Supplementary Figures: ABE8e adenine base editor precisely and efficiently corrects a recurrent *COL7A1* nonsense mutation

Adam Sheriff[1], Ina Guri[1], Paulina Żebrowska[2], Virginia Llopis-Hernandez[1], Imogen R. Brooks[1], Stavroula Tekkela[1], Kavita Subramaniam[1], Ruta Gebrezgabher[1], Gaetano Naso[3], Anastasia Petrova[3], Katarzyna Balon[2], Alexandros Onoufriadis[1], Dorota Kujawa[2], Martyna Kotulska[2], Greg Newby[4,5,6], Łukasz Łaczmański[2], David R. Liu[4,5,6], John A. McGrath[1], Joanna Jacków[1*]

[1] St John's Institute of Dermatology, King's College London, London, UK

[2] Hirszfeld Institute of Immunology and Experimental Therapy, Polish Academy of Sciences, Poland, PL

[3] Molecular and Cellular Immunology Unit, UCL GOS Institute of Child Health, London, UK

[4] Merkin Institute of Transformative Technologies in Healthcare, Broad Institute of Harvard and

MIT, Cambridge, MA, USA

[5]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA, USA

[6]Howard Hughes Medical Institute, Harvard University, Cambridge, MA, USA

*Correspondence should be addressed to:

