

Structural basis for MOF and MSL3 recruitment into the dosage compensation complex by MSL1

Jan Kadlec^{1,2,5}, Erinc Hallacli³⁻⁵, Michael Lipp^{1,2}, Herbert Holz^{3,4}, Juan Sanchez-Weatherby^{1,2,4}, Stephen Cusack^{1,2} & Asifa Akhtar^{3,4}

The male-specific lethal (MSL) complex is required for dosage compensation in *Drosophila melanogaster*, and analogous complexes exist in mammals. We report structures of binary complexes of mammalian MSL3 and the histone acetyltransferase (HAT) MOF with consecutive segments of MSL1. MSL1 interacts with MSL3 as an extended chain forming an extensive hydrophobic interface, whereas the MSL1-MOF interface involves electrostatic interactions between the HAT domain and a long helix of MSL1. This structure provides insights into the catalytic mechanism of MOF and enables us to show analogous interactions of MOF with NSL1. In *Drosophila*, selective disruption of Msl1 interactions with Msl3 or Mof severely affects Msl1 targeting to the body of dosage-compensated genes and several high-affinity sites, without affecting promoter binding. We propose that Msl1 acts as a scaffold for MSL complex assembly to achieve specific targeting to the X chromosome.

Dosage compensation is an essential process that equalizes the expression levels of X-chromosomal genes between males and females. In *Drosophila* males, dosage compensation results in an approximately two-fold upregulation of the transcription of X-linked genes¹. In mammals, the balance in X-chromosomal gene expression is achieved by stochastic inactivation of one of the female X chromosomes². Dosage compensation of X-chromosomal genes in *Drosophila* has been extensively studied as a model of chromosome-wide transcription regulation by histone hyperacetylation^{3,4}.

In *Drosophila*, the process is mediated by the dosage-compensation complex (DCC), also known as the MSL complex, which consists of at least five male-specific lethal proteins (Msl1, Msl2, Msl3, maleless (Mle) and males-absent-on-the-first (Mof)) and two functionally redundant noncoding RNAs (roX1 and roX2)³. In humans, an equivalent complex is formed from the counterparts of at least four *Drosophila* Msl proteins (MSL1, MSL2, MSL3 and MOF), although no associated RNA has yet been identified⁵⁻⁸. The *Drosophila* MSL proteins and roX RNAs are proposed to assemble and coat the X-chromosome in a process involving at least two steps. First, numerous high-affinity sites that are enriched for GA repeat sequences, including the roX genes, are occupied. MSL complexes then spread from these sites to the rest of the X chromosome including many active genes⁹. The MSL complex is responsible for acetylation of histone H4 at lysine 16 (H4K16ac) on the X chromosome mediated by Mof¹⁰⁻¹². The details of the recruitment of the MSL complex to the X chromosome and the exact mechanism of dosage compensation remain poorly understood.

Human and *Drosophila* MSL1 consist of 614 and 1,039 amino acid residues, respectively, with no known globular domains predicted.

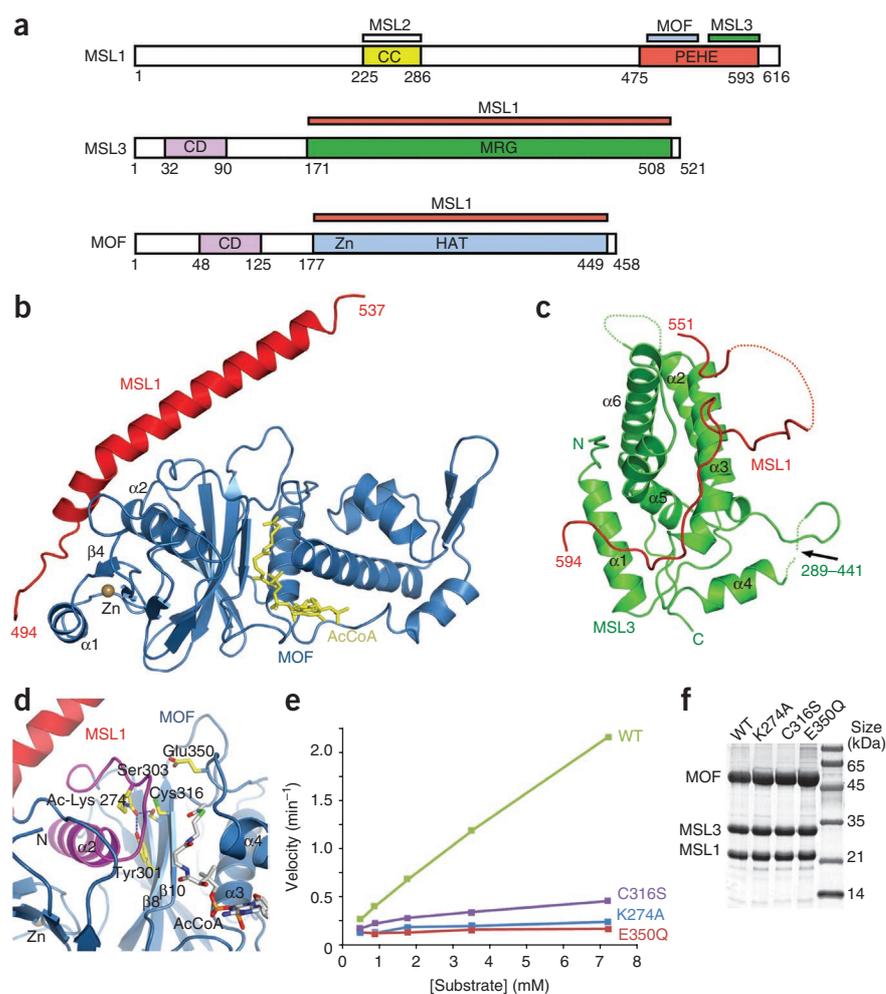
In *Drosophila*, Msl1 was shown to interact with Msl2 via its N-terminal putative coiled coil¹³ and with Msl3 and Mof via a conserved C-terminal region called the PEHE domain¹⁴. Msl3 (521 residues in human) interacts via its MORF4-related gene family (MRG) domain with Msl1 (ref. 15), whereas its N-terminal chromo-barrel domain binds nucleic acids¹⁵ and is required for the spreading of the MSL complex along the X chromosome¹⁶. Mof (458 residues in human) is a member of the MYST family of acetyltransferases¹⁰⁻¹². The Mof chromo-barrel domain is required for RNA binding¹⁷, and the zinc finger within the HAT domain interacts with Msl1 (ref. 14). To have full enzymatic activity and specificity, Mof is required to be in a complex with Msl1 and Msl3 (ref. 14). Recently, it has been found that MOF is also a key component in a second, distinct chromatin-modifying complex called the NSL complex, wherein it interacts with NSL1, which has sequence similarity with MSL1 (refs. 7,18,19).

