The Autoimmune Regulator (AIRE) Is a DNA-binding Protein*

Received for publication, May 29, 2001, and in revised form, August 30, 2001
Published, JBC Papers in Press, August 31, 2001, DOI 10.1074/jbc.M104898200

Pradeep G. Kumar, Malini Laloraya, Cong-Yi Wang, Quin-Guo Ruan, Abdoreza Davoodi-Semiromi, Kuo-Jang Kao, and Jing-Xiong She‡

From the Department of Pathology, Immunology, and Laboratory Medicine, Center for Mammalian Genetics and Diabetes Center of Excellence, College of Medicine, University of Florida, Gainesville, Florida 32610

Autoimmune polyendocrinopathy candidiasis ectodermal dystrophy (APECED),† also known as autoimmune polyglandular syndrome type 1 (APS1), is a rare autosomal recessive disorder common in isolated populations such as Finns, Sardinians, and Iranian Jews (1). This syndrome is characterized by destructive autoimmune diseases of the endocrine organs, chronic candidiasis of mucous membranes, and ectodermal disorders. APECED is caused by mutations in the autoimmune regulator (AIRE) gene on chromosome 21q22.3 (2–4). The AIRE gene has recently been cloned by two independent groups of investigators (5, 6). The AIRE gene consists of 14 exons coding for a 2445-base pair mRNA transcript, and the translated product is expected to have 545 amino acids with a predicted molecular mass of 57.5 kDa. The predicted AIRE protein has several domains indicative of a transcriptional regulator protein (6). AIRE harbors two zinc fingers of plant homeodomain (PHD) type. A putative DNA binding domain named SAND as well as four nuclear receptor binding LXXLL motifs, an inverted LXXLL domain, and a variant of the latter (FXLXXL) hint that this protein functions as a transcription coactivator (5–7). Furthermore, a highly conserved N-terminal 100-amino acid domain in AIRE has a significant homology to the homogenously staining (HSR) domain of Sp100 and Sp140 proteins (7). This domain has been shown to function as a dimerization domain in several Sp-100 related proteins (8). 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, associated with the AIRE homologous nuclear proteins Sp100, Sp140, and Lysp100 (9).

Interestingly, it has recently been shown that AIRE can activate transcription from a reporter gene when fused to a heterologous DNA binding domain. This activation required the full-length protein or the presence of more than one activation domain. A glutathione S-transferase pull-down assay showed that AIRE formed homodimers in vitro, probably through the N-terminal domain (amino acids 1–207) or through the minimal 1–100 amino acid domain resembling Sp100. It has also been shown that AIRE interacts in vitro with CREB-binding protein (CBP) through the CH1 and CH3 conserved domains, which has led the investigators to suggest that the transcriptional activities of AIRE might be mediated through its physical interaction with the common coactivator CBP (10). CBP is a key coactivator that modulates the transcriptional regulation dependent on adenylate cyclase-signaling pathway in eukaryotes. The signal processing is mediated by a family of cyclic AMP-responsive nuclear factors, including CREB, cAMP response element modulator (CREM), and activating transcription factor 1 (ATF-1). These factors contain the basic domain/leucine zipper motifs and bind as dimers to cAMP-responsive elements (CREs). The activation function of CRE-binding proteins in turn is modulated by several kinase-dependent phosphorylations and is mediated by coactivators such as CBP and p300 (11).

Although it is important to note that the interaction of AIRE with CBP might influence signal transduction mediated by cAMP-responsive nuclear factors, some of the structural properties of AIRE are suggestive of its function beyond the suspected transactivation. The presence of two PHD zinc fingers and a leucine zipper was the focus of our attention. The objectives of the present study were 1) to verify whether recombinant human AIRE oligomerizes in vitro and 2) to examine the possibility of AIRE possessing DNA binding properties. In this paper, we report that AIRE forms homodimers and homotetramers in vitro and that both the dimer and tetramer forms of AIRE possess DNA binding activity. We further demonstrate that AIRE exists in oligomerized forms in vivo and that phosphorylation of AIRE by cAMP-dependent PKA and/or PKC could trigger its dimerization.

* This work was supported by National Institutes of Health Grants 1R24DK58778 (NIDDK), 1RO1HD37800 (NICHD), and 1P01AI-42288 (NIAID) and by a grant from the Juvenile Diabetes Research Foundation.
† To whom correspondence should be addressed: Dept. of Pathology, Immunology, and Laboratory Medicine, Box 100275, University of Florida, Gainesville, FL 32610. Tel.: 352-392-0667; Fax: 352-392-3053; E-mail: she@ufl.edu.
‡ The abbreviations used are: APECED, autoimmune polyendocrinopathy candidiasis ectodermal dystrophy; PHD, plant homeodomain; HSR, homogeneously staining region; CREB, cAMP-response element-binding protein; CBP, CREB-binding protein; PAGE, polyacrylamide gel electrophoresis; rhAIRE, recombinant human AIRE; PFD, poly(ADP-ribose) polymerase chain reaction; PKA, protein kinase A; PKC, protein kinase C; EGR, early growth response factor.
Dna Binding Activity of Human Aire Protein

