Dynamics of the intrinsically disordered inhibitor IF7 of glutamine synthetase in isolation and in complex with its partner

José L. Neira[a,b,*], Maria Grazia Ortore[c,*], Francisco J. Florencio[d], M. Isabel Muro-Pastor[d] and Bruno Rizzuti[e]

[a]*Instituto de Biología Molecular y Celular, Universidad Miguel Hernández, Elche (Alicante), Spain;* [b]*Instituto de Biocomputación y Física de Sistemas Complejos (BIFI), Joint Units IQFR-CSIC-BIFI, and GBsC-CSIC-BIFI, Universidad de Zaragoza, Zaragoza, Spain;* [c]*Department of Life and Environmental Sciences, Marche Polytechnic University, Via Brecce Bianche, 60131 Ancona, Italy*; [d]*Instituto de Bioquímica Vegetal y Fotosíntesis, CSIC-Universidad de Sevilla, Seville, Spain;* [e]*CNR-NANOTEC, Licryl-UOS Cosenza and CEMIF.Cal, Department of Physics, University of Calabria, Via P. Bucci, Cubo 31 C, 87036 Arcavacata di Rende, Cosenza, Italy.*

Short title: Dynamics of IF7.

Corresponding authors addresses: José L. Neira, Instituto de Biología Molecular y Celular, Universidad Miguel Hernández, Avda. del Ferrocarril s/n, 03202, Elche (Alicante), Spain. Tel: + 34 966658459. Fax: +34 966658758. E-mail: jlneira@umh.es. Maria Grazia Ortore, Department of Life and Environmental Sciences, Marche Polytechnic University, Via Brecce Bianche, 60131 Ancona, Italy. Tel: +39 0712204608. E-mail: m.g.ortore@univpm.it

Abbreviations used: CSA, chemical shift anisotropy; GS, glutamine synthetase; HSQC, heteronuclear single quantum coherence; IDP, intrinsically disordered protein; IF, inactivating factor; IF7, *Synechocystis* sp. PCC 6803 65-residue-long inactivating factor of GS; IF17, *Synechocystis* sp. PCC 6803 149-residue-long inactivating factor of GS; MD, molecular dynamics; ns-ms, nanosecond-to-millisecond; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; ps-ns, picosecond-to-nanosecond; RSDA, reduced

spectral density approach; SAXS, small angle X-ray scattering; TSP, sodium trimethylsilyl [2,2,3,3-$^{2}H_4$] propionate.

ABSTRACT

Glutamine synthetase (GS) catalyzes the ATP-dependent formation of glutamine from glutamate and ammonia. The activity of *Synechocystis* sp. PCC 6803 GS is regulated, among other mechanisms, by protein-protein interactions with a 65-residue-long, intrinsically disordered protein (IDP), named IF7. IDPs explore diverse conformations in their free states and, in some cases, in their molecular complexes. We used both nuclear magnetic resonance (NMR) at 11.7 T and small angle X-ray scattering (SAXS) to study the size and the dynamics in the picoseconds-to-nanosecond (ps-ns) timescale of: (i) isolated IF7; and (ii) the IF7/GS complex. Our SAXS findings, together with MD results, show: (i) some of the possible IF7 structures in solution; and, (ii) that the presence of IF7 affected the structure of GS in solution. The joint use of SAXS and NMR shows that movements of each amino acid of IF7 were uncorrelated with those of its neighbors. Residues of IF7 with the largest values of the relaxation rates (R_1, R_2 and η_{xy}), in the free and bound species, were mainly clustered around: (i) the C terminus of the protein; and (ii) Ala30. These residues, together with Arg8 (which is a hot-spot residue in the interaction with GS), had a restricted mobility in the presence of GS. The C-terminal region, which appeared more compact in our MD simulations of isolated IF7, seemed to be involved in non-native contacts with GS that help in the binding between the two macromolecules.

Keywords: intrinsically disordered proteins, NMR, protein dynamics, SAXS.

1. INTRODUCTION

Glutamine synthetase (GS) is the first enzyme involved in nitrogen assimilation in bacteria: conversion of inorganic to organic nitrogen by incorporating ammonium to glutamate leading to glutamine formation [1,2], in the presence of Mn^{2+} and Mg^{2+}. Based on sequence similarity and their functional roles, there are three types of GSs: GS type I, the main form present in prokaryotes; GS type II, mostly found in eukaryotes; and GS type III, which is only present in some prokaryotes. In cyanobacteria, GS type I function is modulated at transcriptional and posttranscriptional levels, although the classical adenylylation/deadenylylation regulation of GS has not been observed [2]. However, a protein-protein post-translational regulatory mechanism, involving a 65-residue-long protein (inactivating factor 7 or IF7) and a 149-residue-long one (inactivating factor 17 or IF17), has also been found [2-5]. Analyses of mutant strains lacking IF7, IF17 or both inactivating factors have shown that the two proteins are involved in GS inactivation *in vivo*. A maximal level of inactivation *in vitro* occurs when both proteins are present, although the presence of either IF is enough for GS inactivation [2,3].

Isolated IF7 from *Synechocystis* sp. PCC 6803 is an intrinsically disordered protein (IDP) [6]. Isolated IDPs do not fold into a three-dimensional structure in isolation, adopting instead an ensemble of inter-converting conformations, although in many cases they fold upon binding to their partners [7,8]. The IF7/GS binding reaction is complex, electrostatically-driven, and it involves several GS-bound species of IF7 with different conformations [9]. NMR characterization of isolated IF7 indicates that its N terminus populates helical conformations (residues Thr3-Arg13, Met25-Ala29 and Phe41-Thr43); some of these polypeptide patches comprise the amino acids that initially interact with GS [9], as further tested by site-directed mutagenesis [10], and thus, constitute the hot-spot region of the protein: Arg8, Arg21 and Arg28. Mechanisms by which IDPs recognize their interacting partners are important to describe the conformational landscape of the resulting complex and, in addition, that of the free protein. The conformational space sampled by IDPs is large, and use of complementary techniques is required. Among other techniques, IDPs have been successfully characterized by SAXS and NMR [11-14]. SAXS monitors

long-range interactions among distant polypeptide chain regions, as it can map the whole protein size, and it is sensitive to its overall conformational properties. On the other hand, NMR reports on both local- and long-range interactions, as well as on the timescales of inter-conversion among the different conformers [15]. Hence, while NMR provides a powerful approach for the identification of particular residue flexibility, SAXS, despite its lower resolution, provides structural information on the protein shape and determines how the polypeptide chain stiffness can be affected by different solution conditions [16] and/or the presence or absence of partners.

We were interested in understanding how molecular recognition of IF7 proceeds at the atomic level and how its dynamics is changed upon the binding of GS. We have previously shown that the apparent dissociation constant, K_d, for the complex IF7/GS is 0.3 ± 0.1 μM [6]. There have been recent studies on IDPs that fold upon binding, investigating their dynamics when isolated [17,18] and in the presence of their partners [19-22]. In this work based on SAXS data, we show further evidence that isolated IF7 had a conformation consistent with a disordered polypeptide chain model, but it does not adopt a worm-like model [21]. Furthermore, by combining SAXS measurements with molecular dynamics (MD) simulations, we demonstrate that IF7 did not have a globular shape and that it spanned a variety of different conformations in solution. On the basis of our SAXS data, we can infer that IF7 caused a slight but noticeable compaction in the structure of GS when bound. The ps-ns timescale dynamics sampled by NMR indicates that IF7 was highly mobile even when interacting with GS. The mobility in presence of the latter was decreased especially around residues Arg8, Ala30 and in the C-terminal polypeptide patch. Either in the absence or in the presence of GS, the residues of IF7 did not move collectively, as indicated by the reduced spectral density approach (RSDA), in agreement with the impossibility of fitting SAXS data to a worm-like model. The C-terminal region of IF7, although it has not been recognized as a hot-spot by site-directed mutagenesis, seems to be important to promote non-native contacts with GS and to help in the complex formation, acquiring a more compact conformation than the rest of the protein, as shown by our MD simulations.

2. MATERIALS AND METHODS

2.1. Materials

Deuterium oxide and isopropyl-β-D-1-thiogalactopyranoside were obtained from Apollo Scientific (Stockport, UK). Sodium trimethylsilyl [2,2,3,3-$^{2}H_4$] propionate (TSP), deuterated Tris acid, His-Select HF nickel affinity gel and $^{15}NH_4Cl$ were from Sigma-Aldrich (Madrid, Spain). Dialysis tubing, with a molecular weight cut-off of 3500 Da, was from Spectrapor (VWR, Barcelona, Spain). Amicon centrifugal devices with a molecular weight cut-off of 3500 Da were from Millipore (Barcelona, Spain). Standard suppliers were used for all other chemicals. Water was deionized and purified on a Millipore system.

2.2. Protein expression and purification

GS and IF7 were expressed and purified by using a His-tagged vector [3,5,6,9], with Ni-resin and an additional gel-filtration step performed with a Superdex GS-200 (GE Healthcare) and a Superdex G75, respectively. Both columns were connected to an AKTA-FPLC (GE Healthcare), and we followed absorption at 280 nm. Protein stocks were run in SDS-PAGE gels and found to be > 97% pure. Protein concentration was determined from the absorbance of individual amino acids (each GS monomer has 8 tryptophans and 23 tyrosines, and IF7 has a single tryptophan) [23]. ^{15}N-labelled IF7 was expressed and purified as described [9].

The His-tag of the IF7 species used in our experiments was not removed, as we have previously shown that its presence does not affect the binding between both proteins: the same apparent affinity constant was obtained either in the presence or in the absence of the His-tag [6,9] and using such different techniques as fluorescence and biolayer interferometry. The His-tag of GS used in our experiments was not removed either.

2.3. NMR spectroscopy

The NMR experiments were acquired at 298 K on a Bruker Avance DRX-500 (Bruker GmbH, Germany) spectrometer equipped with a triple-resonance probe and z-gradients. The temperature of the probe was calibrated with methanol [24]. All experiments were carried out in 50 mM Tris (pH 7.2), with a protein

concentration of 300 μM of IF7 and final concentration of GS of 8 μM. This pH value was chosen because at more acidic values GS precipitates. The stoichiometry [25] and the kinetic pathway of the binding reaction [9] between GS and IF7 are complex; the stoichiometry is, in protomer units, [IF7]/[GS] = 0.5 [25], and the kinetic pathway seems to involve several steps [9]. Furthermore, we have previously shown that, at this IF7 concentration [9], there is a compromise between the largest possible number of IF7 signals observed and the progress of the binding [9]. Other NMR dynamic studies of IDPs also report the use of sub-stoichiometric amounts of the largest partner to avoid excessive signal broadening and in some cases. precipitation ([22] and references therein).

