# Molecular Basis for Extender Unit Specificity of Mycobacterial Polyketide Synthases

Anna D Grabowska, Yoann Brison, Laurent Maveyraud, Sabine Gavalda, Alexandre Faille, Virginie Nahoum, Cécile Bon, Christophe Guilhot, Jean-Denis Pedelacq, Christian Chalut, et al.

# Molecular basis for extender unit specificity of mycobacterial polyketide synthases

Anna D. Grabowska,[#] Yoann Brison,[#] Laurent Maveyraud,[#] Sabine Gavalda, Alexandre Faille, Virginie Nahoum, Cécile Bon, Christophe Guilhot, Jean-Denis Pedelacq,* Christian Chalut,* and Lionel Mourey*

**ABSTRACT:** *Mycobacterium tuberculosis* is the causative agent of the tuberculosis disease, which claims more human lives each year than any other bacterial pathogen. *M. tuberculosis* and other mycobacterial pathogens have developed a range of unique features that enhance their virulence and promote their survival in the human host. Among these features lies the particular cell envelope with high lipid content, which plays a substantial role in mycobacterial pathogenicity. Several envelope components of *M. tuberculosis* and other mycobacteria, e.g. mycolic acids, phthiocerol dimycocerosates and phenolic glycolipids, belong to the 'family' of polyketides, secondary metabolites synthesized by fascinating versatile enzymes – polyketide synthases. These megasynthases consist of multiple catalytic domains, among which the acyltransferase domain plays a key role in selecting and transferring the substrates required for polyketide extension. Here, we present three new crystal structures of acyltransferase domains of mycobacterial polyketide synthases and, for one of them, provide evidence for the identification of residues determining extender unit specificity. Unravelling the molecular basis for such specificity is of high importance considering the role played by extender units for the final structure of key mycobacterial components. This work provides major advances for the use of mycobacterial polyketide synthases as potential therapeutic targets and, more generally, contributes to the prediction and bioengineering of polyketide synthases with desired specificity.

*Mycobacterium tuberculosis*, at the origin of tuberculosis, still constitutes a major threat for public health. About a quarter of the world's population has been infected by *M. tuberculosis*, the leading cause of death from a single infectious agent. In 2018, there were an estimated 1.5 million deaths and 10 million people developed tuberculosis.[1] Furthermore, the development of multiple types of resistance to anti-tuberculosis drugs represents a severe health emergency. Among unique features of *M. tuberculosis* and related species is the multilayer structure of their cell envelope.[2, 3] This thick hydrophobic layer has a multifaceted protective role, acting as a permeability barrier against toxic insults, enhancing pathogen adaptation to the hostile environment of infected humans, and makes the mycobacterial cell envelope a well-established drug target.[4] Important components, which contribute to the cell envelope impermeability and exert a key role as effector molecules in the host-pathogen interplay, are long chain lipids produced by polyketide synthases (PKSs).[5] Among such lipids are mycolic acids,[6] phthiocerol dimycocerosates (DIMs/PDIMs) and structurally related phenolic glycolipids (PGLs)[7] (Figure 1a), which have been shown to play an important role in the virulence of *M. tuberculosis*, as well as *M. leprae* and *M. ulcerans,* the etiological agents of leprosy and Buruli ulcer, respectively.

Polyketide biosynthesis is highly related to that of fatty acids. It proceeds through successive decarboxylative condensations between the growing polyketide chain and coenzyme A (CoA)-derived extender units. Similarly to fatty acid synthases (FASs), PKSs consist of catalytic domains, which may be present on a single high-molecular weight protein (type I PKSs and FASs) or are carried by different proteins (type II PKSs and FASs). Type I PKSs can be subdivided in two categories: those, like FASs, acting iteratively through several cycles of synthesis and modular ones where each module is responsible for a condensation reaction in an assembly line. In addition, modular PKSs can be further differentiated into *cis* or *trans*-acyltransferase (AT), depending on AT domains being embedded or not.[8, 9]. Thus, *cis*-AT modules minimally contain three conserved catalytic domains: an AT that selects and helps the transfer of the extender unit, an acyl carrier protein (ACP) to which the extender unit is transferred, and a ketosynthase domain (KS) that catalyzes the condensation between the intermediate and the extender unit resulting in a β-ketoacyl linked to ACP. PKSs may also contain additional domains displaying ketoreductase (KR), dehydratase (DH), or enoylreductase (ER) activity, which modify the β-keto group leading to a hydroxyl group, a double or a single bond, respectively. Most of the current knowledge about structure-function relationships of type I PKSs has been deduced from structures of the related FAS enzymes[10] and individual domains or didomains derived from PKSs.[11, 12] Although there are no atomic

resolution structures of intact PKSs solved so far, low-resolution models have been derived from single-particle electron cryo-microscopy in the case of the PikAIII pikromycin synthase[13] or by hybrid approach combining X-ray crystallography and small angle X-ray scattering for the *M. smegmatis* mycocerosic acid synthase-like PKS.[14]

Around 20 PKS-encoding genes are present in the *M. tuberculosis* genome.[15] The corresponding PKSs mainly participate in the biosynthesis of complex lipid metabolites.[5] The most studied mycobacterial PKS with respect to its structure, function and its potential use as a druggable target is Pks13, an essential enzyme responsible for the final condensation step in mycolic acid biosynthesis[16-20] (Supporting Information Figure 1a). Pks13 has a particular organization, which comprises the mandatory condensing units described above plus an additional ACP domain and a thioesterase (TE) domain involved in the release of the condensation product[19] (Figure 1b). Compared to other PKSs, the AT domain of Pks13 displays a very unusual specificity, selecting atypical substrates that are long carboxylated fatty acid chains. In a previous study, we determined the crystal structure of a 52 kDa AT-containing Pks13 fragment in apo and (carboxy)palmitoylated forms.[17] Given the importance of the AT domain in selecting elongation units, we decided to expand the characterization of this domain in three other mycobacterial PKSs: Mas, PpsA, and PpsC. All three enzymes are involved in the biosynthesis of DIMs and PGLs (Figure 1a), well documented lipid virulence factors.[7] DIMs were recently shown to be transferred from the mycobacterial envelope to macrophage membranes during infection and to promote phagocytosis.[21] PpsA and PpsC (for phenolphthiocerol/phthiocerol synthase A and C) are encoded within the *ppsABCDE* modular PKS gene cluster and together with PpsB, PpsD and PpsE they participate in the formation of the $C_{32}$–$C_{34}$ phenolphthiocerol/phthiocerol chain (Supporting Information Figure 1b). The iterative Mas (for mycocerosic acid synthase) catalyzes the formation of the methyl-branched mycocerosic acids containing 24 to 30 carbon atoms (Supporting Information Figure 1c). During biosynthesis of DIMs and PGLs, mycocerosates are esterified to the two hydroxyl groups of the (phenol)phthiocerol chain. Our selected PKSs differ in domain organization and in the number and type of auxiliary catalytic domains: DH, ER, and KR for Mas and PpsC; DH and KR for PpsA; none for Pks13 (Figure 1b). They also differ with respect to their substrate and extender unit specificity: methylmalonyl-CoA (MM-CoA) for Mas, malonyl-CoA (M-CoA) for PpsA and PpsC, and long acyl-CoA for Pks13. Finally, they carry out various number of elongation/condensation cycles: 2 to 5 iterations for Mas and only one for PpsA, PpsC and Pks13. In the present work, we describe the X-ray structures of AT-containing PpsA, PpsC and Mas fragments. We identified residues of the AT Mas domain responsible for extender unit

specificity and subjected them to site-directed mutagenesis, providing biochemical and structural evidence for molecular-based reprogramming of mycobacterial PKS specificity towards extender units.

## RESULTS AND DISCUSSION

**Structure determination and overall analysis.** The initial choice of construct boundaries for the AT domains of Mas (residues P447-W871) and PpsA (residues V542-P959) relied on a structure-based sequence comparison with the 52 kDa AT-containing fragment of Pks13,[17] hitherto called AT52 and renamed here L-AT-Pks13 (L standing for the KS-to-AT linker). Those for the AT domain of PpsC (residues G546-L876) were identified using the domain trapping technique.[22] Mas and PpsA constructs also contain the L linker whereas that of PpsC only included a 28-residue long peptide stretch from the upstream linker. After recombinant expression, all proteins could be purified to homogeneity and in sufficient quantity for structural analysis. While L-AT-Mas and AT-PpsC gave crystalline materials that allowed solving the corresponding high-resolution structures, L-AT-PpsA remained reluctant to crystallization. Small-angle X-ray scattering (SAXS) analysis of the purified protein revealed some floppy regions (Supporting Information Figure 2). A shorter construct was therefore designed, which included only the core AT domain (i.e. residues N621-H952) of PpsA. This construct allowed to obtain crystals that diffracted X-rays to a resolution of 1.6 Å. Crystals of AT-PpsC and L-AT-Mas diffracted X-rays to 1.9 and 2.4 Å, respectively. All structures were solved using the molecular replacement method (Supporting Information Table 1).

