SUPPLEMENTARY INFORMATION

The same pocket in menin binds both MLL and JunD, but oppositely regulates transcription

Jing Huang, Buddha Gurung, Bingbing Wan, Ke Wan, Xianxin Hua, and Ming Lei

Supplementary Text

Crystallization of menin and the menin-MLL1_{MBM} complex

Because there were no recognizable domains in menin¹, to begin to structurally characterize this protein and its interaction with MLL1, we initially expressed and purified full-length human menin. However, full-length protein was unstable by itself and prone to be degraded (data not shown). Multiple sequence alignment of menin proteins from different organisms clearly revealed a less conserved region, which is dominated by stretches of charged residues and hydrophilic amino acids that favor disordered structures (e.g., glycine, proline, and serine) (Supplementary Fig. 3a). This region in human menin consists of residues 460-550 (Supplementary Fig. 3a). Protein sequence analysis by various programs, including FoldIndex program², predicted that human menin₄₆₀₋₅₄₀ is not folded (Supplementary Fig. 3b). By contrast, except for the termini and two short segments, the remainder of menin was predicted to be folded (Supplementary Fig. 3b). Notably, a strong yeast-two-hybrid interaction was detected between the two regions that are separated by menin₄₆₀₋₅₄₀ (Supplementary Fig. 3c). Thus, these analyses suggested that menin adopts a compact fold with a large 80-residue unstructured loop in the middle of the protein (Supplementary Fig. 3a). For structural studies, a panel of menin constructs with different deletions inside menin₄₆₀₋₅₄₀ were prepared and used in crystallization trials (data not shown). After extensive screening, one such deletion mutant menin $_{\Lambda460-519}$ yielded proteins with good quality and generated crystals suitable for structural determination.

Structural architecture of menin

Although primary sequence analysis failed to identify any known protein motif in menin, strikingly, an unbiased search using Dali³ revealed unequivocal structural resemblance of individual menin domains with well-characterized protein motifs in the database (Supplementary Fig. 4a). The structure of menin thumb domain is most similar to the transglutaminase-like fold of peptide-N-glycanase (PNGase), which characterizes the transglutaminase superfamily⁴. The transglutaminase-fold of PNGase can be superimposed onto the thumb domain with a root-mean-square deviation (rmsd) of \sim 3.6 Å in the positions of 92 equivalent $C\alpha$ atoms (Supplementary Fig. 4b). However, the menin thumb domain lacks the Cys-His-Asp catalytic triad characteristic of the transglutaminase family⁴ (Supplementary Fig. 4c). The corresponding residues $(Ser₁₅₄-His₁₈₁-Glu₁₉₅)$ in the thumb domain are buried inside the structure by a large loop that is absent in PNGase (Supplementary Fig. 4c). Furthermore, when the structure of menin-MLL 1_{MBM} is superimposed with that of PNGase complexed with the substrate

peptide⁵, MLL1_{MBM} and the substrate peptide occupy the same binding site on menin and PNGase (Supplementary Fig. 4b). Therefore, we conclude that the thumb domain of menin is a transglutaminase-like protein-protein interaction module that lacks enzymatic activity.

The palm domain of menin contains a meander of eight α helices that are arranged in a head-to-tail manner and closely resembles the two TPR (tetratrico-peptide repeat) domains of Hop (Hop_{TPR1} and Hop_{TPR2A}) with rmsd values of 2.7 Å and 2.6 Å, respectively (Supplementary Figs. 4d and 4e). TPR domains normally consist of three or more TPR motifs, which are highly degenerate 34-amino-acid repeats⁶. Similar to Hop_{TPR1} and Hop_{TPR2A}, menin_{palm} consists of three TPR motifs and a C-terminal helix, which is also an integral part of the Hop TPR domains (Supplementary Fig. 4f). Notably, the first TPR motif of the palm domain contains a putative nuclear receptor-binding motif $-$ LLWLL⁷ (Supplementary Fig. 4f), which was proposed to mediate the interaction with estrogen receptor- α (ER α)⁸. However, our crystal structures showed that these leucine residues have structural stabilizing roles in a buried environment (Supplementary Fig. 4g), suggesting that the LLWLL sequence of menin is unlikely involved in the interactions with $ER\alpha$ and other nuclear receptors. TPR domains serve as protein-protein interaction modules⁶. The regular TPR helical meanders form cradle-shaped grooves that accommodate peptides in an extended conformation⁹ (Supplementary Figs. 4d and 4e). In contrast, both the palm and the thumb domains in menin are required for $MLL1_{MBM}$ binding. They constitute a palm-like cavity that forces the $MLL1_{MBM}$ peptide to adopt a

highly coiled conformation (Fig. 1d).