Materials and Methods

Plasmid Constructs Protein Expression—The full-length Aire (amino acids 1–545) was amplified by polymerase chain reaction. The fragment was cloned into pET32 containing a thioredoxin tag (Novagen) for 6×His fusion protein expression. Escherichia coli BL21 (DE3) were transformed and selected on ampicillin. Starting with a single colony, 500 ml of an overnight culture of a positive clone was set up. The expression of the AIRE gene was induced with 0.1% isopropyl-β-D-galactopyranoside, and the culture was continued for 3 h. The cells were harvested, and the recombinant protein was purified using the B-PER 6× histidine fusion protein purification system (Pierce) following the manufacturer’s instructions with slight modifications. Briefly, the cells harvested from fresh induction were resuspended in 20 ml of B-PER reagent by pipetting up and down until the cell suspension was homogenous. The suspension was shaken gently on a rocker. The supernatant was separated by centrifugation at 27,000 × g for 15 min at 4 °C. The supernatant was separated into two aliquots of 10 ml each and loaded onto two nickel-chelated columns equilibrated with B-PER reagent. The columns were washed three times with the wash buffers following the instructions of the manufacturer. The elution was performed on an imidazole gradient (0.1–100 mM imidazole). The columns were washed three times with the wash buffers.

Expression and Purification of Aire Protein—The expression of the AIRE gene was induced with 0.1% isopropyl-β-D-galactopyranoside, and the culture was continued for 3 h. The soluble proteins were separated from the insoluble fraction by centrifugation at 27,000 × g for 15 min at 4 °C. The supernatant was separated into two aliquots of 10 ml each and loaded onto two nickel-chelated columns equilibrated with B-PER reagent. The columns were washed three times with the wash buffers following the instructions of the manufacturer. The elution was performed on an imidazole gradient (0.1–100 mM imidazole). The columns were washed three times with the wash buffers.

Immunoprecipitation of Aire from Mouse Thymus—Immunoprecipitation of Aire from Mouse Thymus

DNA Binding Activity of Human Aire Protein

Labeling of Oligonucleotides—The oligonucleotides used in this study are listed in Table I. The oligonucleotides were end-labeled with [γ-32P]ATP after a standard T4 polynucleotide kinase reaction. The phosphorylation reaction was continued for 30 min at 37 °C, after which the reaction was stopped by the addition of 1 µl of 0.5 mM EDTA/10 µl of total reaction volume. The unincorporated label was removed using QIAQUICK nucleotide removal kit (Qiagen) following the instructions of the manufacturer. The labeled oligonucleotides were eluted in 100 µl of TE buffer. The specific activities of the labeled oligonucleotides were measured and were confirmed to be close to 20,000 cpm/100 fmol before they were used in gel-shift assays.

Gel Shift Assays—Gel shift assays were performed with purified Aire (100 µg/reaction), and radiolabeled double-stranded oligonucleotides were prepared as stated above. The oligonucleotides used in this study are listed in Table I. The protein preparations were preincubated in binding buffer (10 µM Tris-HCl (pH 7.5), 50 mM NaCl, 0.5 mM dithiothreitol, 0.5 mM EDTA, 1 mM MgCl2, 4% glycerol) in the presence of 0.05 µg/ml poly(dI-dC)-poly(dI-dC). Competition studies were performed by the addition of 1.75 pmol of unlabeled double-stranded oligonucleotides into the preincubation step. DNA-protein complexes were formed by the addition of 0.175 pmol of probe and resolved by separation from unbound radiolabeled oligonucleotides through 6% non-denaturing polyacrylamide gels (acrylamide:bisacrylamide ratio, 80:1) containing 2.5% glycerol in 50 mM Tris (pH 9.5) and 4.5 mM boric acid, 1 mM EDTA) using a Protean II electrophoresis system (Bio-Rad). The gels were pre-run at 100 V for 30 min before electrophoresis, the gels were recovered and introduced into the transfer module. The proteins were transferred onto a PVDF membrane pre-wet in methanol and equilibrated in the transfer buffer using a Mini Transblot cell (Bio-Rad). The transfer buffer had the following composition: 40% methanol (v/v), 25 mM Tris, and 190 mM glycine (pH 8.2). The transfer was performed at 30 mA of constant current for 12–16 h, after which the PVDF membrane was handled and air-dried.

For the development of blots, the membranes were pre-wet in methanol and then incubated in the Opti–4CN blocking reagent (Bio-Rad) for 1 h at 37 °C followed by extensive washing in phosphate-buffered saline with 0.1% Tween 20 at room temperature. The membranes were incubated with appropriate primary antibodies for 1 h at room temperature. After incubation, the membranes were washed with three changes of phosphate-buffered saline, 0.1% Tween 20 and then incubated with goat anti-rabbit IgG-horseradish peroxidase at a dilution of 1:5000. The membranes were washed extensively after incubation and then developed with amplified Opti–4CN reagents (Bio-Rad) following the recommendations of the manufacturer.

N-Terminal Microsequencing—N-terminal microsequencing of the 54-, 110-, and 220-kDa bands affinity-purified from mouse thymus on an immobilized anti-AIRE antibody were sequenced. They were used in gel-shift assays.
loading of the binding reaction. After loading the samples, the DNA-protein complexes were separated by electrophoresis on 150 V for 3.5 h. All procedures were performed at room temperature. After electrophoresis, the gels were exposed to x-ray film (Fuji Film).

