

# Deciphering Allosterism of an *Escherichia coli* Hexuronate Metabolism Regulator: UxuR

Beatriz C. Almeida<sup>abc</sup>, Sean A. Wirt<sup>d</sup>, Kristala L. J. Prather<sup>d</sup> and Alexandra T. P. Carvalho<sup>abe</sup>

<sup>a</sup> CNC-UC - Center for Neuroscience and Cell Biology, University of Coimbra, Portugal

<sup>b</sup> CIBB - Centre for Innovative Biomedicine and Biotechnology, University of Coimbra, Portugal

<sup>c</sup> University of Coimbra, Institute for Interdisciplinary Research, Doctoral Programme in Experimental Biology and Biomedicine (PDBEB), Portugal

<sup>d</sup> Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

<sup>e</sup> Department of Biocatalysis and Isotope Chemistry, Almac Sciences, Craigavon, Northern Ireland BT63 5QD, U.K.

## Supporting Information

### Index of Content

<b>Methods</b> .....	2
<b>Computational Workflow</b> .....	2
<b>Modelling</b> .....	2
<b>Molecular Docking details</b> .....	2
<b>System Setup</b> .....	3
<b>Conventional Molecular Dynamics details</b> .....	3
<b>Unconstrained Enhanced Molecular Dynamics</b> .....	3
<b>Structural Analyses</b> .....	3
<b>Model validation</b> .....	4
<b>Short- and long-range communication analyses</b> .....	5
<b>Data visualization</b> .....	5
<b>Experimental Workflow</b> .....	5
<b>Microbial Strains</b> .....	5
<b>Sensor Plasmid Construction</b> .....	5
<b>Growth Conditions</b> .....	6
<b>GFP Measurement</b> .....	6
<b>Biosensor validation</b> .....	6
<b>Analyses</b> .....	6
<b>Figure S1-S33</b> .....	7
<b>Table S1-S21</b> .....	43
<b>References</b> .....	51

## Methods

### Computational Workflow

#### Modelling

The UxuR study was performed using the amino acid sequence from *Escherichia coli* str. K-12 substr. MG1655 (GenBank Identifier: AAC77280). The UxuR model used was previously modelled in our last work.<sup>1</sup> As illustrated in Fig. S1, the UxuR model was characterized as a homodimeric protein, in which each monomer is composed of two domains (N-terminal DNA-binding domain and C-terminal effector-binding domain) connected by a disorganized sequence named linker. Different studies have highlighted the role of a linker as preponderant for TFs' function.<sup>2</sup> Due to the level of disorganization of the linker it is difficult to accurately find the right orientation of one domain towards the other (N-terminal domain vs C-terminal domain). Additionally, the few available numbers of FadR C-terminal domain (FCD) structures with DNA make the UxuR modelling even more challenging. Regarding the FadR group, two main conformations can be found: *E. coli* FadR TF (PDB code 1H9T, 3.25 Å resolution), whose structure is characterized by a non-swapped shape, with a kink that turns the linker to the center (Fig. S2 a), and *E. coli* NanR TF (PDB code 6WFQ, 3.90 Å resolution), whose structure is characterized by a non-swapped shape, with linker adopting a parallel orientation with C-terminal domain of the same monomer (Fig. S2 b). To proceed with our work, we compared the structure of the UxuR model with *E. coli* FadR and *E. coli* NanR structures when bonded to DNA and when a ligand is bonded to the C-terminal domain. The *E. coli* FadR crystal structures in complex with DNA or in complex with myristoyl-CoA (PDB code 1H9T, 3.25 Å and PDB code 1H9G, 2.10 Å respectively) maintain the non-swapped shape, with a kink that turns the linker to the center (Fig. S2 a and c). In contrast to the NanR TF bound to DNA (PDB code 6WFQ, 3.90 Å resolution, Fig. S2 b) which linkers are not domain-swapped, in the NanR crystal structure in complex with sialic acid (PDB code 6ON4, 2.10 Å resolution, Fig. S2 d), the linkers cross to form domain-swapped monomers. This would require N-terminal domains to uncross before or upon DNA binding, which according to Horne *et al.*, is plausible given linkers' flexibility.<sup>2</sup> The UxuR model's linkers placement is similar to the NanR structure in complex with sialic acid (Fig. S1 and Fig. S2 d). Nonetheless, it is possible that the UxuR model's linkers adopt the *E. coli* FadR's linkers conformational. Recently, this assumption was made by Singh *et al.* while studying a TF FadR family member, named DgoR.<sup>3</sup> In their study, they used the structure of the FadR dimer (PDB ID: 1HW2, 3.25 Å resolution) as the template for N-terminal domain modelling.<sup>3</sup>

In summary, as a consequence of the scarce information available on the linker's architecture in FadR family members, we used molecular dynamic (MD) simulations. Inspired by the work of Horne *et al.*<sup>2</sup> on NanR TF and Singh *et al.*<sup>3</sup> on DgoR TF, we proposed a possible allosteric model of UxuR functioning. As illustrated in Fig. S3, there might exist four key states: i) Bond state – UxuR is bonded to the operator binding site, repressing gene transcription; ii) "Allosteric" state – the presence and binding of an effector molecule (FrctA) to the C-terminal domain in UxuR disrupts the allosteric communication from one domain to another which leads to a conformational alteration of UxuR. Consequently, the affinity with DNA is disrupted and UxuR is no longer able to bind to the operator binding site; iii) In complex state – upon complete FrctA binding to both monomers, UxuR adopts a final conformation which is no longer compatible with DNA binding affinity. After a decreased concentration of effector molecules inside the cell, the FrctA is released from the Zn(II) pocket and the residue interactions seen in Bond state start to be reestablished; and iv) Free state – UxuR is back to its free form and ready to bind to its DNA operator binding site. In the present study, our UxuR model is analysed in its apo form and in complex with two ligands. Thus, only Free and In complex states are under examination. Since the right placement of the linkers is unknown, UxuR in its apo form (Free state) was simulated, as detailed below, to obtain the closest structure of UxuR suitable for binding to DNA (Bond state).

#### Molecular Docking details

Here, we studied the atomic changes between the free protein and its complexes with the ligands  $\alpha$ -D-fructuronate (FrctA; PubChem Identifier: CID 46878576) and  $\beta$ -D-glucuronate (GlcNA; PubChem Identifier: CID 11877136). For both systems, Molecular Docking was performed with the AutoDock4.2 suite of programs<sup>4</sup> with the Lamarckian Genetic Algorithm (LGA).<sup>5</sup> A grid box (40 × 40 × 40 Å) was centered on the Zn(II) ion, at both chains (chain A: x = 47.741; y = -2.039; z = 21.566 and chain B: x = 46.137; y = 38.828; z = 23.538), for FrctA as well as for GlcNA. A total of 100 LGA runs were carried out for each of the ligand-protein complexes. The population was 300, the maximum number of generations was 27 000 and the maximum

number of energy evaluations was 2 500 000. All rotatable bonds were kept flexible. From the list of results obtained by the molecular docking calculations we visually verify the structural conformations of the ligands with the lowest binding energy of the most populated cluster.

### System Setup

Contrary to the method used in our last work about the UxuR model,<sup>1</sup> the zinc metal coordination was modeled using the empirical scheme of the Extended Zinc AMBER Force Field (EZAFF) model approach according to Zhuoqin Yu and colleagues.<sup>6</sup> Also, we used MCPB.py<sup>7</sup> - a python-based metal center parameter builder - to empirically assign the bonded and angle parameters involving zinc and ligand. The atomic partial charges were calculated by employing the RESP method from B3LYP/6-31G\* single-point energy calculations.<sup>8,9</sup> The missing hydrogen atoms were added using the tleap module in the Molecular Dynamic (MD) software package Amber18.<sup>10</sup> The structures were placed in an octahedral box of water molecules (10.0 Å between the surface of the protein and the box). Counter ions were added to make the system neutral. The general Amber force field<sup>11</sup> and Amber ff14SB force field<sup>12</sup> were used. When present, ligands were described using parameters from the GLYCAM database.<sup>13</sup> We used H++ version 4.0<sup>14</sup> to verify the protonation state of each system at pH 7.

### Conventional Molecular Dynamics details

Conventional Molecular Dynamic (cMD) simulations were performed using the MD software package Amber18.<sup>10</sup> Each system was subjected to two initial energy minimizations and 500 ps of equilibration in an *NVT* ensemble, using Langevin dynamics<sup>15</sup> with small restraints on the protein (10.0 kcal/mol) to heat the systems from 0 K to 310.15 K (37°C, optimal temperature for the bacterial organism). Production simulations were carried out at 310.15 K in the *NPT* ensemble, using Langevin dynamics with a collision frequency of 1 ps<sup>-1</sup>. Constant pressure periodic boundary conditions were imposed with an average pressure of 1 ATM. Isotropic position scaling was used to maintain pressure with a relaxation time of 2 ps. The SHAKE algorithm<sup>16</sup> was applied to all bonds involving hydrogen atoms. The Particle Mesh Ewald method<sup>17</sup> was used to calculate electrostatic interactions with a cutoff distance of 10.0 Å. The total time of cMD simulations was 10 ns for each system (five replicas with different initial velocities each), and the integration time was set to 2 fs. These cMD simulations were useful to check the stability of our systems before performing simulations with enhanced sampling and to obtain representative structures for all analyses when necessary.

### Unconstrained Enhanced Molecular Dynamics

We performed Gaussian accelerated molecular dynamics (GaMD) simulations<sup>34</sup> using the MD software package Amber18.<sup>10</sup> GaMD simulations, an enhanced sampling technique, adding a harmonic boost potential. This technique smooths the biomolecular potential energy surface and reduces the system energy barriers.<sup>18</sup> Since the added boost potential follows a Gaussian distribution, biomolecular free energy profiles can be accurately recovered through cumulant expansion to the second order.<sup>18</sup> As previously described, we performed 10 ns of cMD production. Then, for all replicas of each system, the dual-boost GaMD simulations progressed with 3 ns cMD simulation to collect potential statistics, 50 ns of equilibration after adding the boost potential, and then production runs (check Table S1 to verify all performed simulations and respective production times). All systems followed the same procedure, and GaMD production frames were saved every 0.1 ps for analysis.

### Structural Analyses

The following analyses were computed using the CPPTRAJ<sup>19</sup> from AmberTools18.<sup>10</sup> We make use of all residues of each system except for the first and last ten residues of each monomer, due to the instability associated with the simulations. For all analyses, we make use of the backbone atoms (C, C $\alpha$ , O and N) of each system. To obtain the reference structure of each system, a hierarchical algorithm was applied to cluster the concatenated cMD simulation; the frames were sieved at 10 and the distance cutoff was 2.5 Å. The backbone root-mean-square deviation (RMSD) was computed for each replica of each system, by computing every frame of the entire simulation time. The remaining trajectory post-processing was performed by computing every 30 frames of the concatenated simulation time of each system (the first 50 ns of each replica

were excluded to reduce bias towards reference structures), except for the backbone per-residue root-mean-square fluctuation (RMSF) where we used all frames.

The RMSF was calculated per monomer and per domain, separately, using the most populated structural extracted from hierarchical clustering of the free form's cMD simulations. To identify the overall patterns of motions, we performed Principal Component Analyses (PCA) by setting the cMD lowest structures as a reference, the global translation and rotation were eliminated. A series of eigenvectors representing various modes of conformational change, and their corresponding eigenvalues were produced by diagonalizing the Cartesian covariance matrix of the involved residues. The Free Energy Landscape (FEL) was plotted with the first two principal components (PC1 and PC2). The PyReweighting toolkit<sup>20</sup> was used for reweighting the free energy values of the simulations using cumulant expansion to the 2<sup>nd</sup> order. To assess correlated motions both within individual monomers and between the two monomers, we calculated the Dynamic Cross-Correlation Maps (DCCM) with Bio3D<sup>21</sup> (Rstudio<sup>22</sup>) by considering C $\alpha$  atoms only (all solvent and non-C $\alpha$  atoms were ignored to reduce total file size).

### Model validation

The overtime structural stability of all systems was examined by RMSD. As shown in Fig. S4 and S5, the UxuR model was more flexible than the average proteins, mainly due to the flexibility of the N-terminal DNA-binding domain. As illustrated in Fig. S5, the average distribution of RMSD without the N-terminal domain presents to be less flexible in all systems when compared to the N-terminal domain RMSD. Of note, for this analysis, the linker was included as part of the N-terminal domain, and considering the degree of disorganization of the linker it might be the biggest influence for the high values observed in the N-terminal domain RMSD. The RMSF analyses shown in Fig. S6, suggest that there is a difference in chain behavior over time, especially at the N-terminal domain of UxuR<sup>Apo</sup>. Previous studies have suggested that DNA-binding helices are capable of moving independently of each other, affecting the process of binding affinity search for both specific and nonspecific DNA.<sup>23-25</sup> A careful analysis indicates that the first approach in RMSF analyses when rms fit is performed with the entire protein, revealed that the data obtained illustrates the translation of each domain through time. Here, overall, UxuR<sup>Apo</sup> is the less stable system, meaning that it changes more over time from the reference structure when compared to the systems where a ligand is present. This is in good agreement with the literature, as a repressor in its free form, higher flexibility at the N-terminal domain could translate the search state for the specific DNA sequence (operator) to bind.<sup>23-25</sup> Moreover, to understand how each domain behaves upon ligand binding, the RMSF was calculated for each domain individually (Fig. S7). In good agreement with the RMSD results, the results of RMSF per domain show that the linker is the driver of most of the residue changes over time. In Fig. S7 one can observe that the N-terminal domain for each system is quite stable; whereas the binding of a ligand increases the flexibility of the residues belonging to or near the loops of the barrel (loop1 Glu135-Glu145 and loop2 Asn177-Tyr193). Noteworthy, the fluctuation of the residues in the linker when GlcNA is present is lower when compared with the two other systems (Fig. S7 d). The dominant conformational changes and structural fluctuations that occur within the protein over time were extracted with PCA, focusing on the most important PCs. The plots of the first three eigenvectors (PC1, PC2, and PC3) of all systems were sufficient to describe more than 70 % of the variance (Fig. S8). The first two eigenvectors plot demonstrates that the UxuR<sup>FrctA</sup> progressively shifts the population of conformations, which indicates a more stable and less flexible behavior than UxuR<sup>Apo</sup>. Interestingly, the PC1 vs PC2 plot of UxuR<sup>GlcNA</sup> shows a similar periodic jump from one phase space to another as seen in UxuR<sup>Apo</sup>, suggesting a less stable structure when compared to UxuR<sup>FrctA</sup>. Moreover, the first two PCs of each system were used to estimate the FEL and subsequently determine the minimal energy (native) structures (Fig. S9). Visual inspection of the FEL plots demonstrates that the UxuR model is indeed flexible during the UxuR<sup>Apo</sup> and UxuR<sup>GlcNA</sup> simulations, with a variety of metastable conformations (Fig. S9a and c). On the other hand, when FrctA is bound to the C-terminal domain, it results in a single native structure (Fig. S9b). Analysis of the DCCM revealed correlated motions between the UxuR monomers and within each monomer, as illustrated in Fig. S10. DCCM comparison between UxuR<sup>Apo</sup> and UxuR<sup>FrctA</sup> showed intra- and inter-molecular connection changes, indicating induced alterations in the amino acid correlation upon FrctA binding. On one hand, the movement of residues between domains (N-terminal vs C-terminal) in the same monomer is less correlated when FrctA is present. On the other hand, the movement of residues between monomers in the C-terminal domain, in general, became more correlated upon binding. Of note, helices that maintain dimerization ( $\alpha_5$ -helix and  $\alpha_8$ -helix, coloured in yellow in Fig. S10d) and the ones that interact with them evolve to a different correlation pattern translating change upon binding (yellow rectangle, Fig. S10). Taken together, the correlation motions were stronger (either positive or negative) in UxuR<sup>Apo</sup>, again

underlying the higher flexibility of the UxuR model in its free form. In addition, the notable strong correlation in UxuR<sup>Apo</sup> demonstrates the direct impact of FrctA presence on the overall conformation, which further reduces the flexibility of the UxuR model. Inspection of UxuR<sup>GlcNAc</sup>'s DCCM shows a pattern of ligand binding like FrctA, but the correlations between monomers are more anti-correlated. These dissimilarities indicate that our systems can describe UxuR in different states (free from and in complex with FrctA or GlcNA). More importantly, these results strongly indicate that our model is adequate for studying allosteric events in the TF under study. However, taking into consideration the above mentioned linkers' flexibility and given the impossibility (without a crystal structure) to fully know the UxuR's domain orientation, this study will be only focusing on the residues belonging to the C-terminal domain to avoid misleading results.

### Short- and long-range communication analyses

For all systems, we analyzed differences in hydrogen-bonding interactions between monomers and domains, using a donor-acceptor distance cutoff of 3.5 Å and a donor-hydrogen-acceptor angle cutoff of 135°. Hydrogen bonds (HB) that appeared in fewer than 10 % of the simulation frames were excluded. The protein structure networks and interaction paths between residues, of each system, were created by "Structure ENSEmble NETworks" (SenseNet)<sup>26</sup> with Cytoscape.<sup>27</sup> Here, we combined the contact timelines of carbon atoms and the hydrogen bonds (distance cutoff 3.5 Å; angle cutoff 135°). To reduce the impact of sporadic interactions, edges that depicted interactions within < 10% of the total simulation time were eliminated from the network. The networks were analyzed with the Difference Node Correlation Factor (DNCF) method.<sup>26</sup> This method is based on the presupposition that for a residue to be considered allosteric, the interaction changes within a residue (edge neighbor correlation factor - ECF) must have observable influence in its immediate environment (node correlation factor - NCF). In summary, DNCF quantifies changes between two networks obtained from different MD simulations (target vs reference). The DNCF scores were calculated using the implemented "Correlation" function set to the "Mutual information difference" and "Match Location" modes against the reference system network. Then, the "Degree" function was used to sum over the Neighbor Correlation Factor (ECF) scores calculated in the previous step and normalized using Z-score normalization.

### Data visualization

All plots were generated using Matplotlib<sup>28</sup> and OriginPro version 8.5 OriginLab Corporation (Northampton, MA, USA). Images of protein structures were produced by UCSF Chimera.<sup>29</sup> Finally, to visualize the amino acid positions, we performed alignment with different TFs belonging to the FadR family (InterPro Identifier: IPR008920) and used ESPript<sup>30</sup> (<https://esript.ibcp.fr>) to identify conserved and similar residues among those TFs (Fig. S11).

### Experimental Workflow

#### Microbial Strains

The bacterial strains used in this investigation were variants of *E. coli* K-12 MG1655: MG1655 with *uxuB* and *uxuR* knockouts (MBR), MG1655(DE3) with *gudD*, *uxaC*, *uxuB* and *uxuR* knockouts (M2BR), MG1655 with *uxuB*, *exuR* and *uxuR* knockouts (#3484) and MG1655 with *uxaC*, *uxuR* and *exuR* knockouts (#3491). Additionally, the *E. coli* DH5α strain was used for plasmid construction (Table S2).

#### Sensor Plasmid Construction

An *E. coli* UxuR biosensor was created based on Ni *et al.*'s previous work.<sup>31</sup> However, in this study, we used a binding-site sequence and gene of interest, as detailed in Table S3. The plasmid architecture of the *E. coli* UxuR biosensor is analogous to the P100 design reported by Ni *et al.*<sup>31</sup> Initially, the *uxuR* gene was constitutively expressed under promoter BBa\_J23117 of the Anderson promoter library,<sup>32</sup> whereas the *sfGFP* reporter gene is expressed under an inducible hybrid promoter. This hybrid promoter is an insulated proD promoter<sup>33</sup> with the UxuR operator sequence<sup>34</sup> (Table S3) inserted both between and downstream of the promoter's -35 and -10 RNA polymerase binding sites. Ni's pP100-BsExuR-gfp plasmid, which instead contains the *B. subtilis* ExuR gene and operators, was used as the original DNA template for constructing this biosensor.

First, polymerase chain reaction (PCR) with the BB primer set (Table S4) amplified the backbone of the pP100-BsExuR-gfp biosensor without the *B. subtilis* *exuR* gene. The *E. coli* *uxuR* gene was also amplified from

the pSNR1-UxuR-Upro plasmid designed by Almeida *et al.*,<sup>1</sup> via PCR with the INS primer set (Table S4). Then, the *E. coli uxuR* gene was inserted into the biosensor backbone via Gibson assembly with the NEBuilder HiFi DNA Assembly Cloning Kit (New England BioLabs, Ipswich, MA). Performed twice, PCR amplification of the plasmid with an OP1-2 primer set (Table S4) and subsequent ligation with the KLD Enzyme Mix (New England BioLabs, Ipswich, MA) had replaced each of the two *B. subtilis* ExuR operator sequences in the hybrid promoter with *E. coli* UxuR operator sequences. Thus, creating the pP100-EcUxuR-gfp biosensor. After preliminary results (data not shown), we performed Site-directed Mutagenesis using the Q5<sup>®</sup> Site-Directed Mutagenesis kit (New England BioLabs, Ipswich, MA) to design a plasmid biosensor whose gene of interest (*uxuR*) constitutively expressed by a promoter with higher strength- BBa\_J23101 of the Anderson promoter library<sup>32</sup> (Table S3). The pP100-EcUxuR-gfp biosensor plasmid was amplified via PCR with P1 and P2 primer sets (Table S4) thus creating the UxuR biosensor plasmid (Table S3). All custom primers used for PCR are DNA oligonucleotides purchased from Millipore Sigma (The Woodlands, TX). Throughout this plasmid's construction, the chemical transformation of *E. coli* DH5 $\alpha$  cells was used to propagate DNA, and DNA sequencing was used to verify desired assemblies.

Site-directed Mutagenesis was performed using the Q5<sup>®</sup> Site-Directed Mutagenesis kit (New England BioLabs, Ipswich, MA) to design variants of the UxuR biosensor plasmid. Briefly, the UxuR biosensor plasmid was amplified by PCR using the primers Mut1-20\_F and Mut1-20\_R (Table S4). The amplified PCR product was digested and ligated using the KLD Enzyme mix (New England BioLabs, Ipswich, MA) for 30 min at room temperature. The ligation product was used to transform DH5 $\alpha$  cells. The colonies were sequenced to verify the presence of the desired mutation. All the primers and plasmids used in this study are listed in Tables S4 and S5, respectively.

### Growth Conditions

The strains of interest were transformed with the UxuR biosensor and mutated versions. All transformations were plated on Luria-Bertani (LB) Agar Miller (BD Difco™) with 100  $\mu$ g/mL carbenicillin at 37 °C. Overnight cultures of transformants were grown overnight in Luria-Bertani (LB) Broth Miller (BD Difco™) with 100  $\mu$ g/mL carbenicillin at 37 °C and 250 rpm agitation (Fig. S11a and b).

### GFP Measurement

Fermentation cultures were inoculated from overnight cultures in LB Broth-Miller (BD Difco™) and 100  $\mu$ g/mL carbapenem at a dilution of 1:100 vol/vol. Fermentation was carried out for 24 h in a BioLector (m2p labs GmbH, Baesweiler, Germany). Cultures were grown in FlowerPlates (m2p labs) containing 1 mL medium per well at 37 °C, 1200 rpm agitation, and 80 % relative humidity. The wells contained LB or LB medium supplemented with D-glucuronate (GC  $\geq$ 98 %, Sigma-Aldrich) or D-galacturonate ( $\geq$ 95 % (T), Sigma-Aldrich) (different concentrations were used, specified later); 100  $\mu$ g/mL carbenicillin was added to all cultures. Biomass (backscattered light at 620 nm) and GFP (488 nm excitation / 520 nm emission) measurements were taken by the BioLector every ~20 min. GFP units, as measured by BioLector, were normalized by biomass units from BioLector for data analysis.

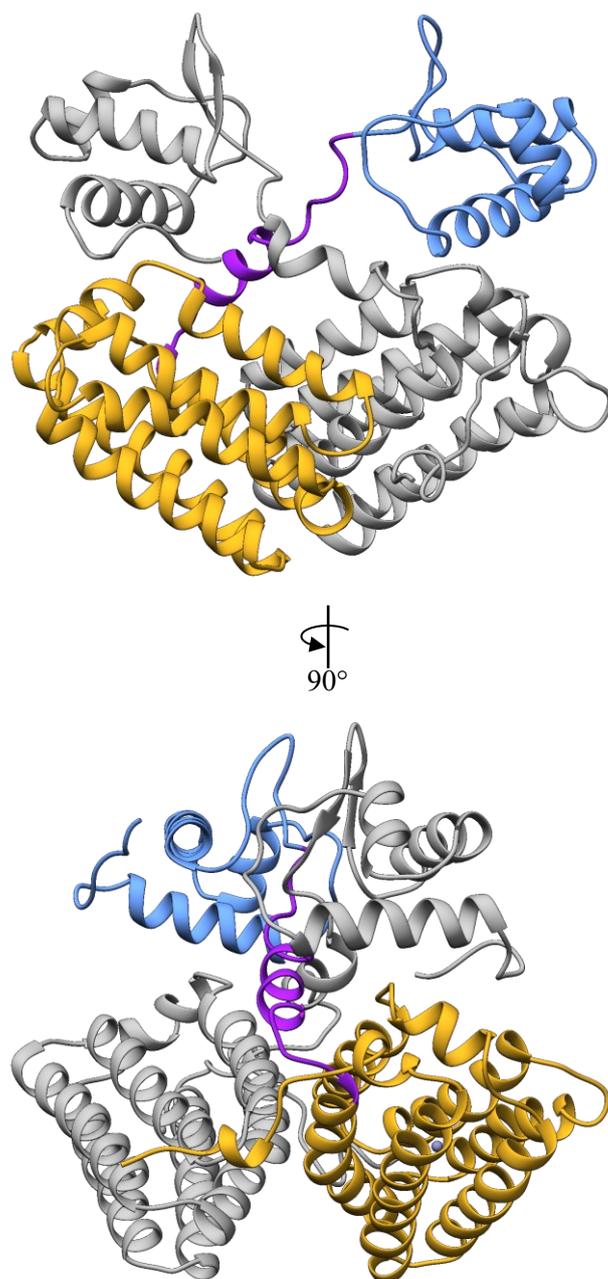
### Biosensor validation

The UxuR biosensor was constructed as illustrated in Fig. S11c and d. We tested different concentrations of GlcNA in the two strains: MBR ( $\Delta$ *uxuB* $\Delta$ *uxuR*) and M2BR ( $\Delta$ *gudD* $\Delta$ *uxaC* $\Delta$ *uxuB* $\Delta$ *uxuR*). The most sensitive results for both strains were obtained using 18 mM GlcNA (Fig. S12a). As expected, the UxuR biosensor responded to FrctA (Fig. S12b), but not to GlcNA (Fig. S12c). The 24h fold change in GFP between MBR and M2BR strains indicates a stable biosensor suitable to test single mutations in the UxuR TF in response to FrctA and GlcNA (Fig. S12d).

