**Mycobacterium tuberculosis** FasR senses long fatty acyl-CoA through a tunnel and a hydrophobic transmission spine

Julia Lara1, Lautaro Diacovich1,2, Felipe Trajtenberg3, Nicole Larrieux3, Emilio L. Malchiodi4, Marisa M. Fernández4, Gabriela Gago1, Hugo Gramajo1,2✉ & Alejandro Buschiazzo3,5✉

*Mycobacterium tuberculosis* is a pathogen with a unique cell envelope including very long fatty acids, implicated in bacterial resistance and host immune modulation. FasR is a TetR-like transcriptional activator that plays a central role in sensing mycobacterial long-chain fatty acids and regulating lipid biosynthesis. Here we disclose crystal structures of *M. tuberculosis* FasR in complex with acyl effector ligands and with DNA, uncovering its molecular sensory and switching mechanisms. A long tunnel traverses the entire effector-binding domain, enabling long fatty acyl effectors to bind. Only when the tunnel is entirely occupied, the protein dimer adopts a rigid configuration with its DNA-binding domains in an open state, leading to DNA dissociation. The protein-folding hydrophobic core connects the two domains, and is completed into a continuous spine when the effector binds. Such a transmission spine is conserved in a large number of TetR-like regulators, offering insight into effector-triggered allosteric functional control.
The complex composition of the cell envelope is a distinctive feature of the *Mycobacterium* genus. *Mycobacterium tuberculosis* (*Mtb*) bears peculiar cell wall lipids that play key roles in pathogenicity, also acting as a barrier against environmental stress, antibiotics and the host’s immune response. A better understanding of the mycobacterial cell wall biogenesis will likely identify drug targets for the development of new antibiotics, badly needed to combat tuberculosis.

*Mtbc*’s outer membrane comprises very long-chain fatty acids (mycolic acids), found in the inner leaflet covalently bonded to the arabinogalactan–peptidoglycan layer, and also in the outer leaflet as non-covalently associated lipids in the form of trehalose mono- and di-mycolates. Mycolic acids, a hallmark of *Mycobacterium*, are synthesised by way of two fatty acid synthase systems, FAS I and FAS II. The multidomain single protein FAS I catalyses de novo biosynthesis of acyl-CoA derivatives. Long-chain acyl-CoAs are used as primers by the FAS II multiprotein system, and iteratively condensed with malonyl-acyl carrier protein (malonyl-ACP) leading to very long-chain meromycolyl-ACPs (up to C_{56}). The latter are eventually converted to FAS I-synthetised C_{24–26} fatty acids to produce mycolic acids. FAS I-derived long-chain acyl-CoAs are used not only as mycolic acid precursors, but also for the biosynthesis of phospholipids, triacylglycerides, polyketides and other complex lipids, relevant for the pathogenicity of *Mtb*.

A complex regulatory network integrating all these pathways must exist in order to maintain lipid homeostasis. Despite the biological relevance of lipid homeostasis, little is known about the environmental signals and the regulation cascades controlling lipid metabolism in *Mtb*.

We had previously identified the transcription factor FasR, as a key activator of *fas* and *acpS* gene expression, coding for FAS I synthase and 4-phosphopantetheinyl transferase (essential to produce functional ACP), form a single operon in *Mtb*. FasR activates the transcription of *fas-acpS*, by binding to three inverted repeats in the operon’s promoter region. FasR-DNA binding is regulated by long-chain acyl-CoAs, which are themselves products of FAS I, namely acyl-CoAs ≥ C_{16} disrupt the interaction of FasR with its cognate DNA. In addition, FasR is essential for *Mycobacterium smegmatis* viability, further highlighting its key role in mycobacterial biology.

FasR sequence reveals homology to members of the TetR family of regulators (TFR), which are one-component sensory transduction proteins, typically dimeric and with each protomer displaying a 2-domain all-helical structure. Sequence alignment of FasR with TFRs with known 3D structures (Supplementary Fig. 1a) predicts FasR’s architecture to comprise a helix–turn–helix DNA-binding domain towards its N-terminus (residues 1–80), and a larger C-terminal domain corresponding to the ligand- or effector-binding domain (residues 82–228). In contrast to the DNA-binding domain, the effector-binding region reveals little or no sequence homology with other TFRs, such sequence diversity being consistent with the large variety of effector signals sensed by different TFRs. The prototype of the TFR is TetR from the Tn10 transposon of *Escherichia coli*, which regulates the expression of the tetracycline efflux pump in Gram-negative bacteria. However, TFR proteins are widely distributed among bacteria, and control a broad range of processes, including fatty acid biosynthesis. Interestingly, the vast majority of TFRs are transcriptional repressors, with very few exceptions acting as activators.

We have now determined the 3D structures of FasR from *Mtb*, in three different states obtained with the protein (i) crystallised alone, (ii) co-crystallised in complex with the fatty acid C_{20}-acyl-CoA, and (iii) in complex with a double-stranded DNA oligonucleotide bearing the specific FasR-binding sequence. The comparison of these crystal structures, together with the functional characterisation of structure-guided FasR point mutants and molecular dynamics computational simulations, uncovered the molecular mechanisms by which long- (C_{16}–C_{20}) and very long-chain (C_{22}–C_{26}) acyl-CoA molecules are sensed by FasR, as well as the means by which such signal disrupts cognate FasR–DNA binding and hence activates *fas-acpS* transcriptional activation. World-wide efforts have disclosed hundreds of protein structures from *Mtb* corresponding to potential drug targets, a valuable input for a number of drug discovery projects. The uncovering of structural and mechanistic insights about a key *Mtb* metabolic regulator, contributes with solid molecular bases for target-based drug discovery, a sensible strategy to combat tuberculosis.

**Results**

Three-dimensional structures of FasR. Recombinant FasR eluted from size-exclusion chromatography suggesting a dimeric structure (~52 kDa), similar to all TFRs. Attempts to crystallise full-length FasR alone failed. We hypothesised that the 33-amino acid segment at the N-terminus is likely flexible (Supplementary Fig. 1b). Multiple sequence alignments revealed high sequence variation of the N-terminal extensions added to absence of a predicted secondary structure, leading to the construction of a truncated FasR lacking the first 33 amino acids (FasRΔ33). FasRΔ33 readily crystallised in the absence of added ligands and also in complex with acyl C_{20}-CoA (arachinoyl- or arachidoyl-CoA). Both crystal forms diffracted X-rays at better than 1.7 Å resolution (Supplementary Table 1), and their structures confirm the dimeric architecture of FasRΔ33, with each protomer organized in two all-helical domains (Fig. 1a), similar to known TFRs.

FasRΔ33-C_{20}-CoA was solved using ab initio methods. The structure exhibited the typical TFR architecture, with a DNA-binding HTH (helix–turn–helix) domain from the N-terminus to residue Ser_{82}, comprising helices α1–α3. A regulatory effector-binding domain (EBD) is located immediately C-terminal to the HTH, from Lys_{83} to the C-terminus, including helices α4–α9. The latter helices are roughly organized in two bundles, the α4–α7 core runs along the long axis of the ellipsoidal regulatory domain, whereas α8–α9, roughly perpendicular to the core, mediate dimerisation by forming a 4-helix bundle with the other protomer’s α8′–α9′ helices (Fig. 1a). The FasRΔ33-C_{20}-CoA dimer is strictly symmetric, with the crystallographic twofold axis relating one protomer to the other.

A striking feature of FasRΔ33-C_{20}-CoA is a tunnel-like cavity, delimited by helices α4, α5, α7 and α8, with its two openings towards the bottom and the top of the EBD. C_{20}-CoA binds within this tunnel, in a parallel orientation with respect to core helices α4, α5 and α7 (Fig. 1a, b). The tunnel is ~28 Å long, with a predominance of hydrophobic residues on its wall pointing their side chains towards the lumen of the tunnel (Fig. 1b). The fatty acid is well defined all along the tunnel (Supplementary Fig. 2a). Towards the cavity’s upper entrance, most of the 4′-phosphopantetheine portion of the CoA cofactor is also observed, the sulfur atom (due to its stronger electron density) was instrumental in positioning the whole C_{20}-CoA moiety. Electron density becomes less clear towards the tip of the pantoic group, likely due to high mobility of the CoA portion, eventually vanishing in the region corresponding to the 3′-phosphoadenosine diphosphate group, which were thus not included in the final model.

