nature immunology

Review article

https://doi.org/10.1038/s41590-026-02483-w

Multimerizing transcription factors FOXP3 and AIRE as chromatin architectural regulators

Received: 23 November 2025

Accepted: 3 March 2026

Published online: 03 April 2026

Check for updates

Fangwei Leng [1,2] ✉, Yu-San Huoh [3] ✉ & Sun Hur [4,5,6] ✉

Central and peripheral immune tolerance depend on distinct transcriptional programs orchestrated by autoimmune regulator (AIRE) and FOXP3, respectively. AIRE promotes the expression of peripheral tissue antigens in medullary thymic epithelial cells for negative selection of autoreactive T cells, whereas FOXP3 enforces the immune-suppressive program of regulatory T cells. Although their immunological roles are well established, the molecular mechanisms by which AIRE and FOXP3 engage the genome and regulate transcription have long been unclear. Recent structural, biochemical and genomic work has revealed an unexpected shared principle: both FOXP3 and AIRE form homomultimers that function as chromatin organizers. Despite functioning in different immunological contexts and possessing distinct modes of genome interaction, both proteins leverage and reinforce pre-existing chromatin landscapes to coordinate broader gene expression programs. In this Review, we summarize recent advances and emerging mechanistic insights into FOXP3 and AIRE, focusing on their multimerization, interactions with repetitive DNA and enhancers and roles as architectural regulators that shape transcriptional programs essential for immune tolerance.

The ability to distinguish self from non-self is fundamental to all immune systems. Although the remarkable capacity of immune receptors to detect pathogen-derived non-self molecules often draws attention, equally important is the immune system's ability to tolerate self, failure of which leads to critical immune disorders. In the vertebrate adaptive immune system, this self-tolerance is achieved through multilayered developmental programs, which allow for the elimination of autoreactive T cells and B cells and the generation of immunosuppressive subsets such as regulatory T (T_{reg}) cells.

Genetic and immunological studies of immune dysregulation in both humans and mice have identified two transcriptional regulators, FOXP3 and AIRE, as key mediators of immune tolerance. Mutations in either of the genes encoding these transcriptional regulators cause severe autoimmune diseases, underscoring their nonredundant roles in maintaining self-tolerance, specifically AIRE in central tolerance and FOXP3 in peripheral tolerance. Despite their established importance, the molecular mechanisms by which FOXP3 and AIRE enforce tolerance have been elusive and the subject of considerable debate over the past two decades. Emerging evidence indicates that their functions extend beyond the conventional paradigm of transcription factors (TFs), which typically regulate gene expression through sequence-specific DNA binding and recruitment of the basal transcriptional machinery.

[1]School of Basic Medicine Sciences, Capital Medical University, Beijing, China. [2]Beijing Key Laboratory of Autoimmune Disease Mechanism Research and Novel Drug Development, Institute for Immunology, Chinese Institutes for Medical Research, Beijing, China. [3]Department of Cell Biology, State University of New York Downstate Health Sciences University, Brooklyn, NY, USA. [4]Howard Hughes Medical Institute, Boston Children's Hospital, Boston, MA, USA. [5]Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, MA, USA. [6]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA, USA. ✉e-mail: fangwei.leng@cimrbj.ac.cn; yu-san.huoh@downstate.edu; sun.hur@crystal.harvard.edu

Recent genomics, biochemistry and structural studies have instead revealed unexpected architectural functions of FOXP3 and AIRE, arising from their unique capacity to form homomultimeric assemblies. These findings not only help redefine the principles underlying immune self-tolerance but also introduce a new and broader conceptual framework for transcriptional regulation.

In this Review, we summarize recent developments in our understanding of the molecular mechanisms and regulatory functions of FOXP3 and AIRE, with a particular focus on emerging structural and biochemical insights.

FOXP3

Immunological and transcriptional functions

FOXP3 is a forkhead (FKH)-family TF that serves as a key player in the development and functions of T_{reg} cells[1–3]. Loss-of-function mutations in *Foxp3* cause severe autoimmune disorders in both mice (scurfy)[4] and humans (IPEX syndrome)[5–8], establishing its indispensable role in immune tolerance. Although once thought to be strictly required for T_{reg} cell lineage specification, subsequent work has revealed a more nuanced view; FOXP3 alone is neither strictly necessary nor sufficient for T_{reg} cell differentiation[9,10]. FOXP3-deficient animals still generate T_{reg}-like cells, but these are heterogeneous, metabolically unstable and functionally impaired. Conversely, ectopic expression of FOXP3 in conventional $CD4^+$ T cells induces only a partial T_{reg} cell-like program, insufficient to recapitulate bona fide T_{reg} cell identity[11,12]. Thus, FOXP3 seems to be a reinforcer of mature T_{reg} cell gene expression and function.

Previous studies have shown that FOXP3 expression is associated with both activation and repression of hundreds of genes[13–16]. Current evidence suggests that most of these genes are mediated indirectly through downstream TFs, such as TCF-1, whereas only a small subset of genes may represent direct FOXP3 targets[13,17,18]. However, the identity of these direct targets remains controversial. For many TFs, direct targets are inferred by intersecting genes whose expression changes following TF perturbation with genes located near TF-binding sites. In the case of FOXP3, such analyses have not yielded a clear picture because genes that respond transcriptionally to FOXP3 are highly context dependent[18–21], whereas the genomic occupancy of FOXP3 is stable across experimental conditions[21].

This disconnect between FOXP3 genomic occupancy and transcriptional regulation was most clearly demonstrated in recent studies using chemically induced FOXP3 degradation mouse models[19,20]. Acute FOXP3 loss affects up to ~500 genes, but both the number and identity of responsive genes vary substantially depending on immunological context. In resting T_{reg} cells, FOXP3 depletion has minimal transcriptional impact, whereas neonatal settings and certain inflammatory conditions markedly amplify FOXP3-dependent regulation[19,20]. Even under inflammatory conditions where FOXP3 depletion has its greatest impact, most FOXP3-occupied loci show little change in expression of nearby genes, and the small fraction of the 'direct' target genes (that is, affected by FOXP3 with nearby FOXP3-occupied site) show modest, often less than two- to threefold, change following FOXP3 depletion[20]. Interestingly, most of these direct target genes are downregulated following FOXP3 expression, consistent with earlier reports showing that some FOXP3-bound sites lose the repressive histone mark H3K27me3 following FOXP3 loss[21] and that FOXP3 may recruit the polycomb repressive complex 2 to silence gene expression[22–27]. However, such an effect is seen only in a small subset of FOXP3-bound sites, with most FOXP3-occupied sites showing little to no change in accessibility or histone marks[28], raising the question of what determines this effect.

Compounding this complexity is uncertainty regarding the mechanisms underlying FOXP3 genomic targeting and sequence specificity. Depending on peak-calling criteria, several thousand to tens of thousands of FOXP3-bound loci have been reproducibly identified across multiple datasets, including CUT&RUN-seq (CNR-seq)[13,29] and chromatin immunoprecipitation followed by high-throughput sequencing (ChIP–seq)[13,30] datasets in T_{reg} cells. A recent ChIP with exonuclease treatment (ChIP–exo)[31] study using ectopically expressed FOXP3 provided higher-resolution mapping of FOXP3-binding sites; however, this study was performed in mouse embryonic stem cells rather than T_{reg} cells. A consistent feature across these datasets is that the canonical FKH motif (TGTTTAC), which most FKH TFs commonly bind, is only modestly enriched, representing roughly 10–20% of occupied sites. Moreover, DNase I protection patterns of these canonical sequences appeared unaffected by FOXP3 expression[28], suggesting that these sites may be occupied by other FKH TFs, such as FOXP1. As discussed later, FOXP3 exhibits distinct sequence preferences and binding modes that diverge from those of other FKH TFs, explaining much of its unique genomic occupancy.

Together, these observations raise the question of whether FOXP3 functions through mechanisms distinct from those of conventional TFs that couple genomic binding to local gene regulation.

Domain architecture

FOXP3 contains a proline-rich N-terminal region, a C2H2 zinc-finger (ZF) and a coiled-coil (CC), a long linker known as the RUNX1-binding region (RBR) and a conserved FKH domain (Fig. 1a). Unlike other FOXP family members (such as FOXP1, FOXP2 and FOXP4), which contain a predicted repression domain, FOXP3 lacks recognizable activator or suppressor domains[32]. Instead, FOXP3 was reported to use the N-terminal region to recruit a wide range of cofactors (including RORα/RORγt[33,34], TRIM28 (also known as KAP1)[35] and TIP60, HDAC7[36] and Eos[37]), which may confer FOXP3 with the ability to assemble both repressive and activating complexes depending on DNAand chromatin context[14].

Adjacent to the N-terminal region, the ZF confers weak DNA affinity, contributing to FKH domain-mediated DNA binding[38], whereas the CC mediates FOXP3 homodimerization through an antiparallel CC interface[39]. The flexible RBR loop that connects the CC and FKH domains is enriched in aromatic and hydrophobic residues. Although initially identified as the RUNX1-interaction region[38,40,41], subsequent work has shown that the RBR also functions as a key oligomerization interface through its hydrophobic residues[42,43].

Finally, the C-terminal FKH domain serves as the sole sequence-specific DNA-binding module of FOXP3. Members of the FKH family, which includes over 50 TFs in humans, share a conserved winged-helix fold, in which helix 3 inserts into the major groove of DNA to recognize the FKH consensus motif (FKHM[44]; Fig. 1b). Although multiple versions of FKH conformation have been proposed (see below), all use the same helix 3 for DNA sequence recognition[38,45,46]. Beyond canonical DNA binding, the FKH domain also mediates protein–protein interactions, including cooperative self-assembly[42,43] and complex formation with the TF NFAT[45,47]. Through these interactions, the FKH domain can promote multimerization and structural organization on chromatin, making the FOXP3 FKH domain a unique example for noncanonical, multifunctional DNA-binding domains that integrate DNA recognition, cofactor interaction and higher-order assembly within a single structural fold.

So far, five distinct structures of FOXP3 in complex with DNA have been reported: a domain-swap (DS) dimer in complex with an isolated FKHM[45,46], a head-to-head (H-H) dimer bound to an inverted-repeat FKHM (IR-FKHM)[38] and three head-to-tail (H-T) multimeric assemblies associated with T2G, T3G and T4G tandem repeats[42,43]. Among these, the DS dimer forms independently of DNA, whereas the others arise through sequence-specific DNA recognition, providing mechanistic insights into the sequence specificity of FOXP3.

