nature communications

Article

https://doi.org/10.1038/s41467-024-54548-9

Structural basis for *C. elegans* pairing center DNA binding specificity by the ZIM/HIM-8 family proteins

Received: 20 May 2024

Accepted: 13 November 2024

Published online: 28 November 2024

Check for updates

Meili Li[1], Chengming Zhu[1], Zheng Xu[2], Mingjing Xu[1], Yan Kuang[1], Xinhao Hou[1], Xinya Huang[1], Mengqi Lv[1], Yongrui Liu[1], Yong Zhang[1], Ziyan Xu[1], Xu Han[1], Suman Wang[1], Yunyu Shi[1] ✉, Shouhong Guang[1] ✉ & Fudong Li[1] ✉

Pairing center (PC) on each chromosome of *Caenorhabditis elegans* is crucial for homolog pairing and initiating synapsis. Within each PC, clusters of 11/12 bp DNA motif recruit one of four paralogous meiosis-specific proteins: ZIM-1, ZIM-2, ZIM-3, or HIM-8. However, the mechanistic basis underlying the specificity of ZIM/HIM-8-PC DNA interaction remains elusive. Here, we describe crystal structures of HIM-8, ZIM-1 and ZIM-2 DNA binding domains (ZF1, ZF2 and CTD) in complex with their cognate PC DNA motifs, respectively. These structures demonstrated the ZF1-2-CTD folds as an integrated structural unit crucial for its DNA binding specificity. Base-specific DNA-contacting residues are exclusively distributed on ZF1-2 and highly conserved. Furthermore, the CTD potentially contributes to the conformational diversity of ZF1-2, imparting binding specificity to distinct PC DNA motifs. These findings shed light on the mechanism governing PC DNA motif recognition by ZIM/HIM-8 proteins, suggesting a co-evolution relationship between PC DNA motifs and ZF1-2-CTD in shaping the specific recognition.

Meiosis is a specialized and essential cell division program that takes place in the germline cells of sexually reproducing eukaryotes. It involves one round of chromosome duplication and two successive cell divisions, commonly referred to as meiosis I and II, leading to the production of haploid gametocytes. In order to ensure faithful reductional segregation, homologous chromosomes need to first locate and recognize each other in a process known as homolog pairing. Despite its critical importance, the mechanisms enabling homologs to recognize each other and establish physical contact remain insufficiently understood[1–4].

It is recognized that rapid prophase chromosome movements (RPMs), which occur during prophase of meiosis I, facilitate homolog pairing. In most organisms (including fungi, plants, and mammals), these RPMs are driven by the telomeres[5]. Meiosis-specific telomere binding proteins link the telomeres to the transmembrane LINC-complex during meiosis, mediating extensive chromosome motions by transferring cytoskeletal forces to the chromosomes. This motion promotes the formation of a meiosis-specific chromosome configuration called "bouquet"[6–9]. The bouquet facilitates homologous recognition and alignment by localizing chromosomes within a limited region of the nuclear volume[10].

However, in the case of *C. elegans*, bouquet formation does not occur, and telomeres do not contribute to RPMs. Instead, short-spaced repeat sequences near one end of each chromosome, known as pairing centers (PCs) or homolog recognition regions (HRRs), mediate chromosome tethering to the nuclear envelope (NE)[11–13]. PCs facilitate pairing and synaptonemal complex assembly, relying on a family of four paralogous proteins, Zinc finger In Meiosis-1 (ZIM-1), ZIM-2, ZIM-3

[1]MOE Key Laboratory for Cellular Dynamics, Hefei National Laboratory for Physical Sciences at the Microscale, The First Affiliated Hospital of USTC, Biomedical Sciences and Health Laboratory of Anhui Province, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, China. [2]Division of Nephrology and Kidney Research Institute, State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, Chengdu, China. ✉e-mail: yyshi@ustc.edu.cn; sguang@ustc.edu.cn; lifudong@ustc.edu.cn

and High Incidence of Males-8 (HIM-8). These proteins are recruited to their respective chromosome's PC region by a conserved 11/12-bp DNA sequence motif clustered within these areas. Specifically, HIM-8 targets the 12-bp motif TTGGTCAGTGCA found in the X chromosome pairing center. ZIM-1 recognizes the 11-bp motif TTGGTCTGCTA on chromosomes II and III, ZIM-2 binds to the 12-bp motif TTGGGCGCTGCT on chromosome V, and ZIM-3 interacts with the 12-bp motif TTGGTTGAGGCT on chromosomes I and IV (Fig. 1a). ZIM/HIM-8 specifically bind to their corresponding chromosome's PC motifs via their two tandem atypical C2H2 zinc fingers (ZF1-2) and adjacent carboxyl-terminal domain (CTD)[13,14]. The accurate recognition of target PC motifs by ZIM/HIM-8 proteins is crucial for proper homolog pairing and segregation[15]. Interrupting the X chromosome PC recognition domain of HIM-8 specifically impaired the segregation of the X chromosomes[12]. Despite a similarity of over 30% in the ZF1-2-CTD domain among ZIM/HIM-8 proteins and a likeness between the 11-bp or 12-bp pairing center motifs they recognize, these four proteins are able to precisely interact with their corresponding PC motifs. This raises intriguing questions regarding the mechanism governing this selectivity.

In this study, we solved the structures of HIM-8, ZIM-1, and ZIM-2 ZF1-2-CTD in complex with their cognate PC DNA motifs. Our structural analysis revealed that ZF1 and ZF2 engage in base-specific contacts deviating from the conventional C2H2 zinc finger paradigm. Notably, the residues responsible for specific base recognition are localized within ZF1-2 and exhibit high conservation across the ZIM/HIM-8 family proteins. Further investigation demonstrated that, although the CTD does not directly participate in DNA base-specific recognition and interacts solely with the phosphate backbone, it is crucial for DNA-binding specificity by forming a cohesive structural unit with ZF1-2. Analysis of the PC motifs revealed that a conserved TG sub-site functions as a registration motif while intervening space between TTGG and TG sub-sites constraining ZIM/HIM-8 binding. These findings shed light on the mechanism by which these paralogous proteins recognize their respective PCs and provide insights into the co-evolution of DNA sequences and DNA-binding proteins to achieve binding specificities.

Results

The HIM-8 ZF1-2-CTD domain exhibits preferential binding to X chromosome PC motif

Previous research, employing computational analysis, fluorescence in situ hybridization (FISH), and Systematic Evolution of Ligands by EXponential enrichment (SELEX) assays, has defined short DNA sequence motifs enriched in PC from each *C. elegans* chromosome which specifically recruit the ZIM/HIM-8 proteins[11–16]. The HIM-8 protein has been found to bind the double-stranded TTGGTCAGTGCA consensus motif, predominantly present in the X chromosome pairing center[14]. To further characterize the in vivo binding behavior of HIM-8, we generated a *him-8::gfp::3*×*flag* worm (Supplementary Data 1 and 2). Two distinct HIM-8 foci in each nucleus were present within the premeiotic region, one or two foci were present in the transition zone, and a single focus was present during pachytene (Fig. 1b), consistent with previous studies[12,17]. The genome-wide chromatin immunoprecipitation experiments coupled with next-generation sequencing (ChIP-seq) profiles indicated a significant enrichment of the HIM-8 protein in the PC region of the X chromosome (Fig. 1c). Further Multiple Em for Motif Elicitation (MEME) analysis yielded a 12-bp consensus motif similar to the previously established X chromosome PC motif (Fig. 1d).

ZIM/HIM-8 proteins contain two atypical C2H2 zinc fingers (ZF1-2) and an adjacent carboxyl-terminal domain (CTD) (Fig. 1e). Both ZF1-2 and CTD in HIM-8 are crucial for the recognition to X chromosome PC motif[14]. To confirm the binding preference of HIM-8 to the X chromosome PC motif, we expressed HIM-8 ZF1-2 domain (amino acids 223-303) and ZF1-2-CTD domain (amino acids 223-361) in *E. coli* cells (Supplementary Data 1 and 3), purified the recombinant protein and performed isothermal titration calorimetry (ITC) experiments[18]. The ITC results showed that ZF1-2-CTD, but not ZF1-2, bound to the 12-bp oligonucleotide with a dissociation constant (K_D) of 0.43 ± 0.01 μM and an *N* value of ~1, supporting the requirement of CTD domain in PC motif binding (Fig. 1f; Supplementary Table 1; Supplementary Data 4). Furthermore, titrating HIM-8 ZF1-2-CTD with pairing center motifs from other chromosomes exhibited minimal heat change (Fig. 1g; Supplementary Table 1; Supplementary Data 4). In summary, these results confirmed HIM-8 ZF1-2-CTD's preferential binding to the X chromosome PC motif.

General features of the HIM-8 ZF1-2-CTD domain

To better understand the molecular mechanism underlying X chromosome PC motif selection by HIM-8, we co-crystallized its ZF1-2-CTD and the double-stranded 12-bp X chromosome PC motif oligonucleotide. The protein-DNA complex crystallized in the *C*2 space group and the structure was determined to a resolution of 2.14 Å. In each crystallographic asymmetric unit, there is one HIM-8 ZF1-2-CTD-DNA complex (Fig. 2a, Supplementary Table 2). The electron density is well defined for HIM-8 ZF1-2-CTD and nucleotides of the DNA (Supplementary Fig. 1a, b), with the exception of the N-terminal 13 and C-terminal 26 residues of HIM-8 ZF1-2-CTD, which is likely due to flexibility.

There were several differences between HIM-8 ZF1-2-CTD and canonical C2H2 ZF proteins. In canonical C2H2 zinc finger protein, the amino acid sequence of each ZF finger is deciphered by a regular expression: X_2-Cys-$X_{2,4}$-Cys-X_{12}-His-$X_{3,4,5}$-His (X stands for any of the 20 natural amino acids)[19]. Notably, HIM-8 ZF1 possesses a longer spacer between the Zn-coordinating Cys and His residue than the conventional C2H2 zinc finger in proteins such as Zif268 (a mouse immediate early protein) and TZAP (a human telomere binding protein stimulating telomere trimming) (Fig. 2b)[20,21]. The Cys-His space of HIM-8 ZF1 spans amino acids 248-266, which leads to an additional short α helix (referred to as η1) between β2 and the conventional recognition α helix (Fig. 2b, c). A superimposition of HIM-8 ZF1 and TZAP ZF11 using TM-align showed a root-mean-square deviation (RMSD) of 1.84 Å and a TM-score of 0.47, indicating their structural diversities (Fig. 2d)[22]. Intriguingly, both η1 and the recognition α-helix establish base-specific contacts within the major groove, facilitating ZF1 in recognizing a greater number of DNA base pairs compared to canonical ZF.

In contrast, the spacing between Zn-coordinating Cys and His of ZF2 is equivalent to that of a canonical C2H2 finger which leads ZF2 to adopt a conventional ββα conformation (Fig. 2b, c)[19]. A superimposition of HIM-8 ZF2 and TZAP ZF11 using TM-align showed an RMSD of 1.30 Å and a TM-score of 0.74, indicating structural similarities between the two ZF domains despite that they only share 21.4% sequence identities (Fig. 2e).

