LETTER

https://doi.org/10.1038/s41586-018-0622-0

A metabolite-derived protein modification integrates glycolysis with KEAP1–NRF2 signalling

Michael J. Bollong[1,6], Gihoon Lee[2,3,6], John S. Coukos[2,3], Hwayoung Yun[1,5], Claudio Zambaldo[1], Jae Won Chang[2,3], Emily N. Chin[1], Insha Ahmad[1], Arnab K. Chatterjee[4], Luke L. Lairson[1,4]*, Peter G. Schultz[1,4]* & Raymond E. Moellering[2,3]*

Mechanisms that integrate the metabolic state of a cell with regulatory pathways are necessary to maintain cellular homeostasis. Endogenous, intrinsically reactive metabolites can form functional, covalent modifications on proteins without the aid of enzymes[1,2], and regulate cellular functions such as metabolism[3–5] and transcription[6]. An important 'sensor' protein that captures specific metabolic information and transforms it into an appropriate response is KEAP1, which contains reactive cysteine residues that collectively act as an electrophile sensor tuned to respond to reactive species resulting from endogenous and xenobiotic molecules. Covalent modification of KEAP1 results in reduced ubiquitination and the accumulation of NRF2[7,8], which then initiates the transcription of cytoprotective genes at antioxidant-response element loci. Here we identify a small-molecule inhibitor of the glycolytic enzyme PGK1, and reveal a direct link between glycolysis and NRF2 signalling. Inhibition of PGK1 results in accumulation of the reactive metabolite methylglyoxal, which selectively modifies KEAP1 to form a methylimidazole crosslink between proximal cysteine and arginine residues (MICA). This posttranslational modification results in the dimerization of KEAP1, the accumulation of NRF2 and activation of the NRF2 transcriptional program. These results demonstrate the existence of direct inter-pathway communication between glycolysis and the KEAP1–NRF2 transcriptional axis, provide insight into the metabolic regulation of the cellular stress response, and suggest a therapeutic strategy for controlling the cytoprotective antioxidant response in several human diseases.

In line with its role in responding to altered cellular conditions, numerous studies have linked deregulated KEAP1–NRF2 signalling in disease, including cancer[9], neurodegenerative disorders[10], chronic inflammatory diseases[11], diabetes[12] and ageing[13]. Efforts to target NRF2 signalling therapeutically have largely focused on covalent small-molecule agonists of KEAP1, including the clinical candidate bardoxolone methyl (BARD)[14]. To discover noncovalent modulators of the KEAP1–NRF2 signalling axis, as well as potentially new mechanisms of action for its regulation, we performed a cell-based, high-throughput phenotypic screen using a NRF2-dependent luciferase reporter (pTI-ARE-LUC) in IMR32 cells[15]. From a library of diverse heterocyclic compounds, we identified a hit compound, CBR-470-0, that did not contain any obvious electrophilic groups, and induced the

