

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Structural and Functional Consequences of the SMA-Linked Missense Mutations of the Survival Motor Neuron Protein: A Brief Update

Wei Li

Abstract

Genetically linked to the survival motor neuron 1 gene *SMN1*, spinal muscular atrophy (SMA) is an autosomal recessive neuromuscular disease with dysfunctional α -motor neurons. As the product of the *SMN1* gene, the survival motor neuron protein (SMN) plays an essential role in the molecular pathogenesis of SMA. On 1 June 2017, a PLoS ONE article reported a set of computational structural analysis to illustrate how do SMA-linked mutations of *SMN1* lead to structurally/functionally deficient variants of SMN. Following this article, this chapter provides a brief update of the structural and functional consequences of the missense mutations of this SMA protein.

Keywords: spinal muscular atrophy, survival motor neuron protein, missense mutation, structural consequence(s), functional consequence(s)

1. Setting the scene up

On 1 June 2017, *PLoS ONE* published an original research article (**Figure 1**) [1] with a title ‘How do SMA-linked mutations of *SMN1* lead to structural/functional deficiency of the SMA protein?’, of which this chapter aims to provide a brief update.

1.1 The genetics of SMA: a brief introduction

SMA is an autosomal recessive neuromuscular disease with α -motor neuron (anterior horn of the spinal cord) dysfunction and muscular atrophy [2]. SMA is caused by loss (~95% of SMA cases) or mutation (~5% of SMA cases) of the survival motor neuron gene 1 *SMN1* (telomeric *SMN*, *telSMN* or *SMN1*, GenBank: U18423, the 5q13 region of human chromosome) [3]. In the 5q13 region of the human chromosome, there is also a nearly identical *survival motor neuron 2* gene *SMN2* (centromeric *SMN*, *cenSMN* or *SMN2*, GenBank: NM_022875) [3]. The two genes (*SMN1* and *SMN2*) have been extensively characterised, and their roles in SMA have been reviewed in detail [2–8].

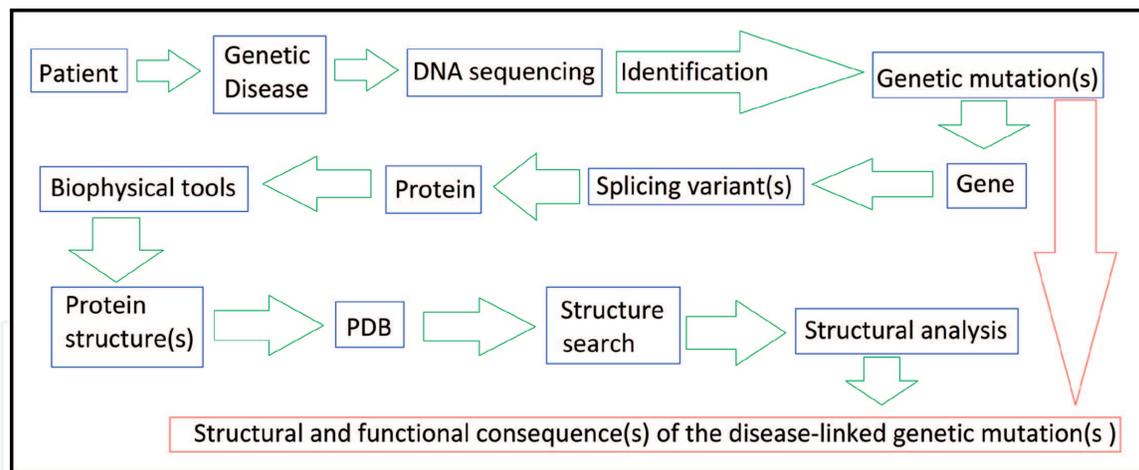


Figure 1.

A flow chart for the computational analysis of structural/functional consequences of clinically identified genetic diseases-linked missense mutation(s) of key gene(s) and protein(s). In [1], SMA was shown as an example of the computational analysis as illustrated here in this figure (<http://biomedical-advances.org/ep-20182-14/>).

1.2 The survival motor neuron protein and its role in SMA

The survival motor neuron (SMN) protein is the product of *SMN1*, the SMA-determining survival motor neuron gene [2, 3]. As a result, SMN is also called the SMA protein. In fact, the 38-kD SMN is the actually affected protein in SMA [9–11], and is a cytoplasmic protein that also occurs in dot-like nuclear structures called gems, which is why SMN is formerly termed Gemin1 [3, 12], too.

In the molecular pathogenesis of SMA, of particular interest is an exon 7-skipping splicing defect identified in the pre-mRNA editing of the *SMN2* gene [5]. Due to this splicing defect, *SMN2* predominantly produces exon 7-skipped transcripts, which encode a truncated isoform of the SMN protein (SMN Δ 7 or SMN2 with 282 residues), in comparison with the full-length SMN protein with 294 residues (SMN1 or FL-SMN).

In pre-mRNA editing, spliceosome is the major functional unit, and spliceosomal small nuclear ribonucleoproteins (snRNPs) are essential components of the nuclear pre-mRNA processing machinery [13–17]. In the pathogenesis of SMA, the SMN protein plays a critical role in pre-mRNA processing, because the biogenesis of spliceosomal snRNPs is promoted by the SMN complex [14, 18, 19], which consists of SMN (Gemin1), Gemin2–8 and UNR-interacting protein (UNRIP) [13, 16, 20]. In the formation of the SMN complex, SMN forms oligomers and directly interacts via its N-terminus with Gemin2 and via its tudor domain with spliceosomal (Sm) proteins [13, 21, 22]. A key component of the SMN complex, SMN first assembles the essential SMN/Gemin complex, which in turn mediates the formation of the Sm core domain of the spliceosomal snRNPs [13, 21, 22].

2. Structural and functional consequences of the SMA-linked missense mutations of SMN

In general, genetic mutation includes missense, nonsense, insertion and deletion mutations. A nonsense mutation is a point mutation in a DNA sequence that results in a premature stop codon, or a nonsense codon in the transcribed mRNA, and in a truncated, incomplete and usually functionally deficient protein product. In contrast, a missense mutation involves substitution of one single amino acid residue, and therefore is able to provide unique access to residue-specific structural insights

into the role of the residue in the structure and function of the target protein, provided that the three-dimensional structure of the target protein is experimentally determined and deposited in the Protein Data Bank. Thus, this chapter focuses on SMA-linked missense mutations of SMN and aims to provide a brief update of their structural and functional consequences with a set of computational structural analysis as described in [1].

2.1 An update of SMA-linked missense mutations of SMN

A set of point mutations (missense and nonsense mutations) have been previously summarised in [1], including A2G [23], nonsense mutation Q15X [24], D30N [25], D44V [25–27], V94G [28], G95R [25], Y130C [29], nonsense mutation Q157X [30], A188S [31], nonsense mutation W190X [32], nonsense mutation L228X [33], P245L [34], L260S [28], S262G and S262I [4, 25], M263T [32], S266P [29], Y272C [4, 35, 36], H273R [29], T274I [4, 35, 36], G275S [32], G279C and G279V [4, 35, 37, 38]. As of 25 September 2018, eight more missense mutations of SMN were summarised and reported, including A2V, Y109C, Y130C, Y130H, P221L, S230L, P244L and R288S [39].

2.2 An update of experimentally determined SMN-related structures

In [1], 11 SMN-related structures were retrieved from the PDB database [40] with 2 search parameters (text search for: survival motor neuron protein and molecule: survival motor neuron protein). In a new search of the PDB database (accessed 25 September 2018) [40] with the same parameters, 14 PDB entries were retrieved, including 1G5V, 1MHN, 2LEH, 4A4E, 4A4G, 4GLI, 4QQ6, 4V98, 5XJL, 5XJQ, 5XJR, 5XJS, 5XJT and 5XJU. In a comparison with the PDB entries in [1], during the past 16 months, six new SMN-related structures were deposited in the Protein Data Bank, including 5XJL (to supersede 3S6N [41]), 5XJQ [42], 5XJR [42], 5XJS [42], 5XJT [42] and 5XJU [42]. While the six PDB entries do contain a set of different yet functionally related protein molecules, including snRNP Sm-D1, snRNP Sm-D2, snRNP E, snRNP F and snRNP G, they also contain a fragment of the survival motor neuron protein (SMN residues 26–62), according to the fasta format data of the six PDB entries [42].

