Revealing the mechanism of repressor inactivation during switching of a temperate bacteriophage

Kim Krighaar Rasmussen, Andrés Palencia, Anders Varming, Habiba El-Wali, Elisabetta Boeri Erba, Martin Blackledge, Karin Hammer, Torsten Herrmann, Mogens Kilstrup, Leila Lo Leggio, et al.

Revealing the mechanism of repressor inactivation during switching of a temperate bacteriophage

Kim Krighaar Rasmussen[1], Andrés Palencia[2], Anders K. Varming[1], Habiba El-Wali[3], Elisabetta Boeri Erba[4], Martin Blackledge[4], Karin Hammer[3], Torsten Herrmann[4], Mogens Kilstrup[3], Leila Lo Leggio[1,*] and Malene Ringkjøbing Jensen[4,*]

[1]Department of Chemistry, University of Copenhagen, Universitetsparken 5, Copenhagen, Denmark
[2]Institute for Advanced Biosciences (IAB), Structural Biology of Novel Drug Targets in Human Diseases, INSERM U1209, CNRS UMR5309, Univ. Grenoble Alpes, Grenoble, France
[3]Metabolic signaling and regulation, Department of Systems Biology, Technical University of Denmark, DK-2800 Lyngby, Denmark
[4]Univ. Grenoble Alpes, CNRS, CEA, IBS, F-38000 Grenoble, France

* To whom correspondence should be addressed

Dr. Malene Ringkjøbing Jensen
E-mail: malene.ringkjobing-jensen@ibs.fr
Tel: +33 4 57 42 86 68

Prof. Leila Lo Leggio
E-mail: leila@chem.ku.dk
Tel: +45 35 32 02 95

Abstract

Temperate bacteriophages can enter one of two life cycles following infection of a sensitive host, the lysogenic or the lytic lifecycle. The choice between the two alternative life cycles is dependent upon a tight regulation of promoters and their cognate regulatory proteins within the phage genome. We investigate the genetic switch of TP901-1, a bacteriophage of *Lactococcus lactis*, controlled by the CI repressor and the MOR anti-repressor and their interactions with DNA. We determine the solution structure of MOR, and we solve the crystal structure of MOR in complex with the N-terminal domain of CI revealing the structural basis of MOR inhibition of CI binding to the DNA operator sites. ^{15}N NMR CPMG relaxation dispersion and rotating frame $R_{1\rho}$ measurements demonstrate that MOR displays molecular recognition dynamics on two different time scales involving a repacking of aromatic residues at the interface with CI. Mutations in the CI:MOR binding interface impair complex formation *in vitro*, and when introduced *in vivo,* the bacteriophage switch is unable to choose the lytic life cycle showing that the CI:MOR complex is essential for proper functioning of the genetic switch. On the basis of sequence alignments we show that the structural features of the MOR:CI complex are likely conserved among a larger family of bacteriophages from human pathogens implicated in transfer of antibiotic resistance.

Significance statement

Temperate bacteriophages can enter one of two life cycles following infection of a host, the lysogenic or the lytic lifecycle. In the lysogenic life cycle, the viral genome is silenced and integrated into the host genome, while in the alternative lytic life cycle, phage replication in the cell leads to production of new phages, cell lysis and release of phage progeny. Using a combination of NMR spectroscopy, X-ray crystallography, biophysics and *in vivo* experiments we have revealed the molecular mechanisms underlying the switch from the lysogenic to the lytic life cycle in the bacteriophage TP901-1 from *Lactococcus lactis*. The molecular mechanisms are likely conserved among a larger family of bacteriophages from human pathogens implicated in transfer of antibiotic resistance.

Temperate bacteriophages are bistable and can enter one of two life cycles following infection of a sensitive host, the lysogenic or the lytic lifecycle. In the lysogenic life cycle (leading to the immune state), the viral genome is silenced and integrated into the host genome to be replicated along with the host DNA during cell division. In the alternative lytic life cycle (the anti-immune state), phage replication in the cell leads to production of new phages, cell lysis and release of phage progeny. The choice between the two alternative life cycles of temperate bacteriophages is dependent upon a tight regulation of promoters and their cognate regulatory proteins within the phage genome.

The temperate bacteriophage TP901-1 infects the gram-positive bacterium *Lactococcus lactis* and the choice between the lytic and lysogenic life cycle is controlled by a genetic switch containing two divergently oriented promoters (the lytic promoter P_L and the lysogenic promoter P_R) that are regulated by the phage-encoded CI protein (encoded by the *cI* gene) and the modulator of repression (MOR, encoded by the *mor* gene) (Fig. 1A). CI contains an N-terminal helix-turn-helix (HTH) domain (CI-NTD), a flexible linker and two C-terminal domains (CI-CTD_1 and CI-CTD_2, Fig. 1B) (1, 2). Full-length CI has been shown to be hexameric (a trimer of dimers) at high protein concentrations, while a shorter construct encompassing residues 1-122 of CI (CI-Δ58) has been shown to be dimeric mediated by dimerization of CI-CTD_1 through a pair of helical hooks (3). MOR is a single domain protein of 72 amino acids of previously unknown structure (Fig. 1B).

Using *in vitro* DNA footprinting and gel retardation experiments, three palindromic operator sites, O_R, O_L and O_D, for binding to CI have been identified on the genome, where O_R is situated between the P_L and P_R promoters (Fig. 1A) (4). In the absence of MOR, binding of CI dimers to all of these operator sites leads to tight repression of the P_L promoter and partial repression of the P_R promoter (5), presumably through formation of a CI hexameric structure looping the DNA out between the operator sites (Fig. 1C). Partial repression of the lysogenic

P_R promoter is believed to result from thermal breathing of the complex by frequent opening and closing of the CI:O_R contact, while a strong repression of the lytic P_L promoter is ensured by interaction of CI at the O_L site thereby efficiently preventing RNA polymerase binding to the lytic promoter. The presence of the O_D operator site is suggested to increase cooperative binding of CI to O_R and O_L and to inhibit transcriptional elongation from P_L (4). In support of the DNA binding capacity of CI, mutations in the HTH DNA recognition motif of CI-NTD (K40A/S41A and Q45A/N48A) lead to transcription from the lytic P_L promoter *in vivo* due to reduced DNA binding affinity (around 400-fold difference in K_D) (6).

While the lysogenic state can be formed without the presence of MOR, the lytic state requires participation by both MOR and CI. In the absence of CI, *e.g.* in a strain carrying a knockout mutation in the *cI* gene, both P_L and P_R are fully active showing that MOR on its own is not able to repress transcription from either of the two promoters (5). In the lytic state, repression of the lysogenic P_R promoter is observed even if O_R, O_L or O_D are mutated, showing that P_R repression is not absolutely dependent upon the CI operator sites (5). Together these experiments argue against a mechanism where MOR alone acts as a repressor and suggest that a CI:MOR complex represses transcription from P_R (Fig. 1D).

Recently, additional evidence in favor of this hypothesis has been presented, first of all by demonstration of a direct interaction between CI and MOR (7). Secondly, a series of mutagenesis experiments verified the role of an evolutionary conserved composite O_{M1}-O_{M2}-O_R-O_{M3} operator site as a plausible binding site for the CI:MOR complex (Fig. 1A). In fact, mutations at all sites (O_{M1}, O_{M2}, O_R and O_{M3}, carried out one by one), decrease repression of P_R in the anti-immune state (7). The O_M sites are non-palindromic, however, O_{M2} and O_{M3} are present on opposite strands to form a palindromic O_{M2}-O_{RL}-O_{RR}-O_{M3} site. Thirdly, mutation of putative DNA binding residues on MOR, was shown to affect repression of P_R in the anti-immune state. Based on these experiments, an updated model of the lytic state was proposed

where CI:MOR complexes of different stoichiometries bind to the O_{M1}-O_{M2}-O_R-O_{M3} operator (Fig. 1D) (7). Importantly, statistical thermodynamical modeling of the TP901-1 genetic switch, supported by measurements of the activity of the lysogenic P_R promoter as function of the concentration of CI in the absence and presence of MOR, suggested that binding of a CI:MOR complex to DNA is not the sole action of MOR and that MOR in addition inhibits CI binding to DNA (8).

In this work, we reveal the structural and dynamic basis of the inhibition of DNA-binding to CI by MOR. We determine the solution structure of MOR using nuclear magnetic resonance (NMR) spectroscopy, and we solve the crystal structure of MOR in complex with CI-NTD revealing how MOR interferes with binding of CI to the DNA operator sites. The mutually exclusive nature of the MOR:CI and CI:O_L complexes is further supported by native mass spectrometry and NMR competition experiments. In addition, ^{15}N NMR relaxation measurements show that MOR displays molecular recognition dynamics on two different time scales involving a repacking of aromatic residues at the interface with CI-NTD. Sequence alignments show that the structural features of the MOR:CI-NTD complex are likely to be conserved among a larger family of bacteriophages from pathogenic species including *Staphylococcus, Enterococcus, Listeria,* and *Streptococcus*. Thus, our results, supported by *in vivo* experiments, demonstrate a novel mechanism of action of this family of temperate bacteriophages that is distinct from the well-studied λ phage, where CI and Cro directly compete for binding to the DNA operator sites (9–11).

Results

The NMR solution structure of MOR reveals a helix-turn-helix fold

The spectral assignment of the backbone resonances of MOR was obtained on the basis of a series of BEST-TROSY triple resonance experiments (*SI Appendix*, Fig. S1) (12). The

secondary chemical shifts identify four well-defined α-helices (H1-H4) and a shorter less well-defined helix (H5) surrounded by two proline residues (Pro59 and Pro63) (Fig. 2A). The solution structure of MOR was determined using the NMR analysis program UNIO (13) on the basis of distance restraints obtained from three-dimensional [^{1}H-^{1}H]-NOESY-^{15}N-HSQC and aliphatic and aromatic [^{1}H-^{1}H]-NOESY-^{13}C-HSQC experiments, combined with dihedral angle restraints obtained from the experimental backbone chemical shifts (*SI Appendix*, Table S1). The central core structure of MOR (residues S6-F66) displays five helices with the residues Q19-N38 adopting a helix-turn-helix (HTH) motif typical of DNA binding proteins (Fig. 2B-D) (10, 14, 15). The N- and C-terminal ends of MOR are dynamic and less well-defined (residues M1-T5 and F67-H72), although NOEs demonstrate a clear contact between the two tails (Fig. 2E).

