Structural Basis for Homodimerization of the Src-associated during Mitosis, 68-kDa Protein (Sam68) Qua1 Domain*

N. Helge Meyer†§, Konstantinos Tripsianes†§, Michelle Vincendeau*, Tobias Madl†§, Fatiha Kateb†§, Ruth Brack-Werner*, and Michael Sattler†§

From the †Institute of Structural Biology, Helmholtz Zentrum München, Ingolstädter Landstrasse 1, 85764 Neuherberg, the §Munich Center for Integrated Protein Science (CIPS®) at Department Chemie, Technische Universität München, Lichtenbergstrasse 4, 85747 Garching, and the ¶Institute of Virology, Helmholtz Zentrum München, Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany

Sam68 (Src-associated during mitosis, 68 kDa) is a prototypical member of the STAR (signal transducer and activator of RNA) family of RNA-binding proteins. STAR proteins bind mRNA targets and modulate cellular processes such as cell cycle regulation and tissue development in response to extracellular signals. Sam68 has been shown to regulate alternative splicing of the pre-mRNAs of CD44 and Bcl-xL, which are linked to tumor progression and apoptosis. Sam68 and other STAR proteins recognize bipartite RNA sequences and are thought to function as homodimers. However, the structural and functional roles of the self-association are not known. Here, we present the solution structure of the Sam68 Qua1 homodimerization domain. Each monomer consists of two antiparallel α-helices connected by a short loop. The two subunits are arranged perpendicular to each other in an unusual four-helix topology. Mutational analysis of Sam68 in vitro and in a cell-based assay revealed that the Qua1 domain and residues within the dimerization interface are essential for alternative splicing of a CD44 minigene. Together, our results indicate that the Qua1 homodimerization domain is required for regulation of alternative splicing by Sam68.

Sam68 (Src-associated during mitosis, 68 kDa) (1) belongs to the STAR (signal transducer and activator of RNA) family of RNA-binding proteins, which also includes Qk1 (quaking 1), SF1 (splicing factor 1), and Gld-1 (germline development defective-1) (2). STAR family proteins link signaling pathways and many aspects of RNA metabolism (splicing, localization, and translation). They are regulated by post-translational modifications such as phosphorylation, acetylation, and arginine methylation (2).

Sam68 acts in post-transcriptional regulation of pre-mRNA splicing in response to extracellular signals (3). It is involved in a variety of pathways, including insulin and T-cell receptor signaling (4), and plays a key role in cell cycle regulation (5). Sam68 exhibits binding specificity for homopolymeric poly(U) RNA and has been shown to recognize UAAA or UUUA sequences with high affinity as determined by Systematic Evolution of Ligands by Exponential Enrichment (SELEX) and in vivo cross-linking (6, 7). Post-translational modifications can regulate Sam68 function by critically affecting the accessibility to RNA (8, 9). Tyrosine phosphorylation by Src kinase during mitosis enhances the interaction of Sam68 with Ras-GAP (10) but prevents its association with RNA. On the other hand, acetylation of lysine residues by histone acetyltransferases enhances RNA binding (11). Finally, overexpression of Sam68 has been linked to prostate cancer, cell proliferation, and survival (12).

Sam68 has been identified as a key determinant in the alternative splicing of various important RNA targets, like CD44 (13) and Bcl-xL (14), which are linked to apoptosis and cancer. In particular, alternative splicing of CD44 impacts embryonic development and immune response (15–18). Up to 10 variant exon sequences can be included in the mature CD44 mRNA. Among them, variable exon 5 (exon v5) inclusion is associated with tumor progression and T-cell activation (17, 19). Sam68 is a target for phosphorylation by extracellular-signal-regulated kinase (ERK), which promotes inclusion of CD44 exon v5 in response to Ras activation by phosphor ester stimulation. Sam68 binds two RNA sequences within exon v5 and the preceding intron, respectively (9, 13). Recent studies indicate that Sam68 interacts with the heterodimeric splicing factor U2AF (U2 small nuclear ribonucleoprotein particle auxiliary factor) (9). Recognition of the 3′ splice site by U2AF is a key step in spliceosome formation. The interaction with Sam68 is thought to stabilize U2AF binding to its cognate RNA elements. Interestingly, phosphorylation of Sam68 by ERK interferes with RNA binding and reduces pre-mRNA occupancy by U2AF. Altogether, these data indicate that Sam68 can regulate alternative splicing in a signal-dependent manner.

