature equilibration is inaccurate. Thus, heating is performed using the canonical NVT ensemble.

```
pmemd.MPI -O -i NVT.in -o NVT.out -p t6odm.prmtop -c min3.rst
-ref min3.rst -r NVT.rst
```

3.3.2 Density Equilibration

Heating at NVT creates void space in the solvent and for a periodic boundary, NPT is the only way to equilibrate the density. Thus, it is necessary to bring the system to the final temperature, using NVT with ntt=3, and then switch to NPT, with ntt=3, for density equilibration.

```
#NPT.in
```

   Equilibration for 1 ns

```
&cntrl
 imin=0
 ntx=5, irest=1,
 ntpr=500, ntwr=500, ntwx=5000, iwrap=1,
 ntc=2, ntf=2, ntb=2, pres0=1.0, ntp=1, taup=2.0, cut=10.0,
 nstlim=500000, dt=0.002, nscm=1000,
 tempi=300.0, temp0=300.0, ntt=3, gamma_ln=0.05,
&end
restrain backbone of the solute
1.0
FIND
* * M *
SEARCH
RES 1 372
END
END
```

ntx=5: Option to read coordinate, velocity, and box information from a restart file.

irest=1: Restart the simulation after reading coordinate and velocity from a previously saved restart file.

ntb=2: To apply constant pressure periodic boundary condition.

ntp=1: MD with isotropic position scaling.

pres0=1.0: Reference pressure [bar] when ntp>0.

taup=2.0: Pressure relaxation time [ps] when ntp>0.

```
pmemd.MPI -O -i NPT.in -o NPT.out -p t6odm.prmtop -c NVT.rst -r
NPT.rst
```

4. Production run

Once the temperature and pressure are equilibrated, a longer production run is performed for T = const., p = const. conditions (NPT ensemble), which are meant to mimic laboratory conditions.

```
#NPT.in
```

*3.3.3 Production Run
5 ns*

```
&cntrl
 imin=0
 ntx=5, irest=1,
 ntpr=5000, ntwr=5000, ntwx=5000, iwrap=1,
 ntc=2, ntf=2, ntb=2, pres0=1.0, ntp=1, taup=2.0, cut=10.0,
 ntr=1,
 nstlim=2500000, dt=0.002, nscm=1000,
 tempi=300.0, temp0=300.0, ntt=3, gamma_ln=0.05,
/
```

The production run is a long MD simulation, and to get an unbiased sampling, the whole production run should be broken into many small consecutive runs rather than attempted in a single long MD run. For example, if we have to run 100 ns, then instead of doing a single 100 ns run, it should be divided into, e.g., 20 consecutive runs of 5 ns each.

```
pmemd.MPI -O -i prod.in -o prod.out -p t6odm.prmtop -c NPT.rst
-r prod.rst
```

*3.4 Analysis of MD
Simulation*

3.4.1 RMSD Analysis

Before we proceed with further analysis of the MD trajectory, we have to check if the trajectory has converged. One method to check the convergence is to look at the RMSD profile along the trajectory. Here, as a reference structure for calculating RMSD, we use the energy minimized structure **min3.pdb** obtained at the end of this section. Files to conduct RMDS analysis are available at 3.MD-simulation/3.4 Analysis of MD simulation/rmsd.

```
#rmsd.in
```

```
trajin prod-1.mdcrd
trajin prod-2.mdcrd
trajin prod-3.mdcrd
strip :Na⁺,Cl⁻,WAT
autoimage origin
reference min3.pdb
rms reference mass out t6odm.rms time 5.0 :1-367@N,CA,C


cpptraj -p t6odm.prmtop <rmsd.in> rmsd.out
```

Above input file uses three MD trajectory files to calculate the RMSD.

strip: Na^+,Cl^-,WAT: Strips all ions and water molecules so that process is faster.

autoimage origin: Centers the periodic boundary trajectory on the first molecule and re-image.

reference min3.pdb: Reference structure to calculate RMSD.

rms reference mass out t6odm.rms time 5.0 :1-367@N,CA,C: Calculate the mass-weighted RMSD of protein backbone atoms using reference structure and write to *t6odm.rms* output file.

Once we get the 3.MD-simulation/3.4 Analysis of MD simulation/rmsd /**t6odm.rms** file, by plotting it, we can check the convergence of the trajectory. Initially, there is a major fluctuation in RMSD, which means changes in the molecule's conformation, and then the RMSD starts to level off. This indicates that our trajectory has converged, i.e., there are no further major changes in conformation; now we can proceed with further analysis.

3.4.2   Cluster Analysis

After MD simulation, we need to retrieve the representative structure(s) of our protein from our MD trajectory. In this tutorial, this is achieved by clustering analysis. This procedure compares the geometries of selected atoms or residues (by calculating RMDS) and generates the clusters of MD frames of similar conformations. Structures belonging to one cluster are similar and vary from the other clusters. Differences in structures which is the basis of clustering are determined by coordinate RMSD. There are many algorithms for clustering, here we are using DBScan. All files are available in SI at 3.MD-simulation/3.4 Analysis of MD simulation/clustering.

Steps in cluster analysis:

1. Estimate the epsilon and minpoint parameters.
2. Perform actual clustering using the above-determined parameters.

Estimation of epsilon and minpoint parameters

To do the clustering analysis using DBScan, we need two parameters: epsilon (distance cutoff to form a cluster) and minpoint (minimum number of points to form a cluster). These parameters can be estimated from a so-called k-dist plot, which can be generated by CPPTRAJ.

```
#k-dist.in

trajin prod-1.mdcrd
trajin prod-2.mdcrd
trajin prod-3.mdcrd
autoimage origin
strip :WAT, Na⁺, Cl-
cluster c1 dbscan kdist 1 rms :368@FE:369@O:370@C1,O1,
O2:372@C19,O2,C6
cluster c2 dbscan kdist 2 rms :368@FE:369@O:370@C1,O1,
O2:372@C19,O2,C6
cluster c3 dbscan kdist 3 rms :368@FE:369@O:370@C1,O1,
O2:372@C19,O2,C6
cluster c4 dbscan kdist 4 rms :368@FE:369@O:370@C1,O1,
O2:372@C19,O2,C6
cluster c5 dbscan kdist 5 rms :368@FE:369@O:370@C1,O1,
O2:372@C19,O2,C6


cpptraj -p t6odm.prmtop <k-dist.in> k-dist.out
```

Cpptraj will make five kdist (**Kdist.1.dat**, **Kdist.2.dat**, **Kdist.3.dat**, **Kdist.4.dat**, and **Kdist.5.dat**) files from three MD trajectories based on five different values of kdist (Å) and RMSD of mask atoms. We considered these atoms as mask atoms because their mutual arrangement is pivotal for the progress of the enzymatic reaction, and we wanted to get a good starting point for the next step of QM:MM calculations.

By plotting these kdist, we can estimate epsilon and minpoint. Initially, the slope of the curve is steep and then becomes flat (increase in density). The distance around which the curve starts to flatten is taken as epsilon value. The kdist at which the shape of the curve does not change too much is considered as minpoint. Generally, after kdist = 4, the shape of the curves does not change too much. But it is suggested to do some clustering experiments with different kdist values. Further, clustering based on different kdist values starting from 1…, the number of clusters decreases with the increase in kdist value. Thus, it helps us to choose appropriate epsilon values for clustering so that we can get enough clusters to get a good representative structure for our MD trajectory.

Clustering

```
#cluster.in

trajin prod-1.mdcrd
trajin prod-2.mdcrd
autoimage origin
strip :WAT, Na+, Cl-

cluster C0
dbscan minpoints 4 epsilon 0.3 sievetoframe
rms :368@FE:369@O:370@C1,O1,O2:372@C19,O2,C6
sieve 10 random
out cnumvtime.dat sil Sil
summary summary.dat
info info.dat
cpopvtime cpopvtime.agr normframe
repout rep repfmt pdb
singlerepout singlerep.nc singlerepfmt netcdf
avgout Avg avgfmt restart
```

Clustering parameters
minpoint = Minimum number of points to form cluster = 2.
epsilon = Distance criteria to form a cluster = 0.3.
Distance matrix options
rms = Use RMSD of mask atom to calculate the distance between frames.

sieve 10 = Every 10 frames are considered for initial clustering then sieved frames added to those clusters, *random* keyword used to select random frames.

Output options
out cnumvtime.dat: This file contains information for cluster number vs. time. Noise frames are assigned to a value of -1.

summary summary.dat: Overall summary of clustering (number of clusters formed).

info info.dat: It contains all the clustering results. This file also has information on DBI and pSF values, which can be later used, e.g., to compare results obtained with different clustering algorithms. Lower DBI and higher pSF values indicate better clustering.

```
cpptraj -p t6odm.prmtop <cluster.in> cluster.out
```

Running the above command generates the following output files. The **summary.dat** (3.MD-simulation/3.4 Analysis of MD simulation/clustering) gives information about the number of clusters formed. Representative structures of each cluster are also

generated, which we can use for our further studies like structure analysis, and as a starting structure for QM or QM:MM calculations. We can also use **cnumvtime.dat** (3.MD-simulation/3.4 Analysis of MD simulation/clustering) to once again check if the simulation converged. We should observe a stable population of the clusters for the phase of simulation where we have observed stable RMSD.

4   Cluster-Based QM Calculations

The first step in the investigation of the reaction pathways is the construction of the appropriate model representing the enzyme active site. Generally, there are two methods to approach this problem. The first is to use the whole protein with a significant portion of the solvent where only active site and residues with reagents are treated with QM level of theory while for the rest of the model, one applies simple molecular mechanics. This approach, known as QM:MM, is described in detail in Subheadings 6 and 7. In the other approach, one cuts from the whole protein-only residues of the active site and regents and applies the QM calculation for the description of all selected atoms. Such fragmentary models are known as "cluster models" (and the link to the statistical clustering analysis described in Subheading 3.4.2 is solely semantic). Such a cluster model of enzyme active site can be based on:

- Crystal structure of the enzyme, preferably with a substrate or substrate analog bound in the active site.
- Structure of enzyme–substrate complex derived from crystal structure that was relaxed during MD simulation, followed by minimization in MM.
- MD-relaxed structure of enzyme–substrate from the homology model.

Alternatively, a substrate can be substituted with a product, and in such a case, the reaction can be followed in reverse (i.e., starting from enzyme–product complex).

4.1 Construction of the Cluster Model

The construction of the cluster model should be done taking the following guidelines into account:

1. All reagents should be included in the model.

2. All residues that are known to be involved in the reaction are included in the model.

3. All solvent molecules that are known to interact via H-bonds with reagents or important protonable residues are included in the model.

4. In case the active site contains the metal ions, all its ligands are introduced into the model ensuring representative electronic surroundings of the metal complex.

5. The second shell of residues that form H-bonds or strong electrostatic interactions with primary residues can be included. Hydrophobic residues that surround the active site can also be included.

6. The residues are trimmed by cleavage at the nonpolarized C–C bond. The dangling bonds are saturated with H atoms.

7. The position of the capping H atoms and adjacent C atoms are frozen by Cartesian constraints.

8. If parts of the enzyme main chain have to be included (e.g., to model oxo anion hole), the cleavage site of the peptide backbone is end-capped by CH_3 groups yielding $-NH-CH_3$ or $OC-CH_3$. The position of C atoms and H atoms along the path of the main chain should be constrained.

9. The position of cofactors (e.g., NADH, FAD, heme) is constrained at the atoms that form covalent bonds with the protein, H-bonds, or other detectable interactions inside the protein (e.g., salt-bridges).

10. Constrains should allow conformational changes that would normally be accessible to residues. However, as the outer shell of protein is removed, sometimes it is necessary to introduce additional constraints preventing artificial changes of geometry during geometry optimization.

11. The size of the cluster model should be kept reasonably low (100–200 atoms) but can be extended if a lower level of theory is used (e.g., semi-empirical methods).

This procedure is demonstrated based on the active site of short-chain dehydrogenase/reductase (SDR) from *Aromatoleum aromaticum* (*S*)-1-phenylethanol dehydrogenase (Fig. 7a, *S*-PEDH) [12, 13]. The SDRs catalyze reversible NAD^+-dependent reaction of alcohols oxidation to ketones/aldehydes (pH optimum at neutral to slightly basic) or NADH-dependent reduction of ketones/aldehydes to alcohols (pH optimum in the range of 5–6).

The industrially interesting reaction, enantioselective reduction of ketones to chiral alcohols, starts with the binding of NADH and ketone. The carbonyl group of the ketone is bound by two H-bonds, formed with Ser (in *S*-PEDH Ser141) and Tyr (Tyr154). This ensures that the proper face of the carbonyl sp^2 atom is turned toward the NADH ring which transfers hydride during the reduction of the ketone to a secondary alcohol. However, to reduce ketone to alcohol, the enzyme also needs to protonate the carbonyl group. This process is achieved by the proton relay system, which consists of a H-bond network involving
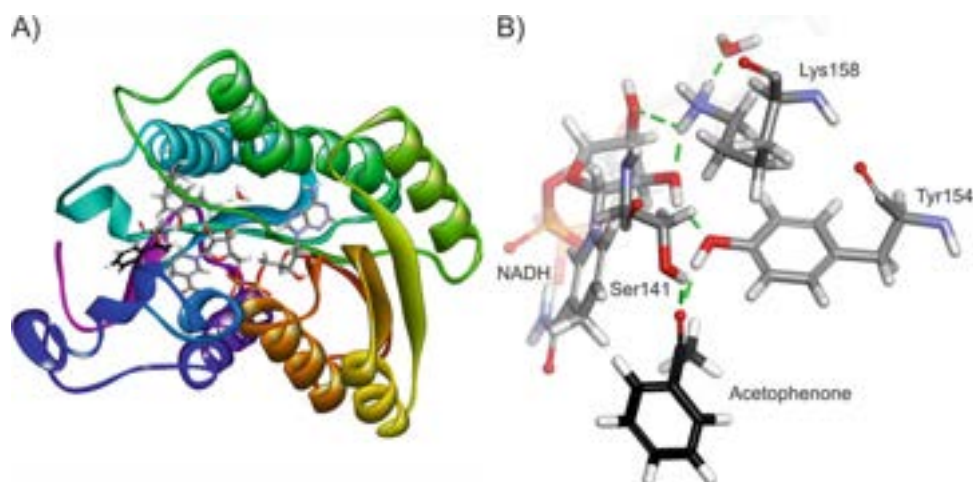
Fig. 7 (**a**) Structure of S-PEDH from *A. aromaticum* [11]; (**b**) active site of the enzyme with docked acetophenone; green dashed lines depict H-bonds (*see* 4.Cluster-QM/PEDH_acetophenon_H2O.pdb)

Tyr154, alcohol group of the NADH ribose, the amine group of Lys (Lys158), and finally water molecule (Fig. 7b). To properly study this reaction, we need to include in the cluster model all these crucial elements:

- Ketone.
- Part of NADH, which contains at least the nicotinamide ring and ribose ring (the rest is not involved in the reaction and works as an anchor for cofactor binding).
- Residues involved in substrate binding and protonation (Ser 141 and Tyr154).
- Residues that transfer a proton to Tyr (ribose, Lys158).
- If we want to model the pH-dependence of the enzyme, the inclusion of water that binds to Lys may be considered. In this case, in acidic conditions (preferential for ketone reduction), the Lys residue is protonated ($-N\varepsilon H_3^+$) and in basic conditions, it is deprotonated ($-N\varepsilon H_2$).

The model is trimmed—in the case of amino acids, bonds between Cα and Cβ are replaced by C-H and their positions are constrained (Fig. 8, yellow). In the case of NADH, the C6–O bond is replaced by C–H, and again their coordinates are frozen. As a result, all residues can rotate along the Cα–Cβ axis, and movement of the ring is not constrained. In the end, we receive a very handy model that contains 94 or 95 atoms (depending on the protonation state of Lys). The 94-atom model has a charge of 0 and the 95-atom model a charge of +1.
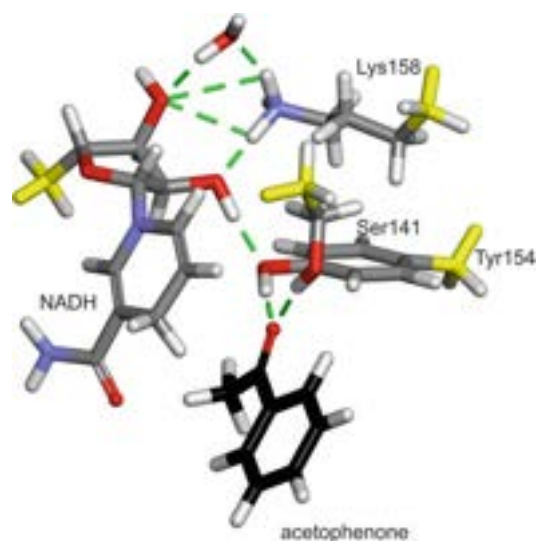
Fig. 8 The structure of the cluster model of the S-PEDH active site [13]. Yellow color indicates the position of the atom with constrained coordinates. The green dashed lines indicate the H-bond network. (4.Cluster-QM/)

4.2 Investigation of the Reaction Pathway with the Cluster Model

Before we describe the steps which lead to the modeling of the reaction mechanism, it is necessary to consider the chemistry of the hypothetical mechanism. It is always useful to draw the sequence of assumed steps. Also, one has to consider if the chosen methodology allows the investigation of the process. In this particular case, the reaction is relatively simple—it involves the transfer of a hydride from NADH to C atom of the keto group and a proton from Tyr to the oxygen atom of the keto group. As a result, we can safely use a spin-restricted approach as radicals are not produced during reactions (no homolytic C–H/O–H cleavage occurs). However, if the reaction involved homolytic C–H activation, one would have to use an unrestricted approach (with the possible breaking of orbital symmetry). Also if the enzyme contains a metal active site that can attain multiple possible spin states, one has to refer to experimental data on the subject or test different possibilities with theoretical calculations. It should be underlined here that type of selected theoretical method may influence the results. For example, it has been demonstrated that certain pure DFT functionals tend to predict the lower spin state of the metal complexes if compared to hybrid-type functionals which contain various proportions of the Hartree–Fock-based exchange part [11].

We can imagine the process of ketone reduction in two manners: hydride and proton can be transferred concomitantly or in two steps. The latter possibility of course can proceed in two ways: first, the hydride is transferred and then the proton or vice versa. It is very difficult to determine experimentally the order of events taking place in the enzyme's active site. Therefore, the theoretician

usually has to consider all chemically viable scenarios and test all of them. Finally, one can compare the energy profiles of all scenarios and select that reaction pathway, which is thermodynamically most probable.

In most of the modeling programs, we can control the position of atoms that are involved in the reaction and move them along predesigned paths. This procedure is called a potential energy scan as it delivers information on how the energy and geometry of our model changes during such transfer (we get energy in the function of one distance). It is possible to move more than one atom at a time. Moving two atoms results in a 2D scan (energy in the function of two distances). However, in this simple case, to study the concomitant transfer of hydride and proton, one needs to control only the position of the hydride (i.e., the distance of H^- to C atom of the carbonyl).

1. After the model is constrained, we conduct a complete geometry optimization of the model at the selected level of theory. As the model involves lots of H-bond, it is good to use a basis set that enables density polarization also on the light atoms (e.g., 6-31g(d,p)) (*see* **Note 2**).

   Example of the root section in Gaussian:

```
# opt=modredundant freq b3lyp/6-31g(d,p) nosymm geom=connec-
tivity pop=regular
```

   The modredundant keyword introduces additional dihedrals associated with H-bonds (four heavy atoms) and between substrate and NADH. This makes the optimization procedure more stable and less prone to errors. The example of the input files is available in the supplement (4.Cluster-QM/**S-PED-H_ES_p-H_opt.gjf**)

2. We perform a vibration analysis of the obtained geometry (freq keyword). This provides two types of information: the number of imaginary frequencies and corrections to electronic energy derived from vibration calculations (zero-point energy, thermal energy, enthalpy, free energy corrections). If the geometry is properly optimized, we should not observe any imaginary frequencies that are not associated with the imposed constraints (i.e., fix points of our model). Additionally, in many programs, it is possible to remove those atoms from vibration analysis so, in the end, we do not see their contributions to thermodynamic corrections. In case we still observe a nonzero number of imaginary frequencies, the optimization should be repeated. If the problem persists, it is good to visualize this vibration mode and potentially change the conformation of the associated part of the model. Then the geometry optimization has to be repeated (*see* **Note 3**).

3. We select reaction coordinate—in this case, it is a distance between the H-atom of the nicotinamide ring and carbonyl C atom.

4. The distance (in this case 2.74 Å) is divided into steps. Some programs only provide even steps scan (e.g., Gaussian, Turbomole), other enable custom steps (e.g., Jaguar). Our goal is to catch the transition state so we need to have a very good sampling of the region where we expect TS to occur. In case we have no prior knowledge about the localization of the transition state, it is best to use even sampling. For H-atom transfer, the step should be no bigger than 0.1 Å. In our case, we scan a range of 2.7 down to 1.1 Å in 16 steps (the endpoint of the scan is standard C-H distance of alcohols, which is ~1.1 Å).

   In order to conduct such a scan with the Gaussian package, the following section is added to the modredundant section at the end of the file (S-PEDH_p-H_scan_TS.gjf):

```
B 77 67 S 16 -0.100000
```

   (B—bond, 77 67—numbers of atoms, S—scan, 16—number of steps, -0.1 step size in Å)

5. In each step (i.e., for a given distance of H---$C^{carbonyl}$), the program conducts geometry optimization and calculates the lowest energy which complies with imposed restrains (i.e., given H---$C^{carbonyl}$ distance).

6. The energy of the optimized geometries along the scanned pathways is analyzed. The geometry with the highest energy on the profile is selected as the closest to the transition state.

7. We conduct a vibrational analysis for the selected geometry (*see* 4.Cluster-QM/**S-PEDH_p-H_preTS_freq.gjf**). The result output file should contain an imaginary vibrational mode associated with our reaction coordinate (in this case transfer of the H$^-$ from NADH to carbonyl). Visualization enables the detection of any coupled motions. In this particular example, the hydride transfer turns out to be associated with a shift of the proton from Tyr to the oxygen atom of the ketone, proton from OH group of ribose to Tyr as well as slight proton shift from the amine group of Lys to OH group of ribose (*see* animated gif: 4.Cluster-QM/**pre_TS_movie.gif**).

8. We conduct optimization of the transition state staring from the selected geometry (*see* 4.Cluster-QM/**S-PEDH_p-H_TS_opt.gjf**). We can read the Hessian needed for the optimization from the already conducted vibration job (guess = read of the checkpoint file from 4.Cluster-QM/**S-PEDH_p-H_preTS_freq.gjf**). It is also possible to calculate it once again (*see* **Note 4**).

9. If the TS optimization finishes with success (we inspect the geometry and the H^- is indeed localized somewhere midway between C atoms, it did not slide to the reactant or product atom), we conduct vibration analysis for the final geometry. In the studied case, the C–H–$C^{carbonyl}$ distances are 1.37 and 1.3 Å, and the frequency of the imaginary vibration is i787 cm^{-1} with high intensity (4200). As for preTS analysis, the hydride transfer is coupled to the motions of the protons in the proton relay system (*see* TS_movie.gif). The TS should not contain any other imaginary frequencies. If they were present and were not caused by our constrains, we had localized higher-order saddle point instead of the transition state.

10. In case the TS optimization was not successful, we should repeat the scan in the vicinity of the maximum of our profile. Perhaps a more detailed sampling is needed (e.g., 0.05 Å). Most probably such investigation will provide slightly different geometry with the highest energy. In such a case we should go back to point 7 and repeat TS optimization (*see* **Note 5**).

11. It is a good practice to check the structures of stationary points connected by TS. It may happen that the geometry we have found does not lie on a path connecting those two minima. The validation can be performed with intrinsic reaction coordinate (IRC) calculations [14]. This procedure requires an initial force constant matrix, which can be read from the checkpoint file from the vibrational analysis of TS (Gaussian keywords: IRC(RCFC)). During the calculations, the atoms are moved along the steepest descent path from the TS (in both directions). After IRC, we can optimize the obtained structures of product and substrate, and conduct vibrational analysis. They should be identical as the stationary points from which we initiated the scan (in our case the structure resulting from 4. Cluster-QM/**S-PEDH_ES_p-H_opt.gjf** calculations) (*see* **Note 6**).

12. Having geometries of all steps (in this case ES, TS, EP), we collect final geometries, energies, and vibration corrections.

13. It has been demonstrated that a good approximation of the geometry can be obtained with a lower basis set but in order to obtain a good description of the energy, it is important to use a bigger basis set than DZVP (in this case 6-311+g(2d,2p)). It is possible to conduct geometry optimization with a very big basis set, but it is not very practical and, in most cases, does not result in significantly better results. Despite the increasing power of the supercomputer centers, it is still practical to introduce those corrections as single-point calculations of energy (*see* **Note 7**).

14. If the geometry optimization was not conducted in a model solvent at this stage, the model solvent corrections can be introduced. In most modeling packages, several different model solvents can be selected (e.g., COSMO, PCM, CPCM, IEFPCM, SDM [15]). An example file (4.Cluster-QM/**S-PEDH_p-H_P_pcm.gjf**) uses the polarizable continuum model (PCM) with $\varepsilon = 4$ (*see* **Note 8**).

15. We construct the energy diagram which describes energies of the enzyme–substrate complex (ES), transition state (TS), and enzyme–product complex (EP) stationary states (Table 1). Of course, for a more complex pathway, we may localize serval consecutive TSs which are separated by respective intermediates (ES, TS1, I1, TS2, I2, ..., EP). For each stationary point, we use electronic energy calculated at the highest available level (e.g., gas-phase with higher basis set) and add corrections for solvatation (provided that the optimization was done without model solvent) and selected correction from the vibration analysis, such as zero-point energy, thermal energy, enthalpy, or Gibbs free energy. In the case of *S*-PEDH modeling, the geometry optimization, vibration correction, and solvent corrections ($\varepsilon = 4$, Polarizable Continuum Model using the integral equation formalism variant IEFPCM) were calculated on the B3LYP [16]/6-31g(d,p) level of theory with a spin-restricted approach. Then the electronic energy was recalculated at the higher level B3LYP/6-311+g(2d,2p), and additionally, a D2 Grimme correction for dispersion forces [17] was introduced (*see* **Note 9**). The energies are referred to as energy of ES stationary point (by calculation of total energy difference between any of the stationary points and ES) and presented in chemical units (kJ/mol or kcal/mol). It is also a good practice to calculate the energy diagram of the pathway at various levels of theory (e.g., for energies calculated at the geometry optimization level and for each introduced correction) (*see* **Note 10**).

Based on the obtained energetics, we should first analyze if the obtained results are in the sensible range. Here we obtain reaction barriers in the range of 87–113 kJ/mol, which can be lowered to ~67–90 kJ/mol if dispersion corrections are included. So regardless of the selected type of "energy descriptors" (be it just electronic energy, ZPE, and solvent-corrected energy or ΔG), we obtain kinetically viable results (the experimental activation energies for these reactions are in the range of 60–80 kJ/mol, depending on the substrate). It is always necessary to validate if the obtained results are chemically sensible (barriers in the range of 250 kJ/mol are under ambient conditions almost impossible to cross). On the other hand, we cannot expect such a small and static model to reproduce the chemical accuracy of the dynamic enzyme.

Table 1
Example of energy profile calculated for acetophenone; hb –base set 6-311+g(2d,2p); solvent—PCM model $\varepsilon = 4$

Substrate		ΔE (kJ/mol)	$\Delta E_{solvent}$ (kJ/mol)	ΔE_{hb} (kJ/mol)	$\Delta E_{hb + solvent}$ (kJ/mol)	$\Delta E_{ZPE + solvent}$ (kJ/mol)	$\Delta H_{+ solvent}$ (kJ/mol)	$\Delta G_{+ solvent}$ (kJ/mol)	ΔvdW corrections (kJ/mol)
Acetophenone	ES	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0
	TS	104.32	101.04	113.99	110.71	92.31	87.59	103.89	−22.65
	EP	12.47	12.08	20.19	19.79	15.23	13.58	16.01	−12.32

4.3 Analysis of Different Reaction Pathways

When modeling the catalytic reaction pathway, we should **always** consider other hypotheses. Just finding one which leads to product formation and yields kinetically accessible barriers does not prove its correctness. As mentioned above, the modeling only provides the answer for the most probable proposed scenarios. In the case of our example, we can test two plausible alternate scenarios:

- Protonation before hydride transfer.
- Hydride transfer before protonation.

In these cases, the reaction should proceed in two steps (two transition states) with one intermediate. In the first scenario, we should follow the following procedure:

1. Scan the shift of the proton from Tyr-OH to oxo group of the ketone, conduct vibrational analysis for the structure with the highest energy, run the TS optimization.
2. Optimize the structure of the intermediate product (protonated ketone).
3. Scan the shift of the hydride from NADH to carbonyl C atom, localize TS for the maximum.
4. Optimize the geometry of enzyme–product.

The second scenario can also be evaluated assuming that **steps 1** and **3** are inverted. Of course, to get the same level of energy, all the corrections applied for the initial mechanism should be calculated and added. Finally, the pathway with the lowest barriers is selected as the most probable.

Sometimes it is possible to exclude one of such alternate mechanisms by optimization of the intermediate. If the intermediate is not stable, the geometry will optimize to the reactant state. This is what was observed in the case of the ketone protonation without the transfer of hydride atom—the proton shifted back to the tyrosyl ion. The geometry was obtained from TS of the concomitant scenario by transferring back the hydride to NADH (at this stage the proton from Tyr was already transferred to ketone—*see* below).

It should be underlined that one can get different advancements of the proton/hydride shifts even during concomitant scenario which was scanned only along the NADH-CO reaction coordinate. For example, for acetophenone and p-hydroxyacetophenone in TS, the proton shift was very advanced with C=O–H distances of 1.09 and 1.07 Å. Meanwhile, for a substrate such as p-nitroacetophenone or p-acetylopyridine, these distances were 1.35 and 1.4 Å while the H–O–Tyr distances were 1.09 and 1.07 Å (i.e., proton still attached to Tyr). This indicates that for the two former cases keto groups were already protonated before TS of hydride, while for the two latter cases the protons were

101

still part of the tyrosine. However, both H atoms were shifted in one step without a discernible intermediate.

5 Correlation of Modeling with Kinetic Experiment

If our mechanistic hypothesis is correct, we can expect that the results of the calculation will correlate with the experiment. It should be underlined here that the calculations can be based on any method described in this chapter, that is, QM-only, QM:MM, or QMMM MD calculation and the selection of the particular method should be guided by the desired level of accuracy as well as computational costs. However, before we go into theory vs. experiment calculations, there are several points which should be taken into account before such a correlation is tried:

- We need several enzyme substrates with different structures and exhibiting different electronic properties in terms of the studied reaction.

- We should have a strong basis to assume that despite the structural diversity of the substrates the catalysis proceeds along the same reaction mechanism.

- It is possible to measure k_{cat}, i.e., the observed kinetics is measured at enzyme saturation by a studies substrate (no influence of different K_m of substrates) and is not modulated by other reagents (like co-substrates, reagents, etc.) or effects (substrate inhibition, substrate-dependent allosteric effect).

- The reaction under study is sensitive to the type of substrate (otherwise we can expect that our kinetic assay is limited by other factors not associated with the modeled process).

- We can collect the experimental values of k_{cat} with an error lower than the differences in k_{cat} for different substrates.

If those conditions are met, one can proceed with the calculation, which requires the modeling of the reaction pathway for each of the substrate. As a result, we get several "kinetics descriptors" (*see* 5.Correlation/correlation_descriptors.xlsx) that can be correlated with ln k_{cat} obtained from the experiment.

For each substrate, we have to identify the rate-limiting barrier, which contributes the most to the observed kinetics. In the presented example, we had only one TS, and this simplifies the problem. We repeated the described procedure for all substrates. Sometimes, for similar substrates we do not need scans—we can directly modify the geometry at TS and leave the program to perform the optimization. Finally, we do the correlation analysis (calculate the linear correlation) between each of the selected descriptors associated with the limiting barrier and ln k_{cat}. If our

mechanistic hypothesis and calculation scheme are correct (as well as the experimental assay which delivered k_{cat}) we should obtain a linear correlation between some of the energy descriptors (Fig. 9, *see* also [18] for another example of successful correlation with experiment). It is advisable to include each of the experimental results as an individual case to account for the experimental error of k_{cat} determination. Unfortunately, we cannot account for the error of modeling because we have no clear way to estimate the level of approximation introduced by the use of the cluster model of the enzyme.

We can also correlate the calculated reaction energies with the experiment. In this case, we use the energy difference between ES and EP stationary state. For the experiment, we can either use the enthalpy of the reaction (measured using calorimetric method) or the equilibrium constants established in the experiment (in this case the enzyme catalyzes reversible process so it is possible to determine reaction equilibrium). A similar correlation procedure is used to obtain correlations. The example of such a correlation is presented in Fig. 10.
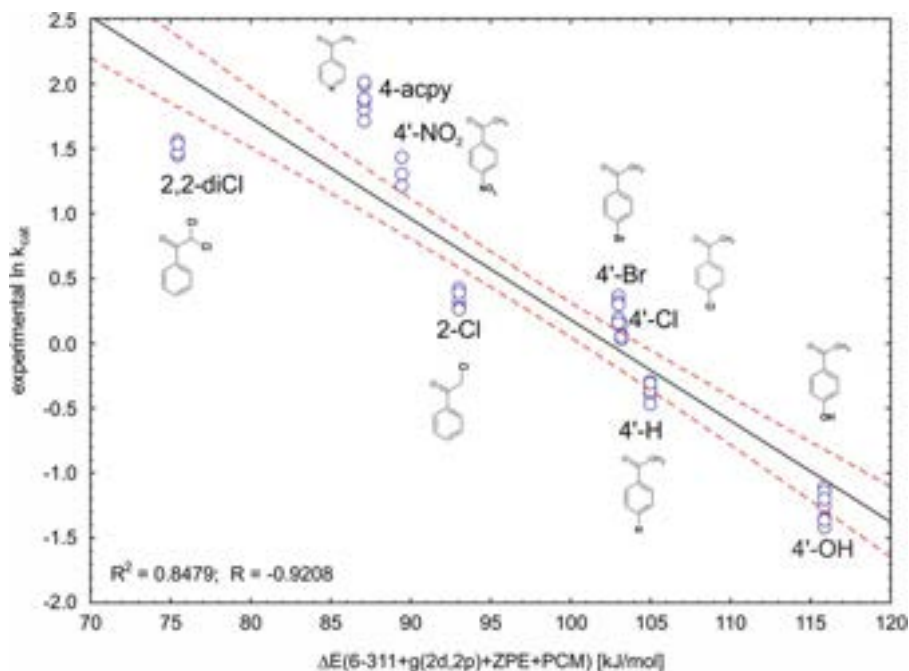
Fig. 9 Correlation analysis—scatter plot linking energy barrier of the TS with experimental ln kcat; the labels refer to substituents of acetophenone (p-H acetophenone). The correlation explains 84.8% of the observed variability [13]
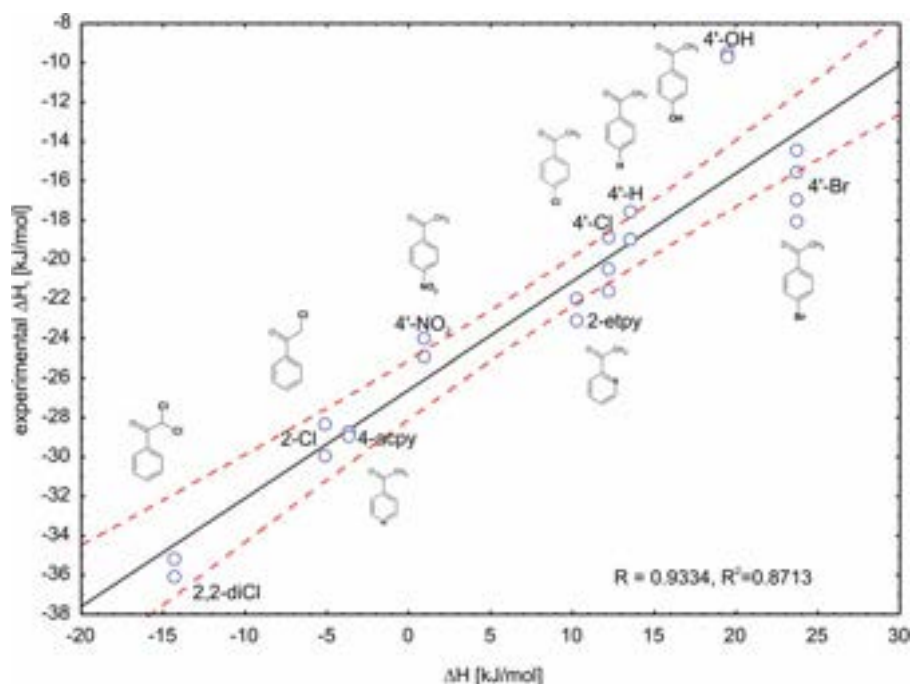
Fig. 10 Correlation plot between theoretical and experimental ΔH_r of reaction [13]

6   QM:MM Modeling

### 6.1   Preparation of the ONIOM Model for QM:MM Calculations

This section describes how to prepare an ONIOM (our own *n*-layered integrated molecular orbital and molecular mechanics) model [19] for an enzyme–substrate complex and how to run and analyze Gaussian ONIOM jobs. As an example, we use the T6ODM system (Fig. 11) already covered in the section dedicated to the MD simulations. We start from the pdb file with an MD representative structure saved, by using cpptraj program, with atomic partial charges.

The procedure to prepare the ONIOM input file can be summarized as follows:

- Selection of the ONIOM model (using VMD).
- Selection of residues that are allowed to relax (using VMD).
- Preparation of an *.atq file containing information on atom types, charges, coordinates, and allocation to a subsystem (using VMD, text editor, and python2 scripts).
- Preparation of header, connectivity list, and parameter sections for Gaussian input and combining them into an input file (text editor, python2 scripts).

Prerequisites:
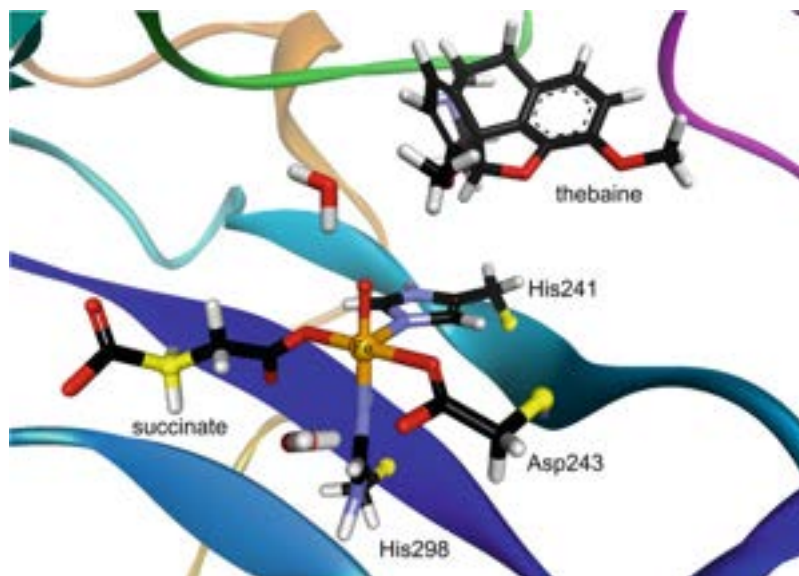
Fig. 11 QM subsystem (shown as sticks) and link atom hosts (shown as yellow balls) for T6ODM ONIOM setup

- VMD, some text editor, python2 interpreter, Linux console, Gaussian.
- Pdb file with atomic charges saved in the occupancy field of the file.
- Tcl, python2, and awk scripts accompanying this section.

6.1.1 Selection of the ONIOM Model and Optimization Zone

The pdb file saved with the atomic partial charges by cpptraj (**t6odm-thb-thb-sin.pdb**) contains coordinates for the whole MD simulation box (protein, substrate, water molecules, counterions—SI *see* files 5.QMMM/1_step_select_opt_fixed_zones/). As the ONIOM calculations are static (geometry optimization, $T = 0$ K) residues that are further than some cutoff distance from the active site usually have to be constrained in order to avoid the local minima artifacts (some random geometry change far from the active site that is not coupled to the reaction coordinate but may affect the calculated energy). Since such remote from the active site and static (frozen) residues have a negligible effect on the computed reaction energy profile, there is no need to include the whole simulation box in the ONIOM model. Hence, in this step, we select as the ONIOM model the whole protein and substrate plus those solvent molecules that have at least one atom within a range of 20 Å from the active site iron. Then, the optimization zone is set, and it includes all residues with at least one atom within 15 Å from the iron.

In what follows, commands executed in the Linux shell begin with the ">" sign, commands executed in the Tcl shell of VMD

begin with the "%" sign, a sequence of choices made from the menu of a program are marked by the "➡" sign, file names are in bold, script files in bold italic.

1. Selection of the ONIOM model.
   (a) Load the pdb file to VMD:

> vmd **t6odm-thb-thb-sin.pdb**

   (b) Open the Tk Console and execute the script that selects residues within a given distance range from a selected atom or group of atoms:
       VMD: Extensions ➡ Tk Console

% source ***select_within_20A.tcl***

       The content of the ***select_within_20A.tcl*** is the following:

```
set sel_0 [atomselect top "all"]
$sel_0 set occupancy -1
# here we provide the radius (20 A) and its origin (atom of
type FE):
set sel_1 [atomselect top "within 20 of type FE"]
set res_id [$sel_1 get residue]
set sel_2 [atomselect top "residue $res_id"]
$sel_2 num
$sel_2 set occupancy 0
# here we provide the name of the pdb fill we want to save:
$sel_2 writepdb resids_within_20A_of_FE.pdb
```

       # residues with at least one atom within a specified radius.
       # from the chosen atom (or residue).
       # will have occupancy set to 0.
       # all other residues will have occupancy set to −1.
       As a result, we get a file: **resids_within_20A_of_FE.pdb**.
   (c) Load the **resids_within_20A_of_FE.pdb** file to VMD:

> vmd **resids_within_20A_of_FE.pdb**

   (d) Visualize the zone within 20 Å of the selected atom:
       VMD: Graphics ➡ Representations ➡ Create Rep ➡ Selected Atoms: Occupancy 0.0000
       Coloring Method: ColorID 0
   (e) Change the selected atoms to include the whole protein plus other molecules with atoms within the above selected (20 Å) range:
       protein or occupancy 0.0000
   (f) Save the last selection to a pdb file:

VMD: File ➙ Save Coordinates ➙ Selected atoms: protein or occupancy 0.0000

File type: pdb [Save ↵]

Filename: …/**t6odm_within_20A.pdb** [OK ↵]

As a result, we get **t6odm_within_20A.pdb** file with all residues and molecules we want to have in the ONIOM model.

2. Selection of residues that are allowed to relax their geometry

    (a) Load the **t6odm_within_20A.pdb** file to vmd and run the *select_within_15A.tcl* (5.QMMM/1_step_select_t_opt_fixed_zones/)

script to select all residues with atoms within 15 Å from Fe:

```
> vmd t6odm_within_20A.pdb
```

VMD: Extensions ➙ Tk Console

```
% source select_within_15A.tcl
```

(the content of this script is analogous to that of *select_within_20A.tcl*, only the radius is set to 15 and the name of the saved file is changed to **resids_within_15A_of_FE.pdb**)

As a result, we get **resids_within_15A_of_FE.pdb** file in which all residues having an atom within 15 Å of Fe have occupancy set to 0.00, whereas all other residues have occupancy set to -1.00. These occupancy values are used later on to distinguish between atoms allowed to relax (value of 0) and atoms with frozen coordinates (value of -1).

    (b) Visualize this pdb file in VMD, check if all residues you want to have in the optimization zone are in the group with occupancy set to 0.00:

```
> vmd resids_within_15A_of_FE.pdb
```

VMD: Graphics ➙ Representations ➙ Create Rep ➙ Selected Atoms: occupancy 0.0000

Coloring method: ColorID

(*see* **Note 11**)

6.1.2 Preparation of the *.atq file (5.QMMM/ 2_step_prepare_atq_file)

The atq file is an auxiliary text file that has a simple format consisting of nine columns (*see* below). The information content of an atq file is exactly the same as of the geometry section of an ONIOM Gaussian input, but the atomic symbol, AMBER atom type, and atomic partial charge are in separate columns (1–3). Column 4 contains either "0" (relaxed atom) or "-1" (frozen atom), columns 5–7 contain x,y,z coordinates (in Å). Column 8 contains either "L" if an atom is in the MM subsystem or "H" if an atom is in the QM

subsystem. For MM atoms that are replaced by H-link atoms during QM calculations (so-called link atom hosts), an extra entry in column 9 is needed. It consists of an atomic symbol ("H"), atom type, and partial charge of the link atom, all separated by "-" and the number of an atom from the QM subsystem to which this link atom is bonded. For historical reasons (some bug in a previous version of Gaussian) the charge of the link atom is set to a tiny, but nonzero value.

Excerpt from an example atq file (column numbers are not included in the atq file):

```
1 2 3 4 5 6 7 8 9
....
N N -0.558200 0 1.21400 6.53200 -2.16800 L
H H 0.319700 0 2.21600 6.63900 -2.23100 L
C CT 0.007200 0 0.61200 5.73400 -1.12800 L H-HC-0.000001 3807
H H1 0.082400 0 -0.31200 5.29700 -1.50600 L
C CT 0.718000 0 1.46500 4.53300 -0.75200 H
H HC -0.140100 0 2.51300 4.77400 -0.93000 H
H HC -0.140100 0 1.39100 4.37100 0.32400 H
C C 0.281900 0 0.98800 3.25600 -1.45300 H
O O2 -0.451800 0 0.32200 3.27500 -2.49300 H
O O2 -0.664600 0 1.44800 2.18700 -0.98900 H
C C 0.443200 0 0.35400 6.57500 0.17200 L
O O -0.501400 0 1.10100 7.48600 0.55200 L
...
```

In this step, we prepare an atq file using the atomic coordinates and fixed/relaxed zone information stored in the pdb file saved at the end of the previous step.

1. Selection of QM subsystem and link atom hosts (atoms that are replaced by H link atoms).
    (a) Load **resids_within_15A_of_FE.pdb** into VMD and make two separate representations for: fragments / atoms that make QM subsystem and link atom hosts
        VMD: Graphics ➡ Representations ➡ Create Rep (Drawing method: licorice):
        (resid 372 439 9791) or (resname FEO OXO) or (resname HIV HIH ASF and sidechain) or (index 5743 to 5748)
        VMD: Graphics ➡ Representations ➡ Create Rep (Drawing method: CPK):
        (index 3776 3804 4664 5749)
        The QM subsystem and link atom hosts are shown in Fig. 11.
    (b) Check if the two representations contain all the atoms that you would like to include in the QM subsystem. Once done, save the visualization state for your later reference

(can be loaded into VMD to easily reproduce the visualization state):

VMD: File ➡ Save Visualization State ➡ Filename: .../**t6odm_hli.vmd** [OK ↵]

(c) Write down atom numbers for atoms from the QM layer that are bound to H-link atoms (3779, 3807, 4667, 5747)

VMD: hit 0 (zero), left-click on a given atom, read out the "index" value from the command line window, as VMD counts index from 0, add 1 to the value)

(d) With text editor, open the script file *oniom_levels.tcl* and from the VMD/Graphical Representations/Selected Atoms window copy the atom masks into the "..." fragments of sel_h and sel_i lines, and in the last line modify the name of the pdb output file to be saved by the script. Save the file.

After editing the content of the *oniom_levels.tcl* file is:

```
set sel_0 [atomselect top "all"]
$sel_0 set beta 0
# here provide info about the QM subsystem:
set sel_h [atomselect top "(resid 372 439 9791) or (resname FEO
OXO) or (resname HIV HIH ASF and sidechain) or (index 5743 to
5748)"]
$sel_h num
$sel_h set beta 2
# here provide info about link atom hosts:
set sel_i [atomselect top "(index 3776 3804 4664 5749)"]
$sel_i num
$sel_i set beta 1
# here give the name of the pdb file to be saved:
$sel_0 writepdb t6odm_oniom_hli.pdb

# saves a pdb file with info which atoms belong to which
subsystem
# beta values are set to: 2 - atoms from the QM subsystem
# 1 - link atom hosts
# 0 - other atoms from the MM subsystem
```

(e) Open the VDM TK Console and execute the *oniom_levels.tcl* script:

```
% source oniom_levels.tcl
```

This generates the **t6odm_oniom_hli.pdb** file, that in the beta factor column stores the information on the subsystem assignment.

(f) Make a copy of the **t6odm_oniom_hli.pdb** file and edit it using a text editor:

```
> cp t6odm_oniom_hli.pdb t6odm_oniom_hli_noX.pdb
```

In the **t6odm_oniom_hli_noX.pdb** file:

- Replace the fifth column ("X") with a space " ".
- Remove the first line of the file, which is not starting with "ATOM".
- Remove the last line, containing "END".
- Revert the nonstandard amino acid names to standard ones: HIH ➡ HIS, HIV ➡ HIS, ASF ➡ ASP.
- From the original pdb file (**t6odm-thb-thb-sin.pdb**), copy the column with atomic partial charges and place them in the last column of the **t6odm_oniom_hli_noX.pdb** file.

(g) With the use of an awk script ***vmd_hli_pdb_plus_charge_to_atq.awk*** generate the atq file using the **t6odm_oniom_hli_noX.pdb** file as input:

```
> awk -f vmd_hli_pdb_plus_charge_to_atq.awk t6odm_oniom_hli_noX.pdb > t6odm_oniom.1st.atq
```

(*see* **Note 12**)

(h) Make a copy of this file and edit the copy:

```
> cp t6odm_oniom.1st.atq t6odm_oniom.2nd.atq
```

Edit the **t6odm_oniom.2nd.atq** file, particularly pay attention to lines beginning with "ATN!", as these contain atoms for which the awk script did not have atom types defined:

- Insert appropriate atom types and format the lines that were marked with "ATN!".
- For atoms coordinated to the metal, which have other than the standard type, change their type consistently with preceding MD simulations (His241:NE2 (atom # 3787) type: NY, His298:NE2 (atom # 4675) type: NX).
- Change the type of the nitrogen atom in the N-terminal amino group to N3 (protonated N-terminal), if this is the case for your system.
- Change the type of oxygen from the C-terminal group from "O" to "O2".
- Change the entries of the last column:
  0 ➡ L
  2 ➡ H

1 ➤ L H-HC-0.000001 XXXX (where XXXX is the number of the QM subsystem atom that is bound to the H-link atom replacing this L-atom).

(i) Check the charges of your system:

```
> ./atq_sum_charges.py t6odm_oniom.2nd.atq
```

The output of this is:
From the atq file I have read 6857 lines
found 87 atoms H, and 4 LAH atoms and 6766 remaining MM atoms
Charge of the H-layer: 1.1938
Charge of LAHs: 0.5297
Charge of other MM atoms: -17.7476
TOTAL charge of the system: -16.0241

(j) Check how the QM-part + link atom hosts look like:

```
> ./xyzn_z_atq.py t6odm_oniom.2nd.atq t6odm_oniom.qm_lah.xyz
```

Visualize the **t6odm_oniom.qm_lah.xyz** file with e.g. VMD.

6.1.3 Preparation of the ONIOM Input File (5.QMMM/3_step_prepare_oniom_input)

In this step, we first use the atq file generated previously and prepare the geometry block of the Gaussian ONIOM input. Then, we prepare a block with a connectivity list and a header section of the input and finally combine all of them into the whole input file.

1. Generate the geometry block of the ONIOM input:

```
> awk -f atq_to_input.awk t6odm_oniom.2nd.atq > t6odm_o-
niom.2nd.atq.for_inp
```

2. Generate the connectivity list:

```
> ./connec_list_for_atq.py t6odm_oniom.2nd.atq t6odm_o-
niom.2nd.atq.connect_1st
> cp t6odm_oniom.2nd.atq.connect_1st t6odm_oniom.2nd.atq.
connect
```

Using a text editor, remove (explicit) bonds between Na+ (atom # 5846) and all other atoms from the **t6odm_oniom.2nd.atq.connect** file. These explicit bonds are removed since in AMBER force field interactions with the Na$^+$ ion are described by nonbonded terms only.

In the connectivity list, each line provides bonding information for a given atom (its number is provided at the beginning of the line); however, information on bonds made by this atom with atoms of lower indexes is provided in lines for these previous atoms and it is not repeated. Thus, to remove all bonds between atom # 5846, search all occurrences of

"5846" in the file and remove all of them except for the line beginning with "5846," which should be truncated to just the "5846" number. For example:

"4491 4492 1.0 4493 1.0 5846 1.0" should be changed to: "4491 4492 1.0 4493 1.0"

"4492 5846 1.0" ➡ "4492"

"4493 5846 1.0" ➡ "4493"

"5846 6012 1.0 6060 1.0 6492 1.0 6501 1.0" ➡ "5846"

3. Prepare the parameters section of the Gaussian input:

The file **t6odm_oniom.params** has to be prepared manually using information that is available in Amber parameter and frcmod files that were used to prepare the *prmtop file used for MD simulations. For example, van der Waals parameters for the Na^+ ion ("VDW Na+ 1.369 0.0874393") are taken from the **frcmod.ionsjc_tip3p** file that is included in the AmberTools files. For the format of this section, *see* the Gaussian web page.

Excerpt from the **t6odm_oniom.params** file with two example lines per VDW, bond, and angle sections:

```
VDW Na+ 1.369 0.0874393
VDW NX 1.8240 0.1700
...
HrmStr1 FE ox 365.0 1.642
HrmStr1 NY FE 61.0 2.110
...
HrmBnd1 CC CV NX 70.0000 120.0000
HrmBnd1 CC CV NY 70.0 120.0
...
```

4. Prepare a file with the head section of the input file: t6odm_oniom.head:

```
%chk=
%mem=24GB
# ONIOM(UB3LYP/def2SVP EmpiricalDispersion=GD3BJ:Amber=Soft-
First) Geom=Connectivity
5d scf=(xqc,maxcycle=350) nosymm


Comment here


-16 5 1 5 1 5
```

The total charge of the system is −16, the total charge of the QM subsystem is +1, multiplicity in both cases is 5—this information is provided in the last line of this file.

5. Join the files (concatenate them) inserting blank lines between the input sections:

```
 > cat t6odm_oniom.head t6odm_oniom.2nd.atq.for_inp <(echo)
t6odm_oniom.2nd.atq.connect <(echo) t6odm_oniom.params
<(echo) > t6odm_oniom.com
```

Edit the **t6odm_oniom.com** file—e.g., insert the appropriate name for the checkpoint file, modify the keyword list, insert a comment.

6. Run Gaussian for the t6odm_oniom.com input.

6.2   How to Analyze the Log File and Recompute Atomic Charges for the QM Subsystem

1. Extracting geometry of the whole system out of the Gaussian output   (5.QMMM/ 4_step_analyze_result_recomput_charges).

   In the *oniom_log_to_xyz2.awk* script file, set the *Ntot* variable to the number of atoms in your system (here 6857); then the script can be used to extract geometry out of the Gaussian output:

```
>awk -f oniom_log_to_xyz2.awk t6odm_oniom.log > t6odm_oniom.
xyz
```

   The last geometry can then be easily extracted with the tail command by providing as a parameter the number of lines by 2 greater than the number of atoms in the system:

```
>tail -6859 t6odm_oniom.xyz > t6odm_oniom.last.xyz
```

2. Visualizing the geometry changes between the starting and optimized structures.
   (a) Load the **t6odm_oniom_hli.pdb** and **t6odm_oniom.last.xyz** into VMD as molecules with ID 0 and 1, respectively, then execute the *rmsd_b.tcl* script:
       VMD: Extensions ➡ Tk Console:

```
% source rmsd_b.tcl
```

   (b) Color-code the atoms accordingly with their displacements:
       VMD: Graphics ➡ Representations:
       Selected Molecule: 1: t6odm_oniom.last.xyz
       Coloring Method: Beta
       Atoms are color-coded according to their displacement, which helps to identify regions of the system that changed their coordinates the most. The colors of the scale, its *Midpoint* and *Offset* can be changed at:
       VMD: Graphics ➡ Colors ➡ Color Scale

3. Recomputing the atomic charges of the QM part of the system and preparing a new atq file.
   (a) Prepare an atq file with coordinates of the ONIOM optimized system:

```
>all_geo_from_xyz_to_atq.py t6odm_oniom.last.xyz t6odm_
oniom.2nd.atq t6odm_oniom.3rd.atq
```

   (b) Using VMD and a text editor, prepare the **t6odm_oniom.omit_q** file with atom indexes (one per line,

index counted from 0, as in VMD) for atoms whose charges should not be included, in the set of atomic charges polarizing the QM system, when calculating new RESP charges of the QM subsystem. In our case, and consistently with the default electronic embedding scheme in Gaussian, these are atoms that are located one bond away from the link atom hosts.

The content of the **t6odm_oniom.omit_q** file is:

```
3774
3777
3789
3802
3805
3812
4662
4665
4677
5750
5751
5752
```

(c) Prepare the **t6odm_oniom.3rd.g_charge** file with coordinates and charges of atoms from the MM subsystem (without LAH:s and those specified in the **t6odm_oniom.omit_q** file):

```
>atq_omit_to_charge.py t6odm_oniom.3rd.atq t6odm_oniom.omit_q t6odm_oniom.3rd.g_charge
```

The first few lines of the **t6odm_oniom.3rd.g_charge** file look like this:

```
-25.39    -3.428    -10.086    -0.664400
-26.068   -3.076    -10.746    0.471500
-24.714   -2.681    -10.157    0.471500
-25.841   -3.482    -9.184     0.471500
-24.854   -4.725    -10.531    0.140300
...
```

(d) Extract the geometry of the QM subsystem (with LAH:s replaced by H atoms).

```
> qm_plus_H_from_xyz.py t6odm_oniom.last.xyz t6odm_oniom.3rd.atq t6odm_oniom.last.qm_H.xyz
```

(e) Prepare files with header and tail for Gaussian ESP calculations in the presence of MM charges:
Content of the gaussian_esp_q.head:

```
%chk=esp_q.chk
%Nproc=24
```

```
%Mem=24GB
# UB3LYP/def2SVP 5d scf=(xqc,maxcycle=350) charge nosymm
Pop=(MK,ReadRadii) iop(6/33=2) iop(6/42=6) iop(6/50=1)
```

SP with MM charges (like in ONIOM-EE) before RESP.

```
1 5
```

gaussian_esp_q.tail:

```
Fe 1.456
gaussian_esp_q.gesp
gaussian_esp_q.gesp
```

(f) Join the files to assemble the Gaussian input file:

> cat **gaussian_esp_q.head t6odm_oniom.last.qm_H.xyz** <(echo) **t6odm_oniom.3rd.g_charge** <(echo) **gaussian_esp_q.tail** > **gaussian_esp_q.com**

From the thus generated **gaussian_esp_q.com** file, remove two lines after charge and multiplicity, which contain number of atoms and comment from the xyz-file.

(g) Run Gaussian for the **gaussian_esp_q.com** input.

(h) Calculate the sum of charges for MM atoms of residues that are split between QM and MM subsystems (from the **t6odm-thb-thb-sin.pdb** file):

```
HIH and HIV: N, H, CA, HA, C, O: 0.0172
ASF: N, H, CA, HA, C, O: -0.2071
SIN: C3, H3, H4, C4, O3, O4: -1.0562
```

(*see* **Note 13**)

(i) Fit RESP charges with charges on H-link atoms fixed to values determined from the above sums: HIH and HIV (atoms # 1 and 20 in the **t6odm_oniom.last.qm_H.xyz** file): 0.0172; ASF (atom # 13): −0.2071; SIN (atom # 40): −0.0562. RESP charge fitting procedure is described in detail in the section on deriving parameters for MD, yet here the symmetry constraints between the atoms are removed completely.

(j) Generate a new atq file (**t6odm_oniom.4th.atq**) with the atomic charges of the QM subsystem as determined in the RESP procedure; **resp_2.punch** is the output of the second passage of the RESP program.

> **add_resp_for_hlayer_to_atq_file.py** resp_2.punch t6odm_oniom.3rd.atq t6odm_oniom.4th.atq

115

The resulting **t6odm_oniom.4th.atq** file contains optimized geometry of the system and atomic charges of the QM subsystem as determined with the RESP procedure in the presence of the MM point charges. It should now be used to prepare a new Gaussian input to run geometry re-optimization with the new set of atomic charges on the QM subsystem. In our experience, such a two-step optimization procedure, with an intervening RESP charge fitting, yields geometries and energies close to those that would be obtained with an electronic embedding scheme (which is usually far more costly and difficult to converge).

7   QMMM MD Simulation

QM cluster and QMMM modeling approaches do not take dynamic effects into account. The results obtained using these methods strongly depend on the initial geometry of the selected model (local minima problem). Using QMMM MD simulation, it is possible to obtain a Free Energy profile of reaction. One of the possibilities is to use an umbrella sampling approach to obtain the potential of mean force (PMF). We present how to generate such a profile at AMBER:AM1 level of theory using the AMBER software package.

7.1   Preparation of QMMM MD Calculations

In the example, we explain how to use the umbrella sampling method to study a reaction catalysed by 3-ketosteroid dehydrogenase from *R. erythropolis* (PDB code 4c3y). The mechanism proposed for this enzyme consists of two steps, but we discuss only the first one (analysis of the second step would be analogous to the first). In this process, a proton is transferred from 3-ketosteroid (in our example: dihydrotestosterone, DHT) to tyrosyl anion.

Procedure:

1. Make sure that the QM method selected for QMMM MD simulation can describe the analyzed process. In our example, we use the semiempirical method AM1. A few semiempirical methods are supported natively by AMBER. It is also possible to use ab inito or DFT methods with AMBER and external computational chemistry programs (there are interfaces for Gaussian, ADF, GAMESS-US, NWChem, Q-Chem, TeraChem, MRCC) [20]. Before MD simulation, we have modeled full reaction path using AMBER:AM1 with QMMM approach—results were comparable with a higher level of theory (AMBER B3LYP)—at least in terms of geometry and chemistry of the predicted mechanism. Barriers obtained with AM1 were significantly overestimated.

2. We start with a structure obtained from MD simulation (after cluster analysis and cooling).

3. To set up QMMM calculations in AMBER, we have to modify the input file setting ifqnt flag in &ctrl section to 1. Only then AMBER is be looking for &qmmm section. Then we have to specify the &qmmm namelist:

```
&qmmm
qmmask="@4542-4555,7294,7295-7344,7384,7385,7386,7402-7428",
qmcharge=-1,
qm_theory='AM1'
&end
```

4. Selection of the QM region—recommendations are the same as for the simple QMMM case. In this example, we selected three residues to QM layer: TYR318, a fragment of FAD cofactor, and dihydrotestosterone. To qmmask parameter, we can assign any valid atom mask, but we have to bear in mind that it cannot be longer than 256 characters (like any other atom mask in AMBER input).

5. Selection of flexible region—it is not necessary to make all atoms of the model flexible. In our example, only residues within 20 Å from dihydrotestosterone are allowed to move during QMMM MD simulation. To select flexible atoms in the AMBER input file, we can define atom mask or group input. For this purpose, using an atom mask is impractical because of its length limit to 256 characters. It is possible to use a distance-based atom mask, but in such a situation, number of frozen residues is dependent on initial coordinates. That is why we use tcl script (SI flexibleSelection.tcl in 6. QMMM_MD directory) to generate group input for AMBER.

6. Now we are ready to start the first QMMM geometry minimization of our model (all input and output files are located in 6. QMMM_MD/minimization in SI).

7. Selection of reaction coordinate—because we want to analyze the transfer of H^+ ion as the reaction coordinate, we choose the distance differences defined as in Fig. 12. If we chose only one distance to control, we would not be sure that after breaking/ forming the bond, one of the reactants would not move away from the active site during MD simulation. We must be sure that we get geometries that match each stage of the reaction. To do this, we have to run a series of independent MD simulations with different values of ξ. To ensure that our coordinate is close to the desired value during MD simulation, we add parabolic potential (i.e., 500 kcal/$Å^2$ mol force constant) which holds H^+ close to the desired ξ. We specify this restraint in distDiff.rst file:
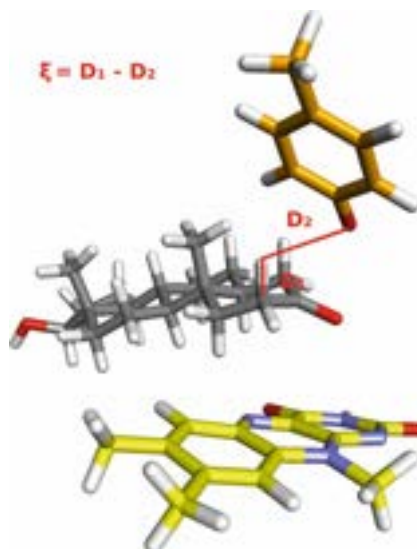
Fig. 12 QM-part of the QMMM MD job with a definition of the studied reaction coordinate ξ

```
&rst
 iat=7310,7337,7337,4551, r1=-20, r2=-1.6 , r3=-1.6 , r4=20 ,
rk2=500,
 rk3=500,rstwt=1.0,-1.0,
 &end
```

(*see* **Note 14**).

7.2 Scanning Atom Transfer with QMMM MD Simulation

1. The next step is to generate the coordinates of the initial geometries that are used in the scan. To do that, we need to run the python script umbrellaSetup.py. This script creates 70 directories (we use 70 windows in our QMMM MD simulation dividing ξ into 70 steps) and slurm script AMBERScan.slurm to run a simulation. In every directory, distDiff.rst file is generated with the appropriate value of reaction coordinate, also some AMBER scripts are copied. Script AMBERScan.slurm performs sequential minimization series. The result coordinates from the first calculation is the input geometry for the next one, etc.

2. The geometries minimized in the previous step are ow used as initial geometries for series of QMMM MD simulations. At each simulation, we add restraint on a reaction coordinate. Each simulation starts with 5 ps of relaxation MD (md_eq.in), then 20 ps of MD simulation (md.in) is performed. For practical reasons, we do not use the heating procedure as in the regular MD simulation. Instead, in the relaxation phase, we set tempi=303 and as the input file contains no velocities, AMBER

assigns them from a Maxwellian distribution. In the MD phase, distDiffDump.dat file is created—it contains values of reaction coordinate of every hundredth frame. To run all these simulations, one has to execute runPMF.slurm script, which was also generated by umbrellaSetup.py. This script works with slurm system.

7.3 Analysis of QMMM MD Simulation

1. The next step is to verify if the values of the reaction coordinate from two consecutive windows overlap. If this condition is not met, it means that we have to decrease the scan step when generating the geometry for QMMM MD simulations or we have to decrease the value of force constant, to ensure that we have continuously sampled the reaction coordinate. To verify this, we need to run the script overlapTest.py. It is recommended to visualize the obtained data on some plots. If we run the two scripts: pmfHist.py and pmfPlots.py, we will create two files: hist.png and plots.png. The first file (Fig. 13a) contains histograms from all windows—as we can see there is no gap between any pair of histograms. The second file contains a plot of the reaction coordinate from all windows (Fig. 13b).

2. To generate a PMF profile from output files from our QMMM MD simulations, we use WHAM (Weighted Histogram Analysis Method). Unfortunately, AMBER does not provide this functionality, and we have to use an external code. We used the version developed by the Grossfield lab [21] (*input—download the wham file from the webpage*). The input file for this program was generated by umbrellaSetup.py script. All we need to do now is to run the following command:

```
wham -2.2 2.2 600 0.000001 303 0 whamInput.dat pmf.dat
```

The first two parameters define the range of ξ for the histogram analysis (-2.2 to 2.2), the next one is the number of bins in the histogram. Then we defined convergence tolerance, temperature (303 K). The value of the next argument is zero because we use aperiodic coordinates. The last two arguments are the input file and the result file (*see* **Note 15**).

Please note that the force constant in WHAM input is twice as big as the one we wrote in AMBER files. This is because wham software multiply entered force constant by 0.5 during calculations.

As a result, we obtain PMF which represents an averaged free energy profile of the proton transfer (Fig. 14).

In addition to the PMF, the above simulation provides a whole range of structures that can be used for other calculations. For example, these geometries can be used for the optimization of the transition states along multiple pathways (i.e., for different
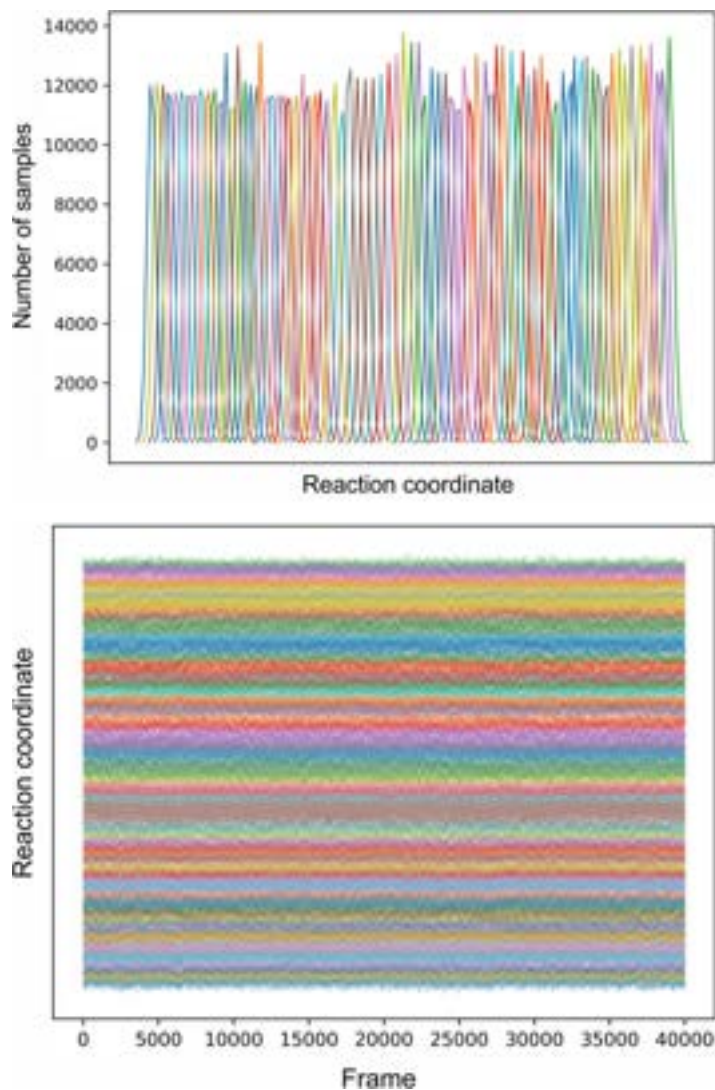
Fig. 13 (**a**) Contours of the histogram from umbrella sampling simulation. (**b**) Reaction coordinate over time for all windows used in umbrella sampling simulation

conformations of the proteins), providing information on the influence of enzyme confirmation on the catalytic process.