FasRΔ33 crystals were also grown in the absence of acyl-CoA with the aim of solving the ligand-free structure. Unexpectedly, additional electron density not corresponding to protein, was visible within the ligand-binding tunnel (Fig. 2a, Supplementary Fig. 2b). It could be part of a polyethylene glycol molecule (PEG 400 was included in the crystallisation mother liquor), but PEG’s...
bridging oxygens are energetically costly if buried within the tunnel’s hydrophobic environment. The density is also consistent with myristic acid (C₁₄), which we hypothesize could likely bind during protein expression in E. coli. Even if the chemical nature of the bound species is not certain, this piece of evidence results in two consequences: FasR₃₃₃₋₃₃₃₋₁₄ is not a true apo form of the protein, but it did disclose by serendipity the structure of FasR with a shorter alkyl chain bound in the effector pocket, as compared with FasR₃₃₃₋₃₃₃₋₁₄. In contrast to the strictly symmetric organization of the FasR₃₃₃₋₃₃₃₋₁₄ dimer, FasR₃₃₃₋₃₃₃₋₁₄ displays one full dimer per asymmetric unit (Fig. 2b), each protomer deviates from strict symmetry with respect to the other. A strong twofold non-crystallographic operator relates nonetheless both protomers, but applying only to the regulatory EBDs. The HTH domains of FasR₃₃₃₋₃₃₃₋₁₄ depart from this relationship, after superimposing one protomer onto the other (Fig. 2b), the EBDs fit together well, while the HTHs are rotated by ~15°. To further characterise this symmetry deviation, two identical FasR₃₃₃₋₃₃₃₋₁₄ dimers were superimposed by maximising the fit between one HTH domain from distinct protomers on each dimer (0.3 Å root-mean-square deviation (rmsd) considering all atoms of the two superimposed HTHs). This rotation operation resulted in >5.5 Å rmsd between the other pair of HTHs. A similar exercise using the EBDs revealed a considerably smaller difference, 0.6 Å on the superimposed pair vs 1.5 Å for the other. Departure from intradimer symmetry is thus largely due to substantial flexibility in the region that joins the regulatory and the DNA-binding domains, not the hinge loop covalently linking both domains, but rather the whole region involving the HTH as a rigid domain plus the lower part of the regulatory domain’s helices that interact with the HTH (mainly the N-terminal half of a4). Such type of flexibility mimics the swinging of a pendulum (Fig. 2b, c), and is consistent with higher atomic displacement parameters and weak electron density in the a₆–a₇ loop as well as in the N-terminal half of helix a₇, features that were not apparent in FasR₃₃₃₋₃₃₃₋₁₄ (which displayed more rigidity including in the HTH domains).

The dimerisation interface is an extremely well conserved structural feature among the entire TFR₁₀, always involving a helical bundle constituted by equivalent helices, a₈ and a₉ from each protomer (according to FasR helix numbering scheme). FasR₃₃₃₋₃₃₃₋₁₄ and FasR₃₃₃₋₃₃₃₋₁₄ dimers were superimposed maximising the fit between the corresponding dimerisation helix bundles. Clear differences between both structures were revealed (Fig. 2c). In FasR₃₃₃₋₃₃₃₋₁₄, the dimer is opened up, with both protomers separating away from each other, compared with the...
Acyl-CoA-binding effect on FasR–DNA association. To test the structure-based hypotheses about acyl binding and its effect on FasR–DNA association, point mutants were designed to block the entrance of the ligand into the hydrophobic tunnel. Tunnel-blocking was expected to abolish ligand-triggered rigidification of the protein and consequent DNA-binding hindrance. Two mutants were constructed: FasR\textsubscript{L106F} substitutes Leu\textsubscript{106} by a phenylalanine at the entrance of the tunnel; whereas FasR\textsubscript{LVL} adds bulky side chains not only on position 106, but also substituting Leu\textsubscript{185} and Val\textsubscript{163} by phenylalanines (Fig. 3a). Point mutations did not affect the dimeric architecture of FasR\textsubscript{L106F} (Supplementary Fig. 3), and while slightly <50% of the triple mutant FasR\textsubscript{LVL} eluted as a monomer, >50% behaved as the wild-type protein.

Electrophoretic mobility shift assays (EMSAs) were performed by pre-incubating FasR, FasR\textsubscript{L106F} and FasR\textsubscript{LVL} (its dimeric form) with C\textsubscript{16}-CoA and C\textsubscript{20}-CoA, and then incubating these reactions with 32P-labelled \textit{fas} promoter (P\textit{fas}MT). Strongly supporting our hypothesis, the acyl-CoA ligands triggered bare or not detectable DNA dissociation in the case of FasR\textsubscript{L106F} and FasR\textsubscript{LVL} mutants, while clearly inhibiting DNA-binding of wild-type FasR (Fig. 3b). FasR–DNA apparent dissociation constants (\textit{K}\textsubscript{\textit{dapp}}) were not significantly affected by the point mutations (Supplementary Fig. 4a). Although direct monitoring of acyl-CoA binding to FasR\textsubscript{L106F} and FasR\textsubscript{LVL} mutants was not feasible due to technical impediments, dose–response analyses by EMSA (Supplementary Fig. 4b) further confirmed that the ligands likely do not bind to the

Fig. 2 FasR\textsubscript{Δ33-ΔC14} exhibits larger protein flexibility including dimer asymmetry. a SigmaA-weighted m\textsubscript{F\textsubscript{obs}}–DF\textsubscript{calc} difference Fourier map contoured at 3.5σ (green mesh) calculated with no ligand bound in the effector-binding tunnel during FasR\textsubscript{Δ33-ΔC14} refinement. The FasR\textsubscript{Δ33-ΔC14} model is shown with grey cartoons, and residues at ≤4 Å from the ligand are depicted with thin sticks. The electron density allowed to model a myristic acid, overlaid within the

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2}
\caption{FasR\textsubscript{Δ33-ΔC14} exhibits larger protein flexibility including dimer asymmetry. a SigmaA-weighted m\textsubscript{F\textsubscript{obs}}–DF\textsubscript{calc} difference Fourier map contoured at 3.5σ (green mesh) calculated with no ligand bound in the effector-binding tunnel during FasR\textsubscript{Δ33-ΔC14} refinement. The FasR\textsubscript{Δ33-ΔC14} model is shown with grey cartoons, and residues at ≤4 Å from the ligand are depicted with thin sticks. The electron density allowed to model a myristic acid, overlaid within the}
\end{figure}
tunnel-occluding mutants. While FasRwt displayed nanomolar-range half-maximal inhibitory concentrations (IC₅₀ = 368 nM) and apparent inhibition constants (Kₐ = 42 nM) with C₂₀-CoA, the tunnel-blocking mutants exhibited significantly increased values. Namely, FasRΔ₁₀₆₋₁₉₆ showed dissociation of the protein–DNA complex only at the highest acyl-CoA concentrations (>4 µM), and FasRVL remained bound to DNA even with 6 µM C₂₀-CoA (Supplementary Fig. 4b). Ligand entrance within the tunnel is thus needed in order to trigger the protein conformational change that precludes binding of the regulator to its cognate DNA site.

