Leukemic T-cell Precursors from T-lineage All Patients Are Characterized by Profound Ku80 Deficiency as well as Functional IK Deficiency with Very Low Ikaros Target Gene Expression Levels

Sanjive Qazi, Zahide Ozer, Cheney Mao, Fatih M. Uckun

1. Systems Immunobiology Laboratory and Developmental Therapeutics Program, Children’s Center for Cancer and Blood Diseases, Children’s Hospital Los Angeles, Los Angeles, CA
2. Department of Biology and Bioinformatics Program, Gustavus Adolphus College, 800 W College Avenue, St. Peter, MN 56082
3. Viva Biotech, Chicago, IL 60612
4. Department of Pediatrics, University of Southern California Keck School of Medicine, Los Angeles, CA 90027, USA
5. Developmental Therapeutics Program, USC Norris Comprehensive Cancer Center, Los Angeles, CA 90089

Abstract We recently found that Ku regulates Ikaros (IK) function and Ku deficiency causes a lymphoproliferative disorder of thymic T-cell precursors with functional IK deficiency. We now examined the expression levels of validated IK target genes that harbor IK binding sites in primary leukemic cells from 1,104 pediatric ALL patients in relationship to Ku expression levels. Hierarchical cluster analysis of cellular gene expression profiles revealed a highly significant correlation between Ku expression level and expression levels of validated IK target genes. Notably, the expression levels of 12 IK-regulated lymphoid priming genes were significantly upregulated in human lymphocyte precursor cells from primary bone marrow specimens of pediatric patients with ALL expressing high levels of Ku. The observed striking Ku-dependency of the IK target gene expression levels taken together with the results of our earlier RNAi experiments indicate that IK function in human lymphocyte precursors is controlled by Ku expression levels. Leukemic T-cell precursors from children with T-lineage ALL were characterized by profoundly diminished Ku80 transcript levels as well as functional IK deficiency with very low IK target gene expression levels. These results extend our previous studies in mice and indicate that Ku-deficiency is also a contributor to the functional IK-deficiency in pediatric T-lineage ALL.

Keywords Ikaros (IK); T-cell precursors; IK deficiency

Introduction Zinc finger (ZF) proteins contain multiple cysteine (Cys) and/or histidine (His) residues and use zinc coordination to stabilize their folding. The Cys2His2 ZF is the most prevalent protein motif in mammalian cells and defines the largest family of sequence-specific DNA binding proteins. Many DNA binding proteins contain several ZF that make tandem contacts along the target DNA segment. Cys2His2 ZF domains participate in protein-DNA, protein-RNA as well as protein-protein interactions (Coleman, 1992; Klug and Schwabe, 1995; Mackay et al., 1998; Nardelli et al., 1992). Ikaros (IK) is a ZF-containing sequence-specific DNA binding protein that plays a pivotal role in immune homeostasis through transcriptional regulation of the earliest stages of lymphocyte ontogeny and differentiation by both (a) gene transcriptional activation via efficient transcription initiation as well as elongation and (b) repression (Georgopoulos, 2002; Cobbs et al., 2000; Thompson et al., 2007; Merkenschlager, 2010; Yoshida et al., 2010; Bottardi et al., 2011). IK is a pleotropic protein with important regulatory functions in gene silencing, DNA replication and genomic
stability, G0/G1 and G1/S transition in cell cycle, and lineage commitment/specification of hematopoietic stem cells. IK also exhibits a tumor suppressor function in lymphocyte precursors, which has been attributed to its ability to repress expression of oncogenic genes via chromatin remodeling in association with the SWI/SNF remodeling complex (Georgopoulos, 2002; Cobbs et al., 2000; Thompson et al., 2007; Merkenschlager, 2010) and recruitment of potentially oncogenic proliferation-promoting genes to pericentromeric heterochromatin (PC-HC) as well as the regulatory control of the homing of the nucleosome-remodelling and histone-diacytelase (NuRD) complex to lymphoid lineage-specific IK target genes that are required for orderly differentiation of lymphocyte precursors (Georgopoulos, 2002; Cobbs et al., 2000; Thompson et al., 2007; Merkenschlager, 2010; Yoshida et al., 2010; Bottardi et al., 2011; Zhang et al., 2011; Dovat et al., 2011). Impaired DNA binding activity of IK has been associated with a release of NuRD from IK target genes to cause both a maturational arrest in lymphocyte ontogeny and an “illegitimate” activation of a network of genes that promote leukemogenesis (Zhang et al., 2011). Functional deficiency of IK due to expression of non-DNA binding dominant-negative IK isoforms caused by aberrant splicing (Sun et al., 1999) or genomic mutations (Mullighan et al., 2009) has been detected in leukemic lymphocyte precursors from patients with acute lymphoblastic leukemia (ALL), the most common form of childhood cancer.

Currently, our knowledge regarding the upstream regulators of IK function is relatively limited (Georgopoulos, 2002; Cobbs et al., 2000; Thompson et al., 2007; Merkenschlager, 2010). IK function, stability, and subcellular localization are generally thought to be regulated by posttranslational modification and heterodimerization with other members of the IK family of DNA binding proteins (Georgopoulos, 2002; Sun et al., 1996). Phosphorylation of IK by casein kinase II (CK2) inhibits its many functions and promotes its degradation via the ubiquitin/proteosome pathway (Gurel et al., 2008). Conversely, dephosphorylation of IK by protein phosphatase 1 is critical for its ability to bind to target DNA sequences, localize to PC-HC in the nucleus, and exert its regulatory functions (Popescu et al., 2009). In a recent study, we identified the spleen tyrosine kinase (SYK) as a posttranslational regulator of IK and showed that SYK-induced activating phosphorylation of IK at unique C-terminal serine phosphorylation sites S358 and S361 is essential for its nuclear localization and optimal transcription factor function (Uckun et al., 2012). Likewise, BTK was identified as a regulator of IK that phosphorylates IK at unique phosphorylation sites S214 and S215 in the close vicinity of its ZF4 within the DNA binding domain, thereby augmenting its sequence-specific DNA binding activity (Ma et al., 2013).

