

# Phosphorylation Drives a Dynamic Switch in Serine/Arginine-Rich Proteins

ShengQi Xiang,<sup>1,9</sup> Vytautas Gapsys,<sup>4,9</sup> Hai-Young Kim,<sup>1</sup> Sergey Bessonov,<sup>5</sup> He-Hsuan Hsiao,<sup>6</sup> Sina Möhlmann,<sup>8</sup> Volker Klaukien,<sup>1</sup> Ralf Ficner,<sup>8</sup> Stefan Becker,<sup>1</sup> Henning Urlaub,<sup>6,7</sup> Reinhard Lührmann,<sup>5</sup> Bert de Groot,<sup>4</sup> and Markus Zweckstetter<sup>1,2,3,\*</sup>

<sup>1</sup>Department of NMR-based Structural Biology, Max Planck Institute for Biophysical Chemistry, Göttingen 37077, Germany

<sup>2</sup>German Center for Neurodegenerative Diseases (DZNE), Göttingen 37073, Germany

<sup>3</sup>Center for Nanoscale Microscopy and Molecular Physiology of the Brain, University Medicine Göttingen, 37073 Goettingen, Germany

<sup>4</sup>Computational Biomolecular Dynamics Group, Max Planck Institute for Biophysical Chemistry, Göttingen 37077, Germany

<sup>5</sup>Department of Cellular Biochemistry, Max Planck Institute for Biophysical Chemistry, Göttingen 37077, Germany

<sup>6</sup>Bioanalytical Mass Spectrometry, Max Planck Institute for Biophysical Chemistry, Göttingen 37077, Germany

<sup>7</sup>Bioanalytics, Department of Clinical Chemistry, University Medical Center, Göttingen 37073, Germany

<sup>8</sup>Abteilung für Molekulare Strukturbiologie, Institut für Mikrobiologie und Genetik, Göttinger Zentrum für Molekulare Biowissenschaften, Georg-August-Universität Göttingen, 37077 Göttingen, Germany

<sup>9</sup>These authors contributed equally to this work

\*Correspondence: [markus.zweckstetter@dzne.de](mailto:markus.zweckstetter@dzne.de)

<http://dx.doi.org/10.1016/j.str.2013.09.014>

## SUMMARY

Serine/arginine-rich (SR) proteins are important players in RNA metabolism and are extensively phosphorylated at serine residues in RS repeats. Here, we show that phosphorylation switches the RS domain of the serine/arginine-rich splicing factor 1 from a fully disordered state to a partially rigidified arch-like structure. Nuclear magnetic resonance spectroscopy in combination with molecular dynamics simulations revealed that the conformational switch is restricted to RS repeats, critically depends on the phosphate charge state and strongly decreases the conformational entropy of RS domains. The dynamic switch also occurs in the 100 kDa SR-related protein hPrp28, for which phosphorylation at the RS repeat is required for spliceosome assembly. Thus, a phosphorylation-induced dynamic switch is common to the class of serine/arginine-rich proteins and provides a molecular basis for the functional redundancy of serine/arginine-rich proteins and the profound influence of RS domain phosphorylation on protein-protein and protein-RNA interactions.

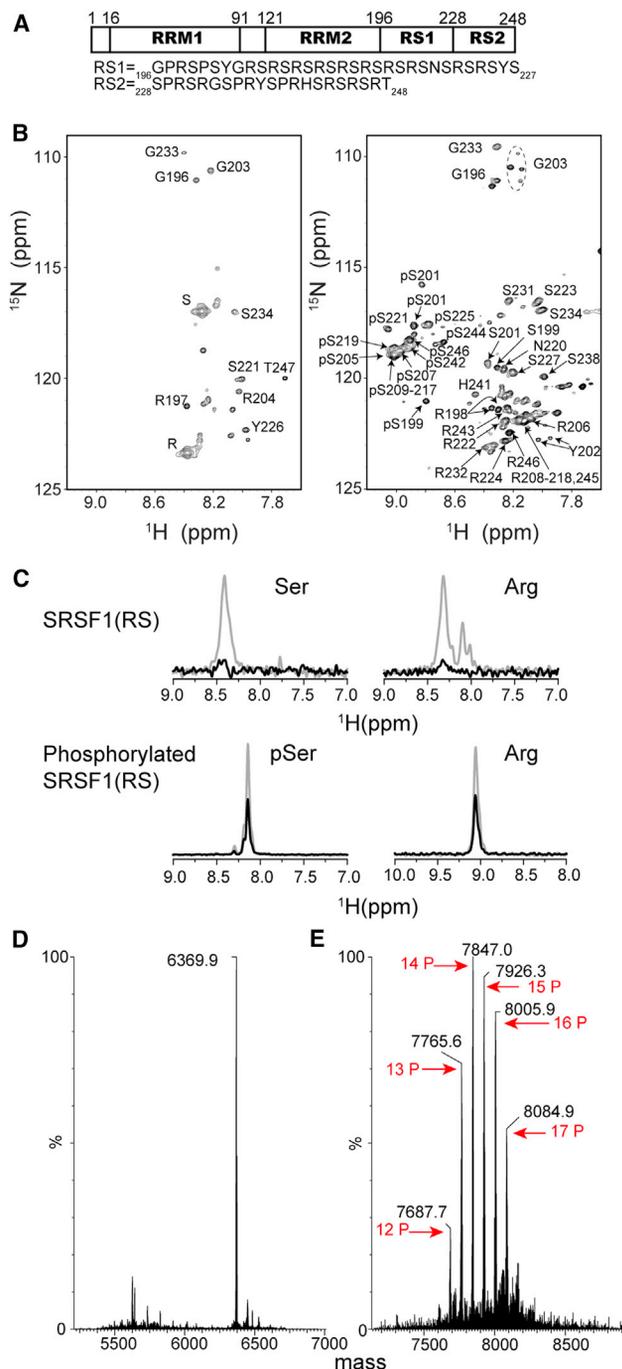
## INTRODUCTION

Serine/arginine-rich (SR) proteins play key roles in RNA metabolism from transcription, RNA splicing, RNA export, and translation to nonsense decay of RNA (Lin and Fu, 2007; Zhong et al., 2009). SR proteins are evolutionary conserved and contain an essential arginine/serine-rich (RS) domain as well as one or two RNA recognition motif (RRM) domains. RRM domains are involved in protein-RNA interactions, whereas RS domains mediate both protein-protein and protein-RNA interactions (Kohtz et al., 1994; Shen and Green, 2004; Valcárcel and Green, 1996;

Wu and Maniatis, 1993). RS and RRM domains are functionally independent in splicing (Cazalla et al., 2002). Serine residues within the RS domains of SR proteins are extensively phosphorylated (Gui et al., 1994). Phosphorylation influences the interactions and the subcellular localization of SR proteins and thus modulates their function. Both hyperphosphorylation and hypophosphorylation of SR proteins are known to inhibit splicing (Kanopka et al., 1998; Prasad et al., 1999; Sanford and Bruzik, 1999). In addition, SR proteins undergo several cycles of phosphorylation and dephosphorylation during splicing. However, little is currently known about the molecular basis of these processes and the phosphorylation-dependent structure of RS domains.

The serine/arginine-rich splicing factor 1 (SRSF1) (also known as ASF/SF2) is a prototypical SR protein. SRSF1 is essential for constitutive splicing and modulates splice site selection. In addition, SRSF1 was implicated in facilitating transcription, maintenance of genome stability, and organization of gene networks (Li and Manley, 2005; Zhong et al., 2009). Deletion of SRSF1 in a chicken cell line led to accumulation of incompletely processed-mRNA and subsequent cell death (Wang et al., 1996). SRSF1 null mice die early due to a defect in the postnatal splicing switch of a calmodulin-dependent kinase transcript (Xu et al., 2005). SRSF1 was also found to be upregulated in several human cancers and HIV-infected cells (Dowling et al., 2008), and the SFRS1 gene has been identified as proto-oncogene (Karni et al., 2007).

SRSF1 contains a RS domain at its C terminus that comprises a repeat of eight consecutive RS dipeptides (Figure 1A). Phosphorylation of the N-terminal half of the RS domain, named RS1, regulates the activity of SRSF1 in constitutive and alternative splicing and modulates the cellular distribution of SRSF1 (Cazalla et al., 2002). Moreover, nuclear entry and speckle formation of SRSF1 requires phosphorylation of RS1 by the serine kinase SRPK1 (Ngo et al., 2005). In line with the functional importance of the RS repeat stretch, a minimal RS domain consisting of only ten consecutive RS dipeptides can replace full-length SRSF1 in splicing of pre-mRNA and nucleocytoplasmic shuttling (Cazalla et al., 2002).



**Figure 1. A Dynamic Switch in the RS Domain of SRSF1**

(A) Domain organization of SRSF1 (RRM, RNA recognition motif; RS, arginine/serine-rich). The primary sequence of the N-terminal (RS1) and C-terminal half (RS2) of the RS domain of SRSF1 are shown.

(B)  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectrum of unphosphorylated (left panel) and phosphorylated SRSF1(RS) (right panel). Resonance assignments are indicated. Note the heterogeneity of phosphorylation, for example for S201 and S199.

(C) Selected traces from  $^{15}\text{N}$ - $^1\text{H}$  hetNOE spectra (reference: gray; saturated spectrum: black) recorded at 298 K for unphosphorylated and phosphorylated SRSF1(RS).

(D) Mass spectrometry of unphosphorylated SRSF1(RS).

(E) Mass spectrometry of phosphorylated SRSF1(RS). Different phosphorylated species are labeled in red.

Genome-wide studies have identified several proteins that contain RS domains but have other RNA-binding domains (Long and Caceres, 2009). In these so-called SR-related proteins, a large number of arginine/aspartic acid and arginine/glutamic acid dipeptide repeats are present (Blencowe et al., 1999; Cazalla et al., 2002). Human Prp28 (hPrp28) is an SR-related protein of the family of DEAD-box RNA helicases and is a component of the U5 small nuclear ribonucleoprotein (snRNP) complex (Teigelkamp et al., 1997). Prp28, (also known as DDX23) is required for the dissociation of U1 snRNP from the 5' splice site before catalytic activation of the spliceosome (Staley and Guthrie, 1999). Phosphorylation of the RS domain of hPrp28 is a key step in assembly of U4/U6.U5 snRNP into the spliceosomal machinery (Mathew et al., 2008).

