Structural basis of STAT2 recognition by IRF9 reveals molecular insights into ISGF3 function

Srinivasan Rengachari[a], Silvia Groiss[a], Juliette M. Devos[a], Elise Caron[b], Nathalie Grandvaux[b,c], and Daniel Panne[a,1]

[a]European Molecular Biology Laboratory, 38042 Grenoble, France; [b]CRCHUM–Research Center, Centre Hospitalier de l'Université de Montréal, Montréal, H2X0A9, QC, Canada; and [c]Department of Biochemistry and Molecular Medicine, Faculty of Medicine, Université de Montréal, Montréal, H3C 3J7, QC, Canada

Edited by George R. Stark, Lerner Research Institute, The Cleveland Clinic Foundation, Cleveland, OH, and approved December 6, 2017 (received for review October 20, 2017)

Cytokine signaling through the JAK/STAT pathway controls multiple cellular responses including growth, survival, differentiation, and pathogen resistance. An expansion in the gene regulatory repertoire controlled by JAK/STAT signaling occurs through the interaction of STATs with IRF transcription factors to form ISGF3, a complex that contains STAT1, STAT2, and IRF9 and regulates expression of IFN-stimulated genes. ISGF3 function depends on selective interaction between IRF9, through its IRF-association domain (IAD), with the coiled-coil domain (CCD) of STAT2. Here, we report the crystal structures of the IRF9–IAD alone and in a complex with STAT2–CCD. Despite similarity in the overall structure among respective paralogs, the surface features of the IRF9–IAD and STAT2–CCD have diverged to enable specific interaction between these family members. We derive a model for the ISGF3 complex bound to an ISRE DNA element and demonstrate that the observed interface between STAT2 and IRF9 is required for ISGF3 function in cells.

JAK/STAT signaling | IRF transcription factor | STAT2 | innate immunity | crystal structure

Cytokine signaling via the JAK–STAT pathway controls the development, differentiation, and regulation of cells in the immune system and is frequently dysregulated in disease (1). JAK–STAT signaling is mediated by four structurally related JAK kinases (JAK1, JAK2, JAK3, TYK2) and seven STAT (1–4, 5a, 5b, 6) proteins (2). A hallmark of cytokine signaling is functional redundancy and extensive pleiotropy, the ability of multiple cytokines to exert overlapping biological activities (3, 4). A critical question is how a limited number of JAK and STAT molecules enable such extensive redundancy and pleiotropy and how gene duplication and divergence among STATs contributes to specificity in cytokine signaling.

JAK-mediated tyrosine phosphorylation of STATs induces dimerization and translocation to the nucleus, where STATs bind the gamma-activated sequence (GAS), a palindromic 9–11 base pair (bp) DNA element, 5′-TTCN$_{2-4}$GAA-3′ in the promoter of target genes (2). An exception occurs in the response to type I and type III IFNs: These cytokines are rapidly induced during viral infection and stimulate activation of a complex termed ISGF3 (IFN-stimulated gene factor 3). ISGF3 contains a STAT1/STAT2 heterodimer that interacts with IRF9, a member of the IRF family of transcription factors (5–8). Mammals contain 10 IRF paralogs that typically bind to the consensus DNA sequence 5′-AANNGAAA-3′ (9–13). As a result of STAT and IRF complex formation, ISGF3 binds to a ∼12–15-bp composite IFN-stimulated response DNA element (ISRE) 5′-G/ANGAAAN$_2$GAAACT-3′. Thus, the physical association of STATs with IRFs contributes to functional specificity in cytokine signaling and enables expression of ISGs (6, 14).

IRFs contain a conserved N-terminal DNA-binding domain (DBD) and a C-terminal IRF-association domain (IAD; Fig. 1*A*). The IAD belongs to the SMAD/FHA domain superfamily (15–17). IRF3 is the best understood IRF family member. Signals emanating from pattern recognition receptors (PRRs) activate the kinase TBK1, which phosphorylates latent IRF3 (13, 18–20). This phosphorylation results in remodeling of an autoinhibitory segment in the IAD of IRF3, leading to dimerization and interaction with the transcriptional coactivators CBP/p300 (21). Although IRF9 contains a structurally related IAD, it does not share the same activation mechanism and coactivator preference with IRF3: IRF9 binds to STAT2 in both unstimulated and type I/III IFN-stimulated cells (22–25). The interaction requires the IRF9–IAD and the coiled-coil domain (CCD) of STAT2 (23), a domain that is conserved among STAT paralogs (Fig. 1*A*). Thus, despite conservation of the IRF9–IAD and the STAT2–CCD among their respective paralogs, only STAT2 and IRF9 interact selectively.

To explore the molecular basis that enables selective IFR9–STAT2 interaction, we have determined the crystal structure of the IRF9–IAD in isolation and in complex with STAT2. As expected, the IRF9–IAD is closely related to IRF3. However, IRF9 lacks the structural elements involved in IRF3 autoinhibition, thus explaining the different activation, cofactor, and oligomerization requirements. IRF9 binds through a hydrophobic patch on the convex side of the IAD to the tip of the STAT2–CCD. The binding interface is conserved in IRF9 and STAT2 orthologs but is divergent in other IRF and STAT paralogs. Point mutations in the conserved interface disrupt STAT2–IRF9 interaction in vitro and in vivo and abolish ISGF3 activity in cells. Overall, our data reveal the molecular basis for selective IRF9 and STAT2 interaction. Comparative structural

Significance

Cytokines interact with their receptors and activate JAK–STAT signaling pathways that lead to changes in gene expression. In mammals, there are seven STATs that have arisen due to gene duplication and genetic drift. STATs have similar DNA binding specificity, and how individual STATs have subfunctionalized to regulate very specific cytokine responses in cells is poorly understood. Here we describe X-ray structures that show how one STAT family member, STAT2, specifically pairs with a member of the IRF family of transcription factors, IRF9. Despite overall structural similarity among STAT and IRF family members, surface features in the interacting domains of IRF9 and STAT2 have diverged to enable specific interaction between these family members and to enable the antiviral response.

Author contributions: S.R., J.M.D., N.G., and D.P. designed research; S.R., S.G., J.M.D., and E.C. performed research; S.R., E.C., and N.G. contributed new reagents/analytic tools; S.R., S.G., N.G., and D.P. analyzed data; and S.R. and D.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.wwpdb.org [PDB ID codes 5OEM (structure of the IRF9–IAD) and 5OEN (structure of the IRF9–IAD/STAT2–CCD complex)].

[1]To whom correspondence should be addressed. Email: panne@embl.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1718426115/-/DCSupplemental.