**Allosteric mechanism of signal transmission.** The shift of the bottom-half of helices α4 and α7, comparing FasRΔ₁₀₆₋₁₉₆, C₁₄, and FasRΔ₁₀₆₋₁₉₆, C₂₀-CoA structures (Fig. 2c), strongly suggested that long enough alkyl chains in the effector-binding tunnel are critical in triggering the rigidification rearrangement, which results in arm-opening. To understand the molecular bases for such effect, the protein region around the distal tip of the ligand acyl chains were analyzed in detail. The very last carbon atoms of C₂₀-CoA interact with mostly bulky hydrophobic residues towards the end of the tunnel, i.e. at the bottom opening of the tunnel that leads to the space separating both promoters in the dimer. Among these hydrophobic residues Leu₁₈₅ (on helix α₄), Phe₁₂₃ (on α₅) and Phe₁₃₈ (on α₆) might play relevant roles (Fig. 3c). The substitution of such voluminous hydrophobic residues by smaller alanine side chains, could uncouple the HTH mobility-restraining effect from ligand-binding. Point mutants FasRΔ₉₈A and FasRΔ₁₂₃A were constructed, which maintained a normal dimeric structure (Supplementary Fig. 2). Both mutants showed significant functional effects, uncoupling ligand-binding and DNA association (Fig. 3d), with a clearer effect observed in the case of FasRΔ₁₂₃A. To further dissect the underlying mechanisms, DNA-binding affinities were analyzed, comparing wild-type vs FasRΔ₉₈A and FasRΔ₁₂₃A mutants (Supplementary Fig. 5). Dissociation constants were quantitated from EMSA data, all in the nanomolar range (Table 1). Compared with FasRwt, FasRΔ₉₈A displayed slightly lower affinity for the Pjₐ₃₅₇ probe and FasRΔ₁₂₃A higher, but neither one exhibiting significant effects. According to the hypothesis that these point mutations would, however, affect efficient transmission of the signal from the effector-binding to the DNA-binding domain, we next calculated IC₅₀ and Kₐ values for the three proteins by evaluating dose–response effects with increasing concentrations of C₂₀-CoA (Supplementary Fig. 6). Indeed, both mutants had significantly lower FasR–DNA-dissociation responses to the ligand (Table 1), particularly so for FasRΔ₁₂₃A. The length of the acyl chain was also critical, with IC₅₀ and Kₐ values all shifted to significantly higher values when C₁₆-CoA was used (Supplementary Fig. 7 and Table 1). These results strongly suggest that residues Leu₁₈₅ and Phe₁₂₃, especially the latter, are key to ensure allosteric signal
transmission while not influencing DNA-binding. That these mutations uncouple signal transmission but do not alter the protein’s affinity for the effector ligand was assessed for the FasR<sub>F123A</sub> mutant (Supplementary Fig. 8). Surface plasmon resonance showed comparable association kinetics of FasR<sub>F123A</sub> to C<sub>20</sub>-CoA as compared to FasR<sub>wt</sub>. In sum, specific residues that are not essential to bind acyl ligands nor DNA, play a key role in transmitting the signal between both domains of the protein once the effector-binding tunnel is fully occupied.

Very long fatty acyl binding anticipates FasR rigidification. Very long fatty acyl moieties are relevant in the biology of Mycobacteriaceae including <i>Mtb</i><sup>16</sup>. Intermediates in the synthesis, and constitutive moieties, of mycolic acids, very long fatty acids are essential components of mycobacterial cell walls. How can such long alkyl chains act as effectors of FasR? A fully extended C<sub>20</sub> acyl chain measures 26.5 Å, within the FasR-C<sub>20</sub>CoA tunnel the acyl shows some bending reducing that length to ~20 Å. The latter magnitude is enough for the C<sub>20</sub> chain to fully occupy the tunnel, its distal tip located immediately beside the bottom opening that leads to the space between protomers in the dimer. We predict that longer acyl chains will be able to accommodate, protruding additional carbon atoms into the interprotomer space. To provide support for this scenario, classical molecular dynamics trajectories were calculated (Supplementary Movie 1) starting from our FasR-C<sub>20</sub>-CoA structure where the C<sub>20</sub>-CoA was substituted in silico by the ~8 Å longer C<sub>26</sub>-CoA (cerotic acyl-CoA), a particularly important biosynthesis intermediate synthesised by FAS I to provide for the α-alkyl chain of mycolic acids. After initial energy minimisation, 10 ns all-atom trajectories were simulated with explicit solvent, showing that the cerotic acyl chains are stable within the tunnel, their 6-carbon extensions indeed protruding into FasR’s interprotomer space (Fig. 4). The available volume in the open form of acyl-bound FasR anticipates even longer acyl chains to be readily accommodated, considering that the protein shows a very stable behaviour with ≤2 Å rmsd within its effector-binding and DNA-binding domains (Supplementary Fig. 9a). Similar trajectories were simulated in the absence of bound effector, suggesting that the effector-binding domain becomes less stable, also exhibiting a larger wiggling of the HTH domains as revealed by calculating the effector-binding tunnel is fully occupied.

The FasR–DNA 3D structure confirms the allosteric mechanism. The crystal structure of full-length FasR was solved by co-crystallisation with a 25-bp double-stranded oligonucleotide bearing the native FasR-binding sequence motif (Fig. 5). A number of crystals and cryo-protections methods were tested, consistently producing strongly anisotropic X-ray diffraction data, reaching 3.85 Å resolution in the best direction (Supplementary Table 1). A form of crystal disorder affected the position of the DNA double helix (details in Methods), occupying equivalent positions in different unit cells while sitting alternatively in both 5′→3′ directions according to a crystallographic twofold.

The protein that was used to grow the FasR–DNA crystals corresponds to full-length FasR. However, the first ~30–35 amino acids at the N-termini are not visible in electron density, indicating they are not bound to the DNA, likely due to high flexibility. Neighbouring FasR dimers related by crystallographic symmetry, are bound to the same DNA fragment on a juxtaposed consecutive manner, at roughly 90° one from the other. Such organization is associated to a pronounced bending of the DNA molecule. The position of the major DNA kink is similar to the one identified in other TFRs, inducing a very similar bending angle as in TetR<sup>17</sup>, or yet similar in magnitude but inverted compared to the archaeal FadR<sup>18</sup> due to the significantly shifted positions of the HTH domains relative to the dimeric effector-binding core (Supplementary Fig. 10). The limited resolution of the FasR–DNA structure, and the presence of crystal disorder affecting the occupancy of the DNA molecules, precluded detailed analyses of protein:DNA interactions. However, the structure did provide two accurate pieces of evidence: (i) the absence of ligand bound within the tunnel of the effector-binding domain, and (ii) a major conformational rearrangement bringing the HTH DNA-binding domains closer together in the dimer (Supplementary Movie 2) such that helices a3 now fit within two successive major grooves on the DNA molecule.

---

**Table 1 FasR association to DNA, and dose-dependent inhibition by effector ligands.**

| Binding to DNA (P<sub>FasMT</sub>) | FasR<sub>wt</sub> | FasR<sub>F123A</sub> | FasR<sub>F123A</sub> |
|-----------------------------------|-----------------|---------------------|---------------------|
| a<sub>K</sub> (nM)                | 83 ± 11         | 155 ± 18            | 39 ± 6              |
| C<sub>20</sub>-CoA dose response  |                 |                     |                     |
| I<sub>C20</sub> (nM)             | 368 ± 105       | 634 ± 100           | 5269 ± 1067         |
| a<sub>K</sub> (nM)                | 42              | 152                 | 571                 |
| C<sub>16</sub>-CoA dose response  |                 |                     |                     |
| I<sub>C16</sub> (nM)             | 5415 ± 1047     | ND                  | ND                  |
| a<sub>K</sub> (nM)                | 589             | ND                  | ND                  |

Comparison of FasR<sub>wt</sub> vs mutants FasR<sub>C44A</sub> and FasR<sub>C52A</sub> that uncouple the allosteric effect. a<sub>K</sub> and I<sub>C50</sub> values are reported as the average of at least three independent experiments ± one standard error of the mean.

ND: not detectable.

