JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

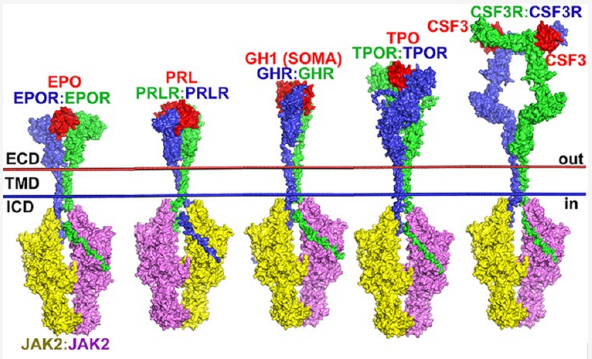
Structural Modeling of Cytokine-Receptor-JAK2 Signaling Complexes Using AlphaFold Multimer

Irina D. Pogozheva, Stanislav Cherepanov, Sang-Jun Park, Malini Raghavan, Wonpil Im, and Andrei L. Lomize*

Cite This: https://doi.org/10.1021/acs.jcim.3c00926 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Homodimeric class 1 cytokine receptors include the erythropoietin (EPOR), thrombopoietin (TPOR), granulocyte colony-stimulating factor 3 (CSF3R), growth hormone (GHR), and prolactin receptors (PRLR). These cell-surface single-pass transmembrane (TM) glycoproteins regulate cell growth, proliferation, and differentiation and induce oncogenesis. An active TM signaling complex consists of a receptor homodimer, one or two ligands bound to the receptor extracellular domains, and two molecules of Janus Kinase 2 (JAK2) constitutively associated with the receptor intracellular domains. Although crystal structures of soluble extracellular domains with ligands have been obtained for all of the receptors except TPOR, little is known about the structure and dynamics of the complete TM complexes that activate the downstream JAK-STAT signaling pathway. Three-dimensional models of five human receptor complexes with cytokines and JAK2 were generated here by using AlphaFold Multimer. Given the large size of the complexes (from 3220 to 4074 residues), the modeling required a stepwise assembly from smaller parts, with selection and validation of the models through comparisons with published experimental data. The modeling of active and inactive complexes supports a general activation mechanism that involves ligand binding to a monomeric receptor followed by receptor dimerization and rotational movement of the receptor TM α-helices, causing proximity, dimerization, and activation of associated JAK2 subunits. The binding mode of two eltrombopag molecules to the TM α-helices of the active TPOR dimer was proposed. The models also help elucidate the molecular basis of oncogenic mutations that may involve a noncanonical activation route. Models equilibrated in explicit lipids of the plasma membrane are publicly available.



1. INTRODUCTION

Cytokines are small secreted glycoproteins that regulate hematopoiesis, neurogenesis, adaptive and innate immunity, lactation, reproduction, growth, and metabolism through binding to cognate receptors.[1] Cytokine receptors are cell-surface glycoproteins with a single transmembrane (TM) α-helix. They lack kinase activity and therefore rely on cytoplasmic tyrosine kinases to mediate intracellular processes. In particular, class 1 and 2 cytokines act via binding to cytokine receptors on the surface of target cells to activate the cytoplasmic nonreceptor protein tyrosine kinase from the Janus Kinases (JAK) family that initiates the downstream JAK-STAT signaling pathway. The cytokine-initiated signaling involves five consecutive steps: (1) binding of a cytokine to a specific receptor and formation of the active receptor dimer; (2) activation of the receptor-associated JAKs by dimerization and *trans*-phosphorylation in their activation loops; (3) phosphorylation of receptor tyrosine residues by the activated JAK; (4) binding of a STAT transcription factor to phosphotyrosines of receptor, leading to phosphorylation, dimer rearrangement, and nuclear translocation of STAT to drive the expression of cytokine-responsive genes; and (5) switching off the activated receptor by tyrosine phosphatases (SHPs), suppressors of cytokine signaling (SOCS), receptor internalization, and downregulation.[2] Human genomes encode more than 50 cytokine receptors, four members of the JAK family (JAK1, JAK2, JAK3, and TYK2), seven STATs (STAT1−4, STAT5a, STAT5b, and STAT6), two SHPs (SHP1 and SHP2), and eight SOCS (SOCS1−7 and CIS).[1−3] Specific members of the JAK, STAT, and SOCS families are linked to individual receptors (Table S1).

Cytokines of the JAK-STAT pathway are α-helical proteins that form either 4-α-helical bundles with up-up-down-down topology (class 1, Figure S1) or structures with 5−6 antiparallel α-helices arranged in an up-down fashion (class

Received: June 19, 2023

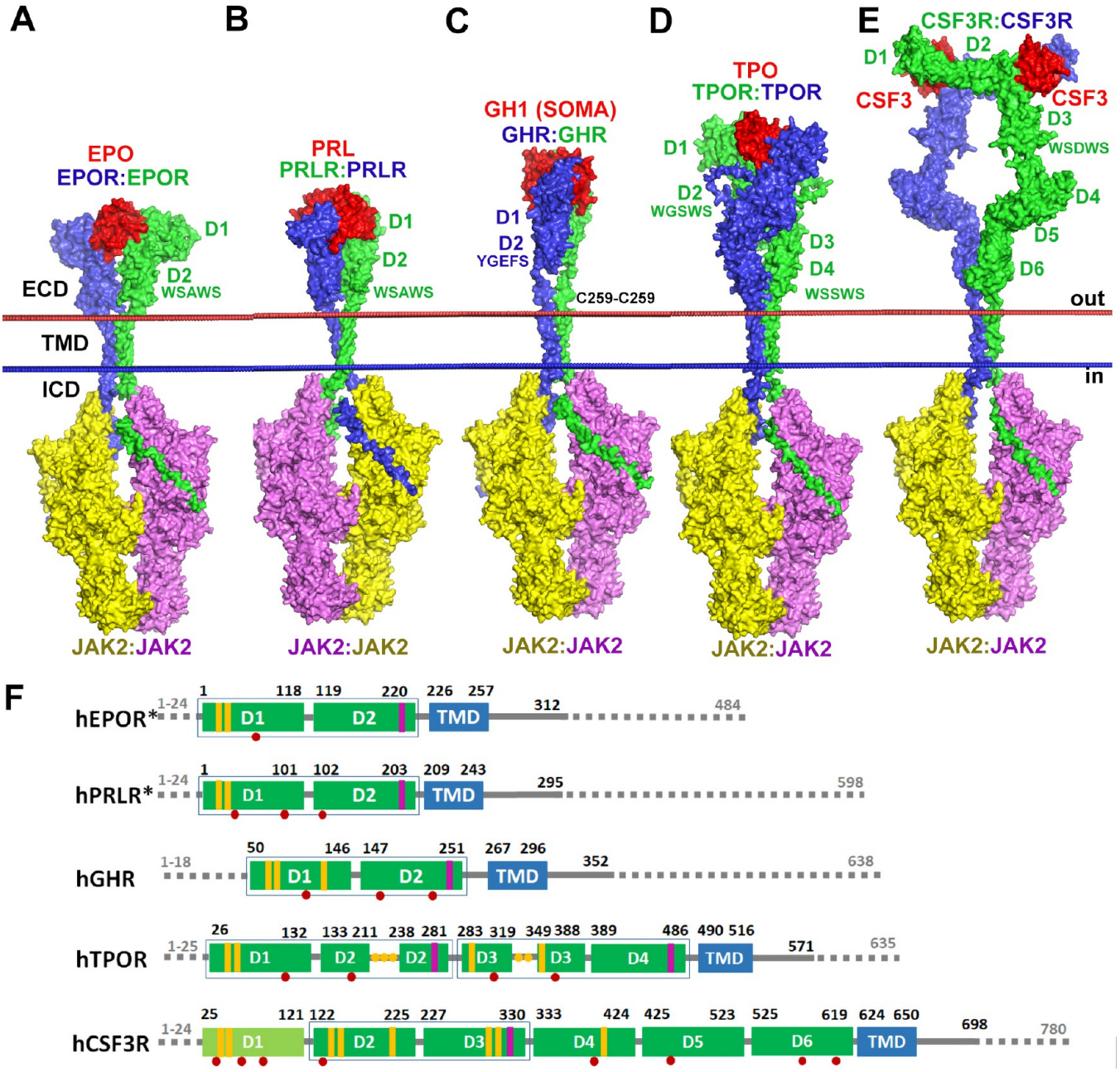


Figure 1. AlphaFold Multimer (AFM)-based models of active signaling complexes of human homodimeric class 1 cytokine receptors: EPOR (A), PRLR (B), GHR (C), TPOR (D), and CSF3R (E). The complexes are composed of receptor homodimers, one (A–D) or two (E) ligands, and two JAK2 molecules bound to the intracellular domains (ICD) of receptors. Molecules are shown by surface representation and colored red for ligand, blue and green for receptor subunits, yellow and pink for JAK2 subunits. The extracellular domains (from D1 to D6) and WSXWS motifs of the receptors are indicated for each complex. The GHR complex has an intermolecular C259-C259 disulfide bond. Hydrophobic membrane boundaries are shown as red (extracellular side) and blue (intracellular side) spheres. (F) Domain architecture of the five cytokine receptors studied. The dark green boxes indicate fibronectin type III (FnIII) domains. The light green boxes indicate immunoglobulin-like (Ig-like) domains. Blue boxes indicate TMDs. The boxes around the two domains indicate cytokine homology modules (CHMs). The gray lines indicate unstructured regions or signal sequences. The yellow bars indicate disulfides. The purple bars indicate the WSXWS motifs. The red circles indicate N-glycosylation sites. The yellow circles indicate cysteine residues in loops of D2 and D3 domains of TPOR, which may form disulfides or metal-bound clusters. The dashed lines indicate disordered regions that have been omitted in the final models but included during some of our AFM calculations. Asterisks indicate receptors with residue numbers corresponding to mature proteins lacking signal sequences.

2). The class 1 cytokine receptors are the largest group of 34 proteins encoded by the human genome.[4] These single-pass TM proteins have different lengths, domain architectures, and quaternary structures. The class 1 receptor family includes five subfamilies: (1) homodimeric receptors that bind one or two ligands per receptor pair; (2, 3, and 4) subfamilies of interleukin (IL) receptor: the IL-12/23, IL-2, and IL-6 receptors forming heterodimers, heterotrimers, or heterotetramers with ligand:receptor stoichiometries of 1:2, 1:3, and 2:4; and (5) a subfamily of IL-3 interleukin receptors forming a 12-meric complex composed of 8 receptors and 4 cytokine molecules.[2]

https://doi.org/10.1021/acs.jcim.3c00926
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

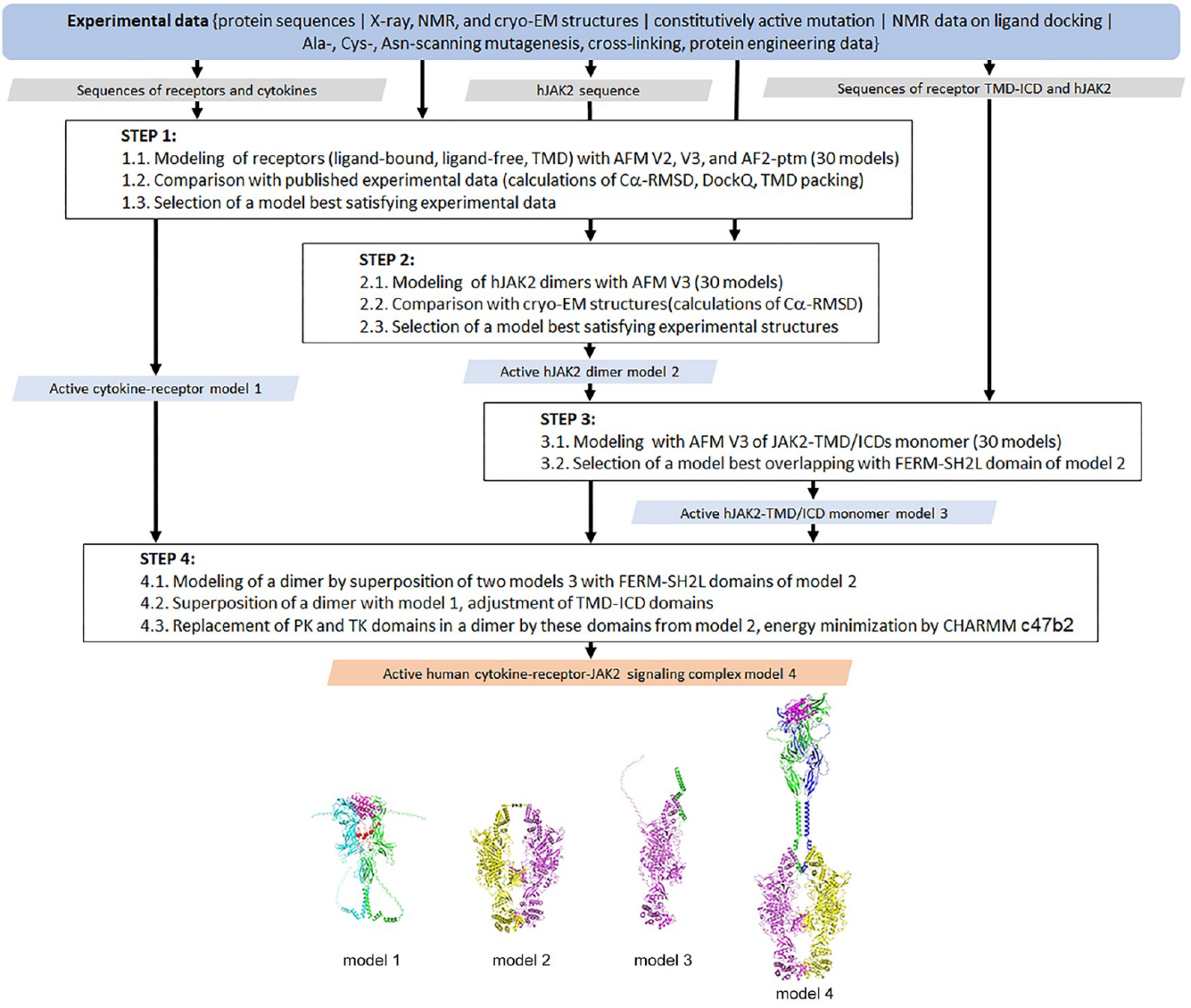


Figure 2. Computational workflow for modeling of cytokine-receptor-JAK2 signaling complexes. Models 1−4 are shown for the human TPO-TPOR-JAK2 (1:2:2) signaling complex.

The class I subfamily of homodimeric receptors includes receptors for erythropoietin (EPOR), growth hormone (GHR), prolactin (PRLR), thrombopoietin (TPOR, also called MPL or CD110), granulocyte colony-stimulating factor 3 (CSF3R), and leptin. These receptors (Figure 1) use two identical chains, each composed of the β-structural extracellular domain (ECD) responsible for ligand binding, a single-helical TM domain (TMD) driving receptor dimerization, and a disordered intracellular domain (ICD) responsible for JAK2 binding and STAT signaling.[2] EPOR, GHR, and PRLR are structurally simple[5−7] with an ECD containing a single cytokine homology module (CHM) formed by two fibronectin type III (FnIII) domains, D1 and D2. The membrane-distal D1 domain carries two conserved disulfides, whereas the membrane-proximal D2 domain features a characteristic WSXWS motif,[8] replaced by the YGEFS motif in GHR (Figure 1C). The TPOR ECD is twice as large as that composed of four FnIII domains: D1 and D2 forming the CHM1, D3 and D4 forming the CHM2.[9,10] The ECD of the long-chain CSF3R contains six domains: the immunoglobulin-like (Ig-like) D1 domain, two FnIII domains, D2 and D3, forming CHM, and three extra FnIII domains, D4, D5, and D6.[11] The long-chain leptin receptor with a more complex domain architecture[12] will not be studied here.

Crystal structures of 1:2 complexes of natural cytokines with soluble receptor ECDs have been solved for human erythropoietin (EPO)-EPOR,[13] human somatotropin (GH1)-GHR,[14] and human prolactin (PRL) with rat PRLR[15] (PDB IDs: 1EER, 3HHR, 3NPZ, respectively). In these crystal structures, two similar receptor chains create an interface for binding the asymmetric surfaces of a cytokine molecule. Site 1 is formed by helices α1 and α4 and the loop connecting α3 and α4, while site 2 is composed of α1 and α3 helices. Crystallographic and biophysical studies demonstrated that a cytokine initially binds to a single receptor via the high-affinity site 1[16,17] and then to the second receptor through the low-affinity site 2.[18] The largest CSF3R-CSF3 complex has a different 2:2 receptor-ligand stoichiometry and represents the association of two 1:1 units.[11] The CSF3 binding site is formed by CHM (D2 and D3) of one chain and the Ig-like domain

https://doi.org/10.1021/acs.jcim.3c00926
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

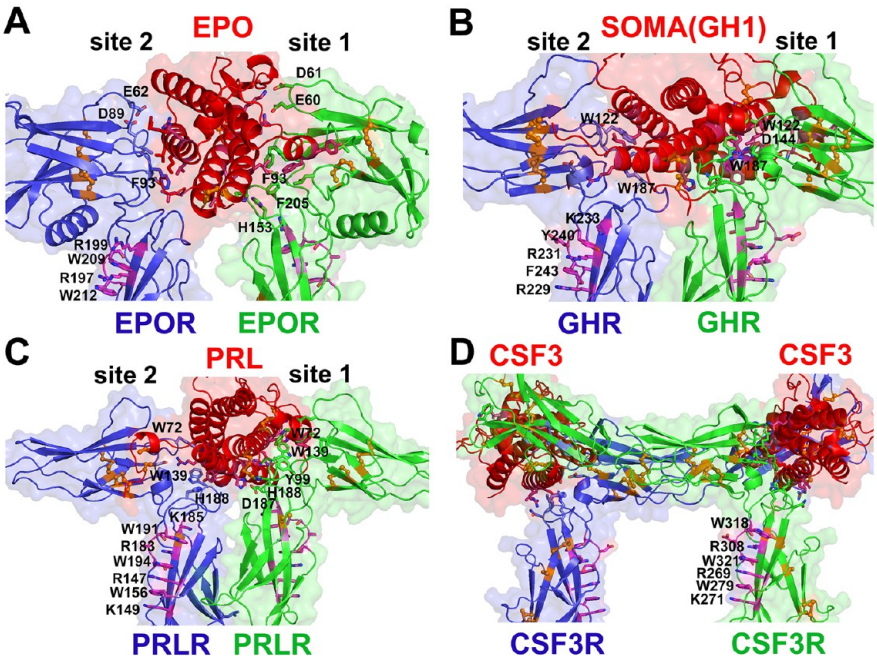


Figure 3. Ligand binding pockets in models of active homodimers of class 1 cytokine receptors: EPOR (A), GHR (B), PRLR (C), and CSF3R (D) with bound cytokine ligands. Protein molecules are shown by semitransparent surfaces and cartoon representations are colored red for ligands, and blue and green for receptor subunits. Interacting receptor and cytokine residues are shown as sticks. Cysteine residues are shown as balls-and-sticks colored orange. Residues involved in the WSXWS signature motif are shown as purple sticks. A set of interdigitated arginine and tryptophan residues from this motif together with neighboring tryptophan, arginine, and lysine residues participate in the network of cation-π interactions.

(D1) of the other chain. At present, there are no experimental structures of the TPOR or its domains. Hence, experimentally based computational models have been proposed for the TPOR TM-ICD in complex with the JAK2 dimer[19] or for the full-length human TPOR in complexes with thrombopoietin (TPO) or oncogenic calreticulin (CRT) mutants that bind to TPOR to cause its aberrant activation in myeloproliferative neoplasms (MPNs).[20,21]

In the absence of experimental atomic-level structures of cytokine receptor signaling complexes, computational modeling provides a valuable alternative. A transformative breakthrough in the protein structure prediction has recently been achieved by developing a new deep learning AlphaFold method that uses coevolutionary and structural information.[22] AlphaFold version 2.0 (AF2) produces models of nearly experimental quality for single-chain proteins and outperforms other methods in predicting contact interfaces of multidomain proteins[23,24] and protein complexes,[25−27] including TM homo- and heterodimers.[28,29] The high speed and quality of predictions by AF2 justified its applications on a proteomic scale.[30] More than 200 million protein models of single-chain proteins from 48 organisms were generated using this method and deposited into the AlphaFold DataBase.[31] The recently released AlphaFold Multimer (AFM) was recognized as the best computational tool for modeling protein complexes.[32]

In this study we used the publicly available AFM ColabFold version[33] to model active signaling complexes of human class 1 homodimeric cytokine receptors, including EPOR, PRLR, GHR, TPOR, and CSF3R. Each signaling complex is composed of one or two (for CSF3R) cytokines bound to a receptor homodimer interacting with the JAK2 dimer. The accuracy of the models was verified through comparison to published experimental data. A comparison of ligand-free and cytokine-bound models revealed molecular mechanisms of receptor activation leading to dimerization and activation of receptor-bound JAK2. These models reveal atomic details of protein-protein interactions, demonstrating conformational changes and structural flexibilities in the ECDs, TMDs, and ICDs of receptors and JAK2 domains in signaling complexes. This understanding aids in deciphering the nature of the oncogenic mutations.

2. RESULTS

2.1. Four-Step Modeling on Cytokine Receptor Signaling Complexes. The direct modeling of complexes composed of 5 or 6 proteins using AFM was not feasible due to their very large size and multidomain architecture. Therefore, for each cytokine receptor, we separately calculated several smaller overlapping parts and assembled them into a complete ligand-receptor-kinase complex. The modeling was performed in four steps (see workflow in Figures 2 and S2).

At the first step, we generated complexes of cytokines with their receptor dimers that included ECDs, TMDs, and parts of ICDs. We also modeled dimers of ligand-free receptors and compared them with the corresponding ligand-bound dimers. In addition, we calculated dimers of peptides representing the TM and juxtamembrane regions (TM-JM). At the second step, we modeled the active dimer of human JAK2. At the third step, we produced complexes of a monomeric JAK2 with a TMD-ICD fragment of each receptor.

At the first three steps, we generated up to 30 various models for each protein complex using different AF2 methods (AFM V2 or V3 and AF2-ptm) with different random seed numbers and numbers of recycles. For each set, we selected a single model that was the most compatible with the available experimental data. We calculated the Cα-atom root-mean-square deviations (Cα-RMSDs) and DocQ scores using available X-ray, NMR, and cryo-EM structures and chose

https://doi.org/10.1021/acs.jcim.3c00926
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

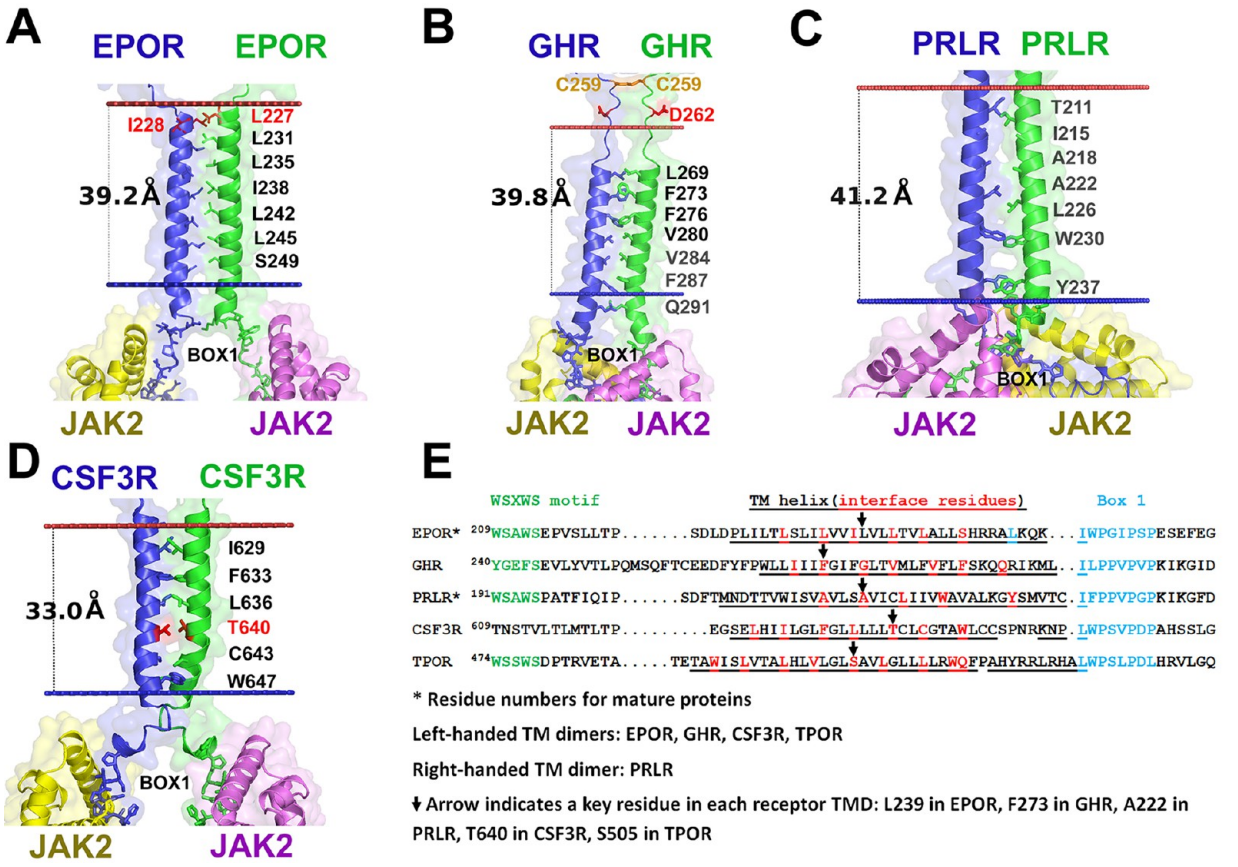


Figure 4. TM α-helical dimers with predicted locations of membrane boundaries in AFM models of signaling complexes of EPOR (A), GHR (B), PRLR (C), and CSF3R (D). Each complex is composed of two receptor molecules (colored blue and green), bound cytokine(s) (not shown), and subunits of a JAK2 homodimer (colored yellow and pink). The TM α-helices form left-handed dimers with positive crossing angles (via extended leucine zipper heptad repeat pattern) for EPOR, GHR and CSF3R, and a right-handed dimer for PRLR (via AxxxA^{222}xxxL motif). Residues from the TMD dimerization interface and Box1 residues are shown as sticks. Cysteine residues are shown as balls-and-sticks colored orange, and C259 residues of the intermolecular disulfide in the active complex of GHR are highlighted. Residues involved in mutations leading to constitutive activation[43,44] are colored red. Protein molecules are shown as semitransparent surfaces and cartoon representations. Membrane boundaries were calculated by the PPM 3.0 method.[34] (E) Sequence alignments of receptor TMDs including juxtamembrane regions. Asterisks indicate receptors with a residue number of mature proteins lacking signal peptides. TM α-helical residues are underlined, and those at the dimerization interface are colored red. The WSXWS and related motifs are colored green. Residues in the "switch" and Box1 motifs are colored blue. Arrows indicate key residues in the TMDs of cytokine receptors.

