Conformational gating in ammonia lyases

Matteo Lambrughini¹, Željka Sanader Maršić¹, Veronica Saez-Jimenez^{1,2}, Valeria Mapelli², Lisbeth Olsson², Elena Papaleo^{1,3*}

¹ Computational Biology Laboratory, Danish Cancer Society Research Center, Copenhagen, Denmark

² Division of Industrial Biotechnology, Department of Biology and Biological Engineering, Chalmers University of Technology, Göteborg, Sweden

³ Translational Disease Systems Biology, Faculty of Health and Medical Sciences, Novo Nordisk Foundation Center for Protein Research University of Copenhagen, Copenhagen, Denmark

Corresponding author: Elena Papaleo, elenap@cancer.dk

ABSTRACT

Ammonia lyases (AL) are enzymes of industrial and biomedical interest. Knowledge of AL structure-dynamics-function relationship would be instrumental for making use of the application potential of these enzymes. We investigated, using microsecond molecular dynamics, the conformational changes in the proximity of the catalytic pocket of a 3-methylaspartate ammonia lyase (MAL) as a model system. In particular, we identified two regulatory elements in the MAL structure, i.e., the $\beta 5$ - $\alpha 2$ loop, and the helix-hairpin-loop subdomain. We showed that they undergo conformational changes switching from 'occluded' to 'open' states. We observed that these rearrangements are coupled to changes in the accessibility of the active site. The $\beta 5$ - $\alpha 2$ loop and the helix-hairpin-loop subdomain modulate the formation of tunnels from the protein surface to the substrate binding site, making the active site more accessible to the substrate when they are in an open state. We pinpointed a sequential mechanism, in which the helix-hairpin-loop subdomain needs to break a subset of intramolecular interactions first, to then allow the opening of the $\beta 5$ - $\alpha 2$ loop and, as a consequence, make the AL catalytic pocket accessible for the substrate.

1 Our data suggest that protein dynamics need to be considered in the design of new
2 AL variants for protein engineering and therapeutic purposes.

3 **Keywords**

4 molecular dynamics simulations, conformational changes, disordered regions,
5 conformational switches, correlated motions

7 **Introduction**

8 Ammonia lyases (ALs) are a broad class of enzymes which catalyze different
9 transformations based on α - and β -amino acid scaffolds, such as the deamination
10 and isomerization of natural amino acids through the reversible cleavage or the
11 shifting of a C-N bond. ALs are highly heterogeneous in their structures and
12 mechanisms of action, as attested by the fact that they cover 31 diverse EC sub-
13 classes and they are characterized by a high stereoselectivity for their substrates [1].
14 These enzymes are commercially appealing for their industrial [2,3] and biomedical
15 applications [1,2]. For example, they have been suggested as potential cancer
16 biotherapeutics [1] due to the fact ALs could prevent the supply of the tumor cells
17 with essential metabolites. One example is the histidine ammonia lyase which
18 influences the growth of ovarian and prostate cancer cells by histidine deamination
19 [1]. The deamination reaction produces urocanic acid and ammonia, eliminating
20 histidine as building block and contributes to preventing protein synthesis, which is
21 essential for the growth of cancer cells. The catalytic mechanism of ALs is well [3–9].
22 Among them, 3-methylaspartate ammonia lyases, or methylaspartases (MALs) [3]
23 catalyze the reversible deamination of 3-methylaspartate to mesaconate. MALs
24 belong to the enolase superfamily and they catalyze a broader range of reactions
25 than other ammonia lyases [1] They are potentially versatile and a promising target

1 to design new variants with different substrate specificity or improved activity and
2 stability. For example, the MAL variant isolated from *Citrobacter amalonaticus*
3 (CaMAL) has been engineered for enantioselective synthesis of N-substituted
4 aspartic acids, which are essential building blocks for pharmaceutical, artificial
5 sweeteners, synthetic enzymes and peptidomimetics [10,11].

6 MAL is a dimeric enzyme. Only a few crystallographic three-dimensional (3D)
7 structures of the dimeric form of MALs are available, deposited in the Protein Data
8 Bank (PDB), as the entries 1KKO, 1KKR [12], 1KD0, 1KCZ [13], along with two
9 mutated variants with PDB entries 3ZVH, 3ZVI [11]. CaMAL is a homodimeric
10 enzyme where each monomer (here referred to as A and B) is composed by 826
11 amino acids and can be divided into two domains: an N-terminal (residues 1–160)
12 and a C-terminal domain (residues 170–413) connected by β -sheet regions (**Figures**
13 **1A-B**). The C-terminal domain folds into a triosephosphate-isomerase (TIM) barrel
14 structure, consisting of eight α -helices and eight parallel β -strands, while the N-
15 terminal domain is composed of a three antiparallel β -strands and four α -helices
16 (**Figure 1B**).

17 The catalytic mechanism of CaMAL has been inferred from the analysis of its 3D
18 structure, the comparison with other members of the enolase family and validated by
19 experimental mutagenesis [14,15]. The mutagenesis studies revealed the
20 importance of specific residues on the structural integrity, activity, and regio- and
21 diastereoselectivity of CaMAL (**Figure 1C**). The reaction catalyzed by CaMAL
22 requires the presence of cations, generally provided by the metal magnesium (Mg^{2+})
23 atom in the catalytic pocket, while the residues K331, H194 and Q329 form the
24 catalytic triad (**Figure 1C**). K331 acts as the base catalyst, while the Mg^{2+} metal ion
25 and the residues H194 and Q329 are responsible for the stabilization of the enolate

1 anion, and binding to the 4-carboxylate group of the substrate [14]. Additionally, Q73,
2 F170, Q172, Y356, T360, C361 and L384 interact with different functional groups of
3 the substrate and are located in the pocket (**Figure 1C**).

4 The N-terminal domain of CaMAL bears, in the proximity of the catalytic pocket, two
5 structural elements, that we refer as $\beta 5$ - $\alpha 2$ loop (residues 70-85) and helix-hairpin-
6 loop subdomain (hhl, residues 12-51) (**Figure 1C**). The $\beta 5$ - $\alpha 2$ loop connects the
7 strand $\beta 5$ with the helix $\alpha 2$ and includes Q73, which form water-mediated hydrogen
8 bonds with the substrate and alters CaMAL k_{cat} when mutated [15]. The hhl
9 subdomain contains a short α -helix (residues 19-25), a small β -hairpin ($\beta 2$ - $\beta 3$
10 strands, residues 29-34) and the loop between the strands $\beta 3$ and $\beta 4$. The hhl
11 subdomain is involved in intermolecular interactions at the interface between the
12 monomers [12].

13 At the best of our knowledge, there is no information available on the conformational
14 changes of MALs that could be related to its function. It is a common feature of
15 several enzymes that disordered or partially structured regions of a folded protein
16 undergo conformational changes that modulate the access of the substrate to the
17 active site or allow for its formation [8,16,25,17–24]. A detailed investigation of the
18 dynamics of MALs is an essential step to understand how to engineer it for
19 applicative purposes.

20 We focused on CaMAL, presenting the first all-atom Molecular Dynamics (MD)
21 investigation of this family of enzyme and its implications for activity. Biomolecular
22 simulations are useful to describe protein functional dynamics, from local to more
23 global motions and the notions of the importance of dynamics for enzyme function is
24 well established [26–31]. This includes local changes in the proximity of the active or
25 substrate binding sites [25,32–37] up to conformational changes of large amplitude

1 associated with the opening and closing of gating loops or domains that can
2 modulate the access to the catalytic site or the substrate binding pocket of an
3 enzyme [16,22,38].

4

5 **Results**

6

7 *CaMAL is a dimer in solution*

8 To select which form of CaMAL to study with MD simulations, we investigated the
9 preferred quaternary structures of CaMAL in solution. In particular, we performed
10 Size-Exclusion Chromatography (SEC) to estimate CaMAL molecular weight (MW).
11 We used proteins with a known MW as calibration standards (see Materials and
12 Methods). The resulting partition coefficients (K_{av}) plotted against the logarithm of
13 MW were fitted with a linear function (correlation coefficient ≈ 0.99) (**Figure 2A**). We
14 identified one single sharp peak for CaMAL, corresponding to a MW of ≈ 80.4 kDa
15 (**Figure 2A**), which is approximately double of the MW estimated for the monomeric
16 form of CaMAL (from its primary sequence, i.e., 45.5 kDa). We concluded that
17 CaMAL is mostly a homodimer when in solution in its substrate-unbound state, in
18 agreement with what previously reported [39].

19 In light of the SEC results, we focused on the CaMAL dimer in solution for our MD
20 simulations. We also collected two additional simulations of the monomers (i.e.,
21 chains A and B from the dimer structure) as a control.

22

23 *CaMAL shows asymmetric dynamic patterns with high flexibility located at the $\beta 5$ - $\alpha 2$*
24 *loop and the hhl subdomain*

1 We employed all-atom explicit solvent one- μ s MD simulations to investigate CaMAL
 2 dynamics-structure-function relationship. In particular, we aimed at identifying
 3 conformational changes that modulate the accessibility of the active site. We
 4 calculated the average per-residue Root Mean Square Fluctuation (RMSF) of the C α
 5 atoms for each monomer of the CaMAL dimer as a measure of protein flexibility. The
 6 peaks in the RMSF profile indicate the protein regions interested by the highest
 7 flexibility (**Figure 2B**). We observed marked fluctuations in the β 5- α 2 loop and the
 8 hhl subdomains in both the monomeric and dimeric states of CaMAL. We found
 9 similar fluctuations also in a MD simulation performed starting from another
 10 crystallographic structure of CaMAL dimer (PDB entry 1KK0), supporting our findings
 11 (see Github repository for more details).