Circular dichroism spectroscopy suggested that the RS domain of SRSF1 is disordered (Labourier et al., 1998). In contrast, molecular dynamics simulations have reported conflicting results for the structural properties of RS dipeptide repeats (Hamelberg et al., 2007; Sellis et al., 2012). For example, a helical conformation was proposed for the unphosphorylated state that is changed into a compact conformation, named arginine-claw, where the phosphate group of one phosphoserine residue is coordinated by at least four arginine side chains (Hamelberg et al., 2007; Sellis et al., 2012). In an X-ray structure of a SRPK1-SRSF1 complex, the N-terminal half of the unphosphorylated RS1 region docked in an extended conformation into a binding groove on SRPK1, whereas the C-terminal half of RS1 remained disordered (Ngo et al., 2008). Currently it is unclear if the proposed conformations and conformational changes occur in unbound RS domains, how phosphorylation changes RS repeats, if similar structural and dynamic properties are found in full-length SR proteins, and what molecular conformations RS domains in SR-related proteins populate.

Using a combination of nuclear magnetic resonance (NMR) spectroscopy and molecular dynamics (MD) simulation, we reveal a phosphorylation-induced conformational switch of the RS dipeptide repeat sequence from a highly disordered state to a partially rigidified, arch-like structure. In addition, we studied phosphorylation of the RS domain of hPrp28 in native U5 snRNP and B complex spliceosomes. We demonstrate that similar phosphorylation-induced changes in dynamics and structure occur in the canonical SR protein SRSF1 and 100 kDa hPrp28. Thus, our study establishes a phosphorylation-specific entropic switch that is common to SR and SR-related proteins and provides a molecular basis for the functional regulation of SR proteins by phosphorylation.

## RESULTS

### A Dynamic Switch in the RS Domain of SRSF1 upon Phosphorylation of Serine Residues

To obtain insight into the structure and dynamics of the RS domain of SRSF1, we recombinantly produced residues 196–248 of SRSF1 [SRSF1(RS)] (Figure 1A). The two-dimensional  $^{15}\text{N}$ - $^1\text{H}$  heteronuclear single quantum coherence (HSQC) spectrum of SRSF1(RS) had a narrow chemical shift dispersion pointing to extensive structural disorder (Figure 1B). Although SRSF1(RS) contains 53 residues, we observed only 33 cross-peaks. Out of these 33 cross-peaks, two had a more than

4-fold higher intensity. Resonance assignment attributes the two cross-peaks to arginine and serine. The high signal intensity suggests that multiple arginine residues—and respectively serine residues—contribute to a single cross-peak and thus have an identical chemical environment. The disordered nature of SRSF1(RS) was further supported by NMR spin relaxation measurements. Steady-state  $^{15}\text{N}$ - $^1\text{H}$  heteronuclear nuclear Overhauser effect (NOE) (hetNOE) values, which probe backbone motions in the pico-to-nanosecond timescale, were below 0.1 (Figure 1C). In addition, large NMR line-widths (Figure 1B) indicate the presence of conformational exchange occurring on the micro-to-millisecond timescale. Taken together, the NMR data demonstrate that SRSF1(RS) is disordered and highly dynamic.

Next, we phosphorylated SRSF1(RS) by the serine kinase SRPK1. Mass spectrometry showed that 12 to 17 phosphate groups were attached to SRSF1(RS) (named SRSF1(RpS) further on) (Figures 1D and 1E). Phosphorylation strongly changed the  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectrum (Figure 1B). The number of NMR signals went up to 52 and chemical shift dispersion was increased. In addition,  $\sim 12$  signals were observed in the amide proton region from 8.6 to 9.2 ppm, in line with the number of phosphate groups detected by mass spectrometry. Despite severe sequence degeneracy in SRSF1(RpS), we were able to sequence specifically assign 43 residues (Figure 1B). The assignment was only possible by using five- and six-dimensional NMR spectra. Sequence-specific resonance assignment unambiguously identified phosphorylated serines in the RS repeat stretches comprising residues 204–219 and 242–247.

To probe the structure and dynamics of SRSF1(RpS), we measured  $^{15}\text{N}$ - $^1\text{H}$  hetNOE and three-bond scalar HN-H $\alpha$  couplings ( $^3J_{\text{HN-H}\alpha}$ ).  $^3J_{\text{HN-H}\alpha}$  reports on the backbone torsion angle  $\phi$  and is therefore a sensitive probe for changes in backbone conformation.  $^3J_{\text{HN-H}\alpha}$  couplings of the phosphorylated serines S209–S217 were on average  $5.2 \pm 0.3$  Hz, values lower than expected for a disordered peptide chain. Similarly, arginine residues in the RS repeats showed decreased  $^3J_{\text{HN-H}\alpha}$  couplings with an average of  $6.3 \pm 0.3$  Hz. The decrease in  $^3J_{\text{HN-H}\alpha}$  suggests that phosphorylation of SRSF1(RS) promotes formation of turn-like structure. Further support for a phosphorylation-induced structural stabilization was provided by  $^{15}\text{N}$ - $^1\text{H}$  hetNOE values that jumped from below 0.1 in SRSF1(RS) to  $0.6 \pm 0.1$  for the phosphorylated serines in SRSF1(RpS) (Figure 1C). Arginine residues in the RS stretch had  $^{15}\text{N}$ - $^1\text{H}$  hetNOE values of  $\sim 0.5 \pm 0.1$ . Taken together, NMR spectroscopy revealed that phosphorylation of the RS dipeptides in SRSF1(RS) causes a switch from a disordered conformation to a partially rigidified structure.

### Phosphorylation Stabilizes the Structure of RS Repeats

Phosphorylation of SRSF1(RS) resulted in a heterogeneous population of phosphorylated species (Figures 1D and 1E) that precluded structural analysis. Taking into account that a stretch of ten consecutive RS dipeptides can partially restore the function of SRSF1 in mRNA splicing (Cazalla et al., 2002), we decided to recombinantly produce residues 200–219 of SRSF1. The region corresponds to SRSF1(RS1) and contains a sequence of eight RS dipeptides (Figure 1A). Phosphorylation of SRSF1(RS1) by SPRK1 allowed preparation of a sample in which six serine residues carried a phosphate group, whereas no other residue was phosphorylated (Figures 2A–2D). Superposition showed

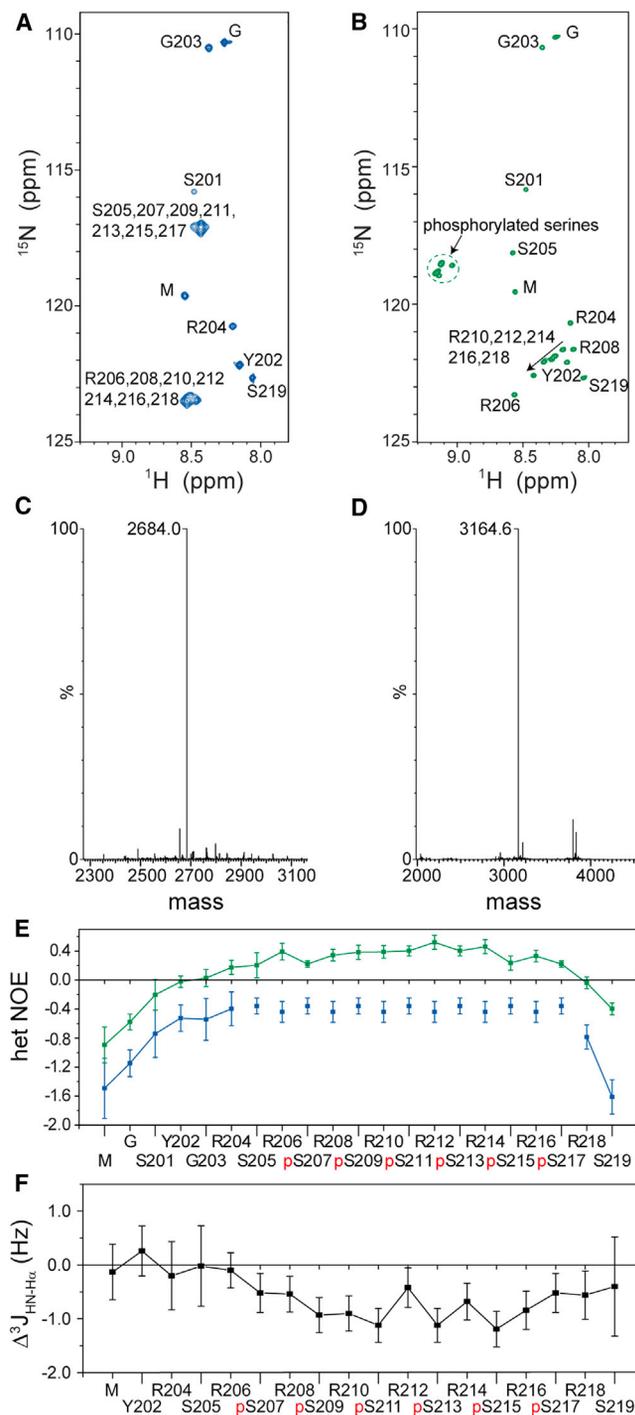
that NMR spectra of unphosphorylated and phosphorylated SRSF1(RS1) matched to the corresponding spectra of SRSF1(RS) (Figures 1, 2, and Figure S1 available online). Thus, SRSF1(RS1) provides a viable model to study phosphorylation-induced structural changes in RS domains. NMR resonance assignment revealed that serines from S207–S217 were phosphorylated to 100% (Figures 2B and 2D). The last and first serine residue of the repeat was not modified. Residues preceding the RS dipeptide stretch did not change upon phosphorylation, indicating that structural changes were restricted to the RS repeat region. In addition, circular dichroism and diffusion NMR spectroscopy demonstrated that SRSF1(RS1) is predominantly monomeric in both the unphosphorylated and phosphorylated state (Figures 2 and S2).

The pK value for the equilibrium between the  $-1$ -charged and  $-2$ -charged phosphate group is  $\sim 6$ . In agreement with the influence of pH on the charge state of the phosphoserine side chain, NMR signals in NMR spectra of SRSF1(RpS1) shifted toward more random coil-like positions with decreasing pH (Figure 3). At pH 4.0, cross-peaks from the RS repeat highly overlapped (Figure 3A), reminiscent of the properties of unphosphorylated SRSF1(RS1). The disordered nature of SRSF1(RpS1) at pH 4.0 was further supported by a  $^{13}\text{C}$ - $^1\text{H}$  HSQC spectrum, in which C $\alpha$  and C $\beta$  chemical shifts of arginine residues were random coil-like (Figures 3B and 3C). In contrast, when increasing the pH from 7.5 to 8.0 no difference in NMR spectra was observed (Figure 3d).

