1 **Structural basis of FANCD2 deubiquitination by USP1-UAF1**

2

3 Martin L. Rennie[1]*, Connor Arkinson[1,2], Viduth K. Chaugule[1,2], Rachel Toth[2], Helen Walden[1,2]*

4

5 [1]Institute of Molecular Cell and Systems Biology, College of Medical Veterinary and Life Sciences,

6 University of Glasgow, Glasgow, UK

7 [2]MRC Protein Phosphorylation and Ubiquitylation Unit, College of Life Sciences, University of Dundee,

8 Dundee, UK

9 *Corresponding authors: Walden, Helen (Helen.Walden@glasgow.ac.uk); Rennie, Martin

10 (Martin.Rennie@glasgow.ac.uk)

11

12 Keywords: ubiquitin, specificity, DNA repair, Fanconi Anemia, cryo-EM

13

14

15

1   **Abstract**

2   Ubiquitin-Specific Protease 1 (USP1), together with the cofactor UAF1, acts during DNA repair processes

3   to specifically to remove mono-ubiquitin signals. The mono-ubiquitinated FANCI-FANCD2 heterodimer is

4   one such substrate and is involved in the repair of DNA interstrand crosslinks via the Fanconi Anemia

5   pathway. Here we determine structures of human USP1-UAF1 with and without ubiquitin, and bound to

6   mono-ubiquitinated FANCI-FANCD2 substrate. Crystal structures of USP1-UAF1 reveal plasticity in USP1

7   and key differences to USP12-UAF1 and USP46-UAF1. A cryoEM reconstruction of USP1-UAF1 in

8   complex mono-ubiquitinated FANCI-FANCD2, highlights a highly orchestrated deubiquitination process

9   with USP1-UAF1 driving conformational changes in the substrate. An extensive interface between UAF1

10   and FANCI, confirmed by mutagenesis and biochemical assays, provides a molecular explanation for

11   their requirement despite neither being directly involved in catalysis. Overall, our data provide

12   molecular details of USP1-UAF1 regulation and substrate recognition.

13

1   **Introduction**

2   Ubiquitination is a reversible post-translation modification involving the attachment of

3   ubiquitin, which acts as a signal to regulate numerous cellular pathways, including protein

4   degradation, DNA replication, and DNA repair[1,2]. The attachment of ubiquitin proceeds via a

5   cascade of enzymes and typically involves formation of an isopeptide bond between a lysine on

6   the substrate the carboxy-terminus of ubiquitin. Removal of the ubiquitin signal is achieved by a

7   group of proteases known as deubiquitinases.

8   Ubiquitin signals are structurally diverse[1,2], composed of either poly-ubiquitin linkages, where

9   different ubiquitin lysines are sequentially modified, or mono-ubiquitination, in which a single

10   ubiquitin is attached to the target protein. Ubiquitin(s) can be directly recognised by reader

11   proteins and/or result in direct conformational changes of the modified target protein for

12   downstream effects. Signal reversal by deubiquitinases involves recognition of these different

13   substrate types in a spatio-temporal manner and the regulation of deubiquitinase activity[2,3].

14   The largest family of deubiquitinases are the Ubiquitin-Specific Proteases (USPs). DNA repair

15   mechanisms including the Fanconi Anemia pathway utilize USP1 for deubiquitination of mono-

16   ubiquitinated targets[4–8]. Loss of USP1 leads to dysfunctional DNA repair and genomic

17   instability[5,7,8], and USP1 has been identified as a potential drug target for overcoming cancer

18   resistance to DNA damaging drug treatment[9–11]. While we now understand how certain

19   deubiquitinases specifically target ubiquitin-ubiquitin linkages, how USP1 targets a specific set

20   of ubiquitin-substrate linkages remains an open question.

1   Three key substrates of USP1 are PCNA, FANCI, and FANCD2, each of which is mono-

2   ubiquitinated at a specific lysine residue during repair of DNA damage[5,6,12,13]. FANCI and

3   FANCD2 can form a heterodimer, while PCNA forms a homotrimer, with both complexes able to

4   form DNA clamps[14–16]. FANCI and FANCD2 are homologous proteins, each with an α-solenoid

5   fold contorted into a saxophone shape[17]. The non-ubiquitinated FANCI-FANCD2 complex is

6   stabilized by extensive interactions across the N-Terminal Domains (NTDs) of each protein that

7   partially bury the ubiquitination site. The C-Terminal Domains (CTDs) extend away from this

8   interface to make an open trough, with the Head Domain connecting the NTD and CTD.

9   Recently, cryogenic electron microscopy (cryo-EM) reconstructions of mono-ubiquitinated

10   FANCI-FANCD2 revealed a closed conformation of the heterodimer, resulting in it encircling

11   DNA[15,16,18]. The ubiquitin conjugated to each subunit makes non-covalent contacts with the

12   NTD of the other subunit, reconfiguring the NTD-NTD interface and creating a CTD-CTD

13   interface. Mono-ubiquitination of the FANCD2 subunit alone is sufficient to induce this closed

14   conformation and results in an enhanced affinity for double stranded DNA[15,16,18,19]. The

15   ubiquitin signal has therefore been proposed to protect DNA during repair or may recruit other

16   repair factors. How this specific ubiquitin signal is removed to unclamp FANCI-FANCD2 post-

17   repair is not well understood.

18   USP1 contains a predicted Ubiquitin-Specific Protease (USP) catalytic fold with two large

19   insertions and an N-Terminal Extension (NTE)[20]. USP1-Associated Factor (UAF1), the

20   catalytically essential cofactor of USP1, as well as USP12 and USP46,[21,22] is composed of a β-

21   propeller domain, an ancillary domain, and a SUMO-like domain (SLD). The β-propeller directly

22   interacts with both USP12[23,24] and USP46[25,26] in the same way, a feature likely conserved for

1   USP1[23]. The SLD is thought to mediate substrate recognition via binding to SUMO-Like domain

2   Interacting Motifs (SLIMs) on substrates[23,25,27].   However, unlike USP1-UAF1, neither USP12-

3   UAF1 nor USP46-UAF1 target mono-ubiquitinated FANCI-FANCD2[22]. Furthermore, FANCI buries

4   a substantial surface of FANCD2's ubiquitin[15,16], which shields it from several deubiquitinases

5   but not USP1-UAF1 *in vitro*[18]. Instead, FANCI phosphorylation is reported to impede USP1-UAF1

6   activity[28,29]; the underlying mechanism of this regulation remains unclear. Recently,

7   biochemical work has shown that the NTE of USP1 is critical for specific removal of FANCD2's

8   ubiquitin[20]. However a detailed understanding of how this specialized deubiquitinase targets

9   ubiquitin sequestered within the FANCI-FANCD2 complex is hampered by the lack of structural

10   information.

11   Here we present structures of human USP1-UAF1 alone, bound to ubiquitin, and finally bound

12   to   a   mono-ubiquitinated   FANCI-FANCD2   substrate,   representing   the   first   human

13   deubiquitinase-substrate complex, providing insight into the catalytic cycle of USP1-UAF1.

14   Crystal structures of truncated USP1-UAF1, with and without ubiquitin, reveal features shared

15   with USP12/46-UAF1 complexes, and those that aid USP1's isopeptidase activity. A cryo-EM

16   structure of full-length USP1-UAF1 bound to FANCI and mono-ubiquitinated FANCD2 reveals an

17   extensive UAF1-FANCI interface that surprisingly does not involve the previously predicted

18   FANCI SLIM. Conformational changes are observed in each entity of the deubiquitinase-

19   substrate complex, in particular FANCD2's ubiquitin disengages from FANCI and associates with

20   USP1. In addition, the structures reveal how engagement of USP1-UAF1 can occur while the

21   FANCI-FANCD2 heterodimer is clamped on DNA. Using mutational analyses and biochemical

22   assays we highlight the importance of the interfaces within the structures.   Overall the

1 structures and assays clarify a highly orchestrated and multifaceted deubiquitination cycle, in

2 which two binding partner proteins, UAF1 and FANCI, along with USP1 contacts with FANCD2,

3 regulate the reversal of this critical ubiquitin signal.

