

## **Disorder Can Exist inside Well-diffracting Crystals**

*Elodie Touzé, Bernard Lorber, Marzanna Deniziak, Hubert Dominique Becker, Daniel Kern, Richard Giegé and Claude Sauter\**

*Architecture et réactivité de l'ARN, Université Louis Pasteur, CNRS, IBMC, 15 rue René Descartes, F-67084 Strasbourg-Cedex, France*

\*To whom correspondence should be addressed: [c.sauter@ibmc.u-strasbg.fr](mailto:c.sauter@ibmc.u-strasbg.fr)

**ABSTRACT:** Unlike other glutaminyl-tRNA synthetases, the one from radioresistant bacterium *Deinococcus radiodurans* (*Dr*-GlnRS) possesses an additional C-terminal extension of 220 residues that shares some homology with the subunit of another enzyme of the translation machinery. *Dr*-GlnRS has been crystallized in an orthorhombic space group. The crystals diffract x-rays to a resolution of  $\sim 2$  Å. The determination of the structure of this atypical GlnRS showed that its N- and C-terminal appendices – that encompass in total one third of the protein's 852 amino acids - are actually disordered in the crystal lattice. This example demonstrates that macromolecule crystallization can tolerate large flexible regions in the solvent channels as long as they don't interfere with the packing contacts. This intriguing case is analyzed and discussed in the light of current crystallogensis strategies.

**KEYWORDS:** crystal packing, disorder, crystal quality, dynamic light scattering, aminoacyl-tRNA synthetase, glutaminyl-tRNA synthetase, *Deinococcus radiodurans*

## 1. Introduction

During the last decade, the massive effort put into genome sequencing and the boom of structural genomics have dramatically changed our way of performing structural investigations.<sup>1</sup> Looking at sequence databases often leads to the identification of intriguing protein variants that are appealing targets for structural biology investigation. We recently got interested in an atypical form of glutaminyl-tRNA synthetase from the radio-resistant bacterium *Deinococcus radiodurans* (*Dr-GlnRS*). Unlike other members of the GlnRS family, it possesses an additional C-terminal extension of 220 residues that shares some homology with another enzyme of the translation machinery. In order to insight into the function of this extra appendix, we crystallized *Dr-GlnRS* and solved its structure. Although the crystals did diffract to a resolution of  $\sim 2$  Å, the C-terminus as well as the N-terminal tail and two internal loops could not be observed in the electron density map. We discuss here an extreme case of disorder (35% of the protein) in well-diffracting crystals.

## 2. Experimental section

The GlnRS from *D. radiodurans* was produced in *E. coli* and purified as described.<sup>2</sup> The homogeneity of the pure protein was analyzed at 20°C by dynamic light scattering using a Dynapro DP801 instrument (Protein Solution, Inc.). Samples at 0.5 to 1.7 mg/ml in 50 mM Tris-HCl pH 7.5, 0.5 mM EDTA, 5 mM DTE were very monodisperse. The diffusion coefficient (extrapolated at zero concentration) was  $D_0 \sim 4.3 \cdot 10^{-7}$  cm<sup>2</sup>/s, corresponding to a particle with a hydrodynamic radius  $R_h \sim 5.0 \pm 0.8$  nm and a derived molecular weight of 143 kDa. The actual molecular weight is 93 kDa, but the overestimation by DLS is due to the elongated shape of the enzyme and the presence of flexible extensions.

Orthorhombic crystals of *Dr-GlnRS* were grown in microbatch<sup>3</sup> in the presence of 8–10% (w/v) PEG 3350.<sup>2</sup> Prior to their analysis by SDS-PAGE or mass spectrometry, they were carefully washed 4 times in 10% (w/v) PEG 3350 to exclude any contamination by the mother liquor, then dissolved in a volatile buffer (100 mM ammonium acetate pH 4.6).

The crystals were analyzed at the European Synchrotron Radiation Facility (ESRF, Grenoble, France). Their main characteristics are given in Table 1. The structure was solved by molecular replacement using the CaspR webservice.<sup>4</sup> Several cycles of model building and refinement smoothly led to a final R-factor and free R-factor of 20.0 and 24.2 %, respectively.<sup>5</sup> No electron density could be found for ~35% of the sequence of the protein, even in low resolution or composite maps. A model of the full-length *Dr*-GlnRS shown in Figure 1A was built based on the structure of the *Dr*-GlnRS core (PDB id: 2HZ7; ref 5). It was manually accommodated in the large solvent channels present in the crystal packing as illustrated in Figure 2. The figures were prepared with PyMol, DeLano Scientific LLC (<http://www.delanoscientific.com>).

