

Supplementary Information

Crystal Structure of the N-terminal region of Human Ash2L Reveals a Winged Helix Motif Involved in DNA Binding

Yong Chen^{1,2,*}, Bingbing Wan^{1,2,*}, Kevin C. Wang^{4,5,*}, Fang Cao³, Yuting Yang²,
Angeline Protacio^{4,5}, Yali Dou^{2,3}, Howard Y. Chang^{4,5}, Ming Lei^{1,2#}

¹Howard Hughes Medical Institute, ²Department of Biological Chemistry, ³Department of Pathology, University of Michigan Medical School, 1150 W. Medical Center Drive, Ann Arbor, MI 48109, USA

⁴Howard Hughes Medical Institute, ⁵Program in Epithelial Biology, Stanford University School of Medicine, CA 94305, USA

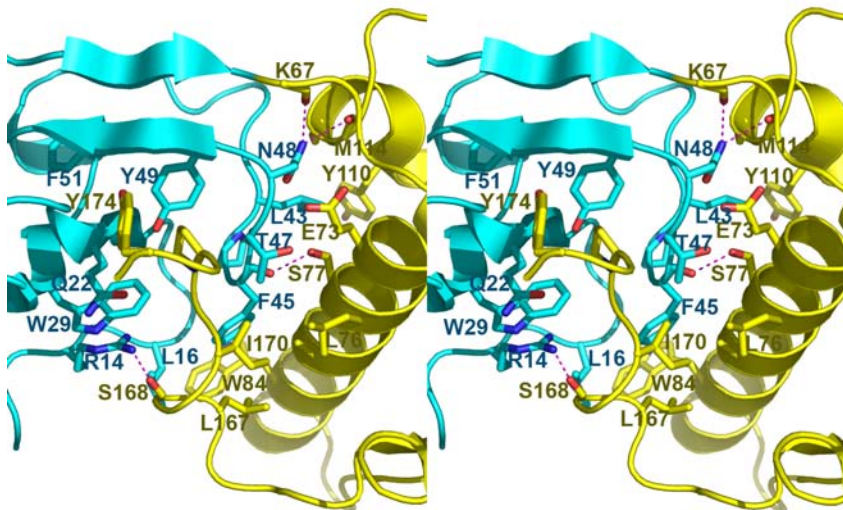
*These authors contribute equally to this work

#Correspondence and requests for materials should be addressed to M.L. (leim@umich.edu).

