RESEARCH ARTICLE

Structural Dynamics of the GW182 Silencing Domain Including its RNA Recognition motif (RRM) Revealed by Hydrogen-Deuterium Exchange Mass Spectrometry

Maja K. Cieplak-Rotowska,¹ Krzysztof Tarnowski,² Marcin Rubin,¹ Marc R. Fabian,^{3,4} Nahum Sonenberg,^{5,6} Michał Dadlez,² Anna Niedzwiecka⁷

¹Division of Biophysics, Institute of Experimental Physics, Faculty of Physics, University of Warsaw, 02-089, Warsaw, Poland

²Laboratory of Mass Spectrometry, Institute of Biochemistry and Biophysics, Polish Academy of Sciences, PL-02106, Warsaw, Poland

³Lady Davis Institute for Medical Research, Jewish General Hospital, Montréal, Québec, Canada

⁴Department of Oncology, McGill University, Montréal, Québec, Canada

⁵Department of Biochemistry, McGill University, Montréal, Québec, Canada

⁶Goodman Cancer Center, McGill University, Montréal, Québec, Canada

⁷Laboratory of Biological Physics, Institute of Physics, Polish Academy of Sciences, Aleja Lotników 32/46, PL-02668, Warsaw, Poland

Abstract. The human GW182 protein plays an essential role in micro(mi)RNA-dependent gene silencing. miRNA silencing is mediated, in part, by a GW182 C-terminal region called the silencing domain, which interacts with the poly(A) binding protein and the CCR4-NOT deadenylase complex to repress protein synthesis. Structural studies of this GW182 fragment are challenging due to its predicted intrinsically disordered character, except for its RRM domain. However, detailed insights into the properties of proteins containing disordered regions can be provided by hydrogen–deuterium exchange mass spectrometry (HDX/MS). In this work, we applied HDX/MS to define the structural state of the GW182 silencing domain. HDX/MS analysis revealed that this domain is clearly divided into a natively unstructured

part, including the CCR4-NOT interacting motif 1, and a distinct RRM domain. The GW182 RRM has a very dynamic structure, since water molecules can penetrate the whole domain in 2 h. The finding of this high structural dynamics sheds new light on the RRM structure. Though this domain is one of the most frequently occurring canonical protein domains in eukaryotes, these results are – to our knowledge – the first HDX/MS characteristics of an RRM. The HDX/MS studies show also that the $\alpha 2$ helix of the RRM can display EX1 behavior after a freezing–thawing cycle. This means that the RRM structure is sensitive to environmental conditions and can change its conformation, which suggests that the state of the RRM containing proteins should be checked by HDX/MS in regard of the conformational uniformity.

Keywords: Hydrogen–deuterium exchange, Mass spectrometry, Structural dynamics, Intrinsically disordered proteins, GW182, RNA recognition motif (RRM)

Received: 5 July 2017/Revised: 8 September 2017/Accepted: 1 October 2017/Published Online: 27 October 2017

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13361-017-1830-9>) contains supplementary material, which is available to authorized users.

Correspondence to: Anna Niedzwiecka; e-mail: annan@ifpan.edu.pl

Introduction

In recent years, there have appeared a significant number of works proving that structural dynamics governs the function of biological molecules to no less extent than the protein sequence and its 3D structure [1–3]. Many biologically important issues related to conformational changes were addressed

by an approach that combined hydrogen–deuterium exchange with mass spectrometry [4–7].

The RRM domain is thought to be one of the most abundant globular protein domains in eukaryotes [8]. To our surprise, among over 115,000 protein structures deposited in the RCSB PDB, only about 340 structures of RRM domains can be found. Among them, more than 200 were resolved by NMR and about 140 by X-ray crystallography, and almost 50% of the RRM structures became known only during the last 5 y. This brief statistic reflects the dynamic structural character of this protein domain that can lead to difficulties in crystallization of these proteins. A canonical RNA recognition motif (RRM) is defined by two conserved ribonucleoprotein motifs (RNPs) within an ~80 amino acid long fold consisting of a four-stranded antiparallel β -sheet supported by two α -helices, arranged in the order of $\beta 1\alpha 1\beta 2\beta 3\alpha 2\beta 4$ [8–10] (Figure 1a). Usually, the $\beta 3$ (RNP1) and $\beta 1$ (RNP2) strands contain conserved aromatic residues that make direct contact to the RNA chain. Interesting exceptions among RRM domains are related to the mRNA 5' cap interacting proteins, i.e., the nuclear cap-binding complex (CBC) [13] and poly(A)-specific ribonuclease (PARN) [14]. The CBC RRM called CBP20 binds the 5' terminus of the polymerase II transcripts chain via Tyr43 residue located at $\beta 1$ in the opposite direction than other RNA-interacting RRM domains [15–17], whereas the PARN RRM binds to the mRNA 5' cap via Trp475 located entirely at its outer surface [18–20].

The GW182 protein is one of the crucial players in miRNA-mediated gene silencing, since it specifically interacts with the Argonaute (Ago) proteins [21–23] and recruits the multi-subunit CCR4-NOT deadenylase complex [11, 24, 25] to the targeted mRNA to trigger both translational repression at the 5' terminus and removal of the mRNA 3' poly(A) tail [26–28]. The first human GW182 protein discovered, now known also as TNRC6A (trinucleotide repeat-containing gene 6 protein A), has many glycine-tryptophan (GW) repeats and a predicted mass of 182 kDa [29]. It was discovered in human cells as the antigen recognized by an autoimmune serum from a patient suffering from motor and sensory polyneuropathy [30]. There are usually three paralogs of GW182 in vertebrates: TNRC6A, TNRC6B, and TNRC6C. Invertebrates usually possess only one protein, e.g., *Drosophila melanogaster* GW182 (also known as Gawky), but three paralogs were found in mosquitos [30, 31]. In a recent phylogenetic study [31], it was found that GW182 proteins are animal specific and a single copy of the gene arose with the emergence of multicellularity. Then three copies of the gene appeared in the last common ancestor of vertebrates as a result of two rounds of whole genome duplication events [31]. TNRC6C is thought to be the founding member of the vertebrate family of GW182 proteins and is the ortholog of the invertebrate GW182 [31]. The three human paralogs are considered to be redundant [32] on the basis that these proteins associate with all four human Ago proteins and a common set of miRNA targets [23, 26, 33], and silencing is efficiently suppressed when at least two of the proteins are co-depleted [34].

GW182s are predicted to be mostly disordered except for two globular domains: the ubiquitin associated domain (UBA) and the RRM (Figure 1b) [30]. The structure of the *D. melanogaster* GW182 RRM domain (residues 1510–1595 in TNRC6C) was solved by NMR [12] and revealed the presence of an additional $\alpha 3$ -helix on the side of the β -sheet that is solvent-exposed in canonical RRM domains (Figure 1a, c) [8]. The $\alpha 3$ -helix is tightly bound to the β -sheet, making it inaccessible to the solvent and ligands. In consequence, the *Dm* GW182 RRM does not bind RNA in vitro [12]. The GW182 RRM domains have a high sequence homology (Figure 1a), which suggests that the human RRM domains are also incapable of RNA binding. It was proposed that the GW182 RRM might be involved instead in protein–protein interactions using the opposite side of the domain, where $\alpha 1$ and $\alpha 2$ helices together with $\beta 3'$ and $\beta 4$ strands form a broader hydrophobic cleft [12], which could potentially be utilized for such interactions, thus contributing to gene silencing. Contrary to these expectations, no evidence of such direct interactions was found, even though the presence of the RRM domain in the GW182 sequence was shown to be required for efficient gene silencing [12, 32]. This observation corresponds better to another putative role that could be fulfilled by RRM domains, i.e., stabilization of a protein structure that enables assembling of a required functional bipartite motif composed of unstructured regions flanking the folded RRM, similarly as, e.g., in case of a bimodular non-classical nuclear localization signal of ADAR1 enzyme [35]. This role is particularly plausible for the RRM belonging to GW182, since four of the remaining regions of the protein, i.e., a glutamine-rich (Q-rich) domain and three GW-rich regions: an N-terminal Argonaute proteins-binding domain, a middle domain (Mid) between Q-rich and RRM, and a C-terminal part [29, 30], are supposed to be disordered (Figure 1b). The N-terminal GW repeats mediate the interaction with Ago through tryptophans [22, 28, 36]. The Q-rich region (residues 1074–1207 in TNRC6C) was implicated to play a role in P-body localization [28, 30]. The Mid region has a highly conserved PAM2 motif that interacts specifically with the cytoplasmic poly(A)-binding protein (PABPC1) [37, 38].

miRNA-dependent gene silencing by the GW182 protein is mediated by its C-terminal half called the silencing domain (SD), which – when artificially tethered – can trigger silencing of reporters bypassing the involvement of miRNAs and Argonautes [26, 33, 39]. The silencing activity of the GW182 SD stems from its ability to bind with deadenylase complexes: CCR4-NOT and Pan2-Pan3, and also with PABP [11, 24, 25, 37, 38]. The interaction of the human GW182 with the CCR4-NOT components is thought to be mediated by the Mid and C-terminal unstructured regions of the SD, among them a region encompassing residues 1286–1306 in TNRC6C, termed the CCR4-NOT interacting motif 1 (CIM1, Figure 1b) [11, 24]. Binding of the Pan2–Pan3 complex requires the interaction of the N- and C-term regions surrounding the RRM [24, 25]. Other GW182 properties were reviewed in more detail by Braun et al. [32].