8   Prediction of Kinetic Isotope Effects

In this section, we demonstrate how we can estimate isotope effects from modeling. This section is general and is not associated with any specific software. The calculations presented here can be carried out using any QM software, which supports the second derivatives
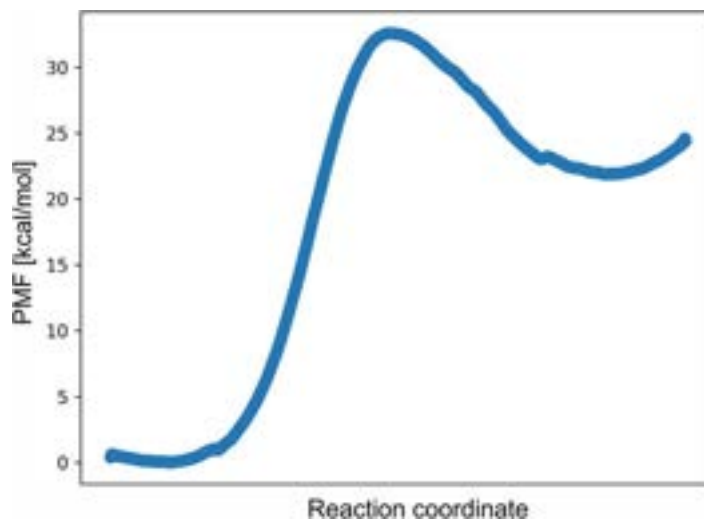
Fig. 14 An averaged free energy profile of the proton transfer

of energy. The procedure is the same for QM cluster or QMMM approach.

First of all, we have to specify what are the isotope effects, and what exactly we can estimate from the calculation. In chemistry, we usually consider kinetic isotope effects (KIE).

The formal definition of KIE is as follows:

$$\text{KIE} = \frac{k_{\text{L}}}{k_{\text{H}}} \tag{1}$$

where k_{L} is the rate constant for the reaction involving reactant substituted with light isotopes and k_{H} is the rate constant measured for a reaction where at least one atom of the substrate was substituted by heavy isotope.

The detailed theory of KIE calculations is very complex (*see* [22, 23] for a comprehensive review). In this section, we present only the basics. Several factors affect KIE value. The first and most important one is the dependency of the free energy activation barrier ($\Delta G_{\text{act}}^{\text{QC}}$) to isotopic substitution.

$$\text{KIE}\big(\Delta G_{\text{act}}^{\text{QC}}\big) = e^{\frac{\Delta G_{\text{act}}^{\text{QC,H}}(T,x) - \Delta G_{\text{act}}^{\text{QC,D}}(T,x)}{\text{RT}}} \tag{2}$$

To explain how to calculate this property, one more time we analyse reaction catalyzed by 3-ketosteroid dehydrogenase from *R. erythropolis*. This time we use the cluster model of this enzyme active site (all input and output files are located in SI in the KIE directory).

To calculate the KIE, we have to:

1. Obtain and optimized structure of reactants and transition state (as described in Subheading 4.2).

2. Calculate the second derivatives of energy for those structures (i.e., frequency calculation) at the desired temperature.

3. Collect thermal corrections to free energy calculated directly from these derivatives. Gaussian software package prints in output from freq job (the thermal correction to Gibbs free energy) and those values can be used to calculate KIE.

4. Calculate the thermal correction to the free energy for the heavy isotopologue—this can be achieved by reading the checkpoint from the freq calculations (freq = readfc) and providing data on the mass of the substituted atom (here substitution of protium to deuterium).

```
H(Iso=2) 0 x y z
```

In this example, we substitute the hydrogen atom, which is transferred from ketosteroid to tyrosyl anion, with deuterium. We can verify that Gaussian correctly set the mass of the selected atom by investigating the output file.

```
Atom 198 has atomic number 1 and mass 1.00783
Atom 199 has atomic number 1 and mass 2.01410
Atom 200 has atomic number 1 and mass 1.00783
```

We collect thermal correction to Gibbs free energy from all output files (Table 2).

Now we can substitute these values into Eq. (2) (after recalculation of units—Hartree to J/mol). Finally, we obtain an estimation of KIE as 4.65. Please note that it does not matter if we use only thermal correction to Gibbs free energy or the sum of the electronic energy and this correction—electronic energy does not depend on the mass of atoms so it does not influence the value of KIE.

There are also external programs that can read the second derivatives from Gaussian outputs and calculate KIE from frequencies (Kinisot [24], ISOEFF [25]).

Another factor that can be important for the KIE calculation is tunneling, i.e., quantum effect associated with traversing the energy barrier even if the atom has too low energy to go over the barrier. The easiest way to estimate the tunneling coefficient is to use Wigner correction [26]:

$$\chi(T) = 1 + \frac{1}{24}\left(\hbar \ \mathrm{Im}\left(\frac{\nu)}{k_b T}\right)\right)^2 \tag{3}$$

As one can see, the tunneling coefficient depends only on the temperature and imaginary frequency of transition state. To calculate tunneling contribution to KIE, we have to calculate the tunneling coefficient for nonsubstituted and isotopically substituted

Table 2
Thermal correction to Gibbs free energy calculated for light (nondeuterated) and heavy (deuterated) substrate [Hartree]

Stationary point	Nondeuterated isotopologue	Deuterated isotopologue
ES	1.66398	1.660581
TS	1.664305	1.66238

reactants. The values of the imaginary frequencies (provided in cm^{-1}) associated with the transition state can be retrieved from the freq output job (calculated for light and heavy reactant). In the case of results calculated in Gaussian, one can use also formatted checkpoint and Gaussview to obtain imaginary frequency for the heavy isotopologue.

Then we can calculate:

$$KIE(\kappa) = \frac{\kappa^L}{\kappa^H} \qquad (4)$$

Finally, we can estimate KIE as:

$$KIE = KIE(\Delta G)KIE(\kappa) \qquad (5)$$

In our example, the values of the imaginary frequencies of TS are presented in Table 3.

When we substitute these values into the Wigner equation, we obtain:

$$KIE(\kappa) = 1.36 \qquad (6)$$

Finally, if we consider tunneling, we can estimate combined KIE as:

$$KIE = KIE(\Delta G)KIE(\kappa) = 6.32 \qquad (7)$$

This estimation yields the so-called intrinsic KIE associated with a particular barrier. In case the effect is strong and we did this calculation for the rate-limiting step in our mechanism, we can expect correlation with the experiment. KIE is the best way to correlate the modeling with the experimentally observed kinetics. However, we have to bear in mind that enzymatic reactions are very often multi-step processes with binding and release of substrate/products and sometimes conformational changes. Each of these processes can be or not sensitive to isotopic substitution. It is very often that if some other rate-limiting step occurs on the reaction pathway the experimentally determined KIE is of lower value than this predicted by theory [27].

Table 3
Imaginary frequencies calculated at TS for a nondeuterated and deuterated substrate

Substrate type	Imaginary frequency (cm^{-1})
Nondeuterated isotopologue	i1256.7172
Deuterated isotopologue	i938.6045

9   Notes

1. In 5O7Y PDB file, the residue number starts from −2, but we renumber PDB file with the residue number starting from 1. Thus, for the whole work, we are using the new numbering scheme, e.g., HIS238 is renumbered to HIS241.

2. At this stage, we should consider if the calculations will be conducted in vacuum and only corrected with model solvent (single point corrections) or the geometry optimization will be also conducted in a model solvent. The latter solution may be of importance if the solvent strongly penetrates the active site and charges are being created during reactions. In some cases, the model solvent can stabilize more charge-separated states of the amino acid (zwitterion not neutral). On the other hand, model solvents can also introduce their own biases. The former solution is definitely faster. Also, at this stage, one has to consider using the dispersion correction. This is especially important if there are many vdW interactions present in the active site of the model.

3. In the case we want to compare the results to a particular temperature at this step, we should provide the temperature for which the thermal energy, enthalpy, or free energy is to be calculated; also one can consider the introduction of the scaling factor appropriate for the functional and basis set selected for calculations.

4. In most packages, it is possible to initiate TS search without prior vibration analysis; however, if the search is not successful, it might be difficult to understand why the search failed. The vibration analysis provides information if the vibration mode is correct, i.e., along the expected direction of the reaction coordinate, and if its frequency and intensity are appropriate for a given type of atom transfer; in this case, the frequency of the initial vibration was i800 cm^{-1} and intensity was very high 4978.

5. In case we are still unable to localize TS, we have to consider that our mechanistic hypothesis is wrong and that there is no TS in this region. Perhaps the order of steps is different. In this particular example, the ketone could have to be initially protonated to enable the reduction of the carbonyl atom.

6. In the case of a several-stage reaction, it is worth to pay attention if the energies of the same intermediate, obtained from different IRCs (and optimized), are comparable. If they differ, then the reason for this should be found. It may be, for example, a change in the conformation of the fragment of the model.

7. The reliability of this approach can be tested by reoptimization of the obtained stationary points in higher basis set; in most cases, no significant geometry changes result from such reoptimization; if significant changes are obtained, this proves that the basis set used was too small to account properly for the chemistry of the studied process.

8. Each of these methods requires several additional parameters to be added, and one should refer to the manual for the program/method specific guidelines. However, for each of them, one has to define the dielectric permeability (ε), which will directly influence the extent to which the model solvent will screen electrostatic interactions between charged species in the model. There are no hard guidelines for the selection of ε value. Usually, values in the range of 4 (average experimental ε of the denaturated protein]) up to 40 are used [28]. In reality, this parameter changes continuously and depends on the degree of exposition of the active site to the water phase (which has ε 78 at 25°C), so in fact, the dielectric constant is not "constant" at all. We recommend using ε of 4 for reactions that are buried deep inside the enzyme and 20 or even higher for reactions taking place in shallow clefts of enzymes that are normally filled with water and if the reagents are in contact with water molecules. Finally one can test with fast single-point calculations the dependence of the energy profile on the value of ε. If big changes of energies are observed for solvent corrections calculated with different values of ε, it is a strong indication that: (1) the geometry optimization and TS detection should be conducted with model solvent or (2) the model is too small and the model solvent penetrates two close the reactants which result with high sensitivity of the reaction to the model solvent. Here it should be underlined that the enzymes usually try to prevent solvent from accessing the reagents to (1) decrease the entropy of binding by the release of reagents solvatation shell and expulsion of solvent molecules bound in the active site and (2) prevent solvent-mediated side reactions.

9. Although it is possible to introduce G correction, one has to bear in mind that those values may introduce an error in entropy estimation associated with the way we constructed the cluster (i.e., constraints).

10. This helps for the detection of potential errors that inevitably occur in the course of modeling and provides additional information on factors stabilizing energy of particular transition states and intermediates (like solvatation correction or higher basis set).

11. The chosen values of radii (of 20 and 15 Å) for selecting the ONIOM model and its optimization zone, as well as the origin of these zones (here Fe atom bound in the active site), can be modified according to the needs of the system studied. For example, one might choose these zones based on the distance from any atom of selected residue(s), e.g., the ones that will be included in the QM-layer and set lower values of the radii.

12. The ***vmd_hli_pdb_plus_charge_to_atq.awk*** script ascribes AMBER atom types to atoms of standard protein residues and water molecules, however, for nonstandard residues, atom types need to be provided manually, which is done in the following step.

13. The values of these charges are used in the next step to constraint the charges of H atoms replacing LAH:s in the RESP charge fitting procedure. This is necessary if we want to update the charges of the QM subsystem and at the same time preserve the correct integer total charge of the whole system.

14. For more details on the meaning of the parameters, *see* the AMBER manual.

15. We do not discuss all possible arguments for WHAM program (e.g., how to use it with periodic coordinate). Please refer to the manual of wham for more details.

Acknowledgments

References

1. Rohman A, van Oosterwijk N, Thunnissen AM, Dijkstra BW (2013) Crystal structure and site-directed mutagenesis of 3-ketosteroid Delta1-dehydrogenase from Rhodococcus erythropolis SQ1 explain its catalytic mechanism. J Biol Chem 288(49):35559–35568. https://doi.org/10.1074/jbc.M113.522771

2. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. J Comput Chem 25

(9):1157–1174. https://doi.org/10.1002/jcc.20035

3. Seminario JM (1996) Calculation of intramolecular force fields from second-derivative tensors. Int J Quantum Chem 60(7):1271–1277. https://doi.org/10.1002/(SICI)1097-461X(1996)60:7<1271::AID-QUA8>3.0.CO;2-W

4. Nakazawa T, Igarashi T, Tsuru T, Kaji Y (2011) Density functional investigation of Fen clusters (n≤6) with Cr substitutions: UB3LYP/LanL2DZ calculation. Comput Mater Sci 50(3):982–990. https://doi.org/10.1016/j.commatsci.2010.10.037

5. Grimme S, Ehrlich S, Goerigk L (2011) Effect of the damping function in dispersion corrected density functional theory. J Comput Chem 32(7):1456–1465. https://doi.org/10.1002/jcc.21759

6. Borowski T, Szaleniec M (2019) Challenges in modelling metalloenzymes. In: Broclawik E, Borowski T, Radoń M (eds) Transition metals in coordination environments: computational chemistry and catalysis viewpoints. Springer International Publishing, Cham, pp 503–525. https://doi.org/10.1007/978-3-030-11714-6_17

7. Nilsson K, Lecerof D, Sigfridsson E, Ryde U (2003) An automatic method to generate force-field parameters for hetero-compounds. Acta Crystallogr D Biol Crystallogr 59(Pt 2):274–289. https://doi.org/10.1107/s0907444902021431

8. Wang J, Wang W, Kollman PA, Case DA (2006) Automatic atom type and bond type perception in molecular mechanical calculations. J Mol Graphics Model 25(2):247–260. https://doi.org/10.1016/j.jmgm.2005.12.005

9. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman P (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 24(16):1999–2012. https://doi.org/10.1002/jcc.10349

10. Lee MC, Duan Y (2004) Distinguish protein decoys by using a scoring function based on a new AMBER force field, short molecular dynamics simulations, and the generalized born solvent model. Proteins Struct Funct Bioinf 55(3):620–634. https://doi.org/10.1002/prot.10470

11. Harvey JN (2006) On the accuracy of density functional theory in transition metal chemistry. Annu Rep Sect C Phys Chem 102:203–226. https://doi.org/10.1039/B419105F

12. Hoffken HW, Duong M, Friedrich T, Breuer M, Hauer B, Reinhardt R, Rabus R, Heider J (2006) Crystal structure and enzyme kinetics of the (S)-specific 1-phenylethanol dehydrogenase of the denitrifying bacterium strain EbN1. Biochemistry 45(1):82–93. https://doi.org/10.1024/bi051596b

13. Stawoska I, Dudzik A, Wasylewski M, Jemiola-Rzeminska M, Skoczowski A, Strzalka K, Szaleniec M (2017) DFT-based prediction of reactivity of short-chain alcohol dehydrogenase. J Comput Aided Mol Des 31(6):587–602. https://doi.org/10.1007/s10822-017-0026-5

14. Fukui K (1981) The path of chemical reactions – the IRC approach. Acc Chem Res 14(12):363–368. https://doi.org/10.1021/ar00072a001

15. Tomasi J, Mennucci B, Cammi R (2005) Quantum mechanical continuum solvation models. Chem Rev 105(8):2999–3093. https://doi.org/10.1021/cr9904009

16. Becke AD (1993) Density-functional thermochemistry. III. The role of exact exchange. J Chem Phys 98(7):5648–5652. https://doi.org/10.1063/1.464913

17. Grimme S, Antony J, Ehrlich S, Krieg H (2010) A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J Chem Phys 132(15):154104. https://doi.org/10.1063/1.3382344

18. Szaleniec M, Salwinski A, Borowski T, Heider J, Witko M (2012) Quantum chemical modeling studies of ethylbenzene dehydrogenase activity. Int J Quantum Chem 112(8):1990–1999. https://doi.org/10.1002/Qua.23143

19. Vreven T, Morokuma K (2006) Chapter 3: Hybrid methods: ONIOM(QM:MM) and QM/MM. In: Spellmeyer DC (ed) Annual reports in computational chemistry, vol 2. Elsevier, Amsterdam, pp 35–51. https://doi.org/10.1016/S1574-1400(06)02003-2

20. Gotz AW, Clark MA, Walker RC (2014) An extensible interface for QM/MM molecular dynamics simulations with AMBER. J Comput Chem 35(2):95–108. https://doi.org/10.1002/jcc.23444

21. Grossfield A. WHAM: the weighted histogram analysis method. http://membrane.urmc.rochester.edu/wordpress/?page_id=126

22. Kohen A, Limbach H-H (2006) Isotope effects in chemistry and biology. Taylor & Francis, Boca Raton

23. Nagel ZD, Klinman JP (2006) Tunneling and dynamics in enzymatic hydride transfer. Chem Rev 106(8):3095–3118. https://doi.org/10.1021/cr050301x

24. Rzepa HS (2015) KINISOT. A basic program to calculate kinetic isotope effects using normal coordinate analysis of transition state and reactants. https://doi.org/10.5281/zenodo.19272

25. Anisimov V, Paneth P (1999) ISOEFF98. A program for studies of isotope effects using Hessian modifications. J Math Chem 26 (1):75–86. https://doi.org/10.1023/A:1019173509273

26. Wigner EP (1932) Crossing of potential thresholds in chemical reactions. Z Phys Chem B19:203–216

27. Glanowski M, Wójcik P, Procner M, Borowski T, Lupa D, Mielczarek P, Oszajca M, Świderek K, Moliner V, Bojarski AJ, Szaleniec M (2021) Enzymatic δ1-Dehydrogenation of 3-Ketosteroids - reconciliation of kinetic isotope effects with the reaction mechanism. ACS Catal 11:8211–8225. https://doi.org/10.1021/acscatal.1c01479

28. Li L, Li C, Zhang Z, Alexov E (2013) On the dielectric "constant" of proteins: smooth dielectric function for macromolecular modeling and its implementation in DelPhi. J Chem Theory Comput 9(4):2126–2136. https://doi.org/10.1021/ct400065j

Publikacja P2

Microbial Cell Factories

Universal capability of 3-ketosteroid Δ^1-dehydrogenases to catalyze Δ^1-dehydrogenation of C17-substituted steroids

Patrycja Wójcik[1], Michał Glanowski[1], Agnieszka M. Wojtkiewicz[1], Ali Rohman[2,3,4] and Maciej Szaleniec[1*]

Abstract

Background: 3-Ketosteroid Δ^1-dehydrogenases (KSTDs) are the enzymes involved in microbial cholesterol degradation and modification of steroids. They catalyze dehydrogenation between C1 and C2 atoms in ring A of the polycyclic structure of 3-ketosteroids. KSTDs substrate spectrum is broad, even though most of them prefer steroids with small substituents at the C17 atom. The investigation of the KSTD's substrate specificity is hindered by the poor solubility of the hydrophobic steroids in aqueous solutions. In this paper, we used 2-hydroxpropyl-β-cyclodextrin (HBC) as a solubilizing agent in a study of the KSTDs steady-state kinetics and demonstrated that substrate bioavailability has a pivotal impact on enzyme specificity.

Results: Molecular dynamics simulations on KSTD1 from *Rhodococcus erythropolis* indicated no difference in ΔG$_{bind}$ between the native substrate, androst-4-en-3,17-dione (AD; − 8.02 kcal/mol), and more complex steroids such as cholest-4-en-3-one (− 8.40 kcal/mol) or diosgenone (− 6.17 kcal/mol). No structural obstacle for binding of the extended substrates was also observed. Following this observation, our kinetic studies conducted in the presence of HBC confirmed KSTD1 activity towards both types of steroids. We have compared the substrate specificity of KSTD1 to the other enzyme known for its activity with cholest-4-en-3-one, KSTD from *Sterolibacterium denitrificans* (AcmB). The addition of solubilizing agent caused AcmB to exhibit a higher affinity to cholest-4-en-3-one (Ping-Pong bi bi $K_{mA} = 23.7$ μM) than to AD ($K_{mA} = 529.2$ μM), a supposedly native substrate of the enzyme. Moreover, we have isolated AcmB isoenzyme (AcmB2) and showed that conversion of AD and cholest-4-en-3-one proceeds at a similar rate. We demonstrated also that the apparent specificity constant of AcmB for cholest-4-en-3-one ($k_{cat}/K_{mA} = 9.25 \cdot 10^6$ M^{-1} s^{-1}) is almost 20 times higher than measured for KSTD1 ($k_{cat}/K_{mA} = 4.71 \cdot 10^5$ M^{-1} s^{-1}).

Conclusions: We confirmed the existence of AcmB preference for a substrate with an undegraded isooctyl chain. However, we showed that KSTD1 which was reported to be inactive with such substrates can catalyze the reaction if the solubility problem is addressed.

Keywords: 3-ketosteroid dehydrogenase, KSTD, Cholest-4-en-3-one Δ^1-dehydrogenase, 3-ketosteroids, 1,2-dehydrogenation, Δ^1-dehydrogenation, Cholest-4-en-3-one, Diosgenone, Cholesterol metabolism

Background

Steroids are widespread compounds in the natural environment. They are essential components of eukaryotic i.e., animal, yeast and fungi and plant, cells. Cholesterol, the parent substance of all sterols, is one of the most

*Correspondence: maciej.szaleniec@ikifp.edu.pl
[1] Jerzy Haber Institute of Catalysis and Surface Chemistry Polish Academy of Sciences, Niezapominajek 8, 30239 Krakow, Poland
Full list of author information is available at the end of the article

Wójcik *et al. Microb Cell Fact*   (2021) 20:119

Page 2 of 12

commonly occurring. Hence bacteria have evolved to use it as a carbon and energy source [1]. The microbial degradation of cholesterol proceeds under either aerobic or anaerobic conditions and the current stage of knowledge was recently reviewed by Rohman et al. [2]. 3-Ketosteroid Δ^1-dehydrogenases (KSTDs) initiate the crucial step of the degradation – steroid nucleus decomposition and as a result are found in aerobic, facultative anaerobic and strictly anaerobic microorganisms [3–5]. KSTDs catalyze regio- and stereoselective dehydrogenation between C1 and C2 atoms of steroid ring A (Fig. 1) [6]. The stage at which the double bond is introduced during cholesterol degradation is still debated and may differ between bacterial species [2].

The wide KSTDs substrate spectrum, the regioselectivity of the catalyzed reaction and mild conditions of the synthesis make those enzymes widely applied for the biologically active Δ^1-dehydrosteroids production [2]. The 1,2-dehydrogenation process can be carried out with either a whole-cell system (native and recombinant bacteria are used) or isolated enzymes [7, 8]. The poor solubility of steroids in aqueous solutions is usually circumvented by the application of pseudo-crystallofermentation technique (i.e., steroid substrate in the solid-state) or by the addition of organic solvents and various solubilizing agents, e.g., vegetable oils, phospholipids, glycols, dextrans or cyclodextrins [8–15]. The biocompatibility of cyclodextrins makes them valuable solubilizing agents for scale-up microbial steroid biotransformation. Importantly, they do not exert a detrimental influence on the stability of the purified enzymes [16, 17].

Although KSTDs substrate scope is generally broad, most of the enzymes prefer steroids with small substituents on the C17 position [2]. Only a few of them exhibited activity towards more branched substrates, such as cholest-4-en-3-one, 3-oxo-5β-cholan-24-oic acid, the methyl ester of 3-oxo-4-cholen-24-oic acid, cortisone acetate, hydrocortisone acetate or 21-acetoxy-pregna-4,9(11),16-triene-3,20-dione [18–23]. The comparative study of substrate specificity is additionally hindered by poor or very bad solubility of steroids in the water phase in which all kinetic enzyme assays are conducted [8].

3-Ketosteroid Δ^1-dehydrogenase from *Streolibacterium denitrificans* (AcmB) is a good example of the enzyme for which the native substrate remains an open question [24]. So far two 3-ketosteroids have been proposed as the AcmB natural substrates. Cholest-4-en-3-one (**2** in Fig. 1) was initially proposed as the AcmB native substrate [21, 25, 26]. However, subsequent research demonstrated a high affinity of the enzyme for C19 – C21 steroids and indicated that AcmB is located on the cytoplasmic side of the inner membranes of *S. denitrificans* [8, 27]. These results suggested that androst-4-en-3,17-dione (AD; **1**) may be the enzyme native substrate [28]. Nevertheless, AcmB's ability to catalyze Δ^1-dehydrogenation of steroids with the aliphatic side chain on the C17 position seemed interesting and quite unusual.

In this work, we demonstrate that substrate bioavailability can significantly influence the kinetic results and lead to misinterpretation of enzyme specificity. We used 2-hydroxypropyl-β-cyclodextrin (HBC) as a solubilizing agent to study the steady-state kinetics for AcmB
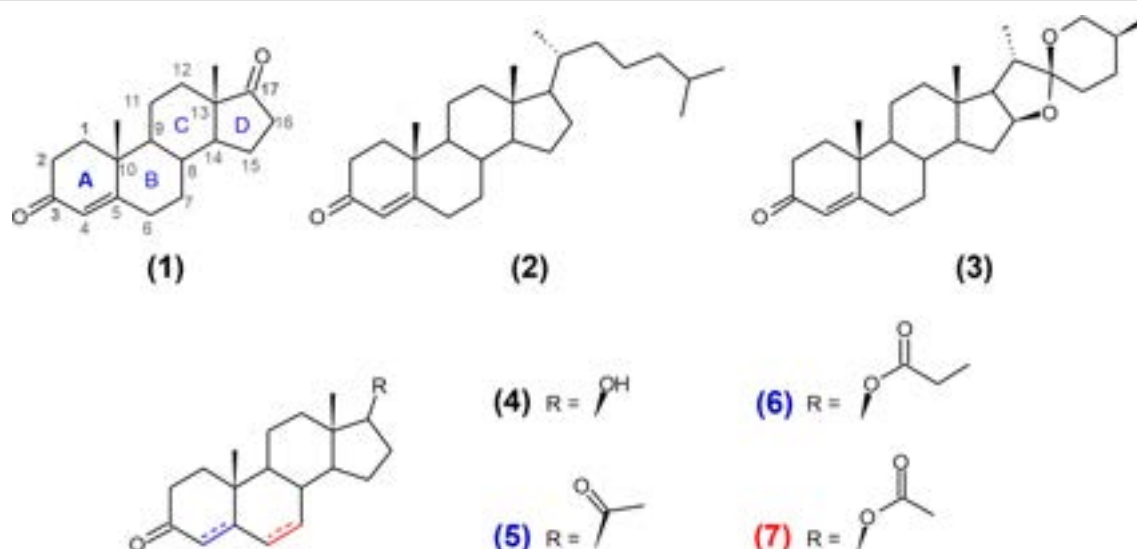
Fig. 1 Structures of compounds tested in this study as substrates: (**1**) androst-4-en-3,17-dione, (**2**) cholest-4-en-3-one, (**3**) diosgenone, (**4**) androstanolone, (**5**) progesterone, (**6**) testosterone propionate and (**7**) 6-dehydrotestosterone acetate

Wójcik *et al. Microb Cell Fact*   (2021) 20:119

Page 3 of 12

and 3-ketosteroid Δ^1-dehydrogenase from *Rhodococcus erythropolis* SQ1 (KSTD1) with AD and cholest-4-en-3-one. In contrast to previous reports [2, 29], we have demonstrated never reported before KSTD1 activity towards cholest-4-en-3-one and even more complex 3-ketosaponin – diosgenone (**3**). Moreover, we proved that AcmB indeed exhibits a significantly higher affinity to cholest-4-en-3-one than to AD while KSTD1 indeed prefers smaller ketosteroids.

Methods
Materials
All chemicals were purchased from Sigma-Aldrich, Tokyo Chemical Industry, Carl Roth or BioShop unless otherwise specified. (25*R*)-Spirost-1,4-dien-3-one (diosgenone) was a gift of Jacek Morzycki from the University of Bialystok.

Proteins expression and purification
The AcmB was cloned, expressed and purified as described previously at Wojtkiewicz et al. work [8].

A gene acmB2 was amplified from genomic *S. denitrificans* DNA using 0.04 μM of primers: forward (5'-TAC TTCCAATCCAATGCTATGAGCGGAGAAACTTTC G-3') and reverse (5'-TTATCCACTTCCAATTCATAC CGCGCTCCG-3'), 1 U Platinum Taq DNA Polymerase High Fidelity (ThermoFisher), 2.4% dimethyl sulfoxide (DMSO), 200 μM dNTP mix, High Fidelity PCR Buffer, 2 mM $MgSO_4$, 160 ng of template DNA and thermocycler (BioRad) program as follows: 94 °C for 2 min, 32 cycles of (94 °C for 15 s, 63 °C for 30 s, 68 °C for 120 s) with a final elongation at 68 °C for 10 min. After gel electrophoresis a band of size ~ 1.7 kbp was excised from agarose gel and purified using Gel-Out kit (A&A Biotechnology) according to the manufacturer's instructions. Next, gene encoding AcmB2 (GenBank: SMB21450.1) was cloned into a pMCSG7 vector. 25 μL pMCSG7 vector (200 ng) was treated with 2.5 μL SspI in a total of 50 μL 1 × Tango Buffer (Thermo Scientific) supplemented with 4 mM dithiothreitol (DTT) for 1.5 h at 37 °C. The linearized, gel-purified vector was prepared for ligase-independent cloning. Namely, 200 ng of pMCSG7 vector were incubated in 40 μL 1 × T4 Polymerase buffer (Novagen) supplemented with 4 mM DTT, 1 μL of T4 Polymerase (Novagen) and 200 μM dGTP (Promega) for 40 min at room temperature, which was followed by a 10 min inactivation of the polymerase at 75 °C. Similarly, the acmB2 gene was prepared, however, the nucleotides were substituted with dCTP (Promega). Then 3.5 μL of vector preparation was reacted with 6 μL of gene preparation and incubated on ice for 3 h. The mixture was transformed into *Escherichia coli* XL10-Gold Ultracompetent Cells (Agilent Technologies) and the cells were grown overnight in 2% (w/v) Lennox Broth (LB) supplemented with ampicillin 100 μg mL^{-1} and 34 μg mL^{-1} chloramphenicol. Afterward, the plasmid was extracted using Plasmid Miniprep DNA Purification Kit (EURx) according to the manufacturer's instructions, sequenced by the Sanger sequencing method (Genomed, Poland) and transformed into calcium chloride chemically competent *E. coli* BL21(DE3)Magic (Creative Biolabs). Bacteria cultivation, harvesting and disruption, as well as enzyme purification were conducted as described for AcmB [8].

A gene encoding 3-ketosteroid Δ^1-dehydrogenase from *R.erythropolis* (KSTD1) was cloned into pET15b vector [30] and used to transform *E. coli* BL21(DE3) Magic cells. An overnight culture of the transformed cells was grown in 2% (w/v) LB supplemented with 100 μg mL^{-1} ampicillin and 50 μg mL^{-1} kanamycin at 37 °C, 180 rpm. The preculture was hundred times diluted in 1 L of fresh LB containing 0.5 M D-sorbitol and antibiotics and grown at 37 °C, 180 rpm. When the OD_{600} reached 0.6, the temperature was reduced to 16 °C and the culture was induced with 100 μM isopropyl β-D-1-thiogalactopyranoside (IPTG). After 48 h the *E. coli* cells were centrifuged for 1 h at 4500 g, 4 °C. The harvested cells were resuspended in a 1:5 ratio (w/v) in the 50 mM Tris–HCl pH 8.5 buffer, 100 mM NaCl, 10% (w/v) glycerol, 5 mM β-mercaptoethanol (BME) and 10 mM imidazole and supplemented with 100 μM phenylmethylsulfonyl fluoride (PMSF). The cells were disrupted by sonication (Sonics Vibra-Cell VCX500, 3 s on, 5 s off, 5 min, 40% amplitude, 150 000 J). The cell-free extract was obtained by centrifugation at 40 000 g, for 1 h at 4 °C. Purification of KSTD1 was carried out on 5 mL HisTrap HP (GE Healthcare) column using FPLC system (BioRad NGC Quest 10 Plus) and linear imidazole gradient 10 – 300 mM in the previously described buffer. The fractions that absorbed UV-light at wavelength 450 nm were collected, desalted using dialysis tubing with 10 kDa cut-off (Thermo Fisher Scientific) according to the manufacturer's instructions and stored at −20 °C.

The concentration of purified AcmB, AcmB2 and KSTD1 was determined using a free FAD extinction coefficient of 11 300 M^{-1} cm^{-1} at 450 nm.

Spectrophotometric activity assay
The steady-state kinetic parameters (K_m, V_{max}, k_{cat}, k_{cat}/K_m) for steroids were determined in a spectrophotometric activity assay using UV-2700 spectrophotometer (Shimadzu) in 0.5 mL quartz cuvettes with 10 mm path length. The measurements were carried out in 0.1 M K_2HPO_4/KH_2PO_4 buffer pH 6.5 (AcmB) or 0.1 M K_2HPO_4/KH_2PO_4 buffer pH 8.0 (KSTD1) with 0.029 – 0.174 mM 2,6-dichloroindophenol (DCPIP), 2% (w/v) 2-hydroxypropyl-β-cyclodextrin (HBC; 14.54 mM),

Wójcik *et al. Microb Cell Fact*   (2021) 20:119

Page 4 of 12

5 to 500 μM concentrations of the steroids dissolved in 2-methoxyethanol (EGME, the final concentration 2% (v/v)) and 0.30 μM of AcmB or 0.23 μM of KSTD1, respectively. The reduction of DCPIP was followed at 700 nm ($\varepsilon_{\text{pH 6.5, HBC}} = 10\ 691\ \text{M}^{-1}\ \text{cm}^{-1}$ or $\varepsilon_{\text{pH 8.0, HBC}} = 11\ 627\ \text{M}^{-1}\ \text{cm}^{-1}$). All measurements were performed in triplicates at 30 °C. The linear slopes were obtained with linear regression fitted to the initial parts (5 s) of the kinetic curves. The Ping-Pong bi bi (non-sequential) mechanism model was fitted to the obtained data for AcmB. The linear function was fitted to the data obtained for KSTD1. All data processing was conducted using OriginPro 2019 software.

HPLC analysis

To confirm KSTD1 activity with diosgenone mini reactor tests were performed. The reaction mixture (1 mL) consisted of 50 mM Tris–HCl buffer pH 8.0, 0.15 mM DCPIP, 2% (w/v) HBC, 0.1 mM diosgenone in dioxane (the final concentration 2% (v/v)) and 87.7 nM KSTD1. The reactions were carried out in duplicate at 30 °C, in the thermoblock at 800 rpm for 20 min. To determine the conversion of the substrate, the reaction mixtures were diluted 1:1 with acetonitrile (ACN), centrifuged at 15 000 g for 5 min and analyzed with HPLC–DAD (Agilent 1100) according to [8].

To confirm AcmB2 activity with AD, cholest-4-en-3-one and diosgenone mini reactor tests were performed. The reaction mixture (1 mL) consisted of 50 mM Tris–HCl buffer pH 8.0, 0.15 mM DCPIP, 2% (w/v) HBC, 0.1 mM of substrate in EGME (the final concentration 2% (v/v)) and 6.55 nM AcmB2 for AD and cholest-4-en-3-one or 65.5 nM for diosgenone. The reaction was carried out in duplicate at 30 °C, in the thermoblock at 800 rpm for 960 h. To determine the conversion of the substrates, the samples were prepared and analyzed as described above.

Determination of HBC/steroid stability constant

The stability constants (K_S) of the inclusion complex between HBC and steroids were determined according to the modified protocol described by Ma et al.[31]. The mixtures of various concentrations of HBC (from 0 to 98 mM), 2% (v/v) EGME and the double molar excess of AD or cholest-4-en-3-one were incubated in the thermoblock at 30 °C, 1000 rpm for 48 h. Afterward, the suspensions were filtered through 0.45 μm membrane filters, diluted 0–200 times with ACN, centrifuged at 15 000 g for 15 min, and concentration of soluble steroid was quantified by HPLC–DAD using calibration on external standards [8]. The linear function was fitted to the data obtained for AD. The stability constant for AD:HBC 1:1

complexes ($K_{1:1}$) was calculated from Eq. (1), where S_0 is an y-intercept [32, 33].

$$K_{1:1} = \frac{\text{slope}}{S_0(1 - \text{slope})} \tag{1}$$

The quadratic function ($[S] = a[HBC]^2 + b[HBC] + S_0$) was fitted to the data attained for cholest-4-en-3-one. The stability constants for cholest-4-en-3-one, HBC 1:1 ($K_{1:1}$) and 1:2 ($K_{1:2}$) complexes, were calculated from the Eq. (2) and (3), respectively [33].

$$K_{1:1} = \frac{b}{S_0} \tag{2}$$

$$K_{1:2} = \frac{a}{S_0 K_{1:1}} \tag{3}$$

S_0 refers to the solubility of steroids in 2% (v/v) EGME/water solution (144.2 ± 7.1 μM for AD and 4.8 ± 0.5 μM for cholest-4-en-3-one).

MD simulations

The structure of KSTD1 was taken from Protein Data Bank (PDB ID: 4c3y [34], resolution 2.3 Å). 1,4-androstadiene-3,17-dione (ANB) and FAD cofactor were already present in the active site of the crystal structure. KSTD1 was determined to be monomeric in solution [34] and, thus, we used the monomeric structure of the enzyme to perform MD simulations. To place structures of androst-4-en-3,17-dione (**1**), cholest-4-en-3-one (**2**) and diosgenone (**3**), androstanolone (**4**), progesterone (**5**), testosterone propionate (**6**) and 6-dehydrotestosterone acetate (**7**) in the enzyme active site, the Kabsch method was used [35] with ANB as a template. The geometry of each substrate was obtained by optimization in the gas phase at B3LYP/6-31G(d,p) level of theory with Gaussian16 [36]. PropKa ver 3.1 [37, 38] and H++packages [39] were used to determine the protonation states of titratable amino acids. The optimal pH for enzyme activity was reported to be about 8.0, so this value was selected for model construction. We assumed that Tyr318 is deprotonated due to the mechanistic hypothesis. AMBER parameters for ketosteroids and tyrosyl anion were obtained with Gaussian 16 software package [36] (at B3LYP/6-31G(d,p) level of theory) and ANTECHAMBER program from AmberTools [40]. Parameters for FAD were obtained from the RESP ESP charge DataBase (R.E.DD.B) [41]. A total amount of 33 sodium cations were added to neutralize the system. Finally, the system was soaked in a $93.5 \times 76.5 \times 72.0\ \text{Å}^3$ box with TIP3P [42] water molecules.

To prepare the model for MD simulation, the system was optimized with Amber package [43] with ff03 force field [44], heated from 0 to 303 K over 100 ps with NVT conditions (canonical ensemble), and equilibrated for

Wójcik *et al. Microb Cell Fact*   (2021) 20:119

Page 5 of 12

100 ps with NPT (isothermal-isobaric) ensemble. Finally, 60 ns of NPT molecular dynamic simulation at 303 K was performed.

The ΔG of substrate binding was estimated using the MM-PBSA algorithm in a Poisson Boltzmann (PB) solvation [45] for 500 frames (snapshots from every 10 ps) from selected 5 ns of each MD simulation. Ionic strength was set to 0.15 M. To select a fragment of trajectory, distances describing the position of steroid were calculated: $d(O_{keto}\text{-}OH_{Tyr487})$, $d(H2\beta\text{-}OH_{Tyr318})$ and $d(H1\alpha\text{-}N_{FAD})$. The first of these parameters is describing the hydrogen bond between the 3-keto group of ketosteroid and Tyr487, the next two are related to the postulated mechanism of the catalyzed reaction, which postulates that H2β and H1α are transferred to Tyr318 and FAD respectively. For MMPBSA calculation we selected 5 consecutive ns, during which the sum of the standard deviation of these parameters was minimal.

Results

Theoretical prediction of substrate specificity

In the crystal structure of KSTD1 in complex with the ligand 1,4-androstadiene-3,17-dione (*cf.* **1**), the steroid ring core (gonane ring system) of the ligand was buried in the active site pocket of the enzyme with its D-ring pointing towards the solvent-accessible area [34]. Likewise, the MD simulation of KSTD1:cholest-4-en-3-one complex revealed that the enzyme binds the steroid ring core while the isooctyl substituent at C17 position protrudes out of the active site and generally does not prevent it from the active binding (Fig. 2).

During the simulation, we also did not observe any significant clashes of the isooctyl side chain with the protein which would explain lack of the enzyme activity toward this class of substrate. From the reaction point of view, the most important interactions between substrate and the enzyme were limited to the 3-keto group of steroid ring A and its closest neighborhood (especially vicinity of C1 to FAD and C2 to Tyr318). However, in the literature, cholest-4-en-3-one was reported as inactive in the reaction with KSTD1 [29]. Therefore, we decided to check if the binding free energies calculated for C17-substituted and non-substituted steroids are consistent with the reported experimental results. Free energies of substrate binding were estimated for KSTD1 with MM-PBSA (Table 1). All studied compounds bound favorably to the active site with ΔG$_{bind}$ in the range of − 11.0 to − 7.9 kcal/mol. Interestingly, there was no difference in ΔG$_{bind}$ between the native substrate of KSTD1 from *R. erythropolis,* androst-4-en-3,17-dione (− 8.02±0.13 kcal/mol, **1**), and cholest-4-en-3-one (− 8.40±0.19 kcal/mol, **2).** Based on this result we formed the hypothesis that in principle it should be possible to convert with KSTD1 not only

Table 1 The estimated free energy of the binding of steroids to the KSTD1 active site

Ligand	ΔG$_{bind}$ [kcal/mol]	Std. err. of ΔG$_{bind}$ [kcal/mol]
Androst-4-en-3,17-dione (**1**)	− 8.02	0.13
Cholest-4-en-3-one (**2**)	− 8.40	0.19
Diosgenone (**3**)	− 6.17	0.15
Androstanolone (**4**)	− 10.97	0.15
Progesterone (**5**)	− 8.23	0.17
Testosterone propionate (**6**)	− 7.85	0.15
6-Dehydrotestosterone acetate (**7**)	− 8.53	0.15
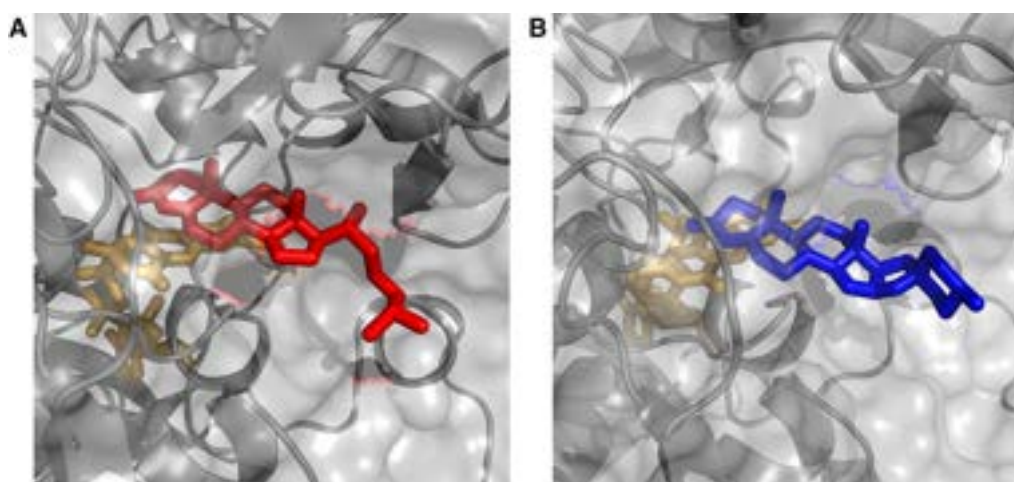
Fig. 2 Frame from MD simulation on KSTD1 with **A** cholest-4-en-3-one (red) and **B** diosgenone (blue) in the active site. FAD is marked with light orange (see SI for pdb structures)

Wójcik *et al. Microb Cell Fact*   (2021) 20:119

Page 6 of 12

Fig. 3 Phase solubility diagrams for androst-4-en-3,17-dione (gray squares) and cholest-4-en-3-one (red circles)
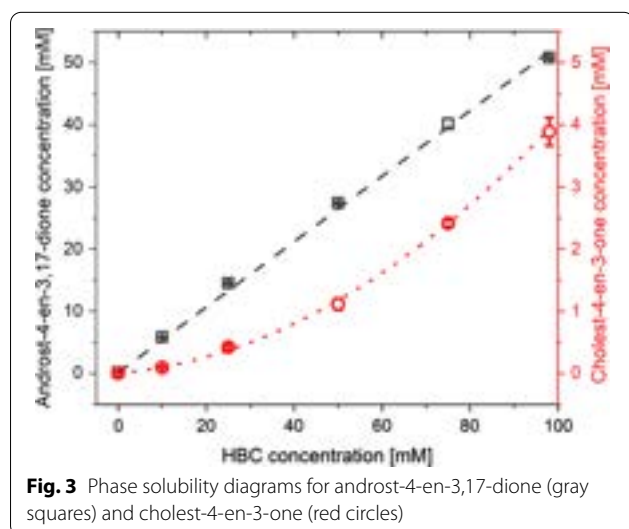
3-ketosteroids with a degraded side chain (**1, 4, 5–7**) but also steroids such as **2** or even 3-ketosaponins as **3** (Fig. 2).

HBC/steroid inclusion complex formation

The phase solubility diagrams of AD (**1**) and cholest-4-en-3-one (**2**) with HBC were determined in the HBC concentration range from 0 to 98 mM at 30 °C, i.e., the temperature of catalytic assays (Fig. 3). As expected, the solubility of both steroids increased with higher HBC concentration. However, the solubility of more hydrophobic **2** in 98 mM HBC solution turned out to be over 10-times lower than the AD.

The concentration of AD in the function of HBC concentration was fitted with the linear function ($R^2 > 0.998$) corresponding to A_L-type phase diagram. The linear course of phase solubility diagram with the slope less than 1 (0.527 ± 0.009) indicates that HBC with AD forms the 1:1 inclusion complex. Meanwhile, the cholest-4-en-3-one/HBC phase solubility diagram turned out to be the A_P-type and the data were fitted with quadratic function ($R^2 > 0.998$). A_P-type phase solubility diagram may indicate the formation of higher-order complexes with respect to cyclodextrin at higher HBC concentrations [31, 32]. The fit parameters are presented in Additional file 1: Tables S1–S4, while the values of $K_{1:1}$ and $K_{1:2}$ were shown in Table 2.

Table 2 Solubility parameters of steroids in the solution of HBC and 2% EGME at 30 °C

Substrate	Solubility curve type	$K_{1:1}$ [M^{-1}]	$K_{1:2}$ [M^{-1}]
Androst-4-en-3,17-dione (**1**)	A_L	7714.4	–
Cholest-4-en-3-one (**2**)	A_P	1277.1	56.8

The obtained constants were used to calculate the percentage of free and complexed steroids in the reaction medium. For androst-4-en-3,17-dione 99% of the substrate has been complexed by HBC. The ratio of free and complexed (1:1 and 1:2) cholest-4-en-3-one has varied slightly depending on the initial steroid concentration (Additional file 1: Table S2) and averaged 2.9% of the free form, 53.5% of the steroid in 1:1 complex and 43.6% of the steroid in 1:2 complex.
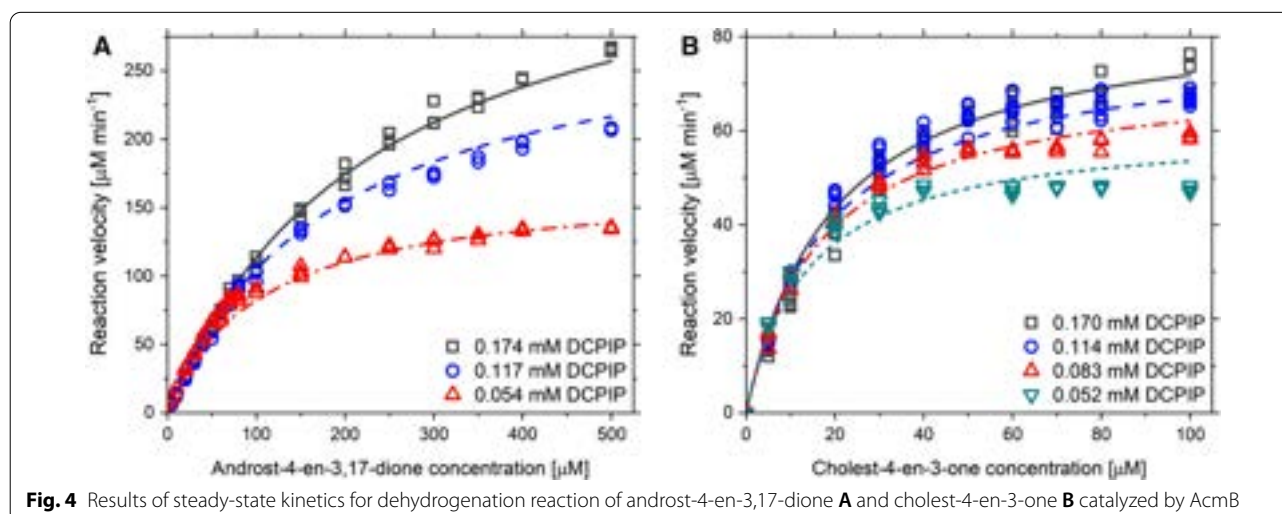
AcmBs native substrate

To determine the substrate specificity of AcmB we conducted steady-state kinetic studies for androst-4-en-3,17-dione and cholest-4-en-3-one. 2,6-Dichloroindphenol was used as an artificial electron acceptor. Since the solubility of cholest-4-en-3-one in aqueous solutions is close to zero, measurements were carried out with 2% HBC and 2% EGME. The reaction velocities were presented in Fig. 4. The Ping-Pong bi bi (non-sequential) mechanism was fitted to the obtained data (4) [46].

$$v = \frac{V_{max}[S]_t[DCPIP]}{K_{mS}[DCPIP] + K_{mDCPIP}[S]_t + [S]_t[DCPIP]} \quad (4)$$

Due to the presence of HBC in the system, the substrate was available to the enzyme mostly in the form of 1:1 and/or 1:2 complexes (99% for S(HBC) AD complex, and 53.52 and 43.55% for S(HBC) and S(HBC)$_2$ complexes with cholest-4-en-3-one, see Additional file 1: Table S2 for details). As a result, only in the case of AD, we could assume that the observed kinetics was mostly coming from the S(HBC) form of the substrate while in the case of cholest-4-en-3-one the kinetics of E:S complex formation was much more complicated and proceeds according to Eq. (5).

$$\text{(5)}$$

E + S(HBC)$_2$ $\xrightarrow{k_1^{S(HBC)2}}$ E:S

$K_{1:2}$

E + S(HBC) $\xrightarrow{k_1^{S(HBC)}}$ E:S $\xrightarrow{k_2}$ E' + DCPIP $\underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}}$ E':DCPIP $\xrightarrow{k_4}$ E + DCPIPH$_2$

$K_{1:1}$

E + S $\underset{k_{-1}^{S}}{\overset{k_1^{S}}{\rightleftharpoons}}$ E:S

Wójcik *et al. Microb Cell Fact*   (2021) 20:119

Page 7 of 12
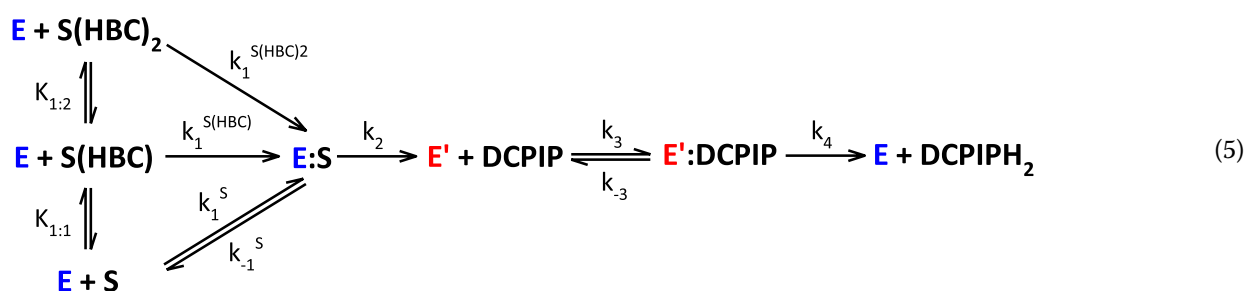
Fig. 4 Results of steady-state kinetics for dehydrogenation reaction of androst-4-en-3,17-dione **A** and cholest-4-en-3-one **B** catalyzed by AcmB

E – enzyme (KSTD); E' – enzyme in its reduced form; S – substrate 1 (steroid); HBC – 2-hydroxypropyl-β-cyclodextrin and DCPIP – substrate 2 (artificial electron acceptor).

Although it is possible to derive the kinetic equation that takes into account all forms of the substrate present in the solution, the correct fitting of all constants requires collecting kinetic data for different concentrations of HBC, steroid and DCPIP. Our experimental data turned out to be significantly underdetermined to handle such a complex non-linear fit. However, as the main purpose of our study was to compare different 3-ketosteroid dehydrogenases under the same conditions (here concentration of HBC which determines the distribution of S:HBC complexes) we have decided to present the kinetic parameters in the function of the total substrate concentration. Due to that, kinetic curves were fitted to the concentration of the total substrate.

Under these conditions, the maximum velocity of reaction for AcmB (V_{max}) was over 8 times higher for AD than for cholest-4-en-3-one. However, AcmB affinity (K_{mS}) to cholest-4-en-3-one turned out to be over 20-times higher compared with AD. As a result, the enzyme catalytic efficiencies (k_{cat}/K_{mS}) for cholest-4-en-3-one was more than two-fold higher than that determined for AD (Table 3).

It should be noted here that DCPIP is also complexed by HBC. This fact however does not impact the comparisons of K_{mS} values, as experiments were conducted for both substrates with the same type and concentrations of oxidation agent (DCPIP) and the HBC concentration was 30 to 84-times higher than DCPIP concentration. The influence of DCPIP complexation (such as decrease of HBC concentration) on obtained kinetic parameters was negligible.

In the genome of *S. denitrificans* we have identified two other putative AcmB homologs, which biological role has not been yet established. For this study, we selected gene encoding "3-oxosteroid 1-dehydrogenase" (GenBank: SMB21450.1) which we named AcmB2 due to its high sequence similarity to AcmB (44% identity and 62% similarity). The recombinant enzyme was expressed in *E. coli* BL21(DE3)Magic and purified on Ni–NTA matrix. To check if AcmB2 also exhibited activity with C17-substituted steroids we performed mini reactor tests with 100 µM of AD, cholest-4-en-3-one or diosgenone with the addition of 2% HBC (Additional file 1: Figs. S1–S8). In a reactor with diosgenone, we introduced a fivefold higher concentration of AcmB2 than for other mini reactors to account for the observed lower rate of the initial tests. After 16 h of reaction, we observed significant conversions of all substrates i.e., 77.2±1.5% for AD, 80.2±2.7% for cholest-4-en-3-one and 53±3.5% for diosgenone. These results show that other 3-ketosteroid dehydrogenase from *S. denitrificans* exhibits similar substrate specificity as AcmB.

KSTD1 substrate spectrum

The kinetic measurements for KSTD1 were performed at the same conditions as described previously for AcmB, except for the pH of the buffer solution. Surprisingly, the affinity of KSTD1 from *R. erythropolis* to AD and cholest-4-en-3-one turned out to be significantly lower than this exhibited by AcmB and it was not possible to reach enzyme saturation conditions in the presence of HBC. The solubility of steroids was limited even with solubilizing additives. On the other hand, a further increase of the HBC concentration results in an apparent K_m shift toward higher values as a consequence of substrate sequestration (i.e., decrease of free substrate

Wójcik *et al. Microb Cell Fact*   (2021) 20:119

Page 8 of 12

concentration and stabilization of the substrate in S(HBC)$_n$ form). Nevertheless, we have decided to estimate the first-order rate coefficients (V_{max}/K_m) for the low substrate concentrations. The linear function was fitted to the reaction velocities determined for 5 – 100 μM steroids concentrations (Fig. 5). To obtain comparable kinetic parameters, the same procedure was applied to AcmB (Additional file 1: Fig. S9).

For the small concentrations of the first substrate ($S << K_m$) the velocity equation becomes invariant to the concentration of the second substrate (DCPIP) and simplifies to a first-order linear equation:

$$v = \frac{V_{max}[S]}{K_m} \tag{6}$$

The apparent catalytic efficiency k_{cat}/K_m of AcmB and KSTD1 for AD was of the same order. KSTD1 proved to be a slightly (22%) better catalyst for the AD Δ^1-dehydrogenation than AcmB. In the case of the reaction with cholest-4-en-3-one, the k_{cat}/K_m value measured for AcmB was almost 20-times higher than for KSTD1. The KSTD1 was 10-times better in the Δ^1-dehydrogenation of AD vs cholest-4-en-3-one (Table 4).

The obtained results confirmed undeniably the ability of KSTD1 to catalyze Δ^1-dehydrogenation of steroids with the aliphatic chain on the C17 position. Due to the low cholest-4-en-3-one bioavailability and the lower KSTD1 affinity to this steroid, it was not possible to determine the actual KSTD1 substrate spectrum without the addition of solubilizing agent [29].

Based on the results of MD simulations for KSTD1 and diosgenone ES complex and the experimentally confirmed activity of AcmB and AcmB2 with diosgenon [8] we decided to investigate if KSTD1 is also able to catalyze the reaction with 6-member ring steroids. The mini reactor tests proved that in presence of 2% HBC 100 μM diosgenone can be converted with excellent 97.3 ± 0.8% conversion in 20 min.

Discussion

The results of theoretical MD simulations for 3-ketosteroid Δ^1-dehydrogenase from *R. erythropolis* SQ1 have shown that neither isooctyl side chain of cholest-4-en-3-one nor polycyclic substituent of diosgenone is a steric hindrance for the reaction catalyzed by KSTD1. Only the gonane ring is recognized by KSTD active site and the additional substituents or rings are protruding into the solvent, outside of the enzyme. This hypothesis was successfully confirmed by experimental kinetic studies and reactor tests. Nevertheless, the affinity and the catalytic efficiency of KSTD1 from *R. erythropolis* for cholest-4-en-3-one turned out to be significantly lower than that determined for AcmB from *S. denitrificans*. Additionally, a more hydrophobic character of the AcmB surface when compared to KSTD1 may also contribute to its higher affinity to more hydrophobic substrates [47].

Our results are in line with reports of Wang et al. [22], who showed that KSTD from *R. erythropolis* WY 1406 strain can also convert substrates with extended C17 substituent, like 21-acetoxy-pregna-4,9(11),16-triene-3,20-dione. In addition to AcmB and the new AcmB2 [8, 21], there were some literature reports about KSTDs activity towards steroids with a more complex side chain on the C17 position [18–20, 23, 48]. Zhang et al. described five KSTD isoenzymes from *Gordonia neofelifaecis* of which four exhibited activity toward cholest-4-en-3-one. Importantly, the authors conducted
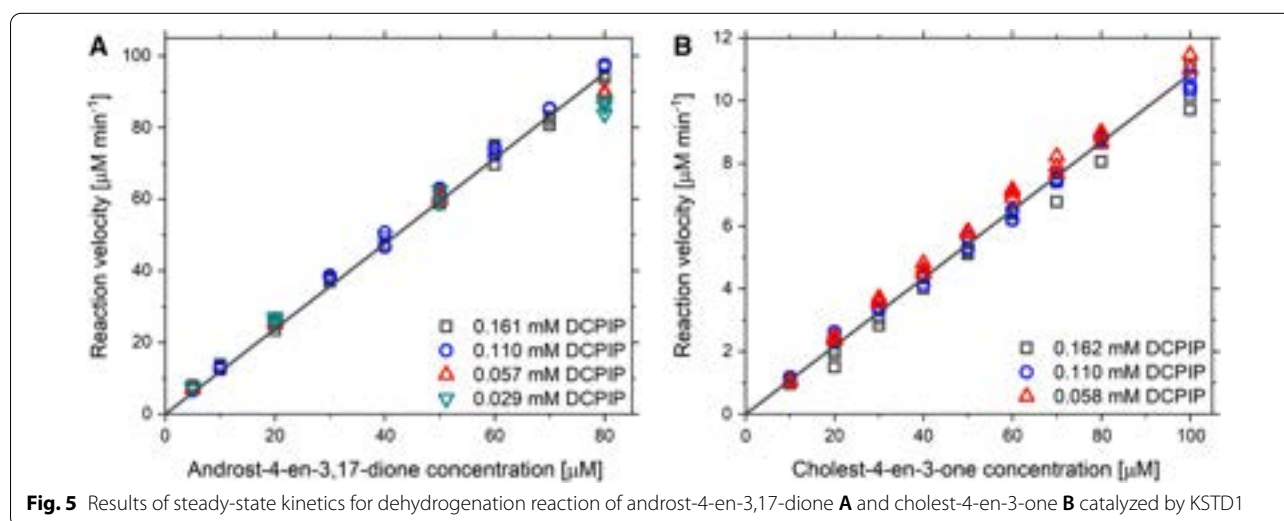
Fig. 5 Results of steady-state kinetics for dehydrogenation reaction of androst-4-en-3,17-dione **A** and cholest-4-en-3-one **B** catalyzed by KSTD1

Wójcik *et al. Microb Cell Fact*   (2021) 20:119

Page 9 of 12

experiments with 0.05% Tween 80 acting as a solubilizing agent [20].

Poor solubility of steroids with the aliphatic or polycyclic substituents at the C17 atom results in their low bioavailability to the enzyme. Compounded with a low affinity of most KSTDs to more complex steroids, these two factors may lead to inaccurate conclusions regarding substrate specificity. The results obtained for KSTD1, AcmB, AcmB2, as well as for KSTDs from *G. neofelifaecis* [20] indicate that the use of cyclodextrins or detergents can influence enzymes substrate specificity analysis. Such analysis, however, requires determination of the HBC/steroid stability constants which in turn help in the determination of the real concentration of the substrate, but complicate the observed kinetics due to different ΔGs of E:S formation for S, S(HBC) or S(HBC)$_2$ forms of the substrate. Furthermore, it should be underlined here that there are several different methods for the determination of these constants (HPLC, NMR, FT-IR, XRD or calorimetry-based, etc.) and they usually differ in the delivered values [49, 50].

Despite these difficulties, the application of solubilizing agent seems indispensable for the determination of the active substrates, especially when KSTD affinity to the compounds with an aliphatic chain on the C17 position is high, as in AcmB (apparent K_{mS} of 23.7 μM). Our results indicate that AD may not be a native substrate of AcmB. The established affinity toward AD is, however, the affinity toward S(HBC) complex (apparent K_{mS} of 529.2 μM) as we know from our previous kinetic studies that K_m value measured in the assay without HBC was an order of magnitude lower (apparent $K_m = 59.6 \pm 3.0$ μM at 200 μM DCPIP and 30 °C) [8] than in the assay with HBC (Tables 3 and 4). The K_{mS} determined in this work for cholest-4-en-3-one is lower than those observed for AD without HBC. Furthermore, as 97% of the cholest-4-en-3-one was in the form of complexes with HBC we can assume that the real K_{mS} for cholest-4-en-3-one (if it was possible to measure it without HBC) is even lower.

The difference in the affinity of KSTDs to C17-substituted steroids might result from their structural differences. The homology modeling of KSTD from *Arthrobacter simplex* (active towards cholest-4-en-3-one) showed the presence of an additional loop close to the enzyme active site, non-observed in the crystal structure of KSTD1 form *R. erythropolis* [34, 51]. Interestingly, the same feature was noticed in the AcmB homology model [52].

Table 4 Apparent kinetic parameters of the purified AcmB and KSTD1

Substrate	AcmB k_{cat}/K_m [M^{-1} s^{-1}]	KSTD1 k_{cat}/K_m [M^{-1} s^{-1}]
Androst-4-en-3,17-dione (**1**)	$(4.00 \pm 0.04) \cdot 10^6$	$(5.17 \pm 0.03) \cdot 10^6$
Cholest-4-en-3-one (**2**)	$(9.25 \pm 0.17) \cdot 10^6$	$(4.71 \pm 0.03) \cdot 10^5$

Our findings have implications for the understanding of the physiological role of AcmB. Initial studies on the anoxic cholesterol metabolism of *S. denitrificans* reported new metabolites, cholest-1,4-diene-3-one and 25-hydroxycholest-1,4-dien-3-one, and suggested that the double bond between the C1 and C2 atoms have had to be introduced at the early stage of cholesterol mineralization. As a result, cholest-4-en-3-one was considered as the AcmB native substrate. [25, 26]. However, subsequent kinetic studies demonstrated the AcmB preference for C19 – C21 steroids [21], which was supported by the localization of the enzyme on the cytoplasmic side of the bacterial inner membrane [27]. Due to that fact, Lin et al. deduced that the AcmB native substrate is rather the AD [27]. Kinetic studies indicating AcmB localization were estimated based on the enzyme activity measured in particular fractions, not by e.g., antibody technique and AcmB specific activity was calculated for the total protein extracted from the membrane. Importantly, the KSTD activity was still detected significantly in the peripheral-periplasmic fraction [27]. In this work, we demonstrated, that a study of the enzyme–substrate specificity and affinity in the case of hydrophobic steroids without the solubilizing agent addition may lead to misinterpretation of achieved results. In the presence of HBC, the AcmB affinity toward cholest-4-en-3-one turned out to be significantly higher than for AD (even taking into consideration all complications introduced by substrate complexation by HBC).

Finally, one has to remember about two other AcmB sequence homologs, from which we confirmed 3-ketosteroid Δ^1-dehydrogenase activity for one of them, AcmB2. Therefore it may well be that *S. denitrificans* deploys different KSTD enzymes on both sides of its inner membrane, which specialize in converting different types of 3-ketosteroids (as reported already for KSTD

Table 3 Apparent kinetic parameters of the purified AcmB

Substrate	K_{mS} [μM]	K_{mDCPIP} [μM]	V_{max} [μM min^{-1}]	k_{cat} [s^{-1}]	k_{cat}/K_{mS} [M^{-1} s^{-1}]
Androst-4-en-3,17-dione (**1**)	529.2 ± 32.5	223.5 ± 16.9	859.2 ± 45.5	46.1 ± 2.3	$(0.87 \pm 0.05) \cdot 10^5$
Cholest-4-en-3-one (**2**)	23.7 ± 1.16	37.3 ± 3.3	104.6 ± 2.7	5.8 ± 0.2	$(2.45 \pm 0.01) \cdot 10^5$

Wójcik *et al. Microb Cell Fact*   (2021) 20:119

Page 10 of 12

isoenzymes from *R. erythopolis* [53, 54]). This aspect definitely requires further studies.

Conclusions

Our results indicate that it is necessary to test KSTD substrates in the presence of solubilizing agents such as HBC, to discern truly inactive steroids from those which under experimental conditions are in too low concentrations. Unfortunately, the presence of HBC introduces additional complications to the kinetic model due to the presence of different forms of the substrate (S and $S(HBC)_n$). Our modeling studies suggest that in principle all KSTDs, which share structural characteristics with KSTD1 should be able to convert more hydrophobic substrates with extended ring system or C17 substituent. This conclusion has an important impact on the biotechnological application of those enzymes.

Abbreviations

AcmB: 3-Ketosteroid Δ^1-dehydrogenase 1 from *Sterolibacterium denitrificans*; AcmB2: 3-Ketosteroid Δ^1-dehydrogenase 2 from *Sterolibacterium denitrificans*; ACN: Acetonitrile; AD: Androst-4-en-3,17-dione; AMBER: Assisted model building with energy refinement; ANB: 1,4-Androstadien-3,71-dione; BME: β-Mercaptoethanol; DAD: Diode array detector; DCPIP: 2,6-Dichloroindophenol; dCTP: Deoxycytidine triphosphate; dGTP: Deoxyguanosine triphosphate; DMSO: Dimethyl sulfoxide; DTT: Dithiothreitol; EGME: 2-Methoxyethanol; ESP: Electrostatic Potential; FAD: Flavin adenine dinucleotide; HBC: 2-Hydroxypropyl-β-cyclodextrin; HPLC: High performance liquid chromatography; IPTG: Isopropyl β-D-1-thiogalactopyranoside; kbp: Kilobase pare; k_{cat}: Catalytic constant/turnover number; K_m: Michaelis constant; K_s: HBC/steroid complex stability constant; KSTD(s): 3-Ketosteroid Δ^1-dehydrogenase(s); KSTD1: 3-Ketosteroid Δ^1-dehydrogenase from *R. erythropolis* SQ1; LB: Lennox Broth; MD: Molecular dynamics; MM-PBSA: Molecular mechanics energies combined with the Poisson–Boltzmann and surface area continuum solvation; MS: Mass spectrometry; NPT: Isothermal-isobaric ensemble; NVT: Canonical ensemble; PDB: Protein Data Bank; PMSF: Phenylmethylsulfonyl fluoride; RESP: Restrained Electrostatic Potential; V_{max}: Maximal reaction velocity.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12934-021-01611-5.

Additional file 1. Fig. S1. Root Mean Square Deviation and selected distances during MD simulation for androst-4-en-3,17-dione. Fragment of trajectory selected for MMPBSA calculations is marked with red rectangles. **Fig. S2.** Root Mean Square Deviation and selected distances during MD simulation for cholest-4-en-3-one. Fragment of trajectory selected for MMPBSA calculations is marked with red rectangles. **Fig. S3.** Root Mean Square Deviation and selected distances during MD simulation for diosgenone. Fragment of trajectory selected for MMPBSA calculations is marked with red rectangles. **Fig. S4.** Root Mean Square Deviation and selected distances during MD simulation for androstanolone. Fragment of trajectory selected for MMPBSA calculations is marked with red rectangles. **Fig. S5.** Root Mean Square Deviation (and selected distances during MD simulation for progesterone. Fragment of trajectory selected for MMPBSA calculations is marked with red rectangles. **Fig. S6.** Root Mean Square Deviation (and selected distances during MD simulation for testosterone propionate. Fragment of trajectory selected for MMPBSA calculations is marked with red rectangles. **Fig. S7.** Root Mean Square Deviation and selected distances during MD simulation for 6-dehydrotestosterone acetate. Fragment of trajectory selected for MMPBSA calculations is marked with red rectangles. **Table S1.** Fit parameters and statistics of

---

phase solubility diagrams of steroids in the solution of HBC and 2% EGME. **Table S2.** Percentage of cholest-4-en-3-one forms dependent on initial substrate concentration; S: cholest-4-en-3-one, HBC: 2-hydroxypropyl-β-cyclodextrin. **Table S3.** Fit statistics of the AcmB Ping-Pong bi bi (non-sequential) mechanism. **Fig. S8.** Progress of the 1,2-dehydrogenation of 0.1 mM AD **A** and cholest-4-en-3-one **B** with 6.55 nM AcmB2 and 0.1 mM diosgenone **C** with 65.5 nM AcmB2 in the presence of 0.15 mM DCPIP in 50 mM Tris-HCl buffer pH 8.0. S–substrate; P–product. **Fig. S9.** Results of steady-state kinetics for dehydrogenation reaction of androst-4-en-3.17-dione **A** and cholest-4-en-3-one **B** calatyzed by AcmB. Reaction velocities were measured in 0.1 M K_2HPO_4/KH_2PO_4 buffer pH 6.5 with 2% HBC. 2% EGME. 0.052–0.174 mM DCPIP. 5–80 µM steroids and 0.30 µM of AcmB at 30 °C. **Table S4.** Fit parameters and statistics of AcmB and KSTD1 steady-state kinetics results.

