The Autoimmune Regulator Protein Has Transcriptional Transactivating Properties and Interacts with the Common Coactivator CREB-binding Protein

Autoimmune polyendocrinopathy candidiasis ectodermal dys trophy, caused by mutations in the autoimmune regulator (AIRE) gene, is an autosomal recessive autoimmune disease characterized by the breakdown of tolerance to organ-specific antigens. The 545 amino acid protein encoded by AIRE contains several structural motifs suggestive of a transcriptional regulator and bears similarity to cellular proteins involved in transcriptional control. We show here that AIRE fused to a heterologous DNA binding domain activates transcription from a reporter promoter, and the activation seen requires the full-length protein or more than one activation domain. At the structural level AIRE forms homodimers through the NH2-terminal domain, and molecular modeling for this domain suggests a four-helix bundle structure. In agreement, we show that the common transcriptional coactivator CREB-binding protein (CBP) interacts with AIRE in vitro and in yeast nuclei through the CH1 and CH3 conserved domains. We suggest that the transcriptional transactivation properties of AIRE together with its interaction with CBP might be important in its function as disease-causing mutations almost totally abolish the activation effect.

Mutations in the recently cloned autoimmune regulator (AIRE) gene lead to the development of a rare autosomal recessive disease, autoimmune polyendocrinopathy candidiasis ectodermal dys trophy (APECED), also known as autoimmune polyglandular syndrome type 1 (1, 2). APECED is characterized by a breakdown of tolerance to organ-specific antigens leading to different combinations of destructive autoimmune phenomena (3). The clinical features of APECED have been extensively described elsewhere (4–8). In summary, APECED patients are characterized by Addison’s disease and hypoparathyroidism, and they may express defects within and outside the endocrine system, mainly as a result of autoimmunity against organ-specific autoantigens. Although a rare disease, APECED is more common in some genetically isolated populations such as Finns (4), Sardinians (9), and Iranian Jews (10). Finally, due to its monogenic etiology APECED could be considered a unique model in studying organ-specific autoimmune diseases at the biochemical level.

The AIRE gene encodes a predicted 57.5-kDa protein carrying a conserved nuclear localization signal, two PHD-type zinc fingers, four LXXLL motifs or nuclear receptor interaction domains, and the recently described SAND and HSR domains (1, 11–15). The expression of AIRE is limited to two types of cells in the thymus as follows: medullary epithelial cells and cells of monocyte-dendritic cell lineage, both cell types representing a population of antigen-presenting cells (3). Lower expression is seen in the spleen, fetal liver, and lymph nodes. At the subcellular level, AIRE can be found in the cell nucleus in a speckled pattern in domains resembling promyelocytic leukemia nuclear...
bodies, also known as ND10, nuclear dots, or potential oncogenic domains (PODs) (13, 16, 17), associated with the AIRE homologous nuclear proteins Sp100, Sp140, and Lysp100 (3). In addition, two POD-interacting proteins, Sp100 and Sp140, contain important sequence homologies with the NH2-terminal region of AIRE (3).

The CREB-binding protein (CBP) (18) functions as a transcriptional coactivator for a variety of transcription factors including Jun, Fos, nuclear receptors, NFκB, and the STAT proteins (19–23). CBP contains three cysteine-histidine-rich segments, referred to as CH1, CH2, and CH3, involved in specific protein-protein interactions. The CH1 domain binds, among others, to STAT2 (24) and HIF-1a (25) linking this region, minimally, to interferon and hypoxia signal transduction. In addition, CH3 mediates CBP association with adenoviral E1A protein and basal transcription factor TFIIIB (19, 26). CBP contains an intrinsic histone acetyltransferase activity (27, 28) and, furthermore, associates with other coactivators such as P/CAF, SRC-1, TIFII, (SRC-2), and ACTR (SRC-3) (21, 29–31). The plethora of cellular and viral proteins that interact with CBP suggest that it may serve as a transcriptional integrator of multiple signaling pathways involved in different aspects of human physiology and disease. Recently, it has been shown that CBP differentially localizes in POD nuclear domains and associates with POD-interacting promyelocytic leukemia protein (PML) (32) and thus links nuclear compartmentalization to transcriptional regulation and human disease (33).

