



# OPEN Epigallocatechin-3-gallate binds tandem RNA recognition motifs of TDP-43 and inhibits its aggregation

Maria Agnese Morando<sup>1,8</sup>, Vito D'Alessandro<sup>1,2,8</sup>, Angelo Spinello<sup>3</sup>, Martina Sollazzo<sup>1,3</sup>, Elisa Monaca<sup>1</sup>, Raffaele Sabbatella<sup>1</sup>, Maria Concetta Volpe<sup>4</sup>, Francesca Gervaso<sup>5</sup>, Alessandro Polini<sup>5</sup>, Sarah Mizielinska<sup>6,7</sup> & Caterina Alfano<sup>1</sup>✉

Transactive response DNA-binding Protein 43 (TDP-43) aggregation is a key pathological feature in Amyotrophic Lateral Sclerosis and related neurodegenerative diseases. This study investigates the inhibitory effects of Epigallocatechin-3-gallate (EGCG), a polyphenol found in green tea, on TDP-43 aggregation. Using a combination of fluorescence assays, NMR spectroscopy, and computational modeling, we demonstrate that Epigallocatechin-3-gallate significantly delays the nucleation phase of TDP-43 aggregation process, thus inhibiting the formation of TDP-43 aggregates *in vitro*. Additionally, we proved a direct interaction of the compound with the RNA recognition motifs of TDP-43 and modeled the mechanism of interaction. Our findings reveal that EGCG stabilizes the RRM domains, counteracting aggregation by interfering with the early stages of the amyloidogenic pathway. Furthermore, EGCG's stability under experimental conditions was ensured using reducing agents, highlighting the importance of maintaining its reduced form for reproducible results. These insights underscore the therapeutic potential of EGCG in TDP-43 proteinopathies and provide a foundation for developing targeted treatments for ALS and related disorders.

**Keywords** EGCG stability, TDP-43, ALS, Protein aggregation, RNA-binding proteins, Protein-ligand interaction

Amyotrophic Lateral Sclerosis (ALS) is a debilitating and fatal neurodegenerative disease characterized by progressive motor neuron degeneration.<sup>1,2</sup> A pathological hallmark of ALS is the formation of cytosolic aggregates composed of poly-ubiquitinated and hyper-phosphorylated Transactive response DNA-binding Protein 43 (TDP-43) within neurons and glial cells.<sup>3–5</sup> The significance of TDP-43 in ALS pathogenesis is further underscored by the identification of more than 50 mutations in the TDP-43 gene (TARDBP) associated with ALS, highlighting the central role of this protein in the disease.<sup>6–12</sup> Notably, TDP-43 mutations are also implicated in frontotemporal lobar dementia (FTD), a neurodegenerative condition that shares clinical and pathological features with ALS. Approximately 45% of FTD cases also exhibit TDP-43 pathology, suggesting a common molecular underpinning between these two diseases.<sup>3,5,13</sup>

TDP-43 is a nuclear factor that plays a crucial role in gene expression, particularly in transcriptional regulation and alternative splicing. Structurally, this protein is a nucleic acid-binding protein that comprises a well-folded N-terminal domain, two conserved RNA recognition motifs (RRM1 and RRM2), and a prion-like C-terminal domain, with the latter harboring the majority of disease-associated mutations.<sup>14–16</sup> In its normal physiological state, TDP-43 primarily resides in the nucleus, where it is involved in RNA metabolism and processing by binding to UG-rich RNA and single-stranded TG-rich DNA sequences through its RRM domains.<sup>17</sup> However, pathologic TDP-43 is mislocalized in the cytoplasm, where it forms pathological aggregates in over 97% of cases.<sup>18,19</sup> These aberrant cytoplasmic inclusions often colocalize with stress granules, which are transient membrane-less organelles composed of protein-RNA complexes formed in response to cellular stress.<sup>20</sup>

<sup>1</sup>Structural Biology and Biophysics Unit, Fondazione Ri.MED, 90133 Palermo, Italy. <sup>2</sup>Department of Mathematics and Physics "E. De Giorgi", University of Salento, 73100 Lecce, Italy. <sup>3</sup>Department of Biological, Chemical and Pharmaceutical Sciences and Technologies (STEBICEF), University of Palermo, 90028 Palermo, Italy. <sup>4</sup>Fondazione Ri.MED, 90133 Palermo, Italy. <sup>5</sup>CNR Nanotec-Institute of Nanotechnology, Campus Ecotekne, 73100 Lecce, Italy. <sup>6</sup>UK Dementia Research Institute at King's College London, London, UK. <sup>7</sup>Department of Basic and Clinical Neuroscience, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. <sup>8</sup>Maria Agnese Morando and Vito D'Alessandro contributed equally to this work. ✉email: calfano@fondazionerimed.com

Multiple studies have identified the RRM domains of TDP-43 as critical players in aggregation and pathology, and amyloidogenic sequence elements are found in these domains.<sup>21–28</sup> Moreover, it has been shown that binding with RNA plays a pivotal role in modulating the aggregation of TDP-43 RRMs, acting as either a facilitator or inhibitor of the process.<sup>29–31</sup> For instance, tight binding with UG-rich RNA aptamers prevents fibrillation of TDP-43 constructs containing the N-terminal domain and both RRMs, whereas weak binding with non-UG-rich aptamers can accelerate its aggregation.<sup>29</sup> The role of the RRMs in TDP-43 aggregation is further emphasized by studies showing that mutations disrupting RNA binding significantly enhance TDP-43 aggregation.<sup>32</sup>

All these findings suggest that stabilizing the RRM domains of TDP-43 is critical for maintaining the protein solubility and preventing its pathological aggregation.

In this study, we explored the potential of Epigallocatechin-3-gallate (EGCG), a polyphenol that constitutes more than 50% of the catechins found in green tea, to stabilize the RRM domains of TDP-43 and inhibit its aggregation. Several studies have reported that EGCG has neuroprotective effects for several neurodegenerative diseases,<sup>33–36</sup> and potently inhibits the formation of huntingtin, abeta, and alpha-synuclein aggregates by redirecting these proteins into non-toxic, unstructured, off-pathway oligomers.<sup>37–40</sup>

In the context of TDP-43 pathological aggregation, previous studies indicate that EGCG may preserve the native conformation of the protein full-length and its C-terminal domain, and prevents TDP-43 degradation and toxic aggregation. Moreover, EGCG-induced TDP-43 aggregates have been reported to be non-toxic.<sup>41,42</sup> However, the specific effects of EGCG on the RRM domains of TDP-43, which are central to its aggregation and function, remain largely unexplored.

Here, we investigate the inhibitory effects of EGCG on TDP-43 aggregation, focusing on its effect on the RRM domains. Using a combination of fluorescence assays, NMR spectroscopy, and computational modeling, we demonstrate that EGCG significantly delays the nucleation phase of TDP-43 aggregation and reduces the formation of amyloid-like aggregates. Our findings reveal that EGCG directly interacts with the RRM domains, stabilizing them in a non-aggregating state. Additionally, we highlight the importance of maintaining EGCG in its reduced and stable form during experiments, as its degradation can confound results. These insights extend our understanding of the mechanism of action of EGCG and its therapeutic potential in TDP-43 proteinopathies.

By elucidating the molecular basis of EGCG's interaction with TDP-43, we aim to contribute to the development of targeted therapies for ALS and related disorders. The findings presented in this study not only reinforce the therapeutic potential of EGCG but also emphasize the importance of understanding the structural and functional dynamics of TDP-43. The interplay between RNA binding, protein aggregation, and cellular stress responses remains a critical area of investigation, with implications for a wide range of neurodegenerative conditions. Future research should focus on integrating biophysical insights with translational studies to develop effective interventions for TDP-43-associated diseases.