Authors' contributions

MSz and PW designed research; MG conducted theoretical modeling; PW and AR conducted research on KSTD1; PW conducted research on AcmB; AW conducted research on AcmB2; PW, MG, AW and MSz analyzed data; PW, MG and A.W. wrote the methods; PW and MSz wrote the introduction, results and discussion. MSz and AW funded the research. All authors read and approved the final manuscript.

Availability of data and materials

All data generated and analyzed during this study are included in this published article and its Additional file 1. Additional raw data (MD trajectories, raw kinetic data) are available on request.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1] Jerzy Haber Institute of Catalysis and Surface Chemistry Polish Academy of Sciences, Niezapominajek 8, 30239 Krakow, Poland. [2] Department of Chemistry, Faculty of Science and Technology, Universitas Airlangga, Surabaya 60115, Indonesia. [3] Laboratory of Proteomics, Research Center for Bio-Molecule Engineering (BIOME), Universitas Airlangga, Surabaya 60115, Indonesia. [4] Laboratory of Biophysical Chemistry, University of Groningen, 9747 AG Groningen, The Netherlands.

Wójcik *et al. Microb Cell Fact*   (2021) 20:119

Page 11 of 12

References

1. Urich K. Sterols and steroids. Comp Anim Biochem. 1994. https://doi.org/10.1007/978-3-662-06303-3_16.
2. Rohman A, Dijkstra BW. The role and mechanism of microbial 3-ketosteroid Delta(1)-dehydrogenases in steroid breakdown. J Steroid Biochem Mol Biol. 2019;191:105366.
3. García JL, Uhía I, Galán B. Catabolism and biotechnological applications of cholesterol degrading bacteria. Microb Biotechnol. 2012;5:679–99.
4. Wang PH, Lee TH, Ismail W, Tsai CY, Lin CW, Tsai YW, et al. An oxygenase-independent cholesterol catabolic pathway operates under oxic conditions. PLoS ONE. 2013;8:e66675.
5. Kreit J. Aerobic catabolism of sterols by microorganisms: key enzymes that open the 3-ketosteroid nucleus. FEMS Microbiol Lett Oxford University Press. 2019;366:1–13.
6. Ringold HJ, Hayano M, Stefanovic V. Concerning the stereochemistry and bacterial of the of steroids. J Biol Chem. 1963. https://doi.org/10.1016/S0021-9258(18)67926-8.
7. Zhang R, Liu X, Wang Y, Han Y, Sun J, Shi J, et al. Identification, function, and application of 3-ketosteroid Δ1-dehydrogenase isozymes in *Mycobacterium neoaurum* DSM 1381 for the production of steroidic synthons. Microb Cell Fact BioMed Central. 2018;17:1–16.
8. Wojtkiewicz AM, Wójcik P, Procner M, Flejszar M, Oszajca M, Hochołowski M, et al. The efficient Δ1-dehydrogenation of a wide spectrum of 3-ketosteroids in a broad pH range by 3-ketosteroid dehydrogenase from Sterolibacterium denitrificans. J Steroid Biochem Mol Biol. 2020;202:105731.
9. Kondo E. Steroid 1-dehydrogenation by a crude enzyme preparation from Arthrobacter simplex. Agric Biol Chem. 1963;27:69–70.
10. Li J, Guan Y, Wang H, Yao S. 17-epoxyprogesterone by encapsulated Arthrobacter simplex cells in an aqueous/organic solvent two-liquid-phase system. J Chem Technol Biotechnol. 2009;84:208–14.
11. Liu X, Zhang R, Bao Z, Yuan C, Cao H, Shi J, et al. Biotransformation of phytosterols to androst-1,4-diene-3,17-dione by Mycobacterium sp. ZFZ expressing 3-ketosteroid-δ1-dehydrogenase. Catalysts. 2020;10:1–12.
12. Wang ZF, Huang YL, Rathman JF, Yang S. Lecithin-enhanced biotransformation of cholesterol to androsta-1,4-diene-3,17-dione and androsta-4-ene-3,17-dione. J Chem Technol Biotechnol. 2002;77:1349–57.
13. Santos RA, Caldeira JCO. Steroid bioconversion in a novel aqueous two-phase system. Biotechnol Lett. 1991;13:349–54.
14. Flygare S, Larsson P. Steroid transformation in aqueous two-phase systems: side-chain degradation of cholesterol by Mycobacterium sp. Enzyme Microb Technol. 1989;11:752–9.
15. Rohman A, Dijkstra BW. Application of microbial 3-ketosteroid Δ1-dehydrogenases in biotechnology. Biotechnol Adv. 2021;49:107751.
16. Hesselink PGM, Van Vliet S, De Vries H, Witholt B. Optimization of steroid side chain deavage by Mycobacterium sp. in the presence of cydodextrins. Enzyme Microb Technol. 1989;11:398–404.
17. Manosroi A, Saowakhon S, Manosroi J. Enhancement of androstadienedione production from progesterone by biotransformation using the hydroxypropyl-beta-cyclodextrin complexation technique. J Steroid Biochem Mol Biol. 2008;108:132–6.
18. Penasse L, Peyre M. Studies of 3-oxo steroid delta-1-oxydo reductase of Arthrobacter simplex. Steroids. 1968;12:525–44.
19. Aries VC, Goddard P, Hill MJ. Degradation of steroids by intestinal bacteria. III. 3-Oxo-5β-steroid Δ1-dehydrogenase and 3-oxo-5β-steroid Δ4-dehydrogenase. Biochim Biophys Acta. 1971;248:482–8.
20. Zhang Q, Ren Y, He J, Cheng S. Multiplicity of 3-ketosteroid Δ1-dehydrogenase enzymes in Gordonia neofelifaecis NRRL B-59395 with preferences for different steroids. Ann Microbiol. 2015;65:1961–71.
21. Chiang YR, Ismail W, Gallien S, Heintz D, Van Dorsselaer A, Fuchs G. Cholest-4-en-3-one-Δ1-dehydrogenase, a flavoprotein catalyzing the second step in anoxic cholesterol metabolism. Appl Environ Microbiol. 2008;74:107–13.
22. Wang X, Feng J, Zhang D, Wu Q, Zhu D, Ma Y. Characterization of new recombinant 3-ketosteroid-Δ1-dehydrogenases for the biotransformation of steroids. Appl Microbiol Biotechnol. 2017;101:6049–60.
23. Mao S, Wang JW, Liu F, Zhu Z, Gao D, Guo Q, et al. Engineering of 3-ketosteroid Δ1-dehydrogenase based site - directed saturation mutagenesis for efficient biotransformation of steroidal substrates. Microb Cell Fact BioMed Central. 2018;17:1–13.
24. Sludge D, Wei ST, Wu Y, Lee T, Huang Y, Yang C, et al. Microbial functional responses to cholesterol catabolism in responses to cholesterol catabolism in. MSystems. 2018;3:1–19.
25. Chiang YR, Ismail W, Müller M, Fuchs G. Initial steps in the anoxic metabolism of cholesterol by the denitrifying *Sterolibacterium denitrificans*. J Biol Chem. 2007;282:13240–9.
26. Chiang Y, Ismail W, Heintz D, Schaeffer C, Van DA, Fuchs G. Study of anoxic and oxic cholesterol metabolism by *Sterolibacterium denitrificans*. J Bacteriol. 2008;190:905–14.
27. Lin CW, Wang PH, Ismail W, Tsai YW, El NA, Yang CY, et al. Substrate uptake and subcellular compartmentation of anoxic cholesterol catabolism in *Sterolibacterium denitrificans*. J Biol Chem. 2015;290:1155–69.
28. Chiang Y, Ismail W. Anaerobic biodegradation of steroids. In: Boll M, editor. Anaerobic utilization of hydrocarbons, oils, and lipids. Cham: Springer; 2017. p. 1–32.
29. Knol J, Bodewits K, Hessels GI, Dijkhuizen L, Van Der Geize R. 3-Keto-5α-steroid Δ1-dehydrogenase from *Rhodococcus erythropolis* SQ1 and its orthologue in *Mycobacterium tuberculosis* H37Rv are highly specific enzymes that function in cholesterol catabolism. Biochem J. 2008;410:339–46.
30. Rohman A, Van Oosterwijk N, Dijkstra BW. Purification, crystallization and preliminary X-ray crystallographic analysis of 3-ketosteroid Δ1-dehydrogenase from Rhodococcus erythropolis SQ1. Acta Crystallogr Sect F Struct Biol Cryst Commun. 2012;68:551–6.
31. Ma YH, Wang M, Fan Z, Shen YB, Zhang LT. The influence of host-guest inclusion complex formation on the biotransformation of cortisone acetate Δ1-dehydrogenation. J Steroid Biochem Mol Biol. 2009;117:146–51.
32. Martin del Valle EM. Cyclodextrins and their uses: a review. Process Biochem. 2004;39:1033–46.
33. Loftsson T, Magnúsdóttir A, Másson M, Sigurjónsdóttir JF. Self-association and cyclodextrin solubilization of drugs. J Pharm Sci. 2002;91:2307–16.
34. Rohman A, Van Oosterwijk N, Thunnissen AMWH, Dijkstra BW. Crystal structure and site-directed mutagenesis of 3-ketosteroid δ1-dehydrogenase from *Rhodococcus erythropolis* SQ1 explain its catalytic mechanism. J Biol Chem. 2013;288:35559–68.
35. Kabsch W. A discussion of the solution for the best rotation to relate two sets of vectors. Acta Crystallogr Sect A. 1978;34:827–8.
36. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, et al. Gaussian 16, Revision B.01. Wallingford, CT: Gaussian, Inc.; 2016.
37. Søndergaard CR, Olsson MHM, Rostkowski M, Jensen JH. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKa values. J Chem Theory Comput. 2011;7:2284–95.
38. Olsson MHM, SØndergaard CR, Rostkowski M, Jensen JH. PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. J Chem Theory Comput. 2011;7:525–37.
39. Anandakrishnan R, Aguilar B, Onufriev AV. H++ 3.0: Automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. Nucleic Acids Res. 2012. https://doi.org/10.1093/nar/gks375.
40. Salomon-Ferrer R, Case DA, Walker RC. An overview of the Amber biomolecular simulation package. WIREs Comput Mol Sci. 2012;3:198–210.
41. Dupradeau FY, Cézard C, Lelong R, Stanislawiak É, Pêcher J, Delepine JC, et al. R.E.DD.B.: a database for RESP and ESP atomic charges, and force field libraries. Nucleic Acids Res. 2008;36:360–7.
42. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys. 1983;79:926–35.
43. Salomon-Ferrer R, Case DA, Walker RC. An overview of the Amber biomolecular simulation package. Wiley Interdiscip Rev Comput Mol Sci. 2013;3:198–210.
44. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, et al. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem. 2003;24:1999–2012.
45. Miller BR, Mcgee TD, Swails JM, Homeyer N, Gohlke H, Roitberg AE. MMPBSA py : an efficient program for end-state free energy calculations. J Chem Theory Comput. 2012;8:3314–21.
46. Itagaki E, Wakabayashi T, Hatta T. Purification and characterization of 3-ketosteroid-delta 1-dehydrogenase from Nocardia corallina. Biochim Biophys Acta Rev Cancer. 1990;1038:60–7.

Wójcik *et al. Microb Cell Fact*   (2021) 20:119

Page 12 of 12

47. Sofińska K, Wojtkiewicz AM, Wójcik P, Zastawny O, Guzik M, Winiarska A, et al. Investigation of quaternary structure of aggregating 3-ketosteroid dehydrogenase from *Sterolibacterium denitrificans*: In the pursuit of consensus of various biophysical techniques. Biochim Biophys Acta Gen Subj. 2019. https://doi.org/10.1016/j.bbagen.2019.03.009.

48. Zhang H, Tian Y, Wang J, Li Y, Wang H, Mao S, et al. Construction of engineered Arthrobacter simplex with improved performance for cortisone acetate biotransformation. Appl Microbiol Biotechnol. 2013;97:9503–14.

49. D'Souza VT, Lipkowitz KB. Cyclodextrins : introduction. Chem Rev. 1998;98:1741–2.

50. Williams RO III, Mahaguna V, Sriwongjanya M. Characterization of an inclusion complex of cholesterol and hydroxypropyl-b-cyclodextrin. Eur J Pharm Biopharm. 1998;46:355–60.

51. Luo J, Cui H, Jia H, Li F, Cheng H, Shen Y, et al. Identification, biological characteristics, and active site residues of 3-ketosteroid δ1-dehydrogenase homologues from arthrobacter simplex. J Agric Food Chem. 2020;69:9496–512.

52. Sofińska K, Wojtkiewicz AM, Wójcik P, Zastawny O, Guzik M, Winiarska A, et al. Investigation of quaternary structure of aggregating 3-ketosteroid dehydrogenase from *Sterolibacterium denitrificans*: In the pursuit of consensus of various biophysical techniques. Biochim Biophys Acta - Gen Subj. 2019;1863:1027–39.

53. Van Der Geize R, Hessels GI, Van Gerwen R, Vrijbloed JW, Van Der Meijden P, Dijkhuizen L. Targeted disruption of the kstD gene encoding a 3-ketosteroid δ1- dehydrogenase isoenzyme of *Rhodococcus erythropolis* strain SQ1. Appl Environ Microbiol. 2000;66:2029–36.

54. van der Geize R, Hessels GI, Dijkhuizen L. Molecular and functional characterization of the kstD2 gene of Rhodococcus *erythropolis* SQ1 encoding a second 3-ketosteroid Δ1-dehydrogenase isoenzyme. Microbiology. 2002;148:3285–92.

Publisher's Note

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

Supporting Information

Universal capability of 3-ketosteroid Δ^1-dehydrogenases to catalyze Δ^1-dehydrogenation of C17- substituted steroids

Patrycja Wójcik[1], Michał Glanowski[1], Agnieszka M. Wojtkiewicz[1], Ali Rohman[2,3,4], Maciej Szaleniec[1]*

*Correspondence: maciej.szaleniec@ikifp.edu.pl

[1] Jerzy Haber Institute of Catalysis and Surface Chemistry Polish Academy of Sciences, Niezapominajek 8, PL30239 Krakow, Poland
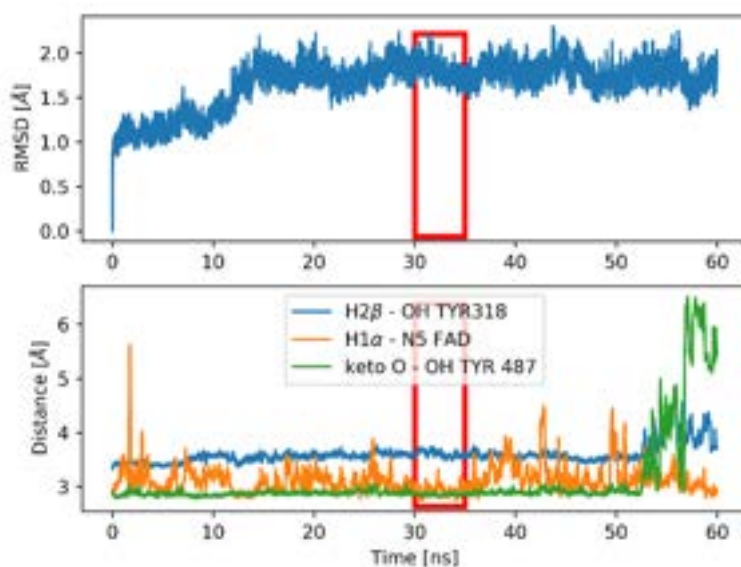
[2] Department of Chemistry, Faculty of Science and Technology, Universitas Airlangga, Surabaya 60115, Indonesia

[3] Laboratory of Proteomics, Research Center for Bio-Molecule Engineering (BIOME), Universitas Airlangga, Surabaya 60115, Indonesia

[4] Laboratory of Biophysical Chemistry, University of Groningen, 9747 AG Groningen, The Netherlands

Results
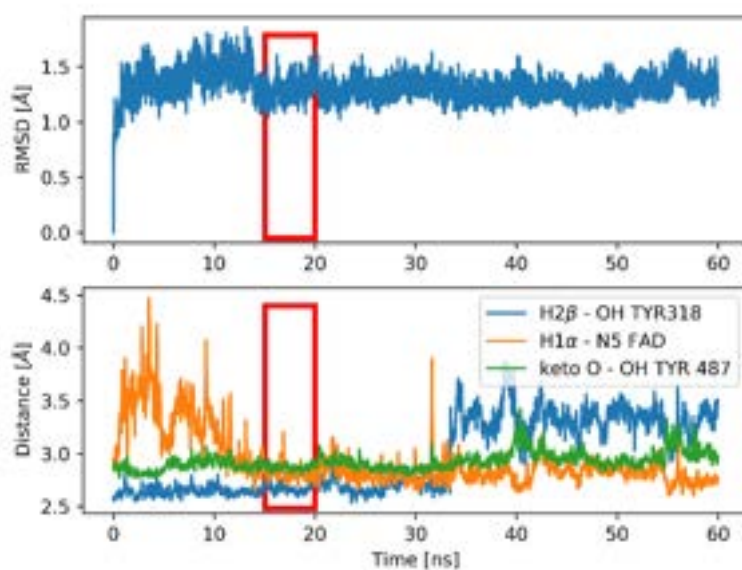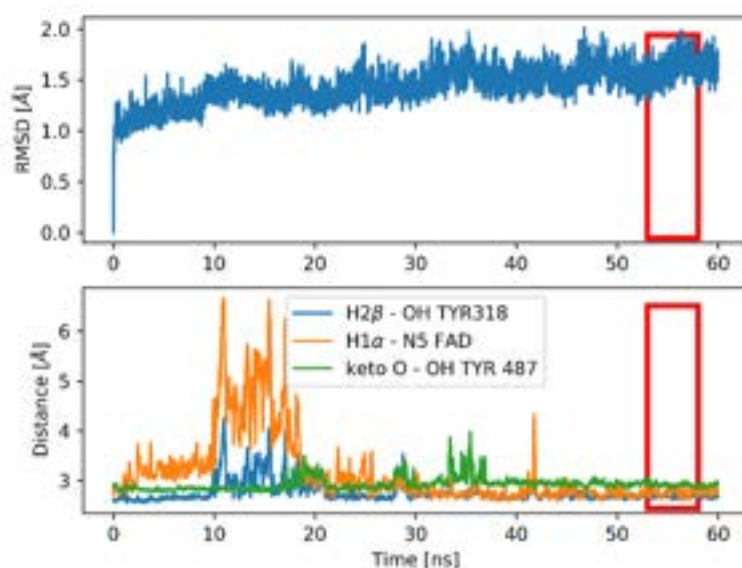
Theoretical prediction of substrate specificity

Fig. S1 Root Mean Square Deviation (top graph) and selected distances (bottom graph, smoothed by moving average) during MD simulation for androst-4-en-3,17-dione. Fragment of trajectory selected for MMPBSA calculations is marked with red rectangles.

Fig. S2 Root Mean Square Deviation (top graph) and selected distances (bottom graph, smoothed by moving average) during MD simulation for cholest-4-en-3-one. Fragment of trajectory selected for MMPBSA calculations is marked with red rectangles.

Fig. S3 Root Mean Square Deviation (top graph) and selected distances (bottom graph, smoothed by moving average) during MD simulation for diosgenone. Fragment of trajectory selected for MMPBSA calculations is marked with red rectangles.

Fig. S4 Root Mean Square Deviation (top graph) and selected distances (bottom graph, smoothed by moving average) during MD simulation for androstanolone. Fragment of trajectory selected for MMPBSA calculations is marked with red rectangles.

Fig. S5 Root Mean Square Deviation (top graph) and selected distances (bottom graph, smoothed by moving average) during MD simulation for progesterone. Fragment of trajectory selected for MMPBSA calculations is marked with red rectangles.

Fig. S6 Root Mean Square Deviation (top graph) and selected distances (bottom graph, smoothed by moving average) during MD simulation for testosterone propionate. Fragment of trajectory selected for MMPBSA calculations is marked with red rectangles.

Fig. S7 Root Mean Square Deviation (top graph) and selected distances (bottom graph, smoothed by moving average) during MD simulation for 6-dehydrotestosterone acetate. Fragment of trajectory selected for MMPBSA calculations is marked with red rectangles.

HBC/steroid inclusion complex formation

Table S1 Fit parameters and statistics of phase solubility diagrams of steroids in the solution

of HBC and 2% EGME at 30°C

Parameter	Substrate	
	Androst-4-en-3,17-dione (**1**)	Cholest-4-en-3-one (**2**)
Equation	$y = ax + b$	$y = ax^2 + bx + c$
a	0.52661 ± 0.00859	0.00035 ± 0.00001
b	0.14420 ± 0.00714*	0.00611 ± 0.00081
c	-	0.00478 ± 0.00054*
R^2	0.99869	0.9989
Adj. R^2	0.99843	0.99835

* measured values

Table S2 Percentage of cholest-4-en-3-one forms dependent on initial substrate

concentration; S - cholest-4-en-3-one, HBC – 2-hydroxypropyl-β-cyclodextrin

$[S]_t$ [μM]	% [S]	% [S(HBC)]	% [S(HBC)$_2$]
5	2.87	53.21	43.93

10	2.87	53.22	43.91
20	2.87	53.24	43.89
30	2.88	53.26	43.86
40	2.88	53.28	43.84
50	2.88	53.30	43.81
60	2.89	53.32	43.79
70	2.89	53.35	43.76
80	2.90	53.37	43.74
100	2.90	53.41	43.69

AcmBs native substrate

Table S3 Fit statistics of the AcmB Ping-Pong bi bi (non-sequential) mechanism

	Substrate	
Parameter	Androst-4-en-3.17-dione (**1**)	Cholest-4-en-3-one (**2**)
Reduced χ^2	32.77842	11.68547
R^2	0.99318	0.9749
Adj. R^2	0.99309	0.97464
AICc	546.15167	484.35427
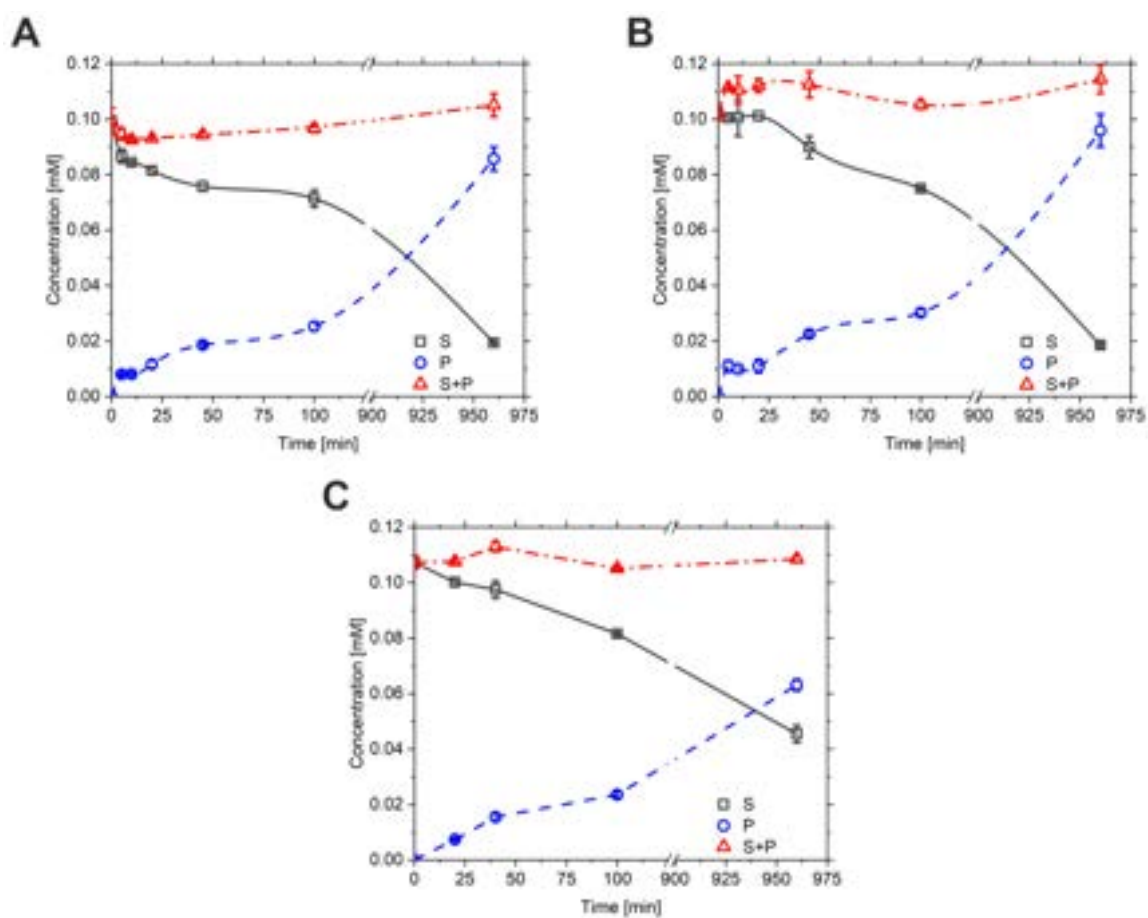
Fig. S8 Progress of the 1.2-dehydrogenation of 0.1 mM AD (**A**) and cholest-4-en-3-one (**B**) with 6.55 nM AcmB2 and 0.1 mM diosgenone (**C**) with 65.5 nM AcmB2 in the presence of 0.15 mM DCPIP in 50 mM Thris-HCl buffer pH 8.0. S – substrate; P – product.

KSTD1 substrate spectrum
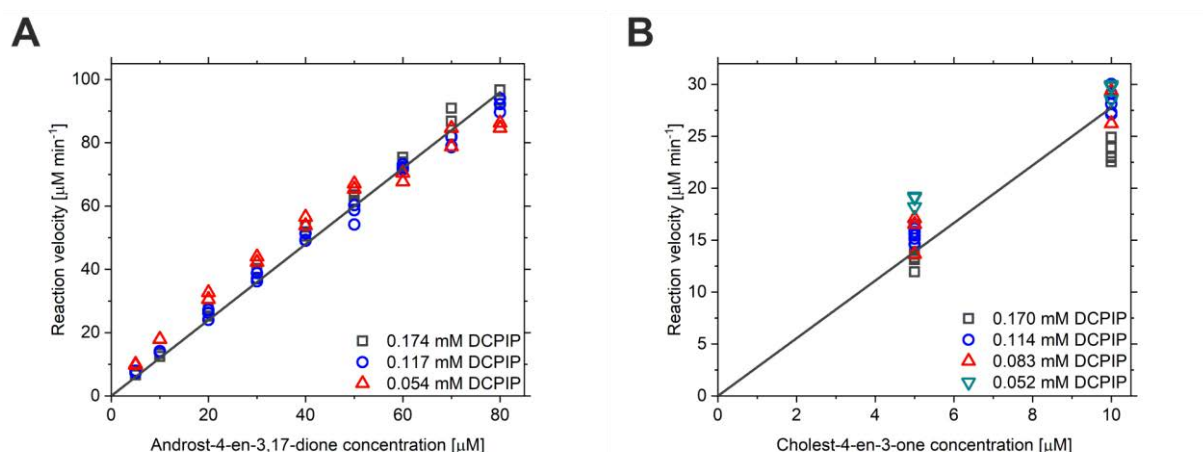
Fig. S9 Results of steady-state kinetics for dehydrogenation reaction of androst-4-en-3.17-dione (**A**) and cholest-4-en-3-one (**B**) calatyzed by AcmB. Reaction velocities were measured in 0.1 M K_2HPO_4/KH_2PO_4 buffer pH 6.5 with 2% HBC. 2% EGME. 0.052 – 0.174 mM DCPIP. 5 – 80 µM steroids and 0.30 µM of AcmB at 30°C.

Table S4 Fit parameters and statistics of AcmB and KSTD1 steady-state kinetics results

<table>
<tr><td colspan="2"></td><td colspan="2" align="center">Substrate</td></tr>
<tr><td></td><td>Parameter</td><td align="center">Androst-4-en-3.17-dione (1)</td><td align="center">Cholest-4-en-3-one (2)</td></tr>
<tr><td rowspan="5">KSTD1</td><td>Equation</td><td align="center">y = ax + b</td><td align="center">y = ax + b</td></tr>
<tr><td>a</td><td align="center">1.18924 ± 0.00672</td><td align="center">0.10844 ± 0.000668728</td></tr>
<tr><td>b</td><td align="center">0</td><td align="center">0</td></tr>
<tr><td>R^2</td><td align="center">0.99755</td><td align="center">0.99704</td></tr>
<tr><td>Adj. R^2</td><td align="center">0.99751</td><td align="center">0.997</td></tr>
<tr><td rowspan="5">AcmB</td><td>Equation</td><td align="center">y = ax + b</td><td align="center">y = ax + b</td></tr>
<tr><td>a</td><td align="center">1.19954 ± 0.0107</td><td align="center">2.77637 ± 0.05598</td></tr>
<tr><td>b</td><td align="center">0</td><td align="center">0</td></tr>
<tr><td>R^2</td><td align="center">0.99368</td><td align="center">0.98597</td></tr>
<tr><td>Adj. R^2</td><td align="center">0.9936</td><td align="center">0.98557</td></tr>
</table>

Publikacja P3

Enzymatic Δ¹-Dehydrogenation of 3-Ketosteroids—Reconciliation of Kinetic Isotope Effects with the Reaction Mechanism

Michał Glanowski, Patrycja Wójcik, Magdalena Procner, Tomasz Borowski, Dawid Lupa, Przemysław Mielczarek, Maria Oszajca, Katarzyna Świderek, Vicent Moliner, Andrzej J. Bojarski, and Maciej Szaleniec*

ACCESS | 📊 Metrics & More | 📖 Article Recommendations | ⓢ Supporting Information

ABSTRACT: Δ¹-Dehydrogenation of 3-ketosteroids catalyzed by flavin adenine dinucleotide (FAD)-dependent 3-ketosteroid dehydrogenases (Δ¹-KSTD) is a crucial step in steroid degradation and synthesis of several steroid drugs. The catalytic mechanism assumes the formation of a double bond in two steps, proton abstraction by tyrosyl ion, and a rate-limiting hydride transfer to FAD. This hypothesis was never verified by quantum-mechanical studies despite contradictory results from the kinetic isotope effect (KIE) reported in 1960 by Jerussi and Ringold [*Biochemistry* **1965**, 4 (10)]. In this paper, we present results that reconcile the mechanistic hypothesis with experimental evidence. Quantum mechanics/molecular mechanics molecular dynamics simulations show that the proposed mechanism is indeed the most probable, but barriers associated with substrate activation (13.4−16.3 kcal·mol⁻¹) and hydride transfer (15.5−18.0 kcal·mol⁻¹) are very close (1.7−2.1 kcal·mol⁻¹), which explains normal KIE values for steroids labeled either at C1 or C2 atoms. We confirm that tyrosyl ion acting as the catalytic base is indeed necessary for efficient activation of the steroid. We explain the lower value of the observed KIE (1.5−3.5) by the nature of the free energy surface, the presence of diffusion limitation, and to a smaller extent, conformational changes of the enzyme upon substrate binding. Finally, we confirm the Ping-Pong bi−bi kinetics of the whole Δ¹-dehydrogenation and demonstrate that substrate binding, steroid dehydrogenation, and enzyme reoxidation proceed at comparable rates.

KEYWORDS: Δ¹-ketosteroid dehydrogenase, 3-ketosteroids, Δ¹-dehydrogenation, kinetic isotope effect, QM/MM, kinetic solvent viscosity effect, Ping-Pong bi−bi mechanism

INTRODUCTION

Steroids are one of the most important groups of drugs on the market, finding their application in treating inflammation and diseases of immune function, such as allergies, asthma, autoimmune diseases, and several mineral metabolism disorders like hyponatremia, hyperkalemia, osteoporosis, and hypotension as well as in birth control.[1] Glucocorticoids (e.g., dexamethasone) were used in clinical practice and are evaluated as agents for the treatment of acute respiratory distress syndrome, which develops in many patients during infection with SARS-CoV-2.[2] As a result, the value of the global steroid market is expected to reach 17 Bn USD by the end of 2025 with extensive growth predicted in the segment of corticosteroids and anabolic steroids (according to QY Research, Inc.).[3] Steroid drugs are synthesized by a combination of chemical and microbiological methods,[4] where the latter utilizes a natural metabolism of cholesterol or phytosterols in fungi or bacteria. Microbial fermentation very often provides steroid active pharmaceutical ingredients (APIs) that can be further functionalized by chemical or biocatalytic means. One of such biocatalytic processes that has been extensively studied during the last 50 years is Δ¹-dehydrogenation of steroids catalyzed by 56 kDa Δ¹-ketosteroid dehydrogenase (Δ¹-KSTD). Δ¹-KSTDs are essential in the production of several steroid drugs, such as betamethasone,[5,6] boldenone,[7,8] prednisone,[5,9] dexamethasone,[10,11] and steroid APIs [4-androstene-3,17-dione (AD), 1,4-androstadiene-3,17-dione (ADD), and 9α-hydroxy-4-androstene-3,17-dione].[12−14] Due to their high importance in steroid metabolism, and consequently for the pharmaceutical industry, the flavin adenine dinucleotide (FAD)-dependent Δ¹-KSTDs are one of the best-studied steroid-degrading
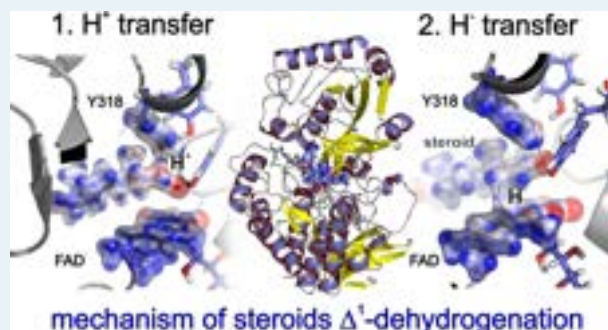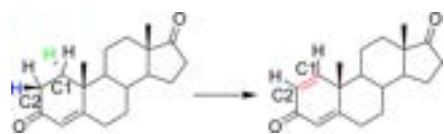
enzymes.[14] They catalyze the regioselective introduction of a double bond between C1 and C2 atoms in ring A of the steroid core (Figure 1).

Figure 1. Dehydrogenation of AD catalyzed by Δ^1-KSTD. Abstracted hydrogen atoms, as marked in blue (H2β) and green (H1α), formed a double bond, in red.

The mechanistic hypothesis describing the catalytic mechanism has been formulated based on (i) the structure of only one KSTD from *Rhodococcus erythropolis* (PDB 4c3x and 4c3y),[15,16] (ii) activity tests of genetically modified enzyme variants, and (iii) numerous kinetic studies that yielded values of the kinetic isotope effect (KIE) for chirally labeled substrates.[15,17−21] The currently accepted hypothesis assumes that the overall reaction proceeds according to the Ping-Pong bi−bi mechanism, with the steroid substrate oxidized during reductive half-reaction (RHR) and FAD reoxidation by the electron acceptor during oxidative half-reaction (OHR).[17] RHR proceeds in a step-wise manner, that is, the steroid is activated by the tyrosyl anion, which abstracts the acidic 2β-H proton from the C2 atom with a high degree of stereoselectivity.[20] In the second, supposedly rate-limiting step, 1α hydride is shifted to FAD resulting in the formation of the double bond (Figure 1).[14,22] Surprisingly, this hypothesis was never validated with quantum chemical calculations despite several intriguing results which did not entirely agree with the undoubtedly sensible mechanistic hypothesis. Over 5 decades ago, Jerussi and Ringold provided values of KIEs obtained for steroids selectively labeled with deuterium at C1 and/or C2 positions. The authors observed KIEs for substrates labeled either at C1 or C2 position and the values of $V_{\mathrm{H}}/V_{\mathrm{D}}$ were in the range of 1.2−2.5, which is relatively low for the primary H/D KIE, especially compared to strictly chemical reactions.[21] More importantly, if the second step of RHR indeed was rate-limiting, why was the substitution of 2β protium with deuterium influencing the observed KIE? Another unresolved question concerns the protonation state of Tyr318. As most of the KSTDs exhibit optimum pH under neutral or basic conditions,[17,23,24] it was sensible to assume that it is a tyrosyl ion that deprotonates acidic C2 atom of the steroid. However, several KSTDs have been recently discovered that exhibit optimum activity under slightly acidic conditions, despite highly conserved composition of the active site.[25−27] Therefore, the question is whether the deprotonation of tyrosine is truly necessary for activation of the substrate. We decided to apply advanced multiscale quantum mechanics/molecular mechanics (QM/MM) methods to finally test the mechanistic hypotheses on the reaction mechanism and to try reconciling it with the results of experiments. In this paper, we present a detailed investigation of Δ^1-dehydrogenation of 17-methyltestosterone (17β-hydroxy-17-methyl-4-androsten-3-one, 17-MT) to metandienone (17β-hydroxy-17-methyl-1,4-androst-dien-3-one, MTD), and dihydrotestosterone (17β-hydroxy-5α-androstan-3-one, DHT) to 1-testosterone (17β-hydroxy-5α-androst-1-en-3-one, 1-TE) by Δ^1-KSTD from *R. erythropolis*. We compared possible reaction pathways and calculated values of the intrinsic KIE for the RHR. These results were

confronted with kinetic experiments providing proof for the Ping-Pong bi−bi mechanism, including observed KIEs from the competitive, pre- and steady-state kinetic as well as kinetic solvent viscosity effect (KSVE).

■ METHODOLOGY

Computational Model Setup. The initial atomic coordinates of KSTD from *R. erythropolis* were obtained from Protein Data Bank (PDB ID: 4c3y, resolution 2.3 Å).[15] This crystal structure contains ADD bound in the active site together with the FAD cofactor. ADD was subsequently replaced by 17-MT or by DHT applying the Kabsch method[28] using coordinates of ADD heavy atoms as a template. Protonation states of titratable amino acids were determined using PropKa ver 3.1[29,30] and H++ software[31] at pH = 8, which was reported as optimal for the enzyme activity. According to the predicted pK_a shifts (see the Supporting Information) and after visual inspection of the geometry, it was concluded that Lys450 should be deprotonated, whereas histidine residues number 162, 328, 362, and 408 were protonated at δ positions. The tyrosine in position 318 was considered in a deprotonated state because, according to the mechanistic hypothesis, it serves as a proton acceptor. Missing AMBER parameters for 17-MT, DHT, and tyrosyl anion were obtained at the B3LYP/6-31G(d,p) level of theory in the gas phase with Gaussian16.[32] Point charges were computed using Merz−Kollman electron density calculations[33] and the RESP procedure as implemented in the Antechamber available in AmberTools[34] (see Supporting Information). Parameters for FAD were adapted from RESP ESP charge DDataBase (R.E.DD.B)[35] (see Supporting Information). The systems were neutralized by adding 33 sodium ions in the electrostatically most favorable positions and soaked in a $93.5 \times 76.5 \times 72.0$ Å^3 box with TIP3P[36] water molecules. Models consisting of just the ketosteroid 17-MT and DHT in a $73.5 \times 73.0 \times 76.2$ Å^3 box of TIP3P water molecules were also prepared as reference models in solution.

Classical MD Simulations. The final models (with a total number of 59737 and 59736 atoms for 17-MT and DHT, respectively) were optimized using Amber package[34] with the ff03 force field.[37] After energy minimization and heating from 0 to 303 K over 100 ps with the *NVT* ensemble, the system was equilibrated during 100 ps with *NPT* conditions. Subsequently, *NPT* nonaccelerated and nonbiased molecular dynamic (MD) simulations at 303 K were conducted for 55 ns using periodic boundary conditions. To improve the time of simulations, cutoffs for nonbonding interactions were applied with a value of 8.0 Å. The temperature during the simulations was controlled using the Langevin thermostat.[38] Root mean square deviation (rmsd) indicates that the systems can be considered as equilibrated after 30 ns (see Supporting Information). Conformations during the last 20 ns of simulations were analyzed to obtain representative structures using cluster analysis based on rmsd of positions of heavy atoms of the side chains of Tyr318, Tyr119, Tyr487, and ring A of the 3-ketosteroid and FAD. Geometry closest to the center of the largest cluster was used as the starting structure for QM/MM calculations.[39] For calculations related to the protein-free models, 8 random frames were selected from the last 20 ns of simulation for each ketosteroid.

QM/MM MD Simulations. The QM subsystem consisted of the ketosteroid substrate, the side chain of Tyr318, and a fragment of FAD, as shown in Figure 2, while the rest of the protein and solvent water molecules were represented by
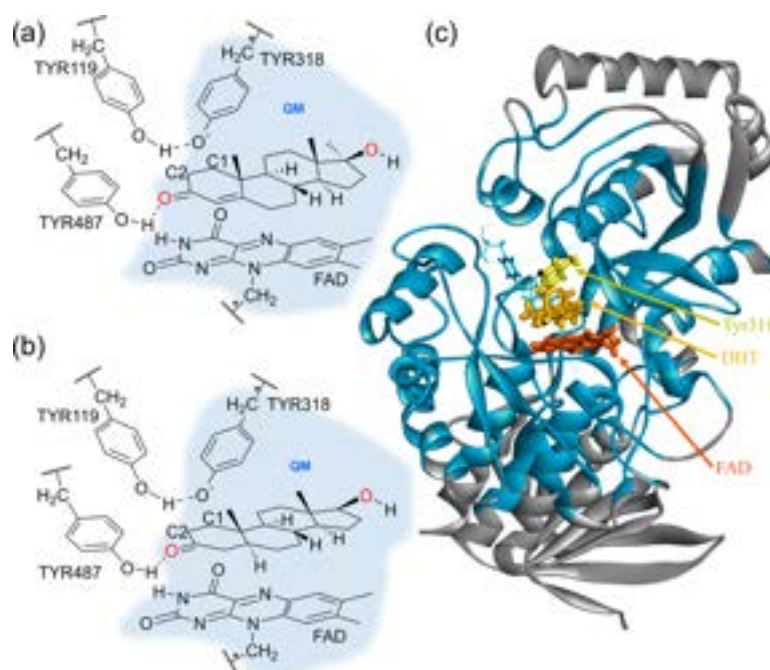
Figure 2. Schematic representation of the active site of the enzyme with substrates: 17-MT (a) and DHT (b). The light blue region contains atoms that are treated quantum mechanically. Quantum link atoms are marked with an asterisk. Overview of the whole model with DHT (c): yellow, orange, and red fragments represent Tyr318, DHT, and FAD, respectively, and flexible part of the model is shown in blue, while the frozen part of the model is in gray. Solvent water molecules are omitted for clarity.

classical AMBER and TIP3P force fields, respectively, as implemented in fDYNAMO library.[40,41] Two hydrogen link atoms[42] were added (Figure 2). Positions of all residues beyond 20 Å from the substrate were fixed. To improve the time of simulations, cutoffs for nonbonding interactions were applied using a smooth switching function between 14.5 and 16 Å. Energies were computed using the standard additive QM/MM scheme. Free energy surfaces (FESs), in terms of potentials of mean force (PMF), were computed for the two chemical steps of the reaction in order to obtain the full free energy profile. The PMFs were computed through the umbrella sampling approach[43] at 303 K.

At first, the potential energy surfaces (PESs) using a selected combination of interatomic distances as a distinguished reaction coordinate were generated. In particular, the antisymmetric combinations of the distances defining the position of the transferred hydrogen, that is, $r_{C2H} - r_{OH}$ for the first step and $r_{C1H} - r_{NH}$ for the second one, were chosen. Then, series of QM/MM MD simulations were carried out in which the distinguished reaction coordinate variable was constrained around the values of the structures generated in the PES with the Umbrella Sampling (US) method,[43] using a parabolic penalty potential with a force constant of 2500 kJ·Å^{-2} mol^{-1}. The values of the variables sampled during the QM/MM simulations were then pieced together to construct a distribution function from which the PMF is obtained using the weighted histogram analysis method.[44] Due to the high computational costs of PMF calculation, the QM Hamiltonian for the QM/MM MD simulations was calculated at the AM1 level of theory. Nevertheless, to reduce the possible errors associated with the semiempirical method, we used an energy spline function defined in terms of interpolated corrections, as described previously.[45] The spline corrections[46] were obtained from single-point calculation at the B3LYP/6-311++G-

(2d,2p)/MM level of theory for structures generated along the AM1/MM PESs. Transition state (TS) structures were then fully optimized and characterized at the AM1/MM and B3LYP/6-31G(d,p)/MM level from structures selected from the quadratic region of the FESs. Each of the structures was optimized with the Baker algorithm[47] using the micro−macro iteration scheme.[45] The gradient norm for optimization was maintained lower than 1 kJ·Å^{-1} mol^{-1} level for the QM region and 0.5 kJ·Å^{-1} mol^{-1} for the remaining movable atoms. Every localized TS structure was verified by the existence of one imaginary frequency as computed using the Hessian matrix containing all the coordinates of the QM subsystem in the presence of the MM environment. Intrinsic reaction coordinates were traced down from located TSs to the connecting valleys in mass-weighted Cartesian coordinates, and the vibration analysis was conducted for the ground states to confirm reaching true minima.

Kinetic and Binding Isotope Effects. Averaged KIEs were calculated for the reaction with deuterium-labeled 17-MT and DHT (Figure 3), and individual contributions of each deuteron (i.e., primary or secondary KIE) are presented in the Supporting Information. From the definition of the free energy of a state and using the transition state theory,[48−50] the total partition function and the zero-point energy of light and heavy
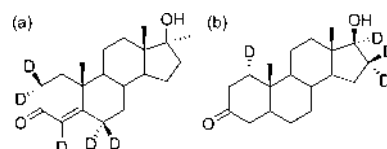
Figure 3. Deuterium-labeled 2,2,4,6,6-d_5-17-MT (a) and 1,16,16,17-d_4-DHT (b) used in the calculations and experiments.

isotopologues were computed from localized E:S, TS1, E:I, and TS2 structures at the AM1/AMBER and B3LYP/AMBER level of theory, as described elsewhere[51−53] (see Supporting Information). At this point, we must keep in mind the nature of the chemical reactions under study, and quantum tunneling effects could have an impact on the quantitative estimation of the computed activation free energies (reducing the effective barrier) and KIEs (increasing the primary deuterium KIE values). Nevertheless, these possible corrections would not influence the obtained trends, as demonstrated previously.[54,55] Analogically, binding isotope effects (BIE)[56] were calculated from a model of the substrate in a water solution and E:S at the B3LYP/AMBER level of theory. For BIE calculation, the QM region was reduced to the steroid molecule.

◼ EXPERIMENTAL METHODOLOGY

Materials. All chemicals were purchased from Sigma-Aldrich (Germany), Tokyo Chemical Industry (Japan), BioShop (Canada), Carl Roth (Germany), or Chempur (Poland) unless otherwise specified. 4-Androsten-17α-methyl-17β-ol-3-one-2,2,4,6,6-d_5 was purchased from CDN Isotopes (Germany), while 17β-hydroxy-5α-androstan-3-one-1,16,16,17-d_4 was obtained from Alsachim (France).

Protein Expression and Purification. A gene encoding 3-ketosteroid Δ^1-dehydrogenase from *R. erythropolis* (KSTD1) was cloned into a pET15b vector[57] and transformed into calcium chloride chemically competent *Escherichia coli (E. coli)* BL21(DE3)Magic cells (Creative Biolabs). An overnight culture of the transformed cells was grown in 2% (w/v) Lennox Broth (LB) supplemented with 100 μg mL^{-1} ampicillin and 50 μg mL^{-1} kanamycin at 37 °C, 180 rpm. The preculture was diluted a hundred times with 1 L of ampicillin and kanamycin supplemented LB with 0.5 M D-sorbitol and grown under the same conditions as the overnight culture until the OD$_{600}$ reached 0.6. Next, the temperature was reduced to 16 °C, and the culture was induced with 100 μM isopropyl β-D-1-thiogalactopyranoside. After 48 h, the *E. coli* cells were harvested by centrifugation at 4500g, 4 °C for 1 h. The cell pellets were resuspended 1:5 (w/v) in the buffer 50 mM Tris-HCl pH 8.5, 100 mM NaCl, 10% (w/v) glycerol, 5 mM β-mercaptoethanol, and 10 mM imidazole and supplemented with 100 μM phenylmethylsulfonyl fluoride. The cell suspension was lysed by sonication (Sonics Vibra-Cell VCX500, 3 s on, 5 s off, 5 min, 40% amplitude, 150 000 J). Cell debris was removed by centrifugation at 40 000g, for 1 h at 4 °C. Purification of KSTD was carried out on a 5 mL HisTrap HP (GE Healthcare) column using an FPLC system (BioRad NGC Quest 10 Plus) and linear imidazole gradient 10−300 mM in the previously described buffer. The fractions that absorbed light at a wavelength of 450 nm were collected, desalted by dialysis, and stored at −20 °C. The concentration of the purified active protein was determined using an extinction coefficient of FAD in KSTD 12 627 M^{-1} cm^{-1} at 450 nm.

To determine the KSTD FAD extinction coefficient (ε_{KSTD}), the protein UV−vis spectrum in the range of 650−300 nm was measured (Shimadzu UV-1280) before and after heat denaturation of the sample at 90 °C for 10 min. The concentration of the free FAD was determined spectrophotometrically using ε_{450nm} of 11 300 M^{-1} cm^{-1}. $\Delta\varepsilon_{KSTD}$ was determined spectrophotometrically as a difference between the spectrum of the KSTD with oxidized FAD and the spectrum of the enzyme with the flavoprotein reduced with the excess

amount of sodium dithionite under anaerobic conditions (98:2 (v/v) N$_2$/H$_2$) divided by the concentration of free FAD. The value of $\Delta\varepsilon_{KSTD}$ at 450 nm equaled 11 368 M^{-1} cm^{-1} and was used to the calculated change in enzyme and substrate concentrations for the global fit of the pre-steady state experiments with 17-MT (see below). The values of ε_{KSTD} for oxidized and reduced KSTD as well as $\Delta\varepsilon_{KSTD}$ in the function of wavelength are provided in the Supporting Information.

Protein MS-Based Identification. 100 μg of KSTD was precipitated with 6-excess of cold acetone and resuspended in 20 μL of 0.5 M triethylammonium bicarbonate buffer with the addition of 2% SDS solution. Then, the protein was reduced by tris-(2-carboxyethyl)-phosphine and alkylated with iodoacetamide. Finally, the protein was digested by 2 μg of Trypsin Gold (from Promega) and desalted on MacroSpin Columns with the C18 resin obtained from The Nest Group, Inc. The obtained peptide maps were evaporated to dryness in a vacuum centrifuge (Labconco). Prior to the nanoLC-MS/MS analysis, the peptides were resuspended in 20 μL of 0.1% formic acid. Nanochromatography combined with tandem mass spectrometry analysis (nanoLC-MS/MS) used to separate protein digests was performed using the Proxeon nanocapillary chromatography system controlled by the Hystar software (Bruker Daltonics). Separations were performed using a PepMap column (15 cm long, 75 μm ID, C18, 3 μm particle size, 100 Å pore size, Thermo-Scientific). The gradient was formed using H$_2$O/0.1% HCOOH (A) and ACN/0.1% HCOOH (B) at a flow rate of 300 nL/min. A gradient was formed from 2 to 35% B at 50 min and up to 90% B at 55 min and then kept until 65 min at 90% B. Fractions eluted from the column were directly deposited with an α-cyano-4-hydroxycinnamic acid matrix on a MALDI target plate by a Proteineer fc II sample collector (Bruker Daltonics). Fifteen-second fractions were collected, 96 fractions for one sample, and spotted on 384 MALDI target plate. The mass spectrometry analyses were performed on ultrafleXtreme (Bruker Daltonics) in positive ion mode.

The acquired mass spectra were analyzed using the Bruker Data Analysis 4.0 software (Bruker Daltonics) and were identified using the Mascot 2.4.1 algorithm (Matrix Science) against the NCBInr protein sequence database ver. 2012. Search parameters were set in the following way: taxonomy: all entries; modification: carbamidomethyl (fixed) or methionine dioxidation (variable); up to 1 missed cleavage; peptide charges: +1; and mass tolerance: 25 ppm for precursor mass and 0.6 Da for fragment mass. Proteins with at least two fragmented, unique peptides detected were considered, and an additional criterion was an ion score higher than 40, which is above the level of false positives ($p \leq 0.05$).

Spectrophotometric Activity Assay. Values of kinetic constants were obtained using a stopped-flow approach. Single wavelength kinetic traces were recorded on the SX20 (Applied Photophysics, UK) stopped-flow spectrophotometer with a 10 mm pathlength and photomultiplier detector. The data were collected and processed using Pro-Data software. All measurements were conducted at 30 °C. The temperature in the stopped-flow experiments was controlled by a Labo Plus (Polyscience, Poland) thermostat bath. Solutions were treated with argon for several minutes to provide anaerobic conditions. All reported concentrations are the final values obtained after mixing and dilution of the reactants. All collected data were

154

processed using OriginPro 2019b software or Octave for global fit.[58]

Steady-State Kinetics. A typical reaction mixture contained 50 mM Tris-HCl buffer pH 8.0, 6.3 nM of KSTD, and varying concentrations of 2,6-dichloroindophenol (DCPIP; 0.025−0.1 mM) and AD (5−120 μM) in isopropanol (the final concentration 1%). In each measurement, one glass syringe of the stopped-flow instrument was filled with the solution of AD and DCPIP and the second one with the enzyme and DCPIP solution. Kinetic traces were followed at 700 nm ($\varepsilon_{700[pH\ 8.0]}$ = 5100 M^{-1} cm^{-1}). The initial rate constants were obtained by fitting a linear function to the first 10 s of the studied reaction. Subsequently, the received data were fitted to the nonsequential Ping-Pong bi−bi model with a nonlinear regression and to an alternate model (e.g., sequential model). The model was selected based on statistical parameters such as R^2, χ^2, AICc, and errors of estimated constants.

Pre-Steady-State Kinetics. The K_D and k_{cat} values for 17-MT were determined in the reaction of 7.1 μM of KSTD with the 17-MT (concentration 25 to 200 μM) using 50 mM Tris-HCl at pH 8.0. In each measurement, one glass syringe of the stopped-flow instrument was filled with the solution of the steroid dissolved in dioxane/buffer solution (the final dioxane concentration 5%), while the second one contained the enzyme solution in the buffer. The reaction was conducted under anaerobic conditions at 30 °C. The kinetics of FAD reduction was followed at 450 nm for 31.25 ms. The obtained traces (8 repetitions for each 17-MT concentration) were globally fitted to a kinetic model (a set of differential equations) consisting of two consecutive reversible steps: formation/dissociation of the ES complex (k_1 and k_{-1}) and FAD reduction/oxidation by the substrate/product (k_2 and k_{-2}): ES ⇌ E'P. To assess the information content of the set of traces and to estimate confidence intervals for fitted kinetics constants, confidence contours on error maps were determined at the level by 10% higher than the minimum determined from unconstrained fitting.[59] Least-squares fitting and confidence contour calculations were performed with an in-house written Octave script,[58,60] which is deposited together with the fitted data at Mendeley Data.[61] A more complicated—three-step model, which included an additional equilibrium (ES ⇌ EI) between substrate binding and the redox step, was also fitted, yet the confidence contours unequivocally showed that thus determined k_1 and k_{-1} are unconstrained by the data and can attain arbitrarily high values. The E + S ⇌ ES ⇌ E'P model was also fitted to the data obtained during the KSVE experiment, but due to lack of enough data for enzyme saturated with the substrate, it was not possible to unequivocally determine k_1 and k_{-1} values.

Kinetic Solvent Viscosity Effect. The KSVE was determined by pre-steady-state kinetics measurements using either glycerol or polyethyleneglycol (PEG) 20000 as viscosigen. The reaction mixture contained 50 mM Tris-HCl buffer pH 8.0, 8.9 μM of KSTD, and varying concentrations of AD (50−200 μM) in 2-methoxyethanol (EGME; the final concentration 5%) and glycerol (at a concentration of 0.27, 1.27, and 2.27 M) or PEG 20000 [at a concentration of 4.2 and 10% (w/v)]. In each measurement, one glass syringe of the stopped-flow instrument was filled with the solution of steroid and the other one with the enzyme solution. The kinetic traces were followed at 450 nm for 35 to 125 ms. The obtained traces were fitted with single exponential functions

$$A = A_0 e^{-\lambda t} + y_0 \tag{1}$$

where λ is the observed decay rate (eigenvalue) of the flavin reduction.

Afterward, the collected data were fitted to the Michaelis−Menten model with nonlinear regression.

The dynamic viscosity of reaction mixtures was determined using the rolling-ball Lovis 2000 M/ME viscometer supplied by Anton Paar at 30 °C. For this purpose, a short capillary tube and a 1.59 mm diameter steel ball were used.

In order to match the hydrodynamic diameter of PEG to that of the KSTD enzyme, the diffusion coefficients of PEG 4000; PEG 8000; PEG 20000 and KSTD were measured in 50 mM Tris-HCl buffer pH 8.0 at 30 °C using the Zetasizer Nano ZS apparatus (Malvern Instruments). Measurements were carried out at a 173° scattering angle using a frequency-doubled diode-pumped solid-state 532 nm laser as an incident light source. Knowing the diffusion coefficient of the molecules, their hydrodynamic diameters (d_H) were calculated using the Stokes−Einstein relationship

$$d_H = \frac{k_B T}{3\pi \eta D} \tag{2}$$

where k_B is the Boltzmann constant, T is the absolute temperature, η is the dynamic viscosity of the medium, and D is the diffusion coefficient of the molecule.

Kinetic Isotope Effect. The KIE at C1 and C2 atoms of the 3-ketosteroid core were determined by a direct method under steady-state or pre-steady-state conditions or in competition experiments for DHT and 17-MT, respectively.

Steady-State Kinetics. The reaction velocities were determined in a spectrophotometric activity assay using a UV-2700 spectrophotometer (Shimadzu) in 0.5 mL quartz cuvettes with a 10 mm path length at 30 °C. The measurements were carried out in 50 mM Tris-HCl buffer pH 8.0 with 200 μM DCPIP, 200 μM steroid dissolved in dioxane [2% (v/v)], and 84.4 nM of KSTD. The reduction of DCPIP was followed at 700 nm ($\varepsilon_{700[pH\ 8.0]}$ = 5100 M^{-1} cm^{-1}). All measurements were performed in triplicates. The initial rate constants were obtained with linear regression fitted to the initial parts of the kinetic curves.

Pre-Steady-State Kinetics. λ_{obs} for C1 was determined in the reaction of 8.9 μM KSTD with either 100 μM DHT or 1,16,16,17-d_4-DHT in EGME/buffer solution (10%). λ_{obs} for the C2 atom was measured with 7.1 μM KSTD and either 250 μM 17-MT or 2,2,4,6,6-d_5-17-MT analogue dissolved in dioxane/buffer solution (6%). Each experiment was repeated at least 12 times. The obtained traces were fitted with single exponential functions, and the ratio of the average $\lambda_{obsH}/\lambda_{obsD}$ was determined.

Competition Experiment. The $^D(V/K)$ kinetic isotope method was measured according to the protocol described previously.[62,63] The determination of the competitive $^D(V/K)$ KIE was based on the fractions of the converted light (x_1) and heavy (x_2) substrate according to the formula

$$^D\left(\frac{V}{K}\right) = \frac{k_1}{k_2} = \frac{\log(1 - x_1)}{\log(1 - x_2)} \tag{3}$$

$^D(V/K)$ was established by a nonlinear fit in Origin 2019 from the reformulated function of $x_1(x_2)$

$$x_1 = 1 - (1 - x_2)^{k_1/k_2} \tag{4}$$

The reactions were carried out in triplicates and the reaction mixtures consisted of 50 mM Tris-HCl buffer pH 8.0, 100 μM DCPIP, equal 100 μM amounts of substrates and their deuterated homologues (17-MT and 2,2,4,6,6-d_5-17-MT or DHT and 1,16,16,17-d_4-DHT) in EGME (the final concentration 2%), and KSTD (0.4 nM for 17-MT or 0.9 nM for DHT). The reactions were run under anaerobic conditions [98:2 (v/v) N_2/H_2] at 30 °C for 18 min. The conversion of each substrate was analyzed with LC-ESI-MS/MS (Agilent 1290 Infinity System equipped with an MS Agilent 6460 Triple Quad Detector) using the positive single-ion monitoring mode (303.3, 308.3, 301.3, and 305.3 m/z signals for $[M + H]^+$ of 17-MT, 2,2,4,6,6-d_5-17-MT, MTD, and 2,4,6,6-d_4-MTD as well as 291.3, 295.3, 289.3, and 292.3 m/z for $[M + H]^+$ of DHT, 1,16,16,17-d_4-DHT, 1-TE, and 16,16,17-d_3-1-TE). The separation was carried out on the Agilent Zorbax Eclipse Plus C18 column (2.1 × 50 mm, 1.8 μm) using the ACN/H$_2$O/HCOOH (40:60:0.1, v/v) mobile phase in the isocratic mode with a flow rate of 0.4 mL/min, 30 °C, and 1 μL injection volume (see Supporting Information for details of the MS method) The quantitation of analytes was conducted using external calibration based on commercial standards or products synthesized by 100% enzymatic conversion of the standards.

■ RESULTS AND DISCUSSION

Kinetic Mechanism. Our kinetic experiments confirmed that Δ^1-dehydrogenation catalyzed by 3-Δ^1-KSTD is indeed proceeded according to the Ping-Pong bi−bi mechanism (Figure 4), as previously observed for transhydrogenation by Itagaki et al.[17]
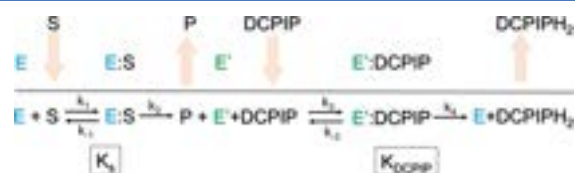
Figure 4. Schematic representation of the nonsequential Ping-Pong bi−bi mechanism.[64−66] K_S and K_{DCPIP} are dissociation constants of complexes E:S and E':DCPIP, respectively. k_2 and k_4 represent k_{cat} for RHR and OHR, respectively.

The steroid substrate (S) binds to the enzyme (E), then it is oxidized, and the dehydrogenated product (P) is released (RHR). Subsequently, the reduced enzyme (E′) binds to an electron acceptor (DCPIP). FADH$^-$ gets oxidized, and reduced DCPIPH$_2$ is released (OHR). The obtained kinetic data for different DCPIP concentrations were fitted to three types of the two-substrate reaction mechanism equations describing sequential random bi−bi, sequential ordered bi−bi, or nonsequential Ping-Pong bi−bi mechanism. The goodness of the fit was assessed by the comparison of the statistical parameters. Because R^2 values do not reflect the best fit for nonlinear models with a different number of parameters, model discrimination was also based on the Akaike information criterion (AICc parameter)[67] with smaller AICc values indicating a better agreement with kinetic data. The nonsequential Ping-Pong mechanism can be described by the following equation

$$V = \frac{V_{max}[S][DCPIP]}{K_S[DCPIP] + K_{DCPIP}[S] + [S][DCPIP]} \quad (5)$$

where V_{max} is equal to $k$catPing-Pong$[E_t]$. Kinetic parameters are provided in Table 1.

Table 1. Kinetic Parameters Obtained from Data Fitted to the Nonsequential Ping-Pong bi−bi Mechanism Model

S	K_S [μM]	K_{DCPIP} [μM]	V_{max} [μM min^{-1}]	$k_{cat}^{Ping-Pong}$ [s^{-1}]
AD	119.6 ± 16.3	40.2 ± 6.9	181.1 ± 17.3	479.1 ± 45.8

RHR mechanism investigation.

We started the analysis of the RHR with pre-steady-state kinetics conducted for 17-MT and global fit of the kinetic data with the simple RHR model assuming reversibility of both steps

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} E: I \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E: P \quad (6)$$

As a result, we were able to estimate values of the kinetic constants (Table 2) describing binding (k_1) and release of the

Table 2. Kinetic Parameters Obtained from Global Data Fitted to the Pre-Steady-State Experiment with 17-MT

|	constant value	stdd	10% confidence range
k_1 [s^{-1} μM^{-1}]	9.0	0.1	6.9−12.3
k_{-1} [s^{-1}]	872	12	676−1259
k_2 [s^{-1}]	422	1.3	327−593
k_{-2} [s^{-1}]	98.2	0.4	77−129

substrate as well as a composite kinetic constant describing jointly two steps of the RHR reaction in the direction of substrate oxidation (k_2) or the reverse direction of substrate reduction (k_{-2}). The ratio of k_{-1} to k_1 yields K_D of 96.7 μM, which is half of the estimated K_D value for AD, indicating a higher affinity of KSTD1 for 17-MT (see Table 3). More

Table 3. Pre-Steady-State Kinetic Parameters for Dehydrogenation of AD at Different Viscosities Introduced by Glycerol (Microviscosity) or PEG 20000 (Macroviscosity)

η [mPa s]	K_D [μM]	k_2 [s^{-1}]	k_2/K_D [M-1 s-1]
		Glycerol	
1.006 (η0)	227.7 ± 13.9	658.4 ± 23.6	(2.9 ± 0.2) × 10^6
1.283	331.2 ± 22.9	571.2 ± 26.3	(1.7 ± 0.1) × 10^6
1.671	339.5 ± 15.0	379.7 ± 11.2	(1.1 ± 0.1) × 10^6
		PEG 20000	
1.011 (η0)	208.0 ± 13.7	533.4 ± 20.0	(2.6 ± 0.2) × 10^6
2.711	210.5 ± 12.2	374.4 ± 12.6	(1.8 ± 0.1) × 10^6
7.636	348.9 ± 31.2	337.9 ± 19.8	(1.0 ± 0.1) × 10^6

important, however, is the fact that the rate of E:S formation is of the same magnitude as the chemical step (k_2 422 s^{-1}), exceeding it only above substrate concentration of 46 μM. This result indicates that any process slowing down the substrate binding (such as diffusion or hindrance of the environment to the conformational change of the enzyme upon enzyme binding) can influence the observed reaction velocity.

In pursuit of an explanation of the low KIE observed for KSTDs by us (see below) and Jerussi and Ringold,[21] and in the light of these results, we decided to evaluate to what extent a substrate diffusion into the enzyme active site or the enzyme conformational change occurring upon substrate binding can

156
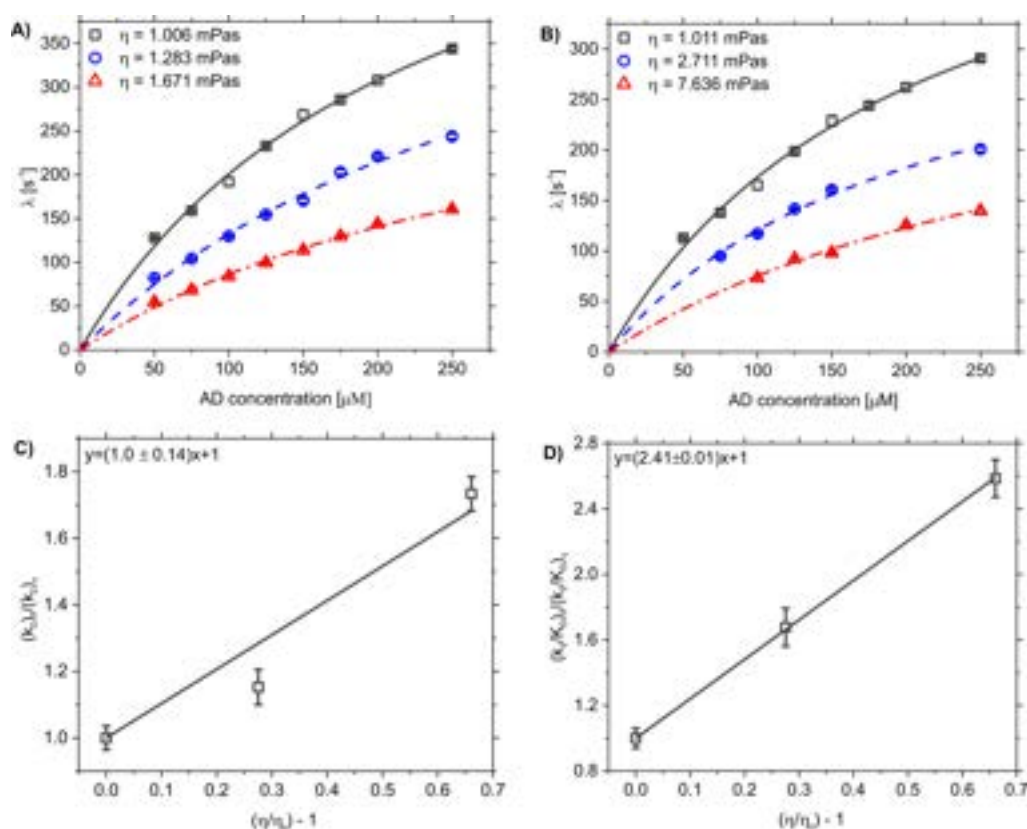
Figure 5. Effect of solvent viscosity on the pre-steady-state kinetics parameters for KSTD with AD as a substrate (A) Michaelis−Menten model fitting kinetics in glycerol, (B) Michaelis−Menten model fitting kinetics in PEG 20 000, (C) effect on the k_2 value in glycerol, (D) effect on the k_2/K_D value in glycerol (see Supporting Information for figures with PEG). Reaction conditions: 50 mM Tris-HCl buffer pH 8.0 with 8.9 μM of KSTD, 50−250 μM of AD in EGME (6%) and 0.27 M ($\square$), 1.27 M ($\square$) or 2.27 M ($\triangle$) glycerol (A) or 0% ($\square$), and 4.2% ($\square$) or 10% ($\triangle$) glycerol PEG 20000 (B), respectively.

affect the kinetic parameters. To address this, we measured the KSVE.[68] The KSVE was determined by rapid mixing of the enzyme with AD under anaerobic conditions in the stopped-flow spectrophotometer. The measurements were carried out at 30 °C using glycerol or PEG 20000 as a viscosigens providing micro- or macroviscosity, respectively. The effect of substrate concentration on the net decay rate of flavin reduction (λ) at increasing glycerol or PEG 20000 concentration is shown in Figure 5. λ can be described by eq 7[69]

$$\lambda = \frac{k_2[S]}{K_D + [S]} \tag{7}$$

where k_2 is the limiting rate of FAD reduction under saturation conditions ($[S] \gg [E]$) and K_D is the dissociation constant of the enzyme−substrate complex. Over the microviscosity range of 1.006−1.671 mPa s, we observed a statistically significant decrease of k_2 (from 658 to 407 s^{-1}) and an increase of the K_D values (from 224 to 372 μM). The obtained data were displayed in a plot of the normalized kinetic parameter: $(k_2)_0/(k_2)_\eta$ and $(k_2/K_D)_0/(k_2/K_D)_\eta$, (Figure 5) as a function of relative viscosity ($\eta/\eta_0 - 1$), and then, the linear function was fitted with the intercept fixed at 1.[68] The value of the linear slope determines the degree of dependence of the normalized kinetic parameters on viscosity. The slope of the line in the plot presented in Figure 5C is 1.0 ± 0.14 and indicates that the process catalyzed by KSTD strongly depends on solution

microviscosity. Meanwhile, in the experiment with PEG 20000, the slope of the line in the plot presented in the Supporting Information is only 0.12 ± 0.05, which indicates minuscule dependence of the catalytic step from the macroviscosity introduced by PEG. Therefore, we can conclude that there is no significant conformational change of the protein during the reaction.