Although MSL complex members have been intensely studied over the last decade, the detailed molecular interactions within the complex remain unknown. As a first step toward achieving an atomic-resolution understanding of the architecture and function of the MSL complex, we determined crystal structures of two mammalian subcomplexes that MSL1 forms with the MOF HAT domain and with the MRG domain of MSL3. Based on the structural results, we carried out mutagenesis in the *Drosophila* orthologs to selectively disrupt the interaction of Msl1 with either Msl3 or Mof *in vivo*, thus demonstrating that the residues critical for the observed protein-protein interactions are conserved throughout evolution. Furthermore, using the sequence similarity between MSL1 and NSL1, we could demonstrate that mutation of corresponding residues in NSL1 also selectively

¹European Molecular Biology Laboratory, Grenoble, France. ²Unit of Virus Host-Cell Interactions, Université Joseph Fourier–European Molecular Biology Laboratory–Centre National de la Recherche Scientifique (UJF-EMBL-CNRS), Unité Mixte Internationale 3265, Grenoble, France. ³European Molecular Biology Laboratory, Heidelberg, Germany. ⁴Present addresses: Max-Planck-Institut für Immunbiologie und Epigenetik, Freiburg im Breisgau, Germany (E.H., H.H., A.A.); Diamond Light Source Ltd., Didcot, UK (J.S.-W.). ⁵These authors contributed equally to this work. Correspondence should be addressed to A.A. (akhtar@immunbio.mpg.de) or S.C. (cusack@embl.fr).

Received 3 March 2010; accepted 20 October 2010; published online 9 January 2011; doi:10.1038/nsmb.1960

Figure 1 Crystal structures of the MSL1–MSL3 and MSL1–MOF subcomplexes. **(a)** Schematic representation of the domain structure of mouse MSL1 (which is essentially identical to human MSL1), MSL3 and MOF. Domain colors correspond to the ribbon diagram in **b** and **c**. The red, blue and green bars indicate MSL1, MOF and MSL3 interacting regions, respectively, as defined in this work and in ref. 13. CC, coiled coil; CD, chromo-barrel domain. **(b)** Ribbon diagram of the mammalian MSL1–MOF–acetyl-CoA complex. The HAT domain of MOF (residues 174–458) is shown in blue. The MOF secondary structures interacting with MSL1 are labeled. **(c)** Ribbon representation of the complex between MSL1 and MSL3. The MSL3 MRG domain (residues 167–288, 442–517) is shown in green, and its secondary structures are labeled. The disordered regions in MSL1 and MSL3 are shown as dots. The arrow indicates the location where residues 289–441 were deleted. **(d)** Ribbon diagram of the catalytic site of the MOF HAT domain. Acetyl-CoA is shown as sticks, with the acetyl moiety close to Cys316 and the catalytic Glu350. The hairpin structure formed by residues 257–281, shown in magenta, interacts with MSL1 (details in **Fig. 2**) and harbors acetylated lysine 274 (Ac-Lys274) that binds to $\beta 8$. **(e)** The acetylation activity of human wild-type (WT), K274A, C316S and E350Q MOF was assessed within the MOF–MSL1_{471–616}–MSL3_{167–289,442–517} complex in a fluorescence based HAT assay (DNTB assay), measuring the production of CoA during the acetylation of a histone H4 N-terminal peptide. The average of two initial velocity measurements is plotted against the substrate concentration. The rates obtained between duplicate runs were typically within 10% of each other. **(f)** Qualitative assessment of the purity of MSL complexes shown by Coomassie staining.



affects its interaction with MOF. Notably, we showed using chromatin immunoprecipitation (ChIP) assays that selective disruption of Msl1 interaction with either Mof or Msl3 severely affects targeting of Msl1 to the coding regions and to the 3' end of X-linked genes, whereas Msl1 binding to promoters is largely independent of Msl3 or Mof *in vivo*. Finally, we showed that high-affinity sites differ in their requirement of Msl3 and Mof for Msl1 recruitment.

RESULTS

Structure of the MSL1–MOF complex

The structure of a complex between the HAT domain of human MOF (residues 174–458), the N-terminal part of the MSL1 PEHE region (residues 470–540) and acetyl-CoA was determined by X-ray crystallography at 2.8-Å resolution and refined to an R_{free} of 25.6% and an R -factor of 22.3% (**Fig. 1** and **Table 1**).

As with other MYST-family HAT domains, the MOF structure consists of a central core that participates in cofactor binding with flanking N- and C-terminal regions (**Fig. 1b** and **Supplementary Fig. 1**). In the N-terminal segment, residues 205–233 form a zinc-binding module, with the absolutely conserved Cys210, Cys213, His226 and Cys230 coordinating the Zn atom.

First, we confirmed that Glu350, in a position corresponding to the putative catalytic Glu338 of Esa1 (ref. 20) (**Fig. 1d**), is also likely to be catalytic residue in MOF, as the E350Q mutation essentially abolished the MOF HAT activity in fluorescence-based acetylation assays on a histone H4 N-terminal tail peptide (**Fig. 1e,f**).

Helix $\alpha 2$ and the downstream chain form a hairpin-like structure (residues 257–281), within which Lys274 showed an additional electron density at its tip that could be modeled as an acetyl group (**Fig. 1d** and **Supplementary Fig. 2a**). Lys274 acetylation was confirmed by mass spectrometry analysis (**Supplementary Fig. 2b**). The acetylated lysine is involved in the otherwise hydrophobic interface of the hairpin and strand $\beta 8$ of the core β -sheet (**Fig. 1d**). Consistent with this, K274A MOF was substantially less stable in a thermal denaturation assay (**Supplementary Fig. 2c,d**) and was not active in acetylation assays (**Fig. 1e,f**). Notably, Lys274 was acetylated even in the E350Q mutants, probably through a residual HAT activity of this mutant (**Supplementary Fig. 3a,b**). Thus, the acetylation of Lys274 seems to be important for structural integrity and proper positioning of the residue 257–281 hairpin, which, in analogy to the structure of the Gcn5 HAT domain in a complex with H3 peptide²¹, might be involved in substrate binding. Structurally important lysine acetylation in the core of the HAT domain appears to be a conserved feature at least among mammalian MYST proteins, as the corresponding lysine residues in Tip60 and MOZ are also modeled as acetylated, forming equivalent hydrogen bonds in their deposited structures (PDB codes 2OU2 and 2OZU, respectively). It remains to be established whether this modification is involved in regulation of the HAT activity of MYST acetyltransferases, as proposed for p300/CBP²².

The next important residue to consider in the catalytic site of MOF was the conserved Cys316. The corresponding cysteine in Esa1 was shown to be acetylated, and a two-step catalytic mechanism was

