Title

Structural insights into substrate recognition by the SOCS2 E3 ubiquitin ligase

Authors

Wei-Wei Kung*, Sarath Ramachandran*, Nikolai Makukhin*, Elvira Bruno, Alessio Ciulli

*These authors contributed equally to this work

Address

Division of Biological Chemistry and Drug Discovery, School of Life Sciences, University of Dundee, James Black Centre, Dow Street, Dundee, DD1 5EH, United Kingdom

Abstract

The suppressor of cytokine signaling 2 (SOCS2) acts as substrate recognition subunit of a Cullin5 E3 ubiquitin ligase complex. SOCS2 binds to phosphotyrosine-modified epitopes as degrons for ubiquitination and proteasomal degradation, yet the molecular basis of substrate recognition has remained elusive. We solved co-crystal structures of SOCS2-ElonginB-ElonginC in complex with phosphorylated peptides from substrates growth hormone receptor (GHR-pY595) and erythropoietin receptor (EpoR-pY426) at 1.98 Å and 2.69 Å, respectively. Both peptides bind in an extended conformation recapitulating the canonical SH2 domain-pY pose, yet capture different conformations of the EF loop via specific hydrophobic interactions. The flexible BG loop, for the first time fully defined in the electron density, does not contact the substrate degrons directly. Cancer-associated SNPs located around the pY pocket weaken substrate-binding affinity in biophysical assays. Our findings reveal insights into substrate recognition and specificity by SOCS2, and provide a blueprint for small molecule ligand design.

Correspondence and requests for materials should be addressed to A.C. (a.ciulli@dundee.ac.uk)

ORCID ID: 0000-0002-8654-1670

Introduction

Cytokines are small glycoproteins that play important roles in the differentiation, development and function of lymphoid and myeloid cells [1]. The Janus kinase (JAK) – signal transducer and activator of transcription (STAT) signaling pathway plays a critical role enabling cells to respond to specific cytokines by regulating gene expression. Suppressor of cytokine signalling (SOCS) proteins, which comprise of cytokine inducible SH2–containing protein (CIS) and SOCS1 – SOCS7, negatively regulate cytokine receptors and inhibit the JAK-STAT signaling pathway [2].

SOCS proteins share a conserved domain architecture comprising of an N-terminal extended SH2 subdomain (ESS) that functions as a bridge at the interface between the SH2 domain and the SOCS box enabling ubiquitination of captured substrate[3–5], followed by a central Src-homology 2 (SH2) domain that recognizes a phosphotyrosine (pY) containing sequence [6], and a C-terminal SOCS box that interacts with the adaptor ElonginB-ElonginC complex (EloBC) [7–9]. All SOCS proteins bind to EloBC and recruit Cullin5 with high specificity, forming different SOCS-EloBC-Cullin5-Rbx2 ($CRL5^{SOCS}$) E3 ligases that catalyse ubiquitin transfer and subsequent proteasomal degradation of specific substrates, as a mechanism to regulate diverse biological processes [10–13]. SOCS proteins serve as substrate recognition modules that impart substrate specificity to each $CRL5^{SOCS}$ E3 complex.

Expression of SOCS proteins is induced by cytokine stimulation. Upon cytokine binding, the oligomerized receptors activate the JAK family kinases that phosphorylate specific tyrosine residues on the receptor, including the docking sites for the STAT proteins. The docked STAT proteins are sequentially phosphorylated, they dimerize and translocate into the nucleus, initiating gene transcription of several downstream proteins including the SOCS proteins. SOCS proteins suppress the JAK-STAT pathway via three distinct but often concomitant mechanisms: 1) KIR mediated direct JAK inhibition [3,14]; 2) Blocking STAT activation by competing for receptor pY sites [15]; 3) Targeting the receptor for proteasomal degradation via SOCS E3 ligase activity [16,17]. Some of the SOCS-substrate interactions have been structurally characterized, including SOCS1-JAK [18], SOCS3-gp130 [4,19], SOCS3-gp130-JAK2 [20] and SOCS6-cKit [21].

SOCS2, one of the members of the SOCS family, is implicated in the disorders in immune system, central nervous system and cancer, and is thus emerging as a promising target for

cancer therapies [22–25]. SOCS2 has been shown as the primary suppressor of growth hormone (GH) pathway where a gigantism phenotype was observed in a *SOCS2*$^{-/-}$mice [26]. Paradoxically, the SOCS2 overexpressed transgenic mice also led to the same phenotype [27]. Attenuation of GHR signalling relies on two phosphorylation sites at GHR that are recognised by SOCS2 [16,28]. The pY487 site of GHR interacts with CRL5^{SOCS2} E3 ligase that targets the GHR for ubiquitination and proteasomal degradation [16]. A downstream pY595 site interacts with SOCS2, STAT5b and SHP2 (SH2 domain-containing phosphatase 2), enabling SOCS2 to inhibit the signaling by blocking this receptor site from STAT5b [15,27,29,30]. Nonetheless, deletion of both sites is required to remove the inhibitory effect of SOCS2 on the GH signaling [16,28]. Analysis of the binding affinity of SOCS2 for these two phosphorylation sites of GHR reveals that the pY595 region exhibits a higher affinity towards SOCS2 (K_D = 1.6 μM) compared to pY487 region (K_D = 11.3 μM) [5,31,32]. An 11-mer phosphorylated peptide spanning the pY595 region of GHR was sufficient to pull down the whole CRL5^{SOCS2} complex from human cell lysates[31] as well as CIS, the closest homologue to SOCS2 from the same family, which plays a role in anti-tumor immunity controlling the differentiation of CD4 T helper cell, and the IL-2 and IL-4 response [31,33]. In addition to GHR and CIS, other substrates have been identified to interact with SOCS2, including the erythropoietin receptor (EpoR) at pY426 [34], the leptin receptor at pY1077 [35], the epidermal growth factor receptor [36] and the insulin-like growth factor-I receptor [37]. The first crystal structure of SOCS2-ElonginB-ElonginC (SBC) was reported in 2006 [32], however the structural basis for substrate recognition to SOCS2 has yet remained elusive.

Here, we determine the first co-crystal structures of SBC in complex with phosphorylated epitope peptides from its physiological targets GHR and EpoR. Our structures reveal the peptides are accommodated in an extended conformation to capture specific interactions with SOCS2. A key flexible region of SOCS2, known as the BG loop, is defined in the electron density for the first time and shown not to contact the bound substrates. Structural analyses supported by biophysical and mutagenesis investigations allowed identification of hotspot residues on the substrate degrons and functional elucidation of disease-relevant single nucleotide polymorphisms (SNPs) of SOCS2. Our findings provide fresh insights into the molecular recognition and selectivity between SOCS2 and target substrates, and provide an important template for future structure-guided ligand design.

Results

To elucidate the molecular basis of substrate recognition by SOCS2, we subjected the SOCS2-ElonginB-ElonginC (SBC) complex to extensive co-crystallization trials with 11-residue phosphopeptides of either EpoR or GHR that span the regions surrounding Tyr426 and Tyr 595 region, respectively. The affinity of SBC for EpoR_pY426 (K_D of 6.9 μM) and GHR_pY595 (K_D of 1.1 μM) was measured by isothermal titration calorimetry (ITC), and found to be consistent with the literature [31,32] (Figure S1). Attempts to co-crystallize wild-type SBC protein constructs [31,32] with either GHR or EpoR peptide were unsuccessful as resulting crystals only diffracted poorly. To improve crystal quality, we engineered a cluster of three mutations K115A/K117A/Q118A on SOCS2 that was predicted to significantly reduce surface conformational entropy and thermodynamically favor crystal packing [38]. Crystallization attempts with this new $S^{KKQ}BC$ triple-mutant construct (K115A/K117A/Q118A on SOCS2) eventually yielded high-resolution datasets.

SBC-EpoR co-crystal structure

The structure of SBC in complex with EpoR_pY426 peptide (SBC-EpoR) was solved and refined at 2.69 Å with 19.64 % R_{work} and 23.51 % R_{free} (Table 1). The overall subunit and domain arrangements of the SBC-EpoR structure is consistent with those of the apo SBC structures [32,39,40] (Figure 1a). Electron density for nine out of eleven non-terminal EpoR_pY426 residues are well defined in the structure (Figure 1b). A classic SH2 domain-pY peptide interaction is observed where the pY residue is anchored at the pY pocket and the flanking residues are extending across the SH2 domain. The pY residue is tightly locked by an intricate hydrogen-bonding network formed by residues Arg73, Ser75, Ser76, Thr83 and Arg96 of SOCS2 (Figure 1c). Additional hydrogen bonds are formed along the backbone of EpoR_pY426 peptide from Glu(-1) to Leu(+3) with SOCS2 residues Thr93, Asn94, Asp107 and one structural water (Figure 1d). Multiple hydrophobic interactions also support the binding of EpoR_pY426 C-terminal residues, Ile(+2), Leu(+3) and Pro(+5) that are well accommodated within a hydrophobic patch created by Leu95, Leu106, Ser108, Ile109, Val112, Leu116 and Leu150 of SOCS2 (Figure 1e).

Crystal structure of SBC-GHR

Encouraged by the success in solving an SBC-EpoR structure, to deepen understanding of the SOCS2 binding epitopes, we co-crystallized SBC with an 11-mer GHR_pY595 peptide (SBC-GHR) and solved the structure at 1.98 Å resolution with 19.00 % R_{work} and 22.66% R_{free} (Table 1). In contrast to SBC-EpoR, which contains one protomer in the asymmetric unit, the SBC-GHR contains two copies of protomer. Alignment of these two

protomers via the backbone atoms of the EloB subunit reveals a hinge motion between the SH2 domain and the SOCS box (Figure S2). Such motion is a common feature in SOCS box and F-box containing proteins and it has been shown to be important as it facilitates accurate orientation and positioning of a target substrate protein relative to the multisubunit CRL complex [39,41]. Contrary to the SBC-EpoR complex, which contains a single copy of peptide per SH2 domain, two copies of GHR_pY595 peptides were found binding per SH2 domain of SOCS2 (giving a total of four copies within the asymmetric unit). The two peptides run in an anti-parallel direction relative to each other across the SH2 domain, with well-defined electron density surrounding them both (Figure 2a). One of the peptides (referred to as peptide A hereafter) binds to SH2 domain in a canonical manner, where the pY is recognized by the positively charged pY pocket between the central β strands and αA (Figure 2b). In contrast, the second peptide, peptide B, has its pY residue exposed to solvent and interacting only with His149 of SOCS2 (Figure 2b).

Specific interaction of the GHR_pY595 phosphopeptide

The unusual simultaneous binding observed for the GHR substrate peptide to SOCS2 SH2 domain is imparted mainly by the region comprising Ser(+2) to Val(+6) from each peptide, which pair such that they form an anti-parallel beta sheet (Figure 2c). Extensive hydrogen bonds are formed between the backbone of the two peptides and backbone residues of SOCS2 and structural waters (Figure 2c). Further hydrophobic interactions appear to reinforce the binding, which impart specificity for GHR. The Ile(+3) and Ile(+5) of peptide A and peptide B settle in a hydrophobic patch of the SH2 domain formed by Leu95, Leu106, Ser108, Leu116 and Leu150 (Figure 2d). Another hydrophobic interaction that is distinct in SBC-GHR compared to SBC-EpoR is that formed by the side chain of Val(-3) of peptide A, that nicely fits into a hydrophobic pocket comprising of Thr88, Ala90, Thr93, Leu95 and Val148 from SOCS2 (Figure 2e). In the SBC-GHR structure, a cobalt ion is modeled at a positive peak that disappeared only at 21 σ level in the unbiased Fo-Fc electron density map. This cobalt ion satisfies the formation of an octahedral coordination geometry with the side chains of His(+4) of peptide B, His149 of SOCS2 and with three surrounding water molecules (Figure 2f).

The BG loop of SOCS2 is observed for the first time in an open conformation

SOCS2 is known to recognize two GHR binding sites at regions around pY487 and pY595, respectively. We therefore hypothesized that the two copies of the GHR peptides bound in the crystal structure might mimic a physiological folded conformation of GHR, presenting

each of the phosphorylated sites bound simultaneously to SOCS2. To test this hypothesis, we utilized an 11-residue GHR_pY487 phosphopeptide (NIDFpYAQVSDI, K_D of 2.3 μM by ITC, Figure S1), mixed with the GHR_pY595 peptide and SBC in equimolar 1:1:1 ratio for co-crystallization. In this crystal structure (hereafter referred to as SBC-GHR_2), still two copies of the GHR_pY595 peptide, but no GHR_pY487, are observed bound, yielding a structure very similar to the previous SBC-GHR structure (Figure S3). However, an important observation in SBC-GHR_2 compared to our other co-crystal structures was that the regions of SOCS2 corresponding to residues 134-162 is now fully visible in the electron density. This is unprecedented, and the first time this region, also called BG loop is structurally defined in full. The BG loop connects the βB and βG strands of an SH2 domain (Figure 3a). The first part of the BG loop (residues 134-148 in SOCS2) differs in length and sequence among SOCS proteins (Figure 3b). We refer herein to this more variable region as the "specificity BG loop", because its conformation, together with the configuration of the adjacent EF loop, governs accessibility of the pY binding pocket and contributes to substrate specificity in SH2 domains [42]. A particular region in the middle of the specificity BG loop (residues 136-145) is found to be disordered in all previously determined SOCS2 structures (PBD code, 2C9W, 4JGH and 5BO4) as well as our other co-crystal structures SBC-GHR and SBC-EpoR. In this SBC-GHR_2 structure, the BG loop is in an open conformation stabilised by crystal contacts, Pro140 (BG loop) to Arg186 (SOCS box) and Pro140 (BG loop) to Ile90 (EloB), as clearly defined by the unbiased omit map at this region (Figure 3c, Table 1).