In the case of the plot of $(k_2/K_D)_0/(k_2/K_D)_\eta$ versus $(\eta/\eta_0 - 1)$, for glycerol, the slope was >2 (see Figure 5D), which suggests that the formation of the enzyme−substrate complex is also viscosity-dependent and may be partially limited by diffusion. As the experiment with PEG on $(k_2/K_D)_0/(k_2/K_D)_\eta$ yielded the slope of 0.26 ± 0.01, we concluded that the conformational change of the enzyme upon binding is only slightly limited by macroviscosity introduced by PEG and the binding is predominantly limited by the diffusion of the substrate.

In a separate experiment, we have confirmed that glycerol has no inhibitory effect on the enzyme (data not shown).

The conclusions drawn from KSVE are in agreement with the analysis of the MD simulations conducted for the study. In particular, we did not observe any significant changes in the enzyme conformation upon substrate binding. However, our MD simulations might not have been long enough to describe such a conformational change, and such phenomena were not the focus of this paper.

Figure 6. Reaction scheme of dehydrogenation of DHT. The substrate is marked in orange, and dashed lines represent hydrogen bond interactions. The first step corresponds to the abstraction of the 2β proton at the C2 position of ketosteroid by tyrosine anion Tyr 318 and the second step is the 1α hydride transfer from the C1 position of the substrate to the N5 atom of the FAD.

The complementary kinetic studies conducted for AD with the steady-state or pre-steady-state technique enabled the comparison of the eigenvalues associated with those two phases.

The k_2 estimated from the rate of FAD reduction (RHR) with AD was in the range of 533−658 s^{-1} (depending on the enzyme batch, Table 3), while $k_{cat}^{Ping-Pong}$ from steady-state reaction with AD and DCPIP was 479.1 ± 45.8 s^{-1}. The k_2 turns out to be even higher (604−742 s^{-1}) if the kinetic results from the pre-steadystate experiment are extrapolated to the same viscosity as that measured for the reaction mixture used in the steady-state experiment (η_{H2O} = 0.896 mPa s). As values of extrapolated k_2 and $k_{cat}^{Ping-Pong}$ were obtained under different conditions, it is very difficult to directly compare their values. However, if one analyzes the relationship between $k_{cat}^{Ping-Pong}$, k_2, and k_4, (eq 8), then one can see that k_2 and k_4 should be of the same order of magnitude.

$$k_{cat}^{Ping-Pong} = \frac{k_2 k_4}{k_2 + k_4} \tag{8}$$

Modeling of the Reaction Mechanism. To test the mechanistic hypothesis for Δ^1-dehydrogenation of 3-ketosteroids by KSTD with deprotonated Tyr318, we first computed QM/MM PESs describing the two process, that is, (i) the abstraction of a proton from C2 by Tyr318 and (ii) the hydride transfer from C1 to FAD, as presented in the Supporting Information. The minimum energy pathway from the reactants (E:S) to products (E:P) leads through the formation of a carbanionic intermediate (E:I) (Figure 6). According to the topology of the PES, a process involving the simultaneous transfer of both hydrogen atoms does not exist. The alternate reaction pathway starting with a hydride transfer also turns out to be extremely unlikely due to the very high energy of the carbocation intermediate (see Supporting Information). Therefore, the formation of the carbanion by proton transfer from C2 is the only thermodynamically possible way to start a reaction. This route leads to the currently accepted mechanism. Then, the FESs were computed in terms of PMFs to obtain the free energy profile of the full reaction. This procedure allowed us to describe the RHR process, as summarized in Figure 7. Analysis of the initial enzyme−substrate complex (E:S) shows how the substrate is localized in the active site pocket with its ring A almost parallel to the isoalloxazine ring system of FAD. The steroid position is stabilized by H-bonds between the 3-keto group and Tyr487 as well as Gly491. Tyr318 anion is interacting by an H-bond with Tyr119. These H-bonds remain stable throughout the whole RHR. In the first step of the reaction, the steroid is activated by

the abstraction of the proton from the 2β position by the Tyr318 anion. Then, when the enolate intermediate is formed (E:I), a new H-bond is created between the 3-keto group and Tyr318 protonated with 2β-H atom. Formation of the intermediate enolate results in a change in the ring A geometry and repositioning of the substrate in the active site, which decreases the distance between 1αH of the substrate and N5$_{FAD}$ atom in E:I. This change of geometry facilitates the next step of the reaction, when the hydride is transferred from the 1α position of the ketosteroid to N5 atom in the FAD residue, resulting in a two-electron-reduced (hydroquinone) unprotonated state of flavin. Finally, the double bond is formed between C1 and C2 atoms of the steroid and the enzyme−product complex is formed (E:P). The resulting FESs obtained for the reaction with DHT (blue line) and 17-MT (orange line) are shown in Figure 8.

The general trend of the FESs obtained for both molecules is similar, that is, the higher barrier is associated with hydride transfer which corresponds to TS2, with free energy barriers of 18.0 and 16.3 kcal·mol^{-1} for DHT and 17-MT, respectively. Both formed intermediates are 3.3 and 2.8 kcal·mol^{-1} less stable than E:S of DHT and 17-MT, respectively. Both TS1 and TS2 barriers are lower for 17-MT than for DHT (13.4 and 15.5 kcal·mol^{-1} vs 16.3 and 18.0 kcal·mol^{-1}, respectively for TS1 and TS2). A detailed description of the calculated reaction mechanism with full distances and geometry analysis is available in Supporting Information. As can be deduced from the FESs, the enolate is quite well stabilized, being only a few kcal·mol^{-1} above the free energy of the enzyme−substrate complex. In contrast to previous assumptions, the free energy barriers for the first and second steps are comparable. Besides that, however, this mechanism is in line with the widely accepted hypothesis. The mutational studies described by Rohman et al.[15] provided useful clues about the mechanism. Replacing Tyr318 with Phe rendered the enzyme completely inactive, while Y119F and Y487F mutants retained only trace activities (0.05 and 2.6% of original activity, respectively). The experimental data suggest that these tyrosines are crucial for enzyme reactivity. Our QM/MM modeling confirmed the role of both residues. Tyr487 takes part in the binding of the ketosteroid during the whole reaction and stabilizes its activated intermediate by a hydrogen bond. Meanwhile, Tyr119 interacts with Tyr318 anion with a hydrogen bond, and this interaction also remains stable throughout the whole RHR. The Tyr318 anion acts as a base that abstracts a proton from the substrate. Surprisingly, after protonation, it becomes a hydrogen bond donor to the intermediate product, similarly to Tyr487 that interacts with the keto group of the steroid.
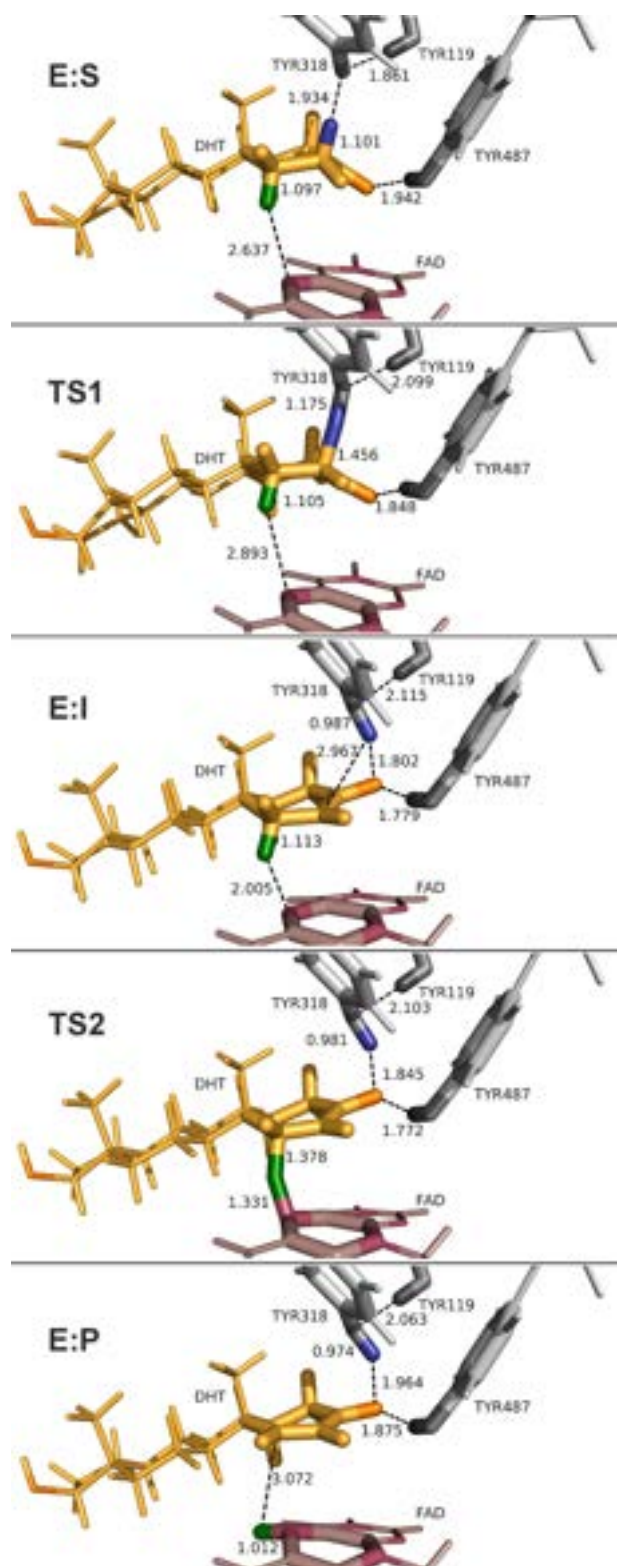
Figure 7. Details of the active site of representative structures of the key-states located along the Δ^1-dehydrogenation of DHT. The distances are given in Å. The transferring atoms are in blue and green.

Therefore, Tyr318 does not only acts as a catalytic base activating the C−H bond in the first step of the reaction but
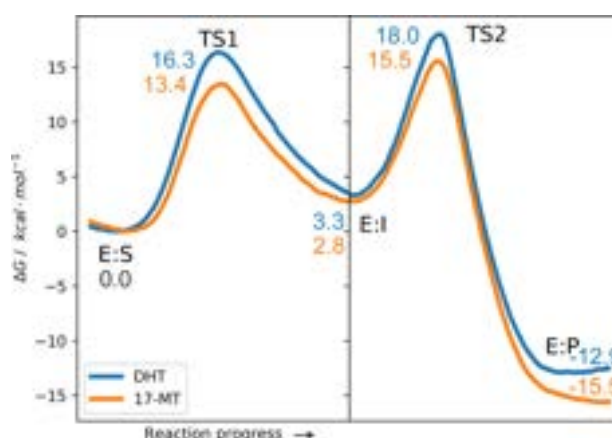
Figure 8. FESs of Δ^1-dehydrogenation of 17-MT (orange line) and DHT (blue line). Results obtained at B3LYP/6-311++g-(2d,2p):AM1/AMBER level of theory at 303 K.

also stabilizes the intermediate enolate. This fact can explain the residual activity of the Y487F, as Tyr318 can stabilize the carboanion intermediate. Up to date, this role was attributed solely to stabilization by H-bond formed by the 3-keto group, Tyr487 and Gly491.[15] A similar situation was postulated for Δ^5-3-ketosteroid isomerase where ketosteroid homoenolate is stabilized by low barrier hydrogen bonds.[70] Finally, the average Mulliken and ChelpG charges confirm that the second step occurs via hydride transfer, leading to the formation of FADH$^-$ (see population analysis deposited in the Supporting Information).

Most of the KSTDs catalyze reaction at neutral to basic pH, which supports the hypothesis of the presence of the tyrosyl ion in the active site. However, some reports indicate that the Δ^1-dehydrogenation can proceed under mild acidic conditions,[25,27] and there were no unequivocal proofs for the deprotonated state of Tyr318. Therefore, we investigated a possible alternative mechanism with protonated Tyr318. However, preliminary results indicate that the potential energies of the first TS and intermediate are extremely high (44.8 and 27.9 kcal·mol^{-1}, respectively); therefore, an alternative mechanism with protonated Tyr318 can be discarded. A detailed description of this alternative pathway is available in Supporting Information.

KIEs and BIEs. The values of KIE were initially estimated with a steady-state technique using 200 μM concentrations of steroids and DCPIP, following the method used by Jerussi and Ringold.[21] Such experiments yielded values of KIE of 1.49 $\pm$ 0.04 and 1.28 $\pm$ 0.02 for DHT and 17-MT, respectively. As our preliminary stopped-flow analysis indicated that the enzyme reoxidation may influence the result of the experiment, we also collected KIEs using the pre-steady-state technique, recording directly the reduction of FAD. Although we have not observed an increase of the KIE for 17-MT, which was calculated as the ratio of $\lambda_{obs}^{H}/\lambda_{obs}^{D} \approx k_2^{H}/k_2^{D}$ and equaled 1.5 $\pm$ 0.04 (see Supporting Information), we measured a significantly higher value for C1-substituted DHT, where the k_2^{H}/k_2^{D} ratio was 3.5 $\pm$ 0.04.

To gain further insight into the effect of binding to the active site of the enzyme of the labeled substrates on KIE, we measured the $^D(V/K)$ in the competition experiment. The obtained values were 2.2 $\pm$ 0.02 for 17-MT and 2.4 $\pm$ 0.07 for DHT, indicating that in the case of DHT, the binding

Table 4. Theoretical QM/MM (B3LYP/AMBER) and Experimental Values of KIEs Obtained for Deuterium-Labeled 17-MT and DHT[a]

	theoretical KIE ± SE	experimental KIE ± SE		
		direct pre-steady-state	competitive $^D(V/K)$	direct steady-state
	17-MT/2,2,4,6,6-d5-17-MT			
E:S → TS1	5.39 ± 0.04			
E:I → TS1	5.13 ± 0.02			
E:I → TS2	0.92 ± 0.02	1.5 ± 0.04	2.2 ± 0.02	1.28 ± 0.02
E:S → TS2	0.98 ± 0.01			
	DHT/1α,16,16,17-d_4-DHT			
E:S → TS1	1.12 ± 0.01			
E:I → TS1	0.97 ± 0.01			
E:I → TS2	4.36 ± 0.05	3.5 ± 0.04	2.4 ± 0.07	1.5 ± 0.04
E:S → TS2	4.95 ± 0.05			

[a]Theoretical values of KIE were computed for each possible individual step of the reaction. The values of direct pre-steady-state KIE correspond to the ratio of the observed λ value (see eq 7); values of direct steady-state KIE correspond to the ratio of activities measured in standard activity assay (200 μM DCPIP/steroid); and $^D(V/K)$ values were obtained with a competition method.

decreases the observed KIE. Surprisingly, we observed a slight increase of the KIE for 17-MT when measured in a competitive experiment compared to a direct measurement method. These results indicate that we can expect an inverse BIE for DHT and a normal (above 1) effect for 17-MT.

To address this issue, we calculated BIE at the B3LYP/AMBER level of theory, obtaining values of 0.97 ± 0.06 and 0.76 ± 0.05, respectively, for DHT and 17-MT. As the values of the calculated BIE were lower than unity for both substrates, we can conclude that under simulation conditions, deuterated substrates are bound better by the enzyme than the non-deuterated ones. These predictions are in qualitative agreement only in the case of DHT.

We should underline that in our MD simulation, we model a limit case, where the substrate is either in complex with enzyme or in pure water. However, steroids are so hydrophobic that it is not possible to dissolve them in pure water in concentrations enabling the saturation of KSTD1 (100–200 μM). In our experiments, the addition of the organic solvent was used (5% dioxane or 10% EGME), which stabilized the substrate in the solution by providing thermodynamically favorable hydrophobic interactions. Therefore, the experimental BIEs should be closer to unity than those predicted in our modeling, and we cannot exclude the possibility of even normal (i.e., above 1) BIE for 17-MT.

We assume that the qualitative agreement of our BIE prediction for DHT with results of V_{max} and V/K KIE stems from the difference in the deuteration pattern between 2,2,4,6,6-d_5-17-MT and 1,16,16,17-d_4-DHT. In the case of 17-MT, the deuteron labels are in ring A or B, which are bound inside the enzyme, while in the case of DHT, three deuterons are located on the ring D, which is exposed to the solvent (see Supporting Information). As a result, the environment of the DHT deuterons changes only in the case of the C1 substituent, while for the rest of the deuterons, the situation remains unchanged (no difference between solvent and enzyme–substrate complex). In consequence, the lack of organic solvent in the simulation will strongly influence the estimation of BIE for 17-MT, while only a small effect will be expected for DHT.

Generally, KIE experiments are in agreement with our estimation of k_2 and k_4. The kinetic assay, based on the process of DCPIP reduction (k_4), is less sensitive to isotope labels, which results in lowering of the observed KIE (i.e., partial masking the intrinsic KIE).

Our experimental KIE results obtained for KSTD from *R. erythropolis* are similar to those reported by Jerussi and Ringold.[21] They performed a series of KIE measurements using KSTD from *B. sphaericus* with 5α-3-keto- and Δ^4-3-ketosteroids deuterated at 1α and 2β positions. Surprisingly, it was found that substitutions at any of these positions affected the V_{max}. Substantial values of KIE were observed mostly for 1α substitution (V_{maxH}/V_{maxD} = 1.9) and much lower for 2β (V_{maxH}/V_{maxD} = 1.2−1.3) in the case of 5α-3-keto steroids. However, in the case of ketosteroid with an additional double bond in ring A (i.e., Δ^4-3-ketosteroids), both substitutions, at 1α or 2β, affected the observed kinetic rate (V_{maxH}/V_{maxD} = 2.4−2.5). Values obtained for 1α and 2β substitutions were too high to be explained by the secondary isotope effect.

To understand this unusual effect we conducted a series of KIE calculations based on QM/MM optimized geometries of E:S, TS1, E:I, TS2, and E:P for 17-MT and DHT. Results of KIEs derived from calculations and experiments are shown in Table 4.

For the Δ^1-dehydrogenation of DHT, isotopic substitutions at 1α should strongly affect the hydride transfer (TS2) but have an almost negligible influence on the proton transfer (TS1). On the other hand, the experiments showed that substitutions at each position (i.e., C1 or C2) affected the kinetics of the RHR and the whole reaction. Such behavior could be explained with an unknown mechanism where both hydrogens are transferred simultaneously, and what would be observed was primary KIE for both substituted positions. This hypothesis was, however, discarded as experimental data,[20,21] and results from our calculations show that hydrogen atoms are removed in a step-wise manner. Another potential explanation of the experimental values would be a situation where the barrier heights of TS1 and TS2 were so close, that both of them affected the reaction rate.[21] As it turns out, our QM/MM free energy profile of the most favorable mechanism combined with calculations of intrinsic KIE confirms this hypothesis (see Figure 8). Finally, the overall value of KIE can be further decreased by binding and other non-chemical steps (e.g., conformational changes of the protein, diffusion limitation of reagents, as demonstrated by KSVE), but for qualitative analysis, we will focus only on the influence of reaction profile and intrinsic KIEs to the overall KIE. The studied reaction is taking place according to the following reaction scheme (9)

$$E: S \underset{k_{I-S}}{\overset{k_{S-I}}{\rightleftharpoons}} E: I \xrightarrow{k_{I-P}} E: P \tag{9}$$

Due to the high barrier between the products and the intermediate (see Figure 8), a reverse process from products to the intermediate (E:P to E:I) can be considered negligible. Then, the expression for k_2 would be

$$k_2 = \frac{k_{S-I} k_{I-P}}{k_{S-I} + k_{I-S} + k_{I-P}} \tag{10}$$

Using eq 10, it is possible to analyze the contribution of each kinetic constant to the overall KIE. We will use the following notation

$$KIE_i = \frac{k_i^H}{k_i^D} \tag{11}$$

where KIE_i is the KIE associated with an individual step characterized by kinetic constant k_i ($i = S-I, I-S, I-P$). KIEs on the k_2 value would be as follows (derivation is given in Supporting Information)

$$KIE_2 = KIE_{S-I} KIE_{I-P} \cdot \frac{\frac{1}{KIE_{S-I}} + \frac{1}{KIE_{I-S}} \frac{k_{I-S}^H}{k_{S-I}^H} + \frac{1}{KIE_{I-P}} \frac{k_{I-P}^H}{k_{S-I}^H}}{1 + \frac{k_{I-S}^H}{k_{S-I}^H} + \frac{k_{I-P}^H}{k_{S-I}^H}} \tag{12}$$

Then, according to eq 12, KIE_2 is a function of 5 variables: KIE_{S-I}, KIE_{I-P}, KIE_{I-S}, k_{I-S}^H/k_{S-I}^H, and k_{H-P}^H/k_{S-I}^H. Three of these variables can be obtained based on the results of QM/MM calculations (KIE_{S-I}, KIE_{I-P}, KIE_{I-S}). For further analysis, we set a range of 10^{-3} to 10^3 for both ratios k_{I-S}^H/k_{S-I}^H and k_{I-P}^H/k_{S-I}^H. These ratios were estimated from the free energy profiles and Eyring equation. Nevertheless, we have to bear in mind that such values are subject to some error. It is also possible to replace k_{I-S}^H/k_{S-I}^H and k_{I-P}^H/k_{S-I}^H in eq 12 by a function of free energies

$$\log \frac{k_i}{k_{S-I}} = \frac{\Delta G_{S-I}^{\ddagger} - \Delta G_i^{\ddagger}}{\ln(10) RT} \tag{13}$$

As a result, KIE_{obs} can be analyzed as a function of KIE_{S-I}, KIE_{I-P}, KIE_{I-S}, $\Delta GS\_I^{\ddagger} - \Delta G_{I-S}^{\ddagger}$, and $\Delta G_{S-I}^{\ddagger} - \Delta G_{I-P}^{\ddagger}$.

Taking into account the above mentioned considerations, we can plot maps of all possible values of KIE_2 for 17-MT and DHT (Figure 9). Black curves correspond to the values of KIE measured in our experiments (with confidence bands based on the experimental standard error of the mean for KIE estimation). Centers of white circles represent KIE_2 values calculated from QM/MM results: 1.09 for 17-MT and 4.78 for DHT. These circles have a radius of 1 kcal·mol^{-1} error (the error of the applied theoretical method) which translates to the confidence ranges of KIE for 0.97–2.0 and 3.31–5.0, respectively for 17-MT and DHT. As can be seen in the plots in Figure 9, the curves representing the experiments for both substrates intersect with the circles corresponding to QM/MM values. It is easy to notice that both circles represent wide ranges of KIE values. Thus, in a mechanism like this one, it is necessary to estimate free energies with very small errors to obtain an accurate theoretical prediction of KIE_{obs}, combined with experiments capable of measuring KIEs also with high accuracy. Nevertheless, a fair agreement between theoretical results and experiment can be found, even without
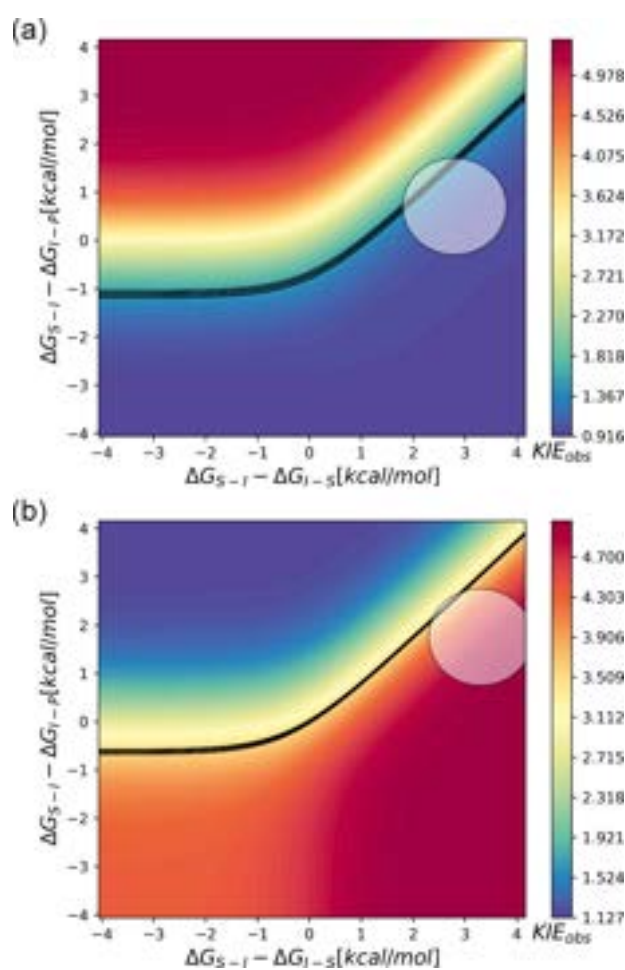
Figure 9. Maps of possible KIE values for 17-MT (a) and DHT (b) as a function of the difference in energy barriers. Black curves represent experimental values (with additional lines indicating experimental errors), and white circles represent the calculated QM/MM KIEs (with 1 kcal·mol^{-1} radius, as the error associated with the method).

considering other factors that may lead to a decrease of measured KIE.

Our analysis demonstrated that only the energy landscape of our mechanism can reconcile the experimental KIE results obtained for substitutions at both C1 and C2 positions, which both lead to KIE values significantly greater than one but much lower than the calculated intrinsic KIE. The main reason for this phenomenon is the relation between individual rate constants leading to a situation where both steps have a significant influence on the overall process.

■ CONCLUSIONS

The reaction catalyzed by Δ^1-KSTD from *R. erythropolis* was investigated using QM/MM computational methods and experimental kinetic measurements of isotope effects. We first proved that dehydrogenation proceeds via a two-step Ping-Pong bi–bi mechanism, which is in line with the widely accepted hypothesis.[16,17] Calculations performed from two different initial states of the enzyme, one with Tyr318 as a tyrosyl anion and the second for a protonated Tyr318, allow us to conclude that the existence of deprotonated tyrosine is essential for initializing the Δ^1-dehydrogenation process. We

161

have shown that the enolate intermediate is stabilized not only by Tyr487 and Gly491 but also by Tyr318. This role of Tyr318 was not reported up to date, but it is in agreement with previous mutational studies revealing the complete inactivation of the enzyme after replacing this residue. However, the retention of minimal activity of the Y487F mutant may indicate that Tyr318 can take over the stabilizing role of Tyr487.[15]

Despite the general agreement with previously proposed mechanisms for the reaction catalyzed by Δ^1-KSTD, the free energy barriers for the first and second steps deduced from our QM/MM simulations turned out to be comparable. This result appears to be at the origin of the surprising KIEs which seemed to contradict the mechanistic hypothesis for over half a century. Our computational results derived from the exploration of the FESs and calculations of BIE and intrinsic KIE, finally explained why deuterium substitution affects the overall reaction rate, regardless of the labeled position. This is the effect of comparable free energy barriers for the first and second step. The difference is so small, that both steps influence the overall reaction rate. Our experimental values of KIE obtained for KSTD1 are consistent with those described for the enzyme from *B. sphaericus*, although we were able to demonstrate that the KIE measured in the pre-steady-state at C1 is significantly higher than the value obtained in the steady-state experiment. Therefore, the conclusions seem to be generally applicable to the whole class of Δ^1-KSTD. Furthermore, we are convinced that our holistic approach to describing enzyme kinetics by experimental (steady- and pre-steady-state kinetics, KIE, and KSVE) and theoretical methods (QM/MM MD, global fit to kinetics, and analysis of potential crossing points of theory and experiments) applies to any enzymatic system and provides a potential guideline for elucidation of similar problems. We also pointed out potential pitfalls which may lead to discrepancies in modeling and experiment due to the influence of additives used to introduce the organic substrate into the assay. Quantitative agreement between computational and experimental results is nontrivial to achieve because, as can be seen from the relationship between k_2, K_D, and viscosity, the relative values of k_1, k_2, and k_4, diffusion-controlled processes and, to smaller effect, conformational changes upon binding, as well as reoxidation of the enzyme play important roles in the overall reactivity of KSTD. Despite these obstacles, however, the application of complementary techniques can lead to a better understanding of the enzymatic mechanism, which has eluded scientists for over 50 years.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acscatal.1c01479.

> PDB files of representative structures optimized at the B3LYP/AMBER level of theory and cartesian coordinates of TS (ZIP)

> Population analysis and coordinates of all stationary structures from the main reaction pathway (XLSX)

> Details on the modeling methodology and analytical LC−MS/MS procedure, a detailed description of the mechanisms, derivation of equations, details on enzyme identification (SDS PAGE, proteomics) and kinetics (figures for Ping-Pong bi−bi fit, KSVE for PEG 20000, pre-steady-state, V_{max} and $^D(V/K)$ KIE), results of DLS

for KSTD and PEG, Cartesian coordinates of TSes, details on the calculation of KIE (per atom contributions) and Wigner tunneling corrections, figures depicting binding mode of deuterated 17-MT and DHT (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Maciej Szaleniec − *Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, 30-239 Krakow, Poland;* ● orcid.org/0000-0002-7650-9263; Email: maciej.szaleniec@ikifp.edu.pl

Authors

Michał Glanowski − *Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, 30-239 Krakow, Poland*

Patrycja Wójcik − *Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, 30-239 Krakow, Poland*

Magdalena Procner − *Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, 30-239 Krakow, Poland*

Tomasz Borowski − *Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, 30-239 Krakow, Poland;* ● orcid.org/0000-0002-3450-3576

Dawid Lupa − *Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, 30-239 Krakow, Poland*

Przemysław Mielczarek − *Laboratory of Proteomics and Mass Spectrometry, Maj Institute of Pharmacology, Polish Academy of Sciences, 31-343 Krakow, Poland*

Maria Oszajca − *Faculty of Chemistry, Jagiellonian University, 30-387 Krakow, Poland;* ● orcid.org/0000-0003-2690-0085

Katarzyna Świderek − *Department of Physical and Analytical Chemistry, Universitat Jaume I, 12071 Castellón, Spain;* ● orcid.org/0000-0002-7528-1551

Vicent Moliner − *Department of Physical and Analytical Chemistry, Universitat Jaume I, 12071 Castellón, Spain;* ● orcid.org/0000-0002-3665-3391

Andrzej J. Bojarski − *Department of Medicinal Chemistry, Maj Institute of Pharmacology, Polish Academy of Sciences, 31-343 Krakow, Poland;* ● orcid.org/0000-0003-1417-6333

Complete contact information is available at:
https://pubs.acs.org/10.1021/acscatal.1c01479

Author Contributions

M.G. conducted all calculations under MSz. V.M. and K.Ś. supervised, analyzed results, developed discussion, co-authored the main text, edited text, and created figures. P.W. overexpressed and purified KSTD, conducted experiments with the KSVE effect and KIE, co-authored methods and results section, edited text, and created figures. M.P. conducted steady-state kinetic experiments and co-authored the method section. P.M. conducted proteomic analysis and co-authored methods. D.L. conducted viscosity and DLS experiments and co-authored methods and results/Supporting Information sections. T.B. developed Octave scripts and conducted global-fit of 17-MT kinetics, M.O. supervised pre-steady-state kinetic experiments and edited text. K.Ś. and V.M. supervised

■ ABBREVIATIONS

Δ^1KSTD, Δ^1-ketosteroid dehydrogenase; 17-MT, 17-methyl-testosterone, 17β-hydroxy-17-methyl-4-androsten-3-one; DHT, dihydrotestosterone, 17β-hydroxy-5α-androstan-3-one; DCPIP, 2,6-dichloroindophenol; AD, 4-androsten-3,17-dione; RHR, reductive half reduction; FES, free energy surface; qm/mm, quantum mechanics/molecular mechanics; md, molecular dynamics

■ REFERENCES

(1) Ericson-Neilsen, W.; Kaye, A. D. Steroids: Pharmacology, Complications, and Practice Delivery Issues. *Ochsner J.* **2014**, *14*, 203−7.

(2) Solinas, C.; Perra, L.; Aiello, M.; Migliori, E.; Petrosillo, N. A Critical Evaluation of Glucocorticoids in the Management of Severe COVID-19. *Cytokine Growth Factor Rev* **2020**, *54*, 8−23.

(3) Global Steroids Industrial Chain Market Growth Overview and Forecast b. 2021, https://www.qyresearch.com/index/detail/701226/global-steroids-industrial-chain-market (accessed Feb 05, 2021).

(4) Fernández-Cabezón, L.; Galán, B.; García, J. L. New Insights on Steroid Biotechnology. *Front. Microbiol.* **2018**, *9*, 958.

(5) James, P.; Rolls, A. Dehydrogenation of Corticoids without Side Chain Degradation by Septomyxa. U.S. Patent 4,088,537 A, May 9, 1978.

(6) Spelling, T. Process for the Overexpression of Dehydrogenases. U.S. Patent 7,416,866 B2, Aug 26, 2008.

(7) Chen, M.-M.; Wang, F.-Q.; Lin, L.-C.; Yao, K.; Wei, D.-Z. Characterization and application of fusidane antibiotic biosynthesis enzyme 3-ketosteroid-Δ1-dehydrogenase in steroid transformation. *Appl. Microbiol. Biotechnol.* **2012**, *96*, 133−142.

(8) Mao, S.; Wang, J.-W.; Liu, F.; Zhu, Z.; Gao, D.; Guo, Q.; Xu, P.; Ma, Z.; Hou, Y.; Cheng, X.; Sun, D.; Lu, F.; Qin, H.-M. Engineering of 3-ketosteroid-Δ1-dehydrogenase based site-directed saturation mutagenesis for efficient biotransformation of steroidal substrates. *Microb. Cell Factories* **2018**, *17*, 141.

(9) Costa, S.; Zappaterra, F.; Summa, D.; Semeraro, B.; Fantin, G. Δ1-Dehydrogenation and C20 Reduction of Cortisone and Hydro-cortisone Catalyzed by Rhodococcus Strains. *Molecules* **2020**, *25*, 2192.

(10) Arthur, N. Process for Production of Dienes by Corynebacteria. U.S. Patent 2,837,464 A, June 3, 1958.

(11) Herráiz, I. Chemical Pathways of Corticosteroids, Industrial Synthesis from Sapogenins. *Methods Mol. Biol.* **2017**, *1645*, 15−27.

(12) Donova, M. V. Transformation of Steroids by Actinobacteria: A Review. *Appl. Biochem. Microbiol.* **2007**, *43*, 1−14.

(13) Donova, M. V.; Egorova, O. V. Microbial Steroid Trans-formations: Current State and Prospects. *Appl. Microbiol. Biotechnol.* **2012**, *94*, 1423−1447.

(14) Rohman, A.; Dijkstra, B. W. The role and mechanism of microbial 3-ketosteroid Δ1-dehydrogenases in steroid breakdown. *J. Steroid Biochem. Mol. Biol.* **2019**, *191*, 105366.

(15) Rohman, A.; Van Oosterwijk, N.; Thunnissen, A.-M. W. H.; Dijkstra, B. W. Crystal Structure and Site-directed Mutagenesis of 3-Ketosteroid Δ1-Dehydrogenase from Rhodococcus erythropolis SQ1 Explain Its Catalytic Mechanism. *J. Biol. Chem.* **2013**, *288*, 35559−35568.

(16) Dijkstra, B. W.; van Oosterwijk, N.; Rohman, A. Structure and Catalytic Mechanism of 3-Ketosteroid Dehydrogenases. *Procedia Chem.* **2016**, *18*, 3−11.

(17) Itagaki, E.; Matsushita, H.; Hatta, T. Steroid Transhydrogenase Activity of 3-Ketosteroid-δ-Dehydrogenase from Nocardia corallina. *J. Biochem.* **1990**, *108*, 122−127.

(18) Hayano, M.; Ringold, H. J.; Stefanovic, V.; Gut, M.; Dorfman, R. I. The Stereochemical Course of Enzymatic Steroid 1,2-Dehydrogenation. *Biochem. Biophys. Res. Commun.* **1961**, *4*, 454−459.

(19) Levy, H. R.; Talalay, P. Bacterial Oxidation of Steroids. *J. Biol. Chem.* **1959**, *234*, 2014−2021.

(20) Ringold, H. J.; Hayano, M.; Stefanovic, V. Concerning the Stereochemistry and Mechanism of the Bacterial C-1,2 Dehydrogen-ation of Steroids. *J. Biol. Chem.* **1963**, *238*, 1960−1965.

(21) Jerussi, R.; Ringold, H. J. The Mechanism of the Bacterial C-1,2 Dehydrogenation of Steroids. III. Kinetics and Isotope Effects*. *Biochemistry* **1965**, *4*, 2113−2126.

(22) Ringold, H. J.; Gut, M.; Hayano, M.; Turner, A. Contribution to the Steric Course of Steroid Reduction and Dehydrogenation. *Tetrahedron Lett.* **1962**, *3*, 835−837.

(23) Shao, M.; Sha, Z.; Zhang, X.; Rao, Z.; Xu, M.; Yang, T.; Xu, Z.; Yang, S. Efficient androst-1,4-diene-3,17-dione production by co-expressing 3-ketosteroid-Δ1-dehydrogenase and catalase inBacillus subtilis. *J. Appl. Microbiol.* **2017**, *122*, 119−128.

(24) Itagaki, E.; Wakabayashi, T.; Hatta, T. Purification and characterization of 3-ketosteroid-Δ1-dehydrogenase from Nocardia corallina. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1990**, *1038*, 60−67.

(25) Wojtkiewicz, A. M.; Wójcik, P.; Procner, M.; Flejszar, M.; Oszajca, M.; Hochołowski, M.; Tataruch, M.; Mrugała, B.; Janeczko, T.; Szaleniec, M. The efficient Δ1-dehydrogenation of a wide spectrum of 3-ketosteroids in a broad pH range by 3-ketosteroid dehydrogenase from Sterolibacterium denitrificans. *J. Steroid Biochem. Mol. Biol.* **2020**, *202*, 105731.

(26) Tataruch, M.; Wójcik, P.; Wojtkiewicz, A. M.; Zaczyk, K.; Szymańska, K.; Szaleniec, M. Application of Immobilized Cholest-4-en-3-one Δ1-Dehydrogenase from Sterolibacterium Denitrificans for Dehydrogenation of Steroids. *Catalysts* **2020**, *10*, 1460.

(27) Chiang, Y.-R.; Ismail, W.; Gallien, S.; Heintz, D.; Van Dorsselaer, A.; Fuchs, G. Cholest-4-En-3-One-Δ 1 -Dehydrogenase, a Flavoprotein Catalyzing the Second Step in Anoxic Cholesterol Metabolism. *Appl. Environ. Microbiol.* **2008**, *74*, 107−113.

(28) Kabsch, W. A discussion of the solution for the best rotation to relate two sets of vectors. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **1978**, *34*, 827−828.

(29) Søndergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pKa Values. *J. Chem. Theory Comput.* **2011**, *7*, 2284−2295.

163

(30) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525−537.

(31) Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0: Automating pK Prediction and the Preparation of Biomolecular Structures for Atomistic Molecular Modeling and Simulations. *Nucleic Acids Res.* **2012**, *40*, W537−W541.

(32) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16*, Revision B.01, 2016.

(33) Singh, U. C.; Kollman, P. A. An Approach to Computing Electrostatic Charges for Molecules. *J. Comput. Chem.* **1984**, *5*, 129−145.

(34) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. An Overview of the Amber Biomolecular Simulation Package. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2013**, *3*, 198−210.

(35) Dupradeau, F.-Y.; Cezard, C.; Lelong, R.; Stanislawiak, E.; Pecher, J.; Delepine, J. C.; Cieplak, P. R.E.DD.B.: A database for RESP and ESP atomic charges, and force field libraries. *Nucleic Acids Res.* **2007**, *36*, D360−D367.

(36) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(37) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* **2003**, *24*, 1999−2012.

(38) Grest, G. S.; Kremer, K. Molecular Dynamics Simulation for Polymers in the Presence of a Heat Bath. *Phys. Rev. A: At., Mol., Opt. Phys.* **1986**, *33*, 3628−3631.

(39) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103*, 227−249.

(40) Field, M. J.; Albe, M.; Bret, C. l.; Proust-De Martin, F.; Thomas, A. The Dynamo Library for Molecular Simulations Using Hybrid Quantum Mechanical and Molecular Mechanical Potentials. *J. Comput. Chem.* **2000**, *21*, 1088−1100.

(41) Krzemińska, A.; Paneth, P.; Moliner, V.; Świderek, K. Binding Isotope Effects as a Tool for Distinguishing Hydrophobic and Hydrophilic Binding Sites of HIV-1 RT. *J. Phys. Chem. B* **2015**, *119*, 917−927.

(42) Field, M. J.; Bash, P. A.; Karplus, M. A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations. *J. Comput. Chem.* **1990**, *11*, 700−733.

(43) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23*, 187−199.

(44) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. THE Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13*, 1011−1021.

(45) Martí, S.; Moliner, V.; Tuñón, I. Improving the QM/MM Description of Chemical Processes: A Dual Level Strategy To Explore the Potential Energy Surface in Very Large Systems. *J. Chem. Theory Comput.* **2005**, *1*, 1008−1016.

(46) Ruiz-Pernía, J. J.; Silla, E.; Tuñón, I.; Martí, S.; Moliner, V. Hybrid QM/MM Potentials of Mean Force with Interpolated Corrections. *J. Phys. Chem. B* **2004**, *108*, 8427−8433.

(47) Baker, J.; Kessi, A.; Delley, B. The Generation and Use of Delocalized Internal Coordinates in Geometry Optimization. *J. Chem. Phys.* **1996**, *105*, 192−212.

(48) Samuel, G.; Laidler Keith, J. E. H. *The Theory of Rate Processes: The Kinetics of I Chemical Reactions, Viscosity, Diffusion and Electrochemical Phenomena*; Mcgraw-Hill Book Compagny: New York, 1941.

(49) Keck, J. C. Variational Theory of Reaction Rates. In *Advances in Chemical Physics*; Prigogine, I., Ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2007; pp 85−121.

(50) Truhlar, D. G.; Garrett, B. C.; Klippenstein, S. J. Current Status of Transition-State Theory. *J. Phys. Chem.* **1996**, *100*, 12771−12800.

(51) Świderek, K.; Arafet, K.; Kohen, A.; Moliner, V. Benchmarking Quantum Mechanics/Molecular Mechanics (QM/MM) Methods on the Thymidylate Synthase-Catalyzed Hydride Transfer. *J. Chem. Theory Comput.* **2017**, *13*, 1375−1388.

(52) Krzemińska, A.; Moliner, V.; Świderek, K. Dynamic and Electrostatic Effects on the Reaction Catalyzed by HIV-1 Protease. *J. Am. Chem. Soc.* **2016**, *138*, 16283−16298.

(53) Świderek, K.; Kohen, A.; Moliner, V. The Influence of Active Site Conformations on the Hydride Transfer Step of the Thymidylate Synthase Reaction Mechanism. *Phys. Chem. Chem. Phys.* **2015**, *17*, 30793−30804.

(54) Wigner, E. Crossing of Potential Thresholds in Chemical Reactions. *Z. Phys. Chem.* **1932**, *B19*, 203−216.

(55) Truhlar, D. G.; Garrett, B. C. Variational Transition State Theory in the Treatment of Hydrogen Transfer Reactions. In *Hydrogen-Transfer Reactions*; John Wiley & Sons, Ltd, 2006; pp 833−874.

(56) Świderek, K.; Paneth, P. Binding Isotope Effects. *Chem. Rev.* **2013**, *113*, 7851−7879.

(57) Rohman, A.; van Oosterwijk, N.; Dijkstra, B. W. Purification, crystallization and preliminary X-ray crystallographic analysis of 3-ketosteroid Δ1-dehydrogenase fromRhodococcus erythropolisSQ1. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2012**, *68*, 551−556.

(58) Eaton, J. W.; Bateman, D.; Hauberg, S.; Wehbring, R. *A High-Level Interactive Language for Numerical Computations*, GNU Octave Version 5.2.0 Manual, 2020.

(59) Johnson, K. A. Fitting Enzyme Kinetic Data with KinTek Global Kinetic Explorer. *Methods Enzymol.* **2009**, *467*, 601−626.

(60) Prinz, H. *Numerical Methods for the Life Scientist: Binding and Enzyme Kinetics Calculated with GNU Octave and MATLAB*; Springer-Verlag Berlin Heidelberg: Berlin, 2011.

(61) Glanowski, M.; Wójcik, P.; Procner, M.; Borowski, T.; Lupa, D.; Mielczarek, P.; Oszajca, M.; Świderek, K.; Moliner, V.; Bojarski, A. J.; Szaleniec, M. *Enzymatic Δ1-Dehydrogenation of 3-Ketosteroids—Reconciliation of Kinetic Isotope Effects with the Reaction Mechanism*; Mendeley Data, 2021.

(62) Rugor, A.; Wójcik-Augustyn, A.; Niedzialkowska, E.; Mordalski, S.; Staroń, J.; Bojarski, A.; Szaleniec, M. Reaction mechanism of sterol hydroxylation by steroid C25 dehydrogenase - Homology model, reactivity and isoenzymatic diversity. *J. Inorg. Biochem.* **2017**, *173*, 28−43.

(63) Melander, L. Evaluation of Rate-Constant Ratios from Experimental Data. *Isotope Effect on Reaction Rates*; The Ronald Press Company: New York, 1960; pp 46−64.

(64) Pérez Galende, P.; Hidalgo Cuadrado, N.; Kostetsky, E. Y.; Roig, M. G.; Villar, E.; Shnyrov, V. L.; Kennedy, J. F. Kinetics of Spanish broom peroxidase obeys a Ping-Pong Bi-Bi mechanism with competitive inhibition by substrates. *Int. J. Biol. Macromol.* **2015**, *81*, 1005−1011.

164

(65) Singh, D.; Chaudhury, S. Single-Molecule Kinetics of an Enzyme in the Presence of Multiple Substrates. *ChemBioChem* **2018**, *19*, 842−850.

(66) Cleland, W. W. The Kinetics of Enzyme-Catalyzed Reactions with Two or More Substrates or Products. *Biochim. Biophys. Acta, Spec. Sect. Enzymol. Subj.* **1963**, *67*, 188−196.

(67) Burnham, K. P.; Anderson, D. R. Multimodel Inference: Understanding AIC and BIC in Model Selection. *Socio. Methods Res.* **2004**, *33*, 261−304.

(68) Gadda, G.; Sobrado, P. Kinetic Solvent Viscosity Effects as Probes for Studying the Mechanisms of Enzyme Action. *Biochemistry* **2018**, *57*, 3445−3453.

(69) Valley, M. P.; Fitzpatrick, P. F. Reductive Half-Reaction of Nitroalkane Oxidase: Effect of Mutation of the Active Site Aspartate to Glutamate†,‡. *Biochemistry* **2003**, *42*, 5850−5856.

(70) Fafarman, A. T.; Sigala, P. A.; Schwans, J. P.; Fenn, T. D.; Herschlag, D.; Boxer, S. G. Quantitative, Directional Measurement of Electric Field Heterogeneity in the Active Site of Ketosteroid Isomerase. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, E299−E308.

Enzymatic Δ^1-Dehydrogenation of 3-Ketosteroids – Reconciliation of Kinetic Isotope Effects with the Reaction Mechanism

Michał Glanowski[1], Patrycja Wójcik[1], Magdalena Procner[1], Tomasz Borowski[1], Dawid Lupa[1], Przemysław Mielczarek[2], Maria Oszajca[3], Katarzyna Świderek[4], Vicent Moliner[4], Andrzej J. Bojarski[2], Maciej Szaleniec[1*]

[1]*Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, Niezapominajek 8, 30-239 Krakow, Poland, [2]Maj Institute of Pharmacology, Polish Academy of Sciences, Smetna 12, 31-343 Krakow, [3]Faculty of Chemistry, Jagiellonian University, Gronostajowa 2, 30-387 Krakow, Poland, [4]Department of Physical and Analytical Chemistry, Universitat Jaume I, 12071 Castellon ,Spain, Poland   e-mail: *maciej.szaleniec@ikifp.edu.pl*

Contents

1 Detailed description of main pathway – deprotonated Tyr318
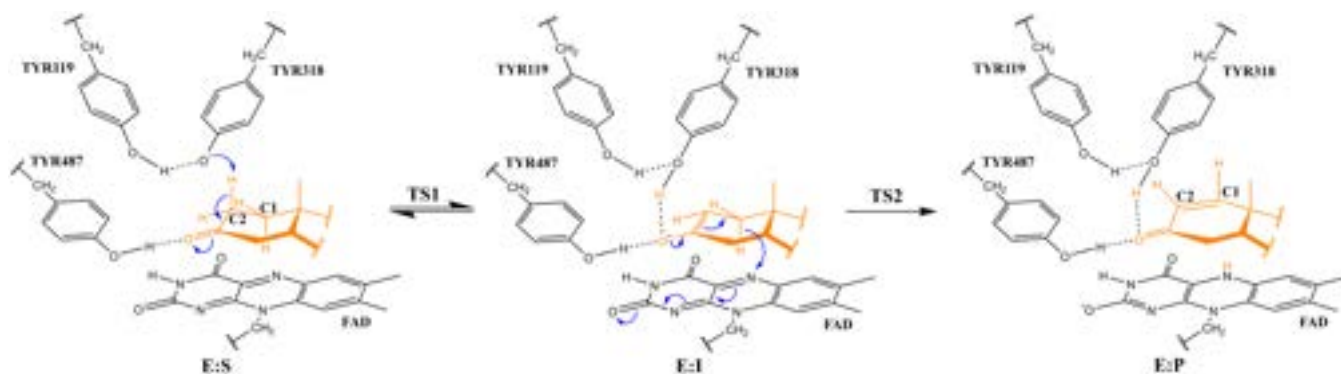

Figure S1: Scheme of main pathway mechanism

The substrate (DHT) is located in the active site pocket near the isoalloxazine ring system of FAD. The ring A of the substrate is almost parallel to the keto-pyrimidine moiety of the FAD ring system, the 3-keto group of substrate lies above the N3 atom of the FAD. DHTs position is stabilized by H-bonds between the 3-keto group and Tyr487 (d(O-H$_{Tyr487}$...O$_{ket}$) = 1.99 Å ± 0.09) as well as Gly491 (d(NH$_{Gly491}$...O$_{ket}$) = 2.25 Å ± 0.02). The initial distance between the Tyr318 anionic oxygen atom (OH$_{Tyr318}$) and the $2\beta H$ atom of the substrate is 1.95 Å ± 0.07. The position of Tyr318 is stabilized by H-bond interactions with Tyr119, with the distance between the hydrogen atom of Tyr119 hydroxyl group (HH$_{Tyr119}$) and (OH$_{Tyr318}$) 1.92 Å ± 0.10. The H-bond between Tyr119 and Tyr318 remains stable throughout the whole RHR.

In the first step of the reaction, the steroid is activated by the abstraction of the proton from the 2β position by the Tyr318 anion. In the transition state **TS1**, the distance between C2 and 2β-H atom is 1.46 Å ± 0.01 for DHT while the distance between 2β-H and OH atom of Tyr318 is 1.17 Å ± 0.01 (see Fig. 7 from the main text). Interestingly, there is no significant shift in the position of the proton H-bond stabilizing the 3-keto group (i.e. with Tyr487) from **E:S** to **TS1**, as the d(O-H$_{Tyr487}$...O$_{ket}$) = 1.92 ± Å 0.06). However, the transfer of a proton from the substrate toward a tyrosyl anion induces a shift of a negative charge in the system with the accumulation of the negative charge at the oxygen atom in the 3-keto group (change of charge from -0.506 to -0.628). Upon completion of the proton transfer to Tyr318, the intermediate enolate structure is formed (**E:I**) and the H-bond between the 3-keto group and Tyr487 indeed gets stronger (d(O-H$_{Tyr487}$...O$_{ket}$) = 1.82 Å ± 0.03. Additionally, a new H-bond is formed between the 3-keto group and protonated with 2β -H atom Tyr318 (d(O-H$_{Tyr318}$...O$_{ket}$) = 1.76 Å ± 0.05. Formation of intermediate enolate results with a change in ring A geometry and repositioning of the substrate in the active site which decreases the distance between 1H of DHT and N5$_{FAD}$ atom is 2.62 Å ± 0.18 in **E:S** to 1.96 Å ± 0.003 in **E:I**. This change of geometry facilitates the next step of the reaction. In the second step of the reaction, the hydride is transferred from the 1 position of ketosteroid to N5 atom in FAD residue, resulting in a 2-electron reduced (hydroquinone) unprotonated state of flavin. In the transition state, **TS2**, the distance between H1 of a substrate and C1 atom of the substrate is 1.38 Å ± 0.01 and between 1H and N5$_{FAD}$ is 1.33 Å ± 0.01. Finally, the double bond is formed between C1 and C2 atoms of steroid and the enzyme-product is formed (E:P, Figure 6 right). At this stage of reaction, FAD is at a two-electron reduced (hydroquinone), deprotonated state.

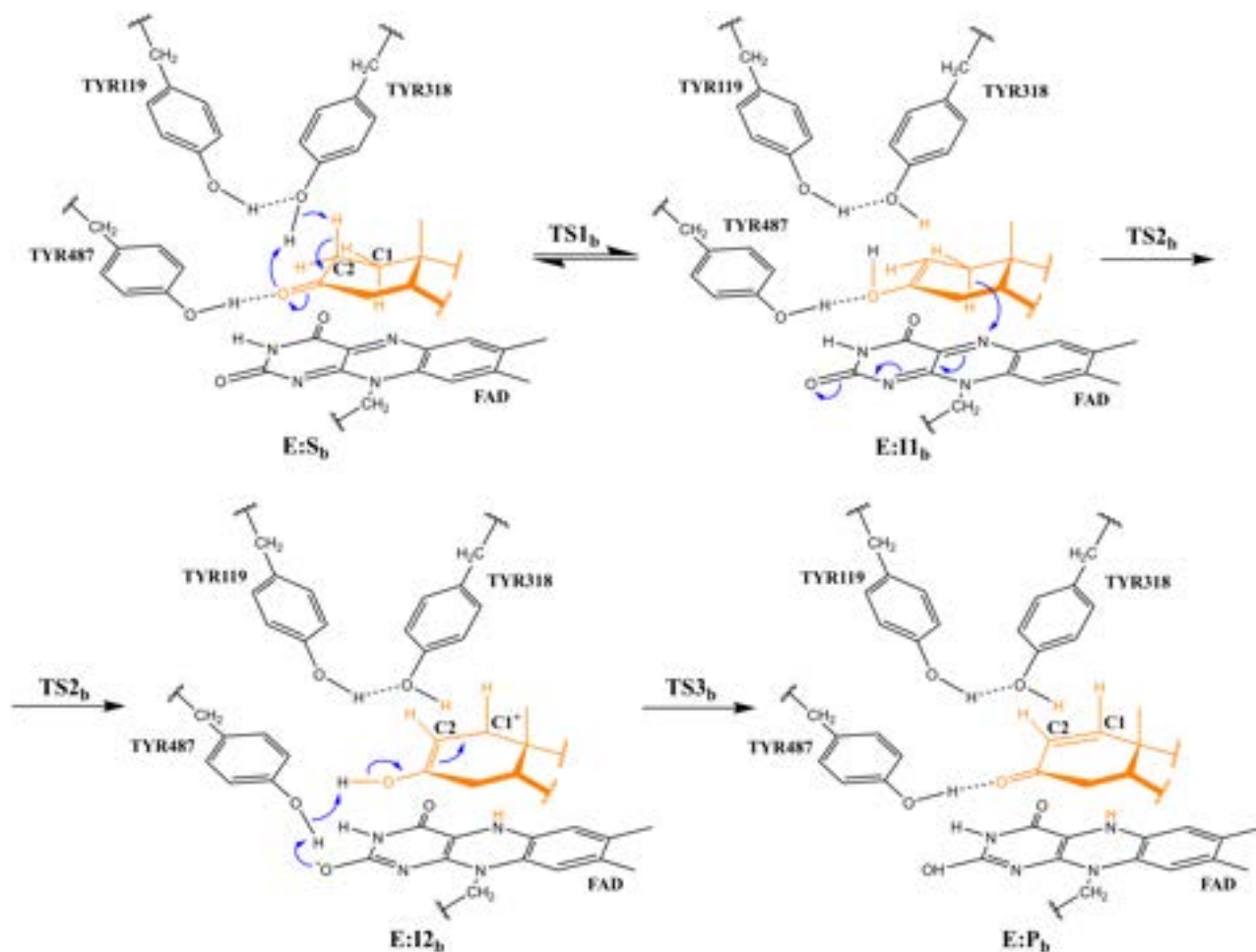## 2 Detailed description of alternative mechanism – protonated Tyr318

Figure S2: Scheme of alternative mechanism

In comparison to **E:S** from main pathway, no significant changes in the active site of the optimized structure of **E:Sb** were noted, except for the state of Tyr318. The distance between the oxygen atom (OH_{Tyr318}) of Tyr318 and the hydrogen atom of Tyr119 hydroxyl group (HH_{Tyr119}) is 1.98 Å. In the first step, activation of C2 atom (**TS1b**) occurs via a double hydrogen shift i.e., we observe a transfer of the 2β proton to Tyr318 with a concomitant shift of HH_{Tyr318} proton to 3-keto group of the substrate. In the optimized **TS1b**, the distance $d(C2_{DHT}\ldots 2\beta H)$ is 1.34 Å and $d(2\beta H\ldots O_{Tyr318})$ = 1.33 Å, while $d(O_{Tyr318}\ldots H_{Tyr318})$ = 1.32 Å and $d(H_{Tyr318}\ldots O_{ket})$ = 1.13 Å. The obtained enol intermediate (**E:Ib**) is stabilized by hydrogen bonds with Tyr318 and Tyr487, similarly to enolate **E:I** in the main pathway. Distance between hydrogen and oxygen atoms in the enol group is 0.982 Å. In the optimized structure of **TS1b**, the hydrogen bond between substrate and Tyr318 is formed by hydrogen from the enol group, while the hydrogen atom bonded to O_{Tyr318} is not directed to steroid. In the second step, the hydride is transferred from 1α position of ketosteroid to N5 atom in FAD residue and an unprotonated hydroquinone state of flavin is formed. The distance between C1 and 1H is 1.63 Å and $N5_{FAD}$ atom is 1.24 Å away from the transferred H atom in **TS2b**. In the last step, there is another concomitant shift of two protons i.e, a proton from the enol group is transferred to Tyr487, and the proton from the OH group of Tyr487 is transferred to the $N1_{FAD}$. In the optimized geometry of **TS3b**, the hydrogen atom from the enol group is 1.30 Å away from Tyr487 OH atom and 1.15 from the oxygen atom of the substrate. The distance between HH_{Tyr487} and OH_{Tyr487} is 1.21. As a result, the reduced and protonated (i.e. uncharged) form of FADH2 is

obtained. Calculations showed that the alternative mechanism is extremely improbable, even the energy of enol intermediate is over 27 kcal/mol higher relative to the substrate. Therefore, we did not perform further analysis as this pathway is thermodynamically improbable.
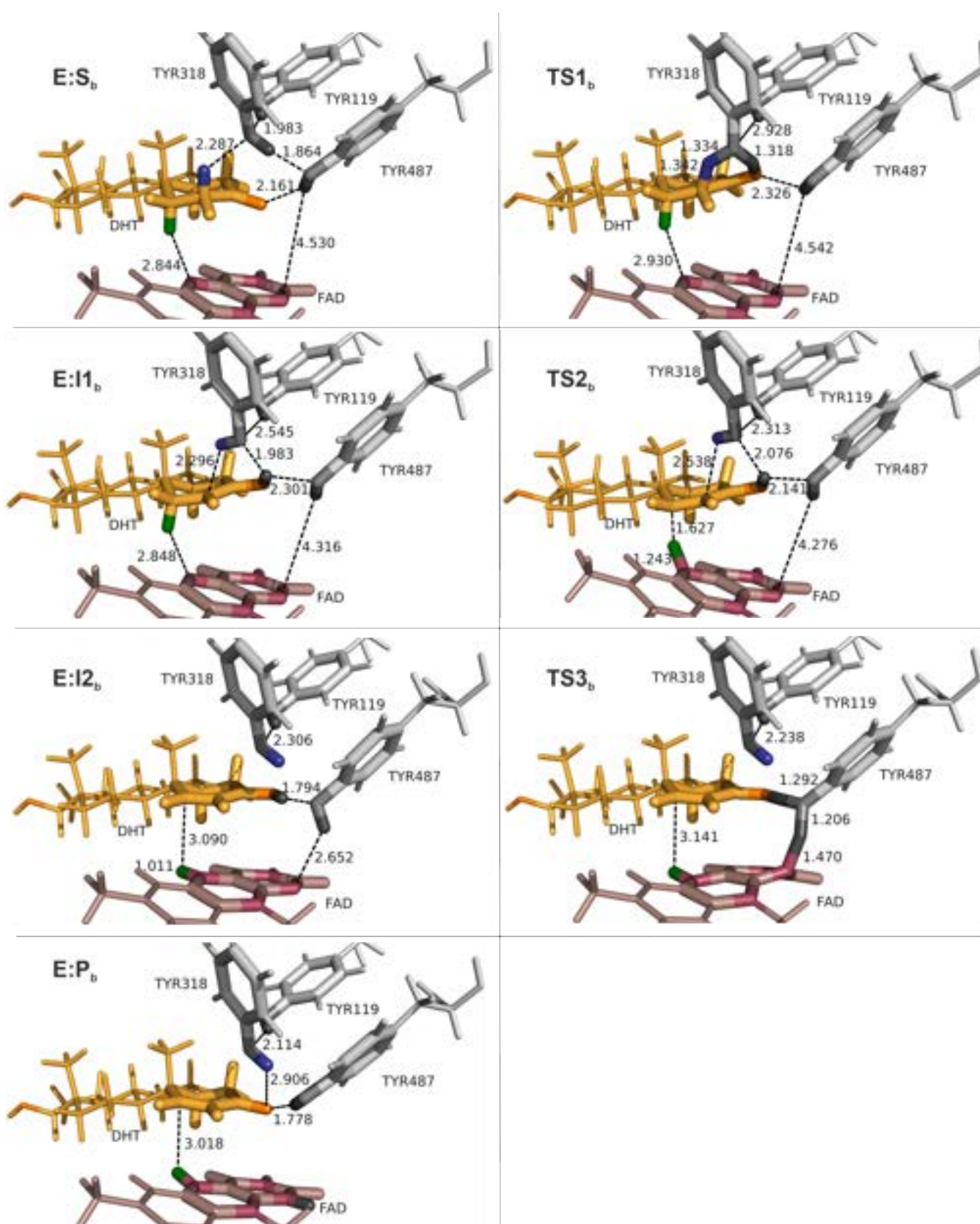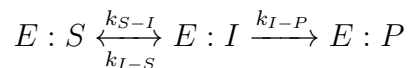

Figure S3: Structures of the stationary geometries along the alternative path. The distances are given in Å.

