Controlling Site-Directed RNA Editing by Chemically Induced Dimerization

Anna S. Stroppel,[a] Ruth Lappalainen,[a] and Thorsten Stafforst*[a]

Abstract: Various RNA-targeting approaches have been engineered to modify specific sites on endogenous transcripts, breaking new ground for a variety of basic research tools and promising clinical applications in the future. Here, we combine site-directed adenosine-to-inosine RNA editing with chemically induced dimerization. Specifically, we achieve tight and dose-dependent control of the editing reaction with gibberellic acid, and obtain editing yields up to 20% and 44% in the endogenous STAT1 and GAPDH transcript in cell culture. Furthermore, the disease-relevant MECP2 R106Q mutation was repaired with editing yields up to 42%. The introduced principle will enable new applications where temporal or spatiotemporal control of an RNA-targeting mechanism is desired.

RNA base editing enables the rewriting of genetic information with high efficiency and without the risk of permanent off-target effects and thus has high prospects for clinical application.[1] Furthermore, the reversibility of an editing event on the transient (m)RNA copy allows to tune the yield of base exchange and might be used to introduce otherwise lethal mutations suddenly and/or temporally restricted.[2] The SNAP-ADAR approach was engineered for site-directed adenosine-to-inosine (A-to-I) RNA editing.[3] For this, the dsRNA binding domains responsible for substrate recognition in wildtype adenosine deaminase acting on RNA (ADAR)[4] are replaced by the self-labeling SNAP-tag. The SNAP-tag binds covalently to guideRNAs carrying its substrate, O^6-benzylguanine[5] (BG, snap-guideRNAs), which then allows for recruitment of the fused ADAR deaminase domain to a specific target via Watson-Crick base pairing. The approach has been shown to be rationally programmable, to achieve high editing yields in cell culture and in vivo,[6] to be very precise,[2] and to be efficient enough to

enable concurrent editing.[6] The extension of the approach by further layers of control is desirable. Recently, we achieved photo-control in developing embryos by application of guide-RNAs carrying a nitropiperonyloxymethyl-protected BG moiety.[7] Here, we now include control of the editing reaction by chemically induced dimerization[8] with a small molecule. This opens many new opportunities to run editing under temporal, spatial or dose control.

Specifically, we decided to use gibberellic acid (GA₃) for chemically induced dimerization. GA₃ is a plant hormone that can be delivered as a cell-permeable prodrug (GA₃–AM), that has been shown to induce the heterodimerization of the two plant proteins GAI (gibberellic acid insensitive) and GID1A (gibberellin insensitive dwarf 1A, Figure 1) on a timescale of seconds to minutes inside live cells.[9] Binding of GA₃ to GID1A induces a conformational change that leads to recruitment of GAI. In order to control the SNAP-ADAR-based editing reaction by GA₃-induced dimerization, the SNAP-tag and the ADAR deaminase domain needed engineering into two separate fusion proteins with GAI and GID1A, respectively. We decided to use a GAI₁₋₉₂–ADAR1 fusion, applying a 92 amino acid fragment of GAI sufficient for dimerization,[9] and a SNAP-GID1A fusion. In our design, we kept the SNAP-tag and ADAR deaminase domain at their respective N- and C-terminal position. We combined the ADAR deaminase domain with the GAI fragment to also place the latter in accordance with its native N-terminal position. The GID1A protein has recently been applied in fusion with an N-terminal eGFP-tag.[9–10] We expected that the exchange of the eGFP-tag with a SNAP-tag will not interfere with the function of the plant protein. Finally, both transgenes have the same size (59 kDa). We engineered four plasmid constructs (**I–IV**) that contain both transgenes in one expression cassette (Figure 2a). This design was chosen to obtain a balanced expression of both transgenes after stable genetic integration of the respective single plasmid into a cell line. Furthermore, under transient expression conditions such constructs would help to reduce the transfection bias and to improve the balance in the expression of both transgenes. The two transgenes were either put consecutively, each with its own CMV promotor and bGH terminator, or they were expressed as a single P2A[11] construct from one promotor using a translational skipping mechanism to create two separate proteins from one transcript in a nearly 1:1 stoichiometry. For the editase fusion, we either chose the catalytic deaminase domain of wildtype human ADAR1 (GAI₁₋₉₂–ADAR1) or of a hyperactive mutant (GAI₁₋₉₂–ADAR1Q), carrying a well-known E>Q single point mutation.[12] To create duo cell lines expressing both transgenes stably under doxycycline induction,