In terms of sequences, Pks13 and Mas are quite distant as they also are with respect to PpsA and PpsC (29–35% sequence identity), while PpsA and PpsC are much closely related (73% sequence identity). Nevertheless, their AT domains adopt the canonical fold found in individual acyltransferases or in AT domains associated with FASs and PKSs, which comprises the large α/β hydrolase subdomain and a smaller ferredoxin-like subdomain (Figure 2a). They also share a high degree of conservation with respect to both secondary and tertiary structure elements **(**Supporting Information Figure 3**)**. This is also true for the upstream KS-to-AT L domain that precedes the AT domain in the resolved structures for Pks13 and Mas. Indeed, despite a low level of sequence identity (22%), the L domains of Pks13 and Mas have remarkably similar secondary structure elements with different orientations with respect to the AT domain (Figure 2a and Supporting Information Figure 3). A rotation of about 60° is necessary to superpose the L domains once having superimposed the AT domains. A search against the Protein Data Bank (PDB) allowed identifying 16 unique AT structures incorporating a L domain, which are

exclusively found in PKSs and in FASs. No correlation could be established between the position of the L domain and factors such as PKS organization, the fragments used for structure determination, extender unit specificity, etc (Supporting Information Table 2). On the other hand, this variability does not seem to be due to the fact that these L-AT systems have been excised from the whole proteins since the linker of L-AT-Mas is found in the same orientation as that observed in the KS-L-AT structure of the Mas-like PKS.[14] The orientation and dynamics of the linker might then be a trait specific to each protein with respect to intra- or inter-domain interfaces. This is in line with the finding that the flanking linkers of AT domains contribute to both efficiency and specificity of the transacylation reaction on the ACP domain.[23]

**Description of the active sites and catalytic cavities.** AT domains are characterized by a Ser/His catalytic dyad located at the interface between the hydrolase and ferredoxin subdomains. This dyad is found in Pks13 (S801/H909), Mas (S623/H727), PpsA (S720/H822) and PpsC (S660/H762) (Figure 2b and Supporting Information Figure 2) and the catalytic serine is part of a consensus sequence: G-X1-S-X2-G, with (X1=Q, X2=L) in Pks13 and (X1=H, X2=M) in Mas, PpsA and PpsC. In L-AT-Mas, the side chain of the catalytic serine adopts a very unusual conformation where it flips back to the main-chain of the consensus motif, making polar interaction with the side-chain of H622 at position X1 (Figure 2b). Histidine, the other catalytic residue, is at hydrogen bond distance of the hydroxyl group of the serine in all structures except that of L-AT-Mas. The histidine optimal position for increasing the reactivity of the catalytic serine[17] occurs through conserved hydrogen bonds with the main-chain carbonyl groups of two residues (Gly962 and His965 in Pks13; Asn776 and Asn779 in Mas; Asn871 and Asn874 in PpsA; and Ans815 and Asn818 in PpsC) (Figure 2b). The arginine residue involved in the selection of carboxylate substrates[17] is also conserved. However, whereas the corresponding Arg826 makes hydrogen bonds with the hydroxyl group of the catalytic serine in L-AT-Pks13, the side chain of its counterpart in L-AT-Mas (Arg648) is displaced by more than 1.0 Å farther resulting in (or because of) hydrogen bonding to the side chain of Asn776 (Figure 2b). In AT-PpsA and AT-PpsC, the corresponding arginines (Arg745 and Arg685, respectively) are even further displaced because of the side chain of the methionine residue found at position X2 (Met721 and Met661, respectively). Opposite to the hydrolase subdomains, the active sites are lined by a stretch of four consecutive residues from the ferredoxin-like subdomain, including the catalytic histidine at the fourth position, which is supposed to play a role in extender unit specificity, i.e. HAFH and YASH for M-CoA and MM-CoA, respectively.[24] As expected for Pks13 ($^{906}$GASH) but not anticipated for Mas ($^{724}$VASH), PpsA ($^{819}$VAPH) and PpsC ($^{759}$VAPH), the residues found at the corresponding positions of

these mycobacterial PKSs are quite different from the canonical HAFH or YASH residues (Figure 2b). From a topographical point of view, the active sites are accessible via a large opening on the surface of the proteins at the frontier between the two subdomains (Supporting Information Figure 4). An electropositive potential area lines the floor of the active site cavities due to the presence of the arginine residue and catalytic histidine described above (Figure 3). While the electropositive character of the active site floors can be noted, the molecular surfaces as a whole are more electronegative, in accordance with the calculated pI of the proteins studied (between 5.0 and 5.1). However, the distribution of charges varies from one protein to another, probably in relation to the intra- or inter-domain interactions involved in their function (Figure 3).

In *Mycobacterium tuberculosis,* preferred Pks13 AT substrates are ($C_{24}$–$C_{26}$) 2-carboxyacyl-CoA, and it was previously shown that long-chain (carboxy)palmitoylated states could be functionally and structurally characterized for L-AT-Pks13.[17] Ester formation did not induce significant structural changes except for R826, whose side chain moves back to establish a bidentate salt bridge with the carboxylate group in the case of carboxy-$C_{16}$-CoA. The acyl chain adopts two conformations, one of which fits inside a long open preformed hydrophobic channel while the other runs perpendicularly in a deep, solvent-exposed cleft (Supporting Information Figure 5a). In L-AT-Mas, the two channels are blocked due to the presence of the side chains of residues Ile655 and Val724 in the major tunnel and of Gln540 in the exposed cleft (Supporting Information Figure 5b). The same situation occurs in AT-PpsA where the side chain of Val819 (resp. Arg638, Ile691, Glu818) would make steric clash with the acyl chain in the major (resp. minor) conformation (Supporting Information Figure 5c). In contrast, only the major cleft is blocked in PpsC through the side chains of Ile757, Val759, and Leu692 (Supporting Information Figure 5d). Thus, there is no hydrophobic tunnel in L-AT-Mas, AT-PpsA and AT-PpsC like those identified in L-AT-Pks13. As mentioned above, a very important change occurs in the position of the methionine residue from the active site G-H-S-M-G consensus motif in both PpsA (Met721) and PpsC (Met661). Such positioning is responsible for the displacement of the arginine side chain involved in bidendate salt bridge formation with the substrate. Furthermore, in the observed conformation, the methionine side chains would make steric clash with the carboxyl group of any carboxylated substrates (Supporting Information Figures 5c,d).

**Mutational analysis of L-AT-Mas substrate specificity**. In order to better understand the impact of the active site environment on substrate specificity, mutants of the L-AT domain of Mas were generated and their binding properties to M-CoA and MM-CoA were evaluated.

Previous studies have shown that the Ser726 residue of the Mas full-length protein plays a major role in substrate recognition and specificity. Indeed, in contrast to the wild-type Mas protein, which specifically incorporates MM-CoA, a S726F Mas mutant protein was shown to preferentially accept M-CoA as extender unit.[25] First using molecular modeling, then based on the crystal structure of L-AT-Mas, we identified additional amino acid residues, not involved in the catalytic machinery but potentially involved in substrate recognition: W544, V593, H622, M624, I655, V687, V722, D723, V724, A757 and N776 (Figure 2b). To confirm these predictions, we generated single-amino acid substitutions of the selected residues (Table 1). Mutations were chosen according to a previous analysis of active site residues and additional conserved residues for the malonate specific AT domains.[26] The generated mutants, as well as the S726F mutant and the wild-type L-AT-Mas proteins, were overexpressed in *E. coli* and purified to homogeneity to compare their ability to bind either MM-CoA or M-CoA. Three mutated proteins (W544R, H622Y, and A757N), which formed insoluble aggregates upon overexpression, could not be analyzed for substrate binding. As expected, wild-type L-AT-Mas exhibited a marked preference for MM-CoA over M-CoA (Table 1 and Figure 4a). On the other hand, whereas most mutations did not change the specificity towards M-CoA, two mutant proteins (S726F and N776H) exhibited a marked change in substrate preference. Indeed, these two proteins could bind MM-CoA and M-CoA with almost equal affinities (Table 1 and Figure 4a). Of note, mutated residues Val724 and Ser726 both belong to the $^{724}$VASH sequence discussed above and their mutation led to sequences (HASH and VAFH) which are hybrids of the HAFH or YASH motifs involved in M-CoA and MM-CoA specificity, respectively. Only the transformation of VASH into VAFH resulted in a switch of specificity, but in contrast to what was previously shown, the single mutation S726F did not allow, at least in our assay, a complete inversion of specificity in L-AT-Mas.[25] It is noteworthy that, besides the specificity switch observed for N776H, the introduced mutation also drastically increased the affinity of L-AT-Mas for the natural substrate. To compare the relative affinity of the S726F and N776H mutant proteins for each substrate, we performed competitive binding assays by incubating the corresponding molecular species with fixed concentrations of radiolabeled M-CoA or MM-CoA and increasing concentrations of unlabeled competitive ligands. Analysis of radiolabeled proteins by SDS-PAGE and autoradiography indicated that the S726F mutant protein has a higher affinity for M-CoA than for MM-CoA whereas the N776H mutant protein has preference for MM-CoA (Figure 4b).

