DNA damage detection in nucleosomes involves DNA register shifting

Syota Matsumoto1,2,7, Simone Cavadini1,2,7, Richard D. Bunker1,2,7, Ralph S. Grand1,2, Alessandro Potenza1,2, Julius Rabl1,2, Junpei Yamamoto3, Andreas D. Schenk1,2, Dirk Schübeler1,2, Shigenori Iwai3, Kaoru Sugasawa4, Hitoshi Kurumizaka5,6 & Nicolas H. Thomä1,2

Access to DNA packaged in nucleosomes is critical for gene regulation, DNA replication and DNA repair. In humans, the UV–damaged DNA–binding protein (UV–DDB) complex detects UV–light–induced pyrimidine dimers throughout the genome; however, it remains unknown how these lesions are recognized in chromatin, in which nucleosomes restrict access to DNA. Here we report cryo–electron microscopy structures of UV–DDB bound to nucleosomes bearing a 6–4 pyrimidine–pyrimidone dimer or a DNA–damage mimic in various positions. We find that UV–DDB binds UV–damaged nucleosomes at lesions located in the solvent–facing minor groove without affecting the overall nucleosome architecture. In the case of buried lesions that face the histone core, UV–DDB changes the predominant translational register of the nucleosome and selectively binds the lesion in an accessible, exposed position. Our findings explain how UV–DDB detects occluded lesions in strongly positioned nucleosomes, and identify slide–assisted site exposure as a mechanism by which high–affinity DNA–binding proteins can access otherwise occluded sites in nucleosomal DNA.

The occlusion of DNA by chromatin impedes DNA repair1,2. Exposure of DNA to UV light induces the formation of pyrimidine dimers—such as 6–4 pyrimidine–pyrimidone photoproducts (6–4PPs) and cyclobutane–pyrimidine dimers—in all regions of the chromatinized mammalian genome3. Nucleosome core particles comprise particles between 145 and 147 base pairs (bp) of B-form DNA wrapped in 1.6–1.7 left-handed turns around a histone octamer protein core, which is formed by two copies each of H2A, H2B, H3 and H4.1,2 This configuration leaves only a small fraction of the nucleosomal DNA accessible to binding factors.

In the global genome repair branch of the nucleotide excision repair pathway, UV lesions are primarily detected by the protein complexes XPC–RAD23–CETN2 and UV–DDB8. UV–DDB, which is composed of the proteins DDB1 and DDB2, operates in conjunction with the DDB1–associated cullin–RING E3 ubiquitin ligase CUL4A (or CUL4B). After recognition of UV damage, the ubiquitin ligase is activated8 and nearby proteins—including histones—are ubiquitinated9. All available data suggest that UV–DDB binding precedes the binding of the chromatin remodelling machinery and of XPC10–12, and that DDB2 engages DNA lesions irrespective of their location in nucleosomes or linker regions13,14. The DDB2 gene is specifically mutated in patients that belong to the genetic complementation group E of xeroderma pigmentosum15, a rare congenital DNA–repair disorder that is characterized by extreme UV sensitivity and an approximately 10,000–fold increased incidence of skin cancer16.

The structural basis of how UV–DDB binds to UV–damaged double–stranded DNA is well understood14,17,18; however, it is currently unclear how UV lesions in the nucleosome are recognized given the access restrictions imposed by the nucleosomal architecture. Two possibilities have been suggested: DNA recognition after DNA looping off from the histone octamer in a process termed site exposure19, and access mediated by ATP–driven nucleosome remodelling.3 Accessing nucleosomal DNA is not only fundamental for DNA repair, but also as pioneer transcription factors20. We aimed to elucidate how UV lesions at various positions in nucleosomal DNA are recognized by UV–DDB and define the principles of protein access to chromatinized DNA templates.

Read–out of an exposed UV lesion

To determine structures of UV–DDB in complex with nucleosomes containing pyrimidine dimers, we assembled octameric nucleosome core particles (NCPs) from synthetic 145–bp human α–satellite DNA and the human core histone proteins H2A, H2B, H3.1 and H4. The forward DNA strand contained either a single 6–4PP lesion (Fig. 1a) or a damage mimic that comprises two consecutive tetrahydrofuran–nucleotide abasic sites (THF2), which is a known UV–DDB ligand21,22. We prepared a 6–4PP nucleosome with the lesion −22/−23 bp from the dyad axis (NCP6–4PP–UV–DDB) to maximize its affinity for UV–DDB22. The complex of NCP6–4PP and full–length recombinant human UV–DDB (NCP6–4PP–UV–DDB) was reconstructed to a resolution of 4.3 Å using single–particle cryo–electron microscopy (cryo–EM) (Fig. 1b, c, Extended Data Fig. 1a–g). The 6–4PP lesion is situated next to histone H3 α–helix α1, near residues 64–84 (Extended Data Fig. 1h, i). The primary contacts are between the DDB2 propeller and the minor groove of the nucleosome at superhelical locations (SHLs) −2 and −3. DDB2 is angled at approximately 60° with respect to a plane parallel to the nucleosomal disc (Fig. 1b), whereas the DDB2 N–terminal region (residues 1–54) is disordered23 (Extended Data Fig. 1j, k). DDB1 projects away from the nucleosome core and does not interact with the nucleosome (Fig. 1b–d).

We next applied focused refinement methods and obtained an improved map at 4.2 Å resolution by masking out DDB1 (Extended Data Fig. 1d). Side chains of positively charged residues in DDB2 (Arg112, Arg332, Arg370, Lys132 and Lys244) engage the DNA backbone within 6 bp (−20 to −25 bp from the dyad axis), together with hydrogen bonding and hydrophobic stacking interactions involving
Gln308, Tyr356, Ile394 and Tyr413 (Fig. 1e, f). DDB2 binds the minor groove around the lesion and inserts the β-hairpin loop composed of residues Phe334-Gln335-His336 as part of its recognition mechanism. Insertion of the β-hairpin requires the 6–4PP to be extruded into an extra-helical conformation, with DDB2 residues Gln335 and His336 stabilizing the two orphaned bases opposite the extruded lesion (adenines 50 and 51) (Fig. 1e, f, Extended Data Fig. 1i). The binding of UV-DDB triggers local DNA distortions around the lesion that are not found for NCP 6–4PP in isolation (Extended Data Fig. 2a–f).

Comparison of the structures of UV-DDB bound to either a 6–4PP-containing nucleosome or an isolated 6–4PP-containing DNA duplex revealed that most of the UV-DDB–DNA contacts are conserved between the two structures. Furthermore, the mechanism of recognition of UV-damaged nucleosomes does not differ from that found for isolated DNA duplexes harbouring pyrimidine dimers (Fig. 2b).

UV-DDB recognizes cyclobutane–pyrimidine dimers, 6–4PP and abasic sites through a shared mechanism. Our 3.9 Å resolution structure of UV-DDB bound to a nucleosome containing a THF2 lesion at position −22/−23 (NCP[THF2]−UV-DDB) reveals a configuration that is essentially indistinguishable from that of NCP[6–4PP]−UV-DDB (map correlation coefficient = 0.94) (Extended Data Fig. 3a–h, Extended Data Fig. 4a–c). UV-DDB also detects cyclobutane–pyrimidine dimers in nucleosomes (NCP[PD]) (Extended Data Fig. 3i, j), and we expect that this recognition proceeds in the same manner as for 6–4PP and THF2.

The DDB1 B-domain—which is the second of the three WD40 propeller domains in the protein—is partially disordered in our nucleosome structure, as is expected considering its rotational mobility. The features of the DDB1 B-domain blades 1, 2, 3, 6, and 7 were sufficiently defined in the cryo-EM map to enable us to place a coordinate model of this domain. This, in turn, enabled positioning of the CUL4 ligase arm, leading to a model of the CRL4DDB2 E3 ligase complex bound to a nucleosome (Fig. 2c). The rotation of the ligase around the nucleosome core explains the ubiquitination of nearby histone tails and diverse core histones, as has been reported previously.

To quantify the binding of UV-DDB to nucleosomes, we developed a fluorescence polarization competition assay. Cy5-labelled 15-bp oligonucleotide with a single THF abasic nucleotide damage site (Cy5–15-bp-DNA[THF]; 10 nM) was incubated with 10 nM UV-DDB; increasing amounts of unlabelled nucleosomes were then added to remove the labelled DNA from the complex by competition. We found that UV-DDB binds nucleosomes carrying 6–4PP at position −22/−23 with a half-maximal effective concentration (EC50) that is only 1.9-fold lower than that for the binding of the identical nucleosome-free DNA (Extended Data Fig. 5a). Additionally, the apparent affinities of UV-DDB for nucleosome-free 145-bp DNA and nucleosomal DNA containing a THF2 double abasic site mimic were indistinguishable (Extended Data Fig. 5b).

Dna damage recognition by UV-DDB at these solvent-exposed nucleosomal positions is therefore not obstructed by the nucleosome.

