



HAL
open science

eIF4G1 N-terminal intrinsically disordered domain is a multi-docking station for RNA, Pab1, Pub1 and self-assembly

Belén Chaves-Arquero, Santiago Martínez-Lumbreras, Nathalie Sibille, Sergio Camero, Pau Bernadó, Ma Ángeles Jiménez, Silvia Zorrilla, José Manuel Pérez-Cañadillas

► To cite this version:

Belén Chaves-Arquero, Santiago Martínez-Lumbreras, Nathalie Sibille, Sergio Camero, Pau Bernadó, et al.. eIF4G1 N-terminal intrinsically disordered domain is a multi-docking station for RNA, Pab1, Pub1 and self-assembly. *Frontiers in Molecular Biosciences*, 2022, <10.3389/fmolb.2022.986121>. <hal-03763836>

HAL Id: hal-03763836

<https://hal.science/hal-03763836v1>

Submitted on 29 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

eIF4G1 N-terminal intrinsically disordered domain is a multi-docking station for RNA, Pab1, Pub1 and self-assembly

Belén Chaves-Arquero¹, Santiago Martínez-Lumbreras², Nathalie Sibille³, Pau Bernadó³, M. Angeles Jimenez⁴, Jose Manuel Pérez-Cañadillas^{4*}

¹Margarita Salas Center for Biological Research, Spanish National Research Council (CSIC), Spain,

²Institute of Structural Biology, Helmholtz Center München, Helmholtz Association of German Research Centres (HZ), Germany, ³INSERM U1054 Centre de Biochimie Structurale de Montpellier, France,

⁴Department of Biological Physical Chemistry, Institute of Physical Chemistry Rocasolano, Spanish National Research Council (CSIC), Spain

Submitted to Journal:
Frontiers in Molecular Biosciences

Specialty Section:
Protein Folding, Misfolding and Degradation

Article type:
Original Research Article

Manuscript ID:
986121

Received on:
04 Jul 2022

Revised on:
22 Aug 2022

Journal website link:
www.frontiersin.org

Conflict of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest

Author contribution statement

BC-A, SM-L and JMP-C made all of the clones, mutants and recombinant proteins. BC-A, SM-L and JMP-C obtained and analyzed NMR data with the help of MJ. Residual Dipolar Couplings were measured by BC-A and NS, and analyzed with the help of PB and NS. SAXS experiments were performed by BC-A and PB and analyzed by SM-L and JMP-C with the advice of PB. Paramagnetic Relaxation Experiments were performed and analyzed by BC-A and JMP-C. Protein structure calculations were done by JMP-C. The project was conceived by JMP-C who wrote the paper with the assistance of BC-A and with contributions other authors.

Keywords

Intrinsically disordered proteins (IDP), eIF4G1, Pab1, Pub1, NMR, Structural ensemble, Paramagnetic relaxation enhancement (PRE), SAXS (small-angle X-ray scattering), biomolecular condensation, Stress granule (SG), RNA binding protein

Abstract

Word count: 153

Yeast eIF4G1 interacts with RNA binding proteins (RBPs) like Pab1 and Pub1 affecting its function in translation initiation and stress granules formation. We present an NMR and SAXS study of the N-terminal intrinsically disordered region of eIF4G1 (residues 1-249) and its interactions with Pub1, Pab1 and RNA. The conformational ensemble of eIF4G11-249 shows an α -helix within the BOX3 conserved element and a dynamic network of fuzzy π - π and π -cation interactions involving arginine and aromatic residues. The Pab1 RRM2 domain interacts with eIF4G1 BOX3, the canonical interaction site, but also with BOX2, a conserved element of unknown function to date. The RNA1 region interacts with RNA through a new RNA interaction motif and with the Pub1 RRM3 domain. This later also interacts with eIF4G1 BOX1 modulating its intrinsic self-assembly properties. The description of the biomolecular interactions involving eIF4G1 to the residue detail increase our knowledge about biological processes involving this key translation initiation factor.

Contribution to the field

In this work we investigated the eIF4G1 N-terminal IDR structure, and revealed key intramolecular and intermolecular eIF4G1 contacts. We combined experimental NMR and SAXS data (in collaboration with Dr. Bernadó's group) and introduced π - π and π -cation interactions as ambiguous restraints to model transient contacts of this structure. To our knowledge, these interactions have never been explicitly used to obtain all-atoms IDP ensembles. Then we map the interactions of RNA, Pub1 and Pab1, that are also principal components of SG cores, over the sequence of eIF4G1 IDP using NMR. We identified new Pab1/Pub1 binding sites in eIF4G1 and solve an NMR structure of an eIF4G1-Pub1 chimera that reveals the molecular basis of recognition. The concerted binding of Pub1 and Pab1 induce conformational rearrangements in eIF4G1 IDP that enhance intrinsic self-association properties. By investigating the interplay between different SG proteins, we believe our work represents a step-forward to disentangle the complexity of the biomolecular condensation.

Funding statement

NMR experiments were performed in the "Manuel Rico" NMR laboratory (LMR) (CSIC). The synchrotron SAXS data was collected at beamline P12 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany). The study was supported by funds from the Spanish MINECO (CTQ2018-84371) to MJ and JMC-P, and Autonomous Community of Madrid (B2017/BMD-3770) to JMP-C. BC-A and SM-L were funded by predoctoral grants from Spanish MINECO (BES-2015-073383) and Autonomous Community of Madrid (CPI/0265/2008) respectively. This work was supported by the Labex EpiGenMed, an « Investissements d'avenir » program (ANR-10-LABX-12-01) awarded to PB. The CBS is a member of France-BioImaging (FBI) and the French Infrastructure for Integrated Structural Biology (FRISBI), 2 national infrastructures supported by the French National Research Agency (ANR-10-INBS-04-01 and ANR-10-INBS-05, respectively).

Ethics statements

Studies involving animal subjects

Generated Statement: No animal studies are presented in this manuscript.

Studies involving human subjects

Generated Statement: No human studies are presented in this manuscript.

Inclusion of identifiable human data

Generated Statement: No potentially identifiable human images or data is presented in this study.

In review

Data availability statement

Generated Statement: The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below:
<http://www.wwpdb.org/>, 6Z29.

In review

eIF4G1 N-terminal intrinsically disordered domain is a multi-docking station for RNA, Pab1, Pub1 and self-assembly

Belén Chaves-Arquero^{1,3}, Santiago Martínez-Lumbreras^{1,4}, Nathalie Sibille², Sergio Camero¹, Pau Bernadó², M^a Ángeles Jiménez¹, Silvia Zorrilla⁵ and José Manuel Pérez-Cañadillas^{1,*}.

¹ Department of Biological Physical Chemistry, Institute of Physical-Chemistry "Rocasolano", CSIC, C/ Serrano 119, 28006 Madrid, Spain

² Centre de Biochimie Structurale (CBS), CNRS, INSERM, Univ. Montpellier, Montpellier, France

³ Present address: Department of Structural and Chemical Biology, Center for Biological Research, CIB, CSIC, Av. Ramiro de Maeztu 9, 28040 Madrid, Spain

⁴ Present address: Institute of Structural Biology, Helmholtz Zentrum München, Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany

⁵ Department of Cellular and Molecular Biology, Biological Research Center, CSIC, Madrid, Spain

*to whom correspondence should be addressed. Tel:+34 91 561 94 00; Fax: +34 91 564 24 31; Email: jmperez@iqfr.csic.es

In review

Abstract

Yeast eIF4G1 interacts with RNA binding proteins (RBPs) like Pab1 and Pub1 affecting its function in translation initiation and stress granules formation. We present an NMR and SAXS study of the N-terminal intrinsically disordered region of eIF4G1 (residues 1-249) and its interactions with Pub1, Pab1 and RNA. The conformational ensemble of eIF4G1₁₋₂₄₉ shows an α -helix within the BOX3 conserved element and a dynamic network of fuzzy π - π and π -cation interactions involving arginine and aromatic residues. The Pab1 RRM2 domain interacts with eIF4G1 BOX3, the canonical interaction site, but also with BOX2, a conserved element of unknown function to date. The RNA1 region interacts with RNA through a new RNA interaction motif and with the Pub1 RRM3 domain. This later also interacts with eIF4G1 BOX1 modulating its intrinsic self-assembly properties. The description of the biomolecular interactions involving eIF4G1 to the residue detail increases our knowledge about biological processes involving this key translation initiation factor.

Keywords: Intrinsically disordered proteins, posttranscriptional control, protein-protein interactions, pi-pi contacts, RRMs, Pab1, Pub1, eIF4G1,

Running title: EIF4G1 IDP structure and interactions

Introduction

The eukaryotic translation initiation factor eIF4G is a central player in the regulation of protein expression. First, it enhances translation initiation, the rate-limiting step, by mRNA 3'/5'-end circularization (Tarun et al., 1997; Wells et al., 1998; Preiss and Hentze, 1999; Sonenberg and Hinnebusch, 2009; Aitken and Lorsch, 2012). Structurally this is achieved by the formation of a “closed-loop” complex (CLC), in which eIF4G forms a stable eIF4F heterotrimer (eIF4G+eIF4E+eIF4A) that recognizes the 5'-cap (via eIF4E) of the mRNA and recruits the 3' poly(A) tail-associated Pab1. On the other hand, eIF4G is involved in the nucleation of stress granule (SG), membrane-less organelles that store components of the translation machineries in an arrested state in response to nutrient starvation, temperature, oxidative or chemical stresses (Kedersha et al., 1999; Hoyle et al., 2007; Buchan et al., 2008, 2011; Jain et al., 2016).

There are two eIF4G genes in *Saccharomyces cerevisiae* that have similar domain architecture (Goyer et al., 1993). EIF4G1 (also referred to as Tif4631) is the most abundant and contains up to 57.4% of its residues in predicted intrinsically disordered regions (IDRs) (MobiDB: <https://mobidb.bio.unipd.it/P39935>): in the N-terminus (residues 1-77;111-392;398-409), middle (residues 481-591) and C-terminus (residues 870-952). The remaining regions contains the domains for interaction with eIF4A (Schütz et al., 2008) and eIF4E (Hershey et al., 1999) that form the eIF4F heterotrimer. EIF4G1 has three RNA binding domains (RNA1, RNA2 and RNA3) within the IDRs (Berset et al., 2003; Park et al., 2011), and a conserved box (BOX3, in the N-terminus IDR) that interacts with Pab1 (Tarun and Sachs, 1996; Kessler and Sachs, 1998) to promote the assembly of the CLC and therefore translation initiation. Conversely, the IDRs also contain the binding sites of translational repressors such as Pub1 (Santiveri et al., 2011) (at the N-terminal IDR); Sbp1 (middle IDR), Scd6 (C-terminal IDR) and Npl3 (middle

and C-term IDRs) (Poornima et al., 2016), and Ded1 (C-terminal IDR) (Hilliker et al., 2011)

All of these eIF4G-interacting proteins are RNA binding proteins as well and contain a combination of folded domains and IDRs. Pab1 and Pub1, those interacting at the eIF4G1 N-terminal IDR, contain four and three RNA Recognition Motifs (RRM), arranged as bead-on-string, and various low complexity domains (LCD). These proteins are constituents of biological condensates and, in response to temperature increase or acidic pH, undergo *in vitro* LLPS in which both the RRM and the LCD participate (Lin et al., 2015; Riback et al., 2017; Kroschwald et al., 2018a). Pab1, Pub1 and eIF4G1 are the protein constituents of the EGP-bodies, the earliest type of SG described in yeast (Hoyle et al., 2007).

Knowing the structural details of the protein/RNA network involving eIF4G1, Pab1 and Pub1 will help to propose integrative models of translation regulation that include activation and repression pathways. However, because these interactions involve IDRs their study is difficult with X-ray crystallography or Cryo-EM. Therefore, we performed an NMR/SAXS structural study of the eIF4G1 N-terminal IDR and generate an “all-atoms” ensemble stabilized by cation- π and π - π transient contacts. Furthermore, we mapped the binding sites of Pab1, Pub1 and RNA, and show the simultaneous binding of the two RNA Binding Proteins (RBP), reconfigure the conformation of eIF4G1 IDR inducing self-assembly through its conserved element BOX1. The implications of these findings for translation regulation and biomolecular condensation are discussed.

RESULTS

The N-terminal eIF4G1 IDR contains residual structural features

The N-terminal IDR of *Saccharomyces cerevisiae* eIF4G1 contains three boxes of about 15-20 conserved residues each, and a conserved RNA binding region (RNA1), within the first 249 amino acids (Park et al., 2011) (Figure 1A). We studied by NMR the conformational properties of eIF4G1₁₋₂₄₉, a construct that is stable for days under different pH conditions (Supplementary Figure 1). NMR is a very powerful technique for the investigation of IDRs and their interactions at the residue level ((Kurzbach et al., 2015; Gibbs et al., 2017; Milles et al., 2018) and references therein).

The ¹H-¹⁵N HSQC NMR spectrum of eIF4G1₁₋₂₄₉ is characteristic of an intrinsically disordered protein (IDP) with low dispersion in the proton dimension and sharp signals (Figure 1B). However, interestingly, several glycines showed relatively broad peaks at $\delta_{\text{NH}} < 8.00$ ppms (e.g. G97, G77 and G83) that are not compatible with a fully disordered state. We identified several minor species (signals labelled in red) that correspond to *cis*Pro conformers, and two uncommon chemical isomerization forms at positions 41 and 76 that were assigned to isoaspartates (Supplementary Figure 2). These variants arise from deamidation of N41 and N76, that lie next to Gly residues in the protein sequence; Asn-Gly sequences are known to have the highest tendency to experience this non-enzymatic deamidation in model peptides (Robinson and Robinson, 2001). The level of deamidation is similar for the two positions (12-14%) and remained constant in different samples and over NMR experimental time, suggesting that these forms might have been generated *in vivo*.

Analysis of eIF4G1₁₋₂₄₉ secondary structure based on ¹³C chemical shifts, T₁/T₂ ¹⁵N relaxation times and residual dipolar couplings (RDCs) revealed the presence of an α -helix within BOX3 (Figure 2A). This finding was confirmed by characteristic sequential

amide-amide NOEs measured in a 3D ^1H - ^{15}N -HSQC-NOESY- ^1H - ^{15}N HSQC spectrum (Supplementary Figure 3). We determined the NMR structure of this α -helix using a BOX3 model peptide (eIF4G1₁₈₇₋₂₃₄) (Supplementary Figure 3). No further standard secondary structure elements were identified in eIF4G1₁₋₂₄₉. However, the broad glycine peaks seen in Figure 1B, suggested the presence of residual higher order structures in this construct. In support of such structures, random coil index” (RCI) values S2 predicted from the chemical shifts (Camilloni et al., 2012), and the lower T_1 , T_2 relaxation times for the conserved boxes suggested that these boxes might be involved in transient contacts that restrict mobility and/or induce chemical exchange processes resulting in short T_2 values (Supplementary Figure 4). The existence of long-range interactions was evidenced by paramagnetic relaxation enhancement (PRE) measurements. The nitroxyl spin-label derivatization of engineered cysteine mutants (eIF4G1₁₋₂₄₉ has no native Cys) indicates long-range PREs for S200C and Q109C mutants (Figure 2B). In a protein such as eIF4G1₁₋₂₄₉, the PREs (calculated as described in (Battiste and Wagner, 2000) and Supplementary Figure 5) are expected to occur within 25-30 Å of the spin label. Therefore, the PRE data suggested the presence of long-range contacts in eIF4G1₁₋₂₄₉ involving BOX1, RNA1 and BOX3.