2.3 An update of the structural and functional consequences of the missense mutations of SMN

2.3.1 Asp44 in the Gemin2-binding domain of SMN

In light of the six new experimentally determined SMN-related structures (Table 1), a new set of computational structural analysis, as previously described in detail in [1], is within the reach of this chapter to provide an update of it. Two aspartates (Asp 35 and Asp44) of SMN stood out in the structural analysis of both intramolecular and intermolecular salt bridges for this SMA protein, as listed in Table 2.

Asp44 is in the exon 2a of *SMN1* (the Gemin2-binding domain), and involved in an SMA-linked Asp44Val (D44V) missense mutation [25], which involves a substitution of Asp44's charged side chain by Val44's hydrophobic side chain. Of extraordinary functional significance is that SMN's Gemin2-binding activity is totally suppressed by the D44V mutation in *SMN1* [41]. Moreover, the D44V SMN (SMND44V) mutant's snRNP assembly activity is lower than that of the wild-type SMN (FL-SMN or SMN1) [27].

PDB ID	Structure title	Method	Release date
5XJL	Crystal structure of the Gemin2-binding domain of SMN, Gemin2 in complex with SmD1/D2/F/E/G from human	X-ray	2 May 2018
5XJQ	Crystal structure of the Gemin2-binding domain of SMN, Gemin2 in complex with SmD1(1-82)/D2/F/E/G from human	X-ray	4 July 2018
5XJR	Crystal structure of the Gemin2-binding domain of SMN, Gemin2dN39 in complex with SmD1(1-82)/D2/F/E/G from human	X-ray	4 July 2018
5XJS	Crystal structure of the Gemin2-binding domain of SMN, Gemin2dN39 in complex with SmD1(1-82)/D2/F/E from human	X-ray	4 July 2018
5XJT	Crystal structure of the Gemin2-binding domain of SMN, Gemin2 in complex with SmD1(1-82)/D2.R61A/F/E/G from human	X-ray	4 July 2018
5XJU	Crystal structure of the Gemin2-binding domain of SMN, Gemin2dN39 in complex with SmD1(1-82)/D2.R61A/F/E/G from human	X-ray	4 July 2018

In this table, X-ray represents X-ray crystallography as a biophysical tool for biomolecular structure determination.

Table 1.

A list of new (compared with those summarised in [1]) experimentally determined SMN-related structures as of 25 September 2018 [40].

In a solid alignment with the computational analysis in [1], a set of salt bridges were structurally identified between SMN's Asp44 (M_Asp_44) and Gemin2's Arg213 (2_Arg_213), as shown in **Table 2**. In particular, four intermolecular salt bridges were identified between the buried side chains (**Table 3**) of these two charged residues, i.e. according to the coordinates data in the PDB entry **5XJL** [42], as shown in **Figure 2**.

Taken together, it is conceivable that the buried side chains of SMN's Asp44 and Gemin2's Arg213 form a salt bridge, which constitutes a favourable electrostatic energy contribution to the SMN-Gemin2 complex structural stability [41], and highlights the functionally indispensable roles of the two residues' charged side chains, considering the experimental observation that the SMN-Gemin2 binding is abrogated by the D44V mutation [41], resulting in a functionally deficient SMA-linked D44V SMN mutant.

In addition to the intermolecular salt bridges formed between SMN's Asp44 and Gemin2's Arg213, a set of intramolecular salt bridges were also identified between side chains of SMN's Asp35 and Lys41 (**Table 2**), which was reported in [1], too, where 15 salt bridges were identified between the side chains of SMN's Asp35 and Lys41 in the salt bridge analysis of the NMR-determined SMN-Gemin2 complex ensemble (PDB ID: 2LEH) [22, 41]. In SMN, Lys41 is a positively charged residue and also a neighbouring residue of Asp44. Functionally different to the SMA-linked D44V mutation, a Lys41Ala (K41A) mutation (not SMA-linked) does not affect SMN-Gemin2 binding [41]. Thus, in another solid agreement with the structural analysis in [1], the structural analysis highlights that the salt bridges between SMN's Asp35 and Lys41 are intramolecular, i.e. within the apo SMN protein, instead of intermolecular, i.e. at the SMN-Gemin2 complex structure interface, which help to explain why the Lys41Ala (K41A) mutation is not SMA-linked [41].

Overall, there is a solid agreement between the old [1] and the new (this chapter) sets of computational structural analysis for both NMR and X-ray SMN-related structures, reflecting the technical maturity of the two main biophysical tools for biomolecular structure determination, particularly in light of the booming number of cryo-electron microscopy (cryo-EM) images uploaded to the Electron

PDB ID	SBnum	Residue A	Atom A	Residue B	Atom B	Distance (Å)
5XJL	4	M_ASP_44	OD1	2_ARG_213	NH1	2.946 (Yellow)
5XJL	4	M_ASP_44	OD1	2_ARG_213	NH2	3.579 (Red)
5XJL	4	M_ASP_44	OD2	2_ARG_213	NH1	3.236 (Brown)
5XJL	4	M_ASP_44	OD2	2_ARG_213	NH2	3.848 (Blue)
5XJQ	3	M_ASP_44	OD1	2_ARG_213	NH1	2.760
5XJQ	3	M_ASP_44	OD1	2_ARG_213	NH2	3.593
5XJQ	3	M_ASP_44	OD2	2_ARG_213	NH1	2.968
5XJR	2	M_ASP_44	OD1	2_ARG_213	NH1	2.385
5XJR	2	M_ASP_44	OD2	2_ARG_213	NH1	2.871
5XJS	3	M_ASP_44	OD1	2_ARG_213	NH1	3.078
5XJS	3	M_ASP_44	OD1	2_ARG_213	NH2	3.670
5XJS	3	M_ASP_44	OD2	2_ARG_213	NH1	2.631
5XJT	3	M_ASP_44	OD1	2_ARG_213	NH1	2.335
5XJT	3	M_ASP_44	OD1	2_ARG_213	NH2	3.386
5XJT	3	M_ASP_44	OD2	2_ARG_213	NH1	3.067
5XJU	2	M_ASP_44	OD1	2_ARG_213	NH1	2.302
5XJU	2	M_ASP_44	OD2	2_ARG_213	NH1	2.989
5XJQ	1	M_ASP_35	OD1	M_LYS_41	NZ	3.921
5XJS	2	M_ASP_35	OD1	M_LYS_41	NZ	3.670
5XJS	2	M_ASP_35	OD2	M_LYS_41	NZ	3.803
5XJT	2	M_ASP_35	OD1	M_LYS_41	NZ	2.416
XJT	2	M_ASP_35	OD2	M_LYS_41	NZ	2.931
5XJU	1	M_ASP_35	OD1	M_LYS_41	NZ	3.274

In this table, the residue naming scheme is **Chain ID_residue name_residue number**, SBnum represents the number of salt bridges computationally identified from the PDB entries listed in this table. In the top four rows for PDB entry 5XJL, Yellow, Red, Brown and Blue represent the colouring scheme for Figure 2. Distance represents the distance between two oppositely charged groups/atoms in Å.

Table 2.

A summary of salt bridge analysis of the six new SMN-related structures as of 25 September 2018 [40].

Microscopy Data Bank (EMDB), where a long way is there to go still for cryo-EM to match NMR spectroscopy and X-ray crystallography in terms of technical maturity and the urgent need of tools for structural model quality validation [45].

2.3.2 Gly95 in the SMN tudor domain

Although not located in the structurally determined region of the six new structures (Table 1), Gly95 is a residue in the SMN tudor domain, and it is involved in a Gly95Arg (G95R) mutation [25]. This G95R mutation significantly reduces SMN's ability to bind Sm proteins, such as Sm-B and Sm-D1 [25], confirming that tudor domain is the essential binding site of SMN to Sm proteins.

In a further inspection of the computational analysis as reported in [1], no salt bridge or hydrogen bond was identified for Gly95. Nonetheless, in the SMN tudor domain NMR ensemble [46], between the side chains of Asp96 and Lys93, 1 salt bridge was found for PDB ID 1G5V [46] with 10 structure models, 18 salt bridges

Residue	SASA (Å ²)	SASA-intrinsic (Å ²)	SASA-Ratio
2_Arg_213	57	238.76	0.238
M_Asp_44	67	140.39	0.477

In this table, SASA, SASA-intrinsic and SASA-ratio represent for SMN's Asp44 and Gemin2's Arg213 the average SASA value calculated by DSSP [43], the intrinsic SASA value [44] and the ratio of SASA divided by SASA-intrinsic, respectively. In this table, the residue naming scheme is Chain ID_residue name_residue number.

