Published in final edited form as:
*Gene*. 2013 January 1; 512(1): 134–142. doi:10.1016/j.gene.2012.09.068.

# Two Single Nucleotide Polymorphisms in the Human Nescient Helix Loop Helix 2 (*NHLH2*) Gene Reduce mRNA Stability and DNA Binding

**Numan AL_Rayyan**, **Umesh Wankhade**, **Korie Bush**, and **Deborah J. Good***
Department of Human Nutrition, Foods and Exercise, Virginia Tech University, Blacksburg, VA 24061

## Abstract

Nescient Helix Loop Helix-2 (NHLH2) is a basic helix-loop-helix transcription factor, which has been implicated, using mouse knockouts, in adult body weight regulation and fertility. A scan of the known single nucleotide polymorphisms (SNPs) in the *NHLH2* gene revealed one in the 3' untranslated region (3'UTR), which lies within an AUUUA RNA stability motif. A second SNP is nonsynonymous within the coding region of *NHLH2*, and was found in a genome-wide association study for obesity. Both of these SNPs were examined for their effect on NLHL2 by creating mouse mimics and examining mRNA stability, and protein function in mouse hypothalamic cell lines. The 3'UTR SNP causes increased instability and, when the SNP-containing *Nhlh2* 3'UTR is attached to luciferase mRNA, reduced protein levels in cells. The nonsynonymous SNP at position 83 in the protein changes an alanine residue, conserved in *NHLH2* orthologs through to *Drosphilia sp.* to a proline residue. This change affects migration of the protein on an SDS-PAGE gel, and appears to alter secondary structure of the protein, as predicted using *in silico* methods. These results provide functional information on two rare human SNPs in the *NHLH2* gene. One of these has been linked to human obese phenotypes, while the other is present in a relatively high proportion of individuals. Given their effects on NHLH2 protein levels, both SNPs deserve further analysis in whether they are causative and/or additive for human body weight and fertility phenotypes.

### Keywords

Nhlh2; NHLH2; NSCL2; mRNA stability; 3'Untranslated Region; SNP; nonsynonymous SNP; DNA binding; obesity; fertility

## 1. Introduction

Nescient helix-loop-helix 2 (Nhlh2, mouse; NHLH2, human) is a member of the large family of basic helix-loop-helix (bHLH) transcription factors (Atchley and Fitch, 1997). The bHLH transcription factors bind DNA through their basic domain at an E-box sequence, denoted as CANNTG, and interact with other transcription factors by forming hetero- and



*To whom correspondence should be addressed: Deborah J. Good, Ph.D., Department of Human Nutrition, Foods and Exercise, 1981 Kraft Drive MC0913, Virginia Tech, Blacksburg, VA, 24061 USA; Telephone: 540-231-0430; Fax: 540-231-5522; goodd@vt.edu.

homodimers through their HLH domains. Targeted deletion of *Nhlh2* in mice results in adult onset obesity accompanied by a phenotype of low physical activity (Coyle et al., 2002), as well as reduced fertility (Good et al., 1997; Johnson et al., 2004), but the role of *NHLH2* in humans has not been examined. One way to assess whether *NHLH2* plays a role in human body weight control is to examine single nucleotide polymorphism(s) in the gene for possible effects on the protein or RNA levels, and to then eventually use population analysis to determine if these SNPs contribute to obesity or related morbidities.

In the dbSNP database (http://www.ncbi.nlm.nih.gov.ezproxy.lib.vt.edu:8080/snp/), only one SNP ( rs144106335) is located in the protein coding region, but this is a synonymous SNP so there is not expected effect on the NHLH2 protein. However, in a moderately sized study of 379 obese and 379 lean individuals, a nonsynonymous mutation in human *NHLH2* was found in 2 obese and 1 lean individuals (Ahituv et al., 2007). Subsequent to that analysis, one of the authors reported that this mutation was only found in obese individuals, and that the frequency of the mutation was 0.001 (0.1 percent) in obese individuals (Goren et al., 2008). However, no studies have been done to further characterize this human *NHLH2* SNP.

While there are multiple SNPs listed in the dbSNP database within non-coding regions of the *NHLH2* and other mRNAs, there are still critical gaps in linking SNPs found in regulatory regions to effects on body weight and other phenotypes. While not as obvious in functional outcome as amino acid substitutions, there is precedent for regulatory differences to vastly change an organism in the absence of amino acid sequence differences. The best example comes from the differences between the human and chimpanzee genomes where there is high divergence in the 5' and 3' untranslated regions (UTRs) of genes, compared to the protein coding regions which potentially contribute to expressional differences between the two species (Kehrer-Sawatzki and Cooper, 2007). In humans, promoter, 5' and 3' UTR SNPs have been linked to obesity or associated phenotypes. For example, a minor allele SNP rs16861194, in the adiponectin promoter was shown to affect in vitro luciferase activity, as well as reduced serum adiponectin levels in carriers (Laumen et al., 2009). A SNP in the 5' UTR of the *IκBα* gene is linked to reduced insulin sensitivity (Miller et al., 2010), however, no in vitro testing of the SNP to identify the mechanism or effect on the *IκBα* gene was done. It is possible that SNPs within a promoter or other regulatory region could be non-causative themselves, but rather linked to another causative mutation. Thus, identification of a possible regulatory SNP must be accompanied by in vitro and/or in vivo experiments to test the regulatory consequences of the polymorphism.

In this study, we characterized one SNP in the 3'UTR of *NHLH2*, which was in close proximity to AUUUA elements, and one nonsynonymous SNP in the coding region of *NHLH2*, which has been linked to obesity in humans. Both SNPs were examined for their effect on the RNA or protein using both *in vitro* and *in silico* methods.

## 2. MATERIALS AND METHODS

### 2.1 Generation of constructs

To analyze the SNP in the coding region of *NHLH2*, an expression construct containing a mouse Nhlh2-myc fusion protein (a generous gift from Dr. Thomas Braun, Max Planck Institute, Bad Nauheim, Germany) was used to create a site-directed mutation resulting in a change of an alanine residue to a proline residue at position 83. Specifically, iProof High fidelity DNA polymerase (BioRad) and a standard *in vitro* mutagenesis assay was used with two primers designed to be homologous to the surrounding DNA sequence, except a change of a guanine residue to a cytosine residue in the exact middle of the primer. The primers are listed in Table 1. The plasmid was sequenced to confirm mutagenesis.

To analyze the SNP in the 3'UTR region of *NHLH2*, two constructs were made using mouse *Nhlh2* expression plasmids previously created or obtained by our lab. For the first construct, a plasmid containing a mouse Nhlh2-myc fusion protein (described above) was used so that the transfected WT mRNA and mutant mRNA would be tagged in cells for QPCR analysis. The mouse 3'UTR was cloned into the construct in place of the SV-40 polyA in the original plasmid. To do this, a 1.2 kb fragment of *Nhlh2* 3'UTR was excised from a vector containing mouse genomic *Nhlh2* sequence using SacI and NotI enzymes. The 3'UTR of *Nhlh2* was cloned into PCS2-MT vector, which already contained the myc tag attached to mouse *Nhlh2* coding sequence. This plasmid was cut with SacI and NotI enzymes to open an insertion sites for the 3'UTR. The 3'UTR was inserted directly after the *Nhlh2* coding sequence. The adenosine residue at position 1563 in the WT mouse *Nhlh2* 3'UTR was replaced with a guanine residue using iProof High fidelity DNA polymerase (BioRad) and a standard *in vitro* mutagenesis assay and primers designed for this purpose (Table 1). The plasmids were sequenced to confirm mutagenesis and correct cloning of all parts. The second construct contained just the *Nhlh2* 3'UTR linked to the luciferase gene. To create this plasmid, a 1338 bp fragment of *Nhlh2* 3'UTR was amplified from mouse *Nhlh2* sequence using Hot Master Taq DNA Polymerase (5-Prime) and cloned into pBluescript. The 3'UTR was then digested by ApaI and SmaI enzymes and cloned into pcDNA3 vector features is driven by the CMV promoter and containing the luciferase reporter gene. The pcDNA3 was cut with ApaI and SmaI to remove the SV-40 poly A tail, and to insert the *Nhlh2* 3'UTR directly downstream of the luciferase coding sequence. The plasmid was subjected to restriction enzyme digestion followed by agarose gel electrophoresis to assess the incidence of insertion followed by sequencing to verify the insertion. This construct was then used as a template to create another plasmid containing the induced mutation at nucleotide position 1563, converting the WT adenosine residue to a guanine residue, mimicking the change found in the Human SNP. All constructs were sequenced to confirm the mutagenesis had occurred correctly.

### 2.2 *In silico* Analysis of SNPs in *NHLH2*

The protein, DNA and RNA sequences for *NHLH2* were analyzed and compared cross-species using BioEdit Sequence Alignment editor for windows, version 7.1.3 (http://www.mbio.ncsu.edu/bioedit/bioedit.html). Accession numbers for each sequence are found in the figure legend. For protein structural analysis, the NHLH2 wildtype and mutant protein sequence, protein sequences were entered into the SAM-T08 protein structure prediction online server at http://compbio.soe.ucsc.edu/SAM_T08/T08-query.html (Katzman et al., 2008). The notor angle prediction program was used in the analysis. The notor alphabet use hydrogen bond properties to predict the torsion angle between the peptide planes (Katzman et al., 2008). The web based server also provides secondary structure prediction, which is presented.