To determine if these contacts are predominantly intra- or intermolecular, we placed the spin label in the non-isotope labeled Q109C mutant, added wild-type ^{15}N -labeled eIF4G1₁₋₂₄₉, and then measured PRE. The PRE fingerprint shows that eIF4G1₁₋₂₄₉ can self-interact through contacts involving BOX1 and RNA1, as these elements “sense” the presence of the spin label in *trans* (Figure 2B lowest graph). However, the magnitude of the effects of the spin label in *trans* is smaller than when it is in *cis* (Figure 2B middle graph), suggesting that there is a small population of the self-associated species. All of the Tyr residues of the construct are contained in these regions involved in eIF4G1₁₋₂₄₉

self-recognition (see Figure 1A). Five of these Tyr residues are included in BOX1 that is a predicted amyloid-like sequence (Supplementary Figure 6). These data suggest that eIF4G₁₋₂₄₉ self-recognition involves Tyr-Tyr interactions.

eIF4G1 IDR conformational ensemble

IDPs, are considered to consist of ensembles of co-existing conformers. To build realistic conformational ensembles of IDPs, it is necessary to identify the possible residual secondary structures and long-range interactions between different regions of the polypeptide (Schwalbe et al., 2014). These structural features are generally sparsely populated and transient in IDPs, which makes them difficult to detect and quantify experimentally thereby undermining the possibilities of calculating conformational ensembles. Nevertheless, several studies have proposed that interactions that favor aggregation, flexibility and/or long-range contacts are prevalent in IDPs (Brangwynne et al., 2015; Vernon et al., 2018; Gomes and Shorter, 2019). In particular on the key role of cation- π and π - π interactions in the “molecular grammar” of phase separation in prion-like IDPs (Wang et al., 2018). There are 11 Arg, 11 Tyr, 3 Phe and 1 Trp in eIF4G₁₋₂₄₉ that are suitable for these types of interactions, and they are mostly located in the conserved boxes (see Figure 1A). Therefore, we hypothesized that cation- π and π - π interactions (involving Arg, Tyr, Phe and Trp) might dominate the long-range contacts in eIF4G₁₋₂₄₉.

Different methods have been used to calculate IDP conformational ensembles using experimental data (NMR, SAXS and others) and/or computational approaches (Bernadó et al., 2005; Kragelj et al., 2015; Kurzbach et al., 2015; Das et al., 2018; Bhattacharya and Lin, 2019; Estaña et al., 2019). Here we used the algorithm in the program Cyana 3.0 (Güntert and Buchner, 2015) for fast generation of eIF4G₁₋₂₄₉ all-atoms structural models. This approach allows a straightforward implementation of: (1) NOE-derived

distance restraints and ^{13}C -derived ϕ/ψ dihedral restraints for the parts of the protein that are well-folded (i.e. BOX3) and (2) ambiguous restraints for the cation- π and π - π interactions (involving Arg and Tyr) that we propose as dominant transient residue-residue contacts. We refer to these latter types of restraints as “knowledge-based” and used cautions to avoid biases in their selection (see materials and methods for specific details). We calculated 80000 eIF4G1₁₋₂₄₉ structures and sorted them using our own greedy algorithm that optimized the fitting to experimental PRE data stepwise. As a seed for the ensemble the protocol chooses the structure with minimum PRE violations and then continues building up the ensemble stepwise using the same criteria (i.e. incorporating the structure that, together with the previously selected structures(s), minimizes violations). The target function reached a minimum value relatively quickly and increased slowly afterwards (Supplementary Figure 7). We arbitrarily selected a final 500-member ensemble to ensure sufficient structural variability while still maintaining good agreement between the back-calculated and the experimental PREs (red line in Figure 2B). Because the intermolecular PRE data showed self-association, we performed a similar protocol for analysis of eIF4G1₁₋₂₄₉ dimers, hypothesizing that Tyr-Tyr interactions are the driving force of dimerization. The SAXS curve evidenced the IDP nature of eIF4G1₁₋₂₄₉ (Figure 2C). To validate the eIF4G1₁₋₂₄₉ ensemble models, we applied the Ensemble Optimization Method (EOM) genetic algorithm (Bernadó et al., 2007; Tria et al., 2015) to model the SAXS curve. Pools for monomeric and dimeric conformations were used without restricting the relative percentages of each set. We performed 10 independent EOM calculations, and each of them resulted in excellent fittings of the experimental curve (Figure 2C). Importantly, similar calculations done with either eIF4G1₁₋₂₄₉ monomers or dimers alone resulted in worse fits. In the SAXS-selected eIF4G1₁₋₂₄₉ ensemble, the monomers dominate (88%). As expected, the back-calculated

PREs from the collection of EOM ensembles showed poorer agreement with the experimental data, but neatly reflected the overall trends regarding the long-range contacts in eIF4G1₁₋₂₄₉ (Supplementary Figure 7).

The eIF4G1₁₋₂₄₉ ensemble of conformers showed a flexible α -helix in BOX3 and, despite the presence of long-range contacts, no predominant tertiary fold was found. The average C α -C α distance maps revealed that local and long-range contacts were prevalent between BOX1 and RNA1-1/RNA1-2 boxes, and between these three elements and BOX3 (Figure 2D left). In contrast, there was a remarkable absence of long-range interactions involving BOX2. Dimerization contacts were dominated by BOX1 and to a lesser extent by the RNA1-1 box (Figure 2D right). The BOX3 region also showed a minimum in the intermolecular C α -C α distance maps, due to the coexistence of intramolecular Arg-Tyr and intermolecular Tyr-Tyr contacts (Figure 2D right). Indeed, these π - π and cation- π interactions tend to appear in networks, rather than in binary mode, probably favored by the planar nature of aromatic and guanidinium groups.

In summary, these data showed that eIF4G1₁₋₂₄₉ is predominantly disordered except for an α -helix in BOX3. Atomistic models were constructed with experimental and knowledge-based restraints and ensembles were built by restraining against experimental NMR and SAXS data. Their analysis showed an intrinsic tendency of eIF4G1₁₋₂₄₉ to dimerize (oligomerize), in which BOX1 plays the chief role.

eIF4G1 interacts with Pab1 and Pub1 through multiple binding sites

eIF4G1, Pub1 and Pab1 that are considered as key components of SG (Buchan et al., 2011; Jain et al., 2016) We studied the structural details of their interaction network. Previous to this work, we showed that Pub1 RRM3 interacts with eIF4G1 (Santiveri et al., 2011). Here we mapped this interaction on eIF4G1₁₋₂₄₉ by analysis of changes on the ¹H-¹⁵N HSQC spectra, which indicated three putative binding sites for RRM3 in eIF4G1₁₋

249 (orange bar chart in Figure 3A): two in RNA1 (RNA1-1 and RNA1-2) and one in BOX1. Strikingly, these sites have a small consensus sequence motif (YNNxxxY), only present in this region of the eIF4G1. We tested the ability of short peptides of eIF4G1 that corresponded to the conserved elements (Figure 3B) to bind to ¹⁵N-labelled Pub1 RRM3, by monitoring their effect on the Pub1 RRM3 ¹H-¹⁵N HSQC spectrum. Only BOX1 and RNA1-1 peptides caused changes in the Pub1 spectrum arising from direct contacts (Figure 3B). The Pub1 binding site in BOX1 overlaps with the amyloid-like sequence (Supplementary Figure 6). The absence of spectral changes upon RNA1-2 peptide titration (Figure 3B) suggested that the effects observed on the spectrum of the corresponding region of eIF4G1₁₋₂₄₉ in complex with Pub1 RRM3 (orange bar chart in Figure 3A) are probably due to conformational rearrangements.

The binding of short eIF4G1 peptides was too weak to obtain structural restraints (e.g. intermolecular NOEs) for calculation of the structure of the complex. To overcome this technical problem, we constructed recombinant chimeras of eIF4G1₃₅₋₄₉ and Pub1 RRM3 (Figure 4 and Supplementary Figure 8). The NMR spectrum of the eIF4G1 peptide fused to the C-terminus of Pub1 was similar to that of Pub1 RRM3 alone (Figure 4A right), whereas that of the N-terminally fused chimera differed significantly (Figure 4A, left), suggesting that the peptide can effectively fold-back into the binding site only in the latter case. Using this latter construct, we obtained enough experimental restraints to calculate the structure of the eIF4G1₃₅₋₄₉-Pub1 RRM3 chimera (PDB: 6Z29), which shed light on the key elements required for molecular recognition (Figure 4B and Supplementary Figure 8). The structure showed that eIF4G1 residues Y₄₁, N₄₂ and N₄₃ (part of the YNNxxxY conserved motif) interact with a shallow cleft in Pub1 RRM3, defined by the contact between helix α 1 and strand β 2. The eIF4G1 Y₄₁ is inserted into a small cavity and contacts I₃₅₈, I₃₆₇ and F₃₇₀ of Pub1, whereas the N₄₂ and N₄₃ of eIF4G1 are more

exposed but contact L₃₆₈ and F₃₆₆ of Pub1. The latter residue was previously shown to be important for eIF4G1-Pub1 interaction (Santiveri et al., 2011). The interaction surface was small, in agreement with a weak eIF4G1-Pub1 interaction.

We next studied the interaction between eIF4G1 and Pab1 using similar approaches. NMR titrations of unlabeled Pab1 RRM12 over ¹⁵N-eIF4G1₁₋₂₄₉ (blue bar chart in Figure 3A) caused similar pattern of perturbations and signal disappearance than Pub1 RRM3 but with additional changes in BOX2 (aa 135-160) and BOX3 (aa 200-234). Experiments with eIF4G1₁₋₈₂ shows little changes for Pab1 RRM12 titration in comparison with Pub1 RRM3 (Supplementary figure 9), suggesting that Pab1 RRM12 does not interact with RNA1 and changes could be due to conformational reorganization of eIF4G1₁₋₂₄₉. For the other elements, only BOX3 has been reported as a Pab1 binding site to date (Kessler and Sachs, 1998). As done for Pub1 RRM3, we studied the interactions of Pab1 RRM12 with eIF4G1 fragments to validate the putative binding sites. As expected, the BOX3 peptide interacted with Pab1 RRM12 causing significant perturbations in the helix1-helix2 interface of RRM2 (Figure 3C right), an equivalent region to that involved in human eIF4G-PABP1 recognition (Safaei et al., 2012). The BOX2 peptide also interacted with RRM2 through a similar interface; but probably weakly because it causes fewer changes than the BOX3 peptide. NMR data showed that the BOX1 peptide does not directly interact with Pab1 RRM12 and that the observed changes in that region of eIF4G1₁₋₂₄₉ (blue bar chart in Figure 3A) are due to reorganization of internal contacts. Interestingly Pub1 and Pab1 used different interfaces of the RRM to interact with eIF4G1 (Supplementary Figure 8C).

These results showed that, although Pub1 and Pab1 interact with eIF4G1₁₋₂₄₉ through multiple sites, most of these interactions are weak because they caused minor chemical shift changes in these RBPs; the exception is the Pab1-BOX3 interaction that showed

changes of a larger magnitude. The presence of two eIF4G1 binding sites for Pub1 and Pab1, combined with their self-association properties, might suggest a possible cooperative recognition mode of eIF4G1.

Pub1 and Pab1 induce BOX1-dependent aggregation

During the course of the study, we found that Pub1 RRM3 interacts differently with different eIF4G1 constructs. Titration with eIF4G1₁₋₈₂, eIF4G1₁₋₃₀₅, eIF4G1₁₋₃₄₈ and eIF4G1₁₋₄₀₂ constructs caused small changes in the Pub1 RRM3 NMR signals equivalent to those described for its interaction with eIF4G1₁₋₂₄₉ and BOX1 and RNA1-1 peptides (Figure 5A left). However, surprisingly, titration with eIF4G1₁₋₁₈₄ cause the disappearance of nearly all of the Pub1 RRM3 ¹H-¹⁵N HSQC crosspeaks (Figure 5A right). This result cannot be explained by changes in the chemical exchange kinetics because signals not affected by binding (i.e. $\Delta\omega=0$) will not experience line broadening. The massive line broadening is either compatible with Pub1 RRM3 being part of a high molecular weight structure (i.e. aggregates), or it is in equilibrium with these large particles in a way that properties associated to their slow tumbling (T_1 , T_2) are transferred to free Pub1 RRM3.

We further investigate the cause of Pub1 RRM3 behavior and discovered that it recover the weak-binding pattern with several eIF4G1₁₋₁₈₄ BOX1 mutants (Δ BOX1, W95A, Y98A/Y99A and Y105A/Y106A), whereas deletion of BOX2 (Δ BOX2) still caused the Pub1 aggregation-like pattern. Thus, BOX1 is the element causing the differential binding mode of Pub1 RRM3, but how. Our hypothesis is that Pub1 RRM3 binding causes conformational rearrangements on eIF4G1₁₋₁₈₄ exposing aggregation-prone BOX1 (supplementary figure 5A and B). Interestingly, eIF4G1₁₋₁₈₄ experiences a complex degradation (supplementary figure 1) that we characterized by NMR (supplementary Figure 10). New C-terminal peaks arise from internal breaks at specific points of the

polypeptide chain and the signals from the BOX1 region disappear (e.g. W95 sidechain and G97), a behavior compatible with aggregation. Macroscopically, eIF4G1₁₋₁₈₄ (and eIF4G1₁₋₁₈₄ ΔBOX1) samples age to hydrogels that bind Congo-red (supplementary figure 5C), a dye used to detect amyloids and protein aggregates (Yakupova et al., 2019), but eIF4G1₁₋₁₈₄ ΔBOX1 does not. As in the case of Pub1 RRM3, these observations suggests that the loss of BOX1 transient interactions to other parts of eIF4G1 (in this case by degradation) trigger its aggregation-prone properties.