SUPPLEMENTARY FIGURES

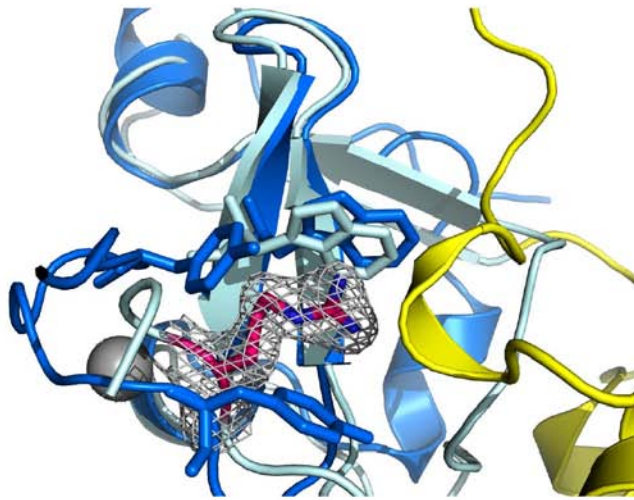
	PHD			SPRY	
Hm	TQAGSVDEENGRQLGEVELQCGICTKWFTADTF----GIDT--SSCLPFM	46	Hm	LYRACLYERVLLALHADRAPQLKISDDRLTVVGEKGYSMVRASHGVRKGAW	334
Mm	TQAGSVDEENGRQLGEVELQCGICTKWFTADTF----GIDT--SSCLPFM	135	Mm	LYRACLYERVLLALHADRAPQLKISDDRLTVVGEKGYSMVRASHGVRKGAW	416
Dm	SAAGVCYCGKERNLNIVELLCATCSRWHVETCV----SYQLGKGLLPI	144	Dm	LYRILVPHSVLLALHADRAPQLKISDRDLAVTGERGYCMVRATHSVNRGCW	416
Ag	----CYCGKERNLNIVELLCATCSRWFHESCI----GFQL--GRLVPFM	60	Ag	LYRVLSPNAVLIALHADRAPQLKVSERDLSITGEKGYCMARASHVYTKGCW	323
Ce	----CYCDGKRELGSVEVVCSTCLKWFHGRCLKEFHEFSR--NSNGVPFM	65	Ce	HYRELLNPTVNVSSNDRAFQLSINGN--SITGFEGYSMARASHVSGKGTW	341
	PHD			SPRY	
	WH			SPRY	
Hm	TNYSFHCNVCHHSNGTYFLRKQANLKEMCLSALANLTWQSRQTQDEHPK--	94	Hm	YFEITVDEMPDDTAARLGWSQPLGNLQAPLGYDKFSYSWRSKKGTKFHQ	382
Mm	TNYSFHCNVCHHSNGTYFLRKQANLKEMCLSALANLTWQSRQTQDEHPK--	183	Mm	YFEITVDEMPDDTAARLGWSQPLGNLQAPLGYDKFSYSWRSKKGTKFHQ	466
Dm	TNYVFCVCKNCSASGLESFRKVSQATISQMCHCAIANMQ--QAASRDGRRQ--	191	Dm	YFEVTIEEMPDGAATRLGWGREYGNLQAPLGYDKFSYSWRSKKGTKFHES	466
Ag	MNYVFCVCKNCSMTGLESFRKVSQATISQMCHCAIANMQ--QAASRDGRRQ--	107	Ag	YWEATVEDMPDGSACRLGWGQEVANLQAPLGYDKFSYSWRSKKGTKFHES	373
Ce	ICYTFTCKQCRPTA--EDWKAKKADLVQMCVTVLATLSAERLKADGKLSAE	114	Ce	YFEVNFDDQDDSHIRIGWSQSYASLQACVGNKFSYGWSRSHKGTKFHEA	391
	WH			SPRY	
Hm	-----TMFSKDKDIIPFIDKYWECMTTRQRPGKMTWPNNIVKMSKE-R	142	Hm	IGKHYS--SGYQGQDVLGFYINLPEDTE-----TAKSLPDTYKDKALI	430
Mm	-----TMFSKDKDIIPFIDKYWECMTTRQRPGKMTWPNNIVKMSKE-R	226	Mm	IGKHYS--SGYQGQDVLGFYINLPEDTE-----TAKSLPDTYKDKALI	507
Dm	-----IQFSKDKDIIPYIEQYWEAMTTMPRLTQSWYSTVQRSVLKDVQ	235	Dm	HGKHYS--DAYEGDVLGFLIELPEEAS-----LDYLPNTFKDRPLV	506
Ag	-----LMFSKDKDIIPYMDHYWEAMTTMARRSTQSWYATVQRSVLKDVQ	151	Ag	HGKHYS--AGYEGDVLGFLIELPEEAS-----LDYLPNTFKDRPLV	413