Affinity, stoichiometry and interaction kinetics of the MOR:CI complex

To study the interaction between MOR and CI, we carried out chemical shift titrations of ^{15}N-labeled CI-Δ58 with MOR (Fig. 3A). The data show that the formation of the complex is slow on the NMR chemical shift time scale. Chemical shift perturbations (CSPs) are observed only within CI-NTD demonstrating that this domain alone is sufficient to mediate interaction (Fig. 3B). Furthermore, we carried out isothermal titration calorimetry (ITC) of the interaction of MOR with both CI-NTD and CI-Δ58 (Fig. 4A,B). The affinities measured for the two constructs are the same within experimental error (K_D = 672 nM for CI-NTD and K_D = 633 nM for CI-Δ58) with a stoichiometry of 1:1 for the MOR:CI-NTD complex and 2:1 for the MOR:CI-Δ58 complex (two MOR molecules binding to one CI-Δ58 dimer). Thus, CI-NTD is sufficient to achieve binding and no cooperativity is observed in the context of the longer, dimeric CI-Δ58 construct. The thermodynamic parameters show very similar pattern for binding of MOR to CI-NTD and CI-Δ58 (Fig. 4C,D). The dissection of the binding energy

into enthalpic and entropic contributions indicates that the nanomolar interaction is mainly driven by the entropic contribution and only few polar interactions as suggested by the low favorable binding enthalpy. In addition, chemical shift perturbations in MOR upon interaction with CI-NTD and CI-Δ58 are almost identical demonstrating that MOR does not establish intermolecular interactions (MOR-MOR) upon binding to the dimeric CI-Δ58 construct (*SI Appendix*, Fig. S2B,E).

The interaction kinetics of the MOR:CI-NTD complex was determined by ^{15}N chemical exchange saturation transfer (CEST) experiments. We employed a sample of ^{15}N-labeled CI-NTD with 18.5 % (molar ratio) of unlabeled MOR, and ^{15}N CEST experiments were recorded for two different saturating B_1 fields (20.8 and 8.4 Hz) (Fig. 4E). The data were analyzed simultaneously for four well-isolated resonances of CI-NTD according to a two-site exchange model (corresponding to the free and MOR-bound state of CI-NTD) yielding an exchange rate (k_{ex}) of 8.9 ± 1.7 s^{-1} and a population of 16.8 ± 2.9 %. The obtained population is slightly lower than what is expected from the dissociation constant obtained from ITC (18.45%). To investigate this effect in more detail, we repeated the analysis of the CEST data by fixing the population, p_B, to values between 10 and 25%. This analysis shows a shallow χ^2 minimum at populations between 16 and 18% with the corresponding k_{ex} values ranging from 9.4 to 8.3 s^{-1}. Thus, an almost equally good fit is obtained for populations approaching 18% in agreement with the ITC data. In addition, as expected for a slow exchange process, complete separation of p_B and k_{ex} is difficult, even using B_1 fields matching the k_{ex} value, however, the apparent k_{on} rate remains well-determined with a value of 1.5 ± 0.3 s^{-1}. To ensure that the ratio of the determined k_{off} and k_{on} rates is equal to the K_D value measured by ITC, we used the apparent k_{on} rate from the CEST experiments and the K_D from ITC to estimate the kinetic constants characteristic of the interaction yielding a slow dissociation rate and an association rate at or slightly beyond the diffusion limit (k_{on} = (9.9 ± 1.8)×10^6 $M^{-1}s^{-1}$ and k_{off} = 6.6 ± 1.2 s^{-1}).

MOR interferes with binding of CI-Δ58 to two adjacent DNA half-sites

To obtain insight into the structural basis of the interaction between MOR and CI-NTD, we determined the crystal structure of the complex. The following regions make up the interface between the two proteins: MOR (Y3-Y5, T42-V52, P59-A69) and CI-NTD (K2-T5, S52-Y59, S68-P78) (Fig. 5A,B), in agreement with previous predictions suggesting that the C-terminal end of CI-NTD is essential for interaction with MOR (16). The interface is stabilized by several polar contacts as detected by the PISA server (17). In particular, residue Q55 of CI-NTD appears to be crucial for the interaction as its side chain inserts into a pocket of MOR formed by Y3, Y5, T42, W43 and F66 where it adopts a conformation that is stabilized by multiple hydrogen bonds (Fig. 5D). In addition to these hydrogen bonds, the complex is stabilized by several hydrophobic contacts formed along the entire interface (Fig. 5C) in agreement with the observed thermodynamic profile of the interaction obtained from ITC experiments (Fig. 4C). We note that the interface observed in the crystal structure is in agreement with experimental chemical shift perturbations in MOR and CI-NTD obtained by assigning the NMR spectra of the complex forms of both MOR and CI-NTD (*SI Appendix*, Fig. S2A-D).

We have previously obtained the crystal structure of CI-NTD in complex with a half O_L operator site (1). On the basis of this and experimental small angle X-ray scattering (SAXS) data, we have obtained a model of CI-Δ58 in complex with the full operator site, where the two N-terminal domains bind to the two adjacent, major grooves on the DNA and a flexible linker connects these domains to the dimeric helical hook structure located at the C-terminus of CI-Δ58 (Fig. 5E) (3). Superimposing the structure of the MOR:CI-NTD complex onto one of the CI-NTD of the CI-Δ58:O_L complex leads to a likely steric clash between MOR and the other CI-NTD (Fig. 5E), suggesting that the formation of the CI-MOR complex interferes with simultaneous binding of CI to the two major grooves of the DNA operator site. To study

in more detail the competition between MOR and DNA for binding to CI, we performed native mass spectrometry (MS) experiments on mixtures of MOR, O_L and CI-Δ58 (*SI Appendix*, Fig. S3A-D). The MS data indicate the formation of two types of hetero-complexes containing CI-Δ58:MOR and CI-Δ58:O_L. No ternary complex of MOR, CI-Δ58 and O_L could be detected. These results strongly support the mutually exclusive nature of the binding of CI-Δ58 to MOR or O_L. This observation was also supported by NMR competition experiments where pre-formation of the CI-Δ58:MOR complex (and similarly the CI-NTD:MOR complex) showed dissociation of MOR upon addition of O_L (*SI Appendix*, Fig. S3E).

MOR shows molecular recognition dynamics involving repacking of aromatic residues

The dynamics of MOR on the pico- to nanosecond (ps-ns) time scale were studied using ^{15}N spin relaxation measurements (18, 19). A model-free analysis of the data using an isotropic diffusion tensor gives a rotational correlation time of 4.5 ns consistent with the molecular weight and a monomeric form of MOR. A complete analysis of the relaxation data shows that the use of a fully anisotropic diffusion tensor is statistically significant (*SI Appendix*, Fig. S4), and the derived order parameters, S^2, demonstrate that MOR is rigid on the ps-ns time scale, except for the first five N-terminal and the last four C-terminal residues (*SI Appendix*, Fig. S4D). In addition, the analysis reveals a number of conformational exchange contributions to the transverse relaxation that have not been efficiently quenched by the spin lock field of 1.5 kHz used in the $R_{1\rho}$ experiments (Fig. 6A,B). These exchange contributions are centered in three different regions of MOR that cluster around Y5 indicating that reorientational dynamics of its side chain are responsible for these exchange contributions (Fig. 6C).

In addition, we measured ^{15}N CPMG relaxation dispersion at two magnetic fields (600 and 950 MHz) to monitor exchange occurring on slower time scales. Several MOR residues display dispersion and by comparing the effective R_2 relaxation rate at high and low CPMG

frequencies with the measured $R_{1\rho}$ data (Fig. 6D), it becomes clear that the dispersion curves asymptotes to the measured $R_{1\rho}$ rates. This means that the relaxation dispersion data can be analyzed independently of the exchange contributions observed for the ring reorientation of Y5, because the Y5 exchange rate is much faster than the applied CPMG field. The residues showing the largest exchange contributions (Fig. 6E) cluster around the side chain of W43 (Fig. 6F), again suggesting that side chain ring reorientations are responsible for the extensive exchange contributions observed in MOR. This conclusion is also supported by the fact that in the ^{1}H-^{15}N HSQC spectrum of MOR, the ^{1}H-^{15}N side chain signal of W43 is not detectable suggesting extreme line broadening due to conformational exchange. We analyzed the CPMG curves simultaneously for all residues showing dispersion according to a two-site exchange model (Fig. 6G). An exchange rate of k_{EX} = 533 ± 16 s^{-1} and a population of the transiently populated (excited) state of W43 of 3.6 ± 0.1 % were obtained from the analysis.

By comparing the NMR structure of MOR to the crystal structure of MOR in complex with CI-NTD, we observe that the side chain of W43 is in two different conformations. The orientation of the W43 side chain in the NMR structure is well-defined by a number of NOEs from the side chain of W43 to L8, L37 and E47 (Fig. 2F, *SI Appendix*, Fig. S5). Some of these NOEs are clearly not compatible with the conformation of the W43 side chain observed in the crystal structure of the complex (*SI Appendix*, Fig. S5). Flipping the side chain of W43 in the crystal structure to the same conformation as observed in the NMR structure leads to van der Waals contacts that are too close between W43 and Y5 of MOR and Q55 of CI-NTD. This suggests that the formation of the complex with CI-NTD requires a repacking of the two aromatic rings of Y5 and W43 in order for Q55 of CI-NTD to insert into the pocket of MOR as depicted in Fig. 5D. The dynamics observed for Y5 and W43, as detected by $R_{1\rho}$ and CPMG relaxation dispersion, most likely represent molecular recognition dynamics (20–22) where the side chains of Y5 and W43 in the free form of MOR transiently flip to

conformations allowing the interaction with CI-NTD. This is in agreement with the measurements of the dynamics of MOR in complex with CI-NTD. The exchange contributions to $R_{1\rho}$ are absent in the complex, and the exchange contributions detected by CPMG relaxation dispersion are absent or significantly reduced in the complex showing that both the Y5 and W43 side chains adopt a rigid conformation in the complex (*SI Appendix*, Fig. S6).

Switch frequencies respond strongly to changes in the MOR:CI-NTD interaction surface and are sensitive measures of CI:MOR and MOR:DNA interactions

To validate our structural data we genetically exchanged specific amino acids in the CI-NTD:MOR interface to alanines or other types of amino acids and quantified the effect on the switch frequencies by recording the color phenotypes of *L. lactis* colonies, transformed by altered TP901-1 switch plasmids. TP901-1 switch plasmids contain the 1 kb DNA region shown in Fig. 1A, which encodes all the functions needed to perform a decision between the lytic and the lysogenic state, after entering an *L. lactis* cell as naked DNA. Because the anti-immune (lytic) P_L promoter is transcriptionally linked as a P_L-*mor*-*lacLM* fusion to the *lacLM* genes, the switch phenotype can be recorded by plating switch plasmid transformants on agar plates containing X-gal and subsequently counting blue (anti-immune phenotype) and white (immune phenotype) colonies. Wildtype TP901-1 switch plasmids yield around 2.6% white colonies in the present series (*SI Appendix*, Table S3), which is equal to the frequency of lysogens after infecting susceptible *L. lactis* with TP901-1 (5).

Initially, we aimed at perturbing the hydrogen bonding network of the MOR:CI-NTD complex mediated by the two residues MOR-Y5 and CI-Q55 (Fig. 5D). Substitution of these amino acids to alanines results in 100% immune phenotype (white colonies) consistent with

MOR being unable to interfere with binding of CI to O_L due to a disruption or weakening of the MOR:CI complex, or alternatively that the MOR-CI complex is unable to repress P_R.