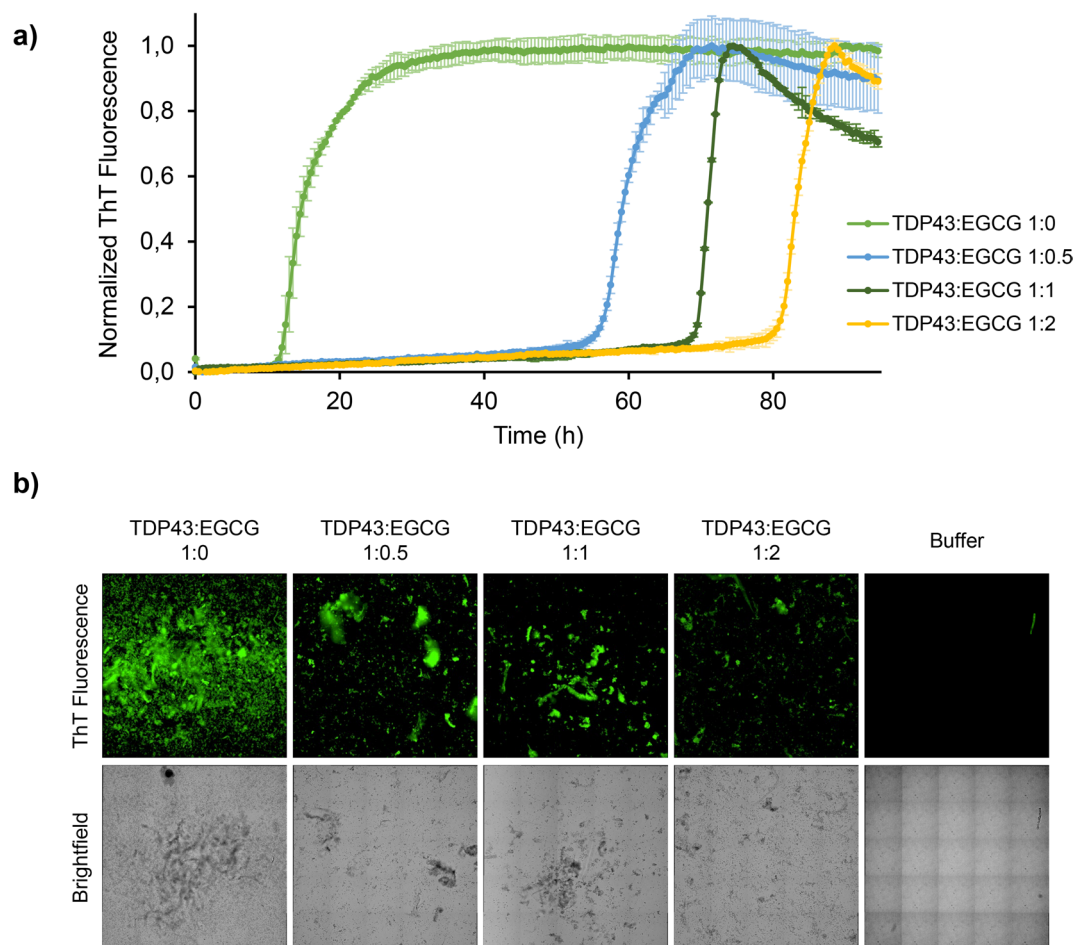
## Results

### EGCG inhibits *in vitro* TDP-43 RRMs aggregation

We first investigated whether EGCG has a modulatory effect on the formation of TDP-43 aggregates *in vitro*. To this purpose, a recombinant TDP-43 construct composed of both RRM1 and 2 (TDP-43 RRMs) was used in a Thioflavin T (ThT)-binding assay, ThT being widely used as a fluorescent probe for detecting amyloid-like aggregates.<sup>43</sup> The aggregation assay was performed in parallel with increasing concentrations of EGCG ranging from 0 to 30  $\mu$ M, corresponding to protein: EGCG ratios 1:0, 1:0.5, 1:1, and 1:2. Although literature reports that oxidized EGCG may enhance its anti-aggregation efficacy,<sup>44–46</sup> in our aggregation assay EGCG was kept in its reduced and stable form by supplementing the sample buffer with 1,4-Dithiothreitol (DTT). This decision was driven by the known instability of EGCG,<sup>47,48</sup> which we investigated further in our experimental conditions by a 1D NMR time course stability analysis. Specifically, we acquired NMR spectra over time on EGCG samples with and without the addition of a reducing agent, which revealed that EGCG rapidly degrades in the absence of a reducing agent while maintaining its reduced and stable form when in the presence of a reducing agent (Suppl. Fig. 1).

The ThT assay results provided compelling evidence that EGCG exerts a significant inhibitory effect on the aggregation of TDP43 RRMs *in vitro*. Both the control and EGCG-treated samples exhibited a characteristic sigmoidal ThT fluorescence curve, indicative of nucleation-mediated aggregation, which is typical for amyloid-forming proteins (Fig. 1a). However, marked differences were observed between the two conditions. Without EGCG, the lag phase of aggregation was approximately 10 h. In contrast, in the presence of EGCG, this phase was drastically extended as a function of the EGCG concentration, clearly indicating that EGCG interferes with a very early step in the amyloid formation pathway and suppresses the assembly of on-pathway amyloidogenic oligomers and protofibrils.<sup>49</sup> Additionally, we inspected the endpoints of the ThT aggregation kinetics using both brightfield and fluorescent microscopy. Large objects several micrometers in size were visible in all samples but reduced in number in the presence of EGCG. Moreover, ThT fluorescence intensity was substantially lower in the EGCG-treated samples, indicating that EGCG reduces the amyloid-like aggregates content (Fig. 1b). These results align with previous studies that have shown the ability of EGCG to modulate protein aggregation.<sup>35,40–42</sup>

Taken together, our findings prove an inhibitory effect of EGCG on TDP-43 RRMs aggregation.<sup>41,42</sup> Furthermore, our results highlighted a critical aspect often overlooked in EGCG research: the compound's stability under experimental conditions. Our 1D <sup>1</sup>H NMR analysis revealed that EGCG begins to degrade within 1 h in the absence of a reducing agent in the sample buffer. This evidence underscores the importance of maintaining EGCG in its reduced form to affect TDP-43 aggregation, as well as the importance of verifying the structure and stability of any component to assure the rigor of the investigation.



**Fig. 1.** Effect of EGCG on TDP-43 RRM aggregation. **(a)** Time evolution of the ThT fluorescence intensity for 15  $\mu$ M TDP-43 RRM samples incubated at 37  $^{\circ}$ C for  $\sim$  90 h with increasing concentration of EGCG ranging from 0 to 30  $\mu$ M, corresponding to protein:EGCG ratio 1:0, 1:0.5, 1:1, and 1:2. Each curve represents the mean of triplicate measurements normalized to the maximum reached value. Error bars in each curve represent the standard deviation of replicates. **(b)** Inspection of the ThT aggregation kinetics endpoints by fluorescent microscopy (top) and brightfield microscopy (bottom).

### EGCG exhibits a direct interaction with TDP-43 RRM

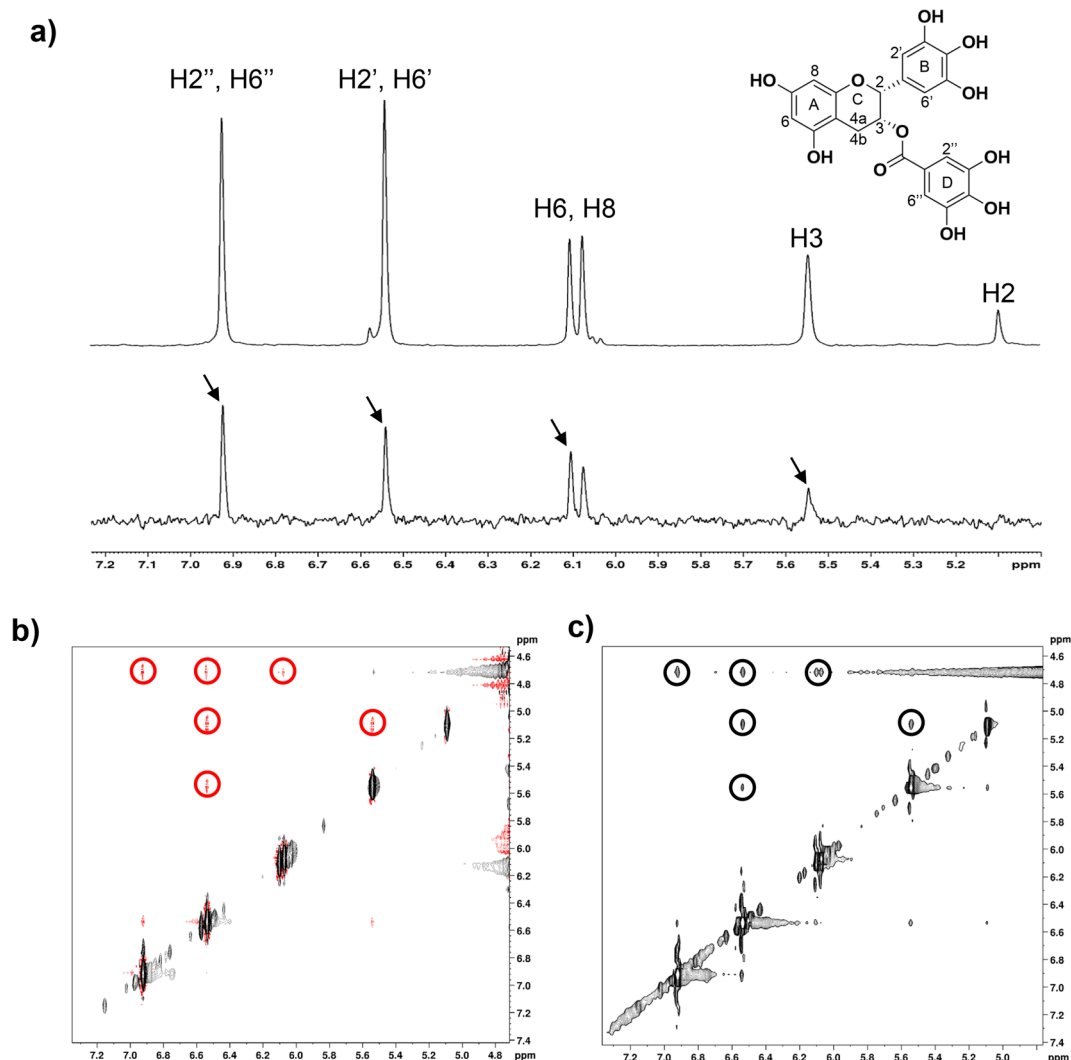
To investigate whether the inhibitory effect of EGCG on TDP-43 RRM aggregation is mediated by a direct interaction between the two molecules, we employed Saturation Transfer Difference (STD) and Transferred Nuclear Overhauser Enhancement (trNOE) NMR experiments. These well-established and versatile techniques are ideal for detecting weak protein-ligand interactions using small quantities of non-labeled proteins.<sup>50</sup>

The STD-NMR spectrum revealed signals from the ligand protons, indicating a direct binding event between EGCG and TDP-43 RRM (Fig. 2a).<sup>51</sup> Further validation was provided by acquiring a 2D NOE Spectroscopy (NOESY) NMR experiment on free EGCG, and a trNOE NMR experiment on EGCG in the presence of TDP-43 RRM. The 2D NOESY spectrum of free EGCG displayed the expected NOE pattern, with cross and diagonal peaks of opposite signs, characteristic of a small molecule in its free state. Conversely, in the trNOE spectrum of EGCG in the presence of TDP-43 RRM, the cross and diagonal peaks showed the same sign, consistent with the behavior expected for a ligand bound to a large molecule (Fig. 2b,c).<sup>52,53</sup>

Taken together, our results indicate a specific, although weak, interaction between EGCG and TDP-43 RRM. This interaction does not induce significant conformational changes in EGCG compared to its unbound form, and the primary sites of interaction on the ligand appear to be the aromatic rings. More investigation is needed to assess the contribution of the hydroxyl groups linked to the aromatic rings.