Ce	HVPEDFTYLSLKDEIVPYMNYMNYMLTAIKQ--KKEWHQNLAPTLLKE-K	161	Ce	KGKHYFGGGFKQDVLGCLLHLPVDKLLQIPANLPSEKLYPVSHKGFENLI	441
	WH			SPRY	
Hm	DVFLVKEHPDPSKDEEDYPKFGLLDQDLSNIGPAYDNQKQSSAVSTSG	190	Hm	KFKSYLYFEEKDFVDKAEKSLKQTPHSEIIFYKNGVNOGVAYKDFEGVY	478
Mm	DVFLVKEHPDPSKDEEDYPKFGLLDQDLSNIGPAYDNQKQSSAVSASG	276	Mm	KFKSYLYFEEKDFVDKAEKSLKQTPHSEIIFYKNGVNOGVAYKDFEGVY	557
Dm	TLFTYEEHAHGA-----MYGLFHQDLRIIPNYESMSKSGALRLTD	277	Dm	KFKSHLYYEDDKDITETLRNLHIHQGSRIEFFKNGQSGVAFEDIYAGSY	556
Ag	TLFSYEESENEQCG-----MYGLANTDLTIQKPTYDEAT-----	184	Ag	KFKSHLYYEDDKDRVNETLKALKVQPGSKIHIFKNGVNOGVAYKDFEGVY	463
Ce	NIFV--QHNDDDD-----LFALAEKNLSLLGPLHEAVLIGKRPIER	201	Ce	SFKANYFEVQEEASADIAKTLVEMPGSYIEFFHNGKSCGKAYENIYAGAY	491
				SPRY	
				SDI	
Hm	NLN-----GGIAAGSSGKGRGAKRK--QDDGGTTGTTKARSDDLFS	238	Hm	FPAISLYKSTVSVNFGPCFKYPP--KDLTYRPMSDMGWAVVEHTLADV	526
Mm	NLN-----GGIAAGSSGKGRGAKRK--QDDGGTTGTTKARSDDLFS	317	Mm	FPAISLYKSTVSVNFGPCFKYPP--KDLTYRPMSDMGWAVVEHTLADV	605
Dm	-----DGYTQASLKNRQRKRFPGTD--SGPTGKGRPSSD--ITAN	316	Dm	FPAISLHKSATVSVNFGPAFKYPEVLVEHKAQGMHVRVEELITEQCLADT	606
Ag	-----ALATQNFSSKSRQKRKLPNSEQSGALGKKSRLGTD--VGAL	223	Ag	YPAISLHKNTVTSVNFVGFKFKHPEVLVEHKAQSMHERVEEMICEQTMADM	513
Ce	ENREPRHI--ELPIEGPKTRGASKR--RHAEAPVTGKKQLAAD---YS	244	Ce	YPSISIFKSATATMNLGPKFRNLP----RGATGIHARAEQQHQEQLTSDM	537
	SPRY			SDI	
Hm	QRLPFGHYPLEHPFNKDGYYRIIAEPDPHAPDPEKLELDC--WAGKPIPGD	286	Hm	LYHVETEVD---GR-RSPP-----WEP-----	534
Mm	QRLPFGHYPLEHPFNKDGYYRIIAEPDPHAPDPEKLELDC--WAGKPIPGD	366	Mm	LYHVETEVD---GR-RSPP-----WEP-----	623
Dm	VKLPHGYPLEHPFNKDGYYRIIAEPDPHAPDPEKLELDC--WAGKPIPGW	366	Dm	LYL--TEHD---GRLRLDN-----MGL-----	623
Ag	VKLPAHGYPLEHPFNKDGYYRIIAEPDPHAPDPEKLELDC--WAGKPIPGW	273	Ag	MYF--TEND---GKRLRDT-----YSI-----	530
Ce	STAAPNGVQIDIPFSKDNRYRYLTVEDPNVPEPDAWNQ--SSAYVIPSF	293	Ce	LYLVSKEVNLHDPFRVREDDDDVDKIKKEIKQEI-----	572