Figure 5C draws attention to a patch of hydrophobic residues, which may be of importance for the overall stability of the MOR:CI-NTD complex. The CI-M73 residue is located in the vicinity of MOR-F67 and when either of these is substituted with an alanine residue (CI-M73A or MOR-F67A) the switching phenotype is 100% immune (white), showing that both amino acids are important for CI:MOR stability. In contrast, plasmids harboring CI-F75A and CI-F75V substitutions lead to switching frequencies of 0 and 1.1%, respectively (*SI Appendix*, Table S3) instead of the normal 2.6% white colonies for the wild type. This suggests that small hydrophobic side chains at the CI-F75 position do not significantly change the stability of the CI:MOR complex. Larger hydrophobic amino acids at the CI-F75 position apparently slightly destabilize the complex as judged from the switching frequencies of 17% and 31% for the CI-F75I and CI-F75Y substitutions, respectively (*SI Appendix*, Table S3). It thus appears that evolution has selected for a CI:MOR binding that is not too strong and still allows for switching. Finally, we also studied a plasmid harboring a MOR-R31A substitution. MOR-R31 is located in the putative DNA recognition helix of MOR with a potential role in the repression of the P_R promoter (*SI Appendix*, Fig. S7). The 100% immune phenotype (white colonies) detected therefore strongly support the prediction that MOR-R31 is important for the anti-immune repression of P_R.

To support the *in vivo* experiments, we tested the four single point mutations of CI-NTD (CI-Q55A, CI-F75A, CI-F75Y and CI-M73A) *in vitro* with respect to their structural integrity and their ability to form a complex with MOR. The CI-NTD variants carrying the above mutations were easily expressed and purified. These variants folded correctly as detected by circular dichroism and NMR (*SI Appendix*, Fig. S8A,B). Furthermore, the variants had similar stability as the wild-type as estimated by nano differential scanning fluorimetry (nDSF) with

T_i between 59.8°C and 62.3°C (wild-type T_i being 60.5°C) and all had similar affinity for the O_L site (K_D = 4-8 μM) as the corresponding wild-type construct. NMR analysis of complex formation with purified MOR showed that CI-M73A severely impairs complex formation with an estimated dissociation constant around 450 μM (*SI Appendix*, Fig. S8C), consistent with its high impact on the switching phenotype. The NMR data also reveal that the two mutants CI-F75A and CI-F75Y are not able to disrupt the complex with MOR (*SI Appendix*, Fig. S8C). This result is in broad agreement with the *in vivo* data showing no or low impact on the switching phenotype for the CI-F75A and CI-F75Y substitutions, respectively. Finally, CI-Q55A shows complex formation similar to the wild-type CI-NTD with almost identical chemical shifts of bound MOR (*SI Appendix*, Fig. S8C). This is unexpected since the 100% switching to the immune state indicates that a MOR:CI-Q55A complex is not able to compete with CI repression. It is possible that the CI-Q55A mutation weakens the complex (but the interaction remains in slow exchange on the NMR chemical shift time scale), so that it is not able to efficiently inhibit CI binding to its DNA operator sites. While the CI-M73A mutation has a very clear effect *in vitro* and is in perfect agreement with the *in vivo* experiments, the mutations at CI-F75 and CI-Q55 are more difficult to interpret probably due to compensatory effects at the binding interface that still allow complex formation.

Interaction of MOR and CI-NTD with the conserved O_{M1}-O_{M2}-O_R-O_{M3} operator site

The CI-MOR complex has been proposed to bind to the evolutionary conserved O_{M1}-O_{M2}-O_R-O_{M3} operator site to repress transcription from the lysogenic promoter, P_R (7). To verify this hypothesis and obtain insight into the importance of the O_M sites, we used ^{15}N-labeled MOR, CI-NTD and pre-formed CI-NTD:MOR complex and added DNA sequences corresponding to O_L (used as a control), O_R, O_{M1}-O_{M2} and O_{M2}-O_R-O_{M3}. The ^{1}H-^{15}N HSQC spectra show that CI-NTD interacts with O_L involving its HTH region in a fast to intermediate exchange regime

as previously shown (1), while its interaction with O_R gives rise to only small chemical shift changes suggesting a weaker interaction with this operator site (*SI Appendix*, Fig. S9), as expected from their sequence differences (1). Surprisingly, we find that CI-NTD interacts specifically with O_{M1}-O_{M2} in a fast to intermediate exchange regime with residues in the HTH motif undergoing the largest chemical shift changes. CI-NTD also interacts with O_{M2}-O_R-O_{M3}, however, additional line broadening of the NMR resonances is observed compared to the interaction with O_{M1}-O_{M2}, probably due to the increased size of the complex (*SI Appendix*, Fig. S9). In agreement with previous data showing interaction with DNA only in the presence of CI (7), MOR does not interact specifically with any of the tested DNA sequences as judged from negligible chemical shift changes (*SI Appendix*, Fig. S10), however, interestingly the interaction between the pre-formed MOR:CI-NTD complex with O_{M2}-O_R-O_{M3} shows that MOR remains in the complex upon interaction with the DNA. Only the flexible tail of CI-NTD is visible in the NMR spectrum of this complex (*SI Appendix*, Fig. S11), probably due to its high molecular weight as well as conformational exchange contributions to the transverse relaxation. Together, these data support the hypothesis that the CI-MOR complex binds to the proposed composite operator site region.

Conservation and evolution of CI:MOR interactions across bacterial genomes

Phylogenetic analysis of the TP901-1 switch region at the DNA level has previously offered structural insights into the CI protein showing that primarily the N-terminal domain of CI that is involved in DNA and MOR binding has been conserved in switches carrying *mor* genes and O_M binding sites (7). To gain more insight into the evolution of the CI:MOR interactions, the primary sequences of CI and MOR were used as templates in the web server ComplexContact (23) to search for similar sequences of the two proteins within the same genome. Because bacterial genomic sequences were used in the search, switch regions from

integrated prophages were detected. Protein sequences and alignments of concatenated CI-MOR amino acid sequences encoded on individual genomes from separate *cI* and *mor* genes were retrieved and annotated according to the bacterial species. Subsequent analysis of these alignments revealed important aspects of the evolution of the CI:MOR interacting pairs. The phylogenetic tree constructed from the CI:MOR alignment (*SI Appendix*, Fig. S12A) shows a branching structure with lower branches containing prophages from *Clostridium* and unidentified gram-positive genomes, from which a larger stem protrudes containing branches of prophage CI and MOR from *L. lactis* and pathogenic species including *Staphylococcus, Enterococcus, Listeria,* and *Streptococcus*. Many of these pathogens have been linked to life-threatening infections and the emergence of antibiotic resistant species is becoming a wide-spread problem. Interestingly, this important branching coincides with the insertion of a nine-amino acid fragment (^{27}SIPFQKKFG^{35}) into CI-NTD, forming the greater part of the scaffolding helix of the DNA-binding HTH domain in the TP901-1 CI protein (shown with filled triangles in *SI Appendix*, Fig. S12, where the left angle of the triangles indicates the point of sequence diversification among the branching members). This region was previously recognized as an extended scaffolding helix in the CI protein of TP901-1, lactococcal phage φ31, enterococcal phage φEf11, and staphylococcal phage TP310-1 (1). The extended scaffolding helix appears to be of minor importance for the interaction with MOR as observed in the crystal structure of the MOR:CI-NTD complex (Fig. 5A,B), although NMR CSPs are observed of residues 33-35 of CI-NTD and at the extreme C-terminus of MOR indicating a contact between these two regions in the complex (*SI Appendix*, Fig. S2).

Logo plots of the amino acid frequencies in CI-NTD and MOR, where the CI repressor carries the extended scaffolding helix, demonstrate that the stem of the phylogenetic tree containing branches of prophage CI and MOR from *L. lactis* and other pathogenic species shows remarkable sequence homology (*SI Appendix*, Fig. S12B,C). This implies a conservation of

the fold of both MOR, CI-NTD and the MOR:CI-NTD complex, which together with previously noted similarities at the DNA level, strongly suggests that the interplay between CI and MOR in the lytic cycle of these human pathogens can be expected to closely resemble TP901-1 from *L. lactis*.

Discussion

We have studied the mechanism of action of the genetic switch from the temperate bacteriophage TP901-1 using an integrated structural biology approach combining NMR spectroscopy, X-ray crystallography with biophysics and native mass spectrometry. In particular, we have characterized the structure and dynamics of MOR using NMR, and we have solved the crystal structure of MOR in complex with the N-terminal domain of CI. On the basis of NMR exchange experiments, we have shown that the formation of the MOR:CI-NTD complex is facilitated by multi-time scale molecular recognition dynamics in the free state of MOR involving transient ring reorientations of the aromatic side chains of Y5 and W43. By superimposing the MOR:CI-NTD crystal structure with a model of the CI-Δ58:O_L complex (3) a likely steric clash is observed between MOR and the other CI-NTD molecule in the dimer (Fig. 5E) revealing how MOR inhibits CI binding to the DNA operator sites and promotes the switch to the lytic life cycle. Using ITC we determined an affinity of 633 nM of the MOR:CI-Δ58 complex (Fig. 4A,B), which is significantly weaker than the estimated affinities of a CI dimer for the O_L site as determined by electrophoretic mobility shift assay (EMSA) (3 nM) (4) suggesting that MOR cannot easily fully displace CI from the DNA operator sites, though the competition may still be sufficient to favor full-length CI-MOR complex P_R repression *in vivo*. Furthermore, we show that while MOR is displaced from the pre-formed CI-NTD:MOR complex upon addition of O_L, it remains bound in the complex when presented with the O_{M2}-O_R-O_{M3} fragment, in agreement with models recently presented.

Unexpectedly, however, CI-NTD binds at least as well to O_{M1}-O_{M2} and O_{M2}-O_R-O_{M3} as it does to the O_R site, something not considered in the recent models.

To further address the significance of the identified complex for the switch mechanism, we carried out *in vitro* and *in vivo* mutagenesis studies. One of our mutants, CI-M73A disrupted the formation of the MOR:CI-NTD complex *in vitro*. This correlated very well with our *in vivo* data since introducing a corresponding mutation in the full-length CI gene *in vivo* leads to a frequency of lysogenization (immune transformants) at 100 % as compared to 2.6 % for the wild type switch showing that a disruption of the MOR:CI complex leads to inability to establish the lytic state. Three additional mutations lead to similar *in vivo* results, further validating that the CI:MOR binding interface is essential for proper functioning of the switch.

Finally, we also tested a mutation MOR-R31A intended to impair MOR-DNA interactions at the O_{M1}-O_{M2}-O_R-O_{M3} operator site (*SI Appendix*, Fig. S7). The *in vivo* data show a frequency of lysogenization at 100 % consistent with an inability of the MOR-CI complex to establish interactions with DNA and repress the P_R promoter. While it is difficult to resolve the apparent discrepancy in affinities with regard to P_L derepression or to reveal the full molecular details of the P_R repression complex at this point, it is important to note that by necessity of the biophysical techniques employed, and lower solubility and stability of full-length CI, much of the presented work is on truncated versions of both proteins and DNA regions.