Conformational changes of the EF and BG loop

The configuration of the EF and BG loops play an important role in governing the accessibility of the binding pocket and specificity toward ligand binding [42]. A comparison of the SOCS2 structures in the presence and absence of peptides bound highlight conformational changes in EF (residue 107-116) and BG loop. In the absence of substrate peptide, the EF loop curls up placing the Ile110 and Cys111 at the hydrophobic SH2 domain (Figure 4a). Upon binding of a substrate peptide, the EF loop opens up forming backbone interactions with GHR_pY595 (Figure 4b), or rearranges itself to allow a specific interaction with EpoR_pY426 (Figure 4c). This specific interaction between EF loop and EpoR_pY426 involves hydrophobic interactions between Ile109 and Val112 of SOCS2, Val112 of a SOCS2 symmetry mate, and Pro(+5) of EpoR_pY426, resulting in a differential binding mode from GHR_pY595 to SOCS2. The BG loop of SOCS2 is observed to be in an open conformation in SBC-GHR_2 and SBC-EpoR (disordered)

structures. A superimposition of the two substrate complex structures suggest that the BG loop opens up further in SBC-GHR structure to accommodate two GHR peptides (Figure 4d).

Biophysical characterization of specificity between GHR_pY595 and SOCS2

To evaluate the specificity of the protein-peptide interaction in solution, we designed single-point mutations on the peptide first, and compared their binding to wild-type peptide by two orthogonal biophysical methods: a direct binding assay using SPR (SBC immobilized on the chip) and ^{19}F ligand-observed displacement NMR assays. In the ^{19}F NMR displacement assay, the fluorine signal of a purposely-designed reporter ligand (also referred to as spy molecule) was monitored as a mean to quantify the extent of the competition between the tested peptides and the spy molecule. A Carr-Purcell-Meibom-Gill (CPMG) pulse sequence was applied to estimate the spin-spin relaxation time (T_2) of the spy molecule in the absence and presence of protein [43–45]. By adding competitor to disrupt the protein-spy interaction, the binding affinity of a competitor can be calculated based on the degree of displacement of the spy[46,47]. The spy molecule used in our assay is phosphate **3**, a fluorinated pY analogue that specifically binds to the pY pocket with a K_D of 50 μM (Scheme 1, and Figure S4). The two assays were found to be robust and reliable where the measured K_D and K_i values at R^2 of 0.74 correlation (Figure S5).

First, we focused on the unique interaction formed by Val(-3) of GHR_pY595, which inserts into a small hydrophobic cavity of SOCS2 (Figure 2e). This interaction was investigated by mutating Val(-3) in the GHR peptide to Tyr and Arg, as representative bulky and charged residues, respectively. We hypothesized this structural change would disrupt the fit at this small hydrophobic pocket. Mutant V(-3)R exhibited between a 6- and a 14-fold loss of binding affinity to SBC, depending on the assay, suggesting the charged group strongly disrupts the interaction (Table 2). By contrast, the V(-3)Y was less disruptive, with only a two-fold loss in affinity.

Next, to map the relative importance and contribution of each individual amino acids to the binding affinity with SBC, alanine scan of the substrate peptides was invoked. The peptide sequences were designed such that individual amino acids were separately mutated into alanine except pY, which is known to abolish binding if mutated even to unphosphorylated Y [31]. The resulting library comprised of the original wild-type sequences, ten derivatives from GHR_pY595 and nine from EpoR_pY426, and was characterized in parallel using

SPR and ^{19}F NMR competition assay (Table 3). Alanine substitution at pY(-3), pY(-1), pY(+3) and pY(+4) of the GHR_pY595 resulted in at least two-fold weakened binding (increase in K_D) compared to the wild type. In contrast, a similar 2-fold weakening in binding affinity was observed in the EpoR_pY426 peptide upon alanine substitution at pY(-1), pY(+2) and pY(+3). These results are consistent with observations from our crystal structures that peptide-SOCS2 binding is mediated by hydrophobic interaction including pY(-3), pY(+3) and pY(+5) on the GHR_pY595 and pY(+2) and pY(+3) on the EpoR_pY426. The binding affinity for each peptide also dropped by at least two-fold with Ala substitution at pY(-1) position, indicating the importance of the residue just upstream of pY.

SNPs study

Several single nucleotide polymorphisms (SNPs) on SOCS2 are reported in the Catalogue of Somatic Mutation in Cancer database (COSMIC) as potentially linked to cancers such as tumours of the lung, breast, and pancreas [48]. We thus next decided to characterize the interaction of selected SNP SOCS2 mutants with substrate peptides GHR_pY595, EpoR_pY426 and GHR_pY487 by SPR. Inspection of our SBC co-crystal structures guided us to select five known SNPs: N94D, R96L, R96Q that are located in the pY-pocket and involved in direct recognition of pY; L106V that is highly conserved at the hydrophobic SH2 domain involving in substrate interaction; and C133Y that participates in the SH2 hydrophobic core (Figure 5a). All mutant proteins expressed and purified similarly to wild-type, and the mutations did not appear to affect the structural integrity and solubility of the constructs, as observed by ^{1}H NMR (Figure S6).

The L106V and C133Y mutations did not affect binding affinities compared to wild type (Figure 5b). In contrast, the N94D and R96L mutations drastically impaired substrate binding, leading to almost undetectable binding response by SPR. Because of the low signal-to-noise, reliable K_D values could not be measured with these protein mutants. For the R96Q mutation no signal response was detected, suggesting that binding was completely abolished and highlighting the most disruptive of the mutation studies. Mutations at R96 site abrogate SOCS2 binding was also reported by Rupp *et al.* with R96C [49]. In addition to SPR, the SNP mutants were characterized using ^{19}F NMR by monitoring the transverse relaxation rate (R_2). L106V and C133Y exhibited similar rate, 10 s^{-1} and 14 s^{-1}, comparable to wild type at 11 s^{-1}, suggesting these mutations did not affect pY recognition. Mutations on N94 and R96 showed a substantial disruption to pY

recognition, resulting in a significant disruption of binding as the no protein control at around 3 to 4 s^{-1}, consistent with the SPR data (Figure 5c).

Discussion

SOCS2 is a substrate binding protein of the CRL5 E3 complex that negatively regulates the JAK-STAT signaling by targeting substrate receptors for degradation and blocking STAT5b activation by competing with receptor pY sites. The details of these interactions have remained elusive and to date structural information remained limited to apo SOCS2. Herein, we have disclosed two novel structures of the SBC in complex with substrate peptides EpoR_pY426 and GHR_pY595. Both peptides recapitulate a canonical substrate-binding mode to SH2 domain of SOCS2. Residues at pY(-1), pY(+1) and pY (+3) positions of the GHR_pY595 and EpoR_pY426 peptide contains similar properties; whereas residues at pY(-3), pY(+2) and pY(+5) position are different in properties and sizes (Figure 6a). In particular, the Val(-3) of GHR_pY595 and Pro(+5) of EpoR_pY426 peptide catch different hydrophobic interactions resulting in exclusive binding modes in SOCS2 comparing with the substrate peptides of SOCS3 and SOCS6 (Figure 6b, 6c) [4,21].

The BG loop of SOCS2 had not been fully revealed in previous published structures. Here, we report an open conformation BG loop, which is contradictory to other SOCS structures with bound peptides, for example SOCS3:gp130 and SOCS6:c-kit [19,21]. In the SOCS3 and SOCS6 peptide-bound structures, the BG loop folds up as a hairpin interacting with the substrate peptide, forming a triple-stranded β sheet structure (Figure 7a, 7b). The corresponding BG loop region in SOCS2 is either fully disordered or in an open conformation (SBC-EpoR and SBC-GHR_2), suggesting that this region does not participate in substrate recognition. Nevertheless, interestingly, a similar triple-stranded β sheet structure is observed in the SBC-GHR_2 structure, where the peptide B replaces the first β-sheet of the BG loop and makes backbone interactions with BG loop and peptide A (Figure 7c). Additionally, a tyrosine phosphatase SHP-2, which contains SH2 domains, features a comparable structure [50]. Its BG loop folds up as a hairpin and forms a triple-stranded β sheet interaction with two bound peptides (Figure 7d). Although the canonical phosphotyrosine binding site is conserved for both SOCS2 and SHP2, the positioning of the non-canonical phosphotyrosine varies significantly. The distance between Cα atoms of phosphotyrosines within the bound peptides are 23 Å and 13 Å in SBC-GHR and SHP2-substrate co-structures, respectively (Figure 7e). Longer distance between the canonical and non-canonical pY enable antiparallel beta strand interactions between four amino acids of GHR peptides unlike antiparallel interactions between two amino acids in SHP-2 substrate peptides. Furthermore, a longer EF loop in SOCS2 pushes the two GHR peptide, away from itself towards the open BG loop, enabling β sheet formation.

The BG loop along with the EF loop forms a hydrophobic channel in SOCS3 and SOCS6. This channel imparts specificity and restricts the binding of substrates. In contrast the open conformation of BG loop in SOCS2 could be critical in enabling SOCS2 to accommodate a wider range of substrates including GHR, EpoR, SOCS1 and SOCS3 amongst others. A comparison of buried surface area of the substrate peptides among SOCS proteins, reveal that EpoR and GHR binds with SOCS2 with only 595 $Å^2$ and 641 $Å^2$, respectively, in contrast to areas of 1714 $Å^2$ for SOCS6/c-KIT and 1761 $Å^2$ for SOCS3/gp130 complexes. Unlike SOCS3 and SOCS6 complex structures, the pY flanking residues from EpoR and GHR do not participate in extensive side-chain hydrogen bonding interactions. Together, these observations are consistent with greater binding affinities of SOCS3 and SOCS6 substrates compared to SOCS2 substrates. The lower potency of affinity for SOCS2 substrates could contribute to its relatively greater promiscuity to multiple substrates.

The observation of the dual-peptide binding mode to SOCS2 was unexpected, however is not unprecedented with SH2 domains, as reported previously with the tyrosine phosphatase SHP-2 [50]. In the co-crystal structure of SHP-2:pY peptide solved by Zhang et al., one pY of the peptide is recognized at the pY pocket and the other one is solvent exposed as in our structures described herein. Besides, two peptides run antiparallel to each other and form an antiparallel four-stranded β sheet with BG loop. Zhang *et al*. suggested that the dimerization of peptide binding in SHP-2 requires at least one pY containing peptide and demonstrates enhanced binding affinity to protein. For our SBC-GHR structure, despite preparing several protein-peptide samples for co-crystallization at 1:1 molar ratio, all dataset collected from crystals were consistent with a 1:2 (protein-peptide) binding mode. In both cases, the binding of the non-canonical binding peptide relies on hydrophobic interaction with the SH2 domain hydrophobic core, and backbone to backbone interaction with the canonical binding peptide and protein (Figure 6d). The 1:1 binding mode for EpoR could be justified by the presence of Pro(+5) in EpoR, that acts as a strand breaker and prevents backbone to backbone interaction with the second peptide. Indeed, from the alanine scan (Table 3), we observe that when Pro(+5) is replaced by alanine, the affinity improves - implying the possibility of secondary peptide binding.

We put forth two distinct models that might explain the dual peptide recognition mode and its role in specific tuning of GHR signaling response. First, a “cis” recognition mode, where the GHR tail folds back as a hairpin structure presenting two binding epitopes around

distinct phosphorylation sites (e.g. pY487 and pY595) for recognition (Figure 8a). However, the crystallography data from our follow-up experiment as described in the SBC-GHR_2 structure is not consistent with this hypothesis, as two instances of the pY595 peptide were found bound despite a molar ratio of 1:1 for pY487 and pY595 peptides being present in the co-crystallization buffer. However, we cannot exclude that simultaneous recognition of the two distinct epitopes would require a loop of the same tail twisting back onto itself to enhance the binding affinity of second epitope. Alternatively, we envisage a "trans" recognition mode, where SOCS2 recognize two separate receptor tails of the activated dimerized GHR receptors at the cell membrane (Figure 8b). SOCS2 might additionally play a role as scaffold bringing two substrates in close proximity, for example by recruiting one instance of phosphorylated substrate to assist the binding of un-phosphorylated one for post-translational modification. This mechanism evokes potential similarities with some phosphodegrons which require two sites to be phosphorylated, utilizing a first kinase to "prime" phosphorylation events, followed by a second kinase for follow-on phosphorylation [51]. An example of such a mechanism is the β-catenin degradation mediated by the β-TrCP [52]. Further biophysical investigation is warranted to address the extent to which these potential mechanisms might be invoked for SOCS2 function.