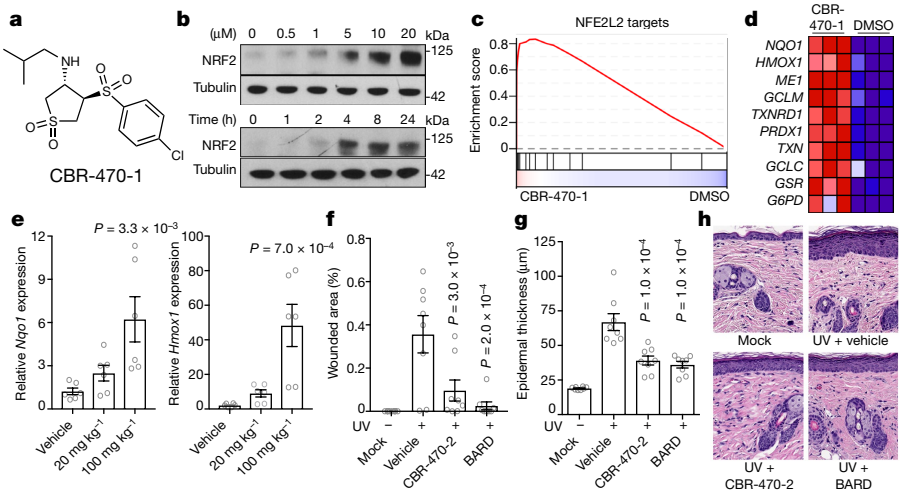


Fig. 1 | CBR-470-series compounds activate NRF2 signalling in vitro and in vivo. a, Structure of CBR-470-1. **b**, NRF2 protein levels from IMR32 cells treated with the indicated concentrations of CBR-470-1 for 4 h (top) or 5 μM CBR-470-1 for the indicated time periods (bottom). Blots are representative of three independent experiments. **c**, Gene set enrichment analysis (GSEA) plot depicting the enrichment of a NRF2 target gene set ('Singh_NFE2L2_Targets' in MSigDB) from IMR32 cells treated for 24 h with 5 μM CBR-470-1 ($n = 3$, $P < 0.0001$, nominal P value in GSEA). **d**, Heat map representation of the leading-edge subset of the most upregulated NRF2-regulated transcripts after CBR-470-1 treatment. Data are biologically independent samples. **e**, Relative *Nqo1* and *Hmox1* transcript levels in mouse skin tissue 24 h after the indicated oral doses of CBR-470-2 ($n = 6$, biologically independent samples). **f**, Quantification of wounded area by automated image analysis from animals of the indicated treatment groups at study end (day 10). **g**, Quantification of epidermal thickness from haematoxylin and eosin (H&E)-stained sections from the indicated groups at study end. **h**, Representative images of H&E-stained skin sections from animals euthanized at day 10 of the study. CBR-470-2, 50 mg kg^{-1} twice a day, orally. BARD, 3 mg kg^{-1} twice a day, orally. UV, 200 mJ cm^{-2}. Data are mean and s.e.m., $n = 8$ animals. Statistical analyses are by one-way analysis of variance (ANOVA) with Dunnett's correction (**e**–**g**).

[1]Department of Chemistry, The Scripps Research Institute, La Jolla, CA, USA. [2]Department of Chemistry, University of Chicago, Chicago, IL, USA. [3]Institute for Genomics and Systems Biology, University of Chicago, Chicago, IL, USA. [4]California Institute for Biomedical Research (Calibr), La Jolla, CA, USA. [5]Present address: College of Pharmacy, Pusan National University, Busan, South Korea. [6]These authors contributed equally: Michael J. Bollong, Gihoon Lee. *e-mail: llairson@scripps.edu; schultz@scripps.edu; rmoellering@uchicago.edu