DS dimer

Unlike FKH domains of other TFs that fold into individual winged-helix monomers (Fig. 1b), the first crystal structure of FOXP3 FKH revealed an unusual DS dimer[45,46] (Fig. 1c). In this conformation, two FKH monomers

https://doi.org/10.1038/s41590-026-02483-w

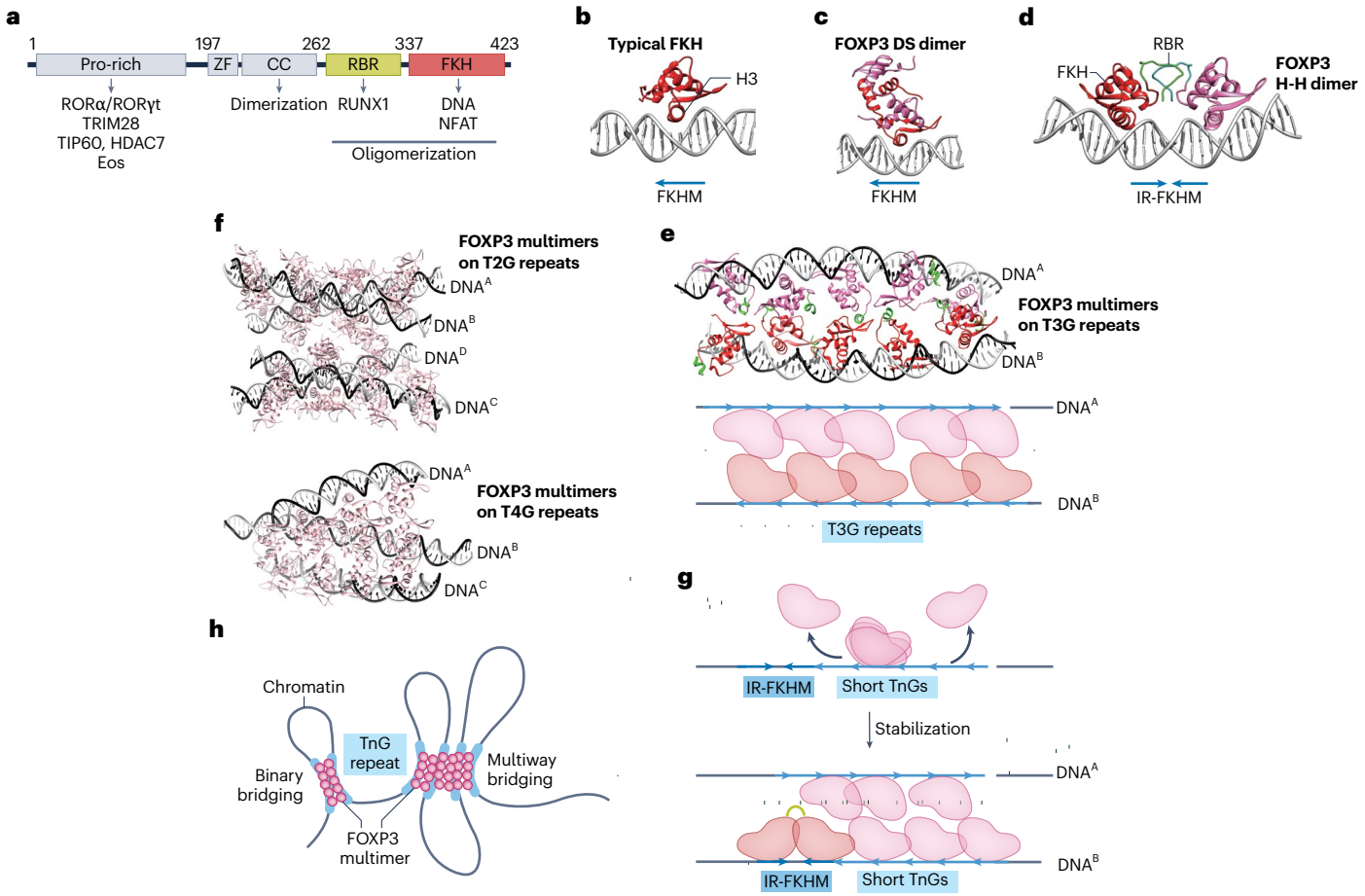

Fig. 1 | Diverse multimeric states of FOXP3 on DNA. a, Domain architecture of FOXP3, composed of an N-terminal proline-rich region, C2H2 ZF, CC, flexible loop region known as the RBR and a conserved FKH DNA-binding domain. Both the RBR and FKH domain mediate DNA-dependent oligomerization, whereas the CC mediates constitutive dimerization. Arrows indicate known interaction partners for individual domains. **b**, Most FKH domains, exemplified by FOXM1 (PDB: 3G73), adopt a monomeric winged-helix conformation in which the signature helix 3 (H3) inserts into the major groove to recognize the FKHM. **c**, A truncated FOXP3 FKH domain forms a DS dimer (PDB: 4WK8). **d**, Longer constructs containing the RBR adopt the canonical winged-helix conformation. Binding to IR-FKHM further enables H-H dimerization mediated by the RBR loop (green/cyan). The ZF and CC are present but are not resolved in the structure (PDB: 7TDW). **e**, FOXP3 forms multimers on extended T3G repeats (PDB: 8SRP), with multimerization mediated by both the FKH domain (red/pink) and RBR (green). The ZF and CC are present but are not resolved. Below is a schematic illustrating the higher-order assembly of FOXP3: FOXP3 subunits form linear arrays along individual T3G repeat DNA molecules, but these arrays are unstable. Stability is achieved when two (or more) arrays associate through protein–protein interactions, resulting in the bridging of separate DNA molecules. **f**, FOXP3 can assemble into various distinct higher-order structures on different TnG repeat DNA. Structures of FOXP3 multimers in complex with T2G repeat DNA (top, PDB: 9D2L) and T4G repeat DNA (bottom, PDB: 9D22) are shown. Note that two to four copies of DNA can be bridged through FOXP3 assemblies. **g**, FOXP3 can use nearby H-H motifs to nucleate or stabilize its multimerization on adjacent TnG repeat DNA. Because FOXP3 multimerization is repeat length dependent, short or suboptimal repeats form unstable assemblies. However, an adjacent H-H motif can act as an anchoring site, compensating for insufficient repeat length and stabilizing FOXP3 multimerization. **h**, FOXP3 multimerization reinforces chromatin architecture by connecting two genomic loci (binary DNA bridging) or multiple loci simultaneously (multiway DNA bridging). Although each individual bridging event may induce only subtle transcriptional changes, the cumulative effect of thousands of such events across the genome can globally modulate gene expression dynamics. This widespread architectural reinforcement may be particularly important under inflammatory or neonatal conditions, where the rapid establishment of new chromatin organization is likely required for appropriate transcriptional responses in T_{reg} cells.

form an intertwined dimeric structure by exchanging their helices composed of hydrophobic residues. Similar DS dimers have also been reported for the isolated FKHs of FOXP1 and FOXP2[47–49], leading to the widely accepted hypothesis that FOXP family TFs may have evolved to adopt a DS dimeric fold. Mutations within or near the DS interface (such as the human mutation M370I or structure-guided variants W348Q, M370T and A372P) lead to loss of FOXP3 function[29,45,50]. These results were interpreted as evidence for the physiological relevance of the domain-swapped dimer.

However, subsequent crystal structures of a longer FOXP3 construct containing the ZF, CC and RBR showed that the FOXP3 FKH actually adopts a nonswapped, monomeric winged-helix conformation[38] (Fig. 1d). In these structures, the same hydrophobic residues previously implicated in domain swapping are instead buried within the folded FKH core and stabilized by the RBR loop. Moreover, RBR–FKH constructs from FOXP1, FOXP2, FOXP3 and FOXP4 were all monomeric[38], suggesting that domain swapping arises artificially from truncation of the RBR, as the intact RBR normally stabilizes the monomeric fold. Accordingly, earlier 'swap-disrupting' mutations may have caused loss of function by destabilizing winged-helix protein folding rather than by disrupting a DS interaction. In keeping with this notion, mutations such as A372G/A372S, which suppress domain swapping without affecting folding, have minimal impact on FOXP3 activity[38]. By contrast, the IPEX-associated R337Q mutation promotes domain swapping and impairs FOXP3 function, which can be partially rescued by introducing A372G/A372S[38]. Together, these findings indicate that the physiological

conformation of FOXP3 is nonswapped and winged-helix, whereas the DS dimer represents a misfolded, nonfunctional state.

H-H dimer

In vitro DNA-binding studies revealed that FOXP3 strongly prefers DNA containing IR-FKHM with a 4-nucleotide spacer over isolated FKHM[38]. This sequence preference of FOXP3 is distinct from most other FKH TFs, including FOXP1, FOXP2 and FOXP4 (refs. 38,51,52), which bind to isolated FKHM without requiring the paired FKHM. Crystal structures showed that FOXP3 binds IR-FKHM as a H-H dimer with direct protein–protein contact between the two subunits[38] (Fig. 1d). These protein–protein interactions stabilize DNA binding and account for the preference of FOXP3 for IR-FKHM. Although the physiological relevance of this H-H binding mode was initially unclear, a recent study showed that IR-FKHM is enriched within FOXP3-occupied genomic loci and that H-H dimerization is physiologically relevant[51].

Pulldown sequencing (PD-seq) using purified FOXP3 and randomized short oligonucleotides further demonstrated that FOXP3 can recognize a broader range of sequences besides IR-FKHM using the same H-H dimeric binding mode[51]. These sequences include (1) TGTTT/TGTTG-like motifs similar to FKHM, and (2) GCAT-based motifs flanked by thymine or cytosine, provided that these sequences are paired with each other in the inverted configuration with a 4-nucleotide gap. As with IR-FKHM, these noncanonical paired motifs are enriched in FOXP3-occupied sites, further supporting the significance of FOXP3 H-H dimerization.

