OXFORD UNIVERSITY PRESS | Nucleic Acids Research

Site-specific isotope-labelling of inosine phosphoramidites and NMR analysis of an inosine-containing RNA duplex

Journal:	*Nucleic Acids Research*
Manuscript ID:	Draft
Manuscript Type:	1 Standard Manuscript
Key Words:	RNA editing, inosine, NMR-spectroscopy, phosphoramidite, isotope-labeling

Site-specific isotope-labelling of inosine phosphoramidites and NMR analysis of an inosine-containing RNA duplex

Andre Dallmann[**,1,2,#] *Alexander V.Beribisky*[**1,2], *Felix Gnerlich*[**,3], *Stefan Schiesser*[3], *Thomas Carell*[*,3] *Michael Sattler*[*1,2]

[1] Institute of Structural Biology, Helmholtz Zentrum München, Ingolstädter Landstraβe. 1, 85764 Neuherberg, Germany. E-mail: sattler@helmholtz-muenchen.de, Homepage: http://www.helmholtz-muenchen.de/stb

[2] Center for Integrated Protein Science Munich at Chair Biomolecular NMR, Department Chemie, Technische Universität München, Lichtenbergstraβe. 4, 85747 Garching, Germany, Homepage: http://www.nmr.ch.tum.de

[3] Center for Integrated Protein Science at the Department of Chemistry, Ludwig-Maximilians-Universität München, Butenandtstraβe 5–13, 81377 Munich (Germany), E-mail: thomas.carell@cup.lmu.de Homepage: http://www.carellgroup.de

* To whom correspondence should be addressed: sattler@helmholtz-muenchen.de. Correspondence may also be addressed to: thomas.carell@cup.lmu.de

** These authors wish it to be known that, in their opinion, they should be regarded as joint first authors.

[#] Present Address: Division of Molecular Structure, National Institute of Medical Research, London, NW7 1AA, UK

Keywords: RNA editing, inosine, NMR-spectroscopy, isotope-labeling, phosphoramidite,

ABSTRACT

Adenosine-to-inosine editing is an important mechanism for posttranscriptional gene regulation. The biochemical and structural features of inosine-edited RNAs are poorly characterized. Solution NMR studies of such RNAs requires the use of ^{13}C and ^{15}N isotope-labelling. However, the commonly used isotope-labelling approach based on *in vitro* transcription of the RNA from a DNA template is not applicable for studies of non-native RNA modifications due to the lack of specific coding for the modified nucleotide.

Here, we report the synthesis of an inosine phosphoramidite with selective ^{13}C and ^{15}N-labeling at the C8 and N7 base position, respectively, and with a uniformly ^{13}C-labeled ribose. We demonstrate that the site-specific isotope-labelling reduces signal overlap considerably and provides unambiguous distance information involving inosine residues from a combination of isotope-edited and -filtered NMR spectra. The inosine-containing RNA duplex has decreased stability compared to the corresponding non-edited RNA and exhibits considerable deviations from A-form geometry, with no indication of the formation of base pairs involving the inosine residues. Selective isotope-labelling of phosphoramidites enables NMR studies to specifically focus on the region of interest in the context of a larger RNA, is not restricted to natural nucleobases and can be combined with segmental labelling approaches.

INTRODUCTION

The conversion of adenosine to inosine (A-to-I editing) in RNA involves deamination of the exocylic N6 amino group in adenine to the non-canonical residue inosine (**Supplementary Fig. S1**). This reaction is catalyzed by a specific class of proteins (1) called Adenosine Deaminases Acting on RNA (ADARs). Since inosine is recognized as guanine by the translational machinery (2), A-to-I editing provides a level of post-transcriptional regulation during gene expression, increases transcript diversity and as a consequence, protein structural and functional diversity. A-to-I editing is involved in various cellular events. Inosine is found in various ion channel-encoding mRNAs as well as viral RNAs (3-6). A-to-I editing was also shown to regulate alternative splicing (7) and antagonize another important RNA-based cellular process, RNA interference (8-10).

In one such example, a sequence found in exons 2 and 3 of the rat α-tropomyosin gene has been subjected to editing (11,12). As a result, a hyper-edited motif with four potential IU base pairs is formed with the sequences IIUI and UUIU (**Fig**. **1A**) in the sense and anti-sense strands, respectively (11,12). A 20-nucleotide dimer fragment (**Fig. 1A**) comprising this motif has later been shown to be a binding target, as well as the potential cleavage substrate of the RNA induced silencing complex (RISC) component Tudor-SN (TSN), also known as SND1 or p100 (13,14). TSN is a multifunctional protein implicated in splicing (15), regulation of transcription (16-18) and apoptosis (19) as well as miRNA processing (20). TSN was shown to specifically bind and cleave this inosine-edited RNA duplex (I-RNA, **Fig. 1A**), while showing no binding or cleavage activity towards an RNA duplex where

the inosine residues have been replaced with guanine (G-RNA) (13). By binding and cleaving the I-RNA duplex, TSN is thought to exclude the RNA from further downstream events in the RNAi pathway (13).

In order to investigate the molecular basis for the TSN discrimination towards I-RNA, structural analysis of the latter is of great importance. So far, one duplex structure containing tandem I:U base-pairs was characterized by X-Ray crystallography (21). The authors propose that I:U base-pairs possess properties similar to their G:U counterparts and their presence does not strongly destabilize A-form RNA helices (21). In contrast, other biophysical studies using solution methods show that I:U base pairs have a more severe effect on the thermodynamic stability of A-form RNA than G:U base pairs (22). NMR analysis of I-RNA can thus provide useful information to characterize the structure and conformational dynamics of this RNA in solution and help to address these ambiguities.

RNA structure determination by NMR is complicated by the high degree of signal overlap, especially for the sugar protons (23). This signal overlap can be greatly reduced by preparation of ^{13}C and/or ^{15}N labelled RNA through the use of ^{13}C and/or ^{15}N labelled nucleoside triphosphates (NTPs) as the substrate for RNA synthesis (24,25) by *in vitro* transcription (26,27). For larger RNA (>30 nt), various selective as well as segmental labelling approaches have been used to further simplify NMR spectra. These approaches involve specific labelling of individual NTPs by chemical (28,29) or enzymatic (30-33) synthesis, metabolic pathways (34-36), cleavage, and re-ligation of an unlabelled 5' or 3' RNA fragment to its 3' or 5' labelled counterpart respectively (37-39) and *in vitro* splicing of a labelled nucleotide into an otherwise unlabelled RNA chain or vice versa (40).

Inosine-specific isotope labelling presents itself with a number of unique challenges. *In vitro* transcription with inosine triphosphate in the place of its guanosine triphosphate counterpart has been performed before (11,12). However, due to the presence of both guanine and inosine residues in edited RNAs including the one studied here, site-specific incorporation of inosine residues by *in vitro* transcription is impossible. Post-transcriptional RNA deamination *in vitro,* is extremely challenging as high levels of editing by ADARs to generate the IIUI motif can only be achieved with longer RNA substrates (11,12). In addition to increasing the size of the RNA and further complicating spectral interpretation, such editing is usually not very specific and may result in the modification of other adenine residues in the edited construct (11,41).

Chemical synthesis of the corresponding phosphoramidites and their subsequent incorporation into RNA, offers an attractive alternative for inosine labelling. Moreover, chemical synthesis allows site-specific isotope labelling of specific atoms within the residues of interest (28,42,43). Here we report the synthesis and incorporation of a selectively labelled inosine into a 20-mer RNA duplex that is implicated in posttranscriptional regulation of gene expression and is recognized by the TSN protein (13,14). We demonstrate that this labelling greatly reduces spectral overlap and provides unambiguous structural information for the central IIUI motif in the I-RNA, showing that the four I:U base-pairs adopt a non-standard conformation.

MATERIAL AND METHODS

Chemical synthesis

Chemicals were purchased from *Sigma-Aldrich, Fluka, ABCR* or *Acros organics* and used without further purification. Solutions were concentrated *in vacuo* on a *Heidolph* rotary evaporator. The solvents for organic syntheses were of reagent grade and purified by distillation. Chromatographic purification of products was accomplished using flash column chromatography on *Merck* Geduran Si 60 (40–63 µm) silica gel (normal phase). Thin layer chromatography (TLC) was performed on *Merck* 60 (silica gel F_{254}) plates. Visualization of the developed chromatogram was performed using fluorescence quenching or staining solutions. ^{1}H-, ^{13}C- and ^{15}N-NMR spectra were recorded in deuterated solvents on *Varian VXR400S*, *Varian Inova 400* and *Bruker AMX 600* spectrometers and calibrated to the residual solvent peak. Multiplicities are abbreviated as follows: s = singlet, d = doublet, t = triplet, m = multiplet, br. = broad and combinations thereof. High-resolution ESI spectra were obtained on a *Thermo Finnigan* LTQ FT-ICR mass spectrometer. High-resolution EI spectra were measured on a *MAT CH 7a* (*Varian*). IR measurements were performed on a *Perkin Elmer Spectrum BX FT-IR* spectrometer with a diamond-ATR (Attenuated Total Reflection) setup. The intensities are described as being weak (w), medium (m) or strong (s).

Synthesis of phosphoramidite

The phosphoramidite building block **1** (**Fig. 2**) was synthesized starting from commercially available $^{13}C_6$-glucose **2**, which was converted into the glycosylation precursor **3** in five steps following literature procedure (44). The limited commercial availability of ^{13}C-containing starting material **3** necessitated these elaborate synthetic steps. The ^{13}C/^{15}N labelled hypoxanthine **4** could be synthesized in one step from [5-^{15}N]-5,6-diamino-4(3*H*)-pyrimidone **5** (45) and [^{13}C]-formic acid. Glycosylation using Silyl-Hilbert-Johnson conditions furnished the ester protected nucleoside **6**. Cleavage of the ester groups yielded the free nucleoside **7** with a yield of 92%. In order to selectively protect the 2'-hydroxy group in satisfactory yields, a two-step procedure was necessary. First the 3'- and 5'-hydroxy groups were silyl protected to compound **8**. Then the 2'-hydroxy functionality was TBS-protected with subsequent deprotection of the 3'- and 5'-hydroxy groups to obtain compound **9**. DMTr-protection of the 5'-hydroxy group to compound **10** and introduction of the phosphoramidite functionality lead to building block **1**, which was incorporated into RNA strands using solid phase synthesis.

