BBA - Proteins and Proteomics 1872 (2024) 141019

Contents lists available at ScienceDirect

BBA - Proteins and Proteomics

journal homepage: www.elsevier.com/locate/bbapap

Conformational and dynamic properties of the KH1 domain of FMRP and its fragile X syndrome linked G266E variant

Flavia Catalano [a], Daniele Santorelli [a], Alessandra Astegno [b], Filippo Favretto [b], Marco D'Abramo [c], Alessandra Del Giudice [c], Maria Laura De Sciscio [c], Francesca Troilo [d], Giorgio Giardina [a], Adele Di Matteo [d,*], Carlo Travaglini-Allocatelli [a,*]

[a] Department of Biochemical Sciences, Sapienza University of Rome, P.le Aldo Moro 5, Rome 00185, Italy
[b] Department of Biotechnology, University of Verona, Strada Le Grazie 15, Verona 37134, Italy
[c] Department of Chemistry, Sapienza University of Rome, P.le Aldo Moro 5, Rome 00185, Italy
[d] CNR Institute of Molecular Biology and Pathology, P.le Aldo Moro 5, Rome 00185, Italy

ARTICLE INFO

Keywords:
Fragile X syndrome
FMRP pathological mutation
KH domain
Conformational states
Structural dynamics

ABSTRACT

The Fragile X messenger ribonucleoprotein (FMRP) is a multi-domain protein involved in interactions with various macromolecules, including proteins and coding/non-coding RNAs. The three KH domains (KH0, KH1 and KH2) within FMRP are recognized for their roles in mRNA binding. In the context of Fragile X syndrome (FXS), over-and-above CGG triplet repeats expansion, three specific point mutations have been identified, each affecting one of the three KH domains (^{R138Q}KH0, ^{G266E}KH1, and ^{I304N}KH2) resulting in the expression of non-functional FMRP. This study aims to elucidate the molecular mechanism underlying the loss of function associated with the ^{G266E}KH1 pathological variant. We investigate the conformational and dynamic properties of the isolated KH1 domain and the two KH1 site-directed mutants ^{G266E}KH1 and ^{G266A}KH1. Employing a combined *in vitro* and *in silico* approach, we reveal that the ^{G266E}KH1 variant lacks the characteristic features of a folded domain. This observation provides an explanation for functional impairment observed in FMRP carrying the G266E mutation within the KH1 domain, as it renders the domain unable to fold properly. Molecular Dynamics simulations suggest a pivotal role for residue 266 in regulating the structural stability of the KH domains, primarily through stabilizing the α-helices of the domain. Overall, these findings enhance our comprehension of the molecular basis for the dysfunction associated with the ^{G266E}KH1 variant in FMRP.

1. Introduction

The hnRNP K homology (KH) domains are small and structurally conserved domains found in various proteins across all kingdoms of life [1–5]. They play multiple functions with a predominant role in regulating transcription [6,7] and translation [8] by specifically recognizing and binding single-stranded nucleic acids [9,10]. These domains exhibit a common structural organization consisting of a central βααβ core decorated by an additional α−helix and β−strand (α' and β') located either at the C- or N-terminus. This distinction defines two primary KH subtypes, *i.e.* the KHI class ($\beta_1\alpha_1\alpha_2\beta_2\beta'\alpha'$) commonly present in eukaryotic proteins, and the KHII class ($\alpha'\beta'\beta_1\alpha_1\alpha_2\beta_2$) typically found in prokaryotic proteins [9]. Therefore, the fundamental structural organization of all KH domains consists of three α-helices packed against a three-stranded β-sheet. The conserved (I/L/V)IGxxGxx(I/L/V) motif, commonly known as the GxxG motif, is frequently located between the two alpha helices within the core of KH domains, [9,10]. It has been proposed to be essential for nucleic acid binding [11]. Notably, KH domains lacking this motif do exist and have been referred to as divergent or degenerated KH [12,13].

The Fragile X messenger ribonucleoprotein (FMRP) is a multidomain protein pivotal in neuronal development and synaptic plasticity. FMRP's structure encompasses i) an N-terminal region housing two Tudor domains primarily responsible for protein-protein interaction along with a degenerated type I KH domain (KH0); ii) two closely positioned type I KH domains (KH1 and KH2) primarily involved in RNA recognition; iii) a C-terminal intrinsically disordered region containing a RGG box known to interact with G-quadruplex structures of RNA [12,14].

Fragile X syndrome (FXS) is an inherited neurodevelopmental disorder characterized by intellectual disability, autistic-like behaviors,

* Corresponding authors.
E-mail addresses: adele.dimatteo@cnr.it (A. Di Matteo), carlo.travaglini@uniroma1.it (C. Travaglini-Allocatelli).

https://doi.org/10.1016/j.bbapap.2024.141019
Received 2 February 2024; Received in revised form 26 March 2024; Accepted 14 April 2024
Available online 17 April 2024

and seizures [15–17]. Typically, FXS results from the absence of FMRP due to CGG triplet repeat expansion in the 5′-untranslated region of *fmr1* gene [14]. Recent research has identified cases of FXS linked to missense variants in FMRP [12,18]. Notably, three pathological missense mutations affect residues in each of the KH domains of FMRP: R138Q in KH0 [19,20], G266E in KH1 [21] and I304N in KH2 [22,23]. While it is widely acknowledged that these mutations primarily impact FMRP's ability to interact with RNAs or other molecular partners [24], the precise molecular mechanisms responsible for these effects remain largely unknown. To address this gap, recent research has focused on studying the KH0 domains of FMRP, providing insight into the aggregation and folding properties of the wild-type (wt) isolated domain compared to its pathological R138Q variant [25].

The present work is focused on characterizing the structural and dynamic properties of the isolated KH1 domain from FMRP (KH1). Given that the G266 to E mutation in KH1 has been shown to impair various functions of FMRP, including its association with polyribosomes, RNA binding and mGluR-mediated AMPA receptor trafficking [21]. Our investigation aims to gain a comprehensive understanding of how this mutation influences the folding of this domain. Based on the available crystal structure of the KH1-KH2 tandem (PDB code: 2qnd, [26]) the widely conserved G266 residue is located at the end of the β_3 strand of the 3-stranded antiparallel β-sheet of KH (Fig. 1 **and Supplementary Fig. S1**). Examination of the three-dimensional structure suggests that replacing G266 with the bulkier and charged glutamic acid residue might disturb the KH1 fold, providing a plausible hypothesis for the malfunction of the variant protein, as previously discussed [21]. To delve deeper into the role played by this residue in the structural dynamics of the KH domain and to unravel the molecular details of the functional impairment associated with the pathological variant of FMRP carrying the G266E substitution, here we present the results of the biophysical characterization of the isolated KH1 and two site-directed mutants, ^{G266E}KH1 and ^{G266A}KH1.