Structural and mutagenesis studies revealed that two distinct regions contribute to FOXP3 H-H dimerization[38,51]. The CC domain mediates constitutive dimerization and facilitates H-H assembly, but it alone is not sufficient and was not resolved in the crystal structures of the H-H dimer[38]. The flexible and hydrophobic RBR is the key determinant of H-H dimerization, promoting H-H dimeric assembly in a DNA sequence-dependent manner. Consistent with this, transplanting the RBR into other FOXP proteins is sufficient to confer H-H dimerization[51], underscoring that the RBR is a unique determinant of FOXP3 within the FOXP family. Together, these studies highlight how RBR, once thought to be a simple protein domain linker, functions as a modular domain with multifaceted roles in FKH domain folding, FOXP3 H-H dimerization and RUNX1 cofactor recruitment.

H-T multimers

Although FOXP3-bound genomic loci show enrichment of H-H motifs, these sites account for only ~10% of the genomic occupancy of FOXP3[51]. To identify additional binding modes missed by short-oligonucleotide PD-seq, longer genomic DNA fragments (~200–300 base pairs) were used for PD-seq[42]. The results revealed a striking enrichment of simple tandem TnG repeats (n = 2–5), a class of microsatellites present in about 18,000 sites within the human genome. Among these repeats, T3G arrays bind FOXP3 with the highest affinity, nearly equivalent to the affinity of FOXP3 for IR-FKHM[42]. FOXP3 binding to TnG repeats is also length dependent, requiring ~40 base pairs of T3G repeats for robust interaction[42]. The strong preference of FOXP3 for long TnG repeat microsatellites was independently validated by CNR-seq[13,29], ChIP–seq[13,30] and ChIP–exo[31]. Notably, allele-specific CNR-seq using F_1 hybrids of two divergent mouse strains revealed that FOXP3 occupancy strongly correlates with the length of TnG repeats[13,42,43], further demonstrating that TnG repeats not only are enriched but also function as drivers of FOXP3 genomic binding.

The cryo-electron microscopy (cryo-EM) structure of FOXP3 in complex with T3G repeats showed that FOXP3 forms H-T multimers on these sequences, with each subunit recognizing TGTTTGT in place of the canonical TGTTTAC motif[42] (Fig. 1e). The repetitive TGTTTGT sequence enables adjacent FOXP3 molecules to align into linear arrays stabilized by direct intersubunit protein contacts. However, isolated filaments on individual DNA were not observed; rather, two to four arrays associate through higher-order FOXP3–FOXP3 interactions, forming multimeric assemblies that bridge multiple DNA molecules (Fig. 1e,f). This suggests that TnG repeat recognition is inseparable from the assembly of the DNA bridging multimeric structure. Single-molecule studies showed that these bridged DNA structures are extremely stable in vitro[43], reflecting the strength of cooperative FOXP3 self-assembly.

FOXP3 also recognizes a broad spectrum of TnG repeat sequences (T2G, T3G, T4G and T5G repeats and their combinations) all by forming cooperative self-assemblies[43]. Because the periodicity of these repeats dictates the spacing and orientation of FOXP3 subunits, this raises an intriguing question: how does a single protein adopt multiple assembly modes with distinct geometries? Cryo-EM analyses of FOXP3 bound to T2G, T3G and T4G repeats revealed that FOXP3 acts as a highly adaptable building block, using 12 distinct intersubunit interfaces to accommodate variable spacing and orientations[43] (Fig. 1e,f). At the center of this plasticity lies the RBR, which mediates flexible 'fuzzy' hydrophobic interactions between adjacent subunits, enabling FOXP3 to assemble into diverse yet ultrastable higher-order complexes.

Consistent with the DNA bridging activity in vitro, chromatin conformation analyses (for example, high-throughput chromosome conformation capture (Hi-C) and high-throughput ChIP with Hi-C)[29,53] suggest that FOXP3 is involved in chromatin looping in T_{reg} cells. FOXP3-sufficient T_{reg} cells display a greater number of loops and increased loop stability compared to FOXP3-deficient counterparts[29,53], with loop anchors frequently containing FOXP3-bound TnG repeats[42]. Reporter assays using paired plasmids (one carrying an enhancer and the other a reporter gene) further demonstrated the capacity of FOXP3 to bridge DNA molecules and thereby facilitate transcription of target genes[43]. Although the magnitude of the effect of FOXP3 on chromatin looping and transcription is modest (typically less than two- to three-fold)[29,43], the widespread nature of these effects supports a model in which FOXP3 helps shape global chromatin architecture rather than act on specific target genes. Given that FOXP3 can simultaneously bridge two to four DNA molecules in vitro[43], it is also possible that FOXP3 can create hub-like structures to orchestrate multiple loci at once. In line with the importance of FOXP3 multimerization, mutations that disrupt multimerization (but not H-H dimerization) impair the ability of FOXP3 to induce T_{reg} cell-associated genes such as *CTLA4* and *CD25* when FOXP3 is ectopically expressed in conventional $CD4^+$ T cells[42,43].

Relationship between H-H dimers and H-T multimers

Both H-H sequences and TnG repeat sequences are enriched within FOXP3-occupied genomic regions, accounting for around 10% and 50% of bound sites, respectively. Intriguingly, approximately one-third of H-H motifs occur immediately adjacent to TnG repeats, and this juxtaposition, particularly when involving short or irregular repeat tracts, appears to enhance the otherwise limited stability of FOXP3 multimers[51]. These observations suggest that H-H dimers may act as nucleation sites for FOXP3 multimerization, initiating cooperative assembly that subsequently propagates along nearby repeat regions (Fig. 1g). In this way, FOXP3 may leverage its H-H dimerization to expand its sequence repertoire for multimerization and thus enhance its higher-order architectural function.

Emerging views and outstanding questions for FOXP3

Together, these findings support a model in which FOXP3 acts as an architectural regulator, shaping chromatin topology rather than functioning as a classical activator or repressor (Fig. 1h). Because FOXP3 primarily binds preaccessible regions marked by H3K27ac, individual looping events may not drive large transcriptional shifts. Instead, its binding and DNA bridging at thousands of sites across the genome may have subtle, yet global, modulatory effects on gene expression dynamics. This is consistent with the widespread yet modest transcriptional and topological effects observed across hundreds of loci, where most FOXP3-bound sites show little nearby transcriptional change, which is

similar to the effect of cohesin, a ring-shaped protein complex involved in DNA looping[54–56].

Future investigations should directly perturb TnG repeats (at single loci or across multiple sites) and combine these manipulations with high-resolution chromatin conformation mapping (for example, micro-C) and transcriptional analysis. Additionally, parallel imaging of FOXP3 multimers, target loci and transcriptional activity could provide single-cell-level insights into how FOXP3-dependent assemblies influence gene regulation and chromatin organization. Because only a subset of FOXP3-bound loci exhibits measurable transcriptional changes in a context-dependent manner, careful selection of loci for perturbation and experimental conditions will be critical for defining the functional contributions of FOXP3.

Equally important is to incorporate FOXP3 interaction partners into this model of its architectural function. Prior imaging studies revealed two types of FOXP3-harboring nuclear clusters: one colocalized with transcriptional activating factors and one with repressive factors[14]. Open questions include what determines these regulatory states, whether they reflect distinct FOXP3 conformations or differential co-occupancy with partner factors and how these complexes assemble, starting from direct partners to downstream effectors. Distinct modes of FOXP3 multimerization may facilitate or inhibit interactions with specific partners by selectively exposing different interaction surfaces. Further investigation is therefore required to elucidate the structural mechanisms underlying FOXP3 multimers in complex with activating factors versus repressive complexes.

More broadly, accumulating evidence indicates that FOXP3 is not unique in recognizing repetitive DNA and bridging chromatin. FOXP1/FOXP2/FOXP4 can also multimerize on TnG repeats and bridge DNA[42], and microsatellite binding may be a widespread property among eukaryotic TFs[57]. Determining whether these behaviors universally depend on multimerization and DNA bridging, as in FOXP3, will help define the general principles of architectural and transcriptional control.

AIRE

Immunological and transcriptional functions

The thymus is the primary organ where developing T cells acquire self-tolerance and establish immune homeostasis[58]. Within the thymic medulla, medullary thymic epithelial cells (mTECs) play a central role by expressing a vast array of peripheral tissue antigens (PTAs)[59], collectively representing 80–90% of protein-coding genes[60]. Through this promiscuous gene expression, mTECs present otherwise tissue-restricted antigens to maturing T cells, enabling the elimination of autoreactive clones by negative selection or their diversion into the T_{reg} cell lineage[61–63].

PTA expression in mTECs occurs through at least two distinct mechanisms during different stages of mTEC development. The first involves the transcriptional regulator AIRE, which promotes stochastic expression of PTAs in a subset of mTECs characterized by high levels of major histocompatibility complex class II[60,64,65]. AIRE-induced PTAs are displayed on the cell surface, allowing recognition and removal of autoreactive T cells[61–63]. The second mechanism operates in a distinct population of mTECs, namely mimetic cells, which do not express AIRE or have lost their expression of AIRE and instead rely on various lineage-defining TFs for PTA expression[66–70].

Although there is yet to be direct evidence that loss of a specific mimetic subset causes autoimmunity, perturbations of mTEC transcriptional programming have been shown to reshape mimetic cell composition, disrupt PTA expression and impair thymic mechanisms enforcing self-tolerance[71,72]. Importantly, loss or dysfunction of AIRE in either mice or humans results in systemic autoimmunity, most notably autoimmune polyendocrine syndrome type 1 (APS-1; also known as autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy)[61,73]. These genetic and immunological findings underscore the unique and nonredundant role of AIRE in immune homeostasis.

Since the discovery of AIRE as a key transcriptional regulator of PTA expression[61], its molecular function in mediating 'promiscuous' gene expression has drawn much attention from both immunological and mechanistic perspectives. However, mechanistic studies of AIRE have been challenging due to the rarity of AIRE$^+$ mTECs (~50,000 isolatable cells per mouse) and the transient nature of AIRE expression during mTEC maturation[74]. Consequently, most mechanistic studies have relied on ectopic expression of AIRE in model cell lines, such as 293T, 1C6 and 4D6 cells[75–79].