To determine the structure of phosphorylated SRSF1(RS1), we measured a large number of structure-sensitive NMR parameters (Figures 2E, 2F, and 4; Table S1): nitrogen and carbon chemical shifts, one-bond C $\alpha$ -H $\alpha$  ( $^1J_{\text{C}\alpha\text{-H}\alpha}$ ), C $\alpha$ -C $\beta$  ( $^1J_{\text{C}\alpha\text{-C}\beta}$ ), and three-bond ( $^3J_{\text{HN-H}\alpha}$ ) scalar couplings, as well as N-H, C $\alpha$ -H $\alpha$ , C $\alpha$ -CO residual dipolar couplings. NOE contacts between side chains were not accessible due to severe resonance degeneracy in  $^{13}\text{C}$ -edited NOE spectra. In  $^{15}\text{N}$ -edited NOE spectra no medium- or long-range contacts were resolved. To define the conformation of the side chains of SRSF1(RpS1), we determined  $^3J_{\text{N-C}\gamma}$  and  $^3J_{\text{CO-C}\gamma}$  scalar couplings. For all NMR parameters deviations from random coil values and from those in the unphosphorylated peptide were observed for the six phosphorylated RS dipeptides (Figures 2E, 2F, and 4). We note that  $^3J_{\text{HN-H}\alpha}$  of the phosphoserines of SRSF1(RpS1) were  $\sim 5.5$  Hz, in line with the couplings observed in phosphorylated SRSF1(RS). The NMR measurements provided a large number of structural parameters for a homogeneously phosphorylated RS repeat sequence and thus formed the basis to reveal phosphorylation-induced conformational changes in SR proteins.

### Ensemble Description of SRSF1(RS1) by Combination of NMR Spectroscopy with MD Simulations

Consistent with the NMR data of the RS domain of SRSF1 (Figure 1C), hetNOE values in SRSF1(RpS1) did not exceed 0.5 (Figure 2E). Thus, despite the phosphorylation-induced conformational switch, the structure of SRSF1(RpS1) retained a large amount of pico-to-nanosecond backbone dynamics and is best described by an ensemble of conformations. In order to determine the structural ensemble of SRSF1(RpS1), NMR measurements were integrated with MD simulations. Unbiased MD simulations were carried out for SRSF1(RpS1) independently



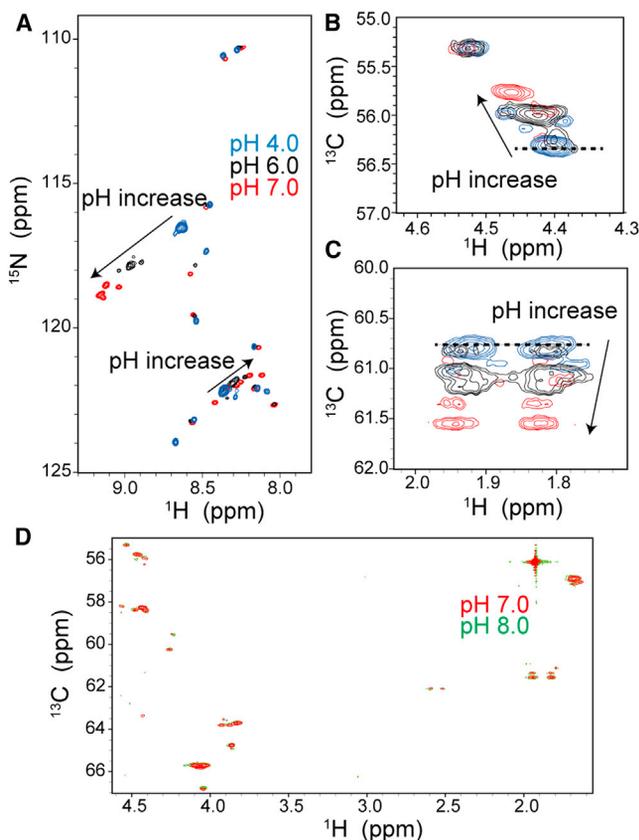
**Figure 2. Phosphorylation Changes the Structure of RS Repeats**

(A and B)  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectrum of the RS1 region of SRSF1 before (A) and after phosphorylation by SRPK1 (B). Six serine residues carried a phosphate group, whereas no other residue was phosphorylated. Resonance assignments are indicated.

(C) Mass spectrometry of unphosphorylated SRSF1(RS1).

(D) Mass spectrometry of phosphorylated SRSF1(RS1). The sample was phosphorylated at six serine residues.

(E)  $^{15}\text{N}$ - $^1\text{H}$  hetNOEs as a function of residue number before (blue) and after phosphorylation (green) of SRSF1(RS1). Values for S207–S217 in unphosphorylated SRSF1(RS1) are identical and correspond to the average



**Figure 3. The Charge State of the Phosphate Group Regulates the Conformational Change**

(A)  $^{15}\text{N}$ - $^1\text{H}$  HSQC of phosphorylated SRSF1(RS1) at pH 4.0 (blue), 6.0 (black), and 7.0 (red).

(B)  $\text{C}\alpha$ - $\text{H}\alpha$  cross-peaks of arginine residues in  $^{13}\text{C}$ - $^1\text{H}$  HSQC spectra. Color coding as in (A).

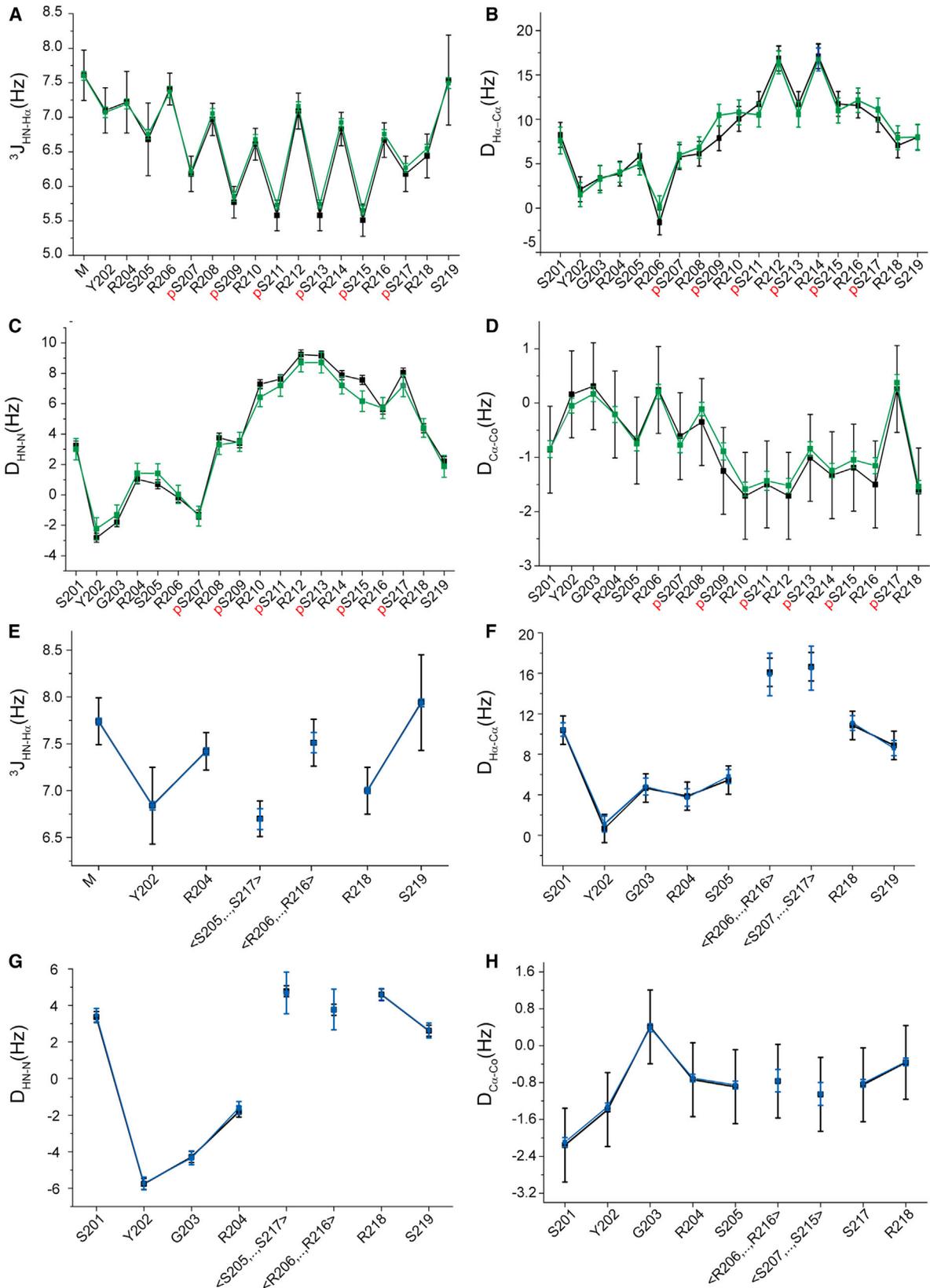
(C)  $\text{C}\beta$ - $\text{H}\beta$  region of arginine. Dashed lines in (B) and (C) mark the random coil values of  $\text{C}\alpha$  and  $\text{C}\beta$ . The random coil-like chemical shifts at pH 4.0 indicate that phosphorylated SRSF1(RS1) is disordered at pH 4.0.

(D) Superposition of  $^{13}\text{C}$ - $^1\text{H}$  HSQC spectra of phosphorylated SRSF1(RS1) at pH 7.0 (red) and pH 8.0 (green).

for the two protonation states of the phosphoserine residues ( $\text{PO}_4^-$  and  $\text{PO}_4^{2-}$ ; simulation details can be found in the [Experimental Procedures](#)). For validation, MD simulations were also performed for the unphosphorylated peptide, which NMR spectroscopy showed to be fully disordered ([Figure 2](#)). To create a representative structural ensemble that is consistent with the experimental NMR data, a subensemble selection method was developed. The procedure combined a Monte-Carlo search with exhaustive scanning (for further details see the [Experimental Procedures](#)). Subensembles containing up to 100

value observed for the strong serine (and arginine, respectively) peak seen in the HSQC spectrum shown in the left panel of A). The N-terminal Met and Gly originate from cloning.