In an attempt to further improve the selectivity of L-AT-Mas for M-CoA, we constructed a series of double mutant proteins containing the S726F mutation in combination with

substitutions that confer slightly or markedly enhanced selectivity to this substrate, i.e. V593T, M624V, V687A, D723S, V724H, and N776H (Table 1). Among these double mutants, M624V-S726F and V687A-S726F demonstrated substantially improved specificity for M-CoA compared with the S726F L-AT-Mas single mutant protein (Table 1 and Figure 4a). However, we noticed that M624V-S726F had reduced capacity to bind M-CoA and MM-CoA at 30 °C (Figure 4a). Substrate-binding capacities of this variant could greatly be improved for both substrates when assays were performed at lower temperatures (4, 12 and 20 °C) over extended periods, suggesting that the double mutation M624V-S726F may affect the ternary structure and/or binding capacity of the protein at elevated temperatures. Finally, single-residue substitutions S726F and N776H and the M624V-S726F double mutation were introduced into the Mas protein to investigate the influence of these mutations on substrate specificity in the context of the full-length protein. In good agreement with results obtained with the L-AT-Mas domain, the S726F and N776H Mas protein mutants demonstrated similar affinities for M-CoA and MM-CoA whereas the M624V-S726F double mutant protein exhibited higher selectivity toward M-CoA (Figure 4c). As observed for L-AT-Mas, the full-length M624V-S726F variant protein also displayed a weaker affinity for MM-CoA and M-CoA than the corresponding wild-type, S726F and N776H Mas proteins at 30 °C, confirming the possible effects of the double mutation on protein stability and/or ligand-binding capacity. Our analysis revealed that an almost complete inversion of specificity in Mas require secondary mutation such as that of methionine 624 (i.e. mutation M624V) that follows the catalytic serine. Overall, our study shows that the two combined mutations are located in two important sequence patterns of acyltransferases: the consensus active site sequence, which includes the catalytic serine, and the sequence previously identified as the selectivity switch.[24, 27] As demonstrated in several studies, a mutation in only the second pattern seems to result in proteins with relaxed specificity whereas a complete switch of specificity would require secondary mutations.[24, 27-29]

**Structural dissection of L-AT-Mas extender unit specificity.** To get further insight into L-AT-Mas extender unit specificity, important efforts were devoted to solve structures of complexes with M- and MM-CoA. Thus, we solved the structures of the wild-type L-AT-Mas enzyme acylated by the natural substrate MM-CoA as well as the S726F mutant protein alone and acylated with M-CoA or MM-CoA (Supporting Information Table 1). In addition, we have been able to solve the structure of the M624V-S726F doubly-mutated protein after soaking with MM-CoA (Supporting Information Table 1) but attempts to solve the structure of the double mutant after soaking with M-CoA were unfortunately unsuccessful. It is noteworthy that

mutations, crystal handling and/or the strong concentrations of extender units used for soaking (final concentration 4.5 mM) had adverse effects on diffraction, leading to resolution around 3.0 Å compared to 2.4 Å for the structure of native L-AT-Mas whose crystals were already fragile. Decrease in resolution correlated very well with an increase in Wilson B factors (Supporting Information Table 1) and somehow with the diminished thermal stability of mutated proteins when assayed for substrate binding. Soaking of crystals of L-AT-Mas and of the S726F mutant protein with the natural substrate MM-CoA induced moderate structural changes in specific locations (Figure 5a). These changes were found exacerbated in the case of S726F L-AT-Mas after soaking with M-CoA (Figure 5b) and for the M624V-S726F double mutant soaked with MM-CoA (Figure 5c). Taken altogether, the most important structural changes were found in loop α11-β6, at the junction between the hydrolase and ferredoxin domains, and in loop α12-β7, helix α13, strand β9, and loop β9-η4 (Figure 5c). The latter are all located in the ferredoxin domain, which alone seems to absorb all the variations linked to mutations and/or soaking and complex formation. L-AT-Mas crystallized in space group $C222_1$ with two molecules in the asymmetric unit. For consistency reason, we will only focus on chain A whose active site is occupied in the different crystallographic structures solved, which is not the case for chain B (Supporting Information Table 1). Upon soaking with MM or M-CoA, the electron density maps clearly indicate that the side chain of the catalytic serine Ser623 is reoriented towards the catalytic histidine His727 whether the acylation has occurred (i.e. methylmalonylated wild type L-AT-Mas and (methyl)malonylated S726F L-AT-Mas) or not (M624V-S726F L-AT-Mas soaked with MM-CoA) (Figure 6 and Supporting Information Figure 6). It is noteworthy that such a switch of the catalytic serine upon acylation has also been described for the AT domain in the third module of ansamitocin PKS (AsmAT3).[30] By searching through the whole PDB, a similar conformation was also found in the structure of the KS-AT di-domain from module CurL of the curacin A polyketide synthase (4MZ0), and in the structure of mammalian FAS (2VZ8 and 2VZ9). When acylation occurs, the side chain of R648 makes polar interactions with the carboxylate group of the ligands (Figures 6a-c and Supporting Information Figures 6a-c). It is also apparent from the electron density that mutation of Ser726 to a phenylalanine does not impede complex formation. This is probably due to the discrete concerted motions observed in the ferredoxin domain, which allow steric adaption and accomodation of both malonate and methylmalonate substrates. These discrete motions accompany the movement of Phe726 to maintain the hydrophobic core of the ferredoxin domain (Figures 6b,c and Supporting Information Figures 6b,c). In contrast, the M624V-S726F double mutation has drastic structural effects, which completely abolished complex formation with

MM-CoA. Indeed, no electron density was observed in the active site after soaking M624V-S726F L-AT-Mas, (Figure 6d and Supporting Information Figure 6d). In fact, the replacement of methionine by valine at position 624 leads to flipping of the guanidinium group of Arg648 to avoid sterical clash. Furthermore, Phe726 is displaced and reoreinted towards the active site so that the phenyl group is found at the position occupied by the methyl branch of the ligand. The movement of Val624 and Phe726 has drastic effect on the positions of several residues from the ferredoxin domain (Figure 6d and Supporting Information Figure 6d).

## CONCLUSIONS

PKSs are essential players in fundamental biological processes contributing to the chemical diversity observed in life, the most obvious example being the biosynthesis of molecules of pharmaceutical interest. Human and animal pathogens also use this macromolecular arsenal extremely efficiently. This is the case of mycobacteria, for which the biogenesis of the highly complex cell envelope involves a large number of PKSs. One key component of PKSs are AT domains, which control the incorporation of building blocks. Thus, strategies used for engineering of PKSs may rely on AT domain swapping/substitution or AT-site directed mutagenesis.[9] With that respect, essential information have been brought by molecular and structural studies of AT domains. However, a search through the whole PDB identified only a very few structures of AT domains trapped in the acyl-bound state, which correspond in all to six publications, of which four are related to PKSs.[17, 30-32] Our work on three new mycobacterial PKSs and, in particular, the resolution of structures of mutants and complexes for one of them, reinforces the numerous efforts undertaken for the molecular understanding of AT domains. They also show that a double mutation on the central units containing the catalytic dyad can reverse the specificity and that plasticity in the AT ferredoxin subdomain plays an essential role for substrate specificity, as evidenced by X-ray crystallography. The observed plasticity exquisitely extends to the PKS family of megasynthases a major conclusion drawn on type II FAS ACP–AT interactions.[33] The information obtained will help better predict the specificity of polyketide synthases, opening the way to the rational manipulation of these systems to obtain new polyketide skeletons and the rational design of inhibitors, a key step in the development of new antibiotics against mycobacterial infections.

**Supporting Information Available:** Detailed information about methods are available free of charge via the internet at http://pubs.acs.org.

## AUTHOR INFORMATION


### Corresponding authors

**Jean-Denis Pedelacq –** Institut de Pharmacologie et de Biologie Structurale, IPBS, Université de Toulouse, CNRS, UPS, 31077 Toulouse, France; Email: jean-denis.pedelacq@ipbs.fr

**Christian Chalut –** Institut de Pharmacologie et de Biologie Structurale, IPBS, Université de Toulouse, CNRS, UPS, 31077 Toulouse, France; Email: christian.chalut@ipbs.fr

**Lionel Mourey –** Institut de Pharmacologie et de Biologie Structurale, IPBS, Université de Toulouse, CNRS, UPS, 31077 Toulouse, France; orcid.org/0000-0002-8259-1259; Email: lionel.mourey@ipbs.fr

### Authors

**Anna D. Grabowska –** Institut de Pharmacologie et de Biologie Structurale, IPBS, Université de Toulouse, CNRS, UPS, 31077 Toulouse, France

Present address: Department of Biophysics and Human Physiology, Faculty of Health Sciences, Medical University of Warsaw, Poland

**Yoann Brison –** Institut de Pharmacologie et de Biologie Structurale, IPBS, Université de Toulouse, CNRS, UPS, 31077 Toulouse, France

Present address: Toulouse White Biotechnology, TWB, INRAE, INSA, CNRS, 31520 Ramonville Saint-Agne, France

**Laurent Maveyraud –** Institut de Pharmacologie et de Biologie Structurale, IPBS, Université de Toulouse, CNRS, UPS, 31077 Toulouse, France

**Sabine Gavalda –** Institut de Pharmacologie et de Biologie Structurale, IPBS, Université de Toulouse, CNRS, UPS, 31077 Toulouse, France

Present address: Toulouse Biotechnology Institute, TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France

**Alexandre Faille –** Institut de Pharmacologie et de Biologie Structurale, IPBS, Université de Toulouse, CNRS, UPS, 31077 Toulouse, France

Present address: Cambridge Institute for Medical Research, University of Cambridge, Cambridge CB2 0XY, UK

**Virginie Nahoum –** Institut de Pharmacologie et de Biologie Structurale, IPBS, Université de Toulouse, CNRS, UPS, 31077 Toulouse, France

**Cécile Bon –** Institut de Pharmacologie et de Biologie Structurale, IPBS, Université de Toulouse, CNRS, UPS, 31077 Toulouse, France

**Christophe Guilhot –** Institut de Pharmacologie et de Biologie Structurale, IPBS, Université de Toulouse, CNRS, UPS, 31077 Toulouse, France

**Author Contributions**

#A.D.G, Y.B. and L.Ma. contributed equally to this work. All authors have given approval to the final version of the manuscript.