---

![Supplementary Fig. 10](image-url)
In addition to the tunnel itself, and in continuity with the bottom opening of it, FasR possesses a cavity delimited by the two protomers. The volume of such cavity in FasR is unusually large (~1800 Å³) compared to many other TFRs: ~600 Å³ (S. enterica Rami PDB 3VYY and M. tuberculosis EthR PDB 5NIO), ~530 Å³ (E. coli RutR PDB 4XX4), ~340 Å³ (E. coli TetR PDB 2XPW), ~135 Å³ (S. acidocaldarius FadR PDB 6EL2), or yet 40 Å³ (P. aeruginosa DesT PDB 3LSJ). This feature is likely relevant, as it anticipates FasR’s ability to accommodate very long fatty acyl chains, such as C30-CoA (synthesised by the Mtb FAS I system26), while maintaining a stable, open configuration (Fig. 4, Supplementary Movie 1). Other fatty acid-sensory TFRs fairly similar to FasR18,21–23, either do not create true continuous tunnels22, or engage a different set of residues running in a perpendicular direction as compared with FasR’s cavity18,21,23. EthR is yet another TFR from Mtb24, intensively investigated as a target to develop anti-tuberculosis medicines. However, EthR is substantially different from FasR (22% sequence identity; ~4.5 Å rmsd after superposition of the effector-binding domains), with a shorter N-terminal extension before the first α helix, and a tunnel displaying wider regions or bulges (Supplementary Fig. 11). Such bulges seem to correlate with EthR’s capacity to bind compounds that include 1–3 aromatic or aliphatic rings.25 At difference with FasR, the physiologic molecules sensed by EthR remain unidentified, despite the >70 EthR crystal structures available, most in complex with surrogate ligands. Among these, only one corresponds to a linear chain (hexadecyl octanoate, in PDBs 1U9N and 1U9O), in this way the most similar to FasR effectors. This ligand is positioned in a similar configuration as the acyl ligands in FasR, but leaving part of EthR’s available tunnel volume unoccupied (Supplementary Fig. 11c), suggesting that physiologic effectors are likely to be of larger size, branched and/or containing bulkier cyclic groups.

FasR is thus equipped to binding very long-chain acyl effectors, but how is such binding coupled to inhibiting DNA association? The type of rearrangements that we have found (Supplementary Movie 2) are consistent with the ones observed in a number of other TFRs, in principle conforming to the mechanistic hypothesis that the binding of effector ligand (the signal) induces an HTH-open conformation, eventually inhibiting TFR association to DNA (the output response). Association of TFRs to DNA indeed requires the HTH domains to close in, in order for the α helices of the two protomers to fit into two successive major grooves (~34 Å apart) on the same side of the cognate DNA26. Such a mechanism has been put forward to explain the workings of E. coli TetR when binding tetracycline27, DesT from Pseudomonas aeruginosa sensing saturated vs unsaturated acyl-CoAs22, RutR recognizing uracil in E. coli28 or yet the multi-drug binding protein QacR from Staphylococcus aureus29, among many others.

Binding the effector stabilises an open configuration of FasR, which does not necessarily imply that the effector mechanically triggers a closed to open transition. If the latter mechanism were true, the ligand-free structure should exhibit a closed, DNA-binding competent configuration. A ligand-free form of FasR could not be crystallised, but turning our attention to the vast number of available TFR crystal structures, those with no effector bound, often correspond to the open form8,10, contradicting the predicted outcome. Among the few apo TFRs that exhibit closed configurations, several reveal crystal packings that fortuitously fix the HTH domains strongly in place (e.g. PDB IDs 2FX0, 1T33, 3VOX and 4JKZ, among others). Moreover, by comparing some of these apo crystals with their DNA-bound counterparts (e.g. 4JKZ vs 4JL3), substantial shifts of the HTH domains can readily be observed, revealing a closed-like configuration, but shifted with respect to the proper DNA-binding-competent one. Reliable information about the true closed configuration has thus largely

---

**Discussion**

FasR is a TFR member that senses long and very long acyl-CoA moieties, subsequently turning off FAS I-mediated fatty acid biosynthesis. Mtb is able to synthesise very long acyl-CoA intermediates (e.g. in the way to synthesising mycolic acids), spanning molecular lengths of 40–50 Å and more. It is known that TFRs possess pockets, sometimes even deeper tunnel-like cavities, which bind to, and enclose the sensed effector ligand. How can FasR deal with the very long effector molecules it senses? It must be stressed that the entire effector-binding domain of FasR measures ~40 Å along its longest axis (and substantially less considering the inner cavity): hence, the fatty acyl effectors can often be longer than the protein’s own physical boundaries. We now answer to this question by revealing a unique hydrophobic tunnel that cuts across the entire effector-binding domain of FasR (Fig. 1), a tunnel that is conspicuously opened on both ends. This singular solution has evolved to lodge the kind of acylated chains of 20 and more carbons that inhibit FasR binding to its cognate DNA6.

This conformational change hampers effector occupation within the tunnel, the latter seems to be constricted by shifted residues (e.g. Phe123) which move their side chains towards the tunnel’s lumen. The FasRΔ3-C14 structure described above proved that acyl-containing compounds from E. coli are invariably associated within the effector-binding cavity, even if specific acyl/acyl-CoA molecules are not added during protein purification and crystallisation. That the FasR tunnel in the FasR–DNA complex is free of bound ligands is supported by unequivocal evidence from difference Fourier maps at both early and late stages of refinement (detailed in Methods). The association to DNA thus correlates to expelling ligands from the effector-binding tunnel of FasR, at least those that attach more loosely within the cavity.
been obtained from crystal structures of TFRs in complex with DNA. Additional evidence further challenges a simple open/closure mechanism: (i) no obvious positional shifts of individual residues can explain the mechanical bases of the alleged pendular movement; (ii) a number of TFR mutants have been identified that either uncouple effector-binding from transcriptional induction\textsuperscript{39}, or invert the effector’s action by triggering a tighter binding to DNA\textsuperscript{31,32}, in both cases often implicating amino acid residues not directly involved in effector-binding.

Finite deformations physics theory seems attractive to highlight allosteric regulation pathways by measuring mechanical strain rather than pairwise atomic position deviations\textsuperscript{33}. Unexpectedly, one of the segments subjected to highest mechanical strain in all TFRs analyzed, corresponds to the loop that connects helices a6 and a7 (Supplementary Fig. 12). A triangle defined by a5, a6 and a7, a conserved feature in all TFR\textsuperscript{8}, harbours the ligand-binding core cavity (which can expand into tunnels with top, bottom and/or lateral openings). Helix a6, associated to a8, is attached to the fixed core, upper-half of the effector-binding domain; but simultaneously, a6’s C-terminal tip and the a6–a7 junction also associate to the moving HTH. In turn, the a6–a7 loop is bound to fixed and moving parts, eventually leading to local deformation. Residues that might explain this strain, in contact with helices a6, the a6–a7 loop and the HTH domain, compose an array of hydrophobic residues that is highly conserved among TFRs (Supplementary Fig. 1 and Supplementary Data 1 and 2), configured in three-dimensions as a continuous spine connecting the two domains of FasR. This spine belongs to, and connects the hydrophobic protein-folding cores of both domains, being interrupted by the ligand-binding cavity in all TFRs analyzed (Fig. 6 and Supplementary Fig. 13). Only in the ligand-bound condition this hydrophobic spine is completed, by the ligand molecule itself at the effector-binding domain, stabilising a rigid and open conformation. Such a mechanism predicts a disordered (flexible) to ordered transition of the TFR protein, which is consistent with the evidence we provide for FasR as well as with available evidence from other TFRs\textsuperscript{8,10,30–32,34}. In particular, fluorescent probes that bind to partially folded proteins in molten globule states\textsuperscript{35}, have been shown to bind promiscuously to apo TFR\textsuperscript{8}. Also, effector-triggered appearance of folding cooperativity between both domains, as well as proteolysis-resistance, have been reported in wild-type and not in allosteric-uncoupled mutants of Tet\textsuperscript{7,36}. Indeed, full-length FasR became more resistant to trypsin proteolysis when pre-incubated with C\textsubscript{59}-CoA (Supplementary Fig. 14), consistent with achieving a more compact fold. Taken together, the extensive body of evidence lends strong support to the transmission spine mechanism as the most consistent interpretation of the effector-mediated allosteric control of TFRs’ DNA-binding function. We cannot, however, exclude that several distinct regulatory mechanisms might have evolved in different subsets of the superfamily, correlated with the broad range of sequence variation of effector-binding domains.

