



HOKKAIDO UNIVERSITY

Title	Structural analysis reveals how tetrameric tyrosine-phosphorylated STAT1 is targeted by the rabies virus P-protein
Author(s)	Sugiyama, Aoi; Minami, Miku; Ugajin, Kaito et al.
Citation	Science signaling, 18(878) https://doi.org/10.1126/scisignal.ads2210
Issue Date	2025-03-18
Doc URL	https://hdl.handle.net/2115/95216
Rights	This is the author's version of the work. It is posted here by permission of the AAAS for personal use, not for redistribution. The definitive version was published in Science signaling on 2025-03-18; DOI: 10.1126/scisignal.ads2210.
Type	journal article
File Information	scisignal.ads2210(6).pdf



Structural analysis reveals how tetrameric tyrosine-phosphorylated STAT1 is targeted by the rabies virus P-protein

Aoi Sugiyama^{1†}, Miku Minami^{1†}, Kaito Ugajin¹, Satomi Inaba-Inoue¹, Nana Yabuno¹, Yuichiro Takekawa¹, Sun Xiaomei¹, Shiho Takei¹, Mina Sasaki², Tomo Nomai², Xinxin Jiang², Shunsuke Kita², Katsumi Maenaka², Mika Hirose³, Min Yao¹, Paul R. Gooley⁴, Gregory W. Moseley⁵, Yukihiro Sugita⁶, Toyoyuki Ose^{1*}

¹Faculty of Advanced Life Science, Hokkaido University, Sapporo, 060-0810 Japan. ²Laboratory of Biomolecular Science, and Center for Research and Education on Drug Discovery, Faculty of Pharmaceutical Sciences, Hokkaido University, Sapporo, Japan. ³Institute for Protein Research, Osaka University, Osaka, 565-0871, Japan. ⁴Department of Biochemistry and Pharmacology, University of Melbourne, Parkville, Victoria 3010, Australia and Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, Victoria 3010, Australia. ⁵Department of Microbiology, Biomedicine Discovery Institute, Monash University, Clayton Campus, Victoria 3800, Australia. ⁶Laboratory of Ultrastructural Virology, Institute for Life and Medical Sciences, Hakubi Center for Advanced Research, Kyoto University, 53 Shogoin Kawahara-cho, Sakyo-ku, Kyoto, 606-8507, Japan.

†These authors contributed equally to this work.

*Corresponding author. Email: ose.toyoyuki@sci.hokudai.ac.jp

Abstract

Members of the signal transducer and activator of transcription (STAT) family are central mediators of signaling in the Janus kinase (JAK)-STAT pathway. STAT molecules are activated by phosphorylation at a conserved tyrosine residue, which is followed by dimerization through the reciprocal interaction between the phosphotyrosine and a Src homology 2 (SH2) domain. Tyrosine-phosphorylated STAT (pY-STAT) then translocates to the nucleus to induce the expression of genes encoding antiviral proteins; thus, it is a target for certain viral proteins that mediate immune evasion. The active and functional forms of STATs are conventionally considered to be dimers, which is based on the previously reported structure of a pY-STAT1 dimer bound to DNA, which comprised only the pY-STAT1 core, lacking the N-terminal domain (NTD) and C-

terminal region. Here, we present the cryo-electron microscopy (cryo-EM) structure of the tetrameric form of intact pY-STAT1 in complex with DNA. This full-length structure represents an important functional complex, because interactions formed between the N-terminal domains of STATs induce oligomerization, which is implicated in regulating transcriptional activity. The tetrameric structure revealed a compact conformation with a previously uncharacterized binding interface: two DNA-bound dimers are two-fold symmetrically aligned that can be seamlessly transformed into a tandem DNA-binding model without NTD dimer separation. Moreover, biochemical analyses indicated that the rabies virus P protein selectively targeted tetrameric pY-STAT1. Concomitant with data showing which regions contribute to the interaction between pY-STAT1 and the P protein, a binding model explaining how P recognizes the pY-STAT1 tetramer was constructed. These data provide insight into how pathogenic viruses regulate the signal transduction system related to the host immune system.

INTRODUCTION

The Janus kinase (JAK)–signal transducer and activator of transcription (STAT) pathway mediates signal transduction important in a diverse array of processes, including immune responses, cellular differentiation, and proliferation (*1–4*). Whereas more than 50 cytokines are implicated in using JAK-STAT pathways to regulate cell functions, there are only seven members of the STAT family in mammals (STAT1 to STAT4, STAT5a, STAT5b, and STAT6) (*5, 6*). STAT proteins exhibit a typical domain structure consisting of an N-terminal domain (NTD), a coiled-coil domain (CCD), a DNA-binding domain (DBD), a linker domain (LD), a Src homology 2 (SH2) domain, and a C-terminal transactivation domain (TAD), which contains a specific conserved tyrosine residue (for example, Tyr⁷⁰¹ in STAT1) that is phosphorylated by JAKs (Fig. 1A). The JAK-STAT system

mediates the transduction of signals from cell surface receptors to the nucleus; specifically, after recognition of a ligand (for example, a cytokine or growth factor) by a specific receptor, cytoplasmic receptor-associated JAKs undergo transphosphorylation and activation, before inducing tyrosine phosphorylation in the cytoplasmic region of the receptor. Phosphorylation of the receptor generates recognition sites for the recruitment of specific STATs, and JAKs then phosphorylate the conserved tyrosine in the STAT TAD (Fig. 1A); this phosphorylation is generally recognized as the activation signal for the STATs.

Tyrosine phosphorylation results in the formation of a "parallel STAT dimer" through reciprocal interactions between the phosphorylated tyrosine residue (pY) of one STAT and the SH2 domain of the other (Fig. 1A). The structure of a parallel dimer was established previously with protein crystallography (7, 8). The transition from an unphosphorylated "anti-parallel dimer" to the "parallel dimer" (9) produces a conformational nuclear localization signal and a specific DNA-binding site, which results in the import of the parallel dimer into the nucleus (10, 11) and the regulation of the transcription of target genes, respectively.

Importantly, because different receptors and JAKs can recruit and activate different STATs, different ligands can induce the formation of specific STAT complexes, resulting in differing signaling outcomes. The diversity of signaling outcomes is expanded by the capability of STATs to form both homo- and hetero-dimers through pY–SH2 domain interactions. This provides one explanation for how STAT proteins can transduce signals from approximately 50 different growth factors and cytokines, including interferons (IFNs), interleukins, and growth hormones, to produce distinct functional outcomes (5, 6, 12). Moreover, whereas pY-STAT dimers are conventionally

considered to be the active form, several STATs also form further interactions between their NTDs, with one co-localization-based assay indicating extensive NTD interactions between homodimers (for STAT1, STAT3, STAT4, and STAT5a) and heterodimers (for STAT1/2 and STAT5a/5b) (*13, 14*). These NTD interactions between unphosphorylated (U)-STATs and pY-STATs (homo- and heterodimers) may also contribute to structural and functional diversity.

STAT1 is a critical mediator of signal transduction in response to type I IFNs (typically, IFN- α and IFN- β), type II interferon (IFN- γ), and type III interferons (for example, IFN- λ 1). STAT1 signaling thus plays crucial roles in immune responses against viruses and other invading microorganisms. Activated STAT1 dimerizes with STAT1, STAT2, STAT3, and STAT5a, and STAT5b through the reciprocal pY-SH2 interaction (*12*). Types I and III IFNs principally stimulate the formation of STAT1 and STAT2 heterodimers, which form the IFN-stimulated gene factor 3 (ISGF3) complex with IFN-regulatory factor 9 (IRF9). ISGF3 activates particular arrays of genes by binding to genomic IFN-stimulated response elements (ISREs) within the gene promoters. Type II IFN stimulates the formation of STAT1 homodimers (gamma-activated factor, or GAF), which activates distinct gene profiles by binding to genomic gamma-activated sequence (GAS) elements.

pY-STAT1 homodimers also interact through their NTDs to form tetramers. NTD-mediated tetramerization is important for DNA binding because of cooperative binding to tandem binding sites such that tetramers have stronger binding affinities than that of pY-STAT1 dimers to cognate DNA with tandem binding sites. This effect was observed in assays in solution (*15*) and correlates with the enhanced expression of target genes that are regulated by promoters harboring a GAS and

an additional binding motif (for example, tandem binding) compared with promoters having only a single GAF-binding site (16). Furthermore, even for single GAF-binding sites in DNA, which are more common than multi-GAF-binding sites in mice, cooperative DNA binding as tetramers is essential for the type II IFN system-induced GAF (17). On the other hand, the role of STAT cooperativity in ISGF3 function remains controversial (17, 18), because tetramerization-incompetent STAT1 (mutated at Phe⁷⁷ and Leu⁷⁸ in the NTD) has little effect on IFN- α /ISGF3-dependent gene expression in mice (17), whereas STAT1 with the same mutation cannot activate ISGF3-targets, such as *OAS2* and *MXI*, in human cells (18). Structural data relating to cooperative DNA binding by STATs is limited. In particular, it is unknown how more than one pY-STAT dimer assembles to form an oligomer and how tetramerization affects DNA binding. The current tetramerization model for STATs is based only on the quantification of interactions using mutants (19, 20).

The critical role of JAK-STAT signaling in antiviral responses has resulted in the evolution by viruses of many mechanisms to inhibit JAK-STAT pathways. These mechanisms are largely mediated by viral proteins known as IFN antagonists, hundreds of which have been identified (21). Mechanisms to inhibit STAT1, STAT2, or both are relatively well-defined, and include cytoplasmic sequestration, dephosphorylation, degradation, inhibition of upstream signaling components, and induction of negative regulators (21). The relationship between viral proteins and STAT3 appears to be more complicated, with reported antagonistic and stimulatory outcomes depending on viral and cellular factors (22, 23). Because of a lack of structural studies on STAT-viral protein complexes, details of the molecular interactions between viral proteins and STATs have remained largely unknown. However, direct structural information has become available,

including for a 21-mer fragment of poxvirus protein 018 with the core domain of STAT1 (the CCD, DBD, LD, SH2 domain, and a part of the TAD), which indicated binding to the SH2 domain (24); non-structural protein 5 (NS5) of Zika virus or dengue virus bound to the NTD and the CCD of human STAT2, respectively (25); protein E27 from rat cytomegalovirus bound to the CCD of rat STAT2 (26); and the C-protein of Sendai virus in complex with the NTD of STAT1 (27). These studies emphasized the diverse nature of viral targeting of STAT molecules, presumably relating to the specific requirements of the viruses, such as differences in the replication cycle and cell-type/host tropism.

Antagonism of IFN-dependent JAK-STAT pathways by rabies virus (RABV) is achieved by the C-terminal domain (CTD) of P-protein (28, 29). RABV is the prototypical member of the lyssavirus genus (family *Rhabdoviridae*, order *Mononegavirales*), which includes several highly pathogenic viruses. Mononegaviruses have a nonsegmented, single-stranded RNA genome in the negative sense. In lyssaviruses, the genome contains five protein-encoding genes: N (nucleoprotein), P (phosphoprotein, P-protein), M (matrix protein), G (glycoprotein), and L (large protein, the RNA-dependent RNA polymerase) (30). RABV causes acute neurological disease in humans with an almost 100% fatality rate, resulting in more than 59,000 deaths every year across 150 countries (31, 32). The capacity to inhibit STAT signaling is a determining factor in pathogenicity (33). Furthermore, and in common with many other viral IFN antagonists, it was reported that P-protein interferes with both IFN induction and signaling (34, 35). To antagonize IFN signaling, P protein interacts directly with STAT1 through its P-CTD (36). Note that the efficient interaction of P protein with STAT1 in cells depends on activation of the cells by IFN- α , IFN- β , or IFN- γ , suggesting that P-CTD binds more efficiently to pY-STAT1 than to

nonphosphorylated STAT1 (28, 29, 33, 36, 37). However, the mechanistic details of how P-CTD recognizes and binds to pY-STAT1, which is likely present in the cell as a tetramer, have not been resolved. This interaction of P-protein and pY-STAT1 inhibits the nuclear accumulation of pY-STAT1 due to a nuclear export sequence (N-NES) in the N-terminal region of the P-protein (28, 29, 38, 39). P-protein also inhibits the nuclear accumulation and signaling of STAT2 in IFN- α/β -treated cells (presumably as part of a pY-STAT1/pY-STAT2 complex) and of STAT3 in cells treated with the interleukin-6 (IL-6) family cytokine oncostatin M (appearing to target STAT1 solely in the context of STAT1-STAT3 complexes). Moreover, P-protein blocks the interaction between activated STAT1 dimers and DNA in extracts of IFN- γ -treated cells (39) and in solution with purified pY-STAT1 dimers (37).

The capacity to bind to STAT1 and inhibit STAT1 signaling is conserved across the P-proteins of the *lyssavirus* genus, albeit with some variation in binding and inhibitory activity (33, 40, 41). Note that STAT1 antagonism enables not only the inhibition of IFN signaling in infected cells but also the deactivation of a pre-established antiviral state (42). Although it is clear that the P-CTD is the principal site of interaction with STAT complexes, a lack of direct structural analysis of the STAT1/P-CTD complex means that our understanding of its binding specificity remains incomplete. Nuclear magnetic resonance (NMR) analysis indicated that there is an extensive interface between STAT1 and P-CTD, including three distinct regions on the P-CTD surface, two of which interact with STAT1-CCD-DBD, whereas the other interacts with the STAT1 NTD, or the C terminal region, or both (37). The P-CTD competes with the binding of pY-STAT1 to target DNA, which suggests a mechanism for the inhibition of IFN signaling in immune evasion (29, 37, 39).

Here, using the intact protein, we report the tetrameric, DNA-bound form of pY-STAT1 by cryo-electron microscopy (cryo-EM). This structure of the full-length protein advances on the previously reported truncated pY-STAT1 dimer-DNA structure which comprised the pY-STAT1 core, but lacks the NTD and C-terminal region (Fig. 1A). These data not only provide structural insights into how oligomeric pY-STAT1 engages target DNA, but also information that can advance our understanding of interactions with other proteins, including viral IFN antagonists. Based on assays using a series of STAT1 constructs including domain truncations and tetramer-incompetent mutants in an unphosphorylated or phosphorylated state, we report insights into how the tetrameric form relates to RABV P protein-binding. Our results indicate that both the dimerization of STAT1 by tyrosine phosphorylation and its tetramerization through the NTD are important for P-CTD binding. Quantitative analyses of the intermolecular interaction imply that there are multiple binding sites on the STAT1 surface and support the existence of several discrete interaction sites on P-CTD (37). Together with SAXS analysis, which demonstrates no substantial conformational rearrangement upon DNA or P-CTD binding to the pY-STAT1 tetramer, the cryo-EM structure of tetrameric pY-STAT1 in complex with DNA was used to predict the P-CTD interaction surfaces.

