Journal of Molecular Biology

Disorder in human Skp1 structure is the key to its adaptability to bind many different proteins in the SCF complex assembly

--Manuscript Draft--

Manuscript Number:	JMB-D-22-00265R2
Article Type:	Full Length Article
Section/Category:	Structure, chemistry, processing and function of biologically important macromolecules and complexes
Keywords:	Skp1; protein structure; conformation; Dynamics, Skp1, SCF complex,
Corresponding Author:	Ashutosh Kumar, Ph.D. Indian Institute of Technology Bombay Mumbai, INDIA
First Author:	Ashutosh Kumar, Ph.D.
Order of Authors:	Ashutosh Kumar, Ph.D.
	Amrita Bhattacharya
	Vaibhav Shukla, PhD
	Nitin Nathubhai Kachariya
	Preeti P
	Praveen Sehrawat
Abstract:	Skp1(S-phase kinase-associated protein 1 - Homo sapiens) is an adapter protein of the SCF complex, which links the constant components (Cul-1-RBX) and the variable receptor (F-box proteins) in Ubiquitin E3 ligase. It is intriguing how Skp1 can recognise and bind to a variety of structurally different F-Box proteins. For practical reasons, previous efforts have used truncated Skp1, and thus it has not been possible to track the crucial aspects of the substrate recognition process. In this background, we report the solution structure of the full-length Skp1 protein determined by NMR spectroscopy for the first time and investigate the sequence-dependent dynamics in the protein. The solution structure reveals that Skp1 has an architecture: β1-β2-H1-H2-L1–H3-L2-H4-H5-H6-H7(partially formed) and a long tail-like disordered C-terminus. Structural analysis using DALI reveals conserved domain structure across species for Skp1. Backbone dynamics investigated using NMR relaxation suggest substantial variation in the motional timescales along the length of the protein. The loops and the C-terminal residues are highly flexible, and the (R 2 /R 1) data suggests ms-μs timescale motions in the helices as well. Further, curved temperature dependence reveals that for several residues across the length of the protein, there exist a native-like low-lying excited states, particularly for several C-terminal residues. Our results provide a rationale for how the protein can adapt itself, bind, and get functionally associated with other proteins in SCF complex utilising its flexibility and conformational sub-states.

To,
The Editor,
Journal of Molecular Biology,

Dear Prof. Wright

We, along with our co-authors, present our response to the reviewers along with the modified manuscript titled **"Disorder in human Skp1 structure is the key to its adaptability to bind many different proteins in the SCF complex assembly"** for publication as an original article in *Journal of Molecular Biology*.

We sincerely thank the editor and the reviewers for giving us the opportunity to improve the quality of our manuscript. All the changes suggested by our reviewers have been duly incorporated, and all the questions have been addressed. We added a new analysis and explanation in response to reviewer no 1 queries and used a variety of biophysical techniques to demonstrate that Human Skp1 is a monomeric protein in response to the question from reviewer no 3. We hope that the manuscript is in significantly better shape and hope that all the queries are resolved now. In our response letter, the reviewer's comments are in black while our responses are in blue.

Now that the concerns of the reviewers are adequately addressed, I hope the present form of our manuscript is suitable for publication. Looking forward to having a favourable response from you.

Best wishes,
Prof. Ashutosh Kumar

Reviewers' comments:

Reviewer #1: The authors have provided clear responses to the majority of points raised previously. However, I remain concerned by the validity of interpreting R2/R1 ratios solely in terms of chemical exchange. I have included a further comment below that I hope more clearly expresses my doubts. If the authors can address this point, then I am happy to recommend the manuscript for publication in JMB.

The authors state that R2/R1 ratios represent exchange contributions (p. 11 l. 23-25), but this is not correct: as originally derived (Kay, Torchia and Bax, Biochemistry (1989), 28, 8972-8979) the ratio is highly sensitive to the rotational correlation time. In the absence of further evidence (e.g., multi-field relaxation measurements and spectral density mapping or model free analysis, or relaxation dispersion measurements) claims to identify chemical exchange in helical regions on the us-ms timescale cannot be justified and should be removed (p. 11 l. 20-32).

Response: We thank the reviewer for suggesting a reference to compare our data. However, in the discussion paragraph, Kay et al[1] clearly state: "In the interpretation of the relaxation data presented herein, we have assumed that the relaxation properties of the ^{15}N spins are governed solely by the ^{1}H-^{15}N dipolar interaction and the CSA interaction. These two interactions are, however, not the only ones that can affect ^{15}N relaxation rates. For example, chemical exchange processes can influence T_2 rates considerably.[2,3]"

"These processes include an exchange between several conformers occurring in specific regions of the protein." Thus, it is suggested the susceptibility of T_2 to chemical exchange processes, which can affect relaxation. Therefore, we have extrapolated our results on this basis. However, as the reviewer pointed out, and we agree, this is not a conclusive measure of exchange processes and is merely indicative that such contributions may exist.

As suggested by the reviewer, we have made appropriate changes: "There is a substantial variation in the ms-μs time scale along the length of the chain, and the helices may exhibit conformational exchange as high R_2 is observed in the helices compared to the surrounding residues.[1,2,3] Higher R_2/R_1 ratios may be indicative of the presence of chemical exchange contributions to R_2 along the protein sequence. A slow timescale chemical exchange process is more likely to exist in the helical region, and it would display relatively high values at the helical regions and low values in the loop regions and at the termini. Conclusively, it emerges from the dynamics data that the protein is highly geared up to adapt structurally, depending upon the needs of the binding partner."

I completely agree that R2 rates are sensitive to chemical exchange, as noted by Kay et al. in the 1989 reference. This contrasts with R1 rates and the heteronuclear NOE, which do not depend on J(0) terms and thus are only sensitive to sub-tc dynamics and tumbling. To a crude approximation, in the macromolecular regime the dependence of R1 and R2 rates on tumbling and chemical exchange can be expressed:
R2 ~ tau_c + Rex
R1 ~ 1/tau_c
The ratio R2/R1 therefore varies approximately as (tau_c^2 + tau_c*Rex). Thus, while there is sensitivity to chemical exchange, there is also very strong dependence on mobility. Of course, a full analysis including internal motion (order parameters) complicates this picture, but the interpretation is also borne out by the authors' own measurements: secondary structure elements have low R1s, high R2s, and high NOEs, while loops have high R1s, low R2s and low NOEs. This trend is acknowledged in the manuscript, but its implications are not, and the effect is merely exaggerated by plotting the ratio R2/R1. Indeed, the Pielak

group have argued for the analysis of the product R1R2 as an alternative to detect chemical exchange in the presence of anisotropic motion (Kneller et al, JACS 124 1852-1853 (2002)).

Ultimately, I may be incorrect in my interpretation, but in the absence of a rigorous analysis of relaxation data, e.g. via model-free analysis, I do not see that the authors can prove this and unambiguously attribute variations in relaxation rates to conformational (chemical) exchange processes rather than local dynamics or anisotropic mobility.

Response: We thank the reviewer for the insightful scientific discussion and concur that the Rex component of R_2 cannot be quantified in the absence of more concrete experimental measurements.

We completely agree with the reviewer that Kneller et.al[4] have quoted that "A commonly used approach for the analysis of heteronuclear relaxation data examines the ratio between transverse and longitudinal rates, R_2/R_1, for identification of residues undergoing chemical exchange and estimation of global correlation times. The R_2/R_1 approach does not distinguish between the effects of motional anisotropy and chemical exchange."

From the R_2/R_1 measurement alone, we cannot rule out that the higher values are due to local dynamics. We have modified the section appropriately and hope that there is no more ambiguity in our discussion of the R_2/R_1 data interpretation. Additionally, we performed a reduced spectral density analysis as appended below.

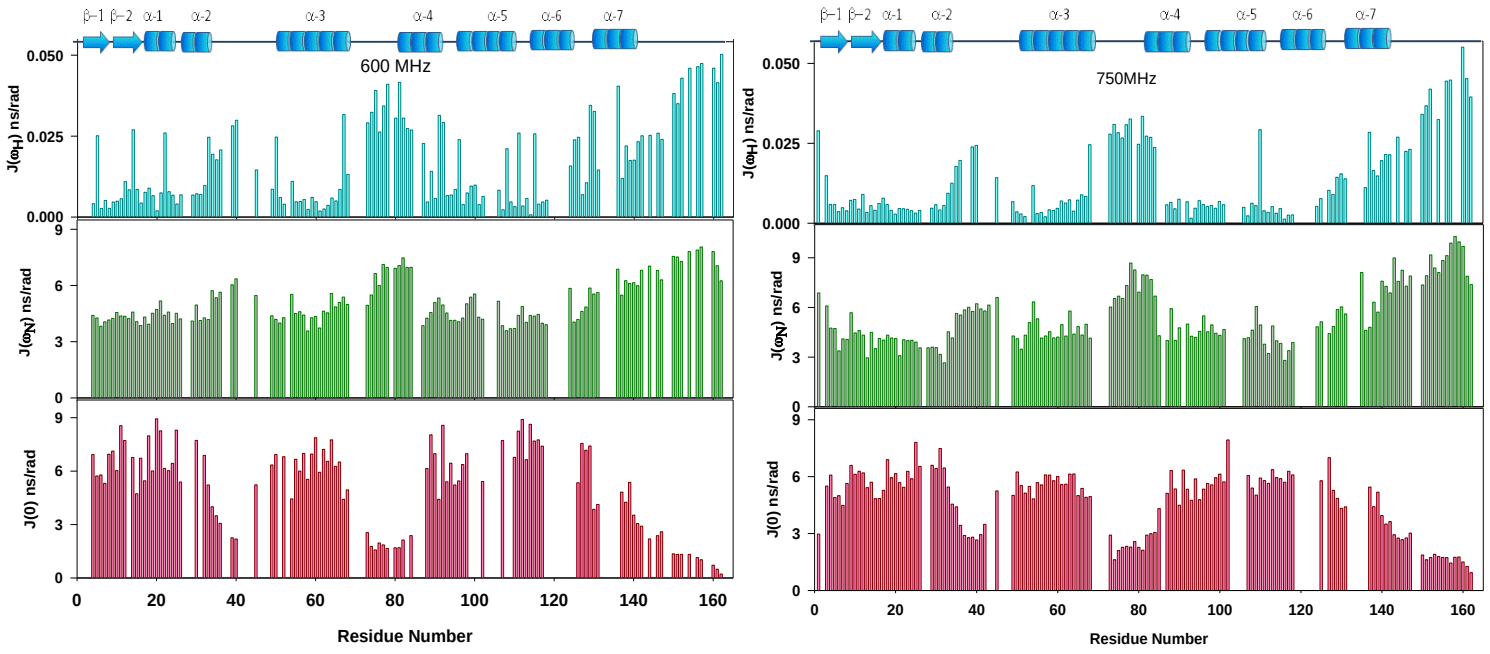


Fig S11**:** Spectral density analysis of Skp1: The figure shows the residue versus $J(\omega_H)$, J(0) and $J(\omega_N)$ values of Skp1 from the backbone dynamics data at two different magnetic fields 750 MHz and 600MHz.

The values of spectral density functions $J(\omega_H)$ are contributed mostly by the $^{1}H–^{15}N$ NOE. On the other hand, R_1 and R_2 largely contribute to the values of $J(\omega_N)$ and J(0) functions, respectively. Higher J(0) values may correspond to the slow rotational motion of the N-H bond vector in the millisecond or even nanosecond time scale. The $J(\omega_N)$ and J(0) follow an inverse correlation in the spectral density values in both fields. Also, the spectral density values mirror the secondary structure of the protein with distinctly patterned spectral density values in the ordered and disordered regions of the protein. This secondary structure pattern agrees with the NMR structure of Skp1 protein, which has a flexible C-terminus and two distinct loop regions.

We have further added the following section to discuss the relaxation data to our main text and modified Fig 4:

"From such data, different strategies have been used to get insights into slow conformational exchange contributions along the sequence of a given protein molecule. These include Model Free analysis, Spectral Density analysis, R_2*R_1 vs. sequence plots, and R_2/R_1 vs. sequence plots. Among these, Model Free analysis is typically used for well-folded proteins and is not suitable for intrinsically disordered regions in the proteins, as in the present case. In an approximate sense, R_2 is linearly dependent on correlation time (τ_c) and R_{ex}, R_1 is inversely dependent on τ_c, and therefore, the product R_2*R_1 should be roughly independent of τ_c in the absence of Rex. Thus, any variation in this product is indicative of sequence-wise variation in Rex contributions. Among the spectral densities, $J(\omega_H)$ is contributed mostly by the $^{1}H–^{15}N$ NOE and reflects on high-frequency motions; $J(\omega_N)$ and J(0) functions have contributions from R_2 and R_1, and among these, only J(0) has contributions from slow motions, which includes chemical exchange. We have used all three approaches, namely, R_2/R_1 vs. sequence, R_2*R_1 vs. sequence, and spectral densities vs. sequence to characterise motions in the Skp1 protein Fig4 (B)."

"Our data shows that the R_2*R_1 product displays substantial sequence-wise variations indicating significant contributions from Rex. The same is true in the R_2/R_1 vs sequence plots, and J(0) vs sequence plot (J(0) is more sensitive to slow motions. Interestingly, the helices seem to exhibit conformational exchange from all three data sets; however, from such data, it is not possible to quantitatively estimate the contributions. We further report that the τ_c of Skp1 in Fig 4 (B) is more or less in agreement with the correlation time of a ~20kDa protein. The overall τ_c of Skp1 is around 10.8 ns. The τ_c value is close to those of other standard globular proteins but shows some deviation as shown in Fig S12, which results from the different R_2/R_1 values of loops and disordered c-term."

Reviewer #2: The authors have addressed all of my concerns. I support this manuscript being accepted for publication in JMB.

Response: We are thankful to the reviewer for the recommendation.

Reviewer #3: The revised manuscript is improved in several ways. The BMRB deposition updating has been done. Data on R1, R2, hNOE, and temperature dependence of chemical shifts are a bit more clearly explained.

Response: We thank the reviewer and are happy to have satisfied their questions and suggestions.

1. The solution structure of free Skp1 is interesting except for the awkward existence of helix-7 out by itself. The comparison with Skp1 complexed with other proteins is useful, but one figure will suffice. The many examples are unconvincingly presented and it is likely that if the relevant parts of Skp1 are overlaid they will be nearly identical.

Response: Our point in displaying the comparison with X-ray, Cryo-EM, and Solution-state dimer structures was to make it clear to the readers that the core domain of the protein remains unchanged; however, none of the structures presented earlier are able to capture the dynamic loops, and the C-terminal H8 region of Skp1.

2. There continue to be major deficiencies. The authors did not satisfactorily address the monomer/dimer status of the molecule, despite strong published evidence that the human Skp1 sequence is a homodimer (kd = ~1 μM), which was confirmed by the recently published solution structure of a similar ortholog. There is a real possibility that they have this wrong but, if they are right, there are simple ways to confirm that.

Response: We understand the concern of the reviewer, and hence we present the following data and explanations to support the monomeric nature of full-length hSkp1. The hSkp1 was reported to be primarily monomeric by Raymond W.M. Ng et al.[5] We used a combination of biophysical techniques to demonstrate that the hSkp1 exists primarily as a monomer. We have added a section to our main text on Page no. 6 with all our biophysical characterization for monomeric Skp1, and the section is as follows:

"The pioneering studies on hSkp1 describe its role in the context of binding to Skp2 and Cul1.[5] In this study, Skp1 was reported to exist as a monomer, although a small amount of dimer population was detected in the biophysical characterization studies. However, *Dictyostelium* Skp1 and its orthologs like OCP2 are known to exist in the dimeric form, and they share considerable sequence similarity to hSkp1.[7-9] To address this, we characterized the protein samples used in our study using various orthogonal biophysical techniques. The purified Skp1 from bacterial culture shows a band around 46 kDa for GST-Skp1 and at 19 kDa after protease

treatment for pure Skp1 on the SDS-PAGE gel. Further, the MALDI-TOF analysis shows a major peak at 19 kDa, indicating the size of the hSkp1 in our studies is 19 kDa which corresponds to monomeric hSkp1, and a small peak corresponding to [2M+H]+ at 38kDa in the MALDI. Next, to investigate the presence of any dimeric Skp1 species, we performed the SV-AUC experiment at three different concentrations to check the concentration-dependent dimerization. Upon increasing the protein concentration in AUC experiments, we observe that the sedimentation coefficient of hSkp1 does not shift, and the sedimentation coefficient is in range of ~2.4 S. The AUC results clearly indicate that no higher molecular weight species are present even with an increase in concentration Fig 1(C). We further recorded 1H-15N HSQC spectra of Skp1 starting from higher (1mM) to lower (50µM) concentrations by diluting the same sample (Fig S1). We observe no concentration-dependent shift in peaks or appearance of peak doublets Fig1(C). Thus, we conclude that Skp1 does not shift from monomer to dimer with the increase in concentration. Hence, we infer that the hSkp1 remains largely in the monomeric state."

For the reference of the reviewer, we add the Fig 1 below:

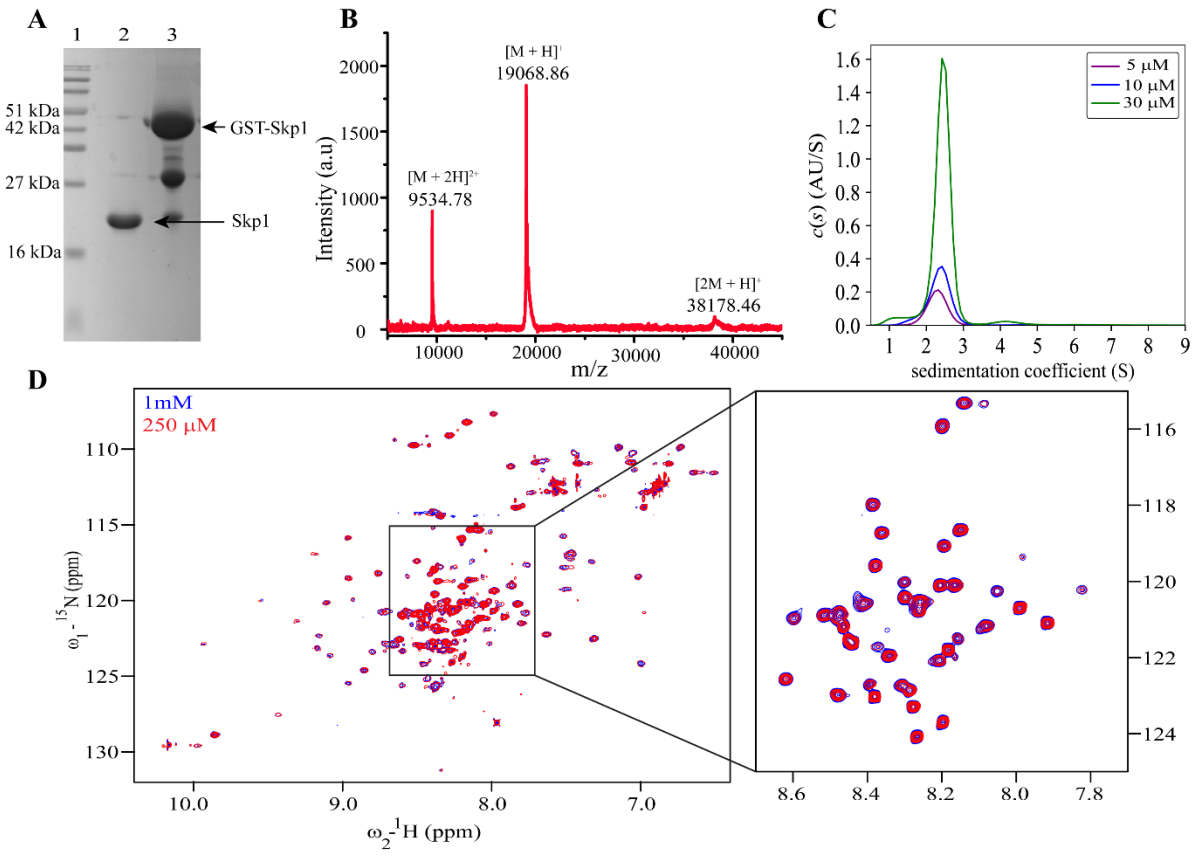
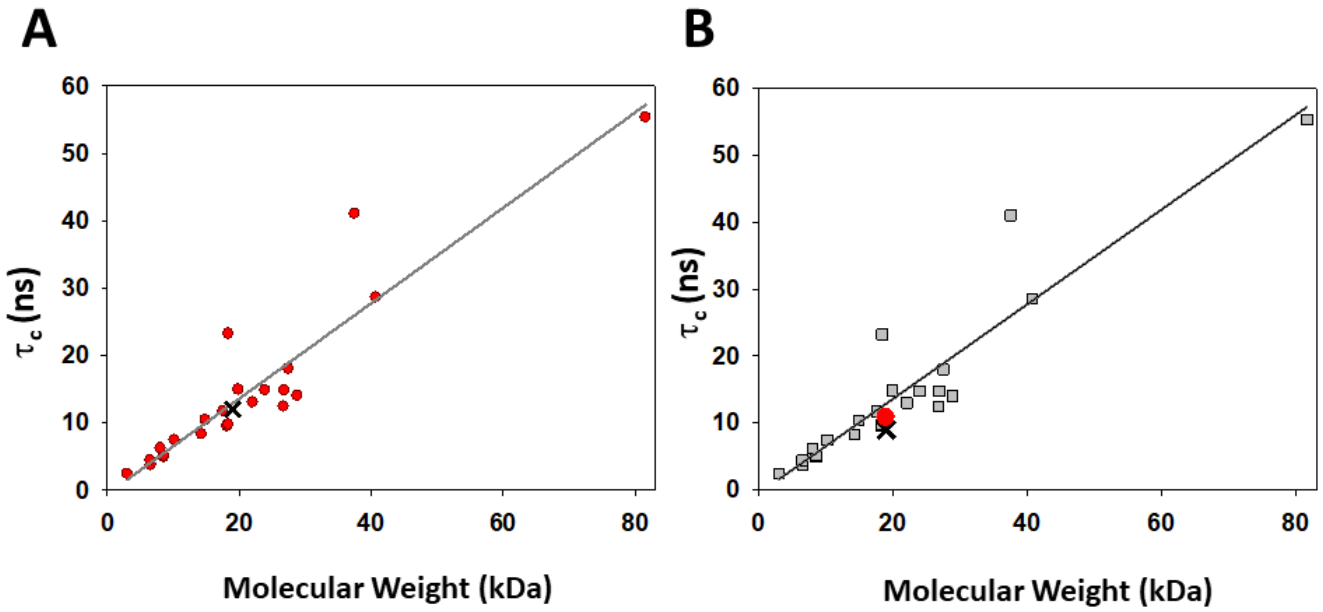


Fig 1: Characterization of Skp1 as a monomer. (A)15% SDS-PAGE Analysis of Skp1 Purification. Lane 1: Protein Marker Lane 2: Eluted Protein. Skp1 Band corresponding to ~19kDa Lane 3: GST-Skp1 protein (band corresponding to ~46 kDa) bound to GSH agarose beads. (B) MALDI-TOF analysis of Skp1. The major peak corresponding to 19068.86 belongs to monomer having z=1 while peak corresponding to 9534.78 is because of z=2. (C) Concentration dependent sedimentation velocity analysis of Skp1. (D) Overlay of ^{1}H-^{15}N HSQC spectra of Skp1 at two different concentrations (250 µM spectra in red and 1 mM spectra in blue).