**Funding**

The authors gratefully acknowledge financial support from the Agence Nationale de la Recherche grant ANR-09-BLAN-0298-01 (to C.G. and L.Mo.).

**Notes**

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

We thank the scientific staff of the European Synchrotron Radiation Facility (Grenoble, France) and SOLEIL (Gif sur Yvette, France). We particularly thank the staff of beamlines SWING and PROXIMA-1 at SOLEIL and ID23-1 and 2 at the European Synchrotron Radiation Facility, where the SAXS and crystallographic experiments were conducted. We thank F. Bergeret for his help with the design of protein constructs and V. Guillet for managing synchrotron radiation trips. The crystallization and macromolecular crystallography equipment used in this study are part of the Integrated Screening Platform of Toulouse (PICT, IPBS, IBiSA).

## FIGURE LEGENDS

**Figure 1.** Mycolic acids, phthiocerol dimycocerosates (DIMs) and phenolic glycolipids (PGLs). (a) Chemical structures. The motifs synthesized by Pks13, PpsA, PpsC and Mas are boxed in gray. (b) Domain organization with corresponding residue boundaries (taken from [5]) and related PDB entries, including those from this work (7AGP, 7AHB, and 7AKC). PDB entries 5BP1, 5BP2 and 5BP3 are from the *M. smegmatis* Mas-like PKS.[14] Abbreviations: ACP, acyl carrier protein; AT, acyltransferase; DH, dehydratase; ER, enoyl reductase; KR, ketoreductase; KS, ketosynthase; L, KS-to-AT linker domain. Note: mycolic acids of mycobacteria display a large diversity of chain lengths and different chemical modifications.[6]

**Figure 2.** Crystal structures of L-AT-Pks13 (PDB code 3TZX), L-AT-Mas, AT-PpsA and AT-PpsC. (a) Overall structures in ribbon representation. The α/β-hydrolase subdomain is in *orange*, the ferredoxin-like subdomain is in *yellow*, the KS-AT L linker is in *cyan*, and the AT-ACP linker is in *red*. Helices and β-strands of Pks13 are numbered as in [17]. The histidine and serine residues forming the catalytic dyad and the arginine involved in substrate binding are shown as sticks and boxed. The same color scheme has been applied throughout all the figures. (b) Detailed description of the active sites. Important residues defining the active sites are shown and labeled. Hydrogen bonds are represented by black dotted lines. Water molecules in proximity of displayed residues are shown as red spheres. L-AT-Mas residues that were mutated in the course of this study are depicted in gray with their labels underlined.

**Figure 3.** Topography of the active site. (a) Molecular surface with electrostatic potential color-coded from red (–6 $k_BT/e$) to blue (+6 $k_BT/e$); white is neutral. Residues of L-AT-Pks13 delineating the groove overhanging the active site and the exit of the channel are highlighted in black as in Supporting Information Figure 4. Corresponding areas of L-AT-Mas, AT-PpsA and AT-PpsC are also shown. (b) Semi-perpendicular view.

**Figure 4.** Substrate specificity of the L-AT-Mas and full-length Mas mutants. (a) Ligand binding assays with L-AT-Mas mutants. Purified proteins were incubated with $^{14}$C-radiolabeled M-CoA or MM-CoA. Samples were then subjected to SDS-PAGE followed by staining with Coomassie blue and autoradiography. The bar graph (bottom) summarizes results from at least three independent experiments for each L-AT-Mas variant. Data shown are the means ± SD of

percentage of M-CoA (light gray bars) or MM-CoA (dark gray bars) binding relative to MM-CoA binding by wild type L-AT-Mas (arbitrary set at 100% in each experiment). Here and on other panels, representative Coomassie blue stained gel (top) and autoradiogram (middle) are shown for each protein. Positions of the L-AT-Mas domains on SDS gels are indicated by an arrowhead. (b) Competition binding assays for the S726F and N776H L-AT-Mas mutants. Proteins were incubated with $^{14}$C-radiolabeled M-CoA (M) or MM-CoA (MM), in the absence or in the presence of increasing concentrations of unlabeled M-CoA (m) or MM-CoA (mm). Samples were then subjected to SDS-PAGE followed by staining with Coomassie blue and autoradiography. Bar graphs (bottom) summarize results of 2 independent experiments for each variant. Data with competitive substrates (light gray bars) are expressed as percentage of binding in the absence of unlabeled compound (set at 100% in each experiment) (dark gray bars). Ratios between labeled and unlabeled substrates in samples are indicated. Positions of S726F and N776H L-AT-Mas on the SDS gels are indicated by arrowheads. (c) Ligand-binding assays with the full-length Mas mutants. Purified proteins were incubated with $^{14}$C-radiolabeled M-CoA or MM-CoA. Samples were then subjected to SDS-PAGE followed by staining with Coomassie blue and autoradiography. The bar graph (bottom) summarizes results of 2 independent experiments for each protein. Data shown are the means ± SD of percentage of M-CoA (light gray bars) or MM-CoA (dark gray bars) binding relative to MM-CoA binding by wild type Mas (arbitrary set at 100% in each experiment). Positions of full-length the Mas proteins on SDS gels are indicated by an arrowhead.

**Figure 5.** Effects of mutations and substrate binding on the L-AT-Mas structure. (a) Comparison of native (color-coded as in Figure 2) and methylmalonylated (black) L-AT-Mas. (b) Comparison of native L-AT-Mas with S726F L-AT-Mas alone (light gray) and after malonylation (middle gray) or methylmalonylation (black). (c) Comparison of native L-AT-Mas with M624V-S726F L-AT-Mas soaked in MM-CoA (black). Top raw, overall structures depicted as ribbon with mostly affected regions, i.e. displacement (δd) greater than 2 × <rmsd> after superimposition of Cα carbon atoms, depicted in magenta on the corresponding MM-CoA structure and labeled. Bottom raw, displacement (δd) analysis after superimposition of Cα carbon atoms of native L-AT-Mas with structures obtained after soaking with MM-CoA: left, methylmalonylated L-AT-Mas; middle, methylmalonylated S726F L-AT-Mas; right, unbound M624V-S726F L-AT-Mas. Continuous and dashed lines are for chain A and chain B,

respectively, the two molecules in the asymmetric unit. Red horizontal lines indicate 2 × <rmsd> .

**Figure 6.** Crystallographic analysis of extender unit specificity for wild type and mutant L-AT-Mas proteins. (a) Structure of (2*S*)-methylmalonylated wild type L-AT-Mas. (b) Structure of (2*S*)-methylmalonylated S726F L-AT-Mas. (c) Structure of malonylated S726F L-AT-Mas. (d) Structure of unbound L-AT-Mas M624V-S726F. Structures are shown as black ribbon and were superimposed with native L-AT-Mas (color-coded as in Figure 2). The ligands are in green and the mutated residues in red. The side chains of mostly affected regions, i.e. displacement (δd) greater than 2 × <rmsd> after superimposition of Cα carbon atoms (as in Figure 5), are depicted in magenta and labeled. Feature-enhanced map (FEM)[34] contoured at 1.0 σ is shown in blue for residues from the active site.

**Table 1. Substrate specificity of L-AT-Mas mutants.**

| Mutation(s) | M vs MM (%)[a] |
|---|---|
| WT | 4.3 ± 1.7 (11)[b] |
| V593T | 20.1 ± 3.4 (3) |
| M624V | 7.8 ± 1.8 (5) |
| I655L | 3.2 ± 2.0 (3) |
| V687A | 8.3 ± 5.4 (3) |
| V722I | 5.1 ± 2.6 (5) |
| D723S | 9.4 ± 1.7 (3) |
| V724H | 5.8 ± 3.6 (3) |
| S726F | 125 ± 17.5 (8) |
| N776H | 95.9 ± 5.4 (4) |
| V593T-S726F | 117.0 ± 12.2 (5) |
| M624V-S726F | 225.0 ± 44.7 (6) |
| V687A-S726F | 216.0 ± 9.9 (3) |
| D723S-S726F | 130.0 ± 12.5 (5) |
| V724H-S726F | 103.0 ± 6.2 (4) |
| N776H-S726F | 91.8 ± 10.8 (4) |

[a]For each protein, data shown are the mean ± SD of percentage of M fixation relative to MM fixation, arbitrarily set at 100%.
[b]Values in parentheses represent the number of independent experiments for each mutant.