RESULTS

Sample preparation and characterization by SEC-MALS

We prepared purified samples of full-length U-STAT1, N-terminal truncated U-ΔN STAT1 (residues 136 to 750), full-length pY-STAT1, N-terminal truncated pY-ΔN STAT1, C-terminal truncated pY-STAT1 ΔC (residues 1 to 709), pY-STAT1 core (residues 136 to 709) comprising

the CCD, DBD, LD, SH2 domain, and a part of the TAD, pY-STAT1 F77A/L78A, as well as RABV (*Nishigahara* strain) P-CTD (Fig. 1B, fig. S1). To characterize the tetramerization state in solution, multiangle light scattering combined with size exclusion chromatography (SEC-MALS) experiments were performed on full-length pY-STAT1, the complex of pY-STAT1 with double-stranded DNA derived from the M67 variant (pY-STAT1/DNA), and the complex of pY-STAT1 with the P-CTD (pY-STAT1/P-CTD) (Fig. 1C). M67 contains a single GAS site, which harbors the pseudo-palindromic consensus sequence 5'-TTCN2-4GAA-3, derived from the *c-fos* promoter (46). The M67 element has been used widely in studies on STAT binding to DNA, including the first crystal structure analyses of the pY-STAT core dimer [pY-STAT1 and pY-STAT3 (7, 8)]. The molecular weight of pY-STAT1 was calculated to be 336.2 kDa (theoretically 362.3 kDa for a tetramer) at the peak position (Fig. 1C), indicating tetramer formation, consistent with previous findings using analytical ultracentrifugation (19). A delay of elution volume of the pY-STAT1/DNA complex [13.40 mL (DNA free) to 13.45 mL (DNA complex)] from the SEC demonstrated that the pY-STAT1 tetramer formed slightly more compact assembly upon DNA binding (Fig. 1C). The presence of P-CTD in the fraction from SEC was confirmed by SDS-PAGE (fig. S1). We prepared a construct of pY-STAT1 in which we introduced the mutations F77A/L78A that disrupt dimerization of the NTD and should prevent tetramerization. This construct produces only dimeric pY-STAT1 F77A/L78A (44), which was confirmed by SEC (Fig. 1D).

Three-dimensional cryo-EM structure reveals how pY-STAT1 assembles into dimer-of-dimers

To elucidate the interaction of full, physiologically relevant pY-STAT1 tetramers with target DNA, we reconstructed 3D cryo-EM maps depicting the dimeric and tetrameric structures of human pY-STAT1 in complex with double-stranded DNA from the same EM data set. Both maps were refined with C2 symmetry. The resolution of the final maps was calculated by the Fourier shell correlation (FSC) method at the FSC 0.143 criterion as 3.07 Å for the dimer and 3.76 Å for the tetramer. Subsequently, atomic models of the pY-STAT1 dimer and tetramer, each with DNA, were structurally refined to each respective final map (Fig. 2, A and B).

Initial refinement was performed on the dimer structure. The density revealed the phosphorylated tyrosine (Tyr⁷⁰¹) and recognition by the SH2 domain of the opposing protomer using residues around Arg⁶⁰² (fig. S2A). Because of the imperfect palindromic nature of the DNA sequence (ACAGTTTCCCGTAAATGC; see Materials and Methods) and the intrinsic inability to uniformly average EM density based on a DNA direction, the EM density corresponding to the nonpalindromic base region represents an average of two base densities related by pseudo-two-fold rotation symmetry. Consequently, the occupancy values of all of the DNA atoms were set to 0.5, and structure refinement was conducted with C2 symmetry. This approach mirrors the reported structures of the pY-STAT core dimer in complex with DNA (7, 8), where the crystallographic two-fold rotation axis is perpendicular to the long axis of pseudo-palindromic DNA (Fig. 2C). The core region of the dimeric pY-STAT1 determined here by cryo-EM can be superimposed well to the reported structures determined by X-ray crystallography, demonstrating an identical recognition mode of the M67 pseudo-palindromic sequence [the same sequence as previously reported (8)]. However, a small difference regarding the angle between the two protomers making up the core was observed (Fig. 2C). Both structures exhibit a protrusion in the

C-terminal loop region (range: 655 to 661 of the SH2 domain), interacting with the identical region from the counterpart protomer (fig. S2, C and D). However, the conformation of these loops differs between the EM structure (the dimer and tetramer) and the previous crystal structure of the pY-STAT1 core with DNA (7, 8). In the crystal structure, the tips of the two protruded loops, related by two-fold symmetry, engage in a "β-sheet-like" interaction involving Ala⁶⁵⁶ and Gln⁶⁵⁷ (fig. S2C). In contrast, the EM structure reveals a shifted tip of the protrusion, with the most substantial interaction between two symmetry mates being a hydrophobic interaction facilitated by Ile⁶⁵⁹ of each protomer (fig. S2D).

For the tetramer structure analysis, the structure refinement with C2 symmetry was applied using two dimer structures. Ultimately, there are three two-fold rotation axes in parallel in the tetramer: a global rotation axis at the center (used in the C2 averaging) and two local axes (Fig. 2D). The refined EM dimeric structure seamlessly aligned with the 3D map of the tetramer without requiring additional local corrections. Due to the absence of residual density, it was impossible to model the NTD spanning residues 1 to 134 and most of the C terminal region spanning residues 710 to 750. Consequently, the final models of both the dimeric and tetrameric structures encompass the core region of pY-STAT1, specifically residues 135 to 711.

In the tetrameric structural configuration, two dimeric units engage in interactions along the two-fold rotational axis. The total surface area utilised for this interaction is 495 Å² per protomer. Notably, hydrophobic regions, including the side chains of Asn⁴⁸³, Ser⁴⁸⁵, and Leu⁴⁸⁸ within the C-terminal loop of the DBD (fig. S3A), a rigid loop connected to β12 of the LD (comprising Thr⁴⁸⁹, Pro⁴⁹⁰, Pro⁴⁹¹, and Cys⁴⁹²), and Leu536 on β13, contribute to the interaction with their respective

two-fold related counterparts (Fig. 2E). Additionally, the formation of an intermolecular salt bridge between Asp⁵³⁴ and Arg⁵⁴¹ serves to strengthen these interactions (fig. S3B). Of note, the S_γ atoms of Cys⁴⁹² from two opposing protomers of the tetramer are found within a distance indicative of potential disulfide bonding (3.8 Å, Fig. 2E). However, the purification and the vitrification for EM observation of the pY-STAT1-DNA complex were carried out in the presence of DTT, making the disulfide bonding in the EM structure less probable, despite prior indications of a role for Cys⁴⁹² in STAT1 aggregation (15). The two-fold rotational axis is positioned equidistantly between the two Cys⁴⁹² residues, resulting in observable rocking movements centered at this point, which became evident from the C2 symmetry averaging procedure (movie S1). In the tetrameric structural conformation, two DNA molecules are arranged in parallel orientation (Fig. 2D). Electrostatic potential calculation of the molecular surface reveals the presence of basic patches not only within the DNA tunnels (DNA binding site) but also along the rims of the DNA tunnels aligned with the DNA direction (Fig. 2F). The formation of these basic belts may facilitate interactions with the DNA backbone, thereby enhancing the efficiency of subsequent sequence scanning by the DNA binding site.

Although modeling was not possible, fragmented low density is located surrounding the CCD. Modelling the location of the NTD based on the crystal structure of U-STAT1 (PDBID: 1YVL) (9) did not align with this density (Fig. 2G, H). Given that the tetramer has a K_D value of 100 nM (tetramer disassociation constant into two dimers) (19) and its formation is induced by the interaction of the NTDs (Fig. 1D), a plausible scenario involves the dimerized NTD settling in this fragmented density (Fig. 2I). In the tetrameric structure with the possible NTD location, NTD dimerization acts as a tether connecting the A chain of the AB dimer and the C chain of the CD

dimer as well as the B chain of the AB dimer and the D chain of the CD dimer (Fig. 2I). The distances between the C-terminus (Gln¹²⁶) of the NTD of the A chain in the indicated NTD dimer (indicated as white arrow) and the N-terminus of the A chain CCD (Met¹³⁵), as well as the distance between the C-terminus (Gln¹²⁶) of the other NTD (of the C chain) in the same NTD dimer and the N-terminus of C chain CCD (Met¹³⁵), are reasonable for each single protomer (15 Å and 10 Å, respectively, Fig. 2J).). The tethering of the B chain and the D chain by the other NTD dimer occurs similarly as related by the C2 symmetry to the A-C tethering. The NTD-integrated tetrameric model was utilised for fitting the SAXS density (below).

Small-angle X-ray scattering (SAXS) demonstrates similar conformations of pY-STAT1, pY-STAT1/DNA, and pY-STAT1/P-CTD

SEC-SAXS experiments were carried out on the pY-STAT1 protein and on complexes (pY-STAT1/DNA, pY-STAT1/P-CTD, Fig. 3) as summarized in table S1. Consistent with the observation on the SEC elution volume (see above), the R_g values (Guinier plot and $P(r)$ function, table S1) revealed that the pY-STAT1/DNA complex is the most compact; pY-STAT1/P-CTD adopts the most stretched conformation. *Ab initio* modeling (electron density averages) for the three molecules is presented in fig. S4B. With the comparable SAXS results ($P(r)$ function and *ab initio* modeling) between pY-STAT1/DNA and pY-STAT1/P-CTD, we assumed that there has been no drastic conformational change in the tetrameric pY-STAT1 upon DNA or P-CTD binding (Fig. 3, fig. S4, table S1). The pY-STAT1 core structure without DNA, is reported to form a similar dimer assembly to the DNA-bound dimer, judged from the crystallographic study of STAT1 fused to a peptide from the parasite *Toxoplasma gondii* (47) (note: the authors attempted crystallisation with DNA but, no density was observed for DNA from the X-ray reflection file deposited in the

PDB (PDBID: 8D3F)). The availability of these structures enabled us to analyze possible interactions of tetrameric pY-STAT1 with proteins, including P-CTD. As P-CTD inhibits DNA binding, we removed two DNA molecules from the EM structure (tetrameric pY-STAT1/DNA) for fitting in pY-STAT1 and pY-STAT1/P-CTD DENSS electron densities (fig. S4A, B), under the assumption that the pY-STAT1 tetramer without DNA adopts a similar conformation to pY-STAT1 with DNA. The location of the NTD of the pY-STAT1 tetramer was predicted from the residual EM map in this study (described above, Fig. 2G, I, J) and the assembly with the NTD was utilised for superimposition. The information obtained from the SAXS experiments, that the three structures have similar conformations, was further used to explore potential target sites for the P-CTD.

Experimental Analysis of Interactions of STAT1 with P-CTD

From the SEC experiments, binding of P-CTD to pY-STAT1 resulted in slightly earlier elution (13.30 mL) (Fig. 1C) than pY-STAT1. The molecular weight of the corresponding SEC-MALS signal of the pY-STAT1/P-CTD complex is estimated to be 353.7 kDa from SEC-MALS. This could indicate that the complex is composed of a pY-STAT1 tetramer plus one P-CTD molecule (theoretically 375.3 kDa), or a pY-STAT1 tetramer plus two P-CTD molecules (theoretically 388.1 kDa), or an even more complicated stoichiometry. The presence of P-CTD in the fraction from SEC was confirmed by SDS-PAGE (fig. S1). To assess interactions of pY-STAT1 (and variants) with P-CTD, and so enable predictions of complex interactions based on the above structural data, we initially attempted isothermal titration calorimetry (ITC) to analyze the interaction between full-length pY-STAT1 and P-CTD. However, precipitation affecting quantification accuracy was observed in the attempted conditions after the titration of P-CTD into pY-STAT1 in the instrument.

Therefore, the fluorescence polarization method was employed instead to determine K_D values for the interaction of P-CTD and constructs of STAT1 (Fig. 1B): U-STAT1, pY-STAT1, pY- Δ N STAT1, pY-STAT1 core, pY-STAT1- Δ C, or pY-STAT1 F77A/L78A (tetramerization deficient mutant) (Fig.4A, B-Table 1).

The K_D value between U-STAT1 and P-CTD was determined as 11 μ M, which is broadly consistent with the reported K_D value between U-STAT1 core and P-CTD derived using analytical ultracentrifugation (10 – 20 μ M) (37). The truncation of NTD from U-STAT1 (U- Δ N-STAT1) gave a K_D value of 6 μ M. The K_D value between pY-STAT1 and P-CTD was determined as 18 nM, increasing to 81 nM in the absence of NTD (pY- Δ N-STAT1). The K_D value for pY-STAT1 F77A/L78A (dimeric pY-STAT1; 38 nM) is intermediate between pY-STAT1 (intact, tetramer) and Δ N-pY-STAT1 (dimer). pY-STAT1- Δ C exhibited a K_D of 46 nM, while the pY-STAT1 core region showed a K_D of 194 nM. Comparison with the K_D values of U- Δ N-STAT1 (6 μ M) and pY- Δ N-STAT1 (81 nM) indicate that the pY-SH2 mediated dimer formation that occurs upon phosphorylation of Tyr⁷⁰¹ is a crucial step for P-CTD recognition. The phosphorylated moiety of Tyr⁷⁰¹ is sequestered within the SH2 domain of its counterpart, concealed from the solvent, as evidenced by the crystallographic (7, 8) and the EM (this study) structures. Consequently, P-CTD cannot directly interact with the phosphorylated moiety of Tyr⁷⁰¹, rather, P-CTD plausibly recognizes two molecular surfaces of pY-STAT1 in proximity upon tyrosine phosphorylation. Moreover, the comparison of the K_D values between tetrameric pY-STAT1 (pY-STAT1 full) and dimeric pY-STAT1 (pY- Δ N STAT1 and pY-STAT1 F77A/L78A) indicates that tetramerization *via* Phe⁷⁷ and Leu⁷⁸ of the NTD further enhances the interaction. A possible, previously unidentified, role of the pY-STAT1 C-terminal region in specific targeting by P-CTD was also

indicated, based on differing K_D values of pY-STAT1 (18 nM) and pY-STAT1 ΔC (46 nM). These insights advance our understanding of the structural biology underpinning the interactions between pY-STAT1 and P-CTD. Combined with the structural information above, these interaction analyses enabled novel predictions about the STAT1/P-CTD complex (see below).

DISCUSSION

In the JAK-STAT system, the phosphorylation of a conserved tyrosine residue located in the loop in the TAD is regarded as a key step for STAT activation, either through homodimerization or heterodimerization, subsequent to ligand recognition by a cell surface receptor and JAK activation. However, interactions mediated by the NTD, independently of tyrosine phosphorylation, have also been identified in several STAT proteins, indicating the potential importance of tetramerization. The disruption of NTD-mediated dimerization of STAT1 has been reported, specifically through generating F77A, L78A, or both mutants (44). These residues, of which Phe⁷⁷ is conserved as a hydrophobic residue (as Leu⁸¹ in STAT2, Val⁷⁷ in STAT3, Leu⁷⁷ in STAT4, Phe⁸¹ in STAT5a/5b, and Ile⁷⁷ in STAT6), and Leu⁷⁸ is strictly conserved in all human STATs, play important roles in this context. Due to NTD-mediated interactions, U-STAT1, U-STAT3, U-STAT4, and U-STAT5a can exist as homodimers, plausibly adopting an anti-parallel conformation, in unstimulated cells (14, 18, 43, 48). Therefore, for a comprehensive understanding of STAT functionality, particularly for phosphorylated STAT, it is important to consider not only pY-SH2 domain interactions, but also interactions mediated by the NTD, whether between homomeric or heteromeric STAT molecules.

From a structural perspective, an important milestone was achieved with the solving of the dimeric structures of the tyrosine-phosphorylated core regions of STAT1 and STAT3 (7, 8). These pioneering insights were derived from crystallographic studies, which included the parallel dimer structures complexed with DNA containing specific STAT1 and STAT3 target sequences (M67). Subsequently, crystal structures of DNA complexes of pY-STAT3 with acetylation of Lys⁶⁸⁵ (49), pY-STAT1 bound with a peptide of the TgIST protein from the parasite *Toxoplasma gondii* (47), and pY-STAT6 in complex with DNA (50) were analyzed. Furthermore, structural characterizations have been extended to the NTD of STAT proteins, revealing its capacity to form homodimers (9, 25). However, note that the linker region connecting the NTD to the core region of STAT proteins, spanning approximately seven amino acid residues, is typically disordered, and cannot be observed in crystal structures of unphosphorylated STATs (9, 25). Consequently, it has not been completely clarified how the NTD tethers pY-STAT1 dimers to generate tetramers or higher-order oligomers. The biological importance of oligomeric formations, particularly in the context of enhancement of promoter binding by the NTD, referred to as cooperative DNA binding, has been reported for various STAT members, most commonly pY-STAT1 (15–18). This phenomenon extends beyond homodimeric STAT interactions to include heterodimeric STAT interactions, such as STAT1-STAT2 and STAT5a-STAT5b (13, 14). Extensive investigations involving tetramerization-deficient mutants of the STAT1 NTD, such as F77A or F77A/L78A, have underscored the indispensable nature of cooperativity for the biological activity of STAT1 (17, 18). Note that both the activation and repression of IFN- γ responses in mouse spleen and embryonic fibroblasts are severely compromised by the F77A mutation in STAT1 (17).