Our recent studies revealed a previously unknown function of the DNA repair protein Ku, which plays an important role in repair of DNA double-strand breaks by nonhomologous end joining (NHEJ) (Suzuki et al., 2010; Walker et al., 2001), as a partner of IK that physically associates with IK thereby augmenting its nuclear localization and sequence-specific DNA binding activity (Ozer et al., 2013). Ku is the only protein outside the IK family of ZF proteins shown to non-enzymatically improve the function of IK as a sequence-specific DNA binding protein via heterodimerization. RNA interference (RNAi) experiments using Ku70-specific or Ku80-specific small interfering RNA (siRNA) further demonstrated that Ku regulates the expression levels of validated IK target genes in a human cell line (Ozer et al., 2013). The interaction of Ku components with IK likely contributes to the anti-leukemic effects of IK as a tumor suppressor, because Ku70 as well as Ku80 haploinsufficiency in mice caused development of a lymphoproliferative disorder (LPD) involving CD2+CD4+CD8+CD1+IL7R+ thymic T-cell precursors with functional IK deficiency (Ozer et al., 2013). It is noteworthy that comprehensive bioinformatic studies on haploinsufficiency have indicated that the Ku80/XRCC5 gene is highly likely to be haploinsufficient (Huang et al., 2010). As Ku70 and Ku80 are not haploinsufficient for DNA double strand break (DSB) repair and even Ku-null cells have normal DNA repair activity due to hyperactive...
alternative NHEJ pathway (Fattah et al., 2010), the development of a T-cell precursor hyperplasia in haplodeficient mice (Ozer et al., 2013) cannot be explained by DNA repair deficits and provides compelling evidence that the interaction of Ku components with IK likely contributes to the anti-leukemic effects of IK as a tumor suppressor. Thus, Ku deficiency may in part be responsible for the IK malfunction in proliferating T-cell precursors in mice, which is a hallmark of pediatric high-risk T-lineage ALL (Sun et al., 1999).

The purpose of the present study was to explore if the IK-regulatory function of Ku might play a role in the biology of pediatric T-lineage ALL. We examined the expression levels of validated IK target genes that harbored IK binding sites (Uckun et al., 2012) in primary leukemic cells from 1104 pediatric ALL patients in relationship to Ku expression levels. Hierarchical cluster analysis of cellular gene expression profiles revealed a highly significant correlation between Ku expression level and expression levels of validated IK target genes. Leukemic T-cell precursors from primary bone marrow specimens of children with T-lineage ALL were characterized by profoundly diminished Ku80 transcript levels as well as functional IK deficiency with very low IK target gene expression levels. These results uniquely implicate Ku-deficiency as a contributor to the functional IK-deficiency in pediatric T-lineage ALL.

1 Results and Discussion

1.1 Ikaros-Ku Molecular Complex

By using multiple assay platforms, including coimmunoprecipitations, pull-down assays, and EMSAs with supershift assays, we recently demonstrated that native as well as recombinant Ku and IK proteins form heterodimers capable of binding IK target DNA sequences with greater affinity that IK homodimers (Ozer et al., 2013). In order to gain further insights into the physical interactions between Ku and IK proteins, we examined the ability of MBP-tagged purified recombinant Ikaros 1 (IK1) and Ikaros 5 (IK5) proteins in solution phase to bind to immobilized recombinant Ku70, Ku80, and Ku70/Ku80 heterodimer by surface plasmon resonance (SPR), which permits direct measurements of the association and dissociation kinetics of binding interactions. Reflective of a high biological affinity, the SPR-based dissociation constant values for the IK1-Ku binding interactions were in the low nanomolar range. MBP-IK1 (but not MBP-IK5) showed high affinity binding to Ku70 (Figure 1A.1), Ku80 (Figure 1A.2) and Ku70/Ku80 heterodimer (Figure 1A.3) with nanomolar KD values. The KD values for the IK1-Ku70, IK1-Ku80, and IK1-Ku70/Ku80 binding interactions were 15.2 nM, 14.1 nM, and 15.3 nM, respectively.

We constructed a molecular model of the Ku-IK1 hetero-complex focusing on the contribution of the N-terminal zinc finger structures, ZF (1-4) of IK (110-256) to the experimentally documented differential binding of Ikaros isoforms to Ku components. Our refined molecular model of the IK1 protein complexed with the Ku70/80 heterodimer is depicted in Figure 1B1-B3. According to our model, Ku80 subunit of Ku protein forms an excellent docking site for IK1 near the entrance of the “ring” along the minor groove that fits well with the contours of the ZF (2-4) domains of IK1 approaching along with the major groove and extending almost a complete turn of DNA duplex. As a result, IK1 and Ku80 form an interface with at least 970 Å2 buried surface. There are 40% polar and 60% non-polar residues involved in the interaction. The following residues of Ku80 are located in the interaction region: E49, K51, R242, H246, P248, R250, R260, Y264, K273-4, T275, T277, K282, T283, L284, K285, K286, E287, E329, K334, S335,E336, K338, A367, R368, D369, D370 and K399. Ku70 is in the middle of the DNA major groove and there is only 194 Å2 contact area for Ku70 from this orientation. However, IK1-ZF (2-4) can readily interact with Ku70 assuming Ku70 is flexible and the segment of the βO-to-α12 turn in Ku70 is capable of adopting a more open conformation. Our model illustrates the importance of IK1 ZF domains for its interactions with Ku and readily explains why IK1 exhibited better binding to Ku components (Ku70 and Ku80) than IK5 lacking ZF (2-4) in SPR assays.

IK1 protein contains four Cys2His2 ZF motifs near its
N terminus that contribute to sequence-specific DNA binding (Georgopoulos, 2002; Thompson et al., 2007). Our study indicates that these motifs, especially ZF (2-4), are also important for its interaction with Ku. To bind all three ZF at once, it would require a protein to have a special helical conformation to fit the extended contours in which three ZFs bind to DNA along its major groove. Ku protein that binds DNA along its minor groove with its Ku80 component fitting the contours of IK1 ZF (2-4) is uniquely suited to bind the ZF of IK1, as confirmed by our modeling studies.