Dr Joanna Jacków, Ph.D

Lecturer in Molecular Gene Expression and Gene Editing

St John's Institute of Dermatology

Faculty of Life Sciences & Medicine

King's College London

9th Floor Tower Wing

Guy's Hospital

Great Maze Pond Road, London SE1 9RT, UK

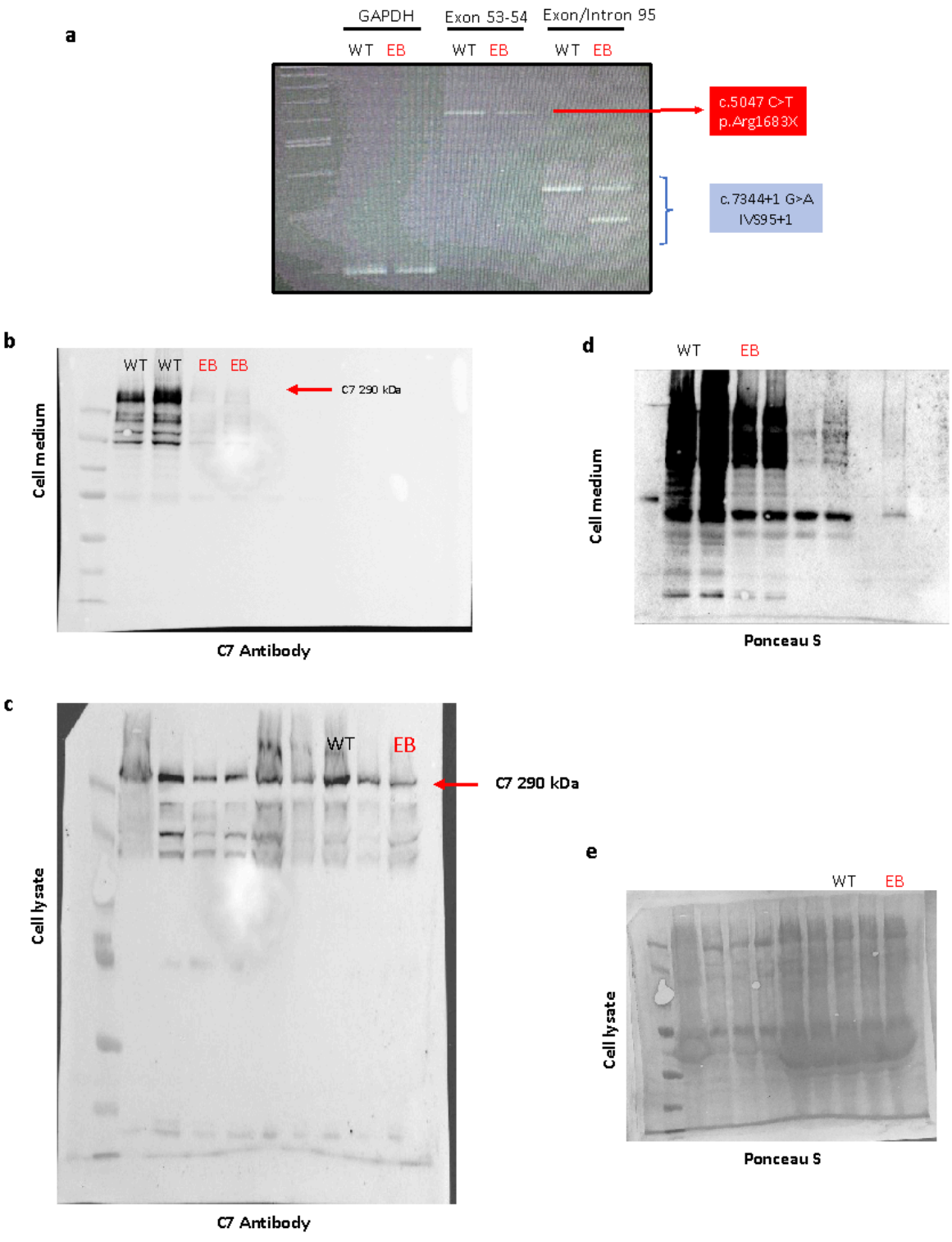
Telephone: +44(0) 20 7188 9613

Email: joanna.jackow@kcl.ac.uk

Short title: Base editing for recessive dystrophic epidermolysis bullosa

Key words: collagen VII, recessive dystrophic epidermolysis bullosa, base editing, off-target

Supplementary Figure S1



Supplementary Figure S1 (a) RT-PCR Transcript analysis of COL7A1 Exons 53-54 and Exon/Intron 95. EB patient cells exhibit reduced transcripts of Exon 53-54 compared to WT due to nonsense mediated decay of the mutated transcript from Allele 1 which harboured the c.5047 C>T mutation. EB patient cells also have an extra smaller sized band of the Exon/Intron 95 transcript compared to WT which was produced by Allele 2 due to to the exon-skipping mutation c.7344+1 G>A. **(b)** Full Western blot of secreted type VII collagen (C7) levels in pre-treated EB fibroblasts compared to Wild-type (WT) fibroblasts. Full-length C7 detected in cell medium using a C7 antibody**. (c)** Full Western blot of intracellular type VII collagen (C7) levels in pre-treated EB fibroblasts compared to Wild-type (WT) fibroblasts. Full-length type VII collagen detected in cell lysates using a C7 antibody. **(d)** Full Western blot of protein in the cell medium stained by Ponceau S. **(e)** Full Western Blot of protein in the cell lysate stained by Ponceau S.

