

Rotational symmetry of the structured Chip/LDB-SSDP core module of the Wnt enhanceosome

Miha Renko^{a,1}, Marc Fiedler^{a,1}, Trevor J. Rutherford^a, Jonas V. Schaefer^b, Andreas Plückthun^b, and Mariann Bienz^{a,2}

^aMedical Research Council Laboratory of Molecular Biology, CB2 0QH Cambridge, United Kingdom; and ^bDepartment of Biochemistry, University of Zurich, 8057 Zurich, Switzerland

Edited by Roeland Nusse, Stanford University School of Medicine, Stanford, CA, and approved September 9, 2019 (received for review July 23, 2019)

The Chip/LIM-domain binding protein (LDB)–single-stranded DNA-binding protein (SSDP) (ChiLS) complex controls numerous cell-fate decisions in animal cells, by mediating transcription of developmental control genes via remote enhancers. ChiLS is recruited to these enhancers by lineage-specific LIM-domain proteins that bind to its Chip/LDB subunit. ChiLS recently emerged as the core module of the Wnt enhanceosome, a multiprotein complex that primes developmental control genes for timely Wnt responses. ChiLS binds to NPFXD motifs within Pygopus (Pygo) and the Osa/ARID1A subunit of the BAF chromatin remodeling complex, which could synergize with LIM proteins in tethering ChiLS to enhancers. Chip/LDB and SSDP both contain N-terminal dimerization domains that constitute the bulk of their structured cores. Here, we report the crystal structures of these dimerization domains, in part aided by DARPin chaperones. We conducted systematic surface scanning by structure-designed mutations, followed by *in vitro* and *in vivo* binding assays, to determine conserved surface residues required for binding between Chip/LDB, SSDP, and Pygo-NPFXD. Based on this, and on the 4:2 (SSDP-Chip/LDB) stoichiometry of ChiLS, we derive a highly constrained structural model for this complex, which adopts a rotationally symmetrical SSDP₂-LDB₂-SSDP₂ architecture. Integrity of ChiLS is essential for Pygo binding, and our mutational analysis places the NPFXD pockets on either side of the Chip/LDB dimer, each flanked by an SSDP dimer. The symmetry and multivalency of ChiLS underpin its function as an enhancer module integrating Wnt signals with lineage-specific factors to operate context-dependent transcriptional switches that are pivotal for normal development and cancer.

Wnt enhanceosome | Chip/LDB1-SSDP | Pygo

Vertebrate LIM-domain binding protein 1 (LDB1, also known as NLI or CLIM) and its *Drosophila* ortholog Chip have pleiotropic functions in controlling embryonic and larval development at multiple stages (1–4), maintenance of intestinal and hematopoietic stem cells (5, 6), and differentiation along erythroid cell lineages (7, 8). Chip/LDB proteins are recruited to remote transcriptional enhancers of pivotal developmental control genes by their C-terminal LIM-interacting domain (LID) that binds directly to LIM-homeodomain DNA-binding proteins, or to GATA and basic-loop-helix (bHLH) DNA-binding proteins via LIM-only adaptors. They facilitate communication between remote enhancers and proximal promoters (2, 9), likely via looping out intervening sequences (8). These long-range enhancer–promoter interactions critically depend on, and are mediated by, self-interaction of LDB proteins (10, 11) through their N-terminal dimerization domain (DD) (12, 13).

Chip/LDB binds to a single-stranded DNA binding protein (SSDP, also known as single-stranded binding protein, SSBP) through the LDB1/Chip conserved domain (LCCD), a short region downstream of the DD (14, 15). The loss-of-function phenotypes of *Drosophila chip* mutants closely resemble those of *ssdp* mutants in various developmental contexts, although the former tend to be stronger and more pleiotropic than the latter (14, 16), especially in the early embryo (17). These phenotypic similarities indicate an intimate cooperation between Chip and SSDP in

mediating enhancer function. Consistent with this, chicken Ssdp1 and Ssdp2 are required to confer transcriptionally active chromatin on loci controlled by Ldb1 (15), and murine Ssbp2 is required for hematopoietic stem cell maintenance, similarly to Ldb1 (18). The molecular basis underlying this intimate cooperation between Chip/LDB and SSDP is not known.

We recently discovered that Chip/LDB and SSDP form a stable complex called ChiLS (Chip/LDB-SSDP), which binds to NPFXD motifs within the nuclear Wnt signaling factor Pygopus (Pygo) and Osa/ARID1A (17), a chromatin-binding subunit of the BAF transcriptional coactivator complex (19). The Wnt enhanceosome is a ChiLS-containing multiprotein complex that is tethered to Wnt-responsive transcriptional enhancers via T cell factors/lymphoid-enhancer binding factors (TCF/LEF) and their associated Groucho/transducin-like enhancer (TLE) corepressors, to prime linked developmental control genes for timely Wnt responses (17, 20). The Wnt response of this complex is conferred by Pygo, which facilitates loading of the Wnt effector β -catenin via an adaptor called BCL9, thereby promoting transcriptional activation (ON state) (17, 20), while its repression (OFF state) appears to depend on Osa (21). The Wnt enhanceosome model envisages a pivotal role of ChiLS in the assembly and function of the complex, implicating ChiLS as its switch module: Its chromatin tethering involves a combination of enhancer-associated

Significance

A structural model is presented for the core Chip/LIM-domain binding protein (LDB)–single-stranded DNA-binding protein (SSDP) (ChiLS) complex, which mediates long-range enhancer–promoter interactions to control the transcription of master regulatory genes with key functions in embryonic development, stem cell maintenance, and differentiation along adult cell lineages, including erythroid maturation. Because of its rotational symmetry and multivalency, ChiLS can bind to multiple combinations of lineage-specific DNA-binding proteins, as well as cofactors responding to extracellular signals such as Wnt, and is thus uniquely poised to integrate these inputs and translate them into transcriptional ON and OFF states of target genes.

Author contributions: M.R., M.F., and M.B. designed research; M.R., M.F., and T.J.R. performed research; J.V.S. and A.P. contributed new reagents/analytic tools; M.R., M.F., T.J.R., and M.B. analyzed data; and M.B. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution License 4.0 (CC BY).

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 6S9R for SSDP, 6D9T for DD-DARPin3, and 6S9S for DD-DARPin10).

¹M.R. and M.F. contributed equally to this work.

²To whom correspondence may be addressed. Email: mb2@mrc-lmb.cam.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1912705116/-DCSupplemental.

First published September 30, 2019.

proteins (including LIM proteins, Pygo, TCF/LEF-associated Groucho/TLE, and the BAF complex) that determine jointly the ON and OFF states of Wnt-controlled downstream genes (17, 20). Importantly, the stoichiometry of the minimal stable ChiLS complex is 2:4 (Chip/LDB:SSDP), as determined by size-exclusion chromatography followed by multiangle light-scattering (SEC-MALS) of purified recombinant proteins (17), implying that ChiLS contains 1 Chip/LDB dimer and 2 SSDP dimers.