Figure 1. **(a)** Sequence alignments of the RRM domains of different GW182 isoforms from human, mouse, and *Drosophila melanogaster*. Secondary structural elements are marked as color blocks corresponding to the α -helices and β -sheets shown in **(c)**; **(b)** schematic representation of human GW182 (TNRC6C): an Ago-binding domain, a ubiquitin associated domain (UBA), a glutamine-rich (Q-rich) region, and the silencing domain (SD). SD contains a CCR4-NOT interacting motif (CIM1) [11], a PABP-interacting motif (PAM2), and RRM; **(c)** predicted human GW182 RRM structure shown as a ribbon colored according to the secondary structural elements shown in Figure **(a)**, aligned with the NMR *D. melanogaster* RRM structure (grey) [12]; **(d)** disorder tendency of the human GW182 silencing domain (SD; residues 1260-1620); CIM1 is marked in mint, PAM2 in dark gold, RRM in light grey

Structural studies of the GW182 silencing domain is challenging because of its intrinsically disordered character. There is a continuum of protein structures ranging in the degree of disorder from close to none in globular proteins through proteins with ordered domains linked via disordered linkers, molten globules to completely unstructured proteins (intrinsically disordered proteins, IDPs; reviewed in [40]). This is linked to a broad range of observed levels of protein structural dynamics. It was shown that the frequency of breaking of main chain H-bonds, which reflects the frequency of deviations of the chain position from the average conformation, may differ by 5–7 orders of magnitude [41]. What we call “protein structure” is just a small window in this spectrum, easy to visualize because of fixed atom positions, whereas the rest of the spectrum of protein states is underappreciated, being more difficult to represent. This phenomenon has been called “a dark-matter in biology” [42].

IDPs have been found to be widespread [43], and involved in such diverse processes as transduction of signals, transport

through nuclear pores, regulation of transcription, translation, and the cell cycle, as well as binding of RNA, DNA, and other small ligands [44]. Many, but not all, IDPs gain structure upon binding a partner [45]. IDPs interact with their binding partners differently than globular proteins do. They are able to interact with different target proteins using the same region, which results in adapting different conformations [46]. These proteins use hydrophobic or aromatic residues to mediate interactions more frequently than ordered proteins do [45, 47]. Such protein structures are neither suitable for X-ray crystallography nor for NMR due to their dynamic character or limited solubility. Moreover, they are still challenging even for a computational approach [48].

However, quite detailed insight into the properties of proteins containing disordered regions can be provided by hydrogen–deuterium exchange mass spectrometry (HDX MS) [5]. This technique allows mapping of the stability of H-bonding networks, since the kinetics of the H/D exchange reflects the local dynamics of the protein structural elements [49–51]. In

this work, we reveal the structural dynamics of the human GW182 silencing domain. We provide the experimental evidence that the GW182 SD is completely unstructured except the RRM, which has also a very dynamic structure. These results are – to our knowledge – the first hydrogen–deuterium exchange mass spectrometry ‘fingerprint’ of an RRM domain.

Materials and Methods

Protein Expression and Purification

Experiments were conducted on the GW182 fragment denoted as the silencing domain (SD, formerly SD10 in [11]), encompassing residues 1260–1620 of the TNRC6C paralog (isoform 1, UniProt Q9HCJ0), with a predicted mass of 39819.9 Da, and confirmed mass of 39822 ± 5 Da. SD was cloned into pET28b(+) His-Sumo vector using *EcoRI* and *XhoI* restriction sites and purified on a His-trap column (GE Healthcare) on AKTA FPLC, followed by in-solution cleavage by CoolCutter (Genecopoeia), and purified again on a His-trap column. Before the experiments, the protein was dialyzed in 50 mM Tris, 150 mM NaCl, 2 mM DTT, 1 mM EDTA, pH 7.0, and concentrated using Amicon Ultra Centrifugal Filters (Merck Millipore) at 4 °C. In most experiments, freshly prepared protein was used. The frozen fraction of the protein, for which the double conformation was found, was stored at –80 °C with 10% glycerol.

Human full length eIF4E was expressed and purified as described previously for the murine homolog [52].

Hydrogen–Deuterium Exchange Experiments

Each set of experiments was performed four times, essentially as described previously [52, 53]. Each exchange and control measurement in the set was repeated at least three times. The protein was used at initial concentrations of 60 or 200 μ M (nonfrozen samples), and 15 or 20 μ M (frozen-thawed samples). A typical set of experiments consisted of mass spectra measurements (LC-MS runs) collected after several time intervals (10 s, 1 min, 20 min, 2 h, and in some cases 24 h) of H/D exchange, triggered by 1:10 dilution in heavy water (99.8%, Cambridge Isotope Laboratories, Inc.)-based buffer (50 mM Tris, 150 mM NaCl, 2 mM DTT, 1 mM EDTA, pH 7.0). The deuterium level in the heavy water-based buffer system after adding the chemicals was 99.27%. pD values were measured on pH-meter (pH 1000 L, pHenomenal) and uncorrected. HDX reaction in a volume of 50 μ L sample was quenched by adding 10 μ L of stop buffer (2 M glycine, 4 M guanidine hydrochloride, 150 mM NaCl, pD 2.5 in 99.8% D₂O). Two control measurements were performed [52, 53]. The first control experiment was related to the residual exchange that might still occur after quenching. This process is called “in-exchange”. The control was obtained by adding stop buffer first, then diluting the sample in the heavy water-based buffer, in order to correct for a possible imperfect arrest of the H/D exchange reaction at pD 2.5. Then the residual exchange level was

measured identically to the main experiments (see below). The experiments in the second set of control, measuring the maximum level of exchange possible to be detected due to a back D/H exchange during LC-MS, were conducted by in-solution pepsin digestion of protein samples in a mildly denaturing condition, followed by HDX and final MS measurements. For this purpose, the immobilized porcine pepsin (Pierce) was prepared by mixing 50 μ L slurry (resin plus supernatant) per one back-exchange sample with a 4-fold excess of wash buffer (200 mM glycine, 0.66 M guanidinium hydrochloride, 150 mM NaCl, pD 2.5 (uncorrected), all in D₂O), centrifuging at 1000 rcf, removing supernatant, and repeating wash three times. Finally, the resin was mixed with wash buffer in a 1:1 ratio and 50 μ L of slurry was used per sample.

Five μ L of protein was diluted into 45 μ L of deuterated reaction buffer, and the reaction was acidified by adding 10 μ L deuterated stop buffer. The solution was incubated with the immobilized pepsin for 4 min at 4 °C with shaking. The sample was centrifuged at maximal speed in a table-top microcentrifuge, and 60 μ L of solution was transferred to a fresh Eppendorf tube. The pD value of the sample was raised to ~ 7 with NaOD. After 2 min of incubation at RT, pD was lowered to ~ 2.5 by adding DCl. The sample was centrifuged once again at maximal speed to get rid of any beads, and 50 μ L was injected into the nanoACQUITY UPLC system (Waters) and analysis continued as for normal exchange samples (see below) except that the pepsin column was taken out.

This enables us to correct directly for the back-exchange of the individual peptides that occurred during the chromatographic separation step performed in the H₂O milieu. The precise correction for the back-exchange also allowed taking into account the effect of the incomplete exchange in the 90% D₂O buffer after dilution.

Mass Spectrometry Measurements

Immediately after HDX quench, samples were passed through an immobilized pepsin column (Poroszyme, Applied Biosystems), placed in HDX Manager system, and kept at 13 °C, with 0.07% formic acid (200 μ L/min flow rate) as a mobile phase. Pepsin-generated peptides were then trapped in a C18 trapping column (Acquity BEH C18 VanGuard Pre-Column, Waters), and then separated using liquid chromatography on a reverse-phase column (Acquity UPLC BEH C18 column, Waters) with a 8% to 40% gradient of acetonitrile in 0.1% formic acid at a 40 μ L/min flow rate using the nanoAcquity Binary Solvent Manager (Waters). The total run time was 13.5 min. The apparatus (excluding the pepsin column) was kept at 0.5 °C. The carryover problem was excluded since the same columns and elution conditions were applied for both the nonfrozen and frozen-thawed samples, and the latter protein was much less concentrated.

Peptides were then passed for MS analysis into the ion source of SYNAPT G2 HDMS (Waters) (Q-ToF) working in Ion Mobility Mode. The mass spectrometer parameters were

set as follows: ESI positive mode, capillary voltage 3 kV, sampling cone 35 V, extraction cone voltage 3 V, source temperature 80 °C, desolvation temperature 175 °C, and desolvation gas flow 800 l/h. Mass spectra were acquired over the m/z range of 350 to 1500 in the IMS mode. Ion-mobility separation conditions in TriWave device were as follows: helium cell flow rate 180 mL/min, ion-mobility nitrogen flow rate 90 mL/min, IMS wave velocity 600 m/s, and IMS wave height 40.0 V.