The construct eIF4G1₁₋₂₄₉ is more stable (supplementary figure 1) and Pub1 RRM3 binds to it weakly (Figure 5A left). This suggests that Pub1 RRM3 is unable to disrupt the BOX3-BOX1 contacts present in longer eIF4G1 forms (eIF4G1₁₋₂₄₉, eIF4G1₁₋₃₀₅, eIF4G1₁₋₃₄₈ and eIF4G1₁₋₄₀₂). Consistent with this hypothesis, titration of Pab1 RRM12 into the eIF4G1₁₋₂₄₉/Pub1 RRM3 complex retrieve the Pub1 RRM3 broad spectrum, suggesting that the Pab1 RRM12 binding to BOX2 and BOX3 acts as a second switch that releases the BOX1 oligomers. NMR data showed that Pub1 RRM3 did not interact with Pab1 RRM12 (Figure 5B middle panel). These data agree with the data described above in Figure 3A, where the presence of the Pab1+Pub1 mixture (Figure 3A, in green), caused larger BOX1 line broadening in the eIF4G1₁₋₂₄₉ spectrum than the presence of either Pub1 or Pab1 alone (Figure 3A, in orange and blue, respectively), which also suggested the existence of BOX1-driven oligomers.

These NMR analyses suggested a two-key mechanism whereby Pub1 and Pab1 bind to eIF4G1 causing conformational changes that promote BOX1 self-assembly (Figure 5C). These two RBPs interact with eIF4G1 elements that contact with BOX1 in the free state (Figure 2D). These contacts likely prevent BOX1 aggregation, whereas the coordinated effect of Pub1/Pab1 binding enhances it.

Pab1-Pub1-eIF4G1 form micrometer-size condensates

Our NMR analysis suggested that simultaneous binding of Pab1 RRM12, eIF4G1₁₋₂₄₉ and Pub1 RRM3 has the potential to form high order structures that cannot be detected by this technique because of their large size. To further investigate this possibility, we determined if different Pub1/Pab1/eIF4G1 combinations could form microscopic condensates that might resemble biological ones. For such experiments we used protein concentrations according to the number of molecules per yeast cell (SGD: <https://www.yeastgenome.org>) and Ficoll 70 (200 g/l) to simulate crowding in the cellular environment. Dynamic light scattering (DLS) indicated that large aggregates were formed in some protein combinations (Figure 7A). In contrast, the individual proteins exhibited autocorrelation functions that did not differ from that of Ficoll-70 alone, suggesting that the individual proteins did not aggregate. Moreover, the curve profiles of the single proteins remained stable for several hours. The Pab1:Pub1 mixture (Figure 7, row 2, column 1) showed the same behavior, but other double protein mixtures and the triple one showed a second phase, evidencing the presence of micrometer-size particles. These particles were present right from the beginning and appeared to reach a maximum within 2 hours of mixing (Figure 7A, row 3, column 3). Because, all of these combinations contained eIF4G1 and at least one RNA binding protein, we concluded that the interactions between eIF4G1 and Pab1/Pub1 promoted aggregation; probably by enhancing the intrinsic propensity of BOX1. Consistent with this using the eIF4G1 Δ BOX1 mutant in the triple mixture showed no aggregation (Figure 7, row 3, column 2). Confocal fluorescence microscopy images of triple mixtures containing Pab1 RRM12, Pub1 RRM123 and Alexa 488 labelled eIF4G1₁₋₂₄₉ in Ficoll 70 (200 g/L) revealed the presence of discrete rounded particles (~1 μ m and smaller, Figure 7 B upper panels). Both, Pab1 and Pub1 were observed to colocalize with eIF4G1 in these assemblies, as observed in fluorescent images in which the proteins were pairwise labelled with

spectrally different dyes (Alexa 488 and Alexa 647, Figure 7B middle and lower panels). Similar structures were observed for binary Pab1/eIF4G1 mixtures (Figure 7C), in good agreement with the DLS measurements.

These results showed that the eIF4G1/Pab1/Pub1 mixtures could form crowding-driven structures resembling those previously described for full length Pab1 (Riback et al., 2017) or Pub1 (Kroschwald et al., 2018), but without the requirement for pH or temperature stress.

eIF4G1 RNA recognition

Besides protein recognition, eIF4G1 binds RNA using three regions RNA1, RNA2 and RNA3 (Berset et al., 2003; Park et al., 2011). The construct eIF4G1₁₋₂₄₉ contains one of them (RNA1, eIF4G1₁₋₈₂). We titrate the eIF4G1₁₋₈₂ construct with three poly(A) probes and found that the strength of the interaction is higher with longer oligos (Figure 6A). The chemical shift changes almost doubled when going from A₁₂ to A₁₄ and a new signal appears in the spectra that corresponds to the side chain Arg guanidinium group (N ϵ -H ϵ) (Figure 6B). These spectral changes evidenced that at least one Arg side-chain is directly involved in RNA contacts, slowing down the otherwise rapid solvent exchange in the free protein (no Arg N ϵ -H ϵ signals). The interaction with poly(A) maps three regions centered around R₃₄, R₅₅ and G₆₅, peaking at R₅₅PH₅₇. In contrast, the canonical RGG box (R₆₀GG), a well-known RNA binding motif (Thandapani et al., 2013; Chong et al., 2018; Chowdhury and Jin, 2022), is less affected by binding. In general, the major changes occur in the region with more Arg/Aromatic density and the observed length-dependence of the binding strength probably reflects the simultaneous interactions with various sites (e.g. Arginines) of A₁₄.

Discussion

During the last 40 years stress granules have been described in multiple organisms and upon different stressors, and have been extensively characterized both morphologically and in their composition (Buchan and Parker, 2009; Buchan et al., 2011; Protter and Parker, 2016; Hofmann et al., 2021; Glauninger et al., 2022). However, despite the intensive scientific interest, there is still no clear functional assignment for this type of biomolecular condensates (Glauninger et al., 2022). It is commonly agreed that stress granules appear upon translation arrest, irrespectively the different pathways that lead to it (e.g. eIF2a phosphorylation, mTOR pathway, or targeting eIF4F function). The blockage of the translation machinery leads to polysome disassembly and ribosomes are probably displaced from the mRNA by RBPs such as Pub1, Scd6, Sbp1, etc (all of them characteristic SG markers). Since these RBPs interacts with a wide range of specificity, it is reasonable to think that the RBP load of a given mRNA depend on its length, which correlates well with the higher abundance of long mRNA in SG (Khong et al., 2017; Namkoong et al., 2018; Matheny et al., 2019). However, mRNP assembly does not look on itself sufficient to nucleate biomolecular condensates and evolve them to SG. Individual mRNPs need to bridge together to build a mesh like structure through protein-protein, protein-RNA and RNA-RNA interactions. Knowing the structural features of these interactions is essential to understand the biophysical grounds of SG formation.

Along this line, our work provides an extensive NMR characterization of eIF4G1₁₋₂₄₉ at the residue level, its complex conformational landscape and its interactions with other SG principal components: the RBPs Pub1 and Pab1. The N-terminus of eIF4G1 contains several short segments (BOXes) that are conserved in the *Saccharomycetales* family and act as short linear interaction motifs (SLiMs) for various purposes: (1) intramolecular and intermolecular self-recognition; (2) specific recognition of RBPs; and (3) RNA recognition.

The eIF4G1₁₋₂₄₉ conformational ensemble, among the few atomistic models of an IDP (Schwalbe et al., 2014; Cordeiro et al., 2019; Kubáň et al., 2019; Shrestha et al., 2019), shows a high degree of structural variability while maintaining a certain degree of compactness due to long-range contacts among several SLiMs. BOX1 stands out by defining a fuzzy network of π - π and π -cation interactions with RNA1 and BOX3 motifs that mask its prion-like tendency. The sequence of BOX1 shows overlapping Pub1 binding (YNN) and amyloid-like (YYNN) motifs, which are statistically underpopulated in the yeast proteome, but relatively abundant among Pub1 and eIF4G1 binding proteins and in SG core components (Supplementary Figure 6). We speculate that these motifs are SG hallmarks.

Other fundamental questions in the field are what are the mechanisms of nucleation of biomolecular condensates or how these small entities grow to large SG (assuming that both are related) (Glauninger et al., 2022). In the light of our *in vitro* results, we can propose a model pretending to address those questions partially. BOX1 self-assembly propensities are the key feature of the system (Figure 5C). Under non-stress situations, eIF4G1 associates with other translation initiation factors and ribosome 40S subunit to promote translation and BOX1 self-assembly is masked by transient contacts with other regions of the eIF4G1 N-terminal IDP. Translation arrest will depopulate polysomes and Pub1 (and other translational repressors) incorporate into the mRNP. The incoming Pub1 would compete out stabilizing interactions of BOX1, triggering condensation. In this model it is interesting to highlight the role of Pub1 RRM3. The interaction with eIF4G1 would not interfere with RNA recognition. Pub1 is the yeast homolog of TIA-1 and TIAR proteins, which are components of mammalian stress granules (Kedersha et al., 1999). We showed that this family of proteins has an RRM3 domain with a unique structure (Santiveri et al., 2011). Now our data suggest that this domain (and no other RRMs) can

form large assemblies with some eIF4G1 constructs. Finding if these assemblies have some degree of long order would be an important field of study for future research.

Acknowledgments

NMR experiments were performed in the “Manuel Rico” NMR laboratory (LMR) (CSIC). The synchrotron SAXS data was collected at beamline P12 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany). The study was supported by funds from the Spanish MINECO (CTQ2018-84371) to MJ and JMC-P, and Autonomous Community of Madrid (B2017/BMD-3770) to JMP-C. BC-A and SM-L were funded by predoctoral grants from Spanish MINECO (BES-2015-073383) and Autonomous Community of Madrid (CPI/0265/2008) respectively. This work was supported by the Labex EpiGenMed, an « Investissements d’avenir » program (ANR-10-LABX-12-01) awarded to PB. The CBS is a member of France-BioImaging (FBI) and the French Infrastructure for Integrated Structural Biology (FRISBI), 2 national infrastructures supported by the French National Research Agency (ANR-10-INBS-04-01 and ANR-10-INBS-05, respectively). We would like to thank Clara M. Santiveri and Francisco Blanco for their helpful suggestions and critical reading of the manuscript and M.T. Seisdedos and G. Elvira (Confocal Laser and Multidimensional Microscopy Facility, CIB-CSIC) for assistance in imaging.

.

Author contributions

BC-A, SM-L, SC and JMP-C made all of the clones, mutants and recombinant proteins. BC-A, SM-L and JMP-C obtained and analyzed NMR data with the help of MJ. Residual Dipolar Couplings were measured by BC-A and NS, and analyzed with the help of PB and NS. SAXS experiments were performed by BC-A and PB and analyzed by SM-L and JMP-C with the advice of PB. Paramagnetic Relaxation Experiments were performed and

analyzed by BC-A and JMP-C. Protein structure calculations were done by JMP-C. Confocal microscopy images were acquired and analyzed by BC-A, JMP-C and SZ with the help of SC. The project was conceived by JMP-C who wrote the paper with the assistance of BC-A and with contributions from other authors.

Declaration of interest

The authors declare no conflicts of interest

MATERIAL AND METHODS

Cloning, protein expression and purification

Plasmids and proteins used in this work are described in the supplementary table 1. DNA fragments corresponding to wild-type constructs of eIF4G1, Pub1 and Pab1 were amplified from *Saccharomyces cerevisiae* genomic DNA using the DNA polymerases KOD or Pfu. These DNA fragments were cloned into a pET28-modified vector that contains an N-terminal thioredoxin A fusion tag, an internal 6xHis tag and a TEV protease site. eIF4G1 mutants were obtained using the Quick-change Lightning Kit and specific DNA primers. Plasmids corresponding to mutant and wild-type proteins were transformed into *E. coli* BL21 (DE3) competent cells and expressed in kanamycin-containing (30 µg/l) LB medium.

For isotope labelling of samples, a K-MOPS derived minimal medium (Neidhardt et al., 1974) was supplemented with ¹⁵NH₄Cl (1 g/l) and/or ¹³C-glucose (4 g/l). Cultures of eIF4G1 and its mutants were grown at 37 °C until OD_{600nm}=0.6-0.8, when they were induced with 0.5 µM IPTG for 4 hours. Pab1 and Pub1 cultures, after reaching OD_{600nm}=0.6, were transferred to 25 °C for induction with IPTG overnight (12-16 hours).

For purification of all recombinant proteins, cell pellets were resuspended in lysis buffer (25 mM potassium phosphate pH 8.0, 300 mM NaCl, 10 mM Imidazole and 1 tablet/50

ml of protease inhibitors cocktail), lysed by sonication and cleared by ultracentrifugation. The supernatant was purified by metal affinity chromatography using a HiTrapTM 5 ml column and elution with 25 mM potassium phosphate buffer pH 8.0, containing 300 mM NaCl and 300 mM imidazole. The samples containing the fusion protein were exchanged into 20 mM Tris pH 8.0 (in the case of the Pab1 construct this buffer was supplemented with 1 mM DTT), and digested overnight at 4 °C with homemade TEV protease. In the case of the Pub1 and Pab1 constructs, the samples were re-loaded onto the HiTrap nickel column to capture the protease, the cleaved N-terminal part of the fusion protein and the undigested protein. The flow through was further purified by ion exchange using an anion exchanger column (Q 5 ml) for all proteins except for Pub1 RRM3 that was purified with a cation exchange column (SP 5 ml). In either case, proteins were eluted with a linear salt gradient (to 1 M NaCl). In the case of the eIF4G1 construct, we observed that the second nickel column negatively affected protein stability and aggregation; we therefore purified the protein away from the uncleaved protein, thioredoxin A and TEV using a cation exchange column (SP 5 ml). Finally, the purified proteins were concentrated and the buffer was exchanged according to their intended use.

Small Angle X-ray Scattering (SAXS) measurements

SAXS experiments were performed using the P12-EMBL beamline at the DESY synchrotron in Hamburg and were analysed with ATSAS software. All data were collected in batch using 25 mM potassium phosphate pH 6.5 and 25 mM NaCl. The concentrations used for analysis of eIF4G1₁₋₂₄₉ were 15 mg/ml, 10 mg/ml, 8 mg/ml, 5 mg/ml, 3 mg/ml and 1 mg/ml. The final SAXS curves were generated using PRIMUS and deposited in the SASBDB under the code SASDP88 (<https://www.sasbdb.org/data/SASDP88/678c641ui8/>).

NMR: resonance assignments and relaxation data

All samples were prepared in NMR buffer (25 mM potassium phosphate pH 6.5, 25 or 150 mM NaCl, 1 mM DTT and 10% D₂O) and experimental data were acquired at 25 °C on a cryoprobe-equipped Bruker AV800 MHz spectrometer. Assignment of the backbone ¹H, ¹⁵N and ¹³C atoms was achieved by following the standard methodology. The 3D HNCA, HNC(O), HN(CO)CA, CBCA(CO)NH and CBCANH experiments were used for backbone assignment and 3D (H)CCH-TOCSY were recorded to assign side chain resonances ((Sattler et al., 1999) and the references therein). Protein concentrations ranged between 100-200 μM. The chemical shifts were deposited in the Biomagnetic Resonance Database (BMRB) with codes 28121 (eIF4G₁₋₂₄₉) and 34517 (eIF4G_{135-49-Pub1 RRM3}). The ¹⁵N backbone amide relaxation T₁ and T₂ parameters were measured with series of ¹H-¹⁵N spectra of standard inversion-recovery and Carr-Purcell-Meiboom-Gill sequences (CPMG). NMR spectra were processed using TOPSPIN v4.1 (Bruker) and NMRPipe, and analyses were done with CcpNmr Analysis.