Table 3. Solvent accessible surface area (SASA) values of SMN's Asp44 and Gemin2's Arg213 (PDB ID: 5XJL) [42].

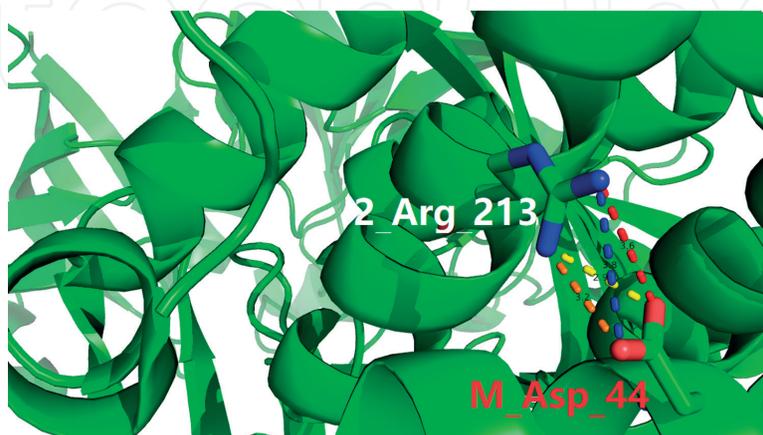


Figure 2. Four salt bridges formed between the buried side chains of SMN's Asp44 (M_Asp_44 in red text) and Gemin2's Arg213 (2_Arg_213 in white text). In this figure, the residue naming scheme is Chain ID_residue name_residue number. In this figure, Asp44's side chain oxygens are coloured red, and Arg213's nitrogen atoms are coloured blue, while all hydrogen atoms are coloured in white, the four dotted lines in four colours represent the four side chain salt bridges formed between the two oppositely charged residues, where the colouring scheme is described in Table 2.

were found for PDB ID 4A4E [47] with 20 structure models (Figure 3) and 16 salt bridges were found for PDB ID 4A4G [47] with 20 structure models. Similarly, 15 salt bridges were also identified between the side chains of Glu147 and Lys97 of SMN (PDB ID: 4A4G [47], with 20 structure models), with the distance between 2 oppositely charged groups being 2.93 ± 0.39 Å.

Quite interestingly, Gly95 sits right between the two oppositely charged neighbouring residues (Asp96 and Lys93), which are the only two charged residues in the

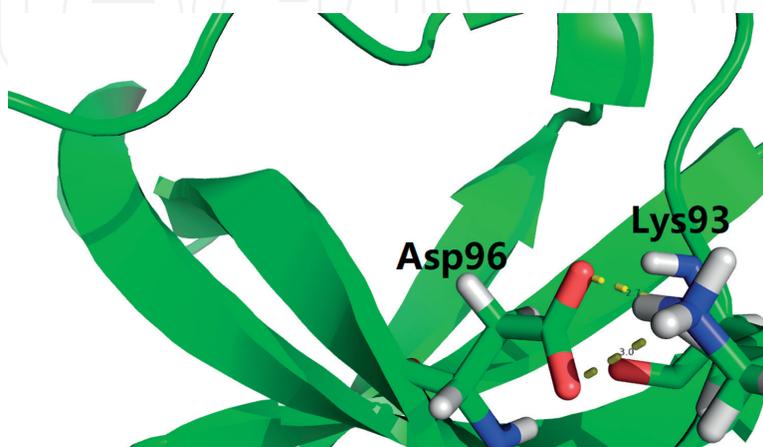


Figure 3. Two salt bridges formed between the side chains of SMN's Asp96 and Lys93 (shown as sticks here) according to a salt bridge analysis of the third structural model of the NMR ensemble (PDB ID 4A4E) [47]. In this figure, Asp96's side chain oxygens are coloured red, and Lys93's side chain nitrogen is coloured blue, while all hydrogen atoms are coloured in white.

tudor domain that are in the spatial proximity of Gly95. Thus, it is conceivable that a G95R mutation disrupts the Asp96-Lys93 salt bridge and/or builds another one (possibly even stronger) between the side chains of Lys95 and Asp96, which either perturbs the structure-stabilising activity of the Asp96-Lys93 salt bridge, and/or makes it energetically more unfavourable for Asp96's side chain to orient towards positively charged side chains in Sm proteins and thereby affect the binding of SMN to Sm proteins. While the potential local electrostatic interaction disruption mechanism here for this SMA-linked G95R mutation is similar to that of the E134K and the Q136E mutations of SMN [1], the former mechanism is dependent on the occurrence of energetically unfavourable electrostatic interaction(s), but the latter mechanism is dependent on the loss of energetically favourable electrostatic interaction(s) for local structural stability of the SMN tudor domain, the essential part of SMN for the Sm protein-binding, which can help explain the reduced Sm core assembly activity of the two SMA-linked SMNE134K and SMNQ136E mutants.

2.3.3 Y109C, Y130C and Y130H in the SMN tudor domain

Among the eight SMN residues with SMA-linked missense mutations [39], only Y109 and Y130 are located in the structurally determined region of SMN [1], according to the updated list of SMN-related structures as of 25 September 2018. Although Y109C, Y130C and Y130H are not located in the structurally determined region of the six new structures, the three missense mutations are located in the structurally determined region of the experimentally determined structures [1].

Tyr130 is a tudor domain hydrophobic residue with a Tyr130Cys (Y130C) mutation [29]. In the computational analysis in [1], no salt bridge or hydrogen bond was identified for Tyr130. Nonetheless, Tyr130 is ~50% buried, with an SASA value of $111.1 \pm 4.18 \text{ \AA}^2$ compared with its standard SASA value at 212.7 \AA^2 , while Tyr109 is deeply buried, with an SASA value of $61.1 \pm 8.43 \text{ \AA}^2$ compared with its standard SASA value at 212.7 \AA^2 . Taken together, the SASA analysis of the three SMA-linked mutations highlights the potential significance of the deeply buried hydrophobic side chains of Tyr109 and Tyr130 in the SMN tudor domain.

PDB File	Acceptor (A)	Donor (D)	Hydrogen (H)	D-A (Å)	H-A (Å)	$\angle ADH$ (*)
0.pdb	OD2, A_ASP_105	OH, A_TYR_109	HH, A_TYR_109	2.73	1.80	13.75
3.pdb	OD2, A_ASP_105	OH, A_TYR_109	HH, A_TYR_109	2.69	1.77	15.61
4.pdb	OD2, A_ASP_105	OH, A_TYR_109	HH, A_TYR_109	2.67	1.72	10.96
5.pdb	OD2, A_ASP_105	OH, A_TYR_109	HH, A_TYR_109	2.77	1.86	17.26
6.pdb	OD2, A_ASP_105	OH, A_TYR_109	HH, A_TYR_109	2.71	1.78	15.09
8.pdb	OD2, A_ASP_105	OH, A_TYR_109	HH, A_TYR_109	2.78	1.87	16.14
12.pdb	OD2, A_ASP_105	OH, A_TYR_109	HH, A_TYR_109	2.83	1.95	20.70
14.pdb	OD2, A_ASP_105	OH, A_TYR_109	HH, A_TYR_109	2.71	1.78	13.76
18.pdb	OD2, A_ASP_105	OH, A_TYR_109	HH, A_TYR_109	2.71	1.76	10.91
19.pdb	OD2, A_ASP_105	OH, A_TYR_109	HH, A_TYR_109	2.63	1.70	13.36

In this table, the names of the PDB files correspond to the single NMR structural model split from the NMR ensemble (PDB entry 4A4E) by a tcl script [1], the residue naming scheme is Chain ID_residue name_residue number, $\angle ADH$ represents the angle formed by acceptor (A), donor (D) and hydrogen (H) ($\angle ADH$).

Table 4.
 The hydrogen bonds formed between the residue side chains between SMN's Tyr109 and Asp105 (PDB entry 4A4E).

