Structure and function analysis of the essential 3'X domain of hepatitis C virus

Jesús Castillo-Martínez, Tamara Ovejero, Cristina Romero-López, Isaías Sanmartín, Beatriz Berzal-Herranz, Elisa Oltra, Alfredo Berzal-Herranz and José Gallego

SUPPLEMENTARY MATERIAL

1. Supplementary Figures

2. Supplementary Tables

1. Supplementary Figures

3'X	5' GGUGGCUCCA UCUUAGCCCU AGUCACGGCU AGCUGUGAAA GGUCCGUGAG CCGCUUGACU GCAGAGAGUG CUGAUACUGG CCUCUCUGCA GAUCAAGU 3'
U3C	5' GG**C**GGCUCCA UCUUAGCCCU AGUCACGGCU AGCUGUGAAA GGUCCGUGAG CCGCUUGACU GCAGAGAGUG CUGAUACUGG CCUCUCUGCA GAUCAAGU 3'
U55C	5' GGUGGCUCCA UCUUAGCCCU AGUCACGGCU AGCUGUGAAA GGUCCGUGAG CCGC**C**UGACU GCAGAGAGUG CUGAUACUGG CCUCUCUGCA GAUCAAGU 3'
U3C/U55C	5' GG**C**GGCUCCA UCUUAGCCCU AGUCACGGCU AGCUGUGAAA GGUCCGUGAG CCGC**C**UGACU GCAGAGAGUG CUGAUACUGG CCUCUCUGCA GAUCAAGU 3'
U3G	5' GG**G**GGCUCCA UCUUAGCCCU AGUCACGGCU AGCUGUGAAA GGUCCGUGAG CCGCUUGACU GCAGAGAGUG CUGAUACUGG CCUCUCUGCA GAUCAAGU 3'
G50C/C52G	5' GGUGGCUCCA UCUUAGCCCU AGUCACGGCU AGCUGUGAAA GGUCCGUGA**C** C**G**GCUUGACU GCAGAGAGUG CUGAUACUGG CCUCUCUGCA GAUCAAGU 3'
U3G/G50C/C52G	5' GG**G**GGCUCCA UCUUAGCCCU AGUCACGGCU AGCUGUGAAA GGUCCGUGA**C** C**G**GCUUGACU GCAGAGAGUG CUGAUACUGG CCUCUCUGCA GAUCAAGU 3'
C29G/A31U	5' GGUGGCUCCA UCUUAGCCCU AGUCACGG**G**U **U**GCUGUGAAA GGUCCGUGAG CCGCUUGACU GCAGAGAGUG CUGAUACUGG CCUCUCUGCA GAUCAAGU 3'
cre26	5' GGAGACAUAU AUCACAGCCU GUCUCC 3'

Figure S1. **RNA sequences studied by NMR spectroscopy and gel mobility assays.** Mutations in the 3'X domain are identified with bold blue font, self-complementary DLS nt are depicted in red, and nt of the *k* and *k'* sequence motifs are underlined in green. The cre26 mutations inserted to increase transcription yield are marked with light blue font.

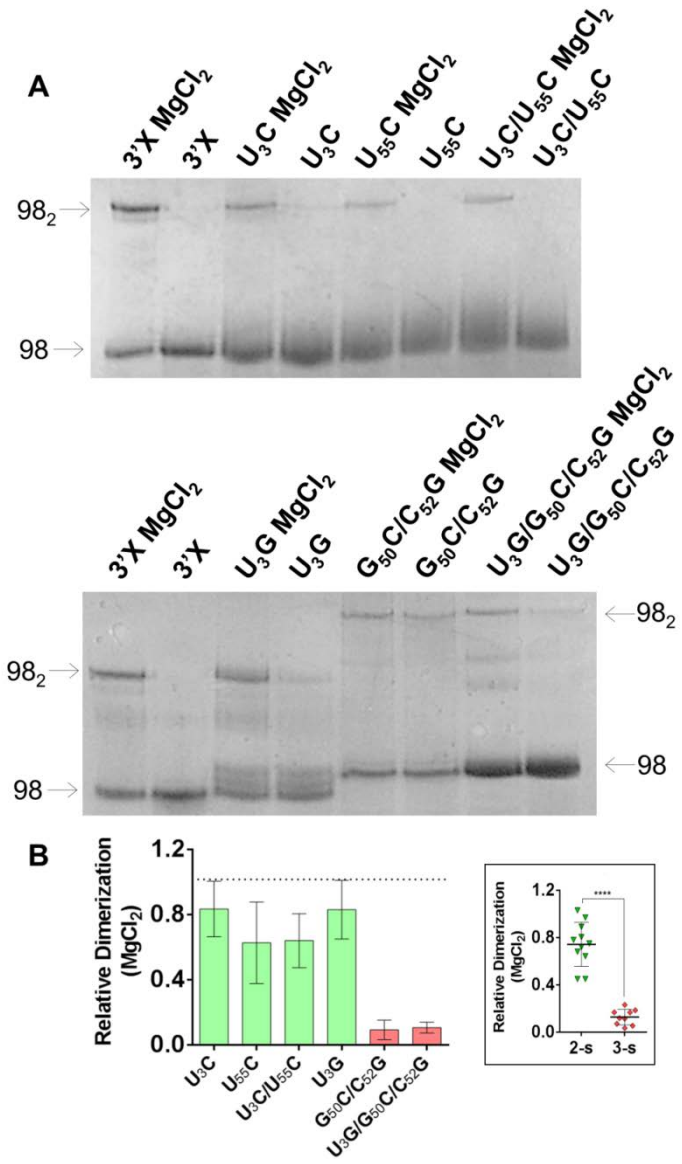

Figure S2. **Impact of Mg^{2+} ions on the conformation of 3'X domain mutants, analysed by native gel mobility assays.** (**A**) Native gels comparing the electrophoretic mobility of wild-type 3'X domain and U_3C, $U_{55}C$, $U_3C/U_{55}C$, U_3G, $G_{50}C/C_{52}G$ and $U_3G/G_{50}C/C_{52}G$ mutants, previously folded in the absence or presence of 2 mM $MgCl_2$. (**B**) Quantification of the homo-dimerization capacity at 2 mM $MgCl_2$ of 3'X domain mutants, obtained by measuring in each lane the fraction of 3'X homodimers (98_2) relative to total 3'X RNA. The results were normalized with respect to the wild-type sequence, which was assigned a value of 1, and the bars represent the average and standard deviation of three independent experiments. Mutants experimentally verified to adopt the wild-type two-stem conformation are represented in green, whereas mutants adopting a different structure are indicated in red. In the scatter plot shown in the inset, the individual values of the two-stem and three-stem mutants are indicated, together with means and standard deviations. The differences between two-stem and three-stem mutants are significant in all cases ($p < 0.005$). Conditions: 10-40 μM RNA, TBM running buffer (2 mM $MgCl_2$).