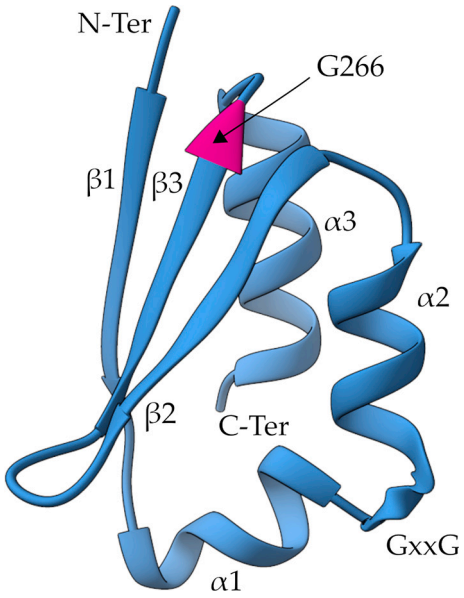

Fig. 1. The KH1 domain of FMRP. A) Ribbon representation of the structure of the KH1 domain of FMRP (res: 216–280; PDB code: 2qnd, Valverde et al. 2007); the position of G266 is highlighted in magenta. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

2. Materials and methods

2.1. Recombinant proteins production

The KH1 domain of FMRP (residues 216–280) and its variants ^{F262W}KH1 (KH1*), ^{G266E}KH1, and $^{F262W/G266A}$KH1 (^{G266A}KH1*) were cloned into the pET28b(+) vector and expressed in *E. coli* BL21 (DE3) strain. Mutants were generated using the QuikChange Lightning Site-Directed Mutagenesis kit (Agilent Technologies, Santa Clara, CA, USA).

Cells expressing the recombinant proteins were grown in LB broth supplemented with 34 μg/mL kanamycin at 37 °C until A_{600} nm reached 0.6/0.7. After the induction with Isopropyl-β-D-1-thiogalactopyranoside (IPTG) at 0.5 mM, cells were grown at 22 °C overnight (KH1, KH1*, ^{G266E}KH1) or for 4 h (^{G266A}KH1*) and then harvested by centrifugation. Purification of all proteins was carried out according to [25] with yields of approximately 10 mg/L for KH1, KH1*, ^{G266A}KH1, and approximately 1.5 mg/L for ^{G266E}KH1.

2.2. Equilibrium experiments

Circular dichroism (CD) experiments were conducted using a Jasco J710 instrument (Jasco Inc., Easton, MD, USA) equipped with a Peltier apparatus. Far-UV spectra (190–250 nm) were obtained using 20 μM proteins in 20 mM sodium phosphate pH 7.2, 100 mM NaCl using a 1 mm quartz cell (scanning speed of 100 nm/min, averaged over three acquisitions). Thermal denaturation profiles were recorded by monitoring the change in the CD signal at 222 nm while increasing temperature (from 20 °C to 90 °C, at a rate of 1 °C/min).

Urea-induced denaturation experiments of KH1* and ^{G266A}KH1* (protein concentration = 6 μM in 20 mM sodium phosphate pH 7.2, 100 mM NaCl) were followed by recording the fluorescence emission between 300 and 400 nm (λ_{ex} = 280 nm) using the FluoroMax-4 spectrofluorometer (Jobin Yvon, Edison, NJ, USA) with a 1 cm path length quartz cuvette. Fluorescence signals at 340 nm were plotted as a function of denaturant concentration and urea-induced denaturation data were analyzed using a two state-equation:

$$Yobs = \frac{(Y_N + (m_N{*}D)) + \left((Y_D + (m_D{*}D)){*}exp\left(\frac{m_{DN}{*}\left(D - D_{\frac{1}{2}}\right)}{0.58}\right)\right)}{1 + exp\left(\frac{m_{DN}{*}\left(D - D_{\frac{1}{2}}\right)}{0.58}\right)}$$

where Y_N and Y_D refer to the y-intercepts and m_N and m_D represent the slopes of the native and denatured baselines, respectively; D is the denaturant concentration. Based on these parameters, ΔG_{DN} was calculated using the equation: $\Delta G^0_{DN} = -\, m_{DN}\, [D]_{1/2}$.

2.3. SAXS analysis

SAXS measurements were conducted at SAXSLab Sapienza using a Xeuss 2.0 Q-Xoom system (Xenocs SAS, Grenoble, France). Sample details and data collection parameters are provided in Supplementary Table S1. The analysis of the scattering profiles was performed using the tools of ATSAS 3.2 [27]. Additional details on data collection methodologies and processing procedures are reported in the Supplementary materials section.

2.4. Molecular dynamics (MD) simulations

The initial structure of the KH1 domain was derived from the KH1-KH2 FMRP crystal structure (PDB code: 2qnd, B monomer [26]) with selenomethionine replaced by methionine. Structures for the ^{G266E}KH1 and ^{G266A}KH1 variants were modeled by AlphaFold [28]. Simulations for wt KH1 and ^{G266E}KH1 were conducted at 300 K, 370 K and 420 K while ^{G266A}KH1 simulations were carried out at 300 K and 420 K. An additional simulation at 450 K was performed for the wt KH1 (Table 1).

Table 1
MD simulations times and temperatures of the KH1 domains investigated in this work.

	300 K	370 K	420 K	450 K
KH1	~ 1.4 μs	~ 2 μs	0.4 μs	0.3 μs
^{G266E}KH1	~ 1.4 μs	~ 2 μs	0.4 μs	–
^{G266A}KH1	~ 1.4 μs	–	0.4 μs	–

Additional information on MD simulations is reported in the Supplementary Materials section.

2.5. Nuclear magnetic resonance (NMR) spectroscopy

^{1}H mono-dimensional (1D) NMR experiments (^{1}H—1D NMR) were acquired on a 600 MHz Bruker AVANCE NEO spectrometer equipped with a triple resonance cryo-probe (Bruker, Karlsruhe, Germany). All the experiments were performed at 298 K, recording 16 transients with a recycle delay of 1.2 s.

For the NMR experiments, proteins were dissolved in 20 mM sodium phosphate pH 7.4, 100 mM NaCl, supplemented with 5% D_2O. [KH1] and [^{G266E}KH1] = 600 μM; [^{G266A}KH1*] = 100 μM.