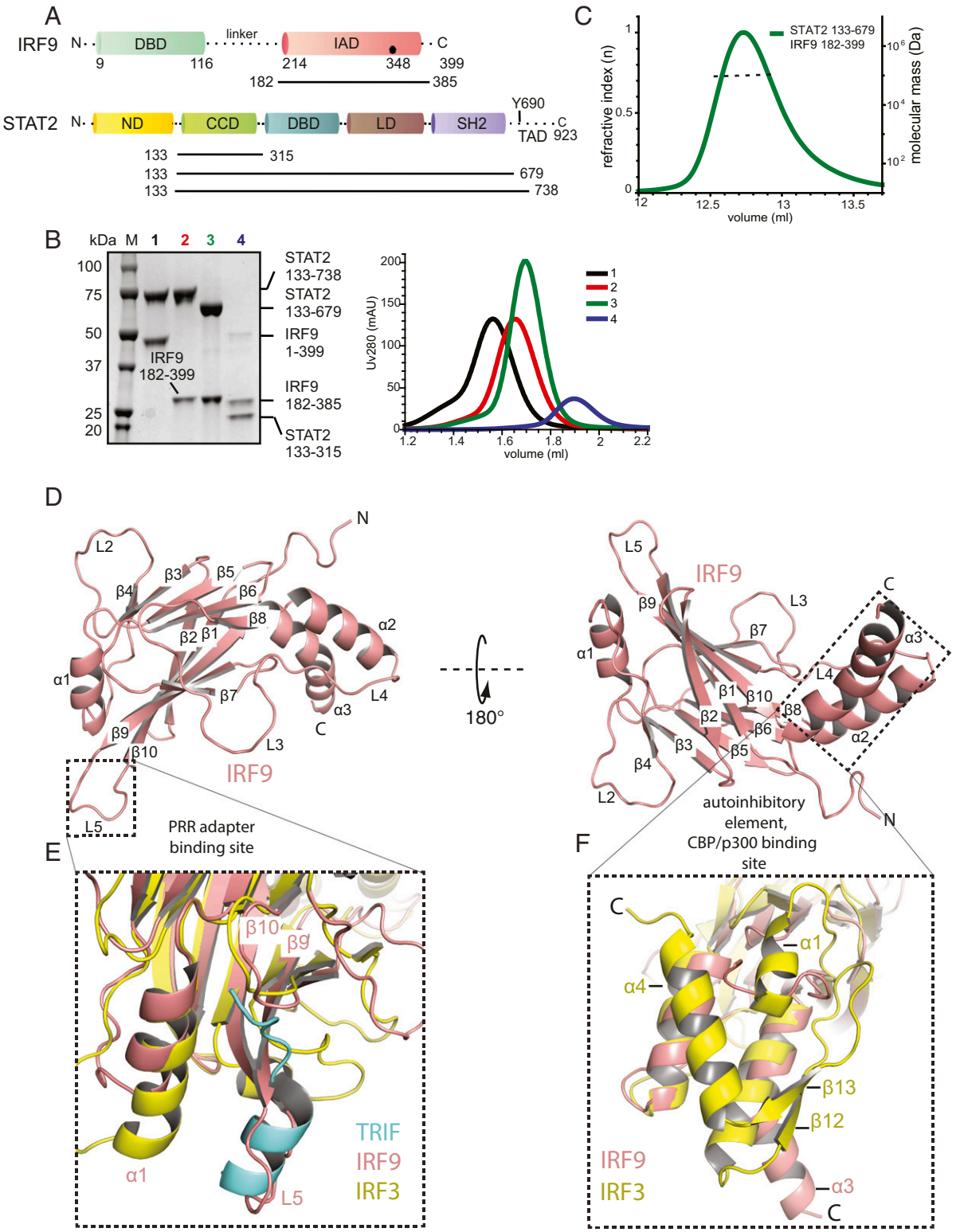

Fig. 1. Characterization of the STAT2–IRF9 complex and structure determination of the IRF9–IAD. (*A*) Domain organization of STAT2 and IRF9. CCD, coiled-coil domain; DBD, DNA-binding domain; IAD, IRF-association domain; LD, linker domain; ND, N-domain; SH2, Src homology domain; TAD, transactivation domain. (*B*) SEC of different STAT2–IRF9 complexes. Fractions were analyzed by SDS–PAGE. Bands corresponding to each protein are indicated. The analyzed samples are 1: $STAT2_{133–738}$–IRF9FL; 2: $STAT2_{133–738}$–$IRF9_{182–399}$; 3: $STAT2_{133–679}$–$IRF9_{182–399}$; 4: $STAT2_{133–315}$–$IRF9_{182–385}$. (*C*) SEC and MALLS profile of the $STAT2_{133–679}$–IRF9–IAD complex. The black line represents the average molecular mass across the peak. (*D*) Ribbon diagram representation of the IRF9–IAD structure (red). The secondary structure elements are labeled. The model is rotated 180° between *Left* and *Right*. (*E*) Structural overlay of IRF9 (red) and IRF3 (yellow). The extended strands β9 and β10 of IRF9 block access of the PRR adaptor TRIF (blue). (*F*) Structural overlay of IRF9 and autoinhibited IRF3–IAD (yellow). The autoinhibitory element of IRF3 is absent in IRF9.

and biochemical analyses yield insights into how gene duplication, evolutionary drift, and natural selection of STATs and IRFs resulted in expansion of the gene regulatory repertoire in cytokine signaling.

Results

Biochemical Basis for STAT2–IRF9 Interaction. STAT2 contains an N-domain (ND), a CCD, a DBD, a linker domain (LD), an SH2 domain, and a C-terminal transactivation domain (TAD) (Fig. 1*A*). Phosphorylation on Y690 enables dimerization of STAT2 (5, 26). IRF9 contains an N-terminal DBD and a C-terminal IAD connected by a flexible linker (Fig. 1*A*). Previous data show that the region spanning amino acids 148–324 of STAT2 and amino acids 217–377 of IRF9 are required for complex formation (8, 23, 27). To further delimit the interacting regions, we coexpressed $STAT2_{133–738}$ and full-length IRF9 (IRF9FL) and purified the complex to homogeneity. We found that IRF9FL copurified with $STAT2_{133–738}$ (Fig. 1*B*, lane 1) and that a C-terminal fragment of IRF9 spanning 182–399 was sufficient for STAT2 binding (Fig. 1*B*, lane 2). Limited proteolysis using Trypsin digestion resulted in C-terminal truncation of STAT2 at position 679 as revealed by liquid chromatography–mass spectrometry (Fig. S1*A*). This shows that the region beyond the SH2 domain is flexibly attached and is not required for complex formation. Coexpression of $STAT2_{133–679}$ with $IRF9_{182–399}$ resulted in a stable complex (Fig. 1*B*, lane 3). Analysis by size exclusion chromatography (SEC) coupled to multiangle laser light scattering (SEC–MALLS) showed a monodisperse complex with a molecular mass of 86.8 ± 0.7 kDa, in agreement with the expected mass of a 1:1 heterodimer (Fig. 1*C*). Chymotryptic digestion resulted in cleavage of the 14 terminal residues of IRF9, further delimiting the interacting region (Fig. S1*B*). Coexpression of $STAT2_{133–679}$ or of the coiled-coil segment $STAT2_{133–315}$ with $IRF9_{182–385}$ resulted in stable and monodisperse heterodimeric complexes (Fig. 1*B*, lane 4). We conclude that the minimal regions required for complex formation comprise $STAT2_{133–315}$ and $IRF9_{182–385}$. Considering that ISGF3 contains a single copy of STAT1 and STAT2 (28), we propose an overall 1:1:1 stoichiometry.

The IRF9–IAD Lacks the Autoinhibitory Element and the PRR Adapter Binding Site. We obtained crystals of the $IRF9_{182–385}$ IAD, which diffracted to 1.9 Å resolution, and determined the structure by molecular replacement. Our final model includes residues 197–385 of IRF9 (Table 1). Like other IAD domains (15, 17, 29), IRF9–IAD has a MH2 domain fold—a central β-sandwich core (β1–β10) flanked by a set of helices and loops (Fig. 1*D*). The domain has a crescent-like shape with a two-helix bundle (α2, α3) on one end, where the N and C termini are located, and helix α1 on the other end. While the β-sandwich core of the IAD domain is conserved among IRFs, the connecting secondary structures and loop regions vary: For example, IRF9 lacks the C-terminal tail of IRF3, constituted by the strands β12 and β13 and helices α1 and α4, which autoinhibit IRF3 in the latent form (Fig. 1*F*). This C-terminal element of IRF3 undergoes a conformational change from a buried, autoinhibitory configuration to an extended coil structure that leads to formation of a domain-swapped dimer upon TBK1-mediated phosphorylation (Fig. S1*C*) (21, 30). The rearranged IAD of IRF3 thereby exposes a hydrophobic binding site for the coactivators CBP/p300 (Fig. S1*C*). The missing autoinhibitory element of IRF9 renders

Table 1. X-ray data collection, phasing, and refinement statistics

Parameters	IRF9–IAD	STAT2–IRF9
Data collection		
Space group	$P3_221$	$P2_1$
Cell dimensions		
a, *b*, *c*, Å	76.77, 76.77, 85.60	30.59, 123.96, 50.99 (β = 92.1°)
Wavelength, λ	0.966	0.972
Resolution, Å	42.8–1.9	47.2–2.9
Total reflections	66,664	22,612
Unique reflections	22,819	7,970
R_{sym} or R_{merge}	0.03 (0.94)	0.19 (0.90)
$I/\sigma I$	14.5 (1.0)	5.9 (1.5)
Completeness, %	97.2 (82.2)	96.0 (98.7)
Redundancy	2.9 (2.1)	2.8 (2.8)
Wilson *B*-factor, Å^2	42.7	34.2
Refinement		
R_{work}/R_{free}	20.6/24.2	25.5/30.7
No. of atoms	1,528	2,769
Protein	1,491	2,698
Ligand	5	N/A
Solvent	32	19
B-factors, Å^2		
Protein	70.3	38.7
Ligand	68.3	N/A
Solvent	55.7	20.5
Rmsd		
Bond lengths, Å	0.014	0.002
Bond angles, °	1.495	0.46
Ramachandran distribution		
Favored	93.05	96.0
Allowed	6.95	4.0
Outliers	0.0	0.0

To calculate R_{free}, 5% of the reflections were excluded from the refinement. R_{sym} is defined as $R_{sym} = \Sigma_{hkl}\Sigma_i|Ii(hkl) - \langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i Ii(hkl)$. Values in parentheses are for the highest resolution shell. N/A, not applicable.

the hydrophobic residues of helices α2 and α3 surface exposed (Fig. 1 *D* and *E*).

In IRF9, the region comprising the helix α1, strands β9 and β10, and the connecting loop L5 correspond to the binding of the pLxIS motif of phosphorylated PRR adaptors such as STING, MAVS, TRIF, or alternatively, the phosphorylated C-terminal tail to form a loop-swapped dimer in IRF3 (Fig. 1*E* and Fig. S1*E*) (21). The extended structure of the strands β9 and β10 and the connecting loop L5 in IRF9 abolish access to the PRR adaptor binding/dimerization site (Fig. 1*E*). The lack of the PRR adaptor binding site and of the autoinhibitory/dimerization element thus explains the different activation and oligomerization properties of IRF9 compared with IRF3 and other IRF family members.