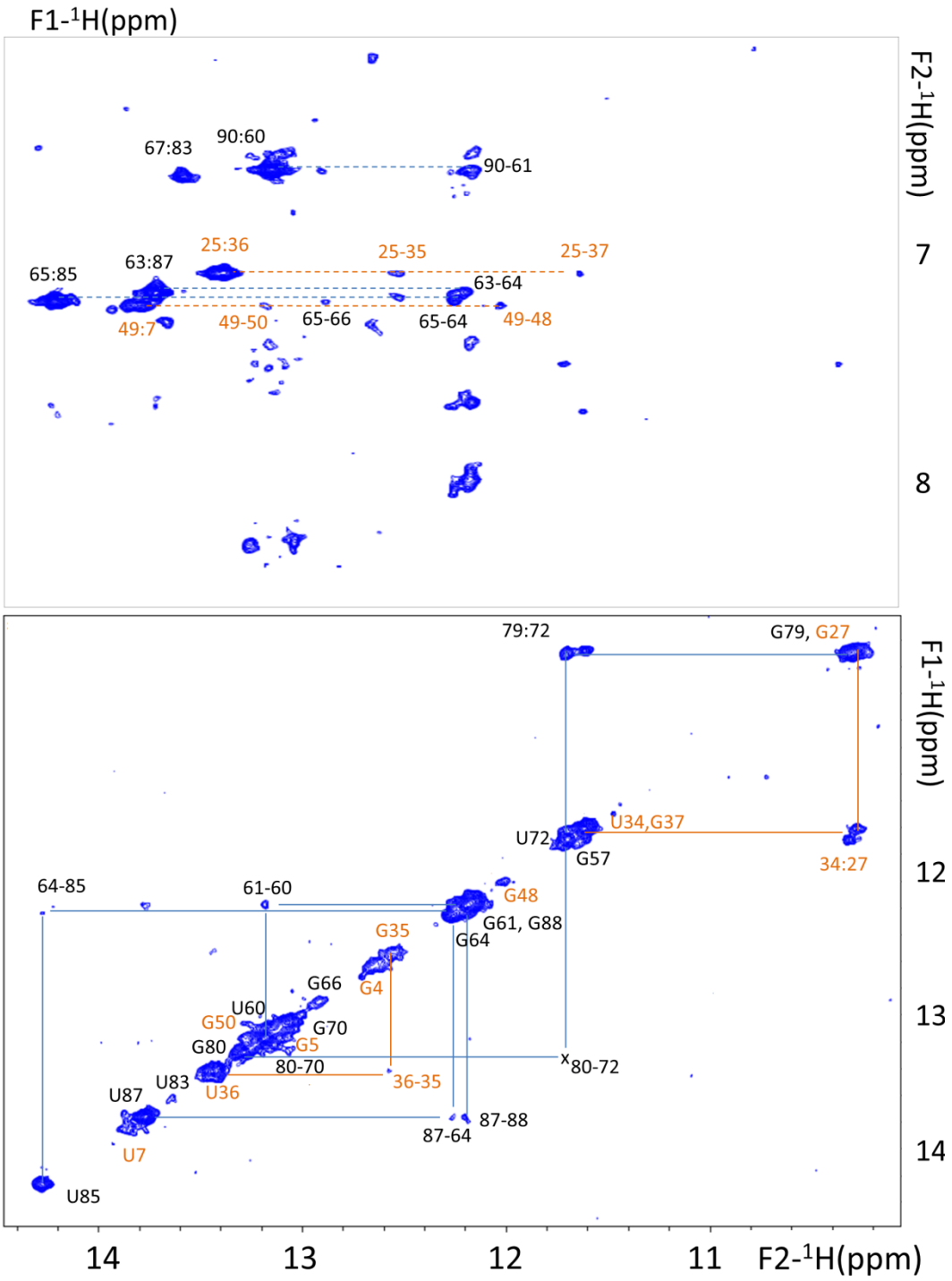

Figure S3. NOESY analysis of the $U_3C/U_{55}C$ 3'X domain mutant. Assignment of the imino-aromatic (top) and imino-imino (bottom) regions of a watergate-NOESY spectrum acquired at 15 ^{o}C with 150 ms mixing time. Residues belonging to the SL1' and SL2' subdomains are labeled in black and orange, respectively. Sequential crosspeaks between imino protons are indicated with solid lines, sequential contacts between A H2 and G H1 protons are marked with dashed lines, and NOE interactions across base pairs are labelled with colons. Assignments of C amino protons have been omitted for clarity, and crosspeaks marked with crosses are visible at a lower contour level. Conditions: 118 µM RNA, 0 mM NaCl/$MgCl_2$.