3. Results

3.1. The ^{G266E}KH1 variant is unfolded in solution

KH1 and ^{G266E}KH1 domains expressed in *E. coli* and purified to homogeneity, were initially examined by far-UV CD spectroscopy. The obtained results (Fig. 2A) highlight stark differences: the KH1 spectrum displayed characteristic features indicative of a properly folded α/β protein as expected. In contrast, the CD spectrum of the ^{G266E}KH1 lacked distinct features in the 210–230 nm region, with a notable minimum at approximately 200 nm, suggesting minimal to no secondary structure content. Additionally, thermal-induced denaturation experiments, monitored *via* far-UV CD spectroscopy, demonstrated a cooperative unfolding process solely in the case of KH1 revealing an apparent Tm at around 51 °C. Conversely, no observable unfolding transition was discernible for the ^{G266E}KH1 domain (Fig. 2B).

To further investigate the structural properties of the ^{G266E}KH1 variant, Small-Angle X-ray Scattering (SAXS) experiments were conducted on both KH1 and ^{G266E}KH1. The resulting scattering profiles displayed noticeable discrepancies (Fig. 2C), reflecting distinct features in their respective pair distance distribution functions (P(r)). The P(r) of KH1 has a compact profile, signifying a globular conformation with a maximum size of approximately 6–7 nm. In contrast, the P(r) obtained for ^{G266E}KH1 is shifted and expanded towards larger distances compared to the wt protein. This shift implies that the pathological variant assumes an overall larger particle size with a more extended conformation. This dissimilarity was further emphasized by the normalized Kratky plot representation (Fig. 2D), highlighting the increased flexibility of the pathological variant. Specifically, the plot for ^{G266E}KH1 deviates from the bell-shaped curve observed for fully globular proteins and instead exhibits a plateau-like behavior more commonly associated with unstructured polypeptides [29]. Moreover, while the wt protein exhibited consistent and overlapping SAXS profiles over the entire 24-h experiment at 20 °C, the ^{G266E}KH1 sample underwent aggregation within a mere 15-h period (**Supplementary Fig. S2**).

The mono-dimensional NMR spectroscopy experiment (^{1}H—1D NMR) conclusively established the profound impact of the G266 to E mutation on the structure of the KH1. In —^{1}H-1D NMR spectra, the signal dispersion in the regions of the amide (6–10 ppm), and methyl (−0.5 to 0.5 ppm) protons provides indications on the folded state of the proteins. In contrast to the distinct features observed in the ^{1}H—1D NMR spectrum of KH1, which include sharp, well-distributed signals along with discernible methyl signals, indicative of a folded protein, the ^{1}H—1D NMR spectrum of the ^{G266E}KH1 showcased characteristics typical of an unstructured protein (Fig. 2E). Notably, the proton chemical shift dispersion in ^{G266E}KH1 spectrum was significantly reduced, particularly evident in the poor chemical shift dispersion of its backbone amidic protons (around 7–10 ppm), a common trait found in unfolded proteins [30,31]. Furthermore, the aliphatic region of the spectrum revealed very sharp signals primarily clustered at approximately 3.6 ppm, suggesting a compromised structural instability in the mutated protein.

3.2. G266 has a key role in the KH fold

Glycine 266 is a conserved residue found in various KH domains of different origins situated at the C-terminal end of the β3 strand [21] [32] (Fig. 1 **and supplementary Fig. S1**). As shown above, its substitution with glutamic acid results in the complete unfolding of the domain. To further investigate the role of glycine at position 266, we engineered a mutant where glycine was substituted with alanine (^{G266A}KH1). Although the alanine substitution introduces minimal steric hindrance, it does reduce the conformational flexibility of the main chain compared to glycine. Moreover, as KH1 lacks Trp residues, a commonly used intrinsic fluorescence probe in folding analyses, we substituted F262 with a Trp residue in both KH1 and ^{G266A}KH1 (referred to as KH1* and ^{G266A}KH1* hereafter). The modification at position 262 did not induce significant alterations in the overall domain conformation as observed by far-UV CD (**Supplementary Fig. S3A**).

Far-UV CD analysis revealed that the G to A substitution at position 266 does not alter the overall secondary structure of the KH1 domain and that the ^{G266A}KH1* variant maintains proper folding (Fig. 3A). Furthermore, SAXS experiments conducted on ^{G266A}KH1* demonstrated that its scattering profile closely aligns with that of the wt protein (Fig. 3B), suggesting that this variant adopts a globular conformation and does not aggregate in solution. Accordingly, the ^{1}H—1D NMR spectrum of the ^{G266A}KH1* variant appeared dispersed, especially in the methyl group region, indicating the ability of the protein to assume a folded structure (**Supplementary Fig. S3B**).

The urea-induced denaturation processes observed for both KH1* and ^{G266A}KH1* were found to be reversible (Fig. 3C), enabling the determination of their respective thermodynamic stabilities (ΔG = 2.08 kcal mol^{-1} for KH1* and ΔG = 1.54 kcal mol^{-1} for ^{G266A}KH1*). Notably, the obtained m-values were remarkably similar (m = 0.80 ± 0.01 kcal mol^{-1} M^{-1} and m = 0.82 ± 0.03 kcal mol^{-1} M^{-1} for KH1* and ^{G266A}KH1*, respectively), strongly suggesting the maintenance of an identical native state structure in both proteins. In summary, while the G to A mutation does not alter the overall domain structure, it does lead to a marginal destabilization of the native state by approximately 0.5 kcal mol^{-1}.

3.3. Unfolding dynamics of KH1 and its variants G266A and G266E by MD simulation

Having observed that the G to A mutation at position 266 marginally destabilizes the protein without significantly altering its conformational and structural integrity, we aimed to elucidate the role of position 266 in modulating the native protein's conformational stability through Molecular Dynamics (MD) investigations. Initially, a simulation lasting 1.4 μs at 300 K was performed starting from the KH1 structure (PDB: 2qnd, [26]), and the ^{G266A}KH1 model obtained by AlphaFold. Additionally, since the unfolding process occurs over an extended time scale in solution, we further analyzed the protein's structural-dynamical behavior at 420 K for 0.4 μs. A structural analysis was carried out on the MD trajectories, and the Root Mean Square Deviation (RMSD), representing the average deviation of the Cα atoms with respect to the initial structures, was calculated for both wt and mutant trajectories (Fig. 4A). The Cα Root Mean Square Fluctuations (RMSF) were also calculated, allowing the identification of the most affected residues (Fig. 4B). These analyses