Figure 1 Ikaros-Ku molecular complex [A] analysis of Ku-Ikaros binding interactions by surface plasmon resonance (SPR)

Note: Representative BIAcore sensorgrams showing the binding of MBP-IK1 vs. MBP-IK5 to Ku70 (A.1), Ku80 (A.2) and Ku70/Ku80 (A.3), that were immobilized on an NTA sensor chip. SPR assays were performed and analyzed, as detailed in Materials and Methods. MBP-IK1 (but not MBP-IK5) exhibited high affinity binding to Ku70, Ku80 as well as Ku70/Ku80. [B] A Model of the Structural Interface Between IK1 and Ku Proteins. [B1] Space filling model of IK1 and Ku70/Ku80 complex based on the crystal structure of the Ku heterodimer bound to DNA and a homology model of IK1 ZF fragment (110-256). Ku70/Ku80-light blue; IK1-white, DNA-atom-type colored (carbon-green, oxygen-red, phosphate-magenta, nitrogen-blue). [B.2&B.3]. Ribbon representation of the IK1 and Ku70/Ku80 complex model is shown in the following colors: Ku70-green, Ku80-orange, IK1-red, DNA-white for backbone and blue for groove. The two subunits of the Ku heterodimer share little sequence identity (15%) with each other. However, the crystal structure of Ku70/Ku80 revealed that both subunits have the same folding, domain structure, and overall tertiary structure except for the terminal regions (shown in different colors). The shared structural features include the N-terminal α/β domain (34-250 and 6-238 for Ku70 and Ku80 respectively) and the β-barrel domain (116 residues). The α/β domains are positioned at the periphery of the Ku heterodimer and contribute little to the DNA binding and domain interface. On the other hand, the β-barrel domains of the heterodimer use seven β-strands to provide a symmetrical circle forming the cradle of the DNA-binding site by pointing one end of each barrel towards the DNA, as revealed by the crystal structure of Ku/DNA complex. The protein dyad axis of the dimer coincides with the bound DNA duplex axis. The DNA is positioned centrally through the ring formed by the protein dyad. All three ZF in IK1 (Ik_ZF2-4) are in close and extensive contact with Ku80. Likewise, a similar interaction is contemplated with Ku70 with the segment of the β0-to-α12 turn in Ku70 and Ku80 C-terminus (541-545) adopting a different conformation, as observed on the other end of the ring.
Figure 2 Effects of Ku Expression Levels on Expression Levels of Lymphoid-Priming Genes in Primary Lymphocyte Precursors from ALL Patients

Note: A one-way agglomerative hierarchical clustering technique was used to organize expression patterns using the average distance linkage method such that IK-regulated lymphoid priming genes (18) (rows) having similar expression across patients were grouped together (average distance metric). The heat map depicts expression values represented by standard deviation units above (red) and below the mean (green). Dendrograms were drawn to illustrate similar gene-expression profiles from joining pairs of closely related gene expression profiles, whereby genes joined by short branch lengths showed most similarity in expression profile across patients.

[B] Co-regulation of IKZF1, Ku and Lymphoid-priming genes in Primary Lymphocyte Precursors from ALL patients. Expression values expressed as Standard Deviation units calculated from 1104 samples were compiled for the 5 studies and rank ordered according to the mean expression of three highly correlated transcripts (208642_s_at (XRCC5), 208643_s_at (XRCC5), 200792_at (XRCC6). These samples were also rank ordered according to IKZF1 expression level (205038_at, 205039_s_at, 216901_s_at, 227344_at and 227346_at; 3 of these were common in all Affymetrix platforms (205038_at, 205039_s_at, 216901_s_at)). T-tests were performed for the combined Standard Deviation units from the 5 datasets (2-sample, Unequal variance correction, p-values<0.05 deemed significant) to reveal 13 transcripts significantly upregulated in specimens with both high Ku and high IKZF1 expression.
1.2 Correlation between expression levels of Ku and Ikaros-regulated lymphoid priming genes

IK has been shown to direct lymphoid priming during the earliest stages of lymphocyte ontogeny by upregulating the expression of specific genes (Yoshida et al., 2010). Notably, 13 transcripts representing 12 IK-regulated lymphoid priming genes were significantly upregulated in human lymphocyte precursor cells from primary bone marrow specimens of pediatric patients with ALL expressing high levels of both IKZF1 and Ku (Figure 2). The 5 most significantly upregulated lymphoid priming genes in specimens with high Ku expression were SOX4 (0.64 SD units, P = 1.4 × 10^{-14}), CSF1R (0.52 SD Units, P = 8.6 × 10^{-11}), IL7R (0.52 SD units, P = 5.1 × 10^{-11}), DNTT (0.51, P = 1.6 × 10^{-10}), NOTCH1 (0.47, P = 1.4 × 10^{-14}) (Figure 2). Hierarchical cluster analysis revealed a subset of 3 genes (RAG1, DNTT and MEF2C (2 transcripts)) that were highly co-regulated with 3 transcripts of IKZF1 (Figure 2). This striking correlation between the expression levels of IK-regulated lymphoid priming genes and Ku expression levels indicates that Ku is involved in the regulation of IK function in primary human lymphocyte precursors.

Table 1 Ikaros target gene expression levels in human lymphocyte precursors according to ikzf1 and ku transcript levels