These model cell systems have successfully recapitulated certain key aspects of AIRE biology, including promiscuous gene activation, nuclear condensate formation, interaction with transcriptional cofactors and disease-associated mutation phenotypes. However, the specific sets of AIRE-bound or AIRE-regulated genes vary between systems, likely because AIRE recognizes chromatin states rather than specific DNA sequences[80,81]. Additionally, AIRE condensate number, size and localization can be sensitive to AIRE expression level and cellular context, and these factors should be considered when interpreting results from ectopic expression models. Nevertheless, these models have enabled detailed mechanistic investigation, but their divergent chromatin landscapes underscore the importance of eventually validating key findings in native mTECs. Recent advances in thymic organoid systems have established conditions to generate AIRE-expressing medullary-like TECs in vitro, while further validations are anticipated, providing new opportunities to dissect AIRE-mediated transcription and T cell selection[82–85].

Domain architecture and insights from APS-1 mutations

AIRE is a 57.5-kDa protein (545 amino acids long in humans) composed of multiple conserved domains (Fig. 2a): an N-terminal caspase activation and recruitment domain (CARD) that mediates homotypic polymerization[78], an Sp100, AIRE, NucP41/NucP75 and DEAF-1 (SAND) domain that interacts with the epigenetic regulatory complex ATF7IP (also known as MBD1)[86] and two tandem plant homeodomain (PHD) fingers, with PHD1 specifically recognizing H3K4me0 (refs. 80,81). The C-terminal tail (CTT) functions as a transactivation domain that recruits transcriptional coactivators[79,87].

Mutations associated with APS-1 span all these domains and provide key mechanistic insights[88]. Point mutations in the CARD (for example, L13P, T16M and L28P) abolish the polymerization activity of AIRE, resulting in recessive loss of function[78,89]. By contrast, mutations in the SAND or PHD1 domain (for example, G228W, R247C, C302Y and C311Y) interfere with AIRE polymerization dynamics, enabling mutant proteins to form inactive nuclear aggregates that frequently colocalize with PML nuclear bodies[78]. Because these mutants retain CARD-mediated polymerization, coexpression with wild-type AIRE leads to co-polymerization into mislocalized, dysfunctional condensates, resulting in dominant-negative effects that are consistent with dominant inheritance in individuals with APS-1[78,89,90]. Finally, mutations in the CTT (for example, pseudoexon insertion at D502[91] or missense mutation P539L[87,92]) also disrupt AIRE function, likely through impairment of the interactions of AIRE with transcriptional coactivators. Unlike SAND or PHD1 mutants, these CTT variants do not interfere with wild-type AIRE, leading to recessive inheritance of APS-1[91,92]. Collectively, these genotype–phenotype relationships point to distinct mechanisms by which each domain contributes to AIRE functions.

AIRE condensate formation

At the center of AIRE function is its unusual ability to polymerize into discrete nuclear foci or 'condensates' (Fig. 2b). Although some nuclear condensates form by liquid–liquid phase separation[93], AIRE condensates arise instead from CARD-mediated polymerization and, in their functional state, do not colocalize with other known nuclear bodies, such as nucleoli, nuclear speckles or PML bodies[94–96]. Early microscopy experiments suggested that AIRE condensates are devoid of active

https://doi.org/10.1038/s41590-026-02483-w

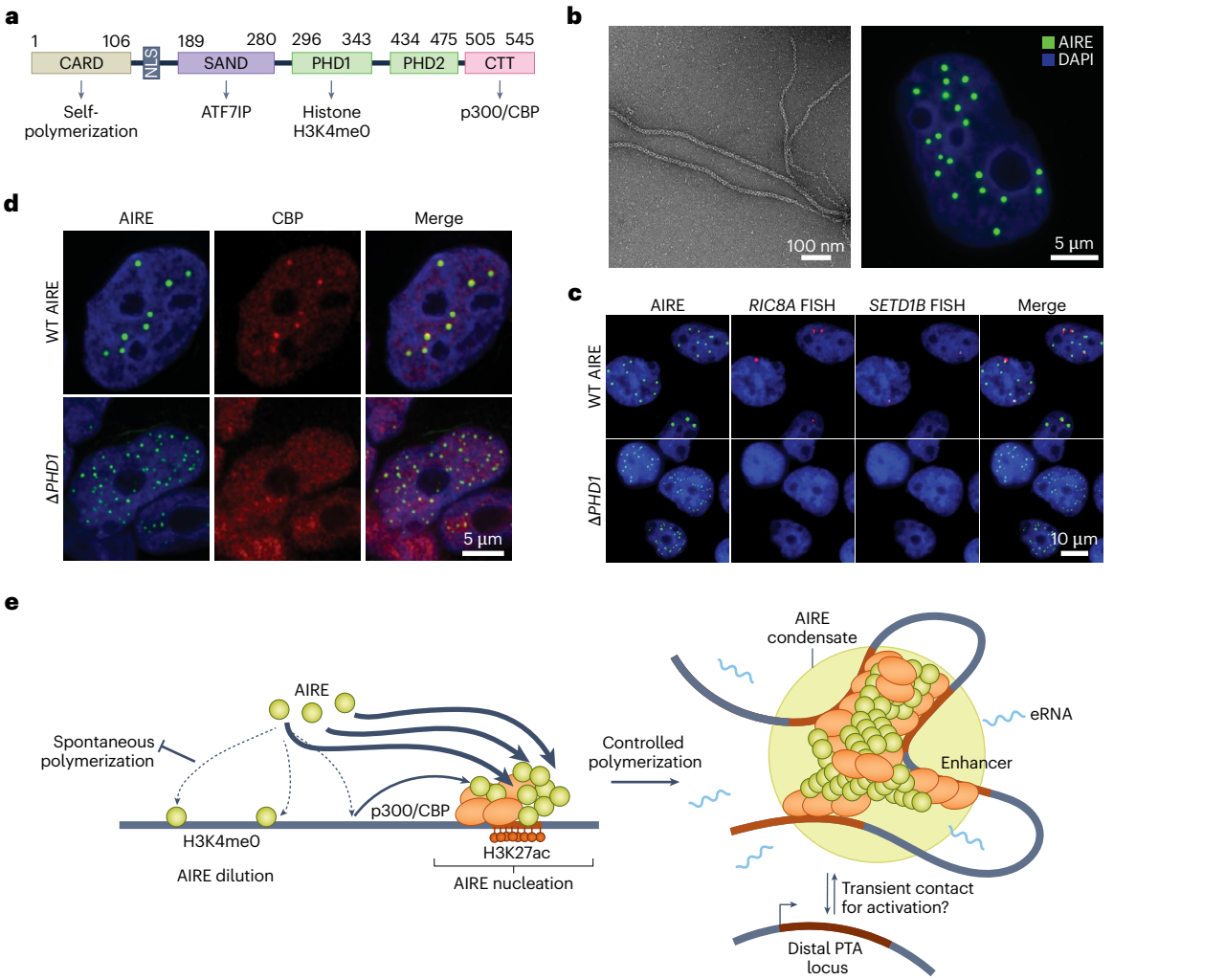

Fig. 2 | Coordinated mechanisms controlling AIRE polymerization. a, Domain architecture of AIRE, including the N-terminal CARD, SAND domain, two tandem PHDs (PHD1 and PHD2) and the CTT. Arrows indicate known interaction partners for individual domains; NLS, nuclear localization signal. **b**, Electron micrograph of isolated human AIRE CARD filaments (left) and AIRE condensates (right) in the nucleus of a 4D6 thymic epithelial cell. **c**, Nuclear condensates formed by wild-type (WT) AIRE correspond to sites of transcriptional activation, as shown by nascent RNA FISH signals for AIRE-induced genes *RIC8A* and *SETD1B*. By contrast, although the Δ*PHD1* mutant also forms nuclear condensates, these structures fail to activate *RIC8A* and *SETD1B*. This loss of function reflects uncontrolled AIRE polymerization in the absence of PHD1-mediated regulation. **d**, Wild-type AIRE condensates are enriched with CBP, whereas Δ*PHD1* condensates lack this association. **e**, Coordinated mechanisms controlling AIRE polymerization and condensate formation. After nuclear entry, the PHD1 of AIRE binds broadly to H3K4me0 across chromatin, preventing spontaneous CARD polymerization outside target enhancers. This inhibition-by-dilution maintains AIRE in a dispersed state until it is recruited to p300/CBP-rich loci, where the local concentration of AIRE overcomes PHD1-mediated inhibition. Through its CTT, AIRE interacts with coactivators p300/CBP and is guided to enhancer sites marked by H3K27ac, which serve as nucleation points for CARD-mediated polymerization. Polymerization creates a positive feedback loop that recruits additional AIRE and p300/CBP molecules, linking multiple AIRE-bound loci into transcriptional condensates. Although AIRE condensates represent active sites of transcription, their overlap with AIRE-induced PTA loci is limited. We propose that AIRE condensates form transient, long-range intra- and interchromatin contacts that enable activation of distal PTA genes. eRNA, enhancer RNA. The left image in **b** was reproduced from ref. 78, Springer Nature Limited. The right image in **b** and images in **c** and **d** were reproduced from ref. 79, Springer Nature Limited.

chromatin[97], implicating that they might function as storage depots. However, subsequent biochemical and genomic analyses revealed that AIRE interacts with multiple transcriptional coactivators[75,97,98] and binds to preactive, preaccessible genomic loci, such as superenhancers[99]. Immunofluorescence combined with nascent RNA fluorescence in situ hybridization (FISH) imaging showed that AIRE condensates in fact colocalize with these AIRE-bound enhancer regions and, in some cases, simultaneously engage with multiple enhancer sites from distinct chromosomes[79]. This colocalization coincides with transcriptional activation at these loci (Fig. 2c), demonstrating that AIRE condensates serve as active sites of transcription or 'transcriptional hubs'. However, not all AIRE condensates are functionally equivalent. Several disease-causing AIRE mutants still form condensates but lack transcriptional activity[78,79,89], indicating that polymerization alone is insufficient for function and that the proper assembly mechanisms must exist to ensure that AIRE polymerizes at the correct genomic or nuclear sites.