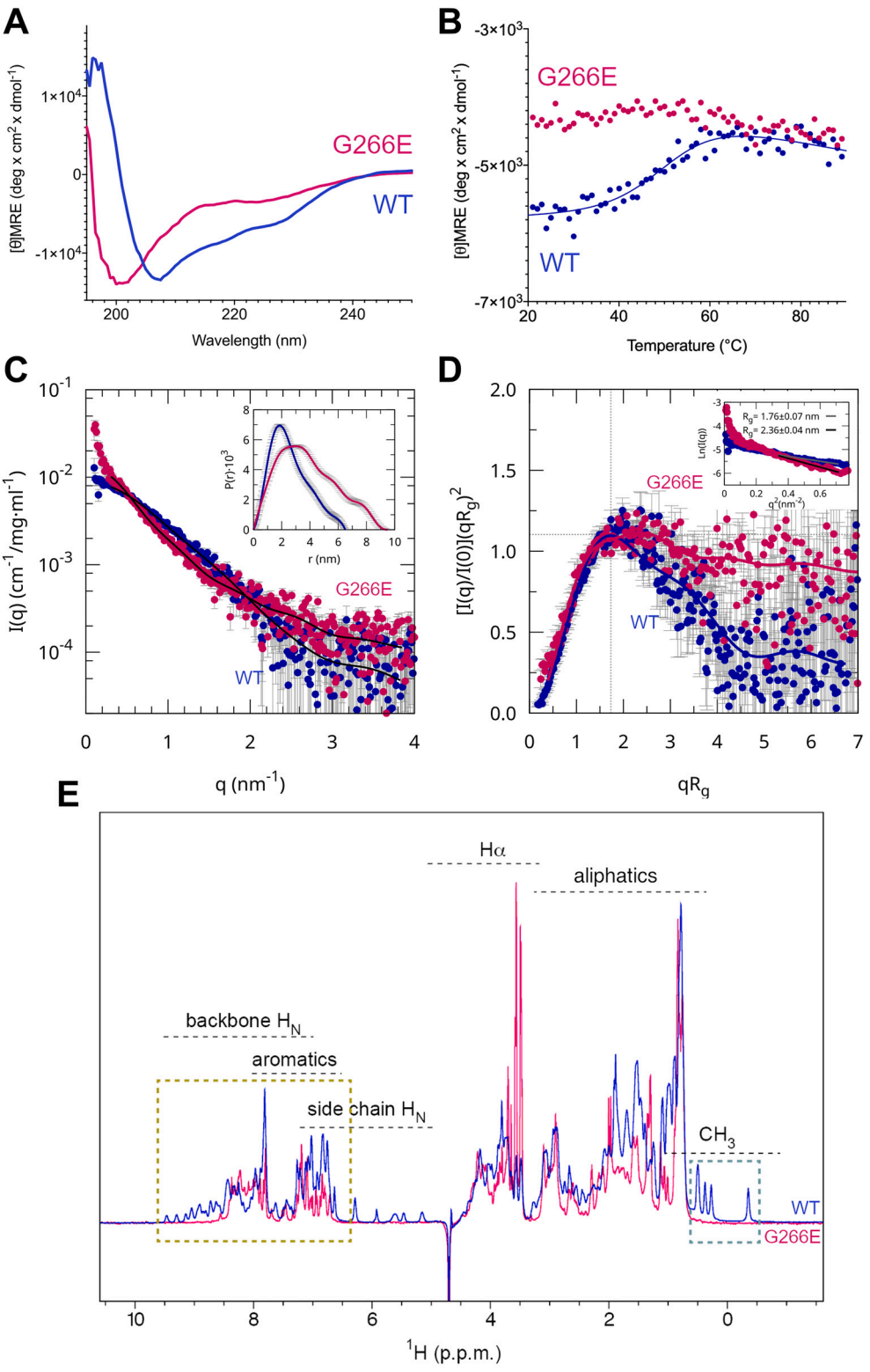

Fig. 2. Structural analysis of wt KH1 and ^{G266E}KH1. In all panels data for the wt KH1 are represented in blue and those of the ^{G266E}KH1 in magenta. A) Far-UV CD spectra at 20°. B) Thermal-induced denaturation monitored by far-UV CD at 222 nm. C) SAXS data fitted by indirect Fourier transform (black solid lines) to obtain the pair distance distribution functions P(r) shown in the inset. D) Dimensionless Kratky plots of SAXS data. The R_g and I(0) values obtained from the Guinier fit shown in the inset were applied. E) Superimposition of the ^{1}H—1D NMR spectra of KH1 and its variant G266E. The two regions characterizing a folded protein are shown in blue square (methyl region) and yellow square (amide protons). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

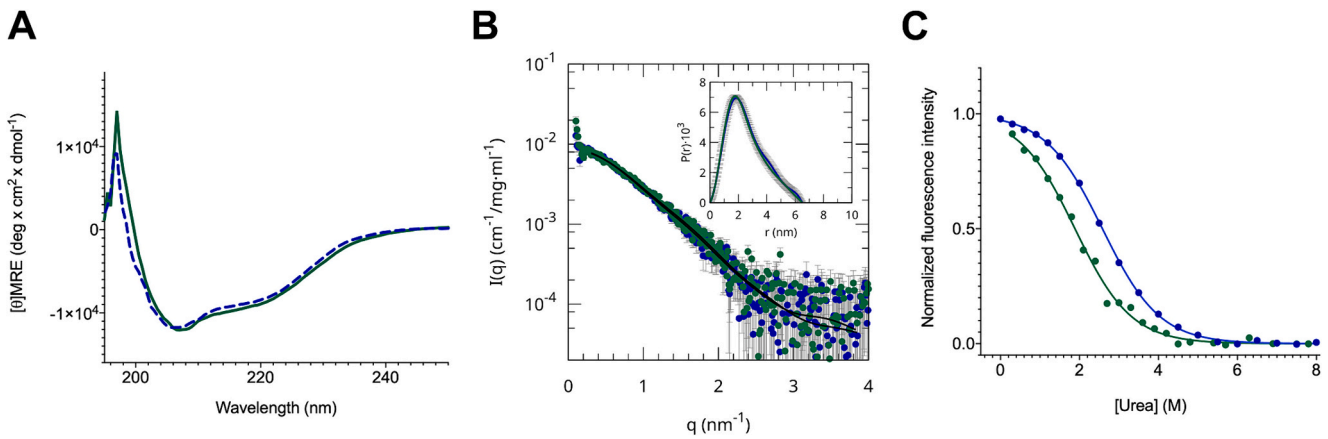

Fig. 3. Spectroscopic analysis of KH1* (blue) and ^{G266A}KH1* (green). A) Far-UV CD spectra of 20 μM samples collected at 20 °C. B) SAXS data and fitting by indirect Fourier transform (black solid lines) to obtain the pair distance distribution functions P(r) shown in the inset. C) Equilibrium unfolding profile at 20 °C obtained by urea-induced denaturation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