Temperate bacteriophages infecting lactic acid bacteria have been placed in the λ supergroup of *Siphoviridae* (24). This classification is based upon the structural proteins and shape of the phage coat where other modules in the phages may be very different. It is therefore not surprising that we observe clear structural and functional differences between the genetic switches of TP901-1 and the well-studied *Escherichia coli* phages λ and 186. In terms of domain organization, the CI proteins from TP901-1, λ and 186 are similar containing an N-

terminal DNA binding domain but displaying a different quaternary structure with full-length CI TP901-1 being hexameric, CI λ octameric and CI 186 heptameric (25, 26). Furthermore, the C-terminal domains are of predominantly β-fold in λ and 186, but helical in TP901-1 (2, 3). In the λ phage, the repressor CI and the anti-repressor Cro compete for binding to multiple operator sites in the divergent promoter region leading to CI-mediated repression of the lytic promoter or Cro repression of the lysogenic promoter. The structure of λ Cro was solved previously using X-ray crystallography showing a dimeric structure composed of three α-helices and three β-strands (27). In contrast, MOR from TP901-1 is monomeric as determined here by NMR spectroscopy and its structure is composed of five α-helices. The structure of MOR (HTH domain) suggests that it is a DNA-binding protein, however, we were not able to show specific binding of MOR to either the O_L or O_R operators sites (*SI Appendix*, Fig. S10). This shows a different mechanism of MOR compared to Cro that instead directly interacts with the λ phage operator site as evidenced by a crystal structure of Cro in complex with an operator site (28).

In phage 186, the promoters are arranged face to face (as opposed to back to back in λ and TP901-1) and CI is inactivated by binding to the anti-repressor protein Tum. Thus, similarly to TP901-1, a protein-protein interaction is the key mechanism for CI inactivation in phage 186, however, the structural features of the CI:Tum complex appear to be completely different to those of the CI:MOR complex solved here. In particular, Tum is a 146 amino acid protein predicted to contain a DinI-like domain at its C-terminus (residues 84-146). The N-terminal domain of Tum is currently of unknown structure and shows poor sequence homology with previously solved structures.

In conclusion, we have unveiled a fundamental structural mechanism that controls the switch of the bacteriophage TP901-1 to the lytic life cycle by providing the atomic details and dynamics of the interaction between CI and MOR proteins. The structural features of the

complex are consistent with the proposed functions of MOR in de-repressing the P_L promoter and co-repressing the P_R promoter, and disruption of the identified CI:MOR interface even by single mutations in either partner was shown to abolish the lytic cycle, as shown by changes of switch frequencies *in vivo*. The findings have much more far-reaching implications than for TP901-1, as indicated by high conservation of the CI-NTD and MOR sequences in a larger family of bacteriophages from pathogenic species including *Staphylococcus, Enterococcus, Listeria,* and *Streptococcus* (*SI Appendix*, Fig. S12). Since the CI:MOR interface is largely conserved, our research provides a guide for understanding the role of the complex in lysogens of several human pathogens, which have been implicated in transfer of antibiotic resistance, and potentially opens new avenues for phage therapy employing temperate bacteriophages (29).

Materials and Methods

Protein expression and purification procedures, details of CD spectroscopy, nDSF, EMSA, ITC and native mass spectrometry and of construction and analysis of genetically altered switch plasmids are described in the *SI Appendix* Materials and Methods.

NMR structure determination of MOR

All NMR experiments were carried out at 25°C in 20 mM Tris buffer, 100 mM NaCl, 1 mM dithiothreitol (DTT) at pH 6.5. The backbone assignment was carried out using a set of BEST-TROSY triple-resonance experiments HNCO, HN(CA)CO, HN(CO)CACB and intra-residue iHNCACB (12). All NMR data were transformed using NMRPipe (30) and analyzed in CcpNmr (31) or SPARKY (32). Sequential connectivities were obtained using NEXUS under CcpNmr followed by manual verification. Secondary structure propensities were calculated from the Cα and Cβ chemical shifts of MOR according to the SSP protocol (33).

Side chain assignments were obtained from a HCCH-TOCSY (34, 35) experiment acquired at a ^{1}H frequency of 950 MHz using a mixing time of 50 ms. Distance restraints for structure determination were obtained from three-dimensional [^{1}H-^{1}H]-NOESY-^{15}N-HSQC (36–38) and aliphatic and aromatic [^{1}H-^{1}H]-NOESY-^{13}C-HSQC spectra. The three NOESY spectra were all acquired with a mixing time of 120 ms at a ^{1}H frequency of 700 MHz. A fresh sample of MOR was used for each of the three NOESY and the HCCH-TOCSY experiments to avoid problems of sample degradation during long acquisition times. Sample concentrations between 0.5 and 0.7 mM were used for the experiments.

A total of 930 meaningful, non-redundant upper NOE-derived distance restraints were automatically generated by the NMR analysis program UNIO (39) in combination with the structure calculation program CYANA (40) following the standard protocol (13). The final structural bundle was calculated in explicit solvent using the program CNS (41). The MOR NMR structure has excellent stereochemistry with 97.3% of the residues in most favored regions and additionally allowed regions, 2.3% in generously allowed regions, and 0.3% in disallowed regions of the Ramachandran plot. Furthermore, the structural ensemble has excellent geometry with no violations of distance restraints greater than 0.5 Å and no dihedral angle violations greater than 5°. With the exception of the N-terminal residues 1 to 4 and the C-terminal residues 67 to 72, the atomic coordinates of the MOR NMR structure is well-defined. The root-mean-square-deviation (RMSD) value about the mean coordinate positions of the backbone atoms for residues 5 to 66 is 0.60 ± 0.07 Å. Detailed NMR and refinement statistics are reported in *SI Appendix*, Table S1. The ^{1}H, ^{13}C and ^{15}N chemical shifts have been deposited in the BioMagResBank under the BMRB ID 34462. The atomic coordinates of the bundle of 20 NMR conformers have been deposited in the PDB database (accession code 6TO6).

NMR chemical shift titrations

The interaction of ^{15}N-labeled MOR with unlabeled CI-NTD was studied by recording ^{1}H-^{15}N HSQC spectra for the following concentrations of MOR and CI-NTD: 106:0 μM (free MOR), 106:226 μM (MOR saturated with CI-NTD) and 106:60 μM (MOR partly saturated with CI-NTD).

The assignments of the NMR spectra of MOR in complex with CI-NTD was carried out using BEST-TROSY triple resonance experiments using a sample of ^{13}C, ^{15}N-labeled MOR (0.7 mM) in complex with unlabeled CI-NTD (1.1 mM). Similarly, the assignments of the NMR spectra of ^{13}C, ^{15}N-labeled CI-NTD (0.4 mM) in complex with unlabeled MOR (0.6 mM) were obtained. Combined ^{1}H and ^{15}N CSPs in MOR or CI-NTD between the free and bound state were calculated as $((\Delta\delta_N/8.5)^2 + (\Delta\delta_H)^2)^{1/2}$. The assignments of the NMR spectra of CI-Δ58 were obtained previously (2).

Interactions of ^{15}N-labeled MOR with CI-NTD mutants were carried out by recording ^{1}H-^{15}N HSQC experiments of MOR in the presence of saturating amounts of CI-M73A, CI-F75A and CI-F75Y (2.3 times excess) and CI-Q55A (3 times excess).

Interactions of ^{15}N-labeled MOR, CI-NTD and pre-formed MOR:CI-NTD complex with DNA sequences corresponding to the O_L, O_R, O_{M1}-O_{M2} and O_{M2}-O_R-O_{M3} operator sites were carried out using concentrations of 126 μM and 128 μM of MOR and CI-NTD, respectively. Concentrations of the DNA sequences were in all cases similar with molar ratios compared to CI-NTD and MOR between 1:0.23 and 1:0.26.

NMR relaxation measurements of MOR and the MOR:CI-NTD complex

^{15}N relaxation rates, R_1, $R_{1\rho}$ and heteronuclear NOEs were measured of a 0.6 mM sample of ^{15}N-labeled MOR at 25°C and a ^{1}H frequency of 700 MHz using standard HSQC-detected pulse sequences (42). The R_1 relaxation rates were obtained by sampling the decay of

magnetization at ten time points, while the $R_{1\rho}$ rates were measured with a spin lock field of 1.5 kHz and with eleven different time points. For each experiment, one time point was repeated for error estimation. The measured $R_{1\rho}$ and R_1 rates were converted into transverse relaxation rates, R_2, by taking into account off-resonance effects (43). CPMG relaxation dispersion experiments (44) were carried out at two magnetic field strengths (600 and 950 MHz) using a constant time relaxation delay of 32 ms and with CPMG frequencies ranging from 31 to 1000 Hz. The data were analyzed simultaneously at the two fields according to a two-site exchange model using the program ChemEx (45).

Similar experiments (R_1, $R_{1\rho}$ and CPMG relaxation dispersion) were carried out of MOR in complex with CI-NTD at 25°C and a ^{1}H frequency of 700 MHz. A sample of ^{13}C, ^{15}N-labeled MOR (0.7 mM) in complex with unlabeled CI-NTD (1.1 mM) was used for these experiments.

CEST experiments

The CEST experiments were carried out at a ^{1}H frequency of 600 MHz on a sample of ^{15}N-labeled CI-NTD (720 μM) with unlabeled MOR (133 μM) corresponding to a molar ratio of 18.5%. The experiments were carried out with an ^{15}N B_1 field strength of 20.8 Hz applied during a constant period of 300 ms. Saturation was carried out in the range from 100 to 129 ppm in steps of 25 Hz. The dataset was complemented by a second experiment using a ^{15}N B_1 field strength of 8.4 Hz during a constant period of 300 ms. Saturation was carried out in the range from 110.5 to 120.5 ppm in steps of 5 Hz and using a reduced ^{15}N sweep width of 6 ppm with the carrier positioned at 116.7 ppm. The CEST data of the resonances D54, N56, S68 and M73, for which large CSPs were observed, were fitted simultaneously for the two different ^{15}N B_1 fields using the analysis program ChemEx according to a two-site exchange model (45).

Crystal structure of the MOR:CI-NTD complex

MOR:CI-NTD was co-crystallized using a sample of MOR at 2mg/mL in a 1:2 molar ratio with CI-NTD (residues 1-89 of CI). We screened for crystallization conditions using the JSCG+ screen (Hampton research). After two days hexagonal crystal plates appeared in several conditions with similar composition 25 % PEG 3350, 0.2 mM salt (e.g. ammonium sulfate and lithium sulfate). A data set for a crystal grown with a reservoir consisting of 0.2 M lithium sulfate, 0.1 M Bis-Tris pH 5.5, 25% PEG 3350 was collected at beam line ID23-1, ESRF, Grenoble, France using a PILATUS detector. XDS/XSCALE (46) was used to process the images into a final almost complete data set extending to 2.28 Å resolution, with space group *P*321 and with cell dimensions of a = 94.41 Å, b = 94.41 Å and c = 30.56 Å.