## REFERENCES


1. WHO. (2019) Global Tuberculosis Report 2019, World Health Organization, Geneva.
2. Daffe, M. (2015) The cell envelope of tubercle bacilli, *Tuberculosis (Edinb) 95 Suppl 1*, S155-158.
3. Daffé, M., and Zuber, B. (2014) The fascinating coat surrounding mycobacteria, In *Bacterial Membranes: Structural and Molecular Biology* (Remau, H., and Fronzes, R., Eds.), pp 179-192, Caister Academic Press.
4. Abrahams, K. A., and Besra, G. S. (2018) Mycobacterial cell wall biosynthesis: a multifaceted antibiotic target, *Parasitology 145*, 116-133.
5. Quadri, L. E. (2014) Biosynthesis of mycobacterial lipids by polyketide synthases and beyond, *Crit. Rev. Biochem. Mol. Biol. 49*, 179-211.
6. Marrakchi, H., Laneelle, M. A., and Daffe, M. (2014) Mycolic acids: structures, biosynthesis, and beyond, *Chem. Biol. 21*, 67-85.
7. Arbues, A., Lugo-Villarino, G., Neyrolles, O., Guilhot, C., and Astarie-Dequeker, C. (2014) Playing hide-and-seek with host macrophages through the use of mycobacterial cell envelope phthiocerol dimycocerosates and phenolic glycolipids, *Front. Cell. Infect. Microbiol. 4*, 173.
8. Keatinge-Clay, A. T. (2017) The uncommon enzymology of cis-acyltransferase assembly lines, *Chem. Rev. 117*, 5334-5366.
9. Musiol-Kroll, E. M., and Wohlleben, W. (2018) Acyltransferases as tools for polyketide synthase engineering, *Antibiotics (Basel) 7*.
10. Maier, T., Leibundgut, M., Boehringer, D., and Ban, N. (2010) Structure and function of eukaryotic fatty acid synthases, *Q. Rev. Biophys. 43*, 373-422.
11. Herbst, D. A., Townsend, C. A., and Maier, T. (2018) The architectures of iterative type I PKS and FAS, *Nat. Prod. Rep. 35*, 1046-1069.
12. Robbins, T., Liu, Y. C., Cane, D. E., and Khosla, C. (2016) Structure and mechanism of assembly line polyketide synthases, *Curr. Opin. Struct. Biol. 41*, 10-18.
13. Dutta, S., Whicher, J. R., Hansen, D. A., Hale, W. A., Chemler, J. A., Congdon, G. R., Narayan, A. R., Hakansson, K., Sherman, D. H., Smith, J. L., and Skiniotis, G. (2014) Structure of a modular polyketide synthase, *Nature 510*, 512-517.
14. Herbst, D. A., Jakob, R. P., Zahringer, F., and Maier, T. (2016) Mycocerosic acid synthase exemplifies the architecture of reducing polyketide synthases, *Nature 531*, 533-537.
15. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence, *Nature 393*, 537-544.
16. Aggarwal, A., Parai, M. K., Shetty, N., Wallis, D., Woolhiser, L., Hastings, C., Dutta, N. K., Galaviz, S., Dhakal, R. C., Shrestha, R., Wakabayashi, S., Walpole, C., Matthews, D., Floyd, D., Scullion, P., Riley, J., Epemolu, O., Norval, S., Snavely, T., Robertson, G. T., Rubin, E. J., Ioerger, T. R., Sirgel, F. A., van der Merwe, R., van Helden, P. D., Keller, P., Bottger, E. C., Karakousis, P. C., Lenaerts, A. J., and Sacchettini, J. C. (2017) Development of a novel lead that targets *Mycobacterium tuberculosis* polyketide synthase 13, *Cell 170*, 249-259 e225.
17. Bergeret, F., Gavalda, S., Chalut, C., Malaga, W., Quemard, A., Pedelacq, J. D., Daffe, M., Guilhot, C., Mourey, L., and Bon, C. (2012) Biochemical and structural study of the atypical acyltransferase domain from the mycobacterial polyketide synthase Pks13, *J. Biol. Chem. 287*, 33675-33690.
18. Gavalda, S., Leger, M., van der Rest, B., Stella, A., Bardou, F., Montrozier, H., Chalut, C., Burlet-Schiltz, O., Marrakchi, H., Daffe, M., and Quemard, A. (2009) The Pks13/FadD32 crosstalk for the biosynthesis of mycolic acids in *Mycobacterium tuberculosis*, *J. Biol. Chem. 284*, 19255-19264.
19. Gavalda, S., Bardou, F., Laval, F., Bon, C., Malaga, W., Chalut, C., Guilhot, C., Mourey, L., Daffe, M., and Quemard, A. (2014) The polyketide synthase Pks13 catalyzes a novel mechanism of lipid transfer in mycobacteria, *Chem. Biol. 21*, 1660-1669.

20. Portevin, D., De Sousa-D'Auria, C., Houssin, C., Grimaldi, C., Chami, M., Daffe, M., and Guilhot, C. (2004) A polyketide synthase catalyzes the last condensation step of mycolic acid biosynthesis in mycobacteria and related organisms, *Proc. Natl. Acad. Sci. U S A 101*, 314-319.
21. Augenstreich, J., Haanappel, E., Ferre, G., Czaplicki, G., Jolibois, F., Destainville, N., Guilhot, C., Milon, A., Astarie-Dequeker, C., and Chavent, M. (2019) The conical shape of DIM lipids promotes Mycobacterium tuberculosis infection of macrophages, *Proc. Natl. Acad. Sci. U S A 116*, 25649-25658.
22. Pedelacq, J. D., Nguyen, H. B., Cabantous, S., Mark, B. L., Listwan, P., Bell, C., Friedland, N., Lockard, M., Faille, A., Mourey, L., Terwilliger, T. C., and Waldo, G. S. (2011) Experimental mapping of soluble protein domains using a hierarchical approach, *Nucleic Acids Res. 39*, e125.
23. Wong, F. T., Chen, A. Y., Cane, D. E., and Khosla, C. (2010) Protein-protein recognition between acyltransferases and acyl carrier proteins in multimodular polyketide synthases, *Biochemistry 49*, 95-102.
24. Reeves, C. D., Murli, S., Ashley, G. W., Piagentini, M., Hutchinson, C. R., and McDaniel, R. (2001) Alteration of the substrate specificity of a modular polyketide synthase acyltransferase domain through site-specific mutations, *Biochemistry 40*, 15464-15470.
25. Trivedi, O. A., Arora, P., Vats, A., Ansari, M. Z., Tickoo, R., Sridharan, V., Mohanty, D., and Gokhale, R. S. (2005) Dissecting the mechanism and assembly of a complex virulence mycobacterial lipid, *Mol. Cell 17*, 631-643.
26. Yadav, G., Gokhale, R. S., and Mohanty, D. (2003) Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases, *J. Mol. Biol. 328*, 335-363.
27. Del Vecchio, F., Petkovic, H., Kendrew, S. G., Low, L., Wilkinson, B., Lill, R., Cortes, J., Rudd, B. A., Staunton, J., and Leadlay, P. F. (2003) Active-site residue, domain and module swaps in modular polyketide synthases, *J. Ind. Microbiol. Biotechnol. 30*, 489-494.
28. Sundermann, U., Bravo-Rodriguez, K., Klopries, S., Kushnir, S., Gomez, H., Sanchez-Garcia, E., and Schulz, F. (2013) Enzyme-directed mutasynthesis: a combined experimental and theoretical approach to substrate recognition of a polyketide synthase, *ACS Chem. Biol. 8*, 443-450.
29. Zhang, F., Shi, T., Ji, H., Ali, I., Huang, S., Deng, Z., Min, Q., Bai, L., Zhao, Y., and Zheng, J. (2019) Structural insights into the substrate specificity of acyltransferases from salinomycin polyketide synthase, *Biochemistry 58*, 2978-2986.
30. Zhang, F., Ji, H., Ali, I., Deng, Z., Bai, L., and Zheng, J. (2020) Structural and biochemical insight into the recruitment of acyl carrier protein-linked extender units in ansamitocin biosynthesis, *Chembiochem 21*, 1309-1314.
31. Li, Y., Zhang, W., Zhang, H., Tian, W., Wu, L., Wang, S., Zheng, M., Zhang, J., Sun, C., Deng, Z., Sun, Y., Qu, X., and Zhou, J. (2018) Structural basis of a broadly selective acyltransferase from the polyketide synthase of splenocin, *Angew. Chem. Int. Ed. Engl. 57*, 5823-5827.
32. Liew, C. W., Nilsson, M., Chen, M. W., Sun, H., Cornvik, T., Liang, Z. X., and Lescar, J. (2012) Crystal structure of the acyltransferase domain of the iterative polyketide synthase in enediyne biosynthesis, *J. Biol. Chem. 287*, 23203-23215.
33. Misson, L. E., Mindrebo, J. T., Davis, T. D., Patel, A., McCammon, J. A., Noel, J. P., and Burkart, M. D. (2020) Interfacial plasticity facilitates high reaction rate of *Escherichia coli* FAS malonyl-CoA:ACP transacylase, FabD, *Proc. Natl. Acad. Sci. U S A 117*, 24224-24233.
34. Afonine, P. V., Moriarty, N. W., Mustyakimov, M., Sobolev, O. V., Terwilliger, T. C., Turk, D., Urzhumtsev, A., and Adams, P. D. (2015) FEM: feature-enhanced map, *Acta Crystallogr. D Biol. Crystallogr. 71*, 646-666.

