Reanalysis of cryo-EM data reveals ALK-cytokine assemblies with distinct stoichiometries

Jan Felix[1,2*], Steven De Munck[1,2], J. Fernando Bazan[3], Savvas N. Savvides[1,3*]

[1]VIB-UGent Center for Inflammation Research, 9052 Ghent, Belgium.

[2] Unit for Structural Biology, Department of Biochemistry and Microbiology, Ghent University, 9052 Ghent, Belgium.

[3] ℏ bioconsulting llc, Stillwater, MN, USA.

*Correspondence: savvas.savvides@ugent.be ; jan.felix@ugent.be

Abstract

Anaplastic lymphoma kinase (ALK) and leukocyte tyrosine kinase (LTK) are receptor tyrosine kinases activated by their cognate ligands ALKAL2 and ALKAL1 resulting in pleiotropic signalling outputs in development, metabolism and cancer. Structural studies by X-ray crystallography[1,2] and electron cryo-microscopy (cryo-EM)[3] recently revealed distinct assemblies corresponding to ALK-ALKAL2 and LTK-ALKAL1 assemblies with either 2:1 stoichiometry[1] versus ALK-ALKAL2 complexes with 2:2 stoichiometry[2,3]. Here, we report the reanalysis of cryo-EM data deposited by Reshetnyak *et al.*[3] (EMPIAR-10930) and show that over half of the particles in the dataset correspond to ALK-ALKAL2 complexes obeying a 2:1 stoichiometry besides the originally reported structure displaying 2:2 stoichiometry. Our workflow critically relied on extensive particle orientation rebalancing in cryoSPARC followed by Blush refinement[4] in Relion leading to a 3D reconstruction of the 2:1 ALK-ALKAL2 complex to 3.2 Å resolution. This cryo-EM structure together with the reported crystal structures of ALK-ALKAL2 and LTK-ALKAL1 complexes with 2:1 stoichiometry, reconcile a common receptor dimerization mode for ALK and LTK poised for the facile juxtapositioning of their intracellular kinase domains for signalling.

Main Text

Anaplastic lymphoma kinase (ALK)[5] and the related leukocyte tyrosine kinase (LTK)[6] are receptor tyrosine kinases (RTKs)[7] that are activated by binding their cognate cytokines, ALKAL1 and ALKAL2 (also called FAM150A/B or AUGβ/α)[8-10]. Their signaling outputs elicit a range of pleiotropic activities in development and metabolism, and their dysregulated intracellular kinases are keenly pursued as therapeutic drug targets in various cancers[11-17]. However, the field had long been hampered by a paucity of insights into the intricate folds of ALK and LTK and their cytokine-bound complexes. A trio of papers recently published in *Nature* [1-3] have aimed to provide the missing structural and mechanistic framework for the cytokine-mediated activation of ALK family receptors but diverged in their proposed assemblies.

Our study by De Munck *et al.*[1] provided representative assemblies for the entire ALK family by reporting crystal structures of both human ALK and LTK in complex with ALKAL2 and ALKAL1, respectively (PDB entries 7nwz and 7nx0). The reported ALK–ALKAL1 and LTK–ALKAL2 complexes comprised the ligand-binding extracellular regions of ALK and LTK without the membrane proximal EGFL domain (ALK_{TG}/LTK_{TG}) and revealed how a single cytokine molecule nucleates homodimeric receptor assembly via distinct site 1 and 2 interaction interfaces, creating a key site 3 receptor-receptor contact (**Figure 1a**). This compact architecture is poised to gather the membrane-proximal EGFL domains at the cell surface, effectively clustering their closely linked transmembrane (TM) helices and enabling the proximity-mediated activation of the intracellular kinase domains of ALK and LTK. However, the absence of membrane proximal receptor EGFL domains in our ALK_{TG}–ALKAL2 and LTK_{TG}–ALKAL1 complex structures could not answer the question of how the EGFL domains might organize proximal to the membrane to allow subsequent formation of receptor-receptor contacts via the TM helices of ALK.

In contrast, Li *et al.*[2] and Reshetnyak *et al.*[3] have focused only on ALK-ALKAL2 complexes comprising the extracellular ligand-binding fragment of ALK and the membrane-proximal EGFL domains, and reported assemblies obeying a 2:2 stoichiometry obtained via X-ray crystallography (PDB entry 7LS0) and electron cryo-microscopy (cryo-EM, PDB entry 7n00), respectively (**Figure 1b**). In this side-by-side architecture, the main

cytokine-receptor interface corresponds to the high affinity site 1 also reported by De Munck et al.[1]. An additional ALK-ALKAL2' interaction, covering 250 Å^2 of buried surface area was proposed to be responsible for bridging the two 1:1 ALK-ALKAL2 subcomplexes into an assembly that could support signaling. The membrane-proximal EGFL domains of $ALK_{TG-EGFL}$ in the 2:2 complex as reported by Reshetnyak *et al.* — and suggested by the 2:2 ALK_{TG}-ALKAL2 complex reported by Li *et al.* — are separated by a distance of 87 Å and thus require fully stretched 13-residue linkers along the membrane surface to pair the ALK TM helices (Reshetnyak *et al.* **Figure 4b** in reference 3) .

Although the crystal structure by Li *et al.* was obtained from crystals containing a fusion protein comprising ALK_{TG} and ALKAL2, thus inherently favoring assemblies that are multiples of a 1:1 stoichiometry (e.g. 2:2), the cryo-EM structure of the 2:2 $ALK_{TG-EGFL}$-ALKAL2 complex reported by Reshetnyak *et al*. demonstrated that such an assembly corresponds to a stable complex in solution. Intrigued by the 2:2 stoichiometry differing from our crystal structures of truncated ALK_{TG} in complex with ALKAL2 (and LTK_{TG} with ALKAL1), we closely examined the cryo-EM data and analysis reported by Reshetnyak *et al*.

At the outset, we identified a number of 3D-classes (Reshetnyak *et al*., Extended Data Figure 4 in reference [3], classes 1, 3,7, 8 and 9) resembling the compact 2:1 ALK_{TG}/LTK_{TG}-cytokine complexes we had reported[1]. To investigate this apparent particle heterogeneity, we leveraged the raw cryo-EM data and uncleaned particle stack as deposited by Reshetnyak *et al.* (access code EMPIAR-10930). To avoid any bias, we accounted for all particles instead of a selected subset, to enable a complete analysis of the particle population in the micrographs. Upon extracting 8x binned particles followed by several iterative cycles of reference-free 2D classification to identify 'junk' particles, we obtained a 'cleaned' set comprising ~11.8 million particles. We note that this number of particles approaches the total number of particles (~12.4 million) reported by Reshetnyak *et al*. Reference-free 2D classification of unbinned particles revealed that 6 of the 10 most populated classes do not correspond to classes shown in the final 2D classification reported in Reshetnyak *et al.* (**Figure 2a**). This prompted us to pursue extensive data processing of particles present in both sets of 2D classes, namely the ones that correspond to clear 2-fold symmetry (**Figure 2a, orange squares**) and the ones that lack apparent symmetry (**Figure 2a, blue squares**).

Ab initio models obtained using a selected subset of most populated classes (1.5 million particles, 3 *Ab Initio* classes, C1 symmetry), revealed two classes, one resembling the 2:1 ALK_{TG}–ALKAL2 complex presented in De Munck *et al.* (**Figure 2b, left**) and one class corresponding to the 2:2 $ALK_{TG-EGFL}$–ALKAL2 complex reported by Reshetnyak *et al.* and Li *et al*. (**Figure 2b, right**), with particles being roughly equally divided between these two distinct stoichiometries. Hetero-refinement of the *Ab Initio* models using the 'cleaned' particle stack of ~11.8 million particles resulted in roughly 4 million particles belonging to 3D classes clearly resembling either a 2:1 (class 0, 1 & 2) or 2:2 (class 6, 7 & 8) ALK-ALKAL2 stoichiometry (**Extended Data Figure 1**). Thus, at least half of the particles present in the dataset deposited in EMPIAR-10930 do not correspond to the assembly reported by Reshetnyak et al. Following hetero-refinement and 3D classification, a final selected set of 334,654 particles was used as input to Non-Uniform (NU) refinement[18] (**Extended Data Figure 1**). However, the final NU-refined map of the 2:1 assembly displayed severe preferential particle orientations leading to a map that is smeared along the Z-direction, a common problem in cryo-EM requiring sample and data collection optimization.

To overcome the apparent anisotropy presented in the 3D reconstructions towards a 2:1 $ALK_{TG-EGFL}$–ALKAL2 complex, we employed exhaustive particle orientation rebalancing in cryoSPARC[19] (v4.5) to exclude particles from over-populated direction bins, followed by 3D refinement with Blush regularization[4] in Relion v5. Blush regularization uses a pre-trained denoising convolutional neural network on pairs of half-set reconstructions as a smoothness prior and is shown to remove artifactual densities resulting from uneven angular distributions or streaky features in overfitted maps[4]. This new regularization approach to resolving vexing cryo-EM structural data was successfully employed to reveal the distinct binding modes of the anti-CRISPR repressor Aca2 with DNA and RNA[20].

We started from a set of selected 3D classes (1.4 million particles) after 3D hetero-refinement of the 'cleaned' dataset, and divided these 3D classes in two sets based on the absence or presence of apparent twofold symmetry. Such grouping resulted in 3D classes resembling either the C1-symmetric 2:1 ALK_{TG}-ALKAL2 crystal structure reported by De Munck et al.[1] (876.806 particles, ~60% of the total), or the deposited C2-symmetric

2:2 $ALK_{TG-EGFL}$-ALKAL2 cryo-EM map reported by Reshetnyak *et al.*[3] (523.467 particles, ~40% of the total) (**Extended Data Figure 2**).

Firstly, we wondered whether our approach would lead to a cryo-EM map of similar resolution and informational content for the C2 symmetric $ALK_{TG-EGFL}$-ALKAL2 complex with 2:2 stoichiometry as reported by Reshetnyak *et al.*[3]. Indeed, further data processing in cryoSPARC, one additional round of 2D classification to remove remaining junk classes, *ab-Initio* model generation (2 classes), hetero-refinement and 3D NU-refinement of the highest resolution class while applying C2 symmetry, resulted in a 2.3 Å resolution map for the 2:2 $ALK_{TG-EGFL}$-ALKAL2 complex (**Extended Data Figure 2 and Methods**). The resulting map has a comparable resolution as the deposited map by Reshetnyak *et al. (*EMD-24095, 2.3 Å) and, importantly, does not contain any particles from 2D classes corresponding to a 2:1 $ALK_{TG-EGFL}$-ALKAL2 stoichiometry (**Extended Data Figure 2**). We note that, similarly as the cryoSPARC refined map reported in Reshetnyak *et* al., map density for the EGFL domains is less defined than for the Relion refined map obtained by Reshetnyak *et al.* at slightly lower resolution.