Interestingly, the CTD (aa 308-335) did not make base-specific contact with the DNA directly. Instead, it folds back and packs against ZF1 and ZF2, to form an integrated structural unit (Fig. 2a). The CTD domain and ZF1-2 associate with each other through both hydrophobic interactions and hydrogen bonds (Fig. 2f), enclosing a total surface area of 2112.1 $Å^2$, as provided by PISA[23]. The first stretch (aa 302-307) of CTD folds back along the DNA phosphate backbone and packs against the ZF2 surface. The stretch is followed by a linker region that bridges the ZF1-CTD-C and ZF2-CTD-N modules. After the linker is the CTD-C portion, containing an α helix and a 3_{10} helix (Fig. 2c). This portion packs against ZF1 through its hydrophobic surface. The contacts between ZF1-2 and CTD include several highly conserved residues: I241, Q242, C243, H244, F245, K260, L263, I264, A267, L268 and H270 at ZF1; L276, M291, R292, Y295 and H298 at ZF2; K302, G306, F307, V309, F319, I322, C326, F327 and Q330 at CTD (Supplementary Fig. 2). This indicates that the ZIM/HIM-8 members may adopt similar overall structure.

https://doi.org/10.1038/s41467-024-54548-9

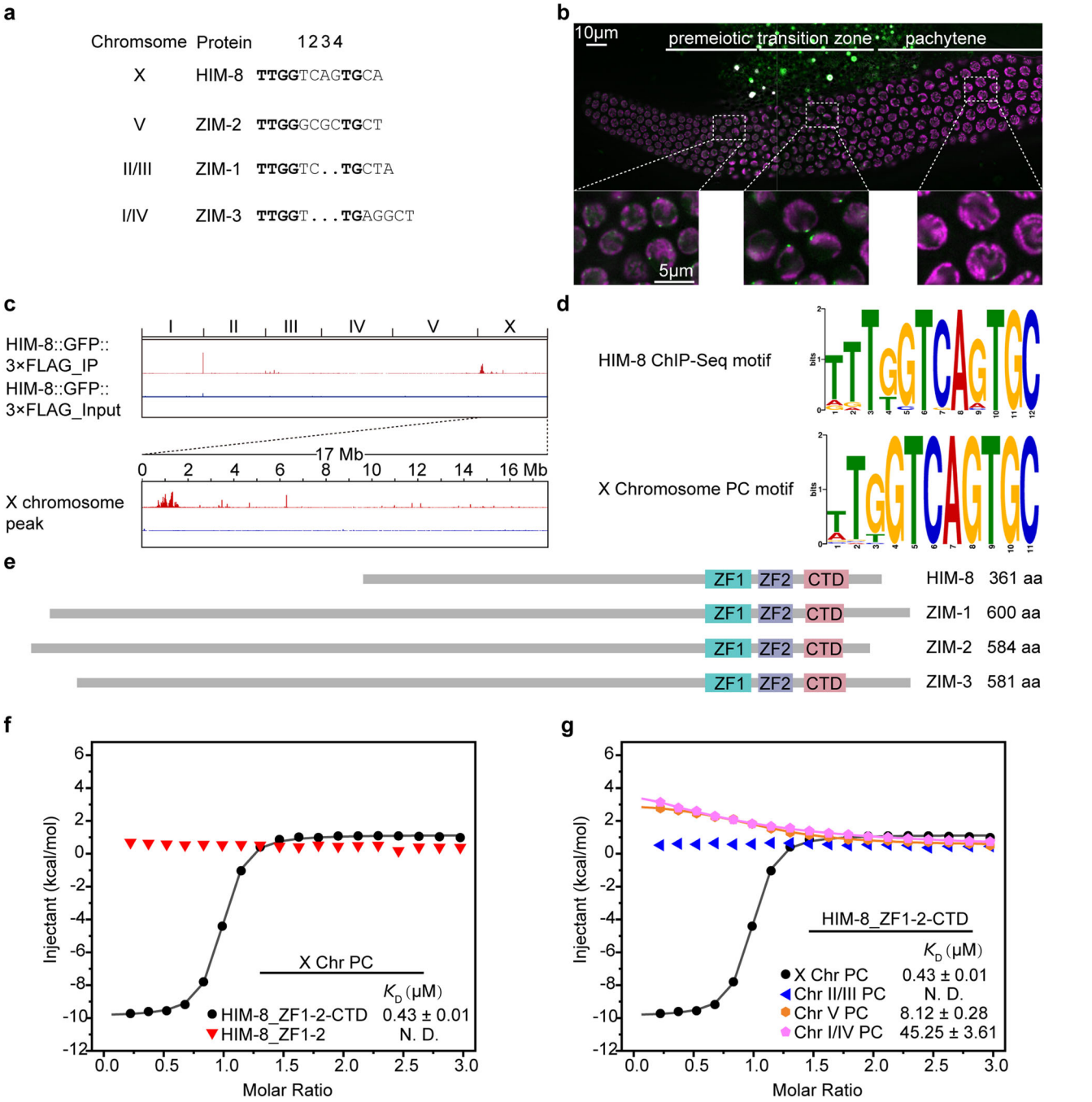


Fig. 1 | Preferential Binding of HIM-8 to the X Chromosome PC DNA. a The sequence alignment of the four PC motifs, based on the ZF1-2-CTD-DNA structures, highlights a major difference among them: the sequence and length between the TTGG and TG sub-sites diverge. **b** Representative images of nuclei expressing HIM-8::GFP and mCherry::H2B during Prophase I progression are presented. In pre-meiotic germ cells, each nucleus exhibits two HIM-8 foci (green). Upon entry into the transition zone (leptotene/zygotene stages), one or two closed HIM-8 foci are observed, typically condensing into a single focus during pachytene. Scale bar: 5 μm. Three times were repeated independently with similar results. **c** ChIP-seq analysis depicts the binding profile of HIM-8::GFP across all six chromosomes in animals at the young adult stage. The binding profile of HIM-8::GFP across X chromosome was zoomed. **d** Utilizing the online MEME suite, a 12-bp consensus binding motif of HIM-8::GFP was identified (top panel), closely resembling the previously established X chromosome PC motif (bottom panel)[14]. **e** Schematic diagram illustrating the two zinc fingers (ZF1 and ZF2) and the carboxyl-terminal domain (CTD) of the ZIM/HIM-8 proteins. **f** Isothermal titration calorimetry (ITC) data demonstrate high-affinity binding of HIM-8 ZF1-2-CTD (depicted in black) to the 12-bp X chromosome PC duplex sequence (TTGGTCAGTGCA), while HIM-8 ZF1-2 (in red) does not exhibit this affinity. "Chromosome" was abbreviated as "Chr", "pairing center" was abbreviated as "PC". K_D, dissociation constant. Data are presented as Mean ± SD. The thermodynamic parameters are presented in Supplementary Table 1. All ITC binding curves are shown in Supplementary Data 4. **g** ITC results reveal the preferential binding of HIM-8 ZF1-2-CTD (black) to the 12-bp X chromosome PC motif over autosomal PC motifs (chr II/III: TTGGTCTGCTA, chr V: TTGGGCGCTGCT, and chr I/IV: TTGGTTGAGGCT).

https://doi.org/10.1038/s41467-024-54548-9

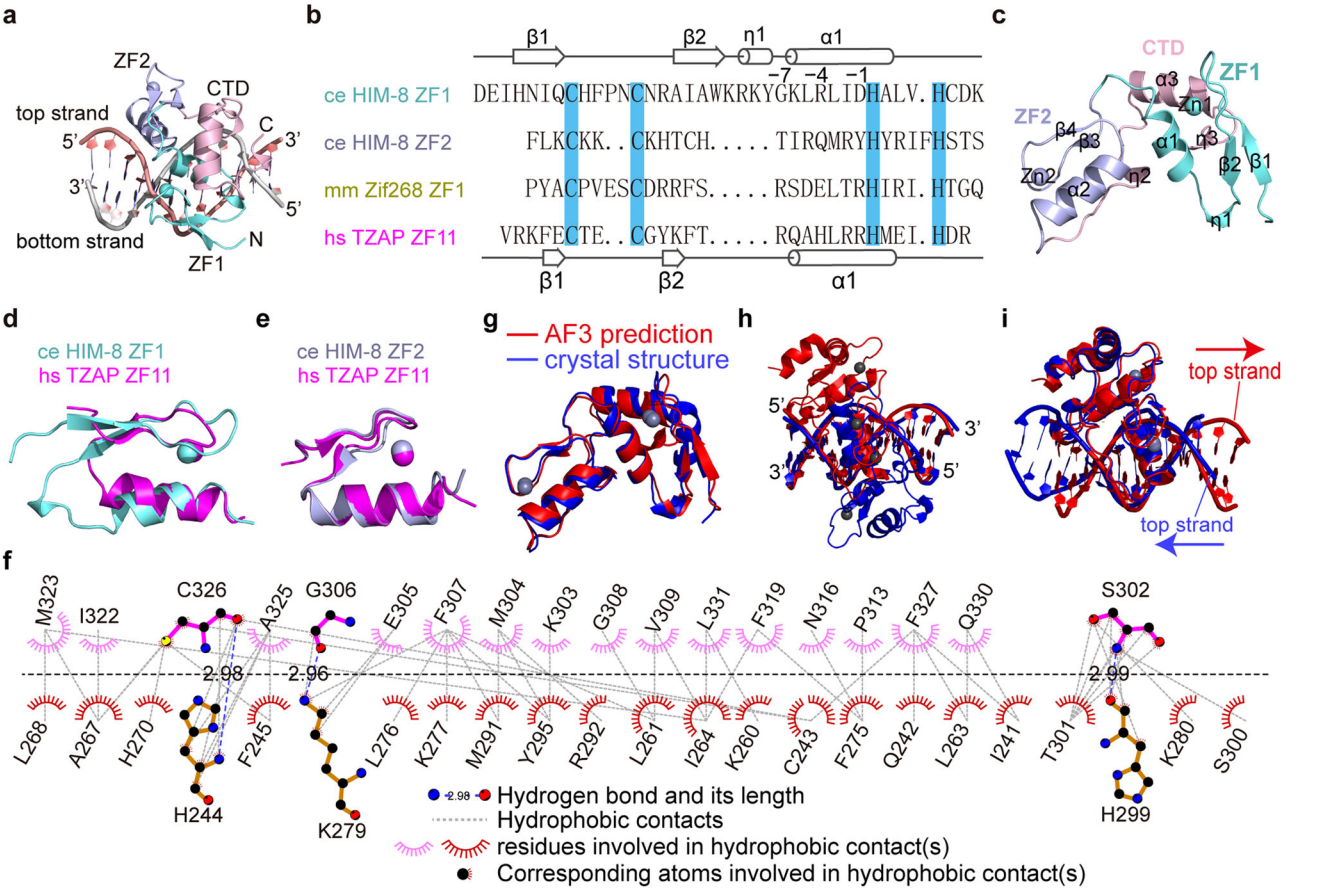


Fig. 2 | Overall structure of HIM-8 ZF1-2-CTD in complex with X chromosome PC DNA. a Overview of the ZF1-2-CTD complex of HIM-8 in association with the 12-bp X chromosome PC motif. ZF1, ZF2, and CTD are depicted in aquamarine, light blue, and light pink, respectively. The top and bottom strands of the motif are colored salmon and gray, respectively. **b** Sequence alignment of HIM-8 ZF1, HIM-8 ZF2, Zif268 ZF1, and TZAP ZF11. Secondary structural elements of HIM-8 ZF1 (top panel) and the other three (bottom panel) are presented. Cysteine and histidine residues crucial for zinc atom binding are highlighted in cyan. Three residues in classical C2H2 zinc fingers that interact specifically with DNA bases are marked as positions −1, −4, and −7. "*C.elegans*" was abbreviated as "Ce", "mouse" was abbreviated as "mm", "human" was abbreviated as "hs". **c** Schematic representation illustrating the HIM-8 ZF1-2-CTD structure. ZF1 comprises a ββηα structure, ZF2 includes a ββα configuration, and CTD contains an ηαη arrangement. CTD folds back and tightly associates with ZF1 and ZF2. **d** Structural comparison between HIM-8 ZF1 and TZAP ZF11 using TM-align, resulting in a root-mean-square deviation (RMSD) of 1.84 Å and a TM-score of 0.47. **e** Structural alignment of HIM-8 ZF2 and TZAP ZF11 using TM-align, resulting in an RMSD of 1.30 Å and a TM-score of 0.74. **f** Hydrophobic contacts and hydrogen-bond interactions at the ZF1-2 and CTD interface are illustrated. This plot was generated using LigPlot+ v. 2.2.8. Residues above the black dashed line belong to the CTD, while those below are part of the ZF1-2 domain. **g** A comparison of the HIM-8 ZF1-2-CTD crystal structure (blue) with its predicted counterpart from AlphaFold 3 (red) reveals an RMSD of 1.048 Å. **h**, **i** The crystal structure of the HIM-8 ZF1-2-CTD complexed with the 12-bp X chromosome PC dsDNA motif (blue) is compared to the predicted protein-DNA complex structure from AF3 (red) via the DNA alignment (**h**) and via the protein alignment (**i**), respectively.

We utilized the AlphaFold 3 (AF3) server to predict the structure of the apo HIM-8 ZF1-2-CTD (Supplementary Fig. 3a, b)[24]. Superimposing this predicted model with our experimentally determined structure resulted in a root-mean square deviation (RMSD) of 1.048 Å, reflecting a high degree of similarity in overall folding and thereby confirming the accuracy of the prediction (Fig. 2g). Further, we generated the structure of the HIM-8 ZF1-2-CTD complexed with the 12-bp X chromosome PC motif using the AF3 server. When we compared the AF3 predictive model (Supplementary Fig. 3c, d) with our structure via DNA duplex alignment, we observed a notable discrepancy in the positioning of the protein relative to the DNA (Fig. 2h). Further superimposition of the HIM-8 ZF1-2-CTD segment from the AF3 model with our structure shows that the orientations of the bound DNA molecules differ significantly (Fig. 2i). This pronounced deviation underscores the AF3 model's limitation in accurately forecasting the spatial arrangement of the protein-DNA complex in this scenario.

Collectively, the HIM-8 ZF1-2-CTD folds as an integrated structural unit, with the CTD domain not forming base-specific contact with the PC motif.

HIM-8 ZF1-2 convey base-specific recognition to X chromosome PC motif

When bound to DNA, the HIM-8 ZF1-2-CTD fits into the DNA major groove with their α helixes oriented toward the DNA (Fig. 2a). Conventionally, C2H2 zinc fingers recognize three consecutive DNA base pairs using the amino acids located at the −1, 3, and 6 positions of the recognition α helix[19]. This rule of one finger-three base recognition does not apply to HIM-8 ZF1 due to the addition of η1. To minimize potential ambiguity, we followed another numbering scheme firstly provided by the Xiaodong Cheng group which used the first Zn-coordination His in each finger as the reference position 0, with residues at sequence positions −1, −4, and −7 corresponding to the structural numbering of 6, 3, and −1, respectively (Fig. 2b)[25]. The nucleotide numbering is demonstrated in Fig. 3a. HIM-8 ZF1-2 runs antiparallel with the top DNA strand and makes base-specific contacts with both strands, with ZF1 binding near the 3′ end and ZF2 binding near the 5′ end (Fig. 3a).