(F) Difference of  $^3J_{\text{HN-H}\alpha}$  couplings between phosphorylated and unphosphorylated SRSF1(RS1). For unphosphorylated S207–S217 one value for all serine and one value for all arginine residues was used. Phosphorylated serines are labeled. Error bars in (E) and (F) represent SD as estimated on the basis of NMR signal intensities.



(legend on next page)

conformers were selected, which were in best agreement with experimental N-H, C $\alpha$ -H $\alpha$ , C $\alpha$ -CO residual dipolar couplings and  $^3J_{\text{HN-H}\alpha}$  scalar couplings (Figures 5A and 5B). For the phosphorylated form, both protonation states were about equally populated in the selected ensembles. Notably, no single conformer could reproduce the experimentally measured couplings (Figures 5A and 5B), in line with the dynamic nature of SRSF1(RpS1). Increasing the ensemble size enabled the structural model to fulfill NMR constraints more closely, as indicated by the decrease in quality factors for residual dipolar couplings and  $^3J_{\text{HN-H}\alpha}$  scalar couplings (Figures 5A and 5B). NMR quality factors saturated for an ensemble size of 30, suggesting that 30 conformations are optimal to avoid both underfitting and overfitting of the experimental data. One hundred independent selections were then performed for this ensemble size. Consistent with a successful selection procedure, the selected ensembles fulfilled the NMR parameters that were used for selection (Figures 4 and S3). Calculation of the Jensen-Shannon divergence between the smoothed projections of the ensembles (Figure 5C), as well as partial least squares analysis (Figure S4), demonstrated that the difference between SRSF1(RS1) and SRSF1(RpS1) was significantly larger than internal differences within each ensemble. The structural ensembles were then validated by back-calculation of a number of NMR observables ( $^1J_{\text{C}\alpha\text{-H}\alpha}$  and  $^1J_{\text{C}\alpha\text{-C}\beta}$  scalar couplings and chemical shifts) that were not used during ensemble generation (Figures 6 and S3). In addition, back-calculated  $^3J_{\text{N-C}\gamma}$  and  $^3J_{\text{CO-C}\gamma}$  couplings were in agreement with experimental values (Figures 6 and S3), although no NMR data, which directly restrict the side chain conformation, were used during ensemble selection. Together the data provide strong support for the conformer ensembles determined by the combined MD/NMR approach.

Although unphosphorylated SRSF1(RS1) is completely disordered (Figure 2), the combined NMR/MD analysis revealed a specific arch-like structure for SRSF1(RpS1) (Figures 7A and 7B). Clustering in internal distance space demonstrated that phosphorylation of RS repeats disfavored predominantly collapsed as well as fully stretched conformations (Figure 7B). Notably, none of the single clusters accurately reproduced the experimental NMR data (Q values for N-H, C $\alpha$ -H $\alpha$ , C $\alpha$ -CO residual dipolar couplings and average unsigned errors for  $^3J_{\text{HN-H}\alpha}$ ,  $^3J_{\text{N-C}\gamma}$  and  $^3J_{\text{CO-C}\gamma}$ ,  $^1J_{\text{C}\alpha\text{-C}\beta}$ , and  $^1J_{\text{C}\alpha\text{-H}\alpha}$ , respectively: 1.5/2.2/1.3/0.9/0.2/0.3/3.0/1.6 for cluster 1, 1.1/1.4/1.3/0.5/0.2/0.4/2.8/0.9 for cluster 2, 0.8/0.9/1.1/0.4/0.2/0.3/2.8/1.1 for cluster 3, 0.8/1.0/0.8/0.4/0.1/0.2/2.8/0.6 for cluster 4, 1.1/1.1/1.2/0.4/0.1/0.5/2.8/0.9 for cluster 5, and 2.1/2.1/3.2/1.2/0.3/0.9/3.1/1.3 for cluster 6), indicating that only the sum of the clusters faithfully characterizes SRSF1(RS1) in solution. In addition, we performed a selection procedure to extract ensembles from the unphosphorylated trajectory using the phosphorylated NMR data as a target, as well as a selection started from the combined pool of

structures (Figures S4A and S4B). The selection based on the inverted pools led to inferior cross-validation statistics, whereas the combined pool gave similar statistics as with the original pools. This suggests that the information of the “own” clusters aids the fulfillment of the NMR data.

Side chains of arginines and phosphoserines showed well-defined preferential orientations in SRSF1(RpS1), although they were smeared out in the reference ensemble of unphosphorylated SRSF1(RS1) (Figure 7C). In addition, analysis of arginine  $\chi_1$  angle distributions demonstrated that phosphorylation induced a population shift from the *gauche+* to *gauche-* side chain conformation (Figure 7D). We note that <1% of phosphoserines were involved in hydrogen bonds with three or more arginine residues (Figure 7; Table S2). Thus, we see no evidence for a previously proposed arginine claw (Hamelberg et al., 2007; Sellis et al., 2012) in the structure of phosphorylated RS repeats.

NMR spectroscopy demonstrated a dramatic phosphorylation-induced change in the flexibility of the RS and RS1 domains of SRSF1 (Figures 1C, 2E, and 2F). We therefore estimated the upper limit of conformational entropy for the determined ensembles using Schlitter's formula (Schlitter, 1993). To this end, conformations from 100 ensembles were superpositioned separately for SRSF1(RS1) and SRSF1(RpS1). Based on covariance matrices and the corresponding eigenvalue spectra, we then estimated the entropy for the backbone of the SRSF1(RpS1) ensemble as 2,049.6 J/(mol K) and for unphosphorylated SRSF1(RS1) as 2,178.4 J/(mol K). Upon inclusion of side chains (with phosphate groups discarded), values increased to 5,584.84 J/(mol K) and 6,490.91 J/(mol K) for SRSF1(RpS1) and SRSF1(RS1), respectively. Thus, phosphorylation strongly reduces the conformational entropy of the RS1 domain of SRSF1, in line with its more ordered structure.

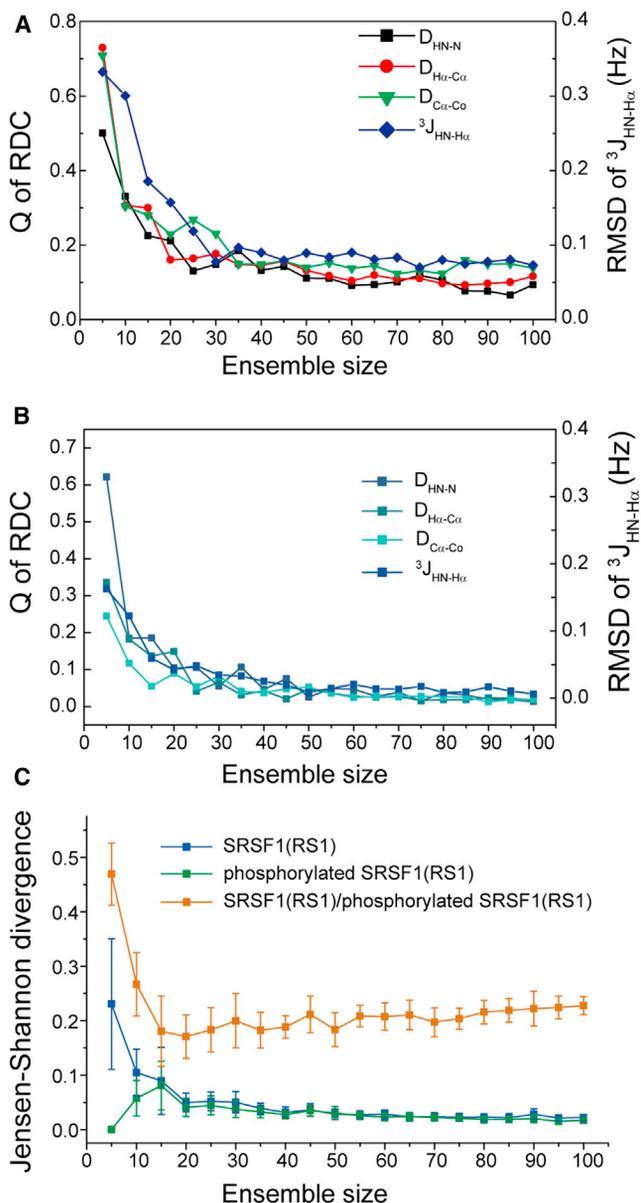
#### Dynamic Tuning in RS-Rich, RE-Rich, and RD-Rich Regions of the 100 kDa Protein hPrp28

hPrp28 is a 100 kDa spliceosomal SR-related protein, which contains an N-terminal domain rich in arginine and serines (Figure 8A). Phosphorylation of hPrp28 plays an important role for assembly of the U4/U6.U5 tri-snRNP and its functional integration into the spliceosomal B complex (Mathew et al., 2008). Sites at which hPrp28 is phosphorylated in native U5 snRNP and spliceosomal B complex, however, were unknown. Here, we analyzed the phosphorylation sites of hPrp28 in U5 snRNPs and assembled B complexes by mass spectrometry. The analysis identified serine residues at positions 14, 23, 38, 39, 41, 65, 67, 106, 107, 109, as well as T25, to be phosphorylated in both complexes (Figure 8B; Table S3). Phosphorylation at S16 and S63 was found only in U5 snRNP.

To investigate the structural consequences of phosphorylation of the RS domain of hPrp28, we performed in vitro phosphorylation of full-length hPrp28(1–820), as well as its RS-rich domains,

#### Figure 4. Validation of NMR-Driven Ensemble Selection of Phosphorylated and Unphosphorylated SRSF1(RS1)

Comparison of experimental NMR data (black), which were used for ensemble selection, with those back-calculated from the conformer ensembles of phosphorylated (A–D) and unphosphorylated (E–H) SRSF1(RS1). Experimental (black) and back-calculated values of  $^3J_{\text{HN-H}\alpha}$  scalar couplings (A and E), C $\alpha$ -H $\alpha$  RDCs (B and F), HN-N RDCs (C and G), and C $\alpha$ -CO RDCs (D and H) are shown. Errors of experimental data represent SD as estimated on the basis of NMR signal intensities and line widths. Errors of back-calculated couplings represent SD values between the mean values of 100 independently selected ensembles. Because every selection already reports on a mean value (average over 30 structures) the SD of the mean values should be called an SE (however, the explicit division by  $\sqrt{n}$  was not applied in any step).



