Radically truncated MeCP2 rescues Rett syndrome–like neurological defects

Rebekah Tillotson1, Jim Selfridge1, Martha V. Koerner1, Kamal K. E. Gadalla2,3, Jacky Guy1, Dina De Sousa1, Ralph D. Hector2, Stuart R. Cobb2 & Adrian Bird1

Heterozygous mutations in the X-linked MECP2 gene cause the neurological disorder Rett syndrome1. The methyl-CpG-binding protein 2 (MeCP2) protein is an epigenetic reader whose binding to chromatin primarily depends on 5-methylcytosine2,3. Functionally, MeCP2 has been implicated in several cellular processes on the basis of its reported interaction with more than 40 binding partners4, including transcriptional co-repressors (for example, the NCoR/SMRT complex5), transcriptional activators6, RNA7, chromatin remodellers8,9, microRNA-processing proteins10 and splicing factors11. Accordingly, MeCP2 has been cast as a multi-functional hub that integrates diverse processes that are essential in mature neurons12. At odds with the concept of broad functionality, missense mutations that cause Rett syndrome are concentrated in two discrete clusters coinciding with interaction sites for partner macromolecules: the methyl-CpG binding domain13 and the NCoR/SMRT interaction domain5. Here we test the hypothesis that the single dominant function of MeCP2 is to physically connect DNA with the NCoR/SMRT complex, by removing almost all amino-acid sequences except the methyl-CpG binding and NCoR/SMRT interaction domains. We find that mice expressing truncated MeCP2 lacking both the N- and C-terminal regions (approximately half of the native protein) are phenotypically near-normal; and those expressing a minimal MeCP2 additionally lacking a central domain survive for over one year with only mild symptoms. This minimal protein is able to prevent or reverse neurological symptoms when introduced into MeCP2-deficient mice by genetic activation or virus-mediated delivery to the brain. Thus, despite evolutionary conservation of the entire MeCP2 protein sequence, the DNA and co-repressor binding domains alone are sufficient to avoid Rett syndrome-like defects and may therefore have therapeutic utility.

The amino-acid sequence of MeCP2 is highly conserved throughout vertebrate species (Fig. 1a), suggesting that most of the protein is under evolutionary selection. Full-length MeCP2 is reported to interact with multiple binding partners and has been implicated in several cellular pathways required for neuronal function12. Rett syndrome (RTT)-causing missense mutations, however, are concentrated in the methyl-CpG binding domain (MBD) and NCoR/SMRT interaction domain (NID)—a small part of the protein—whereas numerous polymorphisms elsewhere in the protein are found in the general population, suggesting that other regions may be dispensable (Fig. 1a). To test whether the MBD and NID might be sufficient for MeCP2 function, we generated mouse lines expressing a stepwise series of deletions of MeCP2. The three regions removed were sequences N-terminal to the MBD (N), C-terminal to the NID (C) and the intervening amino acids between these domains (I) (Fig. 1b). The MeCP2 gene has four exons, with transcripts alternatively spliced to produce two isoforms that differ only at the extreme N termini14. To conserve the MeCP2 gene structure in the knock-in mice, exons 1 and 2 and the first 10 base pairs (bp) of exon 3 (splice acceptor site) were retained, resulting in the inclusion of 29 and 12 N-terminal amino acids from isoforms e1 and e2, respectively (Extended Data Fig. 1a, b). A C-terminal enhanced green fluorescent protein (eGFP) tag was added to facilitate detection and recovery (Fig. 1b). We defined the MBD as residues 72–173 and the NID as residues 272–312 (Extended Data Fig. 1c, d). The intervening region of the ∆NIC allele was replaced by a nuclear localization signal (NLS) from SV40 virus, connected by a short flexible linker. The proportions of native MeCP2 protein sequence retained in ∆N, ∆NC and ∆NIC were 88%, 52% and 32%, respectively.

We tested whether the truncated MeCP2 proteins retained the ability to interact with methylated DNA and the NCoR/SMRT co-repressor complex using cell–culture-based assays. They each immunoprecipitated endogenous NCoR/SMRT complex components when overexpressed in HeLa cells, whereas this interaction was abolished in the negative control NID mutant R306C (Extended Data Fig. 2a). They also localized to methyl-CpG-rich heterochromatic foci in mouse fibroblasts, which is dependent on both DNA methylation2,15 and MBD functionality16, whereas the negative control MBD mutant (R111G) was diffusely distributed (Extended Data Fig. 2b). Finally, we tested whether the truncated derivatives were able to recruit TBL1X, an NCoR/SMRT complex subunit that interacts directly with MeCP2 (refs 5, 17), to methylated DNA (Extended Data Fig. 2c). TBL1X–mCherry accumulates in the cytoplasm, but it is efficiently recruited to heterochromatic foci in the presence of co-expressed wild-type (WT) MeCP2 (ref. 5). All three derivative proteins successfully bridged DNA with TBL1X–mCherry in vivo, whereas the negative control NID mutant (R306C) could not do so (Extended Data Fig. 2c). All truncated proteins therefore retained

---

1 The Wellcome Centre for Cell Biology, University of Edinburgh, Michael Swann Building, King’s Buildings, Max Born Crescent, Edinburgh EH9 3BF, UK. 2 Institute of Neuroscience and Psychology, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G12 8QQ, UK. 3 Pharmacology Department, Faculty of Medicine, Tanta University, Tanta 31527, Egypt.