Supplementary Fig S1. Multiple sequence alignment of Ash2L proteins from five species, *Homo sapiens*, *Mus musculus*, *Drosophila melanogaster*, *Anopheles gambiae*, and *Caenorhabditis elegans*. Domains of Ash2L are highlighted by colored bars above the sequences (PHD in cyan, WH in yellow, SPRY in green and SDI in red).

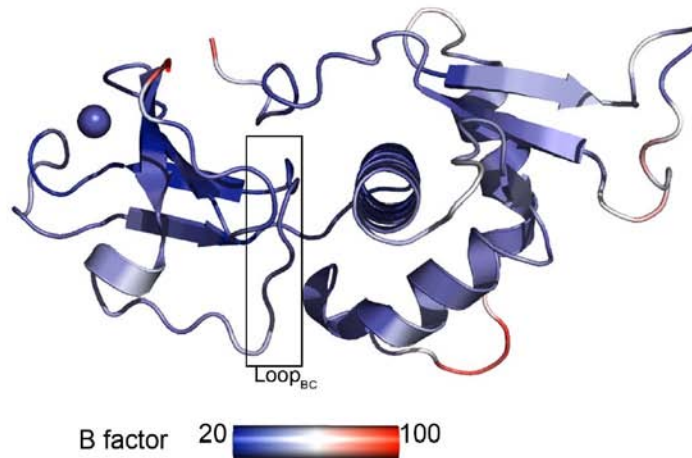


Supplementary Fig S2. Stereo view of the interface between the PHD finger (cyan) and the WH motif (yellow) of Ash2L. Ash2L_{PHD} and Ash2L_{WH} pack across an extensive interface, involving both hydrophobic and electrostatic interactions (shown as magenta dotted lines).