TMD conformations that agree with experimentally identified TMD dimerization modes. We also validated models through published mutagenesis and protein engineering data, disulfide cross-linking, constitutively active mutants, and other data. At each step, we selected one model that best satisfied experimental data (models 1 and 2) or provided the best superposition with the JAK2 model (model 3).

At the fourth step, the full-length active signaling complex of each receptor was assembled from the best models of ligand-bound receptor dimer (model 1) and of the TMD/ICD-bound kinase monomer (model 3) selected in the previous steps. The final model (model 4) was refined by energy minimization and positioned in the membrane by the PPM 3.0 method.[34] The models of five receptor-cytokine signaling complexes were determined with rather high reliability scores for most structural domains but lower reliability for loops and TM helices (Figure S3).

2.2. Step 1: Modeling and Validation of the Receptor Homodimers with and without Ligands. *2.2.1. Complexes of EPOR, GHR, and PRLR Homodimers with Cytokines.* Models of cytokine-receptor complexes were generated by AFM for five human receptors and the extensively studied murine EPOR and validated using available experimental data. The ECDs in the models of five cytokine-bound receptor dimers, EPO-EPOR$_2$, GH1-GHR$_2$, PRL-PRLR$_2$, CSH1-PRLR$_2$, and GH1-PRLR$_2$, superimpose well with the corresponding crystal structures (PDB IDs: 1EER, 3HHR, 3NPZ, 1F6F, and 1BP3, respectively) with the Cα-RMSD less than 2 Å (Table S2). The main residues involved in ligand-receptor interactions (Figure 3) are the same as in the corresponding experimental structures.[13−16,35] For example, hydrophobic and aromatic residues, such as F93, F205, and M150 in hEPOR, W122 and W187 (W104 and W169 in mature protein) in hGHR, and W72 and W139 in hPRLR, contribute significantly to the hormone-receptor interactions. The ligand-receptor complexes were found to be of high quality for site 1 of ligand (DockQ $\geq$ 0.81),[36] and medium quality for site 2 of ligand (0.51 $\leq$ DockQ < 0.81) in all complexes (Table S3).

Models of ligand-bound EPOR, GHR, PRLR, and TPOR show a significant asymmetry of ECDs of two receptor chains. Binding of ligands through two dissimilar surfaces (sites 1 and 2) induces movement of D1-D2 domains of both chains relative to each other in vertical (along the membrane normal, z-axis) and horizontal (in the xy plane) directions. The vertical

https://doi.org/10.1021/acs.jcim.3c00926
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

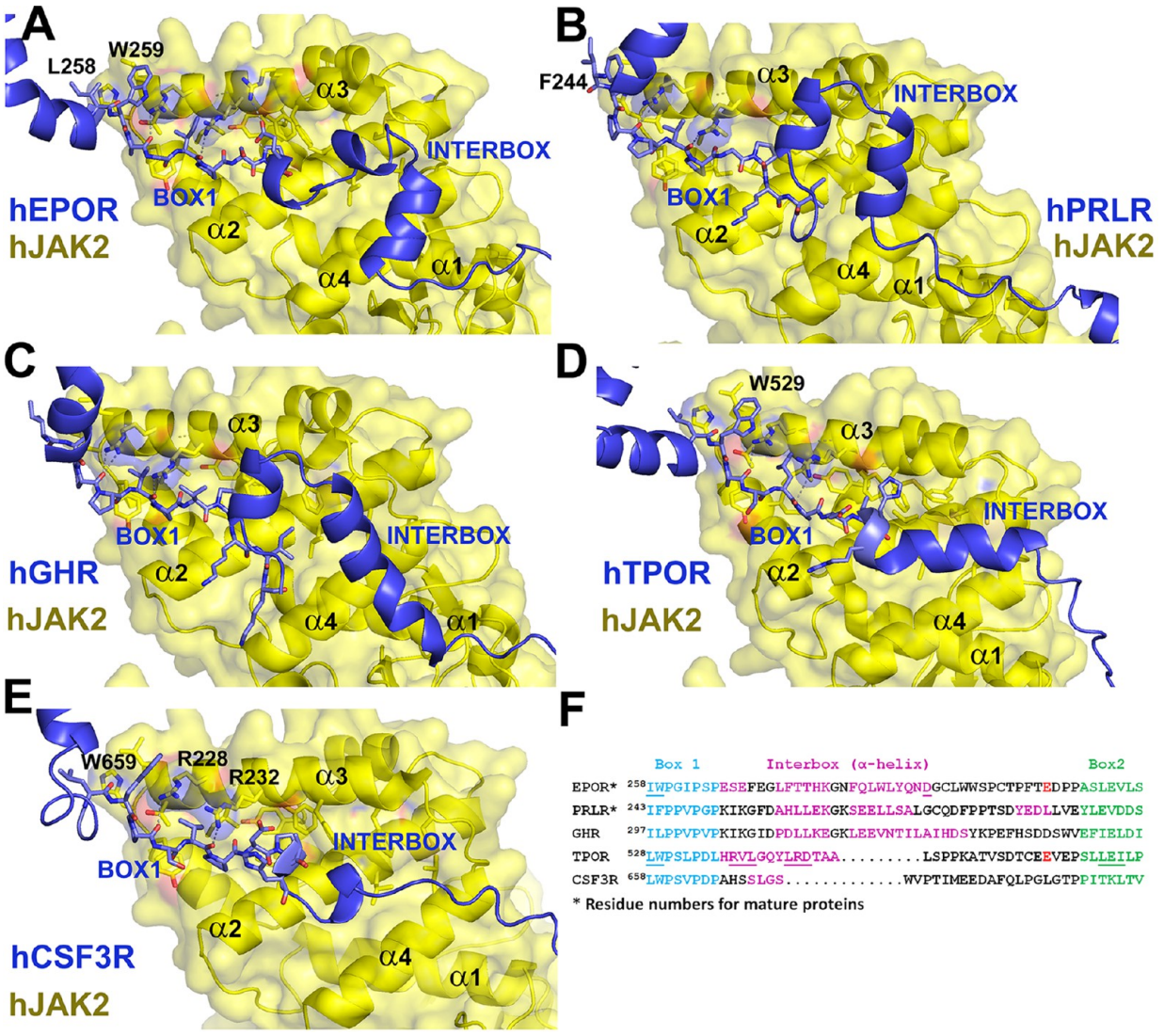


Figure 5. Recognition of the Box1 receptor motifs by the FERM domain of JAK2. Fragments of AFM-models generated at step 3 for cytokine receptor TMD-ICD in complex with JAK2. Specific interactions among ICDs of human EPOR (A), PRLR (B), GHR (C), TPOR (D), and CSF3R (E) and the α1-α4 subdomains of FERM are shown. Pro-rich fragments of the Box1 motifs interact with α3 of FERM; α-helical fragments of the "interbox" regions interact with α2 and α4 of FERM; and Box2 motifs interact with the SH2L domain (see Figure S8). R228 and R232 from the FERM α3 form H-bonds with the main chain carbonyls of the Box1 fragment. (F) Sequence alignments of receptor ICD fragments interacting with the JAK2 FERM-SH2L, based on the AFM models. Residues in the "switch" and Box1 motifs are colored blue. Box2 motifs are colored green, and the interbox regions forming α-helices are colored purple. Underlined residues are known to be important for JAK2 activation.[38,45]

shift (by 5 to 8 Å) of the site 1 ECD is translated to the upward piston movement of the corresponding TM α-helix. Thus, TM helices of the dimer are positioned in the membrane at different heights (Figure 4). The vertical shift is pronounced in models of the short-chain receptors but is not seen in models of multidomain long-chain receptors. The asymmetry is absent in complexes of hEPOR dimers with two similar molecules of peptide mimetics (PDB IDs: 1EBA, 1EBP) and in the ligand-free receptor dimers (see below).

For one of these receptors, hPRLR, we generated models of complexes with three different human hormones known to interact with this receptor *in vivo*: prolactin (PRL), somatotropin (GH1), and placental lactogen (CSH1). These models demonstrate many similarities but also some differences. For example, in all three models, zinc-binding centers are in ligand binding site 1. Zn^{2+} ions may link the α1 and α4 helices of hormones (residues H27 and D183 of hPRL, residues H18 and E174 of hGH1 and hCSH1) with hPRLR (residues D187 and H188), in agreement with experimental studies.[6] The ECD-ligand complexes of hPRLR with hGH1 and hCSH1 are rather similar (with Cα-RMSD of 0.9 Å) but differ from the complex with hPRL (Cα-RMSD of 2.3 and 2 Å, respectively). These differences are likely caused by the width differences of hGH1 and hCSH1compared to hPRL. The separation of the two ECDs to accommodate larger ligands increases the distances between the N-termini of TM helices and slightly shifts the helix crossing point toward the C-terminal end. There is no change in the distances between the helix ends that interact with JAK2.

An important part of the structure is the pair of interacting TM α-helices. These α-helices are rather long (27−36 residues or 40 to 54 Å) (Figure 4A−C), consistent with NMR studies.[37] The polar C-terminal parts of TM α-helices extend from the membrane into the cytosol, where some hydrophobic

https://doi.org/10.1021/acs.jcim.3c00926
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

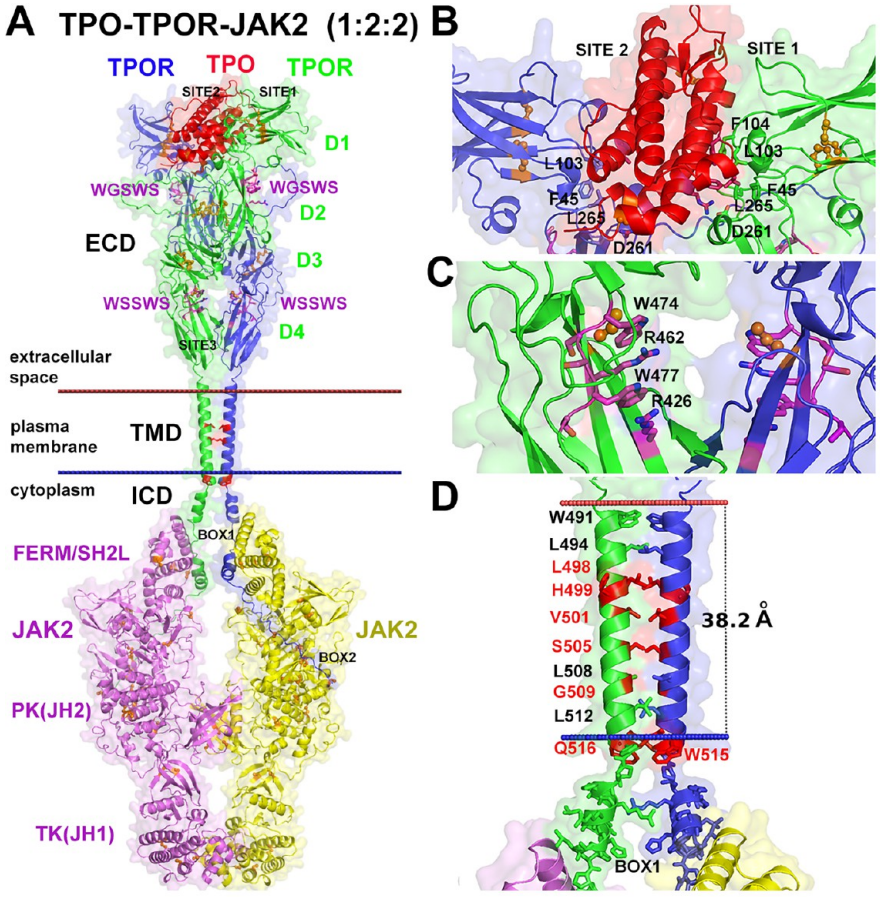


Figure 6. AFM-generated model for the full-length human TPO-TPOR-JAK2 (1:2:2) signaling complex. (A) Overview of the complex in a membrane. Protein molecules are shown as semitransparent surfaces and cartoon representations are colored red for TPO, blue and green for TPOR subunits, and yellow and pink for the JAK2 subunits. Cysteine residues are shown as balls-and-sticks colored orange. Domains of TPOR and JAK2 are indicated. The two WSXWS motifs in D2 and D4 of ECDs, and the Box1 and Box2 regions in ICDs, are highlighted. Disease-associated recurrently mutated residues in the TMD (V501, S505, and W515) which cause constitutive TPOR activation[58,59] are colored red. (B) The TPO binding pocket in the ECDs of TPOR model. TPOR residues interacting with TPO are shown as sticks. F45, L103, F104, D261, and L265 have been previously implicated in TPO binding.[51,52] (C) Close-up of the WSSWS motif in the D4 domain of ECD. Aromatic and basic residues involved in the network of cation-π interactions are shown as purple sticks. (D) Close-up of the TMD and Box1 region of TPOR interacting with JAK2. TM α-helices of TPOR have left-handed arrangements (with a positive crossing angle); residues at the interface are depicted by sticks, residues with natural or engineered mutations (S505N, L498W/H499C,Y, L498W/W515 K, H499G/V501S, H499C,Y/S505N, H499L/G509N, H499L,C,Y/W515 K, V501A/W515L,R, S505N/T487A, S505N/S493C, S505N/V501A,M, S505C/W515L, S505N/Q516, and S505N/V501N/A506 V) associated with constitutive activation of TPOR[10,53,54,56,57,59,60] are colored red. TPOR residues forming the Box1 motif are shown as sticks. Membrane boundaries were calculated by the PPM 3.0 method.[34]

residues from the helix ends along with Box1 residues interact with JAK2 (Figure 5). For example, the $L^{253}xxxI^{257}W^{258}$ "switch" motif in mouse EPOR forms a rigid connection between TMD and ICD, which is critical for the JAK2 activation upon EPO stimulation.[38]

hPRLR is the only homodimeric receptor with a right-handed arrangement of TM α-helices in the model of the active receptor dimer (Table S2). Though AFM calculations generated both right-handed and left-handed TM α-helix arrangements for hPRLR complexes with all three ligands with slightly different helix-helix interfaces, only the right-handed version appeared in the models of constitutively active hPRLR mutants with deleted ECDs (Δ10−186 and Δ1−210).[39,40] The right-handed TMD dimer also had longer TM α-helices in comparison to various left-handed dimers. Therefore, the right-handed dimer of hPRLR was selected as a preferred TMD conformation (Figure 4C). The selected model is similar to one of conformations of the PRLR dimer obtained in multiscale simulations.[41] The right-handed helix dimer is characterized by a negative crossing angle and a tetrad repeat motif. In the AFM model of hPRLR, TM α-helixes cross at the middle of the membrane at A222, while W214 and W230 of both helices are located near the membrane boundaries. The large distances between N−N and C−C termini of interacting TM helices and the presence of adjacent flexible loops may explain lack of effects of Ala- or Gly-insertions at the junctions of hPRLR TMD with ECD or ICD.[42]

In contrast to the right-handed TM dimer of hPRLR, models of hGHR, hEPOR, and mEPOR dimers with bound cytokines demonstrate a left-handed TM helix arrangement, as defined by a positive crossing angle and the $(abcdefg)_n$ heptad repeat motif (where the *a* and *d* positions form the interface).

For example, the model of the active hGHR dimer in complex with hGH1 shows a left-handed TM α-helix arrangement with F273 at the *d*-position of the heptad repeat motif (Figure 4B, Table S2). This helix orientation and the presence of an intermolecular disulfide C259−C259 are consistent with the NMR structure of the active dimer of

https://doi.org/10.1021/acs.jcim.3c00926
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

TM segments of hGHR.[46] The predicted active conformation of the TM dimer is consistent with Cys-scanning mutagenesis and cross-linking studies that localize residues L269, F273, F276, and V280 at the dimerization interface.[47] This TMD packing also agrees with activation of hGHR by the fused coiled-coil dimerization domain of the c-Jun transcription factor that clamps together the TM helices.[47,48]

TM α-helices of the active EPOR dimer also form a leucine zipper with the reference residue, S238 in mEPOR (L239 in hEPOR), occupying the *e*-position of the *heptad repeat motif* (Figure 4A, Table S2). The modeled TM helix arrangement is in good agreement with results of the fusion of the mEPOR TMD with the coiled-coil dimerization domain of the yeast transcription factor *Put3*, where the left-handed dimer *cc-EPOR-III* with S238 in the *e*-position was constitutively active.[49] This helix orientation also explains the constitutive activity of L241N mutation in mEPOR (L242N in hEPOR).[50] The hEPOR L242N mutated residue is located at the dimerization interface (*d*-position) and may stabilize the TM dimer by the formation of intermolecular hydrogen bonds. Furthermore, in the active ligand-bound hEPOR model, L227 and I228 (L226 and I227 in mEPOR) are located at the N-termini of both TM helices close to each other (Figure 4A) and able to form an intermolecular disulfide if substituted by cysteines. Such helix arrangement is consistent with Cys-scanning mutagenesis of mEPOR that found constitutive activity of L226C and I227C mutants forming disulfide-linking dimers.[44]

2.2.2. TPOR Ligand-Receptor Complex. In the structural model of the human TPO-TPOR (1:2) active complex, TPO binds to the D1 (A-B and E-F loops) and D2 (F-G loop) domains via multiple hydrophobic and ionic interactions. Five hTPOR residues, F45, L103, F104, D261, and L265, form multiple contacts with hTPO residues from their α1, α3, and α4 helices and the loop between α1-α2 (Table S5, Figure 6B). These receptor residues were identified in mutagenesis studies as key TPO-binding determinants.[51,52] Two receptor ECDs interact not only with the ligand but also with each other via a long loop within D2 (residues 187−238) and two antiparallel β-strands from D4 (residues 436−438) that form a site 3 between D4 domains (Figure 6A). Nonconserved cysteine residues from the D2 loop form three intramolecular disulfides (C193−C323, C194−C241, and C211−C322) in many AFM V3 models (Table S1). We hypothesize that these disulfides may stabilize a monomeric ECD structure exposed to the extracellular environment, while the D2 loop constrained by disulfide bonds may participate in the ligand binding and in dimer stabilization.

In the model of the active hTPOR dimer, the TM α-helix spans over 28 residues (from T289 to F517 in the RWQF motif) (Figure 6C), similar to long TMDs of human and mouse EPORs.[37] The existence of rather long TMDs encompassing W491 and W515 agrees with NMR studies of TMD dimers.[53,54] However, a helix break at H499, which has been suggested based on NMR data of hTPOR monomers,[55] is not observed in hTPOR, either monomeric or dimeric models generated by AFM. AFM-based models demonstrate the TM α-helix kink at P518. After this helix kink, an additional 3-turn polar helix (A519-L528) extends to the cytosol to interact with JAK2, similarly to the "switch" residues of mEPOR (Figure 5A, Figure 4D). The left-handed α-helix arrangement in the model of the active TM dimer with S505 at the *a*-position and H499 at the *b*-position of the heptad repeat motif (Figure 6D) is consistent with the dimerization interface of the constitutively active *cc-TPOR-I* fusion construct between a dimeric coiled-coil of Put3 and the TMD of mTPOR.[56] This dimerization interface is also supported by Asn-scanning mutagenesis of human and murine TPOR[53] and studies of constitutively active hTPOR mutants.[54,57]

Furthermore, docking of an allosteric ligand eltrombopag to the TM α-helices (Figure S4) also supports the proposed AFM model of the active hTPOR dimer. Two eltrombopag molecules are located at both sides of the TM α-helical dimer and participate in hydrophobic interactions with two sets of W491, I492, I494, V495, T496, L498, and H499 residues near the helix N-termini and ionic interactions between drug carboxyl groups and two R456 residues from the D4 domains. Eltombopag can also form Zn^{2+}-mediated interactions with both H499 residues, like the structurally similar compound SB394725.[55] These positions of eltrombopag are consistent with the previously identified locations of its structural analogues[55,57] and, especially, with a key role for W491 in TPOR activation by eltrombopag.[54]

2.2.3. CSF3R Ligand-Receptor Complex. Unfortunately, AFM was unable to automatically produce models of ligand-receptor complexes of the long-chain hCSF3R with TMDs forming a dimer, even though the D1-D3 domains with bound ligands were superimposable with the corresponding crystal structure (PDB ID: 2D9Q)[11] with Cα-RMSD of around 3 Å (Table S2). Therefore, modeling of receptor-ligand complexes for hCS3FR was done stepwise separately for the ECDs and TMDs. AFM modeling started from a complex of monomeric hCSF3R with bound hCSF3 at a 1:1 ratio. Then, two such models were superimposed with both units of the homodimeric crossover crystal structure of the hCSF3-hCSF3R 2:2 complex composed of two 1:1 units.[11] Superposition demonstrated a good overlap of experimental and calculated 1:1 complexes of D1-D3 domains with hCSF3 (Cα-RMSD of 0.9 Å; Table S2), but different spatial positions of the remaining receptor domains. Small adjustment of the main chain angles in the D3-D4 linker allowed a juxtaposition of TM α-helices to form a dimer.

To define helix orientations in the active ligand-bound state of hCSF3R, we modeled dimers of isolated TM segments with sequences corresponding to native and the constitutively active oncogenic mutant, T640N[61,62] (Table S4). The AFM models of isolated TM segments with native and mutant (T640N) sequences have left-handed TM helix arrangements with T640 (or N640) at the dimerization interface at the *a*-position of the heptad repeat motif (Figure 4D). A similar helix arrangement was predicted by the TMDOCK method.[63] To complete the structure of the full-length active hCSF3R dimer, we combined the model of TM dimers and the model of two multidomain ECDs with two bound CSF3 ligands. In the final model of the receptor-ligand complex, two gain-of-function mutations, T640N and G644E,[61] are located at the TM dimerization interface and can stabilize the TMD dimer via hydrogen bonds. The other oncogenic mutations, T612I, T615A, and T118I,[61] are located at the D6-D6 dimerization interface (the site 3) and may contribute to stabilization of ECD dimers by forming more hydrophobic contacts.

2.2.4. Homodimers of Ligand-Free Receptors and TM Segments. AFM models of ligand-free receptor dimers were generated for human and murine EPOR and human GHR, PRLR, and TPOR. They significantly differ from the corresponding experimental and modeled structures of

https://doi.org/10.1021/acs.jcim.3c00926
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

ligand-bound dimers, with Cα-RMSDs ranging from 5 to 9 Å due to rearrangements of ECDs and TMDS (Tables S2 and S4). Unlike the asymmetric active dimers, the ligand-free models are symmetric, have a larger contact area between ECDs of the two receptors chains, a small-sized ligand binding pocket, and altered mutual orientations of TM α-helixes. However, the reliability scores are lower for ligand-free dimer models (ipTM ranging from 0.2 to 03) compared to the ligand-bound dimer models (ipTM ranging from 0.6 to 0.7).

Models of ligand-free homodimers of hPRLR, hGHR, and mEPOR demonstrate a tighter packing of ECDs than in the ligand-bound dimers with occluded ligand binding pockets (Figure 7). The rearrangement of ECDs in the ligand-free dimers brings two symmetric D1 domains closer together compared with the active dimer models and alters the D2-D2 dimerization interface (site 3); thus, some residues from the ligand binding pocket form receptor-receptor interactions occluding the ligand-binding pocket. For example, the close packing of ECDs in the ligand-free hPRLR brings together D187 and H188 residues, which form a predicted Zn^{2+} binding site with cytokine ligands, hPRL, hGH1, and hCSH1, in the active structure. A new Zn^{2+}-binding site might be formed between two D2 domains of the ligand-free hPRLR (Figure 7C). Molecular dynamic simulations of ECDs of hGHR also pointed to the increased contact of subunits of the ligand-free dimer.[64] Extensive contacts between ECDs of two antiparallel receptor chains are also observed in the crystallographic antiparallel dimer of the ligand-free hEPOR (PDB ID: 1ERN).[65]

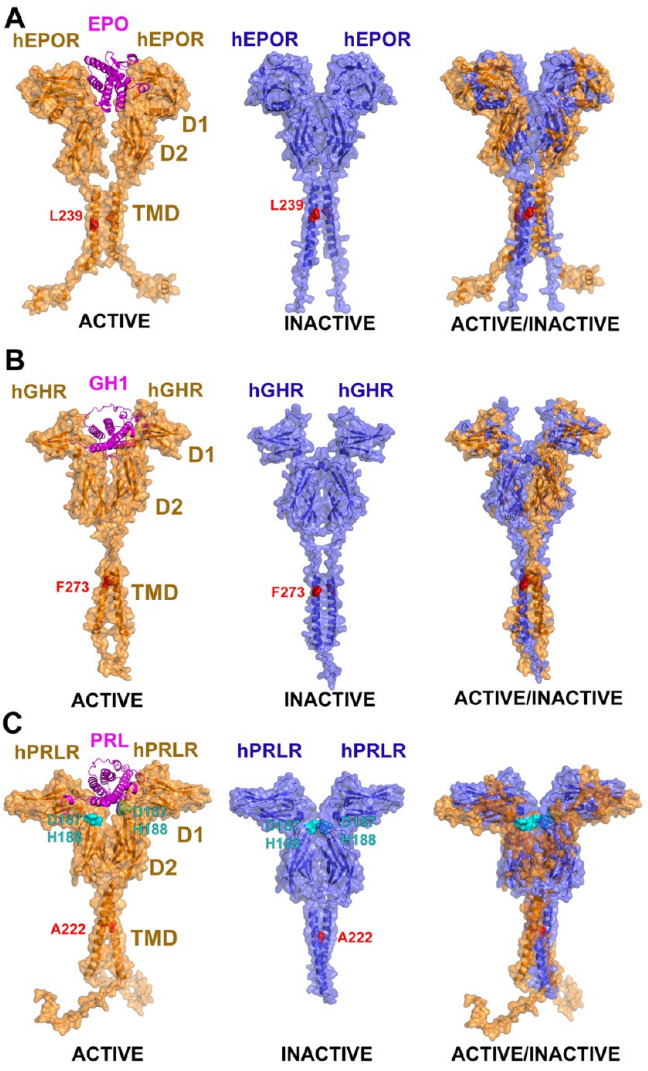


Figure 7. Comparison of AFM-generated models of active and ligand-free dimers of hEPOR (A), hGHR (B), and hPRLR (C). In the ligand-free dimer, D1 domains occlude the ligand binding pockets. Furthermore, the relative positions of the D2 domains are changed, and TM α-helix arrangements are different from those in the active dimers. The molecules are shown by cartoon and semitransparent surface representations colored orange for active dimers and blue for inactive receptor dimers; ligands in the active dimers are colored purple. Reference residues in the TMDs are shown as red spheres. Residue forming a possible Zn^{2+}-binding center in hPRLR (D187 and H188) are shown as cyan spheres.