We then pursued analysis of the particle set with apparent resemblance to $ALK_{TG-EGFL}$-ALKAL2 complexes with 2:1 stoichiometry. We first performed 2D classification to remove remaining junk classes and subjected the resulting particles to automatic Orientation Rebalancing in cryoSPARC v4.5. After *Ab-Initio* model generation (3 classes), hetero-refinement, NU-refinement of the highest resolution class and an additional round of Orientation Rebalancing followed by a final NU-refinement in cryoSPARC v4.5, we obtained a map that was substantially more isotropic yet still displaying anisotropic features even after post-processing via deepEMhancer[21]. Remarkably, upon performing 3D refinement using the orientation rebalanced particle set in Relion v5 with enabled Blush regularization, we were able to generate a massively improved final map at 3.2 Å resolution (**Table 1**) displaying an impressive improvement in isotropy, as evidenced by the 10-fold increase in the cFAR score (**Extended Data Figure 2**). Most importantly, regular post-processing in Relion v4.5 yielded a sharpened map with clear main chain connectivity and density for most side chains, allowing complete model building of a 2:1 $ALK_{TG-EGFL}$-ALKAL2 complex that includes the membrane-proximal EGFL-like domains of ALK. While the majority of the 2:1 $ALK_{TG-EGFL}$-ALKAL2 cryo-EM map has a local resolution to better than 3.5 Å resolution (**Extended Data Figure 3**), the C-terminal tips of the EGFL domains are only resolved at 4 – 4.5 Å resolution pointing to their inherent flexibility.

This new structure of $ALK_{TG-EGFL}$-ALKAL2 with 2:1 stoichiometry is a critical contribution to the present collection of ALK/LTK-cytokine complexes revealed by X-ray crystallography and cryo-EM and is now poised to provide constructive input to ongoing research exploring the role of ALK-cytokine complexes with 2:1 versus 2:2 stoichiometries in signaling. The elucidation of an $ALK_{TG-EGFL}$-ALKAL2 complex with 2:1 stoichiometry based on a dataset that was initially reported to only support $ALK_{TG-EGFL}$-ALKAL2 complexes with 2:2 stoichiometry underscores the power of modern cryo-EM to harness the conformational and stoichiometric diversity harbored in protein samples. Most importantly, though, this new structure is now the most complete ALK-ALKAL2 complex with 2:1 stoichiometry to date and agrees closely with the ALK_{TG}-ALKAL2 with 2:1 stoichiometry reported by De Munck et al.[1] featuring site 1 (ALK-ALKAL2), site 2 (ALK'-ALKAL2), and site 3 (ALK-ALK') interaction interfaces (**Figure 2**, **Extended Data Figure 4**). Accordingly, the root-mean-square deviation (r.m.s.d.) between these two structures is 1.3 Å over 675 aligned Cα-atoms (**Extended Data Figure 4a**). The site 2 interface present in the 2:1 ALK:ALKAL2 complex, but not in the 2:2 ALK-ALKAL2 complex, involves the α1 helix of ALKAL2 with a key histidine residue (H100) interacting with S737, Y739 and L760 of ALK' (**Figure 2e)**. Removal of the α1 helix in ALKAL2 or the combined reverse-charge mutation of four of its residues (K94E, K96E, K99E and H100E) abrogated activation of ALK in cells, and were hypothesized interrupt a proposed interaction of the α1 helix of ALKAL2 with the cell membrane[3]. In addition, an H100E mutation at the center of the site 2 interface was shown to nullify the interaction with ALK'[1,3]. Thus, a 2:1 ALK-ALKAL2 assembly not only provides a direct structural explanation for the functional role of a site 2 interaction, but also provides the structural blueprint for all currently available mutagenesis data[1,3].

The membrane proximal EGFL domain of ALK engaged in the 'site 1' interaction adopts a conformation identical to ALK found in the 2:2 $ALK_{TG-EGFL}$-ALKAL2 complex (r.m.s.d. = 0.554 Å over 345 aligned Cα-atoms, **Extended Data Figure 4c**). Likewise, a 1:1 $ALK_{TG-EGFL}$-ALKAL2 binary complex extracted from the 2:2 $ALK_{TG-EGFL}$-ALKAL2 structure matches the 'site 1' interface found in the 2:1 $ALK_{TG-EGFL}$-ALKAL2 cryo-EM model (r.m.s.d. = 0.631 over 400 aligned Cα-atoms, **Extended Data Figure 4b**). Due to the asymmetric nature of the ALKAL2 monomer, the EGFL domain of the opposing ALK protomer (ALK') in the 2:1 $ALK_{TG-EGFL}$-ALKAL2 complex does

not make any direct contacts at the 'site 2' interaction with ALKAL2 (**Figure 2**). As a result, the ALK' EGFL domain is less resolved at its extremity due to increased flexibility, resulting in the absence of density for the minor β-hairpin in the cryo-EM map (**Figure 2, Extended Data Figure 3**). Moreover, the major β-hairpin of ALK' has an outwards tilted conformation compared to the EGFL domain of ALK engaged in the 'site 1' interaction (r.m.s.d.= 1.717 Å over 331 aligned Cα-atoms, **Extended Data Figure 4d**). Despite this asymmetry, the C-termini of the EGFL domains of the two ALK copies in the 2:1 $ALK_{TG\text{-}EGFL}$-ALKAL2 complex are now spaced ~30 Å apart (**Figure 2**), a distance that would allow facile juxtapositioning of receptor chains and TM helices as already illustrated for many RTK—cytokine assemblies featuring cytokine-mediated receptor dimerization such as KIT–SCF[22,23], CSF-1R–CSF-1[24], Flt3–Flt3R[25], VEGF-VEGFR[26], PDGF-PDGFR[27], EGF–EGFR[28], HER2–HER3–NRG1β[29] and Insulin bound to insulin receptor (IR) family receptors[30,31]. Indeed, when the sequences of ALK (without the N-terminal HBD-MAM1-MAM2 domains and C-terminal unstructured tail) and full-length ALKAL2 are used to predict a 2:1 complex using AlphaFold 3 (AF3)[32], the resulting model contains dimerized TM helices and cytoplasmic kinase domains (**Extended Data Figure 5a & 5b**) suggesting that the geometry of the 2:1 $ALK_{TG\text{-}EGFL}$-ALKAL2 structure is compatible with ALK signaling.

Although the N-terminus of ALKAL2 is dispensable for receptor activation[9], full-length ALKAL2 has the potential to form a disulfide linked homodimer via a cysteine residue (C66) at its N-terminal end. We therefore wondered whether signaling-competent assemblies would be possible via complexes with 2:1 ALK-ALKAL2 stoichiometries based on dimeric ALKAL2. As suggested by Reshetnyak *et al*. (**Extended Data Figure 2 in reference [3]**) with an AlphaFold-2 rendition of ALKAL2, the AF3 model of full-length ALKAL2 predicts the N-terminal region of ALKAL2 to be an α-helix connected to the ALKAL2 three-helix bundle via a long 20-residue flexible linker (**Extended Data Figure 5a & 5c**). Moreover, an AF3 prediction of the complex between two ALKAL2 N-terminal regions (residues 24 – 92) readily results in a confidently-predicted leucine zipper-like coiled-coil with a flexible C-terminus featuring C66 residues of opposing helices closely facing one another (**Extended Data Figure 5c & 5d**). Since the N-terminus (starting at M93) of ALKAL2 in the 2:1 $ALK_{TG\text{-}EGFL}$-ALKAL2 complex is facing the solvent and is completely unobstructed, dimeric ALKAL2 would not sterically hinder the formation of two adjacent 2:1 ALK-ALKAL2 complexes. As can be seen from Extended Data Figure 5, the 20-residue flexible linker between the ALKAL2 three-helix bundle and C66 (involved in S-S linked dimer formation) does not clash with any of the present ALK receptors and has ample space to flare outwards, allowing the formation of two adjacent membrane-embedded complexes with 2:1 stoichiometry tethered by C66-disulfide bridging of short coiled-coil N-terminal helices from neighboring ALKAL2 molecules.

The presence of two distinct stoichiometries of the ALK-ALKAL2 complex in the EMPIAR-10930 dataset (**Figure 3**) is intriguing, and their role in ALK signaling is currently unclear. One hypothesis could be that increasing local cytokine concentrations could shift an ALK-ALKAL2 assembly from a 2:1 to a 2:2 stoichiometry, possibly leading to different signaling outcomes. Interestingly, designed ligands for erythropoietin receptor aimed at tuning the distance between receptor monomers upon dimerization result in modulation of erythropoietin receptor signaling and phosphorylation efficiency of downstream adaptors[33].

Nevertheless, the herein resolved cryo-EM structure of a 2:1 ALK-ALKAL2 complex together with the reported crystal structures of ALK-ALKAL2 and LTK-ALKAL1 complexes with 2:1 stoichiometry[1], reconcile a common receptor dimerization mode for both ALK and LTK that allows the easy pairing of their TM helices, that assures contiguity of intracellular kinase domains for signalling. Further structural, biophysical and cellular studies, including full-length ALK-ALKAL2 complexes, will be required to fully define the possible roles of distinct ALK/LTK-cytokine stoichiometries.