RNA Solid Phase Synthesis

The syntheses of the oligonucleotides were performed on an ABI 394 DNA/RNA Synthesizer (*Applied Biosystems*) using typical reagent concentrations (activator: 0.10 M Activator 42® in MeCN, detritylation: 3% dichloroacetic acid in CH_2Cl_2, oxidation: 25 mM I_2 in MeCN/H_2O/2,6-lutidine (11/5/1), capping: Ac_2O/2,6-lutidine/MeCN (30 ppm H_2O) (20/30/50) and 20% *N*-methylmidazole in MeCN (10 ppm H_2O). The oligonucleotide syntheses were performed on 200 nmol low-volume polystyrene support using 0.1 M RNA TBS-phosphoramidites: A (Bz-A), C (Bz-C), G (dmf-G), U, obtained from *Link Technologies*. The synthesized [$^{13}C_6$^{15}N]-inosine phosphoramidite was incorporated into RNA

using a standard protocol. The coupling times for the modified bases were increased to 20 min to ensure maximum coupling efficiency. The strands were treated with 1 mL ammonium hydroxide (28%)/methyl amine (40%) (1/1 v/v) for 10 minutes at 60 °C to cleave off the solid support and remove the permanent protecting groups. The sample was centrifuged and the supernatant was collected. The pellet was rinsed with 0.25 mL ddH_2O and the combined solvents were evaporated to dryness using a SpeedVac plus CS110A or SPD 111V from *Savant*. The residue was taken up in 115 µL DMSO, 60 µL triethylamine and 75 µL HF·triethylamine. After incubation for 1.5 h at 65 °C 25 µL NaOAc (3 M) and 1 mL *n*-butanol were added and the tube was cooled to –80 °C for 1 h. The sample was centrifuged again, the supernatant was discarded, and the pellet was rinsed twice with 0.25 mL cold ethanol. Analysis and purification of the oligonucleotides were performed on a *Waters* HPLC system (*Waters Alliance* 2695 with PDA 2996, preparative HPLC: 1525EF with 2482 UV detector) with VP 250/10 Nucleosil 100-7 C 18 columns from *Macherey Nagel* using a gradient of 0.1 M triethylamine/acetic acid in water (buffer A) and in 80% acetonitrile (buffer B). The identity of the strands was determined by MALDI-MS.

MALDI-MS

MALDI spectra were recorded on a *Bruker* autoflex II unit with an *MTP* AnchorChip var/384 target. Prior to the measurements the samples were desalted using *MF-Millipore* membrane filters (0.025 µM). *MALDI matrix: HPA Crown*: 3-hydroxypicolinic acid (50 mg), 15-Crown-5 (10 µL), ammonium hydrogencitrate (10 mg) in 500 µL ddH_2O and 500 µL MeCN.

NMR spectroscopy

A sample of 1.2 mM unlabelled chemically synthesized I-RNA purified and desalted was purchased (*IBA*, Goettingen, Germany) and dissolved in 90% H_2O/10% ^2H_2O or 100% ^2H_2O, respectively. A sample of specifically inosine-$^{13}C/^{15}N$-labeled I-RNA synthesized as described above was dissolved in ^2H_2O at 0.2 mM concentration. Final sample volume for all samples was 250 µl (Shigemi). Samples were heated to 95°C for 5 minutes and then cooled slowly to room temperature to promote dimer formation.

NMR spectra of the unlabelled I-RNA were acquired at 278 K or 298 K in H_2O and ^2H_2O, respectively, on a Bruker Avance I 900 MHz spectrometer equipped with a cryogenic probe, apart from the SOFAST HMQC (46) which was acquired on a Bruker Avance III 750 MHz spectrometer. Spectra were processed with Topspin and NMRPipe (47) and analyzed using NMRVIEW (48). Acquisition parameters are summarized in Table 1. For NOESY experiments in ^2H_2O and H_2O, five spectra each were acquired with 50 ms, 100 ms, 150 ms, 200 ms and 300 ms mixing time, respectively. Two TOCSY spectra were recorded with 40 ms and 80 ms mixing times, respectively. The time-domain data were zero-filled to 2048×512 complex data points, followed by apodization using Lorentz-to-Gauss transformation and cosine functions in t_2 and t_1, respectively, before Fourier transformation.

Spectra of the isotope-labelled I-RNA were acquired at 298 K in ^2H_2O on a Bruker Avance III 800 MHz spectrometer equipped with a cryogenic probe (filtered/edited NOESY, (49,50) or on a Bruker Avance III 600 MHz spectrometer equipped with a cryogenic probe (HCCH-TOCSY/COSY).

Spectra were processed with Topspin and NMRPipe (47) and analyzed using NMRVIEW (48). Acquisition parameters described in Table 2 were used. For the NOESY spectra a mixing time of 100 ms was used. The data were zero filled to 2048×512 complex data points, followed by apodization using Lorentz-to-Gauss transformation and cosine functions in t_2 and t_1, respectively, before Fourier transformation.

RESULTS

NMR analysis of the unlabelled inosine-edited RNA duplex

To characterize the structure of the inosine-edited RNA duplex and potential structural differences induced by A-to-I editing we used standard NMR experiments. A two-dimensional ^{1}H-^{1}H NOESY in ^2H_2O was used to assign the I-RNA regions flanking the inosine-containing motif. Based on aromatic-anomeric contacts (23,51), we found that the regions between A1/U40 to A8/U33 and C13/G28 to G20/C21 assume a standard A-form helical conformation. This was further confirmed by assignments of the aromatic-aromatic as well as anomeric-anomeric region and imino-imino NOEs in long mixing time 2D ^{1}H-^{1}H NOESY in ^2H_2O and H_2O (**Fig 3A** and **Supplementary Figure S2** respectively), taking advantage of the pronounced spin diffusion. Some assignments could be obtained for the inosine containing tetranucleotide motif. The aromatic-anomeric walk was used to unambiguously determine H8 (or H6)/H1' intra- as well as inter-residue correlations for I9, I12, on the sense strand, and U29, I30 on the anti-sense strand (**Fig. 1A**) indicating that at least this part of the inosine-containing motif possesses a geometry resembling A-form RNA.

However, further resonance assignment of the inosine-containing motif, even in the aromatic-anomeric region, proved ambiguous. Since the A-form geometry may be distorted in this part of the RNA, standard cross-peaks expected for this conformation (such as inter-residue H8 (or H6)/H1') may not be present. Due to the large extent of spectral crowding, the presence or the absence of such signals is impossible to ascertain. This is the case with potential I10H8/I9H1', U11H6/I10H1' signals. Also the assignment of some inosine H2 resonances, in particular for I9 and I10, is complicated by signal overlap, which makes it difficult to determine whether these protons give rise to an NOE pattern similar to that of adenine H2 (23) or whether additional signals are also observed.

While partial assignment of the aromatic-anomeric region was achieved, assignment of ribose regions of the NOESY spectrum of the I-RNA is impossible due to severe signal overlap (**Fig. 3A**). Of particular importance is the assignment of the ribose resonances of the residues of the inosine-containing motif, as this region of the I-RNA may not adopt a standard A-form helix. To address this issue, we developed a site-specific isotope-labeling scheme for the inosine residues for NMR analysis of the I-RNA.

Choice of isotope-labelling scheme and

The I-RNA labelling scheme was designed with the purpose to provide as much information as possible on the conformation of the inosine residues and their NOE contacts with neighbouring residues and base pairs. While the ribose signals tend to overlap, base signals are better dispersed and provide the most useful information about the duplex conformation. Missing assignments for both

I10 H2 and H8 in the unlabelled I-RNA spectra suggested that their NMR signals might overlap. Hence, we decided to specifically label the C8 and not the C2 position in order to be able to distinguish between the two. In addition to C8, N7 of the inosine base and all ribose carbons were ^{13}C and ^{15}N labelled (**Fig. 1B**). While C8 labelling is useful to resolve NOEs to report on *anti* or *syn* conformation of the base, N7 labelling can be used to check for potential non-Watson-Crick hydrogen bonding interactions such as Hoogsteen base-pairing (52,53). The inosine phosphoramidite was synthesized using the scheme described in Materials and Methods with a yield of 92%.

Inosine-labelling simplifies assignments of the I-RNA duplex

The availability of ^{13}C isotopes enables the use of a ^{1}H, ^{13}C edited NOESY-HSQC experiment to specifically select for inosine correlations as only NOE cross-peaks involving the ^{13}C-labeled ribose or the base H8 protons of inosine (**Fig. 1B**) to any other proton close in space are observed (either bound to ^{13}C or ^{12}C). For the inosine-labelled I-RNA this greatly simplifies spectral assignments of residues in the central IIUI motif and adjacent base pairs by significantly reducing spectral overlap (**Fig. 3B**).

The chemical shift of the H8 proton of I10, which is obscured due to signal overlap in the unlabelled I-RNA NOESY spectrum (**Fig. 3A**) could be unambiguously determined using the NOESY-HSQC experiment (**Fig. 3B**)**.** Moreover, the lack of an inter-residue I10H8/I9H1' cross-peak conclusively demonstrates a clear deviation from A-form geometry in this region of the I-RNA duplex. From the homonuclear NOESY spectrum, this conclusion could not be drawn due to severe spectral overlap. The NOESY-HSQC also allows to conclusively determine the chemical shifts of the inosine H2 resonances. In the homonuclear NOESY experiment, I9 H2 and I10 H2 are overlapped with other signals (**Fig. 3A**), while the I12 H2 and I30 H2 signals are of very weak intensity. The edited NOESY-HSQC unambiguously resolves the location of these three peaks (**Fig. 3B**). All four inosine H2 resonances display an NOE pattern similar to that of the adenine H2 counterparts: One cross-peak to the H1' of the residue located 3' of the inosine in question, and another to the H1' of the residue located 3' of the inosine's counter-base (**Fig. 3B**). In addition, I30 H2 also displays a signal to the H1' of I10. This NOE would correspond to an inter-atom distance which is not in line with A-form helix geometry, further confirming that I10 adopts a non-standard conformation.

Isotope edited and filtered NOESY experiments resolve signal overlap

The ^{1}H, ^{13}C edited NOESY-HSQC (**Fig. 3B, Fig. 4A, left strip**) experiment alone still does not allow for complete I-RNA assignments. The aromatic/ribose region of this RNA still gives rise to a very large degree of spectral crowding, where, for example overlap of the I10 H2 and H8 resonances prohibits unambiguous assignments. These problems can be overcome with the help of isotope-filtered NOESY experiments. Different combinations of isotope-editing and/or filtering methods in the two frequency dimensions of the NOESY experiment enable the separation of NOE cross peaks between ^{13}C and/or ^{12}C bound protons (**Fig. 1C**).