We also investigated the possibility of formation of disulfide-linked dimers by R130C, D133C, and E134C mutants of hEPOR, because it has been shown that the corresponding mutations (R129C, E132C, and E133C) of mEPOR are constitutively active in the absence of ligands.[66,67] In the active ligand-bound hEPOR dimers generated by AFM, as well as in the corresponding crystal structure (PDB ID: 1EER), D133 and E134 from AB loops of D2 domains are close to each other, while R130 from D2 β-strands are rather distant (Cα−Cα distance of 27−30 Å) (Figure S5A). Therefore, these structures are incompatible with the formation of the C130−C130 intermolecular disulfide. However, disulfide bonds between both chains are formed in the ligand-free hEPOR dimeric models of R130C, D133C, and E134C mutants generated by AlphaFold_2.0_ptm. These disulfide-linked dimers (see C130-C130-linked hEPOR dimer in Figure S5B) have the decreased ligand binding pockets and tightly packed parallel TM α-helices, which may bring together associated JAK2 molecules, consistent with their constitutive activity.

Modeling of the ligand-free hTPOR dimer produced two sets of conformations with an open and closed ligand-binding pocket and a different arrangement of the D1-D2 domains in symmetric chains. The closed conformation is too narrow to fit TPO, while the open conformations have a wider space between D1 and D2 that can accommodate TPO after slight domain movements to match asymmetric sides of the ligand.

Importantly, for many of the receptors, the two TM α-helices have different mutual orientations in the inactive (ligand-free) and active (ligand-bound) models generated by AFM (Tables S2, S4, and S6). For example, in hEPOR and hTPOR models, the TM α-helices are loosely packed and have a right-handed helix arrangement in ligand-free dimers but form tightly packed left-handed dimers in the ligand-bound states. The model of the ligand-free hPRLR dimer has a left-handed helix arrangement, compared to the right-handed dimer in the ligand-bound state (Figure 7C). Less significant difference in TM helix orientations is observed between inactive and active states of EPOR and GHR dimers, where the dimerization interface rotates only by ∼50° and ∼100°, respectively (Table S5, Figures 7A,B and S10).

In addition to the calculations of full-length ligand-bound and ligand-free complexes, we modeled dimers formed by TM-JM peptides (Tables S4 and S6). Interestingly, the arrangements of TM α-helices in TM-JM peptide dimers are more similar to TM α-helix packing in active (ligand-bound or constitutively active) than in inactive dimers of full-length receptors (Table S6). For example, AFM predicted the same right-handed TM helix arrangements for the full-length active

https://doi.org/10.1021/acs.jcim.3c00926
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Table 1. Characteristics of the Final Models of Five Homodimeric Class 1 Cytokine Receptor Signaling Complexes

Complex name (stoichiometry)	Residues of ligand, receptor, and JAK2	TM segment (length)	D, Å[b]	TM helix packing	Key residue[c]	Superposition[d] PDB	Superposition[d] Cα-RMSD, Å
hEPO*-hEPOR*-hJAK2 (1:2:2)[a]	1−166*, 1−312*, 36−1132	P226-I258 (33)	39.2 ± 2.2	L+	L239 *e*	1EER	2.1 (592/592)
						8EWY	3.0 (2052/2172)
						TMDOCK	2.8 (58/60)
hGH1*-hGHR-hJAK2 (1:2:2)	1−191*, 50−343, 36−1132	W267-I297 (31)	39.8 ± 1.8	L+	F273 *d*	3HHR	1.9 (571/573)
						8EWY	2.6 (2052/2172)
						5OEK	3.1 (48/48)
						TMDOCK	3.0 (48/48)
hPRL*-hPRLR*-hJAK2 (1:2:2)	1−199*, 1−295*, 36−1132	M208-I243 (36)	41.2 ± 3.3	R−	A...A^{222}...L	3NPZ	1.8 (536/587)
						8EWY	2.6 (2052/2172)
						TMDOCK	3.7 (49/68)[†]
hTPO*-hTPOR- hJAK2 (1:2:2)	1−153*, 26−571, 36−1132	T489-F517 (29) A519-L258 (10)	38.2 ± 4.7	L+	S505 *a* H499 *b*	8EWY	3.0 (2049/2172)
						TMDOCK	2.3 (55/58)
hCSF3-hCSF3R-hJAK2 (2:2:2)	30−207, 25−676, 36−1132	S624-C650 (27) K655-L658 (4)	33.0 ± 1.6	L+	T640 *a*	2D9Q	1.3 (902/933)
						8EWY	2.6 (2052/2172)
						TMDOCK	0.2 (50/53)

[a]Asterisks denote residue numbers for mature proteins (without signal peptide). [b]Intrinsic hydrophobic thickness (D) calculated by the PPM 3.0 method.[34] [c]Letters for left-handed dimers indicate the positons of a reference residue in the $(abcdefg)_n$ heptad repeat motif, where *a*- and *d*-positions are at the dimer interface. [d]Superpositions of Cα-atoms of final computational models with crystal structures of ligand:receptor ECD (1:2) complexes (PDB ID: 1EER, 3HHR, 3NPZ, and 2D9Q) and with the cryo-EM structure of the mJAK1 dimer (PDB ID: 8EWY) were performed by PDBeFold (3NPZ and 2D9Q) and US-Align (others). Superposition with NMR model of the GHR TM active dimer (PDB ID: 5OEK) and TMDOCK models[63] were done by the align method of PyMOL. Cα-RMSD column includes the number of overlapped residues in the structural superposition divided by total number of residues in the structure (in parentheses).

(ligand-bound) hPRLR dimer and its constitutively active mutants, Δ1−186[39] and Δ1−210,[40] which lack large parts of their extracellular domains.

However, the calculated arrangements of TM α-helices are often close to but not exactly the same in dimers of TM-JM peptides and full-length ligand-bound receptors. For example, the model of full-length hTPOR has a left-handed TM α-helix arrangement with S505 at the *a*-position of the heptad repeat motif (Tables 1 and S2), which has been experimentally proven for the full-length human and mouse TPOR[53] and the left-handed dimer of *Put3*-fused *cc-TPOR-I I* construct.[56] However, S505 occupies an alternative *d*-position the left-handed dimer calculated by AFM for TM-JM peptides of constitutively active hTPOR mutants, L498W/H499Y and H499L/W515 K,[54] consistent with the helix orientation found in isolated TM helix dimer of the constitutively active S505N mutant.[57] AFM calculations also reproduced two dissimilar TM dimerization interfaces that were identified in *Put3*-fused constructs of mEPOR[49] and its TMD segments[50] with S238 located at the *e*- or *a*-positions of the heptad repeat motif, respectively. Furthermore, AFM predictions of helix orientations in isolated TMDs of hTPOR, mEPOR, and CSF3R agree with low-energy models generated by the TMDOCK method[63] (Table S6).

2.3. Step 2: Modeling of the Human JAK2 Homodimer. An important component of the active signaling complexes are JAK nonreceptor kinases that are constitutively associated with ICDs of cytokine receptors. Each member of the JAK family is composed of four structural domains: a FERM (four-point-one, ezrin, radixin, moesin) domain, a Src-homology 2-like domain (SH2L), a pseudokinase domain (PK or JH2), and a catalytically active tyrosine kinase domain (TK or JH1) (Figure 8A,B). JAK2 is the main nonreceptor kinase interacting with class 1 homodimeric cytokine receptors. Though experimental crystal structures were obtained for individual domains of JAK2, there is no experimental structure for the full-length JAK2 and its active dimer. Computational models of the full-length JAK2 dimer have been proposed using long-time scale molecular dynamics simulations.[68]

Recently, cryo-EM structures were obtained for the full-length mouse JAK1 active dimers in complex with the ICD fragments of interferon λ receptor1 (INFλR1) stabilized by the oncogenic V657F mutation (analogous to V617F mutation of JAK2) and nanobodies[69,70] (Figure 8A). Both structures (PDB IDs: 7T6F, 8EWY) demonstrate that the dimerization interface is formed between β-structural N-lobes of PK domains. The V657F oncogenic mutation stabilizes the dimeric state by participating in a cluster of contacting aromatic residues at the dimerization interface. Interestingly, these dimer structures demonstrate the different relative positions of TK domains connected by the long flexible loops to PK domains that can be closer together or farther apart from each other. Such positional flexibility of the TK domains may be essential to facilitate their *trans*-phosphorylation at tyrosine residues from the activation loop, the key step in JAK activation, and for the subsequent phosphorylation of tyrosine residues of associated receptors and STAT proteins.

https://doi.org/10.1021/acs.jcim.3c00926
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

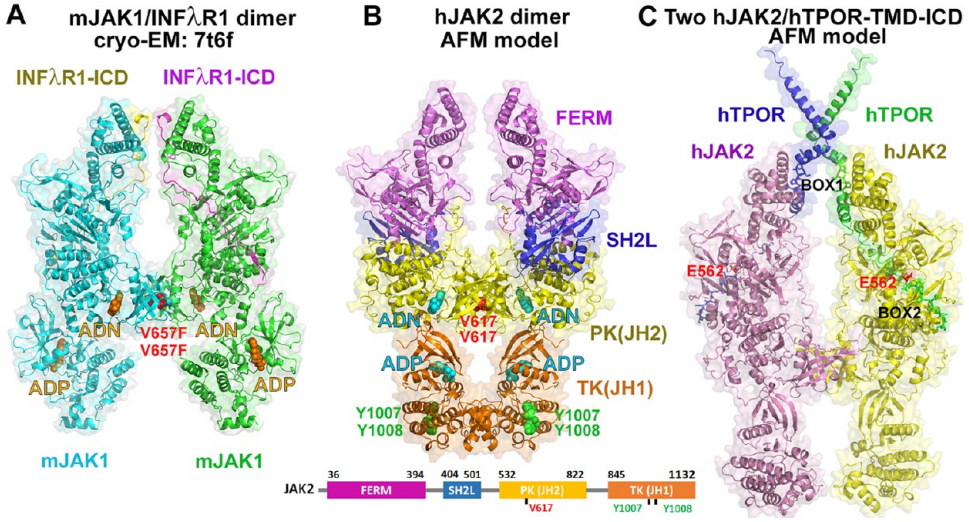


Figure 8. Experimental and computational models of the active dimeric complexes of full-length mouse JAK1 (A) and human JAK2 (B, C). Two JAK1 or JAK2 subunits dimerize via the formation of antiparallel β-strands between β-structural N-lobes of PK domains. These domains contain the oncogenic V657F mutation in mJAK1 or the wild-type V617 residue in hJAK2 (shown by red spheres). (A) cryo-EM-based model (PDB ID: 7T6F) of the mJAK1 dimer in complex with peptides from ICDs of interferon λ receptor1 (INFλR1). JAK1 subunits are colored green and cyan; INFλR1-derived peptides are colored yellow and purple; ligands (adenosine (ADN) and adenosine-5′-diphosphate (ADP)) are shown by orange spheres. (B) AFM-generated model of the human JAK2 dimer with colored FERM, SH2L, PK and TK domains. The cyan spheres indicate ADN in the PK domains and ADP within the TK domains. The green spheres indicate tyrosines (Y1007 and Y1008) from the flexible 15-residue activation loops (residues 993−1017) of TKs that undergo *trans*-phosphorylation during JAK2 activation. The positions of small molecules, ADP and ADN, are similar to those in the cryo-EM structure of mJAK1 (A). The lower panel shows a schematic representation of the hJAK2 domain architecture. (C) Structure of the JAK2 dimer based on AFM-generated models of two hJAK2 monomers, each in complex with a part of ICD domain of hTPOR. The TMDs of hTPOR are also shown. To form the dimer, the calculated monomeric models were superposed with the FERM-SH2L domains of the JAK2 dimer shown in panel B. Each JAK2 subunit (colored yellow and pink) is constitutively bound to the ICDs of TPOR (green or blue) via Box1 and Box2 motifs. E582 (colored red) of TPOR occupies the aberrant binding pocket for phosphorylated tyrosine in the SH2L domain of JAK2. Protein molecules are shown by semitransparent surface and cartoon representations.

At the second step of the AFM modeling, dimers of the full-length human JAK2 were generated with and without short ICD fragments of receptors (Figure S2 and Table S2). The presence of the short receptor fragments did not affect the results of the calculations. The models of the hJAK2 dimer were similar to the cryo-EM-based structures of the mJAK1 dimer (Figure 8B, Table S2), but the distances between two symmetric FERM domains in the models (L224 Cα−Cα distances) varied from 30 to 60 Å (Figure S2). We selected the model of hJAK2 dimer with the minimal distance, similar to 30 Å observed in the experimental structure of the mJAK1 dimer (PDB ID: 7T6F). The dimerization interface in the model was formed by two PK domains, similar to that in the mJAK1 dimer. The dimer is stabilized through association of β-strands connecting the SH2L and PK domains (residues 534−538), two N-lobes, and a C-helix of the PK domain. The oncogenic V617F mutation is located at the PK dimerization interface where two V617F residues of the mutant form a cluster with four aromatic residues (F537 and F595 from each JAK2 subunit), which strengthen PK-PK interactions in the JAK2 dimer (Figure S6). In the selected AFM model of the active hJAK2 dimer (Figure S2, Table 1), TK domains are close to each other, just as in the very recent cryo-EM-based structure of the mJAK1 dimer (PDB ID: 8EWY).[70] The TK-TK interactions in this model can facilitate the *trans*-phosphorylation at tyrosine residues from the activation loops of both TK domains, as discussed by Caveney et al.[70]

2.4. Step 3: Modeling of Monomeric JAK2 Complexes with Receptor TMD and ICDs. The third step included building a complex for each of five receptors composed of a single molecule of JAK2 and receptor TMD and ICD domains; in some cases, a membrane-proximal ECD domain was also included (Figure S2). The AFM models of JAK2 monomers superimpose well with the FERM-SH2L crystal structure of hJAK2[71] and with FERM-SH2L-PK domains of the cryo-EM structure of the mJAK1 dimer[69] or the modeled hJAK2 dimer: the Cα-RMSD values were less than 2.5 Å (Table S2). Interestingly, a few models generated by AFM-V3 represented the more compact autoinhibited (inactive) conformation of JAK2 with the kinase (TK) domain located close to the FERM-SH2L domains and interacting with the PK domain near the kinase active site. This JAK2 domain arrangement is similar to that observed in the crystal structure of PK-TK module of TYK2 (PDB ID: 4OLI).[72]

It has been assumed that the JAK2 FERM-SH2L domains determine the specificity of receptor binding by engaging the receptor Box1 and Box2 cytoplasmic regions.[71] Indeed, AFM-generated models demonstrate that each receptor interacts with JAK2 via the hydrophobic "switch" residues at the TM helix ends, such as L^{253}, I^{257}, and W^{258} in mEPOR,[38] Box1 residues positioned along the α3 of the FERM domain, some α-helical fragments from the interbox region interacting with FERM α2 and α4, and Box2 residues located in the groove in the SH2L domain (Figures 5, 8C, S7, and S8). This membrane-proximal ICD region in cytokine receptors represents the minimal functional core for signal transduction.[45] It was shown that the "P*xx*P*x*P" Box1 motif is essential for binding and activation of JAK2, while the hydrophobic "switch" motif, the acidic and hydrophobic

https://doi.org/10.1021/acs.jcim.3c00926
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

residues from Box2, and several interbox residues are required for JAK2 activation[45] (Figure 4E).

Two interbox α-helices from the JAK2-TMD/ICD model of hEPOR (Figure 5A) overlapped well (Cα-RMSD of 0.9 Å) with the same helices observed in the crystal structure of hEPOR ICD peptide in complex with JAK2 FERM-SH2L domains (PDB ID: 6E2Q).[73] Interbox α-helices found in JAK2-TMD/ICD models of hPRLR and hGHR (Figure 5B,C) are supported by NMR studies of ICD-derived peptides in lipid vesicles.[74,75] Additionally, in the AFM models for TPOR and EPOR, a glutamic acid preceding Box2 (E562 in TPOR and E301 in EPOR) bind to the aberrant phosphotyrosine binding pocket of SH2L (Figures 8C and S8), similar to interactions observed in the crystal structure (PDB ID: 6E2Q).[73]

The unfolded part of receptor after Box2 is not bound to JAK2 and remains highly structurally flexible, which allows tyrosine residues located in this region to enter the catalytic site of TK to be phosphorylated (Figures S7 and S9). Two TPOR tyrosines, Y631 and Y626, were identified as primary and secondary phosphorylation sites, respectively, while phosphorylation of Y591 was shown to participate in receptor downregulation.[10] Interestingly, in the AFM model, the unphosphorylated activation loop of TK (residues 997−1018) partially occludes the catalytic site of the TK domain (Figure S9). Therefore, we suggest that activation of the TK domain after its *trans*-phosphorylation could be induced by the movement of the phosphorylated activation loop away from the catalytic site due to electrostatic interactions between phosphotyrosines (pY1007 and pY1008) and adjacent charged residues. Similarly, phosphorylation of the activation loop in receptor tyrosine kinases relieves the inhibition caused by insertion of unphosphorylated loop into the kinase active site.[76]

2.5. Step 4: Assembly of the Ligand-Receptor-JAK2 Complexes for Five Receptors. Assembly of the final structure of the full-length cytokine-receptor-JAK2 complexes included several substeps (shown for STEP4 in Figures 2 and S2). We first produced the model of the ICD-kinase dimer by superposing two JAK2-TMD/ICD receptor units (model 3) obtained at the step 3 with FERM-SH2L domains of the JAK2 dimer (model 2) selected at the step 2. Second, we joined the models of the ICD-kinase dimer and the ligand-bound complex (model 1 selected at the step 1) by superimposing C-termini of their TM α-helices and adjusting conformations of connecting residues between the TM dimer and Box1 motifs. Third, to improve the PK-PK dimerization interface, we replaced PK and TK domains in the final model by the corresponding PK-TK dimeric structure taken from the active JAK2 dimer model (model 2 selected at the step 2). Finally, we refined the structures using local energy minimizations with CHARMM c47b2.

The final models of cytokine-receptor-JAK2 complexes are close to the corresponding experimental structures of extracellular receptor complexes and JAK2 dimers (Cα-RMSD was from 1.3 to 3 A, Table 1). The models are also consistent with key residues involved in packing of TM α-helices and extracellular domains (Tables 1, S3−S6) and other published experimental data, as described in the Results (see steps 1 to 3 above) and Discussion.

2.6. Setting up All-Atom MD Simulations. AFM modeling uncovered the conformational heterogeneity of cytokine receptors, especially in the region with low reliability such as loops connecting protein domains, ICDs, and TM α-helices (Figure S3). Though we selected one final model of signaling complexes for each receptor studied (Figure 1), other AFM-generated models for active ligand-receptor complexes as well as inhibited and active JAK2 conformations with different positions of TK domains, activation loops, and JAK2-bound receptor ICDs (Figures S7 and S9), may represent different states or snapshots of the conformational dynamics of receptor complexes during their activation. The all-atom molecular dynamics (MD) simulations of these complexes in realistic membranes may shed light on structural transitions between different activation states. Particularly important are the most flexible and the least reliably modeled parts of these complexes, such as loops, ICDs, and ends of TMD regions that may change their conformations upon specific binding of small molecules (e.g., eltrombopag) or interactions with physiologically active lipids (e.g., phosphoinositides).[74,75]

To demonstrate that our models of five cytokine receptor signaling complexes are suitable for all-atom MD simulations in realistic lipid membranes, we built protein-lipid systems for these complexes in an explicit lipid mixture corresponding to the asymmetric mammalian plasma membrane (Table S8). In this study, we used the CHARMM force field for proteins and lipids and TIP3P water model[77,78] with Na^+ and Cl^- ions (see Methods). After successful equilibration of each model in a multicomponent lipid bilayer system, we performed a short production run of 10 ns for each system and deposited the obtained structures together with simulation systems in CHARMM-GUI Archive (https://www.charmm-gui.org/docs/archive/bitopictm). The MD simulations of signaling complexes in realistic membranes used PPM-predicted membrane boundaries. Further studies of the large-scale structural dynamics of these complexes in the plasma membrane using all-atom MD simulations are beyond the scope of this work.

3. DISCUSSION

3.1. AFM-Generated Structures of Signaling Complexes of Homodimeric Cytokine Receptors. We have exploited the power of the Alpha Fold Multimer method[32] to generate three-dimensional (3D) models of active ligand-receptor-kinase signaling complexes of five homodimeric class 1 cytokine receptors. The models are consistent, at the level of atomic details, with various experimental data, such as available crystal structures of ligand-receptor complexes, cryo-EM-based structures of homologous JAK1 dimers, NMR studies of TM helix association, along with data about various protein constructs, constitutively active and disease-associated mutants, Ala-, Cys-, and Asn-scanning mutagenesis, cross-linking, NMR studies of residues of TPOR TM domain interacting with eltrombopag and its analogues, and others (see Results). Moreover, we used experimental data for selection of best models and analysis of molecular details of complex assembly and activation.

The obtained structural models reveal the quaternary structures of full-length protein complexes. The complexes are well-defined continuous structures extending from ligand-stabilized ECDs of receptors to the large intracellular JAK2 dimer via a long membrane-spanning TMD. Ligand, the cornerstone of a complex, holds together two receptor-kinase units, providing rigidity and stability to the whole structure (Figure 1). The first 50−60 residues in the intracellular loops of receptors are bound to a groove on the JAK2 surface and

https://doi.org/10.1021/acs.jcim.3c00926
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

therefore have a fixed structure, while the remaining residues are apparently disordered and flexible, which facilitates the phosphorylation of specific tyrosine residues in the receptors, followed by their binding to SH2 domain-containing proteins from the JAK-STAT signal transduction pathway.

Despite the overall similarity of these five signaling complexes, they are different in many aspects (Figure 1). They have different domain compositions. The receptor-ligand binding stoichiometry is 1:2 for all receptors, but 2:2 for the CSF3-CSF3R complex. Finally, the mutual arrangements of two TM α-helices in the TM dimers are receptor-specific. The significant piston movement of one TM helix relative to the other caused by ligand asymmetry is present only in short-chain receptors where the ligand-binding domains are located close to the membrane. We suggest that this piston movement does not have a significant functional role.

The packing modes of ECDs and TMDs in dimers are receptor-specific and depend on residue compositions and functional states and upon the nature and stoichiometry of the bound ligands (Table 1). For example, the model of the ligand-bound hPRLR complex has a right-handed TM α-helix arrangement, while the packing of TM α-helices in models of the active state of all other receptors is left-handed. Modeling of the same hPRLR receptor with three different ligands demonstrates slightly different helix orientations of the right-handed TMD dimers in the complexes (Table S2). Altered TM α-helix arrangements in the active receptor dimer may induce different physiological responses, as shown using various *Put3*-fusion constructs of mEPOR[49] and mTPOR.[56]

Furthermore, the mutual arrangements of two TM α-helices are different in models of ligand-bound (active) and ligand-free (inactive) complexes (Tables S2, S4, and S6). These results are consistent with experimental observations of distinct rotational positions of the TM helices in the active vs inactive dimers of mEPOR, mTPOR, and hGHR.[47,49,56]

Our modeling confirms the common concept that cytokine receptor signaling complexes undergo significant structural changes during their binding to ligands, dimerization, and activation as well as during constitutive ligand-independent activation of mutants. We explored the flexibility of cytokine receptor complexes by calculating and comparing the following models: (a) ligand-bound and ligand-free receptors, (b) PRLR with three different ligands, (c) ECDs and TMDs of constitutively active receptor mutants, and (d) JAK2 in monomeric and dimeric states with and without bound receptor fragments (Tables S2 and S4, Figures 2, 7, 8, and 10). The largest structural heterogeneity of ligand-bound receptors was observed in packing of TM α-helices, as mentioned above, and in ICDs (e.g., Box1, Box2, interbox, and Tyr-carrying regions), while ligand-free receptors also demonstrated different conformations of ECDs. Modeling of monomeric JAK2 produced structures with different relative orientations of TK and PK (Figure 10), while modeling of dimeric JAK2 demonstrated different distances between FERM domains (Figure S2). Each of the currently selected states represents a snapshot of conformational dynamics of these complexes, which was confirmed by experimental studies. However, the physiological relevance of alternative nonselected AFM-generated structures requires additional experimental validation.

An important structural aspect of membrane proteins, such as single-pass TM cytokine receptors, is their spatial positions in membranes. The lack of membrane boundaries in AFM models as well as in experimental structures of membrane proteins requires an application of computational approaches for the prediction of membrane boundary positions. One of the most advanced method for positioning proteins in membranes is the PPM 3.0 method that orients protein structures in planar and curved membranes by optimizing their transfer energies.[34] We applied PPM 3.0 to calculate the positions of the TM domains of the final models of five signaling complexes of cytokine receptors (Figures 4 and 6). The hydrophobic thicknesses of TM α-helices of EPOR, GHR, PRLR, CSF3R, and TPOR receptors, which matches the distances between lipid carbonyls in two membrane leaflets, are 39.2 Å, 39.8 Å, 41.2 Å, 33.0 Å, and 32.2 Å, respectively. The large hydrophobic lengths of TM α-helices of cytokine receptors may suggest their preferentially localization in membrane rafts characterized by the increased membrane thicknesses.

3.2. Activation Mechanism. The exact molecular mechanism of cytokine receptor activation that triggers JAK2 activation (Figure 9) remains a matter of debate. It is accepted that dimerization is essential but not sufficient for receptor activation,[79] and that receptor dimerization is driven by the association of the TM α-helices.[46,80−82] Moreover, receptor activation requires a specific orientation of receptor TM helices to form a productive dimeric state that brings the ICD-bound JAK2 molecules into positions competent to initiate intracellular signaling.[49,56] However, it remains controversial whether the activation mechanism involves the ligand-induced receptor dimerization (activation model 1, I-III-IV pathway in Figure 9) or conformational changes in preformed inactive receptor dimers upon ligand binding (activation model 2, I-II-IV pathway in Figure 9).[83]

In activation model 1, receptor dimerization occurs only in the presence of appropriate ligands. This model has been recognized for many years and gained additional support in recent studies of dimerization of cytokine receptors in living cells using single-molecule fluorescence microscopy.[19] These studies demonstrated that human TPOR, GHR, and EPOR exist as monomers at the physiological receptor densities in the plasma membrane, while the basal dimerization level is negligible. Evaluation of energy contributions showed that binding of TPO to TPOR provides the main contribution to the total dimerization energy. It was also estimated that the intrinsic dimerization affinity of TPOR-JAK2 subunits is low, but constitutively active oncogenic mutations in the dimerization interface of JAK2 (V617F, M335I, H538L, K539L, H587N, C618R, and N622I) and in the TMD of TPOR (W515 K) provide additive stabilizing free energy contributions which promote TPOR-JAK2 dimerization and formation of the active signaling complex.

