International Journal of Biological Macromolecules 275 (2024) 133663

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

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

Effect of E134K pathogenic mutation of SMN protein on SMN-SmD1 interaction, with implication in spinal muscular atrophy: A molecular dynamics study

Eugenia Polverini [a,*], Pietro Squeri [b], Valeria Gherardi [b,1]

[a] Department of Mathematical, Physical and Computer Sciences, University of Parma, Parco Area delle Scienze 7/A, 43124 Parma, Italy
[b] Department of Chemistry, Life Science and Environmental Sustainability, University of Parma, Parco Area delle Scienze 11/A, 43124 Parma, Italy

ARTICLE INFO

Keywords:
Spinal muscular atrophy
Tudor domain
Sm proteins

ABSTRACT

Spinal muscular atrophy (SMA) is a disease that results from mutations in the Survival of Motor Neuron (SMN) gene 1, leading to muscle atrophy due to motor neurons degeneration. SMN plays a crucial role in the assembly of spliceosomal small nuclear ribonucleoprotein complexes via binding to the arginine-glycine rich C-terminal tails of Sm proteins recognized by SMN Tudor domain. E134K Tudor mutation, cause of the more severe type I SMA, compromises the SMN-Sm interaction without a perturbation of the domain fold. By molecular dynamics simulations, we investigated the mechanism of Tudor-SmD1 interaction, and the effects on it of E134K mutation. It was observed that E134 is crucial to catch the positive dimethylated arginines (DMRs) of the SmD1 tail that, wrapping around the acidic Tudor surface, enters a central DMR into an aromatic cage. The flexible cage residue Y130 must be blocked from the wrapped tail to assure a stable binding.

The charge inversion in E134K mutation causes the loss of a critical anchor point, disfavoring the tail wrapping and leaving Y130 free to swing, leading to DMR detachments and exposition of the C-terminal region of the tail. This could suggest new hypotheses regarding a possible autoimmune response by anti-Sm autoantibodies.

1. Introduction

Spinal Muscular Atrophy (SMA) is an autosomal recessive neurodegenerative disorder considered the first genetic cause of infant mortality [1–3]. It affects especially motor neurons cells, being mostly characterized by degeneration of anterior horn cells of the spinal cord, leading to progressive limb and trunk paralysis and muscular atrophy. Based on the age of onset, the gravity of the clinical picture, and the life span, SMA has been classified into four types (I – IV) by the International SMA Consortium (1992). In the early '90"s the locus of SMA was identified on chromosome 5q13 of the human genome. Actually, this locus contains two genes, SMN1 and SMN2, which produce an identical SMN protein. [2]. Unfortunately, a C-T transition in SMN2 gene specifically modifies the sequence of an exonic splice enhancer [4–6], causing the production of a protein lacking exon 7 in most transcripts (SMNΔ7), protein that is unstable and is rapidly degraded. Therefore, SMN2 does express only 10 % of full-length transcripts, while the vast majority of full-length SMN protein comes from SMN1 gene, which represents the predominant contributor to SMN protein levels [7]. This justifies the fact that over 98 % of all SMA patients carry homozygous mutations or deletions of SMN1 gene. In fact, in these conditions the production of functional SMN protein will depend only on the low production levels of SMN2 gene, highly correlating with severity of SMA [8].

SMN is a 38 KDa − 294 residues protein found in the cytoplasm and nucleus of cells [9–12]. It harbors multiple domains interacting with several different protein partners, in particular: the N-terminal Gemin2 domain, included in the lysine-rich nucleic acid-binding domain; a central Tudor domain; the proline-rich profiline-binding domain and a C-terminal YG domain. Excluding Gemin2, Tudor and YG domain, which are the only experimentally resolved SMN regions [13–19], the remaining large regions of the protein are unstructured [20], therefore the whole SMN structure was only predicted, both by RosettaCM [21] as was made by Singh and co-workers [20], and by AlphaFold [22](AF-Q16637-F1). This multifunctional protein has a role in numerous cellular processes and pathways, mainly regarding various aspects of the RNA metabolism [20,23]. Particularly relevant for this work – and the

* Corresponding author.
E-mail address: eugenia.polverini@unipr.it (E. Polverini).
[1] Present Address: Department of Medicine and Surgery, University of Parma, Via Gramsci 14, 43,126 Parma, Italy.

https://doi.org/10.1016/j.ijbiomac.2024.133663
Received 14 March 2024; Received in revised form 1 July 2024; Accepted 2 July 2024
Available online 3 July 2024

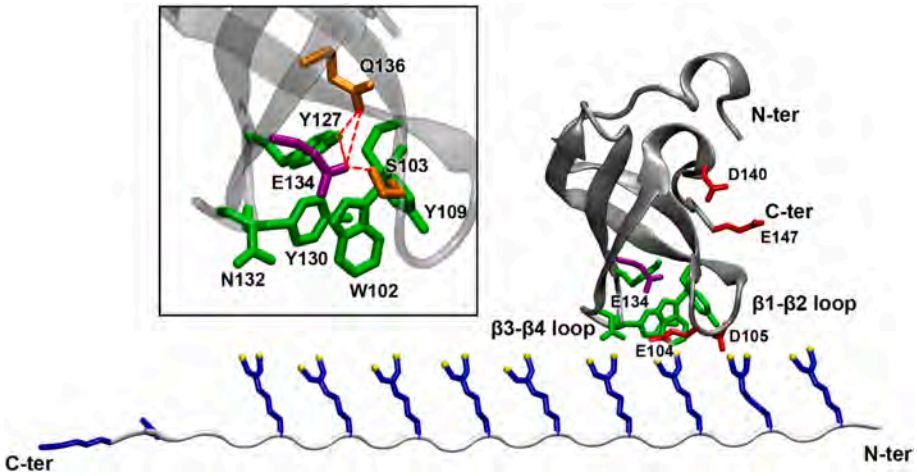


Fig. 1. Starting structure of the wt SMN Tudor domain in the presence of the SmD1 tail. The structure of the SmD1 tail was initially built in extended conformation because it is known that this RG-repeats containing region is unstructured (see Section 2.1). Remarkable residues are labelled and represented in sticks. The mutant residue E134 is purple, the cage residues are green, acidic residues that come in contact with the SmD1 tail are red. In the underlying SmD1tail, DMRs are in blue, with the carbons of the symmetrical dimethylation in yellow balls. The two C-terminal not methylated arginines are also in blue. The cage and the H-bond network residues (Ser103, Tyr127, Glu134 and Gln136, connected with red dashed lines) are shown in the zoomed box.

mainly studied function - it is involved in snRNP assembly, forming a multimeric complex composed of SMN, Gemin proteins (Gemins 2–8) and of UNR-interacting protein (Unrip). The SMN complex mediates the assembly of the seven spliceosomal Sm proteins core (Sm B, D1–3, E, F and G) on their small nuclear RNA (snRNA) binding-sites, to form the assembled Small Nuclear Ribonucleoprotein (snRNP) particles, crucial components of the spliceosomal machinery [12,24–26]. In snRNP assembly, as in several other biological pathways (e.g. in regulation of nuclear architecture, by interacting with coilin protein in Cajal bodies [27], and in regulation of transcription termination by interacting with RNA polymerase II subunit POLR2A [28]), the exon-3 encoded Tudor domain of SMN plays a critical role by interacting with symmetrically dimethylated arginine (sDMR) in the partner proteins. Symmetrical dimethylation of arginine is a very rare posttransaltional modification in humans, generated by placing one methyl group on each of the terminal guanidino nitrogen atoms. In snRNP assembly, the Tudor domain directly binds the symmetrically dimethylated arginines in the C-terminal tail of Sm D1, D3 and B/B′, which are rich in arginine-glycine (RG) repeats [16,17,24,29,30]. It was demonstrated that the symmetrical dimethylation is critically important for SMN binding, strongly enhancing the binding affinity, and that the basic Tudor fold alone is sufficient for binding [17,19,31,32]. The Tudor domain exhibits a structure made by a five-stranded antiparallel β-sheet, strongly bent to form a barrel-like fold (Fig. 1) stabilized by conserved hydrophobic residues. An aromatic cage on the Tudor domain, made up of Tyr127, Tyr130, Asn132, Trp102 and Tyr109 (human SMN sequence numbering), seems to be responsible for mediating the binding with DMR by means of cation-π interactions and hydrophobic interactions. These are reinforced by the symmetrical demethylation, which gives to arginine hydrophobicity and bulkiness, without neutralizing the cationic charge. From titration experiments with single DMR residues, it was shown that the Tudor domain is able to accommodate a DMR residue in the cage, but a cooperativity effect due to the other flanking DMRs was hypothesized to enhance the binding affinity [17,19]. This fact correlates with the presence of several negatively charged residues located at the Tudor surface.

Six of the pathogenic missense mutations causing SMA are located in the SMN Tudor domain, interfering at different levels with the interaction with Sm proteins. In particular, the E134K mutation strongly reduces such interaction, causing the more severe type I SMA [33,34]. This mutation involves one of the negatively charged residues placed at the domain surface and close to the cage residues, and it is of particular interest because, among the six mutations, it is the only one that does not perturbs the domain fold [16,19,29]. Therefore, this mutation would interfere somehow with the direct interaction of Tudor with Sm protein dimethylated tails, compromising the correct assembly of snRNP, which is fundamental for obtaining mature mRNA. Even if the introduction of new FDA-approved therapies such as SM2 splicing modifiers (nusinersen - Spinraza, risdiplam - Evrysdi), and gene replacement therapy (onasemnogene abeparvovec-xioi - Zolgensma) has already shown encouraging results in altering the natural course of the disease, studies have demonstrated that not all patients respond to therapy, especially those with one copy of SMN2 or those who receive post-symptomatic treatment. In addition to this, not all patients have access to the above mentioned therapies due to cost, availability and physical condition, thus making necessary to search for new treatment approaches [35]. In this scenario, delving into the molecular basis of SMN structure and interaction may provide new insight into its disease-causing mutations and offer new ideas for the development of new therapies. Therefore, the present work focused, from a structural biology perspective, on the mechanism by which Tudor SMN interacts with the methylated C-terminal tail of Sm proteins, and on the structural effects of the above mentioned deleterious mutation E134K on this interaction. The direct interference of the E134K mutation in SMN-Sm interaction, with the maintaining of Tudor domain fold, could help to better understand the key points of the binding mechanism. The SmD1 protein was chosen in this study, due to the reported dependence of the strength of the binding from the number of DMRs [17,19]. In fact, the SmD1 tail is characterized by nine consecutive RG repeats, all symmetrically dimethylated. The study was carried out by means of molecular dynamics (MD) simulation techniques, investigating the dynamic behavior of wild type and E134K mutated forms of the SMN Tudor domain, alone and in the presence of the C-terminal tail of SmD1 (residues 97–119, human SmD1 numbering, Fig. 1). With this approach we were able to monitor the formation of the complex between the Tudor and the SmD1 tail, gaining new insights into their interactions, or lack thereof in the mutated protein. The entering of a central DMR of the SmD1 tail in the Tudor aromatic cage and the wrapping of the other tail residues around the Tudor surface were observed. However, in the presence of E134K mutation it was followed by detachments and instability. The contribution of remarkable residues, and the conditions that must occur to stabilize the binding, were pointed out. In particular, the abolishing – with the E134K mutation - of a hydrogen bond triangle made by Tyr127, Glu134 and Gln136, that was hypothesized to impair the binding to DMR [19], was revised in light of

Table 1
Summary of the performed MD simulations and their conventional names. All the simulations are 400 ns long.

