The conserved yeast protein Knr4 involved in cell wall integrity is a multi-domain intrinsically disordered protein

Manon Batista, Ellen I.M. Donker, Cécile Bon, Myriam Guillien, Adriana Caisso, Lionel Mourey, Jean-Marie François, Laurent Maveyraud, Didier Zerbib

The conserved yeast protein Knr4 involved in cell wall integrity is a multi-domain intrinsically disordered protein

Manon Batista [1,2,†], Ellen I.M. Donker [1,2,†], Cécile Bon [2,†], Myriam Guillien [1,2], Adriana Caisso [1], Lionel Mourey [2], Jean Marie François [1], Laurent Maveyraud [2,#] and Didier Zerbib [1,#]

[1] *Toulouse Biotechnology Institute (TBI), Université de Toulouse, CNRS, INRAE, INSA, F-31077 Toulouse, France*

[2] *Institut de Pharmacologie et de Biologie Structurale (IPBS), Université de Toulouse, CNRS, UPS, F-31062 Toulouse, France*

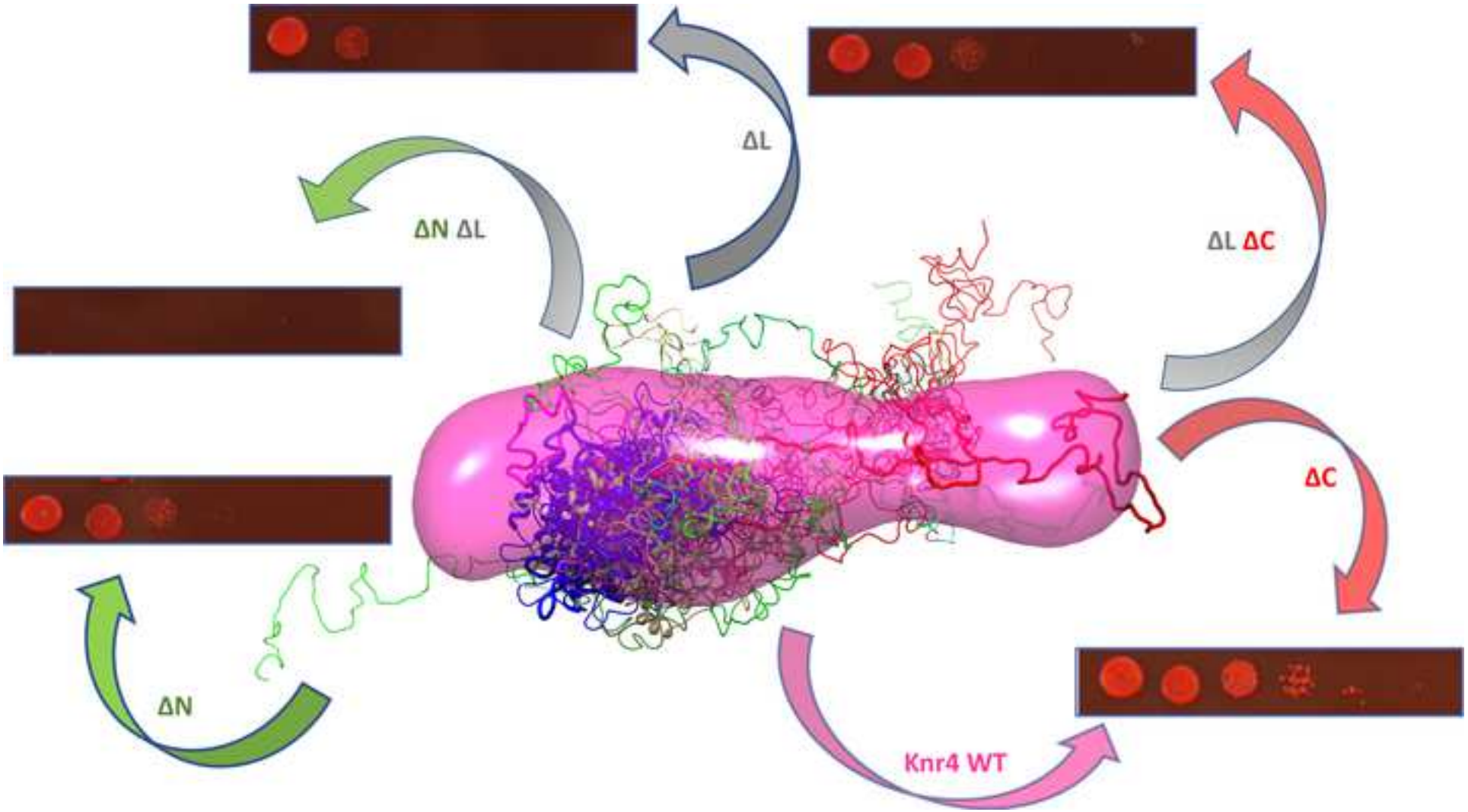
† These authors contributed equally to this work

Correspondence to : Didier Zerbib, *Toulouse Biotechnology Institute (TBI), Université de Toulouse, CNRS, INRAE, INSA, F-31077 Toulouse, France,* didier.zerbib@insa-toulouse.fr , +33 (0) 561 559 966 and Laurent Maveyraud, *Institut de Pharmacologie et de Biologie Structurale (IPBS), Université de Toulouse, CNRS, UPS, F-31062 Toulouse, France,* laurent.maveyraud@ipbs.fr , +33 (0) 561 175 435.

Running title: A complete Structural view of the yeast protein Knr4

Graphical Abstract (for review)

Click here to access/download;Graphical Abstract (for review);Graphical abstract_Batista et al.tif



The yeast protein Knr4 is located at the crossroads of essential cellular processes.

Knr4 consists of three large disordered regions flanking a structured core

Knr4 coordinate domains are critical to its function in parietal stress resistance.

Knr4 is an interesting new target for future antifungal combination therapies

Abstract

Knr4/Smi1 proteins are specific to the fungal kingdom and their deletion in the model yeast *Saccharomyces cerevisiae* and the human pathogen *Candida albicans* results in hypersensitivity to specific antifungal agents and a wide range of parietal stresses. In *S. cerevisiae*, Knr4 is located at the crossroads of several signalling pathways, including the conserved cell wall integrity and calcineurin pathways. Knr4 interacts genetically and physically with several of their components. Its sequence suggests that it contains large intrinsically disordered regions. Here, a combination of small-angle X-ray scattering (SAXS) and crystallographic analysis led to a comprehensive structural view of Knr4. This experimental work unambiguously showed that Knr4 comprises two large intrinsically disordered regions flanking a central globular domain whose structure has been established. The structured domain is itself interrupted by a disordered loop. Using the CRISPR/Cas9 genome editing technique, strains expressing *KNR4* genes deleted from different domains were constructed. The N-terminal domain and the loop are essential for optimal resistance to cell wall-binding stressors. The C-terminal disordered domain, on the other hand, acts as a negative regulator of this function of Knr4. The identification of molecular recognition features, the possible presence of secondary structure in these disordered domains and the functional importance of the disordered domains revealed here designate these domains as putative interacting spots with partners in either pathway. Targeting these interacting regions is a promising route to the discovery of inhibitory molecules that could increase the susceptibility of pathogens to the antifungals currently in clinical use.

Keywords

Intrinsically disordered protein, *Saccharomyces cerevisiae*, Integrative Structural Biology, Knr4/Smi1, Cell wall integrity

Introduction

Proteins that do not have a defined 3D structure over all or part of their sequence, but nevertheless have biological activity, belong to the ubiquitous family of intrinsically disordered proteins (IDPs) [1-6]. These proteins can have different types of disorder which can be restricted to certain regions known as intrinsically disordered protein regions (IDPRs) [2, 3]. IDPs and IDPRs are frequently involved in crucial cellular processes such as signal transduction, gene regulation, and maintenance of homeostasis and are often implicated in human pathologies [1, 7, 8]. Their functions are mainly derived from their ability to interact with a large number of partners [1] through potentially different (dis)ordered structural elements. The plasticity of IDPs and IDPRs, which can exist as ensemble of conformations, gives them structural adaptability and allows them to occupy hub positions in protein-protein interaction (PPI) networks [1, 9, 10]. One of this central protein in *Saccharomyces cerevisiae* is Knr4.

Knr4 contains large IDPRs, is considered as an important hub of the yeast interactome and is located at the crossroad of major parietal stress signalling pathways involved in stress signalling and antifungal sensitivity [11]. In *S. cerevisiae*, Knr4 is the representative of a conserved family of fungus-specific proteins called Knr4/Smi1 [12]. *KNR4* was initially identified in the yeast *Hansenula mrakii* during a search for genes affecting cell wall β-1,3-glucan biosynthesis [13, 14]. Deletion of *KNR4* has multiple physiological consequences: *knr4Δ* mutants are hypersensitive to cell wall-targeting compounds such as caffeine, Congo red (CR) and calcofluor white (CFW) [11, 15], they also show increased sensitivity to high temperatures and to antifungal agents such as caspofungin and cercosporamide [13, 15-17]. These phenotypes appear to be correlated with changes in the cell wall composition and the concentration of chitin, to which CR and CFW bind, increases while that of β-glucan decreases significantly [13].

Two signalling pathways play a central role in cell wall biogenesis, maintenance and stress resistance: the cell wall integrity (CWI) pathway, [18] and the calcineurin (CN) pathway [19]. In the CWI pathway, Knr4 is required for appropriate targeting of the transcription factors Rlm1p and Swi4p by

the mitogen-activated protein kinase (MAPK) Slt2 [20] involved in cell wall remodelling [18, 21]. The link between *KNR4* and the CN pathway was supported by the hyper-sensitive phenotype to calcium stress exhibited by *knr4Δ* mutants [22]. The CWI and CN pathways are not essential individually, but each becomes essential if the other is inactivated [18]. Similarly, the *KNR4* gene is not essential for yeast growth, but its deletion makes both pathways (CWI and CN) essential and renders the cell hypersensitive to cell wall stresses and to antifungal agents [20, 23]. Knr4 physically interacts with the Mitogen-Activated Protein Kinase (MAPK) Slt2 of the CWI pathway and with the Calcineurin phosphatase Cna1 of the CN pathway [20, 24]. This indicates that Knr4 is probably also physically at the crossroad of these two signalling pathways.

More generally, the critical role played by Knr4 on yeast physiology is highlighted by the existence of a very large number of observed genetic synthetic lethal (more than 250) and physical (39 protein partners) interactions [11, 20, 25-30]. Beside the CWI and CN signalling pathways members, Knr4's partners are involved in a variety of central cellular processes related mainly to morphogenesis and stress response. Knr4 partners have functions in bud emergence, polarity development, cell secretion, transcription and ubiquitination [20, 26, 28-30]. Interestingly, *KNR4* orthologs in *C. albicans*, *SMI1* and *SMI1B*, are involved in pathogenesis and in biofilm resistance to antifungal agents [31, 32].

Knr4 has two experimentally-confirmed properties out of the three required for a hub protein: it is at the crossroad of essential cellular pathways and it has a plethora of partners in the yeast PPI network. The third property required for Knr4 to be a hub in PPI networks is intrinsic structural flexibility [33, 34]. The analysis of the Knr4 sequence suggests the presence of two large disordered domains flanking a potentially structured core [11, 23]. When produced ectopically, the structured core is capable of complementing most of the *S. cerevisiae knr4* null mutant phenotypes whereas it is no more able to physically interact with members of the CWI or the CN pathway like the MAPK Slt2 and Cna1. The N- and C-terminal parts seem to respectively ensure or modulate specific protein-protein interactions [11, 23, 24, 28]. It is possible, but not yet demonstrated to date, that these

domains are truly disordered and that PPIs are responsible for specific functions of Knr4 in maintaining parietal integrity [20, 24].

The main objectives of this work were first to obtain a comprehensive structural view of the Knr4 protein, which potentially carries large IDPRs that may explain its status as a hub protein in yeast. Second, we aimed to try to correlate this multi-domain structure with its role in the sensitivity of yeast to parietal integrity disruptors. The multi-domain organisation of Knr4, as well as the ordered or disordered nature of each domain, were previously only predicted from the sequence [23]. We used a combination of small-angle X-ray scattering (SAXS) and crystallographic analysis to decipher the structural organization of Knr4. In addition, using the CRISPR/Cas9 genome editing technique, we constructed strains expressing various genomic deletions of the Knr4 domains at the *KNR4* locus and analysed their sensitivity to well-known cell wall disruptors (CR and CFW).

This work revealed that Knr4 comprises two large IDPRs and a smaller disordered loop and their importance in maintaining cell wall integrity. The identification of molecular recognition features (MoRFs), the possible presence of secondary structure in these disordered domains, as well as the importance of the disordered domains for Knr4 function, suggested that they might be involved in interactions with partners in either pathway.

In eukaryotes, Knr4/Smi1 proteins are specific to the fungal kingdom and their deletion in the yeasts *S. cerevisiae* and *C. albicans* leads to hypersensitivity to certain antifungal agents [11, 31, 32]. Targeting Knr4 interaction regions could lead to promising compounds that could be used in combinatorial therapies to reduce the high doses of antifungal agents currently required [35] to control human pathogenic fungi.

Results

Knr4 presents an ordered core flanked by two disordered domains in solution

The structures in solution of different constructs of Knr4 (Figure 1) were investigated using Small Angle X-ray Scattering (SAXS). More specifically, SAXS data allowed to determine the oligomeric state of Knr4, Knr4$^{\Delta C}$ and Knr4$^{\Delta N\Delta C}$. In addition, the degree of disordered was investigated using Kratky analysis for these molecular species and molecular envelopes were derived from the scattering data. Finally, a comparison was performed between the solution structure and the crystallographic structure (see below) of Knr4$^{\Delta N\Delta C}$ and models of the entire Knr4, Knr4$^{\Delta C}$ and Knr4$^{\Delta N\Delta C}$ entities, considering their flexible regions, were proposed

Knr4, Knr4$^{\Delta C}$ and Knr4$^{\Delta N\Delta C}$ are monomeric in solution. The scattered intensities of Knr4, Knr4$^{\Delta C}$ and Knr4$^{\Delta N\Delta C}$ closely follow the Guinier law in the small-*Q* region and no sign of aggregation is observed (Supplementary Figure 1). As slight proteolysis was apparent for the Knr4$^{\Delta C}$ construct, on-line HPLC was used to isolate an intact fragment. Data with an elution volume corresponding to the intact Knr4$^{\Delta C}$ and stable R_g were merged. A concentration dependence in the small-*Q* region was observed for Knr4$^{\Delta N\Delta C}$ starting at 5 mg.mL^{-1} and above and data at 2 mg.mL^{-1} were chosen for thorough investigations. The molecular masses deduced by forward scattering, Rambo and Tainer analysis and from Porod volume through the SaxsMoW calculator agree with the theoretical molecular masses, and reflect a monomeric behaviour of all variants at the probed concentrations (Table 1).

The core of Knr4 has the same structure in the crystal and in solution. Chain A was extracted from the crystallographic structure of Knr4$^{\Delta N\Delta C}$ and the missing residues (80-83, 149-154, 189-217, 301-309), mainly corresponding to loops, were modelled using Allosmod-FoXS [36]. The best model led to an excellent agreement (χi=0.761, compared to 4.870 without modelling of the missing loops) with the corresponding experimental scattering curve. The superposition of this completed Knr4$^{\Delta N\Delta C}$ structure to its low-resolution model in solution is shown in blue in Figure 2.

Knr4 contains large intrinsically disordered domains. The normalised Kratky plot of the central Knr4$^{\Delta N\Delta C}$ domain displays a characteristic behaviour of a globular domain, with a typical bell-shape having a maximum of about 1.1 near $Q.R_g=\sqrt{3}$ (Figure 3). On the contrary, for a random chain, the curve would rise with increasing $Q.R_g$ to reach a nearly flat region at a value between 1.5 and 2 followed at high-Q values (typically Q > 2-3 nm^{-1}) by a further increase depending on the rigidity of the polypeptide chain [37, 38]. Knr4$^{\Delta C}$ exhibits a slight increase in disorder compared to Knr4$^{\Delta N\Delta C}$, but overall seems relatively ordered. On the contrary, the entire Knr4 protein clearly displays a behaviour of a disordered protein with intermediate flexibility. The normalised distribution of intramolecular distances (Figure 3, inset) shows, in coherence with the Kratky plots, that Knr4$^{\Delta N\Delta C}$ displays a bell shape characteristic of a globular domain, with a most probable intramolecular distance of about 2.3 nm and a maximum dimension of about 8 nm. Knr4$^{\Delta C}$ and Knr4 display most probable intramolecular distances of about 2.5 nm, as reflected by the maximum of the bell, but are much more elongated, with maximum distances of respectively 11.0 and 17.0 nm, as inferred from the most extended conformation significantly present in solution. This indicates that residues 346-505 are disordered. The *Rg* value, as obtained from the Guinier analysis, is for Knr4$^{\Delta N\Delta C}$ of 2.3 ± 0.1 nm. This is similar to the expected theoretical value of 1.9 ± 0.3 nm, and emphasizes that Knr4$^{\Delta N\Delta}$ is close to globular. On the contrary, Rg values of 3.2 ± 0.1 and 4.8 ± 0.3 nm for Knr4$^{\Delta C}$ and Knr4, are intermediate between that of their theoretical globular (2.0 and 2.3 nm, respectively) or IDP counterparts (5.4 and 6.5 nm, respectively) (Table 1).