NMR: Residual dipolar couplings and PRE measurements

The filamentous phage Pf1 was used at a final concentration of 20 mg/ml to induce weak alignment of eIF4G₁₋₂₄₉ (200 μM in 25 mM potassium phosphate pH 6.5 and 25 mM NaCl). NMR experiments were carried out at 298 K in a Bruker Avance III 800 MHz spectrometer equipped with a cryogenic triple resonance probe. Two samples (isotropic and anisotropic) were prepared and couplings (J and J+D) were measured with ¹⁵N-HSQC-DSSE (In Phase Anti Phase IPAP). Experiments were processed using TOPSPIN v2.1 and NMRPipe, and were analyzed with the program CcpNmr Analysis.

For the paramagnetic relaxation enhancement, protein samples from different cysteine-containing eIF4G₁₋₂₄₉ mutants (S200C and Q109C) were chemically modified using the following protocol. Mutant protein samples (600-700 μM) were pre-treated with 5 mM DTT for two hours at room temperature. The DTT was then eliminated by fast buffer

exchange into 25 mM Tris pH 9.0 and 25 mM NaCl using a Nap-5 desalting column. Labelling with 4-(2-Iodoacetamido)-TEMPO was initiated immediately after column elution by adding a tenfold molar excess of the probe dissolved in ethanol (25 mM spin label stock). The reaction was allowed to proceed for 30 minutes at room temperature in the dark. The excess iodoacetamide label was quenched with 10 mM 2-mercaptoethanol for 10 minutes, and afterwards the protein adduct was exchanged into 25 mM potassium phosphate pH 6.5, 25 mM NaCl and 1 mM DTT for later use. The NMR samples were prepared in 5 mm tubes sealed in an N₂ atmosphere to avoid reduction by air, and high resolution ¹H-¹⁵N HSQC spectra were recorded for the oxidized state (active spin label). Subsequently, the spin label was reduced with 10 μM ascorbate (Gillespie and Shortle, 1997), and the reference ¹H-¹⁵N HSQC without paramagnetic relaxation enhancement was recorded. The relaxation effect was calculated as the intensity ratios between peaks in the two spectra.

Structure calculations of eIF4G₁₃₅₋₄₉-Pub1 RRM3 and the eIF4G BOX3 peptide

The NMR structure of the eIF4G₁₈₇₋₂₃₄ construct was determined from NOE-derived distance restraints (2D NOESY spectrum with 60 ms mixing time) and angular restraints (from ¹³C chemical shifts and TALOS+) using the program Cyana. Protein assignments were obtained by comparison with other eIF4G1 constructs and confirmed by triple resonance 3D spectra: CBCA(CO)NH, HNCACB and HNCO.

Two different chimeras of Pub1 with eIF4G₁₃₇₋₅₁ were constructed, with eIF4G₁₃₅₋₄₉ fused either to the N- or C- terminus of Pub1 RRM3. Of these constructs, only the N-terminal fusion (eIF4G₁₃₅₋₄₉) proved to have the right topology and the structure was determined using a similar protocol to the analysis of the structure of Pub1 RRM3 (Santiveri et al., 2011) using distance restraints from a 2D NOESY (60 ms mixing time).

Structure calculation of eIF4G₁₋₂₄₉

eIF4G1₁₋₂₄₉ structures (80000) were calculated with the program Cyana 3.0 using: 1) experimental NOE-derived distance restraints for the BOX3 region; 2) π - π interactions between Tyr, Phe and Trp; and 3) π -cation interactions between Arg and Tyr/Phe/Trp. The latter two interaction types were referred to as knowledge-based constraints (K-BC) and were included, given the importance of these types of contacts for IDP interactions (see main text for references). For each individual structure calculation, the origin residue (Arg/Tyr/Phe/Trp) was randomly selected (80% probability) and ambiguous restraints were generated for the other interaction partner (Arg/Tyr/Phe/Trp). In this way, each of the individual structure calculations contains a unique set of knowledge-based distance restraints. This protocol ensures high variability by avoiding biases of specific pairwise interactions. Thus, the final interactions present on each conformer are freely selected during the calculations. A similar protocol was followed to calculate the structures (80000) of eIF4G1₁₋₂₄₉ dimers, using the same experimental restraints and ambiguous Tyr-Tyr contacts as dimerization driving interactions.

The theoretical PRE-derived intensity ratios were calculated using equations in (Battiste and Wagner, 2000) (Supplementary Figure 5) for the eIF4G1₁₋₂₄₉ monomers and dimers. Correlation time τ_c was estimated from the averaged T_1/T_2 and d was estimated from the distances between amide backbones (N) and Q109/S200 side-chains (C β). We next used a home-made greedy algorithm to select the ensembles that better reproduced the PRE profiles. The algorithm calculates the residual to the experimental data (Q109 and S200) and chooses, as a seed, the conformer that better agrees (lower sum of residuals) with the data. In the next steps, the residuals were computed across 2, 3,..., n structures always choosing the combination with minimal violations. The procedure was repeated until the ensemble size reached 2000 conformers (~2.5% of the original). The first 500 structures

were included in the final eIF4G1₁₋₂₄₉ monomer and eIF4G1₁₋₂₄₉ dimer ensembles (PRE-derived).

We used the EOM protocol to fit the SAXS data. Pools of theoretical SAXS curves were constructed for the eIF4G1₁₋₂₄₉ monomers (2000 conformers) and dimers (2000 conformers). These two pools were combined with the genetic algorithm in the EOM program to model the curve with a fixed size ensemble (50 structures). The percentages of each pool were freely selected by the algorithm and the procedure. The procedure was repeated 10 times, obtaining a 500-member set. It should be noted that some of the structures are repeated between individual EOM calculations. The structural ensembles were deposited in the Protein Ensemble Database (PED) with the codes PED^{XX} and PED^{XX}. The theoretical values of the PREs, and other structural properties, were calculated as averages across the different ensembles using home-made perl scripts.

Dynamic Light Scattering

The DLS measurements were carried out at 25 °C in a DynaPro Titan (Wyatt Technologies) instrument and were analysed with Dynamics V6 software. Protein mixtures (eIF4G1₁₋₂₄₉, eIF4G1₁₋₂₄₉ ΔBOX1, Pub1 RRM123 and Pab1 RRM12) were prepared from extensively centrifuged stocks (>1 h at 15000 RPM; 4 °C), and were filtered (0.22 μm) in PBS buffer and 200 g/l Ficoll 70 stock. Samples were mixed thoroughly and placed in a plastic cuvette (Eppendorf) for measurements. Individual correlation curves were recorded (10 acquisitions of 10s) every 15 minutes over a 5 h period.

Confocal microscopy

eIF4G1₁₋₂₄₉, Pab1 RRM12 and Pub1 RRM123 proteins were purified as described above and were labelled with Alexa Fluor 488 or Alexa Fluor 647 carboxylic acid succinimidyl ester dyes (Molecular Probes), using protein to probe ratios of 1:3. The coupling reaction

was carried out in the dark in PBS (pH 7.4) buffer for 30 minutes on ice and the unreacted probe was removed by size exclusion chromatography using a Nap-5 column. Samples for visualization were prepared by mixing eIF4G₁₋₂₄₉ with Pub1 RRM123 and Pab1 RRM12 in different combinations and at concentrations of 1, 5 and 20 μ M respectively, in PBS, 0.1 mM DTT (pH 7.4) buffer and 200 g/l Ficoll 70. Samples, that contain 1 μ M of fluorescently labelled protein (Alexa Fluor 488-labelled eIF4G1 and/or Alexa Fluor 647-labelled Pub1 or Pab1) for visualization, were placed in silicone chambers that were glued to coverslips and were visualized with Leica TCS SP2 or TCS-SP5 inverted confocal microscopes with a HCX PL APO 63x oil immersion objective (N.A. = 1.4; Leica, Mannheim, Germany). Alexa 488 and Alexa 467 were excited using 488 nm and 633 nm laser excitation lines, respectively. The concentration of the various Alexa-labelled proteins was kept at 1 μ M and the solution was supplemented with unlabeled proteins to reach the above-mentioned concentrations. Various images were registered for each sample, corresponding to different observation fields.

Figure legends

Figure 1. Sequence features and NMR spectrum of eIF4G1₁₋₂₄₉. (A) Schematic representation of *S. cerevisiae* eIF4G1 showing the interaction domains for the other two components of the eIF4F heterotrimer (in black). The N-terminal region containing the conserved boxes (residues 1-249) is highlighted in yellow. The conserved features (BOXes) of eIF4G1₁₋₂₄₉ are represented below including weblogs indicating the conservation of each box across *Saccharomyces* and the position of key side chains (Tyr, Phe, Trp and Arg) capable of π - π and π -cation interactions. (B) ¹H-¹⁵N HMQSC spectrum of eIF4G1₁₋₂₄₉. Each assigned residue is labeled: blue, major form; red, minor species.

Figure 2. NMR structural analysis of eIF4G1₁₋₂₄₉. (A) NMR evidence of residual secondary structure (per residue): percentage of predicted secondary structure (SS pred) calculated using the d2D program (Camilloni et al., 2012) (upper graph), ¹⁵N relaxation T₁/T₂ (middle graph, dotted red line marks the average) and residual dipolar couplings (RDCs) (lower graph). Conserved sequence elements in eIF4G1₁₋₂₄₉ are represented at the top of the figure. (B) Per residue effect of paramagnetic relaxation enhancement (PRE) over relative signal intensities of two individual ¹⁵N-labelled eIF4G1₁₋₂₄₉ mutants and of ¹⁵N-labelled wild-type eIF4G1₁₋₂₄₉ mixed (1:1) with spin-labeled eIF4G1₁₋₂₄₉ Q109C at natural isotopic abundance. Green circles mark the position of spin-labels. The red lines indicate the back-calculated PRE effects across the 500-member ensembles of monomers (upper and middle graphs) and dimers (lower graph). (C) Experimental SAXS curve showing log of intensity (I) versus log of scatter (S) of eIF4G1₁₋₂₄₉ and EOM fitting obtained from the eIF4G1₁₋₂₄₉ atomic models (red curve). Inset: Guinier analysis and the derived Radius of gyration (R_g) and forward scattering intensity (I(0)) values. (D) (left panel) Ensemble-averaged intramolecular C α contact maps obtained from the monomers and dimers (upper and lower triangles, respectively), and (right panel) intermolecular C α

contact maps. The average distances (in Å) were color-coded according to the scale in the middle, and the conserved boxes of eIF4G1₁₋₂₄₉ are indicated along the axes of each map.

Figure 3. Mapping of Pub1 and Pab1 binding sites on eIF4G1₁₋₂₄₉. (A) Sequence dependence of eIF4G1₁₋₂₄₉¹H-¹⁵N HSQC signal intensity ratios between free eIF4G1₁₋₂₄₉ and various eIF4G1₁₋₂₄₉ complexes with Pub1 RRM3 (orange), Pab1 RRM12 (blue) and Pub1 RRM3+Pab1 RRM12 (green). A representative region of the eIF4G1₁₋₂₄₉¹H-¹⁵N HSQC spectrum (in grey) is shown below superimposed with the equivalent spectra of the Pub1 RRM3 (in orange), Pab1 RRM12 (blue) and Pub1 RRM3+Pab1 RRM12 (green) complexes. Specific residues were labelled to illustrate the selective disappearance of eIF4G1₁₋₂₄₉ signals upon complex formation in each case (marked with dots in the bar charts). (B, C) NMR study of the interaction of eIF4G1 peptides with Pub1 RRM3 (B) and Pab1 RRM12 (C) monitored on their ¹H-¹⁵N HSQC spectra. The regions of eIF4G1₁₋₂₄₉ that correspond to each of the five peptides tested are indicated at the top in (B). Black: spectra of the free proteins; Colors: peptide titrations according to the color scheme in (B). Chemical shift perturbations and signal broadening were mapped onto the model structures of Pub1 RRM3 (PDB:2LA4) and Pab1 RRM12 (bottom).

Figure 4. Structural characterization of the eIF4G1-Pub1 interaction. (A) Superposition of 2D ¹H-¹⁵N HSQC of free Pub1 RRM3 (black) on either eIF4G1₃₅₋₄₉-Pub1 RRM3 (red) or Pub1 RRM3-eIF4G1₃₅₋₄₉ (green) chimeric constructs. (B) NMR structure of eIF4G1₃₅₋₄₉-Pub1 RRM3. Yellow, eIF4G1; White, Pub1 RRM3; Blue, Pub1 RRM3 residues contacting eIF4G1. Key interacting residues are labelled on both eIF4G1 and Pub1RRM3 regions of the chimera.

Figure 5. Pub1 RRM3 can interact with eIF4G1 through two different modes. (A) Superposition of the ¹H-¹⁵N HSQC spectra of free Pub1 RRM3 (black) and Pub1 RRM3 in complex with eIF4G1₁₋₃₄₈ (green) or eIF4G1₁₋₁₈₄ (red). The Pub1 RRM3 NMR

crosspeaks show small perturbations in the first complex (left panel) and completely disappear (with the exception of the highly mobile Asn/Gln side chain peaks) in the spectrum of the complex (right panel). The central panel shows eIF4G1 mutants that had similar effects as eIF4G1₁₋₃₄₈ (outlined in green dash) or eIF4G1₁₋₁₈₄ (outlined in red dash) on the Pub1 RRM3 spectrum. **(B)** Study of the effect of eIF4G1₁₋₂₄₉ and/or Pab1 RRM12 titrations on the Pub1 RRM3 ¹H-¹⁵N HSQC spectrum. Superposition of the Pub1 RRM3 spectra before (black signals) and after titration with eIF4G1₁₋₂₄₉ (left, green signals), Pab1 RRM12 (middle, blue signals) and eIF4G1₁₋₂₄₉ + Pab1 RRM12 (right, red signals). A small area of each spectrum is expanded below for a more detailed view. The Pub1 RRM3 signals are unperturbed upon titration with Pab1 RRM12 (blue spectrum), but disappear when Pab1 RRM12 is combined with eIF4G1₁₋₂₄₉ (red spectrum). **(C)** A dual key interaction model between eIF4G1₁₋₂₄₉, Pub1 RRM3 and Pab1 RRM12 to explain oligomerization. Multiple intramolecular interactions between BOX1 and other conserved elements of eIF4G1₁₋₂₄₉ maintain it predominantly monomeric. Weak interactions with Pub1 RRM3 (first key) cancel out some of these transient contacts, but some remain (e.g. BOX1-BOX3) preventing BOX1-driven oligomerization. The interaction with Pab1 RRM12 (second key) further blocks internal contacts to BOX1 triggering its aggregation. Pub1 RRM3 binding to BOX1 oligomeric forms would be reinforced by Pub1-Pub1 interactions.