Gel shift competition assays were performed by the introduction of a synthetic zinc finger peptide (ZIF 285) in the reaction. The basic reaction was set up as stated above, with a modification involving the addition of ZIF 285 at a concentration of 100 pM. This reaction was allowed to incubate for 15 min at room temperature, after which AIRE was introduced at concentrations ranging from 100 pM through 800 pM. Gel shift assays of these reactions were performed in the same fashion as discussed above.

**Oligo Library Screening**—We determined the ability of purified recombinant AIRE to bind DNA employing two oligo libraries. The first library (Nhe oligo library) had randomized a 26-mer stretch flanked by arms for PCR amplification. The second library (G,N,G,G library) had NheGGN,GGN, sequence flanked by adapter sequence for PCR amplification. The libraries were PCR-amplified under standard conditions (melting at 94 °C for 2 min followed by cycling at 94 °C for 30 s, annealing at 63 °C for 30 s, and extension at 72 °C for 1 min; a final hold at 72 °C for 10 min was given for completion of chain elongation). After amplification, the PCR product was suspended in 1× binding buffer and incubated with purified and refolded rhAIRE for 1 h at room temperature with constant shaking. After this, anti-AIRE antibody (1:100 dilution) was added to the mixture, and the incubation was continued for an additional 1 h. Protein A-agarose with preblocked nonspecific binding sites (1:10 dilution) was added to the mixture, and the incubation was allowed to continue for 1 h. At the end of this final incubation, the suspension was centrifuged at 16,000 rpm on an Eppendorf centrifuge.

The supernatant was carefully removed, and the pellet was washed 5 times with 1.5 ml of 1× binding buffer each time. The final pellet was resuspended in 1× TE. A 2-µl aliquot of the resuspended pellet was PCR-amplified under the settings mentioned above. The whole procedure was repeated 5 times as discussed earlier. After the last amplification, the PCR product was separated on a 3% agarose gel, and the band was eluted and cloned using pCR4-TOPO cloning kit for sequencing (Invitrogen). TOP10 cells were transformed adopting a one-shot chemical transformation protocol.

Thirty-two clones each were selected and sequenced from the AIRE-selected Nhe library and the G,N,G,G library. Plasmid minipreps were made from the selected colonies and sequenced using M13 forward primer. The insert sequences were aligned. Two different controls were included in this experimentation. The first control involved the PCR amplification of G,N,G,G library (before selection by AIRE) and its cloning into pCR4-TOPO vector. Forty-four clones were sequenced and the sequences were analyzed to evaluate the bias in the random areas of the oligonucleotides in this library (control 1). The second control involved the cloning of oligonucleotides selected by reagents excluding AIRE from the experimental reaction set-up (control 2).

**Relative Binding Affinity Assays**—Relative binding affinity assays were performed in a manner similar to that outlined for gel-shift assays. In this assay, we constructed four independent sets of four reactions, each using hot oligonucleotides oligo-TGG, oligo-TG, oligo-GG, and oligo-G, respectively. Each set had a gel shift with the given oligonucleotide and competitors with the other three oligonucleotides. The concentrations of the protein, hot oligonucleotide, and competitors were the same as those in the gel-shift assays detailed earlier in this section.

**In Vitro Phosphorylation Assay**—rhAIRE monomer was resuspended in kinase buffer containing 40 mM HEPES (pH 7.4), 10 mM MgCl₂, and 3 mM MnCl₂. 100 µg/ml ATP was added into the suspension. The kinase reaction was initiated by the introduction of the corresponding protein kinases. We tested the phosphorylation of AIRE by Ab1 protein kinase (New England Biolabs), casein kinase I, casein kinase II (Promega Corp.), cAMP-dependent protein kinase A and PKC (Cell Signaling technology). The reactions were incubated for 30 min at 27 °C, after which AIRE was immunoprecipitated with anti-AIRE-agarose. The immunocomplexes were dissociated with 100 mM glycine buffer (pH 2.6), and the anti-AIRE-agarose conjugate was separated by centrifugation. After the supernatant was neutralized with 100 µl Tris buffer (pH 10.0). The eluates were saturated on native PAGE and transferred onto PVDF membrane. The membrane strips were probed with the corresponding anti-Tyr(P) or anti-Ser(P)/Thr(P) antibodies. One of the strips was probed with anti-AIRE antibody to verify the position of AIRE.

**Homology Modeling**—The full-length sequence of human AIRE (GenBank™ accession number NM 000383.1) was subjected to a global threading, and significant threading matches were filtered out considering high global and local Z scores. This process yielded a few templates for further use in homology modeling. The templates used include 11GSA, 1PFD, 1OTCA, 1PRCH, 2RAMB, and 2YHX. Alignments were performed using the program package GCG, and the model was built using Insight II and Discover (Molecular Simulations Inc., San Diego, CA) and evaluated with Swissmodel program (12).