SOCS2 is an attractive therapeutic target due to its links to cancer, diabetes, neurological and inflammatory diseases [23,24,53–57]. Breast, lung, liver and ovarian cancer have been correlated with down-regulation in SOCS2 [58–63]. In addition to the JAK-STAT pathway, a recent study has identified the involvement of SOCS2 in the NF-κB (nuclear factor kappa-light-chain-enhancer of activated B cells) pathway that regulates the immune and inflammatory responses [64,65]. NF-κB is found to be constitutively activated in many types of cancer and influences a diverse array of pro-tumorigenic functions, therefore NK-κB plays an pivotal role in cancer initiation and progression [66]. SOCS2 negatively regulates TNFα induced NF-κB activation by targeting NDR1, a serine-threonine kinase, for proteasomal degradation. Hence SOCS2 deficiency may lead to an increased level of NDR1, which was is reported to result in aggressive behaviour of PC3 prostate cancer cells [64]. These evidences highlight the potential in targeting SOCS2 for drug discovery for inflammation and cancer biology. We have revealed structural insights into the SOCS2-peptide interactions by X-ray crystallography and identified hotspot using alanine scanning, mutation study and SNPs study. This information provides a template to guide the structure-based rational design of SOCS2 ligands that are instrumental in the development

of novel chemical tools to address biological question of SOCS2, and in the quest for novel small molecule ligands binding to SOCS2 as potential therapeutics. SOCS2 binders at the pY binding pocket can be used as inhibitors of the $CRL5^{SOCS2}$, which would be expected to prevent degradation of target substrate receptors, thus prolonging the activity of cytokine signaling pathway and upregulating expression of endogenous STAT5b-responsive gene expression. In a distinct application, a SOCS2 binder could be used as E3 ligase ligand handle for designing new chemical degraders to hijack SOCS2 CRL activity and trigger the degradation of unwanted proteins inside cell [67–69]. This approach, also known as proteolysis targeting chimeras (PROTACs), offers the advantage of inducing rapid and selective intracellular depletion of the target protein, as opposed to mere blockade of a single interaction or activity, which pairs more closely to genetic target validation and often results in greater maximal efficacy of intervention in a signaling pathway. PROTAC-mediated protein degradation has been shown to occur at very low compound concentration (pM to nM range) also potentially allows targeting of intractable protein targets that are beyond the reach of conventional small-molecule approaches that require full occupancy of a target binding site e.g. receptor antagonists and enzyme inhibitors. A limited set of E3 ligases have been targeted so far for PROTACs, notably VHL [70,71] and cereblon [72,73], so extending the approach to other ligandable E3 ligases would be an important advance to the field. The usage of SOCS2 binders for PROTACs may be limited to SOCS2 expressing cells, however, this provides an additional layer of tissue specific degradation of target proteins. PROTACs designed using SOCS2 binders as recruiters would be ideal for degrading targets specifically in cells affected with leukemia and gastrointestinal sarcoma which have been recorded to have upregulated SOCS2 expression [87]. Low levels of SOCS2 in other normal tissues would help minimize toxicity. In addition, a pan-SOCS recruiting PROTACs would help diversifying the target tissue range. Our peptide-bound co-crystal structures suggest that SOCS2 might be ligandable and provide a blueprint for the rational structure-guided design of novel SOCS2 inhibitors and SOCS2 ligand handles for PROTACs.

Materials and Methods

Cloning and protein expression

The human SOCS2 (amino acids 32 – 198) and the ElonginB (amino acids 1 – 104) and ElonginC (amino acids 17 – 112) plasmids were used for protein expression as previously reported [39], and as templates for mutagenesis. SOCS2 mutants with N94D, R96L, R96Q, L106V or C133Y mutation were introduced using a PCR-based method site-directed mutagenesis. SOCS2 wild type and mutants were co-expressed and purified as previously described [31,39].

Crystallization and structure determination of SBC-GHR

To improve crystallization, surface entropy reducing mutations were introduced into SOCS2 construct (amino acids 32 – 198). Three mutation clusters (K63A/E64A/E67A; K113A and K115A/K117A/Q118A) were identified with the SER server [38]. SER-assisted crystallization attempts yielded crystals with the K115A/K117A/Q118A SOCS2-EloBC (S^{KKQ}BC). Five times molar excess of GHR_pY595 (PVPDpYTSIHIV-amide) was incubated with S^{KKQ}BC, followed by removing unbound peptide using a protein concentrator. Sample was concentrated to 22 mg/ml with an additional 0.1 M of sodium cacodylate pH7.2 added to the sample. Diffraction-quality crystals were obtained with 0.005 M Cobalt (II) chloride, 0.1 M MES pH 6.5, 1.0 M ammonium sulphate at 4 °C using hanging drop vapour diffusion method at 2:1 protein:precipitant ratio. Crystals were cryo-protected using 20% MPD prior to vitrification in liquid nitrogen.

Diffraction data were collected at 100 K at Diamond Light Source beamline i04 using Pilatus 6M-F detector at 0.98 Å wavelength. Indexing and integration was processed by XDS [74] and scaling and merging with AIMLESS within the CCP4 program suite [75,76]. The experimental phases was obtained by identifying positions of arsenic atom using MR-SAD Phases in the PHENIX software suite [77,78] with a model provided. The provided template was a lower resolution SBC-GHR co-crystal obtained previously (unpublished work). The structure was reconstructed by AutoBuild [79,80] and manually built in Coot [81]. The resulting structure was refined iteratively with REFMAC5 [82].

Crystallization and structure determination of SBC-EpoR

Five times molar excess of EpoR_pY426 (ASFEpYTILDPS-amide) was incubated with S^{KKQ}BC (5mg/ml). Unbound peptide was removed by a protein concentrator (sartorius Vivaspin) while the mixture was concentrated to 20 mg/ml concentration. Sodium

cacodylate pH7.2 was added to a final concentration of 0.1 M prior to crystallisation. Crystallisation drops were set up in a ratio of 1:1 protein:precipitant in 18% ethanol, 0.1M HEPES pH7.5, 0.1M $MgCl_2$ using hanging drop at 4°C. Crystals were cryo-protected using 20 % PEG400 prior to flash-cooled.

Diffraction data were collected at 100 K on beamline i24 at Diamond Light Source. Data were recorded to Pilatus3 6M-F detector at 0.97 Å wavelength. Data were indexed, integrated, and reduced using XDS [74] and AIMLESS [75,76]. The phase was obtained by molecular replacement (MR) using Phaser [78] with the coordinates of SOCS2-EloB-EloC (PBD ID: 2C9W) as a search model. The presence of the EpoR_pY426 was observed in the initial electron density map. Model building was conducted manually with Coot [81] and refined with cycles of retrained refinement with REFMAC5 [82].

Crystallization and structure determination of SBC-GHR$_2$

GHR_pY595 (PVPDpYTSIHIV-amide) and GHR_pY487 (NIDFpYAQVSDI-amide) were mixed with S^{KKQ}BC at 1:1:1 stoichiometric ratio with a final concentration of 20 mg/ml and additional 0.1 M sodium cacodylate pH7.2. Drops of the complex were mixed 2:1 with 0.005 M cobalt chloride, 0.1 M MES pH6.5 and 1.0 M ammonium sulphate in the sitting-drop vapor diffusion format at 4°C. 20 % MPD was applied to crystal before flash-cooling.

Data collection of the SBC-GHR$_2$ co-crystal was at 100 K on beamline i24 at Diamond Light Source. Images were indexed, intergraded and reduced using XDS [74] and AIMLESS [75,76]. A molecular replacement solution was obtained by Phaser [78] using SBC-GHR as search model. Refinement was performed using REFMAC5 [82] and model building was performed in COOT [81].

Synthetic details

All chemicals, unless otherwise stated were commercially available and used without further purification. Solvents were anhydrous and reactions preformed under positive pressure of nitrogen. Flash column chromatography was performed using a Teledyne Isco Combiflash Rf or Rf200i. As prepacked columns RediSep Rf Normal Phase Disposable Columns were used. NMR spectra were recorded on a Bruker 500 Ultrashield. ^{13}C spectra were ^{1}H decoupled. Chemical shifts (δ) are reported in ppm relative to solvent (CD_3OD: δ_H = 3.31ppm, δ_C = 49.0 ppm) as internal standard. Low resolution MS and analytical HPLC traces were recorded on an Agilent Technologies 1200 series HPLC connected to an

Agilent Technologies 6130 quadrupole LC/MS, connected to an Agilent diode array detector. The column used was a Waters XBridge column (50 mm × 2.1 mm, 3.5 µm particle size) and the compounds were eluted with a gradient of 5−95% acetonitrile/water + 0.1% formic acid over 3 min. Preparative HPLC was performed on a Gilson Preparative HPLC System with a Waters X-Bridge C18 column (100 mm x 19 mm; 5 µm particle size) and a gradient of 5 % to 95 % acetonitrile in water over 10 min, flow 25 mL/min, with 0.1 % formic acid in the aqueous phase.

Cbz-O-bis(dimethylamino)phosphono)-L-tyrosine (**1**)

O-bis(dimethylamino)phosphono)-*L*-tyrosine [83] (485 mg, 1.54 mmol) and $NaHCO_3$ (260 mg, 3.1 mmol) were dissolved in the mixture THF/H_2O = 1:1 (10 mL) and *N*-(benzyloxycarbonyloxy)succinimide (383 mg, 1.54 mmol) was added. The reaction mixture was stirred at room temperature overnight. After the addition of 5% $NaHSO_4$ the product was extracted with ethyl acetate, washed with brine, dried over $MgSO_4$, concentrated by rotary evaporation under reduced pressure. After drying, Cbz-*O*-bis(dimethylamino)phosphono)-*L*-tyrosine **1** (620 mg, 89%) was obtained as pale yellow. ^{1}H NMR (CD_3OD): 2.69 (d, *J* = 10.1 Hz, 12H), 2.92 (dd, *J* = 14.0, 9.3 Hz, 1H), 3.18 (dd, *J* = 14.0, 4.9 Hz, 1H), 4.40 (dd, *J* = 4.9, 9.3 Hz, 1H), 5.03 (s, 2H), 7.06 (d, *J* = 8.5 Hz, 2H), 7.21 (d, *J* = 8.5 Hz, 2H), 7.25-7.36 (m, 5H). ^{31}P NMR (CD_3OD): 18.2.

(S)-2-amino-3-(4-((bis(dimethylamino)phosphoryl)oxy)phenyl)-N-methylpropanamide (**2**)

To a mixture of the compound **1** (160 mg, 0.35 mmol), HATU (135 mg, 0.35 mmol), HOAt (48 mg, 0.35 mmol) and DIPEA (150 µL, 1 mmol) in DMF (1 mL), 2M methylamine solution in THF (0.5 mL) was added under stirring at room temperature. After two hours, LC-MS analysis showed complete conversion of the starting material and formation of the desired product. The mixture was diluted with ethyl acetate, washed with 5% $NaHSO_4$, brine, dried over $MgSO_4$, concentrated by rotary evaporation under reduced pressure. The crude product was dissolved in the mixture ethanol/ethyl acetate= 1:1 (8 mL). Hydrogenation was carried out using H-Cube at 80 ˚C, Pd/C, 1 atm, at 1mL/min. The solvent was evaporated under vacuum to afford **2** (108 mg, 92%) which was directly used in the next step without any further purification. ^{1}H NMR (CD_3OD): 2.66 (s, 3H), 2.71 (d, *J* = 10.1 Hz, 12H), 2.81 (m, 1H), 2.96 (m, 1H), 3.51 (m, 1H), 7.09 (dd, *J* = 8.5, 1.0 Hz, 2H), 7.20 (d, *J* = 8.5 Hz, 2H). ^{31}P NMR (CD_3OD): 18.3.

(S)-4-(3-(methylamino)-3-oxo-2-(2,2,2-trifluoroacetamido)propyl)phenyl dihydrogen phosphate (**3**)

A solution of the compound **2** (80 mg, 0.24 mmol) and DIPEA (85 µL, 0.48 mmol) in DCM (2 mL) was cooled to -78 °C, and trifluoroacetic anhydride (34 µL, 0.24 mmol) was added. The reaction mixture was stirred 1h at -78 °C. After solvent evaporation the residue was dissolved in acetonitrile (0.5 mL) and 2M HCl was added (2 mL). The mixture was stirred at room temperature overnight until no presence of the starting materials was detected by LC-MS. The solvents were evaporated and residue was purified by HPLC to afford phosphate **3** (30 mg, 34%) as a white solid. ^{1}H NMR (CD_3OD): 2.69 (s, 3H), 2.96 (dd, *J* = 13.8, 6.4 Hz, 1H), 3.15 (dd, *J* = 13.8, 6.4 Hz, 1H), 4.57 (dd, *J* = 8.8, 6.4 Hz, 1H), 7.13 (dd, *J* = 8.5, 1.1 Hz, 2H), 7.22 (d, *J* = 8.5 Hz, 2H). ^{13}C NMR (CD_3OD): 26.3, 37.8, 56.5, 117.3 (q, *J* = 286.7 Hz), 121.3 (d, *J* = 4.5 Hz), 131.3, 134.1, 151.9 (d, *J* = 6.8 Hz), 158.7 (q, *J* = 37.5 Hz), 172.4. ^{31}P NMR (CD_3OD): 3.7. ^{19}F NMR (CD_3OD): -75.6.

