

Карманы связывания, докинг

В данном практикуме я работала с предпроцессированным объектом под [номером 17](#).

1. Карманы, изменения

Поверхность данного белка имеет довольно сложную форму (рис. 1, слева): с одной стороны глубокая поперечная борозда, внутри которой находится углубление под лиганд, а помимо борозды — выемка сверху справа, обратная сторона белка относительно ровная (не показано). С такой сложной поверхностью связано разнообразие и неоднозначность выделения карманов. В связанном белке GSiteScorer выделяет два достаточно значимых (simple score > 0.3) кармана: крупный P_0 на всю левую часть борозды и P_1 (карман от правой части борозды под средний “мыс” — выступающую над бороздой часть белка). В P_0 с неплохим скором выделяется более мелкий подкарман P_0_0 вокруг лиганда. В свободном белке отсутствуют — вероятно в связи с подвижностью — участки петель, в частности, “мыса”, борозда шире, и карман под “мысом” отсутствует визуально и не выделяется программно, а выделенные карманы имеют значительно более низкий скор. Зато карману P_0 из связанной формы соответствует два отдельных кармана — P_0 вблизи места связывания лиганда и P_1 слева в борозде (визуально его обособленность от прочей борозды вызывает вопросы).

Если взять для сравнения объёмов (под)карман непосредственно вблизи лиганда, то его объём в свободной форме составляет 602 Å³, а в связанной всего 548 Å³, уменьшаясь таким образом на ~9% — вероятно, не столько за счёт “стягивания” соседних остатков лигандом, сколько за счёт стабилизации петли “мыса” (в свободной форме она может также находиться где-то вблизи и занимать примерно разницу объёмов, однако если её положение разное у разных молекул, мы её не разрешим и запишем весь соответствующий объём как свободный).

В целом разметка GSiteScorer выглядит сравнительно адекватной в тех случаях, когда визуальное определение карманов достаточно однозначно.