**Figure 5. Ensemble Sizes and Interensemble Differences**

Influence of the ensemble size on the agreement between back-calculated and experimental NMR data of phosphorylated (A) and unphosphorylated SRSF1(RS1) (B) and on the difference between structural ensembles of unphosphorylated and phosphorylated SRSF1(RS1) (C). (A) Influence of ensemble size on the agreement between experimental H-N, C $\alpha$ -H $\alpha$ , C $\alpha$ -CO RDCs (Q values; dipolar quality factors) as well as experimental  $^3J_{\text{HN-H}\alpha}$  scalar couplings (root-mean-square deviation [rmsd]) and values back-calculated from selected conformer ensembles of phosphorylated SRSF1(RS1). H-N, C $\alpha$ -H $\alpha$ , C $\alpha$ -CO RDCs, and  $^3J_{\text{HN-H}\alpha}$  scalar couplings were used for ensemble selection. (B) Influence of ensemble size on the agreement between experimental data and back-calculated from selected conformer ensembles of unphosphorylated SRSF1(RS1). H-N, C $\alpha$ -H $\alpha$ , C $\alpha$ -CO RDCs and  $^3J_{\text{HN-H}\alpha}$  scalar couplings were used for ensemble selection. (C) Jensen-Shannon divergence for different ensemble sizes. Notably, interensemble differences are significantly larger than intra-ensemble JS divergence. Errors bars represent SD.

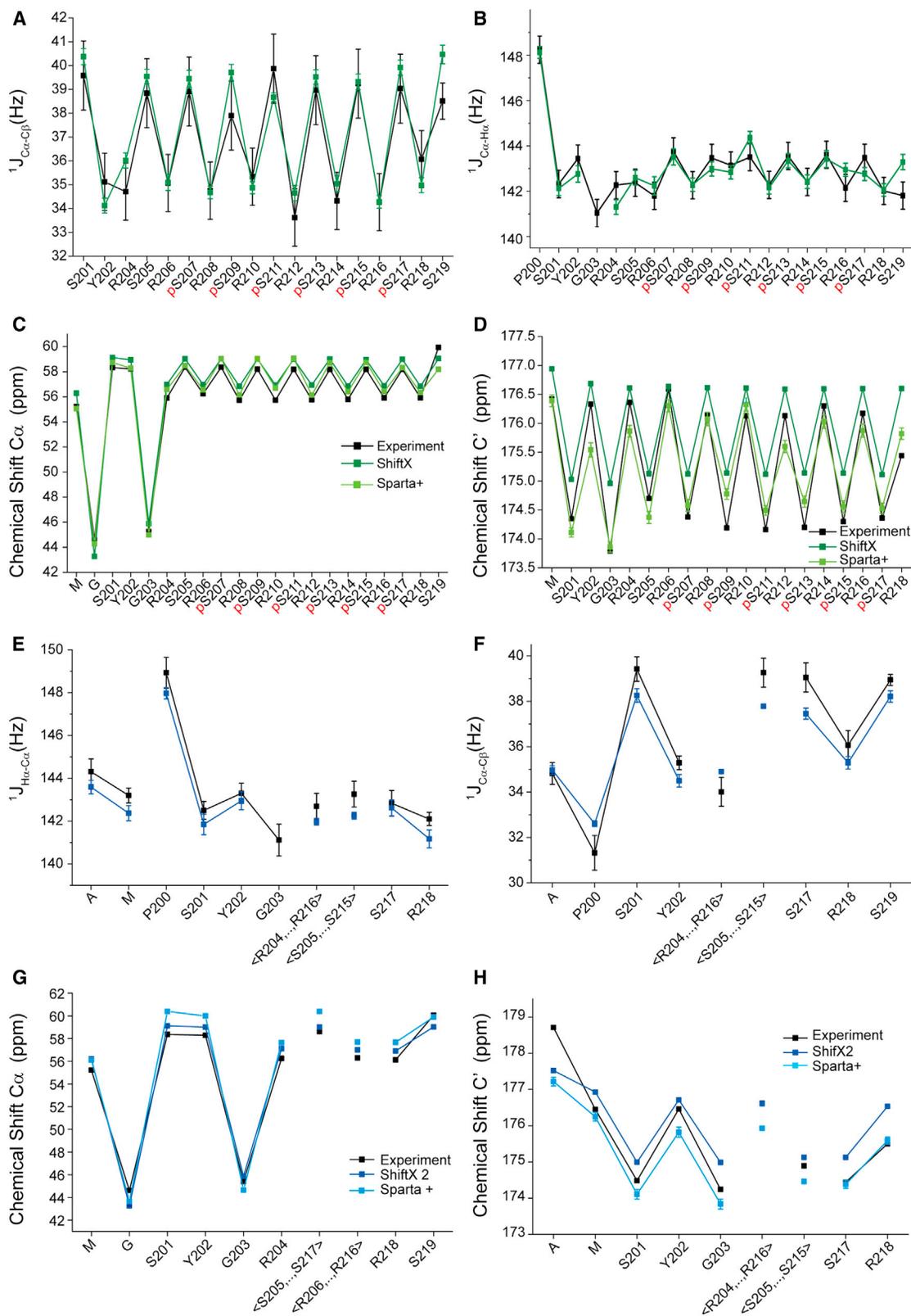
hPrp28(1–257) and hPrp28(1–138) (Figure 8A). Although full-length hPrp28 has an additional 563 residues, the number of NMR signals in a  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectrum was only slightly increased when compared to hPrp(1–257) (Figure 9A). Thus, the globular parts of hPrp28(1–820) tumble to slowly to be detected by solution-state NMR. Nevertheless, the close match of peak positions for the two proteins demonstrates that conformations of the isolated RS domain are retained in the intact protein. Sequence-specific resonance assignment of hPrp28(1–138) phosphorylated by SRPK1 revealed 13 phosphorylation sites including 12 serines and one threonine (Figure 9A). The in vitro phosphorylation pattern is largely consistent with that in native U5 snRNPs and spliceosomal B complexes (Figure 9B). Part of the observed differences might be due to the difficulty to detect highly phosphorylated RS repeats by mass spectrometry under the chosen conditions, i.e., using trypsin for the generation of peptides. Sequence-specific analysis of hetNOE and  $^3J_{\text{HN-H}\alpha}$  scalar couplings in hPrp28(1–138) showed that phosphorylation strongly reduced pico-to-nanosecond timescale motions and led to a preferential population of turn-like conformers (Figures 9B and 9C), in line with our observations for phosphorylated SRSF1(RS). Taken together the data demonstrate that the phosphorylation-induced entropic switch is common to SR proteins and SR-related proteins.

SR-related proteins also contain many arginine/glutamic acid (RE) or arginine/aspartic acid (RD) dipeptides. For example, in hPrp28 a 12 residue-stretch with RE and RD dipeptides is present (Figure 8B). Comparison of the  $^{15}\text{N}$ - $^1\text{H}$  HSQC of unphosphorylated hPrp28(1–820) with that of a synthesized RE repeat peptide—and further supported by carbon chemical shifts in triple-resonance NMR experiments—tentatively identified the NMR signals of RE and RD repeats (Figure 9D). In the two spectral regions, hetNOE values were  $\sim 0.3$ , suggesting that RE and RD repeats are not fully mobile. In contrast, most other signals had disappeared in the saturated spectrum of the hetNOE experiment (Figure 9D). Notably, a comparison of hetNOE values in RE and RD repeats with those in unphosphorylated and phosphorylated RS repeats indicates that pico-to-nanosecond dynamics of RE and RD repeats are in between those of unphosphorylated and phosphorylated RS repeats (Figure 9E).

## DISCUSSION

Although only 20 amino acids are directly encoded in the genetic code, there are various posttranslational modifications of amino acid side chains that are vital for protein function. Reversible phosphorylation is one of the most ubiquitous and significant regulatory mechanisms. Genome analysis revealed that phosphorylation preferentially occurs within intrinsically disordered regions (Iakoucheva et al., 2004). Phosphorylation can change protein conformations in different ways, from local effects such as alignment of side chains of key residues in reaction centers (Cheng et al., 2006) to induction of allosteric changes distant from the phosphate (Banavali and Roux, 2007). In some cases, dramatic conformational changes occur involving either order-to-disorder or disorder-to-order transitions (Espinoza-Fonseca et al., 2007).

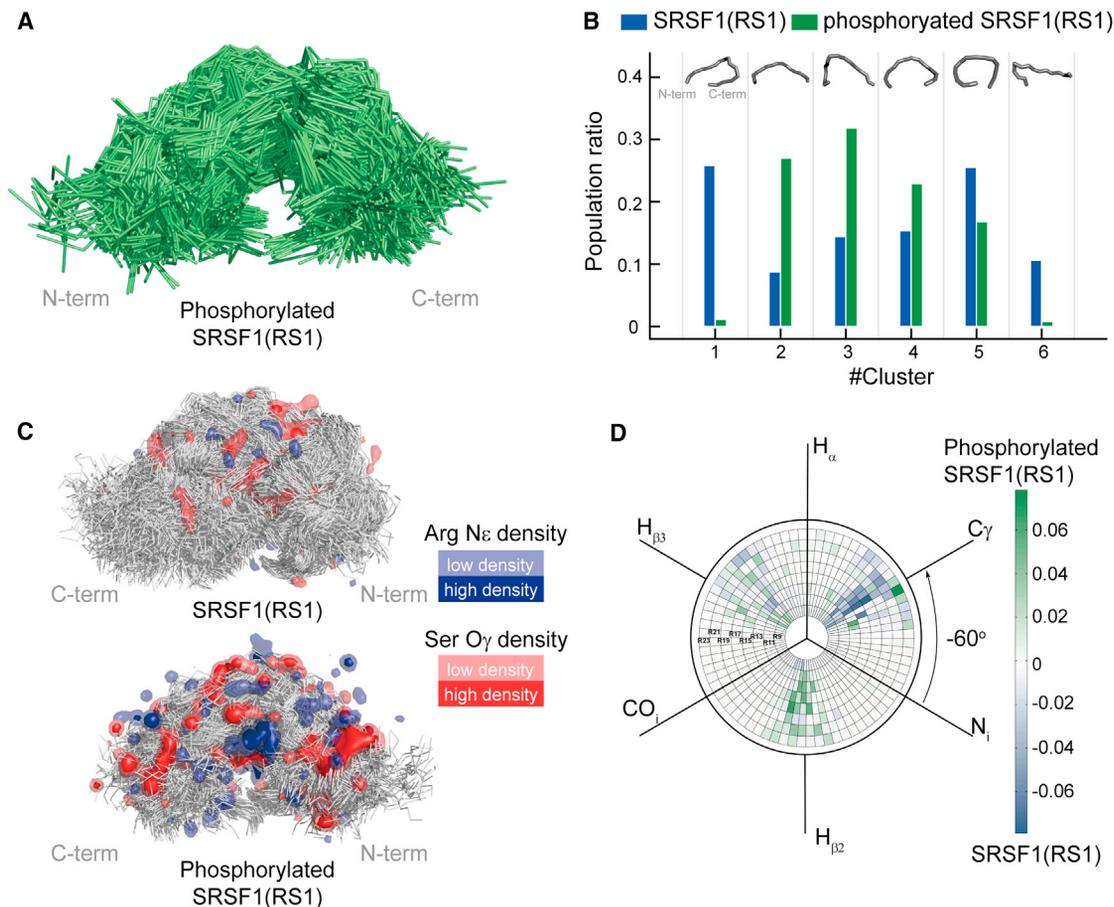
The first SR protein was identified as splicing factor in *Drosophila* in 1991 (Roth et al., 1991). Subsequently, many SR