We have further calculated correlation time τ_c from our backbone dynamics data and it matches closely with the correlation time of a protein of 19kDa. We have also added the following section to the main text on Page no. 11:

"We report that the overall τ_c of Skp1 and it is more or less in agreement with the correlation time of a ~20kDa protein. The overall τ_c of Skp1 is around 10.8 ns which agrees with protein of 19-20 kDa size, which further suggests that Skp1 is monomeric in solution."

We have added the plot here for the reference of the reviewer:



S13: The figure below represents the plot of experimentally determined average rotational correlation times τ_C (as reported in Table 1 of Ryabov et al. (2006)[6] versus molecular weight (determined from the entries of the respective PDB files).

(A) The cross sign shows the position of a 19 kDa globular protein on the standard curve, while the red dots represent the position of proteins reported by Raybov et al.

(B) The grey dots are of the same curve as red dots in (A). The cross sign shows the position of a Skp1 based on τ_C derived from the R_2/R_1 values. The red dot represents the position of Skp1 based τ_C derived from the R_2/R_1 values of the ordered helical regions of Skp1.

We concur with Henzl et al.[7,9] and the work by Kim et al.[8] group on *Dictyostelium* Skp1 dimer that given the very high concentration of Skp1/OCP2 in the human ear, it may tend to dimerize and the dimeric form may have cellular functions. However, we obtained the monomeric form in our studies and have reported our structural and dynamics observations on the monomeric form of Skp1. In this study, our focus is to characterize the IDR regions of Skp1 in the unbound form, unlike the bound X-ray structures.

3.The authors continue to largely ignore published precedent for many of their findings, including the organization of the core structure in solution, and the dynamic nature of the loops and the C-terminal region. They instead state irrelevant details that incorrectly dismiss

the relevance of the previous publications. This violates established norms in scientific publishing.

Ans: We thank the reviewer for pointing out this important point. We are sorry that we missed these important contributions, as indicated by the reviewer. In the revised version, we added all previous work on Skp1, which pioneered the studies to the best of our knowledge. We have included the following changes in our discussion section and hope that credit to previous studies is more conclusive in the current version of the manuscript.

We understand the main concern is regarding the structural features of Skp1, which have been discussed in previous works. Thus, we include the following changes to our text:

The following text has been added to the discussion section on Page no. 16: “The dynamic nature of unstructured regions of Skp1 was initially suggested for OCP2 from CD spectra and HSQC of OCP2.”[7,9]

We have further added a section in our main text: Page no 6 on monomeric hSkp1. There is a previous study first characterized the monomeric Skp1 using biophysical and biochemical experiments.[5]

Our full-length NMR solution structure and backbone dynamics studies substantiate these findings and add residue-level details to the Skp1 understanding.

4. The comparison of BTB structures adds no evident insights to Skp1 function, as previously complained. This was like showing images of all the members of the TPR or Ig family. No useful conclusions were drawn from this.

Response: We elaborated on the importance of the BTB domain and its existence in proteins unrelated to Skp1, as per the suggestions of the other reviewers. The main point we are trying to make here is that it is a versatile conserved domain involved in proteins involved in different pathways across various cellular functions.

5. The extensive discussions of the poorly conserved loop 2 are still unconvincing as there is nothing new presented, since it did not render even in the cryoEM structure cited.

Response:

The IDR regions of free monomeric Skp1 are structural features that have not been captured by any other studies (including cryo-EM). Since this is the first experimental study to elucidate

the poorly conserved loop 2 of Skp1, we focus on the conformation of this loop, as seen in the solution-state structure. Previous studies on Skp1 have presented the folded core of Skp1, including the X-ray structures, Cryo-EM structure, and Skp1 dimer NMR structure .[8]

The pioneering study on human Skp1 by Raymond W.M.Ng et al.[5] for Skp1 interacting with Skp2 defined that the N-terminal regions are also important for binding to Skp2 and deleting only the C-terminal 36 residues did not completely inhibit the Skp1-Skp2 binding. Their construct NΔ114, which had N-terminal residues up to res114(including loop2), showed weak binding to Skp2 as compared to the NΔ67, which had the first 67 N-terminal residues deleted (along with the loop2). Thus, it has been demonstrated that multiple regions in Skp1 participate in F-box binding, and the region (67-114), including loop2, may have an important role to play in the F-Box recognition.

Our discussion on loop2 explores its position in context to other structural elements. The mode by which it is held by the two flanking helices that are part of the folded core (N-terminal region implicated in binding as stated earlier).

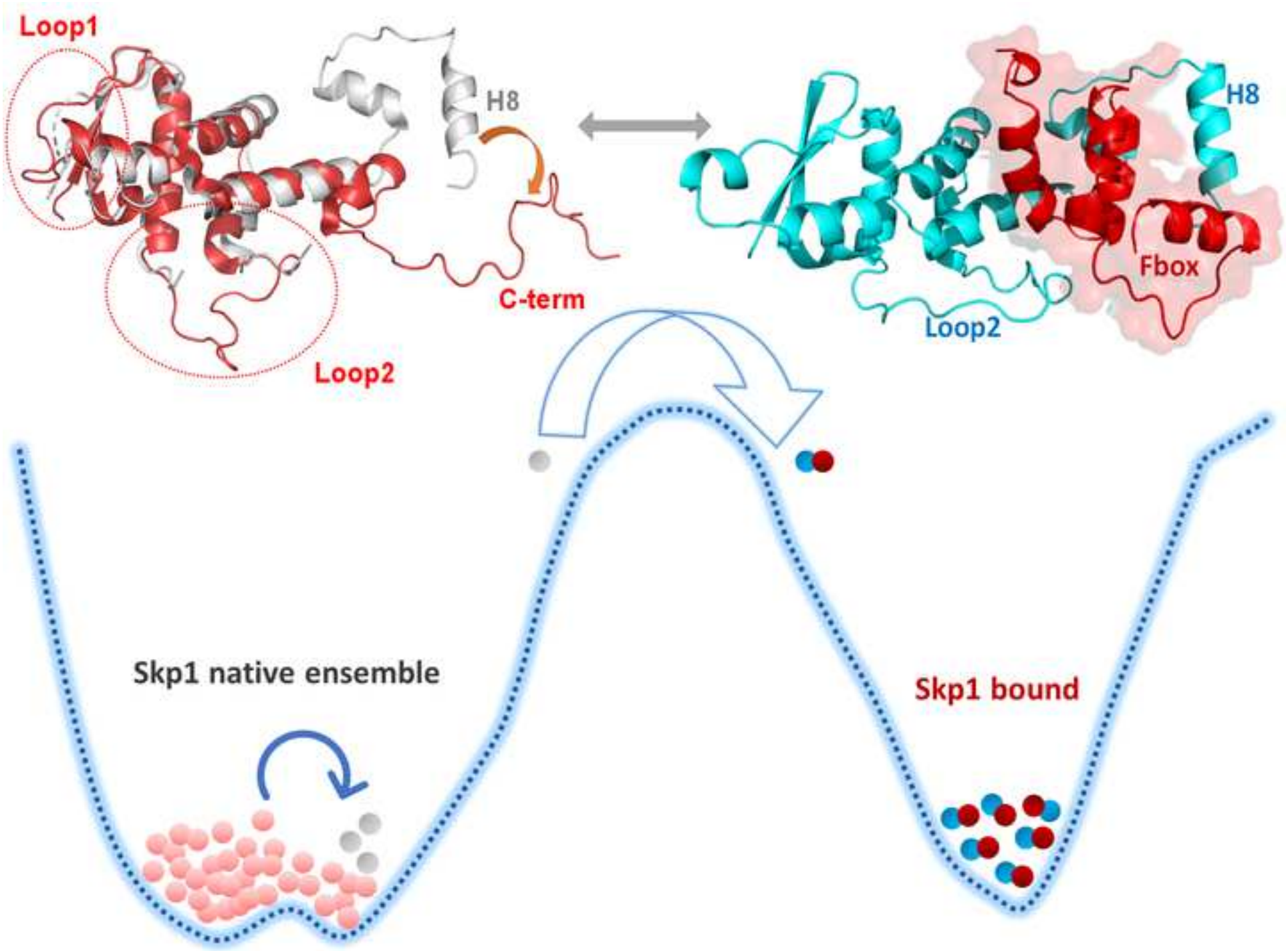
With this new structural insight, we propose that loop2 might be held in a particular way by the two helices as seen in the structure. We further corroborated our previously published MD simulation studies where loop2 was shown to be contributing toward Skp2 binding. This inference was in agreement with the proposition of Raymond W.M.Ng et.al.[5]

6.Writing continues to be redundant, unfocussed, and incorrect in places.

Response: We have tried our best to rectify our writing as suggested by the reviewer and hope that we address the majority of issues that existed in the previous version of the manuscript.

References:

1. Kay, L. E., Torchia, D. A., & Bax, A. (1989). Backbone dynamics of proteins as studied by nitrogen-15 inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry*, *28*(23), 8972-8979.
2. Kaplan, J., & Fraenkel, G. (1980). NMR of Chemically Exchanging Systems (Academic.)
3. Kaplan, J. I., & Fraenkel, G. (1980). Chapter IV-Relaxation. *NMR of chemically exchanging systems*, 27-56.
4. Kneller, J. M., Lu, M., & Bracken, C. (2002). An effective method for the discrimination of motional anisotropy and chemical exchange. *Journal of the American Chemical Society*, *124*(9), 1852-1853.
5. Ng, R. W., Arooz, T., Yam, C. H., Chan, I. W., Lau, A. W., & Poon, R. Y. (1998). Characterization of the cullin and F-box protein partner Skp1. *FEBS letters*, *438*(3), 183-189.
6. Ryabov, Y. E., Geraghty, C., Varshney, A., & Fushman, D. (2006). An efficient computational method for predicting rotational diffusion tensors of globular proteins using an ellipsoid representation. *Journal of the American Chemical Society*, *128*(48), 15432-15444.
7. Henzl, M. T., Thalmann, I., & Thalmann, R. (1998). OCP2 exists as a dimer in the organ of Corti. *Hearing research*, *126*(1-2), 37-46.
8. Kim, H. W., Eletsky, A., Gonzalez, K. J., van der Wel, H., Strauch, E. M., Prestegard, J. H., & West, C. M. (2020). Skp1 dimerization conceals its F-box protein binding site. *Biochemistry*, *59*(15), 1527-1536.
9. Tan, A., Tanner, J. J., & Henzl, M. T. (2008). Energetics of OCP1–OCP2 complex formation. *Biophysical chemistry*, *134*(1-2), 64-71.



Research Highlights:

- Solution structure of free Skp1 reveals the helices, particularly H8 is disordered in its unbound form.

- Structural conservation allows Cul-1 binding at the NTD of Skp1, whereas structural plasticity allows F-box recognition at the CTD of Skp1. Similar mechanisms may be adopted by other adapter proteins, for substrate recognition.

- Conformational sampling of excited sub states by Skp1 residues, which are involved in F-box interactions indicated the excited state may be responsible for recruiting F-box proteins

Disorder in the human Skp1 structure is the key to its adaptability to bind many different proteins in the SCF complex assembly

Amrita Bhattacharya[1‡], Vaibhav Kumar Shukla[2,3‡], Nitin Kachariya[1,4], Preeti[1], Parveen Sehrawat[1], Ashutosh Kumar[1*]

‡ These authors contributed equally

Corresponding author:
Prof. Ashutosh Kumar
Room no. 606, Department of Biosciences and Bioengineering
Indian Institute of Technology Bombay
Powai, Mumbai
India, 400076
Email id: ashutoshk@iitb.ac.in

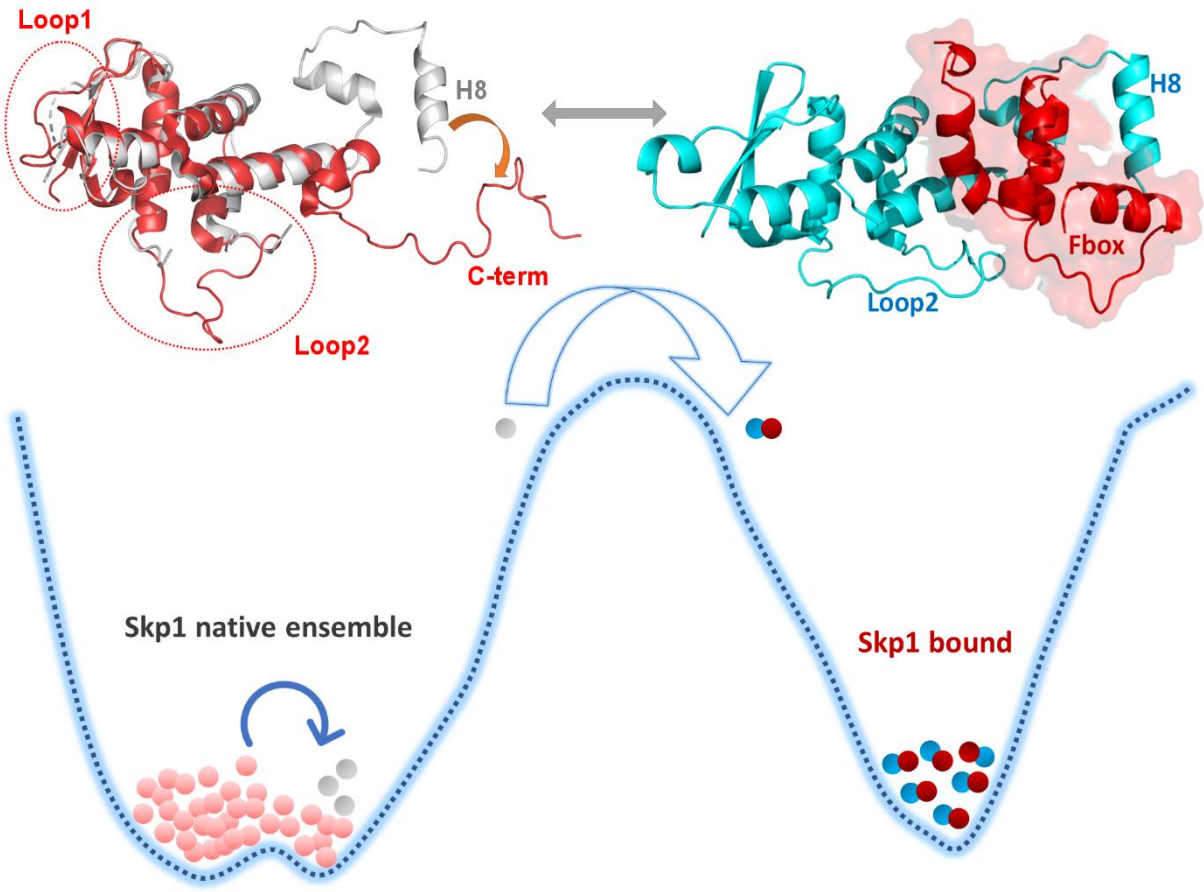
1. Lab No. 606, Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Powai, Mumbai, 400076, India
2. Biophysical Chemistry & Structural Biology Laboratory, UM-DAE Centre for Excellence in Basic Sciences, University of Mumbai, Vidyanagari Campus, Mumbai, 400098, India
3. Present Address: Institute of Structural and Molecular Biology, Division of Biosciences, University College London, London, WC1E 6BT, UK
4. Present Address: Bavarian NMR Center, Department Chemie, Technische Universität München, Munich, Germany

Abstract:

Skp1(S-phase kinase-associated protein 1 - *Homo sapiens*) is an adapter protein of the SCF(Skp1-Cullin1-Fbox) complex, which links the constant components (Cul1-RBX) and the variable receptor (F-box proteins) in Ubiquitin E3 ligase. It is intriguing how Skp1 can recognise and bind to a variety of structurally different F-box proteins. For practical reasons, previous efforts have used truncated Skp1, and thus it has not been possible to track the crucial aspects of the substrate recognition process. In this background, we report the solution structure of the full-length Skp1 protein determined by NMR spectroscopy for the first time and investigate the sequence-dependent dynamics in the protein. The solution structure reveals that Skp1 has an architecture: β1-β2-H1-H2-L1–H3-L2-H4-H5-H6-H7(partially formed) and a long tail-like disordered C-terminus. Structural analysis using DALI (Distance Matrix Alignment) reveals conserved domain structure across species for Skp1. Backbone dynamics investigated using NMR relaxation suggest substantial variation in the motional timescales along the length of the protein. The loops and the C-terminal residues are highly flexible, and the (R_2/R_1) data suggests μs-ms timescale motions in the helices as well. Further, the dependence of amide proton chemical shift on temperature and curved profiles of their residuals indicate that the residues undergo transitions between native state and excited state. The curved profiles for several residues across the length of the protein suggest that there are native-like low-lying excited states, particularly for several C-terminal residues. Our results provide a rationale for how the protein can adapt itself, bind, and get functionally associated with other proteins in the SCF complex by utilising its flexibility and conformational sub-states.

Keywords: Skp1, SCF complex, Protein Structure, NMR Dynamics, Structural adaptability

Graphical Abstract



Research Highlights:

- The solution structure of free Skp1 reveals that the helices, particularly H8, are disordered in the unbound form.
- Structural conservation allows Cul-1 binding at the NTD (N-terminal domain) of Skp1, whereas structural plasticity allows different F-box recognition at the C-terminal helices of Skp1. Other adapter proteins may adopt similar mechanisms for substrate recognition.
- Conformational sampling of excited sub-states by Skp1 residues, which are involved in F-box interactions, indicates that the excited state may be responsible for recruiting F-box proteins.