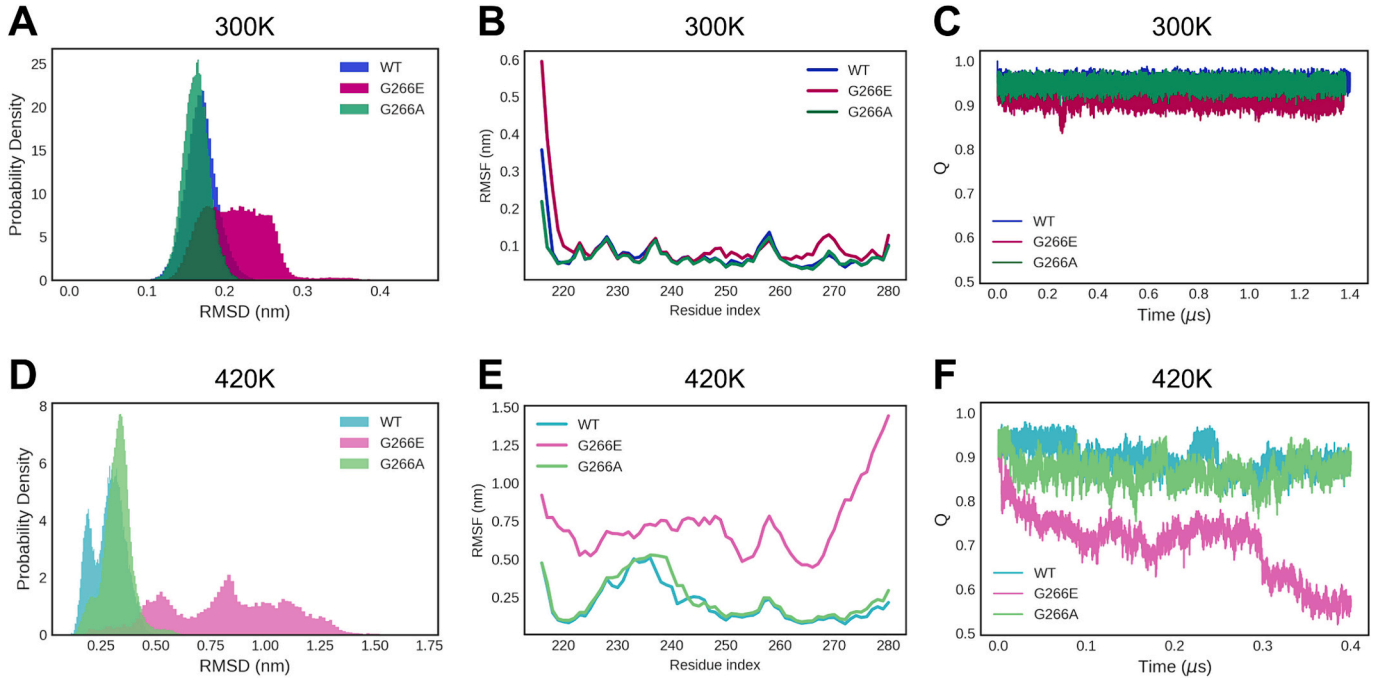

Fig. 4. Structural-dynamic characterization of KH1 (blue), ^{G266E}KH1 (magenta), and ^{G266A}KH1 (green) domain by MD simulations. Cα RMSD distribution, RMSF of Cα atoms, and time evolution of the fraction of native contacts (Q) along the simulation at 300 K (A-C) and 420 K (D—F). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

collectively suggest that the G to A substitution minimally perturbs the structural and dynamic behavior of the mutated protein compared to the wt. Moreover, the evaluation of native contacts (Q) over trajectory reveals a similar profile between the two simulations (Fig. 4C). These observations align with the finding that both proteins explore analogous regions within their essential subspace at 300 K (**Supplementary Fig. S4**). Notably when analyzing the MD simulation at higher temperature (420 K), consistent results were obtained for both wt KH1 and ^{G266A}KH1 (Fig. 4D-F).

Contrary to what was observed for the G266A mutation, our solution-based experiments revealed that substituting glycine with glutamine at position 266, determines the unfolding of KH1. To comprehend the influence of this mutation on the structural dynamics of the domain and to identify the secondary structure elements mainly affected, we conducted MD simulations on ^{G266E}KH1 utilizing the model generated by AlphaFold. Although the starting model aligns perfectly with KH1 (RMSD = 0.8 Å) (Fig. 5), MD trajectories display distinct conformational behavior on the microsecond time scale, and the two proteins explore different regions within their essential subspace even at 300 K (**Supplementary Fig. S4**). Accordingly, the RMSD analysis indicates that the sampled conformations of ^{G266E}KH1, at 300 K, have a higher deviation (from the wt crystal structure) compared to both the wt and the ^{G266A}KH1 (Fig. 4A). This difference becomes particularly pronounced at 420 K, where the pathological variant exhibits an RMSD distribution characteristic of a largely unfolded protein, while KH1 retains a relatively structured conformation (Fig. 4D). Notably, the RMSF analysis highlights that the variation in the conformational dynamics of ^{G266E}KH1 with respect to KH1 primarily originates from higher fluctuations in the region surrounding position 266 and, to a lesser extent, in the N-terminal region and around position 250 (Fig. 4B, E). The examination of native contacts further supports the observation that ^{G266E}KH1 has a greater tendency to disrupt native contacts compared to