[a] A. S. Stroppel, R. Lappalainen, Prof. T. Stafforst
Interfaculty Institute of Biochemistry
University of Tübingen
Auf der Morgenstelle 15, 72076 Tübingen (Germany)
E-mail: thorsten.stafforst@uni-tuebingen.de

Chemistry–A European Journal

Communication
doi.org/10.1002/chem.202101985

Chemistry Europe
European Chemical
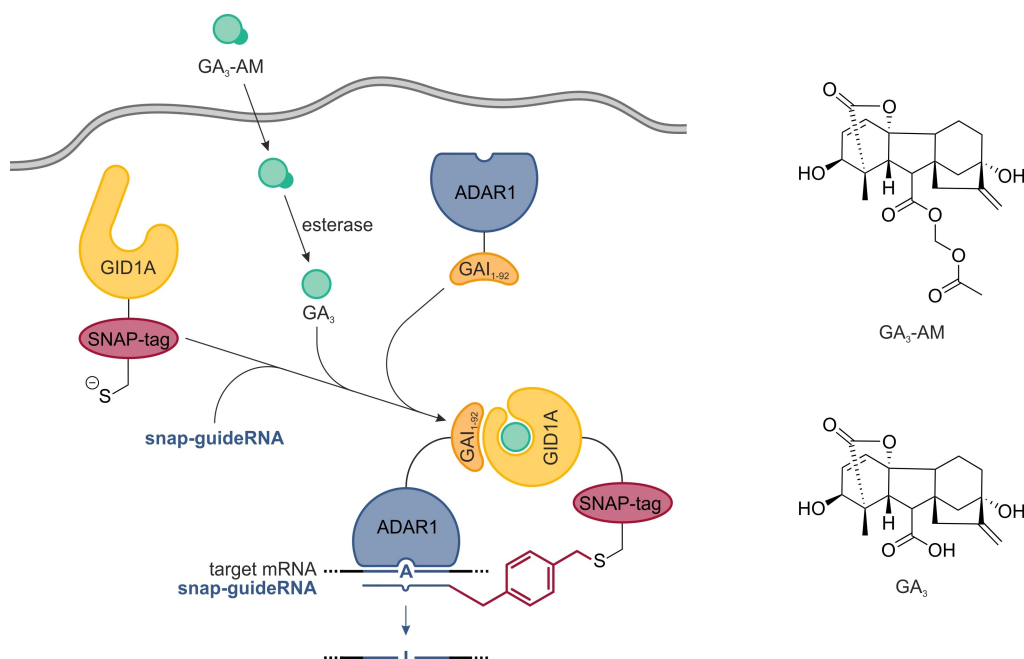Societies Publishing