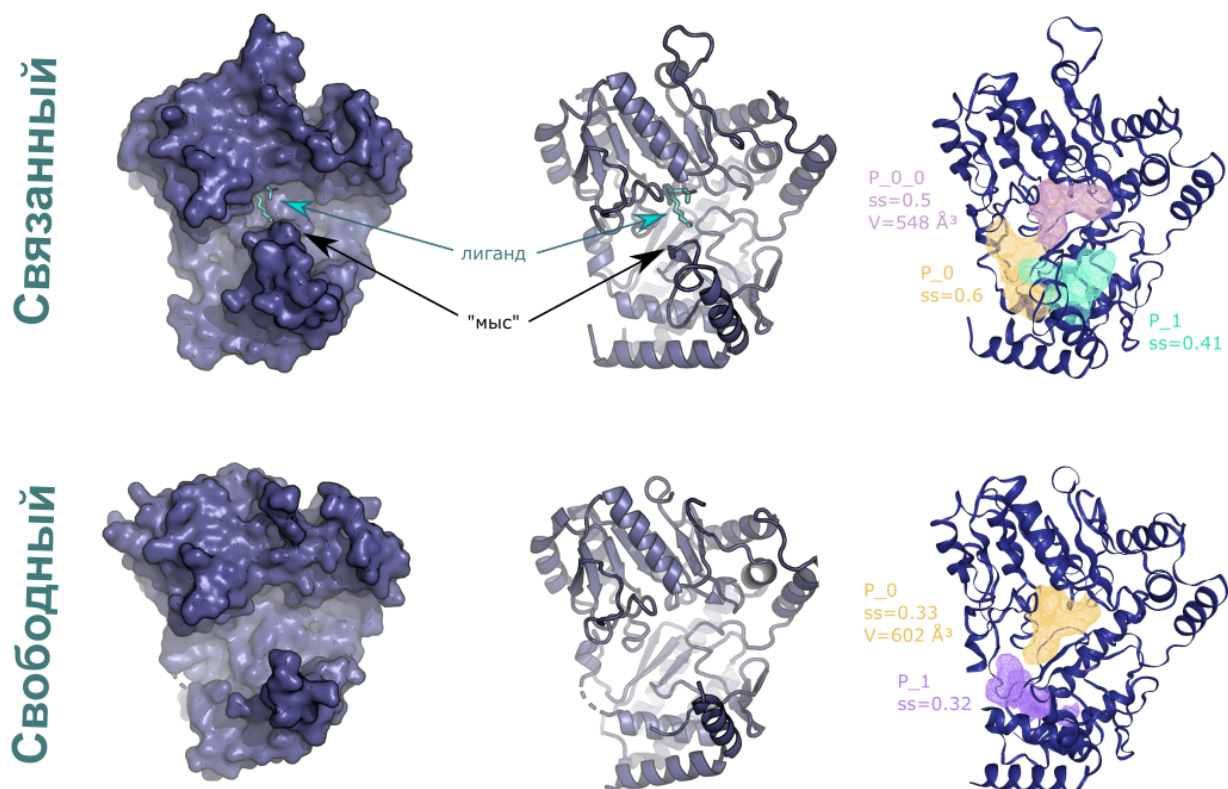


Рис. 1. Карманы на поверхности белка. Слева — поверхность, центр — остов (лиганд показан линиями и окрашен в бирюзовый), справа — остов с нанесёнными карманами с simple score (ss) > 0.3, для кармана связывания лиганда указан объём, цвета карманов на разных картинках никак не соотносятся друг с другом.

2. Докинг для имитации ранних стадий связывания лиганда

Для удобства описания я буду периодически апеллировать к нумерации атомов лиганда, приведённой на рисунке 2.

В сайте связывания лиганд удерживается различными взаимодействиями (рис. 3, слева): ацетильная группа образует солевой мостик и пару водородных связей с Arg-54 (расположенном на кончике "мыса") своей карбоксильной частью, а также водородную связь кислородом OX с Lys-280. Азот пиридинового кольца образует (в предположении его протонированности) солевой мостик с Asp-253 (и, возможно, водородную связь, хотя положение немного кривовато).

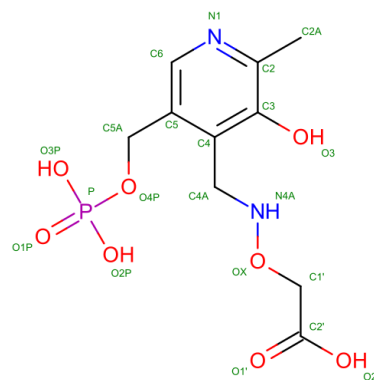


Рис. 2. Лиганд — 4'-дезоксид-4'-ацетамидо-пиридоксаль-5'-фосфат

Фосфатная группа удерживается водородными связями с остовом Gly-145 и остовом и гидроксилом Thr-146, гидрофобная часть пиррольного кольца и радикалов находится в окружении неполярных Val-255 и Tyr-172. В данном случае PoseView вполне успешно справляется с разметкой взаимодействий.

Для моделирования первых этапов связывания лиганда — до индуцированного соответствия — был произведён докинг при помощи DiffDock в свободную структуру. Докинг успешно нашёл (rank1, confidence 0.61) нужный сайт и даже сходную позу, однако многие контакты были изменены или утрачены (рис. 3, центр).

Ацетильная группа не может взаимодействовать с Arg-54, поскольку тот расположен в подвижной петле “мыса” и отсутствует в свободной структуре, Lys-280 присутствует, но развёрнут в противоположную сторону — таким образом, ацетильную группу ничего не стабилизирует. Однако компенсаторно в стабилизации молекулы может участвовать Met-256, расположенный подходящим образом для водородной связи с ОЗ лиганда. Интересно, что в оригинальной структуре этот метионин при небольшом повороте также мог бы образовать такую водородную связь, причём в текущем положении вроде бы ни с чем не взаимодействует, а при потенциальном повороте к лиганду ни с чем бы не мешался — но почему-то этого не делает. Азот пиррольного кольца образует такой же потенциальный солевой мостик с Asp-253, но направлен в середину карбоксигруппы, что не подходит для водородной связи (хотя PoseView зачем-то наоборот рисует к обоим кислородам). Фосфат удерживается водородными связями только с остовом и гидроксилом Thr-146, Gly-145 не рисуется PoseView — и действительно за счёт смещения фосфата больше не подходит для водородной связи. А вот Val-255 и Tyr-172, казалось бы, так же расположены около гидрофобной части лиганда (тирозин перестаёт быть так перпендикулярен, но, поскольку он и так был достаточно смещён в сторону от пиррольного кольца, их взаимодействие и так скорее описывалось как просто гидрофобное, а не Т-стекинг).