Protein variant	Replicas
wt Tudor domain	wt_r1 wt_r2
wt Tudor domain in the presence of SmD1 tail	wt-SmD1_r1 wt-SmD1_r2 wt-SmD1_r3 wt-SmD1_r4 wt-SmD1_r5
Tudor-E134K variant	E134K_r1 E134K_r2
Tudor-E134K variant in the presence of SmD1 tail	E134K-SmD1_r1 E134K-SmD1_r2 E134K-SmD1_r3 E134K-SmD1_r4 E134K-SmD1_r5

a fourth residue involved. The role of the RG repeats flanking the DMR that enters the cage, involving their interaction with the conserved negatively charged patches on the surface of the Tudor domain, was also revealed.

2. Methods

2.1. Molecular dynamics simulations

To investigate the mechanism of interaction between the SMN Tudor domain and the SmD1 C-terminal region, and the effects of the E134K mutation, MD simulations were carried out, in several replicas, on four systems: the Tudor domain alone, the Tudor domain in the presence of the SmD1 tail, the Tudor-E134K variant alone, and the Tudor-E134K variant in the presence of the SmD1 tail. (see Table 1 for number and conventional names of simulations replicas).

For the building of the starting configuration of the system, the structure of the human SMN Tudor domain (residues 84-147) was extracted by the PDB databank [36], id code 4A4E [19]. After a careful comparison of all the available structures, this one was chosen because the cage was conformed to host a DMR, which is the ligand whose binding we want to investigate. Then, the DMR present inside the cage was removed, and the model 1 of the NMR ensemble was used. The E134K variant of the Tudor domain (Tudor-E134K) was built by means of the Swiss-Pdb Viewer software [37], replacing Glu134 with a lysine in the wt Tudor structure.

The C-terminal region of SmD1 protein (also indicated in the text as "SmD1 tail") was built by means of Hyperchem software (Hypercube, Inc.) in extended conformation, because it is known that this RG-repeats containing region is unstructured [38]. The last 23 residues of the SmD1 human sequence (97–119, human SmD1 numbering, Uniprot code: P62314), containing nine RG-repeats, were used:

$G^{97}\underline{R}G\underline{R}^{100}G\underline{R}G\underline{R}G\underline{R}G\underline{R}G\underline{R}^{110}G\underline{R}G\underline{R}GGPRR^{119}$

The underlined R means that this residue is known to be symmetrically dimethylated. Only this region was used, because it is reported that the residues outside the RG-repeats do not contribute to the interaction with the Tudor domain [17]. Symmetrical dimethylation of arginine is a posttransaltional modification generated by placing one methyl group on each of the terminal guanidino nitrogen atoms. The symmetrical dimethylation was added to the indicated arginine residues of the RG-repeats with the same Hyperchem software package.

Then, in the simulations in the presence of the SmD1 C-terminal region, the SmD1 tail was manually placed close to the Tudor cage (Fig. 1) by means of VMD software [39], at a distance that would favor the interaction of the SmD1 tail with the cage region, accelerating the recognition process, but without forcing the entering of a particular DMR and without assuming the formation of predefined interactions.

The protein system (i.e. the Tudor domain alone or in the presence of the SmD1 tail) was then solvated with a water box 1 nm thick around the solute. Chlorine ions were added to the solvent to keep the system neutral. Both termini of the Tudor domain and the N-terminus of the SmD1 tail were kept in a neutral form, considering the absence of a net charge due to the presence of the rest of the protein.

An energy minimization was first performed on the whole system up to a gradient of 100 kJ/(mol nm). Afterwards, a position-restrained dynamics was run for 100 ps to let the solvent relax around the solute. Full MD simulations were then run in multiple replicas 400 ns long at 300 K and 1 bar. A time step of 0.002 ps was used. The NPT ensemble was used, and periodic boundary conditions were applied to the system both for position-restrained and for full MD.

To perform MD simulations, the GROMACS software package [40] with the Gromos96 ffG53a6 force-field [41] was used. The DMR parameters for the force field (available in the Supplementary material) were calculated by means of the ATB server [42] and the partial charges were then manually adjusted, basing on the comparison with Gasteiger distribution [43]. The calculated parameters were compared with those used for analogous constructs and their validation was made with the first control simulation of the Tudor in the presence of the SmD1 tail, in which, in short time, a correct (compared with the available NMR complex) and stable insertion of a DMR inside the cage occurred. Also, the behavior of other DMRs was coherent with the chemical properties of this residue.

Structural analyses were made by the routines embedded in the GROMACS package and by the VMD – Visual Molecular Dynamics software [44]. EP surface were calculated by means of Swiss-PDB Viewer software [37] with a cutoff of −1.8 k_bT/e for the negative value and 1.8 K_bT/e for the positive one; the isosurfaces were then mapped onto protein molecular surface. Graphics were made by OriginPro, Version 2023 (OriginLab Corporation, Northampton, MA, USA). Dynamic contact maps were built by the Conan software [45].

2.2. Binding energy analysis

In the simulations in the presence of the SmD1 tail, the analysis of the binding energy and the per residue energy decomposition was performed by means of the g_mmpbsa software [46] that uses the MM/PBSA method implemented for GROMACS, coupled with the APBS package [47]. The tool calculates the components of the binding energy (vacuum potential energy, polar solvation energy, non-polar solvation energy) in the single trajectory approach, with the exception of the entropic term, the calculation of which is computationally very expensive (in particular here, due to the size of the system) [48]. However, even if the method is unable to give the absolute free binding energy, the tool is suited for calculating relative binding energies. For the calculations, we used the portion of the wt trajectories in which the complex Tudor-SmD1 was stable (looking at the RMSD plot and matrix of both the SmD1 tail and the Tudor domain) and a DMR was present inside the cage. For the analysis of E134K complexes, we chose only the E134K-SmD1_r1 trajectory that was the only one that shows both a DMR temporarily inserted correctly into the cage and the SmD1 tail wrapped around the domain.

3. Results and discussion

3.1. Control simulations of wt and mutated Tudor domain without the SmD1 tail

Previously, a check on the stability of the structure of the wt Tudor domain alone was made, which, as expected, is confirmed by the RMSD plots and matrices, by the secondary structure evolution in time and by the behavior of radius of gyration (Rg) (Figs. S1 A, S2 A and S4 A; the RMSD matrices are available upon request). From the height of the peaks in the RMSF plots (Fig. 2 A) and from the higher oscillations in the RMSD plots of each single cage residue (Fig. 3 A), it appears that the

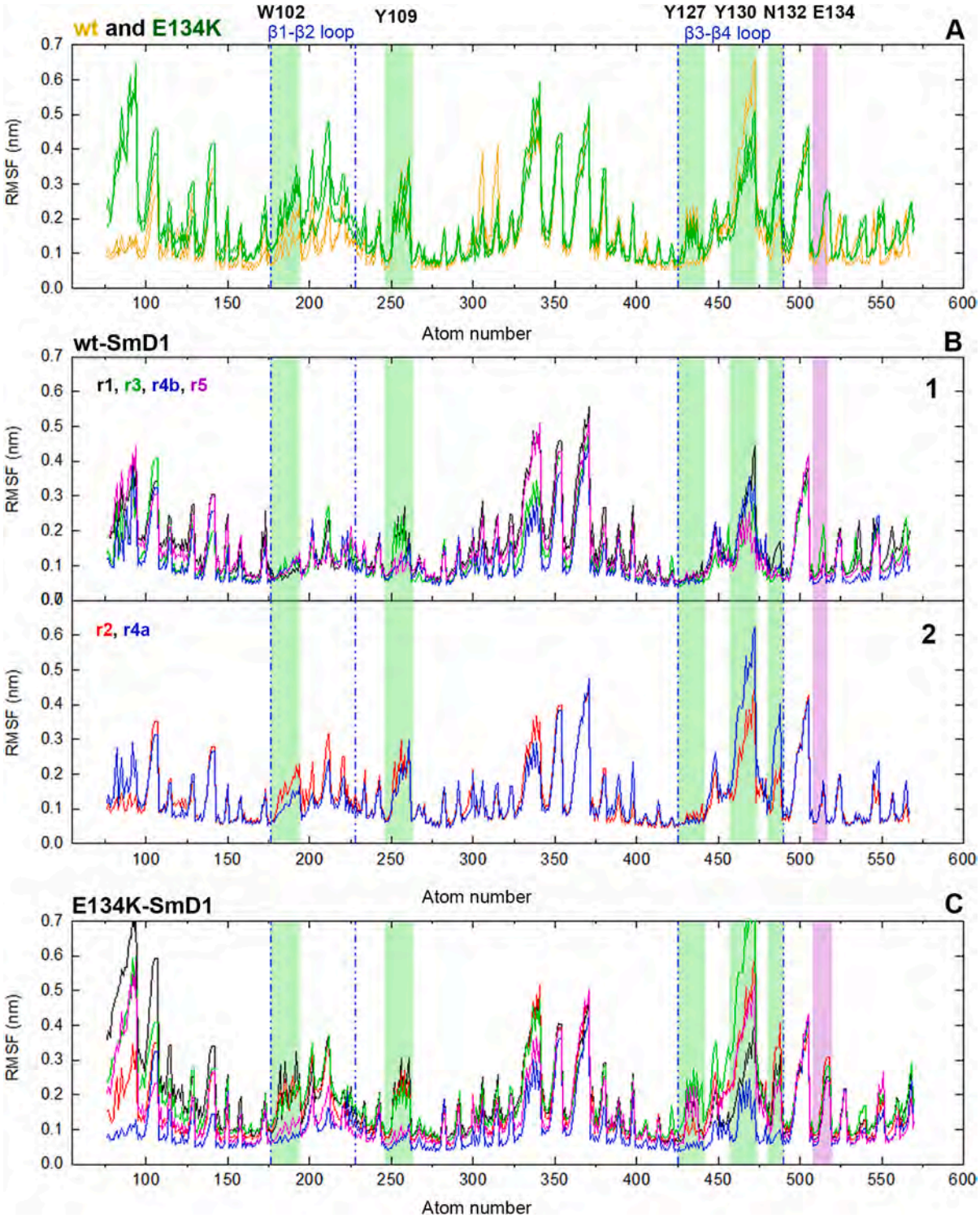


Fig. 2. RMSF plots of the atoms of the protein core (residues 92 to 139) for all the simulations. The alignment to the reference (mean) structure and the subsequent RMSF calculation (performed on all atoms) were restricted to the core region because of the high flexibility of the terminal regions. The RMSF was calculated on the last part of the trajectories, where the Tudor domain reached a stable structure, basing on the RMSD plots and matrices. Panel A shows the RMSF plots of the wt (orange) and E134K (green) simulations in the absence of the SmD1 tail; the slight shift in the last part of the E134K plot (from residue 134) is due to the higher number of atoms in the mutated residue K134. Panel B shows the RMSF in the wt-SmD1 trajectories, splitted in two graphics: the upper referring to the trajectories in which a DMR remains stable in the cage (namely, r1 (black), r3 (green), r5 (purple) and the last part of r4 (blue), named r4b); the lower referring to the trajectories in which a DMR is not stable in the cage (namely, r2 (red) and the first part of r4 (r4a, blue)). The division of r4 trajectory in two parts was made to underline the difference in flexibility of Tyr130 between the first part (very high peak) and the second part (the Tyr130 peak is almost halved), when this residue is blocked in the stacked position with DMR. Panel C shows the RMSF plots of the E134K-SmD1 simulations, progressively colored (r1 is black, r2 is red, r3 is green, r4 is blue, r5 is purple).

The vertical green bands indicate the residues belonging to the cage (see the labels on the top); the purple one indicates the residue 134. The dashed blue lines delimit in order the β1-β2 and the β3-β4 loops. β2-β3 loop, corresponding to the three peaks in the center of the graph, is always very flexible.