Figure 1. Principle of gibberellic acid-induced site-directed RNA editing with the SNAP-ADAR platform. Covalent conjugation of an O^6-benzylguanosine (BG)-modified guideRNA (snap-guideRNA) with the SNAP-tagged deaminase ADAR enables the steering of A-to-I deaminase activity to any arbitrary mRNA to achieve programmable, RNA-guided site-specific RNA editing. To place the process under control of gibberellic acid, the SNAP-ADAR protein is split into a GAI_{1-92}–ADAR and a SNAP-GID1A fusion, separating the editing activity from the RNA-targeting mechanism. Gibberellic acid, delivered in the form of a cell-permeable acetoxymethyl ester (GA_3–AM), enforces heterodimerization of GAI_{1-92} and GID1A by binding to the latter, and thereby recruits the ADAR deaminase to the guideRNA/mRNA substrate duplex.
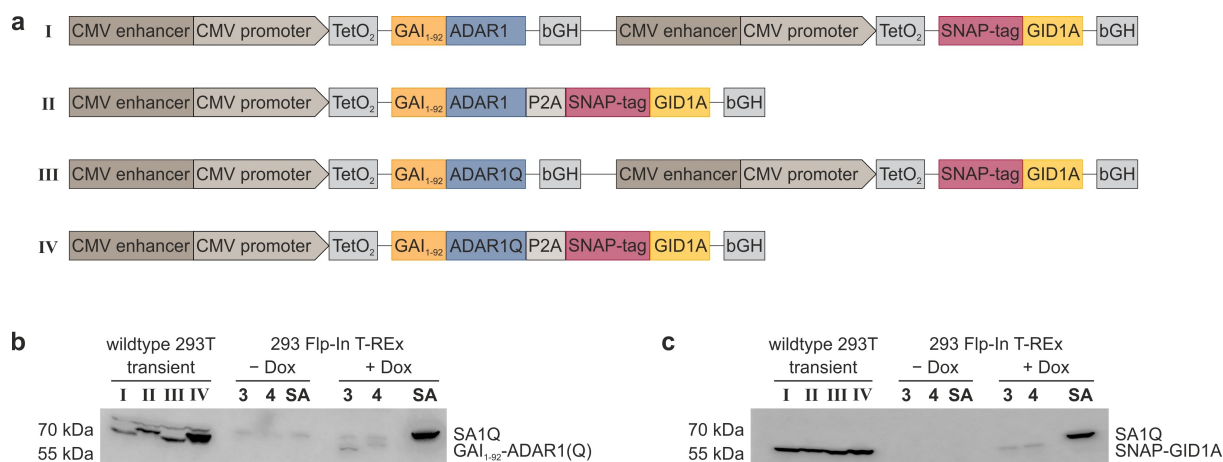

Figure 2. Expression constructs for gibberellic acid-induced RNA editing with SNAP-ADAR and analysis of transgene expression. a) Constructs **I–IV** were designed to create transgenic 293 Flp-In T-REx cell lines **1–4**, stably co-expressing GAI_{1-92}–ADAR1(Q) and SNAP-GID1A from one cassette under doxycycline control. TetO$_2$: tet operator, leads to repression of expression in the absence of a tetracycline;[13] bGH: bovine growth hormone terminator; P2A: porcine teschovirus-1 self-cleaving 2A peptide.[11] The protocol for the generation of stable cell lines **1–4** from constructs **I–IV** and details on the constructs can be found in the Supporting Information. b) Characterization of GAI_{1-92}–ADAR1(Q) expression via Western Blot (α-ADAR1 deaminase domain). Wildtype 293T cells were transiently transfected with constructs **I–IV** and stable cell lines **3** and **4** were examined without (−Dox) and with 24 h (+Dox) doxycycline induction. Previously established SA1Q 293 Flp-In T-REx cell line (**SA**) shown for comparison. c) Same as (b) but for expression analysis of SNAP-GID1A (α-SNAP-tag).

we applied the 293 Flp-In T-REx system. For each construct, **I–IV**, we generated a separate duo cell line, **1–4**, by a plasmid transfection and antibiotic selection procedure, as described before.[6] The doxycycline-dependent expression of both transgenes was confirmed by Western blot with antibodies against ADAR1 deaminase (Figure 2b) and SNAP-tag (Figure 2c). Nota-

bly, the expression levels were comparably low in relation to the stable expression of SNAP-ADAR1Q after integration into the 293 Flp-In T-REx cell line (Figure 2b,c).[6]

First, we tested the editing reaction in cell lines **3** and **4** expressing the hyperactive ADAR1Q fusion. Specifically, we targeted a 5′-U<u>A</u>G codon in the 3′ untranslated region (UTR) of