An important clue into the assembly mechanisms regulating AIRE CARD polymerization came from the surprising finding that deleting the CTT abolished AIRE condensate formation[79]. This was unexpected because this deletion construct still contained an intact CARD, which had previously been thought to be sufficient to drive polymerization. Coimmunoprecipitation followed by mass spectrometry and nuclear magnetic resonance analyses revealed that the CTT of AIRE directly interacts with the highly conserved transcriptional coactivators p300 and CREB-binding protein (CBP)[79]. Functional analyses, including mutational studies and use of p300/CBP inhibitors, suggest that disrupting these interactions prevents both AIRE condensate formation and its transcriptional activity[79]. Notably, unlike traditional TFs that

recruit coactivators to genomic loci, AIRE instead exploits p300/CBP already bound to active enhancers to direct its chromatin targeting. In support of this notion, ChIP–seq analyses of p300 before ectopic AIRE expression in 4D6 cells showed a strong correlation between the pre-existing genomic occupancy of p300 and the AIRE binding profile, and this correlation was compromised following CTT deletion[79].

How, then, does this recruitment lead to AIRE polymerization? Although the mechanism remains unclear, one possible model is that recruitment of multiple AIRE molecules to p300/CBP-rich enhancers increases the local concentration of AIRE, nucleating CARD-mediated polymerization. This process could join multiple AIRE-bound or p300/CBP-bound loci together, amplifying and stabilizing AIRE condensates. Consistent with this model, ChIP–seq analyses of an AIRE variant lacking the CARD revealed similar genomic specificity as wild-type AIRE but with reduced occupancy at AIRE target loci[79]. Importantly, despite partially retaining enhancer targeting, the CARD deletion variant fails to activate PTA transcription, indicating that chromatin localization alone is insufficient and that productive transcription requires CARD-dependent condensate assembly at these sites.

Microscopy studies also showed that p300 and CBP are densely enriched within AIRE condensates, appearing as p300/CBP condensates that are larger and more prominent than those in the absence of AIRE[79] (Fig. 2d). These findings indicate that AIRE not only assembles on p300/CBP-rich loci but also recruits additional p300/CBP molecules. Together, these results support a model in which CTT-mediated recruitment of AIRE to p300/CBP-rich chromatin triggers AIRE polymerization at target enhancers, forming transcriptionally active condensates through a positive feedforward loop between AIRE and p300/CBP.

Regulation of AIRE condensate assembly

Like other CARD-containing proteins, AIRE must prevent spontaneous polymerization to avoid forming aberrant polymers, while also enabling self-assembly at transcriptionally active enhancers. The observation that CTT deletion abolishes AIRE polymerization, even though isolated CARD can spontaneously polymerize, suggests that AIRE requires both a positive nucleation signal from the CTT at enhancer sites and a mechanism that restrains uncontrolled polymerization elsewhere in the nucleus.

Domain deletion analyses pointed to PHD1 as a critical autoinhibitory module for the AIRE CARD[79]. As described above, deleting the CTT alone disrupts AIRE condensate formation, but deleting both PHD1 and CTT restored polymerization, indicating that PHD1 restricts AIRE polymerization. Consistent with this, directly fusing AIRE CARD to PHD1 suppressed polymerization, demonstrating that PHD1 is necessary and sufficient to restrain CARD polymerization. Further protein engineering experiments suggest that the ability of PHD1 to bind unmodified H3K4me0 is essential for inhibiting CARD[79]. These findings support a model in which PHD1 anchors AIRE diffusely across chromatin through recognition of H3K4me0, a pervasive basal chromatin mark[100,101] that effectively dilutes local AIRE concentration, thereby suppressing spontaneous CARD polymerization. This 'inhibition-by-dilution' mechanism ensures that AIRE condensate formation only occurs under specific chromatin contexts, such as p300/CBP-dense enhancers, that can locally concentrate AIRE to overcome its basal chromatin-mediated dispersal. Intriguingly, a similar regulatory strategy was recently described for FOXP2, where nonspecific DNA binding suppresses spontaneous polymerization of its poly(Q) domain[102]. This suggests that inhibition-by-dilution might represent a more general mechanism for preventing aberrant polymerization by TFs.

Recognizing PHD1 as an autoinhibitory module provides new insight into how mutations within this domain disrupt AIRE function. Deletion of *PHD1* causes spontaneous polymerization independent of chromatin binding, leading to complete loss of function in 4D6 cells[79] (Fig. 2c,d). Likewise, APS-1-associated PHD1 mutations cause mislocalized AIRE condensates in 4D6 cells that sequester coexpressed wild-type protein[78], consistent with their dominant-negative and often hypomorphic phenotype in humans. This differs from APS-1 mutations in CARD and CTT, which are generally recessive and highly penetrant[89,103], demonstrating a wide range of APS-1 clinical presentations. Interestingly, a point mutation in PHD1 that disrupts H3K4me0 binding only partially impairs the transcriptional activity of AIRE in mice[104]. The milder phenotype of this mutant, compared to the complete loss of function seen following *PHD1* deletion, may reflect residual PHD1 function or additional regulatory mechanisms present in mTECs but absent in cell line models or a differential reliance on PHD1 for human versus mouse AIRE.

Collectively, these findings reveal a multilayered regulatory mechanism governing AIRE condensate assembly. Following nuclear entry, PHD1 anchors AIRE diffusely through H3K4me0 binding, maintaining AIRE at a low local concentration and preventing spurious polymerization via inhibition-by-dilution. When AIRE CTT engages p300/CBP-enriched genomic regions, local AIRE concentration increases, overcoming PHD1-mediated restraint and promoting CARD polymerization. This coordination between PHD1, CTT and CARD allows for controlled polymerization of AIRE and consequent assembly of transcriptional hubs at specific sites (Fig. 2e).

From condensates to PTA gene activation

It is important to note that, similar to FOXP3, the relationship between AIRE-bound genomic sites and the genes induced by AIRE remains poorly understood. Sites with strong AIRE ChIP–seq signals mainly correspond to active enhancers already occupied by p300/CBP before AIRE expression; these regions colocalize with AIRE condensates but show only modest fold change induction due to their pre-existing transcriptional activity[79]. By contrast, PTA genes, which are initially silent but strongly induced by AIRE, have low AIRE occupancy just above background levels[105]. Although the specific loci belonging to these two groups (AIRE-bound versus AIRE-induced genes) differ across AIRE expression systems, this dichotomy is consistently observed[79,99,105], suggesting shared underlying mechanisms.

This limited overlap between AIRE-bound enhancers and AIRE-induced PTAs raises a key mechanistic question: what roles do AIRE condensates have in PTA induction? AIRE condensates form at AIRE-bound enhancers, yet all domains required for condensate formation, including CARD and CTT, are also essential for AIRE-mediated activation of PTAs[79], suggesting that AIRE leverages its condensates to promote PTA expression. One possibility is that PTA induction occurs indirectly, perhaps through transcriptional cascades, chromatin reorganization or redistribution of limiting coactivators like p300/CBP and Mediator. Alternatively, PTA loci may make transient and infrequent contacts with AIRE condensates, exploiting the transcriptionally permissive environment of AIRE condensates to activate gene expression (Fig. 1e). Such infrequent interactions in a small number of cells could yield strong bursts in transcription that remain barely detectable in averaged ChIP profiles. This is consistent with the findings that AIRE-dependent PTA transcription is stochastic, with most PTAs each expressed in only a small fraction (<10–20%) of mTECs[60,64,65]. In addition, these interactions may occur over long distances that could further complicate detection by methods like ChIP–seq or Hi-C. For instance, at the *Sox2* locus in embryonic stem cells, transcriptional condensates can act over distances up to ~0.5–1 µm (ref. 106).

Other previously proposed mechanisms of AIRE-dependent gene activation can also be integrated into this condensate model. For example, AIRE-induced PTAs often show transcription start site-proximal pausing of RNA polymerase II, and AIRE releases RNA polymerase II, either directly or indirectly, through recruitment of P-TEFb[77,107]. Given that AIRE condensates are enriched in Mediator[79], which can engage the super elongation complex containing P-TEFb[108], transient association with AIRE condensates may therefore facilitate transcriptional

https://doi.org/10.1038/s41590-026-02483-w

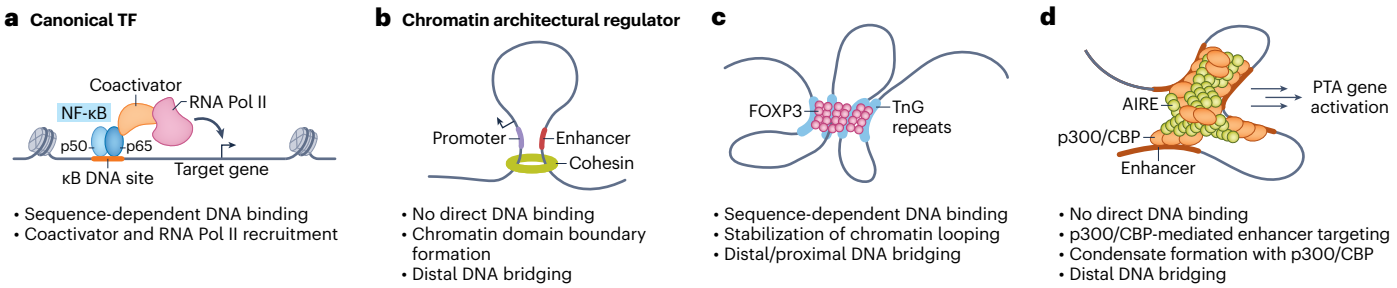

Fig. 3 | FOXP3 and AIRE use distinct architectural modes of transcriptional regulation. a, Canonical TFs, exemplified by NF-κB (p50/p65), bind specific DNA motifs and directly recruit coactivators and RNA polymerase (Pol) II to promote target gene transcription. **b**, By contrast, chromatin architectural regulators do not directly control transcription through recruitment of coactivators or suppressors to target sites. Instead, factors such as cohesin influence transcription by forming chromatin loops, which may include enhancer–promoter loops. **c**, FOXP3 binds dispersed TnG microsatellite repeats across the genome and assembles into multimers that bridge distinct DNA regions on a global scale. Through microsatellite binding and higher-order assembly, FOXP3 stabilizes chromatin looping to indirectly shape transcriptional programs in T_{reg} cells. **d**, AIRE does not function through classical sequence-specific DNA recognition. Instead, it is recruited to active enhancers enriched for p300/CBP. There, AIRE assembles into transcriptional condensates with further recruitment of p300/CBP. These higher-order assemblies mediate distal DNA bridging and enable PTA gene activation through mechanisms that remain to be determined.

elongation. Additional mechanisms (such as the generation of localized DNA damage, potentially through topoisomerase activity[99] or alterations in RNA splicing[75,109,110]) may also fit within this model, as AIRE condensates may harbor various chromatin and splicing regulators and DNA repair machinery reported to interact with AIRE[75,109].