A comparison of NTD of menin with the β-hairpin domains of nucleotide excision repair protein XPA^{10} reveals an intriguing similarity (Supplementary Fig. 4h). Although structure superposition shows poor alignment, they share a common topology. The primary feature of these domains is a large β-hairpin supported by a helical bundle (Supplementary Fig. 4h). The concaved surface of the helical bundle in XPA mediates interactions with DNA molecules¹⁰. In contrast, the same surface on menin_{NTD} binds one end of the palm domain and is essential for protein stability (Fig. 1d).

Unlike the other three domains, the fingers domain of menin has no homology with any known motif in the database. It is composed of two sets of α/β motifs (menin402-459 and menin 548-581), which are closely associated in 3D structure but separated in primary sequence by a large 90-residue loop (Supplementary Fig. 4i). The association between the two α/β motifs is mediated by extensive hydrophobic contacts, a parallel β-strand pair as well as several clusters of electrostatic interactions Supplementary Fig. 4i). This interface between the two α/β motifs is essential for menin integrity as deletion of the C-terminal α/β motif greatly reduced the stability of the protein as attempts to express and purify menin₁₋₄₆₀ only yielded insoluble and aggregate products (data not shown). The importance of this region of menin is further highlighted by the identification of many *MEN1* frameshift mutations that completely lack the second α/β motif and missense mutations that disrupt the interaction between the two motifs (Fig. 1f and Supplementary Table 2).

Another unique feature of the fingers domain is that it contains two nuclear localization signals (NLS1 and NLS2), which play a redundant role in targeting menin to the nucleus (Supplementary Fig. 3a) (Guru et al., 1998). NLS1 (amino acids 479-497), in the middle of the predicted unstructured loop, is not included in the current menin structure (Fig. 1c). In contrast, NLS2 is at the C-terminus of the protein (Fig. 1c), and part of it folds into a short α helix (α 22) that protrudes away from the rest of the protein and binds to a neighboring molecule in the crystals (Supplementary Fig. 4j). The helical structure of NLS2 seen here is very different from the extended-chain conformation adopted by NLS peptides when bound to their associating proteins that impose a particular conformation on the NLS peptides¹¹. As signals for subcellular localization are in general short peptide segments and recognized by their linear sequence, we expect that the menin NLS2 will be an extended chain, not an α helix, when bound to importin molecules. The propensity of NLS2 to form an α helix, however, might also confer this region of menin additional functions after menin is translocated into the nucleus. Further investigation is needed to clarify this possibility.

In addition to NLS1 and NLS2, we previously identified an accessory NLS (NLSa) (residues 546-572) in menin. Notably, NLSa folds into two short α helices and mediates the association between the two α/β motifs in the fingers domain (Supplementary Fig. 4i). Thus, consistent with its name, NLSa only plays an accessory role as a nuclear localization signal¹².