Read–out of occluded UV lesions

We next examined the accessibility of the damaged loci beyond the exposed minor groove position −22/−23 to UV-DDB. We modelled the binding of UV-DDB to the fraction of nucleosomal DNA that covers SHL−1.5 to SHL−2.5, and calculated the extent of steric clashes between UV-DDB and the nucleosome (Fig. 2d, Methods). The overall distribution of clash scores along this superhelix segment is U-shaped, with UV-DDB best accommodated at the solvent-facing minor groove, as described by its cryo-EM structure (Fig. 1). Severe steric clashes are predicted when the lesion is positioned away from the solvent face towards the nucleosome core. We define the high-accessibility locus at the centroid of the U-shaped profile (−21/−22) as position zero, and subsequently refer to the neighbouring −22/−23 position used for structure determination (Fig. 1) as the −1 position.
We next examined THF2 lesions at different sites before and after SHL–2 using electrophoretic mobility shift assays (Extended Data Fig. 5c) and fluorescence polarization competition assays (Fig. 2e, Extended Data Fig. 5d, e). Nucleosomes with lesions positioned 1-bp to 4-bp upstream or downstream of the −1 position retained UV-DDB binding—albeit with a decrease in binding strength (EC\textsubscript{50}) of up to 20-fold—as the lesion is moved away from the solvent-exposed locus towards the histone core (Fig. 2e, Extended Data Fig. 5e). The highest-affinity site was the zero position, which also had the highest predicted accessibility (Fig. 2d). Conversely, UV-DDB showed no relative difference in binding affinities for THF2 sites located on a free 145-bp DNA duplex (Extended Data Fig. 5d, e). Nucleosomes thereby render inward-facing lesions less accessible to DDB2, in line with predicted steric clashes (Fig. 2d), although UV-DDB binding still occurs at these sites.

To elucidate how UV-DDB accommodates these clashes and achieves damage recognition at inward-facing sites, we determined cryo-EM structures of an isolated NCP with THF2 located at position −3 (NCP\textsubscript{THF2}(−3)) and its complex with UV-DDB (NCP\textsubscript{THF2}(−3)–UV-DDB). The resulting cryo-EM maps were calculated at 3.5 Å and 4.1 Å resolution, respectively (Extended Data Fig. 5f–o). In NCP\textsubscript{THF2}(−3)–UV-DDB, DDB2 binds the damaged nucleosome without clashing with the core histones or with the two DNA gyres (Fig. 3a). UV-DDB, however, approaches the nucleosome at an angle that differs from the predicted model by approximately 95° (Extended Data Fig. 5p). Furthermore, the DDB2 position in NCP\textsubscript{THF2}(−3)–UV-DDB did not match the expected DNA register for the human α-satellite DNA repeat that we observed earlier (Fig. 1b, c). Notably, the 145-bp DNA in NCP\textsubscript{THF2}(−3)–UV-DDB is not wrapped symmetrically around the dyad axis of the nucleosome, with 72 bp on either side of 1 bp sitting on the axis. Instead, the DNA in NCP\textsubscript{THF2}(−3)–UV-DDB shows an asymmetric 69-bp and 75-bp divide of the human α-satellite repeat bisected by the nucleosome dyad axis (Fig. 3b). In the repositioned DNA register, the lesion at position −3 (−24/−25 from the dyad axis) instead corresponds to UV-DDB located −21/−22 bp from the dyad axis. Repositioning of the DNA register avoids predicted clashes between UV-DDB and the nucleosome and enables binding (Fig. 3d, e). Notably, the lesion in the repositioned register coincides with the high-affinity and high-accessibility zero-position locus.

To understand whether the presence of the THF2 damage itself is sufficient to trigger the 3-bp translation of the human α-satellite DNA repeat, or whether the shift by 3 bp in the nucleosomal DNA register is promoted by UV-DDB binding, we compared the structures of...
UV-DDB shifts nucleosome DNA registers

We next used an orthogonal approach to examine the effect of UV-DDB on the THF2 nucleosome register in solution. DNase I digestion of lesion-containing NCPs in the presence and absence of UV-DDB followed by next-generation sequencing enabled us to define and count all nucleosome registers within the sample. As predicted from our structural studies, no register shift was observed for UV-DDB binding to the THF2(−1) nucleosome relative to the free THF2(−1) nucleosome (Fig. 3d, Extended Data Fig. 6a–c). Conversely, the addition of UV-DDB to the THF2(−3) nucleosome resulted in a register shift of 3 bp. Sequencing further enabled us to count reads in different registers to determine the population ratio of unshifted (N) versus register-shifted (N_shift) nucleosomes. This revealed that the lesion, all major DNA–histone contacts are maintained but are register-shifted by precisely three nucleotides, as such, in their interactions with histones, nucleotides (n) are thus replaced by nucleotides (n − 3) in the presence of UV-DDB (Fig. 3c).

Register equilibrium at occluded sites

We next studied a cryo-EM structure of UV-DDB bound to a nucleosome with a THF2 lesion at position +1 (NCP_THF2(+1)–UV-DDB; −20/−21 from the dyad axis). Single-particle cryo-EM analysis including 3D classification resulted in two 3D reconstructions, denoted NCP_THF2(+1)–UV-DDB class A and NCP_THF2(+1)–UV-DDB class B (Extended Data Fig. 7a–h), which describe two distinct conformations. In the final reconstruction, NCP_THF2(+1)–UV-DDB class A includes about 2.6 times more particles than NCP_THF2(+1)–UV-DDB class B, and its estimated resolution was 4.5 Å (4.2 Å after focused refinement) (Extended Data Fig. 7c). UV-DDB binds the DNA minor groove of nucleosomes at the +1 position, tilted approximately 10° relative to the histone octamer disc (Fig. 4a). This places DDB2 blade 1 (residues 150–156) and blade 2 (residues 195–200) in very close proximity (6–8 Å) to the neighbouring DNA (Fig. 4b). The binding mode of NCP_THF2(+1)–UV-DDB class A agrees well with the predicted model (Fig. 2d) and did not require a DNA register shift.

The smaller subpopulation of particles used to calculate the NCP_THF2(+1)–UV-DDB class B model was reconstructed to a resolution of 4.8 Å (4.6 Å after focused refinement) (Fig. 4c, Extended Data Fig. 7c). Superposition of NCP_THF2(+1)–UV-DDB classes A and B revealed a difference in the translational setting of the nucleosomal DNA by +1 bp to account for the well-defined DDB2 (Fig. 4d). Thus, UV-DDB in NCP_THF2(+1)–UV-DDB class B is present in an apparent zero register (−21/−22 bp from the dyad axis). The translational setting in NCP_THF2(+1)–UV-DDB class B matches that in the repositioned NCP_THF2(−3)–UV-DDB structure (Extended Data Fig. 7i, j). UV-DDB exists in equilibrium between the +1 position (NCP_THF2(+1)–UV-DDB class A) and a register-shifted population
with the lesion in the highly accessible zero position (NCPTFH2(+1)–UV-DDB class B) (Fig. 2d). A DNase I assay on the UV-DDB-bound +1 nucleosome also identified an increased population of α-satellite DNA molecules shifted by 1 bp towards the dyad axis compared to the −1 nucleosome (Extended Data Fig. 6c), which is consistent with the register-shifted subpopulation seen by cryo-EM.

In NCPTFH2(−3)–UV-DDB, the DNA register shifted by 3 bp towards the dyad axis with respect to the NCPTFH2(−1)–UV-DDB structure, whereas for NCPTFH2(−1)–UV-DDB class B the register moved 1 bp away from the dyad axis (Fig. 4e, Extended Data Fig. 8a). The precise nucleosome register stabilized by UV-DDB and the bidirectional DNA movement thus depend on the position of the lesion.

We next investigated whether the affinity of UV-DDB for nucleosomes is sensitive to the strand on which a DNA lesion is situated. In our previous experiments we used nucleosomes with DNA lesions only on the Watson strand; we therefore investigated analogous nucleosomes that contained a THF2 lesion in the −1 position on the Crick strand. In comparison, UV-DDB-bound these nucleosomes with an approximately 4.7-fold lower EC50, which is indicative of weaker binding (Fig. 5a, b, Extended Data Fig. 8d). Accordingly, a THF2 modelled on the Crick strand binds UV-DDB by approximately 180° relative to the −1 structure and translates it more closely towards the histone octamer (Fig. 5a, b, Extended Data Fig. 8b). We also examined the effect of transposing lesions from the Watson strand in SHL−2.5 to the reverse complement Crick strand in SHL+3.5. The accessibility and affinities of the lesions in the −1 position were indistinguishable for SHL−2.5 and SHL+3.5; conversely, the −3 positions had lower relative affinity and more predicted clashes in SHL+3.5 than in SHL−2.5 (Fig. 5c, Extended Data Fig. 8c, d). We also examined the effect of the nucleosomal DNA positioning strength. In the 601 and 601L positioning sequences—both of which are more strongly positioning than α-satellite DNA—a lesion at the predicted SHL−2 position in 601 or 601L nucleosomes was less tightly bound than its counterpart in α-satellite DNA (Fig. 5d, Extended Data Fig. 8d). The local environment, the accessibility of the lesions and the strength of nucleosome–DNA contacts thus govern the affinity of UV-DDB for damage in chromatin.

**Discussion**

The nucleosomal architecture, with two DNA gyres wrapped around a histone octamer core, imposes a spatial barrier that severely restricts the access of proteins to DNA-embedded pyrimidine dimers (Extended Data Fig. 8e).