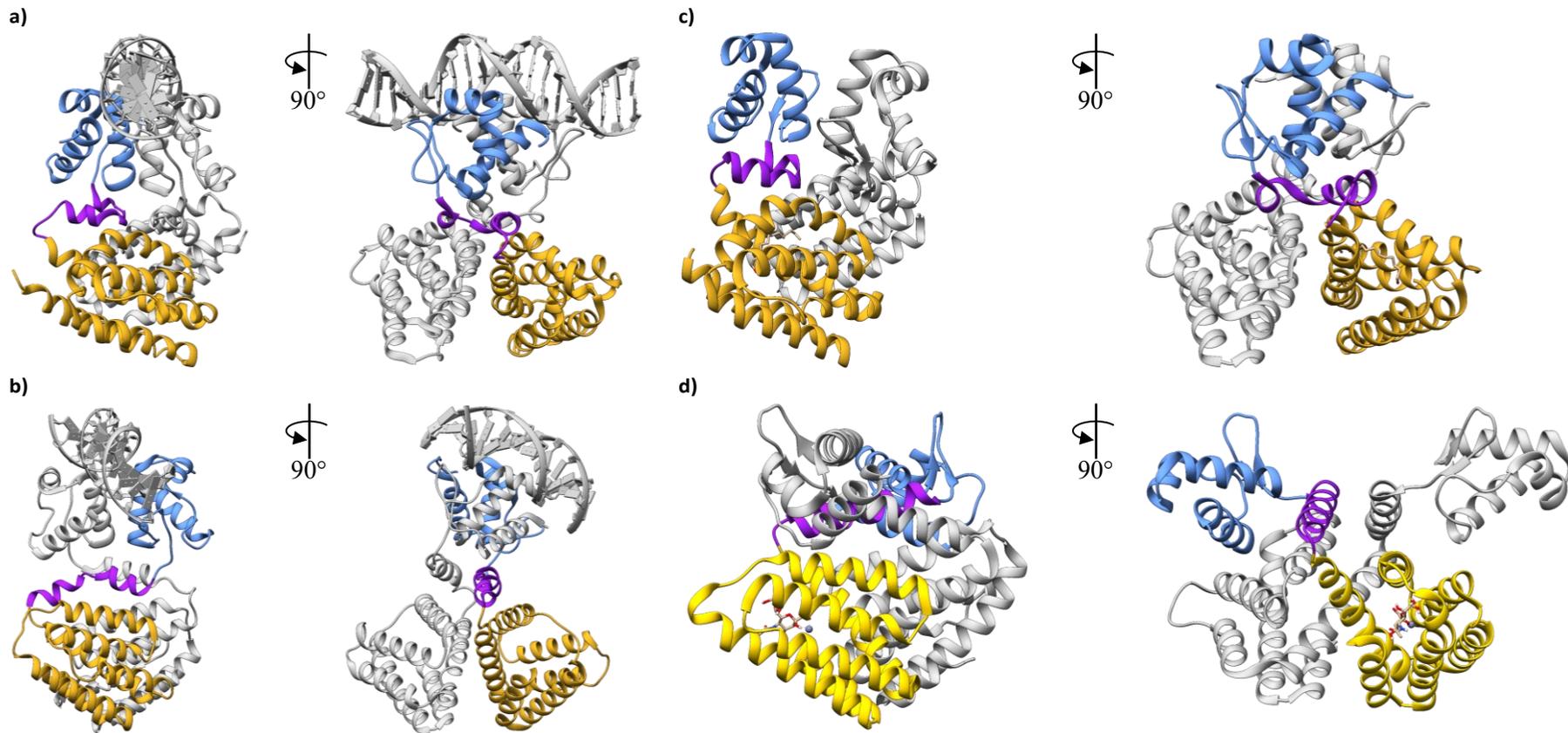
### Analyses

Statistical analyses and plots were performed using GraphPad Prism version 8.0.2 for Windows, GraphPad Software, www.graphpad.com (Boston, Massachusetts USA). An unpaired t-test assuming unequal standard deviations (Welch's correction) was used to compare the normalized GFP signal (GlcNA:LB only) and to compare the fold change in GFP between the mutated versions and the UxuR biosensor (Mut:WT).

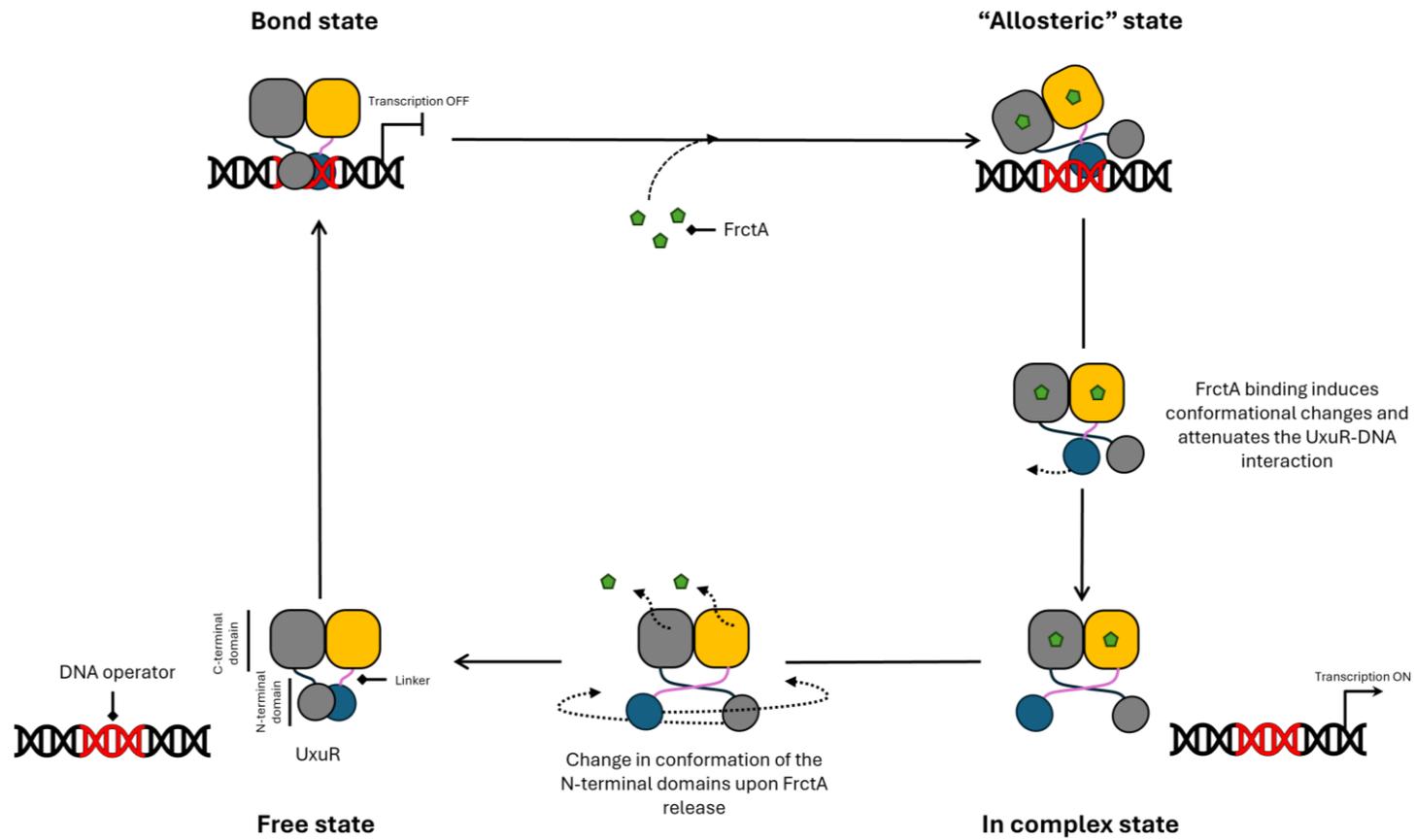
Figure S1-S33



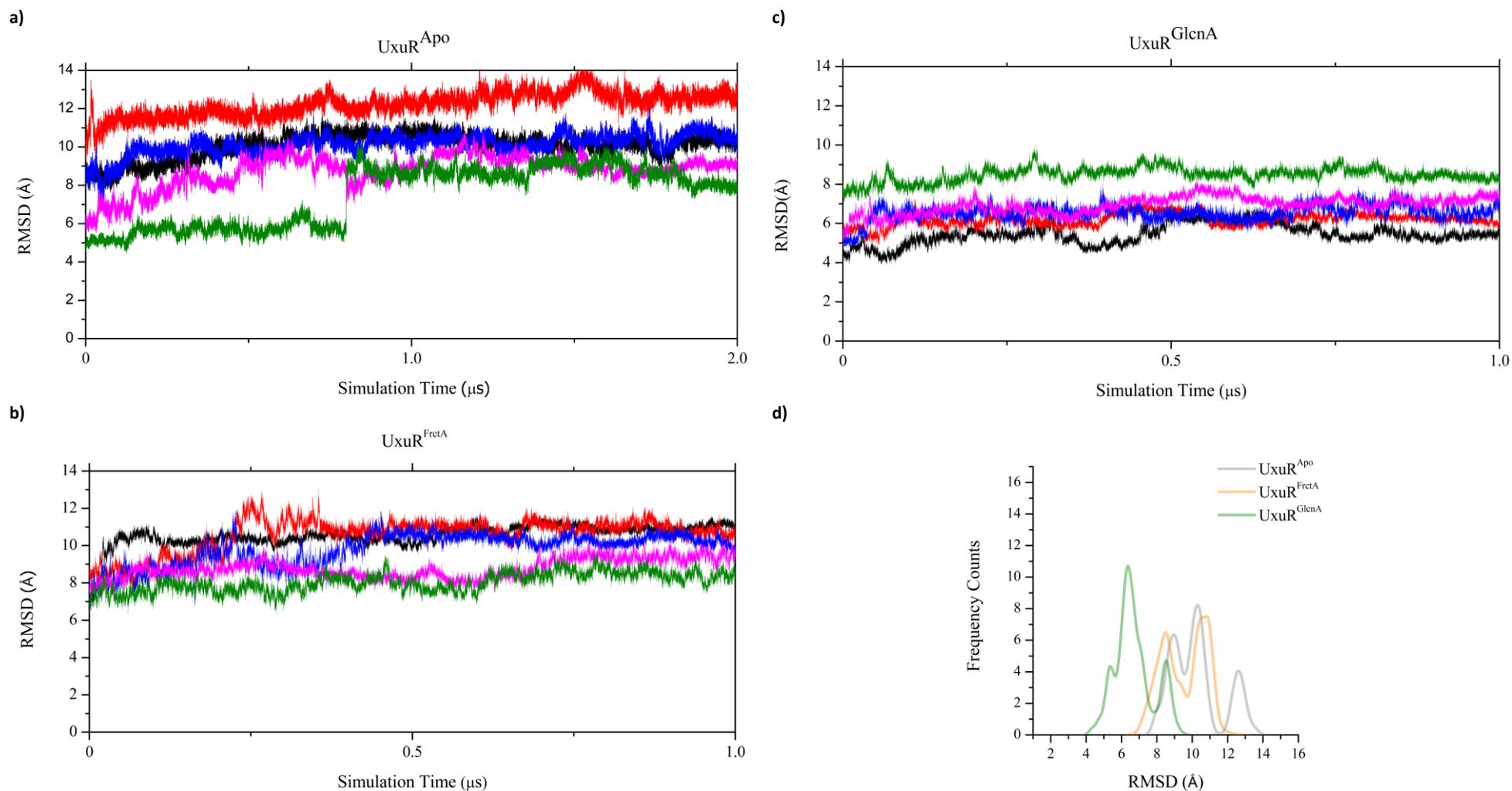
**Fig. S1.** Representative structure of the UxuR TF model in its free form obtained from the hierarchical algorithm cluster of the concatenated cMD simulation. Two perspectives are illustrated front view (top) and side view (bottom). The monomer A was highlighted by different colors: the N-terminal domain is colored in blue, the linker is colored in purple and the C-terminal domain is colored in orange. The monomer B was colored in grey. The Zn(II) ion is represented by a grey sphere.



**Fig. S2** Structure representation of two FCD family members illustrated by two perspectives, front view (left) and side view (right). **a)** FadR from *E. coli* in complex with DNA (PDB code 1H9T), **b)** NanR from *E. coli* in complex with DNA (PDB code 6WFQ), **c)** FadR from *E. coli* in complex with myristoyl-CoA (PDB code 1H9G) and **d)** NanR from *E. coli* in complex with sialic acid (PDB code 6ON4). For each protein, the monomer A was highlighted by different colors: the N-terminal domain is colored in blue, the linker is colored in purple and the C-terminal domain is colored in orange. The monomer B and DNA molecules were colored in grey, the effector molecule is represented by a licorice structure and the Zn(II) ion is represented by a grey sphere.

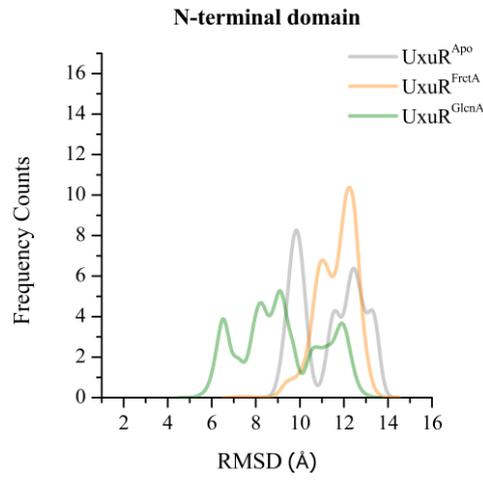


**Fig. S3.** Schematic representation of the proposed putative allosteric model of UxuR functioning.

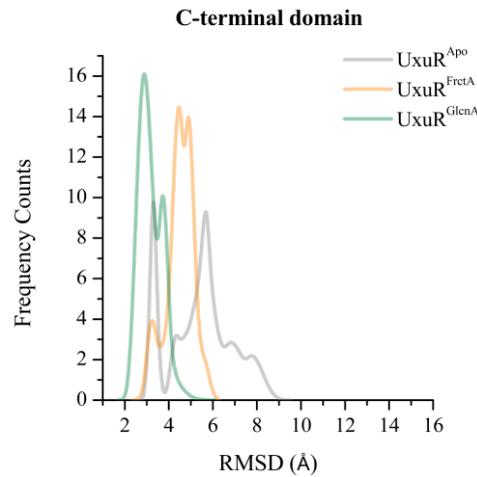


**Fig. S4** The backbone root-mean-square deviation (RMSD) calculated over 5 independent of UxuR model TF in its free form (a), in complex with D-fructuronic acid – FrctA (b) and in complex with D-glucuronic acid - GlcNA (c). The average distribution of root-mean-square deviation (RMSD) of whole protein was calculated over a period of 5.0 μs (5 replicates x 1.0 μs) for the three systems (last microsecond in the free from case). The RMSD of UxuR<sup>Apo</sup> from the cMD's reference structure ranges between 7 and 14 Å during the last 5μs; the UxuR<sup>FrctA</sup> from the cMD's reference structure ranges between 6.5 and 13 Å; and the UxuR<sup>GlcNA</sup> from the cMD's reference structure ranges between 4 and 10 Å. The lowest structure from cMD simulations clustering was used as a representative structure for the RMSD calculations.

a)

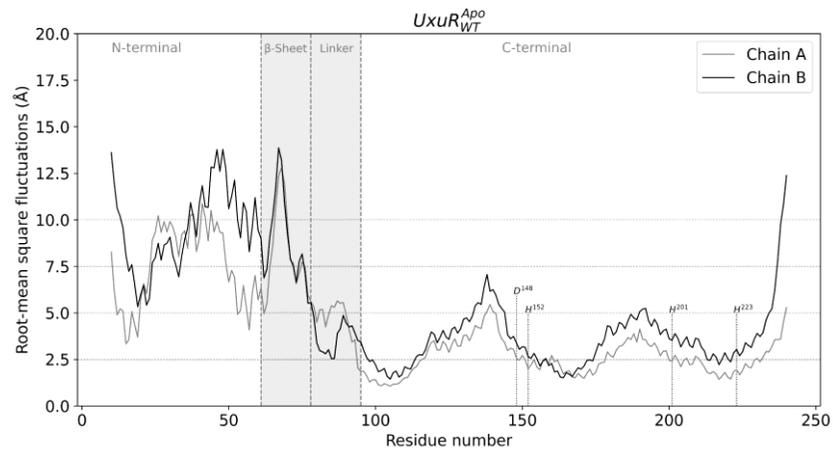


b)

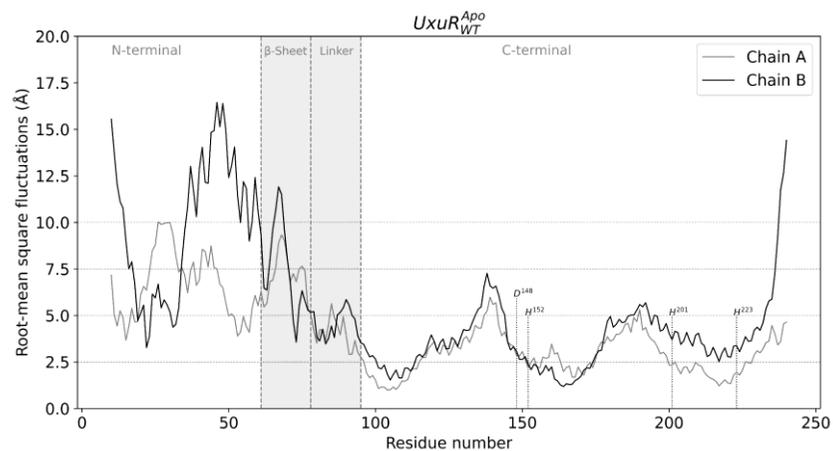


**Fig. S5** Average distribution of root-mean-square deviation (RMSD) calculated over a period of 5.0  $\mu$ s (5 replicates x 1.0  $\mu$ s) for the three systems: free form last microsecond (gray), in complex with D-fructuronic acid – FrctA (orange) and in complex with D-glucuronic acid - GlcnA (green). **a)** N-terminal DNA-binding domain – from 1 to 95 and **b)** C-terminal effector-binding domain – from 96 to 257. Notably, the value range is larger in the UxuR<sup>GlcNA</sup>'s N-terminal domain than in other systems; however, the N-terminal domain changes more from the cMD reference structure in the UxuR<sup>Apo</sup> and UxuR<sup>FrctA</sup> GaMD simulations.

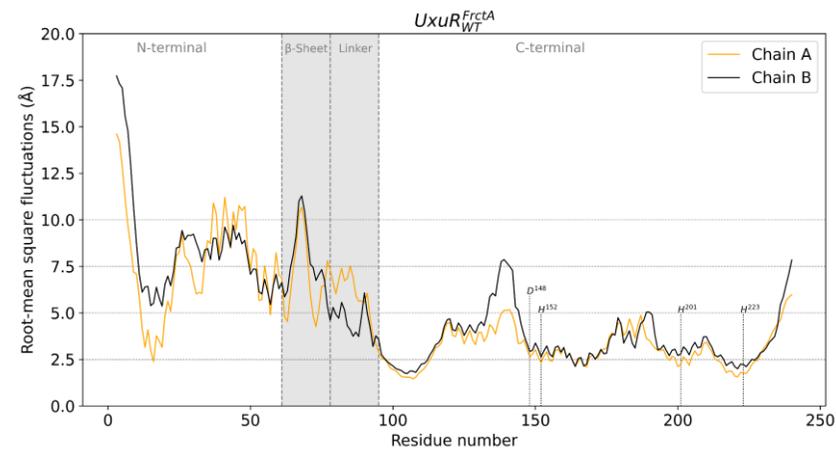
a)



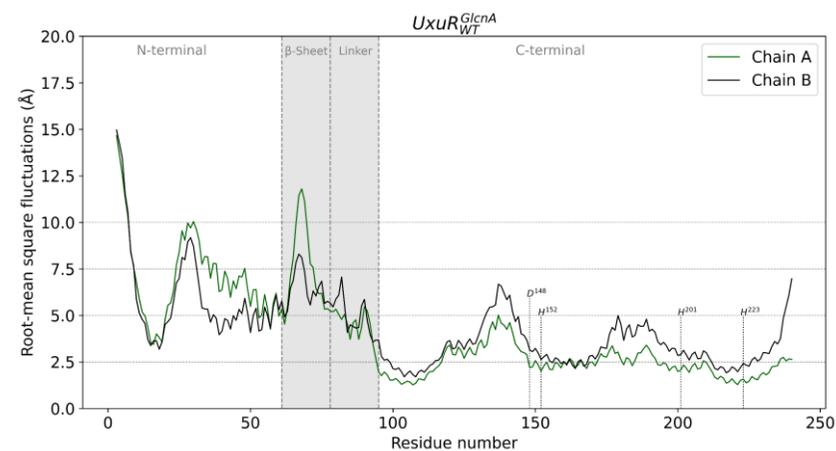
b)



c)

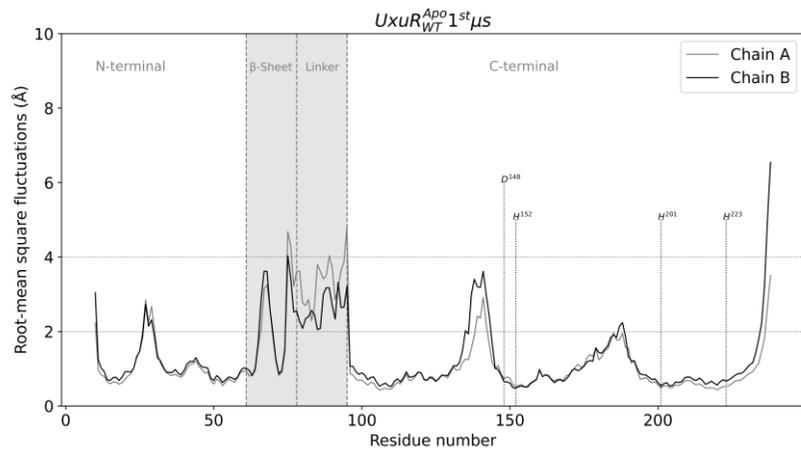


d)

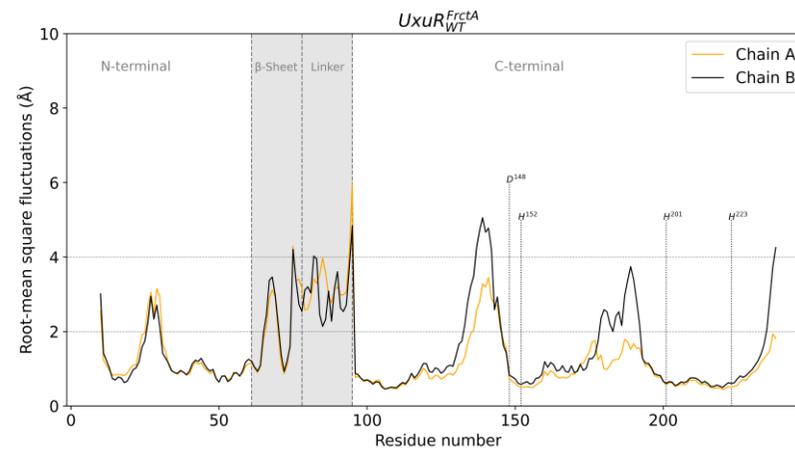


**Fig. S6** Per-residue root-mean-square fluctuation (RMSF) of backbone atoms calculated over a period of 5.0  $\mu$ s (5 replicates x 1.0  $\mu$ s) GaMD simulations for the three systems: Free form first microsecond (a) and last microsecond (b); UxuR in complex with FrctA (c), and UxuR in complex with GlcNA (d).