For example, the ω_1-edited, ω_2-filtered NOESY shows cross-peaks only between a proton bound to ^{13}C in ω_1 and another proton bound to a ^{12}C in ω_2. This allows the detection of an NOE cross-peak between inosine ribose or H8 protons and inosine H2 or any non-inosine protons close in

space (**Fig. 3D and Fig. 4A, center strip**). These signals are mostly inter-residue cross-peaks between inosine ribose protons and various aromatic protons of uracil residues located next to the inosine in question, or protons of residues flanking the motif, which also located next to the labelled inosine moieties. Thus, the edited-filtered NOESY experiment provides unambiguous assignments of the C13H6/I12H2', U31H6/I30H1', A8H2 I9H2' cross-peaks.

The ω_1-edited, ω_2-edited NOESY (**Fig. 3C and Fig. 4A, right strip**), on the other hand gives rise to an NOE between two protons only if both of them are ^{13}C bound. This experiment facilitated the assignment of inosine intra-residue as well as I9 and I10 inter-residue cross-peaks, as these two inosine moieties are located next to each other in the RNA. Using this experiment, cross-peaks such as I10H8/I10H2', I10H8/I9H3', I30H8/I30H3' and I12H8/I12H2' were successfully assigned.

These experiments have also allowed assignment of NMR signals with completely degenerate chemical shifts. For example, I10H8/I10H1' and I10H2/I10H1' cross peaks could be resolved as the I10 H8 proton is ^{13}C-bound, while the I10 H2 proton is not. As a result the first cross-peak appears in the edited-edited NOESY spectrum (**Fig. 4A, right strip**), while the other was observed in its edited-filtered NOESY counterpart (**Fig. 4A, center strip**). Similarly, the completely overlapping, I10H8/I10H2', I10H2/I10H2' and I10H8/I9H2' I10H2/I9H2' correlations can be assigned (**Fig. 4B**). This demonstrates the great utility of the site-specific inosine isotope-labelling combined with isotope-edited and -filtered NMR experiments. Using a combination of these experiments almost all cross-peaks in the aromatic ribose (**Fig. 3A**) and the anomeric/ribose regions (data not shown) could be assigned.

Analysis of inosine base chemical shifts

The specific ^{15}N-labelling of the N7 in inosine residues can be exploited to probe for possible Hoogsteen base pairing of the inosine in the I-RNA duplex (**Fig. 5A**). A distinctive upfield ^{15}N chemical shift of N7 is indicative of Hoogsteen base pairing (53). The absence of observable imino proton signals for the inosine residues in the SOFAST-HMQC (**Fig. 5B**) indicates the lack of stable base pairing interactions. Moreover, the N7 chemical shifts observed in the long-range ^{1}H-^{15}N-HSQC experiment (**Fig. 5C**) demonstrate that the inosine N7 nuclei are not involved in hydrogen bond formation. Additionally the line-broadening observed for NMR signals of the inosine residues flanking the IIUI motif (I9 and I12) compared to the central ones (I10 and I30) suggests that the former are more tightly integrated into the A-form helical stack than the central ones. Nevertheless, I10 and I30 and the uracil residues of the opposite strand (U31 and U11) continue stacking interactions in the RNA duplex region. This can be concluded from the lack of intra-residue H8/H1' or intense H6/H5 NMR signals in the $^{2}H_2O$ NOESY experiment as well as the significant number of inter-residue NOEs that are observed for these two inosine residues.

DISCUSSION

I-RNA is a 20-mer inosine-containing RNA duplex and a known substrate for the RISC component TSN (13,14). To understand the molecular basis for TSN discrimination towards I-RNA, a structural investigation of this RNA is required. Due to highly destablilizing effect of I:U base pairs on

RNA duplex, solution state methods such as NMR spectroscopy are preferable to study not only the structure but also characterize the dynamical features of the I-RNA. Due to the molecular weight of the I-RNA duplex isotopic labelling is required for high resolution structure determination. Enzymatic synthesis of either uniform or selectively labelled $^{13}C/^{15}N$ inosine monophosphate which is a precursor for the generation of guanine and adenine monophosphates has been described by Schultheisz et al (32). However, enzymatic, residue-specific synthesis of $^{13}C/\ ^{15}N$ labelled I-RNA RNA strands is impossible as both inosine and guanosine form base pairs with cytosines in the template strand. Therefore chemical synthesis is the only viable alternative to produce isotopically labelled I-RNA.

Chemical synthesis of isotopically labelled ribophosphoramidites has been described previously. In one such case, ^{13}C labelled ribose moieties were produced using ^{13}C labelled glucose as starting material (28). The appropriate unlabelled base was then attached to this sugar, and the resultant ribophosphoramidite was incorporated into the RNA at various specific positions (28). This kind of ribose labelling technique was shown to simplify NMR spectral assignments of poly CU and S6/S7 RNA constructs both in their free form, as well as when bound to their protein binding partners – polypyrimidine tract binding protein and feminizing locus on X, respectively (28). In other work, synthetic incorporation of a single ^{13}C label at the C6 position of uridine and cytidine has also been outlined (42). A ^{13}C labelled methyl group was also successfully added to the 2'OH of a uridine residue to study the conformational dynamics of various RNA sequences (43). However in neither of these studies, inosine labelling was performed.

The presence of $^{13}C,^{15}N$ labelled inosine residues in I-RNA greatly simplifies the assignment of the IIUI motif of this RNA. With the use of the 2D ^{1}H, ^{13}C NOESY-HSQC, complete assignment of the inosine H2 and H8 regions was achieved. The lack of an I10H8/I9H1' and the presence of non-canonical I30H2/I10H1' signals demonstrates that the inner portion of the I-RNA deviates from ideal A-form. More intricate, editing-editing as well as editing-filtering NOESY experiments reduce spectral overlap even further and facilitate unambiguous spectral assignment of various ribose-aromatic moieties.

It has been postulated before that TSN recognition of the I-RNA occurs due to its distorted helical structure (14). Our data here provide support for this hypothesis as NMR chemical shift assignments of the inosine containing motif of the I-RNA indicate that this region indeed adopts a non-canonical helical conformation.

In conclusion, an isotopically labelled, inosine containing RNA duplex with selective isotopic labels was successfully synthesized and analysed by NMR spectroscopy. This analysis has revealed a clear deviation from A-form geometry for the inosine containing portion of this RNA. Further structural and functional studies need to be conducted to gain additional insight into the exact nature of this deviation, as well its role in TSN recognition.

SUPPLEMENTARY DATA

Supplementary data are available on-line.

ACKNOWLEDGEMENTS

We thank Gerd Gemmecker for assistance with NMR measurements as well as members of the Sattler group for numerous productive discussions.

FUNDING

This work was supported by the *Deutsche Forschungsgemeinschaft* [SFB1035, GRK1721] to M.S., [SFB 1032, SFB749] and Volkswagen Foundation to T.C., the *Bayerisches Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst* to M.S., and the *Fonds der Chemischen Industrie* for a pre-doctoral fellowship to S.S.

REFERENCES

1. Bass, B.L. and Weintraub, H. (1988) An unwinding activity that covalently modifies its double-stranded RNA substrate. *Cell*, **55**, 1089-1098.
2. Basilio, C., Wahba, A.J., Lengyel, P., Speyer, J.F. and Ochoa, S. (1962) Synthetic polynucleotides and the amino acid code. V. *Proceedings of the National Academy of Sciences of the United States of America*, **48**, 613-616.
3. Burns, C.M., Chu, H., Rueter, S.M., Hutchinson, L.K., Canton, H., Sanders-Bush, E. and Emeson, R.B. (1997) Regulation of serotonin-2C receptor G-protein coupling by RNA editing. *Nature*, **387**, 303-308.
4. Bhalla, T., Rosenthal, J.J., Holmgren, M. and Reenan, R. (2004) Control of human potassium channel inactivation by editing of a small mRNA hairpin. *Nature structural & molecular biology*, **11**, 950-956.
5. Bass, B.L., Weintraub, H., Cattaneo, R. and Billeter, M.A. (1989) Biased hypermutation of viral RNA genomes could be due to unwinding/modification of double-stranded RNA. *Cell*, **56**, 331.
6. Polson, A.G., Bass, B.L. and Casey, J.L. (1996) RNA editing of hepatitis delta virus antigenome by dsRNA-adenosine deaminase. *Nature*, **380**, 454-456.
7. Rueter, S.M., Dawson, T.R. and Emeson, R.B. (1999) Regulation of alternative splicing by RNA editing. *Nature*, **399**, 75-80.
8. Scadden, A.D. and Smith, C.W. (2001) RNAi is antagonized by A-->I hyper-editing. *EMBO reports*, **2**, 1107-1111.
9. Yang, W., Wang, Q., Howell, K.L., Lee, J.T., Cho, D.S., Murray, J.M. and Nishikura, K. (2005) ADAR1 RNA deaminase limits short interfering RNA efficacy in mammalian cells. *The Journal of biological chemistry*, **280**, 3946-3953.
10. Kawahara, Y., Zinshteyn, B., Chendrimada, T.P., Shiekhattar, R. and Nishikura, K. (2007) RNA editing of the microRNA-151 precursor blocks cleavage by the Dicer-TRBP complex. *EMBO reports*, **8**, 763-769.
11. Scadden, A.D. and Smith, C.W. (2001) Specific cleavage of hyper-edited dsRNAs. *The EMBO journal*, **20**, 4243-4252.
12. Scadden, A.D. and Smith, C.W. (1997) A ribonuclease specific for inosine-containing RNA: a potential role in antiviral defence? *The EMBO journal*, **16**, 2140-2149.
13. Scadden, A.D. (2005) The RISC subunit Tudor-SN binds to hyper-edited double-stranded RNA and promotes its cleavage. *Nature structural & molecular biology*, **12**, 489-496.
14. Li, C.L., Yang, W.Z., Chen, Y.P. and Yuan, H.S. (2008) Structural and functional insights into human Tudor-SN, a key component linking RNA interference and editing. *Nucleic acids research*, **36**, 3579-3589.