Supplementary References

- 1. Chandrasekharappa, S. C. et al. Positional cloning of the gene for multiple endocrine neoplasia-type 1. *Science* **276**, 404-7 (1997).
- 2. Prilusky, J. et al. FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics* **21**, 3435-8 (2005).
- 3. Holm, L. & Rosenstrom, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38**, W545-9.
- 4. Makarova, K. S., Aravind, L. & Koonin, E. V. A superfamily of archaeal, bacterial, and eukaryotic proteins homologous to animal transglutaminases. *Protein Sci* **8**, 1714-9 (1999).
- 5. Lee, J. H., Choi, J. M., Lee, C., Yi, K. J. & Cho, Y. Structure of a peptide:N-glycanase-Rad23 complex: insight into the deglycosylation for denatured glycoproteins. *Proc Natl Acad Sci U S A* **102**, 9144-9 (2005).
- 6. Lamb, J. R., Tugendreich, S. & Hieter, P. Tetratrico peptide repeat interactions: to TPR or not to TPR? *Trends Biochem Sci* **20**, 257-9 (1995).
- 7. Heery, D. M., Kalkhoven, E., Hoare, S. & Parker, M. G. A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* **387**, 733-6 (1997).
- 8. Dreijerink, K. M. et al. Menin links estrogen receptor activation to histone H3K4 trimethylation. *Cancer Res* **66**, 4929-35 (2006).
- 9. Scheufler, C. et al. Structure of TPR domain-peptide complexes: critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. *Cell* **101**, 199-210 $(2000).$
- 10. Ikegami, T. et al. Solution structure of the DNA- and RPA-binding domain of the human repair factor XPA. *Nat Struct Biol* **5**, 701-6 (1998).
- 11. Cook, A., Bono, F., Jinek, M. & Conti, E. Structural biology of nucleocytoplasmic transport. *Annu Rev Biochem* **76**, 647-71 (2007).
- 12. La, P. et al. Tumor suppressor menin: the essential role of nuclear localization signal domains in coordinating gene expression. *Oncogene* **25**, 3537-46 (2006).

Supplementary Table 2 | Structural analysis of the MEN1-related missense mutations and in-frame deletions.

*Abbreviations:

-
HPT, hyperparathyroidism; PET, endocrine pancreas tumors; PIT, anterior pituitary tumors; NET, thymic/bronchic neuroendocrine tissue tumors; ADR, adrenal gland
tumors; PanNET, pancreatic neuroendocrine tumors; NET-meta,

Supplementary Figure 1 | ITC Measurements of interaction between menin and various N-terminal fragments of MLL1 (related to Fig. 1a).

Supplementary Figure 2 | Characterization of the menin-MLL interaction. a, Sequence alignment of the extreme N-termini of MLL1, MLL2 and TRX proteins from different organisms (*Hs*: *Homo sapiens*; *Mm*: *Mus musculus*; *Xt*: *Xenopus tropicalis*; *Dr*: *Danio rerio*; *Dm*: *Drosophila melanogaster*). The conserved residues and the basic residues are highlighted in yellow boxes. **b,** ITC measurement of the interaction of menin with the MLL 2_{MBM} peptide. Insert is the ITC titration data. The binding curve was fit to a one-binding-site-per-menin model.

Supplementary Figure 3 | Crystallization of the Menin-MLL1_{MBM} Complex. a, Primary sequence alignment of menin family members. Secondary structure assignments are shown as cylinders (α helices) and arrows (β strands), and colored as in Fig. 1c. Red dots denote MEN1-related missense mutations or in-frame deletions, purple dots denote residues involved in MLL1 binding, and cyan dots label the phosphorylation sites. Conserved residues are highlighted in yellow. Nuclear localization sequences (NLS) are boxed. **b,** Protein folding analysis of menin by FoldIndex program. **c,** Interaction between Menin₁₋₄₆₀ and Menin₅₄₀₋₆₁₀, detected by yeast-two-hybrid. **d**, ITC measurement of the interaction of menin_{Δ 460-519} with the MLL1_{MBM} peptide.

Supplementary Figure 4 | Structural architecture of menin. a, The topology diagram of menin that is colored as in Fig. 1c. **b,** Superposition of the structure of Menin_{thumb}-MLL1_{MBM} on that of human PNGase_{TGD}-substrate mimic. Menin_{Thumb}-MLL1_{MBM} and PNGase_{TGD}-substrate mimic are colored in green and yellow,