NMR relaxation data included the acquisition of ^{15}N-R_1, ^{15}N-R_2, ^{15}N-η_{xy} (the ^{1}H-^{15}N dipolar/^{15}N chemical shift anisotropy (CSA) cross-correlated relaxation rate) and ^{1}H-^{15}N NOE experiments. All the relaxation measurements were determined in an interleaved manner to ensure that the experimental conditions were the same for the different relaxation delays. The ^{15}N-R_1, ^{15}N-R_2 and ^{1}H-^{15}N NOE experiments were acquired by using gradient pulse sequences developed by Kay and co-workers [26]. The T_1 (=1/R_1) relaxation time was measured with typically 10 inversion-recovery delays, varying from 5 to 850 ms; T_2 (=1/R_2) was determined by collecting 8 time points ranging from 10 to 400 ms. Two relaxation delays were randomly repeated in each series of rates for each sample (both isolated IF7 and IF7/GS complex) to ensure reproducibility. For the T_1 and T_2 pulse sequences, the delay between scans was 1 s. The ^{1}H-^{15}N NOE experiment was measured by recording interleaved spectra with or without proton saturation. The ^{1}H-^{15}N NOE spectrum recorded in the presence of proton saturation was acquired with a saturation time of 10 s. The ^{1}H-^{15}N NOE spectrum recorded without proton saturation incorporated a relaxation delay of 10 s. The ^{1}H-^{15}N NOE experiment (with and without saturation) was repeated twice for each sample. Spectra were recorded with 2048 × 180 complex matrices in the F_2 and F_1 dimensions, respectively, with typically 32 scans (NOE experiment) and 16 scans (R_1 and R_2 experiments) *per* F_1 experiment. Spectral widths of 1935 and 7000 Hz were used in F_1 and F_2, respectively; the ^{15}N carrier was set at 120 ppm and that of ^{1}H was set on the water signal in all the experiments.

The η_{xy} value was measured with the pulse sequences reported [27], with the same spectral widths described above, and the ^{15}N carrier was set at 120 ppm in all the experiments. The delay between scans was 2 s for each specific relaxation time in both experiments. One set of cross-relaxation (which leads to I_{cross}) and auto-relaxation (which yields I_{auto}) experiments were acquired for each of the relaxation delays. Both sets of experiments consisted typically of 256 experiments in the F_1 dimension and 2 K data points in the F_2 dimension; either 16 or 32 scans were acquired *per* F_1 experiment. The relaxation delays, T (= 2Δ), were 25, 50, 75 and 100 ms; to ensure the reproducibility in each series of samples one of those cross- and auto-relaxation delays were repeated.

All the spectra were zero-filled in the F_1 dimension (two to four times), and processed by using a shifted sine square window function in Topspin 2.1 software package (Bruker). The same window function was used throughout all the T_1 and T_2 experiments. Intensities of the cross-peaks were measured with Topspin 2.1. The R_1 and R_2 values were determined by fitting the measured peak intensities to a two-parameter function: $I(t) = I_0 \exp(-tR_{1,2})$, where $I(t)$ is the peak intensity after a relaxation delay t, and I_0 is the intensity at time zero; uncertainties in the relaxation rates were errors in the fittings to the above equation. The value of η_{xy} was derived from the one-parameter fitting: $\frac{I_{cross}}{I_{auto}} = \tanh(\eta_{xy} T)$; uncertainties on the η_{xy} values were errors in the fitting of data to such equation. Data were fitted to both equations with Kaleidagraph (Synergy Software, Reading, PA, USA). The steady-state ^1H-^{15}N NOE values were determined from the ratios of the peak intensities with and without proton saturation (i.e. NOE = I_{sat}/I_{nonsat}). The uncertainties of the NOE values were determined from the measured background noise levels during repeated experiments.

The experimentally measured R_2 rates can be expressed by: $R_2 = R_2^0 + R_{ex}$, where R_{ex} is the transverse relaxation rate caused by conformational processes (nanosecond-to-millisecond (ns-ms) timescale); and R_2^0 is the transverse relaxation rate, which does not contain any influence from ns-ms motions. Then, to

quantify the value of R_{ex}, it is necessary to obtain R_2^0 independently. The value of R_2^0 was obtained from the value of η_{xy}, as [28,29]:

$$R_2^0 = 1.2589 \cdot \eta_{xy} + 1.3[(NOE - 1) \cdot R_1] \cdot \frac{\gamma_N}{\gamma_H} \quad (1),$$

where γ_N and γ_N are the gyromagnetic ratios of ^{15}N and ^{1}H, respectively. We used this value of R_2^0 to estimate the spectral density function at $\omega = 0$, $J(0)$ (Eq. (2), see below), as it has been described when the RSDA is applied to IDPs [30]. In that way, $J(0)$ is insensitive to chemical exchange, and hence it will only report on ps-ns dynamics.

2.4. Theoretical background for the reduced spectral density approach (RSDA)

The T_1 and T_2 relaxation times and the NOE enhancement of an amide ^{15}N nucleus are dominated by the dipolar interaction of the ^{15}N nucleus with its attached proton and by the CSA. The energy of the CSA and the dipolar interaction has a constant value over the whole spin ensemble [31]. The spectral density function, $J(\omega)$, expresses how the energy of all spins is distributed over all the spectrum of possible frequencies, ω. The measured rates for each N-H bond vector (backbone amide bond) can be approximated (the so-called 'reduced spectral density mapping approach') by the values of the spectral density function at $\omega = 0$, ω_N and $0.87\omega_H$ according to [32-34]:

$$J(0) = \frac{(6R_2 - 3R_1 - 2.72\sigma_{NH})}{(3d^2 + 4c^2)} \quad (2),$$

$$J(\omega_N) = \frac{(4R_1 - 5\sigma_{NH})}{(3d^2 + 4c^2)} \quad (3),$$

$$J(0.87\omega_H) = \frac{(4\sigma_{NH})}{(5d^2)} \quad (4), \text{ and}$$

$$\sigma_{NH} = R_1(NOE - 1)(\gamma_N / \gamma_H) \quad (5),$$

where $c = \frac{(\omega_N)}{\sqrt{3}}(\sigma_{\parallel} - \sigma_{\perp})$ and $d = \frac{(\mu_0 h \gamma_N \gamma_H)}{(8\pi^2 \langle r \rangle^3)}$; μ_0 is the permeability constant of the free space; h is the Planck constant; ω_N is the Larmor frequency of ^{15}N; ω_H is the Larmor frequency of ^{1}H; $\langle r \rangle$ is the length of

the amide bond vector (1.02 Å); and $\sigma_{\parallel}$ and $\sigma_{\perp}$ are, respectively, the parallel and perpendicular components of the CSA tensor ($\sigma_{\parallel} - \sigma_{\perp} = -160$ ppm for a backbone amide [35]). The uncertainties for each specific $J(\omega)$ were the quadrature-weighted sum of errors, which were derived from Eqs. (2) to (5), assuming that errors in the relaxation rates constants were independent [36]. It is important to indicate that $J(0)$ (Eq. (2)) was calculated from the value of R_2^0, which in turn was obtained from η_{xy} (Eq. (1)).

2.5. SAXS experiments

SAXS data were collected at the Austrian beamline at Elettra Synchrotron in Trieste, Italy [37]. Temperature was 298 K during all measurements. SAXS patterns were recorded using the 2D Pilatus3 1M detector system of total area 168.7 × 179.4 mm^2. Data were obtained as a function of the scattering vector Q defined as: $Q = \frac{(4\pi \sin(\theta))}{\lambda}$, where 2θ is the scattering angle, and λ = 0.995 Å, the X-ray wavelength. The incident and transmitted intensities were measured to obtain data on an absolute scale. The sample stage was an autosampler developed in the beamline, whose fixed measurement cell was a 1 mm thick quartz capillary. Each measurement was performed on at least 10 injections of very small sample volumes (15 μL), and was carried out four times for 20 s. Each SAXS spectrum acquisition was followed by 3 s of dead-time. This approach allowed us to minimize the effects of both potential inhomogeneity and radiation damage in the protein sample. After SAXS data comparison, we averaged all the measured scattering data corresponding to the same nominal sample. Normalized SAXS patterns were averaged azimuthally to obtain one-dimensional profiles of scattered intensity. Buffer measurements were always performed before and after sample measurements, and the buffer contribution was subtracted. SAXS curves are obtained from solutions whose concentration was 800 μM for isolated IF7 and 40 μM for GS, whereas the complex of the two proteins was obtained at a concentration of 38 μM for both proteins. All the experiments were repeated at half of their respective concentrations to verify that no oligomeric equilibria affected our measurements (data not shown).

A basic fitting of the experimental macroscopic cross section, $\frac{d\Sigma}{d\Omega}(Q)$, was performed according to both Guinier [38] and Debye approximations [39]. Guinier equation to estimate the gyration radius, R_g, of globular species in solution is:

$$\frac{d\Sigma}{d\Omega}(Q) = \frac{d\Sigma}{d\Omega}(0)\ e^{\left(\frac{-R_g^2 Q^2}{3}\right)} \quad (6).$$

The Debye equation, which is described next, is more appropriate for disordered polymer chains:

$$\frac{d\Sigma}{d\Omega}(Q) = \frac{d\Sigma}{d\Omega}(0)\frac{2}{\left(R_g^4 Q^4\right)}\left(e^{\left(R_g^2 Q^2\right)} - 1 + R_g^2 Q^2\right) \quad (7),$$

where R_g is the estimated gyration radius in the Debye approximation. Further analysis, based on a set of models of the protein structure obtained by MD simulations, was performed by considering SAXS data as the weighted sum of form factors:

$$\frac{d\Sigma}{d\Omega}(Q) = \sum_{i=1}^{N} x_i \frac{d\Sigma}{d\Omega}(Q)\Big|_i \quad (8),$$

where the index i corresponds to a single protein conformation; N is the number of models in the set; x_i is the relative weight of the *i-th* structure, and $\frac{d\Sigma}{d\Omega}(Q)\Big|_i$ is the macroscopic differential scattering cross section obtained by the calculation of the form factor of the *i-th* structure. The calculations of form factors and of the theoretical fittings were obtained using the GENFIT software [40]. The contribution of the solvation shell, whose mass density is supposed to be different from the one of the bulk solvent, was explicitly taken into account. The adimensional parameter considered in the calculation, r_d, is the mass density of the solvation shell relative to the mass density of the bulk water [41]. Due to moderately low sample concentration and to the range of Q monitored in our experiments, the structure factor was considered equal to one.

2.6. Molecular simulations of isolated IF7

The molecular topology of IF7 was built using the Amber 99SB-ILDN force field [42], with protein titratable residues protonated to mimic neutral pH and four Cl^- counterions added to obtain a system with overall zero charge. According to a protocol we previously described [43], IF7 was first built in extended conformation and then collapsed in an MD run performed with the simple 3-point water model TIP3P [44], up to reach R_g = 24 Å in about 10 ns.