Numerical Data Analysis

The peptic peptides were identified in nondeuterated buffer based on mass spectra acquired in MSE mode, with the low-energy function voltage of -4 eV and high-energy function voltage of -28 eV, over the m/z range of 100 to 1950, in either standard or IMS mode. The isotopic envelopes after exchange were assigned to their corresponding peptide sequence tags using DynamX 2.0 or 3.0 (Waters), based on a peptide list obtained from the triple sequencing analysis of unexchanged sample by ProteinLynx Global Server software (Waters) and a randomized database. The following acceptance criteria were used: minimum intensity threshold of 1000, minimum products per amino acid of 0.25, max MH⁺ error of 5 ppm. Identification of isotopic envelopes was conducted with retention time tolerance ± 0.3 min, mass tolerance ± 15 ppm, and drift time deviation ± 2 time bins. The acquired data were verified by visual inspection and exported to Excel, R, or Prism 6 (GraphPad) for further analyses. Bimodal isotopic envelopes were exported directly from MassLynx (Waters) and analyzed by HX-Express2 [54].

The fraction of peptide amide hydrogens exchanged to deuterium, D_f , for each peptide was calculated as:

$$D_f = \frac{M_t - \overline{M}_{in}}{\overline{M}_{bx} - \overline{M}_{in}} \quad (1)$$

where M_t – the average mass of the peptide after a given exchange time, t ; $\overline{M}_{in}$ – the average mass of the peptide in the in-exchange control samples; $\overline{M}_{bx}$ – the average mass of the peptide in the back-exchange control samples.

The time dependence of the fraction exchanged for each pepsin generated peptide was analyzed by a nonlinear, least-squares regression assuming for simplicity the existence of a single or dual population of amide protons of a given peptide, that exchange with the kinetics described by either a mono-exponential or bi-exponential equation, respectively:

$$D_f(t) = D_{inc}(1 - e^{-kt}) \quad (2)$$

or

$$D_f(t) = D_{incf}(1 - e^{-k_f t}) + D_{incs}(1 - e^{-k_s t}) \quad (3)$$

where D_{inc} was a fitting parameter describing the fraction of amide protons (%) that exchange with the rate constant

k (min⁻¹), and D_{incf} , D_{incs} were the fitting parameters for two populations exchanging with the corresponding rate constants, k_f and k_s , respectively. Discrimination between the mono- and bi-exponential models as well as between bi-exponential models with one of the k values fitted or fixed [52, 55, 56] was based on an Akaike's Information Criterion [57] and also on a two-parameter statistical Snedecor's F -test [58] for the models with different numbers of degrees of freedom.

The intrinsic rate constant values, k_{int} , for the investigated peptic peptides were calculated according to the following numerical procedure: the peptide amide hydrogen exchange rates for the individual amino acids pairs were predicted by the SPHERE program [59–61]. These values were used to calculate the expected number of deuteria incorporated by each peptic peptide based on its sequence, and further to determine the exchanged fraction at nine different time points ranging from 0.001 to 60 s. The value of k_{int} for each peptide was then obtained as a fitting parameter according to Equation 2.

Conformational Equilibrium

The Gibbs free energy of activation, ΔG^{++} , defining the height of the barrier for the transition from the closed to open RRM conformation upon the EX1 [62], was estimated from the Eyring-Polanyi transition state theory [63] as:

$$\Delta G^{++} = -R \cdot T \cdot \ln \frac{k^{++} \cdot h}{k_B \cdot T} \quad (4)$$

where R - universal gas constant, T - absolute temperature, h - Planck constant, k_B - Boltzmann constant, and k^{++} was the kinetic rate constant (s⁻¹). The molecules undergo two concurrent processes: the reversible conformational transition and the irreversible HDX according to the exchange-accessibility provided by a particular conformation. The experimentally measured intensities of MS signals for the peptides in the closed conformation that exchange incrementally with the kinetics of the close state come only from this fraction of the closed state population that never once crossed over the conformational barrier to the open state before. Analogously, the intensities of MS data at the characteristic m/z range for the open conformation come not only from the actual open state population but also from these closed state molecules that underwent to the open state and could exchange immediately with the open state kinetics. Thus, the observed apparent decrease in the closed conformation population determined by HX-Express2 after subsequent time intervals of the H/D exchange, $P_{cl}(t)$, is related to the conformational change kinetic rate constant, k^{++} , according to the equation:

$$P_{cl}(t) = P_{cl}(0) \cdot \exp(-k^{++} \cdot t) \quad (5)$$

The difference in the Gibbs free energies of the closed and open conformations, ΔG , was calculated from the initial

population, $P_{cl}(0)$, according to the equation:

$$\Delta G = -R \cdot T \cdot \ln \left(\frac{P_{cl}(0)}{1 - P_{cl}(0)} \right) \quad (6)$$

The Gibbs free energy attributed to the local protein stability (minimum on the local energy landscape) upon the EX2 conditions [62] was calculated as:

$$\Delta G^{local} = -R \cdot T \cdot \ln(p) \quad (7)$$

where p was the protection factor obtained from the intrinsic rate constant, k_{int} , and the measured rate constant, k , corresponding to a given amide hydrogen population, according to the equation:

$$p = k_{int}/k \quad (8)$$

Fluorescence Spectroscopy

Fluorescence spectra were measured on Fluorolog 3.11 (Horiba) for tryptophan (NATA) at 22 μ M in water, 1:1 water:methanol, methanol, propan-2-ol, while for SD at 2 μ M and eIF4E at 2.75 μ M in buffer, in a thermostatted micro-cuvette (Hellma) at 15 $^{\circ}$ C, with excitation wavelength of 280 nm. Concentrations of proteins and NATA were chosen to have equimolar amounts of tryptophan residues in each sample. Energy at maximum of the fluorescence spectra was calculated as:

$$E = h \cdot c / \lambda_{max}, \quad (9)$$

where h was the Planck constant, c was the speed of light in vacuum, λ_{max} was the wavelength of the maximum.

During the tryptophan-based thermofluorescence assay, the fluorescence emission signals at 350 and 330 nm were measured every 10 s with the integration time of 4.6 s, at the excitation wavelength of 280 nm. The temperature gradient was 1 $^{\circ}$ C/min. The actual temperature was measured inside the cuvette with the thermocouple and recorded by the thermostat. The assay for SD was repeated five times at different SD concentrations, from 1 to $\sim$ 13 μ M. The assay for eIF4E was performed at 2.75 μ M. The melting temperature was determined from a non-linear fitting of the Boltzmann sigmoidal curve:

$$F_{350}/F_{330}(T) = FF_i + (FF_f - FF_i) / (1 + \exp((T_m - T)/slope)), \quad (10)$$

where T_m was the melting temperature in $^{\circ}$ C, $slope$ was a cooperativity coefficient, FF_i and FF_f were the initial and final plateau, respectively.

Bioinformatics

The human GW182 RRM structure prediction was done by I-TASSER [64] with default settings. The best model had a C-score of -1.38 and expected TM-score 0.54 ± 0.15 . The RRM figures and contact map were prepared by Discovery Studio v3.5.0 (Accelrys Software Inc.) based on the structures aligned by Swiss-PdbViewer v4.0.4 [65]. Disorder tendency was predicted by the MetaDisorderMD2 meta-server [66] utilizing a variety of algorithms [67–82] to get a well-balanced average prediction. Peptide coverage was drawn using Draw Map [83]. Sequence alignment was performed in Align, Uniprot Clustal Omega [84]. The GW182 flexibility and normal modes were checked by CABS-Flex [85] and WEBnm@ [86] servers.

Circular Dichroism Spectroscopy

Far-UV CD spectra were recorded with the MOS-450/AF-CD spectrometer from Bio-Logic, with a Xe lamp as a source of light, in the wavelength range of 190–260 nm, in a thermostatted 0.1 mm quartz cuvette at 20 $^{\circ}$ C. Samples of both the fresh and frozen-thawed protein at 24.4 μ M in 50 mM Tris, 150 mM NaCl, 2 mM DTT, 1 mM EDTA, pH 7.0, without and with 10% glycerol, respectively, were prepared three times. The measurements for each protein and buffer sample were repeated twice, with an acquisition time of 10 s at each nanometer. The scans were averaged and the reference CD spectra of respective buffers were subtracted from the protein spectra.

Results and Discussion

Structural Dynamics Properties of the GW182 Silencing Domain

Homology modeling of the human GW182 RRM structure (Figure 1c) shows that this domain is similar to the *D. melanogaster* GW182 RRM except for a slight distortion of the shortest $\beta 3'$ -strand. The secondary structural elements corresponding to those identified in the *Dm* RRM are formed by the amino acids in the following ranges of the human protein sequence: $\beta 1$: 1515–1519; $\alpha 1$: 1528–1537; $\beta 2$: 1540–1546; $\beta 3$: 1551–1556; $\alpha 2$: 1559–1569; $\beta 3'$: 1572–1574; $\beta 4$: 1578–2583; and $\alpha 3$: 1587–1594 (Figure 1a, c). Detailed predictions for the whole GW182 silencing domain indicate a mostly disordered character, with the exception of the RRM domain and a few short regions that exhibited a lower disorder tendency (Figure 1d). The PAM2 motif is among them, but not CIM1. However, neither the predicted presence of the well-defined human GW182 SD RRM structure nor the unfolded character of the remaining part of the structure have yet been confirmed experimentally.

