Palanca C.,Rubio V. (2016). Structure of AmtR, the global nitrogen regulator of *Corynebacterium glutamicum*, in free and DNA-bound forms. FEBS J 283(6): 1039-59

Author's last version

**Structure of AmtR, the global nitrogen regulator of
Corynebacterium glutamicum, in free and DNA-bound forms**

Carles Palanca¹ and Vicente Rubio^{1,2}

¹Instituto de Biomedicina de Valencia of the CSIC (IBV-CSIC), Spain

²Group 739 of the Centro de Investigación Biomédica en Red sobre Enfermedades Raras (CIBERER) del Instituto de Salud Carlos III, Spain

***Correspondence to:**

C. Palanca and V. Rubio

Instituto de Biomedicina de Valencia

C/ Jaime Roig 11

Valencia-46010, Spain.

Fax: +34 963 690 800

Tel: +34 963 391 772

E-mails:

cpalanca@ibv.csic.es

rubio@ibv.csic.es

Running title: Structures of DNA-bound and free AmtR from *C. glutamicum*

Abbreviations: ANS, 1-anilino-8-naphthalene sulfonate; DNB domain, DNA binding domain; DTT, dithiothreitol; EDTA, ethylene diamino tetracetic acid; EMSA,

electrophoretic mobility shift assay; HTH, helix-turn-helix; LBD domain, ligand binding and dimerization domain; MBP, maltose binding protein; PDB, Protein DataBank; rmsd, root mean square deviation; SAD, single-wavelength anomalous dispersion;. SeMet, selenomethionine;

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org [PDB ID codes 5DXZ (native AmtR), 5DY1 (SeMet-AmtR) and 5DY0 (AmtR-DNA)].

Keywords: gene repressor, gene regulation, transcriptional regulator, nitrogen regulation, TetR family, X-ray crystallography, crystal structure, PII protein, protein structure.

ABSTRACT

Corynebacterium glutamicum is a bacterium used for industrial amino acid production and understanding its metabolic pathway regulation is of high biotechnological interest. Here, we report crystal structures of AmtR, the global nitrogen regulator of *C. glutamicum*, in apo (2.25Å and 2.65Å resolution) and DNA-bound (3Å resolution) forms. These structures reveal an all- α homodimeric TetR family regulator composed of a helix-turn-helix-hosting N-terminal DNA binding domain and a C-terminal dimerization domain. AmtR presents several unique structural features that appear invariant among AmtR proteins, which may be related to its regulation by the nitrogen-sensing trimeric protein GlnK rather than by small molecule effectors. Compared with other TetR family members, AmtR has an extra C-terminal helix, a large extended external loop that resembles the flexible transducer T-loop of GlnK in sequence, and a large open cavity towards the intersubunit region that changes shape upon DNA binding. The marked kinking of helix 4 decreases in the DNA-bound form. The binding of one AmtR dimer to its DNA operator not only involves the insertion of helices 3 and 3' in adjacent turns of the double-helix major groove, but also the anchoring of 19-residue, arginine-rich and proline-rich N-terminal extensions to two external minor grooves. EMSA assays with a deletion mutant reveal that the 19-residue extension is crucial for AmtR binding to DNA. N-Extension anchoring explains the flanking by AT sequences of the recognized target DNA sequence core. The significance of these findings for the entire TetR family of regulators and for GlnK regulation of AmtR is discussed.

INTRODUCTION

Nitrogen assimilation by *Corynebacterium glutamicum* has great biotechnological importance [1]. This gram-positive bacterium can make and excrete amino acids to the medium, and is used to industrially produce L-glutamate and L-lysine in amounts that have respectively escalated from 1.5 and 0.7 million tons in 2003 [2], to 2.9 and 1.9 million tons in 2013 [3,4]. Consequently, the study of nitrogen-related regulation in this microorganism, including gene expression regulation, also has potential applications [1, 5, 6], making desirable the detailed characterization of the molecular components involved in this regulation.

AmtR is a key component of the nitrogen regulatory system of *C. glutamicum* [7, 8]. This protein is a transcriptional repressor of the TetR family [9] that controls the expression of nitrogen metabolism-related genes [8-10] that are derepressed under ammonia starvation conditions [11,12]. Characteristically, the genes of the AmtR regulon have in their promoters one or various imperfect palindromic sequences (AmtR boxes) to which AmtR binds, which include the central consensus motif CTATn₆ATAG [10,12]. No small molecule effector has been identified as a trigger of AmtR dissociation from DNA [10]. However, electrophoretic mobility shift (EMSA) assays and pull-down experiments revealed that AmtR interaction with the homotrimeric protein of the PII family GlnK [13] triggers the dissociation of AmtR from its target DNA [10]. Indeed, PII proteins sense ammonia abundance by having 2-oxoglutarate (2OG), an inverse indicator of the ammonia levels, as an allosteric effector [13-17]. In fact, under conditions of nitrogen limitation (the same conditions that increase 2OG levels), *C. glutamicum* GlnK is adenylylated at Tyr51 [18], the nearly constant residue of the signal-transducing flexible T-loop of PII that in many bacterial species is uridylylated or adenylylated in a glutamine-controlled process

[13,14,19]. In *C. glutamicum*, only the form of GlnK having Tyr51-AMP in its three subunits has been reported to interact with AmtR and to trigger the dissociation of this repressor from its target DNA [10]. Therefore, this adenylation, rather than 2OG binding to GlnK by itself, appears to trigger AmtR regulon activation. In any case, the sensing of ammonia levels by AmtR appears to be indirect, via the interaction of AmtR with GlnK-AMP₃ [10].

The mechanism by which the derepression of the genes of the AmtR regulon takes place is not well understood, largely because the structure of either AmtR or the AmtR/DNA and AmtR/GlnK complexes remained unknown. As a first step towards this understanding, we have determined here the crystal structures of AmtR in free form and bound to its consensus target DNA sequence. This study has added an important example to the still limited list of TetR family members for which the structures of both the DNA-free and DNA-bound forms have been determined [20-27]. Our present work, in addition to clarifying the structural bases of the mechanism for specific recognition by AmtR of its target DNA, has revealed that an N-terminal 19-residue extension of this protein is a key element for binding of AmtR to DNA through anchorage in the minor groove of the DNA double helix. The importance of this 19-residue extension is confirmed here by deletion mutagenesis and EMSA assays. These findings may help understand the DNA binding of many other TetR family members which apparently have similarly sized or even larger N-terminal extensions [24].

The mechanisms for modulating the ability to bind to DNA of the transcriptional regulators of the TetR family are quite varied [20-26]. Even the way of DNA binding is varied within this family, since some members bind to DNA as pairs of dimers by attaching to the target DNA box cooperatively [21, 23,26,27], whereas other members bind as single

dimers. Our results show that AmtR binds to its DNA box as a single dimer and that its DNA binding domains (DNB domains) experience movements that might be involved in the process of AmtR dissociation from DNA. In turn, several unique structural features of the dimerization domain (abbreviated in this family as LBD domain) of AmtR might be related to AmtR modulation by protein-protein interactions with GlnK. These features include the existence of a C-terminal α -helix, of a 16-residue extended loop that connects helices 8 and 9, and of a large open cavity on the intersubunit boundary that varies in shape upon AmtR binding to DNA. All these newly learned structural features of AmtR open up novel ways for the experimental study of how AmtR and GlnK can associate.

RESULTS

Crystal structures

Crystals of recombinant, N-terminally His-tagged AmtR (His₆-AmtR) grown in the absence of DNA and labeled with SeMet (Apo-SeMet, Table 1) (see Material and Methods), diffracted X-rays at 2.65 Å resolution. After phasing by the single-wavelength anomalous diffraction approach (SAD), they yielded a model at this resolution for one AmtR homodimer (Table 1 and Fig. 1A) with nearly identical subunits (rmsd, 1.02 Å for 221 superimposed C $_{\alpha}$ atoms), except for differences in the kinking of one helix (helix 4) (Fig. 1B).

The crystals of His₆-AmtR not labeled with SeMet (Apo-native, Table 1) diffracted X-rays at 2.25 Å resolution and belonged to a different space group (P3₂21) than those of the Apo-SeMet crystal. Molecular replacement using as template one subunit of Apo-SeMet yielded a solution that consisted in one subunit in the asymmetric unit. The final model for this subunit was obtained at 2.25 Å resolution and was virtually identical to that of subunit

B of the SeMet-AmtR crystal (rmsd, 0.68 Å for 197 superimposed C_α atoms). After applying crystal symmetry, a homodimer was built (Fig. 1C) which closely resembled the dimer observed in the Apo-SeMet crystal (rmsd, 0.97 Å for the superposition of 394 C_α atoms of both dimers). The His₆-tag, residues 1-18 and the two C-terminal residues were not visible in the Apo-SeMet and Apo-native structures, possibly indicating the existence of disorder in these regions. In the Apo-native structure residues 86-89 (which connect helices 4 and 5) were also not observed. (Fig. 1C).

We also determined the structure at 3 Å-resolution of AmtR bound to DNA (Fig. 1D). Crystals were obtained from a mixture of C-terminally His₆-tagged AmtR (AmtR-His₆) and a self-complementary double-helical quasi-palindromic oligonucleotide, 5'ATTATCTATAGATCTATAGATAATGC3' which contained an idealized AmtR box [10,12] (the central, stronger consensus is underlined), followed by a GC pair. In this case the solution obtained by molecular replacement yielded two parallel DNA double helices and two AmtR dimers in the asymmetric unit (Fig. 1D, boxed). The solvent content was very high (82%), and the two AmtR dimers bridged both DNA double helices by coming into mutual contact through the opposite end of their molecules relative to the DNA-bound end (Fig. 1D). This arrangement, in which a single AmtR dimer was bound per AmtR box, contrasts with the binding of two dimers to a single target sequence in some members of the TetR family [21, 23, 26, 27]. In the AmtR-DNA crystal the AmtR dimers and the DNA duplexes of the asymmetric unit were related by approximate non crystallographic 2-fold symmetry around an axis that was perpendicular to the plane containing the molecular two-fold axes of both AmtR dimers (Fig. 1D, E). Actually, the whole crystal consisted of parallel long straight DNA double strands formed thanks to the cohesive end-to-end linkages provided by the terminal CG bases, which were bridged by the contacting AmtR

dimers (Fig.1D). The modest surface buried by the dimer-dimer interactions ($435 \text{ \AA}^2$ per dimer; Fig. 1E) and the thermodynamic predictions of the PISA server ($\Delta G^{\text{diss}} = -5.7$ kcal/mol, [28]) suggested that these dimers of dimers are not very stable. Indeed, gel filtration chromatography showed (Fig. 1F) that upon AmtR-His₆ addition, the DNA peak increased in the mass expected for the binding of a single AmtR-His₆ dimer (52 kDa). Nevertheless, it cannot be ruled out that AmtR dimers of dimers like those observed in the crystal with DNA could be formed *in vivo* in chromosomal regions having high densities of AmtR boxes.