Eight MD simulations were performed, differing by the type of water model and the starting gyration radius. In particular, either the two 4-point water models TIP4P-D [45] or OPC [46] were alternately used, whereas R_g was initially set to 24, 25, 26 or 27 Å. We note that, although we had not previously employed either of these two water models to study IF7, we have already successfully used both to investigate other IDPs in simulation [47, 48]. The two extreme values of the starting protein radius corresponded to the experimental estimate for IF7 obtained using either the Guinier approximation (R_g = 24 Å) or the Debye approximation (R_g = 27 Å) to fit our SAXS results (see section 3.2).

All simulations were run using the package GROMACS [49] for 10 ns in the isothermal-isobaric ensemble. Reference values and coupling times for the thermostat and barostat, treatment of the electrostatics and van der Waals interactions, and time steps for sampling and integration of the equation of motions, were as previously described [50, 51].

3. RESULTS

3.1. Dynamics of IF7 under several conditions as mapped by NMR

The relaxation rates of 23 out of 63 residues (65 total amino acids, minus 1 proline and the N-terminal residue) of IF7 could be measured by using NMR at pH 7.2 (Fig. S1), based on the previous assignment (BRMB accession number 25921 [9]). The same set of residues was observed, although their cross-peaks were broader, in the presence of sub-stoichiometric amounts of GS (Fig. S1). The small number of resonances identified did not map the whole polypeptide chain, but the observation of probes along different regions of the polypeptide chain allowed us to get reliable conclusions.

3.1.1. Dynamics of isolated IF7

The relaxation rates are shown in Fig. 1 and Table ST1. The R_1 values appeared to be between 2.25 and 7.08 s^{-1}, whereas those of R_2 ranged from 1.67 to 4.99 s^{-1}. On the other hand, the η_{xy} values ranged from 1.4 to 7.4 s^{-1}. The largest values of R_1 (Fig 1 A), and those of η_{xy} (Fig. 1 C), were observed for residues located at both termini of the polypeptide chain. By contrast, the R_2 values (Fig. 1 B) were quite uniform along the polypeptide chain.

The NOE values of residues of the isolated IF7 were below 0.52 (the value observed for Ala29) (Fig. 1 D). Keeping in mind that at 11.7 T the maximum value for the NOE is 0.70 [52] in an isotropically tumbling molecule without internal motions, we must conclude that isolated IF7 had a highly flexible backbone. In general terms, residues around Asp40 and those located at the N terminus of the protein were more flexible (i.e., they had lower NOE values).

NOEs are sensitive markers of local structure in IDPs [52], but also η_{xy} (and R_2 rates, when they are not affected by contributions of R_{ex}) can be used to characterize the presence of local order [52,53]. The R_2 (Fig. 1 B) and η_{xy} (Fig. 1 C) rates for isolated IF7 followed a similar trend: they had larger values for residues around Leu10 (with values of 5 s^{-1} for R_2, and 3.29 s^{-1} for η_{xy}), Thr44, and the region around Gln50. The NOE values (Fig. 1D) around Gln50 were among the highest measured, suggesting that the order around those polypeptide patches was larger than the in the rest of the polypeptide chain, and the mobility,

as shown by the values of R_2, was very high on the ps-ns timescale. However, for Leu10 and Thr44, the NOE values were among the smallest measured. This opposite tendency in R_2 and NOE values could look unusual, but there are some IDPs or unfolded proteins in which the NOE value of a particular residue remains unaltered, when compared to those of its neighbors, while the R_2 value of that residue increases (see, for instance, [54,55]). Furthermore, mutations can alter the R_2 rates in unfolded proteins, while the NOE behavior does not mirror those changes [56].

3.1.2. Dynamics of the IF7/GS complex

The addition of GS in submicromolar concentrations to IF7 did not change the chemical shifts of the cross-peaks in the HSQC spectrum of IF7 (Fig. S1, Table ST2), although it induced signal broadening as it happened in the cross-peaks of other IDPs when bound to their partners [22,57, and references therein]. In general, we observed an increase in all the relaxation rates of IF7, when compared to those measured in the absence of GS (Fig. 2).

The R_1 values ranged from 1.77 to 10.98 s^{-1}; the behavior of this rate was completely different from that observed in isolated IF7 (Fig. 1 A), showing the largest values of the relaxation rate not only at both termini, but also around residues Ala30 and Asp40 (Fig. 2 A). On the other hand, the R_2 values (Fig. 2 B) ranged from 1.73 to 17.6 s^{-1}, and the behavior (in contrast to what we observed in the absence of GS, Fig. 1 B) was not uniform: residues Ser27, Ala30, Thr44, Gln46, Ser51, Phe53, Thr56 and Asp58, had the largest R_2 values (and so the largest variations), when compared to those of isolated IF7. Calculation of R_{ex} from R_2^0 (Eq. (1)) for the residues above yielded values larger than 6 s^{-1} except for Ser51 (0.5 s^{-1}) and Thr56 (1.8 s^{-1}). Finally, the trend observed for the η_{xy} values was similar to that of R_1 (i.e., large values at both termini and around Ala30, Fig. 2 C), ranging from 1.02 to 9.53 s^{-1} (Fig. 2).

The NOE values (Fig. 2 D), although presenting differences in specific residues (Tables S1 and S2), were similar to those measured for each amino acid in the absence of GS (Fig. 1 D); for instance, Ala29 had also the largest value: 0.52. All the NOE values showed the same tendency described in the absence of GS, indicating that IF7 kept its disordered state when it is bound.

3.2. The structure of IF7 and its complex with GS as mapped by SAXS

3.2.1. Structure of isolated IF7

SAXS experiments complement and extend NMR results by providing additional information on conformational properties and the size of complexes formed by IDPs, which are governed by both local and long-range interactions. SAXS data of isolated IF7 are shown a Kratky plot (Fig. 3). In the Kratky representations (reporting $(\frac{d\Sigma}{d\Omega}(Q)\times Q^2)$ *versus* Q), a bell-shaped peak indicates that the particles in solution are likely globular or compact, whereas a plateau or linear trend characterizes unfolded/extended conformations [38]. We did not observe a well-defined peak in the Kratky plot of isolated IF7 data, proving that the protein did not have a globular shape. In a first approximation, SAXS data for IF7 were fitted by using either the Guinier or Debye equation (Fig. S2). The two approaches provided a radius of gyration R_g = 24.2 ± 0.2 Å (in the Guinier approach) and R_g = 26.95 ± 0.05 Å (in the Debye approach). The Debye equation is more appropriate for disordered chains than the Guinier one, and accurately fitted our experimental data, as shown by the continuous line (Fig. 3).

Although the fitting obtained with this approach did have small residuals and a good χ^2 value, IF7 cannot be modelled as a single structure with a well-defined R_g. In fact, as many other IDPs, IF7 should be more properly described in terms of a collection of structures spanning a variety of different sizes in solution. For this reason, we performed MD simulations in explicit water to obtain a set of protein structures that, although being far from representing a complete and rigorously defined statistical ensemble, provide a more dynamic picture of IF7. Due to the importance of the solvent for simulating disordered protein conformations in MD, two distinct sets of structures were obtained by using either the IDP-specific water model TIP4P-D [45] or the all-purpose model OPC [46], which have both provided satisfactory results in reproducing unfolded protein states [58, 59].

To test the goodness of our approach, only one representative structure for each gyration radius (i.e., a total of thirteen structures with R_g = 18, 19, ..., 30 Å) was included in each set, thus obtaining the two-fold effect of limiting the complexity in the calculations and avoiding a potential over fitting of our

experimental values. The continuous lines in Fig. 4 correspond to the fitting of SAXS data obtained by using either set of structures. The conformations obtained with both water models did correctly fit the experimental data. The accuracy is similar in both cases, as no significant difference was evident in the agreement with experimental data in terms of the χ^2. The weights, x_i, obtained for each protein conformation, as described in Eq. (8), are reported in Table ST3 for both the IDP-specific water model TIP4P-D and for the model OPC. It is interesting to note that IF7 conformations corresponding to the same R_g, but arising from the two different MD approaches, did not present similar weights in the fitting of the experimental SAXS data. This result confirms the evidence that the SAXS fingerprint provides a wider information in respect to the overall protein dimensions (i.e, R_g), determining protein conformations that more accurately describe the ensemble of IF7 in solution.

Figure 5 summarizes the most relevant IF7 conformations observed in our MD simulations. In particular, Fig. 5 A reports a superposition of the first three structures showing the highest relative weight in the fitting procedure of SAXS data for both simulation sets (the same structures are shown in a separate and more detailed view as Figure S3). The R_g, in decreasing order or relevance, was 18, 24 and 29 Å for the TIP4P-D set, and 22, 27 and 20 Å for the OPC set. These findings clearly indicate that SAXS data resulted from contributions to the scattering of IF7 structures with different sizes; this observation is also in agreement with the distribution of the R_g obtained in our MD simulations (see above). In fact, Fig. 5 B reports the frequency of R_g obtained with the water model TIP4P-D, which is considered the best solvent model to correctly avoid over-compaction in the simulated structure of IDPs [45, 61]. The histogram shows that a variety of gyration radii are readily accessible in the conformational space of isolated IF7 in solution. The structures we obtained (Fig. 5 A) were on average less compact than those we have previously observed for IF7 under acidic conditions [43]. However, all these structures (i.e., the ones belonging to the current TIP4P-D and OPC simulation sets, as well as those previously described at low pH and acquired with the less specific water model TIP3P [44]), share a common feature consisting in an

asymmetric shape, with the N-terminal region more elongated and mobile, and the C-terminal region more globular and compact.

3.2.2. Structure of IF7/GS complex

GS plays a key role in the metabolism of nitrogen, and consequently its structure has been already intensively investigated by using crystals [62,63]. A crystallographic structure of GS from *Synechocystis sp*. PCC 6803 has been released (3NG0 entry in the PDB), but so far no experiments have tested its structure in solution; in this work, we present SAXS data on isolated GS in solution. SAXS experimental data obtained for isolated GS are shown in Fig. 6, together with the theoretical fit (black continuous curve) provided by the calculation of the protein form-factor on the basis of the available crystallographic structure (3NG0), considering also the typical contribution of a hydration shell [41]. It can be easily appreciated that the reconstruction of the form factor fitted the experimental data very satisfactorily even at higher Q values corresponding to the highest structural resolution, thus indicating that structure evidenced by X-ray is similar to that present in solution.