Peptide synthesis

All peptides were prepared via solid-phase peptide synthesis on 10 mmol scale using standard Fmoc chemistry on Rink amide resin (0.68 mmol/g) on an INTAVIS ResPepSL automated peptide synthesizer. *O*-(dibenzylphosphono)-*N*-Fmoc-*L*-tyrosine was synthesised as described below. The peptides were cleaved with 2.5% triisopropylsilane and 2.5% water in TFA. The crude peptides were isolated from the cleavage mixture by precipitation with cold ether, dissolved in the mixture water/DMF=1/1 and purified by preparative HPLC under the following conditions: Waters X-Bridge C18 column (100 mm x 19 mm; 5 µm particle size), gradient of 5-95 % acetonitrile in water over 10 min, flow 25 mL/min, with 0.1 % formic acid in the aqueous phase, UV detection at λobs = 190 and 210 nm). The poor soluble peptides were purified according to the literature procedure [84]: the impurities were extracted by DCM from the solution of peptides in 20% acetic acid. The purity and identity of the peptides were determined by the analytical LCMS on an Agilent Technologies 1200 series HPLC connected to an Agilent Technologies 6130 quadrupole LC/MS linked to an Agilent diode array detector.

O-(dibenzylphosphono)-N-Fmoc-L-tyrosine

To a solution of Fmoc-tyrosine (2 g, 5 mmol) in anhydrous THF (12 mL) *N*-methylmorpholine (540 µL, 5 mmol) and *tert*-butyldimethylsilyl chloride (740 mg, 4.9 mmol) were added. After 15 min 4,5-dicyanoimidazole (1.8 g, 15 mmol) and diisopropylphosphoramidite (3.4 mL, 10 mmol) were added and the reaction mixture was

stirred at room temperature for 4 h. After cooling to 0°C 70% *tert*-butyl hydroperoxide (2 ml, 15 mmol) was introduced. After stirring for 2 h at 0°C, 10% $Na_2S_2O_5$ (20 ml) was added and stirring continued for one more hour. The product was extracted with ethyl acetate, washed with a 5% solution of $KHSO_4$, brine, dried over $MgSO_4$, concentrated by rotary evaporation under reduced pressure, and further purified by column chromatography on silica gel using a gradient elution of 0% to 10% of MeOH in DCM to afford *O*-(dibenzylphosphono)-*N*-Fmoc-*L*-tyrosine (3g, 90%) as a pale yellow solid. NMR spectra were in agreement with the published data [85].

Isothermal titration calorimetry

Experiments were performed with ITC200 instrument (Malvern) in 100mM HEPES pH7.5, 50mM NaCl, 0.5mM TCEP at 298K stirring the sample at 750 rpm. The ITC titration consisted of 0.4 μL initial injection (discarded during data analysis) followed by 19 of 2 μL injections at 120 seconds interval between injections. The GHR_pY595 peptide (PVPDpYTSIHIV-amide, 750 μM), EpoR_pY426 (ASFEpYTILDPS-amide, 750/1500 μM) and GHR_pY487 (NIDFpYAQVSDI-amide, 1500 μM) were directly titrated into SBC (50 μM). Binding data was subtracted from a control titration where peptide was titrated into buffer, and fitted using a one-set-of-site binding model to obtain dissociation constants, binding enthalpy (ΔH), and stoichiometry (N) using MicroCal ITC-ORIGIN Analysis Software 7.0 (Malven).

Surface plasmon resonance

Experiments were performed using Biacore T200 instrument (GH Healthcare) in 20 mM HEPES pH7.5, 150 mM NaCl, 1mM TCEP, 0.005 % Tween20 buffer at 10 °C. Biotinylated wild type SBC and mutants were immobilized onto a chip surface and injected a series of seven concentrations (0.08, 0.25, 0.7, 2.2, 6.7, 20 and 60 μM) of peptide across the sensor surface for 60 sec contact time and 120 sec disassociation time at 30 μl/min flow rate. Data analysis was carried out using Biacore Evaluation Software (GE Healthcare). All data were double-referenced for reference surface and blank injection. The processed sensograms were fit to a steady-state affinity using a 1:1 binding model for K_D estimation.

^{19}F Spy NMR displacement

All NMR experiments were conducted using AV-500 NHz Bruker spectrometer equipped with a 5 mm CPQCI 1H/19F/13C/15N/D Z-GRD cryoprobe) at 298 K. Spectra were recorded using a CPMG pulse sequence that attenuates broad resonances with an

interpulse delay at 0.133 seconds for 80 scans. The transmitter frequency was placed close to the resonance of O_1 = -35451 Hz (-75.3 ppm). Protein was used at 5 μM and spy molecule (phosphate **3**) was used at 100 μM in buffer containing 20 mM HEPES pH8, 50 mM NaCl, 1mM DTT, 20 % D_2O. The GHR related peptides including wild type, V(-3)R, V(-3)Y and alanine scanning were used at10 μM, whereas the EpoR related peptide including wild type and alanine scanning were used at 50 μM. All NMR data were processed and analyzed using TopSpin (Bruker).

Signal contrast between bound and free form of the spy is calculated as below

$$Contrast\ (\%) = \frac{Spy_{Free} - Spy_{Bound}}{Spy_{Free}}$$

where Spy_{free} is the peak integral or T_2 of spy in the absence of protein, while Spy_{bound} is the signal of spy in complex with protein.

For competition studies, the percentage displacement is calculated as below

$$Displacement\ (\%) = \frac{Competitor - Spy_{Bound}}{Spy_{Free} - Spy_{Bound}}$$

where competitor is the peak integral of spy in the presence of competitor, protein and spy.

To calculate the dissociation constant of competitor (Ki), assuming all other experimental conditions are identical in the absence and presence of competitor, the signal intensity is proportional to the concentration of protein-spy complex, and can be expressed as below [47]

$$\frac{Signal(I)}{Signal(0)} = \frac{[Eo] + [Lo] + K_D - [EI] - \sqrt{([Eo] + [Lo] + K_D - [EI])^2 - 4([Eo] - [EI])[Lo]}}{[Eo] + [Lo] + KD - \sqrt{([Eo] + [Lo] + K_D)^2 - 4[Eo][Lo]}}$$

where Signal(I) and Signal(0) are the fluorine intensities in the presence and absence of competitor, respectively. $[E_0]$ is the total concentration of protein. $[L_0]$ is the total concentration of ligand. [EL] is concentration of protein-ligand complex. From the observed ratio of signal intensity, the Ki can be calculated.

T_2 relaxation measurement

Spectra were acquired for each sample at 298K, and 80 scans using a CPMG pulse sequence with a relaxation delay of 0.05, 0.1, 0.2, 0.4 and 0.8 s. Protein was used at 5 μM and spy was used at 100 μM in buffer containing 20 mM HEPES pH8, 50 mM NaCl, 1mM DTT, 20 % D_2O. The T_2 relaxation time was calculated using Prism by fitting magnetisation to a mono exponential function

$$M_{xy}(t) = M_{xy}(0)e^{-t/T_2}$$

where T_2 is the time constant of decay of component magnetisation in the xy-plane (M_{xy}) that is perpendicular to the external field B_0.

Accession Code

The coordinates and structure factors for SBC in complex with EpoR_pY426 peptide, GHR_pY595 peptide, and GHR2_pY595 peptide have been deposited to the Protein Data Bank (PDB) with accession codes 6I4X, 6I5N, and 6I5J, respectively.

Supplemental information

Supplementary Figures S1–S14 are included.

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP7/2007-2013) as a Starting Grant to A.C. (grant agreement No. ERC-2012-StG-311460 DrugE3CRLs). Biophysics and drug discovery activities at Dundee are supported by Wellcome Trust strategic awards (100476/Z/12/Z and 094090/Z/10/Z, respectively). W.K. was supported by a PhD Scholarship by the School of Life Sciences at the University of Dundee. We thank P. Fyfe for support with in-house X-ray facilities, the Diamond Light Source for beamtime (BAG proposals MX14980-13 and MX14980-20) and P. Romano and P. Aller for support at beamlines. We thank de Castro for support with NMR and helpful discussions.

Competing interests: The A.C. laboratory receives sponsored research support from Boehringer Ingelheim and Nurix, Inc. A.C. is a scientific founder, director and shareholder of Amphista Therapeutics, a company that is developing targeted protein degradation therapeutic platforms.

References

1. Sato, N. & Miyajima, A. Multimeric cytokine receptors: common versus specific functions. *Curr. Opin. Cell Biol.* **6**, 174–179 (1994).

2. Starr, R. *et al.* A family of cytokine-inducible inhibitors of signalling. *Nature* **387**, 917–921 (1997).

3. Yasukawa, H. The JAK-binding protein JAB inhibits Janus tyrosine kinase activity through binding in the activation loop. *EMBO J.* **18**, 1309–1320 (1999).

4. Babon, J. J. *et al.* The Structure of SOCS3 Reveals the Basis of the Extended SH2 Domain Function and Identifies an Unstructured Insertion That Regulates Stability. *Mol. Cell* **22**, 205–216 (2006).

5. Bullock, A. N., Rodriguez, M. C., Debreczeni, J. É., Songyang, Z. & Knapp, S. Structure of the SOCS4-ElonginB/C Complex Reveals a Distinct SOCS Box Interface and the Molecular Basis for SOCS-Dependent EGFR Degradation. *Structure* **15**, 1493–1504 (2007).

6. Yoshimura, a *et al.* A novel cytokine-inducible gene CIS encodes an SH2-containing protein that binds to tyrosine-phosphorylated interleukin 3 and erythropoietin receptors. *EMBO J.* **14**, 2816–2826 (1995).

7. Kamura, T. *et al.* The Elongin BC complex interacts with the conserved SOCS-box motif present in members of the SOCS, ras, WD-40 repeat, and ankyrin repeat families. *Genes Dev.* **12**, 3872–3881 (1998).

8. Zhang, J. G. *et al.* The conserved SOCS box motif in suppressors of cytokine signaling binds to elongins B and C and may couple bound proteins to proteasomal degradation. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 2071–2076 (1999).

9. Kamura, T. *et al.* VHL-box and SOCS-box domains determine binding specificity for Cul2-Rbx1 and Cul5-Rbx2 modules of ubiquitin ligases. *Genes Dev.* **18**, 3055–3065 (2004).

10. Piessevaux, J., Lavens, D., Peelman, F. & Tavernier, J. The many faces of the SOCS box. *Cytokine and Growth Factor Reviews* (2008). doi:10.1016/j.cytogfr.2008.08.006

11. Babon, J. J., Sabo, J. K., Zhang, J. G., Nicola, N. A. & Norton, R. S. The SOCS Box Encodes a Hierarchy of Affinities for Cullin5: Implications for Ubiquitin Ligase Formation and Cytokine Signalling Suppression. *J. Mol. Biol.* **387**, 162–174 (2009).

12. Linossi, E. M. & Nicholson, S. E. The SOCS box-adapting proteins for ubiquitination and proteasomal degradation. *IUBMB Life* **64**, 316–23 (2012).

13. Bulatov, E. & Ciulli, A. Targeting Cullin–RING E3 ubiquitin ligases for drug discovery: structure, assembly and small-molecule modulation. *Biochem. J.* **467**, 365–386 (2015).

14. Sasaki, A. *et al.* Cytokine-inducible SH2 protein-3 (CIS3/SOCS3) inhibits Janus tyrosine kinase by binding through the N-terminal kinase inhibitory region as well as SH2 domain. *Genes to Cells* **4**, 339–351 (1999).

15. Ram, P. A. & Waxman, D. J. SOCS/CIS protein inhibition of growth hormone-stimulated STAT5 signaling by multiple mechanisms. *J. Biol. Chem.* **274**, 35553–35561 (1999).

16. Vesterlund, M. *et al.* The SOCS2 ubiquitin ligase complex regulates growth hormone receptor levels. *PLoS One* **6**, (2011).

17. Piessevaux, J., De Ceuninck, L., Catteeuw, D., Peelman, F. & Tavernier, J. Elongin B/C recruitment regulates substrate binding by CIS. *J. Biol. Chem.* (2008). doi:10.1074/jbc.M803742200

18. Liau, N. P. D. *et al.* The molecular basis of JAK/STAT inhibition by SOCS1. *Nat. Commun.* (2018). doi:10.1038/s41467-018-04013-1

19. Bergamin, E., Wu, J. & Hubbard, S. R. Structural Basis for Phosphotyrosine Recognition by Suppressor of Cytokine Signaling-3. *Structure* **14**, 1285–1292 (2006).

20. Kershaw, N. J. *et al.* SOCS3 binds specific receptor-JAK complexes to control cytokine signaling by direct kinase inhibition. *Nat. Struct. Mol. Biol.* **20**, 469–476 (2013).