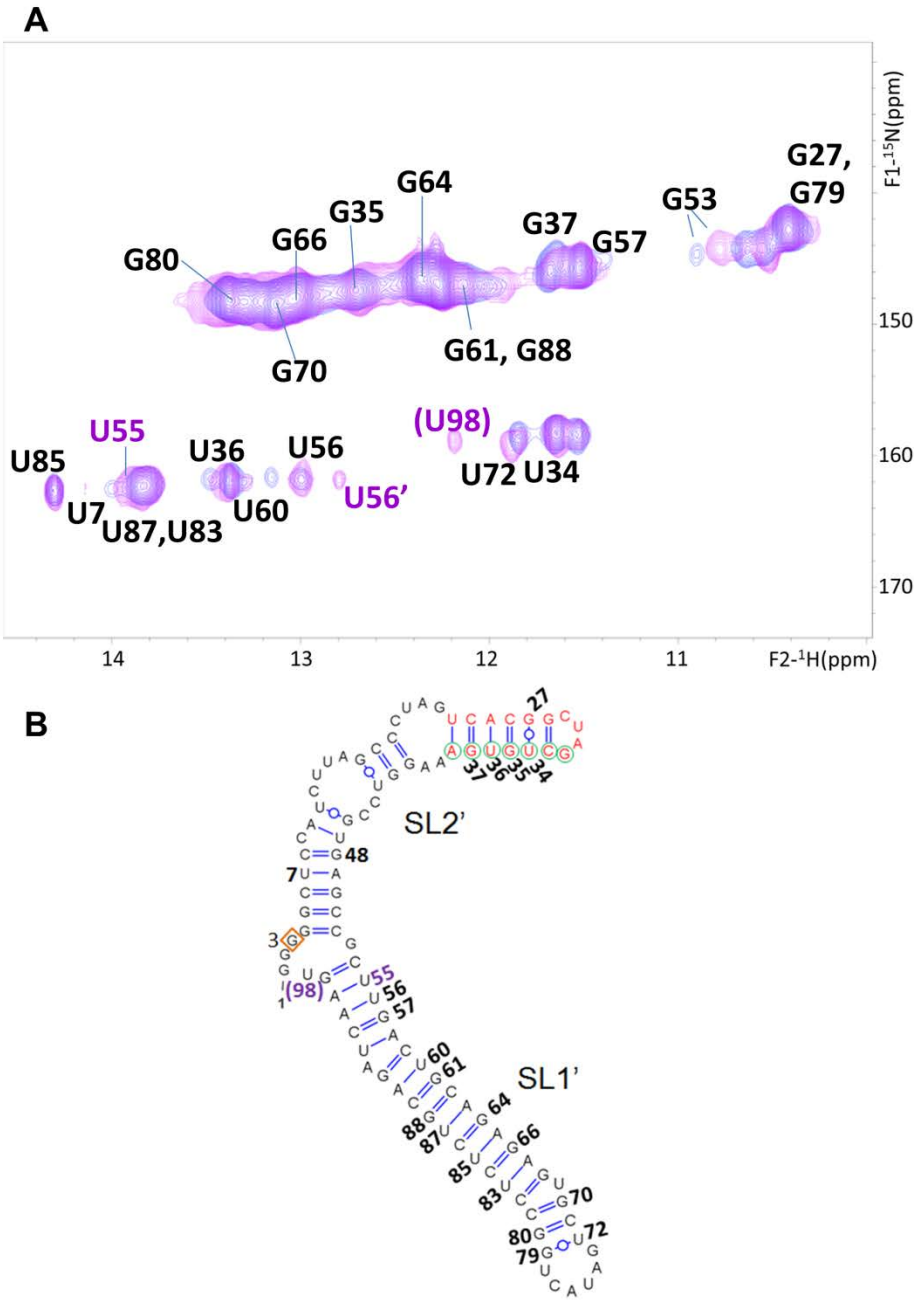

Figure S4. **Structural analysis of the U_3G 3'X mutant.** (**A**) ^{1}H-^{15}N HSQC spectrum of the U_3G mutant at low ionic strength (purple), superposed with that of the wild-type domain (blue) acquired under the same temperature and ionic conditions. All assignments appear in black except for those of mutant signals not matching wild-type resonances, which are labelled in purple. The U56 signal is doubled due to conformational exchange. (**B**) Secondary structure of the U_3G mutant domain supported by the NMR data. The U_3G mutation destabilizing the lower stem of subdomain SL2' is marked with a red square, DLS motif nt are depicted in red, and *k* motif nt are indicated with green circles. Residues with assigned HN resonances are numbered. Conditions: 53 μM RNA, 0 mM $NaCl/MgCl_2$, 25 ^{o}C.