In chromatinized DNA, nucleosomal pyrimidine dimers show a slight positional preference for solvent-exposed minor grooves, where they are also more rapidly repaired. We demonstrate that UV-DDB can identify these lesions directly with high affinity. Conversely, although lesions at inward-facing positions that are inaccessible to solvent are also recognized, the apparent affinities are one or two orders of magnitude lower. In nucleosomes, accessible lesions are bound by UV-DDB directly, whereas occluded sites require additional steps to select for a predominantly register-shifted form of the nucleosome. A nucleosome with a THF2 lesion at the inward-facing −3 position (Fig. 3a, c), for example, exists predominantly in an unshifted, canonical α-satellite DNA register in the absence of UV-DDB, and translates into a predominantly 3-bp-shifted register only when UV-DDB is present (Extended Data Fig. 6f). Partially accessible lesions—for example cases in which DDB2 binds but is placed in close proximity to the nucleosome with potential repulsive interactions, as found for NCPTFH2(+1)–UV-DDB—are present in equilibrium between shifted and unshifted registers.

By exploiting nucleosome-register dynamics, UV-DDB binding enforces a nucleosomal register in a bidirectional manner. The detection of occluded nucleosomal lesions by UV-DDB is independent of ATP-driven chromatin remodellers in vitro and does not require octamer disassembly or looping off of nucleosomal DNA (site exposure). Instead, UV-DDB overcomes the intrinsic phasing power of the strongly positioned human α-satellite DNA repeat and, after sliding, helps to stabilize the lesion in an accessible locus. We designate this mechanism of DNA damage read-out ‘slide-assisted site exposure’ (SASSE).

For lesions located around SHL−2 we find a preferred SASSE repositioning locus at position zero (−20/−21 bp from the dyad axis), where less accessible lesions downstream and upstream reposition when bound by UV-DDB. The same principles seem to apply to SHL+3. We suggest that each superhelix segment, with its U-shaped predicted accessibility profile, has a favoured low-energy locus within its outward-facing minor groove to which lesions reposition through SASSE to minimize steric clashes (Extended Data Fig. 8c). In what is probably attributable to the local asymmetry of nucleosomes, we find that changing the strand (Fig. 5a, b, Extended Data Fig. 8b, d) or the location of a lesion from one SHL to another (Fig. 5c, Extended Data Fig. 8c, d) affects UV-DDB accessibility, the extent of SASSE and the apparent affinity. A similar SASSE mechanism can also be invoked for damage recognition of pyrimidine dimers in the context of multiple nucleosomes (Extended Data Fig. 8f).

We envisage that SASSE activity is shared by other high-affinity DNA-binding proteins. A transcription factor that is able to bind a nucleosomal site with sufficient affinity should similarly be able to enforce its binding register. Whereas remodeller-mediated nucleosome eviction, and nucleosome-dynamic-based site exposure (Extended Data Fig. 6f) are established mechanisms for accessing nucleosomal DNA, our findings point to SASSE as an additional strategy whereby occluded loci can be accessed through changes in the translational register. Our work presents a structural and mechanistic framework for UV-damage recognition in chromatin, and further explains how other
DNA-binding proteins—such as transcription factors—can in principle access nucleosomal DNA.

Online content
Any Methods, including any statements of data availability and Nature Research reporting summaries, along with any additional references and Source Data files, are available in the online version of the paper at https://doi.org/10.1038/s41586-019-1259-3.

Received: 31 October 2018; Accepted: 1 May 2019; Published online 29 May 2019.

1. Thoma, F. Repair of UV lesions in nucleosomes—inntrinsic properties and remodeling. DNA Repair (Amst.) 4, 855–869 (2005).
2. Rodriguez, Y. Hinz, J. M. & Smerdon, M. J. Accessing DNA damage in chromatin: Preparing the chromatin landscape for base excision repair. DNA Repair (Amst.) 32, 113–119 (2015).
3. Hanawalt, P. C. & Spivak, G. Transcription-coupled DNA repair: two decades of progress and surprises. Nat. Rev. Mol. Cell Biol. 9, 958–970 (2008).
4. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature 389, 251–260 (1997).
5. McGinty, R. K. & Tan, S. Nucleosome structure and function. Chem. Rev. 115, 2255–2273 (2015).
6. Sugasawa, K. et al. UV-induced ubiquitylation of XPC protein mediated by UV-DDB-ubiquitin ligase complex. Cell 121, 387–400 (2005).
7. Groissman, R. et al. The ubiquitin ligase activity in the DDB2 and CSA complexes is differentially regulated by the COP9 signalsome in response to DNA damage. Cell 113, 357–367 (2003).
8. Cavadini, S. et al. Cullin–RING ubiquitin E3 ligase regulation by the COP9 signalsome. Nature 531, 596–603 (2016).
9. Wang, H. et al. Histone H3 ubiquitination by the CUL4A-DBD-ROC1 ubiquitin ligase facilitates cellular response to DNA damage. Mol. Cell 22, 383–394 (2006).
10. Lehrman, A. R. DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. Biochimie 85, 1101–1111 (2003).
11. Lujismerburg, M. S. et al. Dynamic in vivo interaction of DDB2 ubiquitin ligase with UV-damaged DNA is independent of damage-recognition protein XPC. J. Cell Sci. 120, 2706–2716 (2007).
12. Puumaalainen, M. R. et al. Chromatin retention of DNA damage sensors DDB2 and XPC through loss of p97 segregate causes genotoxicity. Nat. Commun. 5, 3695 (2014).
13. Fei, J. et al. Regulation of nucleotide excision repair by UV-DDB: prioritization of damage recognition to internucleosomal DNA. PLoS Biol. 9, e1001183 (2011).
14. Fischer, E. S. et al. The molecular basis of CRL4(DDB1)CSN ubiquitin ligase architecture, targeting, and activation. Cell 147, 1024–1039 (2011).
15. Zhu, G. & Chang, E. Xeroderma pigmentosum group E cells lack a nuclear factor that binds to damaged DNA. Science 242, 564–567 (1988).
16. Lehrman, A. R., McGibbon, D. & Stefani, M. Xeroderma pigmentosum. Orphanet J. Rare Dis. 6, 70 (2011).
17. Scrima, A. et al. Structural basis of UV DNA-damage recognition by the DDB1–DDB2 complex. Cell 135, 1213–1223 (2008).
18. Yeh, J. J. et al. Damaged DNA induced UV-damaged DNA-binding protein (UV-DDB) dimerization and its roles in chromatinized DNA repair. Proc. Natl Acad. Sci. USA 109, E2737–E2746 (2012).
19. Li, G., Levitus, M., Bustamante, C. & Widom, J. Rapid spontaneous accessibility of nucleosomal DNA. Nat. Struct. Mol. Biol. 12, 46–53 (2005).
20. Zhu, F. et al. The interaction landscape between transcription factors and the nucleosome. Nature 562, 76–81 (2018).
21. Wittschieben, B. O., Iwai, S. & Wood, R. D. DDB1–DDB2 (xeroderma pigmentosum group E) protein complex recognizes a cyclobutane pyrimidine dimer, mismatches, apurinic/apyrimidinic sites, and compound lesions in DNA. J. Biol. Chem. 280, 39982–39986 (2005).
22. Osakabe, A. et al. Structural basis of pyrimidine–pyrimidone (6–4) photoproduct recognition by UV-DDB in the nucleosome. Sci. Rep. 5, 16330 (2015).

Acknowledgements We thank A. Osakabe and Y. Arimura for discussion and histone preparation. This work was supported by the Novartis Research Foundation and received funding from the Swiss National Science Foundation through Sinergia grant number CRSII3_160734/1 to N.H.T.; from the European Research Council under the European Union’s Horizon 2020 Research and Innovation programme grant agreement, grant numbers 666068 to N.H.T. and 667951 to D.S., and the Innovation programme unter Marie Sklodowska-Curie grant number 705354 and an EMBO Long-Term Fellowship to R.S.G.; in part from the Uehara Memorial Foundation to S.M.; from JSPS KAKENHI grant number JP18H05534 to H.K. and JP16H06307 to K.S.; and from the Platform Project for Supporting Drug Discovery and Life Science Research (BINDS) from AMED under grant number JP18am0101076 to H.K.

Reviewers We thank Orlando Schärer, Bennett Van Houten and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions S.M., S.C. and N.H.T. conceived and planned the experiments. S.M. prepared samples for biochemical and structural studies and performed electrophoretic mobility shift assays and fluorescence polarization assays. S.C. performed cryo-EM and analysis. R.D.B. interpreted the cryo-EM structures and prepared the atomic models, with S.C. providing input. R.D.B., S.C. and N.H.T. analysed the structures. A.P. and J.R. helped to develop the fluorescence polarization assays, S.M. carried out the functional experiments, and J.R., S.M. and N.H.T. analysed the results. A.D.S. contributed to cryo-EM processing and developed CryoFLARE. S.I. and J.Y. synthesized 6–4PP-containing oligonucleotides. S.M. performed biochemical analysis with guidance from K.S. and H.K. R.S.G. developed and analysed the DNase I footprinting assay with guidance from D.S. The research was directed by N.H.T., and all authors contributed to writing the manuscript.

Competing interests The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-019-1259-3.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1259-3.
Reprints and permissions information is available at http://www.nature.com/reprints.

Correspondence and requests for materials should be addressed to N.H.T.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Human histone preparation. Human histones were expressed and purified as described previously23. A DNA fragment encoding the histone H2B(T122C) mutant, in which Thr122 was replaced by cysteine using site-directed mutagenesis, was prepared as described previously23. Reconstitution of the H2A–H2B(T122C) complex, the H3.1–H4 complex, and the histone octamer were performed as described previously23. Lyophilized histones were mixed equally in 20 mM Tris-HCl pH 7.5 buffer, containing 7 M guanidinium hydrochloride and 20 mM 2-mercaptoethanol. Samples were dialysed against 2 M NaCl, 10 mM Tris-HCl pH 7.5, 1 mM EDTA and 2 mM 2-mercaptoethanol. The resulting histone complexes were purified by size-exclusion chromatography (Superdex 200; GE Healthcare).