Previously reported analyses in solution (19, 51) and the cryo-EM structure from the current study demonstrate that pY-STAT1 forms homotetramers. In the tetrameric cryo-EM structure of pY-STAT1 in complex with double-stranded DNA (GAS sequence), the positioning of two double-stranded DNA molecules in the two dimers does not exhibit linearity (Fig. 2D). Additionally, the geometry and the distance between the end of one double-stranded DNA molecule and the most proximal end of the other DNA molecule does not appear to match that expected for reported “tandem GAS” sites that have a 10-bp spacer, indicative of a 360° rotation with a 10-bp translation (approximately 34 Å) of the binding site, as we modeled (Fig. 5A, right). The DNA used in our experiment is an 18-bp oligonucleotide, which is characterized by five bases extending towards the 5’ end and four bases extending towards the 3’ end with the middle GAS consensus element. This particular structural feature mitigates concerns about potential distortion of the pY-STAT1 tetramer due to steric hindrance imposed by the presence of two double-stranded DNA molecules. Note that the MALS and SAXS analyzes (Fig. 1C and Fig. 4, A and D) of the tetrameric assembly suggest a compact, rigid body in solution with bound DNA, which is supported by the tetrameric cryo-EM structure of pY-STAT1 with DNA fitting well to the density calculated from the SAXS experiment (fig. S4). Consequently, it can be inferred that the analyzed tetrameric cryo-EM structure demonstrates a stable state in solution.

The previously proposed tandem-GAS DNA-binding model of the pY-STAT1 tetramer was proposed through binding and reporter-gene experiments using gain-of-function mutations (D92R/E96R) in the NTD (20). In this model, the DNA molecule is positioned centrally, with two NTDs functioning as arms to secure the DNA, whereas the other NTD pair remain unbound (tethering the A- and D-chains and leaving the B- and C-chains unbound; Fig. 5B). To achieve the

transition to this proposed model from the current cryo-EM structure, dissociation of one of the two NTD dimers is required. In this context, the K_D value of 100 nM for the two NTD-NTD interactions of the tetramer formation (19) could be rationalized enabling rearrangement and high-affinity binding to DNA (K_D , 1 nM (15)). To evaluate this model, we constructed a similar model in which two pY-STAT1 dimers are bound to two separate DNA-binding sites with a 10-bp spacer (Fig. 5, A to C). Based on our model, the distance between the two N termini of the CCDs located diagonally is 150 Å (Fig. 5C). Considering that the distance between the two C termini of one NTD dimer is 71 Å (Fig. 5D), it is improbable for one NTD dimer to connect two STAT1 core regions located diagonally, as was suggested previously (20), [a similar model (Fig. 5B) shows where the A- and D-chains are tethered by an NTD], without a sharp bend in the DNA. Instead, we propose that one NTD dimer could tether two STAT1 core regions situated in a parallel manner with the two N termini 70 Å apart (the A- and C-chains and the B- and D-chains; Fig. 5, C and E). The symmetrical positioning of the two NTD dimers enables a seamless transition from the cryo-EM structure to this tandem model without the dissociation and reassociation of the NTD dimer (connecting the A- and C-chains and the B- and D-chains) into monomers, as proposed previously (Fig. 5E) (20).

With these insights into the structure of tetrameric STAT1, we considered how this may apply to other protein interactions of STAT1, including viral IFN antagonist proteins. Whereas there are several reports of viral counteraction of STAT signaling, analysis of STAT structures in complex with viral proteins remains limited. Detailed atomic-level information on how viral proteins disrupt the host immune system can potentially facilitate the rational design of antiviral compounds or attenuated vaccines (37). We thus sought to further characterize the targeting of STAT1 by RABV

P-CTD, by analyzing the interaction of P-CTD with STAT protein constructs, including U-STAT1, U-ΔN STAT1, pY-STAT1, pY-ΔN STAT1, pY-STAT1 core, pY-STAT1 ΔC, or pY-STAT1 F77A/L78A, and using these data together with the tetrameric cryo-EM structure to understand the binding mode of pY-STAT1 and P-CTD. Considering the importance of pY-SH2-mediated dimerization for P-CTD binding [as assessed from the reduction in the K_D value for pY-ΔN-STAT1 (81 nM) compared with that of U-ΔN-STAT1 (6 μM)], it is conceivable that the recognition mode involves one P-CTD unit engaging multiple STAT1 core regions derived from two intertwined STAT1 molecules, rather than through direct binding to the phosphorylated moiety, which is sequestered in the SH2 domain of the counterpart pY-STAT1. Furthermore, interaction analysis with pY-STAT1 (tetramer, K_D = 18 nM) indicated the importance of tetramerization involving Phe⁷⁷ and Leu⁷⁸ of the NTD. Two possibilities exist regarding the effect of tetramerization of the NTD on P-CTD binding. One possibility is the direct interaction between the NTD and core region of the same pY-STAT1 dimer and P-CTD. The other possibility is that the P-CTD molecule interacts with the NTD and the core region derived from two pY-STAT1 dimers; that is, the P-CTD molecule may bridge two pY-STAT1 dimers. We found that the NTD-deficient U-STAT1 and intact U-STAT1 have K_D values of 6 μM [this study; values of 10 to 20 μM were determined previously (37)] and 11 μM, respectively, upon P-CTD interaction. There is also a slightly different interaction affinity among ΔN-pY STAT1 (dimer, K_D = 81 nM), pY-STAT1 F77A/L78A (dimer, K_D = 38 nM), and pY-STAT1 (tetramer, K_D = 18 nM). These data provide evidence for a direct interaction between the STAT1-NTD and P-CTD. We cannot exclude the latter possibility that the lowest K_D value with the pY-STAT1 tetramer does not result from direct a STAT1-NTD/P-CTD interaction, but rather from an interaction of P-CTD with a newly formed interface on the core regions of STAT1 that results from tetramer formation. Because we

were unable to fit STAT1-NTD into the EM-density precisely, it is difficult to build a binding model of the pY-STAT1 tetramer and P-CTD that includes the STAT1-NTD/P-CTD interaction. Therefore, we only built a binding model for the pY-STAT1 tetramer–P-CTD interaction based on the latter possibility using previous information about which residues on the P-CTD surface are important for the interaction (Fig. 6A) (37). Considering the difference in the K_D values between intact pY-STAT1 (tetramer, 18 nM) and pY-STAT1 ΔC (46 nM), the C-terminal region of STAT1, which the EM density also does not cover, likely also contributes to the interaction with P-CTD.

On the P-CTD side, the interaction mechanism between U-STAT1 and P-CTD was previously dissected extensively by NMR spectroscopy and a rigorous mutation-scanning-based reporter gene assay (37). Three distinct regions on the molecular surface of P-CTD, denoted as regions A, B, and C, were identified as critical binding sites for STAT1. Note that the residues in these regions are highly conserved among lyssaviruses. It was suggested that region A (His²⁰³ and Ala²⁰⁶) and region B (Asp²³⁵) contribute to multiple and distinct binding sites on the CCD-DBD of STAT1 because the side chains of the surface residues of the respective regions are oriented in opposite directions. Additionally, region C (Leu²⁷⁶ and Leu²⁷⁷) was proposed to interact with either the NTD or C-terminal region of STAT1.

Assuming that the conformation of the pY-STAT1 tetramer is conserved across pY-STAT1, pY-STAT1/DNA, and pY-STAT1/P-CTD, which is supported by results from the SAXS analysis, we built a binding model using the cryo-EM structure of pY-STAT1/DNA (Fig. 6, A and B). In this model, one P-CTD unit interacts with three STAT1 molecules, bridging two protomers (the A– and B-chains or the C– and D-chains) in the pY-SH2–mediated dimer, and two protomers (the A–

and C- chains or the B- and D- chains), which are connected by the interactions of each NTD. Given the proximity of P-CTD binding to the DNA-binding site, our model explains how P-CTD binding competitively interferes with DNA binding (37), providing a binding stoichiometry of 4:2. Nonetheless, two additional binding sites, bridging two protomers in the pY-mediated dimer, could be justified due to the two-fold symmetry inherent in the primary binding site (Fig. 6, A, D, and E). We generated a 2:2 binding model of pY-STAT1 core/P-CTD with two-fold symmetry, which is related to the K_D value of 194 nM (Fig. 6, D and E). In the tetramer binding model, the two primary binding sites of the P-CTD and two secondary binding sites of the P-CTD are not equivalent (Fig. 6A). This interpretation is consistent with the observed binding stoichiometry of pY- Δ N-STAT1/P3 (residues 53 to 297 of Nish-P), which is one (indicating 2:2) (52), because pY- Δ N-STAT1 (and pY-STAT1 F77A/L78A) forms a pY-SH2-dependent dimer, harboring two equivalent P-CTD-binding sites.

In the context of the RABV P-CTD targeting such as STAT1/STAT1 homodimer, STAT1/STAT2 heterodimer, and STAT1/STAT3 heterodimer, reported data indicate that cell stimulation by RABV infection is a key process to inactivate STAT function (28, 29, 33, 53). It has not been clarified, however, how phosphorylation of STATs is related to the interaction with P-CTD, and whether other factors are necessary for the interaction. In the case of STAT1, we here demonstrated that STAT tetramerization concomitant with tyrosine phosphorylation is the determining factor for STAT1 counteraction by P-CTD. Although our proposed binding model offers insights, further experimental results, including structural analysis of pY-STAT1 in complex with P-CTD, are imperative to obtain a comprehensive understanding of the detailed mechanisms of immune evasion by RABV.

MATERIALS AND METHODS

Expression vector constructs and purification of human STAT1 variants

For U-STAT1 expression, the cDNA coding full-length human STAT1 in fusion with N-terminal $6 \times$ His-tag and GB1 was inserted into pET21d (Merck), and *E. coli* B834 (DE3) was used as the expression host. For all the tyrosine-phosphorylated STAT1 variants, pET21d plasmids harboring STAT1 with N-terminal $10 \times$ His-tag were used for the transformation of BL21(DE3) TKB1 (Agilent Technologies). Based on this plasmid, variant expression vectors of Δ N STAT1 (residues 136 to 750), STAT1 core (residues 136 to 710), STAT1 Δ C (residues 1 to 710), and STAT1 F77A/L78A were constructed. STAT1 F77A/L78A was further mutated after the generation of F77A. The primer sequences for each construct are as follows: Δ N STAT1: 5'-TTA GAC AAA CAG AAA GAG CTT GACAG-3' (forward) and 5'-CAT ATG GCC CTG AAA ATA CAG GTT TTC-3' (reverse); STAT1 Δ C and STAT1 core: 5'-TAA GCG GCC GCA CTC GAG C-3' (forward) and 5'-AGA CAC AGA AAT CAA CTC AGT CTT G-3' (reverse); F77A: 5'-GAATAACGCGTTGCTACAGCATAACATAAGG-3' (forward) and 5'-GTAGCAACGCGTTATTCTCCAAAGAAAAGC-3' (reverse); F77A/L78A: 5'-AGAATAACGCGGCGCTACAGCATAACATAAGG-3' (forward) and 5'-GTAGCGCCGCGTTATTCTCCAAAGAAAAGC-3' (reverse). For the phosphorylation of STAT1 at Tyr⁷⁰¹, BL21(DE3)TKB1 cells were used as the expression host. Purifications were performed as previously reported (51, 54). To compare the tetramerization states of pY-STAT1 and pY-STAT1 F77A/L78A, analysis by size-exclusion column S200 10/300 increase (Cytiva) with purification buffer [10 mM HEPES (pH 7.4), 150 mM NaCl, 1 mM DTT, 10% glycerol] was performed. For standard proteins, a gel filtration standard (BioRad) was used, which

included thyroglobulin (bovine; 670 kD), γ -globulin (bovine; 158 kD), ovalbumin (chicken; 44 kD), myoglobin (horse; 17 kD), and vitamin B12 (1350 Da).

Preparation of rabies virus P-CTD with and without the tetracysteine tag

For fluorescence polarization (FP) experiments, the Nishigahara RABV strain (Nish-) P-CTD protein fused with a C-terminal tetracysteine (TC) tag (amino acid residues: CCPGCC) was prepared. The expression plasmid for P-CTD (amino acid residues 186 to 297, inserted in pET-28(a) (51, 54) was used as the template for PCR with the following primers: Nish-P-CTD_TC forward: CTGCGCGAGCTAAGAATTCGAGCTCCGTCGACAAG; Nish-P-CTD_TC reverse: CAGCCCGGGCAGCAAGATGCATAGCGATTCAAATCATC. The additional amino acid sequence of Nish-P-CTD_TC, CPGCC, as a flexible tail (the original C-terminal residue of Nish-P, Cys²⁹⁷, is used for a part of the TC tag) was used. Purification of Nish-P-CTD with and without the TC tag was performed as reported previously (51, 54).

Fluorescence polarization assays

To assess the binding affinity of the P-CTD for pY-STAT1, P-CTD with a tetracysteine (TC) tag was expressed and fluorescently labelled with FAsH-EDT₂ (ThermoFisher). TC-tagged P-CTD was mixed with FAsH-EDT₂ with a molar ratio of 1:1.6, and FAsH-EDT₂-conjugated P-CTD was purified with a S75 10/300 column (Cytiva). Two independent concentrations of pY-STAT1 were prepared, and each was diluted two-fold on a Black Flat Bottom Polystyrene NB microplate (Corning). In total, 24 wells of different concentrations of pY-STAT1 were prepared, and 63 nM FAsH-EDT₂-fused P-CTD was added to each well. The final volume per well was adjusted to 100 μ l with buffer [10mM HEPES (pH 7.4), 150 mM NaCl, 1 mM DTT, 10% glycerol]. A control

lacking pYSTAT1 was also prepared and used for gain adjustment. Fluorescence polarization measurements were immediately recorded with a SpectraMax i3x plate reader (Molecular Devices) and the assay was repeated in triplicate for each pY-STAT1. One-site specific binding least square fit curves were plotted with software Prism 10.3.1 (GraphPad).

SEC-MALS

Multi-angle light scattering combined with size exclusion chromatography (SEC-MALS) was performed with a high-performance liquid chromatography (HPLC) system, Alliance 2695 (Waters) with DAWN HELEOS II (Wyatt Technology) and a 2414 Refractive Index (RI) detector (Waters). Samples (30 μ l) were injected into a Superose12 increase 10/300 column (Cytiva), equilibrated with a buffer [10 mM HEPES-NaOH, including 150 mM NaCl, 3% glycerol, and 2 mM DTT (pH7.4)], under a flow rate of 0.5 ml/min. Analysis of the acquired data was performed with ASTRA 6.1 software (Wyatt Technology).