Although the exact function of the AIRE protein is still unknown, it has been hypothesized, on the basis of its subcellular expression pattern and structural features, that it might be involved in transcriptional regulation and, in consequence, in the negative selection or anergy induction of self-reactive thymocytes (3). To follow that hypothesis, we analyzed AIRE transcriptional properties in yeast and mammalian systems. We show that AIRE has strong transactivating properties when fused to a heterologous DNA binding domain. We further studied the homodimerization properties of AIRE and built a three-dimensional model of the dimerization-mediating HSR domain. The transactivation data suggested to test AIRE as a transcriptional regulator and, in consequence, we identified AIRE as a novel CBP-associated protein. The transactivation activity and its interaction with CBP reported here strongly support the role of AIRE as a transcriptional regulator.

**EXPERIMENTAL PROCEDURES**

**Plasmid Constructs and Mutagenesis**—The full-length AIRE (amino acids 1–545) and the deletion fragments AIRE (86–545, 175–545, and 292–545) were amplified by polymerase chain reaction. The fragments were then cloned into pGEX1ZT (a gift from Dr. Kalle Salakela, University of Tampere) for GST fusion protein expression. GST-AIRE-(1–207) was created by cloning an EcoRI-BamHI digest from pCAIRE (3) to pGEX1ZT. The same fragment was also subcloned into the pEX vector for *in vitro* translation. Several point mutations were engineered into the GST-AIRE construct, all of which create a premature stop codon into the amino acid sequence as follows: GST-AIRE-(1–138), GST-AIRE-(1–256), GST-AIRE-(1–293), and GST-AIRE-(1–348). The R139X and R257X (plasmids GST-AIRE-(1–138) and GST-AIRE-(1–256), respectively) mutations have been found in APECED patients (1, 9). Two patient mutations affecting the NH2 terminus, L29P and K53E (1), were engineered simultaneously to the full-length AIRE cDNA to yield GST-AIRE L29P/K53E. Also, the L28P mutation was engineered alone into the GST-AIRE construct to make pGST-AIRE L28P. In addition, a mutation was designed to disrupt the second PHD zinc finger motif by replacing a cysteine at position 437 with a proline to make GST-AIRE C437P. All mutations were carried out using the GeneEditor kit (Promega) according to the manufacturer’s instructions.

The pCAIRE plasmid for mammalian expression has been described earlier (3). For *in vitro* translation, the deletion constructs pJEX-AIRE-(85–545), pJEX-AIRE-(175–545), and pJEX-AIRE-(292–545) were cloned into pJEX1 as described above. In *in vitro* translation of full-length AIRE was performed from pCAIRE.

For protein-protein interaction studies in yeast, full-length AIRE was truncated, as described above, into pLexA and pJG-45 (CLONTECH) to give pLex-AIRE and pJG-AIRE, respectively. For transactivation reporter assays and proteins-protein interaction studies in mammalian cells, full-length AIRE (pMA-AIRE-(1–545) and pVP16-AIRE-(1–545)) and the deletion constructs pMA-AIRE-(175–545) and pMA-AIRE-(1–216) were cloned as described above. The stop codon-containing cDNAs were cloned into the pMA (CLONTECH) vector by polymerase chain reaction from GST-AIRE-(1–138), GST-AIRE-(1–256), GST-AIRE-(1–293), and GST-AIRE-(1–349), and the mutated cDNAs from GST-AIRE L28P/K53E and L28P were similarly cloned into pMA and pVP16 (CLONTECH). The cDNA from GST-AIRE C437P was cloned into the pMA vector. The pG5-CAT (CLONTECH) reporter plasmid contains three GAL4 response elements upstream of the E1b minimal promoter. Fig. 1 illustrates all the AIRE truncations used in the experiments described.

The yeast expression vector pGAD424-AIRE contains the full-length AIRE cDNA downstream from the GAL4 activation domain. The yeast expression vectors pGBT9CBP and mutants (2N, 3N, 5N, CH1, CBP4CH1, and CH3) contain the CBP amino acid domains indicated in Fig. 6 in frame with the GALADNA binding domain. The pGEXCBP fusion vectors (2N, 3N, CH3, 5N, EK, KE, CH1, E2, and NT) contain the CBP amino acid domains indicated in Fig. 6 in frame with GST protein.

**Cells and Reporter Assays**—HU7-7 and COS-1 cells were maintained as monolayers in Dulbecco’s modified Eagle’s medium supplemented with 100 units/ml penicillin/streptomycin and 10% bovine calf serum (Life Technologies, Inc.). Cultures were maintained at 37 °C and in 7% CO2.