Some experimental indications whether GW182 silencing domain is intrinsically disordered can be inferred from UV spectroscopy, i.e. environment sensitive fluorescence measurements or circular dichroism. The SD sequence contains 11 tryptophans, 10 of which are outside the RRM domain, and one of them is localized at the border of the structured part, as

Figure 2. (a) Fluorescence spectrum of SD at 2 μM in buffer (thick black line) together with spectra of tryptophan (thin lines) (NATA) at 22 μM in water (red), 1:1 water:methanol (orange), methanol (green), propan-2-ol (blue), and eIF4E at 2.75 μM (thick broken line). Inset: dependence of the energy of the fluorescence spectra maximum on the solvent dielectric constant for tryptophan ($\circ$), SD ($\bullet$), and eIF4E ($*$); (b) thermofluorescence assay for SD at 1 μM (black), NATA at 11 μM (purple), and eIF4E at 2.75 μM (green) with the Boltzmann sigmoidal melting curve (black broken line)

predicted by I-TASSER (Figure 1a). The fluorescence spectrum of SD (Figure 2a) was compared with tryptophan in solvents with decreasing dielectric constants: water, 1:1 water:methanol, methanol, and isopropanol, as well as with eIF4E, a globular protein containing eight tryptophans, where five of them are buried and three of them are water accessible in the *apo* state. The λ_{max} of the SD emission spectrum, 352 nm, resembles that of tryptophan in the water:methanol mixture, contrary to eIF4E with the λ_{max} of 338 nm. The results can be used to semiquantitatively estimate the average dielectric constant experienced by the tryptophan indole rings of the two proteins, which proved to be ~ 52 for SD and ~ 13 for eIF4E. The comparison of the SD thermofluorescence assay with NATA as a standard for a completely water accessible tryptophan and eIF4E as a standard for a globular protein without

disulfide bridges shows that for SD, contrary to eIF4E ($T_m = 42.43 \pm 0.08$ $^{\circ}\text{C}$), no unfolding transition could be detected (Figure 2b). Thus, the GW182 SD tryptophans are, by average, much more water-exposed than in eIF4E, but not totally, since some water protection is visible due to the presence of the SD protein chain.

In order to address the issue of structural dynamics of the GW182 silencing domain in a more precise manner, we performed hydrogen–deuterium exchange experiments for different deuteration times. The reproducible pepsin digestion pattern covering 99% of the human GW182 SD sequence is shown in Figure 3a. The number of pepsin-generated peptides and the sequence coverage were similar for both fresh and frozen protein samples. This complete sequence coverage with a rich redundancy enabled us to conduct a systematic analysis of the susceptibility for the H/D exchange of particular protein regions. The digestion pattern by pepsin indicated preference for cleavage within the α -helices or β -strands of the RRM domain and not at loops, even though the pepsin buffer contained guanidinium hydrochloride at 0.67 M. A similar preference for the peptide bonds within the defined secondary structural elements rather than in unfolded regions was reported earlier for pepsin digestion of the eIF4E protein [52] and for neutrophil elastase digestion of proteins containing RRM domains [87].

The protection against the exchange reflects the stability of the intramolecular hydrogen bonding network and thus describes the protein structural dynamics on a broad time scale, related to the frequency of transient breakages of particular hydrogen bonds and their accessibility to solvent [88]. A comparison of representative mass spectra of selected peptic GW182 SD peptides after different period of the H/D exchange are shown in Figure 3b.

The majority of the identified peptides displayed the complete exchange of all exchangeable amide protons even at the shortest exchange time (10 s, Figure 4a) in the same manner as for the peptide encompassing the CIM1 motif (residues 1286–1306), shown in the left panels of Figure 3b. The results obtained for all peptides analyzed along the GW182 SD sequence at consecutive HDX time points (10 s, 1 min, 20 min, 2 h) are summarized in Figure 4a–d. These results provide the experimental evidence for the following conclusions: (1) CIM1 reveals no detectable secondary structure; and (2) most of the GW182 silencing domain sequence is not engaged in intramolecular hydrogen bonding, which would result in protection against HDX.

Contrary to the unstructured part, a GW182 SD region encompassing the amino acids 1504–1592 clearly displays substantial protection, prominent at the first two time points. This can be associated with the presence of an ordered α/β protein structure, since this region corresponds to the RRM domain of the protein. The right panels of Figure 3b present the spectra of the initially most protected peptide (1531–1543) belonging to the RRM domain (Figures 1b, c and 3a). All the RRM peptides were gradually exchanged upon time, although the details of their exchange kinetics differed significantly. The

(a)

(b)

◀ Figure 3. (a) Reproducible 99% coverage of the GW182 sequence by redundant peptides obtained by pepsin digestion; (b) representative mass spectra of a peptic GW182 peptide with immediately exchanged amide protons (left panels, residues 1286-1306, charge +2) encompassing the CIM1 motif and a peptide with slowly exchanging amide protons (right panels, residues 1531-1543, charge +2) belonging to the most buried part of the RRM domain, at consecutive HDX time points

HDX pattern reflecting the protein structure is much less visible after 20 min, and completely disappears after 2 h of deuteration, except for the peptides encompassing the $\alpha 2$ helix that displayed a residual protection. This means that the overall protection of the RRM residues against the H/D exchange is low, which indicates significant flexibility and remarkable structural dynamics throughout the whole RRM domain.

The HDX results (Figure 4a) indicate also that the secondary structural assignments modelled by I-TASSER (Figure 1a, c) should be extended more towards the N-terminus of the protein, and include the residues down to Arg1511, involved probably in the $\beta 1$ strand, since four peptides within this region (1498-1515, 1504-1515, 1504-1516, and 1508-1515, except for 1504-1510) display some residual protection against the exchange at 10 s.

Kinetic Parameters of the GW182 RRM Hydrogen–Deuterium Exchange

Kinetic plots for the representative peptic peptides covering the whole GW182 silencing domain sequence are shown in Figure 5. Detailed analysis of the bi-exponential deuterium incorporation kinetics for the individual peptides revealed that each RRM peptide is characterized by two populations of amide hydrogens, referred to as fast and slow.

The fast populations, D_{incf} , incorporate deuteria with the k_f rate constants in the range of ca. 6 to 30 min^{-1} , whereas the slow populations, D_{incs} , are much more protected and the exchange occurs with k_s in the range of ca. 0.02 to 2 min^{-1} . The detailed kinetic parameters for seven peptides encompassing the RRM domain are gathered in Table 1. Each peptide is attributed to its secondary structural element marked in Figures 1 and 4. The percentage of the slow population, D_{incs} , shows that the most protected peptides correspond to the $\alpha 1/\text{turn}/\beta 2$ and $\beta 3/\text{turn}/\alpha 2/\text{turn}$ fragments, while 20% of the latter (residues 1553-1571) remains unexchanged even after 2 h.

The local structural dynamics of the protein (without the freezing-thawing cycle) can be analyzed within the EX2 conditions in terms of the Gibbs free energies, $\Delta G_f^{\text{local}}$ and $\Delta G_s^{\text{local}}$, related to the local stability (minimum on the local energy landscape) of peptide fragments containing fast and slowly exchanging protons. These values were calculated from the protection factors based on the site-specific intrinsic rate constants, k_{int} , for given amino acid sequences. The lowest k_{int} value of 300 min^{-1} was obtained for the PAM2 containing

peptide, whereas the average value for the N- and C-terminal GW182 SD parts, flanking the RRM domain was 630 min^{-1} . The most negative $\Delta G_s^{\text{local}}$ value of about -25 kJ/mol was obtained under the native conditions for the $\alpha 2$ helix containing peptide.