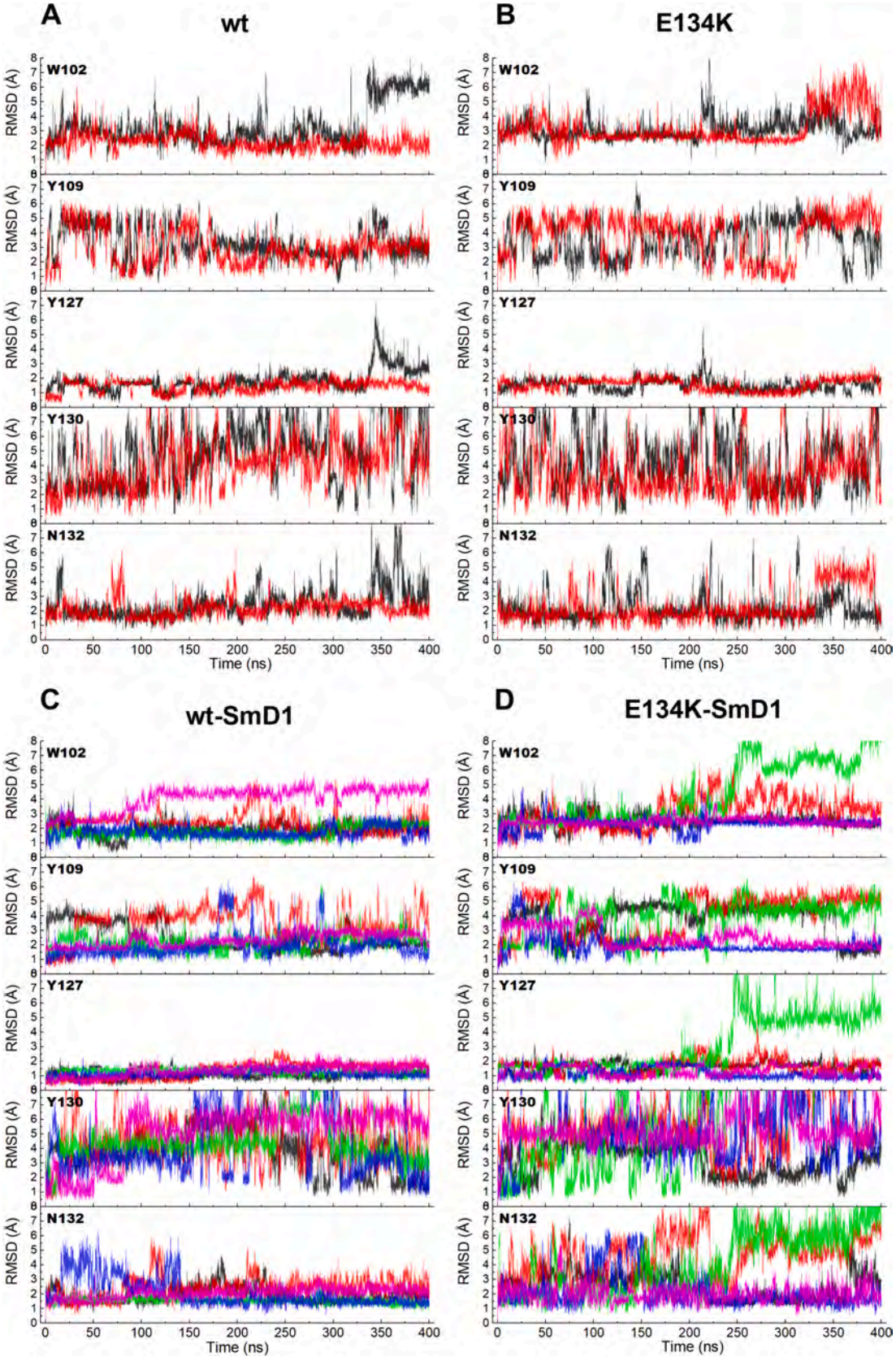
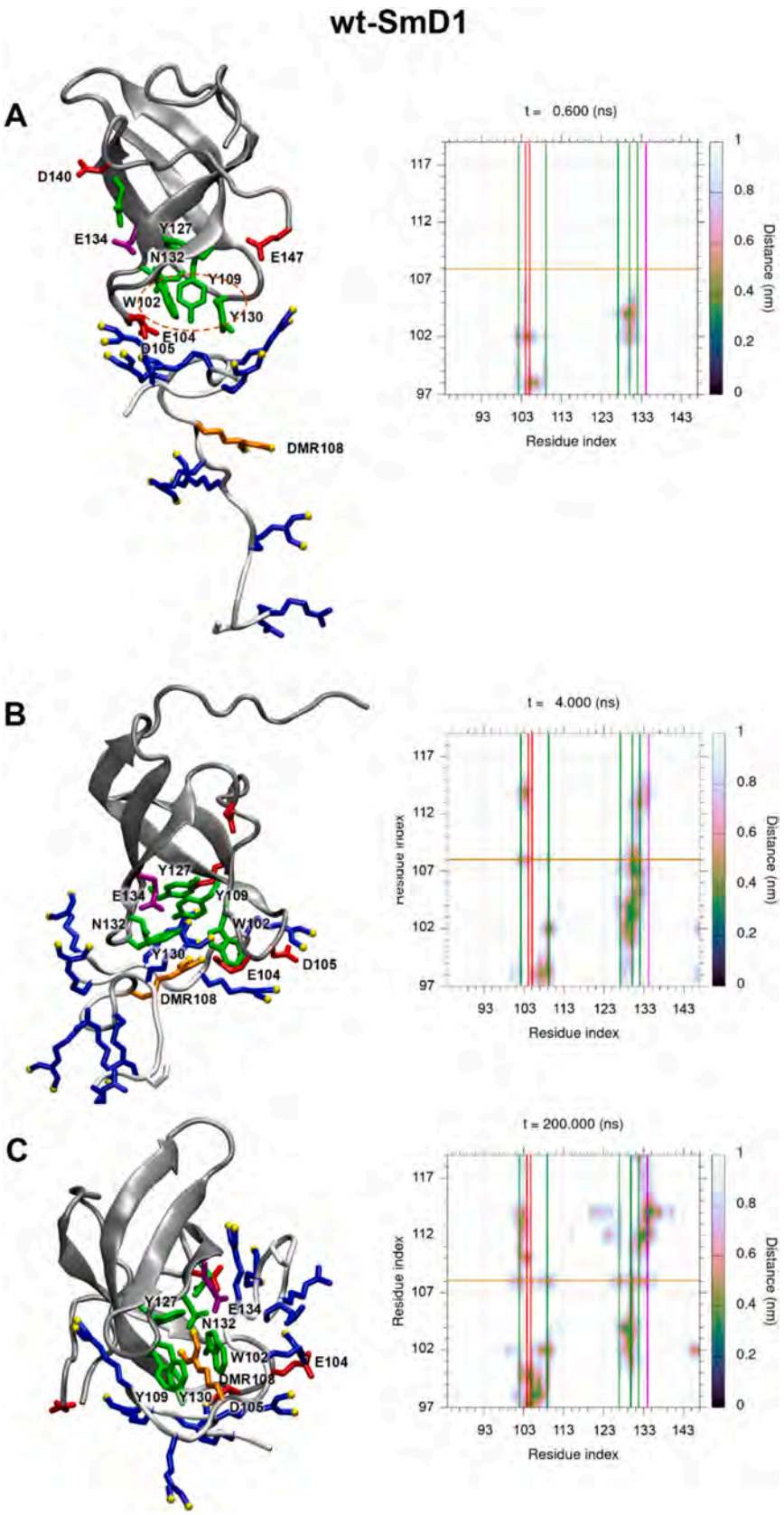


Fig. 3. RMSD plots of each of the five cage residues in the different simulations. The RMSD was calculated on all the atoms of each single residue, after superimposition to the starting structure of the Cα atoms of the protein core (residues 92 to 139). The different kinds of simulations are labelled, following the nomenclature in Table 1: panels A and B correspond to the simulations without the SmD1 tail, panels C and D to the simulations in the presence of the SmD1 tail. The replicas of each simulation are progressively colored (r1 is black, r2 is red, r3 is green, r4 is blue, r5 is purple).



(caption on next page)

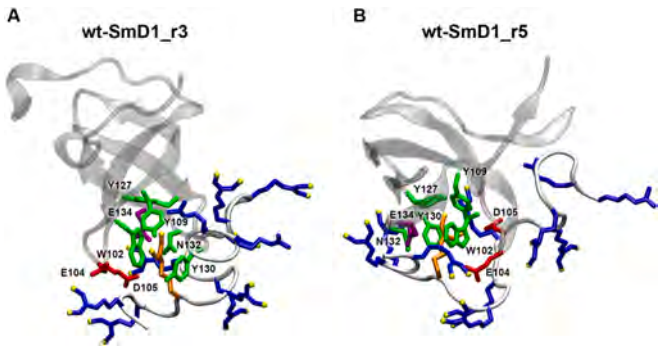


Fig. 5. Final conformation of wt-SmD1_r3 and _r5 trajectories. Remarkable residues are labelled, and colour code is as in Fig. 4; in orange is highlighted the DMR that enters the cage.

empty cage is more flexible than in the presence of the SmD1 tail, as usually happens for binding sites in the absence of their ligand. Exception is the Tyr127 residue that, being located in the inner part of the cage (Fig. 1) and linked in a strong H-bonds network (as we will discuss later), is always very stable. Flexibility concerns all the three loops: the β1-β2, β2-β3 and β3-β4 loops (Fig. 2 A and Fig. S3 A), the first and the last defining the walls of the cage and containing Trp102 and Tyr130 residues, respectively. They appear the more mobile residues of the cage in all the simulations.

However, on the opposite side of the barrel, the structure of the free domain slightly enlarges and becomes less compact (Fig. S3 A), as can also be observed by the slight enhancement in the Rg plot (Fig. S4). This seems due to the break of the H-bond between Ser139 and Gln90 located at the beginning of the two long unordered terminal extremities (nine residues each) that, even if present in the PDB structure, are not actually part of the structured Tudor domain. Such regions are free to move and destabilize the barrel. In particular in wt_r1 simulation, in the last 70 ns of the trajectory, the structure undergoes a deformation due to a random interaction of the N-terminus with the β1-β2-loop (Figs. S1 A, S2 A and S3 A). The behavior of Rg, calculated on the core region (residues 92–139) of the Tudor, confirms that the structured part of the domain maintains a very stable conformation (Fig. S4 A).

As done for the wt structure, the stability of the mutated Tudor domain alone was checked, confirming the preservation of the fold, as was previously reported by NMR experiments [19] (Figs. S1 B, S2 B and S4 B; RMSD matrices are available upon request). Also, the flexibility of the loops and of the empty cage residues is comparable to the one of the wt structure (Figs. 2 A and 3 B), and the widening of the barrel on the opposite side from the cage, due to the oscillations of the N- and C-terminals, is still present (Figs. S3 B and S4 B).

3.2. MD simulations of wt Tudor domain in the presence of the SmD1 tail

To the aim of investigating the mechanism of interaction of SMN with Sm proteins, which triggers the sequence of events leading to the formation of the snRNP particles, MD simulations were performed on a system in which the C-terminal tail of the SmD1 protein, containing nine symmetrically dimethylated RG repeats, was initially placed close to the cage side of the wt SMN Tudor domain (Fig. 1). That SmD1 region was selected because the residues outside the RG-repeats do not contribute to the interaction with the Tudor domain [17].