Methods

Reanalysis of cryo-EM data as reported in EMPIAR-10930

Initial data analysis was carried out in CryoSPARC[19] v3.3.1, and additional steps were performed in either cryoSPARC[19] v4.5 or relion[34,35] v5, specifically for the use of Orientation Rebalancing ('Rebalance Orientations' job) and 3D refinement including Blush regularization[4] respectively. Dose-weighted and motion-corrected micrographs were downloaded from https://www.ebi.ac.uk/empiar/EMPIAR-10930/ and were CTF-corrected using patch CTF estimation. After importing the uncleaned particle stack (~18 million particles) from EMPIAR-10930 and extracting the corresponding particles (8x binned), several iterative cycles of reference-free 2D classification were performed to get rid of 'junk' particles, resulting in a 'cleaned' set comprising ~11.8 million particles. This number of particles approaches the total number of particles (~12.4 million) used by Reshetnyak *et al*. Particle unbinning followed by reference-free 2D classification revealed that almost 2/3 of the ten most populated classes do not correspond to classes shown in the final 2D classification represented in Reshetnyak *et al.* (**Figure 2a**). *Ab initio* models generated on a selected subset of most populated classes (4x binned, 1.5 million particles, 3 *Ab Initio* classes, C1 symmetry) followed by a round of heterorefinement revealed one model that clearly corresponds to the 2:2 ALK–ALKAL2 structure reported by Reshetnyak *et al.,* as well as a model resembling the 2:1 ALK–ALKAL2 structure presented in De Munck *et al.* (**Figure b2 & Extended Data Figure 1**). Next, the three *Ab Initio* classes were used together with the cleaned dataset of ~11.8 million particles (unbinned), as input for heterorefinement (C1, 9 classes). Class 1 and 2 of this heterorefinement were grouped together, and selected particles (~1.2 million) were used for 3D classification without alignment (8 classes). Class 5 after 3D classification was selected (334,654 particles) and used as an input for a final Non-Uniform (NU) refinement[18], While the final map shows the presence of two ALK copies and one ALKAL2 copy (with its characteristic three central α-helices), the map suffers from the severe preferential particle orientations of the 2:1 ALK–ALKAL2 particles, as evidenced by the paucity of different views in the 2D classes as well as the Azimuth versus Elevation plot shown in Extended Data Figure 1. Indeed, Orientation Diagnostics, as implemented in the newest version of cryoSPARC[19] (v4.5), pointed to extreme anisotropy corresponding to a cFAR score of 0.03 (**Extended Data Figure 1)**. Accordingly, the anisotropy present in the obtained NU-refined map of the 2:1 $ALK_{TG\text{-}EGFL}$–ALKAL2 complex prohibited further model building and refinement.

Based on the heterorefinement of 11.849.911 particles, highest resolution 3D classes were grouped in two pools based on the presence or absence of twofold symmetry. Additional 2D classification of the particles without apparent symmetry (876.806 particles) showed that 2D classes mainly consisted of side views of the 2:1 ALK-ALKAL2 complex, with a lack of discernible top views (**Extended Data Figure 2**). We next used Orientation Rebalancing in cryoSPARC v4.5 to trim overpopulated views in the particle stack using a rebalance percentile threshold of 60, resulting in 599.954 remaining particles. These particles were used as an input for an Ab Initio refinement job (C1, 3 classes) followed by heterorefinement (C1, 3 classes). The highest resolution 3D class (class 1) was used for NU-refinement, followed by another round of Orientation Rebalancing, now using a more stringent rebalance percentile threshold of 5. Retained particles (142.986) where used for a second NU-refinement and subsequently exported to Relion v5 for a final round of 3D refinement with enabled Blush regularization. A final 2D classification was performed on the final particle stack (142.986 particles) to demonstrate the appearance of new side and top views not seen in previous 2D classification runs (**Extended Data Figure 2**). Orientation Diagnostics jobs were run on obtained intermediate 3D reconstructions as well as the final map to monitor improvements in map isotropy. The final obtained map has a Gold-Standard Fourier Shell Correlation (FSC) resolution of 3.2 Å at the 0.143 threshold (**Extended Data Figure 3, Table 1**), and was post-processed either in Relion using a sharpening B-factor of -100 $Å^2$, or by using deepEMhancer[21].

Model Building & Refinement

Two copies of $ALK_{TG\text{-}EGFL}$ and one copy of ALKAL2 were extracted from PDB ID 7n00, and rigid-body fitted in the sharpened map of the 2:1 $ALK_{TG\text{-}EGFL}$-ALKAL2 complex (B-factor: -100 $Å^2$) using USCF Chimera[36]. The resulting rigid-body fitted model was subjected to a cycle of manual building in Coot[37], aided by both the regularly sharpened map and a map sharpened using DeepEMhancer[21]. Next, the manually-built and corrected model

was real-space refined against the regularly sharpened map in Phenix v1.19.2-4158[38], using global minimization, local grid search, atomic displacement parameter (ADP) refinement, secondary structure restraints and Ramachandran restraints. Several cycles of manual building in Coot were followed by real-space refinement in Phenix. After the last cycle of manual building in Coot, a nonbonded-weight parameter of 200 was used during refinement in Phenix. The resulting model was subjected to a final refinement against the two half-maps in REFMAC-Servalcat[39] run within CCP-EM suite v1.6[40], with automatic weighting applied. The final model has a map-model FSC of 3.20 Å at the 0.5 threshold, calculated using the unfiltered, unsharpened full map (**Extended Data Figure 3)**. A summary of cryo-EM data collection, processing, refinement and validation statistics can be found in Table 1.

AlphaFold 3 (AF3) Modeling

Alphafold 3[32] modeling was performed using the AlphaFold Server (https://alphafoldserver.com/) using following sequences:

2: 1 ALK-ALKAL2 complex:

- Uniprot entry Q9UM73 (human ALK), residues 681 – 1410, 2 copies:

VHWLFTTCGASGPHGPTQAQCNNAYQNSNLSVEVGSEGPLKGIQIWKVPATDTYSISGYGAAGGKGGKNTMMRSHGVSV
LGIFNLEKDDMLYILVGQQGEDACPSTNQLIQKVCIGENNVIEEEIRVNRSVHEWAGGGGGGGGATYVFKMKDGVPVPLIIA
AGGGGRAYGAKTDTFHPERLENNSSVLGLNGNSGAAGGGGGWNDNTSLLWAGKSLQEGATGGHSCPQAMKKWGWET
RGGFGGGGGGCSSGGGGGGYIGGNAASNNDPEMDGEDGVSFISPLGILYTPALKVMEGHGEVNIKHYLNCSHCEVDECH
MDPESHKVICFCDHGTVLAEDGVSCIVSPTPEPHLPLSLILSVVTSALVAALVLAFSGIMIVYRRKHQELQAMQMELQSPEYKL
SKLRTSTIMTDYNPNYCFAGKTSSISDLKEVPRKNITLIRGLGHGAFGEVYEGQVSGMPNDPSPLQVAVKTLPEVCSEQDELDF
LMEALIISKFNHQNIVRCIGVSLQSLPRFILLELMAGGDLKSFLRETRPRPSQPSSLAMLDLLHVARDIACGCQYLEENHFIHRDI
AARNCLLTCPGPGRVAKIGDFGMARDIYRASYYRKGGCAMLPVKWMPPEAFMEGIFTSKTDTWSFGVLLWEIFSLGYMPYP
SKSNQEVLEFVTSGGRMDPPKNCPGPVYRIMTQCWQHQPEDRPNFAIILERIEYCTQDPDVINTALPIEYGPLVEEEEK

- Uniprot entry Q6UX46 (human ALKAL2), residues 25-152, 1 copy:

GAEPREPADGQALLRLVVELVQELRKHHSAEHKGLQLLGRDCALGRAEAAGLGPSPEQRVEIVPRDLRMKDKFLKHLTGPLY
FSPKCSKHFHRLYHNTRDCTIPAYYKRCARLLTRLAVSPVCMEDKQ

2: 1 ALKΔkinase-ALKAL2 complex:

- Uniprot entry Q9UM73 (human ALK), residues 681 – 1074, 2 copies:

VHWLFTTCGASGPHGPTQAQCNNAYQNSNLSVEVGSEGPLKGIQIWKVPATDTYSISGYGAAGGKGGKNTMMRSHGVSV
LGIFNLEKDDMLYILVGQQGEDACPSTNQLIQKVCIGENNVIEEEIRVNRSVHEWAGGGGGGGGATYVFKMKDGVPVPLIIA
AGGGGRAYGAKTDTFHPERLENNSSVLGLNGNSGAAGGGGGWNDNTSLLWAGKSLQEGATGGHSCPQAMKKWGWET
RGGFGGGGGGCSSGGGGGGYIGGNAASNNDPEMDGEDGVSFISPLGILYTPALKVMEGHGEVNIKHYLNCSHCEVDECH
MDPESHKVICFCDHGTVLAEDGVSCIVSPTPEPHLPLSLILSVVTSALVAALVLAFSGIMIVYRRKHQELQAMQMELQ

- Uniprot entry Q6UX46 (human ALKAL2), residues 25-152, 1 copy:

GAEPREPADGQALLRLVVELVQELRKHHSAEHKGLQLLGRDCALGRAEAAGLGPSPEQRVEIVPRDLRMKDKFLKHLTGPLY
FSPKCSKHFHRLYHNTRDCTIPAYYKRCARLLTRLAVSPVCMEDKQ

2: 1 ALKΔkinase-ALKAL2Δ24-92 complex:

- Uniprot entry Q9UM73 (human ALK), residues 681 – 1074, 2 copies:

VHWLFTTCGASGPHGPTQAQCNNAYQNSNLSVEVGSEGPLKGIQIWKVPATDTYSISGYGAAGGKGGKNTMMRSHGVSV
LGIFNLEKDDMLYILVGQQGEDACPSTNQLIQKVCIGENNVIEEEIRVNRSVHEWAGGGGGGGGATYVFKMKDGVPVPLIIA

AGGGGRAYGAKTDTFHPERLENNSSVLGLNGNSGAAGGGGGWNDNTSLLWAGKSLQEGATGGHSCPQAMKKWGWET
RGGFGGGGGGCSSGGGGGGYIGGNAASNNDPEMDGEDGVSFISPLGILYTPALKVMEGHGEVNIKHYLNCSHCEVDECH
MDPESHKVICFCDHGTVLAEDGVSCIVSPTPEPHLPLSLILSVVTSALVAALVLAFSGIMIVYRRKHQELQAMQMELQ

- Uniprot entry Q6UX46 (human ALKAL2), residues 93-152, 1 copy:

MKDKFLKHLTGPLYFSPKCSKHFHRLYHNTRDCTIPAYYKRCARLLTRLAVSPVCMEDKQ

ALKAL2 N-terminal helix dimer:

- Uniprot entry Q6UX46 (human ALKAL2), residues 25-92, 2 copies :

GAEPREPADGQALLRLVVELVQELRKHHSAEHKGLQLLGRDCALGRAEAAGLGPSPEQRVEIVPRDLR

Data Availability

The cryo-EM map and accompanying structural model for the 2:1 $ALK_{TG\text{-}EGFL}$-ALKAL2 complex have been deposited in the EMDB/PDB with accession codes: EMD-51087/9g5i. The particle .star file (run_data.star) obtained after Blush Refinement in Relion 5 is added to this manuscript as a Supplementary Data File, and can be used to extract particles after running CTF estimation on the Motion Corrected and Dose Weighted (*_DW.mrc) micrographs available from EMPIAR-10930 (https://www.ebi.ac.uk/empiar/EMPIAR-10930/).

