VIP RNA Structures Very Important Paper

Deutsche Ausgabe: DOI: 10.1002/ange.201504014
Internationale Ausgabe: DOI: 10.1002/anie.201504014

Protonation-Dependent Base Flipping at Neutral pH in the Catalytic Triad of a Self-Splicing Bacterial Group II Intron**

*Maria Pechlaner, Daniela Donghi, Veronika Zelenay, and Roland K. O. Sigel**

Dedicated to Helmut Sigel

Abstract: *NMR spectroscopy has revealed pH-dependent structural changes in the highly conserved catalytic domain 5 of a bacterial group II intron. Two adenines with pK_a values close to neutral pH were identified in the catalytic triad and the bulge. Protonation of the adenine opposite to the catalytic triad is stabilized within a G(syn)–AH^+(anti) base pair. The pH-dependent anti-to-syn flipping of this G in the catalytic triad modulates the known interaction with the linker region between domains 2 and 3 (J23) and simultaneously the binding of the catalytic Mg^{2+} ion to its backbone. Hence, this here identified shifted pK_a value controls the conformational change between the two steps of splicing.*

Group II introns are large self-splicing ribozymes.[1] Their most conserved element is domain 5 (D5), a 35-nucleotide hairpin embedded at the core of the intron. D5 is characterized by a highly conserved AGC triplet ("catalytic triad") and an asymmetric internal loop that becomes a two-nucleotide bulge by formation of a G–U base pair (Figure 1). Both are involved in the coordination of two Mg^{2+} ions, which catalyze the two phosphotransesterification reactions of splicing, as well as in long-range tertiary contacts that help to position the 5′ and 3′ splice sites close to the active center.[2–9] The bulge is known as a dynamic structure that undergoes a conformational change upon docking to the rest of the intron.[10–14]

Aside from the same reaction mechanisms of group II introns and the eukaryotic spliceosome, the hypothesis of a common ancestry of the two[15] is fuelled by the striking similarity between D5 and the spliceosomal U6 intramolecular stem loop (ISL; Figure 1).[16] In particular, U6 ISL shares

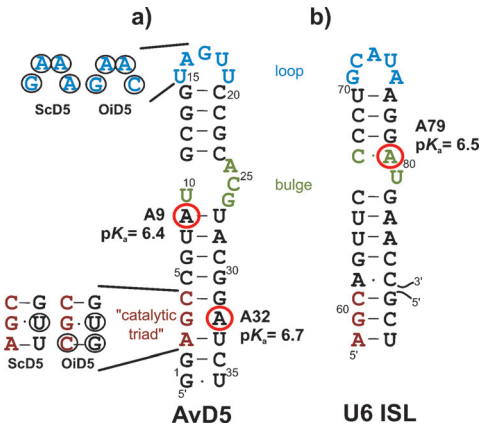

Figure 1. a) Secondary structure of AvD5 indicating sequence differences to the previously studied ScD5 and OiD5. Residues in the loop (blue) and the catalytic triad (red) that differ in the three sequences are circled in black. b) Secondary structure of *S. cerevisiae* U6 ISL.[26]

the highly conserved AGC triad at a 5 nt distance from a metal-coordinating bulge and can be functionally replaced by D5.[17]

Whereas individual nucleotides lack pK_a values in the neutral pH range,[18] in RNA molecules, the local structural and electrostatic environment as well as nearby metal ions can lead to strongly perturbed pK_a values, allowing for nucleobases to participate directly in acid–base catalysis by ribozymes.[19–21] Methods to determine pK_a values in larger nucleic acids are scarce and mostly indirect, as for example, 2-aminopurine fluorescence assays.[21] One example is the catalytic cytosine in the HDV ribozyme with a pK_a value of 6.4.[20,22] In U6 ISL, the bulge nucleotide A79 is protonated with a pK_a value of 6.5, in the course of which an A^+–C base pair is formed and U80 moves to a solvent-exposed conformation.[23,24] In contrast, no nucleotides with perturbed pK_a values have been identified in group II introns thus far, despite several solution studies and molecular dynamics simulations[25] on individual domains, including D5 domains from different organisms.[12,13]