For RNA structural analysis, initially an 800 bp sequence (the limit for the server) was entered, containing nucleotides 1201 through 2001. While there is a slightly different free energy value, the structures appear relatively similar (Figure 3C). The 33 bp RNA sequence shown in Figure 2A from the *NHLH2* wildtype and mutant RNA were entered into the MFold web server with the default settings (Zuker, 2003). The individual structures in graphical output were copied directly from the web server output for both sequences.

### 2.3 Actinomycin D Analysis of mRNA Stability

An Actinomycin D assay was used to measure mRNA stability. The assay was performed on the hypothalamic N29/2 cell line (Cellutions Biosystems, Toronto, Ontario, Canada) (Belsham et al., 2004). The cells were maintained in DMEM-high glucose (4.5 g/liter) medium containing 4.5 g/liter sodium pyruvate and 10% fetal bovine serum, 100 units/ml

penicillin, and 10 μg/ml streptomycin (HyClone, Logan, UT) at 37°C in 5% CO2. Cells were transfected with a total amount of 400 ng of DNA at 40–60 % confluence in 12-well plates using Effectene transfection reagent (Qiagen, Valencia, CA) following the manufacturers protocol. Twenty-four hours following transfection, cells were treated with media containing 5 μg/ ml Actinomycin D (Sigma, Saint Louis, Missouri) for 30, 60, 120 and 240 minutes. Total RNA was extracted using the Trizol™ (Invitrogen, Grand Island, NY) extraction protocol. The RNA used to synthesize cDNA, using standard methods. The *Nhlh2-myc* mRNA levels were measured using the SYBRGreen PCR master mix (Applied Biosystems, Foster city, California), using primers (Table 1) which target the sequence flanking both of myc tag and the *Nhlh2* coding region to differentiate between the endogenous *Nhlh2* which was already expressed in N29/2 cell line and the *Nhlh2* expressed from the transfected DNA. The *Nhlh2* mRNA expression levels were normalized to the level of 18S rRNA reference gene (primers in Table 1). Real-Time PCR experiments were performed on a 7900 HT Fast Real-Time PCR System (Applied Biosystems, Foster City, California).

### 2.4 Luciferase Assays for mRNA Stability Analysis

Transient transfection of N29/2 cells was performed in 12-well tissue culture plates using Effectene transfection reagent (Qiagen, Valencia, CA) following the manufacturer instructions. Cells were seeded into the wells about 24 hours prior to transfection. Each well was seeded with 1.2 ml of media with 1–2 × $10^5$ cells per ml in order to get 40–60 % confluency after 24 hours. Twenty-four hours following cells seeding, each well received 400 ng of luciferase reporter plasmid (WT) or the mutant plasmid. All wells received 35 ng of β-galactosidase (β-gal) expression plasmid as an internal control plasmid. Cells were washed with PBS 24 hours after transfection and lysate were collected for luciferase and β-gal assay. To assay luciferase activity, 5 μl of lysis was placed in the wells of a white 96-well plate and 100 μl luciferase assay buffer (Promega, Madison, WI) was injected sequentially; relative light units were measured by (GLOMAX Multi Detection System) from (Promega, Madison, WI). The β-gal activity was taken to be the rate of increase of absorbance at 450 nm. In order to measure the relative luciferase activity, the luciferase activity of each sample was divided by the corresponding relative β-gal activity. Three independent transfection experiments were performed to confirm the results. In each experiment, triplicate wells were transfected, and luciferase and β-gal activities for each well were determined in triplicates. Luciferase is a long-lived protein and it has been used to study ARE- and miRNA mediated inhibition of translation (Mukhopadhyay et al., 2003). The effect of *Nhlh2* 3'UTR SNP on the luciferase activity was compared with that of the WT construct using paired t tests.

### 2.5 Western Analysis

For $Nhlh2^{WT}$-myc and $Nhlh2^{83P}$ protein analysis, constructs were transfected into N29/2 cells, and 24 hours following transfection, cells were scraped from the tissue culture plates, washed, homogenized in RIPA buffer and processed for Western analysis using standard methods. Equal amounts of protein (40 μg/lane), as determined using BCA Protein Assay (Pierce, Thermo Scientific, Rockford, IL) were separated on a 12% SDS polyacrylamide gel and transferred to nitrocellulose membrane. Western blotting was performed using myc-tag mouse monoclonal primary antibody (9B11; Cell Signaling Technology, Danvers, MA) with rabbit anti-mouse horse-radish peroxidase-linked antibody as a secondary antibody (Abcam, Cambridge, MA). Chemoluminescent signal was detected using the ECL kit (Pierce, Rockford, IL).

### 2.6 Statistical Analysis

For the transfection analysis of the *NHLH2*$^{1563G}$ 3'UTR SNP, all values are expressed as mean ± SEM. Comparison of means between two groups was made using unpaired two-tailed Student's T-test (luciferase assays; Microsoft Excel®, 2008 for Mac), or slope analysis (actinomycin D assays; SAS,). *P* values were calculated using statistical analysis function in Microsoft Excel® (2008 for Mac version 9.2). Significance is expressed at *p≤0.05; **p≤0.01.

## 3. Results

### 3.1 Phylogenetic Analysis of *NHLH2* and Nonsynonymous SNP

A rare SNP in the coding region of *NHLH2* was found with a 0.1% frequency in obese individuals and 0% frequency in lean individuals (Ahituv et al., 2007; Goren et al., 2007). No further analysis of this SNP has been described, nor is this SNP listed in the NCBI database at this time. In addition, there are no other known non-synonymous SNPs in the *NHLH2* coding region. Therefore, we analyzed used phylogenetic analysis to determine if the alanine at position 83 in the NHLH2 protein was conserved across species. As shown in Figure 1A, The nonsynonymous polymorphism (NHLH2$^{83P}$) changes a species-conserved alanine to a proline within the basic DNA binding domain of the bHLH domain of NHLH2 Alanine is present at this position in mammalian and avian classes, as well as the cyprinid (zebrafish) and Tetraodontidae (pufferfish) fish families and in insects (Drosophila). In the human NHLH2$^{83P}$ sequence, the SNP is the result of a guanine to cytosine single nucleotide change as position 247 (counting from ATG, accession # NM_178777) (Figure 1B). This guanine residue is conserved in mouse, while two other polymorphisms in the mouse, relative to the human *NHLH2* sequence do not affect the sequence of the resultant protein. The induced mutation in the mouse Nhlh2 protein was made to mimic the human SNP (Figure 1B) and constructs containing the mutation were used in further studies.

### 3.2 Comparative Analysis of SNPs in the *NHLH2* 3' UTR

We examined all of the known SNPs in the human *NHLH2* gene for possible regulatory effects on *NHLH2* expression. There are a total of 10 SNPs within the proximal promoter of *NHLH2*, 8 SNPs in intronic sequences, and 7 SNPs in either the 3' or 5' untranslated region (UTR) of the mRNA (data not shown, but available on NCBI SNP database). One of the SNPs in the 3' UTR was of interest to us because of its location relative to a putative AUrich element (ARE)-binding protein site (for mRNA stability) (Figure 2A). In a 33 base pair region containing the putative ARE, there are only 5 nucleotide differences (15%) between the mouse and human *NHLH2* genes. The SNP (rs11805084, *NHLH2*$^{A1568G}$, also noted in a separate sequence as being at position 1557) changes a conserved adenosine (A) residue to a guanine (G) residue, which is conserved in both position and sequence in the UTR for the mouse and human *NHLH2* genes (Figure 2B). According to the NCBI SNP database, this SNP has a frequency of 7.1% in a 1000 Genome phase 1 population with 'G' being observed 155 times in the sample population of 629 people (or 1258 chromosomes). The induced mutation in mouse *Nhlh2* (*Nhlh2*$^{1563G}$) was made to mimic the human SNP (Figure 2B), and constructs containing the induced mutation were used in further studies.

### 3.3 Loss of the predicted helical structure in the DNA binding domain of the NHLH2$^{83P}$ protein

Online protein secondary structure analysis was used to determine if the NHLH2$^{83P}$ protein maintained the same predicted secondary structure, compared to the wildtype NHLH2 protein. As shown in Figure 3A, in the mutant protein, a proline at position 83 disrupts an 8 amino acid helical structure within the basic DNA binding domain that precedes the

helixloop- helix functional domain of the protein. The missing helical structure can be seen more clearly in a secondary structure image (Figure 3B, double black arrow). In addition, there appears to be a slight effect in an arginine-rich region, between amino acids 68–70, resulting in a new helical structure in this area (Figure 3B, double grey arrow).

### 3.4 Effects on RNA secondary structure in the *NHLH2*$^{1563G}$ transcript

As shown in Figure 3C, the A/G transition at position 1568 in the *NHLH2* mRNA results in a change in the predicted secondary structure of the RNA within that region. The A/G transition is predicted to result in an additional loop region between nucleotides 1568 and 1577. This appears to cause the AUUUA motif in the RNA to be completely contained within a loop structure in the mutant, compared to just a 3-base pair look in the wildtype. Changes in the structure of AREs can affect RNA binding protein recognition sites (Bolognani and Perrone-Bizzozero, 2008).