15. Yang, J., Valineva, T., Hong, J., Bu, T., Yao, Z., Jensen, O.N., Frilander, M.J. and Silvennoinen, O. (2007) Transcriptional co-activator protein p100 interacts with snRNP proteins and facilitates the assembly of the spliceosome. *Nucleic acids research*, **35**, 4485-4494.
16. Tong, X., Drapkin, R., Reinberg, D. and Kieff, E. (1995) The 62- and 80-kDa subunits of transcription factor IIH mediate the interaction with Epstein-Barr virus nuclear protein 2. *Proceedings of the National Academy of Sciences of the United States of America*, **92**, 3259-3263.
17. Yang, J., Aittomaki, S., Pesu, M., Carter, K., Saarinen, J., Kalkkinen, N., Kieff, E. and Silvennoinen, O. (2002) Identification of p100 as a coactivator for STAT6 that bridges STAT6 with RNA polymerase II. *The EMBO journal*, **21**, 4950-4958.
18. Paukku, K., Yang, J. and Silvennoinen, O. (2003) Tudor and nuclease-like domains containing protein p100 function as coactivators for signal transducer and activator of transcription 5. *Molecular endocrinology*, **17**, 1805-1814.
19. Sundstrom, J.F., Vaculova, A., Smertenko, A.P., Savenkov, E.I., Golovko, A., Minina, E., Tiwari, B.S., Rodriguez-Nieto, S., Zamyatnin, A.A., Jr., Valineva, T. *et al.* (2009) Tudor staphylococcal nuclease is an evolutionarily conserved component of the programmed cell death degradome. *Nature cell biology*, **11**, 1347-1354.
20. Heinrich, E.M., Wagner, J., Kruger, M., John, D., Uchida, S., Weigand, J.E., Suess, B. and Dimmeler, S. (2013) Regulation of miR-17-92a cluster processing by the microRNA binding protein SND1. *FEBS letters*, **587**, 2405-2411.
21. Pan, B., Mitra, S.N., Sun, L., Hart, D. and Sundaralingam, M. (1998) Crystal structure of an RNA octamer duplex r(CCCIUGGG)2 incorporating tandem I.U wobbles. *Nucleic acids research*, **26**, 5699-5706.
22. Serra, M.J., Smolter, P.E. and Westhof, E. (2004) Pronounced instability of tandem IU base pairs in RNA. *Nucleic acids research*, **32**, 1824-1828.
23. Fürtig, B., Richter, C., Wöhnert, J. and Schwalbe, H. (2003) NMR spectroscopy of RNA. *Chembiochem*, **4**, 936-962.
24. Nikonowicz, E.P., Sirr, A., Legault, P., Jucker, F.M., Baer, L.M. and Pardi, A. (1992) Preparation of 13C and 15N labelled RNAs for heteronuclear multi-dimensional NMR studies. *Nucleic Acids Res*, **20**, 4507-4513.
25. Batey, R.T., Inada, M., Kujawinski, E., Puglisi, J.D. and Williamson, J.R. (1992) Preparation of isotopically labeled ribonucleotides for multidimensional NMR spectroscopy of RNA. *Nucleic Acids Res*, **20**, 4515-4523.
26. Milligan, J.F., Groebe, D.R., Witherell, G.W. and Uhlenbeck, O.C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic acids research*, **15**, 8783-8798.
27. Milligan, J.F. and Uhlenbeck, O.C. (1989) Synthesis of small RNAs using T7 RNA polymerase. *Methods in enzymology*, **180**, 51-62.
28. Wenter, P., Reymond, L., Auweter, S.D., Allain, F.H. and Pitsch, S. (2006) Short, synthetic and selectively 13C-labeled RNA sequences for the NMR structure determination of protein-RNA complexes. *Nucleic acids research*, **34**, e79.
29. Jiang, F., Patel, D.J., Zhang, X., Zhao, H. and Jones, R.A. (1997) Specific labeling approaches to guanine and adenine imino and amino proton assignments in the AMP-RNA aptamer complex. *Journal of biomolecular NMR*, **9**, 55-62.
30. Vallurupalli, P., Scott, L., Hennig, M., Williamson, J.R. and Kay, L.E. (2006) New RNA labeling methods offer dramatic sensitivity enhancements in 2H NMR relaxation spectra. *Journal of the American Chemical Society*, **128**, 9346-9347.
31. Schultheisz, H.L., Szymczyna, B.R., Scott, L.G. and Williamson, J.R. (2011) Enzymatic de novo pyrimidine nucleotide synthesis. *Journal of the American Chemical Society*, **133**, 297-304.
32. Schultheisz, H.L., Szymczyna, B.R., Scott, L.G. and Williamson, J.R. (2008) Pathway engineered enzymatic de novo purine nucleotide synthesis. *ACS chemical biology*, **3**, 499-511.

33. Scott, L.G., Tolbert, T.J. and Williamson, J.R. (2000) Preparation of specifically ^{2}H- and ^{13}C-labeled ribonucleotides. *Methods Enzymol.*, **317**, 18-38.
34. Thakur, C.S., Luo, Y., Chen, B., Eldho, N.V. and Dayie, T.K. (2012) Biomass production of site selective 13C/15N nucleotides using wild type and a transketolase E. coli mutant for labeling RNA for high resolution NMR. *Journal of biomolecular NMR*, **52**, 103-114.
35. Thakur, C.S., Sama, J.N., Jackson, M.E., Chen, B. and Dayie, T.K. (2010) Selective 13C labeling of nucleotides for large RNA NMR spectroscopy using an E. coli strain disabled in the TCA cycle. *Journal of biomolecular NMR*, **48**, 179-192.
36. Johnson, J.E., Jr., Julien, K.R. and Hoogstraten, C.G. (2006) Alternate-site isotopic labeling of ribonucleotides for NMR studies of ribose conformational dynamics in RNA. *Journal of biomolecular NMR*, **35**, 261-274.
37. Duss, O., Maris, C., von Schroetter, C. and Allain, F.H. (2010) A fast, efficient and sequence-independent method for flexible multiple segmental isotope labeling of RNA using ribozyme and RNase H cleavage. *Nucleic acids research*, **38**, e188.
38. Tzakos, A.G., Easton, L.E. and Lukavsky, P.J. (2007) Preparation of large RNA oligonucleotides with complementary isotope-labeled segments for NMR structural studies. *Nature protocols*, **2**, 2139-2147.
39. Nelissen, F.H., van Gammeren, A.J., Tessari, M., Girard, F.C., Heus, H.A. and Wijmenga, S.S. (2008) Multiple segmental and selective isotope labeling of large RNA for NMR structural studies. *Nucleic acids research*, **36**, e89.
40. Kawahara, I., Haruta, K., Ashihara, Y., Yamanaka, D., Kuriyama, M., Toki, N., Kondo, Y., Teruya, K., Ishikawa, J., Furuta, H. *et al.* (2012) Site-specific isotope labeling of long RNA for structural and mechanistic studies. *Nucleic acids research*, **40**, e7.
41. Nishikura, K., Yoo, C., Kim, U., Murray, J.M., Estes, P.A., Cash, F.E. and Liebhaber, S.A. (1991) Substrate specificity of the dsRNA unwinding/modifying activity. *The EMBO journal*, **10**, 3523-3532.
42. Wunderlich, C.H., Spitzer, R., Santner, T., Fauster, K., Tollinger, M. and Kreutz, C. (2012) Synthesis of (6-(13)C)pyrimidine nucleotides as spin-labels for RNA dynamics. *Journal of the American Chemical Society*, **134**, 7558-7569.
43. Kloiber, K., Spitzer, R., Tollinger, M., Konrat, R. and Kreutz, C. (2011) Probing RNA dynamics via longitudinal exchange and CPMG relaxation dispersion NMR spectroscopy using a sensitive 13C-methyl label. *Nucleic acids research*, **39**, 4340-4351.
44. Saito, Y., Zevaco, T.A. and Agrofoglio, L.A. (2002) Chemical synthesis of ^{13}C labeled anti-HIV nucleosides as mass-internal standards. *Tetrahedron*, **58**, 9593-9603.
45. Pagano, A.R., Lajewski, W.M. and Jones, R.A. (1995) Syntheses of [6,7-15N]-Adenosine, [6,7-15N]-2'-Deoxyadenosine, and [7-15N]-Hypoxanthine. *Journal of the American Chemical Society*, **117**, 11669-11672.
46. Farjon, J., Boisbouvier, J., Schanda, P., Pardi, A., Simorre, J.P. and Brutscher, B. (2009) Longitudinal-relaxation-enhanced NMR experiments for the study of nucleic acids in solution. *Journal of the American Chemical Society*, **131**, 8571-8577.
47. Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of biomolecular NMR*, **6**, 277-293.
48. Johnson, B.A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods in molecular biology*, **278**, 313-352.
49. Zwahlen, C., Legault, P., Vincent, S.J.F., Greenblatt, J., Konrat, R. and Kay, L.E. (1997) Methods for Measurement of Intermolecular NOEs by Multinuclear NMR Spectroscopy: Application to a Bacteriophage lambda-N-Peptide/boxB RNA Complex. *Journal of the American Chemical Society*, **119**, 6711--6721.

50. Sattler, M., Schleucher, J. and Griesinger, C. (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. NMR Spectrosc.*, **34**, 93-158.
51. Varani, G., Aboul-ela, F. and Allain, F.H.-T. (1996) NMR investigation of RNA structure. *Progr. NMR Spectrosc.*, **29**, 51-127.
52. Nikolova, E.N., Gottardo, F.L. and Al-Hashimi, H.M. (2012) Probing transient Hoogsteen hydrogen bonds in canonical duplex DNA using NMR relaxation dispersion and single-atom substitution. *Journal of the American Chemical Society*, **134**, 3667-3670.
53. Gaffney, B.L., Kung, P.-P., Wang, C. and Jones, R.A. (1995) Nitrogen-15-labeled oligodeoxynucleotides. 8. Use of 15N NMR to probe Hoogsteen hydrogen bonding at guanine and adenine N7 atoms of a DNA triplex. *Journal of the American Chemical Society*, **117**, 12281-12283.

Tables

Table 1: NMR acquisition parameters for the unlabelled I-RNA

Spectrum	Time domain size (F2/F1)	Acquisition times F2/ F1 (ms)	Sweep width F2/F1 (kHz)	Repetition delay (s)	No. of scans	Total experiment time (h)
NOESY $^{2}H_2O$	2048/512	113.9/28.5	8.99/8.99	2	16	5.0
COSY $^{2}H_2O$	2048/256	67.0/8.4	15.29/15.29	1	32	2.5
TOCSY $^{2}H_2O$	2048/256	67.0/8.4	15.29/15.29	1	16	1.25
HSQC $^{2}H_2O$	8192/512	525.9/16.2	7.79/15.85	1	96	21
NOESY H_2O	2048/512	75.8/18.9	13.52/13.52	2	16	5.5
SF-HMQC H_2O	1642/128	49.9/28.1	16.45/2.28	0.2	6144	60

Table 2: NMR acquisition parameters for the specifically isotope-labelled I-RNA

Spectrum	Time domain size (F3/)F2/F1	Acqu. times (F3/)F2/ F1 (ms)	SW (F3/)F2/F1 (kHz)	Repetition delay (s)	No. of scans	Total Exp. time (h)
HCCHCOSY	1536/48/176	159.7/39.8/18.3	4.81/6.04/4.81	1	32	93
HCCHTOCSY	1536/48/176	159.7/39.8/18.3	4.81/6.04/4.81	1	32	93
HSQC	2048/256	160.2/6.4	6.39/20.12	1	16	1
Ed. NOESY	1536/256	120.1/20.0	6.40/6.40	1	256	44
Ed-Ed NOE	1536/256	120.1/20.0	6.40/6.40	1	256	44
Ed-Fi NOE	1536/256	120.1/20.0	6.40/6.40	1	256	44

FIGURE LEGENDS

Figure 1: A) RNA sequence with inosine residues highlighted in green. **B**) Site-selective isotope-labeling scheme and atomic numbering for inosine. Positions in red and blue are isotopically ^{13}C and ^{15}N labelled, respectively **C**) Schematic representation of the spectral editing procedure, showing which signals are expected depending on the type of spectral editing.