Chemistry–A European Journal

Communication
doi.org/10.1002/chem.202101985

Chemistry Europe
European Chemical
Societies Publishing

the endogenous GAPDH transcript (Figure 3a). Beside the snap-GAPDH guideRNA, carrying the BG moiety required for covalent reaction with the SNAP-tag, we also applied an NH_2—GAPDH guideRNA as control, comprising of the same sequence and modification pattern, but lacking the BG moiety, thus incompetent of forming a conjugate. The control guideRNA (NH_2—GAPDH) gave no detectable editing, highlighting the requirement for covalent guideRNA attachment to recruit ADAR deaminase activity. This clean negative control is a hallmark of RNA-targeting with the SNAP-ADAR approach.[2,6] Notably, in the absence of the inducer GA_3—AM, no GAPDH editing above the threshold for accurate detection (5%) was detected with the snap-GAPDH guideRNA. However, in presence of 10 μM GA_3—AM in the medium, editing levels of 29±9% and 44±4% were achieved in cell line **3** and **4**, respectively. We therefore estimate the dynamic change of the editing yield by GA_3—AM induction to be >10 fold. Nevertheless, the editing efficiency

stayed clearly below the one obtained with the analogous SNAP-ADAR1Q cell line,[6] which, as expected, yielded high editing independent of GA_3-AM (74±3% versus 76±3%, Figure 3a). This loss of efficiency might be either due to the low expression of the GAI and GID1A fusion proteins compared to the SNAP-ADAR1Q fusion (Figure 2b,c), or it could be a drawback resulting from the necessity to bring not only one, but two proteins, a guideRNA, and an mRNA together for editing. To make sure that the applied GA_3—AM amount was sufficient to induce maximum editing, we determined the dose-response of the editing yield in cell line **4** over a concentration range from 10 nM to 100 μM GA_3—AM (Figure 3b). We determined the EC_{50} to approximately 290 nM, indicating that the editing yield was already close to saturation at 10 μM GA_3—AM. The determined EC_{50} value fits to earlier reports from different applications in literature.[9]
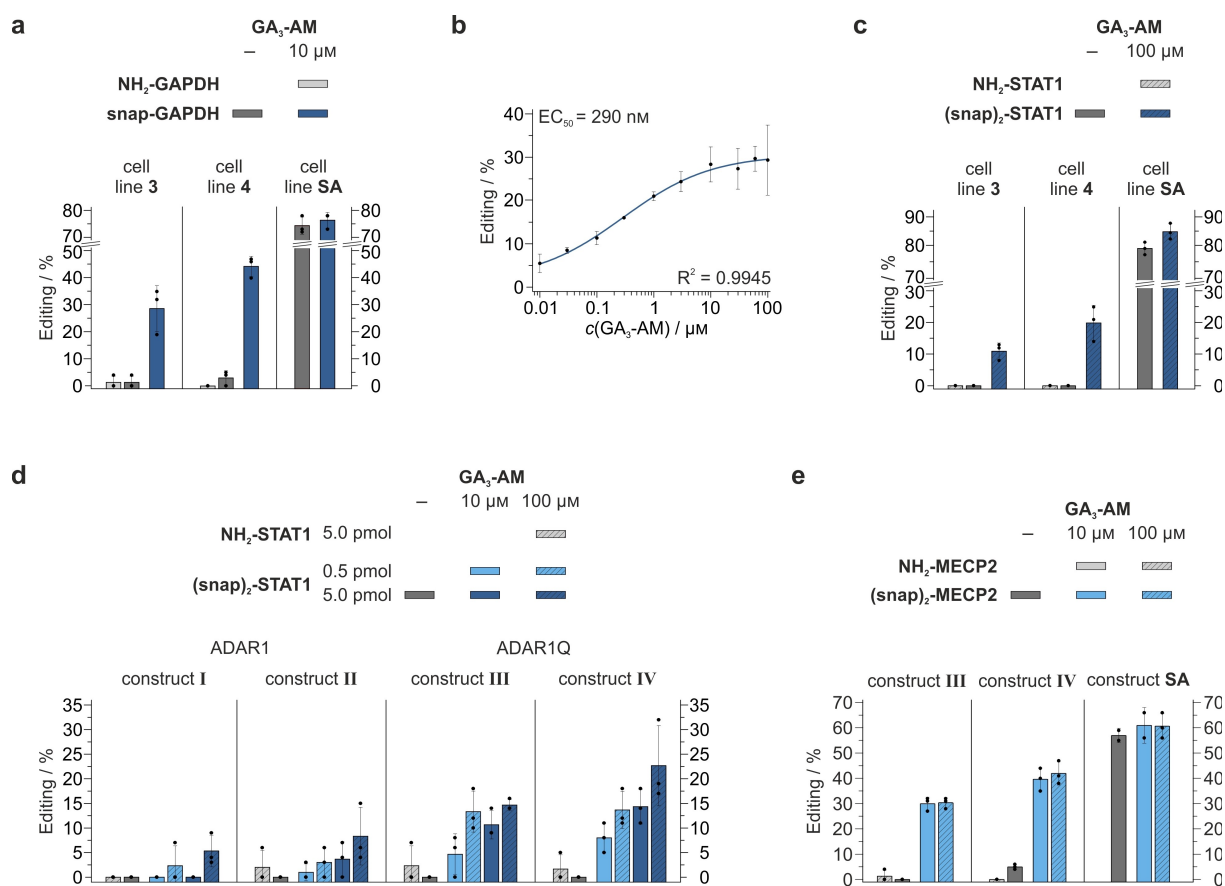