Herein, we report two acid–base equilibria near physiological pH values in the bulge and in the catalytic triad of the bacterial group II intron 5 from *A. vinelandii*. These shifted pK_a values reveal yet another parallel to the spliceosome. The bulge region of *A. vinelandii* domain 5 (AvD5) is identical to the ones of the best studied group II introns, *Sc*.ai5γ D5 (ScD5) from yeast and *O. iheyensis* D5 (OiD5),[13,14] whereas the usual catalytic-triad GU wobble pair is replaced by a GA

[*] Dr. M. Pechlaner, Dr. D. Donghi, Dr. V. Zelenay, Prof. Dr. R. K. O. Sigel
Department of Chemistry, University of Zurich
Winterthurerstrasse 190, 8057 Zürich (Switzerland)
E-mail: roland.sigel@chem.uzh.ch
Homepage: http://www.chem.uzh.ch/rna

[**] Financial support by the Swiss National Science Foundation (project support to R.K.O.S. and to D.D. through an Ambizione Fellowship), an FP7 Marie Curie Intra European Fellowship (to D.D.), and the University of Zurich is gratefully acknowledged. R.K.O.S. holds an ERC starting grant.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201504014. Atomic coordinates for structures are available in the Protein Data Bank (PDB No. 2M57). The NMR chemical shifts can be obtained from the Biological Magnetic Resonance Data Bank (BMRB, accession code 19039).

mismatch in AvD5 (Figure 1). AvD5 forms a 35 nt hairpin with a stable 4 bp upper helix that is capped by a pentaloop and connected to the lower helix via the bulge (Figure 1). At pH/pD 6.7, resonances of protons at or near the GA mismatch in the catalytic triad as well as of the bulge are either not detectable or very broad. The Watson–Crick (WC) base pairs A3–U33 and A9–U27 cannot be confirmed by NOESY and $J_{N,N}$ HNN-COSY spectroscopy. Furthermore, in neither the bulge nor the catalytic triad, imino or amino protons are stabilized by hydrogen bonding, and the resonances of many non-exchangeable protons are broad or absent. Line broadening is a sign of either dynamics or undefined structure and has been also observed in the bulge of other D5s.[10,13]

In AvD5, both catalytic triad and bulge are strongly influenced by pH changes in the neutral range. Upon an increase in pH value, we observed strong chemical-shift changes and distinct sharpening of the previously broad ^{1}H resonances of the bulge. At pH/pD 7.8, line widths are comparable to those of unaffected resonances. On the other hand, lowering the pH value enhances the broadening until the bulge resonances become unobservable at or below pH/pD 6.5. In the catalytic triad, several resonances of protons close to the G4–A32 mismatch and all G4 resonances only become visible at pH/pD values of >7.5 and <6.0, albeit at strongly shifted positions (Figure S1). Even at very high pH values (>9), the catalytic-triad signals remain broad, indicating that no stable structure is formed. In contrast, at pH <5.5, the region seems to become more stable than regular WC base-paired regions, which start to melt because of protonation. Stabilization of the A3–U33 base pair next to the GA mismatch is indicated by the appearing U33 imino resonance with a strong interstrand cross peak to A3-H2 (Figure 2 and S2). Another sign of stabilization is the appearance of resonances corresponding to the imino and amino protons in the GA mismatch itself (G4-H1, A32-H61/H62; Figure 2). Furthermore, at low pH values, characteristic features of a *syn* conformation of G4 in the GA mismatch emerge (the NOE A32-H2/G4-H8 appears and G4-H8/H1′ becomes more intense, Figure S1).

The observed pH dependence indicates one or more acid–base equilibria in the neutral pH range. Protonation at neutral pH is most commonly observed for C-N3 and A-N1, which have pK_a values of approximately 4.2 and 3.6 in single-stranded RNA.[18] H^+ addition has a pronounced effect on the chemical shift of the respective ^{15}N atoms, which provides a direct measure of their protonation states. In AvD5, the ^{13}C and ^{15}N chemical-shift changes in the neutral pH range clearly indicate protonation at both the catalytic triad (A32-N1) and the bulge (A9-N1; Figure 3 a, b). The resonances of A32-N1 and C2 shift by −65.2 and −7.5 ppm, respectively, whereas that of A9-N1 disappears already above pH 7 when the pH value is decreased (starting from pH 9). The resonance of