<table>
<thead>
<tr>
<th>Affy ID</th>
<th>Genesymbol</th>
<th>Ku transcript levels</th>
<th>Ikaros transcript levels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>StdDev Units</td>
<td>T-test P-value</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[High Ku(N=314)]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Low Ku(N=324)]</td>
<td></td>
</tr>
<tr>
<td>205273_s_at</td>
<td>PITRM1</td>
<td>1.21</td>
<td>1.2 × 10^{-44}</td>
</tr>
<tr>
<td>213761_at</td>
<td>MDM1</td>
<td>1.13</td>
<td>2.2 × 10^{-44}</td>
</tr>
<tr>
<td>205884_at</td>
<td>ITGA4</td>
<td>1.19</td>
<td>2.7 × 10^{-46}</td>
</tr>
<tr>
<td>204709_s_at</td>
<td>KIF23</td>
<td>1.13</td>
<td>1.3 × 10^{-46}</td>
</tr>
<tr>
<td>204117_at</td>
<td>PREP</td>
<td>1.06</td>
<td>1.5 × 10^{-41}</td>
</tr>
<tr>
<td>37950_at</td>
<td>PREP</td>
<td>1.04</td>
<td>1.0 × 10^{-38}</td>
</tr>
<tr>
<td>213416_at</td>
<td>ITGA4</td>
<td>1.05</td>
<td>1.9 × 10^{-40}</td>
</tr>
<tr>
<td>220597_at</td>
<td>MDM1</td>
<td>0.98</td>
<td>1.5 × 10^{-36}</td>
</tr>
<tr>
<td>205885_s_at</td>
<td>ITGA4</td>
<td>0.99</td>
<td>3.0 × 10^{-35}</td>
</tr>
<tr>
<td>205352_at</td>
<td>SERPIN11</td>
<td>0.64</td>
<td>2.4 × 10^{-46}</td>
</tr>
<tr>
<td>203474_at</td>
<td>IQGAP2</td>
<td>0.62</td>
<td>3.6 × 10^{-14}</td>
</tr>
<tr>
<td>219748_at</td>
<td>TREML2</td>
<td>0.50</td>
<td>5.8 × 10^{-10}</td>
</tr>
<tr>
<td>200771_at</td>
<td>LAMC1</td>
<td>0.48</td>
<td>2.2 × 10^{-9}</td>
</tr>
<tr>
<td>221250_s_at</td>
<td>MXD3</td>
<td>0.49</td>
<td>2.4 × 10^{-9}</td>
</tr>
<tr>
<td>201034_at</td>
<td>ADD3</td>
<td>0.47</td>
<td>5.9 × 10^{-9}</td>
</tr>
<tr>
<td>201752_s_at</td>
<td>ADD3</td>
<td>0.41</td>
<td>9.2 × 10^{-8}</td>
</tr>
<tr>
<td>213745_at</td>
<td>ATRNLI1</td>
<td>0.44</td>
<td>1.1 × 10^{-7}</td>
</tr>
<tr>
<td>217599_s_at</td>
<td>MDFIC</td>
<td>0.42</td>
<td>3.0 × 10^{-7}</td>
</tr>
<tr>
<td>205882_s_at</td>
<td>ADD3</td>
<td>0.38</td>
<td>1.0 × 10^{-6}</td>
</tr>
<tr>
<td>211675_s_at</td>
<td>MDFIC</td>
<td>0.35</td>
<td>1.2 × 10^{-5}</td>
</tr>
<tr>
<td>207821_s_at</td>
<td>PTK2</td>
<td>0.32</td>
<td>1.2 × 10^{-4}</td>
</tr>
<tr>
<td>203305_at</td>
<td>F13A1</td>
<td>0.30</td>
<td>1.9 × 10^{-4}</td>
</tr>
<tr>
<td>201753_s_at</td>
<td>ADD3</td>
<td>0.26</td>
<td>5.5 × 10^{-4}</td>
</tr>
<tr>
<td>218706_s_at</td>
<td>GRAMD3</td>
<td>0.27</td>
<td>7.4 × 10^{-4}</td>
</tr>
<tr>
<td>200872_at</td>
<td>S100A10</td>
<td>0.25</td>
<td>1.5 × 10^{-5}</td>
</tr>
<tr>
<td>210815_s_at</td>
<td>CALCRL</td>
<td>0.24</td>
<td>2.9 × 10^{-3}</td>
</tr>
<tr>
<td>204672_s_at</td>
<td>ANKRD6</td>
<td>0.23</td>
<td>5.6 × 10^{-3}</td>
</tr>
<tr>
<td>208642_s_at</td>
<td>XRCC5</td>
<td>2.16</td>
<td>1.8 × 10^{-189}</td>
</tr>
<tr>
<td>208643_s_at</td>
<td>XRCC5</td>
<td>2.13</td>
<td>1.1 × 10^{-187}</td>
</tr>
<tr>
<td>200792_at</td>
<td>XRCC5</td>
<td>2.05</td>
<td>1.2 × 10^{-160}</td>
</tr>
<tr>
<td>205038_at</td>
<td>IKZF1</td>
<td>0.95</td>
<td>3.4 × 10^{-36}</td>
</tr>
<tr>
<td>205039_s_at</td>
<td>IKZF1</td>
<td>0.50</td>
<td>1.2 × 10^{-11}</td>
</tr>
<tr>
<td>216901_s_at</td>
<td>IKZF1</td>
<td>0.69</td>
<td>2.0 × 10^{-18}</td>
</tr>
</tbody>
</table>
1.3 Ku-dependent expression of Ikaros target genes in primary human lymphocyte precursors from pediatric ALL Patients

Lymphocyte precursors from mice with IKZF1 null mutation are characterized by >2-fold decreased expression levels of 201 transcripts representing 137 IK target genes, including 61 transcripts representing 45 genes harboring IK binding sites as confirmed by chromatin immunoprecipitation followed by DNA sequencing (Uckun et al., 2012). For 19 of these 45 genes, we identified 27 transcripts in the Affymetrix platforms for human transcriptome that showed a highly significant correlation with IKZF1 gene expression levels in primary human lymphocyte precursors from ALL patients consistent with their status as an IK target gene (Uckun et al., 2012). For the expression level of each one of these 19 genes, a striking correlation was also noted with Ku expression levels (Table 1, Figure 3). Hierarchical cluster analysis identified PITTM1 (1.21 SD units, P = 1.2 × 10^{-49}), MDM1 (1.13 SD units, P = 2.2 × 10^{-49}), ITGA4 (1.19 SD units, P = 2.7 × 10^{-49}), KIF23 (1.13 SD units, P = 3.0 × 10^{-46}) and PREP (1.06 SD units, P = 1.5 × 10^{-42}) as the most significantly upregulated genes in 314 patient samples with high Ku expression (Table 1).

These findings extend our earlier study that demonstrated the ability of both Ku70-specific siRNA and Ku80-specific siRNA to knock down the expression of these validated IK target genes in human 293T cells (Ozer et al., 2013).