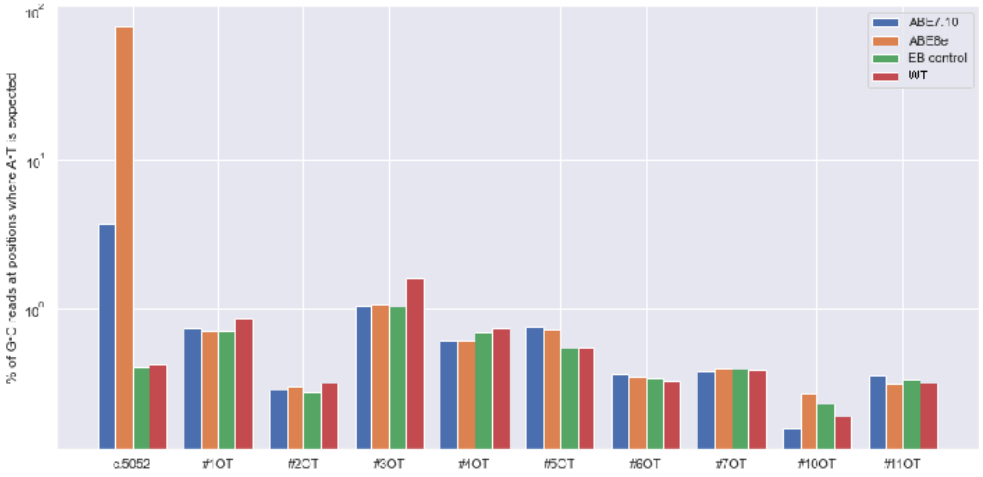
Supplementary Figure S2

gRNA sequence: ACATTTCATCCATCCTCTCCAGG			
Target name	Protospacer sequence	Region	Protein encoded; and its key biological processes
on-target, c.5047 C>T	ACATTTCATCCATCCTCTCCAGG	exon:COL7A1	Collagen alpha-1(VII) chain; Cell adhesion
OT1	GCATCTCAGCCACCCTCTCCAGG	intron:ARL10	ADP-ribosylation factor-like protein 10; GTPase activity, GTP-binding
OT2	ACATCTTGTCCACCCTCTCCAGG	exon:RAB7B	Ras-related protein Rab-7b GTPase activity; Protein transport, Transport
OT3	ACAGTTCACTCACCCTCTCCTGG	exon:ANKFN1	Ankyrin repeat and fibronectin type-III domain-containing protein 1; establishment of mitotic spindle orientation, regulation of establishment of bipolar cell polarity
OT4	ACATTTGAATCATCCTCTCCTGG	intron:ADAMTS12	A disintegrin and metalloproteinase with thrombospondin motifs 12; metal ion binding, metalloendopeptidase activity
OT5	ATACTTCACCCAACCTCTCCTGG	intergenic:SPOCK1-AC106775.1	
OT6	AATTTTCATCCACCCACTCCTGG	intron:GSAP	Gamma-secretase-activating protein; positive regulation of amyloid-beta formation, regulation of proteolysis
OT7	CCATCCCATCCATCCTCTTCTGG	intergenic:AC009410.1-AC074019.1	
OT8	TCATTGCAACCATCCCCTCCTGG	intron:KCNAB1	Voltage-gated potassium channel subunit beta-1; Ion transport, Potassium transport, Transport
OT9	CTATTCCATCCATCCTCTTCTGG	intergenic:AC009977.1-RP11-424G14.1	
OT10	ATATTGCAACCATCCTCACCTGG	intron:DNER	Delta and Notch-like epidermal growth factor-related receptor; Notch signaling pathway
OT11	TCAGGTCATCCACCCTCTCCAGG	intergenic:RP5-884C9.2-LINC01343	

Figure S2 11 putative off-target (OT) sites selected by CRISPOR tool sorted by declining probability of editing. The underlined nucleobases are those in the editing window that were considered in the off-target calculations. The on-target Adenine nucleobase on the positive strand at position c.5047 (chr3: 48580586) is highlighted in green. Highlighted in yellow is the bystander Adenine nucleobase at position c.5052 (chr3: 48580581). For each OT, where they are in the protein coding region (exon or intron), the protein name and its main functions are described. Descriptions for the intergenic regions are omitted.

Supplementary Figure S3

a



b

	ABE7.10	ABE8e	EB control	NHF
c.5052 bystander	3.75	78.49	0.41	0.43
#1OT	0.74	0.71	0.71	0.87
#2OT	0.29	0.30	0.28	0.32
#3OT	1.05	1.07	1.04	1.61
#4OT	0.62	0.62	0.70	0.75
#5OT	0.77	0.73	0.55	0.56
#6OT	0.37	0.35	0.34	0.33
#7OT	0.38	0.40	0.40	0.39
#10OT	0.16	0.27	0.23	0.19
#11OT	0.36	0.32	0.34	0.32

Figure S3. (a) Logarithmic scale comparison of bystander and off-target editing activity in EB-ABE8e, EB-ABE7.10 and untreated EB and WT cells. The percentage of C·G reads where A·T is expected was normalised by the baseline (54.03% of C at c.5047 in EB cells)**.** The bystander mutation at position c.5052 and the 9 of the 10 predicted most likely sites for off-target editing across the genome were interrogated using NGS, illustrating negligible off-target editing activity following ABE7.10 and ABE8e treatment at all 9 sites and a higher efficiency of mutation introduced by ABE8e compared to ABE7.10 at the bystander mutation. **(b)** Height of the bars (%) at the c.5052 position and the 9 OTs, collated in a table.