Dual Conformation of the RRM $\alpha 2$ Helix

In general, the GW182 RRM HDX pattern resembles the shape of the predicted disorder probability profile for the peptides within the RRM domain (Figure 1d). However, two significant divergences are clearly visible after closer inspection, i.e., (1) the fragment predicted to be the best ordered (residues 1540-

Figure 4. (a) Fraction of the exchanged population of amide hydrogens (with vertical standard errors) obtained for pepsin-generated GW182 SD peptides marked as horizontal black bars, after 10 s of deuteration, corrected for back-exchange of the individual peptides, overlaid with secondary structural elements assigned and colored as in Figure 1a, c; (b) the same as in (a) after 1 min, (c) after 20 min, and (d) after 120 min of deuteration

Figure 5. Kinetic plots of hydrogen-deuterium exchange for non-redundant pepsin-generated peptides of GW182 SD; data points represent the fraction of the amide hydrogen population exchanged to deuteria after consecutive time intervals of HDX; experimental uncertainty of the individual data points are shown as vertical bars; mono- or bi-exponential functions were fitted to the experimental data points according to Equation 2 or Equation 3

1550) is relatively the most susceptible to exchange among the peptic RRM peptides (Figure 4a, yellow stripe), and (2) the higher relative local disorder tendency predicted by multiple algorithms for the 1555-1570 region corresponding to the $\alpha 2$ helix is in apparent disagreement with HDX, which shows that this RRM fragment is most protected (Figure 4d). While the first observation can be explained by the fact that this fragment corresponds to the $\beta 2$ strand together with the subsequent tight turn (Figure 1a, c) that both can be concurrently well ordered and partially solvent- and exchange-accessible, the latter

contradiction seemed to be highly puzzling, especially since the inspection of all available unliganded RRMs in the PDB showed that just this $\alpha 2$ helix is the most rigid and precisely structurally defined fragment (with the lowest RMSD values) in the RRM structures resolved by NMR. We checked the low-frequency normal modes and the local flexibility, modeled on known PDB templates. The analyses confirmed that the $\alpha 2$ helix residues should not undergo large amplitude fluctuations (Supplementary Figure 1). Bearing in mind that RRM domains are observed to be unstable at higher concentrations upon

Table 1. Kinetic Parameters of Hydrogen–Deuterium Exchange for Pepsin-Generated Peptides of the GW182 RRM Domain. The Percentage of the Fast and Slow Amide Hydrogen Populations, D_{incf} and D_{incs} , that Incorporate Deuterium with the k_f and k_s Rate Constants, Respectively; the Site-Specific Intrinsic Rate Constants, k_{int} ; and the Gibbs Free Energies Related to the Local Stability in the EX2 Conditions, $\Delta G_{\text{f}}^{\text{local}}$ and $\Delta G_{\text{s}}^{\text{local}}$

GW182	fold	D_{incf} (%)	k_f (min^{-1})	D_{incs} (%)	k_s (min^{-1})	k_{int} (min^{-1})	$\Delta G_{\text{f}}^{\text{local}}$ ($\text{kJ}\cdot\text{mol}^{-1}$)	$\Delta G_{\text{s}}^{\text{local}}$ ($\text{kJ}\cdot\text{mol}^{-1}$)
1504-1515	coil/ β 1	94.8 ± 1.6	30^{a}	10.3 ± 1.7	0.5 ± 0.3	801	-8.0	-18.0
1516-1530	β 1/turn/ α 1	71.5 ± 1.0	9.3 ± 0.4	27.7 ± 1.2	0.056 ± 0.006	358	-8.9	-21.3
1531-1543	α 1/turn/ β 2	24.1 ± 1.1	5.8 ± 0.8	77.4 ± 1.4	0.062 ± 0.002	426	-10.5	-21.5
1544-1552	β 2/turn/ β 3	83.0 ± 1.1	13.1 ± 0.9	18.3 ± 1.4	0.045 ± 0.009	734	-9.8	-23.6
1553-1571	β 3/turn/ α 2/turn	18.7 ± 1.1	12 ± 3	62 ± 2	0.022 ± 0.002	540	-9.3	-24.6
1572-1580	β 3'/turn/ β 4	55.2 ± 1.6	7.4 ± 0.7	44 ± 2	0.032 ± 0.005	403	-9.7	-23.0
1581-1592	β 4/turn/ α 3	55.9 ± 1.8	30^{a}	35.3 ± 1.9	1.08 ± 0.13	336	-5.9	-14.0
<i>after freezing-thawing cycle: double conformation of 1554-1571 peptide (α2) in the EX1 conditions</i>								
<i>closed</i>	β 3/turn/ α 2/turn	15.1 ± 1.9	30^{a}	28 ± 3	0.1^{a}			
<i>open</i>	β 3/turn/ α 2/turn	89.8 ± 1.0	10.4 ± 0.4	5.9 ± 1.2	0.1^{a}			

^aFixed values, according to [52, 55, 56]

longer times, we checked how the α 2 helix conformation could be affected by the freezing-thawing cycle.

The mass spectra recorded for the frozen-thawed samples revealed clearly that the peptides encompassing the α 2 helix (residues 1544-1571) displayed a bimodal isotopic envelope (Figure 6a). The shortest of them was RYSSKEEAQAQKSLHM (1555-1571, Figure 7a). This means that the α 2 helix exists in two distinct conformations. The more exchange-protected conformation of this peptide, denoted to as the closed state, has a similar H/D exchange kinetics (Figure 6b) as the sole form detected for the nonfrozen samples (Figure 5). The more exchange-competent conformation, denoted to as the open state, displays full deuteration after 20 min of the exchange (Figure 6a, b). The populations of the two states at the zero exchange time are comparable (64% versus 36%, Figure 6c), which corresponds to the difference in the Gibbs free energies between these two states is less than RT , $\Delta G = 1.4 \text{ kJ/mol}$ (Figure 6d).

The observed type of bimodal distribution is characteristic for the EX1 kinetics, where the chemical H/D exchange is faster compared with a characteristic time constant for global conformational changes or subunit dissociation that break most of the solvent-exposed hydrogen bonds, before refolding or rebinding takes place [4, 62, 90]. The equilibrium conformation populations remain constant at given temperature, but an apparently observed decrease in the population of the α 2 peptide in the closed conformation and the corresponding apparently observed increase in the open conformation population are visible (Figure 6c) due to the irreversible exchange in excess of D_2O during the time spent by the peptide in the open state before coming back to the closed state (see Materials and Methods). This makes it possible to estimate the kinetic rate constant, k^{++} , for the transition over the barrier from the closed to open conformation within the Eyring-Polanyi transition state theory. Since the observed process is very slow, with k^{++} in the order of 1 h^{-1} , the energy barrier that separates the conformations is very high, about 90 kJ/mol (Figure 6d), which is ~ 36 -fold higher than RT . An average binding free energy contribution from a single non-covalent contact in a water-accessible protein region ranges from -3 to -8 kJ/mol and is similar for salt bridges, hydrogen bonds, van der Waals interactions, and

hydrophobic contacts due to entropic effects related to the presence of counter-ions, surrounding water molecules in buffer and to the penalty of ordering [91]. This means that the transition from the closed to open conformation would require a transient break of more than a dozen structure stabilizing non-covalent contacts within the relatively small RRM domain. Figures 5e and 6a presents the non-covalent side chain contacts formed by the α 2 helix with the residues within the distance of $3.5 \text{ \AA}$ belonging to the other secondary structural elements: β 1, β 3, α 1, and β 4 strands. The peptides of the frozen-thawed protein containing the T1512, S1514, W1515, L1516, L1518, R1519, L1533, C1534, H1537, P1539, I1541, T1542, TILAEF(1578-1583), and E1586 residues (Figure 7a) displayed also dual isotopic envelopes but to a lesser extent.

An example of a rupture of such a great number of contacts leading to a different but energetically equivalent conformational state for a single RRM domain in the absence of a larger globular protein context can be found in the PDB, i.e., the RRM of the poly(A)-specific ribonuclease (PARN) [89]. Full length PARN is a homodimer composed of three domains in each subunit (the nuclease, R3H, and RRM) [19, 92], and its RRM domain is prone to dimerization [93]. When crystallized alone, the truncated RRM domain of PARN forms a dimer with a perfect swapping of just the α 2 helices between the RRM molecules [89] (Figure 7b). Both a putative transition of the α 2 helix as a whole to a swapped conformation, and a cooperative movement with a transient unfolding of intra-helical hydrogen bonds, could yield similar EX1 characteristics. However, it seems hardly possible that the Gibbs free energy of unfolded conformations of the α 2 helix, i.e., the most tightly folded RRM part (according to the PDB NMR data) could be almost equal to those of the helical state (Figure 6d).

Moreover, slow motions related to conformational changes or folding of individual protein residues are in the microsecond to millisecond timescale [94, 95], whereas characteristic times of primary nucleation phase upon protein aggregation are hours [96]. The double isotopic envelope effect related to the presence of two long-lasting conformations observed for the frozen-thawed samples could be thus rather a result of a conformational rearrangement of the whole α 2 helix. The possibility of the existence of α 2 in two conformations could, in turn,

◀ Figure 6. **(a)** Bimodal isotopic envelopes for the GW182 SD RRM $\alpha 2$ peptide (residues 1544-1571, charge +5) after given time of H/D exchange; (—) closed conformation, (---) open conformation; **(b)** Kinetic plots of HDX for the 1544-1571 peptide; data points represent the fraction of the amide hydrogen population exchanged to deuteria after consecutive time intervals of HDX for (●) closed $\alpha 2$ peptide conformation, (○) open conformation, (x) average values; **(c)** apparent populations in the closed and open state as seen by HDX: (●) the fraction of the population of the closed state molecules that never once underwent the change to the open conformation, (○) population of the open state together with the fraction of the population of the closed state molecules that at least once underwent the change to the open conformation; **(d)** Gibbs free energy barrier for the conformational equilibrium of the $\alpha 2$ peptide; **(e)** contact map (within 3.5 Å) for the side chains of the GW182 SD RRM (residues 1509-1595). The α -helix 2 region is marked as gray band. The contacts between the $\alpha 2$ residues with the $\beta 1$, $\beta 3$, $\beta 4$, and $\alpha 1$ are encircled

explain why the disorder tendency prediction (Figure 1d) is higher for the tight $\alpha 2$ helix than for the surrounding area.