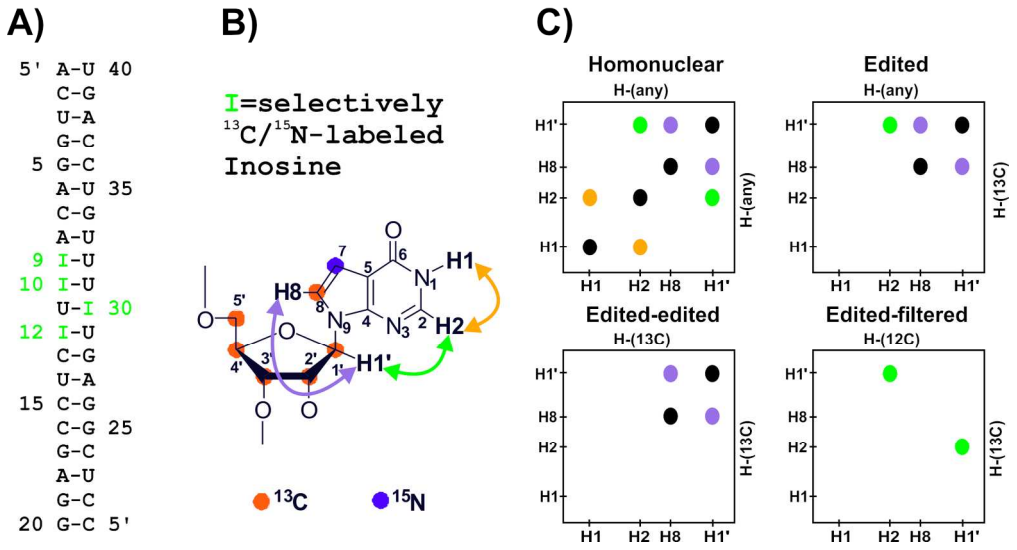
Figure 2: Synthesis of the [$^{13}C_6{}^{15}N$]-Inosine-phosphoramidite building block **1**: Reagents and conditions: **A**) 1. hypoxanthine **4**, HMDS, 2. TMSOTf, DCE, 92%; **B**) NH_3, MeOH, 92%; **C**) 1. $^{t}Bu_2Si(OTf)_2$, imidazole, 2. TBSCl, DMF, 88%; **D**) HF·pyridine, 76%; **E**) DMTrCl, pyridine, 65%; **F**) CEDCl, DIPEA, DCM, 97%; **G**) [^{13}C]-formic acid, H_2SO_4, H_2O, 59%. DCE = 1,2-dichloroethane, DIPEA = *N*,*N*-di*iso*propylethylamine, CEDCl = 2-cyanoethyl-*N,N*-di*iso*propylchlorophosphoramidite.

Figure 3: A) Aromatic-anomeric/sugar proton region of the homonuclear NOESY spectrum. Severe signal overlap especially in the aromatic-sugar region renders unambiguous chemical shift assignment impossible and necessitates the use of editing and filtering approaches. **B**) ω_1-edited, **C**) ω_1-edited ω_2-edited D) ω_1-edited, ω_2-filtered NOESY spectra of I-RNA. Note: All spectra were recorded in $^{2}H_2O$. Cross peaks discussed in the text are indicated by "*".

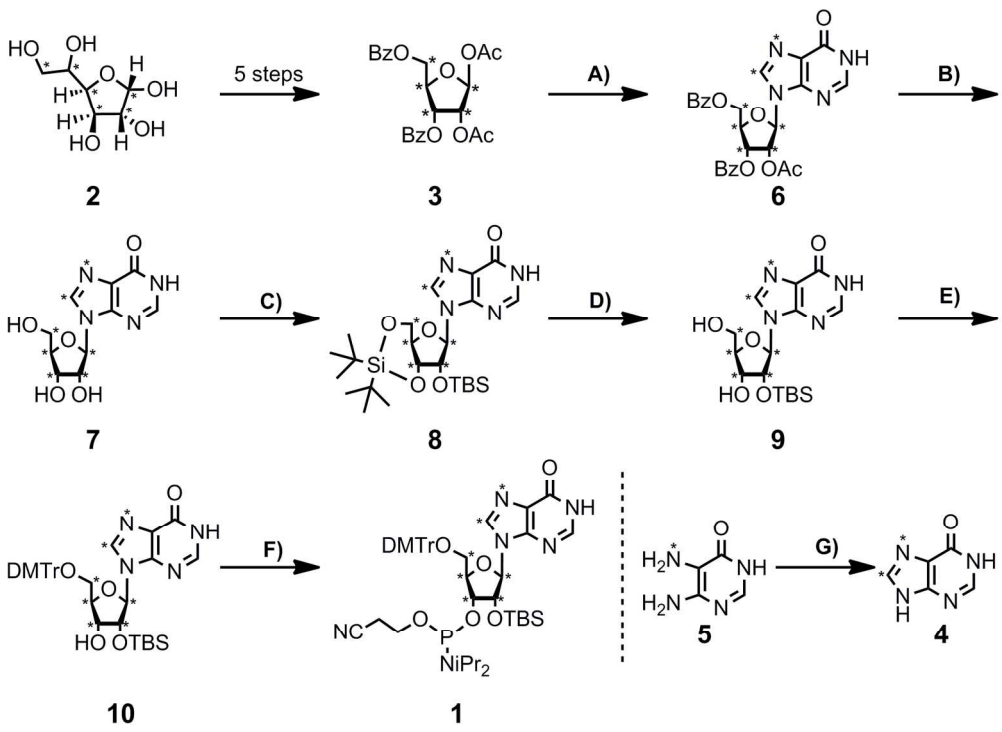
Figure 4: A) Strip plots from edited (black), edited-filtered (red) and edited-edited (blue) NOESY spectra for I-RNA. The two latter spectra are complementary and thus help to resolve overlap especially in the severely overlapped aromatic/sugar region as can be seen with the selected peaks. Chemical shift assignments are indicated. **B**) The distances corresponding to the cross peaks seen in (A) are indicated in a canonical RNA model duplex.

Figure 5: A) Secondary structure information for the inosine duplex RNA from **B**) a natural abundance ^{1}H-^{15}N Imino-sfHMQC, showing that the imino protons for the central IIUI and the adjacent A:U base pair are not observed and **C**) a long-range ^{1}H-^{15}N HSQC spectrum of the labelled inosine RNA, excluding the possibility of Hoogsteen base pair formation due to the absence of upfield N7 chemical shifts (53). Signals indicated by an asterisk "*" originate from an impurity.

Figure 1

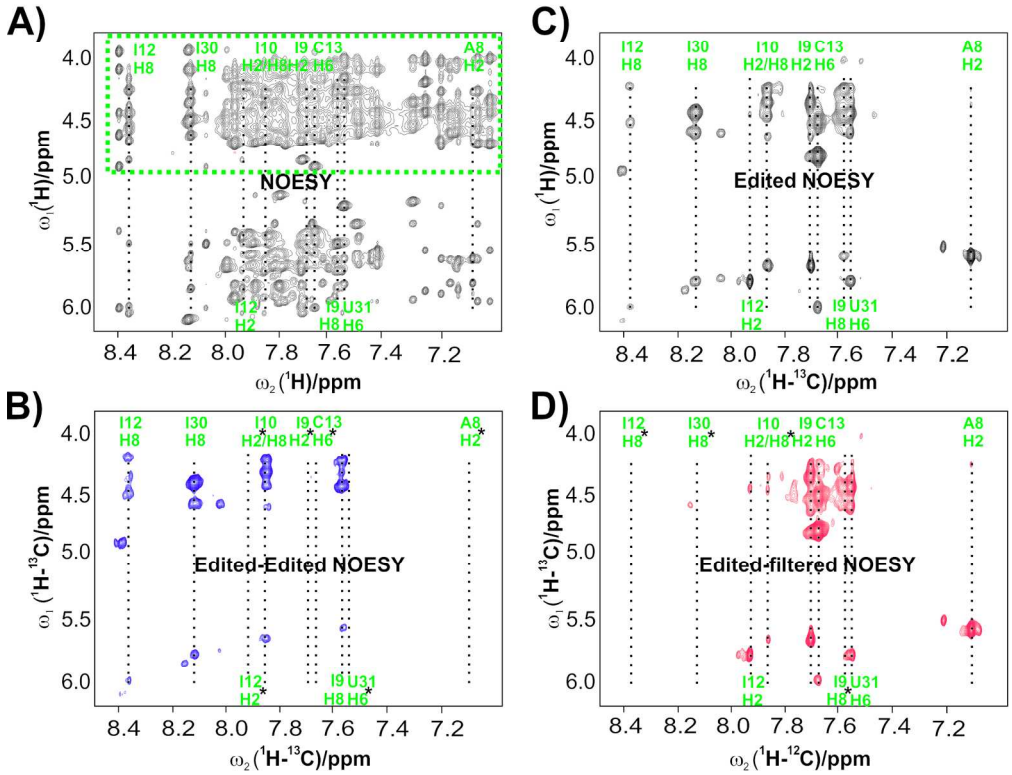


176x104mm (300 x 300 DPI)