Sam68 shares high similarity with SF1, another binding partner of U2AF. The so-called STAR domain, also referred to as GRP33/SAM68/GLD1 domain, consists of an hnRNP K homology (KH) domain flanked by two domains N- and C-terminal of...
the KH domain, referred to as Qua1 and Qua2, respectively (2, 3). Interestingly, SF1 is the only known member devoid of the Qua1 subdomain and functions as a monomer (20). The KH domain is one of the major RNA binding motifs in eukaryotic cells (21). The structural basis for the recognition of single-stranded RNA by the KH-Qua2 domain of SF1 has been described, indicating that the Qua2 domain extends the KH fold and that the KH-Qua2 tandem domain is essential for sequence-specific RNA recognition (22). The solution structure of the free form of the KH-Qua2 region of Quaking showed that the Qua2 helix does not contact the KH domain in the absence of RNA (23). Biochemical data and in vivo studies suggest that the Qua1 domain can oligomerize (24–27). However, the structural basis for oligomerization and the role for the in vivo function of STAR proteins are unknown.

Here, we present the NMR solution structure of the Sam68 Qua1 domain. Qua1 forms a homodimer composed of a perpendicular interaction of two helical hairpins. A network of hydrophobic and electrostatic interactions stabilizes the dimer interface. Based on NMR and biophysical data, we show that the Qua1 domain is sufficient for homodimerization of Sam68 in vitro. Cell-based splicing assays identify critical residues in the dimer interface and reveal that Qua1 is necessary for the function of Sam68 in alternative splicing.

**EXPERIMENTAL PROCEDURES**

**Sample Preparation—Homo sapiens Sam68 Qua1(95–135), Qua1(95–156), KH-Qua2(147–280), and Qua1-KH-Qua2(95–280), which were derived from pcDNA 3.1 HsSam68 described in Ref. 13, were cloned in a modified pETM-11 vector (European Molecular Biology Laboratory) containing an additional N-terminal double Z-tag and expressed in Escherichia coli BL21(DE3).

Point mutations were introduced in pcDNA 3.1 HsSam68 and pETM-11 ZZ Qua1(95–135) using the QuikChange site-directed mutagenesis (Invitrogen) protocol. The corresponding proteins contain two additional residues at the N-terminal tobacco etch virus protease cleavage site, which are derived from the expression vector.
Bacteria were grown in LB medium for preparation of unlabeled sample or M9 medium supplemented with 13C-labeled glucose and/or [15N]H4Cl for uniformly labeled samples, respectively. The protein was expressed at 20 °C for 16 h after induction with 250 mM IPTG.

Proteins were purified using Ni2+/H10001 affinity chromatography (nickel-nitrilotriacetic acid, Qiagen). After cleavage of the His and the Z-tag with tobacco etch virus protease, the tags were separated from the protein by a second Ni2+ affinity step. Qua1 samples were heat-shocked for 5 min at 85 °C. The samples were further purified by size exclusion chromatography on a HiLoad 16/60 Superdex 75 (GE Healthcare) and kept in 10 mM phosphate, pH 6.5, 100 mM NaCl.

An asymmetrically labeled sample to record intermolecular NOEs was prepared as follows. Equimolar amounts of unlabeled and 15N/13C uniformly labeled Qua1 (residues 95–135) were mixed and incubated with 1% SDS at 85 °C for 10 min. After slowly cooling down, SDS was removed by dilution and ultra filtration using a 15-ml Amicon (Millipore) with a molecular weight cutoff of 5,000.