**Figure 6. Cross-Validation of the Ensembles of Phosphorylated and Unphosphorylated SRSF1(RS1)**

Comparison of experimental NMR data (black), which were not used for ensemble selection, with those back-calculated from the conformer ensembles of phosphorylated (A–D; light and dark green) and unphosphorylated (E–H; cyan and blue) SRSF1(RS1). Chemical shifts predicted by ShiftX2 and Sparta+ are

(legend continued on next page)



**Figure 7. Analysis of Selected Subensembles**

(A) Superposition of 100 independently selected ensembles each containing 30 structures of phosphorylated SRSF1(RS1). Only the protein backbone is shown. (B) Cluster analysis of unphosphorylated and phosphorylated SRSF1(RS1) backbone conformations. Three thousand structures of each type were used in the analysis. Clustering based on the principal component analysis identified six clusters in the plane of the first three eigenvectors. For each cluster the population ratio for phosphorylated versus unphosphorylated SRSF1(RS1) is shown.

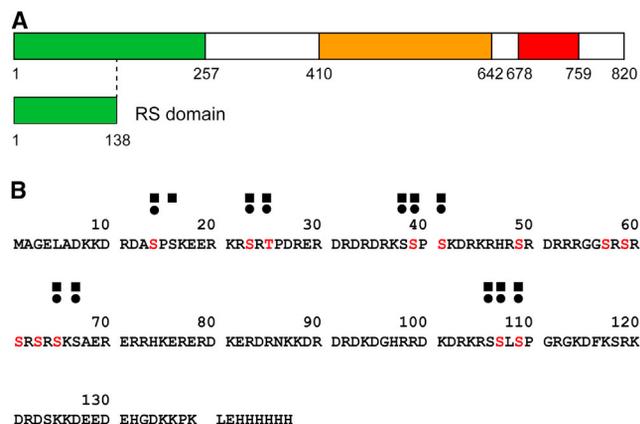
(C) Ensemble of conformations of unphosphorylated (top) and phosphorylated SRSF1(RS1). Three thousand structures from 100 subensembles were used for the analysis. Backbones were superpositioned and are shown in gray. Densities of N $\epsilon$  in arginine side chains (blue) and O $\gamma$  in serine side chains (red) were classified into high- and low-density areas and plotted on the backbone.

(D) Differences in occupancy of arginine side chain dihedral angles  $\chi_1$  between unphosphorylated and phosphorylated SRSF1(RS1). The occupancy of every arginine residue was normalized in such a way that the sum over all angles amounts to 1.0.

proteins were revealed such as human SRSF1 (Ge et al., 1991) and SRSF2 (Fu and Maniatis, 1990) (also known as SC35). SR proteins contain one domain, termed RS domain, that is rich in arginine/serine repeats. Circular dichroism measurements of the RS domain of SRSF1 have suggested that it is predominantly disordered (Labourier et al., 1998). However, as circular dichroism is most sensitive to  $\alpha$  helices and  $\beta$  sheets, and elements of local structure are difficult to detect in the presence of long disordered regions, nonregular structure might be present. Using NMR spectroscopy we showed that the RS domain of SRSF1 is fully disordered (Figure 1). In line with its disordered nature, phosphorylation can efficiently occur in the RS repeats of SRSF1.

Several kinases are responsible for SR protein phosphorylation such as cdc2-like kinases (Manley and Tacke, 1996) and SR protein kinases including SPRK1 (Manley and Tacke, 1996) and SPRK2 (Kuroyanagi et al., 1998). Our mass spectrometric analysis of hPrp28 in native U5 snRNPs and assembled spliceosomal B complexes identified several phosphorylated RS dipeptides (Figure 8B). The phosphorylation pattern was very similar in U5 snRNP and B complexes with the exception of S16 and S63, which were phosphorylated in U5 snRNP but not in B complexes. Because the MS analysis was not done in a quantitative manner, the two sites, which were not detected in B complexes, might have escaped MS detection. In case of the RS domain of

plotted in light and dark green (C and D) and blue (G and H). Errors of experimental data represent SD as estimated on the basis of NMR signal intensities and line widths. Errors of back-calculated couplings represent SD values between the mean values of 100 independently selected ensembles. Because every selection already reports on a mean value (average over 30 structures) the SD of the mean values should be called an SE (however, the explicit division by  $\sqrt{n}$  was not applied in any step).



**Figure 8. Phosphorylation of the SR-Related Protein hPrp28 in Native Spliceosomal Complexes**

(A) Domain organization of hPrp28. The 100 kDa hPrp28 contains three domains: disordered domain (residues 1–257, green), DEAD box domain (residues 410–642, orange), and HELICc domain (residues 678–759, red). Residues 1–138 comprise the RS domain of hPrp28.

(B) Mass spectrometric (MS) determination of phosphorylation sites within the N-terminal RS domain of hPrp28 in native U5 snRNP (square) and functional spliceosomal B complex (circle). Phosphorylation sites determined by NMR after *in vitro* phosphorylation with SRPK1 are in red letters.

SRSF1, SRPK1 preferentially phosphorylates up to 12 serine residues in the N-terminal part (Hagopian et al., 2008; Velazquez-Dones et al., 2005). In line with these studies, our *in vitro* phosphorylation of SRSF1(RS) by SRPK1 resulted in a heterogeneous population of 12 to 17 phosphate groups, precluding structural determination. However, when we only used the N-terminal part of SRSF1(RS), combined with careful purification, we were able to prepare homogeneous samples of SRSF1(RS1) fully phosphorylated at the six serine residues in S207–S217 (Figure 2). A homogeneous and defined state of SRSF1(RS1) was essential for the subsequent structure analysis.

Structure determination of RS domains by X-ray crystallography has been hampered thus far by the dynamic nature of RS domains. In addition, severe sequence degeneracy complicates NMR studies. In order to determine the conformation of phosphorylated SRSF1(RS1), we determined a large number of structural parameters using high-dimensional NMR experiments. Notably, phosphorylated SRSF1(RS1) was not as rigid as secondary structure elements in globular proteins. In line with the possibility of significant mobility in structured molecules, we recently showed that the backbone of a well-defined protein folding intermediate remained highly mobile (Jaremko et al., 2013). To take into account the mobility of phosphorylated SRSF1(RS1), we described SRSF1(RpS1) by ensembles of 30 conformers that are in agreement with a large number of NMR observables. The ensembles revealed an arch-like structure of SRSF1(RpS1) with well-defined preferential orientations of arginine and phosphoserine side chains (Figure 7). The experimentally determined ensemble shows significant deviations from RS repeat models proposed previously by MD simulations (Hamelberg et al., 2007; Sellis et al., 2012). In particular, we did not find any evidence for the compact conformation termed “arginine-claw” (Figure 7; Table S2). The structure of phosphor-

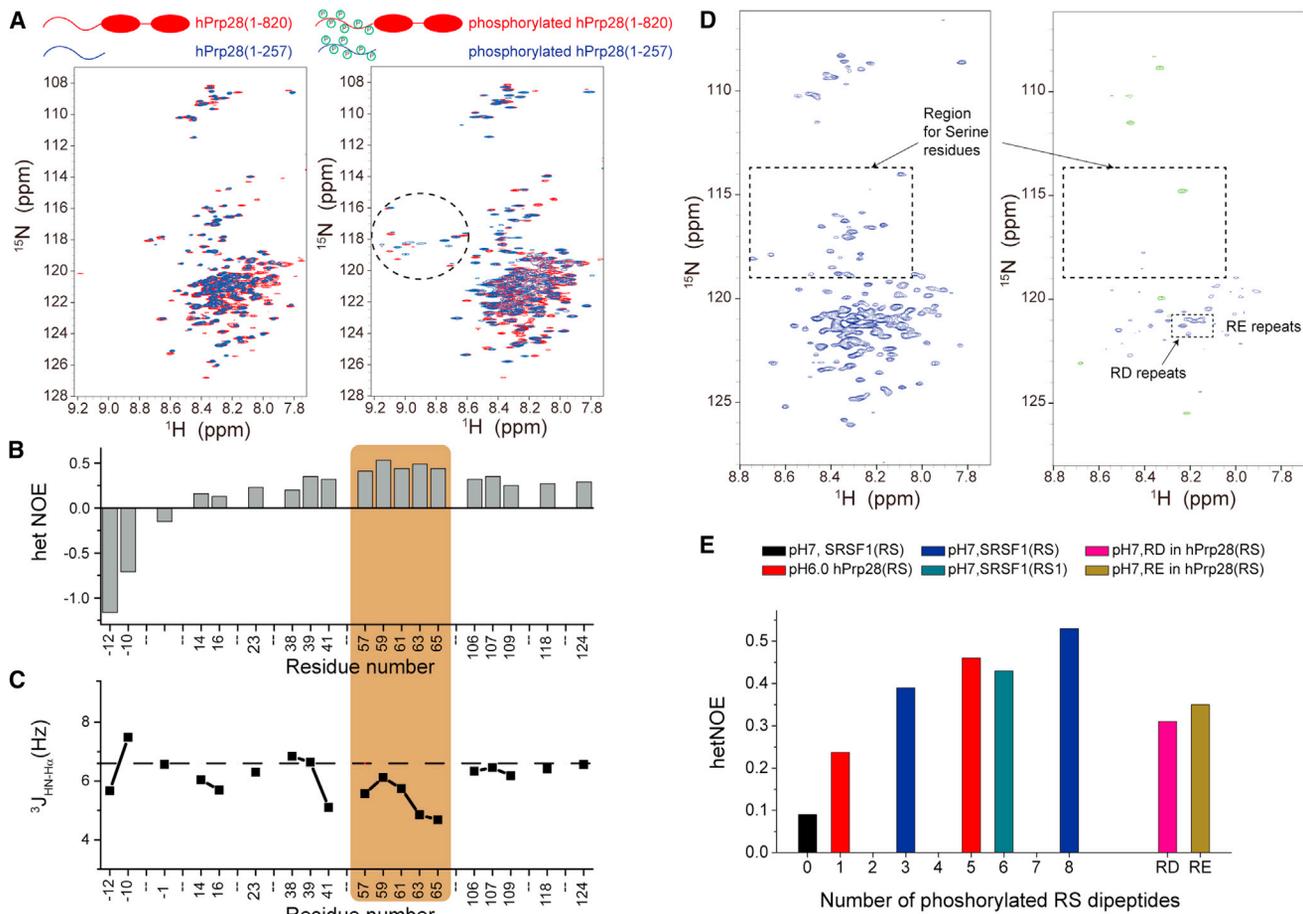
ylated SRSF1(RS1) also differs from the conformation observed for part of unphosphorylated SRSF1(RS1) in complex with SRPK1 (Ngo et al., 2008). Thus, our study provides detailed insights into the structural consequences of SR domain phosphorylation.