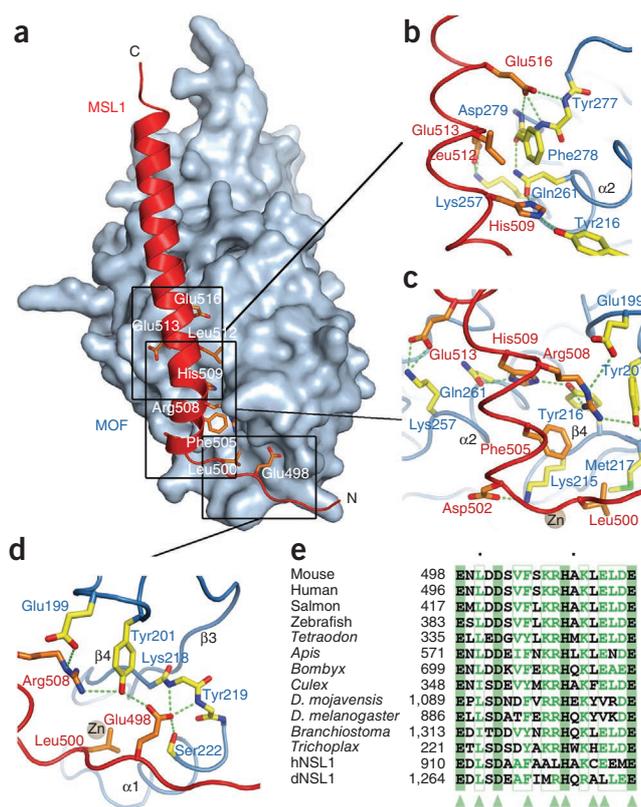


Figure 2 Structure of the MSL1–MOF complex. **(a)** Cartoon representation of MSL1 (residues 471–539, shown in red) in complex with the MOF HAT domain (shown as solvent-accessible surface in blue). MSL1 residues involved in the interaction are drawn as orange sticks. The three overlapping details of the interface shown in **b–d** are localized on the structure by the black boxes. In the detailed figures, MSL1 residues are shown in orange (labeled in red) and MOF residues in yellow (labeled in blue). **(b)** The interaction between MSL1 and the MOF hairpin including helix $\alpha 2$ (residues 257–281). MSL1 Glu516 forms hydrogen bonds with main chain amide groups of Tyr277, Phe278 and Asp279 (for clarity, side chains of Tyr277 and Asp279 are omitted). **(c)** Contacts of MSL1 with helix $\alpha 2$ and strand $\beta 4$ of the MOF zinc finger. Hydrogen bonds of MSL1 Asp502, Arg508, His509 and Glu513 with MOF residues are shown in green. Phe505 inserts into a hydrophobic pocket of MOF. **(d)** Details of the interaction between MSL1 and the zinc finger of MOF ($\alpha 1$, $\beta 4$). Glu498 hydrogen bonds main chain amide groups of Lys218 and Tyr219 as well as hydroxyl groups of Tyr201 and Ser222 of MOF (side chains of Lys218 and Tyr219 are not shown). **(e)** Sequence alignment of the MSL1 and NSL1 proteins. Only the sequence of the fragment involve in the interaction with MOF is shown. Identical residues are in green boxes. Green triangles indicate the interacting residues.

proposed for MYST acetyltransferases, whereby the acetyl moiety is transferred from acetyl-CoA to the substrate lysine via a cysteine residue²³. Contrary to this, a previous biochemical study showed that Cys304 of Esa1 is dispensable for catalysis²⁴. In our MOF structure, Cys316 is not acetylated, and mass spectrometry analysis identified only chymotryptic fragments containing unmodified Cys316 (data not shown). Notably, we showed that the C316S mutant is still partially active in the HAT assay (Fig. 1e,f) and that Lys274 is acetylated in this mutant (Supplementary Fig. 3c). The lower activity of the mutant (Supplementary Fig. 3d–f) might reflect local conformational changes due to the substitution. These data suggest that MOF does not use the two-step catalytic mechanism originally proposed for Esa1.

Next, we analyzed the interaction interface between the MOF HAT domain and MSL1. The MSL1 fragment forms a loop (residues 494–501) followed by a 52-Å-long helix (residues 502–533). Both elements interact extensively with the N-terminal part of the MOF HAT domain (Figs. 1b and 2 and Supplementary Fig. 4), with numerous, mainly charged contacts existing between the two molecules. The complex interface buries 878 Å² of MOF and 1,050 Å² of MSL1. In MOF the interaction involves three regions: helix $\alpha 1$ and strand $\beta 4$ of the zinc finger, an upstream loop connecting $\beta 2$ and $\beta 3$ (residues 197–205), and the aforementioned hairpin (residues 257–281). Details of these interactions are shown in Figure 2. In MSL1, the key interacting residues include Glu498, Asp502, Arg508, His509, Glu513 and Glu516, which form multiple hydrogen bonds and salt-bridge interactions with MOF. Additionally, Leu500, Phe505 and Leu512 are inserted in hydrophobic pockets in the center of the interface. All the MSL1 interacting residues are very well conserved among species, reflecting the importance of this interaction for the functional integrity of the MSL complex (Fig. 2e and Supplementary Fig. 5). Mapping MOF residue conservation across species onto the surface of the MOF HAT domain revealed that the MSL1-interacting region

is rather well conserved, despite the fact that the interface involves several main chain contacts (Supplementary Figs. 4a and 6). The key MSL1 binding residues in MOF are Glu199, Tyr201, Tyr216, Gln261 and Phe278. Because MSL1 fragments interacting with MOF and MSL3 are unstable when expressed alone, further binding-affinity measurements could not be performed *in vitro*.

MYST acetyltransferases function exclusively within multiprotein complexes. Given the high structural similarity among the MYST family members, a corresponding surface including the zinc finger is available for protein-protein interactions in other MYST proteins as well (Supplementary Fig. 4b,d). Indeed, within human MYST acetyltransferases, most of the residues corresponding to the MSL1 binding surface of MOF are conserved (Supplementary Figs. 4c and 7). Consistent with this, a two-hybrid screen has shown that the zinc finger of human HBO1 (MYST2) HAT is essential for the interaction with MCM2 of the minichromosome maintenance complex²⁵. An MCM2 L222A mutant deficient in HBO1 binding could be reverted by mutations of Ile380 in HBO1 (ref. 25), corresponding to MOF Ser222 on helix $\alpha 2$, which is directly involved in the interaction with MSL1 (Supplementary Fig. 7). It is thus likely that HBO1 uses a similar surface for interaction with MCM2.

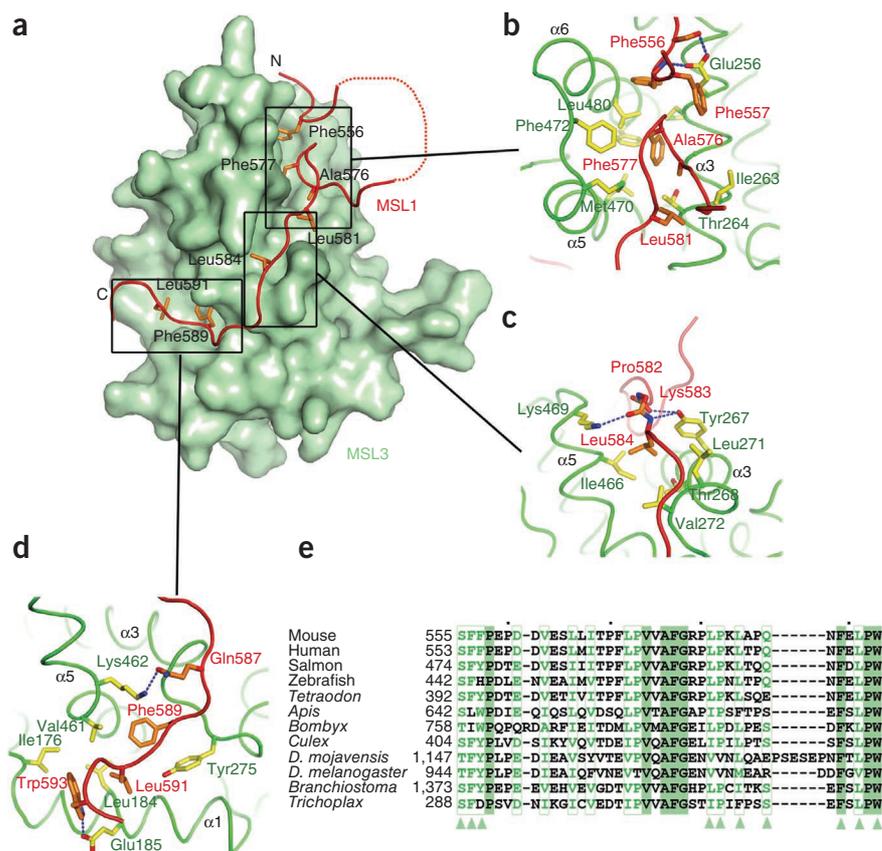
Structure of the MSL1–MSL3 complex

The MSL3 construct used here corresponds to MSL3 isoform c (residues 167–517), including the predicted MRG domain (Fig. 1a). Compared to the sequence of the known structure of the MRG domain of MRG15 (PDB entry 2F5J)²⁶, the human and *Drosophila* MSL3 domains contain two poorly conserved insertions with no predicted secondary-structure elements (residues 223–250 and 290–441 in human MSL3; Supplementary Fig. 8). To obtain diffracting crystals, the longer insertion was replaced with an eight-residue linker. This construct still bound efficiently MSL1 in a binary or ternary complex with MOF (Supplementary Fig. 9). The structure of the complex between the MSL1 PEHE region (residues 545–597) and MSL3_{167–289,442–517} was determined at 3 Å resolution. The refinement resulted in an R_{free} of 25.3% and an R -factor of 23.2% (Table 1).