Here, we report the crystal structures of the DD of SSDP and XLdb1, in the latter case with the help of DARPIn (designed ankyrin repeat proteins) chaperones that mask a hydrophobic surface patch of the DD required for its SSDP binding. Systematic structure-led mutagenesis of conserved solvent-exposed residues of Chip/LDB and SSDP followed by in vitro and in vivo binding assays enabled us to generate an interaction map, and to derive a highly constrained structural model for ChiLS. Its rotationally symmetric SSDP₂-LDB₂-SSDP₂ architecture with 2 predicted NPFxD-binding pockets on either side underscores its function as an integrating core module of the Wnt enhancerosome and other ChiLS-containing enhancer-binding complexes.

Results

SSDP contains an N-terminal LisH domain known to form obligate dimers in other proteins (22), followed by a highly conserved extension and a long nonconserved C-terminal tail likely to be disordered (Fig. 1A). The extended N-terminal dimerization domain (called SSDP-N below; also known as the LUFs domain) (14) can form stable tetramers in vitro following bacterial expression (17). To determine the structure of SSDP-N, we optimized its boundaries and succeeded in obtaining crystals for SSDP₁₋₈₆ that diffracted to 2.4 Å. As we were unable to use molecular replacement with the LisH domain as a search model, we generated 5 different selenomethionine (SeMet)-labeled mutants,

1 of which yielded diffracting crystals that enabled us to determine the structure of SSDP-N (SI Appendix, Table S1).

Structure of the SSDP-N Dimer. The asymmetric crystallographic unit contains a classic LisH dimer-fold in which 2 α-helices (α1, α2) (Fig. 1B) associate to form an antiparallel 4-helix bundle (22), the hydrophobic interface of which is strengthened by additional hydrogen bonds in its periphery (Fig. 1C). A third helix (α3) interacts with its counterpart (α3') from the opposite dimer to form an X-shaped structure, tucked under the 4-helix bundle. It thus resembles TBL1, a subunit of a transcriptional corepressor complex and the only other LisH-containing protein known to form tetramers (23), except that the lengths and angles between α3 and α3' differ in the 2 structures (SI Appendix, Fig. S1). As in other LisH dimers, the dimer interface is extensive (burying ~2,500 Å²), which explains why SSDP forms obligate dimers, and why SSDP monomers are not detectable (17) (see also below). Indeed, the K_d for LisH-mediated self-association of Lis1 was reported to be subfemtomolar (22).

Within the crystal lattice, symmetry-related dimers associate via 2 perpendicular pairs of antiparallel α3 helices. The resulting interlocked bundle of 4 α-helices bury ~1,400 Å² (Fig. 2A), and is held together entirely via hydrophobic contacts, with valine 65 (V65) occupying a central position in the tetramer interface ("α3 interface," below). The same tetramer was observed for human SSBP2 (whose structured region differs from fly SSDP-N only by 2 semiconserved residues) (Fig. 1B) (24) whose structure was reported during the final stages of our work. Intriguingly, some of the hydrophobic residues in the loop between α2 and α3 also mediate association between symmetry-related dimers in the crystal, resulting in alternative tetramer configurations (Fig. 2B). Although these loop-mediated interactions bury smaller interfaces (602 Å² and 624 Å², respectively; "loop interface," below), they proved to be functionally relevant in vivo (see below). We

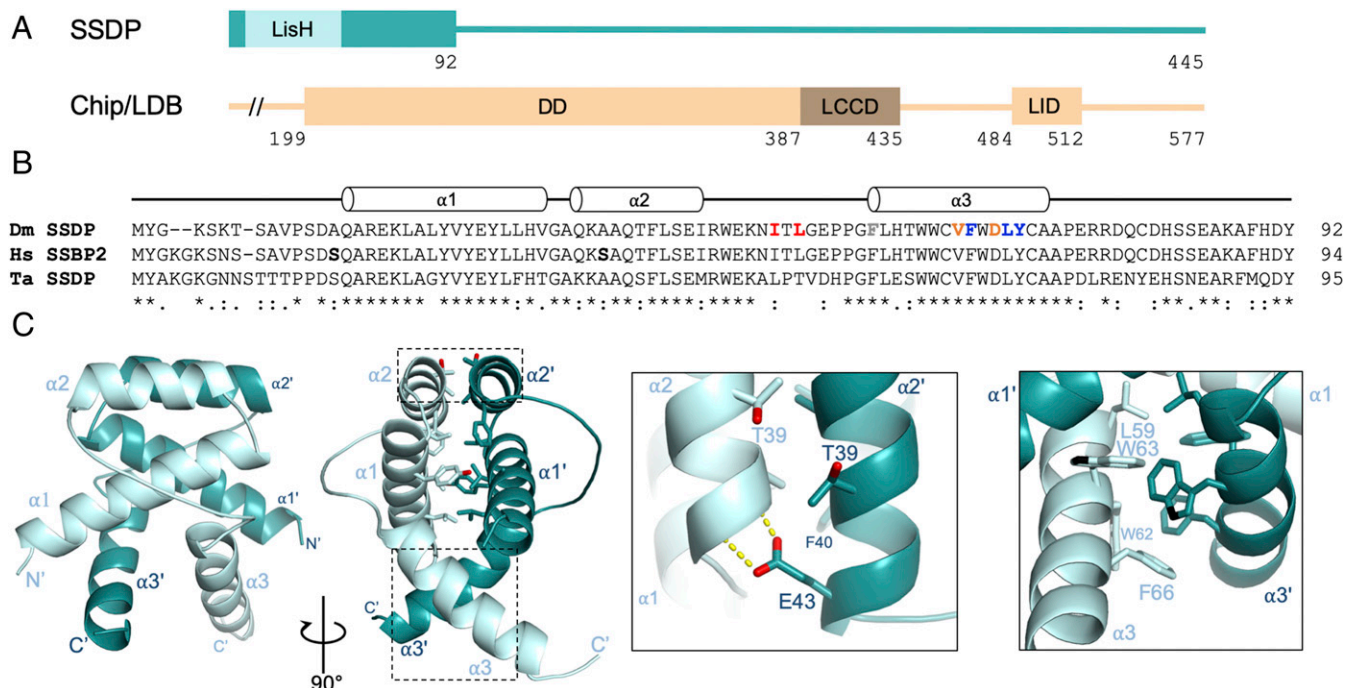


Fig. 1. Structure of the SSDP dimer. (A) Cartoons of SSDP and Chip/LDB and their domains (numbers, *Drosophila* SSDP and Chip). (B) Sequence alignment of the N-terminal dimerization domain of *Drosophila* (Dm), human (Hs), and *Trichoplax adhaerens* (Ta) SSDP, with secondary elements indicated above; bold indicates amino acid variations between Hs SSBP2 and Dm SSDP (structured region only); colored indicates crucial mutated residues. (C) Ribbon representations of SSDP dimer, with helices labeled; (Left Inset) dimerization interface between α2 and α2', with key residues shown in stick, and hydrogen bonds (between α2' E43 and main chain α2 atoms) as yellow dashed lines; (Right Inset) α3/α3' dimerization interface, with key hydrophobic residues shown in stick.