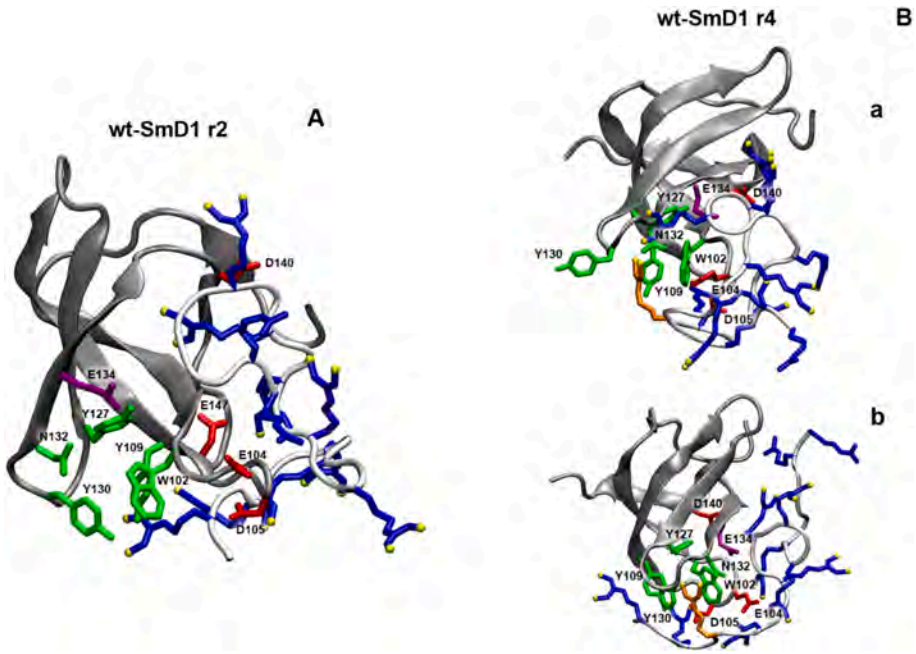


Fig. 6. Final conformation of wt-SmD1_r2 and wt-SmD1_r4 trajectories. Remarkable residues are labelled, and colour code is as in Fig. 4. **(A)** In wt-SmD1_r2 the DMR comes out from the cage and the tail interacts with the acidic wall of the protein. **(B)** The two structures show the two behaviors of Tyr130 in wt-SmD1_r4: **a)** Tyr130 is swinging, leaving the cage open; **b)** Tyr130 is blocked by the tail that maintains it in stacking with the inner DMR (orange), making the binding stable.

Fig. 4. Highlighted steps of the wt-SmD1_r1 trajectory (chosen as a paradigmatic), next to the corresponding contact maps. Each contact map shows the contacts formed between the Tudor and the SmD1 tail at the indicated time. In the Tudor structure shown in cartoon (gray), remarkable residues (in licorice) are labelled (in green are the cage residues, in purple Glu134, in red the acidic residues that come in contact with the tail). The orange dashed circle in the first step (panel A) highlights the inward shift of Trp102 that leads a transient hydrophobic collapse of the cage. In the SmD1 tail the DMRs sidechains are in blue (the yellow balls are the methyl carbons), as well as the two C-terminal not methylated arginines; in orange is highlighted the DMR108 that enters the cage. In the contact maps the remarkable residues are highlighted with the same colour code as in Tudor structure, with the two red lines referring to Glu104 and Asp105. The horizontal orange line refers to the DMR that enters the cage: the spots formed in correspondence to the cage residues are evident.

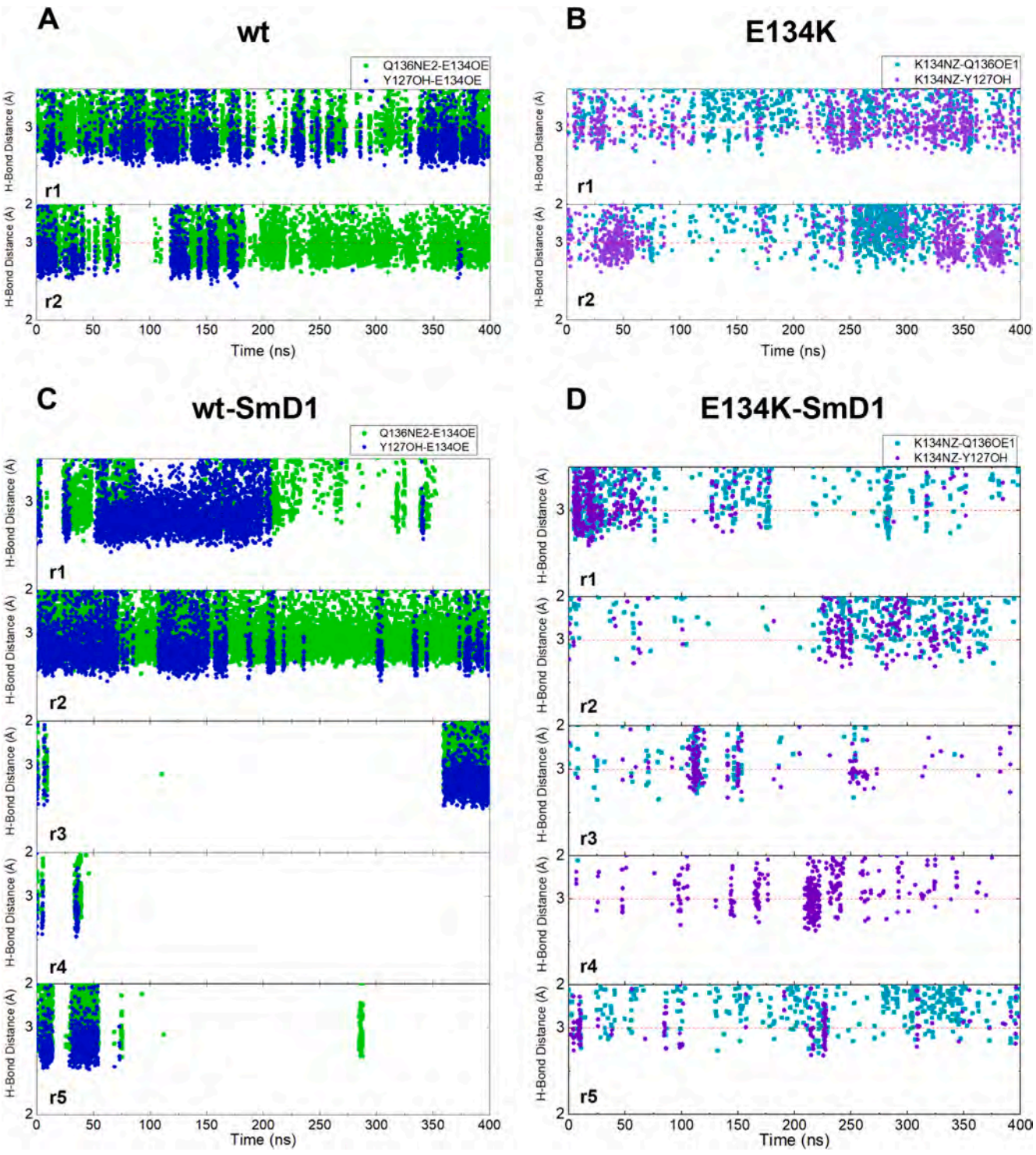


Fig. 7. Distance between donor and acceptor atoms of sidechains of both residues Tyr127 and Gln136, with respect to residue 134, to monitor the formation of H-bonds. **(A)** and **(C)** refer to the wt Tudor trajectories, in the absence and in the presence of the SmD1 tail, respectively. Tyr127 (blue points) and Gln136 (green points) are donor (with atoms OH and NE2, respectively) and Glu134 is acceptor (with both atoms OE1 and OE2). **(B)** and **(D)** refer to the mutated Tudor trajectories, in the absence and in the presence of the SmD1 tail, respectively. Lys134 is donor (with atom NZ) and Tyr127 (violet points) and Gln136 (cyan points) are acceptor (with atoms OH and OE1, respectively).
The horizontal dashed red line indicates the average distance for H-bond in proteins (about 3 Å). Each replica is labelled.

During the 400 ns long trajectories of the complete system, the Tudor fold remains very stable and compact. Its stability was confirmed, also in this case, by the RMSD plot and matrices, by the pattern of secondary structure evolution in time and by the Rg plot (Figs. S1 C and S2 C and S4 C; RMSD matrices are available upon request). Contrarily with what observed for the structure of the Tudor alone, the presence of the SmD1 tail helps to maintain the compactness of the structure and the highly flexible terminals are less free to move (Fig, S5 A).

In a short time, we were able to monitor the entrance of one of the tail's DMRs inside the aromatic cage. At the beginning of the simulation,

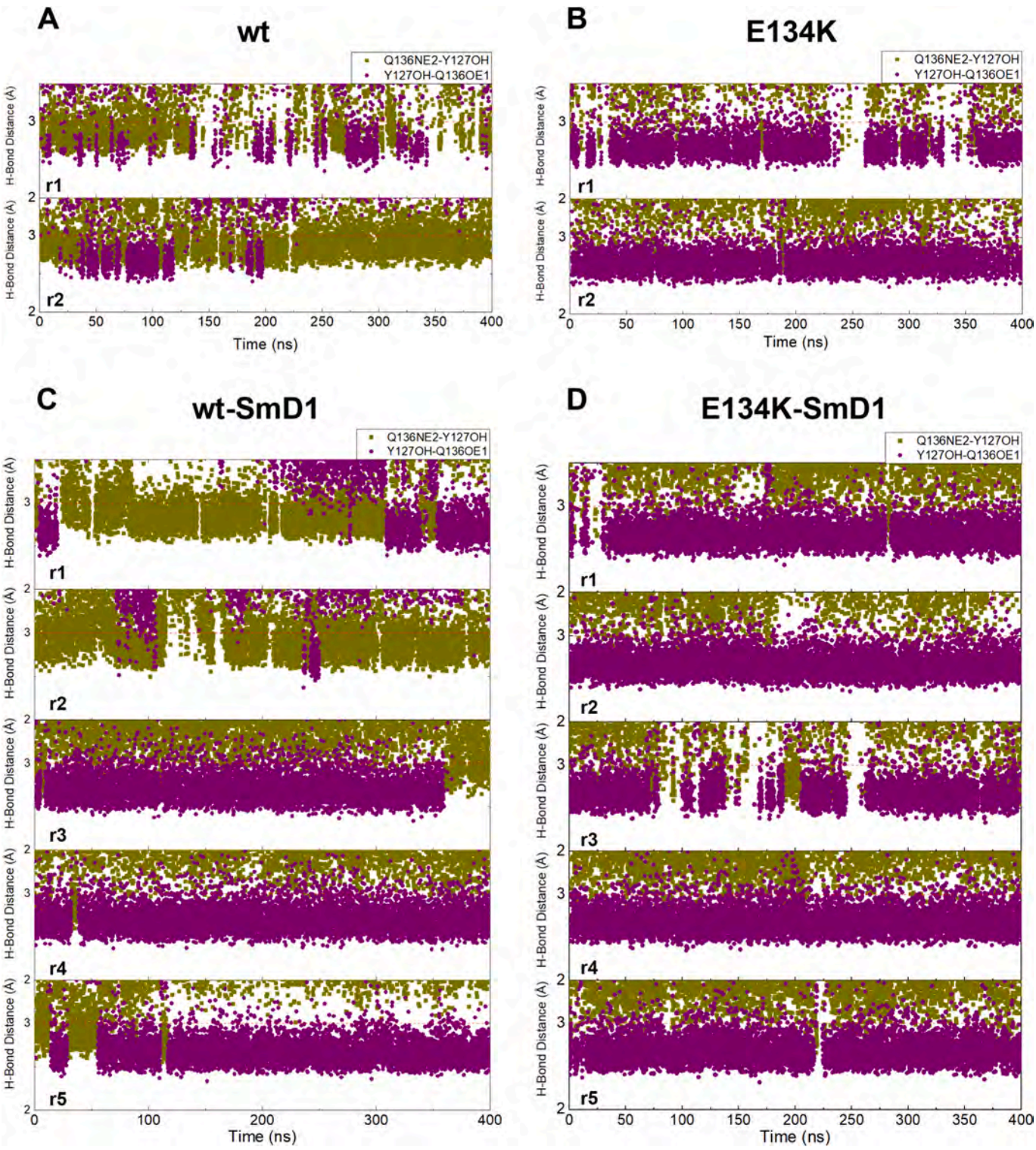


Fig. 8. Distance between donor and acceptor atoms of sidechains of residues Tyr127 and Gln136 in each simulation, to monitor the formation of H-bonds. **(A)** and **(C)** refer to the wt Tudor trajectories, in the absence and in the presence of the SmD1 tail, respectively. **(B)** and **(D)** refer to the mutated Tudor trajectories, in the absence and in the presence of the SmD1 tail, respectively. The distances were calculated both when Gln136 is donor (atom NE2) and Tyr127 acceptor (atom OH), green points, and when Tyr127 is donor (atom OH) and Gln136 acceptor (atom OE1), purple points. The horizontal red line indicates the average distance for H-bond in proteins (about 3 Å). Each replica is labelled.

when the cage is still empty, a hydrophobic collapse of the aromatic residues occurs, where Trp102 in particular occupies the cage space with its bulky sidechain (Fig. 4 A). This behavior explains the observations made by Spranger and co-workers [17] about two different orientations of the aromatic residues of the cage in the NMR and X-ray structures of the free Tudor domain.