An alternative activation model 2 suggests that receptor predimerization occurs in the absence of ligands, and dimer reorganization follows after ligand binding. This model has been proposed based on the extensive structural, biochemical, and mutagenesis studies of different cytokine receptors, including human and mouse EPOR,[49,82] human GHR,[47,79,84] human PRLR,[85] human and mouse TPOR,[53,54,56] and their TMD fragments.[46,50,57] Binding of a specific ligand to the inactive preformed dimer is required for conformational changes and reorientation of receptor TM α-helices to form an active dimeric state that induces proximity and dimerization of associated JAK2 subunits.[83]

There are several experimental observations that challenge the hypothesis of binding of ligands to pre-existing dimers

https://doi.org/10.1021/acs.jcim.3c00926
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

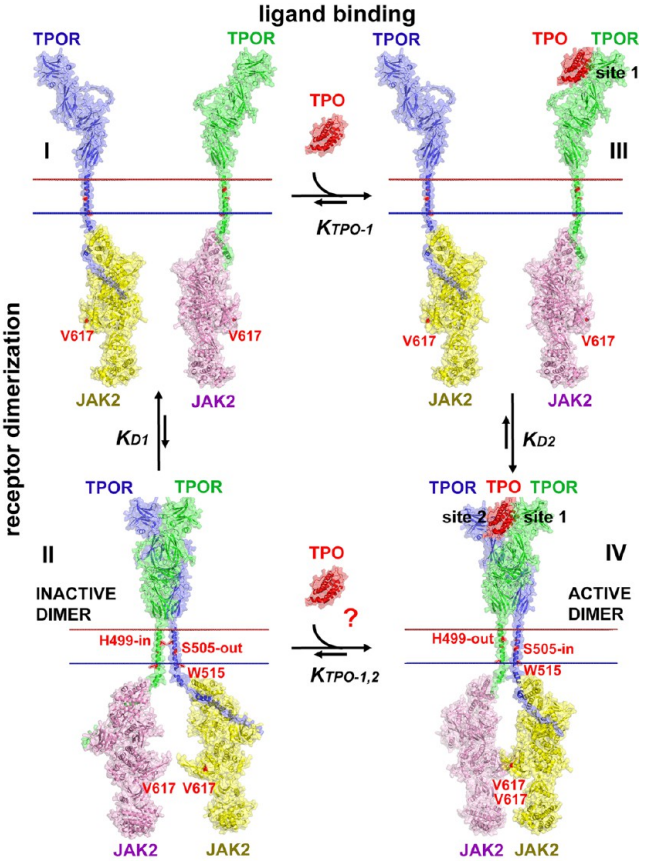


Figure 9. Suggested activation mechanism of homodimeric class 1 cytokine receptors (exemplified by TPOR) based on the AFM modeling and published live-cell dimerization assay.[19] In the absence of the TPO ligand, TPOR receptors are mainly in the monomeric state (state I),[19] even though some dimerization may occur (state II). Ligand binds first via site1 to one receptor chain (state III), and then via site 2 to the second receptor chain. This leads to stabilization of the active receptor dimer (state IV) with specific rotational orientations of TM α-helices whose intracellular ends bring two JAK2 molecules close to receptor ICDs (Box1 and Box2 motifs). This enables the dimerization and activation of JAK2. Protein molecules are shown in semitransparent surface and cartoon representations, colored red for TPO, blue and green for TPOR subunits, yellow and pink for JAK2 subunits. Residues that are involved in the constitutive activation of receptor (S505, W515) and JAK2 (V617) or those that regulate the formation of the active TMD dimer (H499) are colored red.[53,54,60,86,87]

(model 2). First, at physiological concentrations of receptors at the cell surface, the fraction of monomeric receptors (state I) is much higher than of preformed dimers (state II).[19] Second, the bell-shaped dose-dimerization curve[10,19] is consistent with the two-step ligand binding to monomeric receptors (state III): initially via site 1 to one chain, then via site 2 to the recruited second chain, which leads to formation of the ligand-receptor complex (state IV). This dimerization is inhibited by the presence of excess ligand that binds via high-affinity site 1 to receptors, blocking further receptor dimerization via site 2 interactions.

AFM-based modeling provides an insight into possible activation mechanisms. The modeling uncovered that ligand-free dimers could be formed for many receptors, but such structures have occluded ligand binding pockets incapable of accommodating large cytokine molecules along with an unproductive arrangement of TM helices. This is in line with the notion that ligand-free ECDs lock receptors in the inactive states, as PRLR and TPOR variants lacking large parts of ECDs are constitutively active.[85,88] Ligand binding to preformed dimers with a closed ligand binding pocket would require a significant rearrangement of their ECDs and TMDs and possibly even the dissociation of two receptor molecules. Based on these findings, AFM modeling generally provides more support for model 1 of the ligand-induced dimerization and activation, the I-III-IV pathway (Figure 9), at least for class 1 homodimeric cytokine receptor complexes.

Nonetheless, ligand binding to preformed dimers (I-II-IV pathway, Figure 9) could represent an alternative noncanonical activation pathway that is used by some receptors in particular cases. For example, we have recently proposed the two-step TPOR activation by the MPN-associated calreticulin mutants (CRTmut)[20] that are suggested to form a CRTmut-TPOR (2:2) complex.[89] The formation of the CRTmut-TPOR active complex takes place in endoplasmic reticulum (ER) membranes, where the local density of preformed dimers of immature TPOR molecules may be relatively high. We propose that at the first step, a dimer of CRT mutants binds to the preformed TPOR dimer with the occluded ligand binding pocket via interactions with immature mannose-rich glycans linked to N117 residues of both receptor chains. Then, as the second step, CRTmut ligands induce rearrangement of receptor ECDs followed by the insertion of positively charged C-terminal helices of CRTmut into the unlocked binding pocket between D1 and D2 domains of both TPOR chains. A generally similar model for the final active CRTmut-TPOR (2:2) complex has been proposed based on a combination of experimental and computational approaches.[21] The CRTmut-TPOR (2:2) active complex formed by ligand binding to the preformed receptor dimer, together with ICD-associated JAK2 traffics from ER to the plasma membrane via the secretory pathway.[90] Whether other cytokine:receptor complexes might also use the I-II-IV pathway (Figure 9) remains to be further examined.

Importantly, even the ligand-induced dimerization pathway (I-III-IV) implies significant structural changes during the formation of an active signaling complex. The rotational and translational movements of both ECDs relative to each other and side chain rotations are required to adjust the binding pocket for an asymmetric ligand. However, such molecular movements do not change the overall structure of individual monomeric subunits because they are nearly identical in the active and inactive dimers (Cα-RMSD < 0.6 Å). This rotational and piston motions propagate toward the membrane leading to a receptor-specific arrangement of TM α-helices and adjacent ICD Box 1 residues to bring together ICD-bound JAK2 molecules in an orientation appropriate for productive JAK2 dimerization, activation, and subsequent triggering of signaling events. To the contrary, in the ligand-free inactive dimers, receptor-associated JAK2 subunits are spatially separated and have incorrect orientations, which prevents their dimerization.

The results of the modeling are consistent with FRET studies of hGHR signaling.[47] In the ligand-free inactive conformation of the GHR-JAK2 (2:2) complex, JAK2 subunits allow FRET reporters (mCit and mCFP) covalently attached to the receptor C-termini (37 residues below the Box1 motif) to approach each other at a distance of ~47 Å from one side of the JAK2 dimer (Figure S10A). The active GH1-GHR-JAK2

https://doi.org/10.1021/acs.jcim.3c00926
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

(1:2:2) complex has a TM dimer interface different from that in the inactive dimer by ∼100° rotation of F273 toward the dimerization interface (from *e*- to *d*-position of the heptad repeat motif). This TM helix rotation promotes JAK2 dimerization via PK-PK interactions. The tightly packed JAK2 dimer prevents FRET reporters from coming closer to each other (the distance between chromophores is 75 Å) (Figure S10B). Separation of FRET reporters in the active GHR-JAK2 signaling complex is in agreement with experimental data.[47]

3.3. Mapping of Oncogenic Mutant onto Signaling Complexes. Mutations of cytokine receptors and JAK2 have been implicated in dysregulation or chronic activation of cytokine pathways leading to severe pathologies, including hematological malignancies, growth abnormalities, and aberrant immune responses.[5,43,90−93] Disease-associated mutations can be classified as loss-of-function (LOF) and gain-of-function (GOF) mutations. The latter usually cause the constitutive activation of cytokine receptors and JAK kinases.[10,43,94] Mapping of known oncogenic missense mutations onto the AFM-based models of active ligand-receptor-kinase complexes may shed light on possible molecular mechanisms of pathological effects of these mutations.

The majority of GOF mutations in JAK2[94] are located in the PK domain, regions involved in the PK-TK inhibitory interface (Figure 10, spheres colored cyan) and PK-PK dimerization interface (Figure 10, spheres colored blue). Mutations in the PK-PK dimerization interface are mainly associated with MPNs, with a single point mutation, V617F, identified in more than 95% of polycythemia vera (PV) cases and 50−60% of essential thrombocythemia (ET) and primary myelofibrosis (PMF) cases.[95] V617F together with adjacent aromatic residues, F537 and F595, forms a hydrophobic cluster of six aromatic residues from both subunits that stabilizes the JAK2 dimer (Figure S6). Other oncogenic mutations are found near this cluster (M535I, H538L, K539L, and N622I). They strengthen the hydrophobic interactions at the dimerization interface. Indeed, all these mutations induce ligand-independent activation and dimerization of the receptor forming the signaling complex with JAK2.[19] Mutations of the PK-TK inhibitory interface are located between the N-lobes of PK and TK. The interface is formed by hydrophobic and charged residues, including the R683-D873 pair. Mutations of interfacial residues, including this ionic pair, can weaken PK-TK interactions and facilitate movement of TK from the inactive (Figure 10A) to the active conformation (Figure 10B). Activation of JAK2 due to relieved inhibitory function of the PK domain represents a possible mechanism that triggers MPNs, acute myeloid leukemia, and acute megakaryoblastic leukemia caused by these mutations.[94]

Oncogenic hTPOR mutations are found in all receptor domains, with LOF mutations located mainly in ECD and ICD and GOF mutations clustered at dimerization interfaces created by TM α-helices and loops inside the ECD (Figure 10). LOF mutations usually induce thrombocytopenia, while GOF mutations are mainly associated with thrombocytosis and MPNs, such as PMF and ET.[58] LOF mutations, including K39N, R102P, P106L, W154R, R257C, and P635L, show low cell-surface expression, possibly due to defects in receptor folding or trafficking.[10] Another LOF mutation, F104S located in the ligand-binding pocket, impairs TPO binding to the ECDs.[10,58] The most common GOF mutations identified in

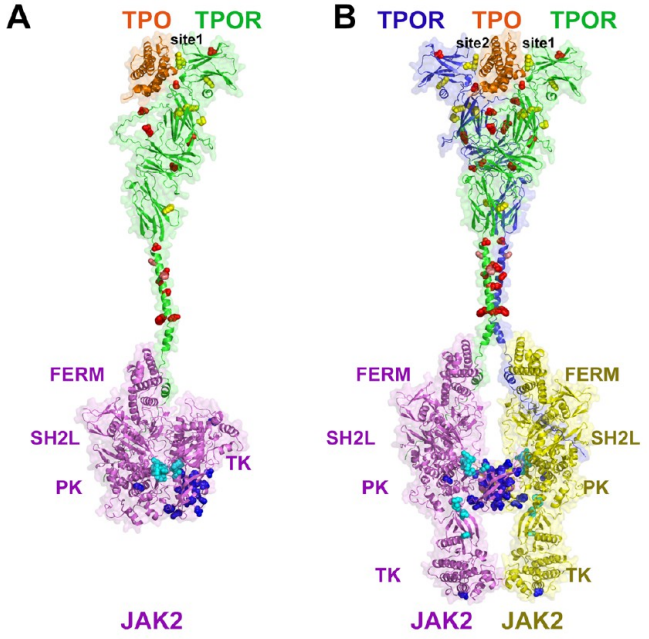


Figure 10. Mapping of disease-associated mutations onto AFM models of TPOR complexes. (A) TPO-TPOR-JAK2 complex (1:1:1). TPO binds to TPOR via its high-affinity site 1, forming a binary (1:1) inactive complex. JAK2 is constitutively associated with the ICD of TPOR. JAK2 is in the autoinhibited (inactive) form, with the PK domain interacting with the TK domain near the kinase active site, which inhibits the TK's catalytic activity. (B) Active signaling complex of TPO-TPOR-JAK2 (1:2:2) with TK in the active state. TPOR dimer is in the TPO-bound (active) conformation with TMDs forming the left-handed dimer. TPOR mutation sites associated with myeloproliferative neoplasms (MPNs)[10] are shown as yellow spheres for lost-of-function mutations, red spheres for gain-of-function mutations, and pink spheres for "enhancer" mutations in the TMD (S493, H299). MPN-associated mutations of JAK2 are shown as blue spheres for mutations at the PK-PK dimerization interface and cyan spheres for mutations at the PK-TK inhibitory interface. Molecules are shown as cartoon and semitransparent surface representations colored orange for TPO, blue and green for TPOR subunits, and yellow and pink for JAK2 subunits.

MF and ET patients, S505N and W515 K/L/A/R, are located within the TMD.[10] These mutations cause constitutive activation of TPOR due to stabilization of the productive TMD dimer. There are also several "enhancer" mutations in TM α-helices that stabilize the active mode of helix dimerization,[54,87] which may enhance the pathological effect.

A more detailed analysis of GOF and LOF mutations in the context of competing structures of active signaling complexes will add to our understanding of the role of disease-associated mutations in cytokine-induced JAK-STAT signaling cascades. Knowing the molecular mechanisms of oncogenic mutations will guide the development of new cancer therapeutic agents.

4. CONCLUSIONS

Using the transformative ability of the AlphaFold2Multimer to predict structures of proteins and their complexes with high accuracy, we generated models of full-length active signaling complexes for human homodimeric cytokine type 1 receptors, EPOR, GHR, PRLR, TPOR, and CSF3R. Analysis of the resulting models of signaling complexes, as well as models of inactive dimers, examines, in a structural context, highly debated questions related to the mechanism of cytokine-

https://doi.org/10.1021/acs.jcim.3c00926
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

initiated activation that triggers the JAK-STAT signaling pathway in cells.

First, we demonstrate that although ligand-free receptors may form stable inactive dimers, the ligand binding pocket in such dimers is occluded, thus preventing ligand binding. At low cell-surface receptor densities, cytokines are more likely to bind and activate monomeric receptors via a two-step process: ligand-induced receptor dimerization accompanied by conformational rearrangements in ECDs and TMDs. This may represent the canonical receptor activation mechanism. A noncanonical activation route through ligand binding to preformed inactive dimers can also occur in specific cases, such as activation of TPOR by the oncogenic CRTmut.[20]

Second, we can picture the complete process of JAK2 activation. The process starts from the receptor-induced proximity of two JAK2 molecules, followed by dissociation of the PK-TK inhibitory complex in each JAK2 molecule. Subsequently, the dimerization of two symmetric PK domains stabilizes the JAK2 dimer leading to *trans*-phosphorylation of both TK activation loops and their movement away from the TK active sites, enabling tyrosine phosphorylation of receptors and other associated proteins.

Many other aspects of receptor conformational dynamics were also clarified, including atomic details of a specific binding of receptor Box1 and Box2 ICD motifs to JAK2, the ancillary role of the piston TM helix movement in short-chain cytokine receptors, and the role of GOF mutations in stabilizing dimerization interfaces in receptor and JAK2 molecules. The mode of interaction of two molecules of eltrombopag, an FDA-approved TPOR agonist, with the TM α-helical dimer of hTPOR is proposed.

The computational modeling described in this study also uncovers certain limitations of the AFM method. The current versions of the AFM program do not allow a direct modeling of large complexes of multidomain proteins. Therefore, such complexes must be assembled from the smaller AFM-generated parts. Moreover, in the case of multiple alternative models produced by AFM, the selection of the correct structures still requires supporting experimental data. When such data are lacking or insufficient, the modeling of complexes also needs to include comparative analysis of models obtained for sequences of different lengths, with different sets of structural domains, mutants, and subunit stoichiometries. It is anticipated that future versions of the AFM program will overcome some of these limitations, allowing predictions of multiprotein complexes directly and with improved accuracy.

Despite the limitations, the computational approaches used in this work can be applied in the future to modeling of cytokine receptor complexes from other families as well as other large functional assemblies of single-pass TM proteins that trigger different intracellular pathways. Knowing 3D structures of such complexes is critical for the development of new drugs and therapeutic strategies.

5. METHODS

5.1. Modeling of signaling complexes with AlphaFold-Multimer (AFM). Modeling of active signaling complexes of five cytokine receptors was performed using AlphaFold_2.0_multimer.v2 1.3.0.version (AFM V2), AlphaFold_2.0_ptm 1.5.2 version (AF2-ptm), and more recent AlphaFold_2.0_multimer.v3 1.5.2.version (AFM V3)[32] implemented through ColabFold notebook.[33] ColabFold was downloaded from Github (https://github.com/YoshitakaMo/localcolabfold) together with the environmental databases (https://colabfold.mmseqs.com) and installed on a local computing cluster. ColabFold was run locally using 12, 24, and 48 recycles, MMseq2 for multiple sequence alignments, refinement with Amber, and no templates. The quality of structural models was characterized by the mean of per residue pLDDT score (predicted Local Distance Difference Test) ranging between 0 and 100 that characterizes local structural accuracy,[22,96] as well as using PAE (Predicted Aligned Error) or PAE-derived pTMscore (predicted TM-score) ranging from 0 to1,[97] which correspond to overall topological accuracy. The confidence of the predicted protein-protein interface is assessed by the interface pTM-score (ipTM) ranging from 0 to 1.[26] For each run, 5 models were generated and ranked by ipTMscores (Tables S2 and S3).

The amino-acid sequences from UniProt[98] were used for modeling the following proteins: human JAK2 (UniProt AC: Q60674), five human receptors, hEPOR, hTPOR, hGHR, hPRLR, and hCSF3R (UniProt ACs: P19235, P40238, P10912, P16471, and Q99062, respectively), mouse EPOR (UniProt AC: P14753), six human cytokines, hEPO, hTPO, hGH1, hPRL, hCSH1, and hCSF3 (UniProt ACs: P01588, P40225, P01241, P01236, P0DML2, and P09919, respectively) (Figure S1), and mouse EPO (UniProt AC: P07321). For hTPO, only the erythropoietin-like N-terminal domain (residues 22−174)[10] was modeled, while the glycan domain was omitted. For an easy comparison with published experimental data, sequences of mature proteins (lacking signal peptides) were used for hEPOR, mEPOR, hPRLR, and cytokines, while full-length sequences of immature proteins (carrying signal peptides) were used in all other cases. Intrinsically disordered regions of receptor ICDs beyond Box1 or Box2 motifs were usually removed. For the four receptors except PRLR, one natural ligand was used (Figure S1). For human PRLR, which can be activated by three human hormones, prolactin (PRL), somatotropin (GH1), and placental lactogen (CSH1),[6] three receptor-hormone pairs were modeled (Table S2). The ligand-receptor stoichiometry was 1:2 for all receptors, except CSF3R, for which the 2:2 complex was modeled.

Additionally, AFM was used to obtain models for ligand-free receptor dimers and receptor fragments composed of TM α-helix with juxtamembrane regions, WSxWS motifs, or adding membrane-proximal domains (D4 for TPOR, D6 and D5-D6 for CSF3R). The complete signaling complexes consist of ligands, receptors, and JAK2 at a stoichiometry 1:2:2 (for EPOR, GHR, PRLR, and TPOR) or 2:2:2 (for CSF3R).

AFM modeling was performed in 4 steps (see Results, Figures 2 and S2 for the workflow). To diversify models, we increased sampling for each complex by running different AF2 versions (AFM V2 or V3 and AF2-ptm), by changing random seed numbers and the number of recycles. Each run with a specified seed and a number of recycles produced 5 different models. Any models with unfolded, noninteracting, or incorrectly oriented monomers were excluded from the subsequent analysis. At every step of the modeling, we selected the single best model based on its agreement with available experimental data. The consistency with experimental data was assessed using Cα-RMSDs with crystal structures, DockQ scores, and agreement with mutagenesis and other data on TM helix packing. We also compared sets of ligand-receptor interacting residues predicted by the models with the corresponding sets obtained by mutagenesis (Table S5, Figures

https://doi.org/10.1021/acs.jcim.3c00926
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

3 and 6) and checked the proper binding of receptor intracellular loops with JAK2 domains: Box-1 residues with the FERM domain (Figure 5), Box-2 residues with the SH2L domain (Figure S8), and specified tyrosine residues with the TK domain (Figures S7 and S9). An important additional criterion was the correct formation of numerous intramolecular disulfide bonds in the extracellular domains of these receptors (Table S1, Figures 3, 4, and 6). During this procedure, our goal was to obtain the most accurate models. Therefore, we were gradually increasing sampling as long as this process improved the consistency of the best model with experimental data. We found that a set of 30 models for each complex was sufficient for this purpose.

For example, the output models of ligand-bound short-chain receptors (EPOR, GHR, and PRLR) demonstrated similar ECD structures, reproducing corresponding crystal structures, but showed diverse arrangements of their TM α-helices (Figures S11 and S12). Models with strongly interacting TM α-helices were selected and compared with available protein engineering and mutagenesis data for helix packing motifs. Then, for each complex, one model that best satisfied the experimental data was chosen for the subsequent modeling of signaling complexes with JAK2 dimers, as described in the Results.

Hence, during the model selection, we relied mostly on the consistency of models with experimental data rather than on predicted ipTM scores. This strategy was based on the previous observations that the predicted contact scores often fail to identify the true models,[99] and experimental verification is usually required to validate and justify the physiological relevance of AF2-predicted conformations of membrane proteins.[100]

We found that the V3 version produced better models for TPOR complexes (Figure S13A), but not for other receptors where the results with V2 and V3 versions were rather similar. AF2-ptm produced better models of disulfide-linked dimers of R130C, D133C, and E134C mutants of hEPOR than AFM V2 and V3. Calculations of full-length CSF3:CSF3R (2:2) tetrameric complexes produced models with spatially separated TM helices (Figure S13B). Therefore, these complexes were produced by a two-step procedure (see the Results).

The final model of each signaling complex was refined using local energy minimization with CHARMM c47b2[101] through the pyCHARMM python module[102] to remove atom hindrances. The minimization was conducted for 1000 steps (dielectric constant $\varepsilon = 63$) with Cα fixation using the CONS HARM command in CHARMM (the force of 20), and fixation of cysteines using the CONS FIX command. The structures were prepared using the PDB input manipulator on CHARMM GUI[103,104] and then converted back using MMTSB convpdb.pl.[105] The models of all complexes are available through the Membranome database.[106]

AFM-generated models were superimposed with each other and with available experimental structures by the *align* method of PyMOL (1.8.4.1) (www.pymol.org) with default parameters (cutoff = 2.0, cycles = 5, transform = 1), PDBeFold (SSM) server,[107] and the US-align web server.[108] Membrane boundaries were calculated by the PPM web tool.[34] The eltrombopag drug was docked to TPOR TM domain manually using PyMOL to satisfy contacts between the drug and specific residues of TPOR identified by NMR.[55,57] The fractions of native contacts (Fnat) and DockQ scores in the final models of the complexes were evaluated using available crystal structures and DockQ program[36] (Table S3). All of the figures were generated by PyMOL.

5.2. Modeling of Lipid Bilayer Systems with Signaling Complexes. The final models of cytokine receptor signaling complexes were embedded into the lipid bilayer composed of explicit lipids forming the asymmetric mammalian plasma membrane.[109] The inner membrane leaflet was composed of phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylinositol (PI), phosphatidylserine (PS), phosphatidic acid (PA), sphingomyelin (SM), and cholesterol (CHOL), while the outer membrane leaflet was composed of PC, PE, SM, CHOL, and glucosylceramide (GlcCer) (Table S7). The TIP3P water model was used to simulate explicit water molecules, while the number of ions (Na^+ and Cl^-) incorporated corresponded to the physiological concentration (150 mM NaCl). Initial membrane structures were built using the CHARMM-GUI *Membrane Builder*.[110−112] The simulations were performed with all-atom CHARMM36m force field,[113] and executed utilizing OpenMM.[114] Following the default equilibration protocol of CHARMM-GUI,[101,104,115] we first applied NVT dynamics with a time step of 1 fs (fs) for 250 ps (ps). Subsequently, we employed the NPT ensemble with a time step of 1 fs and then with a time step of 2 fs. During the equilibration processes, the protein, lipid, and water molecules were subjected to the restraint potentials of their position and dihedral angles. The force constants associated with these potentials were systematically decreased over time. Ten nanoseconds (ns) production runs were performed for each system utilizing a time step of 4 fs and employing the hydrogen mass repartitioning technique[116] in the absence of any restraint potentials. The SHAKE algorithm was employed for managing bonds involving hydrogen atoms.[117] van der Waals interactions were truncated at a cutoff of 12 Å, with a force-switching function applied between 10 and 12 Å,[118] while electrostatic interactions were calculated using the particle-mesh Ewald method.[119] The manipulation of temperature and pressure (at a standard pressure of 1 bar) was achieved by utilizing Langevin dynamics with a friction coefficient of 1 ps^{-1} and a semi-isotropic Monte Carlo barostat, respectively.

■ ASSOCIATED CONTENT

Data Availability Statement

Our models of five cytokine-receptor-JAK2 complexes in explicit lipids of the mammalian plasma membrane and simulation systems are available in CHARMM-GUI Archive (https://www.charmm-gui.org/docs/archive/bitopictm).

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.3c00926.

13 Figures demonstrating molecular details of AFM models of cytokine receptor complexes and 7 Tables presenting quality metrics of different AFM models, characteristics of receptors and their partners, interaction residues in the binding pocket of TPOR, and lipid composition of the mammalian plasma membrane (PDF)

Model of TPOR (PDB)

Model of PRLR (PDB)

Model of GHR (PDB)

Model of EPOR (PDB)

Model of CSF3R (PDB)

https://doi.org/10.1021/acs.jcim.3c00926
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

AUTHOR INFORMATION

Corresponding Author

Andrei L. Lomize − *Department of Medicinal Chemistry, College of Pharmacy, University of Michigan, Ann Arbor, Michigan 48109, United States;* orcid.org/0000-0002-3044-7597; Email: almz@umich.edu

Authors

Irina D. Pogozheva − *Department of Medicinal Chemistry, College of Pharmacy, University of Michigan, Ann Arbor, Michigan 48109, United States;* orcid.org/0000-0002-3024-9574

Stanislav Cherepanov − *Biophysics Program, University of Michigan, Ann Arbor, Michigan 48109, United States;* orcid.org/0000-0001-9459-1849

Sang-Jun Park − *Departments of Biological Sciences and Chemistry, Lehigh University, Bethlehem, Pennsylvania 18015, United States;* orcid.org/0000-0002-7307-3724

Malini Raghavan − *Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, Michigan 48109, United States;* orcid.org/0000-0002-1345-9318

Wonpil Im − *Departments of Biological Sciences and Chemistry, Lehigh University, Bethlehem, Pennsylvania 18015, United States;* orcid.org/0000-0001-5642-6041

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.3c00926

Author Contributions

The paper was written through contributions of all authors. All authors have given approval to the final version of the paper. I. D. Pogozheva and A.L. Lomize wrote and edited the paper. I. D. Pogozheva, A. L. Lomize, and S. Cherepanov performed molecular modeling. S. J. Park and W. Im equilibrated lipid-protein systems for five cytokine-receptor-kinase signaling complexes for MD simulations. M. Raghavan and W. Im edited the paper.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was funded by the Division of Biological Infrastructure of the National Science Foundation (Award # 1855425 for I.P. and A.L.) and (Award # 2011234 for W.I.) and by the National Institutes of Health (R01AI123957 for M.R.). The authors thank S. Todd for local installation of AphaFold Multimer ColabFold version and design of a web server that was used during this work.