References

1. De Munck, S. et al. Structural basis of cytokine-mediated activation of ALK family receptors. Nature 600, 143-147 (2021).
2. Li, T. et al. Structural basis for ligand reception by anaplastic lymphoma kinase. Nature 600, 148-152 (2021).
3. Reshetnyak, A.V. et al. Mechanism for the activation of the anaplastic lymphoma kinase receptor. Nature 600, 153-157 (2021).
4. Kimanius, D. et al. Data-driven regularization lowers the size barrier of cryo-EM structure determination. Nat Methods (2024).
5. Morris, S.W. et al. Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. Science 263, 1281-4 (1994).
6. Ben-Neriah, Y. & Bauskin, A.R. Leukocytes express a novel gene encoding a putative transmembrane protein-kinase devoid of an extracellular domain. Nature 333, 672-6 (1988).
7. Zhang, H. et al. Deorphanization of the human leukocyte tyrosine kinase (LTK) receptor by a signaling screen of the extracellular proteome. Proc Natl Acad Sci U S A 111, 15741-5 (2014).
8. Guan, J. et al. FAM150A and FAM150B are activating ligands for anaplastic lymphoma kinase. Elife 4, e09811 (2015).
9. Reshetnyak, A.V. et al. Identification of a biologically active fragment of ALK and LTK-Ligand 2 (augmentor-alpha). Proc Natl Acad Sci U S A 115, 8340-8345 (2018).
10. Reshetnyak, A.V. et al. Augmentor alpha and beta (FAM150) are ligands of the receptor tyrosine kinases ALK and LTK: Hierarchy and specificity of ligand-receptor interactions. Proc Natl Acad Sci U S A 112, 15862-7 (2015).
11. Borenas, M. et al. ALK ligand ALKAL2 potentiates MYCN-driven neuroblastoma in the absence of ALK mutation. EMBO J 40, e105784 (2021).
12. Hallberg, B. & Palmer, R.H. Mechanistic insight into ALK receptor tyrosine kinase in human cancer biology. Nat Rev Cancer 13, 685-700 (2013).
13. Janostiak, R., Malvi, P. & Wajapeyee, N. Anaplastic Lymphoma Kinase Confers Resistance to BRAF Kinase Inhibitors in Melanoma. iScience 16, 453-467 (2019).
14. Javanmardi, N. et al. Analysis of ALK, MYCN, and the ALK ligand ALKAL2 (FAM150B/AUGalpha) in neuroblastoma patient samples with chromosome arm 2p rearrangements. Genes Chromosomes Cancer (2019).
15. Li, N. et al. Gain-of-function polymorphism in mouse and human Ltk: implications for the pathogenesis of systemic lupus erythematosus. Hum Mol Genet 13, 171-9 (2004).
16. Pospisilik, J.A. et al. Drosophila genome-wide obesity screen reveals hedgehog as a determinant of brown versus white adipose cell fate. Cell 140, 148-60 (2010).
17. Sano, R. et al. An antibody-drug conjugate directed to the ALK receptor demonstrates efficacy in preclinical models of neuroblastoma. Sci Transl Med 11(2019).

18. Punjani, A., Zhang, H. & Fleet, D.J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. Nat Methods 17, 1214-1221 (2020).
19. Punjani, A., Rubinstein, J.L., Fleet, D.J. & Brubaker, M.A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods 14, 290-296 (2017).
20. Birkholz, N. et al. Phage anti-CRISPR control by an RNA- and DNA-binding helix-turn-helix protein. Nature 631, 670-677 (2024).
21. Sanchez-Garcia, R. et al. DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. Commun Biol 4, 874 (2021).
22. Yuzawa, S. et al. Structural basis for activation of the receptor tyrosine kinase KIT by stem cell factor. Cell 130, 323-34 (2007).
23. Krimmer, S.G. et al. Cryo-EM analyses of KIT and oncogenic mutants reveal structural oncogenic plasticity and a target for therapeutic intervention. Proc Natl Acad Sci U S A 120, e2300054120 (2023).
24. Felix, J. et al. Structure and Assembly Mechanism of the Signaling Complex Mediated by Human CSF-1. Structure 23, 1621-1631 (2015).
25. Verstraete, K. et al. Structural insights into the extracellular assembly of the hematopoietic Flt3 signaling complex. Blood 118, 60-8 (2011).
26. Markovic-Mueller, S. et al. Structure of the Full-length VEGFR-1 Extracellular Domain in Complex with VEGF-A. Structure 25, 341-352 (2017).
27. Chen, P.H., Unger, V. & He, X. Structure of Full-Length Human PDGFRbeta Bound to Its Activating Ligand PDGF-B as Determined by Negative-Stain Electron Microscopy. J Mol Biol 427, 3921-34 (2015).
28. Huang, Y. et al. A molecular mechanism for the generation of ligand-dependent differential outputs by the epidermal growth factor receptor. Elife 10(2021).
29. Diwanji, D. et al. Structures of the HER2-HER3-NRG1beta complex reveal a dynamic dimer interface. Nature 600, 339-343 (2021).
30. Uchikawa, E., Choi, E., Shang, G., Yu, H. & Bai, X.C. Activation mechanism of the insulin receptor revealed by cryo-EM structure of the fully liganded receptor-ligand complex. Elife 8(2019).
31. Wang, L., Hall, C., Li, J., Choi, E. & Bai, X.C. Structural basis of the alkaline pH-dependent activation of insulin receptor-related receptor. Nat Struct Mol Biol 30, 661-669 (2023).
32. Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature 630, 493-500 (2024).
33. Mohan, K. et al. Topological control of cytokine receptor signaling induces differential effects in hematopoiesis. Science 364(2019).
34. Scheres, S.H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. J Struct Biol 180, 519-30 (2012).
35. Kimanius, D., Dong, L., Sharov, G., Nakane, T. & Scheres, S.H.W. New tools for automated cryo-EM single-particle analysis in RELION-4.0. Biochem J 478, 4169-4185 (2021).
36. Pettersen, E.F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25, 1605-12 (2004).
37. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60, 2126-32 (2004).
38. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr D Struct Biol 75, 861-877 (2019).
39. Yamashita, K., Palmer, C.M., Burnley, T. & Murshudov, G.N. Cryo-EM single-particle structure refinement and map calculation using Servalcat. Acta Crystallogr D Struct Biol 77, 1282-1291 (2021).
40. Burnley, T., Palmer, C.M. & Winn, M. Recent developments in the CCP-EM software suite. Acta Crystallogr D Struct Biol 73, 469-477 (2017).
41. Chen, S. et al. High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. Ultramicroscopy 135, 24-35 (2013).

Acknowledgements

S.N.S. acknowledges research support from the FWO (grant number G0B4918N) and the Flanders Institute for Biotechnology (VIB, grant number C0101).

Author contributions

J.F and S.N.S. wrote the manuscript with contributions from all authors. J.F performed analysis and processing of the deposited EMPIAR-10930 cryo-EM dataset with contributions from S.D.M. J.F. and S.D.M. performed structural analyses with contributions from S.N.S. J.F. and J.F.B. performed AlphaFold modeling of 2:1 ALK–ALKAL2 complexes.

Competing interests

S.D.M. and S.N.S. are listed as co-inventors on a patent filing EP 21184332.1 with the European Patent Office.

Table 1. CryoEM data collection, processing, refinement and validation statistics

	2:1 $ALK_{TG\text{-}EGFL}$-ALKAL2
Data Collection (EMPIAR-10930, Reshetnyak et al. 2021[3])	
Magnification	15,000
Voltage (kV)	300
Electron exposure (e^{-}/Å^2)	88.4
Defocus range (μm)	0.8 – 1.8
Pixel size (Å)	0.826
Micrographs	13.618
Data Processing (This study)	
Symmetry imposed	C1
Initial particle images (no.)	18,053,705
Final particle images (no.)	142,986
Map resolution (Å)	3.2
FSC threshold	0.143
Map resolution range (Å)	3.0 – 5.5
Refinement (This study)	
Initial model used (PDB code)	7N00
Model resolution (Å)	3.2
FSC threshold	0.5
Map sharpening B factor (Å^2)	-100
Model composition	
Non-hydrogen atoms	5,252
Protein residues	732
B factors (Å^2)	
Protein	165.45
R.m.s. deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.193
Validation	
MolProbity score	1.86
Clashscore	9.85
Poor rotamers (%)	0.00
CaBLAM outliers (%)	2.64
CC (mask/volume) sharpened, unsharpened	0.76/0.71, 0.80/0.88
EMRinger (sharpened/unsharpened)	2.10/1.51
Ramachandran plot	
Favored (%)	95.04
Allowed (%)	4.96
Disallowed (%)	0.00
Ramachandran Z-score	
Whole	-1.92
Helix	-1.17
Sheet	-1.29
Loop	-1.25

1. Reshetnyak, A.V. et al. Mechanism for the activation of the anaplastic lymphoma kinase receptor. Nature 600, 153-157 (2021).

Figure Legends

Figure 1: Comparison of 2:1 ALK/LTK and 2:2 ALK receptor-ligand complexes. (a) Cartoon views of structures of ALK-ALKAL2 and LTK-ALKAL1 complexes with 2:1 stoichiometry as determined by De Munck *et al.*[1] (PDB entries 7nwz and 7nx0) and **(b)** of ALK-ALKAL2 with 2:2 stoichiometry as determined by Reshetnyak *et al.*[3] and Li *et al.*[2] (PDB entries 7n00 and 7ls0). Putative positions of missing EGFL domains in **(a)** are shown as dotted outlines.

