

RESEARCH ARTICLE

Serendipitous SAD Solution for DMSO-Soaked SOCS2-ElonginC-ElonginB Crystals Using Covalently Incorporated Dimethylarsenic: Insights into Substrate Receptor Conformational Flexibility in Cullin RING Ligases

Morgan S. Gadd, Emil Bulatov[‡], Alessio Ciulli*

Division of Biological Chemistry and Drug Discovery, College of Life Sciences, University of Dundee, James Black Centre, Dow Street, Dundee, United Kingdom

[‡] Current Address: Research and Education Centre for Pharmacy, Kazan Federal University, Kazan, Russia
* a.ciulli@dundee.ac.uk



CrossMark
click for updates

OPEN ACCESS

Citation: Gadd MS, Bulatov E, Ciulli A (2015) 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(6): e0131218. doi:10.1371/journal.pone.0131218

Editor: Mark J van Raaij, Centro Nacional de Biotecnología—CSIC, SPAIN

Received: April 14, 2015

Accepted: May 29, 2015

Published: June 29, 2015

Copyright: © 2015 Gadd et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Coordinates have been deposited with the Protein Data Bank (accession number 5BO4).

Funding: This work was supported by European Research Council Starting Grant ERC-2012-StG-311460 DrugE3CRLs (to A. C.), the United Kingdom Biotechnology and Biological Sciences Research Council David Phillips Fellowship BB/G023123/1 (to A. C.), and the Government of the Republic of Tatarstan (Ph.D. studentship to E. B.). Molecular biophysics was supported by a Well- come Trust

Abstract

Suppressor of cytokine signalling 2 (SOCS2) is the substrate-binding component of a Cullin-RING E3 ubiquitin ligase (CRL) complex that targets phosphorylated hormone receptors for degradation by the ubiquitin-proteasome system. As a key regulator of the transcriptional response to growth signals, SOCS2 and its protein complex partners are potential targets for small molecule development. We found that crystals of SOCS2 in complex with its adaptor proteins, Elongin C and Elongin B, underwent a change in crystallographic parameters when treated with dimethyl sulfoxide during soaking experiments. To solve the phase problem for the new crystal form we identified the presence of arsenic atoms in the crystals, a result of covalent modification of cysteines by cacodylate, and successfully extracted anomalous signal from these atoms for experimental phasing. The resulting structure provides a means for solving future structures where the crystals must be treated with DMSO for ligand soaking approaches. Additionally, the conformational changes induced in this structure reveal flexibility within SOCS2 that match those postulated by previous molecular dynamics simulations. This conformational flexibility illustrates how SOCS2 can orient its substrates for successful ubiquitination by other elements of the CRL complex.

Introduction

In structural biology and structure-based drug design the existence of successfully solved and modelled crystallographic protein structures is critical to further advancements. These structures are important to understand how a protein functions and guide the development of small

strategic award 100476/Z/12/Z (to the University of Dundee). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

molecules targeting it. Additionally, for novel homologous or ligand-bound structures, the model itself provides the necessary information to solve the phase problem using molecular replacement (MR) or isomorphous refinement with the model as a template. These approaches rely on the novel structure to be sufficiently similar or isomorphous, respectively, such that the template is an adequate estimate of phases to enable the phase problem to be solved by these methods. In our crystallographic work with a complex of suppressor of cytokine signalling 2 (SOCS2), Elongin C (EloC) and Elongin B (EloB) we have encountered such a difficulty in solving the phase problem.

SOCS proteins such as SOCS2 are negative regulators of cytokine signalling of the Janus kinase/signal transducer and activator of transcription (JAK-STAT) pathway [1]. The SOCS proteins each have a similar domain architecture. They contain a central SH2 domain that specifically recognizes phosphorylated tyrosine residues, a hallmark of cytokine signalling [2]. SOCS proteins also have a C-terminal SOCS box that comprises a BC box that recruits EloB and EloC (EloBC) [3–4] and a Cul5 box that recruits Cullin 5 (Cul5) [5]. All of the SOCS family proteins bind to EloBC using their SOCS box, and both of these are required to recruit Cul5 [6]. Cul5, EloB and EloC are common components of Cul5-based cullin-RING ligases (CRLs).

CRLs are multi-protein complexes that act as E3 ubiquitin ligases [7–8]. They recruit an E2 conjugating enzyme, Cdc34/Ube2r1, carrying ubiquitin and via a substrate-binding protein, such as SOCS2, selectively transfer the ubiquitin to their substrate targets [9]. Using its SH2 domain, SOCS2 interacts with a number of proteins, including growth hormone receptor (GHR), insulin-like growth factor-I receptor, erythropoietin receptor, leptin receptor and epidermal growth factor receptor, when specific tyrosine residues on these proteins are phosphorylated [10–14]. SOCS2 has been shown to bind GHR as a substrate for ubiquitination and thus drive its degradation by the proteasome [15]. Other receptors interacting with SOCS2 are likely to be regulated in the same manner. These targets indicate there may be key roles for SOCS2 in somatic growth, the central nervous system, metabolic regulation, innate immunity and cancer [16]. Thus therapeutic developments targeting the activity of SOCS2 may lead to beneficial outcomes in a range of physiological and pathophysiological processes.

We have identified and validated a number of small molecule binders to a SOCS2:EloC:EloB complex using biophysical approaches. We sought to elucidate the binding sites and binding modes of these small molecules using x-ray crystallography. Despite successfully producing crystals of the complex that were isomorphous to those reported previously (PDB: 2C9W) [17], we found that when soaked with dimethyl sulfoxide (DMSO) alone and DMSO solutions of the compounds these crystals undergo a change in spacegroup and unit cell dimensions. As described above, this change resulted in difficulties that prevented solving the phase problem using MR approaches. We thus sought an alternative means of obtaining phases for the structure of DMSO-treated crystals.

We identified the presence of arsenic atoms in the 2C9W crystal structure as a result of covalent modification of surface cysteine residues with cacodylate buffer. We report here the successful resolution of the phase problem for DMSO-treated SOCS2:EloC:EloB crystals using arsenic single-wavelength anomalous dispersion (As-SAD). The structure reveals significant conformational changes that have resulted from the DMSO treatment that caused the decreased crystal symmetry. These conformational changes illustrate the hinging motion between the two domains of SOCS2 (as well as other SOCS box and F-box containing proteins) that may help to accurately position substrates for ubiquitination by CRL complexes, consistent with previous suggestions based on molecular dynamics simulations.

Results

In our crystallographic efforts on SOCS2:EloC:EloB we were unable to reproduce the crystallization condition previously reported by Bullock *et al.* [17] containing 0.08 M sodium cacodylate (pH 6.5), 1.6 M ammonium sulfate and 0.16 M sodium chloride, with the protein complex in 50 mM Hepes, pH 7.5, 250 mM NaCl, 2.5% (v/v) glycerol and 10 mM dithiothreitol (DTT). Consequently, we attempted our own sparse-matrix screening to identify more readily reproducible crystallization conditions. Crystals were observed in Molecular Dimensions JCSG-*plus* condition 2.11 containing 0.08 M sodium cacodylate, pH 6.5, 0.16 M calcium acetate, 14.4% (w/v) PEG8000 and 20% (v/v) glycerol with the protein complex in 25 mM Hepes, pH 7.5, 250 mM NaCl and 10 mM DTT. Further refinement of the protein complex sample in this solution yielded single crystals that grow up to 500 μm in length in conditions of 0.08 M sodium cacodylate pH 7.2, 0.16 M calcium acetate, 14.5% (w/v) PEG3350 and 20% (v/v) glycerol (Fig 1).