Overall structure and unique features of AmtR

As is characteristic of TetR family members [9], AmtR is a homodimeric all- α protein in which each subunit is composed of two domains (Fig. 1A, C). The N-terminal domain (residues 1-60) consists of a variable (relative to other proteins of the family [24]) 19-residue arginine-rich and proline-rich N-terminal extension which was observed only in the DNA-bound structure (see section on DNA binding); this was followed by the canonical DNB domain (residues 20-60) reported in all structurally well characterized members of this family [9], with its three helices, including a helix-turn-helix (HTH) motif of DNA binding (Fig. 1C, 2A). This DNB domain core superposed extremely well with the DNB domain reported for other members of the same family (Fig. 2A).

The C-terminal domain, known in this family as the ligand-binding and dimerization domain (LBD domain, Fig. 1C), presented some unique features in AmtR compared to other structurally well characterized members of this family. Thus, it had an extra C-terminal helix (helix 10, residues 208-218; Figs. 1C and 2B), extending to 12 the number of helices of the helix bundle conformed by both LBD domains. This extra helix 10 was the

most external one relative to the 2-fold axis of the dimer (Fig. 1C). It interacted with the $\alpha 8$ - $\alpha 9$ loop, which was also a unique AmtR feature (Figs 1C and 2B). Sequence comparisons (Fig. 2C) strongly suggest that helix 10 is constantly found in corynebacterial AmtR, and that the 7-residue turn that connects it to helix 9 is highly conserved, including an invariant proline (in some sequences there is a second proline; shown red over yellow background in Fig. 2C). Thus, there appears to be a strong selective pressure for the existence of this 7-residue turn, and to have, therefore, a separate helix 10. An ancestral longer helix 9, which is frequent among TetR family members (e.g., see TetR from *Escherichia coli*, Protein DataBank (PDB) accession code 1QPI [20]) could have been broken by this 7-residue turn when AmtR emerged.

The already mentioned long $\alpha 8$ - $\alpha 9$ loop (16 residues; amino acids 171-186; Figs. 1C and 2B) was shaped in a way that was reminiscent of two blades of a 3-blade propeller (Fig. 1C). It extended over its own subunit, one blade in contact with the DNB domain (via helix 1) and the other blade in contact with helix 10 (Figs. 1C and 2B), connecting the opposite ends of the subunit. The sequence of this loop is also quite conserved among corynebacteria (Fig. 2C). It contains at least two and even 3 or 4 proline residues, as well as highly polar residues. Once again, this is suggestive of high evolutionary pressure toward maintaining this region as an extended loop, which would be expected if this loop were functionally relevant. A disc-shaped cavity of 8 Å in diameter, rimmed by numerous polar atoms, was conformed in the more distal blade of this loop (labeled in Fig. 3A, left, as *small cavity a*), whereas a less well defined cavity of a similar size was observed (labeled in Fig. 3A, left, as *small cavity b*) between the base of the proximal blade of this loop and helices 8 and 9 of the same subunit. These cavities could accommodate ligands, although these ligands would be small, as confirmed by the lack of significant fluorescence output when the proposed [9]

fluorescent ligand-mimetic compound for TetR family proteins, 1-anilino-8-naphthalene sulfonate (ANS), was added to AmtR (Fig. 3B). In contrast, ANS gave the expected response with a typical TetR family protein, TtgR [9] (Fig. 3B).

TetR family member RutR, an uracil-sensing transcriptional repressor of *E. coli* [29], has overall structural similarity with AmtR although it does not share the unique features of the LBD domain of this protein (Fig. 3C). In particular, the $\alpha 8$ - $\alpha 9$ loop is replaced in RutR with an 8-residue transversal helix ($\alpha 8^*$) that sits on the adjacent subunit (Figs. 3C and close-up in Fig. 3D) to close the uracil site, located in the usual ligand binding region of TetR family proteins [9]. In contrast, the $\alpha 8$ - $\alpha 9$ loop of AmtR came in near-exclusive contacts within its own subunit (Fig. 1C, 3C), and left exposed the region covered by $\alpha 8^*$ in the adjacent subunit of RutR. Thus, a roughly conical cavity of approximately 450 Å³ with a highly polar surface was exposed in AmtR at the intersubunit junction, being a characteristic feature of this protein (Fig. 3A, left panel). This large cavity, defined between the ends of $\alpha 5$ and $\alpha 6$, the first three turns of $\alpha 7$ and most of $\alpha 8$ in one subunit, and the initial part of $\alpha 6'$, the end of $\alpha 8'$ and the beginning of the $\alpha 8'$ - $\alpha 9'$ loop in the other subunit (Fig. 1C), became narrower and perforated at its bottom in the DNA-bound form, a form in which both large cavities of the dimer interconnected to form a real tunnel (Fig. 3A, right panel). Given its different conformations in the DNA-bound and DNA-free forms of AmtR, this cavity could be involved in triggering the dissociation of AmtR from DNA if GlnK binding would stabilize the roughly conical form of this cavity, present in the form of AmtR that is unfit for binding to DNA (see below).

Another non-exclusive but highly characteristic attribute of AmtR which also changed with DNA binding was the variable kinking at Pro77 (an invariant residue among corynebacteria) at the middle of helix 4, the helix that connects the DNB and LBD domains

(Fig. 1B). The magnitude of the angle at this kink (determined with HELANAL [30]) was 27°, 20° and 31° in the Apo-native form and in subunits A and B of the Apo-SeMet form, respectively (Fig. 1B). Since the angle at the kink in all four subunits found in the asymmetric unit of the DNA-bound form was only ~20° (Fig. 1B and see below), this kink appears to be either involved or affected by the conformational processes associated with DNA binding. Indeed, in a protein of the TetR family that senses heme (HrtR) [25], changes in helix 4 are a key mechanism for determining the conformation of the DNB domain and, thus, to modulate DNA binding.

The DNA complex and DNB domain movements.

In the structure of the AmtR-DNA complex, helices 3 and 3', which belong to the HTH motif of each subunit of the dimer, were inserted into the two successive turns of the major groove of the double helix that hosted the quasi-palindromic central part of the AmtR box (Fig. 4A-C). The distance between these helices (measured at the level of the C α of Tyr57, the residue that corresponds to Tyr52 of TetR) was ~37.2 Å (Fig 4C), which was slightly longer than in the corresponding DNA complexes of the TetR family members QacR [21], CgmR [23] and SimR [24] (respective distances, 36.9 Å, 36.3 Å and 36.8 Å). Since the distance between two turns of the major groove is 34 Å in ideal B-DNA, binding of TetR family proteins elongate the double helix somewhat (nearly 10% for AmtR). AmtR also widened the major and minor grooves of DNA (Fig. 4D), by 4.3 % and 5.0 % respectively, compared to ideal B-DNA. The plot of width versus sequence (Fig. 4D) revealed that the anchoring of the N-terminal extensions (see below) to the minor groove widened this groove.

AmtR binding to DNA did not result in substantial DNA bending (Fig. 4A, green

line; corresponds to 5° bending, estimated as in [31]). Accordingly, the average step roll angle was close to zero (0.6°, estimated using the web3DNA server <http://w3dna.rutgers.edu> [32]). Lack of substantial bending might be expected for a transcriptional repressor that acts by preventing the binding of the transcriptional machinery, contrasting with the ~90°-bending induced on the DNA by CRP [33], another homodimeric gene regulatory factor having an HTH motif that is one of the paradigms for bacterial transcriptional activators.

The distance between helices 3 and 3' was longer and more diverse in the two apo forms of AmtR (40.6 Å and 41.5 Å for the Apo-native and the Apo-SeMet forms, respectively) than in the two DNA-bound forms (37.1 Å and 37.3 Å) (Fig. 4C). This reflected the existence of mobility of the DNB domains (Fig. 5A). Relative to the apo forms, the DNB domains exhibited in the DNA-bound form an 8°-12° rigid body rotation (determined using DynDom [34]) around an axis that crossed transversally helices 1 and 4 at their respective residues Ala29 and Leu68 (Fig. 5A). The hinge for this movement was located at the junction between the two domains (i.e., at the connection between helices 3 and 4, residues 60 to 63). This movement, which resulted in the mutual approach of helices 3 and 3' in the dimer by >4 Å (Fig. 4C), placed these helices in position for interaction with the sequence-determining bases of the AmtR box, found in two consecutive turns of the major groove of DNA.

An additional movement, observed when comparing the DNA-bound forms of AmtR with the Apo-native form and with subunit B of the Apo-SeMet form, was the pulling of helix 4 toward the DNB domain, which resulted in not only a smaller angle at the kink found in the middle of this helix (~20° instead of 27°-31°), but also in a vertical displacement of this helix of up to 3.3 Å (Fig. 5B). This brought about an increase in the

longest dimension of the subunit in the DNA-bound form (for example, the distance between the C α atoms of residues 52 and 91 increased by nearly 2 Å) (not shown). This change in length and in helix 4 kinking could be due to the bridging of the two DNA double helices which ran in parallel in the crystal by the AmtR dimer of dimers (Fig. 1D), since one end of each AmtR dimer was anchored to the DNA and the opposite end was anchored to the other AmtR dimer via the mutual interactions of the α 9- α 10 loops of the four subunits (Fig. 1E). However, in the TetR family protein HrtR the changes in helix 4 are key determinants of the binding to DNA of this protein [25]. Therefore, the direct involvement of the change in the kink of helix 4 of AmtR in relation to fixing the DNB domain in a competent conformation for DNA binding cannot be ruled out. In any case, if helix 4 was excluded, no differences in the relative positions of the LBD domains of the two subunits in the dimer were observed in the apo and the DNA-bound forms (Fig. 5C), with very good superposition of their C α atoms (rmsd values for the dimers of the apo and the DNA-bound forms after excluding the DNBs and helix 4, 0.98 ± 0.02 Å for 268 C α atoms). The only element that exhibited some movement was helix 6 of the two subunits (Fig. 5C). This helix interacted with the DNB domain and was a key element in determining the shape of the large cavity shown in Fig. 3A. Since this helix plays a pivotal role in TetR in determining the positions of the DNA binding helices of the neighboring DNB domain [20], the changes observed in this helix in AmtR might have some bearing on the mechanism of gene regulation by this repressor.