Reporter assays were performed using the Mammalian Matchmaker two-hybrid assay kit (CLONTECH) according to the manufacturer’s instructions with some modifications as described (14). Briefly, 1 × 106 HU7-7 or 3 × 105 COS-1 cells were cotransfected with 0.2 pmol of the PM (DNA binding domain fusion) and pVP16 (activation domain fusion) vectors and 2 μg of the pG5-CAT reporter vector using calcium phosphate precipitation (34). In the transactivation assays only the PM and pG5-CAT vectors were transfected. The total amount of DNA transfected was brought up to 1 μg by a carrier plasmid (pBluescript KS). Analysis of chloramphenicol acetyltransferase (CAT) concentration was performed 48 h post-transfection on the CAT enzyme-linked immunosorbent assay kit (Roche Molecular Biochemicals) according to the manufacturer’s instructions. All transfections were performed in triplicate, and mean CAT concentrations were calculated. The CAT results were normalized in respect to total protein concentration. All interaction experiments were confirmed twice.

**In Vitro Binding Assays**—*In vitro* translation was carried out on the TNT T7-coupled Reticulocyte Lysate System (Promega) according to the manufacturer’s instructions using [35S]-cysteine to produce labeled AIRE-(84–545) or the truncated proteins AIRE-(44–545), AIRE-(175–545), or AIRE-(292–545). For AIRE-(1–207), [35S]methionine was used. The translation products were further purified on Sephadex G-50 columns (Amersham Pharmacia Biotech). Expression and purification of GST fusion proteins was performed according to Frangioni and Neel (35). In all assays, 25–100 μl of glutathione-Sepharose with the relevant GST fusion protein was incubated with 10 μl of *in vitro* translated protein in a 500-μl reaction volume overnight at 4 °C. The complexes were washed several times with STE buffer containing 0.5–1% of detergent (Nonidet P-40). Samples were run on SDS-PAGE, and autoradiography was performed using standard techniques.

For Western blotting was assayed as described (36, 37). In summary, 5 μg of GST-CBP fusion proteins were subjected to 10% SDS-PAGE and transferred to nitrocellulose membrane. The transferred proteins were denatured following by renaturation, and the membrane was incubated with a blocking solution containing 5% dry milk in binding buffer (25 mM Tris-Cl (pH 8.0), 100 mM NaCl, 10 mM MgCl2, 2 mM EDTA, 5 mM NaF, 1 mM Na2VO4, 5% glycerol, 0.1% Triton X, 1 mM diethiothreitol, 1 mM mg/ml aprotinin, 1 mg/ml leupeptin, 1 mg/ml papain) A) for 8 h at room temperature. The membrane was incubated with AIRE probe labeled in the TNT coupled reticulocyte lysate system (Promega) with [35S]methionine in binding buffer overnight at 4 °C. After washing twice with binding buffer, the hybridizations were analyzed by image analyzer (LAS2000, Fujix).

**Yeast Two-hybrid Analyses**—The AIRE-AIRE interaction experiments were performed on the Matchmaker LexA two-hybrid system (CLONTECH) according to the manufacturer’s instructions. The pJG-
AIRE and pLex-AIRE plasmids were cotransfected into the CG-1945 yeast strain. Leucine, histidine, and uracil nutritional selection markers were used. For the analysis of β-galactosidase expression, the colony lift filter assay was used. All experiments were performed three times.

For the AIRE-CBP interaction yeast strain Y190 was transformed with a series of pGBT9-CBP wild type and deletion mutants and pACT-empty or pACT-AIRE expression vectors according to the lithium acetate method (38). Yeast transformants were selected on agar plates lacking tryptophan and leucine for 3 days. The selected colonies were inoculated into 1 ml of liquid culture, and liquid β-galactosidase assays were performed as described (38).

**Molecular Modeling of the HSR Domain—** The NH₂ terminus of Aire was modeled based on the structure of the four-helix bundle of proto-oncogene CBP (Protein Data Bank entry 2CBL) (39). Sequences were aligned based on the feature of amino acids in the four-helix bundle structure. Alignment was performed using the program package GCG (40). The model was built using the programs InsightII and Discover (Molecular Simulations, Inc., San Diego, CA). Insertions and deletions were modeled by searching the loops from the selected data base of the Protein Data Bank. The model was refined by energy minimization with Discover in a stepwise manner by using Amber force field. First, all heavy atoms were fixed and then the side chains of built loops were freed followed by the backbone of loops. Finally, only the C-α atoms of conserved regions were constrained. The model was evaluated with the PROCHECK version 3.5 (41) program.