12 The simulations performed on each separated monomer show higher flexibility in the
 13 hhl subdomain in the case of the monomer A than in the dimer, while in the
 14 monomer B this region is more rigid (**Figure S1**). These differences suggest an
 15 asymmetry in the structure and dynamics of the two monomeric subunits of CaMAL,
 16 a trait common to other dimeric enzymes [40–43] (see GitHub repository for more
 17 details). This asymmetry is observed for both the hhl subdomain and the β 5- α 2 loop,
 18 indicating that the dynamic patterns of these regions, in the dimer, are modulated by
 19 the presence of the partner monomer. The RMSF estimated by our MD simulations
 20 is also in agreement with the crystallographic B-factor values associated with the
 21 experimental structures of the CaMAL dimer in its substrate-bound (PDB entry
 22 1KKR) or -unbound (PDB entry 1KKO) states (**Figure 2B**).

23

24 *The β 5- α 2 loop and the hhl subdomain populates ‘open’ and ‘closed’ states in the μ s*
 25 *timescale*

1 Dimensionality reduction techniques help to identify the most important
2 conformational transitions in a MD ensemble and compare simulations [44–47]. In
3 particular, we employed Principal Component Analysis (PCA) [48] and focused on
4 the ‘essential subspace’ described by the first two principal components (PCs)
5 (**Figure 2C-D and Figure S2**). We identified two main conformational states,
6 associated with conformational transitions of the $\beta 5$ - $\alpha 2$ loop and the hhl subdomain
7 (**Figure 2C**). The two regions can populate ‘open’ or ‘occluded/closed’ conformations
8 with respect to the active site (**Figure 2C**) with a progressive opening during the
9 simulation time (**Figure 2D**, color gradient).

10 We observed a different dynamic pattern for the isolated monomer A with a large
11 exploration of the PCA subspace during the simulation time (**Figure 2D**, orange
12 gradient), which is related to its enhanced flexibility (**Figure S1**). On the contrary,
13 CaMAL remains in its closed state in the MD simulations of the isolated monomer B,
14 mostly sampling conformations around the starting structure (**Figure 2D**, light blue
15 gradient), in agreement with the lower RMSF observed for this trajectory (**Figure**
16 **S1**).

17 Moreover, we extended the second replicate of the CaMAL dimer to two μ s since in
18 this replicate the opening of the catalytic pocket was occurring after hundreds of ns
19 (**Figure 2D**, blue and dark orange gradients).

20 We then analyzed more in detail the structural changes, which were associated with
21 the conformational transitions of in the $\beta 5$ - $\alpha 2$ loop and the hhl subdomain (**Figure**
22 **2E**). The opening of the $\beta 5$ - $\alpha 2$ loop were associated with instability of the β -hairpin
23 (residues 29-34), loss of β structure in the $\beta 2$ and $\beta 3$ strands (approximately in 75%
24 of the simulation frames) and changes to bend/coil conformations (**Figure 2E**). The
25 hhl subdomain was stable during the conformational transition, with the residues 19-

1 25 in a stable helical conformation in all the simulations (approximately in more than
2 70% of the simulation frames) (**Figure 2E**).

3

4 *The hhl subdomain and the β 5- α 2 loop modulate the accessibility of the catalytic*
5 *pocket of CaMAL with a sequential mechanism*

6 The conformational transitions observed during CaMAL simulations involved
7 significant rearrangements of the β 5- α 2 loop and hhl subdomain in comparison to
8 their position in the starting structure (**Figure 3A**). To understand the relationship
9 between these rearrangements and changes in other regions of CaMAL, we
10 monitored, over the simulation time, representative pairwise distances between Ca
11 atoms of each of the residues of the β 5- α 2 loop or of the hhl subdomain and the Ca
12 Q329 residue, used as a reference of the pocket (**Figure 3B**). We then estimated the
13 pairwise Pearson correlations of the selected distances with the solvent accessibility
14 of the catalytic pocket of CaMAL, designing a 'correlogram' where only correlations
15 with a p-value lower than 0.01 are highlighted (**Figure 3C**). We calculated the
16 solvent accessible surface area (SASA) of each of the residues of the pocket (i.e.,
17 Q73, F170, Q172, H194, Q329, K331, Y356, T360, C361 and L384), along with their
18 cumulative SASA and the total protein SASA for each monomer. The correlogram
19 turned out to be an efficient approach to investigate possible couplings between
20 changes in the residues of the β 5- α 2 loop or the hhl subdomain and the accessibility
21 of the catalytic pocket of the enzyme. We observed statistically significant positive
22 correlations between the total SASA of the pocket and the distances between Q329
23 and many of the residues of the β 5- α 2 loop and the hhl subdomain (**Figure 3C**).
24 Moreover, we applied a hierarchical cluster approach on the correlation matrices and
25 we identified two main groups of structural measurements with higher positive

1 correlations between them (**Figure 3C**). The first group includes the distances with
 2 the residues of the $\beta 5$ - $\alpha 2$ loop and the residues 12-14, 16, 29-33 and 48-51 of the
 3 hhl subdomains that are correlated with both the total SASA of the pocket and the
 4 accessibility of K331 and T360. These results suggest that the opening of the $\beta 5$ - $\alpha 2$
 5 loop modulates the accessibility of the binding pocket directly, acting of the residues
 6 that are located at its entrance, such as K331 and T360 (**Figure 3B**). The second
 7 group includes the distances with the residues 17-28 and 34-47 of the hhl
 8 subdomains that are correlated with the total protein SASA and the accessibility of
 9 Q73. Q73 is located in the $\beta 5$ - $\alpha 2$ loop, and our analysis suggests that the hhl
 10 subdomains play a role in acting as a conformational gate for the $\beta 5$ - $\alpha 2$ loop (see
 11 below **Figure 4D**).

12 In addition, we observed coupling between several residues of the $\beta 5$ - $\alpha 2$ loop and
 13 the hhl subdomain (**Figure 3C**). Since these two structural elements are localized
 14 spatially close to each other in the CaMAL structure and they form intramolecular
 15 contacts, we estimated if their motions are significantly coupled at the level of
 16 residue-residue displacements. In particular, we calculated the residue-residue
 17 correlated motions along the trajectories using Linear Mutual Information (LMI, [49]).
 18 We calculated the average LMI correlation matrices over time windows of 10 and
 19 100 ns for the MD simulations of CaMAL dimer to assess two different timescales
 20 (**Figure 3D**). We calculated the pairwise differences between the LMI matrices
 21 calculated for each of the two monomers in the same replicate, along with the
 22 differences between the LMI matrices for the monomers A of different replicates
 23 (**Figure 3D** and **Figure S3**). We concluded that the LMI matrices described similar
 24 patterns of correlated fluctuations in the monomers, despite their asymmetry (see
 25 above) with differences in the range of ± 0.3 . We also observed modest differences

1 among LMI matrices averaged over different time windows. Once assessed the
2 robustness of the LMI matrices, we analyze them to identify local couplings in the
3 MD ensemble (**Figure 3D**). In particular, we observed correlations in the motions of
4 the $\beta 5$ - $\alpha 2$ loop and hhl subdomain (**Figure 3D**). Moreover, we identified coupled
5 motions between these two structural elements and most of the N-terminal domain of
6 CaMAL (i.e., residues 1-163), suggesting that the entire N-terminal domain
7 undergoes conformational changes relatively to the C-terminal domain.

8 In summary, we observed couplings between the motions of the $\beta 5$ - $\alpha 2$ loop and the
9 hhl subdomain, along with couplings of these two regions with the accessibility of the
10 catalytic site. We also observed that the opening of the hhl subdomain occur before
11 the opening of the $\beta 5$ - $\alpha 2$ loop in each replicate. Our results thus suggest that the hhl
12 subdomain play a role in constraining the motions of the $\beta 5$ - $\alpha 2$ loop, acting as a
13 conformational gate.

14

15 *The opening of the $\beta 5$ - $\alpha 2$ loop and hhl subdomain shorten the tunnels from the*
16 *protein surface to the active site of CaMAL*

17 We estimate the formation and time-evolution of tunnels towards the catalytic site
18 (**Figure 4**). We used the position of four residues of the active site (i.e., Q329, D238,
19 L384, Y356) as a reference point for the tunnel analysis. We calculated the evolution
20 of the solvent-accessible tunnels from the protein surface to these residues with a
21 probe radius of 1.5 Å, which resembled the size of a water molecule. The analysis
22 identified from eight to ten clusters of tunnels depending on the trajectory (**Figure**
23 **4A**). We calculated the time evolution of the average length and radius of the tunnels
24 in each cluster, which indicate the absence of bottlenecks when a tunnel was
25 identified (**Figure 4B-C**). We observed that the length of the tunnels decreases with

1 the opening of the hhl subdomain and the $\beta 5$ - $\alpha 2$ loop, showing that binding pocket
 2 became more easily accessible (**Figure 4B-C**). We also performed the same
 3 analysis with a probe radius of 2.0 Å to better resemble the size of the natural
 4 substrate (i.e., 3-methylaspartate), obtaining similar results (see Github repository).
 5 In summary, our results indicate that the opening of the hhl subdomain occur first
 6 (**Figure 4D**) and allow the $\beta 5$ - $\alpha 2$ loop outward displacement. Moreover, their
 7 concerted opening provides an efficient access to the active site, shortening the
 8 protein tunnels from the protein surface to the catalytic residues.