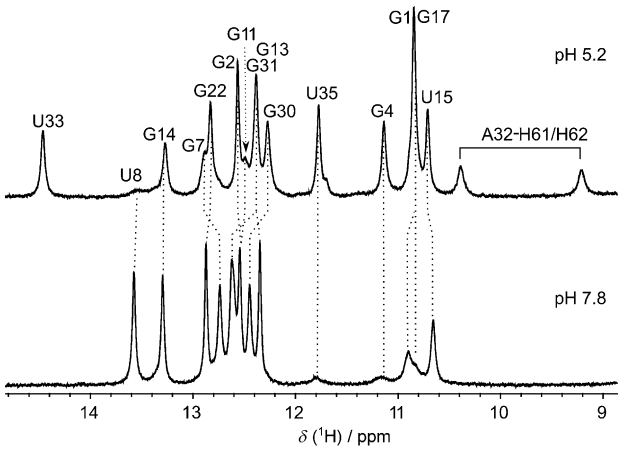

Figure 2. ^{1}H 1D NMR spectra of the imino region recorded at 278 K (700 MHz, 60 mM KCl). Upon decreasing the pH value from pH 7.8 (bottom) to pH 5.2 (top), exchangeable protons at the helix ends (G1-H1, U35-H3) and near the catalytic triad (U33-H3, A32-H61/H62, G4-H1) become observable, whereas the resonances of the only observable imino protons close to the bulge (U8-H3, G11-H1) decrease in intensity.

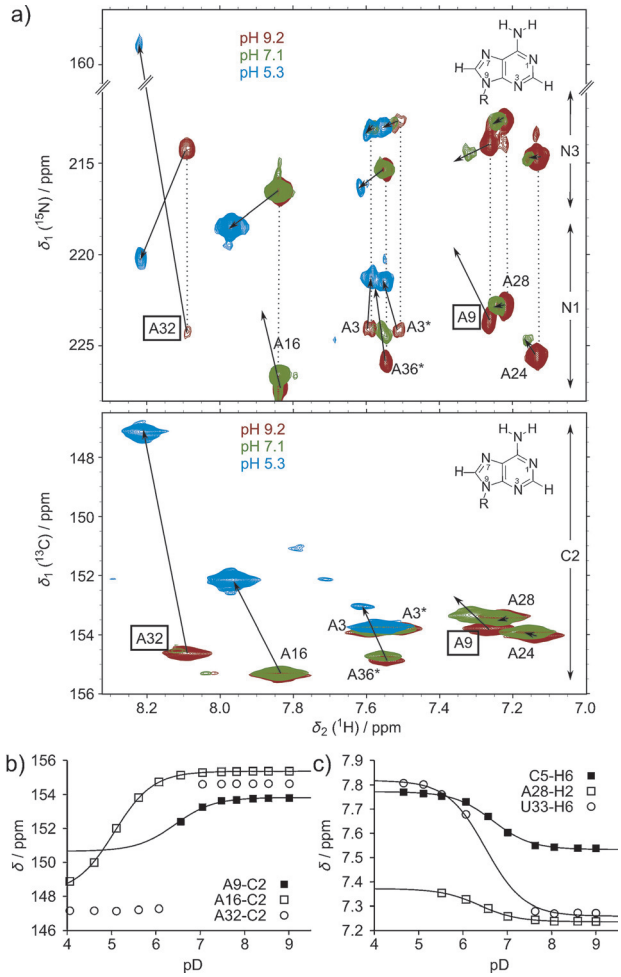

Figure 3. a) $^2J(^1H,^{15}N)$ and $^1J(^1H,^{13}C)$ HSQC spectra showing adenine N1, N3, and C2 assignments at three pH values. A9 and A32, the two adenines that are protonated close to neutral pH, are highlighted in boxes (303 K, 700 MHz, 60 mM KCl). b, c) Chemical-shift changes of A-C2, as an example, and aromatic ^{1}H resonances in bulge, loop, and catalytic triad with changes in the pH value. The lines represent least-squares fits to Eq. (1) (see the Supporting Information). Resonances marked with * are due to 3′ heterogeneity (see the Supporting Information).