Ab initio low resolution modelling of Knr4$^{\Delta N\Delta C}$, Knr4$^{\Delta C}$, and Knr4 was performed using DAMMIN resulting in χ-values of 0.90, 1.46 and 1.07, respectively. In the case of the less disordered Knr4$^{\Delta N\Delta C}$, modelling was also done with GASBOR resulting in a χ-value of 0.92 (Table 1 and Supplementary Table1). All calculated curves were in good agreement with their experimental counterparts, especially for the globular Knr4$^{\Delta N\Delta C}$. The normalised spatial discrepancy (NSD) of the various constructs, lower than 0.7, indicates that the obtained solutions are stable, while the increase of NSD from 0.485 for Knr4$^{\Delta N\Delta C}$ to 0.625 for Knr4 can be attributed to the concomitant increase of disorder. As expected from

the *P(r)*, Knr4$^{\Delta C}$ and Knr4 envelopes are composed of a bulge, to which the crystal structure can be accommodated, and an elongated region, which corresponds to the more disordered N- and C-terminus parts of the protein.

The conformational states in solution of Knr4$^{\Delta N\Delta C}$, Knr4$^{\Delta C}$, and Knr4 were analysed using the Ensemble Optimization Method (EOM). Ensemble analysis for Knr4$^{\Delta C}$ led to a χ-value of 0.93 and revealed a set of discrete conformations with an overall R_g and D_{max} respectively of 3.1 and 11.6 nm and an amplitude of R_g values of 2.2 nm (Tables S1, S2, S3). Ensemble analysis for Knr4 led to a χ-value of 0.88 with an overall R_g and D_{max} respectively of 4.6 and 15.6 nm and revealed a larger flexibility with two broad conformational populations leading to an amplitude of R_g values of 3.6 nm. To validate these analyses, the Knr4$^{\Delta N\Delta C}$ data have also been challenged with the EOM program, and this yielded to a single conformation (χ-value of 0.95, overall R_g and D_{max} respectively of 2.3 and 8.6 nm), in coherence with the low flexibility of this construct. The obtained populations can be seen in Figure 2 either fitted in the corresponding low-resolution envelopes (Figure 2B) or aligned via their SD domain (Figure 2C). As already mentioned in studies on flexible entities [37], envelope density is partially lacking around flexible parts of the models deduced from EOM, due to averaging over the multiple conformations of this flexible region.

Crystallographic structures of the ordered core of Knr4

The central Knr4 domain SD, including or not the disordered loop DL (Knr4$^{\Delta N\Delta C}$ and Knr4$^{\Delta N\Delta L\Delta C}$), has been purified as described before [39] and crystallised. Crystals that belong to the hexagonal space group $P6_2$ (a=b=103.0 Å, c=93.4 Å) with two molecules in the asymmetric unit were obtained with both constructs. The refined structure of Knr4$^{\Delta N\Delta C}$ comprises residues 84-148 (residue numbering refers to mature full length Knr4), 155-188, 218-300 and 310-347 in chain A and 84-186, 217-300 and 310-341 in chain B: a few residues are disordered at the N-terminus of both chains (7 from the N-terminus tag and residues 80 to 83) as well at 6 residues at the C-terminus of chain B (missing residues from the tag). The structure of the ordered core of Knr4 is built of a central 6-stranded β-sheet (strands s1 to s6), surrounded with 7 α-helices (helices h1 to h7) and a small 3-stranded antiparallel β-sheet (strands

s7 to s9) (Figure 4). The DL (residues 188-218) that connects helix h5 to strand s1 is disordered in both chains of the asymmetric unit. The final *R* and R_{free} factors are 0.18486 and 0.22614, respectively (Supplementary Table 4). Structure resolution of the SD domain deleted from DL (Knr4$^{\Delta N\Delta L\Delta C}$) showed that this deletion does not result in significant structural modifications: residues bordering the deletion remain disordered in molecule A of the asymmetric unit of Knr4$^{\Delta N\Delta L\Delta C}$ crystals, whereas they are well defined in the electron density map of molecule B. In that molecule, deletion of DL from SD (Knr4$^{\Delta N\Delta C}$) just results in a minor rearrangement at the C-terminus of helix h5 (Supplementary Figure 3).

Two molecules are found in the asymmetric unit, related by a 2-fold axis. They interact with each other through the loops connecting s2 to s3 and s4 to s5. This interface buries 538 ± 2 Å^2 of accessible surface area of each protomer. A somewhat higher buried surface area (575 ± 30 Å^2) is found between two chains A related by crystallographic symmetry. These potential dimerization interfaces are not predicted to be relevant for complex formation, according to PISA analysis [40].

Knr4 defines a fold in the SCOP database [41] called the Smi1/Knr4-like fold and defined as a 3-layers $\alpha\beta\alpha$ fold. Four structures are found with such fold: two hypothetical proteins from *Listeria innocua* (PDBID 2ICG) and *Pseudomonas syringae* (PDBID 2PAG), a putative glycan synthesis regulator from *Bacteroides fragilis* (PDBID 3D5P) and YobK from *Bacillus subtilis* (PDBID 2PRV). This latter protein is the antitoxin component of a type II toxin-antitoxin system, and inhibits the RNAse activity of YobL, its cognate toxin [42]. Comparison of these structures indicates that although the overall fold is similar, there are significant differences when compared to the Knr4 core structure (Figure 5). The fold of 2ICG, 2PAG, 3D5P and YobK are also made of a central 6-strand β-sheet, but with only 5 strands common to the β-sheet of Knr4: strand s2, which borders the β-sheet in Knr4 is lacking in the other structures (Figure 4, Figure 5). Also, the short helix h2 of Knr4 has no equivalent in the other structures. The 30-residues long disordered loop (residues 188-218) and strand s1 in Knr4 have no equivalent in the other structures, as the helix corresponding to helix h5 in Knr4 is directly connected to the β-strand corresponding to s3 of Knr4, similar to what is observed in the structure of Knr4$^{\Delta N\Delta L\Delta C}$. Finally, as the Knr4 core structure described herein correspond to a 263 residues long protein, whereas 2ICG, 2PAG,

3D5P, and YobK only contain 135 to 158 residues, the C-terminal part of Knr4, encompassing strands s7 to s9 and helix h7, is lacking in the other structures. Hence, the Smi1/Knr4 fold as defined in the SCOP or in PFAM databases corresponds only to a part of the Knr4 core structure presented herein, highlighted with a red frame in Figure 5.

The structure-based sequence alignment identifies only 3 strictly conserved residues: Leu128, at the C-terminal side of helix h3, and Pro133 and Lys37, at the N-terminal side and at the middle of helix h4, respectively. Leu128 and Pro133 are part of a hydrophobic cluster of residues, that also includes Phe132, Val136, Phe140, Val170, and Val171. Although these residues are not conserved in all the proteins, the hydrophobic cluster is found in all cases. In Knr4, the side-chain amine group of Lys137 points at the surface of the protein, at hydrogen bond distance from the side chain of Glu125 (Glu or Gln in the other structures) and Asn134.

Knr4 disordered domains are required for efficient cell-wall stress resistance in yeast

To avoid expression artefacts due to ectopic production of complementation genes, CRISPR/Cas9 technology was used to construct genomic deletions of disordered domains of Knr4. A series of plasmids was constructed allowing to generate PCR fragments for allelic exchange at the *KNR4* locus with genes encoding Knr4 protein deleted of one domain (Knr4$^{\Delta N}$ lacking DD1, Knr4$^{\Delta L}$ lacking DL, and Knr4$^{\Delta C}$ lacking DD2), two domains (Knr4$^{\Delta N\Delta L}$, lacking DD1 and DL, Knr4$^{\Delta N\Delta C}$ lacking DD1 and DD2, and Knr4$^{\Delta L\Delta C}$ lacking DL and DD2) or all three domains (Knr4$^{\Delta N\Delta L\Delta C}$, lacking DD1, DL and DD2) (Figure 1). CR and CFW, known to bind growing chitin chains in the cell wall and affect its assembly [43, 44], were used to assess the role of each Knr4 domain in cell wall stress sensitivity [45, 46]. The *slt2Δ* strain, deleted for the CWI pathway MAPK Slt2 [47], was used as a reference strain as it is hypersensitive to cell wall stressors like CR and CFW. In addition, Slt2 is known to interact with Knr4 [20, 28, 48]. As expected, *slt2Δ* and the *knr4Δ* strains are hypersensitive to CR and CFW, while the wild-type strain is fully resistant in the same conditions (Figure 6(a)). Deletion of DD2 alone does not significantly affect

resistance whereas deletion of DD1 or DL increased susceptibility, to a higher extent in the latter case. Finally, when the DD1-DL or DD1-DL-DD2 domains were deleted, the strain was as susceptible as the *knr4Δ* and *slt2Δ* strains. In all cases, further deletion of DL results in higher susceptibility to CW stresses. More precision was obtained on the effect of each deletion by following the growth curves of these strains (Figure 6B) in the presence or absence of CR or CFW. This was particularly true in the presence of CR where three phenotypic groups emerged, represented in blue, green and orange on figure 6B. The importance of DD1 and DL in stress resistance was still clearly visible, as their deletion resulted in a significantly reduced growth rate, and their combined deletion renders the strain as susceptible as the *knr4* or *slt2* null mutants. The deletion of DD2 alone does not have a visible effect on the growth rate. However, when DD2 is deleted in addition to either DD1 or DL deletion, the growth rate is restored to normal. It therefore appears that deletion of DD2 is able to rescue the effect of the deletion of DD1 or DL. However, the deletion of DD2 could not rescue the combined deletion of DD1 and DL. This suggests that the inhibitory effect of DD2 might be due to an interaction with both domains (DD1 and DL) at the same time. In the presence of CFW, a similar trend was observed. But, as the effect of CFW on growth is weaker, differences in growth rate were small and the results were less clear.

Discussion

Earlier sequence analysis classified Knr4 in the family of proteins containing large IDPRs [2]. Indeed, three of the four identified structural domains were predicted as disordered: DD1 (residues 1-80) at the N-terminus, the central DL (residues 185-210), and DD2 at the C-terminal region (residues 340-505) [11, 23]. SAXS data collected with Knr4, Knr4$^{\Delta C}$, and Knr4$^{\Delta N\Delta C}$ unambiguously showed that all investigated Knr4 species were monomeric in solution and experimentally confirmed this structural organisation. The crystal structures of the ordered core of Knr4, with and without DL were determined: the asymmetric unit of the crystal contained two interacting molecules of Knr4$^{\Delta N\Delta C}$. However, all crystal contact interfaces are predicted to be irrelevant according to PISA analysis, in agreement with our observations. As IDPs are known to be prone to proteolysis [3, 49], a careful attention was given to the extraction of the molecular mass in solution from the SAXS data. All three molecular species displayed molecular masses, either calculated from Guinier, concentration-independent calculation, or based on the Porod volume, in close agreement with the theoretical masses. This reflects that our SAXS data were collected on non-proteolyzed molecular species and rules out that the observed disorder results from degradation.

Kratky plot analysis highlights that Knr4$^{\Delta N\Delta C}$ is globally well folded (with the notable exception of the DL, that encompasses 11% of the Knr4$^{\Delta N\Delta C}$ construct). Knr4$^{\Delta C}$ is found to be slightly flexible, suggesting that DD1 is only partially unfolded, whereas entire Knr4 is by far the most flexible entity, highlighting the higher disorder content of DD2. These observations agree with a previous circular dichroism study [23], which indicated that DD2 lacked any secondary structure in all evaluated constructions, while DD1 displayed some helical content only when DD2 was deleted. However, in the full-length protein, DD1 was also devoid of any structuration: the presence of DD2 seems to prevent secondary structure formation in DD1.

The crystallographic structure of Knr4$^{\Delta N\Delta C}$ suggests that some loops are disordered in the SD domain. While residues 149-154 are only disordered in chain A of the asymmetric unit, residues 185-216 and 301-309 are disordered in both molecule of the asymmetric unit. Residues 188-218

correspond to the DL, which was identified as potentially disordered according to sequence analysis (Durand, 2008 #61;Martin-Yken, 2016 #75). The deletion of residues 189-217, which include DL, does not significantly alter the structure of the ordered part of SD, as shown with the crystal structure of Knr4$^{\Delta N \Delta L \Delta C}$.

It is quite common that IDPs/IDPRs are partially unstructured (and not complete random coils) and contain regions of pre-structured motifs or PreSMos, which may serve as targets for their interaction partners [50]. The potential presence of α-helices in DD1 of Knr4, as evidenced by CD data (Durand, 2008 #61), could be of importance for its hub function as it has been observed previously that 62% of the protein-protein interactions are mediated by α-helical structures at the interface [51]. Search for MoRFs within Knr4 [34] pointed to five regions: one in the DD1 domain (residues 5-10), three in the SD core (residues 91-100, 208-213 and 332-341) and one in the DD2 region (residues 497-505) (Figures 4, Figure 5). When projected on the structure of the Knr4$^{\Delta N \Delta C}$, it appears that residues 91-100 build the central part of helix h1, residues 208-213 are part of the DL loop, and residues 332-441 build the C-terminal part of helix h7. These 3 regions are localized at the surface of the structure of Knr4$^{\Delta N \Delta C}$, and only h1 is present in the Smi1/Knr4 fold as defined in SCOP/PFAM (Figure 4). Indeed, the DL loop and helix h7 are only found in Knr4. Hence, two of these MoRFs could promote interactions specific to Knr4. According to our *in vivo* data, DD1 and DL are important for the biological function of Knr4 tested here, which is not the case for DD2. Indeed, the single deletion of DD1 or DL results in an increased sensitivity to CR and a slower growth, and their simultaneous deletion results in the same phenotype as the Δ*knr4* mutant. On the opposite, deletion of DD2 does not alter the phenotype nor the growth. Our data even suggests that deletion of DD2 might partly compensate the effect of DL deletion: sensitivity to CR is lower and growth is faster for *knr4*$^{\Delta L \Delta C}$ than for *knr4*$^{\Delta C}$. This is particularly clear when growth in the presence of CR was analysed. Simultaneous deletion of DL or DD1 with DD2 did not affect growth, whereas individual deletions of DL or DD1 caused a clear decrease in growth rate. This suggests that DD2 might act as a negative regulator of the function of Knr4. In that respect, CD data obtained previously by Durand and colleagues [23] are particularly interesting as they suggest

that the presence of DD2 decreases the secondary structure content of DD1, which might result in decreased interactions with partners. Indeed, disorder domains of many IDPs have been shown to be involved in interactions with multiple partners [10]. Although the N-terminus of Knr4 becomes essential when CWI pathway is impaired [20] and the absence of DD1 prevents interaction with Slt2 and Cna1 [23], we can hypothesize that Knr4 interacts through this domain with other proteins.

Knr4 stands at the crossroad of the CWI and the CN pathways and we have demonstrated that it includes two large IDPRs and a smaller disordered loop. The importance of DD1 and DL for the proper function of Knr4 leads us to hypothesize that these domains might be involved in interactions with partners in either pathway. The identification of MoRFs in DD1 and DL, as well as the presence of secondary structure elements in DD1, at least in the absence of DD2, corroborate this hypothesis.

As Knr4/Smi1 proteins are specific to the fungal kingdom and their deletion in the yeasts *S. cerevisiae* and *C. albicans* leads to hypersensitivity to certain antifungal agents [11, 31, 32], the development of Knr4 inhibitors directed against these key domains could lead to promising antifungal compounds. Combination therapies using Knr4 inhibitors in conjunction with actual antifungal agents could significantly reduce the high damaging doses currently required to kill human pathogenic fungi [35].

Materials and Methods

Strains and plasmids

Plasmid constructions were performed in MC1061 recA (*E. coli* K12; F^- araD139 Δ(ara-leu)7696 galE15 galK16 recA1 Δ(lac)X74 rpsL (StrR) hsdR2 (r_K^- m_K^+) mcrA mcrB1) using classical cloning procedures and as recommended by enzymes and product manufacturers (Fermentas, Promega, NE Bioloabs). Culture media (Lennox-Broth) were supplemented with ampicillin (150 µg/mL) and agar when needed. The genes encoding, full length Knr4 (Knr4$^{1-505}$), Knr4$^{\Delta N\Delta C}$ (Knr4$^{80-340}$) and Knr4$^{\Delta C}$ (Knr4$^{1-345}$) were cloned into pGEX-6P-3 vector (GE Healthcare), in C-terminal fusion of a cleavable (PreScission™ protease (GE Healthcare)) glutathione S-transferase (GST). Plasmids pGEX-6P-3::Knr4 and pGEX-6P-3:: Knr4$^{\Delta N\Delta C}$ have been described [23, 39]. Plasmid pGEX-6P-3:: Knr4$^{\Delta C}$ was constructed by introducing two in frame stop codons at codons 346 and 347 of *KNR4* in pGEX-6P-3::Knr4 by mutagenesis using self-complementary PCR primers (*KNR4*-Mut-Top and -Bot; Supplementary Table 5). Vector pGEX-6P-3::Knr4$^{\Delta N\Delta L\Delta C}$ was constructed by inverse PCR on pGEX-6P-3 using pGEX-6P-3::Knr4 as matrix and primers allowing the precise deletion of DL (ΔL-Fwd and -Rev; Supplementary Table 5). In order to generate the repair fragments used in the CRISPR/Cas9 genome editing experiments, a genomic fragment from the *KNR4* locus of yeast BY4741 (MATa *his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*) where *KNR4* is flanked by 1096 bp of upstream DNA (promoter) and 341 bp of downstream DNA (terminator) was amplified with specific primers (*KNR4*-Fwd and -Rev; Supplementary Table 5) and then cloned into an intermediate vector (pZE13-MCS, Expressys) between its BamH1 and Xba1 sites. This vector, pZE13::Pro-*KNR4*-Ter, was then used as a template to generate all Knr4 domain deletions shown in Figure 1 by reverse PCR using specific primers (ΔN-Fwd and -Rev, ΔL-Fwd and -Rev and ΔC-Fwd and-Rev; Supplementary Table 5).