Figure 3. Controlling site-directed RNA editing with gibberellic acid. a) A snap-GAPDH guideRNA (5.0 pmol) targeting a 5′-U<u>A</u>G codon in the 3′-UTR of endogenous GAPDH was transfected into cell lines **3** and **4**, as indicated. An NH_2-guideRNA lacking the BG moiety required for covalent conjugation to the SNAP-GID1A fusion served as negative control. Simultaneously, GA_3—AM (10 μM) was added to the medium, as indicated. RNA editing yields were determined 24 h after transfection by RT-PCR and Sanger sequencing, as described in the Supporting Information. Editing clearly depended on GA_3—AM. b) Determination of the dose-response to GA_3—AM effecting the RNA editing yield in cell line **4**. Editing was performed as described in panel a) on endogenous GAPDH, but with GA_3—AM concentrations ranging from 10 nM to 100 μM. The EC_{50} was determined to 290 nM by applying a logistic fit. c) Analogous to panel a), but with a $(snap)_2$-STAT1 guideRNA (5.0 pmol) targeting the phosphorylation site Tyr(Y)701 (5′-U<u>A</u>U codon) in the endogenous STAT1 transcript in cell lines **3** and **4**, induced with 100 μM GA_3—AM, as indicated. d) Editing of Y701 in endogenous STAT1 in wildtype 293T cells and under transient plasmid transfection of the expression cassettes **I**–**IV** (300 ng/well). Cells were treated with GA_3—AM and guideRNAs in indicated concentrations and amounts. e) Repair of transiently plasmid transfected MECP2 R106Q in wildtype 293T cells under transient plasmid transfection of expression cassette **III** or **IV** with a $(snap)_2$-MECP2 guideRNA (1.0 pmol) targeting the disease relevant R106Q mutation, induced with 10 μM or 100 μM GA_3—AM as indicated. In panel (a)–(e), the data is shown as the mean±s.d. of N=3 independent experiments.