As soon as Trp102 moves away, one of the DMR residues of the tail begins to enter the cage (Fig. 4 B). As a note, the residue that enters the cage is not necessarily the one just under the cage in the starting structure; however, a control simulation with a shifted tail was carried

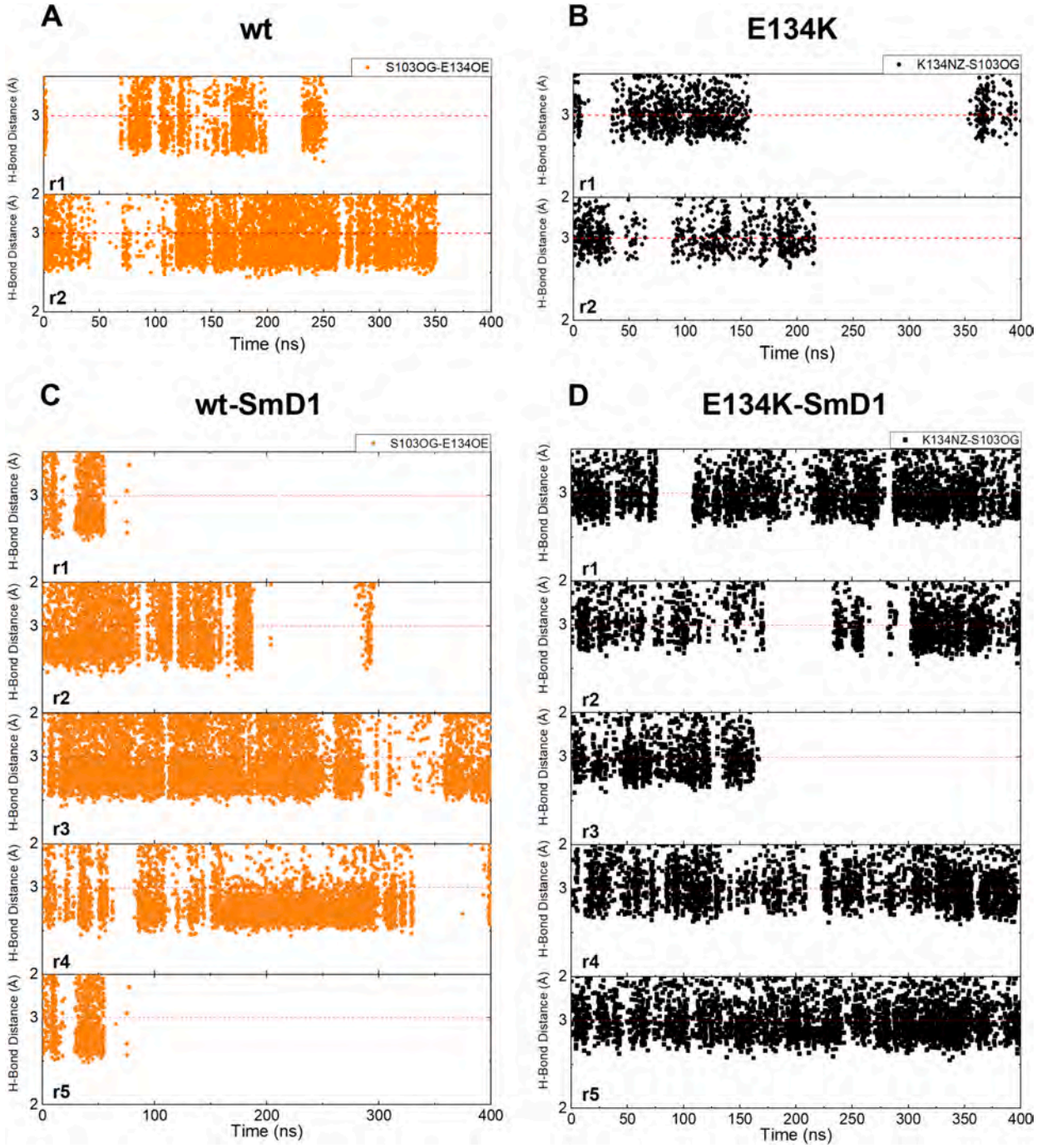


Fig. 9. Distance between donor and acceptor atoms of sidechains of residues 134 and Ser103 in each simulation, to monitor the formation of H-bonds. **(A)** and **(C)** refer to the wt Tudor trajectories, in the absence and in the presence of the SmD1 tail, respectively. Ser103 is donor (atom OG) and Glu134 is acceptor (atoms OE1 and OE2), orange points. **(B)** and **(D)** refer to the mutated Tudor trajectories, in the absence and in the presence of the SmD1 tail, respectively. Lys134 is donor (atom NZ) and Ser103 is acceptor (atom OG), black points. The horizontal red line indicates the average distance for H-bond in proteins (about 3 Å). Each replica is labelled.

out, and the independence of the entering DMR from its initial position was confirmed.

With the entry of the DMR, the cage aromatic residues rearrange, dragging the DMR to penetrate more deeply. The final stable position of the DMR sidechain (Fig. 4 C) recovers the one detected in the NMR structure by Tripsianes and co-workers [19] in their work with single DMRs in solution. The ensemble of net charge, hydrophobicity, and bulkiness of the DMR strengthens the binding by means of cation-π interactions, hydrophobic interactions, and tight fitting in the binding site. The big dimethylated arginine enhances nonelectrostatic contacts with the aromatic ring of the three tyrosines and the tryptophan of the cage. Contemporarily, the numerous cation-π interactions, both in stacking (Tyr130 and Trp102) and in T-shape (Tyr109 and Tyr127), force an orientation that shapes the binding pocket (Fig. 4 C). The "sandwich" conformation Trp102 / DMR / Tyr130 will be particularly important for the stability of the DMR inside the cage. Asn132 completes the walls of the cage even if it is involved more in a network of interactions with surrounding protein residues – in particular residue 134 – than with the DMR inside the cage.

However, the presence of the whole tail of SmD1 is crucial for the formation of the complex and the stability of the binding. In fact, the other DMRs play an active role not only in the first interactions with the

Table 2
Average and standard deviation of all energetic components and of the total binding energy (kJ/mol) calculated for selected trajectory intervals of each simulation.

Energy terms / Simulations	van der Waal	Electrostatic	Polar solvation	Non polar solvation (SASA model)	Total binding energy
wt-SmD1_r1	-294 ± 31	−2722 ± 158	1138 ± 137	−36 ± 4	−1914 ± 101
wt-SmD1_r2	−286 ± 36	−2867 ± 241	1094 ± 184	−36 ± 5	−2094 ± 134
wt-SmD1_r3	−288 ± 26	−2389 ± 148	893 ± 122	−33 ± 4	−1817 ± 101
wt-SmD1_r4	−276 ± 25	−2862 ± 122	1214 ± 104	−36 ± 3	−1959 ± 91
wt-SmD1_r5	−388 ± 32	−2891 ± 131	1421 ± 132	−47 ± 4	−1906 ± 92
E134K-SmD1_r1	−237 ± 22	−1077 ± 96	599 ± 77	−29 ± 2	−743 ± 80

Tudor surface - which lead to the recognition of the binding site - but also in maintaining the binding by wrapping around the acidic surface of the domain (Figs. 4 C, 5, 6 Bb, and Fig. S7 A). The conserved negatively charged patches on the surface of the Tudor domain (Fig. S6 A), and particularly the ones near the cage (indicated by a green arrow in Fig. S6 A), mediates the electrostatic interaction with the cations of the SmD1 tail, which is the driving force of the binding. In fact, as monitored by the dynamic contact maps (Fig. 4, right side of each panel), the very first contacts occur between the basic DMRs and the acidic Glu104 and Asp105 located at the cage entrance in the flexible β1-β2 loop, and they involve quickly Glu134 that, turning itself outwards, catches the DMRs sidechains. In agreement with this, previously NMR experiments [17] report changes in chemical shifts that support the interactions of Glu104, Asp 105 and Glu134 with DMRs. Interestingly, these three acidic residues, together with Asp140 and Glu147, are the only ones that do not form salt-bridge with any other basic residue of the Tudor, and are also the ones with the most favorable binding energy (see the binding energy discussion below). Indeed, Asp140 and Glu147 are also involved in electrostatic interaction with the SmD1 tail, even if the salt bridges of DMRs with Glu105, Asp104 and Glu134 are much more persistent. Jumping from one salt-bridge to another, the tail will move to its final conformation, wrapped around the Tudor surface at the cage-side of the Tudor (Figs. 4 C, 5, 6 Bb and Fig. S7 A). The involvement of the whole DMRs tail in the binding can justify the experimentally observed enhance of affinity for the ligand in the presence of an increasing number of dimethylated RG-repeats [17,19].

However, not in all replicas does a DMR enter the cage in a stable and compact manner. Let's examine wt_SmD1_r2 and wt_SmD1_r4. In wt_SmD1_r2, the SmD1 tail reversed the positions of its terminals, and failed to wrap around the Tudor acidic surface (Fig. 6 A). Glu134 was not able to catch the DMRs tail and Tyr130 was free to swing, as is evident also from the RMSF plot in the region of 130 and 134 (Fig. 2 B2), which shows peaks comparable with those of wt Tudor alone. Therefore, even if a DMR entered the cage, the binding was unstable and the DMR was released quickly. In wt_SmD1_r4, the tail is wrapped around the Tudor, Glu134 is catched (Fig. 2 B and Fig. S8 C) and the DMR is correctly inserted in the site; however the binding becomes stable only when Tyr130 is blocked by the SmD1 tail and stops swinging. Fig. 2 B shows the RMSF of the two parts of r4 trajectory: in the first part Tyr130 is still free to move (Fig. 2 B2 r4a, very high peak of Tyr130), while in the second part the residue is blocked in a stacked position with the inner DMR (Fig. 2 B1 r4b, the Tyr130 peak is almost halved). An example of the two Tyr130 conformations is shown in Fig. 6 B. Also the RMSD per residue shown in Fig. 3 C, with the wide central oscillations of the blue line, confirms this behavior.