### 3.5 The NHLH2 nonsynonymous SNP leads to a defective protein product

Constructs containing Nhlh2$^{WT}$ and Nhlh2$^{83P}$ were used to analyze the effect of the point mutation on Nhlh2 protein. Each construct had a myc-tag linked to the mouse *Nhlh2* gene, which was used to follow the transfected protein within hypothalamic cells (Figure 4A). Transfection of the two constructs into N29/2 hypothalamic cells, and analysis of whole cell extract by western blot revealed a difference in the migration pattern for the mutant Nhlh2$^{83P}$ protein (Figure 4B). In particular, the mutant protein had two bands, one migrating slower than the WT protein (at 19 kDa), and the other appearing to be a breakdown product that still contained the myc-tag moiety.

### 3.6 The *NHLH2* 3'UTR SNP destabilizes the Nhlh2-myc SNP containing mRNA

As the 3'UTR SNP appeared to affect a putative ARE-binding protein site, differences in mRNA stability were measured using several methods. Constructs containing the *Nhlh2* coding region with the wild type (control) 3'UTR for the mouse *Nhlh2* gene, and one containing an induced point mutation at position 1563 were created to measure mRNA levels (Figure 5A, top panel). These constructs included a myc-tag sequence, which is incorporated into the *Nhlh2* mRNA, which allowed for detection of the transfected mRNA. An mRNA decay assay was implemented by transfecting a hypothalamic cell line N29/2 with the control or the induced mutant constructs. Transcription was inhibited using actinomycin D, allowing for RNA harvesting at 30, 60, 120 and 240 minutes after actinomycin D treatment to determine half-life of the mRNA species. At each time point, *Nhlh2* -myc mRNA was measured by real time quantitative PCR (Q-PCR) using a primer set to detect only the myc-tagged mRNA. The presence of SNP in the *Nhlh2* 3'UTR resulted in a initial reduction in *Nhlh2* expression, accompanied by an overall significant decrease in mRNA levels ($P < 0.01$), compared to the control construct (Figure 5B).

### 3.7 The *NHLH2* 3'UTR SNP decreases reporter protein levels

To study the importance of this SNP in regulating Nhlh2 protein levels, a vector with the 3'UTR of *Nhlh2*, plus or minus the SNP, was cloned downstream of a luciferase reporter gene (Figure 5A, bottom panel). In addition, a control the SV-40 Poly A region, which should confer strong stability, was also used. The luciferase protein is a highly stable protein, which is why it is usually used as a reporter in cell-based assays. Each of the vector constructs contained the strong CMV promoter, so that both the SNP containing and control vectors would have high endogenous mRNA expression levels. N29/2 cells were transfected with these constructs and cell extracts were made 24 hours later for luciferase assays. The luciferase activity of the Luciferase-pcDNA3 plasmid which has the CMV promoter and the SV40 poly (A) signal gave relatively similar expression compared to the plasmid which had

wild-type *Nhlh2* 3'UTR instead of SV40 poly A (Figure 5C). These results suggested that substitution of SV40 with the *Nhlh2* 3'UTR did not affect the Luciferase-pcDNA3 plasmid activity, and that generally, the *Nhlh2* UTR is stable and confers good protein translation to the mRNA. However, the luciferase activity of the construct containing the *Nhlh2*$^{1563G}$ 3'UTR were 38.5% lower than the construct without SNP ($P < 0.01$) (Figure 5C). These results suggest that the presence of the SNP affects the amount of protein translated in cells, likely at the level of mRNA stability.

## 4. Discussion

In this study, we have analyzed the functional consequence of two unique SNPs in the human *NHLH2* mRNA by creating mouse mRNA mimics and using *in vitro* assays and *in silico* methods to characterize the mutant versus wild-type versions. We found that both SNPs ultimately have the potential to reduce the amount of functional protein available to individual carriers, although in different ways, predicted by their position within the *NHLH2* mRNA. While each of these SNPs are rare in the human population, they are present and may represent additive genetic contributions to human obesity. With obesity affecting up to 35.7% of Americans (Ogden et al., 2012), or over 111 million individuals, the nonsynonymous mutation in the protein-coding region, which was reported in 0.1 percent of obese individuals could be present in up to 111,000 individuals in the US. Likewise, the NCBI database lists the 3'UTR rs11805084 SNP frequency at 7.1%, suggesting that in the American population (313 million individuals, both obese and non-obese) there could be approximately 22 million carriers.

The nonsynonymous polymorphism (*NHLH2*$^{83P}$) changes a species-conserved alanine to a proline within the basic DNA binding domain of *NHLH2*. As characterized by Atchley and Fitch (Atchley and Fitch, 1997), the basic region for type A bHLH proteins extends from the tyrosine at position 79 to the valine at position 90. As we have shown for NHLH2, and as Atchley and Fitch showed for group A bHLH proteins (Atchley and Fitch, 1997), the alanine at position 83 is highly conserved both across species and other bHLH family members. Furthermore, this position appears to be key amino acid for DNA binding, as predicted using protein structural analysis (Atchley and Zhao, 2007). The aberrant migration of the mutant protein on a polyacrylamide gel, as well as the *in silico* analysis, supports the need for future studies to determine whether NHLH2$^{83P}$ protein can bind to, and transactive known NHLH2 target genes. The SNP also affects the amino acid sequence of NHLH2 in a region where there are predicted post-translational modifications (data not shown). Future work on whether the NHLH2$^{83P}$ protein is post-translationally modified like the NHLH2 protein is warranted.

The SNP in the 3'UTR (rs144106335) is of interest because of its location within an AU-rich sequence elements (ARE) binding protein site (for mRNA stability). Thus, this SNP has the potential to alter the *NHLH2* mRNA and protein level through effects on posttranscriptional regulation mechanisms. In this study, we were able to prove that SNP rs144106335, engineered into mouse *Nhlh2* mRNA affects mRNA stability and the total amount of protein made. mRNA stability, as measured using transcript decay following actinomycin D treatment shows an overall lower level of mRNA transcript even after just 30 minutes of treatment, and statistically lower levels of the mutant mRNA transcript at each time point thereafter. These suggests predict that there will be less mRNA overall available for translation, and the prediction was confirmed by engineering the *Nhlh2* tail region, either as the wildtype sequence, or with the mutation, onto the luciferase protein. As luciferase protein has a 3.68 hour half life in cells (Leclerc et al., 2000), any difference seen in luciferase level is due to the amount of mRNA available for translation. The significant reduction in relative luciferase activity confirms the prediction that less luciferase mRNA is

available for translation when the *Nhlh2*$^{1563G}$ tail, rather than the *Nhlh2*$^{1563A}$ or SV-40 tail is used.

Reduced stability of *Nhlh2*$^{1563G}$ mRNA could be caused by the addition of a destabilizing miRNA site, or by loss of a stabilizing RNA:protein binding site. A search using miRNA Target finder online database, MicroInspector, predicts that mmu-miR-615-5p will bind to either region, and therefore probably is not a contributing factor to the instability. Our data using *in silico* analysis suggests that the overall structure of the *Nhlh2*$^{1563G}$ mRNA could be changed. RNA secondary structures can thereby affect an RNA protein binding site, especially when these occur within AUUUA-rich elements (AREs) (Wilson and Brewer, 1999). The *NHLH2* mRNA 3'UTR has 8 AREs. Yugami and colleagues have shown that hnRNP-U protein can bind to, and stabilize the 3'UTR of the human *NHLH2* (Yugami et al., 2007). The *NHLH2* 3'UTR SNP is in one of these 8 AU rich regions and it is possible that converting the A to G at this position results in eliminating the binding site of stabilizing protein hnRNP-U. Another possibility is that the SNP causes a change in cellular localization of the mRNA, prior to translation. For the endogenous c- MYC gene, a SNP in it's 3'UTR does cause this change, which results in lower overall protein expression (Chabanon et al., 2005).

NHLH2 is a transcription factor, and all work published to date supports its role in neuronal development, reproduction, physical activity, energy metabolism and overall body weight control (Good et al., 1997; Coyle et al., 2002; Johnson et al., 2004; Kruger et al., 2004; Ruschke et al., 2009). SNPs in the *NHLH2* mRNA could ultimately affect any of these physiological processes. For example, we have previously shown that Nhlh2 can interact with the leptin-induced transcription factor Stat3, and that this interaction is necessary for transactivation of the Pc1/3 gene in response to leptin (Fox and Good, 2008). We have also shown that Nhlh2 is required for Mc4R expression (Wankhade and Good, 2011). As *PC1/3* and *MC4R* are two genes already known to contribute to human monogenetic obesity (Jackson et al., 1997; Yeo et al., 1998), our new results suggest it is possible that NLHH2$^{83P}$ would be unable to bind or show significantly reduced binding to the promoter region of *PC1/3* and *MC4R*, and lower overall expression levels of these Nhlh2 targets. Likewise the SNP in the 3'UTR (rs144106335) would also contribute to lower, albeit functionally WT protein. These results suggest that carriers of either SNP may be phenotypically like a heterozygous *Nhlh2* knockout mice, with lower levels of functional NHLH2 protein. Mice that are heterozygous for deletion of *Nhlh2* show a slower onset of obesity and body fat gain, compared to homozygous *Nhlh2* KO mice, but with significant increases compared to WT animals (Coyle et al., 2002). As the *NHLH2*$^{83P}$ SNP was originally identified by a group searching for SNPs affecting human body weight, the results of our study are suggestive that functional mutations in *NHLH2* do contribute to adult-onset obesity, either mono- or polygenetically, in some individuals.