Chemistry–A European Journal

Communication
doi.org/10.1002/chem.202101985

Chemistry
Europe
European Chemical
Societies Publishing

In the past, we found A-to-I RNA editing in the open reading frame (ORF) considerably more challenging compared to editing in the 3′-UTR.[6] Thus, we tested editing in the ORF of the endogenous STAT1 transcript (Figure 3c). Specifically, we designed a guideRNA targeting the phosphorylation site Tyr(Y) 701 (5′-U<u>A</u>U codon), which is important for activation of said transcription factor upon interferon signaling.[14] Again, we found no detectable editing with an NH$_2$-guideRNA lacking the self-labeling moiety. However, we obtained reasonable editing levels when applying the (snap)$_2$–STAT1 guideRNA, able to recruit two SNAP-GID1A proteins per guideRNA. In presence of 100 μM GA$_3$–AM, editing levels of 11±3% and 20±6% were achieved in cell line **3** and **4**, respectively. Again, cell line **4** outperformed cell line **3** by means of editing yields. Due to the lack of detectable editing in absence of GA$_3$–AM, the dynamic range for the induction with GA$_3$–AM was estimated to be very high again. However, the overall editing yields were moderate compared to levels obtained in the analogous SNAP-ADAR1Q 293 Flp-In T-REx cell line (79±2% and 84±3%, without versus with GA$_3$–AM, Figure 3c).[6] We wondered if this was due to the low expression levels of the GAI and GID1A fusion proteins, and if editing could be fostered by stronger expression of the fusion proteins and further optimization of conditions. We thus tested the editing of all four constructs under transient transfection into wildtype 293T cells, and varied the amount of (snap)$_2$–STAT1 guideRNA (0.5 pmol versus 5.0 pmol) and of the inducer (10 μM versus 100 μM, Figure 3d). We made several observations. First, we found wildtype constructs **I** and **II** to give considerably less editing than the hyperactive constructs **III** and **IV**. This is in accordance with literature for the analogous SNAP-ADAR1 293 Flp-In T-REx cells.[6] Second, editing worked better with higher amounts of guideRNA. Third, higher amounts of inducer also fostered editing. Similar to what we had seen for cell line **4** versus cell line **3**, construct **IV** gave better editing yields than construct **III** under most conditions. Taken together, the data suggests that editing yield is boosted by every component that assists in the formation of the tertiary complex (guideRNA+mRNA+two proteins), for example, more guide-RNA, more inducer, and higher protein expression. The latter was supported by Western blot (Figure 2b), showing that more GAI$_{1–92}$–ADAR1Q was expressed by construct **IV** than by construct **III**. Notably, the editing yields at endogenous STAT1 under transient transgene expression did hardly exceed the levels obtained in the stable cell lines (Figure 3c,d). Obviously, the balanced transgene expression in the stable cell lines, even at very low expression levels, is more powerful for targeting endogenous transcripts than the strong, but uneven transgene expression upon plasmid transfection.

Finally, we aimed to repair the R106Q mutation in the transcription factor Methyl CpG Binding Protein 2 (MECP2), which is known to cause Rett syndrome. The underlying G-to-A mutation is located in the DNA binding domain of MECP2 and leads to reduced protein stability and therefore decreased expression levels, as well as reduced binding to heterochromatin.[15] Since healthy expression levels of MECP2 vary between different neural cell types[16] and duplication of MECP2 causes MECP2 duplication syndrome,[17] repair of R106Q

under tight, precisely doseable control at the transcript level is highly desirable. We thus transfected wildtype 293T cells with MECP2 R106Q and either construct **III** or **IV** and tested a guideRNA targeting the R106Q site. Upon induction with 10 μM or 100 μM GA$_3$–AM, we achieved good editing yields for both construct **III** (30±3% and 30±2%, respectively) and construct **IV** (40±5% and 42±5%, respectively, Figure 3e). Contrary to the editing in the STAT1 transcript (Figure 3d), the editing yields for MECP2 in presence of 10 μM and 100 μM GA$_3$–AM were equal, indicating saturation at 10 μM inducer for this target, which fits well to the dose-dependence curve shown for the GAPDH target (Figure 3b). Once again, construct **IV** performed better than construct **III**. Notably, the MECP2 editing levels of our constructs came close to the editing levels with the transfected SNAP-ADAR1Q construct (57±3%, 61±7% and 60±5% without, with 10 μM and with 100 μM GA$_3$–AM, respectively). Importantly, the editing yields obtained with construct **IV** are in the range of editing yields reported to suffice for significant enrichment of heterochromatic MECP2 (37–52%) in vivo in murine neural cells[18] with the λN-ADAR2Q system[19] and therefore might lead to significant diminution of the Rett syndrome phenotype.