Figure 1 | Stepwise truncation of MeCP2 protein to retain only the MBD and NID. a. Human MeCP2 protein sequence with the MBD and the NID, annotated to show single nucleotide polymorphisms in males in the general population (black lines) and RTT-causing missense mutations (red lines). Sequence identity between human and other vertebrate MeCP2 proteins is shown by purple bars and insertions by dark lines. b. Schematic diagram of the deletion series based on the mouse e2 isoforms that were generated in this study, compared with WT-eGFP20.
were compared using Kolmogorov–Smirnov tests: day 1, P = 0.003; day 2, P = 0.402; and day 3, P = 0.015. Crushed performance (learning/worsening) over the 3 days of the experiment was determined using Friedman tests: WT animals, P = 0.601; ΔNIC animals, P = 0.003.

Figure 3 | Additional deletion of the intervening region leads to protein instability and mild RTT-like symptoms. a, Western blot analysis of whole-brain extract showing protein sizes and abundance of MeCP2 in ΔNIC mice and WT-eGFP controls, detected using a GFP antibody. Histone H3 was used as a loading control. *A non-specific band detected by the GFP antibody. For gel source data, see Supplementary Information. b, Flow cytometry analysis of protein levels in nuclei from whole brain (All) and the high-NeuN subpopulation (Neurons) in ΔNIC mice (n = 3) and WT-eGFP controls (n = 3), detected using eGFP fluorescence (AU, arbitrary units). Graph shows mean ± s.e.m. and genotypes were compared by t-test: All, ***P = 0.0002; and Neurons, **P = 0.0001. c, Quantitative PCR analysis of mRNA prepared from whole brain of ΔNIC mice (n = 3) and WT-eGFP controls (n = 3). MeCP2 transcript levels were normalized to cyclophilin A mRNA. Graph shows mean ± s.e.m. (relative to WT-eGFP) and genotypes were compared by t-test: P = 0.110. d, Phenotypic severity scores of ΔNIC mice (n = 10) compared with WT littermates (n = 10) over one year. Graph shows mean scores ± s.e.m. MeCP2-null data (n = 12) are shown for comparison. e, Kaplan–Meier plots showing survival of the cohort shown in d. One ΔNIC animal died at 43 weeks, after receiving phenotypic scores of ≤ 2.5. MeCP2-null data (n = 24) are shown for comparison. f-h, Behavioural analysis of a separate cohort performed at 20 weeks of age: ΔNIC (n = 10) compared with WT littermates (n = 10). Graphs show individual values and medians, and statistical significance as follows: not significant (NS) P > 0.05, *P < 0.05. c, Time spent in the closed and open arms of the elevated plus maze during a 15 min trial. Genotypes were compared using Kolmogorov–Smirnov tests: ΔNIC closed arms, P = 0.988; and open arms, P = 0.759; ΔNC closed arms, P = 0.950; and open arms, P = 0.932. f, Time spent in the central region of the open field test was measured during a 20 min trial. Genotypes were compared using t-tests: ΔNIC, P = 0.822; ΔNC, *P = 0.079. g, Mean latency to fall from the accelerating rotarod in four trials was calculated for each of the 3 days of the experiment. Genotypes were compared using Kolmogorov–Smirnov tests: ΔNIC day 1, P = 0.759; day 2, P = 0.401; and day 3, P = 0.055; ΔNC day 1, P = 0.988; day 2, P = 0.401; and day 3, P = 0.759.

Figure 2 | Deletion of the MeCP2 N- and C-terminal regions has minimal phenotypic consequence. a, b, Phenotypic severity scores of hemizygous male ΔN mice (a, n = 10) and ΔNC mice (b, n = 10), each compared with their WT littermates (n = 10) over one year. Graphs show mean scores ± s.e.m. Mecp2-null data (n = 12) are shown for comparison. c, d, Kaplan–Meier plots of survival of the cohorts shown in a and b. Mecp2-null data (n = 24) are shown for comparison. e-g, Behavioural analysis of separate cohorts performed at 20 weeks of age: ΔN (n = 10) and ΔNC mice (n = 10 for open field/rotarod; 11 for elevated plus maze), each compared with their WT littermates (n = 10). Graphs show individual values and medians, and statistical significance as follows: not significant (NS) P > 0.05, *P < 0.05. c, Time spent in the closed and open arms of the elevated plus maze during a 15 min trial. Genotypes were compared using Kolmogorov–Smirnov tests: ΔN closed arms, P = 0.988; and open arms, P = 0.759; ΔNC closed arms, P = 0.950; and open arms, P = 0.932. c, Time spent in the central region of the open field test was measured during a 20 min trial. Genotypes were compared using t-tests: ΔNIC, P = 0.822; ΔNC, *P = 0.020. g, Mean latency to fall from the accelerating rotarod in four trials was calculated for each of the 3 days of the experiment. Genotypes were compared using Kolmogorov–Smirnov tests: ΔNIC day 1, P = 0.759; day 2, P = 0.401; and day 3, P = 0.055; ΔNC day 1, P = 0.988; day 2, P = 0.401; and day 3, P = 0.759.

the ability to bind methylated DNA and the NCoR/SMRT complex simultaneously.

We generated ΔN and ΔNC knock-in mice by replacing the endogenous Mecp2 allele in embryonic stem (ES) cells, which were used to produce germline-transmitting chimaeras (Extended Data Fig. 3). Truncated proteins were expressed at approximately WT levels in brain and in neurons (Extended Data Fig. 4a–d). To assess phenotypes, knock-in mice were crossed onto a C57BL/6 background (for four generations) and cohorts underwent weekly phenotypic scoring18,19 or behavioural analysis. Although heterozygous female mice are the genetic model for RTT, phenotypes develop late and are mild in the case of hypomorphic Mecp2 mutations20,21. Hemizygous males provide a more sensitive assay of MeCP2 function: MeCP2-null males exhibit severe phenotypes that develop shortly after weaning and median survival is nine weeks21. Both ΔN and ΔNC male mice were viable, fertile and showed phenotypic scores indistinguishable from WT littermates over one year (Fig. 2a–d). ΔN mice had normal body weight (Extended Data Fig. 4e), whereas ΔNC mice were slightly heavier than WT littermates (Extended Data Fig. 4f). This difference was absent in a
more obtunded cohort (Extended Data Fig. 4g), consistent with previous observations that body weight of MeCP2 mutants is affected by genetic background21.