STAT2 Binds to the Convex Surface on the IRF9–IAD. To identify regions of IRF9 that are potentially involved in STAT2 binding, we compiled sequence alignments of IRF9 from divergent vertebrates and mapped the amino acid conservation onto the IAD structure (Fig. 2*A*). We focused on conserved residues of the convex surface of the β-sandwich (IF1) and helix α2 (IF2) and on residues of helix α2 and α3 (IF3) that are involved in CBP/p300 binding in IRF3. We mutated these three putative interfaces in IRF9 and tested binding to STAT2 using a Ni^{2+}-affinity pull-down assay (Fig. 2*B*). Mutations in IF1 reduced or completely abolished IRF9 interaction with STAT2 (Fig. 2*B*). These include IF1-A (R236E, F283A), IF1-B (R236E, L274A, F283A), or IF1-C (L233A, R236E, L274A, F283A). In contrast, mutations in IF2 (R324E, D325K, Q331E, Q333E, P335S) or IF3 (L326A, F330A, I376A) retained STAT2 binding activity (Fig. 2*B*). Thus, our data suggest that STAT2 binds to the convex surface of the IRF9–IAD through residues in IF1.

To analyze the impact of these mutants on STAT2–IRF9 interaction more quantitatively, we measured the equilibrium binding isotherms by isothermal titration calorimetry (ITC). $STAT2_{133–679}$ bound to the IRF9–IAD exothermically, with an equilibrium dissociation constant (K_d) of 10 nM (Fig. 2*C*). As expected, the IF2 and IF3 mutants retained close to wild-type binding affinity (Table 2). The double mutant IF1-A and the triple mutant IF1-B had barely detectable STAT2 binding activity, while binding of the quadruple mutant IF1-C was completely abolished (Fig. 2*C* and Fig. S3 *A* and *B*). Thus, our data indicate that the conserved residues in IF1 synergistically contribute to STAT2 binding. Previous studies have implicated a potential interaction between STAT1 and the C terminus of IRF9, but the relevance for ISGF3 signaling has remained unclear (8, 23, 31). We analyzed the interaction between $STAT1_{136–713}$ and the IRF9–IAD by ITC and found an equilibrium dissociation constant, K_d, of 5 μM (Fig. 2*C* and Table 2). The interaction with STAT1 is also mediated by the IF1 interface of IRF9, as the IF1-C mutant completely abolished STAT1 binding (Fig. S3*C*). The 500-fold higher binding affinity for STAT2 likely explains why IRF9 constitutively interacts with STAT2 but not STAT1 in ISGF3 signaling (23).

Structure of the STAT2–IRF9 Complex. We obtained crystals of $IRF9–IAD_{182–385}$ in complex with the $STAT2–CCD_{133–315}$, which diffracted to 2.9 Å resolution, and we determined the structure by molecular replacement (Table 1). The production of diffraction-quality crystals required surface entropy reduction by introducing mutations Q242A and K243A in an exposed surface loop of STAT2. We located a single copy of the STAT2–CCD in the asymmetric unit bound to the convex side of the IRF9–IAD (Fig. 3*A*). The STAT2–CCD contains a four-helix bundle with a left-handed twist similar to that of other STAT proteins including STAT1, -3, and -5 with which it superposes with a root-mean-square deviation (rmsd) of between 2.7 and 2.9 Å for 151 aligned Cα atoms (Fig. 3*G*). As in other STATs, with the

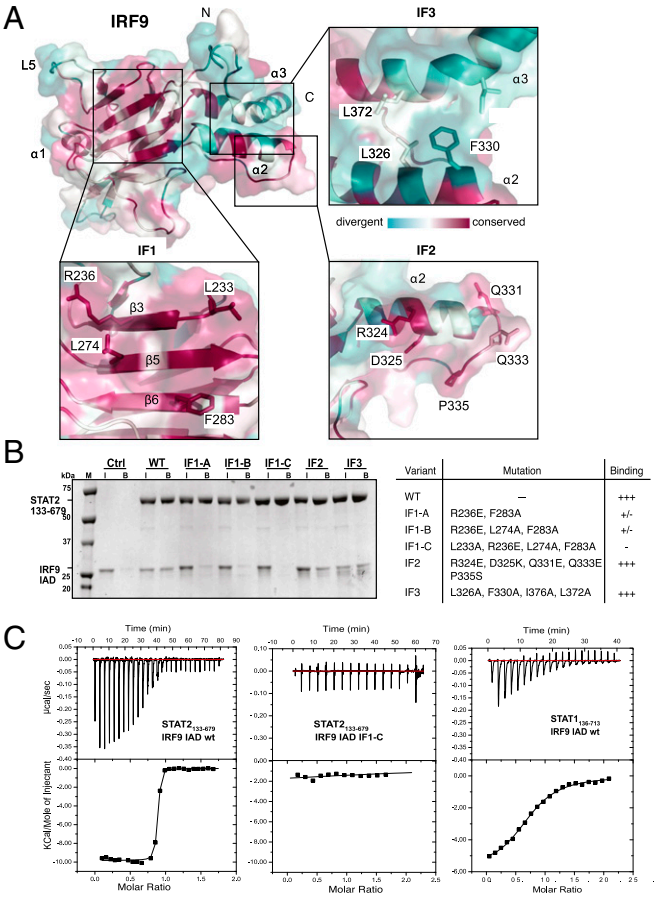

Fig. 2. Characterization of the STAT2 binding interface on IRF9. (*A*) The IRF9–IAD is colored according to amino acid conservation among vertebrate IRF9 orthologs. Magenta, well conserved; light blue, highly variable; for sequence alignment, see Fig. S2. Also shown are zoom-ins of the conserved surfaces boxed in *A*, showing the underlying amino acid residues that were analyzed by mutagenesis. (*B*) Purified His_6-tagged $STAT2_{133–679}$ was incubated with IRF9–IAD variants containing point mutations in the different conserved surfaces. The resulting complexes were incubated with Ni^{2+}-affinity resin and bound proteins analyzed by SDS–PAGE. Lanes are labeled with I (input) and B (bound). Table displays a summary of mutants analyzed: Mutants in the IF1 interface reduced or abolished STAT2 binding. Mutation of residues in the IF2 or IF3 interface did not interfere with STAT2 binding. (*C*) ITC binding curves for the interaction between $STAT2_{133–679}$ and IRF9–IAD, $STAT2_{133–679}$ and IRF9–IAD IF1–C, or $STAT1_{136–713}$ and IRF9–IAD. Ctrl, control.

exception of STAT6, helices α1 and α2 extend beyond the core of the four-helix bundle of the CCD. The β-sandwich core of the IRF9–IAD binds to this extension of α1 and α2 of STAT2 and buries 1,040 $Å^2$ of surface area (Fig. 3*A*). The most significant contribution to the binding interface is made by contacts between α1 of STAT2, which packs against a shallow binding groove on the surface of the IRF9–IAD. STAT2 is the most divergent STAT family member, apparently due to frequent viral targeting that has driven STAT2 divergence (32). Inspection of amino acid conservation among 20 vertebrate STAT2 orthologs revealed a conserved patch on the surface of α1 (Fig. 3*D*). This patch, comprising residues D167, D170, V171, F174, and R175, faces directly into the STAT2–IRF9 binding interface. The lead anchoring residue is F174, which buries the largest solvent-accessible surface area (142 $Å^2$) upon complex formation. F174 binds into a pocket on the surface of IRF9 (Fig. 3 *B* and *C*). Unsurprisingly, a F174D mutation completely abolished complex formation with IRF9 (Fig. S3*D*). V171 and V178 are secondary

Table 2. Summary of ITC data

STAT variant	IRF9 variant	Mutation(s)	K_d
STAT2 WT	IRF9 WT	—	10 nM
	IF1-A	R236E, F283A	n.b.
STAT2 WT	IF1-B	R236E, F283A, L274A	n.b.
	IF1-C	R236E, F283A, L274A, L233A	n.b.
STAT2 WT	IF2	R324E, D325K, Q331E, Q333E, P335S	11 nM
STAT2 WT	IF3	L326A, F330A, I376A	17 nM
STAT2 F174D	IRF9 WT	F174D	n.b.
STAT2 V171E	IRF9 WT	V171E	2.5 μM
STAT1 WT	IRF9 WT	—	5 μM
STAT1 WT	IF1-C	R236E, F283A, L274A, L233A	n.b.
STAT1 E169V	IRF9 WT	E169V	116 nM

n.b., no binding.

anchoring residues and project into two hydrophobic pockets on IRF9. Additional interacting residues of STAT2 include D167, D170, and R175 from helix α1 and Q194, and L197 from helix α2, which form hydrophobic interactions, van der Waals contacts, and salt bridges with conserved surface residues in IRF9 (Fig. 3*C*).