### 3D structural model of RNA-free TDP-43 RRM

To identify the EGCG binding site on TDP-43 RRM, we modelled the protein structure using the deep learning-based prediction tool RoseTTAFold.<sup>54</sup> This was necessary because the only available experimental structure of TDP-43 RRM is in complex with UG-rich RNA (PDB 4bs2).<sup>55</sup> In this structure, the relative orientation of the two RRM domains is stabilized by interactions with the RNA, resulting in a well-defined conformation. However, this RNA-bound conformation differs from the structure of the RNA-free protein in which the highly flexible



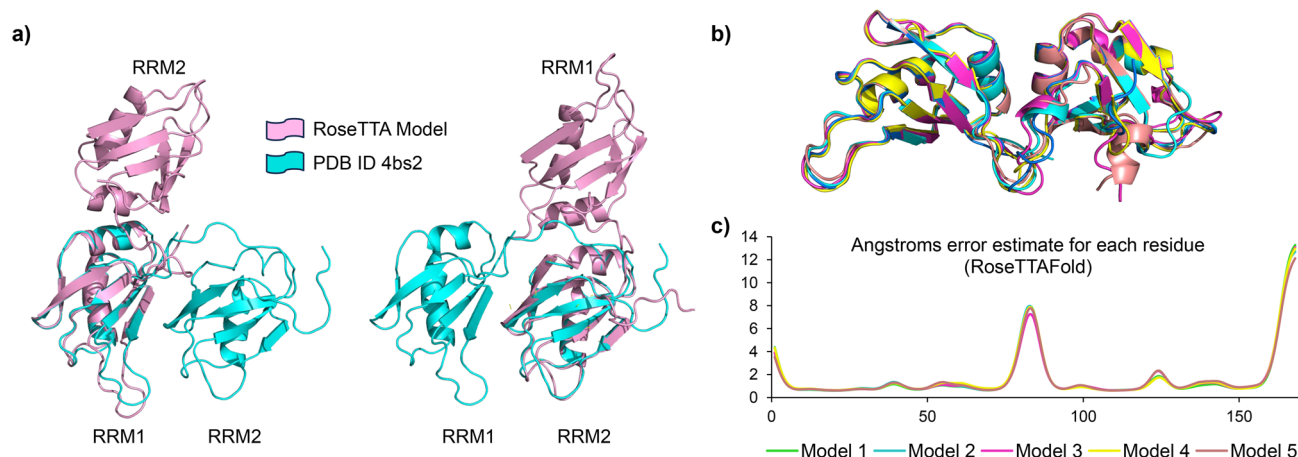
**Fig. 2.** Interaction between EGCG and TDP-43 RRM proved by NMR spectroscopy. **(a)** Comparison between the 1D  $^1\text{H}$  NMR reference spectrum of free EGCG (top) and STD spectrum of EGCG in the presence of TDP-43 RRM (bottom). The structure of EGCG is reported together with the assignment of the proton signals. The presence of EGCG signals in the STD spectrum (black arrows) indicates a binding between the ligand and the protein. **(b)** 2D NOESY spectrum of free EGCG. Cross (highlighted in red circles) and diagonal peaks have opposite signs as expected for a small molecule in a free state. **(c)** trNOE spectrum of EGCG in the presence of TDP-43 RRM. Cross (highlighted in black circles) and diagonal peaks have the same sign as expected for a ligand bound to a large molecule.

loop between the two RRM is thought to confer adaptability to different nucleic acid partners by allowing different orientations of the RRM domains.<sup>56,57</sup>

The RoseTTAFold model obtained for RNA-free TDP-43 RRM maintains the overall structure of the two RRM in the experimental RNA-bound structure, but differs for the mutual orientation of the two RRM, with the RoseTTAFold model displaying a more open conformation (Fig. 3a). The five best-scored models showed high confidence values (0.81) and consistent secondary structure elements, as well as a similar mutual orientation of the two RRM. Higher error values were obtained for the 14-amino acid linker between the two RRM domains, and in the C-terminal tail (Fig. 3b,c). Consistently, residues in these two regions showed lower  $I/I_0$  values in the  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear NOE NMR relaxation experiment, indicating that they experience movements in the nanosecond to picosecond timescale (Suppl. Fig. 2).<sup>58</sup>

To validate the RoseTTAFold model of RNA-free TDP-43 RRM, we measured the amide chemical shift temperature coefficients by recording a series of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra at increasing temperatures (Fig. 4a). This approach evaluates the solvent accessibility of amide protons, and thus their propensity to form inter- and intramolecular hydrogen bonds.<sup>59</sup> Temperature coefficient values lower than  $-5$  ppb/K, typically associated with solvent-exposed amide protons forming intermolecular hydrogen bonds with water, were mapped onto both the RoseTTAFold model and the experimental RNA-bound structure for comparison (Fig. 4b). RoseTTAFold model 1 was arbitrarily selected for this comparison. We found that the RoseTTAFold model is consistent with





**Fig. 3.** 3D structural model of free TDP-43 RRM1 and RRM2. **(a)** Overlap between the RoseTTAFold model 1 of RNA-free TDP-43 RRM1 and RRM2 (pink) and experimental RNA-bound structure (cyan, PDB 4bs2). The alignment was performed by overlapping, in turn, RRM1 (left) and RRM2 (right). The RoseTTAFold model and the experimental RNA-bound structure share the overall structure of the two RRM1 and RRM2. Nonetheless, these display a different mutual orientation. **(b)** Cartoon representation of the 5 best-scored models generated by RoseTTAFold, in overlap. **(c)** Angstrom error estimates versus the position of each residue in the sequence.

the experimental NMR data, as residues with temperature coefficient values lower than  $-5$  ppb/K were exposed to the solvent. In contrast, these same residues were located at the interface between the two RRM1 and RRM2 when mapped onto the RNA-bound experimental structure. These findings suggest that the RNA-free TDP-43 RRM1 and RRM2 adopts a more flexible and open conformation compared to its RNA-bound state, highlighting the dynamic nature of TDP-43 and its potential structural adaptability in the absence of RNA.

### Characterization of EGCG binding sites on TDP-43 RRM1 and RRM2

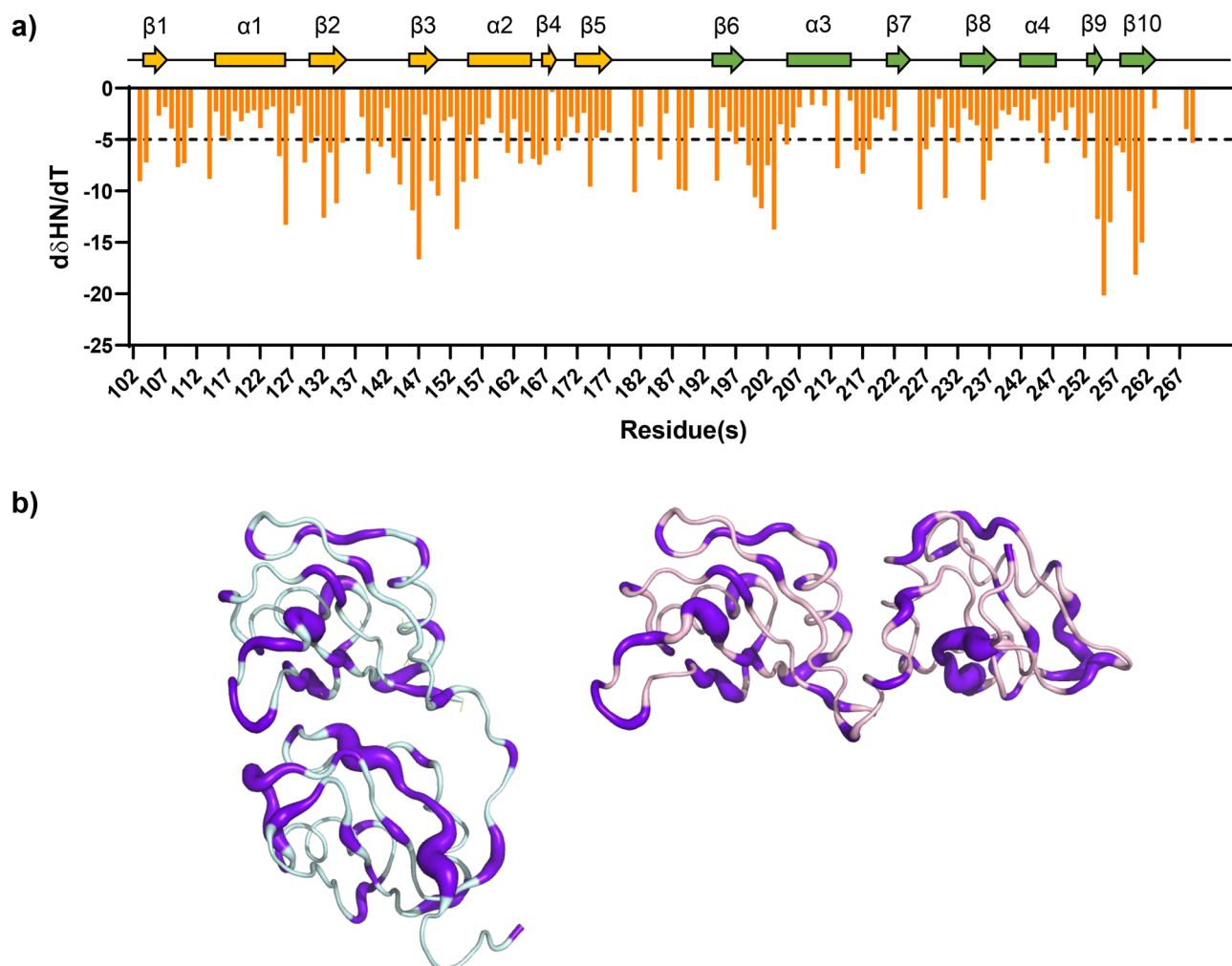
To identify the EGCG binding site on TDP-43 RRM1 and RRM2, we performed Chemical Shift Perturbation (CSP) NMR analysis<sup>60</sup>.  $^{15}\text{N}$ -labeled TDP-43 RRM1 and RRM2 was titrated with increasing concentrations of EGCG, and 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra were recorded for each titration point to track resonance shifts (Fig. 5a). Several peaks shifted smoothly as a function of EGCG concentration, indicating a fast exchange regime between the free and bound states (Fig. 5b). The observed CSPs were quantified, and residues with significant shifts, indicative of potential binding events, were mapped onto the RoseTTAFold model of TDP-43 RRM1 and RRM2 (Fig. 5c,d). Notably, the most substantial shifts were observed for residues localized in two defined regions, each localized on one of the two RRM1 and RRM2, suggesting a protein: ligand stoichiometry of 1:2. Additional CSPs were observed in the flexible 14-amino acid linker. However, these were likely due to conformational changes rather than direct ligand binding, as suggested by  $^{15}\text{N}$  backbone amide relaxation parameters measured after the CSP experiments (Suppl. Fig. 3).