The transmission spine mechanism implies that when the ligand leaves the site, or if it is too short to fully occupy it, the hydrophobic spine is broken, protein folding is sub-optimal and a multitude of conformations of the HTH domain is anticipated (HTH wiggling is illustrated schematically in Fig. 7), including conformations that are competent for DNA-binding. In this line of reasoning, C\textsubscript{14} acyl moiety does not trigger a full-blown disorder-to-order transition, revealed by dimer asymmetry and higher protein flexibility as observed in the FasR\textsubscript{R33–C14} complex. This hypothesis also explains why mutating bulky residues that contribute to building and stabilising the spine (e.g. Leu\textsubscript{a8} and Phe\textsubscript{C123} in FasR), can uncouple the allosteric effect: the ligand is then insufficient to achieve a complete, compact fold in the mutated form (Supplementary Fig. 15). Residues that are not directly involved in effector-binding, but that contribute to building and/or stabilising the spine, will also be able to exert notable effects on allosteric coupling, upholding reported results\textsuperscript{30–32}.

A second conformation, the one bound to DNA, is, however, compactly folded. In this case it is the polynucleotide that pulls on the flexible HTH domains of the dimer bringing them closer together. Correlated to this HTH movement and again transmitted through the hydrophobic spine, several of the bulky residues that line up the effector-binding cavity walls, occlude the cavity (such as Phe\textsubscript{123} in FasR–DNA), completing the spine into a “folded protein-like” core. This occlusion of the ligand-binding pocket in other TFRs when bound to DNA has been described in crystal structures of Tet\textsuperscript{3,39}, Des\textsuperscript{2} and FadR\textsuperscript{32}.

This scenario opens up exciting avenues to be explored. How reversible is the binding of effector compounds, and what triggers their dislodging from the cavity? DNA-binding might likely expel the effector and vice-versa, depending on relative DNA-vs-effector concentrations and affinities. Of note, the assistance of other proteins, in burying the long alkyl moieties when expelled from FasR and other lipid-sensing TFRs should not be ruled out, as they might strongly contribute to the thermodynamics of the allosteric transitions. The hydrophobic transmission spine hypothesis (Fig. 7) may prove instrumental to design better drugs. Eukaryotic protein kinases (ePKs) are not homologous to TFRs, but do exhibit two analogous hydrophobic spines connecting the two domains of ePKs, regulating their activation switch and catalysis\textsuperscript{37}. The spine-mediated regulatory mechanism has been successfully exploited to develop ePK-targeted drugs against cancer and inflammatory diseases\textsuperscript{38,39}. A closer example concerns anti-tuberculosis drug discovery. Comparing EthR structures co-crystallised with different inhibitors\textsuperscript{10}. Compounds simultaneously bearing thienyl and piperidinyl pharmacophores were selected as the most potent among the screening hits. The transmission spine offers a mechanistic explanation as to why the piperidinyl-binding pocket is the crucial region to improving inhibitory activities\textsuperscript{40} in the hit-to-lead development. The piperidinyl- and not the thienyl-interacting pocket engages EthR hydrophobic-spine residues (PDGs 3GIO, 3G1M). Structure-guided drug discovery strategies that exploit the allosteric hydrophobic-spine transmission mechanism might thus prove successful in developing novel medicines against tuberculosis, including multi- and extensively drug-resistant strains.

Methods

Bacterial strains and plasmids. A summarised list of primers, plasmids and cell strains is included as Supplementary Notes in the Supplementary Information. E. coli strain DH5\textalpha cells were used for DNA cloning purposes, transformed according to standard methods. The E. coli BL21 CodonPlus(DE3)-RIL and BL21 λ (DE3) strains were used instead for protein expression. The fasR gene (r5280) was PCR-amplified from Mtb H37Rv genomic DNA using the oligonucleotides F-TevRv2080 (5'-CCCTCCATATGAAACCTGTACTTCCAGGGTATGAGCGATCTCGCC-3') and R-TevRv2080 (5'-GAATTCCTACGAGCGGGTAAGCGG-3') to introduce an NdeI site at the translational start codon and encode a fused Tobacco Etch Virus protease digestion site (TEV), and R-TevRv3208 (5'-GA ATTCTACGAGCGGGTAAGCGG-3') to introduce an EcoRI site at the end of the ORF. To generate a FasR recombinant protein with a hexa-histidine-TEV-tagged fusion, the FasR gene was PCR-amplified from E. coli BL21 λ (DE3) cells, and transforming the PCR product into the plasmid pCR BluntII TOPO (Invitrogen). 

To generate a recombinant protein with a hexa-histidine-TEV-tagged fusion, the FasR gene was PCR-amplified from E. coli BL21 λ (DE3) cells, and transforming the PCR product into the plasmid pCR BluntII TOPO (Invitrogen). 

To generate a recombinant protein with a hexa-histidine-TEV-tagged fusion, the FasR gene was PCR-amplified from E. coli BL21 λ (DE3) cells, and transforming the PCR product into the plasmid pCR BluntII TOPO (Invitrogen).
Expression and purification of proteins. Expressions of all N-terminally hexahistidine-TEV-tagged or hexahistidine-tagged proteins used in this work were carried out following isopropyl-β-thiogalactoside (IPTG) induction in BL21 CodonPlus(DE3)-RIL or BL21 λ (DE3) E. coli. Bacteria were grown at 37 °C in 500 ml LB broth to 0.6–0.7 absorbance at 600 nm. IPTG was then added to 0.3–0.5 mM and the culture was grown for 12 h at 23 °C. Cells were harvested by centrifugation at 2800 × g at 4 °C, resuspended in 30 ml of lysis buffer (50 mM Tris–HCl pH 8, 150 mM NaCl, 5 mM imidazole, 10% glycerol, 10 mM β-mercaptoethanol) and lysed by sonication. After centrifugation (25,000 × g, 4 °C), the supernatant was recovered and FasR, FasRΔ33, or FasR point mutants were separated from whole-cell lysates by Ni-NTA agarose chromatography (Qiagen, Inc). After three washing steps with lysis buffer, 0.5 mg TEV protease and DTT to 1 mM were added and incubated for 3 h at 23 °C and 12 h at 4 °C, with three washing steps with lysis buffer in between. The resin was then washed with lysis buffer containing 250 mM imidazole in lysis buffer, dialysed overnight against FasR Buffer 1 (10 mM Tris–HCl pH 8, 300 mM NaCl), and TEV protease was eluted from the Ni-NTA resin. After three washing steps with lysis buffer, 0.5 mg TEV protease and DTT to 1 mM final concentration were added. The mixture was incubated 2–3 h at 23 °C and 12 h at 4 °C, with three washing steps with lysis buffer in between.

Expression and purification of proteins. Expressions of all N-terminally hexahistidine-TEV-tagged or hexahistidine-tagged proteins used in this work were carried out following isopropyl-β-thiogalactoside (IPTG) induction in BL21 CodonPlus(DE3)-RIL or BL21 λ (DE3) E. coli. Bacteria were grown at 37 °C in 500 ml LB broth to 0.6–0.7 absorbance at 600 nm. IPTG was then added to 0.3–0.5 mM and the culture was grown for 12 h at 23 °C. Cells were harvested by centrifugation at 2800 × g at 4 °C, resuspended in 30 ml of lysis buffer (50 mM Tris–HCl pH 8, 150 mM NaCl, 5 mM imidazole, 10% glycerol, 10 mM β-mercaptoethanol) and lysed by sonication. After centrifugation (25,000 × g, 4 °C), the supernatant was recovered and FasR, FasRΔ33, or FasR point mutants were separated from whole-cell lysates by Ni-NTA agarose chromatography (Qiagen, Inc). After three washing steps with lysis buffer, His6-tagged FasR were eluted from the resin with 250 mM imidazole in lysis buffer, dialysed overnight against FasR Buffer 1 (10 mM Tris–HCl pH 8, 300 mM NaCl), and TEV protease was eluted from the Ni-NTA resin. After three washing steps with lysis buffer, 0.5 mg TEV protease and DTT to 1 mM final concentration were added. The mixture was incubated 2–3 h at 23 °C and 12 h at 4 °C, with three washing steps with lysis buffer in between.