On IRF9, the interacting amino acids are L233, R236, S247, M248, L274, A276, N278, F283, and Q285. The four conserved IRF9 amino acids L233, R236, L274, and F283, which are required for STAT2 binding (Fig. 2 *B* and *C*), contribute a large fraction of the overall buried surface area in this interface. Residues L274 and F283 together with A276, D234, and M278 line the pocket that binds F174 of STAT2.

The IRF9–IAD is closely related to those of IRF3–8 and superposes with the IADs of IRF3–IRF5 with an rmsd of between 2.7 and 2.9 Å for 170 aligned Cα atoms (Fig. S3*G*). Key residues that are involved in STAT2 interaction are conserved in IRF9 orthologs and explain binding selectivity. A critical amino acid residue of IRF9 is A276. A residue with a short side chain at this position is required for formation of the binding pocket for F174 of STAT2. The methyl group of A276 makes hydrophobic contacts with the side chain of F174 of STAT2 and is completely buried upon complex formation. A276 is replaced by a bulky amino acid in other IRFs (Fig. 3*F* and Fig. S2). As a result, the binding pocket is absent in other IRFs, thus preventing association of STAT2 (Fig. S3*G*). Hydrophobic pockets of IRF9 that accommodate the secondary anchoring residues V171 and V178 of STAT2 are also missing in other IRFs (Fig. S3*G*). Comparison of IRF9 in the free and STAT2-bound state shows the conformational rearrangements of IRF9 upon STAT2 binding (Fig. S3*H*). Loop L2 is displaced to accommodate STAT2 on the shallow binding groove of IRF9. This binding groove is occluded in other IRFs due to a longer L2 loop that projects into the binding site (Figs. S2 and S3*H*). In short, despite sharing a similar overall structure with other IRFs, the surface features of the IRF9–IAD have diverged to enable specific recognition of STAT2.

Sequence comparison of STATs shows that F174, the lead-anchoring residue for IRF9 binding of STAT2, is conserved in STAT1 and STAT3 but divergent in other paralogs (Fig. 3*E*). This likely explains why STAT1 can interact with IRF9 (8). The 500-fold lower binding affinity of STAT1 for IRF9 (Fig. 2*C*) is due to divergence of the secondary anchoring residues V171 and V178. In particular, the negatively charged amino acid residues that replace V171 in STAT1 and STAT3 would clash with hydrophobic residues of IRF9. To demonstrate the importance of this interaction, we introduced a V171E mutation in STAT2 and found a 250-fold increased K_d (from 10 nM to 2.5 μM) for IRF9 binding (Fig. S3*E*). Conversely, introduction of E169V into STAT1 resulted in a more than 40-fold reduced K_d (from 5 μM to 116 nM; Fig. S3*A*). Together, while association between STAT2 and IRF9 is driven by F174 of STAT2, the secondary anchor residues and in particular V171 are critical for high binding affinity and specificity of the interaction.

Effects of Point Mutations on ISGF3 Activity. To further analyze the physiological relevance of the IRF9–STAT2 interface for ISGF3 function in cells, we introduced mutations into FLAG-tagged IRF9 and HA-STAT2 and tested their interaction in HEK293 cells. While IRF9WT was able to robustly coimmunoprecipitate with STAT2 (Fig. 4*A*, lane 2), the interaction was reduced to background levels with the IRF9 IF1-A, IF1-B, or the IF1-C mutants (Fig. 4*A*, lanes 3–5). IRF9 IF1-D, a variant containing the mutations L233A, L274A, and F283A, also did not bind STAT2 (Fig. 4*A*, lane 6). Mutation of STAT2 also interfered with IRF9 binding: IRF9WT readily coimmunoprecipitated STAT2WT (Fig. 4*B*, lane 2), but binding of STAT2 carrying a F174D mutation was greatly diminished (Fig. 4*B*, lane 3). Thus, mutation of both the IRF9 and the STAT2 side of the interface disrupts the interaction between full-length STAT2 and IRF9 in vivo. We therefore conclude that the identified binding interface is necessary and sufficient for this interaction.

Next, we tested the impact of the mutations on ISGF3 activity using an *IFIT1*promoter luciferase reporter (*IFIT1*prom-Luc) in $IRF9^{-/-}$/$IRF\text{-}3^{-/-}$ mouse embryonic fibroblasts (MEFs). The luciferase reporter assay showed IFNβ induced *IFIT1*prom-Luc activity in cells reconstituted with FLAG-IRF9WT (Fig. 4*B* and Fig. S4*A*). However, cells expressing the IRF9 IF1 mutants exhibited only partial (IF1-A and IF1-B) or no (IF1-C) *IFIT1*prom-Luc activity in response to IFNβ. The progressive loss of activity observed with IRF9 mutants, IF1-A > IF1-B > IF1-C, correlates with our in vitro binding studies (Fig. 2*B*). Not surprisingly, the IF1-D variant also did not show *IFIT1*prom-Luc activity (Fig. 4*B*). Overall, our results show that mutations to the conserved surface patch disrupt IRF9–STAT2 interaction and abolish ISGF3 function in cells.

Solution Structure of STAT2–IRF9 and Model of the ISGF3 Complex. We also examined the architecture of the STAT2–IRF9 complex in solution by cross-linking/mass spectrometry (XL–MS) and SEC in line with small-angle X-ray scattering (SEC–SAXS). We focused on a STAT2–IRF9 complex containing an extended $STAT2_{133–679}$ construct spanning from the CCD to the SH2 domain. XL–MS showed intermolecular cross-links between K183 in STAT2–CCD and K318 or K381 in IRF9–IAD, in agreement with our structure (Fig. S4*B*). The SAXS scattering curve allowed us to determine the radius of gyration (R_g) of 4.37 ± 0.02 nm, while the distance distribution function *p(r)* showed a curve characteristic of an elongated particle with a maximum diameter of 14.4 ± 0.3 nm (Fig. S4 *C* and *D* and *Inset*). These data allowed rigid body modeling of $STAT2_{133–679}$ spanning from CCD to the SH2 domain in complex with the IRF9–IAD (Fig. S4*B*). The resulting model represented a good fit to the experimental data (Fig. S4*C*). Based on this model, together with available structures of DNA-bound STAT1 and IRF dimers, we propose a composite model for ISGF3, containing a 1:1:1 complex of STAT1:STAT2 and IRF9 bound to the ISRE DNA element of the IFN-inducible adenosine deaminase ADAR1 gene (Fig. 4*C*) (33). In this model, the IRF9–IAD is attached to the STAT2–CCD and is linked by a flexible linker spanning residues 112–198 to the IRF9-DBD. This linker length is adequate to span the predicted distance between the IRF9–IAD and the IRF9-DBD. The ISRE contains two 5′-GAAA-3′ motifs, typical binding elements for IRFs (9) and potential IRF9 binding sites. To position the IRF9-DBD on the ISRE, we performed a series of electrophoretic mobility shift assay (EMSA) experiments and found that a single copy of the IRF9-DBD bound to the DNA substrate (Fig. S5*A*). DNA binding of IRF9-DBD became progressively weaker on ISRE substrates in which the 3′ extension, proximal to the second 5′-GAAA-3′