Таким образом, докинг находит практически правильную форму, отличающуюся контактами в области связывания ацильной группы и немного положением фосфата. Чтобы определить, насколько последнее объясняется неточностью докинга, а не различием положения остатков в свободной и связанной формах, рассмотрим контрольный опыт — докинг в исходную связанную структуру с удалённым лигандом (рисунок 3, справа).

Положение, полученное докингом (rank1, confidence 0.35), немного (на ~0.5-1.3 Å) сдвинуто относительно исходного, и PoseView теряет многие контакты. Тем не менее, даже при таком положении неясно, почему не размечены контакты с Asp-253, Thr-146 и Val-255, визуальнo они выглядят возможными даже при таком сдвиге, так что это, вероятно, неточность PoseView — а вот неточное положение атомов для водородных связей в области фосфата в докинге в свободную структуру (рис. 3, центр) вполне можно объяснить неточностью докинга, а вовсе не слегка отличающимся положением Gly-145 и Thr-146 в свободной структуре.

К некоторому разочарованию, DiffDock не справляется воспроизвести в контроле правильное положение ацетильной группы — вместо правильного Arg-54 с солевым мостиком и водородными связями, направляет её куда-то вбок, где она казалось бы не образует никаких контактов.

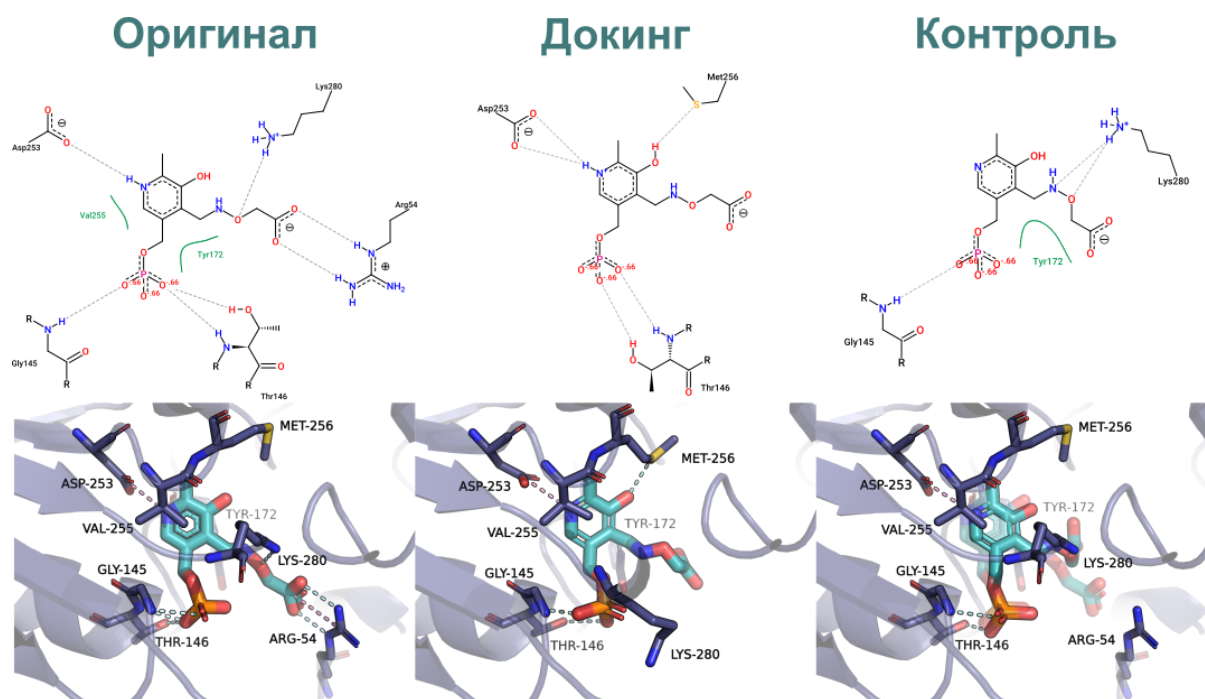


Рис. 3. Взаимодействия в центре связывания. Сверху — результат PoseView, снизу — вручную размеченные в структуре взаимодействия, водородные связи показаны белым, солевые мостики — розовым; оригинальное положение лиганда на картинке с контролем показано полупрозрачно.