List of Abbreviations:

DALI	Distance Matrix Alignment
UPS	Ubiquitin Proteasome System
SCF	Skp1-Cullin1-F-box
Skp1	S-phase Kinase associated Protein-1
RING	Really Interesting New Gene
HECT	Homology to E6AP C-Terminus
VHL	Von Hippal-Lindau
FBP	F-Box Proteins
NMR	Nuclear Magnetic Resonance
CSV	Chemical Shift Variance
IDR	Intrinsically Disordered Region
PDB	Protein Data Bank
NOE	Nuclear Overhauser Effect
RBR	RING-Between RING
Cul-1	Cullin-1
Rbx1	RING Box protein-1
Cks1	Cyclin-dependent Kinases regulatory Subunit
BTB	Broad-Complex, Tramtrack, and Bric à brac
POZ	POx virus and Zinc finger
SOCS	Suppressor Of Cytokine Signalling
RMSD	Root Mean Square Deviation
MALDI-TOF	Matrix Assisted Laser Desorption and Ionization-Time Of Flight
Cryo-EM	Cryo-Electron Microscopy
OCP	Organ of Corti Protein
SV-AUC	Sedimentation Velocity Analytical Ultra Centrifugation
HSQC	Heteronuclear Single Quantum Correlation Spectroscopy

Introduction:

The Ubiquitin proteasome system (UPS) is responsible for tagging proteins with ubiquitin and targeting them for degradation. Multiple enzymes in this pathway carry out the sequential modification, processing, and degradation of proteins via UPS.[1] The degradation process is tightly regulated and ensures quality control.[2,3] Any dysregulation in the process can be deleterious leading to pathological conditions including neurodegenerative diseases[4], cell cycle dysregulation[5], inflammatory diseases[6,7], cancer [8,9] etc.

A sequential three-step cascade of enzymes facilitates the process of ubiquitylation. E3 ligases catalyse the final step of the ubiquitin conjugation process, i.e., covalent attachment of ubiquitin to the substrate.[10,11] This penultimate step is the most important step in terms of rendering substrate specificity, and there exists a large class of E3 ligases to efficiently catalyse this final step of Ubiquitin conjugation.[12] E3 ligases have been classified into three major families; namely Really Interesting New Gene (RING) family, Homology to E6AP C-terminus (HECT) and (Ring Between Ring) RBR family.[13,14]

Among the RING E3 ligases, Skp1-Cullin1-F-box (SCF) RING E3 ligases are the largest RING E3 ligases, and their target substrates are actively involved in many key cellular processes like cell cycle regulation, signal transduction, and DNA replication.[15] Alterations in the ubiquitylation by the SCF ligases or their F-box proteins have resulted in genomic instability therefore these complexes are being actively pursued as anti-cancer targets.[16,17] SCF ligases consist of the following components: A central scaffold protein, Cullin; an adapter protein Skp1; Ring box protein (Rbx1); and substrate binding FBPs. The C-terminus of Cullin is the attachment site for Rbx1 which eventually binds to the E2-Ubiquitin-conjugating enzyme. The N-terminus of the Cullin is the binding site for Skp1 which binds with the substrate-binding proteins.[18,19,20] Skp1 acts as an adapter in the SCF complex connecting the constant parts of the ligase to its variable substrate receptors, the F-box proteins. The human Skp1 is unique as it recognises 69 different F-box proteins and plays a critical role in ensuring fidelity in cell quality control.[21,22] Recent reports have suggested that reduced Skp1 expression is directly associated with the DNA double-strand break and chromosome instability leading to cancer [23,24,25,26]. A functional homologue of Skp1, Elongin C is another adapter protein that has been implicated in tumour suppressor activity and cell cycle regulation via VHL(Von Hippal-Lindau) complex.[27,28,29]

Previous structural studies have presented X-ray structures of the SCF complex, where Skp1 has been crystallized or complexed as part of the multiprotein E3-ligase complex.[30,31] These

studies often used truncated or partially mutated protein constructs for structure determination by crystallography approaches.[32,33,34] However, since the loops are structurally conserved across species, their functional aspects cannot be ignored.[35] To date, no standalone structure of Skp1 is available, which can be compared with the X-ray structures to determine the changes occurring in Skp1 upon F-box binding.

In this background, we report the structural and dynamic features of the human Skp1 investigated by solution-state NMR spectroscopy. Our results reveal a complex dynamic nature of the protein, displaying motions at different time scales in different portions of the chain. The dynamic nature of C-terminal helices could be a key feature of adapter proteins among the E3 ligase, which dictate the recognition and modulation of the conformational dynamics in order to bind F-box proteins and subsequent ubiquitylation. Comparison of the solution structure of the free protein with the structures of the protein in the different complexes reported earlier by different methods reveals residue-specific structural changes of C-terminal helices H6, H7, and H8 to account for adaptation upon substrate binding.

The NMR structure of Skp1 provides the three-dimensional structure of the native state of human origin Skp1. We have further explored the native state conformational ensemble by monitoring the amide proton chemical shift as a function of temperature. Small perturbations in temperature are often used to populate lowly-populated excited alternative conformational states in a protein. Since we noticed that Skp1 could alter its structure in different functional conformations from our structure, we explored the existence of alternative conformations in the native-state ensemble. These alternate states provide a glimpse into the functional states that protein might adopt during substrate binding.

Results

1. Characterization of Skp1 as a monomer

The pioneering studies on Human Skp1 (hSkp1) describe its role in the context of binding to Skp2 and Cul1.[36] In this study, Skp1 was reported to exist as a monomer, although a small amount of dimer population was detected in the biophysical characterization studies. However, *Dictyostelium* Skp1 and its orthologs like OCP2 are known to exist in the dimeric form and share a considerable sequence similarity to hSkp1.[37-44] To address this, we characterized the protein samples used in our study using various orthogonal biophysical techniques. The purified Skp1 from bacterial culture shows a band around 46 kDa for GST-Skp1 and at 19 kDa after protease treatment for pure Skp1 on the SDS-PAGE gel. Further,

the MALDI-TOF analysis shows a major peak at 19 kDa, indicating the size of the hSkp1 in our studies is 19 kDa which corresponds to monomeric hSkp1, and a small peak corresponding to $[2M+H]^+$ at 38kDa in the MALDI. Next, to investigate the presence of any dimeric Skp1 species, we performed the SV-AUC experiment at three different concentrations. Upon increasing the protein concentration, we found that the sedimentation coefficients of hSkp1 do not shift, and it is in the range of ~2.4 S. The AUC results clearly indicate that no higher molecular weight species are forming with an increase in concentration Fig 1(C). We further recorded ^{1}H-^{15}N HSQC spectra of Skp1 from higher (1mM) to lower (50µM) concentrations by diluting the same sample (Fig S1). We observe no concentration-dependent shift in peaks or appearance of peak doublets Fig1(C). From the above observations, we conclude that Skp1 does not shift from monomer to dimer with the increase in concentration and the hSkp1 remains largely in the monomeric state.

2. Solution Structure of human Skp1 (Skp1)

The structure of human Skp1 (hereafter referred as Skp1) in solution has been determined by NMR spectroscopy using standard experiments that collected NOE-based restraints and used energy minimization to obtain an ensemble of lowest energy structures. The structures have been deposited in the PDB under accession number (5XYL), and the details of resonance assignments have been deposited in BMRB under accession no. (26765)[45]. The experimental details for protein sample preparation and NMR data acquisition have been s7ummarized in the previous report[45], and structure calculation has been elaborated in the Methods section. The final structure shown in Fig 2 illustrates an architecture of Skp1 with defined secondary structural elements arrangement as β1-β2-H1-H2-loop1-H3-loop2-H4-H5-H6-H7(partial)- and a long-disordered C-terminal tail. The helix H7 is relatively shorter as compared to the bound form of Skp1 and is absent in some members of the ensemble. Skp1 has two distinct loops: loop1 (residue 35-49), loop2 (residue 67-84). It is worth mentioning that both the loops, especially loop2, are deleted or mutated in all X-ray structures Fig S2(A). Our first high-resolution solution-state structure provides an appropriate module for evaluating the roles of loops and other flexible regions in the recognition mechanism of Skp1 with different partners in the SCF complex assembly. The Skp1 full-length structure shows there are parts of Skp1 that are globular, and loops intersperse these domains, along with a long C-terminal unstructured region. This arrangement distorts the structure from a classical globular state to a unique extended shape. Subsequently, the Skp1 solution structure was explored based on the variability of chemical shifts of the structural ensemble to explore the structural heterogeneity. The C^{α} chemical shifts are highly sensitive to secondary structure changes and

are often used as an observable measure for structural features. Chemical shift variance (CSV) indicates the change in the chemical shift values predicted from the ten energy minimised structures of Skp1 5XYL. The chemical shift values were predicted for C^{α} and backbone H^{N} using ShiftX2. The CSV analysis reveals loop2 is particularly highly variable, and its CSV - C^{α} is comparable to the IDR-like C-terminal. The function of the deleted loop2 was elaborated in previous work from our group, where molecular dynamic simulation of full-length Skp1 bound to Skp2 indicated that loop2 is critical in optimizing the distance needed for ubiquitin transfer to the substrates.[46] The C-terminal region contains helices H7 and H8 as defined in the crystal structures of Skp1 present in different complexes (Fig S3).

3. Structural analysis using DALI (Distance Matrix Alignment)

DALI, an automated method developed by Lisa Holm's group, allows for comparing and clustering all known structures of a particular protein, and selective comparison can also be performed between PDB structures. Different folds and domains are often a common feature between protein families, and DALI also allows for pairwise alignment for structure comparison.[47] The self-comparison score will be maximal in any case due to zero deviations. Larger structures often get a higher score than small structures when compared themselves. If a structure has multiple domains, then the DALI score often allows for better matches for small domains than a partial match to larger domains. DALI also provides a platform that integrates structure comparison with sequence alignment.[48]

Fig 3(A) shows an overlaying NMR structure (5XYL) with one of the X-ray structures. For simplicity, only one X-ray structure and one structure from the NMR ensemble have been shown. We used 5XY as the query structure for the DALI search. In the DALI correlation plot, structural domains of the proteins are expected to show a strong positive correlation within domains and a negative correlation in-between domain. As can be noticed in Fig 3(B), a positive correlation is evident in the N-terminal of Skp1. This indicates that there is a major motif present at the N-terminal region of Skp1, and this corresponds to the BTB-POZ fold domain.[49] To date, only Cullin-1is known to bind at the N-terminus, which indicates that this structural fold is essential to form a conserved interaction interface between Skp1 and Cul1, and the comparison of the conserved residues is shown in Fig 3(C). Interestingly, most of the helical core of Skp1 is highly conserved in terms of hydrophobicity and these hydrophobic residues are replaced in other species by similar amino acids to retain the globular domain of Skp1. The gap in the middle region of the correlation curve is from residues 60-80, which indicates the lack of domain organization in loop2. As expected, there is a relatively low correlation at the C-terminal end, which is indicative of the variable structures at this end. The C-terminus of Skp1 lacks a defined domain organisation due to structural flexibility which might

contribute in recognising the different F-box proteins. Towards the C-terminus, the polar residues increase in number, as shown in blue in Fig 3(C), which probably contributes to the flexible helices and more solvent accessibility of Skp1. The C-terminus does not pack with the globular core and remains bouncing in the surrounding cellular environment, where it can bind and recruit the F-box proteins. Overall, the structural and biological role of Skp1 can be attributed due to this charge conservation. Hence, the interaction of H8 with F-box and the subsequent conformational changes could be facilitated by the highly charged residues present at the C-terminus of Skp1.

Within an RMSD of 2-3 Å and a residue alignment of residues 90-100, several interesting BTB, Kelch family proteins show a high structural similarity with Skp1. This includes proteins like COP9 signalosome complex subunit 1, suppressor of kinetochore protein 1, transcription regulator protein BACH1, Myoneurin, and many adapter proteins like Elongin C. Thus, adapter proteins and several transcription factors show structural similarity to Skp1. Details into structural comparison reveal similarities not only in the conserved helical core of Skp1 but also in the loop regions indicating the importance of flexible regions in the structures of these proteins (Additional Supplementary Files Dali_1). The BTB fold forms the structural core of many proteins, which are not orthologs of Skp1 and are often are not functionally related. The fold has a typical helical conformation that may confirm structural stability and is found in diverse protein structures. The structural and further dynamic studies on Skp1 could shed light on other biologically relevant events like cell cycle, cell signalling, and protein recruitment to multiunit enzyme complexes like SCF.

4. Backbone dynamics of Skp1 shows differential motion across the length

We performed longitudinal relaxation rates (R_1), transverse relaxation rates (R_2), and steady-state heteronuclear NOEs ({^{1}H}-^{15}N NOE) to evaluate the backbone dynamics of Skp1. R_1 relaxation rates report fast ns-ps motions, whereas R_2 measurements capture the contributions from both ns as well as the slower µs-ms exchange motions, while {^{1}H}-^{15}N NOE are sensitive to the fast ps motions.[50] Low NOE values indicate high amplitude ps time scale motions, high R_1 values indicate significant ns time scale motions, and high R_2 values suggest possibilities of conformational exchange contributions at the µs-ms time scale.[51,52] From such data, different strategies have been used to get insights into slow conformational exchange contributions along the sequence of a given protein molecule.[53,54]These include Model Free analysis, Spectral Density analysis, R_2*R_1 vs. sequence plots, and R_2/R_1 vs. sequence plots. Among these, Model Free analysis is typically used for well-folded proteins and is not suitable for intrinsically disordered regions in the proteins, as in the present case.[55,56] In an approximate

sense, R_2 is linearly dependent on correlation time (τ_c) and R_{ex}, R_1 is inversely dependent on τ_c, and therefore, the product R_2*R_1 should be roughly independent of τ_c in the absence of R_{ex}. Thus, any variation in this product is indicative of sequence-wise variation in R_{ex} contributions. Among the spectral densities, $J(\omega_H)$ is contributed mostly by the ^1H–^{15}N NOE and reflects on high-frequency motions; $J(\omega_N)$ and J(0) functions have contributions from R_2 and R_1, and among these, only J(0) has contributions from slow motions, which includes chemical exchange. We have used all three approaches, namely, R_2/R_1 vs. sequence, R_2*R_1 vs. sequence, and spectral densities vs. sequence to characterise motions in the Skp1 Fig 4(B).

We performed calculations for for R_1, R_2, and $\{^1H\}$-^{15}N NOE in a residue-specific manner, excluding the residues with overlapping NMR signals and ten prolines Fig. 4(A). Interestingly, all the relaxation parameters display significant sequence-dependent variations, indicating substantial variations in the internal dynamics along the length of the protein. The heteronuclear NOE displays typical bell-shaped profiles with the dips occurring at the loop regions and the termini. These low NOE values indicate fast motions in the ps time scale.[56,57] The dips at the termini (1-3, 145-163) are quite common, and at the loops (34-45, 73-85) are significant suggesting high amplitude motions in these loops. Wherever the NOEs are low, the R_1 values are high, suggesting significant ns time scale motions in those regions. The bell-shaped pattern is seen in the R_2 values, with the maxima occurring at the helical regions. This pattern almost coincides with the NOE data.

Our data shows that the R_2*R_1 product displays substantial sequence-wise variations indicating possible contributions from R_{ex}. The same is true for the R_2/R_1 vs. sequence plots and the J(0) vs sequence plot, where J(0) is more sensitive to slow motions Fig 4(B). Interestingly, the helices exhibit conformational exchange from all three data sets; however, it is not possible to quantitatively estimate the contributions from such data. Interestingly this is also can be indicated by the curved temperature dependence of amide proton chemical shifts. We report the overall τ_c of Skp1 is around 10.8 ns which is in agreement with protein of 19-20 kDa and Skp1 is monomeric in solution. The τ_c value is close to those of other standard globular proteins[58] but shows some deviation as shown in Fig S13, which results from the different R_2/R_1 values of loops and disordered C-terminal.

In conclusion, it emerges from the dynamics data that Skp1 is geared up to adapt structurally. The C-terminal appears to be particularly adaptable as it is disordered and flexible in solution. Earlier studies proposed a role for the flexible loop1 residues 37- 42 in displacing Cand1 from the Cullin for binding of Skp1.[59] MD simulation studies have also proposed that D82 and D83 residues are involved in interaction with C-terminal tail residues of FBP Skp2.[46] These residues are part of the highly dynamic loop2 which shows motion in the ps-ns timescale Fig 4(A).

5. Curved Temperature dependence of amide proton chemical shifts of Skp1

The chemical shift change of amide proton generally exhibits a linear dependence as a function of temperature.[60] Any non-linearity is attributed to the possibility of sampling conformations of higher energy states.[61] Non-linearity in temperature dependent changes in chemical shift is because a residue makes an excursion between a native state and a higher energy state.[60] The high energy state could be either due to local conformational change along the backbone or due to changes in the surrounding side-chain packing which can influence each other.

The assumption is made that each of the two states can have a linear chemical shift temperature dependence as $\delta_1 = \delta^o_1 + g_1 T$ and $\delta_2 = \delta^o_2 + g_2 T$, where g_1 and g_2 are the gradients of temperature dependence, δ_1 and δ_2 are the chemical shifts of the native and excited states, respectively, and T is the temperature. If P_1 and P_2 are the corresponding populations of the native and the excited states, the observed chemical shift, δ_{obs}, of the amide proton will be given by:

$$\delta_{obs} = \delta_1 P_1 + \delta_2 P_2 \qquad (1)$$

These populations depend on the free energy difference between the two states. If more states contribute, the observed shift will be a weighted average over all accessible states. Thus, the complex dependence of chemical shifts on multiple thermodynamic parameters leads to the non-linear behaviour of the chemical shifts with temperature change.

To understand the influence of these factors, Kumar et al. have performed simulations of H^N chemical shift variation with temperature in the range 288 K–309 K, using a two-state model following the procedure described:[61]

$$\delta_{obs} = \{(\delta_1^0 + g_1 T) + [(\delta_2^0 + g_2 T)\, e^{-(\Delta G/RT)}]\} / [1 + e^{-(\Delta G/RT)}] \qquad (2)$$

where, ΔG is the free-energy difference between the two states, and in $\Delta G = \Delta H - T\Delta S$, ΔH and ΔS are the enthalpy and entropy differences between the two states, respectively.

The δ_{obs} obtained using the above equation were best fitted to a linear equation, and the residuals were plotted against temperature to generate a profile of the deviations. The results show that the curves for ΔG variation, range from 1–4 kcal/mol, with δ_1, δ_2, g_1, and g_2 being kept constant. [61] It is interesting to note that the curvature almost disappears for $\Delta G \geq 3$ kcal/mol in the given temperature range. From the models based on the above simulations by

Kumar et al.[61] the energetics of the lowly populated states depicting curved temperature dependence were obtained and following observations were made.

The energy difference between conformational sub-states relative to the native conformer can be estimated from the temperature dependence of amide proton chemical shift if the alternative states are separated by an energy barrier of 2-3 kcal/mol. Most proteins access low-lying sub-states for $\leq$ 1% of the time, and such excited states are often believed to be crucial for biological function. We have used the above observations and assumptions in our studies of the temperature dependence of amide proton chemical shifts of Skp1 to gain insights into alternate conformational states.

Any structural heterogeneity in the protein conformers would result in curved temperature dependence. In the case of Skp1, structural insights indicate that H8 undergoes a structural transition from helix to random coil in the absence of a binding partner. It does not pack with the helical core of the protein and hence shows greater flexibility in free Skp1. If the structure of the protein is globular and temperature stable, then the temperature dependence will be expected to be perfectly linear. However, if the protein structure is more breathable or dynamic, as in the case of Skp1 or many other proteins, then the temperature dependence deviates from linearity. The increase in the temperature of a solution leads to an increase in the average distance between atoms. This leads to an increase in the interatomic distance between the atoms of the H-bond, consequently weakening the H-bond. Both inter and intra-molecular H-bond influence magnetic bond anisotropy (σ_{ani}), which is directly proportional to the amide chemical shift. If the H-bond is solvent-accessible, then the interaction with the solvent molecule may also influence the amide bond. The chemical environment of the amide bond also influences the H-bond; in particular, the aromatic residues surrounding an H-bond influence chemical shift.

We probed the temperature dependence of the amide proton chemical shift of residues of the Skp1 in the native conditions to gain insight into alternate possible excited state conformations.