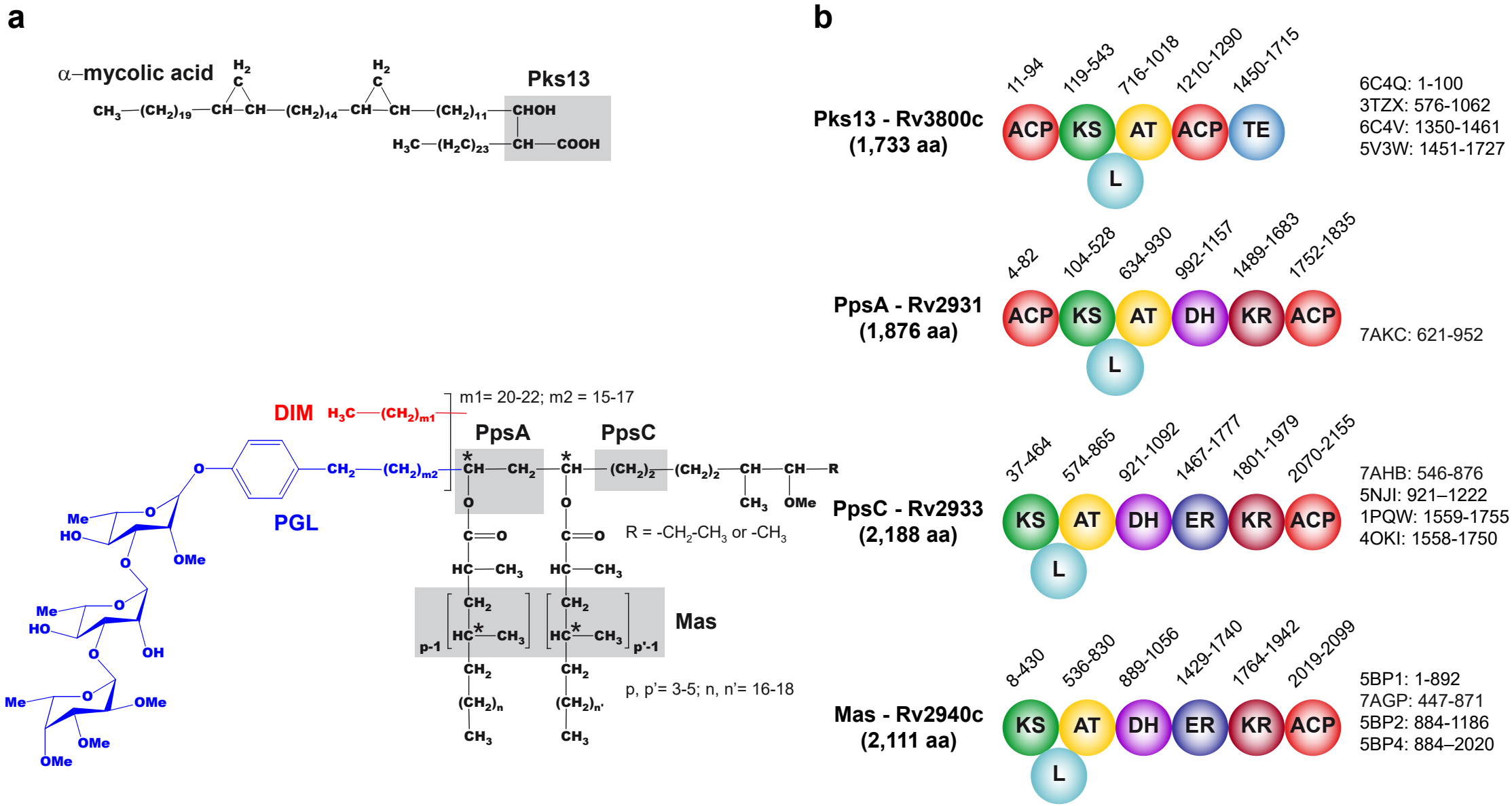
a
α–mycolic acid
Pks13
m1= 20-22; m2 = 15-17
DIM
PpsA
PpsC
PGL
R = -CH2-CH3 or -CH3
Mas
p, p'= 3-5; n, n'= 16-18
b
Pks13 - Rv3800c
(1,733 aa)
11-94
119-543
716-1018
1210-1290
1450-1715
ACP
KS
AT
ACP
TE
L
6C4Q: 1-100
3TZX: 576-1062
6C4V: 1350-1461
5V3W: 1451-1727
PpsA - Rv2931
(1,876 aa)
4-82
104-528
634-930
992-1157
1489-1683
1752-1835
ACP
KS
AT
DH
KR
ACP
L
7AKC: 621-952
PpsC - Rv2933
(2,188 aa)
37-464
574-865
921-1092
1467-1777
1801-1979
2070-2155
KS
AT
DH
ER
KR
ACP
L
7AHB: 546-876
5NJI: 921–1222
1PQW: 1559-1755
4OKI: 1558-1750
Mas - Rv2940c
(2,111 aa)
8-430
536-830
889-1056
1429-1740
1764-1942
2019-2099
KS
AT
DH
ER
KR
ACP
L
5BP1: 1-892
7AGP: 447-871
5BP2: 884-1186
5BP4: 884–2020

Figure 1

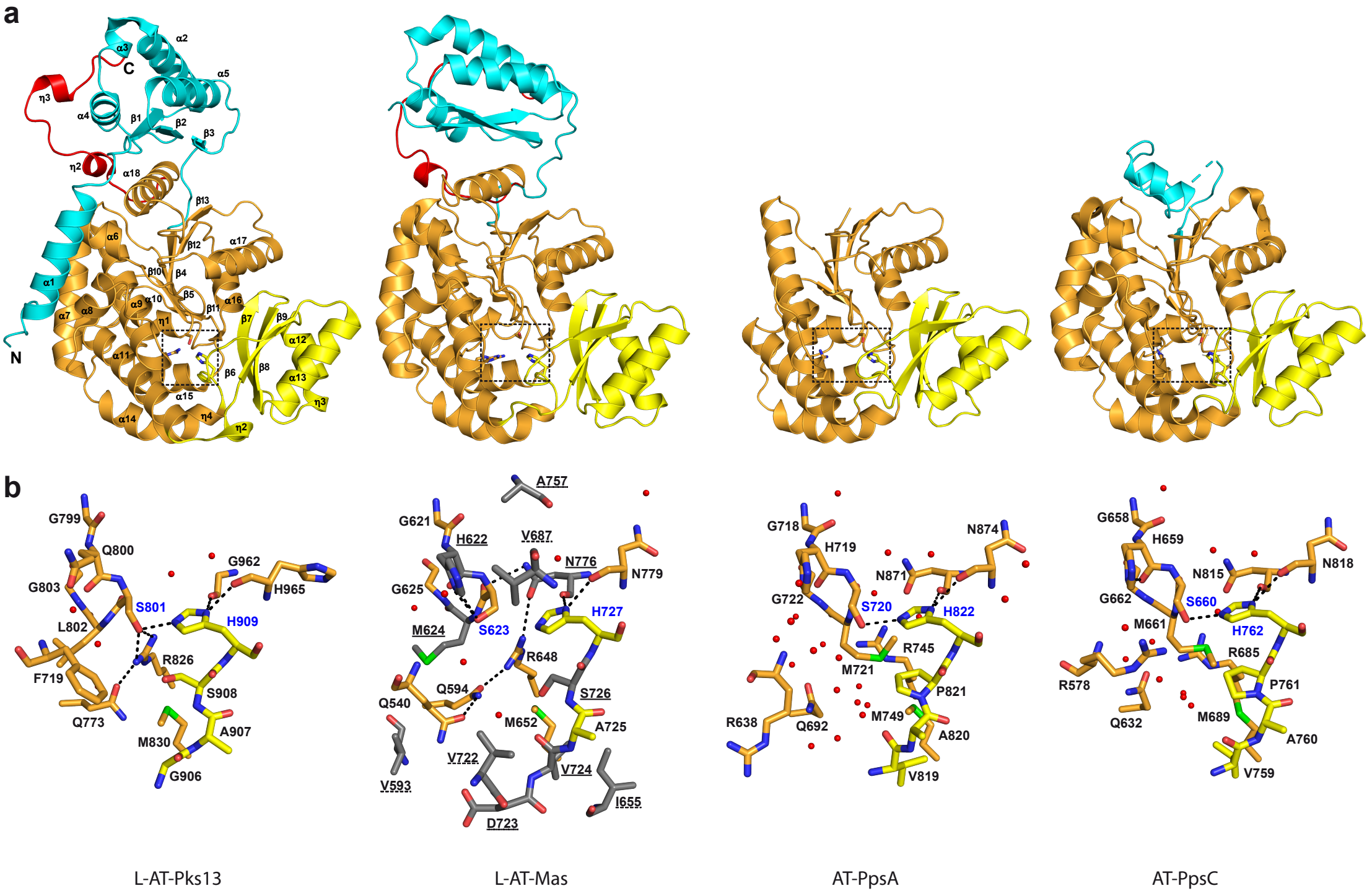
a
C
N
α1
α2
α3
α4
α5
α6
α7
α8
α9
α10
α11
α12
α13
α14
α15
α16
α17
α18
β1
β2
β3
β4
β5
β6
β7
β8
β9
β10
β11
β12
β13
η1
η2
η3
η4
b
G799
Q800
G962
H965
G803
S801
H909
L802
R826
F719
S908
Q773
A907
M830
G906
A757
G621
H622
V687
N776
N779
G625
H727
M624
S623
R648
Q594
S726
Q540
M652
A725
V722
V724
V593
I655
D723
G718
H719
N874
N871
G722
S720
H822
R745
M721
P821
R638
Q692
M749
A820
V819
G658
H659
N818
N815
G662
S660
M661
H762
R685
R578
P761
Q632
M689
A760
V759
L-AT-Pks13
L-AT-Mas
AT-PpsA
AT-PpsC