In summary, the data presented provides functional information on two rare human SNPs in *NHLH2* gene. One of these has been linked to human obese phenotypes, while the other is present in a relatively high proportion of individuals. Based on the results presented, especially with respect to the SNPs effects on relative NHLH2 protein levels in carriers, both deserve further analysis in whether they are causative and/or additive for human body weight and fertility phenotypes.

## Acknowledgments

The authors would like to thank Ms. Risa Pesapane, Ms. Haiyan Zhang, and Ms. Jinhua Zhang for excellent technical assistance. The work was supported in part by funding from the National Institutes of Health (NIH): R01 DK59903 (DJG), NIH 1RC1DK086655-01 and internal departmental and college funds (DJG).

## Abbreviations

| | |
|---|---|
| **A** | adenosine |
| **ARE** | AU-rich element |
| **βgal** | beta galactosidase |
| **bHLH** | basic helix-loop-helix |
| **bp** | base pair |
| **C** | cytosine |
| **dATP** | deoxy adenosine tri-phosphate |
| **DNA** | deoxyribonucleic acid |
| **G** | guanosine |
| **kb** | kilobase |
| **KO** | knockout |
| **MC4R** | melanocortin 4 receptor |
| **μg** | microgram |
| **μl** | microliter |
| **mRNA** | messenger ribonucleic acid |
| **miRNA** | micro ribonucleic acid |
| **NCBI** | National Center for Biotechnology Information |
| **ng** | nanogram |
| **NHLH2** | Nescient helix-loop-helix 2 (human) |
| **Nhlh2** | Nescient helix-loop-helix 2 (mouse) |
| **nt** | nucleotide |
| **oligo** | oligonucleotide |
| **PC1/3** | prohormone convertase 1/3 |
| **QPCR** | quantitative polymerase chain reaction |
| **rRNA** | ribosomal ribonucleic acid |
| **SDS** | sodium dodecyle sulfate |
| **SNP** | single nucleotide polymorphism |
| **SV-40** | simian virus 40 |
| **T** | thymidine |
| **U** | uridine |
| **UTR** | untranslated region |
| **V** | volts |
| **WT** | wild type |

## References

Ahituv N, Kavaslar N, Schackwitz W, Ustaszewska A, Martin J, Hebert S, Doelle H, Ersoy B, Kryukov G, Schmidt S, Yosef N, Ruppin E, Sharan R, Vaisse C, Sunyaev S, Dent R, Cohen J, McPherson R, Pennacchio LA. Medical sequencing at the extremes of human body mass. Am J Hum Genet. 2007; 80:779–791. [PubMed: 17357083]

Atchley WR, Fitch WM. A natural classification of the basic helix-loop-helix class of transcription factors. Proc Natl Acad Sci U S A. 1997; 94:5172–5176. [PubMed: 9144210]

Atchley WR, Zhao J. Molecular architecture of the DNA-binding region and its relationship to classification of basic helix-loop-helix proteins. Mol Biol Evol. 2007; 24:192–202. [PubMed: 17041153]

Belsham DD, Cai F, Cui H, Smukler SR, Salapatek AM, Shkreta L. Generation of a phenotypic array of hypothalamic neuronal cell models to study complex neuroendocrine disorders. Endocrinology. 2004; 145:393–400. [PubMed: 14551229]

Bolognani F, Perrone-Bizzozero NI. RNA-protein interactions and control of mRNA stability in neurons. J Neurosci Res. 2008; 86:481–489. [PubMed: 17853436]

Chabanon H, Mickleburgh I, Burtle B, Pedder C, Hesketh J. An AU-rich stem-loop structure is a critical feature of the perinuclear localization signal of c-myc mRNA. Biochem J. 2005; 392:475–483. [PubMed: 16042622]

Coyle CA, Jing E, Hosmer T, Powers JB, Wade G, Good DJ. Reduced voluntary activity precedes adult-onset obesity in Nhlh2 knockout mice. Physiol Behav. 2002; 77:387–402. [PubMed: 12419415]

Fox DL, Good DJ. Nescient helix-loop-helix 2 interacts with signal transducer and activator of transcription 3 to regulate transcription of prohormone convertase 1/3. Mol Endocrinol. 2008; 22:1438–1448. [PubMed: 18356286]

Good DJ, Porter FD, Mahon KA, Parlow AF, Westphal H, Kirsch IR. Hypogonadism and obesity in mice with a targeted deletion of the Nhlh2 gene. Nat Genet. 1997; 15:397–401. [PubMed: 9090387]

Goren A, Kim E, Amit M, Bochner R, Lev-Maor G, Ahituv N, Ast G. Alternative approach to a heavy weight problem. Genome Res. 2007

Goren A, Kim E, Amit M, Bochner R, Lev-Maor G, Ahituv N, Ast G. Alternative approach to a heavy weight problem. Genome Res. 2008; 18:214–220. [PubMed: 18096750]

Jackson RS, Creemers JW, Ohagi S, Raffin-Sanson ML, Sanders L, Montague CT, Hutton JC, O'Rahilly S. Obesity and impaired prohormone processing associated with mutations in the human prohormone convertase 1 gene. Nat Genet. 1997; 16:303–306. [PubMed: 9207799]

Johnson SA, Marin-Bivens CL, Miele M, Coyle CA, Fissore R, Good DJ. The Nhlh2 transcription factor is required for female sexual behavior and reproductive longevity. Horm Behav. 2004; 46:420–427. [PubMed: 15465527]

Katzman S, Barrett C, Thiltgen G, Karchin R, Karplus K. PREDICT-2ND: a tool for generalized protein local structure prediction. Bioinformatics. 2008; 24:2453–2459. [PubMed: 18757875]

Kehrer-Sawatzki H, Cooper DN. Understanding the recent evolution of the human genome: insights from human-chimpanzee genome comparisons. Hum Mutat. 2007; 28:99–130. [PubMed: 17024666]

Kruger M, Ruschke K, Braun T. NSCL-1 and NSCL-2 synergistically determine the fate of GnRH-1 neurons and control necdin gene expression. EMBO J. 2004; 23:4353–4364. [PubMed: 15470499]

Laumen H, Saningong AD, Heid IM, Hess J, Herder C, Claussnitzer M, Baumert J, Lamina C, Rathmann W, Sedlmeier EM, Klopp N, Thorand B, Wichmann HE, Illig T, Hauner H. Functional characterization of promoter variants of the adiponectin gene complemented by epidemiological data. Diabetes. 2009; 58:984–991. [PubMed: 19074982]

Leclerc GM, Boockfor FR, Faught WJ, Frawley LS. Development of a destabilized firefly luciferase enzyme for measurement of gene expression. Biotechniques. 2000; 29:590–591. 594–596. 598 passim. [PubMed: 10997273]

Miller MR, Zhang W, Sibbel SP, Langefeld CD, Bowden DW, Haffner SM, Bergman RN, Norris JM, Fingerlin TE. Variant in the 3' region of the IkappaBalpha gene associated with insulin resistance

in Hispanic Americans: The IRAS Family Study. Obesity (Silver Spring). 2010; 18:555–562. [PubMed: 19798070]

Mukhopadhyay D, Houchen CW, Kennedy S, Dieckgraefe BK, Anant S. Coupled mRNA stabilization and translational silencing of cyclooxygenase-2 by a novel RNA binding protein, CUGBP2. Mol Cell. 2003; 11:113–126. [PubMed: 12535526]

Ogden, CL.; Carroll, MD.; Kit, BK.; Flegal, KM. Prevalence of obesity in the United States, 2009–2010. In: NCfH, editor. Statistics. Hyattsville, MD: 2012.

Ruschke K, Ebelt H, Kloting N, Boettger T, Raum K, Bluher M, Braun T. Defective peripheral nerve development is linked to abnormal architecture and metabolic activity of adipose tissue in Nscl-2 mutant mice. PLoS One. 2009; 4:e5516. [PubMed: 19436734]

Wankhade UD, Good DJ. Melanocortin 4 receptor is a transcriptional target of nescient helix-loop-helix-2. Mol Cell Endocrinol. 2011; 341:39–47. [PubMed: 21664420]

Wilson GM, Brewer G. The search for trans-acting factors controlling messenger RNA decay. Prog Nucleic Acid Res Mol Biol. 1999; 62:257–291. [PubMed: 9932457]

Yeo GS, Farooqi IS, Aminian S, Halsall DJ, Stanhope RG, O'Rahilly S. A frameshift mutation in MC4R associated with dominantly inherited human obesity. Nat Genet. 1998; 20:111–112. [PubMed: 9771698]

Yugami M, Kabe Y, Yamaguchi Y, Wada T, Handa H. hnRNP-U enhances the expression of specific genes by stabilizing mRNA. FEBS Lett. 2007; 581:1–7. [PubMed: 17174306]

Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 2003; 31:3406–3415. [PubMed: 12824337]

## Highlights

- NHLH2 is a bHLH transcription factor that controls body weight and fertility.
- A 3'UTR and a coding region SNP both reduce functional NHLH2 protein levels.
- The 3'UTR SNP affects RNA secondary structure and mRNA stability.
- The basic domain SNP modifies the protein helical structure.
- Rare SNPs in NHLH2 may contribute to body weight or fertility phenotypes in humans.