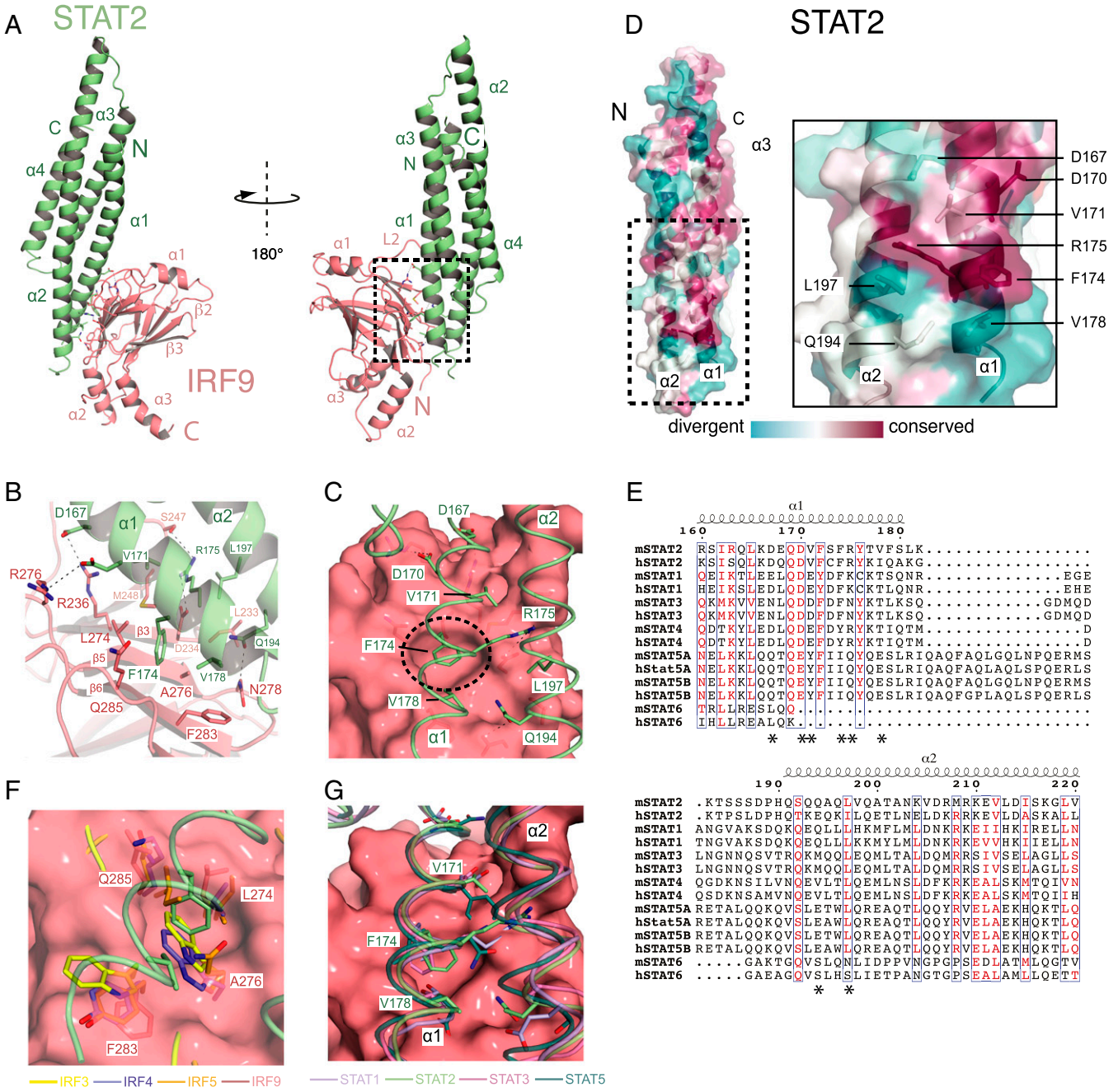

Fig. 3. Structure of the STAT2–IRF9 complex. (*A*) STAT2 (green) and IRF9 (red) are shown as ribbons. The model is rotated 180° between *Left* and *Right*. (*B*) Expanded view of the STAT2:IRF9 interface. Residues in the interface are shown with nitrogen in blue and oxygen in red. Carbon atoms are colored according to residue location. (*C*) Surface drawing of IRF9 with the recognition helices α1 and α2 of STAT2 shown as coils. Residues of STAT2 in the interface are indicated and the key anchoring residue F174 is circled. (*D*) The STAT2–CCD is colored according to amino acid conservation among vertebrate STAT2 orthologs. Magenta, well conserved; light blue, highly variable. Also shown is a zoom-in panel of the conserved surface patch showing the underlying amino acid residues. (*E*) Sequence alignment of mouse and human STAT family members. Residues of STAT2 involved in IRF9 binding are indicated (*). Columns containing residues colored in red indicate positions with a similarity score >0.7 calculated using ESPript, espript.ibcp.fr (56). (*F*) Surface representation of IRF9 superposed with IRF3, -4, and -5. A276 of IRF9, which is critical for accommodation of F174 of STAT2, is replaced by bulky side chains in IRF3, -4, and -5. STAT2 is represented as coils. (*G*) Structural overlay of STAT1, STAT3, and STAT5 onto STAT2. The secondary anchoring residues V171 and V178 are replaced by charged or bulky residues in other STATs.

repeat, was systematically shortened (Fig. S5*A*). This second but not the first 5′-GAAA-3′ motif also contains an additional 5′ flanking AA sequence that is usually recognized through a minor groove contact by a conserved Histidine residue (H44 in mIRF9) and located 2 bp upstream from the core 5′-GAAA-3′ motif (34, 35). We therefore propose that the second 5′-GAAA-3′ motif of the ADAR ISRE contains a single high-affinity binding site for IRF9 and modeled the IRF9-DBD on this position. Next, we purified a STAT1:STAT2 heterodimer and assayed its ability to bind to the ISRE DNA. We observed a single high-affinity DNA binding event (Fig. S5*C*), in agreement with the model that a single STAT1:STAT2 heterodimer binds to the ISRE (28).

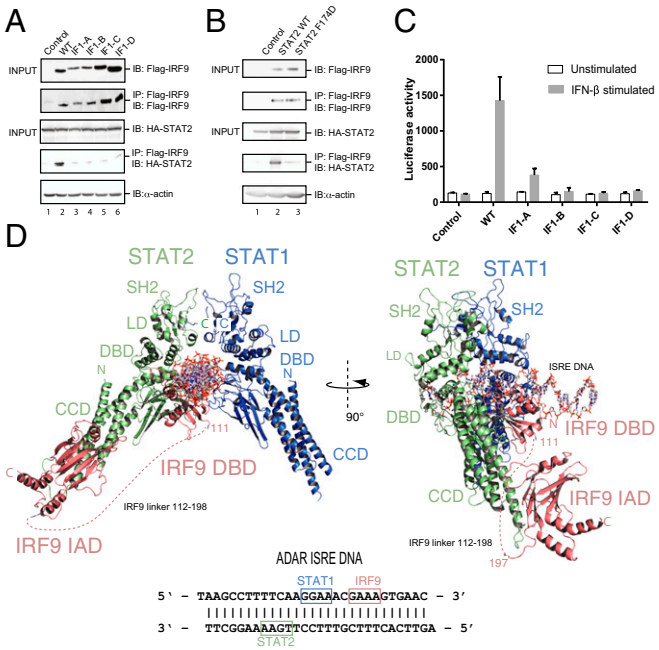

Fig. 4. Effects of point mutations in the STAT2–IRF9 interface on ISGF3 activity. (*A*) Coimmunoprecipitation of IRF9 variants with STAT2: HEK293 cells were cotransfected with FLAG-IRF9 WT or IF1 mutants and HA-STAT2. The empty FLAG vector was used as a control. Immunoprecipitation (IP) was performed with anti-FLAG beads and immunoblotted (IB) for FLAG_IRF9 or HA-STAT2. Actin concentration was used as a loading control for INPUT conditions. Three independent experiments were performed with consistency, and one representative example is shown. (*B*) IP of FLAG IRF9 WT with HA-STAT2 WT or with HA-STAT2 F174D. (*C*) MEF IRF9$^{-/-}$/IRF–3$^{-/-}$ cells were cotransfected with the *IFIT1*prom-*Firefly* luciferase reporter and Renilla luciferase plasmids along with the indicated Flag-IRF9 expression plasmids. Cells were treated with 200 U/mL IFNβ before assaying for luciferase activities expressed as *Firefly*/*Renilla* ratio. Three independent experiments were performed and the mean value ± SEM are shown. (*D*) Model of the ISGF3 complex bound to the ADAR ISRE DNA sequence. A STAT1 dimer bound to a GAS DNA sequence (1BG5) was positioned on the first 5′-GGAA-3′ repeat. A homology model of the IRF9 DBD was positioned on the second GAAA repeat based on the structure of DNA-bound IRF3 (1T2K). The STAT2–IRF9 complex was overlaid with one copy of STAT1 of the STAT1 dimer to obtain the final model. A dashed line indicates the flexible amino acid linker spanning residues 112–198 between the IAD and DBD of IRF9. CCD, coiled-coil domain; IAD, IRF association domain; LD, ligand-binding domain.

STAT-binding sites frequently do not show a sharply defined consensus sequence, but in agreement with DNA selection studies for STAT1 (36) and XL studies on a related ISRE (28), we propose that STAT1 binds to the consensus 5′-GGAA-3′ motif. STAT2 is then positioned upstream on a nonconsensus 5′-TGAA-3′ motif, in agreement with the model that STAT2 does not contribute directly to DNA binding specificity (28). The more generic DNA interaction of STAT2 could be stabilized through the interaction with the IAD domain of IRF9.

Discussion

STATs and IRFs have arisen by gene duplication, evolutionary drift, and natural selection under constant selective pressure from pathogen infection (13, 37, 38). The direct interaction between STAT2 and IRF9 is critical for the function of ISGF3 and the antiviral response (23). Our studies show that the overall architecture of these domains is similar to that of other STAT and IRF paralogs. However, there are several important adaptations: The IRF9–IAD is missing the regulatory apparatus that is used for IRF autoinhibition in the latent form and that in the activated state enables IRF dimerization and interaction with the transcriptional coactivators CBP/p300 (15, 17). In addition, the PRR adaptor-binding site that enables PRR signal-dependent activation of IRF3 is absent in IRF9. IRF9 binds to the tip of the STAT2–CCD using the convex surface of the β-sandwich core of the IAD domain. While the same surface is available in other IRF paralogs, amino acid substitutions at the key anchoring points account for the preferential STAT2 binding. Together, these adaptations explain why IRF9 binds constitutively and selectively to STAT2 (23). IRFs interact with numerous other transcription factors, and we predict that their IAD domains have evolved additional protein-binding interfaces. For example, IRF4 uses another surface of the IAD to interact with the PU.1 transcription factor (29).