The MSL3 MRG domain structure is similar to that of MRG15 (r.m.s. deviation 0.99 Å for 151 C α atoms; Fig. 1c). Electron density is missing for the short insertion connecting $\alpha 2$ and $\alpha 3$ (residues 224–245). The 151-residue region deleted from the MSL3 construct links helices $\alpha 4$ and $\alpha 5$ (Fig. 1c). Interpretable electron density

Figure 3 Structure of the MSL1–MSL3 complex.

(a) Structure of MSL1 (residues 545–597, shown in red) in complex with the MSL3 MRG domain (shown as solvent-accessible surface in green). MSL1 residues involved in the interaction are drawn as orange sticks. The three black boxes correspond to the detailed views of the complex interface shown in **b–d**, where the MSL1 residues are shown in orange (labeled in red) and MSL3 residues in yellow (labeled in green). (b) Ala576 and Phe577 insert into hydrophobic pocket formed by helices $\alpha 3$, $\alpha 5$ and $\alpha 6$. Additionally, the N-terminal Phe556 and Phe557 bind on the top of Phe577, reinforcing this interaction. MSL3 Glu256 of helix $\alpha 3$ hydrogen bonds Ser555 and the main chain of Phe556. (c) The MSL1 binds between helices $\alpha 5$ and $\alpha 3$. Leu584 inserts into a cavity between the helices. MSL1 forms several main chain interaction with MSL3 Tyr267 and Lys469 (side chains of Pro582 and Lys583 are not shown). (d) MSL1 Phe589 and Leu591 insert into pocket formed by hydrophobic residues of helices $\alpha 5$ and $\alpha 1$ and a linker between helices $\alpha 3$ and $\alpha 4$. Additionally, Gln587 and Trp593 form hydrogen bonds with Lys462 and Glu185, respectively. (e) Sequence alignment of the MSL1 fragment involved in the interaction with MSL3. Identical residues are in green boxes. The interacting residues are indicated with green triangles.



was observed for residues 551–558 and 564–594 of MSL1. In the crystal, the N- and C-terminal parts of the MSL1 PEHE region bind to two distinct MSL3 molecules, which we interpret as being a result of nonphysiological domain swapping (see Online Methods and **Supplementary Fig. 10a**). We showed by multiangle laser light scattering (MALLS) that a ternary complex of MOF HAT–MSL1 PEHE–MSL3_{167–289,442–517} is formed with an apparent 1:1:1 stoichiometry (**Supplementary Fig. 10b**).

MSL1 wraps around the MSL3 MRG domain as an extended chain (**Figs. 1c** and **3a**). The complex interface buries 1,307 Å² of MSL3. MSL1 forms numerous hydrophobic as well as several charged interactions with MSL3. The crucial interacting residues of MSL1 are four highly conserved phenylalanines (Phe556, Phe557, Phe577 and Phe589) that insert into different hydrophobic pockets on MSL3 (**Fig. 3**). The C-terminal part of the MSL1 peptide forms a short hairpin harboring Ala576 and Phe577 that inserts into a cavity formed by hydrophobic residues of the helical hairpin $\alpha 5$ – $\alpha 6$ and a perpendicular helix $\alpha 3$. This interface is reinforced by the N-terminal part of the MSL1 peptide, which binds on the top of the MSL1 hairpin with Phe556 positioned between helices $\alpha 6$ and $\alpha 3$. MSL1 then folds around the last turn of $\alpha 3$ to place Phe589 and Leu591 into another hydrophobic surface formed by helices $\alpha 5$ and $\alpha 1$. Details of these interactions are given in **Figure 3**. Essentially all MSL3 and MSL1 residues involved in the interaction are well conserved among species (**Fig. 3e** and **Supplementary Figs. 5, 8** and **10c,d**).

The MSL1–MSL3 structure is, to our knowledge, the first reported MRG domain in complex with its binding partner. Previous studies revealed that mutations of MRG15 residues corresponding to MSL3 Leu480 and Phe484, which form the hydrophobic pocket surrounding MSL1 Phe577 and Phe556, abolish the interaction of MRG15 with MRGBP²⁷. Additionally, the same surface is involved in MRG15 dimerization, in which this hydrophobic pocket accommodates Tyr276 and Leu279 of another protomer²⁶. MRG15 residues

corresponding to the hydrophobic surface binding MSL1 Phe589 are involved in the interaction with the N terminus of PAM14 (ref. 26). Thus, similar hydrophobic surfaces appear to be generally used for protein–protein interactions in MRG domains. Mapping of phylogenetically conserved residues onto the surface of MSL3 revealed an additional highly conserved region formed by charged residues of helices $\alpha 2$ and $\alpha 3$ (**Supplementary Fig. 10e**) that might be involved in another protein–protein or protein–nucleic acid interaction.

The HAT activity of MOF is enhanced by MSL1 and MSL3

In *Drosophila*, the Mof HAT activity is enhanced by the presence of Msl1 and Msl3 (ref. 14). We were interested in identifying the minimal regions required for this activity and in determining whether the fragments used for crystallization are sufficient for enhanced acetylation. We therefore copurified mammalian complexes containing the MOF HAT domain with various MSL1 PEHE and the MSL3 MRG domain fragments (**Supplementary Fig. 9a,b**). The full-length *Drosophila* trimer complex containing Mof, Msl1 and Msl3 was used as a positive control (**Supplementary Fig. 9c**). Consistent with previous observations, the presence of Msl1 and Msl3 enhanced the HAT activity of *Drosophila* Mof on nucleosomal substrates (**Fig. 4a**; compare lanes 11 and 12 with lanes 9 and 10). Notably, the mammalian complex containing the human MOF HAT, MSL1 PEHE domain and MSL3 MRG domain was sufficient for enhancing the acetylation activity (**Fig. 4a**, lanes 5 and 6). However, removing the large insertion within the MSL3 MRG domain (residues 290–441) did not enhance HAT activity to a similar extent, indicating that this segment is required for the full activation potential of MOF (**Fig. 4a**, lanes 7 and 8). These results show that in the mammalian system as well, interaction of MOF with MSL3 via MSL1 enhances the HAT activity of MOF, whereas stimulation is not observed with a subcomplex containing only MOF and