CSP-guided docking calculations by the HADDOCK suite<sup>60,61</sup> were then performed imposing a protein:ligand stoichiometry of 1:2 and binding restraints only for those residues laying in structured regions. The HADDOCK output generated 190 structures grouped into eight clusters, with Clusters 1 and 2 being the most populated and having similar HADDOCK scores:  $-54.6 \pm 1.8$  and  $-54.9 \pm 2.0$ , respectively (Table 1). Both clusters displayed low restraint violation energies, indicating that the CSP-derived restraints were well accommodated. Superimposition of the best structures of each cluster revealed similar EGCG binding sites on RRM1 and RRM2, with recurrent interactions between the proteins and ligands (Table 2). However, no well-defined binding pockets for EGCG were observed, allowing the ligand to adopt slightly different binding poses across simulations (Fig. 6).

To assess the stability of the docked structures, we conducted two separate all-atom molecular dynamics (MD) simulations using GROMACS<sup>62</sup> one for each top-scoring structure from Clusters 1 and 2. Both simulations confirmed the persistence of EGCG binding to the protein throughout the trajectory. The Root Mean Square Deviation (RMSD) analysis of the protein in complex with EGCG revealed fluctuations, primarily due to the dynamic nature of the 14-amino acid linker between the two RRM1 and RRM2, which influences their relative orientation. Fluctuations were also observed when looking at the RMSDs of the two EGCG molecules bound to the protein, as expected when a well-defined binding pocket is missing (Fig. 7 and Suppl. Fig. 4).

To pinpoint the protein residues most involved in binding, we analyzed the frequency of hydrogen bonds and the presence of  $\pi$ - $\pi$  or cation- $\pi$  interactions. The analysis revealed that EGCG engages more stably with RRM1 than RRM2. In RRM1, 10 hydrogen bonds were formed during the simulation, with residues L111, S144, D174, and D185 being the most involved, showing occupancies of 61.57%, 13.21%, 42.38%, and 78.59%, respectively. Additionally, F147 formed a transient  $\pi$ - $\pi$  interaction with EGCG's aromatic rings A and D. In RRM2, the total number of hydrogen bonds fluctuated around 5 throughout the simulation, with residues D247 and K263 having the highest occupancies at 7.21% and 21.98%, respectively. Aromatic side chains of residues F194, F231, and H264 were instead involved in  $\pi$ - $\pi$  interactions with the EGCG aromatic rings A and B.

These findings suggest that, while EGCG binds to both RRM1 and RRM2 of TDP-43, the interaction is more stable with RRM1. Moreover, the structural stability of the protein-ligand complex and the overall binding



**Fig. 4.** Amide chemical shift temperature coefficients analysis. **(a)** Temperature coefficient values versus the position of each residue in the sequence. Values lower than  $-5$  ppb/K (threshold in dotted line), are typically associated with solvent-exposed amide protons forming intermolecular hydrogen bonds with water. **(b)** Worm representation of RNA-bound structure (cyan) and RoseTTAFold model 1 of RNA-free TDP-43 RRM (pink). Residues experiencing temperature coefficient values lower than  $-5$  ppb/K are indicated in purple in both structures. The thickness of the purple traits linearly correlates to the  $d\delta_{HN}/dT$  value.

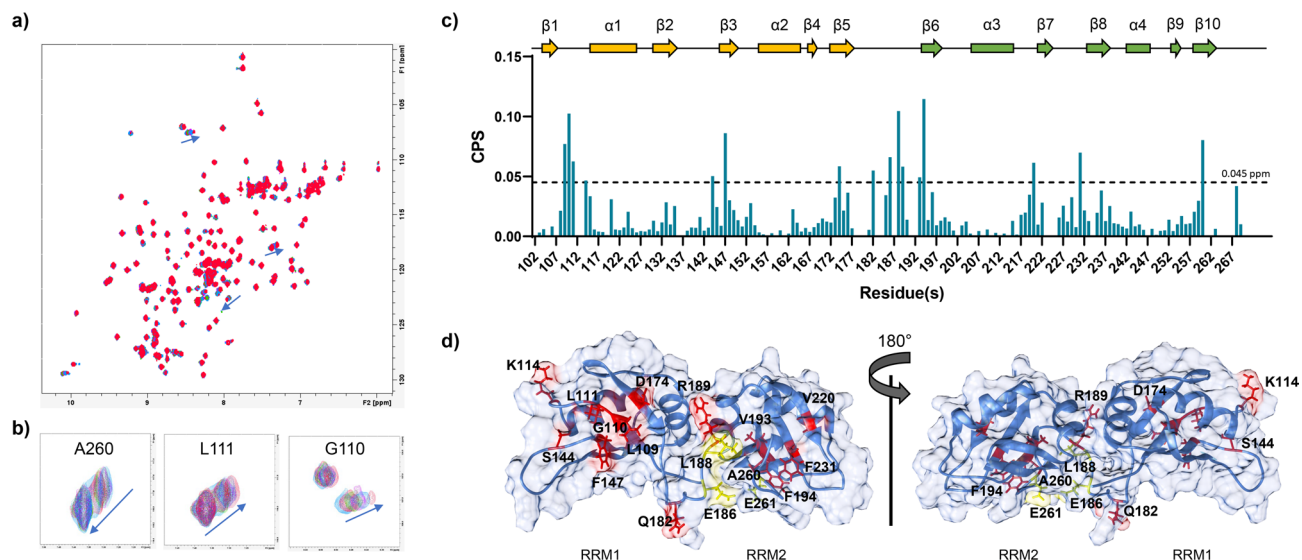
dynamics are influenced by the flexibility of the 14-amino acid linker, which plays a significant role in modulating the relative orientation of the two RRMs.

## Discussion

The present study explores the inhibitory effect of EGCG on the aggregation of TDP-43, a protein critically implicated in the pathogenesis of ALS and other neurodegenerative diseases. Despite extensive research, the precise mechanisms driving TDP-43 aggregation and its toxic effects remain incompletely understood. However, the involvement of specific domains of TDP-43, such as its RRM domains, has been well documented.<sup>21–27,29,32</sup> These RRMs are critical for the protein's interaction with RNA and are also implicated in its aggregation propensity. Consequently, understanding how small molecules like EGCG can modulate the behavior of these domains offers valuable insights into potential therapeutic approaches for ALS and other TDP-43-related pathological conditions.

On the other hand, EGCG has attracted considerable attention as a phytochemical showing promise in modulating protein aggregation toxicity.<sup>33–36,63</sup> The anti-aggregation properties of EGCG are thought to stem from its ability to stabilize proteins in their native or soluble oligomeric states, thereby preventing the formation of toxic aggregates. Moreover, EGCG has been reported to remodel existing aggregates into non-toxic forms, highlighting its potential not only as a preventive but also as a therapeutic agent.<sup>37–40,44</sup>

In the context of TDP-43 self-assembly, previous research has shown that EGCG can modulate the aggregation of both full-length and C-terminal domain of TDP-43.<sup>41,42</sup> However, the specific effects of EGCG on the RRMs of TDP-43, which are involved in the protein's aggregation and function, have not been thoroughly explored until now. Our study provides new insights into how EGCG interacts with these domains and the implications



**Fig. 5.** Chemical shift perturbation analysis. **(a)** Overlay of 2D  $^1\text{H}$ - $^{15}\text{N}$ -HSQC spectra of TDP-43 RRM1 upon EGCG titration. **(b)** Highlight of representative  $^1\text{H}$ - $^{15}\text{N}$ -HSQC peaks exhibiting significant perturbation upon EGCG titration. **(c)** Chemical shift perturbation values for each amino acidic residue of TDP-43 RRM1. **(d)** Cartoon representation of the RoseTTAFold model 1 of TDP-43 RRM1. Highlighted in red those residues experiencing CSP > 0.045 ppm, while in yellow those residues for which the correspondent peak in the  $^1\text{H}$ - $^{15}\text{N}$ -HSQC seems to disappear upon EGCG titration.

	Cluster 1	Cluster 2
Haddock score	$-54.6 \pm 1.8$	$-54.9 \pm 2.0$
Cluster size	91	48
RMSD from overall lowest energy structure	$0.4 \pm 0.0$	$0.3 \pm 0.1$
Van der Waals energy	$-47.3 \pm 2.6$	$-48.2 \pm 2.3$
Electrostatic energy	$-29.1 \pm 18.9$	$-19.4 \pm 24.5$
Desolvation energy	$-6.0 \pm 3.4$	$-5 \pm 3.2$
Restraints violation energy	$16.4 \pm 15.04$	$2.4 \pm 1.29$
Z-score	-1.0	-1.1