A small cluster of C-terminal residues (150-160) shows curved temperature dependence in Skp1. These could indicate an alternative conformation similar to the folded state of H8 at the C-terminal. Apart from this, some other residues of Skp1 also show curved temperature dependence, shown in Fig. 5 and Fig. S11. The plots summarise the most prominent residues with curved temperature dependence and some residues lacking such a profile. Based on theoretical simulations of temperature dependence [60,61], which showed the relationship of ΔG between the ground and excited conformers of a protein with the degree of curvature, we estimated the approximate ΔG between the native and excited states of Skp1. If the ground state chemical shift is upfield compared to the excited state, then the shape of the curve is

concave. Conversely, if the ground state chemical shift is downfield of the excited state, then the shape of the curve is convex. The larger the curvature in temperature dependence, the smaller the value of ΔG between the ground and excited states of Skp1.[60]

Comparing the simulations and experimental curved temperature dependence profiles of Skp1 in the same temperature range, we infer that the alternative conformations are within 2kcal/mol from the ground state. For many residues shown in Fig 5A, residue like R154 shows a distinct convex profile. In cases the curvature profile is very steep, like residue R154, the energy difference lies between 1-2kcal/mol [60,61] (Fig 5). It is interesting to note from Fig 5 that the residues exist in tandem in the H8 region and have convexly curved temperature profiles. In Fig 5B, we notice that the residues in tandem have convex as well as concave profiles. When residues are upfield shifted from their random coil values in the native state, they often show a positive temperature gradient, resulting in concave shapes. Overall, from Fig 5 and Fig S11, we see that Skp1 has several residues or regions which show curved temperature dependence. Thus, alternative conformations are indicated even in the native state ensemble of Skp1.

Discussion

The structure and dynamics at different time scales described above for Skp1 in conjunction with the reported structural studies by X-ray crystallography and Cryo-EM present a comprehensive insight into different complexes formed by Skp1 with other F-Box proteins. In the following paragraphs, we attempt to provide insights into details of the adaptability of Skp1 for binding to a variety of proteins in the SCF complex assembly.

(a) Human Skp1 structure has a shallow energy landscape

The curved temperature dependence of amide proton chemical shifts suggests that many residues in the C-terminal, some in helices, and the loops access low-lying excited states (within 2 kcal /mole). The energy barriers between lowly populated excited states and the ground-state for these residues are not too high and hence can interconvert very rapidly. The R_2 which captures the μs-ms dynamics also substantiates this observation. The shallow energy landscape has a significant implication for the protein as it can rapidly change its structure to facilitate binding to varieties of proteins.

Previously, X-ray crystallography was used to solve structures of Skp1 in complex with various F-Box proteins via its C-terminus, and the comparison of the C-terminal from various available X-ray structures is shown in Fig S4[32-34,62] The C-terminal tail in these X-ray structures are well

folded, comprising of helices H6, H7, and H8 however the length and helicity of H7 and H8 were distinct in various structures. NMR structural details (Table1) indicate that H6 and H7 are partially formed, and the H8 region is completely disordered in the solution structure. Comparative structural analysis suggested a coil-helix transition in the C-terminus of Skp1 upon binding to F-box. The nature of the binding partner or conditions dictates the structure of the C-terminal tail of Skp1 (Fig S4). Evidently, this is facilitated by the dynamics in that region and the shallowness of the native state landscape, as revealed by the curved temperature dependences of amide proton chemical shifts. Next, we performed an RMSD comparison for the well-defined regions of the solution structure with an X-rays structure. An excellent correlation was observed for all well-folded structural regions, listed in Table 2. Many commonalities in the orientation and position of the ordered regions can be seen there in supplementary Fig S6 and S7.

Similarly, the SCF-RBR complex structure was solved using Cryo-EM at a resolution of 3.91Å (Fig S5).[63] Skp1 was bound to Cul1 via its N-terminal and Skp2-Cks1(F-box-substrate protein complex) at the C-terminal (PDB ID: 7B5M).[63] RMSD comparison between solution-state and cryo-EM structures showed many resemblances in the well-structured parts apart from the structural details of loop1 and loop2, probably due to the lack of conclusive electron density maps. Skp1 with Skp2-Cks1 exhibits a similar H8 conformation at the C-terminus as seen in the Skp2-Skp1 complex (PDB ID: 1FQV) (Fig S6).[62,63] Thus, compared to 5XYL, the Cryo-EM structure appeared very similar to the other X-ray structures, except for the information on loop orientations. Although the loops of Skp1 were retained in this structure, they remained largely un-modelled due to their dynamic nature. Nonetheless, as the Cryo-EM structure (PDB ID: 7B5M) offers the structure of Skp1 in the complex with other components of SCF, it offers an insight into the role of loop1 in binding to Cul1. As shown in Fig 7(C), Cul1 binds to the N-terminal end of Skp1, where loop1 participates in stabilizing the complex. Though the electron density of the loop region was undetermined, the modelled portion of this interaction supported the idea of Cul1 interacting with the N-terminal of Skp1, including loop1 and H1. On the other hand, Skp1 helices H5-H7 interact with the subsite-1 and H8 contributes to substite-2 of Skp2.[62] Further, it is interesting to note that in spite of the low electron density for loop2, its structural orientation was placed in the direction of Skp2. Consequently, certain residues showed close proximity with Skp2 advocating a possible interaction between these two proteins. Thus, in a fully intact form, it can be expected that loop2 engaging with Skp2 works as an interaction enhancer. Previous studies on Skp1 with Molecular Dynamics simulations also proposed the functional interaction of loop2 with Skp2 in properly orienting individual proteins at an optimum orientation to facilitate ubiquitin transfer.[46] All these observations

reflect upon the shallowness of the energy landscape of Skp1, which enables rapid adaptation of the structure to bind to different protein partners.

There are three tryptophan residues along the sequence of Skp1. These three residues are essential for the structural features, and the functional aspects of Skp1.[62] Trp 61 and Trp 88 are parts of Helix-3 (H3) and Helix-4(H4), respectively, and these two helices hold the loop2. The orientation of these two helices seems to help in the motion of loop2. Loop2 is not flattened out but behaves like a skipping rope held by H3 and H4, and this is visually evident from our structural ensemble of Skp1(Supplementary Fig S8). The two tryptophan residues have hydrophobic interactions and may play a role in holding loop2 in shape. (Supplementary Figure S9). W159 is known to be a key player in F-box recruitment and interlocks at the hydrophobic F-box domain of Skp2, acting as a "linchpin".[62]

(b) Comparison with solution structure of Skp1 homodimer

The other available NMR structure of Skp1 is from the *Dictyostelium* double truncated Skp1 homodimer (6V88)[37] (Figure S9). The study claims that Skp1 dimer formation obscures the F-box binding subsite1 of Skp1. Authors proposed that the dimer Skp1 is not the biochemically active conformer; rather, it loses its ability to bind to the F-box domain. The biochemical activity is restored upon reversing the dimer formation with the F97E mutation.[37] Human Skp1 has about 66% identity with the *Dictyostelium* Skp1 in full-length proteins (Fig S14). However, identity increases to more than 90% if we compare the C-terminal helices H6, H7, and H8. The dimer NMR structure lacks all the flexible loops and the C-terminal helices H7 and H8. Studies on *Dictyostelium* Skp1 have previously indicated that the C-terminal H8 may be disordered, and a C-terminal proline residue P143 gets glycosylated. P143 glycosylation is known to play oxygen sensing roles in protists.[38,39] Studies have shown that glycosylation increases the helical propensity and may enhance F-box binding. This P143 residue is absent in human Skp1; thus, the above interesting observations may be unique to *Dictyostelium* Skp1.[39,40]

While our study presents the NMR solution structure of complete Skp1(5XYL), depicting the critical disordered C-terminal region as an IDR in human Skp1.The dimerization of Skp1 ortholog OCP2(Organ of Corti Protein 2) has been previously shown. OCP2 is a protein that is highly abundant in the Organ of Corti of the inner ear. [41,42] OCP2 exists as a dimer before being replaced by OCP1(similar to Skp2), and this could be a possible scenario for Skp1 as well in certain physiological conditions.[43,44] The dynamic nature of unstructured regions of Skp1 was also initially suggested for OCP2 from their CD spectra and HSQC spectra of OCP2.[60,61] Our structure and backbone dynamics studies on full-length NMR solution substantiate these findings on the orthologous protein OCP2.

(c) Disorder is the key to recognition in other proteins homologous to Skp1

It is relevant to point out that the functional implications of flexible IDRs seen in this work are not unique to Skp1 alone. Elongin C is an adapter protein in the cullin ring ligase family which is homologous in function to Skp1.[64,65] In terms of structure, Elongin C is shorter than Skp1 as it lacks the H4 helix, H6 helix, and the two-helix C-terminal extension H7, and H8 of Skp1 It also lacks parts of loop2. Skp1 recognizes F-box protein via the c-terminal helices; similarly, the Elongin B and Elongin C complexes recruit the SOCS box proteins.[66,67,68] Fig S10 shows the loops and the disordered termini of Elongin B, Elongin C as a complex, and the interaction of the IDR termini of Elongin with SOCS1. [66,67,68] The problem of structural details of the loops remaining absent in X-ray structures persists in this structure as well (PDB ID:6C5X).[66] However, from the highlighted regions, it is clear that the interacting interfaces consist of loops and IDRs.

6. Conclusion

We have presented a detailed solution-state structural analysis of the full-length human Skp1in free form. Skp1 is the key adapter protein poised to recruit the receptor F-box protein in the SCF complex. The ability to recruit a large variety of F-box proteins is imparted by the intrinsic structural flexibility of Skp1. This flexibility is ensured by the dynamic loops and the C-terminal helices, which undergo a structural adaptation from an IDR to helices. The helicity of the C-terminal helices, especially the H8 helix, is completely lost in the absence of SCF complex formation. This partially disordered C-term of Skp1 is the key to the formation of SCF complexes and a few key residues regulate the process of H8 formation during F-box recognition. Based on our relaxation studies and recent cryo-EM structures, we speculate secondary interactions of Skp1 with F-box proteins may be via loop2 of Skp1.

We have further attempted to obtain a residue-level view of the energy landscape of the native state structural ensemble of Skp1 using the temperature dependence of amide proton. The study provides insights into the alternative conformational states by residues in tandem, indicating structural features forming transiently in alternative conformations. Since we already concluded from the structural analysis that H8 exists in helical and random coil conformers, these two states could be the alternate sub-states being sampled. The ends of helices, such as residues 108-113, show curved temperature profiles that could be due to the structural flexibility of these regions. We observe a close correlation between F-box contact points and the curved temperature dependence of residues. Therefore, we conclude that shallowness in

the energy landscape is vital in rendering structural adaptability to Skp1 in facilitating its binding to a variety of F-Box proteins.

Materials and Methods:

Skp1 expression and Purification

Human Skp1 (Uniprot ID: P63208) is a 163 aa protein of 19 kDa molecular weight. Skp1 gene construct has been inserted into the pGEX-6p-1 vector between *Bam*H1 and *Xho*1 sites.[45] Skp1 is purified as a Glutathione S-transferase (GST) fusion protein, and post purification, the GST tag is cleaved using PreScission protease. The Skp1-pGEX-6p-1 plasmid construct was transformed in *E. coli* BL21 DE3 cells and was grown in M9 minimal media supplemented with ^{13}C-glucose and/or ^{15}N–NH_4Cl (Cambridge Isotope Laboratories, Andover, MA and Eurisotop, France) for up to 0.6-0.8 O.D. (A_{600}) at 37 °C temperature. The Skp1 fusion protein expression was induced overnight by adding 200 µM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 25^{o}C, and the induced culture was harvested. The cells were suspended in sodium phosphate buffer (20 mM, pH 6.0), 1 mM EDTA, 100 mM NaCl, 1mM β-ME, 1mg ml^{-1} lysozyme, 0.1% Triton X-100 followed by sonication for cell lysis. The cell lysate was bound to GSH-sepharose beads, and the beads were washed with a sodium phosphate buffer with an increasing NaCl gradient. The GST tag was removed by incubating with PreScission protease at 4°C overnight. The protein was eluted down, and the fractions were analyzed by 15% SDS-PAGE. The yield of the purified protein was 30 mg/L of culture medium (as described previously in Biomolecular NMR Assignments, 2016)[45]

Mass spectroscopy analysis

The mass spectra of purified Skp1 were recorded on UltrafleXtreme (MALDI-TOF/TOF mass spectrometer, Bruker Daltonics Co.). Briefly, 2 µL of 20 µM protein solution was mixed with 2 µL of matrix solution (sinapinic acid). External calibration was performed with a mixture of protein standards (5–50 kDa).

Analytical Ultracentrifugation

The sedimentation Velocity ultracentrifugation experiment was done on a Beckmann Coulter Optima Ultracentrifuge. Briefly, AUC cells were assembled with 12 mm double-sector Epon charcoal-filled centrepieces and Quartz windows as per the manufacturer's cell assembly instructions. Appropriate amount of samples were loaded in cells, and the whole assembly was equilibrated at 20^{0}C for 2 hr in An60- Ti rotor prior to the run. Data was recorded at 45,000 rpm using absorbance optics at 280 nm wavelength. SEDNTERP[69] was used to estimate the

specific protein volume and the density and viscosity of the buffer. Data were analysed using continuous c(s) distribution model of the SEDFIT program.[70] Frictional ratio, baseline, Fit RI Noise, and Fit Time Independent Noise parameters were floated as variables during fitting. GUSSI program was used to plot data, fit and c(s) distribution plot.[71]

NMR spectroscopic Measurements

$^{15}N/^{13}C$ or ^{15}N labelled Skp1 protein samples were prepared for NMR in 20 mM sodium phosphate, 100 mM NaCl, 1 mM β-ME (pH 6.0). Concentration dependent ^{1}H-^{15}N HSQC of labelled Skp1 were recorded by mixing the protein and buffer in the ratio of 1:1 to make the concentrations as 1mM, 500 µM, 250 µM, 100 µM and 50 µM in same preparations. D_2O was mixed in 90:10 (H_2O/D_2O) ratio before recording spectra. The proton chemical shifts were referenced using 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) as an external calibration agent at 0.0 ppm, and the sample temperature was maintained at 298K. NMR experiments were recorded on a Bruker AVANCE III 750-MHz spectrometer with a 5-mm triple-resonance inverse (TXI) probe with a Z-gradient.

For ^{13}C-NOESY experiments, lyophilized Skp1 sample was dissolved in 100% D_2O, and spectra were recorded at 298K on Bruker Avance 800 MHz spectrometers equipped with cryo-probe. Spectra were processed by using software Topspin 3.2(Bruker) and NMRPipe[72], and analyzed by CARA-1.8.4.[73] and CcpNmr.[74.]

NMR structure calculations

Distance restraints were obtained from 3D ^{15}N-edited NOESY-HSQC spectra (τ_{mix}-150 ms), ^{13}C-edited NOESY-HSQC (τ_{mix}-120 ms), and ^{13}C (aromatic)-edited NOESY-HSQC spectra (τ_{mix}-120 ms) recorded on a Bruker 800MHz NMR spectrometer. All NOE were assigned manually using CARA-1.8.4,[73] and integrated NOE peaks were calibrated and converted to distance restraints with the program CALIBA.[75] The upper distance limits for structure calculations were derived from cross-peak volumes that were calibrated using the CALIBA peak calibration routine in CYANA-3.0. The torsion angle constraints were obtained from assigned backbone chemical shifts using the program TALOS+.[76]

For the structures of Skp1, 200 randomized conformers were generated, out of which 10 conformers with lowest target function and the lowest energy, with no distance violation > 0.5 Å and no angle violations, were selected. These 10 structures with lowest target functions were further refined using the CNS 1.21 program with explicit water as the solvent which improved the Ramachandran Plot statistics and the *Z*-score for the Procheck (phi-psi) and

Procheck (all) for the ordered residues. The program PSVS v1.4 was used to analyze the quality of the structures, and PYMOL[77] was used for generating figures for structures.

Chemical Shift Variance Calculation

Using the published structure, we performed the variance analysis. For this, the chemical shift of the 10 ensemble members in 5XYL was calculated using ShiftX2. Variance is defined as the squared sum of distances of each term (the chemical shift of each ensemble member) from the mean value, divided by the total number of terms in the distribution (N=10, as there are 10 lowest energy ensemble members). Chemical shift (CS) has been widely used to characterize the structural features of a protein, and this approach takes the predicted CS from the structural ensemble.[79-82]

Covariance is a modified version of the variance which measures the total variation of two variables from their expected values (mean). The covariance values tell us whether the variable being compared tends to be similar or distant from the mean value.[82] Then the covariance was calculated considering the formula:

Cov(res)= $\sum$(S1-Si)/N, where S1 is the first ensemble member of 5XYL and in S_i (i=1 to 10) is the i^{th}, and N=10.

NMR Experiments for Backbone Dynamics of Skp1

The backbone dynamics of Skp1 at ps-ms timescale were performed using R_1, R_2, and Heteronuclear nuclear Overhauser effect ({^{1}H}-^{15}N NOE), relaxation experiments. Longitudinal relaxation rates (R_1) were measured using following delay times: 20 ms, 25 ms, 50 ms, 80 ms, 100 ms, 250 ms, 375 ms, 500 ms, 650 ms, and 800 ms, whereas transverse relaxation rates (R_2) were measured at following delay times:10 ms, 25 ms, 50 ms, 75 ms, 100 ms, 125 ms, 150 ms, 200 ms. Here, 25 ms and 125 ms time points were repeated for the error estimation in R_2, and 100 ms and 650 ms time points were repeated for error estimation in R_1. {^{1}H}-^{15}N NOE measurements were performed where proton saturation delay was for 5s. Protein stability was checked at the end of R_1, R_2, and {^{1}H}-^{15}N NOE experiments by recording 2D HSQC spectra. All the NMR experiments were performed using Bruker 750MHz Avance III spectrometer equipped with TXI Wide bore 5mm probe at 298K. All the spectra were processed using TopSpin 3.2. Data was analysed using CcpNmr 2.4.1. R_1 and R_2 rates were obtained by single exponential fitting, and errors in R_1, R_2, and {^{1}H}-^{15}N NOE were calculated from the fit error.

The reduced spectral density was calculated by using only three ^{15}N relaxation parameters to obtain information about the protein motion by assuming that at high frequencies, the spectral density functions: $J(\omega_H) \approx J(\omega_H - \omega_N) \approx J(\omega_H + \omega_N)$ as described by (Peng and Wagner[83] ; Farrow et al.[84]; Lefevre et al.[85] ; Vis et al. [86]). The values and errors were computed as described by Malik et. al.[87]

An estimation of rotational correlation time of a globular protein was obtained through the ratio of collective backbone ^{15}N R_2 and R_1 relaxation rates.[53] $\tau_c \propto R_2/R_1$. The value of ^{15}N R_1 and R_2 are often used to estimate τ_c and it was calculated using the formula:

$$\tau_c = 1/4\pi\nu_N((6*R_2/R_1)-7)^{1/2}$$

Where R_1 and R_2 are the backbone relaxation rates of the protein and ν_N is the frequency of ^{15}N at 750 MHz.

Temperature Dependence of Amide Proton

Temperature dependence of amide proton chemical shift for the Skp1 was performed by recording a series of HSQC spectra at seven different temperatures 286 K, 289 K, 292 K, 295 K, 298 K, 301, 304, 307, 310, 313, and 315 K. The amide proton chemical shift of each residue was plotted in a straight line, and the residuals were plotted to observe the curved temperature profiles.

Accession Numbers: PDB 5XYL

Acknowledgements:

We acknowledge Prof. R. V. Hosur for his support and advice during the drafting of this manuscript.

AK and AB are grateful to DST (Grant: EMR/2016/002798 DTD. 14/03/2017, Grant: SR/SO/BB-0022/2012 DTD- 22/01/2013) for funding during the project and HF-NMR IIT Bombay, IISER Pune, HF-NMR facility at TIFR, Mumbai for instrument time. AB is grateful to Geetanjali A. Dhotre, MALDI-TOF-MS facility, TIFR, Mumbai. AB is also grateful to CSIR-UGC, MHRD and IRCC, IIT Bombay for funding during Ph.D.

References

1. Heinemeyer, W., P. C. Ramos, and R. J. Dohmen. "Ubiquitin-proteasome system." *Cellular and Molecular Life Sciences CMLS* 61.13 (2004): 1562-1578.