What is more, in the computational analysis in [1], 10 side chain hydrogen bonds (**Table 4**) were identified between SMN's Tyr109 and Asp105 in the PDB entry **4A4E** [47], with the donor-acceptor distances ($D - A$ in **Table 4**) at $2.72 \pm 0.06 \text{ \AA}$ and $\angle ADH$ at 14.75 ± 2.93 , no salt bridge was identified for Asp105, and no further hydrogen bonds were identified for Tyr109 and Asp105 for all experimentally determined SMN-related structures as of 25 September 2018.

Taken together, the computational findings here indicate that SMN's Tyr109 and Asp105 contribute to the structural stability of SMN through hydrogen bonding between their side chains, as it is quite clear that if Tyr109 is replaced by Cys109, then the side chain hydrogen bond (**Figure 4, Table 4**) will disappear, and that the negatively charged side chain of Asp105 will gain more geometric freedom due to the disappearance of the hydrogen bond, which can cause a potential disruption of the (either intramolecular and/or intermolecular) electrostatic interaction network, not to mention the possibility of a disrupted disulphide bonding network within the SMN protein, the SMN complex or even the snRNP assembly, which is critical to ensure that pre-mRNA editing of the *SMN1* gene does not go wrong and that its product is the FL-SMN protein, instead of its truncated functionally deficient counterpart.

2.3.4 A structural analysis of the hydrogen bonds formed within the six new SMN-related structures

In light of the six new experimentally determined SMN-related structures (**Table 1**), a new set of hydrogen bonding analysis is conducted according to the details in [1], the result of which is briefly summarised in **Table 5**.

Table 5 shows the four hydrogen bonds formed between snRNP Sm-D2's Asp93 and Gemin2's Arg235 and Arg239. Functionally, Gemin2 is closely linked to SMN (formerly known as Gemin1), and NMR spectroscopy was used to experimentally determine a Gemin1-Gemin2 complex structure (PDB ID: 2LEH) [22, 41], making a closer visual inspection worthwhile of the SMN-related structures (PDB IDs: 5XJS, 5XJT and 5XJU [42], **Table 1**).

From **Figure 5** (PDB ID:5XJS), it is quite clear that the three charged residues (snRNP Sm-D2's Asp93 and Gemin2's Arg235 and Arg239) sit right at the structural interface between Sm-D2 (pink) and Gemin2 (green), with their oppositely

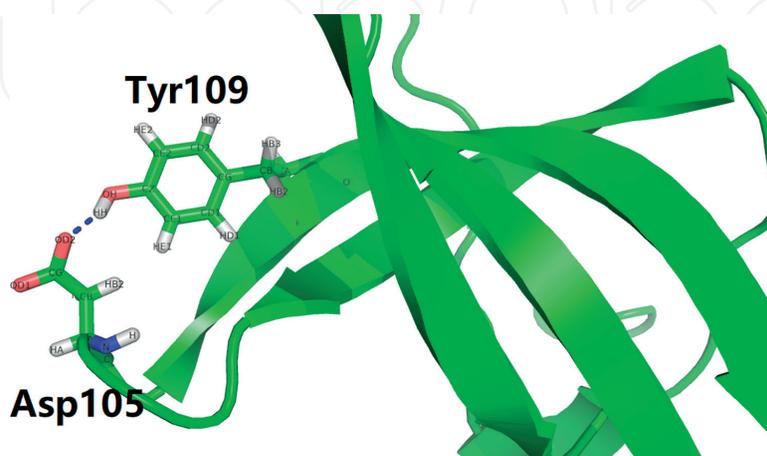


Figure 4. The hydrogen bond (**Table 4**) formed between the side chains of SMN's Tyr109 and Asp105 in the PDB entry **4A4E** [47]. In this figure, SMN's Tyr109 and Asp105 are shown in sticks, all side chain oxygens are coloured red, and side chain nitrogen is coloured blue, while all hydrogen atoms are coloured in white, and all atoms are labelled with their names nearby. The blue dotted line between OD2 of Asp105 and HH of Tyr109 represents the hydrogen bond formed between SMN's Tyr109 and Asp105.

PDB ID	Acceptor (A)	Donor (D)	Hydrogen (H)	D-A (Å)	H-A (Å)	$\angle ADH$ (*)
5XJR	OE1, A_GLN_24	NH2, B_ARG_94	HH21, B_ARG_94	3.00	1.99	1.79
5XJR	OD2, B_ASP_104	NH1, B_ARG_102	HH12, B_ARG_102	2.98	2.07	20.93
5XJS	OD1, B_ASP_93	NE, 2_ARG_235	HE, 2_ARG_235	2.94	1.96	11.86
5XJS	OD1, B_ASP_93	NH2, 2_ARG_239	HH21, 2_ARG_239	2.98	1.98	5.80
5XJS	OD2, B_ASP_60	ND2, B_ASN_64	HD22, B_ASN_64	2.99	2.13	25.53
5XJT	OD1, B_ASP_93	NE, 2_ARG_235	HE, 2_ARG_235	2.65	1.75	21.43
5XJU	OD1, B_ASP_93	NE, 2_ARG_235	HE, 2_ARG_235	2.98	2.08	23.00
5XJU	OD2, B_ASP_60	ND2, B_ASN_64	HD21, B_ASN_64	2.80	1.99	29.63

In this table, the residue naming scheme is Chain ID_residue name_residue number, $\angle ADH$ represents the angle formed by acceptor (A), donor (D) and hydrogen (H) ($\angle ADH$).

Table 5.
 The hydrogen bonds formed between the residue side chains within the six new experimentally determined SMN-related structures.

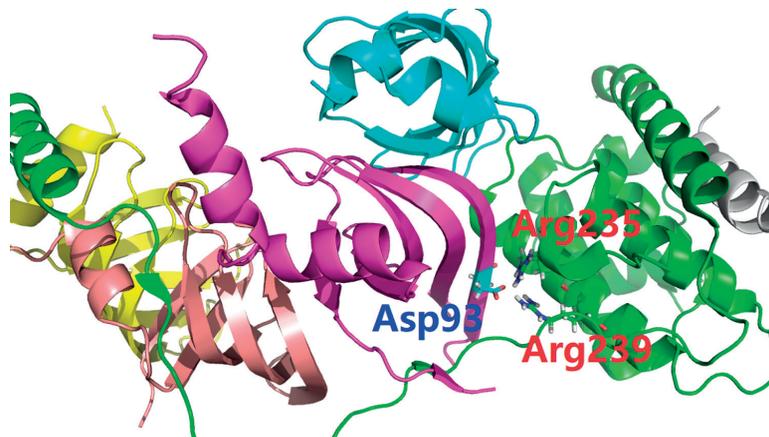


Figure 5.
 Crystal structure of the Gemin2-binding domain of SMN, Gemin2 in complex with SmD1/D2/F/E (PDB ID: 5XJS [42]). In this figure, the whole structure is shown in cartoon and coloured by chain using PyMol [48], where green and pink represent Gemin2 and snRNP Sm-D2, respectively. In this figure, three amino acid residues are shown in sticks and labelled with red and blue texts nearby.

charged side chains closely facing each other, similar to the situation as reported by [1], where the deeply buried side chains of SMN's Lys45 and Asp36 act as two electrostatic clips at the SMN-Gemin2 interface via interactions with both the side chains and the backbone of Gemin2's Gln105, Gln109, His120, His123 and Trp124.

In the subsequent computational salt bridge analysis of the six new SMN-related structures, it turned out that the three charged residues did form salt bridges between their closely facing oppositely charged side chains, as listed in **Table 6** below and illustrated in **Figure 6**.

Collectively, snRNP Sm-D2's Asp93 and Gemin2's Arg235 and Arg239 are three structurally important residues which help stabilise the structural interface through intermolecular electrostatic interactions, including both salt bridges and also hydrogen bonds, similar to the way SMN's Asp44, Gemin2's Arg213 and the two SMN residues (Lys45 and Asp36) play stabilising roles in the SMN-Gemin2 complex structure formation [1].