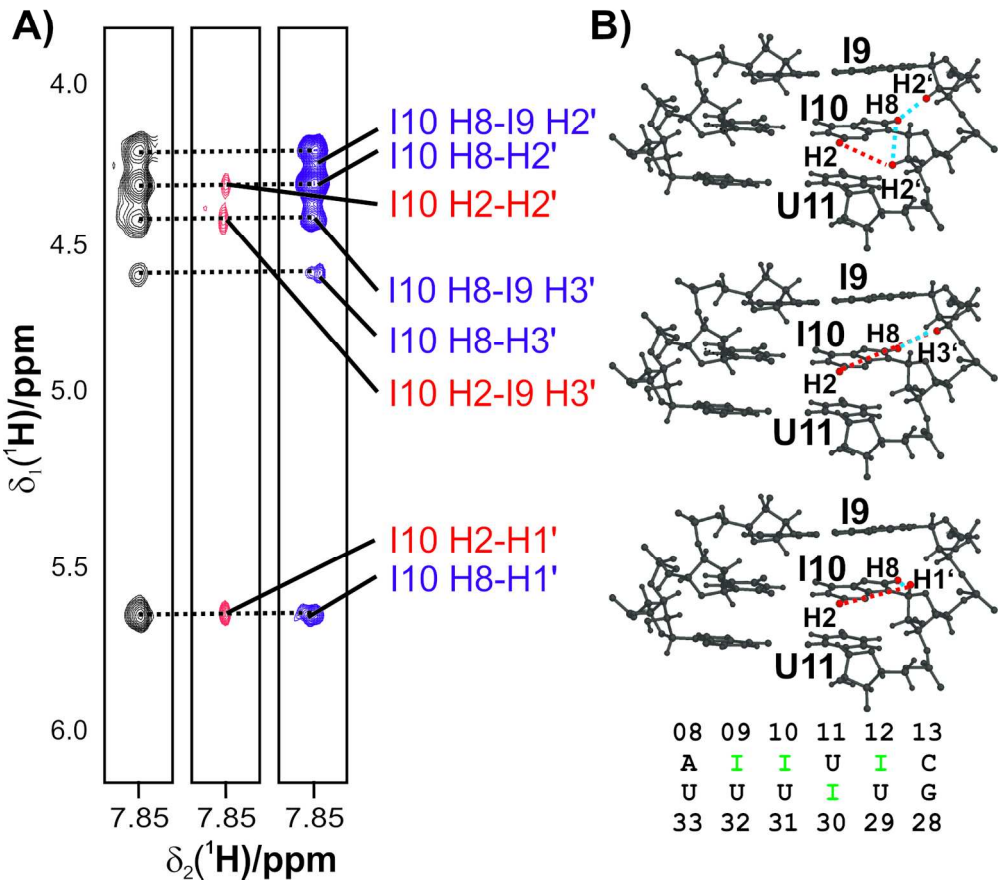
161x117mm (300 x 300 DPI)

Figure 3



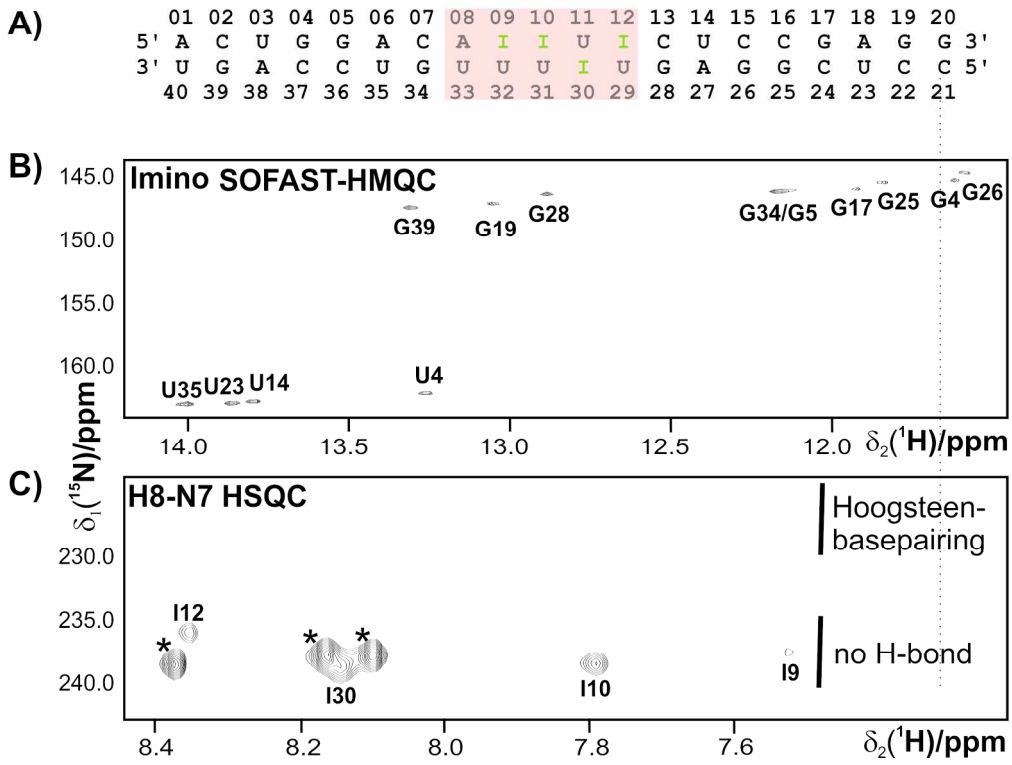
168x140mm (300 x 300 DPI)

Figure 4



156x151mm (300 x 300 DPI)

Figure 5



193x162mm (300 x 300 DPI)

Supplementary Information

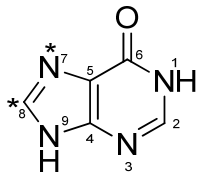
Site-specific isotope-labelling of inosine phosphoramidites and NMR analysis of an inosine-containing RNA duplex

*Andre Dallmann***,[1,2,#] *Alexander Beribisky***[1,2], *Felix Gnerlich***[,3], *Stefan Schiesser*[,3], *Thomas Carell**[,3] *Michael Sattler**[1,2]

Supplementary Methods

Synthesis of phosphoramidite 1

[7-^{15}N, 8-^{13}C]-Hypoxanthine (4)



To a solution of [5-^{15}N]-5,6-diamino-4(*3H*)-pyrimidone (1) (524 mg, 4.12 mmol, 1.0 eq.) in water (5.3 mL) was added 96% H_2SO_4 (0.26 mL, 4.68 mmol, 1.1 eq.) and [^{13}C]-formic acid (0.16 mL, 4.24 mmol, 1.0 eq.). The resulting mixture was stirred at 120 °C for 17 h. After filtration, the filtrate was set to pH 7.5 with a 28% NH_4OH-solution (1.0 mL). The resulting precipitate was collected by filtration and washed with ice cold water (10 mL) to obtain **4** as a pale yellowish solid (335 mg, 2.43 mmol, 59%).

^{1}H-NMR (400 MHz, D_2O, ppm): δ = 8.15 (s, 1 H, H2), 8.05 (dd, J = 204.5 Hz, 10.8 Hz, 1 H, H8).

^{13}C-NMR (101 MHz, D_2O, ppm): δ = 170.9 (CO), 157.0 (C4), 147.5 (C2), 146.4 (d, J = 3.4 Hz, C8), 120.5 (d, J = 7.0 Hz, C5).

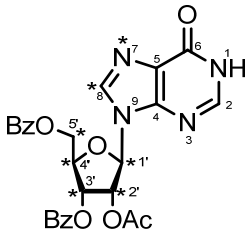
^{15}N-NMR (40 MHz, D_2O, ppm): δ = −159.8 (N7).

HRMS (EI+): calc. for $C_4{}^{13}CH_4N_3{}^{15}NO^+$ [M^+]: 138.0389, found: 138.0393.

IR (ATR): $\tilde{\upsilon}$ (cm^{-1}) = 2805 (w), 2738 (w), 2671 (w), 2622 (w), 2526 (w), 1669 (s), 1574 (m), 1404 (m), 1342 (m), 1203 (m), 1133 (m), 945 (m), 888 (s), 874 (s), 789 (s).

Melting range: > 390 °C decomposition.

[$^{13}C_6{}^{15}N$]-1-(2′-*O*-Acetyl-3′,5′-di-*O*-benzoyl-β-D-ribofuranosyl) hypoxanthine (6)



A round bottom flask equipped with a reflux condenser was charged with [7-^{15}N, 8-^{13}C]-Hypoxanthine **4** (303 mg, 2.20 mmol, 2.2 eq) and $(NH_4)_2SO_4$ (80 mg, 0.60 mmol, 0.6 eq.). After addition of 12 mL hexamethyldisilazane, the resulting slurry was refluxed for 3 h to yield a clear, reddish solution. The liquid was removed by distillation (75 °C, 200 mbar) and the remaining solid was coevaporated with dry toluene. The residue was dissolved in 20 mL 1,2-dichloroethane and thoroughly dried. $^{13}C_5$-1′,2′-Di-*O*-acetyl-3′,5′-di-*O*-benzoyl-β-D-ribofuranose (2) (447 mg, 1.00 mmol, 1.0 eq.) was added. TMSOTf (0.47 mL, 2.50 mmol, 2.5 eq.) is added dropwise over 1 h followed by refluxing the solution for 2 h. TLC analysis showed incomplete consumption of the ribofuranose and heating was continued for further 12 h to ensure complete reaction. The reaction solution was hydrolyzed with sat. $NaHCO_3$ and extracted with CH_2Cl_2 (3 × 80 mL). The combined organic phases were successively washed with sat. $NaHCO_3$ (80 mL) and brine (80 mL) and dried over $MgSO_4$. The crude product was purified by flash chromatography (CH_2Cl_2/MeOH = 20/1 → 10/1) to yield **6** as a clear oil (482 mg, 0.92 mmol, 92%).

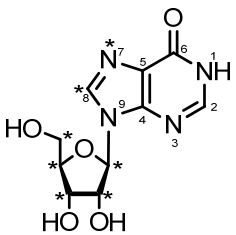
^{1}H-NMR (400 MHz, $CDCl_3$, ppm): δ = 12.90 (s br, 1 H, NH), 8.25 (d, *J* = 11.5 Hz, 1 H, H8), 8.09–8.06 (m, 4 H, Ar), 8.00 (s, 1 H, H2), 7.65–7.55 (m, 2 H, Ar), 7.52–7.42 (m, 4 H, Ar), 6.25 (d, *J* = 169.1 Hz, 1 H, H1'), 6.17 (d, *J* = 153.8 Hz, 1 H, H2'), 6.04 (d, *J* = 160.3 Hz, 1 H, H3'), 4.84 (dd, *J* = 147.9 Hz, 10.1 Hz, 1 H, H4'), 4.73, (d, *J* = 150.3, 1 H, H5'$_\alpha$), 4.65 (dd, *J* = 148.7 Hz, 9.4 Hz, 1 H, H5'$_\beta$), 2.03 (s, 3 H, CH_3).