Figure 6. RNA recognition by eIF4G1₁₋₈₂. **(A)** Chemical shift mapping (CSM) of various poly(A) probes. Aromatics (Y, H and F) and basic residues (R and K) are drawn below the sequence. **(B)** Superposition of the ¹H-¹⁵N-HSQC of the free (black) and A₁₄-bound eIF4G1₁₋₈₂ (red).

Figure 7. Formation of condensates by Pub1/Pab1/eIF4G1 mixtures in crowding conditions. **(A)** DLS analysis of the indicated individual Pub1 (Pub1 RRM123), Pab1

(Pab1 RRM12), and eIF4G1 (eIF4G1₁₋₂₄₉ or eIF4G1₁₋₂₄₉ ΔBOX1) proteins, and of their double and triple mixtures at the concentrations indicated in the figure in the presence of Ficoll (Fc) 70k. Time-dependence (in μs) of the autocorrelation functions $G(t)$ is shown for different protein mixtures and recorded a different time after mixing (panel code shown in the upper left panel). The lower right panel (row 3, column 3) shows the time evolution of the autocorrelation function at 1000 μs for different mixtures of previous graphs, that correspond to the second phase associated to the aggregates. **(B)** Representative confocal fluorescence microscopic images of ternary mixtures eIF4G1₁₋₂₄₉ (eIF4G1) labelled with Alexa 488 dye (A488) and Pab1 RRM12 (Pab1) plus Pub1 RRM123 (Pub1). In this ternary mixture, Pab1 plus Pub1 are either unlabeled (row 1, second column) or one of them is labeled with Alexa 647 dye (A647) as indicated in the figure (2nd and 3rd rows), while the other is unlabeled. **(C)** Confocal images of the mixture of eIF4G1₁₋₂₄₉ (eIF4G1) and Pab1 RRM12 (Pab1), labelled with Alexa 488 and Alexa 647, respectively. In B and C, when present, the final concentration of eIF4G1₁₋₂₄₉, Pab1 RRM12 and Pub1 RRM123 was 1, 20 and 5 μM , respectively. The concentration of labelled proteins was 1 μM and additional unlabeled protein was added to achieve the indicated final concentration. Scale bars, 5 μm . Inset scale bar, 1 μm . In A, B and C, samples contained Ficoll 70 (200 g/L) as a crowding agent.

REFERENCES

- Aitken, C. E., and Lorsch, J. R. (2012). A mechanistic overview of translation initiation in eukaryotes. *Nat Struct Mol Biol* 19, 568–576. doi: 10.1038/nsmb.2303.
- Battiste, J. L., and Wagner, G. (2000). Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. *Biochemistry* 39, 5355–5365. doi: 10.1021/bi000060h.
- Bernadó, P., Blanchard, L., Timmins, P., Marion, D., Ruigrok, R. W., and Blackledge, M. (2005). A structural model for unfolded proteins from residual dipolar couplings and small-angle x-ray scattering. *Proc Natl Acad Sci U S A* 102, 17002–17007. doi: 10.1073/pnas.0506202102.
- Bernadó, P., Mylonas, E., Petoukhov, M. v, Blackledge, M., and Svergun, D. I. (2007). Structural characterization of flexible proteins using small-angle X-ray scattering. *J Am Chem Soc* 129, 5656–5664. doi: 10.1021/ja069124n.
- Berset, C., Zurbriggen, A., Djafarzadeh, S., Altmann, M., and Trachsel, H. (2003). RNA-binding activity of translation initiation factor eIF4G1 from *Saccharomyces cerevisiae*. *RNA* 9, 871–880. doi: 10.1261/rna.5380903.
- Bhattacharya, S., and Lin, X. C. (2019). Recent Advances in Computational Protocols Addressing Intrinsically Disordered Proteins. *Biomolecules* 9. doi: 10.3390/biom9040146.
- Brangwynne, C. P., Tompa, P., and Pappu, R. v (2015). Polymer physics of intracellular phase transitions. *Nature Physics* 11, 899–904. doi: 10.1038/NPHYS3532.
- Buchan, J. R., Muhlrad, D., and Parker, R. (2008). P bodies promote stress granule assembly in *Saccharomyces cerevisiae*. *J Cell Biol* 183, 441–455. doi: 10.1083/jcb.200807043.
- Buchan, J. R., and Parker, R. (2009). Eukaryotic Stress Granules: The Ins and Outs of Translation. *Molecular Cell* 36, 932–941. doi: 10.1016/j.molcel.2009.11.020.
- Buchan, J. R., Yoon, J. H., and Parker, R. (2011). Stress-specific composition, assembly and kinetics of stress granules in *Saccharomyces cerevisiae*. *J Cell Sci* 124, 228–239. doi: 10.1242/jcs.078444.
- Camilloni, C., de Simone, A., Vranken, W. F., and Vendruscolo, M. (2012). Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts. *Biochemistry* 51, 2224–2231. doi: 10.1021/bi3001825.
- Chong, P. A., Vernon, R. M., and Forman-Kay, J. D. (2018). RGG/RG Motif Regions in RNA Binding and Phase Separation. *Journal of Molecular Biology* 430, 4650–4665. doi: 10.1016/j.jmb.2018.06.014.
- Chowdhury, M. N., and Jin, H. (2022). The RGG motif proteins: Interactions, functions, and regulations. *WIREs RNA*. doi: 10.1002/wrna.1748.
- Cordeiro, T. N., Sibille, N., Germain, P., Barthe, P., Boulahtouf, A., Allemand, F., et al. (2019). Interplay of Protein Disorder in Retinoic Acid Receptor Heterodimer and Its Corepressor Regulates Gene Expression. *Structure* 27, 1270-1285.e6. doi: 10.1016/j.str.2019.05.001.

- Das, P., Matysiak, S., and Mittal, J. (2018). Looking at the Disordered Proteins through the Computational Microscope. *Acs Central Science* 4, 534–542. doi: 10.1021/acscentsci.7b00626.
- Estaña, A., Sibille, N., Delaforge, E., Vaisset, M., Cortés, J., and Bernadó, P. (2019). Realistic Ensemble Models of Intrinsically Disordered Proteins Using a Structure-Encoding Coil Database. *Structure* 27, 381–391.e2. doi: 10.1016/j.str.2018.10.016.
- Gibbs, E. B., Cook, E. C., and Showalter, S. A. (2017). Application of NMR to studies of intrinsically disordered proteins. *Archives of Biochemistry and Biophysics* 628, 57–70. doi: 10.1016/j.abb.2017.05.008.
- Gillespie, J. R., and Shortle, D. (1997). Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels. *J Mol Biol* 268, 158–169. doi: 10.1006/jmbi.1997.0954.
- Glauninger, H., Wong Hickernell, C. J., Bard, J. A. M., and Drummond, D. A. (2022). Stressful steps: Progress and challenges in understanding stress-induced mRNA condensation and accumulation in stress granules. *Molecular Cell*. doi: 10.1016/J.MOLCEL.2022.05.014.
- Gomes, E., and Shorter, J. (2019). The molecular language of membraneless organelles. *Journal of Biological Chemistry* 294, 7115–7127. doi: 10.1074/jbc.TM118.001192.
- Goyer, C., Altmann, M., Lee, H. S., Blanc, A., Deshmukh, M., Woolford, J. L., et al. (1993). TIF4631 and TIF4632: two yeast genes encoding the high-molecular-weight subunits of the cap-binding protein complex (eukaryotic initiation factor 4F) contain an RNA recognition motif-like sequence and carry out an essential function. *Mol Cell Biol* 13, 4860–4874. doi: 10.1128/mcb.13.8.4860.
- Güntert, P., and Buchner, L. (2015). Combined automated NOE assignment and structure calculation with CYANA. *J Biomol NMR* 62, 453–471. doi: 10.1007/s10858-015-9924-9.
- Hershey, P. E., McWhirter, S. M., Gross, J. D., Wagner, G., Alber, T., and Sachs, A. B. (1999). The Cap-binding protein eIF4E promotes folding of a functional domain of yeast translation initiation factor eIF4G1. *J Biol Chem* 274, 21297–21304. doi: 10.1074/jbc.274.30.21297.
- Hilliker, A., Gao, Z., Jankowsky, E., and Parker, R. (2011). The DEAD-Box Protein Ded1 Modulates Translation by the Formation and Resolution of an eIF4F-mRNA Complex. *Molecular Cell* 43, 962–972. doi: 10.1016/j.molcel.2011.08.008.
- Hofmann, S., Kedersha, N., Anderson, P., and Ivanov, P. (2021). Molecular mechanisms of stress granule assembly and disassembly. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1868, 118876. doi: 10.1016/J.BBAMCR.2020.118876.
- Hoyle, N. P., Castelli, L. M., Campbell, S. G., Holmes, L. E., and Ashe, M. P. (2007). Stress-dependent relocalization of translationally primed mRNPs to cytoplasmic granules that are kinetically and spatially distinct from P-bodies. *J Cell Biol* 179, 65–74. doi: 10.1083/jcb.200707010.

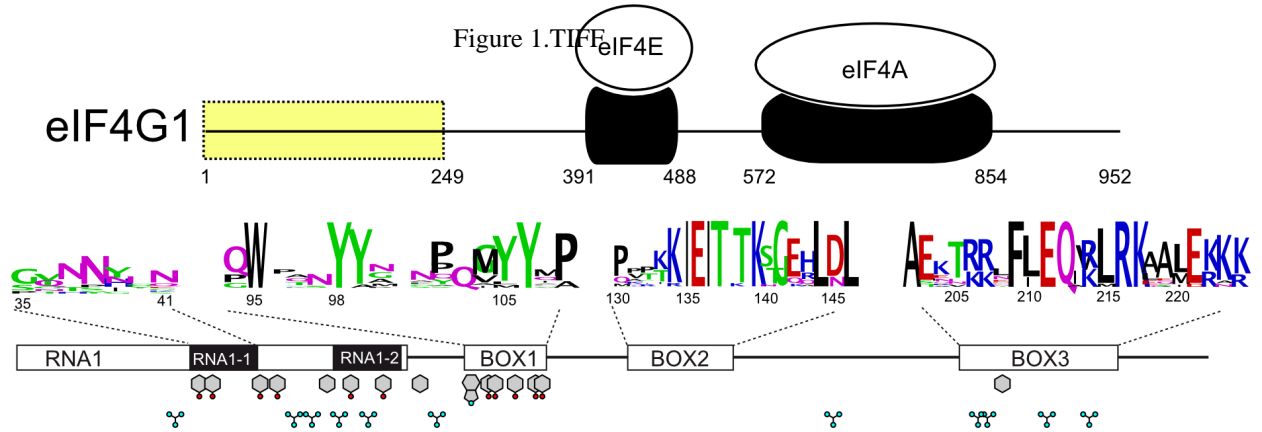
- Jain, S., Wheeler, J. R., Walters, R. W., Agrawal, A., Barsic, A., and Parker, R. (2016). ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. *Cell* 164, 487–498. doi: 10.1016/j.cell.2015.12.038.
- Kedersha, N. L., Gupta, M., Li, W., Miller, I., and Anderson, P. (1999). RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF-2 alpha to the assembly of mammalian stress granules. *J Cell Biol* 147, 1431–1442. doi: 10.1083/jcb.147.7.1431.
- Kessler, S. H., and Sachs, A. B. (1998). RNA recognition motif 2 of yeast Pab1p is required for its functional interaction with eukaryotic translation initiation factor 4G. *Mol Cell Biol* 18, 51–57. doi: 10.1128/mcb.18.1.51.
- Khong, A., Matheny, T., Jain, S., Mitchell, S. F., Wheeler, J. R., and Parker, R. (2017). The Stress Granule Transcriptome Reveals Principles of mRNA Accumulation in Stress Granules. *Molecular Cell* 68, 808-820.e5. doi: 10.1016/j.molcel.2017.10.015.
- Kragelj, J., Blackledge, M., and Jensen, M. R. (2015). Ensemble Calculation for Intrinsically Disordered Proteins Using NMR Parameters. *Adv Exp Med Biol* 870, 123–147. doi: 10.1007/978-3-319-20164-1_4.
- Kroschwald, S., Munder, M. C., Maharana, S., Franzmann, T. M., Richter, D., Ruer, M., et al. (2018). Different Material States of Pub1 Condensates Define Distinct Modes of Stress Adaptation and Recovery. *Cell Reports* 23, 3327–3339. doi: 10.1016/j.celrep.2018.05.041.
- Kubáň, V., Srb, P., Štégnerová, H., Padrta, P., Zachrdla, M., Jaseňáková, Z., et al. (2019). Quantitative Conformational Analysis of Functionally Important Electrostatic Interactions in the Intrinsically Disordered Region of Delta Subunit of Bacterial RNA Polymerase. *J Am Chem Soc* 141, 16817–16828. doi: 10.1021/jacs.9b07837.
- Kurzbach, D., Kontaxis, G., Coudeville, N., and Konrat, R. (2015). NMR Spectroscopic Studies of the Conformational Ensembles of Intrinsically Disordered Proteins. *Adv Exp Med Biol* 870, 149–185. doi: 10.1007/978-3-319-20164-1_5.
- Lin, Y., Protter, D. S., Rosen, M. K., and Parker, R. (2015). Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol Cell* 60, 208–219. doi: 10.1016/j.molcel.2015.08.018.
- Matheny, T., Rao, B. S., and Parker, R. (2019). Transcriptome-Wide Comparison of Stress Granules and P-Bodies Reveals that Translation Plays a Major Role in RNA Partitioning. *Molecular and Cellular Biology* 39. doi: 10.1128/MCB.00313-19.
- Milles, S., Salvi, N., Blackledge, M., and Jensen, M. R. (2018). Characterization of intrinsically disordered proteins and their dynamic complexes: From in vitro to cell-like environments. *Prog Nucl Magn Reson Spectrosc* 109, 79–100. doi: 10.1016/j.pnmrs.2018.07.001.
- Namkoong, S., Ho, A., Woo, Y. M., Kwak, H., and Lee, J. H. (2018). Systematic Characterization of Stress-Induced RNA Granulation. *Molecular Cell* 70, 175-187.e8. doi: 10.1016/j.molcel.2018.02.025.
- Neidhardt, F. C., Bloch, P. L., and Smith, D. F. (1974). Culture medium for enterobacteria. *J Bacteriol* 119, 736–747. doi: 10.1128/jb.119.3.736-747.1974.