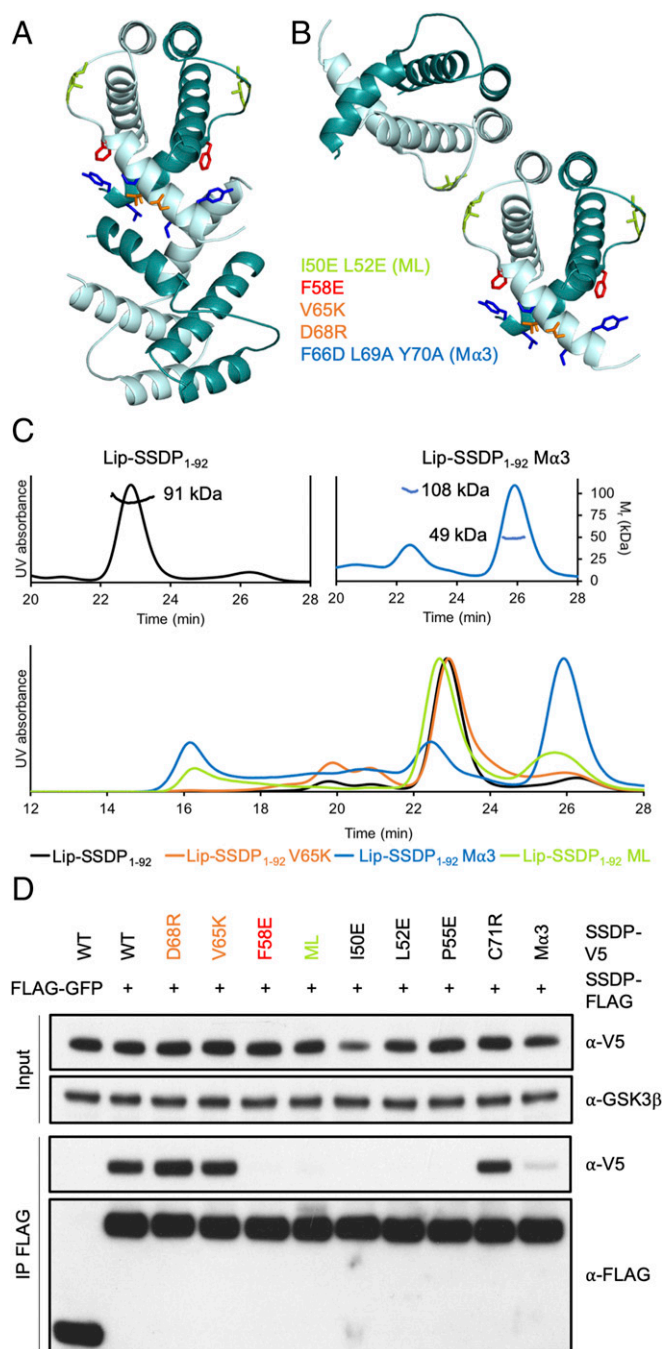


Fig. 2. Two modes of SSDP tetramerization. (A and B) Tetramerization of SSDP through (A) $\alpha 3$ interface or (B) loop interface (between $\alpha 2$ and $\alpha 3$), with key mutated residues in stick (same color in all panels). (C) SEC-MALS profiles of WT and selected Lip-SSDP₁₋₉₂ mutants, as indicated; numbers in Upper, M_r determined by MALS (corresponding to SSDP₂ and SSDP₄; expected M_r , 47 and 94 kDa, respectively). (D) co-IP assays of selected SSDP mutants in transfected LDB1/2 DKO cells (see also *SI Appendix*, Figs. S2 and S3).

note that TBL1 tetramerizes via $\alpha 2$ (*SI Appendix*, Fig. S1) via residues that are not conserved in SSDP.

Alternative Modes of Self-Interaction of SSDP Dimers. To test the functional relevance of the different modes of SSDP self-interactions, we designed repelling amino acid substitutions in solvent-exposed hydrophobic residues in the 2 alternative interfaces of SSDP-N, mutating a centrally located valine in the $\alpha 3$ interface (V65K), and

2 hydrophobic residues in the loop interface (I50E L52E, called ML) that mediate mutual interactions in the crystals (Fig. 1B). We also included a previously designed triple mutation of hydrophobic residues flanking V65 (F66E L69A Y70A, called M α 3) that blocks SSBP2-N tetramerization (24). These mutations were introduced into Lip-SSDP₁₋₉₂ and their self-interaction was tested by SEC-MALS. As expected (24), M α 3 blocks tetramerization of SSDP-N, but V65K does not (Fig. 2C), likely because this single mutation is too weak to do so, given the extensive $\alpha 3$ interface. ML also reduces tetramerization of SSDP-N significantly (Fig. 2C). We conclude that both modes of SSDP-N tetramerization can occur, with $\alpha 3$ being the preferred interface. Tetramerization likely occurs during bacterial expression since SSDP-N dimers and tetramers are stable in solution for >8 h, and do not interconvert following purification.

Next, we designed another 16 substitutions in conserved solvent-exposed residues that should not perturb the fold of SSDP-N (*SI Appendix*, Fig. S24), and introduced these (plus V65K) into full-length V5-tagged SSDP. We then coexpressed these with WT FLAG-tagged SSDP in HEK293T cells from which we deleted LDB1 and LDB2 by CRISPR engineering (below, double-knockout [DKO] cells) (*SI Appendix*, Fig. S3), to test their mutual association by coimmunoprecipitation (co-IP) in the absence of endogenous LDB. As expected, SSDP-FLAG co-IPs efficiently with WT and most mutant SSDP-V5, except for I50E, L52E, P55E, and F58E (each altering the loop surface) (Fig. 2B and *SI Appendix*, Fig. S24), which do not coimmunoprecipitate at all. M α 3 produces a reduced co-IP signal, partly because this triple-mutant is somewhat unstable in cells (Fig. 2D and *SI Appendix*, Fig. S2B). We conclude that full-length SSDP dimers also self-associate in cells upon overexpression, but do so entirely through their loop interface. This implies that the SSDP C terminus blocks self-association of full-length SSDP via its $\alpha 3$ surface.