Figure 2

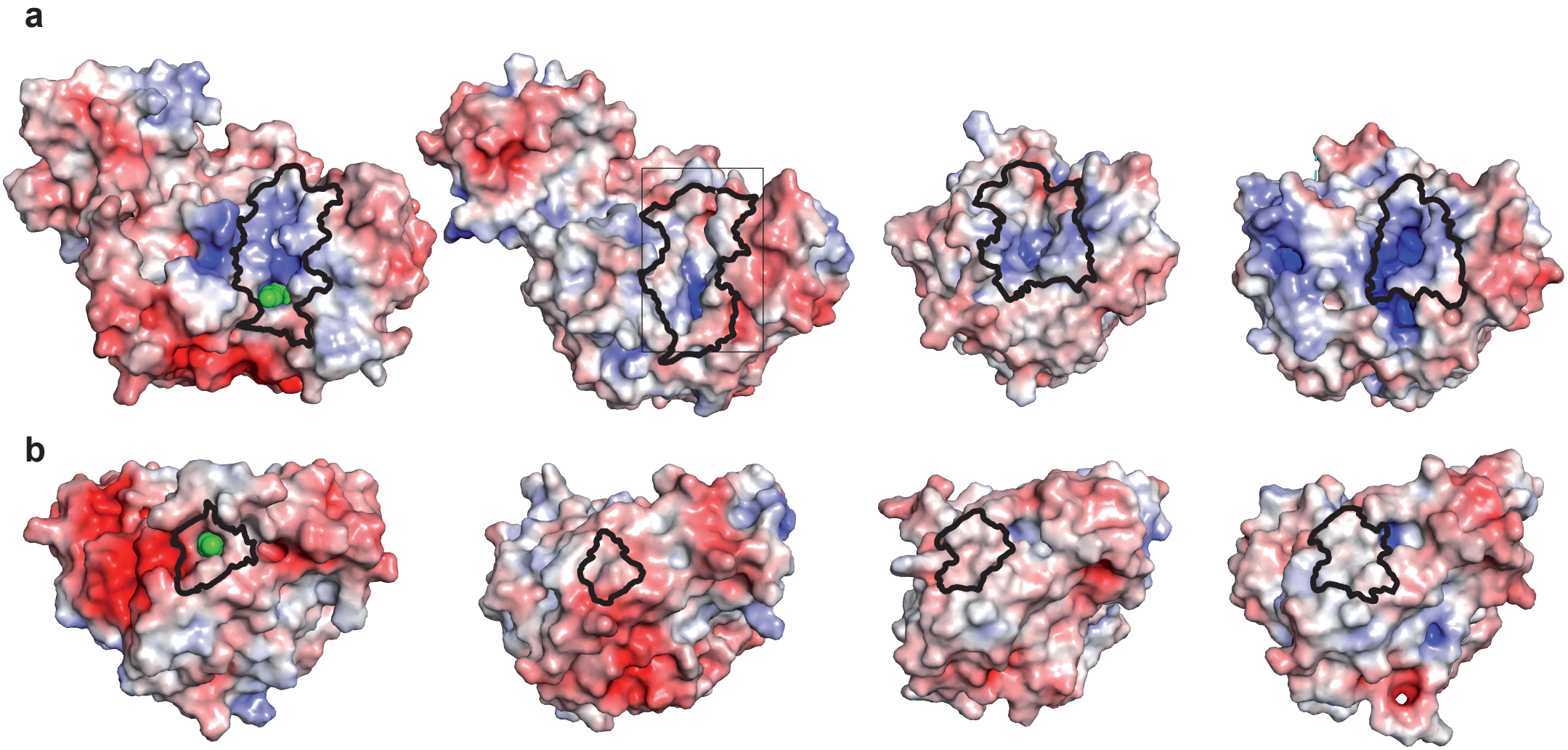


Figure 3

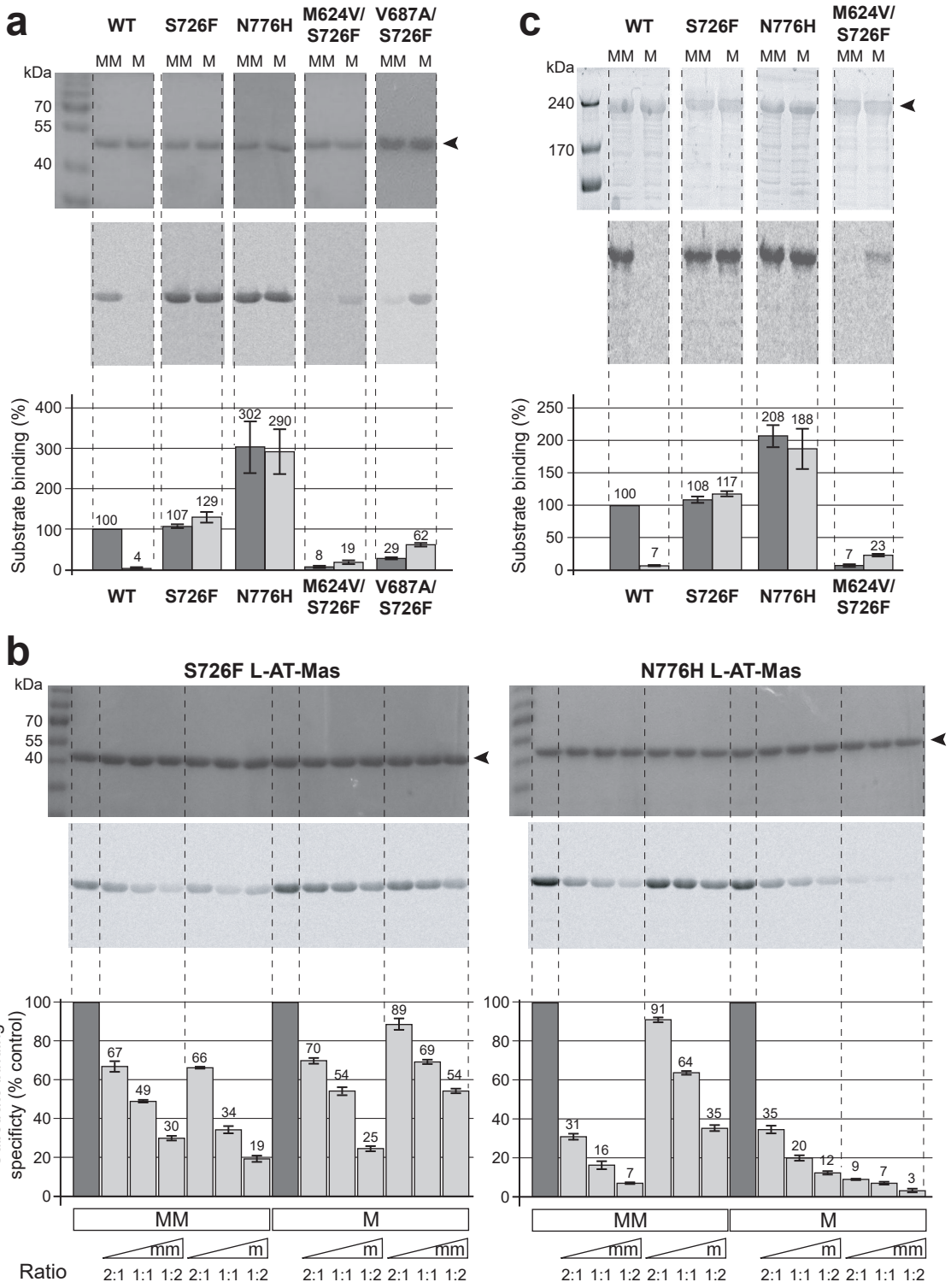
a
WT
S726F
N776H
M624V/ S726F
V687A/ S726F
MM M
kDa
70
55
40
Substrate binding (%)
400
300
200
100
0
100
4
107
129
302
290
8
19
29
62
c
240
170
250
200
150
100
50
100
7
108
117
208
188
7
23
b
S726F L-AT-Mas
N776H L-AT-Mas
Substrate binding specificty (% control)
67
49
30
66
34
19
70
54
25
89
69
54
91
64
35
31
16
7
20
12
9
3
MM
M
mm
m
Ratio
2:1 1:1 1:2 2:1 1:1 1:2 2:1 1:1 1:2 2:1 1:1 1:2