Yeast strain construction by CRIPR/Cas9 genome editing

The yeast strain used as target was BY4741 *knr4Δ*::KanMX4 from the YKO MATa Strain Collection (Open Biosystem). The guide RNA targeting the KanMX4 cassette as well as the Cas9 protein was

provided by the plasmid pGZ110::gKan, a derivative of pML107 (Addgene) generously provided by Bruce Futcher (Stony Brook School of Medicine). This *E. coli*/yeast shuttle plasmid carries the LEU2 gene as a selection marker in yeast. The repair fragments were obtained by amplification of pZE13::Pro-KNR4-Ter derived vectors each carrying one or more domain deletions (Figure 1). BY4741 *knr4Δ*::KanMX4 cells were harvested at early exponential growth (OD_{600nm} between 0.3-0.6), washed (Lithium acetate 100 mM in TE buffer) and then concentrated (10^7 cells in 40 µL) in the same buffer. These cells were transformed in a mixture containing 250 ng of pGZ110::gKan and 1 µg of repair fragment in the presence of 25 µg of carrier DNA (Salmon sperm ssDNA-Sigma) and 50% PEG(8000)/LiAc/TE to make up a total volume of 215 µL. After 30 minutes incubation at 30°C, 13 µL of DMSO (NEB, B0515) were added. After a 10-minutes heat shock at 42°C, cells were washed three times with 1 mL of YPD medium and incubated again for 10 minutes at 30°C. The cells were then centrifuged, resuspended in 150 µL of YPD and plated on YNB-Leu selective plates. The plates were incubated for 48 hours at 30°C and the colonies were transferred to YPD plates. The KNR4 locus was then amplified for verification by sequencing of each deletion (Eurofins).

Yeast Drop Dilution Growth Assay

To perform growth assays on solid media, cell cultures were grown in YPD medium up to an OD_{600} between 0.8 and 1.0 and then adjusted to an OD_{600} of 0.9 in YPD. Tenfold serial dilutions were prepared and spotted onto YNB-CSM agar plates to evaluate their sensitivity to cell wall stress agents. When indicated, YNB-CSM medium was supplemented with 50 mM 2-(N-Morpholino) ethanesulfonic acid (MES) buffered to pH 6. Where indicated, YNB-CSM plates were supplemented with either Congo Red (Sigma-Aldrich), Calcofluor White (ICN Biomedicals), SDS (Euromedex), or Caffeine (Sigma-Aldrich). Plates were incubated for 72h to 96h at 30°C (or 38°C where indicated).

Saccharomyces cerevisiae microplate growth assay

Stock solutions of CR (SIGMA, PCode 102075330, 86.2% purified; $5mg.mL^{-1}$ in water) and CFW (ICN Biomedicals, cat.158067, 4 $mg.mL^{-1}$ in water clarified with 10N sodium hydroxide) were diluted in

YNB/MES to obtain concentrated 10X solutions (300 µg.mL^{-1} and 100 µg.mL^{-1} respectively). Overnight cultures of the tested strains (in triplicates) were collected at OD_{600} ranging from 0.9 to 2.5, centrifuged and resuspended in YNB/MES at a theoretical OD_{600} of 0.05 (OD_{600} =1 corresponds to 1.4 10^7 cells/mL). 20 µL of stressor was added to 180 µL of cells and growth was monitored in 96-well plates in a microplate reader (Epoch2-Biotech) under continuous slow orbital shaking. The turbidity of the cultures was measured at 600 nm every 15 minutes for 60 hours. It should be noted that it was not possible to test higher concentrations of CR than 50 µg.mL^{-1} as its precipitation prevented the reading. Data analysis was performed using GraphPad software (Prism 8.3.0). Triplicates were averaged and baselines (cell-free medium) were subtracted to obtain corrected growth curves. The maximum growth rate (µ max) was calculated by first transforming the baseline corrected data with y=ln(y). This new data was then plotted and a non-linear regression (line through the point (X0, y0)) was added to the linear region of the plotted graph. X0 and Y0 are set to "no constraints" and a linear region of the plotted graph was chosen visually. This region thus varies from strain to strain. The maximum growth rate (µ max) is the slope of this linear region.

Protein purification

pGEX-6P-3 derivatives expressing GST fusions with Knr4, Knr4$^{\Delta N \Delta C}$, Knr4$^{\Delta C}$ or Knr4$^{\Delta N \Delta L \Delta C}$ were introduced into *E. coli* BL21 (E. coli B; F$^-$ ompT gal dcm *l*on hsdS_B($r_B^-m_B^-$) [malB$^+$]$_{K\text{-}12}$ (λ^S)) for expression and purification. Bacteria were cultured at 37°C in LB supplemented with 150 µg/mL ampicillin. Expression was induced at mid-exponential phase with 0.1 mM IPTG and continued at 24°C for 20 hours. Bacteria from 500 mL cultures were pelleted and suspended in 30 mL of lysis buffer (10 mM Tris-HCl pH 8.0, 0, 1 mM EDTA, 150 mM NaCl, 5 mM DTT, 100 µg/mL lysozyme). Complete lysis was achieved by sonication. The lysate was incubated for 15 min in the presence of DNase I (100 U) and 5 mM $MgSO_4$. After the addition of Triton X-100 (2% final), the lysate was clarified by centrifugation (22,000 g). The clear lysate was incubated for 4h at 12°C with 1 mL of 6B glutathione sepharose beads (GE Healthcare) equilibrated in PBS buffer (pH 7.4, 11.9 mM phosphate, 137 mM NaCl, 2.7 mM KCl). The beads were washed with PBS, and then PreScission™ protease (80 U, GE Healthcare) was added

for overnight incubation in cleavage buffer (50 mM TRIS pH 7, 1 mM EDTA, 150 mM NaCl) and then used to pack an open column. Cleaved proteins were recovered in size exclusion chromatography buffer (100 mM MES, 300 mM NaCl, pH 6.0) and fractionated in the same buffer on SEC columns (HiLoad 16/60 Superdex, S75 or S200, Pharmacia). The peak protein fractions were collected and concentrated using ultracentrifugation devices (Vivaspin, Sartorius). The protein concentrations were evaluated by measuring their UV absorption at 280 nM and using theoretical molar extinction coefficients (Knr4, ε = 60,390 $M^{-1}.cm^{-1}$; Knr4$^{\Delta N \Delta C}$, ε = 48,930 $M^{-1}.cm^{-1}$; Knr4$^{\Delta C}$, ε =58,900 $M^{-1}.cm^{-1}$ Knr4$^{\Delta N \Delta L \Delta C}$, ε = 48,930 $M^{-1}.cm^{-1}$). When necessary, proteins were stored at -80°C after flash freezing in liquid nitrogen.

Protein crystallization

Previously determined crystallization conditions [39] repeatedly yielded crystals diffracting at best to 3.5 Å resolution, which nevertheless allowed to obtain Se-SAD phases of poor quality allowing to build a partial model. The new construct described herein allowed improved purification and identification of new crystallisation conditions. Initial crystallisation conditions were identified by screening 1400 conditions using Qiagen commercial kits (Classics, AmSO4, Anions, Cations, JCSG Core I-IV, PEGs, PEG II, MPD, pHClear, pHClear II). The sitting drops were done at 285 K with the help of a crystallisation robot (NanoDrop ExtY) by mixing 200 nL of native protein (10 to 15mg/mL) with the same volume of reservoir solution. The drops were automatically imaged with normal and UV light using an imaging robot (RockImager, RI-1000, Formulatrix). Optimisation of crystallisation conditions was performed manually using the hanging drop vapour diffusion method (2 µL) on siliconized glass slides and with 450 µL reservoirs. Crystals were obtained in the presence of PEG of average molecular weight comprised between 3,000 and 6,000, at a concentration of 15 to 24 % (w/v) at a pH range of 8.0–9.0 in 100 mM bicine buffer. Seleno-methionylated protein crystallised in similar conditions, but resulted in crystals diffracting to lower resolution.

Structure determination

Diffraction data were collected with crystals cooled in a gaseous nitrogen flux at 100 K after a brief immersion in the crystallisation solution supplemented with 20 % ethylene glycol (v/v). Native diffraction data were collected at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) on beamline ID23eh1 while anomalous data were collected at ALBA (Barcelona, Spain) on beamline XALOC. Collected intensities were processed using XDS and XSCALE [52]. SIRAS phasing and initial model building were performed with Phenix [53]. Subsequent model modifications and refinement were performed with Coot [54], Refmac [55], and the CCP4 suite of programs [56]. Data processing and refinement statistics are provided Table S4.

SAXS analysis: sample preparation and data acquisition

Small angle X-ray scattering data were collected at the BM29 beamline (ESRF, Grenoble, France) using a Pilatus 1M detector. The detector-distance of respectively 2.867 m covers a momentum transfer range of $0.04 < Q < 4.94$ nm^{-1}. SAXS experiments were carried out at 20°C with sample concentrations ranging from 1 to 10-15 $mg.mL^{-1}$ and sample volumes of 50 µL in a quartz glass capillary, using the automated sample changer. 2mM fresh DTT was added to each sample and centrifugation was performed just before data collection. The $Knr4^{\Delta C}$ protein, which was more sensitive to proteolysis, was investigated using the online FPLC size exclusion purification system, whereas the $Knr4^{\Delta N\Delta C}$ (SD) and Knr4 proteins were batch collected. Data from the corresponding buffer (last dialysis buffer in case of batch data collection) were used to provide a reference for the scattering background. Each measurement consisted of ten frames for which radiation damage was systematically investigated.

SAXS analysis: data processing

Data processing was performed with the ATSAS suite of programs [57]. The forward scattering I_0 and the radius of gyration R_g were calculated using the Guinier approximation using PRIMUS with $Q.R_g<1.1$, following particular recommendations for IDPs [58]. The calibration of the beamline was made with pure water and in such a way that I_0/c indicates the apparent molecular weight in kDa. In addition, the Rambo and Tainer concentration independent method [59] and the mass evaluation

based on the Porod volume, as implemented in SAXSMoW [60] have been applied. The distance distribution function *P(r)* and maximum dimensions D_{max} were calculated using GNOM, which also provided a R_g calculation based on the whole scattering curve. Dimensionless Kratky plot has been calculated following the recommendations of Durand *et al.* [38]. In the crystallographic structure 5J1B of construct Knr4$^{\Delta N \Delta C}$ (SD), some residues (80-83, 149-154, 189-217, 301-309) were missing, probably due to local disorder. They have been built on an atomic resolution using AllosMod-FoXS [61]. Fitting of the theoretical scattering curves computed from this completed crystallographic structure with the experimental data in solution of the same construct was done using CRYSOL. A molecular envelope of each construct was obtained from 10 independent *ab initio* simulations performed using DAMMIN and GASBOR [62] until a resolution of respectively 3.45 and 4.94 nm^{-1}, then averaged and filtered with DAMAVER [63]. More thorough analysis of the SAXS data was performed using the Ensemble Optimization Method (EOM, [64, 65]). A pool of 10,000 monomeric structures of each construct was constructed with the program RanCh9 by connecting the crystallographic structure of the SD domain with the 80-residues and 165-residues long N-terminal and C-terminal regions. The more compact option was chosen due to the partial flexibility of Knr4$^{\Delta C}$ and Knr4. The pool of structures was submitted to the genetic algorithm (GA) to select the minimal ensemble of conformations that best reproduce the data. Various sizes of the selected sub-ensembles were tested to check for stability as recommended in [66]. The results presented here are from automatic determination of the ensemble size.

Accession numbers

All data generated or analysed during this study are included in this published article, its supplementary information files and publicly available repositories. The SAXS data have been deposited on SASBDB database (https://www.sasbdb.org, [67]) under the draft identifiers SASDPC7, SASDPD7 and SASDPE7 (https://www.sasbdb.org/data/SASDPC7/a44f0aym3x, https://www.sasbdb.org/data/SASDPD7/gzfijr6z6l and https://www.sasbdb.org/data/SASDPE7/m35nignrim). The atomic coordinates for Knr4$^{\Delta N \Delta C}$ and

Knr4$^{\Delta N \Delta L \Delta C}$ have been deposited in the Protein Data Bank (RCSB PDB, https://www.rcbs.org, [68]), under the identifiers PDB ID: 5J1B (10.2210/pdb5J1B/pdb) and PDB ID: 8AJ2 respectively.

Acknowledgements

We thank the scientific staff at the European Synchrotron Radiation Facility (Grenoble, France) and ALBA (Barcelona, Spain) for the use of their excellent data collection facilities. The macromolecular crystallography equipment used in this study are part of the Integrated Screening Platform of Toulouse (PICT, IBiSA). We are grateful to the Bruce Futcher's team (Stony Brook School of Medicine) for the generous gift of plasmipGZ110. Funding was provided by the National Centre for Scientific Research (CNRS-France), the Université Paul Sabatier (Toulouse-France) and the Institut National des Sciences Appliquées de Toulouse (INSAT-Toulouse-France).

References

[1] Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN. Flexible nets. The roles of intrinsic disorder in protein interaction networks. FEBS J. 2005;272:5129-48.

[2] Oldfield CJ, Dunker KA. Intrinsically Disordered Proteins and Intrinsically Disordered Protein Regions. Annual Review of Biochemistry. 2014;83:553-84.

[3] Uversky VN. Protein intrinsic disorder and structure-function continuum. Prog Mol Biol Transl Sci. 2019;166:1-17.

[4] Fung HYJ, Birol M, Rhoades E. IDPs in macromolecular complexes: the roles of multivalent interactions in diverse assemblies. Curr Opin Struct Biol. 2018;49:36-43.

[5] Niklas KJ, Dunker AK, Yruela I. The evolutionary origins of cell type diversification and the role of intrinsically disordered proteins. J Exp Bot. 2018;69:1437-46.

[6] Yang J, Gao M, Xiong J, Su Z, Huang Y. Features of molecular recognition of intrinsically disordered proteins via coupled folding and binding. Protein Sci. 2019;28:1952-65.

[7] Dunker AK, Silman I, Uversky VN, Sussman JL. Function and structure of inherently disordered proteins. Curr Opin Struct Biol. 2008;18:756-64.

[8] Wright PE, Dyson JH. Intrinsically disordered proteins in cellular signalling and regulation. Nature reviews Molecular cell biology. 2015;16:18-29.

[9] Dosztányi Z, Chen J, Dunker KA, Simon I, Tompa P. Disorder and sequence repeats in hub proteins and their implications for network evolution. Journal of proteome research. 2006;5:2985-95.

[10] Haynes C, Oldfield CJ, Ji F, Klitgord N, Cusick ME, Radivojac P, et al. Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. PLoS computational biology. 2006;2.

[11] Martin-Yken H, Francois JM, Zerbib D. Knr4: A Disordered Hub Protein at the Heart of Fungal Cell Wall Signaling. Cellular Microbiology. 2016;18:1217-27.

[12] Wilson D, Madera M, Vogel C, Chothia C, Gough J. The SUPERFAMILY database in 2007: families and functions. Nucleic Acids Res. 2007;35:D308-13.

[13] Hong Z, Mann P, Brown NH, Tran LE, Shaw KJ, Hare RS, et al. Cloning and characterization of KNR4, a yeast gene involved in (1,3)-beta-glucan synthesis. Molecular and cellular biology. 1994;14:1017-25.

[14] Yamamoto T, Uchida K, Hiratani T, Miyazaki T, Yagiu J, Yamaguchi H. In vitro activity of the killer toxin from yeast Hansenula mrakii against yeasts and molds. The Journal of antibiotics. 1988;41:398-403.

[15] Martin H, Dagkessamanskaia A, Satchanska G, Dallies N, Francois J. KNR4, a suppressor of Saccharomyces cerevisiae cwh mutants, is involved in the transcriptional control of chitin synthase genes. Microbiology. 1999;145 (Pt 1):249-58.

[16] Fishel BR, Sperry AO, Garrard WT. Yeast calmodulin and a conserved nuclear protein participate in the in vivo binding of a matrix association region. Proceedings of the National Academy of Sciences of the United States of America. 1993;90:5623-7.

[17] Markovich S, Yekutiel A, Shalit I, Shadkchan Y, Osherov N. Genomic approach to identification of mutations affecting caspofungin susceptibility in Saccharomyces cerevisiae. Antimicrobial agents and chemotherapy. 2004;48:3871-6.