At 20 weeks of age, cohorts were tested for RTT-like behaviours: hypointervisibility, decreased anxiety, and reduced motor abilities. Neither activity (distance travelled in an open field; Extended Data Fig. 4h) nor anxiety (time spent in the open arms of the elevated plus maze; Fig. 2e) was abnormal in ΔN and ΔNC mice, although the latter did spend longer in the centre of the open field (Fig. 2f), indicative of mildly decreased anxiety. Motor coordination was assessed using the accelerating rotarod test over 3 days. Whereas mouse models of RTT show impaired performance that was most striking on the third day20,22, ΔN and ΔNC mice were comparable to WT littermates throughout this test (Fig. 2g). Overall, the results suggest that contributions of the N- and C-terminal regions to MeCP2 function are at best subtle. The result is remarkable given the presence of a neurological phenotype in male mice expressing a slightly more severe C-terminal truncation, which lacks residues beyond T308 (ref. 23). The difference may be explained by tamoxifen treatment. The presence of normal levels of mRNA in ΔNIC mice (Fig. 3c) suggested that deletion of the intervening region compromises protein stability. Despite low protein levels, male ΔNIC mice had a normal lifespan (Fig. 3d and Extended Data Fig. 6a). However, phenotypic scoring over one year detected mild neurological phenotypes (Fig. 3d), predominantly gait abnormalities and partial hindlimb clasping. These symptoms were relatively stable throughout the scoring period. ΔNIC mice also weighed ~40% less than their WT littermates (Extended Data Fig. 6b).

Behavioural analysis of a separate cohort at 20 weeks showed decreased anxiety in male ΔNIC mice, signified by reduced time spent in the closed arms of an elevated plus maze (Fig. 3f), although this phenotype was not detected by the open field test (Fig. 3g). No activity phenotype was detected in the open field (Extended Data Fig. 6c), but, consistent with the gait defects detected by weekly scoring, ΔNIC mice displayed declining motor coordination on the accelerating rotarod over 3 days, culminating in a significantly impaired performance on the third day (Fig. 3h). It is noteworthy that ΔNIC animals are much less severely affected than male mice with the mildest common mutation found in RTT patients, R133C, which had a median lifespan of 26 weeks, higher phenotypic scores and a more pronounced reduction in body weight20 (Extended Data Fig. 7). Reduced protein levels may contribute to the relatively mild phenotype, as mice with ~50% levels of full-length MeCP2 have neurological defects24.

To further test ΔNIC functionality, we asked whether late genetic activation could reverse phenotypic defects in symptomatic MeCP2-deficient mice, as has previously been shown with the full-length protein16. Mice that were MeCP2-deficient through insertion of a floxed transpositional STOP cassette in intron 2 of the ΔNIC gene (Extended Data Figs 5 and 8a, b) resembled MeCP2-nulls (Extended Data Fig. 8c, d). This line was crossed with mice carrying a creER transgene (Cre recombinase fused to a modified oestrogen receptor) to enable removal of the STOP cassette upon tamoxifen treatment. Induced expression of ΔNIC after
the onset of symptoms in STOP creER² mice (Fig. 4a) resulted in high levels of Cre-mediated recombination (Extended Data Fig. 9a) and protein levels similar to those of ΔNIC mice (Extended Data Fig. 9b). ΔNIC activation had a pronounced effect on phenotypic progression, relieving neurological symptoms and restoring normal survival (Fig. 4b, c). Separation of the phenotypic scores into the six tested components showed clear reversal of tremor, hypoactivity and gait abnormalities (Extended Data Fig. 9c). In contrast, control STOP mice lacking the creER² transgene developed severe symptoms and failed to survive beyond 26 weeks. Thus, despite its radically reduced length and relatively low abundance, ΔNIC was able to effectively rescue MeCP2-deficient mice from RTT-like phenotypes.

This finding prompted us to explore whether ΔNIC could be used for gene therapy in MeCP2-null mice. A human version of the ΔNIC gene (hΔNIC) (Fig. 4d), driven by a minimal Mecp2 promoter, was tagged with a short Myc epitope (in place of eGFP) and packaged into a self-complementary adeno-associated viral vector (scAAV). Neonatal mice (postnatal day 1–2) injected intra-cranially with this transgene developed severe symptoms and failed to survive beyond 26 weeks. Thus, despite its radically reduced length and relatively low abundance, ΔNIC was able to effectively rescue MeCP2-deficient mice from RTT-like phenotypes.

This finding prompted us to explore whether ΔNIC could be used for gene therapy in MeCP2-null mice. A human version of the ΔNIC gene (hΔNIC) (Fig. 4d), driven by a minimal Mecp2 promoter, was tagged with a short Myc epitope (in place of eGFP) and packaged into a self-complementary adeno-associated viral vector (scAAV). Neonatal mice (postnatal day 1–2) injected intra-cranially with this transgene developed severe symptoms and failed to survive beyond 26 weeks. Thus, despite its radically reduced length and relatively low abundance, ΔNIC was able to effectively rescue MeCP2-deficient mice from RTT-like phenotypes.