Diffraction data were collected on a single crystal using Cu $K\alpha$ radiation as described. Indexing and merging of the data indicated the spacegroup to be $P3_221$ and the unit cell dimensions to be $a = b = 105.5$, $c = 69.4$, containing one instance of the SOCS2:EloC:EloB complex in the asymmetric unit (ASU). These indexing parameters are equivalent to those described previously [17]. The structure was solved to 2.6 \AA resolution by MR using 2C9W as the search model and refined to an R_{work} of 23.6% and R_{free} of 30.4%. The resulting structure revealed no significant differences to that of 2C9W (RMSD = 0.40 \AA ; data not shown), indicating that both crystallization conditions yield crystals of the same form.

Having successfully produced SOCS2:EloC:EloB crystals we sought to utilize them to elucidate the structural details of small molecule binders identified during screening efforts (to be reported elsewhere). During these investigations we found that, having been soaked with the small molecule solutions in DMSO, the diffraction patterns of the crystals could no longer be indexed with the same space group and unit cell parameters as reported above. These data for DMSO-treated SOCS2:EloC:EloB crystals were instead successfully indexed with the alternative space group $P3_2$ and approximate unit cell dimensions $a = b = 186$ \AA , $c = 67$ \AA . The change in unit cell dimensions and space group represents a six-fold increase in the size of the ASU, thus containing six SOCS2:EloC:EloB complexes instead of a single one.

To solve the phase problem for the new crystallographic parameters, diffraction data collected at resolutions from 3.3–3.8 \AA were used in MR calculations (data not shown).

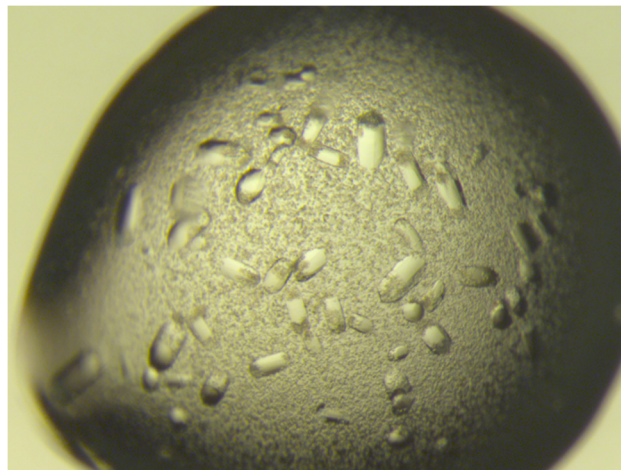


Fig 1. Crystals of the SOCS2:EloC:EloB complex grown after 7 d. The crystals shown here are 200–300 μm in the longest dimension.

doi:10.1371/journal.pone.0131218.g001

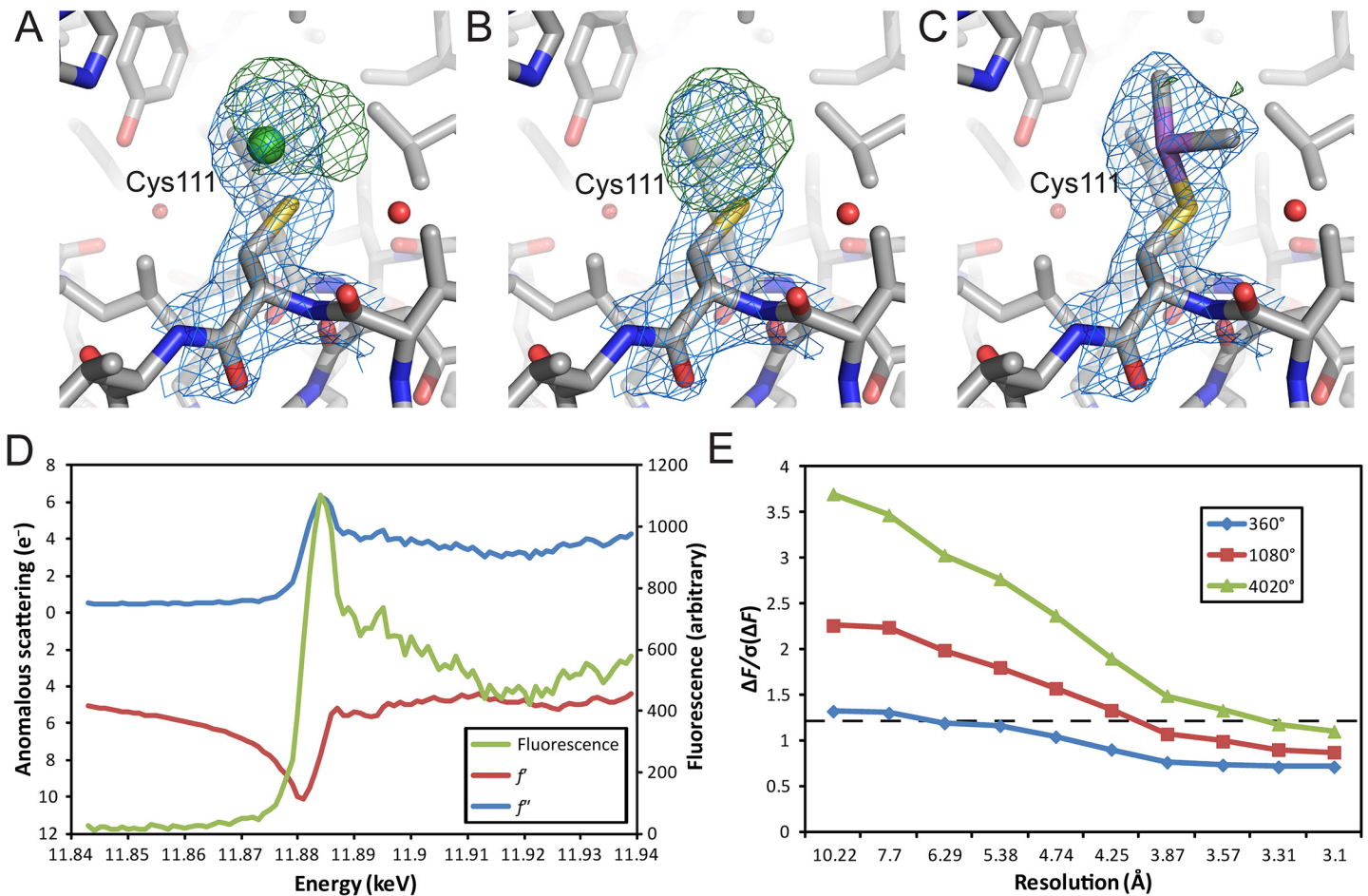


Fig 2. Identification of arsenic as a source of experimental phasing power. (A) The $2F_o - F_c$ (blue mesh) and $F_o - F_c$ (green mesh) maps around modelled $SOCS2_{C111}$ (CPK colored sticks) and adjacent Ni(II) (green sphere) from the PDB structure 2C9W. (B) As in A, with maps recalculated following removal of Ni(II). (C) As in A, with $SOCS2_{C111}$ remodelled as dimethylarsenic cysteine. $2F_o - F_c$ maps are contoured at 1σ and $F_o - F_c$ at 3σ . (D) A fluorescence scan (green line) and measurements of f'_{\neq} and $f'_{=}$ (red and blue lines, respectively) identifying the presence of arsenic and resulting anomalous signals from DMSO-treated SOCS2:EloC:EloB crystals. (E) Observed anomalous signal $[\Delta F/\sigma(\Delta F)]$ in diffraction data plotted against resolution for the As-Peak datasets collected. The cut-off of useful signal (1.2) is indicated by a dashed line [22].