ABBREVIATIONS

AFM, AlphaFold Multimer; CHM, cytokine homology module; CRT, calreticulin; CSF3, granulocyte colony-stimulating factor; CSF3R, granulocyte colony-stimulating factor 3 receptor; CSH1, placental lactogen or chorionic somatomammotropin hormone 1; ECD, extracellular domain; EPOR, erythropoietin receptor; ET, essential thrombocythemia; FERM, four-point-one, ezrin, radixin, moesin; FnIII, fibronectin type III; GH1, somatotropin; GHR, growth hormone receptor; GOF, gain-of-function (mutation); JAK2, Janus Kinase 2; ICD, intracellular domain; IL, interleukin; LOF, loss-of-function (mutation); MD, molecular dynamics; MPN, myeloproliferative neoplasms; NVT, dynamics, constant number, constant-volume, and constant-temperature simulation; NPT ensemble, isothermal-isobaric ensemble; PK, pseudokinase; PMF, primary myelofibrosis; PRLR, prolactin receptor; PV, polycythemia vera; SH2L, Src-homology 2-like SHP, tyrosine phosphatase; SOCS, suppressors of cytokine signaling; TPOR, thrombopoietin receptor; TK, tyrosine kinase; TM, transmembrane; TMD, transmembrane domain; TM-JM, transmembrane and juxtamembrane regions

REFERENCES

(1) Liongue, C.; Sertori, R.; Ward, A. C. Evolution of cytokine receptor signaling. *J. Immunol.* **2016**, *197* (1), 11−18.

(2) Morris, R.; Kershaw, N. J.; Babon, J. J. The molecular details of cytokine signaling via the JAK/STAT pathway. *Protein Sci.* **2018**, *27* (12), 1984−2009.

(3) Schindler, C.; Levy, D. E.; Decker, T. JAK-STAT Signaling: From interferons to cytokines*. *J. Biol. Chem.* **2007**, *282* (28), 20059−20063.

(4) Boulay, J.-L.; O'Shea, J. J.; Paul, W. E. Molecular phylogeny within type I cytokines and their cognate receptors. *Immunity* **2003**, *19* (2), 159−163.

(5) Dehkhoda, F.; Lee, C. M. M.; Medina, J.; Brooks, A. J. The growth hormone receptor: mechanism of receptor activation, cell signaling, and physiological aspects. *Front. Endocrinol.* **2018**, *9*, 35.

(6) Brooks, C. L. Molecular mechanisms of prolactin and its receptor. *Endocr. Rev.* **2012**, *33* (4), 504−525.

(7) Watowich, S. S. The Erythropoietin Receptor: Molecular Structure and Hematopoietic Signaling Pathways. *J. Investig. Med.* **2011**, *59* (7), 1067−1072.

(8) Olsen, J. G.; Kragelund, B. B. Who climbs the tryptophan ladder? On the structure and function of the WSXWS motif in cytokine receptors and thrombospondin repeats. *Cytokine Growth Factor Rev.* **2014**, *25* (3), 337−341.

(9) Varghese, L. N.; Defour, J. P.; Pecquet, C.; Constantinescu, S. N. The thrombopoietin receptor: structural basis of traffic and activation by ligand, mutations, agonists, and mutated calreticulin. *Front. Endocrinol.* **2017**, *8*, 59.

(10) Hitchcock, I. S.; Hafer, M.; Sangkhae, V.; Tucker, J. A. The thrombopoietin receptor: revisiting the master regulator of platelet production. *Platelets* **2021**, *32* (6), 770−778.

(11) Tamada, T.; Honjo, E.; Maeda, Y.; Okamoto, T.; Ishibashi, M.; Tokunaga, M.; Kuroki, R. Homodimeric cross-over structure of the human granulocyte colony-stimulating factor (GCSF) receptor signaling complex. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (9), 3135−3140.

(12) Tsirigotaki, A.; Dansercoer, A.; Verschueren, K. H. G.; Marković, I.; Pollmann, C.; Hafer, M.; Felix, J.; Birck, C.; Van Putte, W.; Catteeuw, D.; Tavernier, J.; Fernando Bazan, J.; Piehler, J.; Savvides, S. N.; Verstraete, K. Mechanism of receptor assembly via the pleiotropic adipokine Leptin. *Nat. Struct. Mol. Biol.* **2023**, *30* (4), 551−563.

(13) Syed, R. S.; Reid, S. W.; Li, C.; Cheetham, J. C.; Aoki, K. H.; Liu, B.; Zhan, H.; Osslund, T. D.; Chirino, A. J.; Zhang, J.; Finer-Moore, J.; Elliott, S.; Sitney, K.; Katz, B. A.; Matthews, D. J.; Wendoloski, J. J.; Egrie, J.; Stroud, R. M. Efficiency of signalling through cytokine receptors depends critically on receptor orientation. *Nature* **1998**, *395* (6701), 511−516.

(14) de Vos, A. M.; Ultsch, M.; Kossiakoff, A. A. Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science* **1992**, *255* (5042), 306−312.

(15) van Agthoven, J.; Zhang, C.; Tallet, E.; Raynal, B.; Hoos, S.; Baron, B.; England, P.; Goffin, V.; Broutin, I. Structural characterization of the stem-stem dimerization interface between prolactin receptor chains complexed with the natural hormone. *J. Mol. Biol.* **2010**, *404* (1), 112−126.

(16) Somers, W.; Ultsch, M.; De Vos, A. M.; Kossiakoff, A. A. The X-ray structure of a growth hormone-prolactin receptor complex. *Nature* **1994**, *372* (6505), 478−481.

(17) Clackson, T.; Ultsch, M. H.; Wells, J. A.; de Vos, A. M. Structural and functional analysis of the 1:1 growth hormone:receptor complex reveals the molecular basis for receptor affinity. *J. Mol. Biol.* **1998**, *277* (5), 1111−1128.
(18) Walsh, S. T.; Jevitts, L. M.; Sylvester, J. E.; Kossiakoff, A. A. Site2 binding energetics of the regulatory step of growth hormone-induced receptor homodimerization. *Protein Sci.* **2003**, *12* (9), 1960−1970.
(19) Wilmes, S.; Hafer, M.; Vuorio, J.; Tucker, J. A.; Winkelmann, H.; Löchte, S.; Stanly, T. A.; Pulgar Prieto, K. D.; Poojari, C.; Sharma, V.; Richter, C. P.; Kurre, R.; Hubbard, S. R.; Garcia, K. C.; Moraga, I.; Vattulainen, I.; Hitchcock, I. S.; Piehler, J. Mechanism of homodimeric cytokine receptor activation and dysregulation by oncogenic mutations. *Science* **2020**, *367* (6478), 643−652.
(20) Desikan, H.; Kaur, A.; Pogozheva, I. D.; Raghavan, M. Effects of calreticulin mutations on cell transformation and immunity. *J. Cell. Mol. Med.* **2023**, *27*, 1032.
(21) Papadopoulos, N.; Nédélec, A.; Derenne, A.; Şulea, T. A.; Pecquet, C.; Chachoua, I.; Vertenoeil, G.; Tilmant, T.; Petrescu, A. J.; Mazzucchelli, G.; Iorga, B. I.; Vertommen, D.; Constantinescu, S. N. Oncogenic CALR mutant C-terminus mediates dual binding to the thrombopoietin receptor triggering complex dimerization and activation. *Nat. Commun.* **2023**, *14* (1), 1881.
(22) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583−589.
(23) Akdel, M.; Pires, D. E. V.; Pardo, E. P.; Jänes, J.; Zalevsky, A. O.; Mészáros, B.; Bryant, P.; Good, L. L.; Laskowski, R. A.; Pozzati, G.; Shenoy, A.; Zhu, W.; Kundrotas, P.; Serra, V. R.; Rodrigues, C. H. M.; Dunham, A. S.; Burke, D.; Borkakoti, N.; Velankar, S.; Frost, A.; Basquin, J.; Lindorff-Larsen, K.; Bateman, A.; Kajava, A. V.; Valencia, A.; Ovchinnikov, S.; Durairaj, J.; Ascher, D. B.; Thornton, J. M.; Davey, N. E.; Stein, A.; Elofsson, A.; Croll, T. I.; Beltrao, P. A structural biology community assessment of AlphaFold2 applications. *Nat. Struct. Mol. Biol.* **2022**, *29* (11), 1056−1067.
(24) Kryshtafovych, A.; Schwede, T.; Topf, M.; Fidelis, K.; Moult, J. Critical assessment of methods of protein structure prediction (CASP)-Round XIV. *Proteins* **2021**, *89* (12), 1607−1617.
(25) Bryant, P.; Pozzati, G.; Elofsson, A. Improved prediction of protein-protein interactions using AlphaFold2. *Nat. Commun.* **2022**, *13* (1), 1265.
(26) Yin, R.; Feng, B. Y.; Varshney, A.; Pierce, B. G. Benchmarking AlphaFold for protein complex modeling reveals accuracy determinants. *Protein Sci.* **2022**, *31* (8), No. e4379.
(27) Gao, M.; Nakajima An, D.; Parks, J. M.; Skolnick, J. AF2Complex predicts direct physical interactions in multimeric proteins with deep learning. *Nat. Commun.* **2022**, *13* (1), 1744.
(28) Hegedüs, T.; Geisler, M.; Lukács, G. L.; Farkas, B. Ins and outs of AlphaFold2 transmembrane protein structure predictions. *Cell. Mol. Life Sci.* **2022**, *79* (1), 73.
(29) Pei, J.; Zhang, J.; Cong, Q. Human mitochondrial protein complexes revealed by large-scale coevolution analysis and deep learning-based structure modeling. *Bioinformatics* **2022**, *38* (18), 4301−4311.
(30) Tunyasuvunakool, K.; Adler, J.; Wu, Z.; Green, T.; Zielinski, M.; Žídek, A.; Bridgland, A.; Cowie, A.; Meyer, C.; Laydon, A.; Velankar, S.; Kleywegt, G. J.; Bateman, A.; Evans, R.; Pritzel, A.; Figurnov, M.; Ronneberger, O.; Bates, R.; Kohl, S. A. A.; Potapenko, A.; Ballard, A. J.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Clancy, E.; Reiman, D.; Petersen, S.; Senior, A. W.; Kavukcuoglu, K.; Birney, E.; Kohli, P.; Jumper, J.; Hassabis, D. Highly accurate protein structure prediction for the human proteome. *Nature* **2021**, *596* (7873), 590−596.
(31) Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; Žídek, A.; Green, T.; Tunyasuvunakool, K.; Petersen, S.; Jumper, J.; Clancy, E.; Green, R.; Vora, A.; Lutfi, M.; Figurnov, M.; Cowie, A.; Hobbs, N.; Kohli, P.; Kleywegt, G.; Birney, E.; Hassabis, D.; Velankar, S. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **2022**, *50* (D1), D439−D444.
(32) Evans, R.; O'Neill, M.; Pritzel, A.; Antropova, N.; Senior, A.; Green, T.; Žídek, A.; Bates, R.; Blackwell, S.; Yim, J.; Ronneberger, O.; Bodenstein, S.; Zielinski, M.; Bridgland, A.; Potapenko, A.; Cowie, A.; Tunyasuvunakool, K.; Jain, R.; Clancy, E.; Kohli, P.; Jumper, J.; Hassabis, D. Protein complex prediction with AlphaFold-Multimer. *bioRxiv [Preprint]* **2021**, DOI: 10.1101/2021.10.04.463034.
(33) Mirdita, M.; Schütze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. ColabFold: making protein folding accessible to all. *Nat. Methods* **2022**, *19* (6), 679−682.
(34) Lomize, A. L.; Todd, S. C.; Pogozheva, I. D. Spatial arrangement of proteins in planar and curved membranes by PPM 3.0. *Protein Sci.* **2022**, *31* (1), 209−220.
(35) Elkins, P. A.; Christinger, H. W.; Sandowski, Y.; Sakal, E.; Gertler, A.; de Vos, A. M.; Kossiakoff, A. A. Ternary complex between placental lactogen and the extracellular domain of the prolactin receptor. *Nat. Struct. Biol.* **2000**, *7* (9), 808−815.
(36) Basu, S.; Wallner, B. DockQ: A Quality Measure for Protein-Protein Docking Models. *PLoS One* **2016**, *11* (8), No. e0161879.
(37) Li, Q.; Wong, Y. L.; Yueqi Lee, M.; Li, Y.; Kang, C. Solution structure of the transmembrane domain of the mouse erythropoietin receptor in detergent micelles. *Sci. Rep.* **2015**, *5*, 13586.
(38) Constantinescu, S. N.; Huang, L. J.-s.; Nam, H.-s.; Lodish, H. F. The erythropoietin receptor cytosolic juxtamembrane domain contains an essential, precisely oriented, hydrophobic motif. *Mol. Cell* **2001**, *7* (2), 377−385.
(39) Lee, R. C.; Walters, J. A.; Reyland, M. E.; Anderson, S. M. Constitutive activation of the prolactin receptor results in the induction of growth factor-independent proliferation and constitutive activation of signaling molecules. *J. Biol. Chem.* **1999**, *274* (15), 10024−10034.
(40) Tan, D.; Huang, K. T.; Ueda, E.; Walker, A. M. S2 deletion variants of human PRL receptors demonstrate that extracellular domain conformation can alter conformation of the intracellular signaling domain. *Biochemistry* **2008**, *47* (1), 479−489.
(41) Minh Hung, H.; Dieu Hang, T.; Nguyen, M. T. Structural investigation of human prolactin receptor transmembrane domain homodimerization in a membrane environment through multiscale simulations. *J. Phys. Chem. B* **2019**, *123* (23), 4858−4866.
(42) Liu, W.; Brooks, C. L. Functional impact of manipulation on the relative orientation of human prolactin receptor domains. *Biochemistry* **2011**, *50* (23), 5333−5344.
(43) Floss, D. M.; Scheller, J. Naturally occurring and synthetic constitutive-active cytokine receptors in disease and therapy. *Cytokine Growth Factor Rev.* **2019**, *47*, 1−20.
(44) Lu, X.; Gross, A. W.; Lodish, H. F. Active conformation of the erythropoietin receptor: random and cysteine-scanning mutagenesis of the extracellular juxtamembrane and transmembrane domains. *J. Biol. Chem.* **2006**, *281* (11), 7002−7011.
(45) Tong, W.; Sulahian, R.; Gross, A. W.; Hendon, N.; Lodish, H. F.; Huang, L. J. The membrane-proximal region of the thrombopoietin receptor confers its high surface expression by JAK2-dependent and -independent mechanisms. *J. Biol. Chem.* **2006**, *281* (50), 38930−38940.
(46) Bocharov, E. V.; Lesovoy, D. M.; Bocharova, O. V.; Urban, A. S.; Pavlov, K. V.; Volynsky, P. E.; Efremov, R. G.; Arseniev, A. S. Structural basis of the signal transduction via transmembrane domain of the human growth hormone receptor. *Biochim. Biophysi. Acta* **2018**, *1862* (6), 1410−1420.
(47) Brooks, A. J.; Dai, W.; O'Mara, M. L.; Abankwa, D.; Chhabra, Y.; Pelekanos, R. A.; Gardon, O.; Tunny, K. A.; Blucher, K. M.; Morton, C. J.; Parker, M. W.; Sierecki, E.; Gambin, Y.; Gomez, G. A.;

https://doi.org/10.1021/acs.jcim.3c00926
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Alexandrov, K.; Wilson, I. A.; Doxastakis, M.; Mark, A. E.; Waters, M. J. Mechanism of activation of protein kinase JAK2 by the growth hormone receptor. *Science* **2014**, *344* (6185), 1249783.

(48) Behncken, S. N.; Billestrup, N.; Brown, R.; Amstrup, J.; Conway-Campbell, B.; Waters, M. J. Growth hormone (GH)-independent dimerization of GH receptor by a leucine zipper results in constitutive activation. *J. Biol. Chem.* **2000**, *275* (22), 17000–17007.

(49) Seubert, N.; Royer, Y.; Staerk, J.; Kubatzky, K. F.; Moucadel, V.; Krishnakumar, S.; Smith, S. O.; Constantinescu, S. N. Active and inactive orientations of the transmembrane and cytosolic domains of the erythropoietin receptor dimer. *Mol. Cell* **2003**, *12* (5), 1239–1250.

(50) Ruan, W.; Becker, V.; Klingmuller, U.; Langosch, D. The interface between self-assembling erythropoietin receptor transmembrane segments corresponds to a membrane-spanning leucine zipper. *J. Biol. Chem.* **2004**, *279* (5), 3273–3279.

(51) Varghese, L. N.; Zhang, J.-G.; Young, S. N.; Willson, T. A.; Alexander, W. S.; Nicola, N. A.; Babon, J. J.; Murphy, J. M. Functional characterization of c-Mpl ectodomain mutations that underlie congenital amegakaryocytic thrombocytopenia. *Growth Factors* **2014**, *32* (1), 18–26.

(52) Chen, W. M.; Yu, B.; Zhang, Q.; Xu, P. Identification of the residues in the extracellular domain of thrombopoietin receptor involved in the binding of thrombopoietin and a nuclear distribution protein (human NUDC). *J. Biol. Chem.* **2010**, *285* (34), 26697–26709.

(53) Leroy, E.; Defour, J. P.; Sato, T.; Dass, S.; Gryshkova, V.; Shwe, M. M.; Staerk, J.; Constantinescu, S. N.; Smith, S. O. His499 regulates dimerization and prevents oncogenic activation by asparagine mutations of the human thrombopoietin receptor. *J. Biol. Chem.* **2016**, *291* (6), 2974–2987.

(54) Levy, G.; Carillo, S.; Papoular, B.; Cassinat, B.; Zini, J. M.; Leroy, E.; Varghese, L. N.; Chachoua, I.; Defour, J. P.; Smith, S. O.; Constantinescu, S. N. Mpl mutations in essential thrombocythemia uncover a common path of activation with eltrombopag dependent on W491. *Blood* **2020**, *135* (12), 948–953.

(55) Kim, M. J.; Park, S. H.; Opella, S. J.; Marsilje, T. H.; Michellys, P. Y.; Seidel, H. M.; Tian, S. S. NMR structural studies of interactions of a small, nonpeptidyl Tpo mimic with the thrombopoietin receptor extracellular juxtamembrane and transmembrane domains. *J. Biol. Chem.* **2007**, *282* (19), 14253–14261.

(56) Staerk, J.; Defour, J.-P.; Pecquet, C.; Leroy, E.; Antoine-Poirel, H.; Brett, I.; Itaya, M.; Smith, S. O.; Vainchenker, W.; Constantinescu, S. N. Orientation-specific signalling by thrombopoietin receptor dimers. *EMBO J.* **2011**, *30* (21), 4398–4413.

(57) Matthews, E. E.; Thévenin, D.; Rogers, J. M.; Gotow, L.; Lira, P. D.; Reiter, L. A.; Brissette, W. H.; Engelman, D. M. Thrombopoietin receptor activation: transmembrane helix dimerization, rotation, and allosteric modulation. *FASEB J.* **2011**, *25* (7), 2234–2244.

(58) Plo, I.; Bellanné-Chantelot, C.; Mosca, M.; Mazzi, S.; Marty, C.; Vainchenker, W. Genetic Alterations of the Thrombopoietin/MPL/JAK2 Axis Impacting Megakaryopoiesis. *Front. Endocrinol.* **2017**, *8*, 234.

(59) Ding, J.; Komatsu, H.; Iida, S.; Yano, H.; Kusumoto, S.; Inagaki, A.; Mori, F.; Ri, M.; Ito, A.; Wakita, A.; Ishida, T.; Nitta, M.; Ueda, R. The Asn505 mutation of the c-MPL gene, which causes familial essential thrombocythemia, induces autonomous homodimerization of the c-Mpl protein due to strong amino acid polarity. *Blood* **2009**, *114* (15), 3325–3328.

(60) Defour, J. P.; Chachoua, I.; Pecquet, C.; Constantinescu, S. N. Oncogenic activation of MPL/thrombopoietin receptor by 17 mutations at W515: implications for myeloproliferative neoplasms. *Leukemia* **2016**, *30* (5), 1214–1216.

(61) Zhang, H.; Coblentz, C.; Watanabe-Smith, K.; Means, S.; Means, J.; Maxson, J. E.; Tyner, J. W. Gain-of-function mutations in granulocyte colony-stimulating factor receptor (CSF3R) reveal distinct mechanisms of CSF3R activation. *J. Biol. Chem.* **2018**, *293* (19), 7387–7396.

(62) Forbes, L. V.; Gale, R. E.; Pizzey, A.; Pouwels, K.; Nathwani, A.; Linch, D. C. An activating mutation in the transmembrane domain of the granulocyte colony-stimulating factor receptor in patients with acute myeloid leukemia. *Oncogene* **2002**, *21* (39), 5981–5989.

(63) Lomize, A. L.; Pogozheva, I. D. TMDOCK: an energy-based method for modeling alpha-helical dimers in membranes. *J. Mol. Biol.* **2017**, *429* (3), 390–398.

(64) Poger, D.; Mark, A. E. Turning the growth hormone receptor on: evidence that hormone binding induces subunit rotation. *Proteins* **2010**, *78* (5), 1163–1174.

(65) Livnah, O.; Stura, E. A.; Middleton, S. A.; Johnson, D. L.; Jolliffe, L. K.; Wilson, I. A. Crystallographic evidence for preformed dimers of erythropoietin receptor before ligand activation. *Science* **1999**, *283* (5404), 987–990.

(66) Watowich, S. S.; Yoshimura, A.; Longmore, G. D.; Hilton, D. J.; Yoshimura, Y.; Lodish, H. F. Homodimerization and constitutive activation of the erythropoietin receptor. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89* (6), 2140–2144.

(67) Watowich, S. S.; Hilton, D. J.; Lodish, H. F. Activation and inhibition of erythropoietin receptor function: role of receptor dimerization. *Mol. Cell. Biol.* **1994**, *14* (6), 3535–3549.

(68) Ayaz, P.; Hammarén, H. M.; Raivola, J.; Sharon, D.; Hubbard, S. R.; Silvennoinen, O.; Shan, Y.; Shaw, D. E. Structural models of full-length JAK2 kinase. *bioRxiv [Preprint]* **2019**, DOI: 10.1101/727727.

(69) Glassman, C. R.; Tsutsumi, N.; Saxton, R. A.; Lupardus, P. J.; Jude, K. M.; Garcia, K. C. Structure of a Janus kinase cytokine receptor complex reveals the basis for dimeric activation. *Science* **2022**, *376* (6589), 163–169.

(70) Caveney, N. A.; Saxton, R. A.; Waghray, D.; Glassman, C. R.; Tsutsumi, N.; Hubbard, S. R.; Garcia, K. C. Structural basis of Janus kinase trans-activation. *Cell Rep.* **2023**, *42* (3), 112201.

(71) Ferrao, R.; Lupardus, P. J. The Janus Kinase (JAK) FERM and SH2 Domains: Bringing Specificity to JAK-Receptor Interactions. *Front. Endocrinol.* **2017**, *8*, 71.

(72) Lupardus, P. J.; Ultsch, M.; Wallweber, H.; Bir Kohli, P.; Johnson, A. R.; Eigenbrot, C. Structure of the pseudokinase-kinase domains from protein kinase TYK2 reveals a mechanism for Janus kinase (JAK) autoinhibition. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (22), 8025–8030.

(73) Ferrao, R. D.; Wallweber, H. J.; Lupardus, P. J. Receptor-mediated dimerization of JAK2 FERM domains is required for JAK2 activation. *eLife* **2018**, *7*, No. e38089.

(74) Araya-Secchi, R.; Bugge, K.; Seiffert, P.; Petry, A.; Haxholm, G. W.; Lindorff-Larsen, K.; Pedersen, S. F.; Arleth, L.; Kragelund, B. B. The prolactin receptor scaffolds Janus kinase 2 via co-structure formation with phosphoinositide-4,5-bisphosphate. *eLife* **2023**, *12*, No. e84645.

(75) Haxholm, G. W.; Nikolajsen, L. F.; Olsen, J. G.; Fredsted, J.; Larsen, F. H.; Goffin, V.; Pedersen, S. F.; Brooks, A. J.; Waters, M. J.; Kragelund, B. B. Intrinsically disordered cytoplasmic domains of two cytokine receptors mediate conserved interactions with membranes. *Biochem. J.* **2015**, *468* (3), 495–506.

(76) Endicott, J. A.; Noble, M. E. M.; Johnson, L. N. The Structural Basis for Control of Eukaryotic Protein Kinases. *Annu. Rev. Biochem.* **2012**, *81* (1), 587–613.

(77) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79* (2), 926–935.

(78) Durell, S. R.; Brooks, B. R.; Ben-Naim, A. Solvent-induced forces between two hydrophilic groups. *J. Phys. Chem.* **1994**, *98* (8), 2198–2202.

(79) Waters, M. J.; Brooks, A. J. JAK2 activation by growth hormone and other cytokines. *Biochem. J.* **2015**, *466* (1), 1–11.