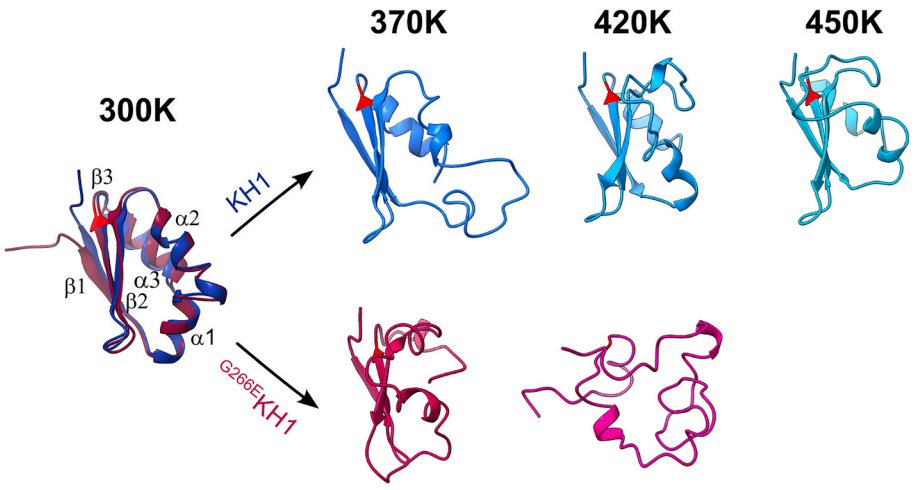

Fig. 5. Structural elements affected by G266E mutation. Extracted snapshots corresponding to the higher RMSD value, at all the investigated temperatures for both wt (blue) and ^{G266E}KH1 (magenta). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the wt construct. While both the wt and the ^{G266A}KH1 maintain native contacts throughout the simulation, even at 420 K, ^{G266E}KH1 starts losing native contacts from the outset of the simulation (Fig. 4C, F).

As the pathologic variant is mainly unfolded at high temperature (420 K), we performed an additional simulation at 370 K, lasting 2 μs, to investigate and compare the conformational dynamics of ^{G266E}KH1 and KH1 at a milder temperature. RMSD, RMSF, and native contacts analysis, reported in **Supplementary Fig. S5**, show all together that the high fluctuations in the mutation region trigger an overall destabilization of the ^{G266E}KH1 structure while, as expected, the wt construct retains almost completely its native folding. To pinpoint which elements are structurally affected by the mutation, we compared the structures corresponding to the maximum value of the Cα RMSD, at all the simulated temperatures (Table 1), for both wt and ^{G266E}KH1 (Fig. 5). Throughout the wt simulations, the β-sheet maintains structural integrity across the entire trajectory. In contrast, both the α1 and α2 helices exhibit an early loss of their native structure even at 370 K (**Supplementary Fig. S5B**), while the α3 helix remains structurally stable at 370 K, necessitating higher temperatures for its unfolding to occur. On the contrary, the ^{G266E}KH1 mutation affects the stability of all three α-helices even at the lowest temperature explored (370 K) while the β-sheet unfolds at a higher temperature. Notably, although the mutation is localized at the C-terminal end of the β3 strand, the major effects concern the unfolding of the helical portion of the domain.

4. Discussion

It is known that specific point mutations situated within the KH domains of FMRP correlate with manifestations of FXS, even in the absence of the CGG-repeat expansion in the *fmr1* gene [18]. In particular, the I304N mutation in the KH2 domain has been linked to severe FXS phenotypes characterized by low IQ, social and behavioral impairment, and macroorchidism [22,23]. This mutation alters the structure of the protein and hinders its function [26]. Similarly, the missense mutation R138Q in the KH0 domain is associated with developmental delay and symptoms related to FXS [20]. While this mutation does not significantly alter the three-dimensional structure of the domain, it does impact its aggregation properties [25].

A third mutation, G266E, located in the KH1 domain of FMRP, was identified in a patient with FXS [21]. This mutation affects various functions of FMRP, including RNA binding and translational regulation, polyribosome association, and mGluR-mediated AMPA receptor trafficking and granule formation [21,33]. The impact of this substitution at the protein level has been hypothesized by Myrick and colleagues, suggesting that the substitution of glycine with glutamic acid might be sterically incompatible, thereby affecting the proper folding of KH1.

Here, we have demonstrated that the G to E mutation does indeed cause the complete unfolding of the variant domain. The far-UV CD spectrum of this pathological variant clearly exhibited features indicative of a random coil, and correspondingly, the thermal denaturation profile did not show a discernable cooperative transition (Fig. 2A, B). The ^{1}H—1D NMR spectrum of ^{G266E}KH1, displaying broad and poorly dispersed peaks typical of an unstructured protein, conclusively highlighted the profound impact of the mutation on the proper folding of KH1 (Fig. 2E). The SAXS analysis of the mutated protein also showed features typicalof non-globular and highly dynamical polypeptides and suggested that the variant protein, but not the wt KH1, has the propensity to aggregate over time. In summary, our results unequivocally indicate that the G266E mutation fully compromises the structure of the KH1 domain of FMRP. It is worth noticing that since KH1 is tandemly connected with KH2 and linked to the N-terminal region of FMRP, it is plausible that the unfolding of this domain may lead to a larger impairment of the entire structured portion of FMRP.

Glycine 266, which is located at the end of the beta-strand β3 at the start of the sharp turn connecting it to the α3 helix, is highly conserved among eukaryotic KH domains (**Supplementary Fig. S1**) [21]. To comprehend the structural significance of the amino acid at this position within the KH domains, we replaced glycine with alanine. The G to A substitution introduces a smaller, non-charged side chain compared to the glutamic acid of ^{G266E}KH1. We found that the G to A variant maintains the overall three-dimensional structure of KH1, causing only a marginal destabilization of the domain (approximately 0.5 kcal mol^{-1}) as demonstrated by chemical denaturation experiments (Fig. 3C). These observations suggest that i) the stability of the KH fold significantly depends on the interactions formed between strand β3 and helix α3 and ii) the steric optimization provided by the smaller side chain of alanine (compared to glutamic acid) rather than the conformational flexibility of glycine is crucial for preserving the folding of the domain.

Our MD simulations, carried out to obtain atomic-level insight into the unfolding dynamics of the KH1 domain, suggested one of the possible unfolding pathways, which can be described by a hierarchical progression initially involving helices α_1, α_2 and the GxxG motif forming the RNA binding cleft [9]. Although position 266 is part of the β-sheet, substantial effects predominantly manifest in the helical portion of the domain. Moreover, while the helix α3 maintains its structure for longer times, only the β-sheet remains correctly folded even at 470 K within our simulation timeframe. It is worth noting that the sequence of events observed during the unfolding dynamics of the KH1 domain mirrors

those reported for the unfolding of KH2, one of the other KH domains of FMRP [34]. This suggests a potentially conserved mechanism among eukaryotic type I KH domains. Investigating whether this process universally applies to KH domains or if variations exist, particularly in the topologically distinct prokaryotic type II KH domains, remains an intriguing area for further exploration.