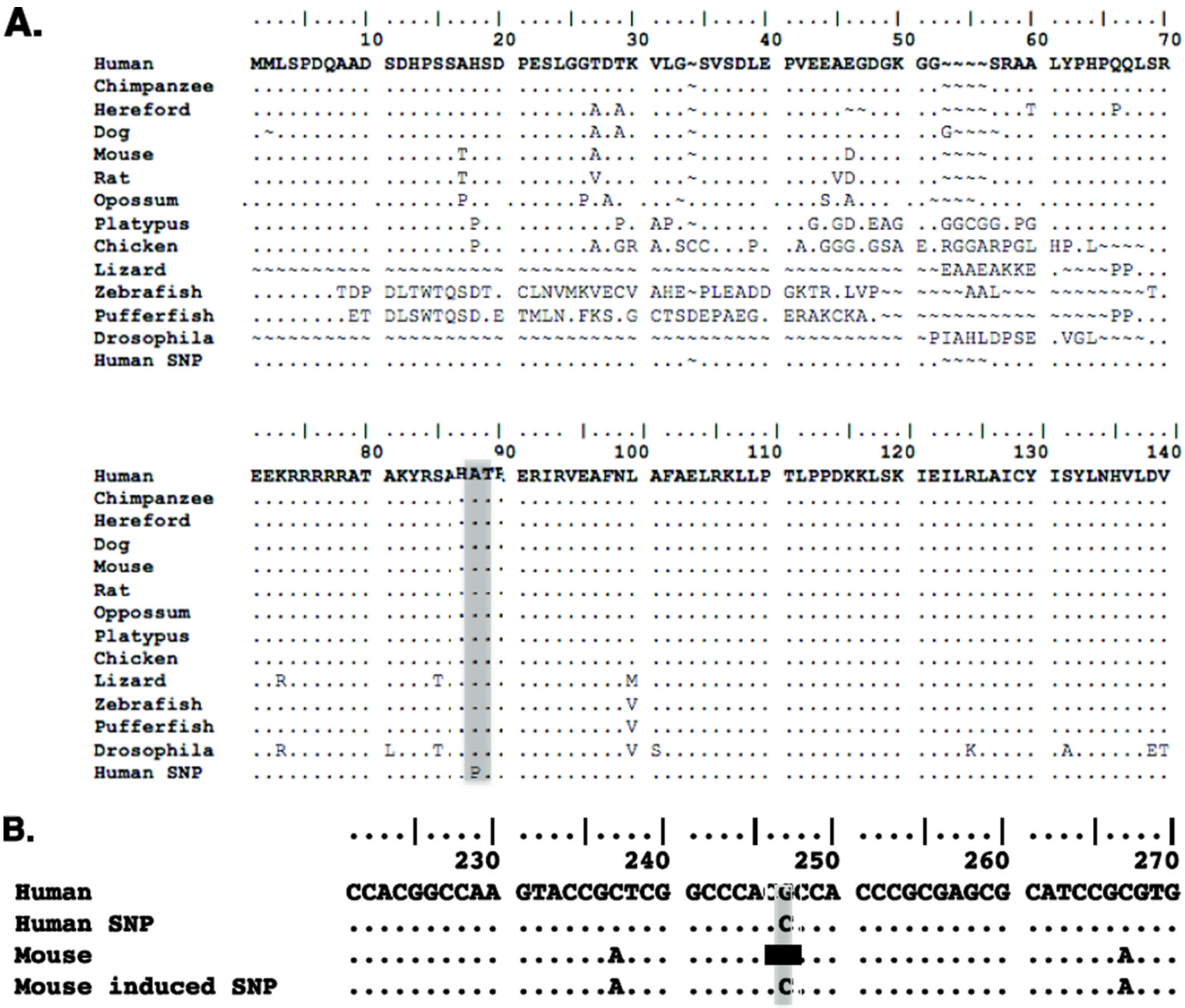



**Figure 1. Phylogenetic comparison of the NHLH2 protein sequence**

(**A**) Bioedit software was used for the comparative analysis with the NHLH2 amino acid sequences from human (*Homo sapiens*, CAI14533.1), chimpanzee (*Pan troglodytes*, XP_003308343.1), cattle (*Bos taurus*, DAA31577.1), dog (*Canis lupus familiarus*, XP_540253.3), mouse (*Mus musculus*, AAH58413.1), rat (*Rattus norvegicus*, NP_001099927.1), oppossum (*Monodelphis domestica*, ENSMODP00000029001.1), platypus *(Ornithorhynchus anatinus*, ENSOANP00000015572.1*)*, lizard (*Anolis carolinensis*, ENSACAP00000003207.2) chicken (*Gallus gallus*, NP_990128.1), zebrafish (*Danio rerio*, NP_991232.1), pufferfish (*Takifugu rubripes*, scaffold 64, nucleotides 60203-60571), and fruit fly (*Drosophila melanogaster*, FBpp0070659.3). The NHLH2$^{83P}$ sequence was derived from CAI14533.1 by substituting a proline for an alanine at position 83. (**B**) The nucleotide sequences of Human *NHLH2* (*Homo sapiens*, NM_005599.3), and mouse *Nhlh2* (*Mus musculus*, NM_178777.3) were used for analysis. The Human SNP, and the engineered mouse SNP are shown. In both panels, the shaded area shows the position of the SNP within the sequence.