We also examined the expression levels of 7 signaling pathway genes (viz.: IL4, IL5, IL10, IL13, STAT4, FGFR4, and VIPR1) previously reported to be regulated by native IK in human cells (35-40) for Ku-dependency. Notably, primary lymphocyte precursors with high Ku expression levels (N=67) exhibited significantly higher expression of the IK1 target gene set than lymphocyte precursors with low Ku expression levels (N= 70) (Figure 4). Borderline significant increases were observed with two IK1 targets (IL13, P= 0.06) and (VIPR1, P=0.09) and significant increases were exhibited with the other 7 IK1 targets. Two transcripts for STAT4 and FGFR4 showed highly significant (P<0.0001) overexpression...
in high Ku samples. Cluster representation of the expression values in standard deviation units revealed two sub-clusters of co-regulated genes: FGFR4, IL13 and VIPR1; and 2 transcripts of IL4 co-regulated with STAT4. The most discriminating differentially expressed IK1 target gene was STAT4 (Z-score: 0.926, P=2.7×10^{-6}) (Figure 4, Panel B). Correlation analysis revealed highly significant (P<0.0001) positive relationships between Ku expression and 6 IK1 target transcripts (both FGFR4 transcripts, both IL4 transcripts, STAT4, IL10) with correlation co-efficients ranging from 0.21 to 0.31. IL5 (r=0.19; P=0.0004) and IL13 (r=0.14; p=0.007) exhibited significant but lower correlation coefficients. The observed striking Ku-dependency of the IK target gene expression levels taken together with the results of our earlier RNAi experiments (Ozer et al., 2013) indicate that IK function in human lymphocyte precursors is controlled in part by Ku expression levels.

Figure 4 Effects of Ku Expression Levels on Expression Levels of Ikaros Target Genes in Primary Lymphocyte Precursors from ALL Patients

Note: Gene Pattern (http://www.broadinstitute.org/cancer/software/genepattern/) was used to extract expression values for the 7 IK-regulated signaling genes (35-40) in the combined data set from 4 studies with a total of 354 primary leukemia samples for further analysis. Gene expression values for 3 Ku transcripts and 9 Ikaros targets from 4 published studies examining 354 ALL patient samples (GSE3912, GSE18497, GSE4698 and GSE7440) were rank ordered according to the mean expression of the 3 Ku transcripts. The gene expression profiles of samples with high Ku expression were compared to those of samples with low Ku expression using a 2-sample T-test (Unequal variances). For each study, the Z-score values were calculated as the SD units from the study mean for all the patients. The datasets were combined to test for consistent differences in the Z-scores for high Ku (>0.75 SD units; N=67 samples) and low Ku (<0.75 SD units; N = 70 samples) groups. For each study the expression values were mean centered with respect to the average expression for each gene taken at diagnosis. A one-way agglomerative hierarchical clustering technique was used to organize expression patterns using the average distance linkage method such that genes (rows) having similar expression across patients were grouped together (average distance metric). Dendrograms were drawn to illustrate similar gene-expression profiles from joining pairs of closely related gene expression profiles, whereby genes joined by short branch lengths showed most similarity in expression profile across patients. The heat map represents the color-coded expression value for 67 samples with high Ku expression reported as mean centered expression value relative to log10 transformed expression values for 70 samples with low Ku expression values.
Leukemic T-cell Precursors from T-lineage All Patients Charaterized by Profound Ku80 Deficiency

Figure 5 Expression levels of Ku, Ikaros, and Ikaros target genes in Normal vs. Leukemic Bone Marrow Cells

Note: We compiled the archived “Microarray in Leukemia (MILE)” gene expression profiling data on primary leukemic cells from 174 T-lineage ALL patients and 74 normal bone marrow specimens (GSE 13159). We compared 74 normal bone marrow specimens vs. 174 leukemic (T-lineage ALL) bone marrow specimens for expression levels of 26 IK target genes represented by 40 transcripts on the Affymetrix human genechips, 3 IKZF1 transcripts, 2 Ku80 (XRCC5) transcripts and 1 Ku70 (XRCC6) transcript. Transcript signal values obtained from hybridization onto the Affymetrix Human Genome U133 Plus 2.0 Arrays were calculated using non-central trimmed mean of differences between perfect match and mismatch intensities with quantile normalization (DQN3). Differential were compared using T-tests utilizing the DQN3 values (2-sample, Unequal variance correction, p-values<0.05 deemed significant). Gene expression values were transformed into standard deviation units calculated from the mean and standard deviation expression values for all the samples in each study and effect sizes were reported using differences standard deviation units between comparison groups. We used a one-way agglomerative hierarchical clustering technique to organize expression patterns using the average distance linkage method such that genes (rows) having similar expression across samples were grouped together (average distance metric). Dendrograms were drawn to illustrate similar gene-expression profiles from joining pairs of closely related gene expression profiles, whereby genes joined by short branch lengths showed most similarity in expression profile across samples. The heat map depicts genes up regulated expression values in red and down regulated expression values in green. (C) Expression of 2 probesets for XRCC5/Ku80 (208642_s_at and 208643_s_at) were compared between 74 normal bone marrow specimens and 174
T-precursor ALL samples compiled from the archived “Microarray Innovations in Leukemia” (MILE) study (GSE13159). Expression values for these probesets hybridized onto the Affymetrix Human Genome U133 Plus 2.0 Arrays were calculated using non-central trimmed mean of differences between perfect match and mismatch intensities with quantile normalization (DQN3, signal normalized with quantiles of the beta distribution with parameters $p=1.2$ and $q=3$). To determine the differential expression of each probeset, T-tests were performed between Normal and T-precursor samples (2-sample, Unequal variance correction) exhibiting significant reductions in XRCC5/Ku80 for both probesets in T-precursor ALL cases.