21. Zadjali, F. *et al.* Structural basis for c-KIT inhibition by the suppressor of cytokine signaling 6 (SOCS6) ubiquitin ligase. *J. Biol. Chem.* **286**, 480–490 (2011).

22. Haan, S. SOCS2 physiological and pathological functions. *Front. Biosci.* **8**, 760 (2016).

23. Keating, N. & Nicholson, S. E. SOCS-mediated immunomodulation of natural killer cells. *Cytokine* (2018). doi:10.1016/j.cyto.2018.03.033

24. Turnley, A. M. Role of SOCS2 in growth hormone actions. *Trends in Endocrinology and Metabolism* (2005). doi:10.1016/j.tem.2005.01.006

25. Rico-Bautista, E., Flores-Morales, A. & Fernández-Pérez, L. Suppressor of cytokine signaling (SOCS) 2, a protein with multiple functions. *Cytokine Growth Factor Rev.* (2006). doi:10.1016/j.cytogfr.2006.09.008

26. Metcalf, D. *et al.* Gigantism in mice lacking suppressor of cytokine signalling-2. *Nature* **405**, 1069–1073 (2000).

27. Greenhalgh, C. J. *et al.* Growth Enhancement in Suppressor of Cytokine Signaling 2 (SOCS-2)-Deficient Mice Is Dependent on Signal Transducer and Activator of Transcription 5b (STAT5b). *Mol. Endocrinol.* **16**, 1394–1406 (2002).

28. Greenhalgh, C. J. *et al.* SOCS2 negatively regulates growth hormone action in vitro and in vivo. *J. Clin. Invest.* **115**, 397–406 (2005).

29. Hansen, J. a, Lindberg, K., Hilton, D. J., Nielsen, J. H. & Billestrup, N. Mechanism of inhibition of growth hormone receptor signaling by suppressor of cytokine signaling proteins. *Mol. Endocrinol.* **13**, 1832–1843 (1999).

30. Stofega, M. R., Herrington, J., Billestrup, N. & Carter-Su, C. Mutation of the SHP-2 binding site in growth hormone (GH) receptor prolongs GH-promoted tyrosyl

phosphorylation of GH receptor, JAK2, and STAT5B. *Mol. Endocrinol.* **14**, 1338–1350 (2000).

31. Bulatov, E. *et al.* Biophysical Studies on Interactions and Assembly of Full-size E3 Ubiquitin Ligase. *J. Biol. Chem.* **290**, 4178–4191 (2015).

32. Bullock, A. N., Debreczeni, J. E., Edwards, A. M., Sundstrom, M. & Knapp, S. Crystal structure of the SOCS2-elongin C-elongin B complex defines a prototypical SOCS box ubiquitin ligase. *Proc. Natl. Acad. Sci.* **103**, 7637–7642 (2006).

33. Yang, X. O. *et al.* The signaling suppressor CIS controls proallergic T cell development and allergic airway inflammation. *Nat. Immunol.* **14**, 732–740 (2013).

34. Eyckerman, S. *et al.* Design and application of a cytokine-receptor-based interaction trap. *Nat. Cell Biol.* **3**, 1114–1119 (2001).

35. Lavens, D. A complex interaction pattern of CIS and SOCS2 with the leptin receptor. *J. Cell Sci.* **119**, 2214–2224 (2006).

36. Goldshmit, Y., Walters, C. E., Scott, H. J., Greenhalgh, C. J. & Turnley, A. M. SOCS2 induces neurite outgrowth by regulation of epidermal growth factor receptor activation. *J. Biol. Chem.* (2004). doi:10.1074/jbc.M312873200

37. Dey, B. R., Spence, S. L., Nissley, P. & Furlanetto, R. W. Interaction of Human Suppressor of Cytokine Signaling (SOCS)-2 with the Insulin-like Growth Factor-I Receptor. *J. Biol. Chem.* **273**, 24095–24101 (1998).

38. Goldschmidt, L., Cooper, D. R., Derewenda, Z. S. & Eisenberg, D. Toward rational protein crystallization: A Web server for the design of crystallizable protein variants. *Protein Sci.* **16**, 1569–1576 (2007).

39. Gadd, M. S., Bulatov, E. & Ciulli, A. Serendipitous SAD solution for DMSO-Soaked SOCS2-ElonginC-ElonginB crystals using covalently incorporated dimethylarsenic: Insights into substrate receptor conformational flexibility in cullin RING ligases. *PLoS One* **10**, (2015).

40. Kim, Y. K. *et al.* Structural basis of intersubunit recognition in elongin BC-cullin 5-SOCS box ubiquitin-protein ligase complexes. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 1587–1597 (2013).

41. Liu, J. & Nussinov, R. The mechanism of ubiquitination in the cullin-RING E3 ligase machinery: Conformational control of substrate orientation. *PLoS Comput. Biol.* **5**, (2009).

42. Kaneko, T. *et al.* Loops govern SH2 domain specificity by controlling access to binding pockets. *Sci. Signal.* **3**, (2010).

43. Carr, H. & Purcell, E. Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments. *Phys. Rev.* **94**, 630–638 (1954).

44. Meiboom, S., Gill, D. & Gillt, D. Modified Spin-Echo Method for Measuring Nuclear Relaxation Times Modified Spin-Echo Method for Measuring Nuclear Relaxation Times*. *Cit. Rev. Sci. Instruments J. Chem. Phys. J. Appl. Phys. J. Chem. Phys. J. Appl. Phys. Spin-Echo Method Meas. Relax. Times Two-Line NMR Spectra Rev.*

Sci. Instrum **29**, 214504–2174 (1958).

45. Dalvit, C., Fagerness, P. E., Hadden, D. T. A., Sarver, R. W. & Stockman, B. J. Fluorine-NMR experiments for high-throughput screening: Theoretical aspects, practical considerations, and range of applicability. *J. Am. Chem. Soc.* **125**, 7696–7703 (2003).

46. de Castro, G. V. & Ciulli, A. Spy vs. spy: selecting the best reporter for 19F NMR competition experiments. *Chem. Commun.* **55**, 1482–1485 (2019).

47. Wang, Y. Sen, Liu, D. & Wyss, D. F. Competition STD NMR for the detection of high-affinity ligands and NMR-based screening. *Magn. Reson. Chem.* **42**, 485–489 (2004).

48. Forbes, S. A. *et al.* COSMIC: Somatic cancer genetics at high-resolution. *Nucleic Acids Res.* **45**, D777–D783 (2017).

49. Rupp, R. *et al.* A Point Mutation in Suppressor of Cytokine Signalling 2 (Socs2) Increases the Susceptibility to Inflammation of the Mammary Gland while Associated with Higher Body Weight and Size and Higher Milk Production in a Sheep Model. *PLoS Genet.* (2015). doi:10.1371/journal.pgen.1005629

50. Zhang, Y. *et al.* Simultaneous binding of two peptidyl ligands by a Src homology 2 domain. *Biochemistry* **50**, 7637–7646 (2011).

51. Mészáros, B., Kumar, M., Gibson, T. J., Uyar, B. & Dosztányi, Z. Degrons in cancer. *Science Signaling* **10**, (2017).

52. Liu, C. *et al.* Control of β-catenin phosphorylation/degradation by a dual-kinase mechanism. *Cell* (2002). doi:10.1016/S0092-8674(02)00685-2

53. Alexander, W. S. & Hilton, D. J. The Role of Suppressors of Cytokine Signaling (SOCS) Proteins in Regulation of the Immune Response. *Annu. Rev. Immunol.* **22**, 503–529 (2004).

54. Yoshimura, A., Naka, T. & Kubo, M. SOCS proteins, cytokine signalling and immune regulation. *Nature Reviews Immunology* **7**, 454–465 (2007).

55. Linossi, E. M., Babon, J. J., Hilton, D. J. & Nicholson, S. E. Suppression of cytokine signaling: The SOCS perspective. *Cytokine and Growth Factor Reviews* **24**, 241–248 (2013).

56. Galic, S., Sachithanandan, N., Kay, T. W. & Steinberg, G. R. Suppressor of cytokine signalling (SOCS) proteins as guardians of inflammatory responses critical for regulating insulin sensitivity. *Biochem. J.* **461**, 177–188 (2014).

57. Inagaki-Ohara, K., Kondo, T., Ito, M. & Yoshimura, A. SOCS, inflammation, and cancer. *JAK-STAT* **2**, e24053 (2013).

58. Wikman, H. *et al.* Identification of differentially expressed genes in pulmonary adenocarcinoma by using cDNA array. *Oncogene* **21**, 5804–5813 (2002).

59. Sutherland, K. D. *et al.* Differential hypermethylation of SOCS genes in ovarian and breast carcinomas. *Oncogene* **23**, 7726–7733 (2004).

60. Farabegoli, F., Ceccarelli, C., Santini, D. & Taffurelli, M. Suppressor of cytokine signalling 2 (SOCS-2) expression in breast carcinoma. *J. Clin. Pathol.* **58**, 1046–1050 (2005).

61. Qiu, X. *et al.* Reduced expression of SOCS2 and SOCS6 in hepatocellular carcinoma correlates with aggressive tumor progression and poor prognosis. *Mol. Cell. Biochem.* **378**, 99–106 (2013).

62. Zhu, J. G. *et al.* Expression of SOCSs in human prostate cancer and their association in prognosis. *Mol. Cell. Biochem.* **381**, 51–59 (2013).

63. Haffner, M. C. *et al.* Favorable prognostic value of SOCS2 and IGF-I in breast cancer. *BMC Cancer* **7**, (2007).

64. Paul, I. *et al.* The ubiquitin ligase Cullin5 SOCS2 regulates NDR1/STK38 stability and NF-κB transactivation. *Sci. Rep.* **7**, (2017).

65. Oeckinghaus, A. & Ghosh, S. The NF-kappaB family of transcription factors and its regulation. *Cold Spring Harbor perspectives in biology* **1**, (2009).

66. Hoesel, B. & Schmid, J. A. The complexity of NF-κB signaling in inflammation and cancer. *Molecular Cancer* **12**, (2013).

67. Hughes, S. J. & Ciulli, A. Molecular recognition of ternary complexes: a new dimension in the structure-guided design of chemical degraders. *Essays Biochem.* **61**, 505–516 (2017).

68. Lucas, X. & Ciulli, A. Recognition of substrate degrons by E3 ubiquitin ligases and modulation by small-molecule mimicry strategies. *Current Opinion in Structural Biology* **44**, 101–110 (2017).

69. Ottis, P. & Crews, C. M. Proteolysis-Targeting Chimeras: Induced Protein Degradation as a Therapeutic Strategy. *ACS Chem. Biol.* **12**, 892–898 (2017).

70. Zengerle, M., Chan, K. H. & Ciulli, A. Selective Small Molecule Induced Degradation of the BET Bromodomain Protein BRD4. *ACS Chem. Biol.* **10**, 1770–1777 (2015).

71. Maniaci, C. *et al.* Homo-PROTACs: bivalent small-molecule dimerizers of the VHL E3 ubiquitin ligase to induce self-degradation. *Nat. Commun.* **8**, 830 (2017).

72. Winter, G. E. *et al.* Phthalimide conjugation as a strategy for in vivo target protein degradation. *Science (80-.).* **348**, 1376 LP-1381 (2015).

73. Lu, J. *et al.* Hijacking the E3 Ubiquitin Ligase Cereblon to Efficiently Target BRD4. *Chem. Biol.* **22**, 755–763 (2015).

74. Kabsch, W. XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 125–132 (2010).

75. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 1204–1214 (2013).

76. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography* **67**, 235–242 (2011).

77. McCoy, A. J., Storoni, L. C. & Read, R. J. Simple algorithm for a maximum-likelihood

SAD function. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 1220–1228 (2004).

78. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

79. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 213–221 (2010).

80. Terwilliger, T. C. *et al.* Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. in *Acta Crystallographica Section D: Biological Crystallography* **64**, 61–69 (2007).

81. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).

82. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 355–367 (2011).

83. Chao, H. G. *et al.* Synthesis and Application of Fmoc-O-[bis(dimethylamino)phosphono]-tyrosine, a Versatile Protected Phosphotyrosine Equivalent. *J. Org. Chem.* (1995). doi:10.1021/jo00129a001

84. Biosystems, A. Cleavage, Deprotection, and Isolation of Peptides after Fmoc Synthesis. *Tech. Bull.* 1–12 (1998). doi:10.1038/nrd.2016.29

85. Perich, J. W. & Reynolds, E. C. The Facile One-Pot Synthesis of Nα-(9-Fluorenylmethoxycarbonyl)-O-(O′, O″-dialkylphosphoro)-L-tyrosines Using Dialkyl N,N-Diethylphosphoramidites. *Synlett* **1991**, 577–578 (1991).