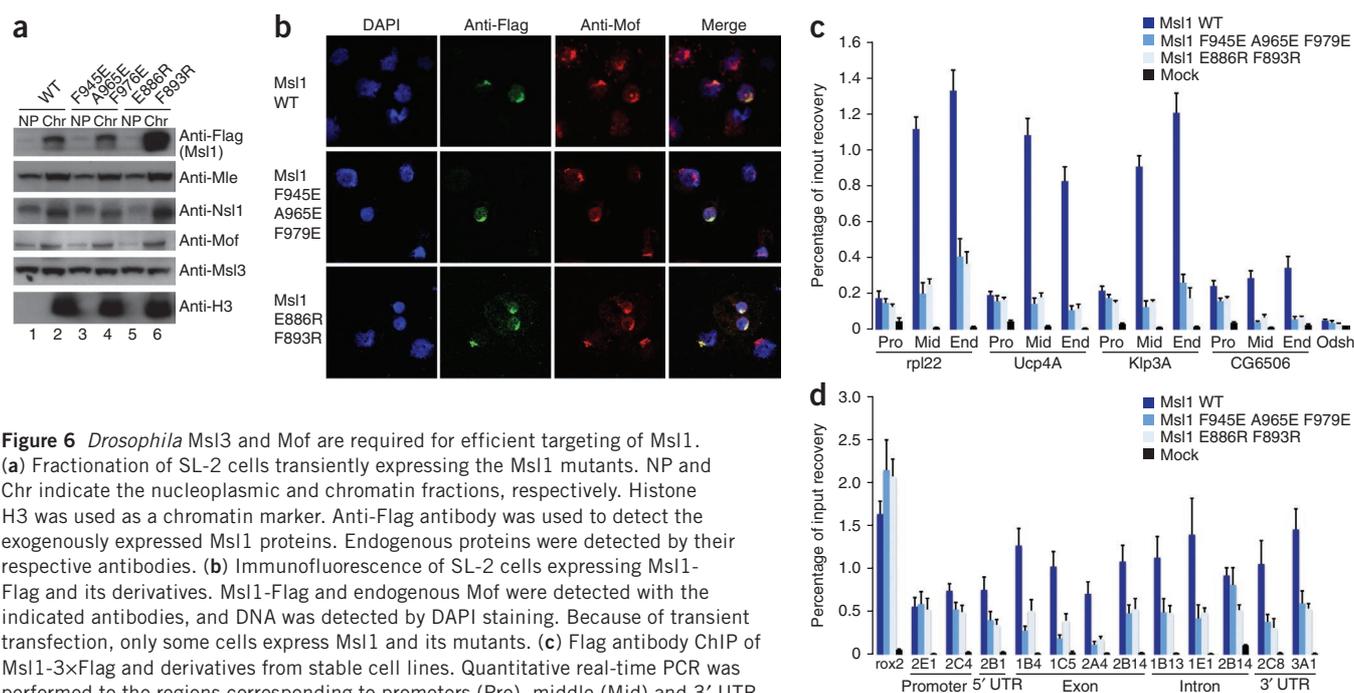


Figure 6 *Drosophila* Msl1 and Mof are required for efficient targeting of Msl1. (a) Fractionation of SL-2 cells transiently expressing the Msl1 mutants. NP and Chr indicate the nucleoplasmic and chromatin fractions, respectively. Histone H3 was used as a chromatin marker. Anti-Flag antibody was used to detect the exogenously expressed Msl1 proteins. Endogenous proteins were detected by their respective antibodies. (b) Immunofluorescence of SL-2 cells expressing Msl1-Flag and its derivatives. Msl1-Flag and endogenous Mof were detected with the indicated antibodies, and DNA was detected by DAPI staining. Because of transient transfection, only some cells express Msl1 and its mutants. (c) Flag antibody ChIP of Msl1-3xFlag and derivatives from stable cell lines. Quantitative real-time PCR was performed to the regions corresponding to promoters (Pro), middle (Mid) and 3' UTR of the genes (End). All tested genes are dosage compensated and located on the X chromosome. ODSH upstream region was used as a negative control for the binding of the MSL complex. Each bar represents the average of four independent immunoprecipitations (IPs), and error bars indicate the s.d. Mock sample corresponds to wild-type cells. (d) Flag antibody ChIP of Msl1-3xFlag and derivatives on high-affinity sites. Thirteen high-affinity sites were chosen from ref. 9 and were grouped according to their genomic positions. The error bars indicate the s.d. of four independent IPs. Note that the two 2B14 sites are different (see **Supplementary Table 1** for the primer pairs used in the study).

lanes 3 and 4), and the interaction was essentially abolished in triple mutants with substitutions that interfere with hydrophobic contacts in the three distinct MSL3 regions (F556E A576E F589E and F556E F577E F589E; **Fig. 5c**, lanes 5 and 6).

To study the integrity of the mutated proteins and the impact of the MSL1 mutations on the incorporation of MOF and MSL3 into the MSL complex *in vivo*, we prepared corresponding mutations in full-length MSL1. Because the MSL proteins are evolutionary conserved and their role is better understood in *Drosophila*, we studied the effect of these mutations using *Drosophila* Msl1 (for the comparison of corresponding amino acid mutations between human and *Drosophila*, see **Fig. 5a**). To assure a complete loss of Msl3 binding, we introduced an additional mutation M973E (L584E in human). We expressed Msl1-Flag proteins in SL-2 cells and immunoprecipitated the corresponding MSL complexes using an anti-Flag antibody-coupled resin. All tested Msl1 mutations in the C-terminal part of the PEHE region were unable to copurify endogenous Msl3, whereas they had no effect on the remaining MSL components, Mof, Msl2 and Mle (**Fig. 5d**, lanes 7–9). In the case of Mof binding-site mutants, a reduced binding of Mof was obtained for the single F893R mutation in the N-terminal portion of the Msl1 PEHE region. Further reduction was observed for the E886R F893R double mutant (**Fig. 5e**, lanes 5 and 6). A partial reduction was also observed for Msl3 and Msl2 incorporation, suggesting that the presence of Mof in the complex might be important for Msl1 stability (**Fig. 5e**). These results indicate that Msl3 and Mof are incorporated into the MSL complex via the Msl1 scaffold and show that at least Msl3 can be disassembled from the complex without any apparent effect on the molecular interactions of other members of the MSL complex.

The fact that the Msl1 F945E A965E F979E triple mutant immunoprecipitates Mof but not Msl3 raises questions about the importance of the interaction between Mof and Msl3 (ref. 28). To further investigate

the putative Mof-Msl3 interaction¹⁴, we performed MSL complex reconstitution assays with the full-length proteins expressed in Sf21 insect cells. In the presence of Msl1, Mof clearly copurifies with Msl3-Flag (**Supplementary Fig. 11d**, lane 4), whereas in the absence of Msl1, the interaction could be seen only using western blot detection (**Supplementary Fig. 11d**, lane 2, and **Supplementary Fig. 11e**, lane 2), suggesting that the Msl3-Mof interaction does occur *in vitro*, albeit substantially more weakly than within the trimeric complex.

NSL1 and MSL1 bind MOF in a similar manner

Notably, the key interacting residues of MSL1 are also conserved in human and *Drosophila* NSL1 (also known as MSLv1 in human) proteins (**Fig. 2e**). Human NSL1 has recently been shown to interact with MOF within the NSL complex, which is involved in acetylation of p53 and H4K16 in male and female cells^{7,18,19}. Indeed, a corresponding fragment of human NSL1 (883–952) is sufficient to form a stable complex with the MOF HAT domain *in vitro* (**Supplementary Fig. 12**). The MSL1 Arg508 forming three hydrogen bonds with MOF (**Fig. 2c,e**) is substituted with Leu920 in human NSL1; however, this arginine remains conserved in *Drosophila* Nsl1. Given the high sequence identity between MSL1 and NSL1 in this region, it is likely that their modes of interaction with MOF are very similar. We did not observe any increase of the MOF HAT activity in the presence of NSL1_{883–952} alone, as had been seen in the presence of MSL1_{470–540} (**Fig. 4a**).