(80) Kubatzky, K. F.; Ruan, W.; Gurezka, R.; Cohen, J.; Ketteler, R.; Watowich, S. S.; Neumann, D.; Langosch, D.; Klingmuller, U. Self assembly of the transmembrane domain promotes signal transduction

https://doi.org/10.1021/acs.jcim.3c00926
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

through the erythropoietin receptor. *Curr. Biol.* **2001**, *11* (2), 110−115.
(81) Kubatzky, K. F.; Liu, W.; Goldgraben, K.; Simmerling, C.; Smith, S. O.; Constantinescu, S. N. Structural requirements of the extracellular to transmembrane domain junction for erythropoietin receptor function. *J. Biol. Chem.* **2005**, *280* (15), 14844−14854.
(82) Constantinescu, S. N.; Keren, T.; Socolovsky, M.; Nam, H.; Henis, Y. I.; Lodish, H. F. Ligand-independent oligomerization of cell-surface erythropoietin receptor is mediated by the transmembrane domain. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (8), 4379−4384.
(83) Atanasova, M.; Whitty, A. Understanding cytokine and growth factor receptor activation mechanisms. *Crit. Rev. Biochem. Mol. Biol.* **2012**, *47* (6), 502−530.
(84) Brown, R. J.; Adams, J. J.; Pelekanos, R. A.; Wan, Y.; McKinstry, W. J.; Palethorpe, K.; Seeber, R. M.; Monks, T. A.; Eidne, K. A.; Parker, M. W.; Waters, M. J. Model for growth hormone receptor activation based on subunit rotation within a receptor dimer. *Nat. Struct. Mol. Biol.* **2005**, *12* (9), 814−821.
(85) Gadd, S. L.; Clevenger, C. V. Ligand-independent dimerization of the human prolactin receptor isoforms: functional implications. *Mol. Endocrinol.* **2006**, *20* (11), 2734−2746.
(86) Defour, J.-P.; Itaya, M.; Gryshkova, V.; Brett, I. C.; Pecquet, C.; Sato, T.; Smith, S. O.; Constantinescu, S. N. Tryptophan at the transmembrane-cytosolic junction modulates thrombopoietin receptor dimerization and activation. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (7), 2540−2545.
(87) Varghese, L. N.; Carreño-Tarragona, G.; Levy, G.; Gutiérrez-López de Ocáriz, X.; Rapado, I.; Martínez-López, J.; Ayala, R.; Constantinescu, S. N. MPL S505C enhances driver mutations at W515 in essential thrombocythemia. *Blood Cancer J.* **2021**, *11* (11), 188.
(88) Sabath, D. F.; Kaushansky, K.; Broudy, V. C. Deletion of the extracellular membrane-distal cytokine receptor homology module of Mpl results in constitutive cell growth and loss of thrombopoietin binding. *Blood* **1999**, *94* (1), 365−367.
(89) Pecquet, C.; Chachoua, I.; Roy, A.; Balligand, T.; Vertenoeil, G.; Leroy, E.; Albu, R. I.; Defour, J. P.; Nivarthi, H.; Hug, E.; Xu, E.; Ould-Amer, Y.; Mouton, C.; Colau, D.; Vertommen, D.; Shwe, M. M.; Marty, C.; Plo, I.; Vainchenker, W.; Kralovics, R.; Constantinescu, S. N. Calreticulin mutants as oncogenic rogue chaperones for TpoR and traffic-defective pathogenic TpoR mutants. *Blood* **2019**, *133* (25), 2669−2681.
(90) Constantinescu, S. N.; Vainchenker, W.; Levy, G.; Papadopoulos, N. Functional Consequences of Mutations in Myeloproliferative Neoplasms. *Hemasphere* **2021**, *5* (6), No. e578.
(91) Bernichtein, S.; Touraine, P.; Goffin, V. New concepts in prolactin biology. *J. Endocrinol.* **2010**, *206* (1), 1−11.
(92) Vainchenker, W.; Plo, I.; Marty, C.; Varghese, L. N.; Constantinescu, S. N. The role of the thrombopoietin receptor MPL in myeloproliferative neoplasms: recent findings and potential therapeutic applications. *Expert Rev. Hematol.* **2019**, *12* (6), 437−448.
(93) Strous, G. J.; Almeida, A. D. S.; Putters, J.; Schantl, J.; Sedek, M.; Slotman, J. A.; Nespital, T.; Hassink, G. C.; Mol, J. A. Growth hormone receptor regulation in cancer and chronic diseases. *Front. Endocrinol.* **2020**, *11*, 597573.
(94) Hammarén, H. M.; Virtanen, A. T.; Raivola, J.; Silvennoinen, O. The regulation of JAKs in cytokine signaling and its breakdown in disease. *Cytokine* **2019**, *118*, 48−63.
(95) Baxter, E. J.; Scott, L. M.; Campbell, P. J.; East, C.; Fourouclas, N.; Swanton, S.; Vassiliou, G. S.; Bench, A. J.; Boyd, E. M.; Curtin, N.; Scott, M. A.; Erber, W. N.; Green, A. R. Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. *Lancet* **2005**, *365* (9464), 1054−1061.
(96) Mariani, V.; Biasini, M.; Barbato, A.; Schwede, T. lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics* **2013**, *29* (21), 2722−2728.
(97) Zhang, Y.; Skolnick, J. Scoring function for automated assessment of protein structure template quality. *Proteins* **2004**, *57* (4), 702−710.
(98) Consortium, T. U. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* **2023**, *51* (D1), D523−D531.
(99) Roy, R. S.; Liu, J.; Giri, N.; Guo, Z.; Cheng, J. Combining pairwise structural similarity and deep learning interface contact prediction to estimate protein complex model accuracy in CASP15. *Proteins* **2023**, DOI: 10.1002/prot.26542.
(100) Del Alamo, D.; Sala, D.; McHaourab, H. S.; Meiler, J. Sampling alternative conformational states of transporters and receptors with AlphaFold2. *eLife* **2022**, *11*. DOI: 10.7554/eLife.75751.
(101) Brooks, B. R.; Brooks, C. L., 3rd; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **2009**, *30* (10), 1545−1614.
(102) Buckner, J.; Liu, X.; Chakravorty, A.; Wu, Y.; Cervantes, L.; Lai, T.; Brooks, C. L., 3rd. pyCHARMM: Embedding CHARMM Functionality in a Python Framework. *ChemRxiv [Preprint]* **2023**, DOI: 10.26434/chemrxiv-2023-j6wt1.
(103) Park, S.-J.; Kern, N.; Brown, T.; Lee, J.; Im, W. CHARMM-GUI PDB Manipulator: various PDB structural modifications for biomolecular modeling and simulation. *J. Mol. Biol.* **2023**, *435*, 167995.
(104) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* **2008**, *29* (11), 1859−1865.
(105) Feig, M.; Karanicolas, J.; Brooks, C. L., 3rd. MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology. *J. Mol. Graph. Model.* **2004**, *22* (5), 377−395.
(106) Lomize, A. L.; Schnitzer, K. A.; Todd, S. C.; Cherepanov, S.; Outeiral, C.; Deane, C. M.; Pogozheva, I. D. Membranome 3.0: Database of single-pass membrane proteins with AlphaFold models. *Protein Sci.* **2022**, *31* (5), No. e4318.
(107) Krissinel, E.; Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D. Biol. Crystallogr.* **2004**, *60* (Pt 12 Pt 1), 2256−2268.
(108) Zhang, C.; Shine, M.; Pyle, A. M.; Zhang, Y. US-align: universal structure alignments of proteins, nucleic acids, and macromolecular complexes. *Nat. Methods* **2022**, *19* (9), 1109−1115.
(109) Pogozheva, I. D.; Armstrong, G. A.; Kong, L.; Hartnagel, T. J.; Carpino, C. A.; Gee, S. E.; Picarello, D. M.; Rubin, A. S.; Lee, J.; Park, S.; Lomize, A. L.; Im, W. Comparative molecular dynamics simulation studies of realistic eukaryotic, prokaryotic, and archaeal membranes. *J. Chem. Inf. Model.* **2022**, *62* (4), 1036−1051.
(110) Jo, S.; Lim, J. B.; Klauda, J. B.; Im, W. CHARMM-GUI Membrane Builder for mixed bilayers and its application to yeast membranes. *Biophys. J.* **2009**, *97* (1), 50−58.
(111) Wu, E. L.; Cheng, X.; Jo, S.; Rui, H.; Song, K. C.; Davila-Contreras, E. M.; Qi, Y.; Lee, J.; Monje-Galvan, V.; Venable, R. M.; Klauda, J. B.; Im, W. CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. *J. Comput. Chem.* **2014**, *35* (27), 1997−2004.
(112) Jo, S.; Kim, T.; Im, W. Automated builder and database of protein/membrane complexes for molecular dynamics simulations. *PLoS One* **2007**, *2* (9), No. e880.
(113) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D., Jr. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **2017**, *14* (1), 71−73.
(114) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L. P.; Simmonett, A. C.; Harrigan,

M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput. Biol.* **2017**, *13* (7), No. e1005659.
(115) Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y.; Jo, S.; Pande, V. S.; Case, D. A.; Brooks, C. L., 3rd; MacKerell, A. D., Jr.; Klauda, J. B.; Im, W. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **2016**, *12* (1), 405−413.
(116) Hopkins, C. W.; Le Grand, S.; Walker, R. C.; Roitberg, A. E. Long-time-step molecular dynamics through hydrogen mass repartitioning. *J. Chem. Theory Comput.* **2015**, *11* (4), 1864−1874.
(117) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23* (3), 327−341.
(118) Steinbach, P. J.; Brooks, B. R. New spherical-cutoff methods for long-range forces in macromolecular simulation. *J. Comput. Chem.* **1994**, *15* (7), 667−683.
(119) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103* (19), 8577−8593.

https://doi.org/10.1021/acs.jcim.3c00926
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Supporting Information

Structural modeling of cytokine-receptor-JAK2 signaling complexes using AlphaFold Multimer

Irina D. Pogozheva[1], Stanislav Cherepanov[2], Sang-Jun Park[3], Malini Raghavan[4], Wonpil Im[3], Andrei L. Lomize[1#]

[1]Department of Medicinal Chemistry, College of Pharmacy, University of Michigan, Ann Arbor, MI 48109, United States;

[2] Biophysics Program, University of Michigan, Ann Arbor, MI 48109, United States;

[3] Departments of Biological Sciences and Chemistry, Lehigh University, Bethlehem, PA 18015, United States;

[4] Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, MI 48109, United States.

[#]Corresponding Author: Andrei L. Lomize; Email: almz@umich.edu

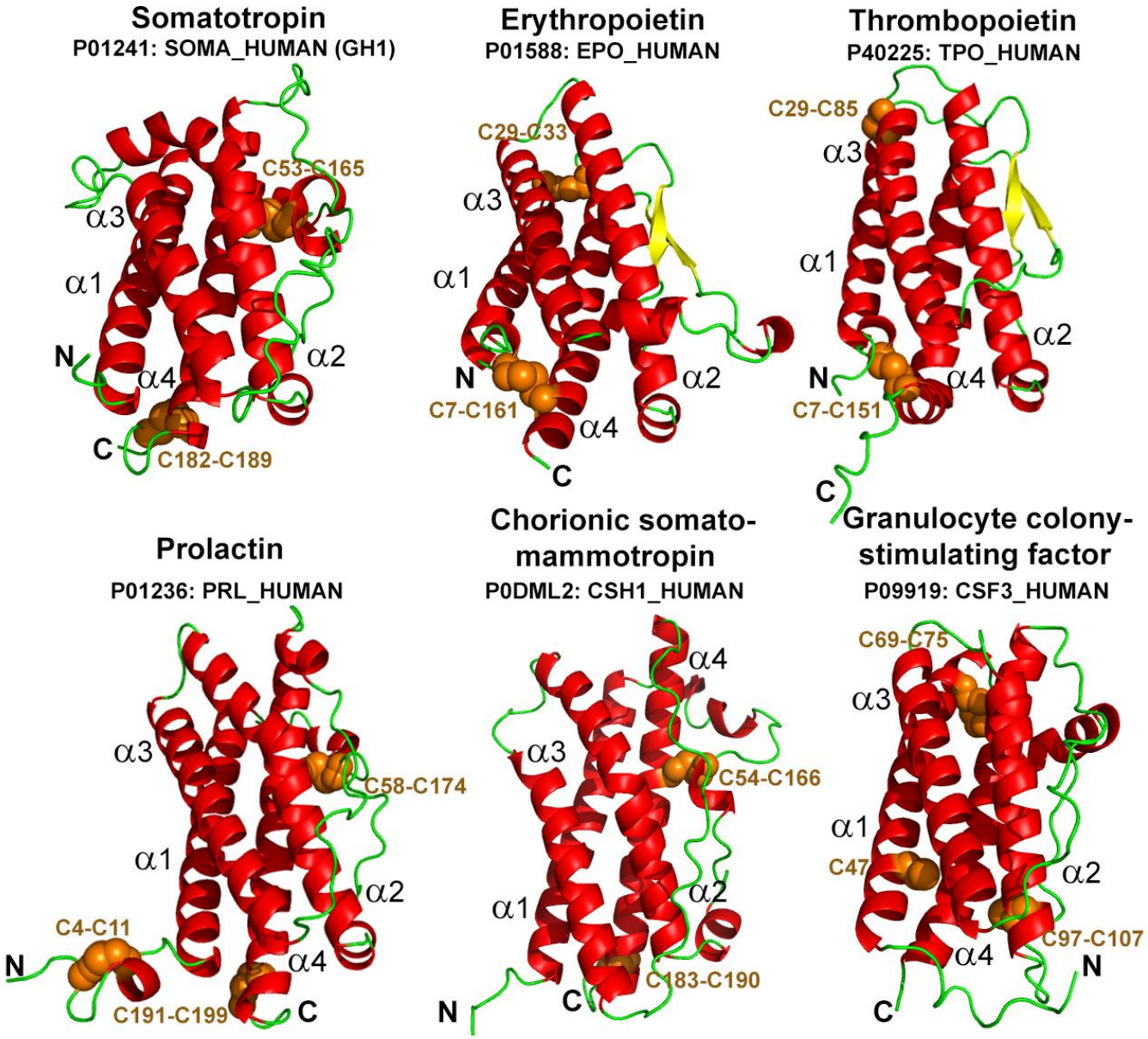


Figure S1. Structures of class 1 cytokines modeled by AlphaFold Multimer (AFM) in complex with their cognate receptors. The depicted cytokines are four-helical α-bundles with up-up-down-down topology and 2-3 disulfides that stabilize loop conformations. Models are shown as cartoons colored by secondary structure: red for α-helix, yellow for β-strand, green for unstructured loops. Cysteine residues and disulfides are shown by orange spheres. Residue numbers are for mature proteins (except CSF3).

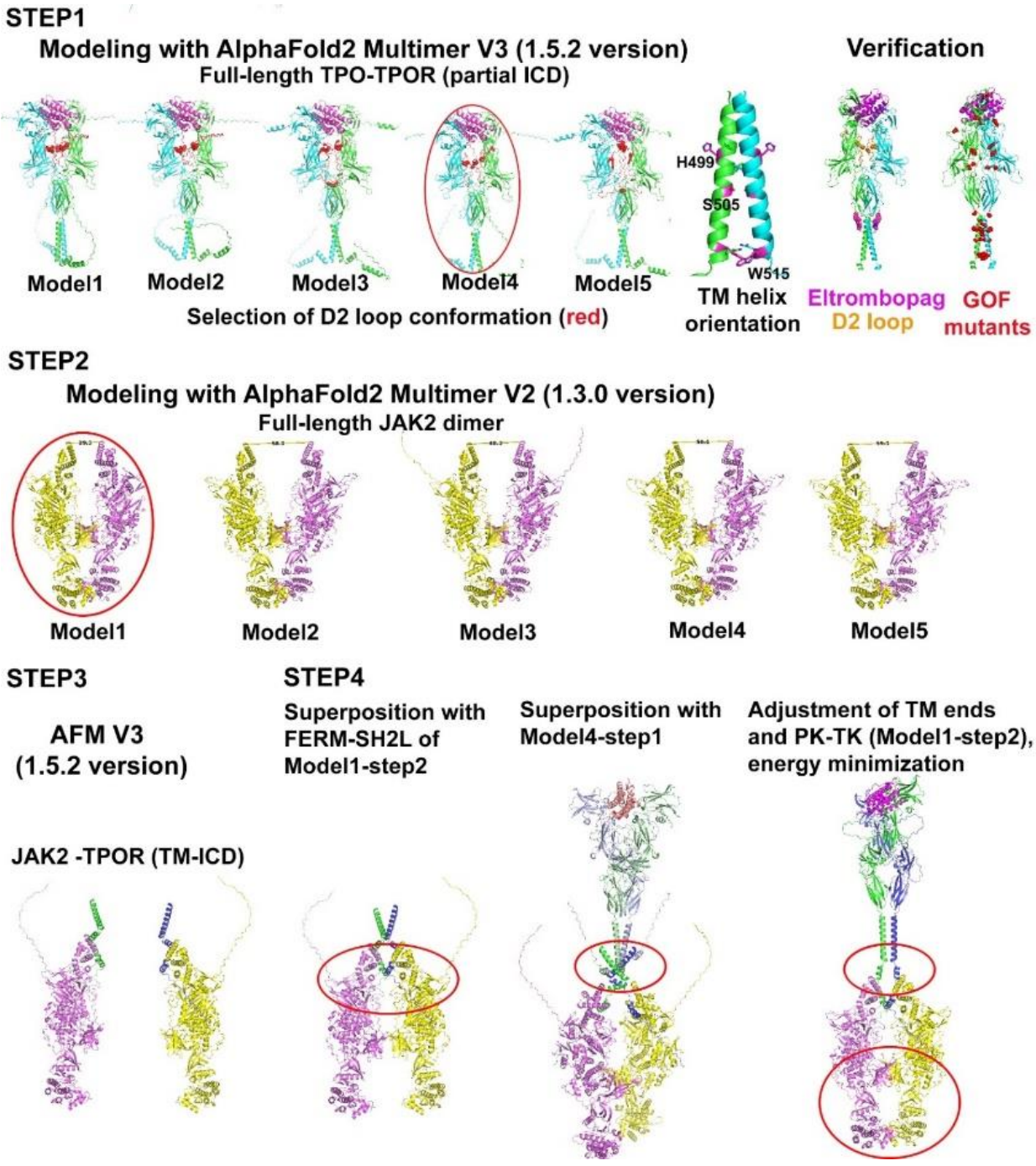


Figure S2. Four-step AMF-modeling of cytokine receptor signaling complexes (exemplified by hTPOR). Step1: Modeling with AFM V3 of TPO-TPOR active (1:2) complex. TPOR subunits are colored green and blue, TPO is colored purple. Five calculated models demonstrate similar conformations of ECDs except for the flexible D2 loop (residues 187-238 colored red) forming different intramolecular disulfides (shown as red spheres). The more frequent are the C193-C323, C194-C241, and C211-C322 disulfides. Four of five models demonstrate left-handed TM helical dimers with S505 at the dimerization interface (*a*-position of the heptad repeat motif) and H499 facing the lipid bilayer (*b*-position). Model 4 containing the most frequent disulfides and TMD helix arrangement was selected for the further calculations. Data supporting model 4 include: (1) the consistency of the TM α-helix arrangement with the *cc-TPOR-I* fusion construct

between the dimeric coiled-coil of *Put3* and TMD of mTPOR that displays constitutive activity [34]; (2) the docking of two eltrombopag molecules (shown by purple spheres) to the TMD; and (3) the localization of the majority of GOF mutations (shown by red spheres) at flexible structural elements (D2 loop and TMDs) participating in dimerization interfaces. Step2: modeling of JAK2 dimers. JAK2 subunits are colored pink and yellow. AFM V2 and V3 models produce similar structures of JAK2 dimers with the PK-PK dimerization interface but different distances between FERM domains (Cα-Cα distances between two L244 varies between 30 and 60 Å). The model 1 with the minimal distance between FERM domains was selected for the further calculations. Step3. Modeling of the JAK2 monomer (colored pink or yellow) with TPOR TM α-helix and the ICD domain (colored blue and green). AFM V2 and V3 produced rather similar models. Step4. Structural superposition of two monomer models (from step 3) with FERM-SH2L domains of JAK2 dimer (model 1 from step2) followed by docking of the ligand-bound dimer (model 4 from step 1), adjustment of TM helix ends, substitution of PK and TK domains by those from JAK2 dimer (model 1 from step 2), and model refinement by energy minimization.

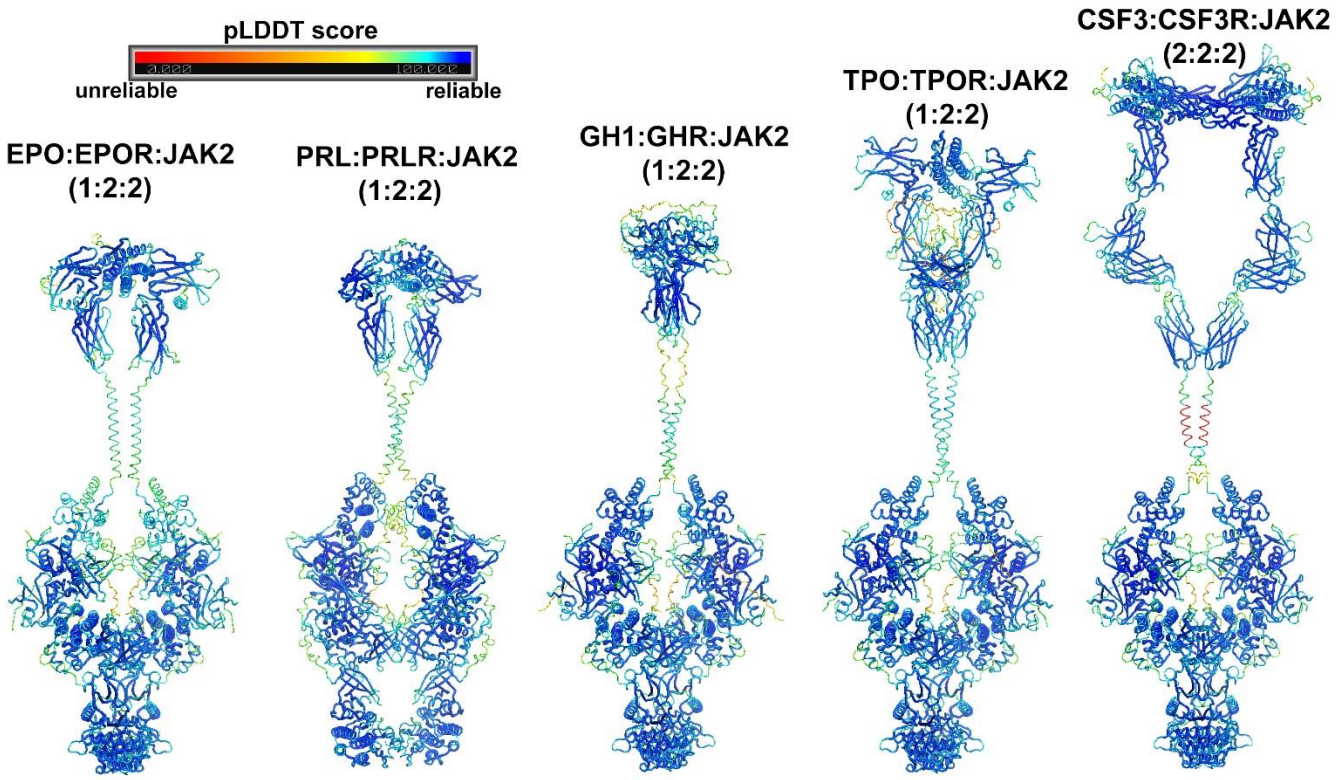


Figure S3. AFM-predicted models of five cytokine receptor signaling complexes colored by per residue pLDDT score.

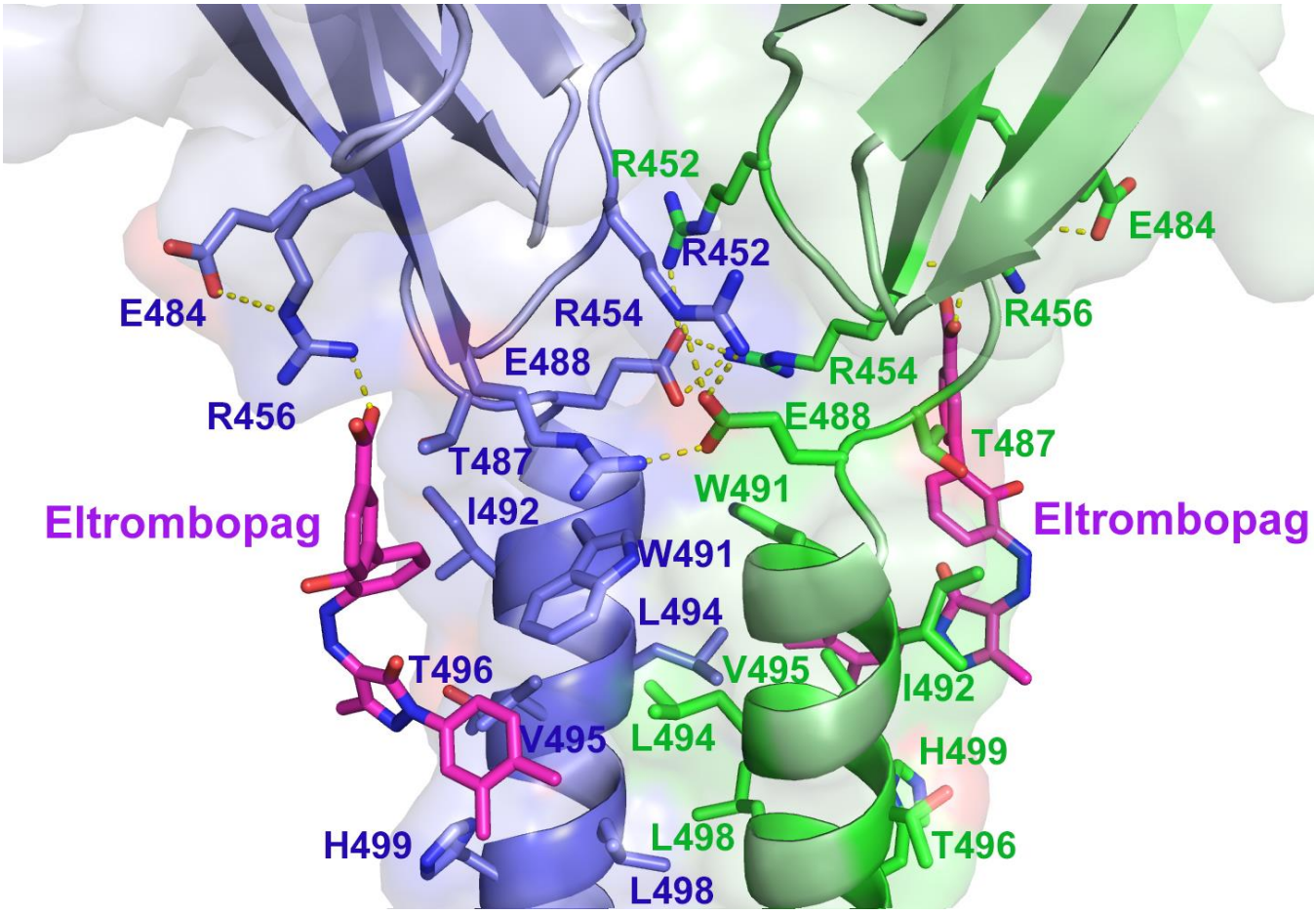


Figure S4. Suggested binding mode and interactions of the small molecule agonist eltrombopag with TM α-helices of the active human TPO-TPOR-JAK2 (1:2:2) signaling complex. Two eltrombopag molecules interact with the N-terminal part of the TM α-helical dimer (residues W491, I492, L494, V495, T496, L498, and H499) and R456 from the D4 domain. Neighboring residues (E488, R452, and R454) form a hydrogen bond network between D4 domains of TPOR (shown by yellow dashes). Molecules are shown by cartoon and semi-transparent surface representations colored blue and green for receptor subunits. Eltrombopag (colored magenta for Cα-atoms) and neighboring residues (colored blue and green for Cα-atoms) are shown as sticks.