1.4 Low Level Ku Expression is a Unique Hallmark of Childhood T-Precursor Leukemia and It is Associated with Diminished Expression Levels of Ikaros Target Genes

We next examined and compared the gene expression profiles of normal bone marrow cells in 74 healthy bone marrow specimens vs. leukemic bone marrow cells in primary leukemia specimens 174 patients with T-lineage ALL for differential expression of validated IK target genes. Thirty of 40 (75%) transcripts representing 20 of 26 validated (77%) IK target genes as well as all 3 IKZF1 transcripts were expressed at markedly lower levels in leukemic bone marrow specimens from T-lineage ALL patients (Figure 5 A and B). These results are in accord with our original report that children with T-lineage ALL with NCI high-risk criteria express dysfunctional dominant-negative IK isoforms (Sun et al., 1999). Notably, leukemic bone marrow specimens from T-lineage ALL patients also exhibited profoundly reduced expression levels for both of the 2 transcripts encoding the XRCC5/Ku80 component of the heterodimeric DNA repair protein Ku (Figure 5C). The documented association of diminished IK target gene expression in T-lineage ALL with markedly reduced Ku80 transcript levels uniquely indicates that Ku80 deficiency may play a previously unknown causal role in IK malfunction that has been implicated in leukemogenesis of pediatric high risk T-lineage ALL. A role of Ku80 deficiency in T-lineage ALL is also supported by our recent study, which demonstrated that Ku80 haploinsufficiency in mice causes development of an LPD involving CD2+CD4+CD8+IL7R+ thymic T-cell precursors with functional IK deficiency (Ozer et al., 2013).

2 Materials and Methods

2.1 Expression of Recombinant Human Ku in Sf21 Insect Ovary Cells

The 2.0-kb ku70 and 2.2-kb ku80 cDNA fragments including their protein coding segments (Gell and Jackson, 1999) were individually cloned into the Ncol/KpnI site of the 4.9-kb pFastBacHT (PFBH) donor vector (Life Technologies) containing a 6×-histidine (6×His) tag to construct recombinant PFBH-ku70 and PFBH-ku80 plasmids. PFBH-ku70 and PFBH-ku80 were used to generate recombinant baculoviruses by site-specific transposition in Escherichia (E.) coli DH10Bac competent cells (Life Technologies), which harbor a baculovirus shuttle vector (bmacid), bMON14272 with a mini-attTn7 target site for site-specific transposition using previously reported procedures (Uckun et al., 2010, 2011; Mahajan et al., 2001). The bacterial colonies containing recombinant bacmids were identified by disruption of the lacZa gene. High molecular weight miniprep DNA was prepared from selected E.coli clones containing recombinant bacmid and transfected into Sf21 cells using the Cellfectin reagent (Life Technologies) as previously described (Ozer et al., 2013; Uckun et al., 2010, 2011; Mahajan et al., 2001). Sf21 cells were infected with both recombinant baculoviruses to produce the Ku70/Ku80 heterodimer (Ozer et al., 2013). Cultures were incubated at 28°C and stirred at 80–100 rpm using a magnetic stirrer (Belleco Glass, Inc.,Vineland, NJ) for 48 h. Infected cells were harvested by gentle centrifugation in a Beckman GS-6 centrifuge at 500 × g for 7 min at room temperature. Cells from 1-liter cultures were flash-frozen at -80°C and stored at -80°C until purification of recombinant Ku70 and Ku80 proteins. Frozen Sf21 insect cells were lysed in 1× Triton X-100 extraction buffer (1% Triton X-100, 10 mM Tris, 130 mM NaCl, 10 mM NaF, 10 mM sodium phosphate, pH 7.5) (1 mL lysis buffer per 20 × 106 cells). One pellet of CompleteTM protease inhibitors (Roche Molecular Biochemicals)was added for each 25 mL of the lysate and the mixture was rotated for 2 h at 4°C. The cell pellets were centrifuged at 45 000 rpm × 1 h a Beckman Optima LE-80K ultracentrifuge.
using a 45 Ti rotor. Following centrifugation, the clarified supernatant was filtered through a 0.22 μm membrane filter (S100) and dialyzed for 4 h into buffer A (20 mM sodium phosphate, 10% glycerol, pH 7.2). For purification of Ku70 and Ku80 proteins as well as Ku70/Ku80 heterodimer, dialyzed supernatants were applied to a Nickel-chelation column (Pharmacia) that was equilibrated with buffer B (20 mM sodium phosphate, 500 mM NaCl, 0.5 M imidazole, 10% glycerol, pH 7.2). Ku70 or Ku80 enriched fractions were dialyzed against buffer C (20 mM Tris, pH 8.0 with 1 M DTT, 10% glycerol) overnight to remove imidazole and then applied to a Sepharose Q HP26/10 ion exchange column (column volume 50 mL; Amersham Pharmacia Biotech) followed by size exclusion chromatography on a Superdex 200 HR 10/30 column (Pharmacia) for further purification (Ozer et al., 2013).

2.2 Monitoring of Binding Interactions Using Surface Plasmon Resonance (SPR) Technology

A BIAcore X surface plasmon resonance-based biosensor system (Amersham Pharmacia Biotech Biosensor AB) was used to measure the kinetic parameters for the interactions between soluble recombinant MBP-IK1 vs. MBP-IK5 fusion proteins (=analytes) and the immobilized 6xHis-tagged recombinant Ku (Ku70/80 heterodimer), Ku70, Ku80 proteins captured via metal chelation (= ligands), as previously described (Ozer et al., 2013; Mahajan et al., 2001; Rajamohan et al., 2001). The nitrilotriacetic acid (NTA) sensor chip was saturated by injection 20 μl of 500 mM NiCl2 solution at a flow rate of 20 μL/min in NTA buffer. Ligands (200 nM in NTA buffer) were immobilized on the sensor chip by injecting 25 μL of the ligand solution (20 μL/min). A 50-μL sample of either MBP-IK1 or MBP-IK5 fusion protein (100 nM) was injected at 25°C at a flow rate of 20 μL/min onto the sensor chip surface on which a specific ligand had been immobilized. Between samples, the binding surface was regenerated by injection of 40 μL of EDTA-containing regeneration buffer (10 mM HEPES, 0.15 M NaCl, 0.35 M EDTA, 0.005 Surfactant P20, pH: 8.3) at a flow rate 20 μL/min to remove metal ions. The primary data were analyzed and association rate constant (kon)/dissociation rate constant (koff) values were determined using the BIAevaluation software (Version 3.0) supplied with the instrument (Biacore, Inc.), as previously described (Mahajan et al., 2001; Rajamohan et al., 2001). The dissociation constant KD = koff/kon and association constant KA = kon/koff were also calculated.