The SAXS experiment with a stoichiometric ratio 1:1 of GS and IF7 was acquired to provide insight into the interaction of the two proteins. Experimental data are shown and scaled in Fig. 6 (cyan continuous line) and a theoretical curve obtained by the sum of the SAXS fingerprints of both GS and IF7 alone, weighted according to their respective concentration (equimolar) in the sample, and assuming no complex formation, is reported for comparison (yellow continuous line); the same experimental data are also reported in unscaled form in Fig. S4. Because GS and IF7 sizes are very different, it is clear that GS scattering contribution dominates SAXS data. However, the fact that the experimental curve, corresponding to the simultaneous presence of GS and IF7 in solution, was different from the sum of the single scattering contributions of the two isolated proteins, confirms the presence of a well-defined assembly arrangement between IF7 and GS. Although the presence of the IF7/GS complex in solution is well-documented [6,9,25], our SAXS results provide a direct estimation of some of its structural features at a low resolution. In fact, a Guinier fitting approach (Fig. S5) provides an assessment of the average

dimensions of both isolated GS and of the GS/IF7 complex in solution. In particular, according to the Guinier approach, the values obtained were $R_g = (56.52 \pm 0.08)$ Å for isolated GS, we found the $R_g = (53.1 \pm 0.1)$ Å for the complex. Guinier fitting approach was performed within the limit $R_g * Q_{max} \leq 1.3$, where Q_{max} is the maximum value of Q for which the fitting has been performed. At higher Q values (Fig. S5), the fitting fails because both proteins are not exactly spherical. The Guinier approximation considers a fixed contrast between the protein in solution and the bulk solvent and hence, although results indicate just a slight decrease of the dimension of IF7/GS complex with respect to GS, we can ascribe this change to a compaction in the structure of GS due to the interaction with IF7.

3.3. Reduced spectral density approach (RSDA) for isolated IF7 and its complex with GS

3.3.1. Analyses of the spectral density functions

The RSDA is used to obtain additional information on protein dynamics for IF7 in isolation or when it was bound to GS. The RSDA provides insight into the motion of the N-H bond vector by the calculation of the $J(\omega)$ values at three frequencies: 0, ω_N and $0.87\omega_H$ [30, 32-34]. It could be thought that the RSDA can be only used for well-folded proteins, but since it does not rely on any specific assumption about the overall rotation of the molecule, it has been recently applied to interpret the relaxation measurements in IDPs [20,21,30, 64-67]. In fact, this approach has been shown to be successful in describing the dynamics behavior of IDPs, without the need to use an extensive equipment for long periods of time [20,21,68]. In the following, we describe the results for IF7 in the two explored conditions: in isolation and in the GS-bound species.

***J*(0).** The $J(0)$ values were obtained from the R_2^0 values (which in turn they were calculated from the η_{xy} rate, Eq. (1)). A small $J(0)$ value indicates internal flexibility of the N-H amide bond: the smaller the value of $J(0)$, the larger the flexibility of the N-H bond vector. Conversely, a large $J(0)$ value indicates residues with slower fluctuations. There was a great variation among the $J(0)$ values of the different residues (Fig. 7 A, black dots) for isolated IF7, but in general the values were very small for most of the residues (< 1 ns rad^{-1}): only Gln50 and Ser51 had values around 1.5 ns rad^{-1}. It is interesting to note that these residues

are at the C-terminus of the protein, where in our MD simulations (Fig. 5 A) we observed compaction of the polypeptide chain. All these findings indicate a large flexibility over most of the polypeptide chain.

Upon complex formation, a slower rotational diffusion and therefore an increase in the $J(0)$ value should be expected [37]. In the presence of GS, $J(0)$ increased in general, indicating that the mobility of IF7 in the presence of GS was reduced, due to the larger size of the complex. Residues at the C terminus of the protein, as well as Lys19 and Ser27 showed the largest values (on average two times larger than those for other amino acids, Fig. 7 A, red dots), with distinct $J(0.87\omega_H)$ spectral density values (Fig. 7 C), especially for residues at the C-terminal region. This means that these residues were more rigid (with slower movement) in the presence of GS.

To summarize, the isolated IF7 had a large flexibility in the ns-ps timescale, which was reduced, but not quenched, in the presence of GS.

***J*(ω_N).** For small proteins, $J(\omega_N)$ usually decreases with an increase of the internal flexibility; however, for large proteins $J(\omega_N)$ increases with a larger protein flexibility [30, 32-34]. In isolated IF7, we observed larger values of $J(\omega_N)$ for Arg8 (at the N terminus), Ser51 and Asp58 (both at the C terminus of the protein), since most of the residues had a fairly constant, small value of $J(\omega_N)$ (Fig. 7 B, black dots). Therefore, considering IF7 as a small protein, such small value of $J(\omega_N)$ for the majority of the residues further pinpoints the high flexibility of the polypeptide chain.

Upon addition of GS, all the $J(\omega_N)$ increased, although the variations were not very large (on average, a 10-15 % variation), further benchmarking the large flexibility of the protein even when it was bound to GS. The largest variations were observed for Ser27, Glu38 and Asp58 (Fig. 7 B, red dots), indicating that these residues had a restricted mobility in the presence of GS.

***J*(0.87ω_H).** The high frequency $J(0.87\omega_H)$ value reports on fast (ps timescale) motions in flexible regions of proteins: the largest values are found in highly mobile regions, and the smallest for rigid polypeptide patches [32,33]. In general, for isolated IF7, the $J(0.87\omega_H)$ values were slightly larger at both main chain termini, whereas for residues in the middle of the polypeptide chain they had a fairly constant value (Fig. 7

C, black dots). In the presence of GS, the $J(0.87\omega_H)$ values increased, indicating that IF7 had a large flexibility even though it was bound to its partner. In the complex, the regions of IF7 with the highest values of $J(0.87\ \omega_H)$ were at Arg8 and at the C terminus of the protein (Fig. 7 C red dots). These results indicate that those regions were highly mobile, even when forming the complex, confirming the results obtained by analyzing the other two values of the spectral density function.

Recently, by using molecular simulations it has been possible to estimate the correlation times of the movements of polypeptide patches of some IDPs in isolation [16,17,55, 67,69,70]. We have also recently used the plots of the spectral density functions at different frequencies against each other to allow for an easier and direct identification of residues affected by ps-ns timescale motions in another IDP in the presence of some of its partners [21]. These plots allow one to assess changes in protein dynamics without assuming any explicit, physical model. The only assumption is that the value of the spectral density function at a particular frequency can be expressed as a linear combination of the values of the such function at other frequencies [32,33,70].

Then, we represented $J(0.87\omega_H)$ *versus* $J(0)$ to elucidate whether the motions involving IF7 backbone were collectively organized [65,66,70]. If the movements of all IF7 residues were to occur collectively (either in the absence or in the presence of GS), then both spectral density values should be linearly correlated [33]. Residues not involved in such collective motions should deviate from the linearity in the plots of one spectral density function at a particular frequency *versus* another value, or should appear clustered in different sections, along distinct straight lines, indicating independent motions of some polypeptide regions. Under both conditions, the plots of $J(0.87\omega_H)$ *versus* $J(0)$ for IF7 were not linearly correlated (Fig. 8), indicating that most of the residues did not move collectively. However, as the values of $J(0)$ and $J(0.87\omega_H)$ were different in the two environments, the uncorrelated motions of the same residue in the two conditions were different. That is, binding of GS changed the conformational movements of IF7, but the movements of each amino acid were uncorrelated with those of its neighbors.

4. DISCUSSION

Recent studies try to address the dynamics of various IDPs by using several NMR fields at different conditions [16,17,67,71]. Furthermore, the dynamics of IDPs has been also studied in the presence of some of their partners [18,19], although in most of the examples when the IDPs acquire a well-folded structure upon binding. However, we do not know how the dynamics of IDPs changes upon binding to their partners. In this work, we have tried to address these questions by using IF7, a 65-residue-long IDP, and studying the binding to its partner, GS, by using NMR measurements at a single field strength (11.7 T) and synchrotron SAXS. It is important to note that since with NMR we are measuring exclusively the dynamics of the backbone amide protons, we cannot speculate on the side-chain mobility of IF7 and its contribution to the mobility of either the isolated protein or in the complex with GS.

We have previously shown that the polypeptide patches comprising residues Thr3-Arg13, Met25-Ala29, and Phe41-Thr43 have a certain helical propensity [9]. In isolated IF7, except for Arg8, whose R_1 rate had a large value (Fig. 1 A), none of the other residues belonging to those polypeptide patches had a large value for any of the measured relaxation rates. On the other hand, Ala29 showed the largest NOE value (Figs. 1 D, 2 D), either in the absence or in the presence of GS, indicating that the presence of helical structure might induce some local rigidity. As it happens with the relaxation rates, the values of the spectral density functions at $\omega = 0$, ω_N and $0.87\ \omega_H$ for those polypeptides patches, did not show large values, except Arg8 which had some of the largest ones for $J(\omega_N)$ and $J(0.87\ \omega_H)$ (Fig. 7 B, C). Therefore, in isolated IF7, the regions with a certain helical propensity showed a high mobility, but Arg8 appeared to have a restricted mobility (a large $J(0)$ value, Fig 7 A (black dots)). It is interesting to note that mutational studies indicate that substitution of Arg8, Arg21, and Arg28 with glutamic residues abolishes the ability of the mutant protein to inactivate GS [10]. Furthermore, molecular modeling of the binding of GS to an N-terminal, 38-residue-long fragment of IF7, suggests that residues Arg8, Arg21, and Arg28 are involved in binding to GS [72]. Therefore, we could consider those residues as hot-spots of IF7, and then, from our mapping of the dynamics of the isolated protein, we can conclude that Arg8 did have a restricted motion (even though its

NOE value was not very large, Fig. 1 D). This restricted motion at Arg8 was kept (and it did not increase) when IF7 was bound to GS (Fig. 2 D, and Fig. 7 A, red dots). These findings do agree with those observed for other small IDPs [21], since also the hot-spots of those other proteins had a reduced mobility when the polypeptide chain is in isolation. It has been suggested that the difference in the mobility of the hot-spots in IDPs when compared to the other residues in their polypeptide chains could allow a "free-for-all" binding for the different partners [73]. In the case of IF7, its sole function so far described is the inhibition of GS, and then the mobility of the hot-spot region (and thus, the small NOE value) does not need to slow down to a larger extent, especially if the binding is directed by other regions of the protein (see below). This biological behavior among different IDPs could be also responsible for the behavior of the relaxation rates in the presence or in the absence of their partners: for instance, in an IDP which is involved in forming ternary complexes with different macromolecules, only the η_{xy} values change in the presence of the partners [21].