This finding prompted us to explore whether ΔNIC could be used for gene therapy in MeCP2-null mice. A human version of the ΔNIC gene (hΔNIC) (Fig. 4d), driven by a minimal Mecp2 promoter, was tagged with a short Myc epitope (in place of eGFP) and packaged into a self-complementary adeno-associated viral vector (scAAV). Neonatal mice (postnatal day 1–2) injected intra-cranially with this transgene developed severe symptoms and failed to survive beyond 26 weeks. Thus, despite its radically reduced length and relatively low abundance, ΔNIC was able to effectively rescue MeCP2-deficient mice from RTT-like phenotypes.

This finding prompted us to explore whether ΔNIC could be used for gene therapy in MeCP2-null mice. A human version of the ΔNIC gene (hΔNIC) (Fig. 4d), driven by a minimal Mecp2 promoter, was tagged with a short Myc epitope (in place of eGFP) and packaged into a self-complementary adeno-associated viral vector (scAAV). Neonatal mice (postnatal day 1–2) injected intra-cranially with this transgene developed severe symptoms and failed to survive beyond 26 weeks. Thus, despite its radically reduced length and relatively low abundance, ΔNIC was able to effectively rescue MeCP2-deficient mice from RTT-like phenotypes.
Design of the truncated MeCP2 proteins. The MBD and NID were defined as residues 72–173 and 272–312, respectively. All three constructs retained the key N-terminal sequences encoded by exons 1 and 2, present in isoforms e1 and e2, respectively. They also include the first three amino acids of exons 3 (EKE) to preserve the splice acceptor site. The intervening region (I) was replaced in ΔNIC by the NLS of SV40 preceded by a flexible linker. The sequence of the NLS was PKKKRKV (DNA sequence: CCCAAGAAAAGCGGAAGGTG) and of the linker was GSSSSG (DNA sequence: GGATCCGGAGGCTCCTGGAG). All three proteins were C-terminally tagged with eGFP connected by a linker. To be consistent with a previous study tagging full-length MeCP2 (ref. 20), the linker sequence CKDPVAT (DNA sequence: TGGACATGATCCGGAGGCTCCTGGAG) was used to connect the C terminus of ΔN to eGFP. To connect the NID to the eGFP tag in ΔNC and ΔNIC, the flexible GSSSSG linker was used instead (DNA sequence: GGAGGATCCGGAGGCTCCTGGAG). For expression in cultured cells, cDNA sequences encoding e2 isoforms of the MeCP2 deletion series were synthesized (GeneArt, Thermo Fisher Scientific) and cloned into the pEGFP-N1 vector (Clontech) using XhoI and NotI restriction sites (NEB). Point mutations (R111G and R306C) were inserted into the WT-eGFP plasmid using a QuickChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies). Primer sequences for R111G were as follows: forward TGGACACGAGAAACCATGAAAGGGAAGGCTCGG and reverse GGCGCACTCCCTTGTTAAGCTTTCGTGTCCA; and for R306C: forward CTCCCCGCGTCTGCAGCTTCTTGATGAGGGA and reverse TCCCCATACAGAAGTGGCAAGAGCAAGGCGGAGG. For ES cell targeting, genomic sequences encoding exons 3 and 4 of the eGFP-tagged truncated proteins were synthesized (GeneArt, Thermo Fisher Scientific) and cloned into the pEGFP-N1 vector and then targeted using Flp-mediated FLP recombination targeting (LoxP sites ‘floxed’) in intron 2.

Cell culture. HeLa and NIH-3T3 cells were originally from a standard repository (such as American Type Culture Collection, ATCC) and maintained in the Bird laboratory for many years. HeLa and NIH-3T3 cells were grown in DMEM (Gibco) supplemented with 10% fetal bovine serum (FBS; Gibco) and 1% penicillin–streptomycin (Gibco). ES cells were grown in Glasgow MEM (Gibco) supplemented with 10% FBS (Gibco, batch tested), 1% non-essential amino acids (Gibco), 1% sodium pyruvate (Gibco), 0.1% β-mercaptoethanol (Gibco) and 1,000 units ml−1 LIF (ESGRO). Low-passage ES cells were a gift from A. Smith, Institute for Stem Cell Research, University of Edinburgh. They were received at passage 12 and frozen in aliquots (mycoplasma-free). Laboratory stocks of HeLa and NIH-3T3 cells were periodically tested for mycoplasma.

Immunoprecipitation. HeLa cells were transfected with pEGFP-N1–MeCP2 plasmids using JetPEI (PolyPlus Transfection) and harvested after 24–48 h. Nuclear extracts were prepared using Benzonase (Sigma) and 150 mM NaCl, and MeCP2–eGFP complexes were captured using GFP-Trap_A beads (Chromotek) as recommended by the manufacturer’s instructions for cultured cells. Genomic DNA was purified from ES cells using Puregene Core Kit A (Qiagen) according to the manufacturer’s instructions for cultured cells. Genomic DNA was digested with restriction enzymes (NEB), separated by agarose gel electrophoresis and transferred onto ZetaProbe membranes (BioRad). DNA probes homologous to either exon 4 or the end of the 3′ homology arm were radioactively labelled with [α-32P]dCTP (Perkin Elmer) using the Prime-a-Gene Labelling System (Promega). Blots were probed overnight, washed and exposed in Phosphorimager cassettes (GE Healthcare) before scanning on a Typhoon FLA 7000. Bands were quantified using ImageQuant software.

Protein levels in whole-brain crude extracts were quantified using western blotting. Extracts were prepared as described previously27, and blots were probed with antibodies against GFP (NEB N296) at a dilution of 1:1,000, followed by L1-COR secondary antibodies (listed above).Histone H3 (Abcam ab1791) was used as a loading control (dilution 1:10,000). Levels were quantified using Image Studio Lite version 4.0 software and compared using t-tests. WT-eGFP mice28 were used as controls. Three biological replicates of each genotype were analysed, except for the reversal experiment where five STOP creER2 + Tmx animals and one STOP + Tmx animal were analysed.