Суммируя описанное выше, можно предположить следующую схему связывания лиганда с данным белком: сначала практически без изменений в собственной конформации первичное взаимодействие обеспечивают Asp-253 с азотом пиррола (солевой мостик?), гидрофобные Val-255 и Tyr-172, а также Gly-145 и/или Thr-146 (вероятно, лишь какая-то одна из их водородных связей с фосфатом). Возможно также временное участие Met-256. Далее происходит как сравнительно мелкая подгонка фосфата и Gly-145 и Thr-146 для образования всех трёх водородных связей, поворот радикала Lys-280 в правильную сторону и, наконец, приближение и стабилизация петли “мыса” с образованием прочного контакта ацетильной группы с Arg-54. Возможно, при этом по какой-то (неочевидной стерической?) причине теряется взаимодействие с Met-256.

Подгонку положения Gly-145, Thr-146 и Lys-280 при очень большом желании можно назвать индуцированным соответствием, но в целом, нельзя сказать, что для данного комплекса оно характерно. Движение “мыса” с Arg-54 практически наверняка происходит уже после связывания как его следствие.

3. Деанонимизация и последствия

Поскольку за предыдущие практикумы я привыкла для общего развития читать, с каким объектом работаю, я деанонимизировала структуры с использованием сервера Dali, выяснив, что комплексу соответствует структура 2ukv (цепь C), а свободной форме — структура 6iz9 (цепь A).

Белок представляет собой бактериальную аминотрансферазу (бета-трансаминазу).

Опыт показал, что поиском по структурному сходству (structure similarity) в PDB можно определить принадлежность к аминотрансферазам, но не найти оригинальную структуру (её нет даже среди первых 100 находок при сортировке по score).

Важно, что этот белок, согласно литературным данным, существует в растворе в виде гомодимера как свободный, так и связанный [1,2] — что объясняет такую своеобразную форму поверхности — как раз по поперечной “бороздке” и верхней “выемке” происходит контакт двух субъединиц (рисунок 4). Димер визуально имеет уже относительно ровную поверхность с выраженным крупным и глубоким карманом под лиганд (по одному на субъединицу), выглядящим шире в случае свободного белка и уже в случае связанного.

Тем не менее, GSiteScorer делит этот оформленный крупный карман на более мелкие, по-разному в случае свободной и связанной формы (рис. 4, слева в центре и слева снизу), но в обоих случаях выделяя подкарман вокруг лиганда (P_2_0 в связанном, P_0_0 в свободном). И если сравнить их объёмы, то получится результат, обратный результату на мономере — при связывании объём кармана не уменьшается, а напротив увеличивается на 16%, с 464 Å³ до 537 Å³ (притом что погрешность определения объёма кармана между расчётом по мономеру и по димеру в случае связанной формы, где он более чётко определён — как визуально, так и по скорам — составляет всего порядка 10 Å³, т.е. такой погрешностью разницу в 70 Å³ не объяснить).

Кажется единственный вывод, который тут можно сделать — в свободной форме карманы автоматически выделяются недостаточно чётко и уверенно, чтобы можно было сравнивать их объёмы со связанной формой.