DNA preparation. Oligonucleotides containing 6–4 photoproduct (6–4PP) or cyclobutane–pyrimidine dimers (CPDs) were synthesized from 6–4PP24 or CPD building block (Glen Research). Oligonucleotides containing THF were purchased from Sigma. By mixing with complementary oligonucleotides in a 1:1 ratio, damaged double-strand DNA was formed. After heating to 95 °C for 5 min, the DNA was annealed by slow cooling to room temperature; damaged and undamaged DNA duplexes were then ligated by T4 DNA ligase (New England BioLabs)25. Ligated DNA was purified by native polyacrylamide gel electrophoresis using a Prep Cell (Bio-Rad) with TE buffer (10 mM Tris-HCl pH 7.5 and 100 μM EDTA). The location of damage of these DNAs is described in Fig. 1a. The –5 position in SHL (−26−−27) is structurally equivalent to the −5 position at the −16/−17 position in SHL-1. Because the −26/−27 could not be stably assembled, we used −16/−17 in the neighbouring SHL instead (designated −5x).

Nucleosome assembly. The DNA and histone octamer complex were mixed in a 1:1.5 molar ratio in the presence of 2 M KCl. The samples were dialysed against refolding buffer (RB) high (10 mM Tris-HCl pH 7.5, 2 M KCl, 1 mM EDTA and 1 mM dithiothreitol (DTT)). The KCl concentration was gradually reduced from 2 M to 0.25 M using a peristaltic pump with RB low (10 mM Tris-HCl pH 7.5, 250 mM NaCl, 1 mM EDTA and 1 mM DTT) at 4 °C. Samples were further dialysed against RB low buffer at 4 °C overnight. Reconstituted nucleosomes were incubated at 55 °C for 2 h followed by purification on native polyacrylamide gel gel electrophoresis using a Prep Cell apparatus (Bio-Rad) in TCS buffer (10 mM Tris-HCl pH 7.5 and 250 μM TCEP) to remove non-specific complexes formed between free DNA and histones.

Preparation of UV-DDB. Human wild-type DDB1, wild-type DDB2, and a DDB2 variant lacking amino acids 1–40 (DDB2(−1−40)) were subcloned into pAG-derived vectors23. The complex of DDB1 and DDB2(ΔN) is referred to as UV-DDB(ΔN). Recombinant proteins were expressed in 8-litre cultures of E. coli. Cultures were grown in LB medium containing the Bac-to-Bac system (Thermo Fisher). Cells were cultered at 27 °C. Cultured 2 days after infection, resuspended in lysis buffer (50 mM HEPES pH 7.4, 250 mM KCl, 1 mM EDTA, 1% SDS) and chilled on ice for 10 min. Samples were treated with Proteinase K (10 μg/mL) followed by purification on a DNA preparation.

Negative stain EM. For structure determination of the NCPTHF2(1−3)–UV-DDB complex, after gel filtration a 3-bp shift (N_Shift) in the nucleosome dyad position) and purified human UV-DDB were mixed in a 1:2 molar ratio in GF buffer (50 mM HEPES pH 7.4, 50 mM NaCl and 250 μM TCEP) and incubated on ice for 30 min. NCP complexes were treated with a titration (0.12 μg/mL, 0.06 μg/mL, 0.015 μg/mL) of UV-DDB and incubated on ice for 30 min. NCP complexes were chosen for the competitive titration assays the UV-DDB bound to the fluorescent oligo (Cy5−15−bp–DNA(THF) was titrated to saturation with increasing concentrations of UV-DDB. The curves were numerically fitted to a model describing the binding isotherm as implemented in DynaFit25. For fitting, we fixed the concentration of UV-DDB and the label, and refined the Kd value, the offset, and the gain. We then performed the same titration using 10 mM Cy5−15−bp–DNA(THF) and UV-DDB and titrated the complex with increasing concentrations of NCP(THF)(1−3)–UV-DDB. The two isotherms—for UV-DDB binding to the label and for UV-DDB engaging the unlabelled complex competing off the label—were implemented in DynaFit25. Numerical fitting was carried out with the Kd value for Cy5−15−bp–DNA(THF) fixed, and the Kd values for the NCP(THF)(1−3)–UV-DDB complex, as well as gain and offset, were fitted. Given the tight binding of the UV-DDB to Cy5−15−bp–DNA(THF) with affinity constants approximately tenfold lower than the concentration of the label, these values need to be viewed as an estimation, in line with the literature36.

DNase I nucleosome footprinting assay. NCPs with undamaged α-satellite DNA or carrying THF2 at position −1, −3 or +1 and purified human UV-DDB were mixed in a 1:1 molar ratio in GF buffer (50 mM HEPES pH 7.4, 50 mM NaCl and 250 μM TCEP) and incubated on ice for 30 min. NCP complexes were treated with a titration (0.12 μg/mL, 0.06 μg/mL, 0.015 μg/mL) of DNase I (New England BioLabs, M0303S) in the presence of 2.5 mM MgCl₂ and 0.5 mM CaCl₂ (for 5 min at 37 °C). The reaction was stopped by adding an equal volume of Stop Buffer (200 mM NaCl, 30 mM EDTA, 1% SDS) and chilled on ice for 10 min. Samples were treated with a solution of RNase A, 10 μg/mL, and treated with Proteinase K and RNAase A for 30 min. The resulting DNA fragments were purified by gel filtration (Superdex 200; GE Healthcare) in 50 mM HEPES pH 8.0, 500 mM NaCl and 250 μM TCEP and stained with 2% (w/v) uranyl acetate. Data were collected using a Tecnai T12 electron microscope (Thermo Fisher) operating at 100 kV with a pixel size of 3.08 Å at the specimen level. Images were recorded with a TVIPS TemCam F416 with varying defocus (−220 m to −150 m). All particles (12,774) were selected using e2boxer.py and processed with SPARX25. After two-dimensional (2D) classification with iterative stable alignment and classification in SPARX, the best 115 2D class averages were used for 3D ab initio model generation with ssxview.py from SPARX.

Cryo-EM data acquisition. For structure determination of the NCP(THF)(1−3)–UV-DDB complex, after gel filtration a 3-bp sample (around 0.5 mg/mL) was applied to Lacey carbon grids (Ted Pella). All other NCP–UV-DDB assemblies were reconstituted by mixing NCP and UV-DDB in a 1:1 molar ratio in GF buffer. After 30 min incubation on ice, a 3 μL sample (between around 1 mg/mL and 10 mg/mL) was transferred to the specimen grid (Quantifoil Micro Tools), blotted, and stained with 2% (w/v) uranyl acetate. Data were collected using a Tecnai T12 electron microscope (Thermo Fisher) operating at 100 kV with a pixel size of 3.08 Å at the specimen level. Images were recorded with a TVIPS TemCam F416 with varying defocus (−0.5 μm to −2.0 μm). All particles (12,774) were selected using e2boxer.py and processed with SPARX25. After two-dimensional (2D) classification with iterative stable alignment and classification in SPARX, the best 115 2D class averages were used for 3D ab initio model generation with ssxview.py from SPARX.
around 2.5 mg ml⁻¹ was applied to either R1 2.1/3 UltraAoil Quantifoil grids (Quantifoil Micro Tools) or R2/2 Quantifoil holey carbon grids (Quantifoil Micro Tools). Glow discharging was carried out in a Solarus plasma cleaner (Gatan) for 12 s in an H₂O₂ environment. After 2–3 s blotting and 1 s post-blot incubation, the grids were vitrified using a Leica EM GP plunger (Leica Microsystems) operated at 4°C and 85% humidity. Data were collected automatically with EPU (Thermo Fisher) on a Cs-corrected (CEOS GmbH, Heidelberg, Germany) Titan Krios (Thermo Fisher) electron microscope at 300 kV. Zero-energy-loss micrographs were recorded using a Gatan K2 summit direct electron detector (Gatan) located after a Quantum-LS energy filter (slit width of 20 eV). For the NCPT²⁻(1–)

- UV-DDB complex, acquisition was performed in energy-filtered transmission electron microscopy (EFETM) super-resolution mode at a nominal magnification of 105,000×, yielding a pixel size of 0.55 Å at the specimen level. For all other datasets the acquisition was performed in EFETM counting mode at a nominal magnification of 130,000×, yielding a pixel size of 0.86 Å at the specimen level. Except for the NCPT carrying THF2 at position –3 (–24–25) and NCFT–UV-DDB with THF2 at position –1 (–22–23), all other datasets were collected with a Volta phase plate (VPP). The C2 aperture was 50 μm and, for non VPP data, the objective aperture was 100 μm. Data for different NCPTs and UV-DDB assemblies have been recorded with dose rates of between 3.5 and 5 e⁻ per pixel per second and total doses of between 40 and 60 e⁻ Å⁻² (details in Extended Data Tables 1, 2). The exposures were fractionated into 40 frames. The targeted defocus values ranged from 0.5 μm to 3 μm for non VPP datasets and from 0.4 μm to 0.6 μm from VVP datasets. The phase plate position was changed regularly (every 80–100 exposures) to target phase shifts between 20° and 130°.