**RESULTS**

**AIRE Purification, Antibody Production, and Refolding Assays**—The full-length construct of human AIRE (hAIRE) in pet32 showed a very low basal level of expression of AIRE, which was augmented severalfold after induction with isopropyl-1-thio-β-D-galactopyranoside. FIG. 1A, lanes 2 and 3). Although a minor fraction of AIRE appeared at the estimated molecular size of 56 kDa, as represented in lane 4 (FIG. 1A), which showed a purified fraction of AIRE. This fraction is termed AIRE monomer for descriptive purposes.

The processing of the rhAIRE under oxidation and refolding conditions brought about a major shift in the pattern of migration of this molecule (FIG. 1A, lane 5). Although a minor fraction of AIRE still appears at the 56-kDa position, the major bands appear at estimated molecular sizes of 110 and 220 kDa when the refolded rhAIRE was analyzed on native PAGE. Both bands migrated to the 56-kDa position on SDS-PAGE in the presence of 8 M urea (FIG. 1A, lane 6), suggesting the formation of dimers and tetramers of AIRE during the refolding process.

**AIRE Forms Dimers and Tetramers in Vivo**—Polyclonal antibodies raised in rabbit in our laboratory against the full-
length rhAIRE recognized the AIRE in mouse thymus extracts on Western blots (Fig. 1B, lane 1). Along with the expected 56-kDa band, we also observed the presence of anti-rhAIRE-positive bands at molecular weight positions of 110 and 220 kDa, suggesting the possible oligomerization of AIRE in vivo (Fig. 1B, lane 1). Both these high molecular weight bands were microsequenced, confirming that they are derivatives of AIRE (data not shown).

**AIRE Is Phosphorylated in Vivo**—A computer-based prediction (www.cbs.dtu.dk/services/NetPhos) of the properties of AIRE had indicated that this molecule is quite possibly phosphorylated at serine and/or at threonine at several locations, with one possible tyrosine phosphorylation site at Tyr-394. To verify this prediction, we immunoprecipitated AIRE from total tissue extracts (Fig. 1). More precisely, rhAIRE had indicated that this molecule is quite possibly phosphorylated at serine and/or at threonine at several locations, with one possible tyrosine phosphorylation site at Tyr-394. To verify this prediction, we immunoprecipitated AIRE from total tissue extracts (Fig. 1). Both these high molecular weight bands were microsequenced, confirming that they are derivatives of AIRE (data not shown).

To verify this prediction, we immunoprecipitated AIRE from total tissue extracts (Fig. 1). More precisely, rhAIRE had indicated that this molecule is quite possibly phosphorylated at serine and/or at threonine at several locations, with one possible tyrosine phosphorylation site at Tyr-394. To verify this prediction, we immunoprecipitated AIRE from total tissue extracts (Fig. 1). Both these high molecular weight bands were microsequenced, confirming that they are derivatives of AIRE (data not shown).

To verify this prediction, we immunoprecipitated AIRE from total tissue extracts (Fig. 1). More precisely, rhAIRE had indicated that this molecule is quite possibly phosphorylated at serine and/or at threonine at several locations, with one possible tyrosine phosphorylation site at Tyr-394. To verify this prediction, we immunoprecipitated AIRE from total tissue extracts (Fig. 1). Both these high molecular weight bands were microsequenced, confirming that they are derivatives of AIRE (data not shown).

To verify this prediction, we immunoprecipitated AIRE from total tissue extracts (Fig. 1). More precisely, rhAIRE had indicated that this molecule is quite possibly phosphorylated at serine and/or at threonine at several locations, with one possible tyrosine phosphorylation site at Tyr-394. To verify this prediction, we immunoprecipitated AIRE from total tissue extracts (Fig. 1). Both these high molecular weight bands were microsequenced, confirming that they are derivatives of AIRE (data not shown).

To verify this prediction, we immunoprecipitated AIRE from total tissue extracts (Fig. 1). More precisely, rhAIRE had indicated that this molecule is quite possibly phosphorylated at serine and/or at threonine at several locations, with one possible tyrosine phosphorylation site at Tyr-394. To verify this prediction, we immunoprecipitated AIRE from total tissue extracts (Fig. 1). Both these high molecular weight bands were microsequenced, confirming that they are derivatives of AIRE (data not shown).

**AIRE Can Be Phosphorylated in Vitro**—Five different kinases (Abl protein kinase, CK1, CK2, cAMP-dependent PKA and PKC) were selected considering the predicted affinity (www.cbs.dtu.dk/services/NetPhos) of one or all of them to phosphorylate recombinant human AIRE produced by prokaryotic expression. Abl protein kinase and the casein kinases used in this experiment could not bring about phosphorylation of rhAIRE in vitro (Fig. 1C, lanes 2–4). Interestingly, cAMP-dependent PKA brought about phosphorylation of rhAIRE (Fig. 1C, lane 5). To our surprise, we observed that the rhAIRE band shifted toward the 100-kDa position, indicating a possibility of dimerization of this molecule in response to phosphorylation. Protein kinase C also brought about phosphorylation of rhAIRE. Again, rhAIRE appeared to dimerize in response to phosphorylation although the monomer form of rhAIRE was also visible after the termination of the phosphorylation reaction (Fig. 1, lane 6).