Figure 2: Re-analysis of cryo-EM data from EMPIAR entry EMPIAR-10930 deposited by Reshetnyak *et al.,* Nature (2021). (a) 10 most populated 2D class averages after performing 2D classification on 11.8 million selected particles. Classes corresponding to 2:1 ALK–ALKAL2 particles are highlighted in blue, while classes corresponding to 2:2 ALK–ALKAL2 particles are highlighted in orange.
(b) *Ab initio* model generation in cryoSPARC[19] results in classes corresponding to both a 2:1 ALK-–ALKAL2 and 2:2 ALK–ALKAL2 assembly. Hetero-refinement of the 2:1 and 2:2 ALK–ALKAL2 classes results in the two models displaying 2:1 and 2:2 ALK–ALKAL2 assemblies, respectively. The ALK–ALKAL2 complexes with 2:1 (PDB 7nwz) and 2:2 (PDB 7n00) stoichiometries are shown fitted inside the maps, with ALK colored pink/grey and ALKAL2 colored blue/purple. **(c)** 3D refinement with Blush regularization[4] in Relion v5 of a trimmed particle set after extensive automatic rebalancing of particle orientations as implemented in cryoSPARC 4.5 (see also Methods & Table1). A front view is shown of a map sharpened in Relion using a B-factor of-100 $Å^2$, with ALK colored pink/grey and ALKAL2 colored blue. Front, side and top views are shown of a transparent map with fitted structural model. **(d & e)** Zoomed-in insets illustrating the quality of map density corresponding to the two major interaction interfaces (site 1 & 2) in the ALK:ALKAL2 complex. Key interacting residues are annotated and shown as sticks.

Figure 3: Side-by-side comparison of 2:1 and 2:2 ALK-ALKAL2 complexes present in the EMPIAR-10930 cryo-EM dataset. Models with pdb code 9g5i (this study) and 7n00 (Reshetnyak *et al*., 2021) are shown as cartoons with ALK/ALK' colored pink/grey and ALKAL2 colored blue in the 2:1 ALK-ALKAL2 complex, or blue/purple in the 2:2 ALK-ALKAL2 complex. The distance between the C-termini of membrane-proximal EGFL domains is annotated.

Extended Data Figure 1: (a) Initial processing workflow of cryo-EM dataset EMPIAR-10930. All initial processing steps were performed in CryoSPARC[19] v3.3.1. **(b)** Gold-standard Fourier Shell Correlation (FSC) plot corresponding to the final map obtained after non-uniform refinement (C1 symmetry). Curves are shown after applying no mask (blue), a loose mask (green), or a tight mask (red) to the two half maps before calculating FSCs. The corrected FSC (purple) is calculated using the tight mask with correction by noise substitution[41], and the resolution at FSC=0.143 is annotated via a dotted purple line. **(c)** conical FSC summary plot generated via 'Orientation Diagnostics' in cryoSPARC v4.5. **(d)** Left: Azimuth plot showing the distribution of orientations over Azimuth (X-axis) and Elevation (Y-axis) angles for the particle set corresponding to the NU-refined map shown in (A). Right: Plot showing Relative Signal amount vs. Viewing Direction.

Extended Data Figure 2: (a) Final processing workflow of cryo-EM dataset EMPIAR-10930. Final processing steps were performed in CryoSPARC[19] v4.5 or Relion[34,35] v5. Post-processing was performed in Relion or using deepEMhancer[21]. **(b & c)** Gold-standard Fourier Shell Correlation (FSC) plots corresponding to the final maps of the 2:1 **(b)** and 2:2 **(c)** ALK:ALKAL2 complexes. The corrected FSCs (purple lines) are calculated after performing correction by noise substitution[41], and the resolution at FSC=0.143 is annotated via dotted purple lines. **(d & e)** conical FSC summary plots corresponding to the final maps of the 2:1 **(d)** and 2:2 **(e)** ALK:ALKAL2 complexes, generated via 'Orientation Diagnostics' in cryoSPARC v4.5. **(f)** Azimuth plot showing the distribution of orientations over Azimuth (X-axis) and Elevation (Y-axis) angles for the particle set corresponding to the final

map of the 2:1 ALK-ALKAL2 complex, obtained after Orientation Rebalancing in cryoSPARC v4.5 and 3D refinement with Blush regularization[35] in Relion v5. **(f)** Display of the 20 most populated classes after 2D classification of the final particle stack corresponding to the 2:1 ALK-ALKAL2 complex. New side & top views not observed in initial 2D classification runs are shown as blue squares.

Extended Data Figure 3: (a) Local Resolution Estimation performed in Relion v5 of the final map for the 2:1 ALK-ALKAL 2 complex. **(b)** Gold-standard Fourier Shell Correlation (FSC) plot corresponding to the final map of the 2:1 ALK:ALKAL2 complex. The corrected FSC (purple lines) is calculated after performing correction by noise substitution[41], and the resolution at FSC = 0.143 is annotated via a dotted purple line. A map-to model FSC curve calculated using the fitted model and an unsharpened, unfiltered full map is shown (grey line) with the resolution at FSC=0.5 annotated via a dotted grey line. **(c)** conical FSC summary plot corresponding to the final map of the 2:1 ALK:ALKAL2 complex, generated via 'Orientation Diagnostics' in cryoSPARC v4.5. **(d)** Plot showing Relative Signal amount vs. Viewing Direction.

Extended Data Figure 4: Structural alignments of ALK-ALKAL2 X-ray and cryo-EM models. (a) Structural alignment of the 2:1 ALK-ALKAL2 X-ray structure (PDB 7nwz) with the cryo-EM structure of the 2:1 ALK-ALKAL2 complex obtained in this study (PDB 9G5I). **(b)** Structural alignment of a 1:1 ALK-ALKAL2 complex, extracted from the cryo-EM structure of the 2:2 ALK-ALKAL2 complex (PDB 7n00), with the cryo-EM structure of the 2:1 ALK-ALKAL2 complex obtained in this study. **(c)** Structural alignment of one copy of ALK, extracted from the cryo-EM structure of the 2:2 ALK-ALKAL2 complex (PDB 7n00), with the first ALK molecule (pink) present in the cryo-EM structure of the 2:1 ALK-ALKAL2 complex obtained in this study. **(d)** Structural alignment of one copy of ALK, extracted from the cryo-EM structure of the 2:2 ALK-ALKAL2 complex (PDB 7n00), with the second ALK molecule (ALK', grey) present in the cryo-EM structure of the 2:1 ALK-ALKAL2 complex obtained in this study.

Extended Data Figure 5: AlphaFold 3 prediction of the full-length 2:1 ALK-ALKAL2 complex. (a) AlphaFold 3[32] prediction of a 2:1 ALK-ALKAL2 complex using the sequence of full-length ALK (excluding N-terminal HBD-MAM1-MAM2 domains and C-terminal unstructured tail) and ALKAL2 (see Methods). Side and front cartoon views are shown with ALK colored pink/grey and ALKAL2 colored blue. Cysteine 66 (C66) of ALKAL2 is shown as yellow spheres. **(b)** Predicted Aligned Error (PAE) plot of the AF3 prediction of the 2:1 ALK-ALKAL2 complex, with annotated regions in the plot corresponding to interfaces at ALK-ALKAL2 site 1 (orange), ALK-ALKAL2 site 2 (blue), ALK-ALK' site 3 (red), the TM helix-TM helix interface (purple) and the interface between dimerized kinase domains (brown). pTM and iPTM scores are shown for the 2:1 ALK-ALKAL2 prediction shown in **(a)** as well as a prediction with omitted ALK kinase domains (annotated with *) and omitted ALK kinase domains and N-terminal residues 24-92 of ALKAL2 (annotated with **). **(c)** Alphafold 3 prediction of a dimer of N-terminal ALKAL2 residues 24 – 92. Interacting residues are annotated as sticks, with cysteine 66 (C66) shown as yellow spheres. **(d)** Alphafold 3 prediction shown in **(c)**, colored corresponding to the per-residue model confidence (pLDDT: predicted local distance difference test) score.

Figures

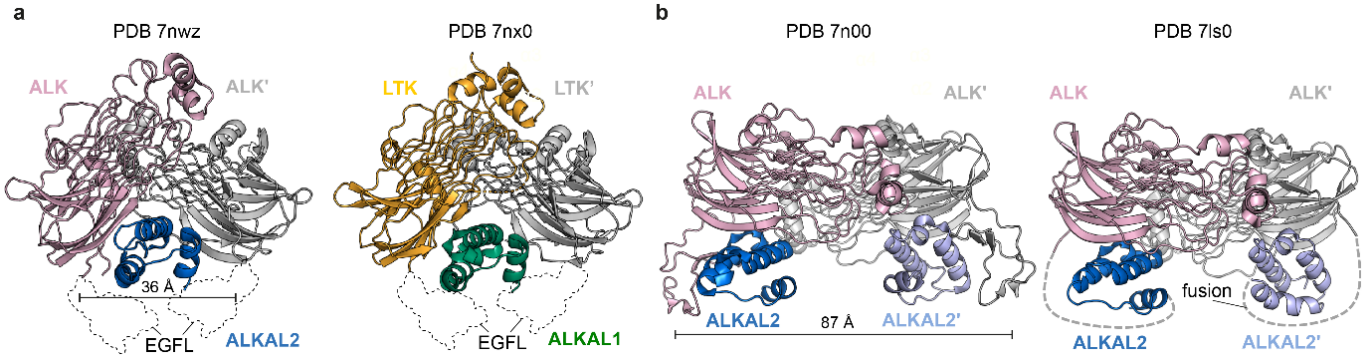


Figure 1: Comparison of 2:1 ALK/LTK and 2:2 ALK receptor-ligand complexes. (a) Cartoon views of structures of ALK-ALKAL2 and LTK-ALKAL1 complexes with 2:1 stoichiometry as determined by De Munck *et al.*[1] (PDB entries 7nwz and 7nx0) and **(b)** of ALK-ALKAL2 with 2:2 stoichiometry as determined by Reshetnyak *et al.*[3] and Li *et al.*[2] (PDB entries 7n00 and 7ls0). Putative positions of the EGFL domains absent in **(a)** are shown as dotted outlines.