Cryo-EM sample preparation and data acquisition

The palindromic GAS sequence double-stranded oligonucleotides, termed M67, from the *c-fos* promoter (ACAGTTTCCCGTAAATGC/TGCATTTACGGGAAACTG) was dissolved in buffer [10 mM HEPES-NaOH (pH 7.4), 150 mM NaCl, 10% glycerol, 10 mM MgCl₂, 0.5 mM EDTA]. The oligonucleotide mixture was incubated at 95°C for 3 min and cooled to room temperature before being added to the pY-STAT1 solution. The pY-STAT1–DNA complex was further prepared by mixing pY-STAT1 and DNA at a molar ratio of 1:1.5 (protein dimer: DNA duplex). P-CTD was further added with a molar ratio of 1:1.5 (pY-STAT1:P-CTD), which was expected to generate a pY-STAT1/P-CTD complex. The sample was SEC-purified with a Superose 12 10/300

column (Cytiva) and the peak fraction (0.6 mg/ml) was used for grid preparation. Three microliters of the sample were applied onto a glow-discharged [30 s, 30 mA on both sides using PIB-10 (Vacuum Device Inc.)] 300-mesh R0.6/1.0 Quantifoil Au grid. Grids were blotted for 5 s in 100% humidity at 4°C and plunge-frozen in liquid ethane with a Vitrobot Mark IV (Thermo Fisher Scientific). Data were collected on a Titan Krios cryo-transmission electron microscope (Thermo Fisher Scientific) equipped with a thermal field emission electron gun operated at 300 kV, a Cs corrector (CEOS GmbH), an energy filter with a 20-eV slit width, and a K3 direct electron detector camera (Gatan) at the Institute for Protein Research, Osaka University. For automated data acquisition, SerialEM software was used to collect cryo-EM images as movies with a 3 x 3 beam image shift scheme. Movie frames were recorded with the K3 camera at a nominal magnification of 81,000 $\times$, corresponding to a pixel size of 0.87 Å in EFTEM mode, with a defocus range set between -0.8 and -1.8 μ m. The data were collected with a total exposure time of 4.7 s, fractionated into 54 frames, maintaining a cumulative dose of approximately 50 electrons per Å² in counting and correlated double sampling (CDS) imaging mode. A total of 6993 movies were collected and used for the subsequent steps of image processing.

Image processing, 3D reconstruction, and structure refinement

A gain reference image was generated employing all movie frames through the “*relion_estimate_gain*” command within the software package RELION 4beta (55). Subsequent image-processing steps were performed with cryoSPARC software (56). After motion correction and gain normalization with the “Patch motion correction” algorithm, the contrast transfer function (CTF) was estimated with the “Patch CTF Estimation” job. Images with poor CTF fits were excluded, resulting in 6417 images. Particle coordinates were registered using the template-based

particle-picking algorithm, and the selected particles were extracted into a 240×240-pixel box with a pixel size of 1.76 Å. After particle extraction, 2,649,037 particles were selected after three rounds of two-dimensional (2D) classification. These selected particles were subjected to reference-based three-dimensional (3D) classification with a “Heterogeneous Refinement” job to roughly segregate pY-STAT1 dimers and tetramers. Subsequently, four rounds of “Ab-initio Reconstruction” jobs were performed to remove junk particles in the dimer and tetramer classes, respectively. The particles were then combined, and a pool of 495,675 particles was again subjected to a “Heterogeneous Refinement” job to classify the dimers and tetramers. Particles in each class were subjected to “Ab-initio Reconstruction” jobs to remove junk particles and select 3D classes with well-resolved reconstruction of the coiled-coil domains. For the dimers, the particles were re-centered and extracted into a 240×240-pixel box with a pixel size of 1.32 Å. To visualize the structural diversity of the tetramer, “3D Flexible Refinement” was performed before the final “Ab-initio Reconstruction.” The final 3D reconstruction involved the selected dimers comprising 115,824 particles and tetramers with 17,869 particles using “Homogeneous Refinement” and “Non-uniform Refinement” jobs, achieving reconstructions at resolutions of 3.07 and 3.84 Å, respectively. Resolution was determined using the "gold standard" Fourier Shell Correlation (FSC) method, comparing two maps reconstructed independently from particle sets that had been randomly divided. The local resolution of the final maps was calculated with the “Local Resolution Estimation” job. The image processing workflow is depicted in fig. S5.

Cryo-EM model building and refinement

For atomic model building, the cryo-EM maps were trimmed to remove unnecessary regions outside the molecules with the map trimming tool within Servalcat software (57). Manual model

fitting was performed with Coot (58) which was followed by model refinement with Servalcat with the DNA restrain introduced by LIBG (59). C2 symmetry was applied during the model refinement of both the dimer and tetramer. For dimer refinement, the occupancy values of all of the DNA atoms were set to 0.5 because the two-fold rotation axis is located at the center of the double-strand DNA. STAT1 residues 1 to 134 and 712 to 750 in the dimer, as well as the tetramer, had no resolvable density and were omitted from modeling. Pymol (Schrödinger, LLC) and ChimeraX (60) were used for protein figure preparation. Protein data collection, data processing, and model-building statistics are summarized in table S2.

SEC-SAXS

All of the data of small-angle X-ray scattering connected to size exclusion chromatography (SEC-SAXS) were collected at beamline BL-10C at the photon factory, KEK (Tsukuba, Japan) (61). The X-ray energy was set at 12.4 keV (wavelength: 1.00 Å), and the sample to detector distance was 2081 mm. The chromatography was performed with a HPLC system, Prominencei (SHIMADZU) with a Superdex 200 increase 10/300 column (Cytiva) equilibrated with buffer (10 mM HEPES-NaOH, 150 mM NaCl, 3% glycerol, and 2 mM DTT (pH 7.4)]. Samples (150 µl each) containing 6.8 mg/ml pY-STAT1, 9.2 mg/ml pY-STAT1/P-CTD, and 19 mg/ ml pY-STAT1/DNA were injected, and the flow rate was set to 0.05 ml/min. The detailed beamline setting parameters are shown in table S1. Serial scattering images were accumulated with 20-s exposures. Several hundred SAXS images were recorded in PILATUS3 2M (DECTRIS). The 20 images measured before the sample fraction were averaged as a background profile. A fiber spectrometer, QE65pro (Ocean Insight), set at an angle of 45° to the sample cell, was used to obtain the concentration for each frame (62). All scattering images were averaged at each azimuthal angle to convert them to

one-dimensional scattering intensity data, and then subtraction using the background profile was also performed. The absolute intensity calibration was performed with water as a standard. These processes were calculated using SAngler 2.1.65 (63). Scattering profiles greater than half the intensity of the elution peaks were averaged with MOLASS 1.0.11 (64). The radius of gyration (R_g) and forward scattering intensity (I_0) were calculated from the Guinier approximation by AutoGunier 0.7.1 (65), and the maximum particle dimension ($D_{\max}$) and the pair distribution function ($P(r)$) were calculated with MOLASS (64). These parameters are summarized in table S1. To calculate low-resolution electron density maps, the DENSSWeb server was used with the default parameters (66). The 20 reconstructed electron density maps were aligned and averaged for each dataset. The electron density maps were drawn with ChimeraX (60).

Supplementary Materials

Figs. S1 to S5.

Tables S1 and S2.

Movie S1.

References and Notes

1. J. J. O'Shea, S. M. Holland, L. M. Staudt, JAKs and STATs in Immunity, Immunodeficiency, and Cancer. *N. Engl. J. Med.* **368**, 161–170 (2013).
2. J.-L. Casanova, S. M. Holland, L. D. Notarangelo, Inborn Errors of Human JAKs and STATs. *Immunity* **36**, 515–528 (2012).
3. H. Yu, D. Pardoll, R. Jove, STATs in cancer inflammation and immunity: a leading role for STAT3. *Nat. Rev. Cancer* **9**, 798–809 (2009).
4. J. F. Bromberg, M. H. Wrzeszczynska, G. Devgan, Y. Zhao, R. G. Pestell, C. Albanese, J. E. Darnell, Stat3 as an Oncogene. *Cell* **98**, 295–303 (1999).
5. P. J. Murray, The JAK-STAT Signaling Pathway: Input and Output Integration. *J. Immunol.* **178**, 2623–2629 (2007).
6. G. M. Delgoffe, P. J. Murray, D. A. Vignali, Interpreting mixed signals: the cell's cytokine conundrum. *Curr. Opin. Immunol.* **23**, 632–638 (2011).
7. S. Becker, B. Groner, C. W. Müller, Three-dimensional structure of the Stat3 β homodimer bound to DNA. *Nature* **394**, 28101 (1998).
8. X. Chen, U. Vinkemeier, Y. Zhao, D. Jeruzalmi, J. E. Darnell, J. Kuriyan, Crystal Structure of a Tyrosine Phosphorylated STAT-1 Dimer Bound to DNA. *Cell* **93**, 827–839 (1998).

9. X. Mao, Z. Ren, G. N. Parker, H. Sonderrmann, M. A. Pastorello, W. Wang, J. S. McMurray, B. Demeler, J. E. Darnell, X. Chen, Structural Bases of Unphosphorylated STAT1 Association and Receptor Binding. *Mol Cell* **17**, 761–771 (2005).
10. T. Sekimoto, N. Imamoto, K. Nakajima, T. Hirano, Y. Yoneda, Extracellular signal-dependent nuclear import of Stat1 is mediated by nuclear pore-targeting complex formation with NPI-1, but not Rch1. *EMBO J.* **16**, 7067–7077 (1997).
11. K. M. McBride, G. Banninger, C. McDonald, N. C. Reich, Regulated nuclear import of the STAT1 transcription factor by direct binding of importin- α . *EMBO J.* **21**, 1754–1763 (2002).
12. G. M. Delgoffe, D. A. A. Vignali, STAT heterodimers in immunity. *JAK-STAT* **2**, e23060 (2013).
13. X. Li, S. Leung, C. Burns, G. R. Stark, Cooperative binding of Stat1–2 heterodimers and ISGF3 to tandem DNA elements. *Biochimie* **80**, 703–710 (1998).
14. A. Begitt, S. Krause, J. R. Cavey, D. E. Vinkemeier, U. Vinkemeier, A family-wide assessment of latent STAT transcription factor interactions reveals divergent dimer repertoires. *J. Biol. Chem.* **299**, 104703 (2023).
15. U. Vinkemeier, S. L. Cohen, I. Moarefi, B. T. Chait, J. Kuriyan, J. E. Darnell, DNA binding of in vitro activated Stat1 α , Stat1 β and truncated Stat1: interaction between NH₂-terminal domains stabilizes binding of two dimers to tandem DNA sites. *The EMBO J.* **15**, 5616–5626 (1996).
16. J. Staab, C. Herrmann-Lingen, T. Meyer, Clinically Relevant Dimer Interface Mutants of STAT1 Transcription Factor Exhibit Differential Gene Expression. *PLoS ONE* **8**, e69903 (2013).
17. A. Begitt, M. Driescher, T. Meyer, C. D. Schmid, M. Baker, F. Antunes, K.-P. Knobloch, M. R. Owen, R. Naumann, T. Decker, U. Vinkemeier, STAT1-cooperative DNA binding distinguishes type 1 from type 2 interferon signaling. *Nat. Immunol.* **15**, 168–176 (2014).
18. A. Göder, T. Ginter, T. Heinzel, S. Stroh, J. Fahrner, A. Henke, O. H. Krämer, STAT1 N-terminal domain discriminatively controls type I and type II IFN signaling. *Cytokine* **144**, 155552 (2021).
19. N. Wenta, H. Strauss, S. Meyer, U. Vinkemeier, Tyrosine phosphorylation regulates the partitioning of STAT1 between different dimer conformations. *Proc. Natl. Acad. Sci.* **105**, 9238–9243 (2008).
20. J. Staab, T. Riebeling, V. Koch, C. Herrmann-Lingen, T. Meyer, The two interfaces of the STAT1 N-terminus exhibit opposite functions in IFN γ -regulated gene expression. *Mol. Immunol.* **67**, 596–606 (2015).
21. A. R. Harrison, G. W. Moseley, The Dynamic Interface of Viruses with STATs. *J. Virol.* **94** (2020).
22. A. A. R. Suarez, N. V. Renne, T. F. Baumert, J. Lupberger, Viral manipulation of STAT3: Evade, exploit, and injure. *PLoS Pathog.* **14**, e1006839 (2018).
23. Z. Chang, Y. Wang, X. Zhou, J.-E. Long, STAT3 roles in viral infection: antiviral or proviral? *Futur. Virol.* **13**, 557–574 (2018).
24. C. Talbot-Cooper, T. Pantelejevs, J. P. Shannon, C. R. Cherry, M. T. Au, M. Hyvönen, H. D. Hickman, G. L. Smith, Poxviruses and paramyxoviruses use a conserved mechanism of STAT1 antagonism to inhibit interferon signaling. *Cell Host Microbe* **30**, 357–372.e11 (2022).
25. B. Wang, S. Thurmond, K. Zhou, M. T. Sánchez-Aparicio, J. Fang, J. Lu, L. Gao, W. Ren, Y. Cui, E. C. Veit, H. Hong, M. J. Evans, S. E. O’Leary, A. García-Sastre, Z. H. Zhou, R. Hai, J. Song, Structural basis for STAT2 suppression by flavivirus NS5. *Nat Struct Mol Biol* **27**, 875–885 (2020).

26. V. T. K. Le-Trilling, S. Banchenko, D. Paydar, P. M. Leipe, L. Binting, S. Lauer, A. Graziadei, R. Klingen, C. Gotthold, J. Bürger, T. Bracht, B. Sitek, R. J. Lebbink, A. Malyshkina, T. Mielke, J. Rappsilber, C. M. Spahn, S. Voigt, M. Trilling, D. Schwefel, Structural mechanism of CRL4-instructed STAT2 degradation via a novel cytomegaloviral DCAF receptor. *EMBO J.* **42**, e112351 (2023).
27. K. Oda, Y. Matoba, T. Irie, R. Kawabata, M. Fukushi, M. Sugiyama, T. Sakaguchi, Structural Basis of the Inhibition of STAT1 Activity by Sendai Virus C Protein. *J Virol* **89**, 11487–11499 (2015).
28. K. Brzózka, S. Finke, K.-K. Conzelmann, Inhibition of interferon signaling by rabies virus phosphoprotein P: activation-dependent binding of STAT1 and STAT2. *J. Virol.* **80**, 2675–83 (2006).
29. A. Vidy, M. Chelbi-Alix, D. Blondel, Rabies Virus P Protein Interacts with STAT1 and Inhibits Interferon Signal Transduction Pathways. *J. Virol.* **79**, 14411–14420 (2005).
30. O. Delmas, E. C. Holmes, C. Talbi, F. Larrous, L. Dacheux, C. Bouchier, H. Bourhy, Genomic diversity and evolution of the lyssaviruses. *Plos One* **3**, e2057 (2008).
31. K. Hampson, L. Coudeville, T. Lembo, M. Sambo, A. Kieffer, M. Attlan, J. Barrat, J. D. Blanton, D. J. Briggs, S. Cleaveland, P. Costa, C. M. Freuling, E. Hiby, L. Knopf, F. Leanes, F.-X. Meslin, A. Metlin, M. E. Miranda, T. Müller, L. H. Nel, S. Recuenco, C. E. Rupprecht, C. Schumacher, L. Taylor, M. A. N. Vigilato, J. Zinsstag, J. Dushoff, G. A. for R. C. P. for R. Prevention, Estimating the global burden of endemic canine rabies. *Plos Neglect Trop D* **9**, e0003709 (2015).
32. W. H. Organization, WHO Expert Consultation on Rabies. Second report. *World Heal. Organ. Tech. Rep. Ser.*, 1–139, back cover (2013).
33. L. Wiltzer, K. Okada, S. Yamaoka, F. Larrous, H. V. Kuusisto, M. Sugiyama, D. Blondel, H. Bourhy, D. A. Jans, N. Ito, G. W. Moseley, Interaction of Rabies Virus P-Protein With STAT Proteins is Critical to Lethal Rabies Disease. *J. Infect. Dis.* **209**, 1744–1753 (2014).
34. N. ITO, G. W. MOSELEY, M. SUGIYAMA, The importance of immune evasion in the pathogenesis of rabies virus. *J. Vet. Méd. Sci.* **78**, 1089–1098 (2016).
35. K. Brzózka, S. Finke, K.-K. Conzelmann, Identification of the Rabies Virus Alpha/Beta Interferon Antagonist: Phosphoprotein P Interferes with Phosphorylation of Interferon Regulatory Factor 3. *J. Virol.* **79**, 7673–7681 (2005).
36. L. Wiltzer, F. Larrous, S. Oksayan, N. Ito, G. A. Marsh, L. F. Wang, D. Blondel, H. Bourhy, D. A. Jans, G. W. Moseley, Conservation of a unique mechanism of immune evasion across the Lyssavirus genus. *J Virol* **86**, 10194–9 (2012).
37. Md. A. Hossain, F. Larrous, S. M. Rawlinson, J. Zhan, A. Sethi, Y. Ibrahim, M. Aloï, K. G. Lieu, Y.-F. Mok, M. D. W. Griffin, N. Ito, T. Ose, H. Bourhy, G. W. Moseley, P. R. Gooley, Structural Elucidation of Viral Antagonism of Innate Immunity at the STAT1 Interface. *Cell Reports* **29**, 1934–1945.e8 (2019).
38. N. Ito, G. W. Moseley, D. Blondel, K. Shimizu, C. L. Rowe, Y. Ito, T. Masatani, K. Nakagawa, D. A. Jans, M. Sugiyama, Role of Interferon Antagonist Activity of Rabies Virus Phosphoprotein in Viral Pathogenicity. *J. Virol.* **84**, 6699–6710 (2010).
39. A. Vidy, J. E. Bougrini, M. K. Chelbi-Alix, D. Blondel, The nucleocytoplasmic rabies virus P protein counteracts interferon signaling by inhibiting both nuclear accumulation and DNA binding of STAT1. *J. Virol.* **81**, 4255–63 (2007).