A9-C2 starts to shift at a higher pH value than any other Ade-C2 or Cyt-C6 resonances in AvD5 (apart from A32-C2; see the Supporting Information, Figure S3 and Table S1). In particular, in the bulge, neither A24-C2 nor C25-C6 move significantly in the neutral pH range. Protonation of A9 is rather unexpected, because in contrast to the unpaired bulge nucleotides A24 and C25, A9 can form a WC pair to U27 according to structural data of other D5 domains.[7,13,14]

The pK_a values of A32 and A9 were determined by non-linear least-squares fits of the pH-dependent chemical-shift changes of the ^{1}H and ^{13}C resonances between pD 5.5 and 9 (see Figure 3b,c and Tables S2 and S3). A protonation event influences not only resonances of the concerned nucleotide but also those of its neighbors, and fits of resonances from nuclei close to the protonation sites should return the same pK_a values. Indeed, values for the bulge (residues U8–G11 and C23–A28) and the catalytic triad (residues A3–C5 and G31–U33) agree very well with each other (see Tables S2 and S3). The arithmetic mean for each protonation site gives the final values, $pK_{a,A9,bulge} = 6.38 \pm 0.12$ for protonation at A9-N1 and $pK_{a,A32,cat.} = 6.69 \pm 0.10$ at A32-N1.

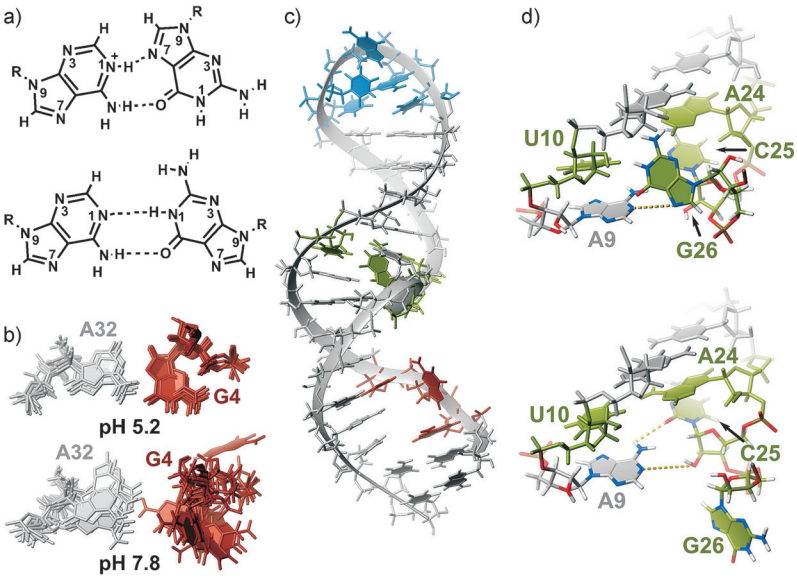

Figure 4. a) G(*syn*)–AH$^+$(*anti*) (top) and G(*anti*)–A(*anti*) (bottom) base pairs. b) The GA mismatch in the catalytic triad of AvD5 at low and high pH values (the ten lowest-energy structures are shown). c) Solution structure of AvD5 (the lowest-energy structure at pH 7.8 is shown; see Figure S5 for the ensemble). Blue: loop, green: bulge, red: catalytic triad. d) Representative structures for the bulge conformation at pH 7.8 with G26 in the minor and major groove, respectively. Dashed yellow lines indicate interactions that could favor protonation of A9N1.

Outside the immediate environment of the bulge and the catalytic triad, no resonances display significant changes in chemical shift between pH 5.5 and 9.0 (Table S1). In particular, the pentaloop is hardly affected by pH changes: A16 starts to show signs of protonation only below pH 6 (Figure 3a,b; for more details on the AvD5 loop, see the Supporting Information and Figure S4).

On the basis of distance restraints derived from ^{1}H,^{1}H NOESY spectra at either pH 7.8 or 5.2, we calculated two ensembles of solution structures (Figures 4b,c and S5, Table S4). The ensemble of the bulge at pH > 7 resembles the previously published solution structures (see the Supporting Information).[12,13] The bulge structure at acidic pH remains unclear because of the lack of resonances, and we can only assume that there is an ensemble of interconverting structures. The proximity of C25-O2, C25-O2′, or G26-N7 to the amino and imino moieties of A9, as observed in some of the structures of the AvD5 ensemble, could favor the protonated state (Figure 4d).