**AIRE Can Bind to a Zinc Finger Binding Consensus**—We screened seven consensus sequences known to be recognized by zinc finger/flexucine zipper proteins (Table I) to examine the possibility of AIRE interacting with DNA through its zinc finger domains. In our initial attempts, we conducted gel shift assays with monomeric AIRE. None of the oligonucleotides brought about a gel shift when we used AIRE monomer in the assays (data not shown). Interestingly, the refolded recombinant protein showed a strong gel shift with EGR consensus, producing two bands corresponding to the respective dimer and tetramer positions of AIRE (Fig. 2A, lane 2). The addition of a 100 nM excess of cold consensus oligonucleotide brought about significant reduction in the intensity of the gel-shifted bands, signifying the specificity of the interaction (Fig. 2A, lane 3). Furthermore, the use of a mutant EGR consensus with base replacements at the underlined positions (GAATTCCAGGGGGGAGGGGCGGGA, GG → TA) did not bring about competitive removal of the gel-shifted bands (Fig. 2A, lane 4). Moreover, hot mutant EGR consensus also failed to show a gel-shifted band, confirming this observation (Fig. 2A, lane 5). This indicated the possible requirement of those underlined guanine residues in this oligonucleotide for its binding to AIRE.

Furthermore, the use of a mutant EGR consensus with base replacements at the underlined positions (GAATTCCAGGGGGGAGGGGCGGGA, GG → TA) did not bring about competitive removal of the gel-shifted bands (Fig. 2A, lane 4). Moreover, hot mutant EGR consensus also failed to show a gel-shifted band, confirming this observation (Fig. 2A, lane 5). This indicated the possible requirement of those underlined guanine residues in this oligonucleotide for its binding to AIRE.

To further confirm that the zinc finger domain of AIRE is involved in this interaction, we performed a competition assay where AIRE was made to interact with EGR consensus in the presence of a known synthetic zinc finger peptide ZIF 268 showing strong affinity toward the EGR consensus. In the absence of the competitor peptide, AIRE brought about a gel shift as shown in Fig. 2A, lane 2. In the presence of 100 nM ZIF 268, AIRE could not bring about visible mobility shifts over a concentration range of 100–300 nM (Fig. 2B, lanes 3–5), but a 4-fold excess of AIRE (400 nM) restored the mobility shift (Fig. 2B, lane 6). A further increase in the concentration of AIRE resulted in the appearance of the lower gel-shifted band (corresponding to AIRE dimer) as well (Fig. 2B, lanes 7–10).

**Consensus Sequence Recognized by AIRE Is Partly Different from That of EGR**—We screened an oligo library with pre-selected guanine doublets with a random seven-nucleotide spacer (G2N7G2 library) and a random oligo library (N 26 library) to identify the AIRE binding consensus sequence. The libraries were subjected to five generations of pull-down assays, and the products after the fifth amplification were cloned and sequenced.

We analyzed 59 clones selected out by AIRE from the G2N7G2 library by direct sequencing using M13 forward primer. The sequences were aligned and shaded using Genesys software (Fig. 3A). Alignment of the sequences from 32 clones is presented because of software limitations. A quantity-based generation of consensus from the aligned sequences was performed using Genedoc software (Fig. 3A). The pre-selected G-doublets are shaded gray in the alignment. It is clear that AIRE preferred an A/T-rich neighborhood around GG for its binding. A careful examination of the consensus line showed that this sequence contained two ATGGTTA motifs (underlined in double lines) and a weak TATA-box (TTAGTTA) (underlined in single line).

To verify whether this observation was because of a bias in the library used, we cloned the unselected library into TAg-
 Relative Binding Affinity of AIRE to TTATTA and to ATTTGTTA Motifs—To address the relative binding affinities of AIRE to the TTATTA (T-box) motif and to the ATTTGTTA motif (G-box), we performed gel shift assays and competitions with various combinations of four oligonucleotides: oligo-TGG, oligo-TG, oligo-GG, and oligo-G (Table I). Two major gel shift patterns emerged from these analyses. When the labeled oligo-GG and oligo-G are used, two shifted bands were observed (Fig. 4 lanes 9–16). The upper band corresponds to the AIRE-tetramer position, and the lower band corresponds to the dimer position. Interestingly, a higher band was observed when both labeled oligos with a T-box (oligo-TGG and oligo-TG) were analyzed (Fig. 4, lanes 1–8).

AIRE showed strong binding with hot oligo-TGG (Fig. 4, lane 4), which could be competed out substantially with cold oligo-TG (Fig. 4, lane 1). Cold oligo-GG brought about partial inhibition in gel shift (Fig. 4, lane 2), but cold oligo-G did not reduce the shifted band of hot oligo-TGG (Fig. 4, lane 3). Oligo-TG also showed good binding with AIRE protein (Fig. 4, lane 8). This binding could be totally abolished in competitions with oligo-TGG (lane 5), oligo-GG (lane 6), or oligo-G (lane 7). Furthermore, the binding of AIRE to oligo-GG (lane 12) could be blocked significantly by oligo-TGG (lane 9) and to a lesser extent by oligo-TG (lane 10) and oligo-G (lane 11). The binding of AIRE to oligo-G showed a pattern similar to that of oligo-GG (Fig. 4, lanes 13–16). It is interesting to note that the shifted AIRE dimer is not detectable without competitor oligos (Fig. 4, lanes 12 and 16) but appears in the presence of competitors. This is because the labeled oligos only become available for binding with the AIRE dimers when the competitor oligos occupied the binding sites of the AIRE tetramers. These results further suggest that the AIRE tetrampers have higher binding affinity than the AIRE dimers.