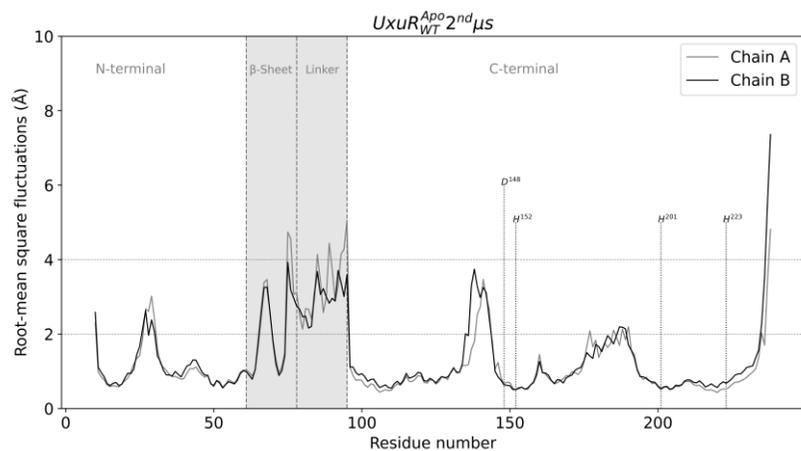
a)



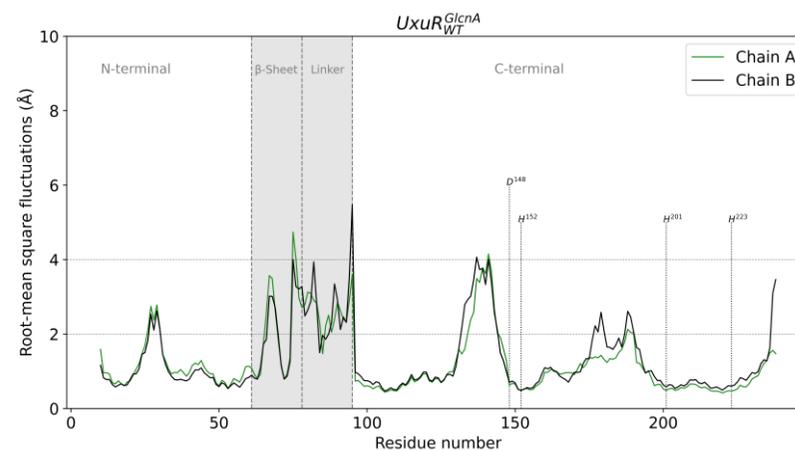
c)



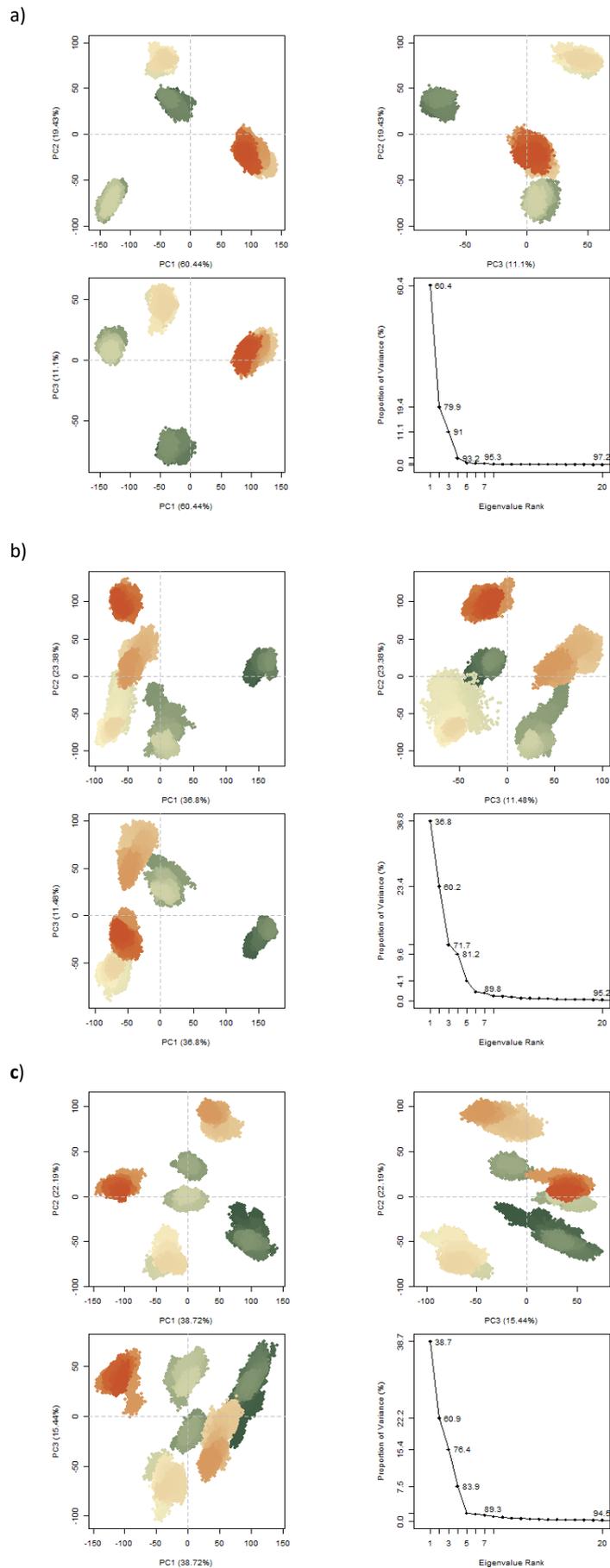
b)



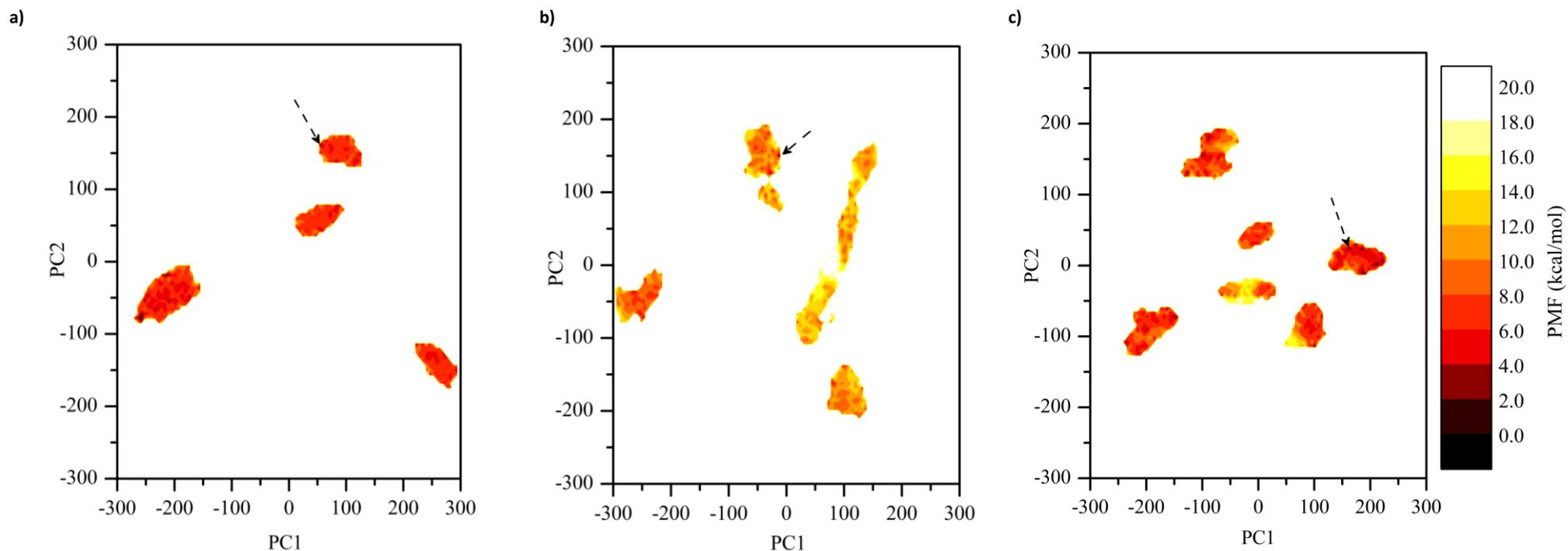
d)



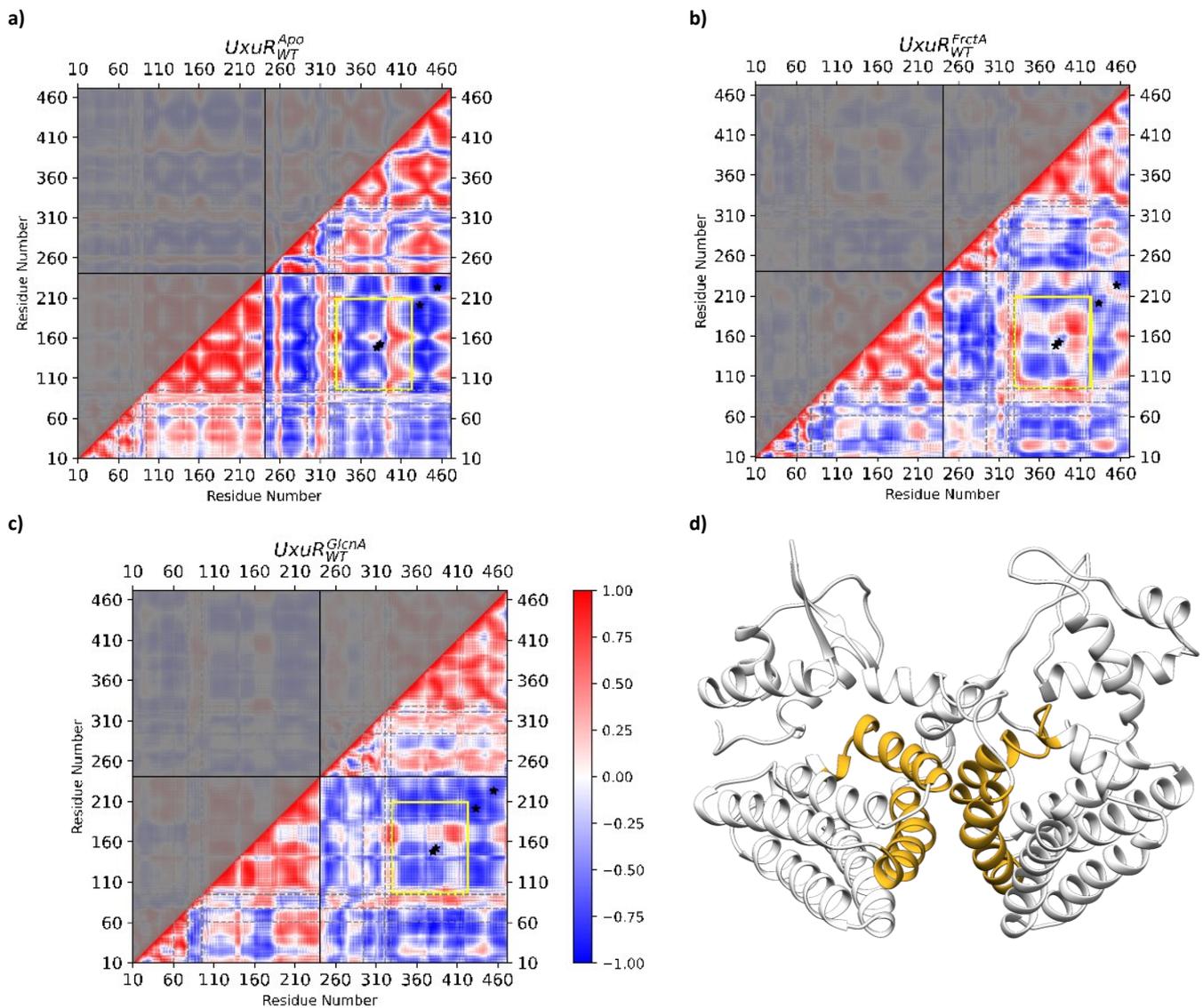
**Fig. S7.** Per-residue root-mean-square fluctuation (RMSF) of backbone atoms calculated per domain over a period of  $5.0 \mu\text{s}$  (5 replicates  $\times$   $1.0 \mu\text{s}$ ) GaMD simulations for the three systems: Free form first microsecond (a) and last microsecond (b); UxuR in complex with FrctA (c), and UxuR in complex with GlcNAc (d).



**Fig. S8** The principal component analysis (PCA) of the free form of UxuR – UxuR<sup>Apo</sup> during the last microsecond (a), UxuR in complex with D-fructuronic acid - FrctA (b), and UxuR in complex with D-glucuronic acid - GlcNA (c) The color gradients (green to yellow to orange) represent the periodic jumps between conformations.



**Fig. S9** Two-dimensional Free energy landscapes (FEL in kcal/mol) obtained from the first two eigenvectors (PC1 and PC2) at 300 K. **a)** UxuR TF model in free form (UxuR<sup>apo</sup>); **b)** UxuR TF model in complex with D-fructuronic acid - FrctA and **c)** UxuR TF model in complex with D-glucuronic acid - GlcNA. The deeper color areas in the maps represent the metastable and native states. The native states' structural coordinates (indicated by a black arrow) were extracted from the lowest energy state from a specific time frame.



**Fig. S10** Dynamic cross-correlation maps (DCCM) of UxuR model TF in its free from - UxuR<sup>Apo</sup> (a), complex with D-fructuronic acid - UxuR<sup>FrctA</sup> (b) and complex with D-glucuronic acid - UxuR<sup>GlcNA</sup> (c). The bottom left and top right squares of each plot reveal correlated motions within individual monomers of the dimer, while the top left and bottom right squares display correlated motions between the two monomers of the dimer. Positive values (in red) represent residues that move in the same direction (correlated), whereas negative values (in blue) represent residues that move in the opposite direction (non-correlated). Deeper colors indicate the stronger positively (in red) or negatively (in blue) correlated residue motion. Black stars indicate the four conserved residues (Asp148, His152, His201 and His223) known to form the Zn<sup>2+</sup> binding site in UxuR. The yellow square delineates the area where the correlations change the most upon ligand binding. d) Minimal energy (native) structure of UxuR TF model in complex with FrctA obtained from the reweight free energy landscape. Helices colored in yellow highlight the helices  $\alpha_5$  and  $\alpha_8$  of each monomer.

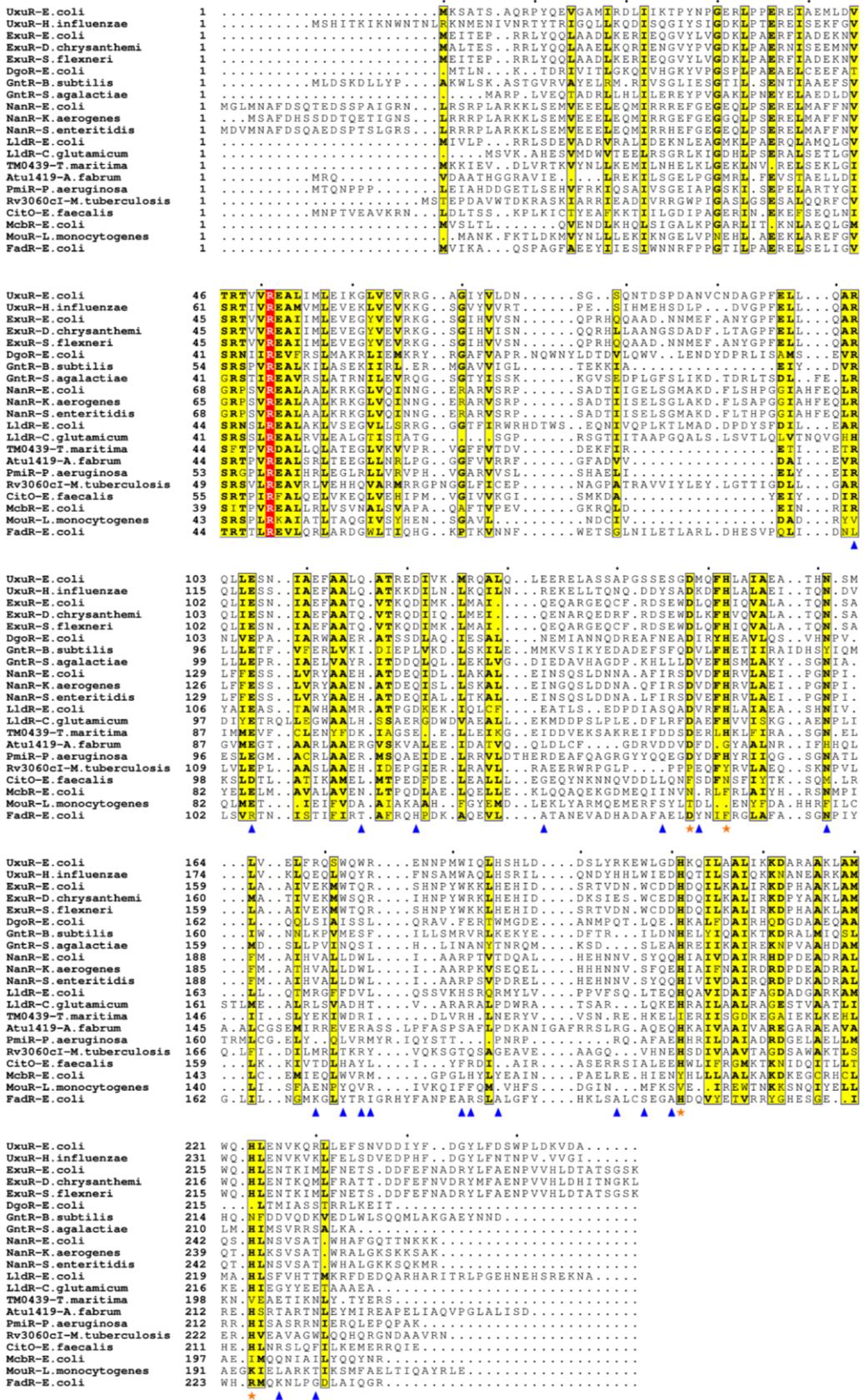
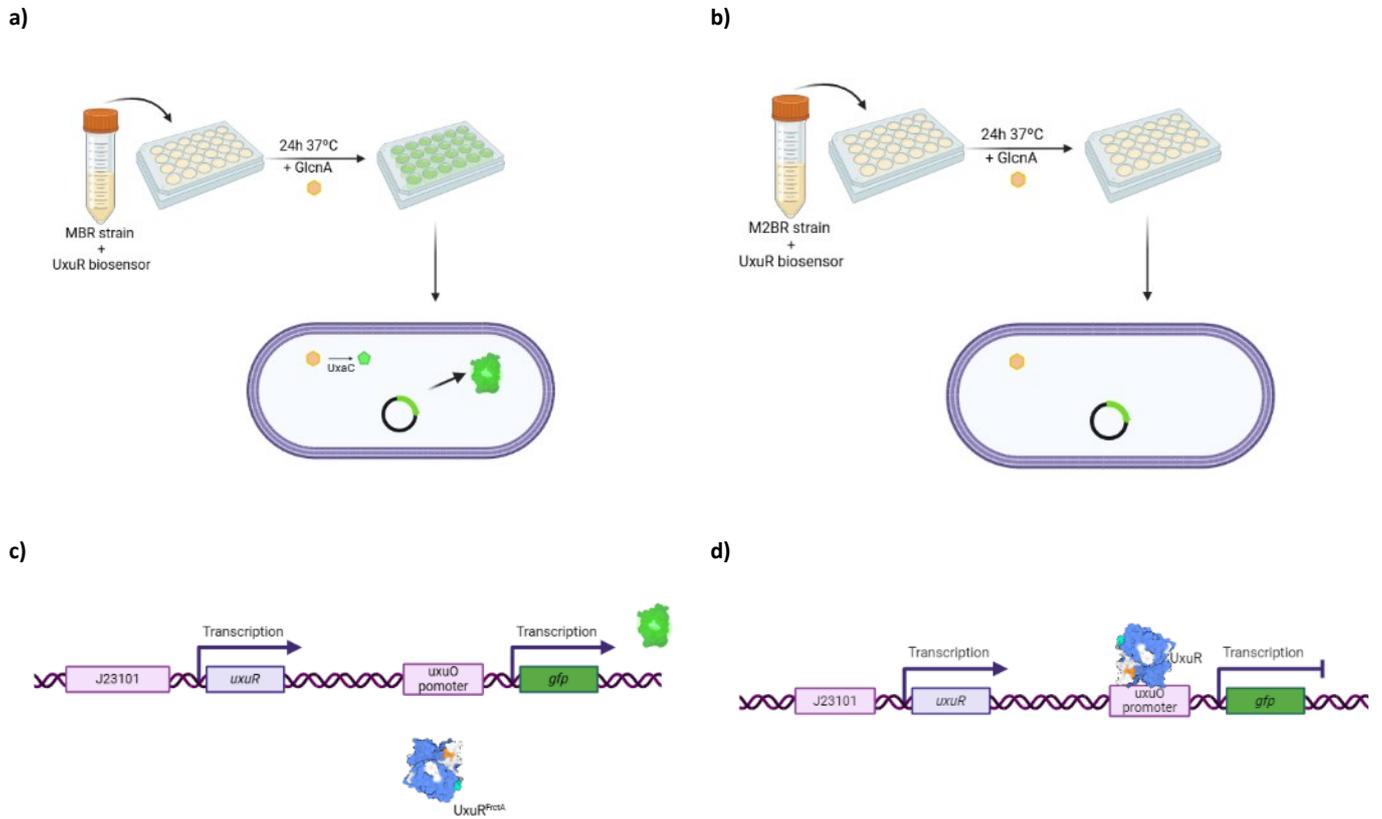
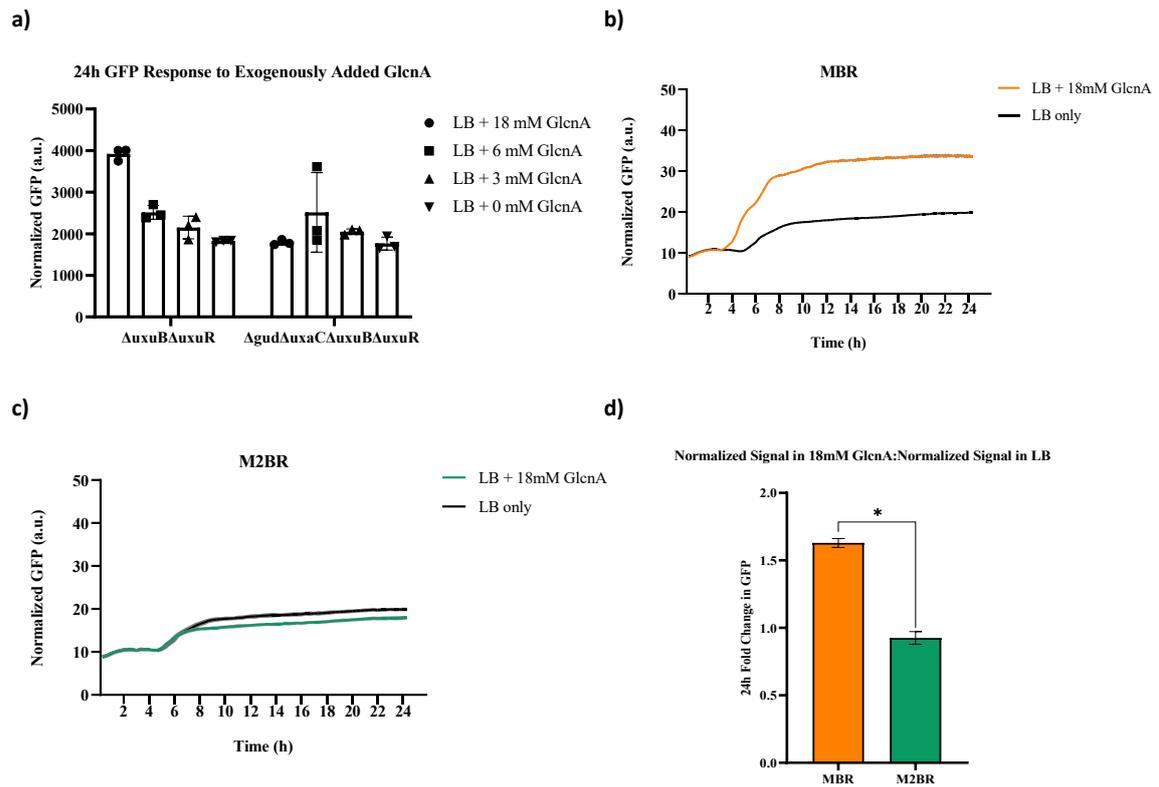


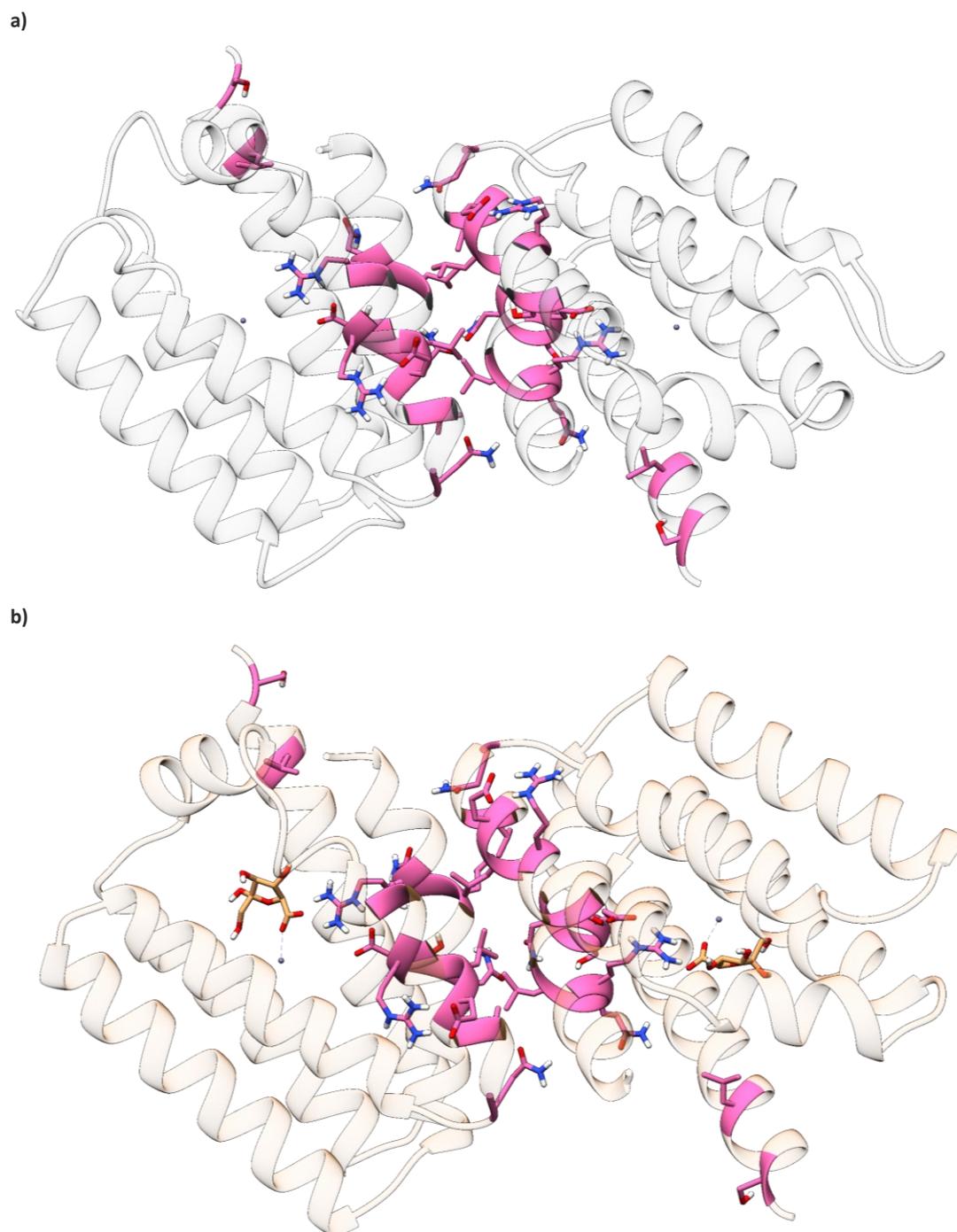
Fig. S11 Sequence alignment of *E. coli* UxuR and other FCD family members. The conserved and similar residues are indicated by red-shaded boxes and yellow-shaded boxes, respectively. The residues involved in Zn(II) binding in *E. coli* UxuR are indicated by orange stars. Blue triangles indicate all *in vitro* tested residues in this study.



**Fig. S12** Schematic representation of growth conditions for two strains: MBR -  $\Delta uxuB\Delta uxuR$  (a) and M2BR -  $\Delta gud\Delta uxaC\Delta uxuB\Delta uxuR$  (b), accompanied by a schematic representation of the WT biosensor response to external addition of D-glucuronic acid (GlcA) for both strains: MBR (c) and M2BR (d). Illustration in a) emphasizes the isomerization of D-glucuronic acid into D-fructuronic acid by the UxaC enzyme.