Considering the intimate functional relationship between Gemin2 and SMN, a further set of structural analysis was conducted for the hydrogen bond and the salt bridge for Arg235 and Arg239 of PDB entry 2LEH [22, 41], and it turned out that

PDB ID	SBnum	Residue A	Atom A	Residue B	Atom B	Distance (Å)
5XJL	3	B ASP_93	OD1	2_ARG_239	NH1	3.734
5XJL	3	B ASP_93	OD1	2_ARG_239	NH2	3.052
5XJL	3	B ASP_93	OD2	2_ARG_239	NH2	3.052
5XJQ	3	B ASP_93	OD1	2_ARG_239	NH1	3.817
5XJQ	3	B ASP_93	OD1	2_ARG_239	NH2	3.059
5XJQ	3	B ASP_93	OD2	2_ARG_239	NH2	3.004
5XJR	3	B ASP_93	OD1	2_ARG_239	NH1	3.811
5XJR	3	B ASP_93	OD1	2_ARG_239	NH2	3.022
5XJR	3	B ASP_93	OD2	2_ARG_239	NH2	2.938
5XJS	3	B ASP_93	OD1	2_ARG_239	NH1	3.688
5XJS	3	B ASP_93	OD1	2_ARG_239	NH2	2.983
5XJS	3	B ASP_93	OD2	2_ARG_239	NH2	3.092
5XJT	2	B ASP_93	OD1	2_ARG_239	NH2	3.251
5XJT	2	B ASP_93	OD2	2_ARG_239	NH2	3.163
5XJU	3	B ASP_93	OD1	2_ARG_239	NH1	3.634
5XJU	3	B ASP_93	OD1	2_ARG_239	NH2	3.084
5XJU	3	B ASP_93	OD2	2_ARG_239	NH2	3.089
5XJL	2	B ASP_93	OD1	2_ARG_235	NH2	3.657
5XJL	2	B ASP_93	OD2	2_ARG_235	NH2	3.475
5XJQ	2	B ASP_93	OD1	2_ARG_235	NH2	3.686
5XJQ	2	B ASP_93	OD2	2_ARG_235	NH2	3.647
5XJR	2	B ASP_93	OD1	2_ARG_235	NH2	3.847
5XJR	2	B ASP_93	OD2	2_ARG_235	NH2	3.800
5XJS	2	B ASP_93	OD1	2_ARG_235	NH2	3.548
5XJS	2	B ASP_93	OD2	2_ARG_235	NH2	3.379
5XJT	2	B ASP_93	OD1	2_ARG_235	NH2	3.258
5XJT	2	B ASP_93	OD2	2_ARG_235	NH2	3.766
5XJU	1	B ASP_93	OD1	2_ARG_235	NH2	3.996

In this table, the residue naming scheme is Chain ID_residue name_residue number, SBnum represents the number of salt bridges computationally identified from the PDB entries listed in this table. Distance represents the distance between two oppositely charged groups/atoms in Å.

Table 6.

A summary of salt bridge analysis of the six new SMN-related structures as of 25 September 2018 [40].

the two arginines did not form any intermolecular electrostatic interaction with SMN, neither salt bridge nor hydrogen bond. Instead, the 2 arginines of Gemin2 only formed 2 hydrogen bonds with Gln272 and His231 of Gemin2, and 1 stable salt bridge with Asp274 of Gemin2, where 16 salt bridges were identified for the 32 NMR structural models (**Table 7**), according to the structural analysis of PDB entry **2LEH** [22, 41].

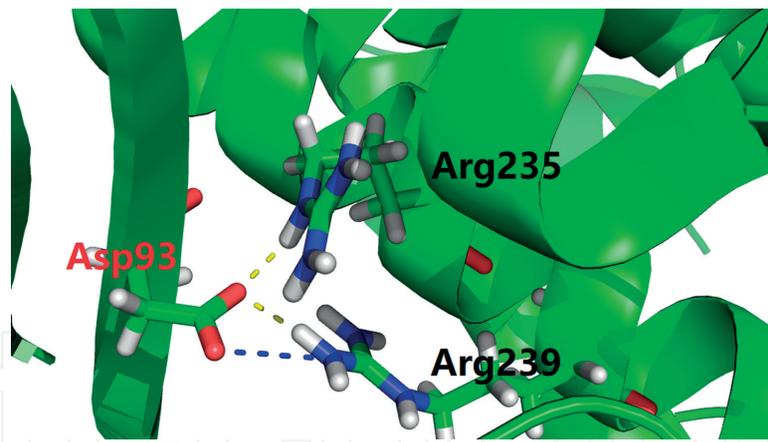


Figure 6. Crystal structure of the Gemin2-binding domain of SMN, Gemin2 in complex with SmD1/D2/F/E (PDB ID: 5XJS) [42]. In this figure, the yellow dotted lines represent two examples of the hydrogen bonds formed between Asp93 and Arg235, while the blue dotted line represents an example of the salt bridge formed between Asp93 and Arg239 (Table 6).

PDB ID	SBnum	Residue A	Atom A	Residue B	Atom B	Distance (Å)
24.pdb	2	A_ASP_274	OD1	A_ARG_239	NH2	3.797
24.pdb	2	A_ASP_274	OD2	A_ARG_239	NH2	2.738
01.pdb	1	A_ASP_274	OD2	A_ARG_235	NH2	3.577
02.pdb	1	A_ASP_274	OD2	A_ARG_235	NH2	3.164
04.pdb	1	A_ASP_274	OD2	A_ARG_235	NH2	3.195
09.pdb	1	A_ASP_274	OD2	A_ARG_235	NH2	2.692
16.pdb	1	A_ASP_274	OD2	A_ARG_235	NH1	3.871
17.pdb	1	A_ASP_274	OD2	A_ARG_235	NH1	2.868
19.pdb	1	A_ASP_274	OD2	A_ARG_235	NH1	3.718
21.pdb	2	A_ASP_274	OD2	A_ARG_235	NH1	3.457
21.pdb	2	A_ASP_274	OD2	A_ARG_235	NH2	3.429
23.pdb	1	A_ASP_274	OD2	A_ARG_235	NH2	3.798
24.pdb	1	A_ASP_274	OD2	A_ARG_235	NH2	2.888
26.pdb	3	A_ASP_274	OD1	A_ARG_235	NH2	3.770
26.pdb	3	A_ASP_274	OD2	A_ARG_235	NH1	3.836
26.pdb	3	A_ASP_274	OD2	A_ARG_235	NH2	2.442
31.pdb	1	A_ASP_274	OD2	A_ARG_235	NH2	3.403

In this table, the residue naming scheme is **Chain ID_residue name_residue number**, SBnum represents the number of salt bridges computationally identified.

Table 7. A summary of salt bridge analysis of PDB entry 2LEH [22, 41].

3. Concluding remarks

Given SMN's critical role in the maturation of snRNP and in the development of SMA [2, 6, 11], it is necessary for the structure-activity relationship (SAR) characterisation to continue for the SMA protein. With various biophysical tools available for structural determination, for SMN-related proteins and biological complexes,

such as the SMN complex and snRNPs, their structure determination and functional characterisation will undoubtedly continue to advance, which will be helpful both in further understanding of SMN's role in SMA from a molecular structural point of view. In practice, however, advancements do not come easy. For instance, although both full-length structures of FL-SMN (with 294 residues) and SMN Δ 7 (with 282 residues) were already experimentally determined using X-ray crystallography and deposited in the database (PDB IDs: 4NL6 and 4NL7), they were subsequently withdrawn by the author because the sample used for the structure determination was wrong. Otherwise, these two full-length SMN structures would constitute the very first step towards a comprehensive picture of the structural and functional insights into SMN's role in the molecular pathogenesis SMA.

As of 25 September 2018, there is still no full-length SMN (or the SMN complex or the snRNP assembly) structure deposited in the wwPDB website [40], although it contains six new experimentally determined SMN-related structures, in addition to those reported in [1]. In terms of amino acid sequence, those SMN-related structures are still only SMN fragments, ranging from Gly26 to Lys51, and from Asn84 to Glu147. In between, there is still structurally not-determined-yet regions (referred to as **structural gaps** below) consisting of 204 SMN residues. Sixteen months have passed since the publication of [1], the structural gaps still remain, literally zero progress has been made to bridge them in spite of the six newly deposited structures, calling **again** [1] for further comprehensive structural determination and functional research for this SMA protein.

4. A residue-specific distributional analysis of the structural gaps in the Protein Data Bank

As a 38-kD protein, SMN is essentially a small one in terms of molecular weight, in comparison with all proteins whose structures have been deposited in the Protein Data Bank (PDB), a primary database for experimentally determined structures of biological molecules [40]. As discussed above, even for a protein as small as SMN, experimental structure determination does not seem simple or easy, especially when it has to be done in a full-length and gapless manner. Therefore, to test whether any residue-specific statistical pattern (not known yet before this chapter) exists in the structural gaps in the whole Protein Data Bank (accessed 25 September 2018), this chapter presents a set of residue-specific distributional analysis of all structural gaps throughout PDB.