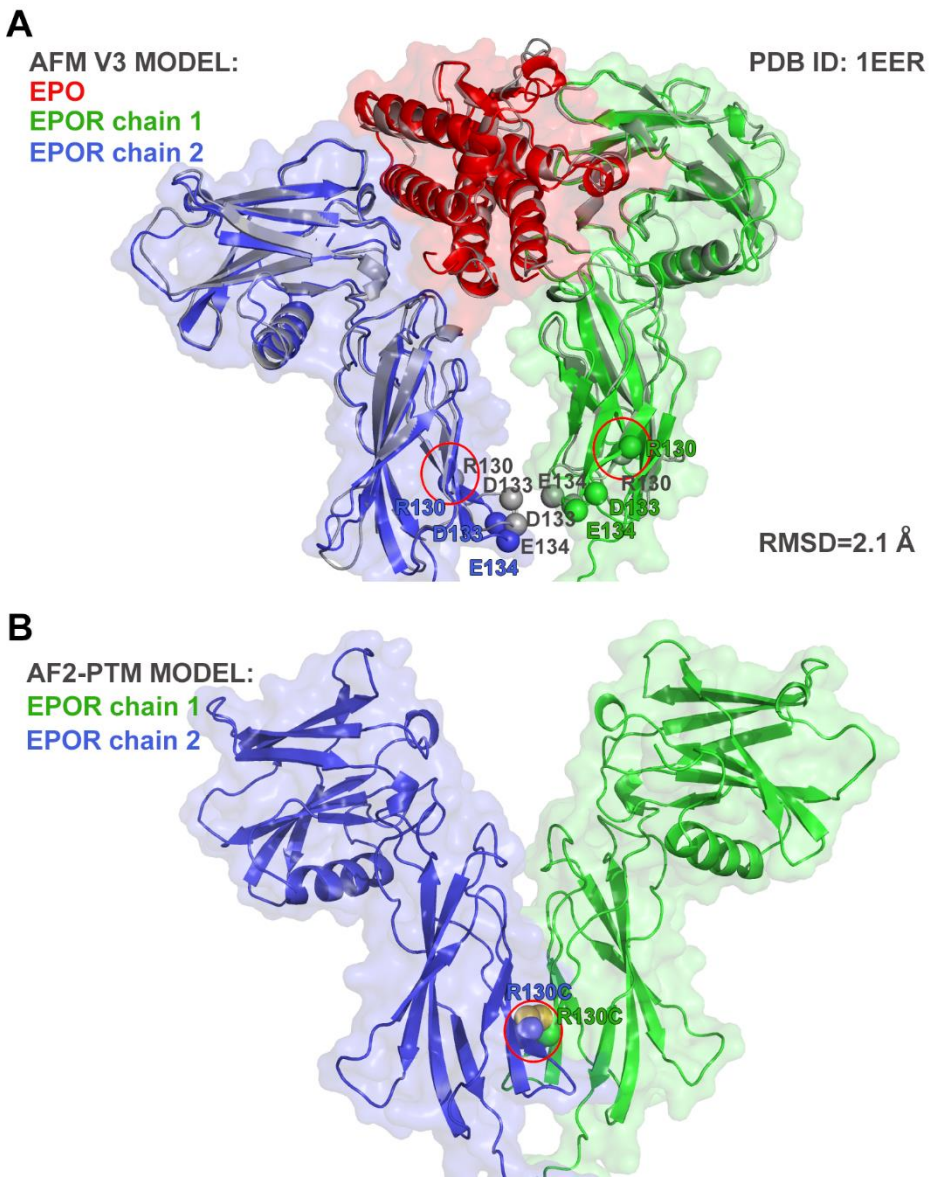


Figure S5. Spatial arrangement of R130, D133, and E134 residues of hEPOR corresponding to residues involved in formation of constitutively active disulfide-linked dimers of mEPOR upon substitution by cysteine. (**A**) Mapping of mutated residues in the crystal structure (PDB ID: **1EER**) and the AFM model of the ligand-bound hEPOR dimer. The ligand-bound AFM model and crystal structure are very similar (RMSD of 2.1 A) and provide proximity of D133 and E134, but not of R130 residues from both chains. (**B**) Mapping of C130-C130 intermolecular disulfide in the AFM2-ptm model of the ligand-free R130C hEPOR mutant. C130-C130 disulfide bond produces significant structural changes in the model of the ligand-free R130C homodimer (RMSD with 1EER is 9.7 Å). Proteins are shown by cartoon and semitransparent surface representations colored by chain (green and blue) for the AFM and AF2-ptm models and by gray for the X-ray structure; EPO ligand in the AFM model is colored red. Cα-atoms of residues forming intermolecular disulfides, R130, D133, and E134, are shown by spheres. The red circles enclose R130 in (**A)** and C130-C130 disulfide in (**B)**.

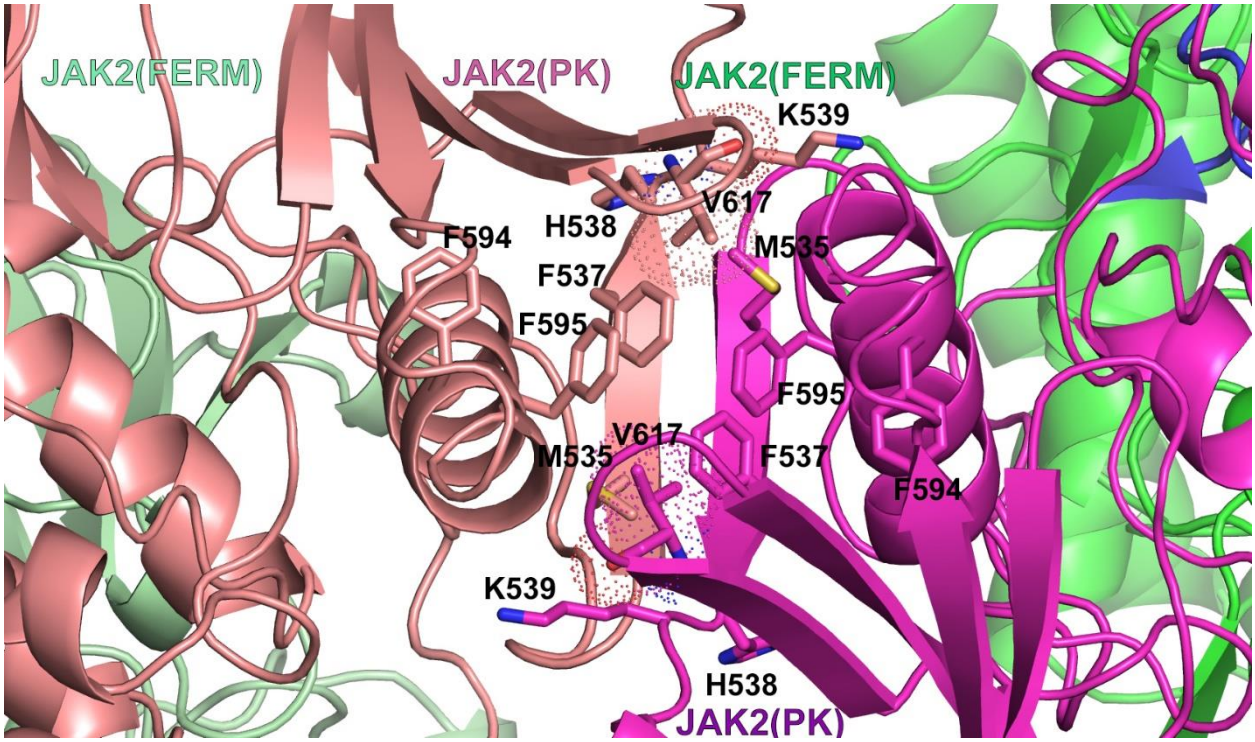


Figure S6. Close-up of the PK dimerization interface in the active human JAK2 dimer generated by AFM. The interface is formed by two antiparallel β-strands from PK N-lobes (cartoon representation colored pink and purple) and contains the hydrophobic cluster with V617 surrounded by aromatic residues, F537 and F595, as well as by PK residues involved in oncogenic mutations, M535, H538, K539 (shown by sticks). The oncogenic mutations (V617F, M535L, H538L, and K539L) stabilize the PK dimer by enhancing hydrophobic contacts and shape complementarity at the dimerization interface.

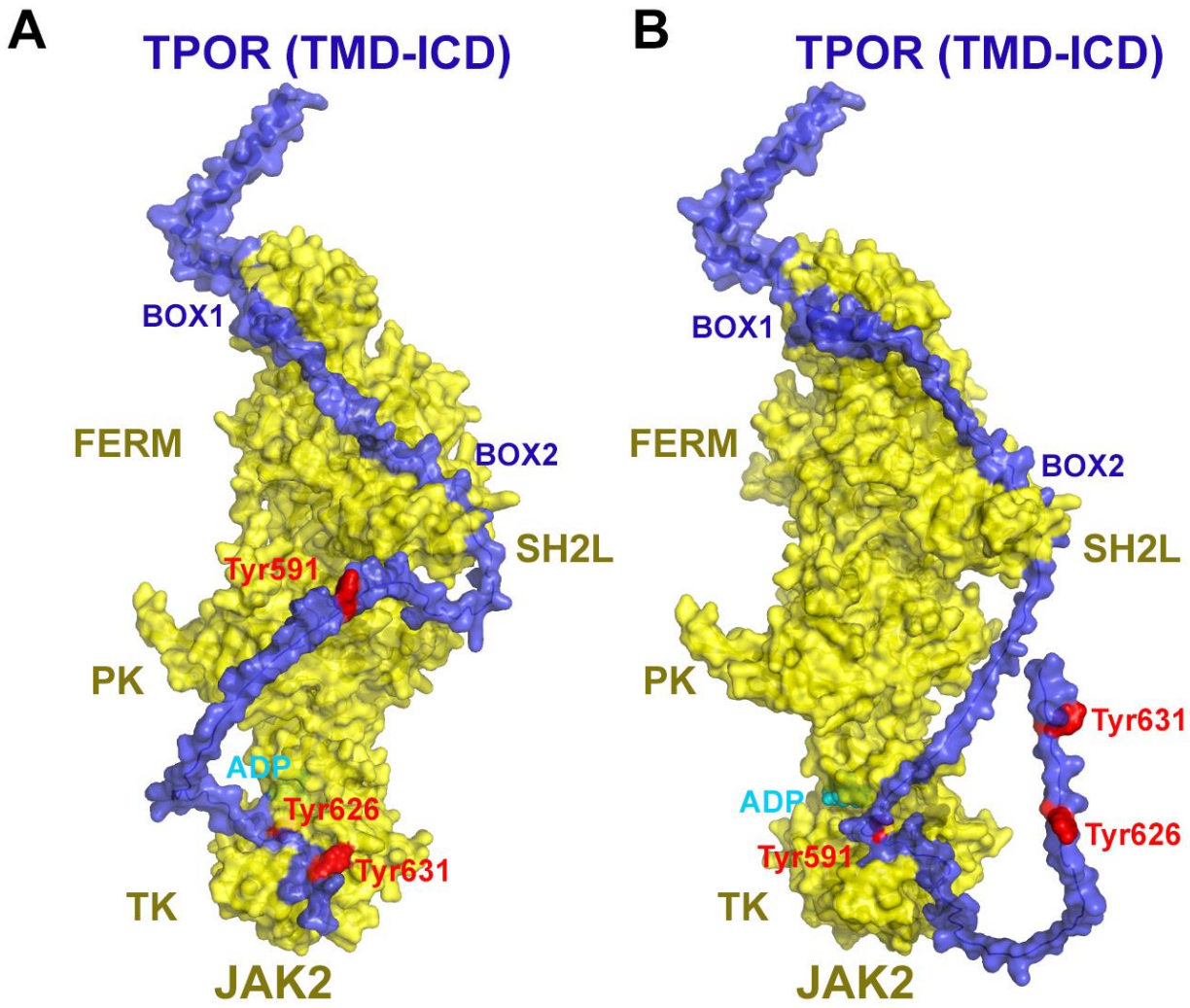


Figure S7. AFM modeling of interactions between TK domain of JAK2 and C-terminal tyrosine residues of TPOR that undergo phosphorylation by JAK2; Y626 (**A**) and Y591 (**B**). Due to the high flexibility of the unstructured ICD, different tyrosine residues can bind to the ligand binding pocket of the TK domain of JAK2. Protein molecules are shown using semi-transparent surfaces and cartoon representations and are colored yellow for JAK2, blue for TPOR. The ADP ligand is shown in cyan and TPOR tyrosine residues (Y591, Y626, and Y631) are colored red.

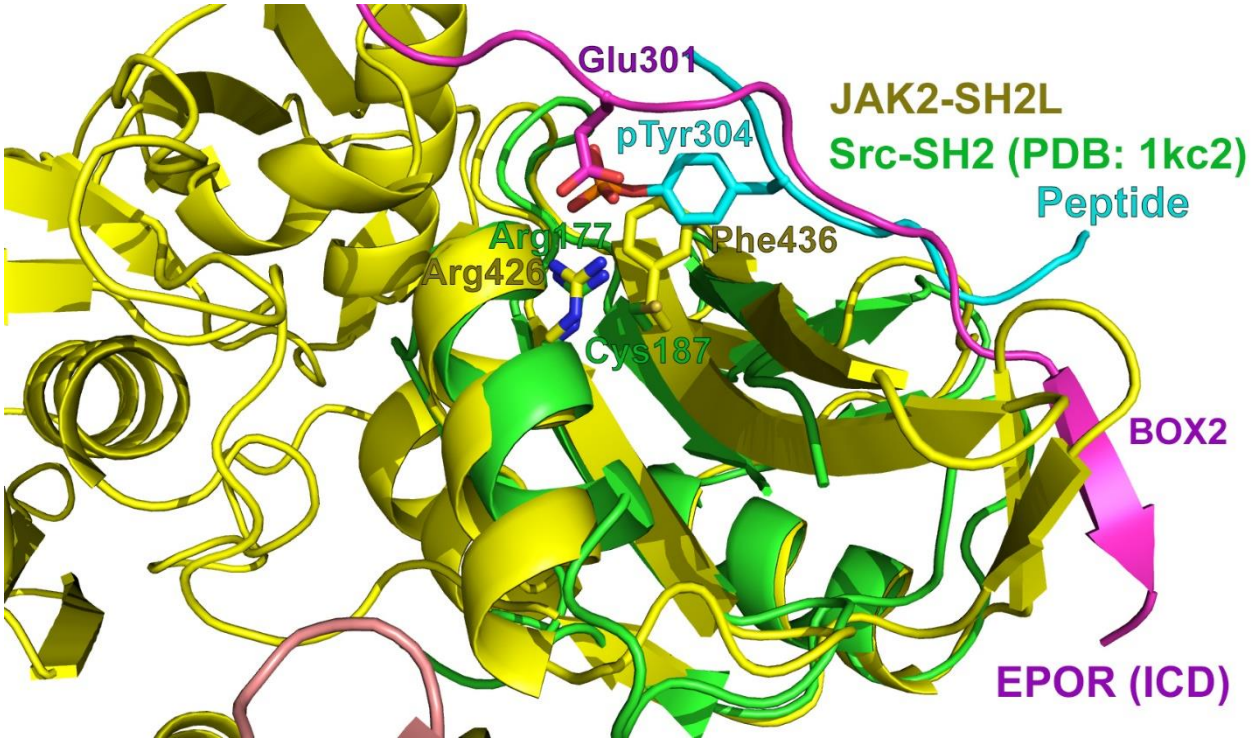


Figure S8. Close up of the SH2-like (SH2L) domain of human JAK2 (colored yellow) with a bound fragment of the human EPOR ICD (colored purple) from the AFM model superposed with the crystal structure of the Src-SH2 domain (colored green) with bound peptide (colored cyan) containing a phosphorylated tyrosine, pTyr304 (PDB ID: **1KC2**). SH2L domain of JAK2 has an aberrant binding site for phosphorylated tyrosine. This aberrant site carries Arg426 for binding negatively charged groups (i.e. phosphates), but lacks space for the tyrosine aromatic ring because the conserved cysteine is substituted by a bulky Phe436 residue. Therefore, the SH2L domain can specifically bind the negatively charged Glu301 from hEPOR Box 2 motif by forming ionic interactions with Arg426. Similar interactions were observed in the crystal structure of the JAK2 FERM-SH2L domain with the EPOR ICD peptide (PDB ID: **6E2Q**). It was suggested that these interactions contribute to the specificity of receptor binding [1, 2].

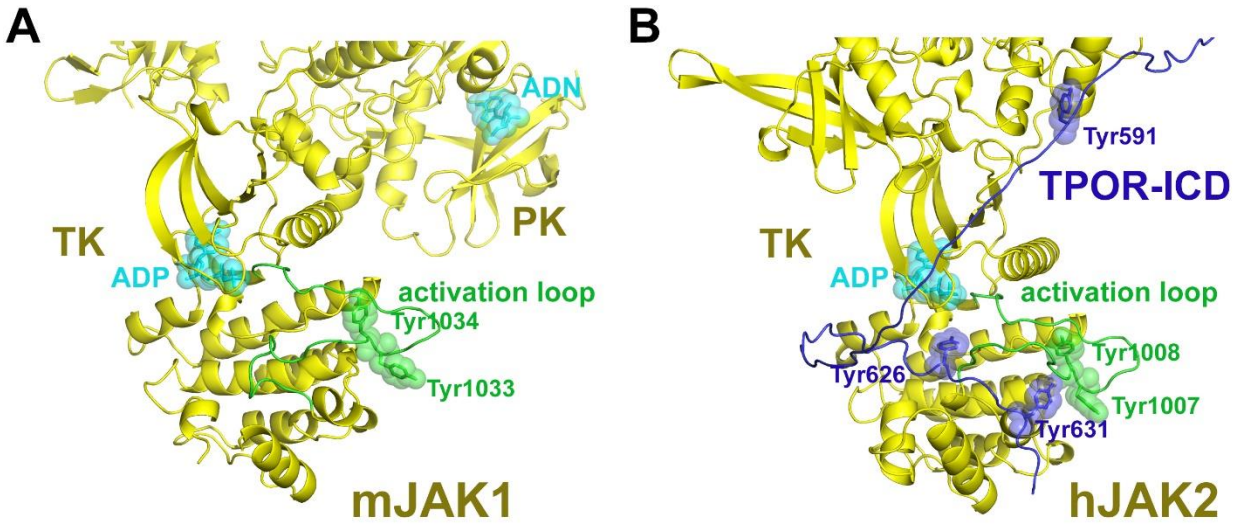


Figure S9. Activation loops in the TK domains of JAKs. (**A**) TK and partial PK domains from the cryo-EM-based model of mJAK1 (PDB ID: **7T6F**). (**B**) The TK domain from the AFM-generated model of hJAK2 with the fragment of TPOR ICD interacting with PK. The flexible activation loop may partially occlude the ligand binding pocket of the TK domain. Protein molecules are shown in cartoon representations colored blue for TPOR ICD, yellow for JAK2 with the activation loop of the TK domain colored green; ligands (ADP and ADN) are shown as sticks and semi-transparent cyan spheres. Tyrosine residues of TPOR, including Y626, occupying the ligand binding pocket of TK, are shown as blue sticks and semi-transparent spheres (**B**).

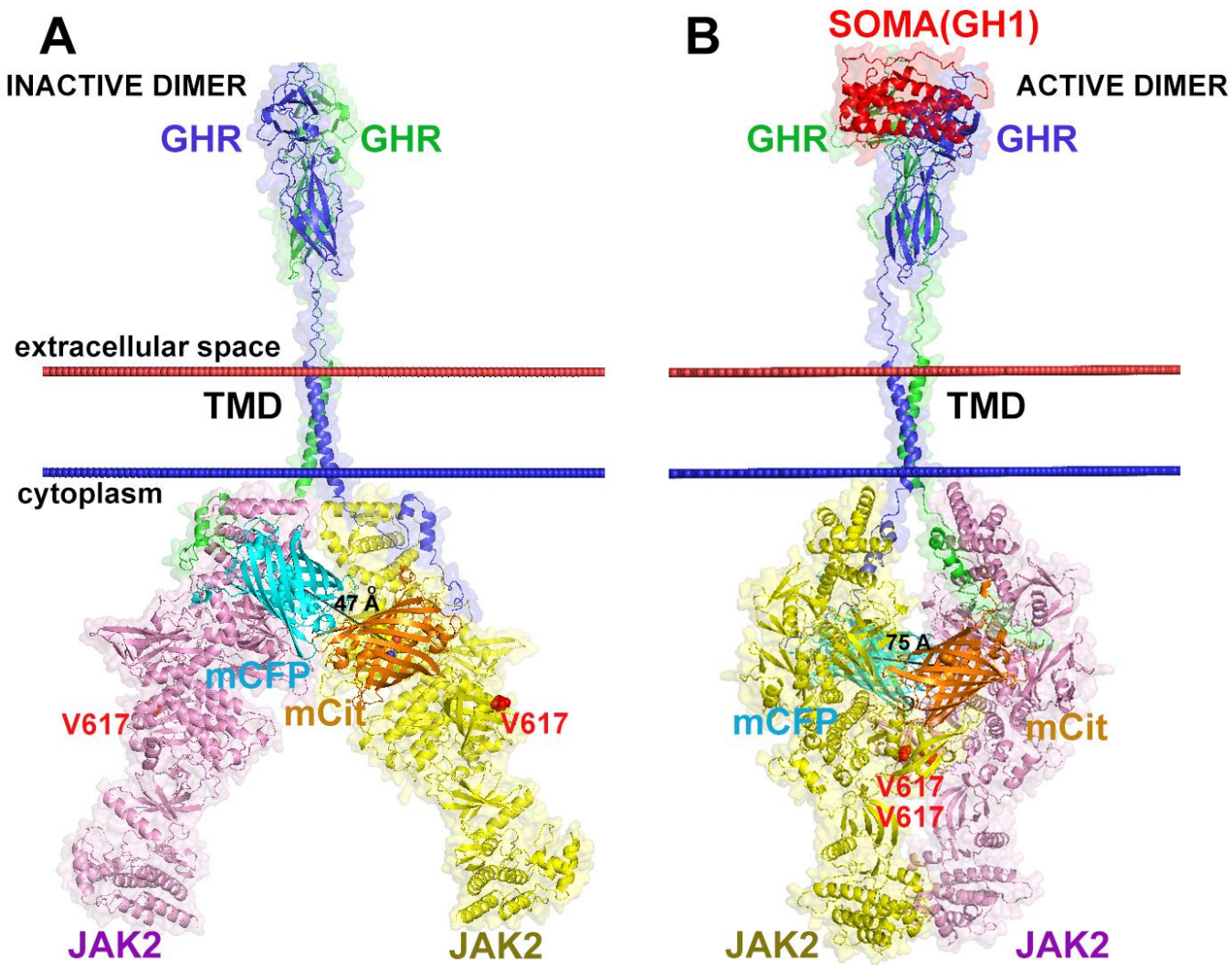


Figure S10. Suggested ligand-induced activation process of the GHR-JAK2 complex. AFM-generated models of ligand-free GHR-JAK2 dimer (**A**), and GHR-JAK2 dimer in the presence of GH1 ligand (**B**). Comparison with the results of FRET between GHR molecules labeled by FRET reporters (mCit and mCFP) positioned at C-terminus, 37 residues below the Box1 motif [3]. In the inactive receptor state, FRET reporters are located at the same side of the JAK2 dimer (distance between chromophores is 47 Å). However, in the active state, FRET reporters are located at the opposite sides of the JAK2 dimer (distance between chromophores is 75 Å). Protein molecules are shown as semi-transparent surfaces and cartoon representations, colored red for GH1, blue and green for GHR subunits, yellow and pink for JAK2 subunits. FRET reporters mCit and mCFP (shown as orange and cyan β-barrels, respectively) were modeled using the available CFP structure (PDB ID: **3ZFT**). V617 residue located at the dimerization interface of the PK domain is shown as red spheres.

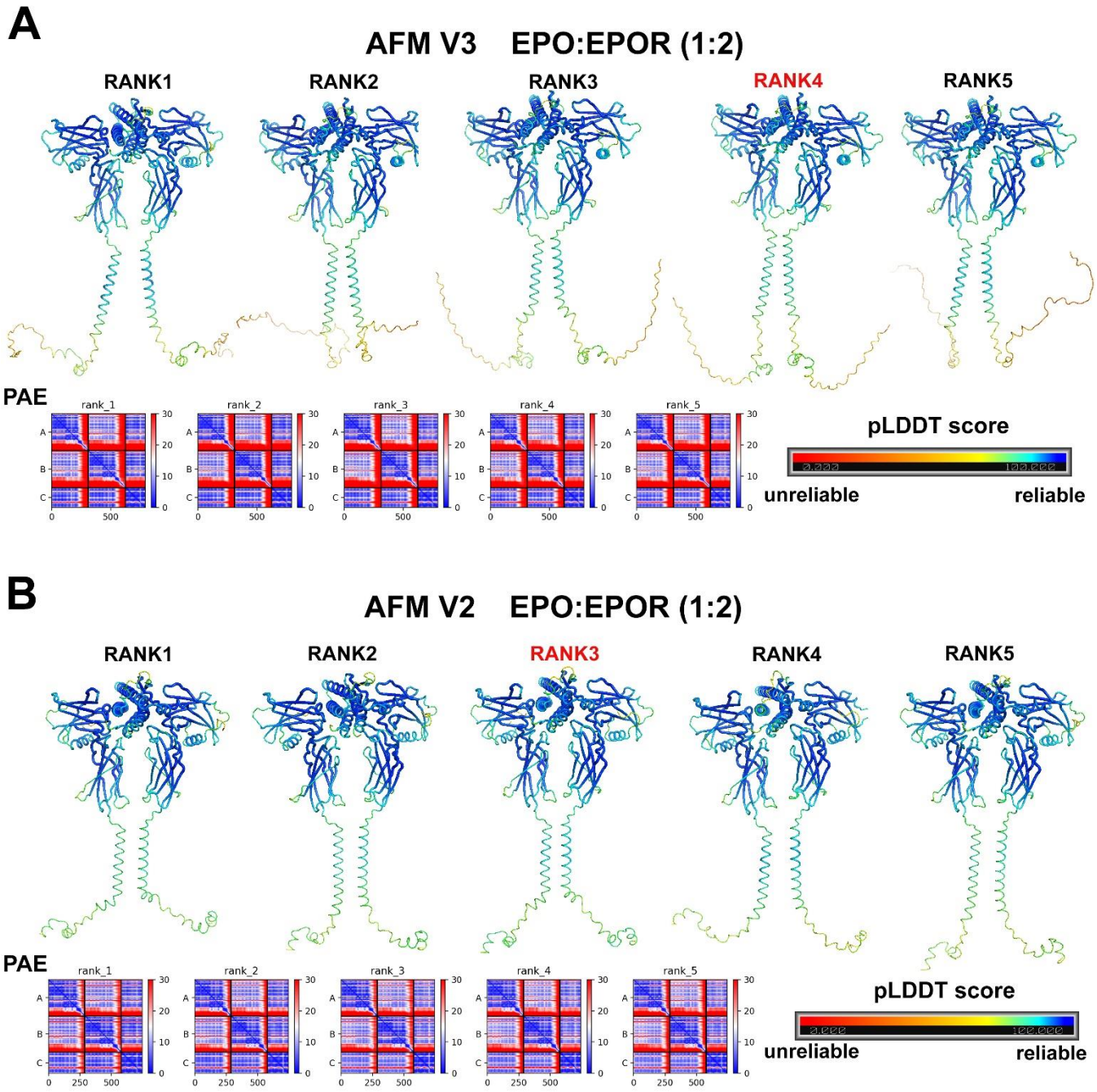


Figure S11. Quality metrics of AFM predicted models. (**A**) Five models predicted for EPO-EPOR (1:2) complex by AFM V3 (A) and AFM V2 (**B**). Models are colored by per residue pLDDT scores. Selected model ranks are marked by red.