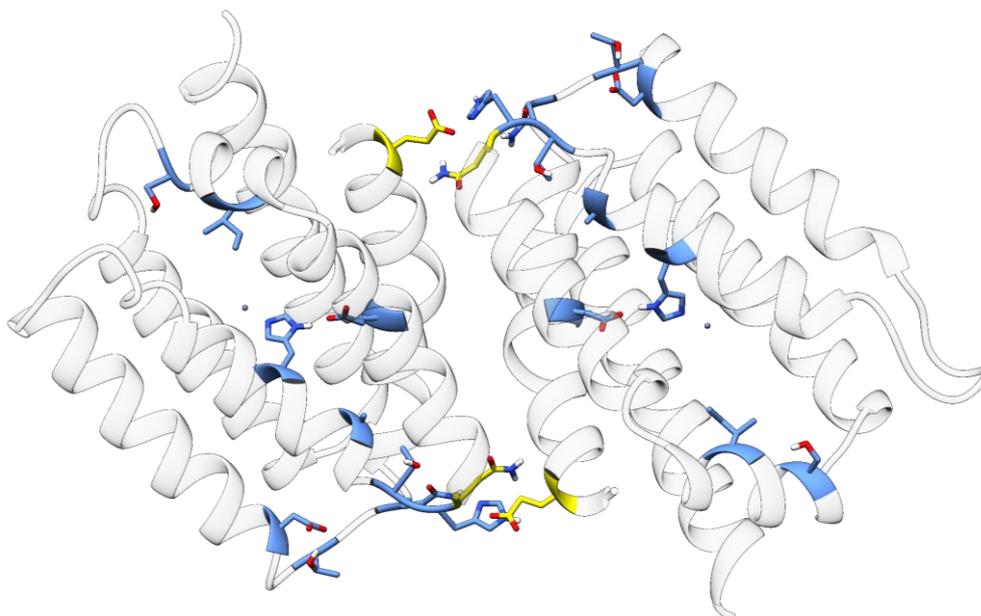


**Fig. S13** Normalized GFP signal in MBR ( $\Delta uxuB\Delta uxuR$ ) and M2BR ( $\Delta gudD\Delta uxaC\Delta uxuB\Delta uxuR$ ) strains. **a)** 24h GFP Response to MBR and M2BR strains transformed with WT UxuR biosensor in LB supplemented with different concentrations of D-glucuronic acid - GlcNA. Error bars represent the standard deviation of three biological replicates; **b)** 24h GFP Response to MBR strain transformed with WT UxuR biosensor in LB plus 18mM GlcNA and in LB only. Data shown is an average of three biological replicates with the standard deviation shown as the shaded region; **c)** 24h GFP Response to M2BR strain transformed with WT UxuR biosensor in LB plus 18mM GlcNA and LB only. Data shown is an average of three biological replicates with the standard deviation shown as the shaded region.; **d)** The fold change difference in normalized GFP signal in the *uxaC* gene knockout strain (M2BR) implies that the WT UxuR biosensor cannot function in the presence of only GlcNA but depends on its isomerization to fructuronic acid. Error bars represent the SEM of three biological replicates \* indicates normalized GFP significantly higher ( $p < 0.05$ ) in MBR when 18 mM GlcNA is supplemented.

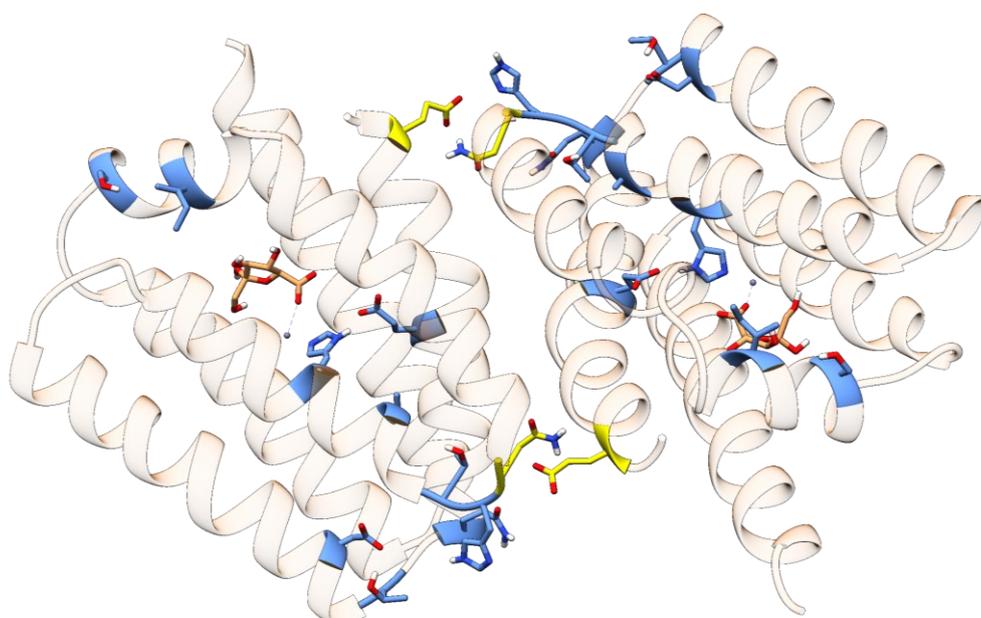


**Fig. S14.** Minimal energy (native) structure representation of the UxuR TF model obtained from the reweighted free energy landscape. **a)** UxuR in its Apo form and **b)** UxuR in complex with D-fructuronic acid (FrctA). The structures are illustrated from the top view, and the N-terminal domains and linkers were omitted for clear visualization of the C-terminal domain with focus on interactions Arg102-Glu106, Gln103-Ser107, Leu104-Asn108, Asn161-Leu164, Glu166-Arg169, Leu167-Ser171 and Leu231-Ser235 (pink residues). The FrctA is represented by an orange licorice structure and the Zn(II) ion is represented by a grey sphere.

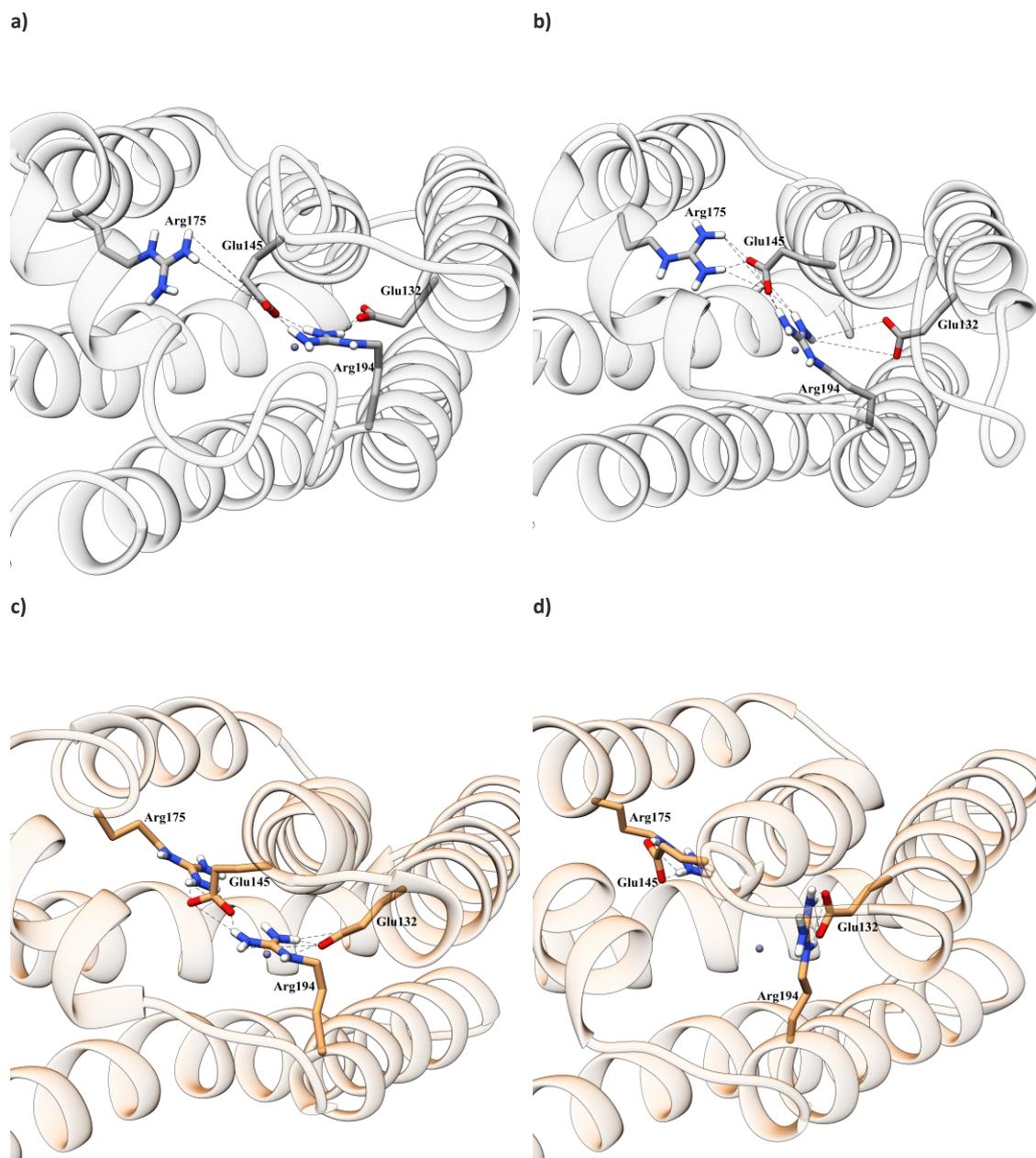
a)



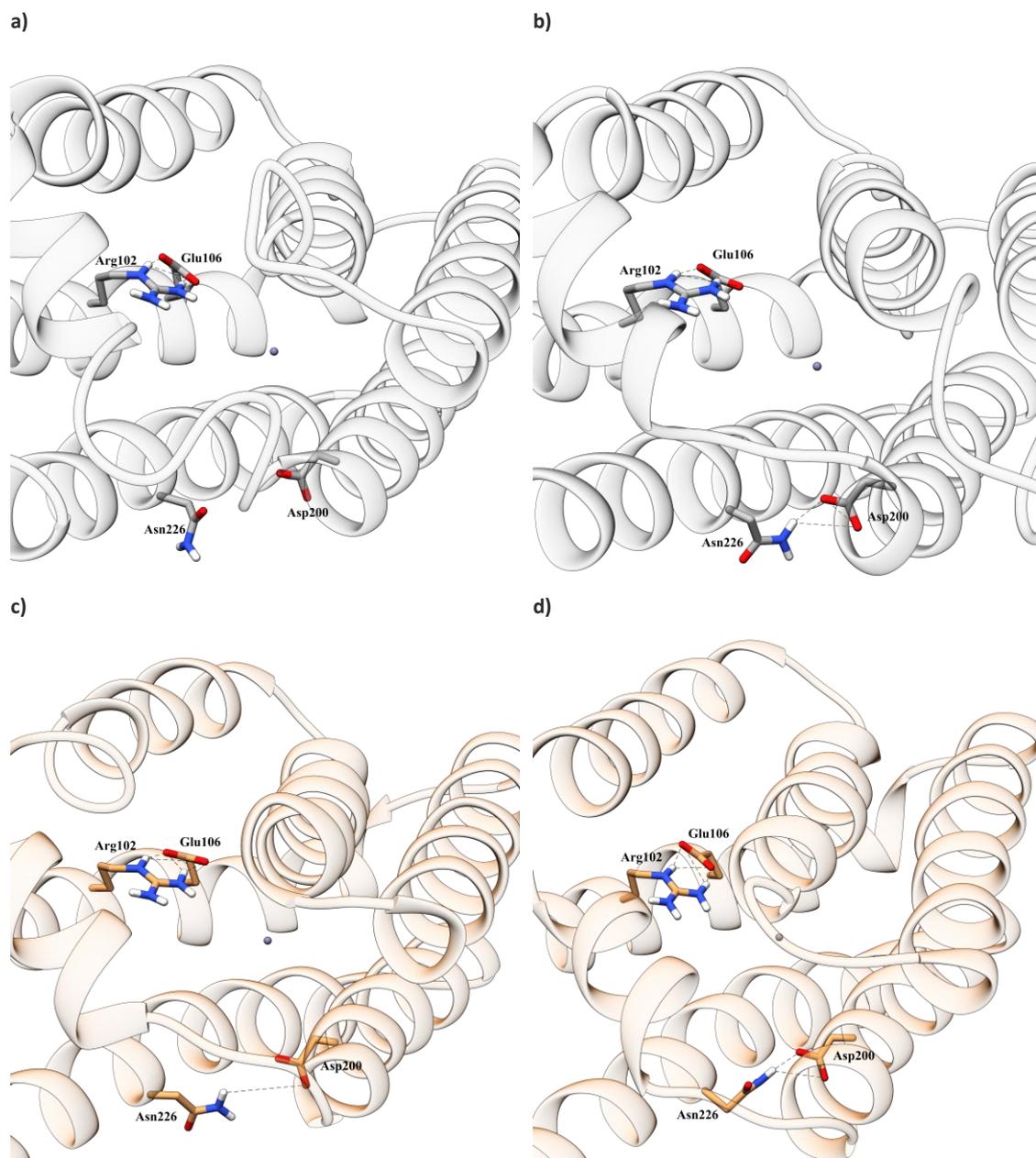
b)



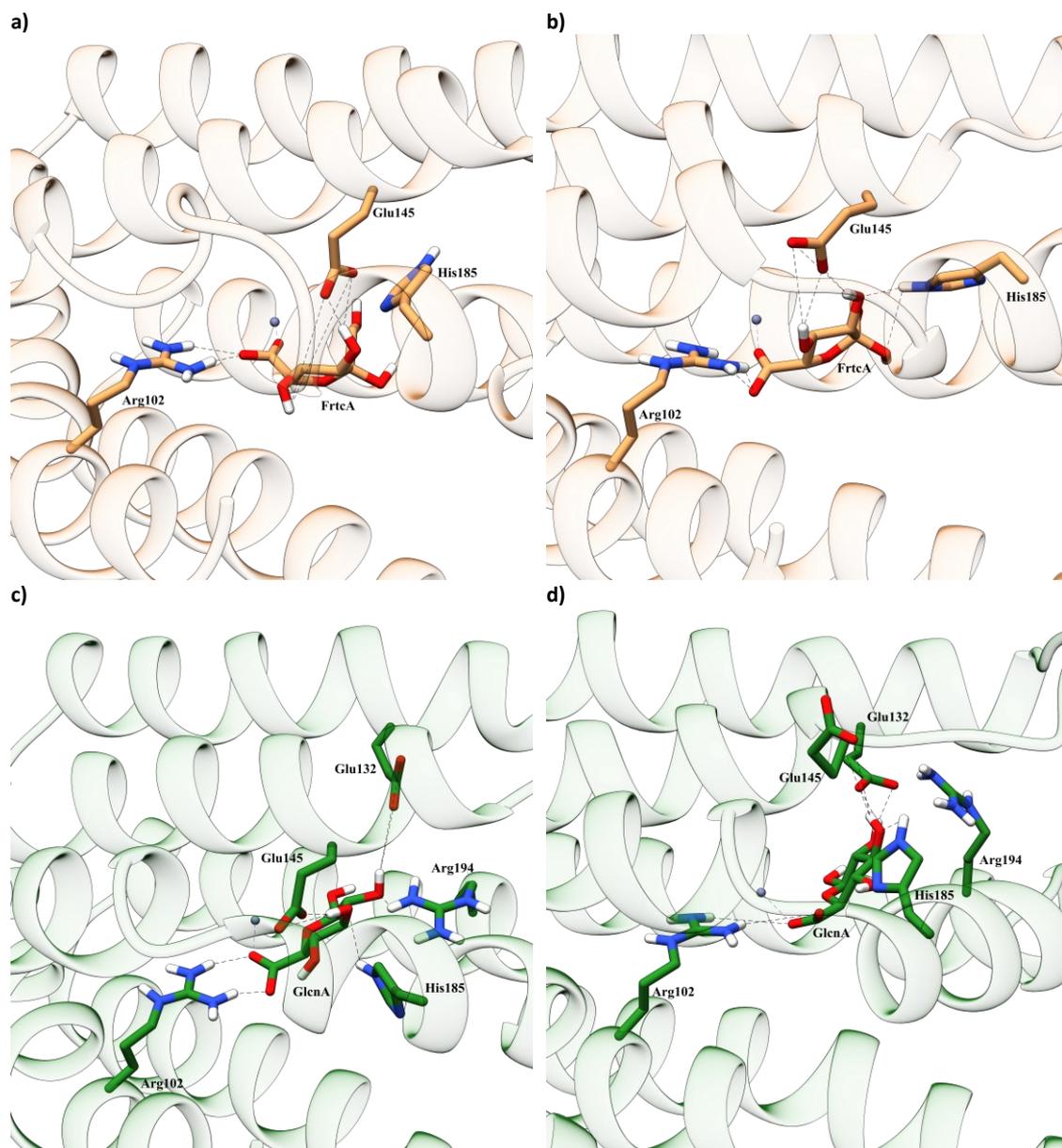
**Fig. S15.** Minimal energy (native) structure representation of the UxuR TF model obtained from the reweighted free energy landscape. **a)** UxuR in its Apo form and **b)** UxuR in complex with D-fructuronic acid (FrctA). The structures are illustrated from the top view, and the N-terminal domains and linkers were omitted for clear visualization of the C-terminal domain with focus on interactions Glu106-His152, Gln116-His160, Asp121-Thr118, Ala156-Thr159 and Ile182-Ser186 (blue residues) and Glu97-Asn161' interaction (yellow residues). The FrctA is represented by an orange licorice structure and the Zn(II) ion is represented by a grey sphere.



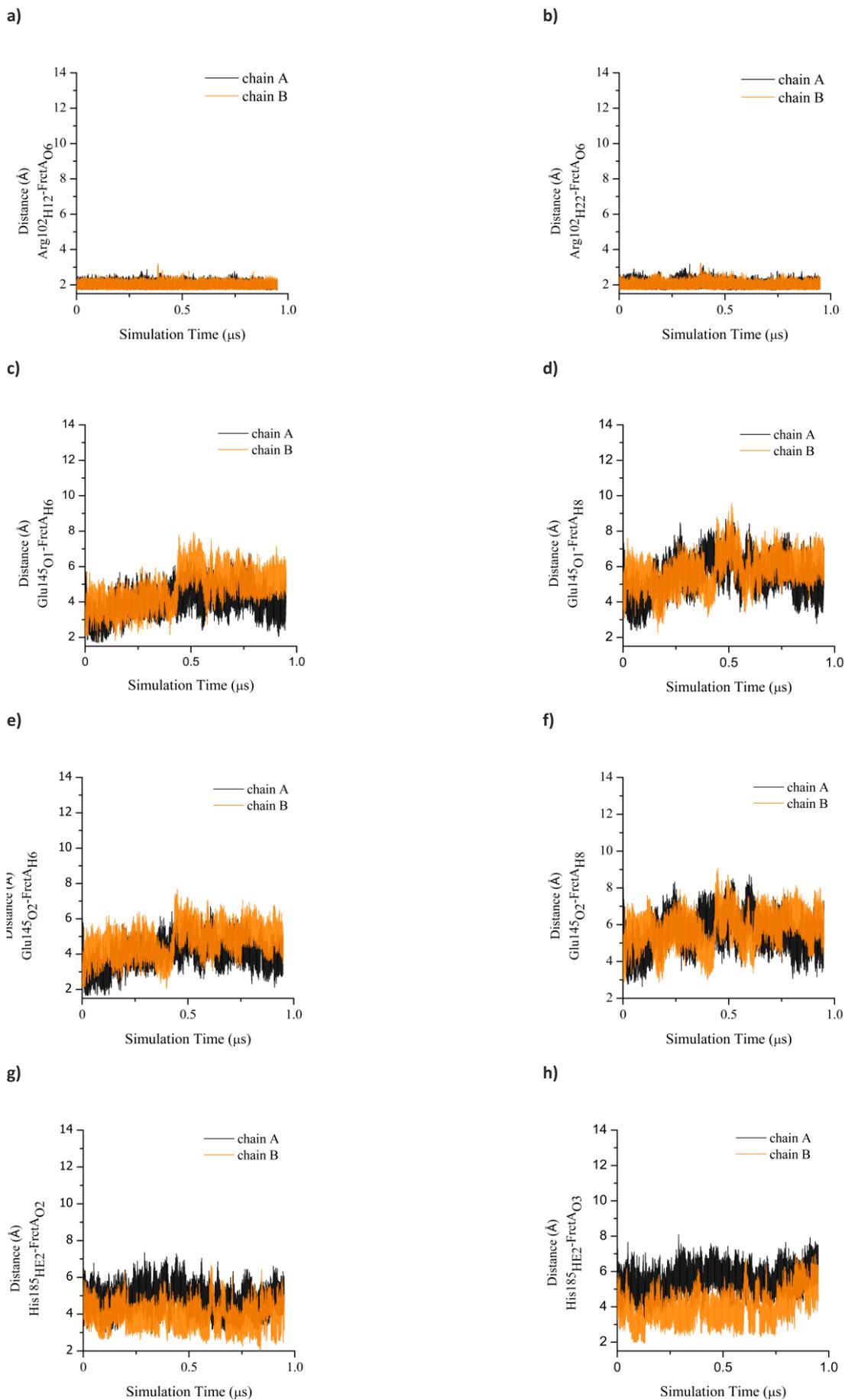
**Fig. S16.** Minimal energy (native) structure representation of the UxuR TF model obtained from the reweighted free energy landscape. UxuR in its free form (**a** and **b**) and UxuR in complex with D-fructuronic acid (FrctA, **c** and **d**), with focus on the Zn(II) ion and residues Glu132, Glu145, Arg175 and Arg194 at C-terminal domain. The FrctA was omitted from the structures **c**) and **d**) to facilitate visual inspection. The Zn(II) ion is represented by a grey sphere.



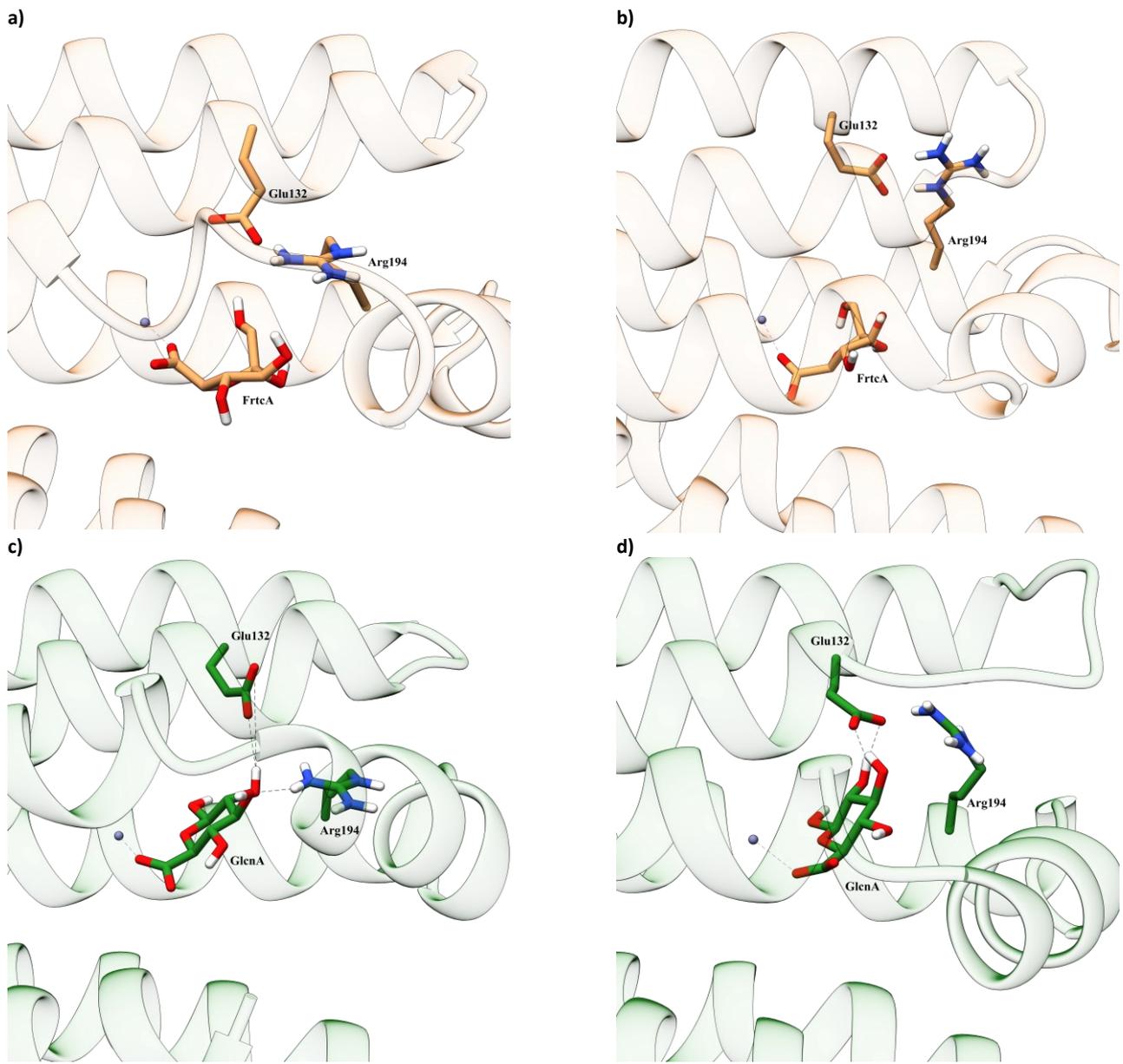
**Fig. S17.** Minimal energy (native) structure representation of the UxuR TF model obtained from the reweighted free energy landscape. UxuR in its free form (**a** and **b**) and UxuR in complex with D-fructuronic acid (**c** and **d**), with focus on the Zn(II) ion and residues Arg102, Glu106, Asp200 and Asn226 at C-terminal domain. The D-fructuronic acid was omitted from the structures **c**) and **d**) to facilitate visual inspection. The Zn(II) ion is represented by a grey sphere.