Molecular Replacement (MR) was carried out with Phaser.PHENIX(47) and experimental data were phased using the crystal structure of CI-NTD (PDB: 5A7L), placing it as fixed model, and using the NMR model with lowest energy of MOR truncated for flexible tails (N– and C-terminal) for a further MR search. A model of MOR created using the webserver PHYRE2 (48) also allowed phasing. Phaser-MR found a solution with rotational function Z-score (RFZ) of 2.8, a top translation function Z-score (TFZ) of 9.1, and a log-likelihood (LLG) of 313.722. Initial refinement successfully improved phases and Rwork/Rfree went from 43.26%/42.25% to 29.76%/37.70%. Remaining parts of the proteins not included in the search models were manually built or not observable, presumably due to intrinsic flexibility (residues 81-89 of CI-NTD and his-tag). The final structure was refined to Rwork/Rfree of 18.9%/21.7% with good geometry using Phenix.refine. Data and model statistics are summarized in *SI Appendix*, Table S2 and the structure has been deposited in the PDB (accession code 6TRI).

Acknowledgments

This work was supported by the Danish Council for Independent Research (Grant 4002-00107) to L.L.L, M.R.J. and K.H. We acknowledge the platforms of the Grenoble Instruct-ERIC center (ISBG; UMS 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB). Platform access was supported by FRISBI (ANR-10-INBS-05-02) and GRAL, a project of the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). The IBS acknowledges integration into the interdisciplinary Research Institute of Grenoble (IRIG, CEA). This work has been supported by iNEXT, grant number 653706, funded by the Horizon 2020 programme of the European Commission. Travel to synchrotrons was supported by the DANSCATT program, funded by the Danish Ministry for Higher Education and Science (Grant 7055-00001B) and the European Community's Seventh Framework Programme (FP7/2007–2013) (to L.L.L.). Financial support is acknowledged from the French Agence Nationale de la Recherche through ANR T-ERC MAPKassembly (to M.R.J.). We thank Caroline Mas for assistance and access to the biophysical platform and Luca Signor for technical assistance with analyzing intact proteins in denaturing conditions and useful discussions. K.K.R., A.V. and L.L.L. are members of ISBUC, Integrative Structural Biology at the University of Copenhagen (www.isbuc.ku.dk).

Author contributions

K.K.R., L.L.L., K.H., M.K. and M.R.J. designed research, K.K.R expressed and purified proteins for NMR, X-ray crystallography, native mass spectrometry and ITC experiments, A.K.V. expressed, purified and characterized mutants of CI, K.K.R., M.R.J and A.P carried out and analyzed the ITC experiments, E.B.E. carried out and analyzed native mass spectrometry experiments, K.K.R., M.B., T.H and M.R.J. acquired and analyzed all NMR

experiments, K.K.R. and L.L.L. determined the crystal structure of the MOR:CI-NTD complex, H.E.-W. and M.K. carried out the *in vivo* experiments, L.L.L and M.R.J supervised and coordinated the project. K.K.R., M.K., L.L.L. and M.R.J. wrote the manuscript with input from all authors.

References

1. K. H. Frandsen, *et al.*, Binding of the N-Terminal Domain of the Lactococcal Bacteriophage TP901-1 CI Repressor to Its Target DNA: A Crystallography, Small Angle Scattering, and Nuclear Magnetic Resonance Study. *Biochemistry* **52**, 6892–6904 (2013).

2. K. K. Rasmussen, *et al.*, Structural and dynamics studies of a truncated variant of CI repressor from bacteriophage TP901-1. *Sci Rep* **6**, 29574 (2016).

3. K. K. Rasmussen, *et al.*, Structural basis of the bacteriophage TP901-1 CI repressor dimerization and interaction with DNA. *FEBS Lett.* **592**, 1738–1750 (2018).

4. A. H. Johansen, L. Brøndsted, K. Hammer, Identification of operator sites of the CI repressor of phage TP901-1: evolutionary link to other phages. *Virology* **311**, 144–156 (2003).

5. M. Pedersen, K. Hammer, The role of MOR and the CI operator sites on the genetic switch of the temperate bacteriophage TP901-1. *J. Mol. Biol.* **384**, 577–589 (2008).

6. M. Pedersen, L. Lo Leggio, J. G. Grossmann, S. Larsen, K. Hammer, Identification of quaternary structure and functional domains of the CI repressor from bacteriophage TP901-1. *J. Mol. Biol.* **376**, 983–996 (2008).

7. M. Pedersen, *et al.*, Repression of the lysogenic PR promoter in bacteriophage TP901-1 through binding of a CI-MOR complex to a composite OM-OR operator. *Sci Rep* **10**, 8659 (2020).

8. A. Alsing, M. Pedersen, K. Sneppen, K. Hammer, Key players in the genetic switch of bacteriophage TP901-1. *Biophys. J.* **100**, 313–321 (2011).

9. H. Eisen, P. Brachet, L. Pereira da Silva, F. Jacob, Regulation of repressor expression in lambda. *Proc. Natl. Acad. Sci. U.S.A.* **66**, 855–862 (1970).

10. R. G. Brennan, B. W. Matthews, The helix-turn-helix DNA binding motif. *J. Biol. Chem.* **264**, 1903–1906 (1989).

11. A. B. Oppenheim, O. Kobiler, J. Stavans, D. L. Court, S. Adhya, Switches in bacteriophage lambda development. *Annu. Rev. Genet.* **39**, 409–429 (2005).

12. Z. Solyom, *et al.*, BEST-TROSY experiments for time-efficient sequential resonance assignment of large disordered proteins. *J. Biomol. NMR* **55**, 311–321 (2013).

13. P. Guerry, V. D. Duong, T. Herrmann, CASD-NMR 2: robust and accurate unsupervised analysis of raw NOESY spectra and protein structure determination with UNIO. *J. Biomol. NMR* **62**, 473–480 (2015).

14. I. B. Dodd, J. B. Egan, Improved detection of helix-turn-helix DNA-binding motifs in protein sequences. *Nucleic Acids Res.* **18**, 5019–5026 (1990).

15. L. Aravind, V. Anantharaman, S. Balaji, M. M. Babu, L. M. Iyer, The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol. Rev.* **29**, 231–262 (2005).

16. M. Pedersen, M. Ligowska, K. Hammer, Characterization of the CI repressor protein encoded by the temperate lactococcal phage TP901-1. *J. Bacteriol.* **192**, 2102–2110 (2010).

17. E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).

18. G. Lipari, A. Szabo, Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J. Am. Chem. Soc.* **104**, 4559–4570 (1982).

19. A. M. Mandel, M. Akke, A. G. Palmer, Backbone dynamics of Escherichia coli ribonuclease HI: correlations with structure and function in an active enzyme. *J. Mol. Biol.* **246**, 144–163 (1995).

20. Q. Zhang, A. C. Stelzer, C. K. Fisher, H. M. Al-Hashimi, Visualizing spatially correlated dynamics that directs RNA conformational transitions. *Nature* **450**, 1263–1267 (2007).

21. O. F. Lange, *et al.*, Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science* **320**, 1471–1475 (2008).

22. E. Delaforge, *et al.*, Large-Scale Conformational Dynamics Control H5N1 Influenza Polymerase PB2 Binding to Importin α. *J. Am. Chem. Soc.* **137**, 15122–15134 (2015).

23. S. Wang, S. Sun, Z. Li, R. Zhang, J. Xu, Accurate De Novo Prediction of Protein Contact Map by Ultra-Deep Learning Model. *PLOS Computational Biology* **13**, e1005324 (2017).

24. H. Brüssow, F. Desiere, Comparative phage genomics and the evolution of Siphoviridae: insights from dairy phages. *Mol. Microbiol.* **39**, 213–222 (2001).

25. I. B. Dodd, *et al.*, Cooperativity in long-range gene regulation by the lambda CI repressor. *Genes Dev.* **18**, 344–354 (2004).

26. P. J. Neufing, K. E. Shearwin, J. Camerotto, J. B. Egan, The CII protein of bacteriophage 186 establishes lysogeny by activating a promoter upstream of the lysogenic promoter. *Mol. Microbiol.* **21**, 751–761 (1996).

27. D. H. Ohlendorf, D. E. Tronrud, B. W. Matthews, Refined structure of Cro repressor protein from bacteriophage lambda suggests both flexibility and plasticity. *J. Mol. Biol.* **280**, 129–136 (1998).

28. R. A. Albright, B. W. Matthews, Crystal structure of lambda-Cro bound to a consensus operator at 3.0 A resolution. *J. Mol. Biol.* **280**, 137–151 (1998).

29. R. Monteiro, D. P. Pires, A. R. Costa, J. Azeredo, Phage Therapy: Going Temperate? *Trends in Microbiology* **27**, 368–378 (2019).

30. F. Delaglio, *et al.*, NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).

31. W. F. Vranken, *et al.*, The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696 (2005).

32. T. D. Goddard, D. G. Kneller, SPARKY 3. *University of California, San Francisco*.

33. J. A. Marsh, V. K. Singh, Z. Jia, J. D. Forman-Kay, Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: implications for fibrillation. *Protein Sci.* **15**, 2795–2804 (2006).

34. A. Bax, G. M. Clore, A. M. Gronenborn, 1H-1H correlation via isotropic mixing of 13C magnetization, a new three-dimensional approach for assigning 1H and 13C spectra of 13C-enriched proteins. *Journal of Magnetic Resonance (1969)* **88**, 425–431 (1990).

35. L. E. Kay, G. Y. Xu, A. U. Singer, D. R. Muhandiram, J. D. Formankay, A Gradient-Enhanced HCCH-TOCSY Experiment for Recording Side-Chain 1H and 13C Correlations in H2O Samples of Proteins. *Journal of Magnetic Resonance, Series B* **101**, 333–337 (1993).

36. D. Marion, *et al.*, Overcoming the overlap problem in the assignment of proton NMR spectra of larger proteins by use of three-dimensional heteronuclear proton-nitrogen-15 Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1.beta. *Biochemistry* **28**, 6150–6156 (1989).

37. D. Marion, L. E. Kay, S. W. Sparks, D. A. Torchia, A. Bax, Three-dimensional heteronuclear NMR of nitrogen-15 labeled proteins. *J. Am. Chem. Soc.* **111**, 1515–1517 (1989).

38. E. R. P. Zuiderweg, S. W. Fesik, Heteronuclear three-dimensional NMR spectroscopy of the inflammatory protein C5a. *Biochemistry* **28**, 2387–2391 (1989).

39. P. Guerry, T. Herrmann, Comprehensive automation for NMR structure determination of proteins. *Methods Mol. Biol.* **831**, 429–451 (2012).

40. P. Güntert, C. Mumenthaler, K. Wüthrich, Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* **273**, 283–298 (1997).

41. A. T. Brunger, Version 1.2 of the Crystallography and NMR system. *Nat Protoc* **2**, 2728–2733 (2007).

42. N.-A. Lakomek, J. Ying, A. Bax, Measurement of ^{15}N relaxation rates in perdeuterated proteins by TROSY-based methods. *J. Biomol. NMR* **53**, 209–221 (2012).