86. Armougom, F. *et al.* Expresso: Automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic Acids Res.* (2006). doi:10.1093/nar/gkl092

87. Hruz T., Laule O., Szabo G., Wessendorp F., Bleuler S., Oertle L., Widmayer P., Gruissem W. and Zimmermann P. Genevestigator V3: a reference expression database for the meta-analysis of transcriptomes. *Advances in Bioinformatics* (2008). 420747

Figures and Schemes

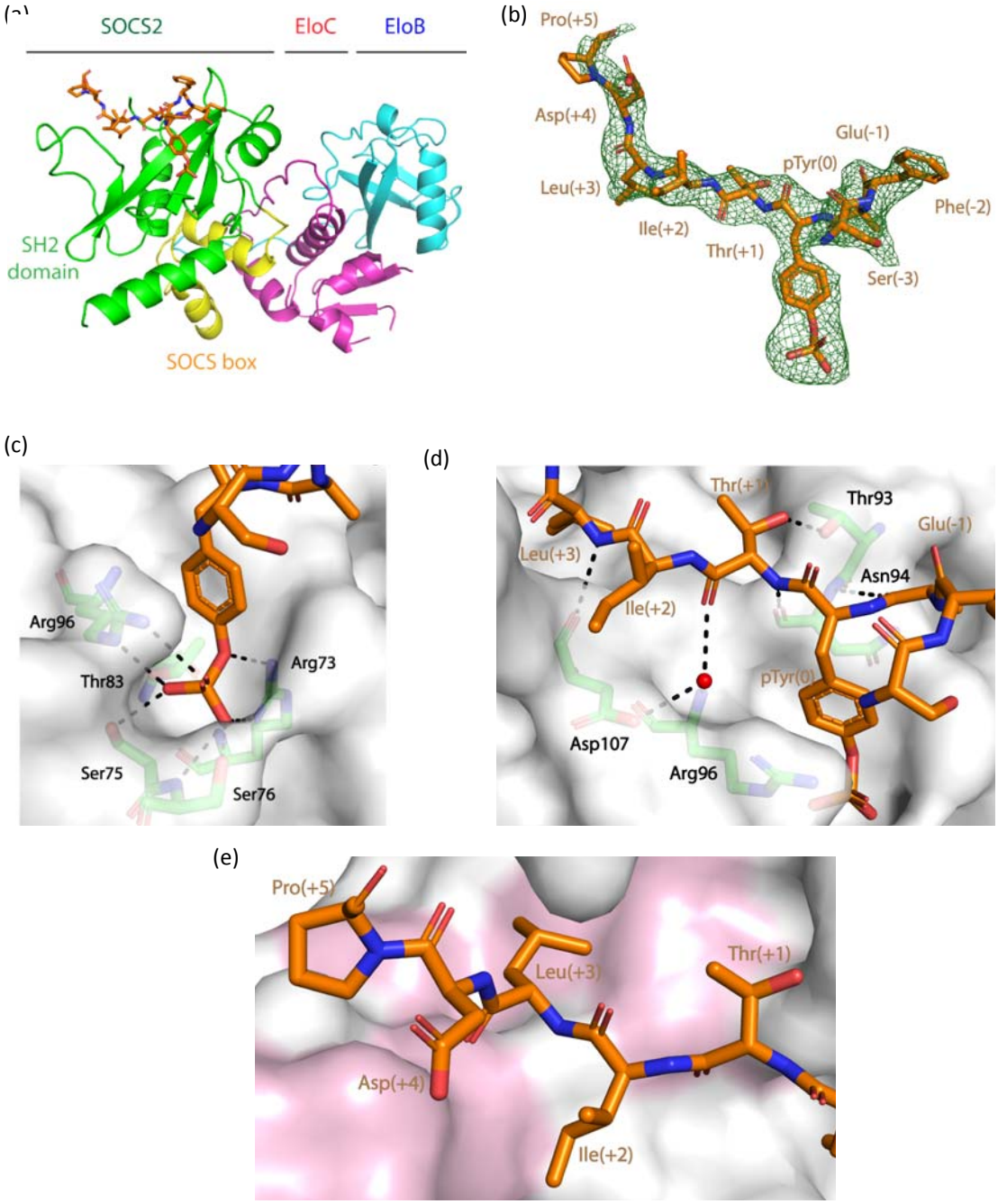


Figure 1. Structural and interaction detail of the SBC-EpoR co-crystal
(a) Domain and protein arrangement of the SBC-EpoR co-crystal. Protein chains are shown in cartoon, with EloB (cyan), EloC (magenta) and SOCS2, comprising of SOCS box (yellow) and SH2 domain (green). The EpoR_pY426 peptide is shown in orange stick. (b) The Fo-Fc ligand omit map of the EpoR_pY426 peptide (green mesh) contoured at 2.0 σ level to highlight densities for the EpoR_pY426 peptide (orange stick) (c) Hydrogen bond interactions (dash) between the pY of EpoR_pY426 peptide (orange stick) and SOCS2 (green). (d) Hydrogen bond interaction (dash) between the EpoR_pY426 peptide (orange stick), SOCS2 (green stick) and water (red sphere). (e) Hydrophobic interaction between EpoR_pY426 peptide (orange stick) and SOCS2 (surface). Hydrophobic residues on SOCS2 are colored in pink.

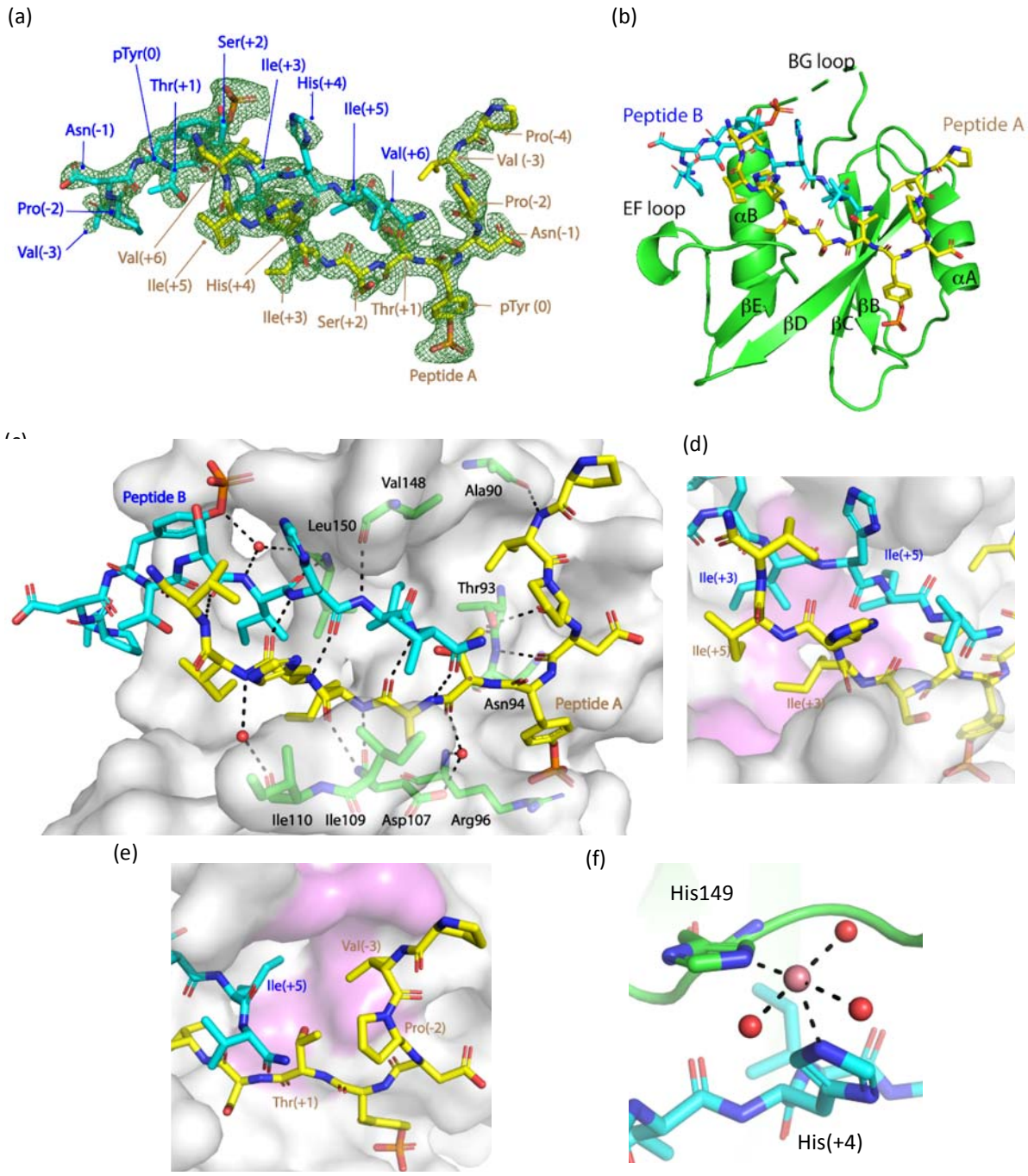


Figure 2. Structural and interaction detail of the SBC-GHR co-crystal

Two copies of GHR_pY595 were observed in the co-crystal. One copy is shown as peptide A (yellow stick) and the other one as peptide B (cyan stick). (a) The Fo-Fc ligand omit map of the peptides (green mesh) contoured at 2.5 σ level to highlight densities for the peptide A and peptide B (yellow and cyan stick). (b) Cartoon diagram of the SH2 domain of SOCS2 (green) with peptides A and peptide B. (c) Hydrogen bond interactions (dash) among peptide A, peptide B, SOCS2 (green stick) and water (red sphere). (d) (d)(e) Hydrophobic interaction of the peptides with C-terminal half and N-terminal half of the SH2 domain (surface), respectively. SOCS2 residues involved in hydrophobic interactions are coloured in pink. (f) The coordination of cobalt ion (pink sphere) with His149 of SOCS2 (green), His(+4) of peptide B and water molecules (red sphere).

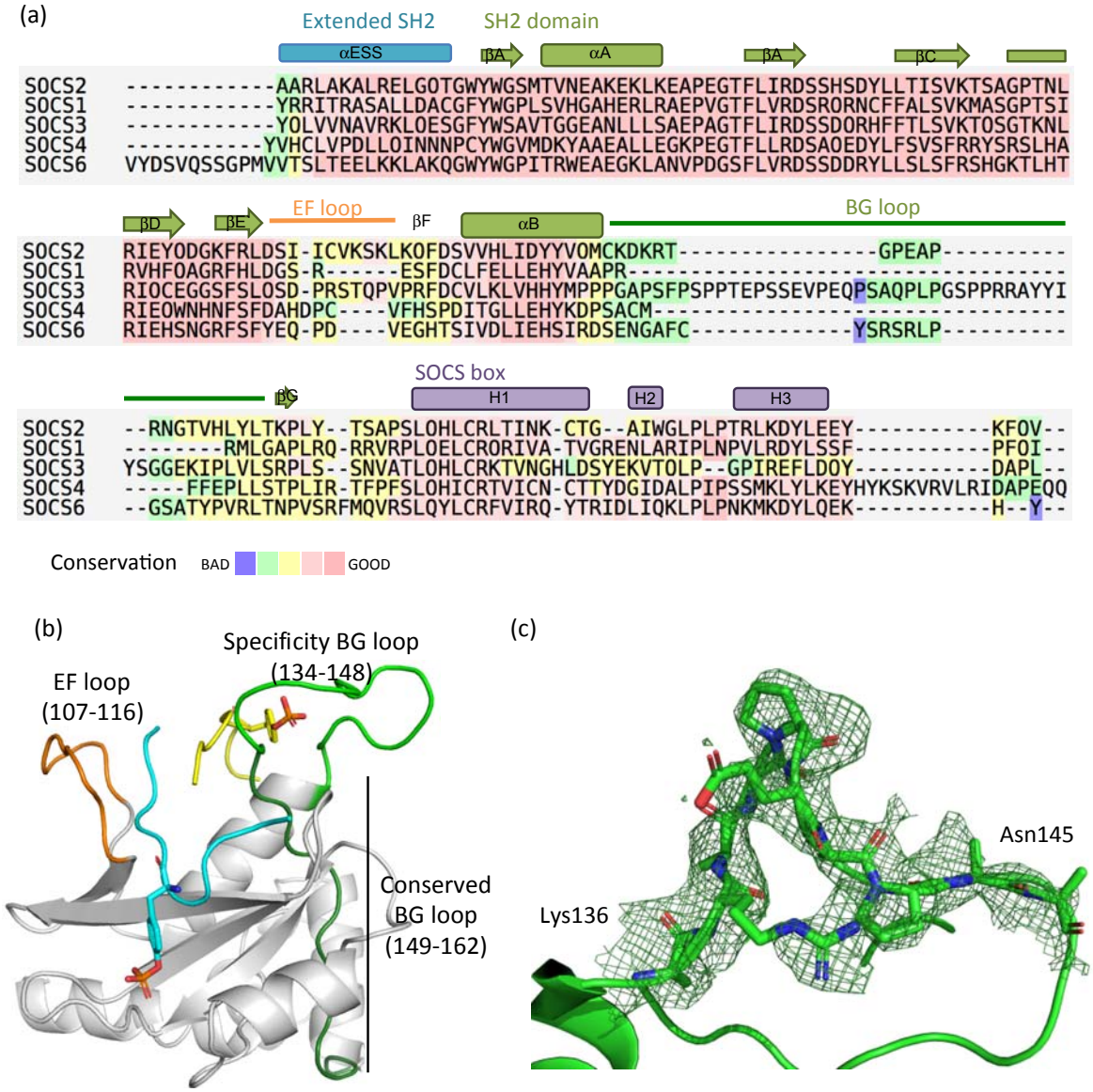


Figure 3. The BG loop of SOCS2

(a) Secondary structure elements in SOCS2 are shown above the sequence alignment. SOCS proteins with structure available were aligned using T-Coffee expresso mode for sequence alignment with structural information [86]. (b) Locations of the EF loop (orange) and the BG loop (green) on SOCS2 (white). (c) The Fo-Fc ligand omit map of the previously disordered BG loop (green mesh) contoured at 1.5 σ level.