While the number of experimentally determined protein structures keeps increasing in the PDB, with the number of cryo-EM structures [49] on the rise, X-ray crystallography and NMR spectroscopy remain to date the two main (**Table 8**) supplementary biophysical tools in structural biology, both with strengths and weaknesses [50, 51].

In PDB-format data, the atomic coordinates presented in ATOM records in a PDB file may not exactly match the sequence in the SEQRES records. However, these amino acids will often be included in the SEQRES records, since the portion of the chain was present during the experiment. In these cases, a 'REMARK 465' entry will be included in the header of the PDB file to identify each missing residue. For X-ray crystallography data, the ends of chains and mobile loops are often not observed in crystallographic experiments, and as a result, atomic coordinates are not included as ATOM records in the file, leading to the occurrence of gaps for structure determined by X-ray crystallography. Among currently available biophysical tools, NMR spectroscopy is able to provide unique access to atomic-level structural dynamic behaviour of protein molecules in solution under physiological conditions (such as temperature, pH, etc.). As a result, this chapter focuses on

Experimental method	Proteins	Nucleic acids	Protein/NA complex	Other	Total
X-ray	121,081	1958	6257	10	129,306
NMR	10,848	1256	250	8	12,362
Electron microscopy	1750	31	623	0	2404
Other	244	4	6	13	267
Multi-method	117	5	2	1	125

Table 8.
 A summary of the number of experimentally determined biomolecular structures in PDB as of 25 September 2018.

the structural gaps within protein structures determined by NMR spectroscopy, and aims to test whether any residue-specific statistical pattern exists in them. Here, structural gaps are defined as protein fragments with residues which exist in the originally studied molecule as shown in the SEQRES records, but not in the observed structure/atomic coordinates.

As of 20 September 2018, 10,844 NMR-determined protein structures have been deposited in the Protein Data Bank, according to a structure search with two parameters (molecule type = protein, experimental method = NMR). After the 10,844 PDB files were downloaded from the PDB website, the numbers of the total and the missing amino acid residues were extracted with an in-house python script for all proteins, as listed in **Table 9**.

Residue	Missing no.	Total no.	Ratio = Missing no./Total no.
A	1782	75,627	0.023
C	152	25,777	0.00589673
E	1811	79,729	0.022
D	1390	59,908	0.023
G	3321	81,347	0.041
F	615	38,993	0.015
I	640	55,070	0.011
H	5146	26,182	0.196
K	1442	75,766	0.019
M	1203	23,652	0.050
L	1439	90,833	0.015
N	831	43,080	0.019
Q	1273	44,594	0.028
P	1518	46,205	0.032
S	3159	73,904	0.042
R	1234	52,761	0.023
T	1112	56,787	0.019
W	138	13,211	0.010
V	980	70,252	0.013
Y	626	32,905	0.019
Sum	29,812	1,066,583	0.027

Table 9.
 The numbers of the total and the missing amino acid residues in NMR-determined protein structures as of 25 September 2018.

In total, the 10,844 protein structures contains 1,066,583 amino acid residues, ~2.8% of which (29812) are missing, i.e. the atomic positions of the 29,841 residues were not experimentally determined by NMR spectroscopy, although they were present in the NMR sample during the structural determination process.

From **Figure 7**, it can be seen that for 19 residues (excluding histidine), the missing ratio is well below or pretty close to 5%, while the missing ratio is 19.6% for histidine, as shown by the blue sharp peak on **Figure 7**. In a statistical one sample t-test analysis of the 19 missing ratios, it turned out 100% acceptable ($P=1$) that the average of ratio is 0.0231, and that the fitness between the 19 missing ratios and the red horizontal line (**Figure 8**) is 100% acceptable ($P=1$), according to a statistical Chi-square test, as revealed by **Figure 8**.

While a missing ratio of 5% might be considered statistically insignificant, a missing ratio of 19.6% is clearly not to be ignored here, raising one obvious question: what on earth is so special about histidine that makes it so special among the 20 naturally occurring amino acids in this residue-specific distributional analysis of the structural gaps?

Similar to the other 19, histidine is a naturally occurring amino acid that is used in the biosynthesis of proteins. Also similar to the other 19, it contains an amino group (which is in the protonated —NH_3^+ form under biological conditions) and a carboxylic acid group (which is in the deprotonated —COO^- form under biological conditions). In particular, histidine has an imidazole side chain (which is partially protonated), classifying it as a positively charged amino acid at physiological pH (~7.4). That is, among the 20 naturally occurring amino acids, five (Arg, Lys, His, Glu and Asp) possess ionisable side chains. Among the five, histidine is the only one whose side chain has an ionisable (with an intrinsic pKa at 6.04) [52, 53] imidazole ring structure, which can exist in two inter-convertible tautomeric states. While at a pH of 7.0, the imidazole ring is mostly deprotonated (proton occupancy = 9.88%), at a pH of 6.0, the imidazole ring is largely protonated (proton

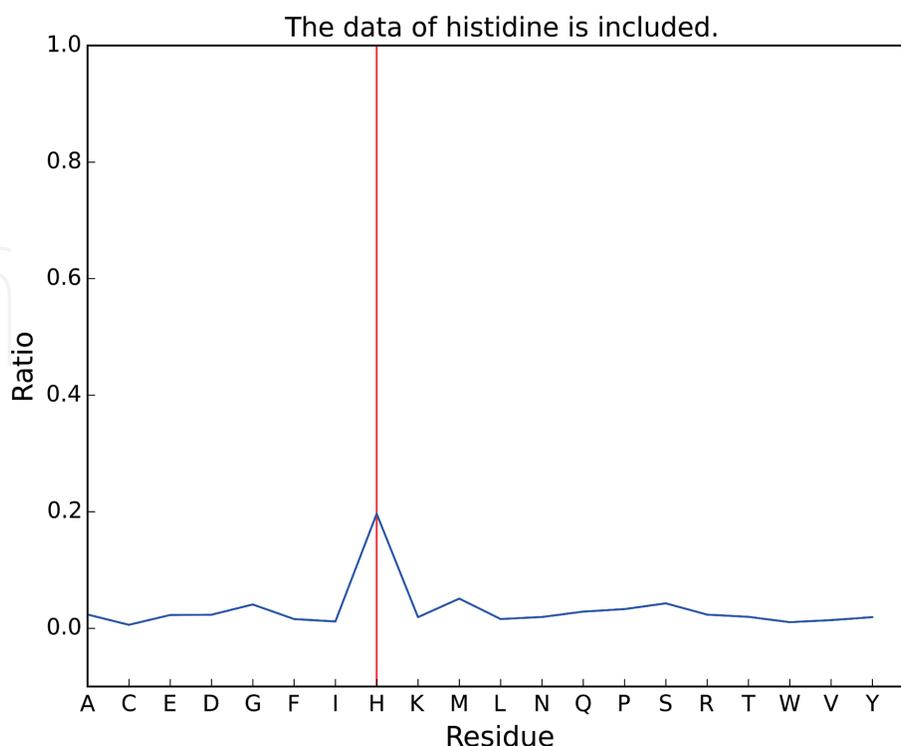


Figure 7.

A residue-specific distribution of the missing residues in NMR-determined protein structures as of 25 September 2018. In this figure, x-axis represents the one-letter codes for amino acid residues, and y-axis represents the residue-specific ratio of missing versus total residues in those NMR structures. The red vertical line highlights histidine as a particular residue with an outstanding missing ratio.