43. M. Akke, A. G. Palmer, Monitoring Macromolecular Motions on Microsecond to Millisecond Time Scales by R1ρ−R1 Constant Relaxation Time NMR Spectroscopy. *J. Am. Chem. Soc.* **118**, 911–912 (1996).

44. D. F. Hansen, P. Vallurupalli, L. E. Kay, An improved 15N relaxation dispersion experiment for the measurement of millisecond time-scale dynamics in proteins. *J Phys Chem B* **112**, 5898–5904 (2008).

45. P. Vallurupalli, G. Bouvignies, L. E. Kay, Studying "invisible" excited protein states in slow exchange with a major state conformation. *J. Am. Chem. Soc.* **134**, 8148–8161 (2012).

46. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).

47. A. J. McCoy, *et al.*, Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674 (2007).

48. R. M. Bennett-Lovsey, A. D. Herbert, M. J. E. Sternberg, L. A. Kelley, Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. *Proteins* **70**, 611–625 (2008).

Figures legends

Fig. 1. The genome of TP901-1 functions as a genetic switch. (*A*) Schematic representation of the fragment of the TP901-1 genome that functions as a genetic switch. Gray boxes indicate the positions of the genes *cI* and *mor*. The positions and directions of the two promoters P_L and P_R are shown as arrows. The operator sites, O_R, O_L and O_D, which bind to CI are shown as boxes (wheat color) along with their DNA sequences. Orange boxes indicate the position of the evolutionary conserved O_{M1}, O_{M2} and O_{M3} operator sites that are putative binding sites for the CI-MOR complex (see panel D). (*B*) Domain organization of the two proteins CI (top) and MOR (bottom). The N- and C-terminal domains of CI are connected by a flexible linker (gray), and crystal structures of CI-NTD and CI-CTD_1 are shown, as determined previously (2, 3). (*C*) In the immune state, CI is believed to form a hexameric structure around which the DNA loops by binding to the three operator sites. (*D*) In the anti-immune state, MOR (red) interacts with CI and inhibits binding of CI to the DNA allowing transcription from the lytic promoter, P_L. In addition, CI:MOR complexes of different stoichiometries interact with the O_{M1}-O_{M2}-O_R-O_{M3} operator site through an unknown binding mode (represented by dashed lines) to repress transcription from the lysogenic promoter, P_R.

Fig. 2. NMR solution structure of MOR. (*A*) Secondary structure propensity (SSP) calculated from experimental Cα and Cβ chemical shifts. Positive values correspond to helical propensity, while negative values represent β-strand propensity. (*B*) Topology diagram of the solution structure of MOR where H2-loop-H3 constitute the HTH motif, H2 being the scaffolding (green) and H3 the recognition helix (blue). (*C*) Structure of MOR shown as a bundle of the 20 structures with lowest energy. (*D*) Structure of MOR shown as a single structure of lowest energy after water refinement. (*E*) Plane of the aromatic [^{1}H-^{1}H]-NOESY-^{13}C-HSQC spectrum at the ^{13}C resonance frequency of Y3Cδ (133.25 ppm) showing NOEs

from the Y3Hδ protons. Clear NOEs are observed from Y3 to A68 and A69 located in the C-terminal tail of MOR. (*F*) Planes of the aromatic [^{1}H-^{1}H]-NOESY-^{13}C-HSQC spectrum at the ^{13}C resonance frequencies of W43Cδ (120.90 ppm) and W43Cζ (113.57 ppm) showing NOEs from the W43Hδ1 and W43Hζ2 protons. Clear NOEs are observed from W43Hδ1 to E47 and from W43Hζ2 to the side chains of L8 and L37. These NOEs define the orientation of W43 in the NMR structure.

Fig. 3. Mapping the interaction site of MOR on CI. (*A*) Superposition of a region of the ^{1}H-^{15}N HSQC spectrum of ^{15}N CI-Δ58 (red) and ^{15}N CI-Δ58 in complex with MOR (blue). Assignments of the HSQC spectrum of CI-Δ58 were obtained previously (red labels) (2). All resonances corresponding to CI-CTD$_1$ (residues 90-122) remain unperturbed by the interaction with MOR. (*B*) Combined ^{1}H and ^{15}N CSPs in ^{15}N CI-Δ58 upon interaction with MOR.

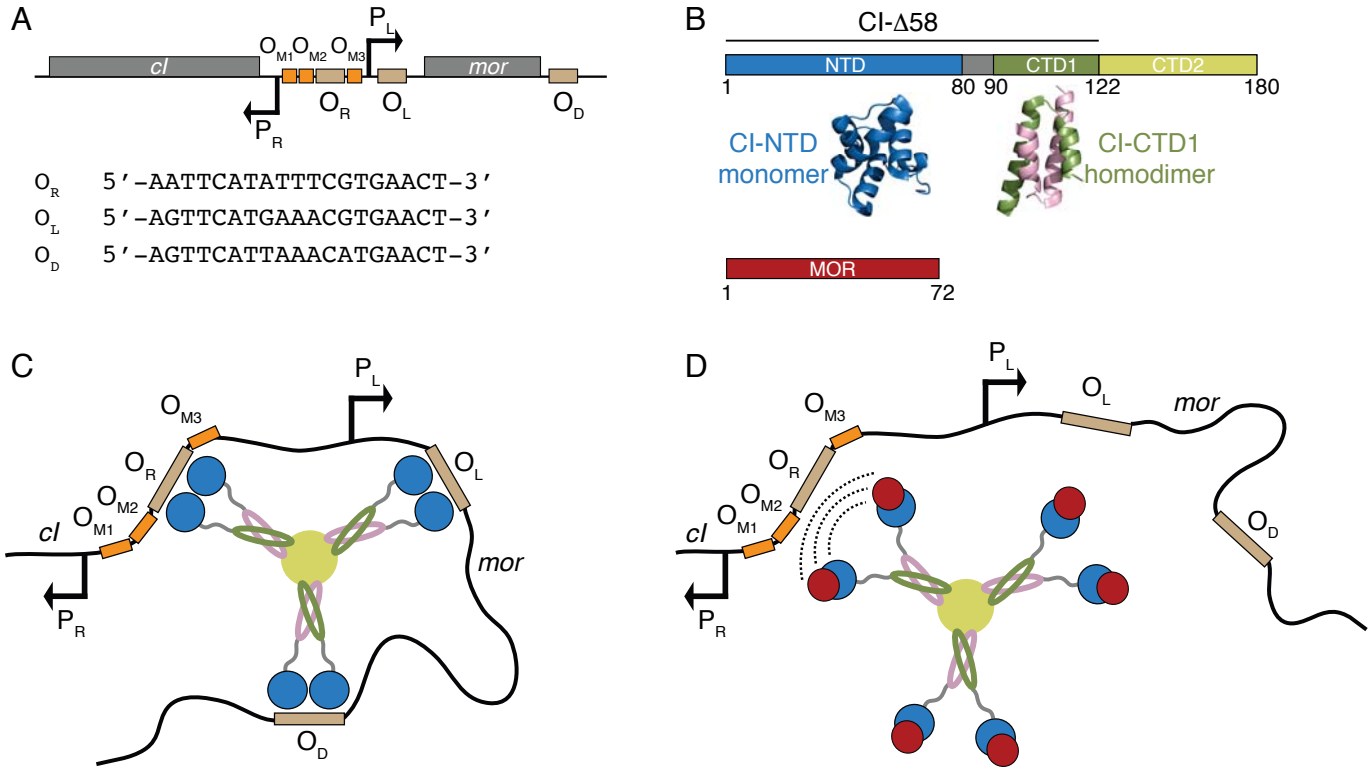
Fig. 4. Binding affinity and interaction kinetics of the MOR:CI-NTD complex. (*A*) ITC data of the MOR:CI-NTD complex. Representative data are shown with raw injections heats (top) and the corresponding specific binding isotherms (bottom). (*B*) ITC data of the MOR:CI-Δ58 complex. The dimer concentration of CI-Δ58 is reported. The analysis of the ITC data was carried out according to a model of *n* identical and independent binding sites. The results of the ITC data are reported as the mean of two independent measurements along with the standard error of the mean. (*C*) Dissection of binding free energies into the enthalpic and entropic contributions for the interaction of MOR with CI-NTD. (*D*) Dissection of binding free energies into the enthalpic and entropic contributions for the interaction of MOR with CI-Δ58. In both panel *C* and *D*, the values are reported per independent binding site. (*E*) CEST data used to determine the interaction kinetics of the MOR:CI-NTD complex. Data were

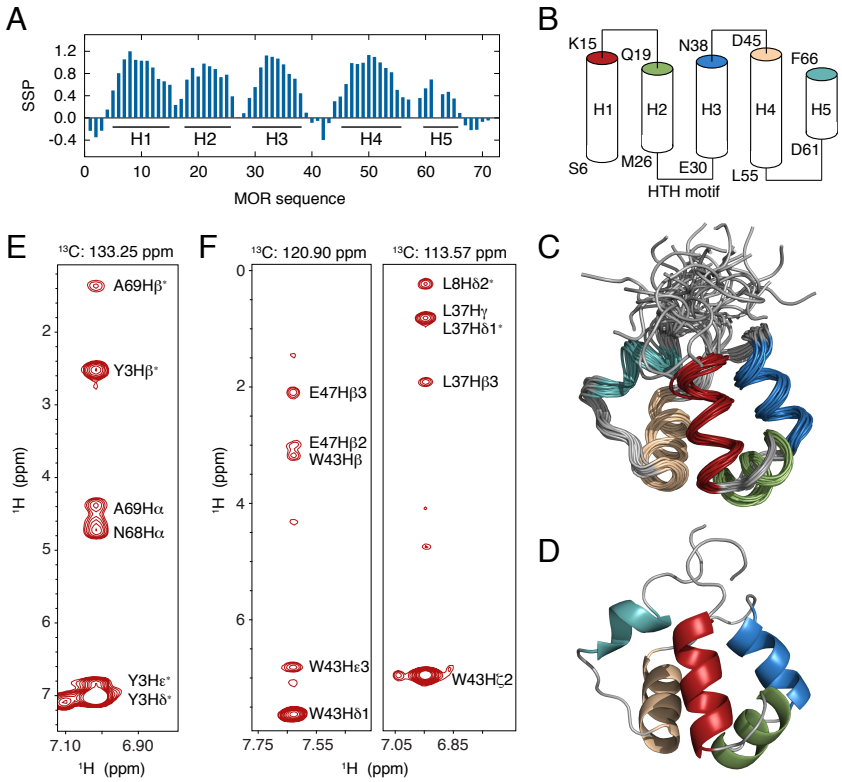
acquired for a sample of ^{15}N-labeled CI-NTD with 18.5% (molar ratio) of MOR for two different ^{15}N B_1 field strengths (red: 20.8 Hz and blue: 8.4 Hz). Experimental data are shown as points, while lines correspond to a simultaneous analysis of all of the data shown according to a two-site exchange model.