Structure of the LDB Dimer. Next, we purified fragments spanning the DD or DD-LCCD from human LDB1, linked to a cleavable His₆ and Lipoyl (Lip) tag, but could not crystallize these. Screening LDB orthologs from different species, we found that *Xenopus laevis* Ldb1 (XLdb1) (1) yielded the most stable protein after removal of solubility tags. XLdb1 is closely related to human LDB1 (*SI Appendix*, Fig. S4), but we did not succeed in obtaining diffracting crystals with this protein either, largely because the DD and DD-LCCD undergo nonspecific aggregation over time. We therefore resorted to chaperone-based approaches, selecting nanobodies (25) or high-affinity DARPins, genetically engineered antibody-mimetics based on consensus ankyrin repeat proteins and selected from diverse synthetic libraries (26), against the purified DD or ChiLS complex (*SI Appendix*, *Supplementary Methods*). None of 35 selected nanobodies facilitated crystallization, but we obtained DD crystals with 5 of 15 selected DARPins (*SI Appendix*, Fig. S5) that diffracted to 2.0 to 2.59 Å (*SI Appendix*, Table S1), depending on the DARPIn (*SI Appendix*, Fig. S64). None of these 5 DARPins bind to the DD-LCCD nor to the assembled ChiLS complex.

Next, we determined the crystal structures of the DD (XLdb1 20-200) bound to DARPins 2, -3, -5, -7, or -10 (*SI Appendix*, Fig. S6). Remarkably, each of the 5 DARPins (despite their different sequences) recognizes the same lateral surface patch of the DD, perpendicular to its dimerization interface and facing outwards (*SI Appendix*, Fig. S6B). This patch is highly hydrophobic (e.g., Y81, I83, and L87 engage in direct DARPIn contacts), which explains why its masking by DARPins was crucial to prevent non-specific aggregation of the DD via this patch (called “hydrophobic patch,” below).

The DD adopts a cone-shaped $\alpha + \beta$ barrel-fold, composed of a long, highly curved antiparallel β -sheet (formed by $\beta 1$ to $\beta 6$) whose cavity is filled by 3 α -helices ($\alpha 1$ to $\alpha 3$) that run roughly parallel to the β -sheet (Fig. 3A). This structure is capped by a

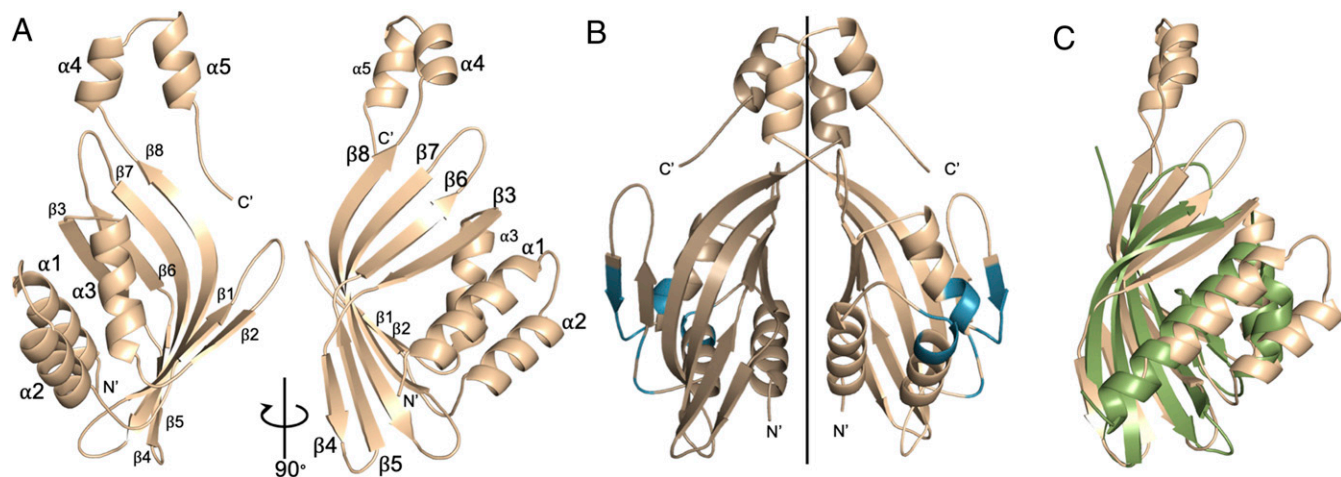


Fig. 3. Structure of the DD. (A and B) Ribbon representations of (A) the DD monomer (with α -helices labeled) and (B) the DD dimer; vertical line is the symmetry axis; blue is the DARPin-binding patch. (C) Superimposition of the DD (wheat) and NTF2 (green; 1JB5) (see also *SI Appendix, Fig. S7*).

short C-terminal extension at the apex of the DD, formed by 2 intertwined α -helices ($\alpha 4$ and $\alpha 5$) that contribute to the dimer interface (Fig. 3B). The whole dimer interface is extensive (burying $\sim 1,500 \text{ \AA}^2$), which explains why the DD dimer is highly stable in solution.

Curiously, the closest structural relative of the DD is scytalone dehydratase (27) (rmsd $\sim 2 \text{ \AA}$), a bacterial ketosteroid isomerase (*SI Appendix, Fig. S7A*). The DD also resembles the fold of other members of this enzyme family (28), which are among the most efficient enzymes known (29). The DD fold is also found in eukaryotes; for example, in nuclear transport factor-2 (NTF-2) (30) (rmsd $\sim 2.2 \text{ \AA}$) (Fig. 3C). Whether any of these structural similarities have functional significance is unclear. The oligomeric state of the DD folds is variable, ranging from monomeric to tetrameric, and their dimerization modes also vary, with some folds dimerizing via their curved β -sheets (*SI Appendix, Fig. S7B*). However, none of the known DD folds contain an intertwined apical extension, as seen in the XLdb1 DD (Fig. 3B and *SI Appendix, Fig. S7B*). This apex may help to determine the dimerization mode of the DD, and stabilize the Chip/LDB dimer, given that a region spanning this apex is essential for β -globin transcription during erythroid maturation (albeit not for LDB1 dimerization per se, as the DD core may be partially competent to dimerize by itself) (11).

Binding Between Chip/LDB and SSDP. Next, we attempted to crystallize the ChiLS complex, using bicistronic coexpression of SSDP₁₋₉₂ and DD-LCCD from XLdb1. LCCD is a conserved 49-amino acid stretch (Fig. 1A and *SI Appendix, Fig. S4*) required for binding to SSDP (14, 15). It spans 2 predicted α -helices ($\alpha 6$ and $\alpha 7$), whereby $\alpha 6$ is separated from the DD by a linker of 9 conserved residues (*SI Appendix, Fig. S4*). We designed 2 DD fragments with C-terminal extensions, namely DD-LCCD1 (20-244, spanning $\alpha 6$ and $\alpha 7$) or DD-LCCD2 (20-226, spanning only $\alpha 6$). Both form a stable complex with SSDP₁₋₉₂, implying that $\alpha 6$ suffices for SSDP binding at high protein concentrations during bacterial expression. We purified both complexes, but did not obtain diffracting crystals despite testing a wide range of conditions, as well as cocrystallization with nanobodies or DARPins.