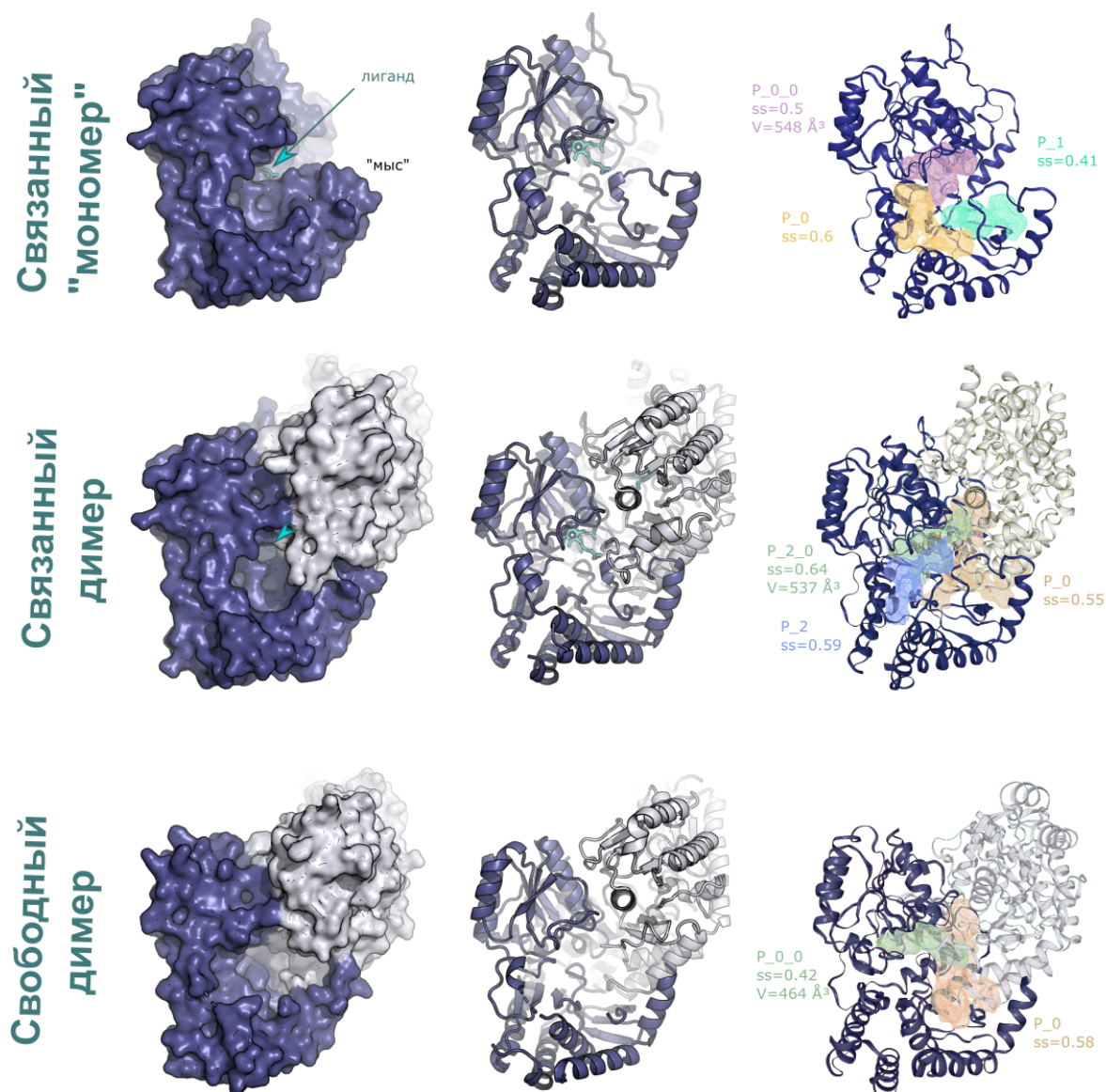


Рис. 4. Карманы на поверхности димера. Слева — поверхность, центр — остов (лиганд показан линиями и окрашен в бирюзовый), справа — остов с нанесёнными карманами с *simple score* (*ss*) > 0.3, для кармана связывания лиганда указан объём, цвета карманов на разных картинках никак не соотносятся друг с другом. Верхний ряд повторяет верхний ряд рисунка 1 в новом ракурсе для удобства соотнесения, в среднем и нижнем ряду белым/светло-серым показана вторая цепь димера. Голубая стрелка указывает положение лиганда — в случае димера его всё ещё можно различить с поверхности, пусть и с трудом.

Более того, петля второй субъединицы (S — в случае свободного белка уже присутствует в структуре, в случае связанного восстановлена из C по кристаллической симметрии) контактирует (рисунок 4, центр) с лигандом первой (C) — в частности, как раз в области той самой неправильно поворачивающейся при докинге карбоксигруппы, с которой вторая цепь образует водородные связи за счёт остова Gly-313 и гидроксигруппы Thr-314 (рисунок 5 слева). Может быть, неспособность DiffDock правильно воспроизвести исходное состояние лиганда связано с тем, что мы подаём ему искусственно мономерное состояние с неполным набором взаимодействующих аминокислот?