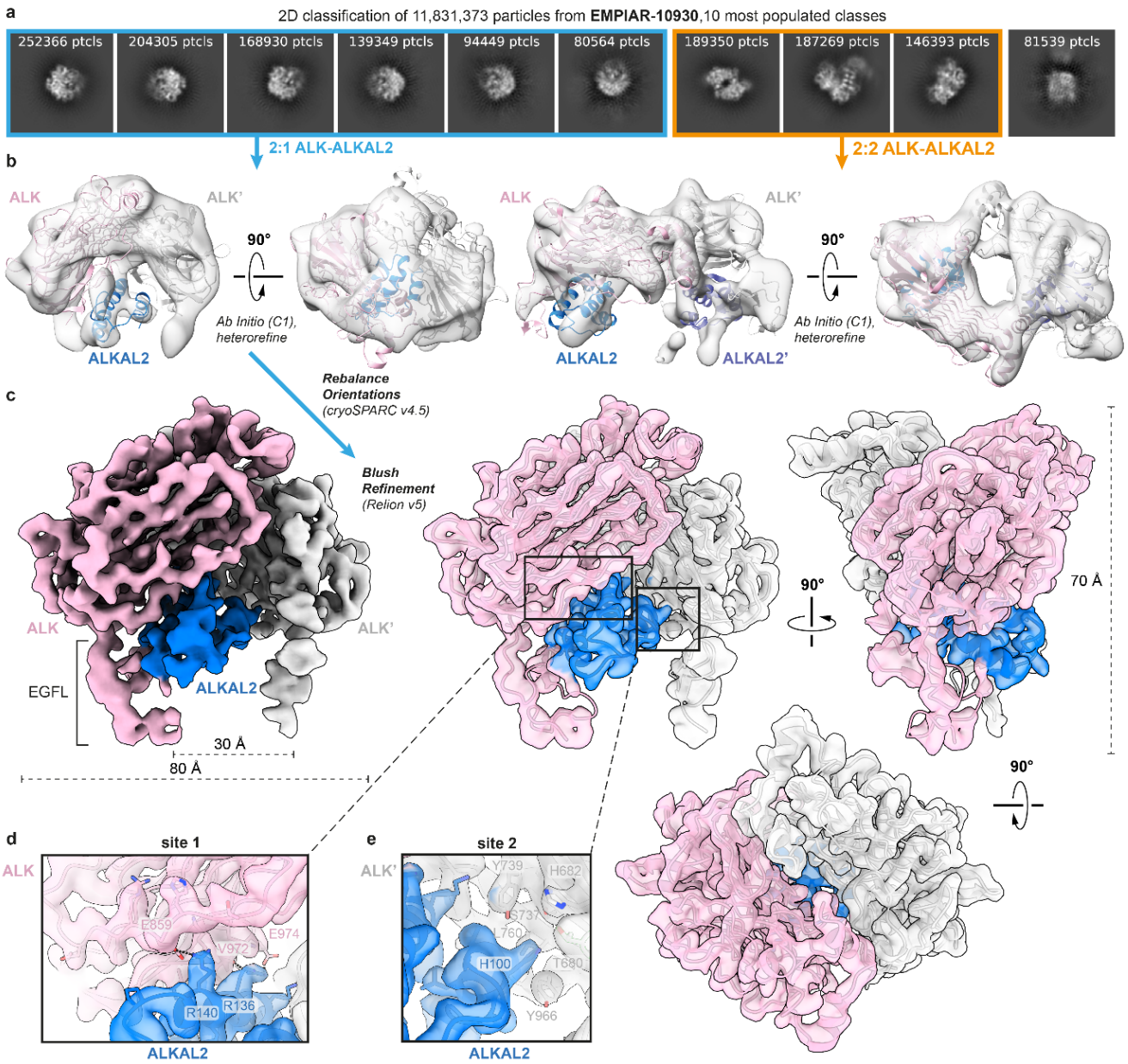


Figure 2: Re-analysis of cryo-EM data from EMPIAR entry EMPIAR-10930 deposited by Reshetnyak *et al.*, Nature (2021). (a) 10 most populated 2D class averages after performing 2D classification on 11.8 million selected particles. Classes corresponding to 2:1 ALK–ALKAL2 particles are highlighted in blue, while classes corresponding to 2:2 ALK–ALKAL2 particles are highlighted in orange.
(b) *Ab initio* model generation in cryoSPARC[19] results in classes corresponding to both a 2:1 ALK-–ALKAL2 and 2:2 ALK–ALKAL2 assembly. Hetero-refinement of the 2:1 and 2:2 ALK–ALKAL2 classes results in the two models displaying 2:1 and 2:2 ALK–ALKAL2 assemblies, respectively. The ALK–ALKAL2 complexes with 2:1 (PDB 7nwz) and 2:2 (PDB 7n00) stoichiometries are shown fitted inside the maps, with ALK colored pink/grey and ALKAL2 colored blue/purple. **(c)** 3D refinement with Blush regularization[4] in Relion v5 of a trimmed particle set after extensive automatic rebalancing of particle orientations as implemented in cryoSPARC 4.5 (see also Methods & Table1). A front view is shown of a map sharpened in Relion using a B-factor of-100 Å^2, with ALK colored pink/grey and ALKAL2 colored blue. Front, side and top views are shown of a transparent map, sharpened using deepEMhancer, with fitted structural model. **(d & e)** Zoomed-in insets illustrating the quality of map density corresponding to the two major interaction interfaces (site 1 & 2) in the ALK:ALKAL2 complex. Key interacting residues are annotated and shown as sticks.

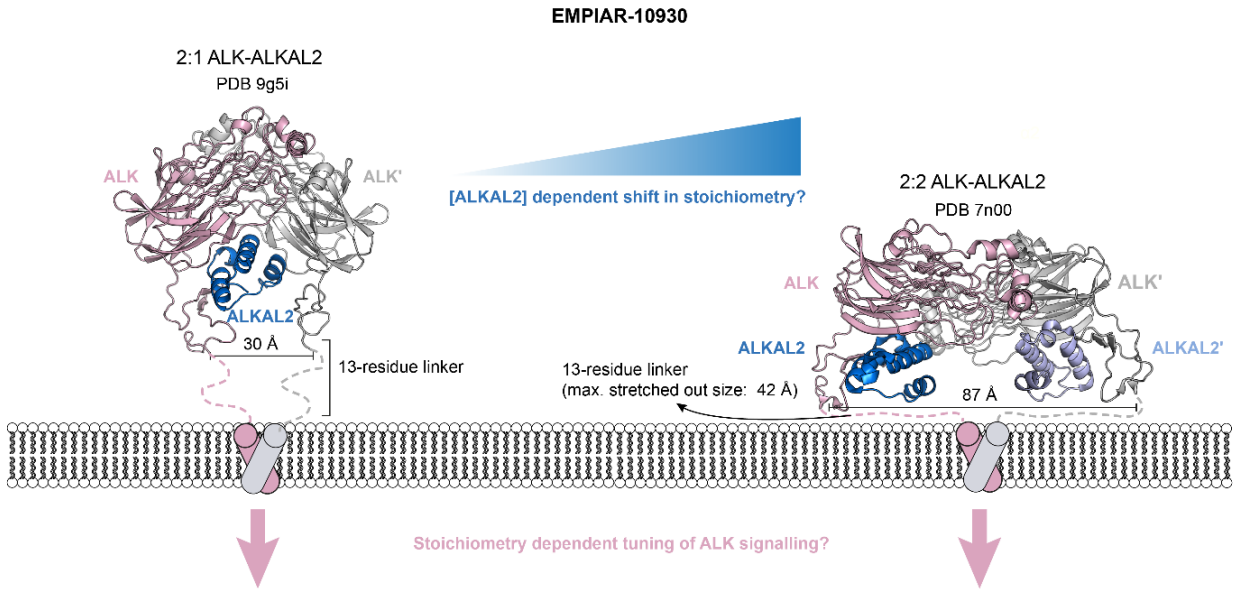
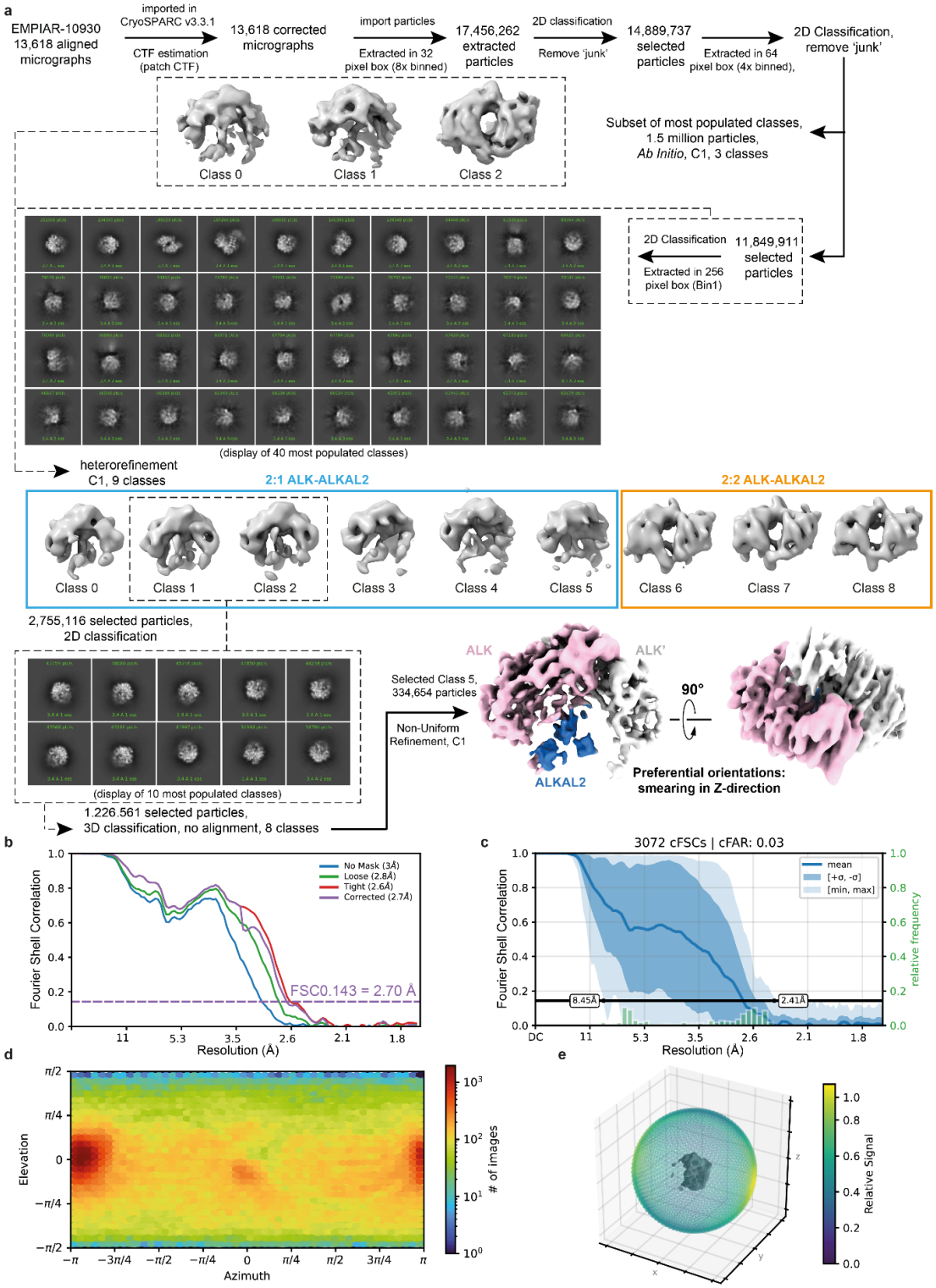
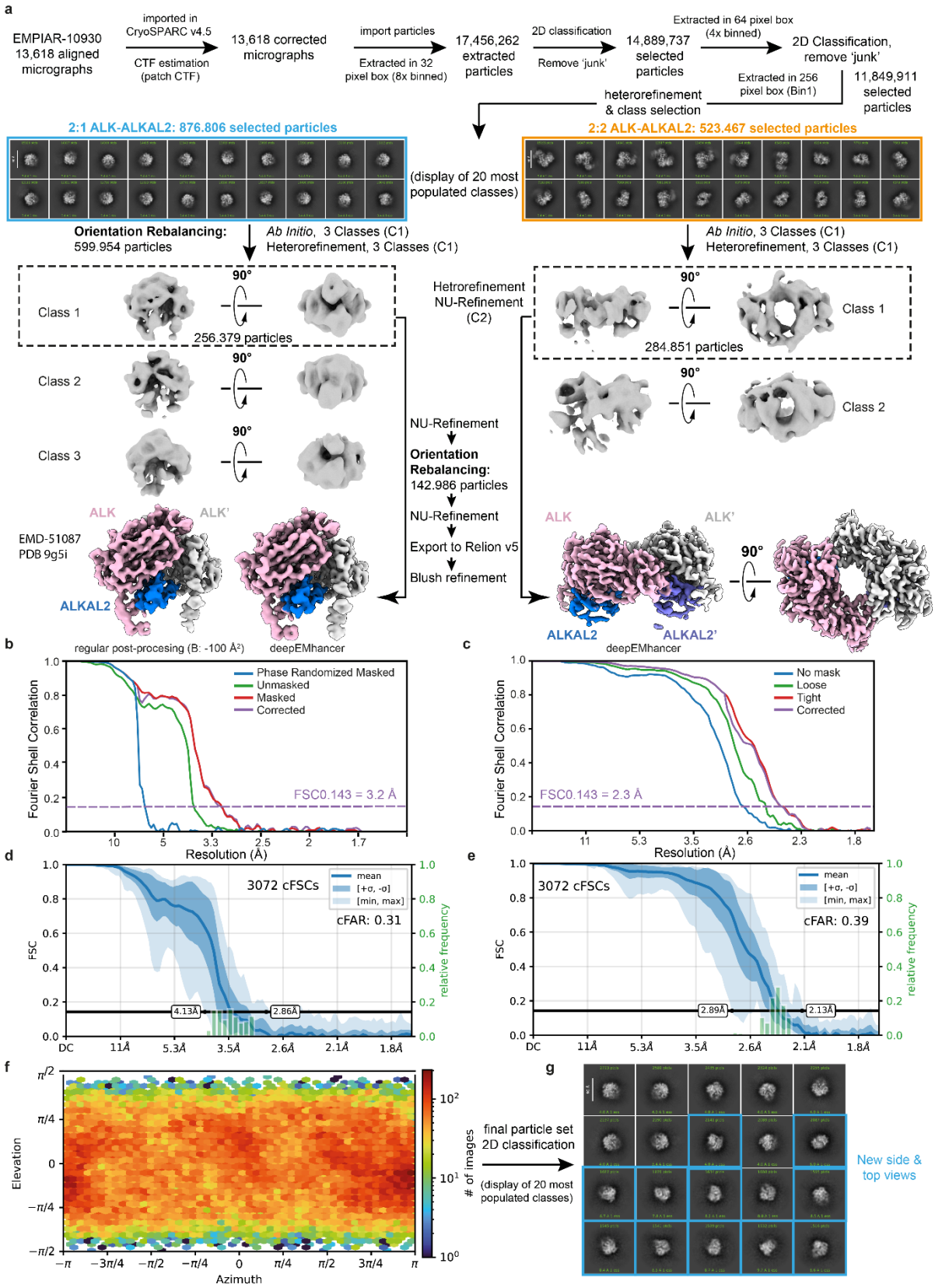