The dynamics of IF7 was mainly dominated by motions in the ps-ns timescale, with uncorrelated motions among all the residues, as suggested by the absence of linearity in the $J(0.87\omega_H)$ *versus* $J(0)$ plots (Fig. 8). Residues showing the largest values of $J(0)$ (with a reduced mobility) were those around Arg8 (which is involved in the binding of GS, see above) and at the C terminus of the protein (Fig. 4 C). We have previously hypothesized that the C-terminal region of IF7 (which is highly hydrophobic, as it contains the majority of Asp and Glu residues of the protein [4,43]) is key in achieving the inhibitory function of the protein (although it is not important in the encounter complex between the two macromolecules) [74]. In the presence of GS, residues at the C terminus had the largest variations in the $J(0)$, $J(\omega_N)$ and $J(0.87\omega_H)$ values (Fig. 7) when compared to those of the isolated IF7; moreover, the C-terminal region of the GS-bound protein showed the largest values of the relaxation rates (R_1, R_2 and η_{xy}, Fig. 2). Then, taken together, our NMR results in this work, together with previous results obtained with a short, N-terminal fragment of IF7 [74], indicate that not only Arg8, Arg21 and Arg28 (the two latter ones could not be observed in our experiments due to fast solvent-exchange) are key residues in the binding to GS, but also that residues at

the C terminus are important in such reaction. Recognition in IDPs occurs through conformational heterogeneity, with conformations of the IDP adopting different contact topologies and having different plasticity [75,76], which are modulated by residues outside the hot-spot regions: this regulation is based on the so-called non-native contacts of the IDP. These non-native contacts in IF7 would involve residues at its C-terminal region. The fact that different regions of IF7 are involved at different stages in the binding to GS (and therefore, different residues intervene at different stages in the binding between the two macromolecules) is also supported by previous titration experiments with GS [9]. However, it could be thought that the large $J(0)$ values around the C terminus were due to the high hydrophobicity of this region [54,56,77]. Currently, for IDPs or unfolded proteins there is not a clear relationship between hydrophobicity and variations in R_2 rates (or $J(0)$ values): for some unfolded proteins relaxation rates increase in hydrophobic regions [53,54,77], but for others, the two parameters are uncorrelated [56]. We think that the large values of $J(0)$ at the C terminus of IF7 correspond to a more compact conformation of the polypeptide chain (as suggested by our MD simulations, Fig. 5) to pave the way for the binding to GS.

Our SAXS results clearly show that IF7 is an IDP, which in solution adopted a disordered conformation when isolated. When bound to GS, IF7 did not adopt the same conformation as when it was isolated in solution, but it forms fuzzy complexes. There are several pieces of evidence confirming this result. First, SAXS data clearly show that IF7/GS structure was not a simple addition of the structures of the two isolated proteins (in fact, GS structure became more compact after IF7 binding). And second, we observed the absence of changes in the chemical shifts for the cross-peaks of the spectra of IF7, when GS was present (Fig. S1). It has been recently suggested that fuzziness in an IDP is due to the minimization of the unfavorable free energy of folding, while the protein establishes interactions with the partner involving backbone and side-chain atoms [78]. Since we are monitoring the backbone amide protons of IF7, we cannot speculate on the side-chain mobility of IF7 and their contribution to the mobility of the complex with GS.

The absence of linearity in the $J(0.87\omega_H)$ *versus* $J(0)$ plots indicates that the movements of all residues were uncorrelated, and then, in contrast with what is observed in other IDPs [21], we could not estimate local correlation times of IF7, which could have explained movements of segments of the polypeptide chain. In addition, SAXS data of isolated IF7 in solution were successfully fitted by a set of structures from MD, while a fitting of SAXS curve by a worm-like model, that was successfully used for other IDPs [21], did not provide satisfactory results. Then, it follows that IF7 could not be considered as a chain with a well-defined persistence length determining its stiffness, hence the arrangement and the movement of all residues is uncorrelated, as the dynamical results provided by RSDA indicate (Fig. 8). We hypothesize that this behavior was also related to the function of the protein: IF7 is a small protein when compared to GS (a dodecameric protein, formed by 450-residue-long monomers); then, the residues of IF7 do not need to restrict its conformational flexibility to a large extent (which could be provided by the presence of correlated motions), as IF7 does not have to bind other proteins of similar size.

AUTHOR CONTRIBUTIONS

JLN and MGO conducted the experiments. BR modelled the protein in simulation. FJF and MIMP provided materials. JLN, BR and MGO analyzed and interpreted the data. JLN, BR, FJF, MIMP and MGO wrote, edited and reviewed the manuscript.

The authors declare no competing financial interest.

ACKNOWLEDGEMENTS AND FUNDING

We than Paul Fitzpatrick for handling the manuscript. We thank two anonymous reviewers for helpful comments and discussion. MGO thanks Heinz Amenitsch for technical assistance in SAXS experiments. BR acknowledges the kind hospitality and use of computational resources in the European Magnetic Resonance Center (CERM), Sesto Fiorentino (Florence), Italy.

This work was supported by Spanish Ministry of Economy and Competitiveness [RTI 2018-097991-BI00 (to JLN)] with European ERDF funds (MCIU/AEI/FEDER, EU).

SUPPLEMENTARY MATERIAL

The Supplementary Material contains three tables (Tables ST1, ST2 and ST3) with the IF7 rates measured for the different conditions (ST1 and ST2), and data from the simulations (ST3); and five figures (NMR spectra (Fig. S1), SAXS (Fig. S2,S4 and S5) and MD (Fig. S3)).

5. REFERENCES

[1] J. A. Leigh, J. A. Dodsworth, Nitrogen regulation in bacteria and archaea. Annu. Rev. Microbiol. 61 (2007) 349-377.

[2] P. Bolay, M. I. Muro-Pastor, F. J. Florencio, S. Klähn, The distinctive regulation of cyanobacterial glutamine sythetase. Life 8 (2018) 52.

[3] M. García-Domínguez, J. C. Reyes, F. J. Florencio, Glutamine synthetase inactivation by protein-protein interaction. Proc. Natl. Acad. Sci. USA 96 (1999) 7161-7166.

[4] J. C. Reyes, F. J. Florencio, A novel mechanism of glutamine synthetase inactivation by ammonium in the cyanobacterium *Synechocystis* sp. PCC 6803: involvement of an inactivating protein. FEBS Lett. 367 (1995) 45-48.

[5] C. V. Galmozzi, L. Saelices, F. J. Florencio, M. I. Muro-Pastor, Posttranscriptional regulation of glutamine synthetase in the filamentous cyanobacterium *Anabaena* sp. PCC 7120: differential expression between vegetative cells and heterocysts, J. Bacteriol. 192 (2010) 4701-4711.

[6] M. I. Muro-Pastor, F. N. Barrera, J. C. Reyes, F. J. Florencio, J. L. Neira, The inactivating factor of glutamine synthetase IF7 is a natively unfolded protein, Protein Sci. 12 (2003) 1443-1454.

[7] P. E. Wright, H. J. Dyson, Intrinsically disordered proteins in cellular signalling and regulation, Nat. Rev. Mol. Cell Biol. 16 (2015) 18-29.

[8] P. Tompa, Intrinsically unstructured proteins. Trends Biochem. Sci. 27 (2002) 527-533.

[9] D. Pantoja-Uceda, J. L. Neira, L. Saelices, R. Robles-Rengel, F. J. Florencio, M. I. Muro-Pastor, J. Santoro, Dissecting the binding between glutamine synthetase and its two natively unfolded protein inhibitors, Biochemistry 55 (2017) 3370-3382.

[10] L. Saelices, C. V. Galmozzi, F. J. Florencio, M. I. Muro-Pastor, Mutational analysis of the inactivating factors IF7 and IF17 from *Synechocystis* sp. PCC 6803: critical role of arginine amino acid residues for glutamine synthetase inactivation. Mol. Microbiol. 82 (2011) 964-975.

[11] V. Receveur-Bréchot, J. M. Bourhis, V. N. Uversky, B. Canard, S. Longhi, Assessing protein disorder and induced folding, Proteins 62 (2006) 24-45.

[12] P. Bernadó, E. Mylonas, M. V. Petoukhov, M. Blackledge, D. I. Svergun, Structural characterization of flexible protein using small-angle X-ray scattering, J. Am. Chem. Soc. 129 (2007) 5656-5664.

[13] N. Rezaei-Ghaleh, M. Blackledge, M. Zweckstetter, Intrinsically disordered proteins: from sequence and conformational properties toward drug discovery, ChemBiochem 13 (2012) 930-950.

[14] T. N. Cordeiro, F. Herranz-Trillo, A. Urbanek, A. Estana, J. Cortes, N. Sibile, P. Bernardo, Small-angle scattering studies of intrinsically disordered proteins and their complexes. Curr. Opin. Struct Biol. 42 (2017) 15-23.

[15] M. R. Jensen, P.R. L. Markwick, S. Meier, C. Griesinger, M. Zweckstetter, S. Grzesiek, P. Bernadó, M. Blackledge, Quantitative determination of the conformational properties of partially folded and intrinsically disordered proteins using NMR dipolar couplings, Structure 17 (2009) 1169-1185.

[16] H. D. T. Mertens, D. I. Svergun, Combining NMR and small angle X-ray scattering for the study of biomolecular structure and dynamics, Arch. Biochem. Biophys. 628 (2017) 33-41.

[17] N. Salvi, A. Abyzov, M. Blackledge, Multi-time-scale dynamics in intrinsically disordered proteins from NMR relaxation and molecular simulation, J. Phys. Chem. 7 (2016) 2483-2489.

[18] A. Abyzov, N. Salvi, R. Schneider, D. Maurin, R.W. H. Ruigrok, M. Ringkjøbing Jensen, M. Blackledge, Identification of dynamic modes in an intrinsically disordered protein using temperature-dependent NMR relaxation, J. Am. Chem. Soc. 138 (2016) 6240-6251.

[19] R. Schneider, D. Maurin, G. Communie, J. Kragelj, D.M. Hansen, R.W.H. Ruigrok, M. Ringkjøbing Jensen, M. Blackledge, Visualizing the molecular recognition of an intrinsically disordered protein using multinuclear relaxation dispersion NMR, J. Am. Chem. Soc. 137 (2015) 1220-1229.

[20] M. L. Gill, R. A. Byrd, A. G. III Palmer, Dynamics of GCN4 facilitates DNA interaction: a model-free analysis of an intrinsically disordered region, Phys. Chem. Chem. Phys. 18 (2016), 5839-5849.

[21] J. L. Neira, M. Palomino-Schätzlein, C. Ricci, M. G. Ortore, B. Rizzuti, J. L. Iovanna, Dynamics of the intrinsically disordered protein NUPR1 in isolation and in its fuzzy complexes with DNA and prothymosin α. Biochim Biophys. Acta Proteins Proteom. 1867 (2019), 140252, doi: 10.1016/j.bbapap.2019.07.005.

[22] R. Schneider, M. Blackledge, M. Ringkjøbing Jensen, Elucidating binding mechanisms and dynamics of intrinsically disordered protein complexes using NMR spectroscopy, Curr. Opin. Struct. Biol. 54 (2018) 10-18.

[23] S. C. Gill, P. H. von Hippel, Calculation of protein extinction coefficients from amino acid sequence data. Anal. Biochem. 182 (1989), 319-326.

[24] J. Cavanagh, W. J. Fairborther, A.G. III Palmer and J. Skelton, Protein NMR spectroscopy, Academic Press, NY, 1996.

[25] L. Saelices, C.V. Galmozzi, F.J. Florencio, M.I. Muro-Pastor, J.L. Neira, The inactivating factor of glutamine synthetase IF17 is an intrinsically disordered protein, which folds upon binding to its target, Biochemistry 50 (2011) 9767-9778.

[26] N. A. Farrow, R. Muhandiram, A. U. Singer, S. M. Pascal, C.M., Kay, G. Gish, S. E. Shoelson, T. Pawson, T., J. D. Forman-Kay, L.E. Kay, Backbone dynamics of a free and phosphopeptide-complexed src homology 2 domain studied by ^{15}N NMR relaxation, Biochemistry 33 (1994) 5984-6003.