Structural bases of the DNA binding specificity of AmtR

The structure of the DNA-AmtR complex (Fig. 6A) revealed the key role of helix 3 in determining the specificity of this repressor for the central and most conserved part of the

AmtR box, $^{-7}\text{CTATnnnnnnATAG}^{+7}$ [10] (where n is any base, and bases are numbered negative and positive by respectively counting upstream or downstream from the mid-point of the box; see also the numbering in Fig. 4A and on the X-axis of Fig. 4D). Helix 3, encircled by the specificity-determining bases of the major groove of the double helix, interacted with the DNA via six DNA sequence-restricting residues, Arg52, Gln53, Ala54, Ser55, Tyr57 and Tyr58 (Fig. 6A). A C-to-T change at position -7 (which corresponds to base 6 of the oligonucleotide used in the present studies) would cause a clash between the methyl group of the thymine with the side-chain of Ser55 (firmly held in place by hydrogen bonding with the -7 to -8 phosphate bridge), thus cytosine would be preferred at the -7 position. Pyrimidines are favored at positions T^{-6} and T^{+5} because of the steric complementarity with the side chain of Ala54. A thymine is favored at position -6 because its methyl group was virtually clamped between the side-chains of Ala54 and Arg52. In this clamp the side chain of Arg52 was firmly fixed in place by charged hydrogen bonding with the -7 to -6 phosphate bridge. Gln53 favored an adenine over any other base at the +4 position, by making hydrogen bonds via its side chain with the amine group and the N7 atom of this adenine. The side chains of tyrosine residues 57 and 58, fixed in place by hydrogen bonding to phosphates via their hydroxyl groups, acted as base-holding platforms. Thus, Tyr57 contacted the N7 and C8 of A^{+4} and the methyl group of T^{+5} and rendered unfavorable the T-to-C substitution at the +5 position. Similar interactions of Tyr58 with T^{-8} and A^{-9} (the two bases upstream of the start of the central and most conserved part of the consensus) selected for thymine and against cytosine at the -8 position, which agrees with the more extended version of the consensus [12].

Despite not being sequence-determining, the DNA-protein interactions mediated by the ribose-phosphate backbone of the -3 to +3 nucleotides (the central n_6 stretch of the

AmtR box), and of nucleotides found upstream from the consensus (see below) should greatly contribute to the binding affinity. Thus (Fig. 6A) in addition to the already detailed phosphate-mediated interactions, the +2 to +3 phosphate bridge interacted with Lys63 and Thr42 (both non-helix 3 residues), the -7 to -8 phosphate bridge was hydrogen-bonded to main-chain N atoms of Arg52 and Arg10, the -8 to -9 phosphate bridge made hydrogen bonds with nitrogen atoms of the imidazolic and guanidino group of His59 and Arg9, while this last arginine also made extensive stacking contacts with the ribose at position -7.

A 19-residue N-terminal extension anchors AmtR on the minor groove of DNA

Preceding helix 1, AmtR has the 19-residue N-terminal extension of sequence MAGAVGRPRRSAPRRAGKN. This extension, which is very rich in positively charged residues and hosts three proline residues, was not visible in the structures of the apo forms, indicating that it was disordered in these forms, as predicted by the RONN server (www.bioinformatics.nl/~berndb/ronn.html) [35] (Fig. 6B). In contrast, this extension was well-defined in the structures of the four AmtR subunits in the AmtR-DNA complex (Figs. 1D, 4A,B and Figs. 6C-F), with 6 residues near its N-terminus (³GAVGRP⁸; these residues will henceforth be called "the anchor") anchored to the minor groove of DNA (Fig. 6E) that flanks laterally the turn of the major groove that accommodates helix 3 or 3' (Fig. 4B). In subunits A, C and D of the DNA-containing crystal, the anchor sat in the cognate double helix. However, in subunit B it was inserted into an adjacent DNA double helix which sat in a normal direction near the cognate DNA (Fig. 6C). The interactions of the anchor with the DNA were identical whether it sat in the cognate DNA as in subunits A, C and D, or in the nearby double helix from another DNA-AmtR complex, as in subunit B (Fig. 6D). The interactions of the distal part of the N-terminal extension with the remainder of the protein

(which exclusively involved contacts with the DNB domain of the same subunit, Fig. 6F) were also identical in the cognate DNA-binding form and in the neighbor DNA-binding form of the extension (not shown). The differences between these two forms resulted from a rigid body motion of the more N-terminal part of the extension relative to its distal part (or vice versa, Figs. 6C, D). This motion was due to rotations in the main chain bonds at amino acids 11-13. Given its 3:1 frequency in the DNA crystal, and since it appears the more likely *in vivo*, the structure of the anchoring to the cognate DNA will be discussed here.

The structure of the N-terminal extension was identical in subunits A, C and D (rmsd < 0.5 Å for 15 C α atoms). The anchor, of which a very prominent part was the fully extended side-chain of Arg7, filled the minor groove between base pairs A⁻¹²:T⁺¹² and C⁻⁷:G⁺⁷, (Figs. 4B, and 6E,F), and was sandwiched between both sugar phosphate chains, and making intimate contacts via Val5 and Gly6 with base pairs T⁻¹¹:A⁺¹¹ and T⁻¹⁰:A⁺¹⁰, respectively, and via Arg7 with base pairs A⁻⁹:T⁺⁹ and T⁻⁸:A⁺⁸ (Figs. 6E,F). Such contacts do not favor the substitution of these bases by C:G pairs, since these three anchor residues would clash with the amino group of the guanine in the event of any of these A:T pairs being substituted for a G:C pair. These interactions of the anchor with the bases, together with the aforementioned interactions of Tyr58 with T⁻⁸, cooperated in conferring AmtR selectivity for AT bases at the positions preceding and following the major groove recognition sequences of the AmtR box, which are characteristic of the extended recognition sequence for AmtR [12].

The N-terminal anchor was linked to the DNB domain by a narrow connector (residues 8-10, Fig. 4B) that crossed the medial rim of the minor groove (the phosphoribose chain) making the already indicated connections with DNA phosphates via Arg9 and Arg10

(Fig. 6A), while residues 9-18 were glued to the core of the DNB domain via interactions with helices 1 and 2 and with the $\alpha 2$ - $\alpha 3$ connector (Fig. 6F). The richness in arginine residues of this extension appeared particularly important for gluing this anchor to both DNA and the DNB domain core, since Arg7 was inserted into the minor groove, Arg9 and Arg10 interacted with DNA phosphates and protein residues, including an ion pair between Arg9 and Glu23 (Fig. 6F), Arg14 capped the negative C-end of helix 2, and Arg15 formed an ion pair with a negatively charged patch (Asp26, Glu30) on the outward side of helix 1 (Fig. 6F).

The importance of the N-terminal extension for DNA binding was corroborated by EMSA assays in which the deletion of residues 2 to 15 abolished the binding of AmtR to a 170-bp double helical DNA (the promoter of the *amtB-glnK-glnD* operon, bases -88 to -258 relative to the translation initiation point) hosting three non-overlapping AmtR binding boxes [12]. No binding was observed even when the concentration of the mutant repressor was 8-fold greater than that needed for saturation of all three boxes by wild-type AmtR (Fig. 6G).

DISCUSSION

Although AmtR was purified and crystallized [36], no crystal structure of this transcriptional repressor had been reported. The present structures confirm the sequence-derived inference that AmtR is a member of the TetR family. However, AmtR presents some unique features that may provide targets for further investigation of the mechanisms by which its dissociation from DNA is triggered.

An aspect that has been fully clarified structurally here is how AmtR binds to DNA and the mechanisms used to ensure specificity. As other members of its family, such as

TetR [20], or as other homodimeric transcriptional regulators that utilize an HTH DNA binding motif such as CRP [33], AmtR utilizes for DNA binding a precisely spaced and positioned pair of symmetrical helices from its DNB domains (helices 3 and 3'). These helices are inserted into two adjacent turns of the major groove of the DNA double helix which host the two halves of the quasi-palindrome that is recognized by AmtR. Unlike some other members of the TetR family [21,23,26,27], a single AmtR dimer binds to its DNA box instead of binding as two dimers. The functionality *in vivo* of the DNA-binding machinery revealed by the present structures is supported by the observation that the AmtR residues of helix 3 identified here as involved in the recognition of the DNA target sequence, Arg52, Gln53, Ala54, Ser55, Tyr57 and Tyr58, had all been proven to be important for DNA binding in prior site-directed mutagenesis studies to alanine (or to glycine in the case of Ala54) that showed that their mutations negatively affected the binding affinity of AmtR for its target DNA [35]. The need to precisely position in the DNA the two half-palindromes of the AmtR box recognized by the sequence-determining residues of helix 3 explains the exact requirement for 6 intervening nucleotides along the double helix, as best illustrated by the reactivation of the inactive remnant AmtR box of the *amtR* promoter by the artificial lengthening of its intervening sequence from 5 to 6 nucleotides [12].

An important difference in the mechanism of DNA binding of AmtR compared to other dimeric transcriptional regulators that utilize an HTH DNA binding motif, including the paradigm of its own family, TetR [9,20-23, 25-27], or CRP family members [33], is the important role played in AmtR by its 19-residue N-terminal extension. This extension is disordered while the protein is not DNA-bound, becoming organized and acting as an anchor that hooks on the minor groove of DNA when AmtR attaches to the DNA double

helix. Since most of the AmtR helix 3 residues that are involved in DNA binding make Van der Waals contacts rather than hydrogen bonds with the specificity-conferring bases, extra strength may be needed for DNA binding, which could be provided by anchoring the N-terminal extension to the minor groove of DNA. Indeed, the abolition of DNA binding when residues 2-15 of this extension were deleted strongly supports the importance of this anchoring. The reduced DNA binding affinity previously observed with the Glu30Ala mutation of AmtR [37] could also reflect the importance of this anchor, since, as revealed by the structure of the DNA-AmtR complex, Glu30 makes a salt bridge with Arg15, and should be expected to play an important role in the gluing of the N-terminal extension to the DNB domain core.