Figure 3: Side-by-side comparison of 2:1 and 2:2 ALK-ALKAL2 complexes present in the EMPIAR-10930 cryo-EM dataset. Models with pdb code 9g5i (this study) and 7n00 (Reshetnyak *et al.*, 2021) are shown as cartoons with ALK/ALK' colored pink/grey and ALKAL2 colored blue in the 2:1 ALK-ALKAL2 complex, or blue/purple in the 2:2 ALK-ALKAL2 complex. The distance between the C-termini of membrane-proximal EGFL domains is annotated.

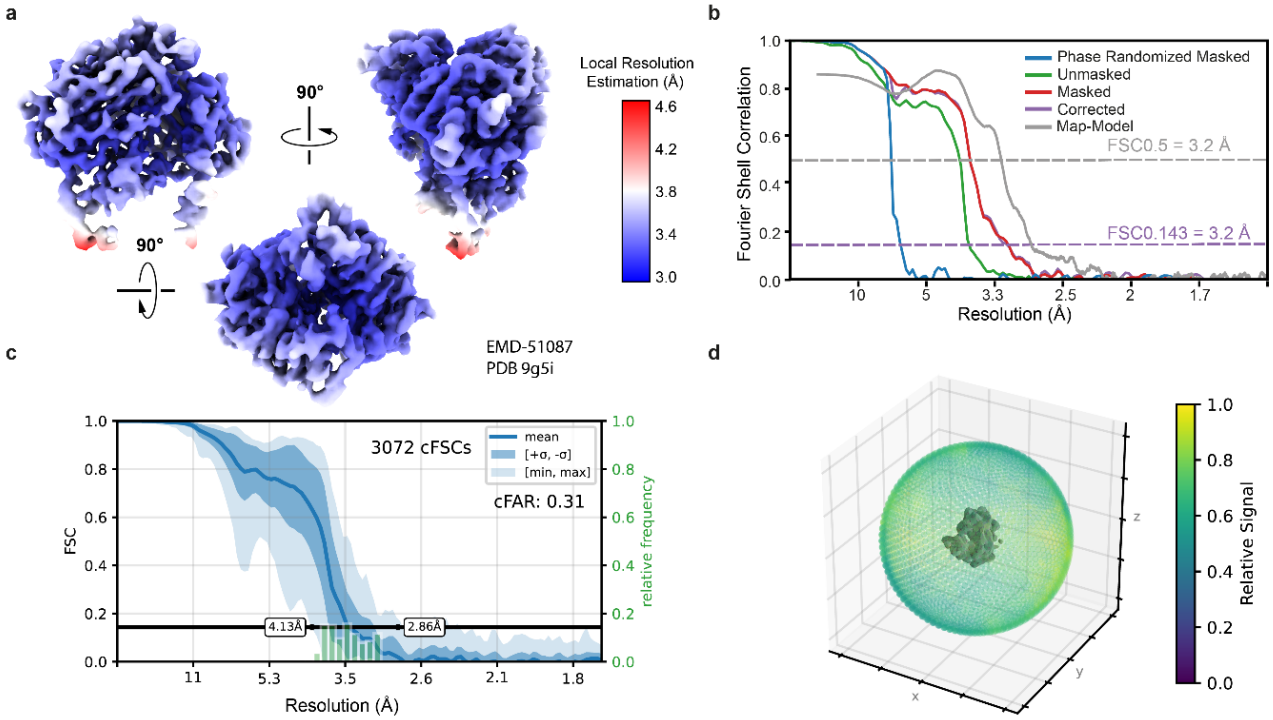


Extended Data Figure 1: (a) Initial processing workflow of cryo-EM dataset EMPIAR-10930. All initial processing steps were performed in CryoSPARC[19] v3.3.1. **(b)** Gold-standard Fourier Shell Correlation (FSC) plot corresponding to the final map obtained after non-uniform refinement (C1 symmetry). Curves are shown after applying no mask (blue), a loose mask (green), or a tight mask (red) to the two half maps before calculating FSCs. The corrected FSC (purple) is calculated using the tight mask with correction by noise substitution[41], and the resolution at FSC=0.143 is annotated via a dotted purple line. **(c)** conical FSC summary plot generated via 'Orientation Diagnostics' in cryoSPARC v4.5. **(d)** Left: Azimuth plot showing the distribution of orientations over Azimuth (X-axis) and Elevation (Y-axis) angles for the particle set corresponding to the NU-refined map shown in (A). Right: Plot showing Relative Signal amount vs. Viewing Direction.

