The FEBS Journal

FEBS PRESS science publishing by scientists

The induction mechanism of the flavonoid-responsive regulator FrrA

Nadine Werner[1,*], Sebastiaan Werten[2], Jens Hoppen[1], Gottfried J. Palm[1], Michael Göttfert[3] and Winfried Hinrichs[1]

1 Institute for Biochemistry, Department Molecular Structural Biology, University of Greifswald, Germany
2 Institute of Biological Chemistry, Biocenter, Medical University of Innsbruck, Austria
3 Institute of Genetics, Dresden University of Technology, Germany

Keywords
flavonoid; isoflavonoid; SAXS; TetR-family repressor; X-ray crystallography

Correspondence
W. Hinrichs, Institute for Biochemistry, University of Greifswald, Felix-Hausdorff-Straße 4, 17489 Greifswald, Germany
Tel.: +49 17654172520
E-mail: winfried.hinrichs@uni-greifswald.de

***Present address**
Centre for Integrative Biology, Department of Integrated Structural Biology, IGBMC (Institute of Genetics and of Molecular and Cellular Biology), 1 rue Laurent Fries, Illkirch, France

Nadine Werner and Sebastiaan Werten contributed equally to this work.

(Received 16 May 2021, revised 13 July 2021, accepted 26 July 2021)

doi:10.1111/febs.16141

Bradyrhizobium diazoefficiens, a bacterial symbiont of soybean and other leguminous plants, enters a nodulation-promoting genetic programme in the presence of host-produced flavonoids and related signalling compounds. Here, we describe the crystal structure of an isoflavonoid-responsive regulator (FrrA) from *Bradyrhizobium*, as well as cocrystal structures with inducing and noninducing ligands (genistein and naringenin, respectively). The structures reveal a TetR-like fold whose DNA-binding domain is capable of adopting a range of orientations. A single molecule of either genistein or naringenin is asymmetrically bound in a central cavity of the FrrA homodimer, mainly via C–H contacts to the π-system of the ligands. Strikingly, however, the interaction does not provoke any conformational changes in the repressor. Both the flexible positioning of the DNA-binding domain and the absence of structural change upon ligand binding are corroborated by small-angle X-ray scattering (SAXS) experiments in solution. Together with a model of the promoter-bound state of FrrA our results suggest that inducers act as a wedge, preventing the DNA-binding domains from moving close enough together to interact with successive positions of the major groove of the palindromic operator.

Introduction

Rhizobia are Gram-negative soil bacteria, particularly known for their ability to form an endosymbiotic nitrogen-fixing association with legumes [1, 2]. Attracted by substances that are released by the plant, especially (iso)flavonoids and related compounds, rhizobia colonize their host while activating hundreds of genes that promote the formation of root nodules. Simultaneously, efflux systems are activated that mitigate the toxicity of the flavonoids. How rhizobia sense the presence of these molecules and react to them is still not fully understood.

The TetR-family comprises a large number of dimeric transcription repressors, commonly involved in the regulation of bacterial responses to small organic compounds, solvents and metals [3,4]. These proteins have N-terminally a highly conserved DNA-binding domain, whereas the more divergent C-terminal regulatory domain forms the dimer interface and one or two ligand-binding pockets [3]. The interaction with ligand molecules causes conformational changes which have allosteric effects on the repressor incompatible

Abbreviations
CHES, 2-(cyclohexylamino)ethanesulfonic acid; DBD, DNA-binding domain; HTH, helix–turn–helix; SeMet, L-selenomethionine.

for DNA binding, inducing the transcription of target genes.

Although database searches reveal that rhizobial genomes encode numerous TetR-like proteins [4], only few have been studied. In *Rhizobium etli*, CasR negatively regulates the expression of *casA*, which encodes a calmodulin-like protein [5]. The *emrAB* genes of *Sinorhizobium meliloti*, likely to encode an efflux system, are regulated by EmrR [6,7]. EmrR was shown to bind to palindromic sequences in the *emrAB-emrR* intergenic region. Binding was affected by luteolin, which is a known signal molecule in the symbiosis between *S. meliloti* and alfalfa (*Medicago sativa*).

In *Bradyrhizobium diazoefficiens* USDA 110 (formerly *B. japonicum*), two TetR family members, FrrA and BdtR, have been characterized in detail [8,9]. Both repress their own transcription as well as the expression of genistein-inducible efflux systems encoded by neighbouring genes in the opposite orientation. Genistein and daidzein (Fig. 1A), two known signal molecules secreted by soybean seedlings [10], interfere with DNA binding by FrrA, whereas BdtR responds in similar fashion to genistein but not to daidzein [8,9]. For FrrA, it was shown that coumestrol (Fig. 1A) also induces gene expression, albeit not as efficiently as genistein and daidzein [8]. Naringenin (Fig. 1A), a flavonoid released by pea seedlings [11], had no effect at all [8]. However, a detailed understanding of the origins of this specificity is still lacking. Here, we report the comprehensive structural analysis of a flavonoid-responsive regulator by X-ray crystallography and small-angle scattering (SAXS) in solution, revealing how ligands are recognized and enabling us to propose a model for the induction of this class of repressors.

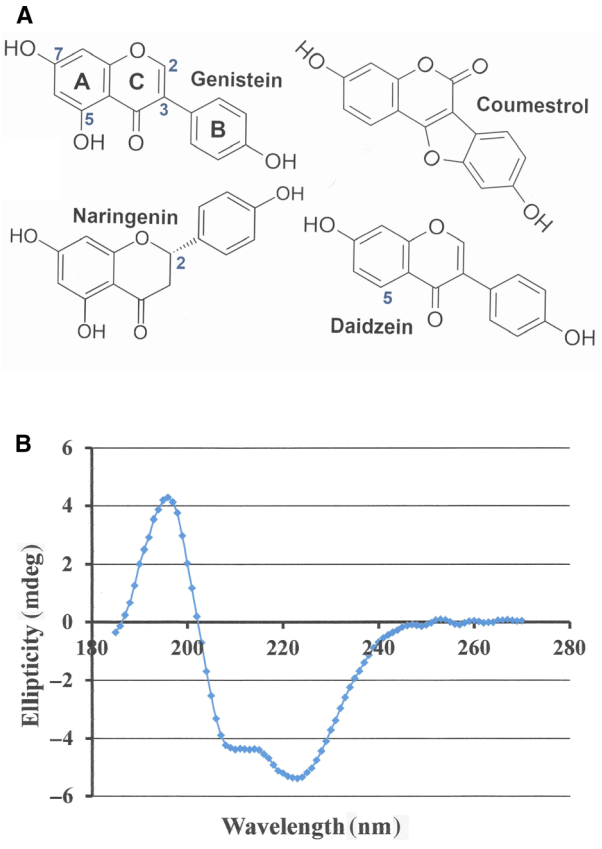

Fig. 1. (A) Chemical structure of isoflavonoid genistein (IUPAC nomenclature indicated) and flavonoid (*S*)-naringenin. In daidzein, the hydroxyl group 5 of genistein is replaced by hydrogen. Note, that coumestrol is a derivative of genistein. (B) CD-spectrum of FrrA.