**NMR Spectroscopy**—NMR measurements were carried out at 298 K on a Bruker Avance III 750-MHz spectrometer equipped with a TXI cryo-probe head, a 600-MHz spectrometer equipped with a TCI cryo-probe head or on an Avance 900 instrument equipped with a TXI cryo-probe head. Spectra were processed with NMRPipe (28) and analyzed with Sparky (29). Backbone assignment was done semi-automatically using MARS (30). For backbone and side chain assignment CBCA(CO)NH, CBCANH and (H)CCH-TOCSY spectra were recorded (31). Distance information was obtained from 15N- and 13C-edited NOESY spectra with a mixing time of 70 ms. To distinguish inter- and intramolecular NOEs, a set of isotope-edited and filtered NOESY spectra was recorded (31, 32). Experiments were carried out using a 1 mM uniformly 15N/13C-labeled or 2 mM asymmetrically 15N/13C-labeled Qua1(95–156) sample, respectively. 15NR1 and R2 relaxations rates and {1H}-15N heteronuclear NOE data were measured at a 750-MHz proton Larmor frequency and 298 K as described (33). H9/C9, N-C, and H9-N residual dipolar couplings were recorded using HNCO-based NMR experiments (34) with a 1 mM Qua1(95–135) sample that was aligned in a medium containing 15 mg/ml Pf1 phage (Profos AG, Regensburg, Germany) as described (35). Paramagnetic relaxation enhancements were measured from saturation recovery 1H,15N HSQC experiments (recovery times between 0.01 and 4.0 s) at concentrations of 0, 1, 2, 3, 4, 5, 7, and 10 mM of the soluble paramagnetic agent gadolinium diethylenetriaminepenta-acetic acid bismethylamide (Gd(DTPA-BMA)). Back-calculation and data analysis were carried out according to Ref. 36.

---

**FIGURE 2. NMR characterization of the Sam68 Qua1 domain.** 13C secondary chemical shifts, 15N R1, R2 relaxation rates, [1H]-15N heteronuclear NOE, and 1H-15N residual dipolar couplings (RDCs) are plotted versus Sam68 Qua1 residue numbers. The secondary structure is indicated at the top. Error bars for 15N relaxation data are estimated from the noise level as described in Ref. 33. Error bars for RDCs are estimated based on spectral resolution.
Structure Calculation—Automated NOE cross-peak assignment was performed using the software CYANA 3.0 (37). Automatically assigned NOEs and completeness of the NOE cross-peaks were manually inspected. Homodimer symmetry is explicitly taken into account for network anchoring, and identical conformation of the two monomers is ensured by dihedral angle difference restraints for all corresponding torsion angles. Additionally, a symmetric relative orientation of the two monomers is maintained by distance difference restraints between symmetry-related intermolecular Cα–Cα distances (38). Distance restraints from the CYANA calculation and TALOS+(39)-derived torsion angles and the residual dipolar coupling restraints were used in a water refinement calculation (40) with Aria 2.2 (41). The quality of the structure ensemble was validated using the iCING3 web server as well as PROCHECK (43) and WHATCHECK (44). Molecular images were generated using PyMOL (45).

Circular Dichroism (CD) Spectroscopy—Temperature series of far-UV (190–250 nm) CD spectra were recorded on 100 μM Qua1 wild type or mutant proteins in 20 mM phosphate, pH 6.5, 50 mM NaCl using a JASCO J-715 spectropolarimeter.

Cell Culture and Transfections—U138MG cells were cultured as described previously (46). All overexpression experiments were performed in 6-well plates. Cells were seeded at a density of 2 x 10^5 cells/well 1 day prior to transfection and cultured for 30 h after transfection. Transfections were performed using FuGENE® HD transfection reagent (Roche Diagnostics) according to the manufacturer’s instructions. For co-transfection experiments, 2 μg of CD44 v5 minigene-plasmid DNA (13) was co-transfected with 300 ng of plasmids expressing Sam68 or wild type or mutants. As a control, the CD44 v5 DNA (13) was co-transfected with 300 ng of plasmids expressing Sam68 or wild type or mutants. As a control, the CD44 v5 DNA was co-transfected with 300 ng of the pcDNA 3.1 expression plasmid. Expression of all proteins was checked by Western blot analysis (data not shown).