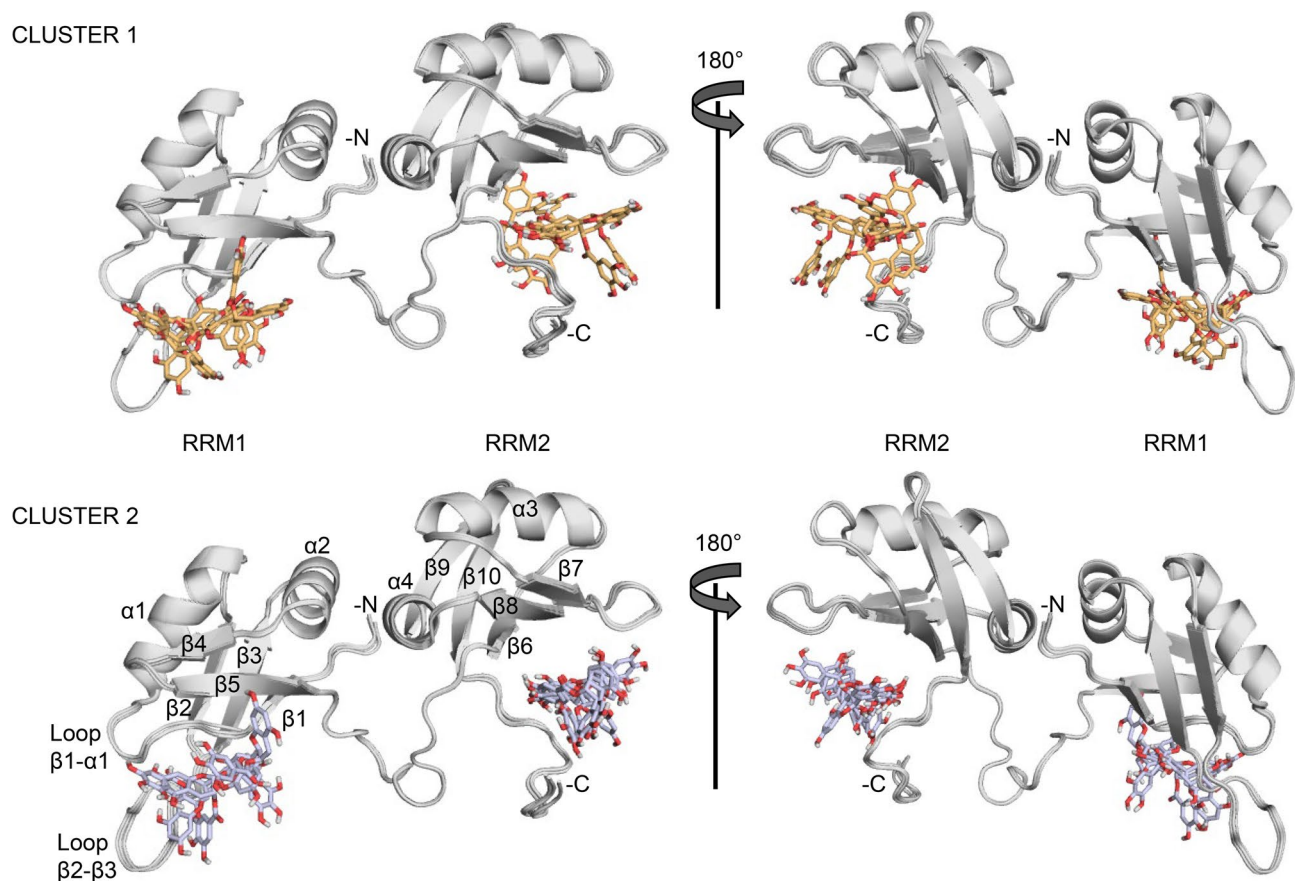
**Table 1.** HADDOCK outputs of the most populated clusters obtained by TDP-43 RRM1/EGCG docking.

for TDP-43 aggregation. More specifically, our study revealed that EGCG exerts a significant inhibitory effect on the aggregation of TDP-43 RRM1 *in vitro*. Using a ThT-binding assay, we demonstrated that EGCG considerably delays the nucleation phase of aggregation and reduces the overall formation of ThT-positive aggregates. The extension of the lag phase and the reduction of the maximum fluorescence intensity observed in our ThT assay suggest that EGCG interferes with the early stages of TDP-43 aggregation, potentially by binding to the protein and stabilizing it in a non-amyloidogenic form. Accordingly, our study revealed a direct albeit weak interaction between EGCG and TDP-43 RRM1.

The 3D structural analysis of the RNA-free TDP-43 RRM1, predicted using RoseTTAFold, provided further context for the observed interactions. The model revealed that TDP-43 RRM1 adopts a more open and flexible conformation in the absence of RNA than in the RNA-bound state. This flexibility is particularly pronounced in the 14-amino acid linker between RRM1 and RRM2, which is likely to influence the overall dynamics of the protein and its interaction with ligands like EGCG. Our docking studies and molecular dynamics simulations supported this notion, showing that EGCG can bind to both RRM1 and RRM2, with a preference for RRM1. This suggests that RRM1 may play a more crucial role in the aggregation process of TDP-43, and that targeting this domain with small molecules like EGCG could be particularly effective in preventing TDP-43 aggregation. Given that RRM1 plays a key role in RNA recognition, this raises the question of whether EGCG competes with RNA for binding. Our findings suggest that while some overlap exists between EGCG binding sites and RNA-binding regions, direct competition is highly unlikely. The dynamic nature of the TDP-43/EGCG interaction, combined with the flexible conformation of the 14-amino acid linker between RRM1 and RRM2, allows for multiple binding poses without necessarily displacing RNA. This observation aligns with previous studies showing that EGCG can engage structurally distinct regions of proteins without fully occluding their functional sites.<sup>64</sup> The dynamic nature of the linker thus emerged as a critical factor influencing the binding of EGCG and the structural stability of the protein-ligand complex. The linker's flexibility allows the two RRM1 to adopt

	Cluster 1	Cluster 2
RRM1	Hydrogen bond G110 (HN) L111 (HN) W113 (NH) H143 (NH) S144 (C=O) K145 (C=O) G146 (C=O) R165 (NH2) R171 (NH2) D174 (C=O) $\pi$ -cation R171	Hydrogen bond G110 (C=O) L111(C=O) (HN) H143 (NH) S144 (C=O) (OH) K145 (C=O) G146 (C=O) R171 (NH2) D174 (C=O) $\pi$ -cation R171
RRM2	Hydrogen bond N259 (C=O) E261 (C=O) (HN) K263 (NH3) H264 (NH) $\pi$ -cation K263 $\pi$ - $\pi$ F194 F231	Hydrogen bond K192 (NH3) E261 (C=O) (HN) K263 (HN) (NH3) H264 (NH) R268 (NH2)

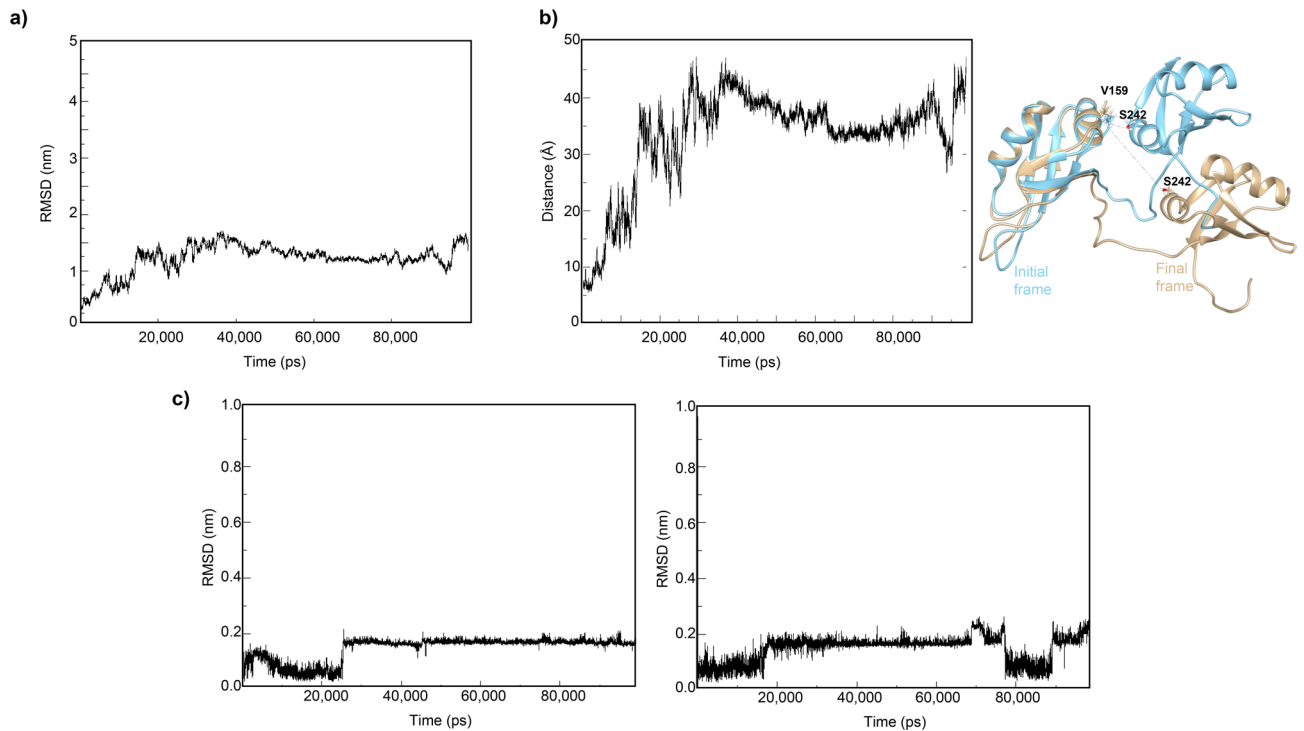
**Table 2.** Recurrent interactions observed in HADDOCK Clusters 1 and 2.



**Fig. 6.** Docking models. Cartoon superimposition of 4 of the most representative models from both HADDOCK Cluster 1 (top) and 2 (bottom). The two clusters show similar EGCG binding sites on RRM1 and RRM2. However, no well-defined binding pockets for EGCG are observed, allowing the ligand to adopt slightly different binding poses across simulations.

various relative orientations, which in turn affects how EGCG interacts with the protein. This observation is important because it highlights the potential for allosteric effects in the regulation of TDP-43 aggregation. By modulating the conformation of the linker, it may be possible to influence the overall aggregation behavior of TDP-43, providing a novel therapeutic strategy.





**Fig. 7.** Parameters from molecular dynamics simulation on HADDOCK Cluster 1. **(a)** RMSD for the protein structure over 100ns of simulation from the starting HADDOCK model. **(b)** Distance between Ca of V159 and Ca of S242 measured along the simulation. The two RRM domains assume different orientations with respect to the initial frame, moving away from each other. This distance correlated with the RMSD fluctuations observed in **(a)**. **(c)** RMSD along the 100ns simulation of EGCG molecule bound to RRM1 (left) and RRM2 (right).

Our findings reinforce the potential of EGCG as a therapeutic agent in neurodegenerative diseases. By demonstrating that EGCG can directly interact with and inhibit the aggregation of TDP-43 RRMs, we provide a mechanistic basis for its neuroprotective effects observed in other models. This adds up to the growing body of evidence supporting the use of EGCG and other polyphenolic compounds in neurodegenerative disease therapy.