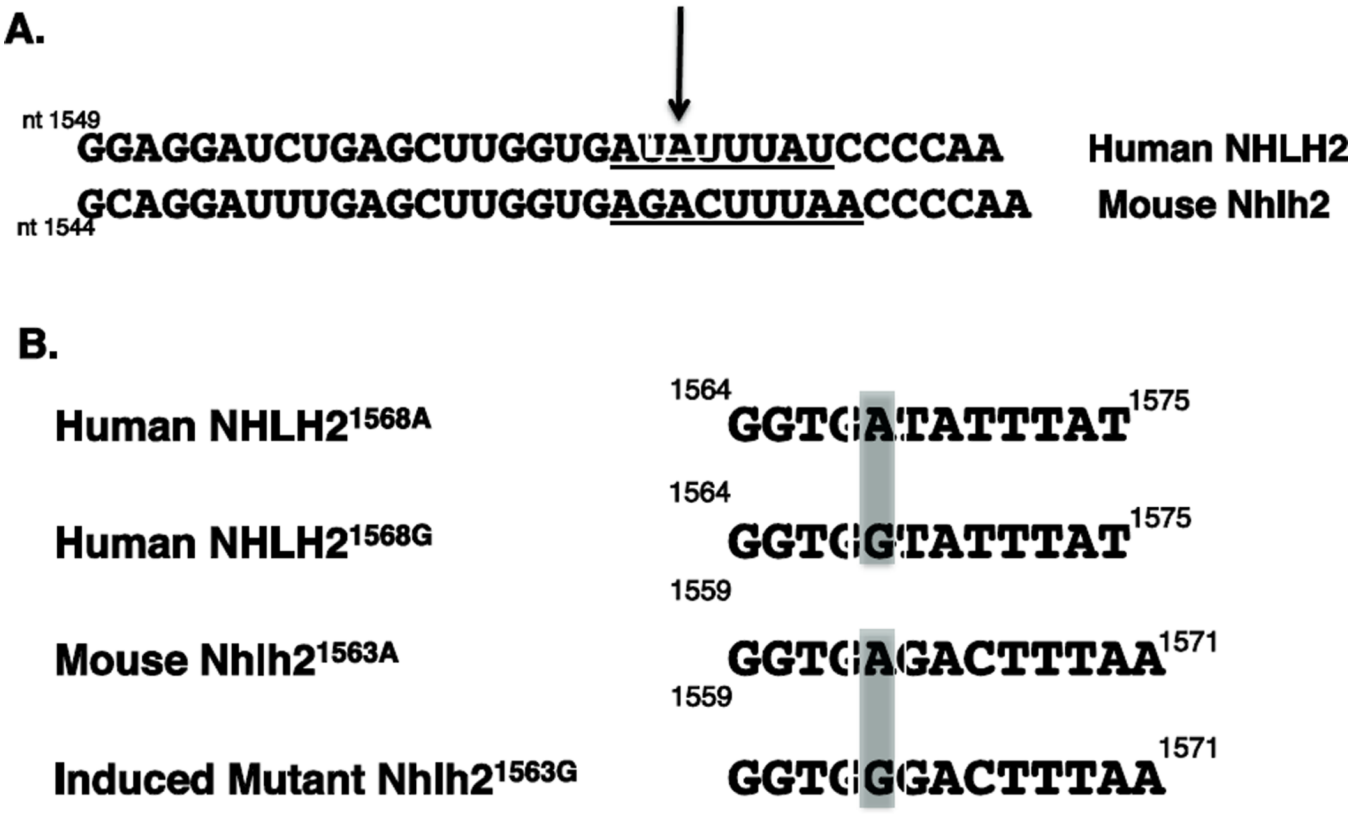
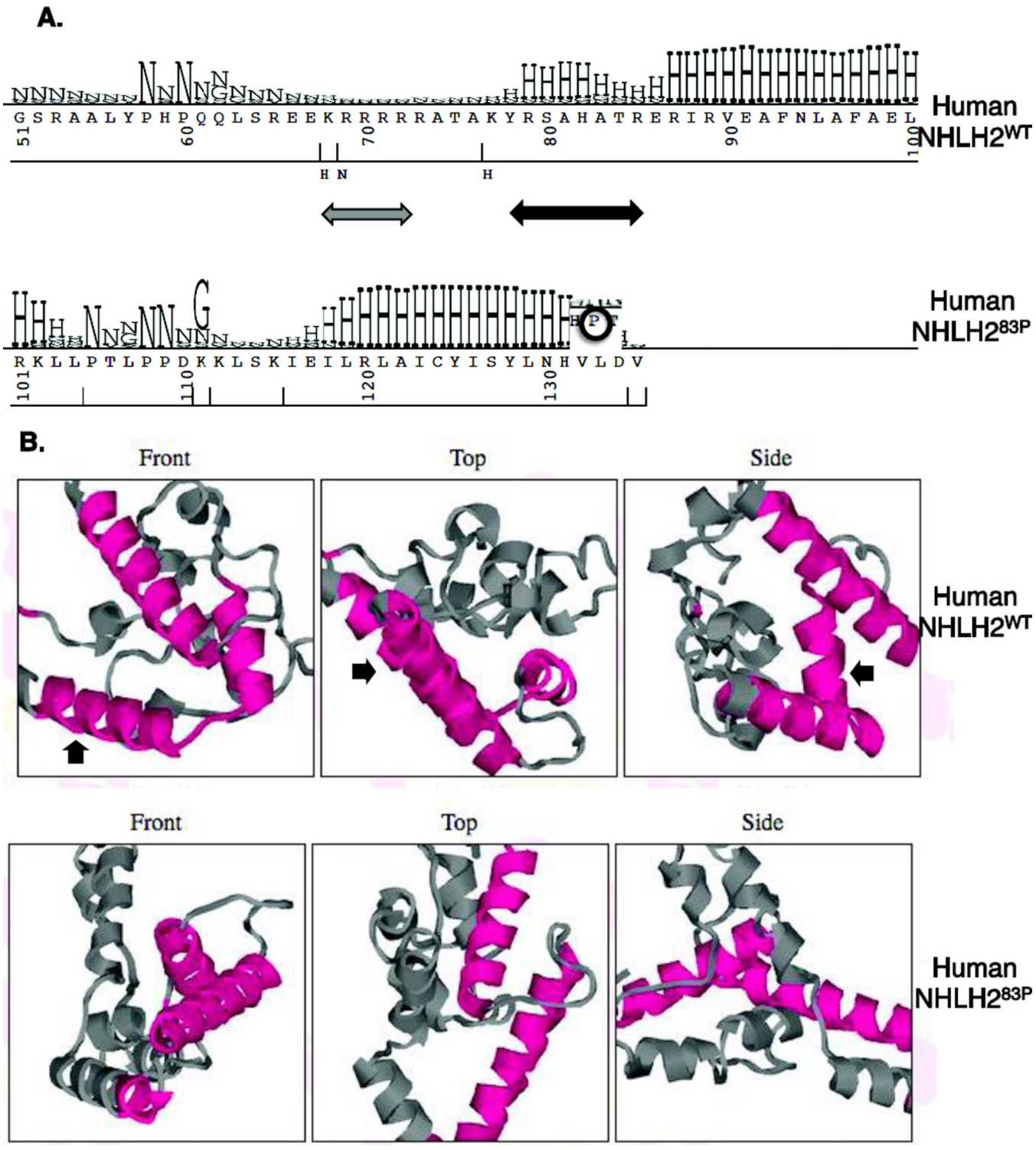



**Figure 2. Comparison of Human rs144106335 with mouse *Nhlh2***

(**A**) Comparison of the Human *NHLH2* and mouse *Nhlh2* mRNAs within the region containing the human SNP rs144106335. The location of the SNP within the human *NHLH2* mRNA is indicated by an arrow, with the corresponding "A" also present in mouse *Nhlh2*. A region containing a putative mRNA stability motif is underlined in both sequences. (**B**) Genomic DNA sequence for the WT Human *NHLH2* (Human *NHLH2*$^{1568A}$), Human *NHLH2* containing the SNP ((Human *NHLH2*$^{1568G}$), WT mouse *Nhlh2* (Mouse *Nhlh2*$^{1563A}$) or mouse *Nhlh2* containing the induced mutation to mimic the human SNP (Mouse *Nhlh2*$^{1563G}$). The shaded area shows the position of the SNP within the sequence.



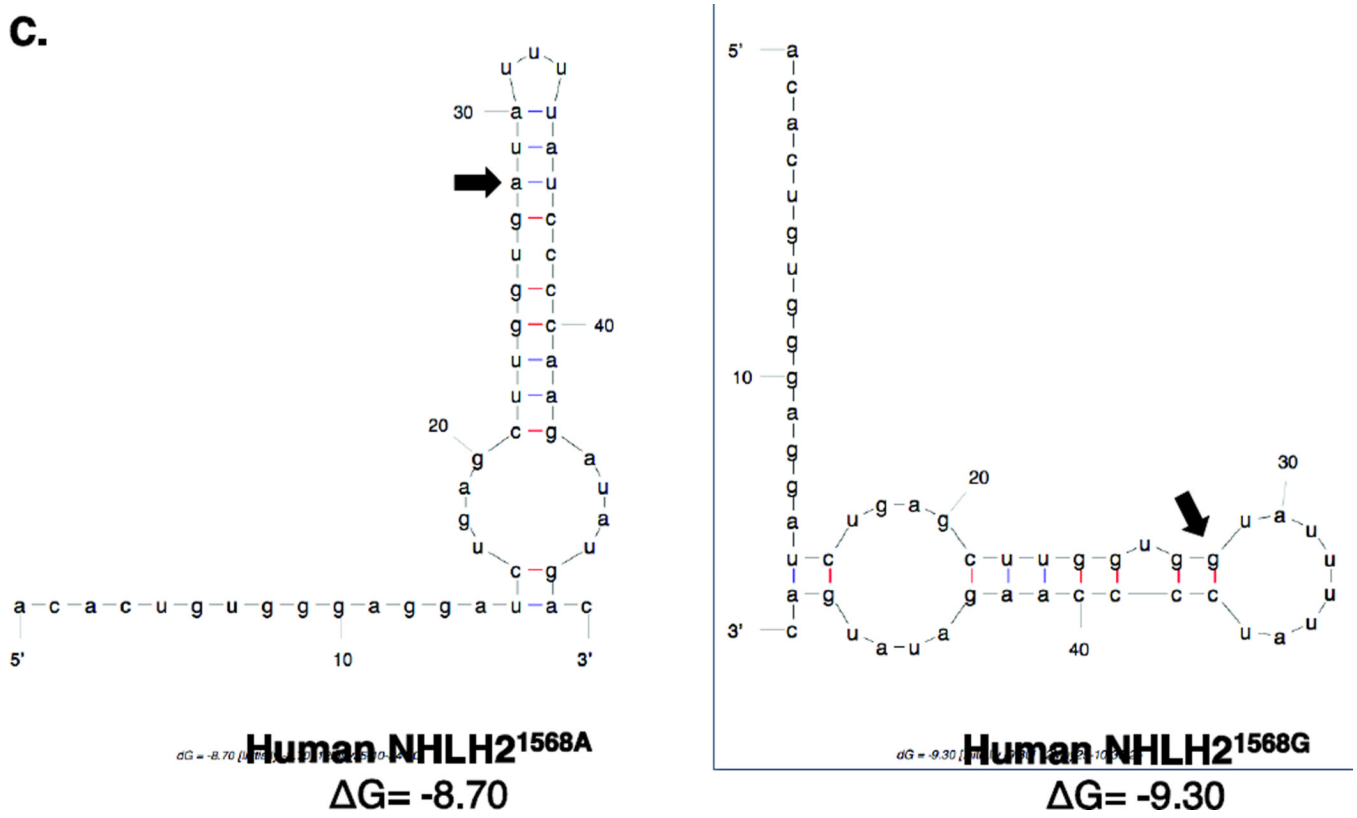



**Figure 3.** ***In silico*** **protein and mRNA structural analysis**

(**A**) The wildtype and mutant protein sequences were submitted to the SAM-T08 protein structure prediction online server. The figure shows the results for the region between amino acids 51–100, where “G” indicates separation +3 (3–10 helix), “H” indicates separation +4 (alpha helix), and “N” indicates that no hydrogen bond is formed. The double-headed arrows indicate the regions affected by the proline substitution, which is circled (**B**) Screen shots from secondary structural predications from the SAM-T08 protein structure prediction online server. The arrows represent the helical structure that is present in the WT protein, but absent in NHLH2$^{83P}$. (**C**) Screen shots from the RNA secondary structure prediction, using MFold web server. The arrow indicates the position of the SNP for the WT (in *NHLH2*$^{1568A}$) and mutant (in *NHLH2*$^{1568G}$) mRNAs.