[18] Levin DE. Cell wall integrity signaling in Saccharomyces cerevisiae. Microbiology and molecular biology reviews : MMBR. 2005;69:262-91.

[19] Cyert MS. Calcineurin signaling in Saccharomyces cerevisiae: how yeast go crazy in response to stress. Biochemical and biophysical research communications. 2003;311:1143-50.

[20] Martin-Yken H, Dagkessamanskaia A, Basmaji F, Lagorce A, Francois J. The interaction of Slt2 MAP kinase with Knr4 is necessary for signalling through the cell wall integrity pathway in Saccharomyces cerevisiae. Molecular Microbiology. 2003;49:23-35.

[21] Heinisch JJ, Lorberg A, Schmitz HP, Jacoby JJ. The protein kinase C-mediated MAP kinase pathway involved in the maintenance of cellular integrity in Saccharomyces cerevisiae. Molecular microbiology. 1999;32:671-80.

[22] Zhao Y, Du J, Zhao G, Jiang L. Activation of calcineurin is mainly responsible for the calcium sensitivity of gene deletion mutations in the genome of budding yeast. Genomics. 2013;101:49-56.

[23] Durand F, Dagkessamanskaia A, Martin-Yken H, Graille M, Tilbeurgh H, Uversky VN, et al. Structure–function analysis of Knr4/Smi1, a newly member of intrinsically disordered proteins family, indispensable in the absence of a functional PKC1–SLT2 pathway in Saccharomyces cerevisiae. Yeast. 2008;25:563-76.

[24] Dagkessamanskaia A, Durand F, Uversky VN, Binda M, Lopez F, El Azzouzi K, et al. Functional dissection of an intrinsically disordered protein: understanding the roles of different domains of Knr4 protein in protein-protein interactions. Protein Sci. 2010;19:1376-85.

[25] Costanzo M, Baryshnikova A, Bellay J, Kim Y, Spear ED, Sevier CS, et al. The genetic landscape of a cell. Science. 2010;327:425-31.

[26] Goehring AS, Mitchell DA, Tong AH, Keniry ME, Boone C, Sprague GF. Synthetic lethal analysis implicates Ste20p, a p21-activated potein kinase, in polarisome activation. Molecular biology of the cell. 2003;14:1501-16.

[27] Lesage G, Shapiro J, Specht CA, Sdicu A-MM, Ménard P, Hussein S, et al. An interactional network of genes involved in chitin synthesis in Saccharomyces cerevisiae. BMC genetics. 2005;6:8.

[28] Basmaji F, Martin-Yken H, Durand F, Dagkessamanskaia A, Pichereaux C, Rossignol M, et al. The 'interactome' of the Knr4/Smi1, a protein implicated in coordinating cell wall synthesis with bud emergence in Saccharomyces cerevisiae. Molecular genetics and genomics : MGG. 2006;275:217-30.

[29] Tong AH, Lesage G, Bader GD, Ding H, Xu H, Xin X, et al. Global mapping of the yeast genetic interaction network. Science (New York, NY). 2004;303:808-13.

[30] Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, et al. A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. Nature. 2000;403:623-7.

[31] Nobile CJ, Johnson AD. Candida albicans Biofilms and Human Disease. Annual review of microbiology. 2015;69:71-92.

[32] Martin-Yken H, Bedekovic T, Brand AC, Richard ML, Znaidi S, d'Enfert C, et al. A conserved fungal hub protein involved in adhesion and drug resistance in the human pathogen Candida albicans. The Cell Surface. 2018;4:10-9.

[33] Uversky VN. Intrinsic disorder-based protein interactions and their modulators. Curr Pharm Des. 2013;19:4191-213.

[34] Yan J, Dunker AK, Uversky VN, Kurgan L. Molecular recognition features (MoRFs) in three domains of life. Mol Biosyst. 2016;12:697-710.

[35] Van Daele R, Spriet I, Wauters J, Maertens J, Mercier T, Van Hecke S, et al. Antifungal drugs: What brings the future? Med Mycol. 2019;57:S328-S43.

[36] Weinkam P, Pons J, Sali A. Structure-based model of allostery predicts coupling between distant sites. Proc Natl Acad Sci U S A. 2012;109:4875-80.

[37] Receveur-Brechot V, Durand D. How random are intrinsically disordered proteins? A small angle scattering perspective. Curr Protein Pept Sci. 2012;13:55-75.

[38] Durand D, Vives C, Cannella D, Perez J, Pebay-Peyroula E, Vachette P, et al. NADPH oxidase activator p67(phox) behaves in solution as a multidomain protein with semi-flexible linkers. J Struct Biol. 2010;169:45-53.

[39] Julien S, Tondl P, Durand F, Dagkessamanskaia A, van Tilbeurgh H, Francois JM, et al. Crystallographic studies of the structured core domain of Knr4 from Saccharomyces cerevisiae. Acta Crystallogr F Struct Biol Commun. 2015;71:1120-4.

[40] Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. J Mol Biol. 2007;372:774-97.

[41] Murzin AG, Brenner SE, Hubbard T, Chothia C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol. 1995;247:536-40.

[42] Holberger LE, Garza-Sanchez F, Lamoureux J, Low DA, Hayes CS. A novel family of toxin/antitoxin proteins in Bacillus species. FEBS Lett. 2012;586:132-6.

[43] Herth W. Calcofluor white and Congo red inhibit chitin microfibril assembly of Poterioochromonas: evidence for a gap between polymerization and microfibril formation. J Cell Biol. 1980;87:442-50.

[44] Pringle JR. Staining of bud scars and other cell wall chitin with calcofluor. Methods Enzymol. 1991;194:732-5.

[45] Ram AFJ, Klis FM. Identification of fungal cell wall mutants using susceptibility assays based on Calcofluor white and Congo red. Nature Protocols. 2006;1:2253-6.

[46] Roncero C, Durán A. Effect of Calcofluor white and Congo red on fungal cell wall morphogenesis: in vivo activation of chitin polymerization. Journal of Bacteriology. 1985;163:1180-5.

[47] González-Rubio G, Sastre-Vergara L, Molina M, Martín H, Fernández-Acero T. Substrates of the MAPK Slt2: Shaping Yeast Cell Integrity. Journal of Fungi. 2022;8:368.

[48] Martin-Yken H, Dagkessamanskaia A, Talibi D, Francois J. KNR4 is a member of the PKC1 signalling pathway and genetically interacts with BCK2, a gene involved in cell cycle progression in Saccharomyces cerevisiae. Current genetics. 2002;41:323-32.

[49] Hubbard SJ, Beynon RJ, Thornton JM. Assessment of conformational parameters as predictors of limited proteolytic sites in native protein structures. Protein Eng. 1998;11:349-59.

[50] Kim DH, Han KH. PreSMo Target-Binding Signatures in Intrinsically Disordered Proteins. Mol Cells. 2018;41:889-99.

[51] Bullock BN, Jochim AL, Arora PS. Assessing helical protein interfaces for inhibitor design. J Am Chem Soc. 2011;133:14220-3.

[52] Kabsch W. Xds. Acta Crystallogr D Biol Crystallogr. 2010;66:125-32.

[53] Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010;66:213-21.

[54] Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010;66:486-501.

[55] Vagin AA, Steiner RA, Lebedev AA, Potterton L, McNicholas S, Long F, et al. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. Acta Crystallogr D Biol Crystallogr. 2004;60:2184-95.

[56] Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr. 2011;67:235-42.

[57] Franke D, Petoukhov MV, Konarev PV, Panjkovich A, Tuukkanen A, Mertens HDT, et al. ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. J Appl Crystallogr. 2017;50:1212-25.

[58] Borgia A, Zheng W, Buholzer K, Borgia MB, Schuler A, Hofmann H, et al. Consistent View of Polypeptide Chain Expansion in Chemical Denaturants from Multiple Experimental Methods. J Am Chem Soc. 2016;138:11714-26.

[59] Rambo RP, Tainer JA. Accurate assessment of mass, models and resolution by small-angle scattering. Nature. 2013;496:477-81.

[60] Piiadov V, Ares de Araujo E, Oliveira Neto M, Craievich AF, Polikarpov I. SAXSMoW 2.0: Online calculator of the molecular weight of proteins in dilute solution from experimental SAXS data measured on a relative scale. Protein Sci. 2019;28:454-63.

[61] Schneidman-Duhovny D, Hammel M, Tainer JA, Sali A. FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. Nucleic Acids Res. 2016;44:W424-9.

[62] Svergun DI. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys J. 1999;76:2879-86.

[63] Volkov VV, Svergun DI. Uniqueness of ab initio shape determination in small-angle scattering. Journal of Applied Crystallography. 2003;36:860-4.

[64] Tria G, Mertens HD, Kachala M, Svergun DI. Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. IUCrJ. 2015;2:207-17.

[65] Bernadó P, Mylonas E, Petoukhov MV, Blackledge M, Svergun DI. Structural characterization of flexible proteins using small-angle X-ray scattering. Journal of the American Chemical Society. 2007;129:5656-64.

[66] Sagar A, Svergun D, Bernado P. Structural Analyses of Intrinsically Disordered Proteins by Small-Angle X-Ray Scattering. Methods Mol Biol. 2020;2141:249-69.

[67] Kikhney AG, Borges CR, Molodenskiy DS, Jeffries CM, Svergun DI. SASBDB: Towards an automatically curated and validated repository for biological scattering data. Protein Sci. 2020;29:66-75.

[68] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. Nucleic Acids Res. 2000;28:235-42.

[69] Skolnick J, Jaroszewski L, Kolinski A, Godzik A. Derivation and testing of pair potentials for protein folding. When is the quasichemical approximation correct? Protein Sci. 1997;6:676-88.

[70] Bernado P, Blackledge M. A self-consistent description of the conformational behavior of chemically denatured proteins from NMR and small angle scattering. Biophys J. 2009;97:2839-45.

Table 1. Theoretical and observed protein parameters

Protein		Knr4$^{\Delta C}$	Knr4$^{\Delta N\Delta C}$	Knr4
R_g (nm)	Expected [1,2]	2.0/5.4 ± 0.3	1.9/4.6 ± 0.3	2.3/6.5 ± 0.4
	Experimental [3,4]	3.2/3.2 ± 0.1	2.3/2.3 ± 0.1	4.8/5.0 ± 0.3
D_{max} (nm)		11.0 ± 0.3	8.0 ± 0.2	17.0 ± 1.0
Theoretical MW (kDa)		39.6	30.8	57.1
Experimental MW (kDa)	Guinier [5]	39.1 ± 0.7	27.5 ± 0.5	55.0 ± 1.0
	Rambo Tainer [6]	38.0 ± 0.8	25.8 ± 7.1	61.3 ± 0.3
	SAXSMoW [7]	43.8	32.8	55.8

[1,2] Expected radius of gyration (R_g) for globular proteins[1] or for full IDP[2] have been calculated respectively as described by [69] and [70].

[3,4] Experimental R_g were respectively obtained from Guinier[3] and P(r) analysis[4]. Theoretical molecular weight (MW) were calculated from the primary sequences.

[5,6,7] Experimental MW were obtained respectively from Guinier analysis[5], Rambo and Tainer concentration-independent method[6] and SAXSMoW calulator[7] [60].

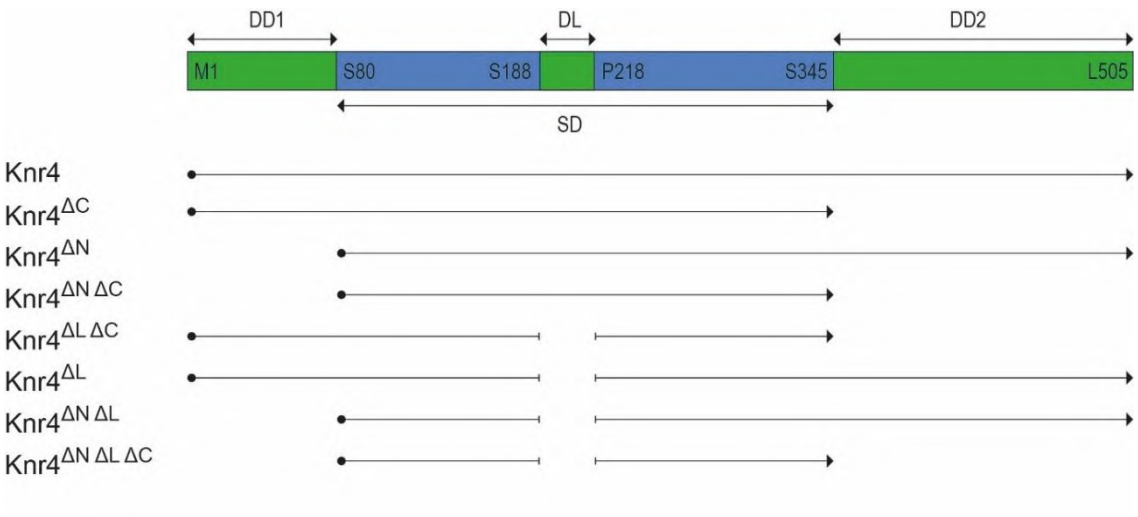


Figure 1. Schematic representation of the domain organisation of Knr4. Putative disordered domains (DD1, DL and DD2) are shown in green and ordered regions of SD are in blue. Residues at the domain boundaries are indicated. The different constructions used in the paper are depicted bellow with solid lines. N-terminus are represented by a close circle, C-terminus by an arrowhead and DL deletion boundaries by vertical bars.

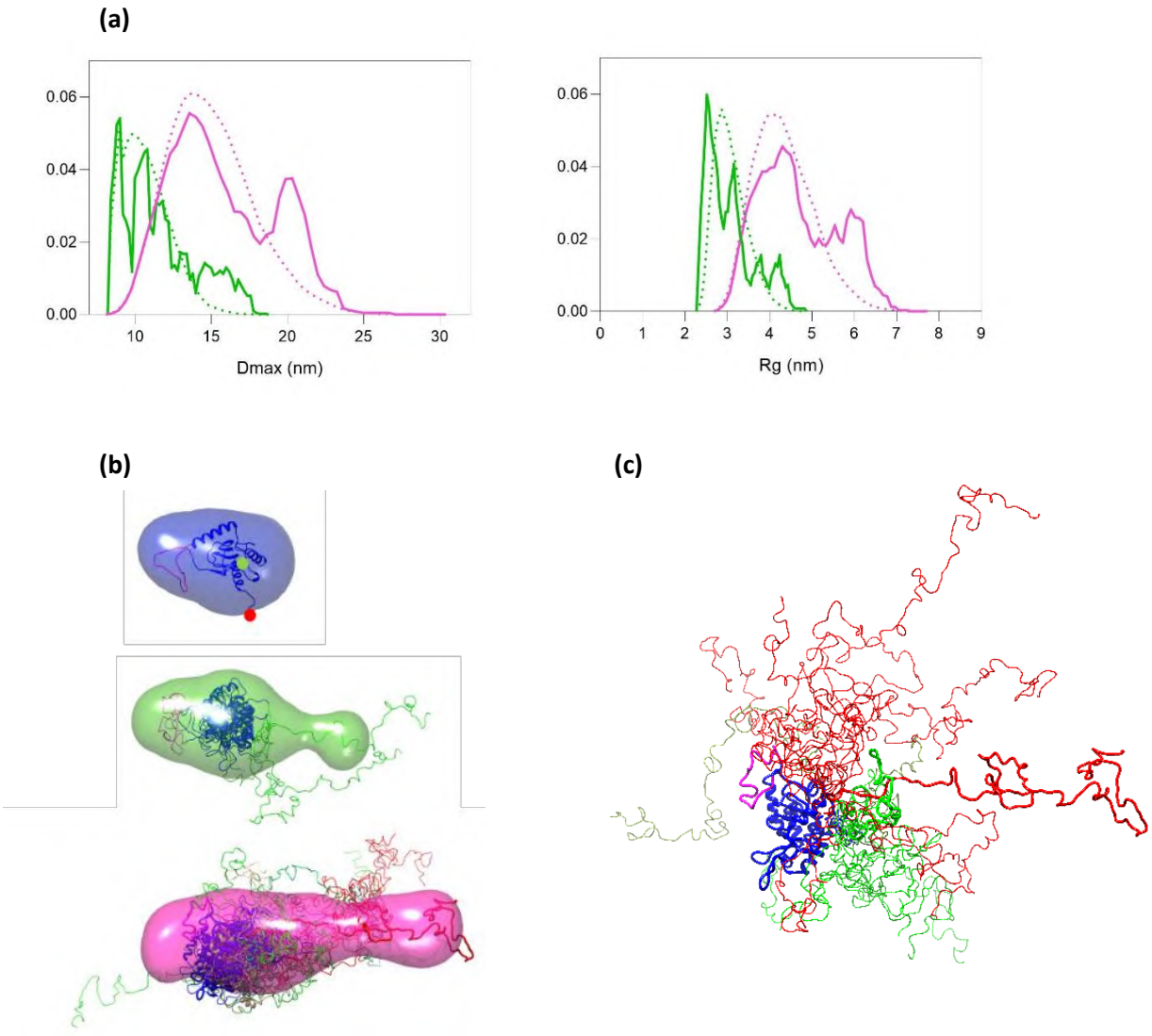


Figure 2. SAXS data analysis **(a).** Histograms of the distribution of R_g and D_{max} values of the Knr4$^{\Delta C}$ construct and full length Knr4 obtained from ensemble optimization modelling (EOM). **(b)**. Envelopes of the various constructs (blue: Knr4$^{\Delta N\Delta C}$, green: Knr4$^{\Delta C}$, red: Knr4) obtained at 3.45 nm^{-1}. The hybrid modelling of the full-length protein is shown in deep colour, with the Knr4$^{\Delta N\Delta C}$(SD) completed structure in blue (with N- and C-terminus residue highlighted in green and red respectively), DL loop in magenta and most probable conformations aligned with the envelope of the full-length protein. **(c).** Most probable conformations in full length protein aligned via their SD domain. The most highly populated conformation (21%) is depicted with a thick line