ZF1 has G259, R262, and D265 at base-interacting positions (−7, −4, and −1) of conventional C2H2 zinc finger (Fig. 2b). However, only

https://doi.org/10.1038/s41467-024-54548-9

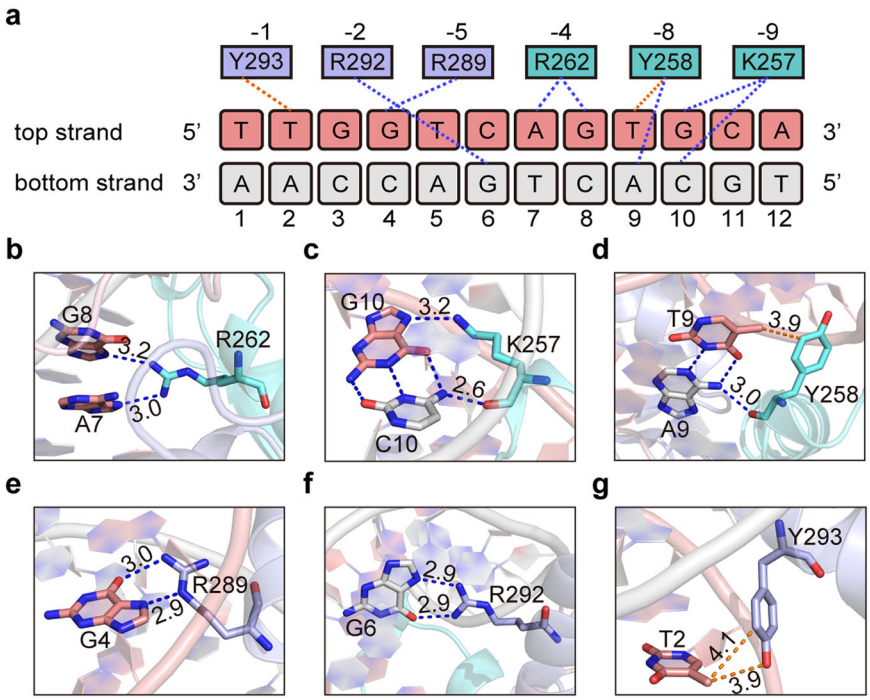


Fig. 3 | Sequence-specific DNA Binding by HIM-8 ZF1-2-CTD. a General schematic diagram of detailed base-specific interactions between HIM-8 ZF1-2-CTD and the 12-bp X chromosome PC DNA sequence. Color codes of ZF1, ZF2, top strand, and bottom strand elements are defined as in Fig. 2a. Direct hydrogen bonds (blue) and Van der Waals contacts (orange) are depicted with dashed lines. **b**–**g** Base-specific contacts down to the atom are highlighted. Interatomic distances are measured in angstroms.

R262 at position −4 is involved in base-specific contact. R262 forms hydrogen bonds with both the A at base pair 7 and the G at base pair 8 (Fig. 3a, b). Interestingly, D265 at position −1 does not contact with any base but meditates the intra-molecular interactions with ZF2, by forming hydrogen bonds with the main chain of R289 in ZF2, which may help to stabilize the ZF1-2-CTD overall structure (Supplementary Fig. 4a). Since G259 lacks a side chain, it does not make contacts with any base.

The additional η1 in ZF1 is involved in base-specific interactions with base pairs 9 and 10. K257 at position −9 forms hydrogen bonds with the G and the C at base pair 10 through its side chain and main chain, respectively (Fig. 3a, c). Y258 at position −8 makes Van der Waals contacts with the T at base pair 9 and forms a hydrogen bond with the A at base pair 9 through its main chain (Fig. 3a, d). These interactions provide highly specific recognition of base pairs 9 and 10.

ZF2 makes base-specific contacts with DNA using amino acids occupying positions −5, −2, and −1 (Fig. 3a), but not at positions −7 and −4. R289 at position −5 forms hydrogen bonds with the G at base pair 4. Its terminal N atom donates an H-bond to the guanine O6 atom (Fig. 3a, e). Two terminal N atoms of R292 at position −2 donate H-bonds to the O6 and N7 atoms of the guanine at base pair 6, forming a bidentate base-specific interaction to the guanine, a pattern specific to guanine (Fig. 3a, f). Thereby, HIM-8 ZF2 employs two arginines to recognize two guanines (at base pair 4 and 6) in the 5′ portion of the HIM-8 PC motif. Additionally, Y293 at position −1 makes Van der Waals contacts to the T at base pair 2 (Fig. 3a, g).

In addition to these base contacts, DNA binding involves the interaction of the DNA backbone with residues A253, K260, R289, Q290, Y293, G306, F307, V309, and S310 (Supplementary Fig. 4b–k). These residues distribute across ZF1, ZF2, and CTD, together enhancing the binding affinities to the X chromosome PC motif.

In summary, HIM-8 binding to the X chromosome PC motif involves both base contacts and backbone interactions, and the base-specific interaction mode is obviously deviating from one finger-three base mode of conventional C2H2 zinc finger[26,27], revealing DNA recognition diversities of C2H2 zinc finger proteins.

HIM-8 mutants reduced the binding affinities to the X chromosome PC motif and increased male production

The significance of the identified hydrogen bonding and Van der Waals interactions in the binding of HIM-8 ZF1-2-CTD to the X chromosome PC motif was validated through ITC experiments. Mutations in base-contacting residues within HIM-8 ZF1-2-CTD, including K257A, Y258A, R262A, R289A, R292A, or Y293A, significantly diminished affinities for the X chromosome PC motif (Fig. 4a, b; Supplementary Table 1; Supplementary Data 3 and 4).

To investigate the in vivo functional significance of the interaction between HIM-8 ZF1-2-CTD and the X chromosome PC motif, we attempted to generate HIM-8 mutant worms based on the *him-8::gfp::3 × flag* worm model. Mutations K257A, Y258A, R262A, R289A, R292A, or Y293A were introduced respectively using oligonucleotide-guided CRISPR/Cas9 gene editing technology (Fig. 4c, Supplementary Data 1 and 2). Subsequently, two mutant worm strains, R289A and Y293A, were successfully created (Fig. 4d). The remaining mutant strains failed to exhibit observable GFP signals, and subsequent qPCR and Western Blot experiments confirmed the absence of HIM-8 mutant expression in these worms (Supplementary Fig. 5a–c; Supplementary Data 1). An *him-8* knockout allele (*him-8 (−/−)*), which deletes the sequence from amino acids 14 to 361 of HIM-8, served as a null allele for comparative analysis (Supplementary Data 1 and 2).

In the HIM-8_R289A::GFP::3 × FLAG worm—a mutant displaying an 8-fold reduction in X chromosome PC motif binding affinity (Fig. 4b)—two distinct HIM-8 foci remained visible in some prophase I nuclei (Fig. 4d). To quantify the presence of these foci during the pachytene stage, we utilized the SYP-1 protein as a marker which is a core component of the synaptonemal complex and is primarily expressed during this stage (Supplementary Fig. 6a, b)[12]. In the HIM-8::GFP::3 × FLAG;tagRFP::SYP-1 strain, 6.83% of pachytene

https://doi.org/10.1038/s41467-024-54548-9

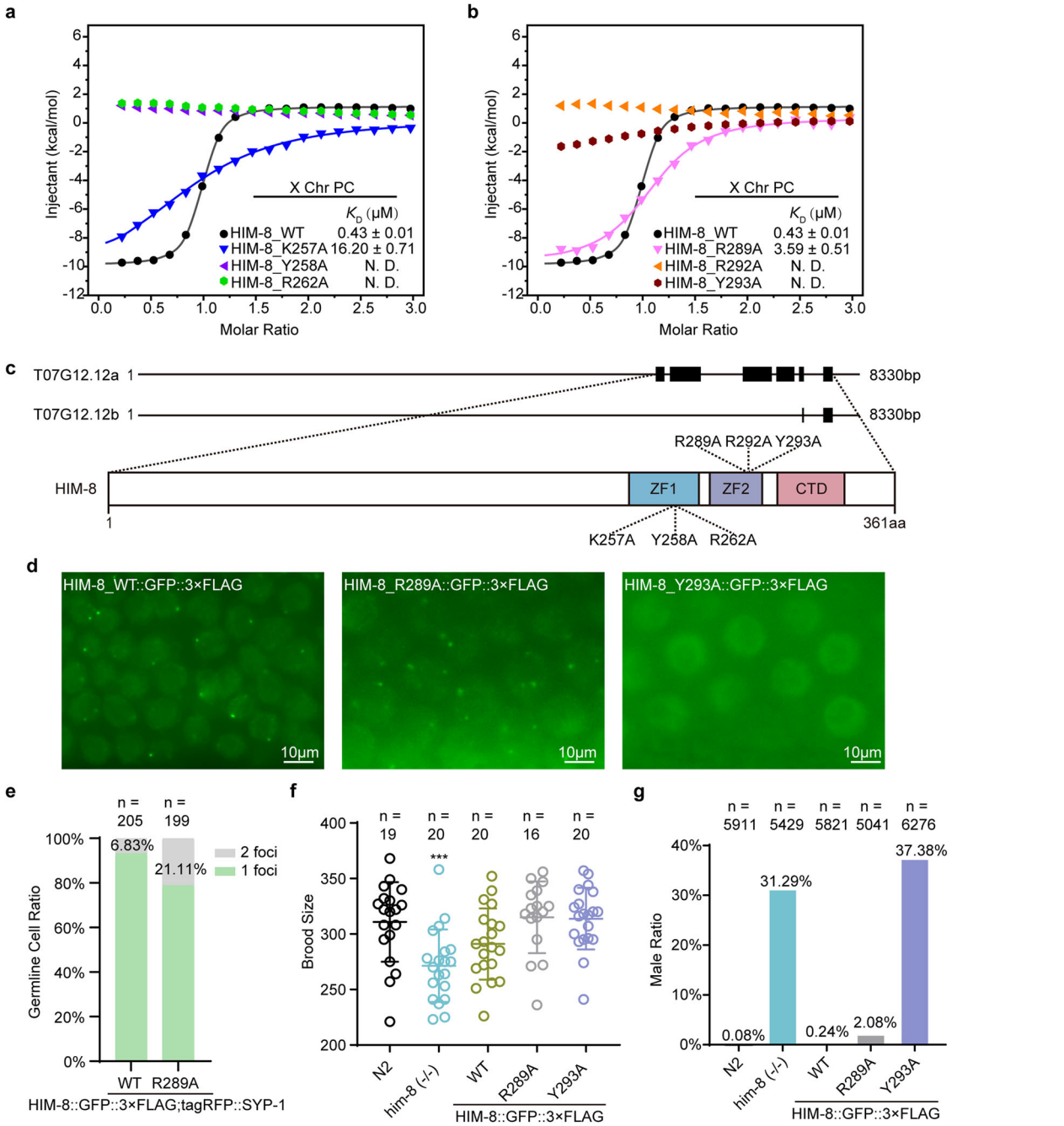


Fig. 4 | Mutational analysis of HIM-8 residues responsible for base-specific binding in vitro and in vivo. a Single-point mutations of base-specific residues on ZF1 are shown to reduce the binding affinities to the 12-bp X chromosome PC DNA sequence, as quantified by ITC. **b** Single-point mutations of base-specific residues on ZF2 lead to reduced binding affinities to the 12-bp X chromosome PC DNA sequence, as assessed by ITC. **c** A schematic diagram of *him-8* gene and HIM-8 protein. There are two variants of the *him-8* gene denoted by numbers T07G12.12a and T07G12.12b. Exons are represented by black boxes on the line. Single mutations were labeled on the HIM-8 protein. **d** Observation within the pachytene nucleus reveals typically one HIM-8 focus in HIM-8_WT::GFP::3 × FLAG worms. Conversely, an increase in unpaired foci is observed in HIM-8_R289A::GFP::3 × FLAG mutant worms, while the foci in HIM-8_Y293A::GFP::3 × FLAG mutant nematodes appear diffused. Three times were repeated independently with similar results. **e** Quantification of the ratio of worm pachytene germline cells exhibiting either one or two HIM-8 foci. The variable "*n*" denotes the number of cells analyzed. **f** The brood size of the indicated animals at 20 °C is depicted. Data are presented as Mean ± SD. A two-sided *t*-test was performed. Statistical significance was denoted by $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. WT vs N2, $p = 0.0770$; *him-8(−/−)* vs N2, $p = 0.0010$; R289A vs N2, $p = 0.7189$; Y293A vs N2, $p = 0.7741$. The "*n*" represents the number of P0 generation animals, whose self-progeny were tested. **g** The male ratio of self-progeny produced by wild-type N2, *him-8(−/−)*, HIM-8_WT::GFP::3 × FLAG, HIM-8_R289A::GFP::3 × FLAG and HIM-8_Y293A::GFP::3 × FLAG hermaphrodites serves as a measure of X chromosome nondisjunction, with "*n*" representing the number of animals counted. Source data are provided as a Source Data file.