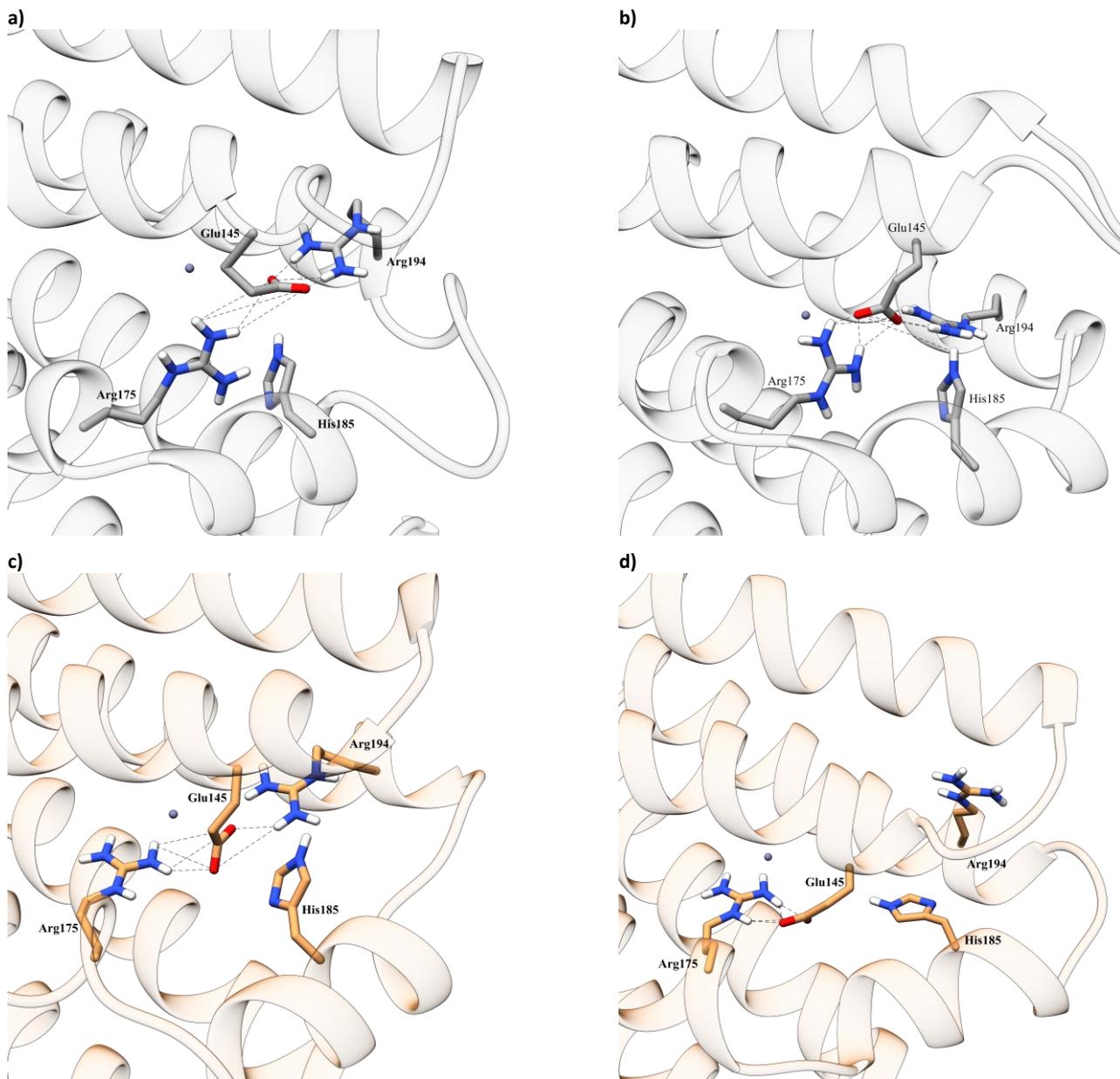
**Fig. S18.** Minimal energy (native) structure representation of the UxuR TF model obtained from the reweighted free energy landscape. UxuR in complex with D-fructuronic acid (FrctA) with focus on residues Arg102, Glu145 and His185 at C-terminal domain (**a** and **b**) and UxuR in complex with D-glucuronic acid (GlcA) with focus on residues Arg102, Glu132, Glu145, His185 and Arg194 at C-terminal domain (**c** and **d**). The FrctA and GlcA are represented by orange and green licorice structures, respectively. The Zn(II) ion is represented by a grey sphere.



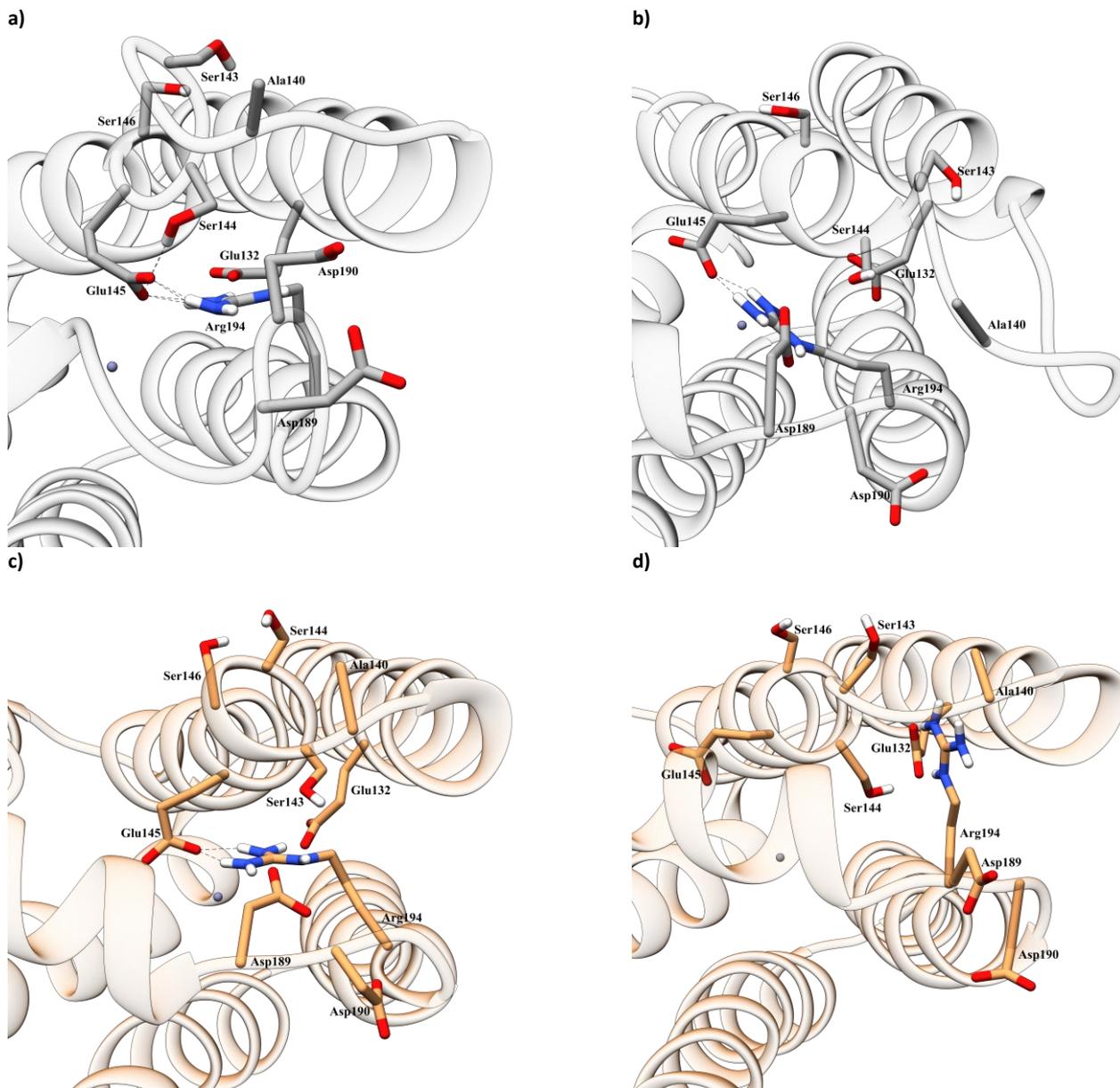
**Figure S19.** Average distances (Å) between D-fructuronic acid and the residues Arg102 (a and b), Glu15 (c - f) and His185 (g and h). The atomic coordinates were based on HB analyses.



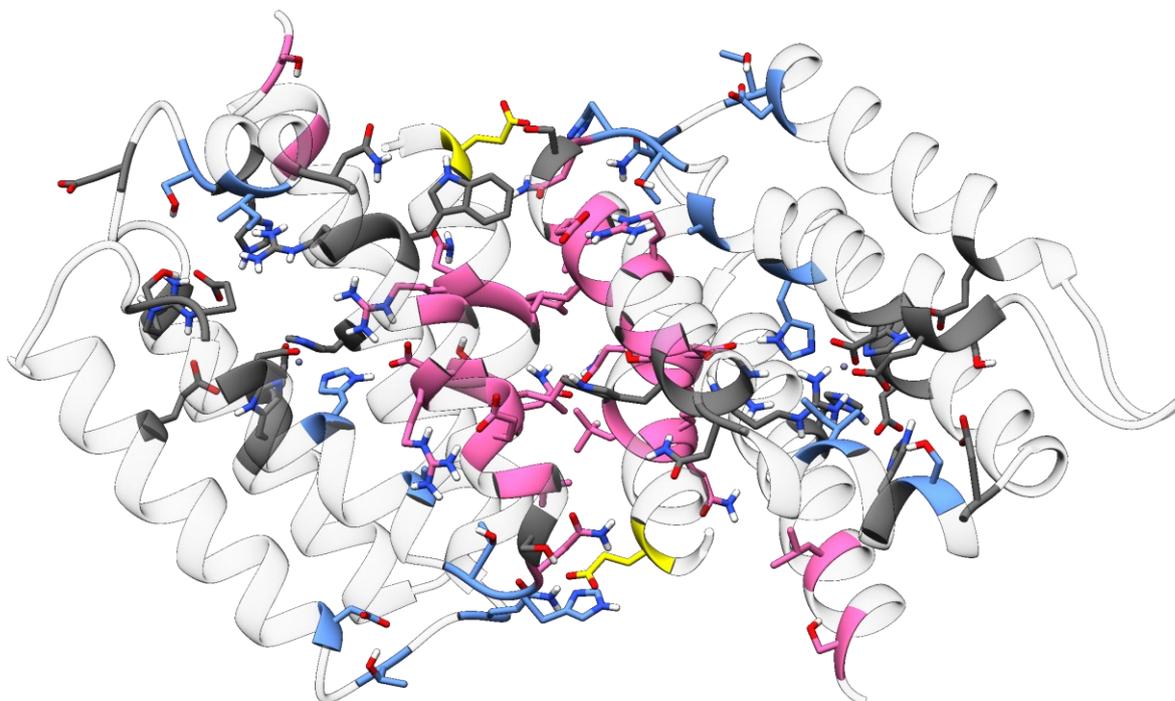
**Fig. S20.** Minimal energy (native) structure representation of the UxuR TF model obtained from the reweighted free energy landscape. UxuR in complex with D-fructuronic acid (FrctA, **a** and **b**) and UxuR in complex with D-glucuronic acid (GlcNA, **c** and **d**), with focus on residues Glu132 and Arg194 at C-terminal domain. The FrctA and GlcNA are represented by orange and green licorice structures, respectively. The Zn(II) ion is represented by a grey sphere.



**Fig. S21.** Minimal energy (native) structure representation of the UxuR TF model obtained from the reweighted free energy landscape. UxuR in its free form (**a** and **b**) and UxuR in complex with D-fructuronic acid (FctA, **c** and **d**), with focus on Zn(II)-binding site and residues Glu145, Arg175, His185 and Arg194 at C-terminal domain. The FrctA was omitted from the structures **c**) and **d**) to facilitate visual inspection. The Zn(II) ion is represented by a grey sphere.

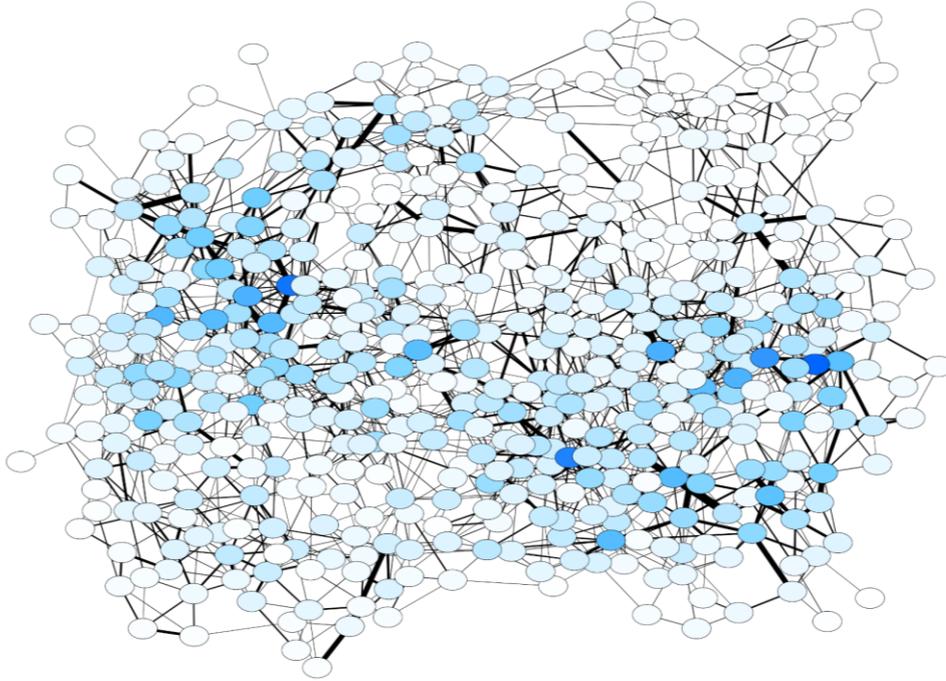


**Fig. S22.** Structure representation of the UxuR TF model obtained from the reweighted free energy landscape. UxuR in its free form (**a** and **b**) and UxuR in complex with D-fructuronic acid (FrctA, **c** and **d**), with focus on Arg194 at the C-terminal domain. The FrctA was omitted from the structures **c**) and **d**) to facilitate visual inspection. The Zn(II) ion is represented by a grey sphere.

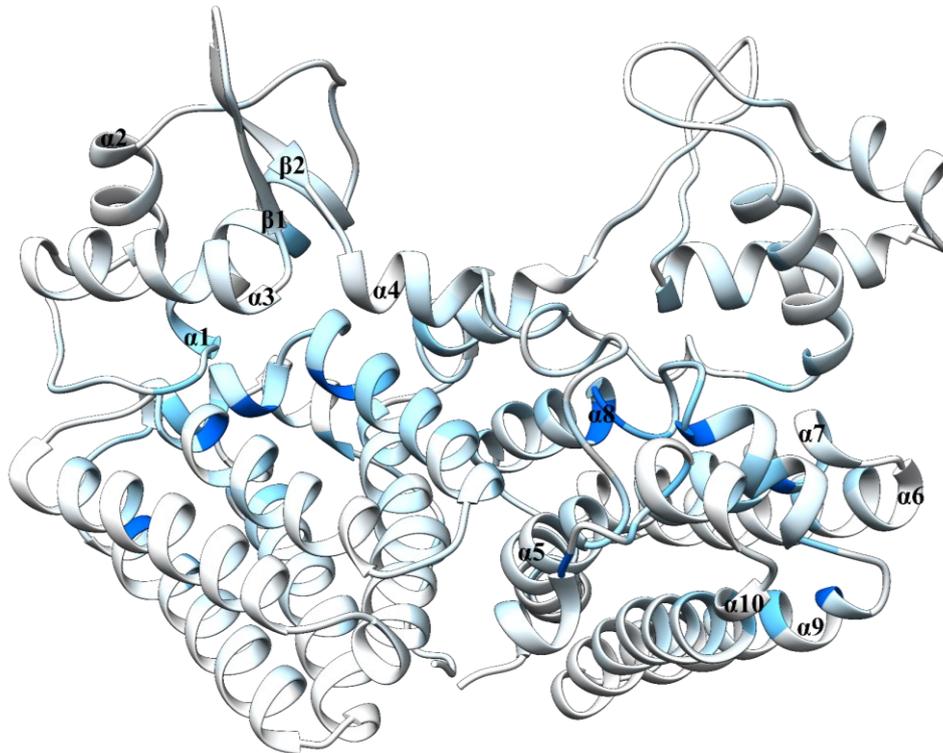


**Fig. S23.** Minimal energy (native) structure representation of the UxuR TF model in its free form, obtained from the reweighted free energy landscape. The structures are illustrated from the top view, and the N-terminal domains and linkers were omitted for clear visualization of the C-terminal domain with a focus on residues Glu97, Arg102, Gln103, Leu104, Glu106, Ser107, Asn108, Gln116, Thr118, Asp121, Glu132, Ser144, Glu145, Ala156, Thr159, His160, Asn161, Ser162, Leu164, Glu166, Leu167, Arg169, Ser171, Trp174, Arg175, Asn178, Ile182, His185, Ser186, Asp189, Arg194, Asp200, Leu231, and Ser235. The Zn(II) ion is represented by a grey sphere.

a)

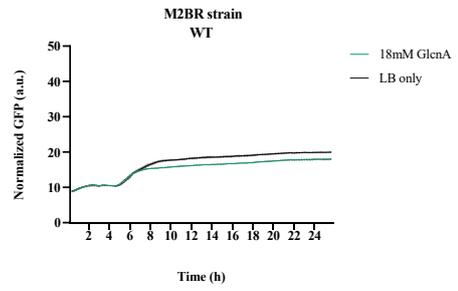
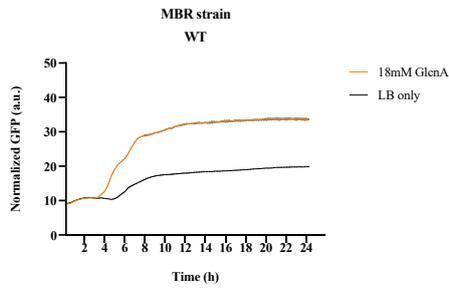


b)

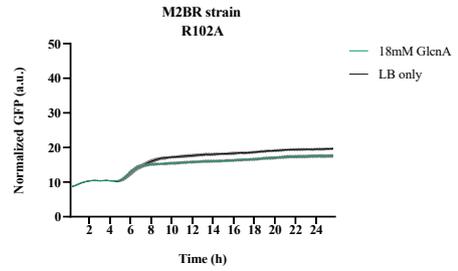
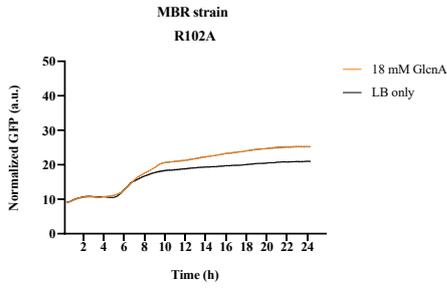


**Figure S24.** (a) Allosteric predictions of the NCF model mapped to the UxuR model TF in complex with D-fructuronic acid - UxuR<sup>FrctA</sup>. Nodes colored from low (white) to high (red) scores. For visual clarity, only edges occurring in 10% of simulation time are shown. (b) Representation of the native structure from the reweight free energy landscape of UxuR<sup>FrctA</sup>, colored (white to blue) according to the z-score normalized DNCF scores.

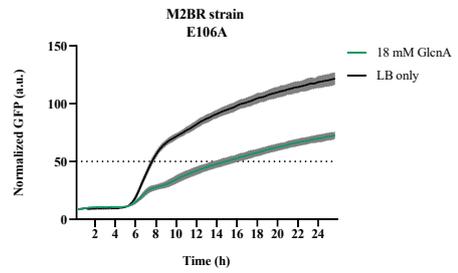
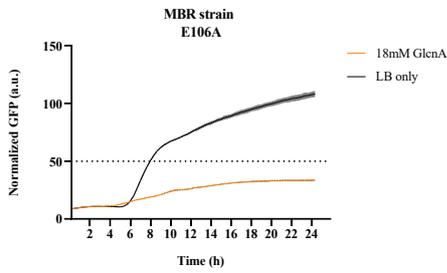
a)



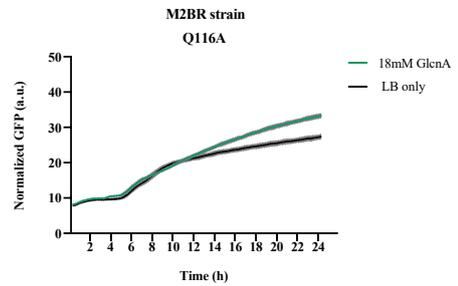
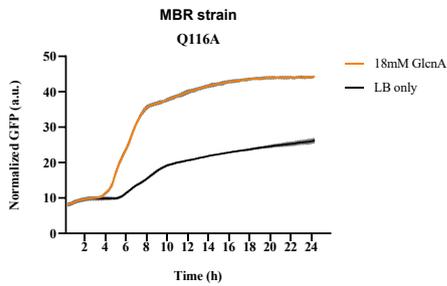
b)



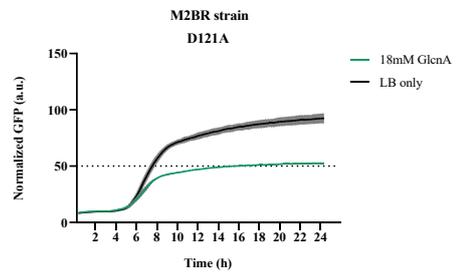
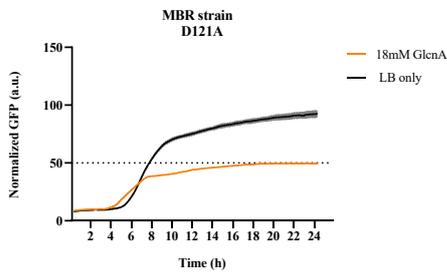
c)



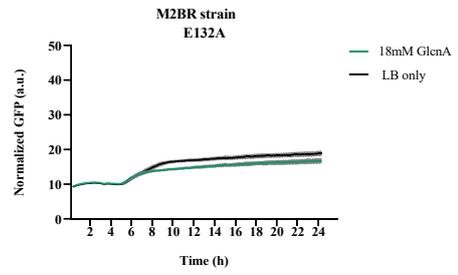
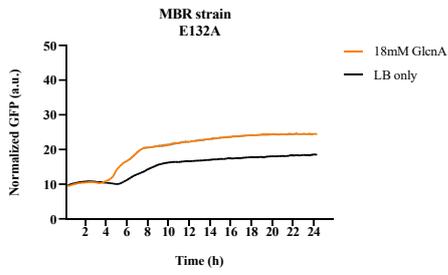
d)



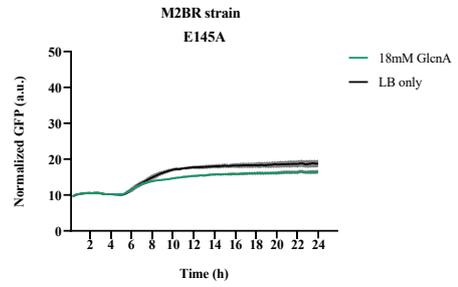
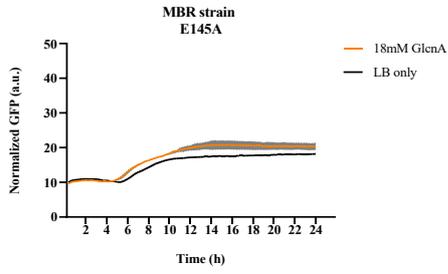
e)



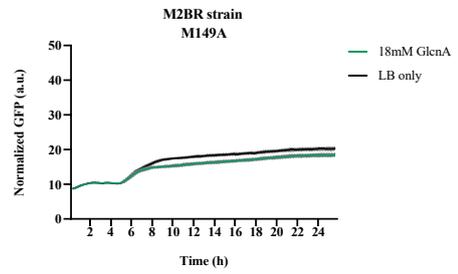
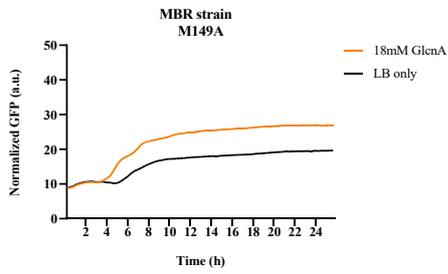
f)



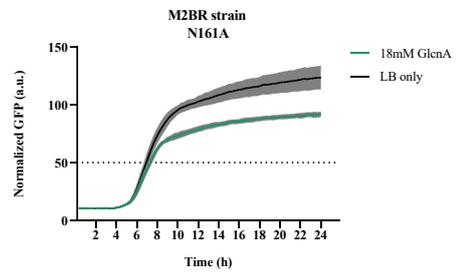
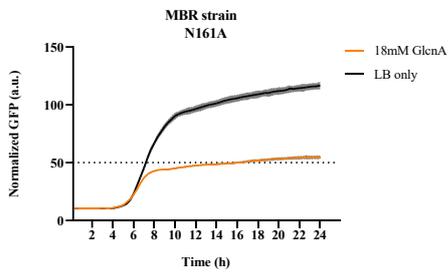
g)



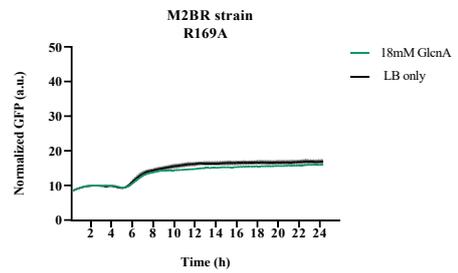
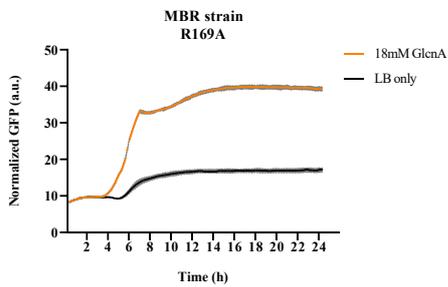
h)



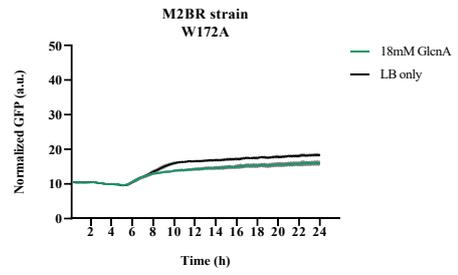
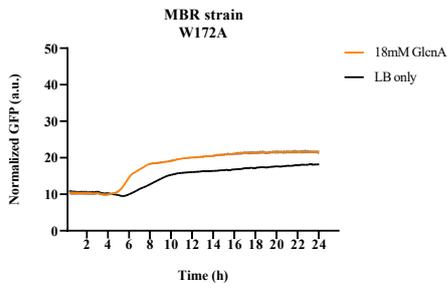
i)



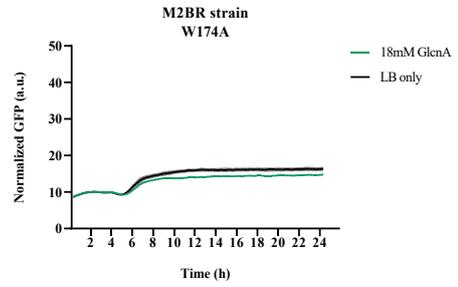
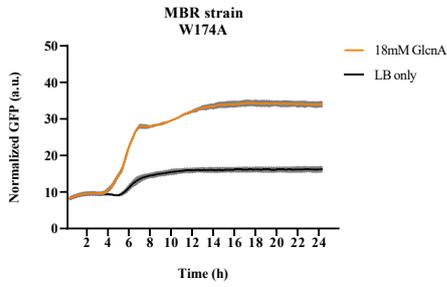
j)



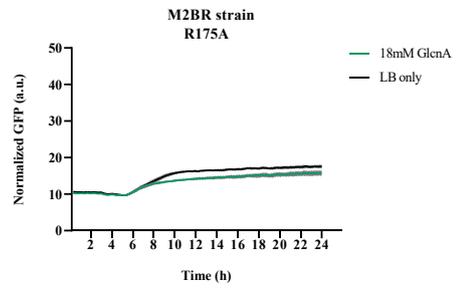
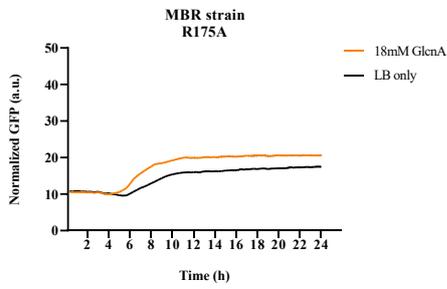
k)