RT-PCR Analysis—Cytoplasmic RNA was prepared using the PARIS kit (Ambion) according to the manufacturer’s protocol. 1 units/μg of RNA was transcribed with Superscript II (Invitrogen) according to the manufacturer’s protocol, using random hexamers. RT-PCR analyses were carried out as described in Ref. 13 using 25 PCR cycles. RT-PCR bands were quantified densitometrically using the ImageJ Software. Three independent experiments were performed.

RESULTS

Qua1 Is Sufficient for Sam68 Homodimerization—It has been shown that the STAR domain of Sam68 (Fig. 1A) and other members of the STAR family dimerize (24–27). The apparent molecular mass estimated from size exclusion chromatography of ~40 kDa is consistent with a dimer of the Sam68 STAR domain (not shown). To identify the region that mediates dimerization, we expressed proteins comprising the Qua1, KH-Qua2, and Qua1–KH-Qua2 domains in E. coli. An overlay of ^1H,^15N correlation NMR spectra of the Qua1 (residues 95–135) and KH-Qua2 (residues 147–280) subdomains with that of the STAR domain (residues 95–280) shows that chemical shifts

| TABLE 1 Structural statistics of the Sam68 Qua1 homodimer |
|-------------------------------------------|
| NOE-based distance restraints† | Intraresidual, sequential | 950 | | |
| | Medium range (2 ≤ | 1 | 2 ≤ | 1 | ≤ 4) | 504 | | |
| | Long range (1 | 2 | 1 | ≤ 5) | 316 | | |
| | Intermolecular | 456 | | |
| | Total | 2226 | | |
| Other restraints | φ-ψ dihedral angle restraints | 128 | | |
| | Residual dipolar coupling restraints | 128 | | |
| | (Hα-N, N-Cα, Cα-Cα) | 128 | | |
| Coordinate precision r.m.s.d. | Backbone (Å) | 0.32 ± 0.09 | | |
| | Heavy atom (Å) | 0.97 ± 0.16 | | |
| Consistency (structure vs. restraints) | r.m.s.d. (Å) from experimental distance restraints† | 0.019 ± 0.002 | | |
| | r.m.s.d. (*) from experimental torsion angle restraints‡ | 0.911 ± 0.104 | | |
| | RDC Q-factor‡ | 0.230 ± 0.005 | | |

Structure Z-score

WHATCHECK‡

First generation packing quality 0.497
Second generation packing quality 0.0028
Ramachandran plot appearance 0.001
x1/x2 rotamer normality 2.057
Backbone conformation 1.325

Ramachandran plot‡

Most favored regions 96.0%
Allowed regions 4.0%
Generously allowed regions 0.0%
Disallowed regions 0.0%

† Distance restraints were employed with a soft square well potential using an energy constant of 30 kcal mol^-1Å^-2. No distance restraint was violated by more than 0.5Å.
‡ Torsion angle restraints derived from TALOS (39) were applied to φ, ψ backbone angles using energy constants of 200 kcal mol^-1 radians^-2. No dihedral angle restraint was violated by more than 5°.
§ Residual dipolar couplings (RDCs) were employed with a harmonic potential using an energy constant of 0.5 kcal mol^-1 Hz^-2. Q-factor as defined in Ref. 42.
¶ PROCHECK (43) and WHATCHECK (44) were used to determine the quality of the structure. Positive WHATCHECK Z-scores indicate that structure is better than average.

seen in the smaller domains are very similar to those of common residues in the STAR domain. This indicates that there are no strong contacts between Qua1 and KH-Qua2 and that these two regions are structurally autonomous (Fig. 1B).