Funding

The Sapienza Research Infrastructure is acknowledged for the SAXS measurements at SAXSLab Sapienza, funded by the Large Equipment Project 2015-C26J15BX54. A.D.G. acknowledges co-financing of Sapienza University of Rome and the European Union—FSE-REACT-EU, PON Research and Innovation 2014–2020 DM1062/2021 for the RTD-A contract 26-G-15245-1. We ackonwledge ISCRA for awarding this project access to the LEONARDO supercomputer, owned by EuroHPC Joint Undertaking, hosted by CINECA (Italy).

This work was further partially funded by the European Union - NextGenerationEU through the ItaliaDomani PNRR projects "Potentiating the Italian Capacity for Structural Biology Services in Instruct-ERIC" (ITACA.SB, no. IR0000009) and #NEXTGENERATIONEU and the Italian Ministry of University and Research, National Recovery and Resilience Plan (PNRR), project MNESYS (PE0000006).

CRediT authorship contribution statement

Flavia Catalano: Writing – review & editing, Investigation, Formal analysis. **Daniele Santorelli:** Investigation, Formal analysis. **Alessandra Astegno:** Writing – review & editing, Investigation, Formal analysis. **Filippo Favretto:** Investigation, Formal analysis. **Marco D'Abramo:** Writing – review & editing, Investigation, Formal analysis. **Alessandra Del Giudice:** Writing – review & editing, Investigation, Funding acquisition, Formal analysis. **Maria Laura De Sciscio:** Writing – review & editing, Investigation, Formal analysis. **Francesca Troilo:** Writing – review & editing, Investigation, Formal analysis. **Giorgio Giardina:** Writing – review & editing, Supervision. **Adele Di Matteo:** Writing – review & editing, Writing – original draft, Supervision, Investigation, Funding acquisition, Conceptualization. **Carlo Travaglini-Allocatelli:** Writing – review & editing, Writing – original draft, Supervision, Investigation, Conceptualization.

Declaration of competing interest

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.

Data availability

Data will be made available on request.

Acknowledgements

We thank the Centro Piattaforme Tecnologiche of the University of Verona for providing access to the spectroscopic platform.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbapap.2024.141019.

References

[1] H. Siomi, M.J. Matunis, W.M. Michael, G. Dreyfuss, The pre-mRNA binding K protein contains a novel evolutionary conserved motif, Nucleic Acids Res. 21 (1993) 1193–1198, https://doi.org/10.1093/nar/21.5.1193.

[2] M. Olejniczak, X. Jiang, M.M. Basczok, G. Storz, KH domain proteins: another family of bacterial RNA matchmakers? Mol. Microbiol. 117 (2022) 10–19, https://doi.org/10.1111/mmi.14842.

[3] N.V. Grishin, KH domain: one motif, two folds, Nucleic Acids Res. 29 (2001) 638–643, https://doi.org/10.1093/nar/29.3.638.

[4] Y. Zhang, Y. Ma, R. Liu, G. Li, Genome-wide characterization and expression analysis of KH family genes response to ABA and SA in Arabidopsis thaliana, Int. J. Mol. Sci. 23 (2022), https://doi.org/10.3390/ijms23010511.

[5] D. Santorelli, S. Rocchio, F. Fata, I. Silvestri, F. Angelucci, F. Imperi, D. Marasco, C. Diaferia, L. Gigli, N. Demitri, L. Federici, A. Di Matteo, C. Travaglini-Allocatelli, The folding and aggregation properties of a single KH-domain protein: ribosome binding factor a (RbfA) from Pseudomonas aeruginosa, Biochim. Biophys. Acta Gen. Subj. 1865 (2021) 129780, https://doi.org/10.1016/j.bbagen.2020.129780.

[6] L.M. Quinn, FUBP/KH domain proteins in transcription: Back to the future, Transcription 8 (2017) 185–192, https://doi.org/10.1080/21541264.2017.1293595.

[7] F. Sánchez-Jiménez, V. Sánchez-Margalet, Role of Sam68 in post-transcriptional gene regulation, Int. J. Mol. Sci. 14 (2013) 23402–23419, https://doi.org/10.3390/IJMS141223402.

[8] A. Ostareck-Lederer, D.H. Ostareck, Control of mRNA translation and stability in haematopoietic cells: the function of hnRNPs K and E1/E2, Biol. Cell. 96 (2004) 407–411, https://doi.org/10.1016/J.BIOLCEL.2004.03.010.

[9] R. Valverde, L. Edwards, L. Regan, Structure and function of KH domains, FEBS J. 275 (2008) 2712–2726, https://doi.org/10.1111/j.1742-4658.2008.06411.x.

[10] G. Nicastro, I.A. Taylor, A. Ramos, KH-RNA interactions: Back in the groove, Curr. Opin. Struct. Biol. 30 (2015) 63–70, https://doi.org/10.1016/j.sbi.2015.01.002.

[11] D. Hollingworth, A.M. Candel, G. Nicastro, S.R. Martin, P. Briata, R. Gherzi, A. Ramos, KH domains with impaired nucleic acid binding as a tool for functional analysis, Nucleic Acids Res. 40 (2012) 6873–6886, https://doi.org/10.1093/nar/gks368.

[12] I. D'Annessa, F. Cicconardi, D. Di Marino, Handling FMRP and its molecular partners: structural insights into fragile X syndrome, Prog. Biophys. Mol. Biol. 141 (2019) 3–14, https://doi.org/10.1016/j.pbiomolbio.2018.07.001.

[13] M.H.K. Cheng, R.P. Jansen, A jack of all trades: the RNA-binding protein vigilin, Wiley Interdiscip rev, RNA 8 (2017) 1–15, https://doi.org/10.1002/wrna.1448.

[14] J.D. Richter, X. Zhao, The molecular biology of FMRP: new insights into fragile X syndrome, Nat. Rev. Neurosci. 22 (4) (2021) 209–222, https://doi.org/10.1038/s41583-021-00432-0.

[15] R.J. Hagerman, E. Berry-Kravis, H.C. Hazlett, D.B. Bailey, H. Moine, R.F. Kooy, F. Tassone, I. Gantois, N. Sonenberg, J.L. Mandel, P.J. Hagerman, Fragile X syndrome, Nat. Rev. Dis. Primers 3 (2017) 17065, https://doi.org/10.1038/NRDP.2017.65.

[16] D.D. Protic, R. Aishworiya, M.J. Salcedo-Arellano, S.J. Tang, J. Milisavljevic, F. Mitrovic, R.J. Hagerman, D.B. Budimirovic, Fragile X syndrome: from molecular aspect to clinical treatment, Int. J. Mol. Sci. 23 (2022) 1935, https://doi.org/10.3390/ijms23041935.