9

10 *Complex networks of contacts modulate the conformational changes from 'closed' to* 11 *'open' states of CaMAL*

12 To dissect the mechanism of opening of the substrate binding pocket through the
 13 conformational changes of the hhl subdomain and the $\beta 5$ - $\alpha 2$ loop, we also used an
 14 approach based on contact maps [50]. We measured pairwise intramolecular
 15 contacts between residues, their persistency along the trajectory time and estimated
 16 the evens of breaking and formation of the interactions. We observed that the
 17 residues of the $\beta 5$ - $\alpha 2$ loop and the hhl subdomain in the open states are generally
 18 involved in a lower number of contacts at a lower persistence then in the closed
 19 states (see Github repository).

20 A cluster analysis based on the RMSD between the contact maps along the
 21 trajectory identified five sequential clusters (**Figure 5A**). The clusters allowed to
 22 study the temporal evolution of the contacts along the trajectories during the
 23 transition from closed to open states of the $\beta 5$ - $\alpha 2$ loop and the hhl subdomain. We
 24 estimated the percentage increase in the pairwise distances between the residues in
 25 the $\beta 5$ - $\alpha 2$ loop and the hhl subdomain with the ones in the rest of the protein,

1 comparing the average contact maps of the first and fifth cluster. We identified
2 several pairs of residues for which we observed a marked increase in the distances
3 along the simulation time. For these residues, we measured the distances with
4 respect to residues of the $\beta 5$ - $\alpha 2$ loop (**Figure 5B**) and the hhl subdomain (**Figure**
5 **5C**), as distance networks.

6 The distance network analysis for the $\beta 5$ - $\alpha 2$ loop (**Figure 5B**) showed a significant
7 increase of distances with residues at both side of the entrance of the pocket: L196,
8 N198, Y240, M389, G359-S367 (containing two catalytic residues, i.e., T360 and
9 C361) on one side and, D307-T312 and K331-D334 (containing the catalytic residue
10 K331), on the other side. Moreover, we observed significant changes in the
11 distances with amino acids at the bottom of the loop, such as H88-I90, L93, and
12 H127-R131, along with several residues in the hhl subdomain or its proximity, as
13 S13-Y16, D18-A22, K24, T40-T44 and E51-S54 (**Figure 5B**).

14 The distance network for the hhl subdomain (**Figure 5C**) showed an increase in the
15 distances with residues around the entrance of the pocket, such as for G174, L196,
16 N198, N199, Y240, D277, E308, C361-E365 (containing the catalytic residue C361),
17 S367, A368, and K386-G390. Furthermore, significant increases in the distance
18 network are associated with the $\beta 5$ - $\alpha 2$ loop, such as the residues D69-G78, R80,
19 L83, L85, and A86.

20 The mapping of the residues interested by changes in the distance network during
21 the opening of the hhl subdomain and the $\beta 5$ - $\alpha 2$ loop (**Figure 5D**) showed that the
22 hhl subdomain is involved in modulating a large area of the protein, while the $\beta 5$ - $\alpha 2$
23 loop is mostly associated with the regions around the active site. In summary, the
24 contact-based distance network analysis allowed to identify the most significant

1 molecular changes associated with the opening of the hhl subdomain and the $\beta 5$ - $\alpha 2$
2 loop.

3 **Discussion**

4 An essential step toward the application of enzymes to industrial processes, protein
5 engineering and biomedical treatments, is to understand their structural properties,
6 dynamics and unveil the link with enzyme function and activity. ALs have been
7 proposed as therapeutic agents against cancer, but despite their potential, low
8 stability and rapid clearance hindered their applications [1]. The development of new
9 modified variants of ALs will open up for broad applications, but it requires more
10 detailed knowledge on their structural properties and functional mechanisms. We
11 contributed to fill this gap using computational structural biology. Our study
12 highlighted two important regulatory elements for CaMAL, i.e., the $\beta 5$ - $\alpha 2$ loop and
13 hhl subdomain. They are highly flexible and undergo conformational changes on the
14 μ s timescale, switching from closed to open states with respect to the catalytic site.
15 The hhl subdomain and the $\beta 5$ - $\alpha 2$ loop modulate the accessibility of the catalytic
16 pocket of the enzyme, shortening the tunnels to access the active site, in a
17 sequential mechanism where hhl subdomain acts as a conformational gate and
18 needs to open first to promote the detachment of the $\beta 5$ - $\alpha 2$ loop from the residues at
19 the entrance of the pocket. This mechanism might have broader relevance for ALs,
20 as attested by the fact that other similar modulators have been reported, for example
21 in aspartate, phenylalanine and histidine ammonia lyases, [4,8,18] These ALs do
22 not share similarities in sequence and are structurally unrelated but, the presence of
23 loops acting as opening/closing ‘clamp’ above the active site pinpoint interesting
24 similarities with the molecular mechanisms that we identified in CaMAL (**Figure 6**).

1 Our results also provide a framework for the future design of mutant variants of
2 CaMAL to entrap the enzyme in its closed state to experimentally validate the effect
3 of the opening/closing transitions on the enzymatic activity of CaMAL. For example,
4 one strategy to pursue could be the engineering of disulphide bridges or mutations
5 that stabilize the interaction networks here identified. A suitable area for mutation is,
6 for example, the region above the catalytic pocket in the proximity of Y240, or the
7 residues G359-S367, which are localized at the side-entrance of the pocket. Our
8 study also opens new directions to understand, at a deeper level, how the substrate
9 recognition and binding happen in CaMAL and how this is regulated by events of
10 conformational changes of the 'gate' elements promoted by the substrate.

11

12 ***Materials and Methods***

13 *CaMAL expression and purification*

14 We retrieved the gene sequence encoding for MAL from *Citrobacter amalonaticus*
15 (CaMAL) (NCBI accession numbers AB005294) from the National Center for
16 Biotechnology Information database and codon optimized for expression in
17 *Escherichia coli* [51]. CaMAL was synthesized by GenScript (Piscataway, NJ, USA)
18 and cloned in the vector pET28b so that a HisTag was introduced in the C-terminal
19 region of the gene. The pET-28b plasmid, which contains the CaMAL gene, was
20 transformed into *Escherichia coli* BL21 (DE3) for protein expression. The cells
21 containing the plasmid were grown in LB media supplemented with 50 µg/ml of
22 kanamycin and autoinduction media with lactose [52] at 30°C at 180 rpm for 16 h.
23 We solubilized the recollected cells in 50 mM Tris-HCl (pH 8) with 300 mM NaCl and
24 10 mM imidazole and stored at -20°C. The thawed cells were sonicated (Branson

1 Digital Sonifier, model 250) using a maximum amplitude of 30%, in 7 cycles of 30 s
2 and centrifuged for 20 min at 13500 rpms. We purified the supernatants with the
3 expressed protein by immobilized metal ion affinity chromatography (IMAC), using a
4 1-ml HisTrap column (GE Healthcare, Uppsala, Sweden) and 50 mM phosphate pH
5 8, 300 mM NaCl, 20mM imidazole buffer. A gradient up to 500 mM imidazole was
6 used to elute the purified protein.

7

8 *Size exclusion chromatography experiments*

9 We estimate the molecular weight (MW) of CaMAL by SEC using a Superdex 200
10 HiLoad 16/600 column (GE Healthcare, Uppsala, Sweden), 50 mM sodium
11 phosphate pH 8.0 with 150 mM sodium chloride as the mobile phase, a flow rate of 1
12 ml min⁻¹, and UV (280 nm) detection. We used blue dextran to determine the column
13 void volume (V_o), which was 43.4 ml. We selected four proteins as molecular mass
14 standards from the gel filtration calibration kit HMW (GE Healthcare, Uppsala,
15 Sweden), i.e. ovalbumin (MW ~ 44 000), conalbumin (MW ~ 75 000), aldolase (MW
16 ~ 158 000) and ferritin (MW ~ 440 000), with elution volumes of 79.8 mL, 73.4 mL,
17 65.6 mL and 53.4 mL, respectively. The calibration curve was constructed as the
18 partition coefficient (K_{av}) versus Log (MW). We calculated K_{av} using the equation K_{av}
19 = (V_e - V_o) / (V_t - V_o), where V_t is the total bed volume.

20

21 *Preparation of the initial structures for MD simulations*

22 We used the X-ray structure of the homodimer of CaMAL in complex with its natural
23 substrate, i.e. the 3-methyl-aspartic acid (PDB entry 1KKR, [12]) and in the unbound
24 state (PDB entry 1KK0, [12]) as starting structures for the MD simulations. We
25 removed the coordinates of the substrate from the X-ray structure of the bound state

1 of CaMAL. We used Modeller version 9.16 [53] to replace the seleno-methionines in
 2 the PDB entries with methionines and we modeled the missing residues at the N-
 3 and C-terminal of the protein. In the re-modeling step, we restrained all the protein
 4 atoms except for the two residues immediately next to each selenomethionine in the
 5 amino acid sequence. We generated 200 different models and we selected the one
 6 with the lowest Root Mean Square Deviation (RMSD) calculated on all the
 7 methionine atoms using as a reference their original orientation in the X-ray
 8 structure. We retained three crystallographic water molecules, which have been
 9 suggested to be important for the interaction between the enzyme and its substrate
 10 [12].