We therefore took an alternative approach, namely systematic mutational analysis of conserved solvent-exposed residues of each ChiLS dimer subunit (Fig. 4A), to identify their surfaces required for mutual interaction. Since DARPin binding to their cognate patch of the DD (Fig. 3B) blocks its interaction with SSDP, we initially focused on this hydrophobic patch, designing a set of repelling mutations (*SI Appendix, Fig. S8*), which might

block binding between DD-LCCD1 and SSDP. Indeed, 2 of them do so (L87D R90D and Y81D L87D R90D) (*SI Appendix, Table S2*). Since Y81, L87, and R90 are direct DARPin-interacting

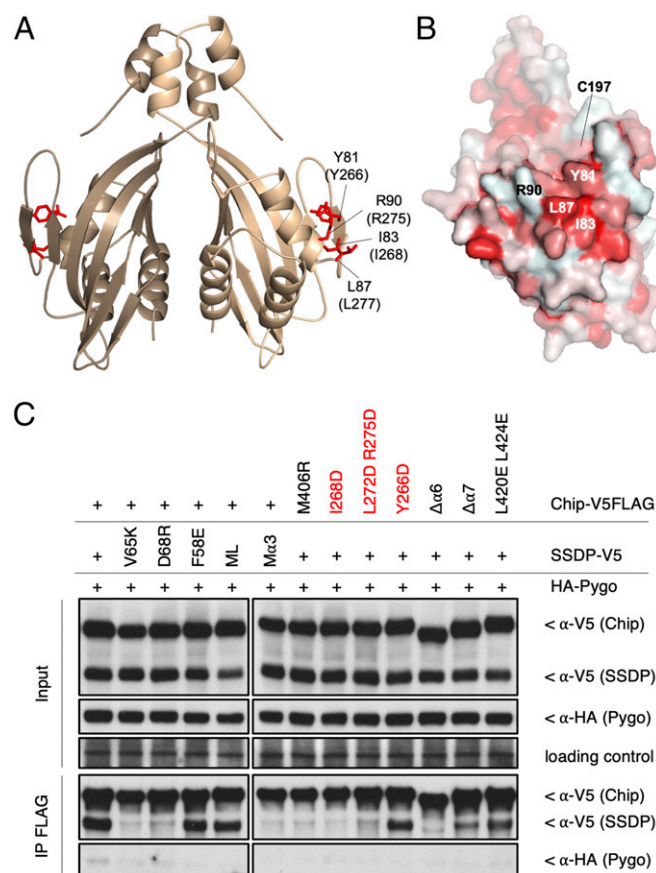


Fig. 4. Mutual interactions between Chip/LDB, SSDP, and Pygo. (A) Functionally relevant solvent-exposed residues in the DD (Chip residue numbers in parenthesis), tested for interaction with SSDP and Pygo. (B) Surface representation of DD, with hydrophobic patch mediating interaction with SSDP highlighted; Y81, L87, and R90, DARPin-binding residues; C197 marks the C terminus of the structured part of the DD, and start of LCCD. (C) co-IP assays of selected mutants in transfected HEK293T cells (see also *SI Appendix, Table S3*).

residues (Fig. 4B), this implicates the hydrophobic patch in the interaction between SSDP and the DD.

Next, we designed additional mutations in the apical, lateral, and basal surfaces of the DD dimer, in LCCD and its linker to DD (*SI Appendix*, Fig. S8), and we also generated internal deletions that remove LCCD $\alpha 6$ or $\alpha 7$ ($\Delta\alpha 6$, $\Delta\alpha 7$), for co-IP assays in transfected HEK293T cells aimed at testing the binding between full-length proteins. To minimize cross-reaction with endogenous LDB, we introduced these mutations into Chip-V5FLAG, and coexpressed them with SSDP-V5. This confirmed the functional importance of Y81, L87, and R90, but also identified additional residues near the hydrophobic patch as critical for co-IP with SSDP (Fig. 4C and *SI Appendix*, Table S3), marking a contiguous hydrophobic surface in the DD (Fig. 4B). In contrast, repelling point mutations in the apical or basal surface of the DD did not affect co-IP between Chip and SSDP (*SI Appendix*, Table S3). We conclude that a lateral hydrophobic surface patch of the DD is crucial for its association with SSDP. This implies that a single Chip/LDB dimer can accommodate 2 SSDP dimers.

In addition to this hydrophobic patch, we also found that $\alpha 6$ is crucial for co-IP between Chip and SSDP, while $\Delta\alpha 7$ also reduces it; indeed, mutation of a single methionine in $\alpha 6$ (M406R) reduces co-IP between the 2 proteins to background levels (Fig. 4C). M402E also blocks co-IP, and L405D reduces it, as does a double mutation in $\alpha 7$; however, none of the mutations in the conserved linker between the DD and LCCD affect co-IP (*SI Appendix*, Table S3). Thus, $\alpha 6$ is essential for binding between Chip and SSDP, and M402 and M406 may directly contact SSDP. Consistent with this, a 10-amino acid stretch spanning these methionines in Chip is essential for its binding to SSDP (14), as is a similar 6-amino acid stretch in chicken Ldb1, whereby the latter is also critical for motoneuron specification in the chicken embryo (15), indicating the physiological importance of LCCD $\alpha 6$. Either $\alpha 6$ constitutes the SSDP-binding surface of Chip/LDB, together with the lateral hydrophobic patch, or this patch binds primarily to $\alpha 6$, which in turn binds to SSDP, perhaps the more likely scenario, for reasons discussed below. In contrast, $\alpha 7$ appears to be auxiliary in mediating the Chip/LDB-SSDP interaction.

To define the Chip-binding surface of SSDP, we conducted co-IP assays between WT Chip-V5FLAG and our panel of 17 SSDP-V5 mutants (Fig. 2D and *SI Appendix*, Fig. S24), following coexpression in HEK293T cells. Most SSDP mutations have no effect (including ML), but D68R reduces co-IP, as does M $\alpha 3$, while V65K essentially blocks co-IP (Fig. 4C). We conclude that SSDP binds to Chip/LDB via its $\alpha 3$ surface (Fig. 24).

Dependence of Pygo Binding on ChiLS Complex Assembly. Next, we tested our SSDP and Chip mutants for interaction with *Drosophila* Pygo upon coexpression in HEK293T cells. This interaction is relatively weak, and not easily detectable by co-IP (17). However, we consistently obtained a co-IP signal between Pygo and WT Chip and SSDP, but with none of the mutants that block the Chip-SSDP interaction (Fig. 4C and *SI Appendix*, Table S3). This suggests that the integrity of the ChiLS complex is essential for Pygo binding, consistent with our previous evidence that peptides spanning NPFXD from Pygo or human Pygo2 bind to the assembled recombinant ChiLS complex but not to its subunits alone (17). Notably, F58E blocks Pygo binding to ChiLS while it barely affects binding between Chip and SSDP, which suggests that SSDP F58 contributes critically to the NPFXD-binding pocket of ChiLS.