stage cells exhibited two HIM-8 foci, whereas this percentage increased to 21.11% in the HIM-8_R289A::GFP::3 × FLAG;tagRFP::SYP-1 strain (Fig. 4e). This result suggests that while the mutant maintains the ability to bind the X chromosome PC, it may be still deficient in promoting effective pairing and synapsis, underscoring the necessity of high binding affinity. Remarkably, HIM-8 foci were conspicuously absent in the HIM-8_Y293A::GFP::3 × FLAG worm (Fig. 4d), a mutant near entirely blocked affinity for the X chromosome PC motif (Fig. 4b), indicating an inability to localize to the X chromosome pairing center.

https://doi.org/10.1038/s41467-024-54548-9

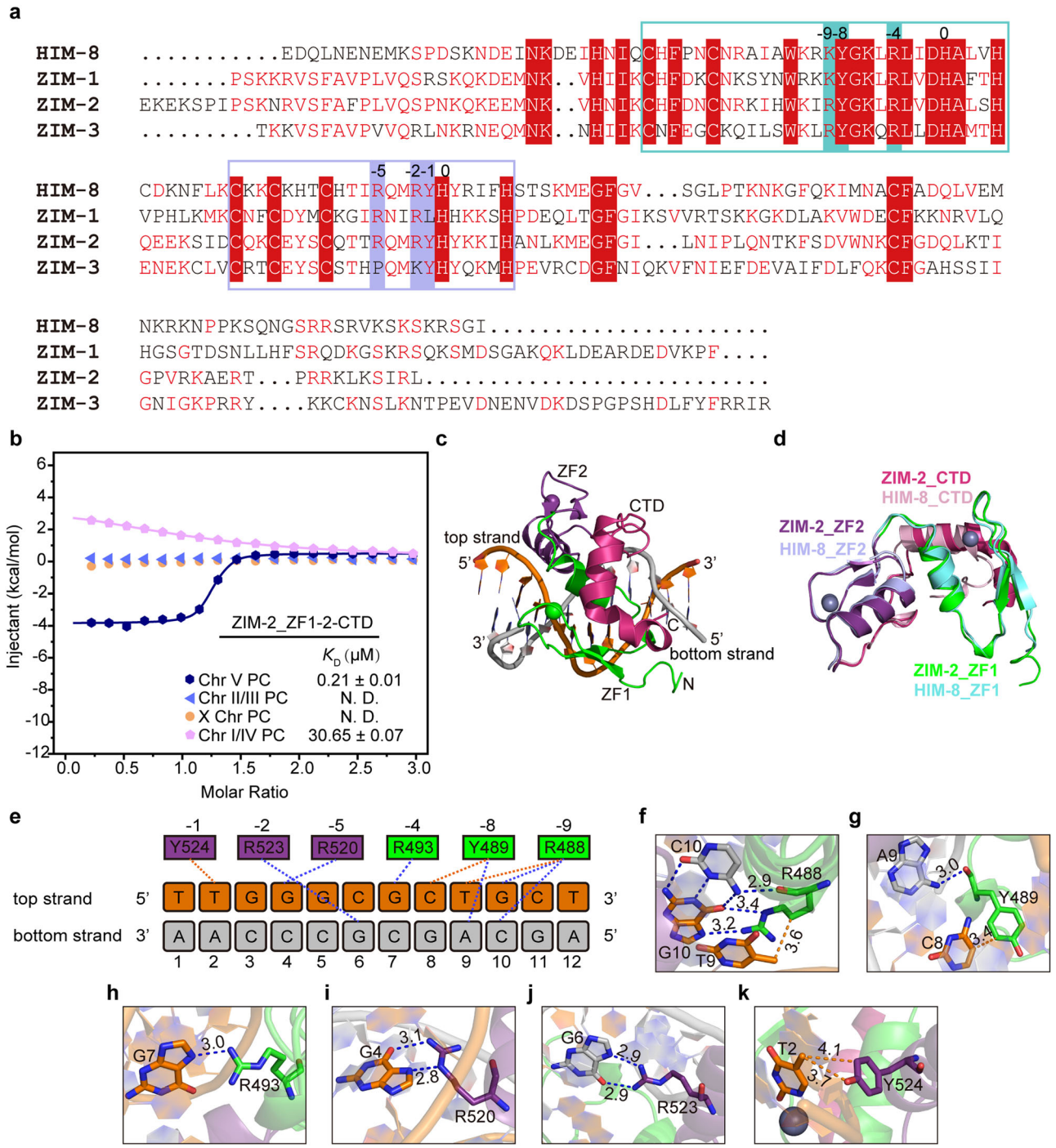


Fig. 5 | Structure of ZIM-2 ZF1-2-CTD in complex with chromosome V PC DNA. **a** The sequence alignment of ZF1-2-CTD domains from HIM-8, ZIM-1, ZIM-2, and ZIM-3 proteins was generated using the ESPript3 server, with identical residues highlighted in a red box. ZF1 residues are enclosed within an aquamarine box, while ZF2 residues are delineated by a light blue box. DNA base-contacting residues at positions −4, −8, and −9 of ZF1, and positions −1, −2, and −5 of ZF2, are indicated. **b** ITC assays demonstrate the preferential binding of ZIM-2 ZF1-2-CTD (navy blue) to the 12-bp chromosome V PC motif over PC motifs from other chromosomes (chr II/III, X chr, and chr I/IV). **c** Depiction of the overall structure of ZIM-2 ZF1-2-CTD in complex with the 12-bp chromosome V PC motif. ZF1, ZF2, and CTD are colored green, violet purple, and warm pink, respectively. The top strand and bottom strand are colored orange and gray, respectively. **d** Structure superposition of ZIM-2 ZF1-2-CTD and HIM-8 ZF1-2-CTD alignment reveals an RMSD of 0.899 Å. **e** General schematic diagram illustrating the detailed base-specific interactions between ZIM-2 ZF1-2-CTD and the 12-bp chromosome V PC DNA sequence. Color codes for ZF1, ZF2, and elements of the top and bottom strands are consistent with those defined in Fig. 5c. Direct hydrogen bonds are depicted in blue, while Van der Waals contacts are represented by orange dashed lines. **f**–**k** Further highlighting base-specific contacts at the atomic level, with interatomic distances measured in angstroms.

Upon assessing the brood sizes of these HIM-8 mutant worms, we observed minimal changes (Fig. 4f). The HIM-8_WT::GFP::3 × FLAG worm displayed a slight increase in male production, from 0.08% ($n = 5911$) to 0.24% ($n = 5821$) (Fig. 4g). In contrast, the HIM-8_Y293A::GFP::3 × FLAG worm exhibited a substantial rise in male production, reaching 37.38% ($n = 6276$), a level comparable to that observed in the *him-8 (−/−)* worm. This finding strongly suggests that the Y293A mutation entirely abolished binding to the X chromosome

PC in vivo. Additionally, we noted a modest increase in male production in the HIM-8_R289A::GFP::3 × FLAG worm (2.08%), likely due to reduced pairing center binding, resulting in ineffective pairing and synapsis (Fig. 4g).

Collectively, these observations reinforce the significance of X chromosome PC motif recognition by HIM-8 ZF1-2-CTD in ensuring accurate X chromosome pairing, synapsis, and segregation.

Recognition and specificity of PC binding by ZIM-2 ZF1-2-CTD

The sequence alignment among ZIM/HIM-8 members highlighted the conservation of amino acids responsible for the base-specific recognition of the X chromosome PC motif binding across these proteins (Fig. 5a)[28]. However, this raises a pertinent question regarding how the four paralogues distinctly recognize their specific PC motifs. To address this, we embarked on studying the structures of the ZIM proteins. Our investigation through ITC assays revealed that ZIM-2 ZF1-2-CTD exhibited a preference for the 12-bp double-stranded chromosome V PC motif, driven by a substantial negative ΔH, mirroring the interaction observed between HIM-8 and the X chromosome PC motif (Fig. 5b; Supplementary Table 1; Supplementary Data 3 and 4).

Subsequently, we obtained the crystal structure of ZIM-2 ZF1-2-CTD in complex with the chromosome V PC motif (resolved at 1.8 Å) (Fig. 5c, Supplementary Fig. 7a, b, Supplementary Table 2). Notably, ZIM-2 ZF1-2-CTD exhibited a remarkable structural similarity to HIM-8 ZF1-2-CTD, with an RMSD of 0.899 Å (Fig. 5d). A detailed examination of base-specific contacts unveiled that ZIM-2 ZF1-2-CTD recognizes its PC motif in a manner highly akin to HIM-8's recognition of the X chromosome PC motif (Figs. 5e, 3a). Specifically, R488 (−9) of ZF1 interacts with the TG sub-site and the C at base pair 10 (Fig. 5e, f). Y489 (−8) forms a hydrogen bond with the A at base pair 9 through its main chain and makes Van der Waals contacts to the C at base pair 8 (Fig. 5e, g). R493 (−4) of ZF1 recognizes the G at base pair 7 (Fig. 5e, h). R520 (−5) of ZF2 forms hydrogen bonds with the G at base pair 4 (Fig. 5e, i), and R523 (−2) of ZF2 forms hydrogen bonds with the G at base pair 6 (Fig. 5e, j). Additionally, Y524 (−1) of ZF2 makes Van der Waals contacts to the T at base pair 2 (Fig. 5e, k). These recognition patterns demonstrate a striking similarity between HIM-8 and ZIM-2. However, HIM-8 and ZIM-2 can specifically recognize the X chromosome and chromosome V PC motif, respectively.

We further utilized the AF3 server to predict the structure of the ZIM-2 ZF1-2-CTD in complex with the chromosome V PC motif. Upon comparison, the model generated by AF3 displayed a significant divergence from the corresponding experimentally resolved crystal structure (Supplementary Fig. 7c–f), akin to the discrepancies observed in the HIM-8-DNA complex. This underscores the challenges faced by AF3 in accurately predicting these specific protein-DNA interactions.

Recognition and specificity of PC binding by ZIM-1 ZF1-2-CTD

To further access the DNA-binding preference of the ZIM/HIM-8 family, we tried to investigate the underlying mechanism governing ZIM-1's recognition of the PC DNA. ZIM-1 ZF1-2-CTD exhibited a preference for chromosome II/II PC motif as revealed by ITC experiments. Nevertheless, this binding is predominantly driven by entropy, which differs from that observed for HIM-8 and ZIM-2 (Fig. 6a; Supplementary Table 1; Supplementary Data 3 and 4). We then obtained the crystal structure of ZIM-1 ZF1-2-CTD in complex with the chromosome II/III PC motif, resolved at a resolution of 1.9 Å (Fig. 6b, Supplementary Fig. 8a, b, Supplementary Table 2). A comparative analysis between the structures of ZIM-1 and HIM-8 revealed a lesser degree of similarity compared to that observed between ZIM-2 and HIM-8, showcasing a structural divergence with an RMSD of 2.4 Å. However, when individually assessing ZF1 and ZF2, both displayed high similarity to their counterparts in HIM-8. Intriguingly, the alignment of ZF1 within the overall structure revealed a distinct orientation of ZF2 concerning ZF1, a feature facilitating its specific binding to the chromosome II/III PC motif (Fig. 6c). The observed distinct conformation of ZF2 may account for the entropy-driven nature of ZIM-1's binding to its PC DNA motif. Despite our efforts, we have not yet succeeded in obtaining the structure of the apo form of ZIM-1 ZF1-2-CTD. Consequently, further investigation is required to determine whether conformational rearrangement occurs upon DNA binding.

The ZIM-1 PC motif exhibits a significant divergence from the binding motifs of HIM-8 and ZIM-2, particularly in the number of intervening sequence bases between the TTGG and TG sub-sites. Specifically, the ZIM-1 PC motif contains two base pairs, whereas the HIM-8 and ZIM-2 PC motifs comprise four base pairs (Fig. 1a). Upon examining base-specific features, it was observed that in the ZIM-1 DNA complex structure, K475 (−9) of ZF1 formed base-specific contacts with the TG sub-site and the C at base pair 8 (Fig. 6d, e). Y476 (−8) of ZF1 forms a hydrogen bond with the A at base pair 7 through its main chain and makes Van der Waals contacts to the T at base pair 5 (Fig. 6d, f), a mechanism akin to that employed by HIM-8 and ZIM-2. However, the side chain of R480 (−4) of ZF1 adopted a distinct conformation to recognize the final G of the TTGG sub-site (Fig. 6d, g), a base typically recognized by ZF2 of HIM-8 or ZIM-2 (Figs. 6h, 3e). Furthermore, while R507 (−5) of ZF2 formed hydrogen bonds with the G at base pair 3 (Fig. 6d, i), R510 (−2) of ZF2 did not engage in base-specific contacts but formed hydrogen bonds with the DNA phosphate backbone of the A at base pair 7 (Supplementary Fig. 9). Besides, L511 (−1) makes Van der Waals contacts to the T at base pair 2 (Fig. 6d, j).