Phosphorylation and dephosphorylation of RS domains regulates the activity of SR proteins in spliceosome assembly and splice site selection (Kanopka et al., 1998; Prasad et al., 1999; Sanford and Bruzik, 1999). For example, phosphorylation of SRSF1 enhances the interaction between SRSF1 and U1-70K proteins in A complex formation (Xiao and Manley, 1997). Often these interactions involve binding between two RS domains. In addition, RS domains can directly contact pre-mRNA in a phosphorylation-dependent manner (Shen and Green, 2004, 2007; Shen et al., 2004). Phosphorylation changes the charge of RS repeats from highly positive to highly negative and therefore leads to electrostatic repulsion with RNA (Xiao and Manley, 1997). Beyond a simple charge effect, however, the conformer ensemble of phosphorylated SRSF1(RS1) revealed an arch-like backbone structure with preferential side chain orientations (Figure 7), which influences the molecular recognition.

In intrinsically disordered proteins such as RS domains (Figure 1), the required decrease in entropy from the highly dynamic unbound state to a rigidified complex structure results in lower affinity of interactions and severely influences the binding process. A key finding of our study is that the internal dynamics in the RS domain of SRSF1 strongly decrease upon phosphorylation. Phosphorylation of six consecutive RS dipeptides switches the RS domain of SRSF1 from a fully disordered state to a partially stabilized conformation with a better defined chemical environment (Figures 1 and 2). The phosphorylation-induced stabilization strongly decreases the conformational entropy, as indicated by entropy estimates from the determined structural ensembles. Because of the important contribution of entropy to binding, the decrease in internal dynamics upon phosphorylation will favor binding of RS domains to a wide variety of interaction partners.

An entropy-driven change in SR domain interactions is further supported by the NMR data on RE and RD repeats. RE and RD repeats can partially replace phosphorylated RS repeat sequences in spliceosomal assays (Cazalla et al., 2002). NMR hetNOE values of RE and RD signals in hPrp28 indicate that RE/RD repeats are partially stabilized (Figure 9E). In addition, shorter RS repeat sequences (Figure 9E) or a lower charge state of the phosphate groups (Figure 3) lead to less rigidification when compared to a consecutive stretch of six or eight phosphorylated RS dipeptides. In line with the gradual decrease in dynamics—and therefore in conformational entropy—with increasing RpS repeat length (Figure 9E), the potency of SR proteins correlates with the number of RS dipeptides (Graveley et al., 1998). Indeed, in a single phosphorylated RS dipeptide the  $^{15}\text{N}$ - $^1\text{H}$  hetNOE value was only slightly increased to  $\sim 0.2$  (Figure 9E), which is consistent with the observation that a single RS dipeptide is not sufficient for RS protein function (Heinrichs and Baker, 1997). Taken together, our study provides strong support for an entropy-driven modulation of SR domain interactions.

SR proteins have many redundant functions. For example, both constitutive and alternative splicing showed little impact



**Figure 9. Entropic Tuning in RS-Rich, RE-Rich, and RD-Rich Regions of the 100 kDa Protein hPrp28**

(A) Superposition of  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectra of full-length hPrp28 (red) and hPrp28(1–257) (blue) before (left panel) and after phosphorylation by SRPK1 (right panel). (B)  $^{15}\text{N}$ - $^1\text{H}$  hetNOE values of sequence-specifically assigned residues of phosphorylated hPrp28(1–138). The RS repeat region is highlighted. (C)  $^3J_{\text{HN-H}\alpha}$  couplings of phosphorylated hPrp28(1–138). The dashed line marks the random coil value of serine. (D)  $^{15}\text{N}$ - $^1\text{H}$  hetNOE spectra (reference: left; saturated spectrum: right) of hPrp28(1–257). (E) Averaged  $^{15}\text{N}$ - $^1\text{H}$  hetNOE values in phosphorylated RS repeats of increasing RS dipeptide length. Proteins of origin are specified. Averaged hetNOE values in RE and RD repeats of hPrp28(1–257) are also shown.

in single SR protein knockout cells (Kawano et al., 2000; Longman et al., 2000). The experiments suggested that a particular exonic splicing enhancer sequence may be recognized by several SR proteins (Tacke et al., 1998), or that one exon may contain several exonic splicing enhancers that are recognized by different SR proteins (Schaal and Maniatis, 1999; Lynch and Maniatis, 1995). The data indicate that RS domains in SR proteins are exchangeable or at least similar in activity. A functional redundancy of SR proteins was further supported by RS domain substitution experiments (Shin et al., 2005) and by the finding that a stretch of ten consecutive RS dipeptides partially restored the function of SRSF1 in mRNA splicing (Cazalla et al., 2002). Our study provides insights into the molecular basis of the functional redundancy of SR proteins. In both the canonical SR protein SRSF1 and the SR-related protein hPrp28 the RS domain is disordered in its native state (Figures 1 and 9). Upon RS phosphorylation, a dynamic switch occurs from the disordered state to a partially rigidified structure with preferential side chain orientations (Figure 7). We conclude that the basis for the profound

influence of phosphorylation of RS domains on protein-protein and protein-RNA interactions is the disordered nature of the unphosphorylated state and the dynamic switch to a partially rigidified structure upon phosphorylation.

#### EXPERIMENTAL PROCEDURES

SRSF1(RS) and SRSF1(RS1) were cloned into a pET22 1a plasmid using NcoI and BamHI restriction sites. Recombinant SRSF1(RS) and SRSF1(RS1) were produced in *Escherichia coli* BL21 (DE3) induced by 1 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) at 37°C for 4 hr. Uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled proteins were produced by growing cells in M9 minimal medium using  $^{15}\text{N}$ - $\text{NH}_4\text{Cl}$ ,  $^{13}\text{C}_6$ -glucose as sole nitrogen and carbon sources. NMR experiments were measured at 288 K and 298 K on Bruker 900, 800, 700, and 600 MHz spectrometers equipped with cryoprobes and a 600 MHz spectrometer equipped with a room temperature probe. For MD simulations the Amber99sb\* force field (Best and Hummer, 2009; Hornak et al., 2006) was employed. The topology of phosphoserines was modified by including additional backbone dihedral angles in accordance to the topology of a serine residue to adapt it to the Amber99sb\* force field. To derive ensembles that best match the experimental NMR data, structural ensembles generated by MD were subjected to

a subensemble selection procedure. A full description of the methods is given in the [Supplemental Experimental Procedures](#).

#### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, eight figures, and two tables and can be found with this article online at <http://dx.doi.org/10.1016/j.str.2013.09.014>.

#### ACKNOWLEDGMENTS

We thank Johanna Lehne, Monika Raabe, and Jasmin Corso for excellent technical help in MS and Karin Giller for generation of expression constructs. This work was supported by the DFG Collaborative Research Center 860, Projects A1 (R.L.), A2 (R.F.), A10 (H.U.), and B2 (M.Z.). S.-Q.X. was supported by a stipend from the Max Planck Society-Chinese Academy of Science Doctoral Training Program. V.G. was supported by an International Max Planck Research School for Physics of Biological and Complex Systems stipend.