Results

FrrA is a member of the TetR family

According to blue native acrylamide gel electrophoresis [9] and SAXS data (see below), FrrA is a homodimer. CD spectroscopy revealed an exclusively α-helical protein (Fig. 1B), confirmed by X-ray crystal structure analysis. The ligand-free FrrA and its complexes with genistein and naringenin crystallized isomorphously with two almost identical dimers in the asymmetric unit. The domain organization of the FrrA homodimer shows a typical TetR-like fold (Fig. 2A). Previously, sequence analysis of FrrA revealed high sequence identity of the N-terminal DNA-binding domain [9] typically observed for the TetR/AcrR transcriptional regulator family [4,12]. This domain consists of a three-helix bundle (α1 – α3, residues 19 – 63) plus a longer helix (α4, residues 65 – 82) that connects it to the regulatory domain. The typical helix–turn–helix (HTH) motifs (α2, α3 and their symmetry mates) bind to two adjacent positions of the major groove of the palindromic operator DNA [13,14]. In brief, both monomers participate in dimer formation with six α-helices (α5 – α10) representing the ligand-binding regulatory domain. Most of these α-helices of the regulatory domain are longitudinally oriented to the dimer axis. Helix α6 and α8 are perpendicular to the dimer axis. Helices α9 and α10 are antiparallel and form a four-helix bundle with their symmetry mates at an angle of ~60 °, a typical dimerization motif of the TetR family [15].

A search for structures comparable to FrrA using the DALI server [16,17] identified a plenitude of transcriptional regulators of the TetR-family with highly similar structures. The DALI search revealed 64 structures with a *Z*-score ≥ 10.0, but a sequence identity of less than 25%. Top examples are a repressor from *Mesorhizobium loti* MAFF 303099 (PDB ID: 3BHQ, Z-score

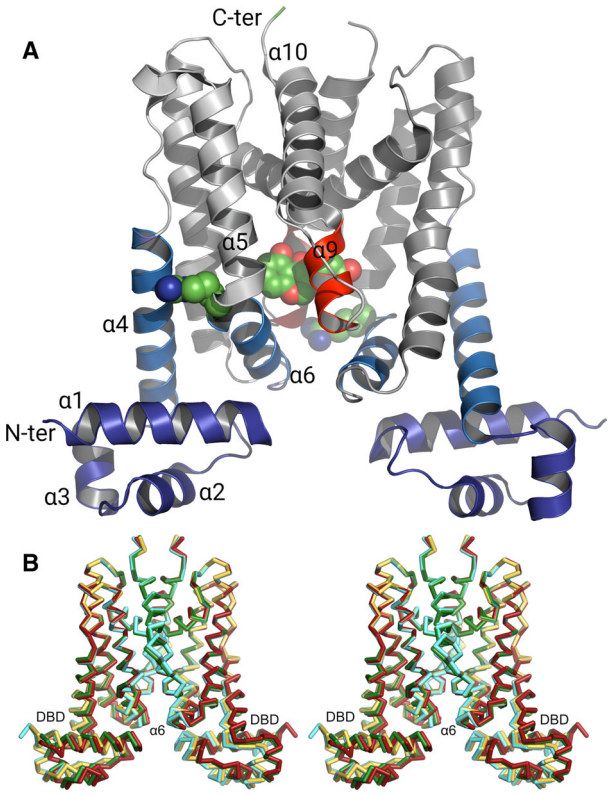

Fig. 2. (A) All-helical FrrA structure with vertical dimer axis. The N and C termini are labelled in monomer B. The DNA-binding domains (helices α1 – α3) are shown in blue, helices α4 and α6 light blue. Atoms of the central genistein and the side chains of Arg111 are represented as spheres. Arg111 is located in the short connecting turn of helices α5/α6. In monomer A Arg111 is rotated inwards, in monomer B outwards. Genistein is sandwiched by both helices α9 (red) of the FrrA dimer. The figure was produced using PYMOL. (B) Stereo view (wall eyed) showing the superposed FrrA homodimers AB (yellow and green) and CD (brown and cyan) as Cα-traces. Both regions that appear to be mobile, the DNA-binding domain (DBD) and helix α6, have been labelled. The figure was produced using PYMOL.

20.6, RMSD on Cα atoms 2.7 Å), QacR [18] (PDB ID: 3BQZ, Z-score 15.4, RMSD on Cα atoms 3.6 Å) and CmeR [19] (PDB ID: 2QCO, Z-score 14.8, RMSD on Cα atoms 3.1 Å). The only difference in the polypeptide fold is found for the C-terminal extension of helix α4. In FrrA, this helix is two turns shorter resulting in an extended conformation for residues 83 – 89.

Asymmetry of the FrrA dimers in the crystal reveals mobile regions

The secondary structure and global fold of the four FrrA monomers in the crystallographic asymmetric unit (A, B, C and D) are almost identical. However, chain-inverted superposition of the two independent homodimers onto themselves (AB onto BA and CD onto DC, yielding RMSD values for all Cα atoms of 1.4 Å and 1.9 Å, respectively) reveals that both dimers are remarkably asymmetric. In each case, the asymmetry is mainly brought about by a pendulum-like movement of the DNA-binding domain (α1 – α4), accompanied by a shift of helix α6 (residues 113 – 122) within the ligand-binding domain (Fig. 2B). In sharp contrast, the remainder of each FrrA dimer (residues 84 – 110 of helix α5 and residues 128 – 214 of helices α7 – α10) is nearly symmetrical, with chain-inverted superposition yielding RMSD values of 0.51 Å (dimer AB) and 0.57 Å (dimer CD) on Cα atoms. Moreover, this part of the structure is virtually identical in both dimers (superposition can be achieved with RMSD of 0.40 Å on Cα atoms). Taken together, these observations suggest that most of the ligand-binding regulatory domain of FrrA forms a rigid dimeric core, while the DNA-binding domain and helix α6 have a certain freedom to move with respect to it. In the context of the crystal, this allows the DNA-binding domain to adapt its orientation to the lattice environment.

An eye-catching difference of the broken dimer symmetry is located at residues 111 – 113 which form a short connection between helices α5 and α6. The superposition of the monomers within the dimers (A on B and C on D) revealed a difference of about 5 Å of the corresponding Arg111Cα positions. In monomers A and C, the side chain of Arg111 points into the dimer core, whereas in monomers B and D the side chain of Arg111 is directly opposed and points outwards (Fig. 2A). In both conformations of Arg111 direct interaction with the ligand is impossible.

FrrA specifically recognizes isoflavonoids

Unlike many other proteins of the TetR family, FrrA has only one ligand-binding pocket, which is located at the dimer interface of the regulatory domain. This is in contrast to TetR, which has two separated symmetry-related ligand-binding sites. Other members of the TetR-family are also induced by a single ligand molecule, for example QacR [20]. The observed ligand binding underlines the asymmetry of the FrrA dimers, because only one genistein or naringenin is observed close to the two-fold axis of the homodimer. Due to the size of the pocket and the asymmetric positioning of the ligand, an unoccupied volume (~ 20%) would allow an alternative orientation of the ligand with respect to the two-fold axis, but was not observed.

Genistein and naringenin are completely buried in the hydrophobic binding-pocket and the major protein/ligand contacts can be described as van der Waals interaction. Between the ligands and the protein surface in the binding pocket is no space for water molecules and we assume a significant entropic term for the binding constant.

Both ligands have several oxygen substituents suitable as hydrogen bond acceptors and donors. Surprisingly, only a single hydrogen bond of the phenolic hydroxyl (OH7) of the chromen-4-one core and the carboxylate of Asp113 (B and D monomer, respectively) is observed at a distance of 3.3 – 3.5 Å (Fig. 3). With naringenin, this interaction is only observed in the AB dimer.

Several amino acid side chains with aromatic or alkyl C–H point towards the π-systems of the ligands at distances of 3.6 – 3.8 Å (Fig. 3A). This kind of hydrogen bond with a C–H moiety as donor and a π-system as acceptor, has been described frequently [21–23]. In contrast, more polar interactions are observed for enzymes which modify flavonoids [24–26].