4

5 **Results**

6 *Architecture and plasticity of USP1*

7 In order to understand how USP1 is regulated by UAF1 and how USP1 targets ubiquitin, we

8 sought to determine structures of the multiple stages in USP1's catalytic cycle. We determined

9 structures of USP1-UAF1 with and without ubiquitin using X-ray crystallography. Full-length

10 constructs of USP1-UAF1 did not yield well-diffracting crystals. In order to obtain USP1-UAF1

11 crystals that diffract to better than 4 Å, we deleted several flexible regions of USP1, predicted

12 to be disordered, and the SLD of UAF1 (Fig. 1a, Methods). We determined a ubiquitin-bound

13 structure of the same construct by reacting USP1 with propargylated ubiquitin[30] (Fig. 1b). The

14 USP1 structure exhibits a typical "fingers-palm-thumb" architecture (Fig. 1a). Ubiquitin sits in

15 the palm, making contacts with all three subdomains, while only the fingers interact with

16 UAF1's β-propeller domain, as observed for USP12-UAF1[23,24] and USP46-UAF1[25]. Small angle X-

17 ray scattering (SAXS) measurements confirmed the oligomeric arrangement observed in the

18 crystal structures (Fig 1c, Supplementary Fig. S1a-c). The deleted Inserts 1 and 2 are positioned

19 on the opposite face of USP1 to the ubiquitin binding site, while the deleted N-terminus

20 extends from the distal tip of USP1, furthest from UAF1 (Fig. 1B).

6

1   Comparison of the ubiquitin bound and unbound structures of USP1-UAF1 reveals

2   conformational changes in USP1. In the ubiquitin-bound structure, "Binding Loops" BL1 and BL2

3   are re-arranged with respect to the ubiquitin-free structure, which allows threading of the

4   ubiquitin tail between the palm and thumb (Fig. 1d), while the globular body of ubiquitin

5   primarily contacts the fingers, similar to other USPs. In addition to these local changes, USP1

6   exhibits flexibility in the fingers, with a 30° rotation relating the palm and thumb of the two

7   crystal structures (Fig. 1a-b). Furthermore, each crystal had two USP1-UAF1 complexes per

8   asymmetric unit both with similar conformations within the same asymmetric unit

9   (Supplementary Fig. S1d). Such structural plasticity has been also observed for crystal structures

10   of USP12-UAF1 and USP46-UAF1 and may facilitate substrate engagement.

11   Although USP1, USP12, and USP46 are all regulated by UAF1, only USP12 and USP46 activities

12   are enhanced by another cofactor, WDR20[31,32]. Several structural features of the ubiquitin-free

13   USP1-UAF1 structure correspond more closely to WDR20-bound USP12 and USP46 structures

14   than WDR20-free USP12[23,26] (Fig. 1e-g). A conserved phenylalanine is buried between the

15   fingers and palm, for USP1-UAF1, USP12/46-UAF1-WDR20, but not USP12-UAF1 (Fig. 1e).

16   Correlating with this, USP1's BL1 is ordered in USP1-UAF1, USP12/46-UAF1-WDR20 but not

17   USP12-UAF1 (Fig. 1f), however crystal packing effects of the ubiquitin-free USP1-UAF1 structure

18   may contribute to this stabilization. Therefore, the USP fold of USP1 may be primed for efficient

19   catalysis without the additional requirement of WDR20. Overall, our crystal structures reveal

20   typical USP-ubiquitin and USP-UAF1 interactions and highlight both the differences between

21   USP1 and USP12/46 along with the positioning of the inserts and NTE of USP1.

22

Fig. 1. Structural characterization of USP1-UAF1. (**a**) Crystal structure of USP1, lacking the NTE and Insert 1 and Insert 2, in complex with UAF1 lacking the SLD. Subdomains of USP1 and domains of UAF1 are indicated. Deleted regions are indicated in the primary structure as dashed lines (see also Methods). (**b**) Crystal structure of USP1 lacking the N-terminal region and Insert 1 and Insert 2 covalently modified with propargylated ubiquitin in complex with UAF1 lacking the SLD. The termini and approximate

1   positions where the NTE, Insert 1, and Insert 2 would be are indicated. The rotation axis relating the

2   palm and thumbs of the two structures is shown. (**c**) Fit of the ubiquitin-bound crystal structure with

3   solution state SAXS measurements. (**d**) Comparison of USP1 catalytic region between the ubiquitin-free

4   (gray) and ubiquitin-bound (colored) crystal structures. Structures were aligned by the USP1 subunit. (**e**)

5   A phenylalanine, conserved in USP1, USP12, and USP46, occupies a pocket between the palm and

6   fingers in ubiquitin-free USP1-UAF1 structure, and in USP46-UAF1-WDR20 (6JLQ)[26] and USP12-UAF1-

7   WDR20 (5K1C) but not USP12-UAF1 (5K1A)[23]. (**f**) The BL1 is ordered in ubiquitin-free USP1-UAF1

8   structure, and in USP46-UAF1-WDR20 (6JLQ)[26] and USP12-UAF1-WDR20 (5K1C) but not USP12-UAF1

9   (5K1A)[23]. (**g**) Multiple sequence alignment of USP1, USP12, and USP46 in the region including the

10   conserved phenylalanine (indicated by *) and the BL1. Sequence alignment was performed using Clustal

11   Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/) and visualized using BOXSHADE.

12
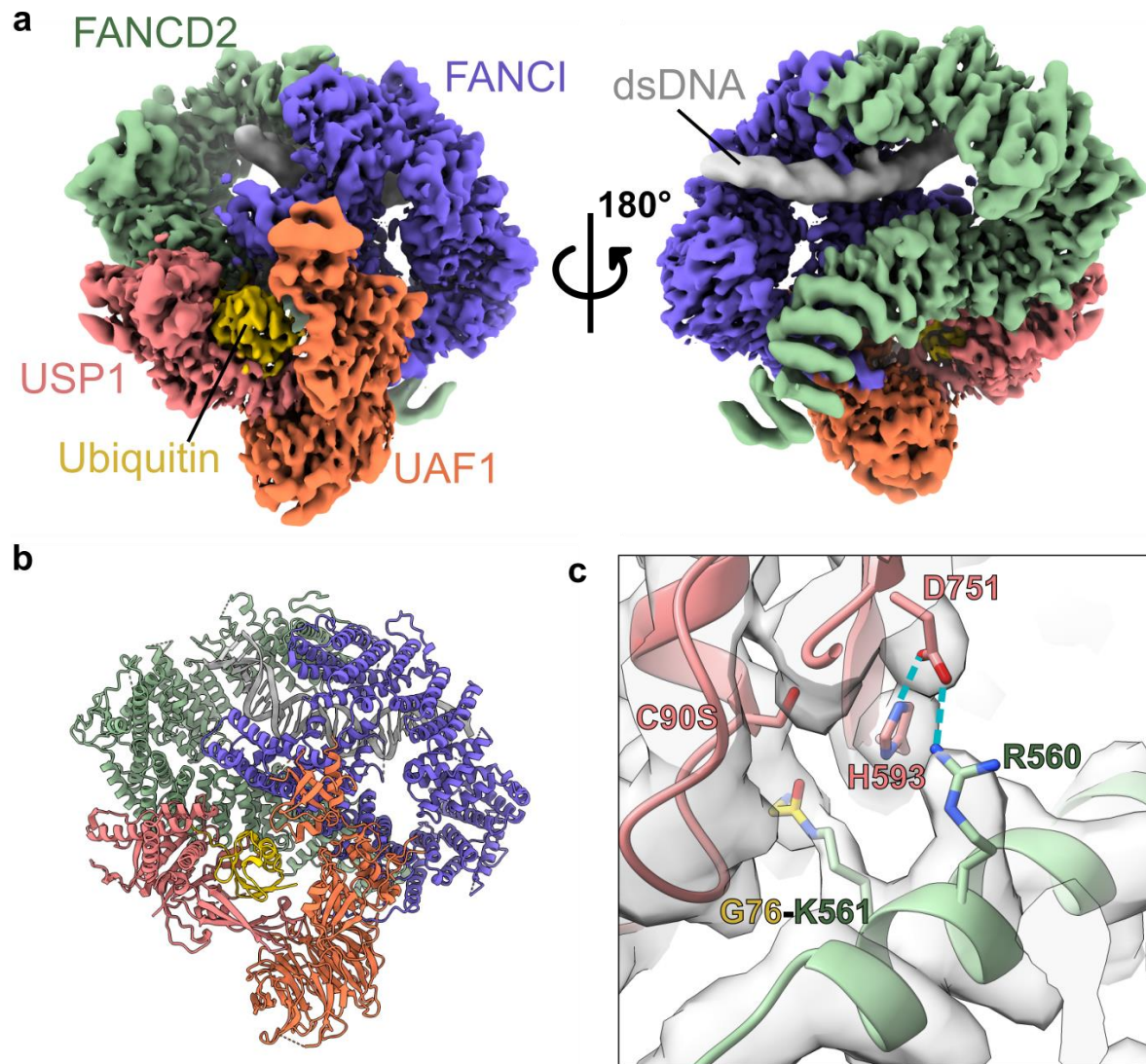
13   *Structure of USP1-UAF1 bound to FANCI-FANCD2 substrate*

14   Our USP1-UAF1 structures allow comparisons to USP12/46 complexes, demonstrate plasticity in

15   the USP fold and show of USP1 binds ubiquitin. However, in order to understand how USP1-