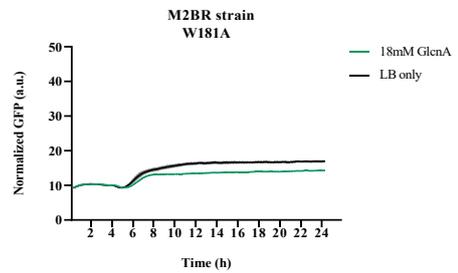
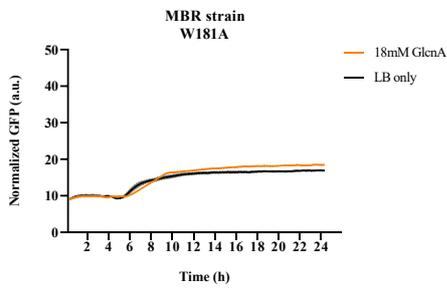
l)



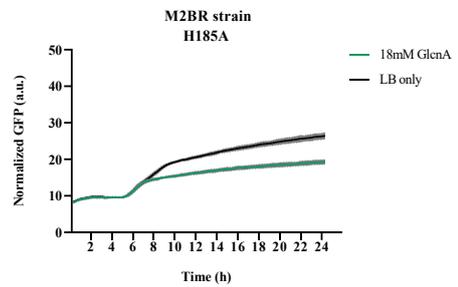
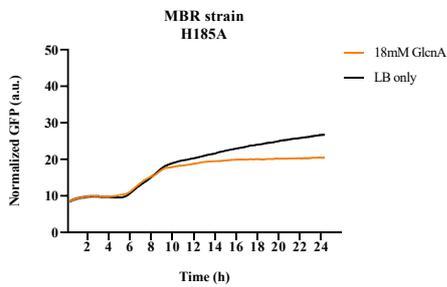
m)



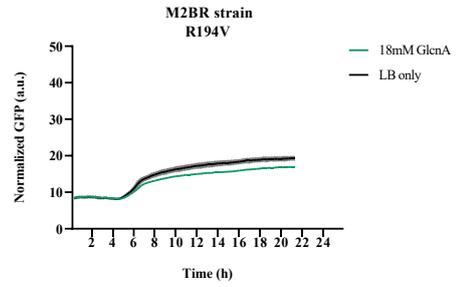
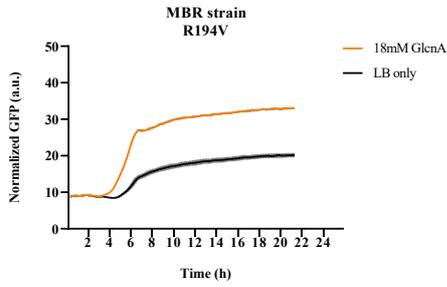
n)



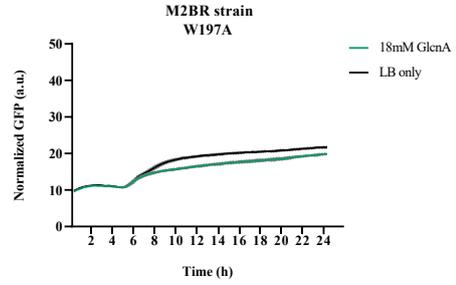
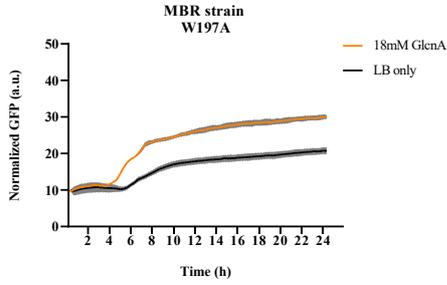
o)



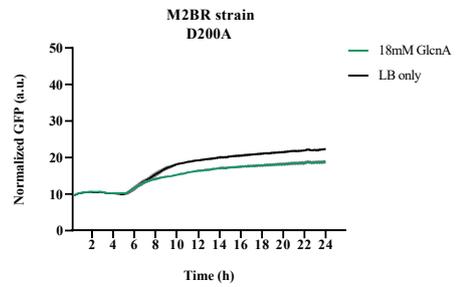
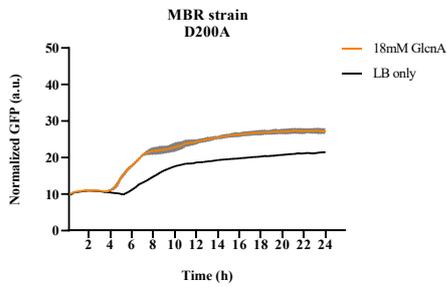
p)



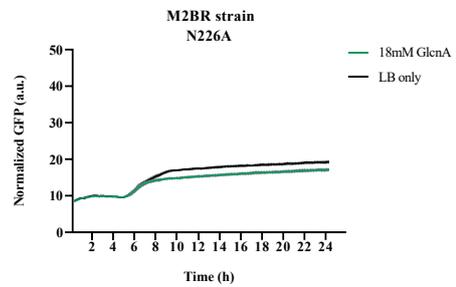
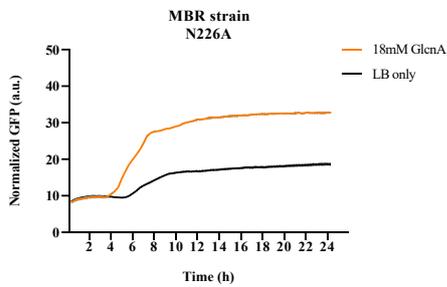
q)



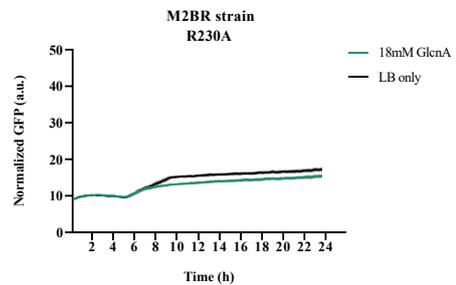
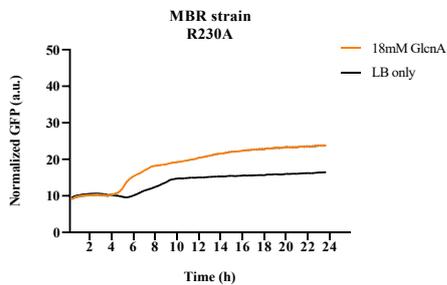
r)



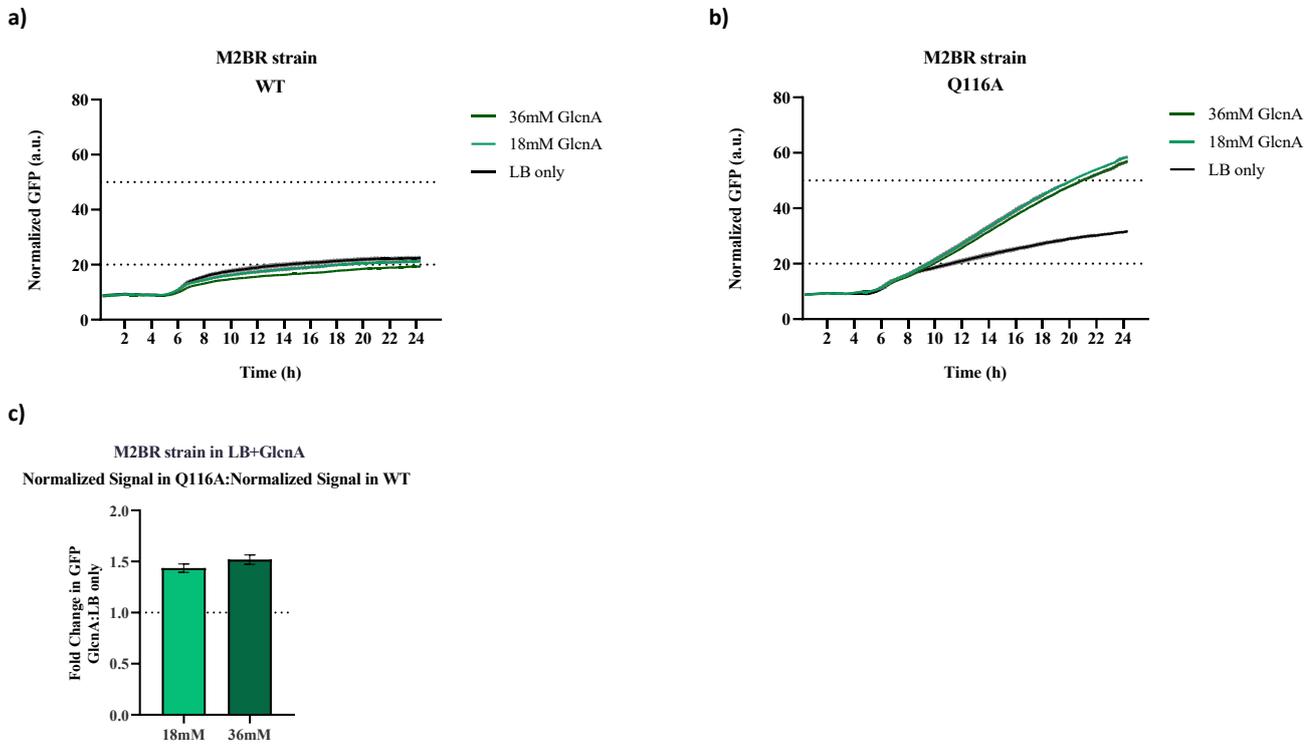
s)



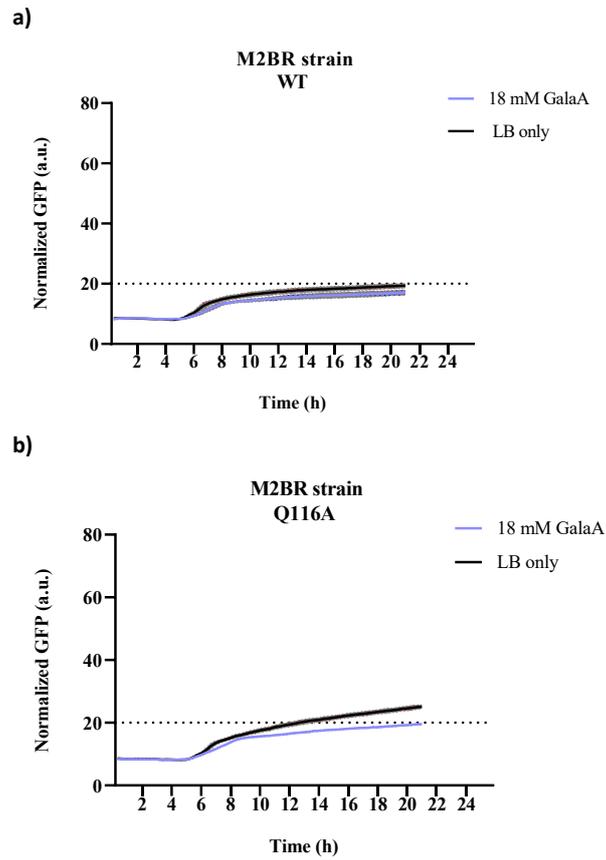
t)



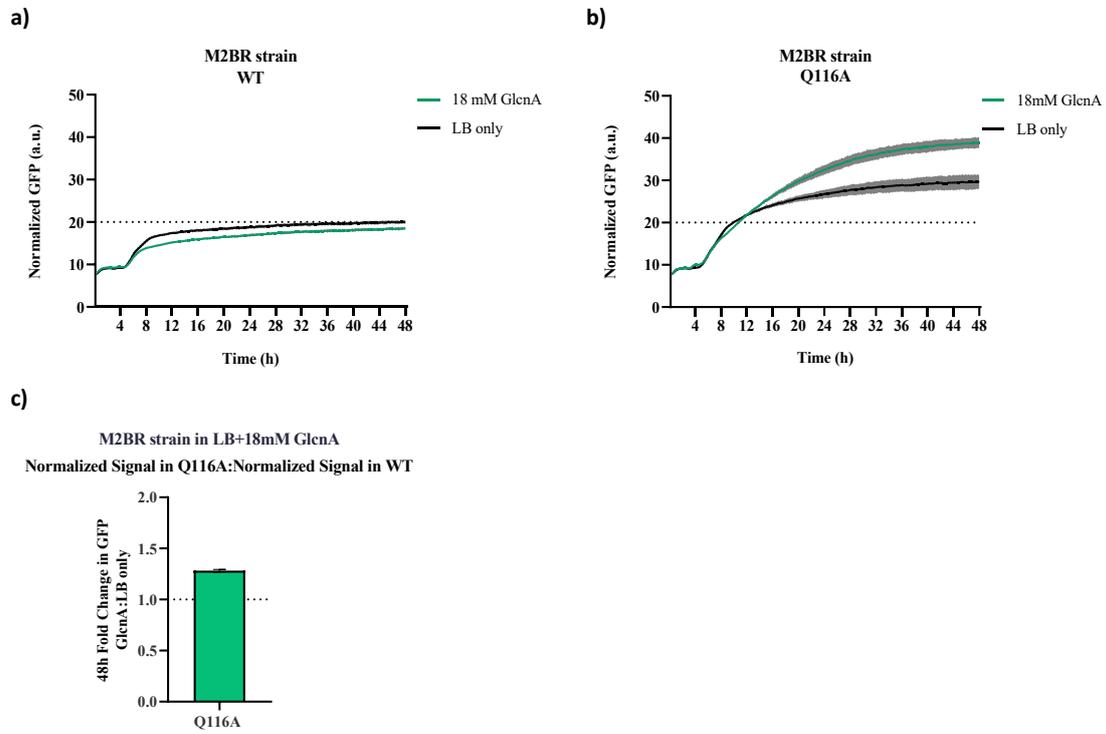
**Fig. S25.** Normalized GFP expression following induction by 18 mM of D-glucuronic acid (GlcNA) in two strains: first column - MBR ( $\Delta luxuB\Delta luxuR$ ) and second column - M2BR ( $\Delta gud\Delta luxuA\Delta luxuB\Delta luxuR$ ). The data shown is an average of three biological replicates with the standard deviation shown as the shaded region.



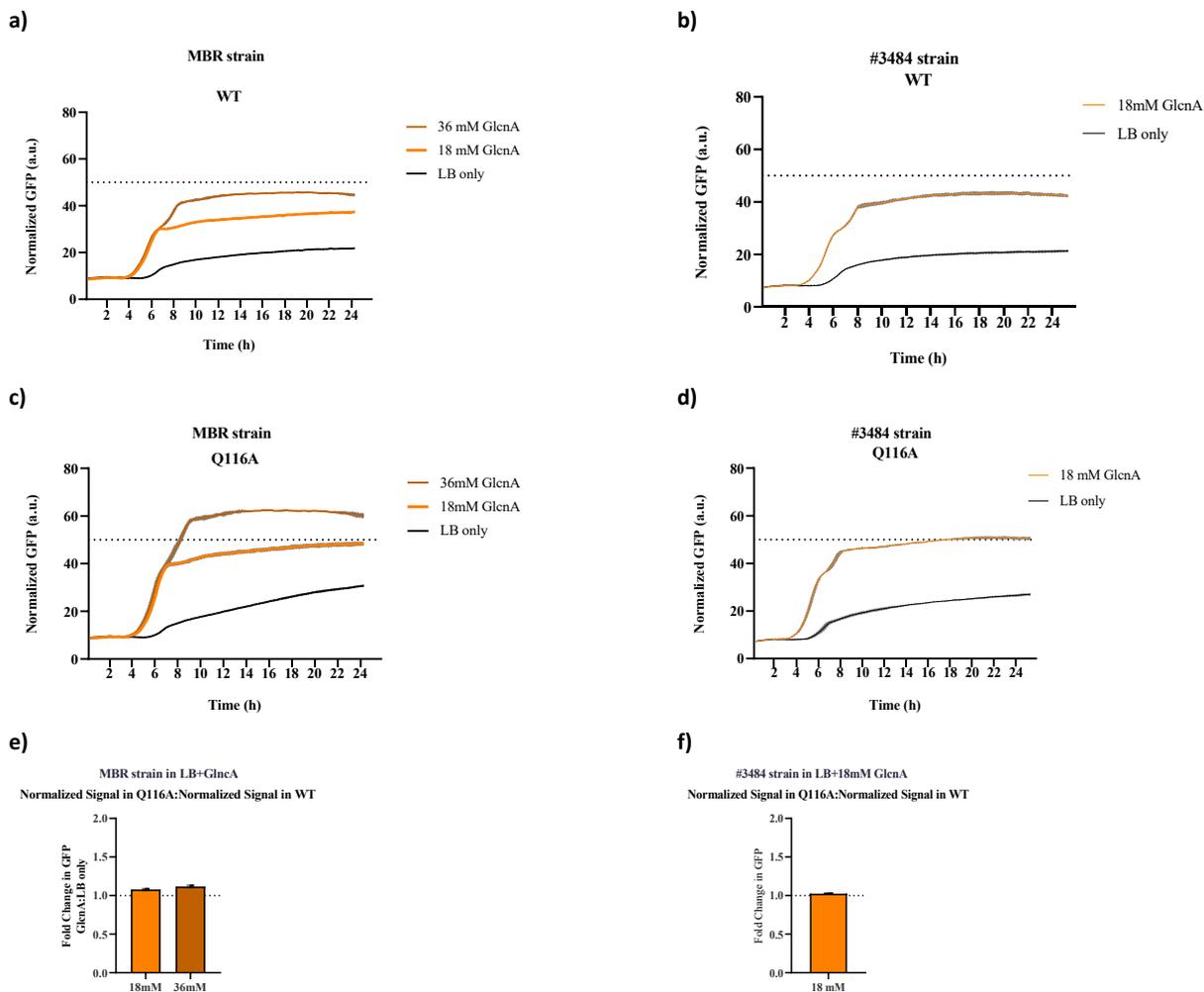
**Fig. S26** Normalized GFP expression following induction by 18 mM and 36 mM of D-glucuronic acid (GlcNA) in M2BR ( $\Delta gud\Delta uxaC\Delta uxuB\Delta uxuR$ ) strain transformed with WT biosensor (**a**) and with the Q116A variant biosensor (**b**). The data shown is an average of three biological replicates with the standard deviation shown as the shade region. 24h Fold change in normalized GFP in the presence of LB supplemented with 18 mM (first bar) and 36 mM (second bar) of GlcNA (**c**). Error bars represent  $\pm 1$  SEM from the mean of three replicate cultures.



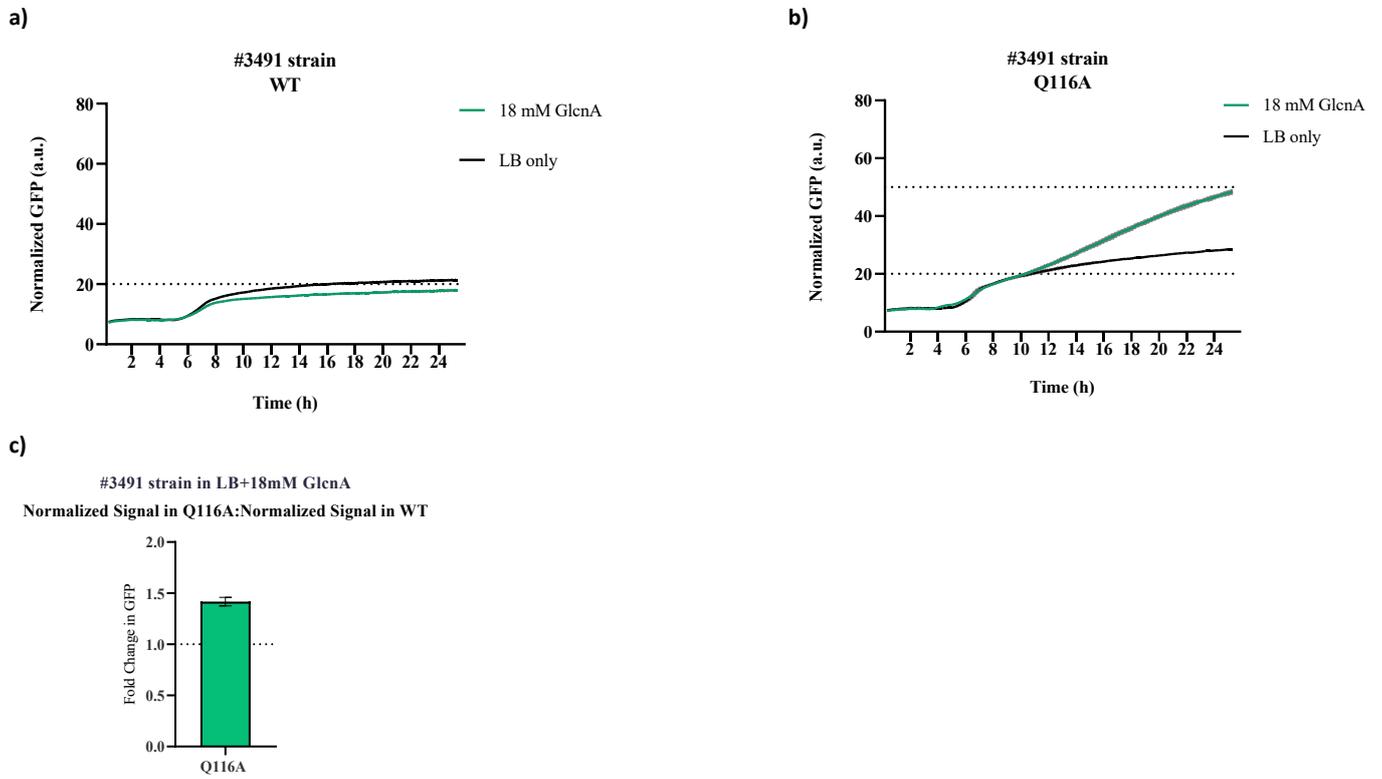
**Fig. S27** Normalized GFP expression following induction by 18 mM of D-galacturonic acid (GalaA) in M2BR ( $\Delta gud\Delta uxaC\Delta uxuB\Delta uxuR$ ) strain transformed with WT biosensor (**a**) and with the Q116A variant biosensor (**b**). The data shown is an average of three biological replicates with the standard deviation shown as the shade region.



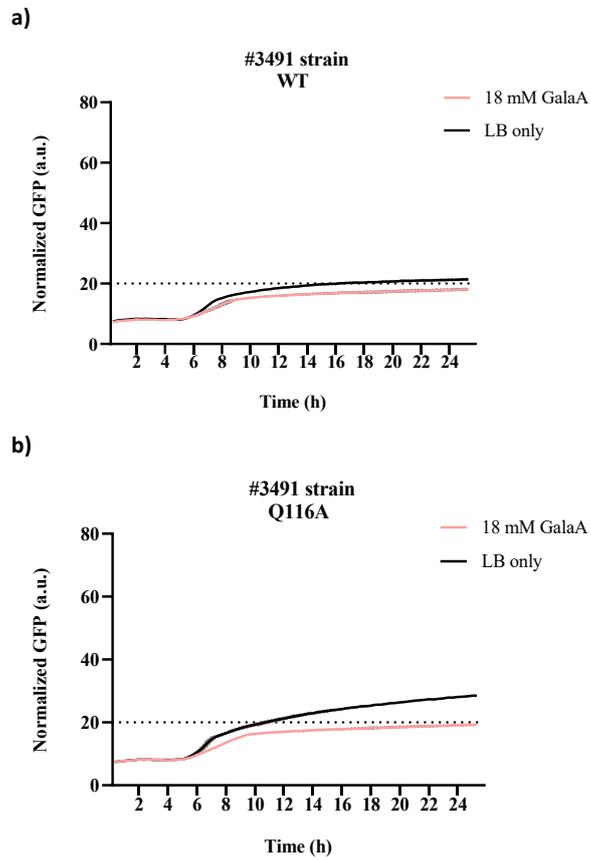
**Fig. S28** Normalized GFP expression following induction by 18 mM of D-glucuronic acid (GlcNA) in M2BR ( $\Delta gud\Delta uxa\Delta C\Delta uxuB\Delta uxuR$ ) strain transformed with WT biosensor (a) and with the Q116A variant biosensor (b). The data shown is an average of three biological replicates with the standard deviation shown as the shade region. 48h Fold change in normalized GFP in the presence of LB supplemented with 18 mM of GlcNA (c). Error bars represent  $\pm 1$  SEM from the mean of three replicate cultures.



**Fig. S29** Normalized GFP expression following induction by 18 mM and 36 mM of D-glucuronic acid (GlcA) in MBR strain ( $\Delta luxuB\Delta luxuR$ ) transformed with WT biosensor (a) and with the Q116A variant biosensor (c). Normalized GFP expression following induction by 18 mM of GlcA in #3484 strain ( $\Delta luxuB\Delta luxuR\Delta dexuR$ ) transformed with WT biosensor (b) and with the Q116A variant biosensor (d). The data shown is an average of three biological replicates with the standard deviation shown as the shade region. 24h Fold change in normalized GFP in the presence of LB supplemented with 18 mM (first bar) and 36 mM (second bar) of GlcA in the transformed MBR strain (e). 24h Fold change in normalized GFP in the presence of LB supplemented with 18 mM of GlcA in the transformed #3484 strain (f) Error bars represent  $\pm 1$  SEM from the mean of three replicate cultures.

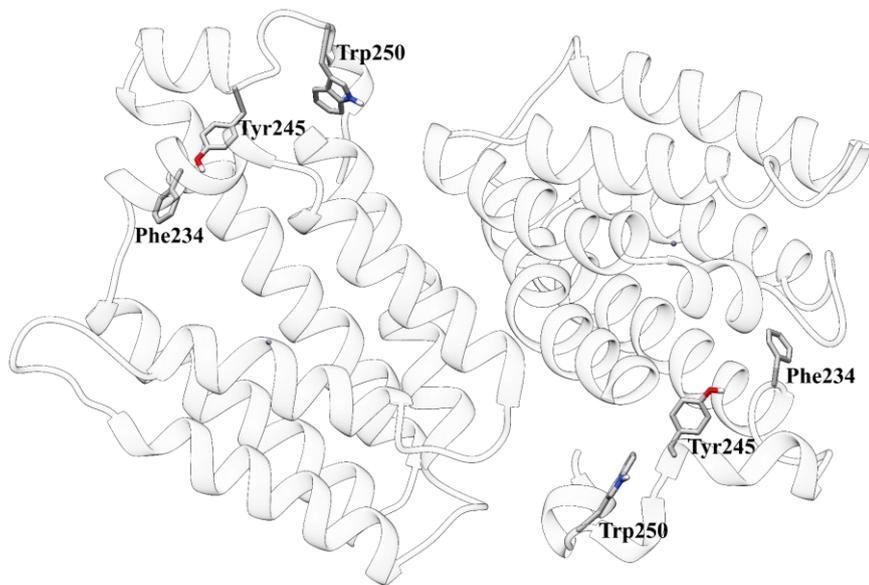


**Fig. S30** Normalized GFP expression following induction by 18 mM of D-glucuronic acid (GlcA) in #3491 strain ( $\Delta uxaC\Delta uxuR\Delta exuR$ ) strain transformed with WT biosensor (a) and with the Q116A variant biosensor (b). The data shown is an average of three biological replicates with the standard deviation shown as the shade region. 24h Fold change in normalized GFP in the presence of LB supplemented with 18 mM of GlcA (c). Error bars represent  $\pm 1$  SEM from the mean of three replicate cultures.