**RESULTS**

**AIRE Has Transcriptional Transactivation Activity—** To determine the ability of AIRE to regulate transcription, we performed a series of GAL4 system reporter assays in HUH-7 and COS-1 cells. The full-length AIRE cDNA (pM-AIRE-(1–545)) and the fragments from pM-AIRE (residues 1–138, 1–216, 1–256, 1–293, 1–348, and 175–545), covering the full-length sequence (Fig. 1), were expressed as fusion proteins with the GAL4 DNA binding domain, and the transcriptional activation was measured as CAT expression from the pG5-CAT reporter. Fold activity is compared with base-line activity as determined by the amount of CAT expressed in transfected with the empty pM vector. Mock, non-transfected COS-1 cells. Wild type AIRE-(1–545) (wt) and mutants AIRE-(1–216) and AIRE-(175–545) are fused to GAL4 DBD (pM vector). Expression vectors are used as indicated.

**FIG. 1. A schematic representation of the AIRE protein.** The important conserved motifs and the AIRE truncation mutants used in the experiments are indicated. PHD, PHD-type zinc finger domain; PRR, proline-rich region.

**FIG. 2. Wild type AIRE has transcriptional transactivation properties.** Transient transfection was conducted in COS-1 cells using pG5-CAT as reporter. Fold activity is compared with base-line activation (base) as determined by the amount of CAT expressed in transfected with the empty pM vector. Mock, non-transfected COS-1 cells. Wild type AIRE-(1–545) (wt) and mutants AIRE-(1–216) and AIRE-(175–545) are fused to GAL4 DBD (pM vector). Expression vectors are used as indicated.

AIRE and pLex-AIRE plasmids were cotransfected into the CG-1945 yeast strain. Leucine, histidine, and uracil nutritional selection markers were used. For the analysis of β-galactosidase expression, the colony lift filter assay was used. All experiments were performed three times.

For the AIRE-CBP interaction yeast strain Y190 was transformed with a series of pGBT9-CBP wild type and deletion mutants and pACT-empty or pACT-AIRE expression vectors according to the lithium acetate method (38). Yeast transformants were selected on agar plates lacking tryptophan and leucine for 3 days. The selected colonies were inoculated into 1 ml of liquid culture, and liquid β-galactosidase assays were performed as described (38).

Molecular Modeling of the HSR Domain—The NH₂ terminus of Aire was modeled based on the structure of the four-helix bundle of proto-oncogene CBP (Protein Data Bank entry 2CBL) (39). Sequences were aligned based on the feature of amino acids in the four-helix bundle structure. Alignment was performed using the program package GCG (40). The model was built using the programs InsightII and Discover (Molecular Simulations, Inc., San Diego, CA). Insertions and deletions were modeled by searching the loops from the selected data base of the Protein Data Bank. The model was refined by energy minimization with Discover in a stepwise manner by using Amber force field. First, all heavy atoms were fixed and then the side chains of built loops were freed followed by the backbone of loops. Finally, only the C-α atoms of conserved regions were constrained. The model was evaluated with the PROCHECK version 3.5 (41) program.

**RESULTS**

AIRE Has Transcriptional Transactivation Activity—To determine the ability of AIRE to regulate transcription, we performed a series of GAL4 system reporter assays in HUH-7 and COS-1 cells. The full-length AIRE cDNA (pM-AIRE-(1–545)) and the fragments from pM-AIRE (residues 1–138, 1–216, 1–256, 1–293, 1–348, and 175–545), covering the full-length sequence (Fig. 1), were expressed as fusion proteins with the GAL4 DNA binding domain, and the transcriptional activation was measured as CAT expression from the pG5-CAT reporter. Fold activity is compared with base-line activity as determined by the amount of CAT expressed in transfected with the empty pM vector. Mock, non-transfected COS-1 cells. Wild type AIRE-(1–545) (wt) and mutants AIRE-(1–216) and AIRE-(175–545) are fused to GAL4 DBD (pM vector). Expression vectors are used as indicated.

AIRE and pLex-AIRE plasmids were cotransfected into the CG-1945 yeast strain. Leucine, histidine, and uracil nutritional selection markers were used. For the analysis of β-galactosidase expression, the colony lift filter assay was used. All experiments were performed three times.

For the AIRE-CBP interaction yeast strain Y190 was transformed with a series of pGBT9-CBP wild type and deletion mutants and pACT-empty or pACT-AIRE expression vectors according to the lithium acetate method (38). Yeast transformants were selected on agar plates lacking tryptophan and leucine for 3 days. The selected colonies were inoculated into 1 ml of liquid culture, and liquid β-galactosidase assays were performed as described (38).