11

12 *MD simulations*

13 We carried out all-atom MD simulations of CaMAL in explicit solvent with *Gromacs*
 14 version 5.1.2 [54]. We collected two one- μ s MD simulations of CaMAL in its
 15 monomeric form, i.e., considering the chain A and chain B separately. Moreover, we
 16 run three replicates of the substrate-unbound form CaMAL dimer from the 1KKR
 17 structure, using different random seed for the simulations and one replicate starting
 18 from the 1KK0 structure to assess the reproducibility of our findings. We employed
 19 the CHARMM22 force field with CMAP corrections (namely CHARMM27) [55] and
 20 the TIPS3P water model [56]. We predicted the pKa of the histidine residues with
 21 PROPKA 3.1 [57] and analyzed the pattern of possible hydrogen bonds for the
 22 histidine imidazole nitrogens. We modelled H351 with the hydrogen on the δ nitrogen
 23 of the ring, while all the other histidines were modelled with hydrogens on the ϵ
 24 nitrogen.

1 Dimer simulations were done using a dodecahedral box of water with a minimum
2 distance between protein and box edges of 15 Å applying periodic boundary
3 conditions, while box edges of 30 Å were used for monomers. We used a larger box
4 for the simulations of the monomers to avoid that possible large rearrangements in
5 their structures, which could be caused by the removal of the partner monomer,
6 created artifacts associated with periodic boundary conditions. We gradually
7 equilibrated the system through a series of energy minimization, solvent
8 equilibration, thermalization and pressurization steps using a protocol previously
9 applied to other cases of study [58,59]. In each simulation, we used a 9 Å switch
10 cutoffs for Van der Waals and short-range Coulomb interactions. We employed the
11 Particle-mesh Ewald switch summation method for long-range electrostatic
12 interactions [60]. We stored the conformations every 20 ps, and we used the LINCS
13 algorithm to constrain the heavy atom bond lengths [61]. We collected two one-μs
14 replicates and the third replicate was extended till two μs (see Result and Discussion
15 for more details).

16 We calculate the minimum distance between the atoms of the protein and its periodic
17 images to identify potential periodic boundary artefacts. In all the monomer
18 simulations the minimal distance was always higher than 33 Å and in the dimer 39 Å.

19

20 *Principal Component Analysis*

21 We used a dimensionality reduction technique based on Principal Component
22 Analysis (PCA) [44,45,48] to compare the space sampled by the different MD
23 simulations. The PCA was carried out on a concatenated trajectory of the MD
24 simulations of both monomeric and dimeric states of CaMAL (isolating each
25 monomer from the dimer first) to compare them in the same essential subspace

1 [36,47]. We use the the atoms belonging to the secondary structure of the TIM barrel
 2 (i.e., residues 177-187, 209-225, 235-238, 251-265, 270-273, 281-297, 303-306,
 3 313-322, 327-330, 338-351, 355-357, 365-378, and 382-384) for the fitting. We used
 4 these residues as a reference for structural superimposition since they are located in
 5 the core regions of the TIM barrel domain and do not undergo significant changes in
 6 the simulations. We performed the PCA calculations for different subsets of atoms of
 7 CaMAL: i) all the protein residues, ii) the region 70-85 (i.e., the $\beta 5$ - $\alpha 2$ loop), and iii)
 8 the region 12-51 (i.e., the hhl subdomain) (**Figure 2** and **Figure S2**). We focused on
 9 the first two principal components, which account for approximately 43% of the
 10 variance in the whole protein scan, and approximately 60% in the $\beta 5$ - $\alpha 2$ loop and hhl
 11 subdomain PCA analysis.

12

13 *Protein flexibility*

14 We calculated the per-residue C α -atoms Root Mean Square Fluctuation (RMSF) as
 15 an index of flexibility. RMSF was calculated over non-overlapping time windows of
 16 10 ns along the trajectories and then averaged, with a protocol used for other
 17 proteins [62–64].

18

19 *Changes in secondary structures*

20 We used the DSSP dictionary [65] to estimate the different classes of secondary
 21 structures attained for each residue of the $\beta 5$ - $\alpha 2$ loop and hhl subdomain. DSSP
 22 defines eight classes of secondary structure elements: 3_{10} helix, α -helix, π -helix, β -
 23 sheet, β -bridge, helix turn, bend and coil. For each residue of interest a persistence
 24 degree of secondary structure was calculated for each class [66]. The DSSP

1 algorithm is based on calculating the patterns of hydrogen bonds between mainchain
2 carbonyl and amide groups.

3

4 *Pairwise distances between selected residues*

5 We calculated the time series of distances between the C α atom of a reference
6 residue in the binding pocket, i.e. Q329, and the C α atoms of all the residues of the
7 β 5- α 2 loop and hhl subdomain.

8

9 *Solvent accessible surface area (SASA)*

10 We estimated the SASA for all the atoms of each residues of the catalytic pocket
11 which was either involved in the catalysis or in the interaction with the substrate (i.e.,
12 Q73, F170, Q172, H194, Q329, K331, Y356, T360, C361 and L384). We also
13 calculated their cumulative SASA and the total protein SASA per monomer.

14

15 *Correlogram*

16 We collected correlogram plots for each trajectory to investigate the coupling
17 between the pairwise selected distances and the different SASA measurements. We
18 used an in-house R script based on the *PerformanceAnalytics*, *corrplot* and *Hmisc* R
19 packages and the *rcorr()* R function to calculate dependencies between all the pairs
20 of variables. We measured linear dependencies using the Pearson correlation
21 coefficient as a metric for correlation estimate and we calculated the p-value
22 associated to each correlation. We derived the correlograms using the *corrplot()* R
23 function and a hierarchical clustering approach to rank the correlations depending on
24 the degree of couplings between each pair of measurements. We calculated for each
25 measurement (distances and accessibilities) a bivariate scatter plot with the fitted

1 line and the correlations coefficients with the significance levels at p-values of 0,
2 0.001, 0.01, 0.05, 0.1, 1. We performed the correlogram analysis for each monomer
3 A and B, separately as derived from the replicates of the CaMAL dimer.

4

5 *Correlated motions*

6 We filtered the trajectories on the C α -atoms and fitted on the C α -atoms of the same
7 subset of residues used as a reference for the PCA calculation, i.e., belonging to the
8 secondary structure of the TIM barrel (see above). We calculated C α LMI correlation
9 matrices as previously described [67,68] using the *Wordom* package version 0.22
10 [69]. We then converted the *Wordom* outputs to the *xPyder* [70] compatible format
11 using the *wordomdccm2dat* tool. LMI values range from 0 (i.e., uncorrelated
12 motions) to 1 (i.e., fully correlated motions). We defined a cutoff of 0.5 to identify
13 significant correlations and remove spurious correlations, as previously applied to
14 other cases of study [58,67]. We also employed a cutoff of distance in the sequence
15 of +/- three residues to filter out the correlations between residues that are next to
16 each other in the primary sequence. We used two non-overlapping time windows of
17 10 and 100 ns, respectively to estimate average LMI correlation matrices. We
18 calculated the pairwise differences between the LMI matrices calculated for each of
19 the three different dimer replicates, to identify regions in the structure with
20 similar/dissimilar correlations. We compared the LMI matrices calculated for each
21 replicate of the dimer on the monomer A and the monomer B. Moreover, we
22 evaluated the differences between the LMI matrices of the monomer A and the
23 monomer A of different replicate (**Figure S3**). We plotted LMI matrices as heatmaps
24 using an in-house R script based on the ggplot2, reshape2 and viridis R functions.

25

Tunnel analyses

We used the CAVER Analyst 1.0 [71] to determine the optimal parameters for the calculations and identification of tunnels along the MD trajectories of CaMAL. We then used the standalone version of CAVER 3.02 software [72] to identify tunnels, which could mediate the access to the active site of CaMAL during the MD simulations. We used as a reference point for the tunnel search the position of the catalytic residues of CaMAL. In particular, we used: i) the nitrogen atom in the amide group of the side chain of Q329; ii) the oxygen atom of the carboxylic group of the side chain of D238; iii) the carbon atom in the methyl group of the side chain of L384 and, iv) the oxygen atom in the hydroxyl group of the side chain of Y356. We performed the tunnel calculations on 1000 representative frames along each trajectory. We removed water molecules and the Mg^{2+} atoms before the calculation. We carried out the tunnel analysis on the MD replicates of the CaMAL dimer, analyzing separately the monomers A and B. We tested different parameters to perform tunnel search and clustering with CAVER, following an optimization procedure previously published [72]. In particular, we selected two set of parameters. The first set aimed to monitor the evolution of solvent-accessible tunnels from the protein surface towards the catalytic site, where we used a probe radius of 1.5 Å, comparable to the size of a water molecule [73], a shell radius of 5 Å, shell depth of 4 Å, and a clustering threshold of 7 Å. In the second set of parameters, we used the same parameters as the first set but with a probe radius of 2.0 Å to resemble the size of the natural substrate 3-methylaspartate. For this purpose, we used as a reference the radius of gyration of 3-methylaspartate (i.e., 2.23 Å). We used for the calculation the conformation of the 3-methylaspartate in the bound crystallographic structure of CaMAL (PDB entry 1KKR [12]).

1 The tunnels were clustered using the hierarchical average-link clustering method,
2 which permit to estimate the dissimilarity (i.e., the distance) of two measured tunnels,
3 dividing each tunnel pathway by sequence of N points starting from the reference
4 site and then evaluating their pairwise distance in their corresponding order. We
5 used PyMOL for visual inspection and graphical representation of the tunnels.