Для проверки этой гипотезы я повторила контрольный докинг с “естественной” димерной структурой. Как можно видеть на рисунке 5 слева, добавление остатков из второй цепи в сайт связывания не спасло — задоченный лиганд всё также немного сдвинут относительно оригинального положения, а ацильная группа всё также сгибается вместо выпрямления для оптимального контакта с Arg-54, только сгибается теперь в другую сторону (что позволяет образовать водородную связь с Gly-313 второй цепи) и часть контактов с Arg-54 “боком”.

Похоже, у DiffDock просто почему-то проблема с “тройным” контактом (солевой мостик + две водородные связи) аргинина с карбоксигруппой.

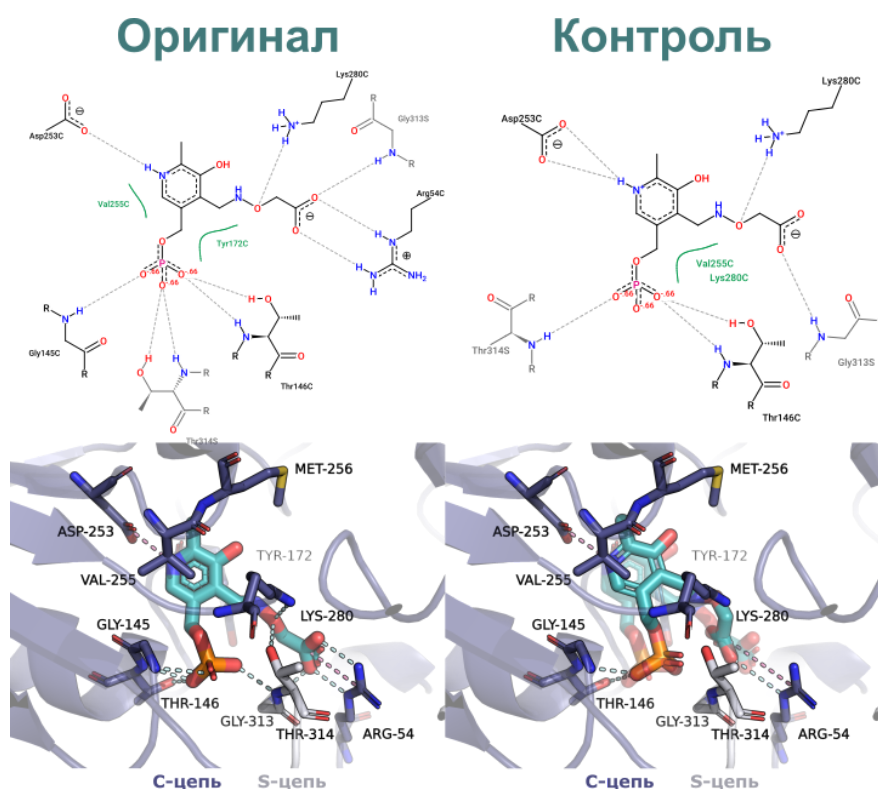


Рис. 5. Взаимодействия в центре связывания с учётом второй цепи димера (показана серым). Сверху — результат PoseView, снизу — вручную размеченные в структуре взаимодействия, водородные связи показаны белым, солевые мостики — розовым; оригинальное положение лиганда на картинке с контролем показано полупрозрачно.

Выводы

- PoseView хорошо размечает контакты в кристаллической структуре, но начинает странно их терять (при визуальном некритичном изменении) на структурах из докинга
- GSiteScorer нормально выделяет карманы, когда simple score > 0.5, с меньшими (в данном случае — у свободной формы) есть вопросы
- DiffDock вроде бы дочит неплохо, но иногда странно не справляется
- Учёт димерности вместо того, чтобы всё спасти и объяснить, только запутал

Список литературы

- 1 Kwon, S. and Park, H. H. (2019) Crystal structure of the apo form of a β -transaminase from *Mesorhizobium* sp. strain LUK. *Protein Science* 28, 964–970.
- 2 Wybenga, G. G., Crismaru, C. G., Janssen, D. B. and Dijkstra, B. W. (2012) Structural Determinants of the β -Selectivity of a Bacterial Aminotransferase *. *Journal of Biological Chemistry, Elsevier* 287, 28495–28502.