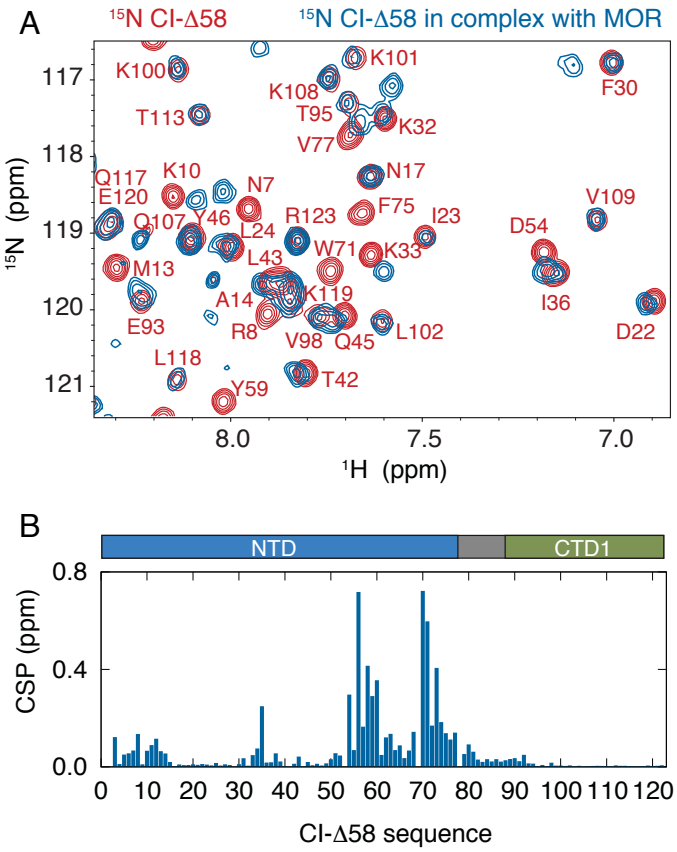
Fig. 5. Structure of the MOR:CI-NTD complex and impact on DNA binding. (*A*) Cartoon representation of the crystal structure of the MOR:CI-NTD complex (180° rotation shown). Residues implicated in the interaction between the two proteins are shown as pink, blue and green (for MOR) and teal, red and wheat (for CI-NTD). (*B*) Sequences of CI-NTD and MOR with residues implicated in the interaction between the two proteins color-coded as in panel *A*. (*C*) The interface between MOR and CI-NTD is dominated by hydrophobic interactions as shown by analysis of the crystal structure of the complex by the software LigPlot. (*D*) Zoom on the polar interactions mediated by Q55 and N56 (red) in CI-NTD at the interface with MOR. A number of hydrogen bonds are formed between Q55 and N56 and the residues Y3, Y5, T42, W43 and F66 in MOR. (*E*) Model of the dimeric CIΔ58-O_L structure obtained on the basis of the crystal structure of CI-NTD in complex with the O_L half-site and the crystal structure of CI-CTD_1 combined with experimental SAXS data (3). The linker connecting CI-NTD with CI-CTD_1 is indicated as gray dashed lines. The crystal structure of MOR:CI-NTD was superimposed onto the structure of one of the CI-NTD domains (marked with a *). MOR is shown as a transparent surface demonstrating that MOR interferes with binding of the CI dimer to two adjacent half-sites on the DNA.

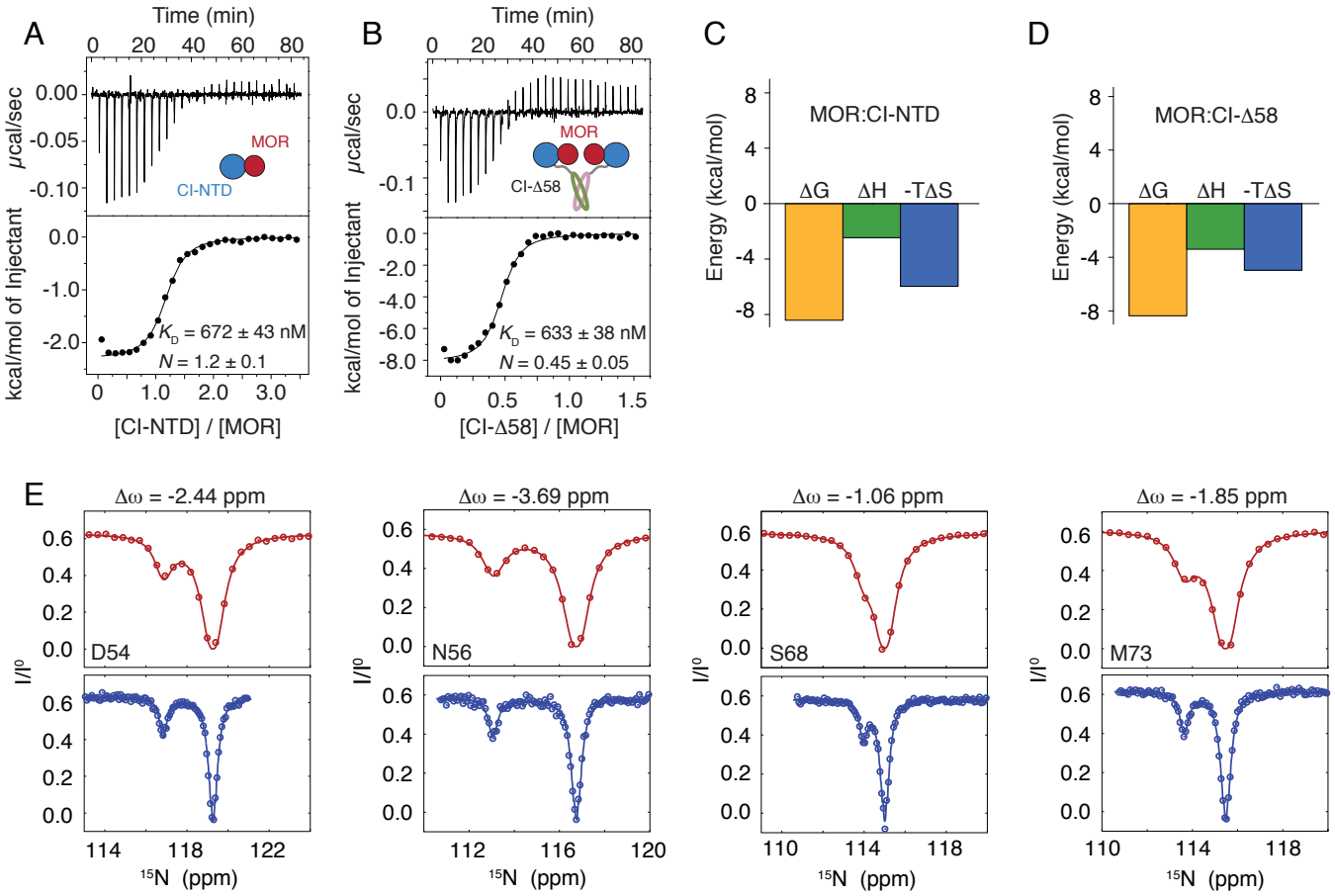
Fig. 6. Multi-timescale dynamics of MOR from ^{15}N relaxation rates and CPMG relaxation dispersion. (*A*) Ratio of ^{15}N R_2 and R_1 in MOR measured at 25°C and a ^{1}H frequency of 700 MHz. (*B*) Exchange contributions, R_{EX}, to transverse relaxation derived from a Lipari-Szabo

model-free analysis of ^{15}N R_1, R_2 and $\{^1H\}$-^{15}N NOEs (*SI Appendix*, Fig. S4). Exchange contributions are observed in three different regions of MOR defined as cluster 1 (red), 2 (blue) and 3 (green). (*C*) Residues experiencing Lipari-Szabo exchange contributions mapped onto the solution structure of MOR. The residues cluster around residue Y5 (shown as cyan sticks). (*D*) Comparison of ^{15}N $R_{1\rho}$ relaxation rates (blue) with effective R_2 rates obtained from relaxation dispersion experiments at high (1 kHz, red) and low (31 Hz, green) CPMG frequencies. All rates were obtained at 25°C and a ^1H frequency of 600 MHz. (*E*) Residues experiencing exchange contributions as detected by CPMG relaxation dispersion. The plot shows the difference between effective R_2 rates measured at low (31 Hz) and high (1 kHz) CPMG frequencies at 600 MHz. The residues cluster in three different regions of MOR (cluster 1, 2 and 3). (*F*) Residues experiencing exchange contributions as detected by CPMG relaxation dispersion mapped onto the solution structure of MOR. The residues cluster around W43 (shown as cyan sticks). (*G*) Examples of CPMG relaxation dispersion curves at 600 and 950 MHz for representative residues in MOR. The full drawn lines correspond to a simultaneous analysis of all the data according to a two-site exchange model.

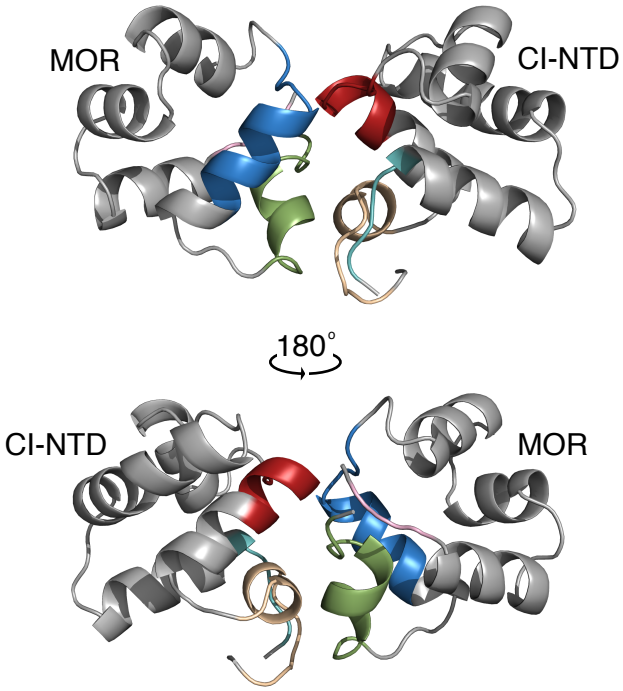

A
cI
mor
O_{M1} O_{M2} O_{M3}
P_L
P_R
O_R
O_L
O_D
O_R 5′-AATTCATATTTCGTGAACT-3′
O_L 5′-AGTTCATGAAACGTGAACT-3′
O_D 5′-AGTTCATTAAACATGAACT-3′
B
CI-Δ58
NTD
CTD1
CTD2
1
80 90
122
180
CI-NTD monomer
CI-CTD1 homodimer
MOR
1
72
C
P_L
O_{M3}
O_R
O_{M2}
O_{M1}
cI
P_R
O_L
mor
O_D
D
P_L
O_{M3}
O_L
mor
O_R
O_{M2}
O_{M1}
cI
O_D
P_R