Superposition of the dimers AB and CD shows a virtually identical genistein binding mode in both dimers. The ligand with its chromen-4-one core is closely positioned to the B and D monomer of the respective dimer. In these monomers, the Arg111 side chain points outwards, away from the ligand, and is located in a pocket formed by the C-terminal residues of helix α4 (carboxylate Asp83, Ser79Oγ, carboxylate Glu76). In the other monomers (A and C), the side chain of Arg111 is oriented towards the C termini of helices α6 and α9, both of the complementary monomers (B and D).

For naringenin, in one homodimer the chroman-4-one body of naringenin binds in the same orientation as is observed for chromen-4-one of genistein. However, the phenyl group points into another direction (~ 60 °) due to the different substitution patterns (Fig. 3B). In the other dimer, the long axis of naringenin is tilted by about ~ 60 ° in such a way that the three rings roughly cover the space otherwise occupied by genistein (Fig 3C). Also, the conformations of the ligands differ with respect to the plane-to-plane orientation of the phenyl substituent to the core. In genistein, the phenyl ring is rotated by about ~ 25 °, whereas in naringenin the phenyl is out of plane by ~ 60 °. In all observed orientations of the ligands, the phenyl substituent is surrounded by a hydrophobic environment.

The side chain of Arg176 (α8, C-terminal) of monomer B is close to the hydroxyl OH5 and carbonyl O6 of the ligand, but the guanidinium hydrogen bonds to the phenolic Tyr104 hydroxyl and Gly177 carbonyl of the other monomer, indicating a more polar interaction. The hydroxyl OH5 faces the guanidinium π-system of Arg176 (Nε, 3.2 Å). This interaction is

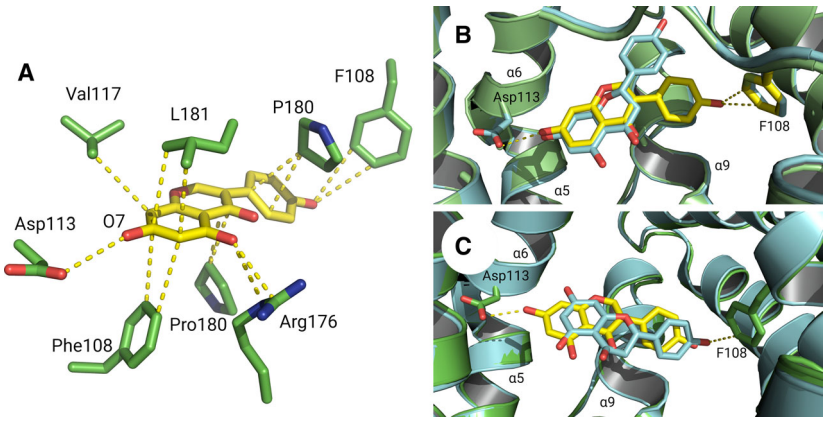

Fig. 3. (A) Genistein recognition by amino acid side chains of monomer A (one-letter code) and monomer B (three-letter code). The C-H hydrogen bonds with the π-system are shown by dotted lines in the range of 3.6 – 3.8 Å. The carboxylate of Asp113 is hydrogen bonded to genistein hydroxyl O7 (3.3 Å). Oxygen and nitrogen atoms are shown in blue and red, respectively. Hydrogen atoms are omitted for clarity. The C-H···π hydrogen bond of Pro180 side chain (CH_2 of Cγ) to the aromatic C3 of genistein is not possible with naringenin, because it has a methylene group in the corresponding position. (B) Ligand-binding pocket with superposition of the genistein (yellow) or (*S*)-naringenin (grey blue) bound in the AB dimer of FrrA. (C) Superposition of the CD dimer with bound genistein or (*R*)-naringenin. Note that in both dimers genistein is in the same orientation. The hydrogen bond of the carboxylate of Asp113 (monomer A, C) to genistein and naringenin hydroxyl O7 (3.3 Å) and the C-H interactions of F108 (monomer B, D) with the phenolic hydroxyl of the B ring (3.6 – 3.8 Å) are shown by dotted lines. Other interactions are omitted for clarity. Genistein and naringenin are sandwiched by helices α9 of both monomers, but only helix α9 of monomer A is shown for clarity. The figures were produced using PYMOL.

apparently less important, because this hydroxyl is missing in another inducing isoflavonoid (daidzein). The Arg176 of monomer A is without ligand contact.

Both monomers sandwich the ligand with Phe108 (C-terminal of α5) and Pro180 (N-terminal of α9). Monomer A interacts with the phenyl substituent, whereas the corresponding residues of monomer B contact the chromen-4-one core. Also, one of the methyl groups of Val117 (monomer B) and both methyl groups and Cβ of Leu181 (monomer A) are at distances of 3.7 Å and 3.8 Å, respectively. The side chains of Leu181 of monomer B and Val117 (of monomer A) are at least 4 Å apart of the ligand.

These interactions differ only for the phenyl substituent of naringenin with its flavone core compared to the isoflavone core of genistein, daidzein and coumestrol. These isoflavonoids were characterized as inducers, but not naringenin [9]. Obviously, the isoflavone framework is essential for induction.

It should be mentioned that at the observed resolution limit of the naringenin complex it is not possible to discriminate between the enantiomers of naringenin. (*S*)-naringenin fits best to the electron density maps of the AB dimer, whereas (*R*)-naringenin was modelled in the CD dimer. In the observed naringenin conformation, a superposition of the enantiomers differs only by about 1 Å of the asymmetric C-atom, due to the puckering of the C-ring [27].

The structure of FrrA in solution

Small-angle X-ray scattering (SAXS) measurements for FrrA in the absence of inducers are shown in Fig. 4A. Comparison of the extrapolated forward scattering (I_0) to that of a bovine serum albumin (BSA) reference suggests a particle mass of 42 kDa. This corresponds to approximately twice the mass of FrrA (23.8 kDa), indicating that the repressor exists as a homodimer in solution. This is consistent with data obtained by blue native acrylamide gel electrophoresis [9].

Surprisingly, neither of the homodimers in the asymmetric unit of the FrrA crystal can explain our SAXS data to within experimental error. Analysis using CRYSOL [28] and either the AB or the CD dimer as a model yields a χ^2-value of 4.9, indicative of a considerable lack of fit. We argued that this could be related to the marked asymmetry of the homodimers, which is most likely caused by crystal packing forces. To find a model that more accurately reflects the situation in solution, we generated symmetric structures by superimposing two copies of a single chain onto the invariant dimer core. However, none of these four symmetrised models (AA, BB, CC and DD) satisfactorily explained the experimental SAXS data either, yielding χ^2-values ranging from 4.4 (CC) to 4.9 (DD). We next used the simulated annealing approach implemented in CORAL [20] to optimize the fit to the experimental data by varying the orientation of the DNA-binding domain as well as the position of helix α6. Resulting models from several independent CORAL runs explain the data almost perfectly, with χ^2-values near unity. Compared to the crystallographic dimers, the optimized models invariably show a rotation of the DNA-binding domain away from the core and a shift of helix α6 towards the resulting gap (a typical example is shown in Fig. 4B). Thus, in solution FrrA appears to be considerably less compact than in the crystal lattice.