For flow cytometry analysis, fresh brains were harvested from 12-week-old animals and Dounce-homogenized in 5 ml homogenization buffer (320 mM sucrose, 5 mM CaCl2, 3 mM MgCl2, 10 mM Tris HCl pH 7.5, 0.1 mM EDTA, 0.1% NP40, 0.1 mM PMSF, 14.3 mM β-mercaptoethanol, protease inhibitors (Roche)), and 5 ml of 50% OptiPrep gradient centrifugation medium (50% OptiPrep (Sigma D1556:250ML), 5 mM CaCl2, 5 mM MgCl2, 10 mM Tris HCl pH 7.8, 0.1 mM PMSF, 14.3 mM β-mercaptoethanol) was added. This was layered on top of 10 ml of 29% OptiPrep solution (v/v in H2O, diluted from 60% stock) in Ultra-Clear Beckman Coulter centrifuge tubes, and samples were centrifuged at 10,100g for 30 min, 4°C. Pelleted nuclei were resuspended in resuspension buffer (20% glycerol in DPBS ( Gibco) with protease inhibitors (Roche)). For flow cytometry analysis, nuclei were pelleted at 600g (5 min, 4°C), washed in 1 ml PBTB (5% w/v BSA, 0.1% Triton X-100 in DPBS with protease inhibitors (Roche)) and then resuspended in 250 μl PBTB. To stain for NeuN, NeuN antibody (Millipore MAB377) was conjugated to Alexa Fluor 647 (APEX Antibody Labelling Kit, Invitrogen A10475), added at a dilution of 1:125 and incubated under rotation for 45 min at 4°C. Flow cytometry (BD LSRSortrSeA OR using FACSDIVA version 8.0.1 software) was used to obtain the mean eGFP fluorescence for the total nuclei (n = 50,000 per sample) and the high-NeuN (neuronal) subpopulation (n > 8,000 per sample). Three biological replicates of each genotype were analysed. The protein levels of the novel mouse lines were compared with WT-eGFP controls using t-tests. To compare protein levels in WT-eGFP mice with WT littermates, nuclei were also stained with an MeCP2 antibody (Sigma M7443) conjugated to Alexa Fluor 568 (APEX Antibody Labelling Kit, Invitrogen A10494) at a dilution of 1:125.

To determine mRNA levels, RNA was purified and reverse transcribed from brain (harvested at 11 weeks of age); and MeCP2 and cyclophilin A transcripts were analysed by qPCR using LightCycler 480 SW 1.5 software as previously described29. Three biological replicates of each genotype were analysed, each pipetted in triplicate. mRNA levels in ΔNIC mice were compared with WT-eGFP controls using a t-test.
Phenotypic characterization of knock-in mice. Consistent with a previous study\textsuperscript{20}, mice were backcrossed for four generations to reach ~94% C57BL/6j before undergoing phenotypic characterization. Two separate cohorts, each consisting of 10 mutant animals (11 for ∆NC elevated plus maze) and 10 WT littermates, were produced for each novel knock-in line. One cohort was scored and weighed regularly from 4 to 52 weeks of age as previously described\textsuperscript{18,19}. Survival was graphed using Kaplan–Meier plots. Preliminary outbred (75% C57BL/6j) cohorts of ∆NC and ∆NIC strains were also analysed: ∆NC cohort, ∆NC n = 7 and WT littermates, n = 9; ∆NIC cohort, ∆NIC n = 10 and WT littermate, n = 1. Previously published\textsuperscript{20} data for Mecp2-null and R133C-eGFP (both backcrossed onto C57BL/6j) were included for comparison. The second backcrossed cohorts underwent behavioural analysis at 20 weeks of age (see refs 19, 20 for detailed protocols). Tests were performed over a 2 week period: elevated plus maze on day 1, open field test on day 2, and accelerating rotarod test on days 6–9 (1 day of training followed by 3 days of trials). All analysis was performed blind to genotype. Animals were randomly assigned to the two backcrossed cohorts and the order in which they were analysed was randomized. In a previous study, cohorts of this size were sufficient to show statistically significant differences between mice expressing patient mutations and their WT littermates in the same behavioural tests\textsuperscript{20}. This cohort size was therefore considered sufficient to be confident of an absence of the same phenotypes in the asymptomatic novel lines.

Statistical analysis. Growth curves were compared using repeated measures ANOVA (the animals that died within the experimental period—one WT in each ∆NC cohort and one ∆NIC in their cohort—were excluded from this analysis to enable a balanced design). Survival curves were compared using a Mantel–Cox test. For behavioural analysis, when all data fitted a normal distribution (open field centre time and distance travelled), genotypes were compared using t-tests (unpaired, two-tailed). For non-parametric datasets (elevated plus maze time in arms/centre and accelerating rotarod latency to fall), genotypes were compared using Kolmogorov–Smirnov tests. Change in performance over time in the accelerating rotarod test was determined using Friedman tests. All analysis was performed using GraphPad Prism 7 software.