A
SSP
1.2
0.8
0.4
0.0
-0.4
H1
H2
H3
H4
H5
0
10
20
30
40
50
60
70
MOR sequence
B
K15
Q19
N38
D45
F66
H1
H2
H3
H4
H5
S6
M26
E30
L55
D61
HTH motif
C
D
E
^{13}C: 133.25 ppm
A69Hβ*
Y3Hβ*
A69Hα
N68Hα
Y3Hε*
Y3Hδ*
^{1}H (ppm)
7.10
6.90
F
^{13}C: 120.90 ppm
^{13}C: 113.57 ppm
E47Hβ3
E47Hβ2
W43Hβ
W43Hε3
W43Hδ1
L8Hδ2*
L37Hγ
L37Hδ1*
L37Hβ3
W43Hζ2
7.75
7.55
7.05
6.85


A
15N CI-Δ58
15N CI-Δ58 in complex with MOR
K100
K101
K108
T113
T95
K32
V77
F30
Q117
K10
N7
N17
E120
V109
Q107
Y46
R123
F75
I23
D54
L24
W71
K33
L43
M13
A14
K119
I36
E93
R8
V98
Q45
L102
D22
L118
Y59
T42
117
118
119
120
121
15N (ppm)
8.0
7.5
7.0
1H (ppm)
B
NTD
CTD1
0.8
0.4
0.0
CSP (ppm)
0 10 20 30 40 50 60 70 80 90 100 110 120
CI-Δ58 sequence


A
Time (min)
0 20 40 60 80
μcal/sec
0.00
-0.05
-0.10
MOR
CI-NTD
kcal/mol of Injectant
0.0
-1.0
-2.0
K_D = 672 ± 43 nM
N = 1.2 ± 0.1
0 1.0 2.0 3.0
[CI-NTD] / [MOR]
B
Time (min)
0 20 40 60 80
μcal/sec
0.0
-0.1
MOR
CI-Δ58
kcal/mol of Injectant
0.0
-2.0
-4.0
-6.0
-8.0
K_D = 633 ± 38 nM
N = 0.45 ± 0.05
0 0.5 1.0 1.5
[CI-Δ58] / [MOR]
C
MOR:CI-NTD
Energy (kcal/mol)
8
4
0
-4
-8
ΔG ΔH -TΔS
D
MOR:CI-Δ58
Energy (kcal/mol)
8
4
0
-4
-8
ΔG ΔH -TΔS
E
Δω = -2.44 ppm
D54
Δω = -3.69 ppm
N56
Δω = -1.06 ppm
S68
Δω = -1.85 ppm
M73
I/I⁰
0.6
0.4
0.2
0.0
114 118 122
112 116 120
110 114 118
^{15}N (ppm)

A



B

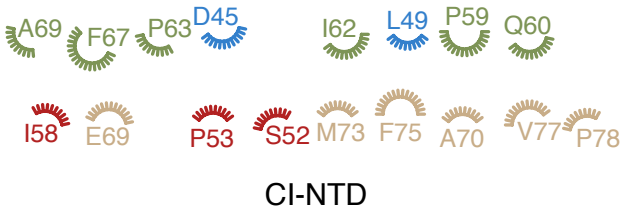
CI-NTD

^{1}MKTDTSNRLK ^{11}QIMAERNLKQ ^{21}VDILNLSIPF
^{31}QKKFGIKLSK ^{41}STLSQYVNSV ^{51}QSPDQNRIYL
^{61}LAKTLGVSEA ^{71}WLMGFDVPMV

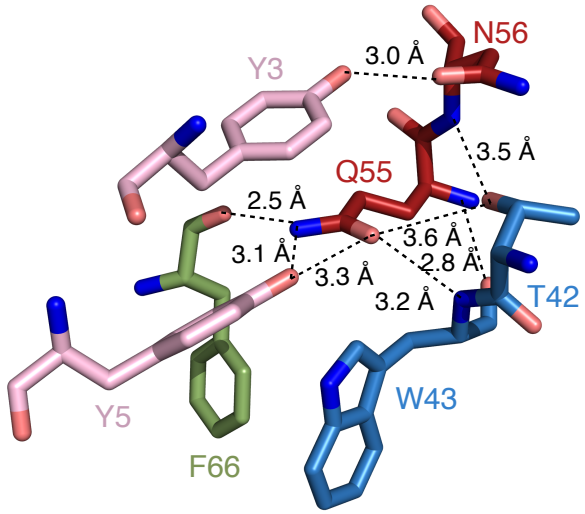
MOR

^{1}MSYDYSSLLG ^{11}KITEKCGTQY ^{21}NFAIAMGLSE
^{31}RTVSLKLNDK ^{41}VTWKDDEILK ^{51}AVHVLELNPQ
^{61}DIPKYFFNAK ^{71}VH

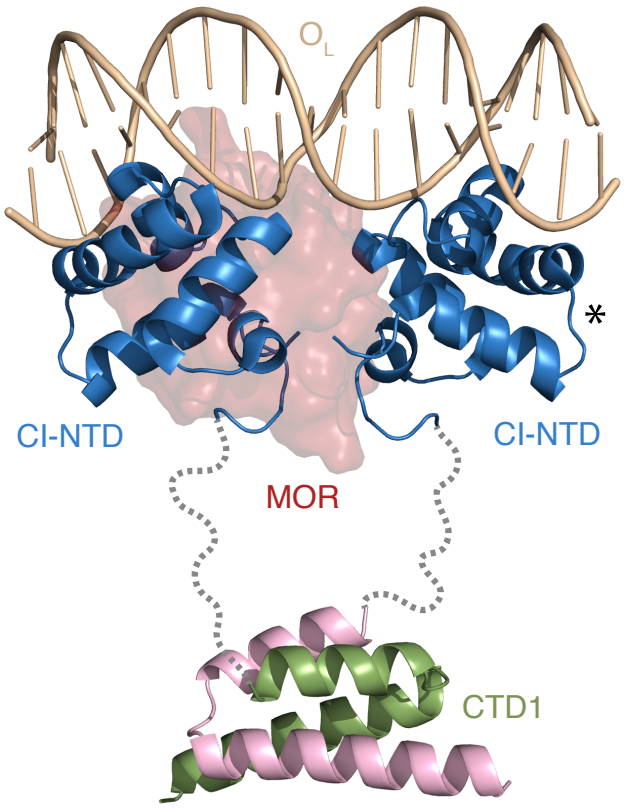
C

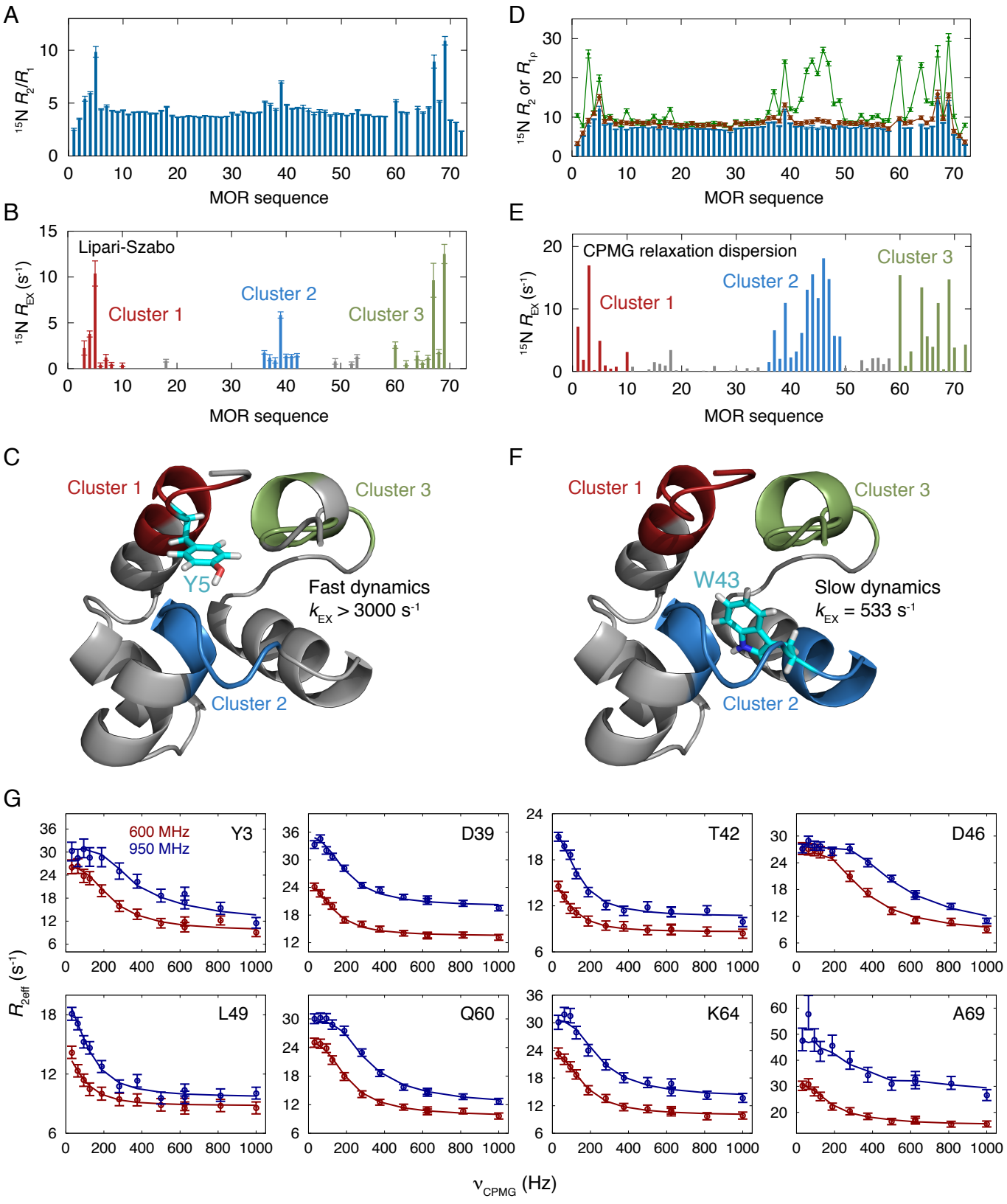


D



E




A
^{15}N R_2/R_1
MOR sequence
B
Lipari-Szabo
^{15}N R_{EX} (s^{-1})
Cluster 1
Cluster 2
Cluster 3
MOR sequence
C
Cluster 1
Cluster 3
Y5
Fast dynamics
k_{EX} > 3000 s^{-1}
Cluster 2
D
^{15}N R_2 or $R_{1\rho}$
MOR sequence
E
CPMG relaxation dispersion
^{15}N R_{EX} (s^{-1})
Cluster 1
Cluster 2
Cluster 3
MOR sequence
F
Cluster 1
Cluster 3
W43
Slow dynamics
k_{EX} = 533 s^{-1}
Cluster 2
G
600 MHz
950 MHz
Y3
D39
T42
D46
L49
Q60
K64
A69
R_{2eff} (s^{-1})
ν_{CPMG} (Hz)