3   Derivation of relationship between KIE$_2$ and intrinsic KIEs

For two step reaction with reversible first step:

$$E : S \xrightleftharpoons[k_{I-S}]{k_{S-I}} E : I \xrightarrow{k_{I-P}} E : P$$

we can write rate constant for the overall process:

$$k_2 = \frac{k_{S-I}k_{I-P}}{k_{S-I} + k_{I-S} + k_{I-P}}$$

For ratio between rate constant for reactants not isotopically substituted and substituted with deuterium, we use notation:

$$KIE_i = \frac{k_i^H}{k_i^D}$$

Kinetic isotope effect of the overall process can be expressed as follows:

$$
\begin{aligned}
KIE_2 &= \frac{k_2^H}{k_2^D} \\[2mm]
&= \frac{k_{S-I}^H k_{I-P}^H}{k_{S-I}^H + k_{I-S}^H + k_{I-P}^H} \cdot \frac{k_{S-I}^D + k_{I-S}^D + k_{I-P}^D}{k_{S-I}^D k_{I-P}^D} \\[2mm]
&= KIE_{S-I} \cdot KIE_{I-P} \cdot \frac{k_{S-I}^D + k_{I-S}^D + k_{I-P}^D}{k_{S-I}^H + k_{I-S}^H + k_{I-P}^H}
\end{aligned}
$$

Such obtained equation can be further modified to obtain:

$$
\begin{aligned}
KIE_2 &= KIE_{S-I} \cdot KIE_{I-P} \cdot \frac{\dfrac{k_{S-I}^H}{KIE_{S-I}} + \dfrac{k_{I-S}^H}{KIE_{I-S}} + \dfrac{k_{I-P}^H}{KIE_{I-P}}}{k_{S-I}^H + k_{I-S}^H + k_{I-P}^H} \\[3mm]
&= KIE_{S-I} \cdot KIE_{I-P} \cdot \frac{\dfrac{1}{KIE_{S-I}} + \dfrac{1}{KIE_{I-S}} \dfrac{k_{I-S}^H}{k_{S-I}^H} + \dfrac{1}{KIE_{I-P}} \dfrac{k_{I-P}^H}{k_{S-I}^H}}{1 + \dfrac{k_{I-S}^H}{k_{S-I}^H} + \dfrac{k_{I-P}^H}{k_{S-I}^H}}
\end{aligned}
$$

4 Methodology details

4.1 Prediction of titrable aminoacids pKa

Table S 1: Predicted pKa shifts for KSTD from *Rhodococcus erythropolis*

Residue	pKa$^{\text{propka}}$	pKa$^{\text{H++}}$	Residue	pKa$^{\text{propka}}$	pKa$^{\text{H++}}$	Residue	pKa$^{\text{propka}}$	pKa$^{\text{H++}}$
GLU-7	4.04	4.182	ASP-156	3.57	<0.000	ARG-344	12.47	>12.000
CYS-8	12.04	>12.000	ARG-157	10.86	>12.000	GLU-345	4.05	3.613
ASP-9	2.89	2.162	ASP-161	2.77	2.535	CYS-353	11.83	>12.000
TYR-24	10.45	>12.000	HIS-162	5.52	5.894	LYS-361	10.72	11.133
GLU-37	3.15	<0.000	ARG-171	10.56	>12.000	HIS-362	4.3	6.239
LYS-38	10.13	>12.000	ARG-176	11.28	>12.000	GLU-364	4.38	4.193
ASP-40	3.81	2.807	LYS-186	10.38	10.803	ASP-372	3.93	3.702
ARG-41	12.89	>12.000	GLU-188	4.05	3.245	GLU-375	2.57	2.631
TYR-48	13.05	>12.000	ARG-190	12.74	>12.000	GLU-376	4.79	4.155
GLU-62	4.73	4.425	GLU-192	4.87	4.405	LYS-380	10.45	11.412
ARG-63	13.98	>12.000	GLU-201	4.67	4.575	ASP-386	3.52	3.992
ASP-68	2.55	0.301	ASP-202	3.91	3.704	ARG-389	14.35	>12.000
GLU-71	3.27	3.178	ARG-204	13.05	>12.000	GLU-393	4.68	4.562
ARG-74	15.6	>12.000	GLU-209	4.79	3.967	LYS-394	11.38	>12.000
TYR-76	13.96	>12.000	GLU-211	4.2	5.185	ASP-397	3.95	3.802
ARG-78	13.15	>12.000	GLU-215	4.65	4.727	LYS-400	10.55	11.122
ASP-83	3.87	3.942	ARG-218	12.34	>12.000	ASP-404	1.74	<0.000
GLU-85	4.06	2.323	LYS-220	10.54	>12.000	GLU-405	4.6	4.454
GLU-87	4.22	4.170	ARG-223	12.6	>12.000	GLU-406	4.02	3.599
ARG-88	11.39	>12.000	GLU-233	6.22	<0.000	HIS-408	6.84	8.407
ASP-90	2.39	0.691	GLU-237	4.28	5.044	ARG-409	11.51	>12.000
TYR-92	14.51	>12.000	ARG-239	12.77	>12.000	GLU-411	4.8	5.387
GLU-94	4.13	3.500	GLU-240	4.59	4.454	ASP-412	2.87	<0.000
GLU-104	2.84	1.269	LYS-247	10.14	11.241	TYR-414	14.02	>12.000
GLU-110	4.75	5.039	ASP-261	3.07	1.728	ASP-415	5.68	3.872
GLU-112	4.21	3.234	ASP-277	3.62	3.029	CYS-419	10.57	>12.000
ARG-114	13.08	>12.000	CYS-282	13.47	>12.000	GLU-433	4.81	4.561
ASP-118	5.33	<0.000	GLU-286	4.64	3.127	TYR-438	10.14	>12.000
TYR-119	17.57	>12.000	ASP-289	2.81	2.542	ARG-441	12.83	>12.000
TYR-120	12.7	>12.000	ARG-299	13.5	>12.000	ASP-446	4.11	<0.000
LYS-121	10.57	11.446	ASP-305	3.36	3.248	LYS-450	6.49	>12.000
GLU-123	4.68	4.546	GLU-309	5.2	6.151	ASP-456	3.64	0.561
ARG-125	13.69	>12.000	ARG-310	15.29	>12.000	ARG-460	13.64	>12.000
ASP-127	4.0	3.142	TYR-311	12.66	>12.000	ARG-463	12.33	>12.000
ARG-130	14.06	>12.000	GLU-314	3.46	<0.000	ASP-465	2.19	2.823
ASP-136	3.85	2.971	TYR-318	22.84	>12.000	ASP-470	3.04	2.862
ASP-138	3.43	3.319	ASP-319	4.46	<0.000	TYR-473	13.47	>12.000
ASP-141	4.34	4.666	ARG-323	13.07	>12.000	ARG-485	12.34	>12.000
ASP-144	3.25	3.510	ASP-326	4.94	4.012	TYR-487	14.21	>12.000
LYS-148	10.64	11.396	HIS-328	6.97	9.197	TYR-502	10.93	11.684
ARG-150	13.86	>12.000	ASP-329	3.39	2.427	ARG-503	11.77	>12.000
GLU-152	4.99	1.029	ASP-330	4.01	4.131	ASP-507	3.93	2.678
ASP-154	4.31	3.723	ASP-342	3.11	0.837	LYS-510	10.63	10.246

4.2   MM parameters of non-standard residues

Table S 2: Atom types and charges of 17-MT

Atom name	x coord	y coord	z coord	AMBER atom type	charge
C1	-47.4343	-13.4170	39.4166	c3	0.368358
C2	-48.5942	-14.3963	39.8152	c3	-0.183339
C3	-49.8146	-14.0734	38.9085	c3	-0.140011
C4	-49.2680	-13.0592	37.8837	c3	-0.025215
C5	-48.2006	-12.2597	38.6688	c3	0.182344
C6	-48.8245	-11.3306	39.7367	c3	-0.319715
C7	-47.3961	-11.4189	37.6641	c3	-0.014596
C8	-48.3175	-10.5081	36.8245	c3	-0.173206
C9	-49.4559	-11.2796	36.1126	c3	0.004675
C10	-50.2533	-12.1848	37.0952	c3	-0.030842
C11	-51.2714	-13.0327	36.3213	c3	-0.073365
C12	-52.2112	-12.1563	35.4819	c3	-0.037738
C13	-51.4624	-11.2117	34.5726	c2	0.057068
C14	-51.7908	-11.1264	33.2667	ce	-0.411307
C15	-51.2206	-10.1454	32.3280	c	0.625409
O1	-51.5354	-10.1236	31.1456	o	-0.505666
C16	-50.2722	-9.1271	32.9363	c3	-0.307608
C17	-49.4832	-9.7175	34.1084	c3	0.042530
C18	-50.3655	-10.3549	35.2152	c3	0.137544
C19	-51.0495	-9.2359	36.0451	c3	-0.299956
H1	-48.2696	-15.4358	39.7115	hc	0.054870
H2	-48.8397	-14.2609	40.8736	hc	0.054870
H3	-50.2233	-14.9676	38.4298	hc	0.053253
H4	-50.6301	-13.6302	39.4925	hc	0.053253
H5	-48.7185	-13.6405	37.1267	hc	0.032234
H6	-48.0339	-10.7840	40.2605	hc	0.070106
H7	-49.4081	-11.8793	40.4829	hc	0.070106
H8	-49.4963	-10.5877	39.3026	hc	0.070106
H9	-46.6423	-10.8108	38.1777	hc	0.008127
H10	-46.8529	-12.0826	36.9826	hc	0.008127
H11	-48.7401	-9.7272	37.4655	hc	0.058512
H12	-47.7038	-9.9880	36.0824	hc	0.058512
H13	-48.9694	-11.9739	35.4081	hc	0.010677
H14	-50.8109	-11.5535	37.8019	hc	0.044180
H15	-50.7343	-13.7277	35.6599	hc	0.032561
H16	-51.8622	-13.6463	37.0108	hc	0.032561
H17	-52.8980	-12.7709	34.8930	hc	0.041323
H18	-52.8361	-11.5619	36.1658	hc	0.041323
H19	-52.5604	-11.7719	32.8485	ha	0.151083
H20	-49.6098	-8.7534	32.1498	hc	0.090629
H21	-48.8102	-10.4911	33.7152	hc	0.012863
H22	-51.7241	-8.6456	35.4177	hc	0.081255
H23	-50.3087	-8.5522	36.4692	hc	0.081255
H24	-51.6441	-9.6373	36.8705	hc	0.081255
H25	-50.8786	-8.2713	33.2661	hc	0.090629
H26	-48.8444	-8.9427	34.5448	hc	0.012863
O2	-46.6867	-12.9675	40.5494	oh	-0.621754
H27	-47.3096	-12.6419	41.2115	ho	0.378715
C20	-46.3815	-14.1265	38.5597	c3	-0.265460
H28	-45.5578	-13.4502	38.3200	hc	0.072201
H29	-46.7975	-14.5195	37.6280	hc	0.072201
H30	-45.9695	-14.9651	39.1288	hc	0.072201

Table S 3: MM parameters for 17MT

| Nonbon			Bonds			Angles			Dihedrals				Dihedrals			
c	1.908	0.086	c2-ce	547.3	1.346	c2-c3-c3	63.4	111.56	c2-ce-c-o	2.175	180.0	2	oh-c3-c3-hc	0.25	0.0	1
c2	1.908	0.086	c2-c3	326.8	1.51	c2-ce-ha	49.6	119.94	c3-c3-c3-oh	0.156	0.0	3	c3-c3-oh-ho	0.16	0.0	3
ce	1.908	0.086	c3-c2	326.8	1.51	o-c-c3	67.4	123.2	o-c-c3-hc	0.8	0.0	1	ho-oh-c3-c3	0.25	0.0	1
hc	1.487	0.0157	oh-c3	316.7	1.423	c2-c3-hc	47.0	110.36	c-c3-c3-c3	0.156	0.0	3	c3-c2-ce-c	6.65	180.0	2
o	1.6612	0.21	c-o	637.7	1.218	oh-c3-c3	67.5	110.19	ho-oh-c3-c3	0.16	0.0	3	oh-c3-c3-hc	0.0	0.0	3
ha	1.459	0.015	c3-hc	330.6	1.097	c3-c2-ce	64.0	123.15	oh-oh-c3-c3	0.144	0.0	3	c3-c3-c2-c3	0.0	0.0	2
c3	1.908	0.1094	oh-ho	371.4	0.973	c3-c3-oh	67.5	110.19	hc-c3-c3-c3	0.16	0.0	3	c3-c3-c3-c3	0.2	180.0	1
ho	0	0	c3-c3	300.9	1.538	ce-c-c3	62.9	116.44	c3-c3-c3-hc	0.16	0.0	3	o-c-ce-ha	2.175	180.0	2
oh	1.721	0.2104	ce-c	354.5	1.482	ho-oh-c3	47.4	107.26	o-c-c3-hc	0.0	0.0	3	ce-c2-c3-c3	0.0	0.0	2
			ce-ha	342.5	1.088	c2-ce-c	65.5	120.42	c2-c3-c3-c3	0.156	0.0	3	c3-c2-c3-c3	0.0	0.0	2
			c3-oh	316.7	1.423	c3-c2-c3	62.9	115.65	o-c-c3-hc	0.08	180.0	3	c3-c3-c3-c3	0.25	180.0	2
			c-c3	313.0	1.524	ce-c2-c3	64.0	123.15	ce-c-c3-hc	0.0	180.0	2	c-ce-c2-c3	6.65	180.0	2
						c3-oh-ho	47.4	107.26	hc-c3-c3-oh	0.25	0.0	1	oh-c3-c3-c3	0.156	0.0	3
						c3-c3-c2	63.4	111.56	c3-c3-c3-c2	0.156	0.0	3	hc-c3-c3-hc	0.15	0.0	3
						c-c3-c3	63.3	111.04	hc-c3-c3-oh	0.0	0.0	3	c2-c3-c3-hc	0.156	0.0	3
						ce-c-o	68.8	123.2	c-c3-c3-hc	0.156	0.0	3	ce-c-c3-c3	0.0	180.0	2
						c-ce-ha	46.5	116.46	c3-c3-c3-c3	0.18	0.0	3	c3-c3-c2-ce	0.0	0.0	2
						c-c3-hc	46.9	108.77	c3-c-ce-ha	2.175	180.0	2	c2-ce-c-c3	2.175	180.0	2
						c3-c3-hc	46.3	109.8	oh-c3-c3-oh	1.175	0.0	2	c3-c3-oh-ho	0.25	0.0	1
						hc-c3-hc	39.4	107.58	ce-c2-c3-hc	0.0	0.0	2	Impropers			
						c3-c3-c3	62.9	111.51	c3-c2-c3-hc	0.0	0.0	2	c3-ce-c-o	10.5	180.0	2
									c3-c2-ce-ha	6.65	180.0	2	c-c2-ce-ha	1.1	180.0	2
													c3-c3-c2-ce	1.1	180.0	2
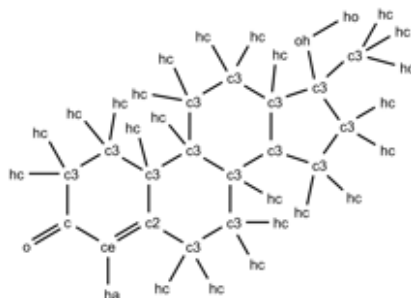
Figure S4: 17MT structure and atom types.

Table S 4: Atom types and charges of DHT

Atom name	x coord	y coord	z coord	AMBER atom type	charge
C1	74.1034	31.0552	43.1953	c3	0.257735
C2	74.8885	31.1937	44.5327	c3	-0.254944
C3	74.3486	32.4806	45.2228	c3	-0.029405
C4	73.1712	32.9164	44.3234	c3	0.010159
C5	73.5995	32.4935	42.8989	c3	0.123720
C6	74.7636	33.3293	42.3274	c3	-0.335742
C7	72.3719	32.5859	41.9817	c3	-0.062103
C8	71.7424	33.9931	42.0188	c3	-0.145315
C9	71.4055	34.4820	43.4503	c3	0.032464
C10	72.6273	34.3487	44.4112	c3	0.019151
C11	72.2449	34.7266	45.8514	c3	-0.075114
C12	71.5710	36.0990	45.9408	c3	-0.120649
C13	70.3647	36.1842	44.9968	c3	0.180131
C14	69.5793	37.4994	45.1877	c3	-0.305659
C15	68.3456	37.5683	44.3013	c	0.542985
O1	67.2575	37.9124	44.7225	o	-0.483056
C16	68.5716	37.1774	42.8490	c3	-0.216899
C17	69.4164	35.8951	42.6953	c3	0.038091
C18	70.7402	35.9034	43.5060	c3	0.131804
C19	71.6848	36.9841	42.9314	c3	-0.319106
H1	74.7684	30.3014	45.1539	hc	0.072301
H2	75.9609	31.2767	44.3201	hc	0.072301
H3	74.0350	32.2988	46.2539	hc	0.025273
H4	75.1206	33.2578	45.2610	hc	0.025273
H5	72.3262	32.2568	44.5848	hc	-0.006993
H6	75.0555	32.9313	41.3496	hc	0.081655
H7	75.6470	33.3165	42.9739	hc	0.081655
H8	74.4944	34.3774	42.1794	hc	0.081655
H9	72.6326	32.3193	40.9504	hc	0.021437
H10	71.6272	31.8489	42.3139	hc	0.021437
H11	72.4202	34.7039	41.5323	hc	0.054212
H12	70.8334	33.9820	41.4093	hc	0.054212
H13	70.6446	33.7864	43.8432	hc	-0.015720
H14	73.4176	35.0396	44.0825	hc	0.018803
H15	71.5561	33.9643	46.2467	hc	0.024321
H16	73.1334	34.7000	46.4936	hc	0.024321
H17	71.2421	36.2890	46.9703	hc	0.035255
H18	72.2919	36.8916	45.7004	hc	0.035255
H19	67.5970	37.0738	42.3641	hc	0.064777
H20	68.8131	35.0366	43.0215	hc	-0.002658
H21	71.2707	37.9907	43.0336	hc	0.073713
H22	71.8575	36.8206	41.8632	hc	0.073713
H23	72.6607	36.9897	43.4237	hc	0.073713
H24	69.0843	38.0173	42.3616	hc	0.064777
H25	69.6197	35.7422	41.6315	hc	-0.002658
H26	73.2249	30.4171	43.3516	h1	0.052841
O2	74.8201	30.4116	42.1534	oh	-0.620352
H27	75.6768	30.8500	42.0700	ho	0.387925
H28	70.2208	38.3595	44.9478	hc	0.077459
H29	69.2591	37.6179	46.2266	hc	0.077459
H30	69.6805	35.3710	45.2885	hc	-0.015611

9

Table S 5: MM parameters for DHT

	Nonbon			Bonds			Angles			Dihedrals				Dihedrals		
oh	1.721	0.2104	oh-c3	316.7	1.423	c3-c3-c3	62.9	111.51	c3-c3-c3-c3	0.2	180.0	1	o-c-c3-hc	0.0	0.0	2
ho	0	0	c3-c3	300.9	1.538	c-c3-hc	46.9	108.77	hc-c3-c3-oh	0.0	0.0	3	ho-oh-c3-c3	0.16	0.0	3
c3	1.908	0.1094	oh-ho	371.4	0.973	c3-c3-h1	46.4	109.56	c3-c3-c3-c3	0.18	0.0	3	o-c-c3-hc	0.8	0.0	1
h1	1.387	0.0157	c-c3	313.0	1.524	c3-c3-c	63.3	111.04	oh-c3-c3-oh	1.175	0.0	2	c3-c3-c3-c	0.156	0.0	3
o	1.6612	0.21	c3-hc	330.6	1.097	c3-c-o	67.4	123.2	c3-c3-oh-ho	0.16	0.0	3	oh-c3-c3-h1	0.0	0.0	3
c	1.908	0.086	c3-oh	316.7	1.423	o-c-c3	67.4	123.2	o-c-c3-c3	0.0	180.0	2	c3-c3-c-c3	0.0	180.0	2
hc	1.487	0.0157	c3-h1	330.6	1.097	c3-c3-hc	46.3	109.8	c3-c3-c3-hc	0.16	0.0	3	oh-c3-c3-h1	0.25	0.0	1
			c-o	637.7	1.218	c3-c-c3	62.0	116.5	h1-c3-oh-ho	0.167	0.0	3	oh-c3-c3-oh	0.144	0.0	3
			c3-c	313.0	1.524	c3-oh-ho	47.4	107.26	ho-oh-c3-c3	0.25	0.0	1	oh-c3-c3-c3	0.156	0.0	3
						ho-oh-c3	47.4	107.26	hc-c3-c3-hc	0.15	0.0	3	c3-c-c3-c3	0.0	180.0	2
						h1-c3-oh	50.9	110.26	o-c-c3-hc	0.08	180.0	3	c-c3-c3-hc	0.156	0.0	3
						oh-c3-h1	50.9	110.26	h1-c3-c3-c3	0.156	0.0	3	h1-c3-c3-oh	0.0	0.0	3
						h1-c3-h1	39.2	108.46	c3-c3-c3-h1	0.156	0.0	3	h1-c3-c3-oh	0.25	0.0	1
						c3-c3-oh	67.5	110.19	h1-c3-c3-h1	0.156	0.0	3	hc-c3-c3-h1	0.156	0.0	3
						hc-c3-hc	39.4	107.58	c-c3-c3-c3	0.156	0.0	3	c3-c3-c3-c3	0.25	180.0	2
						c-c3-c3	63.3	111.04	ho-oh-c3-h1	0.167	0.0	3	hc-c3-c3-oh	0.25	0.0	1
						h1-c3-c3	46.4	109.56	c3-c3-c3-oh	0.156	0.0	3	c3-c-c3-hc	0.0	180.0	2
						oh-c3-c3	67.5	110.19	c3-c3-c-o	0.0	180.0	2	Impropers			
									c3-c3-oh-ho	0.25	0.0	1	c3-c3-c-o	1.1	180.0	2
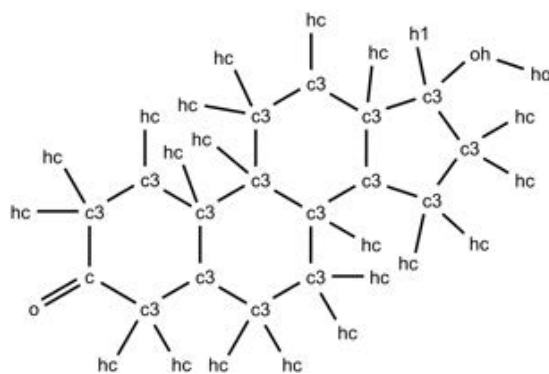
Figure S5: DHT structure and atom types.

Table S 6: Atom types and charges of FAD

Atom name	x coord	y coord	z coord	AMBER atom type	charge
O3'	-36.6980	-10.8920	22.6570	oh	-0.597900
H3T	-37.3610	-11.5690	23.1100	ho	0.422500
C1'	-36.6480	-9.3810	20.0630	c3	0.100100
H1'	-35.9420	-10.1960	19.9040	h2	0.143800
C2'	-36.5040	-8.8850	21.4370	c3	0.142600
H2'	-36.7910	-7.8370	21.5270	h1	0.052500
C3'	-37.4000	-9.7810	22.2180	c3	0.118000
H3'	-37.8690	-9.2520	23.0480	h1	0.057500
C4'	-38.4280	-10.2690	21.2200	c3	0.148400
H4'	-38.5150	-11.3540	21.2720	h1	0.074400
C5'	-39.8050	-9.6090	21.5030	c3	0.107000
H5'1	-40.3490	-9.4620	20.5700	h1	0.041700
H5'2	-39.6650	-8.6480	21.9970	h1	0.041700
O4'	-37.9530	-9.8760	19.9330	os	-0.362800
O2'	-35.1140	-9.1130	21.8370	oh	-0.606900
H2T	-35.0060	-8.9080	22.8610	ho	0.431500
N6	-34.7420	-6.4690	15.4300	nh	-0.728900
H61	-35.1530	-5.5630	15.8230	hn	0.363200
H62	-34.2240	-6.4600	14.4940	hn	0.363200
C6	-34.8730	-7.5930	16.0970	ca	0.569800
C5	-35.5400	-7.5170	17.2690	ca	0.061500
N7	-36.1390	-6.5030	17.9290	nc	-0.602900
C8	-36.6560	-6.9930	19.0500	cd	0.158100
H8	-37.2100	-6.4320	19.8030	h5	0.145100
N9	-36.3720	-8.3080	19.0870	na	-0.029100
C4	-35.6950	-8.6200	17.9920	ca	0.460700
N3	-35.2050	-9.7940	17.5600	nb	-0.753200
C2	-34.5270	-9.8970	16.3310	ca	0.550300
H2	-34.1580	-10.8580	15.9730	h5	0.038400
N1	-34.3390	-8.7240	15.5780	nb	-0.734900
O5A	-40.5280	-10.5000	22.3580	os	-0.411100
PA	-41.4820	-9.9730	23.4350	p5	1.040000
OA1	-40.7670	-8.8800	24.1350	o	-0.790800
OA2	-41.8820	-11.1440	24.2370	o	-0.790800
OP	-42.6890	-9.2030	22.5450	os	-0.445800
PB	-44.2010	-9.8180	22.1960	p5	1.040000
OB1	-44.9790	-8.6560	21.7040	o	-0.790800
OB2	-44.2000	-11.0710	21.4410	o	-0.790800
O5B	-44.6640	-10.2210	23.6080	os	-0.411100
C1	-47.0510	-8.4520	28.9110	c3	-0.045300
H11	-48.0490	-8.2150	28.5410	h1	0.116500
H12	-46.5050	-7.5280	29.1020	h1	0.116500
C2	-46.3130	-9.2800	27.8920	c3	0.191800
H2	-46.9650	-10.1040	27.6000	h1	0.057700
O2	-45.1110	-9.8470	28.4390	oh	-0.589400
HO2	-44.4750	-9.0760	28.7620	ho	0.371700
C3	-45.9650	-8.4910	26.6270	c3	0.019700
H3	-45.0420	-7.9270	26.7600	h1	0.124600
O3	-47.0450	-7.6280	26.2630	oh	-0.645000
HO3	-47.9060	-8.2010	26.0810	ho	0.426400
C4	-45.7880	-9.5810	25.6130	c3	0.113600
H4	-46.7560	-9.8560	25.1940	h1	0.134500
O4	-45.2000	-10.7320	26.3010	oh	-0.576000
HO4	-44.2730	-10.4590	26.7110	ho	0.402000
C5	-44.8610	-9.1220	24.5050	c3	-0.036900
H51	-45.3100	-8.2850	23.9710	h1	0.085700
H52	-43.9050	-8.8140	24.9270	h1	0.085700
N1	-48.7040	-10.6800	29.3680	nd	-0.518600

C2	-49.4920	-11.8140	29.6510	c	0.608100
O2	-50.2690	-12.1950	28.7860	o	-0.569600
N3	-49.5000	-12.4960	30.8300	n	-0.321300
H3	-50.1160	-13.2620	30.9300	hn	0.310500
C4	-48.7360	-12.1290	31.7620	c	0.372300
O4	-48.7680	-12.6750	32.8740	o	-0.520700
N10	-47.1450	-9.1440	30.1160	na	0.061400
C10A	-47.9410	-10.2700	30.2670	cc	0.165400
C4A	-47.9170	-10.9880	31.5170	cc	0.437200
N5	-47.2000	-10.6290	32.4720	nd	-0.554900
C5A	-46.3770	-9.6290	32.3860	ca	0.301200
C9A	-46.2580	-8.8720	31.2290	ca	-0.004200
C9	-45.2480	-7.9160	31.2020	ca	-0.295800
H9	-45.0880	-7.3270	30.3390	ha	0.165200
C8	-44.4140	-7.7450	32.3040	ca	0.116400
C7	-44.5950	-8.4810	33.4730	ca	0.065200
C6	-45.5900	-9.4260	33.4990	ca	-0.216000
H6	-45.7320	-10.0370	34.3590	ha	0.158700
C7M	-43.7450	-8.3450	34.7110	c3	-0.227400
H7M1	-43.2030	-7.3960	34.7380	hc	0.084500
H7M2	-43.0160	-9.1780	34.7160	hc	0.084500
H7M3	-44.3800	-8.4740	35.6010	hc	0.084500
C8M	-43.3410	-6.7820	32.2010	c3	-0.208300
H8M1	-43.4030	-6.2230	31.2570	hc	0.081100
H8M2	-42.3860	-7.3210	32.2650	hc	0.081100
H8M3	-43.3810	-6.1090	33.0670	hc	0.081100

Figure S6: FAD structure and atom types.

Table S 7: MM parameters for FAD

Nonbon			Bonds			Angles			Dihedrals				Dihedrals			
nb	1.824	0.17	c-o	637.7	1.218	h1-c3-os	50.8	109.78	c3-os-c3-na	0.383	0.0	3	c3-c3-oh-ho	0.16	0.0	3
nc	1.824	0.17	oh-c3	316.7	1.423	c-n-c	63.7	127.08	os-c3-c3-oh	0.144	0.0	3	ca-nc-cd-h5	4.75	180.0	2
nh	1.824	0.17	c3-os	308.6	1.432	c-cc-cc	63.6	122.69	c3-os-p5-os	1.2	0.0	2	os-c3-c3-oh	1.175	0.0	2
oh	1.721	0.2104	cd-h5	351.8	1.082	ca-ca-c3	63.5	120.77	c3-os-c3-c3	0.1	180.0	2	o-c-cc-cc	2.875	180.0	2
p5	2.1	0.2	ca-ha	345.8	1.086	o-c-cc	69.1	123.93	o-c-cc-nd	2.875	180.0	2	h1-c3-na-cc	0.0	0.0	2
ho	0	0	ca-ca	461.1	1.398	ca-ca-nb	68.8	122.94	os-p5-os-p5	0.8	0.0	2	c3-na-cc-cc	1.7	180.0	2
hc	1.487	0.0157	nd-c	416.9	1.387	h1-c3-h1	39.2	108.46	nh-ca-nb-ca	4.8	180.0	2	na-ca-nb-ca	4.8	180.0	2
na	1.824	0.17	c3-hc	330.6	1.097	h1-c3-na	49.8	108.78	c3-c3-c3-hc	0.16	0.0	3	os-p5-os-c3	0.25	0.0	3
h5	1.359	0.015	c-n	427.6	1.379	ha-ca-ca	48.2	119.88	hn-nh-ca-ca	1.05	180.0	2	h1-c3-oh-ho	0.167	0.0	3
c	1.908	0.086	os-p5	330.6	1.615	nc-ca-ca	69.5	119.72	hc-c3-c3-oh	0.25	0.0	1	nd-cc-cc-nd	4.0	180.0	2
os	1.6837	0.17	cc-nd	525.4	1.317	c-c3-hc	46.9	108.77	h2-c3-na-cd	0.0	0.0	2	hc-c3-c3-hc	0.15	0.0	3
cc	1.908	0.086	nd-cc	525.4	1.317	o-c-c3	67.4	123.2	c3-os-p5-o	0.8	0.0	2	c3-c3-na-cc	0.0	0.0	2
c3	1.908	0.1094	ca-nb	488.0	1.339	p5-os-c3	77.7	119.54	c-nd-cc-na	4.75	180.0	2	c3-c3-c3-h1	0.156	0.0	3
cd	1.908	0.086	c3-oh	316.7	1.423	na-ca-nb	69.8	127.09	na-cc-cc-nd	4.0	180.0	2	hn-nh-ca-nb	1.05	180.0	2
n	1.824	0.17	c3-h1	330.6	1.097	c3-oh-ho	47.4	107.26	p5-os-c3-c3	0.383	0.0	3	ca-ca-ca-ca	3.625	180.0	2
hn	0.6	0.0157	nh-ca	417.9	1.386	nb-ca-h5	51.9	115.82	c3-c3-c3-os	0.156	0.0	3	ca-ca-nc-cd	4.8	180.0	2
ca	1.908	0.086	ca-h5	343.2	1.088	os-p5-os	45.0	101.84	ca-ca-ca-na	3.625	0.0	2	cd-nc-ca-ca	4.8	180.0	2
h1	1.387	0.0157	p5-os	330.6	1.615	c-nd-cc	66.7	120.49	ca-ca-nb-ca	4.8	180.0	2	ca-ca-c3-hc	0.0	0.0	2
nd	1.824	0.17	c3-h2	331.7	1.096	os-c3-h1	50.8	109.78	h1-c3-na-ca	0.0	0.0	2	ha-ca-ca-ca	3.625	180.0	2
o	1.6612	0.21	n-hn	403.2	1.013	cc-na-ca	67.4	113.15	ca-ca-ca-nb	3.625	180.0	2	ha-ca-ca-c3	3.625	180.0	2
ha	1.459	0.015	c3-na	327.7	1.463	nd-c-n	71.6	117.11	c-c3-c3-hc	0.156	0.0	3	n-c-cc-cc	2.875	180.0	2
h2	1.287	0.0157	oh-ho	371.4	0.973	na-cc-cc	68.6	117.77	os-c3-na-cd	0.0	0.0	2	na-ca-ca-ca	3.625	180.0	2
			c-c3	313.0	1.524	nd-cc-cc	71.6	112.56	n-c-cc-nd	2.875	180.0	2	oh-c3-c3-hc	0.25	0.0	1
			ca-nc	467.7	1.352	oh-c3-h1	50.9	110.26	p5-os-c3-h1	0.383	0.0	3	o-c-n-c	2.5	180.0	2
			na-ca	420.5	1.384	n-c-o	74.2	123.05	nd-ca-ca-ha	3.625	180.0	2	cc-na-ca-ca	0.3	180.0	2
			cc-cc	419.8	1.428	ca-c3-hc	46.8	110.47	c3-na-ca-nb	0.3	180.0	2	o-c-c3-c3	0.0	180.0	2
			c-cc	371.0	1.468	hn-n-c	48.3	117.55	nd-c-n-hn	2.5	180.0	2	na-ca-ca-nd	3.625	180.0	2
			os-c3	308.6	1.432	n-c-cc	69.1	112.7	ca-nb-ca-h5	4.8	180.0	2	c3-c3-c3-na	0.156	0.0	3
			c3-c3	300.9	1.538	ca-ca-na	69.1	118.34	oh-c3-c3-hc	0.0	0.0	3	oh-c3-c3-na	0.156	0.0	3
			p5-o	479.5	1.487	c3-c3-h1	46.4	109.56	c3-c3-os-p5	3.95	180.0	1	nc-ca-ca-na	3.625	180.0	2
			n-c	427.6	1.379	h2-c3-na	50.2	107.31	nd-c-n-c	2.5	180.0	2	os-c3-c3-os	1.175	0.0	2
			nc-cd	525.4	1.317	nb-ca-nb	70.9	127.26	oh-c3-c3-oh	0.144	0.0	3	o-c-nd-cc	4.0	180.0	2
			nh-hn	404.6	1.012	nd-ca-ca	69.5	119.72	cc-cc-na-ca	1.7	180.0	2	c3-na-cd-h5	1.7	180.0	2
			cd-na	425.8	1.38	p5-os-p5	98.4	126.25	ca-nc-cd-na	4.75	180.0	2	c-nd-cc-cc	4.75	180.0	2
			ca-c3	321.0	1.516	c3-na-cd	61.9	126.46	c3-os-c3-na	0.65	0.0	2	c-cc-cc-na	4.0	180.0	2
			na-cc	425.8	1.38	nc-cd-na	74.9	112.22	c3-os-c3-c3	0.383	0.0	3	os-c3-na-ca	0.0	0.0	2
			nd-ca	467.7	1.352	h1-c3-c3	46.4	109.56	c3-c3-c3-c3	0.18	0.0	3	o-c-c3-hc	0.08	180.0	3
			nb-ca	488.0	1.339	c-cc-nd	67.6	121.88	o-p5-os-c3	0.8	0.0	2	h1-c3-c3-h1	0.156	0.0	3
						h5-ca-nb	51.9	115.82	h2-c3-c3-h1	0.156	0.0	3	hc-c3-c3-c3	0.16	0.0	3
						oh-c3-c3	67.5	110.19	oh-c3-c3-os	0.144	0.0	3	c3-na-cc-nd	1.7	180.0	2
						na-ca-ca	69.1	118.34	p5-os-p5-o	0.8	0.0	2	nd-ca-ca-ca	3.625	180.0	2

Angle	k	θ
ca-ca-ha	48.2	119.88
c3-c3-na	65.5	112.88
cd-na-ca	67.4	113.15
ca-nb-ca	68.3	117.22
hc-c3-hc	39.4	107.58
nd-c-o	73.9	123.18
os-c3-c3	68.0	107.97
c3-c3-oh	67.5	110.19
ca-nc-cd	72.5	104.88
c3-c3-hc	46.3	109.8
o-c-n	74.2	123.05
c3-na-cc	61.9	126.46
o-p5-os	43.9	115.46
h2-c3-c3	46.2	110.22
hn-nh-ca	48.4	116.07
ho-oh-c3	47.4	107.26
hn-nh-hn	40.1	115.12
c3-os-c3	62.7	112.48
cc-cc-nd	71.6	112.56
c3-c3-c3	62.9	111.51
c-c3-c3	63.3	111.04
ca-ca-ca	66.6	120.02
cc-nd-ca	72.5	104.88
c3-na-ca	62.3	124.36
ca-ca-nc	69.5	119.72
c3-c3-os	68.0	107.97
c3-os-p5	77.7	119.54
nh-ca-nb	72.7	116.94
nh-ca-ca	68.3	120.95
h1-c3-oh	50.9	110.26
c-n-hn	48.3	117.55
o-p5-o	45.8	115.8
h5-cd-na	49.6	121.55
h2-c3-os	50.9	109.58
nc-cd-h5	50.6	125.52
os-p5-o	43.9	115.46
nd-cc-na	74.9	112.22
os-c3-na	71.3	109.03

Dihedral	v	γ	n
c-c3-c3-c3	0.156	0.0	3
c3-c3-os-c3	0.1	180.0	2
c3-os-p5-o	0.55	0.0	3
o-c-c3-hc	0.0	0.0	2
h1-c3-os-p5	0.383	0.0	3
ca-nb-ca-nb	4.8	180.0	2
nc-cd-na-ca	1.7	180.0	2
h5-cd-na-ca	1.7	180.0	2
nd-cc-na-ca	1.7	180.0	2
h2-c3-c3-c3	0.156	0.0	3
c3-os-p5-os	0.25	0.0	3
h2-c3-na-ca	0.0	0.0	2
os-p5-os-c3	1.2	0.0	2
oh-c3-c3-os	1.175	0.0	2
nc-ca-ca-nb	3.625	180.0	2
os-c3-c3-h1	0.25	0.0	1
cc-cc-nd-ca	4.75	180.0	2
c-n-c-cc	2.5	180.0	2
c3-c3-os-c3	0.383	0.0	3
c3-c3-os-p5	0.383	0.0	3
nh-ca-ca-ca	3.625	180.0	2
os-c3-c3-c3	0.156	0.0	3
c3-ca-ca-c3	3.625	180.0	2
hn-n-c-o	2.5	180.0	2
h1-c3-c3-oh	0.25	0.0	1
c3-c3-na-cd	0.0	0.0	2
c3-c3-oh-ho	0.25	0.0	1
os-c3-c3-os	0.144	0.0	3
h1-c3-c3-os	0.0	0.0	3
nd-cc-cc-c	4.0	180.0	2
oh-c3-c3-h1	0.0	0.0	3
c3-os-c3-h1	0.383	0.0	3
ho-oh-c3-c3	0.16	0.0	3
o-c-n-hn	2.0	0.0	1
c3-c3-c3-c3	0.2	180.0	1
o-p5-os-c3	0.55	0.0	3
o-p5-os-p5	0.8	0.0	2
c-n-c-o	2.5	180.0	2
na-ca-ca-ha	3.625	180.0	2
c3-c3-c3-oh	0.156	0.0	3
h2-c3-os-c3	0.383	0.0	3
n-c-nd-cc	4.0	180.0	2
hc-c3-c3-oh	0.0	0.0	3
o-c-c3-hc	0.8	0.0	1

Dihedral	v	γ	n
cd-na-ca-nb	0.3	180.0	2
c3-c3-na-ca	0.0	0.0	2
o-c-n-hn	2.5	180.0	2
h1-c3-c3-na	0.156	0.0	3
oh-c3-c3-oh	1.175	0.0	2
ho-oh-c3-h1	0.167	0.0	3
p5-os-p5-os	0.8	0.0	2
hn-n-c-cc	2.5	180.0	2
ho-oh-c3-c3	0.25	0.0	1
os-c3-c3-h1	0.0	0.0	3
oh-c3-c3-c3	0.156	0.0	3
c3-na-cd-nc	1.7	180.0	2
ca-ca-ca-ha	3.625	180.0	2
ca-ca-ca-c3	3.625	180.0	2
cc-nd-ca-ca	4.8	180.0	2
hn-n-c-o	2.0	0.0	1
p5-os-c3-c3	3.95	180.0	1
ca-ca-na-cd	0.3	180.0	2
nh-ca-ca-nc	3.625	180.0	2
h1-c3-c3-oh	0.0	0.0	3
h2-c3-c3-oh	0.156	0.0	3
c-cc-nd-ca	4.75	180.0	2
h1-c3-c3-c3	0.156	0.0	3
c3-c3-c3-os	0.25	0.0	1
oh-c3-c3-h1	0.25	0.0	1
c3-na-ca-ca	0.3	180.0	2
c3-c3-c3-c3	0.25	180.0	2

Impropers

Improper	v	γ	n
ca-ca-ca-nd	1.1	180.0	2
n-cc-c-o	10.5	180.0	2
ca-nb-ca-nh	1.1	180.0	2
c3-ca-na-cd	1.1	180.0	2
ca-ca-ca-na	1.1	180.0	2
ca-na-ca-nb	1.1	180.0	2
c3-ca-na-cc	6.1	180.0	2
h5-nb-ca-nb	10.5	180.0	2
ca-ca-ca-nc	1.1	180.0	2
ca-hn-nh-hn	1.1	180.0	2
h5-na-cd-nc	1.1	180.0	2
ca-ca-ca-ha	1.1	180.0	2
c-c-n-hn	1.1	180.0	2
ca-ca-ca-c3	1.1	180.0	2
nd-n-c-o	10.5	180.0	2
c3-ca-ca-ca	1.1	180.0	2

4.3   Root Mean Square Deviation (RMSD) plots from MD simulations

Figure S7: RMSD of KSTD with 17-MT (TYR318 is deprotonated), results from 55 ns of MD simulation.

Figure S8: RMSD of KSTD with DHT (TYR318 is deprotonated), results from 55 ns of MD simulation.

Figure S9: RMSD of KSTD with 17-MT (alternative pathway - TYR318 is protonated), results from 55 ns of MD simulation.
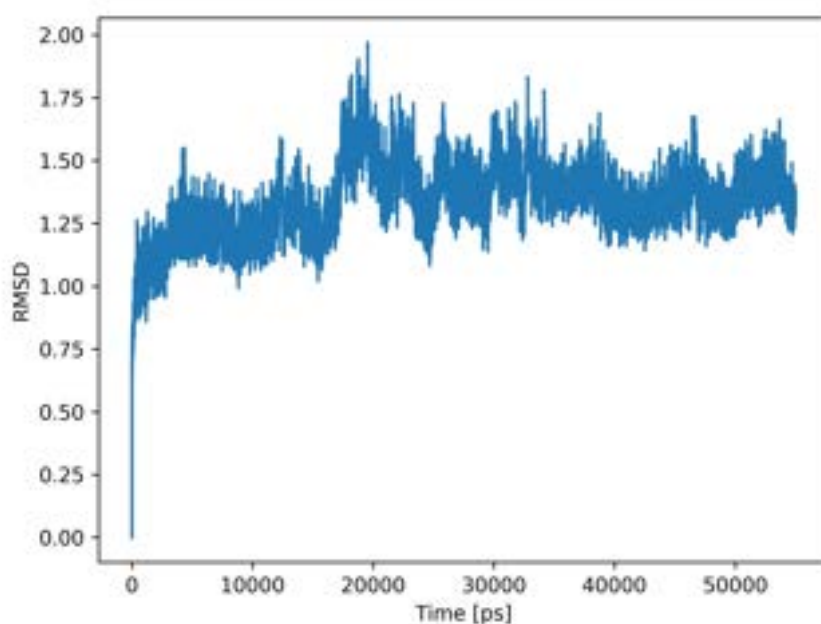
Figure S10: RMSD of KSTD with DHT (alternative pathway - TYR318 is protonated), results from 55 ns of MD simulation.

4.4   Alternative mechanism – model preparation

Model setup was identical like for the main pathway, but Tyr318 was kept in its standard protonation state. The system was neutralized by adding 32 sodium ions. Calculation was conducted for 17MT and DHT. Structure was relaxed by 55 ns of MD. QM layer was enlarged by including TYR487.

16

5 Results details

5.1 Two dimensional Potential Energy Surface
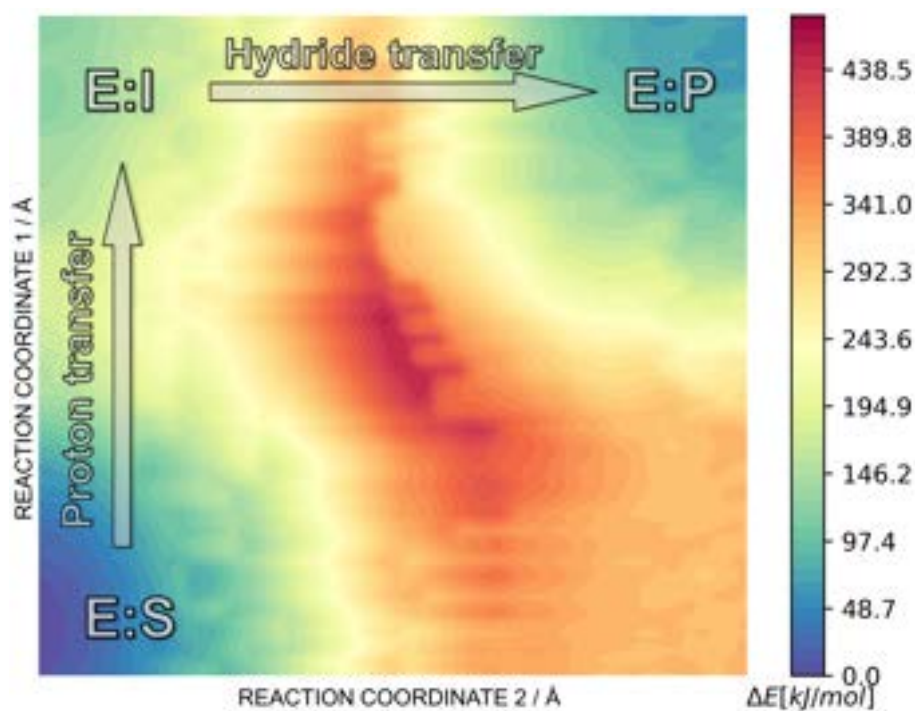
Figure S11: PES for Δ^1-dehydrogenation of 17MT at AM1/AMBER level of theory
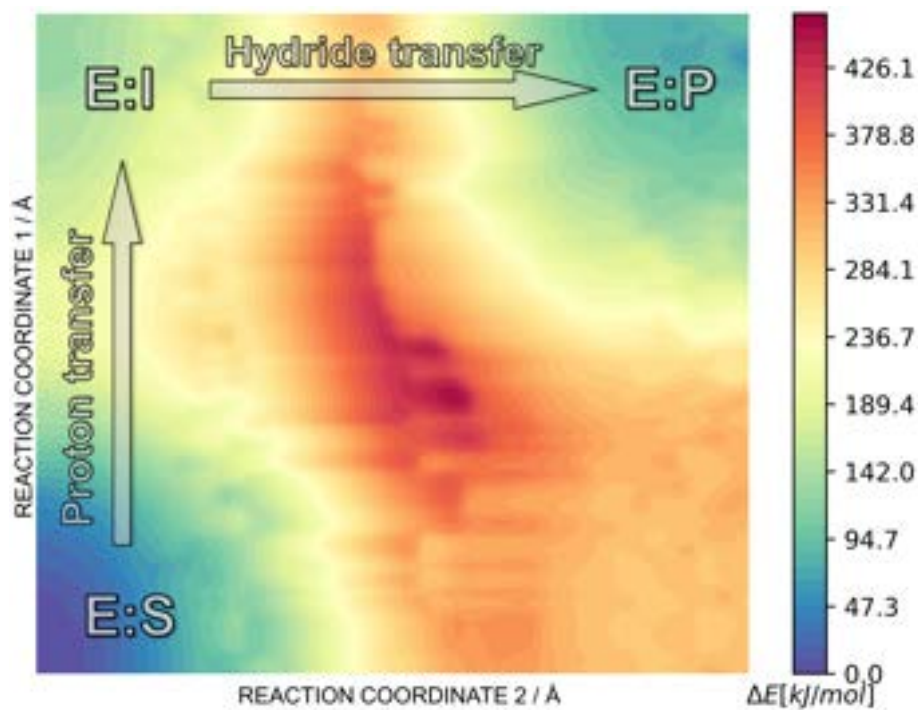
Figure S12: PES for Δ^1-dehydrogenation of DHT at AM1/AMBER level of theory
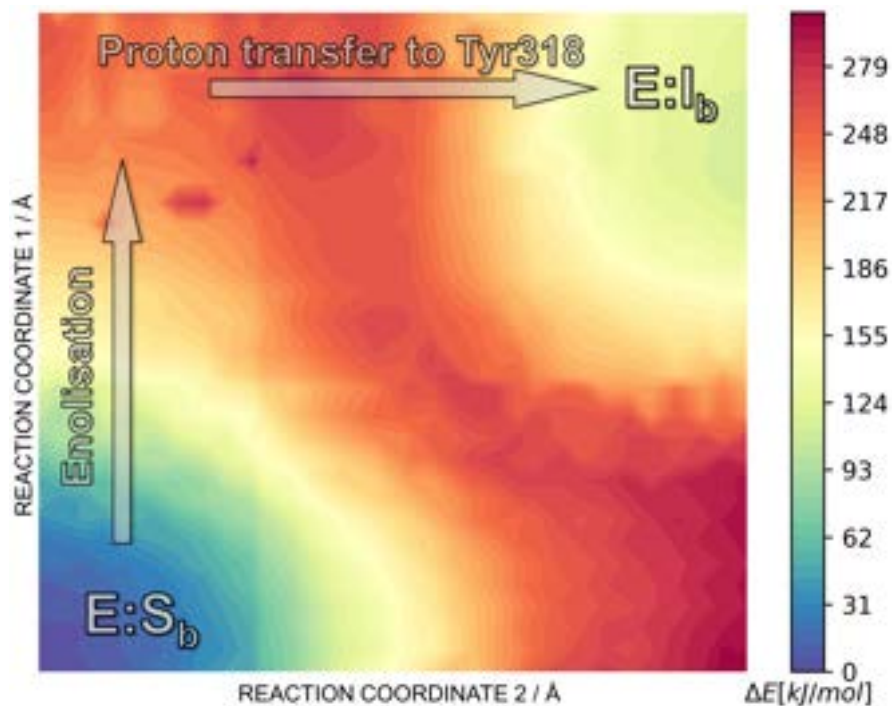
Figure S13: PES for 17MT for first step of alternative mechanism. Presented energies are obtained from single point calculations at B3LYP/AMBER level at theory. Structures were optimized at AM1/AMBER level of theory

5.2   Potential of Mean Force
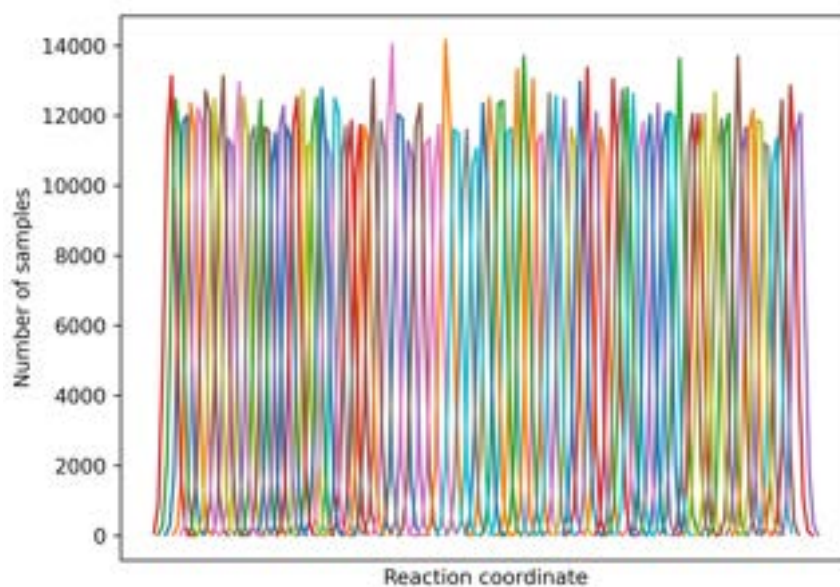
Figure S14: Reaction coordinate distribution for each window of PMF calculations of first stage for 17MT
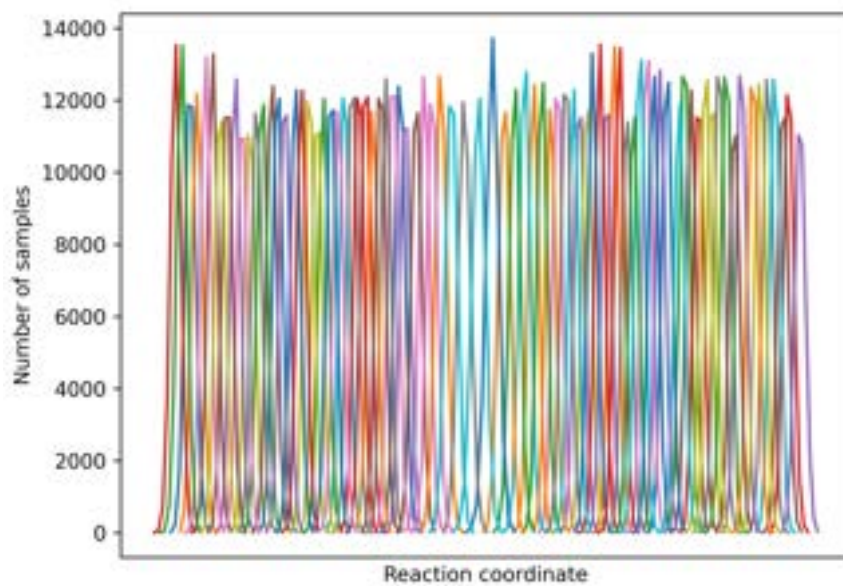
Figure S15: Reaction coordinate distribution for each window of PMF calculations of second stage for 17MT
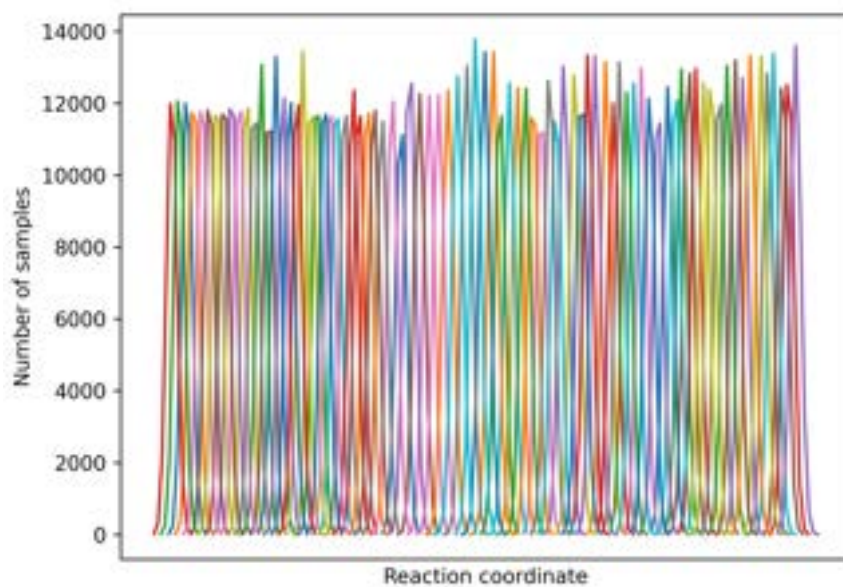
Figure S16: Reaction coordinate distribution for each window of PMF calculations of first stage for DHT
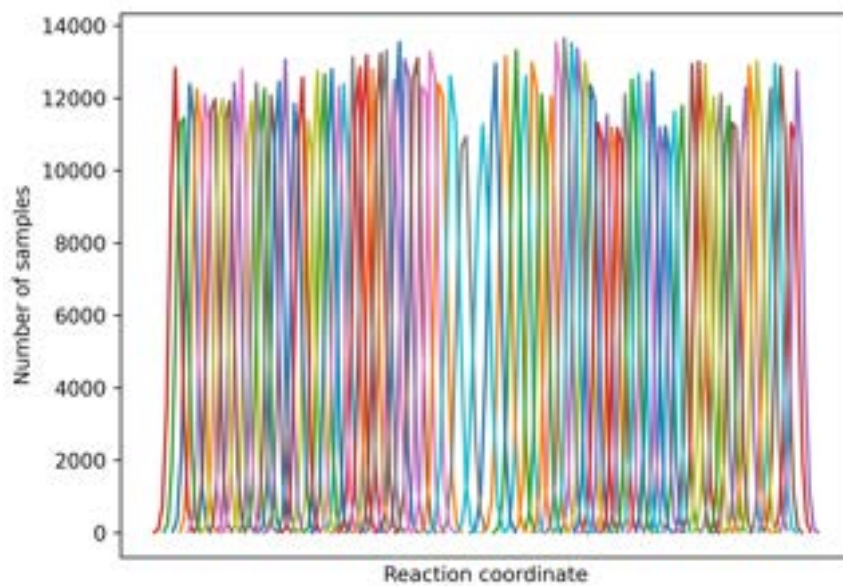
Figure S17: Reaction coordinate distribution for each window of PMF calculations of second stage for DHT

| H | 16.274015 | 45.489601 | 73.616988 | | | | |

5.11 17-MT – computed KIE details

Table S 20: Overview of computed KIE values for 17-MT.

Stage	AM1:AMBER	B3LYP:AMBER	B3LYP:AMBER Wigner corr.
E:S⟶TS1	**4.6931** ± 0.52063	**5.3885** ± 0.12346	**7.16** ± 0.18
E:S⟶TS2	**1.1168** ± 0.05405	**0.97662** ± 0.042572	**0.977** ± 0.043
E:I⟶TS1	**3.6687** ± 0.62819	**5.126** ± 0.068516	**6.81** ± 0.11
E:I⟶TS2	**0.97408** ± 0.051163	**0.91915** ± 0.06294	**0.920** ± 0.063
E:S⟶E:I	**1.3056** ± 0.20534	**1.0512** ± 0.020787	**1.051** ± 0.031
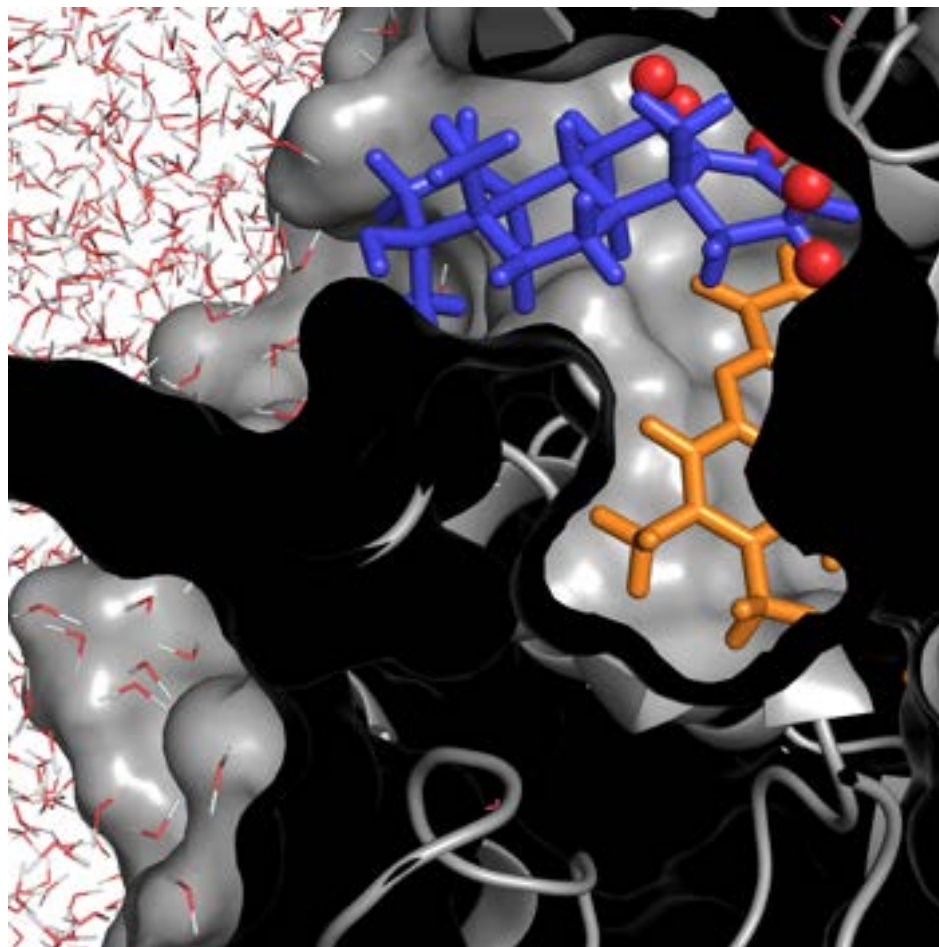
Figure S33: Representative structure of DHT (blue sticks)–enzyme (gray) complex. Atoms substituted with deuterium are marked as red spheres.

5.12 DHT – computed KIE details

Table S 21: Overview of computed KIE values for DHT.

Stage	AM1:AMBER	B3LYP:AMBER	B3LYP:AMBER Wigner corr.
E:S⟶TS1	**1.0298** ± 0.02599	**1.1241** ± 0.035588	**1.124** ± 0.036
E:S⟶TS2	**4.7988** ± 0.14875	**4.9532** ± 0.13808	**6.85** ± 0.20
E:I⟶TS1	**1.0013** ± 0.019547	**0.96991** ± 0.036058	**0.970** ± 0.036
E:I⟶TS2	**4.6291** ± 0.14308	**4.3646** ± 0.15518	**6.04** ± 0.22
E:S⟶E:I	**1.0288** ± 0.027973	**1.16** ± 0.046048	**1.159** ± 0.057

44
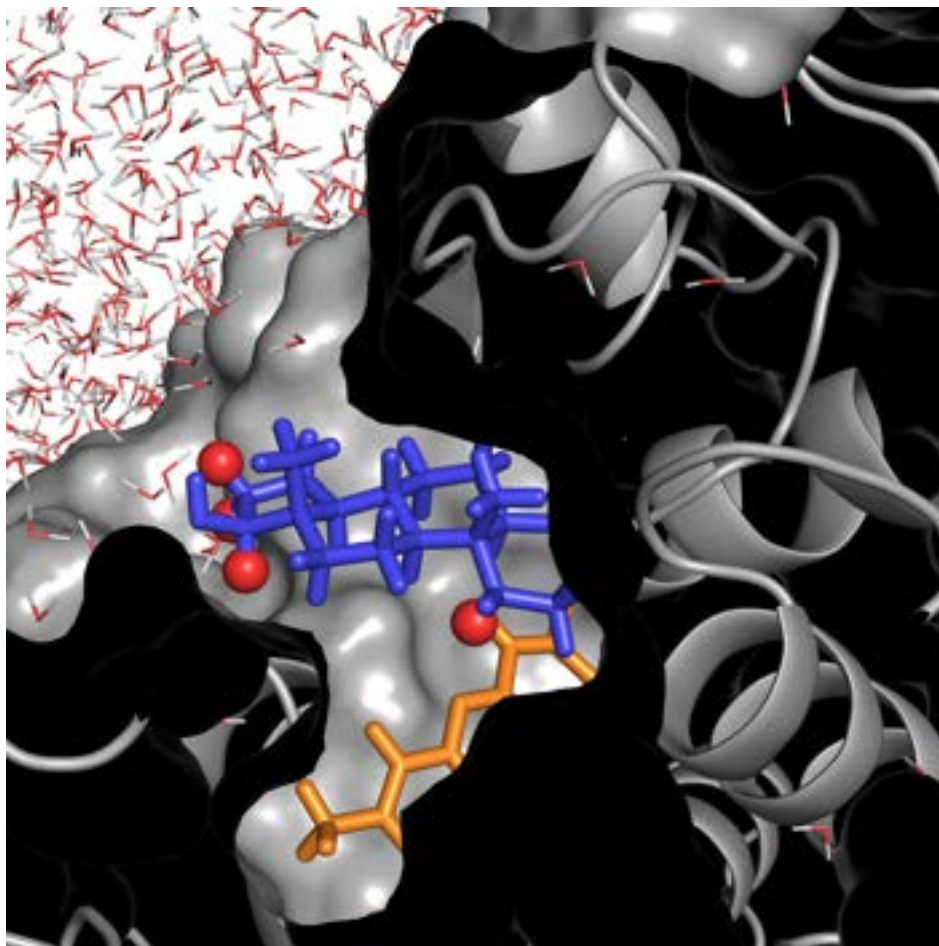
Figure S34: Representative structure of DHT (blue sticks)–enzyme (gray) complex. Atoms substituted with deuterium are marked as red spheres.

5.13 KIE contribution from individual atoms

Table S 22: KIE contribution from individual deuterium atoms for 17MT, S⟶TS1
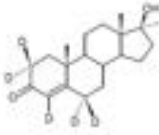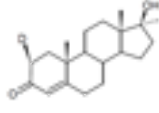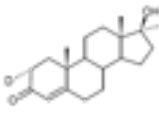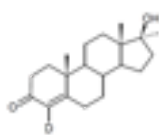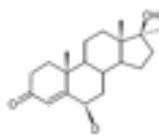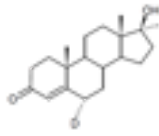
AM1:AMBER			B3LYP:AMBER			Molecule
Mean	Stdev	max-min	Mean	Stdev	max-min	
4.6931	0.52063	1.9084	**5.3885**	0.12346	0.37894	
4.6228	0.39968	1.3104	**4.7951**	0.043492	0.11177	
1.0407	0.023459	0.11594	**1.0863**	0.0046135	0.010637	
1.0038	0.023266	0.1025	**1.0144**	0.013856	0.04413	
0.97875	0.008086	0.034514	**0.99855**	0.0040321	0.012843	
0.98051	0.01121	0.041743	**1.0079**	0.0078959	0.025763	

Table S 23: KIE contribution from individual deuterium atoms for 17MT, I⟶TS1
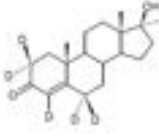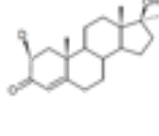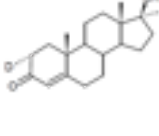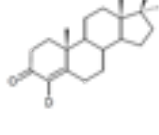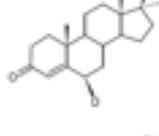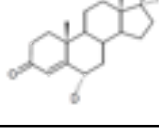
AM1:AMBER			B3LYP:AMBER			Molecule
Mean	Stdev	max-min	Mean	Stdev	max-min	
3.6687	0.62819	2.5582	**5.126**	0.068516	0.18964	
3.8069	0.72602	2.7384	**5.4398**	0.041293	0.11866	
0.9799	0.024778	0.11592	**0.94661**	0.0051138	0.011629	
0.99987	0.023654	0.10345	**1.0083**	0.019101	0.060745	
0.99537	0.0082126	0.038895	**0.97261**	0.0025893	0.0068353	
0.99488	0.014379	0.06276	**1.017**	0.01187	0.034772	

Table S 24: KIE contribution from individual deuterium atoms for 17MT, I⟶TS2

AM1:AMBER			B3LYP:AMBER			Molecule
Mean	Stdev	max-min	Mean	Stdev	max-min	
0.97408	0.051163	0.24457	**0.91915**	0.06294	0.17839	
0.948	0.021182	0.097175	**0.96439**	0.023296	0.069148	
1.0003	0.015425	0.075839	**0.94904**	0.0040438	0.011448	
1.0012	0.016481	0.080304	**0.99364**	0.02392	0.071782	
1.0137	0.006864	0.032632	**1.0053**	0.0075283	0.021251	
1.0114	0.0087242	0.036826	**1.004**	0.016958	0.046937	

Table S 25: KIE contribution from individual deuterium atoms for 17MT, S⟶TS2
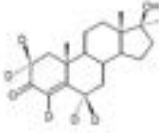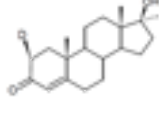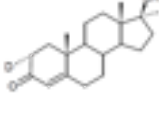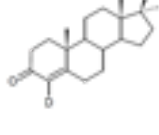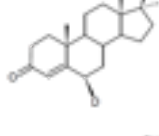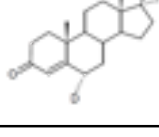
AM1:AMBER			B3LYP:AMBER			Molecule
Mean	Stdev	max-min	Mean	Stdev	max-min	
1.1168	0.05405	0.26734	**0.97662**	0.042572	0.1196	
0.99541	0.014175	0.064241	**0.86155**	0.014016	0.041283	
1.0779	0.024716	0.12212	**1.0761**	0.0047021	0.012027	
1.0128	0.019793	0.087102	**0.99659**	0.018107	0.053371	
1.0025	0.0081102	0.03564	**1.0313**	0.004622	0.014392	
1.0119	0.0061296	0.026297	**1.0098**	0.0099398	0.030114	

Table S 26: KIE contribution from individual deuterium atoms for DHT, S⟶TS1
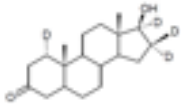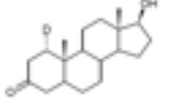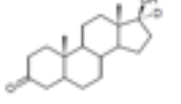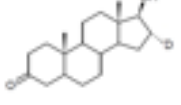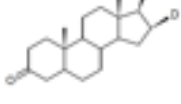
AM1:AMBER			B3LYP:AMBER			Molecule
Mean	Stdev	max-min	Mean	Stdev	max-min	
1.0298	0.02599	0.13132	**1.1241**	0.035588	0.11303	
1.0255	0.004388	0.020789	**1.1163**	0.017658	0.050495	
0.99818	0.019915	0.090244	**0.99982**	0.042435	0.12247	
0.99734	0.014464	0.059107	**0.99827**	0.01491	0.048042	
1.0088	0.016951	0.07317	**1.0095**	0.018289	0.0593	

Table S 27: KIE contribution from individual deuterium atoms for DHT, I⟶TS1
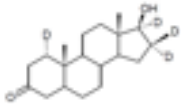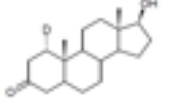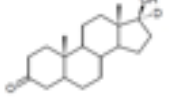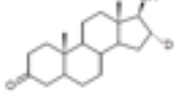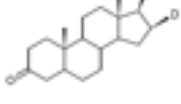
| AM1:AMBER | | | B3LYP:AMBER | | | Molecule |
Mean	Stdev	max-min	Mean	Stdev	max-min	
1.0013	0.019547	0.1041	**0.96991**	0.036058	0.10383	
0.99874	0.006973	0.028978	**0.96001**	0.0096493	0.028594	
0.99902	0.018107	0.081268	**1.0026**	0.041578	0.11882	
0.99946	0.014193	0.056181	**1.0011**	0.015128	0.048911	
1.0042	0.013827	0.05828	**1.007**	0.017792	0.054939	

51

Table S 28: KIE contribution from individual deuterium atoms for DHT, I⟶TS2
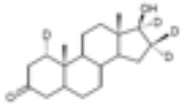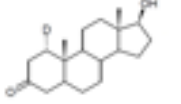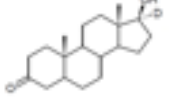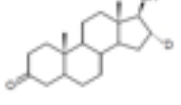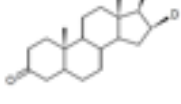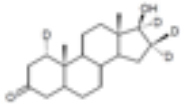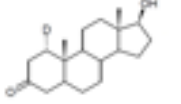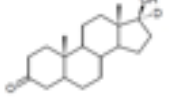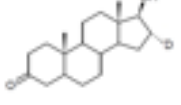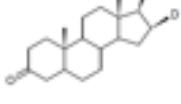
AM1:AMBER			B3LYP:AMBER			Molecule
Mean	Stdev	max-min	Mean	Stdev	max-min	
4.6291	0.14308	0.71294	**4.3646**	0.15518	0.46853	
4.6182	0.06111	0.26868	**4.3422**	0.036866	0.085493	
0.9995	0.018838	0.097834	**1.0029**	0.049484	0.15634	
1.0031	0.01578	0.066468	**1.004**	0.020212	0.063215	
0.99978	0.021735	0.111	**0.99927**	0.026952	0.083667	

Table S 29: KIE contribution from individual deuterium atoms for DHT, S⟶TS2

AM1:AMBER			B3LYP:AMBER			Molecule
Mean	Stdev	max-min	Mean	Stdev	max-min	
4.7988	0.14875	0.74841	**4.9532**	0.13808	0.43339	
4.7554	0.060651	0.24244	**4.9123**	0.072691	0.17124	
0.9975	0.021063	0.11136	**0.99848**	0.044729	0.1345	
0.99945	0.015214	0.063753	**0.99797**	0.015977	0.051953	
1.0124	0.022037	0.10359	**1.0127**	0.023873	0.07295	

53

Publikacja P4

CCC
RightsLink

🏠 Home   ❓ Help ⌄   💬 Live Chat   👤 Sign in   👤+ Create Account

Structure, Mutagenesis, and QM:MM Modeling of 3-Ketosteroid Δ1-Dehydrogenase from Sterolibacterium denitrificans—The Role of a New Putative Membrane-Associated Domain and Proton-Relay System in Catalysis

ACS Publications
Most Trusted. Most Cited. Most Read.

Author: Patrycja Wójcik, Michał Glanowski, Beata Mrugała, et al

Publication: Biochemistry

Publisher: American Chemical Society

Date: Feb 1, 2023

Copyright © 2023, American Chemical Society

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms and Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from {COMPLETE REFERENCE CITATION}. Copyright {YEAR} American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your RightsLink request. No additional uses are granted (such as derivative works or other editions). For any uses, please submit a new request.

If credit is given to another source for the material you requested from RightsLink, permission must be obtained from that source.