Summing up, from the analysis of all the wt trajectories we have extrapolated some indispensable requirements for a stable binding: 1) the acidic residue 104, 105 and 134 must catch the basic DMRs on the SmD1 tail; 2) both the extremities of the SmD1 tail must be wrapped around the surface by means of salt-bridge interactions, in particular with the more acidic side of the surface; 3) the stacking Trp102 / DMR / Y130 must be compact, and the DMR "correctly" oriented inside the cage with the guanidinium group pointing towards Tyr109 (e.g. as in Fig. 4 C); 4) the very flexible cage residue Tyr130, located in the β3-β4 loop, must be blocked by the SmD1 tail to ensure the stability of the cage around the ligand and the stacking previously described (as e.g. in Figs. 4 C and 5). Tyr130, in fact, is the only residue not placed in a beta structure but in a very flexible loop, therefore is capable of wide oscillations (Figs. 2 and 3) that, opening the wall of the cage, highly destabilize the binding of the inner DMR. In most cases, the observed blockage is efficiently obtained by cation-π interactions of the external DMRs with Tyr130 ring. Alternatively, also the steric hindrance of the tail is also sufficient.

A network of H-bonds was reported to involve Tyr127, Glu134 and Gln136 and was hypothesized to be related to the strength of the binding [19]. In the starting structure (Fig. 1, zoom box) Tyr127 is a donor to Glu134, while Gln136 is a donor to both Glu134 and Tyr127. However, in wt-SmD1 trajectories Glu134 protrudes to the solvent to catch the DMRs tail, in this time breaking the H-bond with the other two residues. This is particularly evident in Fig. 7, where the distance of Glu134 as H-bond acceptor from Gln136 and Tyr127 are monitored. In Fig. 7 C, which is related to the wt-SmD1 trajectories, they break for wt-SmD1_r1, r3, r4 and r5, and not, as we said above, for r2. When this happens, Gln136 flips and becomes acceptor from Tyr127. Therefore, only the H-bond between Gln136 and Tyr127 is always present (Fig. 8 C). MD simulations carried out in the absence of the SmD1 tail confirm that the intrinsic flexibility of Glu134 (Fig. S8 A) makes its H-bonds with the other residues of the network not so persistent (Fig. 7 A), even if, in this case, the observed deformation of the barrel affects the whole H-bond network (Fig. 8 A). Such flexibility of Glu134 is in agreement also with its variable orientation observed in the NMR structure ensemble of the free Tudor domain (PDB id code 1G5V [16]) and with the high value of its beta factor in the analogous X-ray structure (PDB id code 1MHN [17]).

Interestingly, the analysis of the trajectories revealed another H-bond, between Glu134 and Ser103, that is present in the starting structure (Fig. 1) and quite persistent in all the simulations both with and without the SmD1 tail (Fig. 9), bringing to four the number of residues involved in the H-bond network. The formation of this H-bond likely allows the cage region to remain compact. However, in the simulations with the Tudor alone this H-bond is not always present (Fig. 9 A, r1), probably because, as we said, the structure is destabilized by the strong oscillations of the terminal regions. In addition, its presence is less needed when the SmD1 tail is able to strongly wrap around the cage region (Fig. 9 C, e.g. r1 and r5 trajectories).

To better compare the stability of the binding and the contribution of specific residues, an analysis of the energy of binding was attempted (see Section 2.2. for details). To compare the results, we selected for the analysis the part of each trajectory in which all the conditions 1 to 4 indicated above are satisfied. In such conditions, all the replicas of the wt complex has comparable binding energy (Table 2), with, as expected, a very high electrostatic contribution.

The per residue energy decomposition was plotted for all wt trajectories to show their similar behavior (Fig. 10 A) and more deeply analyzed for r1 trajectory to make a direct comparison with the E134K-SmD1 complex (Fig. 10 B), as we will discuss later. In Fig. S9 the different per residue contribution to the binding energy are shown. The dashed line in Fig. 10 and Fig. S9 separates the Tudor domain from the SmD1 tail. All the plots confirm the clear preponderant electrostatic interaction, evident from the highly favorable energy of all the negatively charged residues in the Tudor domain and all the positive DMRs in the SmD1 tail. In fact, all the energy absolute values greater than 100 kJ/

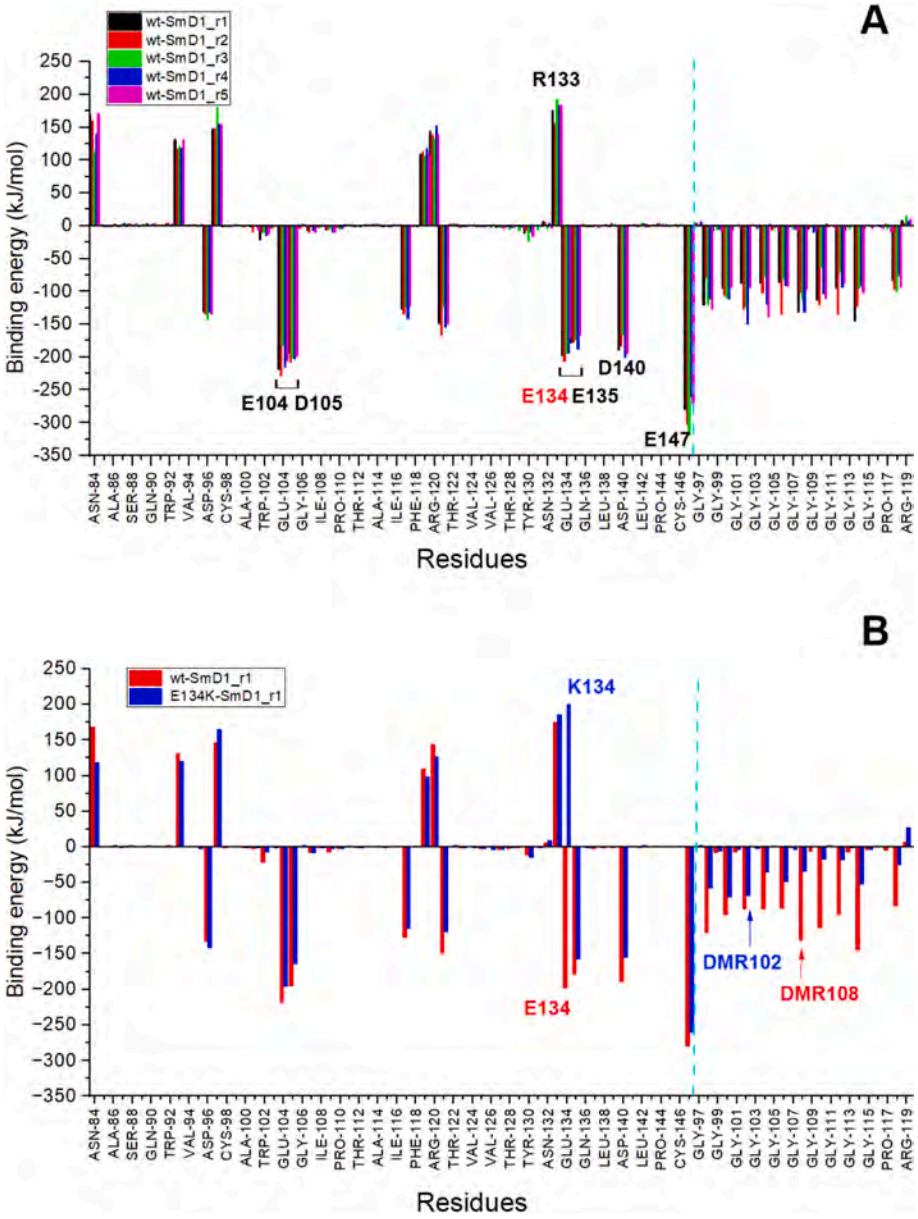


Fig. 10. **(A)** Per residue decomposition of average binding energy for wt trajectories. **(B)** Comparison between the per residue energy decomposition of wt-SmD1_r1 and E134K-SmD1_r1 trajectories. Relevant residues are labelled. The cyan dashed line separates the Tudor domain from the SmD1 tail. All the bars in the Tudor domain with an absolute value greater than 100 kJ/mol correspond to acidic (negative values) or basic (positive values) residues. In panel A, all the bars in the SmD1 tail with a negative value greater than 70 kJ/mol correspond to DMRs or arginine residues. In panel B, the less favored contribution of DMRs and arginines is evident; the two DMRs that enter the cage in the two simulations are indicated: DMR108 in wt-SmD1_r1 trajectory; DMR102 in E134K-SmD1_r1 one.

mol correspond to the charged residues present in the Tudor domain: negative values for acidic ones and positive values for basic ones. In particular, Glu104, Asp105 and Glu134, together with Asp140 and Glu147, have the lowest energy value, confirming their strong involvement in the binding. From the energy plots, also Glu135 shows a strong interaction, but its contact with the SmD1 tail takes place only in the second part of the trajectories, when the tail stabilizes by wrapping around the Tudor domain. The magnitude of DMRs contributions (again with an absolute value greater than 70 kJ/mol) depends on which of them fits into the cage, driving the positioning of the rest of the tail. As an example, in wt_r1 trajectory (Fig. 10 B) DMR108 is the one that enters the cage.

3.3. MD simulations of Tudor-E134K variant in the presence of the SmD1 tail

To investigate the effects of the mutation E134K on the interaction of SMN with SmD1, MD simulations were performed on the Tudor-E134K variant in the presence of the SmD1 tail.

In the first part of the trajectories, a DMR tries to enter the cage (Fig. 11 A), but it is unable to form a correct, stable and compact binding because the tail of DMRs is unable to wrap both their extremities around the cage. In fact, the charge inversion (from −1 to +1) of residue 134 not only causes the loss of the second point of attachment, but also repulses the positively charged DMRs. Residue 134, in fact, is adjacent to the basic residue Arg133, so the mutation E134K highly increases the basicity of that region and, contemporarily, decreases the high acidity of that side of the Tudor surface (Fig. S6 B, yellow arrow). Due to the fact