2.3 Molecular Model of Ikaros-Ku Complex

We used the known cDNA sequence of IK1 and the already published crystal structure of the Ku70/80 heterodimer bound to DNA in order to construct a molecular model to explore possible modes of interaction between IK1 and Ku proteins. The IK1 molecule from residue 110 to residue 256 was modeled based on the homology with two zinc finger proteins (PDB access codes: 1AAY, 1G2D) using InsightII (Molecular Simulation Inc. San Diego, CA) and MOE (Chemical Computing Group Inc., Montreal, Canada), as previously described (Brady and Stouten, 2000; Vig et al., 1998). The interaction surface analyses were done with GRASP (Nicholls et al., 1991). We first explored the interaction of the Ku70/Ku80 heterodimer bound to DNA (PDB access code: 1JEY) and the IK1 zinc-finger (ZF) domains using our IK1 homology model. The initial positions of both Ku heterodimer and IK1 ZF domains were placed manually. The DNA major groove contour and base pair positions were used as a guide to properly position IK1 ZF domains with respect to Ku. The initial 55-nucleotide DNA element from the crystal structure of the Ku heterodimer was extended in a standard B conformation and fixed in the subsequent docking calculation. Major steric collisions with Ku were removed by manually adjusting the torsion angles of the residues on the interface. We created a definitive binding set of Ku residues in the binding pocket to move as a 4.0 Å shell around the manually docked IK1 ZF domains. This general position of the Ku, IK1 and DNA were used as a starting point for subsequent automatic docking trials and energy minimization. The docking calculation was repeated several times for several slightly different initial positions of the complex towards each other to ensure the stability of the end result and avoid local minimum. The calculations used a consistent valence force field
in the Discovery program and a Monte Carlo strategy in the Affinity program. Each energy-minimized final docking position of the complex was evaluated using the interactive score function in the LUDI module, visual inspection and interface analyses including contact surface area, steric tension, improper rotamer positions by GRASP, CHAIN and O. The final binding position of the complex was determined based on the evaluation of favorable binding interactions using the LUDI score function. The parameters used in this docking strategy included searching for five unique structures, 1 000 minimization steps for each structure, energy range of 10.0 kcal/mol, maximum translation of the ligand of 3.0 Å, maximum rotation of the ligand of 10°, and an energy tolerance of 1500 kcal/mol.

2.4 Bioinformatics and Statistical Analysis of Gene Expression Profiles

The publicly available archived GSE32311 database was used to compare gene expression changes in CD4+CD8+ double-positive wild type (N=3; GSM800500, GSM800501, GSM800502) vs. IKZF1 null mouse thymocytes (N=3, GSM800503, GSM800504, GSM800505) from the same genetic background of (C57BL/6 x 129S4/SvJae). Probe level RMA signal intensity values were obtained from the mouse 430_2.0 Genome Array. Up-regulated and down-regulated transcripts in IKZF1 knockout mice were identified by filtering changes greater than 2-fold and T-test P-values less than 0.05 (T-test, Unequal Variances, Excel formula). We identified 1 158 transcripts representing 924 genes that were down-regulated in IKZF1 null mice with a subset of 201 transcripts representing 137 genes exhibiting >2-fold decreased expression levels. By cross-referencing this IK-regulated gene set with the archived ChIPseq data (GSM803110) using the Integrative Genomics Browser (Robinson et al., 2011), we further identified 45 Ikaros target genes that harbored IK binding sites (Uckun et al., 2012). The Gene Pattern web based software (http://www.broadinstitute.org/cancer/software/genepattern) was used to extract expression values from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database to compile gene expression profiles of human lymphocyte precursors in 1 104 primary leukemia specimens from pediatric ALL patients (GSE3912, N=113; GSE18497, N=82; GSE4698, N=60; GSE7440, N=99; GSE13159, N=750). We focused our analysis on 45 validated IK target genes (Uckun et al., 2012). Expression values expressed as Standard Deviation units were compiled for the 5 studies and rank ordered according to the mean expression of three highly correlated transcripts (208642_s_at (XRCC5), 208643_s_at (XRCC5), 200792_at (XRCC6). Prospective power analysis was utilized to determine the Standard Deviation cut-off for “high Ku expression” and “low Ku expression” in the data sets. To control for False Positive Rate (FPR) to detect for differences in 3 Ku transcripts out of approximately 20 000 transcripts common across the 5 Affymetrix platforms, we set the unadjusted P-value at 2.5×10⁻⁶ (FPR = 0.05). Sample size greater than 132 would provide sufficient to detect a difference of 1 standard deviation units with 99.9% power. Therefore, samples were assigned to the “high Ku expression” group if their expression level was >0.5 standard deviations units higher than the mean expression level (N=314) and to the “low Ku expression” group if their expression level was >0.5 standard deviations units lower than the mean expression level (N=324). These samples were also rank ordered according to IKZF1 expression level (205038_at, 205039_s_at, 216901_s_at, 227344_at and 227346_at; 3 of these were common in all Affymetrix platforms - 205038_at, 205039_s_at, 216901_s_at) resulting in 302 ALL samples with high IKZF1 expression and 318 samples with low IKZF1 expression. T-tests were performed for the combined Standard Deviation units from the 5 datasets (2-sample, Unequal variance correction, p-values<0.05 deemed significant) and revealed 27 transcripts representing 19 IK target genes (Table 1) and 13 transcripts representing 12 lymphoid-priming genes (Uckun et al., 2012; Ma et al., 2013) that were significantly up-regulated in samples with both high Ku and high IKZF1 expression levels. We used a one-way agglomerative hierarchical clustering technique to organize expression patterns using the average distance linkage method such that genes (rows) having similar expression across patients were grouped together (average distance metric). Dendrograms were drawn to illustrate similar gene-expression profiles from joining pairs of closely
related gene expression profiles, whereby genes joined by short branch lengths showed most similarity in expression profile across samples (Uckun et al., 2012; Ma et al., 2013).