Homology Modeling—A threading of the sequence of the full-length human AIRE protein returned several templates that were used to generate an energy-minimized model for the full-length protein (Fig. 5). The model depicts an important conformational ingenuity in that both the PHD domains and the SAND domain are arranged in a linear fashion, providing maximal DNA binding surface. The positioning of the leucine zipper well within the middle DNA binding domain (PHD1) would permit the zip-locking of the protein dimer onto the DNA. The leucine-zipper skeleton is shown in red in Fig. 5.

FIG. 2. Electrophoretic mobility shift assays with rhAIRE. A, incubation of rhAIRE with EGR-1 consensus brought about a weak (corresponding to label D) and a strong (corresponding to label T) gel-shifted bands (lane 2). The relative positions of the bands indicated that the lower and the upper bands resulted from the binding of AIRE dimer and tetramer to the oligonucleotide, respectively. The use of a 10-fold excess of cold consensus oligonucleotide brought about a total reduction of the gel-shifted bands (lane 3). However, the use of a 50-fold excess of non-competitor oligonucleotide (EGR-nonconsensus) did not compete out the gel-shifted bands (lane 4). Use of a hot EGR nonconsensus (Table I) did now show gel shift (lane 5). Lane 1 shows the free label, competition assays with zinc finger peptide ZIF 268. Though ZIF 268 would bind EGR consensus, the complex moved along with the free label, resulting in the absence of any distinct gel-shifted band (lane 2). In the presence of 100 μM ZIF 268, AIRE could not bring about visible mobility shifts over a concentration range of 100–300 pM (lanes 3–5), but a 4-fold excess of AIRE (400 pM) restored the mobility shift (lane 6). A further increase in the concentration of AIRE (500–800 pM) resulted in the appearance both the upper and the lower gel-shifted bands (lanes 7–10). D, dimer; T, tetramer.
in vitro very much like Sp100 (10). Recombinant full-length hAIRE from the E. coli expression system did not dimerize, probably because post-translational processing of AIRE may be prerequisite for dimerization. Because it is known that proteins could be unfolded and refolded successfully in vitro, our initial attempts were to examine whether we could develop a strategy to refold rhAIRE in vitro. Under the conditions stated in this paper, AIRE refolded and dimerized as expected. At the same time, we observed the formation of tetramers as well (Fig. 1). If tetramer formation were not an artifact, it would imply that AIRE has dimerization and tetramerization domains in its structure. This appears to be the case if we combine the observation of Pitkanen et al. (10) that AIRE could dimerize utilizing its N-terminal HSR domain and the fact that AIRE does have a leucine zipper extending over amino acids 319–341. At this moment, the possibility of oligomerization of AIRE through its leucine zipper is not experimentally proven. But it is established that leucine zipper proteins dimerize, engaging the zipper domains of the interacting partners. In that event, it is quite likely that AIRE might dimerize using the leucine zipper motif and would tetramerize using the HSR domain. The functional significance of the formation of dimers and tetramers of AIRE would depend on whether such molecular interactions take place in vivo or not. To address this issue, we performed a native PAGE followed by Western blot analysis of human thymus extracts. Anti-rhAIRE antibody recognized three bands at molecular weight positions of 56, 110, and 220 kDa. These are the expected positions of monomer, dimer, and tetramer.
In order to gain insight into the structural architecture of this molecule in space, we performed molecular modeling and threading of the AIRE sequence against the available Protein Data Bank templates. This analysis revealed that the leucine zipper (Fig. 1C) and the zinc finger domains are surface-positioned in this molecule, making them available for DNA-protein interactions. The leucine zipper (drawn in a red skeleton) is within the PHD1. The SP-100 HSR is free on the surface. From this model, we interpret that AIRE could dimerize using the leucine zipper motif and could tetramerize using the HSR domain.

The importance of the leucine zipper domain and the zinc finger domains in deciding the functional competence of AIRE was established by microsequencing analysis. Thus, the potential presence of dimer and tetramer forms of AIRE in human thymus is established.

Homo- and/or hetero-oligomerization of proteins is important in the modulation of their function. In vitro, proteins oligomerize in response to stimuli of intrinsic or extrinsic origin. Phosphorylation or dephosphorylation of proteins at specific targets by respective kinases or phosphatases is a common pathway through which biological signal transduction propagates. Because recombinant AIRE did not possess the capability to autodimerize, we suspected a possible activation of this molecule by phosphorylation that might be required for its dimerization. A computer-based prediction of the possible phosphorylation sites in AIRE pointed toward a high chance of phosphorylation of serine/threonine at multiple locations. In the light of this prediction, we tried a series of kinases for their ability to phosphorylate AIRE in vitro. Interestingly, we found that cAMP-dependent protein kinase A and PKC could heavily phosphorylate AIRE (Fig. 1C). It is also important to state at this point that the phosphorylation of AIRE by cAMP-dependent PKA and PKC also led to the dimerization of this molecule. Although the relevance of this in vitro demonstration cannot be applied to the in vivo scenario, we feel that there is a strong possibility that AIRE is a substrate for cAMP-dependent protein kinase A and/or PKC. Furthermore, phosphorylation of AIRE appears to be the trigger for its dimerization in vitro and possibly in vivo. This observation, in conjunction with the report that AIRE interacts with CREB-binding protein (10) suggest that AIRE could act as a downstream modulator of cAMP-dependent signal transduction and transcription regulation. Because the human CREB protein contains the leucine zipper motif (14), it is also possible that there is a direct interaction between CREB and AIRE.