16   UAF1 recognises mono-ubiquitinated substrate, we determined a cryo-EM structure of full

17   length USP1-UAF1 bound to the FANCI-FANCD2^{Ub}-DNA substrate with ubiquitin only on

18   FANCD2. This singly mono-ubiquitinated substrate is readily deubiquitinated by USP1-UAF1

19   even in the presence of DNA, unlike the doubly mono-ubiquitinated complex, FANCI^{Ub}-

20   FANCD2^{Ub} [18,20,33]. We employed a catalytically compromised C90S mutant of USP1, and our

21   previous strategy for generating the mono-ubiquitinated substrate[34,35] to reconstitute the

22   USP1^{C90S}-UAF1-FANCI-FANCD2^{Ub} complex, with full length versions of all four proteins

23   (Supplementary Fig. S2a). We then determined a cryo-EM reconstruction of this enzyme-

24   substrate complex, bound to double-stranded DNA, at 3.8 Å global resolution (FSC=0.143) and

25   local resolution 3.6-10 Å (FSC=0.5) (Fig. 2a and Supplementary Fig. S2b-d). Using our crystal

1   structure of ubiquitin-bound USP1-UAF1 described above (Fig. 1), and the structures of mono-

2   ubiquitinated human FANCI-FANCD2[Ub] (PDB ID: 6VAF) and full-length UAF1 (PDB ID: 5K1A) we

3   built a model of approximately 70% of the residues in the complex including the K561-G76

4   isopeptide linkage (Fig. 2b-c).

5   Globally, USP1-UAF1 and the ubiquitin of FANCD2 are in a similar arrangement to the ubiquitin-

6   bound crystal structure described above (Fig. 1). FANCI-FANCD2 are in a closed conformation

7   encircling DNA, similar to that observed for the singly and doubly monoubiquitinated

8   complexes[15,16,18]. The enzyme-substrate complex is arranged such that USP1 contacts FANCD2

9   while UAF1 contacts FANCI in a 1:1:1:1 stoichiometry.

10

Fig. 2. Structural characterization of full-length USP1-UAF1 bound to the FANCI-FANCD2^{Ub} substrate. (**a**) Locally filtered cryo-EM map of the complex containing USP1 (pink), UAF1 (orange), FANCI (violet), and ubiquitinated (yellow) FANCD2 (green) at a threshold of 0.07. (**b**) Model of the enzyme-substrate complex. (**c**) The catalytic site of USP1 with the FANCD2^{K561}-ubiquitin^{G76} isopeptide linkage overlaid with the DeepEMhancer[36] map at a threshold of 0.1. Hydrogen bonds are shown as blue dashed lines.

UAF1 and FANCI form a scaffold for deubiquitination

11

1   UAF1 and FANCI have an extensive interface that involves interactions from the β-propeller,

2   ancillary, and SLD of UAF1 and the FANCI NTD and HD, and buries ~1600 $Å^2$ (Fig. 3a).

3   Interestingly, a FANCI SLIM[27], previously proposed to mediate interactions with UAF1, was not

4   located at the interface observed here. Instead, a FANCI loop ([251]DELLDVV[257]) directly interacts

5   with the SLD of UAF1 (Fig. 3a), with L254 inserting into the SLD hydrophobic groove. The UAF1-

6   FANCI interface observed here likely explains how FANCI favours USP1-UAF1 targeting over

7   more promiscuous deubiquitinases[18].

8   To further test this hypothesis, we compared the deubiquitination activity of USP1

9   reconstituted with either full-length UAF1, or UAF1 truncations lacking the SLD ($UAF1^{ΔSLD}$) or

10   comprising just the β-propeller domain ($UAF1^{β\text{-propeller}}$). Both truncations severely impaired

11   deubiquitination of $FANCI\text{-}FANCD2^{Ub}$ compared to full-length UAF1 (Fig. 3b). To confirm that

12   this was not due generic reduction in activity we also assessed deubiquitination of $FANCD2^{Ub}$ in

13   the absence of FANCI (Supplementary Fig. S3a-d). For the UAF1 truncations deubiquitination of

14   $FANCD2^{Ub}$ alone was only slightly reduced compared to full-length, as also reported for USP46-

15   $UAF1^{β\text{-propeller}}$ acting on fluorescently tagged ubiquitin[25]. Indeed, even at enzyme concentrations

16   resulting in comparable deubiquitination of $FANCD2^{Ub}$ alone, activity against $FANCI\text{-}FANCD2^{Ub}$

17   for the $UAF1^{β\text{-propeller}}$ was reduced compared to full-length UAF1, confirming the importance of

18   the UAF1-FANCI interaction (Supplementary Fig. 3d). In the absence of UAF1 (USP1 alone) we

19   did not detect deubiquitination of either substrate, consistent with the requirement of UAF1

20   for deubiquitinase activity of USP1 (Supplementary Fig. S3a-b). Overall, these data suggest that

21   the SLD and Ancillary domain are important for deubiquitination in the context of the FANCI-

22   FANCD2 heterodimer.

1

2
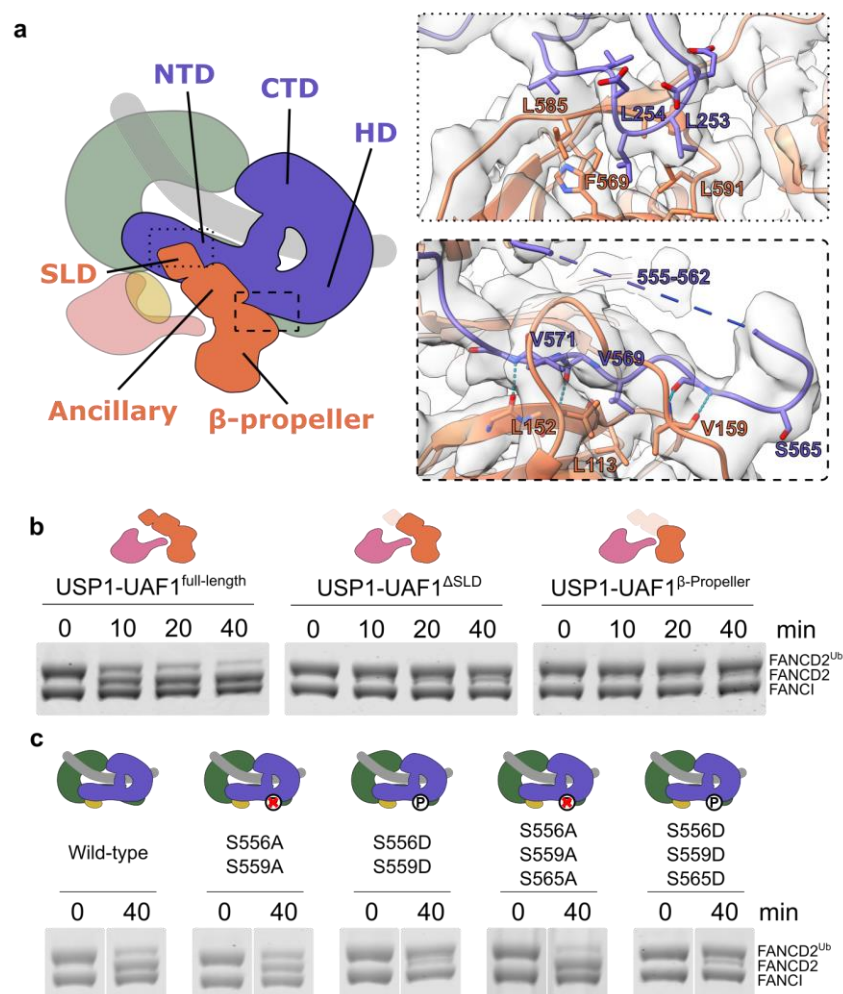
Fig. 3. FANCI-UAF1 interactions are important for deubiquitination of FANCI-FANCD2^{Ub}. (**a**) The interface between FANCI (violet) and UAF1 (orange). The FANCI$^{249-260}$ loop inserts into the hydrophobic pocket of the SLD of UAF1, while the FANCI$^{547-576}$ region interacts with the β-propeller domain of UAF1. Hydrophobic side chains involved in the interaction are shown and backbone hydrogen bonds are highlight as blue dashed lines. (**b**) Truncations of UAF1 reduce FANCI-FANCD2^{Ub} deubiquitination. Deubiquitination of FANCD2^{Ub} (1 μM) in complex with non-ubiquitinated FANCI (1 μM) by full-length USP1 reconstituted with UAF1 or truncations of UAF1 (200 nM enzyme complex) was assessed by SDS-PAGE and Coomassie staining (see also Supplementary Fig. S3). (**c**) Phosphorylation within FANCI$^{547-576}$ regulates deubiquitination. Deubiquitination of FANCD2^{Ub} (1 μM) in complex with non-ubiquitinated FANCI phosphorylation mimics (1 μM) by full-length USP1 reconstituted with full-length UAF1 (100 nM

13

1   enzyme complex) was assessed using SDS-PAGE and Coomassie staining. All assays were in the presence

2   of 4 μM 61 base pair dsDNA, and performed at least twice (two technical replicates).