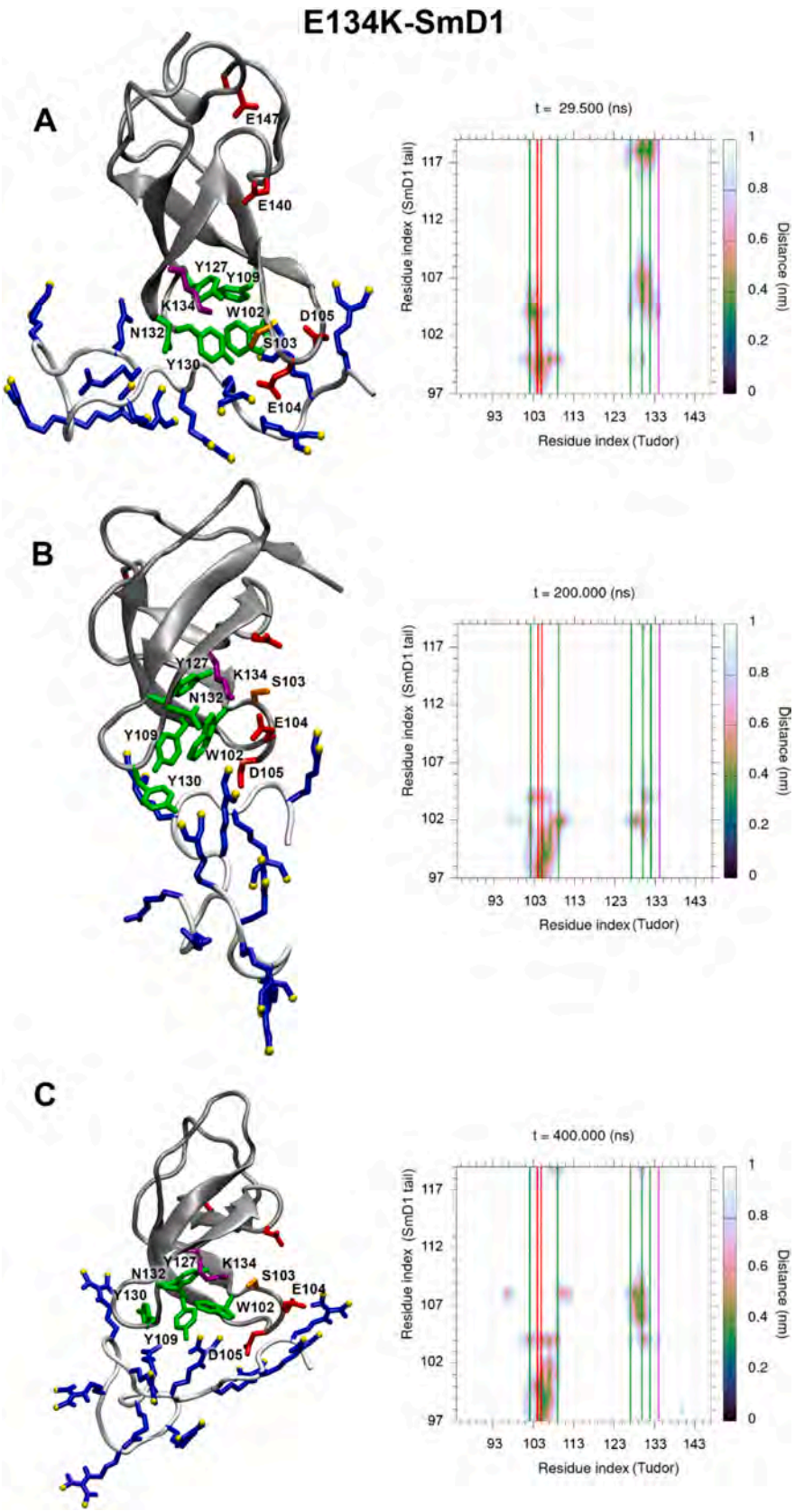


Fig. 11. Highlighted steps of the E134K-SmD1_r1 trajectory next to the corresponding contact maps. Each contact map shows the contact formed between the Tudor and the SmD1 tail at the indicated time. In the Tudor structure shown in cartoon (gray), remarkable residues (in licorice) are labelled and colored as in Fig. 4 (except that here in purple is Lys134), with the addition of Ser103 in orange. The SmD1 tail and the contact maps have the same colour code as in Fig. 4 too.

that the other side of the domain surface does not present an evident character of acidity or basicity, this new electrostatic configuration is unfavorable for the wrapping of the basic SmD1 tail. While the contacts of the N-terminal region of the tail with Glu104 and Asp105 residues in β1-β2 loop are strong and preserved, the tail C-terminal region shows several temporary detachments from the Tudor surface, causing the DMR to leave the site (Fig. 11 B). Then, the C-terminal of the tail tends to roll up or to wrap the surface catching acidic residues as Asp96, Asp140 or Glu147, at the opposite side with respect to Lys134 and Tyr130 (Fig. 11 C and Fig. S7 B). In Fig. 11, the trajectory E134K-SmD1_r1 is used as an example of the behavior described above.

In all the trajectories, Tyr130 is free to swing widely (Fig. 3 D). Such oscillations, together with the interactions of the tail with the residues at the β1-β2 loop region, are able to destabilize all the cage, enlarging the binding cavity sometimes in an irreversible manner, as happens in E134K-SmD1_r3 (Figs. S5, S10 and Fig. 3 D green line).

The oscillations of cage residues appear higher with respect to the wt structures (Figs. 2 C and 3 D), in particular regarding residues Trp102, Tyr130 and Asn132. However, the Tudor fold is stable (Figs. S2 D and S4 D). Lys134 forms, in fact, new interactions with the opposite β1-β2 loop, in particular involving Asp104 (salt-bridge), Trp102 (cation-π) and Ser103 (H-bond); the latter is also more persistent then in the wt structures (Fig. 9) and help to preserve the folding of the domain.

The mutation of residue 134 contemporarily causes a perturbation in the H-bonds network. Lys134 participates only sporadically in the H-bonds triangle with Gln136 and Tyr127 (Fig. 7 B and D), in particular in the presence of the SmD1 tail. On the contrary, its interaction with Ser103 is preserved and likely helps to prevent the unfolding of the cage and likely of the whole structure. As reported above, exception is E134K-SmD1_r3, in which 134–103H-bond breaks from about 160 ns (Fig. 9 D) causing a strong opening of the barrel at the cage side (Fig. S10). In addition, as already observed in the wt structures, the flip of Gln136 in the absence of residue 134 ensures the persistence of Tyr127 → Gln136 H-bond (Fig. 7 B and D). The perturbation of the H-bonds network by E134K mutation was previously hypothesized to be related to the decrease of binding affinity to Sm proteins' tail [19]. Actually, the presence of the fourth interaction with Ser103, together with the results obtained for the wt simulations, showed that the involvement of residue 134 in the network is not so critical, and its flexibility is important for the interaction with the SmD1 tail.

The analysis of the binding energy in the complexes with mutated Tudor was more difficult due to the instability of the binding. Therefore, for comparison with the wt structures, we chose for the calculations only the trajectory E134K-SmD1_r1 that shows, temporarily, a DMR correctly inserted inside the cage and the tail wrapped around the domain. In fact, as for the other trajectories, E134K-SmD1_r2 replica shows a completely uncorrected insertion of the DMR in the cage (pointing towards Trp102 in stacking with Tyr109) that eventually destroys the cage; E134K-SmD1_r3 and _r4 trajectories present only one of these two conditions (when a DMR is inside the cage, the tail is not wrapped and vice versa), while in E134K-SmD1_r5 no DMR is able to enter the cage. The E134K-SmD1_r1 total binding energy is highly less favorable than in the wt complexes. The direct comparison of the wt-SmD1_r1 and the E134K-SmD1_r1 per residue energy decomposition (Fig. 10 B and Fig. S9 B) clearly shows the reversal contribution of mutated residue K134 that, with its positive charge, is very unfavorable to a binding with the highly positively charged DMRs' tail. This affects not only all the contributions of the DMRs, that become less favorable as well (the more negative energy corresponding to the DMR that temporarily enters the cage), but also the van der Waals + electrostatic contribution of the other negatively charged residues (Fig. S9).

4. Conclusions

The interaction between the SmD1 tail and the Tudor domain is primarily electrostatic, driven by the presence of acidic residues near the entrance of the aromatic cage and stabilized by the interactions between the positively charged DMR residues of the SmD1 tail and the negatively charged residues on the Tudor surface. Specifically, Glu104, Asp105, and Glu134 serve as important anchoring points for the DMRs tail and are essential for the initial "jumping" of the salt-bridges. In the native protein, a central DMR of the SmD1 tail rapidly and stably enters the cage, forming a network of both hydrophobic and cation-π interactions, in both stacking and T-shaped orientation. In particular, the stacking of this DMR with Trp102 and Tyr130 and the correct position of DMR pointing towards Tyr 109 contribute highly to stabilize the binding. The other DMRs of the RG-repeats play an active role in stabilizing the complex, wrapping around the acidic surface - and finally the cage region - of the Tudor domain. The stability of the binding is positively correlated with the centrality of the entering DMR, as the two regions of the tail on either side of this DMR wrap around the surface more effectively. In addition, the flexible cage residue Y130 must be blocked from the wrapped tail, to assure that the DMR remains inserted in the binding site.

The E134K mutation results in charge inversion and destabilizes the binding, preventing the attachment of the DMRs tail to the surface and leading to an unstable or absent entrance of a DMR in the binding site. This causes temporary detachments of the C-terminal region of the tail.

Residue E134 was reported to be involved in a hydrogen bonds triangle together with Q136 and Tyr127, the latter belonging to the aromatic cage. The perturbation of the 134–136-127H-bonds network caused by the E134K mutation finally resulted not critical: the intrinsic flexibility of such a residue, evident also in wt protein, is necessary for the interaction with the SmD1 tail. Instead, is the always present interaction between residues 127 and 136 that helps to preserve an excessive flexibility of the cage. A fourth H-bond interaction involving residue 134 with Ser103 is found both in the native and mutated form, which seems to be relevant to avoid the unfolding of the cage region, and definitely of the whole domain, particularly when the SmD1 tail does not wrap around the cage surface.

The loss of a strong Tudor-SmD1 interaction, if by one side causes the loss of a functional splicing machinery, by the other side causes the exposition of the detached Sm tails. This exposure could stimulate the recognition by anti-Sm autoantibodies, as reported in other diseases such as lupus erithematosus [49]. Further studies are now required to investigate whether the exposition of the Sm tails could lead to an autoimmune response, as this would imply new aspects of the pathogenetic mechanism - that leads to neurodegeneration - in some types of SMA.

CRediT authorship contribution statement

Eugenia Polverini: Writing – original draft, Visualization, Supervision, Project administration, Methodology, Formal analysis, Conceptualization. **Pietro Squeri:** Writing – review & editing, Visualization, Investigation, Formal analysis. **Valeria Gherardi:** Writing – review & editing, Investigation, Formal analysis.

Declaration of competing interest

Authors state no conflict of interest.

Data availability

Data will be made available on request.

Acknowledgements

This research benefits from the HPC (High Performance Computing) facility of the University of Parma, Italy. We thank Dr. Martina Simeone for her contribution to the wt Tudor simulations.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijbiomac.2024.133663.