doi:10.1371/journal.pone.0131218.g002

Unfortunately a solution could not be found despite multiple attempts with a variety of different methods, including efforts to separately place SOCS2 and EloBC, as had been used to solve similar structures previously [17–20]. The difficulty of these MR calculations is most probably because of the six complexes (18 protein chains) to place in the rotation-translation function, but may be exacerbated by potential conformational changes that have driven this loss in crystallographic symmetry, the lower resolution of the diffraction data, as well as the fact that the complex is not a packed, globular unit.

In spite of these challenges we sought an alternative way to solve the phase problem for this structure. In the deposited model for 2C9W there are three Ni(II) ions modelled around the single complex in the ASU. Careful analysis of structure and electron density using the deposited structure factors for 2C9W suggested that the two Ni(II) ions modelled as being bound to cysteine residues ($SOCS2_{C111}$ and $EloBC_{89}$) inadequately satisfy the electron density in these parts of the structure (Fig 2A). Having removed the Ni(II) ions we identified that a covalent modification to the cysteine residues was a more likely explanation for the additional electron density (Fig 2B). By re-examining the crystallization and sample buffer conditions we

determined that the presence of cacodylate and dithiothreitol (DTT) may have resulted in dimethylarsenic modifications to these thiol residues. Such a covalent modification was first observed and characterized by the reactivity of As(III)-thiolate intermediates (generated from DTT or β -mercaptoethanol) with the *EcoRI* methylase complex [21]. Following remodelling and refinement we observed that these modified sidechains fully satisfied the additional electron density at these cysteine residues (Fig 2C).

Arsenic sits adjacent to selenium in the periodic table and is a potential source of anomalous signal for experimental phasing. We performed a fluorescence scan to confirm the presence of arsenic and identified a wavelength of 11.8847 keV ($\lambda = 1.04323 \text{ \AA}$) for maximum production of f' and f'' from the arsenic peak (Fig 2D). In an effort to conduct single-wavelength anomalous dispersion (SAD) phasing for the crystals we collected a series of datasets with increasing redundancy (Fig 2E). Due to the weakness of the observed anomalous signal, ultimately a dataset of 4020 degrees (redundancy ≈ 120) was used to ensure a SAD solution could be achieved (Table 1).

Experimental phases for the structure of DMSO-treated SOCS2:EloC:EloB were determined by SAD using diffraction data collected at 11.8847 keV to 3.1 \AA resolution. 12 individual arsenic peaks were found in the resulting experimental map. Automated model building produced a series of chains representing an incomplete, fragmented structure with an R_{work} of 37.9% and R_{free} of 43.4%. The experimental map and partial structure were used to align regions of the template structure 2C9W to place the six copies of the complex in the ASU. The 12 unique arsenic peaks were confirmed as sites of cacodylate modifications on cysteine residues (Fig 3A). Peaks for the six SOCS2_{C111} residues can still be observed in the anomalous difference map at approximately 20σ , whereas those for EloB_{C89} can only be observed at 5σ .

The six complexes in the asymmetric unit were refined against a native dataset processed to 2.9 \AA resolution. The final model was refined to an R_{work} value of 22.9% and R_{free} of 28.7% (Table 1). Each complex in the model maintains the same overall structural assembly observed in 2C9W, however there are also differences. In EloB all residues but the initiating methionine are observed in the 2C9W structure. In our structure, however, residues 18 and 19 and six C-terminal residues of chain E, as well as residues 15–19, 32–37 and 78–96 of chain N are not visible in the electron density. In EloC residues 46–57 and 85–88 are disordered and not visible in the 2C9W structure, however the final loop (85–88) is traceable in the electron density within chains C, I and L of DMSO-treated SOCS2:EloC:EloB.

When aligned via the backbone atoms of the EloB proteins, significant deviations in the gross structure of the complex can be observed by analysing SOCS2 on the opposite side of the complex (Fig 3B). These deviations are well illustrated by residues 92–100 of SOCS2 when all complexes are aligned via the backbone atoms of EloB (Fig 3C). When analysing each component of the DMSO-treated SOCS2:EloC:EloB structure in this alignment we find that EloB, EloC and the SOCS2 box of SOCS2 are well aligned (average RMSD of backbone atoms = 0.62 ± 0.25), whereas significant deviations are observed for the SH2 domain of SOCS2 (average RMSD of backbone atoms = 2.1 ± 0.9) (Fig 3D).

Discussion

Experimental phasing is a key element of crystal structure resolution, not only because it can provide a solution to the phase problem where MR proves insufficient, but also it eliminates phase bias in any successful MR solution. Sources of anomalous signal in native crystals include sulfur or phosphorous atoms, co-ordinated metal ions or unnatural amino acids, such as selenomethionine. Each of these approaches has drawbacks, for example sulfur-SAD requires precisely measured reflections to a reasonable resolution [23], metal ions are restricted to those

Table 1. X-ray diffraction data and model statistics for DMSO-treated SOCS2:EloC:EloB.

Scaling statistics	As-Peak	Native
Wavelength (Å)	1.04323	0.95372
Space group	$P3_2$	$P3_2$
Unit-cell parameters (Å)	$a = b = 185.9, c = 67.3$	$a = b = 185.2, c = 67.2$
Resolution limits (Å)	54.5–3.1 (3.27–3.10)	46.3–2.9 (2.98–2.90)
Completeness (%)	100 (100)	99.2 (100)
Wilson B factor (Å ²)	64.2	59.6
Observed reflections	5659871 (744593)	219197 (17340)
Unique reflections	47227 (6906)	56676 (4449)
Redundancy	119.8 (107.8)	3.9 (3.9)
R_{merge}^a (%)	12.0 (74.4)	6.7 (62.2)
$I/\sigma(I)$	52.6 (13.4)	15.6 (2.3)
CC _{1/2} (%)	100 (98.9)	99.9 (67.3)
Phasing statistics		
FOM _{PHASER}	0.311	
FOM _{RESOLVE}	0.310	
Model refinement		
R_{work}^b (%)		22.9 (33.5)
R_{free}^c (%)		28.7 (41.5)
No. of reflections used in refinement		53812 (4013)
No. of reflections in the test set		2862 (203)
Protein atoms		14871
Solvent molecules		33
RMSD bond length (Å)		0.006
RMSD bond angles (°)		1.01
Mean protein B factor (Å ²)		77.5
Ramachandran plot, residues in		
Favored regions (%)		94.1
Allowed regions (%)		5.7
Disallowed regions (%)		0.2

Values in parentheses are for the highest resolution shell.