### 3. Results and Discussion

**3.1 Design and characterization of the target.** Sequence analysis of *D. radiodurans* genome<sup>6</sup> revealed that the encoded GlnRS consists of the classical polypeptide fold, as found in the *Escherichia coli* GlnRS,<sup>7</sup> and of a 220 residue-long C-terminal extension that is specific to *Dr*-GlnRS (Figure 1). The latter extension presents a region predicted to adopt the same fold as the C-terminal domain of the GatB subunit of tRNA dependent amidotransferase (AdT). AdT is actually an enzyme that participates in an alternative pathway of the glutamylation of tRNA<sup>Gln</sup> in organisms that are deprived of GlnRS.<sup>8</sup> No sequence homology was found between the polypeptide chain that connects the core of GlnRS to the GatB-like domain but its amino acid composition suggests that it might be an unstructured and flexible linker. Despite the presence of this 60–80 residue-long connector that was a priori not favorable for crystallization, we decided to keep the entire extension in our construct because it is indispensable for an optimal enzymatic activity.<sup>5</sup> A first encouraging result was that the full-length protein was homogeneous and monodisperse when analyzed by dynamic light scattering. A second one was that crystallization trials readily produced well-diffracting crystals in the presence of PEG 3350. They were used to determine a structure of the enzyme at 2.3 Å resolution (Figure 1).

**3.2 Crystals with much disorder in their packing.** The three-dimensional structure of *Dr-GlnRS* was solved by molecular replacement using the structure of the *E. coli* ortholog (PDB id 1O0B, 554 residues displaying 44% identity with the core of *Dr-GlnRS*). In the orthorhombic crystals characterized by a solvent content of 44% (v/v) the greater part of the GlnRS core is ordered in the packing. By contrast, the N-terminal tail, two loops of the core and the C-terminal extension are too mobile to be observable in the electron density map. Overall, 297 residues amongst 852 (*i.e.* 35%) are missing in the final model, although their presence in the crystals was confirmed by SDS-PAGE and mass spectrometry analyses (data not shown). A model of the complete enzyme was built<sup>5</sup> in which the floppy regions of the polypeptide chain can easily be accommodated in the large solvent channels of the crystal. Additional data on experimental procedures and crystal packing are given in the supplementary material. Small substrates of the enzyme (glutamine and a synthetic glutaminyladenylate analog) were used in co-crystallization assays to attempt to stabilize the floppy parts of the protein. The packing geometry and the diffraction properties of the substituted crystals did not change significantly, and no major stabilizing effect was observed despite a slight alteration of cell parameters (data not shown).

**3.3. Lessons for protein crystallogenesi.** This work on *Dr-GlnRS* is encouraging for structural biologists faced with difficult crystallization problems. Indeed, significant disorder can exist inside well-diffracting crystals that are suitable for structure determination. In other words, proteins with large flexible and/or unstructured appendices might be crystallized. Nowadays, the analysis of genomic sequences often detects open reading frames that code for proteins made of folded domains and polypeptides deprived of any apparent structure. As suggested by the work on *Dr-GlnRS*, three prerequisites must be fulfilled to allow the crystallization of these types of proteins and the growth of high quality crystals. First, they must be homogeneous in solution according *e.g.* to light scattering criteria. Second, flexible appendices, if present, should not interfere with crystal packing. These are often detrimental to crystallization, as seen with the aspartyl-tRNA synthetase (AspRS) from yeast, where the presence of a 70 residue-long N-terminal extension (representing

~12% of AspRS monomer) did interfere with intermolecular contacts. Indeed, the original poor-diffracting crystals could eventually be enhanced by the removal of the extension.<sup>9</sup> Third, the size of the solvent channels should be large enough to accommodate the floppy domains. In the present case, disordered regions can easily find some empty room in the packing since the visible atoms only correspond to 43% of the unit cell content (Table 1). This might even be a more general feature taking into account that protein crystals can contain up to 80% of solvent.

#### 4. Conclusion

Well-diffracting crystals of this atypical GlnRS illustrate the fact that a highly ordered packing can accommodate quite dynamic and flexible macromolecular species. In the present case, electron density is missing for approximately one-third of the protein's residues, presumably because of the inherent flexibility of those regions. Thus, at least in this example, macromolecular crystal packing appeared to be tolerant that might be expected. Also, this shows that the removal of flexible regions is not always necessary to ensure successful crystallization as long as the protein sample is homogeneous in solution, for instance according to light scattering measurements.