In separate analyses of gene expression profiles of lymphocyte precursors with high vs low Ku expression levels, we compiled the gene expression profiles of 354 primary leukemia specimens from newly diagnosed or relapsed ALL patients (GSE3912, N=113; GSE18497, N=82; GSE4698, N=60; GSE7440, N=99). We focused our analysis on the genes for the regulatory cytokines/peptides (FGFR4, IL10, IL13, IL4, IL5, STAT4 and VIPR1) comparing samples with high versus low Ku expression. These genes were identified as IK-regulated genes in human cells (Yu et al., 2002; Umetsu and Winandy, 2009; Mary et al., 2009; Gregory et al., 2006; Yap et al., 2005; Dorsam and Goetzl, 2002). For each of the 4 studies, the expression values of the IK target gene transcripts common in across the Affymetrix platforms (204579_at (FGFR4), 211237_s_at (FGFR4), 207433_at (IL10), 207844_at (IL13), 207538_at (IL4), 207539_s_at (IL4), 207952_at (IL5), 206118_at (STAT4), and 205019_s_at (VIPR1)) were standardized relative to mean expression value across all samples in that study (Z-scores in standard deviation units). Expression values expressed as Z-scores were compiled for the 4 studies and rank ordered according to the mean expression of three highly correlated transcripts (208642_s_at (XRCC5), 208643_s_at (XRCC5), 200792_at (XRCC6)). Prospective power analysis was utilized to determine Z-score cut-off for “high” Ku expression and “low” Ku expression in the data sets. To control for False Positive Rate (FPR) to detect for differences in 3 Ku transcripts out of approximately 20 000 transcripts common across the 4 Affymetrix transcripts, we set the unadjusted P-value at $1 \times 10^{-5}$ (FPR = 0.07). Sample size greater than 102 would provide sufficient power to detect a difference of 1.4 standard deviation units at the 99% level. Therefore, samples were defined as “high” Ku expression if their expression was greater than 0.7 standard deviations units and “low” Ku expression if the sample exhibited expression level lower than 0.7 standard deviation units and intermediate expression for the remaining samples (N=217). To determine the differential expression of each gene transcript, T-tests were performed for the combined Z-scores from the 4 datasets (2-sample, Unequal variance correction, p-values<0.05 deemed significant) to compare the expression of high Ku expression (N=67) versus low Ku expression (N=70) samples for the IK targets. We used a one-way agglomerative hierarchical clustering technique to organize expression patterns using the average distance linkage method such that genes (rows) having similar expression across patients were grouped together (average distance metric). Dendrograms were drawn to illustrate similar gene-expression profiles from joining pairs of closely related gene expression profiles, whereby genes joined by short branch lengths showed most similarity in expression profile across samples. Pairwise correlations (r; JMP Software (SAS, Cary, NC)) of the average expression of the 3 Ku transcripts versus each of the 9 transcripts of IK target genes were performed using the 354 samples compiled from the 4 studies. The correlation co-efficient and un-adjusted P-values (P<0.05 deemed significant) were reported to identify IK target genes driven by Ku expression.

Acknowledgments

The project described was supported in part by DHHS grants P30CA014089, U01-CA-151837, and R01CA-154471 (to FMU). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health. This work was also supported in part by Children’s Hospital Los Angeles Institutional Endowment and Special Funds (FMU), 2011 V-Foundation Translational Research Award (FMU), Ronald McDonald House Charities of Southern California (FMU), Couples Against Leukemia Foundation (FMU). KU70 and Ku80 cDNA were kindly provided by Dr. Stephen P. Jackson (Wellcome Trust/Cancer Research UK Gurdon Institute).

Authors’ Contributions

F.M.U conceived and supervised this study and wrote the final manuscript. All authors have contributed to the design and conduct of the research. All authors reviewed the paper. The authors have declared no
competing interests.

References


Huang N., Lee I., Marcotte E.M., and Hurles M.E., 2010, Characterizing and predicting haploinsufficiency in the genome, PLOS Genetics http://dx.doi.org/10.1371/journal.pgen.1001154


Ma H., Quzi S., Ozer Z., Zhang J., Ishkhanian R., and Uckun F.M., 2013, Regulatory Phosphorylation of Ikaros by BTK, PLOS One, 8(8): e71302 http://dx.doi.org/10.1371/journal.pone.0071302 PMid:23977012 PMCid:PMC3747153


Mahajan S., Vassilev A., Sun N., Ozer Z., Mao C., and Uckun FM., 2001, Transcription factor STAT5A is a substrate of Bruton’s tyrosine kinase in B cells, J. Biol. Chem., 276: 31216-31228 http://dx.doi.org/10.1074/jbc.M104874200 PMid:11413148


Ozer Z., Qazi S., ishkhanian R., Hasty P., Ma H., and Uckun F.M., 2013, Ku Haploinsufficiency Causes a Lymphoproliferative Disorder of Immature T-cell Precursors Due to Ikaros Malfunction, International Journal of Molecular Medical Science, 3: 50-64


Rajamohan F., Ozer Z., Mao C., Uckun F.M., 2001, Active center cleft residues of pokeweed antiviral protein mediate its high-affinity binding to the ribosomal protein L3, Biochemistry, 40(31): 9104-9114 http://dx.doi.org/10.1021/bi002851p PMid:11478877


Umetsu S.E., and Winandy S., 2009, Ikaros is a regulator of II10 expression in CD4⁺ T cells, J. Immunol., 183(9): 5518-5525
http://dx.doi.org/10.4049/jimmunol.0901284
PMid:19828627 PMCid:PMC2778601

http://dx.doi.org/10.1016/S0968-0896(98)00108-4

http://dx.doi.org/10.1038/35088000 PMid:11493912

http://dx.doi.org/10.1016/j.febslet.2005.07.018
PMid:16081070

http://dx.doi.org/10.1016/j.coi.2010.02.011
PMid:20299195

http://dx.doi.org/10.1210/me.16.5.1069 PMid:11981041

http://dx.doi.org/10.1038/ni.2150 PMid:22080921