3

4   In addition, the β-propeller domain of UAF1 interacts with a loop connecting the NTD and HD of

5   FANCI (FANCI[547-576]), the phosphorylation of which inhibits deubiquitination by USP1-UAF1[28,29]

6   (Fig. 3a). Although density for the FANCI phosphorylation targets (S556, S559, and S565) is weak

7   in our reconstruction, neighbouring residues contribute hydrophobic interactions and backbone

8   hydrogen bonds to the interaction. We confirmed a role of FANCI phosphorylation in regulating

9   FANCI-FANCD2[Ub] deubiquitination using different FANCI mutants (Fig. 3c, Supplementary Fig.

10   3e). The addition of FANCI phosphomimic S556D-S559D or S556D-S559D-S565D mutants

11   resulted in reduced FANCD2 deubiquitination compared to addition of wild-type FANCI. On the

12   other hand, phosphodead S556A-S559A or S556A-S559A-S565A mutants were similar to wild-

13   type. Finally, we note that mutation of the previously proposed SLIM motif of FANCI[27] (I683A-

14   L685A) resulted in similar deubiquitination to wild-type FANCI, confirming its redundancy for

15   FANCI-FANCD2[Ub] deubiquitination *in vitro* (Supplementary Fig. S3e). We propose that

16   phosphorylation alters the conformation of FANCI[547-576] such that it has reduced competency

17   for binding to UAF1. Taken together, the extensive UAF1-FANCI interface, along with FANCI

18   phosphorylation influence efficient USP1-UAF1 activity and reveal the multi-faceted regulation

19   underlying FANCD2 deubiquitination.

20

21   *Ubiquitin is poised for removal from FANCD2*

1 The ubiquitin conjugated to FANCD2 undergoes a large rearrangement in the enzyme-

2 substrate, USP1-UAF1-FANCI-FANCD2^{Ub}, complex relative to the substrate only, FANCI-

3 FANCD2^{Ub}, complex (Fig. 4a, Supplementary Movie S1). In the FANCI-FANCD2^{Ub} complex, the

4 ubiquitin tail is bent in a conformation incompatible with USP1 binding, with ubiquitin R72

5 forming intramolecular interactions with FANCD2 and the ubiquitin hydrophobic patch forming

6 intermolecular interactions with FANCI[15,16]. In the USP1-UAF1-FANCI-FANCD2^{Ub} complex,

7 ubiquitin is rotated by 160° and extracted from its non-covalent interaction with FANCI to sit in

8 the hand of USP1, with the ubiquitin tail extended to sit between the palm and thumb (Fig. 4a,

9 Supplementary Fig. S4a). The resulting USP1-ubiquitin interaction is analogous to the ubiquitin-

10 bound crystal structure above (Fig. 1b), however the palm and thumb are slightly more

11 contracted with respect to the fingers (Supplementary Fig. S4b).

12 While the USP1-ubiquitin interaction buries ~1800 Å^2, the observed USP1-FANCD2 interaction is

13 significantly smaller, burying ~600 Å^2. Within the USP1-FANCD2 interface, a conserved arginine

14 on FANCD2 (R560), adjacent to the ubiquitinated lysine, hydrogen bonds with D751 of the USP1

15 catalytic triad (Fig. 2c). The catalytic cysteine, mutated to serine in our structure, appears to

16 face away from the isopeptide bond, indicating local changes in the active site are required

17 before hydrolysis can occur.

18 Sitting at the crux between FANCI, FANCD2, and the BL1 of USP1 is weak density

19 (Supplementary Fig. S4c). Given its location it could belong to the NTE of USP1, the BL1 of USP1

20 or the N-terminus of UAF1, however we were not able to unambiguously assign this density.

21 Contributing to this uncertainty, the NTE density terminates adjacent to the USP domain and

22 FANCD2 (Supplementary Fig. S4d), leaving 74 residues unmodelled. Density for Inserts 1 and 2

15

1   of USP1 is absent in our structure (Supplementary Fig. S4e), suggesting that these are not

2   tightly associated to the USP domain and consistent with their disposability for FANCI and

3   FANCD2 deubiquitination[20].

4
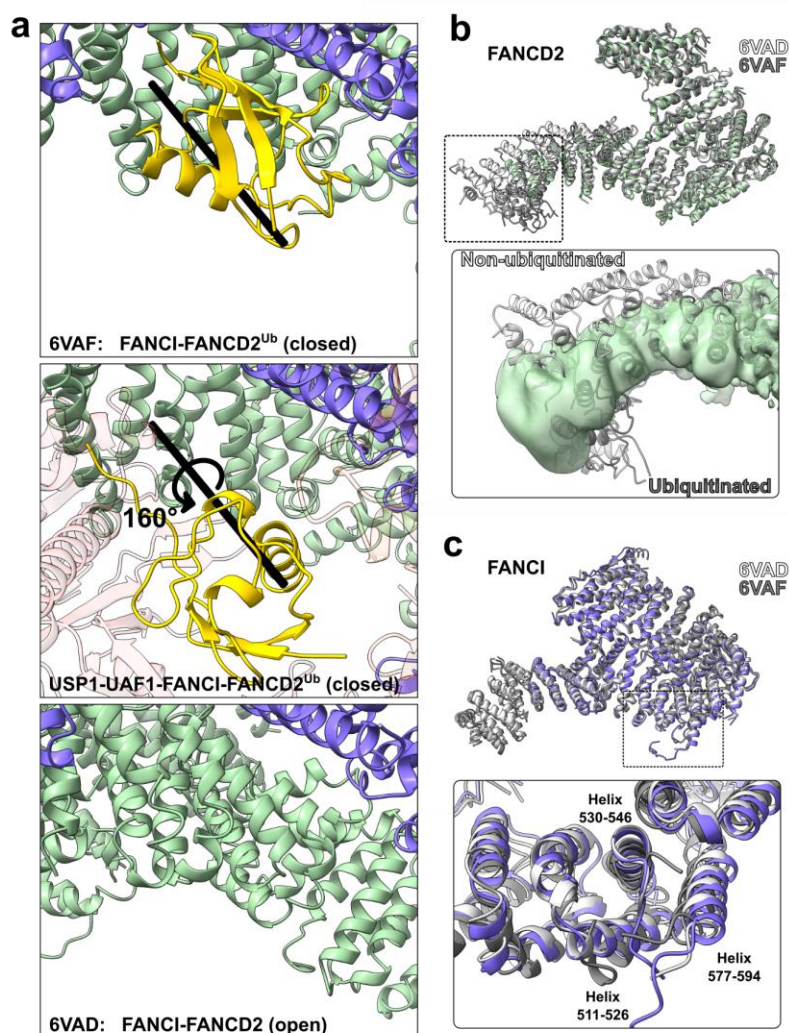
5

6   **Fig. 4.** Conformational changes in FANCI (green) and FANCD2 (violet) during the deubiquitination cycle.

7   (**a**) Conformational changes of ubiquitin (yellow) associated with USP1-UAF1 binding. Structures of the

8   FANCI-FANCD2^{Ub} substrate of deubiquitination (6VAE) and the FANCI-FANCD2 product (6VAD)[15] were

9   aligned to the USP1-UAF1-FANCI-FANCD2^{Ub} structure by the FANCI subunit. The axis of rotation for

10   ubiquitin is shown as a black cylinder. In the deubiquitinated state (bottom panel) FANCD2 and USP1

16

1 would sterically clash. (**b**) Aligned FANCD2 subunits from FANCI-FANCD2^{Ub} (6VAF; dark gray), USP1-

2 UAF1-FANCI-FANCD2^{Ub} (green), and FANCI-FANCD2 (6VAD; light gray). The N-terminus of FANCD2 in the

3 USP1-UAF1-FANCI-FANCD2^{Ub} structure (locally filter map at a threshold of 0.033) is in an intermediate

4 conformation between the FANCI-FANCD2^{Ub} and FANCI-FANCD2 states. (**c**) Aligned FANCI subunits from

5 FANCI-FANCD2^{Ub} (dark gray), USP1-UAF1-FANCI-FANCD2^{Ub} (violet), and FANCI-FANCD2 (light gray).

6 Although all three states are globally very similar, FANCI of USP1-UAF1-FANCI-FANCD2^{Ub} more closely

7 resembles the non-ubiquitinated FANCI-FANCD2 state in the HD close to the UAF1 interface.