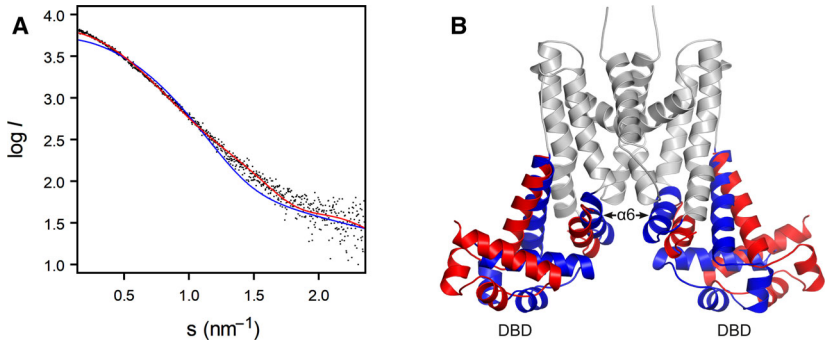

Fig. 4. (A) Logarithmic plot of the experimental SAXS data for ligand-free FrrA, shown as black dots. The blue curve shows the theoretical scattering intensity for one of the homodimers in the crystal structure (chains A and B), computed using CRYSOL (χ^2 = 4.9). The red curve corresponds to the theoretical scattering for the optimized model shown in panel b (χ^2 = 1.2). (B) Comparison of the structure of the AB homodimer in the crystal and an optimized model produced by CORAL [20]. The invariable core of the homodimer is shown in grey. The DNA-binding domain and helix α6, whose positions were optimized against the experimental data, are shown in blue (crystallographic structure) and red (SAXS, optimized model). For clarity, dummy residues in the CORAL model are not shown. The figure was produced using PYMOL.

SAXS data recorded in the presence of saturating amounts of either naringenin or genistein are indistinguishable from those obtained for the protein alone, with correlation map scores [29] of 10 ($P = 0.80$) and 11 ($P = 0.54$), respectively. These observations are consistent with our crystallographic analysis of the two FrrA-ligand complexes and confirm that neither inducer causes major structural changes in the repressor.

FrrA promoter binding and induction by isoflavonoids

The absence of detectable conformational changes in FrrA upon ligand binding suggests that inducers do not alter the free state of the repressor, but act instead by interfering with a structural transition required for DNA binding. Since no structures of FrrA bound to its target site are as yet available, we used DALI [17] to search the PDB for closely related folds. The most similar repressor whose complex with DNA has been structurally characterized was found to be QacR, a TetR-like regulator of multidrug resistance from *Staphylococcus aureus* [30]. At the sequence level, the two proteins exhibit 18% amino acid identity, while the dequalinium-bound (induced) form of QacR, PDB ID: 1JT6 [31], can be superposed onto FrrA with RMSD of 3.2 Å on Cα atoms. Figure 5A,B show a superposition of FrrA onto promoter-bound QacR, PDB ID: 1JT0 [32]. As this comparison demonstrates, a pendulum-like movement of the DNA-binding domains of the FrrA dimer towards each other is required in order to fit the two instances of helix α3, the recognition helix of the HTH motif, into consecutive positions of the major groove on the same face of the double helix. This DNA-binding mode, which is highly conserved in all TetR-like repressors, requires a specific distance between the HTH motifs as well as a precise orientation of the DNA-binding domains in order to align both α3 helical axes with the major groove of the nucleic acid [13,14,33]. Interestingly, the squeezing movement of the DNA-binding domains of the FrrA dimer towards their DNA-binding positions would inevitably push both α6 helices into the inducer-binding pocket, but the inducing flavonoid prop the domains in an open conformation (Fig. 5C). This observation readily explains why DNA binding and inducer binding are mutually exclusive despite the apparent absence of flavonoid-induced structural changes to the free repressor. Consistent with this movement of helices α6 the inducer-binding pockets of both FrrA complexes (~ 2650 Å^3) decrease by about 21% in apo-FrrA (~ 2100 Å^3).

According to the corresponding short operator sequences of about 15 base pairs, it is not expected that FrrA binds as a dimer of dimers like the QacR/DNA complex [32].

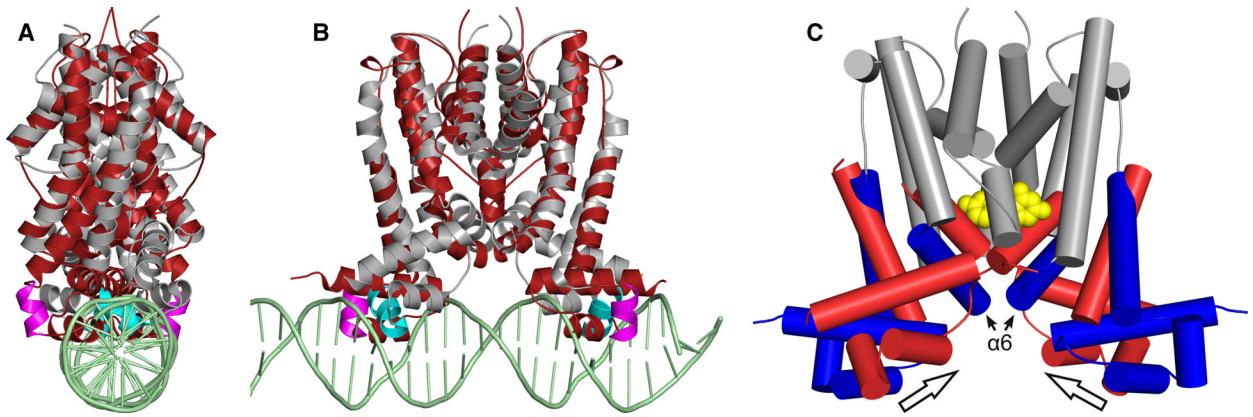

Fig. 5. Model for FrrA DNA-binding and induction by isoflavonoids. (A) Superposition of FrrA (dimer AB, red) onto the QacR-DNA complex, viewed down the helical axis of the nucleic acid (PDB ID: 1JT0, QacR dimer AC shown in grey, DNA in green). A pendulum-like, inward rotation of the FrrA DNA-binding domain is required to position helix α3 (blue) inside the major groove, allowing sequence-specific interaction with the base pairs. (B) The same superposition, view perpendicular to the helical axis of the DNA. (C)Model for induction of FrrA by isoflavonoids. The inward movement of the DNA-binding domains towards each other (arrows), is essential for interaction with the nucleic acid as shown in panels A and B. The bound inducer props helix α6 of both monomers and prevents conformational changes required for DNA binding. The rigid core of the FrrA dimer is shown in grey, the DNA-binding domain and helix α6 in blue. A model of DNA-bound FrrA, shown in red, was generated by independently superposing each DNA-binding domain, together with helix α6, onto the QacR-DNA structure (PDB ID: 1JT0). Genistein is represented as a surface model (yellow). The figures were produced using PYMOL.

Discussion

A systematic overview of the folding pattern and sequence conservation of the TetR family is provided in earlier work [3,4,15]. In contrast to the strong conservation of the DNA-binding domain, sequence identity within the regulatory domain of the TetR family is weak. This variability reflects the requirements of signal-molecule specificity, determined by differently shaped pockets as well as numerous selective side chain contacts.

It is instructive to compare the ligand-binding pocket of FrrA to those of similar repressors that were identified in our DALI search, such as QacR [31] and CmeR [34]. These proteins share sequence identities with FrrA of 18% and 21%, respectively. Although their overall fold is highly similar to that of FrrA and, consequently, the ligand-binding pockets are in roughly the same position, major differences are apparent. The QacR dimer binds one molecule of various cationic lipophilic drugs in a pocket which overlaps with the FrrA binding cavity. The central cavity of CmeR is separated into two distinct binding sites, which are specific for bile acids [34]. In contrast to QacR and FrrA, CmeR binds two inducer molecules symmetrically with respect to the homodimer axis, which is common for the TetR family. These observations indicate that TetR-like repressors can accommodate a wide variety of inducers through evolutionary adaptation of their central ligand-binding cavity. As was proposed for TetR [35], changes in binding specificity would be expected to leave the allosteric mechanism intact, enabling rapid incorporation of the repressors into nascent regulatory pathways.