Genetic activation of minimal MeCP2 (∆NIC). Transcriptionally silent minimal Mecp2 (∆NIC) was activated in symptomatic null-like STOP mice following the procedure used in ref. 18. In short, the ∆NIC Mecp2 allele was inactivated by the retention of the NeoSTOP cassette in intron 2 by mating chimaeras with WT females instead of CMV-cre deleter mice. Resulting STOP/+ females were crossed with heterozygous creER\textsuperscript{T2} transgenic males (JAX stock 004682) to produce males of four genotypes (87.5% C57BL/6j). A cohort consisting of all four genotypes WT (n = 4), WT creER\textsuperscript{T2} (n = 4), STOP (n = 9) and STOP creER\textsuperscript{T2} (n = 9), was scored and weighed weekly from 4 weeks of age. When STOP and STOP creER\textsuperscript{T2} mice displayed RTT-like symptoms, all individuals were given a series of tamoxifen injections: one at 6 weeks of age, one at 7 weeks of age, and five daily injections at 8 weeks of age, each at a dose of 100 μg (body weight). Brain tissue from tamoxifen-treated STOP creER\textsuperscript{T2} (n = 8), WT (n = 1) and WT creER\textsuperscript{T2} (n = 1) animals was harvested at 28 weeks of age (after successful symptom reversal in STOP creER\textsuperscript{T2} mice) for biochemical analysis. Brain tissue from one tamoxifen-treated STOP mouse (harvested at its humane end-point) was also included in the biochemical analysis (methods described above).

Vector delivery of minimal MeCP2 (∆NIC). The AAV vector expressing minimal Mecp2 (∆NIC) was tested in Mecp2-null and WT mice maintained on a C57BL/6j background. Self-complementary AAV (scAAV) particles, comprising AAV2 ITR-flanked genomes packaged into AAV9 capsids, were generated at the University of North Carolina Vector Core. Particles were produced as previously described\textsuperscript{42} by transfection of HEK293 cells with helper plasmids (pXX6-80, pGSK2/9) and a plasmid containing the ITR-flanked construct in the presence of polyethyleneimine (Polysciences). For translational relevance, the ∆NIC-expressing construct used the equivalent human MECP2-e1 coding sequence tagged with a small C-terminal Myc epitope to replace the eGFP tag used in knock-in experiments. The transgene was under the control of an endogenous Mecp2 promoter fragment as previously described\textsuperscript{21}. Vector was formulated in high-salt PBS (containing 350 mM total NaCl) supplemented with 5% sorbitol. Virus (3 μL per site; dose = 1 × 10\textsuperscript{11} viral genomes per mouse) was injected bilaterally into the neuropil of unaanaesthetized postnatal day (P)1/2 mice, as described previously\textsuperscript{42}. Control injections used the same diluent without vector (vehicle control). Animals were selected at random for injection with the vehicle or vector control. The injected pups were returned to the home cage and assessed weekly from 5 weeks of age, as described above (performed blind to genotype). Cohorts were as follows: WT + vehicle (n = 15); Mecp2-null + vehicle (n = 20; 19 of which were scored as one reached its humane end-point early); WT + ∆NIC (n = 14); and Mecp2-null + ∆NIC (n = 17).

To validate the expression of virally delivered ∆NIC, mice were deeply anaesthetized with pentobarbitone (50 mg, intraperitoneally) and transcardially perfused with 4% (w/v) paraformaldehyde (in 0.1 M PBS). A vibrating microtome (Leica VT1200) was used to obtain 70 μm sections of the brain. Sections were washed three times in 0.3 M PBS followed by blocking using 5% (w/v) normal goat serum with 0.3% (v/v) Triton X-100 in 0.3 M PBS (PBST) for 1 h at room temperature. Samples were then incubated for 48 h on a shaker at 4 °C with primary antibodies against Myc (Abcam ab19106; 1:500 dilution) and NeuN (Abcam 104224; 1:500 dilution). Samples were washed three times with 0.3 M PBST and incubated with secondary antibodies (Alexa Fluor 594 goat anti-rabbit (Abcam 150080; 1:500 dilution) and Alexa Fluor 647 goat anti-mouse (Stratech scientific LTD, 115-605-003JFR; 1:500 dilution)) at 4 °C overnight. Finally, sections were incubated with DAPI (Sigma; 1:1,000) for 30 min at room temperature before mounting with Fluoroshield with DAPI (Sigma, F6057). Z-series at 0.6–1.3 μm intervals were captured using a Zeiss LSM710 or Zeiss Axiovert LSM510 laser confocal microscope (40× objective). To estimate transduction efficiency, the ratio of Myc-positive nuclei to DAPI-stained nuclei was calculated from random sections of hippocampus (CA1), layer 5 of primary motor cortex, thalamus, hypothalamus and brainstem (n = 3 mice per genotype, 27 fields from each brain region). Mecp2-null + ∆NIC mice were analysed after reaching their humane endpoints (aged 33, 35 and 36 weeks). WT + ∆NIC mice were harvested for analysis at 4 weeks of age.

Data availability. All data are available from the corresponding author upon reasonable request. Source Data underlying all graphs and full scans of all western and Southern blots are included in the online version of the paper.

31. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas9 systems. Science 339, 819–823 (2013).
32. Clément, N. & Grieger, J. C. Manufacturing of recombinant adeno-associated viral vectors for clinical trials. Mol. Ther. Methods Clin. Dev. 3, 16002 (2016).
33. Gadalla, K. K. E. et al. Improved survival and reduced phenotypic severity following AAV9/MECP2 gene transfer to neonatal and juvenile male Mecp2 knockout mice. Mol. Ther. 21, 18–30 (2013).
34. Tao, J. et al. Phosphorylation of MeCP2 at serine 80 regulates its chromatin association and neurological function. Proc. Natl Acad. Sci. USA 106, 4882–4887 (2009).
35. Ebert, D. H. et al. Activity-dependent phosphorylation of MeCP2 threonine 308 regulates interaction with NCoR. Nature 499, 341–345 (2013).
36. Ho, K. L. et al. MeCP2 binding to DNA depends upon hydration at methyl-CpG. Mol. Cell 29, 525–531 (2008).
37. Lyst, M. J., Connelly, J., Merusi, C. & Bird, A. Sequence-specific DNA binding by AT-hook motifs in MeCP2. FEBS Lett. 590, 2927–2933 (2016).
**Extended Data Figure 1 | Design of the MeCP2 deletion series.**