In summary, the amino acids side chain conformation of ZF1 and the docking geometry of ZF2 on the DNA diversify to confer ZIM-1 specificity to chromosome II/III PC motif. We remain uncertain whether the observed conformational differences are induced by the PC DNA. Our comparison between the ZIM-1 structure in the complex and the AF3-predicted apo ZIM-1 structure revealed no significant conformational differences, with an RMSD of 0.727 Å. However, the absence of an experimental structure for the apo ZIM-1 ZF1-2-CTD leaves open the question of whether DNA induces conformational changes, warranting further investigation. Nevertheless, conventional tandem C2H2 zinc fingers often undergo conformational changes upon DNA binding, suggesting a potential cooperative effect between the PC motif and ZF1-2-CTD domains. Furthermore, similar to the predictions for HIM-8 and ZIM-2, the AF3 server is unable to accurately predict the interactions between ZIM-1 ZF1-2-CTD and the PC DNA motif (Supplementary Fig. 10a–d).

TG sub-site functions as the registration motif, and the spacing and sequence between TTGG and TG sub-site constrain ZIM/HIM-8 Binding

Consistent with the structures, mutations in the TG sub-site (TG to CA) within the X chromosome PC motif led to complete disruption of the HIM-8 ZF1-2-CTD binding, as revealed by ITC experiments. Likewise, mutation of the TG registration site of the chromosome V PC motif abolished the ZIM-2 ZF1-2-CTD binding. Additionally, ZIM-1 ZF1-2-CTD significantly reduced the binding affinity to mutated TG sub-site (Fig. 7a; Supplementary Table 1; Supplementary Data 3 and 4). These findings highlight the critical role of the TG sub-site in PC DNA recognition.

Upon analysis of the DNA complex structures of HIM-8, ZIM-2, and ZIM-1, a prominent discrepancy among the four PC motifs emerges, specifically in the spacing between the TTGG and TG sub-sites. Notably, the HIM-8 and ZIM-2 binding PC motifs exhibit a four-base separation between these sub-sites, each displaying a distinct sequence in the intervening bases. In contrast, the ZIM-1 and ZIM-3 binding PC motifs reveal only two and one intervening bases, respectively (Fig. 1a). Hence, it becomes apparent that the diversity in the sequence between the TTGG and TG sub-sites plays a pivotal role in determining their respective binding specificities.

https://doi.org/10.1038/s41467-024-54548-9

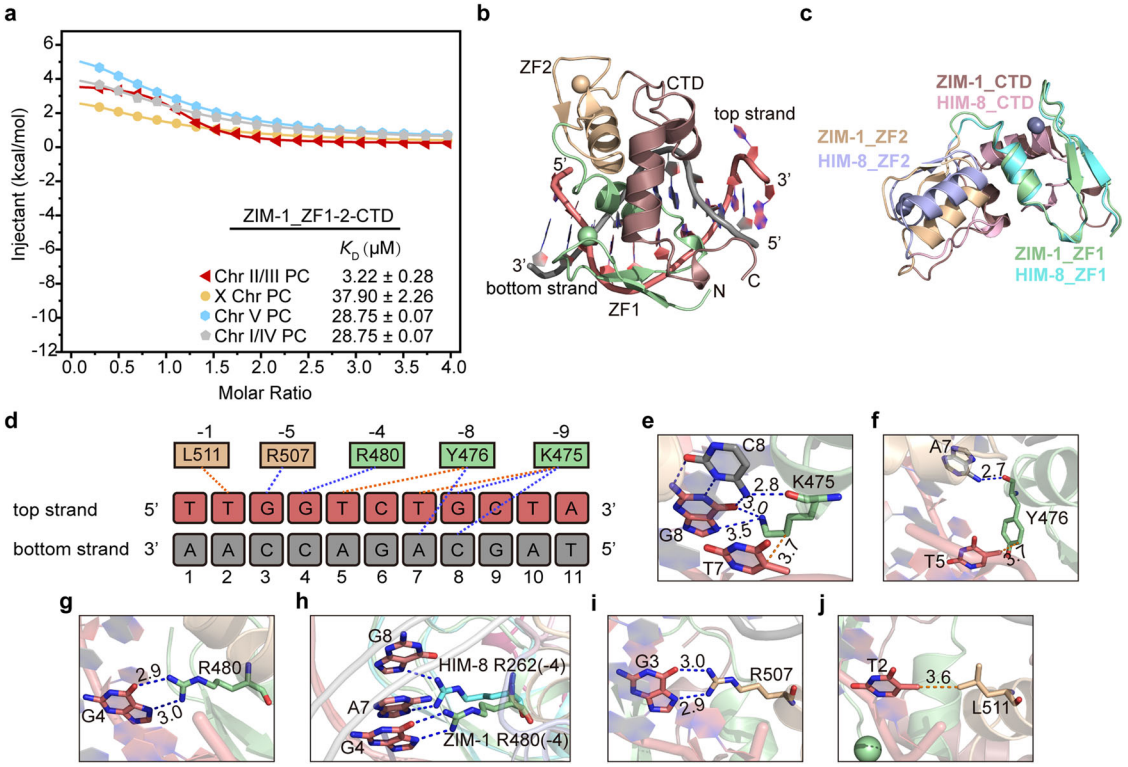


Fig. 6 | Structure of ZIM-1 ZF1-2-CTD in complex with chromosome II/III PC DNA. a ITC experiments reveal the preferential binding of ZIM-1 ZF1-2-CTD (depicted in red) to the 11-bp chromosome II/III PC motif over motifs from other chromosomes (X chr, chr V, and chr I/IV). **b** Depiction of the overall structure of ZIM-1 ZF1-2-CTD in complex with the 11-bp chromosome II/III PC motif. ZF1, ZF2, and CTD are colored in pale green, wheat, and dirty violet, respectively. The top strand and bottom strand are represented in deep salmon and gray, respectively. **c** Structural superposition of ZIM-1 ZF1-2-CTD and HIM-8 ZF1-2-CTD alignment using their ZF1 domains within the overall structure reveals an RMSD of 2.4 Å, suggesting an altered ZF2 docking geometry. **d** General schematic diagram illustrating the detailed base-specific interactions between ZIM-1 ZF1-2-CTD and the 11-bp chromosome II/III PC DNA sequence. Color codes for ZF1, ZF2, and elements of the top and bottom strands adhere to those defined in (**b**). Direct hydrogen bonds are represented in blue, while Van der Waals contacts are depicted with orange dashed lines. **e**–**g** Further highlighting base-specific contacts of ZF1 at the atomic level, with interatomic distances measured in angstroms. **h** The side chain of ZIM-1 ZF1 R480 (−4) recognizes the G at the base pair 4 on the chromosome II/III PC motif TTGG sub-site. However, the side chain of HIM-8 ZF1 R262 (−4) corresponding to ZIM-1 ZF1 R480 (−4) recognizes the A at base pair 7 and the G at base pair 8, but not the G at the base pair 4 on the X chromosome motif. **i**, **j** Further highlighting base-specific contacts of ZF2 at the atomic level, with interatomic distances measured in angstroms.

To validate this hypothesis, ITC experiments were conducted using the 12-bp X chromosome PC motif, wherein alterations were made to the sequence of the intervening bases. Modifying A:T_7 to G:C and G:C_8 to C:G drastically reduced its binding affinity to HIM-8 ZF1-2-CTD (Fig. 7b; Supplementary Table 1; Supplementary Data 3 and 4). Notably, this DNA variant exhibited an enhanced binding to ZIM-2 ZF1-2-CTD, akin to the affinity observed with the chromosome V PC motif (Fig. 7c; Supplementary Table 1; Supplementary Data 3 and 4). These outcomes strongly corroborate the pivotal role of the intervening bases sequence, particularly the 7th and 8th base pairs, as determining factors influencing the preferences of HIM-8 and ZIM-2.

Subsequently, additional ITC experiments were performed using the X chromosome PC sequence, wherein the spacing between the TTGG and TG sub-sites was altered to 3 (Δ7A or Δ8G), 2 (ΔAG), and 1 (ΔCAG). As expected, the binding affinities of HIM-8 ZF1-2-CTD to these mutations showed a notable reduction in all cases (Fig. 7d; Supplementary Table 1; Supplementary Data 3 and 4).

Collectively, these findings support a model in which the intervening sequence, influenced by both its length and sequence, co-evolves with ZIM/HIM-8 members, resulting in the emergence of highly specific preferences.

CTD contributes to PC motif selection

In pursuit of a comprehensive understanding of the inherent modularity and potential evolutionary trajectory in PC DNA selection, we set out to engineer ZIM/HIM-8 family proteins to finely adjust their DNA-binding specificity. Particularly noteworthy is that, among the six DNA base-contacting residues of HIM-8 and ZIM-2, the residue at position −9 of ZF1 differs, being a Lysine (K257) in HIM-8 and an Arginine (R488) in ZIM-2. Interestingly, this mutant, ZIM-2 R488K, also exhibited an absence of binding to X chromosome PC DNA, while retaining binding to chromosome V PC DNA, albeit with a moderately reduced affinity (Fig. 8a; Supplementary Table 1; Supplementary Data 3 and 4). This indicates that direct base-specific residues do not determine the selection of the PC motif, and variations outside the DNA-contacting residues may play a pivotal role in PC DNA selection.

Sequence alignment analysis revealed that the CTD portion exhibits greater divergence than ZF1-2 (Fig. 8b)[29]. Furthermore, previous SELEX studies have demonstrated a comparable DNA recognition pattern in ZF1-2 among ZIM/HIM-8 members, consistent with conserved critical DNA recognition residues on ZF1-2. Upon CTD addition, HIM-8 exhibits specific recognition of the X chromosome pairing center, underscoring the importance of CTD in DNA sequence recognition[14]. Therefore, we postulated that the CTD evolves to mediate PC DNA selection. To substantiate this hypothesis, we initially attempted to interchange the CTD portion between HIM-8 and ZIM-2. The resulting chimera, designated as HIM-8 ZF1-2-CTD^{ZIM-2} (Fig. 8c), exhibited reduced binding to the X chromosome PC motif and increased affinity for the chromosome V PC motif (Fig. 8d; Supplementary Table 1; Supplementary Data 3 and 4). Moreover, these two

https://doi.org/10.1038/s41467-024-54548-9

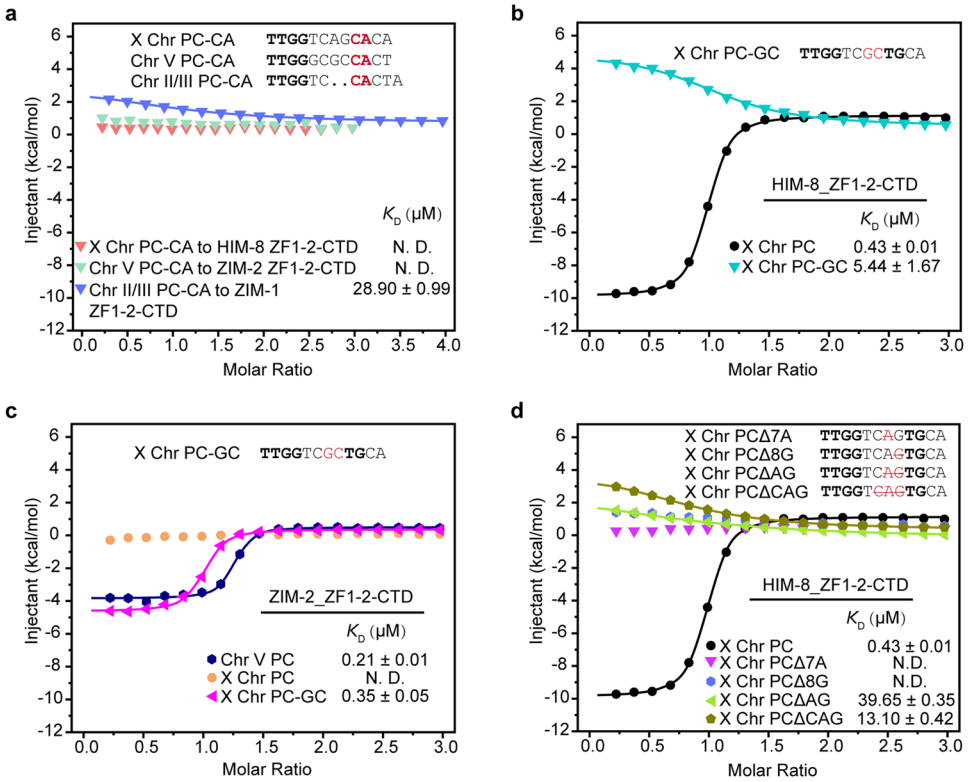


Fig. 7 | Sequence features of PC DNA. a Mutations in the TG sub-site (changing TG to CA) within the X, V, and II/III chromosome PC motifs result in the complete disruption of their binding to HIM-8 ZF1-2-CTD, ZIM-2 ZF1-2-CTD, and ZIM-1 ZF1-2-CTD, respectively. Nucleotides marked in red are mutating nucleotides. **b** Mutation of A_7G_8 to GC in the X chromosome PC motif leads to reduced binding to HIM-8 ZF1-2-CTD. Nucleotides marked in red are mutating nucleotides. **c** Mutation of A_7G_8 to GC in the X chromosome PC motif leads to enhanced binding to ZIM-2 ZF1-2-CTD. These results indicate that the sequence between the TTGG and TG sub-sites determines the binding specificity of HIM-8 and ZIM-2. Nucleotides marked in red are mutating nucleotides. **d** Reducing the number of base pairs between the TTGG and TG sub-sites to 3, 2, and 1 results in a significant reduction in their binding affinities to HIM-8 ZF1-2-CTD. Nucleotides labeled in red deletion line represent the deleting nucleotides.

bindings are both driven by entropy, which differs from that observed for wild-type HIM-8 and ZIM-2. Then we generated a mutant worm strain carrying HIM-8_CTD$^{ZIM\text{-}2}$::GFP::3 × FLAG chimeric protein in situ. In this strain, we observed a pronounced reduction in HIM-8 foci within the germline cells (Fig. 8e), with only 14.05% of the cells maintaining detectable HIM-8 foci (Fig. 8f). The brood size remained normal (Fig. 8g) compared to that of the control group. However, there was a significant increase in the male progeny ratio, escalating to 38.72% (Fig. 8h). These findings provide support for the pivotal role played by CTD in selecting the PC motif.