To corroborate our co-IP results with in vitro binding assays between recombinant proteins, we used NMR spectroscopy (17), probing an ^{15}N -labeled 27 amino acid peptide spanning NPFXD from human Pygo2 (^{15}N -Lip-NPFXD_{Pygo2}) with purified recombinant ChiLS. This revealed that WT ChiLS interacts with ^{15}N -Lip-NPFXD_{Pygo2}, even without LCCD $\alpha 7$ (*SI Appendix*, Fig. S9), indicating that the latter is dispensable for ChiLS binding to NPFXD at high protein concentration. We also used this assay to determine

whether excess unlabeled Lip-NPFXD peptide from Pygo can compete for binding of purified ChiLS to an ^{15}N -labeled 26 amino acid peptide from Osa (^{15}N -Lip-NPFXD_{Osa}), which was the case (*SI Appendix*, Fig. S10). This implies that the same ChiLS pocket can accommodate the NPFXD motif of either ChiLS ligand.

A Highly Constrained Structural Model of ChiLS. Next, we constructed a model of the ChiLS complex, taking into account its constituent dimer structures (Figs. 1C and 3B), the results from our mutational analysis (Fig. 4 and *SI Appendix*, Tables S2 and S3) and the previously determined 4:2 (SSDP:Chip/LDB) stoichiometry of the ChiLS complex (17). The only ChiLS configuration that is consistent with all our data corresponds to an SSDP₂-LDB₂-SSDP₂ architecture with rotational symmetry (Fig. 5A). In this model, the lateral hydrophobic surface patches of each DD subunit (Fig. 5A, blue) interface with their downstream-adjacent LCCD $\alpha 6$ (Fig. 5A, orange) and the $\alpha 3$ surface of an SSDP dimer (Fig. 5A, dark cyan). Because of its inherent rotational symmetry, the complex can accommodate two NPFXD ligands (Fig. 5A, gray rods), through opposite pockets whose positions are defined by proximity to the LCCD $\alpha 6$ methionines (Fig. 5A, orange sticks) and SSDP F58 (Fig. 5A, red patch).

To corroborate our model, we asked whether a minimal LCCD fragment could bind SSDP, and if so, whether this would be sensitive to mutations in its $\alpha 3$ or loop surface. We therefore conducted SEC-MALS of Chip-LCCD (residues 384 to 436, of which 47 are identical in human LDB1) (*SI Appendix*, Fig. S4) tagged with maltose binding protein (MBP) coexpressed with WT, V65K, M $\alpha 3$, or ML Lip-SSDP₁₋₉₂. As a further control, we also included a triple mutant of LCCD (M402E L405D M406R, M $\alpha 6$) expected to block binding to SSDP (Fig. 4C). Indeed, the WT proteins form predominantly a single complex eluting in 1 main SEC peak, containing similar amounts of LCCD and SSDP as judged by PAGE of the corresponding fractions (Fig. 5B). Based on its molar mass ($M_r = 190$ kDa) as determined by MALS, the complex in this peak most likely corresponds to LCCD-SSDP₄-LCCD: That is, 2 LCCD-SSDP₂ complexes interacting via the SSDP loop interface (Fig. 5B, cartoon). We also observed a smaller peak ($M_r = 90$ kDa), likely corresponding to LCCD-SSDP₂, in addition to minor peaks with considerably higher molar masses, likely corresponding to unspecific aggregates (Fig. 5B, asterisks). As expected, none of the mutants show the 190-kDa complex: SSDP M $\alpha 3$ neither binds to LCCD nor tetramerizes, and so only forms dimers; SSDP V65K and LCCD M $\alpha 6$ cannot bind to their partner subunits, and so the 2 main peaks observed with these mutants correspond to LCCD and SSDP₄ (Fig. 5B, cartoons). Importantly, SSDP ML can interact with LCCD through $\alpha 3$, and so forms LCCD-SSDP₂ (possibly in addition to coeluting SSDP₄) (Fig. 5B). This confirms that 1) LCCD can bind to SSDP directly via $\alpha 6$, and 2) SSDP dimers can undergo mutual interactions via their loop surfaces if bound to LCCD. Thus, in the presence of LCCD, the $\alpha 3$ surface of SSDP prefers to bind to LCCD rather than itself.

Discussion

Our work has led to a highly constrained structural model of ChiLS, the core complex of the Wnt enhanceosome (17, 20). Its rotational symmetry implies that ChiLS contains 2 structurally identical NPFXD-binding pockets, each bordered by an SSDP dimer and LCCD $\alpha 6$ (and possibly DD residues) (Fig. 5A), and each binding either Pygo2 or Osa NPFXD (*SI Appendix*, Fig. S10). Thus, 1 single ChiLS core complex, via its 2 NPFXD pockets, could accommodate Pygo as well as the Osa/ARID1 subunit of the BAF complex simultaneously. This is consistent with the notion that Pygo and the BAF complex are constitutive components of the Wnt enhanceosome (20).

ChiLS is also found in other enhancer-binding complexes, most notably those binding to the remote LCR enhancer that controls β -globin genes during erythroid maturation (8, 18, 31),