8

9 *The closed conformation of FANCI-FANCD2 is recognised by USP1-UAF1*

10 The FANCI-FANCD2 heterodimer remains in the closed conformation that encircles DNA when

11 USP1-UAF1 is bound, despite disruption of the non-covalent interaction between FANCD2's

12 ubiquitin and FANCI (Fig. 2a). The open, trough shaped FANCI-FANCD2 conformation would be

13 incompatible with the mode of USP1-UAF1 binding identified here due to steric clashes with

14 USP1 (Fig. 4a, bottom panel). However, the FANCD2 subunit adopts an intermediate

15 conformation to the ubiquitinated and non-ubiquitinated extremes (Fig. 4b). The NTD of

16 FANCD2 was previously shown to flex with respect to the CTD, adopting an extended

17 conformation in unmodified FANCD2 and bent when it is ubiquitinated[15]. When bound to USP1-

18 UAF1, the NTD of FANCD2 becomes more extended than the ubiquitinated extreme, however

19 retains some flexibility as illustrated by 3D variability analysis[37] (Supplementary Movies S2-S3).

20 Although residues 1-180 of FANCI had weak density and were not modelled for our structure,

21 the rest of FANCI remains in a globally similar conformation regardless of ubiquitination state or

22 binding of USP1-UAF1 (Fig. 4c). However, there are slight changes in the region 437-790,

23 encompassing part of the NTD and most of the HD, including the FANCI ubiquitination site. The

17

1   USP1-UAF1 bound state more closely aligns with the non-ubiquitinated state, particularly for

2   helices 511-526, 530-546, and 577-594 (Fig. 4c), which are adjacent UAF1 β-propeller binding

3   site. Interestingly, the other face of this region contacts the FANCD2 NTD. Therefore UAF1

4   interaction with FANCI may allosterically regulate the interaction between FANCI and FANCD2.

5

6   **Discussion**

7   The structures described here provide insight into the catalytic cycle of USP1-UAF1

8   deubiquitination and mechanisms by which the cycle is controlled (Fig. 5). The ubiquitin-free

9   crystal structure (Fig. 1a) represents the free enzyme (E), while the cryo-EM structure (Fig. 2b)

10   represents a substrate-bound state (ES), in which the enzyme has globally distorted the

11   substrate. Local changes in the active site are further required to distort the isopeptide bond

12   into the transition state before hydrolysis occurs. Subsequent dissociation of deubiquitinated

13   FANCI-FANCD2 in the open conformation (P) likely occurs prior to dissociation of ubiquitin, as

14   occurs for other USPs[38]. We reason that this would occur due to the narrow size of the pocket

15   through which ubiquitin could leave the in the USP1-UAF1-FANCI-FANCD2^{Ub} structure, and the

16   conformational changes occurring that appear to be restoring the open conformations of the

17   FANCI and FANCD2 subunits. Finally, ubiquitin-bound enzyme (EP), represented here by the

18   ubiquitin-bound crystal structure (Fig. 1b), must release the ubiquitin to complete catalytic

19   cycle.

20

18
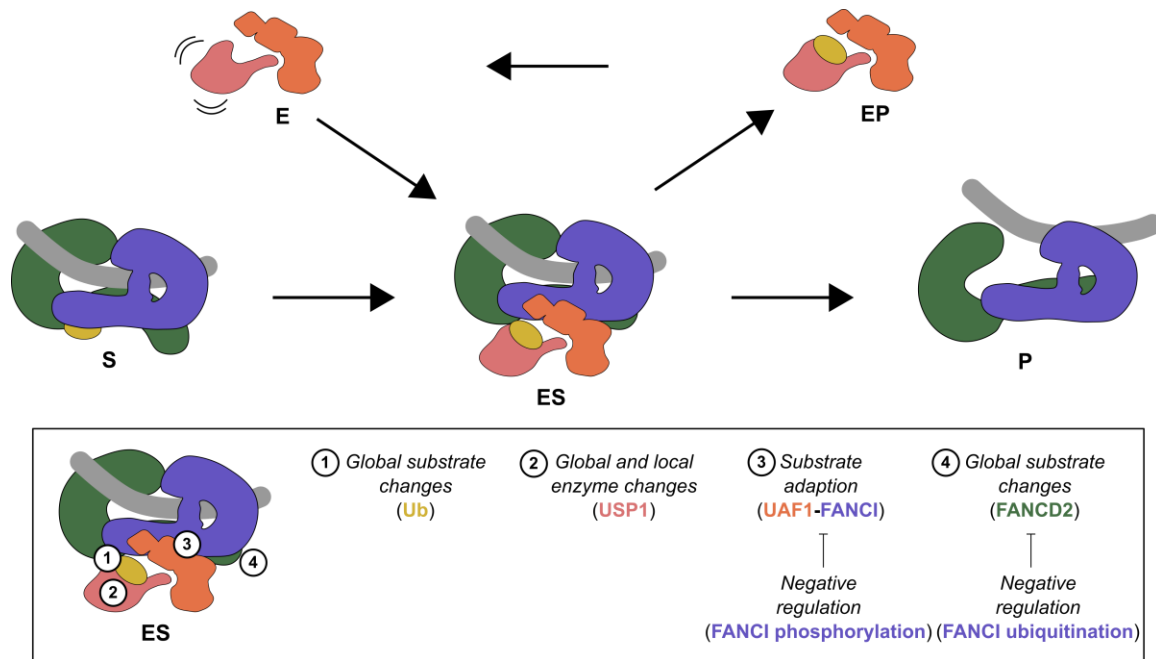
Fig. 5. Schematic representation of the deubiquitination cycle of FANCI-FANCD2 by USP1-UAF1. Key features of the reaction are highlighted.

Enzymes acting on large proteins typically utilize various regions of the substrate and the enzyme distal to the active site, to facilitate catalysis[39]. The USP1-UAF1 enzyme and FANCI-FANCD2^{Ub} substrate are prime examples of this complexity, with regulation occurring at multiple places (Fig. 5). Furthermore, in this system, two relatively rigid subunits, UAF1 of the enzyme complex and FANCI of the substrate complex, appear to act as a scaffold, while the catalytic subunit, USP1, and the ubiquitin and conjugated protein, FANCD2, are much more flexible. Such flexibility may be an essential feature of the machinery, allowing for gross conformational changes to occur, much in the same way local rearrangements of a catalytic site are required in enzymes acting on small molecules.

1 Although we were unable to unambiguously locate the NTE of USP1, which we have previously

2 shown is necessary for efficient deubiquitination of FANCD2[20], we did identify a candidate

3 region (Supplementary Fig. S4c). Curiously, the region identified sits between FANCI and

4 FANCD2 with two aromatic residues in close proximity (FANCI^{F277} and FANCD2^{Y520}). Both of

5 these residues are buried by the NTD-NTD interface in the open, non-ubiquitinated FANCI-

6 FANCD2 state. As such the occupant of the density may maintain the closed conformation while

7 FANCI and FANCD2 return to their respective open conformations. The proximity of the BL1 of

8 USP1, which contains a small insertion compared to USP12 and USP46 (Fig. 3g) suggests the

9 possibility of a *cis* interaction between the NTE of USP1 and its BL1. However, further work is

10 required to confirm these hypotheses. Unlike the NTE of USP1, which facilitates

11 deubiquitination of FANCD2 in general, the UAF1-FANCI interaction facilitates deubiquitination

12 in the context of the FANCI-FANCD2 heterodimer.

13 Our structure demonstrates that USP1 can recognize the FANCI-FANCD2 complex assembled on

14 DNA and in the closed conformation. However, the role of DNA in regulation of

15 deubiquitination is controversial with studies reporting DNA-mediated inhibition[18,20,33] and

16 enhancement[40,41] of deubiquitination. Such discrepancies may originate from binding to FANCI-

17 FANCD2 and USP1-UAF1, respectively. Despite addition of dsDNA after formation of the

18 enzyme-substrate complex we did not find evidence of DNA binding to USP1-UAF1 suggesting

19 dsDNA preferentially binds to FANCI-FANCD2. Upon dissociation of USP1-UAF1, return of the

20 FANCI-FANCD2 to the open conformation may unlock it from DNA.

21 Within the USP1-UAF1-FANCI-FANCD2^{Ub} structure there is space to accommodate ubiquitin

22 conjugated to FANCI. As such we expect a similar mechanism for recognition of the FANCI^{Ub}-

20

1   FANCD2[Ub], for FANCD2 deubiquitination, as FANCI-FANCD2[Ub]. FANCD2 is deubiquitinated

2   slower in the FANCI[Ub]-FANCD2[Ub] complex with I44 of FANCI's ubiquitin playing an important

3   role[18,20,33]. The interaction between FANCI's ubiquitin and FANCD2's NTD[15] may restrict flexing

4   of the NTD observed in the USP1-UAF1-FANCI-FANCD2[Ub] structure (Fig. 4b). Alternatively, the

5   conjugation of ubiquitin to FANCI may reduce the propensity for USP1-UAF1 to return of the

6   FANCI subunit to its open conformational state (Fig. 4c). Either or both of these would provide a

7   mechanism for the protection of FANCD2 from deubiquitination that is achieved by FANCI

8   ubiquitination. Although the order of deubiquitination of FANCI[Ub]-FANCD2[Ub] is not known, the

9   FANCD2 subunit is more rapidly deubiquitinated in biochemical assays and therefore expected

10 to be deubiquitinated first. How FANCI is deubiquitinated remains an open question and will

11 require further enzyme-substrate structures.