**a.** Diagram of the genomic DNA sequences encoding WT and ΔNIC MeCP2, showing the retention of the extreme N-terminal amino acids encoded in exons 1 and 2 and the first 10 bp of exon 3, the deletion of the N- and C-terminal regions, the replacement of the intervening region with a linker and SV40 NLS, and the addition of the C-terminal eGFP tag. Colour key: 5′ untranslated region, white; MBD, blue; NID, pink; other MeCP2 coding regions, grey; SV40 NLS, orange; linkers, dark grey; eGFP, green. **b.** The N-terminal ends of the sequences of all three truncated proteins (e1 and e2 isoforms) showing the fusion of the extreme N-terminal amino acids to the MBD (starting with P72). **c, d.** Protein sequence alignment of the MBD (**c**) and NID (**d**) regions using ClustalWS, shaded according to BLOSUM62 score. Both alignments are annotated with RTT-causing missense mutations (http://mecp2.chw.edu.au/) (red), activity-dependent phosphorylation sites (orange), sequence conservation, interaction domains and known/predicted (https://npsa-prabivibicpfr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_phd.html) structure. Interaction sites: methyl-CpG binding domain (residues 78–162 (ref. 13)), AT hook 1 (residues 183–195 (ref. 37)), AT hook 2 (residues 257–272 (ref. 28)), NCoR/SMRT interaction domain (residues 285–309 (ref. 5)). The bipartite NLS is also shown (residues 253–256 and 266–271). The regions retained in ΔNIC are: MBD resides 72–173 (highlighted by the blue shading in **c**) and NID resides 272–312 (highlighted by the pink shading in **d**). Residue numbers correspond to that of mammalian e2 isoforms.
Extended Data Figure 2 | Truncated MeCP2 proteins retain the ability to bind methylated DNA and the NCoR/SMRT complex. 

**a**, eGFP-tagged truncated proteins immunoprecipitate components of the NCoR/SMRT co-repressor complex: NCoR, HDAC3 and TBL1XR1. WT and R306C were used as positive and negative controls for binding, respectively. In, input; IP, immunoprecipitate. For gel source data, see Supplementary Information. 

**b**, Representative images showing localization of eGFP-tagged truncated MeCP2 proteins to mCpG-rich heterochromatic foci when they are overexpressed in mouse fibroblasts (NIH-3T3 cells). WT and R111G were used as controls to show focal and diffuse localization, respectively. Scale bars, 10 μm. 

**c**, Representative images showing recruitment of TBL1X–mCherry to heterochromatin by eGFP-tagged truncated proteins when they are co-overexpressed in NIH-3T3 cells. WT and R306C were used as positive and negative controls for TBL1X–mCherry recruitment, respectively. Scale bars, 10 μm. 

Quantification (right) shows the percentage of cells with focal TBL1X–mCherry localization, evaluated relative to WT using Fisher’s exact tests: R306C, ****P < 0.0001; ΔN, P = 0.071; ΔNC, P = 0.604; ΔNIC, P = 0.460. Total numbers of cells counted: WT, n = 117; R306C, n = 119; ΔN, n = 113; ΔNC, n = 119; ΔNIC, n = 125; over three independent transfection experiments.
Extended Data Figure 3 | Generation of ΔN and ΔNC mice. Diagrammatic representation of ΔN (a) and ΔNC (b) knock-in mouse line generation. The endogenous MeCP2 allele was targeted in male ES cells. The site of Cas9 cleavage in the WT sequence is shown by the scissors symbol (used for production of ΔN knock-in ES cells). The selection cassette was removed in vivo by crossing chimaeras with deleter (CMV-cre) transgenic mice. Southern blot analysis shows correct targeting of ES cells and successful cassette deletion in the knock-in mice. The solid black line represents the sequence encoded in the targeting vector and the dotted lines indicate the flanking regions of mouse genomic DNA. For gel source data, see Supplementary Information.
Extended Data Figure 4  See next page for caption.
Extended Data Figure 4 | ΔN and ΔNC knock-in mice express truncated proteins at approximately WT levels and display minimal phenotypes. a, Western blot analysis of whole-brain extract showing protein sizes and abundance of MeCP2 in ΔN and ΔNC mice and WT-eGFP controls, detected using a GFP antibody. Histone H3 was used as a loading control. *A non-specific band detected by the GFP antibody. For gel source data, see Supplementary Information. b, Flow cytometry analysis of protein levels in nuclei from whole brain (All) and the high-NeuN subpopulations (Neurons) in WT-eGFP (n = 3), ΔN (n = 3) and ΔNC (n = 3) mice, detected using eGFP fluorescence. Graph shows mean ± s.e.m. and genotypes were compared with WT-eGFP controls by t-test: All, P = 0.338; ΔN, *P = 0.003; and Neurons ΔN, P = 0.672; ΔNC, *P = 0.014. au, arbitrary units. c, Flow cytometry analysis of protein levels in WT (n = 3) and WT-eGFP (n = 3) mice, detected using an MeCP2 antibody. Graph shows mean ± s.e.m. and genotypes were compared by t-test: All, P = 0.214; and Neurons, P = 0.085.