The conformational exchange of a protonated and an unprotonated species in the bulge region at neutral pH is shared by the spliceosomal U6 ISL.[23] In that case, it was possible to show that the exchange involves an expulsion of U80 from the helix,[24] which is reminiscent of the expulsion of D5 bulge residues in the intron-docked form of D5.[5,6,14] In contrast, in AvD5, a protonated A9 residue is incompatible with the docked D5 conformation, where A9 is not only paired with U27, but also flanked by base pairs on either side.[14] The small fraction of protonated A9 under physiological conditions is a sign of the inherent flexibility of the bulge, which facilitates docking.

The resonances in the catalytic triad that experience the strongest shifts are broadened beyond detection near the pK_a (e.g., A32-C2 and U33-H6 in Figure 3b,c). The underlying exchange process might be the flipping of G4 from a *syn* to an *anti* conformation. Our NOE data in combination with A32 protonation suggests the formation of a G(*syn*)–AH$^+$(*anti*) base pair (Figure 4a,b). The pH-dependent switching of a single GA mismatch between a protonated G(*syn*)–AH$^+$-(*anti*) base pair and an unprotonated G(*anti*)–A(*anti*) base pair has previously been observed in DNA, but, to the best of our knowledge, never in RNA.[27–29] In AvD5, the low intensity of the G4-H8/H1′ peak suggests an *anti* conformation at basic pH; however, structural information is scarce owing to comparatively broad peaks and unobservable imino protons. The structure is much less defined at basic pH than at acidic pH (local RMSD at pH 5.2 of 0.9 Å vs. 2.2 Å; Table S4, Figure 4b) and both stacked or looped-out G4 conformations are possible at pH 7.8 (Figure 4b). On the opposite strand, A32 is stacked inside the helix irrespective of the pH value as is confirmed by a strong interstrand A32-H2/C5-H1′ cross peak.

The stabilizing hydrogen bond G4-N7 in the G(*syn*)–AH$^+$(*anti*) base pair formed upon protonation explains the elevated pK_a value of A32-N1. In the crystal structure of a group IIC intron, the respective GU wobble pair forms a base triple with a conserved G in the linker between domains 2 and 3 (J23).[14] Marcia and Pyle suggested that this interaction had to be broken between the two splicing steps to allow for substrate exchange.[6] This structural change between the two splice steps is thought to be accompanied by a transient disassembly of the $[Mg^{2+}]_2$ core of the active site. The phosphate of the catalytically active G4 coordinates

these Mg^{2+} ions.[6] Moreover, in a group IIC intron crystal structure representing the pre-catalytic state, this G has been mutated to an A, the latter modelled in a *syn* conformation.[30] This structure is reminiscent of the *syn*-G4 conformation that we observe in AvD5, and strongly resembles the intermediate observed by Marcia and Pyle.[6] The pH-dependent *anti*-to-*syn* flipping of G4 in AvD5 would disrupt the interaction with J23 and modulate the G4 backbone, thus influencing the geometry of the Mg^{2+} ion-binding site and allowing for the conformational changes necessary between the two steps of splicing. It could then play a role both in the clearance of the active site and possibly in substrate positioning for the second step of splicing by new contacts made via the free WC face of G4 in the G(*syn*)–AH^+(*anti*) pair.

In conclusion, we have identified and characterized the presence of two acid–base equilibria near neutral pH within the catalytic core of a self-splicing ribozyme. The pK_a perturbation in the bulge region is likely to play a role in the folding process of the group II intron. In contrast, the acid–base equilibrium in the catalytic triad can directly influence catalysis: Comparison with available crystal structures suggests a protonation-dependent flipping of G4 between the G(*syn*)–AH^+(*anti*) base pair and an interaction with the J23 linker; G4 flipping thus plays an important role in active site clearance and substrate positioning between the two steps of splicing.