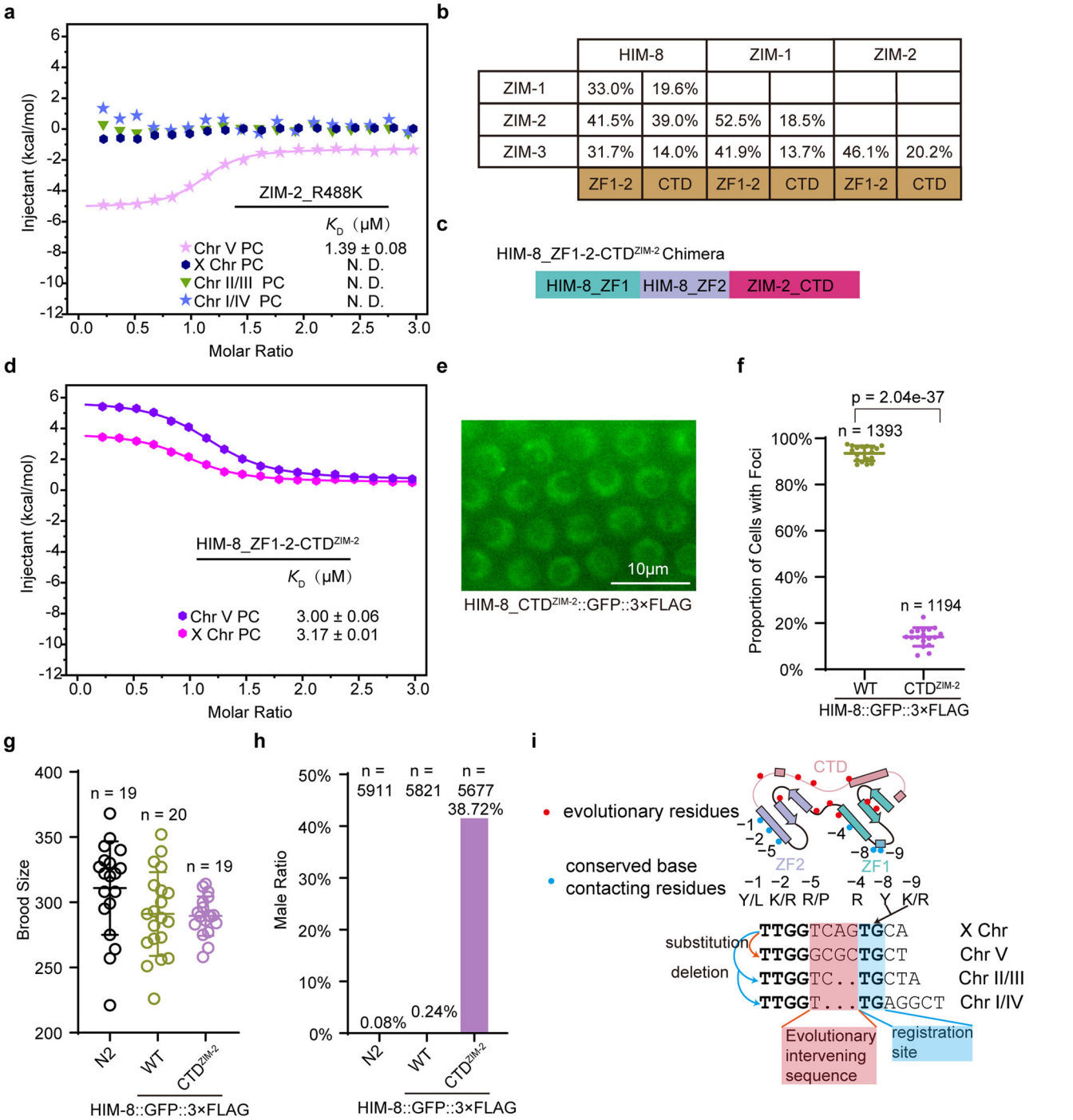
Subsequently, we tried to interchange the CTD between HIM-8 and ZIM-1 to assess their respective binding specificities. However, attempts to substitute HIM-8 CTD with ZIM-1 CTD or vice versa resulted in the failure of the two chimera proteins to fold effectively upon expression in *E. coli*. This observation further suggests the necessity for co-evolution between CTD and ZF1-2 to stabilize the overall fold of ZF1-2-CTD.

In summary, these outcomes underscore two significant findings: (1) direct DNA base-contacting residues do not significantly exert significant influence on the selection of the PC motif, and (2) the evolution of CTD and the amino acids governing ZF1, ZF2, and CTD interactions is instrumental in conferring binding preferences to distinct PC motifs.

Discussion

Investigating the molecular basis, we delineated the sequence-specific PC DNA binding attributed to the ZF1-2-CTD domain within ZIM/HIM-8 proteins. Extensive studies have elucidated how a single conventional C2H2 zinc finger domain establishes base-specific contacts with three consecutive DNA base pairs via three amino acids at canonical "recognition" positions on its helix[19]. Moreover, diverse quantities of such zinc fingers can be sequentially linked to recognize DNA strands of varying lengths[30,31]. Typically, interactions between zinc fingers are infrequent, allowing flexibility in tandem fingers prior to their DNA binding[26]. Notably, abundant interactions between ZF1-2 and the CTD domain of ZIM/HIM-8 proteins drive their conformational folding into an integral structural unit, crucial for PC DNA selection.

It's a long-standing question that how transcription factors evolved to acquire new binding specificities[32]. Like the homeodomain family cases, alterations in DNA-contacting residues underlie the emergence of their novel sequence-specific DNA-binding capabilities[33]. Yet, variations in sequence-specific DNA binding can stem not only from alterations in DNA-contacting residues but also from changes outside the DNA-binding residues or even the DNA-binding domain. Recent studies have established that the addition of a basic N-terminal or C-terminal motif to the C2H2 zinc finger confers additional binding within the minor groove. Analysis of DNA-binding specificity mechanisms in transcription factors possessing two C2H2 zinc fingers, each with identical DNA-contacting residues, revealed that an N-terminal region augments the binding affinity to sequences containing the AT dinucleotide[34]. Notably, TZAP, a newly characterized telomeric DNA-binding protein, interacts with the TTAGGG sequence using its 11th zinc finger to recognize the GGG site and a conserved basic C-terminal arm to establish minor groove contacts with the TA site[21]. Moreover, Kaizo, another C2H2 zinc finger protein, possesses a C-terminal extension facilitating minor groove contacts[35]. Conversely, the ZIM/HIM-8 CTD does not establish base-specific contacts with the PC motif; rather, it forms phosphate contacts and engages with ZF1



and ZF2, potentially modulating the ZF1-ZF2 docking conformation to the DNA, thereby imparting its binding specificity.

Another mechanism altering DNA-binding specificity in ZFs involves inter-ZF linkers and amino acids mediating inter-ZF interactions, inducing alternative DNA docking geometries that subsequently alter their binding specificity[34,36,37]. This mechanism bears a resemblance to that observed in ZIM/HIM-8 members. While the DNA-contacting residues within ZIM/HIM-8 members remain highly conserved, ZF2 of ZIM-1 adopts an alternative binding mode specific to the chromosome II/III PC motif. Prior SELEX studies demonstrated a comparable DNA recognition pattern in ZF1-2 among ZIM/HIM-8 members, consistent with our structural findings identifying conserved critical DNA recognition residues on ZF1-2[14]. Upon CTD addition, HIM-8 exhibits specific recognition of the X chromosome pairing center, emphasizing the importance of CTD in DNA sequence recognition. Moreover, the CTD region exhibits more substantial sequence diversity than ZF1-ZF2, underscoring the evolutionary modulation of CTD sequences governing ZF1-2 conformational diversity, thereby facilitating ZIM/HIM-8 to discern variations in PC sequences.

The length and composition of the intervening sequence between TTGG and TG sub-sites are crucial for the specific recognition by ZIM/HIM-8. This mechanism bears a resemblance to the mechanism employed by the nuclear receptor family. While different nuclear receptors have inherently similar DNA recognition sequences, when forming dimers to recognize DNA, they display specific recognition for

Fig. 8 | CTD contributes to PC motif selection. a ZIM-2 R488K mutant showed a moderately reduced binding to chromosome V PC and showed no binding to X chromosome, chromosome II/III, and chromosome I/IV PC DNA. **b** The sequence similarity alignment of ZF1-2 and CTD from HIM-8, ZIM-1, ZIM-2, and ZIM-3 proteins was generated using the EMBOSS Needle Pairwise Sequence Alignment (PSA) server. **c** Schematic diagram illustrating the two zinc fingers (ZF1 and ZF2) and the carboxyl-terminal domain (CTD) of the HIM-8 ZF1-2-CTD^{ZIM-2} protein. **d** The substitution of HIM-8 CTD with ZIM-2-CTD (HIM-8 ZF1-2-CTD^{ZIM-2}) resulted in reduced binding to the X chromosome PC motif and increased affinity for the chromosome V PC motif. **e** Images of representative meiotic germ cells from the HIM-8_CTD^{ZIM-2}::GFP::3 × FLAG animals illustrate that the HIM-8 foci have disappeared in most germline cells within this worm strain. Three times were repeated independently with similar results. **f** The quantification of the germline cell ratio with or without HIM-8 foci in HIM-8_CTD^{ZIM-2}::GFP::3 × FLAG animals. Here, "*n*" denotes the number of cells tested. Data are presented as Mean ± SD. A two-sided *t*-test was performed. Source data are provided as a Source Data file. **g** The brood size of the indicated animals at 20 °C is displayed, with "*n*" representing the number of P0 generation animals whose self-progeny was analyzed. Data are presented as Mean ± SD. A two-sided *t*-test was performed. WT vs N2, $p = 0.0770$; CTD^{ZIM-2} vs N2, $p = 0.2616$. **h** The male ratio of self-progeny produced by wild-type N2, HIM-8_WT::GFP::3 × FLAG, and HIM-8_CTD^{ZIM-2}::GFP::3 × FLAG hermaphrodites is presented as a measure of X chromosome nondisjunction, with "*n*" representing the number of animals counted. **i** The model depicts key features of the ZIM/HIM-8 ZF1-2-CTD binding to chromosome PCs.TG sub-sites as registration sites are highlighted in the light blue box. Evolutionary intervening sequences are highlighted in the pink box. The orange curved arrow shows the intervening sequences substitution from the X chromosome to the chromosome V PC motif. The light blue curved arrows indicate the intervening sequence deletion from the X chromosome to chromosome II/III and I/IV PC motif. Potential residues on ZF1-2 and CTD to affect DNA binding are highlighted (red dots). Conserved base-contacting residues are highlighted with light blue dots. Conserved residues type of ZF1 position −4, −8, and −9 and ZF2 position −1, −2, and −5 recognizing bases are shown. Residues at ZF1 positions −8 and −9 are responsible for TG registration sites (black arrows). Source data are provided as a Source Data file.

the number of intervening sites between consecutive DNA-binding domains (DBDs)[38–40]. A similar mechanism has been identified in plant ARF family transcription factors[41]. However, ZIM/HIM-8, besides exhibiting specific recognition for inter-site spacing, also exhibits clear selectivity for the sequence within the spacing.