References

[1] D.F. Roberts, J. Chavez, S.D.M. Court, The genetic component in child mortality, Arch. Dis. Child. 45 (1970) 33–38, https://doi.org/10.1136/adc.45.239.33.
[2] S. Lefebvre, L. Bürglen, S. Reboullet, O. Clermont, P. Burlet, L. Viollet, B. Benichou, C. Cruaud, P. Millasseau, M. Zeviani, Identification and characterization of a spinal muscular atrophy-determining gene, Cell 80 (1995) 155–165, https://doi.org/10.1016/0092-8674(95)90460-3.
[3] M.R. Lunn, C.H. Wang, Spinal muscular atrophy, Lancet 371 (2008) 2120–2133, https://doi.org/10.1016/S0140-6736(08)60921-6.
[4] L. Cartegni, A.R. Krainer, Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1, Nat. Genet. 30 (2002) 377–384, https://doi.org/10.1038/ng854.
[5] C.L. Lorson, E.J. Androphy, An exonic enhancer is required for inclusion of an essential exon in the SMA-determining gene SMN, Hum. Mol. Genet. 9 (2000) 259–265, https://doi.org/10.1093/hmg/9.2.259.
[6] U.R. Monani, C.L. Lorson, D.W. Parsons, T.W. Prior, E.J. Androphy, A.H. Burghes, J.D. McPherson, A single nucleotide difference that alters splicing patterns distinguishes the SMA gene SMN1 from the copy gene SMN2, Hum. Mol. Genet. 8 (1999) 1177–1183, https://doi.org/10.1093/hmg/8.7.1177.
[7] M.S. Weiss, K. Diederichs, R.J. Read, S. Panjikar, G.D. Van Duyne, A.G. Matera, U. Fischer, C. Grimm, A critical examination of the recently reported crystal structures of the human SMN protein, Hum. Mol. Genet. 25 (2016) 4717–4725, https://doi.org/10.1093/hmg/ddw298.
[8] S. Lefebvre, P. Burlet, Q. Liu, S. Bertrandy, O. Clermont, A. Munnich, G. Dreyfuss, J. Melki, Correlation between severity and SMN protein level in spinal muscular atrophy, Nat. Genet. 16 (1997) 265–269, https://doi.org/10.1038/ng0797-265.
[9] Q. Liu, G. Dreyfuss, A novel nuclear structure containing the survival of motor neurons protein, EMBO J. 15 (1996) 3555–3565.
[10] D.D. Coovert, T.T. Le, P.E. McAndrew, J. Strasswimmer, T.O. Crawford, J. R. Mendell, S.E. Coulson, E.J. Androphy, T.W. Prior, A.H. Burghes, The survival motor neuron protein in spinal muscular atrophy, Hum. Mol. Genet. 6 (1997) 1205–1214, https://doi.org/10.1093/hmg/6.8.1205.
[11] G. Battaglia, A. Princivalle, F. Forti, C. Lizier, M. Zeviani, Expression of the SMN gene, the spinal muscular atrophy determining gene, in the mammalian central nervous system, Hum. Mol. Genet. 6 (1997) 1961–1971, https://doi.org/10.1093/hmg/6.11.1961.
[12] Q. Liu, U. Fischer, F. Wang, G. Dreyfuss, The spinal muscular atrophy disease gene product, SMN, and its associated protein SIP1 are in a complex with spliceosomal snRNP proteins, Cell 90 (1997) 1013–1021, https://doi.org/10.1016/S0092-8674(00)80367-0.
[13] R. Zhang, B.R. So, P. Li, J. Yong, T. Glisovic, L. Wan, G. Dreyfuss, Structure of a key intermediate of the SMN complex reveals Gemin2's crucial function in snRNP assembly, Cell 146 (2011) 384–395, https://doi.org/10.1016/j.cell.2011.06.043.
[14] K.L. Sarachan, K.G. Valentine, K. Gupta, V.R. Moorman, J.M. Gledhill, M. Bernens, C. Tommos, A.J. Wand, G.D. Van Duyne, Solution structure of the core SMN-Gemin2 complex, Biochem. J. 445 (2012) 361–370, https://doi.org/10.1042/BJ20120241.
[15] R. Martin, K. Gupta, N.S. Ninan, K. Perry, G.D. Van Duyne, The survival motor neuron protein forms soluble glycine zipper oligomers, Structure 20 (2012) 1929–1939, https://doi.org/10.1016/j.str.2012.08.024.
[16] P. Selenko, R. Sprangers, G. Stier, D. Bühler, U. Fischer, M. Sattler, SMN Tudor domain structure and its interaction with the Sm proteins, Nat. Struct. Biol. 8 (2001) 27–31, https://doi.org/10.1038/83014.
[17] R. Sprangers, M.R. Groves, I. Sinning, M. Sattler, High-resolution X-ray and NMR structures of the SMN tudor domain: conformational variation in the binding site for symmetrically dimethylated arginine residues, J. Mol. Biol. 327 (2003) 507–520, https://doi.org/10.1016/S0022-2836(03)00148-7.
[18] Y. Liu, A. Iqbal, W. Li, Z. Ni, Y. Wang, J. Ramprasad, K.J. Abraham, M. Zhang, D. Y. Zhao, S. Qin, P. Loppnau, H. Jiang, X. Guo, P.J. Brown, X. Zhen, G. Xu, K. Mekhail, X. Ji, M.T. Bedford, J.F. Greenblatt, J. Min, A small molecule antagonist of SMN disrupts the interaction between SMN and RNAP II, Nat. Commun. 13 (2022), https://doi.org/10.1038/s41467-022-33229-5.
[19] K. Tripsianes, T. Madl, M. Machyna, D. Fessas, C. Englbrecht, U. Fischer, K. M. Neugebauer, M. Sattler, Structural basis for dimethylarginine recognition by the Tudor domains of human SMN and SPF30 proteins, Nat. Struct. Mol. Biol. 18 (2011) 1414–1420, https://doi.org/10.1038/nsmb.2185.
[20] R.N. Singh, M.D. Howell, E.W. Ottesen, N.N. Singh, Diverse role of survival motor neuron protein, Biochim. Biophys. Acta - Gene Regul. Mech. 2017 (1860) 299–315, https://doi.org/10.1016/j.bbagrm.2016.12.008.
[21] Y. Song, F. Dimaio, R.Y.R. Wang, D. Kim, C. Miles, T. Brunette, J. Thompson, D. Baker, High-resolution comparative modeling with RosettaCM, Structure 21 (2013) 1735–1742, https://doi.org/10.1016/j.str.2013.08.005.
[22] M. Varadi, D. Bertoni, P. Magana, U. Paramval, I. Pidruchna, M. Radhakrishnan, M. Tsenkov, S. Nair, M. Mirdita, J. Yeo, O. Kovalevskiy, K. Tuny, A. Žídek, H. Tomlinson, D. Hariharan, J. Abrahamson, T. Green, J. Jumper, D. Hassabis, S. Velankar, AlphaFold Protein Structure Database in 2024:, (2024) 1–8. doi:https://doi.org/10.1093/nar/gkab1061.
[23] H. Chaytow, Y.T. Huang, T.H. Gillingwater, K.M.E. Faller, The role of survival motor neuron protein (SMN) in protein homeostasis, Cell. Mol. Life Sci. 75 (2018) 3877–3894, https://doi.org/10.1007/s00018-018-2849-1.
[24] A.H.M. Burghes, C.E. Beattie, Spinal muscular atrophy: why do low levels of survival motor neuron protein make motor neurons sick? Nat. Rev. Neurosci. 10 (2009) 597–609, https://doi.org/10.1038/nrn2670.
[25] L. Pellizzoni, N. Kataoka, B. Charroux, G. Dreyfuss, A novel function for SMN, the spinal muscular atrophy disease gene product, in pre-mRNA splicing, Cell 95 (1998) 615–624, doi:S0092-8674(00)81632-3.
[26] S. Paushkin, A.K. Gubitz, S. Massenet, G. Dreyfuss, The SMN complex, an assemblyosome of ribonucleoproteins, Curr. Opin. Cell Biol. 14 (2002) 305–312, https://doi.org/10.1016/S0955-0674(02)00332-0.
[27] F.M. Boisvert, J. Côté, M.C. Boulanger, P. Cléroux, F. Bachand, C. Autexier, S. Richard, Symmetrical dimethylarginine methylation is required for the localization of SMN in Cajal bodies and pre-mRNA splicing, J. Cell Biol. 159 (2002) 957–969, https://doi.org/10.1083/jcb.200207028.
[28] D. Yanling Zhao, G. Gish, U. Braunschweig, Y. Li, Z. Ni, F.W. Schmitges, G. Zhong, K. Liu, W. Li, J. Moffat, M. Vedadi, J. Min, T.J. Pawson, B.J. Blencowe, J. F. Greenblatt, SMN and symmetric arginine dimethylation of RNA polymerase II C-terminal domain control termination, Nature 529 (2016) 48–53, https://doi.org/10.1038/nature16469.
[29] D. Bühler, V. Raker, R. Lührmann, U. Fischer, Essential role for the tudor domain of SMN in spliceosomal U snRNP assembly: implications for spinal muscular atrophy, Hum. Mol. Genet. 8 (1999) 2351–2357, https://doi.org/10.1093/hmg/8.13.2351.
[30] W.J. Friesen, G. Dreyfuss, Specific sequences of the Sm and Sm-like (Lsm) proteins mediate their interaction with the spinal muscular atrophy disease gene product (SMN), J. Biol. Chem. 275 (2000) 26370–26375, https://doi.org/10.1074/jbc.M003299200.
[31] W.J. Friesen, S. Paushkin, A. Wyce, S. Massenet, G.S. Pesiridis, G. Van Duyne, J. Rappsilber, M. Mann, G. Dreyfuss, The methylosome, a 20S complex containing JBP1 and pICln, produces dimethylarginine-modified Sm proteins, Mol. Cell. Biol. 21 (2001) 8289–8300, https://doi.org/10.1128/MCB.21.24.8289-8300.2001.
[32] H. Brahms, L. Meheus, V. De Brabandere, U. Fischer, R. Lührmann, Symmetrical dimethylation of arginine residues in spliceosomal Sm protein B/B′ and the Sm-like protein LSm4, and their interaction with the SMN protein, Rna 7 (2001) 1531–1542, https://doi.org/10.1017/S135583820101442X.
[33] P. Mohaghegh, N.R. Rodrigues, N. Owen, C.P. Ponting, T.T. Le, A.H.M. Burghes, K. E. Davies, Analysis of mutations in the tudor domain of the survival motor neuron protein SMN, Eur. J. Hum. Genet. 7 (1999) 519–525, https://doi.org/10.1038/sj.ejhg.5200346.
[34] S. Lefebvre, L. Bürglen, J. Frézal, A. Munnich, J. Melki, The role of the SMN gene in proximal spinal muscular atrophy, Hum. Mol. Genet. 7 (1998) 1531–1536, https://doi.org/10.1093/hmg/7.10.1531.
[35] I. Angilletta, R. Ferrante, R. Giansante, L. Lombardi, A. Babore, A. Dell'Elice, E. Alessandrelli, S. Notarangelo, M. Ranaudo, C. Palmarini, V. De Laurenzi, L. Stuppia, C. Rossi, Spinal muscular atrophy: an evolving scenario through new perspectives in diagnosis and advances in therapies, Int. J. Mol. Sci. 24 (2023), https://doi.org/10.3390/ijms241914873.
[36] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I. N. Shindyalov, P.E. Bourne, The protein data bank, Nucleic Acids Res. 28 (2000) 235–242.
[37] N. Guex, M.C. Peitsch, SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling, Electrophoresis 18 (1997) 2714–2723.
[38] C. Kambach, S. Walke, R. Young, J.M. Avis, E. De La Fortelle, V.A. Raker, R. Lührmann, J. Li, K. Nagai, Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs, Cell 96 (1999) 375–387, https://doi.org/10.1016/S0092-8674(00)80550-4.
[39] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38, https://doi.org/10.1016/0263-7855(96)00018-5.
[40] D. van der Spoel, E. Lindhal, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: fast, flexible and free, J. Comput. Chem. 26 (2005) 1701–1718, https://doi.org/10.1002/jcc.20291.
[41] C. Oostenbrink, A. Villa, A.E. Mark, W.F. van Gunsteren, A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6, J. Comput. Chem. 25 (2004) 1656–1676, https://doi.org/10.1002/jcc.20090.
[42] K.B. Koziara, M. Stroet, A.K. Malde, A.E. Mark, Testing and validation of the automated topology builder (ATB) version 2.0: prediction of hydration free enthalpies, J. Comput. Aided Mol. Des. 28 (2014) 221–233, https://doi.org/10.1007/s10822-014-9713-7.
[43] J. Gasteiger, M. Marsili, Iterative partial equalization of orbital electronegativity - a rapid access to atomic charges, Tetrahedron 36 (1980) 3219–3228, https://doi.org/10.1016/0040-4020(80)80168-2.
[44] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38, https://doi.org/10.1016/0263-7855(96)00018-5.
[45] D. Mercadante, F. Gräter, C. Daday, CONAN: a tool to decode dynamical information from molecular interaction maps, Biophys. J. 114 (2018) 1267–1273, https://doi.org/10.1016/j.bpj.2018.01.033.

[46] R. Kumari, R. Kumar, A. Lynn, g_mmpbsa-a GROMACS tool for high-throughput MM-PBSA calculations, J. Chem. Inf. Model. 54 (2014) 1951–1962, https://doi.org/10.1021/ci500020m.
[47] N.A. Baker, D. Sept, S. Joseph, M.J. Holst, J.A. McCammon, Electrostatics of nanosystems: application to microtubules and the ribosome, Proc. Natl. Acad. Sci. USA 98 (2001) 10037–10041, https://doi.org/10.1073/pnas.181342398.
[48] B.R. Miller 3rd, T.D.J. McGee, J.M. Swails, N. Homeyer, H. Gohlke, A.E. Roitberg, MMPBSA.Py: an efficient program for end-state free energy calculations, J. Chem. Theory Comput. 8 (2012) 3314–3321, https://doi.org/10.1021/ct300418h.
[49] H. Brahms, J. Raymackers, A. Union, F. de Keyser, L. Meheus, R. Lührmann, The C-terminal RG dipeptide repeats of the spliceosomal Sm proteins D1 and D3 contain symmetrical dimethylarginines, which form a major B-cell epitope for anti-Sm autoantibodies, J. Biol. Chem. 275 (2000) 17122–17129, https://doi.org/10.1074/jbc.M000300200.