40. C. Deffrasnes, M.-X. Luo, L. Wiltzer-Bach, C. T. David, K. G. Lieu, L.-F. Wang, D. A. Jans, G. A. Marsh, G. W. Moseley, Phenotypic Divergence of P Proteins of Australian Bat Lyssavirus Lineages Circulating in Microbats and Flying Foxes. *Viruses* **13**, 831 (2021).
41. K. G. Lieu, A. Brice, L. Wiltzer, B. Hirst, D. A. Jans, D. Blondel, G. W. Moseley, The rabies virus interferon antagonist P protein interacts with activated STAT3 and inhibits Gp130 receptor signaling. *J. Virol.* **87**, 8261–5 (2013).
42. G. Manokaran, M. D. Audsley, H. Funakoda, C. T. David, K. A. Garnham, S. M. Rawlinson, C. Deffrasnes, N. Ito, G. W. Moseley, Deactivation of the antiviral state by rabies virus through targeting and accumulation of persistently phosphorylated STAT1. *PLoS Pathog.* **18**, e1010533 (2022).
43. J. Braunstein, S. Brutsaert, R. Olson, C. Schindler, STATs Dimerize in the Absence of Phosphorylation*. *J Biol Chem* **278**, 34133–34140 (2003).
44. X. Chen, R. Bhandari, U. Vinkemeier, F. van den Akker, J. E. Darnell, J. Kuriyan, A reinterpretation of the dimerization interface of the N-terminal Domains of STATs. *Protein Sci* **12**, 361–365 (2003).
45. M. Zhong, M. A. Henriksen, K. Takeuchi, O. Schaefer, B. Liu, J. ten Hoeve, Z. Ren, X. Mao, X. Chen, K. Shuai, J. E. Darnell, Implications of an antiparallel dimeric structure of nonphosphorylated STAT1 for the activation–inactivation cycle. *P Natl Acad Sci Usa* **102**, 3966–3971 (2005).
46. B. J. Wagner, T. E. Hayes, C. J. Hoban, B. H. Cochran, The SIF binding element confers sis/PDGF inducibility onto the c-fos promoter. *EMBO J.* **9**, 4477–4484 (1990).
47. Z. Huang, H. Liu, J. Nix, R. Xu, C. R. Knoverek, G. R. Bowman, G. K. Amarasinghe, L. D. Sibley, The intrinsically disordered protein TgIST from *Toxoplasma gondii* inhibits STAT1 signaling by blocking cofactor recruitment. *Nat. Commun.* **13**, 4047 (2022).
48. M. I. Ndubuisi, G. G. Guo, V. A. Fried, J. D. Etlinger, P. B. Sehgal, Cellular Physiology of STAT3: Where's the Cytoplasmic Monomer?*. *J Biol Chem* **274**, 25499–25509 (1999).
49. Y. Belo, Z. Mielko, H. Nudelman, A. Afek, O. Ben-David, A. Shahar, R. Zarivach, R. Gordan, E. Arbely, Unexpected implications of STAT3 acetylation revealed by genetic encoding of acetyl-lysine. *Biochimica Et Biophysica Acta Bba - Gen Subj* **1863**, 1343–1350 (2019).
50. J. Li, J. P. Rodriguez, F. Niu, M. Pu, J. Wang, L.-W. Hung, Q. Shao, Y. Zhu, W. Ding, Y. Liu, Y. Da, Z. Yao, J. Yang, Y. Zhao, G.-H. Wei, G. Cheng, Z.-J. Liu, S. Ouyang, Structural basis for DNA recognition by STAT6. *Proc National Acad Sci* **113**, 13015–13020 (2016).
51. Y. Nagano, A. Sugiyama, M. Kimoto, T. Wakahara, Y. Noguchi, X. Jiang, S. Saijo, N. Shimizu, N. Yabuno, M. Yao, P. R. Gooley, G. W. Moseley, T. Tadokoro, K. Maenaka, T. Ose, The Measles Virus V Protein Binding Site to STAT2 Overlaps That of IRF9. *J. Virol.* **94** (2020).
52. A. Sethi, S. M. Rawlinson, A. Dubey, C.-S. Ang, Y. H. Choi, F. Yan, K. Okada, A. M. Rozario, A. M. Brice, N. Ito, N. A. Williamson, D. M. Hatters, T. D. M. Bell, H. Arthanari, G. W. Moseley, P. R. Gooley, Structural insights into the multifunctionality of rabies virus P3 protein. *Proc. Natl. Acad. Sci. United States Am.* **120**, e2217066120 (2023).
53. A. R. Harrison, K. G. Lieu, F. Larrous, N. Ito, H. Bourhy, G. W. Moseley, Lyssavirus P-protein selectively targets STAT3-STAT1 heterodimers to modulate cytokine signalling. *PLoS Pathog.* **16**, e1008767 (2020).
54. A. Sugiyama, T. Nomai, X. Jiang, M. Minami, M. Yao, K. Maenaka, N. Ito, P. R. Gooley, G. W. Moseley, T. Ose, Structural comparison of the C-terminal domain of functionally divergent lyssavirus P proteins. *Biochem. Biophys. Res. Commun.* **529**, 507–512 (2020).

55. S. H. W. Scheres, Processing of Structurally Heterogeneous Cryo-EM Data in RELION. *Methods Enzym.* **579**, 125–57 (2016).
56. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
57. K. Yamashita, M. Wojdyr, F. Long, R. A. Nicholls, G. N. Murshudov, GEMMI and Servalcat restrain REFMAC5. *Acta Crystallogr. Sect. D* **79**, 368–373 (2023).
58. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **66**, 486–501 (2010).
59. A. Brown, F. Long, R. A. Nicholls, J. Toots, P. Emsley, G. Murshudov, Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta Crystallogr. Sect. D* **71**, 136–153 (2015).
60. E. C. Meng, T. D. Goddard, E. F. Pettersen, G. S. Couch, Z. J. Pearson, J. H. Morris, T. E. Ferrin, UCSF ChimeraX : Tools for structure building and analysis. *Protein Sci.* **32**, e4792 (2023).
61. N. Shimizu, T. Mori, Y. Nagatani, H. Ohta, S. Saijo, H. Takagi, M. Takahashi, K. Yatabe, T. Kosuge, N. Igarashi, BL-10C, the small-angle x-ray scattering beamline at the photon factory. *AIP Conf. Proc.* **2054**, 060041 (2019).
62. P. Bernadó, N. Shimizu, G. Zaccai, H. Kamikubo, M. Sugiyama, Solution scattering approaches to dynamical ordering in biomolecular systems. *Biochimica et Biophys. acta Gen. Subj.* **1862**, 253–274 (2017).
63. N. Shimizu, K. Yatabe, Y. Nagatani, S. Saijo, T. Kosuge, N. Igarashi, Software development for analysis of small-angle x-ray scattering data. *AIP Conf. Proc.* **1741**, 050017 (2016).
64. K. Yonezawa, M. Takahashi, K. Yatabe, Y. Nagatani, N. Shimizu, MOLASS: Software for automatic processing of matrix data obtained from small-angle X-ray scattering and UV–visible spectroscopy combined with size-exclusion chromatography. *Biophys. Physicobiology* **20**, e200001 (2023).
65. M. V. Petoukhov, P. V. Konarev, A. G. Kikhney, D. I. Svergun, ATSAS 2.1 – towards automated and web-supported small-angle scattering data analysis. *J. Appl. Crystallogr.* **40**, s223–s228 (2007).
66. T. D. Grant, Ab initio electron density determination directly from solution scattering data. *Nat. Methods* **15**, 191–193 (2018).

Acknowledgments: We thank N. Shimizu for assistance in SAXS data collection at Photon Factory. **Funding:** Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research KAKENHI Grants 17K07296 (to T.O.); Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research KAKENHI 21H02408 (to Y.S. and T.O.); Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research KAKENHI 21J12357 (to A.S.); Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research KAKENHI JP20H05873 (to K.M.); Japan Science and Technology Agency (JST) SPRING JPMJSP2119 (to M.M.); Japan Science and Technology Agency (JST) FOREST JPMJFR214S (to Y.S.); Takeda Science Foundation (K.M.); Hokkaido University Biosurface project (K.M.); Japan Agency for Medical Research and Development (AMED) JP17am0101093 (to K.M.); Japan Agency for Medical Research and Development (AMED) JP20ae0101047 (to K.M.); Japan Agency for Medical Research and Development (AMED) JP21fk0108463 (to K.M.); Japan Agency for Medical Research and Development (AMED) JP22ama121037 (to K.M.); Japan Agency for Medical Research and Development (AMED) 22gm1810004 (to K.M.); Japan Agency for Medical Research and Development (AMED) JP223fa627005 (to K.M.); Hokkaido University-Hitachi

Joint Cooperative Support Program for Education and Research (to A.S.); Platform Project for Supporting in Drug Discovery and Life Science Research from AMED JP21am0101072 (support no. 1941) (to T.O.); Platform Project for Supporting in Drug Discovery and Life Science Research from AMED JP23ama121001 (support no. 5925) (to T.O.); and Global Facility Center and Pharma Science Open Unit funded by MEXT under "Support Program for Implementation of New Equipment Sharing System" in Hokkaido University. **Author contributions:** Conceptualization: A.S. and T.O. Methodology: A.S., M.M., K.U., S.I.-I., Y.T., S.X., S.T., M.S., T.N., X.J., S.K., M.H., M.Y., Y.S., and T.O. Investigation: A.S., M.M., K.U., Y.S., and T.O. Visualization: A.S., M.M., Y.S., and T.O. Supervision: G.W.M., P.R.G., Y.S., and T.O. Writing—original draft: A.S., Y.S., and T.O. Writing—review and editing: G.W.M., P.R.G., and T.O. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The cryo-EM density maps and corresponding atomic models have been deposited in the Electron Microscopy Data Bank (EMDB) and the Protein Data Bank (PDB) with accession numbers EMD-39679 and 8YYU for the tetramer and EMD-39680 and 8YYV for the dimer. Additionally, the raw electron microscopy movies have been archived in the Electron Microscopy Public Image Archive (EMPIAR) with accession number EMPIAR-12034. The SAXS information is deposited in the small angle scattering biological data bank (SASBDB; <https://www.sasbdb.org/>). Deposition IDs are SASDUL4 for pY-STAT1, SASDUN4 for pY-STAT1/DNA, and SASDUM4 for pY-STAT1/P-CTD. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Fig. 1. Domain structure and oligomeric state of tyrosine-phosphorylated STAT1 (pY-STAT1).

(A) Schematic of the domain structure of human STAT1 with conserved domains and Tyr⁷⁰¹ (to be phosphorylated) is indicated. Note that the core region comprises the coiled-coil, DNA binding, linker, SH2 domains, and a part of TA domain. N indicates the N-terminal domain and C indicates the C-terminal region. **(B)** Oligomerization states of STAT1, with the pY/SH2 domain interactions and inter-N-terminal domain interactions shown for U-STAT1 (dimer), U-ΔN STAT1 (monomer), pY-STAT1 (tetramer), pY-ΔN STAT1 (dimer), pY-STAT1 ΔC (tetramer), pY-STAT1 F77A/L78A (dimer) and the pY-STAT1 core (dimer). **(C)** SEC-MALS chromatograms of pY-STAT1 (red), pY-STAT1/P-CTD (magenta), and pY-STAT1/DNA (blue). Traces are displayed with the light scattering (left axis) for elution volumes between 12 and 15 ml. Molecular weights of pY-STAT1 (red) and pY-STAT1/P-CTD (magenta) are also indicated by the horizontal lines with respect to the value of the right axis. The molecular weight of pY-

STAT1/DNA was not calculated because of the difference of the optical constant of protein and DNA. The elution volumes of each peak position are indicated. **(D)** The SEC chromatogram of wild pY-STAT1 (orange) and the pY-STAT1 Y77A/L78A (blue) are shown, indicating the effects of the mutations on the oligomerization state. The chromatogram of the protein standards is shown in gray. Retention volumes of the STAT proteins and the protein standard proteins are indicated with their molecular weights. The vertical value is arbitrary.

Fig. 2. Structure of pY-STAT1 in complex with DNA and the predicted location of the NTD of pY-STAT1. **(A)** EM density (3.07 Å resolution, gray) and the model of pY-STAT1 dimer in complex with a double-stranded DNA molecule. The view is parallel to the two-fold axis used in the C2 averaging. The different pY-STAT1 protomers are colored blue and green. The backbone of DNA is colored in red. **(B)** EM density (3.67 Å resolution, gray) and the model of pY-STAT1 tetramer in complex with two double-stranded DNA molecules. The different pY-STAT1 protomers are colored-blue, green, magenta, and yellow. The backbones of the DNA molecules are colored red. The view is from the two-fold axis used in the C2 averaging. **(C)** The displacement between the EM structure (green and light blue) and the reported crystal structure of the pY-STAT1 core/-DNA complex (PDBID, 1BF5; yellow and magenta) based on the superimposition of the left side protomers. **(D)** Ribbon diagram of the tetrameric pY-STAT1 in complex with DNA (this study) to represent inter-dimer interaction. **(E)** Zoomed view of the dimer-dimer interface of the LD of the tetrameric pY-STAT1 shown in (D). The 2-fold rotation axis is parallel to the surface of the page. The amino acid residues contributing to the interaction are labeled. **(F)** Electrostatic potential surface representation of tetrameric pY-STAT1 (red negative, blue positive). The bound two DNA molecules are shown in orange coils. **(G)** Low-level contoured EM density for pY-

STAT1/DNA from the direction of the two-fold rotation axis. Each protomer is colored as in (D). Residual densities, around the CCDs, are predicted to be the location of the NTDs (indicated with white arrows). (H) NTD dimer location from the crystal structure of U-STAT1 (PDBID, 1YVL) are indicated with white arrows. They do not fit to the residual density. U-STAT1 structure (one protomer of U-STAT1 of the antiparallel dimer (PDBID, 1YVL)) containing the NTD is superimposed on protomers (magenta and blue) of the tetrameric EM structure based on the STAT1 core region. Two NTDs are shown by white arrows (magenta and blue). The EM density contouring and the direction of the view is the same as (G). (I) Manual modeled two NTD dimers in the residual EM density (cyan). The EM density contouring and the direction of the view is the same as (G). Each protomer of the pY-STAT1 core is labeled as A, B, C and D-chains. (J) The location of the NTD dimer tethering the protomers A and C. The C-terminus (Gln¹²⁶) of the NTD of the A chain (cyan) and the N-terminus of the A chain (magenta) CCD (Met¹³⁵) are indicated. The C-terminus (Gln¹²⁶) of the C chain in the same NTD dimer (cyan) and the N-terminus of C chain (magenta) CCD (Met¹³⁵) are also indicated. The distances between each pair are 15 Å and 10 Å, respectively.