^{13}C-NMR (101 MHz, $CDCl_3$, ppm): δ = 169.4 (CO^{Ac}), 166.1 (CO^{Bz}), 165.3 (CO^{Bz}), 158.7 (C6), 148.7 (C2), 145.4 (C4), 143.2 (CH-Ar), 143.1 (CH-Ar), 139.1 (C8), 133.8 (C-Ar), 133.5 (C-Ar), 129.7–128.5 (CH-Ar), 129.2 (dd, *J* = 111.6 Hz, 8.0 Hz, C5), 86.6 (d, *J* = 43.9, C1'), 80.9 (dd, *J* = 42.6 Hz, 38.4 Hz, C4'), 73.1 (t, *J* = 39.6 Hz, C2'), 71.2 (td, *J* = 39.4 Hz, 39.5 Hz, 3.1 Hz, C3'), 63.4 (dd, *J* = 43.6 Hz, 24.9 Hz, C5').

HRMS (ESI−): calc. for $C_{20}{}^{13}C_6H_{21}N_3{}^{15}NO_8^-$ $[M-H]^-$: 524.1537, found: 524.1538.

IR (ATR): $\tilde{\upsilon}$ (cm^{-1}) = 3064 (w), 2916 (w), 1739 (m), 1926 (s), 1704 (s), 1585 (w), 1451 (w), 1371 (w), 1272 (s), 1225 (m), 1110 (m), 1048 (m), 1025 (m), 789 (m).

[$^{13}C_5{}^{15}N$]-Inosine (7)



The protected nucleoside **6** (461 mg, 1.30 mmol) was stirred in 7 N methanolic ammonia. TLC analysis showed complete conversion after 48 h. The solvent was removed *in vacuo* and the product was obtained as an off-white solid (330 mg, 1.20 mmol, 92%).

^{1}H-NMR (400 MHz, DMSO, ppm): δ = 8.29 (dd, *J* = 214.7 Hz, 12.0 Hz, 1 H, H8), 8.03 (s, 1 H, H2), 5.83 (dd, *J* = 165.4 Hz, 3.5 Hz, 1 H, H1'), 5.43–5.38 (m br, 2 H, 2'-OH, 3'-OH), 5.14 (s br., 1 H, 5'-OH), 4.44 (d, *J* = 143.9 Hz, 1 H, H2'), 4.08 (d, *J* = 145.3 Hz, 1 H, H3'), 3.89 (d, *J* = 151.4 Hz, 1 H, H4'), 3.57 (m, 2 H, H5').

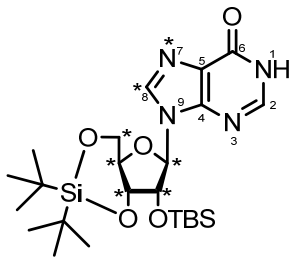
^{13}C-NMR (101 MHz, DMSO, ppm): δ = 157.0 (C6), 146.3 (C2), 142.9 (d, *J* = 9.5 Hz, C4), 139.1 (C8), 124.8 (C5), 87.8 (d, *J* = 42.4 Hz, C1'), 86.0 (dd, *J* = 41.5 Hz, 38.7 Hz, C4'), 74.5 (dd, *J* = 42.3 Hz, 37.7 Hz, C2'), 70.7 (td, *J* = 38.0 Hz, 3.6 Hz, C3'), 61.7 (d, 41.9 Hz, C5').

HRMS (ESI−): calc. for $C_4{}^{13}C_6H_{11}N_3{}^{15}NO_5^-$ $[M-H]^-$: 274.0907, found: 274.0908.

IR (ATR): $\tilde{\upsilon}$ (cm^{-1}) = 3541 (w), 3304 (m), 3098 (m), 2892 (m), 2722 (m), 1684 (s), 1592 (m), 1411 (w), 1370 (w), 1214 (m), 1108 (m), 1048 (s), 1022 (s), 882 (m), 820 (m), 785 (m).

Melting range: 220–224°C.

2'-(*O'*-^{t}Butyldimethylsilyl)-3',5'-(*O*-di-^{t}butylsilandiyl)-[$^{13}C_6{}^{15}N$]-inosine (8)



Nucleoside **7** (310 mg, 1.13 mmol, 1.0 eq.) was dissolved in dry DMF (40 mL) and cooled to 0 °C. Di-^{t}butylsilylbistriflate (547 mg, 1.24 mmol, 1.1 eq.) was added dropwise over 20 min. After stirring for further 30 min the reaction was allowed to warm to room temperature and imidazole (384 mg, 5.65 mmol, 5.0 eq.) was added. After further 30 min at room temperature TBSCl (204 mg, 1.36 mmol, 1.2 eq.) was added and the reaction mixture was heated to 60 °C for 2 h. Then, another 100 mg of TBSCl (0.67 mmol, 0.6 eq) were added and stirring was continued for another 2 h. The reaction was hydrolyzed by addition of 10 mL H_2O and the aqueous layer was extracted with EtOAc (2 × 50 mL). The combined organic phases were subsequently washed with sat. NaCl solution and dried over $MgSO_4$.The crude oil was purified via silica gel chromatography (CH_2Cl_2/MeOH = 40/1 → 20/1) to yield **8** (521 mg, 1.00 mmol, 88%) as a colourless oil.

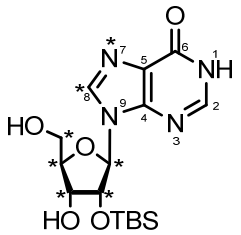
^{1}H-NMR (600 MHz, $CDCl_3$, ppm): δ = 8.20 (s, 1 H, H2), 8.17 (dd, *J* = 214.7 Hz, 11.6 Hz, 1 H, H8), 6.02 (dd, *J* = 167.1 Hz, 3.0 Hz, 1 H, H1'), 4.52 (d, *J* = 150.9 Hz, 1 H, H3'), 4.45 (d, *J* = 154.1 Hz, 1 H, H2'), 4.11 (d, *J* = 148.9 Hz, 1 H, H4'), 3.93–3.64 (m, 2 H, H5'), 1.04 (s, 18 H, 2 × $C(CH_3)_3$), 0.91 (d, *J* = 23.3 Hz, 9 H, $C(CH_3)_3$), 0.90 (s, 6 H, 2 × CH_3), 0.16 (d, *J* = 2.4 Hz, 3 H, CH_3), 0.08 (d, *J* = 6.1 Hz, 3 H, CH_3).

^{13}C-NMR (151 MHz, $CDCl_3$, ppm): δ = 158.8 (C6), 142.9 (C4), 142.6 (dd, *J* = 379.4 Hz, 10.3 Hz, C2), 138.7 (C8), 124.7 (C5), 89.0 (d, *J* = 41.2 Hz, C1'), 85.2 (dd, *J* = 46.0 Hz, 43.9 Hz, C4'), 75.5 (dd, *J* = 41.1 Hz, 38.1 Hz, C2'), 71.2 (dd, *J* = 38.4, 38.1 Hz, C3'), 62.1 (d, *J* = 43.8 Hz, C5'), 27.4 (6 C, 2 × ($C(\mathbf{C}H_3)_3$), 25.9 (3 C, 1 × ($C(\mathbf{C}H_3)_3$), 19.9 (2 C, 2 × ($\mathbf{C}(CH_3)_3$), 18.1 (1 × ($\mathbf{C}(CH_3)_3$), −4.7 (CH_3), −5.4 (CH_3).

HRMS (ESI+): calc. for $C_{18}{}^{13}C_6H_{43}N_3{}^{15}NO_5Si_2^+$ $[M+H]^+$: 530.2938, found: 530.2933.

IR (ATR): $\tilde{\upsilon}$ (cm^{-1}) = 3379 (w), 2948 (w), 2856 (w), 1663 (s), 1470 (m), 1251 (m), 1232 (m), 1091 (m), 1049 (m), 824 (s), 803 (m), 777 (s).

2'-(*O*-^{t}Butyldimethylsilyl)-[$^{13}C_6{}^{15}N$]-inosine (9)



Treatment of silylated compound **8** (750 mg, 1.44 mmol, 1.0 eq.) with 70% HF·pyridine (205 mg, 1.44 mmol, 1.0 eq.) in 1 mL pyridine was performed in dry CH_2Cl_2 at 0 °C for 2.5 h. The reaction was ended by addition of excess TMSOMe to yield the 2'-OH-TBS protected substance **9** (420 mg, 1.08 mmol, 76%) after flash chromatography (CH_2Cl_2/MeOH = 20/1 → 5/1) as a clear oil.

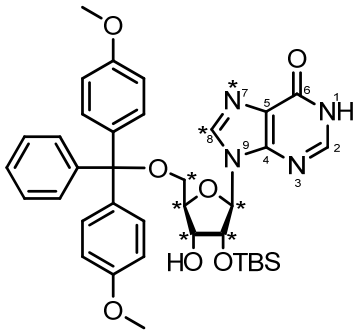
^{1}H-NMR (400 MHz, CD_3OD, ppm): δ = 8.38 (dd, *J* = 215.6 Hz, 11.5 Hz, 1 H, H8), 8.08 (s, 1 H, H2), 6.01 (dd, *J* = 166.8 Hz, 3.6 Hz, 1 H, H1'), 4.74 (dd, *J* = 146.6 Hz, 3.5 H, H2'), 4.31 (dd, *J* = 194.2 Hz, 3.3 Hz, 1 H, H3'), 4.15 (d, *J* = 155.9 Hz, 1 H, H4'), 3.99–3.57 (m, 2 H, H5'), 0.82 (s, 9 H, $C(CH_3)_3$), −0.02 (s, 3 H, CH_3), −0.13 (s, 3 H, CH_3).

^{13}C-NMR (101 MHz, CD_3OD, ppm): δ = 157.7 (C6), 145.5 (C4), 143.4 (d, *J* = 10.4 Hz, C2), 139.6 (s, C8), 124.7 (C5), 89.2 (d, *J* = 45.7 Hz, C1'), 86.5 (dd, *J* = 46.0 Hz, 39.7 Hz, 39.9 Hz, C4'), 76.4 (dd, *J* = 43.6 Hz, 38.2 Hz, C2'), 71.0 (dd, *J* = 38.3, 38.1 Hz, C3'), 61.6 (d, *J* = 41.3 Hz, C5'), 24.7 (3 C, $(C(\mathbf{C}H_3)_3)$, 17.5 ($\mathbf{C}(CH_3)_3$), −6.3 (CH_3), −6.6 (CH_3).

HRMS (ESI−): calc. for $C_{10}{}^{13}C_6H_{25}N_3{}^{15}NO_5Si^-$ $[M-H]^-$: 388.1771, found: 388.1767.

IR (ATR): $\tilde{\upsilon}$ (cm^{-1}) = 3852 (w), 3743 (w), 2928 (m), 2361 (w), 1707 (s), 1684 (s), 1506 (m), 1471 (w), 1206 (w), 1126 (w), 1062 (w), 838 (m), 786 (m).