Values for the highest resolution shell are given in parentheses.

$$^a R_{\text{merge}} = \frac{\sum_{hkl} \sum_i (|I_i(hkl)| - [I(hkl)])}{\sum_{hkl} \sum_i I_i(hkl)}$$

^b $R_{\text{work}} = |F_{\text{obs}} - F_{\text{calc}}| / |F_{\text{obs}}|$, where F_{obs} and F_{calc} are the observed and calculated structure amplitudes, respectively.

^c R_{free} is R_{work} for the 5% validation set.

doi:10.1371/journal.pone.0131218.t001

proteins that bind them such as metalloproteins, and selenomethionine incorporation is not always practical. In this case the phasing atom, arsenic, has been incorporated as an essential step in the crystallization process, and thus the native crystals facilitate anomalous phasing.

Arsenic is relatively underused as a potential source of anomalous signal for phasing. To date only five structures have been solved using the atom as a source of anomalous signal and deposited in the PDB, four of which were a product of cacodylate incorporation [24–27]. Compared to isomorphous replacement methods that soak heavy atoms, this approach results in only minor changes to the protein surface and happens during the crystallization process, and thus is more likely to facilitate crystal growth rather than inhibit it or disrupt grown crystals. Indeed, in our experience both SOCS2:EloC:EloB and a similar complex where SOCS2 is

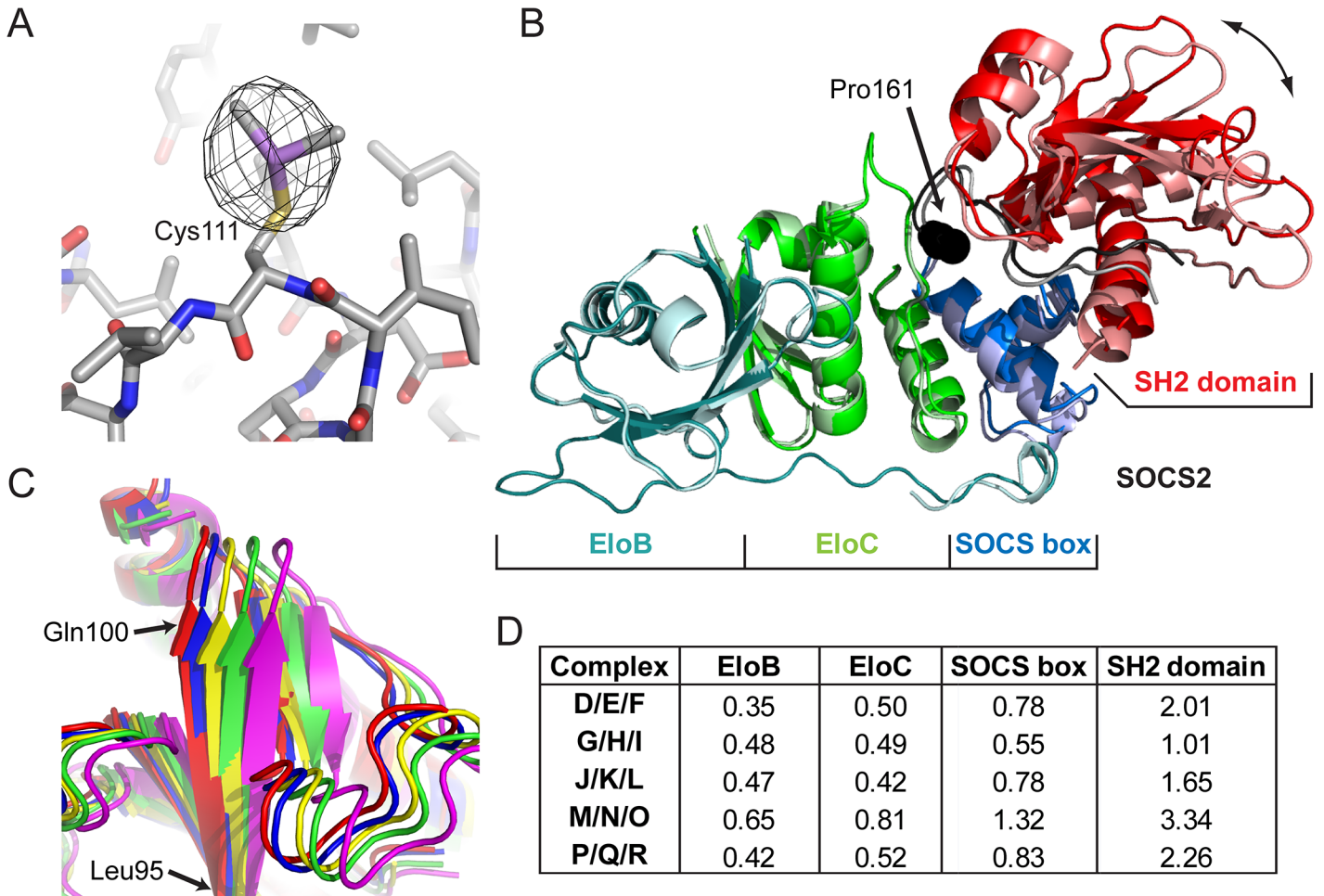


Fig 3. The structure of DMSO-treated SOCS2:EloC:EloB. (A) Anomalous difference map (black mesh) shown at 5σ calculated from As-Peak diffraction data in proximity to SOCS2_{C111} residue of Chain J (CPK colored sticks) with the dimethylarsenic modification. (B) Structure of the two most conformationally different complexes of DMSO-treated SOCS2-EloBC (A/B/C and M/N/O), aligned via the backbone atoms of EloB. Protein chains are shown as ribbons, with EloB (cyan), EloC (green), SOCS2 SOCS box (blue) and SOCS2 SH2 domain (red). Chains A/B/C are in darker colors and M/N/O lighter. The linker/hinging region between SOCS2 SOCS box and SH2 domain is shown in black/grey and residue Pro161 as spheres. (C) SH2 domains of DMSO-treated SOCS2-EloBC illustrating the five different conformations observed in the ASU. Chains A/B/C (red), D/E/F (green), G/H/I (blue), J/K/L (yellow) and M/N/O (magenta). Complexes have been aligned via the backbone atoms of EloB from each complex and superposed. Complex P/Q/R is not shown as its conformation is identical to that of D/E/F. (D) Root mean square deviation (RMSD) values (in Å) for the backbone atoms of EloB, EloC, SOCS2 SOCS-box domain (residues 162–198) and SOCS2 SH2 domain (32–134) when complexes are superposed as in C. RMSD values are between each complex and complex A/B/C.

doi:10.1371/journal.pone.0131218.g003

substituted by the von Hippel–Lindau tumor suppressor (pVHL) require both DTT and cacodylate to be present in the crystallization experiment for successful crystal growth [28]. For the crystallization of the catalytic domain of HIV-1 integrase the presence of both DTT and cacodylate was also observed to be an absolute requirement, and a mechanism by which the reaction took place was proposed (Fig 4) [29].