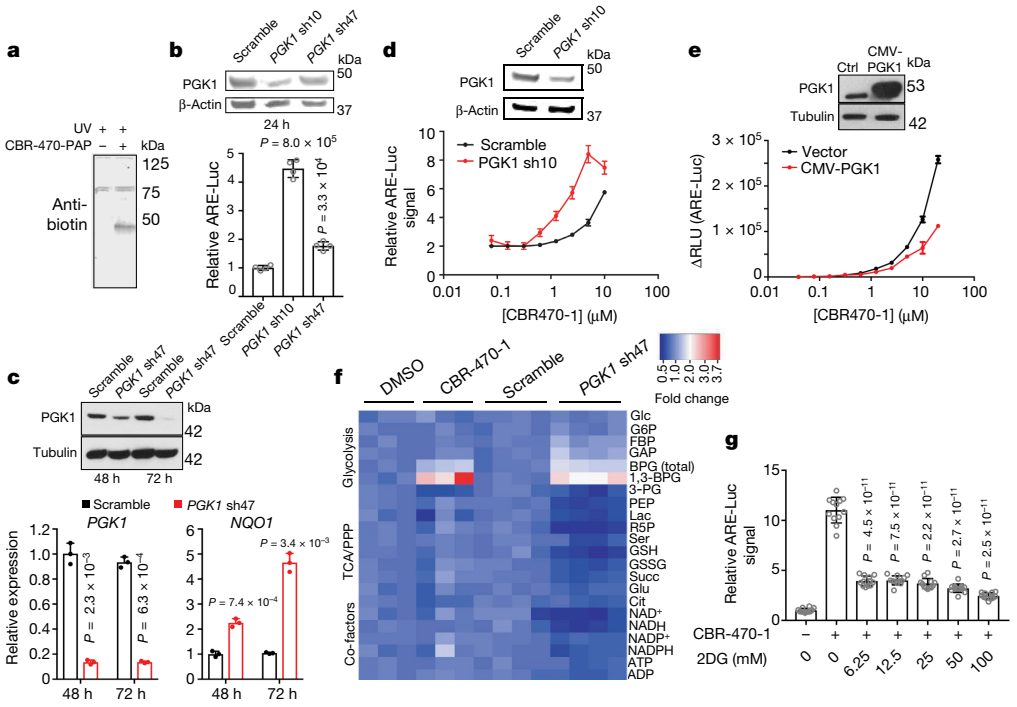


Fig. 2 | CBR-470-1-dependent inhibition of glycolysis activates NRF2 signalling. a, Anti-biotin western blot analysis of IMR32 cells treated with CBR-470-PAP (10 μM) for 1 h and exposed to UV light to induce photocrosslinking (representative shown from $n = 4$ biological replicates). **b**, Transient transfection of shRNA constructs targeting *PGK1* (sh10 and sh47) in HEK293T cells activates the ARE-LUC reporter. PGK1 and β-actin protein levels shown from representative experiments ($n = 4$ biological replicates). **c**, Viral shRNA knockdown of *PGK1* induces *NQO1* mRNA levels in IMR32 cells. PGK1 and tubulin protein levels are shown from representative experiments ($n = 3$). **d**, **e**, CBR-470-1 activation of ARE-LUC reporter in HEK293T cells with transient knockdown (**d**) or overexpression (**e**) of *PGK1* demonstrates opposing effects on compound potency. PGK1, actin and tubulin protein levels are shown from representative experiments ($n = 3$). ΔRLU, change in relative light units. **f**, Heat map depiction of relative metabolite levels in IMR32 cells treated for 30 min with CBR-470-1 (left) or viral shRNA knockdown of *PGK1* (right) relative to DMSO and scramble shRNA controls, respectively. BPG refers to both 2,3-BPG and 1,3-BPG, whereas 1,3-BPG refers specifically to the 1,3-isomer. 3-PG; 3-phosphoglycerate; Cit, citrate; FBP, fructose-1,6-bisphosphate; G6P, glucose 6-phosphate; GAP, glyceraldehyde-3-phosphate; Glc, glucose; Glu, glutamate; GSSG, glutathione disulfide; Lac, lactate; $NADP^+$, nicotinamide adenine dinucleotide phosphate; NAD^+, nicotinamide adenine dinucleotide; NADH, nicotinamide adenine dinucleotide (reduced); NADPH, nicotinamide adenine dinucleotide phosphate (reduced); PEP, 2-phosphoenolpyruvate; PPP, pentose phosphate pathway; R5P, ribose-5-phosphate; Succ, succinate; TCA, tricarboxylic acid. **g**, ARE-LUC reporter activity in IMR32 cells co-treated with CBR-470-1 (5 μM) and 2-deoxyglucose (2DG) for 24 h ($n = 12$). Statistical analyses are by univariate two-sided *t*-tests (**b**, **c**, **g**). Data are mean and s.d. of biologically independent samples.

transcriptional activity of antioxidant-response elements (AREs) to a similar magnitude to the previously reported NRF2 activators tert-butylhydroquinone (TBHQ) and AI-1 (Extended Data Fig. 1a–c). Structure–activity relationship analysis around the cyclic sulfone scaffold led to the identification of CBR-470-1 (Fig. 1a), an isobutylamine-substituted analogue that was not reactive in glutathione assays and had a half-maximum effective concentration (EC_{50}) value of approximately 1 μM in the cellular ARE-LUC assay (Extended Data Fig. 1d, e). CBR-470-1 treatment resulted in a dose- and time-dependent accumulation of NRF2 protein in IMR32 cells (Fig. 1b), and increased both mRNA and protein levels of the NRF2-responsive genes *NQO1* and *HMOX1* (Extended Data Fig. 1f, g; Extended Data Table 1). Expression profiling of IMR32 cells exposed to compound for 24 h revealed that the most significantly enriched gene set was 'NFE2L2 targets', which consisted of NRF2 target genes (Fig. 1c, d); the expression changes of these target transcripts were validated by focused quantitative reverse transcription PCR (qRT–PCR) analysis (Extended Data Fig. 1f). CBR-470-1 also induced transcript levels of *NQO1* and *HMOX1*, and enhanced NRF2 protein stabilization in HEK293T, SH-SY5Y and primary human lung fibroblasts (Extended Data Fig. 1h, i), indicating that these effects are not restricted to a specific cell type. Genetic depletion of NRF2 protein using short hairpin RNA (shRNA) inhibited the ability of CBR-470-1 and TBHQ to induce luciferase expression, indicating that CBR-470-1 activity is dependent on NRF2 (Extended Data Fig. 1j). Finally, CBR-470-1 treatment induced a cytoprotective NRF2 phenotype in vitro, as shown by the protection of SH-SY5Y cells challenged with the cell-permeable peroxide tert-butyl hydroperoxide (TBHP) (Extended Data Fig. 1k).