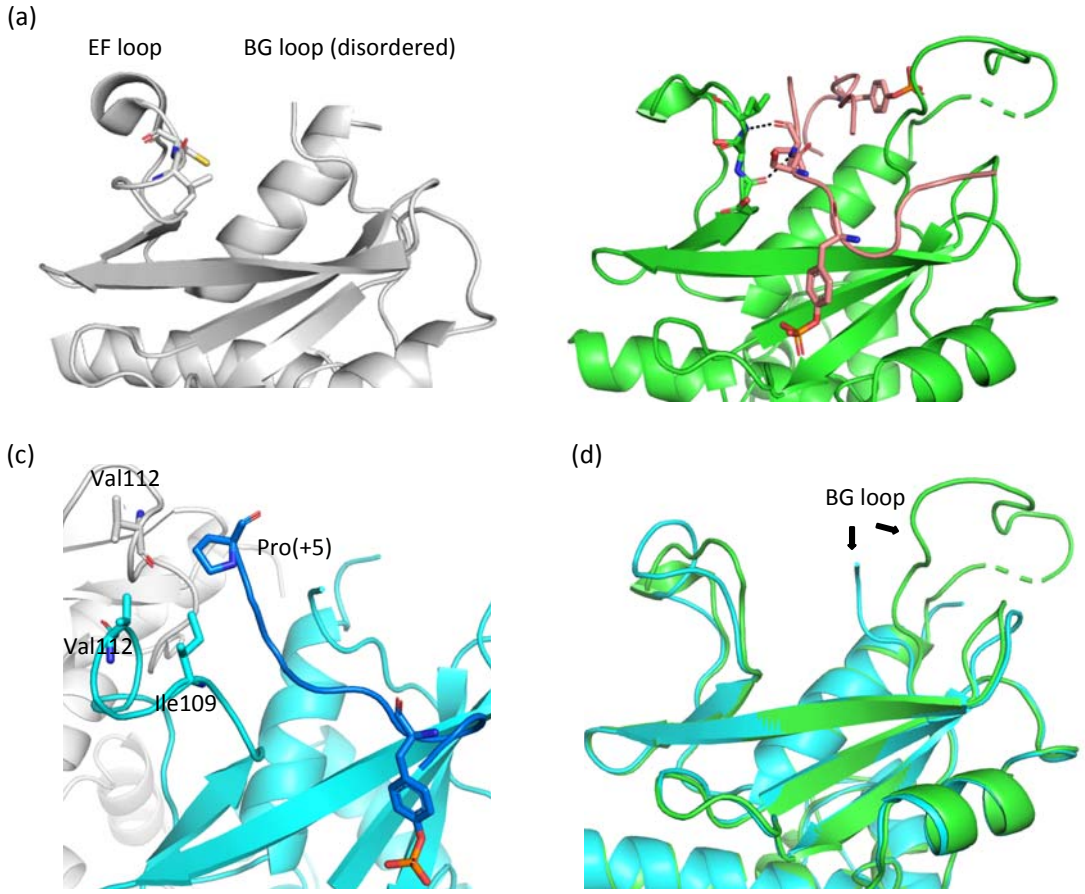
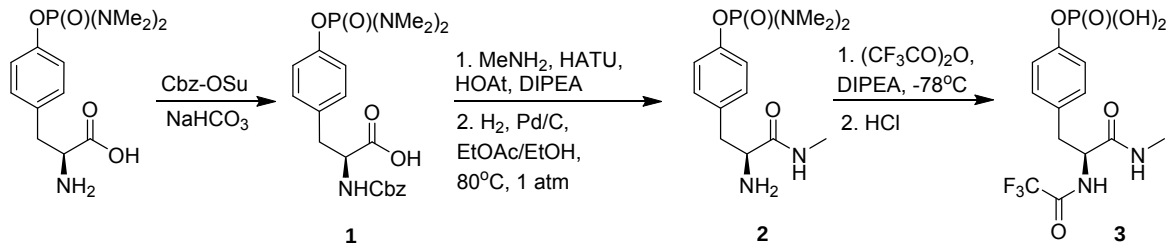


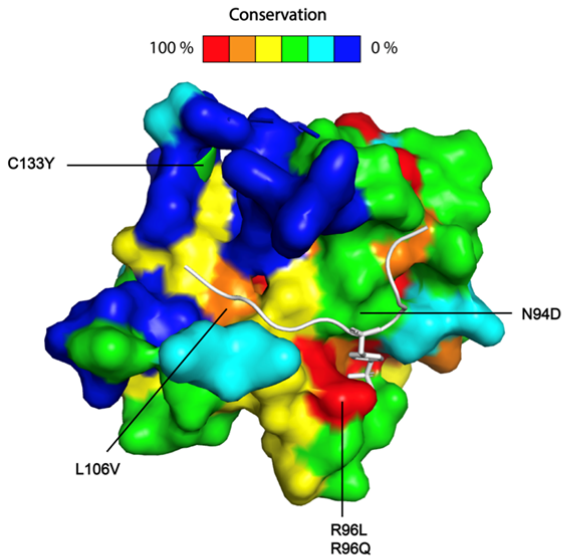
Figure 4. Conformation changes of the EF and BG loops

(a) The EF loop curls up in the apo SOCS2 (white, PDB code: 2C9W), placing Ile110 and Cys111 (stick) at the SH2 hydrophobic pocket. (b) The EF loop makes backbone interactions with GHR peptide (pink) in the SBC-GHR_2 structure. (c) The Ile109 and Val 112 of SOCS2 (cyan) and Val112 of SOCS2 symmetry mate (white) make unique hydrophobic interaction with Pro(+5) of EpoR peptide (blue). (d) Superposition of the SOCS2 from SBC-GHR_2 (green) and SBC-EpoR (cyan) displays conformational changes of the BG loop.



Scheme 1. Synthesis of spy molecule 3

(a)



(b)

Protein	SPR K_D (μM)		
	GHR_pY595	GHR_pY487	EpoR_pY426
Wild type	1.4 ± 0.12	5.1 ± 0.77	12.5 ± 0.63
L106V	2.4 ± 0.11	8.9 ± 0.70	14.1 ± 1.23
C133Y	2.2 ± 0.10	7.7 ± 0.71	13.4 ± 0.82
N94D	n.d.	weak	n.d.
R96L	weak	weak	n.d.
R96Q	n.d.	n.d.	n.d.

(c)

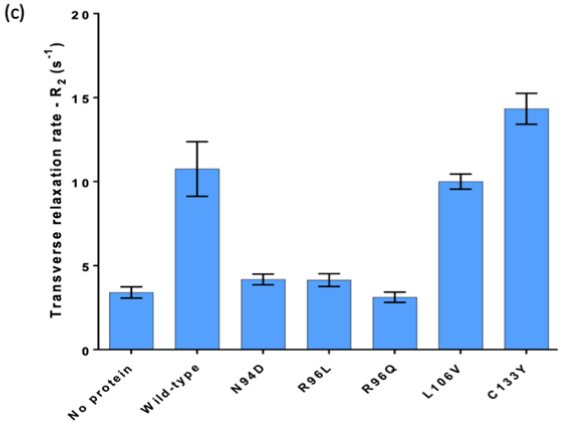


Figure 5. Characterization of SNP containing SOCS2 by SPR and ^{19}F-NMR

(a) Sequence conservation mapped onto the SH2 domain of SOCS2 surface. Conservation surface representation based upon the ClustalW multiple sequence alignment of CIS and SOCS1 – SOCS7 sequences where highly conserved residues are shown in red/orange colour and variable residue positions coloured blue. (b) The SBC proteins binding affinity for phosphorylated substrate peptides were determined by SPR. Weak denotes signal detected, however, saturation was not achieved. Hence, binding too weak to be detected. N.d. denotes signal not detected. Values reported are the means ± s.e.m. from four independent experiments (c) Transverse relaxation rate (R_2) of SBC proteins. Error bars are R_2 errors generated from five-point data fitting.

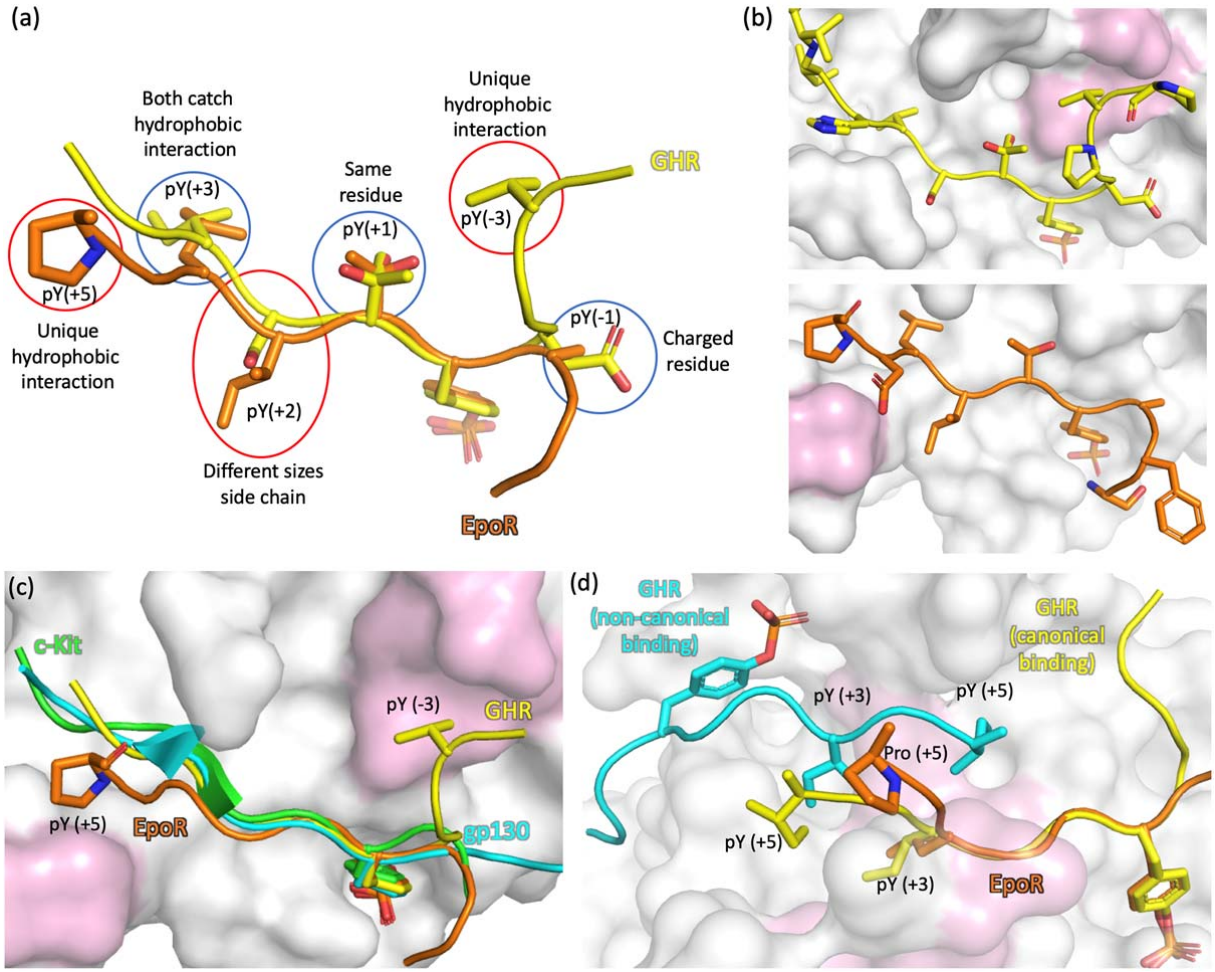


Figure 6. Comparison of the GHR and EpoR binding

(a) Overlay of the GHR_pY595 (yellow) and EpoR (orange) peptides. Side chains circled in blue have similar properties and red have different properties. (b) Different hydrophobic interactions (pink) caught by GHR_pY595 (yellow) and EpoR (orange) peptide. (c) Overlay of the substrate peptide of SOCS proteins. The GHR_pY595 (yellow) and EpoR (orange) peptides catch distinct hydrophobic cavities (pink) on SOCS2. These interactions differentiate SOCS2 substrate binding mode from other substrates of SOCS proteins. The SOCS3 substrate peptide gp130 is in cyan (PDB ID: 2HMH) and SOCS6 substrate peptide c-Kit is in green (PDB ID: 2VIF). (d) The binding of the non-canonical GHR (cyan) relies on backbone to backbone interaction and hydrophobic interaction from pY(+3) and pY(+5) of the GHR_pY595 to the SH2 domain hydrophobic core (pink). The binding mode of EpoR (orange) makes it unfavourable to establish these interactions.