- Park, E.-H. H., Walker, S. E., Lee, J. M., Rothenburg, S., Lorsch, J. R., Hinnebusch, A. G., et al. (2011). Multiple elements in the eIF4G1 N-terminus promote assembly of eIF4G1•PABP mRNPs in vivo. *The EMBO Journal* 30, 302–316. doi: 10.1038/emboj.2010.312.
- Poornima, G., Shah, S., Vignesh, V., Parker, R., and Rajyaguru, P. I. (2016). Arginine methylation promotes translation repression activity of eIF4G-binding protein, Scd6. *Nucleic Acids Res* 44, 9358–9368. doi: 10.1093/nar/gkw762.
- Preiss, T., and Hentze, M. W. (1999). From factors to mechanisms: translation and translational control in eukaryotes. *Curr Opin Genet Dev* 9, 515–521. Available at: <https://www.ncbi.nlm.nih.gov/pubmed/10508691>.
- Protter, D. S. W., and Parker, R. (2016). Principles and Properties of Stress Granules. *Trends Cell Biol* 26, 668–679. doi: 10.1016/j.tcb.2016.05.004.
- Riback, J. A., Katanski, C. D., Kear-Scott, J. L., Pilipenko, E. v, Rojek, A. E., Sosnick, T. R., et al. (2017). Stress-Triggered Phase Separation Is an Adaptive, Evolutionarily Tuned Response. *Cell* 168, 1028-1040 e19. doi: 10.1016/j.cell.2017.02.027.
- Robinson, N. E., and Robinson, A. B. (2001). Prediction of protein deamidation rates from primary and three-dimensional structure. *Proc Natl Acad Sci U S A* 98, 4367–4372. doi: 10.1073/pnas.071066498.
- Safaei, N., Kozlov, G., Noronha, A. M., Xie, J. W., Wilds, C. J., and Gehring, K. (2012). Interdomain Allosteric Promotes Assembly of the Poly(A) mRNA Complex with PABP and eIF4G. *Molecular Cell* 48, 375–386. doi: 10.1016/j.molcel.2012.09.001.
- Santiveri, C. M., Mirassou, Y., Rico-Lastres, P., Martínez-Lumbreras, S., Manuel Pérez-Canadillas, J., Martínez-Lumbreras, S., et al. (2011). Pub1p C-Terminal RRM Domain Interacts with Tif4631p through a Conserved Region Neighbouring the Pab1p Binding Site. *Plos One* 6, e24481. doi: 10.1371/journal.pone.0024481.
- Sattler, M., Schleucher, J., and Griesinger, C. (1999). Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Progress in Nuclear Magnetic Resonance Spectroscopy* 34, 93–158. doi: 10.1016/s0079-6565(98)00025-9.
- Schütz, P., Bumann, M., Oberholzer, A. E., Bieniossek, C., Trachsel, H., Altmann, M., et al. (2008). Crystal structure of the yeast eIF4A-eIF4G complex: an RNA-helicase controlled by protein-protein interactions. *Proc Natl Acad Sci U S A* 105, 9564–9569. doi: 10.1073/pnas.0800418105.
- Schwalbe, M., Ozenne, V., Bibow, S., Jaremko, M., Jaremko, L., Gajda, M., et al. (2014). Predictive atomic resolution descriptions of intrinsically disordered hTau40 and α -synuclein in solution from NMR and small angle scattering. *Structure* 22, 238–249. doi: 10.1016/j.str.2013.10.020.
- Shrestha, U. R., Juneja, P., Zhang, Q., Gurumoorthy, V., Borreguero, J. M., Urban, V., et al. (2019). Generation of the configurational ensemble of an intrinsically disordered protein from unbiased molecular dynamics simulation. *Proc Natl Acad Sci U S A* 116, 20446–20452. doi: 10.1073/pnas.1907251116.

- Sonenberg, N., and Hinnebusch, A. G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* 136, 731–745. doi: 10.1016/j.cell.2009.01.042.
- Tarun, S. Z., and Sachs, A. B. (1996). Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. *EMBO J* 15, 7168–7177. Available at: <https://www.ncbi.nlm.nih.gov/pubmed/9003792>.
- Tarun, S. Z., Wells, S. E., Deardorff, J. A., and Sachs, A. B. (1997). Translation initiation factor eIF4G mediates in vitro poly(A) tail-dependent translation. *Proc Natl Acad Sci U S A* 94, 9046–9051. doi: 10.1073/pnas.94.17.9046.
- Thandapani, P., O'Connor, T. R., Bailey, T. L., and Richard, S. (2013). Defining the RGG/RG Motif. *Mol Cell* 50, 613–623. doi: 10.1016/j.molcel.2013.05.021.
- Tria, G., Mertens, H. D., Kachala, M., and Svergun, D. I. (2015). Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ* 2, 207–217. doi: 10.1107/S205225251500202X.
- Vernon, R. M., Chong, P. A., Tsang, B., Kim, T. H., Bah, A., Farber, P., et al. (2018). Pi-Pi contacts are an overlooked protein feature relevant to phase separation. *Elife* 7. doi: 10.7554/eLife.31486.
- Wang, J., Choi, J. M., Holehouse, A. S., Lee, H. O., Zhang, X., Jahnel, M., et al. (2018). A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell* 174, 688–699 e16. doi: 10.1016/j.cell.2018.06.006.
- Wells, S. E., Hillner, P. E., Vale, R. D., and Sachs, A. B. (1998). Circularization of mRNA by eukaryotic translation initiation factors. *Mol Cell* 2, 135–140. Available at: <https://www.ncbi.nlm.nih.gov/pubmed/9702200>.
- Yakupova, E. I., Bobyleva, L. G., Vikhlyantsev, I. M., and Bobylev, A. G. (2019). Congo Red and amyloids: history and relationship. *Bioscience Reports* 39. doi: 10.1042/BSR20181415.

Figure 1.TIFF

A



B

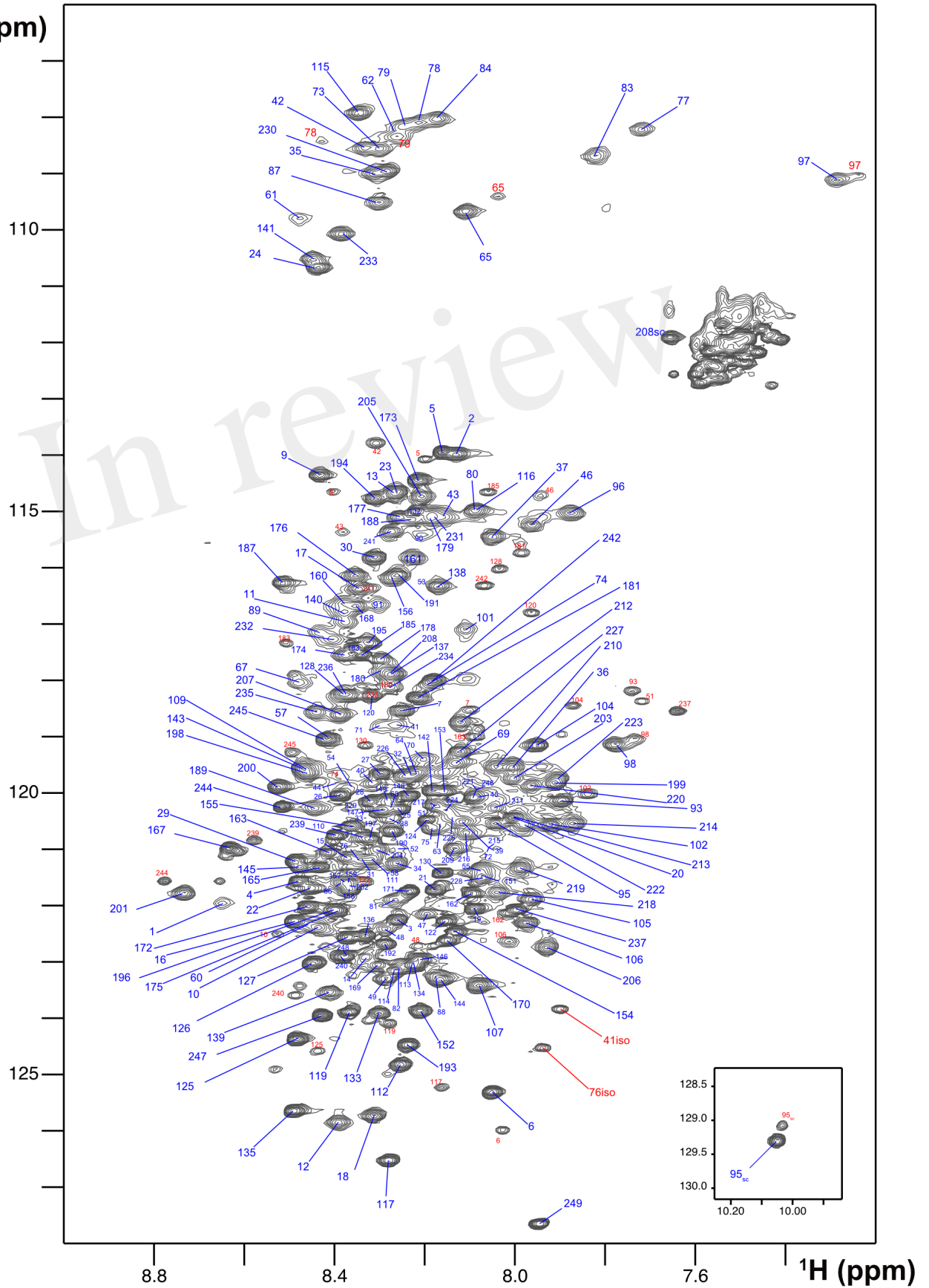
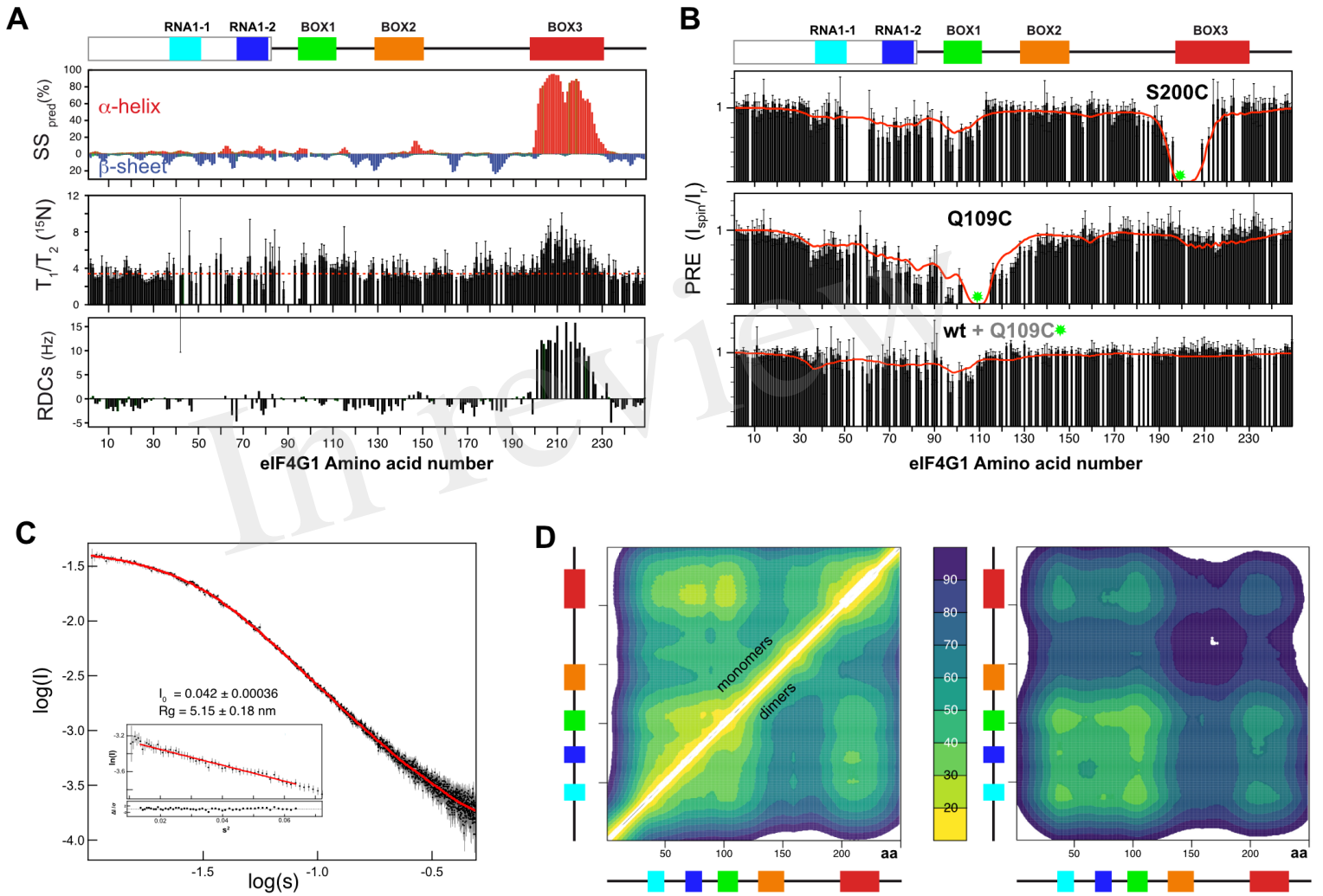
 ^{15}N (ppm)