6

7 *Contact-based analyses*

8 For the contact-based analyses, we used the CONtact ANalysis (CONAN) software
9 [50]. CONAN permits to obtain a statistical analysis of intramolecular contacts in
10 proteins and to investigate how they evolve along the MD trajectories. We used as
11 cutoffs a r_{cut} value of 10 Å, and r_{inter} and $r_{\text{high-inter}}$ values of 5 Å, as previously used
12 [50]. We performed a contact-based cluster analysis using a k-medoid clustering
13 approach based on the RMSD between the contact maps calculated for each frame
14 along each trajectory. The number of maximum clusters was set to five. We
15 observed that CONAN identifies five clusters that are sequentially distributed over
16 the simulation time. We used this analysis to study the temporal evolution of the
17 contacts along the trajectories. We calculated the average distance maps for each
18 cluster, accounting for each pair of residues over all the frames assigned to the
19 same cluster. We analyzed the average distance maps using in-house R scripts. We
20 calculated the percentage increase (or decrease) in the pairwise distances between
21 residues comparing the average contact maps of the first and fifth cluster. We
22 retained as significant only the pairs of residues with a percentage increase in their
23 distance higher than 50%. We analyzed the data using network analysis in
24 *Cytoscape*. The movie of the time-evolution of the contact maps are reported in the
25 Github repository associated to the publication.

1

2 *Data and Software Availability*

3 All the software used are freely available. The R and bash scripts, input files for the
4 simulations, input and output files from the modeling step, along with the results
5 generated will be soon freely available in a GitHub repository associated to our
6 publication (https://github.com/ELELAB/MAL_MD). The entire MD trajectories are
7 available upon request.

8

9 **Acknowledgements**

10 The study was funded by The Swedish Research Council for Environment,
11 Agricultural Sciences and Spatial Planning (FORMAS) (project number 942-2015-
12 1628), the NovoNordisk Foundation under the program for Biotechnology-based
13 Synthesis and Production Research (reference number NNF-17OC0027588) to EP
14 and LO groups and a European Biophysical Society Association (EBSA) bursary
15 granted to VSJ to visit EP group in 2017. EP lab is part of the Center of Excellence in
16 Autophagy, Recycling and Disease (CARD), funded by the Danish National
17 Research Foundation (DNRF125) and also supported by The Danish Council for
18 Independent Research, Natural Science Research Project 1 (102517). The
19 calculations described in this paper were performed using resources provided by the
20 Swedish National Infrastructure for Computing (SNIC) at PDC Center for High
21 Performance Computing under the project SNIC 2017/1-537 and the DeiC National
22 Life Science Supercomputer at DTU.

23

24 **Author Contributions**

1 E.P. and M.L. designed the research; E.P. supervised the work; V.S-J. performed
2 the experiments; M.L, Z.S.M. and V.S-J. performed the modeling and simulations;
3 M.L. and Z.S.M. performed the analyses of the simulations; E.P, M.L, Z.S., V.S-J.,
4 V.M., L.O. discussed the results; M.L. and Z.S. provided scripts and code for the
5 study; M.L, Z.S. and E.P. wrote the manuscript with suggestions from all the
6 coauthors.

7

8 **Figure Legends**

9

10 **Figure 1. The 3-methylaspartase ammonia lyase from *Citrobacter amalonaticus***
11 **(CaMAL) has two structural elements in the proximity of the catalytic pocket,**
12 **i.e., the helix-hairpin-loop (hhl) and the $\beta 5$ - $\alpha 2$ loop.** A) The homodimeric structure
13 of CaMAL (PDB entry 1KKR) is shown as white surface and the monomer B is
14 shown as cartoon for sake of clarity. The hhl subdomain and the the $\beta 5$ - $\alpha 2$ loop are
15 in gold and brown, respectively. B) The monomer A of CaMAL dimer is shown as
16 cartoon and represented with a gradient of color from N-terminal (brown) to C-
17 terminal (purple) with indicated the secondary structural elements in the proximity of
18 the hhl subdomain ($\alpha 1$, $\beta 1$ - $\beta 4$) and the $\beta 5$ - $\alpha 2$ loop ($\alpha 2$ and $\beta 5$). C) The residues in
19 the catalytic pocket are shown as sticks with the following color coding: Q73 red,
20 F170 orange, Q172 green, H194 forest green, Q329 aquamarine, K331 dark cyan,
21 Y356 cornflower blue, T360 navy blue, C361 purple and L384 violet red. The hhl
22 subdomain and the $\beta 5$ - $\alpha 2$ loop are colored in gold and brown, respectively.

23

24 **Figure 2. CaMAL is a dimer in solution with two highly flexible regions.** A) Size
25 exclusion chromatography experiment shows single absorption peak at the

1 molecular mass of 80.4kDa which corresponds to the dimeric form of CaMAL. B) The
2 B-factor (in the left) and RMSF (in the right) values are shown as tube representation
3 on the structure of CaMAL dimer. The thickness and the color spectrum of the tube
4 are proportional to the flexibility of the residues, with thicker red tubes representing
5 highly flexible regions and thinner blue tubes indicating more rigid regions of the
6 enzyme. The highest B-factor and RMSF values can be observed in the hhl
7 subdomain and the $\beta 5$ - $\alpha 2$ loop. C) 2D distribution plot calculated using as reaction
8 coordinates the first and the second principal components from the PCA analysis
9 performed on all the protein atoms. We identified two densely populated subsets of
10 conformations (indicated as 'open' and 'closed' states). One representative structure
11 from each state is shown as cartoon in the lower part of the panel for illustrative
12 purposes. The 'closed' structural ensemble included conformations of the CaMAL
13 monomers similar to the initial X-ray structure. The 'open' ensemble accounted for
14 conformations of the CaMAL monomers with both the $\beta 5$ - $\alpha 2$ loop and the hhl
15 subdomain in open states. D) We highlighted the PCA subspace sampled by each
16 trajectory. For the MD simulations of the dimer we showed the monomer A and B
17 separately for sake of clarity. Simulations of the monomers and dimers are indicted
18 as 'mono' and 'dim', respectively and the MD replicates of the dimeric form are
19 numbered sequentially from one to three. The time evolution of each trajectory in the
20 conformational space is represented as shade of color, from dark (0 ns) to light (end
21 of the trajectory). We observed the transition from open to closed states in the dimer
22 simulations. The differences observed for the simulations of the monomer forms are
23 discussed in the text. E) We estimated the occurrence of different secondary
24 structural classes in the regions of the hhl subdomain and $\beta 5$ - $\alpha 2$ loop during the
25 simulations. The color bars indicate the structural classes in the X-ray structure, as a

1 reference. For sake of clarity only data obtained for the first replicate of CaMAL
2 dimer are reported. The other results are in the Github repository associated to the
3 publication.

4

5 **Figure 3. The opening of the hhl subdomain and the $\beta 5$ - $\alpha 2$ loop is correlated**
6 **with changes in the accessibility of the catalytic pocket of CaMAL.** The residues
7 of the catalytic pocket are shown with the following color coding: Q73 red, F170
8 orange, Q172 green, H194 forest green, Q329 aquamarine, K331 dark cyan, Y356
9 cornflower blue, T360 navy blue, C361 purple and L384 violet red. The hhl
10 subdomain and the $\beta 5$ - $\alpha 2$ loop are showed in gold and brown, respectively. A)
11 Surface representation of two representative structures belonging to the ‘closed’ (in
12 the left) and ‘open’ (in the right) states of the enzyme. The open conformations
13 promote a higher accessibility of the residues in the catalytic pocket. B) We showed
14 the distances between the C α atom of Q329 in the pocket and the C α atoms of all
15 the residues in the $\beta 5$ - $\alpha 2$ loop and in the hhl subdomain that we used in the
16 analyses. The CaMAL is shown as white cartoon and the residues in the binding
17 pocket, for which the solvent accessibility surface area (SASA) was calculated, are
18 shown as sticks. C) The correlogram shows the Pearson correlation coefficients
19 between each pair of measurements (i.e., the selected distances and SASAs) at a p-
20 value < 0.01. The color intensity and the size of the circle are proportional to the
21 correlation coefficients, i.e. the strength of the correlations, between each pair of
22 measurements, from dark red (negative correlations) to dark blue (positive
23 correlations). The correlogram shows a clear coupling between changes in the the
24 $\beta 5$ - $\alpha 2$ loop and in the hhl subdomain and accessibility of specific residues of the
25 catalytic binding pocket. The details are discussed in the text. D) The heatmap

1 shows the average correlated motions in the MD trajectories, using Linear Mutual
2 Information (LMI). The LMI correlations are shown as a gradient from 0 (i.e.,
3 uncorrelated motions shown in yellow) to 1 (i.e., fully correlated motions in dark
4 blue). The LMI correlation matrix shows a coupling in the motions of the $\beta 5$ - $\alpha 2$ loop
5 and the hhl subdomain, which are also coupled to the entire N-terminal domain. For
6 sake of clarity, we reported the results for the monomer A of the first MD replicate of
7 the CaMAL dimeric form. The other results are in the Github repository associated to
8 the publication.