[27] N. Tjandra, A. Szabo, A. Bax, Protein backbone dynamics and ^{15}N chemical shift anisotropy from quantitative measurements of relaxation interference effects, J. Am. Chem. Soc. 118 (1996) 6986-6991.

[28] N. Tjandra, P. Wingfield, S. Stahl, A. Bax, Anisotropic rotational diffusion of perdeuterated HIV protease from ^{15}N NMR relaxation measurements at two magnetic fields, J. Biomol. NMR 8 (1996) 273-284.

[29] N. Rezaei-Ghaleh, K. Giller, S. Becker, M. Zweckstetter, Effect of zinc binding on β-amyloid structure and dynamics: implications for Aβ aggregation, Biophys. J. 101 (2011) 1202-1211.

[30] P. Kaderávek, V. Zapletal, Z. Rabatinová, L. Krasny, V. Sklenár, L. Zídek, Spectral density mapping protocols for analysis of molecular motions in disordered proteins, J. Biomol. NMR 58 (2014) 193-207.

[31] A. Abragam, Principles of nuclear magnetism, Clarendon Press, Oxford, UK, 1961.

[32] J. W. Peng, G. Wagner, Investigation of protein motions via relaxation measurements, Methods Enzymol. 239 (1994) 563-595.

[33] K. T. Dayie, G. Wagner, J. F. Lefèvre, Theory and practice of nuclear spin relaxation in proteins. Annu. Rev. Phys. Chem. 47 (1996) 243-282.

[34] N.A. Farrow, O. Zhang, A. Szabo, D.A. Torchia, L. E. Kay, Spectral density function mapping using ^{15}N relaxation data exclusively, J. Biomol. NMR 6 (1995) 153-162.

[35] Y. Hiyama, C. Niu, J. V. Silverton, A. Bavaso, D. A. Torchia, Determination of the ^{15}N chemical shift tensor via ^{15}N-^{2}H dipolar coupling in Boc-glycylglycyl [^{15}N] glycine benzyl ester, J. Am. Chem. Soc. 110 (1988) 2378-2383.

[36] I. G. Hughes, T. P. A Hase, Measurements and their uncertainties: a practical guide to modern error analysis, Oxford University Press, Oxford, UK, 2010.

[37] H. Amenitsch, M. Rappolt, M. Kriechbaum, H. Mio, P. Laggner, S. Bernstorff, First performance assessment of the small-angle X-ray scattering beamline at ELETTRA, J. Synchrotron Radiation 5 (1998) 506-508.

[38] O. Glatter, O. Kratky, Small angle X-ray scattering. Academic Press, London/New York, 1982.

[39] P. Debye, Molecular-weight determination by light scattering. J. Phys. Chem. 51 (1947) 18-32.

[40] F. Spinozzi, C. Ferrero, M.G. Ortore, A. De Maria Antolinos, P. Mariani, GENFIT: software for the analysis of small-angle X-ray and neutron scattering data of macromolecules in solution, J. Appl. Crystallogr. 47 (2014) 1132-1139.

[41] M.G. Ortore, F. Spinozzi, P. Mariani, A. Paciaroni, L.R.S. Barbosa, H. Amenitsch, M. Steinhart, J. Ollivier, D. Russo. Combining structure and dynamics: non-denaturing high-pressure effect on lysozyme in solution J. Roy. Soc. Interface 6 (2009) S619-S634.

[42] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J.L. Klepeis, R.O. Dror, D.E. Shaw, Improved side-chain torsion potentials for the Amber ff99SB protein force field, Proteins 78 (2010) 1950-1958.

[43] C. Cozza, J. L. Neira, F. J. Florencio, M. I. Muro-Pastor, B. Rizzuti, Intrinsically disordered inhibitor of glutamine synthetase is a functional protein with random-coil-like pK_a values, Protein Science 26 (2017) 1105-1115.

[44] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926-935.

[45] S. Piana, A.G. Donchev, P. Robustelli, D.E. Shaw, Water dispersion interactions strongly influence simulated structural properties of disordered protein states, J. Phys. Chem. B 119 (2015) 5113-5123.

[46] S. Izadi, R. Anandakrishnan, A.V. Onufriev, Building water models: a different approach, J. Phys. Chem. Lett. 5 (2014) 3863-3871.

[47] D. Pantoja-Uceda, J.L. Neira, L.M. Contreras, C.A. Manton, D.R. Welch, B. Rizzuti, The isolated C-terminal nuclear localization sequence of the breast cancer metastasis suppressor 1 is disordered, Arch. Biochem. Biophys. 664 (2019) 95-101.

[48] P. Santofimia-Castaño, B. Rizzuti, O. Abián, A. Velázquez-Campoy, J.L. Iovanna, J.L. Neira, Amphipathic helical peptides hamper protein-protein interactions of the intrinsically disordered chromatin nuclear protein 1 (NUPR1), BBA-Gen. Subjects 1862 (2018) 1283-1295.

[49] M.J. Abraham, T. Murtola, R. Schulz, S. Pall, J.C. Smith, B. Hess, E. Lindahl, GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1-2 (2015) 19–25.

[50] S. Evoli, R. Guzzi, B. Rizzuti, Molecular simulations of beta-lactoglobulin complexed with fatty acids reveal the structural basis of ligand affinity to internal and possible external binding sites, Proteins 82 (2014) 2609-2619.

[51] J.L. Neira, B. Rizzuti, J.L. Iovanna, Determinants of the pK_a values of ionizable residues in an intrinsically disordered protein, Arch. Biochem. Biophys. 598 (2016) 18-27.

[52] L. E. Kay, D. A. Torchia, A. Bax, Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease, Biochemistry 28 (1989) 8972-8979.

[53] M.-K. Cho, H.Y. Kim, P. Bernadó, C.O. Fernández, M. Blackledge, M. Zweckstetter, Amino acid bulkiness defines the local conformations and dynamics of natively unfolded α-synuclein and Tau, J. Am. Chem. Soc. 129 (2007) 3032-3033.

[54] J. Wirmer, W. Peti, H. Schwalbe, Motional properties of unfolded ubiquitin: a model for a random coil protein, J. Biomol. NMR 35 (2006) 175-186.

[55] N. Salvi, A. Abyzov, M. Blackledge, Atomic resolution conformational dynamics of intrinsically disordered proteins from NMR spin relaxation, Prog. Nucl. Magn. Reson. Spectrosc. 102-103 (2017) 43-60.

[56] J. Wirmer, C. Schlörb, J. Klein-Seetharaman, R. Hirano, T. Ueda, T. Imoto, H. Schwalbe, Modulation of compactness and long-range interactions of unfolded lysozyme by single-point mutations, Angew. Chem. Int. Ed. Engl. 43 (2004) 5780-5785.

[57] T. Mittag, S. Orlicky, W.-Y. Choy, X. Tangm, H. Ling, F. Sicheri, L. E. Kay, M. Tyers, J. Forman-Kay, Dynamics equilibrium engagement of a polyvalent ligand with a single-site receptor, Proc. Natl. Acad. Sci. USA 105 (2008) 17772-17777.

[58] J. Henriques, M. Skepö, Molecular dynamics simulations of intrinsically disordered proteins: on the accuracy of the TIP4P-D water model and the representativeness of protein disorder models, J. Chem. Theory Comput. 12 (2016) 3407-3415.

[59] P.S. Shabane, S. Izadi, A.V. Onufriev, General purpose water model can improve atomistic simulations of intrinsically disordered proteins, J. Chem. Theory Comput. 15 (2019) 2620-2634.

[60] M. Shatsky, R. Nussinov, H.J. Wolfson, A method for simultaneous alignment of multiple protein structures. Proteins 56 (2004) 143-156.

[61] J. Henriques, M. Skepö, Molecular dynamics simulations of intrinsically disordered proteins: on the accuracy of the TIP4P-D water model and the representativeness of protein disorder models, J. Chem. Theory Comput. 12 (2016) 3407-3415.

[62] D. Eisenberg, H.S. Gill, G.M.U. Pfluegl, S.H. Rotstein, Structure-function relationships of glutamine synthetases, BBA-Protein Struct. M. 1477 (2000) 122-145.

[63] M. Cottevieille, E. Larquet, S. Jonic, M.V. Petoukhov, G. Caprini, S. Paravisi, D.I. Svergun, M.A. Vanoni, N. Boisset, The subnanometer resolution structure of the glutamate synthase 1.2-MDa hexamer by cryoelectron microscopy and its oligomerization behavior in solution. J. Biol. Chem. 283 (2008) 8237-8249.

[64] P. D. Vise, B. Baral, A. J. Latos, G. W. Daughdrill, NMR chemical shift and relaxation measurements provide evidence for the coupled folding and binding of the p53 transactivation domain, Nucl. Acids Res. 33 (2005) 2061-2077.

[65] A. Atkinson, B. Kiefer, The role of protein motions in molecular recognition: insights from heteronuclear NMR relaxation measurements, Prog. Nucl. Magn. Reson. Spec. 44 (2004) 141-187.

[66] M. Andrec, G. Montelione, R. M. Levy, Lipari-Szabo mapping: a graphical approach to Lipari-Szabo analysis of NMR relaxation data using reduced spectral density mapping,. J. Biomol. NMR 18 (2000) 83-100.

[67] S. N. Khan, C. Charlier, R. Augustyniak, N. Salvi, V. Déjean, G. Bodenhausen, O. Lequin, P. Pelupessy, F. Ferrage, Distribution of pico- and nanosecond motions in disordered proteins from nuclear spin relaxation, Biophys. J. 109 (2015) 988-999.

[68] K. Melková, V. Zapletal, S. Jansen, E. Nomilner, M. Zachrdla, J. Hritz, J, Nováček, M. Zweckstetter, M. Ringkjøbing Jensen, M. Blackledge, L. Zídek, Functionally specific binding regions of microtubule-associated protein 2c exhibit distinct conformations and dynamics, J. Biol. Chem. 293 (2018) 13297-13309.

[69] N. Salvi, A. Abyzov, M. Blackledge, Analytical description of NMR relaxation highlights correlated dynamics in intrinsically disordered proteins, Angew Chem. Int. Ed. Engl. 56 (2017) 14020-14024.

[70] F. Ochsebein, R. Guerois, J.-M. Neumann, A. Sanson, E. Guittet, C. van Heijenoort. ^{15}N NMR relaxation as a probe for helical intrinsic propensity: the case of the unfolded D2 domain of annexin I, J. Biomol. NMR 19 (2001) 3-18.

[71] A. Borgia, M. B. Borgia, K. Bugge, V. M. Kissling, P. O. Heidarsson, C. B. Fernandes, A. Sottini, A. Soranno, K. J. Buholzer, D. Nettels, B. B. Kregalund, B.B. Best, B. Schuler, Extreme disorder in an ultrahigh-affinity protein complex,. Nature 555 (2018) 61-66.