Figure 2.TIFF



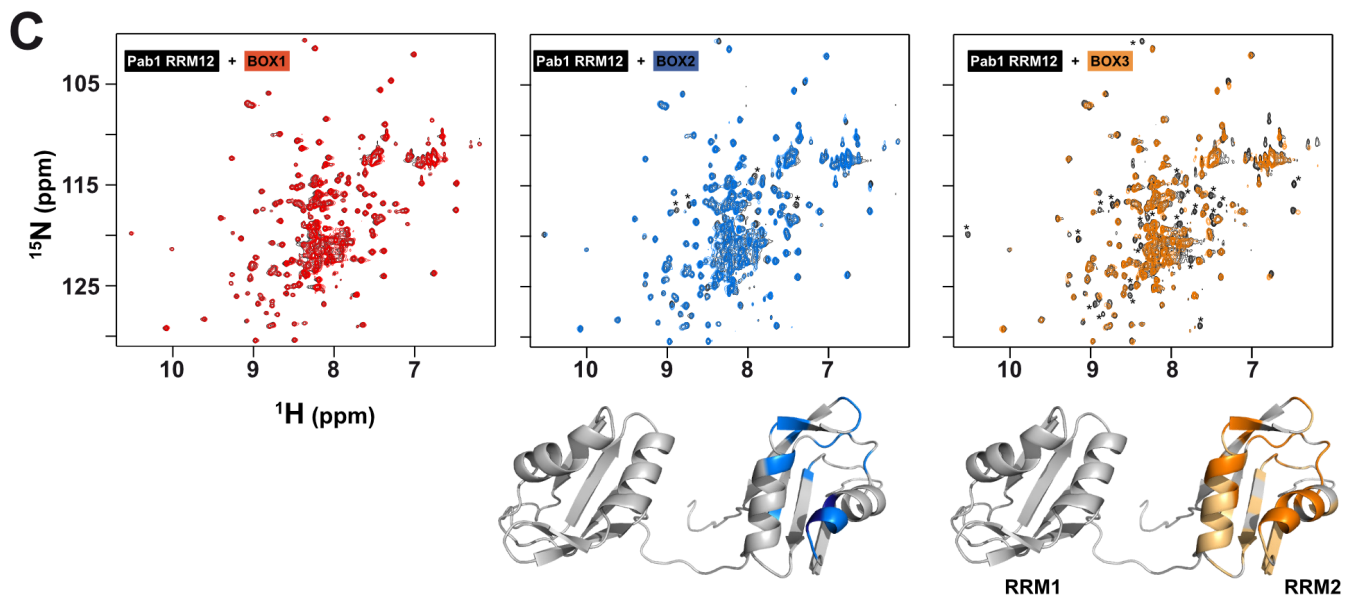
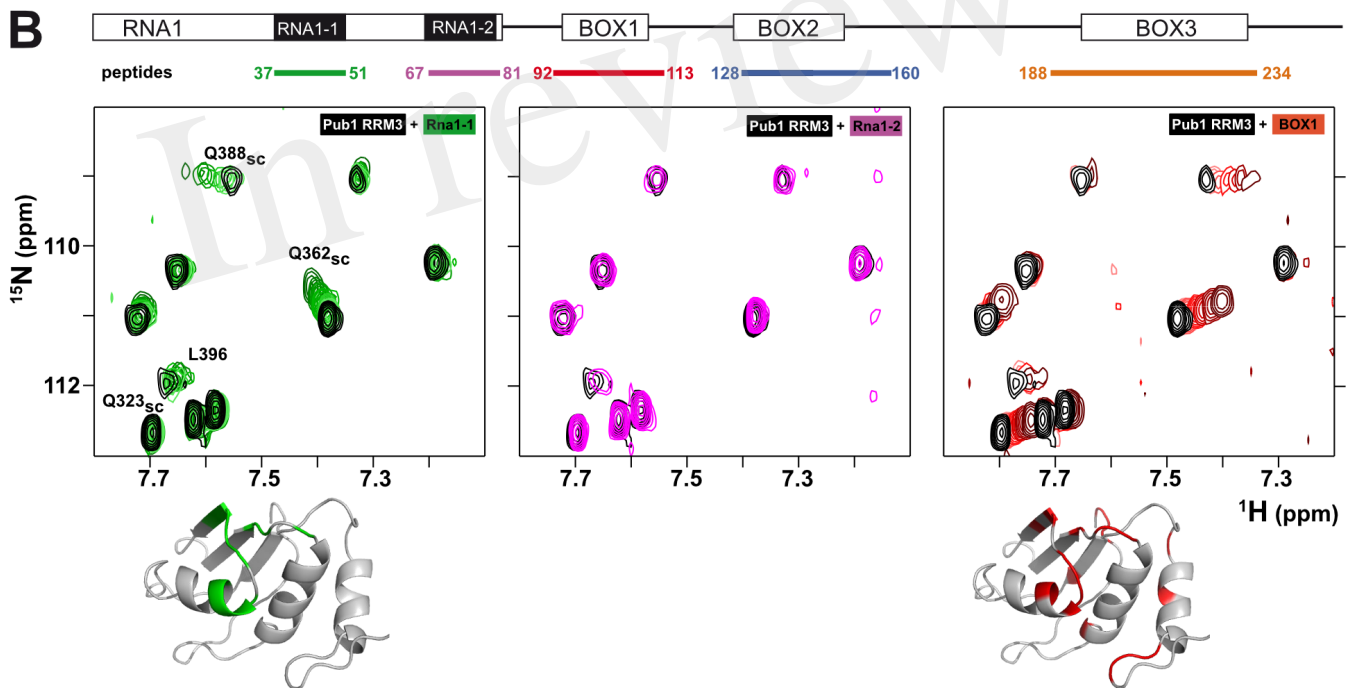
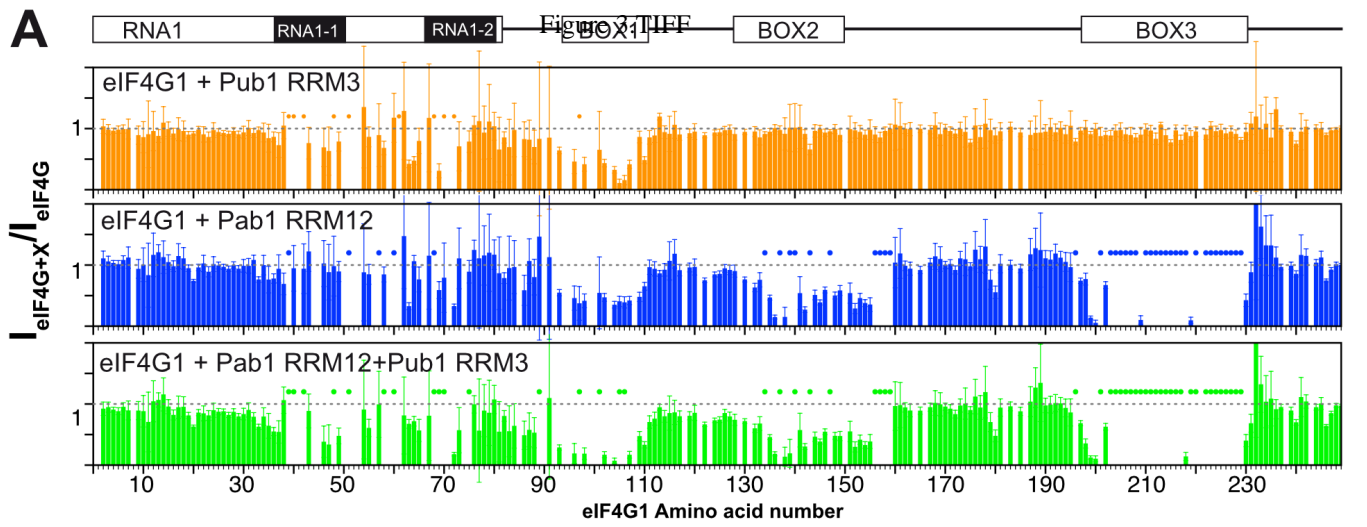
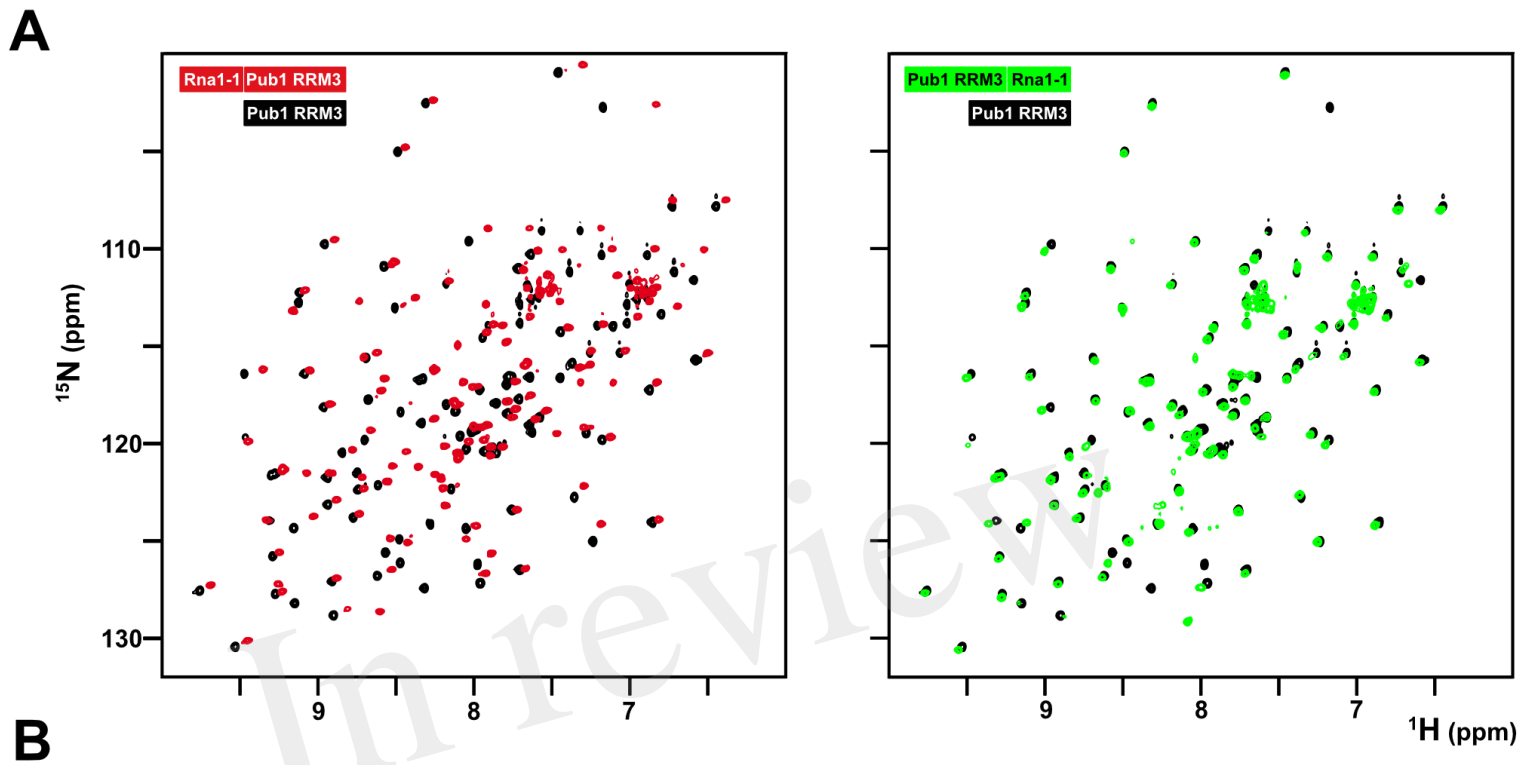