PTA selectivity

Another unresolved issue concerns PTA specificity: what determines which genes are activated by AIRE through this stochastic activation mechanism? Because AIRE lacks a canonical sequence-specific DNA-binding domain, its target gene specificity likely depends on chromatin states or structures rather than DNA sequences. Analyses of genomic features near AIRE-induced genes have yielded several hypotheses. These include H3K4me0, recognized by the AIRE PHD1 domain[80,81], and other repressive marks such as H3K27me3 or H3K9me3 (the latter generated by ATF7IP–SETDB1 (ref. 86)), both enriched at AIRE-dependent PTA loci[60,111]. However, not all regions carrying these marks are AIRE responsive, suggesting involvement of additional chromatin or topological cues. Other studies found that many AIRE target genes are bound by RNA polymerase II but are transcriptionally inactive because the polymerase is stalled in a promoter-proximal paused state[77]. This observation raises the possibility that AIRE may use promoter-proximal pausing as one feature for PTA selectivity.

Recent computational analyses have further refined the mechanisms that govern the choice of AIRE target genes. Promoters of AIRE-induced genes are enriched for $(CA)_n$ repeats, which can adopt left-handed Z-DNA conformations[105]. Because these regions are prone to torsional stress and localized double-strand breaks, Z-DNA formation frequently coincides with DNA repair and topoisomerase complexes, some of which interact directly with AIRE[105]. Z-DNA formation at these promoters correlates strongly with AIRE-dependent transcriptional activation[105], leading to the model that AIRE may exploit chromatin features (for example, bound DNA repair factors) and/or DNA structural features (Z-DNA) to identify its targets. In parallel, chromatin accessibility outside the main accessible regions, so called accessibility 'noise', was also implicated in guiding the AIRE target selection. A recent study[112] showed that mTECs repress p53 before AIRE expression and that this p53 suppression indirectly increases 'noise' in chromatin accessibility. These noisy regions are enriched for nucleosome-disfavored AT-rich tracts and flank many AIRE-dependent PTAs, where elevated accessibility noise predicts ectopic expression.

Exactly which combination of these chromatin features enables AIRE to selectively target PTAs and the biochemical basis for this specificity remain to be determined. Even so, current evidence highlights the unconventional nature of AIRE's mode of action.

AIRE and mimetic cells

Beyond its immediate molecular functions in condensate formation and PTA activation, AIRE also appears to influence the broader developmental trajectory of mTECs (see ref. 113 for an extensive review). AIRE-expressing mTECs give rise to most 'mimetic' cells, which partially resemble peripheral tissues in morphology and gene expression[66–70]. As AIRE is downregulated during this transition, these post-AIRE mTECs begin expressing lineage-specific TFs that drive tissue-restricted gene expression, thereby producing PTAs through mechanisms similar to those in their peripheral counterparts.

Although AIRE influences mTEC maturation and the development of some mimetic cell subsets (for example, microfold, corneo, ciliated and neuroendocrine mTECs), its precise role in these processes remains unclear. The expression of some of the lineage-specific TFs in mimetic cells is modulated by AIRE but is not absolutely dependent on it[68–70]. One possibility is that AIRE activity primes chromatin states that facilitate the action of lineage-defining TFs, making prior AIRE expression beneficial for subsequent mimetic differentiation. However, sustained AIRE expression may not be required and could even be incompatible with the differentiation process itself. Clarifying how the transient activity of AIRE affects mimetic cell development will be an important goal for future work.

AIRE beyond the thymus

In addition to its well-established role in mTECs, AIRE is also expressed in rare extrathymic antigen-presenting cell (APC) populations, collectively referred to as extrathymic AIRE-expressing cells (eTACs)[114,115]. These eTACs reside in secondary lymphoid organs and comprise distinct hematopoietic APC subsets, including dendritic cell-like populations as well as RORγt$^+$ APCs[115–117]. Functionally, extrathymic AIRE expression has been linked to the regulation of peripheral tolerance through interactions with autoreactive T cells and the promotion of tolerogenic immune programs[115]. Notably, AIRE expression in RORγt$^+$ APCs has also been shown to contribute to host defense, particularly mucosal antifungal immunity[118]. Despite these emerging functional insights, the molecular functions and mechanisms by which AIRE operates in eTACs remain poorly defined. Interestingly, AIRE also forms nuclear speckles in eTACs[118], suggesting that its polymerization and condensate-forming properties are conserved features of its activity.

Emerging views and outstanding questions for AIRE

Although recent discoveries have reshaped our understanding of AIRE, several fundamental questions remain unresolved. How AIRE condensates that formed at preactive enhancers activate distal PTA loci is still unclear. Similarly, reconciling the apparent disconnect

between genomic loci with high AIRE occupancy and PTA loci exhibiting AIRE-dependent expression will require single-cell and spatial analyses capable of capturing transient or stochastic enhancer–promoter contacts. Approaches such as spatial transcriptomics, multiplexed imaging of nascent RNA and high-resolution three-dimensional genome-mapping techniques could help visualize how AIRE-driven transcriptional hubs interact with AIRE-dependent targets in situ. Likewise, live-cell microscopy tracking AIRE and its cofactors over time may illuminate the dynamics of condensate nucleation, maturation and dissolution. Progress in developing physiologically relevant model systems, such as thymic organoids[82–85] or induced pluripotent stem cell-derived thymic epithelial cells[119], will further enable these questions to be addressed in more native chromatin contexts. Together, these questions underscore the complexity of AIRE biology and highlight the need for continued mechanistic dissection of how AIRE coordinates gene regulation, cell fate and immune tolerance.

Concluding perspectives

FOXP3 and AIRE represent distinct yet converging paradigms for enforcing immune tolerance through unconventional modes of transcriptional regulation that rely on higher-order molecular assemblies. Rather than functioning as classical TFs that bind discrete sequences near promoters or enhancers to directly activate or repress nearby genes (Fig. 3a), we propose that both FOXP3 and AIRE operate as architectural regulators more like cohesin (Fig. 3b) that reshape the transcriptional landscape through multimerization and chromatin organization.

FOXP3 binds thousands of TnG microsatellite repeats dispersed across the genome, assembling into multimers that act as molecular scaffolds to bridge distal regulatory elements, thereby stabilizing chromatin loops and reinforcing pre-existing regulatory architectures (Fig. 3c). Despite inducing minimal changes to local histone marks or chromatin accessibility, these subtle architectural effects of FOXP3 may collectively modulate the global transcriptional network that governs T_{reg} cell development and function, thereby ensuring immune homeostasis.

AIRE also leverages pre-existing chromatin landscapes by binding p300/CBP-rich active enhancers, consolidating these loci into transcriptional condensates densely enriched with coactivators (Fig. 3d). Although these condensates induce little local change in histone modifications or accessibility, they promote stochastic activation of distal PTA loci, potentially through transient physical contacts. PTA selectivity has also been linked to low-complexity DNA elements, such as $(CA)_n$ and AT-rich sequences, although how they are recognized either directly or indirectly by AIRE remains unclear.

These mechanisms, although divergent but centering on multimerization, may not be unique to FOXP3 and AIRE. Other transcriptional regulators, including members of the SP100 family to which AIRE belongs, contain an analogous polymerization domain and form nuclear condensates[120]. Furthermore, microsatellite binding appears to be a widespread property among eukaryotic TFs[42,57]. As we uncover more about the architectural principles underlying transcriptional regulators, FOXP3 and AIRE may serve as compelling models for gene regulation.

References

1. Fontenot, J. D., Gavin, M. A. & Rudensky, A. Y. FoxP3 programs the development and function of $CD4^+CD25^+$ regulatory T cells. *Nat. Immunol.* **4**, 330–336 (2003).
2. Hori, S., Nomura, T. & Sakaguchi, S. Control of regulatory T cell development by the transcription factor *Foxp3*. *Science* **299**, 1057–1061 (2003).
3. Khattri, R., Cox, T., Yasayko, S. A. & Ramsdell, F. An essential role for Scurfin in $CD4^+CD25^+$ T regulatory cells. *Nat. Immunol.* **4**, 337–342 (2003).
4. Brunkow, M. E. et al. Disruption of a new forkhead/winged-helix protein, scurfin, results in the fatal lymphoproliferative disorder of the scurfy mouse. *Nat. Genet.* **27**, 68–73 (2001).
5. Chatila, T. A. et al. *JM2*, encoding a fork head-related protein, is mutated in X-linked autoimmunity-allergic disregulation syndrome. *J. Clin. Invest.* **106**, R75–R81 (2000).
6. Bennett, C. L. et al. The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of *FOXP3*. *Nat. Genet.* **27**, 20–21 (2001).
7. Wildin, R. S. et al. X-linked neonatal diabetes mellitus, enteropathy and endocrinopathy syndrome is the human equivalent of mouse scurfy. *Nat. Genet.* **27**, 18–20 (2001).
8. Bacchetta, R. & Roncarolo, M. G. IPEX syndrome from diagnosis to cure, learning along the way. *J. Allergy Clin. Immunol.* **153**, 595–605 (2024).
9. Lin, W. et al. Regulatory T cell development in the absence of functional Foxp3. *Nat. Immunol.* **8**, 359–368 (2007).
10. Gavin, M. A. et al. Foxp3-dependent programme of regulatory T-cell differentiation. *Nature* **445**, 771–775 (2007).
11. Hill, J. A. et al. Foxp3 transcription-factor-dependent and -independent regulation of the regulatory T cell transcriptional signature. *Immunity* **27**, 786–800 (2007).
12. Hill, J. A. et al. Retinoic acid enhances Foxp3 induction indirectly by relieving inhibition from $CD4^+CD44^{hi}$ cells. *Immunity* **29**, 758–770 (2008).
13. van der Veeken, J. et al. The transcription factor Foxp3 shapes regulatory T cell identity by tuning the activity of *trans*-acting intermediaries. *Immunity* **53**, 971–984 (2020).
14. Kwon, H.-K., Chen, H.-M., Mathis, D. & Benoist, C. Different molecular complexes that mediate transcriptional induction and repression by FoxP3. *Nat. Immunol.* **18**, 1238–1248 (2017).
15. Marson, A. et al. Foxp3 occupancy and regulation of key target genes during T-cell stimulation. *Nature* **445**, 931–935 (2007).
16. Zheng, Y. et al. Genome-wide analysis of Foxp3 target genes in developing and mature regulatory T cells. *Nature* **445**, 936–940 (2007).
17. Yoshida, H. et al. The *cis*-regulatory atlas of the mouse immune system. *Cell* **176**, 897–912 (2019).
18. Zemmour, D. et al. Single-cell analysis of FOXP3 deficiencies in humans and mice unmasks intrinsic and extrinsic $CD4^+$ T cell perturbations. *Nat. Immunol.* **22**, 607–619 (2021).
19. Hu, W. et al. Temporal and context-dependent requirements for the transcription factor Foxp3 expression in regulatory T cells. *Nat. Immunol.* **26**, 2059–2073 (2025).
20. Jager, C. Inducible protein degradation reveals inflammation-dependent function of the T_{reg} cell lineage-defining transcription factor Foxp3. *Sci. Immunol.* **10**, eadr7057 (2025).
21. Arvey, A. et al. Inflammation-induced repression of chromatin bound by the transcription factor Foxp3 in regulatory T cells. *Nat. Immunol.* **15**, 580–587 (2014).
22. Peeters, J. G. C., Silveria, S., Ozdemir, M., Ramachandran, S. & DuPage, M. Hyperactivating EZH2 to augment H3K27me3 levels in regulatory T cells enhances immune suppression by driving early effector differentiation. *Cell Rep.* **43**, 114724 (2024).
23. DuPage, M. et al. The chromatin-modifying enzyme EZH2 is critical for the maintenance of regulatory T cell identity after activation. *Immunity* **42**, 227–238 (2015).
24. Yang, X. P. et al. EZH2 is crucial for both differentiation of regulatory T cells and T effector cell expansion. *Sci. Rep.* **5**, 10643 (2015).
25. Sarmento, O. F. et al. The role of the histone methyltransferase enhancer of zeste homolog 2 (EZH2) in the pathobiological mechanisms underlying inflammatory bowel disease (IBD). *J. Biol. Chem.* **292**, 706–722 (2017).
26. Zhang, Y. et al. The polycomb repressive complex 2 governs life and death of peripheral T cells. *Blood* **124**, 737–749 (2014).