12 Although USP1 and FANCI compete for the hydrophobic patch of ubiquitin, the surface area of

13 ubiquitin buried by FANCI (1450 Å^2)[15] is less than that of USP1. As the FANCI binding surface is

14 remains accessible in our structure, the larger surface buried by USP1 may allow for

15 preferential binding of ubiquitin to USP1 over FANCI. In the doubly mono-ubiquitinated

16 complex, the surface area of the ubiquitin conjugated to FANCI, buried by FANCD2 is 1820 Å^2 [15].

17 Therefore, it may be more difficult for USP1-UAF1 to extract the ubiquitin conjugated to FANCI

18 from the complex, consistent with the observation that deubiquitination of FANCD2's ubiquitin

19 occurs more rapidly than FANCI's ubiquitin[18,20,33].

20 Finally, we have revealed the mechanism FANCI recognition by UAF1 (Fig. 3). Unlike the NTE of

21 USP1, which facilitates deubiquitination of FANCD2 in general[20], the UAF1-FANCI interface

22 facilitates deubiquitination in the context of the FANCI-FANCD2 heterodimer. This interface

21

1   appears to act as a switchable substrate adaptor, allowing USP1 to target ubiquitinated FANCD2

2   when FANCI is dephosphorylated. In addition it contributes to the ability of USP1 to overcome

3   the masking of the hydrophobic patch of ubiquitin by FANCI. The FANCI region containing the

4   S556, S559, and S565 phosphorylation sites appears to be a key player in this interaction.

5   Interestingly, crystal structures of mouse FANCI-FANCD2 (3S4W) and FANCI alone (3S51) show

6   this region two different conformations[17], while in cryo-EM structures of human FANCI-FANCD2

7   without UAF1 the region is not well ordered (6VAA, 6VAD, 6VAE, 6VAF)[15]. Therefore, the

8   dynamics of this region may be important for regulating recognition by UAF1. Overall, our

9   results reveal a sophisticated deubiquitination process in which multiple regions of the

10   substrate are recognized to coordinate both ubiquitin removal and substrate conformation.

11

1 **Methods**

2 *Protein expression and purification*

3 Proteins were prepared described previously[18,20,34]. Briefly, His-TEV-USP1$^{\Delta1-66,\Delta125-138,\Delta229-408,\Delta608-737}$-

4 UAF1$^{\Delta564-677}$ (USP1$^{\Delta N\Delta1\Delta2}$-UAF1$^{\Delta SLD}$) was co-expressed in *Sf*21 insect cells and purified by Ni-NTA affinity,

5 anion exchange, TEV protease treatment, subtractive Ni-NTA, and gel filtration[18]. His$_6$-TEV-

6 USP1^{G670A,G671A}, His$_6$-TEV-USP1^{G670A,G671A,C90S}, and His$_6$-TEV-UAF1$^{27-359}$ (β-propeller domain) were expressed

7 in *Sf*21 cells and purified by Ni-NTA affinity, anion exchange, TEV protease treatment, subtractive Ni-

8 NTA, and gel filtration[20]. His$_6$-3C-UAF1 and His$_6$-TEV-UAF1$^{2-580}$ (ΔSLD) were expressed in *Sf*21 cells and

9 purified by Ni-NTA affinity, anion exchange, and gel filtration (without removal of the tag)[20].

10 Propargylated ubiquitin was prepared from a ubiquitin-intein-CBD construct via reaction with MESNa

11 then proparglamine at 4°C, and reacted with USP1$^{\Delta N\Delta1\Delta2}$-UAF1$^{\Delta SLD}$ as described previously[20].

12 His$_6$-3C-FANCD2, His$_6$-FANCI, His$_6$-TEV-V5-FANCI, and His$_6$-TEV-V5-FANCI mutants were expressed in *Sf*21

13 cells and purified by Ni-NTA affinity, anion exchange, and gel filtration[34]. FANCD2 was ubiquitinated and

14 purified using an engineered Ube2T and SpyCatcher-SpyTag setup[34]. The His$_6$-3C tag was removed

15 during preparation of ubiquitinated FANCD2, while for FANCI tags were not removed.

16 Protein concentrations were determined using the predicted extinction coefficients at 280 nm[42] and

17 absorbance via a Nanodrop. The ratio of 260nm/280nm was ≤0.65 for all protein batches used in

18 subsequent experiments.

19 *Crystallization, data collection, and processing*

20 Crystals used to solve the structures were grown by hanging drop at 19°C. Purified USP1$^{\Delta N\Delta1\Delta2}$-UAF1$^{\Delta SLD}$

21 was concentrated to 11-15 mg/mL and drops of 4.5 μL were setup using a reservoir of 10% w/v

22 PEG4000, 100 mM MES/Imidazole pH 6.0-6.5, 30 mM CaCl$_2$, 30 mM MgCl$_2$, 20-25% glycerol v/v and a

23

1   ratio of 2 volumes protein to 1 volume reservoir. Purified USP1$^{\Delta N \Delta L1 \Delta L2}$-UAF1$^{\Delta SLD}$ was reacted with

2   propargylated ubiquitin (USP1$^{\Delta N \Delta 1 \Delta 2\text{-Ub-prg}}$-UAF1$^{\Delta SLD}$) was concentrated to 4-7 mg/mL and drops of 3-4.5

3   μL were setup using a reservoir of 8-13% w/v PEG3350, 0.1 M citric acid/Bis-Tris propane pH 4.1 and a

4   ratio of 2 volumes protein to 1 volume reservoir. For cryo-protection, crystals were transferred into

5   reservoir solution supplemented with 25% glycerol and incubated for 5 min up to 12 hours, followed by

6   vitrification in liquid nitrogen.

7   X-ray diffraction data were collected on a PILATUS 6M-F detector at Diamond Light Source, Beamline

8   I04. For USP1$^{\Delta N \Delta L1 \Delta L2\text{-Ub-prg}}$-UAF1$^{\Delta SLD}$, indexing and integration were performed using iMosflm[43], and

9   scaling and merging using AIMLESS[44]. For USP1$^{\Delta N \Delta L1 \Delta L2}$-UAF1$^{\Delta SLD}$ diffraction was anisotropic and was

10   processed using the STARANISO web server (http://staraniso.globalphasing.org/cgi-bin/staraniso.cgi).

11   For USP1$^{\Delta N \Delta 1 \Delta 2\text{-Ub-prg}}$-UAF1$^{\Delta SLD}$, molecular replacement was performed using PHASER[45], with 5CVN chain B

12   (with side chains removed) and 5CVL used as search models for USP1 and UAF1 respectively. One USP1

13   molecule and two UAF1 molecules were placed. Strong density positive difference density for the Zn

14   atom in the USP1 fingers subdomain was used to validate the molecular replacement solution. Iterative

15   manual modeling building and automated refinement were performed using COOT[46] and Refmac5[47],

16   respectively. During automated refinement UAF1 from 51KA (chain B) was used to provide reference

17   restraints. After several rounds of refinement, another USP1 molecule and two ubiquitin molecules were

18   placed using rigid body fitting. For USP1$^{\Delta N \Delta 1 \Delta 2}$-UAF1$^{\Delta SLD}$, molecular replacement was performed using

19   USP1$^{\Delta N \Delta 1 \Delta 2\text{-Ub-prg}}$-UAF1$^{\Delta SLD}$ chains A and B, and refined as above. Final refinement steps were performed

20   in phenix[48]. Crystallography data and model statistics are reported in Supplementary Table S1.

21   ***SAXS data collection and processing***

22   Size-exclusion chromatography (SEC) coupled to SAXS measurements were performed at the Diamond

23   Light Source, Beamline B21 via the mail-in service. USP1$^{\Delta N \Delta 1 \Delta 2\text{-Ub-prg}}$-UAF1$^{\Delta SLD}$ at ~10 mg/mL was thawed

1   and injected at 0.075 mL/min onto a Superose 6 Increase 3.2/300 column equilibrated with 20 mM Tris

2   pH 8.0, 200 mM NaCl, 5% glycerol, 5 mM DTT. SAXS data were collected on an Eiger 4M detector. Buffer

3   subtraction and averaging were performed using chromixs[49]. Buffer subtracted, averaged scattering was

4   analyzed using ATSAS[49] and RAW[50]. Crysol was used to fit PDB models to the SAXS data.