d, Example histogram (of one WT-eGFP sample) showing how the Neuronal subpopulation was defined according to NeuN-AF647 staining. e–g, Growth curves of the backcrossed scoring cohorts (e, f; see Fig. 2a–d) and an outbred (g; 75% C57BL/6j) cohort of ΔNC mice (n = 7) and WT littermates (n = 9). Graphs show mean values ± s.e.m. Genotypes were compared using repeated measures ANOVA: ΔN, P = 0.362; ΔNC, ****P < 0.0001; ΔNC (outbred), P = 0.739. Mecp2-null data (n = 20) are shown for comparison to the backcrossed cohorts. h, Behavioural analysis of ΔN (n = 10) and ΔNC mice (n = 10) each compared with their WT littermates (n = 10) at 20 weeks of age (see Fig. 2e–g). Total distance travelled in the open field test was measured during a 20 min trial. Graphs show individual values and medians. Genotypes were compared using t-tests: ΔN, P = 0.691; ΔNC, P = 0.791. NS, not significant.
Extended Data Figure 5 | Generation of ΔNIC and STOP mice.
Diagrammatic representation of ΔNIC and STOP mouse line generation. The endogenous MeCP2 allele was targeted in male ES cells. The site of Cas9 cleavage in the WT sequence is shown by the scissors symbol. The selection cassette was removed in vivo by crossing chimaeras with deleter (CMV-cre) transgenic mice to produce constitutively expressing ΔNIC mice, or retained to produce STOP mice. Southern blot analysis shows correct targeting of ES cells and successful cassette deletion in the ΔNIC knock-in mice. The solid black line represents the sequence encoded in the targeting vector and the dotted lines indicate the flanking regions of mouse genomic DNA. For gel source data, see Supplementary Information.
Extended Data Figure 6 | ΔNIC mice have a normal lifespan and no activity phenotype but decreased body weight. a, Kaplan–Meier plot showing survival of an outbred (75% C57BL/6J) cohort of ΔNIC mice (n = 10) and their WT littermate (n = 1). b, Growth curve of the backcrossed cohort used for phenotypic scoring (see Fig. 3d, e). Graph shows mean ± s.e.m. Genotypes were compared using repeated measures ANOVA ****P < 0.0001. Mecp2-null data (n = 20) are shown for comparison. c, Behavioural analysis of ΔNIC mice (n = 10) compared with their WT littermates (n = 10) at 20 weeks of age (see Fig. 3f–h). Total distance travelled the open field test was measured during a 20 min trial. Graph shows individual values and medians. Genotypes were compared using a t-test. P = 0.333. NS, not significant.
Extended Data Figure 7 | ΔNIC mice have a less severe phenotype than the mildest mouse model of RTT, R133C. a–c, Repeat presentation of phenotypic analysis of ΔNIC mice and WT littermates in Fig. 3d, e and Extended Data Fig. 6b, this time including eGFP-tagged R133C mice (n = 10) for comparison. a, Phenotypic severity scores (mean ± s.e.m.). b, Growth curve (mean ± s.e.m.). c, Survival (Kaplan–Meier plot).
Extended Data Figure 8 | STOP mice with transcriptionally silenced ΔNIC resemble Mecp2-nulls. a, Western blot analysis of whole-brain extract showing protein sizes and abundance of MeCP2 in STOP mice compared with WT-eGFP and ΔNIC controls, detected using a GFP antibody. Histone H3 was used as a loading control. *A non-specific band detected by the GFP antibody. For gel source data, see Supplementary Information. b, Flow cytometry analysis of protein levels in nuclei from whole brain (All) and the high-NeuN subpopulation (Neurons) in WT-eGFP (n = 3), ΔNIC (n = 3) and STOP (n = 3) mice, detected using eGFP fluorescence. Graph shows mean ± s.e.m. and genotypes were compared using t-tests: ****P < 0.0001. au, arbitrary units. c, Phenotypic scoring of STOP mice (n = 22) compared with published Mecp2-null data (n = 12)\(^{20}\). Graph shows mean scores ± s.e.m. d, Kaplan–Meier plot showing survival of STOP mice (n = 14) compared with Mecp2-null data (n = 24)\(^{20}\).
Successful activation of ΔNIC in tamoxifen-injected STOP creER mice leads to symptom reversal. a, Southern blot analysis of genomic DNA to determine the level of recombination mediated by CreER² in tamoxifen-injected (+ Tmx) STOP creER² animals. WT, WT creER², ΔNIC and STOP samples, with or without tamoxifen injection, were included as controls. (Bsu36I digestion, see restriction map in Extended Data Fig. 5.) b, Protein levels in tamoxifen-injected STOP creER² animals were determined using western blotting (upper, n = 5) and flow cytometry (lower, n = 3). Constitutively expressing ΔNIC mice (n = 3) were used for comparison. Graphs show mean values ± s.e.m. (quantification by western blotting is shown normalized to ΔNIC). Genotypes were compared using t-tests: western blotting, P = 0.434; flow cytometry All nuclei, P = 0.128; and Neuronal nuclei, *P = 0.016. For gel source data, see Supplementary Information. c, Heatmap of the phenotypic scores of the tamoxifen-injected STOP creER² (top; n = 9) and STOP (bottom; n = 9 until 8 weeks of age, see survival plot in Fig. 4c) animals divided into the six categories. The plot is shaded according to the mean score for each category.
Virus-encoded humanized ΔNIC is expressed in brain and does not have adverse consequences in WT mice. a, b, Representative confocal images from thalamus and brainstem of MeCP2-null + hΔNIC (a) and WT + hΔNIC (b) mice; stained with an antibody against the Myc epitope (red) and the neuronal marker NeuN (green). Nuclei are stained with DAPI (blue). Scale bars, 20 μm. Graphs show transduction efficiency (mean ± s.e.m.) in different brain regions (n = 3 mice per genotype, 27 fields from each brain region). c. Phenotypic scoring (mean ± s.e.m.) of scAAV-injected and control mice from 5 to 30 weeks: WT + vehicle (n = 15), MeCP2-null + vehicle (n = 20) and WT + hΔNIC (n = 14). d. Kaplan–Meier plot showing survival of the cohort shown in c. One WT + hΔNIC animal was culled owing to injuries at 28 weeks of age (shown by a tick). An arrow indicates the timing of the viral injection.