The induction model that we propose for FrrA involves a pendulum-like motion of the DNA-binding domain together with helix α4 and helix α6 upon ligand binding. A similar movement has been postulated for QacR [31]. In several crystal structures of QacR-ligand complexes, the asymmetric unit comprises a remarkably asymmetric ligand-bound QacR dimer, but also a symmetrical dimer of apo-QacR (e.g. PDB ID: 1jt6) [31]. In relation to the ligand-free QacR, the asymmetric FrrA conformation is favoured in apo-FrrA by crystal packing contacts. Although it seems tempting to speculate that the asymmetric FrrA dimer in our crystals corresponds to the induced state, we do not have any evidence to that effect. Indeed, the asymmetric dimer structure was found to be incompatible with our SAXS data, irrespective of the presence of ligands.

Taken together, our crystallographic structures and SAXS measurements in solution demonstrate that the DNA-binding domain (helices α1 – α4) and helix α6 of FrrA have a considerable freedom to move with respect to the ligand-binding core of the repressor dimer. A comparable situation has been observed for TetR [33,36]. Interestingly, neither of our experimental approaches have revealed any structural differences between ligand-free FrrA and complexes with flavonoids. This suggests that inducers do not function by forcing a change in the conformation of free FrrA, but, on the contrary, prevent a structural transition required for promoter binding. Based on the known DNA-binding mode of other TetR-like repressors and modelling of the DNA-bound form of FrrA, we propose a model of induction by isoflavonoids in which the inducer, acting as a wedge, directly obstructs an inward rotation of the DNA-binding domains that is critical for their interaction with the closely spaced half-sites of the palindromic operator. This remarkably simple allosteric mechanism, which is likely to play a role in many other TetR-type repressors, is particularly well-suited for regulatory circuits requiring a broad inducer specificity, as it hardly places any constraints on the precise chemical nature of the ligand. Given that their specificity can be altered without affecting the allosteric mechanism, TetR-type repressors are highly attractive as molecular switches in biotechnological applications and synthetic biology [37,38].

Material and methods

Recombinant protein expression and purification

A synthetic *frrA* gene, codon-optimized for expression in *Escherichia coli*, was purchased from Eurofins Genomics (Ebersberg, Germany). The gene was integrated into pET-28a, which adds a His_6-tag followed by a TEV-protease cleavage site (MGSSHHHHHHSSENLYFQ↓GH) to the N-terminal end of the protein. FrrA was produced by overexpression in *E. coli* strain BL21 (DE3). To solve the crystallographic phase problem, the methionines of FrrA were substituted by L-selenomethionine (SeMet) [39]. A preculture was grown in LB medium with kanamycin (final concentration 50 $\mu g \cdot mL^{-1}$) and 1% glucose. The preculture was centrifuged at 2150 × *g* and 277 K for 20 min and resuspended in M9 medium. The main culture was grown at 310 K in M9 medium supplemented with kanamycin (final concentration 50 $\mu g \cdot mL^{-1}$). At an OD_{600} of 0.8 the culture was supplemented with 20 mL L-SeMet (1 g l^{-1}) and 5 mL of an amino acid mixture (L-Lys, L-Phe, L-Thr, D, L-Ile, D,L-Val each 100 mM, 50 mM L-Leu) [39]. The protein expression was induced 15 min later by the addition of IPTG to an end concentration of 1 mM. After 3 h of incubation at 310 K, the cells were harvested by centrifugation at 10 000 × ***g*** and 277 K for 20 min. The pellets were resuspended in lysis buffer (20 mM Na_2HPO_4, pH 7.5,

250 mM NaCl, 5 mM imidazole). The cells were disrupted by 15 × 1 min sonification and centrifuged at 40 000 × ***g*** and 277 K for 45 min. The supernatant containing the His_6-tagged protein was injected onto POROS 20 MC metal chelate affinity resin (Applied Biosystems, Weiterstadt, Germany) loaded with Ni^{2+} ions. The column was washed with lysis buffer and the protein eluted with elution buffer (20 mM Na_2HPO_4, pH 7.5, 250 mM NaCl, 500 mM imidazole). Protein-containing fractions were pooled and the His_6-tag was cleaved off with TEV protease at 291 K, leaving the N-terminal sequence GH followed by the target sequence. The protein was then loaded again onto the affinity column to remove the His_6 peptide, the His_6-tagged TEV protease as well as any contaminants that had previously coeluted. The flow-through contained pure FrrA, which was dialysed against 20 mM Na_2HPO_4, pH 7.5, 50 mM NaCl and 50 mM imidazole. The purified protein appeared as a single band on Coomassie-stained SDS/PAGE gels.

The SeMet variant of FrrA was used throughout for all experimental approaches and will be referred to in the following as FrrA.

Small-angle X-ray scattering

Small-angle X-ray scattering (SAXS) data were collected at beamline P12 of the EMBL Outstation at PETRA III, DESY, Hamburg [40], using a PILATUS 2 M pixel detector, a sample-to-detector distance of 3.1 m and a wavelength of 1.24 Å. Measurements covered the momentum transfer range 0.008–0.47 $Å^{-1}$ ($s = 4\pi \cdot \sin(\theta) / \lambda$, where 2θ is the scattering angle and λ is the X-ray wavelength). All experiments took place at a temperature of 283 K. A capillary flow cell was used to obtain 20 successive 50 ms exposures for each sample, revealing no significant change in the course of any one experiment. Several FrrA dilutions (0.94–8.3 $mg \cdot mL^{-1}$) were analysed in a buffer containing 20 mM Na_2HPO_4, pH 7.5 and 100 mM NaCl. All data were normalized to the intensity of the transmitted beam and radially averaged. Scattering of the buffer was subtracted and the resulting curves were scaled to unity protein concentration (1 $mg \cdot mL^{-1}$). For further data analysis, version 3.0.3 of the ATSAS package was used [41]. Theoretical scattering patterns based on high-resolution structural models were fitted to the experimental data using the program CRYSOL, which accurately accounts for protein surface hydration [28]. The program CORAL [20] was used for SAXS-based modelling of the domain positioning within FrrA as well as for adding dummy residues to represent a short N-terminal segment missing in the crystal structure. In all CORAL calculations, twofold symmetry of the homodimer was imposed. Both, the DNA-binding domain (residues 20–83) and the region comprising helix α6 (residues 111–127) were allowed to move as rigid bodies, while the distance between residues 83 and 84 was restrained to 3 Å to favour pendulum-like movements of the former.

Circular dichroism spectroscopy

Circular dichroism (CD) was measured with a J-810 CD spectrometer (JASCO, Tokyo, Japan) in a wavelength range of 185–270 nm and corrected for buffer contributions. The protein sample contained 50 $\mu g \cdot mL^{-1}$ FrrA and 20 mM Na_2HPO_4 (pH 7.5).

Protein crystallization

At concentrations suitable for crystallization SeMet-labelled FrrA is stable in buffers with an imidazole concentration of at least 50 mM. SeMet-labelled FrrA was crystallized by means of the hanging drop vapour diffusion technique at 291 K. For the hanging drops, 1 µL protein solution (6 $mg \cdot mL^{-1}$ FrrA in 20 mM Na_2HPO_4, 100 mM imidazole, pH 7.5, 50 mM NaCl) was mixed with 1 µL reservoir solution (20% PEG8000, and 0.1 M 2-(*N*-cyclohexyl amino)ethane sulfonic acid (CHES), pH 9.5) and equilibrated against 500 µL reservoir solution.