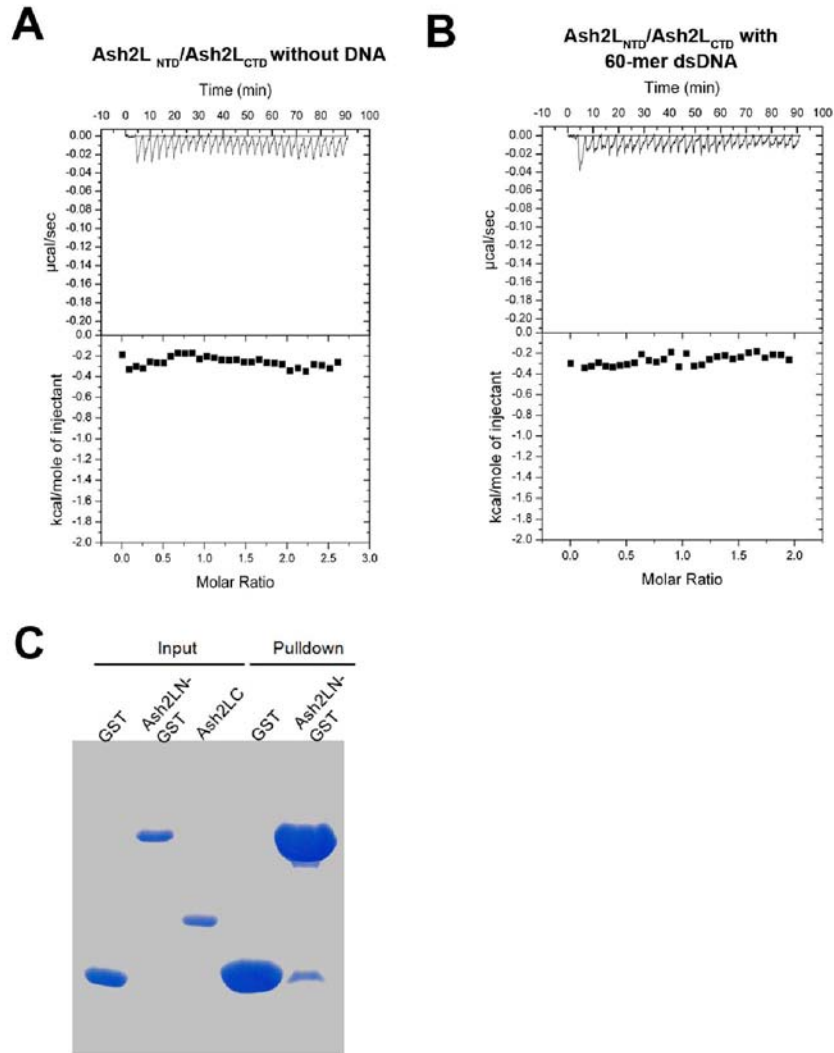
A



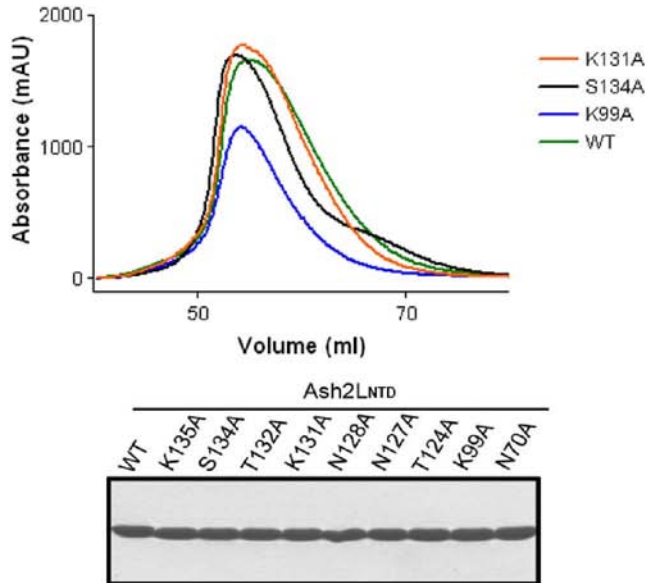
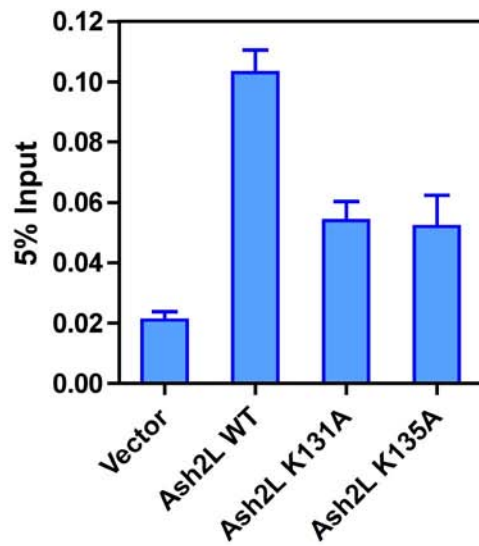
B



Supplementary Fig S3. Ash2L_{NTD} is not a histone-binding module. (A) Structure superimposition of Ash2L_{PHD} (pale cyan) and PHF8_{PHD} (marine). The electron density (2Fo-Fc) map, contoured at 1.0 σ is shown for Arg14 of Ash2L_{NTD}. (B) The structure of Ash2L_{NTD} is colored by B-factor. The lowest B factor is in blue and the highest in red. The loop between strands β B and β C has low B factor and its conformation is well defined.



Supplementary Fig S4. Ash2L_{NTD} does not directly interact Ash2L_{CTD}. (A) *In vitro* ITC measurement of the interaction between Ash2L_{NTD} and Ash2L_{CTD} showed that there is no direct interaction between these two domains of Ash2L. (B) *In vitro* ITC measurement of the interaction between Ash2L_{NTD} and Ash2L_{CTD} in the presence of 60mer dsDNA. (C) *In vitro* pull-down assay showed no interaction between Ash2L_{NTD} and Ash2L_{CTD} in the presence of DNA. GST-Ash2L_{NTD} was incubated with Ash2L_{CTD} in the present of 60-mer dsDNA.

A**B**

Supplementary Fig S5. Mutational studies of Ash2L. (A) Preparation of mutant Ash2L_{NTD} proteins. Upper panel: gel filtration chromatography profiles (Hiload Superdex 75) of wild-type and three representative mutant Ash2L_{NTD}. Lower panel: SDS-PAGE of wild-type and nine mutant Ash2L_{NTD} proteins corresponding to the peak fractions in the gel filtration profiles. (B) Mutants K131A and K135A decreased Ash2L localization to the *HOXC8* locus in cells, as shown by ChIP assay. Error bars in the graph represent standard deviation.

Supplementary Table S1. Data collection, phasing and refinement statistics for human Ash2L_{NTD}

	Se-Met Ash2L _{NTD}
Data collection	
Space group	<i>P</i> 3 ₁ 21
Cell dimensions	
a, b, c (Å)	49.984, 49.984, 165.519
a, b, γ (°)	90, 90, 120
Wavelength (Å)	0.9785 (Se peak)
Resolution (Å)	100-2.1
R _{merge}	0.063(0.525)*
I / σ I	73.7(3.8)
Completeness (%)	99.4(97.9)
Redundancy	19(10.2)
Refinement	
Resolution (Å)	35-2.1
No. reflections	26822
R _{work} / R _{free} (%)	21.3/25.4
No. atoms	
Protein	1336
Ligand (Zn)	1
Water	75
B-factors (Å ²)	
Protein	44.73
Ligand (Zn)	35.21
Water	44.57
R.m.s deviations	
Bond lengths (Å)	0.011
Bond angles (°)	1.117

*Values in parentheses are for highest-resolution shell.