Cryo-EM image processing. Real-time evaluation together with acquisition by EPU (Thermo Fisher) was performed with CryoFLARE (in-house development; http://www.cryoflare.org). This pre-processing step includes drift correction of micrograph stacks, contrast transfer function (CTF) determination and particle autopicking. Drift correction was performed either with Unblur⁴⁰ or MotionCor²⁴¹. For MotionCor2, a sum of 40 frames was generated with and without applying a dose-weighting scheme and CTF was fitted using GCTF⁴² on the non-dose-weighted sums. Particles were picked using Gautomatch (K. Zhang, MRC Laboratory of Molecular Biology, Cambridge) on the dose-weighted sums. For Unblur, the motion-corrected averages obtained after whole-image drift correction were used for CTF estimation (GCTF), and further processing included particle polishing in RELION⁴³. In this case, movie processing was performed and the number of frames included in the final reconstructions was chosen to target a total dose of around 28 e⁻ Å⁻². For details of image processing for all samples, see Extended Data Figs. 1–3, 7 and Extended Data Tables 1, 2.

The resolution values reported for all reconstructions are based on the gold-standard Fourier shell correlation curve (FSC) at 0.143 criterion⁴⁴ and restraints for secondary structure and to reference models for UV-DDB derived refinement was performed with phenix.real_space_refine with supplementary template models into the cryo-EM maps, as well as model building, were carried out using phenix.real_space_refine with supplementary restraints for secondary structure and to reference models for UV-DDB derived from PDB entries 4E54 and 3EI1 after re-refinement. MOLPROBITY⁴⁵ was used for model validation. Owing to twofold rotational ambiguity in our cryo-EM reconstructions without UV-DDB, we modelled the DNA damage regions in the NCPT²⁻(1–) and NCFT²⁻(1–) structures (Extended Data Figs. 2b, 5) symmetrically across the nucleosome dyad axis, assuming a superposition of damaged and undamaged DNA at equal occupancy. Data collection and refinement statistics are presented in Extended Data Tables 1, 2.

Density map segmentation and figure preparation. Cryo-EM maps were segmented using phenix.map_box from the PHENIX suite. Structural figures were produced using PyMOL (The PyMOL Molecular Graphics System, v2.0 Schrödinger, LLC).

Clash scores and contact surface area calculation. Clash scores for NCFT–UV-DDB models were calculated using a PyMOL script developed in-house (available from https://doi.org/10.8604/m9.figshare.7969655.v4). In brief, a UV-DDB probe containing an appropriately positioned DNA fragment for superimposing on a nucleosome template model was placed in all possible binding positions and clash scores for each map were taken as the total number of residues in UV-DDB that contained atoms closer than 1 Å to nucleosomes. Accessible contact surface area calculations were carried out using AREAIMOL from the CCP4 suite⁴⁶, assuming an inward-facing probe sphere.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Atomic coordinates and cryo-EM maps are deposited in the Protein Data Bank (PDBe) (D Etwatr, Stanford Microscopy Data Bank (EMDB) under accession codes 6R8Y and EMD-4762 (NCPT²⁻–UV-DDB); 6R8Z and EMD-4763 (NCFT²⁻(1–)–UV-DDB); EMD-6981 and EMD-4765 (NCFT²⁻(1–)–UV-DDB); 6R89 and EMD-4764 (NCPT²⁻(1–)–UV-DDB class A); 6R89 and EMD-4766 (NCPT²⁻(1–)–UV-DDB class B); 6R83 and EMD-4767 (NCPT²⁻(3–)).

31. Iwai, S., Shimizu, M., Kamiya, H. & Ohtsuika, E. Synthesis of a phosphoramidite coupling unit of the pyrimidine (6–4) pyrimidine photocyst and its incorporation into oligodeoxynucleotides. J. Am. Chem. Soc. 118, 7642–7643 (1996).

32. Abdulkhaiman, W. et al. A set of baculovirus transfer vectors for screening of antibody tags and parallel expression strategies. Anal. Biochem. 385, 383–385 (2009).

33. Marks, B. D. et al. Multiparameter analysis of a screen for progesterone receptor ligands: comparing fluorescence lifetime and fluorescence polarization measurements. Assay Drug Dev. Technol. 3, 613–622 (2005).

34. Kuzmič, P. DynaFiT—a software package for enzymology. Methods Enzymol. 467, 247–280 (2009).

35. Thoma, N. & Goody, R. S. in Kinetic Analysis of Macromolecules: A practical approach (ed. Johnson, K. A.) 153–170 (Oxford Univ. Press, 2003).

36. Reardon, J. T. et al. Comparative analysis of binding of human damaged DNA-binding protein (XPC) and Escherichia coli damage recognition protein (UvrA) to the major ultraviolet photoproducts: T[T]C[T], T[4T]C[T], T[4T]T[T] and T[T][Dewar][T]. J. Biol. Chem. 286, 21301–21306 (2011).

37. Gaidatzis, D., Lerch, A., Hahne, F. & Stadler, M. B. Quasiligand quantification and annotation of short reads in Bioinformatics 31, 1130–1132 (2015).

38. Tang, G. et al. EMAN2: an extensible image processing suite for electron microscopy. J. Struct. Biol. 157, 47–55 (2007).

39. Hohn, M. et al. SPARX, a new environment for cryo-EM image processing. J. Struct. Biol. 175, 157–169 (2012).

40. Grant, T. & Grigorieff, N. Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. eLife 4, e06980 (2015).

41. Li, X. et al. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. Nat. Methods 10, 584–590 (2013).

42. Zhang, K. Gctf: Real-time CTF determination and correction. J. Struct. Biol. 193, 1–12 (2016).

43. Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. J. Struct. Biol. 180, 519–532 (2010).

44. Radermacher, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. J. Mol. Biol. 333, 721–745 (2003).

45. Chen, S. et al. High-resolution noise substitution to measure overfitted and validate resolution in 3D structure determination by single particle electron cryomicroscopy. Ultramicroscopy 135, 24–35 (2015).

46. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D 66, 213–221 (2010).

47. de la Rosa-Trevín, J. M. et al. Xnipp 3.0: an improved software suite for image processing in electron microscopy. J. Struct. Biol. 184, 321–328 (2013).