BACK   CLOSE WINDOW

Biochemistry

Article

Structure, Mutagenesis, and QM:MM Modeling of 3-Ketosteroid Δ1-Dehydrogenase from *Sterolibacterium denitrificans*—The Role of a New Putative Membrane-Associated Domain and Proton-Relay System in Catalysis

Patrycja Wójcik, Michał Glanowski, Beata Mrugała, Magdalena Procner, Olga Zastawny, Monika Flejszar, Katarzyna Kurpiewska, Ewa Niedziałkowska, Wladek Minor, Maria Oszajca, Andrzej J. Bojarski, Agnieszka M. Wojtkiewicz, and Maciej Szaleniec*

Read Online

ACCESS | 📊 Metrics & More | 📰 Article Recommendations | 🗒 Supporting Information

ABSTRACT: 3-Ketosteroid Δ1-dehydrogenases (KstD) are important microbial flavin enzymes that initiate the metabolism of steroid ring A and find application in the synthesis of steroid drugs. We present a structure of the KstD from *Sterolibacterium denitrificans* (AcmB), which contains a previously uncharacterized putative membrane-associated domain and extended proton-relay system. The experimental and theoretical studies show that the steroid Δ1-dehydrogenation proceeds according to the Ping−Pong bi−bi kinetics and a two-step base-assisted elimination (E2cB) mechanism. The mechanism is validated by evaluating the experimental and theoretical kinetic isotope effect for deuterium-substituted substrates. The role of the active-site residues is quantitatively assessed by point mutations, experimental activity assays, and QM/MM MD modeling of the reductive half-reaction (RHR). The pre-steady-state kinetics also reveals that the low pH (6.5) optimum of AcmB is dictated by the oxidative half-reaction (OHR), while the RHR exhibits a slight optimum at the pH usual for the KstD family of 8.5. The modeling confirms the origin of the enantioselectivity of C2-H activation and substrate specificity for Δ4-3-ketosteroids. Finally, the cholest-4-en-3-one turns out to be the best substrate of AcmB in terms of ΔG of binding and predicted rate of dehydrogenation.
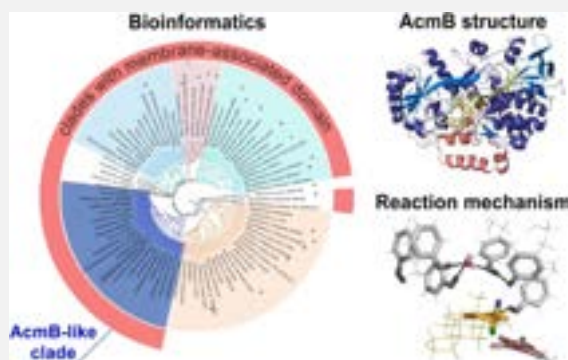
INTRODUCTION

Steroids belong to a class of lipid triterpenes characterized by a structure of an aliphatic tetracyclic system with a low number of functional groups in its core, complex spatial structure, and low solubility in water.[1] They are widespread in plants, insects, vertebrates, yeasts, and fungi.[2] Cholesterol-derived steroids are used as hormones and regulate various aspects of the metabolism. As a result, steroid hormones and their analogues are one of the most important groups of drugs produced by the pharmaceutical industry.[3]

The chemical and biotechnological modification of steroids is of utmost importance for the development of new functionalities of steroid drugs. One such important modification is the introduction of a double bond between the atoms C1 and C2 of ring A, so-called Δ1-dehydrogenation, which increases the potency and selectivity of steroid drugs. For example, the oxidation of hydrocortisone to prednisolone results in increased anti-inflammatory activity while reducing salt-retaining activity.[4,5] Δ1-Dehydrogenation, which starts the central degradation pathway in bacteria (the opening of ring A leading to the complete mineralization of the steroids), is

catalyzed by a microbial flavin enzyme, 3-ketosteroid Δ1-dehydrogenase (KstD, EC 1.3.99.4). The genes encoding KstDs, characterized by a rather high sequence diversity, are abundant among various bacteria and fungi.[6] As a result, microbial systems are routinely used by the pharmaceutical industry in the biotransformation of various sterols and steroids.[7,8]

Steroid-transforming bacteria frequently have several isoenzymes at their disposal. These isoenzymes usually differ in their substrate preference and are up-regulated in the presence of different steroids.[6] Multiple isoenzymes in a genome may help microorganisms degrade steroids with different structural characteristics.[6] Recently, several reports confirmed that KstDs
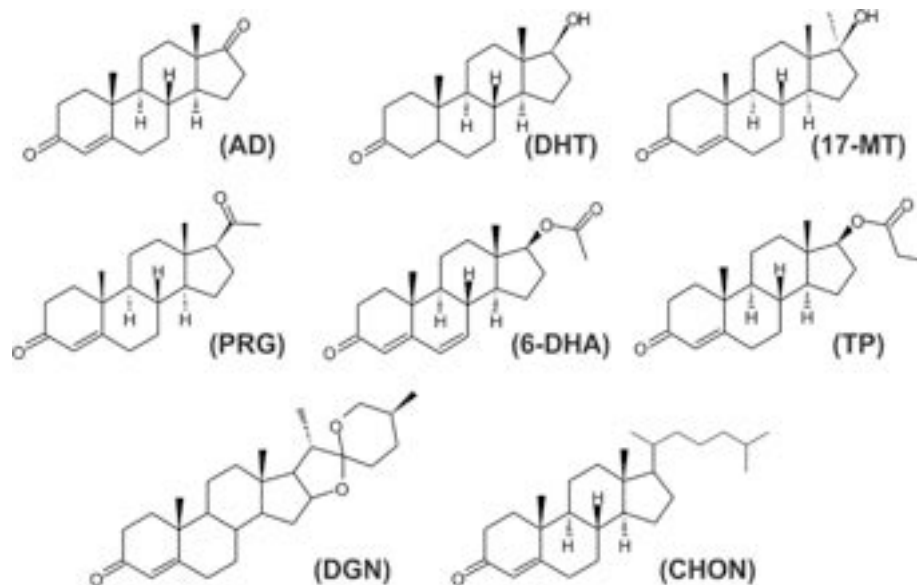
Figure 1. Steroids used in the study; AD—androst-4-en-3,17-dione, DHT—dihydrotestosterone, 17-MT—17-methyltestosterone, PRG—progesterone, 6-DHA—6-dehydrotestosterone acetate, TP—testosterone propionate, DGN—diosgenone, and CHON—cholest-4-en-3-one.

are capable of dehydrogenating steroids with undegraded isooctyl C17 substituents.[9,10] One such nonstandard KstDs is AcmB from *Sterolibacterium denitrificans* Chol-1, a suitable catalyst for the dehydrogenation of C17-extended ketosteroids and 3-keto-saponins. Interestingly, it also exhibits multiple pH optima, which depend on the electron acceptor used. As a result, AcmB can catalyze the 1,2-dehydrogenation of 3-ketosteroids at a slightly acidic or basic pH.[11,12]

Furthermore, the sequence of AcmB also differs from the only structurally characterized representative of the KstDs, namely, the KstD1 from *Rhodococcus erythropolis* (PDB: 4C3X and 4C3Y).[13] Our bioinformatic analysis indicates that KstD1 from *R. erythropolis* (GenBank accession code AAF19054.1) is shorter by approximately 50 amino acids when compared to the majority of the know KstD sequences. The longest difference in the amino acid sequence, 38 amino acids (compared to the sequence of AcmB, see Figure S1), was previously referred to as the "loop".[10] As a result, the crystal structure of KstD1 from *R. erythropolis* represents only a fraction of the known enzymes in the KstD family. In this work, we structurally characterize KstD from *S. denitrificans* (AcmB, GenBank accession code ABV59992), which, based on its amino acid sequence composition, represents most of the known KstDs. The structural characterization is combined with kinetic and modeling studies addressing the reaction mechanism with several steroid substrates (Figure 1). With pre-steady-state kinetics, we explained the surprising origin of two pH optima of AmcB.[11] We have also successfully used site-directed mutagenesis and QM/MM MD calculations to quantitatively explain the influence of active-site residues on reaction kinetics. Finally, we discussed the biological role of the "loop". We discuss its possible involvement in the enzyme−membrane association, the substrate binding, and its potential responsibility for the substrate specificity of AcmB.

METHODS

Materials. All chemicals were purchased from Sigma-Aldrich, Tokyo Chemical Industry, Carl Roth, or BioShop unless otherwise specified. 2,2,4,6,6-d_5-4-Androsten-17α-meth-yl-17β-ol-3-one was purchased from CDN Isotopes, while 1,16,16,17-d_4-17β-hydroxy-5α-androstan-3-one was obtained from Alsachim.

Protein Purification and Crystallization. AcmB was expressed in *Escherichia coli* BL21(DE3)Magic (Creative Biolabs) and purified by Ni-affinity chromatography as previously described by Wojtkiewicz et al.[11] The removal of (His)$_6$-AcmB fusion protein was carried out during overnight dialysis with recombinant (His)$_6$-TEV protease. The protein was then loaded onto a HiLoad Superdex 200 16/600 pg (Cytiva) size-exclusion column and eluted with 50 mM Tris-HCl pH 8.5, 150 mM NaCl, 0.5 mM tris(2-carboxyethyl)-phosphine (TCEP), and 0.2% (v/v) Tween 20. All chromatographic experiments were performed using an NGC Quest 10 Plus instrument (BioRad).

The protein was concentrated to 20 mg mL^{-1} using an Amicon Ultra-15 30 kDa (Millipore) and incubated on ice with the excess amount of androsta-1,4-diene-3,17-dione (ADD) for approximately 3 h. Subsequently, the solution was centrifuged at 14,000 g, 4 °C for 5 min. Crystallization conditions were screened using MCSG1, MCSG2, MCSG3, Top96, SuperCOMBI, PurePEG (Anatrace), Index, Silver Bullets, and PEG/Ion HT (Hampton Research) with success in the MCSG1 screen. Crystallization experiments were performed using the sitting-drop vapor diffusion method by mixing a 1:1 screen solution with the protein solution. Finally, a diffraction-quality crystal grew in 0.1 M HEPES NaOH pH 7.5, 25% (w/v) PEG 3350 (MCSG1). For XRD experiments, the crystal was cryoprotected by drying over 1 M NaCl and flash-cooled in liquid nitrogen.

Data Collection and Analysis. Initial crystallization and structure determination were performed in MINORlab with data collected in Argonne. Later, a better crystal was used to obtain 1.84 Å diffraction data collected at a temperature of 100 K on the BESSY 14.1 beamline (Helmholtz-Zentrum Berlin, Germany) using a Dectris PILATUS 6 M detector. The data set was recorded at a wavelength of 0.9184 Å. Data were collected and processed using CrysAlis[Pro][14] and XDS.[15] The structure solution and model building were carried out with

Phenix (Autobuild)[16] and CCP4 (ARP/wARP, BUCCA-NEER).[17] The structure was solved by molecular replacement with the structure of *R. erythropolis* (PDB: 4C3Y)[13] as a search model. The resulting model was then refined with REFMAC 5.8[18] and manually rebuilt with WinCoot 0.8.9.2.[19] The quality of the model was evaluated using Molprobity[20] and the wwPDB Validation Service.[21] The data were deposited in the Protein Data Bank (PDB) with PDB code 7P18. The quality of diffraction data and structure refinement is present in Table S3 (Supporting Information).

The structure was further analyzed using PyMOL,[22] Chimera 1.15,[23] and Discovery Studio 2018.[24] The membrane-associated region was predicted with Orientations of Protein in Membranes database and PPM 3.0 using the bacterial Gram-negative inner membrane model.[25]

Site-Directed Mutagenesis. The pMCSG7-acmb plasmid[26] was used as a template for site-directed mutagenesis introducing point mutations (Y115F, Y118F, Y363F, Y467F, Y536F, and G540P) into the AcmB gene. The details of the procedure are available in Supporting Information. The FAD content in the expressed enzymes was estimated as the ratio of the spectrophotometrically measured AcmB-FAD concentration ($\varepsilon_{\text{AcmB 450 nm}}$ = 12 094 M^{-1} cm^{-1}, determined as in ref 27) to the total protein concentration determined according to the Bradford method.[14]

Kinetic Assays. *Stopped-Flow Spectrophotometric Activity Assay.* The pre-steady-state and steady-state kinetics aimed at establishing the kinetic mechanisms of the enzyme were collected using a stopped-flow spectrophotometer SX20 (Applied Photophysics). The solutions were treated with argon for several minutes to provide anaerobic conditions. All reported concentrations are the final values obtained after mixing and diluting the reactants. All measurements were performed in triplicate. Data were collected using Pro-Data software and processed using OriginPro 2019b software.

RHR and OHR pH Optimum. The pre-steady-state kinetics was used to determine the pH optimum of the reductive half-reaction (RHR) and oxidative half-reaction (OHR) catalyzed by AcmB. The following buffers were used: 50 mM Tris-HCl, 150 mM NaCl, and 5% (w/v) glycerol (pH 7.5−8.5) and 50 mM glycine-NaOH, 150 mM NaCl, and 5% (w/v) glycerol (pH 9.0−10.0). The enzyme purified with Ni-affinity chromatography[11] was transferred to the respective buffer using Econo-Pac 10DG desalting columns (BioRad) according to the instruction manual. Due to the rapid aggregation of the enzyme at low pH observed under the concentration required for the stopped-flow experiment, it was impossible to measure the reduction rate below pH 7.5.

For RHR, the FAD reduction was followed at 450 nm at 20 °C in the reaction of 9.3 μM of AcmB with 100 μM progesterone using the buffers described above. In each measurement, one glass syringe of the stopped-flow instrument was filled with the solution of the steroid dissolved in EGME (the final concentration of 8%) and buffer solution, while the second one contained the enzyme solution in the buffer.

In the case of the OHR experiment, the enzyme was first reduced under anaerobic conditions with a subequivalent amount of progesterone (1:0.8), thus avoiding the excess progesterone that would interfere with the enzyme reoxidation. In each measurement, one glass syringe was filled with the buffer solution of the reduced AcmB, and the second syringe was filled with the buffer solution of 2,6-dichloroindophenol (DCPIP) and 2-methoxyethanol (EGME). The solutions were

treated with argon for several minutes before the experiment. The OHR was followed at 616 nm and 20 °C in 50 μM DCPIP in 8% EGME with 10 μM AcmB in the buffer solution. The obtained traces were fitted with double-exponential functions, yielding eigenvalues λ_{obs}.

Kinetic Studies. The steady-state kinetic studies leading to establishing the kinetic mechanism were conducted with the stopped flow according to the previously described methodology.[27] The reaction mixture after mixing in stopped flow contained 100 mM K_2HPO_4/KH_2PO_4 buffer pH 6.5, 52.2 nM of AcmB with 50% FAD content, 1% isopropyl alcohol (IPA), and varying concentrations of DCPIP (0.1−0.3 mM) and progesterone (2.5−50 μM). In order to establish the kinetic mechanism, the received data were fitted with nonlinear regression to three two-substrate kinetic models (i.e., sequential ordered bi−bi and random sequential bi−bi and the nonsequential Ping−Pong bi−bi models). The best model was selected based on statistical parameters (i.e., R^2, χ^2, AICc, and errors of estimated constants).

Activity Assay for Mutated Variants and the Kinetic Isotope Effect. The kinetic assays of mutein activities, as well as steady-state values for the direct kinetic isotope effect, were obtained using a UV-2700 spectrophotometer (Shimadzu) in 0.5 mL quartz cuvettes with a 10 mm path. The reduction of DCPIP was followed at 700 nm ($\varepsilon_{700[\text{pH 6.5}]}$ = 4 576 M^{-1} cm^{-1} or $\varepsilon_{700[\text{pH 8.5}]}$ = 5 190 M^{-1} cm^{-1}) at 30 °C, kinetic curves' initial parts (5−10 s).

The specific activities of wild-type AcmB and its mutants were determined in the reaction that consisted of 0.1 M KH_2PO_4/K_2HPO_4 buffer pH 6.5, 200 μM DCPIP, and 100 μM androst-4-en-3,17-dione (AD) in IPA (the final concentration of 2%) and 0.15−0.42 μM of AcmB.

HPLC Activity Assay. The activities of muteins were additionally confirmed with HPLC analysis. The reaction mixture consisted of 0.1 M KH_2PO_4/K_2HPO_4 buffer pH 6.5, 400 μM DCPIP, 200 μM AD in IPA (2%), and 24.6−38.9 μM of AcmB. The reactions were carried out in a thermoblock at 30 °C and 800 rpm for 30 min. The reaction progress was stopped after 30 min by mixing the samples with acetonitrile (1:1), followed by centrifugation at 14,000 g for 5 min and analysis with LC DAD-HPLC (Agilent 1100) according to Wojtkiewicz et al.[11]

Kinetic Isotope Effect. The kinetic isotope effect (KIE) was determined by direct and competition methods. In the direct approach, the reaction rates were determined in the spectrophotometric activity assay described above. Measurements were carried out in 100 mM K_2HPO_4/KH_2PO_4 pH 6.5 or 50 mM Tris-HCl pH 8.5 with 200 μM DCPIP, 100 μM (pH 6.5), or 200 μM (pH 8.5) steroid (17-MT, 2,2,4,6,6-d_5-17-MT, DHT and 1,16,16,17-d_4-DHT) dissolved in dioxane (the final concentration 1 or 2%, respectively) and 0.17 μM of AcmB.

For the competitive kinetic isotope effect, $^D(V/K)$ was measured according to the previously described protocol.[28,29] The value of $^D(V/K)$ KIE was established based on the fractions of the converted nondeuterated (x_1) and deuterated (x_2) substrate according to the formula

$$\left(\frac{V}{K}\right) = \frac{k_1}{k_2} = \frac{\log(1 - x_1)}{\log(1 - x_2)}$$

KIE was established by a nonlinear fit to the reformulated function of $x_1(x_2)$

$$x_1 = 1 - (1 - x_2)^{(k_1/k_2)}$$

The reaction mixtures consisted of 50 mM K_2HPO_4/KH_2PO_4 pH 6.5 or 50 mM Tris-HCl pH 8.5, 100 μM DCPIP, and equal 100 μM amounts of substrates and their deuterated homologues (17-MT and 2,2,4,6,6-d_5-17-MT or DHT and 1,16,16,17-d_4-DHT) in EGME (the final concentration 2%) and AcmB (1.9 nM for pH 6.5 or 4.7 nM for pH 8.5). The reactions were carried out in triplicate under anaerobic conditions [98:2 (v/v) N_2/H_2] at 30 °C for 18 min. The conversion of each substrate was analyzed by LC-ESI-MS/MS (Agilent 1290 Infinity System equipped with an MS Agilent 6460 Triple Quad Detector). The separation was conducted on a Zorbax Eclipse Plus C18 column (1.8 μm, 2.1 × 50 mm, Agilent Technologies) at 30 °C in the isocratic mode using an ACN/H_2O/HCOOH [60:40:0.1 (v/v/v)] mobile phase at a 0.4 mL/min flow rate. The MS signals were collected in the positive single-ion monitoring mode (303.3, 308.3, 301.3, and 305.3 m/z signals for $[M + H]^+$ of 17-MT, 2,2,4,6,6-d_5-17-MT, methandienone (MTD), and 2,4,6,6-d_4-MTD, respectively, as well as 291.3, 295.3, 289.3, and 292.3 m/z for $[M + H]^+$ of DHT, 1,16,16,17-d_4-DHT, 1-testosterone (1-TE), and 16,16,17-d_3-1-TE, respectively); see the Supporting Information, Table S1 and Figure S2. The quantitation of analytes was conducted according to a previously established protocol.

Bioinformatics and Computation. *Phylogenetic Tree.* For template sequences, KstD1 (WP_020909157) from *R. erythropolis* SQ1 and AcmB (WP_154715887) and AcmB2 (WP_067169324) from *S. denitrificans* Chol-1, we run protein blast using the Protein BLAST algorithm (NCBI) for 5000 hits. From the alignment, we selected sequences with >90% query cover and >30% identity from different microorganisms. Selected sequences together with templates (81 sequences in total) were further aligned using multiple sequence alignment (Clustal Omega).[30] The obtained phylogenetic tree was visualized using FigTree v1.4.4. The sequence alignment was analyzed with Jalview.[31]

Model Setup. The structure of chain A of AcmB in complex with the reaction product (androsta-1,4-diene-3,17-dione, ADD) available in the PDB (code 7P18) was used to prepare the model for further simulations. ADD was replaced with other ketosteroids, that is, AD, 17-methyltestosterone (17-MT), or dihydrotestosterone (DHT), with the use of the Kabsch algorithm.[32] The protonation states of the titratable amino acids were determined with propKa3.1[33,34] for pH 6.5. The missing molecular mechanics (MM) charges for ketosteroids were obtained with Gaussian16[35] at the B3LYP/6-31G(d,p) level of theory.[36] The FAD parameters were taken from RESP ESP charge DataBase (R.E.DD.B).[37] The charge of the protein, ketosteroid, and FAD combined was +1, which was neutralized by the addition of one Cl^- ion. The models were soaked with TIP3P water molecules in a 94.4 × 78.8 × 78.4 $Å^3$ box.

MD Simulations. For all MD simulations, the AMBER package with ff03 forcefield was used.[38,39] Each model was first optimized and then heated from 0 to 303 K with the *NVT* ensemble. Then, 100 ps with *NPT* conditions were simulated to equilibrate the system. Finally, 60 ns of MD simulation with periodic boundary conditions, controlled with a Langevin thermostat were conducted. A cutoff for nonbonding interactions was set to 8.0 Å.

Binding Free Energy and Interaction Energy. The MM-PBSA algorithm[40] was used to estimate the change of Gibbs free energy of substrate binding (ΔG_b). From each trajectory, two ΔG_b estimates were received: total ΔG_b for the whole simulation and ΔG_{best} estimated for the 5 ns simulation with the best geometry parameters of the bound substrate. These results were compared with MM-PBSA analysis of enzyme: substrate MD simulations of KstD1 from *R. erythropolis*,[12] which were analyzed according to the above protocol.

The average interaction energies between the ligand and each amino acid of AcmB were calculated for the MD simulation fragment related to ΔG_{best}. Interaction energies were computed as the difference between the energy of the ligand-amino acid pair and the sum of energies for the separated ligand and residue.

QM/MM MD Simulations. QM/MM modeling was performed with four different substrates: AD, 17-MT, CHON, and DHT. The first one was used not only with the wild type of AcmB but also with mutants (Y118F, Y115F, Y467F, and Y536F) in order to match the kinetic experiments (Figure S3). All of the models were treated according to the procedure described below.

The QM layer consisted of the steroid substrate, Y363, and the FAD fragment (similar to our last study[27]), while other residues were treated with the AMBER forcefield as implemented in the fDynamo library.[41,42] The positions of residues beyond 20 Å from the substrate were fixed. Two antisymmetric combinations of distances were selected as reaction coordinates to describe the transfer of both hydrogen atoms: rC2H-rOH (proton abstraction) and rC1H-rNH (hydride transfer). One-dimensional potential energy scans were performed to generate the initial structures for QM/MM MD simulations. The reaction was studied with the umbrella sampling method;[43] at each window of simulation, a parabolic penalty potential was added with a force constant of 2500 kJ/$Å^2$ mol. One QM/MM MD simulation consisted of 5 ps of system relaxation and then 20 ps of simulation, the results of which were used for further analysis. The weighted histogram analysis method[44] was used to combine the QM/MM MD results into a distribution function. QM/MM MD simulations were run at the AM1/AMBER level of theory.[45] The obtained profiles were corrected with the energy spline function defined in terms of interpolated corrections.[46−48] Single-point calculations were performed at the B3LYP/6-311++G-(2d,2p)/AMBER level of theory for geometries used for initial structures for QM/MM MD simulations. After that, stationary-state structures were optimized with the Baker algorithm[49] and using the micro−macro iteration scheme[50] at AM1/AMBER and B3LYP/6-31G(d,p)/AMBER levels of theory. Every optimized structure was verified by calculating the Hessian and checking the number of imaginary frequencies. Kinetic isotope effects (KIEs) for DHT and 17-MT were calculated from the definition of the free energy of a state.[51] For each of the stationary states, E:S, TS1, E:I, and TS2, three structures were optimized. Then, the KIEs for every elementary step were calculated as described elsewhere.[52,53]

RESULTS

Bioinformatic Analysis. We have performed an exhaustive analysis of KstD sequences based on multiple sequence alignment. We selected 83 sequences from a diverse set of organisms from over 5000 retrieved sequences that exhibited >90% query cover and >30% identity with respect to KstD1 (WP_020909157) from *R. erythropolis* SQ1, AcmB (WP_154715887), and AcmB2 (WP_067169324) from *S.*

denitrificans (see Supporting Information). The phylogenetic analysis showed that the analyzed sequences can be divided into five general clades, which we named after representative KstDs with characterized substrate specificity such as ReKstD1-like from *R. erythropolis*, PsKstD3-like and PsKstD4-like from *Pimelobacter simplex*, and finally AcmB-like and AcmB2-like from *S. denitrificans* (Figures 2 and S4, Table S2).
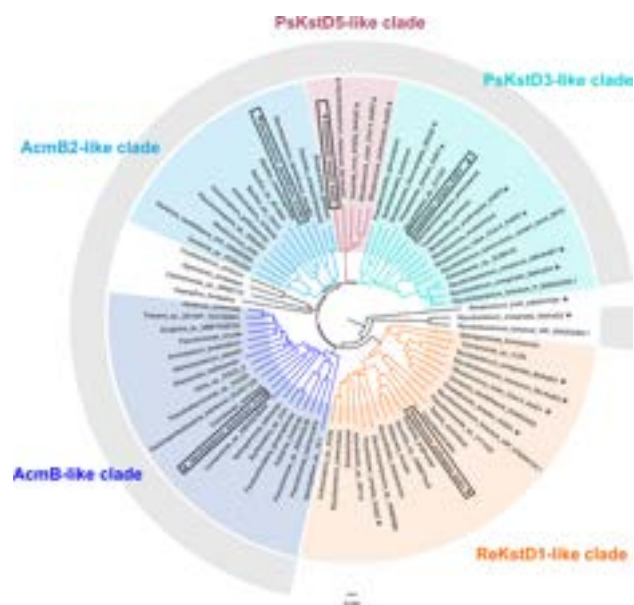
Figure 2. Phylogenetic analysis of the KstDs from *Sterolibacterium denitrificans* and representative orthologues in other species. The tree was rooted to Δ⁴-KstD from *Rhodococcus josti*.[57] The scale length was set as 0.04; the asterisk marks enzymes with identified activity; KstDs in frames indicate enzymes representative of the clade with characterized structure or substrate specificity. The gray rim indicates sequences with an additional "loop".

KstDs with shorter sequences lacking the "loop" aggregated solely into the ReKstD1-like clade which contained mostly Actinobacteria such as KstD from *Mycobacterium* spp.,[54] KstD1 from *R. ruber* Chol-4,[55] and *P. simplex* KstD2 and KstD4.[10] The two clades represented by *P. simplex*, PsKstD3-like and PsKstD5-like, contained several well-characterized dehydrogenases, such as KstD4 from *Mycobacterium smegmatis*,[56] and several enzymes from *Rhodococcus* spp., *P. simplex*, *Streptomyces flavovariabilis*, and *Nocardia nova*. Both clades were composed of KstDs from Actinobacteria with rare examples of proteobacteria. Meanwhile, in clades AcmB- and AcmB2-like, we could not identify any other biochemically characterized enzymes. The AcmB-like clade was mostly composed of sequences from proteobacteria, while in the case of the AcmB2-like clade, we identified numerous members of Actinobacteria and less numerous Firmicutes and proteobacteria. Interestingly, the KstD2 from *M. smegmatis* was classified as a phylogenetically different sequence, indicating the existence of yet another phylogenetic clade that was not represented by our selection of sequences.

Remarkably, if the same analysis is conducted only for the "loop" region, we obtain a very similar phylogenetic tree (Figure S4). This indicates that the loop region is an important determinant of KstDs sequence diversity.

AcmB Crystal Structure. In the course of the crystallization experiments, a single yellow crystal of wild-type AcmB was obtained in the trigonal form with two molecules in the asymmetric unit, denoted as chains A and B in the PDB structure. Although a few hydrogen bonding interactions exist between the molecules, the size-exclusion chromatography experiments and atomic force microscopy imaging[26] showed that AcmB occurs as a monomer in solution (Figure S5). The high similarity of monomers A and B was proven by the low value of rms deviations for Cα atoms (0.174 Å). The subtle differences between chains A and B are correlated with the different number of bound ligands. Analysis of the crystal structure proved that a FAD molecule is present in both monomers, while an ADD ligand occupies the active site of molecule A. The electron density maps have very good quality for almost the entire protein molecules, except for short disordered regions at the N-terminus of both monomers (MSI of A and MSIE of B), as well as for side chains of a few residues (L194A, M180B, L194B, and L198B). These regions were excluded from the refinement. Data collection and refinement statistics are summarized in Table S3.

Overall Structure. AcmB has an α/β fold formed by three tightly packed domains. The molecule has an elongated shape with dimensions of 52 × 55 × 72 Å³; its volume equals 69,445 Å³ and it has a surface area of 21,035 Å². The overall architecture is very similar to KstD1 with a RMS deviation of 0.927 Å calculated for Cα. The largest, FAD-binding domain comprises helices H2-H5, H10-H13, and H21-H22 (Figure 3A). This domain also contains three β-sheets: the largest one, antiparallel β-sheet A (B1-B2, B10-B13, and B23), antiparallel β-sheet B (B3-B5), and small two-stranded β-sheet E (B14-B21). The second catalytic domain (A320-T498) consists of a six-stranded antiparallel β-sheet F (strands B15-B20) located in the core of the molecule decorated with seven helices (H14-H20). The third domain (Y153-R204) comprises a unique part, not observed in the KstD1 structure, which we believe serves as a region responsible for the association of the enzyme to the cell membrane.

Putative Membrane-Associated Domain. Initially, we expected that the fragment Y153-R204 would be disordered since this region is missing in homologous proteins of microbial origin and our homologous modeling did not predict any secondary structure elements in the long loop.[26] However, our structural data showed for the first time that the domain Y153-R204 of AcmB not only forms secondary structures but also closes around the product, further tightening its binding (Figure 3E). The domain is composed of four helices (H6-H9) together with an antiparallel β-sheet D (B7-B8). The hypothesis on the function of this domain was corroborated by the prediction of membrane-embedded residues of 178−198 (H8 and H9) by the PPM 3.0 server (Figures 3C and S6). The predictor classified AcmB as a peripheral protein with a ΔG_{transfer} of −12 kcal/mol. This important difference from KstD1 may also result in the concomitant stabilization of the ligands (see below). As a result, we named the Y153-R204 fragment the putative membrane-associated domain.

AcmB is an amphipathic protein located in the cytoplasm or periplasmic space, and it binds to the cytoplasmic membrane through weak dispersion interactions.[9,26] We also observed a high aggregation tendency of AcmB, which was not reported in the case of KstD1.[26] Thus, it could be hypothesized that an additional domain is responsible for anchoring the protein to
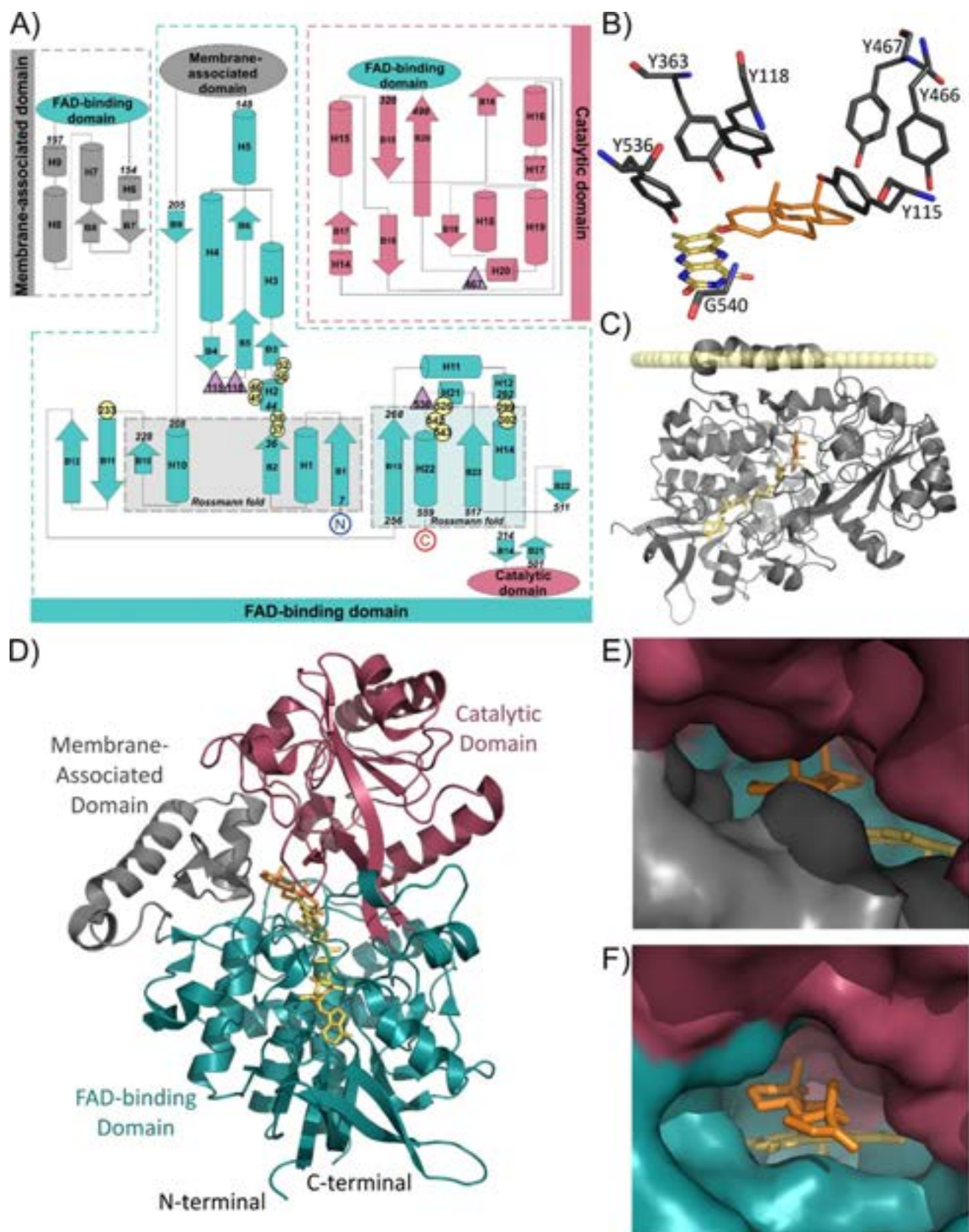
Figure 3. (A) Topology of AcmB chain A and its overall structure; residues involved in FAD coordination are marked as yellow circles, residues involved in catalysis are marked by purple triangles; (B) AcmB active site with catalytically crucial residues, FAD in yellow, ADD in orange; (C) alignment of AcmB at the membrane; (D) overall fold of the AcmB molecule with the FAD-binding domain (green), putative membrane-associated domain (gray), and catalytic domain (red); (E) entrance to the AcmB active site; (F) entrance to the active site of KstD1.

the bacterial cytoplasmic membrane. Probably, the binding of the protein to the membrane occurs through a longer amphipathic α-helix (H9) (Figure S5A). Since helix H9 is parallel to the lipid bilayer, its hydrophobic part is separated from the water environment, whereas the polar surface of the helix interacts with the water phase and negatively charged phosphoryl groups, and the membrane-associated domain can penetrate the area of the polar bilayer groups. Such an orientation of the enzyme would also locate the active site near the membrane, allowing steroids to dissolve in it. The association of proteins with the membrane is quite common among steroid-transforming enzymes such as cholesterol oxidase.[58]

Furthermore, the analysis of the B factors and our MD modeling showed that the membrane-associated domain is the most flexible and mobile element of the structure of AcmB (Figures S6B and S7). These facts indicate that this domain or a part of it could function as a lid that closes the active site after the substrate binds to the protein. However, unequivocal confirmation of this hypothesis would require the determination of the apo-enzyme structure.

To better understand the role of the unique domain Y153-R204, the surface of the protein was analyzed in terms of its electrostatic potential. In Figure S8, the surface of AcmB is shown with the positively charged amino acids (arginine, lysine, and histidine) marked in blue, and the negatively charged amino acid residues (aspartic acid and glutamic acid) are colored red. The highest concentration of positively charged amino acids is observed within the putative membrane-associated domain (marked with a black dashed line). This observation further supports the suggested role of domain Y153-T204 in the interaction with the lipid bilayer of membranes.

FAD-Binding Site. The FAD-binding domain comprises regions E4-A152, K205-W319, and K499-K561 (Figure 3A) and presents a conserved fold previously observed in KstD1. The main part of the FAD-binding domain adopts the supersecondary structure, the characteristic of the Rossmann fold (Figure S9). Its arrangement differs from a basic topology, and it is a typical variation of the fold noticed for dehydrogenases.[59] The fold that binds dinucleotides such as FAD involves two Rossmann binding motifs, which typically form a six-stranded parallel β-sheet flanked by α-helices. As was observed in KstD1, the first half of the Rossmann fold in AcmB (B1-H1-B2-B10-B10) is very similar to the basic topology, while the second half (B13-B13-B23-H22) is slightly different from the basic arrangements since the third β-strand is not present. Besides the missing β-strand, the variation of the fold observed for dehydrogenases includes a three-stranded β-meander (B11-B12 and part of B13) connecting two halves of the Rossmann fold, instead of the crossover α-helix, as in the AcmB structure. In addition to the Rossmann fold, the FAD-binding domain is decorated with several secondary structure elements.

Due to the numerous insertions within the previously described Rossmann motif (especially between B2 and H10), it can be concluded that AcmB belongs to the GR2 subfamily of the structural family of glutathione reductase (GR). The aforementioned insertions make alignment of the amino acid sequence possible only for the first 30 amino acid residues from the N-terminus. The most conserved structural motif of the GR family is the sequence xhxhGxGxxGxxxhxxh(x)$_8$hxhE-(D) (where x is any amino acid and h is a hydrophobic amino acid), which is part of the Rossmann fold and located at the N-

terminal part of the protein.[59] The main differences between the basic version of the analyzed fragment and both known KstD structures are the presence of alanine residue instead of the glycine residue at the position of the third conserved glycine and the number of amino acids between the two hydrophobic amino acids in the final part of the sequence, which is seven for AcmB (xhxhGxGxxAxxxhxxh(x)$_7$hxhE(D): ^{10}VIVVGSGAGAMLAAARAHDLGLSVLVVE[37]). This fragment forms the initial region of the Rossmann motif (B1-H1-B2) (Figure S9). There are hydrophobic interactions between α-helix and β-sheets. The negatively charged, conserved glutamic acid residue that terminates the motif (E37) forms hydrogen bonds with the hydroxyl groups of adenosine monophosphate ribose. The FAD coenzyme in an extended conformation occupies the elongated cavity of the largest domain (Figures 3D and S10A). The adenine end of the cofactor reaches the floor of the cavity formed by H13, B2, and B11. The planar isoalloxazine ring binds at the edge of the FAD-binding domain and reaches the catalytic domain. Similar to the position revealed by the KstD1 structure, the isoalloxazine ring interacts with the H2/B3 loop of the FAD-binding domain and the catalytic domain. The observed yellow color of the crystals suggests the presence of noncovalently bound FAD in an oxidized state. The FAD molecule is stabilized at the active site by a network of hydrogen bonds formed by residues: E37, K38, T45, S46, G50, A52, F233, N299, D302, N526, T542, and L543, as shown in Figure S10A. The analysis of AcmB-FAD interactions also reveals numerous interactions of the π-alkyl (V13, K38, I496, and F338), π-sigma (A270 and L543), and alkyl−alkyl (A294 and I496).

Active Site of AcmB. The active site of the AcmB binds the flavin part of the FAD and the 3-ketosteroid (here, ADD product) and is located in the area of contact of all three domains. Figure S10 shows that most of the amino acids surrounding ADD are nonpolar. The axial methyl groups on C10 and C13 carbons face the catalytic domain, and the sterane A ring is almost parallel to isoalloxazine (Figure 3B). The active site contains six conserved tyrosine residues: Y115, Y118, and Y536 from the FAD-binding domain as well as Y363, Y466, and Y467 from the catalytic domain. Similar to the active site of KstD1 from *R. erythropolis*, Y363 is positioned close to the C2 atom of the steroid, and we assume that it is in the tyrosyl anion form. In this form, it can act as a catalytic base during substrate activation (see below). Y363 is accompanied by Y118, which, based on the short distance of their tyrosyl O atoms (2.8 Å), forms an H-bond with Y363. The Y536 hydroxyl and the G540 NH of the peptide bond form hydrogen bonds with carbonyl oxygen at carbon C3 of ADD (Figure 3B). Additionally, the steroid core forms π-alkyl or alkyl−alkyl interactions with the amino acids A52, F338, Y363, and A539.

These three tyrosines (Y363, Y118, and Y536) are directly involved in catalysis and are connected via water (HOH701) with a proton-relay system composed of Y115, Y467, and Y466. Our bioinformatic analysis showed that these six tyrosines are highly conserved in the majority of the analyzed KstD sequences (65%), while the second most common motif consists of only catalytic tyrosines (Y118, Y367, and Y536) encountered in 24% of the 82 analyzed sequences (Figure S3 and Table S2). This type of active-site organization without an extended proton-relay system is present in the other known structure of KstD1 (Figure S11).
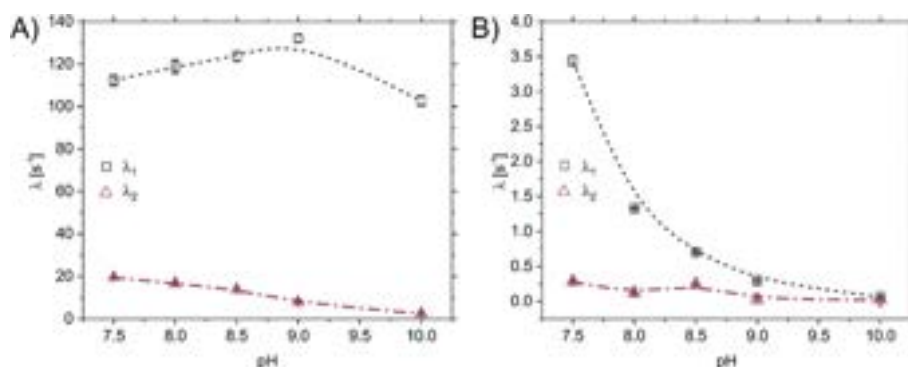
Figure 4. pH optima of the (A) RHR (AcmB^{ox} + AD) and (B) OHR (AcmB^{red} + DCPIP) of AcmB obtained in pre-steady-state kinetics. λ_1 and λ_2 represent eigenvalues of the double-exponential model used to fit the observed kinetics. The error bars represent the standard deviation of the experiment.

Table 1. Estimated Free Energy of the Binding (ΔG_b) of Ketosteroids to the AcmB Active Site and Average Interaction Energies (IEs) between Different Substrates and AcmB or the Membrane-Associated One

substrate	total ΔG_b [kcal/mol]	ΔG_{best} [kcal/mol]	IE [kcal/mol]	IE with res. 153−204 [kcal/mol]
androst-4-en-3,17-dione (AD)	−4.1 ± 0.2	−6.1 ± 0.1	−40.1 ± 0.2	−4.93 ± 0.05
cholest-4-en-3-one (CHON)	**−10.0±0.3**	**−12.4±0.1**	**−61.4±0.2**	**−15.68±0.08**
dihydrotestosterone (DHT)	−6.6 ± 0.1	−7.0 ± 0.1	−41.9 ± 0.1	−10.63 ± 0.06
progesterone (PRG)	−6.3 ± 0.2	−7.3 ± 0.1	−42.0 ± 0.2	−7.95 ± 0.04
17-methyltestosterone (17-MT)	−7.6 ± 0.1	−8.7 ± 0.1	−46.3 ± 0.2	−4.53 ± 0.04
testosterone propionate (TP)	−4.9 ± 0.3	−7.1 ± 0.1	−49.2 ± 0.2	−10.02 ± 0.07
6-dehydrotestosterone acetate (6-DHA)	−3.8 ± 0.5	−7.2 ± 0.1	−53.4 ± 0.2	−11.10 ± 0.07
diosgenone (DGN)	−3.4 ± 0.3	−5.7 ± 0.2	**−60.0±0.2**	**−13.48±0.07**

AcmB Kinetic Mechanism. Kinetic experiments have confirmed that the reaction catalyzed by AcmB proceeds according to Ping−Pong bi−bi mechanisms, as recently reported by us for KstD1 from *R. erythropolis*.[27] AcmB exhibited a high affinity for progesterone with a K_{mA} value of 4.4 ± 0.3 μM and a 20-fold lower affinity for DCPIP (K_{mB} 79.3 ± 6.9 μM). These data are consistent with the apparent kinetic parameters reported previously.[11] The established V_{max} of the enzyme was 33.4 ± 1 μM min^{-1}, which corresponds to k_{cat} of 21.3 s^{-1}.

pH Optimum of the Half-Reactions. Under the Ping−Pong bi−bi mechanism, it is possible to independently study the kinetics of the RHR, that is, binding of the steroid substrate to the enzyme active site, oxidation, and reduction of the enzyme's FAD followed by the release of the product, and OHR, that is, binding of DCPIP to the active site, oxidation of FADH$^-$ to FAD, and release of the reduced DCPIPH$_2$. Therefore, we employed pre-steady-state kinetics to investigate the pH dependence of both the RHR and OHR. The aim was to explain our previous observation that the enzyme exhibits two reaction pH optima, one at pH 6.5 when the reaction was conducted with DCPIP and the second at pH 8.5 when the reaction was conducted with PMS or PMS and DCPIP.[11] Initially, we hypothesized that under acidic conditions, the reaction catalyzed by AcmB might proceed according to a different mechanism from that of the other KstDs. However, later on, we suspected that it is the sluggish OHR that is responsible for the observed pH optimum of the steady-state reaction.

To test that, we have examined the pH dependence of the RHR and OHR in the range of 7.5−10.0. The kinetic curves were fitted with a double-exponential model, yielding two eigenvalues λ_1 and λ_2 (Figure 4). For the RHR, the λ_1 was in

the range of 100 to 130 s^{-1} with a slight optimum at pH 9.0, while λ_2 was in the range of 0−20 s^{-1} and linearly decreased from pH 7.5 to 10.0. In the OHR with DCPIP, the observed eigenvalues were much lower than those observed for the RHR and the highest λ_1 was observed at 7.5, indicating the pH optimum at lower pH values, which were unfortunately beyond the experimentally accessible range due to the rapid aggregation of the enzyme.

Substrate Specificity. Recently, we have reported that the kinetics of the AcmB point to cholest-4-en-3-one (CHON) as a native substrate.[12] AcmB differs from KstD1 by the presence of a 40 amino acid fragment (153−204) that was formerly referred to as a "loop" but that in fact, forms a putative membrane-associated domain at the entrance of the active site (Figure 3E,F). Besides anchoring AcmB to the membrane and positioning the active-site entrance toward it, this domain seems to be involved in the binding of substrates with extended C17 substituents, which stick out of the KstD1 active site. We assume that the presence of this fragment may be responsible for the unique substrate specificity of AcmB with respect to enzymes devoid of this sequence. To check this hypothesis, we decided to conduct a series of MD simulations and estimate the free energy of binding for AcmB: steroid complexes, as well as calculate the interaction energies of substrates with the whole protein and membrane-associated domain (Table 1).

Theoretical predictions of ΔG_b obtained for AcmB turned out to be similar to those published for KstD1[12] (Tables S4− S6). The estimated free energy of binding for the effective enzyme−substrate complexes (ΔG) for all analyzed substrates ranged between −10 and −3.4 kcal/mol (Table S4), but the ΔGs of the substrates with a degraded C17 substituent were in the range of −7.6 to −3.7 kcal/mol. The best-bound substrate turned out to be cholest-4-en-3-one (−10 kcal/mol), while
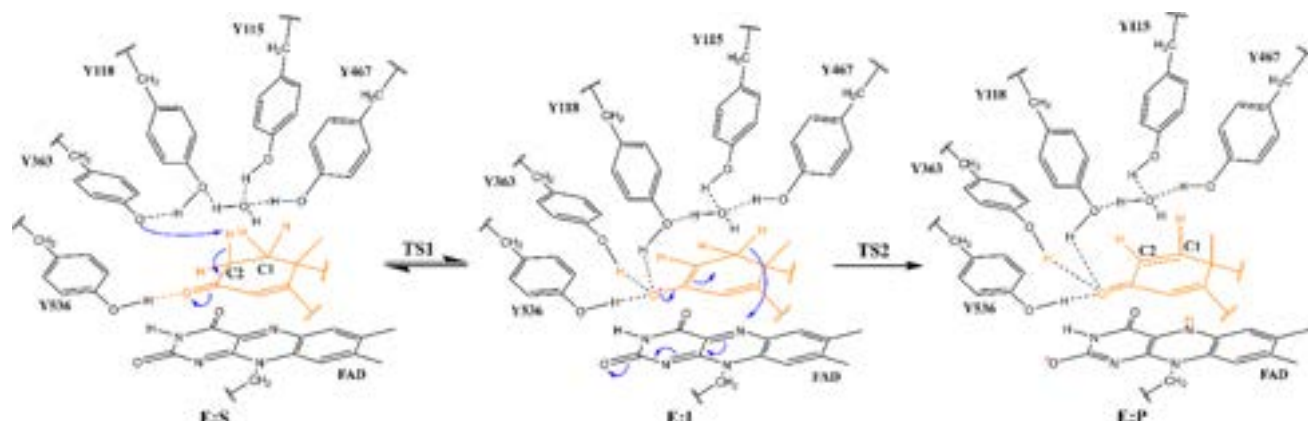
Figure 5. Scheme of the mechanism for the reaction catalyzed by AcmB.

diosgenone (DGN) was bound only with ΔG of −3.4 kcal/mol. The ΔG_{best} calculated for the 5 ns of the best **E:S** geometries led to similar conclusions (range from −8.7 to −6.1, −12.4 kcal/mol for CHON). The analogical analysis conducted for the KstD1 from *R. erythropolis* yielded values of ΔG_b in the range of −9.3 to −3.8 kcal/mol. Cholest-4-en-3-one was not different from other substrates (ΔG_b of −6.3 kcal/mol), but diosgenin exhibited better stabilization than for AcmB (−5.8 kcal/mol). These results confirm the preferential binding of cholest-4-en-3-one by the enzyme with an additional "loop" and explain the better apparent activity of KstD1 with DGN compared to AcmB.[12] This result also suggests that the membrane-associated domain plays a role in the differentiation of substrates with an extended C17 substituent (i.e., preferential binding of CHON over DGN).

Similarly, substrates with short C17 substituents had average interaction energy with putative membrane-associated domain in the range of −11.1 to −4.5 kcal/mol, while for cholest-4-en-3-one and diosgenone −15.7 and −13.5 kcal/mol, respectively (Figure S12). Accordingly, the interaction with the whole protein was in the range of −53 to −40 kcal/mol, except for cholest-4-en-3-one and diosgenone, for which it was approximately −60 kcal/mol. Closer analysis indicates that the beginning of the membrane-associated domain is primarily responsible for the stabilization of the substrates. The T156 universally stabilizes both smaller and bigger substrates, while the hydrophobic residues between M164 and A169 are involved in the stabilization of substrates with extended C17 substituents. These results suggest that the membrane-associated domain is involved in the stabilization of enzyme complexes with bigger substrates.

The most important interaction inside the binding site (Figure S13) is between the 3-keto group of the substrates and Y536 and G540, which form H-bond interactions (i.e., in the range of −6.4 to −2.4 kcal/mol) as well as via hydrophobic interactions F338 and A539. Interestingly, due to the lack of a double bond in ring A and the different position of the 3-keto group, DHT exhibits a different binding pattern, forming effective H-bond interaction only with G540 (IE −3.24 kcal/mol). The distance between the C3═O atom and H atom of the Y536 hydroxy group is significantly larger in the case of DHT compared to the native substrate CHON (median of 3.48 Å vs 2.94 Å, respectively).

Reaction Mechanism. The modeling revealed that the reaction mechanism catalyzed by AcmB proceeds according to the classical mechanism postulated in the literature, which was

recently confirmed by our calculations for KstD1 from *R. erythropolis*[27] (Figure 5). The enzyme binds a steroid substrate (**E:S**), so its ring A is positioned almost parallel to the isoalloxazine ring system of FAD. Y536 and G540 form H-bond interactions with the 3-keto group of the substrate, while the deprotonated Y363 is positioned in close vicinity (1.9−2.2 Å) to the 2βH atom, prepared to deprotonate the substrate (Figure S13). Y363 forms a direct H-bond interaction with Y118, which seems to be stabilizing its tyrosyl form. The Y363-Y118 tandem is connected via water to four other tyrosines (Y115, Y463, Y466, and Y467) that provide H-bonding interactions for several water molecules (Figure S14). These tyrosines, together with water molecules, form a proton relay that enables the proton's swift transfer from the active site to the solvent. The water molecule connecting the Y363-Y118 tandem with the proton-relay system is not only present in the crystal structure (HOH701 in chain A), but its position is also stable during most of the MD simulations.

In the first step of the reaction (**TS1**), the proton from the 2βC position is abstracted by a tyrosyl anion (Y363), which results in the formation of an enolate intermediate product **E:I**, as indicated by the shortening of the C2−C3 bond from 1.5 Å in **E:S** to 1.38 Å in **E:I**. At this point, the hydrogen bond network reorganizes, as Y363 and Y118 join the Y536 in H-bond stabilization of the negative charge at the 3-keto group of the steroid enolate. In the second step (**TS2**), a hydride anion from the 1αC position is transferred to the N5 atom of FAD yielding reduced FADH⁻ and 1-dehydro product (**E:P**). The stationary states of the pathway are presented in Figure S15.

As in our previous calculations for KstD1,[27] both barriers are of similar height (13.5 and 15.6 kcal/mol), with **TS1** slightly lower (2.3 kcal/mol) than **TS2** (Table 2 and Figure S18). The **E:I** is very well stabilized (2.7 kcal/mol), and the final product **E:P** is exergonic (−9.2 kcal/mol).

The modeling showed that the energy profile for AD is very similar to that of 17-MT (Figure S20). This was expected because the only structural difference between AD and 17-MT (substitution of C17) is pointed toward the entrance to the active site (Figure S16). However, in the case of DHT, which has no double bond in ring A, we observed an elevation of both barriers and energy of **E:I** by approx. 3 kcal/mol. The saturated A ring of the substrate resulted in a slightly later **TS1** and an earlier **TS2** compared to the geometries obtained with AD (Figures S17−S20) and higher charge separation in the **E:I**. These results are consistent with our steady-state kinetic experiments, which yielded, respectively, 71 and 19% specific

205

Table 2. Free Energy Barriers ($\Delta G^{\ddagger}$ in kcal/mol) for Reactions Catalyzed by AcmB with 17-MT, DHT, CHON and AD and Its Mutants in Reaction with AD as well as Experimental Relative Specific Activities (rSA) with Respect to Reaction Conducted with the WT Enzyme and AD at pH 6.5 with DCPIP

AcmB variant	substrate	TS1	I	TS2	P	rSA [%]
WT 2βH	AD	13.2	2.7	15.5	−9.2	100
WT 2αH	AD	16.3	2.7	15.5	−9.2	
WT	17-MT	13.6	2.8	15.6	−9.2	71
WT	DHT	16.5	5.9	19.0	−10.4	19
WT	CHON	12.8	2.8	14.1	−9.5	n.d.[a]
Y118F	AD	10.8	4.7	21.6	−1.4	trace
Y115F	AD	12.2	5.2	18.9	−0.9	32
Y467F	AD	14.0	5.3	16.7	−6.1	58
Y536F	AD	9.9	9.4	28.7	2.4	0.6
Y363F	AD					0

[a]Activity assay required the addition of cyclodextrin, which makes it incomparable to the assay without one.[12]

activity of AcmB with 17-MT and DHT with respect to AD (Table 2). Interestingly, we obtained the best PES for the AcmB native substrate (Figure S20), CHON. The C−H activation was very similar to AD (Figure S18), yielding only a slightly lower barrier (12.8 kcal/mol). However, the hydride transfer to FAD proceeded with a barrier of only 14.1 kcal/mol. This result suggests that the RHR for CHON should proceed with the highest rate of all investigated substrates. Unfortunately, due to the low water solubility of CHON, it was not possible to directly compare the AcmB dehydrogenation rate of CHON and AD.[12]

It was also observed that 1,2-dehydrogenation can proceed even if a 2β position is substituted by a hydroxy group.[60] This suggested that 3-ketosteroid can also be activated by an abstraction of the equatorial 2αH atom by Y363. We decided to test if such activation yields a kinetically accessible energy profile. To abstract the 2αH atom, Y363 has to break the H-bond with Y118, which in turn forms an H-bond with the 3-keto group (Figure S15). Despite the increased nucleophilic character of the tyrosyl ion, **TS1** is reached at the later stage and the barrier is higher by 3 kcal/mol compared to the barrier of 2βH and also higher than the barrier of hydride transfer (Figure S18). This demonstrates that although the enzyme has a preference for the abstraction of the axial 2βH, the enantioselectivity is not obligatory.

Effect of Mutations. As expected, the Y363F mutation of AcmB that eliminates the tyrosyl residue from the active site renders the enzyme completely inactive (Tables 3 and S7). The second most severe mutation is Y118F, which interrupts the H-bond with Y363. We were not able to determine the activity of the mutant enzyme in a spectrophotometric assay, but we detected the product after overnight incubation. The QM/MM modeling (Figures S22 and S25) showed that the

Y118F mutation facilitates proton removal from the substrate (**TS1** lowered by 2.3 kcal/mol), but the enolate intermediate is less stable and the barrier of the hydride transfer is increased by 6.1 kcal/mol. This can be explained by an increased nucleophilic character of the Y363 ion which is not moderated by the H-bond with Y118, as depicted by a much lower charge on the Y363 oxygen atom in Y118F versus the WT enzyme (q^{APT} −0.9 vs −0.46, respectively). However, the lack of an additional H-bond with Y118 and enolic intermediate's 3-keto group results in increased energy of **E:I** and **TS2** due to worse enolate stabilization (i.e., higher charge of the O atom bound to C3 and at the C2 atom in Y118F than in the WT).

Mutation of the Y536 also has a severe impact on enzyme activity (0.6% of the WT). In our modeling, G540 took over the role of Y536 in the stabilization of substrate binding, forming an H-bond with a 3-keto group [d(C3═O---HN-G540 is 2.1 Å, Figure S24)]. The position of the substrate was very similar to the WT, and the proton abstraction turned out to be very easy (**TS1** of 9.85 kcal/mol, Figure S25). However, upon deprotonation, the conformation of ring A in the substrate changed, shifting the C3═O group away from G540 and breaking the H-bond. As a result, only two tyrosine residues are involved in the interaction with enolate resulting in a small stabilization of the intermediate (by 0.35 kcal/mol) and difficult hydride transfer (**TS2** 35 kcal/mol).

Furthermore, we have assessed mutations of the tyrosines involved in the proton relay using Y115F and Y467F mutants (Figures S21 and S23). These variants exhibited decreased specific activities of 32 and 58% with respect to the WT enzyme. These mutations practically did not change the height of the **TS1** barrier (differences within 1 kcal/mol) but lowered the stabilization of **E:I** by 2.6 kcal/mol and slightly increased the energy of **TS2** (1.1−3.4 kcal/mol). This effect may be associated with the decreased polarity of the active site upon substitution of the tyrosine residue by phenylalanine and with partial disruption of the proton-relay system, which utilizes both Y115 and Y467 (Figure S14).

Finally, we studied the mutation G540P, which aimed at removing the auxiliary H-bond between the main chain amino group and the 3-keto group of the substrate. However, structural interference introduced by the proline resulted in misfolding of the enzyme. Despite this, we were still able to detect trace activity after overnight incubation of the enzyme with the substrate. Unfortunately, we had to assume that the introduced mutation resulted in a severe change in the protein structure, which does not allow for the modeling of this effect nor gives any insight into the structure of the **E:S** complex.

Kinetic Isotope Effect. To gain better insight into the reaction mechanism catalyzed by AcmB, we decided to determine the KIE at pH 6.5 and 8.5 (Table 3). We applied two methods: (i) a direct method that compared enzyme activities with the unlabeled or labeled substrates measured independently under steady-state and substrate saturation conditions and (ii) a competitive method where the equimolar

Table 3. Results of the Experimental KIE Obtained for C1-Substituted DHT and C2-Substituted 17-MT

method	direct method		competitive method	
deuterated atom	C1	C2	C1	C2
	1,16,16,17-d_4-DHT	2,2,4,6,6-d_5-17-MT	1,16,16,17-d_4-DHT	2,2,4,6,6-d_5-17-MT
pH 6.5	1.05 ± 0.04	1.17 ± 0.02	1.50 ± 0.02	1.21 ± 0.01
pH 8.5	1.33 ± 0.08	1.06 ± 0.03	1.23 ± 0.02	1.28 ± 0.01

mixture of the unlabeled and labeled substrate was converted by the enzyme while we analyzed the composition of the product with LC—MS (Figure S25). The former method yields a ratio of k_{cat} in the steady-state which is dependent on both substrate oxidation and enzyme reoxidation, while the latter method also takes into account the differences in K_m of the substrate isotopologues, yielding $^D(V/K)$. We observed a very small KIE at pH 6.5 regardless of the substitution position (1.05 and 1.17, respectively, for substrate deuterated at C1 and C2, Figure S26). At pH 8.5, the observed KIE for C1 substituted substrate turned out to be slightly higher (1.33) compared to the C2 substituted substrate (1.06). The competitive experiments yielded $^D(V/K)$ values slightly higher but still in the range of only 1.2−1.5.

The experimental results on the KIE were confronted with the modeling and theoretical prediction of the intrinsic KIE associated with particular molecular steps, as well as the overall KIE associated with the RHR process estimated with the use of free energy barriers and the Eyring equation[61] (Tables S8−S10).

In the case of 2,2,4,6,6-d_5-17-MT, we predicted a high KIE (Tables S8−S10) associated with the deprotonation of the C2 atom (**S → TS1**, 5.3 ± 0.19) and a transition from the intermediate back to the substrate (**I → TS1**, 3.55 ± 0.20), as well as an inverse KIE resulting from hydride transfer (**I → TS2**, 0.7 ± 0.03).

For C1-substituted 1,16,16,17-d_4-DHT, calculations suggest a high value of KIE for the hydride transfer (**I → TS2**, 4.23 ± 0.03) and a much lower one (close to unity) for the other effects associated with the deprotonation at C2. The overall KIE for the whole RHR estimated with 1 kcal/mol accuracy turned out to be in the range of 1.07−2.14 for C2-substituted 17-MT and 4.42−4.96 for C1-substituted DHT. These results indicate that the experimentally observed kinetics is limited by some other process not associated with the RHR, as the KIE, although still noticeable, is severely decreased.

■ DISCUSSION

The analysis of the AcmB crystal structure reveals that its structure is similar to that reported previously for KstD1 from *R. erythropolis*, with the exception of a 40-amino acid long putative membrane-associated domain that is localized in close vicinity of the active site previously referred to as the "loop". Helixes 8 and 9 of this domain exhibit amphipathic character, enabling its anchoring to the membrane. Theoretical prediction indicates that such anchoring positions the enzyme active site oriented toward the membrane surface, which may facilitate the formation of the enzyme–substrate complex even with highly hydrophobic steroids. This domain also narrows the entrance to the active site and strongly interacts with the extended alkyl sidechain in cholest-4-en-3-one, a native substrate of AcmB.[12] Prediction of $\Delta G_{binding}$ and the energy of interaction supports this hypothesis, especially compared to the results calculated for KstD1, which lacks preference for binding cholest-4-en-3-one over steroids with the degraded C17 substituent. The presence of a loop within the active center of KstD1 that may affect the enzyme-catalyzed reactions was previously reported for several homologous models by Luo et al.[10] The long loops (which could also form elements of secondary structure) may also negatively affect the enzyme activity by hindering substrate access to the active site.[58]

The secondary structure of the putative membrane-associated domain, a combination of β-sheets and α-helices,

turned out to be much more complex than previously predicted by homology modeling.[10,26] The bioinformatic analysis of KstD sequences with known biological functions shows the prevalent presence of this domain in the enzyme class. Furthermore, the sequences differ significantly, forming separate clades. Although the data on the substrate specificity of the representatives of each clade is still not plentiful, we suspect that the differences in the sequence of the putative membrane-associated domain are partially responsible for the reported differences in the substrate specificity. Enzymes belonging to the ReKstD1-like clade all lack a membrane-associated domain and are reported to be inactive or slightly active with C17-substituted steroids.[10,12] We already demonstrated that ReKstD1 converts cholest-4-en-3-on if proper solubilization is provided. Still, ReKstD1 has a definitively higher preference for smaller substrates. Until now, the only characterized members of AcmB and AmcB2-like clads are AcmB and AcmB2, which are both capable of converting C17-substituted substrates, although with different structural characteristics.[12] Moreover, KstD3 from *P. simplex*, a representative of the PsKstD3-like clade, has been reported as active with cholest-4-en-3-one along with several other isoenzymes for which this activity was not observed. Interestingly, a similar phylogenetic tree can be obtained when only sequences of the putative membrane-associated domain are taken into consideration (Figure S23). As a result, we shall assume that the sequence of the putative membrane-associated domain strongly influences the divergence of the KstD family and may be responsible for the different biological roles of particular enzymes. However, this issue needs further research and a reanalysis of enzyme activities for cholest-4-en-3-on with the use of a cyclodextrin solubilizer.

As expected, the substrate specificity is not solely controlled by the putative membrane-associated domain but by the structure of the active site. We have observed differences in ΔG binding even between steroid substrates with relatively similar structures. The sequence alignment analysis of 82 KstDs, as well as the previous report,[6] revealed that many of the residues in the 5 Å radius of the steroid ligand are highly conserved among all clads of the KstDs. The anchoring motive of the 3-ketosteroid group (^{536}YhGhG^{540}) is highly conserved as well as catalytically critical Y363 and Y118. The rest of the tyrosine residues forming the proton-relay systems are less conserved, with Y115, Y466, and Y467 present in 75, 68, and 69% of the analyzed sequences, respectively. Interestingly, the substitution of Y466 and Y467 (predominantly by F) strongly correlates with the lack of the putative membrane-associated domain. The hydrophobic residues that interact with the substrate (A52, F338, and A539) are usually substituted by other residues with similar biochemical properties (G52, L338, and P538 in 86, 55, and 52% sequences, respectively). Our modeling confirmed the structure-based hypothesis that both Y536 and the main chain of G540 are involved in the directional binding of 3-ketosteroids. However, the structure of **E:S** appears to be far less rigid than could be expected for such a bulky substrate, and sometimes, the substrates are bound less tightly by only one of these residues. This is especially true for DHT, which, due to the different conformation of the ring A from that in Δ^4-steroids, does not form a very strong H-bond interaction with Y536 such as AD or CHON. The slightly less effective stabilization of enolate is additionally compounded by the lack of a double bond conjugated with the C3 keto group that results in higher charge separation observed in **E:I** when
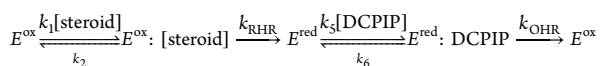
compared to Δ^4-unsaturated substrates. These two factors seem to be responsible for the overall increase of ΔG for both barriers and the intermediate product (**E:I**). Although we were only able to verify these calculations with steady-state kinetics, which revealed a 5-fold slower oxidation of DHT with respect to AD, we can assume that in the case of DHT, it is the RHR which predominantly controls the observed enzyme activity instead of the OHR, as is the case with the other substrates which are oxidized faster. Therefore, we propose that this may explain the observed differences in KstD specificity toward solely Δ^4-3-ketosteroid.[6] The modeling also confirmed our previous kinetic analysis which pointed out at cholest-4-en-3-one as a native substrate.[12] Not only is CHON preferentially bound by the enzyme (the lowest total and the best ΔG and IE) but also its RHR is associated with the lowest barriers for both C–H activation and H$^-$ transfer.

Our combined kinetic and modeling studies confirmed that, despite its peculiarities and seemingly different pH of operation from the majority of KstDs, AcmB catalyzes Δ^1-dehydrogenation according to the Ping–Pong bi–bi kinetics, while the RHR proceeds according to the classical two-step mechanisms. The QM/MM MD modeling confirmed an E2cB mechanism for RHR. We have shown that the reaction can proceed under simulated slightly acidic pH of 6.5, provided Y363 is in a deprotonated state, and the obtained free energy surfaces (FESs) are very similar to those reported previously for the dehydrogenation of 17-MT and DHT by KstD1 from *R. erythropolis*. We have demonstrated that the stereoselectivity of activation at the C2 atom is mostly kinetic with the abstraction of the 2βH atom approximately 6.5-fold faster than 2αH, which is in qualitative agreement with the 10% yield of 2β-hydroxy-androstenedione reported by Hayano et al. over 60 years ago.[60] We have further supported this conclusion by showing how deuteration at the 2β position can decrease the enantioselectivity of C–H activation in 1,2-hydrogenation catalyzed by AcmB in D$_2$O.[62]

There are two essential questions arising from our experiment with isotope-labeled substrates. The first question is why does isotopic substitution at any of the C1 or C2 positions result in a measurable KIE? Second, if RHR is the only isotope-sensitive part of the reaction to such substitution and if the OHR is indeed much slower than the RHR (as our preliminary pre-steady-state kinetics suggests), should the KIE be observable at all under the steady-state conditions?

The answer to the first question is related to the shape of the free energy profile obtained for AcmB. By considering it within the precision of computational methods (ca. 1 kcal/mol), we found that both substitutions, at 2β and 1α positions, significantly reduced the reaction rate of the RHR. The free energy barriers are so close to each other that both of them have a significant impact on the RHR rate. A detailed explanation of a similar situation can be found in our previous paper related to KstD1 from *R. erythropolis*.[27]

Investigation of the second question requires an analysis of the kinetic equations for the Ping–Pong bi–bi mechanism. In our case (only RHR-sensitive for isotopic substitution), the expressions for steady-state k_{cat} and KIE(k_{cat}) are as follows[63]

$$E^{ox} \underset{k_2}{\overset{k_1[steroid]}{\rightleftharpoons}} E^{ox}: [steroid] \xrightarrow{k_{RHR}} E^{red} \underset{k_6}{\overset{k_5[DCPIP]}{\rightleftharpoons}} E^{red}: DCPIP \xrightarrow{k_{OHR}} E^{ox}$$

$$k_{cat\ steady\text{-}state} = \frac{k_{RHR} \cdot k_{OHR}}{k_{RHR} + k_{OHR}}$$

$$KIE_{steady\text{-}state} = \frac{KIE_{RHR} + \frac{k_{RHR}}{k_{OHR}}}{1 + \frac{k_{RHR}}{k_{OHR}}}$$

As we were not able to obtain an experimental value of the KIE for the RHR due to the too-high rate of the process, it was not possible to determine the experimental ratio of k_{RHR}/k_{OHR}. Still, we could use our computational estimations, so in Table 4, we presented calculated values of KIE$_{steady\text{-}state}$ in a function of potential KIE$_{RHR}$ (in the range of 1.07–4.96) and the ratio between k_{RHR} to k_{OHR}.

Table 4. Calculated Values of KIE$_{steady\text{-}state}$ in the Function of KIE$_{RHR}$ and k_{RHR}/k_{OHR}

KIE$_{RHR}$	k_{RHR}/k_{OHR}				
	1	10	25	50	100
1.07	1.04	1.01	1.00	1.00	1.00
1.50	1.25	1.05	1.02	1.01	1.01
2.00	1.50	1.09	1.04	1.02	1.01
3.00	2.00	1.18	1.08	1.04	1.02
4.00	2.50	1.27	1.12	1.06	1.03
4.96	2.98	1.36	1.15	1.08	1.04

Although our analysis is approximate, it demonstrates that it is possible to observe the KIE in the steady-state experiment, even when the process related to the OHR is 100 times slower than the RHR. Based on our preliminary pre-steady-state kinetics, we expect that at an optimal pH of 6.5, the OHR should be at least 10 times slower than the RHR, which would explain the experimental KIE in the range of 1.05–1.33. We have seen a similar masking effect in our recent study of KstD1, but k_{OHR} and k_{RHR} were estimated as of the same magnitude.[27] As a result, we observed higher values of the steady-state KIE for KstD1 (1.28 for 17-MT and 1.5 for DHT) than for AcmB (1.06–1.17 for 17-MT and 1.05–1.33 for DHT).

We were also able to investigate the effects of Y to F mutations in the active site of AcmB on the free energy profile of the RHR. Our mutations introduced a minimal disturbance to the protein structure, only removing the tyrosyl OH group. These mutations can be divided into two groups, those targeting the proton-relay system (Y115F and Y467F) and those directly interfering with the substrates' activation and enolate stabilization (Y536F, Y363F, and Y118F). Our bioinformatic analysis showed (KstDs from ReKstD1-like clade) that tyrosines of the proton-relay system are not obligatory for KstD activity. Their substitution to phenylalanine results in a moderate decrease (2–3 fold) in the specific activity and slightly increases the barrier (**TS2**) of the hydride transfer to FAD. On the other hand, any mutation of tyrosines involved in catalysis leaves KstD either unable to activate the substrate (Y363F) or, while making the abstraction of the 2Hb proton easier, prohibitively increases the barrier of hydride transfer, rendering KstD virtually inactive.