The ^1H NMR frequencies in the ^1H,^15N HSQC spectrum of Qua1 (residues 95–156) are mainly found within a small region centered around 8 ppm, indicating the presence of mostly helical and random coil conformation. Analysis of the ^13Cαβ secondary chemical shifts reveals that residues 95–135 comprise two α-helices that are interconnected by a short loop. The 20 C-terminal amino acids (residues 136–156) do not exhibit any secondary structure. ^15N R1 and R2 relaxation rates as well as ^1H T1 and T2 relaxations (28) show that the C terminus is highly flexible, whereas the loop connecting the two helices has a slightly increased flexibility when compared with the helical segments (Fig. 2). Thus, only residues 95–135 define the globular fold of the Qua1 domain. A tumbling correlation time of τc = 9.5 ns was estimated from the ^15N R2/R1 relaxation rates ratio for Qua1, consistent with the molecular mass expected for a homodimer (14 kDa). Even at concentrations of 10 μM, there is no significant change in the ^15N relaxation rates and thus τc (data not shown), indicating that the dimerization constant is at least in the low micromolar to nanomolar range. Additionally,

3 G. W. Vuister, J. F. Doreleijers, and A. W. Sousa da Silva, manuscript in preparation.
the apparent molecular weight determined by size exclusion chromatography matches well with a homodimer (supplemental Fig. 1). These observations suggest that the Qua1 domain is sufficient for homodimerization.

For structural analysis and to further confirm that Qua1 is a homodimer, we attempted to identify intermolecular NOEs. In the case of a symmetric homodimer, the NOESY spectrum contains both intramolecular and intermolecular cross-peaks. To unambiguously distinguish intermolecular NOEs for asymmetrically isotope-labeled dimers, isotope-edited/filtered NMR experiments will detect exclusively intermolecular NOEs for symmetrically isotope-labeled dimers. Isotope-edited/filtered NOESY experiments (Fig. 1) were recorded on an asymmetrically15N/13C-labeled sample from intramolecular NOEs, isotope-filtered experiments derived from13C- and15N-edited NOESY HSQC spectra as well as13C/15N-filtered,15N- or13C-edited NOESY-HSQC spectra. A summary of the structural and restraint statistics is given in Table 1. The ensemble of the 20 lowest energy structures obtained after water refinement is shown in Fig. 3A. The quality of the structures was further validated by comparison of measured and back-calculated relaxation rate enhancements upon the addition of the paramagnetic co-solvent Gd(DTPA-BMA) (36), which showed excellent agreement (supplemental Fig. 2).

Each Qua1 monomer comprises two α-helices that are aligned in an antiparallel fashion with a tilt angle of ~30°, connected by a short loop. Hydrophobic residues are spaced every 3–4 residues apart on each helix and interact with residues arrayed on the other helix of the same monomer, i.e. Leu-104, Leu-107, Lys-111 and Leu-114 in helix α1 contact Ile-132, Ile-129, Leu-125, and Met-122 in helix α2, respectively (supplemental Fig. 3). This arrangement is reminiscent of a short “zipper” where the extensive hydrophobic contacts lead to a tight packing of the two helices of each monomer. A potential side-chain hydrogen bond between the invariant Tyr-103 (helix α1) and Glu-128 (helix α2) further stabilizes the helical hairpin. The lack of the corresponding hydrogen bond may be consistent with a lethal phenotype of the E48G mutation in the mouse Qk1 paralog (Glu-48 corresponds to Sam68 Glu-128). Another important feature of this topology is the highly conserved Pro-116 in the loop connecting the two helices, which allows the reversal of the peptide backbone.

The Qua1 dimer is formed by perpendicular stacking of the two monomers with a C2 symmetry (Fig. 3, A and B). Numerous hydrophobic contacts, which mainly involve the loop region connecting the two helices, stabilize the dimer interface. Below this largely hydrophobic contact area, cross-subunit hydrogen bonds and electrostatic contacts between the N-terminal end of helix α2 and the α1′/α2′ helices of the other monomer provide additional intersubunit interactions (Fig. 3). The interface area covers 624
An arrangement of two antiparallel helices is a common feature in a variety of different proteins and protein folds. However, structural similarity searches (47) did not reveal any structure in the Protein Data Bank (PDB) database with a similar perpendicular arrangement of two helical hairpins, indicating that the Qua1 homodimer adopts a novel four-helix dimer topology.