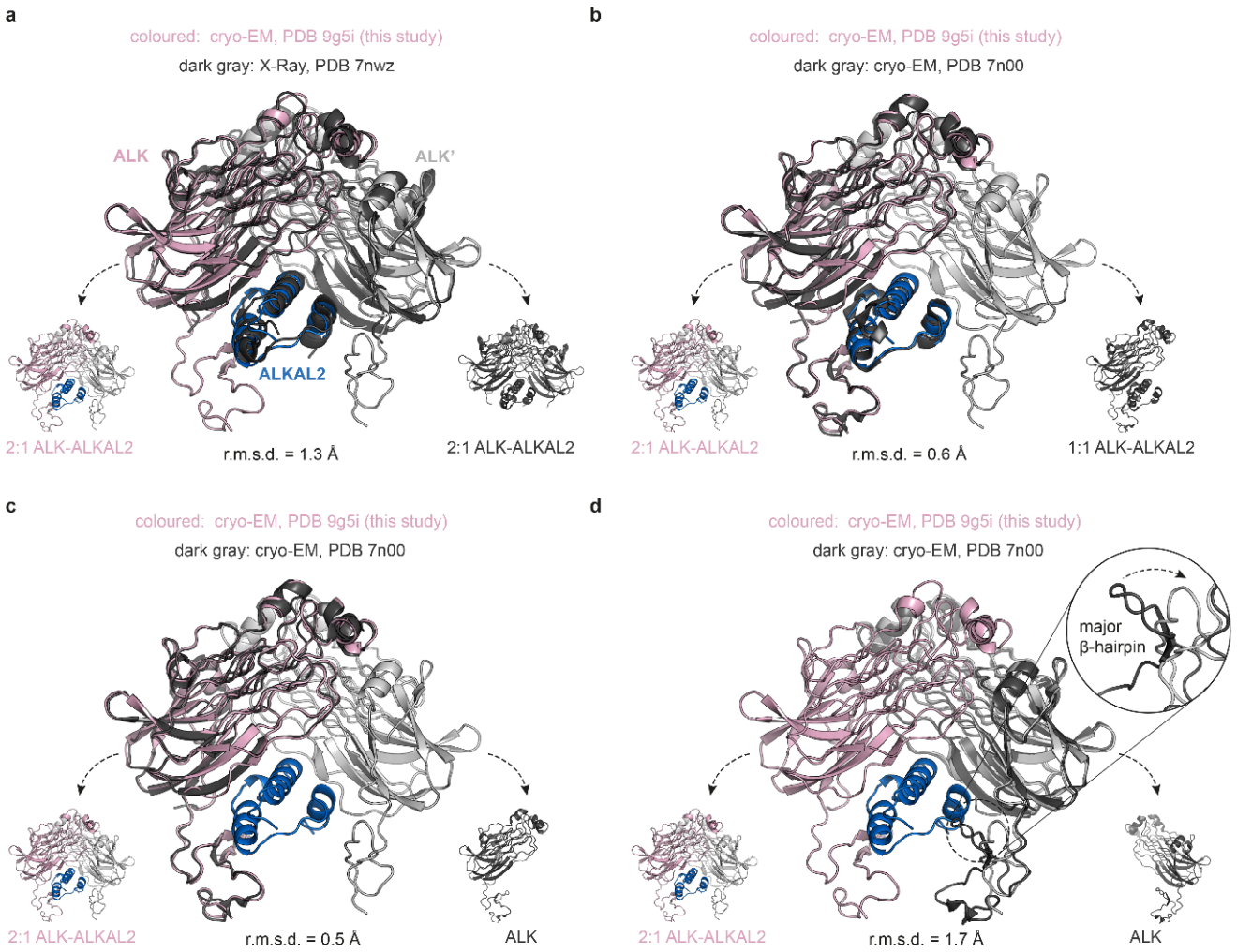


Extended Data Figure 2: **(a)** Final processing workflow of cryo-EM dataset EMPIAR-10930. Final processing steps were performed in CryoSPARC[19] v4.5 or Relion[34,35] v5. Post-processing was performed in Relion or using deepEMhancer[21]. **(b & c)** Gold-standard Fourier Shell Correlation (FSC) plots corresponding to the final maps of the 2:1 **(b)** and 2:2 **(c)** ALK:ALKAL2 complexes. The corrected FSCs (purple lines) are calculated after performing correction by noise substitution[41], and the resolution at FSC=0.143 is annotated via dotted purple lines. **(d & e)** conical FSC summary plots corresponding to the final maps of the 2:1 **(d)** and 2:2 **(e)** ALK:ALKAL2 complexes, generated via 'Orientation Diagnostics' in cryoSPARC v4.5. **(f)** Azimuth plot showing the distribution of orientations over Azimuth (X-axis) and Elevation (Y-axis) angles for the particle

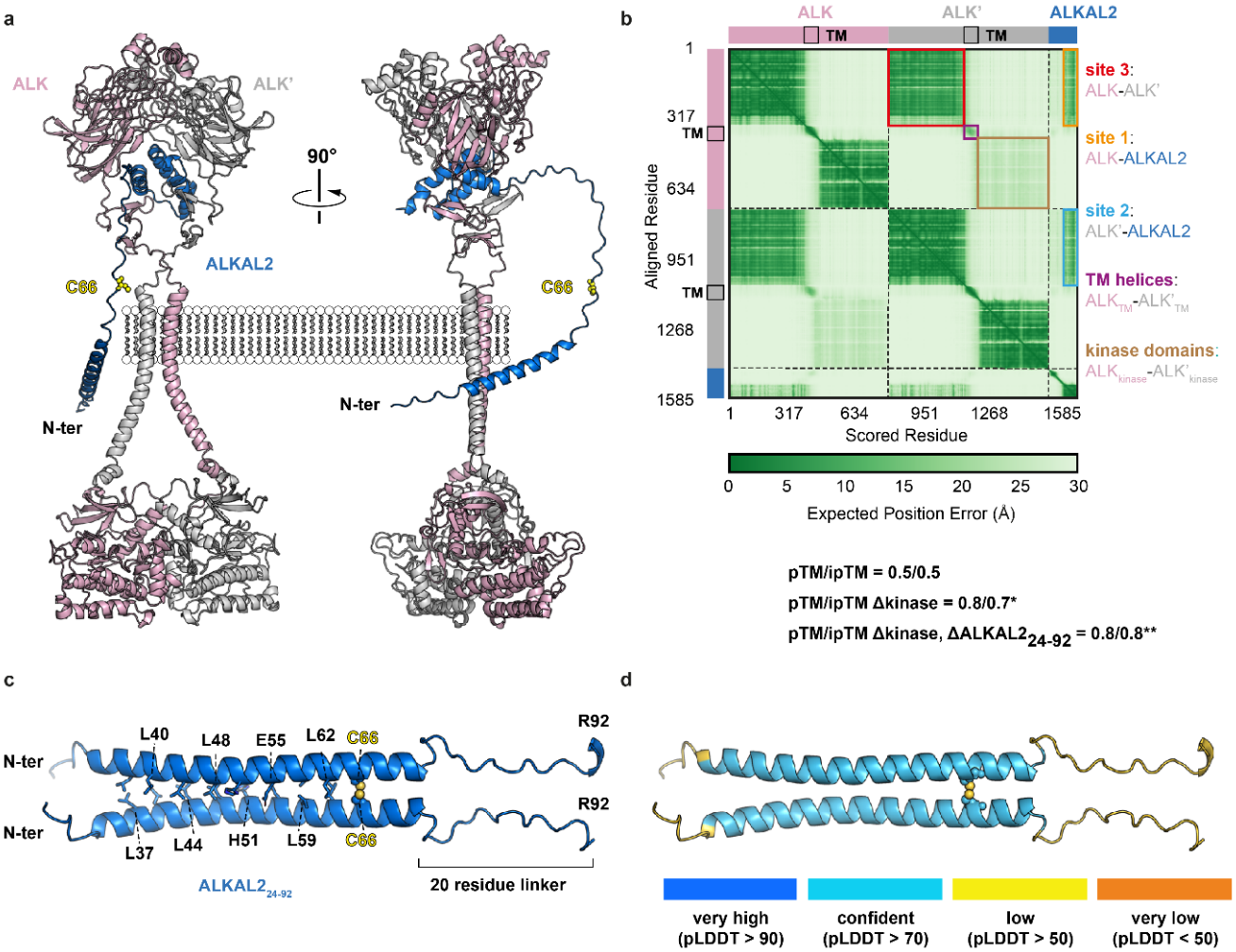
set corresponding to the final map of the 2:1 ALK-ALKAL2 complex, obtained after Orientation Rebalancing in cryoSPARC v4.5 and 3D refinement with Blush regularization[35] in Relion v5. **(f)** Display of the 20 most populated classes after 2D classification of the final particle stack corresponding to the 2:1 ALK-ALKAL2 complex. New side & top views not observed in initial 2D classification runs are shown as blue squares.



Extended Data Figure 3: (a) Local Resolution Estimation performed in Relion v5 of the final map for the 2:1 ALK-ALKAL 2 complex. **(b)** Gold-standard Fourier Shell Correlation (FSC) plot corresponding to the final map of the 2:1 ALK:ALKAL2 complex. The corrected FSC (purple lines) is calculated after performing correction by noise substitution[41], and the resolution at FSC=0.143 is annotated via a dotted purple line. A map-to model FSC curve calculated using the fitted model and an unsharpened, unfiltered full map is shown (grey line) with the resolution at FSC=0.5 annotated via a dotted grey line. **(c)** conical FSC summary plot corresponding to the final map of the 2:1 ALK:ALKAL2 complex, generated via 'Orientation Diagnostics' in cryoSPARC v4.5. **(d)** Plot showing Relative Signal amount vs. Viewing Direction.



Extended Data Figure 4: Structural alignments of ALK-ALKAL2 X-ray and cryo-EM models. (a) Structural alignment of the 2:1 ALK-ALKAL2 X-ray structure (PDB 7nwz) with the cryo-EM structure of the 2:1 ALK-ALKAL2 complex obtained in this study (PDB 9G5I). **(b)** Structural alignment of a 1:1 ALK-ALKAL2 complex, extracted from the cryo-EM structure of the 2:2 ALK-ALKAL2 complex (PDB 7n00), with the cryo-EM structure of the 2:1 ALK-ALKAL2 complex obtained in this study. **(c)** Structural alignment of one copy of ALK, extracted from the cryo-EM structure of the 2:2 ALK-ALKAL2 complex (PDB 7n00), with the first ALK molecule (pink) present in the cryo-EM structure of the 2:1 ALK-ALKAL2 complex obtained in this study. **(d)** Structural alignment of one copy of ALK, extracted from the cryo-EM structure of the 2:2 ALK-ALKAL2 complex (PDB 7n00), with the second ALK molecule (ALK', grey) present in the cryo-EM structure of the 2:1 ALK-ALKAL2 complex obtained in this study.



Extended Data Figure 5: AlphaFold 3 prediction of the full-length 2:1 ALK-ALKAL2 complex. (a) AlphaFold 3[32] prediction of a 2:1 ALK-ALKAL2 complex using the sequence of full-length ALK (excluding N-terminal HBD-MAM1-MAM2 domains and C-terminal unstructured tail) and ALKAL2 (see Methods). Side and front cartoon views are shown with ALK colored pink/grey and ALKAL2 colored blue. Cysteine 66 (C66) of ALKAL2 is shown as yellow spheres. **(b)** Predicted Aligned Error (PAE) plot of the AF3 prediction of the 2:1 ALK-ALKAL2 complex, with annotated regions in the plot corresponding to interfaces at ALK-ALKAL2 site 1 (orange), ALK-ALKAL2 site 2 (blue), ALK-ALK' site 3 (red), the TM helix-TM helix interface (purple) and the interface between dimerized kinase domains (brown). pTM and iPTM scores are shown for the 2:1 ALK-ALKAL2 prediction shown in **(a)** as well as a prediction with omitted ALK kinase domains (annotated with *) and omitted ALK kinase domains and N-terminal residues 24-92 of ALKAL2 (annotated with **). **(c)** Alphafold 3 prediction of a dimer of N-terminal ALKAL2 residues 24 – 92. Interacting residues are annotated as sticks, with cysteine 66 (C66) shown as yellow spheres. **(d)** Alphafold 3 prediction shown in **(c)**, colored corresponding to the per-residue model confidence (pLDDT: predicted local distance difference test) score.