The STAT2–CCD is closely related to that of other STATs, with the exception of STAT6, which is missing the IRF9 binding site entirely. The lead-anchoring residue of STAT2, F174, is conserved in STAT1 and STAT3. Amino acid residues at this position are critical for stabilizing an "antiparallel" and apparently inactive dimer structure of STAT1 (F172) and STAT5 (I174) and thus play an important role in the STAT activation and inactivation cycle (39, 40) (Fig. 3*E*). As IRF9 binds to this interface, it is not possible that STAT2 adopts such an antiparallel conformation when bound to IRF9.

The secondary anchoring residues V171 and V178 of STAT2 are divergent. Due to this divergence, STAT1 binds with 500-fold weaker affinity to IRF9 compared with STAT2, contributing to differential binding energetics. Previous studies have reported an interaction between STAT1 and IRF9 under certain experimental conditions (8), which is likely due to this residual binding affinity between STAT1 and IRF9. However, other experiments could not reproduce these results (23). Nevertheless, STAT1 and IRF9 appear to interact functionally on certain promoters even in the absence of STAT2 (41, 42). Overall, our data are in agreement with the model that IRF9 selectively and constitutively interacts with STAT2. Further work is required to investigate the physiological relevance of a potential interaction with STAT1.

Together, functional divergence of the STAT and IRF paralogs is due to amino acid substitutions at key anchoring points in the binding interface that enable select family members (STAT2–IRF9) to interact with high affinity and thereby restrict interaction between other STAT and IRF paralogs. Our study thus provides evidence for how gene duplication and divergence can contribute to the evolution of tightly integrated systems by slight structural variation of key amino acids of primordial variants. These variations resulted in new protein–protein interactions that enable a significant increase in the regulatory repertoire of the mammalian cytokine response.

Methods

Constructs, Expression, and Purification. Sequences encoding murine IRF9, STAT1, and STAT2 and any variants described in the text were amplified by PCR from cDNA clones and inserted between the *NcoI* and *HindIII* sites in the first ORF of the pETDuet vector. Coexpression was achieved by cloning respective cDNA into the second ORF between *NdeI* and *KpnI* sites. Mutants of IRF9 and STAT2 were generated using the QuikChange site-directed mutagenesis kit (Agilent). Sequences of all expression constructs were confirmed by DNA sequencing. Individual proteins and protein complexes were expressed in *Escherichia coli* BL21 (DE3) cells and induced with IPTG. Upon reaching an $OD_{600\ nm}$ of 0.6, *E. coli* cultures were shifted from a temperature of 37 °C to 16 °C, for 12 h. Cells were pelleted at 9,000 *g* (JLA–8.1; Beckman) followed by resuspension in buffer A [20 mM Hepes (pH 7.5), 300 mM NaCl, and 0.5 mM TCEP] containing 10 mM imidazole supplemented with protease inhibitors (Roche) and lysed using a microfluidizer (Microfluidics). Lysates were centrifuged at 24,000 *g* for 30 min, using a JA–25.5 rotor (Beckman), and the resultant supernatant passed over Ni^{2+}-conjugated IMAC Sepharose resin (GE Healthcare). Columns were subsequently washed using buffer A containing 20 mM imidazole and eluted with buffer A containing 500 mM imidazole. For all purifications involving His-TEV proteins, the Ni^{2+} eluate was incubated overnight at 4 °C with His-tagged TEV protease during dialysis against buffer A with 20 mM imidazole. Cleaved tags, TEV protease, and uncleaved protein were removed by subtractive purification over Ni^{2+}

resin. All proteins were further purified by SEC using a HiLoad 16/60 Superdex 200 prep-grade column (GE Healthcare) equilibrated in buffer A.

Limited Proteolysis. To identify stable variants of STAT2 and IRF9, we performed limited proteolysis by Trypsin and α-chymotrypsin. The purified $STAT2_{133–738}$–$IRF9_{182–399}$ complex (120 μg) was titrated against decreasing concentrations of Trypsin ranging from a ratio of 1:125–1:1,000 (wt/wt) and incubated for up to 30 min at 4 °C. The samples were then analyzed by SDS–PAGE, acid hydrolysis, and MS. The stable $STAT2_{133–679}$–$IRF9_{182–399}$ (120 μg) complex identified by Trypsin digestion was further proteolyzed by α-chymotrypsin for 72 h at a ratio of 1:5,000 (wt/wt) at 4 °C. The fragments were analyzed as above.

Crystallization. Initial trails of $IRF9_{182–385}$ yielded crystals with inherent pathologies like twinning, translational pseudosymmetry, and high copy number. To overcome these problems, we mutated surface residues to induce alternate crystal packing. An E348A mutant yielded crystals in 0.1 M Hepes (pH 6.8), 1.5 M ammonium phosphate monobasic, and 0.1 M ammonium sulfate that diffracted to a minimum Bragg spacing of 1.9 Å. Production of diffraction quality crystals of the STAT2–IRF9 complex also required mutation of surface residues Q242A and K243A of STAT2 and of E347A and E348A of IRF9. Crystals of the STAT2–IRF9 complex (10 $mg{\cdot}mL^{-1}$) were obtained in 0.2 M potassium formate and 20% PEG3350 and appeared within 3 d after setup. These were further optimized by microseeding, yielding crystals that diffracted to 2.7 Å resolution. We used 25% glycerol as a cryoprotectant before flash cooling in liquid nitrogen. IRF9–IAD diffraction data were collected at 100 K at an X-ray wavelength of 0.966 Å at beamline ID30A-1 (MASSIF-1) of the European Synchrotron Radiation Facility (ESRF) with a Pilatus 6M-F detector (43). Diffraction data for crystals of the STAT2–IRF9 complex were collected at a wavelength of 0.972 Å at the beamline ID23-1 of ESRF with a Pilatus 6M-F detector. Indexing and scaling of the data were performed with XDS and XSCALE (44). The structure was solved by molecular replacement using the IRF5–IAD [Protein Data Bank (PDB) ID code 3DSH] structure as a search model (45). STAT2–IRF9 data were indexed with XDS and scaled using AIMLESS (46). The structure of the IRF9–IAD and STAT1–CCD (PDB ID code 1BF5) were used as search models for molecular replacement. Final models were produced by iterative rounds of automatic and manual model building and refinement, using Coot and PHENIX (47, 48). The final IRF9–IAD model contained residues 197–385 and was refined to a resolution of 1.9 Å with an R_{work} and an R_{free} of 20.6% and 24.2%, respectively (Table 1). Analysis of the refined structure in MolProbity showed that there are no residues in disallowed regions of the Ramachandran plot. The MolProbity all atom clash score was 4.7, placing the structure in the 97th (best) percentile of structures ($n = 773$) refined at comparable resolution (49). The final STAT2–IRF9 model contained residues 141–315 of STAT2 and 206–376 of IRF9 and was refined to a resolution of 2.9 Å with R_{work} and an R_{free} of 25.5% and 30.7%, respectively (Table 1). The MolProbity all atom clash score was 0.37, placing the structure in the 100th (best) percentile of structures ($n = 175$). Figures displaying the structures were generated using PyMol (50).

SEC–MALLS and SAXS Analysis. SEC was performed at 20 °C with a Superdex 200 10/300 GL column (GE Healthcare) equilibrated in buffer A. We injected 50 μL of $STAT2_{133–679}$/IRF9–IAD at 7 mg/mL, and the sample eluted at a flow rate of 0.5 mL/min. MALLS was recorded with a laser emitting at 690 nm using a DAWN-EOS detector (Wyatt Technology Corp.). The refractive index was measured using a RI2000 detector (Schambeck SFD). Data analysis was performed with the ASTRA software (Wyatt Technology Corp.). The averaged molecular mass represents the measurements across the elution peak.