Keywords: group II introns · NMR spectroscopy · ribozymes · RNA · structure–function relationships

How to cite: *Angew. Chem. Int. Ed.* **2015**, *54*, 9687–9690
Angew. Chem. **2015**, *127*, 9823–9826

[1] K. Lehmann, U. Schmidt, *Crit. Rev. Biochem. Mol. Biol.* **2003**, *38*, 249.
[2] M. Boudvillain, A. de Lencastre, A. M. Pyle, *Nature* **2000**, *406*, 315.
[3] A. de Lencastre, S. Hamill, A. M. Pyle, *Nat. Struct. Mol. Biol.* **2005**, *12*, 626.
[4] A. de Lencastre, A. M. Pyle, *RNA* **2008**, *14*, 11.
[5] N. Toor, K. Rajashankar, K. S. Keating, A. M. Pyle, *Nat. Struct. Mol. Biol.* **2008**, *15*, 1221.
[6] M. Marcia, A. M. Pyle, *Cell* **2012**, *151*, 497.
[7] A. R. Robart, R. T. Chan, J. K. Peters, K. R. Rajashankar, N. Toor, *Nature* **2014**, *514*, 193.
[8] M. Marcia, S. Somarowthu, A. Pyle, *Mobile DNA* **2013**, *4*, 14.
[9] R. K. O. Sigel, A. Vaidya, A. M. Pyle, *Nat. Struct. Biol.* **2000**, *7*, 1111.
[10] N. V. Eldho, K. T. Dayie, *J. Mol. Biol.* **2007**, *365*, 930.
[11] O. H. Gumbs, R. A. Padgett, K. T. Dayie, *RNA* **2006**, *12*, 1693.
[12] M. Seetharaman, N. V. Eldho, R. A. Padgett, K. T. Dayie, *RNA* **2006**, *12*, 235.
[13] R. K. O. Sigel, D. G. Sashital, D. L. Abramovitz, A. G. Palmer III, S. E. Butcher, A. M. Pyle, *Nat. Struct. Mol. Biol.* **2004**, *11*, 187.
[14] N. Toor, K. S. Keating, S. D. Taylor, A. M. Pyle, *Science* **2008**, *320*, 77.
[15] K. T. Dayie, R. A. Padgett, *RNA* **2008**, *14*, 1697.
[16] K. S. Keating, N. Toor, P. S. Perlman, A. M. Pyle, *RNA* **2010**, *16*, 1.
[17] G. C. Shukla, R. A. Padgett, *Mol. Cell* **2002**, *9*, 1145.
[18] A. Domínguez-Martín, S. Johannsen, A. Sigel, B. P. Operschall, B. Song, H. Sigel, A. Okruszek, J. M. González-Pérez, J. Niclós-Gutiérrez, R. K. O. Sigel, *Chem. Eur. J.* **2013**, *19*, 8163.
[19] B. Lippert, *Chem. Biodiversity* **2008**, *5*, 1455.
[20] J. L. Wilcox, A. K. Ahluwalia, P. C. Bevilacqua, *Acc. Chem. Res.* **2011**, *44*, 1270.
[21] J. L. Wilcox, P. C. Bevilacqua, *J. Am. Chem. Soc.* **2013**, *135*, 7390.
[22] B. Gong, J.-H. Chen, E. Chase, D. M. Chadalavada, R. Yajima, B. L. Golden, P. C. Bevilacqua, P. R. Carey, *J. Am. Chem. Soc.* **2007**, *129*, 13335.
[23] A. Huppler, L. J. Nikstad, A. M. Allmann, D. A. Brow, S. E. Butcher, *Nat. Struct. Biol.* **2002**, *9*, 431.
[24] N. J. Reiter, H. Blad, F. Abildgaard, S. E. Butcher, *Biochemistry* **2004**, *43*, 13739.
[25] M. Pechlaner, R. K. O. Sigel, W. F. van Gunsteren, J. Dolenc, *Biochemistry* **2013**, *52*, 7099.
[26] J. E. Burke, D. G. Sashital, X. Zuo, Y.-X. Wang, S. E. Butcher, *RNA* **2012**, *18*, 673.
[27] X. Gao, D. J. Patel, *J. Am. Chem. Soc.* **1988**, *110*, 5178.
[28] C. Carbonnaux, G. A. Van der Marel, J. H. Van Boom, W. Guschlbauer, G. V. Fazakerley, *Biochemistry* **1991**, *30*, 5449.
[29] A. N. Lane, T. C. Jenkins, D. J. Brown, T. Brown, *Biochem. J.* **1991**, *279*, 269.
[30] R. T. Chan, A. R. Robart, K. R. Rajashankar, A. M. Pyle, N. Toor, *Nat. Struct. Mol. Biol.* **2012**, *19*, 555.

Received: May 2, 2015
Published online: June 26, 2015