A key question in this study concerns how EGCG, a single small molecule, can engage with structurally distinct regions of TDP-43, namely the structured RRMs and the intrinsically disordered C-terminal domain. EGCG is known for its promiscuous binding properties, primarily mediated by  $\pi$ - $\pi$  stacking and hydrogen bonding with aromatic residues, which can easily be engaged in both structured RRMs and the disordered C-terminal domain despite their structural differences. This mode of binding is not uncommon for amyloid-modulating small molecules. EGCG has been shown to bind proteins differently depending on their conformational state: in structured proteins, it preferentially interacts with dynamic, exposed loops, while in intrinsically disordered proteins, it binds more diffusely along the sequence<sup>64</sup>. This dual interaction model provides a plausible explanation for EGCG's ability to engage multiple regions of TDP-43.

Our study highlights the importance of targeting specific domains of TDP-43 in therapeutic development. The differential binding of EGCG to RRM1 and RRM2 suggests that these domains may contribute to TDP-43 aggregation in distinct ways, confirming previous evidence.<sup>27</sup> Therefore, therapies that selectively target RRM1 could potentially be more effective in preventing or reversing TDP-43 aggregation. Additionally, the role of the 14-amino acid linker in modulating the overall dynamics of TDP-43 suggests that this region could be another promising target for therapeutic intervention. Compounds that stabilize the linker in a specific conformation could alter the aggregation propensity of TDP-43, providing a new approach to disease treatment. The open and flexible conformation of the RNA-free TDP-43 RRMs confirms that the protein is highly dynamic and capable of adopting multiple conformations depending on its interaction partners. This flexibility may be a key factor in its aggregation behavior and its ability to form different types of aggregates under different pathological conditions.

Additionally, we highlight the importance of maintaining EGCG in its reduced and stable form during experiments, as its degradation can confound results. Adding any component without a clear understanding of its structure and stability can completely undermine the rigor of the investigation. Therefore, ensuring that EGCG remains in its stable and reduced form allowed us to draw reliable conclusions about its effects on TDP-43 aggregation. These insights extend our understanding of the mechanism of action of EGCG and its therapeutic potential in TDP-43 proteinopathies.

Future research should focus on bridging the gap between *in vitro* findings and clinical applications to better understand the therapeutic potential of EGCG. The development of novel EGCG derivatives with enhanced binding affinity and specificity for TDP-43 RRMs could also be a promising avenue for therapeutic development.

Finally, understanding the interplay between TDP-43 aggregation, RNA binding, and cellular stress responses will be crucial for developing comprehensive therapeutic strategies for ALS and related disorders.

## Methods

### Production of Recombinant TDP-43 RRM<sub>s</sub>

In the present study we used a recombinant TDP-43 deletion mutant comprising residues K102–Q269, which includes both RRM1 and RRM2 (TDP-43 RRM<sub>s</sub>). The recombinant protein was expressed and purified as previously described with minor adjustments.<sup>27</sup> Briefly, TDP-43 RRM<sub>s</sub> was expressed in *E. coli* Rosetta2(DE3) cells as a fusion protein with a TEV cleavable N-terminal SUMO solubilization domain carrying a 6xHis tag. Protein expression was induced by adding 0.5 mM Isopropyl-β-D-1-thiogalactopyranoside, and carried out overnight at 18 °C. Cells were then collected by centrifugation and resuspended in lysis buffer (10 mM potassium phosphate buffer pH 7.2, 150 mM KCl, 5% v/v glycerol, 1 mg/ml lysozyme, Roche complete™ EDTA-free Protease Inhibitor, 5 mM MgCl<sub>2</sub>, 1 μg/ml DNase I, and 1 μg/ml RNaseA). Cells disruption was performed by 1 h incubation at RT followed by probe-sonication. Soluble 6xHis-SUMO tagged TDP-43 RRM<sub>s</sub> was recovered in the supernatant of the lysate by centrifugation at 70,000 rcf for 45 min at 4 °C. Isolation of the target protein was performed by immobilized metal ion affinity chromatography loading the clarified lysate onto a Cytiva 5 ml pre-packed Ni Sepharose affinity column for capturing. Protein elution was performed by a 0–500 mM imidazole linear gradient. The 6xHis-SUMO protein was then cleaved by incubation with TEV protease (1:2 TEV: Protein molar ratio) for 2 h at RT, followed by overnight at 4 °C, while dialyzing the mixture against 10 mM potassium phosphate buffer pH 7.2, 150 mM KCl. The cleaved protein was recovered into the flow-through of a second nickel affinity chromatography column, and loaded onto a Cytiva HiTrap Heparin column to remove unwanted nucleic acids. The protein was eluted with potassium phosphate buffer with 1.5 M KCl, and submitted to size-exclusion chromatography using a HiLoad 16/60 Superdex 75 prep grade pre-equilibrated in 10 mM potassium phosphate buffer pH 7.2, 150 mM KCl. Protein identity and purity were checked by SDS-PAGE analysis.

For <sup>15</sup>N-labelled TDP-43 RRM<sub>s</sub> samples, cell growth and protein expression were conducted in minimal media supplemented with <sup>15</sup>N-NH<sub>4</sub>Cl as sole source of nitrogen.

### Thioflavin T aggregation assay

TDP-43 RRM<sub>s</sub> aggregation kinetics was monitored by Thioflavin T (ThT, Sigma-Aldrich T3516) fluorescence emission.<sup>65–67</sup> ThT fluorescence intensity measurements were performed at 37 °C on protein samples concentrated 15 μM, in 20 mM Hepes buffer pH 7.5 supplemented with 150 mM NaCl, 1mM Dithiothreitol (DTT) and 26 μM ThT. All samples underwent Size Exclusion Chromatography (SEC) immediately before the ThT aggregation assay, to avoid the impact of possible seeds of aggregation and to ensure sample quality. Measurements were performed at least in triplicate under constant agitation at 500 rpm using a CLARIOstar Plus MicroPlate reader (BMG LabTech), incubating the samples in optically clear-bottomed 96-well black plates (Perkin Elmer, 6005225), sealed with a clear adhesive to prevent samples evaporation (Perkin Elmer, 6050185). ThT emission was measured up to 90 h every 30 min at 480 nm under excitation at 440 nm, from the plate bottom. EGCG was added to the protein samples from stock solution at concentrations ranging from 0 to 30 μM. The stock solution was prepared by dissolving EGCG as powder (Sigma-Aldrich, E4143) in 20 mM Hepes buffer pH 7.5, 150 mM NaCl, 1mM DTT, and filtering the solution with 0.22 μm cellulose acetate syringe filters.

### Microscopy images acquisition

Endpoints from ThT aggregation assay were transferred to black, optically clear flat-bottom, 96-well PhenoPlate plates (PerkinElmer, 6055700). Images, both in brightfield and fluorescence mode, were acquired using an Operetta CLS High content screening microscope with a 40x objective. A minimum of 25 fields were acquired per well.

### Ligand-based NMR measurements

Saturation Transfer Difference (STD) and Transferred Nuclear Overhauser Enhancement (trNOE) NMR experiments were acquired at 298 K on a Bruker AVANCE 800 MHz NMR spectrometer equipped with cryoprobe.

In the STD-NMR analysis, a 1D <sup>1</sup>H NMR reference spectrum was first recorded on a sample containing unlabeled TDP-43 RRM<sub>s</sub> with a 100-fold excess of EGCG. A second 1D <sup>1</sup>H NMR spectrum was acquired selectively saturating the protein signals for 3 s, on resonance at 0.55 ppm and off resonance at –40 ppm (saturated spectrum). A spin-lock filter of 40ms was applied to eliminate the protein signals. Subtracting the saturated spectrum from the reference spectrum produced the STD-NMR spectrum.

The NMR experiments 2D NOESY on free EGCG, and trNOE on EGCG in the presence of TDP-43 RRM<sub>s</sub> were acquired by using standard pulse sequence from Bruker. Mixing time was set at 600 ms and 100 ms for free EGCG and EGCG/protein, respectively.

All spectra were recorded in 10 mM potassium phosphate buffer pH 7.2 and 150 mM KCl supplemented with 0.5 mM of tris(2-carboxyethyl)phosphine (TCEP), and analyzed by Bruker Topspin software.

### Generation of RNA-free TDP-43 RRM<sub>s</sub> structural model

The RoseTTAFold suite was exploited to quickly and accurately predict the three-dimensional structure of RNA-free TDP-43 RRM<sub>s</sub>.<sup>54</sup> Through a “three-track” neural network, this web-based tool is able to process sequence, distance, and coordinate information simultaneously, resulting in the generation of accurate structural models. In the case of RNA-free TDP-43 RRM<sub>s</sub>, the tool generated five best-scored models with a high confidence value of 0.81.

### NMR backbone assignment

To perform the NMR resonance assignment of the backbone atoms of RNA-free TDP-43 RRM1s, 2D  $^{15}\text{N}$ -HSQC, 3D  $^{15}\text{N}$ -NOESY-HSQC, and 3D  $^{15}\text{N}$ -TOCSY-HSQC NMR experiments were acquired at 298 K using a Bruker AVANCE 800 MHz NMR spectrometer equipped with a cryoprobe. All spectra were recorded on a sample of 0.2 mM  $^{15}\text{N}$ -labelled TDP-43 RRM1s in 10 mM potassium phosphate buffer pH 7.2, 150 mM KCl, and 0.5 mM TCEP. Data were processed using NMRpipe<sup>68</sup> and analyzed by CcpNmr software.<sup>69</sup>

The assignment at pH 6.5 already available on Biological Magnetic Resonance Bank (BMRB, ID 27613),<sup>70</sup> was used as starting point. About 70% of the amide resonances were transferred by visual inspection. The remaining resonances were assigned by analysis of the 3D experiments.

### NMR thermal coefficient measurements

Amide chemical shift temperature coefficients were calculated for 0.2 mM of  $^{15}\text{N}$ -labelled TDP-43 RRM1s in 10 mM potassium phosphate buffer pH 7.2, 150 mM KCl, 0.5 mM TCEP, and 10%  $\text{D}_2\text{O}$ . A series of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra were recorded at 283, 288, 293, 298, and 303 K using a Bruker AVANCE 800 MHz NMR spectrometer equipped with a cryoprobe. All spectra were referenced to the water signal for each temperature, processed using Bruker Topspin suite, and analyzed by CcpNmr software.<sup>69</sup>

The chemical shift values  $\delta_{\text{HN}}$  of all residues were plotted as a function of the temperature. The data were analyzed assuming that residues with  $d\delta_{\text{HN}}/dT < -5$  ppb/K form weaker H-bonds and should be considered secondary structure breakpoints or, if in an unstructured region, as promoters of H-bonds with water. Values of  $d\delta_{\text{HN}}/dT > -5$  ppb/K correspond to the formation of tighter bonds likely involved in secondary structure elements.