It is well known that zinc finger proteins represent an important class of regulatory DNA-binding proteins. Different varieties of zinc finger proteins are all linked by the utilization of zinc ions to add a structural element to the binding component. One characteristic of this motif is that it recognizes G-rich DNA (15–19). Because AIRE contained both leucine zipper and zinc finger domains, we decided to characterize the potential DNA binding activity of AIRE using oligonucleotides known to be recognized by proteins having these distinct domains. Thus, we performed gel shift assays with AIRE and consensus sequences recognized by a number of leucine zipper/zinc finger proteins (Table I). The results showed that AIRE introduced band shift with the EGR consensus but not with the EGR nonconsensus oligonucleotide. The band shift could be competed out with an excess of cold EGR oligonucleotide and also with ZIF 268 synthetic peptide, indicating that the band shift is introduced by the interaction of zinc finger motifs in AIRE and EGR consensus (Fig. 2). Although AIRE monomer failed to bring about band shift, the AIRE dimer and tetramer forms brought about strong band shifts under our assay conditions, implicating the requirement of oligomerization of this molecule for it to gain the DNA binding activity. Pull down assays with an oligonucleotide library with sequence of N₆GGN₆GGN₆ indicated the selection of two motifs in the oligonucleotides that would bind AIRE. The first motif is a weak TATA-box (TTATTA), and the second motif is a tandem repeat of ATTGTTTA (G-box). Gel shift competition assays with oligonucleotides oligo-TGG, oligo-TG, oligo-GG, and oligo-G showed that oligo-TGG exhibited the highest AIRE binding capability that could not be abolished by oligo-G or oligo-G (Fig. 4, lanes 1–4). The binding of AIRE to oligo-TG appeared to be highly unstable, as all the three competitors used in this study could abolish the gel shift (Fig. 4, lanes 5–8). The binding of AIRE to oligo-GG (Fig. 4, lanes 9–12) or oligo-G (Fig. 4, lanes 13–16) appeared to be comparable. Based on the binding affinity of AIRE with these oligonucleotides under the given conditions of competition, we conclude that we could arrange them in the following descending order: oligo-TGG > oligo-GG > oligo-G > oligo-TG. It is noteworthy to state that the complexes between AIRE and oligo-TGG and oligo-TG migrated to a larger molecular mass position compared with the expected positions of AIRE dimer and AIRE tetramer. It is likely that high affinity binding between AIRE and the TGG or TG oligonucleotides produce complexes with more than one tetramer unit of AIRE per oligonucleotide molecule in our assay system.

The construction of a three-dimensional model for the full-length human AIRE protein with the available templates gives us an insight into the structural architecture of this molecule in space. As could be interpreted from this model, this protein has the following visible domains: SAND (197–262), PHD (298–341 and 433–476), the N-terminal β-helix-loop-helix (HSR), and the leucine zipper (319–341). In the folded form, both the HSR and the leucine zipper domains are available for molecular interaction. A recent determination of the three-dimensional structure of the SAND domain from Sp100b represented a novel α/β-fold, in which a conserved KDWD sequence motif is found within an α-helical, positively charged surface patch (20). In our model for AIRE, the proposed DNA binding SAND and the known DNA binding zinc finger domains are also exposed for DNA-protein interactions. These interpretations match well with our experimental observations.

The importance of the leucine zipper domain and the zinc finger domains in deciding the functional competence of AIRE...
becomes evident if we correlate the known AIRE mutations with the APECED disease. So far more than 20 different mutations of the AIRE gene have been identified in APECED patients (21). The PHD domains in AIRE are located at amino acids 299–340 (PHD1) and 434–475 (PHD2). The leucine zipper is located at position 319–341. The major mutations identified so far include the Finnish mutation R257X and the 13-bp deletion in exon 8 (nucleotides 1094–1106), resulting in a premature termination codon at amino acid 479 and a protein that is 66 amino acids shorter and a T→C transition at +2 of intron 3 (22). A recent report revealed novel frameshift mutations of the AIRE gene. These include an insertion of a cytosine at nucleotide 29,635 at exon 10 (29635insC), which should lead to a premature termination at the codon 520, yielding a truncated protein missing the third LXXLL motif, and a deletion of a guanine at nucleotide 33,031 at exon 13 (33031delG), which should result in a premature termination at the codon 371, producing a truncated protein missing the second plant homeodomain-type zinc finger motif and the third LXXLL motif, and a deletion of a guanine at nucleotide 33,031 at exon 13 (33031delG), which should result in a premature termination at the codon 520, yielding a truncated protein missing the third LXXLL motif (21). The majority of the mutations listed above does affect the PHD domain(s) and/or the leucine zipper domain, although two of the mutations affect only the last LXXLL motif. In addition to what was thought about the implications of these mutations to date, our data strongly indicate that mutations affecting the leucine zipper domain and/or the zinc finger domain would make AIRE incompetent to bind DNA. This would affect the expression of gene(s) under the regulatory influence of AIRE.