Fig. 3. Solution state of pY-STAT1 and its complexes analyzed by SEC-SAXS. (A to C) Experimental SAXS profiles of (A) pY-STAT1, (B) pY-STAT1/DNA, and (C) pY-STAT1/P-CTD. Guinier plots for experimental SAXS curves are shown in insets. (D) Pair distribution $P(r)$ function for the experimental SAXS curves of pY-STAT1 (blue), pY-STAT1/DNA (green), and pY-STAT1/P-CTD (purple).

Fig. 4. Identification of critical regions of STAT1 for its interaction with P-CTD. (A) Fluorescence polarization analysis of FlAsH-EDT₂ labeled P-CTD and pY-STAT1 (green), pY-ΔN-STAT1 (purple), U-STAT1 (red), and U ΔN-STAT1(blue). (B) Fluorescence polarization analysis-of P-CTD with pY-STAT1 full (green), pY-STAT1 core (gray), pY-ΔN-STAT1 (purple), pY-STAT1-ΔC (red), or pY-STAT1 Y77A/L78A (blue). Data in (A) and (B) are shown in anisotropy value as mean ± SEM from three independent experiments (n = 3) for each construct. Data were analyzed with GraphPad Prism using the one site specific binding least squares fit function to calculate the $B_{\max}$ and K_D values (Table 1).

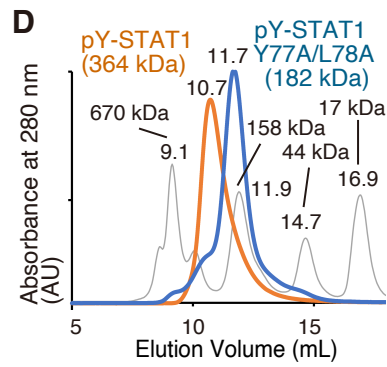
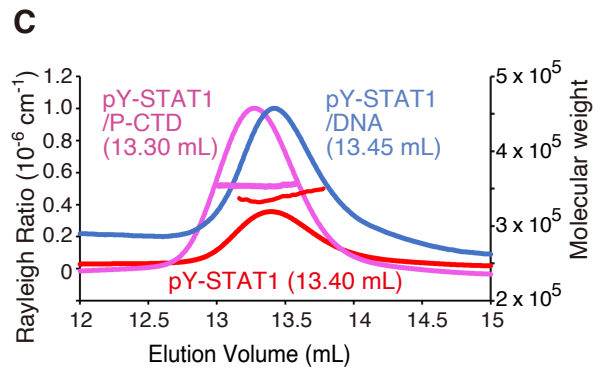
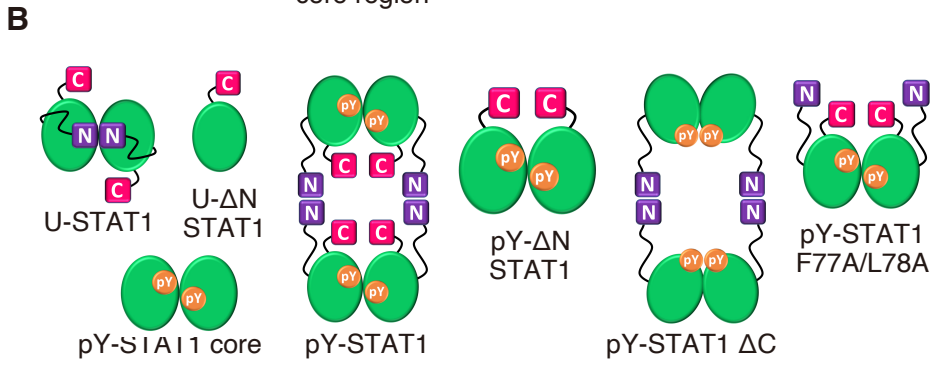
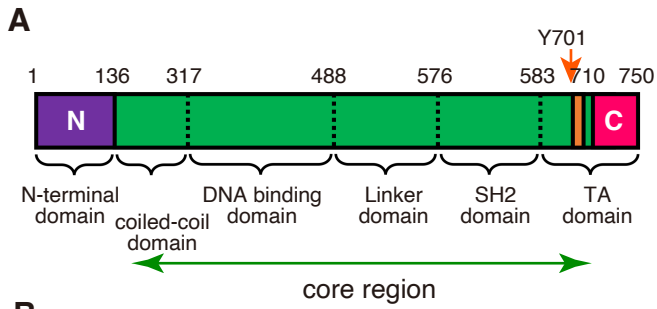
Fig. 5. Model of the cryo-EM structure of pY-STAT1/DNA converted to the binding of a single linear strand of DNA harboring two GAS sites. (A) A conversion model from the cryo-EM STAT1 tetramer structure bound to 18-mer DNA (left, from Fig. 2 (I)) to a STAT1 tetramer bound to a single double-stranded DNA molecule containing tandem GAS sites (right). The naming of each protomer of the pY-STAT1 core is the same as in Fig. 2 (I). In the right structure, two pY-STAT1 dimer cores bind to a DNA molecule with two GAS sequence recognition sites; the two GAS sequences with a 10-bp gap were generated. (B) The-NTD tethered model of tetrameric pY-STAT1 proposed previously (20). The protomers, A and D, are tethered by an NTD dimerization, whereas the NTDs of B and C are separated. The DNA-bound model of tetrameric pY-STAT1 cores is the same as in (A) (right). (C) The distances are shown to predict which two pY-STAT1 protomers are tethered by one NTD dimer, respectively. The model is the same as in (A) (right). Two numbers indicate the distances between the two N-termini of CCD of protomers A (shown in magenta) and C (shown in blue), or protomers A and D (shown in green). (D) The distance between the two C-termini of the NTD dimer (cyan), each of which is linked to an N-

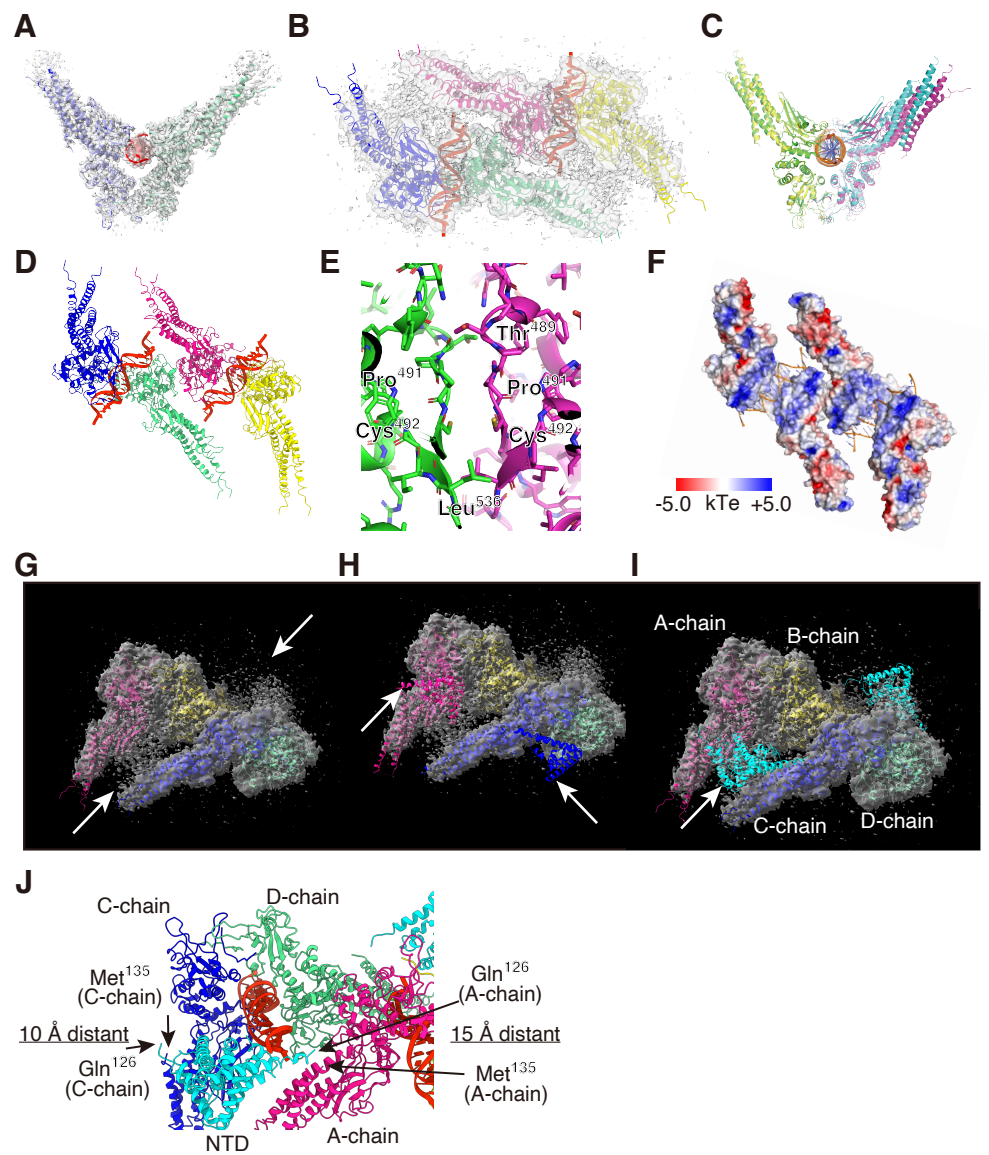
terminus of a CCD, is calculated as 71 Å. **(E)** Newly proposed model in this study in which two NTD dimers are tethering two pY-STAT1 dimers (A and C, B and D). The conversion from the EM tetrameric structure to this model is possible without the dissociation of NTD dimers as proposed previously (20).

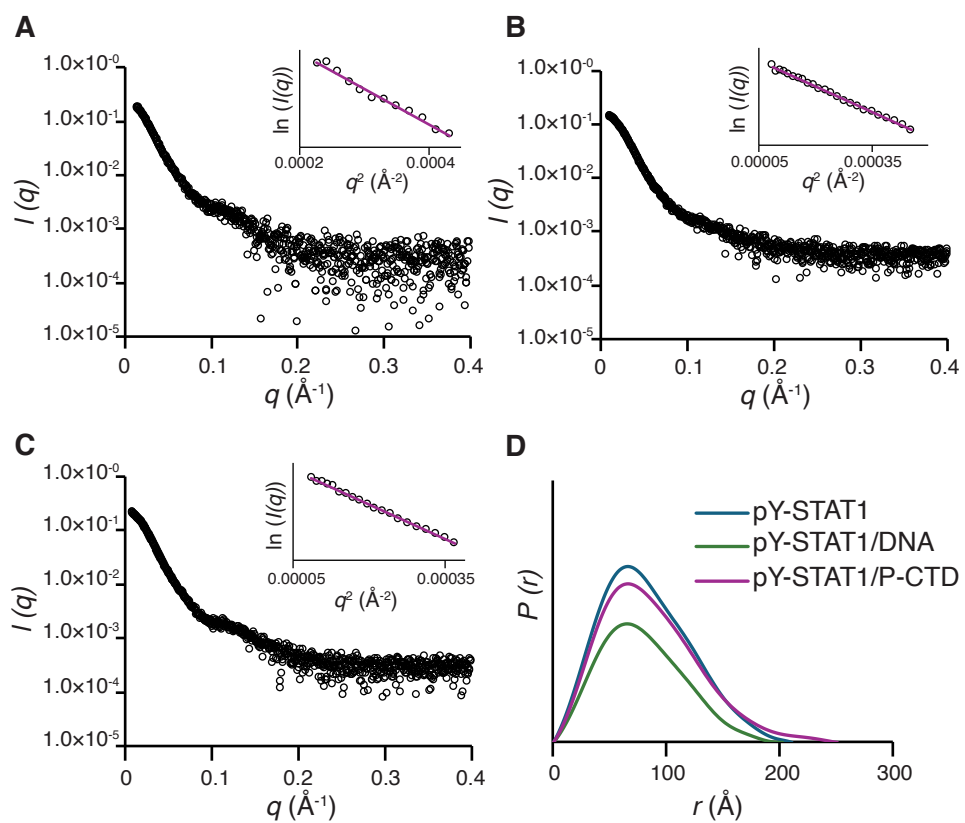
Fig. 6. Predicted structural model of P-CTD-bound pY-STAT1. **(A)** Each protomer of the pY-STAT1 core is represented with a semi-transparent surface model (A chain, magenta; B chain, yellow; C chain, blue; D chain, green). The two NTD dimers as modeled in Fig. 2 (I) are in cyan. Two P-CTDs in putative primary binding sites are shown with grey surface model on which Leu²⁷⁶ and Leu²⁷⁷ of the P-CTD are shown in orange. The putative two secondary binding sites, which are two-fold symmetry related to the primary two binding sites as a STAT1 dimer unit, are indicated as dotted ovals. **(B)** The modeling of P-CTDs at the primary binding sites as shown in (A), rotated by 90° around x. **(C)** Zoomed view of the primary binding site. P-CTD (grey) is interacting with the A chain (magenta), the B chain (yellow), and the D chain (green) of the pY-STAT1 core. Leu²⁷⁶ and Leu²⁷⁷ are shown in orange, and His²⁰³, Ala²⁰⁵ are in blue, and Asp²³⁵ is in red. **(D)** The dimer representation of pY-STAT1 (AB chain, magenta and yellow) with two-fold symmetry-related P-CTDs (the primary and the secondary binding sites). **(E)** The same model as (C), rotated by 90° around x.