A question arises whether – alternatively – the GW182 RRM region that displays the bimodal isotope envelopes might serve as a nucleation region for protein aggregation during the freezing-thawing cycle, and thus, if the observed EX1 kinetics of conformational changes might reflect a process of associated retarded re-solubilization. However, if this were the case, the sole conformation detected for the soluble, nonfrozen protein would correspond to the open conformation, whereas the opposite is true. The H/D exchange kinetics (residues 1553-1571 in Figure 5) detected for fresh samples is similar to the closed conformation of the 1554-1571 peptide (Figure 6b) in the frozen samples (for kinetic parameters, see Table 1).

To further check whether the bimodal isotope envelopes in HDX MS could be evidence of conformational changes at the secondary structure level that could lead to protein aggregation, we measured the far-UV CD spectra of GW182 SD (Figure 8).

The main negative peak at 202 nm and a shoulder at 220 nm point to the prevailing contribution coming from disordered residues with much lesser contributions of the residues in the α -helical, antiparallel β -sheet, and β -turn conformations. The spectra of both the freshly prepared and frozen-thawed protein are very similar. The small difference observed in the range of the shortest wavelengths (<200 nm) might be attributed to a decrease in β -turn or coil content, but the putative change is in the order of the experimental uncertainty. This means that the populations of the amino acid residues in varying secondary structure conformations present in the fresh and frozen samples are essentially unchanged, and the putative slight decrease in β -turn would be in agreement with the movement of the whole $\alpha 2$ helix (Figure 7). We can thus conclude again that the dual isotopic envelopes in the HDX MS of the frozen GW182 samples are related to the conformational change at the RRM tertiary structure level rather than to direct aggregation or misfolding. The exact numerical determination of the secondary structure contents by means of available CD software failed due to the mainly unstructured character of the protein, while the software are focused and trained on protein sets with known crystal structures.

Detailed analyses of the molecular properties of proteins containing intrinsically disordered regions are still a challenging task, where HDX MS appears to be the method of choice. Conventional approaches that address protein structure-related issues based on NMR and crystallography require high concentrations of proteins. Not all proteins can be produced in such quantities, while many aggregate or precipitate. HDX MS is a far less protein- and time-consuming approach than multidimensional NMR. Moreover, IDPs do not exhibit a stable 3D conformation in physiological conditions, and thus are not suitable for crystallography. HDX MS can provide an insight into the structural dynamics and intermolecular interactions of such proteins, since this method detects local stability of the hydrogen-bonding network, which reflects protein flexibility, conformational transitions, and – indirectly – solvent accessibility, at a peptide level. This is a significantly higher resolution compared with spectroscopic techniques such as circular

Figure 7. **(a)** Human GW182 SD RRM model with the shortest $\alpha 2$ peptide (RYSSKEEAAKAQSLHM) that displays the bimodal isotopic envelope, marked yellow. Residues in non-covalent contacts with $\alpha 2$ (≤ 3.5 Å, excluding internal $\alpha 2$ contacts) belonging to $\alpha 1$, $\beta 1$, $\beta 3$, $\beta 4$ (T15, S17, W18, L19, L21, R22, L36, C37, H40, P42, I44, T45, TILAEEF(81-86), E89) shown as CPK balls-and-sticks; **(b)** crystal structure of a single RRM domain truncated from human PARN, with the $\alpha 2$ helix that is swapped in the RRM dimer (PDB id code: 3CTR [89]) marked gold

Figure 8. Far-UV CD spectra of GW182 silencing domain; (○) nonfrozen protein, (■) protein after the freezing-thawing cycle

dichroism or steady-state fluorescence and correlation spectroscopy that give structurally averaged data. Even for globular domains like the RRM, HDX MS is advantageous in studying the long time-scale conformational changes, which would be hardly observable by other methods, including mutation-based labeling for FRET or DEER ESR. HDX MS can detect conformational inhomogeneity based on bimodal isotope envelopes from 1-d measurements, directly after protein purification. Our results show that the influence of the environmental conditions on the RRM domain structure should not be neglected, since this may have an impact on interactions with other molecules.

Since GW182 SD is a mainly intrinsically disordered protein and is available in limited amounts, HDX-MS was deemed to be the most suitable method for studying its structural dynamics in solution. The results described here form a crucial biophysical basis for further experiments that address the nature of molecular interactions of GW182 SD with the CNOT1 subunit of the CCR4-NOT de-adenylation machinery in the process of miRNA-mediated gene silencing.

Conclusions

Our work reveals the hydrogen–deuterium exchange mass spectrometric description of an RNA Recognition Motif, which is one of the most frequently occurring canonical protein domains. We show that the RRM structure reveals the alternating HDX pattern, composed of the regions containing amide hydrogens that are more or less exchange-competent, even when the RRM domain is studied in the context of a larger fragment of the generally disordered protein, which is not suitable for NMR due to its size and aggregation at higher concentrations. The experimentally determined structural dynamics of the GW182 SD RRM correlates well, in general, with the predicted α/β content. We found that the RRM structure is highly dynamic, since the complete H/D exchange is achieved in 2 h. These results indicate that water molecules can penetrate the

whole domain. The HDX studies suggest also that the $\alpha 2$ helix of the RRM can change its conformation after a freezing-thawing cycle, which might be a clue for explanation of a possible RRM aggregation pathway. These studies have provided the experimental proof that the silencing domain (SD) of GW182, except the RRM domain, is indeed natively unstructured. In particular, the CCR4-NOT interacting motif 1 does not display any evidence of a secondary structure in the HDX MS timescale.

Acknowledgments

The authors thank Professor K. Hall for helpful discussion, Professor E. Darzynkiewicz for access to his laboratory at the Division of Biophysics, University of Warsaw, and R. Borowski from IP PAN for eIF4E expression and purification. The work was partially supported by ERDF through the NanoFun Project POIG.02.02.00-00-025/09 and the Foundation for Polish Science International PhD Projects Program at the Faculty of Physics, University of Warsaw; grants no. UMO-2013/11/N/NZ1/02387 to M.C.R., UMO-2014/14/A/NZ1/00306 to M.D., and UMO-2016/22/E/NZ1/00656 to A.N. from the National Science Center, Poland.

Open Access

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Berendsen, H.J., Hayward, S.: Collective protein dynamics in relation to function. *Curr. Opin. Struct. Biol.* **10**, 165–169 (2000)
- Hensen, U., Meyer, T., Haas, J., Rex, R., Vriend, G., Grubmüller, H.: Exploring protein dynamics space: the dyasome as the missing link between protein structure and function. *PLoS One* **7**, e33931 (2012)
- Smith, C.A., Ban, D., Pratihar, S., Giller, K., Paulat, M., Becker, S., Griesinger, C., Lee, D., de Groot, B.L.: Allosteric switch regulates protein–protein binding through collective motion. *Proc. Natl. Acad. Sci.* **113**, 201519609 (2016)
- Yan, X., Watson, J., Ho, P.S., Deinzer, M.L.: Mass spectrometric approaches using electrospray ionization charge states and hydrogen–deuterium exchange for determining protein structures and their conformational changes. *Mol. Cell. Proteomics* **3**, 10–23 (2004)
- Balasubramaniam, D., Komives, E.A.: Hydrogen-exchange mass spectrometry for the study of intrinsic disorder in proteins. *Biochim. Biophys. Acta* **1834**, 1202–1209 (2013)
- Nakazawa, S., Ahn, J., Hashii, N., Hirose, K., Kawasaki, N.: Analysis of the local dynamics of human insulin and a rapid-acting insulin analog by hydrogen/deuterium exchange mass spectrometry. *Biochim. Biophys. Acta* **1834**, 1210–1214 (2013)
- Zhu, S., Shala, A., Bezginov, A., Slijoka, A., Audette, G., Wilson, D.J.: Hyperphosphorylation of intrinsically disordered tau protein induces an amyloidogenic shift in its conformational ensemble. *PLoS One* **10**, 1–15 (2015)
- Maris, C., Dominguez, C., Allain, F.H.-T.: The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J.* **272**, 2118–2131 (2005)