### NMR chemical shift perturbation experiments

To identify the EGCG binding surface on TDP-43 RRM1s, a series of 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra were recorded on 0.2 mM  $^{15}\text{N}$ -labelled TDP-43 RRM1s titrated with increasing concentration of EGCG (used ligand:protein molar ratios 0.0, 0.15, 0.25, 0.35, 0.5, 0.75, 1.0, 1.25, 1.5, 2.0). Chemical Shift Perturbation (CSP) was followed and calculated using the interactive CSP analysis module of CcpNmr software.<sup>69</sup> Peaks in overlap were excluded from the analysis. The threshold value was set equal to  $2\sigma$ , indicated as a standard practice.<sup>71</sup>

### Relaxation heteroNOE NMR experiments

Relaxation hetNOE NMR experiments were acquired on 0.2 mM TDP-43 RRM1s in 10 mM potassium phosphate buffer pH 7.2, 150 mM KCl, 0.5 mM TCEP, and 10%  $\text{D}_2\text{O}$ , both in the absence of EGCG and in the presence of 2-fold excess of EGCG. For each sample, two bi-dimensional  $^{15}\text{N}$  spectra were recorded: a reference spectrum and a spectrum recorded with an initial proton saturation of 8s. The NOE values were then determined from the ratios of the average intensities of the peaks with and without proton saturation. The data were analyzed by the Dynamic Center tool provided by Bruker topspin platform.

### Molecular docking

The TDP-43 RRM1s model 1 predicted by RoseTTAFold was arbitrarily selected to perform experimentally-driven 3-body-docking calculation with two molecules of EGCG (PubChem ID 65064). The docking was performed on the integrative modelling platform HADDOCK2.4.2.<sup>60,61</sup> The simulation was launched by setting as active those residues with a CSP value higher than  $2\sigma$  (*vide infra*) and forming a contiguous surface on each RRM domain. More specifically, residues L109, G110, L111, F147 and D174 laying on RRM1 were specified to interact with one molecule of EGCG, while V193, F194, V220, F231, A260 laying on RRM2 were specified to interact with a second molecule of EGCG. Two sets of experimental AIR restraints were introduced in TBL format, generated using HADDOCK server tool <http://alcazar.science.uu.nl/services/GentBL/>. The restraints were implemented at different stages of the calculations. During the rigid-body docking, the two CSP-guided binding regions on the protein and both the ligand molecules were selected as active. It followed the flexible refinement stage, where the two CSP-guided binding regions on the protein were turned inactive. After a third stage of water refinement, the RMSD was implemented as clustering method with a cut-off of 2.0 Å. For each pose, residues located at a maximum distance of 3 Å from EGCG structure were analysed using PyMOL suite.<sup>72</sup>

### Molecular dynamic simulation

Classical molecular dynamics (MD) simulations were performed with the GROMACS 2022 software package (<https://doi.org/10.5281/zenodo.6103835>), using CHARMM36 all-atom force field to generate the protein topology.<sup>73</sup> The ligand topology was obtained through a web tool based on the same force field <https://cgenff.com/>. The best structure of selected HADDOCK clusters 1 and 2 were used to build the systems. Structures were solvated in a cubic box with explicit water molecules (TIP3P). Counterions were added to neutralize the charge of the system. The systems were subjected to two equilibration steps after a short minimization using the steepest descent algorithm with a force convergence criterion set to 1000 kJ/mol-nm<sup>2</sup>. The first step was carried out to equilibrate and it was performed by applying positional restraints on the protein and the ligands. The second step was carried out to equilibrate the pressure and it was performed by coupling the systems with the Parrinello-Rahman barostat.<sup>74,75</sup> The production run was performed at constant pressure (1 atm) and temperature (300 K) for 100 ns using periodic boundary conditions and the velocity-rescaling thermostat.<sup>76</sup> LINCS algorithm<sup>77</sup> was used to constrain the bonds involving hydrogen atoms, while the particle mesh Ewald method<sup>78</sup> was used to account for long-range electrostatic interactions with a cutoff of 10 Å. An integration time step of 2 fs was used in all simulations. All the trajectories were inspected and analysed with Gromacs utilities. Root-mean-square deviation (RMSD) and root-mean-square fluctuations (RMSF) were measured with tools implemented in GROMACS 2022. Hydrogen bond occupancy and distances were measured using VMD.<sup>79</sup>

## Data availability

All data are available from the corresponding author on reasonable request.

Received: 9 January 2025; Accepted: 9 May 2025

Published online: 23 May 2025

## References

- Longinetti, E. & Fang, F. Epidemiology of amyotrophic lateral sclerosis: an update of recent literature. *Curr. Opin. Neurol.* **32**, 771–776 (2019).
- Hardiman, O. et al. Amyotrophic lateral sclerosis. *Nat. Rev. Dis. Primers.* **3**, 17085 (2017).
- Arai, T. et al. TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem. Biophys. Res. Commun.* **351**, 602–611 (2006).
- Mackenzie, I. R. A. & Rademakers, R. The role of transactive response DNA-binding protein-43 in amyotrophic lateral sclerosis and frontotemporal dementia. *Curr. Opin. Neurol.* **21**, 693–700 (2008).
- Neumann, M. et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**, 130–133 (2006).
- Van Deerlin, V. M. et al. TARDBP mutations in amyotrophic lateral sclerosis with TDP-43 neuropathology: a genetic and histopathological analysis. *Lancet Neurol.* **7**, 409–416 (2008).
- Pesiridis, G. S., Lee, V. M. Y. & Trojanowski, J. Q. Mutations in TDP-43 link glycine-rich domain functions to amyotrophic lateral sclerosis. *Hum. Mol. Genet.* **18**, R156–162 (2009).
- Sreedharan, J. et al. TDP-43 mutations in Familial and sporadic amyotrophic lateral sclerosis. *Science* **319**, 1668–1672 (2008).
- Gitcho, M. A. et al. TDP-43 A315T mutation in Familial motor neuron disease. *Ann. Neurol.* **63**, 535–538 (2008).
- Buratti, E. Functional significance of TDP-43 mutations in disease. *Adv. Genet.* **91**, 1–53 (2015).
- Chen-Plotkin, A. S., Lee, V. M. Y. & Trojanowski, J. Q. TAR DNA-binding protein 43 in neurodegenerative disease. *Nat. Rev. Neurol.* **6**, 211–220 (2010).
- Da Cruz, S. & Cleveland, D. W. Understanding the role of TDP-43 and FUS/TLS in ALS and beyond. *Curr. Opin. Neurobiol.* **21**, 904–919 (2011).
- Karch, C. M. et al. Selective genetic overlap between amyotrophic lateral sclerosis and diseases of the frontotemporal dementia spectrum. *JAMA Neurol.* **75**, 860–875 (2018).
- Prasad, A., Bharathi, V., Sivalingam, V., Girdhar, A. & Patel, B. K. Molecular mechanisms of TDP-43 misfolding and pathology in amyotrophic lateral sclerosis. *Front. Mol. Neurosci.* **12**, 25 (2019).
- Johnson, B. S. et al. TDP-43 is intrinsically aggregation-prone, and amyotrophic lateral sclerosis-linked mutations accelerate aggregation and increase toxicity. *J. Biol. Chem.* **284**, 20329–20339 (2009).
- François-Moutal, L. et al. Structural insights into TDP-43 and effects of Post-translational modifications. *Front. Mol. Neurosci.* **12**, 301 (2019).
- 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).
- Ling, S. C., Polymenidou, M. & Cleveland, D. W. Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. *Neuron* **79**, 416–438 (2013).
- Hergesheimer, R. C. et al. The debated toxic role of aggregated TDP-43 in amyotrophic lateral sclerosis: a resolution in sight? *Brain* **142**, 1176–1194 (2019).
- Wolozin, B. & Ivanov, P. Stress granules and neurodegeneration. *Nat. Rev. Neurosci.* **20**, 649–666 (2019).
- Johnson, B. S., McCaffery, J. M., Lindquist, S. & Gitler, A. D. A yeast TDP-43 proteinopathy model: exploring the molecular determinants of TDP-43 aggregation and cellular toxicity. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 6439–6444 (2008).
- Wang, Y. T. et al. The truncated C-terminal RNA recognition motif of TDP-43 protein plays a key role in forming proteinaceous aggregates. *J. Biol. Chem.* **288**, 9049–9057 (2013).
- Shodai, A. et al. Conserved acidic amino acid residues in a second RNA recognition motif regulate assembly and function of TDP-43. *PLoS One.* **7**, e52776 (2012).
- Shodai, A. et al. Aberrant assembly of RNA recognition motif 1 links to pathogenic conversion of TAR DNA-binding protein of 43 kDa (TDP-43). *J. Biol. Chem.* **288**, 14886–14905 (2013).
- Chiang, C. H. et al. Structural analysis of disease-related TDP-43 D169G mutation: linking enhanced stability and caspase cleavage efficiency to protein accumulation. *Sci. Rep.* **6**, 21581 (2016).
- Guenther, E. L. et al. Atomic-level evidence for packing and positional amyloid polymorphism by segment from TDP-43 RRM2. *Nat. Struct. Mol. Biol.* **25**, 311–319 (2018).
- Zacco, E., Martin, S. R., Thorogate, R. & Pastore, A. The RNA-Recognition motifs of TAR DNA-Binding protein 43 May play a role in the aberrant Self-Assembly of the protein. *Front. Mol. Neurosci.* **11**, 1–11 (2018).
- Dang, M., Li, Y. & Song, J. Tethering-induced destabilization and ATP-binding for tandem RRM domains of ALS-causing TDP-43 and hnRNPA1. *Sci. Rep.* **11**, 1034 (2021).
- Zacco, E. et al. RNA as a key factor in driving or preventing self-assembly of the TAR DNA-binding protein 43. *J. Mol. Biol.* **431**, 1671–1688 (2019).
- Rengifo-Gonzalez, J. C. et al. The cooperative binding of TDP-43 to GU-rich RNA repeats antagonizes TDP-43 aggregation. *eLife.* **10**, e67605 (2021).
- Ozguney, B., Mohanty, P. & Mittal, J. RNA binding tunes the conformational plasticity and intradomain stability of TDP-43 tandem RNA recognition motifs. *Biophys. J.* **123**, 3844–3855 (2024).
- Chen, H. J. et al. RRM adjacent TARDBP mutations disrupt RNA binding and enhance TDP-43 proteinopathy. *Brain* **142**, 3753–3770 (2019).
- Gonçalves, P. B., Soderro, A. C. R. & Cordeiro, Y. Green tea Epigallocatechin-3-gallate (EGCG) targeting protein misfolding in drug discovery for neurodegenerative diseases. *Biomolecules* **11**, 767 (2021).
- Singh, N. A., Mandal, A. K. A. & Khan, Z. A. Potential neuroprotective properties of epigallocatechin-3-gallate (EGCG). *Nutr. J.* **15**, 60 (2015).
- Nagle, D. G., Ferreira, D. & Zhou, Y. D. Epigallocatechin-3-gallate (EGCG): chemical and biomedical perspectives. *Phytochemistry* **67**, 1849–1855 (2006).
- Payne, A., Nahashon, S., Taka, E., Adinew, G. M. & Soliman, K. F. A. Epigallocatechin-3-Gallate (EGCG): New therapeutic perspectives for neuroprotection, aging, and neuroinflammation for the modern age. *Biomolecules* **12**, 371 (2022).
- Masuda, M. et al. Small molecule inhibitors of  $\alpha$ -Synuclein filament assembly. *Biochemistry* **45**, 6085–6094 (2006).
- Shal, B., Ding, W., Ali, H., Kim, Y. S. & Khan, S. Anti-neuroinflammatory potential of natural products in Attenuation of Alzheimer's disease. *Front. Pharmacol.* **9**, 548 (2018).
- Ehrnhoefer, D. E. et al. Green tea (–)-epigallocatechin-gallate modulates early events in Huntingtin misfolding and reduces toxicity in Huntington's disease models. *Hum. Mol. Genet.* **15**, 2743–2751 (2006).