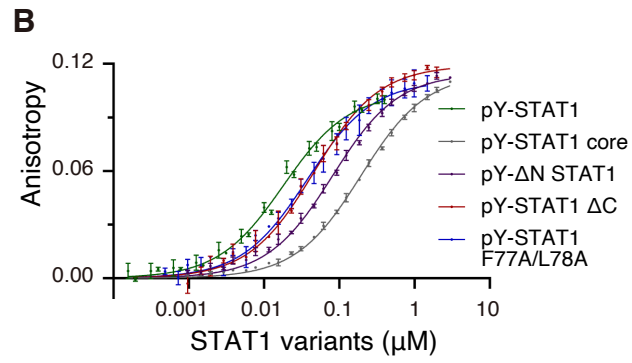
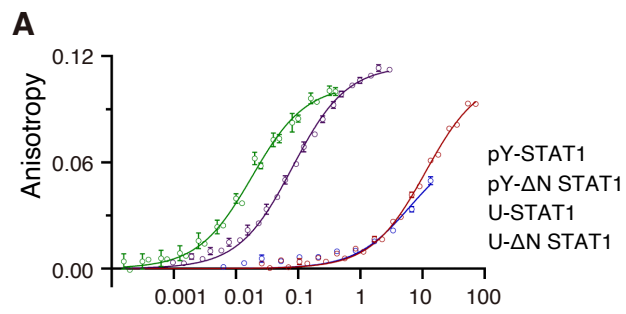
Table 1. Interaction analysis of STAT1 with P-CTD as determined by FP. Summary table of data shown in Fig. 4, A and B, with mean $B_{\max}$ and K_D values calculated for the interaction of P-CTD with indicated STAT1 constructs. The 95% confidence interval (CI) for each of the K_D values is provided, indicating the variability on the estimated effects of binding.

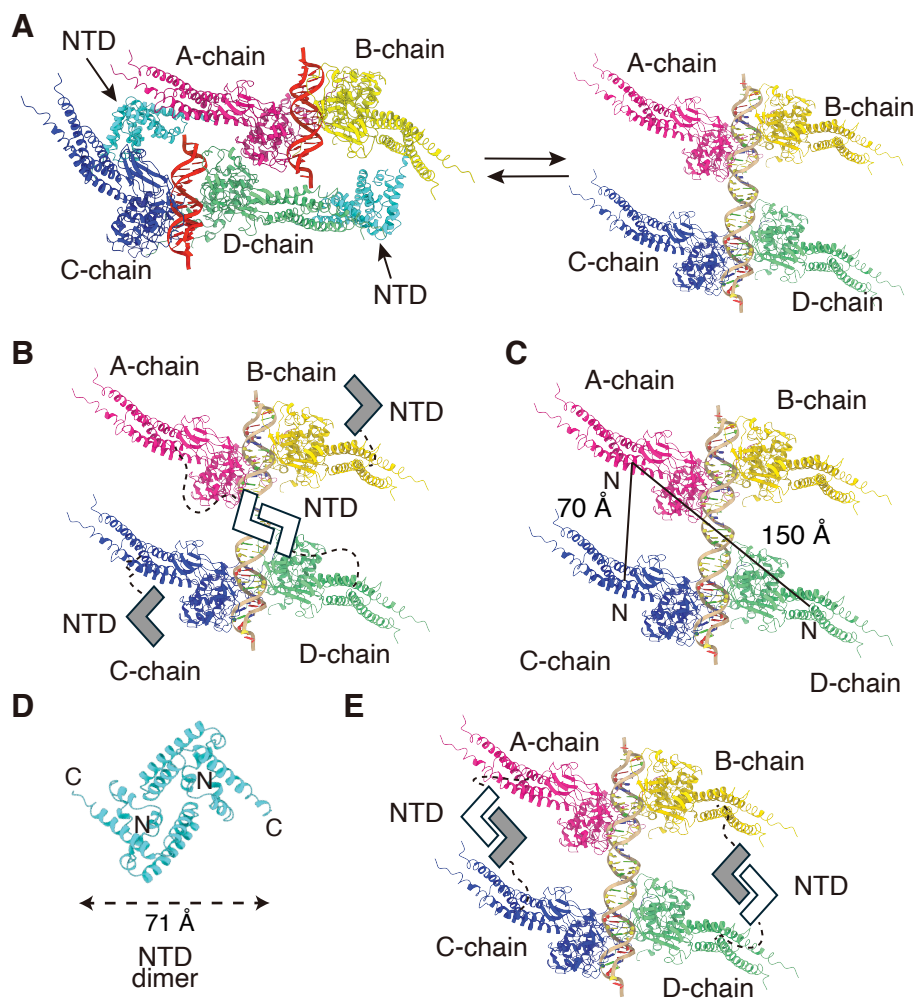
STAT variants	$B_{\max}$	95% CI (profile likelihood)	K_D (nM)	95% CI (profile likelihood)
U-STAT1	0.1089	0.1056 - 0.1124	11,180	10,230 - 12,220
U-ΔN STAT1	0.06921	0.05543 - 0.09203	6198	3,710 - 10,920
pY-STAT1	0.1038	0.1005 - 0.1072	18.10	15.98 - 20.50
pY-ΔN STAT1	0.1146	0.1127 - 0.1166	81.44	75.80 - 87.49
pY-STAT1 ΔC	0.1198	0.1163 - 0.1220	46.42	41.73 - 51.61
pY-STAT1 F77A/L78A	0.1106	0.1063 - 0.1151	37.6	31.77 - 44.42
pY-STAT1 core	0.1148	0.1130 - 0.1165	194.2	184.2 - 204.7

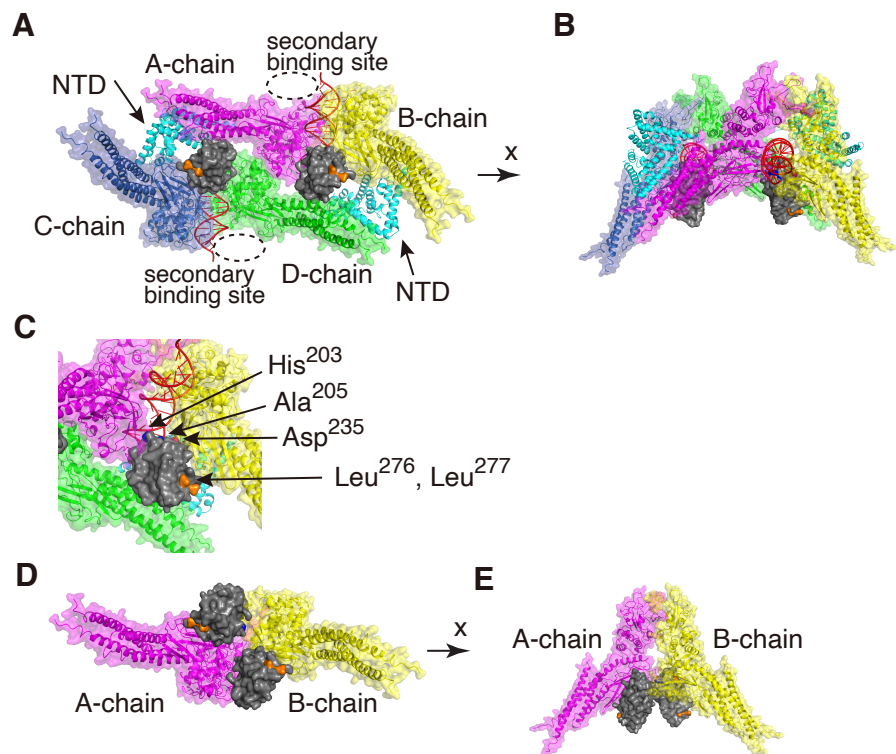












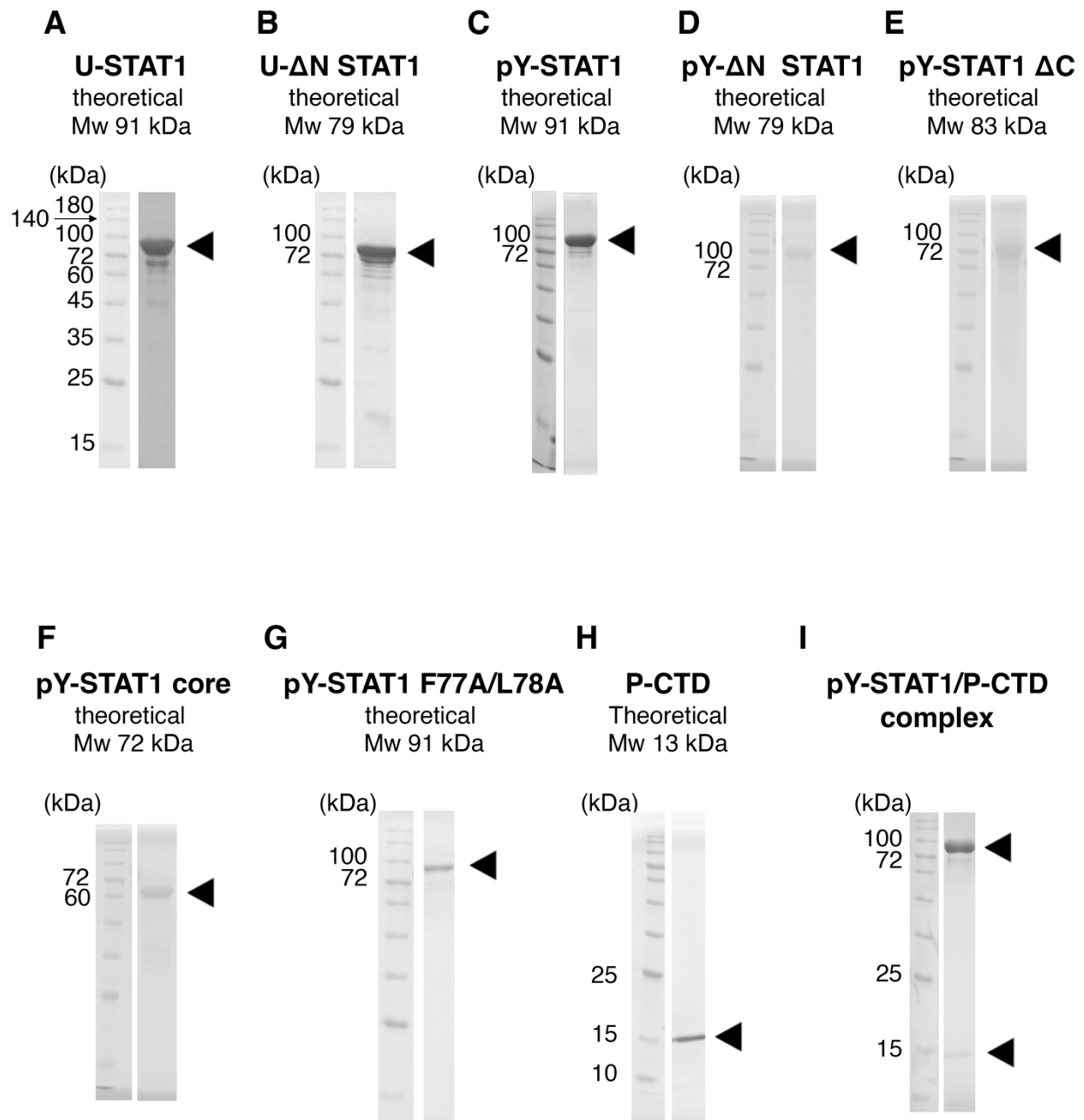


Fig. S1. SDS-PAGE analysis of purified STAT1 variants. (A to I) SDS-PAGE analysis of (A) U-STAT1, (B) U-ΔN STAT1, (C) pY-STAT1, (D) pY-ΔN STAT1, (E) pY-STAT1-ΔC, (F) pY-STAT1 core, (G) pY-STAT1 F77A/L77A, (H) RABV (*Nishigahara* strain)-P-CTD, and (I) pY-STAT1 in complex with RABV-RVP-CTD after isolation by size-exclusion chromatography. The molecular weight (Mw) of each SDS-PAGE marker protein is shown.

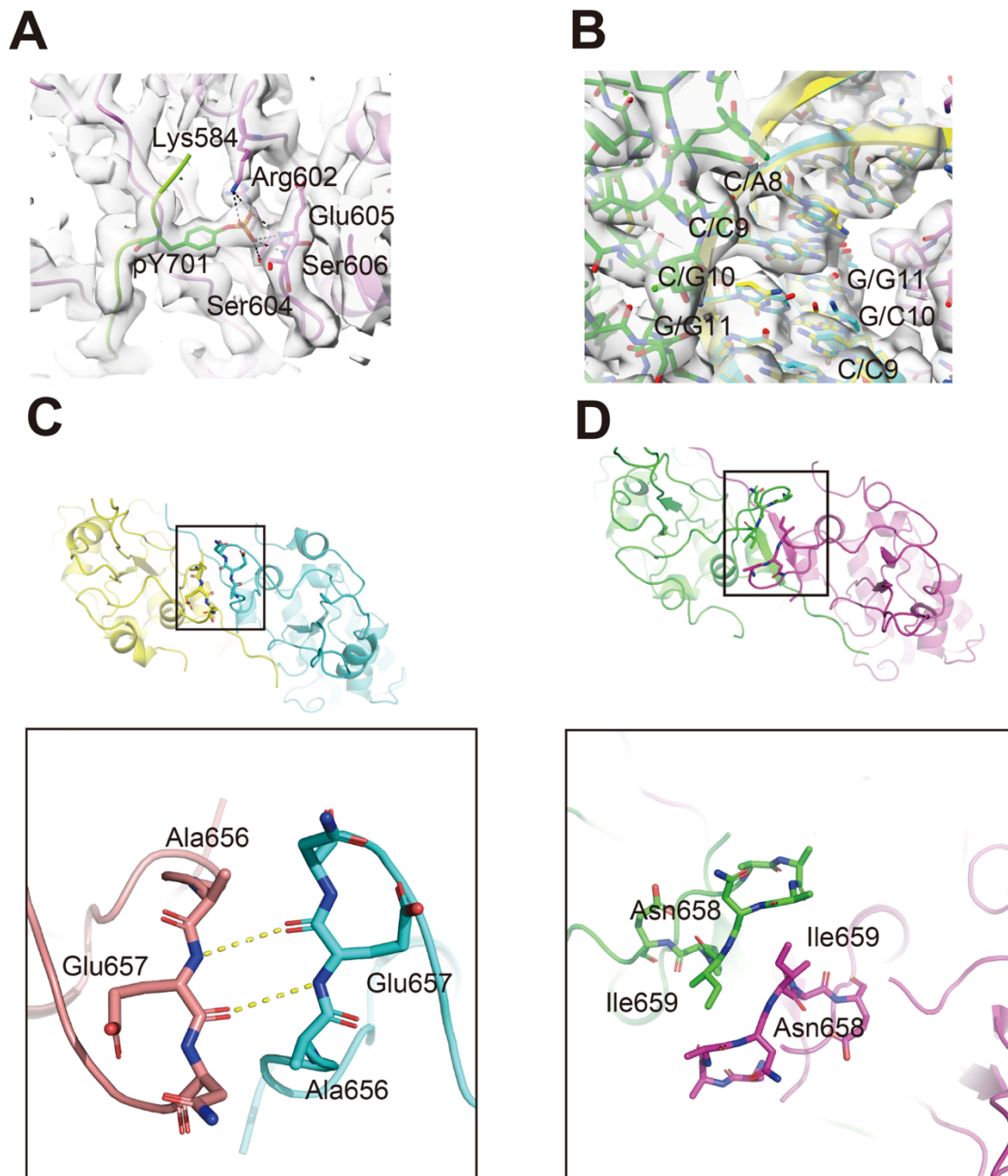
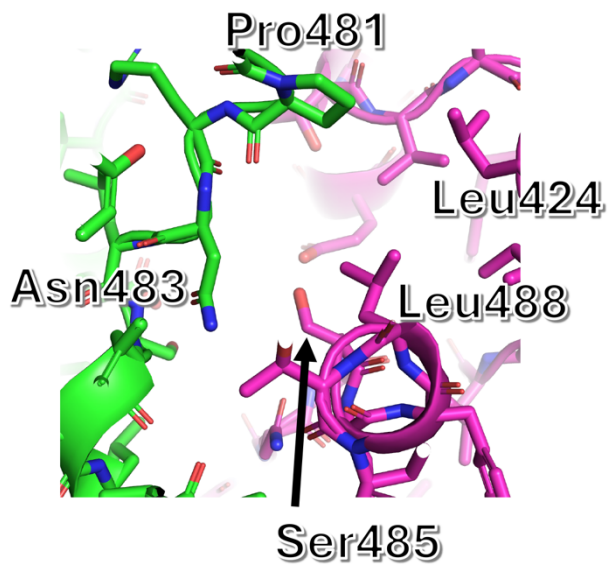


Fig. S2. Characterization of the dimer interface of the pY-STAT1/DNA complex. (A) EM density (dimer analysis in this study) of phosphorylated Tyr⁷⁰¹ in one protomer (green with orange-colored phospho moiety) recognized by the reciprocal SH2 domain of another monomer (magenta). The side chains of Lys⁵⁸⁴, Arg⁶⁰², and Ser⁶⁰⁴ of the magenta protomer are labeled, together with the main chain amide nitrogen atoms of Glu⁶⁰⁵ and Ser⁶⁰⁶. (B) The DNA molecule modeled in the EM density (dimer analysis in this study) of the pY-STAT1 dimer (green and magenta). DNA bases representing two possibilities (each occupancy value of their atoms is assigned as 0.5) are labeled and shown in yellow and cyan. For the labelling of the DNA bases, those from the yellow chain are written before the backslash and those from the cyan chain after the backslash (yellow/cyan). (C and D) The difference of the SH2 domain-mediated interactions between the previous crystal structure (PDB ID: 1BF5) in (C) and our current EM structure in (D) from the same orientation. The magnified views of the interaction with amino acid residues involved are shown in the insets.