9. Cléry, A., Blatter, M., Allain, F.H.-T.: RNA recognition motifs: boring? Not quite. *Curr. Opin. Struct. Biol.* **18**, 290–298 (2008)
10. Daubner, G.M., Cléry, A., Allain, F.H.-T.: RRM-RNA recognition: NMR or crystallography...and new findings. *Curr. Opin. Struct. Biol.* **23**, 100–108 (2013)
11. Fabian, M.R., Cieplak, M.K., Frank, F., Morita, M., Green, J., Srikumar, T., Nagar, B., Yamamoto, T., Raught, B., Duchaine, T.F., Sonenberg, N.: miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4–NOT. *Nat. Struct. Mol. Biol.* **18**, 1211–1217 (2011)
12. Eulalio, A., Tritschler, F., Büttner, R., Weichenrieder, O., Izaurralde, E., Truffault, V.: The RRM domain in GW182 proteins contributes to miRNA-mediated gene silencing. *Nucleic Acids Res.* **37**, 2974–2983 (2009)
13. Gonatopoulos-Pournatzis, T., Cowling, V.H.: Cap-binding complex (CBC). *Biochem. J.* **457**, 231–242 (2014)
14. Virtanen, A., Henriksson, N., Nilsson, P., Nissbeck, M.: Poly(A)-specific ribonuclease (PARN): an allosterically regulated, processive, and mRNA cap-interacting deadenylase. *Crit. Rev. Biochem. Mol. Biol.* **48**, 192–209 (2013)
15. Mazza, C., Segref, A., Mattaj, I.W., Cusack, S.: Large-scale induced fit recognition of an m7GpppG cap analogue by the human nuclear cap-binding complex. *EMBO J.* **21**, 5548–5557 (2002)
16. Worch, R., Niedzwiecka, A., Stepinski, J., Mazza, C., Jankowska-Anyszka, M., Darzynkiewicz, E., Cusack, S., Stolarski, R.: Specificity of recognition of mRNA 5' cap by human nuclear cap-binding complex. *RNA*. **11**, 1355–1363 (2005)
17. Worch, R., Jankowska-Anyszka, M., Niedzwiecka, A., Stepinski, J., Mazza, C., Darzynkiewicz, E., Cusack, S., Stolarski, R.: Diverse role of three tyrosines in binding of the RNA 5' cap to the human nuclear cap binding complex. *J. Mol. Biol.* **385**, 618–627 (2009)
18. Nilsson, P., Henriksson, N., Niedzwiecka, A., Balatsos, N.A.A., Kokkoris, K., Eriksson, J., Virtanen, A.: A multifunctional RNA recognition motif in poly(A)-specific ribonuclease with cap and poly(A) binding properties. *J. Biol. Chem.* **282**, 32902–32911 (2007)
19. Wu, M., Nilsson, P., Henriksson, N., Niedzwiecka, A., Lim, M.K., Cheng, Z., Kokkoris, K., Virtanen, A., Song, H.: Structural basis of m7GpppG binding to poly(A)-specific ribonuclease. *Structure*. **17**, 276–286 (2009)
20. Niedzwiecka, A., Nilsson, P., Worch, R., Stepinski, J., Darzynkiewicz, E., Virtanen, A.: Molecular recognition of mRNA 5' cap by 3' poly(A)-specific ribonuclease (PARN) differs from interactions known for other cap-binding proteins. *Biochim. Biophys. Acta Protein Proteomics*. **1864**, 331–345 (2016)
21. Meister, G., Landthaler, M., Peters, L., Chen, P.Y., Urlaub, H., Lühmann, R., Tuschl, T.: Identification of novel Argonaute-associated proteins. *Curr. Biol.* **15**, 2149–2155 (2005)
22. Takimoto, K., Wakiyama, M., Yokoyama, S.: Mammalian GW182 contains multiple Argonaute-binding sites and functions in microRNA-mediated translational repression. *RNA*. **15**, 1078–1089 (2009)
23. Landthaler, M., Gaidatzis, D., Rothballer, A., Chen, P.Y., Soll, S.J., Dinic, L., Ojo, T., Hafner, M., Zavolan, M., Tuschl, T.: Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs. *RNA*. **14**, 2580–2596 (2008)
24. Chekulaeva, M., Mathys, H., Zipprich, J.T., Attig, J., Colic, M., Parker, R., Filipowicz, W.: miRNA repression involves GW182-mediated recruitment of CCR4–NOT through conserved W-containing motifs. *Nat. Struct. Mol. Biol.* **18**, 1218–1226 (2011)
25. Braun, J.E., Huntzinger, E., Fauser, M., Izaurralde, E.: GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Mol. Cell*. **44**, 120–133 (2011)
26. Lazzaretti, D., Toumier, I., Izaurralde, E.: The C-terminal domains of human TNRC6A, TNRC6B, and TNRC6C silence bound transcripts independently of Argonaute proteins. *RNA*. **15**, 1059–1066 (2009)
27. Eulalio, A., Huntzinger, E., Izaurralde, E.: GW182 interaction with Argonaute is essential for miRNA-mediated translational repression and mRNA decay. *Nat. Struct. Mol. Biol.* **15**, 346–353 (2008)
28. Eulalio, A., Helms, S., Fritsch, C., Fauser, M., Izaurralde, E.: A C-terminal silencing domain in GW182 is essential for miRNA function. *RNA*. **15**, 1067–1077 (2009)
29. Eystathiou, T., Chan, E.K.L., Tenenbaum, S.A., Keene, J.D., Griffith, K., Fritzler, M.J.: A phosphorylated cytoplasmic autoantigen, GW182, associates with a unique population of human mRNAs within novel cytoplasmic speckles. *Mol. Biol. Cell*. **13**, 1338–1351 (2002)
30. Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., Izaurralde, E.: mRNA degradation by miRNAs and GW182 requires both CCR4–NOT deadenylase and DCP1–DCP2 decapping complexes. *Genes Dev.* **20**, 1885–1898 (2006)
31. Zieleszinski, A., Karlowski, W.M.: Early origin and adaptive evolution of the GW182 protein family, the key component of RNA silencing in animals. *RNA Biol.* **12**, 761–770 (2015)
32. Braun, J.E., Huntzinger, E., Izaurralde, E.: The role of GW182 proteins in miRNA-mediated gene silencing. *Adv. Exp. Med. Biol.* **768**, 147–163 (2013)
33. Zipprich, J.T., Bhattacharyya, S., Mathys, H., Filipowicz, W.: Importance of the C-terminal domain of the human GW182 protein TNRC6C for translational repression. *RNA*. **15**, 781–793 (2009)
34. Huntzinger, E., Braun, J.E., Heimstädt, S., Zekri, L., Izaurralde, E.: Two PABPC1-binding sites in GW182 proteins promote miRNA-mediated gene silencing. *EMBO J.* **29**, 4146–4160 (2010)
35. Barraud, P., Banerjee, S., Mohamed, W.I., Jantsch, M.F., Allain, F.H.: A bimodular nuclear localization signal assembled via an extended double-stranded RNA-binding domain acts as an RNA-sensing signal for transportin 1. *Proc Natl Acad Sci U S A*. **111**, E1852–E1861 (2014)
36. El-Shami, M., Pontier, D., Lahmy, S., Braun, L., Picart, C., Vega, D., Hakimi, M., Jacobsen, S.E., Cooke, R., Lagrange, T.: Reiterated WG/GW motifs form functionally and evolutionarily conserved Argonaute-binding platforms in RNAi-related components service. *Genes Dev.* **21**, 2539–2544 (2007)
37. Fabian, M.R., Mathonnet, G., Sundermeier, T., Mathys, H., Zipprich, J.T., Svitkin, Y.V., Rivas, F., Jinek, M., Wohlschlegel, J., Doudna, J.A., Chen, C.-Y.A., Shyu, A.-B., Yates, J.R., Hannon, G.J., Filipowicz, W., Duchaine, T.F., Sonenberg, N.: Mammalian miRNA RISC recruits CAF1 and PABP to affect PABP-dependent deadenylation. *Mol. Cell*. **35**, 868–880 (2009)
38. Zekri, L., Huntzinger, E., Heimstädt, S., Izaurralde, E.: The silencing domain of GW182 interacts with PABPC1 to promote translational repression and degradation of microRNA targets and is required for target release. *Mol. Cell Biol.* **29**, 6220–6231 (2009)
39. Li, S., Lian, S.L., Moser, J.J., Fritzler, M.L., Fritzler, M.J., Satoh, M., Chan, E.K.L.: Identification of GW182 and its novel isoform TNGW1 as translational repressors in Ago2-mediated silencing. *J. Cell Sci.* **121**, 4134–4144 (2008)
40. Tompa, P.: Intrinsically disordered proteins: a 10-year recap. *Trends Biochem. Sci.* **37**, 509–516 (2012)
41. Morgan, C.R., Engen, J.R.: Investigating solution-phase protein structure and dynamics by hydrogen exchange mass spectrometry. *Curr. Protoc. Protein Sci.* **17**(6), 1–17 (2009)
42. Ross, J.L.: The dark matter of biology. *Biophys. J.* **111**, 909–916 (2016)
43. Ward, J.J., Sodhi, J.S., McGuffin, L.J., Buxton, B.F., Jones, D.T.: Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.* **337**, 635–645 (2004)
44. Wright, P.E., Dyson, H.J.: Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **293**, 321–331 (1999)
45. Mészáros, B., Tompa, P., Simon, I., Dosztányi, Z.: Molecular principles of the interactions of disordered proteins. *J. Mol. Biol.* **372**, 549–561 (2007)
46. Tompa, P.: Intrinsically unstructured proteins. *Trends Biochem. Sci.* **27**, 527–533 (2002)
47. Gunasekaran, K., Tsai, C.-J., Nussinov, R.: Analysis of ordered and disordered protein complexes reveals structural features discriminating between stable and unstable monomers. *J. Mol. Biol.* **341**, 1327–1341 (2004)
48. Rauscher, S., Gapsys, V., Gajda, M.J., Zweckstetter, M., de Groot, B.L., Grubmüller, H.: Structural ensembles of intrinsically disordered proteins depend strongly on force field: a comparison to experiment. *J. Chem. Theory Comput.* **11**, 5513–5524 (2015)
49. Miller, D.W., Dill, K.A.: A statistical mechanical model for hydrogen exchange in globular proteins. *Protein Sci.* **4**, 1860–1873 (1995)
50. Hoofnagle, A.N., Resing, K.A., Ahn, N.G.: Protein analysis by hydrogen exchange mass spectrometry. *Annu. Rev. Biophys. Biomol. Struct.* **32**, 1–25 (2003)
51. Skinner, J.J., Lim, W.K., Bédard, S., Black, B.E., Englander, S.W.: Protein dynamics viewed by hydrogen exchange. *Protein Sci.* **21**, 996–1005 (2012)
52. Rutkowska-Włodarczyk, I., Stepinski, J., Dadlez, M., Darzynkiewicz, E., Stolarski, R., Niedzwiecka, A.: Structural changes of eIF4E upon binding