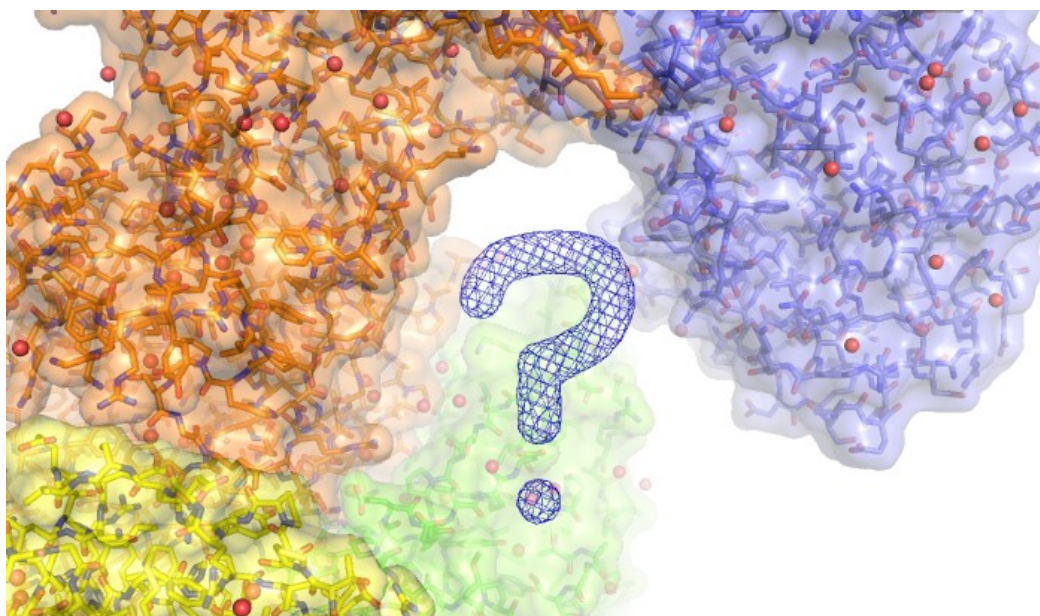
**Acknowledgements.** This work was supported by Université Louis Pasteur (Strasbourg), Centre National de la Recherche Scientifique, and grants from Association pour la Recherche sur le Cancer (ARC) and the French Ministry for Research (ACI-BCMS 042358). The authors acknowledge the teams of ID14 and BM30 beamlines at ESRF (Grenoble, France) for assistance during data collection. C.S. was the recipient of a Marie Curie European Reintegration Grant (MERC-CT-2004-004898). This work was presented at the 11<sup>th</sup> International Conference on Crystallization of Biological Macromolecules (ICCBM-11) held in Québec-City, Canada (August 16–21, 2006).