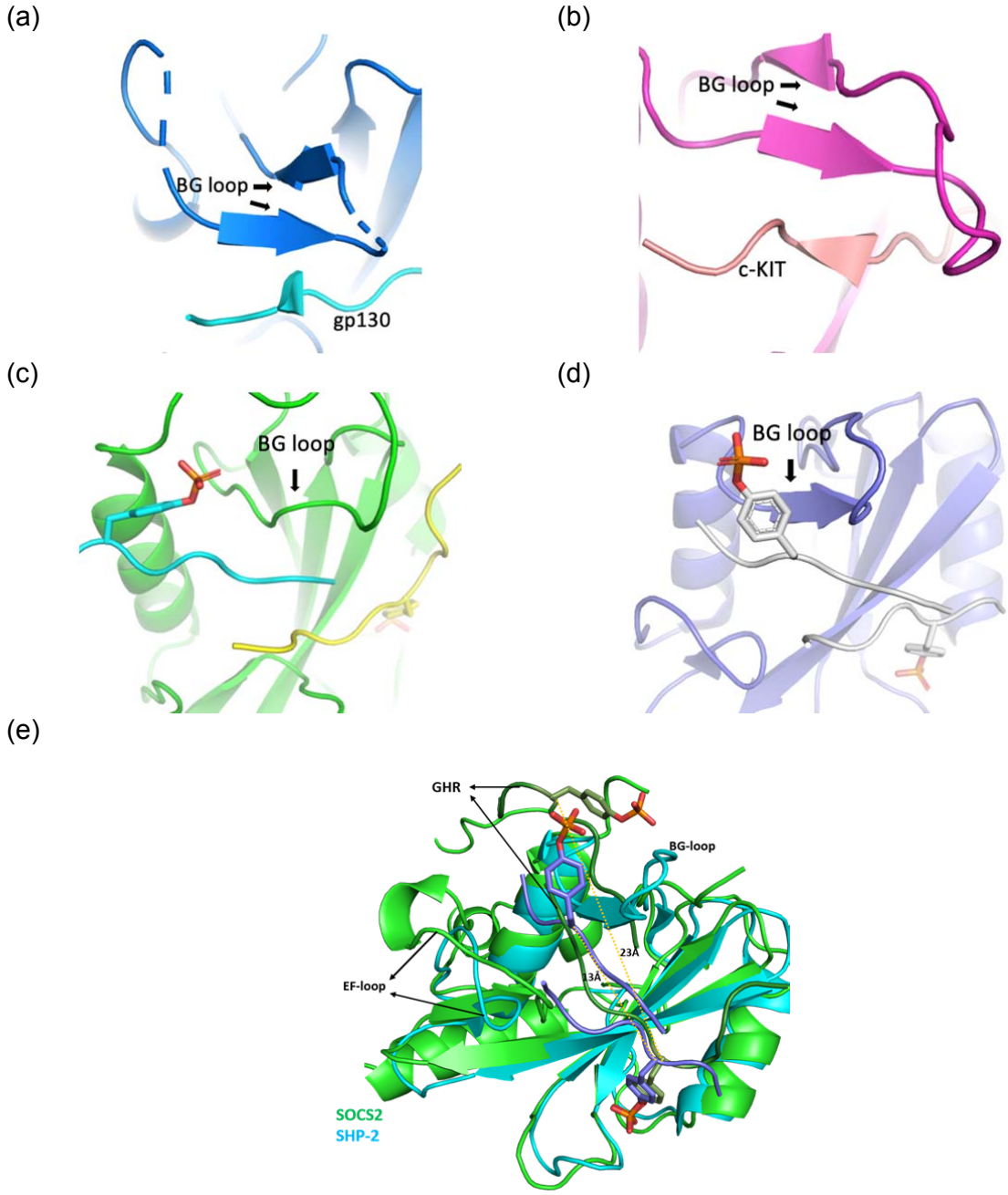


Figure 7. The triple-stranded β-sheet interaction between substrate peptide and protein

(a) The SOCS3 (blue) in complex with gp130 (cyan) (PDB code: 2HMH) (B) The SOCS6 (magenta) in complex with c-Kit (pink) (PBD code: 2VIF) (C) The SOCS2 (green) in complex with GHR_pY595 peptides. Peptide A is shown in yellow and peptide B in cyan. (d) The SHP2 (purple) in complex with peptides (white). (e) overlap of SH2 domains from SOCS2-GHR (green) and SHP2-pY peptide (cyan) highlights the difference in positioning of non-canonical pY from GHR and SHP2-substrate peptide

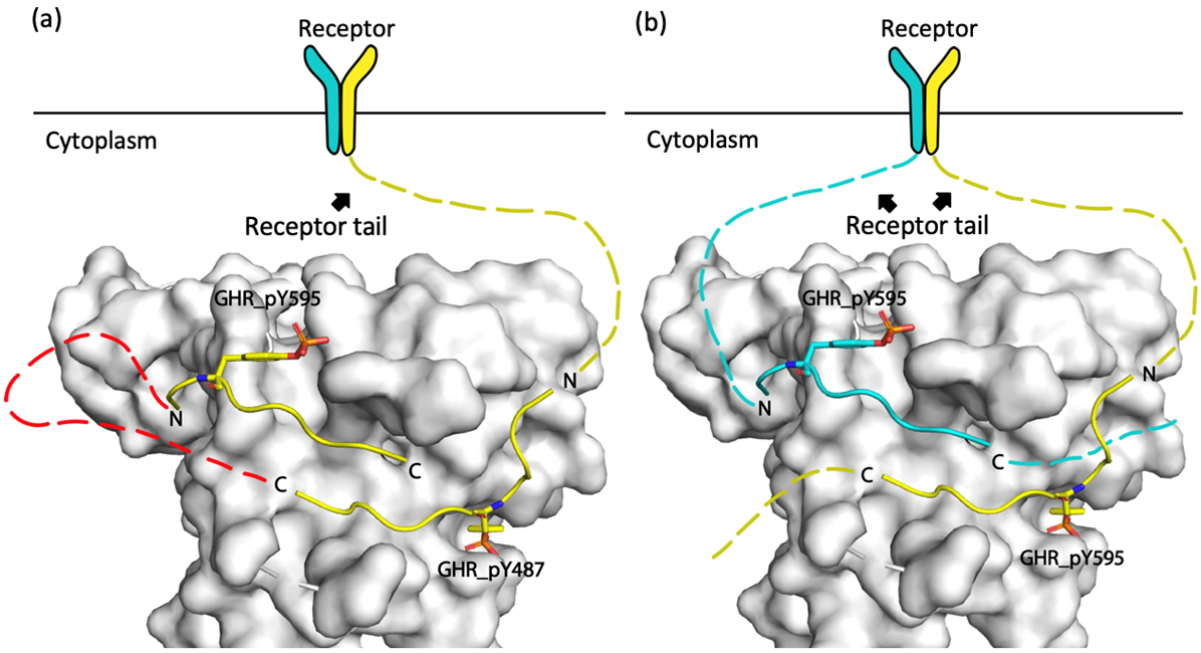


Figure 8. Illustration of GHR peptide binding modes

(a) Illustration of the "cis" recognition binding mode where the two binding sites on GHR, pY487 and pY595, folded into a hairpin structure for SOCS2 binding. SOCS2 is shown in white surface, the hairpin structure is illustrated by connecting the two peptides (yellow cartoon) with red dash lines. (b) Illustration of the "trans" binding mode where SOCS2 (white surface) could grab two receptor tails (cyan and yellow dash line) at the same time after dimerization of receptors at the membrane.

Tables

Table 1. Data collection and refinement statistics

	SBC-EpoR	SBC-GHR	SBC-GHR$_2$
PDB code	6I4X	6I5N	6I5J
Data collection			
Wavelength (Å)	0.9686	0.9795	0.9686
Space group	I 1 2 1	P 2_1 2_1 2	P 2 2_1 2_1
Cell dimensions			
a, b, c (Å)	41.29, 56.33, 203.39	113.18, 156.76, 57.57	57.83, 113.71, 156.94
α, β, γ (°)	90.00, 91.53, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00
Molecules/ASU	1	2	2
Resolution	29.36–2.69 (2.82–2.69)	113.55–1.98 (2.01–1.98)	92.08–2.80 (2.95–2.80)
R_{merge} (%)	10.8 (51.4)	9.4 (103.5)	19 (72.1)
$<I/\sigma(I)>$	9.6 (2.4)	17.1 (2.2)	7.3 (2.7)
Completeness (%)	93.2 (63.3)	100 (100)	100 (100)
Redundancy	4.9 (4.2)	13.3 (13.2)	7.8 (7.8)
$CC_{1/2}$	0.99 (0.82)	1.0 (0.9)	0.98 (0.74)
Refinement			
Resolution (Å)	2.69	1.98	2.8
Unique reflections	12265 (791)	72170(7105)	26275(2606)
R_{work}/R_{free} (%)	19.64/23.51	19.00/22.66	20.96/26.38
Wilson B factor (Å^2)	44.1	24.2	42.5
Average B factor (Å^2)	46.9	32.0	42.5
No. non-hydrogen atoms Protein/ligand/water	2688/83/11	6433/344/512	5775/327/33
R.M.S.D.			
Bond lengths (Å)	0.002	0.010	0.002
Bond angles (°)	0.451	1.064	0.412
Ramachandran analysis			
Preferred regions (%)	96.12	97.82	96.83
Allowed regions (%)	3.58	2.18	3.17
Outliers (%)	0.3	0.00	0.00

Values in parentheses are for the highest resolution shell

Table 2. Binding affinities of the GHR peptide and Val(-3) analogues

GHR peptide	Sequence	K_D by SPR (μM)	K_i by NMR (μM)
Wild type	PVPDpYTSIHIV	1.5 ± 0.09	1.5 ± 0.06
V(-3)R	P**R**PDpYTSIHIV	9.0 ± 1.05	20.6 ± 2.09
V(-3)Y	P**Y**PDpYTSIHIV	3.3 ± 0.13	3.3 ± 0.24

Residue mutated in bold
Values reported are the means ± s.e.m. from four independent experiments

Table 3. Affinity measurement for the GHR_pY595 and EpoR_pY426 derivatives by SPR and ^{19}F NMR

Sequence	Ala. position	SPR K_D ± SE (μM)	NMR K_i (μM)
PVPDpYTSIHIV	GHR wild type	1.5 ± 0.09	1.5 ± 0.06
AVPDpYTSIHIV	pY(-4)	1.7 ± 0.14	0.9 ± 0.17
P**A**PDpYTSIHIV	pY(-3)	2.7 ± 0.19	3.4 ± 0.25
PV**A**DpYTSIHIV	pY(-2)	1.4 ± 0.12	4.1 ± 0.75
PVP**A**pYTSIHIV	pY(-1)	6.2 ± 0.37	21.7 ± 3.34
PVPDpY**A**SIHIV	pY(+1)	1.0 ± 0.09	0.5 ± 0.07
PVPDpYT**A**IHIV	pY(+2)	1.4 ± 0.07	1.2 ± 0.06
PVPDpYTS**A**HIV	pY(+3)	3.7 ± 0.22	3.4 ± 0.66
PVPDpYTSI**A**IV	pY(+4)	6.7 ± 0.78	18.5 ± 2.67
PVPDpYTSIH**A**V	pY(+5)	2.4 ± 0.13	1.4 ± 0.36
PVPDpYTSIHI**A**	pY(+6)	1.9 ± 0.09	1.6 ± 0.21
ASFEpYTILDPS	EpoR wild type	13.3 ± 0.59	7.1 ± 1.01
A**A**FEpYTILDPS	pY(-3)	14.2 ± 0.65	7.5 ± 1.04
AS**A**EpYTILDPS	pY(-2)	16.3 ± 1.02	7.2 ± 0.42
ASF**A**pYTILDPS	pY(-1)	21.7 ± 1.49	23.5 ± 2.99
ASFEpY**A**ILDPS	pY(+1)	17.1 ± 0.38	10.3 ± 1.28
ASFEpYT**A**LDPS	pY(+2)	30.8 ± 1.24	17.2 ± 0.89
ASFEpYTI**A**DPS	pY(+3)	32.3 ± 1.42	20.4 ± 0.88
ASFEpYTIL**A**PS	pY(+4)	18.0 ± 0.91	14.1 ± 1.74
ASFEpYTILD**A**S	pY(+5)	8.9 ± 1.25	6.9 ± 1.39
ASFEpYTILDP**A**	pY(+6)	16.4 ± 0.29	11.3 ± 1.29

Residue mutated in bold
Values reported are the means ± s.e.m. from four independent experiments