Mutational Analysis of the Dimer Interface—We confirmed the dimer contacts seen in the structure by mutational analysis. Various single point mutations were designed, and 1H,15N HSQC experiments of the corresponding proteins were recorded to monitor the influence of the mutation on the overall fold. The capability to form homodimers was determined by analysis of apparent local correlation times, $\tau_c^{\text{app}}$, estimated from the 15N relaxation data. The $\tau_c^{\text{app}}$ values reflect the size of the protein tumbling in solution, and the distribution of $\tau_c^{\text{app}}$ therefore allows estimating the ratio between the monomeric and the dimeric forms (supplemental Fig. 4).

Mutation of Phe-118 (F118S) leads to a virtually complete loss of structure and dimerization (Fig. 4A). Remarkably, at concentrations above 1 mM, a small portion of the protein can still dimerize, as indicated by weak NMR signals in the 1H,15N HSQC, which reflect the folded dimer (data not shown). The H120K variant is in slow exchange between unfolded protein and structurally intact dimer, as indicated by weak NMR signals in the 1H,15N HSQC spectra (Fig. 4A). The Y103S and E110A variants maintain the overall fold, and the protein still forms dimers (Fig. 4B). However the distribution of local correlation times, $\tau_c^{\text{app}}$, derived from the NMR relaxation data (supplemental Fig. 4) is slightly decreased in both mutants when compared with the wild type, consistent with a fast exchange between monomeric and dimeric species. The effect of the E110A mutation on dimerization is more pronounced as the average $\tau_c^{\text{app}}$ value is lower than for the Y103S mutant. The E110A Qua1 domain cannot form intermolecular hydrogen bonds to the backbone amides of Ala-121 and His-120, which are indicated by weak NMR signals in the 1H,15N HSQC spectrum of the E110A mutant (Fig. 4B). Thermal denaturation of the Qua1 mutants followed by CD spectroscopy demonstrates that the thermal stability of the Y103S and E110A variants is decreased when compared with

A key determinant of the dimer formation is the almost invariant Phe-118 in the $\alpha1$-$\alpha2$ loop. Its aromatic side chain stacks with the Leu-114 and Ala-121, Leu-124, and Leu-125 mediate additional hydrophobic interactions stabilizing the dimer interface. Near the intersection of the $\alpha2$/$\alpha2'$ helices, the His-120' side chain contacts the aromatic ring of Tyr-103 in the other monomer. Apart from this hydrophobic network, cross-subunit hydrogen bonds between Glu-110 and the backbone amide protons of Ala-121' and His-120' are observed (Fig. 3E, left). This is reflected by large downfield shifts in the 1H NMR frequencies of the backbone amides of Ala-121 and His-120 (Fig. 4). Furthermore, the side chain of His-120' is in close proximity to the side chains of Glu-106 and Gln-123', implying electrostatic stabilization (Fig. 3E, left). Charged residues are distributed at the periphery and at the opposite side of the interface, giving an overall negative charge to the Qua1 homodimer (Fig. 3D).

\[^{\text{2}}\text{A}^2/\text{monomer}, which accounts for more than 18% of the total surface area of one monomer (Fig. 3C).

An arrangement of two antiparallel helices is a common feature in a variety of different proteins and protein folds. However, structural similarity searches (47) did not reveal any structure in the Protein Data Bank (PDB) database with a similar perpendicular arrangement of two helical hairpins, indicating that the Qua1 homodimer adopts a novel four-helix dimer topology.
the wild type protein (supplemental Fig. 5), although both mutants are still mainly dimeric. In contrast, the F118S and H120K mutants are highly unstable or already unfolded at room temperature (data not shown).

In summary, the structure-guided mutagenesis shows a good correlation between the importance of contacts in the interface and dimer formation. Specifically, the F118S mutation completely disrupts the dimer, whereas the other mutations appear to have more local effects only destabilizing the Qua1 dimer.