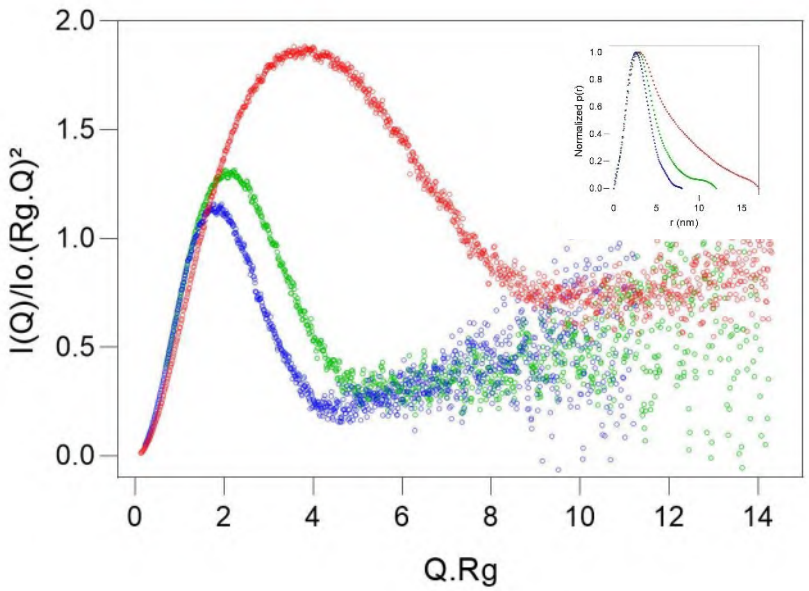


Figure 3. Dimensionless Kratky plots of the various constructs. Knr4$^{\Delta N \Delta C}$ is in blue, Knr4$^{\Delta C}$ in green and Knr4 in red. The corresponding normalized distribution of intramolecular distances are shown in the inset.

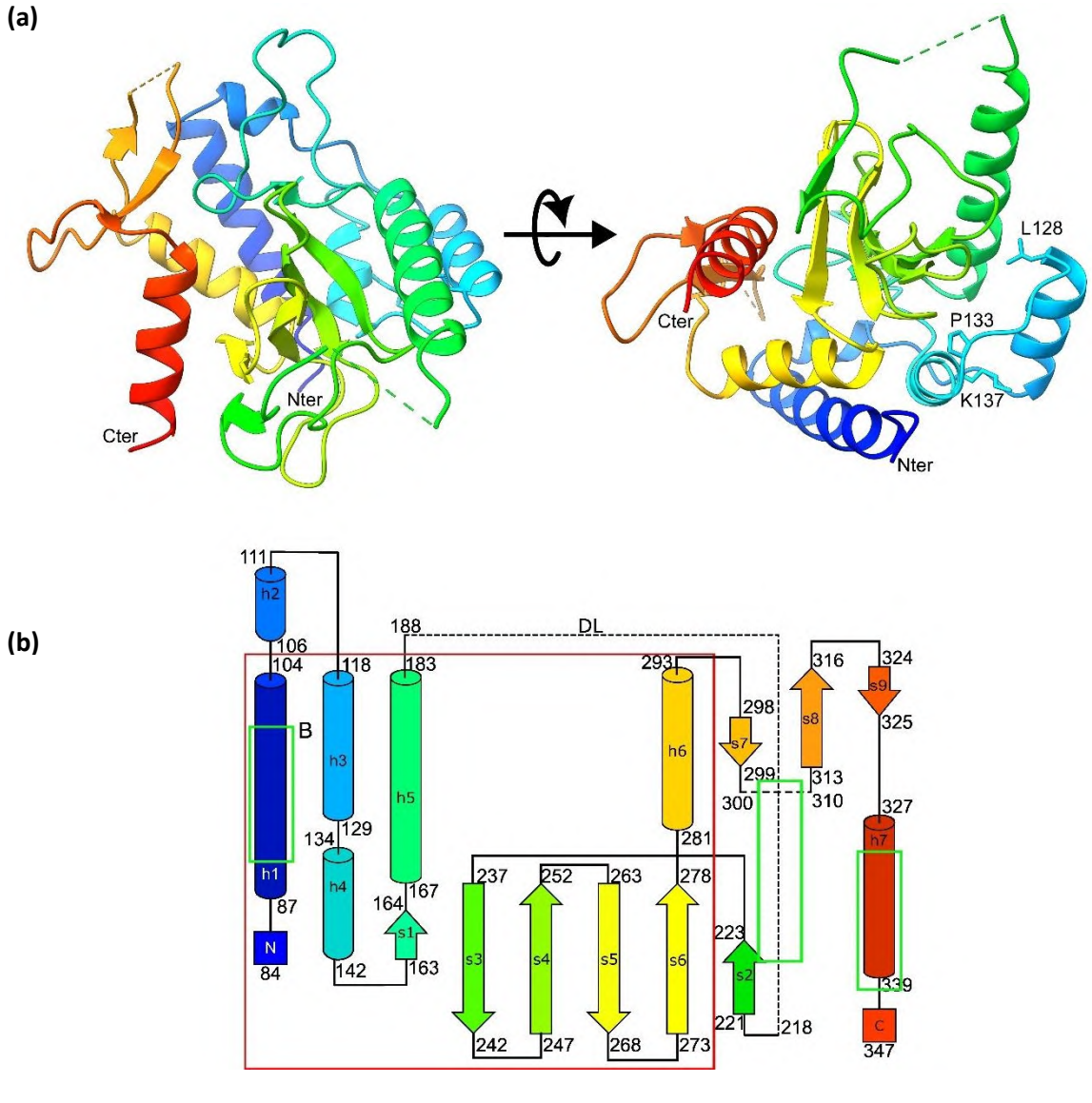


Figure 4. **(a)** schematic representation of the structure of the ordered core of Knr4 SD. The structure is represented as a rainbow-colored ribbon, starting from blue for the N-terminal and ending in red for the C-terminal. Stretches of disordered residues with no associated electron density are represented with a discontinuous line. Conserved residues are depicted as sticks and labelled. **(b).** Topology of Knr4. α-helices are symbolized with cylinders and β-strands with arrows. Boundaries of secondary structure elements and of disordered parts of the structure are indicated. Helices are named h1 to h7 and strands s1 to s9. Colour coded is the same as in A. Red box indicates the fold common to the structure of the Smi1/Knr4-like fold. Identified MORFs are indicated with a green box.

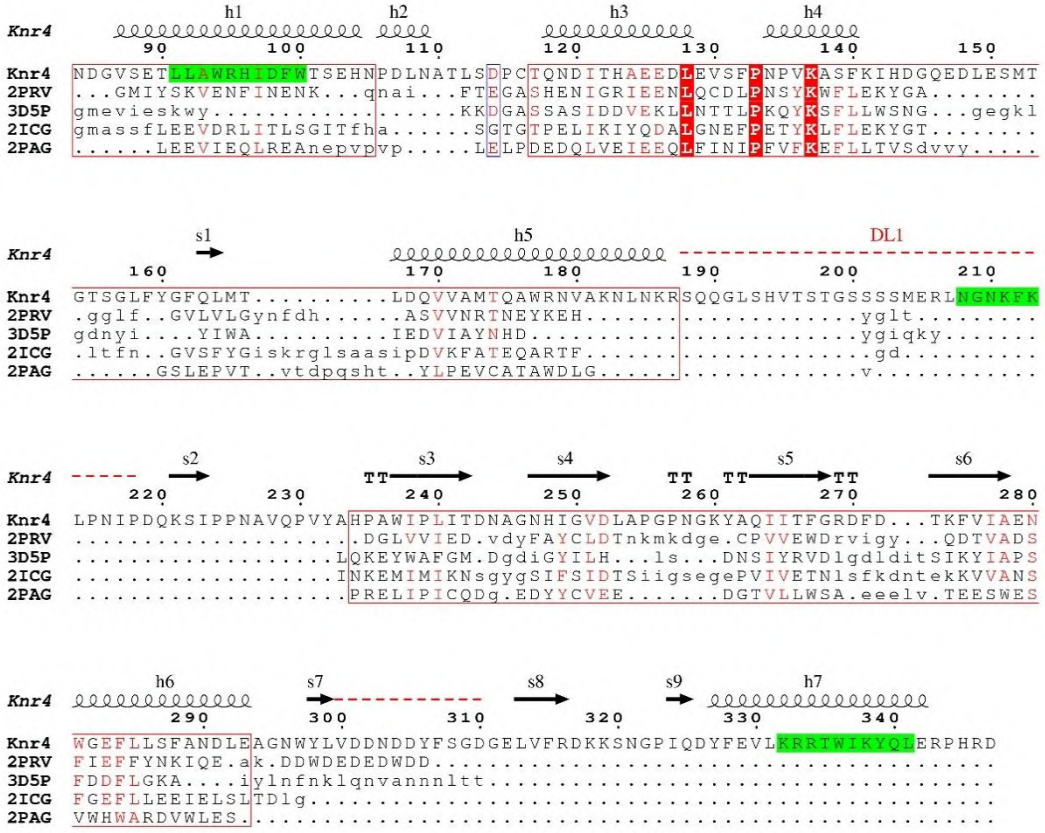


Figure 5. structure-based sequence alignment of proteins identified as displaying the Smi1/Knr4 fold in the SCOP database. Knr4 secondary structure elements are displayed above the alignment. Strictly conserved residues are indicated with white letters on red background. Conserved residues are indicated in red. Lower case letters indicate regions where the structures do not align with Knr4. Disordered loops in Knr4 structures are indicated with a dashed red line in the secondary structure scheme. The red frames highlight the structurally conserved regions in the five structures. Identified MORFs are indicated with a green background.

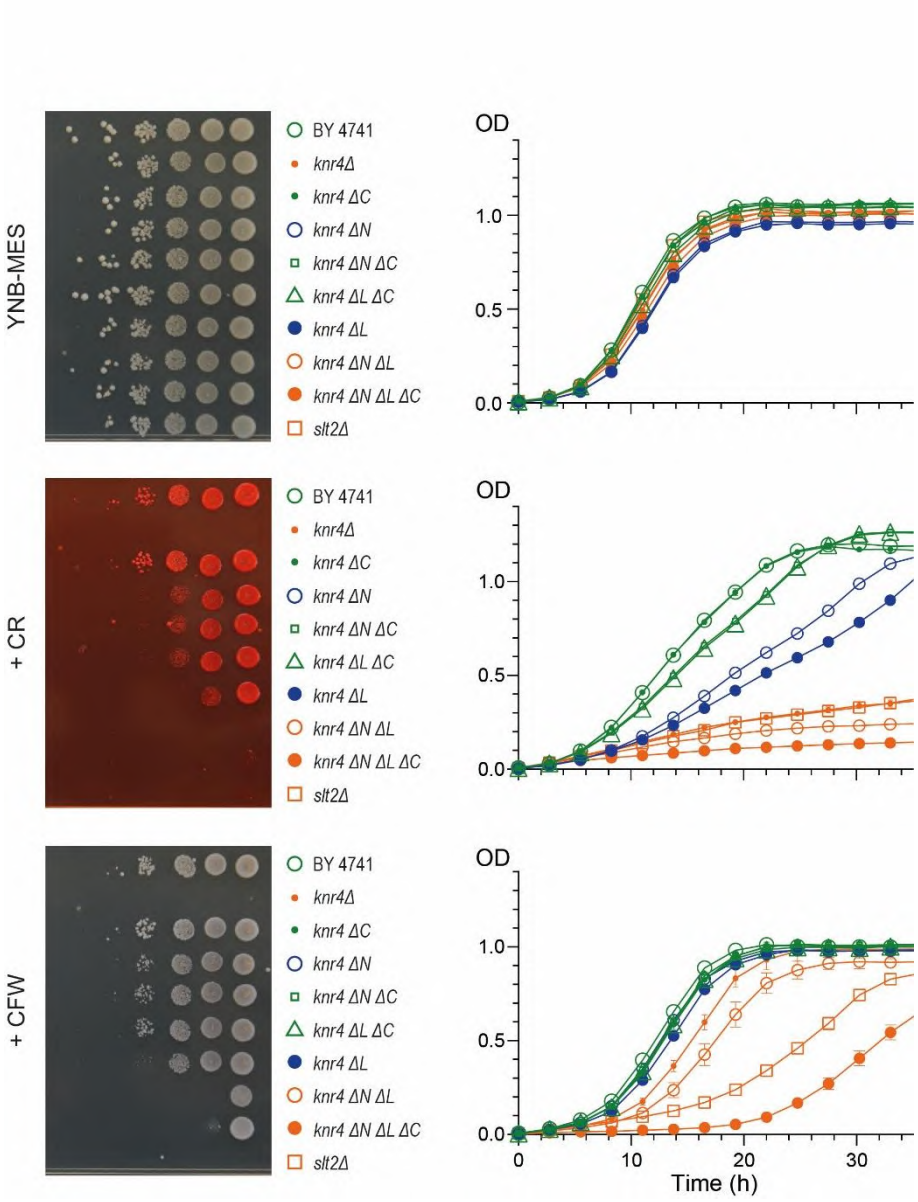


Figure 6. Effect of CR and CWF on growth of Knr4 mutant strains. **(Left column).** Growth on YNB-MES plates of serial dilutions (10X) of wild-type BY4741 (WT) and its isogenic deletion mutants of either a full-length gene (*knr4Δ* and *slt2Δ*) or specified domains of knr4 (knr4$^{\Delta C}$, knr4$^{\Delta N}$, Knr4$^{\Delta L}$, knr4$^{\Delta N\Delta C}$, knr4$^{\Delta L\Delta C}$, knr4$^{\Delta N\Delta L}$ and knr4$^{\Delta N\Delta L\Delta C}$) in the presence of Congo red (CR, 50 µg.mL^{-1}) or calcofluor white (CFW, 30 µg.mL^{-1}). **(Right column).** Growth curves of wild-type BY4741 (WT) and its isogenic deletion mutants of either a full-length gene (*knr4Δ* and *slt2Δ*) or specified domains of knr4 in absence or presence of CR (30 µg.mL^{-1}) or CFW (30 µg.mL^{-1}). Symbol legend is indicated on the figure.