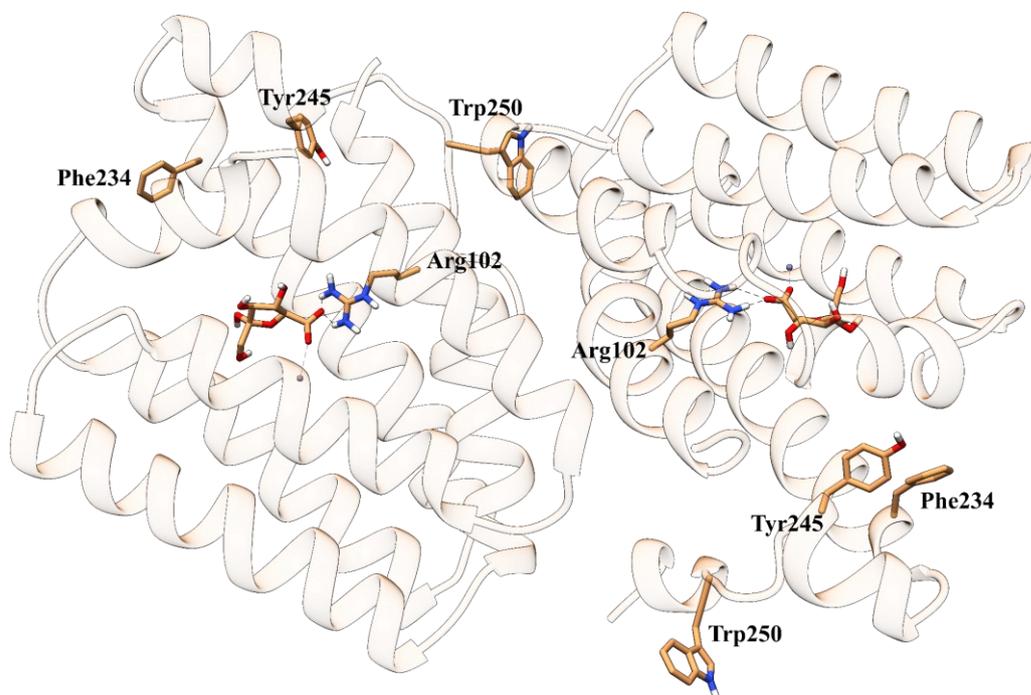


**Fig. S31** Normalized GFP expression following induction by 18 mM of D-galacturonic acid (GalaA) in #3491 strain ( $\Delta uxaC\Delta uxuR\Delta exuR$ ) strain transformed with WT biosensor (a) and with the Q116A variant biosensor (b). The data shown is an average of three biological replicates with the standard deviation shown as the shade region. 24h Fold change in normalized GFP in the presence of LB supplemented with 18 mM of GalaA (c). Error bars represent  $\pm 1$  SEM from the mean of three replicate cultures.

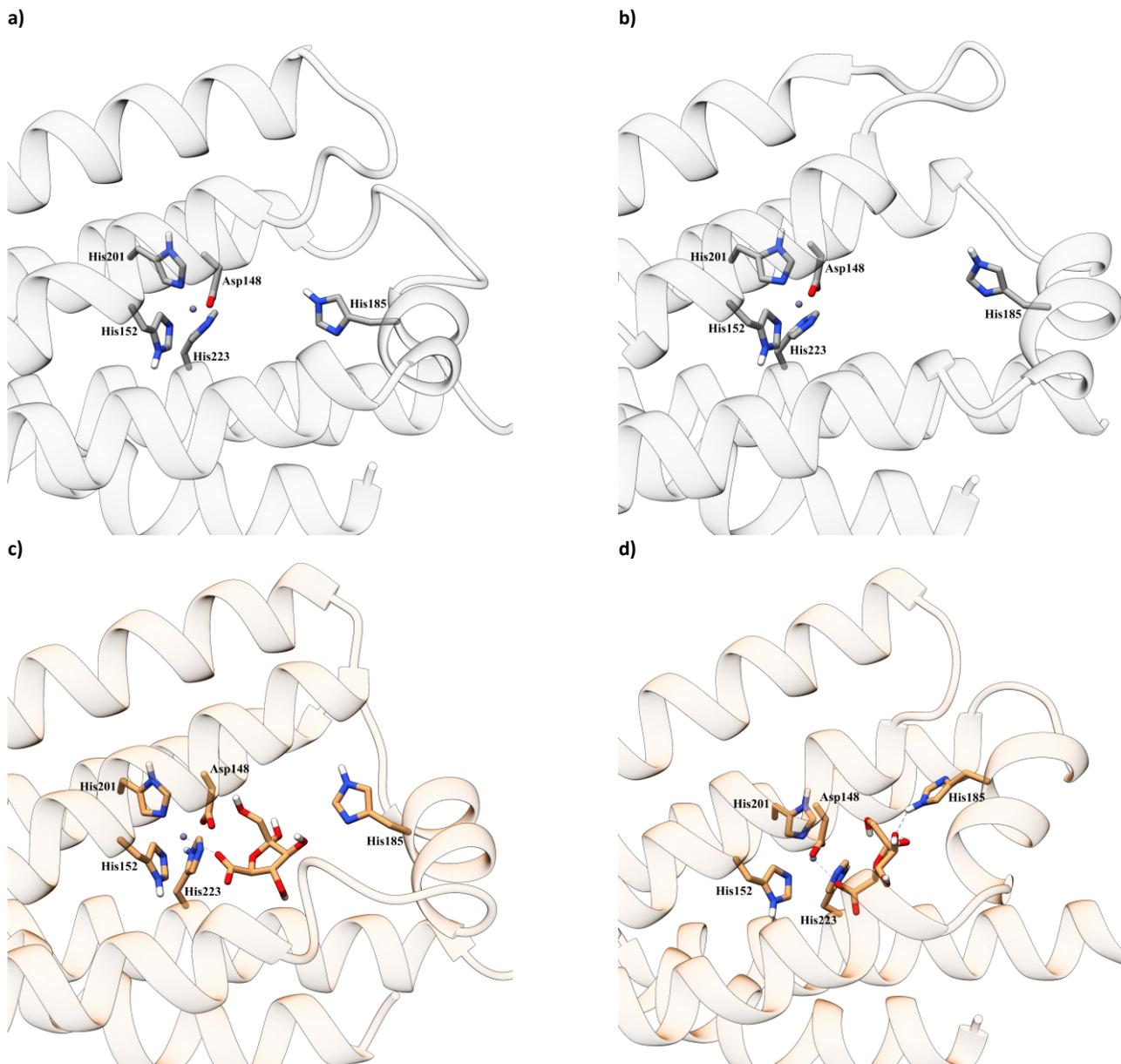
a)



b)



**Fig. S32.** Minimal energy (native) structure representation of the UxuR TF model obtained from the reweighted free energy landscape. UxuR in its free form (a) and UxuR in complex with D-fructuronic acid (b). The structures are illustrated from the top view, and the N-terminal domains and linkers were omitted for clear visualization of the C-terminal domain with a focus on residues Arg102, Phe234, Tyr245 and Trp250. The FrctA is represented by an orange licorice structure and the Zn(II) ion is represented by a grey sphere.



**Figure S33.** Minimal energy (native) structure representation of the UxuR TF model obtained from the reweighted free energy landscape. UxuR in its free form (**a** and **b**) and UxuR in complex with D-fructuronic acid (FctA, **c** and **d**), with focus on Zn(II)-binding site and residue His185 at C-terminal domain. The FrctA is represented by an orange licorice structure and the Zn(II) ion is represented by a grey sphere.

## Table S1-S21

**Table S1.** Molecular Dynamic Simulations details.

State	Content	Simulation time		Replicates
		cMD	GaMD	
UxuR <sup>Apo</sup>	Free form	10 ns	2000 ns	5
UxuR <sup>FrctA</sup>	D-fructuronate	10 ns	1000 ns	5
UxuR <sup>GlcNA</sup>	D-glucuronate	10 ns	1000 ns	5

**Table S2.** Strains used in this study.

Strain ID	Relevant genotype	Description	Reference
DH5 $\alpha$	<i>F</i> <sup>-</sup> $\phi$ 80 <i>lacZ</i> $\Delta$ M15 $\Delta$ ( <i>lacZYA-argF</i> )U169 <i>recA1 endA1</i> <i>hsdR17</i> ( <i>r<sub>K</sub><sup>-</sup>, m<sub>K</sub><sup>+</sup>) <i>phoA</i> <i>supE44</i> <math>\lambda</math>-<i>thi-1</i> <i>gyrA96</i> <i>relA1</i></i>	Cloning strain.	Prather Lab
#MBR	$\Delta$ <i>uxuB</i> $\Delta$ <i>uxuR</i>	Removing <i>uxuB</i> prevents the consumption of DFU and removing <i>uxuR</i> prevents interference from endogenous UxuR.	1
#M2BR	$\Delta$ <i>uxaC</i> $\Delta$ <i>uxuR</i>	Removing <i>uxaC</i> prevents the isomerization of GLU into DFU and removing <i>uxuR</i> prevents interference from endogenous UxuR.	1
#3484	$\Delta$ <i>uxuB</i> $\Delta$ <i>exuR</i> $\Delta$ <i>uxuR</i>	Removing <i>uxuB</i> prevents the consumption of DFU and removing <i>exuR</i> and <i>uxuR</i> prevents interference from endogenous ExuR and UxuR, respectively.	Prather Lab
#3491	$\Delta$ <i>uxaC</i> $\Delta$ <i>uxuR</i> $\Delta$ <i>exuR</i>	Removing <i>uxaC</i> prevents the isomerization of GLU into DFU and, removing <i>exuR</i> and <i>uxuR</i> prevents interference from endogenous ExuR and UxuR, respectively.	Prather Lab

**Table S3.** Genetic parts of UxuR biosensor.

Name of genetic part	Sequence	Reference
BBa_J23101 (promoter)	tttacagctagctcagtcctaggtattatgctagc	32
<i>E. coli</i> WT <i>uxuR</i> (ORF)	atgaaatctgccacctctgcgcaaagacctaccaggaagtcggggcgatgatccgcatctgatcataaagacgc cgtacaatctggcgaacggctgccccggagcgtgaaattgcagaaatgcttgatgtcacgacggacgggtgtagc tgaagcgtgatcatgctggagatcaaaggctggggaagtagccggggtgccggtatctatgttcttgataact caggcagccagaacacagacagtcaggatgccaacgtctgcaacgatgccggtcctttgagctgttacaggcgcg gcagttattggagagcaacatcgccagtttccgcttgcaggctaccgcaagatatcgtaaaatgcgtcagg cattgcaactggaagagcgtgaaactggctccagtcgcccggcagcagcgaagcggtagatgcagttccatct cgctattgccgaagcaacgataacagcatgctggtagctgtccgctcagtcctggcagtgccgggaaaacaat ccaatgtggctccagttgcacagccatctggatgacagcctgtatcgaagagtggtggcgatcacaacaga tcctcggcgttaatacaaaaagatgccgagcggcgaagctggcaatgtggcagcatctgaaaacgttaagca acgtctgctggaattctgaacgttgacgatattttttagtgctatctgttgattcatggccgctggataaagtcg acgcctga	-

<b>BBa_B1002</b> (terminator)	cgcaaaaaaccccgcttcggcggggtttttcgc	32
<b>Hybrid promoter, with UxuR operator sites in bold</b>	tttac <b>aaattggtataccaattt</b> tataatatattcaggga <b>aaattggtataccaattt</b> acaataattttgtaaact tt	34
<b>sfGFP (ORF)</b>	atgcgtaaaggcgaagagctgttcaactgggtcgtccctattctggggaactggatggatgtcaacggtcataa gtttccgtgctggcggagggtgaaggtgacgcaactaatggtaactgacgctgaagttcatctgtactactggta aactccggttccttggccgactctggaacgacgctgacttatgggtttcagtgctttgctcgttatccggaccat gaagcagcatgacttctcaagtccgcatgcccgaaggctatgtgcaggaacgcacgatttcttaaggtgacg gcacgtacaaaacgctgcggaagtgaatttgaaggcgataccctggtaaacgcattgagctgaaggcattg actttaagaggacggcaatatcctggccataagctggaatacaatttaacagccacaatgtttacatcaccgcc gataacaaaaaatggcattaagcgaattttaaattcgccacaacgtggaggatggcagcgtgacgctggct gatcactaccagcaaaactccaatcggatggtctgttctgctgcagacaatcactatctgagcagcaaaag cgttctgtctaaagatccgaacgagaaacgcatcatatggttctgctggagttcgaaccgcagcgggcatcacgc atggtatggatgaactgtacaaatga	-
<b>BBa_B0015</b> (terminator)	ccaggcatcaataaaaacgaaaggctcagtcgaaagactgggcctttctgtttatctgtttgtcgtggaacgctc tctactagagtcacactggctcaccttcgggtggcctttctgctttata	32

**Table S4.** List of primers used in this study.

ID	Description	Forward sequence(_F)	Reverse sequence(_R)
<b>BB</b>	-	TGACCCTTGAGACCATGAAAAGTAAAAATCCTA ACTCGAGCGCAAAAAAC	CAGAGGTGGCAGATTTTCATATTGTACTA CCTTAGATTAGTCCGAGC
<b>INS</b>	-	CTAATCTAAGGTAGTACAATATGAAATCTGCCA CCTCTGC	CGAGTTAGGATTTTCACITTCATGGTCTC AAGGGTCAGGCGTCGACTTTATCCAG
<b>OP1</b>	Promoter core	TACCAATTTTATAATATATTCAGGGAAAATTGG TATACC	TACCAATTTTCGTAAGTTATCCAGCAACC
<b>OP2</b>	Downstream	TACCAATTTACAAATAATTTTGTTAACTTTTCAG C	TACCAATTTTCCTGAATATATTATAATG TTAACGT
<b>P1</b>	-	CCGCCCTTTTACAGCTAGCT	CCGCACGGTTATCCAC
<b>P2</b>	-	CAGTCCTAGGTATTATGCTAGCCGT	AGCTAGCTGTAAGGGGC
<b>Mu t1</b>	R102A	GTTACAGGCGgcGcAGTTATTGGAGAG	AGCTCAAAGGACCGGCA
<b>Mut 2</b>	E106A	CAGTTATTGGcGAGCAACATCGCC	CCGCGCTGTAACAGCTC
<b>Mut 3</b>	Q116A	TGCCGCTTTGgcGGCTACCCGC	AACTCGGCGATGTTGCTC
<b>Mut 4</b>	D121A	ACCCGCGAAGcaATCGTCAAATGC	AGCTGCAAAGCGGCAAA
<b>Mut 5</b>	E132A	TTGCAACTGGcAGAGCGTGAA	TGCCTGACGATTTTGAC
<b>Mut 6</b>	E145A	GGCAGCAGCGcAAGCGGTGAC	CGGCGCACTGGAAGCCAG
<b>Mut 7</b>	M149A	AAGCGGTGACgcGcAGTTCCATCTCGCTATTGC CGAAGC	TCGCTGCTGCCCGGCGCA
<b>Mut 8</b>	N161A	AGCAACGCATgcCAGCATGCTG	TCGGCAATAGCGAGATGG

<b>Mut 9</b>	R169A	GGAGCTGTTcgcgCAGTCCTGGC	ACCAGCATGCTGTTATG
<b>Mut 10</b>	W172A	CCGTCAGTCCgcgCAGTGCGG	AACAGTCCACCAGCATG
<b>Mut 11</b>	W174A	GTCCTGGCAGgcgCGGGAAAACA	TGACGGAACAGCTCC
<b>Mut 12</b>	R175A	CTGGCAGTGGgcGAAAAACAATCCAATG	GACTGACGGAACAGCTCC
<b>Mut 13</b>	W181A	CAATCCAATGgcgATCCAGTTGCAC	TTTTCCCGCCACTGC
<b>Mut 14</b>	I182L	TCCAATGTGGcTCCAGTTGCAC	TTGTTTTCCCGCCACTGC
<b>Mut 15</b>	H185A	ATCCAGTTGCCcAGCCATCTG	CCACATTGGATTGTTTTCC
<b>Mut 16</b>	R194V	CAGCCTGTATgtCAAAGAGTGGTTG	TCATCCAGATGGCTGTGC
<b>Mut 17</b>	W197A	TCGCAAAGAGgcGTTGGGCGATC	TACAGGCTGTATCCAGATG
<b>Mut 18</b>	D200A	TGGTTGGGCGcaCACAAACAGATC	CTCTTTGCGATACAGGCTG
<b>Mut 19</b>	N266A	GCATCTGGAAgcCTTAAGCAACGTCTG	TGCCACATTGCCAGCTTC
<b>Mut 20</b>	R230A	CGTTAAGCAAgccCTGCTGGAATTCTCG	TTTTCCAGATGCTGCCAC
<b>Seq</b>	Sequencing	CTGCGTTATCCCTGATTCTG	CGTAAAGTTATCCAGCAACC

**Table S5.** List of plasmids used in this study.

Plasmid ID	Sequence	Reference
<b>UxuR biosensor</b>	Sensor plasmid contains fructuronic acid sensitive transcriptional regulator, <i>uxuR</i> , to regulate <i>gfp</i> expressed from the hybrid promoter. Carb <sup>R</sup>	This study
<b>UxuR biosensor mutant</b>	UxuR biosensor with single mutation listed in Table S2. Carb <sup>R</sup>	This study

**Table S6.** Asp148's atomic interactions in UxuR model TF.

#Acceptor	DonorH	Fraction	#Acceptor	DonorH	Fraction
<b>UxuR<sup>Apo</sup> – chain A</b>			<b>UxuR<sup>Apo</sup> – chain B</b>		
-	-	-	-	-	-
<b>UxuR<sup>FrctA</sup> – chain A</b>			<b>UxuR<sup>FrctA</sup> – chain B</b>		
AP1_148@OD1	ARG_175@HH22	0.1379	AP1_148'@OD1	ARG_175'@HH22	0.1697
AP1_148@OD1	ARG_175@HH12	0.1096	AP1_148'@OD1	ARG_175'@HH12	0.1201
<b>UxuR<sup>GlcA</sup> – chain A</b>			<b>UxuR<sup>GlcA</sup> – chain B</b>		
AP1_148@OD1	ARG_175@HH22	0.5288	AP1_148'@OD1	ARG_175'@HH22	0.4755
AP1_148@OD1	ARG_175@HH12	0.4706	AP1_148'@OD1	ARG_175'@HH12	0.4159

**Table S7.** His152's atomic interactions in UxuR model TF.

#Acceptor	DonorH	Fraction	#Acceptor	DonorH	Fraction
<b>UxuR<sup>Apo</sup> – chain A</b>			<b>UxuR<sup>Apo</sup> – chain B</b>		
GLU_106@OE1	HD1_152@HD1	0.5673	GLU_106'@OE1	HD1_152'@HD1	0.512
GLU_106@OE2	HD1_152@HD1	0.4048	GLU_106'@OE2	HD1_152'@HD1	0.4615
<b>UxuR<sup>FrctA</sup> – chain A</b>			<b>UxuR<sup>FrctA</sup> – chain B</b>		
GLU_106@OE1	HD1_152@HD1	0.5655	GLU_106'@OE1	HD1_152'@HD1	0.468
GLU_106@OE2	HD1_152@HD1	0.3826	GLU_106'@OE2	HD1_152'@HD1	0.4677
<b>UxuR<sup>GlcA</sup> – chain A</b>			<b>UxuR<sup>GlcA</sup> – chain B</b>		
GLU_106@OE2	HD1_152@HD1	0.4028	GLU_106'@OE2	HD1_152'@HD1	0.4682
GLU_106@OE1	HD1_152@HD1	0.3975	GLU_106'@OE1	HD1_152'@HD1	0.3334

**Table S8.** His201's atomic interactions in UxuR model TF.

#Acceptor	DonorH	Fraction	#Acceptor	DonorH	Fraction
<b>UxuR<sup>Apo</sup> – chain A</b>			<b>UxuR<sup>Apo</sup> – chain B</b>		
GLU_132@OE2	HD2_201@HD1	0.2092	GLU_132'@OE2	HD2_201'@HD1	0.2692
GLU_132@OE1	HD2_201@HD1	0.198	GLU_132'@OE1	HD2_201'@HD1	0.246
<b>UxuR<sup>FrctA</sup> – chain A</b>			<b>UxuR<sup>FrctA</sup> – chain B</b>		
GLU_132@OE2	HD2_201@HD1	0.18	-	-	-
GLU_132@OE1	HD2_201@HD1	0.1622	-	-	-
<b>UxuR<sup>GlcNA</sup> – chain A</b>			<b>UxuR<sup>GlcNA</sup> – chain B</b>		
-	-	-	-	-	-

**Table S9.** His223's atomic interactions in UxuR model TF.

#Acceptor	DonorH	Fraction	#Acceptor	DonorH	Fraction
<b>UxuR<sup>Apo</sup> – chain A</b>			<b>UxuR<sup>Apo</sup> – chain B</b>		
ASP_200@OD1	HD3_223@HD1	0.2417	ASP_200'@OD2	HD3_223'@HD1	0.2742
ASP_200@OD2	HD3_223@HD1	0.18	ASP_200'@OD1	HD3_223'@HD1	0.2592
<b>UxuR<sup>FrctA</sup> – chain A</b>			<b>UxuR<sup>FrctA</sup> – chain B</b>		
ASP_200@OD2	HD3_223@HD1	0.2751	ASP_200'@OD2	HD3_223'@HD1	0.3282
ASP_200@OD1	HD3_223@HD1	0.258	ASP_200'@OD1	HD3_223'@HD1	0.2584
<b>UxuR<sup>GlcNA</sup> – chain A</b>			<b>UxuR<sup>GlcNA</sup> – chain B</b>		
ASP_200@OD1	HD3_223@HD1	0.2135	ASP_200'@OD2	HD3_223'@HD1	0.2149
ASP_200@OD2	HD3_223@HD1	0.1576	ASP_200'@OD1	HD3_223'@HD1	0.2131

**Table S10.** Arg175's atomic interactions in UxuR model TF.

#Acceptor	DonorH	Fraction	#Acceptor	DonorH	Fraction
<b>UxuR<sup>Apo</sup> – chain A</b>			<b>UxuR<sup>Apo</sup> – chain B</b>		
GLU_145@OE1	ARG_175@HH22	0.1554	GLU_145'@OE2	ARG_175'@HH12	0.2885
GLU_145@OE1	ARG_175@HH12	0.1328	GLU_145'@OE1	ARG_175'@HH12	0.1791
GLU_145@OE1	ARG_175@HH11	0.1072	GLU_145'@OE2	ARG_175'@HH22	0.1704
GLU_145@OE2	ARG_175@HH11	0.0963	GLU_145'@OE1	ARG_175'@HH22	0.1248
<b>UxuR<sup>FrctA</sup> – chain A</b>			<b>UxuR<sup>FrctA</sup> – chain B</b>		
GLU_145@OE1	ARG_175@HH12	0.1763	GLU_145'@OE2	ARG_175'@HH21	0.2014
GLU_145@OE1	ARG_175@HH22	0.1753	GLU_145'@OE1	ARG_175'@HH21	0.192
GLU_145@OE2	ARG_175@HH12	0.1538	GLU_145'@OE1	ARG_175'@HH21	0.192
GLU_145@OE2	ARG_175@HH22	0.1503	AP1_148'@OD1	ARG_175'@HH22	0.1697
AP1_148@OD1	ARG_175@HH22	0.1379	GLU_145'@OE1	ARG_175'@HE	0.1543
GLU_145@OE1	ARG_175@HH21	0.1256	GLU_145'@OE2	ARG_175'@HE	0.133
AP1_148@OD1	ARG_175@HH12	0.1096	AP1_148'@OD1	ARG_175'@HH12	0.1201
GLU_145@OE2	ARG_175@HH21	0.0997	GLU_145'@OE1	ARG_175'@HH22	0.1113
			GLU_145'@OE2	ARG_175'@HH12	0.1086
<b>UxuR<sup>GlcNA</sup> – chain A</b>			<b>UxuR<sup>GlcNA</sup> – chain B</b>		
AP1_148@OD1	ARG_175@HH22	0.5288	AP1_148'@OD1	ARG_175'@HH22	0.4755
AP1_148@OD1	ARG_175@HH12	0.4706	AP1_148'@OD1	ARG_175'@HH12	0.4159
GLU_145@OE2	ARG_175@HH21	0.4273	GLU_145'@OE2	ARG_175'@HH21	0.3362
GLU_145@OE1	ARG_175@HH21	0.3603	GLU_145'@OE1	ARG_175'@HH21	0.3154
GLU_145@OE1	ARG_175@HE	0.2648	GLU_145'@OE1	ARG_175'@HE	0.1868
GLU_145@OE2	ARG_175@HE	0.1741	GLU_145'@OE2	ARG_175'@HE	0.1498

**Table S11.** Glu132's atomic interactions in UxuR model TF.

#Acceptor	DonorH	Fraction	#Acceptor	DonorH	Fraction
<b>UxuR<sup>Apo</sup> – chain A</b>			<b>UxuR<sup>Apo</sup> – chain B</b>		
GLU_132@OE1	ARG_194@HH11	0.2672	GLU_132'@OE2	HD2_201'@HD1	0.2692
GLU_132@OE2	ARG_194@HH11	0.2368	GLU_132'@OE1	HD2_201'@HD1	0.246
GLU_132@OE2	HD2_201@HD1	0.2092	GLU_132'@OE2	ARG_194'@HH11	0.1803
GLU_132@OE1	HD2_201@HD1	0.198	GLU_132'@OE1	ARG_194'@HH11	0.1674