Supplementary Figure S4

All genes which were significantly differentially expressed ($p<0.05$)

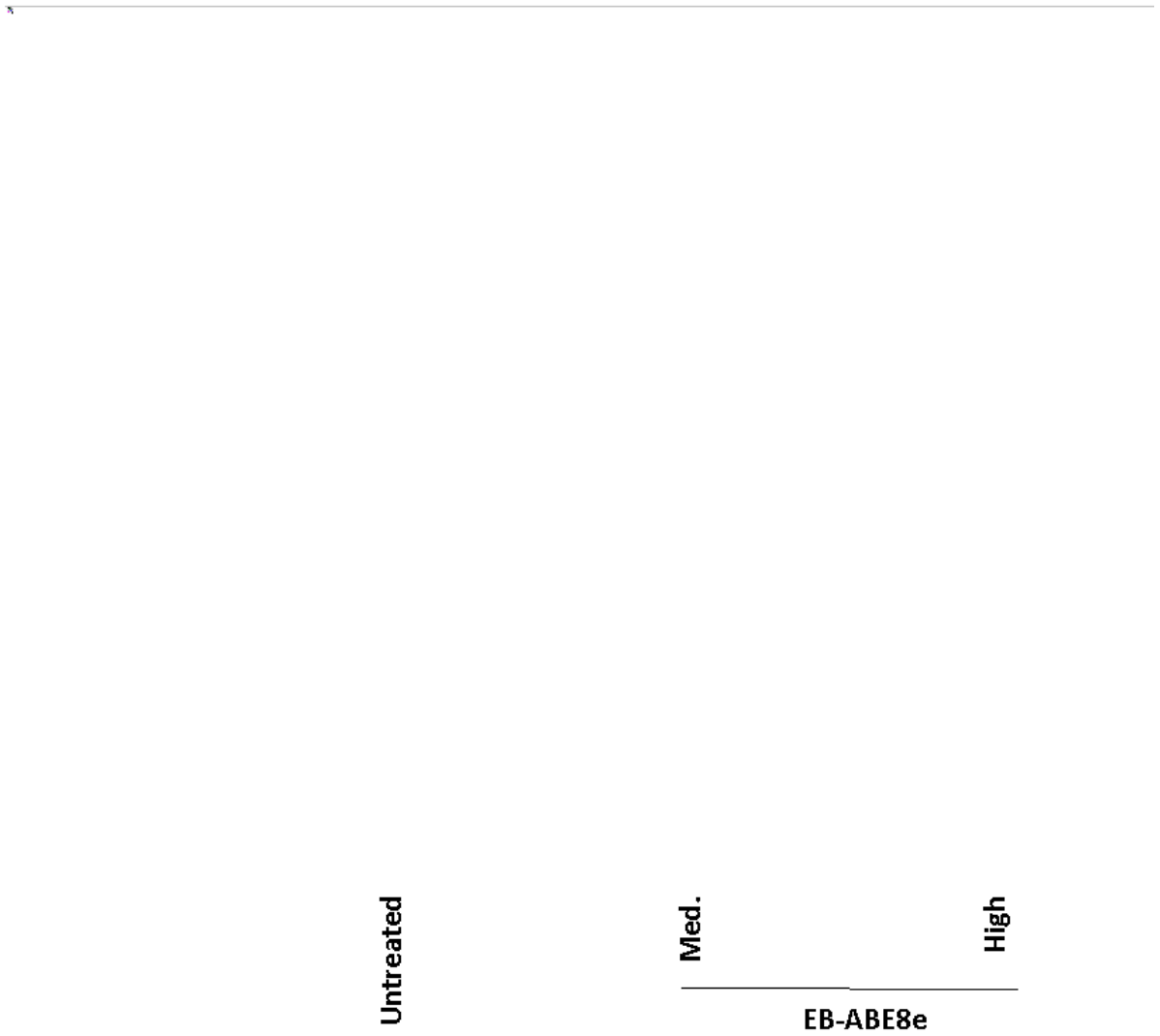
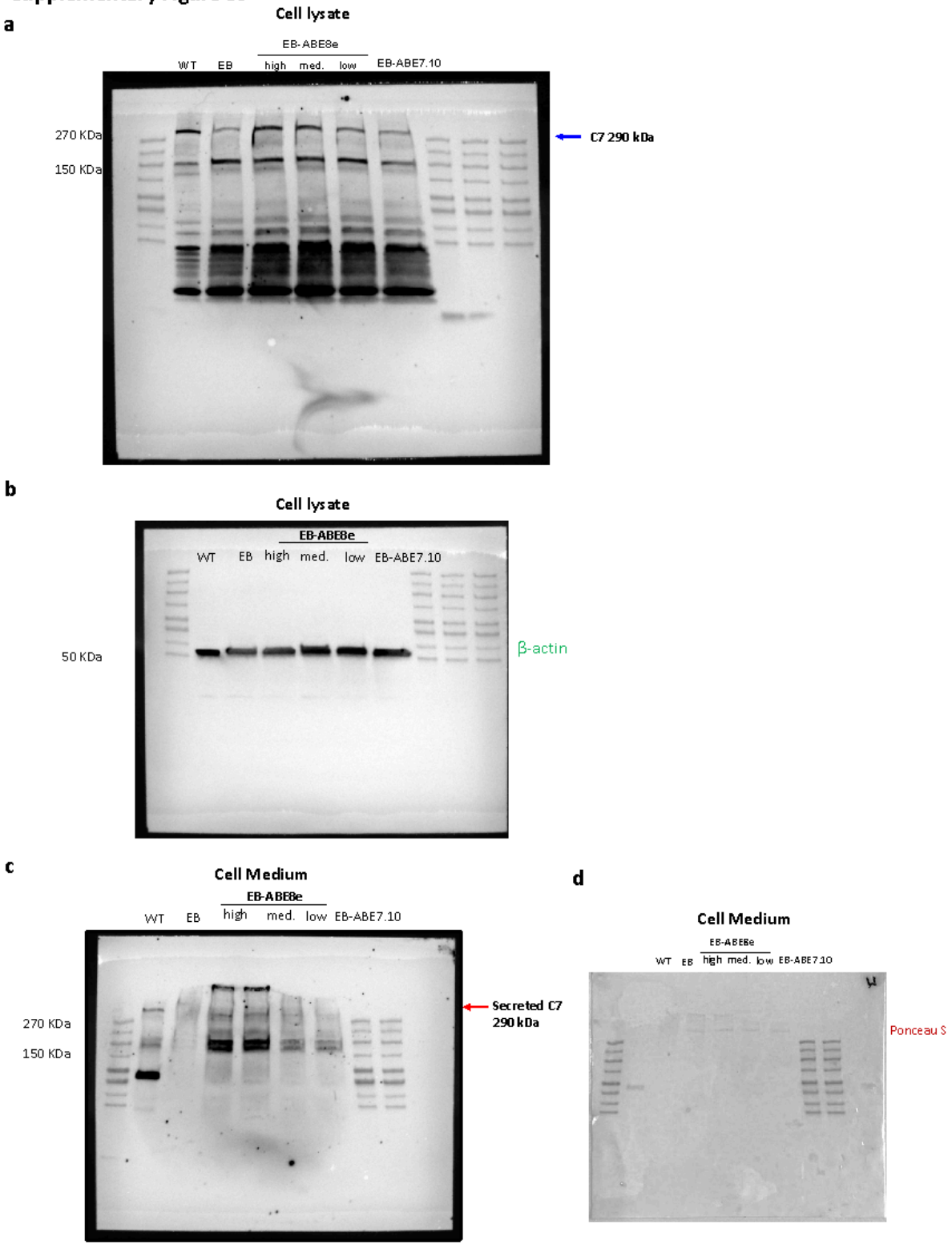


Figure S4. Full heatmap of 443 differentially expressed genes and regions detected ($p<0.05$). Following RNA-Seq, normalized fragments per kilobase of transcript per million mapped reads (FPKM) were used to shortlist genes which were significantly differentially expressed, using a p-value of < 0.05 which yielded 443 differentially expressed genes. This was compared between untreated EB cells, EB-ABE8e high dose and EB-ABE8e medium dose. A heatmap generated using the clustermap function (https://seaborn.pydata.org/generated/seaborn.clustermap.html) from Python's seaborn library from the log2 FPKM of all 443 differentially expressed genes is shown.

Supplementary Figure S5



Supplementary Figure S5. (a) Full Western Blot of WT, untreated and treated fibroblasts. Polyclonal antibody used to stain for C7 in the cell lysate. **(b)** Full Western Blot of WT, untreated and treated fibroblasts. Monoclonal antibody used to stain for C7 in the cell lysate to normalize differential values of C7 intensity. **(c)** Full Western Blot of WT, untreated and treated fibroblasts. Polyclonal antibody used to stain for C7 in the cell medium. **(d)** Full Western Blot of WT, untreated and treated fibroblasts, Ponceau S used to stain for total protein in the cell medium as loading control.