Supplemental data for :

The conserved yeast protein Knr4 involved in cell wall integrity is a multi-domain intrinsically disordered protein

Manon Batista [1,2,†], Ellen I.M. Donker [1,2,†], Cécile Bon [2,†], Myriam Guillien [1,2], Adriana Caisso [1], Lionel Mourey [2], Jean Marie François [1], Laurent Maveyraud [2,#] and Didier Zerbib [1,#]

Supplementary Table 1. Essential details about SAXS samples

Protein	**Knr4$^{\Delta C}$**	**Knr4$^{\Delta N \Delta C}$ (SD)**	**Knr4**
Organism	*Saccharomyces cerevisiae*		
Source	Recombinant (*E. coli*)		
UniProt ID	P32566	P32566	P32566
Construct boundaries	1-345	80-340	1-505
N-Terminal Tag	GST, cleaved	GST, cleaved	GST, cleaved
MW (kDa)	39.6	30.9	57.5
A_{280nm} 0.1 % (w/v)	1.586	1.488	1.050
$\bar{\nu}$ ($cm^3\ g^{-1}$)	0.724	0.728	0.720
ρ_M, ρ_S, $\Delta\bar{\rho}$ (all in $10^{10}\ cm^{-2}$)	12.475, 9.542, 2.933	12.420, 9.542, 2.879	12.547, 9.542, 3.006
	SEC-SAXS	**Batch-SAXS**	
Column	Superdex 200 5/150 GL	-	-
C_{load}, V_{inject}	5.5 $mg.mL^{-1}$, 45 µL	-	-
Flow rate	0.3 $mL.min^{-1}$	-	-
Concentration range	-	1.0-9.4 $mg.mL^{-1}$	1.1-9.0 $mg.mL^{-1}$
Solvent source	SEC flow-through prior to elution of protein	Last-step dialysis	Last-step dialysis

Supplementary Table 2. Essential details about SAXS data collection parameters acquisition

Protein	**Knr4$^{\Delta C}$**	**Knr4$^{\Delta N\Delta C}$ (SD)**	**Knr4**
Source	BM29 ESRF	BM29 ESRF	BM29 ESRF
Detector	Pilatus	Pilatus	Pilatus
Wavelength (Å)	0.9919	0.9919	0.9919
Distance sample-detector (m)	2.867	2.872	2.872
Q-range (nm^{-1})	0.051-4.945	0.051-4.945	0.051-4.945
Temperature (°C)	20.0	20.0	20.0
Exposure time (s)	1.5	1	1
Number of frames	Continuous data measurements of SEC elution	10	10
Monitoring of radiation damage	Frame-by-frame comparison	Frame-by-frame comparison	Frame-by-frame comparison
Absolute scaling method	Water calibration	Water calibration	Water calibration

Supplementary Table 3. Essential details about SAXS data reduction, analysis and interpretation

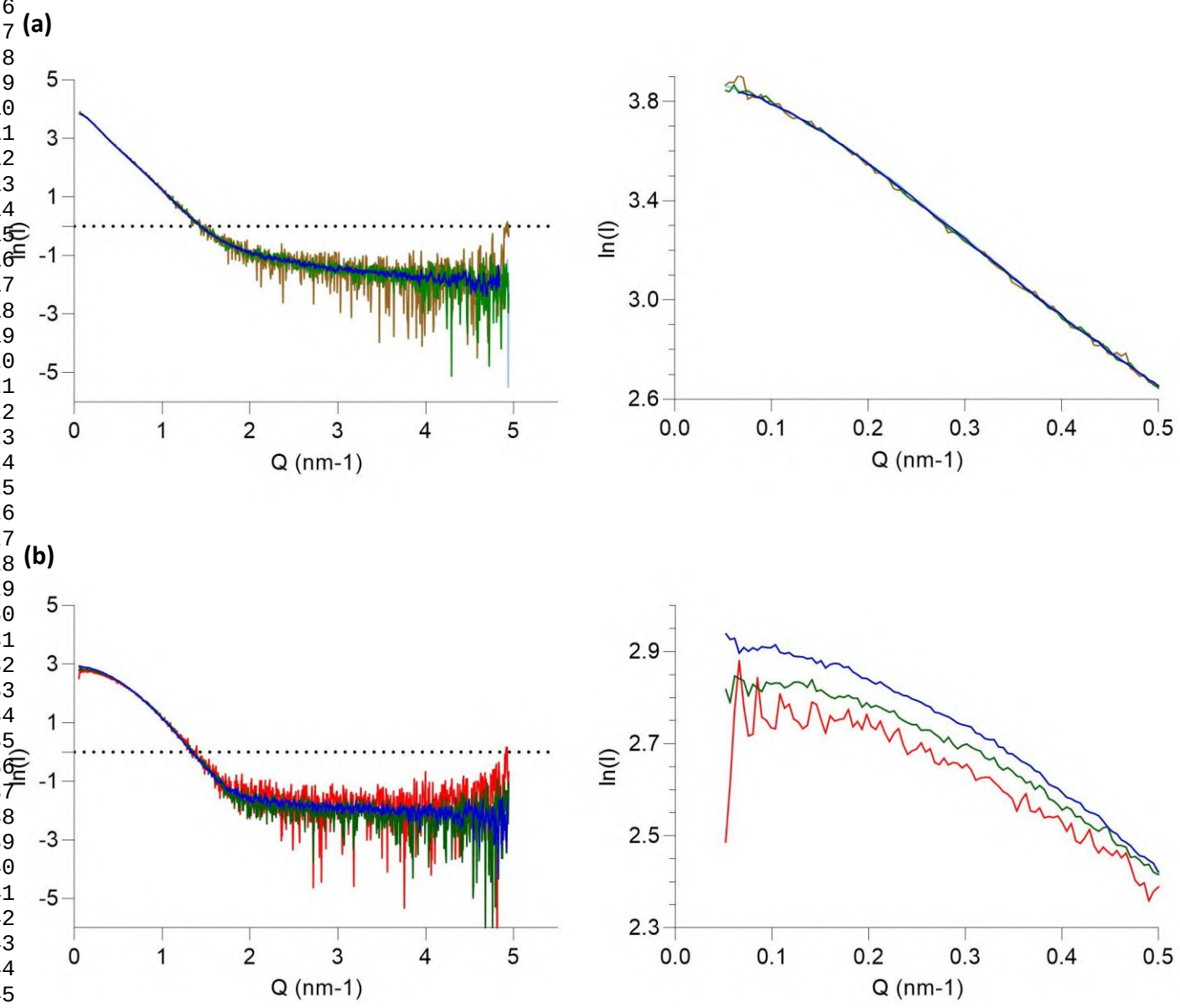
Protein	$Knr4^{\Delta C}$	$Knr4^{\Delta N\Delta C}$ (SD)	Knr4
Data reduction Software	Primusqt	Primusqt	Primusqt
Mean relative errors until 3 nm^{-1} (%)	3.3	15.0	10.1
	Basic analyses		
Guinier Fidelity[1] (%)	82	95	100
R_g (nm)	2.81 ± 0.10 nm	2.29 ± 0.01 nm	4.99 ± 0.19 nm
I_o	22.94 ± 0.05	27.38 ± 0.03	48.08 ± 0.06
P(r) total estimate	0.738	0.887	0.476
R_g, Dmax (nm)	3.17 ± 0.02, 11.0 ± 0.3	2.31 ± 0.01, 8.0 ± 0.2	5.03 ± 0.01, 18.0 ± 1.0
I_o	31.74 ± 0.01	27.43 ± 0.01	47.61 ± 0.05
	Shape modelling DAMMIN		
χ, Normalised spatial discrepancy	1.46, 0.553 ± 0.010	0.90, 0.485 ± 0.016	1.07, 0.625± 0.008
DAM volume (10^3 nm^3)	88.49	69.01	142.7
CorMap p-value	0.03	0.37	0.08
	Shape modelling GASBOR		
χ		0.92	
Total excluded DRM volume (10^3 nm^3)		46.75	
CorMap p-value		0.49	
	Conformational states, EOM		
χ	0.93	0.95	0.88
R_g (nm), D_{max} (nm), % of population	2.5 9.1 9% 2.5 9.4 36% 3.1 12.3 18% 3.3 11.4 9% 3.4 11.7 9% 3.7 15.2 9% 4.7 18.6 9% Final ensemble: 3.1 11.6	2.3 8.6 100%	3.5 10.6 7% 3.6 12.0 14% 3.9 15.0 7% 4.1 14.3 7% 4.2 13.0 14% 4.2 13.7 7% 4.5 17.3 7% 4.7 16.3 7% 5.7 18.2 21% 7.1 25.8 7% Final ensemble: 4.6 15.6

Supplementary Table 4. Crystallographic data processing and refinement statistics

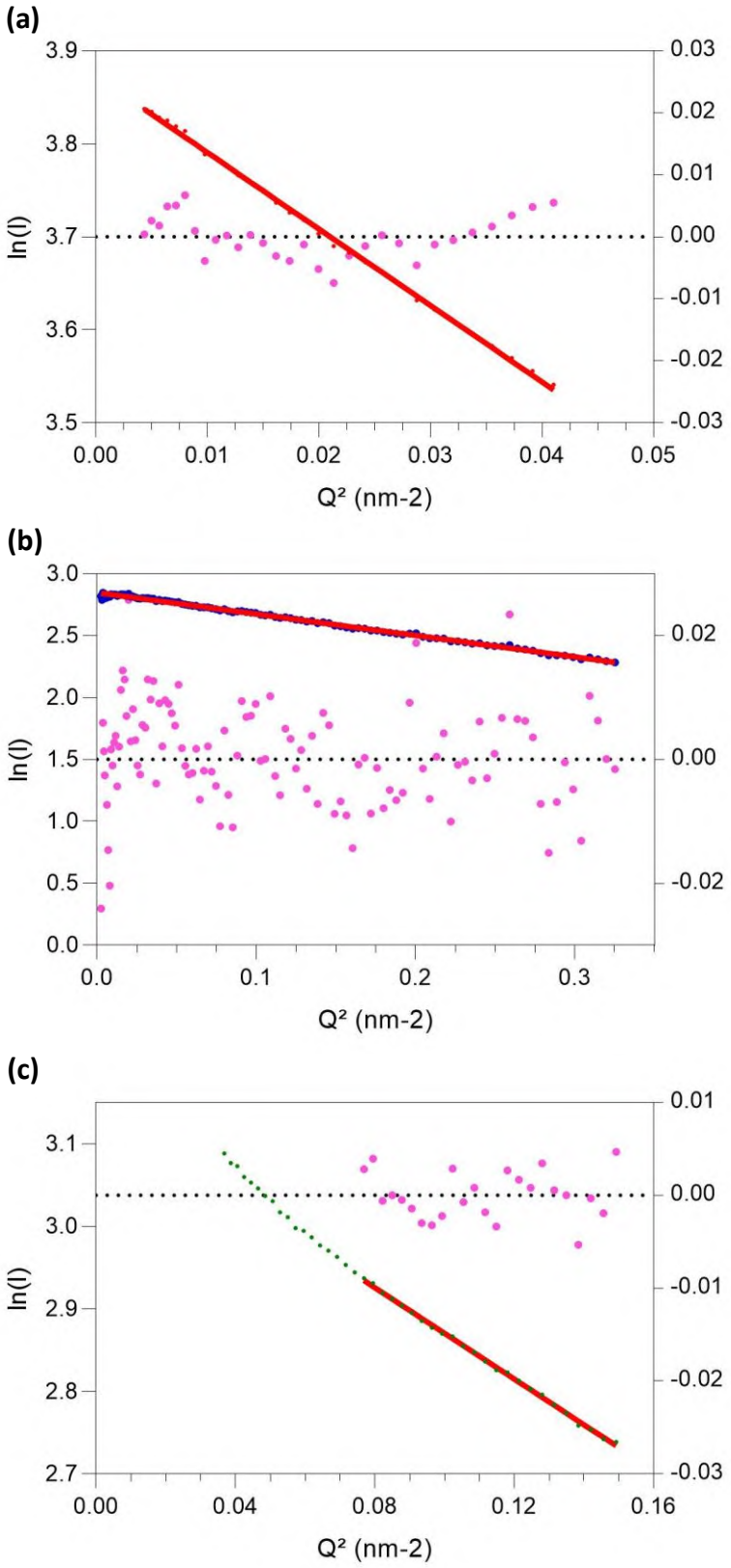
Protein	Knr4$^{\Delta N\Delta C}$-Se	Knr4$^{\Delta N\Delta C}$	Knr4$^{\Delta N\Delta L\Delta C}$
PDB ID		5J1B	8AJ2
Beamline	ALBA, xaloc	ESRF, ID23eh1	ESRF, ID23eh1
wavelength (Å)	0.97949	0.972	0.972
Spacegroup	$P6_2$	$P6_2$	$P6_2$
Cell parameters (Å)		a=b=103.00, c=93.38	a=b=103.14, c=93.68
Resolution (Å)	3.20 (3.25 – 3.20)	2.50 (2.65 – 2.50)	2.20 (2.33-2.20)
Number observations	190,462 (7,343)	49,956 (7,939)	324,810 (52,494)
Number unique	18,522 (853)	19,048 (3,051)	28,788 (4,593)
Multiplicity	10.3 (8,6)	2.6 (2.6)	11.3 (11.4)
Completeness (%)	99.9 (97.4)	97.5 (97.5)	99.9 (99.7)
Rsym	0.174 (1.570)	0.040 (0.908)	0.071 (1.142)
Rmeas	0.183 (1.667)	0.054 (1.139)	0.075 (1.196)
CC1/2	99.7 (42.5)	99.9 (38.4)	99.9 (82.3)
Anomalous correlation	40	-	-
SigAno	1.265	-	-
<I/s>	12.98 (1.41)	12.7 (1.0)	18.7 (2.5)
Resolution		35.00 – 2.50	
Number of reflections		18,066	
Rfactor/ Rfree		0.18486/0.22614	
Nb atoms		3,417	
Rms bond length (Å)		0.008	
Rms bond angle (°)		1.132	

Supplementary Table 5. Oligonucleotides used in this study

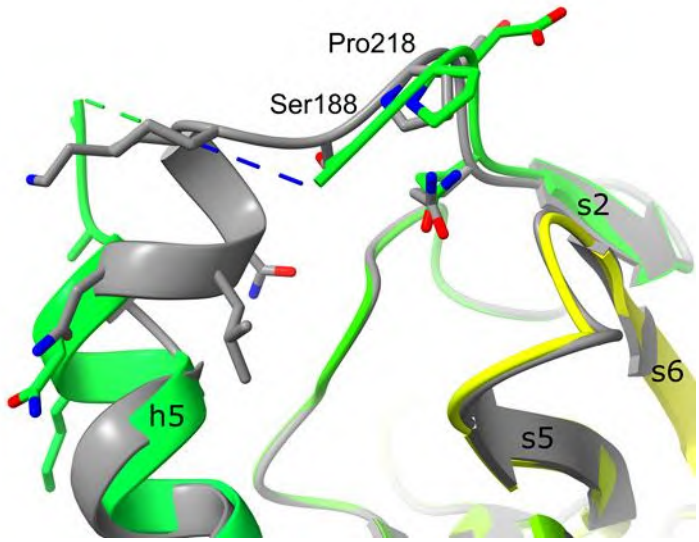
Primer	Sequence (5' --> 3')	Purpose
***KNR4*-Mut-Top**	CAAGAAAACTTGAGATCTCAATAATAAAAA TCTCAACCTG	pGEX-6P-3:: Knr4$^{\Delta C}$ construction
***KNR4*-Mut-Bot**	CAGGTTGAGATTTTTATTATTGAGATCTCAA GTTTTCTTG	pGEX-6P-3:: Knr4$^{\Delta C}$ construction
***KNR4*-Fwd**	TACCATGGGGATCCACTAGATTTTCATCCCT	pZE13::Pro-*KNR4*-Ter construction
***KNR4*-Rev**	TAAGCTTCGCGGTGGCGGC	pZE13::Pro-*KNR4*-Ter construction
pZE13-Fwd	GTGGATCCCCATGGTACGCGTGC	pZE13::Pro-*KNR4*-Ter construction
pZE13-Rev	CCACCGCGAAGCTTATCGATACCGTCGACC	pZE13::Pro-*KNR4*-Ter construction
***KNR4*-Rep-Fwd**	CAAGCCCTAAAGCACGTGAC	Repair Fragment
***KNR4*-Rep-Rev**	CTTCGTAGTGGCCTCAAACC	Repair Fragment
ΔN-Fwd	TCCACGGAGTCAAACGATG	DD1 deletion in pZE13::Pro-*KNR4*-Ter
ΔN-Rev	CATTTTATACTAAAAAATTCTGCCAAGTTG	DD1 deletion in pZE13::Pro-*KNR4*-Ter
ΔL-Fwd	CCAGATCAAAAATCTATTCCTCCAAATG	DL deletion in pZE13::Pro-*KNR4*-Ter
ΔL-Rev	AGATCTTTTGTTTAGGTTCTTTGCG	DL deletion in pZE13::Pro-*KNR4*-Ter
ΔC-Fwd	TGAAATATCACAATTAACATTCTACAACC	DD2 deletion in pZE13::Pro-*KNR4*-Ter
ΔC-Rev	TTGTGATCTCAAGTTTTCTTGATACTTG	DD2 deletion in pZE13::Pro-*KNR4*-Ter



Supplementary Figure 1. Analysis of concentration dependence for SAXS data collected in batch mode. Left, SAXS intensity normalised by concentration. Right, zoom at low Q-value. (a) protein Knr4 at 1.1 mg.mL^{-1} (brown), 2.6 mg.mL^{-1} (green), 5.8 mg.mL^{-1} (dark blue), 9.0 mg.mL^{-1} (light blue). (b) protein Knr4$^{\Delta N \Delta C}$ (SD) at 0.97 mg.mL^{-1} (red), 2.1 mg.mL^{-1} (dark green), 5.7 mg.mL^{-1} (light green), 8.08 mg .mL^{-1} (dark blue).



Supplementary Figure 2. Guinier analyses of constructs. **(a).** Knr4. **(b).** Knr4$^{\Delta C}$. **(c).** Knr4$^{\Delta N \Delta C}$.



Supplementary Figure 3. Superimposition of the structures of Knr4$^{\Delta N\ \Delta C}$ (same colour as in Figure 2) and Knr4$^{\Delta N\ \Delta L\ \Delta C}$ (in grey). Residues Ser188 and Pro218, that borders the deleted region in Knr4$^{\Delta N\ \Delta L\ \Delta C}$ are labelled, as are the secondary structure elements.

Table 1. Theoretical and observed protein parameters

Protein		**Knr4$^{\Delta C}$**	**Knr4$^{\Delta N\Delta C}$**	**Knr4**
R_g (nm)	Expected [1,2]	2.0/5.4 ± 0.3	1.9/4.6 ± 0.3	2.3/6.5 ± 0.4
	Experimental [3,4]	3.2/3.2 ± 0.1	2.3/2.3 ± 0.1	4.8/5.0 ± 0.3
D_{max} (nm)		11.0 ± 0.3	8.0 ± 0.2	17.0 ± 1.0
Theoretical MW (kDa)		39.6	30.8	57.1
Experimental MW (kDa)	Guinier [5]	39.1 ± 0.7	27.5 ± 0.5	55.0 ± 1.0
	Rambo Tainer [6]	38.0 ± 0.8	25.8 ± 7.1	61.3 ± 0.3
	SAXSMoW [7]	43.8	32.8	55.8

[1,2] Expected radius of gyration (R_g) for globular proteins[1] or for full IDP[2] have been calculated respectively as described by {Skolnick, 1997 #1294} and {Bernado, 2009 #2441}.