Apart from increasing the affinity for DNA, the N-terminal extension might also guide AmtR to find its canonical box. The sequence restrictions revealed by the present structure for fitting the N-terminal extension anchor in the small groove of DNA justifies the lengthening of the AmtR box by several A:T pairs on both ends of the core sequence which is recognized in the major groove of DNA [12]. Perhaps one role of the N-terminal extensions is to also search for AT-rich islands in the minor groove of the GC-rich *C. glutamicum* chromosome (>50% GC content, <http://insilico.ehu.es/oligoweb/index2.php?m=all>). The DNA-exploring ability of the N-terminal extension could result from the displacement of both anchors (one per subunit) along the DNA minor groove until a proper homing site is found. Because of its length, in its disordered state the N-terminal extension would allow exploration of a volume around AmtR generously exceeding the size of an AmtR dimer. Even in its folded state a substantial space could be sensed, given the ability of rotation of the DNA-binding part relative to the DNB domain core-gluing part of the N-terminal extension (Fig. 6C,D).

In any case, since bioinformatic analyses [24] suggest that arginine-rich N-terminal extensions as long as that in AmtR exist in >25% of TetR family gene regulators, the observations made on AmtR could be applied to many other TetR family members. Indeed, an even longer N-terminal extension was proven important for DNA binding in SimR, a TetR family factor from *Streptomyces antibioticus* (another organism with high GC content) that, like AmtR, largely uses van der Waals contacts for the major groove interactions between DNA and its helix 3 residues [24]. Despite the very important differences found in the sequence and length of the N-terminal extensions of AmtR and SimR, the comparison of the structures of the DNA-bound forms of these two proteins allow us to infer as common traits which could be present in other TetR family members, the involvement of an extended arginine as part of the DNA anchor, which in both cases sat on the minor groove of DNA; the gluing of a part of the extension on the DNB domain core; and the importance of the positively charged residues of the extension for this gluing and for interacting with DNA phosphates.

While the present results have largely clarified how AmtR is bound to DNA, they have failed to shed light as to how AmtR is released from DNA to allow the transcription of the AmtR-regulated genes. Since AmtR binding to DNA requires both helix 3 and the N-terminal extension, hampering the binding to DNA of any one or both these elements could release AmtR from the double helix. Stabilization of a conformation in which helices 3 and 3' occupied an inappropriate position, and/or were at an inadequate mutual distance for binding on the target DNA box has been the mechanism revealed to date for the release from DNA of the few TetR family members for which the regulatory process has been well characterized structurally [20, 22, 24, 25]. Given the movement of the DNB domain around a hinge at the junction between domains LBD and DNB, fixing of this junction could freeze

the DNB domains in an inappropriate position for DNA binding. However, the release mechanism from DNA of SimR is a 16° rotation of one subunit with respect to the other subunit causing the misplacing of the DNA-binding helices [24], rather than the hinge motions of the DNB domain that are similar in this protein to those in AmtR. If AmtR dissociation from DNA were caused similarly to that of SimR, the rotation would have to be triggered by the binding to AmtR of its allosteric effector GlnK. Since this effector has to be in the form in which Tyr51 of its flexible T-loops is adenylylated [10], and given the fact that the Tyr51Phe GlnK mutation abolishes the sensitivity of *C. glutamicum* to nitrogen limitation by [38], it appears that GlnK uses its T-loops to interact with AmtR, as might have been expected, given other known examples of interactions of PII proteins with their protein targets [39-44]. Interestingly, since the length and sequence of the T-loop and that of the $\alpha 8$ - $\alpha 9$ loop of AmtR are somewhat similar (Fig. 2D), a T-loop from PII could conceivably replace the $\alpha 8$ - $\alpha 9$ loop, causing the release of the $\alpha 8$ - $\alpha 9$ loop from its binding on its own subunit. This could allow the interaction of this loop with the other subunit to thus trigger an altered relation between both subunits of the AmtR dimer that could be similar to the subunit rotation in SimR. In fact, in other TetR family members such as RutR [29] and SimR [24] the sequence that connects the helices that correspond to helices 8 and 9 of AmtR is projected on the other subunit, interacting with it, wrapping in SimR the other subunit and being centrally involved in the subunit rotation that triggers the dissociation of this protein from DNA [24].

It is also noteworthy that, in crystal structures of *E. coli* GlnK with its protein target AmtB, this ammonia channel was obliterated by Arg47 of GlnK, which adopted a fully extended conformation [39,40]. This extended arginine might replace the extended arginine of the N-terminal extension of AmtR which is inserted in the minor groove of DNA, thus

preventing the anchoring of the N-terminal extension to the DNA. The nearby T loop-linked AMP could, in turn, interact with the arginines (Arg9 and Arg10) of the N-terminal extension, which interact with the phosphate groups at the rim of the minor groove of DNA in the AmtR-DNA complex. In this way, GlnK would act as a decoy for preventing the homing of the N-terminal extension on DNA, thus triggering AmtR dissociation from the DNA.

Finally, as already indicated, changes in helix 4 and helix 6 have been identified [20, 22] in other proteins of this family as being mediators to trigger the DNB domain changes that result in DNA dissociation. Since these two helices are the only elements of the LBD domain of AmtR that exhibit conformational differences between the DNA-free and DNA-bound forms, it is conceivable that these conformational changes might also mediate DNA dissociation from AmtR. If this were the case, the triadenylylated form of GlnK would have to stabilize the DNA binding-incompetent conformation of these elements of the LBD domain. The present structures, and the identification as unique elements of AmtR of both the $\alpha 8$ - $\alpha 9$ loop and helix 10, render these elements as potential candidate points of GlnK interaction with AmtR that merit experimental investigation.

Whichever the mechanism of GlnK association with AmtR, this association poses the symmetry problem of linking a homodimeric protein to a homotrimeric ligand, which would suggest that the complex consists of two PII trimers connected by three intervening AmtR dimers. The two other proteins of the TetR family that have been shown to interact with other proteins, DhaS of *L. lactis* [45] and SlmA [27,46] of *E. coli*, do not provide any clues about GlnK binding to AmtR. The second of these examples seems irrelevant for GlnK-AmtR binding since it is likely to represent an exclusively structural mechanism to prevent the cytokinesis orchestrated by FtsZ (the partner protein that interacts with SlmA).

The DhaS transcriptional activator of *L. lactis* appears to be a closer example relative to AmtR since this transcriptional activator of the TetR family senses the dihydroxyacetone levels via its interaction with a paralog of dihydroxyacetone kinase, the homodimeric protein DhaQ, with which it makes a stable complex [45]. This resembles the sensing of nitrogen abundance by GlnK with AmtR as the effector [10]. With the DhaS/DhaQ pair, binding of dihydroxyacetone to DhaQ results in the activation of the transcription of the Dha operon by DhaS [45]. Although the structures of DhaS and DhaQ have been determined [45], no structure of the DhaS/DhaQ complex has yet been reported. Therefore, no inferences can be made that could be extrapolated to the AmtR-GlnK complex.

In summary, the present studies have clarified the structure of AmtR bound to DNA, have highlighted the importance and determined the role of its long N-terminal extension in DNA binding, and have also uncovered some unique structural features of the AmtR protein itself, which may be related in as yet non-clarified ways to the utilization by this repressor of another signaling protein, GlnK, as its allosteric effector. Further experimental studies, particularly the determination of the structure of the complex of AmtR with adenylylated GlnK, at high resolution or by utilizing lower resolution techniques such as SAXS or electron microscopy, may be necessary to clarify how GlnK can control the function of AmtR.

MATERIALS AND METHODS

Cloning of the *amtR* gene

The AmtR protein used in these studies is encoded by the *amtR* gene from *C. glutamicum* ATCC 13869 (genomic DNA for this strain was kindly provided by Dr. José Antonio Gil, University of León, Spain). The sequencing of this gene revealed that the

AmtR protein of this strain exhibits four amino acid changes (Arg34His, Val141Ile, Thr87Ile and Pro203Ser), relative to the Uniprot KB database deposited sequence (**Uniprot Protein Database ID: H7C699** accession number) for AmtR of *C. glutamicum*, which corresponds to strain ATCC 13032 of this microorganism.

For His₆-AmtR production, the *amtR* gene from *C. glutamicum* ATCC 13869 genomic DNA was PCR-cloned into pLIC-SGC1 (a gift from Dr. Opher Gileadi, Structural Genomics Consortium and Nuffield Department of Medicine, University of Oxford, UK) using a high-fidelity thermostable DNA polymerase (Deep Vent, from New England Biolabs, Ipswich, Ma, USA) and the oligonucleotide pair 5'TACTTCCAATCCATGGCAGGAGCAGTGG3' (forward) and 5'TATCCACCTTTACTGTTATTTTCGCGTCAGCCTG3' (reverse). These primers included vector-complementary sequences (underlined) for ligation-independent cloning [47] into BseRI-digested pLIC-SGC1, using T4 DNA polymerase (from New England Biolabs). This cloning generates AmtR N-terminally fused to the His₆ tag-hosting sequence MH₆SSGVDLG TENLYFQS (abbreviated here as His₆-AmtR).

For AmtR-His₆ production, PCR amplification was carried out as above with the respective forward and reverse primers, 5'GAATTACCATATGGCAGGAGCAGTGG3' and 5'CTAATCTCTCGAGTTTCGCGTCAGCCTG3', digesting the amplified gene with NdeI and XhoI (respective sites underlined in the primers), and then inserting the gene directionally using T4 ligase into the corresponding sites of pET26 (from Novagen, Madison, WI, USA). AmtR produced in this way has the C-terminal tag LEH₆. The Δ^{2-15} deletion mutant was generated as the wild type protein, except for the utilization in the initial cloning of the forward primer 5'AGGAAGTAGAATTACCATATGGCAGGCAAGAATCC3'.