**Molecular Modeling of the HSR Domain—** The NH₂ terminus of Aire was modeled based on the structure of the four-helix bundle of proto-oncogene CBP (Protein Data Bank entry 2CBL) (39). Sequences were aligned based on the feature of amino acids in the four-helix bundle structure. Alignment was performed using the program package GCG (40). The model was built using the programs InsightII and Discover (Molecular Simulations, Inc., San Diego, CA). Insertions and deletions were modeled by searching the loops from the selected data base of the Protein Data Bank. The model was refined by energy minimization with Discover in a stepwise manner by using Amber force field. First, all heavy atoms were fixed and then the side chains of built loops were freed followed by the backbone of loops. Finally, only the C-α atoms of conserved regions were constrained. The model was evaluated with the PROCHECK version 3.5 (41) program.

**RESULTS**

AIRE Has Transcriptional Transactivation Activity—To determine the ability of AIRE to regulate transcription, we performed a series of GAL4 system reporter assays in HUH-7 and COS-1 cells. The full-length AIRE cDNA (pM-AIRE-(1–545)) and the fragments from pM-AIRE (residues 1–138, 1–216, 1–256, 1–293, 1–348, and 175–545), covering the full-length sequence (Fig. 1), were expressed as fusion proteins with the GAL4 DNA binding domain, and the transcriptional activation was measured as CAT expression from the pG5-CAT reporter. Fold activity is compared with base-line activity as determined by the amount of CAT expressed in transfected with the empty pM vector. Mock, non-transfected COS-1 cells. Wild type AIRE-(1–545) (wt) and mutants AIRE-(1–216) and AIRE-(175–545) are fused to GAL4 DBD (pM vector). Expression vectors are used as indicated.
modimerization, we performed GST pull-down experiments using various AIRE deletion constructs. A series of truncation mutants were expressed as GST fusions (AIRE-(1–138, 1–256, 1–293, and 1–348)) and binding with in vitro translated, labeled wild type AIRE was tested. We found that of the truncated proteins AIRE-(1–256, 1–293, and 1–348) from the NH₂-terminal region clearly bound labeled full-length AIRE (Fig. 4A). Moreover, AIRE-(1–138) bound to some extent but gave a less prominent band. We confirmed NH₂-terminal homodimerization by testing the binding of the full-length protein and AIRE-(1–207) (as GST fusions) with in vitro translated, labeled AIRE-(1–207). We found that labeled AIRE-(1–207) bound GST-AIRE-(1–545 and 1–207) but not GST alone (Fig. 4A). To study the possibility of AIRE-AIRE homodimerization within other regions, we tested the binding of expressed GST-AIRE-(1–545) fusion with in vitro translated, labeled fragments AIRE-(175–545, 84–545, and 292–545). However, we did not see any interactions with these deletion constructs (data not shown). Taken together, these results show that the NH₂-terminal domain (aa 1–207) of AIRE, and probably the minimal 1–100 amino acid domain, mediates AIRE homodimerization.

**Patient Mutations Are Deficient for Transcriptional Activation and Homodimerization**—To assess the role of disease-causing NH₂-terminal missense mutations found in APECED patients, we tested two AIRE proteins carrying a single and a double point mutation, AIRE L28P and AIRE L28P/K83E, respectively, for transcriptional activation. We found that both mutants were very weak in activating transcription from the reporter compared with wild type AIRE GAL4 DNA binding domain fusion protein (Fig. 4B). Indeed, pM-AIRE L28P and pM-AIRE L28P/K83E caused an activation approximately 10% and 3%, respectively, of that caused by the wild type pM-AIRE-(1–545) construct (Fig. 4B). To assess the role of the PHD zinc finger domains in the activating function, we tested the mutant pM-AIRE C437P in which the second cysteine of the most COOH-terminal zinc finger (Fig. 1) is replaced by a proline. We found that it was deficient in activating transcription yielding an activation approximately 30% of the activation presented by wild type AIRE (Fig. 4C). We then tested these mutants for dimerization with wild type AIRE and also for self-interaction, L28P-L28P and L28P/K83E-L28P/K83E, respectively, using the appropriate DBD and activation domain fusion constructs. We found that neither AIRE L28P nor AIRE L28P/K83E were able to homodimerize with themselves (Fig. 4B, right panel), but both could to some extent interact with the wild type protein (Fig. 4B, left panel). We conclude that AIRE mutant proteins, present in patients carrying missense point mutations in the conserved HSR domain (1–100 aa), express severely reduced transactivation and homodimerization properties compared with wild type AIRE.