[3,4] Experimental R_g were respectively obtained from Guinier[3] and P(r) analysis[4]. Theoretical molecular weight (MW) were calculated from the primary sequences.

[5,6,7] Experimental MW were obtained respectively from Guinier analysis[5], Rambo and Tainer concentration-independent method[6] and SAXSMoW calulator[7] {Piiadov, 2019 #2445}.

Supplementary Table 1. Essential details about SAXS samples

Protein	**Knr4$^{\Delta C}$**	**Knr4$^{\Delta N\Delta C}$ (SD)**	**Knr4**
Organism	*Saccharomyces cerevisiae*		
Source	Recombinant (*E. coli*)		
UniProt ID	P32566	P32566	P32566
Construct boundaries	1-345	80-340	1-505
N-Terminal Tag	GST, cleaved	GST, cleaved	GST, cleaved
MW (kDa)	39.6	30.9	57.5
A_{280nm} 0.1 % (w/v)	1.586	1.488	1.050
$\bar{\nu}$ ($cm^3\ g^{-1}$)	0.724	0.728	0.720
ρ_M, ρ_S, $\Delta\bar{\rho}$ (all in $10^{10}\ cm^{-2}$)	12.475, 9.542, 2.933	12.420, 9.542, 2.879	12.547, 9.542, 3.006
	SEC-SAXS	**Batch-SAXS**	
Column	Superdex 200 5/150 GL	-	-
C_{load}, V_{inject}	5.5 $mg.mL^{-1}$, 45 µL	-	-
Flow rate	0.3 $mL.min^{-1}$	-	-
Concentration range	-	1.0-9.4 $mg.mL^{-1}$	1.1-9.0 $mg.mL^{-1}$
Solvent source	SEC flow-through prior to elution of protein	Last-step dialysis	Last-step dialysis

Supplementary Table 2. Essential details about SAXS data collection parameters acquisition

Protein	**Knr4$^{\Delta C}$**	**Knr4$^{\Delta N\Delta C}$ (SD)**	**Knr4**
Source	BM29 ESRF	BM29 ESRF	BM29 ESRF
Detector	Pilatus	Pilatus	Pilatus
Wavelength (Å)	0.9919	0.9919	0.9919
Distance sample-detector (m)	2.867	2.872	2.872
Q-range (nm^{-1})	0.051-4.945	0.051-4.945	0.051-4.945
Temperature (°C)	20.0	20.0	20.0
Exposure time (s)	1.5	1	1
Number of frames	Continuous data measurements of SEC elution	10	10
Monitoring of radiation damage	Frame-by-frame comparison	Frame-by-frame comparison	Frame-by-frame comparison
Absolute scaling method	Water calibration	Water calibration	Water calibration

Supplementary Table 3. Essential details about SAXS data reduction, analysis and interpretation

Protein	$Knr4^{\Delta C}$	$Knr4^{\Delta N\Delta C}$ (SD)	Knr4
Data reduction Software	Primusqt	Primusqt	Primusqt
Mean relative errors until 3 nm^{-1} (%)	3.3	15.0	10.1
	Basic analyses		
Guinier Fidelity[1] (%)	82	95	100
R_g (nm)	2.81 ± 0.10 nm	2.29 ± 0.01 nm	4.99 ± 0.19 nm
I_o	22.94 ± 0.05	27.38 ± 0.03	48.08 ± 0.06
P(r) total estimate	0.738	0.887	0.476
R_g, Dmax (nm)	3.17 ± 0.02, 11.0 ± 0.3	2.31 ± 0.01, 8.0 ± 0.2	5.03 ± 0.01, 18.0 ± 1.0
I_o	31.74 ± 0.01	27.43 ± 0.01	47.61 ± 0.05
	Shape modelling DAMMIN		
χ, Normalised spatial discrepancy	1.46, 0.553 ± 0.010	0.90, 0.485 ± 0.016	1.07, 0.625± 0.008
DAM volume (10^3 nm^3)	88.49	69.01	142.7
CorMap p-value	0.03	0.37	0.08
	Shape modelling GASBOR		
χ		0.92	
Total excluded DRM volume (10^3 nm^3)		46.75	
CorMap p-value		0.49	
	Conformational states, EOM		
χ	0.93	0.95	0.88
R_g (nm), D_{max} (nm), % of population	2.5 9.1 9% 2.5 9.4 36% 3.1 12.3 18% 3.3 11.4 9% 3.4 11.7 9% 3.7 15.2 9% 4.7 18.6 9% Final ensemble: 3.1 11.6	2.3 8.6 100%	3.5 10.6 7% 3.6 12.0 14% 3.9 15.0 7% 4.1 14.3 7% 4.2 13.0 14% 4.2 13.7 7% 4.5 17.3 7% 4.7 16.3 7% 5.7 18.2 21% 7.1 25.8 7% Final ensemble: 4.6 15.6

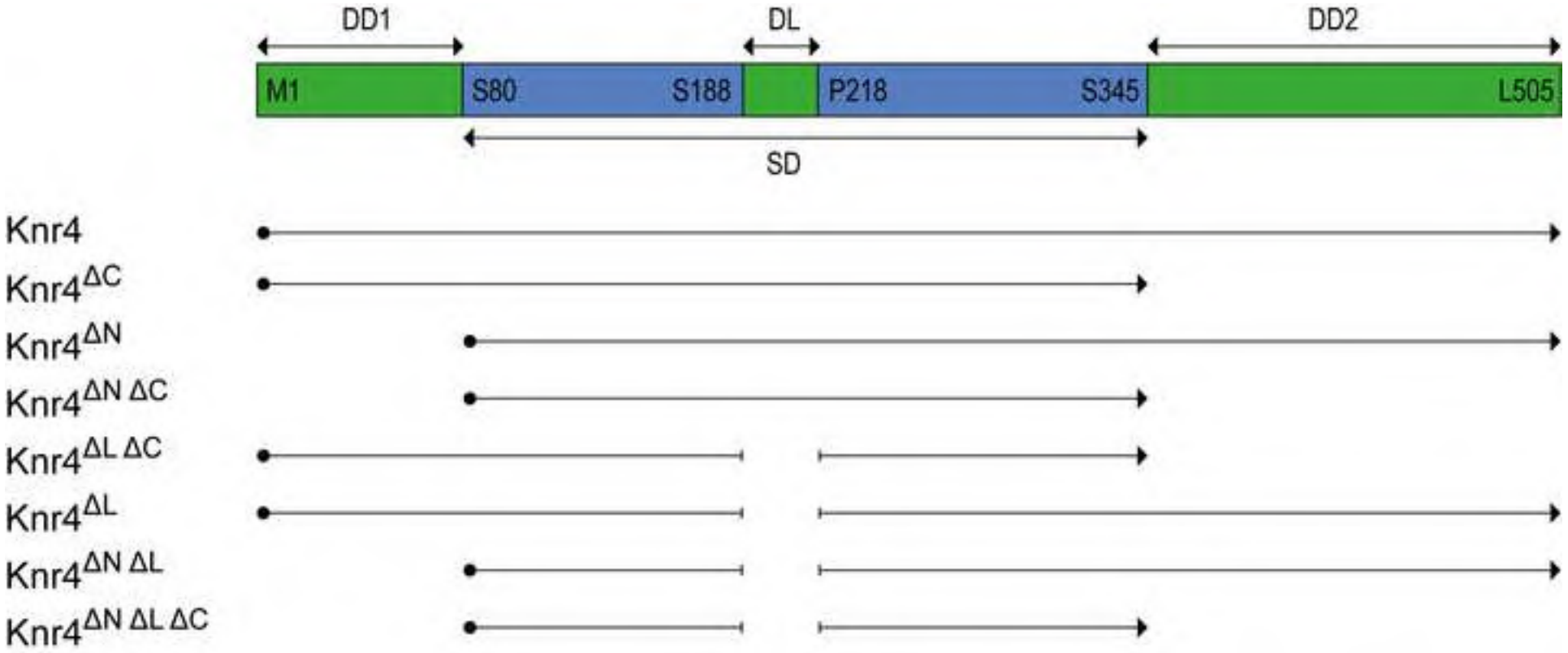
Supplementary Table 4. Crystallographic data processing and refinement statistics

Protein	**Knr4$^{\Delta N\Delta C}$-Se**	**Knr4$^{\Delta N\Delta C}$**	**Knr4$^{\Delta N\Delta L\Delta C}$**
PDB ID		5J1B	8AJ2
Beamline	ALBA, xaloc	ESRF, ID23eh1	ESRF, ID23eh1
wavelength (Å)	0.97949	0.972	0.972
Spacegroup	*P*6$_2$	*P*6$_2$	*P*6$_2$
Cell parameters (Å)		a=b=103.00, c=93.38	a=b=103.14, c=93.68
Resolution (Å)	3.20 (3.25 – 3.20)	2.50 (2.65 – 2.50)	2.20 (2.33-2.20)
Number observations	190,462 (7,343)	49,956 (7,939)	324,810 (52,494)
Number unique	18,522 (853)	19,048 (3,051)	28,788 (4,593)
Multiplicity	10.3 (8,6)	2.6 (2.6)	11.3 (11.4)
Completeness (%)	99.9 (97.4)	97.5 (97.5)	99.9 (99.7)
Rsym	0.174 (1.570)	0.040 (0.908)	0.071 (1.142)
Rmeas	0.183 (1.667)	0.054 (1.139)	0.075 (1.196)
CC1/2	99.7 (42.5)	99.9 (38.4)	99.9 (82.3)
Anomalous correlation	40	-	-
SigAno	1.265	-	-
<I/s>	12.98 (1.41)	12.7 (1.0)	18.7 (2.5)
Resolution		35.00 – 2.50	
Number of reflections		18,066	
Rfactor/ Rfree		0.18486/0.22614	
Nb atoms		3,417	
Rms bond length (Å)		0.008	
Rms bond angle (°)		1.132	

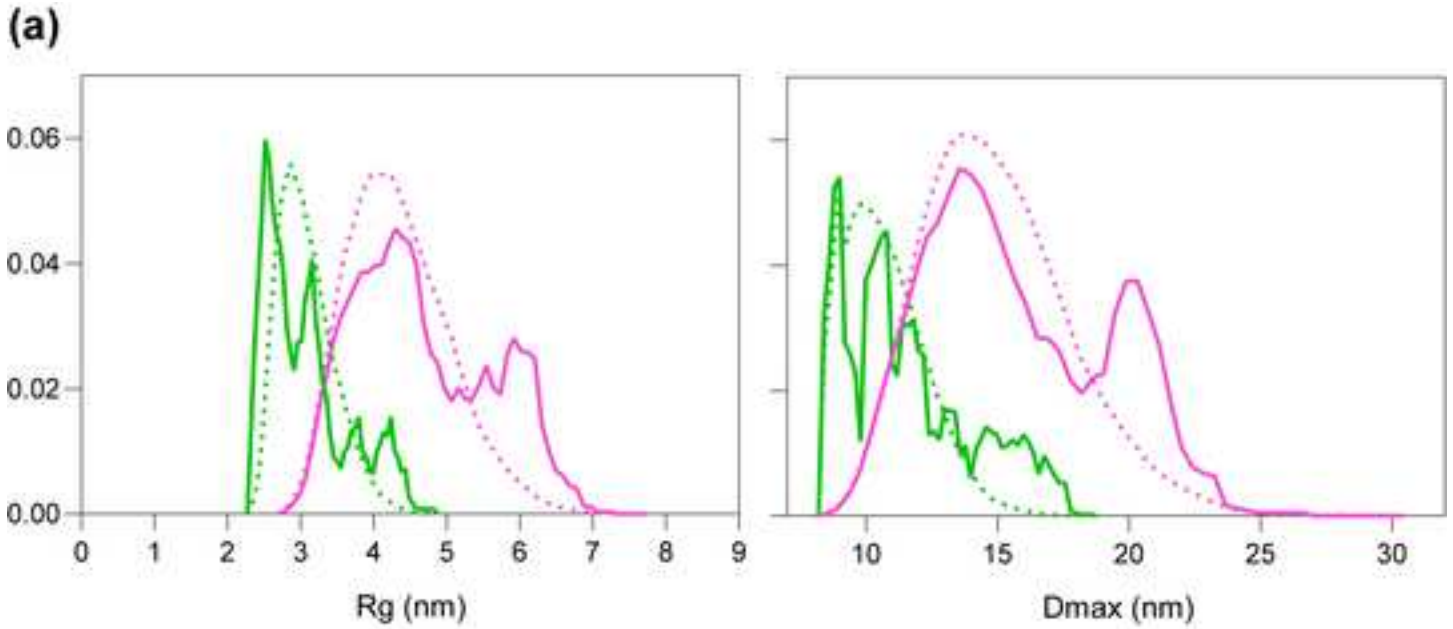
Supplementary Table 5. Oligonucleotides used in this study

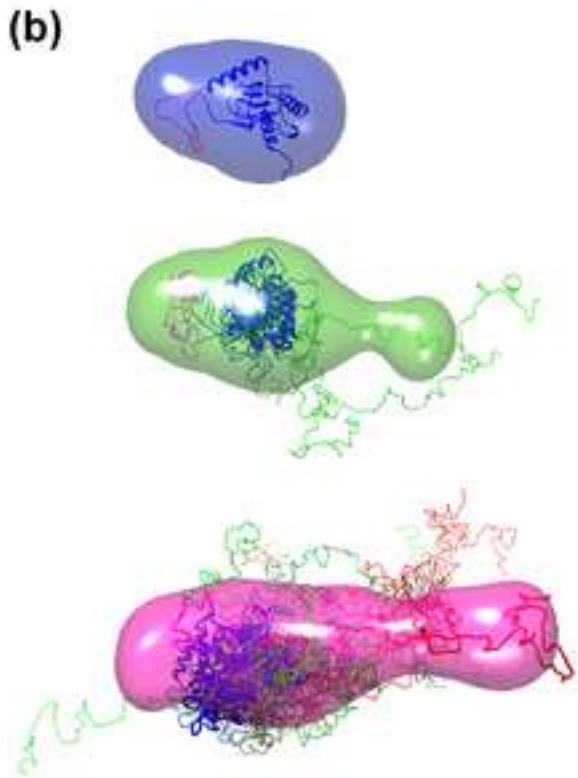
Primer	Sequence (5' --> 3')	Purpose
***KNR4*-Mut-Top**	CAAGAAAACTTGAGATCTCAATAATAAAAA TCTCAACCTG	pGEX-6P-3:: Knr4$^{\Delta C}$ construction
***KNR4*-Mut-Bot**	CAGGTTGAGATTTTTATTATTGAGATCTCAA GTTTTCTTG	pGEX-6P-3:: Knr4$^{\Delta C}$ construction
***KNR4*-Fwd**	TACCATGGGGATCCACTAGATTTTCATCCCT	pZE13::Pro-*KNR4*-Ter construction
***KNR4*-Rev**	TAAGCTTCGCGGTGGCGGC	pZE13::Pro-*KNR4*-Ter construction
pZE13-Fwd	GTGGATCCCCATGGTACGCGTGC	pZE13::Pro-*KNR4*-Ter construction
pZE13-Rev	CCACCGCGAAGCTTATCGATACCGTCGACC	pZE13::Pro-*KNR4*-Ter construction
***KNR4*-Rep-Fwd**	CAAGCCCTAAAGCACGTGAC	Repair Fragment
***KNR4*-Rep-Rev**	CTTCGTAGTGGCCTCAAACC	Repair Fragment
ΔN-Fwd	TCCACGGAGTCAAACGATG	DD1 deletion in pZE13::Pro-*KNR4*-Ter
ΔN-Rev	CATTTTATACTAAAAAATTCTGCCAAGTTG	DD1 deletion in pZE13::Pro-*KNR4*-Ter
ΔL-Fwd	CCAGATCAAAAATCTATTCCTCCAAATG	DL deletion in pZE13::Pro-*KNR4*-Ter
ΔL-Rev	AGATCTTTTGTTTAGGTTCTTTGCG	DL deletion in pZE13::Pro-*KNR4*-Ter
ΔC-Fwd	TGAAATATCACAATTAACATTCTACAACC	DD2 deletion in pZE13::Pro-*KNR4*-Ter
ΔC-Rev	TTGTGATCTCAAGTTTTCTTGATACTTG	DD2 deletion in pZE13::Pro-*KNR4*-Ter