27. Wang, D. et al. Targeting EZH2 reprograms intratumoral regulatory T cells to enhance cancer immunity. *Cell Rep.* **23**, 3262–3274 (2018).
28. Samstein, R. M. et al. Foxp3 exploits a pre-existent enhancer landscape for regulatory T cell lineage specification. *Cell* **151**, 153–166 (2012).
29. Liu, Z., Lee, D. S., Liang, Y., Zheng, Y. & Dixon, J. R. Foxp3 orchestrates reorganization of chromatin architecture to establish regulatory T cell identity. *Nat. Commun.* **14**, 6943 (2023).
30. Kitagawa, Y. et al. Guidance of regulatory T cell development by SATB1-dependent super-enhancer establishment. *Nat. Immunol.* **18**, 173–183 (2017).
31. Arora, S. et al. Joint sequence and chromatin neural networks characterize the differential abilities of forkhead transcription factors to engage inaccessible chromatin. Preprint at *bioRxiv* https://doi.org/10.1101/2023.10.06.561228 (2023).
32. DelRosso, N. et al. Large-scale mapping and mutagenesis of human transcriptional effector domains. *Nature* **616**, 365–372 (2023).
33. Du, J., Huang, C., Zhou, B. & Ziegler, S. F. Isoform-specific inhibition of RORα-mediated transcriptional activation by human FOXP3. *J. Immunol.* **180**, 4785–4792 (2008).
34. Ichiyama, K. et al. Foxp3 inhibits RORγt-mediated *IL-17A* mRNA transcription through direct interaction with RORγt. *J. Biol. Chem.* **283**, 17003–17008 (2008).
35. Huang, C. et al. Cutting edge: A novel, human-specific interacting protein couples FOXP3 to a chromatin-remodeling complex that contains KAP1/TRIM28. *J. Immunol.* **190**, 4470–4473 (2013).
36. Li, B. et al. FOXP3 interactions with histone acetyltransferase and class II histone deacetylases are required for repression. *Proc. Natl Acad. Sci. USA* **104**, 4571–4576 (2007).
37. Pan, F. et al. Eos mediates Foxp3-dependent gene silencing in $CD4^+$ regulatory T cells. *Science* **325**, 1142–1146 (2009).
38. Leng, F. et al. The transcription factor FoxP3 can fold into two dimerization states with divergent implications for regulatory T cell function and immune homeostasis. *Immunity* **55**, 1354–1369 (2022).
39. Song, X. et al. Structural and biological features of FOXP3 dimerization relevant to regulatory T cell function. *Cell Rep.* **1**, 665–675 (2012).
40. Ono, M. et al. Foxp3 controls regulatory T-cell function by interacting with AML1/Runx1. *Nature* **446**, 685–689 (2007).
41. Rudra, D. et al. Runx–CBFβ complexes control expression of the transcription factor Foxp3 in regulatory T cells. *Nat. Immunol.* **10**, 1170–1177 (2009).
42. Zhang, W. et al. FOXP3 recognizes microsatellites and bridges DNA through multimerization. *Nature* **624**, 433–441 (2023).
43. Leng, F. et al. Ultrastable and versatile multimeric ensembles of FoxP3 on microsatellites. *Mol. Cell* **85**, 1509–1524 (2025).
44. Dai, S., Qu, L., Li, J. & Chen, Y. Toward a mechanistic understanding of DNA binding by forkhead transcription factors and its perturbation by pathogenic mutations. *Nucleic Acids Res.* **49**, 10235–10249 (2021).
45. Bandukwala, H. S. et al. Structure of a domain-swapped FOXP3 dimer on DNA and its function in regulatory T cells. *Immunity* **34**, 479–491 (2011).
46. Chen, Y. et al. DNA binding by FOXP3 domain-swapped dimer suggests mechanisms of long-range chromosomal interactions. *Nucleic Acids Res.* **43**, 1268–1282 (2015).
47. Wu, Y. et al. FOXP3 controls regulatory T cell function through cooperation with NFAT. *Cell* **126**, 375–387 (2006).
48. Chu, Y. P. et al. Solution structure and backbone dynamics of the DNA-binding domain of FOXP1: insight into its domain swapping and DNA binding. *Protein Sci.* **20**, 908–924 (2011).
49. Stroud, J. C. et al. Structure of the forkhead domain of FOXP2 bound to DNA. *Structure* **14**, 159–166 (2006).
50. Van Gool, F. et al. A mutation in the transcription factor *Foxp3* drives T helper 2 effector function in regulatory T cells. *Immunity* **50**, 362–377 (2019).
51. Leng, F. et al. FoxP3 forms a head-to-head dimer *in vivo* and stabilizes its multimerization on adjacent microsatellites. *Cell Rep.* **44**, 116633 (2025).
52. Nakagawa, S., Gisselbrecht, S. S., Rogers, J. M., Hartl, D. L. & Bulyk, M. L. DNA-binding specificity changes in the evolution of forkhead transcription factors. *Proc. Natl Acad. Sci. USA* **110**, 12349–12354 (2013).
53. Ramirez, R. N., Chowdhary, K., Leon, J., Mathis, D. & Benoist, C. FoxP3 associates with enhancer–promoter loops to regulate T_{reg}-specific gene expression. *Sci. Immunol.* **7**, eabj9836 (2022).
54. Nora, E. P. et al. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* **169**, 930–944 (2017).
55. Ghavi-Helm, Y. et al. Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression. *Nat. Genet.* **51**, 1272–1282 (2019).
56. Rao, S. S. P. et al. Cohesin loss eliminates all loop domains. *Cell* **171**, 305–320 (2017).
57. Horton, C. A. et al. Short tandem repeats bind transcription factors to tune eukaryotic gene expression. *Science* **381**, eadd1250 (2023).
58. Xing, Y. & Hogquist, K. A. T-cell tolerance: central and peripheral. *Cold Spring Harb. Perspect. Biol.* **4**, a006957 (2012).
59. Derbinski, J., Schulte, A., Kyewski, B. & Klein, L. Promiscuous gene expression in medullary thymic epithelial cells mirrors the peripheral self. *Nat. Immunol.* **2**, 1032–1039 (2001).
60. Sansom, S. N. et al. Population and single-cell genomics reveal the *Aire* dependency, relief from Polycomb silencing, and distribution of self-antigen expression in thymic epithelia. *Genome Res.* **24**, 1918–1931 (2014).
61. Anderson, M. S. et al. Projection of an immunological self shadow within the thymus by the Aire protein. *Science* **298**, 1395–1401 (2002).
62. Malchow, S. et al. Aire enforces immune tolerance by directing autoreactive T cells into the regulatory T cell lineage. *Immunity* **44**, 1102–1113 (2016).
63. Yang, S., Fujikado, N., Kolodin, D., Benoist, C. & Mathis, D. Regulatory T cells generated early in life play a distinct role in maintaining self-tolerance. *Science* **348**, 589–594 (2015).
64. Brennecke, P. et al. Single-cell transcriptome analysis reveals coordinated ectopic gene-expression patterns in medullary thymic epithelial cells. *Nat. Immunol.* **16**, 933–941 (2015).
65. Meredith, M., Zemmour, D., Mathis, D. & Benoist, C. Aire controls gene expression in the thymic epithelium with ordered stochasticity. *Nat. Immunol.* **16**, 942–949 (2015).
66. Bautista, J. L. et al. Single-cell transcriptional profiling of human thymic stroma uncovers novel cellular heterogeneity in the thymic medulla. *Nat. Commun.* **12**, 1096 (2021).
67. Bornstein, C. et al. Single-cell mapping of the thymic stroma identifies IL-25-producing tuft epithelial cells. *Nature* **559**, 622–626 (2018).
68. Givony, T. et al. Thymic mimetic cells function beyond self-tolerance. *Nature* **622**, 164–172 (2023).
69. Michelson, D. A., Hase, K., Kaisho, T., Benoist, C. & Mathis, D. Thymic epithelial cells co-opt lineage-defining transcription factors to eliminate autoreactive T cells. *Cell* **185**, 2542–2558 (2022).
70. Miller, C. N. et al. Thymic tuft cells promote an IL-4-enriched medulla and shape thymocyte development. *Nature* **559**, 627–631 (2018).
71. Michelson, D. A., Zuo, C., Verzi, M., Benoist, C. & Mathis, D. HNF4 activates mimetic-cell enhancers to recapitulate gut and liver development within the thymus. *J. Exp. Med.* **220**, e20230461 (2023).