X-ray scattering data were collected using an inline HPLC setup at the BioSAXS beamline (BM29) of the European Synchrotron Radiation Facility. Inline SEC was performed at a temperature of 10 °C using a Superdex Increase 200 10/300 GL column equilibrated in SEC buffer. Data were collected with a photon-counting Pilatus 1M detector at a sample–detector distance of 2.86 m, a wavelength of $\lambda = 0.991$ Å, and an exposure time of 1 s per frame. A momentum transfer range of 0.008–0.47 $Å^{-1}$ was covered ($q = 4\pi \sin\theta/\lambda$, where θ is the scattering angle and λ the X-ray wavelength). Data collected across the peak were subtracted from buffer scattering, and the frames 1,904–2,264 showing a constant R_g were merged for further analysis. R_g values were obtained from the Guinier approximation $sR_g < 1.3$ using Primus (51). Distance distribution functions *p(r)* and the Porod volumes Vp were computed from the entire scattering curve using GNOM (51). CORAL from the ATSAS suite was used to model the $STAT2_{133–679}$–IRF9–IAD complex using the $STAT2_{133–679}$:IRF9–IAD homology model and the IRF9–IAD structure as the input files. The final model conforms well to the scattering data, with a $\chi^2 = 1.64$. The model for the ISGF3 complex bound to a ISRE DNA sequence 5′-GGGAAATGGAAACT-3′ was obtained by positioning a STAT1 dimer bound to a GAS DNA sequence (PDB ID code 1BG5) on the first 5′-GGAA-3′ repeat. A homology model of the IRF9-DBD was positioned on the second GAAA repeat based on the structure of DNA-bound IRF3 (PDB ID code 1T2K). The final model was obtained by overlaying the STAT2–IRF9 complex onto the distal copy of the STAT1 dimer.

XL–MS. XL was performed using the $STAT2_{133–679}$/IRF9–IAD complex by incubating with isotope-labeled disuccinimidyl suberate (DSS) as described previously (52). Protein digestion was performed at 37 °C with LysC for 4 h followed by Trypsin digestion overnight; digested peptides were enriched by SEC. Fractions were injected onto a nanoAcquity ultraperformance LC column connected to a LTQ Orbitrap Velos Pro instrument (Thermo Scientific) for LC-based MS measurements. Data processing was performed using xQuest/xProphet. Identified cross-links were mapped onto the SAXS model of the $STAT2_{133–679}$/IRF9–IAD heterodimer and analyzed using Xlink analyzer (53).

ITC. Proteins were extensively dialyzed against ITC buffer [20 mM Hepes (pH 7.5), 300 mM NaCl, 2% glycerol] and subjected to calorimetry using a MicroCal ITC200 system (Malvern Instruments). $STAT2_{133–679}$ at 30 μM was titrated with different IRF9–IAD variants at concentrations ranging from 150 μM to 320 μM. For IRF9–IADWT, titrations were carried out by injection of 1.5 μL of IRF9–IAD (ITC200) every 180 s into the sample cell containing $STAT2_{133–679}$ variants. For mutant IRF9–IAD, 16 successive injections of 2.5 μL were done every 240 s. $STAT1_{136–713}$ variants or $STAT2_{133–679}$ V171E were titrated against IRF9–IAD every 120 s for 20 injections. Enthalpy change data for titrations were double background-corrected via subtraction of protein into buffer measurement. The data were fit using MicroCal Origin 7.0 software (OriginLab).

Cell Culture. All media and supplements were from Gibco, except where indicated. $IRF9^{-/-}$/IRF–$3^{-/-}$ MEFs, kindly provided by K. Mossman, McMaster University, Hamilton, Canada, were immortalized using the 3T3 protocol and cultured in MEM medium supplemented with nonessential amino acid, sodium pyruvate, 1% L-glutamine, and 10% heat-inactivated FBS (HI-FBS). HEK293 cells (ATCC) were cultured in DMEM containing 1% L-glutamine and 10% Fetalclone III (HyClone). All cultures were performed without antibiotics and controlled for the absence of mycoplasma contamination using the MycoAlert *Mycoplasma* Detection Kit (Lonza).

Luciferase Reporter Assays. MEF cells were cotransfected with the pRL-null Renilla (*Renilla* luciferase, internal control), the *IFIT1*prom-pGL3 *firefly* luciferase reporter (54), and the indicated FLAG-tagged IRF9WT or mutant expression plasmid using the *Trans*IT-LT1 transfection reagent (Mirus). At 8 h posttransfection, cells were stimulated for 16 h with 200 U/mL murine IFNβ (PBL Assay Science). Luciferase activities were quantified using the dual luciferase reporter assay kit (Promega). Relative luciferase activities were calculated as the *firefly* luciferase/*Renilla* ratio. Protein extracts were subjected to SDS–PAGE electrophoresis and analyzed by immunoblot using the anti-Flag M2 (F1804; Sigma-Aldrich) and anti-actin (A5441; Sigma-Aldrich) antibodies. Immunoreactive bands were visualized using the Western Lightning Chemiluminescence Reagent Plus (Perkin-Elmer Life Sciences) acquired on an ImageQuant LAS 4000mini apparatus (GE Healthcare).

Coimmunoprecipitation. HEK293 cells were transfected with FLAG-tagged IRF9WT or mutants together with HA-tagged STAT2WT or F174D mutant encoding plasmids using the calcium phosphate method. Cells were lysed by sonication in 50 mM Hepes (pH 7.4), 150 mM NaCl, 5 mM EDTA, 10% glycerol, 1% triton, 10 μg/mL aprotinin, 10 μg/mL leupeptin, 5 mM NaF, 1 mM activated Na_3VO_4, 2 mM p-nitrophenyl phosphate, and 10 mM β-glycerophosphate (pH 7.5). Cell lysates (1 mg) were subjected to immunoprecipitation using 2 μg anti-FLAG M2 antibodies (F1804; Sigma-Aldrich) for 3 h at 4 °C. Elution of immunocomplexes was performed by incubation on ice for 2 h in lysis buffer containing 100 μg/mL FLAG-peptide (F3290; Sigma-Aldrich). Immunocomplexes were analyzed by immunoblot using the anti-FLAG M2, anti-HA (ab9110; Abcam), and anti-actin (A5441; Sigma-Aldrich) antibodies as described above.

Purification of the STAT1–STAT2 Complex. STAT1 (136–713) with an N-terminal His tag containing a cleavable site for Tobacco Etch Virus (TEV) protease and STAT2 (133–748) containing a C-terminal intein–chitin binding fusion domain were cloned into the ORFs 1 and 2 of the pETDuet-1 vector, respectively. STAT1 and STAT2 were coexpressed with an Elk receptor tyrosine kinase domain in *E. coli* BL21(DE3) TKB1 cells (Agilent) (55). Cells were harvested by centrifugation at 9,000 *g* and were resuspended in buffer A containing 10 mM imidazole. Subsequently the cells were lysed in a microfluidiser

(Microfluidics Corp.), and the soluble fraction was obtained by centrifugation at 24,000 *g* for 30 min. The supernatant was first passed over Ni^{2+}-conjugated IMAC Sepharose resin (GE Healthcare) and washed with buffer A containing 20 mM imidazole. The protein was eluted in buffer A containing 500 mM imidazole and loaded onto chitin resin following instructions by the manufacturer (New England Biolabs). The column was washed in buffer A and incubated at room temperature overnight in buffer A containing 50 mM DTT. The cleaved STAT1–STAT2 heterodimer was eluted in the same buffer and further purified by SEC in buffer A.

EMSA. Analysis of DNA binding by STAT1–STAT2 and IRF9 was performed by incubating the indicated amounts of purified proteins with 5 μM of 23 bp, 25 bp, 27 bp, or 29 bp ISRE DNA in a buffer containing 20 mM Hepes (pH 7.5), 150 mM NaCl, 1 mM DTT, and 5% glycerol. The samples were incubated at room temperature for 5 min and analyzed using 6% polyacrylamide gels. The gels were stained with ethidium bromide and imaged by UV fluorescence.

Data Availability. Coordinates for the IRF9–IAD and the IRF9–STAT2 complex are available from PDB under ID codes 5OEM and 5OEN, respectively.

ACKNOWLEDGMENTS. We thank Joanna Kirkpatrick and the proteomic core facility at the European Molecular Biology Laboratory (EMBL) for processing and analysis of cross-linked samples. We thank the Partnership for Structural Biology (Grenoble) for providing access to their biophysical platform. We thank Dr. K. Mossman (McMaster University) for the MEF cells used in this study and Dr. Thomas Decker (Max F. Perutz Laboratories, University of Vienna) for comments on the manuscript. We thank the staff at ESRF Beamlines BM29 and ID30a-1,3 (Massively Automated Sample Selection Integrated Facility 1) for their support during data collection. This work was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Grant 355306 (to N.G.). S.R. was supported by a postdoctoral fellowship from the Fondation Association pour la Recherche sur le Cancer and has received support from the Fondation Innovations en Infectiologie. N.G. is a recipient of a Research Chair in signaling in virus infection and oncogenesis from Université de Montréal.