To investigate the interaction between NSL1 and MOF, we prepared mutations E910R, F917R and H921R in human NSL1 and first tested the ability of these mutants to interact with the MOF HAT domain *in vitro*. H921R abolished and E910R considerably reduced the NSL1 binding to MOF (**Fig. 5f** and **Fig. 5g**, lanes 1 and 3). Next, we investigated whether NSL1 uses the same interaction surface for MOF contact *in vivo*. For this purpose, amino acids predicted to interact with Mof

Table 1 Data collection and refinement statistics

	MSL1-MOF-AcCoA	MSL1-MSL3
Data collection		
Space group	$I4_122$	$P2_1$
Cell dimensions		
a, b, c (Å)	180.9, 180.9, 80.7	75.8, 127.1, 79.6
α, β, γ (°)	90, 90, 90	90, 118.4, 90
Resolution (Å)	48–2.7 (2.8–2.7) ^a	46–3.0 (3.12–3.0)
R_{merge}	8.6 (80.7)	3.4 (51.9)
$I/\sigma I$	14.05 (2.16)	16.37 (2.04)
Completeness (%)	99.1 (98.6)	93.5 (94.9)
Redundancy	4.0 (4.1)	2.2 (2.2)
Refinement		
Resolution (Å)	42–2.7	46–3.0
No. reflections	17,569	23,794
$R_{\text{work}} / R_{\text{free}}$	22.3 / 25.6	23.2 / 25.3
No. atoms		
Protein	2,641	6,492
Ligand	50	–
Zinc ion	1	–
Water	10	–
<i>B</i> -factors		
Protein	61	100
Ligand	45	–
Ion	56	–
Water	39	–
R.m.s. deviations		
Bond lengths (Å)	0.006	0.009
Bond angles (°)	0.945	1.118

^aValues in parentheses are for highest-resolution shell.

were mutated in full-length *Drosophila* Nsl1, and the mutant proteins were expressed in SL-2 cells. Remarkably, the *Drosophila* Nsl1 mutant E1264R F1271R showed a severe loss of Mof interaction (Fig. 5h, lane 5), whereas the Mbdr2-Nsl3 interaction was preserved. This indicated that MOF uses similar surfaces for the integration into either NSL or MSL complexes in *Drosophila* cells. We therefore propose that specificity of Mof targeting depends on the differential interactions of Msl1 or Nsl1 with other members of the respective complexes.

Targeting of MSL1 to X-linked genes requires MOF and MSL3

On the basis of our high-resolution structures, we were able to design mutations in Msl1 that selectively disrupt its interaction with Mof or Msl3. We next investigated the impact of such mutations on the recruitment of Msl1 to X-chromosomal target genes. For these *in vivo* experiments, we selected Msl1 F945E A965E F979E, which lacks the interaction with Msl3, and Msl1 E886R F893R, which shows compromised Mof binding. To test whether the wild-type Msl1-Flag and its mutant derivatives are incorporated or targeted to chromatin, we initially performed cell fractionation assays (see **Supplementary Methods**). In wild-type cells, all MSL members have both nucleoplasmic and chromatin distributions, with an enrichment in the chromatin-bound pool (Supplementary Fig. 13a,b). Transiently expressed Msl1-Flag and the Msl1 F945E A965E F979E and E886R F893R mutants were detected mostly in the chromatin fractions, indicating that our constructs were incorporated into chromatin much like endogenous MSL complexes (Fig. 6a). Next, we tested whether X-chromosomal targeting is affected by the disruption of Msl1 interaction with either Mof or Msl3. So as not to exceed physiological protein levels, we expressed the constructs under the control of

the copper-inducible MtnB promoter in uninduced conditions and used an anti-Flag antibody to visualize the exogenous Msl1-Flag and mutant derivatives by immunofluorescence (IF) microscopy. All constructs were correctly targeted to the X chromosome, showing colocalization with endogenous Mof (Fig. 6b).

Because immunofluorescence microscopy does not provide sufficient resolution to observe targeting to individual loci, we performed ChIP of the Msl1 derivatives on X-linked genes. For this purpose, we generated stable SL-2 cell lines that express 3×Flag epitope-tagged Msl1 and confirmed that Msl1-3×Flag showed binding profiles similar to those of the endogenous Msl1 (Supplementary Fig. 13c,d). ChIP experiments revealed that in contrast to the wild-type Msl1, Msl1 E886R F893R and Msl1 F945E A965E F979E showed considerably reduced binding on the body of X-linked genes (Fig. 6c). We also observed a low but consistent Msl1 signal around promoters of X-linked genes. Notably, this signal remained largely unaffected in the Msl1 mutants (Fig. 6c).

We next asked whether the compromised chromatin binding of the Msl1 mutants is restricted to low-affinity sites or whether targeting to high-affinity sites is also impaired. For this purpose, we chose 13 recently mapped high-affinity sites⁹ and compared the binding profiles of wild-type Msl1 and its mutant derivatives to these sites. The roX2 gene was used as a control because it is a high-affinity site for MSL complex assembly and Msl1 binding on this site is independent of Msl3 or Mof²⁹. Because high-affinity sites are located at different loci on the X chromosome, we separated them into positional categories (promoter proximal, 5' untranslated region (UTR), exon, intron and 3' end) in order to investigate any site-specific differences. Notably, this analysis revealed that disruption of Mof or Msl3 interactions also affected optimal binding of Msl1 to these high-affinity sites, especially those located away from promoter regions (Fig. 6d). However, for sites that are promoter proximal, such as 2E1 and 2C4, MSL1 mutants remained bound at comparable levels to the wild-type Msl1 (Fig. 6d).

DISCUSSION

In this study, we report the first crystal structures of two mammalian MSL subcomplexes containing two consecutive fragments of MSL1. These structures reveal how MSL1 uses short interacting peptides for the recruitment of MSL3 and MOF into the dosage-compensation complex. The conserved interactions of MOF with MSL1 and NSL1 explain how MOF can interact with two separate proteins complexes. It is tempting to speculate whether other proteins may also interact with MOF in a similar fashion.

Mutations of key interacting residues in *Drosophila* Msl1 revealed important information regarding its targeting to X-linked genes. In particular, we draw three important conclusions from these data. First, there appears to be a separate mode of Msl1 binding at the promoter region of target genes (regardless of low or high affinity) that seems to be independent of Msl3 or Mof interaction. These observations are noteworthy because Msl1 binding to promoters has not been appreciated previously. It will be interesting to investigate in future how other members of the MSL complex, such as Msl2, may influence targeting to specific chromatin regions in a similar manner.

Second, Msl1 binding to the coding regions of genes, as well as to high-affinity sites, requires efficient interaction with Msl3 or Mof. The latter observation is intriguing because high-affinity sites traditionally have been defined as sites where partial complexes of Msl1-Msl2 can bind in the absence of other components such as Msl3 or Mof^{1,9}. However, our data revealed that optimal Msl1 binding to some high-affinity sites also requires Msl3 or Mof, suggesting that there is a more complex targeting mechanism than has been appreciated to date.

Third, these data reveal that there are qualitative differences among the high-affinity sites and that one possible cause of these variations could be

the positional clue within the chromatin environment. These site-specific chromatin characteristics, such as nucleosome-depleted promoter regions or differential chromatin marks within the coding regions¹, may work in combination with the postulated high-affinity site sequences⁹.