The final mixtures from the above cloning approaches were transformed into thermocompetent *E. coli* TOP10 cells (from Invitrogen, Carlsbad, CA, USA) by heat shock, isolating the recombinant plasmids from single colonies exhibiting the appropriate antibiotic resistance (either ampicillin for pLIC-SGC1 or kanamycin for pET26) using the Wizard *Plus* SV Minipreps DNA purification system (from Promega, Madison, WI, USA). The isolated plasmids were used to transform electrocompetent *E. coli* BL21(DE3) (Novagen) cells by electroporation [48]. These cells were used in all the expression assays.

Production of recombinant proteins

His₆-AmtR labeled with SeMet was produced by growing the appropriately transformed cells at 37°C in a modified minimal medium in which methionine was replaced by SeMet [49], which was supplemented with 0.1 mg·ml⁻¹ ampicillin. For expression of native (no SeMet labeling) His₆-AmtR and its Δ^{2-15} deletion mutant, and AmtR-His₆, Luria Bertoni (LB) medium supplemented with either 0.1 mg·ml⁻¹ ampicillin (His₆-AmtR and its mutant) or 50 µg·ml⁻¹ kanamycin (AmtR-His₆) was used. In all cases, 1 mM isopropyl β-D-thiogalactopyranoside was added when cell densities reached an OD₆₀₀ of 0.6 -1, and growth was continued for 3-4 h before cell harvesting by centrifugation. All subsequent steps were at 4°C.

The cells from 1-2 L of culture, suspended in 20-40 ml of a solution of 0.1 M HEPES, pH 7.5, 0.5 M NaCl, 1 mM dithiothreitol (DTT), 1 mM phenylmethylsulfonyl fluoride and 1 mM EDTA, were disrupted by sonication. The suspension was clarified by centrifugation and the supernatant was applied to a 1-ml HisTrap column (GE Healthcare, Pittsburgh, PA, USA) fitted in a ÄKTA FPLC (GE Healthcare) that had been equilibrated with 25 mM HEPES, pH 7.5, 0.5 M NaCl, 5 mM MgCl₂, 1 mM DTT and 20 mM imidazole. After

washing with this solution until negligible UV-absorbing material emerged from the column, a 20-ml linear gradient of 20-500 mM imidazole in the same buffer was applied. His₆-tagged AmtR was eluted at 130-170 mM imidazole. Centrifugal ultrafiltration (Amicon 10K from Millipore, Bedford, MA, USA) was used to concentrate the different proteins to 5-10 mg·ml⁻¹ and to place them in a solution of 25 mM HEPES, pH 7.5 and 150 mM (NH₄)₂SO₄. In the case of SeMet-His₆-AmtR this solution contained also 1 mM EDTA and 1 mM DTT .

TtgR from *Pseudomonas putida* was expressed in *E. coli* BL21(DE3) cells transformed with plasmid pET29a(+)-TtgR (kindly provided by Drs. Ana María Fernández-Escamilla and Tino Krell, Estación Experimental del Zaidín, CSIC, Granada), which carries the gene for this protein [50,51]. Expression and purification were as for AmtR-His₆.

When the concentrations of the gene regulators are given in molar terms, values correspond to the molarities of the dimers.

Crystallization and data collection

Crystallization was carried out at the Robotized Crystallization Facility of the Institute of Biomedicine of Valencia (Spain) by vapor diffusion at 21°C in sitting drops of 0.4/0.4 µL of crystallization solution/protein solution using commercial screens (from Qiagen, Molecular Dimensions and Rigaku). The best crystals of SeMet-labeled His₆-AmtR had acicular appearance and a maximal dimension of 0.1 mm. They were obtained in 1 week using 6% PEG 8000, 0.2 M NaCl and 0.1 M NaKPO₄, pH 7, as crystallization solution, and 7 mg·ml⁻¹ AmtR in the final protein suspension buffer (see previous section), supplemented with 1 mM CaCl₂. Crystals of native His₆-AmtR had cubic habit and ~0.1 mm maximum dimension. They grew in two weeks using 0.2 M Li₂SO₄ and 20% PEG 400

as crystallization solution and $6.2 \text{ mg}\cdot\text{ml}^{-1}$ AmtR.

For the crystallization of AmtR-His₆ complexed to DNA, $6 \text{ mg}\cdot\text{ml}^{-1}$ ($120 \text{ }\mu\text{M}$) of AmtR-His₆, in a solution of 25 mM HEPES, pH 7.5, and 150 mM $(\text{NH}_4)_2\text{SO}_4$, was incubated for 10 min at 24°C with $250 \text{ }\mu\text{M}$ (as single strands) of the 26-mer self-hybridizing (except for the GC overhangs at both ends) synthetic (Sigma-Aldrich, Desenhofen, Germany) oligonucleotide 5'ATTATCTATAGATCTATAGATAATGC3'. This mixture was used as the protein-DNA solution in crystallization trials as described for His₆-AmtR. The best crystals, having a rectangular cuboid habit and up to 0.5 mm maximum dimension, were obtained after 2 weeks by using 0.1 M HEPES, pH 6.8, and 1.2 M trisodium citrate as the crystallization solution.

Crystals were flash-cooled with liquid N_2 in the corresponding crystallization solutions supplemented with 25% PEG 400. All the diffraction experiments were carried out at 100 K .

The SAD diffraction experiments with AmtR apo were carried out on the ID29 tunable line of the ESRF Synchrotron (Grenoble, France). Presence of Se was first confirmed by a fluorescence scan, which showed a peak at $0.9791 \text{ }\text{\AA}$. This wavelength was used for diffraction. We collected 2,400 images covering 360° , at a resolution of $2.65 \text{ }\text{\AA}$, using a PILATUS 6M detector. This crystal belonged to the H3 space group (see Table 1), with cell dimensions 160.5 , 160.5 and $52.4 \text{ }\text{\AA}$, and angles of 90° , 90° and 120° . Datasets of native His₆-AmtR were collected at the ESRF ID 23-2 line, at $0.8726 \text{ }\text{\AA}$ wavelength, using a MAR225 CCD detector. These crystals diffracted X-rays at $2.25 \text{ }\text{\AA}$ resolution. They belonged to the P3₂21 space group, with cell dimensions $69.3 \text{ }\text{\AA}$, $69.3 \text{ }\text{\AA}$ and $92.7 \text{ }\text{\AA}$. For X-ray diffraction by the crystal of the AmtR-DNA complex, a wavelength of $0.9795 \text{ }\text{\AA}$ was used on the I-04 line of the Diamond Synchrotron (Oxford, UK), collecting a dataset at $3 \text{ }\text{\AA}$

resolution with a PILATUS 6M detector. The crystals belonged to the $P222_1$ space group, with cell dimensions of 101.8 Å, 137.4 Å and 239.9 Å (Table 1). Data collected with PILATUS and CCD detectors were processed, respectively, with XDS [52] and IMOSFLM [53]. All data were scaled using AIMLESS [54].

Phasing, model building and refinement

Initial single-wavelength (0.9791 Å, corresponding to the peak of the X-ray fluorescence spectrum of selenium) anomalous diffraction (SAD) phases for the SeMet-labelled AmtR apo were obtained with AutoSol [55]. This program found eight Se atoms in the asymmetric unit, and automatically generated an initial model that comprised 335 residues encompassed in 19 different protein segments belonging to two different subunits. This initial solution was improved by iterative cycles of manual model building with COOT [56] followed by restrained refinement using REFMAC [57]. Non-crystallographic symmetry restraints and B-factor restraints were used and gradually released as refinement progressed. All the diffraction data were used throughout the refinement process, except for the 5% randomly selected data used to calculate R_{free} . For the SeMet AmtR structure TLS was used in the last refinement step [58], choosing the TLS groups with the TLSMD server (<http://skuld.bmsc.washington.edu/~tlsmd/>, [59]). Table 1 summarizes the data on the refinement process. The final model at 2.65 Å resolution contained one AmtR dimer in the asymmetric unit.

For phasing of the His₆-AmtR dataset molecular replacement with Phaser [60] was performed, using as search model one subunit of SeMet-AmtR. Model building and refinement were as for SeMet-AmtR. The final model at 2.65 Å resolution contained a single AmtR subunit in the asymmetric unit. This model lacked residues 1-18, 86-89 and

221-222. The same molecular replacement approach was used for phasing of the DNA-containing crystal, although in this case the search models were the subunit of the AmtR apo and a single strand of DNA from the QacR-DNA structure (PDB accession code: 1JT0, [21]). The initial solution consisted of two AmtR dimers, one double-stranded DNA molecule and enough electron density for a second DNA molecule. Model building and refinement was as for AmtR apo. The final model at 3 Å resolution (Table 1) included two AmtR dimers, each bound to a double-stranded DNA molecule. Solvent content was 82%.

Stereochemical analysis with RAMPAGE [61] revealed excellent values for all the models, with no outliers in the Ramachandran plot. The buried surface areas and contacts between subunits were calculated with a probe of 1.4 Å radius using PISA [28]. Subdomain movements were analyzed with DynDom [33]. Figures of the protein structures were generated using PyMOL (<http://www.pymol.org>). Geometry characterization of helices was carried out with HELANAL Plus [30]. DNA structure parameters were calculated with 3D-DART [31] and 3DNA [32].

Electrophoretic mobility shift assays (EMSA)

EMSA was carried out with a PCR-generated 171-bp DNA fragment using *C. glutamicum* ATCC13869 genomic DNA as template, Deep-Vent polymerase, and the oligonucleotide primers 5'GTTGTCTGGCAGGTGGTTTTC3' and 5'CCAGCTCTAACTATAGACCTACAGAACTAATTTC3'. The amplified fragment, that encompassed bases -258 to -88 upstream of the translation initiation point for the *amtB* gene of *C. glutamicum*, was purified using a commercial column kit (GeneClean kit, from MP Biomedicals, Solon, OH, USA), γ - ^{32}P end-labeled [62] with $[\gamma$ - $^{32}\text{P}]$ ATP (3000 Ci·mmol $^{-1}$, from Perkin Elmer, Wellesley, MA, USA) and separated from the free

nucleotide by gel filtration in TE buffer. This labeled fragment, at 0.27 nM concentration (as a double helix), was incubated for 30 min at 24 °C with the indicated concentrations of His₆-AmtR (as a dimer), either the wild type form or the form carrying the Δ^{2-15} deletion, in 50 mM HEPES, pH 7.5, 3 mM MgCl₂, 5% glycerol, 1 mM DTT, 0.025 mg·ml⁻¹ bovine serum albumin and 0.5 mg·ml⁻¹ salmon sperm DNA. Then equal volumes of each incubation solution and of a solution of 62.5 mM Tris-HCl, pH 7.5, 10% glycerol and 0.02% bromphenol blue, were mixed and applied to a 6% polyacrylamide gel which was prepared and run in 0.5 × TBE. Electrophoresis was run at 100 V and 4 °C for 90 min. The gel was then dried for 1 h by heating under vacuum using a Bio-Rad 583 gel drying system fitted with a Hydrotech pump, and was then revealed overnight to an imaging plate and scanned with a Fujifilm FLA-5000 imaging system.