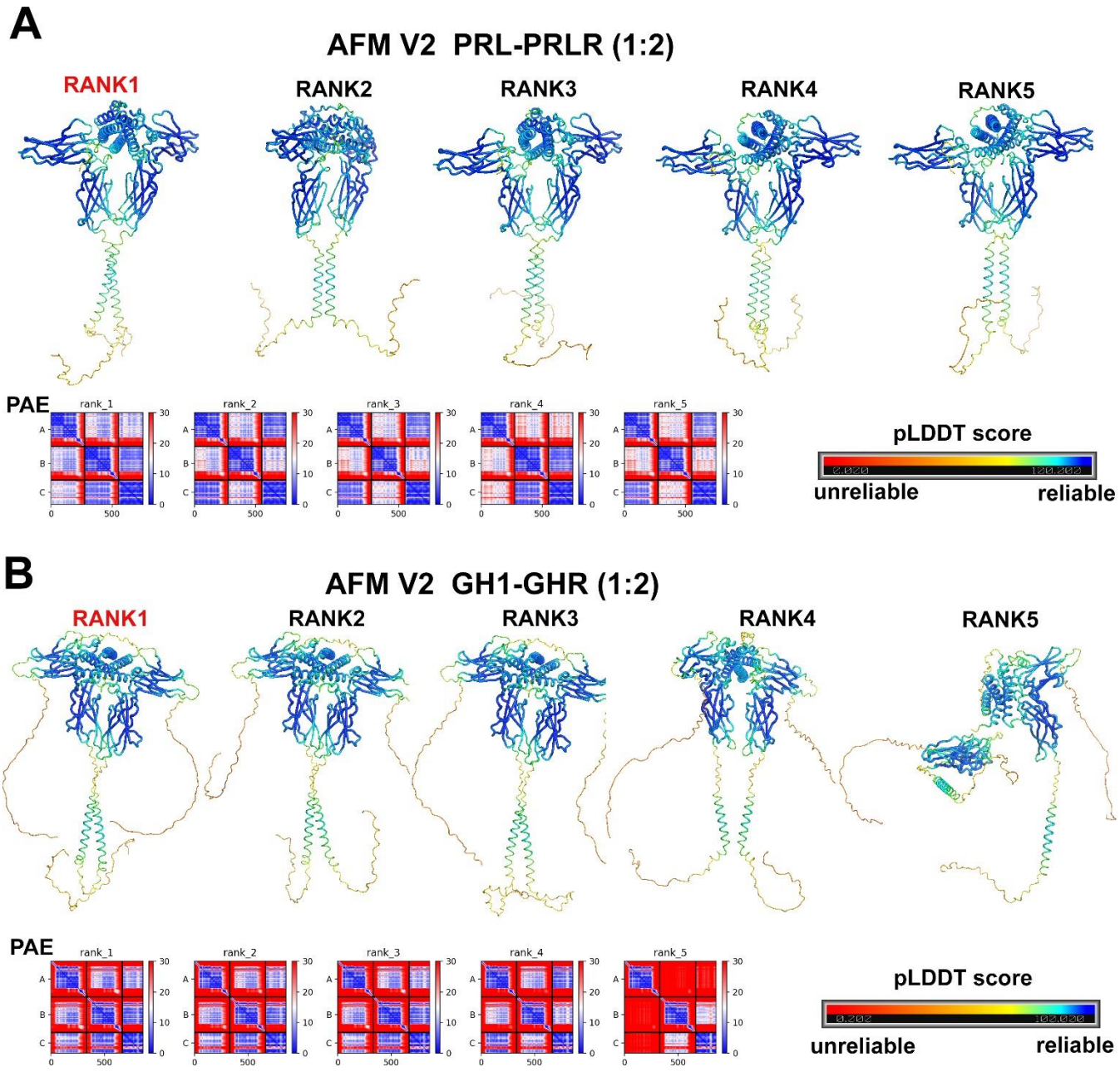


Figure S12. Quality metrics of AFM V2 predicted models. (**A**) Five models generated for PRL-PRLR (1:2) complex. (**B**) Five models generated for GH1-GHR (1:2) complex. Models are colored by per residue pLDDT scores. Selected model ranks are marked by red.

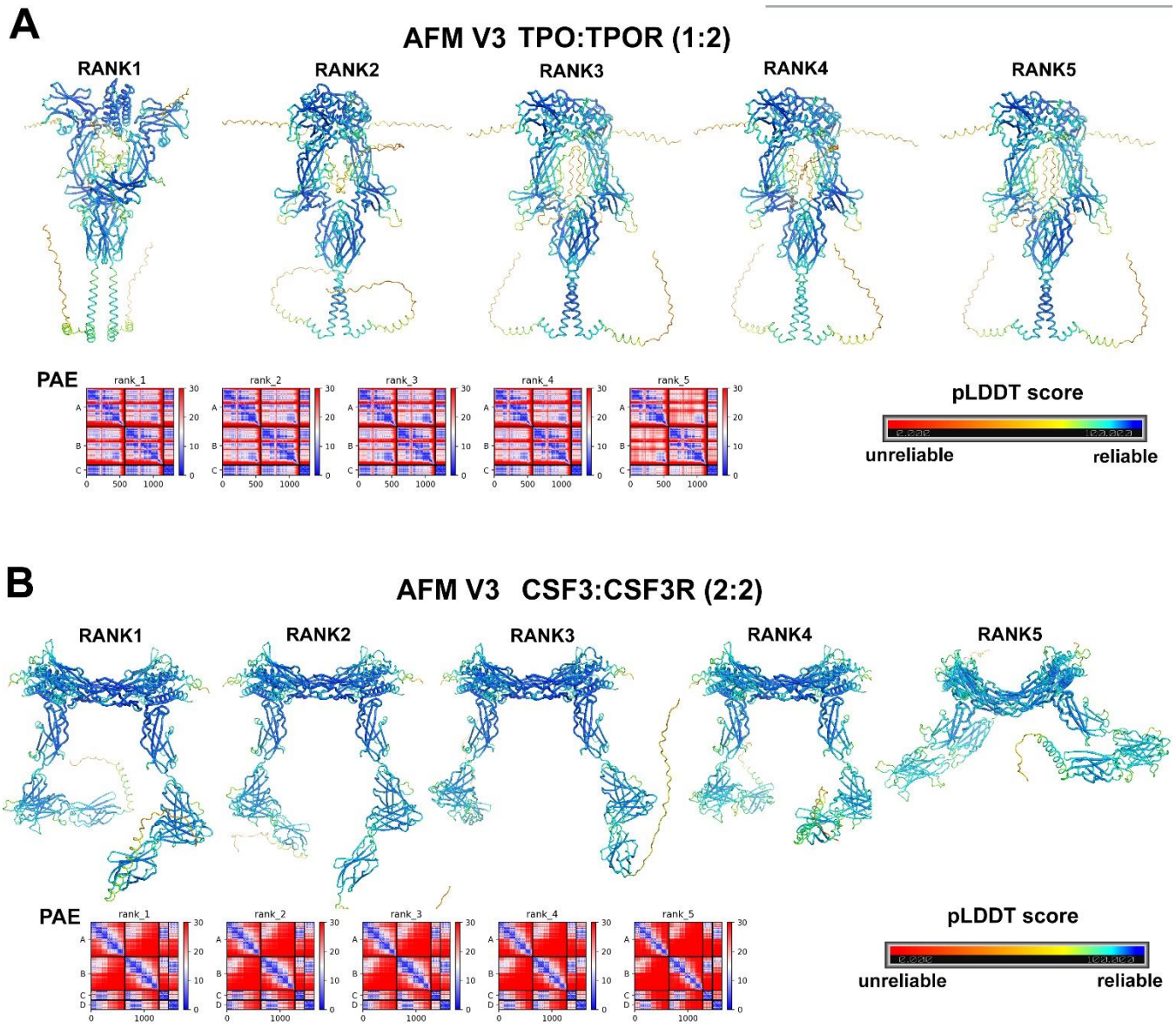


Figure S13. Quality metrics of AFM predicted models. (**A**) Five models predicted for TPO-TPOR (1:2) complex by AFM V3. (**B**) Five models predicted for CSF3-CSF3R (2:2) complex by AFM V3. Models are colored by per residue pLDDT scores.

Table S1. Structural features and interacting partners of class 1 homodimeric cytokine receptors from the JAK-STAT signaling pathway [4-6].

Receptor	Sequence length (signal peptide)	Disulfides	N-glycosylation sites	WSXWS motifs	Natural ligands	JAKs	STATs	SOCSs
EPOR*	508 (24)	28-38, 67-83	52	[209]WSAWS	**EPO**	**JAK2** LYN	**STAT5A** STAT3, STAT1	SOCS3
PRLR*	622 (24)	12-22, 51-62	35, 80, 108	[191]WSAWS	**PRL,** CSH1, CSH2, GH1	**JAK2**	**STAT5A,** STAT5B STAT3, STAT1	SOCS2
GHR	638 (18)	56-66, 101-112, 126-140, 259-259†	115, 156, 200	[240]YGEFS	**GH1 (SOMA)**	**JAK2,** LYN	**STAT5B** STAT3, STAT1	SOCS2
TPOR	635 (25)	40-50, 77-93, 291-301, 334-352, 193-323#, 194-241#, 211-322#	117, 178, 298, 358	[269]WGSWS [474]WSSWS	**TPO,** CRTmut**	**JAK2,** TYK2	**STAT5A**, STAT3, STAT1	SOCS3
CSF3R	780 (24)	26-52, 46-101, 131-142, 167-218, 177-186, 248-295, 266-309, 388-395	51, 93, 128, 134, 389, 474, 571, 610	[318]WSDWS	**CSF3**	**JAK1,** JAK2, JAK3	**STAT3,** STAT5 STAT1	SOCS3

* Residue numbers are for mature proteins (lacking signal peptide).

† Intermolecular disulfide

Disulfides possibly formed between loops inside D2 and D3 domains

** CRTmut, calreticulin mutants related to MPNs

Bold characters indicate the main interacting protein.

Table S2. Parameters of AFM-generated models of homodimers of ligand-bound receptors, dimers of TM segments for constitutively active mutants, JAK2 dimers, and monomeric JAK2-receptor complexes. These were used as structural blocks for building complete models of active ligand-receptor-kinase signaling complexes.

Name	Residues	pLDDT	pTM	ipTM	TM helix packing	PDB code or model	Cα-RMSD, Å
Ligand-bound receptor homodimers							
hEPO*-hEPOR* (1:2)	1-166* (lig), 1-288*	81.7	0.737	0.680	**L+, L^{239} *e***	1EER	1.6 (560/592)
mEPO*-mEPOR (1:2)	1-166* (lig) 1-314	78.5	0.687	0.645	**L+, S^{238} e**	1EER	1.06 (462/592)
hGH1*-hGHR (1:2)	1-191* (lig) 1-332	70.3	0.648	0.639	**L+, F^{273} *d*-**	3HHR 5OEK	1.2 (549/573) 3.1 (48/48)
hPRL*-hPRLR* (1:2)	1-199* (lig) 1-276*	80.2	0.680	0.635	**R-, AxxxA^{222}xxxL**	3NPZ	1.8 (536/587)
hCSH1*-hPRLR* (1:2)	1-191* (lig) 1-276*	80.8	0.665	0.584	**R-, LxxxW^{230}xxxL**	1F6F	1.9 (403/593)
hGH1*-hPRLR* (1:2)	1-191* (lig) 1-276*	82.2	0.713	0.641	**R-, CxxxV^{229}xxxA**	1BP3	1.1 (322/383)
hTPO*-hTPOR (1:2)#	1-153* (lig) 1-635	75.1	0.67	0.639	**L+, S^{505} *a*, H^{499} *b***	None	
hCSF3-hCSF2R (1:1)	30-207 (lig) 1-667	83.1	0.664	0.788	N/A	2D9Q	0.9 (411/466)
hCSF3-hCSF2R (2:2)	30-207 (lig) 1-667	86.9	0.684	0.643	N/A	2D9Q	3.2 (567/933)
JAK2 homodimer							
hJAK2 dimer	1-1132	79.5	0.743	0.672	224-224 distance 29.3 Å	7T6F 8EWY	2.5 (1898/2164) 3.0 (2052/2172)
JAK2-receptor complexes (1:1)							
hJAK2-hEPOR* (D2-TM-ICD)	1-1132, 120-372*	82.9	0.769	0.631	L+	7T6F (FERM-SH2L-PK)	2.48 (618/762)
hJAK2-hGHR (D2-TM-ICD)	1-1132, 148-390	82.5	0.754	0.634	L+	7T6F (FERM-SH2L-PK)	2.47 (620/762)
hJAK2-hPRLR* (TM-ICD)	1-1132, 205-295*	85.6	0.838	0.727	R-	7T6F (FERM-SH2L-PK)	2.45 (620/762)
hJAK2-hTPOR (TM-ICD)	1-1132, 488-635	82.8	0.808	0.552	L+	7T6F (FERM-SH2L-PK)	2.63 (636/762)
hJAK2-hCSF3R (TM-ICD)	1-1132, 606-708	84.6	0.826	0.612	L+	7T6F (FERM-SH2L-PK)	2.58 (621/762)

* Residue number are for mature proteins. # calculated by AFM V3. Other models were calculated by AFM V2.

The AFM parameters (pLDDT, pTM, ipTM) are provided for a single model selected for further modeling and analysis out of 5 models generated by AFM. Type of helix arrangement in dimers (L+ or R-) as defined by the sign of the crossing angle: L+, left-handed (coiled coil) dimer with a positive crossing angle and heptad repeat, or R-, right-handed dimer with a negative crossing angle and the tetrad (i.e. GxxxG) repeat motif. Letters for left-handed dimers indicate the position of a reference residue in the $(abcdefg)_n$ heptad repeat motif (*a* and *d* positions are located at the helix-helix interface). RMSD column includes the number of overlapped residues in the structural superposition divided by total number of residues in the structure (in parentheses). RMSD values were calculated by the align method of PyMOL.

Bold characters indicate a TM helix packing consistent with the final structure of the active receptor dimer.

Table S3. Fraction of natural contacts (F_{nat}), LRMS, iRMS, and DockQ values for extracellular domains of cytokine-receptor complexes.

Protein	PDB ID	Subunits	F_{nat}	LRMS	IRMS	DockQ
Site 1						
EPOR	1EER	A-C	0.96	2.6	1.3	0.81
GHR	3HHR	A-B	0.84	2.9	1.1	0.80
PRLR	3NPZ	A-B	0.79	2.9	0.8	0.82
CSF3R	2D9Q	A-B	0.95	2.1	0.9	0.88
CSF3R	2D9Q	C-D	0.95	2.1	0.9	0.88
Site 2						
EPOR	1EER	A-B	0.79	3.5	1.2	0.75
GHR	3HHR	A-C	0.81	3.5	1.4	0.73
PRLR	3NPZ	A-C	0.59	6.6	2.0	0.52
CSF3R	2D9Q	A-D	0.58	2.1	1.8	0.64
CSF3R	2D9Q	B-C	0.58	2.1	1.8	0.64

DocQ, a protein-protein docking model quality, is derived by combining $F_{nat,}$ LRMS, and iRMS; medium quality models on the CAPRI-set have ($0.51 \leq DockQ < 0.81$), high quality model have DocQ>0.81 [13]. AFM models of cytokine-receptor complexes shows high quality for site1 of ligand and medium quality for site 2 of ligand.

Table S4. Parameters of AFM-generated models of ligand-free receptor dimers and dimers of TM α-helical segments.

Name	Residues	pLDDT	pTM	ipTM	TM helix packing	PDB entry or another model	Cα-RMSD, Å*
Ligand-free receptor homodimers							
hEPOR* #	1-269*	76.8	0.456	0.154	L+, L^{239} *b*	1EER	8.94 (412/592)
hEPOR(R130C)* ##	1-264*	81.6	0.690	0.592	Parallel, V^{237}	1EER	9.74 (416/592)
hEPOR(R130C)* ##	1-295*	78.5	0.642	0.551	Parallel, V^{237}	1EER	9.65 (425/592)
hEPOR(D133C)* ##	1-295*	78.9	0.469	0.238	Parallel, V^{237}	1EER	5.24 (370/592)
hEPOR(E134C)* ##	1-295*	74.9	0.417	0.176	Parallel, V^{237}	1EER	5.26 (411/592)
mEPOR*	1-190*	75.7	0.496	0.258	L+, S^{238} *a*	1EER	5.6 (408/592)
hGHR	1-310	70.8	0.503	0.352	L+, F^{273} *e*	3HHR	7.5 (384/573)
hPRLR*	2-279*	82.4	0.474	0.163	L+, A^{222} *a*	3NPZ	8.6 (391/587)
hTPOR#	26-550 1-552	72.4	0.526	0.444	R- H^{499} out	Final model	4.8 (952/1104)
hCSF3R (D5-D6-TM segment)	421-660	79.2	0.389	0.120	R-, T^{640} out	Final model	8.1 (434/480)
Homodimers of TM α-helices*							
hEPOR*	209-288*	54.6	0.393	0.337	L+, L^{239} *a*	Final model TMDOCK	2.9 (57/58) 1.7 (53/58)
mEPOR*	208-287*	52.8	0.364	0.306	L+, S^{238} *a*	Final model TMDOCK	3.1 (56/58) 2.0 (56/58)
hGHR	240-325	52.0	0.394	0.356	L+, F^{273} *a*	5OEK Final model	3.6 (53/54) 1.6 (51/54)
hPRLR*	191-249*	62.7	0.516	0.467	**R-, AxxxA^{222}xxxL**	Final model TMDOCK	2.7 (68/68) 2.2 (48/50)
hPRLR* CAM (Δ10-186)	1-9+187-276	49.7	0.384	0.360	**R-, AxxxA^{222}xxxL**	Final model TMDOCK	1.9(67/68) 3.3(49/50)
hPRLR* CAM (Δ1-210)	211-276	62.0	0.492	0.461	**R-, AxxxA^{222}xxxL**	Final model TMDOCK	1.7 (60/68) 3.6 (50/50)
hTPOR CAM H499L/W515K	488-549	71.9	0.463	0.373	L+, S^{505} *d*, H^{499} *e*	Final model TMDOCK	1.2 (52/54) 1.4 (48/58)
hTPOR CAM (L498W/H499Y)	488-549	73.5	0.49	0.404	L+, S^{505} *d*, H^{499} *e*	Final model TMDOCK	1.2 (52/54) 1.4 (48/58)
mTPOR	456-533	63.6	0.47	0.397	L+, S^{498} *d*, L^{492} *e*	Final model TMDOCK	2.8 (54/54) 1.9 (56/58)
hCSF3R	606-669	60.1	0.483	0.442	**L+, T^{640} *a***	Final model TMDOCK	0.8 (49/54) 0.8 (41/50)
hCSF3R CAM (T640N)	606-669	58.8	0.439	0.389	**L+, T^{640} *a***	Final model TMDOCK	0.7 (48/54) 0.9 (44/50)

See legend for Table S2. Bold characters indicate TM helix arrangement similar to that in the models of final ligand-receptor-kinase signaling complexes. CAM, constitutively active mutants.
calculated by AFM V3.
calculated by AF2-ptm
Other models were calculated by AFM V2.

Table S5. Interactions of hTPO with hTPOR in the AFM-generated model of the ligand-bound hTPOR dimer.*

Site 1		Site2	
TPO*	**TPOR**	**TPO***	**TPOR**
L16	I263	**R10**	**E261**
D45	R102	V11	**F104**
F46	**L103, F104**	**K14**	E160
L48	**F45, L103, L265**	**R17**	D163
K52	E46	**R98**	E99
H133	F126	L99	**L103, F104**
K136	D128	L101	F105
R140	**D261**		
F141	**F104**, F164, **L265**		
L144	F164		

* Residue numbers are for mature protein (TPO).

Bold characters indicate hTPOR residues (F45, L103, F104, D261, and L265) that are involved in hTPO binding, in accordance with mutagenesis studies [7, 8] and hTPO residues (R10, K14, R17, K52, R98, H133, K136, F141, L144) that are essential for binding to hTPOR [9, 10].

Table S6. TM helix arrangement in final models of complete ligand-receptor-kinase complexes, in models of ligand-bound and ligand-free receptor dimers, and in dimers of isolated TM segments calculated by AFM [11] and TMDOCK [12].

Name	Complete complex	Ligand-bound receptor dimer	Ligand-free receptor dimer	TM helix dimer (AFM)	TM helix dimer (TMDOCK)
hEPOR*	L+, L^{239} *e*	L+, L^{239} *e*	R-	L+, L^{239} *a*	no association
mEPOR*	L+, S^{238} *e*	L+, S^{238} *e*	L+, S^{238} *a*	L+, S^{238} *a*	L+, S^{238} *a*
hGHR	L+, F^{273} *d*	L+, F^{273} *d*	L+, F^{273} *e*	L+, F^{273} *a*	L+, F^{273} *c*
hPRLR*	R-, $AxxxA^{222}xxxL$	R-, $AxxxA^{222}xxxL$	L+ A^{222} *a*	R-, $AxxxA^{222}xxxL$ (WT,CAM**)	R-, $SxxxC^{225}xxxV$
hCSF3R	L+, T^{640} *a*	N/A	N/A	L+, T^{640} *a* (WT,CAM**)	L+, T^{640} *a*
hTPOR	L+, S^{505} *a*, H^{499} *b*	L+, S^{505} *a*, H^{499} *b*	R-	L+, S^{505} *d*, H^{499} *e* (CAM)**	L+, S^{505} *d*, H^{499} *e*

Cells marked by gray indicate helix arrangements similar to that in the final ligand-receptor-kinase complexes.

Type of helix arrangement in dimers (L+ or R-) is defined by the sign of the crossing angle: L+, left handed dimer (coiled coil) with positive crossing angle and the $(abcdefg)_n$ heptad repeat motif, and R-, right-handed dimer with a negative crossing angle and the tetrad (i.e. GxxxG) repeat motif. Letters for left-handed dimers (bold character) indicate the position of a reference residue in the heptad repeat motif, where *a* and *d* positions form the dimerization interface.

* Residue numbers are for mature proteins.

** Calculated for TMDs of constitutively active mutants (CAM): hPRLR (Δ10-186, Δ1-210), hCSF3R (T640N), and TPOR (H499L/W515K and L498W/H499Y).

Table S7: Lipid composition of the mammalian plasma membrane (number of specified lipid molecules in each leaflet).

Lipid name	Lipid Head / Tail	GHR		CSF3R, EPOR, PRLR, TPOR	
		Outer	Inner	Outer	Inner
POPC	PC(16:0/18:1(9Z))	64	28	48	21
PLPC	PC(16:0/18:2(9Z,12Z))	88	44	66	33
PAPE	PE(16:0/20:4(5Z,8Z,11Z,14Z))	12	48	9	36
POPE	PE(16:0/18:1(9Z))	12	56	9	42
POPI	PI(16:0/18:1(9Z))	0	20	0	15
PAPS	PS(16:0/20:4(5Z,8Z,11Z,14Z))	0	44	0	33
POPA	PA(16:0/18:1(9Z))	0	4	0	3
SSM	SM(d18:1/18:0)	44	20	33	15
NSM	SM(d18:1/24:1)	44	20	33	15
CMH	GlcCer(d18:1/16:0)	16	0	12	0
CHOL	Cholesterol	148	116	111	87
Total		428	400	321	300

Lipids head groups: PC, phosphatidylcholine; PE, phosphatidylethanolamine; PI, phosphatidylinositol; PS, phosphatidylserine; SM, sphingomyelin; GlcCer, glucosylceramide.

REFERENCES

1. Ferrao, R., Lupardus, P.J. (2017). The Janus Kinase (JAK) FERM and SH2 Domains: Bringing Specificity to JAK-Receptor Interactions. *Front. Endocrinol.* **8**, 71.
2. Ferrao, R.D., Wallweber, H.J., Lupardus, P.J. (2018). Receptor-mediated dimerization of JAK2 FERM domains is required for JAK2 activation. *eLife* **7**, e38089.
3. Brooks, A.J., Dai, W., O'Mara, M.L., Abankwa, D., Chhabra, Y., Pelekanos, R.A., et al. (2014). Mechanism of activation of protein kinase JAK2 by the growth hormone receptor. *Science* **344**, 1249783.
4. Dehkhoda, F., Lee, C.M.M., Medina, J., Brooks, A.J. (2018). The Growth Hormone Receptor: Mechanism of Receptor Activation, Cell Signaling, and Physiological Aspects. *Front. Endocrinol.* **9**, 35.
5. Liongue, C., Sertori, R., Ward, A.C. (2016). Evolution of Cytokine Receptor Signaling. *J. Immunol.* **197**, 11-18.
6. Morris, R., Kershaw, N.J., Babon, J.J. (2018). The molecular details of cytokine signaling via the JAK/STAT pathway. *Protein Sci.* **27**, 1984-2009.
7. Varghese, L.N., Zhang, J.-G., Young, S.N., Willson, T.A., Alexander, W.S., Nicola, N.A., et al. (2014). Functional characterization of c-Mpl ectodomain mutations that underlie congenital amegakaryocytic thrombocytopenia. *Growth Factors* **32**, 18-26.
8. Chen, W.M., Yu, B., Zhang, Q., Xu, P. (2010). Identification of the residues in the extracellular domain of thrombopoietin receptor involved in the binding of thrombopoietin and a nuclear distribution protein (human NUDC). *J. Biol. Chem.* **285**, 26697-26709.
9. Pearce, K.H., Jr., Potts, B.J., Presta, L.G., Bald, L.N., Fendly, B.M., Wells, J.A. (1997). Mutational analysis of thrombopoietin for identification of receptor and neutralizing antibody sites. *J. Biol. Chem.* **272**, 20595-20602.
10. Hitchcock, I.S., Kaushansky, K. (2014). Thrombopoietin from beginning to end. *Br. J. Haematol.* **165**, 259-268.
11. Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., et al. (2021). Protein complex prediction with AlphaFold-Multimer. *bioRxiv* 2021.2010.2004.463034 [Preprint], doi: https://doi.org/10.1101/2021.10.04.463034
12. Lomize, A.L., Pogozheva, I.D. (2017). TMDOCK: an energy-based method for modeling alpha-helical dimers in membranes. *J. Mol. Biol.* **429**, 390-398.
13. Basu, S., Wallner, B. DockQ: A Quality Measure for Protein-Protein Docking Models. *PLoS One* **2016**, *11* (8), e0161879. DOI: 10.1371/journal.pone.0161879.