### References

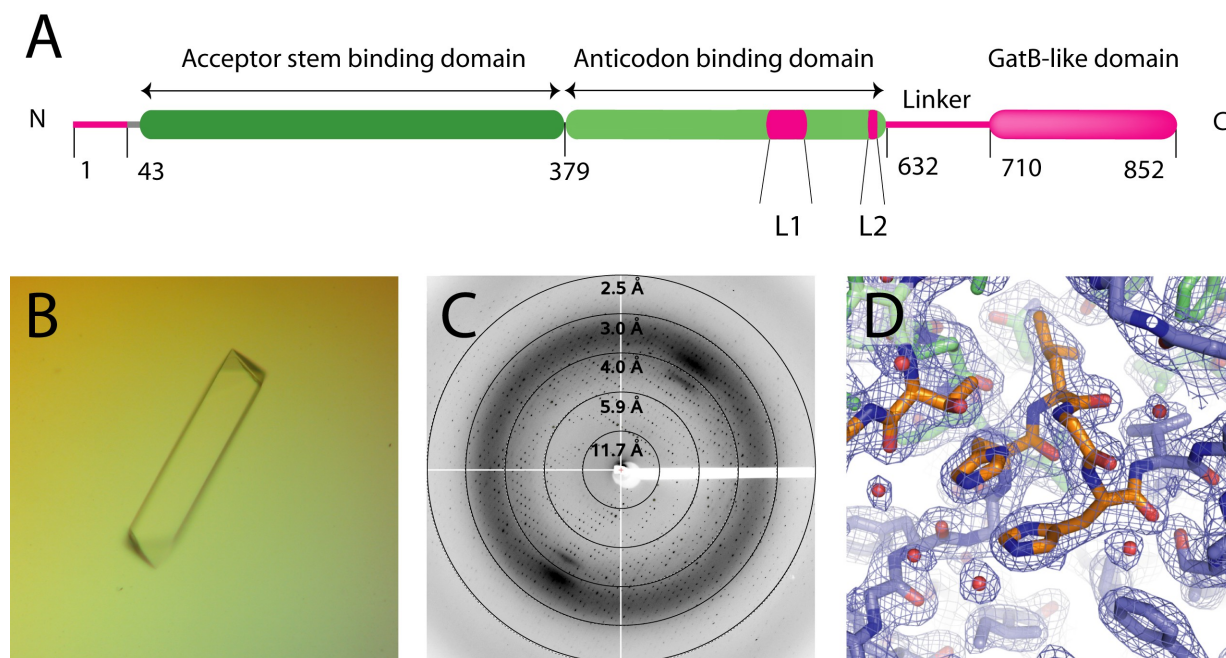
- (1) Pusey, M. L.; Liu, Z. J.; Tempel, W.; Praissman, J.; Lin, D.; Wang, B. C.; Gavira, J. A.; Ng, J. D. *Prog. Biophys. Mol. Biol.* **2005**, *88*, 359–386.
- (2) Deniziak, M.; Sauter, C.; Becker, H. D.; Giegé, R.; Kern, D. *Acta Crystallogr. D* **2004**, *D60*, 2361–2363.
- (3) Chayen, N. E.; Shaw Stewart, P. D.; Blow, D. M. *J. Crystal Growth* **1992**, *122*, 176–180.
- (4) Claude, J.-B.; Suhre, K.; Notredame, C.; Claverie, J.-M.; Abergel, C. *Nucleic Acids Res.* **2004**, *32*, W606-W609.
- (5) Deniziak, M., Sauter C., Becker H. D., Paulus C., Giegé R.; Kern D. *Nucleic Acid Res.* **2007**, *35*, 1421–1438.
- (6) White, O.; Eisen, J. A.; Heidelberg, J. F.; Hickey, E. K.; Peterson, J. D.; Dodson, R. J.; Haft, D. H.; Gwinn, M. L.; Nelson, W. C.; Richardson, D. L.; Moffat, K. S.; Qin, H.; Jiang, L.; Pamphile, W.; Crosby, M.; Shen, M.; *et al.* *Science* **1999**, *286*, 1571–1557.
- (7) Perona, J. **2005**. In *Aminoacyl-tRNA synthetases*. Ibba, M.; Francklyn, C.; Cusack, S. (eds) Landes Bioscience, Georgetown, pp. 73–88.
- (8) Curnow, A. W.; Hong, K.-W.; Yuan, R.; Kim, S.; Martins, O.; Winkler, W.; Henkin, T. M.; Söll, D. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 11819–11826.
- (9) Sauter, C.; Lorber, B.; Théobald-Dietrich, A.; Giegé, R. *J. Crystal Growth* **2001**, *232*, 399–408.

### Table of content graphic and synopsis

Excellent diffraction properties do not necessarily mean that there is much order in crystals made of biological macromolecules. We report an extreme case of a 90 kDa enzyme that forms orthorhombic crystals diffracting to  $\sim 2$  Å resolution but in which one third of the protein is disordered and invisible in the electron density map.