In the structure presented here the modified residue SOCS2_{C111} has the equivalent or lower thermal displacement (*B* factors) on the arsenic and methyl atoms than other atoms in the side-chains from the six complexes. On the other hand EloB_{C89} has significantly higher *B* factors for the arsenic and methyl atoms. This likely corresponds with the observation that SOCS2_{C111} contacts a surface formed by residues 87–91 of an adjacent SOCS2 molecule on a two-fold symmetry axis and is observable at 20σ in the anomalous difference map, whereas EloB_{C89} forms

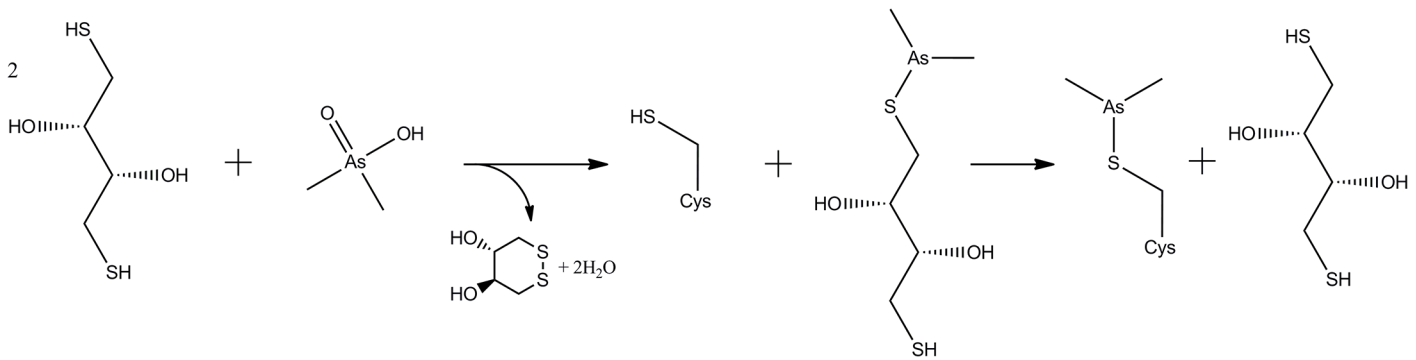


Fig 4. Mechanism for dimethylarsenic modification of cysteine residues proposed by Maignan *et al.* [29]. DTT reduces As(V) of cacodylate in its acid form ($pK_a = 6.3$) into an As(III)-containing dimethylarsenic-dithiothreitol conjugate. This adduct then reacts with a reduced cysteine sidechain to result in dimethylarsenic cysteine and a replenished molecule of DTT.

doi:10.1371/journal.pone.0131218.g004

no similar crystal contacts and the arsenic atom is observable only at 5σ . SOCS2_{C111} therefore forms a highly ordered contact that may explain the beneficial effect of the dimethylarsenic modifications to the growth of these crystals. Thus the modification with cacodylate essentially generates a new sample of the same protein with a different potential for the formation of crystal contacts and thus successful crystallization.

To date 159 structures in the Protein Data Bank contain dimethylarsenic cysteine or dimethylarsinoyl cysteine (the two modifications that can result from cacodylate incorporation) as modelled residues. The usefulness of this buffer component for both potential phasing purposes and facilitating crystallization would suggest that it is extremely underutilized. Anecdotally this would appear to be because of the toxicity of cacodylate due to its reactivity with thiols [21, 30]. However, cacodylate is a common buffer component in the more popular sparse-matrix crystallization screens, and thus will continue to be identified as a potential crystallization hit [31–32].

Our structure of DMSO-treated SOCS2:EloC:EloB illustrates important features of conformational dynamics within this important regulatory complex. Differences between the individual complexes in the ASU of the structure are well illustrated when a superposition is performed via the backbone atoms of the EloB molecules (Fig 3B, 3C and 3D). In all cases the chains of EloB, EloC and the SOCS-box domain of SOCS2 are well aligned, with low RMSD values, indicating these parts of the complex form a rigid unit in relation to one another. In contrast the range of RMSD values for the SH2 domain of SOCS2 indicates a significant degree of conformational freedom between this domain and the remainder of the complex.

Molecular dynamics simulations have suggested that the highly conserved Pro161 residue of SOCS2, which lies just before the SOCS box (162–198), is a key component in imparting conformational flexibility at this hinge region (Fig 3B) [33]. The ensemble of models found in the DMSO-treated SOCS2:EloC:EloB structure reinforces this observation with the substrate-binding SH2 domain of SOCS2 alone exhibiting a clear hinging motion relative to the remainder of the complex (Fig 3B and 3C), consistent with the computational predictions [33]. A similar effect is observed when comparing the crystal structures of SOCS2:EloC:EloB when bound and not bound to Cul5 [17, 34]. However, the observed flexion between the SOCS box and SH2 domain in the Cul5-bound structure lies outside the range observed in the ensemble of the DMSO-treated SOCS2:EloC:EloB structure.

The computationally observed hinging motion is replicated in other SOCS-box and F-box substrate-binding proteins, such as pVHL and S phase kinase-associated protein 2 (Skp2), and appears to be regulated by the binding of substrate and/or adaptor proteins [33, 35]. A

similar analysis to that above with published crystal structures of these proteins reveals similar effects. For the pVHL:EloC:EloB complex the hinging motion of the substrate-binding domain ranges with motion at one end for the pVHL:EloC:EloB:HIF-1 α complex [36–37], then the four protomers of the pVHL:EloC:EloB apo complex [38] and at the other extreme the pVHL:EloC:EloB:Cul2 structure [39]. All seven of these structures show a similarly rigid alignment of EloB, EloC and the SOCS box of pVHL. When aligned via Skp1, the published structures of F-box containing proteins Skp2 and cell division control protein 4 (Cdc4) reveal the same rigidity between Skp1 and the F-box domain, but relative hinging motion for the leucine-rich repeat and WD40 repeat like domains in Skp1 and Cdc4, respectively [40–42]. As a common feature in SOCS-box and F-box containing proteins, the hinging motion and its regulation by substrate- and adaptor-protein binding is thus postulated to be critical for ensuring accurate positioning and orientation of the target substrate [33, 35]. These motions tie into global motions identified in CRL complexes that bring the substrate into proximity for ubiquitination by Cdc34 at the other end of the CRL complex [43].

The observed hinging motion between the SH2 domain and the remainder of the complex is likely to be a key driver of the change in crystal form from 2C9W to the structure presented here, as each complex (except for D/E/F and P/Q/R) has a slightly different orientation of the SH2 domain. The degree of disorder at unstructured loops in each complex differs (e.g. chain N lacks a total of 28 residues found in every other chain of EloB) and is likely to be another factor in the loss of crystallographic symmetry within the crystals. Due to these rather significant conformational changes the structure was not solvable using the data initially available to us (lower than 3.3 Å resolution) by MR using 2C9W or its individual components as search models. Experimental phasing using the arsenic atoms of the dimethylarsenic modifications was critical in providing a solution to the phase problem for this structure. While we were eventually successful in generating a MR solution from our subsequently collected data to 2.9 Å resolution, this structure revealed a more limited range of conformations which do not reflect the extent of hinging motions observed in our experimentally phased structure, suggesting phase bias in the former.