1. Stark GR, Darnell JE, Jr (2012) The JAK-STAT pathway at twenty. *Immunity* 36: 503–514.
2. Levy DE, Darnell JE, Jr (2002) Stats: Transcriptional control and biological impact. *Nat Rev Mol Cell Biol* 3:651–662.
3. Paul WE, Seder RA (1994) Lymphocyte responses and cytokines. *Cell* 76:241–251.
4. Kishimoto T, Taga T, Akira S (1994) Cytokine signal transduction. *Cell* 76:253–262.
5. Schindler C, Shuai K, Prezioso VR, Darnell JE, Jr (1992) Interferon-dependent tyrosine phosphorylation of a latent cytoplasmic transcription factor. *Science* 257:809–813.
6. Levy DE, Kessler DS, Pine R, Reich N, Darnell JE, Jr (1988) Interferon-induced nuclear factors that bind a shared promoter element correlate with positive and negative transcriptional control. *Genes Dev* 2:383–393.
7. Fu XY, Kessler DS, Veals SA, Levy DE, Darnell JE, Jr (1990) ISGF3, the transcriptional activator induced by interferon alpha, consists of multiple interacting polypeptide chains. *Proc Natl Acad Sci USA* 87:8555–8559.
8. Horvath CM, Stark GR, Kerr IM, Darnell JE, Jr (1996) Interactions between STAT and non-STAT proteins in the interferon-stimulated gene factor 3 transcription complex. *Mol Cell Biol* 16:6957–6964.
9. Escalante CR, Yie J, Thanos D, Aggarwal AK (1998) Structure of IRF-1 with bound DNA reveals determinants of interferon regulation. *Nature* 391:103–106.
10. Furui J, et al. (1998) Solution structure of the IRF-2 DNA-binding domain: A novel subgroup of the winged helix-turn-helix family. *Structure* 6:491–500.
11. Takahasi K, et al. (2003) X-ray crystal structure of IRF-3 and its functional implications. *Nat Struct Biol* 10:922–927.
12. Panne D, Maniatis T, Harrison SC (2007) An atomic model of the interferon-beta enhanceosome. *Cell* 129:1111–1123.
13. Yanai H, Negishi H, Taniguchi T (2012) The IRF family of transcription factors: Inception, impact and implications in oncogenesis. *OncoImmunology* 1:1376–1386.
14. Reich N, et al. (1987) Interferon-induced transcription of a gene encoding a 15-kDa protein depends on an upstream enhancer element. *Proc Natl Acad Sci USA* 84: 6394–6398.
15. Qin BY, et al. (2003) Crystal structure of IRF-3 reveals mechanism of autoinhibition and virus-induced phosphoactivation. *Nat Struct Biol* 10:913–921.
16. Hiscott J, Lin R (2005) IRF-3 releases its inhibitions. *Structure* 13:1235–1236.
17. Chen W, et al. (2008) Insights into interferon regulatory factor activation from the crystal structure of dimeric IRF5. *Nat Struct Mol Biol* 15:1213–1220.
18. Takeuchi O, Akira S (2010) Pattern recognition receptors and inflammation. *Cell* 140: 805–820.
19. Ablasser A, et al. (2013) cGAS produces a 2′-5′-linked cyclic dinucleotide second messenger that activates STING. *Nature* 498:380–384.
20. Kato H, Takahasi K, Fujita T (2011) RIG-I-like receptors: Cytoplasmic sensors for non-self RNA. *Immunol Rev* 243:91–98.
21. Zhao B, et al. (2016) Structural basis for concerted recruitment and activation of IRF-3 by innate immune adaptor proteins. *Proc Natl Acad Sci USA* 113:E3403–E3412.
22. Banninger G, Reich NC (2004) STAT2 nuclear trafficking. *J Biol Chem* 279:39199–39206.
23. Martinez-Moczygemba M, Gutch MJ, French DL, Reich NC (1997) Distinct STAT structure promotes interaction of STAT2 with the p48 subunit of the interferon-alpha-stimulated transcription factor ISGF3. *J Biol Chem* 272:20070–20076.
24. Kraus TA, Lau JF, Parisien JP, Horvath CM (2003) A hybrid IRF9-STAT2 protein recapitulates interferon-stimulated gene expression and antiviral response. *J Biol Chem* 278:13033–13038.
25. Lau JF, Parisien JP, Horvath CM (2000) Interferon regulatory factor subcellular localization is determined by a bipartite nuclear localization signal in the DNA-binding domain and interaction with cytoplasmic retention factors. *Proc Natl Acad Sci USA* 97: 7278–7283.
26. Improta T, et al. (1994) Transcription factor ISGF-3 formation requires phosphorylated Stat91 protein, but Stat113 protein is phosphorylated independently of Stat91 protein. *Proc Natl Acad Sci USA* 91:4776–4780.
27. Veals SA, Santa Maria T, Levy DE (1993) Two domains of ISGF3 gamma that mediate protein-DNA and protein-protein interactions during transcription factor assembly contribute to DNA-binding specificity. *Mol Cell Biol* 13:196–206.
28. Qureshi SA, Salditt-Georgieff M, Darnell JE, Jr (1995) Tyrosine-phosphorylated Stat1 and Stat2 plus a 48-kDa protein all contact DNA in forming interferon-stimulated-gene factor 3. *Proc Natl Acad Sci USA* 92:3829–3833.
29. Remesh SG, Santosh V, Escalante CR (2015) Structural studies of IRF4 reveal a flexible autoinhibitory region and a compact linker domain. *J Biol Chem* 290:27779–27790.
30. Panne D, McWhirter SM, Maniatis T, Harrison SC (2007) Interferon regulatory factor 3 is regulated by a dual phosphorylation-dependent switch. *J Biol Chem* 282: 22816–22822.
31. Bluyssen HA, et al. (1995) Combinatorial association and abundance of components of interferon-stimulated gene factor 3 dictate the selectivity of interferon responses. *Proc Natl Acad Sci USA* 92:5645–5649.
32. Chowdhury FZ, Farrar JD (2013) STAT2: A shape-shifting anti-viral super STAT. *JAK-STAT* 2:e23633.
33. Stark GR, Kerr IM, Williams BR, Silverman RH, Schreiber RD (1998) How cells respond to interferons. *Annu Rev Biochem* 67:227–264.
34. Panne D, Maniatis T, Harrison SC (2004) Crystal structure of ATF-2/c-Jun and IRF-3 bound to the interferon-beta enhancer. *EMBO J* 23:4384–4393.
35. Fujii Y, et al. (1999) Crystal structure of an IRF-DNA complex reveals novel DNA recognition and cooperative binding to a tandem repeat of core sequences. *EMBO J* 18: 5028–5041.
36. Horvath CM, Wen Z, Darnell JE, Jr (1995) A STAT protein domain that determines DNA sequence recognition suggests a novel DNA-binding domain. *Genes Dev* 9:984–994.
37. Liongue C, O'Sullivan LA, Trengove MC, Ward AC (2012) Evolution of JAK-STAT pathway components: Mechanisms and role in immune system development. *PLoS One* 7:e32777.
38. Randall RE, Goodbourn S (2008) Interferons and viruses: An interplay between induction, signalling, antiviral responses and virus countermeasures. *J Gen Virol* 89: 1–47.
39. Mao X, et al. (2005) Structural bases of unphosphorylated STAT1 association and receptor binding. *Mol Cell* 17:761–771.
40. Neculai D, et al. (2005) Structure of the unphosphorylated STAT5a dimer. *J Biol Chem* 280:40782–40787.
41. Majumder S, et al. (1998) p48/STAT-1alpha-containing complexes play a predominant role in induction of IFN-gamma-inducible protein, 10 kDa (IP-10) by IFN-gamma alone or in synergy with TNF-alpha. *J Immunol* 161:4736–4744.
42. Rauch I, et al. (2015) Noncanonical effects of IRF9 in intestinal inflammation: More than type I and type III interferons. *Mol Cell Biol* 35:2332–2343.
43. Svensson O, Malbet-Monaco S, Popov A, Nurizzo D, Bowler MW (2015) Fully automatic characterization and data collection from crystals of biological macromolecules. *Acta Crystallogr D Biol Crystallogr* 71:1757–1767.
44. Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* 66:125–132.
45. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40: 658–674.
46. Winn MD, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67:235–242.
47. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132.
48. Adams PD, et al. (2011) The Phenix software for automated determination of macromolecular structures. *Methods* 55:94–106.
49. Chen VB, et al. (2010) MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66:12–21.
50. DeLano WL (2002) *The PyMOL Molecular Graphics System* (DeLano Scientific, San Carlos, CA).
51. Petoukhov MV, et al. (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Crystallogr* 45:342–350.
52. Herzog F, et al. (2012) Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. *Science* 337:1348–1352.
53. Kosinski J, et al. (2015) Xlink analyzer: Software for analysis and visualization of cross-linking data in the context of three-dimensional structures. *J Struct Biol* 189:177–183.
54. Grandvaux N, et al. (2002) Transcriptional profiling of interferon regulatory factor 3 target genes: Direct involvement in the regulation of interferon-stimulated genes. *J Virol* 76:5532–5539.
55. Becker S, Corthals GL, Aebersold R, Groner B, Müller CW (1998) Expression of a tyrosine phosphorylated, DNA binding Stat3beta dimer in bacteria. *FEBS Lett* 441: 141–147.
56. Robert X, Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* 42:W320–W324.