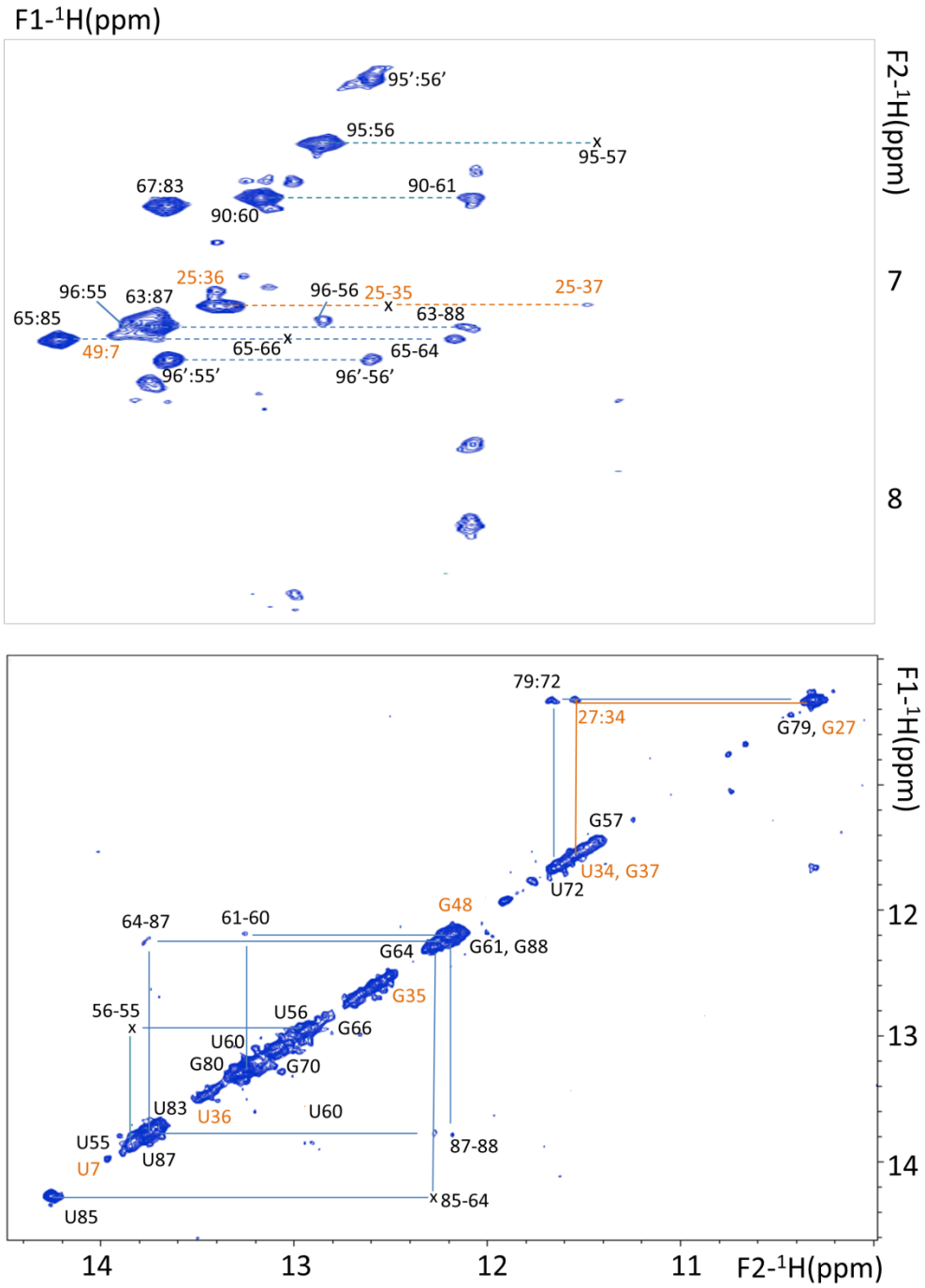

Figure S5. NOESY analysis of the U_3G 3'X mutant. Assignment of the imino-aromatic (top) and imino-imino (bottom) regions of a watergate-NOESY spectrum acquired at 15 ^{o}C with 150 ms mixing time. Residues belonging to the SL1' and SL2' subdomains are labeled in black and orange, respectively. Sequential crosspeaks between imino protons are indicated with solid lines, sequential contacts between A H2 and G H1 protons are marked with dashed lines, and NOE interactions across base pairs are labelled with colons. The U55 and U56 H3 as well as A95 and A96 H2 resonances are doubled due to conformational exchange. Assignments of C amino protons have been omitted for clarity, and crosspeaks marked with crosses are visible at a lower contour level. Conditions: 117 µM RNA, 0 mM NaCl/$MgCl_2$.

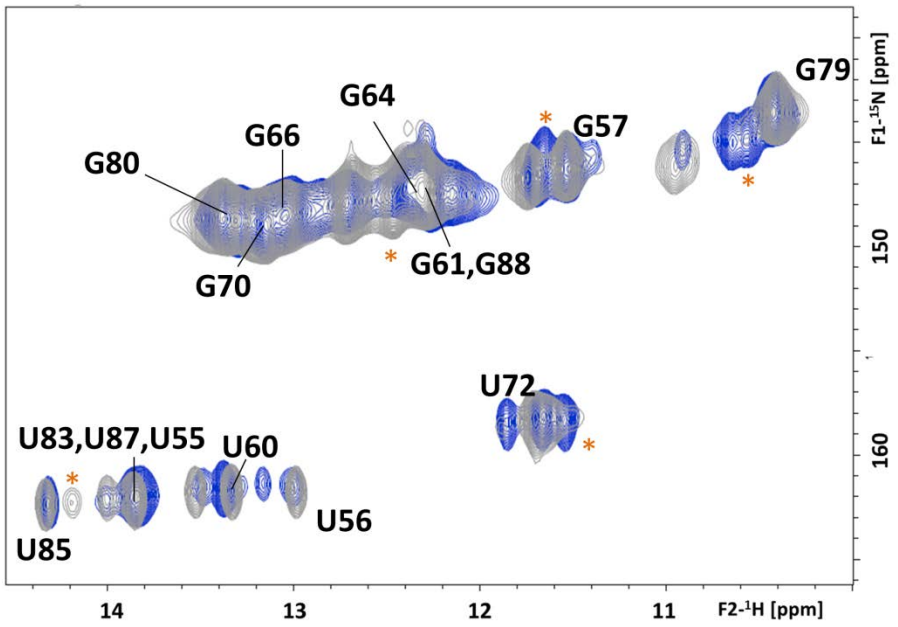

Figure S6. **HSQC analysis of the $U_3G/G_{50}C/C_{52}G$ 3'X mutant.** The ^{1}H-^{15}N HSQC spectrum of the $U_3G/G_{50}C/C_{52}G$ mutant at low ionic strength (black) is superposed with that of the wild-type domain (blue) acquired under the same temperature and ionic conditions. The most important differences between both spectra are marked with red asterisks, and wild-type crosspeaks not matching mutant signals are labelled in blue. Conditions: 64 μM RNA, 0 mM NaCl/$MgCl_2$, 27 ^{o}C.