Expression and purification of proteins. Expressions of all N-terminally hexahistidine-TEV-tagged or hexahistidine-tagged proteins used in this work were carried out following isopropyl-β-thiogalactoside (IPTG) induction in BL21 CodonPlus(DE3)-RIL or BL21 λ (DE3) E. coli. Bacteria were grown at 37 °C in 500 ml LB broth to 0.6–0.7 absorbance at 600 nm. IPTG was then added to 0.3–0.5 mM and the culture was grown for 12 h at 23 °C. Cells were harvested by centrifugation at 2800 × g at 4 °C, resuspended in 30 ml of lysis buffer (50 mM Tris–HCl pH 8, 150 mM NaCl, 5 mM imidazole, 10% glycerol, 10 mM β-mercaptoethanol) and lysed by sonication. After centrifugation (25,000 × g, 4 °C), the supernatant was recovered and FasR, FasRΔ33, or FasR point mutants were separated from whole-cell lysates by Ni-NTA agarose chromatography (Qiagen, Inc). After three washing steps with lysis buffer, His6-tagged FasR were eluted from the resin with 250 mM imidazole in lysis buffer, dialysed overnight against FasR Buffer 1 (10 mM Tris–HCl pH 8, 300 mM NaCl), and TEV protease was eluted from the Ni-NTA resin. After three washing steps with lysis buffer, 0.5 mg TEV protease and DTT to 1 mM final concentration were added. The mixture was incubated 2–3 h at 23 °C and 12 h at 4 °C, with three washing steps with lysis buffer in between.

Expression and purification of proteins. Expressions of all N-terminally hexahistidine-TEV-tagged or hexahistidine-tagged proteins used in this work were carried out following isopropyl-β-thiogalactoside (IPTG) induction in BL21 CodonPlus(DE3)-RIL or BL21 λ (DE3) E. coli. Bacteria were grown at 37 °C in 500 ml LB broth to 0.6–0.7 absorbance at 600 nm. IPTG was then added to 0.3–0.5 mM and the culture was grown for 12 h at 23 °C. Cells were harvested by centrifugation at 2800 × g at 4 °C, resuspended in 30 ml of lysis buffer (50 mM Tris–HCl pH 8, 150 mM NaCl, 5 mM imidazole, 10% glycerol, 10 mM β-mercaptoethanol) and lysed by sonication. After centrifugation (25,000 × g, 4 °C), the supernatant was recovered and FasR, FasRΔ33, or FasR point mutants were separated from whole-cell lysates by Ni-NTA agarose chromatography (Qiagen, Inc). After three washing steps with lysis buffer, His6-tagged FasR were eluted from the resin with 250 mM imidazole in lysis buffer, dialysed overnight against FasR Buffer 1 (10 mM Tris–HCl pH 8, 300 mM NaCl), and TEV protease was eluted from the Ni-NTA resin. After three washing steps with lysis buffer, 0.5 mg TEV protease and DTT to 1 mM final concentration were added. The mixture was incubated 2–3 h at 23 °C and 12 h at 4 °C, with three washing steps with lysis buffer in between.

Expression and purification of proteins. Expressions of all N-terminally hexahistidine-TEV-tagged or hexahistidine-tagged proteins used in this work were carried out following isopropyl-β-thiogalactoside (IPTG) induction in BL21 CodonPlus(DE3)-RIL or BL21 λ (DE3) E. coli. Bacteria were grown at 37 °C in 500 ml LB broth to 0.6–0.7 absorbance at 600 nm. IPTG was then added to 0.3–0.5 mM and the culture was grown for 12 h at 23 °C. Cells were harvested by centrifugation at 2800 × g at 4 °C, resuspended in 30 ml of lysis buffer (50 mM Tris–HCl pH 8, 150 mM NaCl, 5 mM imidazole, 10% glycerol, 10 mM β-mercaptoethanol) and lysed by sonication. After centrifugation (25,000 × g, 4 °C), the supernatant was recovered and FasR, FasRΔ33, or FasR point mutants were separated from whole-cell lysates by Ni-NTA agarose chromatography (Qiagen, Inc). After three washing steps with lysis buffer, His6-tagged FasR were eluted from the resin with 250 mM imidazole in lysis buffer, dialysed overnight against FasR Buffer 1 (10 mM Tris–HCl pH 8, 300 mM NaCl), and TEV protease was eluted from the Ni-NTA resin. After three washing steps with lysis buffer, 0.5 mg TEV protease and DTT to 1 mM final concentration were added. The mixture was incubated 2–3 h at 23 °C and 12 h at 4 °C, with three washing steps with lysis buffer in between.

Expression and purification of proteins. Expressions of all N-terminally hexahistidine-TEV-tagged or hexahistidine-tagged proteins used in this work were carried out following isopropyl-β-thiogalactoside (IPTG) induction in BL21 CodonPlus(DE3)-RIL or BL21 λ (DE3) E. coli. Bacteria were grown at 37 °C in 500 ml LB broth to 0.6–0.7 absorbance at 600 nm. IPTG was then added to 0.3–0.5 mM and the culture was grown for 12 h at 23 °C. Cells were harvested by centrifugation at 2800 × g at 4 °C, resuspended in 30 ml of lysis buffer (50 mM Tris–HCl pH 8, 150 mM NaCl, 5 mM imidazole, 10% glycerol, 10 mM β-mercaptoethanol) and lysed by sonication. After centrifugation (25,000 × g, 4 °C), the supernatant was recovered and FasR, FasRΔ33, or FasR point mutants were separated from whole-cell lysates by Ni-NTA agarose chromatography (Qiagen, Inc). After three washing steps with lysis buffer, His6-tagged FasR were eluted from the resin with 250 mM imidazole in lysis buffer, dialysed overnight against FasR Buffer 1 (10 mM Tris–HCl pH 8, 300 mM NaCl), and TEV protease was eluted from the Ni-NTA resin. After three washing steps with lysis buffer, 0.5 mg TEV protease and DTT to 1 mM final concentration were added. The mixture was incubated 2–3 h at 23 °C and 12 h at 4 °C, with three washing steps with lysis buffer in between.
Fig. 7 FasR-mediated long acyl-CoA sensing and response mechanism. The model combines information from the three crystal structures presented in this report (FasR_{A53-C16}, FasR_{A53-C20-CoA} and FasR-DNA complexes). A free structure of FasR with no bound effectors nor DNA has not been determined experimentally, and is not currently known whether it builds up to detectable concentrations within the living cell. The dotted arrows reflect plausible equilibria. See Supplementary Movie 3 to better grasp the anticipated dynamics.

The equilibrium dissociation constant ($K_D$) is a quantitative measurement to assess the affinity of biological interactions. For the FasR-DNA EMSA binding experiments described in this work we define the $K_D$ as the concentration of FasR for which 50% of the DNA is in complex with the protein. The relationship between $K_D$ and affinity is reciprocal, lower $K_D$s correspond to higher affinities. After performing binding reactions in which the protein is titrated, the fraction of DNA bound at each concentration of protein is calculated and the data are adjusted to a binding equation using non-linear regression. The density of DNA bands was assessed with GelPro Analyzer software considering background subtraction. The fraction of bound DNA was plotted as a function of protein concentration and fitted to Eq. (1) using GraphPad Prism software to perform non-linear regression:

$$\text{Fraction bound} = \frac{B_{\text{max}}}{1 + 100 \times \left( E_i - \log(C_{\text{aff}}) \right)}$$

Knowing $[P]$ = protein concentration, the apparent dissociation constant ($appK_D$) can be quantified as well as the maximal fraction bound ($B_{\text{max}}$) plateau.