72. Sin, J. H. et al. Ikaros is a principal regulator of Aire^{+} mTEC homeostasis, thymic mimetic cell diversity, and central tolerance. *Sci. Immunol.* **8**, eabq3109 (2023).
73. Husebye, E. S., Anderson, M. S. & Kampe, O. Autoimmune polyendocrine syndromes. *N. Engl. J. Med.* **378**, 1132–1141 (2018).
74. Metzger, T. C. et al. Lineage tracing and cell ablation identify a post-Aire-expressing thymic epithelial cell population. *Cell Rep.* **5**, 166–179 (2013).
75. Abramson, J., Giraud, M., Benoist, C. & Mathis, D. Aire's partners in the molecular control of immunological tolerance. *Cell* **140**, 123–135 (2010).
76. Berger, A. H. et al. High-resolution transcriptional impact of AIRE: effects of pathogenic variants p.Arg257Ter, p.Cys311Tyr, and polygenic risk variant p.Arg471Cys. *Front. Immunol.* **16**, 1572789 (2025).
77. Giraud, M. et al. Aire unleashes stalled RNA polymerase to induce ectopic gene expression in thymic epithelial cells. *Proc. Natl Acad. Sci. USA* **109**, 535–540 (2012).
78. Huoh, Y. S. et al. Dual functions of Aire CARD multimerization in the transcriptional regulation of T cell tolerance. *Nat. Commun.* **11**, 1625 (2020).
79. Huoh, Y. S. et al. Mechanism for controlled assembly of transcriptional condensates by Aire. *Nat. Immunol.* **25**, 1580–1592 (2024).
80. Koh, A. S. et al. Aire employs a histone-binding module to mediate immunological tolerance, linking chromatin regulation with organ-specific autoimmunity. *Proc. Natl Acad. Sci. USA* **105**, 15878–15883 (2008).
81. Org, T. et al. The autoimmune regulator PHD finger binds to non-methylated histone H3K4 to activate gene expression. *EMBO Rep.* **9**, 370–376 (2008).
82. Lim, S. et al. Derivation of functional thymic epithelial organoid lines from adult murine thymus. *Cell Rep.* **43**, 114019 (2024).
83. Hubscher, T. et al. Thymic epithelial organoids mediate T-cell development. *Development* **151**, dev202853 (2024).
84. Major, V. et al. Establishment of a microwell-array-based miniaturized thymic organoid model suitable for high-throughput applications. *Cell Rep.* **44**, 115579 (2025).
85. Ramos, S. A. et al. Generation of functional thymic organoids from human pluripotent stem cells. *Stem Cell Reports* **18**, 829–840 (2023).
86. Waterfield, M. et al. The transcriptional regulator Aire coopts the repressive ATF7ip-MBD1 complex for the induction of immunotolerance. *Nat. Immunol.* **15**, 258–265 (2014).
87. Meloni, A., Incani, F., Corda, D., Cao, A. & Rosatelli, M. C. Role of PHD fingers and COOH-terminal 30 amino acids in AIRE transactivation activity. *Mol. Immunol.* **45**, 805–809 (2008).
88. Bruserud, O., Oftedal, B. E., Wolff, A. B. & Husebye, E. S. *AIRE*—mutations and autoimmune disease. *Curr. Opin. Immunol.* **43**, 8–15 (2016).
89. Oftedal, B. E. et al. Dominant mutations in the autoimmune regulator *AIRE* are associated with common organ-specific autoimmune diseases. *Immunity* **42**, 1185–1196 (2015).
90. Abbott, J. K. et al. Dominant-negative loss of function arises from a second, more frequent variant within the SAND domain of autoimmune regulator (*AIRE*). *J. Autoimmun.* **88**, 114–120 (2018).
91. Ochoa, S. et al. A deep intronic splice-altering *AIRE* variant causes APECED syndrome through antisense oligonucleotide-targetable pseudoexon inclusion. *Sci. Transl. Med.* **16**, eadk0845 (2024).
92. Meloni, A. et al. Delineation of the molecular defects in the *AIRE* gene in autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy patients from Southern Italy. *J. Clin. Endocrinol. Metab.* **87**, 841–846 (2002).
93. Sabari, B. R., Dall'Agnese, A. & Young, R. A. Biomolecular condensates in the nucleus. *Trends Biochem. Sci.* **45**, 961–977 (2020).
94. Bjorses, P. et al. Localization of the APECED protein in distinct nuclear structures. *Hum. Mol. Genet.* **8**, 259–266 (1999).
95. Rinderle, C., Christensen, H. M., Schweiger, S., Lehrach, H. & Yaspo, M. L. *AIRE* encodes a nuclear protein co-localizing with cytoskeletal filaments: altered sub-cellular distribution of mutants lacking the PHD zinc fingers. *Hum. Mol. Genet.* **8**, 277–290 (1999).
96. Su, M. A. et al. Mechanisms of an autoimmunity syndrome in mice caused by a dominant mutation in Aire. *J. Clin. Invest.* **118**, 1712–1726 (2008).
97. Pitkanen, J. et al. Cooperative activation of transcription by autoimmune regulator AIRE and CBP. *Biochem. Biophys. Res. Commun.* **333**, 944–953 (2005).
98. Yoshida, H. et al. Brd4 bridges the transcriptional regulators, Aire and P-TEFb, to promote elongation of peripheral-tissue antigen transcripts in thymic stromal cells. *Proc. Natl Acad. Sci. USA* **112**, E4448–E4457 (2015).
99. Bansal, K., Yoshida, H., Benoist, C. & Mathis, D. The transcriptional regulator Aire binds to and activates super-enhancers. *Nat. Immunol.* **18**, 263–273 (2017).
100. Noma, K., Allis, C. D. & Grewal, S. I. Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. *Science* **293**, 1150–1155 (2001).
101. Rea, S. et al. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**, 593–599 (2000).
102. Saad, S. et al. DNA binding and mitotic phosphorylation protect polyglutamine proteins from assembly formation. *Cell* **188**, 2974–2991 (2025).
103. Oftedal, B. E. et al. Dominant-negative heterozygous mutations in *AIRE* confer diverse autoimmune phenotypes. *iScience* **26**, 106818 (2023).
104. Koh, A. S., Kingston, R. E., Benoist, C. & Mathis, D. Global relevance of Aire binding to hypomethylated lysine-4 of histone-3. *Proc. Natl Acad. Sci. USA* **107**, 13016–13021 (2010).
105. Fang, Y., Bansal, K., Mostafavi, S., Benoist, C. & Mathis, D. Aire relies on Z-DNA to flag gene targets for thymic T cell tolerization. *Nature* **628**, 400–407 (2024).
106. Du, M. et al. Direct observation of a condensate effect on super-enhancer controlled gene bursting. *Cell* **187**, 2595–2598 (2024).
107. Oven, I. et al. AIRE recruits P-TEFb for transcriptional elongation of target genes in medullary thymic epithelial cells. *Mol. Cell. Biol.* **27**, 8815–8823 (2007).
108. Takahashi, H. et al. Human mediator subunit MED26 functions as a docking site for transcription elongation factors. *Cell* **146**, 92–104 (2011).
109. Giraud, M. et al. An RNAi screen for Aire cofactors reveals a role for Hnrnpl in polymerase release and Aire-activated ectopic transcription. *Proc. Natl Acad. Sci. USA* **111**, 1491–1496 (2014).
110. Keane, P., Ceredig, R. & Seoighe, C. Promiscuous mRNA splicing under the control of AIRE in medullary thymic epithelial cells. *Bioinformatics* **31**, 986–990 (2015).
111. Handel, A. E. et al. Comprehensively profiling the chromatin architecture of tissue restricted antigen expression in thymic epithelial cells over development. *Front. Immunol.* **9**, 2120 (2018).
112. Gamble, N. et al. Thymic epithelial cells amplify epigenetic noise to promote immune tolerance. *Nature* **646**, 724–733 (2025).
113. Michelson, D. A. & Mathis, D. Thymic mimetic cells: ontogeny as immunology. *Annu. Rev. Cell Dev. Biol.* **40**, 283–300 (2024).
114. Gardner, J. M. et al. Deletional tolerance mediated by extrathymic Aire-expressing cells. *Science* **321**, 843–847 (2008).
115. Gardner, J. M. et al. Extrathymic Aire-expressing cells are a distinct bone marrow-derived population that induce functional inactivation of CD4^{+} T cells. *Immunity* **39**, 560–572 (2013).
116. Akagbosu, B. et al. Novel antigen-presenting cell imparts T_{reg}-dependent tolerance to gut microbiota. *Nature* **610**, 752–760 (2022).

117. Wang, J. et al. Single-cell multiomics defines tolerogenic extrathymic Aire-expressing populations with unique homology to thymic epithelium. *Sci. Immunol.* **6**, eabl5053 (2021).
118. Dobes, J. et al. Extrathymic expression of Aire controls the induction of effective T_H17 cell-mediated immune response to *Candida albicans*. *Nat. Immunol.* **23**, 1098–1108 (2022).
119. Pretemer, Y. et al. An iPSC-based in vitro model recapitulates human thymic epithelial development and multi-lineage specification. *Nat. Commun.* **16**, 7680 (2025).
120. Fraschilla, I. & Jeffrey, K. L. The speckled protein (SP) family: immunity's chromatin readers. *Trends Immunol.* **41**, 572–585 (2020).

Acknowledgements

S.H. acknowledges the NIH (R01AI180137) and is a HHMI investigator.

Author contributions

F.L., Y.H. and S.H. wrote the paper and prepared the figures.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Fangwei Leng, Yu-San Huoh or Sun Hur.

Peer review information *Nature Immunology* thanks Ye Zheng, James Gardner and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Laurie A. Dempsey, in collaboration with the rest of the *Nature Immunology* team.