Other assays

Size exclusion chromatography was carried out on a Superdex 200 column (10/300 Increase column from GE Healthcare) mounted on an ÄKTA FPLC (GE Healthcare) and ran at 24°C and at a 0.75 ml·min⁻¹ flow rate using HEPES 25 mM pH 7.5 and NaCl 0.2 M as the running solution, monitoring the absorbance at 280 nm. Protein standards of known molecular mass were from Amersham Biosciences or Sigma.

To monitor ANS binding [63], 0.5 µM AmtR-His₆ or TtgR-His₆ was mixed with 50 µM ANS (from Sigma-Aldrich) in a solution of TrisCl 50 mM pH 7.5, NaCl 100 mM, determining the fluorescence emission spectrum in a Perkin Elmer LS 45 Fluorescence Spectrometer using an excitation wavelength of 340 nm.

Protein concentrations were determined with Coomassie blue [64] using a commercial reagent (from Bio-Rad, Hercules, CA, USA) and bovine serum albumin as a

standard.

ACKNOWLEDGEMENTS

This work was supported by grants from the Valencian (PrometeoII/2014/029) and Spanish Governments (BFU2011-30407 and BFU2014-58229-P). C.P. is a JAE-Predoc fellow of the CSIC. We thank Dr. Jose A. Gil for providing the genomic DNA of *Corynebacterium glutamicum* ATCC 13869 used for cloning, Drs. Tino Krell and Ana María Fernández Escamilla for providing expression plasmids of TtgR from *P. putida*, and Ismael Mingarro (University of Valencia) for access to the spectrofluorometer. We are also grateful to Diamond Light Source for access to beamline I-04 (mx8035) and to the ESRF for use of beamlines ID29 and ID23-2, which contributed to the results presented here. ALBA synchrotron's beamline XALOC was used to screen and improve initial crystals and has to be also thanked. We also thank the collaboration of the members of the staffs of these three synchrotrons. The research that has led to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under BioStruct-X (grant agreement No. 283570), within proposal 7687. We thank Ignacio Fita (IMB-CSIC, Barcelona) for critically reading the revised version of this manuscript.

AUTHOR CONTRIBUTIONS

C.P. performed the experiments. C.P. and V.R. planned the experiments, analyzed the data and wrote the paper.

REFERENCES

1. Burkovski A (2015) Trends in *Corynebacterium glutamicum*. Research and application. In *Corynebacterium glutamicum: from systems biology to biotechnological applications* (Burkovski A, ed.), pp 1–10. Caister Academic press, Norfolk, UK.
2. Hermann T (2003) Industrial production of amino acids by coryneform bacteria. *J Biotechnol* **104**, 155–172.
3. Ajinomoto Co., Inc. (2013) Feed-use amino acids business.
<http://www.ajinomoto.com/en/ir/pdf/Feed-useAA-Oct2013.pdf>.
4. Ajinomoto Co., Inc. (2013b) First quarter-FY2013 market and other information.
http://www.ajinomoto.com/en/ir/pdf/FY13_Data_E.pdf.
5. Rehm N & Burkovski A (2011) Engineering of nitrogen metabolism and its regulation in *Corynebacterium glutamicum*: influence on amino acid pools and production. *Appl Microbiol Biotechnol* **89**, 239–248.
6. Rehm N, Bürger J & Burkovski A (2015) Manipulation of nitrogen metabolism and alternative nitrogen sources for *Corynebacterium glutamicum*. In (Burkovski A, ed.) *Corynebacterium glutamicum: from systems biology to biotechnological applications*, pp 71-82. Caister Academic press, Norfolk, UK.
7. Jakoby M, Nolden L, Meier-Wagner J, Krämer R & Burkovski A (2000) AmtR, a global repressor in the nitrogen regulation system of *Corynebacterium glutamicum*. *Mol Microbiol* **37**, 964–977.
8. Burkovski A (2007) Nitrogen control in *Corynebacterium glutamicum*: proteins, mechanisms, signals. *J Microbiol Biotechnol* **17**, 187–194.

9. Yu Z, Reichheld SE, Savchenko A, Parkinson J & Davidson AR (2010) A comprehensive analysis of structural and sequence conservation in the TetR family transcriptional regulators. *J Mol Biol* **400**, 847–864.
10. Beckers G, Strösser J, Hildebrandt U, Kalinowski J, Farwick M, Krämer R & Burkovski A (2005) Regulation of AmtR-controlled gene expression in *Corynebacterium glutamicum*: mechanism and characterization of the AmtR regulon. *Mol Microbiol* **58**, 580–595.
11. Rehm N, Georgi T, Hiery E, Degner U, Schmiedl A, Burkovski A & Bott M (2010) L-Glutamine as a nitrogen source for *Corynebacterium glutamicum*: derepression of the AmtR regulon and implications for nitrogen sensing. *Microbiology* **156**, 3180–3193.
12. Hasselt K, Rankl S, Worsch S & Burkovski A (2011) Adaptation of AmtR-controlled gene expression by modulation of AmtR binding activity in *Corynebacterium glutamicum*. *J Biotechnol* **154**, 156–162.
13. Leigh JA & Dodsworth JA (2007) Nitrogen regulation in bacteria and archaea. *Annu Rev Microbiol* **61**, 349–377.
14. Forchhammer K (2008) P_{II} signal transducers: novel functional and structural insights. *Trends Microbiol* **16**, 65–72.
15. Radchenko M & Merrick M (2011) The role of effector molecules in signal transduction by PII proteins. *Biochem Soc Trans* **39**, 189–194.
16. Maier S, Schleberger P, Lü W, Wacker T, Pflüger T, Litz, C & Andrade SLA (2011) Mechanism of disruption of the Amt-GlnK complex by PII-mediated sensing of 2-oxoglutarate. *PLoS ONE* **6**, e26327.

17. Zeth K, Fokina O & Forchhammer K (2014) Structural basis and target-specific modulation of ADP sensing by the *Synechococcus elongatus* PII signaling protein. *J Biol Chem* **289**, 8960–8972.
18. Strösser J, Lüdke A, Schaffer S, Krämer R & Burkovski A (2004) Regulation of GlnK activity: modification, membrane sequestration and proteolysis as regulatory principles in the network of nitrogen control in *Corynebacterium glutamicum*. *Mol Microbiol* **54**, 132–147.
19. Forchhammer K & Lüddecke J (2015) Sensory properties of the P_{II} signalling protein family. *FEBS J* *in press*.
20. Orth P, Schnappinger D, Hillen W, Saenger W & Hinrichs W (2000) Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system. *Nat Struct Biol* **7**, 215–219.
21. Schumacher MA, Miller MC, Grkovic S, Brown MH, Skurray RA & Brennan, RG (2002) Structural basis for cooperative DNA binding by two dimers of the multidrug-binding protein QacR. *EMBO J* **21**, 1210–1218.
22. Miller DJ, Zhang YM, Subramanian C, Rock CO & White SW (2010) Structural basis for the transcriptional regulation of membrane lipid homeostasis. *Nat Struct Mol Biol* **17**, 971–975.
23. Itou H, Watanabe N, Yao M, Shirakihara Y & Tanaka I (2010) Crystal structures of the multidrug binding repressor *Corynebacterium glutamicum* CgmR in complex with inducers and with an operator. *J Mol Biol* **403**, 174–184.
24. Le TBK, Schumacher MA, Lawson DM, Brennan RG & Buttner MJ (2011) The crystal structure of the TetR family transcriptional repressor SimR bound to DNA

- and the role of a flexible N-terminal extension in minor groove binding. *Nucleic Acids Res* **39**, 9433–9447.
25. Sawai H, Yamanaka M, Sugimoto H, Shiro Y & Aono S (2012) Structural basis for the transcriptional regulation of heme homeostasis in *Lactococcus lactis*. *J Biol Chem* **287**, 30755–30768.
 26. Kim Y, Kang Y & Choe J (2013) Crystal structure of *Pseudomonas aeruginosa* transcriptional regulator PA2196 bound to its operator DNA. *Biochem Biophys Res Commun* **440**, 317–321.
 27. Tonthat NK, Milam SL, Chinnam N, Whitfill T, Margolin W & Schumacher MA (2013) SlmA forms a higher-order structure on DNA that inhibits cytokinetic Z-ring formation over the nucleoid. *Proc Natl Acad Sci USA* **110**, 10586–10591.
 28. Krissinel E & Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774–797.
 29. Nguyen Le Minh P, de Cima S, Bervoets I, Maes D, Rubio V & Charlier D (2015) Ligand binding specificity of RutR, a member of the TetR family of transcription regulators in *Escherichia coli*. *FEBS Open Bio* **5**, 76–84.
 30. Kumar P & Bansal M (2012) HELANAL-Plus: a web server for analysis of helix geometry in protein structures. *J Biomol Struct Dyn* **30**, 773–783.
 31. van Dijk M & Bonvin AMJJ (2009) 3D-DART: a DNA structure modelling server. *Nucleic Acids Res* **37**, W235–9.
 32. Zheng G, Lu XJ & Olson WK (2009) Web 3DNA—a web server for the analysis, reconstruction, and visualization of three-dimensional nucleic-acid structures, *Nucleic Acids Res* **37**, W240–W246