Acyl-CoA effectors trigger FasR-DNA dissociation, and can thus be treated as non-competitive binder inhibitors. Dose–response experiments were performed to assess inhibition activity, using a modified EMSA protocol. A constant and saturating concentration of protein was equilibrated with 0.3 mM labelled DNA and increasing concentrations of acyl-CoAs in equilibration buffer. The EMSA data were collected as above and fitted to the sigmoidal dose response Eq. (2) in GraphPad Prism to determine the acyl-CoA concentration required to displace half of the bound FasR-DNA complex ($IC_{50}$):

$$\text{Fraction bound} = \frac{B_{\text{max}}}{1 + 100 \times \left( E_i - \log(C_{\text{aff}}) \right)}$$

$E_i = \log$ of acyl-CoA concentration.

EMSA experiments with FasR$_{wt}$ and selected FasR mutants were performed by triplicate, to express $appK_D$ and $IC_{50}$ as average values ± one standard error of the mean.

Apparent inhibition constants ($appK_i$) were also calculated to correct $IC_{50}$ figures taking into account the $K_D$, as well as the concentrations of labelled DNA ([D]) and protein ([P]) according to the Lin and Riggs conversion Eq. (3):

$$appK_i = \frac{2 \times IC_{50}}{[P]}$$

$IC_{50}/[P]$ according to the Lin and Riggs conversion Eq. (3):

Controlled proteolysis. For pre-incubation, 2 mM C$_{20}$-CoA was mixed with 0.2 mM wild-type His$_{6}$-tagged FasR in 10 mM Tris.HCl pH 8, 0.3 M NaCl for 1 h at 25 °C. FasR (16 μM), either pre-incubated or not with C$_{20}$-CoA, was incubated with trypsin (32 nM; Promega, V511C) at 37 °C in the same buffer. At different time points, aliquots were drawn, the digestion stopped by adding SDS sample buffer.
and immediately boiled and analysed by SDS-PAGE. Gels were stained with Coomassie Brilliant Blue, digitalized with a Typhoon FFL 7000 scanner (GE) and densitometry. The quantification of the full-length band was performed with ImageJ. Observed degradation was fitted with an exponential function to compare decay rates.

**Crystallisation and data collection.** FasR\textsubscript{33-C20-CoA} (5 mg/ml) crystallised at 20 °C, mixing 2 + 2 μl of protein and reservoir solution (0.1 M MES monohydrate pH 6.0, 22% v/v polyethylene glycol 400) using 1 ml reservoir on a hanging-drop vapour-diffusion setup. Transferred to mother liquor with 20% (v/v) glycerol as cryo-protector, crystals were mounted in cryo-loops (Hampton Research), and flash cooled in liquid nitrogen (see Supplementary Notes for extended details and references). X-ray diffraction data were collected at ~163 °C at the in-house Protein Crystallography Facility of the Institut Pasteur de Montevideo (Uruguay) with a MicroMax-007 rotating anode and Rigaku X-MARS image plate (marXperts) controlled with the proprietary mar345db software.

FasR\textsubscript{33-C20-CoA} was crystallised in complex with C20-CoA using a 1:1 molar stoichiometry of the acyl-CoA ligand in the crystallisation drops (2 µl protein:ligand + 2 µl mother liquor 2.2 M NaCl, 0.1 M Na-acetate trihydrate pH 4.7). A hanging-drop vapour-diffusion setup was performed with 20% (v/v) mother liquor as reservoir. Crystals were transferred to mother liquor with 25% (v/v) glycerol, mounted in cryo-loops and flash cooled in liquid N\textsubscript{2}. X-ray diffraction data were collected at ~173 °C at SOLEIL synchrotron (PROXIMA 1 beamline, France), using a PILATUS 6 M detector (Dectris) controlled with the open source MXcube software.

FasR\textsubscript{33-C20-CoA} was crystallised in complex with DNA by a vapour-diffusion crystallisation. Double-stranded DNA was generated by co-incubating the two complementary oligonucleotides FwPf25ntm (5′-TACGGTACGTCGACCTGCAAT-3′) and RvPf25ntm (5′-TACCGGGGATCCGCCGATT-3′) under standard slow-cooling hybridisation conditions. Double-stranded DNA was mixed in a 1:1 stoichiometric ratio with FasR\textsubscript{33-C20-CoA} with 2 µl mother liquor 2.2 M NaCl, 0.1 M Na-acetate pH 7.0, 0.1 M imidazole over 1 ml mother liquor as reservoir solution. Crystals were transferred to mother liquor with 35% (v/v) PEG 400, mounted in cryo-loops and flash cooled in liquid N\textsubscript{2}. X-ray diffraction data were collected at ~173 °C at Diamond Light Source synchrotron (I04-1 beamline, UK), using a PILATUS 6M detector (Dectris) controlled with the open source Generic Data Acquisition software.

Bragg diffraction intensities were integrated with XDS\textsuperscript{45}, and scaled and reduced to amplitudes with Aimless and Cryosolver\textsuperscript{46}.

**Structure determination and refinement.** The structure of FasR\textsubscript{33-C20-CoA} was solved ab initio with Arcimboldo\textsuperscript{53} which uses Phaser\textsuperscript{46} as molecular replacement (MR) engine to place α-helices, and Shelxe\textsuperscript{57} for density modification and chain-trace extension. The structure of FasR\textsubscript{33-C14} was solved by MR\textsuperscript{46} using the refined FasR\textsubscript{33-C20-CoA} model as search probe. Buster\textsuperscript{60} was used to refine both FasR\textsubscript{33-C20-CoA} and FasR\textsubscript{33-C14} atomic models, iterating with manual model rebuilding and validation with Coot\textsuperscript{49}. Final validation was done with MolProbity\textsuperscript{61}. OMIT maps were calculated for the FasR\textsubscript{33-C20-CoA} and FasR\textsubscript{33-C14} structures using phenix.re

Molecular dynamics simulations.** The FasR\textsubscript{33-C20-CoA} complex was built using the FasR\textsubscript{33-C20-CoA} (PDB 606N) model as template. The bound acyl-CoA was manually extended by six carbons using PyMol\textsuperscript{62}. C20-CoA was optimized and 10,000 rotamers were generated with RDKit (http://www.rdkit.org). Energy minimisation was performed with the Rosetta suite\textsuperscript{63} using dihedral symmetry constraints, harmonic restraints were used to preserve the ligands positions as observed in the crystal structure, and 10,000 models were generated. The best complex was selected based on Rosetta energy score, with optimal steric contacts and no clashes. The selected model was used as starting structure for classical molecular dynamics simulations using Gromacs 2018_cud8.0 and GROMOS96 43a1 force field\textsuperscript{64}. An octahedron box was solvated and charge-balancing counterions were included to neutralise charges\textsuperscript{65}. Initially, the system was relaxed by energy minimisation, and then equilibrated for 200 ps using a reference temperature of 27 °C. Simulations were performed for 10 ns with no constraints, recording snapshots every 5 ps for analysis. Identical strategy was repeated for the protein alone (the effector ligand was removed from the starting model) and root-mean-squared deviation in atom positions were calculated for different domains as defined in the text and figure captions.

**Data availability** Data supporting the findings of this manuscript are available from the corresponding authors upon reasonable request. A Reporting Summary for this Article is available as a Supplementary Information file. Source data are provided with this paper.

Macromolecular 3D structural data (model coordinates and coordinate file structures) presented in this study have been deposited in the wwPDB with accession codes 6J9P, 6J9Q and 6J9R. X-ray diffraction data for each one of the three structures have been deposited in SBGrid with Digital Object Identifiers 10.15785/SBGRID/648, 10.15785/SBGRID/647 and 10.15785/SBGRID/649, respectively.