Figure 4.TIFF



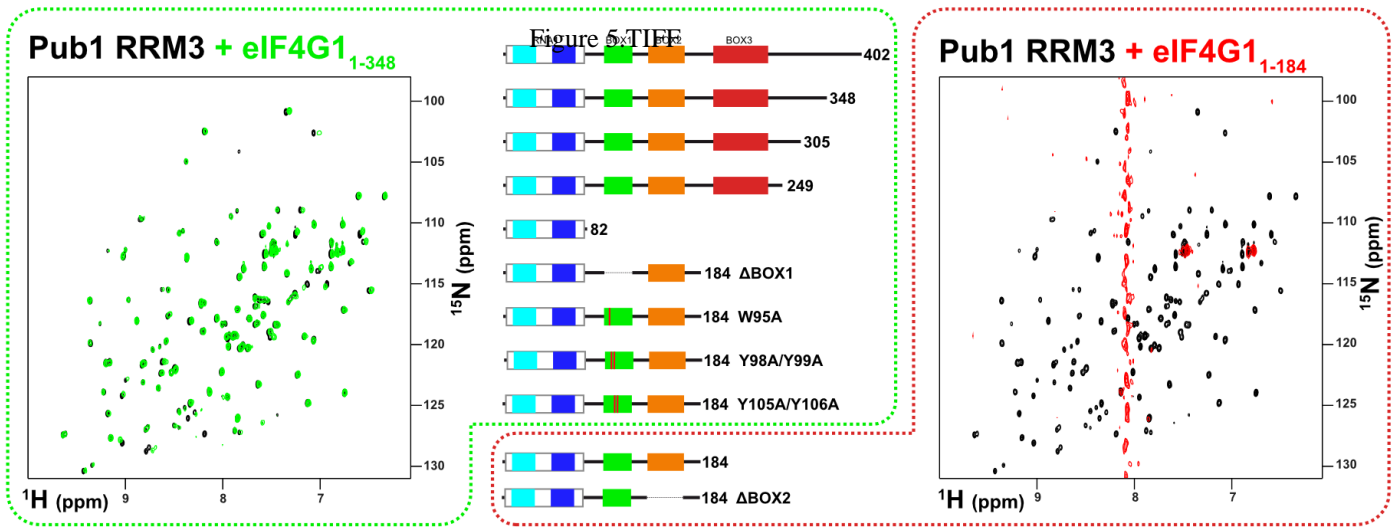
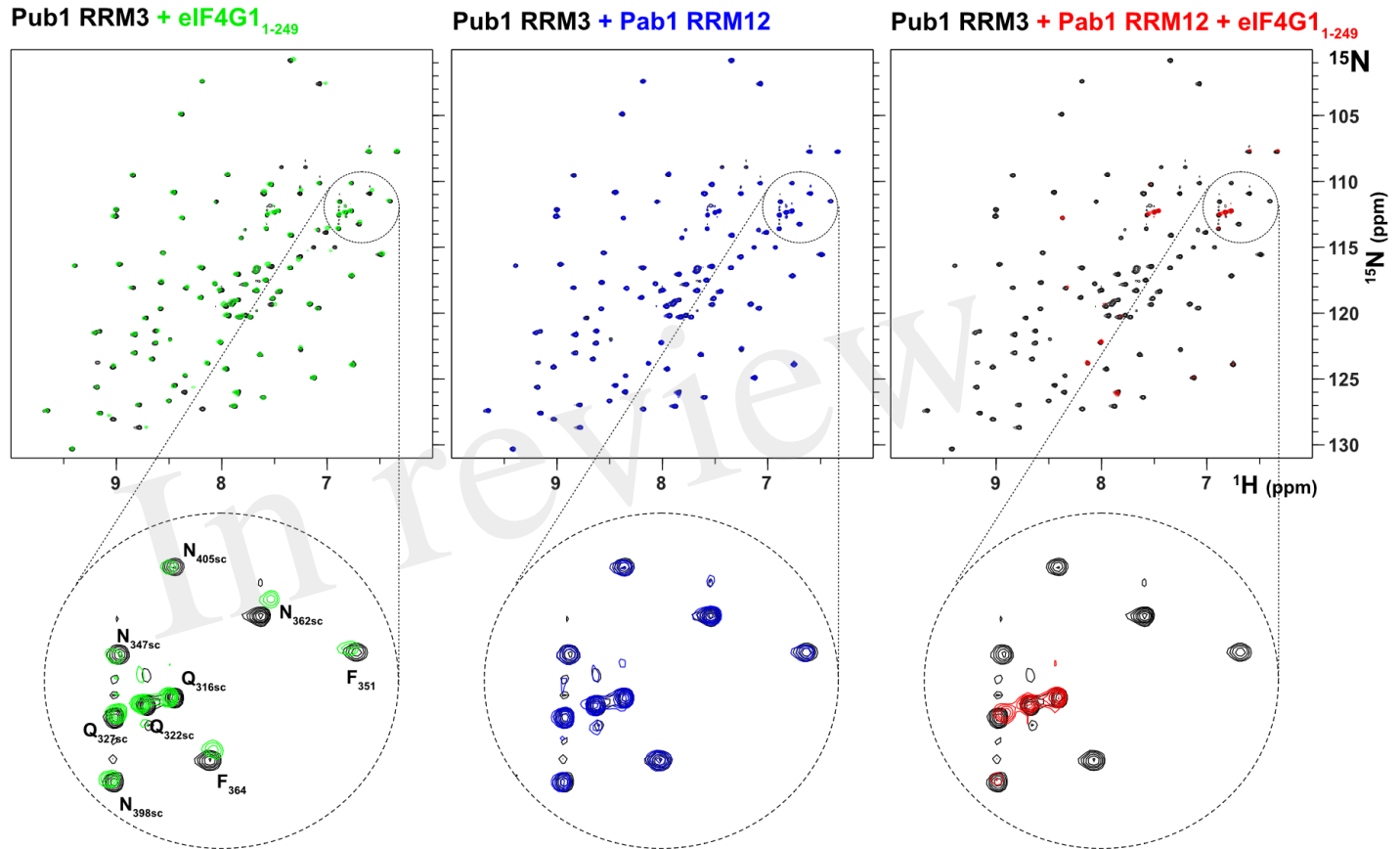
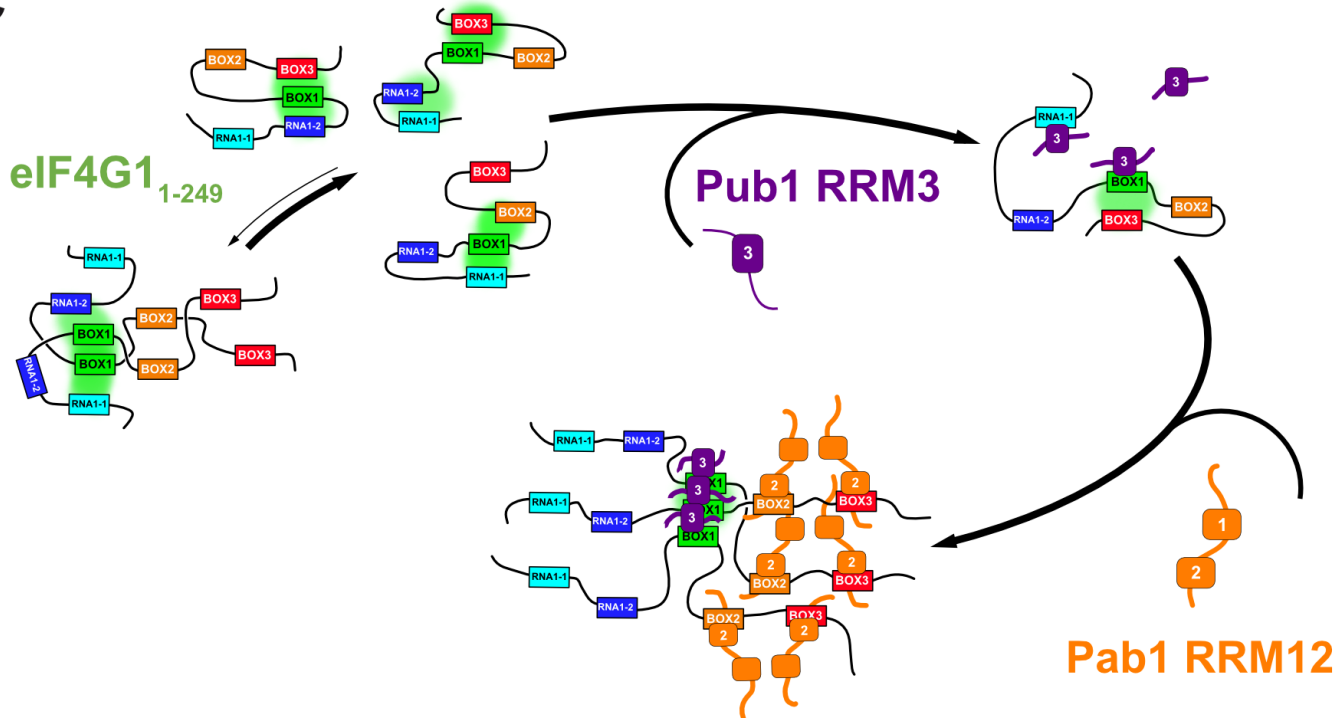
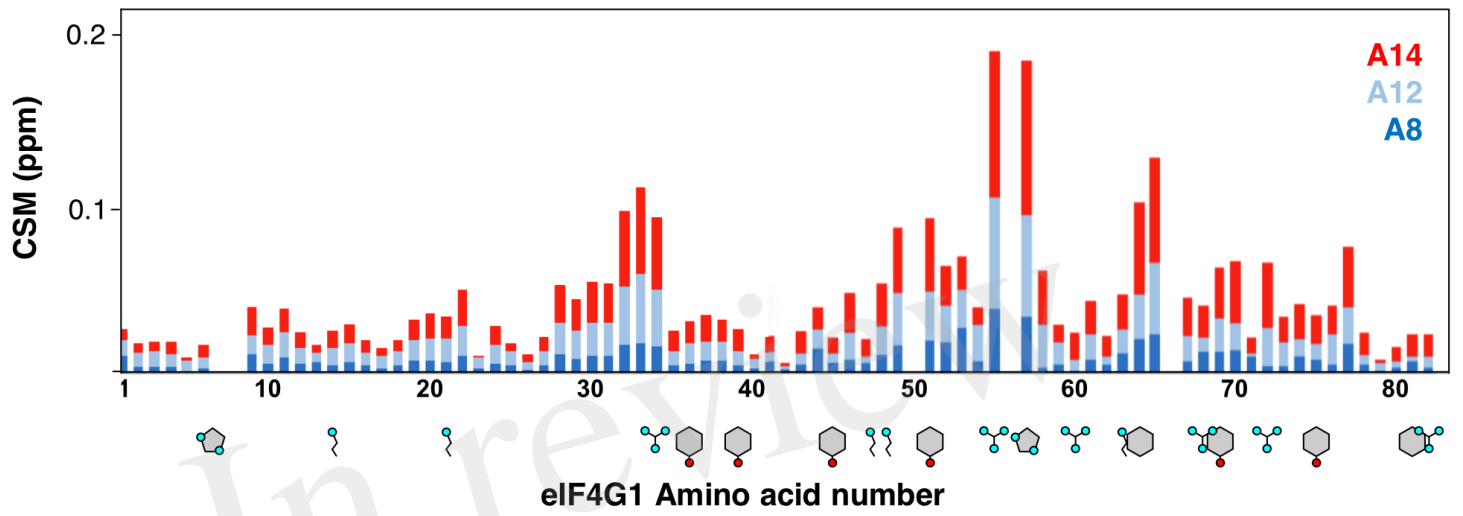
A**B****C**

Figure 6.TIFF

A



B

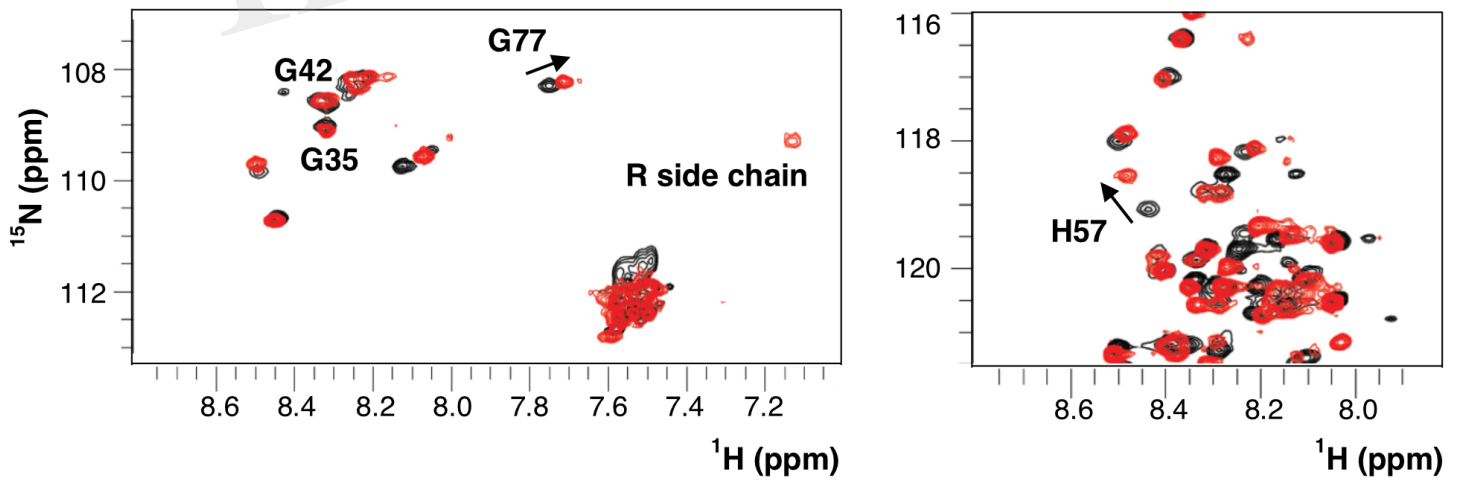
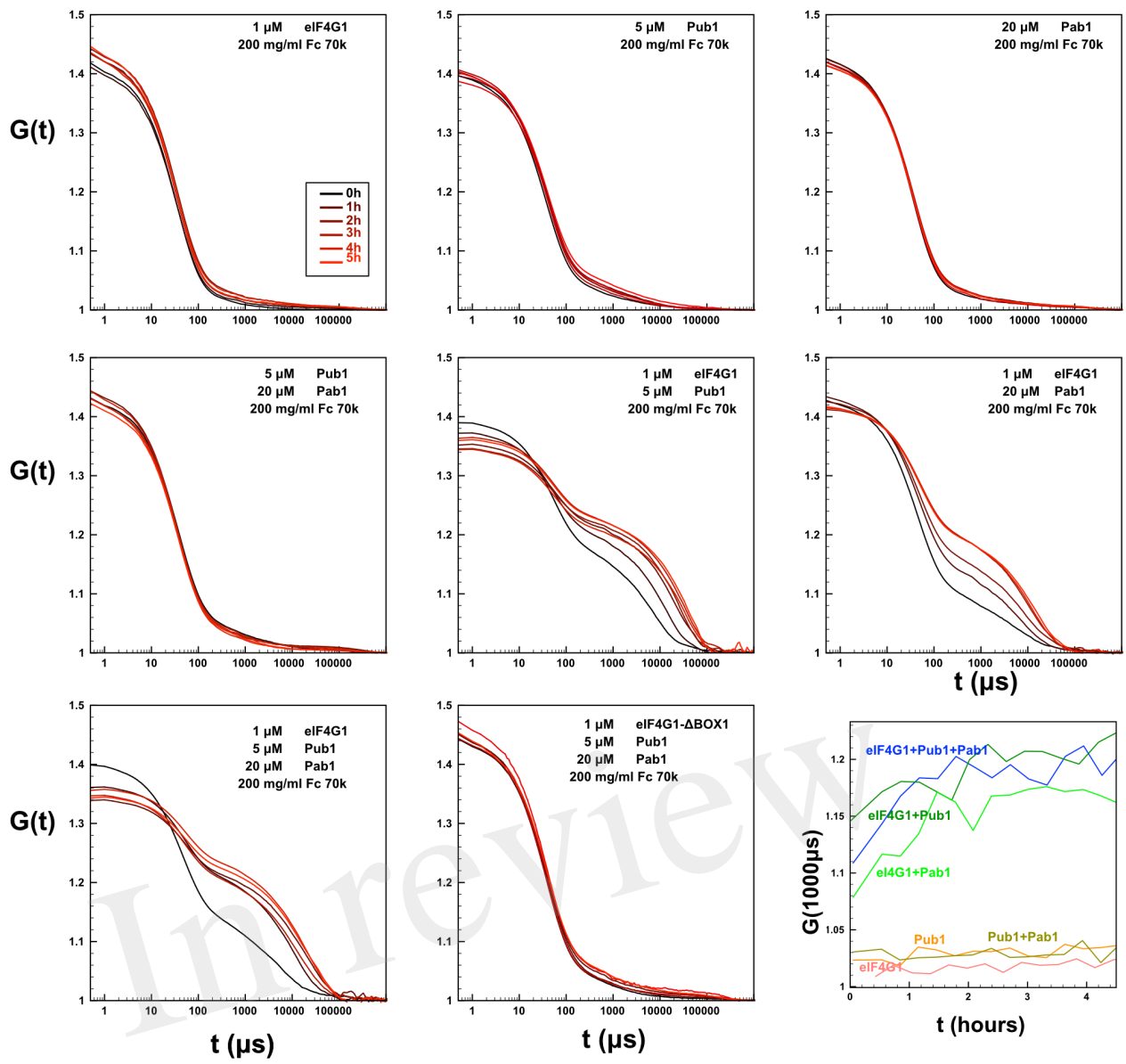
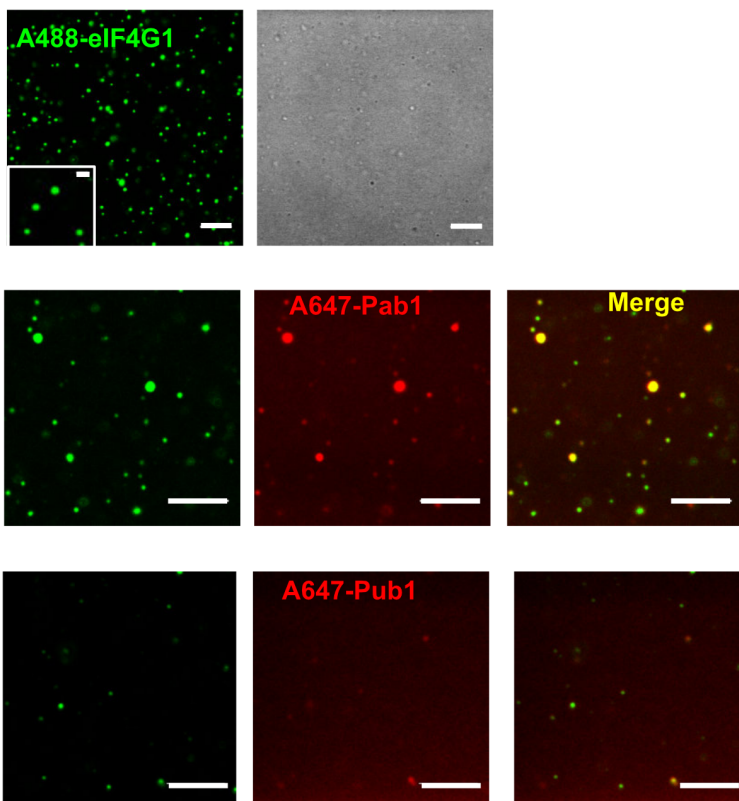


Figure 7.TIFF

A



B



C