33. Schultz SC, Shields GC & Steitz TA (1991) Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees. *Science* **30**, 1001–1007.
34. Hayward S & Lee RA (2002) Improvements in the analysis of domain motions in proteins from conformational change: DynDom version 1.50. *J Mol Graph Model* **21**, 181–183.
35. Yang ZR, Thomson R, McNeil P & Esnouf RM (2005) RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. *Bioinformatics* **21**, 3369–3376.
36. Hasselt K, Sevvana M, Burkovski A & Muller YA (2009) Crystallization and preliminary crystallographic analysis of the global nitrogen regulator AmtR from *Corynebacterium glutamicum*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **65**, 1123–1127.
37. Muhl D, Jessberger N, Hasselt K, Jardin C, Sticht H & Burkovski A (2009) DNA binding by *Corynebacterium glutamicum* TetR-type transcription regulator AmtR. *BMC Mol Biol* **10**, 73.
38. Nolden L, Ngouoto-Nkili CE, Bendt AK, Krämer R & Burkovski A. (2010) Sensing nitrogen limitation in *Corynebacterium glutamicum*: the role of *glnK* and *glnD*. *Mol Microbiol* **42**, 1281–95.
39. Gruswitz F, O'Connell J & Stroud RM (2007) Inhibitory complex of the transmembrane ammonia channel, AmtB, and the cytosolic regulatory protein, GlnK, at 1.96 Å. *Proc Natl Acad Sci USA* **104**, 42–47
40. Conroy MJ, Durand A, Lupo D, Li XD, Bullough PA, Winkler FK & Merrick M (2007) The crystal structure of the *Escherichia coli* AmtB-GlnK complex reveals

- how GlnK regulates the ammonia channel. *Proc Natl Acad Sci USA* **104**, 1213–1218.
41. Llacer JL, Contreras A, Forchhammer K, Marco-Marn C, Gil-Ortiz F, Maldonado R, Fita L & Rubio V. (2007). The crystal structure of the complex of PII and acetylglutamate kinase reveals how PII controls the storage of nitrogen as arginine. *Proc Natl Acad Sci USA* **104**, 17644–17649.
 42. Mizuno Y, Moorhead GB & Ng KK (2007) Structural basis for the regulation of N-acetylglutamate kinase by PII in *Arabidopsis thaliana*. *J Biol Chem* **282**, 35733–40.
 43. Llacer JL, Espinosa J, Castells MA, Contreras A, Forchhammer K & Rubio V (2010) Structural basis for the regulation of NtcA-dependent transcription by proteins PipX and PII. *Proc Natl Acad Sci USA* **107**, 15397–15402.
 44. Zhao MX, Jiang YL, Xu BY, Chen Y, Zhang CC & Zhou CZ (2010) Crystal structure of the cyanobacterial signal transduction protein PII in complex with PipX. *J Mol Biol* **402**, 552–529.
 45. Christen S, Srinivas A, Bhler P, Zeller A, Pridmore D Bieniossek C, Baumann U & Erni B (2006). Regulation of the Dha operon of *Lactococcus lactis*: a deviation from the rule followed by the Tetr family of transcription regulators. *J Mol Biol* **281**, 23129–23137.
 46. Tonthat NK, Arold ST, Pickering BF, Van Dyke MW, Liang S, Lu Y, Beuria TK, Margolin W & Schumacher MA (2011) Molecular mechanism by which the nucleoid occlusion factor, SlmA, keeps cytokinesis in check. *EMBO J* **30**, 154–164.
 47. Gileadi O, Burgess-Brown NA, Colebrook SM, Berridge G, Savitsky P, Smee CE, Loppnau P, Johansson C, Salah E & Pantic NH (2008) High throughput production

- of recombinant human proteins for crystallography. *Methods Mol Biol* **426**, 221–246.
48. Green RM & Sambrook J (2012) Transformation of *E. coli* by electroporation. In *Molecular cloning. A laboratory manual 4th Edition*, pp. 177-182. Cold Spring Harbor, New York, NY
49. Doublie S (2007). Production of selenomethionyl proteins in prokaryotic and eukaryotic expression systems. *Methods Mol Biol* **363**, 91–108.
50. Espinosa-Urgel M, Serrano L, Ramos JL & Fernández-Escamilla AM (2015) Engineering biological approaches for detection of toxic compounds: a new microbial biosensor based on the *Pseudomonas putida* TtgR repressor. *Mol Biotechnol* **57**, 558–564.
51. Fernández-Escamilla AM, Fernandez-Ballester G, Morel B, Casares-Atienza S & Ramos JL (2015) Molecular binding mechanism of TtgR repressor to antibiotics and antimicrobials. *PLoS ONE* **10**, e0138469
52. Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* **66**, 125–132.
53. Battye TGG, Kontogiannis L, Johnson O, Powell HR & Leslie, AGW (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* **67**, 271–281.
54. Evans PR & Murshudov, GN (2013) How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* **69**, 1204–1214.
55. Terwilliger TC, Adams PD, Read RJ, McCoy AJ, Moriarty NW, Grosse-Kunstleve RW, Afonine PV, Zwart PH, & Hung LW (2009) Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr D Biol Crystallogr* **65**, 582–601.

56. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132.
57. Murshudov GN, Vagin AA & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240–255.
58. Winn MD, Murshudov GN & Papiz MZ (2003) Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol* **374**, 300–321.
59. Painter J & Merritt EA (2006) TLSMD web server for the generation of multi-group TLS models. *J Appl Crystallogr* **39**, 109–111.
60. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674.
61. Lovell SC, Davis IW, Arendall WB, de Bakker PIW, Word JM, Prisant MG, Richardson JS & Richardson DC (2003) Structure validation by C _{α} geometry: ϕ , ψ and C _{β} deviation. *Proteins* **50**, 437–450.
62. Tabor S (2001) Phosphatases and kinases. In *Current Protocols in Molecular Biology*. Greene Publishing Associates and Wiley Interscience, John Wiley and Sons, NY.
63. Tahlan K, Yu Z, Xu Y, Davidson AR & Nodwell JR (2008) Ligand recognition by ActR, a TetR-like tegulator of sctinorhodin export. *J Mol Biol* **383**, 753–761.
64. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248–254.
65. Baker NA, Sept D, Joseph S, Holst MJ & McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci*

USA 98, 10037–10041.

Table 1. X-Ray data collection and refinement statistics of AmtR crystals

Crystal PDB file identifier	Apo-SeMet 5DY1	Apo-native 5DXZ	DNA-complexed 5DY0
<i>Data collection^a</i>			
Synchrotron	ESRF	ESRF	DLS
Beamline	ID29	ID23-2	I-04
Wavelength (Å)	0.9791	0.8726	0.9795
Space group	H3	P3 ₂ 21	P222 ₁
Unit cell parameters			
a, b, c (Å)	160.46, 160.46, 52.44	69.31, 69.31, 92.72	101.73, 137.42, 239.87
α , β , γ , (°)	90.00, 90.00, 120.00	90.00, 90.00, 120.00	90.00, 90.00, 90.00
Solvent content (%)	49.1	48.6	82.0
Matthews coefficient (Å ³ /Da)	2.42	2.39	6.16
Resolution range (Å)	49.06 - 2.65	36.69 - 2.25	119.94 - 3.00
Reflections, total/unique	47,544/14,116 (4,134/1,790)	139,352/12,703 (12,890/1,158)	295,546/67,304 (20,906/4,525)
Completeness (%)	96.4 (82.8)	100 (100)	99.0 (99.8)
Multiplicity	3.4 (2.3)	11.0 (11.1)	4.4 (4.6)
I/ σ _I	10.4 (2.1)	22.5 (6.5)	17.1 (2.1)
CC _{1/2} (%)	99.5 (70.3)	99.9 (95.8)	99.9 (77.4)
Wilson B factor (Å ²)	41.20	38.76	75.80
R _{pim} ^b (%)	6.1 (30.1)	2.9 (18.3)	2.9 (35.8)
Phasing statistics for Se-Met substitution			
Se atoms	8	-	-
RCR _{anom} ^c	1.336	-	-
FOM ^d	0.27	-	-
<i>Refinement</i>			
Resolution	49.06 - 2.65	36.69 - 2.25	119.94 - 3.00
Reflections, work/test	13,410/706	12,068/635	63939/3365
R _{factor} ^e /R _{free} ^f (%)	20.9/24.2	22.0/24.9	20.6/22.4
rmsd from ideal			
Bond length (Å)	0.005	0.014	0.009
Bond angle (°)	1.00	1.50	1.38
Number of			
Polypeptide chains	2	1	4
Protein atoms	3,120	1,531	6,600
DNA molecules	0	0	4
Solvent molecules	12	35	0
Average B-factor (Å ²)			
Protein	52.5	44.5	94.0
DNA	-	-	96.0
H ₂ O	30.5	42.8	-

Ramachandran plot^g (%)

Favoured	97.5	99.0	96.8
Allowed	2.5	1.0	3.2
Outliers	0	0	0

^aValues in parentheses are the data for the highest resolution shell.

^b $R_{pim} = \sum_{hkl} [1/(N-1)]^{1/2} \sum_i |I_{hkl,i} - \langle I_{hkl} \rangle| / \sum_{hkl} \langle I_{hkl} \rangle$

^cCalculated by Aimless [54]

^dFigure of merit after density modification

^e $R_{factor} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$ where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively.

^f R_{free} is the same as R_{factor} , but was calculated for the 5% data omitted from the refinement.

^gCalculated using RAMPAGE [61]

FIGURE LEGENDS

Figure 1. AmtR apo and DNA complexes. Cartoon representations of: (A) the apo-SeMet AmtR dimer; (B) helix 4 in the specified crystal structures, indicating the angle defined in each structure by the kink of the helix; (C) the apo-native dimer generated by applying the crystal symmetry to the subunit found in the asymmetric unit (colored), marking the two domains and the different secondary structure elements and highlighting the $\alpha 8$ - $\alpha 9$ loop in magenta; (D) the long DNA double helical chains formed in the DNA-AmtR crystal, showing the regular bridging of different DNA double helices by AmtR dimers of dimers. The box limited by a continuous line isolates the contents of the asymmetric unit. The broken line box marks the interdimeric junction, which is shown in (E) expanded and with the final part of helix 9 and the initial part of helix 10 in cylinder representation. (F) Size exclusion chromatography of AmtR-His₆ and of its complex with DNA, which at the top shows the semilogarithmic plot of mass versus the elution volume for protein standards (filled circles; thyroglobulin, 669 kDa; ferritin, 443 kDa; β -amylase, 200 kDa; aldolase, 158 kDa; alcohol dehydrogenase tetramer, 147 kDa; ovalbumin, 44 kDa; maltose-binding protein, 42 kDa; and carbonic anhydrase, 29 kDa). The position of the AmtR peak is indicated with an open blue circle, positioned for the sequence-deduced mass of ~52 kDa for the dimer. The colored lines project the elution positions of DNA, either alone or mixed with AmtR-His₆, to show the increase in the mass of the DNA peak as expected for the binding of one AmtR dimer. At the bottom, elution profiles are monitored as optical absorption at 280 nm for AmtR alone (blue; plotted on the scale of the right-hand y-axis), for DNA alone (red) and for the DNA-AmtR mixture (magenta-colored broken line). DNA and AmtR were used in amounts of 2.5 nmol each (molarity of the double helix or the protein dimer), in an injection volume ≤ 0.1 ml.