Finally, we have shown that the RHR proceeds at a constant rate in a wide pH range (7.5–10) with a slight optimum at pH 9.0. Based on our preliminary pre-steady-state, it seems that the OHR with DCPIP is slower than the RHR and, as a result, controls the steady-state kinetics and determines a steady-state pH optimum at 6.5. Unfortunately, unlike KstD1 from *R. erythropolis*, AcmB turned out to be a poor subject for pre-steady-state kinetics. Its RHR under substrate-saturated

208

conditions proceeds too fast for the stopped-flow equipment, while the concentrated, reduced enzyme exhibited fast aggregation below pH 7.5, preventing the investigation of OHR kinetics at 6.5.

CONCLUSIONS

We determined the structure of KstD from *S. denitrificans* (AcmB), demonstrating for the first time the structure of the so-called "loop" which is a characteristic motif for the majority of the known KstD sequences. This putative membrane-associated domain may be responsible for anchoring the enzyme to the cytoplasmic membrane, positioning the enzyme active site toward the source of the substrate, as well as stabilization of the **E:S** complex with C17-substituted 3-ketosteroids. With QM/MM MD modeling and kinetic studies, we confirmed that the 1,2-dehydrogenation catalyzed by AcmB proceeds according to the Ping−Pong bi−bi mechanism, while the RHR proceeds according to the accepted two-step elimination mechanism. We showed that the OHR, not the RHR, is responsible for the low values of the KIE observed in the steady-state experiments as well as peculiar pH optima. As in the case of KstD1 from *R. erythropolis*, the nature of FES is responsible for the KIE observed for both C1- and C2-labeled substrates, which once again demonstrates that our findings are general for KstD class. Our modeling also provided quantitative insight into the role of active-site tyrosines, the influence of the Δ^4-double bond in the steroid on its activity, and corroborated experimental evidence on nonobligatory enantioselectivity during C−H activation.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.biochem.2c00576.

> Extended experimental procedure including site-directed mutagenesis, kinetic isotope effect-competition method, and QMMM model setup; phylogenetic tree for "loop" sequences of KstDs and six-tyrosine motif analysis; diffraction data collection and refinement statistics; size-exclusion chromatograms of AcmB; figures and mobility analyses of the putative membrane-associated domain; figures of AcmB structure; binding sites; details on MM-PBSA $\Delta G_{binding}$ and IEs for AcmB and KstD1; details on MD trajectories; figures of stationary states of all mechanisms and figures depicting PES profiles; detailed data on activities of mutants; and experimental and theoretical KIEs (PDF)

> PDB files of representative structures optimized at the B3LYP/AMBER level of theory and sequence alignment of 82 KstDs (ZIP)

> Population analysis (ZIP)

AUTHOR INFORMATION

Corresponding Author

Maciej Szaleniec − *Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, 30-239 Kraków, Poland;* ⓞ orcid.org/0000-0002-7650-9263; Email: maciej.szaleniec@ikifp.edu.pl

Authors

Patrycja Wójcik − *Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, 30-239 Kraków, Poland*

Michał Glanowski − *Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, 30-239 Kraków, Poland*

Beata Mrugała − *Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, 30-239 Kraków, Poland*

Magdalena Procner − *Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, 30-239 Kraków, Poland; Jerzy Maj Institute of Pharmacology Polish Academy of Sciences, 31-343 Kraków, Poland*

Olga Zastawny − *Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, 30-239 Kraków, Poland*

Monika Flejszar − *Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, 30-239 Kraków, Poland; Department of Physical Chemistry, Faculty of Chemistry, Rzeszow University of Technology, 35-959 Rzeszów, Poland;* ⓞ orcid.org/0000-0001-8014-3543

Katarzyna Kurpiewska − *Faculty of Chemistry, Jagiellonian University, 30-387 Kraków, Poland*

Ewa Niedziałkowska − *Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, Virginia 22908, United States*

Wladek Minor − *Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, Virginia 22908, United States;* ⓞ orcid.org/0000-0001-7075-7090

Maria Oszajca − *Faculty of Chemistry, Jagiellonian University, 30-387 Kraków, Poland*

Andrzej J. Bojarski − *Jerzy Maj Institute of Pharmacology Polish Academy of Sciences, 31-343 Kraków, Poland;* ⓞ orcid.org/0000-0003-1417-6333

Agnieszka M. Wojtkiewicz − *Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, 30-239 Kraków, Poland*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.biochem.2c00576

Author Contributions

P.W. purified and crystalized the enzyme, determined and refined the enzyme structure, developed LC−MS methods, conducted kinetics (pH, mutated variants, and KIE), analyzed results, coauthored and edited the manuscript, and created figures. M.G. conducted all calculations, analyzed results, developed discussion, coauthored the main text, edited text, and created figures. B.M. refined the structure of AcmB and coauthored the manuscript. M.P. conducted stopped-flow steady-state and pre-steady-state kinetics, and coauthored the Methods section. O.Z. developed AcmB mutant variants and expressed and purified the enzymes. M.F. conducted stopped-flow steady-state Ping−Pong kinetics and analyzed the results. K.K. refined the structure of AcmB and coauthored the manuscript (crystallographic sections). M.O. provided assistance in pre-steady-state kinetics and edited the manuscript. A.B. supervised P.W., provided funding, and edited the manuscript. E.N. developed conditions for enzyme crystallization and edited the manuscript. W.M. oversaw initial experiments and edited the manuscript. A.M.W. purified the enzyme, conducted the bioinformatic analysis, coauthored the

bioinformatic analysis section, and created figures. M.S. designed the study, provided funding, supervised M.G, P.W., M.P., O.Z., and M. F., analyzed and curated all data, coauthored, and edited the manuscript. The manuscript was written with the contributions of all authors. All authors have approved the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ ABBREVIATIONS

AcmB, anaerobic cholesterol metabolism enzyme B; KstD, 3-ketosteroid dehydrogenase; AD, androst-4-en-3-one; DHT, dihydrotestosterone; 17-MT, 17-methyltestosterone; PRG, progesterone; 6-DHA, 6-dehydrotestosterone acetate; TP, testosterone propionate; DGN, diosgenone; CHON, cholest-4-en-3-one

■ REFERENCES

(1) Hosta-Rigau, L.; Zhang, Y.; Teo, B. M.; Postma, A.; Städler, B. Cholesterol - A Biological Compound as a Building Block in Bionanotechnology. *Nanoscale* **2013**, *5*, 89−109.

(2) Costa, S.; Zappaterra, F.; Summa, D.; Semeraro, B.; Fantin, G. Δ1-Dehydrogenation and C20 Reduction of Cortisone and Hydrocortisone Catalyzed by Rhodococcus Strains. *Molecules* **2020**, *25*, 2192.

(3) Tong, W.; Dong, X. Microbial Biotransformation: Recent Developments on Steroid Drugs. *Recent Pat. Biotechnol.* **2009**, *3*, 141−153.

(4) Khan, M.; Lee, F.; Lee, H. J. Synthesis and Pharmacology of Anti-Inflammatory Steroidal Antedrugs. *Chem. Rev.* **2008**, *108*, 5131−5145.

(5) Zhang, H.; Tian, Y.; Wang, J.; Li, Y.; Wang, H.; Mao, S.; Liu, X.; Wang, C.; Bie, S.; Lu, F. Construction of Engineered Arthrobacter Simplex with Improved Performance for Cortisone Acetate Biotransformation. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 9503−9514.

(6) Rohman, A.; Dijkstra, B. W. The Role and Mechanism of Microbial 3-Ketosteroid Δ1-Dehydrogenases in Steroid Breakdown. *J. Steroid Biochem. Mol. Biol.* **2019**, *191*, 105366.

(7) Donova, M. V.; Egorova, O. V. Microbial Steroid Transformations: Current State and Prospects. *Appl. Microbiol. Biotechnol.* **2012**, *94*, 1423−1447.

(8) Rohman, A.; Dijkstra, B. W. Application of Microbial 3-Ketosteroid Δ1-Dehydrogenases in Biotechnology. *Biotechnol. Adv.* **2021**, *49*, 107751.

(9) Chiang, Y. R.; Ismail, W.; Gallien, S.; Heintz, D.; Van Dorsselaer, A.; Fuchs, G. Cholest-4-En-3-One-Δ1-Dehydrogenase, a Flavoprotein Catalyzing the Second Step in Anoxic Cholesterol Metabolism. *Appl. Environ. Microbiol.* **2008**, *74*, 107−113.

(10) Luo, J. M.; Cui, H. L.; Jia, H. C.; Li, F.; Cheng, H. J.; Shen, Y. B.; Wang, M. Identification, Biological Characteristics, and Active Site Residues of 3-Ketosteroid Δ1-Dehydrogenase Homologues from Arthrobacter Simplex. *J. Agric. Food Chem.* **2020**, *68*, 9496−9512.

(11) Wojtkiewicz, A. M.; Wójcik, P.; Procner, M.; Flejszar, M.; Oszajca, M.; Hochołowski, M.; Tataruch, M.; Mrugała, B.; Janeczko, T.; Szaleniec, M. The Efficient Δ1-Dehydrogenation of a Wide Spectrum of 3-Ketosteroids in a Broad PH Range by 3-Ketosteroid Dehydrogenase from Sterolibacterium Denitrificans. *J. Steroid Biochem. Mol. Biol.* **2020**, *202*, 105731.

(12) Wójcik, P.; Glanowski, M.; Wojtkiewicz, A. M.; Rohman, A.; Szaleniec, M. Universal Capability of 3-ketosteroid Δ1 Dehydrogenases to Catalyze Δ1-dehydrogenation of C17-substituted Steroids. *Microb. Cell Factories* **2021**, *20*, 119.

(13) Rohman, A.; van Oosterwijk, N.; Thunnissen, A. M. W. H.; Dijkstra, B. W. Crystal Structure and Site-Directed Mutagenesis of 3-Ketosteroid Δ1-Dehydrogenase from Rhodococcus Erythropolis SQ1 Explain Its Catalytic Mechanism. *J. Biol. Chem.* **2013**, *288*, 35559−35568.

(14) Oxford Diffraction. *CrysAlis Pro Oxford Diffraction Ltd*, Version 1.171.36.20 (Release 27-06-2012 CrysAlis171.NET): Abingdon, England, 2006.

(15) Kabsch, W. XDS. *Acta Crystallogr., Sect. D: Struct. Biol.* **2010**, *66*, 125−132.

(16) Adams, P. D.; Afonine, V.; Bunkóczi, V. B.; Chen, W.; Davis, N.; Echols, N. W.; Headd, R. J.; Hung, D. C.; Kapral, S.; Grosse-Kunstleve, C. PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 213−221.

(17) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A. J.; McNicholas, S.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 Suite and Current Developments. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 235−242.

(18) Murshudov, G. N.; Skubák, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. REFMAC 5 for the Refinement of Macromolecular Crystal Structures. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 355−367.

(19) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 486−501.

(20) Chen, V. B.; Arendall, W. B., III; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: All-Atom Structure Validation for Macromolecular Crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2009**, *66*, 12−21.

(21) Young, J. Y.; Westbrook, J. D.; Feng, Z.; Sala, R.; Peisach, E.; Oldfield, T. J.; Sen, S.; Gutmanas, A.; Armstrong, D. R.; Berrisford, J. M.; Chen, L.; Chen, M.; Di Costanzo, L.; Dimitropoulos, D.; Gao, G.; Ghosh, S.; Gore, S.; Guranovic, V.; Hendrickx, P. M. S.; Hudson, B. P.; Igarashi, R.; Ikegawa, Y.; Kobayashi, N.; Lawson, C. L.; Liang, Y.; Mading, S.; Mak, L.; Mir, M. S.; Mukhopadhyay, A.; Patwardhan, A.; Persikova, I.; Rinaldi, L.; Sanz-Garcia, E.; Sekharan, M. R.; Shao, C.;

210

Swaminathan, G. J.; Tan, L.; Ulrich, E. L.; van Ginkel, G.; Yamashita, R.; Yang, H.; Zhuravleva, M. A.; Quesada, M.; Kleywegt, G. J.; Berman, H. M.; Markley, J. L.; Nakamura, H.; Velankar, S.; Burley, S. K. OneDep: Unified WwPDB System for Deposition , Biocuration , and Validation of Macromolecular Structures in the PDB Archive. *Structure* **2017**, *25*, 536−545.

(22) *The PyMOL Molecular Graphics System*, Version 1.2r3pre, Schrödinger, LLC.

(23) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera — A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25*, 1605−1612.

(24) BIOVIA. *Dassault Systèmes, Discovery Studio*, V21.1.0.20298; Dassault Systèmes: San Diego, 2021.

(25) Lomize, M. A.; Pogozheva, I. D.; Joo, H.; Mosberg, H. I.; Lomize, A. L. OPM Database and PPM Web Server: Resources for Positioning of Proteins in Membranes. *Nucleic Acids Res.* **2012**, *40*, D370−D376.

(26) Sofińska, K.; Wojtkiewicz, A. M.; Wójcik, P.; Zastawny, O.; Guzik, M.; Winiarska, A.; Waligórski, P.; Cieśla, M.; Barbasz, J.; Szaleniec, M. Investigation of Quaternary Structure of Aggregating 3-Ketosteroid Dehydrogenase from Sterolibacterium Denitrificans: In the Pursuit of Consensus of Various Biophysical Techniques. *Biochim. Biophys. Acta, Gen. Subj.* **2019**, *1863*, 1027−1039.

(27) Glanowski, M.; Wójcik, P.; Procner, M.; Borowski, T.; Lupa, D.; Mielczarek, P.; Oszajca, M.; Świderek, K.; Moliner, V.; Bojarski, A. J.; Szaleniec, M. Enzymatic Δ1-Dehydrogenation of 3-Ketosteroids—Reconciliation of Kinetic Isotope Effects with the Reaction Mechanism. *ACS Catal.* **2021**, *11*, 8211−8225.

(28) Rugor, A.; Wójcik-Augustyn, A.; Niedzialkowska, E.; Mordalski, S.; Staroń, J.; Bojarski, A.; Szaleniec, M.Reaction Mechanism of Sterol Hydroxylation by Steroid C25 Dehydrogenase − Homology Model, Reactivity and Isoenzymatic Diversity. *J. Inorg. Biochem.***2017**, *173*. DOI: 10.1016/j.jinorgbio.2017.04.027

(29) Melander, L.*Isotope Effect on Reaction Rates*; Ronald Press Co.: New York, 1960; pp 46−64.

(30) Madeira, F.; Park, Y. M.; Lee, J.; Buso, N.; Gur, T.; Madhusoodanan, N.; Basutkar, P.; Tivey, A. R. N.; Potter, S. C.; Finn, R. D.; Lopez, R. The EMBL-EBI Search and Sequence Analysis Tools APIs in 2019. *Nucleic Acids Res.* **2019**, *47*, W636−W641.

(31) Waterhouse, A. M.; Procter, J. B.; Martin, D. M. A.; Clamp, M.; Barton, G. J. Jalview Version 2—a Multiple Sequence Alignment Editor and Analysis Workbench. *Bioinformatics* **2009**, *25*, 1189−1191.

(32) Kabsch, W. A Discussion of the Solution for the Best Rotation to Relate Two Sets of Vectors. *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr.* **1978**, *34*, 827−828.

(33) Søndergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of PKa Values. *J. Chem. Theory Comput.* **2011**, *7*, 2284−2295.

(34) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical PKa Predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525−537.

(35) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; WilliamsDing, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J.*G16_C01, Gaussian 16, Revision C.01*; Gaussian, Inc.: Wallingford, 2016.

(36) Becke, A. D. Density-Functional Thermochemistry, III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(37) Dupradeau, F. Y.; Cézard, C.; Lelong, R.; Stanislawiak, É.; Pêcher, J.; Delepine, J. C.; Cieplak, P. R.E.DD.B.: A Database for RESP and ESP Atomic Charges, and Force Field Libraries. *Nucleic Acids Res.* **2008**, *36*, D360−D367.

(38) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. An Overview of the Amber Biomolecular Simulation Package. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2013**, *3*, 198−210.

(39) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* **2003**, *24*, 1999−2012.

(40) Miller, B. R.; McGee, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. MMPBSA.Py : An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314−3321.

(41) Field, M. J.; Albe, M.; Bret, C.; Proust-De Martin, F.; Thomas, A. The Dynamo Library for Molecular Simulations Using Hybrid Quantum Mechanical and Molecular Mechanical Potentials. *J. Comput. Chem.* **2000**, *21*, 1088−1100.

(42) Krzemińska, A.; Paneth, P.; Moliner, V.; Świderek, K. Binding Isotope Effects as a Tool for Distinguishing Hydrophobic and Hydrophilic Binding Sites of HIV-1 RT. *J. Phys. Chem. B* **2015**, *119*, 917−927.

(43) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23*, 187−199.

(44) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. THE Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13*, 1011−1021.

(45) Dewar, J. S.; Zoebisch, M.; Healy, E.; Stewart, J. Development and Use of Quantum Mechanical Molecular Models. 76. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **2002**, *107*, 3902−3909.

(46) Ruiz-Pernía, J. J.; Silla, E.; Tuñón, I.; Martí, S.; Moliner, V. Hybrid QM/MM Potentials of Mean Force with Interpolated Corrections. *J. Phys. Chem. B* **2004**, *108*, 8427−8433.

(47) Chuang, Y.-Y.; Corchado, J. C.; Truhlar, D. G. Mapped Interpolation Scheme for Single-Point Energy Corrections in Reaction Rate Calculations and a Critical Evaluation of Dual-Level Reaction Path Dynamics Methods. *J. Phys. Chem. A* **1999**, *103*, 1140−1149.

(48) Świderek, K.; Tuñón, I.; Martí, S.; Moliner, V. Protein Conformational Landscapes and Catalysis. Influence of Active Site Conformations in the Reaction Catalyzed by L-Lactate Dehydrogenase. *ACS Catal.* **2015**, *5*, 1172−1185.

(49) Baker, J.; Kessi, A.; Delley, B. The Generation and Use of Delocalized Internal Coordinates in Geometry Optimization. *J. Chem. Phys.* **1996**, *105*, 192−212.

(50) Martí, S.; Moliner, V.; Tuñón, I. Improving the QM/MM Description of Chemical Processes: A Dual Level Strategy To Explore the Potential Energy Surface in Very Large Systems. *J. Chem. Theory Comput.* **2005**, *1*, 1008−1016.

(51) Ruggiero, G. D.; Guy, S. J.; Martí, S.; Moliner, V.; Williams, I. H. Vibrational Analysis of the Chorismate Rearrangement: Relaxed Force Constants, Isotope Effects and Activation Entropies Calculated for Reaction in Vacuum, Water and the Active Site of Chorismate Mutase. *J. Phys. Org. Chem.* **2004**, *17*, 592−601.

(52) Świderek, K.; Tuñón, I.; Williams, I. H.; Moliner, V. Insights on the Origin of Catalysis on Glycine N-Methyltransferase from Computational Modeling. *J. Am. Chem. Soc.* **2018**, *140*, 4327−4334.

(53) Krzemińska, A.; Moliner, V.; Świderek, K. Dynamic and Electrostatic Effects on the Reaction Catalyzed by HIV-1 Protease. *J. Am. Chem. Soc.* **2016**, *138*, 16283−16298.

211

(54) Wang, X.; Feng, J.; Zhang, D.; Wu, Q.; Zhu, D.; Ma, Y. Characterization of New Recombinant 3-Ketosteroid-Δ1-Dehydro-genases for the Biotransformation of Steroids. *Appl. Microbiol. Biotechnol.* **2017**, *101*, 6049−6060.

(55) Guevara, G.; Fernández de las Heras, L.; Perera, J.; Navarro Llorens, J. M. Functional Differentiation of 3-Ketosteroid Δ1-Dehydrogenase Isozymes in Rhodococcus Ruber Strain Chol-4. *Microb. Cell Factories* **2017**, *16*, 42.

(56) Brzostek, A.; Śliwiński, T.; Rumijowska-Galewicz, A.; Korycka-Machała, M.; Dziadek, J. Identification and Targeted Disruption of the Gene Encoding the Main 3-Ketosteroid Dehydrogenase in Mycobacterium Smegmatis. *Microbiology* **2005**, *151*, 2393−2402.

(57) van Oosterwijk, N.; Knol, J.; Dijkhuizen, L.; van der Geize, R.; Dijkstra, B. W. Structure and Catalytic Mechanism of 3-Ketosteroid-Δ4-(5α)- Dehydrogenase from Rhodococcus Jostii RHA1 Genome. *J. Biol. Chem.* **2012**, *287*, 30975−30983.

(58) Urich, K.*Sterols and Steroids BT-Comparative Animal Bio-chemistry*; Urich, K., Ed.; Springer Berlin Heidelberg: Berlin, Heidelberg, 1994; pp 624−656.

(59) Dym, O.; Eisenberg, D. Sequence-Structure Analysis of FAD-Containing Proteins. *Protein Sci.* **2001**, *10*, 1712−1728.

(60) Hayano, M.; Ringold, H. J.; Stefanovic, V.; Gut, M.; Dorfman, R. I. The Stereochemical Course of Enzymatic Steroid 1,2-Dehydrogenation. *Biochem. Biophys. Res. Commun.* **1961**, *4*, 454−459.

(61) Eyring, H. The Activated Complex in Chemical Reactions. *J. Chem. Phys.* **1935**, *3*, 107−115.

(62) Wojtkiewicz, A. M.; Glanowski, M.; Waligórski, P.; Janeczko, T.; Szaleniec, M. 1,2-Hydrogenation and Transhydrogenation Catalyzed by 3-Ketosteroid Δ1-Dehydrogenase from Steroli-bacterium Denitrificans—Kinetics, Isotope Labelling and QM:MM Modelling Studies. *Int. J. Mol. Sci.* **2022**, *23*, 14660.

(63) Karsten, W. E.; Cook, P. F.*Substrate Dependence of Isotope Effects*. In *Isotope Effects In Chemistry and Biology*; Kohen, A., Limbach, H.-H., Eds.; Springer-Verlag, 2019; pp 796−802.

📖 Recommended by ACS

Structural and Functional Characterization of the Ureidoacrylate Amidohydrolase RutB from *Escherichia coli*

Markus R. Busch, Reinhard Sterner, *et al.*

JANUARY 04, 2023
BIOCHEMISTRY   READ 📄

Structural Insights into the Substrate Range of a Bacterial Monoamine Oxidase

Samantha N. Muellers, Karen N. Allen, *et al.*

JANUARY 20, 2023
BIOCHEMISTRY   READ 📄

Compound I Formation and Reactivity in Dimeric Chlorite Dismutase: Impact of pH and the Dynamics of the Catalytic Arginine

Daniel Schmidt, Christian Obinger, *et al.*

JANUARY 27, 2023
BIOCHEMISTRY   READ 📄

Structural and Functional Analysis of a Highly Active Designed Phosphotriesterase for the Detoxification of Organophosphate Nerve Agents Reveals an Unpredicted...

Laura Job, Arne Skerra, *et al.*

FEBRUARY 08, 2023
BIOCHEMISTRY   READ 📄

Get More Suggestions >

Structure, mutagenesis and QM:MM modelling of 3-ketosteroid Δ^1-dehydrogenase from *Sterolibacterium denitrificans* – the role of new putative membrane-associated domain and proton-relay system in catalysis

Patrycja Wójcik[1], Michał Glanowski[1], Beata Mrugała[1], Magdalena Procner[1,2] Olga Zastawny[1], Monika Flejszar[1,3], Katarzyna Kurpiewska[4], Ewa Niedziałkowska[5], Wladek Minor[5], Maria Oszajca[4], Andrzej J. Bojarski[2], Agnieszka M. Wojtkiewicz[1], Maciej Szaleniec[1,*]

[1]Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, Niezapominajek 8, 30-239 Kraków, Poland

[2]Jerzy Maj Institute of Pharmacology Polish Academy of Sciences, Smętna 12, 31-343 Kraków, Poland

[3]Department of Physical Chemistry, Faculty of Chemistry, Rzeszow University of Technology, Al. Powstańców Warszawy 6, 35-959 Rzeszów, Poland

[4] Faculty of Chemistry, Jagiellonian University, Gronostajowa 2, 30-387 Kraków, Poland

[5] Department of Molecular Physiology and Biological Physics, University of Virginia, 1340 Jefferson Park Av., Pinn Hall, Charlottesville, VA 22908, USA

*email: maciej.szaleniec@ikifp.edu.pl

S1

INDEX

QMMM model setup

Figure S3. Overview of AcmB active site. Atoms inside the blue area are treated with QM methods. Link atoms are marked with asterisks.

Putative membrane-associated domain
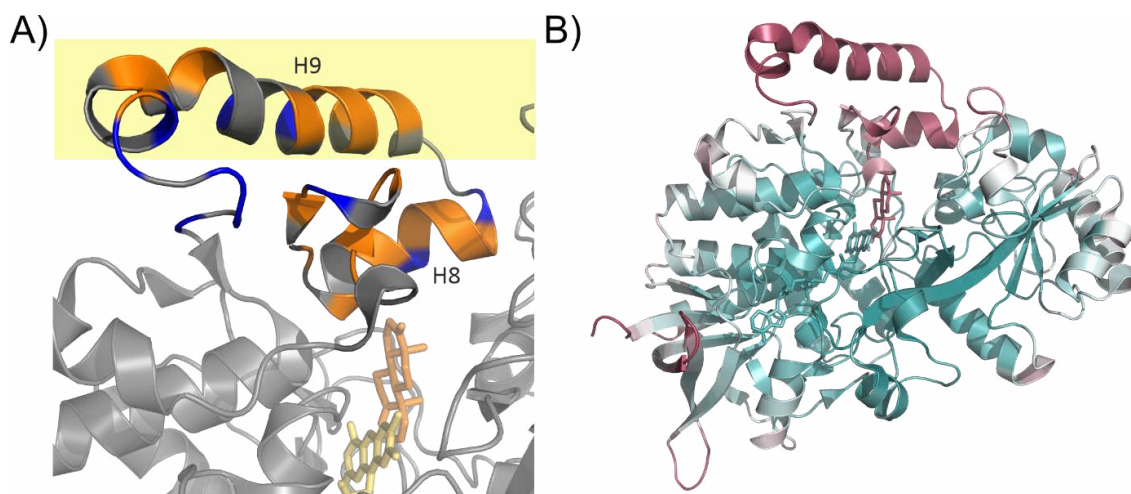
Figure S6. AcmB structure with a putative membrane-associated domain: A) with hydrophobic residues (Kyte and Doolittle[5] hydrophobicity index > 0) in orange and positively charged residues in blue; B) structure of AcmB with B-factor colouring indicating relative mobility/disorder of the protein regions (reddish – B-factor > 50, high mobility, blue B-factor <20, low mobility)
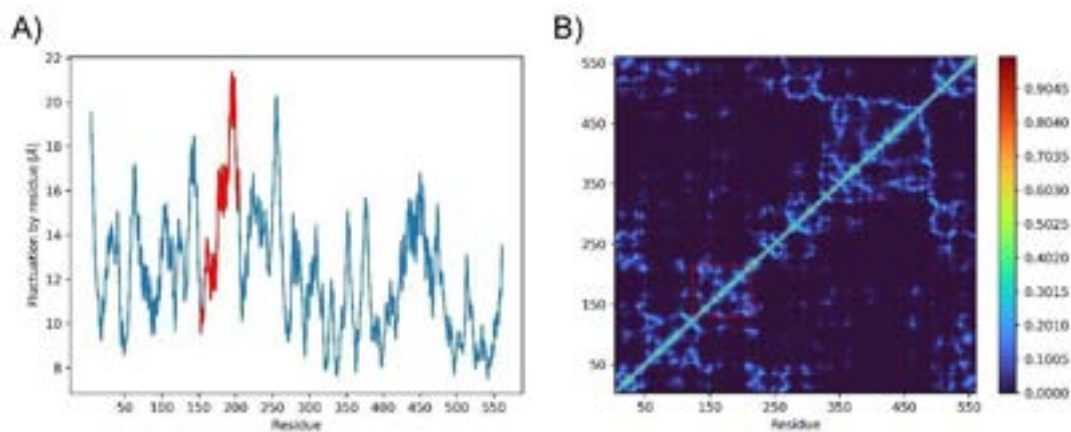
Figure S7. A) Fluctuation of AcmB residues in MD simulation and B) correlation map of the AmcB region movements during simulation. The red color marks the putative membrane-associated region (Y153–R204).

MD modelling

Table S5 Summary of MD simulation for **E:S** complexes. The RMSD and crucial distances indicate catalytically competent poses of the substrate. The red frame indicates the selected fragment of simulation used for the calculation of ΔG_{best}.

Androst-4-en-3,17-dione (AD)

Dihydrotestosterone (DHT)
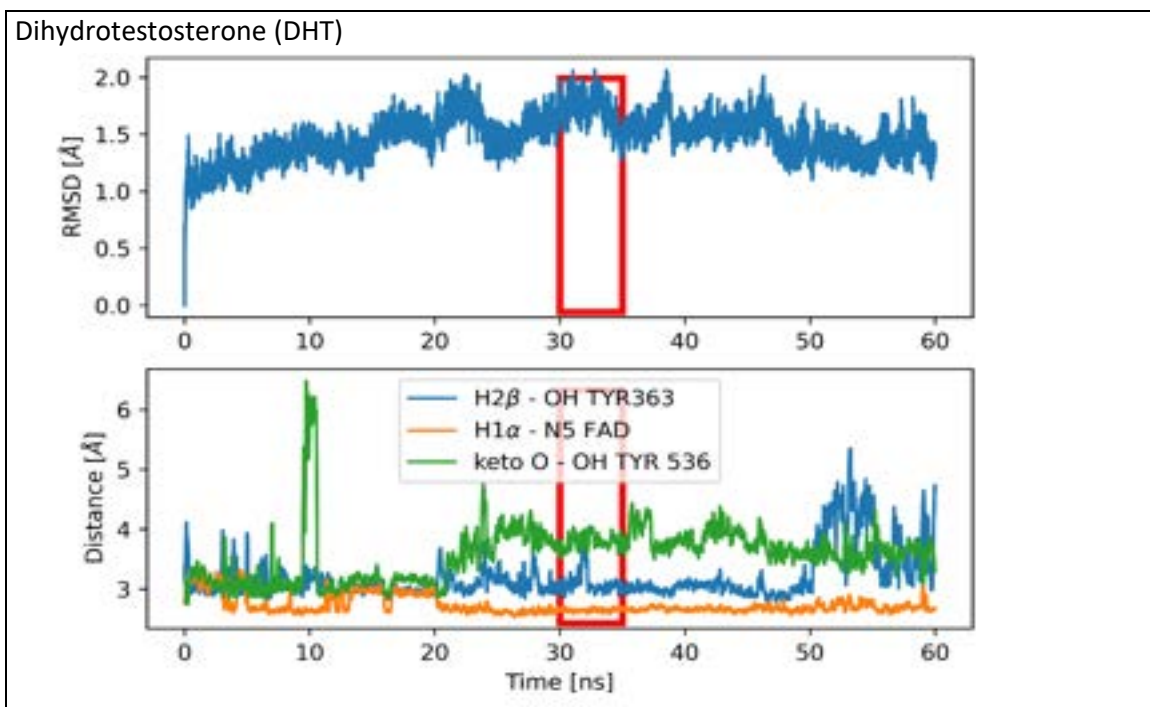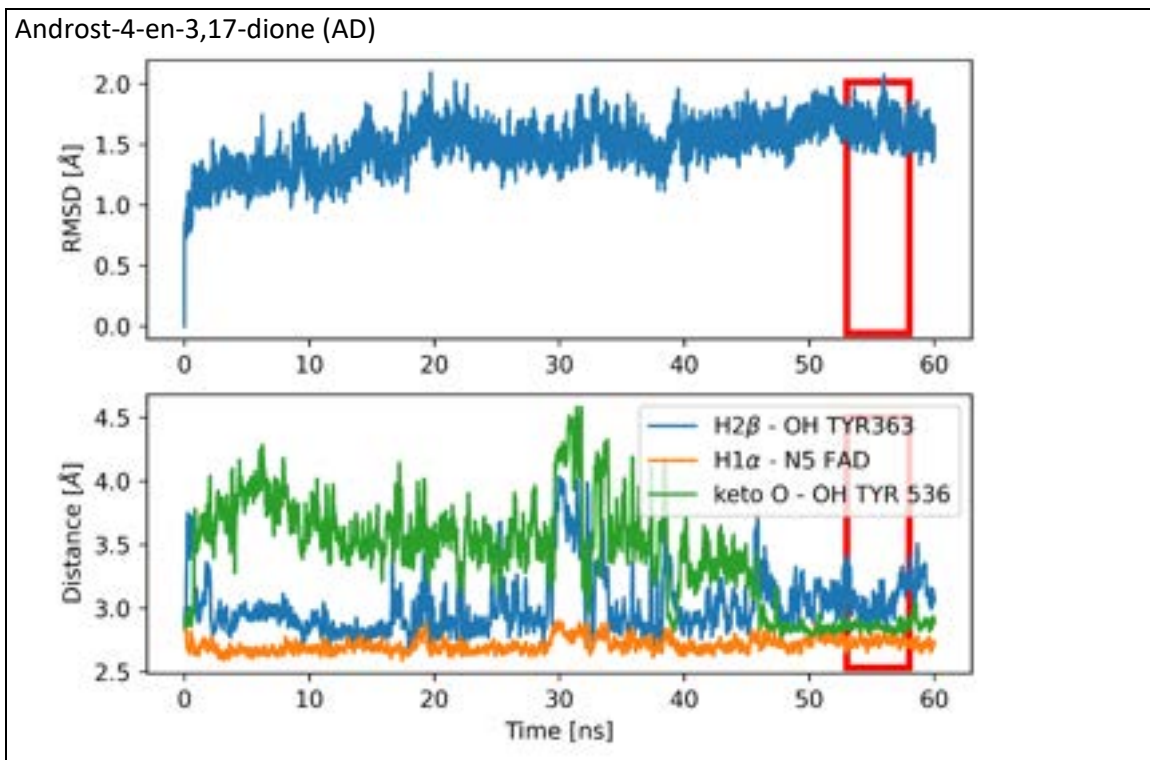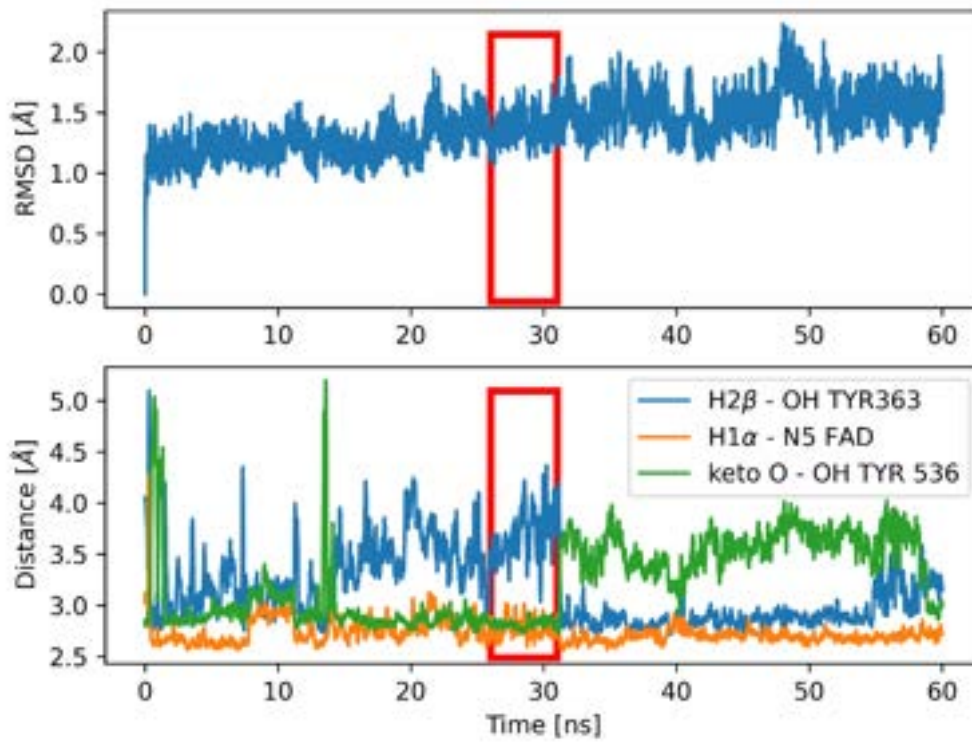
Progesterone (PRG)

17-methyltestosterone (17-MT)
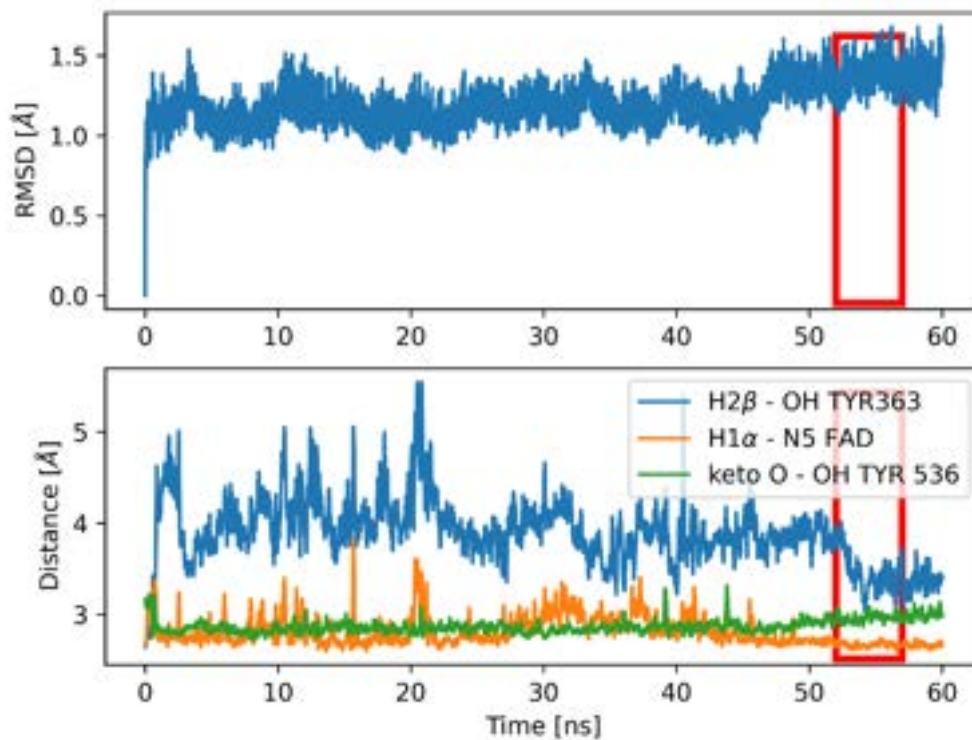
Testosterone propionate (TP)

6-Dehydrotestosterone acetate (6-DHA)

S21
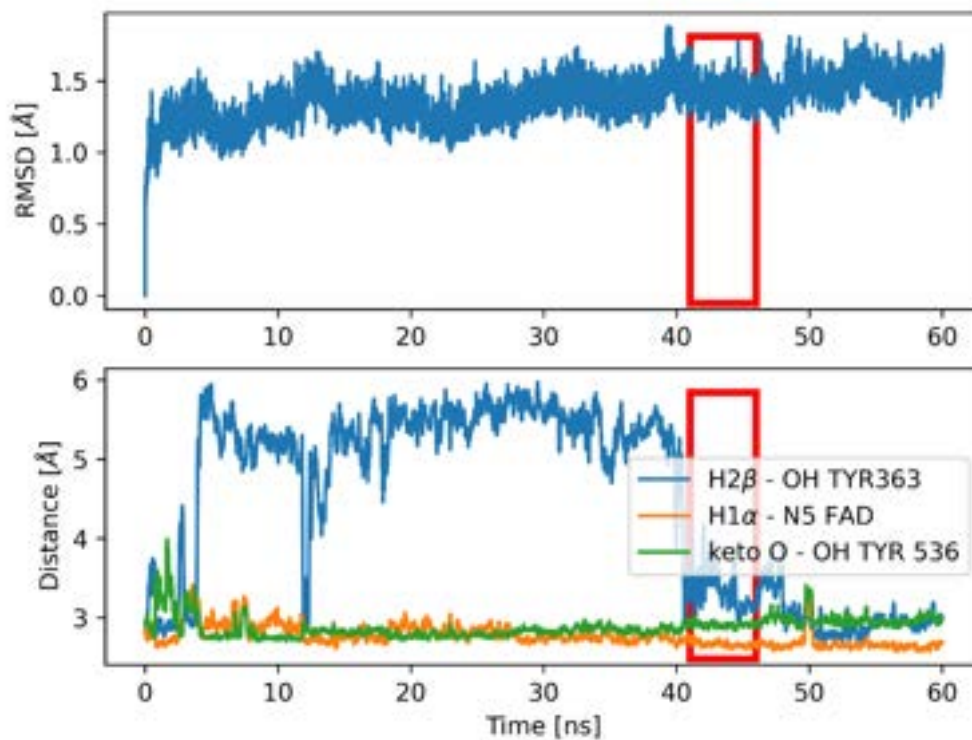
Diosgenone (DGN)

Cholest-4-en-3-one

S22
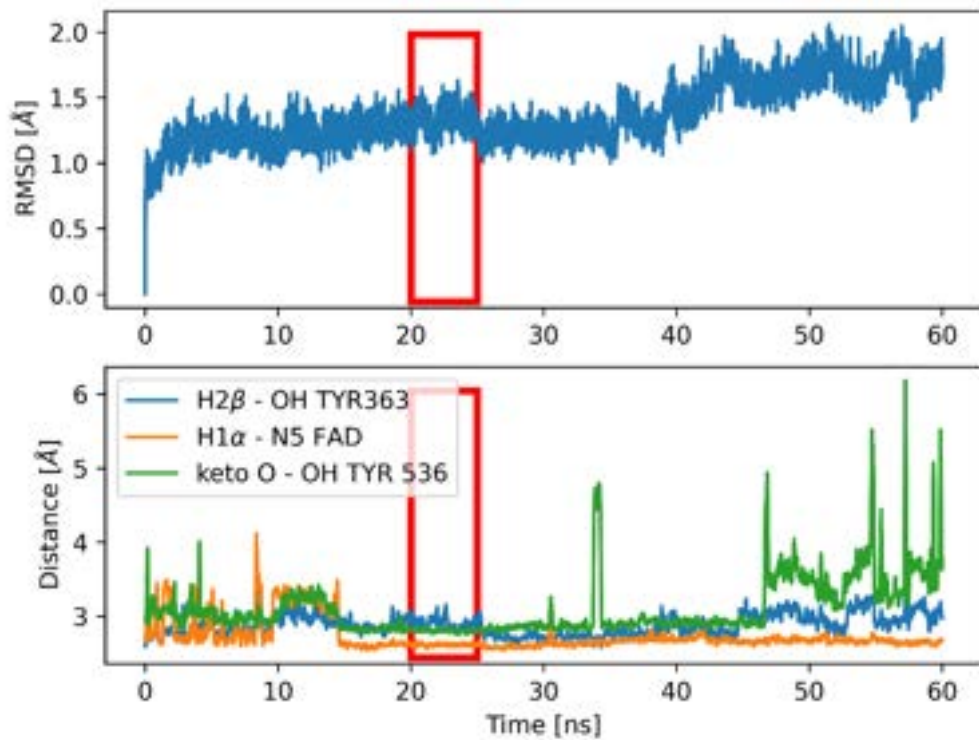
Table S6. Summary of MMPBSA ΔG calculation for MD simulation for **E:S** complexes.
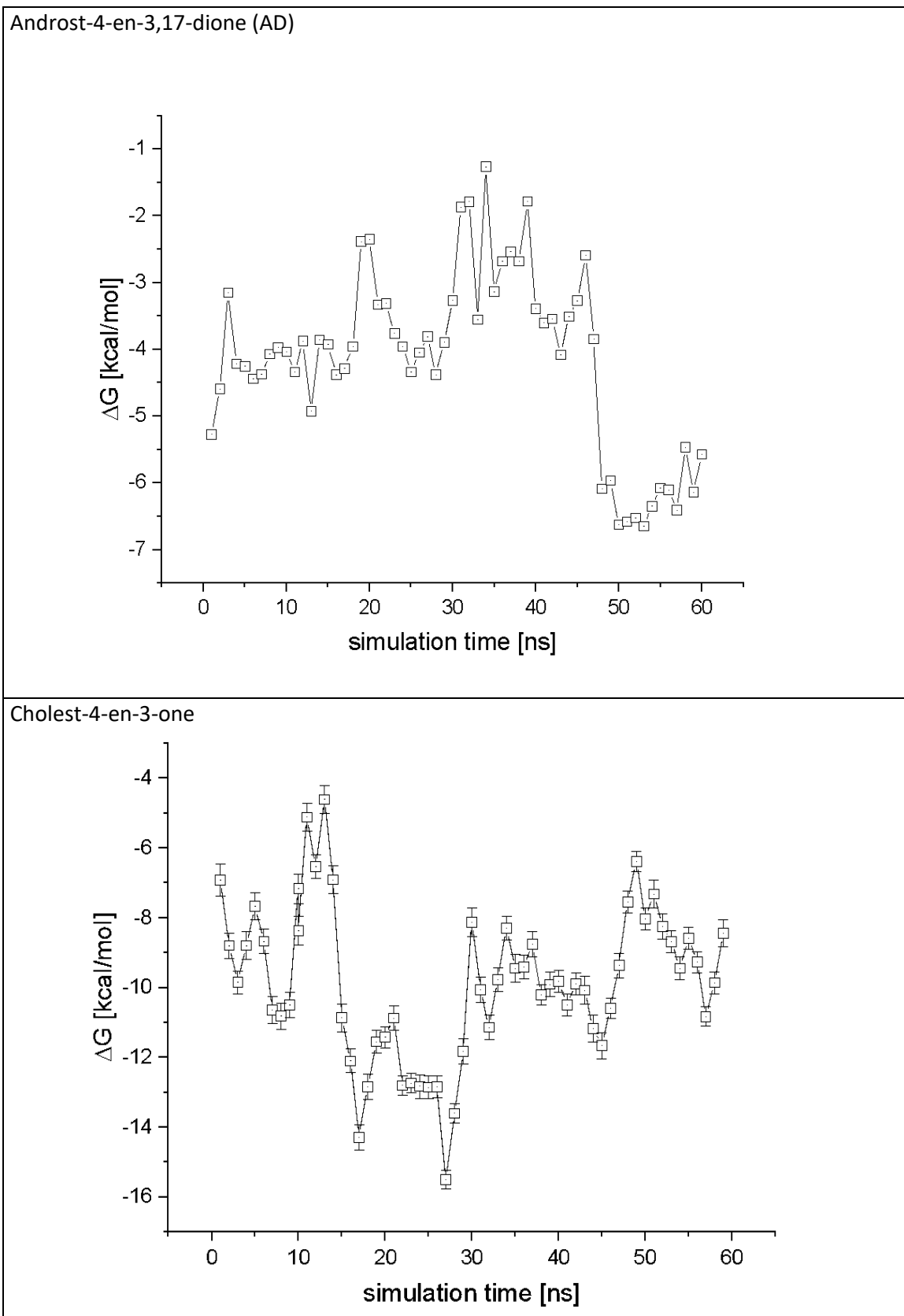
Androst-4-en-3,17-dione (AD)

Cholest-4-en-3-one

Dihydrotestosterone (DHT)
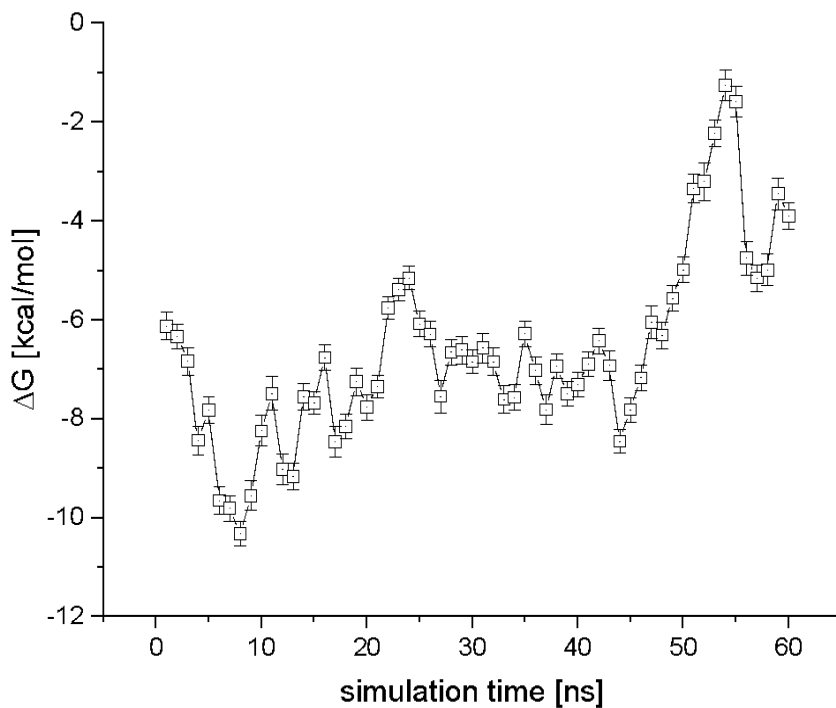
Progesterone (PRG)
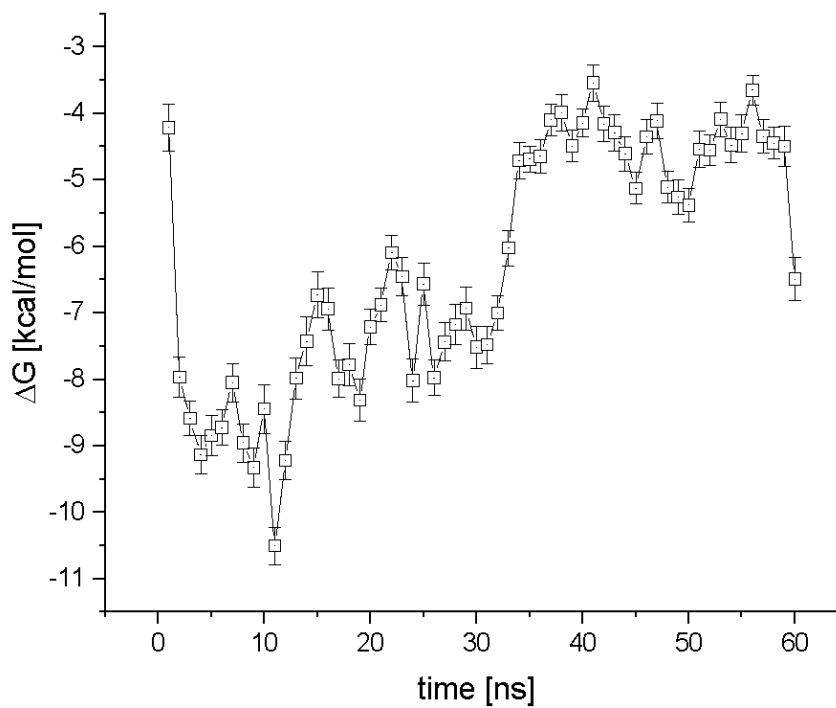
17-methyltestosterone (17-MT)
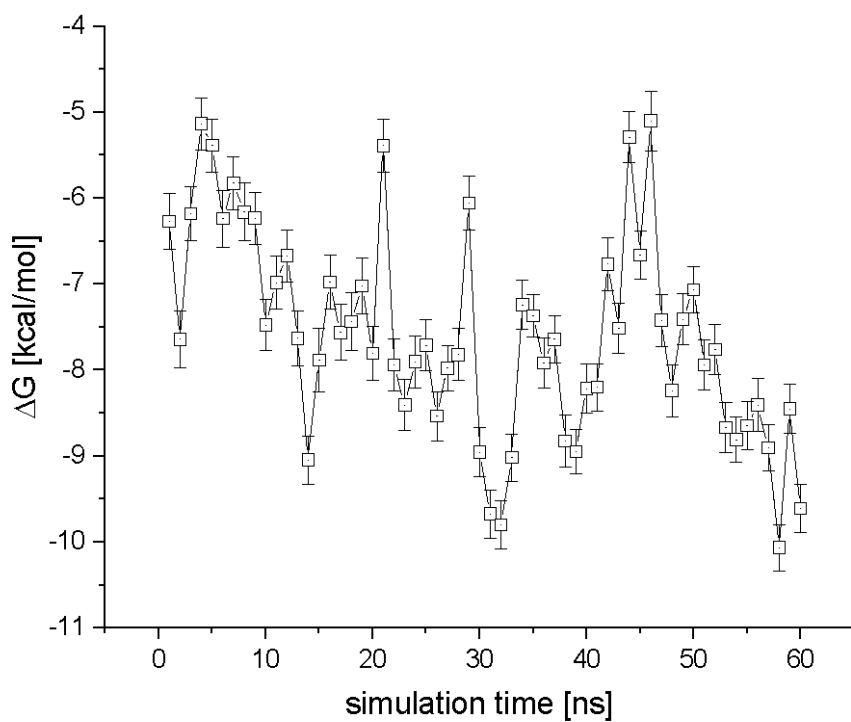
Testosterone propionate (TE-P)

S25
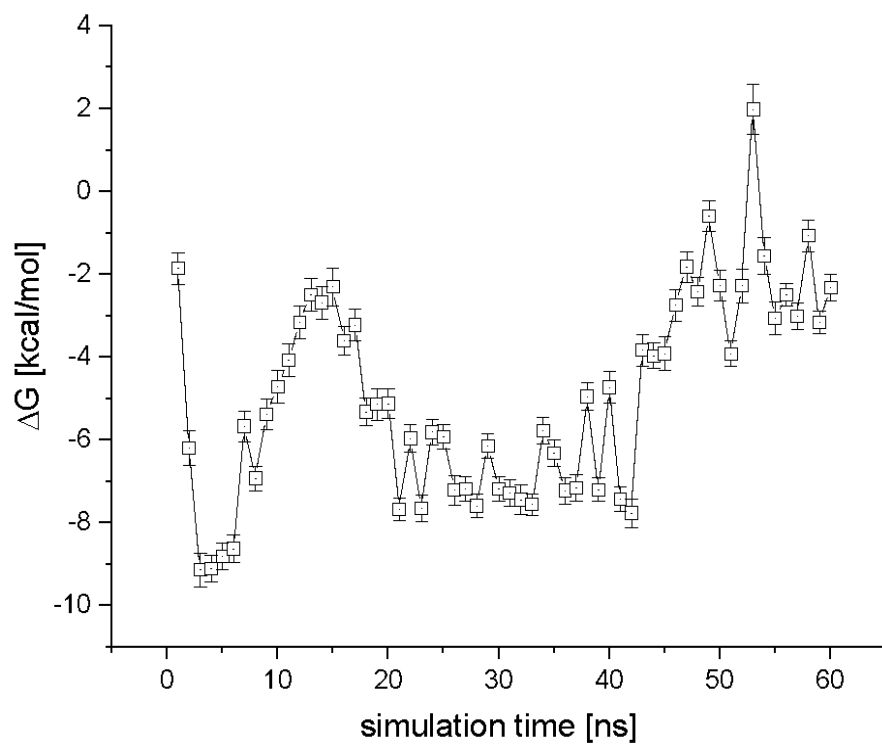
6-Dehydrotestosterone acetate (6-DHA)
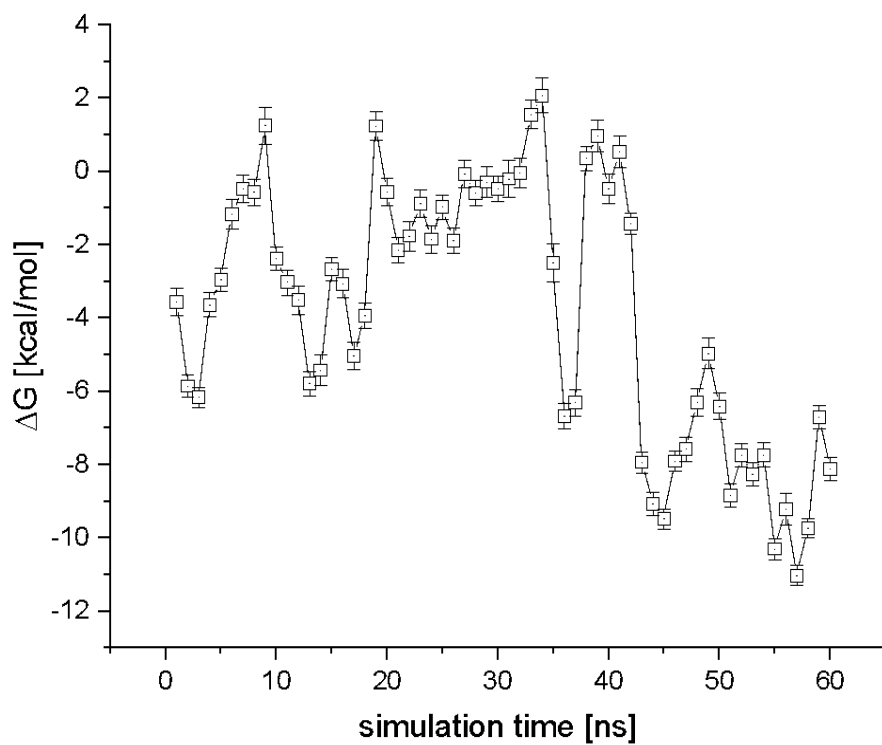
Diosgenone
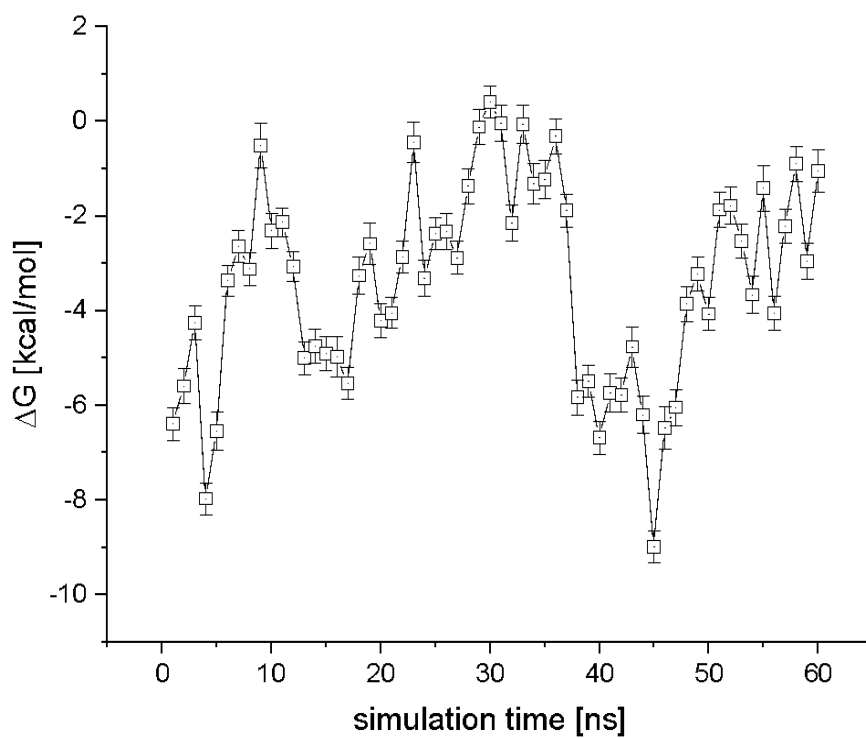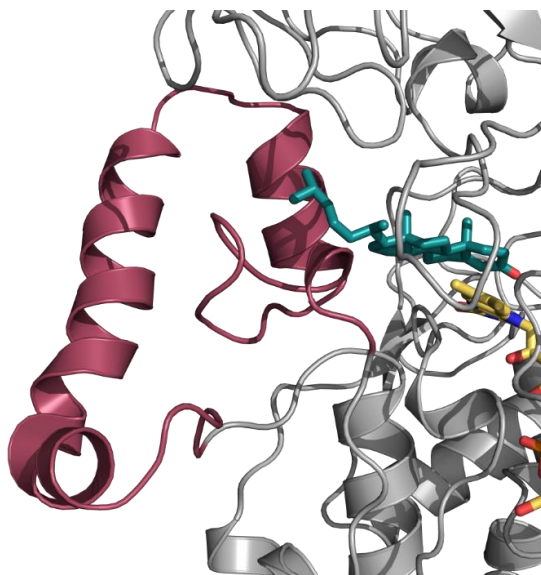
Figure S12. Results of MD simulation for AcmB with cholest-4-en-3-one (purple) in the active site. Domain unique for AcmB marked with red.
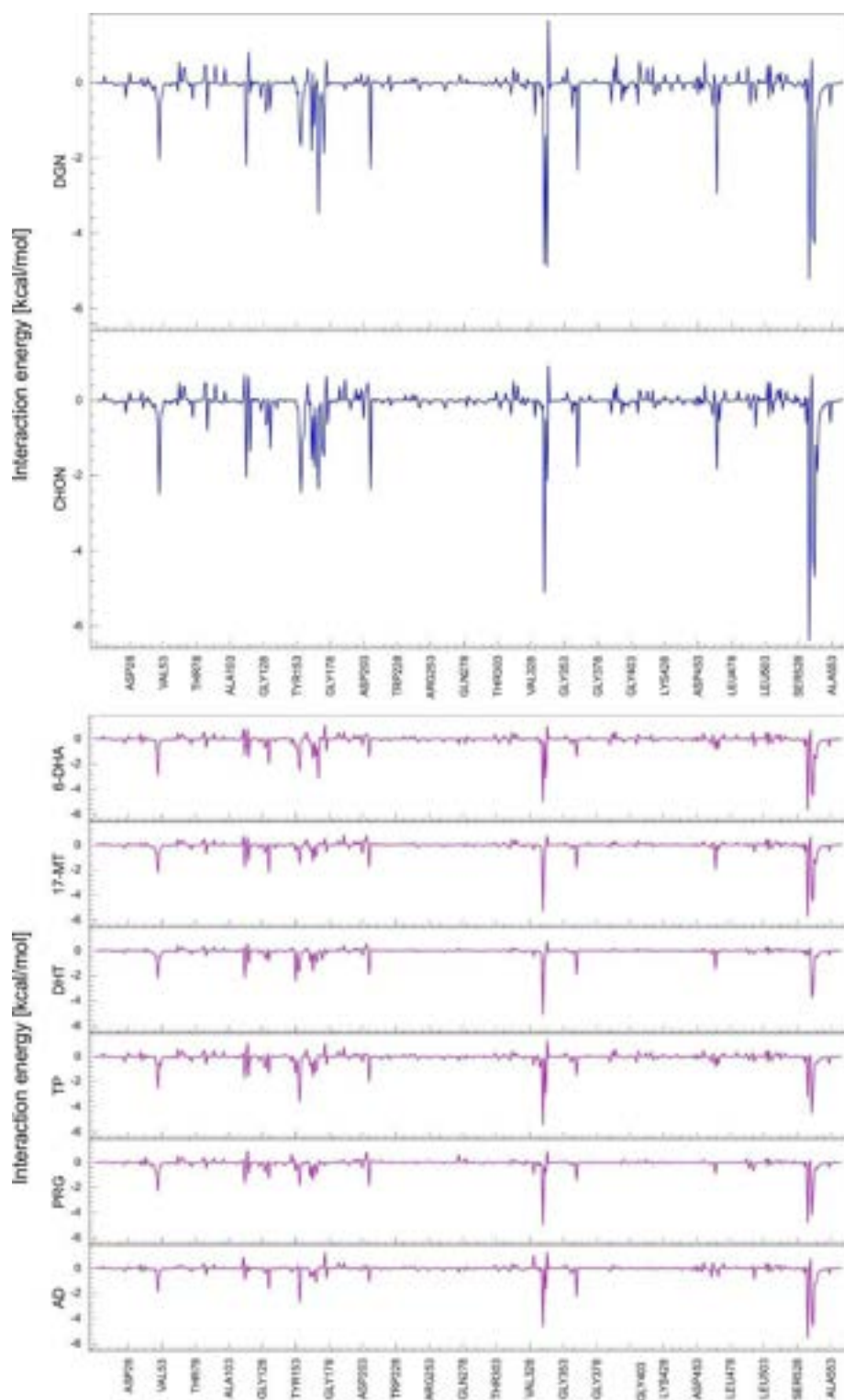
Figure S13. Interaction energies between AcmB and its substrates calculated for the best-fit phase of MD simulations (red frames in Table. S5).

Reaction mechanism
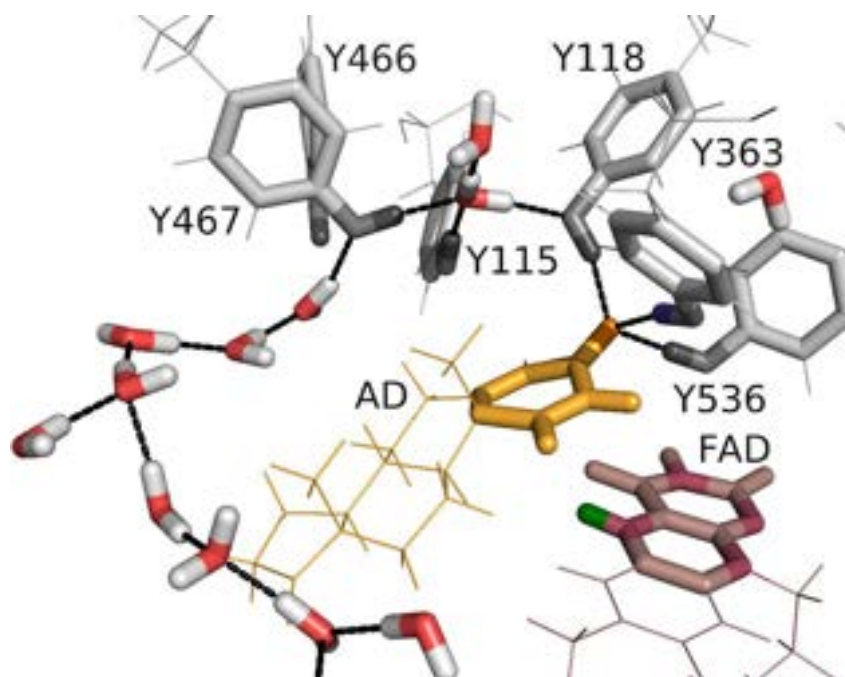
Figure S14. A proton relay system depicted for **E:P** with ADD. H-bonds are marked by dashed green and blue lines.

S29
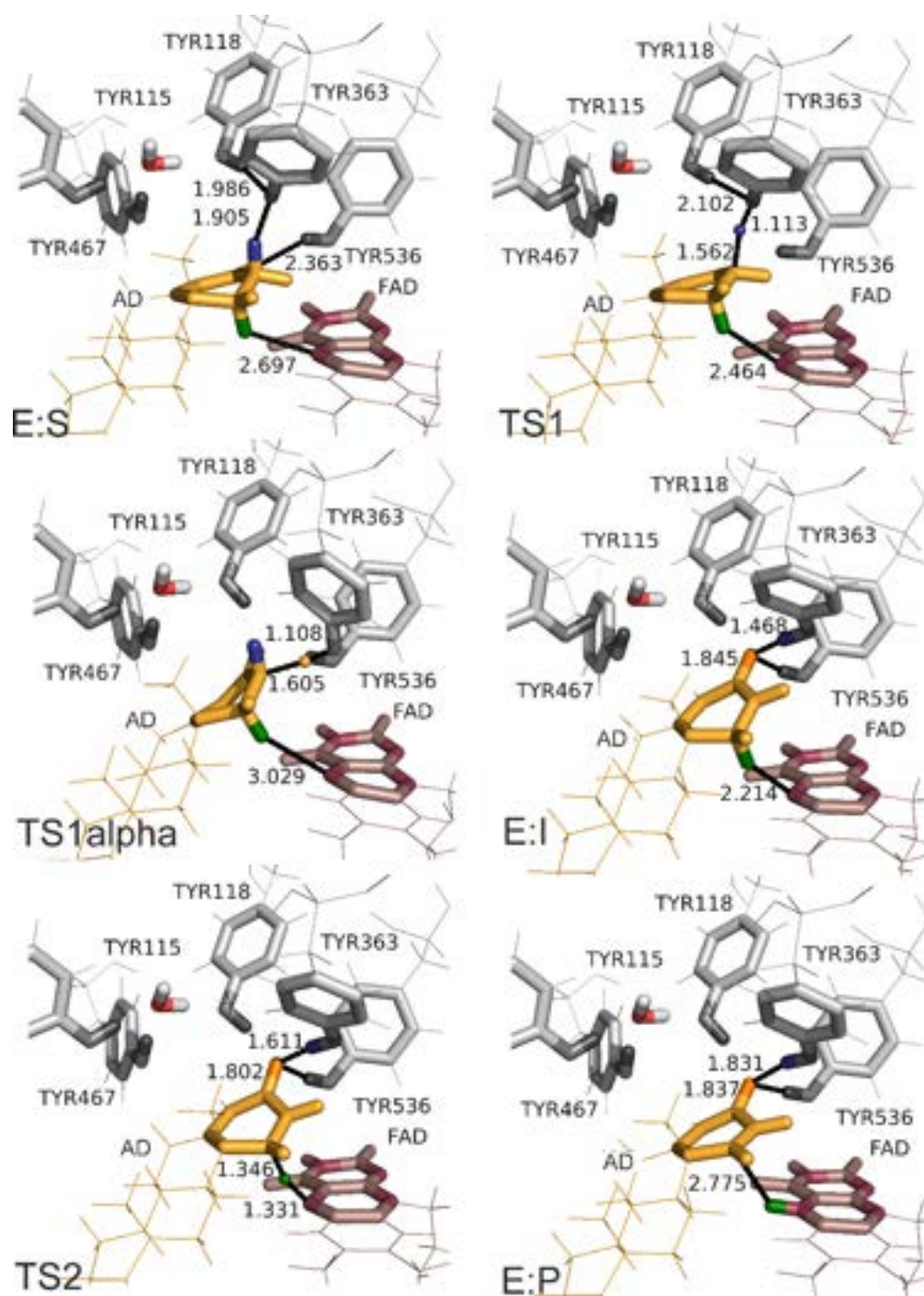
Figure S15. Reaction mechanism of Δ^1-dehydrogenation of androst-4-en-3,17-dione (AD). Two potential mechanisms of substrate activation are depicted as TS1 (2βH abstraction) or TS1alpha (2αH abstraction). Distances are provided in Å.