2. Hershko, Avram. "Ubiquitin: roles in protein modification and breakdown." *Cell (Cambridge)* 34, no. 1 (1983): 11-12.
3. Ciechanover, Aaron, and Yong Tae Kwon. "Degradation of misfolded proteins in neurodegenerative diseases: therapeutic targets and strategies." *Experimental & molecular medicine* 47.3 (2015): e147-e147.
4. Zheng, Qiuyang, et al. "Dysregulation of ubiquitin-proteasome system in neurodegenerative diseases." *Frontiers in aging neuroscience* 8 (2016): 303.
5. Nakayama, Keiichi I., and Keiko Nakayama. "Ubiquitin ligases: cell-cycle control and cancer." *Nature Reviews Cancer* 6.5 (2006): 369-381.
6. Wang, Jingsong, and Michael A. Maldonado. "The ubiquitin-proteasome system and its role in inflammatory and autoimmune diseases." *Cell Mol Immunol* 3.4 (2006): 255-261.
7. Marfella, Raffaele, et al. "The ubiquitin-proteasome system and inflammatory activity in diabetic atherosclerotic plaques: effects of rosiglitazone treatment." *Diabetes* 55.3 (2006): 622-632.
8. Liu, Han, Emily H-Y. Cheng, and James J-D. Hsieh. "Bimodal degradation of MLL by SCFSkp2 and APCCdc20 assures cell cycle execution: a critical regulatory circuit lost in leukemogenic MLL fusions." *Genes & development* 21.19 (2007): 2385-2398.
9. Juven-Gershon, Tamar, and Moshe Oren. "Mdm2: the ups and downs." *Molecular medicine* 5.2 (1999): 71-83.
10. Deshaies, Raymond J., and Claudio AP Joazeiro. "RING domain E3 ubiquitin ligases." *Annual review of biochemistry* (2009): 78.
11. Berndsen, Christopher E., and Cynthia Wolberger. "New insights into ubiquitin E3 ligase mechanism." *Nature structural & molecular biology* 21.4 (2014): 301-307.
12. Komander, David, and Michael Rape. "The ubiquitin code." *Annual review of biochemistry* 81 (2012): 203-229.
13. Wei, Dongping, and Yi Sun. "Small RING finger proteins RBX1 and RBX2 of SCF E3 ubiquitin ligases: the role in cancer and as cancer targets." *Genes & cancer* 1.7 (2010): 700-707.
14. Buetow, L., & Huang, D. T. (2016). Structural insights into the catalysis and regulation of E3 ubiquitin ligases. *Nature reviews Molecular cell biology*, *17*(10), 626-642.
15. Cardozo, Timothy, and Michele Pagano. "The SCF ubiquitin ligase: insights into a molecular machine." *Nature reviews Molecular cell biology* 5.9 (2004): 739-751.
16. Thompson LL, Rutherford KA, Lepage CC, McManus KJ. The SCF Complex Is Essential to Maintain Genome and Chromosome Stability. *International Journal of Molecular Sciences*. 2021; 22(16):8544.
17. Russell, Adrian, et al. "Cyclin D1 and D3 associate with the SCF complex and are coordinately elevated in breast cancer." *Oncogene* 18.11 (1999): 1983-1991.

18. Skaar, Jeffrey R., Julia K. Pagan, and Michele Pagano. "SnapShot: F box proteins I." *Cell* 137.6 (2009): 1160-1160.
19. Skaar, Jeffrey R., Julia K. Pagan, and Michele Pagano. "Mechanisms and function of substrate recruitment by F-box proteins." *Nature reviews Molecular cell biology* 14.6 (2013): 369-381.
20. Petroski, Matthew D., and Raymond J. Deshaies. "Function and regulation of cullin–RING ubiquitin ligases." *Nature reviews Molecular cell biology* 6.1 (2005): 9-20.
21. Bai, Chang, et al. "SKP1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box." *Cell* 86.2 (1996): 263-274.
22. Thompson, Laura L., et al. "The SCF Complex Is Essential to Maintain Genome and Chromosome Stability." *International Journal of Molecular Sciences* 22.16 (2021): 8544.
23. Thompson, Laura L., et al. "Reduced SKP1 expression induces chromosome instability through aberrant cyclin E1 protein turnover." *Cancers* 12.3 (2020): 531.
24. Casimiro, Mathew C., et al. "Overview of cyclins D1 function in cancer and the CDK inhibitor landscape: past and present." *Expert opinion on investigational drugs* 23.3 (2014): 295-304.
25. Nakayama, Keiichi I., and Keiko Nakayama. "Ubiquitin ligases: cell-cycle control and cancer." *Nature Reviews Cancer* 6.5 (2006): 369-381.
26. Santra, Manas K., Narendra Wajapeyee, and Michael R. Green. "F-box protein FBXO31 mediates cyclin D1 degradation to induce G1 arrest after DNA damage." *Nature* 459.7247 (2009): 722-725.
27. Takagi, Yuichiro, et al. "Identification of elongin C sequences required for interaction with the von Hippel-Lindau tumor suppressor protein." *Journal of Biological Chemistry* 272.43 (1997): 27444-27449.
28. Benmaamar, Ramla, and Michele Pagano. "Involvement of the SCF complex in the control of Cdh1 degradation in S-phase." *Cell Cycle* 4.9 (2005): 1230-1232.
29. Botuyan, Maria Victoria, et al. "Solution structure and dynamics of yeast elongin C in complex with a von Hippel-Lindau peptide." *Journal of molecular biology* 312.1 (2001): 177-186.
30. Zheng, Ning, et al. "Structure of the Cul1–Rbx1–Skp1–F box Skp2 SCF ubiquitin ligase complex." *Nature* 416.6882 (2002): 703-709.
31. Latres, E., D. S. Chiaur, and M. Pagano. "The human F box protein β-Trcp associates with the Cul1/Skp1 complex and regulates the stability of β-catenin." *Oncogene* 18.4 (1999): 849-854.
32. Gorelik, M., Manczyk, N., Pavlenco, A., Kurinov, I., Sidhu, S. S., & Sicheri, F. (2018). A structure-based strategy for engineering selective ubiquitin variant inhibitors of Skp1-Cul1-F-box ubiquitin ligases. Structure, 26(9), 1226-1236..

33. Hao, Bing, et al. "Structure of a Fbw7-Skp1-cyclin E complex: multisite-phosphorylated substrate recognition by SCF ubiquitin ligases." *Molecular cell* 26.1 (2007): 131-143.
34. Kuchay, Shafi, et al. "GGTase3 is a newly identified geranylgeranyltransferase targeting a ubiquitin ligase." *Nature structural & molecular biology* 26.7 (2019): 628-636.
35. Ng, R. W., Arooz, T., Yam, C. H., Chan, I. W., Lau, A. W., & Poon, R. Y. (1998). Characterization of the cullin and F-box protein partner Skp1. *FEBS letters*, *438*(3), 183-189.
36. Kishan, K. V., & Agrawal, V. (2005). SH3-like fold proteins are structurally conserved and functionally divergent. Current Protein and Peptide Science, 6(2), 143-150.
37. Kim, Hyun W., et al. "Skp1 dimerization conceals its F-Box protein binding site." *Biochemistry* 59.15 (2020): 1527-1536.
38. Sheikh, M. Osman, David Thieker, Gordon Chalmers, Christopher M. Schafer, Mayumi Ishihara, Parastoo Azadi, Robert J. Woods et al. "O2 sensing–associated glycosylation exposes the F-box–combining site of the Dictyostelium Skp1 subunit in E3 ubiquitin ligases." *Journal of Biological Chemistry* 292, no. 46 (2017): 18897-18915.
39. West, Christopher M., Zhuo A. Wang, and Hanke van der Wel. "A cytoplasmic prolyl hydroxylation and glycosylation pathway modifies Skp1 and regulates O2-dependent development in Dictyostelium." *Biochimica et Biophysica Acta (BBA)-General Subjects* 1800, no. 2 (2010): 160-171.
40. Xu, Xianzhong, Alexander Eletsky, M. Osman Sheikh, James H. Prestegard, and Christopher M. West. "Glycosylation promotes the random coil to helix transition in a region of a protist Skp1 associated with F-box binding." *Biochemistry* 57, no. 5 (2018): 511-515.
41. Henzl, Michael T., Isolde Thalmann, John D. Larson, Elena G. Ignatova, and Ruediger Thalmann. "The cochlear F-box protein OCP1 associates with OCP2 and connexin 26." *Hearing research* 191, no. 1-2 (2004): 101-109.
42. Henzl, Michael T., Julie O'Neal, Richard Killick, Isolde Thalmann, and Ruediger Thalmann. "OCP1, an F-box protein, co-localizes with OCP2/SKP1 in the cochlear epithelial gap junction region." *Hearing research* 157, no. 1-2 (2001): 100-111.
43. Tan, Anmin, John J. Tanner, and Michael T. Henzl. "Energetics of OCP1–OCP2 complex formation." *Biophysical chemistry* 134, no. 1-2 (2008): 64-71.
44. Tan, Anmin, and Michael T. Henzl. "Conformational stabilities of guinea pig OCP1 and OCP2." *Biophysical chemistry* 144, no. 3 (2009): 108-118.
45. Kachariya, Nitin Nathubhai, Sarath Chandra Dantu, and Ashutosh Kumar. "Backbone and side schain assignments of human cell cycle regulatory protein S-phase kinase-associated protein 1." *Biomolecular NMR assignments* 10.2 (2016): 351-355.

46. Chandra Dantu, Sarath, Nitin Nathubhai Kachariya, and Ashutosh Kumar. "Molecular dynamics simulations elucidate the mode of protein recognition by Skp1 and the F-box domain in the SCF complex." *Proteins: Structure, Function, and Bioinformatics* 84.1 (2016): 159-171.
47. Holm, Lisa. "Using Dali for protein structure comparison." *Structural Bioinformatics*. Humana, New York, NY (2020): 29-42.
48. Holm, L. “DALI and the persistence of protein shape”. *Protein Science 29*(1), (2020) :128-140.
49. Stogios, Peter J., et al. "Sequence and structural analysis of BTB domain proteins." *Genome biology* 6.10 (2005): 1-18.
50. Kumar, Sandeep, et al. "Folding and binding cascades: dynamic landscapes and population shifts." *Protein science* 9.1 (2000): 10-19.
51. Dyson, H. Jane, and Peter E wright. "Insights into the structure and dynamics of unfolded proteins from nuclear magnetic resonance." *Advances in protein chemistry* 62 (2002): 311-340.
52. Werbeck, Nicolas D., et al. "A distal regulatory region of a class I human histone deacetylase." *Nature communications* 11.1 (2020): 1-9.
53. Shukla, V. K., Kabra, A., Maheshwari, D., Yadav, R., Jain, A., Tripathi, S., ... & Arora, A. (2015). Solution structures and dynamics of ADF/cofilins UNC-60A and UNC-60B from Caenorhabditis elegans. *Biochemical Journal*, *465*(1), 63-78.
54. Kay, L. E., Torchia, D. A., & Bax, A. (1989). Backbone dynamics of proteins as studied by nitrogen-15 inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry*, *28*(23), 8972-8979.
55. Kaplan, J., & Fraenkel, G. (1980). NMR of Chemically Exchanging Systems (Academic)
56. Kaplan, J. I., & Fraenkel, G. (1980). Chapter IV-Relaxation. *NMR of chemically exchanging systems*, 27-56.
57. Kneller, J. M., Lu, M., & Bracken, C. (2002). An effective method for the discrimination of motional anisotropy and chemical exchange. *Journal of the American Chemical Society*, *124*(9), 1852-1853.
58. Ryabov, Y. E., Geraghty, C., Varshney, A., & Fushman, D. (2006). An efficient computational method for predicting rotational diffusion tensors of globular proteins using an ellipsoid representation. *Journal of the American Chemical Society*, *128*(48), 15432-15444.
59. Pierce, Nathan W., et al. "Cand1 promotes assembly of new SCF complexes through dynamic exchange of F box proteins." *Cell* 153.1 (2013): 206-215.
60. Baxter, N. J., & Williamson, M. P. “Temperature dependence of 1H chemical shifts in proteins”. *Journal of biomolecular NMR*, *9*(4), (1997) : 359-369.

61. Kumar, A., Srivastava, S., & Hosur, R. V.” NMR characterization of the energy landscape of SUMO-1 in the native-state ensemble.” *Journal of molecular biology*, *367*(5) (2007):1480-1493.
62. Schulman, Brenda A., et al. "Insights into SCF ubiquitin ligases from the structure of the Skp1–Skp2 complex." *Nature* 408.6810 (2000): 381-386.
63. Horn-Ghetko, D., Krist, D. T., Prabu, J. R., Baek, K., Mulder, M. P., Klügel, M., ... & Schulman, B. A. (2021). Ubiquitin ligation to F-box protein targets by SCF–RBR E3–E3 super-assembly. *Nature*, *590*(7847), 671-676.
64. Okumura, F., Nakatsukasa, K., & Kamura, T. “Role of Elongin BC-containing ubiquitin ligases.” *Frontiers in oncology*, *2*, (2012):10.
65. Liau, Nicholas PD, et al. "The molecular basis of JAK/STAT inhibition by SOCS1." *Nature communications* 9.1 (2018): 1-14.
66. Bullock, A. N., Debreczeni, J. É., Edwards, A. M., Sundström, M., & Knapp, S. “Crystal structure of the SOCS2–elongin C–elongin B complex defines a prototypical SOCS box ubiquitin ligase.” *Proceedings of the National Academy of Sciences, 103*(20), (2006): 7637-7642.
67. Kipreos, E. T., & Pagano, M. “The F-box protein family.” *Genome biology*, *1*(5), (2000):1-7.
68. Liau, N. P., Laktyushin, A., Lucet, I. S., Murphy, J. M., Yao, S., Whitlock, E., ... & Babon, J. J. “The molecular basis of JAK/STAT inhibition by SOCS1.” *Nature communications*, *9*(1), (2018): 1-14.
69. Laue, T. M. (1992). Computer-aided interpretation of analytical sedimentation data for proteins. *Analytical ultracentrifugation in biochemistry and polymer science.*, 90-125.
70. Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophysical journal*, *78*(3), 1606-1619.
71. Brautigam, C. A. (2015). Calculations and publication-quality illustrations for analytical ultracentrifugation data. In *Methods in enzymology* (Vol. 562, pp. 109-133). Academic Press.
72. Delaglio, Frank, et al. "NMRPipe: a multidimensional spectral processing system based on UNIX pipes." *Journal of biomolecular NMR* 6.3 (1995): 277-293.
73. Keller, Rochus Leonhard Josef. "Optimizing the process of nuclear magnetic resonance spectrum analysis and computer aided resonance assignment." PhD diss., ETH Zurich, 2005.

74. Skinner, S. P., Goult, B. T., Fogh, R. H., Boucher, W., Stevens, T. J., Laue, E. D., & Vuister, G. W. (2015). Structure calculation, refinement and validation using CcpNmr Analysis. *Acta Crystallographica Section D: Biological Crystallography*, *71*(1), 154-161.
75. Güntert, Peter, Werner Braun, and Kurt Wüthrich. "Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA." *Journal of molecular biology* 217.3 (1991): 517-530.
76. Shen, Y., Delaglio, F., Cornilescu, G., & Bax, A. (2009). TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *Journal of biomolecular NMR*, *44*(4), 213-223.
77. DeLano, W. L. (2002). PyMOL.
78. Kay, Lewis E., Dennis A. Torchia, and Ad Bax. "Backbone dynamics of proteins as studied by nitrogen-15 inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease." *Biochemistry* 28, no. 23 (1989): 8972-8979.
79. Wishart, David S., and Brian D. Sykes. "Chemical shifts as a tool for structure determination." *Methods in enzymology* 239 (1994): 363-392.
80. Wishart, D. S., Sykes, B. D., & Richards, F. M. (1991). Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *Journal of molecular biology*, *222*(2), 311-333.
81. Mielke, Steven P., and Viswanathan V. Krishnan. "Characterization of protein secondary structure from NMR chemical shifts." *Progress in nuclear magnetic resonance spectroscopy* 54, no. 3-4 (2009): 141.
82. Cochran, William G. "Analysis of covariance: its nature and uses." *Biometrics* 13, no. 3 (1957): 261-281.
83. Peng JW, Wagner G (1992) Mapping of spectral density functions using heteronuclear NMR relaxation measurements. J Magn Reson 98:308–332.
84. Farrow NA, Zhang O, Szabo A et al (1995) Spectral density function mapping using 15 N relaxation data exclusively. J Biomol NMR 6:153–162.
85. Lefevre JF, Dayie KT, Peng JW, Wagner G (1996) Internal mobility in the partially folded DNA binding and dimerization domains of GAL4: NMR analysis of the N–H spectral density functions. Biochemistry 35:2674–2686.
86. Vis H, Vorgias CE, Wilson KS et al (1998) Mobility of NH bonds in DNA-binding protein HU of shape *Bacillus stearothermophilus* from reduced spectral density mapping analysis at multiple NMR fields. J Biomol NMR 11:265–277.
87. Malik, N., Kumar, A. (2016). Resonance assignment of disordered protein with repetitive and overlapping sequence using combinatorial approach reveals initial structural propensities and local restrictions in the denatured state. *J Biomol NMR* (66)**,** 21–35.

Table 1. Experimental restraints and structural statistics for final ensemble of 10 structures of Skp1

Completeness of resonance assignments (%)	
Backbone	94%
Side chain	68%
Aromatic	15%

Conformationally restricting restraints	
Distance Restraint List	
Total	2290
Intraresidue (i = j)	963
Sequential (ji _ jj = 1)	517
medium-range	486
long-range	324
Hydrogen bonds	24
Dihedral angle restrains (ϕ and φ)	207
Disulfide restraints	0
No. of restraints per residue	14.04
No. of long-range restraints per residue	1.99
Residual restraint violations	
Average no. of distance violations per structure	
0.1–0.2 Å	0
0.2–0.5 Å	0
>0.5 Å	0
Average no. of dihedral angle violations per structure	
1–10^{o}	0
>10 ^{o}	0
Model quality	
Rmsd backbone atoms (Å)	0.7 Å
Rmsd heavy atoms (Å)	1.2 Å
Rmsd bond lengths (Å)	0.004 Å
Rmsd bond angles (^{o})	0.7°
MolProbity Ramachandran statistics[a]	
Most favoured region (%)	97.4%
Allowed region (%)	2.6%
Additionally allowed region (%)	0.0%
Disallowed region (%)	0.0%
Global quality scores (raw/Z score)	
Verify3D	0.26 /-3.21
PROCHECK (φ-ψ)	0.07 /0.59

PROCHECK (all)	-0.18/-1.06
MolProbity clash score	37.07/-4.84
Model contents	
Ordered residue ranges	3-34,38-40,43-72,87-126,134-139
Total no. of residues	163
BMRB accession number	26765
PDB ID code	5XYL

[a] region, Skp1: *3-8,11-17,18-33,52-65,87-93,97-110,113-127*

Table 2

RMSD comparison of the Skp1 solution structure ensemble (10 members) with the Skp1 from different complexes for the selected ordered region (helical core).