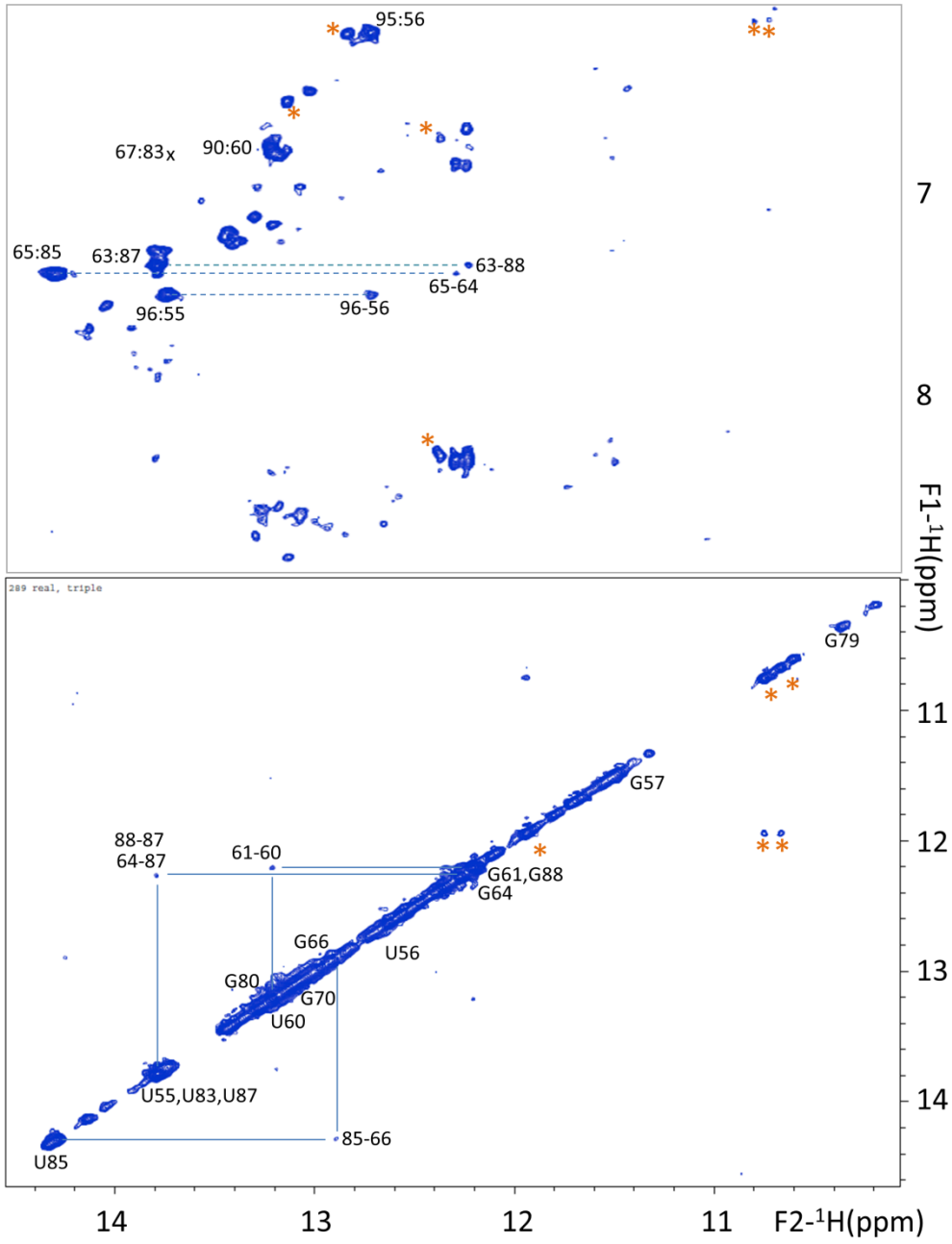

Figure S7. NOESY analysis of the $U_3G/G_{50}C/C_{52}G$ 3'X mutant. Assignment of the imino-aromatic (top) and imino-imino (bottom) regions of a watergate-NOESY spectrum acquired at 15 ^{o}C with 150 ms mixing time. Sequential crosspeaks between imino protons are indicated with solid lines, sequential contacts between A H2 and G H1 protons are marked with dashed lines, and NOE interactions across base pairs are labelled with colons. Assignments of C amino protons have been omitted for clarity, and crosspeaks marked with crosses are visible at a lower contour level. Crosspeaks that do not appear in the NOESY spectra of the wild-type domain or the mutants adopting the two-stem conformation are marked with red asterisks. Conditions: 106 μM RNA, 0 mM $NaCl/MgCl_2$.