Received: 20 June 2019; Accepted: 30 June 2020; Published online: 24 July 2020
References

1. Daffe, M. & Draper, P. in Advances in Microbial Physiology Vol. 39 (ed. Poole, R.) 131–203 (Academic Press, US, 1997).
2. Bhatt, A., Mollé, V., Besra, G. S., Jacobs, W. R. Jr. & Kremer, L. The Mycobacterium tuberculosis FAS-II condensing enzymes: their role in mycolic acid biosynthesis, acid-fastness, pathogenesis and in future drug development. Mol. Microbiol. 64, 1442–1454 (2007).
3. Matsuki, H., Laneelle, M. A. & Daffe, M. Mycolic acids: structural biofilm and beyond. Clin. Microbiol. 21, 67–85 (2014).
4. Daniel, J. et al. Induction of a novel class of diacylglycerol acyltransferases and triacylglycerol accumulation in Mycobacterium tuberculosis as it goes into a dormancy-like state in culture. J. Bacteriol. 186, 5017–5030 (2004).
5. Deb, C. et al. A novel in vitro multiple-stress dormancy model for Mycobacterium tuberculosis generates a lipid-loaded, drug-tolerant, dormant pathogen. PLoS ONE 9, e6077 (2009).
6. Mondino, S., Gago, G. & Gramajo, H. Transcriptional regulation of fatty acid biosynthesis in mycobacteria. Mol. Microbiol. 89, 372–387 (2013).
7. Ulrich, L. E., Koonin, E. V. & Zhihun, I. B. One-component systems dominate signal transduction in prokaryotes. Trends Microbiol. 13, 52–56 (2005).
8. Yu, Z., Reichfeld, S. E., Savchenko, A., Parkinson, J. & Davidson, A. R. A comprehensive analysis of structural and sequence conservation in the TetR family transcriptional regulators. J. Mol. Biol. 400, 847–864 (2010).
9. Rozewicki, J., Li, S., Amada, K. M., Standley, D. M. & Katoh, K. MAFFT-DASH: integrated protein sequence and structural alignment. Nucleic Acids Res. 37, W5–W10 (2009).
10. Cuthbertson, L. & Lodwell, J. R. The TetR family of regulators. Microbiol. Mol. Biol. Rev. 77, 440–475 (2013).
11. Orth, P., Schnappinger, D., Hillen, W., Saenger, W. & Hinrichs, W. Structural basis of gene regulation by the tetracycline inducible TetR repressor-operator system. Nat. Struct. Biol. 7, 215–219 (2000).
12. Zhang, Y. M. & Rock, C. O. Transcriptional regulation in bacterial membrane lipid synthesis. J. Lipid Res. 50(Suppl), S115–S119 (2009).
13. Chim, N. et al. The TB Structural Genomics Consortium: a decade of progress. Tuberculosis 91, 155–172 (2011).
14. Ioerger, T. R. & Sacchettini, J. C. Structural genomics approach to drug perspectives against tuberculosis. Mol. Biol. Rev. 69, 5017–5030 (2004).
15. Willand, N. et al. Synthetic EthR inhibitors boost antibacterial activity of ethionamide. Nat. Med. 15, 537–544 (2009).
16. Heffler, M. A., Walters, R. D. & Kugel, J. F. Using electrophoretic mobility shift assays to measure equilibrium dissociation constants: GAL4-p33 binding DNA as a model system. Biochem. Mol. Biol. Educ. 40, 383–387 (2012).
17. Lin, S. Y. & Riggs, A. D. Lac repressor binding to non-operator DNA: detailed studies and a comparison of equilibrium and rate competition methods. J. Mol. Biol. 72, 671–690 (1972).
18. Evans, P. R. An introduction to data reduction: space-group determination, scaling and intensity statistics. Acta Crystallogr. D Biol. Crystallogr. 65, 282–292 (2011).
19. McCoy, A. J. et al. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674 (2007).
20. Winn, I. & Sheldrick, G. M. An introduction to experimental phasing of macromolecules illustrated by SHELX: new autootracing features. Acta Crystallogr. D Biol. Crystallogr. 74, 106–118 (2018).
21. Bricegno, G. et al. BUSTER v2.10.3 (Global Phasing Ltd., 2017).
22. Emxely, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of COOT. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501 (2010).
23. Williams, C. J. et al. Molsoft: an electronic tool to deliver further reference data for improved all-atom structure validation. Protein Sci. 27, 293–315 (2018).
24. Lieschsher, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr. D Biol. Crystallogr. 75, 861–877 (2019).
25. Brown, A. et al. Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. Acta Crystallogr. D Biol. Crystallogr. 71, 136–153 (2015).
26. Lebedev, A. A. & Iusupov, M. N. Space-group and origin ambiguity in macromolecular structures with pseudo-symmetry and its treatment with the program KAMAD. Acta Crystallogr. D Biol. Crystallogr. 70, 2430–2441 (2014).
27. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. D Biol. Crystallogr. 68, 352–367 (2012).
28. Brown, A. & et al. Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. Acta Crystallogr. D Biol. Crystallogr. 71, 136–153 (2015).
29. Witschel, E. & Henkric, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. D Biol. Crystallogr. 60, 2256–2268 (2004).
30. Notredame, C., Higgins, D. G. & Heringa, J. T-Coffee: a novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. 302, 265–275 (2000).
31. Finn, R. D. et al. HHMER web server: 2015 update. Nucleic Acids Res. 43, W30–W38 (2015).
62. Huang, Y., Niu, B., Gao, Y., Fu, L. & Li, W. CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* **26**, 680–682 (2010).

63. Wallace, I. M., O’Sullivan, O., Higgins, D. G. & Notredame, C. M-Coffee: combining multiple sequence alignment methods with T-Coffee. *Nucleic Acids Res.* **34**, 1692–1699 (2006).

64. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132 (1982).

65. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324 (2014).

66. Das, R. & Baker, D. Macromolecular modeling with rosetta. *Annu. Rev. Biochem.* **77**, 363–382 (2008).

67. Abraham, M. J. et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).

68. Schmit, J. D., Kariyawasam, N. L., Needham, V. & Smith, P. E. SLTCAP: a simple method for calculating the number of ions needed for MD simulation. *J. Chem. Theory Comput.* **14**, 1823–1827 (2018).

Acknowledgements

We thank Stanislas Leibler, Michael Mitchell and Pablo Sartori for sharing initial strain analysis scripts; Matias Machado for assistance in molecular dynamics; Frank Lehmann for initial cloning efforts; Sebastian Klinke (Fundación Leloir) and the staff at Proxima 1 beamline (Soleil synchrotron) and at I04-1 beamline (Diamond synchrotron) for assistance with data collection. We acknowledge computational and storage services (TARS cluster) provided by the Institut Pasteur IT Dept (Paris). We thank the CCP4/CeBEM Macromolecular Crystallography School (USP@São Carlos, 2018), especially Isabel Usón and Paul Emsley for helping us, respectively, with ShelX and Coot, in dealing with low-resolution density modification and model building. J.L. traineeships at IPasteur-Montevideo were funded by CeBEM (www.cebem-lat.org). Support to A.B. from Institut Pasteur (grant 761-International_Joint_Research_Unit-IMiZA-2016), to G.G. from ANPCyT (grant PICT 2015-0796) and to H.G. from ANPCyT (grants PICT 2012-0168 and 2022) and NIH (grant 1R01AI095183-01) are acknowledged.

Author contributions

J.L., L.D., N.L., M.F. and F.T. carried out the experiments and acquired the data; J.L., L.D., G.G., F.T., E.M., H.G. and A.B. conceived the study and performed data analyses and interpretation; J.L., H.G. and A.B. wrote the paper; H.G. and A.B. substantially revised the paper and coordinated the project. All authors gave final approval for publication.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-17504-x.

Correspondence and requests for materials should be addressed to H.G. or A.B.

Peer review information Nature Communications thanks Rene Wintjens and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020