PDB ID	Information on type of Protein Complex	Method Used to solve Structure	RMSD Range (for comparison with 10 ensemble members)	Average RMSD Å	Error Å
6v88	Skp1A (truncated dimer)	Solution NMR	1.970-2.594	2.30	0.20
6o60	Skp1-FBXL2-GGTase	X-RAY 2.503 Å	2.549-4.059	3.09	0.56
7B5M	SCF-RBR E3-E3 super-assembly: CUL1-RBX1-**SKP1**-SKP2-CKSHS1-p27~Ub~ARIH1. Transition State 2	Electron Microscopy 3.91 Å	2.522-4.391	3.11	0.68

Figures and Artwork

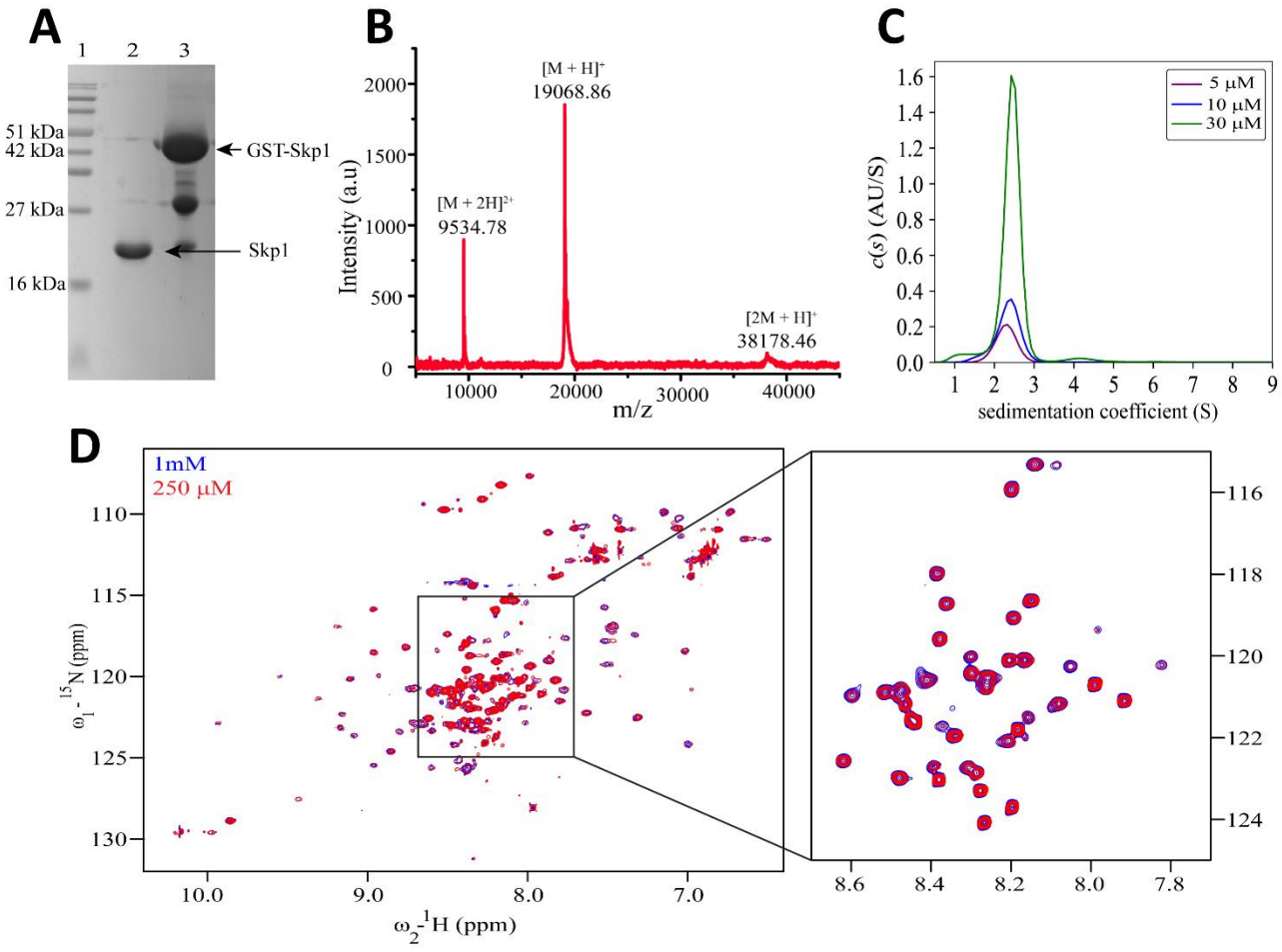


Fig 1: Characterization of Skp1 as a monomer. (A)15% SDS-PAGE Analysis of Skp1 Purification. Lane 1: Protein Marker Lane 2: Eluted Protein. Skp1 Band corresponding to ~19kDa Lane 3: GST-Skp1 protein (band corresponding to ~46 kDa) bound to GSH agarose beads. (B) MALDI-TOF Analysis of Skp1. The major peak corresponding to 19068.86 belongs to monomer having z=1 while peak corresponding to 9534.78 is because of z=2. (C) Concentration dependent sedimentation velocity Analysis of Skp1. As the sedimentation coefficient is almost same at three different concentrations (5μM, 10μM, 30 μM) implying single monomer conformation of Skp1. (D) Overlay of ^{1}H-^{15}N HSQC spectra of Skp1 at two different concentrations (250 μM spectra in red and 1 mM spectra in blue).

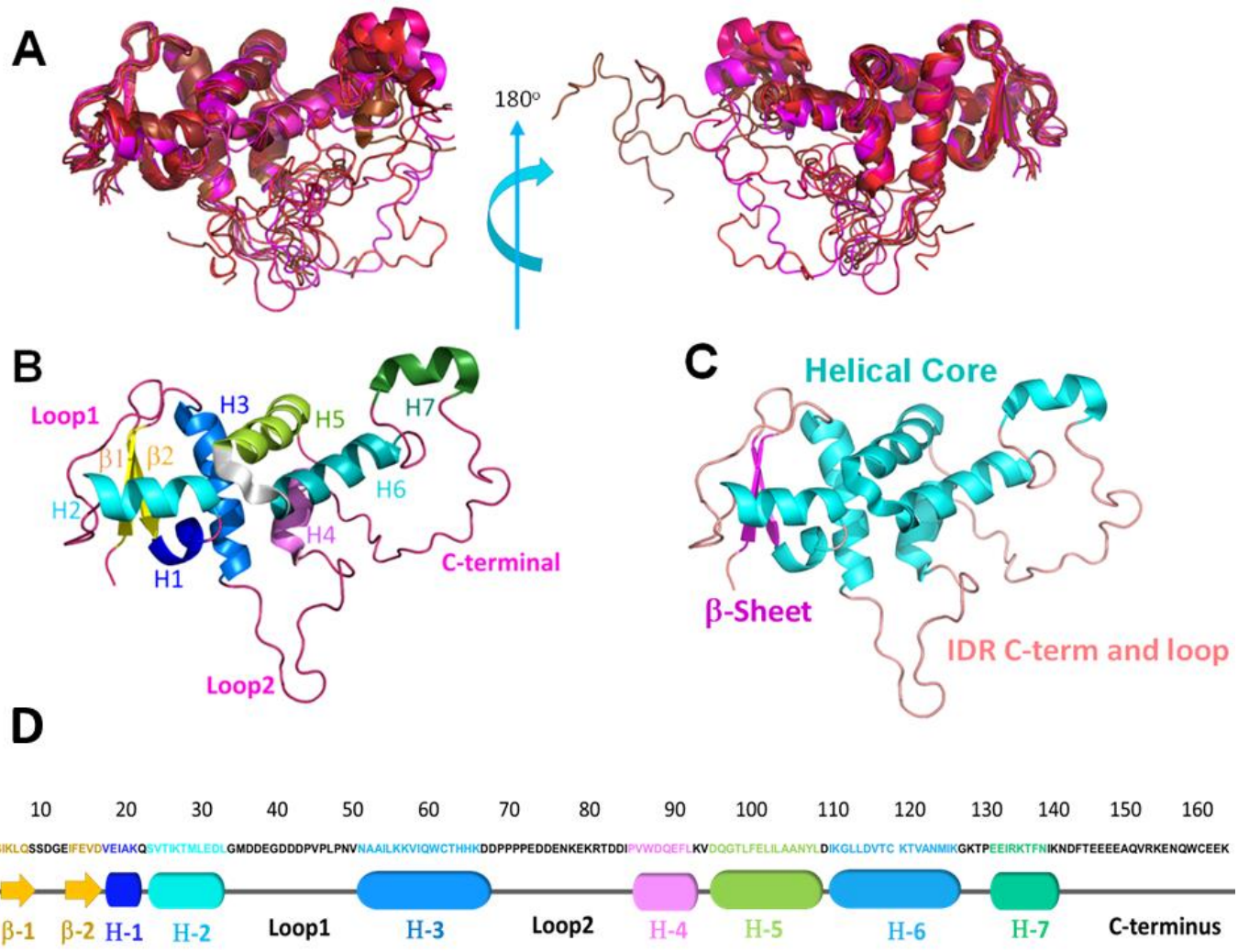


Fig 2: Solution Structure of Skp1 monomer. (A) Ensemble of 10 lowest energy structures of Skp1 derived using solution state NMR restraints; (B) The different structural features of Skp1 in 5XYL; (C) Skp1 full length protein in solution state has a helical core interspersed by two loops and has a disordered C-terminal end; (D) A schematic representation of the secondary structure of Skp1 as obtained from the solution NMR structure 5XYL.

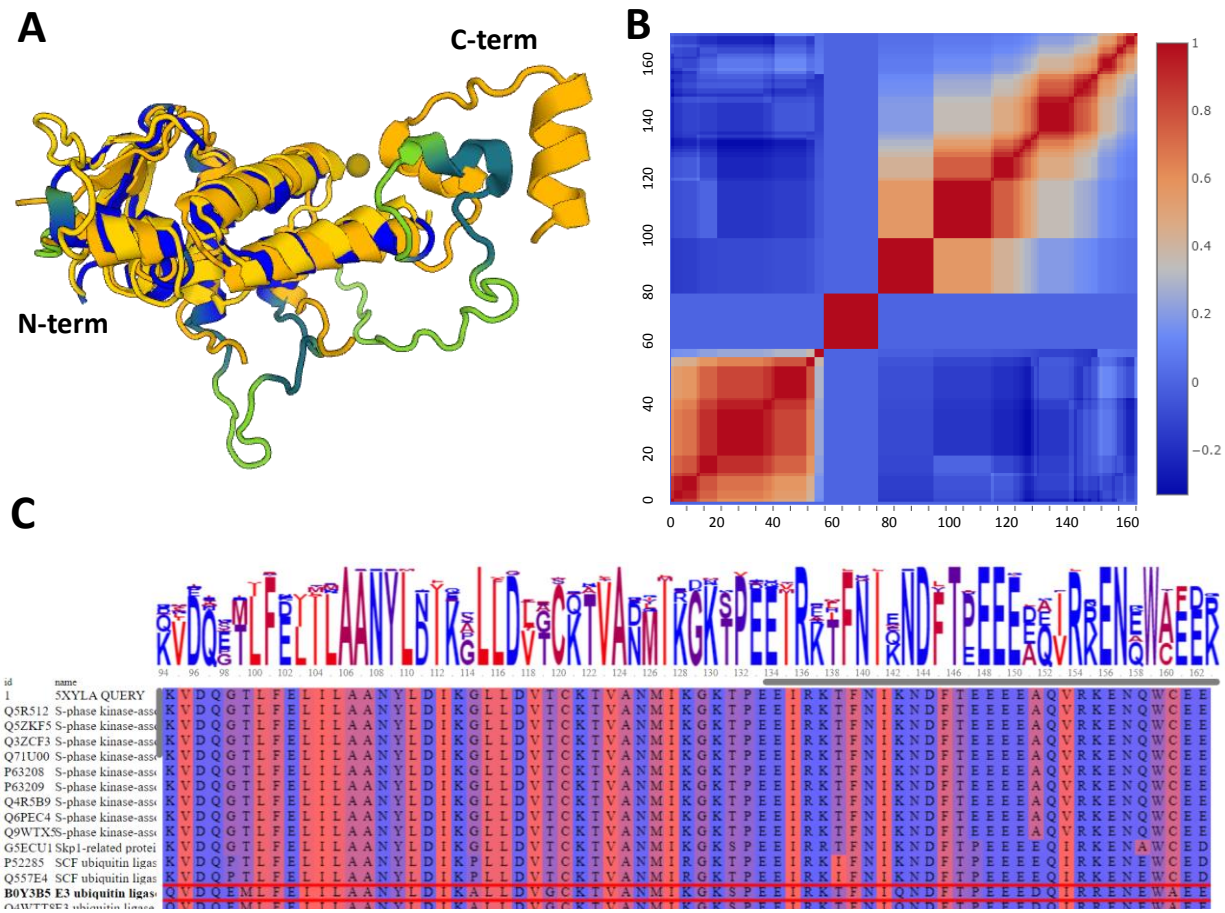
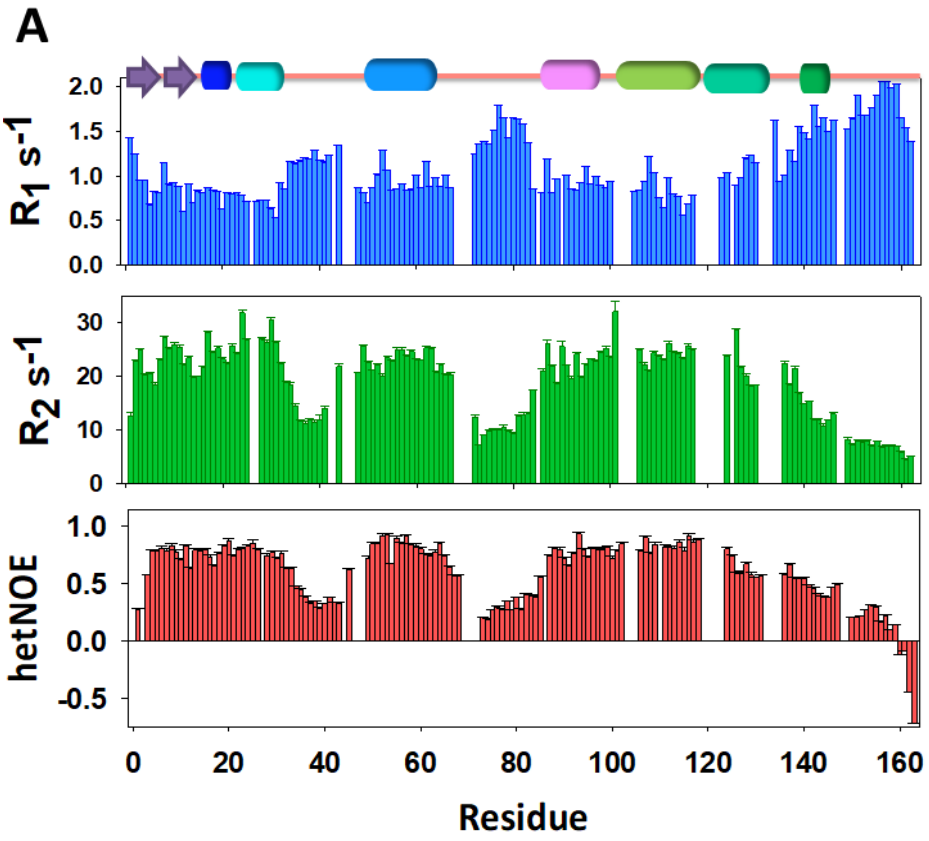


Fig 3: Integrated structure and sequence analysis of Skp1 on the DALI webserver. (A) Structural alignment of 5XYL with X-ray structures. (B) This plot shows the correlation of matched structures along with positions of the query structure. Correlation is undefined if a residue position is empty; in these cases, correlation is set to one or zero (otherwise). X-axis labels are PDB residue numbers, and Y-axis labels are sequential residue numbers (C). These sequences are of the Skp1 C term. The colour indicates the hydrophobicity of residues, with red being hydrophobic and blue indicating blue polar residues. The representations are for 94-162 residues and show that the residues are highly charged conserved. The Pfam analysis of the DALI structural matrix reveals domains that show maximum structural match from different protein families are the BTB, BTB_2, POZ, BACK, and MATH domains to Skp1. Skp1 is a known BTB family protein; however, there are interesting new results from the DALI Pfam, e.g., likeness to ion channels and signalling proteins, for example, the BTB_2 and Ion trans domains in voltage-gated potassium channel beta2-subunit (Additional Supplementary File Pfam_Dali_2).


A
R1 s-1
2.0
1.5
1.0
0.5
0.0
R2 s-1
30
20
10
0
hetNOE
1.0
0.5
0.0
-0.5
0
20
40
60
80
100
120
140
160
Residue

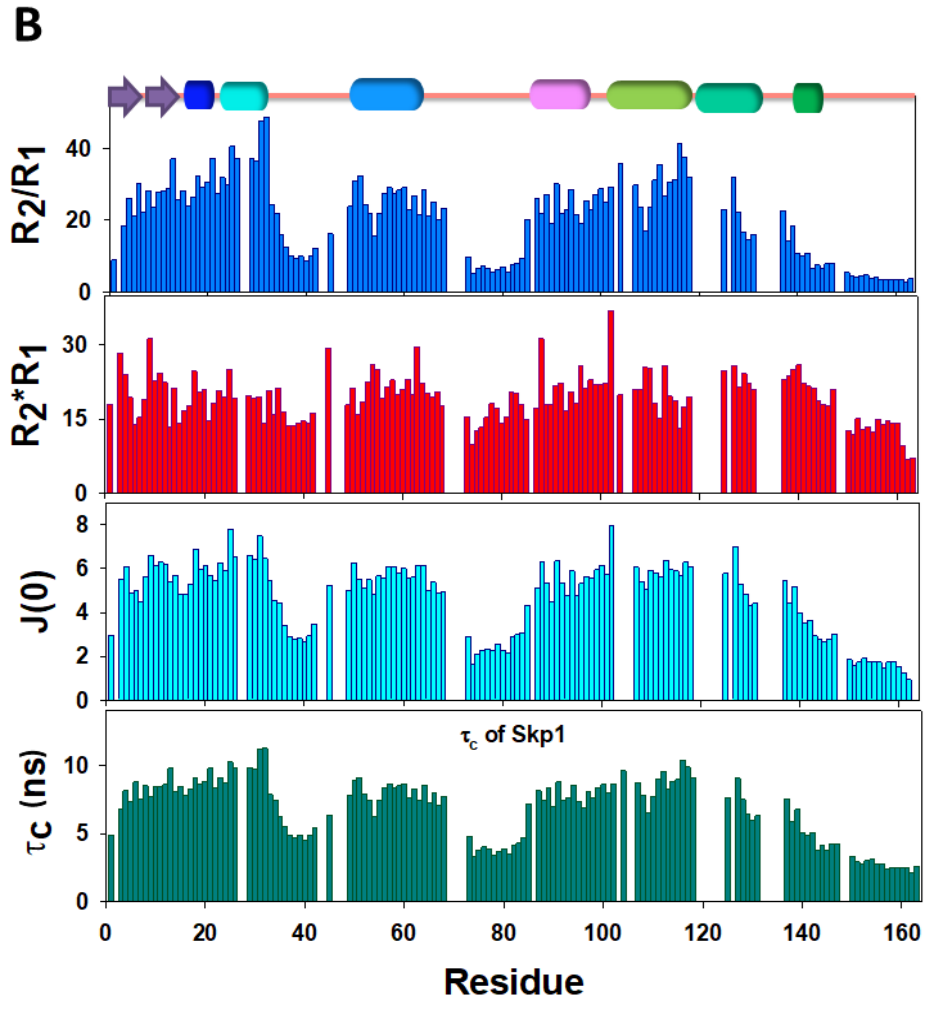

B
R2/R1
40
20
0
R2*R1
30
15
0
J(0)
8
6
4
2
0
τc of Skp1
τc (ns)
10
5
0
0
20
40
60
80
100
120
140
160
Residue

Fig 4: The backbone dynamics of Skp1: The cartoon above each set of plots is indicated as follows: helices as cylinders, beta-sheets as arrows. The loops and IDR C-term are indicated in orange. (A) The backbone relaxation experiments were recorded at 750MHz. The first bar graph indicates the R_1 values(blue) with error. The second bar graph indicates the R_2 values (in green) with error. The third bar graph represents R_2/R_1 values (in pink) along with error. (B) The R_2/R_1 plot is shown in blue and the R_2*R_1 plot is shown in red. The third plot in cyan is of reduced spectral density J(0), while the last plot in dark cyan is of rotational correlation time of Skp1 from R_2/R_1.