With the identification of AIRE as a DNA-binding protein, two relevant points to be addressed are the factors influencing AIRE expression and the gene(s) under the regulatory influence of AIRE. It appears that the first question has already been addressed, since a recent report states that AIRE was absent in RelB-deficient mouse (23). The target genes of AIRE are yet to be identified.

REFERENCES

1. Ahonen, P. (1985) Clin. Genet. 27, 535–542
2. Aaltonen, J., Bjorjes, P., Sandkuijl, L., Perheentupa, J., and Peltonen, L. (1984) Nat. Genet. 8, 83–87
3. Bjorjes, P., Aaltonen, J., Vikman, A., Perheentupa, J., Ben Zion, G., Chiumello, G., Dahl, N., Heideman, P., Hoorweg-Nijman, J. J., Mathivon, L., Mullis, P. E., Poh, M., Ritzen, M., Romee, G., Shapiro, M. S., Smith, C. S., Solyom, J., Zlotogora, J., and Peltonen, L. (1996) Am. J. Hum. Genet. 59, 879–886
4. Heino, M., Scott, H. S., Chen, Q., Peterson, P., Maebu, U., Papasavvas, M. P., Mittaz, L., Burag, C., Rossler, C., Chrousos, G. P., Stratakis, C. A., Nagamine, K., Kudoh, J., Shimizu, N., Maclaren, N., Antonarakis, S. E., and Krohn, K. (1999) Hum. Mutat. 13, 69–74
5. The Finnish-German Autoimmune Polyendocrinopathy Candidiasis Ectodermal Dystrophy Consortium (1997) Nat. Genet. 17, 399–403
6. Nagamine, K., Peterson, P., Scott, H. S., Kudoh, J., Minoshima, S., Heino, M., Krohn, K. J., Lalioti, M. D., Mullis, P. E., Antonarakis, S. E., Kawasaki, K., Asakawa, S., Ita, P., and Shimizu, N. (1997) Nat. Genet. 17, 393–398
7. Gibson, T. J., Ramu, C., Gemund, C., and Aasland, R. (1998) Trends Biochem. Sci. 23, 242–244
8. Seeler, J. S., Marchio, A., Sitterlin, D., Transy, C., and Dejean, A. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 7316–7321
9. Bjorjes, P., Pelto-Huikko, M., Kaukonen, J., Aaltonen, J., Peltonen, L., and Ulmanen, I. (1999) Hum. Mol. Genet. 8, 259–266
10. Piikkanen, J., Doucas, V., Sterneck, T., Nakajima, T., Aratani, S., Jensen, K., Will, H., Vahamurto, P., Ollila, J., Vihinen, M., Scott, H. S., Antonarakis, S. E., Kudoh, J., Shimizu, N., Krohn, K., and Peterson, P. (2000) J. Biol. Chem. 275, 16892–16899
11. De Cesare, D. and Sassone-Corsi, P. (2000) Prog. Nucleic Acid Res. Mol. Biol. 64, 343–369
12. Schwede, T., Niemand, A., Guex, N., and Peitsch, M. C. (2000) Res. Microbiol. 151, 107–112
13. Bjorjes, P., Halonen, M., Palvimo, J. J., Kolmer, M., Aaltonen, J., Ellenon, P., Perheentupa, J., Ulmanen, I., and Peltonen, L. (2000) Am. J. Hum. Genet. 67, 378–392
14. Dwarki, V. J., Montminy, M., and Verma, I. M. (1990) EMBO J. 9, 225–232
15. Desjarlais, J. R., and Berg, J. M. (1992) Proteins 12, 101–104
16. Desjarlais, J. R., and Berg, J. M. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 7345–7349
17. Desjarlais, J. R., and Berg, J. M. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 2256–2260
18. Desjarlais, J. R., and Berg, J. M. (1994) Proc Natl. Acad. Sci. U. S. A. 91, 11099–11103
19. Pavletich, N. P., and Babo, C. O. (1991) Science 252, 809–817
20. Bottomley, M. J., Collard, M. W., Haggevik, J. I., Liu, Z., Gibson, T. J., and Sattler, M. (2001) Nat. Struct. Biol. 8, 568–570
21. Ishii, T., Suzuki, Y., Ando, N., Matsuo, N., and Ogas, T. (2000) J. Clin. Endocrinol. Metab. 85, 2922–2926
22. Wang, C. Y., Davoudi-Semnani, A., Huang, W., Connor, E., Shi, J. D., and She, J. X. (1998) Hum. Mutat. 103, 681–685
23. Heino, M., Peterson, P., Sillanpaa, N., Guerin, S., Wu, L., Anderson, G., Scott, H. S., Antonarakis, S. E., Kudoh, J., Shimizu, N., Jenkinson, E. J., Nasquet, P., and Krohn, K. J. (2000) Eur. J. Immunol. 30, 1884–1893