In summary, we have revealed that the binding specificity is determined by both DNA motifs and amino acids. First, the differences in the number and type of spacer sequences of the four chromosomes PC motifs are important for specific binding. Second, the ZF1-2-CTD domain of the ZIM/HIM-8 family forms an integrated structural fold crucial for recognizing their respective PC motifs. The DNA base-contacting residues, primarily distributed on ZF1 and ZF2, remain highly conserved. The CTD and its interactions with ZF1 and ZF2 collectively contribute to shaping the overall structure, enabling the recognition of diverse PC motifs. Our study sheds light on the co-evolution of the ZIM/HIM-8 family with their target chromosome PC motifs. PC motifs may have evolved by changing the number and type of spacer sequences between the TTGG and TG sites. Meanwhile, the ZIM/HIM-8 proteins have evolved to adapt to the corresponding DNA sequences. There is an adaptable evolutionary relationship between them (Fig. 8i).

Methods

Plasmid construction

The DNA segments encoding the ZF1-2-CTD regions of *C. elegans* HIM-8 (amino acids 223-361), ZIM-1 (amino acids 437-560), and ZIM-2 (amino acids 450-570) were amplified from a cDNA library of the N2 nematode using PCR (Supplementary Data 1). Subsequently, the HIM-8 and ZIM-1 ZF1-2-CTD constructs were cloned into a modified pGEX-3C vector, respectively, wherein the thrombin site was replaced by a PreScission protease cleavage site. Meanwhile, the ZIM-2 ZF1-2-CTD construct was inserted into a modified pET-28a vector (Novagen), featuring an additional SUMO tag following the 6×His tag. Point mutations were introduced into the HIM-8 and ZIM-2 sequences utilizing the Blunting Kination Ligation (BKL) Kit (Takara, 6127A).

Protein expression and purification

The proteins were expressed in *E. coli* BL21-GOLD (DE3) cells (Novagen). The cells were cultured at 37 °C in Luria–Bertani (LB) medium supplemented with 0.1 mM ZnSO4. Induction of protein expression was initiated when the OD_{600} reached approximately 0.9 by adding 0.4 mM isopropyl β-D-thiogalactopyranoside (IPTG), followed by incubation at 16 °C for 24 h. Subsequently, the cells were harvested through centrifugation, and the resulting pellets were resuspended in Buffer A (20 mM Tris, pH 7.5, and 1 M NaCl). DNase A (Thermo Fisher Scientific, EN0521) and RNase A (Sangon Biotech, B694345) were added at a final concentration of 10 μg/mL to eliminate nonspecific-binding nucleic acids.

GST-tagged fusion protein-containing supernatant was subjected to glutathione 4B column chromatography (GE Healthcare) followed by overnight PreScission protease cleavage in Buffer B (20 mM Tris, pH 7.5 and 300 mM NaCl) at 4 °C. SUMO-tagged fusion protein-containing supernatant was purified using a Ni^{2+}-chelating column (GE Healthcare) followed by SUMO Protease cleavage in Buffer B. Target proteins and GST/SUMO tag were separated by the HiTrap SP FF cation exchange column (GE Healthcare) on an AKTA PURE system (GE Healthcare). Target proteins were further purified by size-exclusion chromatography on a HiLoad 16/60 Superdex 75 column (GE Healthcare) in Buffer C (20 mM Tris-HCl, pH 7.5, 150 mM NaCl) for subsequent crystallization or Buffer B for ITC experiments.

Protein–DNA complex preparation

The synthetic single-stranded DNA fragment (from General Biology, Anhui, China) was dissolved in Buffer C. Equal molar concentrations of two complementary single-stranded DNAs were mixed and heated in a water bath, then allowed to naturally cool to room temperature, forming double-stranded DNA (dsDNA). The annealed dsDNA was combined with the target protein with a ratio 1.2:1 and then dialyzed into Buffer C at 4 °C for 12 h to form the protein–DNA complex. The complex was further purified by size-exclusion chromatography on a HiLoad 16/60 Superdex 75 column in Buffer C and finally concentrated to 0.5 mM in preparation for crystallization.

Crystallization and data collection

Crystals were grown at 20 °C using the hanging-drop vapor diffusion method. Different protein-DNA complexes were mixed with unique crystallization buffers in a 1:1 ratio to form specific crystals. For HIM-8 ZF1-2-CTD in complex with X chromosome PC motif, the crystallization buffer contained 25% w/v PEG 3350, 0.1 M Bis-Tris pH 5.5, and 0.1 M Ammonium Sulfate. For ZIM-1 ZF1-2-CTD in complex with chromosome II/III PC motif, the buffer contained 20% w/v PEG 600 and 0.1 M HEPES, pH 7.4. For ZIM-2 ZF1-2-CTD in complex with chromosome V PC motif, the buffer contained 30% w/v PEG 4000 and 0.2 M Imidazole malate, pH 6.0.

All crystals were harvested within their respective crystallization solutions and supplemented with 25% glycerol as a cryoprotectant, subsequently flash-frozen in liquid nitrogen. The X-ray diffraction datasets were collected at the Shanghai Synchrotron Radiation Facility (SSRF, Shanghai, China) beamline BL17U1 and BL19U1, utilizing a diffraction wavelength of 0.979 Å. The X-ray diffraction data are shown in Supplementary Table 2.

Structure determination and refinement

All data sets were indexed, integrated, and scaled using the *HKL*-2000 program suite[42]. Subsequently, the crystallographic phases of the HIM-8/Chromosome X PC motif sets were determined by zinc single-wavelength anomalous dispersion. Specifically, the zinc atom sites were found by the SHELX C/D program, while phasing and density modification were conducted using the *AutoSol* program within the *PHENIX* suite[43,44]. The HIM-8 ZF1-2-CTD protein and the DNA model were built with the aids of the Buccaneer program in the *CCP4* suite and *Coot*[45–47]. Crystallographic phasing of ZIM-2 and ZIM-1 data sets was performed with the PHASER program within the *CCP4* suite[48], employing the HIM-8 ZF1-2-CTD structure as the search model. All structural refinements were performed with the *PHENIX.Refine* program[49]. Detailed crystallographic parameters are presented in Supplementary Table 2. All the figures of structures were generated using PyMOL.

Isothermal titration calorimetry

ITC experiments were conducted at 25 °C utilizing a MicroCal PEAQ-ITC instrument. Double-stranded oligo deoxynucleotides were loaded into the syringe at a concentration of approximately 750 μM or 1 mM. Meanwhile, proteins were loaded into the sample cell at approximately 50 μM. The titration protocol involved 19 injections, starting with a single initial 1 μL injection, followed by 18 injections of 2 μL of DNA sample into the protein sample cell. The intervals between injections were maintained at 120 s and the reference power was set at 5 μcal/s. Thermodynamic data were analyzed with a one-site binding model through the MicroCal PEAQ-ITC Analysis Software. The ITC binding parameters were provided in Supplementary Table 1. The protein sequences used in ITC experiments are shown in Supplementary Data 3. The ITC binding curves twice repeated independently were provided in Supplementary Data 4. The data and figures were analyzed and visualized by OriginPro 2021b.

C. elegans strain construction

The Bristol strain N2 served as the native wild-type strain, and all the strains used in this study were cultivated consistently at 20 °C. Detailed information regarding the strains utilized is provided in Supplementary Data 2. To generate the HIM-8::GFP::3 × FLAG worm, approximately 1500-bp of genomic DNA sequence, including upstream and downstream homologous arm sequences (left and right arms), along with a linker sequence (GGAGGTGGAGGTGGAGCT), and the *gfp::3* × *flag* coding sequence, were PCR-amplified. These sequences were co-constructed into an optimized pCFJ151 vector (pCFJOA) using the ClonExpress MultiS One Step CloningKit (Vazyme, C113-02). The *gfp::3* × *flag* sequence was inserted before the stop codon of the *him-8* gene via the 18-bp linker sequence, yielding the plasmid pCFJOA_*him-8::gfp::3* × *flag*. Small guide RNAs targeting the *him-8* gene were designed via the Integrated DNA Technologies (IDT) website and were cloned into the Pu6 plasmid. The sg-RNA sequences are listed in Supplementary Data 1. The plasmids mix containing pDD162 (50 ng/μL), the pCFJOA_*him-8::gfp::3* × *flag* repair plasmid (50 ng/μL), pCFJ90 (5 ng/μL), and five sgRNAs (30 ng/μL) was injected into the gonad of young adult N2 animals.

The six single-point mutant nematodes in the HIM-8::GFP::3 × FLAG background were also created employing the CRISPR/Cas9 system[50]. These transgenic strains and mutants were screened through PCR amplification, microimaging, and confirmed by sequencing. The *him-8* null allele mutant was conducted by the CRISPR/Cas9 system without a repair plasmid, screened via PCR amplification, and confirmed by sequencing.

Brood size and male ratio assay

Brood size assays were performed at 20 °C. Adult animals (P0 generation) were selected and transferred onto an NGM plate to generate the first progeny (F1 generation). When the F1 worms reached in L3–L4 stage, 20 F1 hermaphrodites were singled to 20 new NGM plates and then transferred daily until no embryo production was observed on the plates. The hatched F2 generation worms, including the number of males and hermaphrodites were counted and summarized, respectively. The data and figures were analyzed and visualized by GraphPad Prism 8.

Quantification of HIM-8::GFP foci

For the quantification of HIM-8::GFP foci number in Fig. 4e, approximately 200 pachytene germline cells were selected from at least 15 independent animals. In each gonad, cells within an area where foci were clearly visible were circled in the pachytene region, with each circle containing about 10 cells. The number of cells exhibiting one or two foci was then counted. For the quantification of germline cells with HIM-8::GFP foci in Fig. 8f, a total of at least 1000 germ cells were analyzed from 18 independent animals. The data and figures were analyzed and visualized by GraphPad Prism 8.

ChIP experiments

ChIP experiments were performed as previously described[51,52]. Young adult worm samples were cross-linked with 2% formaldehyde for 30 min. Fixation was quenched with 125 mM glycine for 5 min at room temperature. Subsequently, the samples were sonicated for 20 cycles (30 s on and 30 s off per cycle) at medium output using a Bioruptor 200 machine. Lysates from HIM-8::GFP::3 × FLAG worm were pre-cleared and immunoprecipitated with 1.5 μL (1.5:1000) of rabbit anti-GFP antibody (Abcam, ab290) overnight at 4 °C. Chromatin/antibody complexes were recovered with Dynabeads Protein A (Invitrogen, 10002D) and were subjected to sequential extensive washes with 150 mM, 500 mM, and 1 M NaCl. The crosslinks were reversed overnight at 65 °C. Only the input DNA sample was treated with RNase (Thermo Fisher Scientific, EN0531) for 30 min at 65 °C, and then all DNA samples were purified using a QIAquick PCR purification kit (Qiagen, 28104).

ChIP-Seq

The ChIP DNA samples were deep sequenced at Novogene Bioinformatics Technology Co., Ltd. (Tianjin, China). Briefly, 10–300 ng of ChIP DNA sample was combined with End Repair Mix (Novogene Bioinformatics Technology Co., Ltd. (Tianjin, China)) and incubated at 20 °C for 30 min, then DNA was purified with a QIAquick PCR purification kit (Qiagen) and incubated with A-tailing mix at 37 °C for 30 min. The 3′-end-adenylated DNA was incubated with the adapter in the ligation mix at 20 °C for 15 min. The adapter-ligated DNA was amplified through several rounds of PCR and purified in a 2% agarose gel to recover the target fragments. The average length was analyzed using an Agilent 2100 Bioanalyzer instrument (Agilent DNA 1000 Reagents) and quantified by qPCR (TaqMan probe). The libraries were further amplified on a cBot system to generate clusters on the flow cell and sequenced by a single-end 50 method on a HiSeq1500 system.

ChIP-Seq data analysis

Part I of ChIP-seq analysis is genomic alignment. ChIP-seq clean reads were aligned to the ce11(wbcel235) assembly of the *C. elegans* genome using Bowtie 2 (version 2.3.5.1) with default settings[53]. The samtools (version 0.1.19) "view" utility was used to convert the alignments to Binary Alignment Map (BAM) format[54]. The BAM format files were sorted with samtools "sort" utility, and the "index" utility was used to index the BAM files. In Part II, ChIP-seq peaks were called using MACS2 version 2.1.1 with a permissive 0.01 *q*-value cutoff against summed ChIP-seq input[55]. Part III of ChIP-seq analysis is visualization with Integrative Genomics Viewer (IGV) genome browser[56]. Running MACS2 "bdgcmp" utility to generate logLR track. Peaks overlapping blacklisted regions were discarded. Deeptools subcommand bamCoverage

(version 3.5.0) was used to produce bigWig track from bam files[57]. The bigWig file could be displayed by the IGV genome browser to visualize signals and peaks. Part IV for ChIP-seq analysis is the binding motif searching. The default output file format of MACS2 is the narrow peak file, which is essentially Browser Extensible Data (BED) format. Running Bedtools "getfasta" utility to generate a FASTA file from the BED file. Then the binding motif discovery could be performed by inputting the FSATA format file in "meme" utility of MEME suite[58].

Microscopy imaging

Images were captured using a Leica DM4 B microscope with Leica LAS X Microscope Software. Gonads were dissected in PBS buffer supplemented with 0.2 mM levamisole.