5'-*O*-Dimethoxytrityl-2'-(*O*-^{t}butyldimethylsilyl)-[$^{13}C_6{}^{15}N$]-inosine (10)



Compound **9** (410 mg, 1.05 mmol, 1.0 eq.) was dissolved in 10 mL dry pyridine and dimethoxytrityl chloride (390 mg, 1.16 mmol, 1.1 eq.) was added at 0 °C and the resulting mixture was stirred at 0 °C for 16 h. The solvent was removed *in vacuo* at room temperature and the resulting solid was redissolved in CH_2Cl_2 and submitted to flash column chromatography (*iso*-hexane/acetone/CH_2Cl_2 = 35/40/25). The nucleoside **10** was obtained as a pale yellow foam (470 mg, 0.68 mmol, 65%).

^{1}H-NMR (400 MHz, $CDCl_3$, ppm): δ = 8.00 (d, J = 20.3 Hz, 1 H, H8), 7.94 (ddd, J = 213.3 Hz, 12.1 Hz, 11.8 Hz, 1 H, H2), 7.40–7.14 (m, 9 H, Ar), 6.78–6.73 (m, 4 H, Ar), 5.93 (d, J = 166.4 Hz, 1 H, H1'), 4.82 (d, J = 122.3 Hz, 1 H, H2'), 4.60 (d, J = 125.9 Hz, 1 H, H3'), 4.49 (d, J = 85.7 Hz, H4'), 3.73 (s, 6 H, 2 × OCH_3), 3.57 (dd, J = 7.3 Hz, 2.5 Hz, 1 H, H5'$_\alpha$), 3.33 (d, J = 2.4 Hz, 1 H, H5'$_\beta$), 0.80 (d, J = 13.1 Hz, 9 H, $C(CH_3)_3$), −0.02 (s, 3 H, CH_3), −0.18 (s, 3 H, CH_3).

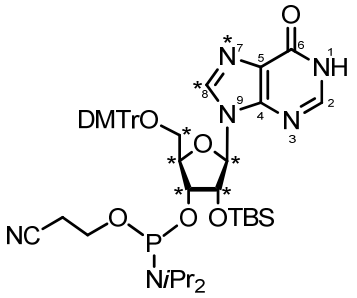
^{13}C-NMR (101 MHz, $CDCl_3$, ppm): δ = 158.5 (2 × **C**OCH_3), 157.7 (C6), 145.5 (C4), 141.0 (dd, J = 68.6 Hz, 10.1 Hz, C2), 139.9 (C-Ar), 138.1 (d, J = 21.4 Hz, C8), 135.5 (2 × C-Ar), 130.0 (2 × CH-Ar), 129.9 (2 × CH-Ar), 128.0 (CH-Ar), 127.8 (2 × CH-Ar), 126.9 (2 × CH-Ar), 126.8 (C5), 113.1 (4 × CH-Ar), 88.5 (ddd, J = 94.1 Hz, 42.8 Hz, 2.9 Hz, C1'), 86.8 (O**C(**C-Ar)$_3$), 84.2 (ddd, J = 43.3 Hz, 39.0 Hz, 4.7 Hz, C4'), 75.4 (ddd, J = 42.0 Hz, 36.7 Hz, 4.7 Hz, C2'), 74.7 (ddd, J = 38.7 Hz, 37.3 Hz, 3.1 Hz, C3'), 63.1 (dd, J = 43.5 Hz, 11.7 Hz, C5'), 55.2 (2 × OCH_3), 29.3 (**C**$(CH_3)_3$), 25.6 (C(**C**$H_3)_3$), −4.8 (d. J = 12.7 Hz, $SiCH_3$), −5.2 (d, J = 19.7 Hz, $SiCH_3$).

HRMS (ESI+): calc. for $C_{31}{}^{13}C_6H_{45}N_3{}^{15}NO_7Si^+$ $[M+H]^+$: 692.3224, found: 692.3228.

IR (ATR): $\tilde{\upsilon}$ (cm^{-1}) = 3114 (w), 3071 (w), 2929 (w), 2853 (w), 1699 (s), 1605 (w), 1686 (w), 1509 (s), 1249 (m), 1178 (m), 1132 (w), 1058 (m), 1028 (m), 975 (w), 871 (w), 832 (s).

Melting range: 163–164 °C.

3'-(Di*iso*propylamino-*O*-β-cyanoethoxyphosphino)-5'-*O*-dimethoxytrityl-2'-(*O*-^{t}butyldimethylsilyl)-[$^{13}C_6{}^{15}N$]-inosine (1)



In a Schlenk tube DMT-protected nucleoside **10** (120 mg, 0.17 mmol, 1.0 eq.), 2-cyanoethyl *N,N*-di*iso*propylchlorophosphoramidite (164 mg, 0.69 mmol, 4.0 eq.) and *N,N*-di*iso*propylethylamine (89 mg, 0.69 mmol, 4.0 eq.) were dissolved in 2 mL rigorously degassed CH_2Cl_2 and the solution was degassed three more times (freeze, pump, thaw). The solution was stirred at room temperature for 16 h. The reaction solution was directly applied to silica gel chromatography using deactivated silica (CH_2Cl_2 → CH_2Cl_2/MeOH =10/1). The product was isolated as a di*iso*propylamine-adduct (3.2 eq.) as determined from ^{1}H-NMR. Reaction yielded 176 mg (0.13 mmol, 77%) of a yellow resin as a mixture of two diastereomers on P.

^{31}P-NMR (81 MHz, $CDCl_3$, ppm): δ = 151.3, 150.4.

HRMS (ESI+): calc. for $C_{40}{}^{13}C_6H_{60}N_5{}^{15}NO_8PSiCl^-$ $[M+Cl]^-$: 926.3923, found: 926.3930.

Sequences of synthesized RNA Strands

Oligonucleotide 1: 5' – ACU GGA CA**I** **IUI** CUC CGA GG - 3'

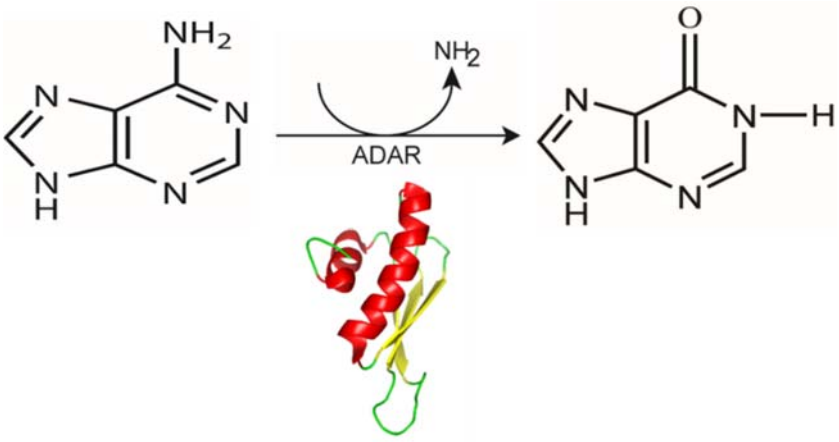
Oligonucleotide 2: 5' – CCU CGG AGU **I**UU UGU CCA GU - 3'

Supplementary references

1. Pagano, A.R., Lajewski, W.M. and Jones, R.A. (1995) Syntheses of [6,7-15N]-Adenosine, [6,7-15N]-2'-Deoxyadenosine, and [7-15N]-Hypoxanthine. *J. Am. Chem. Soc.*, **117**, 11669-11672.
2. Saito, Y., Zevaco, T.A. and Agrofoglio, L.A. (2002) Chemical synthesis of ^{13}C labeled anti-HIV nucleosides as mass-internal standards. *Tetrahedron*, **58**, 9593-9603.

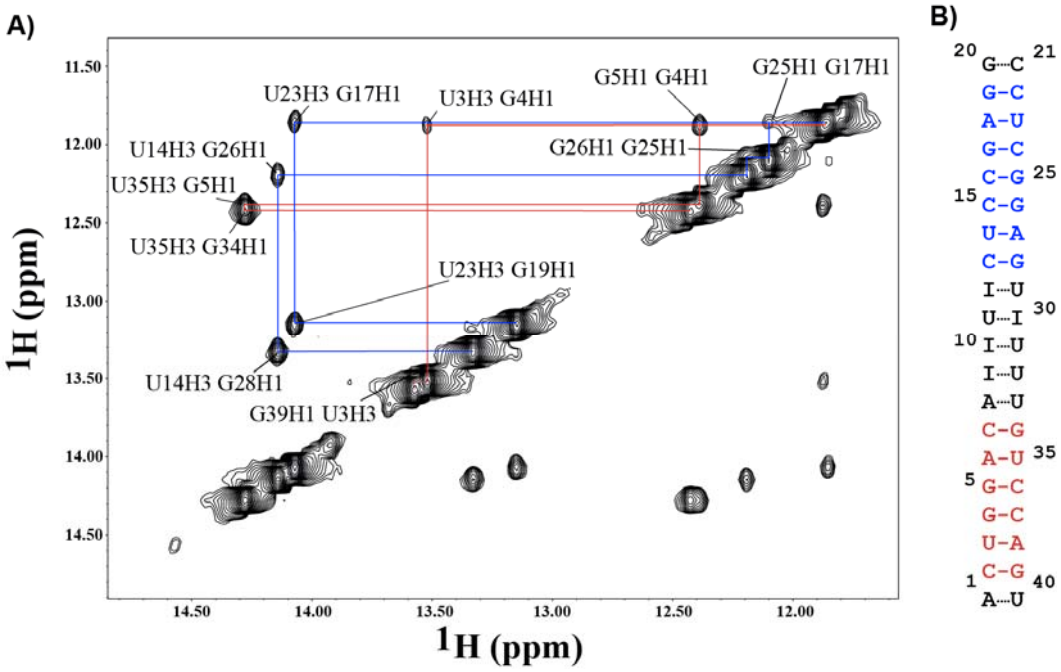
Supplementary Figures

Supplementary Fig. S1



Suppl. Fig. S1 The conversion of adenine to inosine by deamination of the exocyclic N6 amino group by adenosine deaminases acting on RNA (ADARs).

Supplementary Fig. S2



Suppl. Fig. S2 A) Imino-imino regions in a 2D water NOESY showing the cross-peaks for the regions flanking the inosine containing motif in the I-RNA duplex. The imino-imino walk for the base-pairs below and above the inosine containing motif are marked in red and in blue respectively. **B**) I-RNA secondary structure with the base-pairs for which the imino-imino walk could be performed are marked with the same colour code as in the NOESY spectrum.