FrrA was co-crystallized with equimolar ratio of genistein and (*R*,*S*)-naringenin. Both ligands were dissolved in 70% ethanol. The racemate of naringenin was used in view of spontaneous racemisation of flavanones reported in earlier work [42]. The crystallization conditions were the same as for apo-FrrA.

The crystals appeared after 2–3 days, were cryoprotected with 16% PEG8000, 20% PEG400 and 0.16 M CHES, pH 9.5 and flash-frozen in liquid nitrogen prior to X-ray data collection.

X-ray diffraction data collection

Multiple-wavelength anomalous diffraction (MAD) data sets of SeMet-labelled FrrA crystals were collected at wavelengths near the selenium absorption K-edge (12.6578 keV, λ = 0.9795 Å). Data of peak and inflection points were collected at 100 K at beamline BL14.1, operated by the Joint Berlin MX-Laboratory at the BESSY II electron storage ring (Berlin Adlershof, Germany) [43] and processed with XDSAPP [44,45]. Diffraction data of FrrA in complex with genistein and naringenin were collected at 100 K at beamline P13 [46], operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany). Diffraction data were processed with XDS [47]. Details are listed in Table 1.

Structure determination and refinement

Initial phasing of SeMet-labelled FrrA was based on multiple anomalous diffraction of peak and inflection datasets (2WL-MAD) collected at the absorption K-edge of Se [25].

Table 1. Data collection statistics. Values in parentheses correspond to the highest resolution shell.

Protein, complex	FrrA		FrrA-genistein	FrrA-naringenin
PDB ID	6G87		6G8G	6G8H
radiation source, beamline	BESSY II, BL 14.1		PETRA III, BL 13, EMBL c/o DESY	
Detector	Pilatus 6 M		Pilatus 6 M	
Temperature (K)	100	100	100	100
Wavelength (Å)	0.9796 (inflection)	0.9790 (peak)	0.9763	0.9763
Resolution range (highest shell) (Å)	47.576 – 3.08 (3.26 – 3.08)	47.57 – 2.92 (3.10 – 2.92)	119.1 – 2.60 (2.76 – 2.60)	118.71 – 2.60 (2.67 – 2.60)
Unit cell parameters a = b; c (Å)	119.64; 78.49	119.59; 78.48	119.04; 78.31	118.71; 78.35
Space group	$P4_3$		$P4_3$	$P4_3$
Rotation range (°) / increment (°)	360 / 0.1	360 / 0.1	180 / 0.1	180 / 0.1
measured reflections	282840 (44671)	330097 (53675)	225481 (35078)	223019 (14858)
unique reflections*	40276 (6469)*	47028 (7487)*	65566 (10480)*	65529 (4795)*
Redundancy	7.02 (6.91)	7.02 (7.17)	3.44 (3.35)	3.40 (3.10)
Mean $I/\sigma(I)$	13.26 (1.95)	15.17 (1.96)	11.52 (0.82)	9.55 (0.74)
Completeness (%)	99.7 (99.5)	99.6 (99.0)	99.1 (97.7)	99.5 (98.4)
R_{meas} (%)	12.6 (101)	10.9 (100)	9.1 (152)	9.3 (188)
Wilson *B*-factor (Å^2)	73.1	70.5	76.3	79.8
CC 1/2	0.998 (0.716)	0.999 (0.704)	0.998 (0.279)	0.998 (0.183)

*Friedel pairs are treated as different reflections.

We used the Auto-Rickshaw server [48,49] of the EMBL/Hamburg (http://www.embl-hamburg.de/Auto-Rickshaw/), which utilizes programs of the CCP4 suite [50], SHELX-C/D/E [51–53], DM [54] and ARP/wARP [55,56].

The best solution was identified for space group $P4_3$ with twelve Se positions with occupancies in the range of 0.87 to 0.40 above a noise threshold of 0.27. According to the self-rotation function, this is in agreement with three SeMet in the FrrA sequence assuming four monomers of FrrA in the asymmetric unit. The same procedure for the alternative enantiomorph space group $P4_1$ revealed only nine heavy atoms with occupancies 0.41 to 0.29 above a threshold of 0.21.

After density modification and applying noncrystallographic symmetry (NCS), the electron density maps showed long helical motifs. Tracing by ARP/wARP [55,56] provided 38 polypeptide fragments with 503 residues belonging to four protein chains (A, B, C, D) with 196 residues each. These fragments were suitable for complete model building with COOT [57] and subsequent refinement with REFMAC5 [58]. The numbering of the polypeptide refers to the construct after TEV-protease cleavage of the expression tag. Thus, the used FrrA (amino acid residues 1 – 214) has two additional residues N-terminal compared to the wild-type sequence, verified by mass spectrometry. The crystallized FrrA comprises amino acid residues 19 – 214.

Both FrrA complexes and apo-FrrA crystallized isomorphously in the tetragonal space group $P4_3$. Phasing of diffraction data of FrrA in complex with genistein and naringenin was successful by molecular replacement with

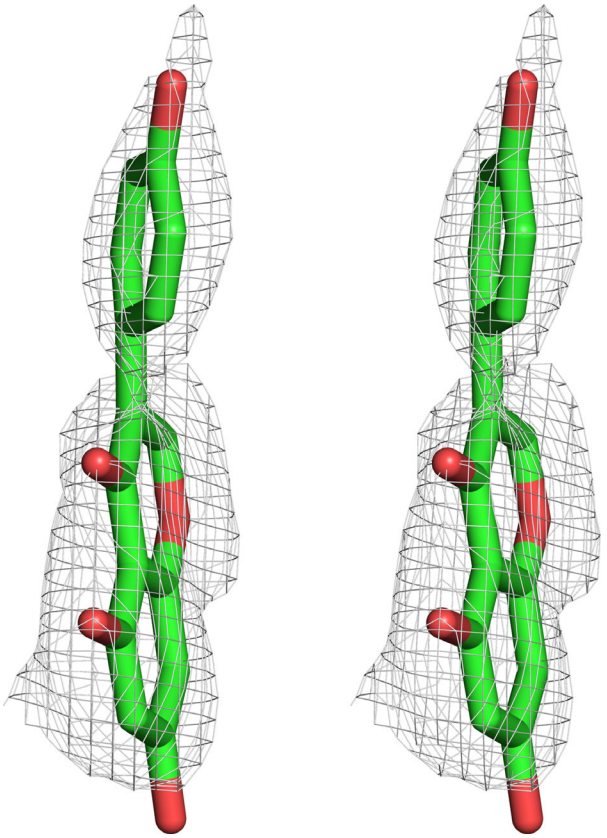

Fig. 6. Wall eyed stereo view of genistein in the binding site of the AB-dimer with mF_o - DF_c electron density omit-map contoured at 2.5 σ level. The figure was produced using PYMOL.

Table 2. Refinement statistics. Values in parentheses correspond to the highest resolution shell.