Figure 4

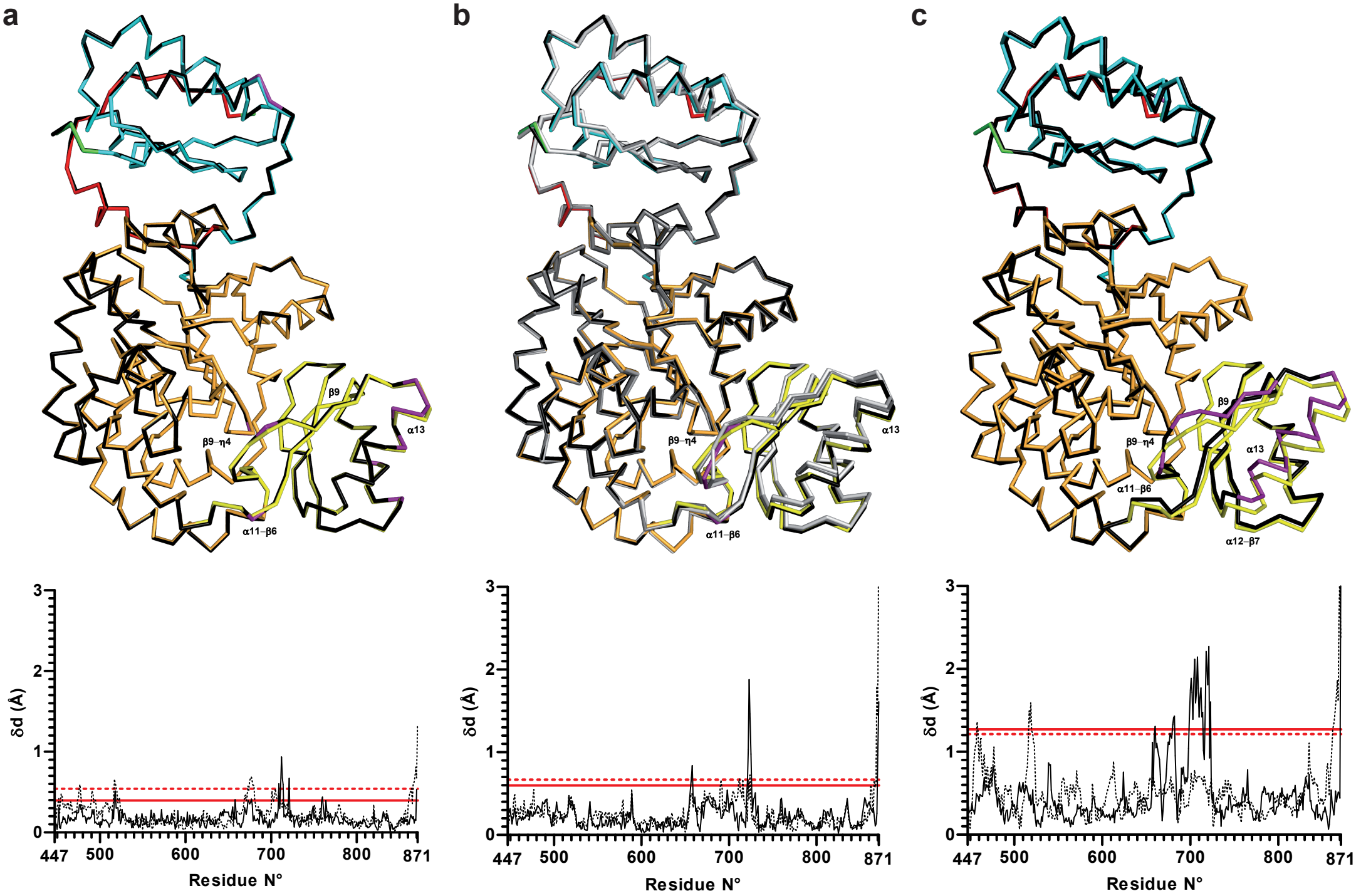
a
b
c
β9
α13
β9–η4
α11–β6
α12–β7
δd (Å)
Residue N°
447
500
600
700
800
871
0
1
2
3

Figure 5

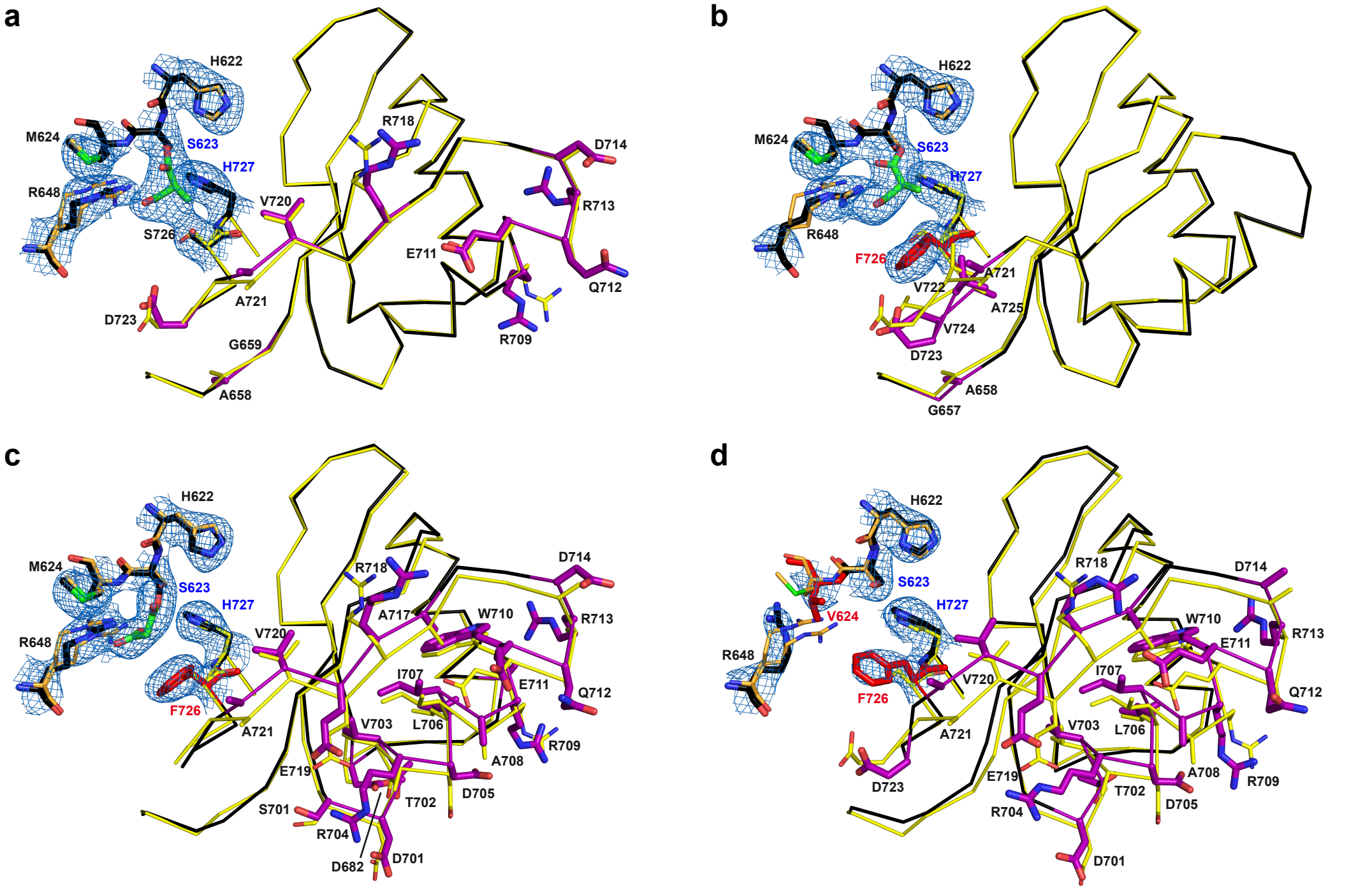
a
H622
M624
S623
H727
R718
D714
R648
V720
R713
S726
E711
Q712
A721
D723
R709
G659
A658
b
H622
M624
S623
H727
R648
F726
A721
V722
A725
V724
D723
A658
G657
c
H622
M624
S623
H727
R718
A717
W710
D714
R713
R648
V720
I707
E711
Q712
F726
A721
V703
L706
R709
A708
E719
D705
T702
S701
R704
D682
D701
d
H622
S623
R718
D714
V624
H727
W710
R713
E711
R648
I707
V720
Q712
F726
A721
V703
L706
A708
R709
D723
E719
T702
D705
R704
D701

Figure 6