- to the mRNA 5' monomethylguanosine and trimethylguanosine Cap. *Biochemistry*. **47**, 2710–2720 (2008)
53. Tamowski, K., Fituch, K., Szczepanowski, R.H., Dadlez, M., Kaus-Drobek, M.: Patterns of structural dynamics in RACK1 protein retained throughout evolution: a hydrogen–deuterium exchange study of three orthologs. *Protein Sci.* **23**, 639–651 (2014)
 54. Weis, D.D., Engen, J.R., Kass, I.J.: Semi-automated data processing of hydrogen exchange mass spectra using HX-Express. *J. Am. Soc. Mass Spectrom.* **17**, 1700–1703 (2006)
 55. Mandell, J.G., Baerga-Ortiz, A., Akashi, S., Takio, K., Komives, E.A.: Solvent accessibility of the thrombin-thrombomodulin interface. *J. Mol. Biol.* **306**, 575–589 (2001)
 56. Gemmecker, G., Jahnke, W., Kessler, H.: Measurement of fast proton exchange rates in isotopically labeled compounds. *J. Am. Chem. Soc.* **115**, 11620–11621 (1993)
 57. Akaike, H.: A new look at the statistical model identification. *IEEE Trans. Automat. Contr.* **19**, 716–723 (1974)
 58. Beyer, W.H.: CRC standard mathematical tables, 28th edn. CRC Press, Boca Raton (1987)
 59. Zhang, Y.-Z.: Protein and peptide structure and interactions studied by hydrogen exchange and NMR. Ph.D. Thesis. Structural Biology and Molecular Biophysics, University of Pennsylvania, PA, USA (1995)
 60. Bai, Y., Milne, J.S., Mayne, L., Englander, S.W.: Primary structure effects on peptide group hydrogen exchange. *Proteins*. **17**, 75–86 (1993)
 61. Connelly, G.P., Bai, Y., Jeng, M.F., Englander, S.W.: Isotope effects in peptide group hydrogen exchange. *Proteins*. **17**, 87–92 (1993)
 62. Wales, T.E., Engen, J.R.: Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrom. Rev.* **25**, 158–170 (2006)
 63. Eyring, H., Polanyi, M.: Über einfache Gasreaktionen. *Z. Phys. Chem. Abt. B*. **12**, 279–311 (1931)
 64. Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., Zhang, Y.: The I-TASSER suite: protein structure and function prediction. *Nat. Methods*. **12**, 7–8 (2015)
 65. Guex, N., Peitsch, M.C.: SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*. **18**, 2714–2723 (1997)
 66. Kozłowski, L.P., Bujnicki, J.M.: MetaDisorder: a meta-server for the prediction of intrinsic disorder in proteins. *BMC Bioinf.* **13**, 111 (2012)
 67. Linding, R., Jensen, L.J., Diella, F., Bork, P., Gibson, T.J., Russell, R.B.: Protein disorder prediction: implications for structural proteomics. *Structure*. **11**, 1453–1459 (2003)
 68. Obradovic, Z., Peng, K., Vucetic, S., Radivojac, P., Dunker, A.K.: Exploiting heterogeneous sequence properties improves prediction of protein disorder. *Proteins Struct. Funct. Genet.* **61**, 176–182 (2005)
 69. Su, C.T., Chen, C.Y., Hsu, C.M.: IPDA: Integrated protein disorder analyzer. *Nucleic Acids Res.* **35**, 465–472 (2007)
 70. Cheng, J., Sweredoski, M.J., Baldi, P.: Accurate prediction of protein disordered region by mining protein structure data. *Data Min. Knowl. Disc.* **11**, 213–222 (2005)
 71. Linding, R., Russell, R.B., Neduva, V., Gibson, T.J.: GlobPlot: exploring protein sequences for globularity and disorder. *Nucleic Acids Res.* **31**, 3701–3708 (2003)
 72. Dosztanyi, Z., Csizmok, V., Tompa, P., Simon, I.: IUPred: Web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics*. **21**, 3433–3434 (2005)
 73. Hirose, S., Shimizu, K., Kanai, S., Kuroda, Y., Noguchi, T.: POODLE-L: a two-level SVM prediction system for reliably predicting long disordered regions. *Bioinformatics*. **23**, 2046–2053 (2007)
 74. Ishida, T., Kinoshita, K.: PrDOS: prediction of disordered protein regions from amino acid sequence. *Nucleic Acids Res.* **35**, 460–464 (2007)
 75. Vullo, A., Bortolami, O., Pollastri, G., Tosatto, S.C.E.: Spritz: a server for the prediction of intrinsically disordered regions in protein sequences using kernel machines. *Nucleic Acids Res.* **34**, W164–W168 (2006)
 76. Yang, Z.R., Thomson, R., McNeil, P., Esnouf, R.M.: RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. *Bioinformatics*. **21**, 3369–3376 (2005)
 77. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J.: Gapped BLAST and PS I-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997)
 78. Söding, J.: Protein homology detection by HMM-HMM comparison. *Bioinformatics*. **21**, 951–960 (2005)
 79. Jaroszewski, L., Rychlewski, L., Li, Z., Li, W., Godzik, A.: FFAS03: a server for profile–profile sequence alignments. *Nucleic Acids Res.* **33**, 284–288 (2005)
 80. Lobley, A., Sadowski, M.I., Jones, D.T.: pGenTHREADER and pDomTHREADER: new methods for improved protein fold recognition and superfamily discrimination. *Bioinformatics*. **25**, 1761–1767 (2009)
 81. Wallner, B., Elofsson, A.: Pcons5: combining consensus, structural evaluation, and fold recognition scores. *Bioinformatics*. **21**, 4248–4254 (2005)
 82. Kelley, L.A., Sternberg, M.J.E.: Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc.* **4**, 363–371 (2009)
 83. Kavan, D., Man, P.: MSTools - Web based application for visualization and presentation of HXMS data. *Int. J. Mass Spectrom.* **302**, 53–58 (2011)
 84. Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., Higgins, D.G.: Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011)
 85. Jamroz, M., Kolinski, A., Kmiecik, S.: CABS-flex predictions of protein flexibility compared with NMR ensembles. *Bioinformatics*. **30**, 2150–2154 (2014)
 86. Tiwari, S.P., Fuglebakk, E., Hollup, S.M., Skjærven, L., Cragolini, T., Grindhaug, S.H., Tekle, K.M., Reuter, N.: WEBnm@ v2.0: Web server and services for comparing protein flexibility. *BMC Bioinformatics*. **15**, 427 (2014)
 87. Kistowski, M., Dębski, J., Karczmarski, J., Paziewska, A., Oledzki, J., Mikula, M., Ostrowski, J., Dadlez, M.: A strong neutrophil elastase proteolytic fingerprint marks the carcinoma tumor proteome. *Mol. Cell. Proteomics*. **16**, 213–227 (2017)
 88. Konermann, L., Pan, J., Liu, Y.H.: Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem. Soc. Rev.* **40**, 1224–1234 (2011)
 89. Monecke, T., Schell, S., Dickmanns, A., Ficner, R.: Crystal structure of the RRM domain of poly(A)-specific ribonuclease reveals a novel m7G-cap-binding mode. *J. Mol. Biol.* **382**, 827–834 (2008)
 90. Englander, S.W.: Hydrogen exchange and mass spectrometry: a historical perspective. *J. Am. Soc. Mass Spectrom.* **17**, 1481–1489 (2006)
 91. Nick Pace, C., Martin Scholtz, J., Grimsley, G.R.: Forces stabilizing proteins. *FEBS Lett.* **588**, 2177–2184 (2014)
 92. Wu, M., Reuter, M., Lilie, H., Liu, Y., Wahle, E., Song, H.: Structural insight into poly(A) binding and catalytic mechanism of human PARN. *EMBO J.* **24**, 4082–4093 (2005)
 93. Niedzwiecka, A., Lekka, M., Nilsson, P., Virtanen, A.: Global architecture of human poly(A)-specific ribonuclease by atomic force microscopy in liquid and dynamic light scattering. *Biophys. Chem.* **158**, 141–149 (2011)
 94. Maynard, C.M., Hall, K.B.: Interactions between PTB RRM domains induce slow motions and increase RNA binding affinity. *J. Mol. Biol.* **397**, 260–277 (2010)
 95. Chung, H.S., Eaton, W.A.: Single-molecule fluorescence probes dynamics of barrier crossing. *Nature*. **502**, 685–688 (2013)
 96. Knowles, T.P.J., Vendruscolo, M., Dobson, C.M.: The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **15**, 384–396 (2014)