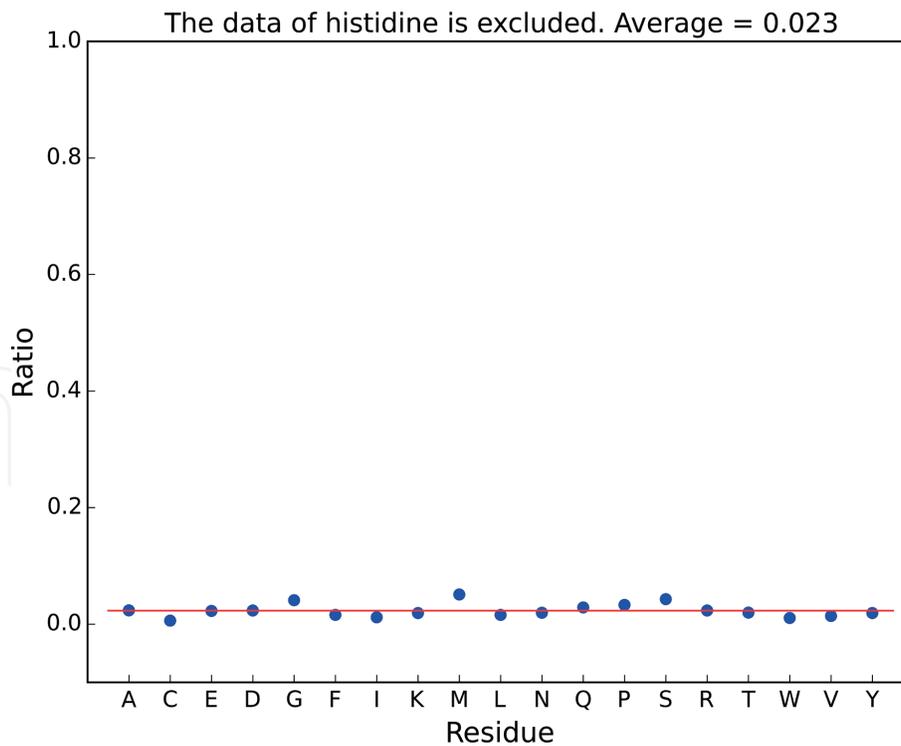


Figure 8.

A residue-specific scatter plot of the missing residues in NMR-determined protein structures as of 25 September 2018. In this figure, x-axis represents the one-letter codes for amino acid residues, and y-axis represents the residue-specific ratio of missing versus total residues in those NMR structures. The red horizontal line represents the average missing ratio level of the 19 residues.

occupancy = 52.30%), as defined by the classical Henderson-Hasselbalch equation [50], where the positively charged imidazole ring bears two NH bonds and has a positive electric charge, which is equally distributed between both nitrogens. As the pH increases, the imidazole ring loses the positive charge, and the remaining proton of the neutral imidazole ring can reside on either nitrogen, giving rise to two tautomeric states of the histidine side chain [52, 54, 55].

To sum up, it is probable that the missing ratio of histidine is much higher than the other 19 because it has a special side chain with special dynamic structural and physicochemical properties (such as stacking interaction [56]), and with a special imidazole ring in constant protonation-deprotonation equilibrium [57] and two tautomeric states [52, 54, 55], making its NMR-observables (chemical shift for instance) difficult to be experimentally observed and measured by NMR spectroscopy and structurally calculated by NMR-related software in the structural determination of proteins. To address this issue of PDB-wide structural gaps, selective isotope labelling of histidine residues (the side chains in particular) can be a useful approach in biomolecular structural determination by NMR spectroscopy, not just alone, but also in collaboration with other biophysical tools, not just for the special histidine, but also for its 19 siblings in the fundamental building block of life.

IntechOpen

IntechOpen

Author details

Wei Li

Institute of Nautical Medicine, Nantong University, Nantong City, Jiangsu Province, China P. R.

*Address all correspondence to: liweiqidong@stu.edu.cn

IntechOpen

© 2018 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Li W. How do SMA-linked mutations of SMN1 lead to structural/functional deficiency of the SMA protein? *PLoS One*. 2017;**12**(6):e0178519
- [2] Burghes AHM. When is a deletion not a deletion? When it is converted. *American Journal of Human Genetics*. 1997;**61**(1):9-15
- [3] Lefebvre S, Bürglen L, Reboullet S, Clermont O, Burlet P, Viollet L, et al. Identification and characterization of a spinal muscular atrophy-determining gene. *Cell*. 1995;**80**(1):155-165
- [4] Lorson CL, Strasswimmer J, Yao JM, Baleja JD, Hahnen E, Wirth B, et al. SMN oligomerization defect correlates with spinal muscular atrophy severity. *Nature Genetics*. 1998;**19**(1):63-66
- [5] Hua Y, Vickers TA, Okunola HL, Bennett CF, Krainer AR. Antisense masking of an hnRNP A1/A2 intronic splicing silencer corrects SMN2 splicing in transgenic mice. *American Journal of Human Genetics*. 2008;**82**(4):834-848
- [6] Burghes AH, Beattie CE. Spinal muscular atrophy: Why do low levels of survival motor neuron protein make motor neurons sick? *Nature Reviews Neuroscience*. 2009;**10**(8):597-609
- [7] Hua Y, Sahashi K, Hung G, Rigo F, Passini MA, Bennett CF, et al. Antisense correction of SMN2 splicing in the CNS rescues necrosis in a type III SMA mouse model. *Genes & Development*. 2010;**24**(15):1634-6144
- [8] Hua Y, Sahashi K, Rigo F, Hung G, Horev G, Bennett CF, et al. Peripheral SMN restoration is essential for long-term rescue of a severe spinal muscular atrophy mouse model. *Nature*. 2011;**478**(7367):123-216
- [9] Liu Q, Dreyfuss G. A novel nuclear structure containing the survival of motor neurons protein. *EMBO Journal*. 1996;**15**(14):3555-3565
- [10] Carvalho T, Fcalapez AA, Lafarga M, Berciano M, Carmo FM. The spinal muscular atrophy disease gene product, SMN: A link between snRNP biogenesis and the Cajal (coiled) body. *Journal of Cell Biology*. 1999;**147**(4):715-727
- [11] Young PJ, Le TT, Dunckley M, Nguyen TM, Burghes AH, Morris GE. Nuclear gems and Cajal (coiled) bodies in fetal tissues: Nucleolar distribution of the spinal muscular atrophy protein, SMN. *Experimental Cell Research*. 2001;**265**(2):252-261
- [12] Young P, Ntlorson MC, Le T, Androphy E, Burghes A, Morris G. The exon 2b region of the spinal muscular atrophy protein, SMN, is involved in self-association and SIP1 binding. *Human Molecular Genetics*. 2000;**9**(19):2869-8277
- [13] Seng CO, Magee C, Young PJ, Lorson CL, Allen JP. The SMN structure reveals its crucial role in snRNP assembly. *Human Molecular Genetics*. 2015;**24**(8):2138-2146
- [14] Grimm C, Chari A, Pelz JP, Kuper J, Kisker C, Diederichs K, et al. Structural basis of assembly chaperone-mediated snRNP formation. *Molecular Cell*. 2013;**49**(4):692-703
- [15] Pellizzoni L, Charroux B, Dreyfuss G. SMN mutants of spinal muscular atrophy patients are defective in binding to snRNP proteins. *Proceedings of the National Academy of Sciences*. 1999;**96**(20):11167-11172
- [16] Pellizzoni L. Essential role for the SMN complex in the specificity of snRNP assembly. *Science*. 2002;**298**(5599):1775-1779
- [17] Workman E, Saieva L, Carrel TL, Crawford TO, Liu D, Lutz C, et al. A

SMN missense mutation complements SMN2 restoring snRNPs and rescuing SMA mice. *Human Molecular Genetics*. 2009;**18**(12):2215-2229

[18] Otter S, Grimm M, Neuenkirchen N, Chari A, Sickmann A, Fischer U. A comprehensive interaction map of the human survival of motor neuron (SMN) complex. *Journal of Biological Chemistry*. 2007;**282**(8):5825-8533

[19] Chari A, Golas MM, Klingenhäger M, Neuenkirchen N, Sander B, Englbrecht C, et al. An assembly chaperone collaborates with the SMN complex to generate spliceosomal snRNPs. *Cell*. 2008;**135**(3):497-509

[20] Gubitz AK, Feng W, Dreyfuss G. The SMN complex. *Experimental Cell Research*. 2004;**296**(1):51-56

[21] Liu Q, Fischer U, Wang F, Dreyfuss G. The spinal muscular atrophy disease gene product, SMN, and its associated protein {SIP1} are in a complex with spliceosomal snRNP proteins. *Cell*. 1997;**90**(6):1013-1021

[22] Sarachan KL, Valentine KG, Gupta K, Moorman VR, Gledhill JM, Bernens M, et al. Solution structure of the core SMN-Gemin2 complex. *Biochemical Journal*. 2012;**445**(3):361-370