5   ***Gel filtration analysis***

6   To prepare the USP1^{C90S}-UAF1-FANCI-FANCD2^{Ub} complex the four individually purified subunits were

7   thawed on ice and mixed 1.2:1.2:1:1. The assembled complex was diluted to ~5 mg/mL and a total

8   volume of 250 μL with 20 mM Tris pH 8.0, 25 mM NaCl, 5% glycerol, 3 mM MgCl2, 1 mM DTT to yield a

9   final NaCl concentration of approximately 200 mM. The sample was then injected at 0.5 mL/min onto a

10   Superose 6 Increase GL 10/300 column equilibrated in EM buffer (20 mM Tris pH 8.0, 150 mM NaCl, 2

11   mM DTT).

12   ***Cryo-EM sample preparation and data collection***

13   The USP1^{C90S}-UAF1-FANCI-FANCD2^{Ub} complex was prepared by mixing the four individually purified

14   subunits 1:1:1:1. The complex was exchanged into EM buffer using a Bio-Spin P-30 column (Bio-Rad).

15   The concentration of complex was estimated from absorbance at 280 nm and 1 equivalent of dsDNA (61

16   base pairs; TGATCAGAGGTCATTTGAATTCATGGCTTCGAGCTTCATGTAGAGTCGACGGTGCTGGGAT; IDT) per

17   protein complex was added. Immediately prior to preparing grids the sample was pre-equilibrated to

18   room temperature for 5 min. Quantifoil 1.2/1.3 300 mesh grids were glow discharged twice and 3.6 μL

19   of at 8.8 μM complex was applied. The grids were blotted for 2.5 s and vitrified in liquid ethane using a

20   Vitrobot operating at ~95% humidity at 15°C.

21   ***Cryo-EM data collection and processing***

25

1   A CRYO ARM 300 (JEOL) equipped with a DE64 detector was used to collect a total of 4912 movies using

2   beam-Image shift at the Scottish Centre for Macromolecular Imaging (SCMI). All movies were collected

3   in counting mode with a calibrated pixel size of 1.015 Å using SerialEM[51]. Movies were as 59 frames for a

4   total dose of ~65 e$^-$/Å^2 over 14.9 s. A subset of the movies were additionally gain corrected using

5   relion_estimate_gain[52]. Subsequent processing was performed in Cryosparc v2.13.2[53] (fig. SB-D). Patch

6   motion correction, patch CTF estimation, and manual curation was performed resulting in 4593 dose-

7   weighted, motion corrected images. Blob picking was performed with minimum and maximum particle

8   diameters of 100 Å and 220 Å respectively, using an elliptical blob. 2D classification was used to

9   generate classes for template picking and select particles for *ab initio* reconstruction with 2 classes, from

10   which a single class with clear protein features was obtained. Template picking was performed on the

11   entire dataset and 1,485,510 boxes were extracted. To remove non-particles, multiple rounds of

12   heterogenous refinement were performed with one good starting model and low pass filtered to 20 Å

13   and 2 or 3 "junk" starting models not representing the protein complex of interest. This cleaning

14   resulted in 391,552 particles for which homogeneous refinement, followed by non-uniform, local

15   refinement using the dynamic mask generated from homogenous refinement, were performed.

16   The local resolution was performing using FSC threshold of 0.5 and Adaptive Window Factor of 8, which

17   was then used for local filtering. DeepEMhancer (version 0.13)[36] via COSMIC2[54] was used to post-process

18   the half-maps using the highRes learning model to aid interpretation.

19   ***Model building***

20   Initially rigid body fitting was performed with USP1 and ubiquitin from the ubiquitin-free structure

21   (chains A and C), UAF1 from 5K1A (chain B), and FANCI and FANCD2 from 6VAF (chains A and B) using

22   UCSF Chimera[55]. The N-terminal region of FANCD2 (residues 45 to 311) was fitted as a rigid body into the

23   locally filtered map[46]. Manual model editing was performed using COOT and incorporating torsion,

1   planar peptide, trans peptide, and Ramachandran restraints. Where appropriate secondary structure

2   restraints were also included. Automated refinement against a globally sharpened map, with the B-

3   factor estimated from the Guinier plot, was performed using phenix real-space refinement[56]. A non-

4   bonded weight of 250 was used to reduce steric clashes. Bond and angle restraints for the USP1 Zinc

5   finger and the K561-G76 isopeptide bond were incorporated. UAF1 from 51KA (chain B) was used to

6   provide reference restraints. The dsDNA was modelled using rigid body fitting of chains S and T from

7   6VAE. Cryo-EM data and model statistics are reported in Supplementary Table S2.

8   Superposition of atomic models was performed using "MatchMaker" in UCSF Chimera with the default

9   settings unless otherwise specified. Rotations were estimated using the "measure rotation" command in

10   UCSF Chimera. Buried surface area calculations were performed using PISA[57]. Figures were produced

11   with UCSF Chimera[55] or ChimeraX[58].

12   ***Deubiquitination assays***

13   Deubiquitination reactions were performed by preparing a 2x substrate mix and a 2x enzyme mix, and

14   mixing these 1:1 to initiate the reaction. Both mixes were setup on ice, and then incubated at room

15   temperature 30 min prior to reaction initiation and during the reaction. 5 μL aliquots of the reaction

16   were terminated at the indicated timepoints by addition of 5 μL 2x NuPAGE LDS buffer (Thermo Fisher)

17   supplemented with 200 mM DTT.

18   The 2x substrate mixes were prepared by diluting stocks (≥30 μM) of FANCD2$_{Ub}$, His$_6$-V5-TEV-FANCI (or

19   matched buffer), and dsDNA (61 base pairs) with DUB buffer (20 mM Tris pH 8.0, 75 mM NaCl, 5%

20   glycerol, 1 mM DTT). The resulting 2x mixes were composed 2 μM FANCD2$_{Ub}$, 2 μM FANCI, 8 μM dsDNA.

21   The 2x enzyme mixes were prepared by diluting concentrated stocks (≥30 μM) of USP1, and UAF1 (full-

22   length or β-propeller domain) with DUB buffer. The resulting 2x mixes were composed of 200 nM USP1

23   and 200 nM UAF1 (full-length) or 400 nM USP1 and 400 nM UAF1 (β-propeller domain). SDS-PAGE,

27

1   Coomassie staining, and western blotting was performed as described previously [18]. For western

2   blotting, 1:1000 dilution Anti-FANCD2 (ab108928) and 1:8000 dilution Anti-V5 (ab27671; 1 in 8000

3   dilution) were used for detection of FANCD2 and FANCI, respectively. After labelling secondary anti-

4   bodies at 1:10,000 dilution, bands were visualized on an Odyssey CLX Infrared Imaging System (Li-COR).

5   Each band was quantified using Image Studio with horizontal background subtraction (Li-COR). Replicate

6   assays were performed from different thawed protein aliquots, each from the same protein

7   preparation.

8

9   **Data availability**

10   The atomic coordinates and structure factors of ubiquitin-free and ubiquitin-bound USP1-UAF1

11   are being deposited to the PDB. The atomic coordinates and cryo-EM maps, including locally

12   filtered and sharpened and DeepEMhancer maps, are being deposited to the PDB and EMDB.

Acknowledgements

Author contributions

M.L.R., C.A., V.K.C. and H.W. conceived this work; C.A., M.L.R. and V.K.C. purified proteins; C.A. and R.T. generated various expression vectors and performed mutagensis; C.A. performed crystallography; M.L.R. performed cryo-EM data processing; M.L.R. and C.A. performed model building and refinement; M.L.R. performed assays; M.L.R. and C.A. wrote the manuscript with contributions from all other authors; H.W. secured funding and supervised the project.

Competing interests

The authors declare no competing interests.

29

References

1. Yau, R. & Rape, M. The increasing complexity of the ubiquitin code. *Nat. Cell Biol.* **18**, 579–586 (2016).

2. Mevissen, T. E. T. & Komander, D. Mechanisms of deubiquitinase specificity and regulation. *Annu. Rev. Biochem.* **86**, 159–192 (2017).

3. Grabbe, C., Husnjak, K. & Dikic, I. The spatial and temporal organization of ubiquitin networks. *Nat. Rev. Mol. Cell Biol.* **12**, 295–307 (2011).

4. Nijman, S. M. B. *et al.* The deubiquitinating enzyme USP1 regulates the fanconi anemia pathway. *Mol. Cell* **17**, 331–339 (2005).

5. Huang, T. T. *et al.* Regulation of monoubiquitinated PCNA by DUB autocleavage. *Nat. Cell Biol.* **8**, 339–347 (2006).

6. Sims, A. E. *et al.* FANCI is a second monoubiquitinated member of the Fanconi anemia pathway. *Nat. Struct. Mol. Biol.* **14**, 564–567 (2007).