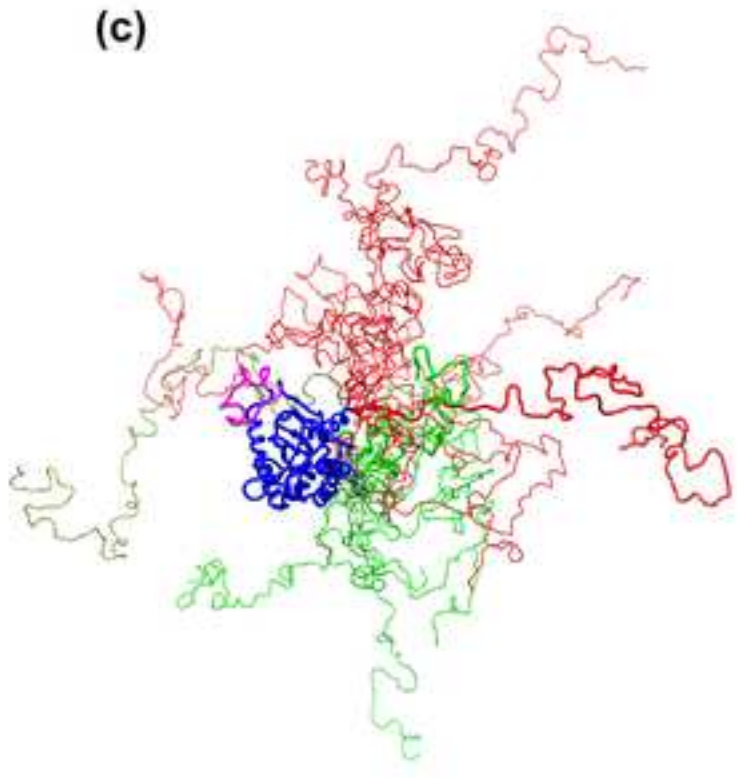
Figure



Figure

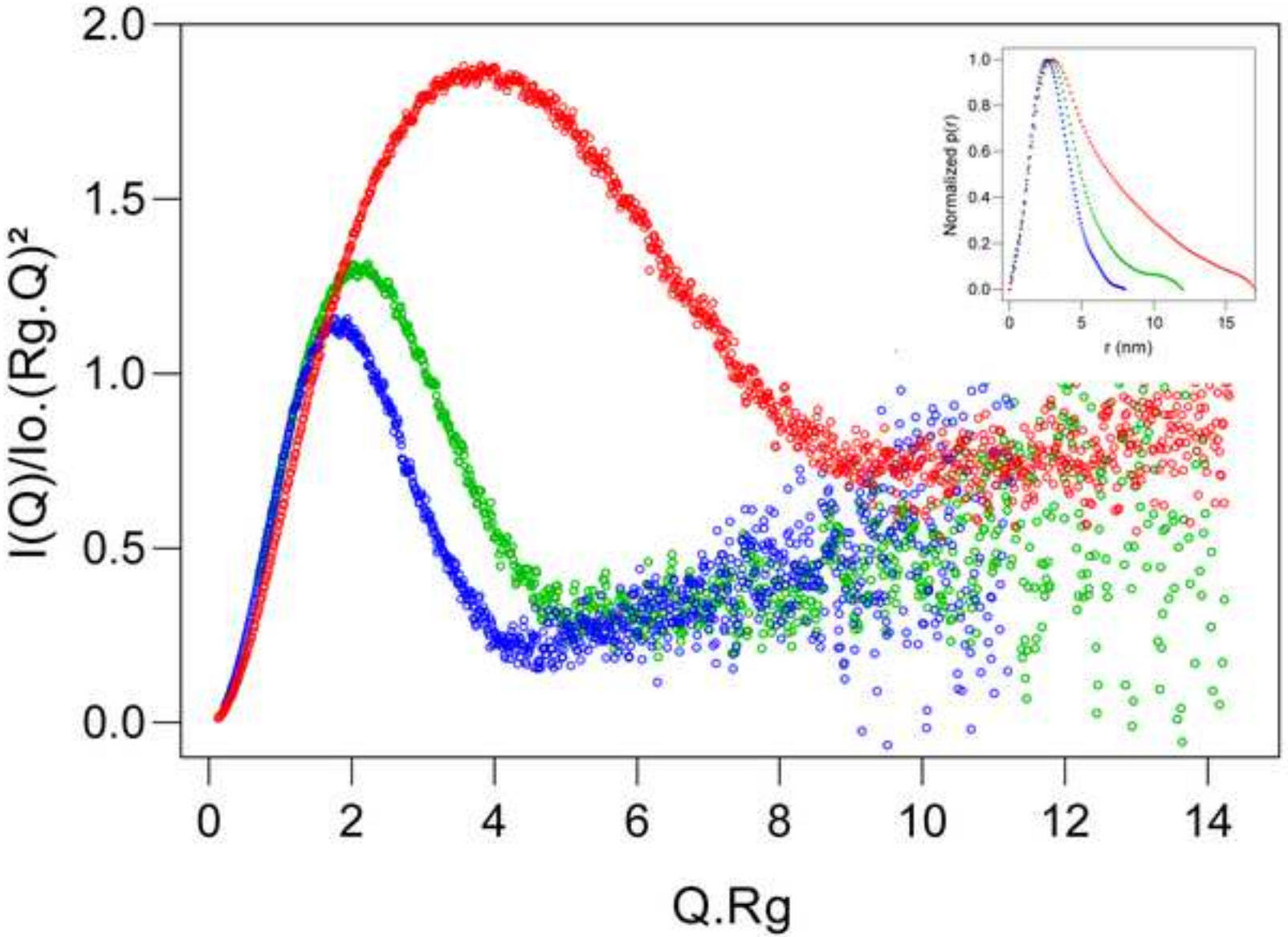




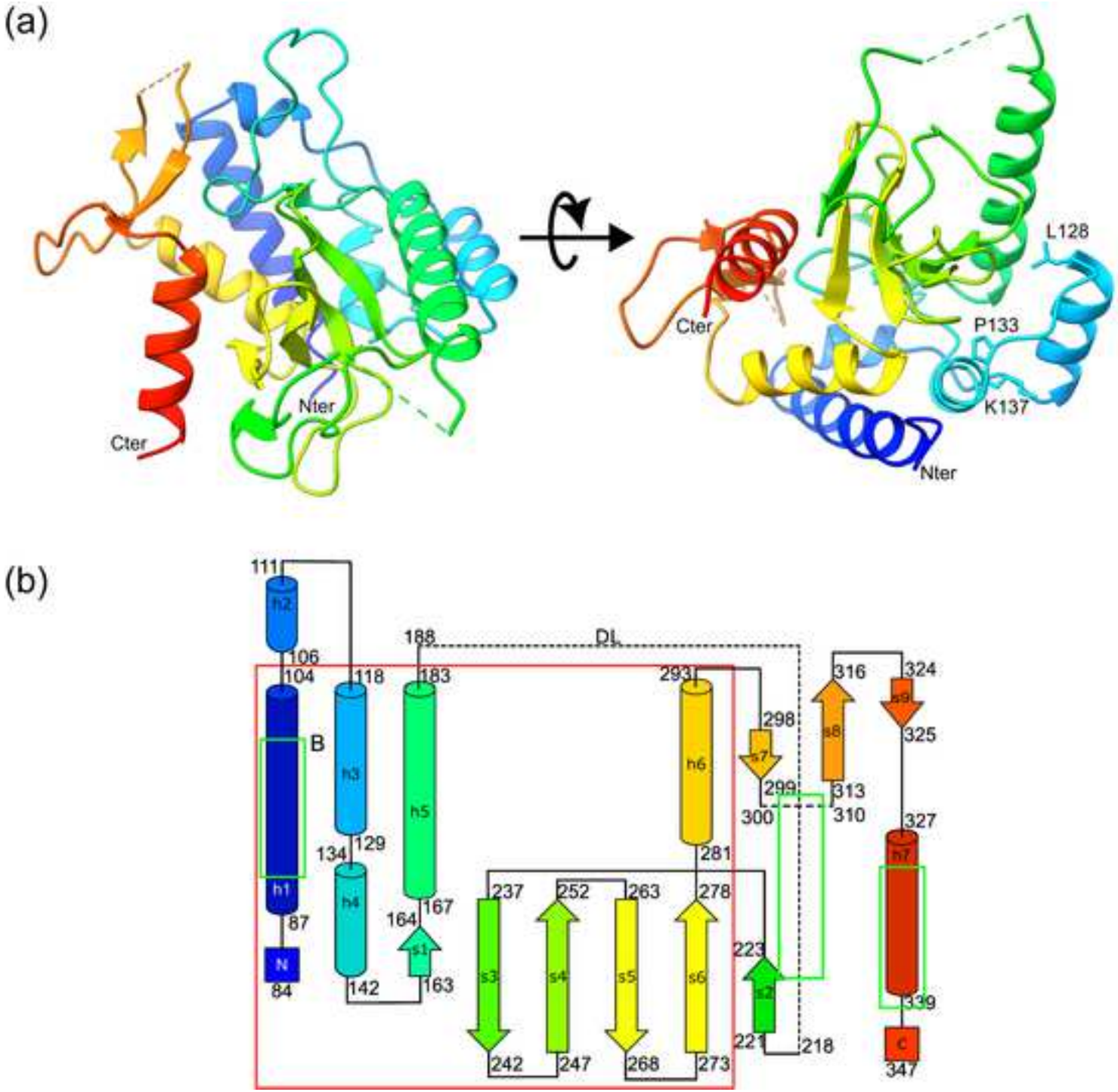


Figure

Click here to access/download;Figure;Figure 3_Batista et al.tif

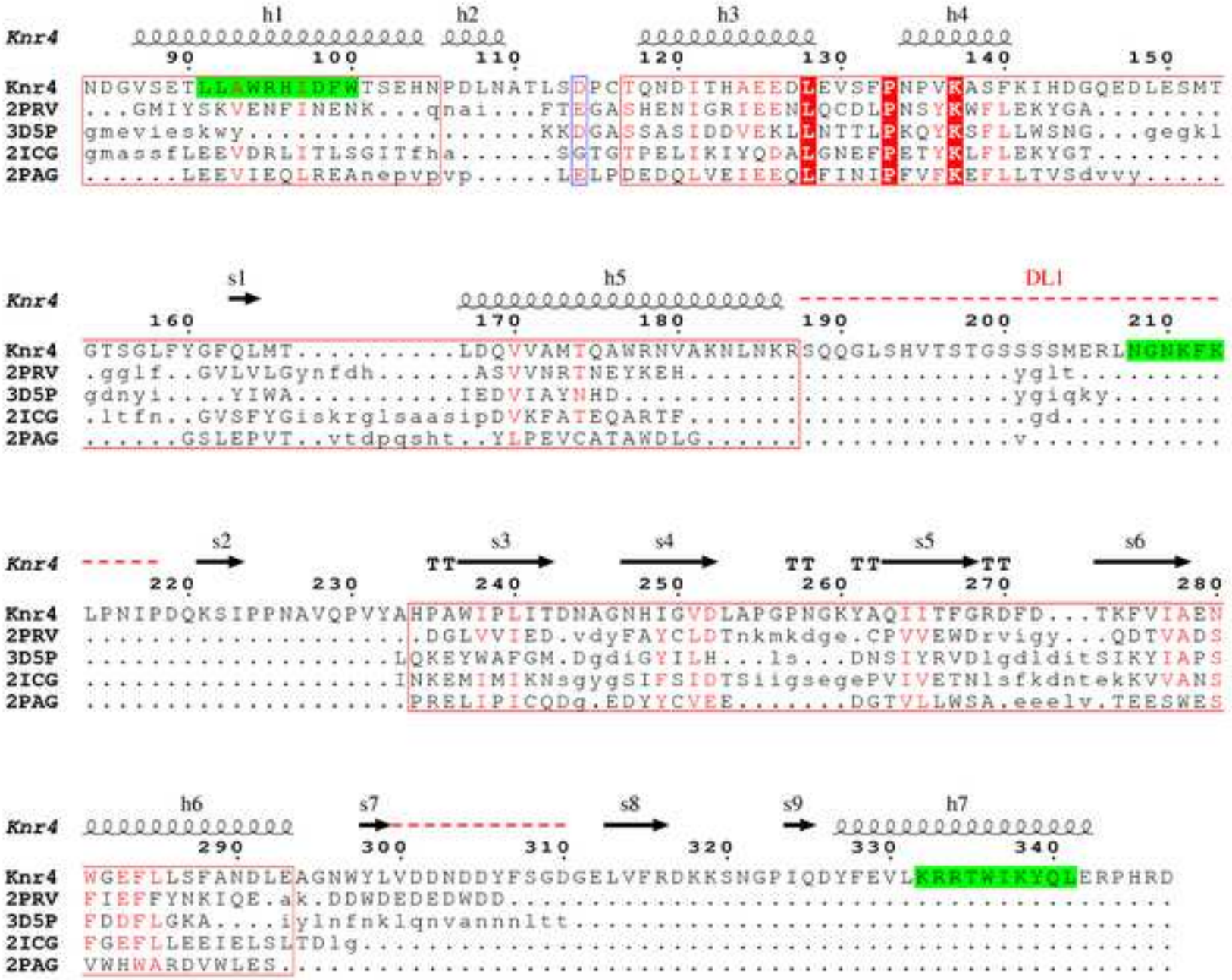


Figure



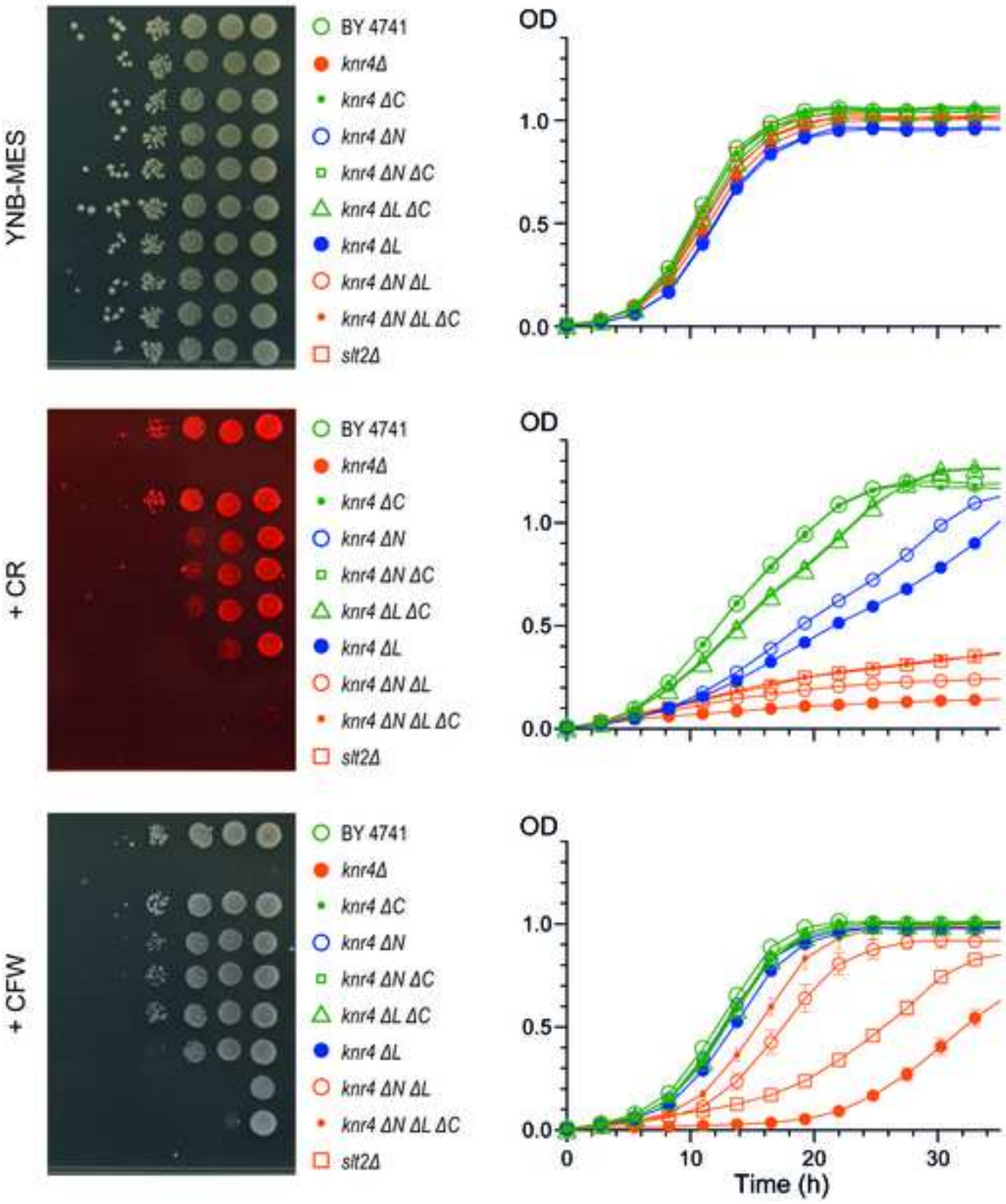
Figure

Click here to access/download;Figure;Figure 5_Batista et al.tif



Figure

Click here to access/download;Figure;Figure 6_Batista et al.tif



Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Credit author statement

Manon Batista (MB): Investigation (SAXS and Crystallography), Ellen Donker (ED): Investigation (*in vivo* experiments), Cecile Bon (CB): Investigation (SAXS), -writing original draft, -writing review and editing, Myriam Guillen (MG): investigation (protein purification), Adriana Caisso (AC): investigation (protein purification and crystallization), Lionel Mourey (LM): supervision, Jean Marie François (JMF): supervision, Laurent Maveyraud (LM): investigation (crystallography structure solving), -writing original draft, -writing review and editing, Didier Zerbib (DZ): conceptualization, supervision, -writing original draft,- writing Review and editing.