[72] L. Saelices, R. Robles-Rengel, F. J. Florencio, M. I. Muro-Pastor, A core of three amino acids at the carboxyl-terminal region of glutamine synthetase defines its regulation in cyanobacteria Mol. Microbiol. 96 (2015) 483-496.

[73] M. Fuxreiter, Fold or not to fold upon binding-does it really matter?, Curr. Opin. Struct Biol. 54 (2019) 19-25.

[74] J. L. Neira, F. J. Florencio, M. I. Muro-Pastor, The isolated, twenty-three-residue-long, N-terminal region of the glutamine synthetase inactivating factor binds to its target. Biophys Chem. 228 (2017) 1-9.

[75] A. Toto, C. Camilloni, R. Giri, M. Brunori, M. Vendruscolo, S. Gianni, Molecular recognition by templated folding of an intrinsically disordered protein, Sci Rep. 6 (2016) 21994.

[76] M. Schwarten, Z. Solyom, S. Feuerstein, A. Aladag, S. Hoffmann, D. Willbold, B. Brutscher, Interaction of non structural protein 5A of the hepatitis C virus with Src homology 3 domains using non canonical binding sites, Biochemistry 52 (2013) 6160-6168.

[77] J. Klien-Seetheraman, M. Oikawa, S. B. Grimshaw, J. Wirmer, E. Duchardt, T. Ueda, T. Imoto, L. J. Smith, C. M. Dobson, H. Schwalbe, Long range interactions within a non-native protein, Science 295 (2002) 1719-1722.

[78] S. Hadzi, A. Mernik, C. Podlipnik, R. Loris, J. Lah, The thermodynamic basis of the fuzzy interaction of an intrinsically disordered protein, Angew. Chem. Int. Ed. 56 (2017) 14494-14497.

FIGURE LEGENDS

FIGURE 1: **Experimental relaxation data for isolated IF7 at 298 K**. (A) R_1; (B) R_2; (C) η_{xy}; and (D) heteronuclear NOE. Fitting errors to the exponential function for R_1 and R_2, and for the hyperbolic tangent function (for η_{xy}) are indicated. Errors in heteronuclear NOEs were determined from the measurements of the intensity of regions of two repeated spectra.

FIGURE 2: **Experimental relaxation data for IF7/GS complex at 298 K**. (A) R_1; (B) R_2; (C) η_{xy} and (D) heteronuclear NOE. Fitting errors to the exponential function for R_1 and R_2, and for the hyperbolic tangent function (for η_{xy}) are indicated. Errors in heteronuclear NOEs were determined from the measurements of the intensity of regions of two repeated spectra.

FIGURE 3: **Kratky plot of isolated IF7 at 298 K**. Kratky plot of SAXS experimental data (only half of the points was plotted for clarity), together with the theoretical fitting line obtained by a Debye approach corresponding to a disordered chain (continuous line, Eq. (7)).

FIGURE 4: **SAXS data relative to isolated IF7 at 298 K with fitting curves arising from protein models obtained in MD simulations**. The continuous lines correspond to theoretical fits from a set of structures obtained with the water model TIP4P-D (cyan line) and OPC (yellow line).

FIGURE 5**: Structures of IF7 obtained in MD simulations.** (A) The best three structures obtained in MD runs with the water model TIP4P-D (shades of blue) and OPC (shades of red); structures with the highest relative weight in the fit of the SAXS data are represented as having the most intense colour. The N-terminal region is on top and the C-terminal at the bottom. MultiProt [60] was used for molecule superposition. For a separated and more detailed view of the different structures, see Fig. S3

(Supplementary Material). (B) Histogram of the distribution of protein R_g in simulation with the water model TIP4P-D.

FIGURE 6: **Comparison between SAXS curves for GS alone and in complex with IF7**. From bottom to top: experimental SAXS points corresponding to isolated GS in solution and relative theoretical fitting (black continuous line) obtained from the crystallographic structure (3NG0 in the PDB); SAXS experimental data corresponding to GS in equimolar ratio with IF7 in solution (cyan line); SAXS theoretical curve obtained from the weighted sum of SAXS curves of IF7 and GS, weighted according to their respective concentration (equimolar) in the sample, assuming no complex formation (yellow line). Values on the y-axis refer to the experimental points and their theoretical fitting, whereas the other curves are up-shifted for clarity.

FIGURE 7: **RSDA results for IF7 in the absence and in the presence of GS**. (A) $J(0)$; (B) $J(\omega_N)$; (C) $J(0.87\ \omega_H)$. Errors in the value of each function were obtained from propagation errors of Equations (2-5).

FIGURE 8: **The plots of $J(0.87\ \omega_H)$ *versus* $J(0)$**. (A) isolated IF7; and (B) complex with GS (red). The $J(0)$ values were obtained from R_2^0 values. Errors in the values of each spectral density function were obtained from propagation errors of Equations (2-5).

Fig. 1(Neira et al.)

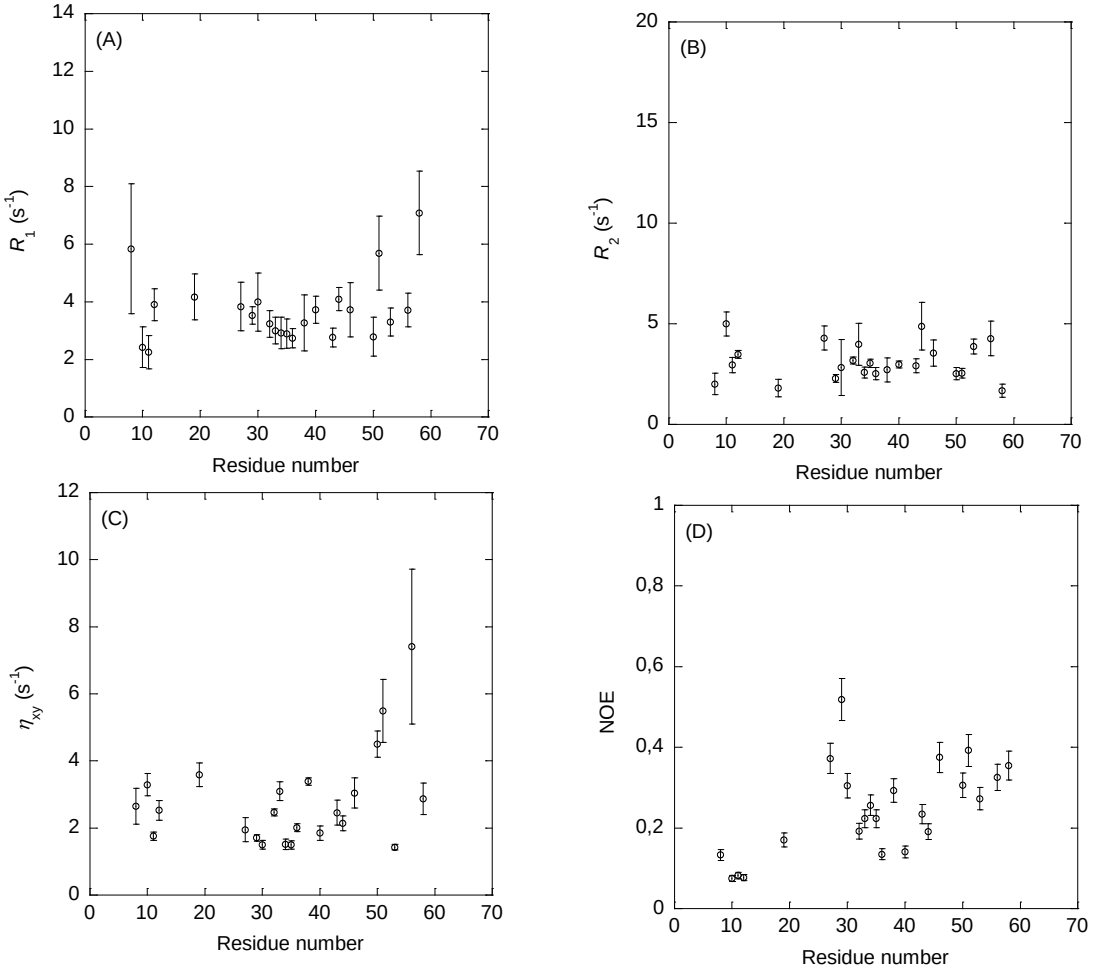

Fig. 2 (Neira et al.)

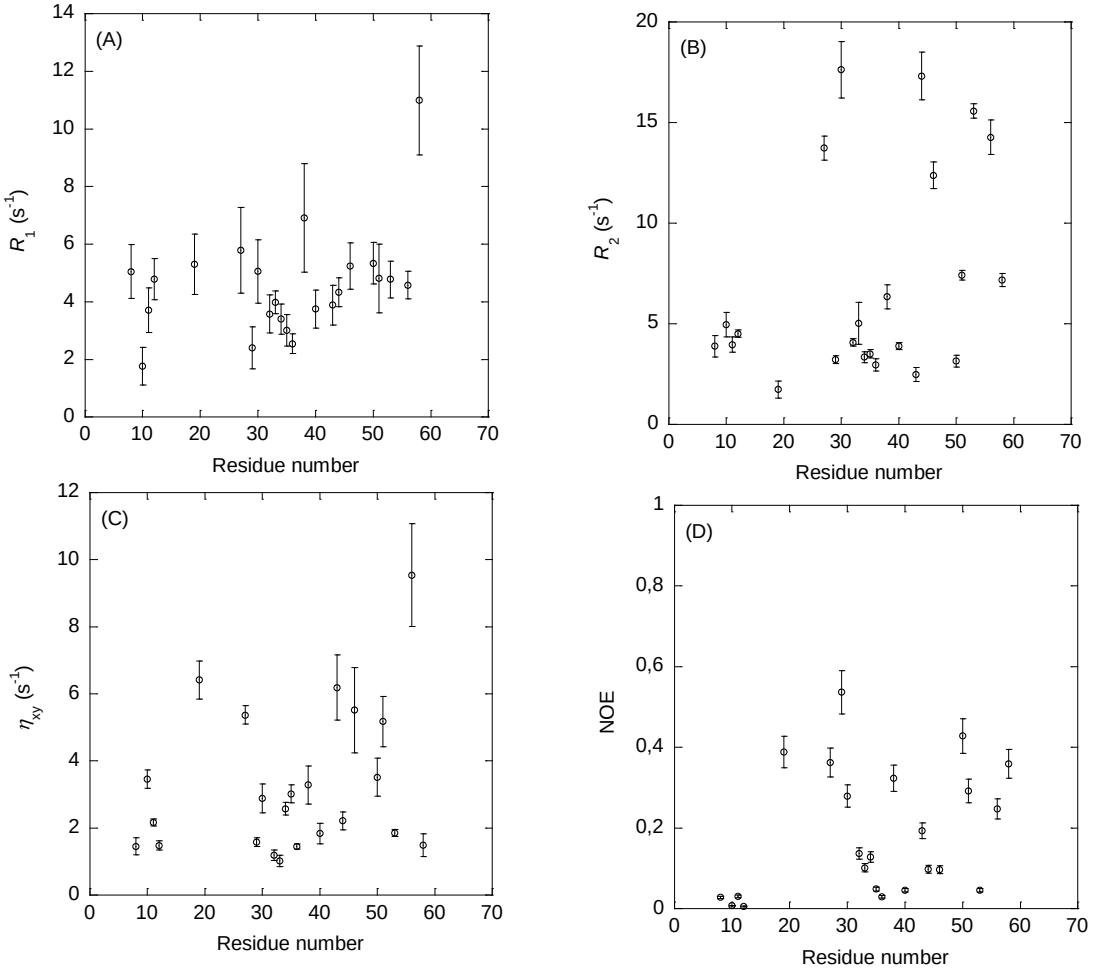

Fig. 3 (Neira et al.)