7. Oestergaard, V. H. *et al.* Deubiquitination of FANCD2 Is Required for DNA Crosslink Repair. *Mol. Cell* **28**, 798–809 (2007).

8. Murai, J. *et al.* The USP1/UAF1 Complex Promotes Double-Strand Break Repair through Homologous Recombination. *Mol. Cell. Biol.* **31**, 2462–2469 (2011).

9. Chen, J. *et al.* Selective and cell-active inhibitors of the USP1/ UAF1 deubiquitinase complex reverse cisplatin resistance in non-small cell lung cancer cells. *Chem. Biol.* **18**, 1390–1400 (2011).

10. García-Santisteban, I., Peters, G. J., Giovannetti, E. & Rodríguez, J. A. USP1 deubiquitinase: Cellular functions, regulatory mechanisms and emerging potential as target in cancer therapy. *Mol. Cancer* **12**, 1–12 (2013).

11. Liang, Q. *et al.* A selective USP1-UAF1 inhibitor links deubiquitination to DNA damage responses. *Nat. Chem. Biol.* **10**, 298–304 (2014).

12. Garcia-Higuera, I. *et al.* Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. *Mol. Cell* **7**, 249–262 (2001).

13. Smogorzewska, A. *et al.* Identification of the FANCI Protein, a Monoubiquitinated FANCD2

30

Paralog Required for DNA Repair. *Cell* **129**, 289–301 (2007).

14. Krishna, T. S. R., Kong, X. P., Gary, S., Burgers, P. M. & Kuriyan, J. Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell* **79**, 1233–1243 (1994).

15. Wang, R., Wang, S., Dhar, A., Peralta, C. & Pavletich, N. P. DNA clamp function of the monoubiquitinated Fanconi anaemia ID complex. *Nature* (2020). doi:10.1038/s41586-020-2110-6

16. Alcón, P. *et al.* FANCD2–FANCI is a clamp stabilized on DNA by monoubiquitination of FANCD2 during DNA repair. *Nat. Struct. Mol. Biol.* **27**, 240–248 (2020).

17. Joo, W. *et al.* Structure of the FANCI-FANCD2 complex: Insights into the Fanconi anemia DNA repair pathway. *Science* **333**, 312–316 (2011).

18. Rennie, M. L. *et al.* Differential functions of FANCI and FANCD2 ubiquitination stabilize ID2 complex on DNA. *EMBO Rep.* e50133 (2020). doi:10.15252/embr.202050133

19. Tan, W. *et al.* Monoubiquitination by the human Fanconi Anemia core complex clamps FANCI:FANCD2 on DNA in filamentous arrays. *Elife* **9**, e54128 (2020).

20. Arkinson, C., Chaugule, V. K., Toth, R. & Walden, H. Specificity for deubiquitination of monoubiquitinated FANCD2 is driven by the N-terminus of USP1. *Life Sci. Alliance* **1**, e201800162 (2018).

21. Cohn, M. A. *et al.* A UAF1-Containing Multisubunit Protein Complex Regulates the Fanconi Anemia Pathway. *Mol. Cell* **28**, 786–797 (2007).

22. Cohn, M. A., Kee, Y., Haas, W., Gygi, S. P. & D'Andrea, A. D. UAF1 is a subunit of multiple deubiquitinating enzyme complexes. *J. Biol. Chem.* **284**, 5343–5351 (2009).

23. Li, H. *et al.* Allosteric Activation of Ubiquitin-Specific Proteases by β-Propeller Proteins UAF1 and WDR20. *Mol. Cell* **63**, 249–260 (2016).

24. Dharadhar, S., Clerici, M., van Dijk, W. J., Fish, A. & Sixma, T. K. A conserved two-step binding for the UAF1 regulator to the USP12 deubiquitinating enzyme. *J. Struct. Biol.* **196**, 437–447 (2016).

25. Yin, J. *et al.* Structural Insights into WD-Repeat 48 Activation of Ubiquitin-Specific Protease 46. *Structure* **23**, 2043–2054 (2015).

26. Zhu, H., Zhang, T., Wang, F., Yang, J. & Ding, J. Structural insights into the activation of USP46 by

WDR48 and WDR20. *Cell Discov.* **5**, 1–4 (2019).

27. Yang, K. *et al.* Regulation of the Fanconi anemia pathway by a SUMO-like delivery network. *Genes Dev.* **25**, 1847–1858 (2011).

28. Cheung, R. S. *et al.* Ubiquitination-Linked Phosphorylation of the FANCI S/TQ Cluster Contributes to Activation of the Fanconi Anemia I/D2 Complex. *Cell Rep.* **19**, 2432–2440 (2017).

29. Tan, W., van Twest, S., Murphy, V. J. & Deans, A. J. ATR-Mediated FANCI Phosphorylation Regulates Both Ubiquitination and Deubiquitination of FANCD2. *Front. Cell Dev. Biol.* **8**, 1–11 (2020).

30. Ekkebus, R. *et al.* On terminal alkynes that can react with active-site cysteine nucleophiles in proteases. *J. Am. Chem. Soc.* **135**, 2867–2870 (2013).

31. Kee, Y. *et al.* WDR20 regulates activity of the USP12·UAF1 deubiquitinating enzyme complex. *J. Biol. Chem.* **285**, 11252–11257 (2010).

32. Dahlberg, C. L. & Juo, P. The WD40-repeat Proteins WDR-20 and WDR-48 Bind and activate the Deubiquitinating Enzyme USP-46 to promote the abundance of the Glutamate Receptor GLR-1 in the ventral Nerve Cord of Caenorhabditis elegans. *J. Biol. Chem.* **289**, 3444–3456 (2014).

33. van Twest, S. *et al.* Mechanism of Ubiquitination and Deubiquitination in the Fanconi Anemia Pathway. *Mol. Cell* **65**, 247–259 (2017).

34. Chaugule, V. K., Arkinson, C., Toth, R. & Walden, H. Enzymatic preparation of monoubiquitinated FANCD2 and FANCI proteins. *Methods Enzymol.* **618**, 73–104 (2019).

35. Chaugule, V. K. *et al.* Allosteric mechanism for site-specific ubiquitination of FANCD2. *Nat. Chem. Biol.* **16**, 291–301 (2020).

36. Sanchez Garcia, R. *et al.* DeepEMhacer: a deep learning solution for cryo-EM volume post-processing. *bioRxiv* (2020). doi:10.1101/2020.06.12.148296

37. Punjani, A. & Fleet, D. J. 3D variability analysis : Directly resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM images. *bioRxiv* (2020). doi:10.1101/2020.04.08.032466.

38. Clerici, M., Luna-Vargas, M. P. A., Faesen, A. C. & Sixma, T. K. The DUSP-Ubl domain of USP4

enhances its catalytic efficiency by promoting ubiquitin exchange. *Nat. Commun.* **5**, (2014).

39. Rennie, M. L., Chaugule, V. K. & Walden, H. Modes of allosteric regulation of the ubiquitination machinery. *Curr. Opin. Struct. Biol.* **62**, 189–196 (2020).

40. Lim, K. S. *et al.* USP1 Is Required for Replication Fork Protection in BRCA1-Deficient Tumors. *Mol. Cell* **72**, 925-941.e4 (2018).

41. Liang, F. *et al.* DNA requirement in FANCD2 deubiquitination by USP1-UAF1-RAD51AP1 in the Fanconi anemia DNA damage response. *Nat. Commun.* **10**, 1–8 (2019).

42. Pace, N. C., Vajdos, F., Fee, L., Grimsley, G. & Gray, T. How to Measure and Predict the Molar Absorption Coefficient of a Protein. *Protein Sci.* **4**, 2411–2423 (1995).

43. Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. *Acta Cryst. D* **67**, 271–281 (2011).

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

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

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

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

48. Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr. Sect. D Struct. Biol.* **75**, 861–877 (2019).

49. Franke, D. *et al.* ATSAS 2.8: A comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* **50**, 1212–1225 (2017).

50. Hopkins, J. B., Gillilan, R. E. & Skou, S. BioXTAS RAW: Improvements to a free open-source program for small-angle X-ray scattering data reduction and analysis. *J. Appl. Crystallogr.* **50**, 1545–1553 (2017).

51. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).

52. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, e42166 (2018).

53. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).

54. Cianfrocco, M. A., Wong-Barnum, M., Youn, C., Wagner, R. & Leschziner, A. COSMIC2: A science gateway for cryo-electron microscopy structure determination. in *Practice & Experience in Advanced Research Computing* New Orleans, LA Jul 9-13 (2017). doi:10.1145/3093338.3093390

55. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

56. Afonine, P. V. *et al.* Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 531–544 (2018).

57. Krissinel, E. & Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **372**, 774–797 (2007).

58. Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).