Received: August 9, 2013

Revised: September 25, 2013

Accepted: September 27, 2013

Published: October 31, 2013

#### REFERENCES

- Banavali, N.K., and Roux, B. (2007). Anatomy of a structural pathway for activation of the catalytic domain of Src kinase Hck. *Proteins* 67, 1096–1112.
- Best, R.B., and Hummer, G. (2009). Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. *J. Phys. Chem. B* 113, 9004–9015.
- Blencowe, B.J., Bowman, J.A.L., McCracken, S., and Rosonina, E. (1999). SR-related proteins and the processing of messenger RNA precursors. *Biochem. Cell Biol.* 77, 277–291.
- Cazalla, D., Zhu, J., Manche, L., Huber, E., Krainer, A.R., and Cáceres, J.F. (2002). Nuclear export and retention signals in the RS domain of SR proteins. *Mol. Cell Biol.* 22, 6871–6882.
- Cheng, Y., Zhang, Y., and McCammon, J.A. (2006). How does activation loop phosphorylation modulate catalytic activity in the cAMP-dependent protein kinase: a theoretical study. *Protein Sci.* 15, 672–683.
- Dowling, D., Nasr-Esfahani, S., Tan, C.H., O'Brien, K., Howard, J.L., Jans, D.A., Purcell, D.F., Stoltzfus, C.M., and Sonza, S. (2008). HIV-1 infection induces changes in expression of cellular splicing factors that regulate alternative viral splicing and virus production in macrophages. *Retrovirology* 5, 18.
- Espinoza-Fonseca, L.M., Kast, D., and Thomas, D.D. (2007). Molecular dynamics simulations reveal a disorder-to-order transition on phosphorylation of smooth muscle myosin. *Biophys. J.* 93, 2083–2090.
- Fu, X.D., and Maniatis, T. (1990). Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. *Nature* 343, 437–441.
- Ge, H., Zuo, P., and Manley, J.L. (1991). Primary structure of the human splicing factor ASF reveals similarities with Drosophila regulators. *Cell* 66, 373–382.
- Graveley, B.R., Hertel, K.J., and Maniatis, T. (1998). A systematic analysis of the factors that determine the strength of pre-mRNA splicing enhancers. *EMBO J.* 17, 6747–6756.
- Gui, J.F., Lane, W.S., and Fu, X.D. (1994). A serine kinase regulates intracellular localization of splicing factors in the cell cycle. *Nature* 369, 678–682.
- Hagopian, J.C., Ma, C.T., Meade, B.R., Albuquerque, C.P., Ngo, J.C., Ghosh, G., Jennings, P.A., Fu, X.D., and Adams, J.A. (2008). Adaptable molecular interactions guide phosphorylation of the SR protein ASF/SF2 by SRPK1. *J. Mol. Biol.* 382, 894–909.
- Hamelberg, D., Shen, T., and McCammon, J.A. (2007). A proposed signaling motif for nuclear import in mRNA processing via the formation of arginine claw. *Proc. Natl. Acad. Sci. USA* 104, 14947–14951.
- Heinrichs, V., and Baker, B.S. (1997). In vivo analysis of the functional domains of the Drosophila splicing regulator RBP1. *Proc. Natl. Acad. Sci. USA* 94, 115–120.
- Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006). Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* 65, 712–725.
- Iakoucheva, L.M., Radivojac, P., Brown, C.J., O'Connor, T.R., Sikes, J.G., Obradovic, Z., and Dunker, A.K. (2004). The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* 32, 1037–1049.
- Jaremko, M., Jaremko, L., Kim, H.Y., Cho, M.K., Schwieters, C.D., Giller, K., Becker, S., and Zweckstetter, M. (2013). Cold denaturation of a protein dimer monitored at atomic resolution. *Nat. Chem. Biol.* 9, 264–270.
- Kanopka, A., Mühlemann, O., Petersen-Mahrt, S., Estmer, C., Ohmalm, C., and Akusjärvi, G. (1998). Regulation of adenovirus alternative RNA splicing by dephosphorylation of SR proteins. *Nature* 393, 185–187.
- Karni, R., de Stanchina, E., Lowe, S.W., Sinha, R., Mu, D., and Krainer, A.R. (2007). The gene encoding the splicing factor SF2/ASF is a proto-oncogene. *Nat. Struct. Mol. Biol.* 14, 185–193.
- Kawano, T., Fujita, M., and Sakamoto, H. (2000). Unique and redundant functions of SR proteins, a conserved family of splicing factors, in *Caenorhabditis elegans* development. *Mech. Dev.* 95, 67–76.
- Kohtz, J.D., Jamison, S.F., Will, C.L., Zuo, P., Lührmann, R., Garcia-Blanco, M.A., and Manley, J.L. (1994). Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. *Nature* 368, 119–124.
- Kuroyanagi, N., Onogi, H., Wakabayashi, T., and Hagiwara, M. (1998). Novel SR-protein-specific kinase, SRPK2, disassembles nuclear speckles. *Biochem. Biophys. Res. Commun.* 242, 357–364.
- Labourier, E., Rossi, F., Gallouzi, I.E., Allemand, E., Divita, G., and Tazi, J. (1998). Interaction between the N-terminal domain of human DNA topoisomerase I and the arginine-serine domain of its substrate determines phosphorylation of SF2/ASF splicing factor. *Nucleic Acids Res.* 26, 2955–2962.
- Li, X., and Manley, J.L. (2005). Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell* 122, 365–378.
- Lin, S., and Fu, X. (2007). SR proteins and related factors in alternative splicing. *Adv. Exp. Med. Biol.* 623, 107–122.
- Long, J.C., and Cáceres, J.F. (2009). The SR protein family of splicing factors: master regulators of gene expression. *Biochem. J.* 417, 15–27.
- Longman, D., Johnstone, I.L., and Cáceres, J.F. (2000). Functional characterization of SR and SR-related genes in *Caenorhabditis elegans*. *EMBO J.* 19, 1625–1637.
- Lynch, K.W., and Maniatis, T. (1995). Synergistic interactions between two distinct elements of a regulated splicing enhancer. *Genes Dev.* 9, 284–293.
- Manley, J.L., and Tacke, R. (1996). SR proteins and splicing control. *Genes Dev.* 10, 1569–1579.
- Mathew, R., Hartmuth, K., Möhlmann, S., Urlaub, H., Ficner, R., and Lührmann, R. (2008). Phosphorylation of human PRP28 by SRPK2 is required for integration of the U4/U6-U5 tri-snRNP into the spliceosome. *Nat. Struct. Mol. Biol.* 15, 435–443.
- Ngo, J.C., Chakrabarti, S., Ding, J.H., Velazquez-Dones, A., Nolen, B., Aubol, B.E., Adams, J.A., Fu, X.D., and Ghosh, G. (2005). Interplay between SRPK and Clk/Sty kinases in phosphorylation of the splicing factor ASF/SF2 is regulated by a docking motif in ASF/SF2. *Mol. Cell* 20, 77–89.
- Ngo, J.C.K., Giang, K., Chakrabarti, S., Ma, C.T., Huynh, N., Hagopian, J.C., Dorrestein, P.C., Fu, X.D., Adams, J.A., and Ghosh, G. (2008). A sliding docking interaction is essential for sequential and processive phosphorylation of an SR protein by SRPK1. *Mol. Cell* 29, 563–576.
- Prasad, J., Colwill, K., Pawson, T., and Manley, J.L. (1999). The protein kinase Clk/Sty directly modulates SR protein activity: both hyper- and hypophosphorylation inhibit splicing. *Mol. Cell Biol.* 19, 6991–7000.
- Roth, M.B., Zahler, A.M., and Stolk, J.A. (1991). A conserved family of nuclear phosphoproteins localized to sites of polymerase II transcription. *J. Cell Biol.* 115, 587–596.

- Sanford, J.R., and Bruzik, J.P. (1999). Developmental regulation of SR protein phosphorylation and activity. *Genes Dev.* *13*, 1513–1518.
- Schaal, T.D., and Maniatis, T. (1999). Multiple distinct splicing enhancers in the protein-coding sequences of a constitutively spliced pre-mRNA. *Mol. Cell Biol.* *19*, 261–273.
- Schlitter, J. (1993). Estimation of absolute and relative entropies of macromolecules using the covariance-matrix. *Chem. Phys. Lett.* *215*, 617–621.
- Sellis, D., Drosou, V., Vlachakis, D., Voukkalis, N., Giannakouros, T., and Vlassi, M. (2012). Phosphorylation of the arginine/serine repeats of lamin B receptor by SRPK1—insights from molecular dynamics simulations. *Biochim. Biophys. Acta* *1820*, 44–55.
- Shen, H., and Green, M.R. (2004). A pathway of sequential arginine-serine-rich domain-splicing signal interactions during mammalian spliceosome assembly. *Mol. Cell* *16*, 363–373.
- Shen, H., and Green, M.R. (2007). RS domain-splicing signal interactions in splicing of U12-type and U2-type introns. *Nat. Struct. Mol. Biol.* *14*, 597–603.
- Shen, H., Kan, J.L., and Green, M.R. (2004). Arginine-serine-rich domains bound at splicing enhancers contact the branchpoint to promote prespliceosome assembly. *Mol. Cell* *13*, 367–376.
- Shin, C., Kleiman, F.E., and Manley, J.L. (2005). Multiple properties of the splicing repressor SRp38 distinguish it from typical SR proteins. *Mol. Cell Biol.* *25*, 8334–8343.
- Staley, J.P., and Guthrie, C. (1999). An RNA switch at the 5' splice site requires ATP and the DEAD box protein Prp28p. *Mol. Cell* *3*, 55–64.
- Tacke, R., Tohyama, M., Ogawa, S., and Manley, J.L. (1998). Human Tra2 proteins are sequence-specific activators of pre-mRNA splicing. *Cell* *93*, 139–148.
- Teigelkamp, S., Mundt, C., Achsel, T., Will, C.L., and Lührmann, R. (1997). The human U5 snRNP-specific 100-kD protein is an RS domain-containing, putative RNA helicase with significant homology to the yeast splicing factor Prp28p. *RNA* *3*, 1313–1326.
- Valcárcel, J., and Green, M.R. (1996). The SR protein family: pleiotropic functions in pre-mRNA splicing. *Trends Biochem. Sci.* *21*, 296–301.
- Velazquez-Dones, A., Hagopian, J.C., Ma, C.T., Zhong, X.Y., Zhou, H., Ghosh, G., Fu, X.D., and Adams, J.A. (2005). Mass spectrometric and kinetic analysis of ASF/SF2 phosphorylation by SRPK1 and Clk/Sty. *J. Biol. Chem.* *280*, 41761–41768.
- Wang, J., Takagaki, Y., and Manley, J.L. (1996). Targeted disruption of an essential vertebrate gene: ASF/SF2 is required for cell viability. *Genes Dev.* *10*, 2588–2599.
- Wu, J.Y., and Maniatis, T. (1993). Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell* *75*, 1061–1070.
- Xiao, S.H., and Manley, J.L. (1997). Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing. *Genes Dev.* *11*, 334–344.
- Xu, X., Yang, D., Ding, J.H., Wang, W., Chu, P.H., Dalton, N.D., Wang, H.Y., Bermingham, J.R., Jr., Ye, Z., Liu, F., et al. (2005). ASF/SF2-regulated CaMKII $\delta$  alternative splicing temporally reprograms excitation-contraction coupling in cardiac muscle. *Cell* *120*, 59–72.
- Zhong, X.Y., Wang, P., Han, J., Rosenfeld, M.G., and Fu, X.D. (2009). SR proteins in vertical integration of gene expression from transcription to RNA processing to translation. *Mol. Cell* *35*, 1–10.