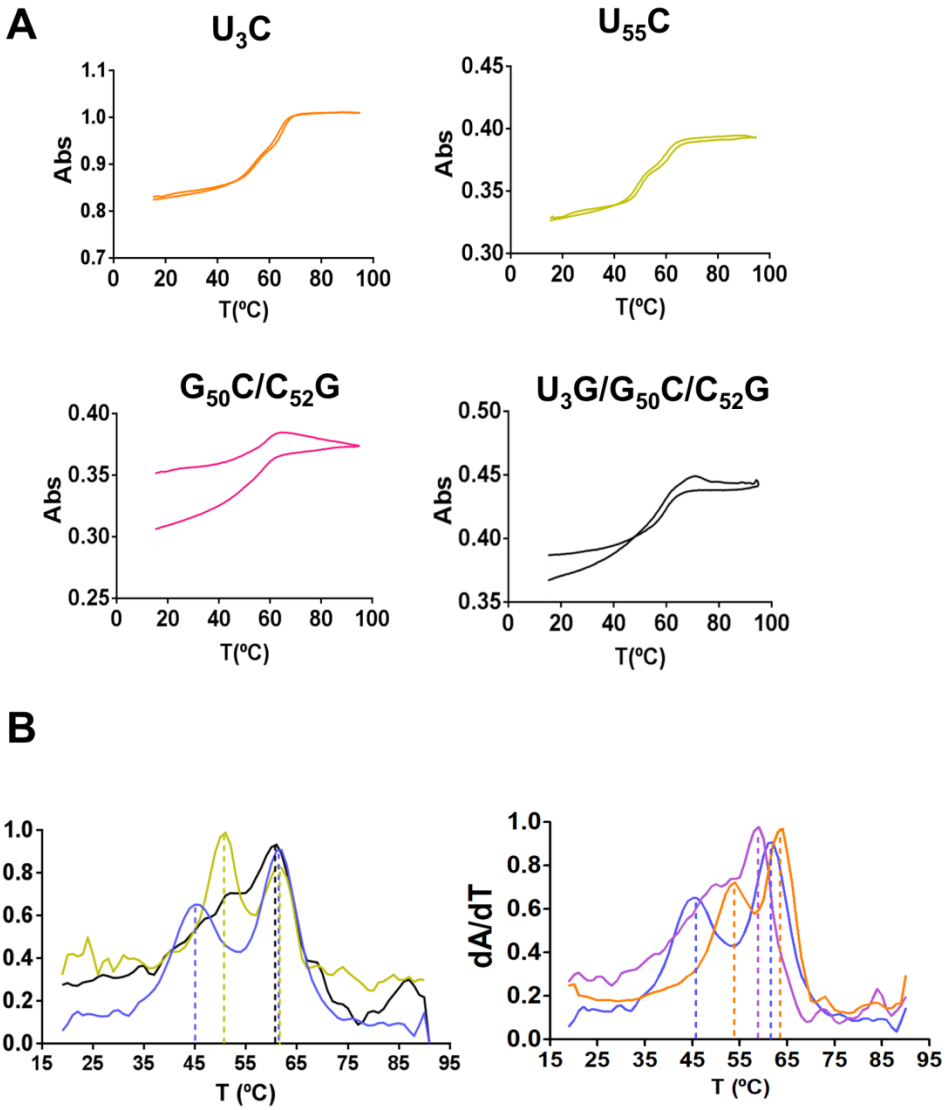

Figure S8. Thermal stability of the U_3C, $U_{55}C$, $G_{50}C/C_{52}G$ and $U_3G/G_{50}C/C_{52}G$ 3'X mutant domains. (**A**) Representative UV-monitored thermal denaturation curves. Regardless of the temperature gradient used for the experiments, the heating and cooling curves did not superpose for mutants $G_{50}C/C_{52}G$ and $U_3G/G_{50}C/C_{52}G$. (**B**) Superposition of the first derivative of absorbance as a function of temperature for wild-type, $U_{55}C$ and $U_3G/G_{50}C/C_{52}G$ domains, respectively represented in blue, green and black colors (left); and wild-type, U_3C and $G_{50}C/C_{52}G$ domains, respectively represented in blue, orange and pink colors (right). Conditions: 10 mM sodium phosphate (pH 6.0), 0.1 mM EDTA and 0 mM $NaCl/MgCl_2$.

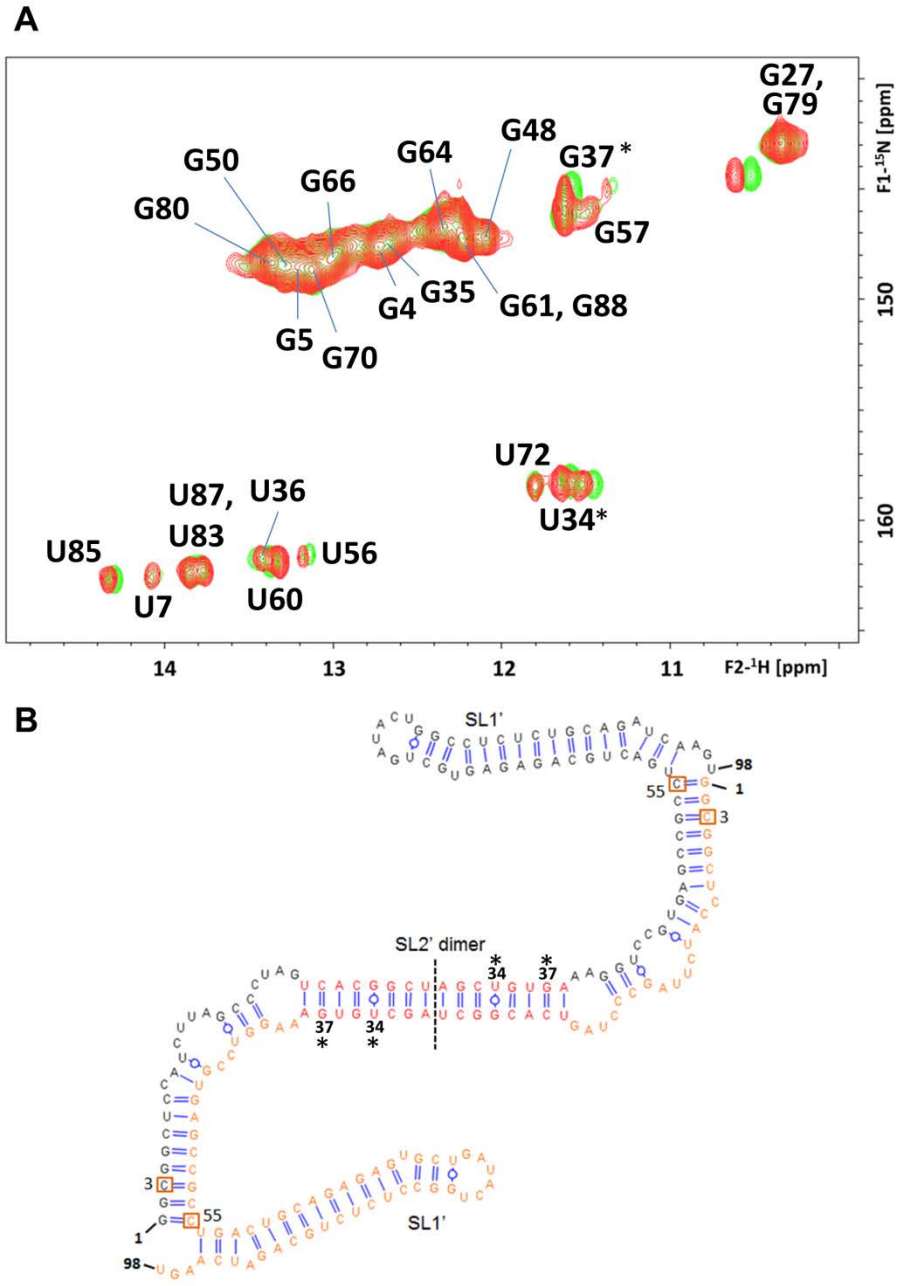

Figure S9. **Impact of NaCl on the structure of the $U_3C/U_{55}C$ 3'X mutant analysed by NMR spectroscopy.** (**A**) Superposition of ^{1}H-^{15}N HSQC spectra of mutant $U_3C/U_{55}C$ acquired in the absence (green) and presence (red) of 100 mM NaCl. (**B**) Secondary structure of the extended homodimer likely formed by $U_3C/U_{55}C$ in conditions of higher ionic strength. The dimer structure comprises two symmetrical halves (separated by the dashed line), and the NMR signals of each of the two halves are equivalent. Assigned nucleotides whose imino protons or imino nitrogens undergo chemical shift variations greater than two standard deviations from the mean perturbation upon the addition of 100 mM NaCl are marked with asterisks ($\Delta\delta \geq 0.05$ and 0.5 ppm for H and N, respectively). Mutations relative to wild-type domain are indicated with red squares. Conditions: 69 μM RNA, 25 °C.