48. Zhang, K. S., Grimes, J. J. T. & Davey, C. A. DNA stretching and extreme kinking in the nucleosome core. J. Mol. Biol. 368, 1067–1074 (2007).
51. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. Acta Crystallogr. D 60, 2126–2132 (2004).
52. Nicholls, R. A., Long, F. & Murshudov, G. N. Low-resolution refinement tools in REFMAC5. Acta Crystallogr. D 68, 404–417 (2012).
53. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D 66, 12–21 (2010).
54. Winn, M. D. et al. Overview of the CCP4 suite and current developments. Acta Crystallogr. D 67, 235–242 (2011).
55. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods 14, 331–332 (2017).
56. Ekundayo, B., Richmond, T. J. & Schalch, T. Capturing structural heterogeneity in chromatin fibers. J. Mol. Biol. 429, 3031–3042 (2017).
Extended Data Fig. 1 | See next page for caption.
Extended Data Fig. 1 | Classification and refinement procedures for NCP6–4PP–UV-DDB. a, Representative negative-stain micrograph (left) and reference-free 2D class averages (right) obtained with ssisac.py (SPARX) for the NCPTHF2(−1)–UV-DDB complex. b, Ab initio model generated with sxviper.py in SPARX for the complex shown in a. c, Representative Volta phase plate (VPP) cryo-EM micrographs (left) and reference-free 2D class averages (right) for NCP6–4PP–UV-DDB. d, Classification and refinement procedures for the cryo-EM reconstruction of NCP6–4PP–UV-DDB. Three different microscope sessions (4,530 micrographs) were collected under identical imaging conditions and processed independently before merging the best particles to obtain the final high-resolution reconstruction. For each session, a small dataset was manually selected to obtain 2D class averages that were used for particle autopicking with RELION. Several rounds of 2D and 3D classification were necessary to obtain homogeneous datasets. The model shown in b was low-pass-filtered to 60 Å resolution and used as an initial model for the first round of 3D classification of each session. Given the total dose of 40 e− Å−2 over 40 frames, only frames 1 to 28 were included for movie refinement and particle polishing in RELION. To improve the resolution, the best particles from the three sessions were pooled and subjected to 3D classification into four classes. Refinement of the particles included in class II using a soft mask around the entire complex produced a 4.3 Å resolution map. Refinement of the same set of particles with a soft mask that excluded DDB1 produced a 4.2 Å resolution map. e, Gold-standard FSC curves for NCP6–4PP–UV-DDB (red) and for the same complex after masking out DDB1 (blue). f, Angular distribution of the particles included in the final models. g, Local-resolution-filtered map for NCP6–4PP–UV-DDB coloured by resolution (MonoRes47). h, The 6–4PP lesion is located next to H3 α–helix α1. i, Orphaned bases are stabilized by β-hairpin loop insertion. j, Fluorescence polarization dose–response curves using 10 nM Cy5-labelled 15-bp oligonucleotide with a single THF damage site (Cy5–15-bp-DNATHF) mixed with UV-DDB (0.3–200 nM). The interaction is measured and plotted as described in Methods. All data are shown as mean ± s.d. of three technical replicates. k, A 10 nM solution of a Cy5–15-bp-DNA3HF was mixed with 10 nM wild-type UV-DDB or the UV-DDB(ΔN) variant lacking residues 1–40 of DDB2, and counter-titrated with NCPTHF2(−1). Although the DDB2 N-terminal region (residues 1–40) contributes to nucleosome binding in biochemical assays, we did not find interpretable density for this segment; we also found no evidence of UV-DDB dimerization in our cryo-EM structures18.
Extended Data Fig. 2 | Classification and refinement procedures for NCP$^{6–4PP}$. a, Representative VPP cryo-EM micrograph (left) and reference-free 2D class averages (right) for NCP$^{6–4PP}$. b, Four microscope sessions with a total of 5,187 micrographs were collected under identical imaging conditions. All dose-fractionated micrograph stacks were subjected to beam-induced motion correction with MotionCor2. Initial processing was carried out with unweighted MotionCor2-corrected sums (including all frames). A small dataset was manually selected to obtain initial 2D class averages used for autopicking in RELION. A few rounds of 2D classification led to a dataset with 209,324 particles. A nucleosome map cut out from the 4.2 Å resolution map shown in Extended Data Fig. 1d was low-pass-filtered to 60 Å and used as an initial model. Three-dimensional classification into four classes allowed us to discard 54% of the particles (class I and class III). The remaining 106,417 particles were refined to 4.7 Å resolution. Re-extraction of the particles from dose-weighted micrographs and re-refinement with a mask led to a 4.1 Å resolution map. Particles were subjected to a final round of 2D classification before the last refinement to obtain a map at 4 Å resolution. c, Gold-standard FSC curves. d, Angular distribution of the particles included in the final model. e, Local-resolution-filtered map for NCP$^{6–4PP}$. f, Middle and right, the 4 Å resolution cryo-EM map of NCP$^{6–4PP}$ shown in two different views, with the 6–4PP (orange) tentatively assigned. Left, cryo-EM density (grey surface) of NCP6–4PP–UV-DDB, illustrating the undistorted nature of the nucleosome. At this resolution, the $\alpha$-satellite nucleosome carrying the damage is pseudo-symmetrical. Both SHL$^+2$ and SHL$^−2$ loci show continuous density around the expected site of the lesion. Whereas both half-sites are probably averaged in the process of classification, the continuous density in both loci suggests that the 6–4PP lesion is present predominantly in an apparent helical, non-extruded conformation.
Extended Data Fig. 3 | See next page for caption.
Extended Data Fig. 3 | Classification and refinement procedures for NCP\textsuperscript{THF2}(-1)–UV-DDB. a, Representative cryo-EM micrograph and reference-free 2D class averages for the NCP\textsuperscript{THF2}(-1)–UV-DDB complex. b, A total of 2,811 micrographs were collected over three sessions using identical imaging conditions. For each session, the model shown in Extended Data Fig. 1b was low-pass-filtered to 60 Å and used as initial model for the first round of 3D classification. For each session, a small dataset was manually picked to obtain initial 2D class averages used for autopicking in RELION. Several rounds of 2D and 3D classification were necessary to obtain homogeneous datasets. Given the accumulated dose of 60 e\textsuperscript{-} Å\textsuperscript{−2} equally distributed over 40 frames, frames 1 to 18 were retained during movie refinement and particle polishing in RELION. Further 3D classification into four models was performed by masking out DDB1. Next the particles included in class I and class IV were pulled together and subjected to refinement with a mask around the entire complex leading to a 3.9 Å resolution map. Refinement with a mask that excluded DDB1 led to an improved 3.8 Å resolution map. c, Gold-standard FSC curves for the 3.9 Å (blue) and 3.8 Å (purple) resolution map, respectively. d, Angular distribution of the particles included in the final models. e, Local-resolution–filtered maps for NCP\textsuperscript{6–4PP}–UV-DDB and NCP\textsuperscript{THF2}(-1)–UV-DDB coloured by resolution (MonoRes\textsuperscript{47}). f, g, The NCP\textsuperscript{6–4PP}–UV-DDB model fitted into the NCP\textsuperscript{THF2}(-1)–UV-DDB cryo-EM map, illustrating that the two models are structurally identical given the resolution. h, Damage location and sequence of NCP\textsuperscript{THF2}(-1)–UV-DDB and NCP\textsuperscript{CPD}(-1)–UV-DDB are indicated. i, As in Extended Data Fig. 1k with increasing amounts of competing nucleosomes containing 6–4PP, THF2, or CPD. All data are shown as mean ± s.d. of three technical replicates. j, Gel electrophoretic mobility shift assays were carried out with 40 nM Alexa Fluor 488-labelled undamaged nucleosome or nucleosomes containing 6–4PP, THF2, or CPD lesions. Gels were imaged using the Alexa Fluor 488 signal with a Typhoon Image Analyzer (GE Healthcare).
Extended Data Fig. 4 | Representative NCP–UV-DDB cryo-EM map segments. a, Different views for the unsharpened local-resolution-filtered map of the NCP<sup>6–4PP</sup>(−1)–UV-DDB complex. Together with the class averages (Extended Data Fig. 1a), this argues that the density observed is fully accounted for by a single UV-DDB complex engaged with a single damaged nucleosome. b, Representative, sharpened local-resolution-filtered maps of the NCP<sup>6–4PP</sup>(−1)–UV-DDB complex for which the central DDB1 B-domain features were masked out, showing map segments for the damaged DNA duplex; the DDB2–DNA–histone H3 interface, and H3/H4 position surrounding the damage, and the DDB2 β-hairpin loop involved in damage recognition. c, As in b but with map segments from NCP<sup>THF2</sup>(−1)–UV-DDB. d, Sharpened local-resolution-filtered map obtained for NCP<sup>6–4PP</sup>(−1)–UV-DDB with DDB1 masked out. e, Segments of H2A, H2A, H3, H4, DDB2 β-sheet and DDB2–6–4PP interface excised from the 4.2 Å resolution NCP<sup>6–4PP</sup>(−1)–DDB2 map shown in d.
Extended Data Fig. 5 | See next page for caption.
Extended Data Fig. 5 | Biochemical and biophysical characterization of NCP\(^{THF2}(−3)\) and NCP\(^{THF2}(−3)\)-UV-DDB. a, b, A 10 nM solution of Cy5–15-bp-DNATHF was mixed with 10 nM UV-DDB and counter-titrated with nucleosome or 145-bp DNA containing 6–4PP (a) or THF2 (b) at the −1 (−22/−23) position. Undamaged 145-bp DNA was used as a negative control. All data are shown as mean ± s.d. of three technical replicates. c, Gel electrophoretic mobility shift assays were carried out by mixing 40 nM of nucleosomes containing THF2 at different positions, −5∗ (−16/−17) to +3 (−18/−19), with increasing amounts of UV-DDB (0 to 320 nM). Gels were imaged by Alexa Fluor 488 with Typhoon Image Analyzer. d, As in b, but counter-titrated with different 145-bp DNA containing THF2 at the −5∗ (−16/−17) to +3 (−18/−19) positions. e, Estimation of EC\(_{50}\) from the counter-titration experiments (a, b and Fig. 2e). Data are shown as mean ± s.d. of three technical replicates. Lesions placed at position zero show the highest binding affinity, with an EC\(_{50}\) around 1.1-fold higher than those at the −1 (−22/−23) position used for structure determination in Fig. 1. However, the EC\(_{50}\) decreased approximately 2.5-fold for nucleosomes with THF2 lesions at the +1 (−20/−21) site, around eightfold at the +2 (−19/−20) site and around 15-fold at the +3 (−18/−19) site. A similar drop in affinity is found for lesions placed successively in the other direction from the −2 (−23/−24) site through to the −3 (−24/−25) and −4 (−25/−26) sites (Fig. 2e). f, Representative VPP cryo-EM micrographs (left) and reference-free 2D class averages (right) for the NCP\(^{THF2}(−3)\)-UV-DDB complex. g, Ab initio model generated with RELION for the complex shown in f, h. Two different microscope datasets were collected under identical imaging conditions leading to 3,890 micrographs. All dose-fractionated micrograph stacks were subjected to beam-induced motion correction with MotionCor2\(^{25}\). All frames (1–40) were included during this step. Further processing was carried out using MotionCor2-corrected sums that were filtered according to exposure dose (1 e\(^−\) Å\(^−2\) per frame). A small dataset was manually picked to obtain 2D class averages used for autopicking within RELION. The model shown in g was low-pass-filtered to 60 Å and used as initial model for the first round of 3D classification. Several rounds of 2D and 3D classification were necessary to obtain homogeneous datasets. The last 3D classification divided the dataset into six models. Refinement of the best particles with a soft mask around the entire complex led to a 4.1 Å resolution map. i, Representative conventional (no VPP) cryo-EM micrograph (left) and reference-free 2D class averages (right) for the isolated NCP\(^{THF2}(−3)\). j, A total of 2,433 micrographs were collected and a small dataset was manually picked to obtain initial 2D class averages followed by autopicking in RELION. Four rounds of 2D classification led to a homogeneous dataset. The density for UV-DDB was removed from the model shown in g. The resulting map was low-pass-filtered to 60 Å and used as initial model for the first round of 3D refinement leading to a map at 4.1 Å resolution after polishing. Given the accumulated dose of 40 e\(^−\) Å\(^−2\) spanning 40 frames, frames 1–28 were included during movie refinement and particle polishing in RELION. To improve the resolution, we performed 3D classification into three classes. Refinement with a mask of class II led to a 3.6 Å resolution map. Per particle CTF refinement improved the map to 3.5 Å resolution. k, l, Local-resolution-filtered map for NCP\(^{THF2}(−3)\)-UV-DDB (k) and NCP\(^{THF2}(−3)\) (l). m, Gold-standard FSC curves for NCP\(^{THF2}(−3)\)-UV-DDB (blue) and NCP\(^{THF2}(−3)\) (orange). n, o, Angular distribution for NCP\(^{THF2}(−3)\)-UV-DDB (n) and NCP\(^{THF2}(−3)\) (o). p, Overlay of the predicted NCP\(^{THF2}(−3)\)-UV-DDB model (red) with its cryo-EM structure (yellow), the difference between the two is reconciled by nucleosomal register shifting.
Extended Data Fig. 6 | See next page for caption.
Extended Data Fig. 6 | Thermodynamic dissection of UV-DDB binding with slide-assisted site-exposure. a, DNase I digestion of undamaged (UD) nucleosomes with a range of enzyme concentrations (0.12 U to 0.015 U) show identical sensitive sites (i), data shown are the average of two technical replicates per enzyme concentration and are highly reproducible across replicates (ii). Note that an example of the data reproducibility is shown for the undamaged construct, but the data for all constructs were highly reproducible (Pearson correlations of $R > 0.95$). b, Ratio of counts at position $n$ versus $n-3$ for undamaged NCP, NCP$^{THF2}(-1)$, NCP$^{THF2}(-3)$ in the absence and presence of UV-DDB. c, (i) Peaks for UD NCP, NCP$^{THF2}(-1)$, NCP$^{THF2}(-3)$ and NCP$^{THF2}(+1)$ in the presence of UV-DDB used to quantify the register-shift ratios ($n = 2$). Note that in the case of NCP$^{THF2}(+1)$ we expect only 30–40% of molecules to shift by a maximum of 1 bp. Our data are consistent with an increased population of $\alpha$-satellite DNA in the $+1$ nucleosome shifted by 1 bp towards the dyad axis compared to the $-1$ nucleosome (iii), $n = 4$, mean ± s.d.; however, the width and the overlap of the shifted or unshifted peak prevented further detailed analysis. d, Titration of a 10 nM solution of a Cy5–15-bp-DNA$^{THF}$ with increasing concentrations of UV-DDB. The resulting curve was fitted with DynaFit resulting in a $K_d$ of around 0.15 nM. Given the tight binding indicated by a low $K_d$ relative to the high concentration of the Cy5–15-bp–DNA$^{THF}$ label, the value should be viewed as an estimate. These are, however, in line with the results of previous experiments. All data are shown as mean ± s.d. of three technical replicates. e, The complex of 10 nM Cy5–15-bp–DNA$^{THF}$ and UV-DDB was back-titrated with NCP$^{THF2}(-3)$ and the data fitted numerically in DynaFit. All data are shown as mean ± s.d. of three technical replicates. f, Thermodynamic binding scheme invoking induced-fit binding, in which the register shift is induced on UV-DDB binding, and a conformational preselection branch (selected-fit), in which a pre-equilibrium exists that is shifted by UV-DDB binding. g, Equations describing the thermodynamic binding process, and approximations used to derive $K_1$, $K_2$, $K_3$ and $K_4$. 
Extended Data Fig. 7 | See next page for caption.
Extended Data Fig. 7 | Classification and refinement procedures for NCP<sub>THF2</sub>(+1)–UV-DDB. a, Representative VPP cryo-EM micrograph (left) and reference-free 2D class averages (right) obtained with <i>cis</i>TEM<sup>48</sup> for NCP<sub>THF2</sub>(+1)–UV-DDB. b, Six ab initio models were generated with <i>cis</i>TEM. c, Five datasets were collected with VPP under identical imaging conditions obtaining a total of 11,675 micrographs. All dose-fractionated micrograph stacks were subjected to beam-induced motion correction with MotionCor2<sup>25</sup>. Further processing was carried out using dose-weighted MotionCor2-corrected sums (including all frames, 1 e<sup>−</sup> Å<sup>−2</sup> per frame). A small dataset was manually selected to obtain 2D class averages and used for autopicking in RELION. Class III shown in b was low-pass-filtered to 60 Å and used as an initial model for the first round of 3D classification for all five sessions independently. Several rounds of 2D and 3D classification were necessary to obtain homogeneous datasets. The best particles from each session were pooled and subjected to 3D classification into four models, revealing two different conformations for the complex. Thus, additional rounds of 3D classification were necessary to obtain a homogeneous subset of particles. Particles included in class I, class II and class III were merged and divided into four 3D classes. Owing to their structural similarity, particles in class V, class VI and class VII were pooled and divided into 3D classes. A homogeneous dataset was obtained after merging particles in class IX and class X, leading to a 4.2 Å resolution map (class A) with a mask that excluded DDB1. Particles in class IV and class VIII were pooled and subjected to 3D classification. The best particles led to a 4.6 Å resolution map after masking out DDB1 (class B). d, Gold-standard FSC curves for class A (blue) and class B (red) respectively. e, Angular distribution for the particles included in class A (e) and class B (f). g, h, Local-resolution-filtered maps for class A (g) and class B (h) (MonoRes<sup>25</sup>). i, Magnified view showing the interactions between DDB2 (green) and different cryo-EM models, demonstrating that essentially identical interactions between the protein and the DNA are maintained despite different type of lesions and rotational settings. j, Superimposition of atomic models for the repositioned NCP<sub>THF2</sub>(+1)–UV-DDB (class B) and the repositioned NCP<sub>THF2</sub>(−3)–UV-DDB (light grey, Extended Data Fig. 5h, k), illustrating repositioning to a common position.
Extended Data Fig. 8 | See next page for caption.
Extended Data Fig. 8 | Placing the damage on the opposite strand exposes UV-DDB to a different binding environment. a, DNA sequences and position of all complex structures with different lesions as shown in Fig. 4a. The relative translational register shift is indicated. b, DNA sequence for NCPTF2(−1)–UV-DDB with THF2 placed at position −22/−23 bp from the dyad axis on the reverse strand (3′ to 5′). c, Comparison of clash scores between SHL−2.5 (Watson, grey) and SHL+3.5 (Crick, orange) shown in Fig. 5c. d, Estimation of EC_{50} from the counter-titration experiments shown in Fig. 5a, c, d. All data are shown as mean ± s.d. of three technical replicates. e, Accessible contact surface area for nucleosomal DNA calculated as indicated across a range of probe radii. Most of the DNA surface (more than 93%) is inaccessible to probes with radii exceeding 4 Å. f, A mechanism invoking UV-DDB-induced nucleosome register-shifting is also conceivable for damage recognition of pyrimidine dimers in multiple nucleosomes. Modelling of available di- and tetranucleosome structures suggests that UV-DDB can access a substantial fraction of the outward-facing minor grooves in these more heterochromatic model substrates56, with similar predicted accessibility to corresponding positions in mononucleosomes (Fig. 5e).
## Extended Data Table 1 | NCP–UV-DDB complex cryo-EM data collection, refinement and validation statistics