Protein, complex	FrrA	FrrA-genistein	FrrA-naringenin
PDB ID	6G87	6G8G	6G8H
Resolution range (Å)	47.56 – 2.92 (3.00 – 2.92)	119.04 – 2.60 (2.67 – 2.60)	118.71 – 2.60 (2.67 – 2.60)
No. of reflections	24189 (1749)	32099 (2426)	32100 (2367)
R_{cryst}	0.183	0.188	0.183
R_{free} / % of data	0.211 / 4.5	0.216 / 5.1	0.219 / 4.8
Figure of Merit	0.84	0.79	0.80
RMSD bond lengths (Å)	0.012	0.013	0.011
RMSD bond angles (°)	1.55	1.68	1.79
Ramachandran parameters favoured / allowed (%)	98 / 2	99 / 1	97.3 / 2.7
Average *B*-factors (Å^2) main chain/side chain/ligand	75.2/82.3/ –	70.6/80.3/97.0	76.0/84.5/102.8

PHASER [59] using an apo-FrrA dimer (chains AB) as search model. Again, electron density map inspection and final model corrections were done with COOT [57], followed by refinement with REFMAC5 [58]. In the course of refinement, the ligands were identified in different electron density maps (Fig. 6) and subsequently included for final refinement. Some molecules of the buffering agent (CHES) of the crystallization setup were identified in all structures, always in the neighbourhood of positively charged amino acid residues. In the refinement of all structures, NCS restraints were applied to all four monomers of FrrA in the asymmetric unit. Further details are listed in Table 2.

Figures showing molecular models were prepared using PYMOL [60]. The volume of inducer-binding pockets was calculated with HotSpot Wizard v3.1 [61].

Acknowledgements

The beamline staff of BESSY/Berlin and the EMBL outstation at DESY/Hamburg are gratefully acknowledged for their kind assistance during data collection.

Conflict of interest

The authors declare no conflict of interest.

Peer Review

The peer review history for this article is available at https://publons.com/publon/10.1111/febs.16141.

Data accessibility

Structure factors and final coordinates of apo-FrrA, the complexes of FrrA with genistein and FrrA with naringenin have been deposited in the Protein Data Bank under accession codes 6G87, 6G8G and 6G8H, respectively. The SAXS data and models of FrrA have been deposited in the Small-Angle Scattering Biological Data Bank (SASBDB http://www.sasbdb.org) with entries SASDKA6 (FrrA), SASDKB6 (FrrA with genistein) and SASDKC6 (FrrA with naringenin).

Author contributions

NW protein purification, crystallization, crystallography, structure refinement; SW evaluation and interpretation of SAXS data, modelling and analysis of the induction mechanism, manuscript writing; JH cloning, protein expression, X-ray data collection; GJP X-ray data collection; MG cloning, initiated the study; WH conceived and initiated the study, final crystallographic refinement, manuscript writing.

References

1 Poole P, Ramachandran V & Terpolilli J (2018) Rhizobia: from saprophytes to endosymbionts. *Nat Rev Microbiol* **16**, 291–303.
2 Masson-Boivin C & Sachs JL (2018) Symbiotic nitrogen fixation by rhizobia - the roots of a success story. *Curr Opin Plant Biol* **44**, 7–15.
3 Cuthbertson L & Nodwell JR (2013) The TetR family of regulators. *Microbiol Mol Biol Rev* **77**, 440–475.
4 Ramos JL, Martinez-Bueno M, Molina-Henares AJ, Teran W, Watanabe K, Zhang XD, Gallegos MT, Brennan R & Tobes R (2005) The TetR family of transcriptional repressors. *Microbiol Mol Biol Rev* **69**, 326–356.
5 Xi CW, Schoeters E, Vanderleyden J & Michiels J (2000) Symbiosis-specific expression of *Rhizobium etli casA* encoding a secreted calmodulin-related protein. *Proc Natl Acad Sci USA* **97**, 11114–11119.
6 Rossbach S, Kunze K, Albert S, Zehner S & Göttfert M (2014) The Sinorhizobium meliloti EmrAB efflux system is regulated by flavonoids through a TetR-Like

regulator (EmrR). *Mol Plant-Microbe Interact* **27**, 379–387.
7 Santos MR, Marques AT, Becker JD & Moreira LM (2014) The *Sinorhizobium meliloti* EmrR regulator is required for efficient colonization of medicago sativa root nodules. *Mol Plant-Microbe Interact* **27**, 388–399.
8 Han F, He X, Chen W, Gai H, Bai X, He Y, Takeshima K, Ohwada T, Wei M & Xie F (2020) Involvement of a Novel TetR-Like Regulator (BdtR) of *Bradyrhizobium diazoefficiens* in the efflux of isoflavonoid genistein. *Mol Plant-Microbe Interact* **33**, 1411–1423.
9 Wenzel M, Lang K, Günther T, Bhandari A, Weiss A, Lulchev P, Szentgyörgyi E, Kranzusch B & Göttfert M (2012) Characterization of the Flavonoid-responsive regulator FrrA and its binding sites. *J Bacteriol* **194**, 2363–2370.
10 Kosslak RM, Bookland R, Barkei J, Paaren HE & Appelbaum ER (1987) Induction of Bradyrhizobium japonicum common nod genes by isoflavones isolated from Glycine max. *Proc Natl Acad Sci USA* **84**, 7428–7432.
11 Makarova LE, Dudareva LV, Petrova IG & Vasil'eva GG (2016) Secretion of phenolic compounds into root exudates of pea seedlings upon Inoculation with *Rhizobium leguminosarum* bv. *viceae* or *Pseudomonas siringae* pv. *pisi*. *Appl Biochem Microbiol* **52**, 205–209.
12 Deng W, Li C & Xie J (2013) The underling mechanism of bacterial TetR/AcrR family transcriptional repressors. *Cell Signal* **25**, 1608–1613.
13 Brennan RG & Matthews BW (1989) The helix-turn-helix DNA-binding motif. *J Biol Chem* **264**, 1903–1906.
14 Harrison SC (1991) A structural taxonomy of DNA-binding domains. *Nature* **353**, 715–719.
15 Yu Z, Reichheld SE, Savchenko A, Parkinson J & Davidson AR (2010) A comprehensive analysis of structural and sequence conservation in the TetR family transcriptional regulators. *J Mol Biol* **400**, 847–864.
16 Holm L & Sander C (1995) DALI - a network tool for protein-structure comparison. *Trends Biochem Sci* **20**, 478–480.
17 Holm L (2020) DALI and the persistence of protein shape. *PROTEIN Sci* **29**, 128–140.
18 Peters KM, Schuman JT, Skurray RA, Brown MH, Brennan RG & Schumacher MA (2008) QacR−cation recognition is mediated by a redundancy of residues capable of charge neutralization. *Biochemistry* **47**, 8122–8129.
19 Gu R, Su C-C, Shi F, Li M, McDermott G, Zhang Q & Yu EW (2007) Crystal structure of the transcriptional regulator CmeR from *Campylobacter jejuni*. *J Mol Biol* **372**, 583–593.
20 Petoukhov MV, Franke D, Shkumatov AV, Tria G, Kikhney AG, Gajda M, Gorba C, Mertens HDT, Konarev PV & Svergun DI (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Crystallogr* **45**, 342–350.
21 Newberry RW & Raines RT (2019) Secondary forces in protein folding. *ACS Chem Biol* **14**, 1677–1686.
22 Makuvaza JT, Kokkin DL, Loman JL & Reid SA (2019) C-H/pi and C-H-O interactions in concert: a study of the anisole-methane complex using resonant ionization and velocity mapped ion imaging. *J Phys Chem A* **123**, 2874–2880.
23 Brandl M, Weiss MS, Jabs A, Suhnel J & Hilgenfeld R (2001) C-H···pi-interactions in proteins. *J Mol Biol* **307**, 357–377.
24 Jez JM, Bowman ME, Dixon RA & Noel JP (2000) Structure and mechanism of the evolutionarily unique plant enzyme chalcone isomerase. *Nat Struct Biol* **7**, 786–791.
25 Ngaki MN, Louie GV, Philippe RN, Manning G, Pojer F, Bowman ME, Li L, Larsen E, Wurtele ES & Noel JP (2012) Evolution of the chalcone-isomerase fold from fatty-acid binding to stereospecific catalysis. *Nature* **485**, 530–U147.
26 Thomsen M, Tuukkanen A, Dickerhoff J, Palm GJ, Kratzat H, Svergun DI, Weisz K, Bornscheuer UT & Hinrichs W (2015) Structure and catalytic mechanism of the evolutionarily unique bacterial chalcone isomerase. *Acta Crystallogr Sect D Biol Crystallogr* **71**, 907–917.
27 Shin W & Lah MS (1986) Structure of (R, S)-Naringenin. *Acta Crystallogr Sect C - Cryst Struct Commun* **42**, 626–628.
28 Svergun DI, Baberato C & Koch M (1995) CRYSOL - a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr* **28**, 768–773.
29 Franke D, Jeffries CM & Svergun DI (2015) Correlation Map, a goodness-of-fit test for one-dimensional X-ray scattering spectra. *Nat Methods* **12**, 419–422.
30 Brown MH & Skurray RA (2001) Staphylococcal multidrug efflux protein QacA. *J Mol Microbiol Biotechnol* **3**, 163–170.
31 Schumacher MA, Miller MC, Grkovic S, Brown MH, Skurray RA & Brennan RG (2001) Structural mechanisms of QacR induction and multidrug recognition. *Science* **294**, 2158–2163.
32 Schumacher MA, Miller MC, Grkovic S, Brown MH, Skurray RA & Brennan RG (2002) Structural basis for cooperative DNA binding by two dimers of the multidrug-binding protein QacR. *EMBO J* **21**, 1210–1218.
33 Orth P, Schnappinger D, Hillen W, Saenger W & Hinrichs W (2000) Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system. *Nat Struct Biol* **7**, 215–219.
34 Lei H-T, Shen Z, Surana P, Routh MD, Su C-C, Zhang Q & Yu EW (2011) Crystal structures of CmeR-bile acid complexes from *Campylobacter jejuni*. *Protein Sci* **20**, 712–723.