When protein crystals are soaked in solutions with molecules not found in the crystallization experiment they will sometimes be damaged or destroyed altogether, potentially cracking, losing order or resolubilising, particularly in the presence of organic solvents [44]. In a case such as this the disruption of the crystal lattice has been limited enough to maintain useful crystallographic diffraction, albeit with reduced crystallographic symmetry within the crystal. Similar effects to crystal properties have been observed previously. In the literature we identified one case where the addition of new molecules (NaBr) in a soaking condition for a peroxidase resulted in a decrease in crystallographic symmetry, from $I4$ to $P4_2$ [45]. On the other hand three cases of increased symmetry were identified: for crystals of the DNA mismatch repair protein MutS the presence of $MnCl_2$ in the soaking solution results in an increase in symmetry from $P2_1$ to $P2_12_12_1$ and increased resolution of diffraction [46]; soaking of glutathione S-transferase crystals with zinc protoporphyrin resulted in a change of spacegroup from $P2_12_12_1$ to $C222_1$ [47]; and soaking again with NaBr on crystals of a pepstatin-insensitive carboxyl proteinase resulted in a change of symmetry from $P2_12_12_1$ to $P6_122$ [48]. In all cases these soaking atoms were not visible in the electron density map (although in the latter case Br atoms were weakly visible in an anomalous difference map), suggesting that their influence on the physical properties of the soaking conditions was likely to be driving the observed alterations in crystal symmetry.

When crystal-soaking conditions alter but do not destroy the crystal properties and thus resultant x-ray diffraction, it is important to salvage such cases where possible. While typically a pre-existing solved structure should enable the phase problem to be addressed using MR, this

was not the case here. The change in space group and unit cell dimensions in the presence of DMSO have resulted in a large increase in the number of protein molecules, as well as general disordering of the complex, making it very difficult to establish an MR solution with any ease. Thus, this structure provides a key avenue for the solution of ligand-bound structures of SOCS2:EloC:EloB where DMSO (and potentially other solvents) is required to solubilize ligands for successful soaking experiments. In turn these ligand-bound structures will be critical for the future development of chemical tools or potential small molecule therapeutics that modulate SOCS2-based CRL complexes.

Materials and Methods

Protein expression, purification and crystallization

SOCS2:EloC:EloB was expressed and purified as described previously [49]. Briefly, the complex was co-expressed in *Escherichia coli* BL21(DE3) from the pLIC (His₆-SOCS2) and pCDF (EloBC) plasmids. The sequences are: human SOCS2 (residues 32–198; a gift from A. Bullock, Structural Genomics Consortium, Oxford, UK); EloB (amino acids 1–104); and EloC (amino acids 17–112). Protein expression was induced with isopropyl β-D-1-thiogalactopyranoside at 18°C for 12 h. SOCS2:EloC:EloB was extracted from lysed cells in the soluble fraction and purified by affinity chromatography using a HisTrap column (GE Healthcare). The His₆-tag fused to SOCS2 was removed by cleavage with tobacco etch virus protease and the complex reapplied to the HisTrap column as a non-binding species. Finally the complex was purified by size-exclusion chromatography on a Superdex 75 16/600 column (GE Healthcare) in a buffer of 25 mM HEPES, pH 7.5, 250 mM NaCl and 10 mM DTT.

Crystallization conditions for SOCS2:EloC:EloB were identified using commercially available sparse-matrix screens (Qiagen and Molecular Dimensions). A 96-channel Phoenix high-speed liquid handling system (Art Robbins Instruments) was used to set up crystallization experiments using the sitting-drop vapor-diffusion method at room temperature. Crystallization drops were set up in ratios of 1:1 or 1:2 protein:liquor, respectively, and incubated at room temperature. Crystal formation and growth were monitored manually or by automated imaging using a Rock Imager 1000 (Formulatrix). Optimization of crystallization conditions was performed manually in 2-μl drops using the hanging- and sitting-drop vapor-diffusion methods at room temperature.

X-ray diffraction data collection

SOCS2:EloC:EloB crystals were flash-frozen in liquid nitrogen without cryoprotection. For determining isomorphism with 2C9W, diffraction data were recorded on a Saturn 944HG + CCD detector using x-rays produced by a Rigaku M007HF generator (Cu Kα). To obtain anomalous dispersion data, DMSO-treated SOCS2:EloC:EloB crystals were also flash-frozen without cryoprotection. Diffraction data were collected on BM14 at the European Synchrotron Radiation Facility at 100 K on a single crystal. A fluorescence scan was performed to identify the presence of arsenic atoms and determine the wavelength for maximum dispersive effects. The As-Peak dataset was collected at a wavelength of 1.04323 Å for 4020 frames with an oscillation angle of 1° per frame. The native dataset was collected at a wavelength of 0.95372 Å for 120 frames with an oscillation angle of 1° per frame.

Experimental phasing, model building and refinement

Diffraction data were indexed and integrated using Mosflm [50] and scaled and merged with Aimless in CCP4 [51–52]. The structure solution pipeline AutoSol in the PHENIX software

suite [53–54] was used to establish experimental phases by identifying arsenic atom positions and generate initial phases (PHASER) [55], then to perform density modification, solvent flattening and automated model building into the resulting map (RESOLVE) [56]. The incomplete model generated by RESOLVE was completed with the aid of the previously published SOCS2:EloC:EloB structure by a series of alignments and manual building in Coot [57]. The resulting structure was refined with REFMAC5 [58] using isotropic temperature factors, TLS groups and NCS restraints. The MOLPROBITY server [59] was used to validate the structure by identifying steric clashes and geometric problems. The coordinates and structure factors of DMSO-treated SOCS2:EloC:EloB have been deposited in the Protein Data Bank with the PDB code 5BO4.

Acknowledgments

We thank Dr. Paul Fyfe for support with the in-house x-ray facility in the College of Life Sciences at the University of Dundee. We thank the EMBL staff Dr. Hassan Belrhali and Dr. Babu A. Manjasetty for providing support on the beamline and EMBL-DBT for providing access to the BM14 beamline at the ESRF (proposal mx1481).

Author Contributions

Conceived and designed the experiments: MSG EB AC. Performed the experiments: MSG EB. Analyzed the data: MSG EB AC. Wrote the paper: MSG AC.

References

1. Starr R, Willson TA, Viney EM, Murray LJ, Rayner JR, Jenkins BJ, et al. A family of cytokine-inducible inhibitors of signalling. *Nature*. 1997 Jun 26; 387(6636):917–21. PMID: [9202125](#)
2. Yoshimura A, Ohkubo T, Kiguchi T, Jenkins NA, Gilbert DJ, Copeland NG, 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*. 1995 Jun 15; 14(12):2816–26. PMID: [7796808](#)
3. Kamura T, Sato S, Haque D, Liu L, Kaelin WG Jr., Conaway RC, 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*. 1998 Dec 15; 12(24):3872–81. PMID: [9869640](#)
4. Zhang JG, Farley A, Nicholson SE, Willson TA, Zugaro LM, Simpson RJ, 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*. 1999 Mar 2; 96(5):2071–6. PMID: [10051596](#)
5. Kamura T, Maenaka K, Kotoshiba S, Matsumoto M, Kohda D, Conaway RC, et al. VHL-box and SOCS-box domains determine binding specificity for Cul2-Rbx1 and Cul5-Rbx2 modules of ubiquitin ligases. *Genes Dev*. 2004 Dec 15; 18(24):3055–65. PMID: [15601820](#)
6. Babon JJ, Sabo JK, Zhang JG, Nicola NA, Norton RS. The SOCS box encodes a hierarchy of affinities for Cullin5: implications for ubiquitin ligase formation and cytokine signalling suppression. *J Mol Biol*. 2009 Mar 20; 387(1):162–74. PMID: [19385048](#)
7. Willems AR, Schwab M, Tyers M. A hitchhiker's guide to the cullin ubiquitin ligases: SCF and its kin. *Biochim Biophys Acta*. 2004 Nov 29; 1695(1–3):133–70. PMID: [15571813](#)
8. Bulatov E, Ciulli A. Targeting Cullin–RING E3 ubiquitin ligases for drug discovery: structure, assembly and small-molecule modulation. *Biochem J*. 2015 May 1; 467(3):365–86. doi: [10.1042/BJ20141450](#) PMID: [25886174](#)
9. Feldman RM, Correll CC, Kaplan KB, Deshaies RJ. A complex of Cdc4p, Skp1p, and Cdc53p/cullin catalyzes ubiquitination of the phosphorylated CDK inhibitor Sic1p. *Cell*. 1997 Oct 17; 91(2):221–30. PMID: [9346239](#)
10. Dey BR, Spence SL, Nissley P, Furlanetto RW. Interaction of human suppressor of cytokine signaling (SOCS)-2 with the insulin-like growth factor-I receptor. *J Biol Chem*. 1998 Sep 11; 273(37):24095–101. PMID: [9727029](#)
11. Eyckerman S, Verhee A, der Heyden JV, Lemmens I, Ostade XV, Vandekerckhove J, et al. Design and application of a cytokine-receptor-based interaction trap. *Nat Cell Biol*. 2001 Dec; 3(12):1114–9. PMID: [11781573](#)