The NH₂ Terminus of AIRE Has a Predicted α-Helical Four-helix Bundle Structure—In the light of our findings in the in vitro and in vivo interaction assays, which implicated the NH₂-terminal portion of AIRE as mediator of homodimerization, we have performed a structural prediction and built a three-dimensional model of that region. The only intact protein motif within the first 207 amino acids is the HSR domain of approximately 100 NH₂-terminal amino acids. Thus, we chose the first 95 amino acids for the modeling. Fig. 5 shows the predicted three-dimensional structure of the HSR domain. Sequence analysis of the NH₂ terminus of AIRE against the Protein Data Bank did not indicate any significant similarities. Nevertheless, results from PSI-BLAST (42) at NCBI contained mainly α-helical proteins, including four-helix bundle-containing proteins. Secondary structure prediction with the Protein Data Bank did not indicate any significant similarities. Nevertheless, results from PSI-BLAST (42) at NCBI contained mainly α-helical proteins, including four-helix bundle-containing proteins. Secondary structure prediction with the Protein Data Bank (43) suggested strongly that the first 95 amino acids of AIRE were α-helical. The amino-terminal region of proto-oncogene CBL contains a four-helix bundle with similar lengths of α-helical regions as AIRE (39). HelicalWheel (40) prediction indicated amphipathic regions in the α-helices in AIRE. The hydrophobic side chains point to the interior of the structure to form a four-helix bundle. This information was used to refine the alignment produced with the GCG program package (40). Despite the low overall sequence similarity, the alignment has good agreement of the bundle-forming residues. All the loops between the α-helices are shorter in AIRE than in CBL. The structure was validated and found to have good stereochemistry.

**AIRE Interacts with CBP**—Given that AIRE expresses strong transcriptional activities in mammalian cells, we investigated whether AIRE could directly interact with known co-factors involved in transcriptional activation. We initially carried out a series of in vitro experiments using the most common co-activator, the CREB-binding protein (CBP), as the first potential target of AIRE protein. A series of CBP deletion mutants covering almost the full-length CBP (Fig. 6A, I) were produced as GST fusion proteins. Equal amounts of GST alone and CBP (2N, 3N, CH3, and 5N) mutant fusion proteins were analyzed in a 10% SDS-PAGE, and samples were tested for interaction with in vitro translated wild type AIRE protein in a far Western experiment (Fig. 6B, panel I). As shown, AIRE interacts specifically with 2N (117–737 aa) and CH3 (1680–1890 aa) CBP fragments and not with GST alone or the 3N (738–1625 aa) and 5N (2389–2260 aa) domains. The result above suggested that AIRE interacts minimally with two independent domains of CBP, one of which was definitely the CH3 region. To delineate further the second interaction domain of AIRE-CBP located in the 2N region, we tested additional GST-CBP fusion proteins covering the CBP region 1–737 aa (Fig. 6A, panel II). Indeed, we found that AIRE binds specifically to the minimal CH1 domain (344–451 aa) and not to the EK-(1–270), E2-(452–721), or NT-(451–521) (Fig. 6B, II). We conclude that AIRE interacts in vitro with CBP through the CH1 and CH3 conserved domains.

---

**Fig. 3. AIRE forms homodimers.** A, AIRE homodimerizes in vitro. In vitro labeled AIRE-(1–545) binds to GST-AIRE-(1–545) and not to GST alone. Increasing amounts of the GST-AIRE-(1–545) fusion protein (25, 50, and 100 μl) and 50 μl of GST alone were incubated with the same amount of in vitro labeled AIRE protein. Input, 33% of the labeled protein used in the binding experiments. All samples were analyzed in 10% SDS-PAGE. B, AIRE homo-interaction in the mammalian system. Experiment is conducted as in Fig. 2. The presence of pVPI6-AIRE-(1–545), 5-fold from pM-AIRE-(1–138), and 30–60-fold from the AIRE mutants pM-AIRE-(1–256, 1–293, and 1–348). None of the pM-(1–138, 1–256, 1–293, or 1–348) mutants alone stimulate reporter gene expression.
Given the results of the in vitro binding studies, it was important to determine whether AIRE and CBP interact also in the yeast nuclei. A series of Y190 transformations were carried out with wild type CBP and mutants (2N, CH1, D CH1, 3N, CH3, and 5N) fused to the GAL4 DNA binding domain and wild type AIRE fused to the GAL4 activation domain or empty GAL4 activation domain vectors (Fig. 6C). Indeed, as shown in the in vitro studies, AIRE interacts specifically also in yeast nuclei with CBP through two independent sites (CH1 and D CH1). We suggest that the transcriptional activities of AIRE might be mediated through its physical interaction with the common coactivator CBP.