9

10 **Figure 4. The opening of the $\beta 5$ - $\alpha 2$ loop and hhl subdomain modulate the**
11 **length and the apperance of tunnels toward the active site of CaMAL.** A) We
12 showed the main clusters of tunnel pathways calculated from the MD ensembles on
13 the structure of the CaMAL monomer. In particular, the tunnels calculated for each
14 conformation are shown as pathway centerlines and grouped into clusters. For each
15 conformation, if more than one tunnel pathway is grouped in the same cluster, only
16 the pathway with the lowest cost is shown. B-C) We showed, as heatmaps, the
17 temporal evolution of the two main cluster of tunnels in blue and green, respectively.
18 The plots report the average length and the radius (colors in the plot) of the tunnels
19 in the cluster for each trajectory frame. For sake of clarity, we illustrated only the
20 results for two clusters for the monomer A of the first MD replicate of the CaMAL
21 dimer form. The other results are in the Github repository associated to the
22 publication. D) We showed the temporal evolution of three distances between the Ca
23 atoms of Q329 (in the pocket) and S75 (in the $\beta 5$ - $\alpha 2$ loop, in brown), P38 and A49
24 (in the hhl subdomain, in gold and dark gold, respectively). The origin of the y-axis is
25 set to the lowest value observed for the three selected distances (10.77 Å). We

1 showed only the results for the monomer A of the first MD replicate of the CaMAL
2 dimer form for sake of clarity. The other results are in the Github repository
3 associated to the publication.

4

5 **Figure 5. Rearrangements of intramolecular contacts during the opening of the**
6 **hhl subdomains and the $\beta 5$ - $\alpha 2$ loop.** A) We performed a cluster analysis based on
7 the RMSD between the contact maps collected at different time points during the
8 simulation, setting the maximum number of clusters as five. The plot shows the time
9 evolution of the identified five clusters with dots colored according to the distance
10 (measured in RMSD) of each conformation from the cluster medoid, which is
11 indicated with a blue dot. B-C) Circular plots showing the pairwise distance network
12 for the $\beta 5$ - $\alpha 2$ loop (B) and the hhl subdomains (C). We estimated the percentage
13 increase of the pairwise distances between the residues of the $\beta 5$ - $\alpha 2$ loop and the
14 hhl subdomain with the residues of the rest of the protein, comparing the average
15 contact maps of the first and fifth cluster. In the graphs, the edges indicate
16 percentage increase above 50% in the network and connect pairs of residues, which
17 represent the nodes. The values of percentage increase are colored as a gradient
18 from red (50%) to purple (300%). The node corresponding to each residue are
19 shown as points and they are colored on the base of their degree in the network,
20 from low (yellow) to high (dark blue). D) We mapped the network of residues on a
21 representative open conformation of the CaMAL dimer as green and blue sticks for
22 the $\beta 5$ - $\alpha 2$ loop and the hhl subdomain, respectively. The hhl subdomain and the $\beta 5$ -
23 $\alpha 2$ loop are highlighted in gold and brown, respectively and both the monomers are
24 shown. For sake of clarity, we illustrated the results for the monomer A of the first

1 replicate of the CaMAL dimeric form. The other results are in the Github repository
2 associated to the publication.

3

4 **Figure 6. Other ammonia lyases show conformationally heterogeneous**
5 **structural elements in the proximity of their catalytic sites, which could**
6 **undergo opening/closing transitions.** We showed the structures of: i) CaMAL with
7 the hhl subdomain and the $\beta 5$ - $\alpha 2$ loop highlighted in gold and brown, respectively; ii)
8 aspartate (AAL, PDB entry 3R6V), phenylalanine (PAL, 1W27), and histidine (HAL,
9 1GKM) ammonia lyases with their disordered regions colored in red in the proximity
10 of the catalytic site, which is indicated by a blue star.

11

12 **References**

- 13 [1] F. Parmeggiani, N.J. Weise, S.T. Ahmed, N.J. Turner, Synthetic and
14 Therapeutic Applications of Ammonia-lyases and Aminomutases, Chem. Rev.
15 (2017) acs.chemrev.6b00824. doi:10.1021/acs.chemrev.6b00824.
- 16 [2] M.M. Heberling, B. Wu, S. Bartsch, D.B. Janssen, Priming ammonia lyases
17 and aminomutases for industrial and therapeutic applications, Curr. Opin.
18 Chem. Biol. 17 (2013) 250–260. doi:10.1016/j.cbpa.2013.02.013.
- 19 [3] M. De Villiers, V.P. Veetil, H. Raj, J. De Villiers, G.J. Poelarends, Catalytic
20 mechanisms and biocatalytic applications of aspartate and methylaspartate
21 ammonia lyases, ACS Chem. Biol. 7 (2012) 1618–1628.
22 doi:10.1021/cb3002792.
- 23 [4] J. Zhang, Y. Liu, A QM/MM study of the catalytic mechanism of aspartate
24 ammonia lyase, J. Mol. Graph. Model. 51 (2014) 113–119.
25 doi:10.1016/j.jmglm.2014.05.003.

- 1 [5] A. Heine, G. Herrmann, T. Selmer, F. Terwesten, W. Buckel, K. Reuter, High
2 resolution crystal structure of *Clostridium propionicum* β -alanyl-CoA: AMmonia
3 lyase, a new member of the “hot dog fold” protein superfamily, *Proteins Struct.*
4 *Funct. Bioinforma.* 82 (2014) 2041–2053. doi:10.1002/prot.24557.
- 5 [6] K.T. Watts, B.N. Mijts, P.C. Lee, A.J. Manning, C. Schmidt-Dannert, Discovery
6 of a Substrate Selectivity Switch in Tyrosine Ammonia-Lyase, a Member of the
7 Aromatic Amino Acid Lyase Family, *Chem. Biol.* 13 (2006) 1317–1326.
8 doi:10.1016/j.chembiol.2006.10.008.
- 9 [7] N.J. Turner, Ammonia lyases and aminomutases as biocatalysts for the
10 synthesis of α -amino and β -amino acids, *Curr. Opin. Chem. Biol.* 15 (2011)
11 234–240. doi:10.1016/j.cbpa.2010.11.009.
- 12 [8] A.L. Seff, S. Pilbák, I. Silaghi-Dumitrescu, L. Poppe, Computational
13 investigation of the histidine ammonia-lyase reaction: A modified loop
14 conformation and the role of the zinc(II) ion, *J. Mol. Model.* 17 (2011) 1551–
15 1563. doi:10.1007/s00894-010-0849-7.
- 16 [9] Y. Asano, Y. Kato, C. Levy, P. Baker, D. Rice, Structure and Function of
17 Amino Acid Ammonia-lyases, *Biocatal. Biotransformation.* 22 (2004) 133–140.
18 doi:10.1080/10242420410001703496.
- 19 [10] W.J. Quax, H. Raj, P.G. Tepper, G.J. Poelarends, V. Puthan Veetil, M. de
20 Villiers, F.J. Dekker, Enantioselective Synthesis of N-Substituted Aspartic
21 Acids Using an Engineered Variant of Methylaspartate Ammonia Lyase,
22 *ChemCatChem.* 5 (2013) 1325–1327. doi:10.1002/cctc.201200906.
- 23 [11] H. Raj, W. Szymański, J. De Villiers, H.J. Rozeboom, V.P. Veetil, C.R. Reis, M.
24 De Villiers, F.J. Dekker, S. De Wildeman, W.J. Quax, A.M.W.H. Thunnissen,
25 B.L. Feringa, D.B. Janssen, G.J. Poelarends, Engineering methylaspartate

- 1 ammonia lyase for the asymmetric synthesis of unnatural amino acids, *Nat.*
- 2 *Chem.* 4 (2012) 478–484. doi:10.1038/nchem.1338.
- 3 [12] C.W. Levy, P.A. Buckley, S. Sedelnikova, Y. Kato, Y. Asano, D.W. Rice, P.J.
- 4 Baker, Insights into enzyme evolution revealed by the structure of
- 5 methyiaspartate ammonia lyase, *Structure.* 10 (2002) 105–113.
- 6 doi:10.1016/S0969-2126(01)00696-7.
- 7 [13] M. Asuncion, W. Blankenfeldt, J.N. Barlow, D. Gani, J.H. Naismith, The
- 8 structure of 3-methyiaspartase from *Clostridium tetanomorphum* functions via
- 9 the common enolase chemical step, *J. Biol. Chem.* 277 (2002) 8306–8311.
- 10 doi:10.1074/jbc.M111180200.
- 11 [14] H. Raj, B. Weiner, V.P. Veetil, C.R. Reis, W.J. Quax, D.B. Janssen, B.L.
- 12 Feringa, G.J. Peolarends, Alteration of the diastereoselectivity of 3-
- 13 methyiaspartate ammonia lyase by using structure-based mutagenesis,
- 14 *ChemBioChem.* 10 (2009) 2236–2245. doi:10.1002/cbic.200900311.
- 15 [15] H. Raj, G.J. Poelarends, The roles of active site residues in the catalytic
- 16 mechanism of methyiaspartate ammonia-lyase, *FEBS Open Bio.* 3 (2013)
- 17 285–290. doi:10.1016/j.fob.2013.07.002.
- 18 [16] E. Papaleo, V. Ranzani, F. Tripodi, A. Vitriolo, C. Cirulli, P. Fantucci, L.
- 19 Alberghina, M. Vanoni, L. De Gioia, P. Coccetti, An Acidic Loop and Cognate
- 20 Phosphorylation Sites Define a Molecular Switch That Modulates Ubiquitin
- 21 Charging Activity in Cdc34-Like Enzymes, *PLoS Comput. Biol.* 7 (2011).
- 22 doi:10.1371/journal.pcbi.1002056.
- 23 [17] E. Papaleo, N. Casiraghi, A. Arrigoni, M. Vanoni, P. Coccetti, L. De Gioia,
- 24 Loop 7 of E2 Enzymes: An Ancestral Conserved Functional Motif Involved in
- 25 the E2-Mediated Steps of the Ubiquitination Cascade, *PLoS One.* 7 (2012).