12. Lavens D, Montoye T, Piessevaux J, Zabeau L, Vandekerckhove J, Gevaert K, et al. A complex interaction pattern of CIS and SOCS2 with the leptin receptor. *J Cell Sci.* 2006 Jun 1; 119(Pt 11):2214–24.
13. Goldshmit Y, Walters CE, Scott HJ, Greenhalgh CJ, Turnley AM. SOCS2 induces neurite outgrowth by regulation of epidermal growth factor receptor activation. *J Biol Chem.* 2004 Apr 16; 279(16):16349–55. PMID: [14764607](#)
14. Greenhalgh CJ, Rico-Bautista E, Lorentzon M, Thaus AL, Morgan PO, Willson TA, et al. SOCS2 negatively regulates growth hormone action in vitro and in vivo. *J Clin Invest.* 2005 Feb; 115(2):397–406. PMID: [15690087](#)
15. Vesterlund M, Zadjali F, Persson T, Nielsen ML, Kessler BM, Norstedt G, et al. The SOCS2 ubiquitin ligase complex regulates growth hormone receptor levels. *PLoS One.* 2011; 6(9):e25358. doi: [10.1371/journal.pone.0025358](#) PMID: [21980433](#)
16. Rico-Bautista E, Flores-Morales A, Fernandez-Perez L. Suppressor of cytokine signaling (SOCS) 2, a protein with multiple functions. *Cytokine Growth Factor Rev.* 2006 Dec; 17(6):431–9. PMID: [17070092](#)
17. Bullock AN, Debreczeni JE, Edwards AM, 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 U S A.* 2006 May 16; 103(20):7637–42. PMID: [16675548](#)
18. Muniz JR, Guo K, Kershaw NJ, Ayinampudi V, von Delft F, Babon JJ, et al. Molecular architecture of the ankyrin SOCS box family of Cul5-dependent E3 ubiquitin ligases. *J Mol Biol.* 2013 Sep 9; 425(17):3166–77. doi: [10.1016/j.jmb.2013.06.015](#) PMID: [23806657](#)
19. Woo JS, Imm JH, Min CK, Kim KJ, Cha SS, Oh BH. Structural and functional insights into the B30.2/SPRY domain. *EMBO J.* 2006 Mar 22; 25(6):1353–63. PMID: [16498413](#)
20. Stanley BJ, Ehrlich ES, Short L, Yu Y, Xiao Z, Yu XF, et al. Structural insight into the human immunodeficiency virus Vif SOCS box and its role in human E3 ubiquitin ligase assembly. *J Virol.* 2008 Sep; 82(17):8656–63. doi: [10.1128/JVI.00767-08](#) PMID: [18562529](#)
21. Tsao DH, Maki AH. Optically detected magnetic resonance study of the interaction of an arsenic(III) derivative of cacodylic acid with EcoRI methyl transferase. *Biochemistry.* 1991 May 7; 30(18):4565–72. PMID: [2021649](#)
22. Schneider TR, Sheldrick GM. Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr.* 2002 Oct; 58(Pt 10 Pt 2):1772–9.
23. Wang J, Dauter M, Dauter Z. What can be done with a good crystal and an accurate beamline? *Acta Crystallogr D Biol Crystallogr.* 2006 Dec; 62(Pt 12):1475–83.
24. Brooks MA, Ravelli RB, McCarthy AA, Strub K, Cusack S. Structure of SRP14 from the Schizosaccharomyces pombe signal recognition particle. *Acta Crystallogr D Biol Crystallogr.* 2009 May; 65(Pt 5):421–33.
25. Retailleau P, Prange T. Phasing power at the K absorption edge of organic arsenic. *Acta Crystallogr D Biol Crystallogr.* 2003 May; 59(Pt 5):887–96.
26. Greenwald J, Le V, Butler SL, Bushman FD, Choe S. The mobility of an HIV-1 integrase active site loop is correlated with catalytic activity. *Biochemistry.* 1999 Jul 13; 38(28):8892–8. PMID: [10413462](#)
27. Liu X, Zhang H, Wang XJ, Li LF, Su XD. Get phases from arsenic anomalous scattering: de novo SAD phasing of two protein structures crystallized in cacodylate buffer. *PLoS One.* 2011; 6(9):e24227. doi: [10.1371/journal.pone.0024227](#) PMID: [21912678](#)
28. Galdeano C, Gadd MS, Soares P, Scaffidi S, Van Molle I, Birced I, et al. Structure-guided design and optimization of small molecules targeting the protein-protein interaction between the von Hippel-Lindau (VHL) E3 ubiquitin ligase and the hypoxia inducible factor (HIF) alpha subunit with in vitro nanomolar affinities. *J Med Chem.* 2014 Oct 23; 57(20):8657–63. doi: [10.1021/jm5011258](#) PMID: [25166285](#)
29. Maignan S, Guilloteau JP, Zhou-Liu Q, Clement-Mella C, Mikol V. Crystal structures of the catalytic domain of HIV-1 integrase free and complexed with its metal cofactor: high level of similarity of the active site with other viral integrases. *J Mol Biol.* 1998 Sep 18; 282(2):359–68. PMID: [9735293](#)
30. Kenyon EM, Hughes MF. A concise review of the toxicity and carcinogenicity of dimethylarsinic acid. *Toxicology.* 2001 Mar 7; 160(1–3):227–36. PMID: [11246143](#)
31. Jancarik J, Kim SH. Sparse-Matrix Sampling—a Screening Method for Crystallization of Proteins. *J Appl Crystallogr.* 1991 Aug 1; 24:409–11.
32. Newman J, Egan D, Walter TS, Meged R, Berry I, Ben Jelloul M, et al. Towards rationalization of crystallization screening for small- to medium-sized academic laboratories: the PACT/JCSG+ strategy. *Acta Crystallogr D Biol Crystallogr.* 2005 Oct; 61(Pt 10):1426–31.
33. Liu J, Nussinov R. The mechanism of ubiquitination in the cullin-RING E3 ligase machinery: conformational control of substrate orientation. *PLoS Comput Biol.* 2009 Oct; 5(10):e1000527. doi: [10.1371/journal.pcbi.1000527](#) PMID: [19798438](#)