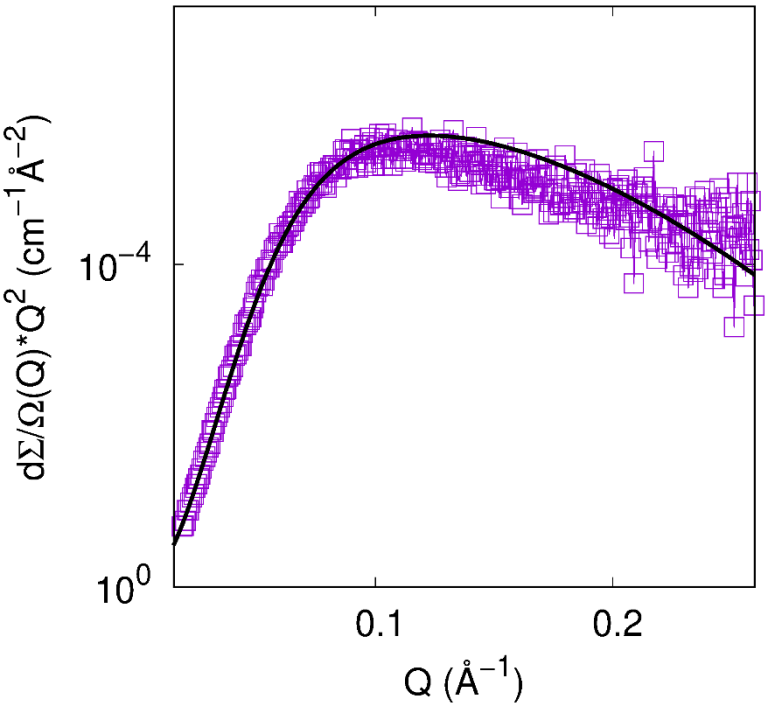

Fig. 4 (Neira et al.)

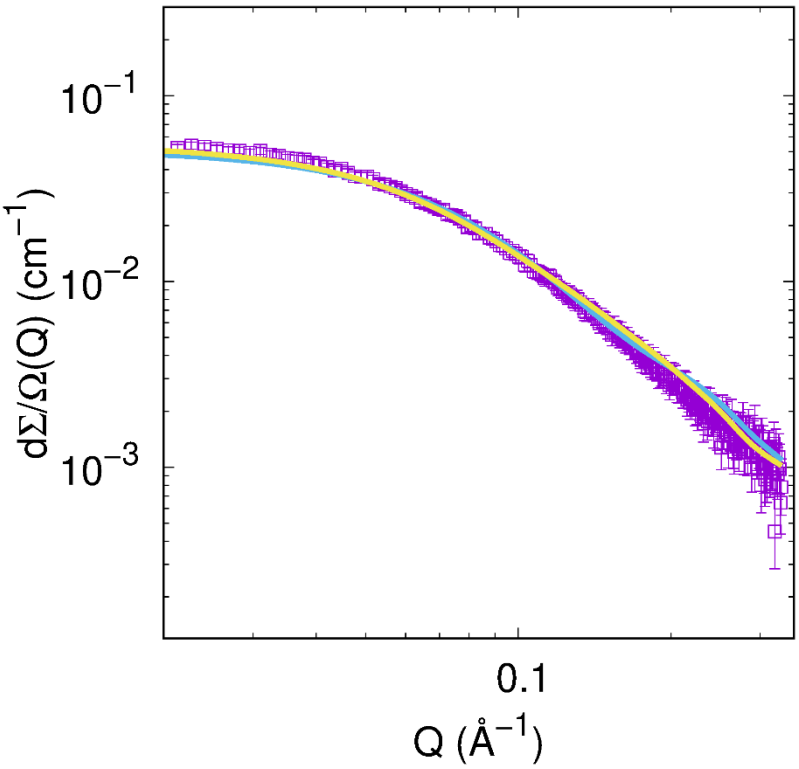

Fig. 5 (Neira et al.)

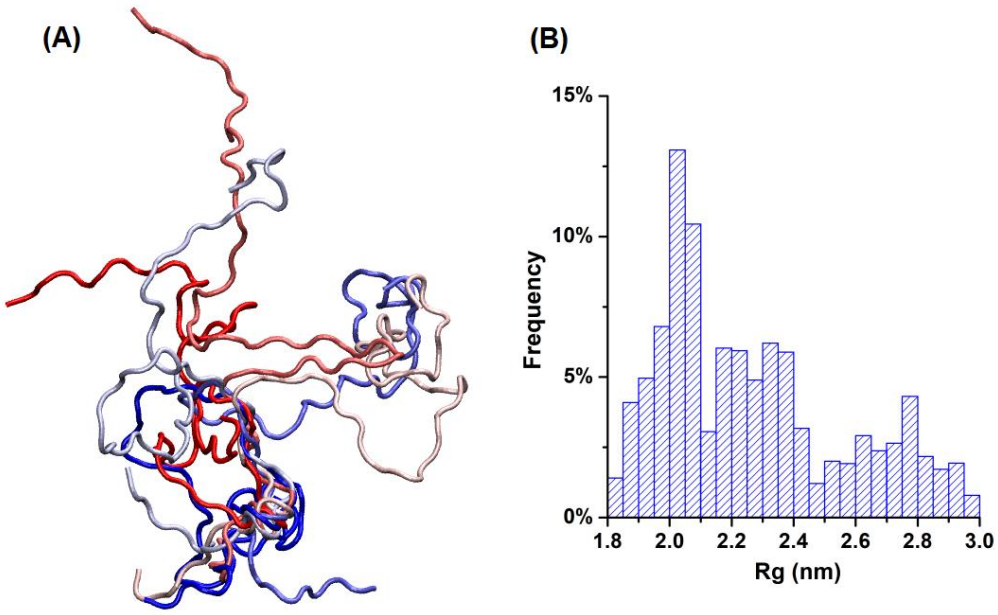

Fig. 6 (Neira et al.)

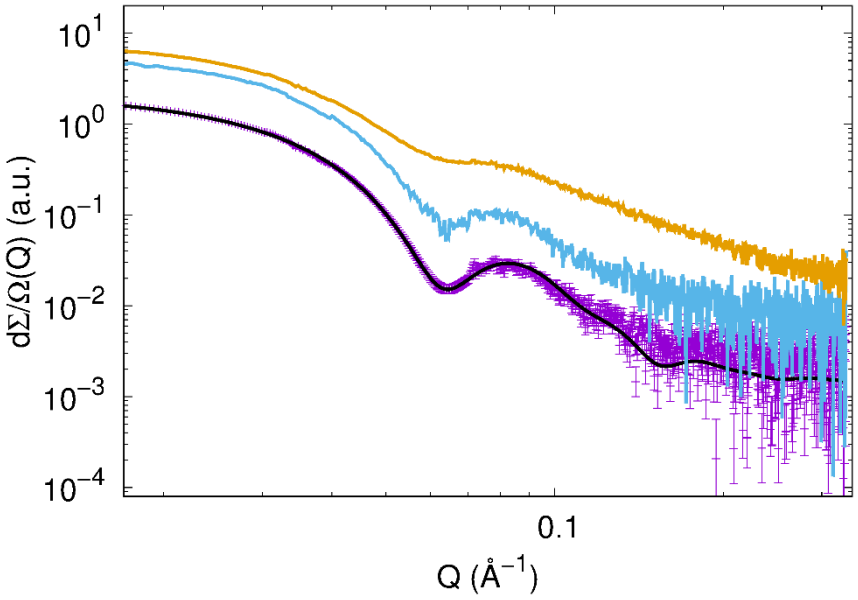

Fig. 7 (Neira et al.)

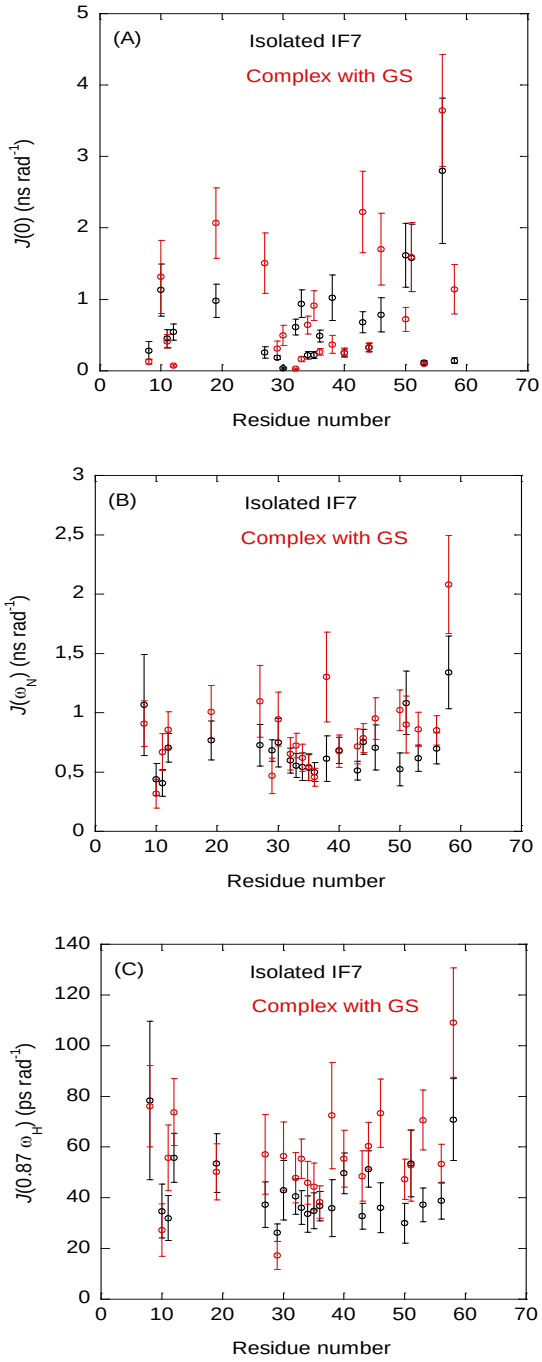

Fig. 8 (Neira et al.)