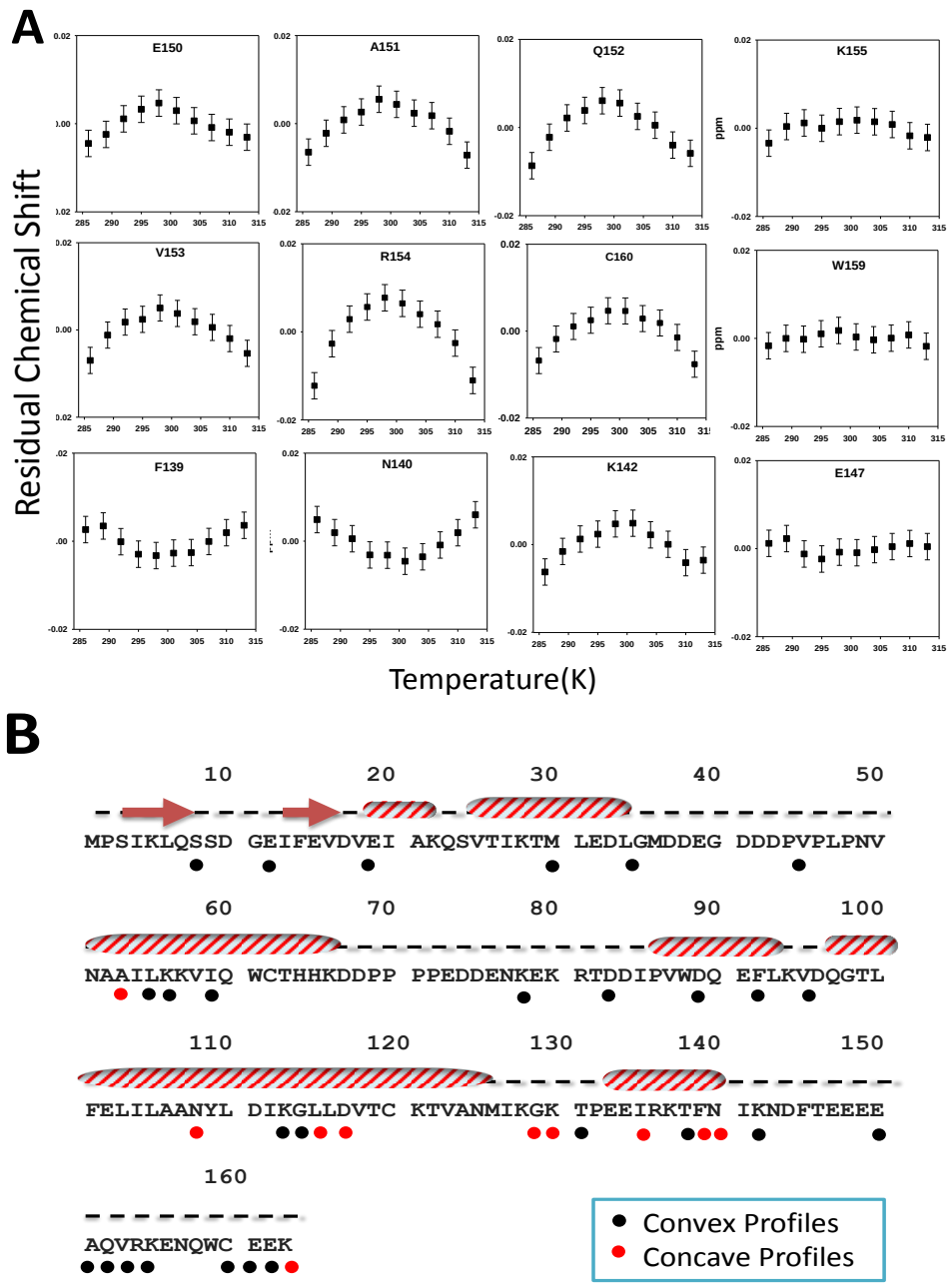

A
Residual Chemical Shift
E150
A151
Q152
K155
V153
R154
C160
W159
F139
N140
K142
E147
ppm
Temperature(K)
B
10 20 30 40 50
MPSIKLQSSD GEIFEVDVEI AKQSVTIKTM LEDLGMDDEG DDDPVPLPNV
60 70 80 90 100
NAAILKKVIQ WCTHHKDDPP PPEDDENKEK RTDDIPVWDQ EFLKVDQGTL
110 120 130 140 150
FELILAANYL DIKGLLDVTC KTVANMIKGK TPEEIRKTFN IKNDFTEEEE
160
AQVRKENQWC EEK
Convex Profiles
Concave Profiles
C

Fig 5. Curved Temperature Dependence of Skp1. (A) The residues of the H8 region in the C-term show curved temperature dependence. Some of the distinct residues with concave, convex, and straight-line profiles are highlighted. (B) Across the Skp1 structural landscape (helix and beta-sheet indicated in cylinder and arrow, respectively), the residues which show concave (red dots) and convex (black dots) profiles are highlighted with dots. C) on the Skp1 NMR structure, the residues in tandem of more than three are indicated in pink dot clusters.

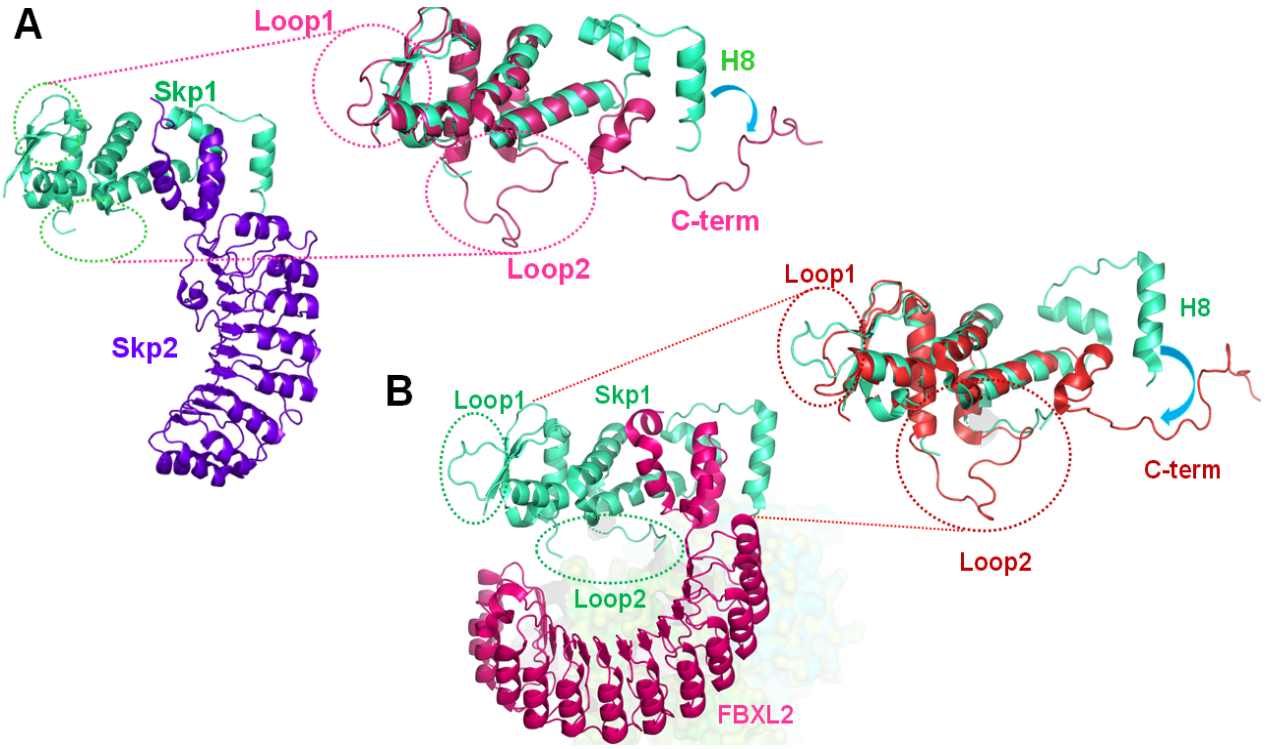


Fig 6: Comparison of X-ray and NMR structures of Skp1 (A) Comparison of Skp1 from 1FQV(Green) vs 5XYL (Dark Pink). (B) Comparison of Skp1 from 6o60(Green) vs 5XYL(Red).

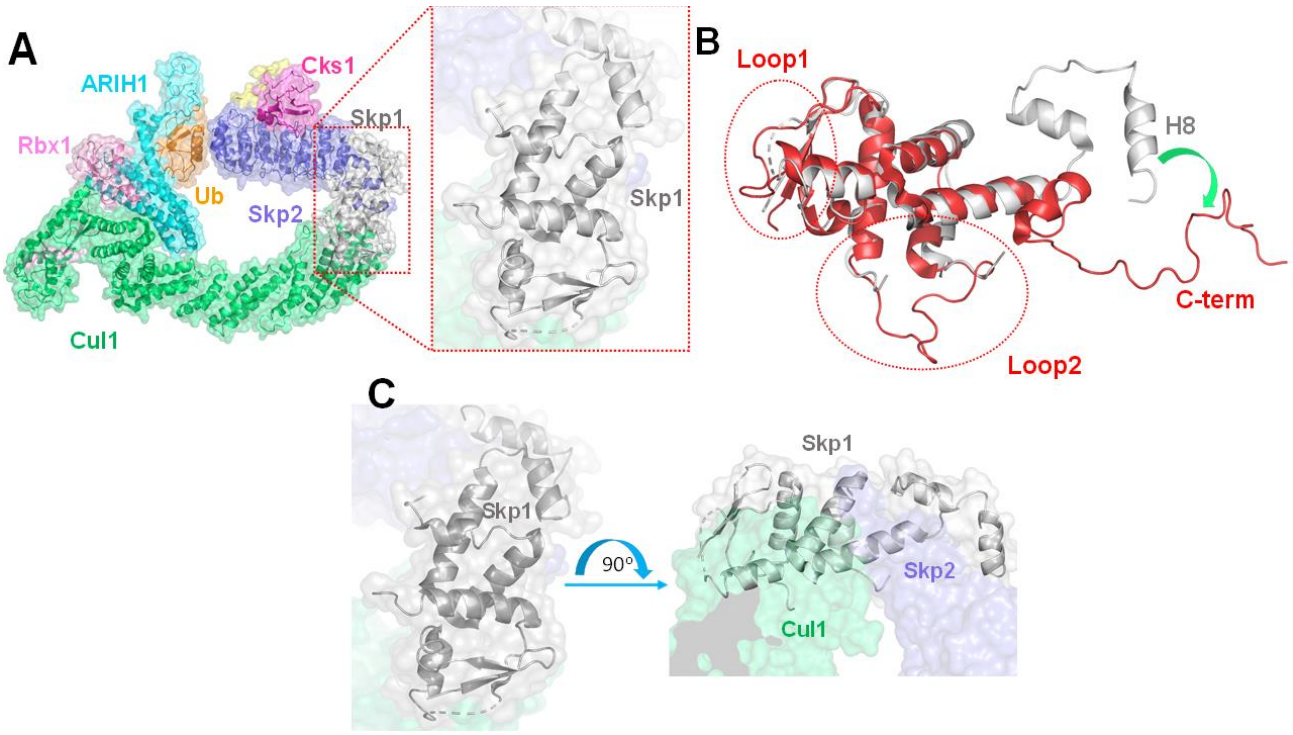


Fig 7: Comparison of the Cryo-Em and NMR Structures of Skp1(A) Skp1(Grey) in the SCF complex assembly as solved by cryo-EM 7B5M. (B) Skp1 from cryo-EM(Grey) vs Skp1 5XYL(Red). (C) Skp1 in interaction with Cul1 and Skp2 as part of SCF complex.

Table 1. Experimental restraints and structural statistics for final ensemble of 10 structures of Skp1

Completeness of resonance assignments (%)	
Backbone	94%
Side chain	68%
Aromatic	15%
Conformationally restricting restraints	
Distance Restraint List	
Total	2290
Intraresidue (i = j)	963
Sequential (ji _ jj = 1)	517
medium-range	486
long-range	324
Hydrogen bonds	24
Dihedral angle restrains (ϕ and φ)	207
Disulfide restraints	0
No. of restraints per residue	14.04
No. of long-range restraints per residue	1.99
Residual restraint violations	
Average no. of distance violations per structure	
0.1–0.2 Å	0
0.2–0.5 Å	0
>0.5 Å	0
Average no. of dihedral angle violations per structure	
1–10°	0
>10 °	0
Model quality	
Rmsd backbone atoms (Å)	0.7 Å
Rmsd heavy atoms (Å)	1.2 Å
Rmsd bond lengths (Å)	0.004 Å
Rmsd bond angles (°)	0.7°
MolProbity Ramachandran statistics[a]	
Most favored region (%)	97.4%
Allowed region (%)	2.6%
Additionally allowed region (%)	0.0%
Disallowed region (%)	0.0%
Global quality scores (raw/Z score)	
Verify3D	0.26 /-3.21

PROCHECK (φ-ψ)	0.07 /0.59
PROCHECK (all)	-0.18/-1.06
MolProbity clash score	37.07/-4.84
Model contents	
Ordered residue ranges	3-34,38-40,43-72,87-126,134-139
Total no. of residues	163
BMRB accession number	26765
PDB ID code	5XYL

[a] region, Skp1: *3-8,11-17,18-33,52-65,87-93,97-110,113-127*

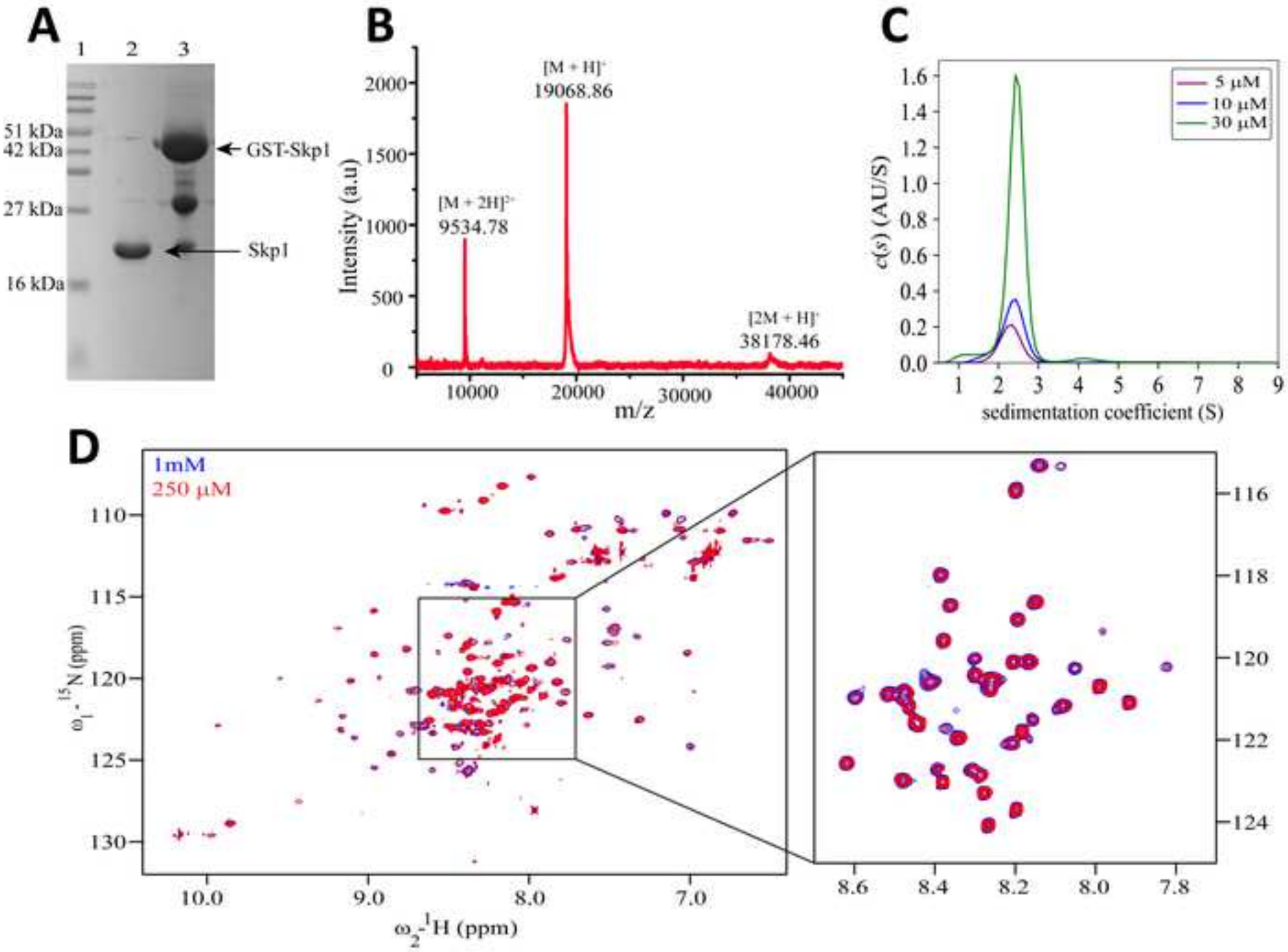
Table 2

RMSD comparison of the Skp1 solution structure ensemble (10 members) with the Skp1 from different complexes for the selected ordered region (helical core).

PDB ID	Information on type of Protein Complex	Method Used to solve Structure	RMSD Range (for comparison with 10 ensemble members)	Average RMSD Å	Error Å
6v88	Skp1A (truncated dimer)	Solution NMR	1.97-2.59	2.29	0.203
6o60	Skp1-FBXL2-GGTase	X-RAY 2.50 Å	2.55-4.06	3.09	0.569
7B5M	SCF-RBR E3-E3 super-assembly: CUL1-RBX1-**SKP1**-SKP2-CKSHS1-p27~Ub~ARIH1. Transition State 2	Electron Microscopy 3.91 Å	2.52-4.39	3.11	0.68