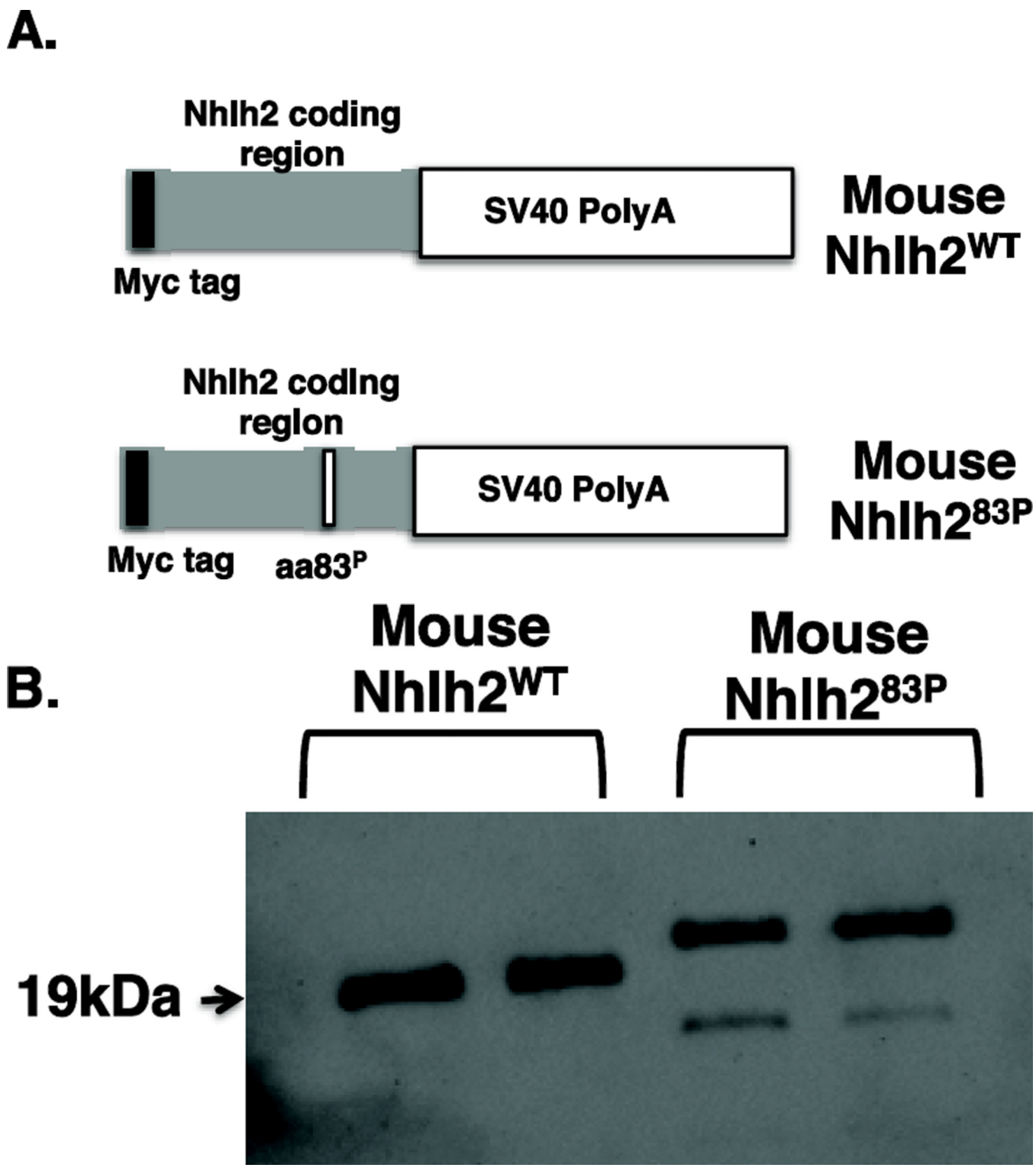
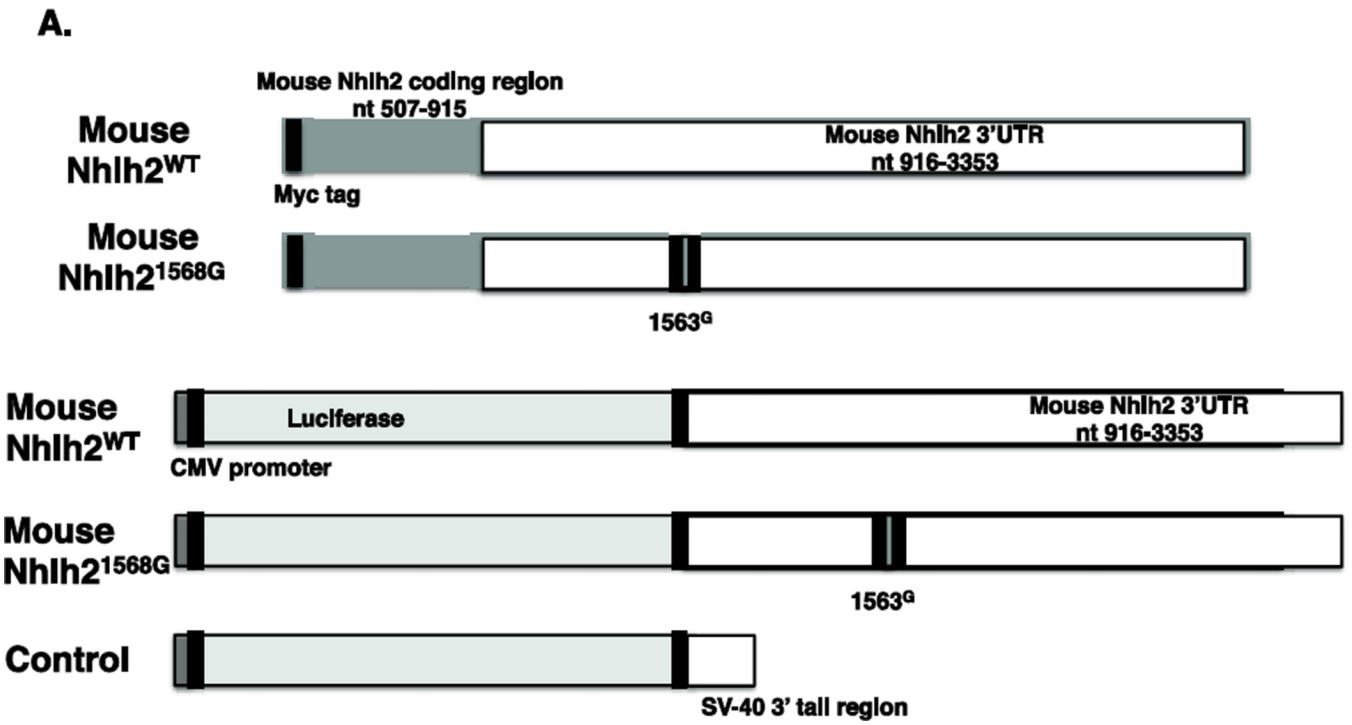
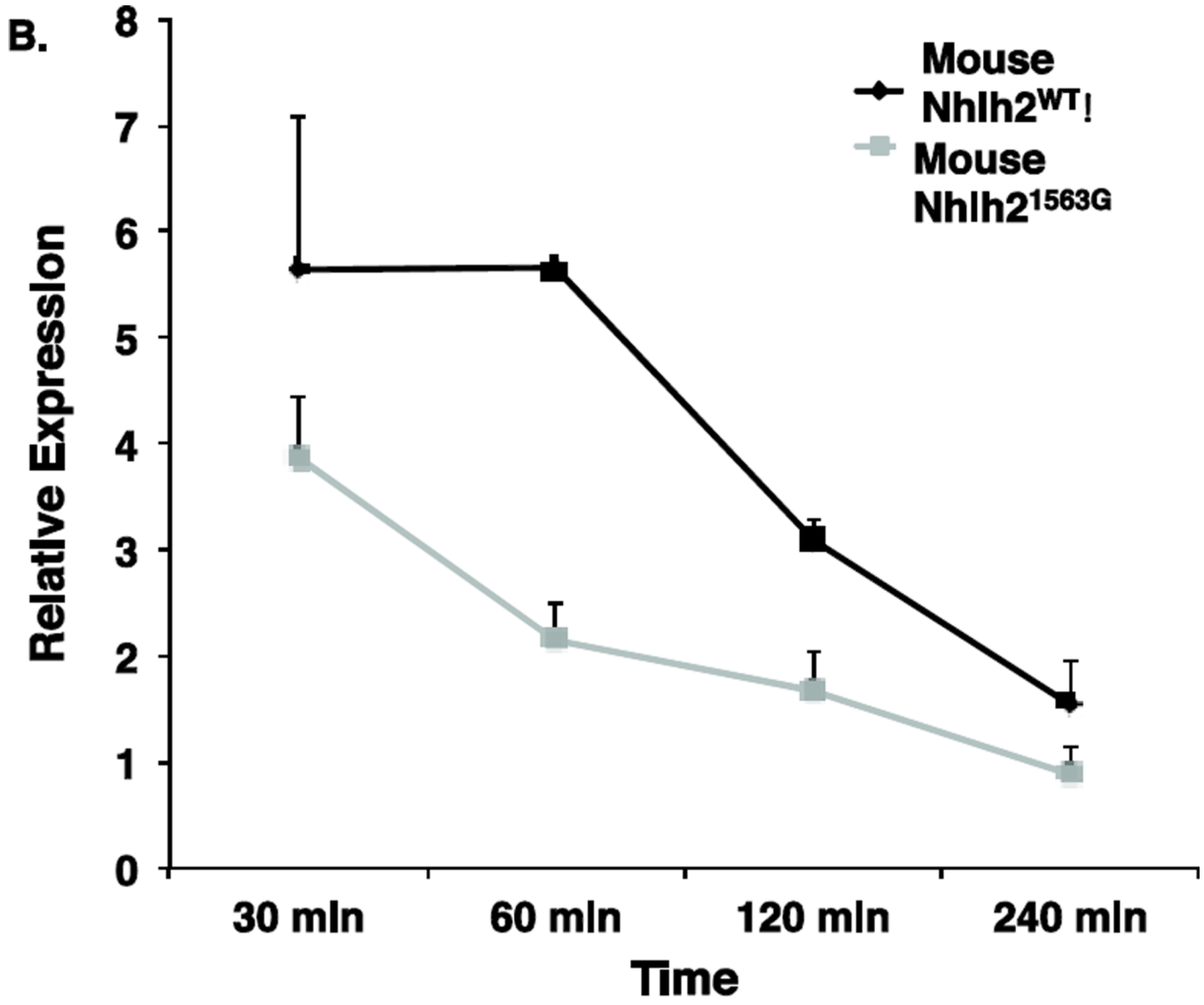