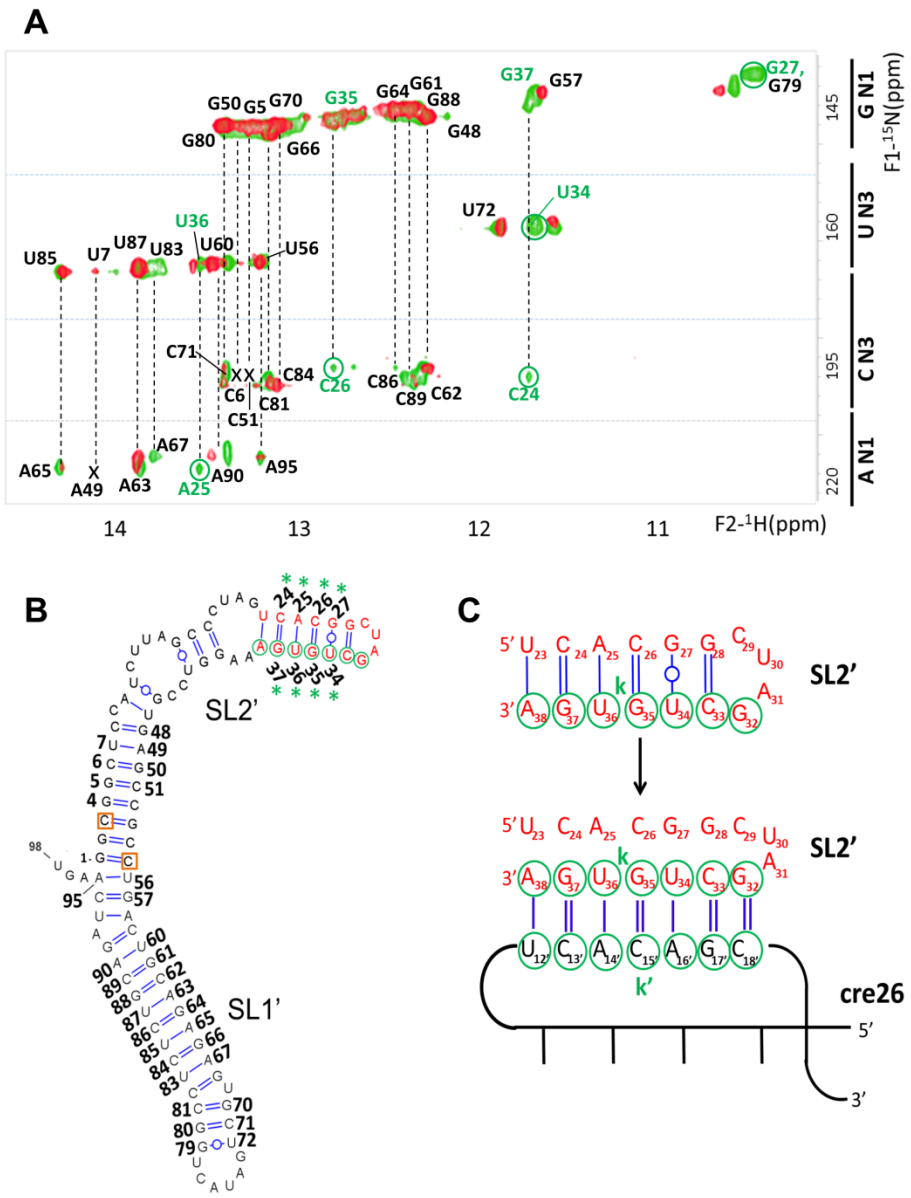

Figure S10. **NMR spectroscopy analysis of the distal interaction formed between the $U_3C/U_{55}C$ 3'X mutant and 5BSL3.2 hairpin cre26.** (**A**) Superposition of ^{1}H-^{15}N HNN-COSY spectra of $U_3C/U_{55}C$, acquired in the absence (green) and presence (red) of one molar equivalent of unlabelled cre26 sequence. Since the experiments used unlabelled cre26, all of the HN and HNN crosspeaks correspond to $U_3C/U_{55}C$. Residues whose HNN crosspeaks disappear or weaken with the addition of cre26 are labelled in green. Crosspeaks marked with crosses are visible at a lower contour level. (**B**) Secondary structure of the $U_3C/U_{55}C$ mutant supported by the NMR data. Residues with assigned HN and HNN resonances are numbered, and those whose HNN-COSY crosspeak patterns undergo changes upon addition of cre26 are marked with green asterisks. (**C**) Schematic representation of the base-pairs formed by SL2' nt in the absence and presence of cre26. In (**B**) and (**C**), DLS motif nt are depicted in red, and *k* and *k'* motif nt are identified with green circles. Conditions: 50 μM RNA, 2 mM $MgCl_2$, 25 ^{o}C.