and light-blue and red, respectively. c , Structural alignment of Menin_{Thumb} and $PNGase_{TGD}$ that shows the superposed catalytic triad as well as the loop L58 that buries the triad in menin structure. **d**, Structural alignment between the menin_{palm}-MLL1_{MBM} and the Hop $_{TPR1}$ -peptide complexes. menin_{palm}-MLL1_{MBM} and Hop $_{TPR1}$ -peptide are colored in blue and yellow, and pink and magenta, respectively. **e,** Structural superposition between the complexes menin_{Palm}-MLL1_{MBM} and Hop_{TPR2A}/peptide. menin_{Palm}-MLL1_{MBM} and Hop_{TPR2A}-peptide are colored in blue-yellow and pink-magenta, respectively. **f**, Structure-based sequence alignment of the TPR domains among menin and other TPR-containing proteins. The TPR consensus residues are highlighted in yellow. **g,** The LLWLL motif of menin is located within a buried environment. **h,** Superposition of menin_{NTD} (orange) and XPA_{MBD} (pink) highlights the common β-sheet and the α 1 and α 2 helices. **i,** Interactions between the two α/β motifs in the fingers domain of menin. The two α/β motifs are colored in palecyan and teal, respectively. Interacting residues of the two motifs are represented as stick models, and hydrogen bonds are denoted as magenta dash lines. **j,** NLS2 adopts a helical conformation and binds to a neighboring molecule in the crystal. NLS2 is colored cyan, and the interacting residues in the other molecule are colored orange (NTD) or green (Thumb domain).

Supplementary Figure 5 | ITC measurements of interaction between menin and MLL1_{MBM} mutants (related to Fig. 2b).

Supplementary Figure 6 | ITC measurements of interaction between MLL1_{MBM} and **menin mutants (related to Fig. 2c).**

Supplementary Figure 7 | In vivo functional analysis of the menin-MLL1 interaction. a, Expression level of wild-type and mutant menin detected by western blot, related to Figs. 2h-2j. **b,** Expression of wild-type and mutant menin in menin-null MEFs does not significantly affect histone H3 distribution at the *Hoxc8* promoter region.

Supplementary Figure 8 | Characterization of the interaction between LEDGF and the menin-MLL1 complex. a, ITC measurement of the interaction between LEDGF_{IBD} and the menin-MLL $1_{MBM-LBM}$ complex. **b**, ITC measurement of the interaction between $LEDGF_{IBD}$ and menin. We could not perform the ITC measurement of the interaction between LEDGF_{IBD} and MLL1_{MBM-LBM}, as the MLL1_{MBM-LBM} peptide is prone to be degraded by itself. We tried the co-expression and co-purification of $LEDGF_{IBD}$ and MLL1MBM-LBM, but could not get the protein complex (data not shown). **c,** $MLL1_{MBM-LBM}\Delta$ could form a stable protein complex with menin and LEDGF_{IBD}, as shown in the SDS-PAGE. The staining difference between MLL1_{MBM-LBM} Δ and LEDGF_{IBD} results from the composition difference of basic residues. **d**, Comparison of $MLL1_{MBM}$ in the structures of menin-MLL1_{MBM} (red) and menin-MLL1_{MBM-LBM}Δ-LEDGF_{IBD} (yellow). **e**, Comparison of the domain organizations between MLL1 and MLL2. MLL2 lacks an LBM fragment.

Supplementary Figure 9 | Characterization of the interaction between menin and JunD_{MBM}. a, Summary of ITC analysis of the interaction between menin and various JunD peptides. **b**, ITC measurement of the interaction between menin and JunD₁₋₁₀₀. **c**, ITC measurement of the interaction between menin and JunD_{MBM}. **d**, ITC measurement

of the interaction between menin and JunD_{MBM}Δ. **e**, ITC measurement of the interaction between menin and JunD₂₇₋₃₆. **f**, The menin-JunD_{MBM} interface. JunD_{MBM} (magenta) and the interacting residues in menin (green: residues in the thumb domain; blue: residues in the palm domain) are presented as stick models. The Menin-Jun D_{MBM} intermolecular hydrogen bonds are shown as pink dash lines.

Supplementary Figure 10 | ITC measurements of interactions between menin and $JunD_{MBM} mutants (related to Fig. 4d).$

Supplementary Figure 11 | Immunoblot analysis of endogenous menin following fractionation on a Superpose 6 gel-filtration column. Positions of molecular weight markers are indicated.