Taken together, our data provide concrete structural evidence of the scaffolding role of Msl1 in the assembly and function of the MSL complex in *Drosophila* and mammals. Moreover, insights gained at atomic resolution provide us with the unique possibility of investigating the importance and mechanism of individual MSL complex members for transcription regulation and dosage compensation. Future structural investigations, including examination of additional components of the MSL complex, promise to broaden our understanding of how this chromatin remodeling machine is targeted to the X chromosome and upregulates transcription by two-fold.

METHODS

Methods and any associated references are available in the online version of the paper at <http://www.nature.com/nsmb/>.

Accession codes. Atomic coordinates and structure factors for the MSL1–MSL3 and MSL1–MOF complexes have been deposited with the Protein Data Bank under accession codes respectively 2Y0N and 2Y0M.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We thank K. Dzyek of EMBL Heidelberg and I. Berard of the Institut de Biologie Structurale (IBS) Grenoble for their help with mass spectrometry; the High Throughput Crystallisation Laboratory (HTX) group of EMBL Grenoble for performing initial screening crystallization trials; M. Jamin's group for help with the MALLS experiment; the European Synchrotron Radiation Facility (ESRF)-EMBL Joint Structural Biology Group for access to and assistance on the ESRF synchrotron beamlines; P. Tropberger from R. Schneider's laboratory (Max Planck Institute of Immunology, Freiburg) for providing nucleosomes; and D. Panne and members of both laboratories for critical reading of the manuscript. E.H. is a Darwin trust fellow.

AUTHOR CONTRIBUTIONS

J.K., E.H., S.C. and A.A. designed the project and wrote the manuscript; J.K. and M.L. produced and crystallized the proteins; J.K. solved, refined and analyzed the structures, prepared and tested the mutants *in vitro*, and carried out HAT assays on the H4 peptide; J.S.-W. and J.K. performed dehydration of MSL1–MSL3 crystals; H.H. performed HAT assays on nucleosomes; and E.H. carried out all *in vivo* experiments.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at <http://www.nature.com/nsmb/>.

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>.

1. Straub, T. & Becker, P.B. Dosage compensation: the beginning and end of generalization. *Nat. Rev. Genet.* **8**, 47–57 (2007).
2. Heard, E. & Disteche, C.M. Dosage compensation in mammals: fine-tuning the expression of the X chromosome. *Genes Dev.* **20**, 1848–1867 (2006).
3. Mendjan, S. & Akhtar, A. The right dose for every sex. *Chromosoma* **116**, 95–106 (2007).
4. Gelbart, M.E. & Kuroda, M.I. *Drosophila* dosage compensation: a complex voyage to the X chromosome. *Development* **136**, 1399–1410 (2009).
5. Smith, E.R. *et al.* A human protein complex homologous to the *Drosophila* MSL complex is responsible for the majority of histone H4 acetylation at lysine 16. *Mol. Cell. Biol.* **25**, 9175–9188 (2005).
6. Marin, I. Evolution of chromatin-remodeling complexes: comparative genomics reveals the ancient origin of “novel” compensasome genes. *J. Mol. Evol.* **56**, 527–539 (2003).
7. Mendjan, S. *et al.* Nuclear pore components are involved in the transcriptional regulation of dosage compensation in *Drosophila*. *Mol. Cell* **21**, 811–823 (2006).
8. Taipale, M. *et al.* hMOF histone acetyltransferase is required for histone H4 lysine 16 acetylation in mammalian cells. *Mol. Cell. Biol.* **25**, 6798–6810 (2005).
9. Alekseyenko, A.A. *et al.* A sequence motif within chromatin entry sites directs MSL establishment on the *Drosophila* X chromosome. *Cell* **134**, 599–609 (2008).
10. Smith, E.R. *et al.* The *Drosophila* MSL complex acetylates histone H4 at lysine 16, a chromatin modification linked to dosage compensation. *Mol. Cell. Biol.* **20**, 312–318 (2000).
11. Akhtar, A. & Becker, P.B. Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in *Drosophila*. *Mol. Cell* **5**, 367–375 (2000).
12. Hilfiker, A., Hilfiker-Kleiner, D., Pannuti, A. & Lucchesi, J.C. mof, a putative acetyltransferase gene related to the Tip60 and MOZ human genes and to the SAS genes of yeast, is required for dosage compensation in *Drosophila*. *EMBO J.* **16**, 2054–2060 (1997).
13. Scott, M.J., Pan, L.L., Cleland, S.B., Knox, A.L. & Heinrich, J. MSL1 plays a central role in assembly of the MSL complex, essential for dosage compensation in *Drosophila*. *EMBO J.* **19**, 144–155 (2000).
14. Morales, V. *et al.* Functional integration of the histone acetyltransferase MOF into the dosage compensation complex. *EMBO J.* **23**, 2258–2268 (2004).
15. Morales, V., Regnard, C., Izzo, A., Vetter, I. & Becker, P.B. The MRG domain mediates the functional integration of MSL3 into the dosage compensation complex. *Mol. Cell. Biol.* **25**, 5947–5954 (2005).
16. Sural, T.H. *et al.* The MSL3 chromodomain directs a key targeting step for dosage compensation of the *Drosophila melanogaster* X chromosome. *Nat. Struct. Mol. Biol.* **15**, 1318–1325 (2008).
17. Akhtar, A., Zink, D. & Becker, P.B. Chromodomains are protein-RNA interaction modules. *Nature* **407**, 405–409 (2000).
18. Li, X., Wu, L., Corsa, C.A., Kunkel, S. & Dou, Y. Two mammalian MOF complexes regulate transcription activation by distinct mechanisms. *Mol. Cell* **36**, 290–301 (2009).
19. Raja, S.J. *et al.* The nonspecific lethal complex is a transcriptional regulator in *Drosophila*. *Mol. Cell* **38**, 827–841 (2010).
20. Yan, Y., Barlev, N.A., Haley, R.H., Berger, S.L. & Marmorstein, R. Crystal structure of yeast Esa1 suggests a unified mechanism for catalysis and substrate binding by histone acetyltransferases. *Mol. Cell* **6**, 1195–1205 (2000).
21. Rojas, J.R. *et al.* Structure of *Tetrahymena* GCN5 bound to coenzyme A and a histone H3 peptide. *Nature* **401**, 93–98 (1999).
22. Wang, L., Tang, Y., Cole, P.A. & Marmorstein, R. Structure and chemistry of the p300/CBP and Rtt109 histone acetyltransferases: implications for histone acetyltransferase evolution and function. *Curr. Opin. Struct. Biol.* **18**, 741–747 (2008).
23. Yan, Y., Harper, S., Speicher, D.W. & Marmorstein, R. The catalytic mechanism of the ESA1 histone acetyltransferase involves a self-acetylated intermediate. *Nat. Struct. Biol.* **9**, 862–869 (2002).
24. Berndsen, C.E., Albaugh, B.N., Tan, S. & Denu, J.M. Catalytic mechanism of a MYST family histone acetyltransferase. *Biochemistry* **46**, 623–629 (2007).
25. Burke, T.W., Cook, J.G., Asano, M. & Nevins, J.R. Replication factors MCM2 and ORC1 interact with the histone acetyltransferase HBO1. *J. Biol. Chem.* **276**, 15397–15408 (2001).
26. Zhang, P. *et al.* The MRG domain of human MRG15 uses a shallow hydrophobic pocket to interact with the N-terminal region of PAM14. *Protein Sci.* **15**, 2423–2434 (2006).
27. Bowman, B.R. *et al.* Multipurpose MRG domain involved in cell senescence and proliferation exhibits structural homology to a DNA-interacting domain. *Structure* **14**, 151–158 (2006).
28. Buscaino, A. *et al.* MOF-regulated acetylation of MSL-3 in the *Drosophila* dosage compensation complex. *Mol. Cell* **11**, 1265–1277 (2003).
29. Kelley, R.L. *et al.* Epigenetic spreading of the *Drosophila* dosage compensation complex from roX RNA genes into flanking chromatin. *Cell* **98**, 513–522 (1999).