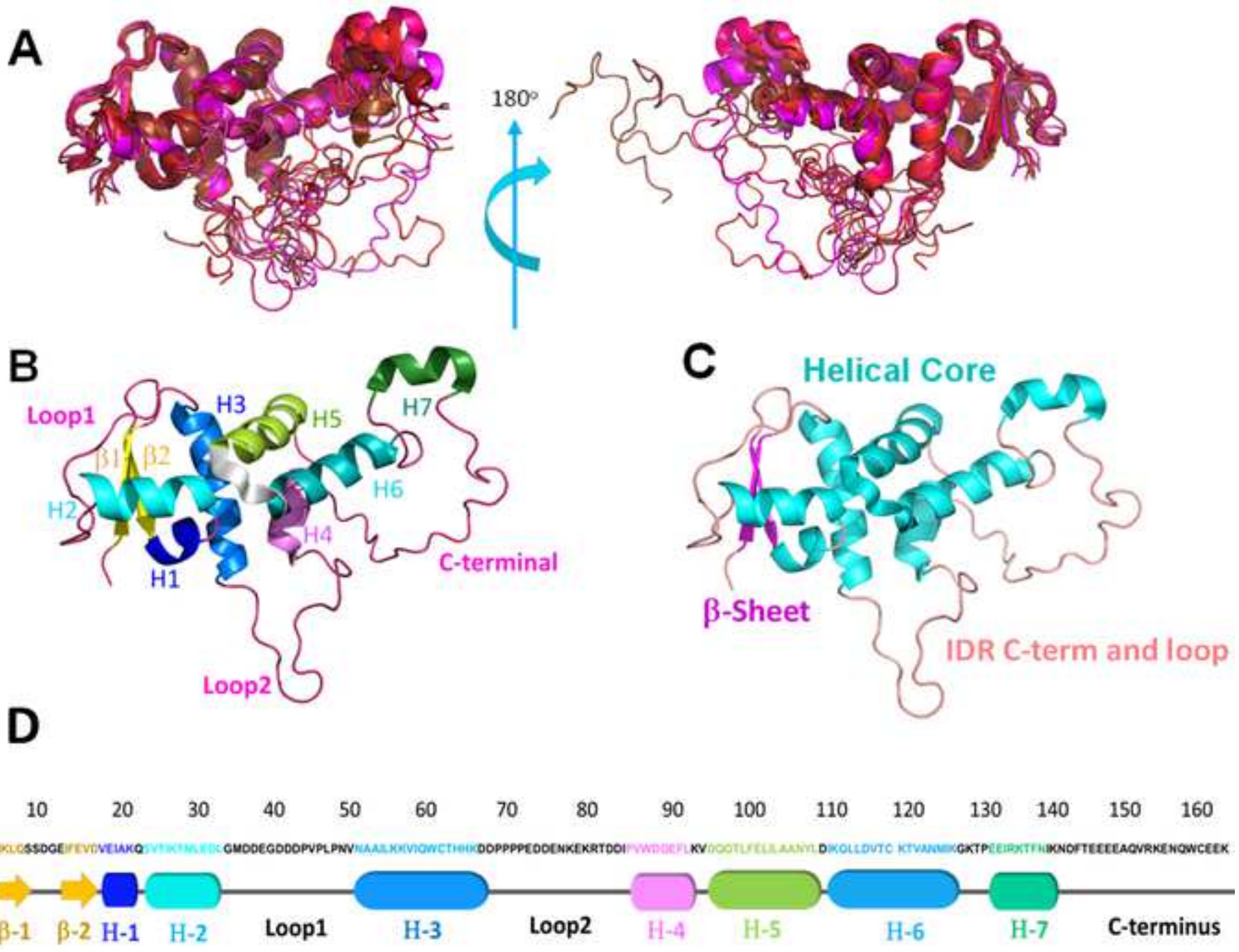
Figure

Click here to access/download;Figure;Fig 1.tif



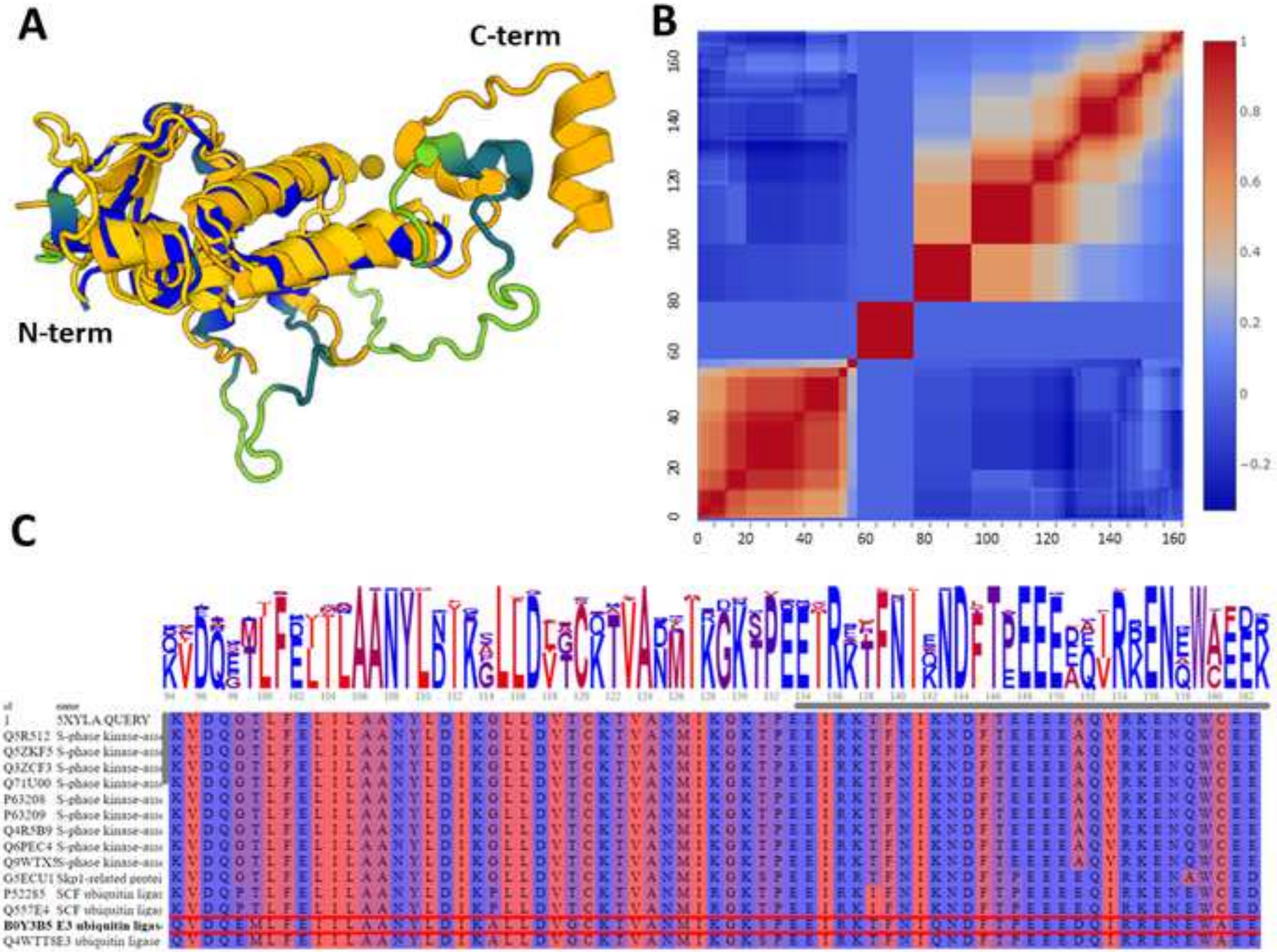
Figure

Click here to access/download;Figure;Fig 2.tif



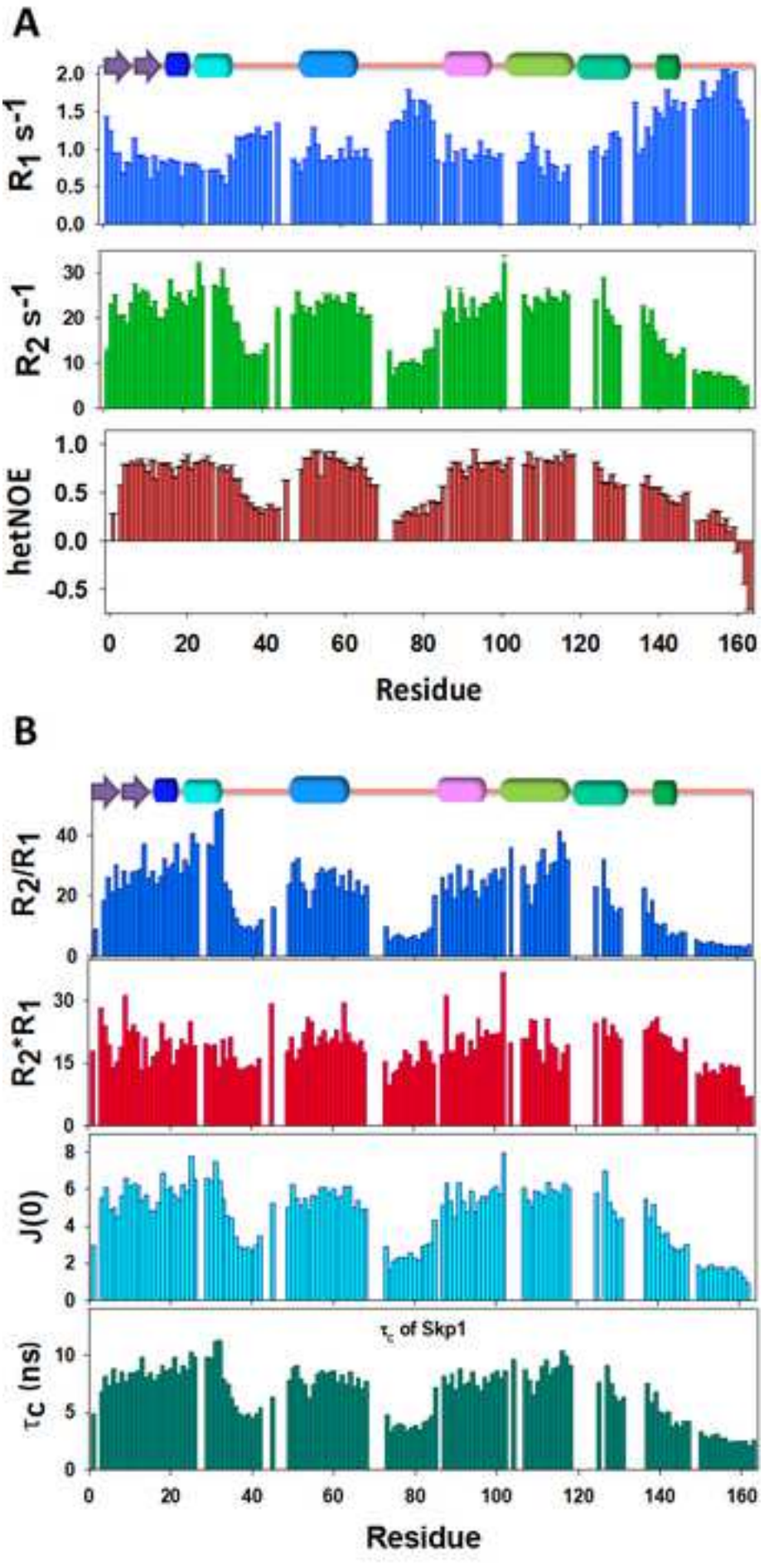
Figure

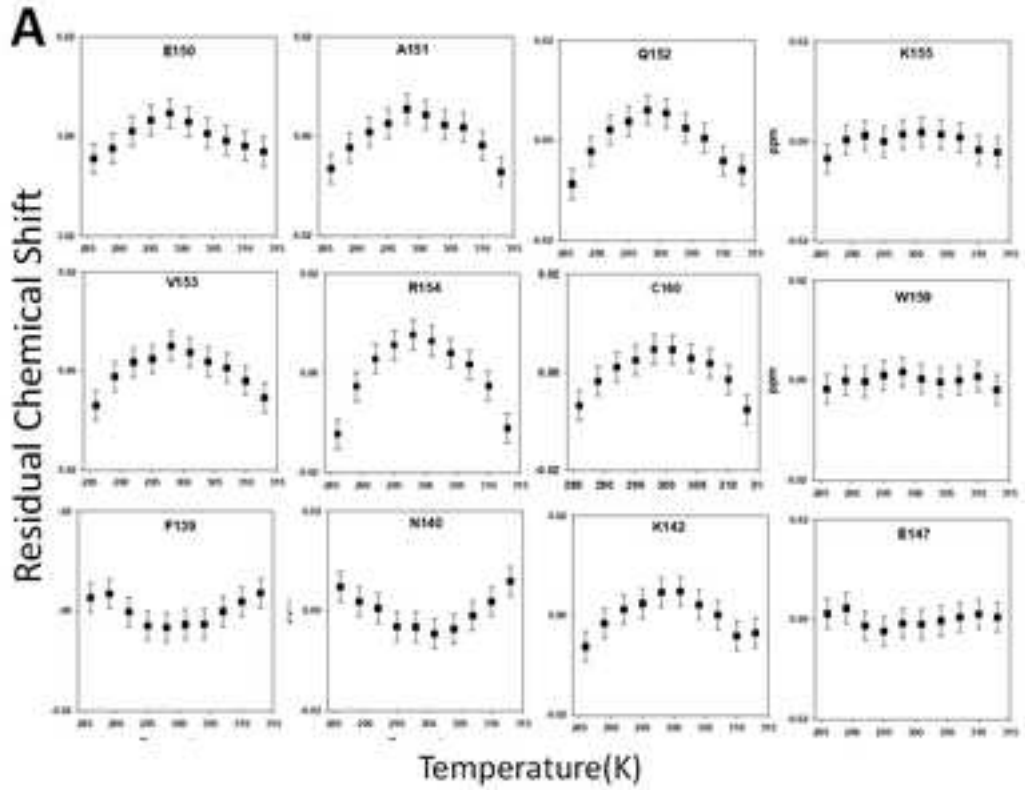
Click here to access/download;Figure;Fig 3.tif

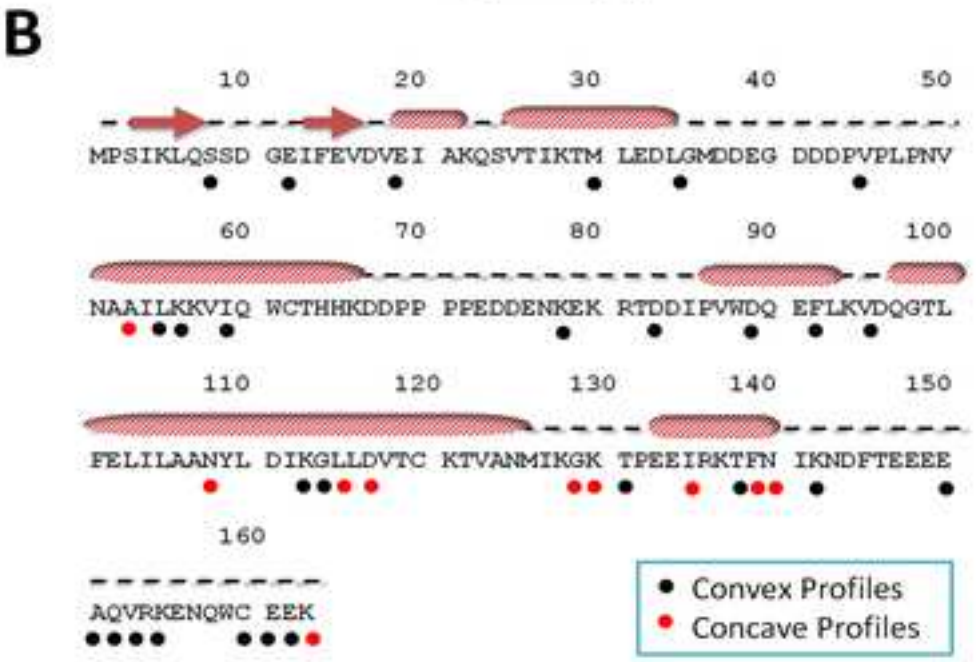


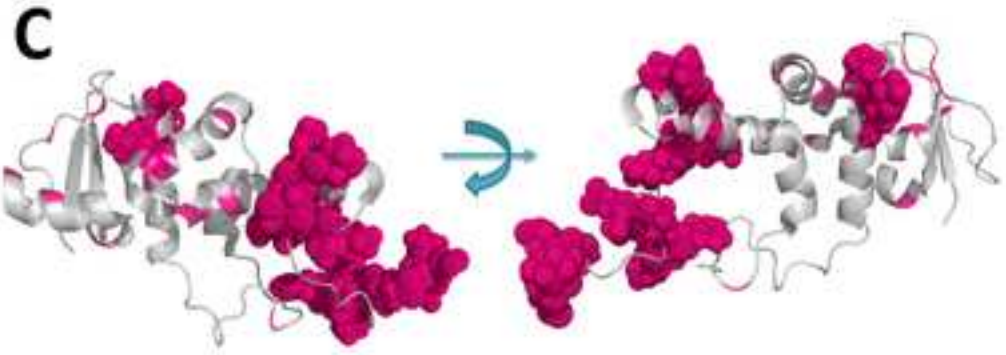
Figure

Click here to access/download;Figure;Fig 4.tif



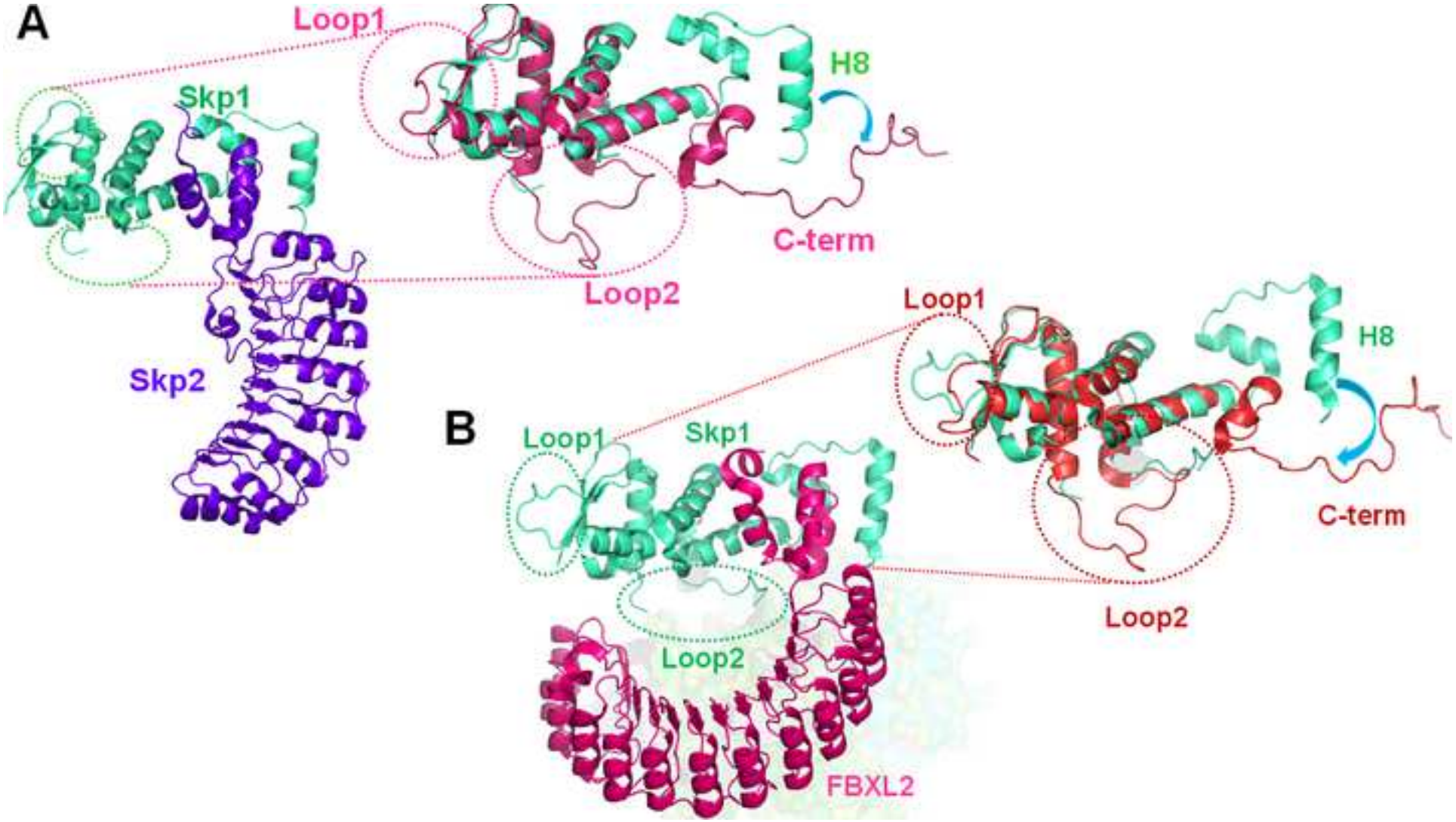






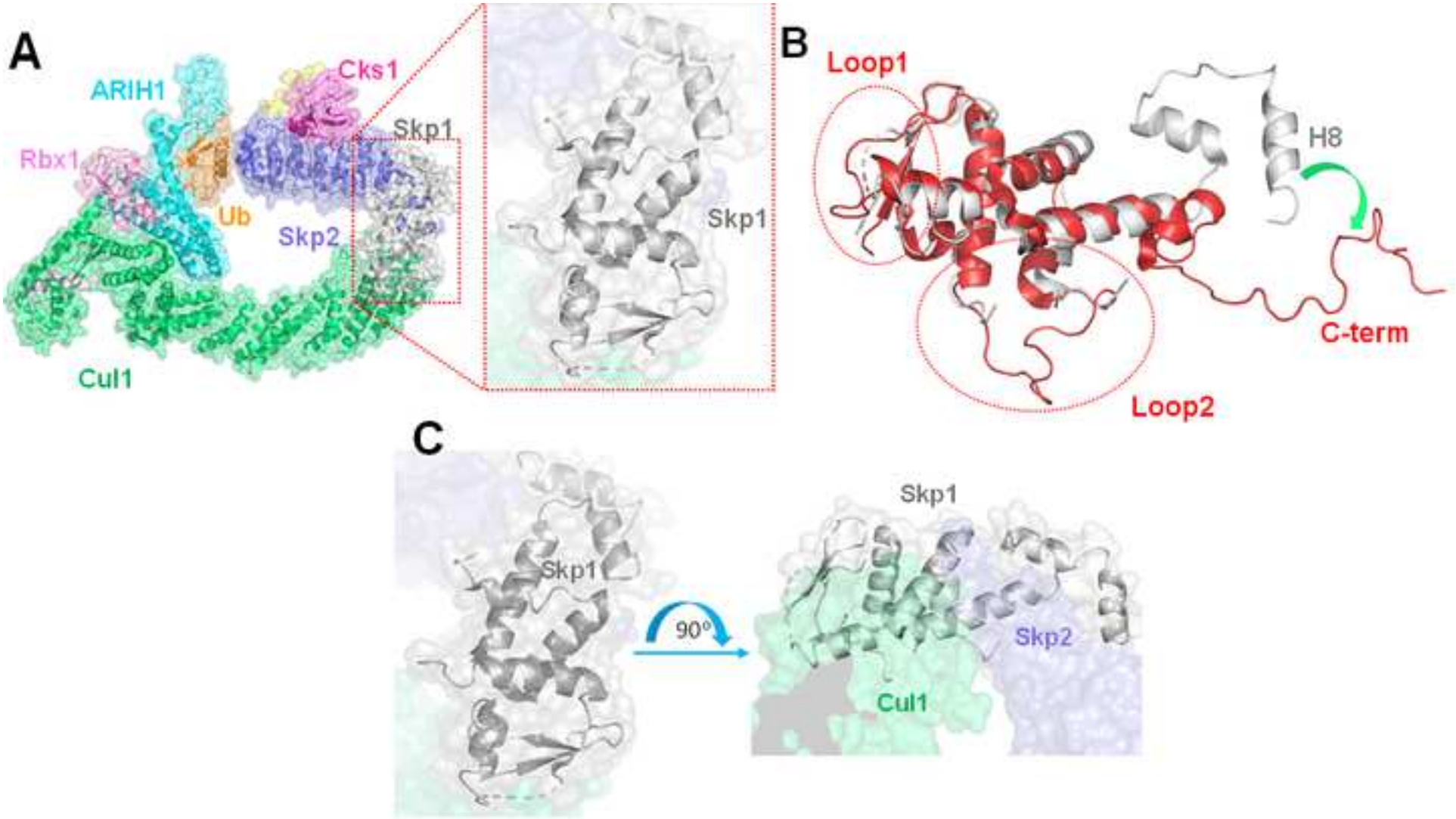
Figure

Click here to access/download;Figure;Fig 6.png



Figure

Click here to access/download;Figure;Fig 7.png



Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

NA

Author Credit Statement

Amrita Bhattacharya.: Sample preparation, Data curation, Methodology, Conceptualization, Writing- Original draft preparation, Writing revised Draft, Preparation of figures and tables, Visualization, Investigation.

Ashutosh Kumar*:* Supervision, Conceptualization, Reviewing, Investigation

Vaibhav Shukla: Methodology, Data Curation for Structure, Software, Validation.

Nitin Kachariya: Sample preparation, Data acquisition

Preeti: Sample preparation, Data acquisition, figure for review

Praveen Sehrawat: Data curation and validation for AUC