35 Werten S, Schneider J, Palm GJ & Hinrichs W (2016) Modular organisation of inducer recognition and allostery in the tetracycline repressor. *FEBS J* **283**, 2102–2114.
36 Orth P, Cordes F, Schnappinger D, Hillen W, Saenger W & Hinrichs W (1998) Conformational changes of the Tet repressor induced by tetracycline trapping. *J Mol Biol* **279**, 439–447.
37 Urlinger S, Baron U, Thellmann M, Hasan MT, Bujard H & Hillen W (2000) Exploring the sequence space for tetracycline-dependent transcriptional activators: novel mutations yield expanded range and sensitivity. *Proc Natl Acad Sci USA* **97**, 7963–7968.
38 Schmidt S, Berens C & Klotzsche M. (2014) A novel TetR-regulating peptide turns off rtTA-mediated activation of gene expression. *PLoS One* **9** , e96546.
39 Doublié SBT (1997) Preparation of selenomethionyl proteins for phase determination. Methods in Enzymology – Macromolecular Crystallography Part A, pp. 523–530. Academic Press.
40 Blanchet CE, Spilotros A, Schwemmer F, Graewert MA, Kikhney A, Jeffries CM, Franke D, Mark D, Zengerle R, Cipriani F *et al.* (2015) Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). *J Appl Crystallogr* **48**, 431–443.
41 Franke D, Petoukhov MV, Konarev PV, Panjkovich A, Tuukkanen A, Mertens HDT, Kikhney AG, Hajizadeh NR, Franklin JM, Jeffries CM *et al.* (2017) ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J Appl Crystallogr* **50**, 1212–1225.
42 Mol JNM, Robbins MP, Dixon RA & Veltkamp E (1985) Spontaneous and enzymic rearrangement of naringenin chalcone to flavanone. *Phytochemistry* **24**, 2267–2269.
43 Müller U, Darowski N, Fuchs MR, Foerster R, Hellmig M, Paithankar KS, Puehringer S, Steffien M, Zocher G & Weiss MS (2012) Facilities for macromolecular crystallography at the Helmholtz-Zentrum Berlin. *J Synchrotron Radiat* **19**, 442–449.
44 Sparta KM, Krug M, Heinemann U, Mueller U & Weiss MS (2016) XDSAPP2.0. *J Appl Crystallogr* **49**, 1085–1092.
45 Krug M, Weiss MS, Heinemann U & Mueller U (2012) XDSAPP: a graphical user interface for the convenient processing of diffraction data using XDS. *J Appl Crystallogr* **45**, 568–572.
46 Cianci M, Bourenkov G, Pompidor G, Karpics I, Kallio J, Bento I, Roessle M, Cipriani F, Fiedler S & Schneider TR (2017) P13, the EMBL macromolecular crystallography beamline at the low-emittance PETRA III ring for high- and low-energy phasing with variable beam focusing. *J Synchrotron Radiat* **24**, 323–332.
47 Kabsch W (2010) XDS. *Acta Crystallogr Sect D - Biol Crystallogr* **66**, 125–132.
48 Panjikar S, Parthasarathy V, Lamzin VS, Weiss MS & Tucker PA (2005) Auto-Rickshaw: an automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. *Acta Crystallogr Sect D - Struct Biol* **61**, 449–457.
49 Panjikar S, Parthasarathy V, Lamzin VS, Weiss MS & Tucker PA (2009) On the combination of molecular replacement and single-wavelength anomalous diffraction phasing for automated structure determination. *Acta Crystallogr Sect D - Struct Biol* **65**, 1089–1097.
50 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr Sect D - Struct Biol* **67**, 235–242.
51 Schneider TR & Sheldrick GM (2002) Substructure solution with SHELXD. *Acta Crystallogr Sect D - Struct Biol* **58**, 1772–1779.
52 Sheldrick GM (2010) Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr Sect D - Struct Biol* **66**, 479–485.
53 Thorn A & Sheldrick GM (2013) Extending molecular-replacement solutions with SHELXE. *Acta Crystallogr Sect D - Struct Biol* **69**, 2251–2256.
54 Cowtan K & Main P (1998) Miscellaneous algorithms for density modification. *Acta Crystallogr Sect D - Struct Biol* **54**, 487–493.
55 Perrakis A, Harkiolaki M, Wilson KS & Lamzin VS (2001) ARP/wARP and molecular replacement. *Acta Crystallogr Sect D - Struct Biol* **57**, 1445–1450.
56 Morris RJ, Zwart PH, Cohen S, Fernandez FJ, Kakaris M, Kirillova O, Vonrhein C, Perrakis A & Lamzin VS (2004) Breaking good resolutions with ARP/wARP. *J Synchrotron Radiat* **11**, 56–59.
57 Emsley P & Cowtan K (2004) COOT: model-building tools for molecular graphics. *Acta Crystallogr Sect D - Biol Crystallogr* **60**, 2126–2132.
58 Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F & Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr Sect D - Biol Crystallogr* **67**, 355–367.
59 >McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674.
60 Schrödinger L (2015) The PyMOL molecular graphics system. *Version* **1**, 8.
61 Sumbalova L, Stourac J, Martinek T, Bednar D & Damborsky J (2018) HotSpot Wizard 3.0: web server for automated design of mutations and smart libraries based on sequence input information. *Nucleic Acids Res* **46**, W356–W362.