**DISCUSSION**

**AIRE Activates Transcription**—The data presented above provide evidence that AIRE can remarkably activate transcription from a reporter gene when fused to a heterologous DNA binding domain. Mutagenesis analysis could not accurately pinpoint any activation domain; none of the COOH- or NH2-terminal truncation mutants tested had the transcriptional transactivating activity. The mutants pM-AIRE L28P and pM-AIRE L28P/K83E almost totally lack the activity of the wild type pM-AIRE(1–545) amounting to approximately 10 and 3% of wild type (wt), respectively. The mutants can dimerize with wild type pVP16-AIRE(1–545) but slightly less effectively than wild type pM-AIRE(1–545) with wild type pVP16-AIRE, even when the basal activation caused by pM-AIRE(1–545) is taken into account (left panel). The mutants are, however, unable to form either AIRE L28P-AIRE L28P or AIRE L28P/K83E-AIRE L28P/K83E dimers (right panel). C, the AIRE C437P mutant is deficient in activating transcription from the promoter causing an activation of 30% compared with wild type AIRE, suggesting that the second PHD zinc finger is important for the activation function.

**AIRE Homodimerizes through the NH2 Terminus**—The GST pull-down and two-hybrid experiments in yeast and mammalian systems indicated that amino acids 1–207 of AIRE mediate the homodimer formation. This is in agreement with earlier reports as this NH2-terminal region contains the HSR domain (1–100 aa), shared by AIRE and Sp100 proteins, which has been shown to mediate Sp100 homodimerization and subnuclear targeting (14).

**Missense Mutations Affect Transcriptional Activation and Homodimerization**—We found that AIRE L28P fused with the GAL4 DBD retained only 10% of the transcriptional transactivating activity of the wild type protein. The double mutant (AIRE L28P/K83E) behaved essentially in a similar manner (Fig. 4B). Neither of the mutants was able to form homodimers but both did, to some extent, dimerize with wild type AIRE (Fig. 4B). It is possible that the conformational change caused by these mutations is not enough to abolish completely dimerization with the wild type protein having the intact HSR domain, but homodimerization between mutant forms is effectively destroyed. The C437P mutant was also deficient for transcriptional activation resulting in 30% activation of wild type AIRE that supports the involvement of the second PHD zinc finger in transcriptional activation. We conclude, based on these findings, that the HSR domain of AIRE is needed, but is not sufficient, for transcriptional transactivation activity. It
appears likely that the dimerization through the HSR region is necessary for the transcriptional transactivation function. Furthermore, at least the second of the PHD zinc fingers is also needed for the activation function.

The AIRE HSR domain consists of approximately 100 NH$_2$-terminal amino acids and has previously been predicted to have an α-helical structure (11). We performed a structural prediction of the HSR domain and built a three-dimensional molecular model of this region. The sequence analysis of AIRE suggested the 95 NH$_2$-terminal residues to have a four-helix bundle conformation. Several bundle structures with tightly packed α-helices have been determined (51) and have been implicated in dimer formation (52, 53). The helices in the compact domain are linked by loops of variable length. The helices are generally amphipathic in nature, and this property was used to align the sequences. The inner core of the protein is formed by hydrophobic residues including several leucines.

The three-dimensional structure of the four-helix bundle was used to interpret the binding and transactivation data as well as information from APECED-causing mutations in this domain. The Leu-28 residue locates in the α-helix 2 and substitution by proline would cause a bend to the α-helix and thereby affect the loop between the first two helices. The local structural change of helix 2 and its connection to the first helix will strongly weaken the packing of the whole domain. Interestingly enough, all APECED-causing missense mutations described so far are within the HSR domain. The mutations in other regions of AIRE are mostly nonsense mutations and frameshift-causing changes, indicating that the HSR region is the most vulnerable domain for conformational changes.

**AIRE Interacts with CBP**—The data presented above provide evidence that the autoimmune regulator protein physically interacts with common coactivator CBP and support a model in which this interaction may contribute to AIRE transcriptional properties. The direct interaction of AIRE and CBP, both in vitro and in yeast nuclei, suggests that AIRE may function as a novel CBP cofactor. Indeed, AIRE is a strong transcriptional activator when fused to a heterologous DNA binding domain.