- 1 doi:10.1371/journal.pone.0040786.
- 2 [18] S. Pilbák, A. Tomin, J. Rétey, L. Poppe, The essential tyrosine-containing loop
- 3 conformation and the role of the C-terminal multi-helix region in eukaryotic
- 4 phenylalanine ammonia-lyases, FEBS J. 273 (2006) 1004–1019.
- 5 doi:10.1111/j.1742-4658.2006.05127.x.
- 6 [19] R.C. Wade, M.E. Davis, B.A. Luty, J.D. Madura, J.A. McCammon, Gating of
- 7 the active site of triose phosphate isomerase: Brownian dynamics simulations
- 8 of flexible peptide loops in the enzyme., Biophys. J. 64 (1993) 9–15.
- 9 doi:10.1016/S0006-3495(93)81335-3.
- 10 [20] H.-X. Zhou, S.T. Wlodek, J.A. McCammon, Conformation gating as a
- 11 mechanism for enzyme specificity, Proc. Natl. Acad. Sci. 95 (2002) 9280–
- 12 9283. doi:10.1073/pnas.95.16.9280.
- 13 [21] M. Kokkinidis, N.M. Glykos, V.E. Fadoulglou, Protein flexibility and enzymatic
- 14 catalysis, 2012. doi:10.1016/B978-0-12-398312-1.00007-X.
- 15 [22] E. Papaleo, G. Saladino, M. Lambrugh, K. Lindorff-Larsen, F.L. Gervasio, R.
- 16 Nussinov, The role of protein loops and linkers in conformational dynamics and
- 17 allostery., Chem. Rev. 116 (2016) 6391–6423.
- 18 doi:10.1021/acs.chemrev.5b00623.
- 19 [23] M.K. Rout, B.L. Lee, A. Lin, W. Xiao, L. Spyropoulos, Active Site Gate
- 20 Dynamics Modulate the Catalytic Activity of the Ubiquitination Enzyme E2-25K,
- 21 Sci. Rep. 8 (2018) 1–15. doi:10.1038/s41598-018-25476-8.
- 22 [24] A. Gora, J. Brezovsky, J. Damborsky, Gates of Enzymes, Chem. Rev. 113
- 23 (2013) 5871–5923. doi:10.1021/cr300384w.
- 24 [25] M. Pasi, L. Riccardi, P. Fantucci, L. De Gioia, E. Papaleo, Dynamic properties
- 25 of a psychrophilic alpha-amylase in comparison with a mesophilic homologue.,

- 1 J. Phys. Chem. B. 113 (2009) 13585–95. doi:10.1021/jp900790n.
- 2 [26] D. Tobi, I. Bahar, Structural changes involved in protein binding correlate with
- 3 intrinsic motions of proteins in the unbound state., Proc. Natl. Acad. Sci. U. S.
- 4 A. 102 (2005) 18908–13. doi:10.1073/pnas.0507603102.
- 5 [27] Z. Kurkcuoglu, A. Bakan, D. Kocaman, I. Bahar, P. Doruker, Coupling between
- 6 Catalytic Loop Motions and Enzyme Global Dynamics, PLoS Comput. Biol. 8
- 7 (2012) e1002705. doi:10.1371/journal.pcbi.1002705.
- 8 [28] A. Tousignant, J.N. Pelletier, Protein motions promote catalysis, Chem. Biol.
- 9 11 (2004) 1037–1042. doi:10.1016/j.
- 10 [29] J.L. Klepeis, K. Lindorff-Larsen, R.O. Dror, D.E. Shaw, Long-timescale
- 11 molecular dynamics simulations of protein structure and function., Curr. Opin.
- 12 Struct. Biol. 19 (2009) 120–7. doi:10.1016/j.sbi.2009.03.004.
- 13 [30] R.O. Dror, R.M. Dirks, J.P. Grossman, H. Xu, D.E. Shaw, Biomolecular
- 14 simulation: a computational microscope for molecular biology., Annu. Rev.
- 15 Biophys. 41 (2012) 429–52. doi:10.1146/annurev-biophys-042910-155245.
- 16 [31] T. Saleh, C.G. Kalodimos, Enzymes at work are enzymes in motion, Science
- 17 (80-.). 355 (2017) 247–248. doi:10.1126/science.aal4632.
- 18 [32] E. Papaleo, L. Sutto, F.L. Gervasio, K. Lindorff-Larsen, Conformational
- 19 changes and free energies in a proline isomerase, J. Chem. Theory Comput.
- 20 10 (2014) 4169–4174. doi:10.1021/ct500536r.
- 21 [33] R.N. Shinde, S. Karthikeyan, B. Singh, Molecular dynamics studies unravel
- 22 role of conserved residues responsible for movement of ions into active site of
- 23 DHBPS, Sci. Rep. 7 (2017) 1–10. doi:10.1038/srep40452.
- 24 [34] P. Mereghetti, L. Riccardi, B.O. Brandsdal, P. Fantucci, L. De Gioia, E.
- 25 Papaleo, Near native-state conformational landscape of psychrophilic and

- 1 mesophilic enzymes: probing the folding funnel model, J. Phys. Chem. B. 114
- 2 (2010) 7609–7619. doi:10.1021/jp911523h.
- 3 [35] F. del Caño-Ochoa, A. Grande-García, M. Reverte-López, M. D’Abramo, S.
- 4 Ramón-Maiques, Characterization of the catalytic flexible loop in the
- 5 dihydroorotase domain of the human multi-enzymatic protein CAD, J. Biol.
- 6 Chem. 293 (2018) 18903–18913. doi:10.1074/jbc.RA118.005494.
- 7 [36] I. Valimberti, M. Tiberti, M. Lambrugh, B. Sarcevic, E. Papaleo, E2 superfamily
- 8 of ubiquitin-conjugating enzymes: constitutively active or activated through
- 9 phosphorylation in the catalytic cleft, Sci. Rep. 5 (2015) 14849.
- 10 doi:10.1038/srep14849.
- 11 [37] E. Papaleo, L. Riccardi, C. Villa, P. Fantucci, L. De Gioia, Flexibility and
- 12 enzymatic cold-adaptation: A comparative molecular dynamics investigation of
- 13 the elastase family, Biochim. Biophys. Acta - Proteins Proteomics. 1764 (2006)
- 14 1397–1406. doi:10.1016/j.bbapap.2006.06.005.
- 15 [38] E. Papaleo, G. Renzetti, Coupled motions during dynamics reveal a tunnel
- 16 toward the active site regulated by the N-terminal α -helix in an acylaminoacyl
- 17 peptidase, J. Mol. Graph. Model. 38 (2012) 226–234.
- 18 doi:10.1016/j.jmgm.2012.06.014.
- 19 [39] Y. Kato, Y. Asano, Purification and properties of crystalline 3-methylaspartase
- 20 from two facultative anaerobes, Citrobacter sp. strain YG-0504 and Morganella
- 21 morganii strain YG-0601., Biosci. Biotechnol. Biochem. 59 (1995) 93–9.
- 22 <http://www.ncbi.nlm.nih.gov/pubmed/7765982> (accessed March 15, 2019).
- 23 [40] E. Papaleo, G. Renzetti, G. Invernizzi, B. Ásgeirsson, Dynamics fingerprint and
- 24 inherent asymmetric flexibility of a cold-adapted homodimeric enzyme. A case
- 25 study of the Vibrio alkaline phosphatase, Biochim. Biophys. Acta - Gen. Subj.

- 1 1830 (2013) 2970–2980. doi:10.1016/j.bbagen.2012.12.011.
- 2 [41] E. Moroni, D.A. Agard, G. Colombo, The Structural Asymmetry of
- 3 Mitochondrial Hsp90 (Trap1) Determines Fine Tuning of Functional Dynamics,
- 4 J. Chem. Theory Comput. 14 (2018) 1033–1044.
- 5 doi:10.1021/acs.jctc.7b00766.
- 6 [42] J.M. Flynn, P. Mishra, D.N. a. Bolon, Mechanistic Asymmetry in Hsp90 Dimers,
- 7 J. Mol. Biol. (2015). doi:10.1016/j.jmb.2015.03.017.
- 8 [43] T.H. Kim, P. Mehrabi, Z. Ren, A. Sljoka, C. Ing, A. Bezginov, L. Ye, R. Pomès,
- 9 R.S. Prosser, E.F. Pai, The role of dimer asymmetry and protomer dynamics in
- 10 enzyme catalysis, Science (80-.). 355 (2017). doi:10.1126/science.aag2355.
- 11 [44] E. Papaleo, P. Mereghetti, P. Fantucci, R. Grandori, L. De Gioia, Free-energy
- 12 landscape, principal component analysis, and structural clustering to identify
- 13 representative conformations from molecular dynamics simulations: The
- 14 myoglobin case, J. Mol. Graph. Model. 27 (2009) 889–899.
- 15 doi:10.1016/j.jmglm.2009.01.006.
- 16 [45] I. Daidone, A. Amadei, Essential dynamics: foundation and applications, Wiley
- 17 Interdiscip. Rev. Comput. Mol. Sci. 2 (2012) 762–770. doi:10.1002/wcms.1099.
- 18 [46] M. Tiberti, E. Papaleo, T. Bengtsen, W. Boomsma, K. Lindorff-Larsen,
- 19 ENCORE: software for quantitative ensemble comparison, PLoS Comput. Biol.
- 20 11 (2015) e1004415.
- 21 [47] F. Martín-García, E. Papaleo, P. Gomez-Puertas, W. Boomsma, K. Lindorff-
- 22 Larsen, Comparing molecular dynamics force fields in the essential subspace.,
- 23 PLoS One. 10 (2015) e0121114. doi:10.1371/journal.pone.0121114.
- 24 [48] A. Amadei, A.B. Linssen, H.J. Berendsen, Essential dynamics of proteins.,
- 25 Proteins. 17 (1993) 412–25. doi:10.1002/prot.340170408.