UxuR <sup>FrctA</sup> – chain A			UxuR <sup>FrctA</sup> – chain B		
GLU_132@OE2	HD2_201@HD1	0.18	GLU_132'@OE2	ARG_194'@HH21	0.2535
GLU_132@OE1	ARG_194@HH21	0.1728	GLU_132'@OE2	ARG_194'@HE	0.2215
GLU_132@OE1	HD2_201@HD1	0.1622	GLU_132'@OE1	ARG_194'@HH21	0.221
GLU_132@OE2	ARG_194@HE	0.1491	GLU_132'@OE1	SER_144'@HG	0.2178
GLU_132@OE1	ARG_194@HH11	0.1332	GLU_132'@OE1	ARG_194'@HE	0.2169
GLU_132@OE2	ARG_194@HH21	0.1202	GLU_132'@OE2	SER_144'@HG	0.2085
			GLU_132'@OE2	ARG_194'@HH22	0.1271
			GLU_132'@OE1	ARG_194'@HH12	0.1161
			GLU_132'@OE2	ARG_194'@HH12	0.1154
			GLU_132'@OE1	ARG_194'@HH22	0.1064
UxuR <sup>GlnA</sup> – chain A			UxuR <sup>GlnA</sup> – chain B		
GLU_132@OE1	SER_144@HG	0.2593	GLU_132'@OE2	ML1_259'@H8	0.1717
GLU_132@OE2	SER_144@HG	0.2464	GLU_132'@OE2	SER_144'@HG	0.1362
GLU_132@OE1	ML1_259@H8	0.1382	GLU_132'@OE1	ARG_194'@HH21	0.1332
GLU_132@OE2	ML1_259@H8	0.1272	GLU_132'@OE1	SER_144'@HG	0.1331
GLU_132@OE1	ARG_194@HH11	0.1246	GLU_132'@OE2	ARG_194'@HH21	0.1326
GLU_132@OE2	ARG_194@HH11	0.1164	GLU_132'@OE1	ML1_259'@H8	0.1295
			GLU_132'@OE2	ARG_194'@HE	0.1063

**Table S12.** Glu106's atomic interactions in UxuR model TF.

#Acceptor	DonorH	Fraction	#Acceptor	DonorH	Fraction
UxuR <sup>Apo</sup> – chain A			UxuR <sup>Apo</sup> – chain B		
GLU_106@OE1	HD1_152@HD1	0.5673	GLU_106'@OE1	HD1_152'@HD1	0.512
GLU_106@OE2	ARG_102@HE	0.5533	GLU_106'@OE2	ARG_102'@HE	0.4962
GLU_106@OE1	ARG_102@HH21	0.526	GLU_106'@OE1	ARG_102'@HH21	0.4806
GLU_106@OE2	HD1_152@HD1	0.4048	GLU_106'@OE2	HD1_152'@HD1	0.4615
GLU_106@OE1	ARG_102@HE	0.3937	GLU_106'@OE1	ARG_102'@HE	0.4456
GLU_106@OE2	ARG_102@HH21	0.387	GLU_106'@OE2	ARG_102'@HH21	0.4307
UxuR <sup>FrctA</sup> – chain A			UxuR <sup>FrctA</sup> – chain B		
GLU_106@OE1	HD1_152@HD1	0.5655	GLU_106'@OE2	ARG_102'@HH21	0.4758
GLU_106@OE2	ARG_102@HE	0.5564	GLU_106'@OE1	ARG_102'@HE	0.4682
GLU_106@OE1	ARG_102@HH21	0.5562	GLU_106'@OE2	HD1_152'@HD1	0.468
GLU_106@OE2	HD1_152@HD1	0.3826	GLU_106'@OE1	HD1_152'@HD1	0.4677
GLU_106@OE2	ARG_102@HH21	0.3798	GLU_106'@OE2	ARG_102'@HE	0.4638
GLU_106@OE1	ARG_102@HE	0.3768	GLU_106'@OE1	ARG_102'@HH21	0.4636
UxuR <sup>GlnA</sup> – chain A			UxuR <sup>GlnA</sup> – chain B		
GLU_106@OE1	ARG_102@HE	0.4525	GLU_106'@OE1	ARG_102'@HE	0.5052
GLU_106@OE2	ARG_102@HE	0.448	GLU_106'@OE2	ARG_102'@HH21	0.4994
GLU_106@OE2	HD1_152@HD1	0.4028	GLU_106'@OE2	HD1_152'@HD1	0.4682
GLU_106@OE2	ARG_102@HH21	0.3975	GLU_106'@OE1	ARG_102'@HH21	0.4035
GLU_106@OE1	HD1_152@HD1	0.3975	GLU_106'@OE2	ARG_102'@HE	0.3865
GLU_106@OE1	ARG_102@HH21	0.3882	GLU_106'@OE1	HD1_152'@HD1	0.3334
			GLU_106'@OE1	SER_171'@HG	0.1247

**Table S13.** Asp200's atomic interactions in UxuR model TF.

#Acceptor	DonorH	Fraction	#Acceptor	DonorH	Fraction
UxuR <sup>Apo</sup> – chain A			UxuR <sup>Apo</sup> – chain B		
ASP_200@OD1	HD3_223@HD1	0.2417	ASP_200'@OD2	HD3_223'@HD1	0.2742
ASP_200@OD2	HD3_223@HD1	0.18	ASP_200'@OD1	HD3_223'@HD1	0.2592
ASP_200@OD1	ASN_226@HD22	0.1206	ASP_200'@OD2	ASN_226'@HD22	0.1727
			ASP_200'@OD1	ASN_226'@HD22	0.1428
UxuR <sup>FrctA</sup> – chain A			UxuR <sup>FrctA</sup> – chain B		
ASP_200@OD2	HD3_223@HD1	0.2751	ASP_200'@OD2	HD3_223'@HD1	0.3282
ASP_200@OD1	HD3_223@HD1	0.258	ASP_200'@OD1	HD3_223'@HD1	0.2584
ASP_200@OD2	ASN_226@HD22	0.1195	ASP_200'@OD2	ASN_226'@HD22	0.219
			ASP_200'@OD1	ASN_226'@HD22	0.1837

UxuR <sup>GlcNA</sup> – chain A			UxuR <sup>GlcNA</sup> – chain B		
ASP_200@OD1	ASN_226@HD22	0.308	ASP_200'@OD1	HD3_223'@HD1	0.2149
ASP_200@OD2	ASN_226@HD22	0.2499	ASP_200'@OD2	HD3_223'@HD1	0.2131
ASP_200@OD1	HD3_223@HD1	0.2135	ASP_200'@OD1	ASN_226'@HD22	0.1438
ASP_200@OD2	HD3_223@HD1	0.1576	ASP_200'@OD2	ASN_226'@HD22	0.1323

**Table S14.** Ser144's atomic interactions in UxuR model TF.

#Acceptor	DonorH	Fraction	#Acceptor	DonorH	Fraction
UxuR <sup>Apo</sup> – chain A			UxuR <sup>Apo</sup> – chain B		
SER_144@O	ARG_194@HH22	0.1178	GLU_135'@OE2	SER_144'@HG	0.2111
UxuR <sup>FrctA</sup> – chain A			UxuR <sup>FrctA</sup> – chain B		
ASP_189@OD2	SER_144@HG	0.1731	GLU_132'@OE1	SER_144'@HG	0.2178
ASP_189@OD1	SER_144@HG	0.1449	GLU_132'@OE2	SER_144'@HG	0.2085
UxuR <sup>GlcNA</sup> – chain A			UxuR <sup>GlcNA</sup> – chain B		
GLU_132@OE1	SER_144@HG	0.2593	GLU_132'@OE2	SER_144'@HG	0.1362
GLU_132@OE2	SER_144@HG	0.2464	GLU_132'@OE1	SER_144'@HG	0.1331
			SER_144'@O	ARG_194'@HH22	0.1398

**Table S15.** FrctA's atomic interactions with UxuR binding site.

#Acceptor	DonorH	Fraction	#Acceptor	DonorH	Fraction
UxuR <sup>FrctA</sup> – chain A			UxuR <sup>FrctA</sup> – chain B		
ML1_259@O6	ARG_102@HH12	0.7969	ML1_259'@O6	ARG_102'@HH22	0.8069
ML1_259@O6	ARG_102@HH22	0.7899	ML1_259'@O6	ARG_102'@HH12	0.7993
GLU_145@OE1	ML1_259@H6	0.3083	GLU_145'@OE2	ML1_259'@H6	0.2815
GLU_145@OE2	ML1_259@H6	0.3068	GLU_145'@OE1	ML1_259'@H6	0.2645
GLU_145@OE1	ML1_259@H8	0.1262	GLU_145'@OE2	ML1_259'@H8	0.1779
GLU_145@OE2	ML1_259@H8	0.1227	GLU_145'@OE1	ML1_259'@H8	0.17
			ML1_259'@O3	HIE_185'@HE2	0.13
			ML1_259'@O2	HIE_185'@HE2	0.1097

**Table S16.** GlcNA's atomic interactions with UxuR binding site.

#Acceptor	DonorH	Fraction	#Acceptor	DonorH	Fraction
UxuR <sup>GlcNA</sup> – chain A			UxuR <sup>GlcNA</sup> – chain B		
ML1_259@O7	ARG_102@HH12	0.8957	ML1_259'@O7	ARG_102'@HH12	0.6048
ML1_259@O7	ML1_259@H7	0.535	ML1_259'@O7	ML1_259'@H7	0.5679
ML1_259@O6	ARG_102@HH22	0.4447	ML1_259'@O6	ARG_102'@HH22	0.2801
GLU_145@OE2	ML1_259@H6	0.3197	GLU_145'@OE2	ML1_259'@H6	0.2949
GLU_145@OE1	ML1_259@H6	0.3162	GLU_145'@OE1	ML1_259'@H6	0.2393
ML1_259@O4	ARG_194@HH12	0.1864	GLU_132'@OE2	ML1_259'@H8	0.1717
GLU_132@OE1	ML1_259@H8	0.1382	GLU_132'@OE1	ML1_259'@H8	0.1295
GLU_132@OE2	ML1_259@H8	0.1272	ML1_259@O2	HIE_185@HE2	0.1022
ML1_259@O2	HIE_185@HE2	0.1214			

**Table S17.** Arg102's atomic interactions in UxuR model TF.

#Acceptor	DonorH	Fraction	#Acceptor	DonorH	Fraction
UxuR <sup>Apo</sup> – chain A			UxuR <sup>Apo</sup> – chain B		
GLU_106@OE2	ARG_102@HE	0.5533	GLU_106'@OE2	ARG_102'@HE	0.4962
GLU_106@OE1	ARG_102@HH21	0.526	GLU_106'@OE1	ARG_102'@HH21	0.4806
GLU_106@OE1	ARG_102@HE	0.3937	GLU_106'@OE1	ARG_102'@HE	0.4456
GLU_106@OE2	ARG_102@HH21	0.387	GLU_106'@OE2	ARG_102'@HH21	0.4307
UxuR <sup>FrctA</sup> – chain A			UxuR <sup>FrctA</sup> – chain B		

ML1_259@O6	ARG_102@HH12	0.7969	ML1_259'@O6	ARG_102'@HH22	0.8069
ML1_259@O6	ARG_102@HH22	0.7899	ML1_259'@O6	ARG_102'@HH12	0.7993
GLU_106@OE2	ARG_102@HE	0.5564	GLU_106'@OE2	ARG_102'@HH21	0.4758
GLU_106@OE1	ARG_102@HH21	0.5562	GLU_106'@OE1	ARG_102'@HH21	0.4682
GLU_106@OE2	ARG_102@HH21	0.3798	GLU_106'@OE2	ARG_102'@HE	0.4638
GLU_106@OE1	ARG_102@HE	0.3768	GLU_106'@OE1	ARG_102'@HE	0.4636
<b>UxuR<sup>GlcNA</sup> – chain A</b>			<b>UxuR<sup>GlcNA</sup> – chain B</b>		
ML1_259@O7	ARG_102@HH12	0.8957	ML1_259'@O7	ARG_102'@HH12	0.6048
GLU_106@OE1	ARG_102@HE	0.4525	GLU_106'@OE1	ARG_102'@HE	0.5052
GLU_106@OE2	ARG_102@HE	0.448	GLU_106'@OE2	ARG_102'@HH21	0.4994
ML1_259@O6	ARG_102@HH22	0.4447	GLU_106'@OE1	ARG_102'@HH21	0.4035
GLU_106@OE2	ARG_102@HH21	0.3975	GLU_106'@OE2	ARG_102'@HE	0.3865
GLU_106@OE1	ARG_102@HH21	0.3882	ML1_259'@O6	ARG_102'@HH22	0.2801

**Table S18.** Glu145's atomic interactions in UxuR model TF.

#Acceptor	DonorH	Fraction	#Acceptor	DonorH	Fraction
<b>UxuR<sup>Apo</sup> – chain A</b>			<b>UxuR<sup>Apo</sup> – chain B</b>		
GLU_145@OE2	ARG_194@HH22	0.3892	GLU_145'@OE1	ARG_194'@HH22	0.4388
GLU_145@OE2	ARG_194@HH12	0.3441	GLU_145'@OE2	ARG_194'@HH22	0.3021
GLU_145@OE1	ARG_194@HH22	0.2736	GLU_145'@OE2	ARG_175'@HH12	0.2885
GLU_145@OE1	ARG_194@HH12	0.2266	GLU_145'@OE1	ARG_194'@HH12	0.2457
GLU_145@OE1	ARG_175@HH22	0.1554	GLU_145'@OE2	ARG_194'@HH12	0.2067
GLU_145@OE1	ARG_175@HH12	0.1328	GLU_145'@OE1	ARG_175'@HH12	0.1791
GLU_145@OE1	ARG_175@HH11	0.1072	GLU_145'@OE2	ARG_175'@HH22	0.1704
			GLU_145'@OE1	ARG_175'@HH22	0.1248
			GLU_145'@OE1	HIE_185'@HE2	0.1248
<b>UxuR<sup>FrctA</sup> – chain A</b>			<b>UxuR<sup>FrctA</sup> – chain B</b>		
GLU_145@OE2	ML1_259@H6	0.3083	GLU_145'@OE2	ML1_259'@H6	0.2815
GLU_145@OE1	ML1_259@H6	0.3068	GLU_145'@OE1	ML1_259'@H6	0.2645
GLU_145@OE1	ARG_175@HH12	0.1763	GLU_145'@OE2	ARG_175'@HH21	0.2014
GLU_145@OE1	ARG_175@HH22	0.1753	GLU_145'@OE1	ARG_175'@HH21	0.192
<b>GLU_145@OE2</b>	<b>ARG_194@HH22</b>	<b>0.1657</b>	GLU_145'@OE1	ARG_175'@HH21	0.192
GLU_145@OE2	ARG_175@HH12	0.1538	GLU_145'@OE2	ML1_259'@H8	0.1779
GLU_145@OE2	ARG_175@HH22	0.1503	GLU_145'@OE1	ML1_259'@H8	0.17
<b>GLU_145@OE1</b>	<b>ARG_194@HH22</b>	<b>0.1489</b>	GLU_145'@OE1	ARG_175'@HE	0.1543
GLU_145@OE1	ML1_259@H8	0.1262	GLU_145'@OE2	ARG_175'@HE	0.133
GLU_145@OE1	ARG_175@HH21	0.1256	GLU_145'@OE1	ARG_175'@HH22	0.1113
GLU_145@OE2	ML1_259@H8	0.1227	GLU_145'@OE2	ARG_175'@HH12	0.1086
<b>UxuR<sup>GlcNA</sup> – chain A</b>			<b>UxuR<sup>GlcNA</sup> – chain B</b>		
GLU_145@OE2	ARG_175@HH21	0.4273	GLU_145'@OE2	ARG_175'@HH21	0.3362
GLU_145@OE1	ARG_175@HH21	0.3603	GLU_145'@OE1	ARG_175'@HH21	0.3154
GLU_145@OE2	ML1_259@H6	0.3197	GLU_145'@OE2	ML1_259'@H6	0.2949
GLU_145@OE1	ML1_259@H6	0.3162	GLU_145'@OE1	ML1_259'@H6	0.2393
GLU_145@OE1	ARG_175@HE	0.2648	GLU_145'@OE1	ARG_175'@HE	0.1868
GLU_145@OE2	ARG_175@HE	0.1741	GLU_145'@OE2	ARG_175'@HE	0.1498

**Table S19.** His185's atomic interactions in UxuR model TF.

#Acceptor	DonorH	Fraction	#Acceptor	DonorH	Fraction
<b>UxuR<sup>Apo</sup> – chain A</b>			<b>UxuR<sup>Apo</sup> – chain B</b>		
-	-	-	GLU_145'@OE1	HIE_185'@HE2	0.1248
<b>UxuR<sup>FrctA</sup> – chain A</b>			<b>UxuR<sup>FrctA</sup> – chain B</b>		
-	-	-	ML1_259'@O3	HIE_185'@HE2	0.13
			ML1_259'@O2	HIE_185'@HE2	0.1097
<b>UxuR<sup>GlcNA</sup> – chain A</b>			<b>UxuR<sup>GlcNA</sup> – chain B</b>		
ML1_259@O2	HIE_185@HE2	0.1214	ML1_259@O2	HIE_185@HE2	0.1022

**Table S20.** Arg194's atomic interactions in UxuR model TF.

#Acceptor	DonorH	Fraction	#Acceptor	DonorH	Fraction
<b>UxuR<sup>Apo</sup> – chain A</b>			<b>UxuR<sup>Apo</sup> – chain B</b>		
GLU_145@OE2	ARG_194@HH22	0.3892	GLU_145'@OE1	ARG_194'@HH22	0.4388
GLU_145@OE2	ARG_194@HH12	0.3441	GLU_145'@OE2	ARG_194'@HH22	0.3021
GLU_145@OE1	ARG_194@HH22	0.2736	ASP_189'@OD2	ARG_194'@HH21	0.2813
GLU_132@OE1	ARG_194@HH11	0.2672	ASP_189'@OD1	ARG_194'@HH21	0.2659
GLU_132@OE2	ARG_194@HH11	0.2368	GLU_145'@OE1	ARG_194'@HH12	0.2457
GLU_145@OE1	ARG_194@HH12	0.2266	GLU_145'@OE2	ARG_194'@HH12	0.2067
ASP_190@OD1	ARG_194@HH21	0.2165	ASP_189'@OD1	ARG_194'@HE	0.2027
ASP_189@OD1	ARG_194@HE	0.1957	ASP_189'@OD2	ARG_194'@HE	0.1847
ASP_189@OD1	ARG_194@HH21	0.1887	GLU_132'@OE2	ARG_194'@HH11	0.1803
SER_146@O	ARG_194@HH12	0.1525	GLU_132'@OE1	ARG_194'@HH11	0.1674
ASP_190@OD2	ARG_194@HH21	0.145			
ASP_189@OD2	ARG_194@HH21	0.141			
ASP_190@OD1	ARG_194@HE	0.1407			
ASP_189@OD2	ARG_194@HE	0.1302			
ASP_190@OD2	ARG_194@HE	0.1301			
SER_144@O	ARG_194@HH22	0.1178			
ASP_190@O	ARG_194@HH21	0.1026			
<b>UxuR<sup>FrctA</sup> – chain A</b>			<b>UxuR<sup>FrctA</sup> – chain B</b>		
ASP_189@OD1	ARG_194@HH21	0.2711	GLU_132'@OE2	ARG_194'@HH21	0.2535
ASP_189@OD2	ARG_194@HH21	0.2355	GLU_132'@OE2	ARG_194'@HE	0.2215
GLU_132@OE1	ARG_194@HH21	0.1728	GLU_132'@OE1	ARG_194'@HH21	0.221
GLU_145@OE2	ARG_194@HH22	0.1657	GLU_132'@OE1	ARG_194'@HE	0.2169
ASP_189@OD1	ARG_194@HE	0.1608	ASP_189'@OD2	ARG_194'@HE	0.1313
GLU_132@OE2	ARG_194@HE	0.1491	GLU_132'@OE2	ARG_194'@HH22	0.1271
GLU_145@OE1	ARG_194@HH22	0.1489	GLU_132'@OE1	ARG_194'@HH12	0.1161
GLU_132@OE1	ARG_194@HH11	0.1332	GLU_132'@OE2	ARG_194'@HH12	0.1154
ASP_189@OD2	ARG_194@HE	0.1303	ASP_189'@OD1	ARG_194'@HH21	0.1104
GLU_132@OE2	ARG_194@HH21	0.1202	ALA_140'@O	ARG_194'@HH12	0.1071
			GLU_132'@OE1	ARG_194'@HH22	0.1064
			SER_143'@OG	ARG_194'@HH22	0.1034
<b>UxuR<sup>GlcNA</sup> – chain A</b>			<b>UxuR<sup>GlcNA</sup> – chain B</b>		
ASP_189@OD2	ARG_194@HH21	0.2758	ASP_189'@OD2	ARG_194'@HH21	0.175
ASP_189@OD1	ARG_194@HH21	0.2183	ASP_189'@OD1	ARG_194'@HH21	0.1601
ML1_259@O4	ARG_194@HH12	0.1864	ASP_189'@OD2	ARG_194'@HE	0.1496
ASP_189@O	ARG_194@HE	0.1558	SER_144'@O	ARG_194'@HH22	0.1398
ASP_189@OD2	ARG_194@HE	0.1505	GLU_132'@OE1	ARG_194'@HH21	0.1332
ASP_189@OD1	ARG_194@HE	0.1422	GLU_132'@OE2	ARG_194'@HH21	0.1326
GLU_132@OE1	ARG_194@HH11	0.1246	ASP_189'@OD1	ARG_194'@HE	0.1318
GLU_132@OE2	ARG_194@HH11	0.1164	GLU_132'@OE2	ARG_194'@HE	0.1063

**Table S21.** Prediction of key residues, according to descending DNCF scores, contributing to the allosteric mechanism in the UxuR TF model.

<b>Key residues with allosteric role</b>	
Arg194, Trp181, Trp174, Trp172, Glu145, Arg175, Trp250, Trp197, Tyr245, His185, Arg19, Phe234, Arg9, Asp189, Met180, Leu231, Pro85, Gln12, Leu184, Met149, Phe247, Glu13, Asp148, Glu132, Met163, Phe96, Glu166, Leu188, Ser144, Gln173, Gln170, Arg169, Asn178, Lys60 and Glu176.	
<b>Key residues in direct contact with D-fructuronic acid (FrctA)</b>	
chain A	chain B
Arg194, Trp172, Glu145, Arg175, Trp197, His185 and Asp148.	Trp181', Arg194', Arg175', Glu145', His185', Trp197', Met149', Trp172', Leu188' and Ser144'.
Underlined residues were tested <i>in vitro</i> in this study. Residues in <b>bold</b> were pinpointed in both chains.	

## References

1. Almeida, B. C., Kaczmarek, J. A., Figueiredo, P. R., Prather, K. L. J. & Carvalho, A. T. P. Transcription factor allosteric regulation through substrate coordination to zinc. *NAR Genomics and Bioinformatics* **3**, lqab033 (2021).
2. Horne, C. R. *et al.* Mechanism of NanR gene repression and allosteric induction of bacterial sialic acid metabolism. *Nat Commun* **12**, 1988 (2021).
3. Singh, S. *et al.* Molecular mechanisms underlying allosteric behavior of Escherichia coli DgoR, a GntR/FadR family transcriptional regulator. *Nucleic Acids Research* **53**, gkae1299 (2025).
4. Morris, G. M. *et al.* AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J Comput Chem* **30**, 2785–2791 (2009).
5. Morris, G. M. *et al.* Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry* **19**, 1639–1662 (1998).
6. Yu, Z., Li, P. & Merz, K. M. Extended Zinc AMBER Force Field (EZAFF). *J. Chem. Theory Comput.* **14**, 242–254 (2018).
7. Li, P. & Merz, K. M. MCPB.py: A Python Based Metal Center Parameter Builder. *J. Chem. Inf. Model.* **56**, 599–604 (2016).
8. Bayly, C. I., Cieplak, P., Cornell, W. & Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **97**, 10269–10280 (1993).
9. Cieplak, P., Cornell, W. D., Bayly, C. & Kollman, P. A. Application of the multimolecule and multiconformational RESP methodology to biopolymers: Charge derivation for DNA, RNA, and proteins. *Journal of Computational Chemistry* **16**, 1357–1377 (1995).
10. Case, D. A. *et al.* The Amber biomolecular simulation programs. *Journal of Computational Chemistry* **26**, 1668–1688 (2005).
11. Hornak, V. *et al.* Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics* **65**, 712–725 (2006).
12. Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).
13. Kirschner, K. N. *et al.* GLYCAM06: A Generalizable Biomolecular Force Field. Carbohydrates. *J Comput Chem* **29**, 622–655 (2008).
14. Anandakrishnan, R., Aguilar, B. & Onufriev, A. V. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Research* **40**, W537–W541 (2012).

15. Pastor, R. W. Techniques and Applications of Langevin Dynamics Simulations. in *The Molecular Dynamics of Liquid Crystals* (eds. Luckhurst, G. R. & Veracini, C. A.) 85–138 (Springer Netherlands, Dordrecht, 1994). doi:10.1007/978-94-011-1168-3\_5.
16. Hauptman, H. A. [1] Shake-and-bake: An algorithm for automatic solution ab initio of crystal structures. in *Methods in Enzymology* vol. 277 3–13 (Academic Press, 1997).
17. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. (1998) doi:10.1063/1.464397.
18. Miao, Y., Feher, V. A. & McCammon, J. A. Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation. *J. Chem. Theory Comput.* **11**, 3584–3595 (2015).
19. Roe, D. R. & Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **9**, 3084–3095 (2013).
20. Miao, Y. *et al.* Improved Reweighting of Accelerated Molecular Dynamics Simulations for Free Energy Calculation. *J Chem Theory Comput* **10**, 2677–2689 (2014).
21. Grant, B. J., Rodrigues, A. P. C., ElSawy, K. M., McCammon, J. A. & Caves, L. S. D. Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics* **22**, 2695–2696 (2006).
22. R: The R Project for Statistical Computing. <https://www.r-project.org/>.
23. Halford, S. E. & Marko, J. F. How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res* **32**, 3040–3052 (2004).
24. Liao, Q. *et al.* Long Time-Scale Atomistic Simulations of the Structure and Dynamics of Transcription Factor-DNA Recognition. *J. Phys. Chem. B* **123**, 3576–3590 (2019).
25. Lükling, M., Elf, J. & Levy, Y. Conformational Change of Transcription Factors from Search to Specific Binding: A lac Repressor Case Study. *J. Phys. Chem. B* **126**, 9971–9984 (2022).
26. Schneider, M. & Antes, I. SenseNet, a tool for analysis of protein structure networks obtained from molecular dynamics simulations. *PLOS ONE* **17**, e0265194 (2022).
27. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498–2504 (2003).
28. Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering* **9**, 90–95 (2007).
29. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612 (2004).
30. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research* **42**, W320–W324 (2014).

31. Ni, C., Fox, K. J. & Prather, K. L. J. Substrate-activated expression of a biosynthetic pathway in *Escherichia coli*. *Biotechnology Journal* **17**, 2000433 (2022).
32. Registry of Standard Biological Parts. <https://parts.igem.org/Promoters/Catalog/Anderson>.
33. Davis, J. H., Rubin, A. J. & Sauer, R. T. Design, construction and characterization of a set of insulated bacterial promoters. *Nucleic Acids Res* **39**, 1131–1141 (2011).
34. Suvorova, I. A. *et al.* Comparative Genomic Analysis of the Hexuronate Metabolism Genes and Their Regulation in Gammaproteobacteria. *Journal of Bacteriology* **193**, 3956–3963 (2011).