|                      | NCP<sup>6-PP</sup>-UV-DDB (EMD-4762) | NCP<sup>THF2(-1)</sup>-UV-DDB (EMD-4763) | NCP<sup>THF2(-3)</sup>-UV-DDB (EMD-4765) | NCP<sup>THF2(+1)</sup>-UV-DDB class A (EMD-4764) | NCP<sup>THF2(+1)</sup>-UV-DDB class B (EMD-4766) |
|----------------------|--------------------------------------|----------------------------------------|---------------------------------------------|-----------------------------------------------|-----------------------------------------------|
| **Data collection and processing** |                                      |                                        |                                             |                                               |                                               |
| Magnification        | 130,000                              | 105,000                                | 130,000                                     | 130,000                                       | 130,000                                       |
| Voltage (kV)         | 300                                  | 300                                    | 300                                         | 300                                           | 300                                           |
| Electron exposure (e-/Å²) | 40                                 | 60                                     | 40                                          | 40                                            | 40                                            |
| Defocus range (µm)   | -0.4 – -0.7                          | -1.0 – -3.0                            | -0.4 – -0.7                                 | -0.4 – -0.7                                   | -0.4 – -0.7                                   |
| Pixel size (Å)       | 0.86                                 | 0.55 (super resolution)               | 1.72 (data 2 × binned)                     | 0.86                                          | 0.86                                          |
| **Symmetry imposed** | C1                                   | C1                                     | C1                                          | C1                                            | C1                                            |
| Initial particle images (no.) | 1,355,846                      | 511,505                                | 719,698                                     | 1,078,133                                     | 1,078,133                                     |
| Final particle images (no.) | 84,000                             | 129,986                                | 119,309                                     | 128,763                                       | 48,925                                        |
| Map resolution (Å)   | 4.3                                  | 3.9                                    | 4.2                                         | 4.5                                           | 4.8                                           |
| FSC threshold        | 0.143                                | 0.143                                  | 0.143                                       | 0.143                                         | 0.143                                         |
| Map resolution range (Å) | 3.0–12                        | 3.0–12                                  | 4.0–7                                       | 4.0–12                                        | 4.0–12                                        |
| **Refinement**       |                                      |                                        |                                             |                                               |                                               |
| Initial models used (PDB codes) | 4ZUX, 5Y0C, 4E54, 3E14           | 4ZUX, 5Y0C, 4E54, 3E14                  | 4ZUX, 5Y0C, 4E54, 3E14                   | 4ZUX, 5Y0C, 4E54, 3E14                  | 4ZUX, 5Y0C, 4E54, 3E14                   |
| Model resolution (Å) | 3.9                                  | 3.8                                    | 4.4                                         | 4.2                                           | 4.5                                           |
| FSC threshold        | 0.143                                | 0.143                                  | 0.143                                       | 0.143                                         | 0.143                                         |
| Map sharpening B factor (Å<sup>2</sup>) | -120                             | -233                                   | -180                                        | -160                                          | -200                                          |
| Model composition    |                                      |                                        |                                             |                                               |                                               |
| Non-hydrogen atoms   | 21,717                               | 21,617                                 | 21,477                                      | 21,268                                        | 21,500                                        |
| Protein residues     | 1996                                 | 1984                                   | 1969                                        | 1944                                          | 1974                                          |
| Nucleotides          | 288                                  | 288                                    | 288                                         | 288                                           | 288                                           |
| B factors (Å<sup>2</sup>) |                                      |                                        |                                             |                                               |                                               |
| Protein             | 85                                   | 118                                    | 257                                         | 156                                           | 261                                           |
| DNA                  | 73                                   | 90                                     | 264                                         | 155                                           | 244                                           |
| DNA damage           | 64                                   | 74                                     | 229                                         | 143                                           | 249                                           |
| R.m.s. deviations    |                                      |                                        |                                             |                                               |                                               |
| Bond lengths (Å)     | 0.012                                | 0.009                                  | 0.005                                       | 0.005                                         | 0.007                                         |
| Bond angles (°)      | 1.301                                | 0.932                                  | 0.944                                       | 0.873                                         | 1.107                                         |
| Validation           |                                      |                                        |                                             |                                               |                                               |
| MolProbity score     | 1.44                                 | 1.48                                   | 1.43                                        | 1.41                                          | 1.58                                          |
| Clashscore           | 2.71                                 | 1.98                                   | 2.15                                        | 2.00                                          | 4.28                                          |
| Poor rotamers (%)    | 0.47                                 | 0.18                                   | 0.65                                        | 0.24                                          | 0.36                                          |
| Ramachandran plot    |                                      |                                        |                                             |                                               |                                               |
| Favored (%)          | 94.0                                 | 95.0                                   | 95.0                                        | 96.6                                          | 94.7                                          |
| Allowed (%)          | 5.7                                  | 4.6                                    | 4.6                                         | 3.1                                           | 5.0                                           |
| Disallowed (%)       | 0.3                                  | 0.4                                    | 0.5                                         | 0.3                                           | 0.4                                           |
| **Model-to-data fit** |                                      |                                        |                                             |                                               |                                               |
| CCmask               | 0.76                                 | 0.75                                   | 0.52                                        | 0.72                                          | 0.70                                          |
| CCbox                | 0.78                                 | 0.79                                   | 0.64                                        | 0.76                                          | 0.83                                          |
| CCpeaks              | 0.68                                 | 0.71                                   | 0.47                                        | 0.61                                          | 0.60                                          |
| CCvolume             | 0.74                                 | 0.74                                   | 0.53                                        | 0.71                                          | 0.71                                          |