Functional Qua1 Domain Is Required for Splicing Regulation by Sam68—Sam68 has previously been shown to increase the inclusion of the CD44 exon v5 in mRNAs produced from a CD44v5 minigene (Fig. 5A) (13). Here, we used this assay to investigate whether the Qua1 domain is required for regulation of splicing by Sam68. These experiments were performed in the human astrocytic cell line U138MG because human astrocytes were previously reported to express relatively low levels of endogenous Sam68 (48) (supplemental Fig. 6). Overexpression of wild type full-length Sam68 led to a more than 3-fold increase of exon v5 inclusion (Fig. 5, B and C). As expected, a mutant lacking the KH region, which mediates RNA binding, failed to increase exon v5 inclusion. Interestingly, the mutant lacking the Qua1 domain was similarly ineffective as the ΔKH mutant in this assay. This demonstrates that the Qua1 domain is essential for Sam68 activity in vivo. The same effect is observed for the F118S variant, which in vitro, in the context of the Qua1 domain, does not dimerize. H120K, which is disturbed in dimerization in vitro (supplemental Fig. 4), has significantly reduced splicing activity. In contrast, the E110A mutant is only slightly destabilized in vitro, and, consistently, supports alternative splicing almost as well as the wild type protein. Although mutation of Tyr-103 has the least effect on dimerization in vitro, it seems to be crucial for Sam68 activity as the Y103S variant is virtually inactive in the splicing assay.

DISCUSSION

We show that the Sam68 Qua1 domain adopts a unique four-helix dimer fold with the two monomers stacking perpendicularly. The dimer interface is stabilized by a combination of hydrophobic interactions and cross-subunit hydrogen bonds. Residues that mediate these interactions are highly conserved (Fig. 1A). We demonstrated that mutations of critical residues that build the dimer interface impair the quaternary arrangement. The Qua1 domain is highly conserved within the STAR family, with the exception of the loop region, implying that the Qua1 dimeric structures of these proteins are similar.

In fact, while this manuscript was in preparation, the crystal structure of the Gld-1 Qua1 dimer (PDB 3K6T) was reported (49) with the same topology described here for Sam68. The two structures superimpose very well (backbone coordinate r.m.s.d. 1.1 Å; supplemental Fig. 7), although local differences are seen for intermolecular contacts in the loop region, presumably linked to its different composition.

Our results demonstrate that the Qua1 domain alone is sufficient for dimerization of the Sam68 STAR domain. Although we cannot strictly exclude additional contributions from the KH-Qua2 domain to the dimer interface (25), it seems likely that the Qua1 domain is the main determinant for the dimerization of Sam68. In any case, our results unequivocally show that the Qua1 domain is required for the functional activity in splicing regulation of Sam68 in vivo. Thus, dimerization of Sam68 is at least as important for splicing as the binding to the target RNA itself, which is mediated by the KH-Qua2 domain. As the surface of Qua1 is negatively charged (Fig. 3D), it is unlikely that the Qua1 domain contributes directly to RNA binding. Instead, Qua1 homodimerization may bring together...
two KH-Qua2 domains for recognition of bipartite RNA sequences (Fig. 6), a conserved feature within the STAR family.

With respect to alternative splicing of the CD44 v5 exon, Qua1 dimerization may serve two functions. Firstly, it could stabilize binding to two independent RNA binding sites, one within exon v5 and one in the preceding intron, thus promoting splice site definition by other factors. For instance, Sam68 stabilizes binding of U2AF65 to the CD44 pre-mRNA (Fig. 6), which has no canonical U2AF65 binding site downstream of the branch point RNA sequence (9). Secondly, Sam68 appears to promote rearrangement of the spliceosome by leaving the protein-RNA complex. It has been postulated that this is facilitated by a reduced RNA binding affinity linked to phosphorylation (9).

Phosphorylation of Tyr-103 might act as a possible switch in disassembling the Sam68 dimer and thus releasing the protein from the CD44 pre-mRNA (Fig. 6). This is required for recruiting constitutive splicing factors and spliceosome assembly. Phosphorylation of Tyr-103 might contribute to the release of Sam68 from the pre-mRNA by destabilizing the dimerization via Qua1.

edge of the dimer interface, it is in principle accessible to phosphorylation by a kinase. Phosphorylation could sterically interfere with the dimerization. Additionally, the intermolecular hydrogen-bond network, which we have shown to be crucial for dimerization, could be disturbed by additional hydrogen-bond acceptors and donors provided by the phosphate group. Finally, the additional negative charge of the phosphate might lead to repulsion of the two monomers and thereby destabilize the dimer. Phosphorylation of Tyr-103 might thus play a role for the Sam68 splicing activity, beyond this residue being important for the stability of the Qua1 dimer. Future studies should focus on understanding the molecular mechanisms of the function of the Sam68 STAR domain and its modulation by phosphorylation.