Figure 2. Structures of the DNB and LBD domains and presence of the $\alpha 8$ - $\alpha 9$ loop and of helix 10 in AmtR from different corynebacteria. (A) Stereo view of the superposition of the backbone (in straw representation) of the core of the DNB domains of AmtR (red; residues 20-62), QacR (green; residues 3-45; from PDB accession code, 1JT0), TetR (gray; residues 5-47; PDB 1QPI), CgmR (magenta; residues 5-47; PDB 2YVH) and RutR (blue; residues 19-61; PDB 4X1E). (B) Stereo view of the superposition of the cartoon representations of the LBD domains of AmtR (red) and QacR (green, PDB 1JT0) (C) Alignment of the $\alpha 8$ - $\alpha 9$ loop and of the $\alpha 9$ - $\alpha 10$ turn and helix 10 of *C. glutamicum* ATCC 13869 (structural elements schematized above the sequences), with the corresponding sequences of the putative AmtR proteins of other corynebacteria. Identities and conservative replacements are shown with black or gray background, respectively, when the prevalence of identical or conserved residues is ≥ 8 among the aligned sequences. When not highlighted in this way, proline residues are shown red on a yellow background. Dots indicate non-illustrated amino acids. The dots at the end indicate the continuation of the sequence to the C-terminal residue (numbering at the end of each sequence). The sequence of the T-loop of GlnK from *C. glutamicum* (Uniprot KB file **Uniprot Protein Database ID: H7C694**) is aligned at the top, with identical or conserved residues with the $\alpha 8$ - $\alpha 9$ loop of *C. glutamicum* AmtR ATCC 13869 enclosed in squares. The sequences of AmtR proteins, from second to bottom, correspond to the following sources (species / Uniprot KB entry) *C. glutamicum* strain 13032 / **Uniprot Protein Database ID: Q79VH8**, *C. efficiens* / **Uniprot Protein Database ID: Q8FR27**, *C. callunae* / **Uniprot Protein Database ID: M1UY13**, *C. diphtheriae* / **Uniprot Protein Database ID: A0A0D6FQN8**, *C. diphtheriae* strain ATCC 700971 / **Uniprot Protein Database ID: Q6NID1**, *C. ulcerans* / **Uniprot Protein**

Database ID: G0CY51, *C. accolens* / Uniprot Protein Database ID: E0MWK4, *C. pseudotuberculosis* strain 1002 / Uniprot Protein Database ID: D9Q7A0, *C. striatum* / Uniprot Protein Database ID: C2CM41, *C. aurimucosum* / Uniprot Protein Database ID: C3PF10, *C. camporealensis* / Uniprot Protein Database ID: A0A0F6TAZ3, *C. singulare* / Uniprot Protein Database ID: A0A0B6EPL5, *C. marinum* / Uniprot Protein Database ID: A0A0B6TK39, *C. uterequi* / Uniprot Protein Database ID: A0A0G3HFI8, *C. testudinoris* / Uniprot Protein Database ID: A0A0G3H639, *C. kutscheri* / Uniprot Protein Database ID: A0A0F6TEA6, and *C. pseudotuberculosis* strain C231 / Uniprot Protein Database ID: D9Q9A8.

Figure 3. Polar cavities and the $\alpha 8$ - $\alpha 9$ loop. (A) Surface representations of the AmtR dimer in both the native-apo and DNA-bound forms, illustrating the existence of a large cavity on the intersubunit boundary, and of two smaller cavities defined by the $\alpha 8$ - $\alpha 9$ loop (schematized in black as the protein backbone). Electrostatic positive and negative surface potentials, calculated with APBS [65], are colored blue or red as indicated in the figure. The scale gives the range of values for the dimensionless coefficient given by APBS to calculate the surface potential. (B) The fluorescence emission spectrum of mixtures of 0.5 μ M TtgR or AmtR (as dimers) and 50 μ M ANS. Values have been corrected for the very small fluorescence of 50 μ M ANS in the absence of protein. For further details, see Materials and methods. (C) Superposition (cartoon representation) of the LBD domain of Apo-native AmtR (in red) on the dimer of the same domain of RutR ([29]; PDB file 4X1E), with one subunit in blue and the other in gray, showing how helix 8* of RutR (in lighter blue) is projected on the other subunit, whereas the corresponding $\alpha 8$ - $\alpha 9$ loop of AmtR (in pink) recedes over its own subunit. (D) Detail of the superposition in (C) showing the

subunit of RutR (colored gray) in surface representation, to illustrate the cavity obliterated in RutR by helix 8* from the adjacent subunit, but which is fully exposed in AmtR.

Figure 4. AmtR dimer binding to the DNA double helix. (A) The DNA is shown in backbone representation, with each chain in a different color, marking the C3' atom of each ribose with a sphere, and identifying by name the DNA bases, numbering them relative to the AmtR box center (upstream and downstream bases are given increasing negative and positive values, respectively; see the correspondence with the 5'-to-3' numbering of the oligonucleotide on the X-axis of panel D). The green line corresponds to the central axis of the double helix, illustrating the very slight DNA bending. AmtR is shown in cartoon representation with each subunit in a different red hue. (B) The DNA is shown in spheres representation, with one chain colored white and the other colored yellow, although P and O atoms are respectively orange and red in both chains. The AmtR dimer is shown in surface representation, with one subunit in gray while the other one is transparent and blue excepting the N-terminal extension, which is colored magenta, whereas alpha helices are seen across the surface as cylinders. The N-terminal extension elements, some helices and domains DNB and LBD (subunit to the left) are labeled. (C) DNA double helix backbone and helices 3 and 3' of the AmtR dimer in the DNA-bound form (blue) or superimposed from the apo forms, either native (green) or SeMet (orange). The helices are represented as semitransparent cylinders. The distances between the C α atoms of Tyr42 and Tyr42' (from the 3 and 3' helices of each dimer) are shown in the color of the corresponding dimer, with a broken line of the same color connecting these C α atoms. (D) Groove widths in ideal B-DNA (dotted black traces, as indicated) or measured manually (mean $\pm$ SD for the two DNA double helices in the asymmetric unit) at each base pair step of the AmtR-DNA

crystal (colored traces) are shown. The position of the N-terminal extension anchor is indicated.

Figure 5. Movements of the DNB domain and of helix 4, and lack of movement of other AmtR molecule elements, upon DNA binding. This figure depicts stereoviews of superpositions of the indicated protein regions, shown in backbone representations. (A) A subunit of the Apo-SeMet form (black) and of the AmtR-DNA complex (red). In this case only the LBD domains (residues 64-220) have been superposed, to highlight the movement of the DNB domain. The hinge (residues 60-63, marked in green) and the axis (gray line) around which the DNB domain rotates as described in the text, are shown, and the direction of the rotation is indicated by arrows. The broken arrow indicates the movement of helix 3. (B) The movement of both the DNB domain and helix 4 when all other elements (not shown) of subunits B of Apo-SeMet and AmtR-DNA are superposed to highlight the movement of helix 4. The blue dotted line marks the upstream movement of the helix (shown for the $C\alpha$ of Thr79). (C) Representation of the part of the dimer formed by both LBD domains, not including helix 4, for the structures of Apo-native (green), Apo-SeMet (orange) and both AmtR-DNA dimers (blue and red).

Figure 6. Major and minor groove interactions of AmtR in the AmtR-DNA crystal.

Throughout the figure, red broken lines illustrate hydrogen bonds between protein atoms or between the protein and the DNA. Oxygen, nitrogen and phosphorus atoms are colored red, blue and orange, respectively except in the DNA bases of panel F, where N and O atoms are not differentially colored. (A) Stereo representation of the insertion of $\alpha 3$ in the major groove of DNA. The DNA and the protein side-chains are shown in sticks representation,

and carbon atoms are colored yellow (DNA), green (helix 3 and part of the $\alpha 3$ - $\alpha 4$ connector), magenta (N-terminal extension residues) or dark blue (Thr42). (B) Estimated probability of disorder along the AmtR polypeptide sequence, as predicted by RONN [35]. The boxed residues are those of the N-terminal extension. (C) The C-terminal extensions of subunits A (in blue) and B (in red) bound to the minor grooves of DNA to which they are anchored (subunit A to its cognate DNA; subunit B to a DNA outside its asymmetric unit), with residues 15-20 superposed to show the divergence of their N-terminal halves. Protein chains are shown in backbone representation, while DNA is shown as cartoon and colored orange. (D) The same as (C) except for the residues that are superposed, which in this case are residues 4-11 to highlight the fact that the N-terminal end is conformationally identical in both subunits. (E) Stereo view of the interactions of the N-terminal extension with the minor groove of DNA. In the sticks representation of DNA and protein, the C atoms of DNA are colored white or yellow depending on the chain, those of the N-terminal extension are colored magenta, and those of other protein parts are in blue. (F) Overview of all the interactions of an AmtR subunit (subunit A of the AmtR-DNA crystal) with DNA (in spheres representation, with the C atoms of each chain in yellow or white). The N-terminal extension is in backbone representation and is colored in magenta, with some side chains shown in sticks or spheres representation. The DNB domain core is shown in cartoon representation, with helices as cylinders and in blue, and with some side chains shown in sticks representation. (G) Electrophoretic mobility shift assays of AmtR-H₆, either wild-type or with residues 2-15 deleted, with its operator DNA. For further details see Materials and Methods.

Fig. 1

Fig. 2

Fig. 3

Fig. 4

Fig. 5

Fig. 6