ONLINE METHODS

Expression, purification and crystallization. The MSL1–MOF and MSL1–MSL3 complexes were produced by coexpression in bacteria and purified as described in **Supplementary Methods**. The MSL1–MOF crystals grew at 20 °C in a solution containing 0.1 M sodium acetate (pH 5.0) and 1.0 M sodium formate and were cryoprotected with 30% (v/v) glycerol. The MSL1–MSL3 crystals were obtained at 15 mg ml⁻¹ in a solution containing 0.1 M ADA (pH 6.5), 0.1 M Li₂SO₄ and 0.9 M MgSO₄ and initially diffracted to 8-Å resolution. After dehydration using the HC1b humidity control device and freezing in the presence of perfluoropolyether PFO-X125/03 (Lancaster Synthesis), the diffraction limit was extended to 2.8 Å³⁰.

Data collection and structure determination. Crystals of the MOF_{174–458}–MSL1_{470–540} complex belong to the space group *I*₄22 with the unit cell dimensions *a*, *b* = 180.9 Å and *c* = 80.7 Å. The asymmetric unit contains one complex and has a solvent content of 70%. A complete native dataset was collected to a resolution of 2.8 Å on beamline ID14-EH4 at the ESRF (Grenoble, France). The data were processed using XDS³¹. Phases were obtained by molecular replacement using PHASER³² with the deposited structure of the human MOF HAT domain (PDB code: 2GIV) as a search model. The initial map was improved using the prime-and-switch density modification option of RESOLVE³³. After manual model rebuilding with COOT³⁴, the structure was refined using Refmac5 (with TLS refinement)³⁵ to a final *R*-factor of 22.3% and *R*_{free} of 25.6% (**Table 1**) with all residues in allowed (97.7% in favored) regions of the Ramachandran plot, as analyzed by MOLPROBITY³⁶. A representative part of the 2*F*_o – *F*_c electron density map calculated using the refined model is shown in **Supplementary Figure 14a**.

Crystals of the MSL3_{167–289,442–517}–MSL1_{545–597} complex belong to the space group *P*2₁ with four complexes per asymmetric unit. A complete native dataset was collected to a resolution of 3.0 Å on beamline ID14-EH2 at the ESRF. The data was processed using XDS³¹. The structure was solved by molecular replacement with PHASER³² using the structure of the human MRG15 MRG domain (PDB code: 2F5J²⁶) as a search model. Probably because of weak intensity of reflections above 3.5 Å, the obtained electron density appeared to be of a lower resolution than 3 Å. *B*-factor sharpening was therefore used to improve map quality³⁷. The structure was built in COOT³⁴ and refined with Refmac5 (using TLS refinement)³⁵ to a final *R*-factor of 23.2% and *R*_{free} of 25.3% (**Table 1**) with 99.87% of residues in allowed (95.9% in favored) regions of the Ramachandran plot, as analyzed by MOLPROBITY³⁶. A representative part of the –60 Å² *B*-factor–sharpened 2*F*_o – *F*_c electron density map calculated using the refined model is shown in **Supplementary Figure 14b–d**. The Wilson *B*-value for this dataset was determined to be 107 Å² using a maximum likelihood–based method as implemented in PHENIX³⁸. Accordingly, the mean value for the isotropic individual *B*-factor for the final model is 100 Å². Interpretable electron density is observed for residues 551–558 and 564–594 of MSL1. In the crystal, the N- and C-terminal parts of MSL1 bind to two distinct MSL3 molecules, a result that we interpret as a result of domain swapping. It is unclear whether the MSL1 N terminus is swapped between two or four molecules (see **Supplementary Fig. 10a**). Buried surface areas of protein–protein interactions were calculated using the PISA web server at the European Bioinformatics Institute (http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html)³⁹.

His-tag pulldown assays. The MSL1 and NSL1 and their mutated versions were coexpressed with MSL3 or MOF in bacteria, and the resultant complexes were purified using Ni²⁺ resin; see **Supplementary Methods**.

Thermal denaturation assay. Thermal denaturation (thermal shift) assays were performed as described in **Supplementary Methods**.

Expression of Msl1 and Nsl1 mutants in *Drosophila* SL-2 cells. *Drosophila* Msl1 was expressed as a C-terminal Flag fusion in the pAc5.1/V5-His A vector. Nsl1 was expressed as an N-terminal Flag fusion in a modified pBSactTAP vector. Transient transfection of SL-2 cells was done with Qiagen Effectene Transfection Reagent. Details are given in **Supplementary Methods**.

HAT assays. The activity of human MSL subcomplexes and MOF mutants in the active site was assessed within the MOF–MSL1–MSL3 complex in DTNB (5,5'-dithio-bis(2-nitrobenzoic acid) HAT assays. The activity of human and *Drosophila* MSL subcomplexes was also assayed using native mono- and dinucleosomes obtained from MCF-7 cells. Further details can be found in **Supplementary Methods**.

In vitro reconstitution assay. The reconstitution of *Drosophila* MSL subcomplexes produced in insect cells is described in **Supplementary Methods**.

Fractionation of SL-2 cells. SL-2 cell fractionation was based on swelling the cells in hypotonic buffer and vortexing in the presence of mild detergent. The supernatant is cytoplasmic fraction and the pellet is nuclei. The nucleoplasmic fraction was obtained by salt extraction, and the remaining chromatin was solubilized by benzonase treatment. See **Supplementary Methods** for details.

Generation of SL-2 stable cells lines. The cells were transfected with 0.5 µg of DNA with Qiagen Effectene Transfection Reagent, and selection was carried out with 1 mg ml⁻¹ Geneticin for 2 weeks. Details are given in **Supplementary Methods**.

Immunofluorescence for SL-2 cells. SL-2 cells were swollen in 500 µl 0.5% sodium citrate for 7 min and loaded through a single-chamber Cytospin tunnel. The cells were spun for 10 min at 900 r.p.m. in a Cytospin (Thermo Shandon) and visualized with a 63× objective. See **Supplementary Methods** for details.

ChIP protocol from SL-2 cells. ChIP protocol was carried out as in ref. 19 with minor modifications. See **Supplementary Methods** for a detailed protocol.

30. Sanchez-Weatherby, J. *et al.* Improving diffraction by humidity control: a novel device compatible with X-ray beamlines. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 1237–1246 (2009).
31. Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* **26**, 795–800 (1993).
32. McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C. & Read, R.J. Likelihood-enhanced fast translation functions. *Acta Crystallogr. D Biol. Crystallogr.* **61**, 458–464 (2005).
33. Terwilliger, T.C. Maximum-likelihood density modification. *Acta Crystallogr. D Biol. Crystallogr.* **56**, 965–972 (2000).
34. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
35. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 (1997).
36. Davis, I.W., Murray, L.W., Richardson, J.S. & Richardson, D.C. MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res.* **32**, W615–W619 (2004).
37. Brunger, A.T., DeLaBarre, B., Davies, J.M. & Weiss, W.I. X-ray structure determination at low resolution. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 128–133 (2009).
38. Adams, P.D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1948–1954 (2002).
39. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).