*Correlation coefficients were calculated using maps filtered based on local-resolution estimation with MonoRes<sup>47</sup>, sharpened with phenix.auto_sharpen<sup>46</sup>, and did not take into account the poorly resolved DDB1 B-domain (DDB1-8PB) portion of the map.
Extended Data Table 2 | NCP cryo-EM data collection, refinement and validation statistics

| Data collection and processing | NCP<sup>4PP</sup> (EMD-4767) | NCP<sup>THF2(-3)</sup> (EMD-4768) |
|-------------------------------|-------------------------------|-------------------------------|
| Magnification                 | 130,000                       | 130,000                       |
| Voltage (kV)                  | 300                           | 300                           |
| Electron exposure (e-/Å²)     | 40                            | 40                            |
| Defocus range (μm)            | -0.4 to -0.7                  | -1.0 to -3.0                  |
| Pixel size (Å)                | 0.86                          | 0.86                          |
| Symmetry imposed              | C1                            | C1                            |
| Initial particle images (no.) | 827,045                       | 266,180                       |
| Final particle images (no.)   | 98,387                        | 78,672                        |
| Map resolution (Å)            | 4.0                           | 3.5                           |
| FSC threshold                 | 0.143                         | 0.143                         |
| Map resolution range (Å)      | 3.1–12                        | 3.0–12                        |

| Refinement                   | NCP<sup>4PP</sup> (PDB 6R93) | NCP<sup>THF2(-3)</sup> (PDB 6R94) |
|-------------------------------|-------------------------------|-------------------------------|
| Initial models used (PDB codes) | 4ZUX, 5Y0C                  | 4ZUX, 5Y0C                  |
| Model resolution (Å)          | 3.8                           | 3.4                           |
| FSC threshold                 | 0.143                         | 0.143                         |
| Map sharpening B factor (Å²)  | -110                          | -140                          |
| Model composition             |                               |                               |
| Non-hydrogen atoms            | 12,844                        | 12,372                        |
| Protein residues              | 775                           | 803                           |
| Nucleotides                   | 290                           | 290                           |
| B factors (Å²)                |                               |                               |
| Protein                       | 63                            | 68                            |
| DNA                           | 99                            | 111                           |
| DNA damage                    | 103                           | 118                           |
| R.m.s. deviations             |                               |                               |
| Bond lengths (Å)              | 0.008                         | 0.008                         |
| Bond angles (°)               | 1.238                         | 1.012                         |
| Validation                    |                               |                               |
| MolProbity score              | 1.15                          | 1.15                          |
| Clashscore                    | 2.04                          | 1.07                          |
| Poor rotamers (%)             | 0.00                          | 0.00                          |
| Ramachandran plot             |                               |                               |
| Favored (%)                   | 97.0                          | 97                            |
| Allowed (%)                   | 2.5                           | 3.0                           |
| Disallowed (%)                | 0.5                           | 0.4                           |

| Model-to-data fit*            | NCP<sup>4PP</sup> (PDB 6R93) | NCP<sup>THF2(-3)</sup> (PDB 6R94) |
|-------------------------------|-------------------------------|-------------------------------|
| CCmask                        | 0.67                          | 0.69                          |
| CCBox                         | 0.73                          | 0.74                          |
| CCpeaks                       | 0.62                          | 0.66                          |
| CCvolume                      | 0.66                          | 0.70                          |

*Correlation coefficients were calculated using the maps filtered based on local-resolution estimation with MonoRes<sup>+</sup> and sharpened with phenix.auto_sharpen<sup>+</sup>.
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- n/a
- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. $F$, $t$, $r$) with confidence intervals, effect sizes, degrees of freedom and P value noted
  Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen’s $d$, Pearson’s $r$), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection
- EPU1.x (Thermo Fischer), PHERAtstar (BMG Labtech)

Data analysis
- CryoFLARE (www.cryoflare.org), RELION2, RELION3, Cctf1.06, Gautomatch0.3, Urblur1.0.2, XMIPP1.2, cisTEM1.0.0-beta, EMAN2.2, Pymol2.0.3, PHENIX dev_3318, REFMACS.0238, Typhoon Image Analyzer, Dynafit4 (BioKin Ltd.), Prism8 (GraphPad), CCP4, scanDamage.py (https://github.com/bunkrich27/pymol-scripts/blob/master/scanDamage.py), MotionCor2 1.05, QuasR 3.8, NMF package

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Cryo-EM maps and models are deposited, in the PDB (and EMDB) under accession codes are 6R8Y (4762), 6R8Z (4763), 6R91 (4765), 6R90 (4764), 6R92 (4766), 6R93 (4767) and 6R94 (4768).
**Life sciences study design**

All studies must disclose on these points even when the disclosure is negative.

- **Sample size**: No sample size calculations were performed. We chose a n=3 size for the biochemical fluorescence binding experiments in question. This is based on previous tests regarding reproducibility in the laboratory.

- **Data exclusions**: No data where excluded from the analysis in the biochemical binding experiments.

- **Replication**: All attempts to replicate the biochemical data were successful.

- **Randomization**: Not relevant to the biochemical experiments reported.

- **Blinding**: No blinding was used in the biochemical studies, given the small sample size.

**Reporting for specific materials, systems and methods**

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

| n/a | Involved in the study |
|-----|------------------------|
| ☒   | Antibodies             |
| ☒   | Eukaryotic cell lines  |
| ☒   | Palaeontology          |
| ☒   | Animals and other organisms |
| ✘   | Human research participants |
| ✘   | Clinical data          |

### Methods

| n/a | Involved in the study |
|-----|------------------------|
| ☒   | ChIP-seq               |
| ☒   | Flow cytometry         |
| ✘   | MRI-based neuroimaging |

**Eukaryotic cell lines**

Policy information about [cell lines](#)

- **Cell line source(s)**: SF9 and High Five cells were used in this study.
- **Authentication**: None of cell lines were authenticated.
- **Mycoplasma contamination**: The cell lines were not tested for Mycoplasma contamination.
- **Commonly misidentified lines** (See [ICLAC register](#)): None of commonly misidentified lines were used in this study.