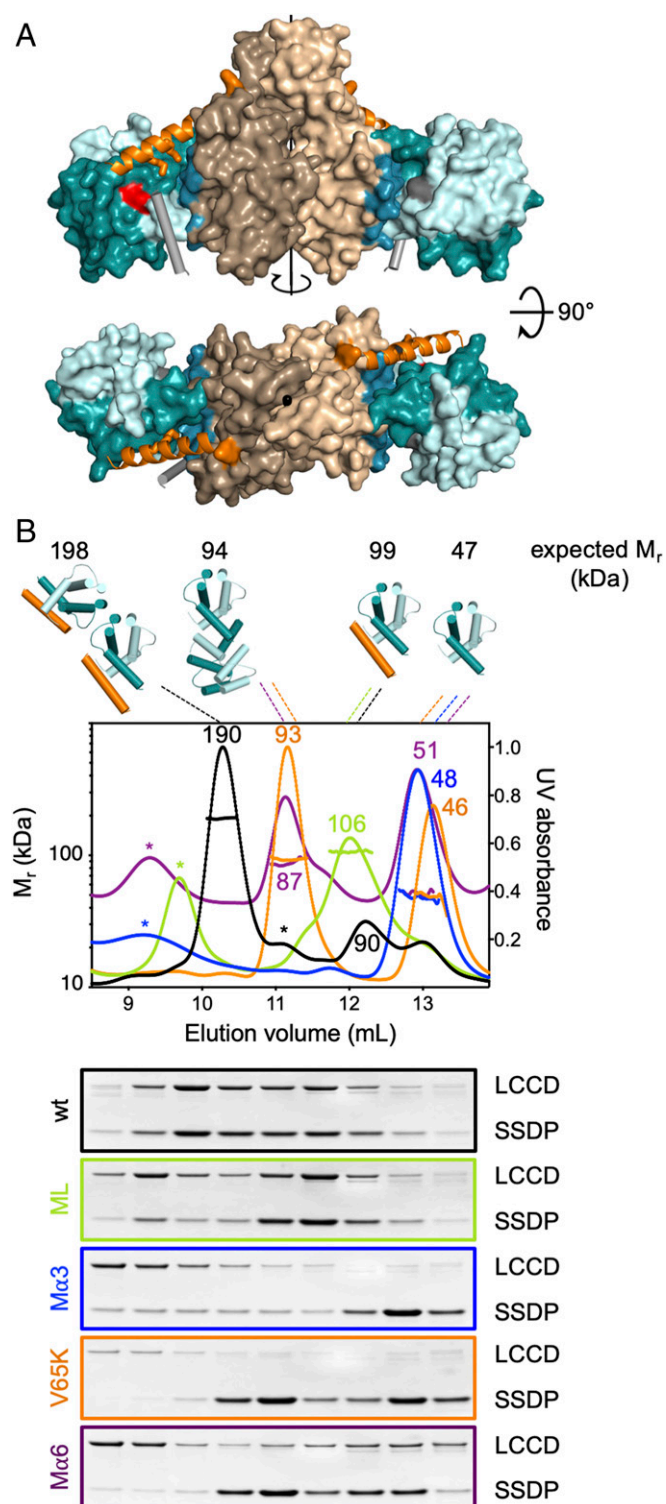


Fig. 5. Structural model of the ChiLS complex. (A) Structural model of ChiLS, derived from stoichiometry of complex and mutational analysis (see text); (Upper) side view; (Lower) view from top, revealing rotational symmetry of ChiLS and its binding sites for SSDP and NPFxD (whose precise positions and angles relative to the DD are arbitrary in this model). SSDP dimers (cyan) were docked manually onto the DARPin-binding patch (blue) in each lateral surface of the DD (wheat); (orange) predicted LCCD $\alpha 6$, with SSDP-binding residues (M402, M406) in stick; positions of the pockets for NPFxD (gray rods) are restrained by proximity to LCCD $\alpha 6$ (orange) and SSDP F58 (red); note that SSDP F58 (gray) differs from SSDP F58 (red) regarding its structural environment. (B) SEC-MALS profiles of complexes (cartoons above panels) formed between coexpressed WT and mutant MBP-LCCD_{chip} and

but also to transcriptional enhancers of developmental control genes (1–4), and of genes that control stem cell maintenance (5, 6) and normal as well as malignant erythroid differentiation (7, 32, 33). N-terminal domains of Chip/LDB proteins self-associate (12, 13), and this property is crucial for mediating long-range interactions between remote enhancers and proximal promoters (2, 8, 10, 11). The DD dimer, as revealed by our structural analysis (Fig. 3B), provides the molecular basis for this function of Chip/LDB in mediating long-range enhancer–promoter interactions. From its structure, it is difficult to see how DD could oligomerize, as previously proposed (2, 8, 13), but we note that recombinant DD has a tendency to aggregate in vitro, likely via its hydrophobic DARPin-binding patch.

Recombinant SSDP-N clearly tetramerizes via its $\alpha 3$ surface (24) (Fig. 2A and C). However, in our co-IP assays involving expression of full-length proteins in cells, SSDP uses its $\alpha 3$ surface exclusively for Chip/LDB binding, while it self-associates via its loop surface (Fig. 2C). Furthermore, our evidence suggests that LCCD binding to SSDP may promote loop-mediated self-association of SSDP (Fig. 5B), possibly by inducing a conformational change of its loop surface. Indeed, Chip/LDB-bound SSDP dimers that self-associate via their loop interfaces (Fig. 2B–D) could promote dimerization of the core ChiLS complex (Fig. 5B), assembling higher-order oligomers that could be instrumental for the function of ChiLS in mediating long-range enhancer–promoter contacts. Whatever the case, our results strongly support the notion that SSDP is essential for the function of Chip/LDB proteins in transcriptional activation by remote enhancers (14, 17, 18, 31, 34, 35).

Our structural model of ChiLS provides mechanistic insight into how ChiLS-containing enhancer complexes integrate multiple inputs from signaling and lineage factors: Because of its symmetrical SSDP₂-Chip/LDB₂-SSDP₂ architecture, a single ChiLS core complex can bind simultaneously to 2 different NPFxD ligands (17, 20) (Fig. 5A) and to 2 sets of distinct enhancer-binding proteins via its LID (binding to LIM-containing proteins, or GATA and bHLH factors) (3, 4, 7, 8). Therefore, ChiLS is uniquely poised as an integrating core module of multiprotein complexes that are tethered to transcriptional enhancers by specific combinations of DNA-binding proteins and their associated signal-responsive cofactors. Furthermore, by exchanging some of these factors, ChiLS can switch enhancer complexes between ON and OFF states.

Materials and Methods

Plasmids. Plasmids used for cell-based assays and bacterial expression were used as described previously (17), including His₈ (for DARPins), His₆-Lip (for XLdb1 and SSDP mutants), and bicistronic expression vectors (for coexpression of SSDP_{1–92} and MBP-LCCD).

Protein Purification. Proteins were expressed in BL21(DE3)-RIL, and purified with Ni-NTA resin followed by gel filtration. For crystallization, the solubility tags were removed by gel filtration following digestion with tobacco etch virus protease.

Functional Assays in Human Cells. HEK293T cells were grown in DMEM, supplemented with 10% FBS and transfected in 6-well plates with PEI. One-hundred nanograms per well of Chip, 400 ng per well of SSDP, and 500 ng per well of Pygo constructs were used for all transfections.

DARPin Selection. Ribosome display selections against the DD or ChiLS complex were carried out as detailed in *SI Appendix, Supplementary Methods*.

NMR Spectroscopy. Proteins were expressed in minimal medium supplemented by ¹⁵N-ammonium chloride and purified as described above. NMR

Lip-SSDP_{1–92}; numbers are M_r values determined by MALS (within panels), or as expected (above panels); asterisks are large aggregates; below, PAGE revealing proteins in corresponding preparative SEC fractions.

spectra were recorded with a Bruker Avance III spectrometer at 600 MHz ^1H , as described previously (17).

SEC-MALS. Purified proteins were analyzed with an Agilent 1200 Series chromatography system connected to a Dawn Heleos II 18-angle light-scattering detector combined with an Optilab rEX differential refractometer (Wyatt). Samples were loaded onto a Superdex-200 10/300 gel-filtration column (GE Healthcare) at 2 mg/mL and run at 0.5 mL/min in buffer (PBS, 1 mM DTT). Data were processed with Astra V software.

Crystallization. Concentrated proteins (10 to 20 mg/mL) were used for initial screens with ~1,500 different crystallization conditions in 100 + 100-nL drops in a 96-well sitting-drop format. Crystals emerged under multiple conditions

(*SI Appendix, Table S1*) after growing for several days at 19 °C by the vapor-diffusion method, and were directly flash-frozen in liquid nitrogen. For determination of structures, see *SI Appendix, Supplementary Methods*.