In summary, we achieved tight control of site-directed RNA editing by chemically induced dimerization with a small molecule plant hormone. The dimerization occurs promptly (seconds to minutes) after addition of the inducer[9] and elicits a tunable, dose-dependent response. This temporal and dose-dependent control of the RNA editing reaction may break new ground for attractive applications, for example, to trigger targeted editing during embryogenesis after microinjection of all components,[7] to trigger editing to measure RNA lifetimes with RNA timestamp approaches more accurately,[20] or to modulate the pharmacological (adverse) effect of targeted editing.[21] Importantly, we demonstrated that our engineered system, based on the SNAP-ADAR approach in combination with gibberellic acid-induced GID1A-GAI heterodimerization, works not only via transient overexpression, but also under stable genetic integration of the components, which, as we had shown before, reduces artifacts[22] and global off-target editing.[2,6] Furthermore, the editing reaction was strongly dependent on the small molecule – virtually lacking any reaction in the absence of gibberellic acid. Even though the splitting of the SNAP-ADAR editing enzyme into two separate fusion proteins was required to engineer small molecule control, editing of lowly expressed endogenous transcripts was possible in reasonable yields, as demonstrated for the editing of the functionally important phosphorylation site Y701 in STAT1. This is even more remarkable given the comparably low expression levels of the engineered components. We assume that the SNAP-ADAR RNA-targeting approach is particularly suited for the engineering of this kind of small molecule-control as the covalent conjugation of guideRNA and SNAP-tag pre-organizes two components permanently, thus reducing the number of components which need to encounter for editing. Additionally, the disease-relevant R106Q mutation in MECP2 could be repaired to an extent that has been reported[18] to significantly enhance MECP2 function. We expect that the

Chemistry–A European Journal

Communication
doi.org/10.1002/chem.202101985

Chemistry Europe
European Chemical
Societies Publishing

approach can be further improved, for example the editing yields may be amplified by optimizing the expression levels of the fusion proteins. Furthermore, the approach could be extended by one- or two photon-decaging of gibberellic acid[10] to enable spatiotemporal control in the future.[8,23] Finally, the design principle could be included into further tools that apply RNA-guided proteins to manipulate the (epi)transcriptome[24] or could be integrated into existing SNAP-tag-based sensors[25] to include a further layer of control.

Experimental Section

Detailed experimental procedures for Western Blotting (including full blots) and editing experiments, as well as further details on constructs **I–IV**, the generation of stable cell lines and guideRNAs along with their sequences can be found in the Supporting Information.

Conflict of Interest

T.S. holds patents on site-directed RNA editing.