34. Kim YK, Kwak MJ, Ku B, Suh HY, Joo K, Lee J, et al. Structural basis of intersubunit recognition in elongin BC-cullin 5-SOCS box ubiquitin-protein ligase complexes. *Acta Crystallogr D Biol Crystallogr*. 2013 Aug; 69(Pt 8):1587–97.
35. Liu J, Nussinov R. Molecular dynamics reveal the essential role of linker motions in the function of cullin-RING E3 ligases. *J Mol Biol*. 2010 Mar 12; 396(5):1508–23. doi: [10.1016/j.jmb.2010.01.022](https://doi.org/10.1016/j.jmb.2010.01.022) PMID: [20083119](https://pubmed.ncbi.nlm.nih.gov/20083119/)
36. Min JH, Yang H, Ivan M, Gertler F, Kaelin WG, Jr., Pavletich NP. Structure of an HIF-1alpha-pVHL complex: hydroxyproline recognition in signaling. *Science*. 2002 Jun 7; 296(5574):1886–9. PMID: [12004076](https://pubmed.ncbi.nlm.nih.gov/12004076/)
37. Hon WC, Wilson MI, Harlos K, Claridge TD, Schofield CJ, Pugh CW, et al. Structural basis for the recognition of hydroxyproline in HIF-1 alpha by pVHL. *Nature*. 2002 Jun 27; 417(6892):975–8. PMID: [12050673](https://pubmed.ncbi.nlm.nih.gov/12050673/)
38. Stebbins CE, Kaelin WG Jr., Pavletich NP. Structure of the VHL-ElonginC-ElonginB complex: implications for VHL tumor suppressor function. *Science*. 1999 Apr 16; 284(5413):455–61. PMID: [10205047](https://pubmed.ncbi.nlm.nih.gov/10205047/)
39. Nguyen HC, Yang H, Fribourgh JL, Wolfe LS, Xiong Y. Insights into Cullin-RING E3 Ubiquitin Ligase Recruitment: Structure of the VHL-EloBC-Cul2 Complex. *Structure*. 2015 Mar 3; 23(3):441–9. doi: [10.1016/j.str.2014.12.014](https://doi.org/10.1016/j.str.2014.12.014) PMID: [25661653](https://pubmed.ncbi.nlm.nih.gov/25661653/)
40. Schulman BA, Carrano AC, Jeffrey PD, Bowen Z, Kinnucan ER, Finnin MS, et al. Insights into SCF ubiquitin ligases from the structure of the Skp1-Skp2 complex. *Nature*. 2000 Nov 16; 408(6810):381–6. PMID: [11099048](https://pubmed.ncbi.nlm.nih.gov/11099048/)
41. Hao B, Zheng N, Schulman BA, Wu G, Miller JJ, Pagano M, et al. Structural basis of the Cks1-dependent recognition of p27(Kip1) by the SCF(Skp2) ubiquitin ligase. *Mol Cell*. 2005 Oct 7; 20(1):9–19. PMID: [16209941](https://pubmed.ncbi.nlm.nih.gov/16209941/)
42. Orlicky S, Tang X, Willems A, Tyers M, Sicheri F. Structural basis for phosphodependent substrate selection and orientation by the SCFCdc4 ubiquitin ligase. *Cell*. 2003 Jan 24; 112(2):243–56. PMID: [12553912](https://pubmed.ncbi.nlm.nih.gov/12553912/)
43. Liu J, Nussinov R. The role of allostery in the ubiquitin-proteasome system. *Crit Rev Biochem Mol Biol*. 2013 Mar-Apr; 48(2):89–97. doi: [10.3109/10409238.2012.742856](https://doi.org/10.3109/10409238.2012.742856) PMID: [23234564](https://pubmed.ncbi.nlm.nih.gov/23234564/)
44. Abad-Zapatero C, Oliete R, Rodriguez-Puente S, Pous J, Martinelli L, Johnson ME, et al. Humidity control can compensate for the damage induced in protein crystals by alien solutions. *Acta Crystallogr Sect F Struct Biol Cryst Commun*. 2011 Oct 1; 67(Pt 10):1300–8.
45. Li S, Peterson NA, Kim MY, Kim CY, Hung LW, Yu M, et al. Crystal Structure of AhpE from *Mycobacterium tuberculosis*, a 1-Cys peroxiredoxin. *J Mol Biol*. 2005 Mar 4; 346(4):1035–46. PMID: [15701515](https://pubmed.ncbi.nlm.nih.gov/15701515/)
46. Zhang H, Xu Q, Lu M, Xu X, Wang Y, Wang L, et al. Structural and functional studies of MutS2 from *Deinococcus radiodurans*. *DNA Repair (Amst)*. 2014 Sep; 21:111–9.
47. Asojo OA, Ceccarelli C. Structure of glutathione S-transferase 1 from the major human hookworm parasite *Necator americanus* (Na-GST-1) in complex with glutathione. *Acta Crystallogr F Struct Biol Commun*. 2014 Sep; 70(Pt 9):1162–6.
48. Dauter Z, Li M, Wlodawer A. Practical experience with the use of halides for phasing macromolecular structures: a powerful tool for structural genomics. *Acta Crystallogr D Biol Crystallogr*. 2001 Feb; 57(Pt 2):239–49.
49. Bulatov E, Martin EM, Chatterjee S, Knebel A, Shimamura S, Konijnenberg A, et al. Biophysical studies on interactions and assembly of full-size E3 ubiquitin ligase: suppressor of cytokine signaling 2 (SOCS2)-elongin BC-cullin 5-ring box protein 2 (RBX2). *J Biol Chem*. 2015 Feb 13; 290(7):4178–91. doi: [10.1074/jbc.M114.616664](https://doi.org/10.1074/jbc.M114.616664) PMID: [25505247](https://pubmed.ncbi.nlm.nih.gov/25505247/)
50. Powell HR, Johnson O, Leslie AG. Autoindexing diffraction images with iMosflm. *Acta Crystallogr D Biol Crystallogr*. 2013 Jul; 69(Pt 7):1195–203.
51. Evans PR, Murshudov GN. How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr*. 2013 Jul; 69(Pt 7):1204–14.
52. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr*. 2011 Apr; 67(Pt 4):235–42.
53. Terwilliger TC, Adams PD, Read RJ, McCoy AJ, Moriarty NW, Grosse-Kunstleve RW, et al. Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr D Biol Crystallogr*. 2009 Jun; 65(Pt 6):582–601.
54. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*. 2010 Feb; 66(Pt 2):213–21.
55. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr*. 2007 Aug 1; 40(Pt 4):658–74.

56. Terwilliger TC. Maximum-likelihood density modification. *Acta Crystallogr D Biol Crystallogr*. 2000 Aug; 56(Pt 8):965–72.
57. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*. 2010 Apr; 66(Pt 4):486–501.
58. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr*. 2011 Apr; 67(Pt 4):355–67.
59. Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*. 2010 Jan; 66(Pt 1):12–21.