- 1 [49] O.F. Lange, H. Grubmüller, Generalized correlation for biomolecular dynamics,
2 Proteins Struct. Funct. Bioinforma. 62 (2005) 1053–1061.
3 doi:10.1002/prot.20784.
- 4 [50] D. Mercadante, F. Gräter, C. Daday, CONAN: A Tool to Decode Dynamical
5 Information from Molecular Interaction Maps, Biophys. J. 114 (2018) 1267–
6 1273. doi:10.1016/j.bpj.2018.01.033.
- 7 [51] P. Puigbò, E. Guzmán, A. Romeu, S. Garcia-Vallvé, OPTIMIZER: a web
8 server for optimizing the codon usage of DNA sequences., Nucleic Acids Res.
9 35 (2007) W126-31. doi:10.1093/nar/gkm219.
- 10 [52] F.W. Studier, Protein production by auto-induction in high-density shaking
11 cultures, Protein Expr. Purif. 41 (2005) 207–234.
12 doi:10.1016/J.PEP.2005.01.016.
- 13 [53] N. Eswar, B. Webb, M. a Marti-Renom, M.S. Madhusudhan, D. Eramian, M.-Y.
14 Shen, U. Pieper, A. Sali, Comparative protein structure modeling using
15 MODELLER., Curr. Protoc. Protein Sci. Chapter 2 (2007) Unit 2.9.
16 doi:10.1002/0471140864.ps0209s50.
- 17 [54] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl,
18 GROMACS: High performance molecular simulations through multi-level
19 parallelism from laptops to supercomputers, SoftwareX. 2 (2015) 19–25.
20 doi:10.1016/j.softx.2015.06.001.
- 21 [55] P. Bjelkmar, P. Larsson, M.A. Cuendet, B. Hess, E. Lindahl, Implementation of
22 the CHARMM force field in GROMACS: analysis of protein stability effects
23 from correction Maps, virtual interaction sites, and water models., J. Chem.
24 Theory Comput. 6 (2010) 459–66. doi:10.1021/ct900549r.
- 25 [56] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein,

- 1 Comparison of simple potential functions for simulating liquid water, J. Chem.
- 2 Phys. 79 (1983) 926. doi:10.1063/1.445869.
- 3 [57] M.H.M. Olsson, C.R. Sndergaard, M. Rostkowski, J.H. Jensen, PROPKA3:
- 4 Consistent Treatment of Internal and Surface Residues in Empirical pK_a
- 5 Predictions, J. Chem. Theory Comput. 7 (2011) 525–537.
- 6 doi:10.1021/ct100578z.
- 7 [58] M. Tiberti, G. Invernizzi, E. Papaleo, (Dis)similarity Index To Compare
- 8 Correlated Motions in Molecular Simulations, J. Chem. Theory Comput. 11
- 9 (2015) 4404–4414. doi:10.1021/acs.jctc.5b00512.
- 10 [59] M. Nygaard, T. Terkelsen, A.V. Olsen, V. Sora, J. Salamanca, F. Rizza, S.
- 11 Bergstrand, M. Di Marco, M. Vistesn, M. Lambrugh, M. Jaattela, T. Kallunki,
- 12 E. Papaleo, The mutational landscape of the oncogenic MZF1 SCAN domain
- 13 in cancer, Front. Mol. Biosci. 3 (2016) 1–18. doi:10.3389/fmolb.2016.00078.
- 14 [60] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A
- 15 smooth particle mesh Ewald method, J. Chem. Phys. 103 (1995) 8577.
- 16 doi:10.1063/1.470117.
- 17 [61] B. Hess, H. Bekker, H. Berendsen, J. Fraaije, LINCS: A linear constraint solver
- 18 for molecular simulations, J. Comput. Chem. 12 (1993) 1463–1472.
- 19 doi:10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-
- 20 H.
- 21 [62] K.R. skarsson, M. Nygaard, B. Ellertsson, S.H. Thorbjarnardottir, E. Papaleo,
- 22 M.M. Kristjnsson, A single mutation Gln142Lys doubles the catalytic activity
- 23 of VPR, a cold adapted subtilisin-like serine proteinase, Biochim. Biophys.
- 24 Acta - Proteins Proteomics. 1864 (2016) 1436–1443.
- 25 doi:10.1016/j.bbapap.2016.07.003.

- 1 [63] A. Di Rita, A. Peschiaroli, P. D'Acunzo, D. Strobbe, Z. Hu, J. Gruber, M.
2 Nygaard, M. Lambrugh, G. Melino, E. Papaleo, J. Dengjel, S. El Alaoui, M.
3 Campanella, V. Dötsch, V. V. Rogov, F. Strappazzon, F. Cecconi, HUWE1 E3
4 ligase promotes PINK1/PARKIN-independent mitophagy by regulating
5 AMBRA1 activation via IKK α , Nat. Commun. 9 (2018) 3755.
6 doi:10.1038/s41467-018-05722-3.
- 7 [64] D. Michetti, B.O. Brandsdal, D. Bon, G.V. Isaksen, M. Tiberti, E. Papaleo, A
8 comparative study of cold- and warm-adapted Endonucleases A using
9 sequence analyses and molecular dynamics simulations, PLoS One. 12 (2017)
10 e0169586. doi:10.1371/journal.pone.0169586.
- 11 [65] W. Kabsch, C. Sander, Dictionary of protein secondary structure: Pattern
12 recognition of hydrogen-bonded and geometrical features, Biopolymers. 22
13 (1983) 2577–2637. doi:10.1002/bip.360221211.
- 14 [66] E. Papaleo, M. Pasi, M. Tiberti, L. De Gioia, Molecular dynamics of mesophilic-
15 like mutants of a cold-adapted enzyme: insights into distal effects induced by
16 the mutations., PLoS One. 6 (2011) e24214.
17 doi:10.1371/journal.pone.0024214.
- 18 [67] G. Invernizzi, M. Tiberti, M. Lambrugh, K. Lindorff-Larsen, E. Papaleo,
19 Communication Routes in ARID Domains between Distal Residues in Helix 5
20 and the DNA-Binding Loops, PLoS Comput. Biol. 10 (2014) e1003744.
21 doi:10.1371/journal.pcbi.1003744.
- 22 [68] O.F. Lange, H. Grubmüller, Full correlation analysis of conformational protein
23 dynamics., Proteins. 70 (2008) 1294–312. doi:10.1002/prot.21618.
- 24 [69] M. Seeber, A. Felling, F. Raimondi, S. Muff, R. Friedman, F. Rao, A. Caflisch,
25 F. Fanelli, Wordom: a user-friendly program for the analysis of molecular

- 1 structures, trajectories, and free energy surfaces., J. Comput. Chem. 32 (2011)
- 2 1183–94. doi:10.1002/jcc.21688.
- 3 [70] M. Pasi, M. Tiberti, A. Arrigoni, E. Papaleo, xPyder: a PyMOL plugin to analyze
- 4 coupled residues and their networks in protein structures., J. Chem. Inf. Model.
- 5 279 (2012) 1–6. doi:10.1021/ci300213c.
- 6 [71] B. Kozlikova, E. Sebestova, V. Sustr, J. Brezovsky, O. Strnad, L. Daniel, D.
- 7 Bednar, A. Pavelka, M. Manak, M. Bezdeka, P. Benes, M. Kotry, A. Gora, J.
- 8 Damborsky, J. Sochor, CAVER Analyst 1.0: graphic tool for interactive
- 9 visualization and analysis of tunnels and channels in protein structures,
- 10 Bioinformatics. 30 (2014) 2684–2685. doi:10.1093/bioinformatics/btu364.
- 11 [72] E. Chovancova, A. Pavelka, P. Benes, O. Strnad, J. Brezovsky, B. Kozlikova,
- 12 A. Gora, V. Sustr, M. Klvana, P. Medek, L. Biedermannova, J. Sochor, J.
- 13 Damborsky, CAVER 3.0: A Tool for the Analysis of Transport Pathways in
- 14 Dynamic Protein Structures, PLoS Comput. Biol. 8 (2012) e1002708.
- 15 doi:10.1371/journal.pcbi.1002708.
- 16 [73] M. Manak, M. Zemek, J. Szkandera, I. Kolingerova, E. Papaleo, M. Lambrugh, i,
- 17 Hybrid Voronoi diagrams, their computation and reduction for applications in
- 18 computational biochemistry, J. Mol. Graph. Model. 74 (2017) 225–233.
- 19 doi:10.1016/j.jmglm.2017.03.018.
- 20

CaMAL

substrate

$\beta 5$ - $\alpha 2$ loop

hhl subdomain

A **B** **C** **D**

A **B** $\beta 5-\alpha 2$ loop **C** hhl subdomain **D**

CaMAL

AAL

317-328 loop

PAL

102-124 loop

HAL

39-80 loop