[23] Monani UR, Pastore MT, Gavrilina TO, Jablonka S, Le TT, Andreassi C, et al. A transgene carrying an A2G missense mutation in the SMN gene modulates phenotypic severity in mice with severe (type I) spinal muscular atrophy. *Journal of Cell Biology*. 2003;**160**(1):41-52

[24] Wirth B. An update of the mutation spectrum of the survival motor neuron gene *SMN1* in autosomal recessive spinal muscular atrophy (SMA). *Human Mutation*. 2000;**15**(3):228-237

[25] Sun Y, Grimm M, Schwarzer V, Schoenen F, Fischer U, Wirth

B. Molecular and functional analysis of intragenic SMN1 mutations in patients with spinal muscular atrophy. *Human Mutation*. 2005;**25**(1):64-71

[26] Zhang ML, Lorson CL, Androphy EJ, Zhou J. An in vivo reporter system for measuring increased inclusion of exon 7 in SMN2 mRNA: Potential therapy of SMA. *Gene Therapy*. 2001;**8**(20):1532-1538

[27] Ogawa C, Usui K, Aoki M, Ito F, Itoh M, Kai C, et al. Gemin2 plays an important role in stabilizing the survival of motor neuron complex. *Journal of Biological Chemistry*. 2007;**282**(15):11122-11134

[28] Clermont O, Burlet P, Benit P, Chanterau D, Saugier Veber P, Munnich A, et al. Molecular analysis of SMA patients without homozygous SMN1 deletions using a new strategy for identification of SMN1 subtle mutations. *Human Mutation*. 2004;**24**(5):417-427

[29] Prior TW. Spinal muscular atrophy diagnostics. *Journal of Child Neurology*. 2007;**22**(8):952-956

[30] Brichta L, Garbes L, Jedrzejowska M. Nonsense-mediated messenger RNA decay of survival motor neuron 1 causes spinal muscular atrophy. *Human Genetics*. 2008;**123**(2):141-153

[31] Zapletalová E, Hedvičáková P, Kozák L, Vondráček P, Gaillyová R, Maří'ková T, et al. Analysis of point mutations in the SMN1 gene in SMA patients bearing a single SMN1 copy. *Neuromuscular Disorders*. 2007;**17**(6):476-481

[32] Alías L, Bernal S, Fuentes-Prior P, Barceló MJ, Also E, Martínez-Hernández R, et al. Mutation update of spinal muscular atrophy in Spain: Molecular characterization of 745 unrelated patients and identification of four novel mutations in the SMN1 gene. *Human Genetics*. 2008;**125**(1):29-39

- [33] Tsai C, Jong Y, Hu C, Chen C, Shih M, Chang C, et al. Molecular analysis of SMN, NAIP and P44 genes of SMA patients and their families. *Journal of the Neurological Sciences*. 2001;**190**(2):35-40
- [34] Rochette CF, Surh LC, Ray PN, McAndrew PE, Prior TW, Burghes AHM, et al. Molecular diagnosis of non-deletion SMA patients using quantitative PCR of SMN exon 7. *Neurogenetics*. 1997;**1**(2):141-147
- [35] Shpargel KB, Gall JG. Gemin proteins are required for efficient assembly of Sm-class ribonucleoproteins. *Proceedings of the National Academy of Sciences of the United States of America*. 2005;**102**(48):17372-73717
- [36] Hahnen E, Schönling J, Rudnik-Schöneborn S, Raschke H, Zerres K, Wirth B. Missense mutations in exon 6 of the survival motor neuron gene in patients with spinal muscular atrophy (SMA). *Human Molecular Genetics*. 1997;**6**(5):821-285
- [37] Talbot K, Ponting CP, Theodosiou AM, Rodrigues NR, Surtees R, Mountford R, et al. Missense mutation clustering in the survival motor neuron gene: A role for a conserved tyrosine and glycine rich region of the protein in RNA metabolism? *Human Molecular Genetics*. 1997;**6**(3):497-500
- [38] Wang CH, Papendick BD, Bruinsma P, Day JK. Identification of a novel missense mutation of the SMN^T gene in two siblings with spinal muscular atrophy. *Neurogenetics*. 1998;**1**(4):273-276
- [39] Sneha P, Zenith TU, Habib USA, Evangeline J, Kumar DT, Doss CGP, et al. Impact of missense mutations in survival motor neuron protein (SMN1) leading to spinal muscular atrophy (SMA): A computational approach. *Metabolic Brain Disease*. 2018:1-12.
- <https://link.springer.com/article/10.1007%2Fs11011-018-0285-4>
- [40] Berman H, Henrick K, Nakamura H. Announcing the worldwide protein data bank. *Nature Structural Biology*. 2003;**10**(12):980
- [41] Zhang R, So BR, Li P, Yong J, Glisovic T, Wan L, et al. Structure of a key intermediate of the SMN complex reveals Gemin2's crucial function in snRNP assembly. *Cell*. 2011;**146**(3):384-395
- [42] Yi ZRH. Structures of 7S mutant complexes. Available from: <http://www.rcsb.org/structure/5XJQ> [to be published]
- [43] Kabsch W, Sander C. Dictionary of protein secondary structure: Pattern recognition of hydrogenbonded and geometrical features. *Biopolymers*. 1983;**22**(12):2577-2637
- [44] Hubbard SJ, Thornton JM. Naccess. Computer Program. Vol. 2. London: Department of Biochemistry and Molecular Biology, University College London; 1993
- [45] Baker M. Cryo-electron microscopy shapes up. *Nature*. 2018;**561**(7724):565-567
- [46] Sattler M, Selenko P, Sprangers R, Stier G, Bühler D, Fischer U. SMN tudor domain structure and its interaction with the Sm proteins. *Nature Structural Biology*. 2001;**8**(1):27-31
- [47] Tripsianes K, Madl T, Machyna M, Fessas D, Englbrecht C, Fischer U, et al. Structural basis for dimethylarginine recognition by the tudor domains of human SMN and SPF30 proteins. *Nature Structural & Molecular Biology*. 2011;**18**(12):1414-1420
- [48] DeLano WL. Pymol: An open-source molecular graphics tool. *CCP4 Newsletter On Protein Crystallography*. 2002;**40**:82-92

- [49] Wu J, Yan Z, Li Z, Qian X, Lu S, Dong M, et al. Structure of the voltage-gated calcium channel $\text{Ca}_v1.1$ at 3.6 Å resolution. *Nature*. 2016;**537**(7619):191-196
- [50] Li W. Gravity-driven pH adjustment for site-specific protein pKa measurement by solution-state NMR. *Measurement Science and Technology*. 2017;**28**(12):127002
- [51] Li W. NMR-observed atomic bond length stability supports a dimensionality shift in protein main chain 3D structure description and representation. *Current Research Bioorganic and Organic Chemistry*. 2018;**2018**(2):1-4
- [52] Pelton JG, Torchia DA, Meadow ND, Roseman S. Tautomeric states of the active-site histidines of phosphorylated and unphosphorylated IIIgIc, a signal-transducing protein from *Escherichia coli*, using two-dimensional heteronuclear NMR techniques. *Protein Science*. 2008;**2**(4):543-558
- [53] Hansen AL, Kay LE. Measurement of histidine pKa values and tautomer populations in invisible protein states. *Proceedings of the National Academy of Sciences*. 2014;**111**(17):E1705-E1712
- [54] Li S, Hong M. Protonation, tautomerization, and rotameric structure of histidine: A comprehensive study by magic-angle-spinning solid-state NMR. *Journal of the American Chemical Society*. 2011;**133**(5):1534-1544
- [55] Wolff N. Histidine pKa shifts and changes of tautomeric states induced by the binding of gallium-protoporphyrin IX in the hemophore HasASM. *Protein Science*. 2002;**11**(4):757-765
- [56] Wang L, Sun N, Terzyan S, Zhang X, Benson DR. A histidine/tryptophan π -stacking interaction stabilizes the heme-independent folding core of microsomal apocytochrome b5 relative to that of mitochondrial apocytochrome b5. *Biochemistry*. 2006;**45**(46):13750-13759
- [57] Hass MAS, Yilmaz A, Christensen HEM, Led JJ. Histidine side-chain dynamics and protonation monitored by ^{13}C CPMG NMR relaxation dispersion. *Journal of Biomolecular NMR*. 2009;**44**(4):225-233