40. Ehrnhoefer, D. E. et al. EGCG redirects amyloidogenic polypeptides into unstructured, off-pathway oligomers. *Nat. Struct. Mol. Biol.* **15**, 558–566 (2008).
41. Wang, I. F. et al. The self-interaction of native TDP-43 C terminus inhibits its degradation and contributes to early proteinopathies. *Nat. Commun.* **3**, 766 (2012).
42. Meshram, V. D. et al. In silico and in vitro studies suggest epigallocatechin gallate (EGCG), a polyphenol in green tea, can bind and modulate the aggregation and cytotoxicity of the full-length TDP-43 protein implicated in TDP-43 proteinopathies. 12.22.573011 (2023). <https://doi.org/10.1101/2023.12.22.573011> (2023).
43. Biancalana, M. & Koide, S. Molecular mechanism of Thioflavin-T binding to amyloid fibrils. *Biochim. Biophys. Acta.* **1804**, 1405–1412 (2010).
44. Sternke-Hoffmann, R., Peduzzo, A., Bolakhri, N., Haas, R. & Buell, A. K. The aggregation conditions define whether EGCG is an inhibitor or enhancer of  $\alpha$ -Synuclein amyloid fibril formation. *Int. J. Mol. Sci.* **21**, 1995 (2020).
45. Sneideris, T. et al. The environment is a key factor in determining the Anti-Amyloid efficacy of EGCG. *Biomolecules* **9**, 855 (2019).
46. An, T. T., Feng, S. & Zeng, C. M. Oxidized Epigallocatechin gallate inhibited lysozyme fibrillation more strongly than the native form. *Redox Biol.* **11**, 315–321 (2017).
47. Mizooku, Y., Yoshikawa, M., Tsuneyoshi, T. & Arakawa, R. Analysis of oxidized Epigallocatechin gallate by liquid chromatography/mass spectrometry. *Rapid Commun. Mass. Spectrom.* **17**, 1915–1918 (2003).
48. Wei, Y. et al. Certain (–)-epigallocatechin-3-gallate (EGCG) auto-oxidation products (EAOPs) retain the cytotoxic activities of EGCG. *Food Chem.* **204**, 218–226 (2016).
49. Rochet, J. C. & Lansbury, P. T. Amyloid fibrillogenesis: themes and variations. *Curr. Opin. Struct. Biol.* **10**, 60–68 (2000).
50. Maity, S., Gundampati, R. K. & Suresh Kumar, T. K. NMR methods to characterize protein–ligand interactions. *Nat. Prod. Commun.* **14**, 1934578X19849296 (2019).
51. Walpole, S., Monaco, S., Nepravishta, R. & Angulo, J. STD NMR as a technique for ligand screening and structural studies. *Methods Enzymol.* **615**, 423–451 (2019).
52. Campbell, A. P. & Sykes, B. D. The two-dimensional transferred nuclear overhauser effect: theory and practice. *Annu. Rev. Biophys. Biomol. Struct.* **22**, 99–122 (1993).
53. Leone, M., Freeze, H. H., Chan, C. S. & Pellecchia, M. The nuclear overhauser effect in the lead identification process. *Curr. Drug Discov. Technol.* **3**, 91–100 (2006).
54. Baek, M. et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).
55. Lukavsky, P. J. et al. Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. *Nat. Struct. Mol. Biol.* **20**, 1443–1449 (2013).
56. Auweter, S. D. et al. Molecular basis of RNA recognition by the human alternative splicing factor Fox-1. *EMBO J.* **25**, 163–173 (2006).
57. Furukawa, Y. et al. A molecular mechanism realizing sequence-specific recognition of nucleic acids by TDP-43. *Sci. Rep.* **6**, 20576 (2016).
58. Kay, L. E., Torchia, D. A. & Bax, A. Backbone dynamics of proteins as studied by  $^{15}\text{N}$  inverse detected heteronuclear NMR spectroscopy: application to Staphylococcal nuclease. *Biochemistry* **28**, 8972–8979 (1989).
59. Cierpicki, T. & Otłowski, J. Amide proton temperature coefficients as hydrogen bond indicators in proteins. *J. Biomol. NMR.* **21**, 249–261 (2001).
60. Dominguez, C., Boelens, R. & Bonvin, A. M. J. J. HADDOCK: a protein–protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **125**, 1731–1737 (2003).
61. van Zundert, G. C. P. et al. The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes. *J. Mol. Biol.* **428**, 720–725 (2016).
62. Van Der Spoel, D. et al. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).
63. Bieschke, J. Natural compounds May open new routes to treatment of amyloid diseases. *Neurotherapeutics.* **10**, 429–439 (2013).
64. Fusco, G. et al. Molecular determinants of the interaction of EGCG with ordered and disordered proteins. *Biopolymers.* **109**, e23117 (2018).
65. Ban, T., Hamada, D., Hasegawa, K., Naiki, H. & Goto, Y. Direct observation of amyloid fibril growth monitored by thioflavin T fluorescence. *J. Biol. Chem.* **278**, 16462–16465 (2003).
66. Sabaté, R. & Saupe, S. J. Thioflavin T fluorescence anisotropy: an alternative technique for the study of amyloid aggregation. *Biochem. Biophys. Res. Commun.* **360**, 135–138 (2007).
67. Groenning, M. Binding mode of thioflavin T and other molecular probes in the context of amyloid fibrils-current status. *J. Chem. Biol.* **3**, 1–18 (2010).
68. Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR.* **6**, 277–293 (1995).
69. Vranken, W. F. et al. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696 (2005).
70. Scott, D. D., Francois-Moutal, L., Kumirov, V. K. & Khanna, M.  $^1\text{H}$ ,  $^{15}\text{N}$  and  $^{13}\text{C}$  backbone assignment of Apo TDP-43 RNA recognition motifs. *Biomol. NMR Assign.* **13**, 163–167 (2019).
71. Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **73**, 1–16 (2013).
72. Yuan, S., Chan, H. C. S. & Hu, Z. Using PyMOL as a platform for computational drug design. *WIREs Comput. Mol. Sci.* **7**, e1298 (2017).
73. Huang, J. et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods.* **14**, 71–73 (2017).
74. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
75. Nosé, S. & Klein, M. L. Constant pressure molecular dynamics for molecular systems. *Mol. Phys.* **50**, 1055–1076 (1983).
76. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
77. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).
78. Essmann, U. et al. A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).
79. Humphrey, W., Dalke, A. & Schulten, K. Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).

## Acknowledgements

This work was funded under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.4 - Call for tender No. 3138 of 16 December 2021, rectified by Decree n.3175 of 18 December 2021 of the Italian Ministry of University and Research funded by the European Union–NextGenerationEU. Project code CN\_00000033, Concession Decree No. 1034 of 17 June 2022 adopted by the Italian Ministry of University and Research, CUP B73C21001300006, Project title “National Biodiversity Future Center - NBFC”. VD was sup-

ported by Programma Operativo Nazionale Ricerca e Innovazione 2014-2020 (CCI 2014IT16M2OP005), Fondo Sociale Europeo, Azione I.1 “Dottorati Innovativi con caratterizzazione Industriale”, CUP F85F20000260006. The authors acknowledge the ATeN Center of the University of Palermo for infrastructure support, and Annalisa Pastore for useful discussion.

### Author contributions

C.A. conceptualization, supervision, resources, and funding acquisition; V.D., E.M., and R.S. samples preparations; M.S. and M.C.V. fluorescence-based measurements; V.D. and M.A.M. NMR acquisition and analysis; A.S. and M.A.M. docking and molecular dynamics; C.A., V.D. writing-original draft preparation; C.A., M.A.M., V.D., A.S., F.G., A.P., and S.M. writing-review and editing. All authors have read and agreed to the published version of the manuscript.

### Declarations

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-02035-6>.

**Correspondence** and requests for materials should be addressed to C.A.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025