**Figure 4. Analysis of Nhlh2$^{83P}$ protein**

(**A**) Schematic of the myc-tagged constructs used in the functional analysis of Nhlh2$^{83P}$. (**B**) Western analysis using anti-cmyc antibody on whole cell extracts. Cell extracts from two

transfections for each construct are shown. The size, in kilodaltons (KDa) is shown for the WT protein.





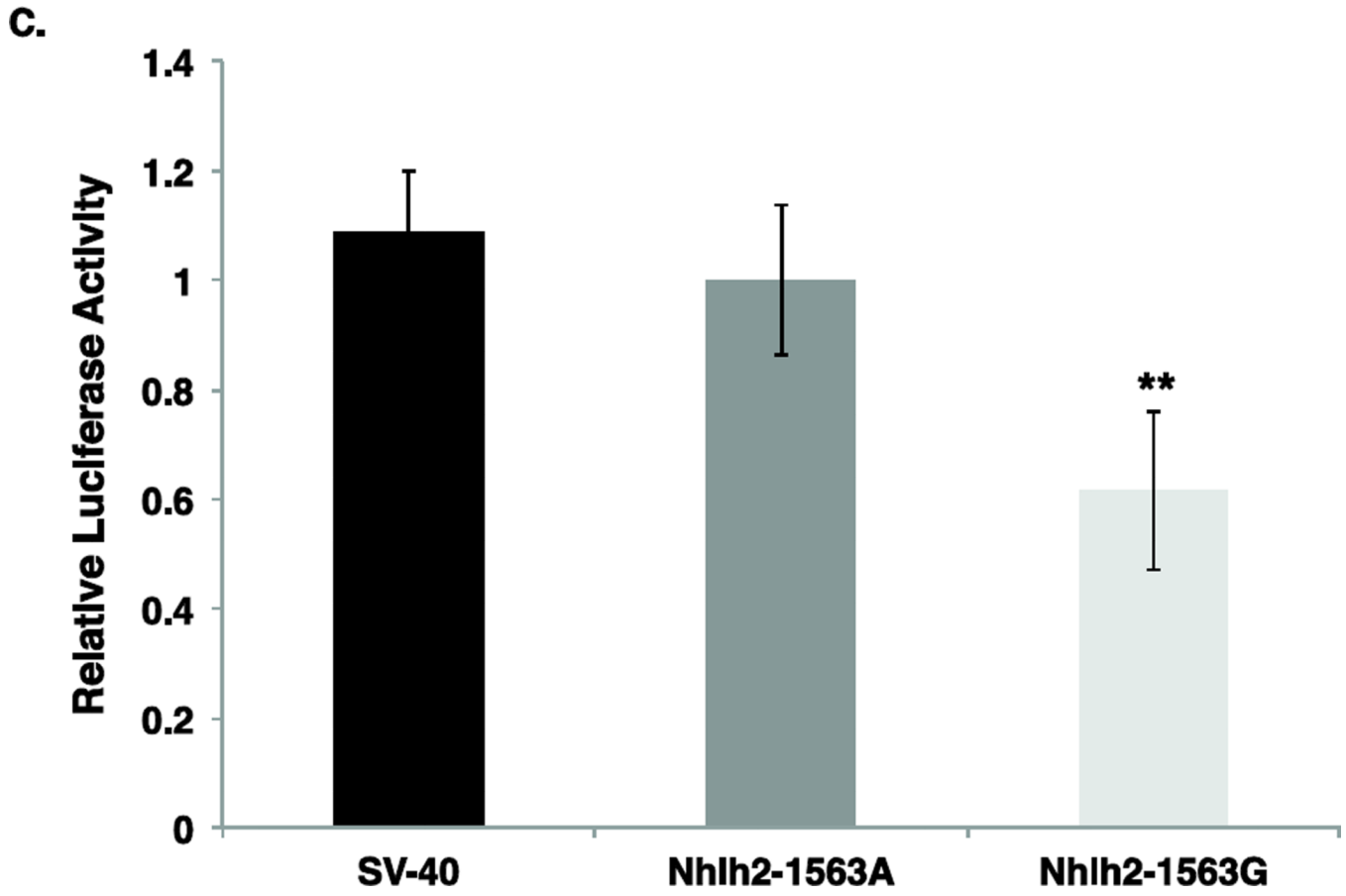



**Figure 5. mRNA stability analysis for *Nhlh2*$^{1563G}$**

(**A**) Schematic representation of the constructs used in the study. The position of the site-directed mutation is shown for the mutant constructs. (**B**) N29/2 cells transfected with the construct containing Myc-tagged *Nhlh2* with the *Nhlh2*$^{WT}$ 3'UTR (black diamond) or with the same plasmid containing Myc-tagged *Nhlh2*$^{1563G}$ with the mutant 3'UTR (grey squares). Cells were treated with actinomycin D and RNA was harvested at different time points after the Actinomycin D treatment. The 18S rRNA reference gene was used as a control. The data shown are the means ± SEM. The overall trend (slope analysis/area under the curve) between WT and Mutant construct was highly significant ($P<0.01$). (**C**) Luciferase activity for transfections using constructs containing the SV-40 3'UTR (SV-40), the *Nhlh2*$^{WT}$ 3'UTR, or the *Nhlh2* mutant 3'UTR construct (*Nhlh2*$^{1563G}$), all fused downstream of the luciferase gene. All values are reported relative to β-galactocidase activity, which was used as a control for the transfection efficiency. The data shown are the means ± SEM, ** $P < 0.01$.

**Table 1**

Oligonucleotides used in methods

| OLIGO SEQUENCE | GENE NAME/PURPOSE |
|---|---|
| 5'-ATCGGCCCACCCCACCCGCGAG-3' | In vitro mutagenesis, forward primer, Nhlh2 |
| 5'-TCGCGGGTGGGGTGGGCCGATC-3’ | In vitro mutagenesis, reverse primer, Nhlh2 |
| 5'-GCAGGATTTGAGCTTGGTGGGACTTTAACCCCAAGATAG-3’ | In vitro mutagenesis, forward primer, Nhlh2 |
| 5’- CTATCTTGGGGTTAAAGTCCCACCAAGCTCAAATCCTGC-3’ | In vitro mutagenesis, reverse primer, Nhlh2 3’ |
| 5’-ATGGAGAGCTTGGGCGACCTCA-3’ | Amplification forward primer, Nhlh2-myc |
| 5’-TTGGTCCGGACTCAGCATCATCGAAT-3’ | Amplification reverse primer, Nhlh2-myc |
| 5’-CGCCGCTAGAGGTGAAATTC-3’ | Amplification forward primer, 18S |
| 5’-TTGGCAAATGCTTTCGCTC-3’ | Amplification reverse primer, 18S |
| 5'-CTGCCAAGCAGTTGGGATTAAGGG-3' | EMSA analysis, forward primer, necdin |
| 5'-CCCTTAATCCCAACTGCTTGGCAG-3' | EMSA analysis, reverse primer, necdin |
| 5'-CTTGGGATCTTGGTATCCATATTCAACCGAGGCGGTTCTCTT-3' | EMSA analysis, forward primer, Pc1/3 |
| 5'-AAGAGAACCGCCTCGGTTGAATATGGATACCAAGATCCAAG-'3 | EMSA analysis, reverse primer, Pc1/3 |
| 5'-GCAGAAACTGCAAATGGAGAAACAGCT-3' | EMSA analysis, forward primer, Mc4R |
| 5'-AGCTGTTTCTCCATTTGCAGTTTCTGC-3' | EMSA analysis, reverse primer, Mc4R |