[1] H. A. Rees, D. R. Liu, *Nat. Rev. Genet.* **2018**, *19*, 770–788.
[2] P. Vogel, T. Stafforst, *Curr. Opin. Biotechnol.* **2019**, *55*, 74–80.
[3] T. Stafforst, M. F. Schneider, *Angew. Chem. Int. Ed.* **2012**, *51*, 11166–11169; *Angew. Chem.* **2012**, *124*, 11329–11332.
[4] K. Nishikura, *Annu. Rev. Biochem.* **2010**, *79*, 321–349.
[5] A. Keppler, S. Gendreizig, T. Gronemeyer, H. Pick, H. Vogel, K. Johnsson, *Nat. Biotechnol.* **2003**, *21*, 86–89.
[6] P. Vogel, M. Moschref, Q. Li, T. Merkle, K. D. Selvasaravanan, J. B. Li, T. Stafforst, *Nat. Methods* **2018**, *15*, 535–538.
[7] A. Hanswillemenke, T. Kuzdere, P. Vogel, G. Jékely, T. Stafforst, *J. Am. Chem. Soc.* **2015**, *137*, 15875–15881.
[8] S. Voß, L. Klewer, Y.-W. Wu, *Curr. Opin. Chem. Biol.* **2015**, *28*, 194–201.
[9] T. Miyamoto, R. DeRose, A. Suarez, T. Ueno, M. Chen, T.-p. Sun, M. J. Wolfgang, C. Mukherjee, D. J. Meyers, T. Inoue, *Nat. Chem. Biol.* **2012**, *8*, 465–470.
[10] K. M. Schelkle, T. Griesbaum, D. Ollech, S. Becht, T. Buckup, M. Hamburger, R. Wombacher, *Angew. Chem. Int. Ed.* **2015**, *54*, 2825–2829; *Angew. Chem.* **2015**, *127*, 2867–2871.
[11] J. H. Kim, S.-R. Lee, S.-H. Li, H.-J. Park, J.-H. Park, K. Y. Lee, M.-K. Kim, B. A. Shin, S.-Y. Choi, *PLoS One* **2011**, *6*, e18556.
[12] A. Kuttan, B. L. Bass, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E3295–E3304.
[13] F. Yao, T. Svensjö, T. Winkler, M. Lu, C. Eriksson, E. Eriksson, *Hum. Gene Ther.* **1998**, *9*, 1939–1950.
[14] J. E. Darnell, I. M. Kerr, G. R. Stark, *Science* **1994**, *264*, 1415–1421.
[15] J. R. Sinnamon, S. Y. Kim, G. M. Corson, Z. Song, H. Nakai, J. P. Adelman, G. Mandel, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E9395–E9402.
[16] M. D. Shahbazian, B. Antalffy, D. L. Armstrong, H. Y. Zoghbi, *Hum. Mol. Genet.* **2002**, *11*, 115–124.
[17] H. Van Esch, M. Bauters, J. Ignatius, M. Jansen, M. Raynaud, K. Hollanders, D. Lugtenberg, T. Bienvenu, L. R. Jensen, J. Gécz, C. Moraine, P. Marynen, J.-P. Fryns, G. Froyen, *Am. J. Hum. Genet.* **2005**, *77*, 442–453.
[18] J. R. Sinnamon, S. Y. Kim, J. R. Fisk, Z. Song, H. Nakai, S. Jeng, S. K. McWeeney, G. Mandel, *Cell Rep.* **2020**, *32*, 107878.
[19] M. F. Montiel-Gonzalez, I. Vallecillo-Viejo, G. A. Yudowski, J. J. C. Rosenthal, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 18285–18290.
[20] S. G. Rodriques, L. M. Chen, S. Liu, E. D. Zhong, J. R. Scherrer, E. S. Boyden, F. Chen, *Nat. Biotechnol.* **2020**, *39*, 320–325.
[21] V. M. Rivera, T. Clackson, S. Natesan, R. Pollock, J. F. Amara, T. Keenan, S. R. Magari, T. Phillips, N. L. Courage, F. C. Jr, D. A. Holt, M. Gilman, *Nat. Med.* **1996**, *2*, 1028–1032.
[22] P. Vogel, A. Hanswillemenke, T. Stafforst, *ACS Synth. Biol.* **2017**, *6*, 1642–1649.
[23] N. Ankenbruck, T. Courtney, Y. Naro, A. Deiters, *Angew. Chem. Int. Ed.* **2018**, *57*, 2768–2798; *Angew. Chem.* **2018**, *130*, 2816–2848.
[24] a) X.-M. Liu, J. Zhou, Y. Mao, Q. Ji, S.-B. Qian, *Nat. Chem. Biol.* **2019**, *15*, 865–871; b) K. Rau, L. Rösner, A. Rentmeister, *RNA* **2019**, *25*, 1311–1323; c) B. S. Zhao, I. A. Roundtree, C. He, *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 31–42.
[25] a) Q. Yu, L. Xue, J. Hiblot, R. Griss, S. Fabritz, C. Roux, P.-A. Binz, D. Haas, J. G. Okun, K. Johnsson, *Science* **2018**, *361*, 1122–1126; b) T. Podewin, J. Ast, J. Broichhagen, N. H. F. Fine, D. Nasteska, P. Leippe, M. Gailer, T. Buenaventura, N. Kanda, B. J. Jones, C. M'Kadmi, J.-L. Baneres, J. Marie, A. Tomas, D. Trauner, A. Hoffmann-Röder, D. J. Hodson, *ACS Cent. Sci.* **2018**, *4*, 166–179.