Acknowledgments—We thank Harald König, Karlsruhe Institute of Technology (KIT), for providing Sam68 expression vectors and CD44 minigene constructs and Serge Storz for cloning of the initial Qua1 construct. We thank the Bavarian NMR Centre (BNMRZ) for NMR time.

REFERENCES

1. Taylor, S. J., and Shalloway, D. (1994) Nature 368, 867–871
2. Lukong, K. E., and Richard, S. (2003) Biochim. Biophys. Acta 1653, 73–86
3. Vernet, C., and Artzt, K. (1997) Trends Genet 13, 479–484
4. Najib, S., Martín-Romero, C., González-Yanes, C., and Sánchez-Margalet, V. (2005) Cell Mol. Life Sci. 62, 36–43
5. Pillay, I., Nakano, H., and Sharma, S. V. (1996) Cell Growth Differ. 7, 1487–1499
6. Lin, Q., Taylor, S. J., and Shalloway, D. (1997) J. Biol. Chem. 272, 27274–27280
7. Itoh, M., Haga, I., Li, Q. H., and Fujisawa, J. (2002) Nucleic Acids Res. 30, 5452–5464
8. Wang, L. L., Richard, S., and Shaw, A. S. (1995) J. Biol. Chem. 270, 2010–2013
9. Tisserant, A., and König, H. (2008) PLoS One 3, e1418
10. Guitard, E., Barlat, I., Maurier, F., Schweighoffer, F., and Tocque, B. (1998) Biochem. Biophys. Res. Commun. 245, 562–566
11. Babic, I., Jakymiw, A., and Fujita, D. J. (2004) Oncogene 23, 3781–3789
12. Busà, R., Paronetto, M. P., Farini, D., Pierantozzi, E., Botti, F., Angelini, D. F., Attisani, F., Vespasiani, G., and Sette, C. (2007) Oncogene 26, 4372–4382
13. Matter, N., Herrlich, P., and König, H. (2002) Nature 420, 691–695
14. Paronetto, M. P., Achsel, T., Massiello, A., Chalfant, C. E., and Sette, C. (2007) J. Cell Biol. 176, 929–939
15. Arch, R., Wirth, K., Hofmann, M., Ponta, H., Matzku, S., Herrlich, P., and Zöller, M. (1992) Science 257, 682–685
16. Cooper, D. L., and Dougherty, G. J. (1995) Nat. Med. 1, 635–637
17. Sherman, L., Wainwright, D., Ponta, H., and Herrlich, P. (1996) Curr. Top. Microbiol. Immunol. 213, 249–269
18. Schuster, L., Wainwright, D., Ponta, H., and Herrlich, P. (1998) Genes Dev. 12, 1058–1071
19. Weg-Remers, S., Ponta, H., Herrlich, P., and König, H. (2001) EMBO J. 20, 4194–4202
20. Berglund, J. A., Chua, K., Abovich, N., Reed, R., and Rosbash, M. (1997) Cell 89, 781–787
21. Bud, C. G., and Dreyfuss, G. (1994) Science 265, 615–621
22. Liu, Z., Luyten, I., Bottomley, M. J., Messias, A. C., Hougnignon-Molango, S., Sprangers, R., Zanier, K., Krämer, A., and Sattler, M. (2001) Science 294, 1098–1102
23. Maguire, M. L., Guler-Gane, G., Nietlispach, D., Raine, A. R., Zorn, A. M., Standart, N., and Broadhurst, R. W. (2005) J. Mol. Biol. 348, 265–279
24. Zorn, A. M., and Krieg, P. A. (1997) Genes Dev. 11, 2176–2190
25. Chen, T., Damaj, B. B., Herrera, C., Lasko, P., and Richard, S. (1997) Mol. Biol. Cell.
Structure and Function of the Sam68 Qua1 Domain

28901