[17] C.J. Westmark, Fragile X and APP: a decade in review, a vision for the future, Mol. Neurobiol. 56 (2019) 3904–3921, https://doi.org/10.1007/s12035-018-1344-x.

[18] A. Quartier, H. Poquet, B. Gilbert-Dussardier, M. Rossi, A.S. Casteleyn, V. Des Portes, C. Feger, E. Nourisson, P. Kuentz, C. Redin, J. Thevenon, A.L. Mosca-Boidron, P. Callier, J. Muller, G. Lesca, F. Huet, V. Geoffroy, S. El Chehadeh, M. Jung, B. Trojak, S. Le Gras, D. Lehalle, B. Jost, S. Maury, A. Masurel, P. Edery, C. Thauvin-Robinet, B. Gérard, J.L. Mandel, L. Faivre, A. Piton, Intragenic FMR1 disease-causing variants: a significant mutational mechanism leading to fragile-X syndrome, Eur. J. Hum. Genet. 25 (2017) 423–431, https://doi.org/10.1038/ejhg.2016.204.

[19] S.C. Collins, S.M. Bray, J.A. Suhl, D.J. Cutler, B. Coffee, M.E. Zwick, S.T. Warren, Identification of novel FMR1 variants by massively parallel sequencing in developmentally delayed males, Am. J. Med. Genet. A 152 (A) (2010) 2512–2520, https://doi.org/10.1002/ajmg.a.33626.

[20] L.K. Myrick, P.Y. Deng, H. Hashimoto, Y.M. Oh, Y. Cho, M.J. Poidevin, J.A. Suhl, J. Visootsak, V. Cavalli, P. Jin, X. Cheng, S.T. Warren, V.A. Klyachko, Independent role for presynaptic FMRP revealed by an FMR1 missense mutation associated with intellectual disability and seizures, Proc. Natl. Acad. Sci. USA 112 (2015) 949–956, https://doi.org/10.1073/pnas.1423094112.

[21] L.K. Myrick, M. Nakamoto-Kinoshita, N.M. Lindor, S. Kirmani, X. Cheng, S. T. Warren, Fragile X syndrome due to a missense mutation, Eur. J. Hum. Genet. 22 (2014) 1185–1189, https://doi.org/10.1038/ejhg.2013.311.

[22] K. De Boulle, A.J.M.H. Verkerk, E. Reyniers, L. Vits, J. Hendrickx, B. Van Roy, F. Van Den Bos, E. de Graaff, B.A. Oostra, P.J. Willems, A point mutation in the FMR-1 gene associated with fragile X mental retardation, Nat. Genet. 3 (1993) 31–35, https://doi.org/10.1038/ng0193-31.

[23] H. Siomi, M. Choi, M.C. Siomi, R.L. Nussbaum, G. Dreyfuss, Essential role for KH domains in RNA binding: impaired RNA binding by a mutation in the KH domain of FMR1 that causes fragile X syndrome, Cell 77 (1994) 33–39, https://doi.org/10.1016/0092-8674(94)90232-1.

[24] J.A. Suhl, S.T. Warren, Single-nucleotide mutations in FMR1 reveal novel functions and regulatory mechanisms of the fragile X syndrome protein FMRP, J. Exp. Neurosci. 2015 (2015) 35–41, https://doi.org/10.4137/JEn.s25524.

[25] D. Santorelli, F. Troilo, F. Fata, F. Angelucci, N. Demitri, G. Giardina, L. Federici, F. Catalano, A. Di Matteo, C. Travaglini-Allocatelli, Folding mechanism and aggregation propensity of the KH0 domain of FMRP and its R138Q pathological variant, Int. J. Mol. Sci. 23 (2022), https://doi.org/10.3390/ijms232012178.

[26] R. Valverde, I. Pozdnyakova, T. Kajander, J. Venkatraman, L. Regan, Fragile X mental retardation syndrome: structure of the KH1-KH2 domains of fragile X

mental retardation protein, Structure 15 (2007) 1090–1098, https://doi.org/10.1016/j.str.2007.06.022.
[27] K. Manalastas-Cantos, P.V. Konarev, N.R. Hajizadeh, A.G. Kikhney, M. V. Petoukhov, D.S. Molodenskiy, A. Panjkovich, H.D.T. Mertens, A. Gruzinov, C. Borges, C.M. Jeffries, D.I. Svergun, D. Frankea, ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis, J. Appl. Crystallogr. 54 (2021) 343–355, https://doi.org/10.1107/S1600576720013412.
[28] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A. A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature 596 (7873) (2021) 583–589, https://doi.org/10.1038/s41586-021-03819-2.
[29] V. Receveur-Bréchot, D. Durand, How Random Are Intrinsically Disordered Proteins? A Small Angle Scattering Perspective, Curr. Protein Pept. Sci. 13 (2012) 55–75, https://doi.org/10.2174/138920312799277901.
[30] C. Göbl, M. Resch, M. Strickland, C. Hartlmüller, M. Viertler, N. Tjandra, T. Madl, Increasing the chemical-shift dispersion of unstructured proteins with a covalent lanthanide shift reagent, Angew. Chem. Int. Ed. 55 (2016) 14847–14851, https://doi.org/10.1002/ANIE.201607261.
[31] E.B. Gibbs, E.C. Cook, S.A. Showalter, Application of NMR to studies of intrinsically disordered proteins, Arch. Biochem. Biophys. 628 (2017) 57–70, https://doi.org/10.1016/J.ABB.2017.05.008.
[32] H.A. Lewis, K. Musunuru, K.B. Jensen, C. Edo, H. Chen, R.B. Darnell, S.K. Burley, Sequence-Specific RNA Binding by a Nova KH Domain: Implications for Paraneoplastic Disease and the Fragile X Syndrome, Cell 100 (3) (2000) 323–332, https://doi.org/10.1016/s0092-8674(00)80668-6.
[33] E.L. Starke, K. Zius, S.A. Barbee, FXS causing missense mutations disrupt FMRP granule formation, dynamics, and function, PLoS Genet. 18 (2022), https://doi.org/10.1371/JOURNAL.PGEN.1010084.
[34] D. Di Marino, T. Achsel, C. Lacoux, M. Falconi, C. Bagni, Molecular dynamics simulations show how the FMRP Ile304Asn mutation destabilizes the KH2 domain structure and affects its function, J. Biomol. Struct. Dyn. 32 (2014) 337–350, https://doi.org/10.1080/07391102.2013.768552.