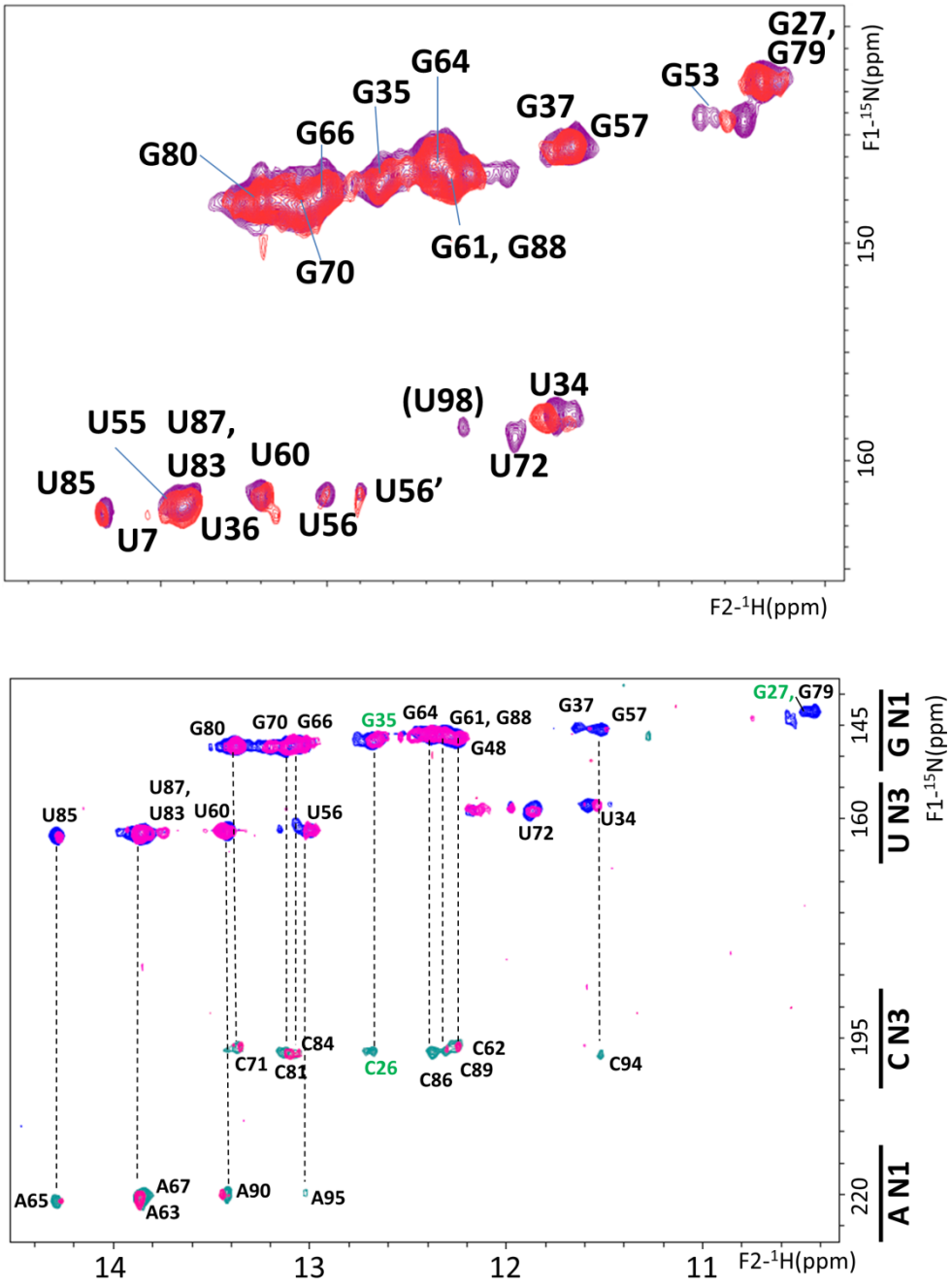

Figure S11. **Impact of NaCl and cre26 on the structure of the U_3G 3'X mutant analysed by NMR spectroscopy.** (**A**) Superposition of ^{1}H-^{15}N HSQC spectra of mutant U_3G acquired in the absence (purple) and presence (red) of 100 mM NaCl. Conditions: 53 μM RNA, 25 ^{o}C. (**B**) Superposition of ^{1}H-^{15}N HNN-COSY spectra of U_3G, acquired in the absence (blue) and presence (purple) of one molar equivalent of unlabelled cre26 sequence. Since the experiments used unlabelled cre26, all of the HN and HNN crosspeaks correspond to U_3G. Residues whose HNN crosspeaks disappear or weaken with the addition of cre26 are labelled in green. Conditions: 45 μM RNA, 2 mM $MgCl_2$, 25 ^{o}C.

2. Supplementary tables

Table S1. List of primers used in this study. *Eco*RI restriction site and T7 RNA polymerase promoter sequences are indicated with lower case and italics letters, respectively. Mutations relative to the wild-type sequence are underlined.

U3C forward	gaattc*TAATACGACTCACTATA*GG**C**GGCTCCATCTTAGCCCTA GTCACGGCTAGC
U3C reverse	GCTAGCCGTGACTAGGGCTAAGATGGAGCC**G**CC*TATAGTGA GTCGTATTA*gaattc
U3G forward	gaattc*TAATACGACTCACTATA*GG**G**GGCTCCATCTTAGCCCTA GTCACGGCTAGC
U3G reverse	GCTAGCCGTGACTAGGGCTAAGATGGAGCC**C**CC*TATAGTGA GTCGTATTA*gaattc
U55C forward	GTGAAAGGTCCGTGAGCCGC**C**TGACTGCAGAGAGTGCTGAT ACTGG
U55C reverse	CCAGTATCAGCACTCTCTGCAGTCA**G**GCGGCTCACGGACCT TTCAC
G50C/C52G forward	CTAGCTGTGAAAGGTCCGTGA**C**C**G**GCTTGACTGCAGAGAGT GCTGATACTGGCCTCTC
G50C/C52G reverse	GAGAGGCCAGTATCAGCACTCTCTGCAGTCAAGC**C**G**G**TCAC GGACCTTTCACAGCTAG
U3C forward (pFK and pGL)	CCTCTTTTTTTCCTTTTCTTTCCTTTGG**C**GGCTCCATCTTAGC CCTAGTCACGGCTAGC
U3C reverse (pFK and pGL)	GCTAGCCGTGACTAGGGCTAAGATGGAGCC**G**CCAAAGGAAA GAAAAGGAAAAAAAGAGG
U3G forward (pFK and pGL)	CCTCTTTTTTTCCTTTTCTTTCCTTTGG**G**GGCTCCATCTTAGC CCTAGTCACGGCTAGC
U3G reverse (pFK and pGL)	GCTAGCCGTGACTAGGGCTAAGATGGAGCC**C**CCAAAGGAAA GAAAAGGAAAAAAAGAGG
C29G-A31U forward (pFK and pGL)	GCTCCATCTTAGCCCTAGTCACGGGTTGCTGTGAAAGGTCC GTGAGCCGCTTGA
C29G-A31U reverse (pFK and pGL)	GTCAAGCGGCTCACGGACCTTTCACAGCAACCCGTGACTAG GGCTAAGATGGAGC
pGL sequencing	GACGATAGTCATGCCCCGCG
pFK sequencing	GCCACCTGACGTCTAAGAAACC