![Molecular model of the HSR domain showing a four-helix bundle](image-url)
These findings are supported further by the CBP mutagenesis that depicted two independent CBP-AIRE interaction domains. The first is the minimal 344–451 aa, CH1 domain, that mediates CBP interaction with the STAT2 or interferon signaling pathway and HIF-1α or hypoxia-inducible factor (24, 25). The second is the minimal 1680–1890 aa, CH3 domain, that mediates CBP interaction with adenoviral E1A and basal transcription factor TFIIIB (19, 26). This finding is of great importance as AIRE could be used as a bridging molecule between the basal transcriptional machinery and signal-activated trans-regulators. Indeed, AIRE can be found in the cell nucleus in a speckled pattern in domains resembling promyelocytic leukemia nuclear bodies (3) which have been directly implicated in transcription regulation of AP-1 and nuclear receptor signaling pathways (32, 33, 54). Taken together, the data presented above may lead to a new consideration as to the potential contribution of AIRE protein and of AIRE-associated nuclear structures in transcription control.

Even though further analysis will be necessary to elucidate the minimal domain of AIRE involved in CBP interaction and its role in transcription control, our data suggest that the transcriptional properties of AIRE and its interaction with CBP might be important in its function. We have shown earlier that AIRE is expressed in antigen-presenting epithelial cells that are responsible for the negative selection of the thymus medulla (3). The process of negative selection by which the immune system deletes autoreactive T-cells occurs through the apoptosis of CD4+8 immature thymocytes, and it is induced by medullary epithelial cells (55, 56). The induction of the apoptotic signal is elicited by interaction of the T-cell receptor and major histocompatibility complex-self-peptide complex but also by potential costimulatory interactions such as CD40-CD40L, CD30-CD30L and Fas-FasL. Given the restricted expression of AIRE in cells mediating the negative selection and the autoimmune phenotype of APECED patients, AIRE might well be involved in regulation of the apoptotic mechanisms in thymus medulla. The described role of AIRE as an activator protein and its interaction with common coactivator CBP thus gives a strong basis to study the transcriptional regulation of genes involved in thymic negative selection and could provide a model for studying organ-specific autoimmune diseases at the biochemical level.

Acknowledgments—We are grateful to Ulla Kiiskinen for excellent technical assistance and to Y. Dunant, P. Suter, and D. Trono for their contributions.

REFERENCES
1. Nagamine, K., Peterson, P., Scott, H. S., Kudoh, J., Minoshima, S., Heino, M., Krohn, K. J., Laloti, M. D., Mullis, P. E., Antonarakis, S. E., Kawasaki, K., Asakawa, S., Ito, F., and Shimizu, M. (1997) Nat. Genet. 17, 393–398
2. The Finnish-German APECED Consortium (1997) Nat. Genet. 17, 399–403
3. Heino, M., Peterson, P., Kudoh, J., Nagamine, K., Lagerstedt, A., Ovod, V., Ranki, A., Rantala, I., Nieminen, M., Tuukkanen, J., Scott, H. S., Antonarakis, S. E., Shimizu, N., and Krohn, K. J. (1999) J. Med. Genet. 36, 281–285
4. Ahonen, P. (1985) Clin. Genet. 27, 535–542
5. Ahonen, P., Myllärniemi, S., Sipila, I., and Perheentupa, J. (1996) N. Engl. J. Med. 322, 1829–1836
6. Perheentupa, J. (1996) Horm. Metab. Res. 28, 353–356
7. Obermayer-Straub, P., and Manns, M. P. (1998) Bailliere’s Clin. Gastroenterol. 12, 293–315
8. Aaltonen, J., and Björres, P. (1999) Ann. Med. 31, 111–116
9. Rosatelli, M. C., Meloni, A., Meloni, A., Devoto, M., Cao, A., Scott, H. S., Peterson, P., Heino, M., Krohn, K. J., Nagamine, K., Kudoh, J., Shimizu, N., and Antonarakis, S. E. (1998) Hum. Genet. 103, 429–434
10. Zlotogora, J., and Shapiro, M. S. (1992) J. Med. Genet. 29, 824–826
11. Mittaz, L., Rossier, C., Heino, M., Peterson, P., Krohn, K. J., Gos, A., Morris, M. A., Kudoh, J., Shimizu, N., Antonarakis, S. E., and Scott, H. S. (1999) 6, 299–306.