A



B

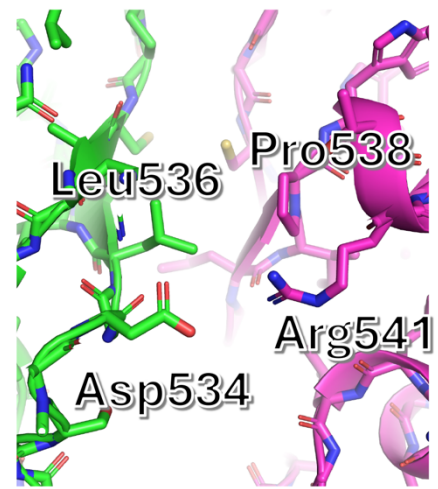


Fig. S3. Magnified view of the dimer-dimer interface of the DBD. (A and B) Expansion of the regions of the DBD (A) and LD (B) shown in Fig. 2D. Additional regions of the LD that form the tetramer are shown in Fig. 2E. The different pY-STAT1 protomers related to C2 symmetry is colored green and magenta.

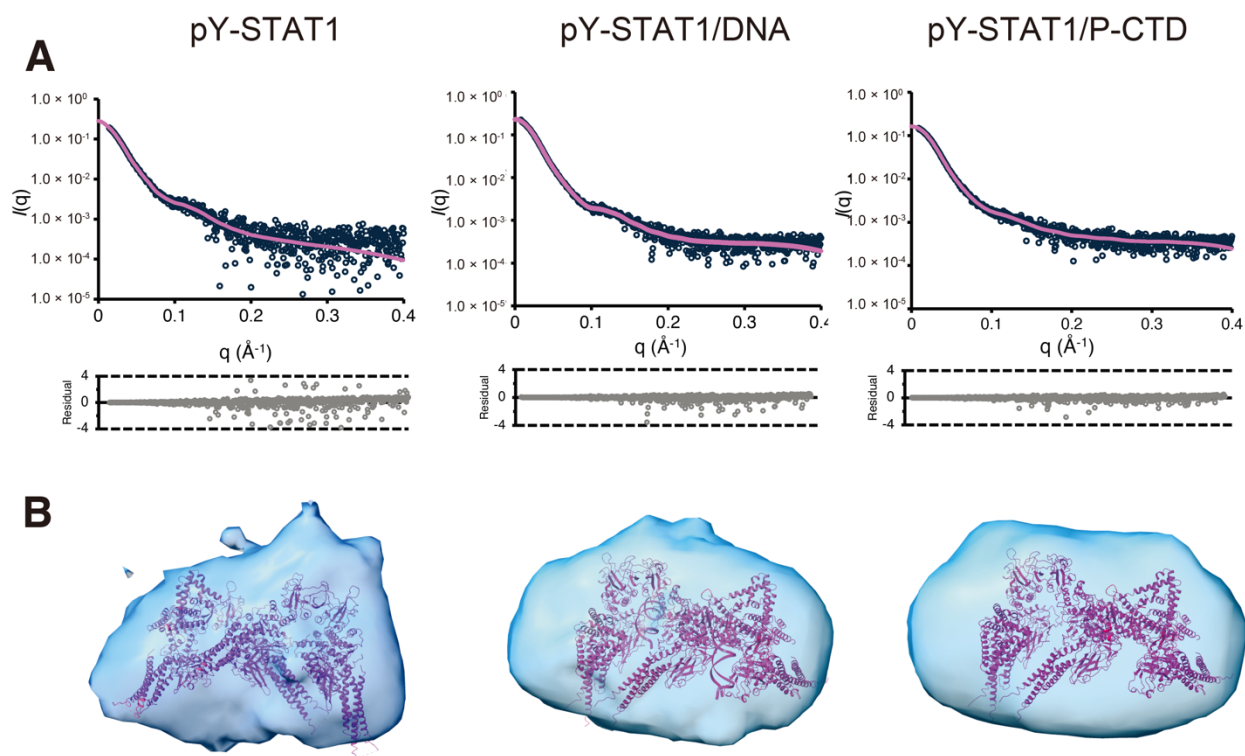


Fig. S4. SAXS analysis of pY-STAT1, and pY-STAT1/DNA, and pY-STAT1/P-CTD. (A) Top: Calculated scattering intensities of the indicated DENSS densities (solid lines) against the experimental scattering curves (dotted plots). Bottom: Plots of the residuals between experimental and calculated scattering. (B) Electron densities for the three molecular species calculated with DENSS from the SEC-SAXS experiments of pY-STAT1 (left), pY-STAT1/DNA (middle), and pY-STAT1/P-CTD (right). The models were fitted in the DENESS density with ChimeraX (60)

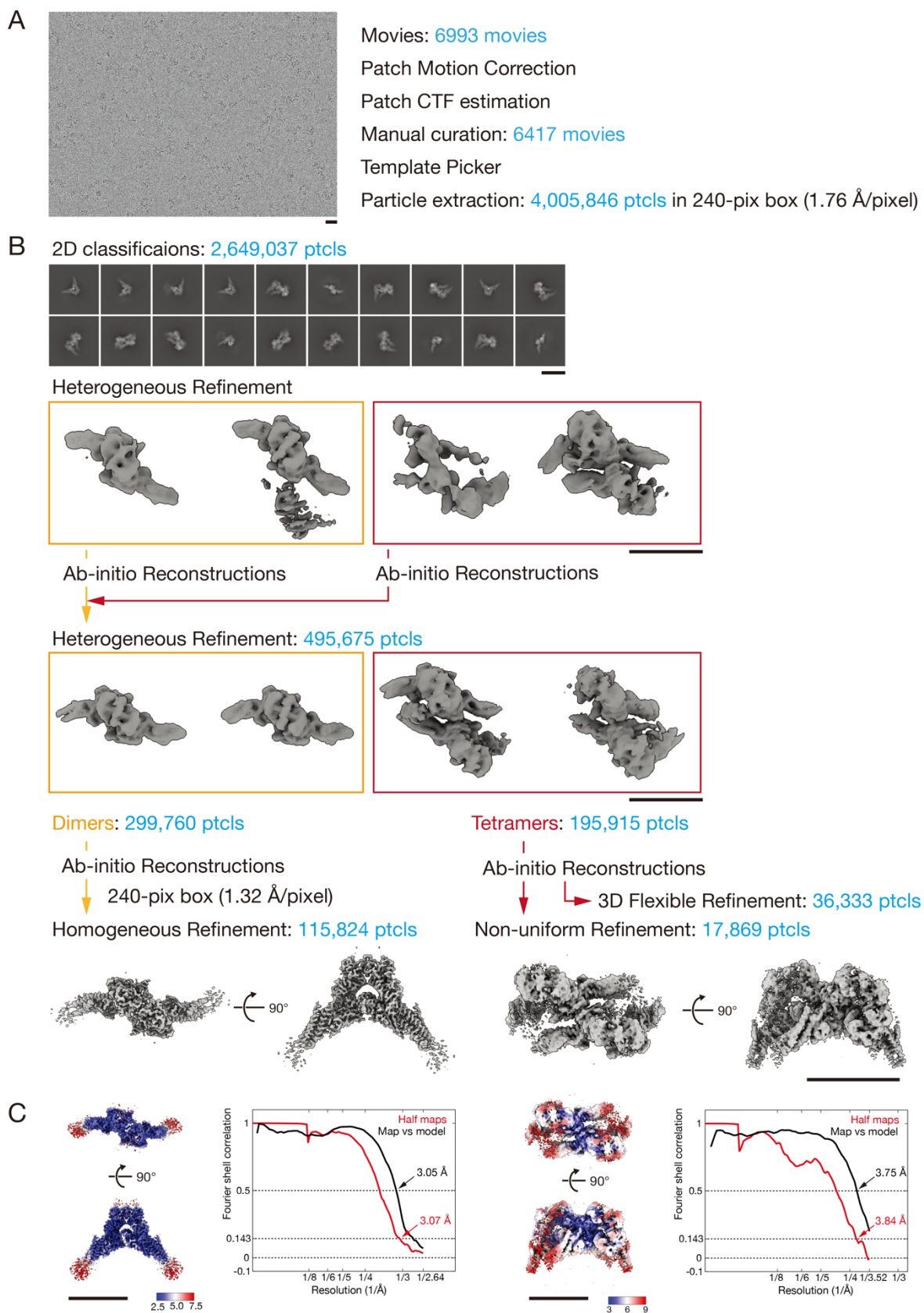


Fig. S5.

Fig. S5. Cryo-EM image processing workflow for the STAT1 dimer and tetramer. (A) A representative cryo-EM micrograph and the preprocessing steps undertaken, illustrating the initial stages of data acquisition and preparation for single-particle analysis. (B) The subsequent image-processing workflow for single-particle analysis, highlighting the sequence of computational steps from 2D classification to 3D reconstruction. (C) The local resolution maps and cross-validation Fourier Shell Correlation (FSC) curves, comparing half-map reconstructions for internal consistency and map-vs.-model validation to assess the resolution and accuracy of the final atomic models for the STAT1 dimer (left) and tetramer (right). Scale bar, 100 Å.

Table S1. SAXS data collection and analysis of pY-STAT1 complexes.

Sample details	pY-STAT1	P-CTD
Organism	<i>Homo sapiens</i> <i>E. coli</i> expressed	Rabies virus nishigahara strain <i>E. coli</i> expressed
Source		
Uniprot ID	P42224	Q9IPJ8

Sample details	pY-STAT1	pY-STAT1/RVP-CTD	pY-STAT1/DNA
SASDBD ID	SASDUL4	SASDUM4	SASDUN4
Extinction (A_{280})	1.259	1.255	2.180
P.S.V ($\text{cm}^3 \text{g}^{-1}$)	0.737	0.734*	0.733
Contrast (cm^{-2})	2.845	2.864*	2.906
Theoretical M_w (Da)	349278	381127.94*	371554
Initial conc. (mg/ml)	6.8	9.2	19†
Injection volume (μl)	155	130	165
Conc. range (mg/ml)	0.255-0.503	0.514-1.02	0.467-0.905
Concentration method	UV-Visible spectroscopy		
Solvent composition	10 mM HEPES-NaOH pH 7.4, 150 mM NaCl, 3% glycerol, 2 mM DTT		

*Stoichiometry is assumed as 4:2 †Approximate value

SAS data collection parameters			
Sample	pY-STAT1	pY-STAT1/P-CTD	pY-STAT1/DNA
Source, instrument		Photon Factory BL-10C	
Wavelength ($\text{\AA}$)		1.0	
Camera length (mm)		2081	
Beam geometry (μm)		V180×H630, Bent cylindrical mirror + 2 slits + 1 pinhole	
Detector		PILATUS3 2M (Dectris)	
q -measurement range ($\text{\AA}^{-1}$)		0.0149-0.5460	
Absolute scaling method		Comparison with scattering from pure H_2O	
Basis for normalization to constant counts		Normalized to incident intensity by μ ion chamber	
Method for monitoring radiation damage		data frame-by-frame comparison	
Exposure time		20 s	
No. of images	309	279	252
Sample temperature (K)		293	
Flow rate of HPLC during exposure (ml/min)		0.05	

Software for SAS data analysis	
SAXS data reduction	SAngher 2.1.65
Basic analysis (Guinier, P(r), molecular mass)	MOLASS 1.0.11, GNOM (ATSAS)
Ab initio modelling	DENSS with DENSSWeb

Structural parameters			
Guinier Analysis			
R_g (Å)	62.3 ± 1.2	65.2 ± 0.414	60.7 ± 0.328
$I(0)$ (cm ⁻¹)	0.247 ± 0.0041	0.229 ± 0.00105	0.16 ± 0.000644
Q range	0.014 to 0.401	0.00779 to 0.401	0.00883 to 0.401
$\max Q \times R_g$	1.3	1.26	1.29
P(r) analysis			
$I(0)$ (cm ⁻¹)	0.251 ± 0.00691	0.227 ± 0.00191	0.16 ± 0.00249
R_g (Å)	64.5 ± 1.97	64.8 ± 0.489	61.3 ± 1.28
$D_{\max}$ (Å)	211	194	203
Q -range (Å ⁻¹)	0.014 to 0.401	0.078 to 0.1225	0.0088 to 0.1314
M from volume of correlation (*)	368790 (1.056)	398538 (1.046)	313119 (0.843)
Porod volume (Å ⁻³) (ratio to M)	462634 (1.325)	787369 (2.066)	657199 (1.769)
V , M using MoW (ratio of M to expected)	320620, 264522 (0.757)	330398, 272589 (0.715)	226554, 186915 (0.503)

*Ratio to the expected value.

Table S2. Cryo-EM data collection, refinement, and validation statistics of the pY-STAT1 dimer and tetramer.

	pY-STAT1 dimer (EMDB-39680) (PDB 8YYV)	pY-STAT1 tetramer (EMDB-39679) (PDB 8YYU)
Data collection and processing		
Magnification	81 K	81 K
Voltage (kV)	300	300
Electron exposure (e-/Å ²)	50	50
Defocus range (μm)	-0.8 to -1.8	-0.8 to -1.8
Pixel size (Å)	0.87	0.87
Symmetry imposed	C2	C2
Initial particle images (no.)	4,005,846	
Final particle images (no.)	115,824	17,869
Map resolution (Å)	3.07	3.84
FSC threshold		
Refinement		
Initial model used (PDB code)	1BF5, 1YVL	1BF5, 1YVL
Model resolution (Å)		
FSC threshold	3.05	3.75
Model composition		
Non-hydrogen atoms	5,249	10103
Protein residues	553	1058
Nucleic acids	36	72
<i>B</i> factors (Å ²)		
Protein	113.917	265.479
Nucleic acids	126.650	185.726
R.m.s. deviations		
Bond lengths (Å)	0.0138	0.0133
Bond angles (°)	2.77	2.64
Validation by MolProbity		
MolProbity score	3.54	3.38
Clashscore	31.43	28.66
Poor rotamers (%)	11.07	8.66
Ramachandran plot		
Favored (%)	77.39	81.24
Allowed (%)	15.62	13.93
Disallowed (%)	6.99	4.84

Movie S1. Three-dimensional variation analysis of the tetramer density.