S30
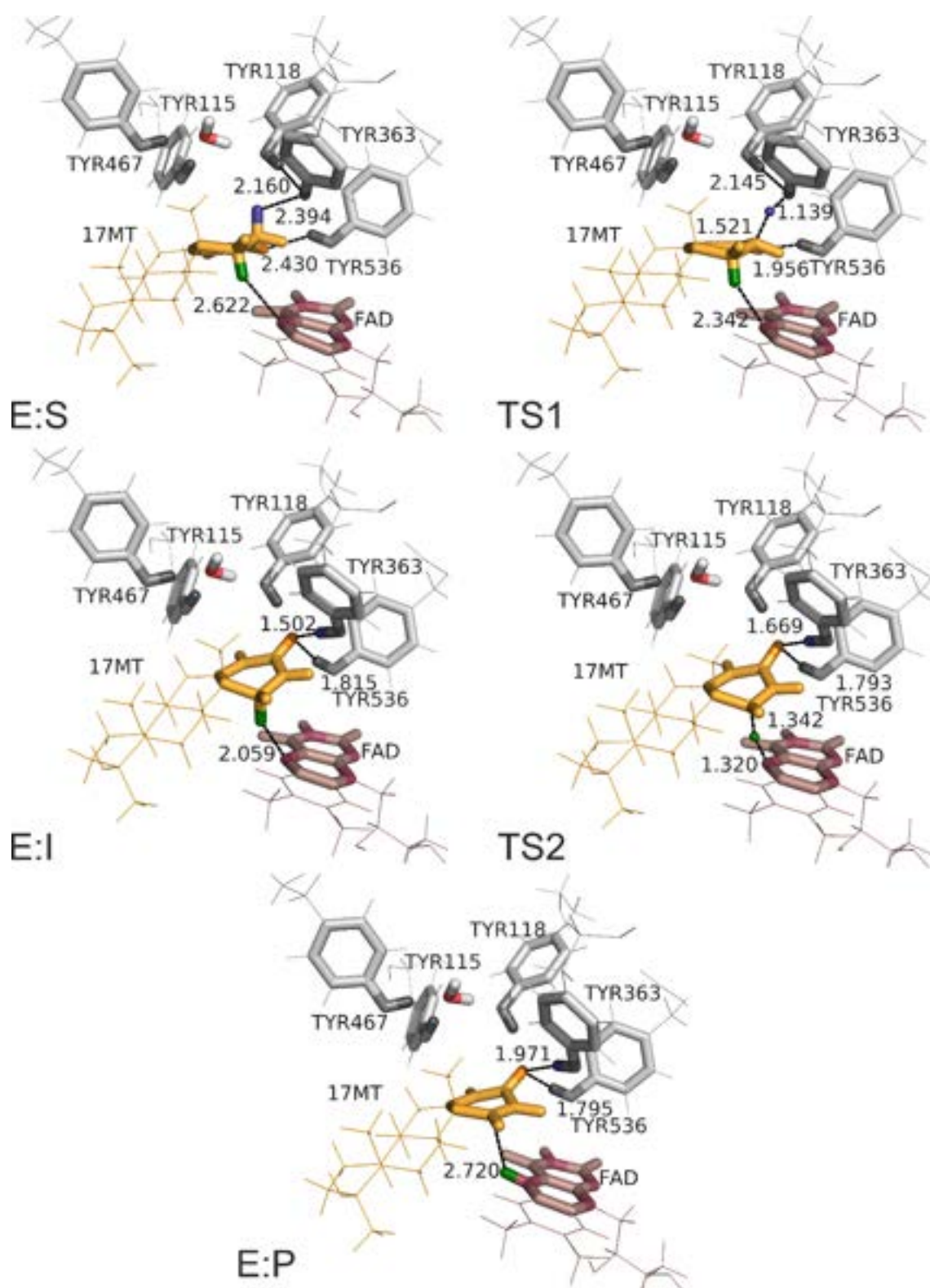
Figure S16. Reaction mechanism of Δ^1-dehydrogenation of 17-methyltestosterone (17-MT). Distances are provided in Å.
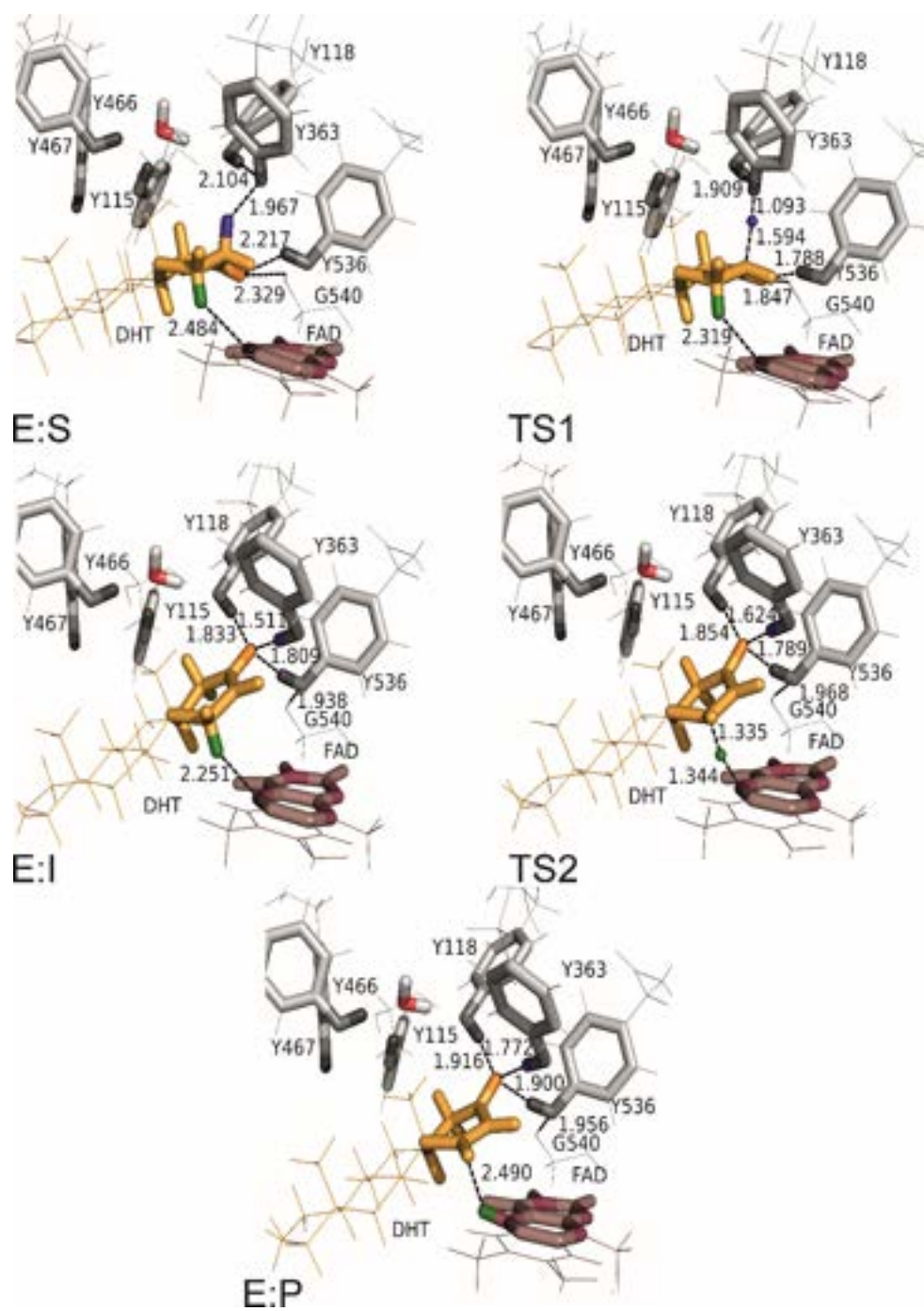
Figure S17. Reaction mechanism of Δ^1-dehydrogenation of dihydrotestosterone (DHT). Distances are provided in Å.
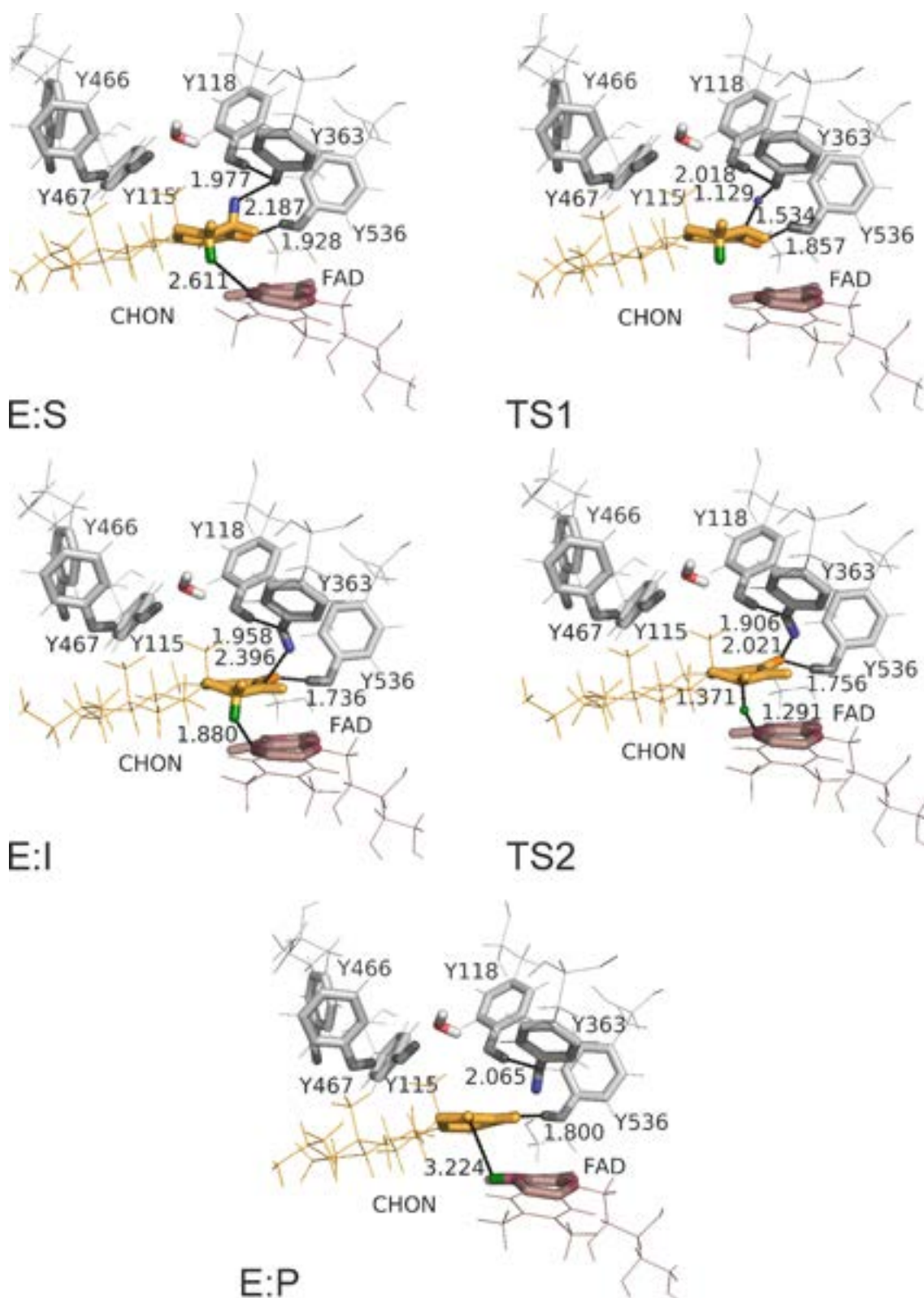
Figure S18. Reaction mechanism of Δ^1-dehydrogenation of cholest-4-en-3-one (CHON). Distances are provided in Å.
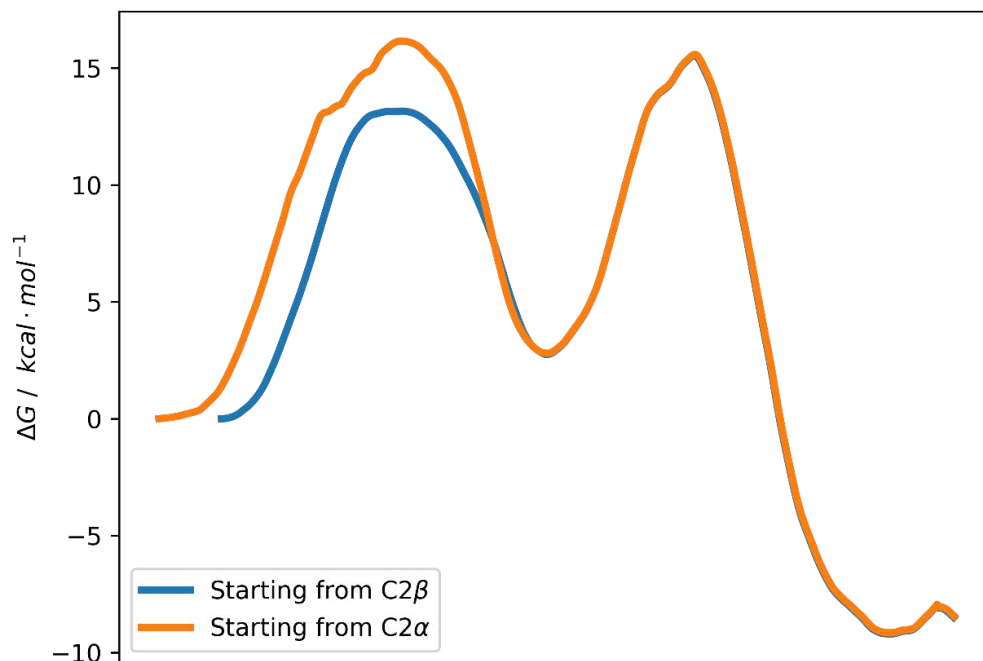
Figure S19. Free energy surfaces of Δ^1-dehydrogenation of androst-4-en-3,17-dione (AD) by AcmB with activation of 2Hα (orange line) or 2Hβ (blue line). Results obtained at B3LYP/6-311++g-(2d,2p):AM1/AMBER level of theory at 303 K.
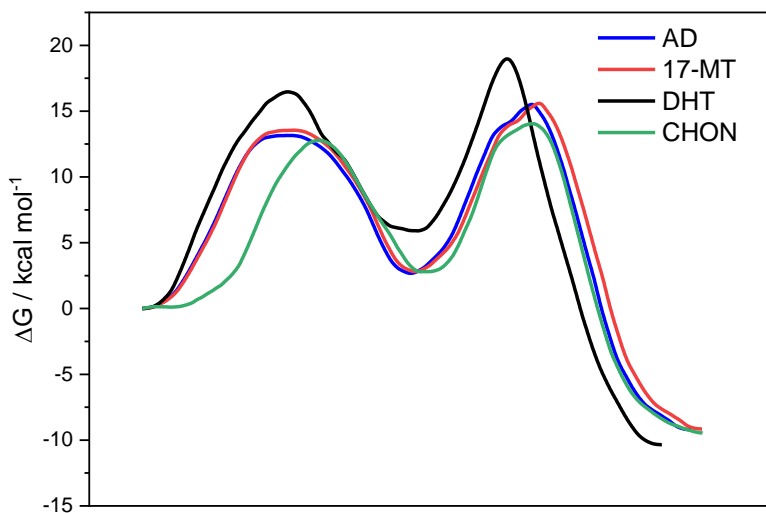
Figure S20. Free energy surfaces of Δ^1-dehydrogenation by AcmB of androsterone (AD, blue line), 17-methyltestosterone (17-MT, red line), dihydrotestosterone (DHT black line) and cholest-4-en-3-one (CHON green line). Results obtained at B3LYP/6-311++g-(2d,2p):AM1/AMBER level of theory at 303 K.

S34
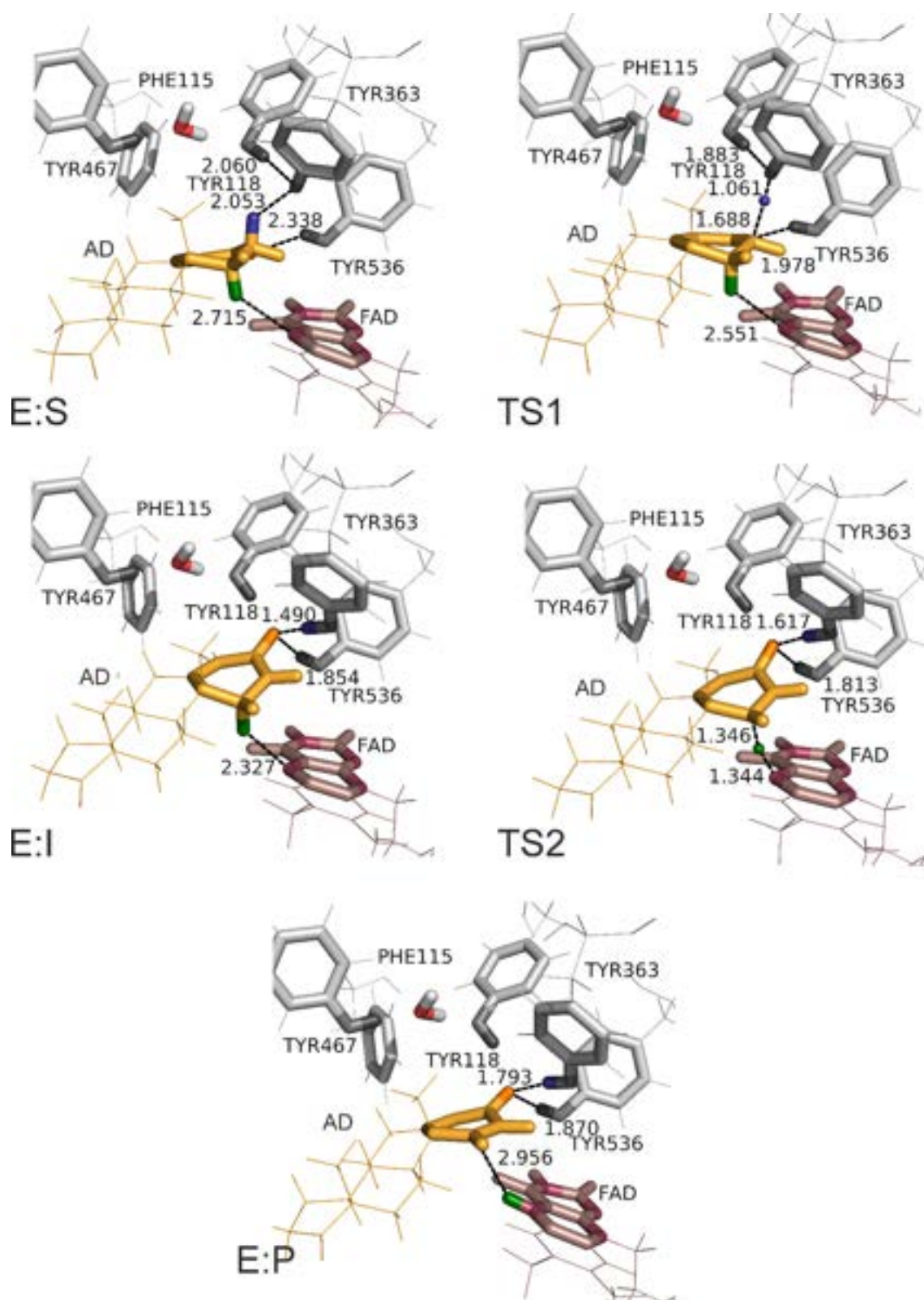
Figure S21. Reaction mechanism of Δ^1-dehydrogenation of androst-4-en-3,17-dione (AD) for mutated version of AcmB (Y115F). Distances are provided in Å.
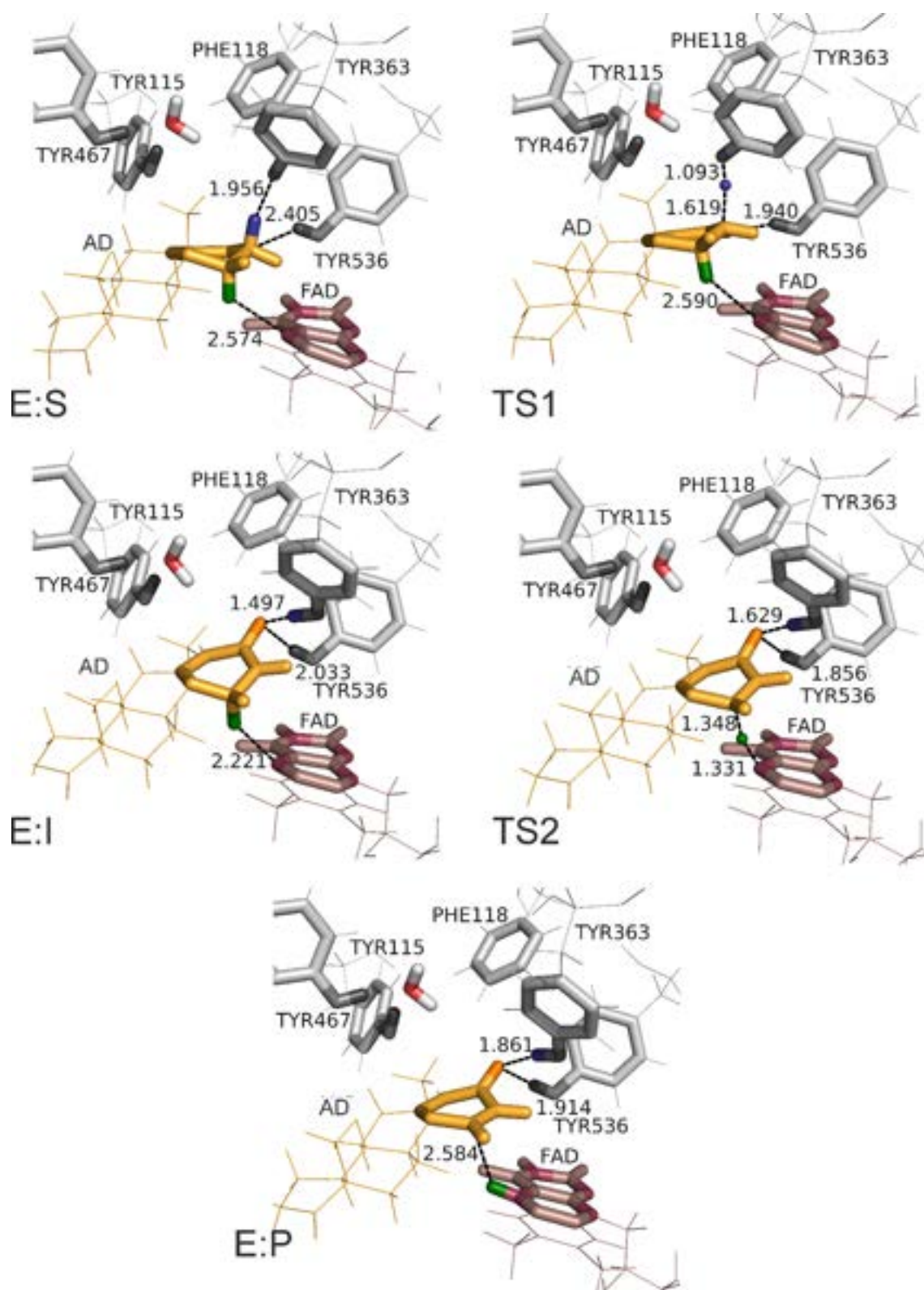
Figure S22. Reaction mechanism of Δ^1-dehydrogenation of androst-4-en-3,17-dione (AD) for mutated version of AcmB (Y118F). Distances are provided in Å.
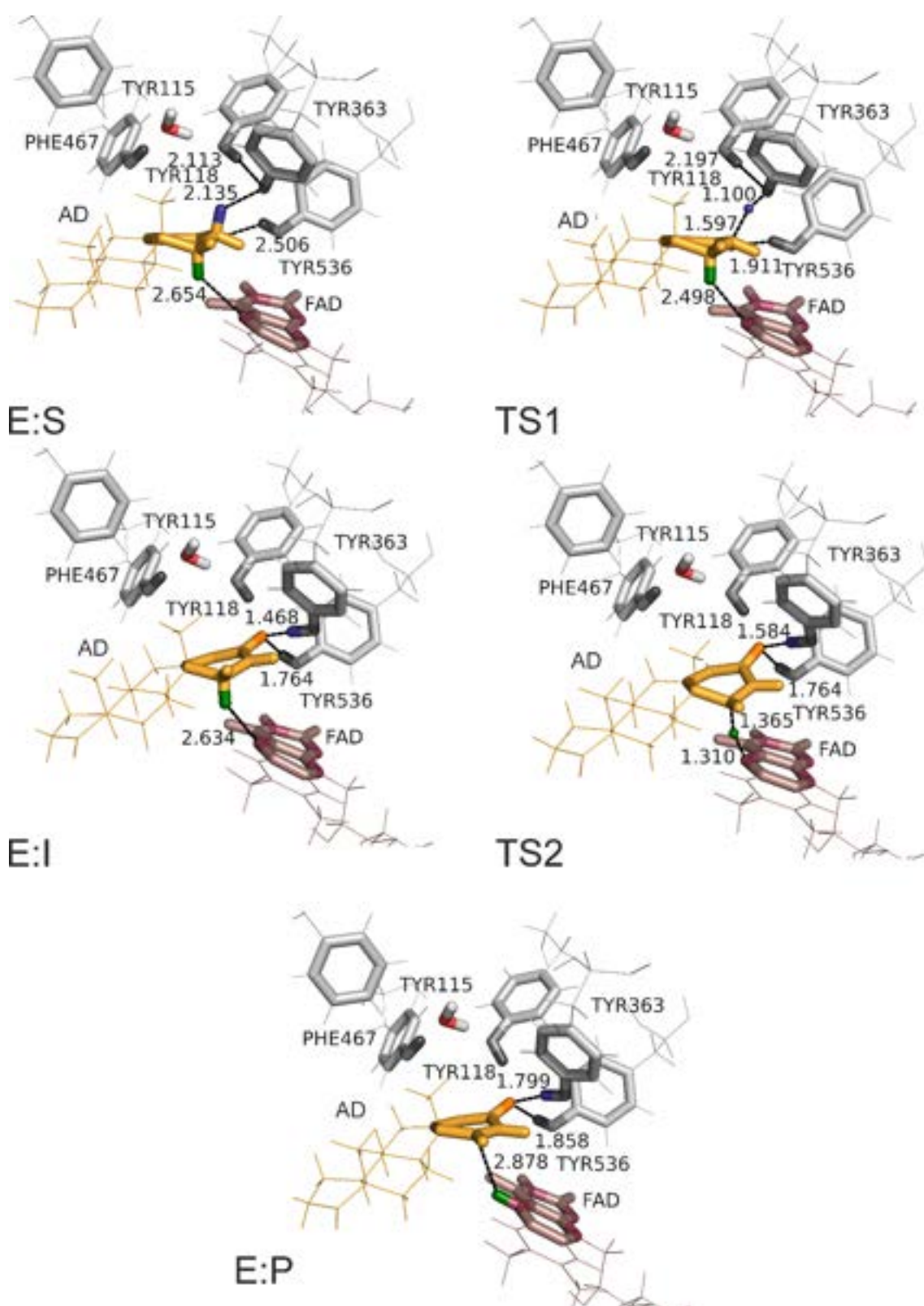
Figure S23. Reaction mechanism of Δ^1-dehydrogenation of androst-4-en-3,17-dione (AD) for mutated version of AcmB (Y467F). Distances are provided in Å.
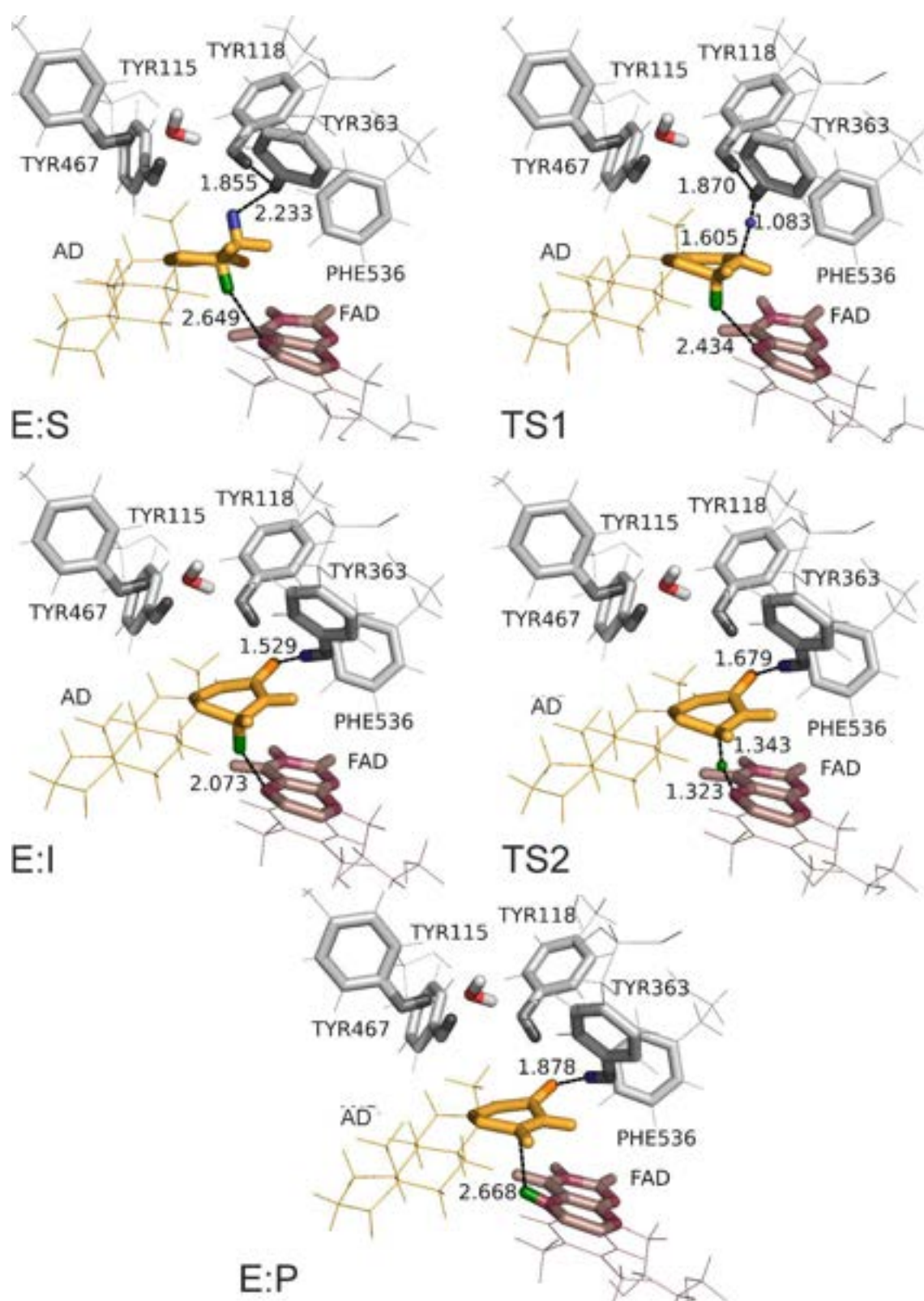
Figure S24. Reaction mechanism of Δ1-dehydrogenation of androst-4-en-3,17-dione (AD) for mutated version of AcmB (Y536F). Distances are provided in Å.
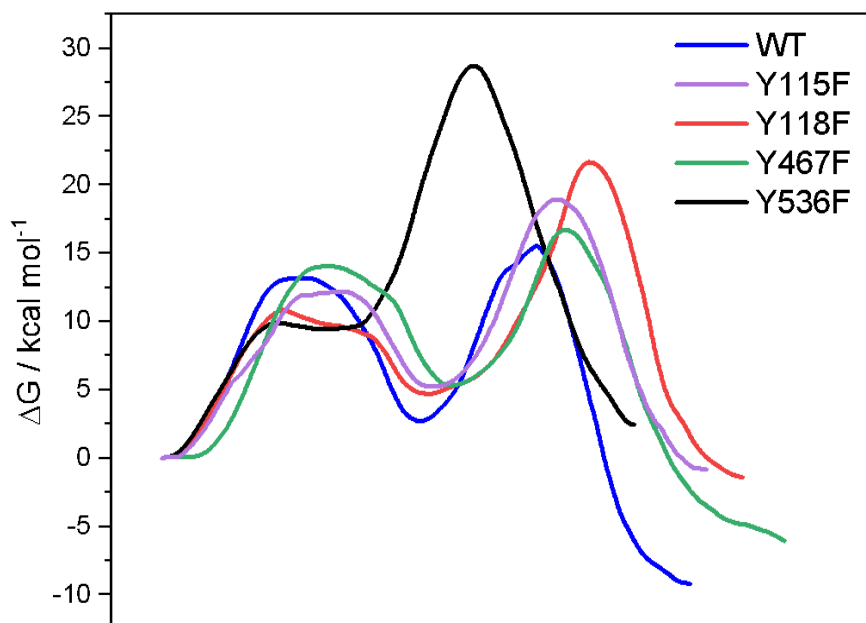
Figure S25. Free energy surfaces of Δ^1-dehydrogenation of of androst-4-en-3,17-dione (AD) by WT AcmB (blue) and Y115F (violet), Y118F (red), Y467F (green), Y536F (black). Results obtained at B3LYP/6-311++g-(2d,2p):AM1/AMBER level of theory at 303 K.

Table S8. Theoretical QM/MM (B3LYP/AMBER) of KIEs obtained for deuterium-labelled 17-MT and DHT.

	theoretical KIE	Std error
	17-MT/2,2,4,6,6-d5-17-MT	
S→TS1	5.31	0.19
I→TS1	3.55	0.20
I→TS2	0.70	0.03
S→TS2	1.05	0.02
	DHT/1α,16,16,17-d$_4$ -DHT	
S→TS1	1.12	0.02
I→TS1	0.95	0.02
I→TS2	4.23	0.03
S→TS2	4.97	0.04

Table S9. Detailed statistics of computed KIE for DHT.

Stage and substituted atoms	Mean KIE	Median KIE	STD KIE
B3LYP S -> TS1 all	1.12057	1.11806	0.02202
B3LYP S -> TS2 all	4.96645	4.95770	0.03655
B3LYP I -> TS1 all	0.95494	0.95519	0.01819
B3LYP I -> TS2 all	4.23235	4.23154	0.02363
B3LYP S -> TS1 H20	1.12428	1.12418	0.01567
B3LYP S -> TS2 H20	4.98720	4.99028	0.05481
B3LYP I -> TS1 H20	0.95685	0.95669	0.01331
B3LYP I -> TS2 H20	4.24452	4.24744	0.04650
B3LYP S -> TS1 H26	1.00562	1.00241	0.01251
B3LYP S -> TS2 H26	1.00978	1.00831	0.00989
B3LYP I -> TS1 H26	0.99699	0.99365	0.00819
B3LYP I -> TS2 H26	1.00111	1.00133	0.00327
B3LYP S -> TS1 H1	0.99198	0.99337	0.00834
B3LYP S -> TS2 H1	0.98561	0.98628	0.00879
B3LYP I -> TS1 H1	1.00209	1.00304	0.00982
B3LYP I -> TS2 H1	0.99566	0.99495	0.01017
B3LYP S -> TS1 H2	0.99907	0.99899	0.00315
B3LYP S -> TS2 H2	1.00052	1.00071	0.00303
B3LYP I -> TS1 H2	0.99918	0.99984	0.00421
B3LYP I -> TS2 H2	1.00063	1.00201	0.00411

Table S10. Detailed statistics of computed KIE for 17-MT.

Stage and substituted atoms	Mean KIE	Median KIE	STD KIE
B3LYP S -> TS1 all	5.30720	5.23805	0.19380
B3LYP S -> TS2 all	1.04553	1.04550	0.01804
B3LYP I -> TS1 all	3.54875	3.57063	0.19672
B3LYP I -> TS2 all	0.69911	0.70650	0.03154
B3LYP S -> TS1 H25	4.32011	4.25166	0.16708
B3LYP S -> TS2 H25	0.83143	0.83261	0.01064
B3LYP I -> TS1 H25	3.73748	3.72849	0.20545
B3LYP I -> TS2 H25	0.71930	0.73225	0.02955
B3LYP S -> TS1 H20	1.15814	1.15884	0.01287
B3LYP S -> TS2 H20	1.08402	1.08204	0.00879
B3LYP I -> TS1 H20	1.02740	1.02956	0.00895
B3LYP I -> TS2 H20	0.96164	0.96145	0.00411
B3LYP S -> TS1 H19	1.00263	1.00288	0.00685
B3LYP S -> TS2 H19	1.04205	1.04165	0.00762
B3LYP I -> TS1 H19	0.96566	0.96619	0.00520
B3LYP I -> TS2 H19	1.00362	1.00228	0.00601
B3LYP S -> TS1 H18	1.00752	1.00675	0.01035
B3LYP S -> TS2 H18	1.03679	1.03629	0.00843
B3LYP I -> TS1 H18	0.97958	0.97674	0.01208
B3LYP I -> TS2 H18	1.00804	1.00551	0.01070
B3LYP S -> TS1 H17	1.03462	1.03709	0.01823
B3LYP S -> TS2 H17	1.05877	1.06155	0.01846
B3LYP I -> TS1 H17	0.97692	0.97718	0.00356
B3LYP I -> TS2 H17	0.99972	1.00028	0.00264

Acknowledgement

References

(1) Sofińska, K.; Wojtkiewicz, A. M.; Wójcik, P.; Zastawny, O.; Guzik, M.; Winiarska, A.; Waligórski, P.; Cieśla, M.; Barbasz, J.; Szaleniec, M. Investigation of Quaternary Structure of Aggregating 3-Ketosteroid Dehydrogenase from Sterolibacterium Denitrificans: In the Pursuit of Consensus of Various Biophysical Techniques. *Biochim. Biophys. Acta - Gen. Subj.* **2019**, *1863* (6). https://doi.org/10.1016/j.bbagen.2019.03.009.

(2) Glanowski, M.; Wojcik, P.; Procner, M.; Borowski, T.; Lupa, D.; Mielczarek, P.; Oszajca, M.; Swiderek, K.; Moliner, V.; Bojarski, A. J.; Szaleniec, M. Enzymatic Δ1-Dehydrogenation of 3-Ketosteroids-Reconciliation of Kinetic Isotope Effects with the Reaction Mechanism. *ACS Catal.* **2021**, *11*, 8211–8225. https://doi.org/10.1021/acscatal.1c01479.

(3) Bradford, M. M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal. Biochem.* **1976**, *72*, 248–254. https://doi.org/S0003269776699996 [pii].

(4) Chen, V. B.; Arendall III, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: All-Atom Structure Validation for Macromolecular Crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2009**, *66*, 12–21. https://doi.org/10.1107/S0907444909042073.

(5) Kyte, J.; Doolittle, R. F. A Simple Method for Displaying the Hydropathic Character of a Protein. *J. Mol. Biol.* **1982**, *157* (1), 105–132. https://doi.org/http://dx.doi.org/10.1016/0022-2836(82)90515-0.

(6) Rohman, A.; Van Oosterwijk, N.; Thunnissen, A. M. W. H.; Dijkstra, B. W. Crystal Structure and Site-Directed Mutagenesis of 3-Ketosteroid Δ1-Dehydrogenase from Rhodococcus Erythropolis SQ1 Explain Its Catalytic Mechanism. *J. Biol. Chem.* **2013**, *288* (49), 35559–35568. https://doi.org/10.1074/jbc.M113.522771.

(7) Wójcik, P.; Glanowski, M.; Wojtkiewicz, A. M.; Rohman, A.; Szaleniec, M. Universal Capability of 3-Ketosteroid Δ1-Dehydrogenases to Catalyze Δ1-Dehydrogenation of C17-Substituted Steroids. *Microb. Cell Fact.* **2021**, *20* (1), 1–12. https://doi.org/10.1186/s12934-021-01611-5.

Publikacja P1

Oświadczenia autorów

Oświadczenie

Ja niżej podpisany, Michał Glanowski, współautor publikacji:

M. Glanowski, S. Kachhap, T. Borowski, M. Szaleniec, "Model Setup and Procedures for Prediction of Enzyme Reaction Kinetics with QM-Only and QM:MM Approaches", Computational Methods for Estimating the Kinetic Parameters of Biological Systems. Methods in Molecular Biology, 2385 (2022) 175-236, doi: 10.1007/978-1-0716-1767-0_10.

Oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- Autorstwo sekcji 2.1 Parameterization of Nonstandard Residues, 2.2 Parametrization of Nonstandard Amino Acid, 7 QMMM MD Simulation oraz 8 Prediction of Kinetic Isotope Effects

Michał Glanowski

Oświadczenie

Ja niżej podpisana, Sangita Kachhap, współautorka publikacji:

M. Glanowski, S. Kachhap, T. Borowski, M. Szaleniec, "Model Setup and Procedures for Prediction of Enzyme Reaction Kinetics with QM-Only and QM:MM Approaches", Computational Methods for Estimating the Kinetic Parameters of Biological Systems. Methods in Molecular Biology, 2385 (2022) 175-236, doi: 10.1007/978-1-0716-1767-0_10.

Oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- Autorstwo sekcji 2.3 Parameterization of the Metal Center, 3 MD Simulation oraz 9 Notes

Sangita Kachhap

Oświadczenie

Ja niżej podpisany, Tomasz Borowski, współautor publikacji:

M. Glanowski, S. Kachhap, T. Borowski, M. Szaleniec, "Model Setup and Procedures for Prediction of Enzyme Reaction Kinetics with QM-Only and QM:MM Approaches", Computational Methods for Estimating the Kinetic Parameters of Biological Systems. Methods in Molecular Biology, 2385 (2022) 175-236, doi: 10.1007/978-1-0716-1767-0_10.

Oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- Autorstwo sekcji 6 QM:MM Modeling oraz 9 Notes

Tomasz Borowski

Oświadczenie

Ja niżej podpisany, Maciej Szaleniec, współautor publikacji:

M. Glanowski, S. Kachhap, T. Borowski, M. Szaleniec, "Model Setup and Procedures for Prediction of Enzyme Reaction Kinetics with QM-Only and QM:MM Approaches", Computational Methods for Estimating the Kinetic Parameters of Biological Systems. Methods in Molecular Biology, 2385 (2022) 175-236, doi: 10.1007/978-1-0716-1767-0_10.

Oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- Opracowanie ogólnej koncepcji rozdziału
- Autorstwo sekcji 4 Cluster-Based QM Calculations, 5 Correlation of Modeling with Kinetic Experiment oraz 9 Notes
- Edycja całości manuskryptu i odpowiedzi na pytania recenzentów

Maciej Szaleniec

Publikacja P2

Oświadczenia autorów

Oświadczenie

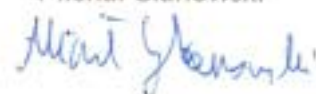Ja niżej podpisana, Patrycja Wójcik, współautorka publikacji:

Patrycja Wójcik, Michał Glanowski, Agnieszka M. Wojtkiewicz, Ali Rohman, Maciej Szaleniec
„Universal capability of 3-ketosteroid Δ^1-dehydrogenases to catalyze Δ^1-dehydrogenation of C17-substituted steroids"
Microb. Cell Fact., 20 (2021) 119, DOI: 10.21203/rs.3.rs-317042/v1

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- zaprojektowanie eksperymentów,
- produkcję oraz oczyszczanie AcmB i KSTD1,
- przeprowadzenie eksperymentów katalitycznych z wykorzystaniem AcmB oraz KSTD1,
- wyznaczanie stałych trwałości kompleksów HBC–steroid,
- analizę otrzymanych danych
oraz
- redakcję wstępu, metod, wyników i dyskusji.

Patrycja Wójcik

Oświadczenie

Ja niżej podpisany, Michał Glanowski, współautor publikacji:

Patrycja Wójcik, Michał Glanowski, Agnieszka M. Wojtkiewicz, Ali Rohman, Maciej Szaleniec
„Universal capability of 3-ketosteroid Δ1-dehydrogenases to catalyze Δ1-dehydrogenation of C17-substituted steroids"
Microb. Cell Fact., 20 (2021) 119, DOI: 10.21203/rs.3.rs-317042/v1

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:

przeprowadzenie obliczeń i ich analize (symulacje dynamiki molekularnej oraz obliczeń MMPBSA), opis metodologii obliczeniowej.

Michał Glanowski

Oświadczenie

Ja niżej podpisana, Agnieszka Wojtkiewicz, współautorka publikacji:

Patrycja Wójcik, Michał Glanowski, Agnieszka M. Wojtkiewicz, Ali Rohman, Maciej Szaleniec
„Universal capability of 3-ketosteroid Δ1-dehydrogenases to catalyze Δ1-dehydrogenation of C17-substituted steroids"
Microb. Cell Fact., 20 (2021) 119, DOI: 10.21203/rs.3.rs-317042/v1

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- przygotowanie enzymu AcmB2 oraz przeprowadzenie eksperymentów z enzymem
- opis części poświęconych izoenzymowi AcmB tj. AcmB2
- edycję manuskryptu
- udział w napisaniu odpowiedzi dla recenzentów

PODPIS

Author contribution statement:

I , Ali Rohman, co-author of the paper:

state that my contribution was following:
- Development of KSTD1 expression system
- Edition of the manuscript text

Ali Rohman

Oświadczenie

Ja niżej podpisany, Maciej Szaleniec, współautorka publikacji:

Patrycja Wójcik, Michał Glanowski, Agnieszka M. Wojtkiewicz, Ali Rohman, Maciej Szaleniec
„Universal capability of 3-ketosteroid Δ1-dehydrogenases to catalyze Δ1-dehydrogenation of C17-substituted steroids"
Microb. Cell Fact., 20 (2021) 119, DOI: 10.21203/rs.3.rs-317042/v1

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- Sformułowanie wstępnej hipotezy
- Zaplanowanie badań przedstawionych w publikacji
- Przeanalizowanie wyników uzyskanych w eksperymencie i modelowaniu
- Współautorstwo oryginalnego manuskryptu (wstęp, wyniki, dyskusja)

Prof. dr hab. M. Szaleniec

Publikacja P3

Oświadczenia autorów

Oświadczenie

Ja niżej podpisany, Michał i Glanowski, współautor publikacji:

Michał Glanowski, Patrycja Wójcik, Magdalena Procner, Tomasz Borowski, Dawid Lupa, Przemysław Mielczarek, Maria Oszajca, Katarzyna Świderek, Vicent Moliner, Andrzej J. Bojarski, and Maciej Szaleniec „Enzymatic Δ1-Dehydrogenation of 3-Ketosteroids – Reconciliation of Kinetic Isotope Effects with the Reaction Mechanism" ACS Catal. 11 (2021) 8211–8225 DOI: 10.1021/acscatal.1c01479

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- przeprowadzenie wszystkich obliczeń pod opieką M. Szaleńca, K. Świderek i V. Molinera,
- analiza wyników i opracowanie dyskusji,
- współautorstwo głównego tekstu,
- przygotowanie rysunków.

Imię i nazwisko (podpis)

Oświadczenie

Ja niżej podpisana, Patrycja Wójcik, współautor publikacji:

Michał Glanowski, Patrycja Wójcik, Magdalena Procner, Tomasz Borowski, Dawid Lupa, Przemysław Mielczarek, Maria Oszajca, Katarzyna Świderek, Vicent Moliner, Andrzej J. Bojarski, and Maciej Szaleniec „Enzymatic Δ1-Dehydrogenation of 3-Ketosteroids – Reconciliation of Kinetic Isotope Effects with the Reaction Mechanism" ACS Catal. 11 (2021) 8211–8225 DOI: 10.1021/acscatal.1c01479

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- ekspresję i oczyszczanie KSTD
- przeprowadzenie eksperymentów KSVE i KIE
- analizę danych eksperymentalnych
- współautorstwo w tworzeniu sekcji Metody i Wyniki
- edycję tekstu
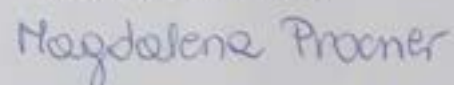- przygotowywanie wykresów

Oświadczenie

Ja niżej podpisany, Magdalena Procner, współautor publikacji:

Michał Glanowski, Patrycja Wójcik, Magdalena Procner, Tomasz Borowski, Dawid Lupa, Przemysław Mielczarek, Maria Oszajca, Katarzyna Świderek, Vicent Moliner, Andrzej J. Bojarski, and Maciej Szaleniec „Enzymatic Δ1-Dehydrogenation of 3-Ketosteroids – Reconciliation of Kinetic Isotope Effects with the Reaction Mechanism" ACS Catal. 11 (2021) 8211–8225 DOI: 10.1021/acscatal.1c01479

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- współautorstwo w wykonaniu i opracowaniu eksperymentów techniką zatrzymanego przepływu (stopped-flow),
- współautorstwo w opracowaniu sekcji „Metody".

Magdalena Procner

Magdalena Procner

Oświadczenie

Ja niżej podpisany, TOMASZ BOROWSKI, współautor publikacji:

M. Glanowski, P. Wójcik, M. Procner, T. Borowski, D. Lupa, P. 0Mielczarek, M. Oszajca, K. Świderek, V. Moliner, A.J. Bojarski, 0M. Szaleniec 0"Enzymatic Δ1-Dehydrogenation of 03-Ketosteroids - Reconciliation of Kinetic Isotope Effects with 0the Reaction Mechanism"

ACS Catal., 2021, 11, 13, 8211–8225
DOI: 10.1021/acscatal.1c01479

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- współautorstwo części dotyczącej modelowania kinetyki enzymatycznej

PODPIS

Oświadczenie

Ja niżej podpisany, Dawid Lupa, współautor publikacji:

Michał Glanowski, Patrycja Wójcik, Magdalena Procner, Tomasz Borowski, Dawid Lupa, Przemysław Mielczarek, Maria Oszajca, Katarzyna Świderek, Vicent Moliner, Andrzej J. Bojarski, and Maciej Szaleniec „*Enzymatic Δ1-Dehydrogenation of 3-Ketosteroids – Reconciliation of Kinetic Isotope Effects with the Reaction Mechanism*" ACS Catal. 11 (2021) 8211–8225 DOI: 10.1021/acscatal.1c01479

Oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:

- Przeprowadzenie pomiarów współczynnika dyfuzji cząsteczek dehydrogenazy ketosteroidowej oraz poli(glikolu etylenowego) o różnych masach cząsteczkowych

- Wyznaczenie wartości lepkości dynamicznej roztworów dehydrogenazy ketosteroidowej, glicerolu oraz poli(glikolu etylenowego) o różnych masach cząsteczkowych

- Ilościowe opracowanie otrzymanych wyników badań (Figure 5 oraz Fig. S26)

- Współudział w redakcji manuskryptu oraz *Supporting Information*
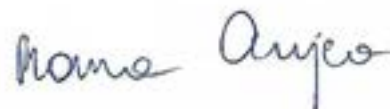
Imię i nazwisko (podpis)

Oświadczenie

Ja niżej podpisany, PRZEMYSŁAW MIELCZAREK, współautor publikacji:

Michał Glanowski, Patrycja Wójcik, Magdalena Procner, Tomasz Borowski, Dawid Lupa, Przemysław Mielczarek, Maria Oszajca, Katarzyna Świderek, Vicent Moliner, Andrzej J. Bojarski, and Maciej Szaleniec „Enzymatic Δ1-Dehydrogenation of 3-Ketosteroids – Reconciliation of Kinetic Isotope Effects with the Reaction Mechanism" ACS Catal. 11 (2021) 8211–8225 DOI: 10.1021/acscatal.1c01479

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- Lista wkładów wg Author Contributions w publikacji

Przemysław Mielczarek
Imię i nazwisko (podpis)

Oświadczenie

Ja niżej podpisana, Maria Oszajca, współautor publikacji:

Michał Glanowski, Patrycja Wójcik, Magdalena Procner, Tomasz Borowski, Dawid Lupa, Przemysław Mielczarek, Maria Oszajca, Katarzyna Świderek, Vicent Moliner, Andrzej J. Bojarski, and Maciej Szaleniec „Enzymatic Δ1-Dehydrogenation of 3-Ketosteroids – Reconciliation of Kinetic Isotope Effects with the Reaction Mechanism" ACS Catal. 11 (2021) 8211–8225 DOI: 10.1021/acscatal.1c01479

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- nadzór nad wykonaniem badań kinetycznych z wykorzystaniem techniki zatrzymanego przepływu,
- edycja oryginalnego tekstu publikacji

Imię i nazwisko (podpis)

2021-06-21

Author contribution statement:

I , Vicent Moliner, co-author of the paper:

state that my contribution was following:
- supervision of calculation,
- co-authoring method section
- text edition
- project funding
- response to reviewers' answers

VICENTE|
MOLINER|
IBAÑEZ

Firmado digitalmente por
VICENTE|MOLINER|IBAÑEZ
Fecha: 2021.06.21 11:58:28
+02'00'

Vicent Moliner

(signature)

Author contribution statement:

I , Katarzyna Świderek, co-author of the paper:

 Michał Glanowski, Patrycja Wójcik, Magdalena Procner, Tomasz Borowski, Dawid Lupa, Przemysław Mielczarek, Maria Oszajca, Katarzyna Świderek, Vicent Moliner, Andrzej J. Bojarski, and Maciej Szaleniec „Enzymatic Δ1-Dehydrogenation of 3-Ketosteroids – Reconciliation of Kinetic Isotope Effects with the Reaction Mechanism" ACS Catal. 11 (2021) 8211–8225 DOI: 10.1021/acscatal.1c01479

state that my contribution was following:
- supervision of calculation,
- co-authoring method section
- text edition
- project funding
- response to reviewers' answers

KATARZYNA PATRYCJA| SWIDEREK

Elektronicznie podpisany przez KATARZYNA PATRYCJA|SWIDEREK Data: 2021.06.21 12:38:12 +02'00'

Katarzyna Świderek

(signature)

Oświadczenie

Ja niżej podpisany, Andrzej Bojarski, współautor publikacji:

Michał Glanowski, Patrycja Wójcik, Magdalena Procner, Tomasz Borowski, Dawid Lupa, Przemysław Mielczarek, Maria Oszajca, Katarzyna Świderek, Vicent Moliner, Andrzej J. Bojarski, and Maciej Szaleniec „Enzymatic Δ1-Dehydrogenation of 3-Ketosteroids – Reconciliation of Kinetic Isotope Effects with the Reaction Mechanism" ACS Catal. 11 (2021) 8211–8225 DOI: 10.1021/acscatal.1c01479

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- opiekę naukową nad MG
- współfinansowanie badań
- współredakcję i korektę manuskryptu

Andrzej Bojarski

Oświadczenie

Ja niżej podpisany, Maciej Szaleniec, współautor publikacji:

 Michał Glanowski, Patrycja Wójcik, Magdalena Procner, Tomasz Borowski, Dawid Lupa, Przemysław Mielczarek, Maria Oszajca, Katarzyna Świderek, Vicent Moliner, Andrzej J. Bojarski, and Maciej Szaleniec „Enzymatic Δ1-Dehydrogenation of 3-Ketosteroids – Reconciliation of Kinetic Isotope Effects with the Reaction Mechanism" ACS Catal. 11 (2021) 8211−8225 DOI: 10.1021/acscatal.1c01479

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- Zaplanowanie projektu badawczego
- Zaplanowanie obliczeń i eksperymentów
- Nadzór nad Michałem Glanowskim, Patrycją Wójcik, Magdaleną Procner,
- Analizę wyników teoretycznych i eksperymentalnych
- Współautorstwo tekstu publikacji
- Edycję publikacji
- Odpowiedzi na recenzje
- Zapewnienie finansowania projektu

Maciej Szaleniec

Publikacja P4

Oświadczenia autorów

Oświadczenie

Ja niżej podpisana, Patrycja Wójcik, współautorka publikacji:

Wójcik, Patrycja; Glanowski, Michał; Mrugała, Beata; Procner, Magdalena; Zastawny, Olga; Flejszar, Monika; Kurpiewska, Katarzyna; Niedzialkowska, Ewa; Minor, Wladek; Oszajca, Maria; Bojarski, Andrzej; Wojtkiewicz, Agnieszka; Szaleniec, Maciej, "Structure, mutagenesis and QM:MM modelling of 3-ketosteroid Δ^1-dehydrogenase from *Sterolibacterium denitrificans* – the role of new putative membrane-associated domain and proton-relay system in catalysis", Biochemistry 2022

DOI: 10.1021/acs.biochem.2c00576

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- Oczyszczenie i krystalizację enzymu, oraz rozwiązanie i udokładnienie jego struktury.
- Opracowanie metod LC-MS oraz przeprowadzenie następujących eksperymentów kinetycznych: określenie optimum pH dla reakcji połówkowej redukcji, pomiar aktywności mutantów, pomiar kinetycznego efekt izotopowego.
- Analizę uzyskanych przeze mnie wyników.
- Udział w edycji manuskryptu.
- Stworzenie grafiki do rysunków 3 oraz 4.

Patrycja Wójcik

Oświadczenie

Ja niżej podpisany, Michał Glanowski, współautor publikacji:

Wójcik, Patrycja; Glanowski, Michał; Mrugała, Beata; Procner, Magdalena; Zastawny, Olga; Flejszar, Monika; Kurpiewska, Katarzyna; Niedzialkowska, Ewa; Minor, Wladek; Oszajca, Maria; Bojarski, Andrzej; Wojtkiewicz, Agnieszka; Szaleniec, Maciej, "Structure, mutagenesis and QM:MM modelling of 3-ketosteroid Δ^1-dehydrogenase from *Sterolibacterium denitrificans* – the role of new putative membrane-associated domain and proton-relay system in catalysis", Biochemistry 2022

DOI: 10.1021/acs.biochem.2c00576

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:

- Udział w edycji manuskryptu
- Przygotowanie grafik do Figure 5. w tekście głównym manuskryptu oraz do SI
- Przeprowadzenie wszystkich obliczeń opisanych w publikacji: symulacje MD, MMPBSA, QMMM i QMMM MD i analiza ich wyników
- Udział w dyskusji wyników

Kraków 2023-01-11

Oświadczenie

Ja niżej podpisana, Beata Mrugała, współautor publikacji:

Wójcik, Patrycja; Glanowski, Michał; Mrugała, Beata; Procner, Magdalena; Zastawny, Olga; Flejszar, Monika; Kurpiewska, Katarzyna; Niedziałkowska, Ewa; Minor, Wladek; Oszajca, Maria; Bojarski, Andrzej; Wojtkiewicz, Agnieszka; Szaleniec, Maciej, "Structure, mutagenesis and QM:MM modelling of 3-ketosteroid Δ¹-dehydrogenase from *Sterolibacterium denitrificans* – the role of new putative membrane-associated domain and proton-relay system in catalysis", Biochemistry 2022

DOI: 10.1021/acs.biochem.2c00576

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- Przeprowadziłem eksperyment krystalizacji i kierowałam optymalizacją
- Rozwiązałem strukturę białka AcmB

Author Contributions

P.W. purified and crystalized the enzyme, determined and refined the enzyme structure, developed LC-MS methods, conducted kinetics (pH, mutated variants, kinetic isotope effect), analysed results, co-authored and edited manuscript, and created figures

M.G. conducted all calculations, analysed results, developed discussion, co-authored the main text, edited text, and created figures

B.M. refined the structure of AcmB, co-authored the manuscript

M.P. conducted stopped-flow steady-state and pre-steady state kinetics, co-authored method section

O.Z. developed AcmB mutant variants, expressed and purified the enzymes,

M.F. conducted stopped-flow steady-state Ping Pong kinetics, analysed results

K.K. refined the structure of AcmB, co-authored the manuscript (crystallographic sections)

M.O. provided assistance in pre-steady state kinetics, edited the manuscript

A.B. supervised P.W., provided funding, edited the manuscript

E.N. developed conditions for enzyme crystallization, edited the manuscript

W.M. oversaw initial experiments and edited the manuscript

A.M.W. purified the enzyme, conducted the bioinformatic analysis, co-authored the bioinformatic analysis section, created figures

M.S. designed the study, provided funding, supervised M.G, P.W., M.P., O.Z., M. F., analysed and curated all data, co-authored and edited the manuscript,

Beata Mrugała

Oświadczenie

Ja niżej podpisana, Magdalena Procner, współautor publikacji:

Wójcik, Patrycja; Glanowski, Michał; Mrugała, Beata; Procner, Magdalena; Zastawny, Olga; Flejszar, Monika; Kurpiewska, Katarzyna; Niedzialkowska, Ewa; Minor, Wladek; Oszajca, Maria; Bojarski, Andrzej; Wojtkiewicz, Agnieszka; Szaleniec, Maciej, "Structure, mutagenesis and QM:MM modelling of 3-ketosteroid Δ^1-dehydrogenase from *Sterolibacterium denitrificans* – the role of new putative membrane-associated domain and proton-relay system in catalysis", Biochemistry 2022

DOI: 10.1021/acs.biochem.2c00576

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:

- przeprowadzenie eksperymentów techniką zatrzymanego przepływu (pomiary kinetyki dla stanu stacjonarnego i przedstacjonarnego)
- współautorstwo sekcji Metody

Magdalena Procner

Oświadczenie

Ja niżej podpisana, Olga Zastawny, współautor publikacji:

Wójcik, Patrycja; Glanowski, Michał; Mrugała, Beata; Procner, Magdalena; Zastawny, Olga; Flejszar, Monika; Kurpiewska, Katarzyna; Niedzialkowska, Ewa; Minor, Wladek; Oszajca, Maria; Bojarski, Andrzej; Wojtkiewicz, Agnieszka; Szaleniec, Maciej, "Structure, mutagenesis and QM:MM modelling of 3-ketosteroid Δ^1-dehydrogenase from *Sterolibacterium denitrificans* – the role of new putative membrane-associated domain and proton-relay system in catalysis", Biochemistry 2022

DOI: 10.1021/acs.biochem.2c00576

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- uzyskałam zmutowane warianty białka AcmB;
- przeprowadziłam ekspresję oraz oczyszczanie zmienionych form enzymu.

Olga Zastawny

Oświadczenie

Ja niżej podpisana, MONIKA FLEJSZAR, współautorka publikacji:

Wójcik, Patrycja; Glanowski, Michał; Mrugała, Beata; Procner, Magdalena; Zastawny, Olga; Flejszar, Monika; Kurpiewska, Katarzyna; Niedzialkowska, Ewa; Minor, Wladek; Oszajca, Maria; Bojarski, Andrzej; Wojtkiewicz, Agnieszka; Szaleniec, Maciej, "Structure, mutagenesis and QM:MM modelling of 3-ketosteroid Δ¹-dehydrogenase from *Sterolibacterium denitrificans* – the role of new putative membrane-associated domain and proton-relay system in catalysis", Biochemistry 2022

DOI: 10.1021/acs.biochem.2c00576

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- przeprowadzenie pomiarów techniką zatrzymanego przepływu (ang. stopped flow) mających na celu zbadanie kinetyki reakcji dwusubstratowej oraz określenie jej typu mechanizmu („Ping-Pong"), a także analizę uzyskanych wyników.

Monika Flejszar

Oświadczenie

Ja niżej podpisana, Katarzyna Kurpiewska, współautor publikacji:

Wójcik, Patrycja; Glanowski, Michał; Mrugała, Beata; Procner, Magdalena; Zastawny, Olga; Flejszar, Monika; Kurpiewska, Katarzyna; Niedzialkowska, Ewa; Minor, Wladek; Oszajca, Maria; Bojarski, Andrzej; Wojtkiewicz, Agnieszka; Szaleniec, Maciej, "Structure, mutagenesis and QM:MM modelling of 3-ketosteroid Δ^1-dehydrogenase from *Sterolibacterium denitrificans* – the role of new putative membrane-associated domain and proton-relay system in catalysis", Biochemistry 2022

DOI: 10.1021/acs.biochem.2c00576

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- Brałam udział w opracowaniu krystalograficznej części badań (udokładnienie struktury enzymu) oraz przygotowaniu odpowiedniej części manuskryptu (badania strukturalne)
- Stworzyłam grafikę do rysunku 3A.

Oświadczenie

Ja niżej podpisana, Ewa Niedziałkowska, współautor publikacji:

Wójcik, Patrycja; Glanowski, Michał; Mrugała, Beata; Procner, Magdalena; Zastawny, Olga; Flejszar, Monika; Kurpiewska, Katarzyna; Niedziałkowska, Ewa; Minor, Wladek; Oszajca, Maria; Bojarski, Andrzej; Wojtkiewicz, Agnieszka; Szaleniec, Maciej, "Structure, mutagenesis and QM:MM modelling of 3-ketosteroid Δ^1-dehydrogenase from *Sterolibacterium denitrificans* – the role of new putative membrane-associated domain and proton-relay system in catalysis", Biochemistry 2022

DOI: 10.1021/acs.biochem.2c00576

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
-
- Zoptymalizowałam początkowe warunki krystalizacji białka i brałam udział w edycji manuskryptu

Ewa Niedziałkowska

Oświadczenie

Ja niżej podpisany, Wladek Minor, współautor publikacji:

Wójcik, Patrycja; Glanowski, Michał; Mrugała, Beata; Procner, Magdalena; Zastawny, Olga; Flejszar, Monika; Kurpiewska, Katarzyna; Niedzialkowska, Ewa; Minor, Wladek; Oszajca, Maria; Bojarski, Andrzej; Wojtkiewicz, Agnieszka; Szaleniec, Maciej, "Structure, mutagenesis and QM:MM modelling of 3-ketosteroid Δ^1-dehydrogenase from *Sterolibacterium denitrificans* – the role of new putative membrane-associated domain and proton-relay system in catalysis", Biochemistry 2022

DOI: 10.1021/acs.biochem.2c00576

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:

oversaw initial experiments and edited the manuscript

Author Contributions

P.W. purified and crystalized the enzyme, determined and refined the enzyme structure, developed LC-MS methods, conducted kinetics (pH, mutated variants, kinetic isotope effect), analysed results, co-authored and edited manuscript, and created figures
M.G. conducted all calculations, analysed results, developed discussion, co-authored the main text, edited text, and created figures
B.M. refined the structure of AcmB, co-authored the manuscript
M.P. conducted stopped-flow steady-state and pre-steady state kinetics, co-authored method section
O.Z. developed AcmB mutant variants, expressed and purified the enzymes,
M.F. conducted stopped-flow steady-state Ping Pong kinetics, analysed results
K.K. refined the structure of AcmB, co-authored the manuscript (crystallographic sections)
M.O. provided assistance in pre-steady state kinetics, edited the manuscript
A.B. supervised P.W., provided funding, edited the manuscript
E.N. developed conditions for enzyme crystallization, edited the manuscript
W.M. oversaw initial experiments and edited the manuscript
A.M.W. purified the enzyme, conducted the bioinformatic analysis, co-authored the bioinformatic analysis section, created figures
M.S. designed the study, provided funding, supervised M.G, P.W., M.P., O.Z., M. F., analysed and curated all data, co-authored and edited the manuscript,
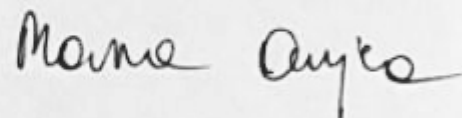
Wladek Minor
University of Virginia

Oświadczenie

Ja niżej podpisana, Maria Oszajca, współautor publikacji:

Wójcik, Patrycja; Glanowski, Michał; Mrugała, Beata; Procner, Magdalena; Zastawny, Olga; Flejszar, Monika; Kurpiewska, Katarzyna; Niedzialkowska, Ewa; Minor, Wladek; Oszajca, Maria; Bojarski, Andrzej; Wojtkiewicz, Agnieszka; Szaleniec, Maciej, "Structure, mutagenesis and QM:MM modelling of 3-ketosteroid Δ^1-dehydrogenase from *Sterolibacterium denitrificans* – the role of new putative membrane-associated domain and proton-relay system in catalysis", Biochemistry 2022

DOI: 10.1021/acs.biochem.2c00576

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- nadzór nad wykonaniem badań kinetycznych z wykorzystaniem techniki zatrzymanego przepływu,
- edycja oryginalnego tekstu publikacji

PODPIS

Oświadczenie

Ja niżej podpisany, Andrzej Bojarski, współautor publikacji:

Wójcik, Patrycja; Glanowski, Michał; Mrugała, Beata; Procner, Magdalena; Zastawny, Olga; Flejszar, Monika; Kurpiewska, Katarzyna; Niedzialkowska, Ewa; Minor, Wladek; Oszajca, Maria; Bojarski, Andrzej; Wojtkiewicz, Agnieszka; Szaleniec, Maciej, "Structure, mutagenesis and QM:MM modelling of 3-ketosteroid Δ^1-dehydrogenase from *Sterolibacterium denitrificans* – the role of new putative membrane-associated domain and proton-relay system in catalysis", Biochemistry 2022

DOI: 10.1021/acs.biochem.2c00576

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- Nadzorowałem pracę Patrycji Wójcik i Michała Glanowskiego
- Brałem udział w edycji i korekcje manuskryptu
- Brałem udział w zapewnieniu finansowania

Oświadczenie

Ja niżej podpisana, Agnieszka M. Wojtkiewicz, współautorka publikacji:

Wójcik, Patrycja; Glanowski, Michał; Mrugała, Beata; Procner, Magdalena; Zastawny, Olga; Flejszar, Monika; Kurpiewska, Katarzyna; Niedzialkowska, Ewa; Minor, Wladek; Oszajca, Maria; Bojarski, Andrzej; Wojtkiewicz, Agnieszka; Szaleniec, Maciej, "Structure, mutagenesis and QM:MM modelling of 3-ketosteroid Δ¹-dehydrogenase from *Sterolibacterium denitrificans* – the role of new putative membrane-associated domain and proton-relay system in catalysis", Biochemistry 2022

DOI: 10.1021/acs.biochem.2c00576

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- Przeprowadziłam oczyszczanie enzymu
- Przeprowadziłam analizę bioinformatyczną
- Napisałam sekcję analizy bioinformatycznej
- Brałam udział w edycji manuskryptu
- Stworzyłam grafikę do rysunków analizy bioinformatycznej
- Przeprowadziłam obliczenia w ramach analizy bioinformatycznej

Oświadczenie

Ja niżej podpisana, Maciej Szaleniec, współautor publikacji:

Wójcik, Patrycja; Glanowski, Michał; Mrugała, Beata; Procner, Magdalena; Zastawny, Olga; Flejszar, Monika; Kurpiewska, Katarzyna; Niedzialkowska, Ewa; Minor, Wladek; Oszajca, Maria; Bojarski, Andrzej; Wojtkiewicz, Agnieszka; Szaleniec, Maciej, "Structure, mutagenesis and QM:MM modelling of 3-ketosteroid Δ^1-dehydrogenase from *Sterolibacterium denitrificans* – the role of new putative membrane-associated domain and proton-relay system in catalysis", Biochemistry (2022)

DOI: 10.1021/acs.biochem.2c00576

oświadczam, że moje zaangażowanie w powstanie w/w publikacji obejmowało:
- Zaprojektowanie badań
- Zapewnienie finansowania
- Nadzór merytoryczny nad pracą Michała Glanowskiego, Patrycji Wójcik, Magdaleny Procner, Olgi Zastawny, Moniki Flejszar,
- Analizę danych
- Współautorstwo pierwotnej wersji manuskryptu
- Edycję manuskryptu
- Korespondencję z redakcją, odpowiedzi dla recenzentów

Maciej Szaleniec