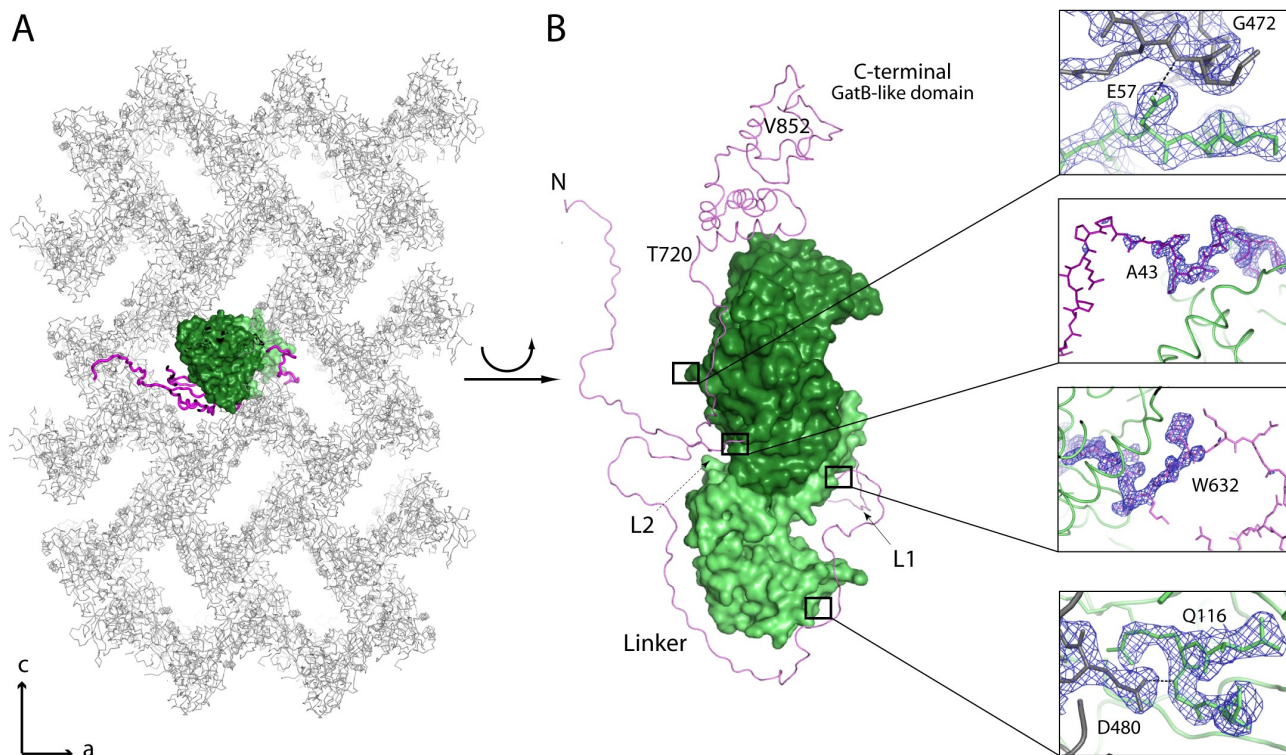


## Figures and Tables



**Figure 1.** Crystals of a protein with almost one-third of floppy regions can yield good diffraction: (A) Schematic representation of the *D. radiodurans* GlnRS primary structure. Regions that are disordered in the crystal packing are indicated in pink. L1 et L2 correspond to loops 537-565 and 614-619 that are mobile in the GlnRS core. The acceptor stem and tRNA anticodon binding domains of the enzyme are shown in dark and light green, respectively; (B) A typical prismatic crystal measuring ~0.3 mm in length; (C) Diffraction image (0.5° oscillation) collected on beamline ID14-1, ESRF, France. The resolution was 1.9 Å at the corner of the ADSC quantum 4 detector; (D) Close-up view showing the electron density map ( $2Fo-Fc$  map, 2.3 Å resolution, contour level of 1.3 sigma) in the neighborhood of the protein active site. This region is in the acceptor stem binding domain (or catalytic domain), close to where glutamine and ATP are bound to form glutaminyladenylate, prior to transfer of glutamine at the 5' end of the tRNA acceptor stem.





**Figure 2.** Views of the crystalline packing of *Dr*-GlnRS: (A) Crystal packing with large solvent channels oriented along the b direction. (B) Model of the full-length GlnRS represented with the same color code as in Figure 1: the GlnRS core is shown in green and disordered regions in pink (including the N-terminal tail, loops L1 and L2, the C-terminal extension). The GatB-like C-terminal domain (residues T720-V852) was modeled on the basis of the structure of an homolog domain from *Bacillus subtilis* (PDB id: 1NG6). On the right, close-up views of the boxed areas showing the abrupt transition from ordered to disordered regions in N- and C-terminus (neighborhood of residues A43 and W632, respectively) of the polypeptide chain in the electron density map. As a comparison, two views of contact regions (involving residues E57-G472, Q116-D480) are also depicted.

**Table 1. Characteristics of *Dr*-GlnRS crystals**

Average crystal length	0.2–0.4 mm
Space group	$P2_12_12_1$
Cell parameters a, b, c	74.1 Å – 95.9 Å – 115.7 Å
Resolution limit	2.3 Å
Asymmetric unit	1 monomer (852 residues)
Solvent content	44 % (v/v)
Matthews coefficient	2.2 Å <sup>3</sup> /Da
Number of ordered residues	555 (65%)
Number of disordered residues	297 (35%)
Ordered cell volume <sup>a</sup>	43%

(<sup>a</sup>) Percentage of the cell volume occupied by ordered residues