ACKNOWLEDGMENTS. We thank Roger Williams, Chris Johnson, and Melissa Gammons for technical advice and discussions; Thomas Reinberg and other members of the High-Throughput Binder Selection Facility for their contributions to generating the DARPins used in this study; Jan Steyaert, Els Pardon, and Tomasz Uchanski for nanobody production (funded by Instruct-ERIC); and the Diamond Light Source (beamlines I04 and I24) and European Synchrotron Radiation Facility (beamline ID23-1) for beamtime and assistance. This work was supported by Cancer Research UK (Grants C7379/A15291 and C7379/A24639, to M.B.) and the Medical Research Council (Grant U105192713, to M.B.).

1. A. D. Agulnick *et al.*, Interactions of the LIM-domain-binding factor Ldb1 with LIM homeodomain proteins. *Nature* **384**, 270–272 (1996).
2. D. Dorsett, Distant liaisons: Long-range enhancer-promoter interactions in *Drosophila*. *Curr. Opin. Genet. Dev.* **9**, 505–514 (1999).
3. J. M. Matthews, J. E. Visvader, LIM-domain-binding protein 1: A multifunctional co-factor that interacts with diverse proteins. *EMBO Rep.* **4**, 1132–1137 (2003).
4. R. Bronstein, D. Segal, Modularity of CHIP/LDB transcription complexes regulates cell differentiation. *Fly (Austin)* **5**, 200–205 (2011).
5. I. Dey-Guha, M. Mukhopadhyay, M. Phillips, H. Westphal, Role of Ldb1 in adult intestinal homeostasis. *Int. J. Biol. Sci.* **5**, 686–694 (2009).
6. L. Li *et al.*, Nuclear adaptor Ldb1 regulates a transcriptional program essential for the maintenance of hematopoietic stem cells. *Nat. Immunol.* **12**, 129–136 (2011).
7. P. E. Love, C. Warzecha, L. Li, Ldb1 complexes: The new master regulators of erythroid gene transcription. *Trends Genet.* **30**, 1–9 (2014).
8. G. Liu, A. Dean, Enhancer long-range contacts: The multi-adaptor protein LDB1 is the tie that binds. *Biochim. Biophys. Acta. Gene Regul. Mech.* **1862**, 625–633 (2019).
9. P. Morcillo, C. Rosen, M. K. Baylies, D. Dorsett, Chip, a widely expressed chromosomal protein required for segmentation and activity of a remote wing margin enhancer in *Drosophila*. *Genes Dev.* **11**, 2729–2740 (1997).
10. W. Deng *et al.*, Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* **149**, 1233–1244 (2012).
11. I. Krivega, R. K. Dale, A. Dean, Role of LDB1 in the transition from chromatin looping to transcription activation. *Genes Dev.* **28**, 1278–1290 (2014).
12. L. W. Jurata, G. N. Gill, Functional analysis of the nuclear LIM domain interactor NLI. *Mol. Cell. Biol.* **17**, 5688–5698 (1997).
13. A. J. Cross, C. M. Jeffries, J. Trehwella, J. M. Matthews, LIM domain binding proteins 1 and 2 have different oligomeric states. *J. Mol. Biol.* **399**, 133–144 (2010).
14. D. J. van Meyel, J. B. Thomas, A. D. Agulnick, Ssdp proteins bind to LIM-interacting co-factors and regulate the activity of LIM-homeodomain protein complexes in vivo. *Development* **130**, 1915–1925 (2003).
15. B. Lee, S. Lee, A. D. Agulnick, J. W. Lee, S. K. Lee, Single-stranded DNA binding proteins are required for LIM complexes to induce transcriptionally active chromatin and specify spinal neuronal identities. *Development* **143**, 1721–1731 (2016).
16. R. Bronstein *et al.*, Transcriptional regulation by CHIP/LDB complexes. *PLoS Genet.* **6**, e1001063 (2010).
17. M. Fiedler *et al.*, An ancient Pygo-dependent Wnt enhanceosome integrated by Chip/LDB-SSDP. *eLife* **4**, e09073 (2015).
18. J. Li *et al.*, Requirement for ssbp2 in hematopoietic stem cell maintenance and stress response. *J. Immunol.* **193**, 4654–4662 (2014).
19. C. Kadach, G. R. Crabtree, Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Sci. Adv.* **1**, e1500447 (2015).
20. L. M. van Tienen, J. Mieszczynek, M. Fiedler, T. J. Rutherford, M. Bienz, Constitutive scaffolding of multiple Wnt enhanceosome components by Legless/BCL9. *eLife* **6**, e20882 (2017).
21. R. T. Collins, J. E. Treisman, Osa-containing Brahma chromatin remodeling complexes are required for the repression of wingless target genes. *Genes Dev.* **14**, 3140–3152 (2000).
22. M. H. Kim *et al.*, The structure of the N-terminal domain of the product of the lissencephaly gene Lis1 and its functional implications. *Structure* **12**, 987–998 (2004).
23. J. Oberoi *et al.*, Structural basis for the assembly of the SMRT/NCOR core transcriptional repression machinery. *Nat. Struct. Mol. Biol.* **18**, 177–184 (2011).
24. H. Wang, Z. Wang, Q. Tang, X. X. Yan, W. Xu, Crystal structure of the LUF5 domain of human single-stranded DNA binding protein 2 (SSBP2). *Protein Sci.* **28**, 788–793 (2019).
25. E. Pardon *et al.*, A general protocol for the generation of nanobodies for structural biology. *Nat. Protoc.* **9**, 674–693 (2014).
26. A. Plückthun, Ribosome display: A perspective. *Methods Mol. Biol.* **805**, 3–28 (2012).
27. T. Lundqvist *et al.*, Crystal structure of scytalone dehydratase—A disease determinant of the rice pathogen, *Magnaporthe grisea*. *Structure* **2**, 937–944 (1994).
28. E. M. Westbrook, O. E. Piro, P. B. Sigler, The 6-A crystal structure of delta 5-3-keto-steroid isomerase. Architecture and location of the active center. *J. Biol. Chem.* **259**, 9096–9103 (1984).
29. T. L. Amyes, J. P. Richard, Specificity in transition state binding: The Pauling model revisited. *Biochemistry* **52**, 2021–2035 (2013).
30. T. L. Bullock, W. D. Clarkson, H. M. Kent, M. Stewart, The 1.6 angstroms resolution crystal structure of nuclear transport factor 2 (NTF2). *J. Mol. Biol.* **260**, 422–431 (1996).
31. N. Meier *et al.*, Novel binding partners of Ldb1 are required for haematopoietic development. *Development* **133**, 4913–4923 (2006).
32. I. A. Wadman *et al.*, The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. *EMBO J.* **16**, 3145–3157 (1997).
33. T. H. Rabbitts, LMO T-cell translocation oncogenes typify genes activated by chromosomal translocations that alter transcription and developmental processes. *Genes Dev.* **12**, 2651–2657 (1998).
34. L. Chen *et al.*, Ssdp proteins interact with the LIM-domain-binding protein Ldb1 to regulate development. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 14320–14325 (2002).
35. N. Nishioka *et al.*, Ssdp1 regulates head morphogenesis of mouse embryos by activating the Lim1-Ldb1 complex. *Development* **132**, 2535–2546 (2005).