Western blot

Worms at the young adult stage were harvested, washed three times with 1 × M9 buffer, boiled in SDS loading dye supplement with 100 mM DL-Dithiothreitol (DTT) at 95 °C for 10 min, and then resolved by SDS–PAGE. Proteins were transferred to a 0.2 μm nitrocellulose blotting membrane (Cytiva, A30377262) and blocked with 5% skim milk in 1 × TBST buffer. The membrane was incubated with primary antibodies overnight at 4 °C and subsequently incubated with secondary antibodies.

Primary antibodies for western blots were anti-FLAG monoclonal antibodies produced in mouse (Sigma, F1804) at 1:1000, and rabbit anti-β-actin monoclonal antibodies (Beyotime, AF5003) used at 1:5000. The secondary antibodies were goat anti-mouse (Beyotime, A0216) used at 1:5000 and goat anti-rabbit (Beyotime, A0208) antibodies at 1:5000.

Quantitative real-time PCR (qRT-PCR)

RNAs were isolated from young adult animals by freezing and dissolving repeatedly in liquid nitrogen and 37 °C water bath in TRIzol solution followed by DNase I digestion (Thermo Fisher Scientific, EN0521) and isopropanol precipitation. RNAs were reversed transcribed to cDNAs via HiScript III RT SuperMix for qPCR (Vazyme, R323) with Oligo (dT) 23VN primers. cDNAs were quantified with ChamQ Universal SYBR qPCR Master Mix (Vazyme, Q111-02) via qRT-PCR reaction, which was performed by LightCycler 96 (Roche) machine and analyzed by the LightCycler 96 Software (Roche). Primers for qRT-PCR analyses are shown in Supplementary Data 1. The data and figures were analyzed and visualized by GraphPad Prism 8.

Statistics and reproducibility

For each microscopy result in Figs. 1b, 4d, and 8e; Supplementary Figs. 5a and 6b, independent experiments were performed three times. For ITC assays in Figs. 1f, g, 4a, b, 5b, 6a, 7a–d, and 8a, d from two independent experiments performed in duplicate.

The mean and standard deviation of the results are presented in graphs with error bars. All of the experiments were conducted with independent *C. elegans* animals or the indicated number of replicates (*N*). Statistical analysis was performed with the two-tailed Student's *t*-test as indicated.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The atomic coordinates and structure factors for the HIM-8 ZF1-2-CTD::X chromosome PC motif, ZIM-2 ZF1-2-CTD::chromosome V PC motif and ZIM-1 ZF1-2-CTD::chromosome II/III PC motif structures have been deposited into the Protein Data Bank (PDB) under accession codes 8YV9, 8YVA, and 8YVB, respectively. The raw sequence data of HIM-8::GFP have been deposited in the Genome Sequence Archive of China National Center for Bioinformation National Genomics Data Center under submission number CRA015547. Source data are provided with this paper.

References

1. Bhalla, N. & Dernburg, A. F. Prelude to a division. *Annu. Rev. Cell Dev. Biol.* **24**, 397–424 (2008).
2. Láscarez-Lagunas, L., Martinez-Garcia, M. & Colaiácovo, M. SnapShot: meiosis – prophase I. *Cell* **181**, 1442–1442.e1441 (2020).
3. Klutstein, M. & Cooper, J. P. The Chromosomal Courtship Dance – homolog pairing in early meiosis. *Curr. Opin. Cell Biol.* **26**, 123–131 (2014).
4. Zickler, D. & Kleckner, N. Meiosis: dances between homologs. *Annu. Rev. Genet.* **57**, 1–63 (2023).
5. Lee, C. Y. et al. Mechanism and regulation of rapid telomere prophase movements in mouse meiotic chromosomes. *Cell Rep.* **11**, 551–563 (2015).
6. Shibuya, H., Hernandez-Hernandez, A., Morimoto, A., Negishi, L., Hoog, C. & Watanabe, Y. MAJIN links telomeric DNA to the nuclear membrane by exchanging telomere cap. *Cell* **163**, 1252–1266 (2015).
7. Chikashige, Y. et al. Membrane proteins Bqt3 and -4 anchor telomeres to the nuclear envelope to ensure chromosomal bouquet formation. *J. Cell Biol.* **187**, 413–427 (2009).
8. Zhang, F. et al. The F-box protein ZYGO1 mediates bouquet formation to promote homologous pairing, synapsis, and recombination in rice meiosis. *Plant Cell* **29**, 2597–2609 (2017).
9. Wang, Y. et al. The meiotic TERB1-TERB2-MAJIN complex tethers telomeres to the nuclear envelope. *Nat. Commun.* **10**, 564 (2019).
10. Tsai, J. H. & McKee, B. D. Homologous pairing and the role of pairing centers in meiosis. *J. Cell Sci.* **124**, 1955–1963 (2011).
11. MacQueen, A. J. et al. Chromosome sites play dual roles to establish homologous synapsis during meiosis in *C. elegans*. *Cell* **123**, 1037–1050 (2005).
12. Phillips, C. M. et al. HIM-8 binds to the X chromosome pairing center and mediates chromosome-specific meiotic synapsis. *Cell* **123**, 1051–1063 (2005).
13. Phillips, C. M. & Dernburg, A. F. A family of zinc-finger proteins is required for chromosome-specific pairing and synapsis during meiosis in C. elegans. *Dev. Cell* **11**, 817–829 (2006).
14. Phillips, C. M. et al. Identification of chromosome sequence motifs that mediate meiotic pairing and synapsis in *C. elegans*. *Nat. Cell Biol.* **11**, 934–942 (2009).
15. Rog, O. & Dernburg, A. F. Chromosome pairing and synapsis during *Caenorhabditis elegans* meiosis. *Curr. Opin. Cell Biol.* **25**, 349–356 (2013).
16. Sanford, C. & Perry M. D. Asymmetrically distributed oligonucleotide repeats in the Caenorhabditis elegans genome sequence that map to regions important for meiotic chromosome segregation. *Nucleic Acids Res.* **29**, 2920–2926 (2001).
17. Wynne, D. J., Rog, O., Carlton, P. M. & Dernburg, A. F. Dynein-dependent processive chromosome motions promote homologous pairing in *C. elegans* meiosis. *J. Cell Biol.* **196**, 47–64 (2012).
18. Bastos, M. et al. Isothermal titration calorimetry. *Nat. Rev. Method Prim.* **3**, 1–23 (2023).
19. Pabo, C. O., Peisach, E. & Grant, R. A. Design and selection of novel Cys_2His_2 zinc finger proteins. *Annu. Rev. Biochem.* **70**, 313–340 (2001).
20. Pavletich, N. P. & Pabo, C. O. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* **252**, 809–817 (1991).
21. Zhao, Y. et al. The 11th C2H2 zinc finger and an adjacent C-terminal arm are responsible for TZAP recognition of telomeric DNA. *Cell Res.* **28**, 130–134 (2018).

22. Zhang, Y., & Skolnick, J. TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.* **33**, 2302–2309 (2005).
23. Krissinel, E. & Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **372**, 774–797 (2007).
24. Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
25. Liu, Y., Zhang, X., Blumenthal, R. M. & Cheng, X. A common mode of recognition for methylated CpG. *Trends Biochem. Sci.* **38**, 177–183 (2013).
26. Choo, Y. & Klug, A. Physical basis of a protein-DNA recognition code. *Curr. Opin. Struct. Biol.* **7**, 117–125 (1997).
27. Sera, T. Zinc-finger-based artificial transcription factors and their applications. *Adv. Drug Deliv. Rev.* **61**, 513–526 (2009).
28. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324 (2014).
29. Madeira, F. et al. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.* **50**, W276–W279 (2022).
30. Klug, A. The discovery of zinc fingers and their development for practical applications in gene regulation and genome manipulation. *Q. Rev. Biophys.* **43**, 1–21 (2010).
31. Persikov, A. V. & Singh, M. De novo prediction of DNA-binding specificities for Cys_2His_2 zinc finger proteins. *Nucleic Acids Res.* **42**, 97–108 (2014).
32. Gera, T., Jonas, F., More, R. & Barkai, N. Evolution of binding preferences among whole-genome duplicated transcription factors. *eLife* **11**, e73225 (2022).
33. Berger, M. F. et al. Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. *Cell* **133**, 1266–1276 (2008).
34. Siggers, T., Reddy, J., Barron, B. & Bulyk, M. L. Diversification of transcription factor paralogs via noncanonical modularity in C2H2 zinc finger DNA binding. *Mol. Cell* **55**, 640–648 (2014).
35. Buck-Koehntop, B. A. et al. Molecular basis for recognition of methylated and specific DNA sequences by the zinc finger protein Kaiso. *Proc. Natl Acad. Sci. USA* **109**, 15229–15234 (2012).
36. Hamed, M. Y., Siam, R. & Zaid, R. The role of zinc finger linkers in zinc finger protein binding to DNA. *J. Comput. Aided Mol. Des.* **35**, 973–986 (2021).
37. Garton, M. et al. A structural approach reveals how neighbouring C2H2 zinc fingers influence DNA binding specificity. *Nucleic Acids Res.* **43**, 9147–9157 (2015).
38. Umesono, K., Murakami, K. K., Thompson, C. C. & Evans, R. M. Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D3 receptors. *Cell* **65**, 1255–1266 (1991).
39. Kurokawa, R. et al. Regulation of retinoid signalling by receptor polarity and allosteric control of ligand binding. *Nature* **371**, 528–531 (1994).
40. Siggers, T. & Gordan, R. Protein-DNA binding: complexities and multi-protein codes. *Nucleic Acids Res.* **42**, 2099–2111 (2014).
41. Boer, D. R. et al. Structural basis for DNA binding specificity by the auxin-dependent ARF transcription factors. *Cell* **156**, 577–589 (2014).
42. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
43. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **75**, 861–877 (2019).
44. Sheldrick, G. M. A short history of SHELX. *Acta Crystallogr. Sect. A* **64**, 112–122 (2008).
45. Agirre, J. et al. The CCP4 suite: integrative software for macromolecular crystallography. *Acta Crystallogr. Sect. D. Struct. Biol.* **79**, 449–461 (2023).
46. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **66**, 486–501 (2010).
47. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **62**, 1002–1011 (2006).
48. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
49. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **68**, 352–367 (2012).
50. Dickinson, D. J., Ward, J. D., Reiner, D. J. & Goldstein, B. Engineering the Caenorhabditis elegans genome using Cas9-triggered homologous recombination. *Nat. Methods* **10**, 1028–1034 (2013).
51. Huang, X. et al. A chromodomain protein mediates heterochromatin-directed piRNA expression. *Proc. Natl. Acad. Sci. USA* **118**, e2103723118 (2021).
52. Hou, X. et al. Systematic characterization of chromodomain proteins reveals an H3K9me1/2 reader regulating aging in *C. elegans*. *Nat. Commun.* **14**, 1254 (2023).
53. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
54. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
55. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
56. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
57. Ramirez, F., Dundar, F., Diehl, S., Gruning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–W191 (2014).
58. Bailey, T. L., Johnson, J., Grant, C. E. & Noble, W. S. The MEME Suite. *Nucleic Acids Res.* **43**, W39–W49 (2015).

Acknowledgements

This work was financially supported by grants from the Strategic Priority Research Program of the Chinese Academy of Sciences [XDB39010300] (to Y.S.) and [XDB39010600] (to S.G.); Chinese National Natural Science Foundation [32293210, 92149302, 32090040, 31870760] (to Y.S.), [32371282, 32171222, U1932122] (to F.L.) and [32230016] (to S.G.); the National Key R&D Program of China [2022YFA1302700] (to S.G.), the Fundamental Research Funds for the Central Universities [WK9100000062] (to F.L.). We thank the staff of the BL17U1 and BL19U1 beamlines at Shanghai Synchrotron Radiation Facility and the National Center for Protein Science Shanghai, People's Republic of China, for assistance during data collection. We also thank Dr. Huichao Ou at the Experiment Center for Life Science of USTC for his valuable suggestions.

Author contributions

Fudong Li and Shouhong Guang conceptualized the research. Fudong Li, Yunyu Shi, and Shouhong Guang designed the research. Meili Li performed protein expression and purification, crystallographic experiments, and ITC with the help of Zheng Xu, Mengqi Lv, Yongrui Liu, Yong Zhang, Ziyan Xu, Xu Han, and Suman Wang. Fudong Li determined the structures. Meili Li performed the construction of nematode strains, microscopy imaging, brood size and male ratio assay, quantification of HIM-8::GFP foci, ChIP-seq experiments, western blot, and qRT-PCR with the help of Chengming Zhu, Zheng Xu, Mingjing Xu, Yan Kuang, Xinhao Hou, and Xinya Huang. Fudong Li, Meili Li, Shouhong Guang, and Yunyu Shi analyzed the data and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-54548-9.

Correspondence and requests for materials should be addressed to Yunyu Shi, Shouhong Guang or Fudong Li.

Peer review information *Nature Communications* thanks Owen Davies and the other, anonymous, reviewers for their contribution to the peer review of this work. A peer review file is available.