We next sought to determine whether CBR-470-1 or related analogues induce the activation of NRF2 signalling in vivo. Oral dosing of BALB/c mice with 20 or 100 mg kg^{-1} (twice a day, orally) of a glycine-substituted analogue, CBR-470-2, which is equally potent to CBR-470-1 and has more favourable bioavailability (Extended Data Fig. 2a–e), induced NRF2 target gene expression in several organs, with dose-dependent increases in *Nqo1* and *Hmox1* transcript levels observed in the skin (Fig. 1e). Published studies in *Nrf2* (also known as *Nfe2l2*)-knockout mice have demonstrated that NRF2 is essential to protect against photo-ageing phenotypes and skin carcinogenesis[16] resulting from ultraviolet (UV) irradiation[17]. We therefore evaluated CBR-470-2 activity in this acute UV damage mouse model. Mice were prophylactically dosed with CBR-470-2 (50 mg kg^{-1}, twice a day, orally) or BARD (3 mg kg^{-1}, orally) for five days before exposure to a single dose of UV irradiation. Mice were then dosed for an additional five days, euthanized and UV damage was quantified. CBR-470-2 and BARD treatment resulted in comparable beneficial effects on erythema histological scores and total wounded area (Fig. 1f, Extended Data Fig. 2f, g). Both CBR-470-2 and BARD were also found to decrease epidermal thickness in response to UV exposure, consistent with activation of the NRF2 cytoprotective program[18] (Fig. 1g, h). The combined pharmacodynamic and efficacy data indicate that CBR-470-2 treatment is capable of modulating NRF2 signalling in vivo, despite this compound series operating via an apparent mechanism that is independent of direct KEAP1 binding.

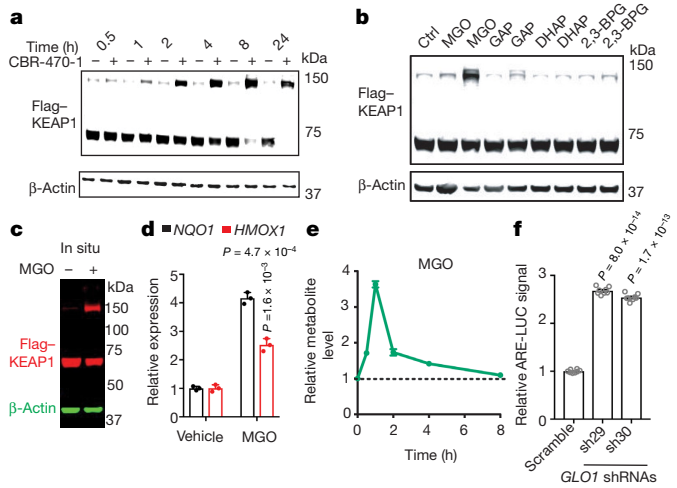


Fig. 3 | Methylglyoxal modifies KEAP1 to form a covalent, high-molecular mass dimer and activate NRF2 signalling. a, Time-course, anti-Flag western blot analysis of whole-cell lysates from HEK293T cells expressing Flag–KEAP1 treated with DMSO or CBR-470-1. **b**, Western blot monitoring of Flag–KEAP1 migration in HEK293T lysates after incubation with central glycolytic metabolites in vitro (1 and 5 mM, left and right for each metabolite). **c**, Flag–KEAP1 (red) and β-actin (green) from HEK293T cells treated with MGO (5 mM) for 8 h. **d**, Relative *NQO1* and *HMOX1* mRNA levels in IMR32 cells treated with MGO (1 mM) or water control ($n = 3$). **e**, LC–MS/MS quantitation of cellular MGO levels in IMR32 cells treated with CBR-470-1 relative to DMSO ($n = 4$). **f**, ARE-LUC reporter activity in HEK293T cells with transient shRNA knockdown of *GLO1* ($n = 8$). Statistical analyses are by univariate two-sided *t*-test (**d**, **f**). Data are mean ± s.e.m. of biologically independent samples.

To determine the mechanism by which CBR-470-1 activates NRF2 signalling, we generated a photo-affinity probe that contained biotin and diazirine substituents, termed CBR-470-PAP, which retained cellular activity in ARE-LUC reporter assays ($EC_{50} = 2.4\,\mu M$; Extended Data Fig. 3a, b). Treatment of IMR32 cells with 5 μM CBR-470-PAP for 1 h, followed by UV irradiation and subsequent anti-biotin western blot analysis of cellular lysates (Fig. 2a), together with biochemical fractionation and liquid chromatography–tandem mass spectrometry (LC–MS/MS) analysis, identified the enzyme phosphoglycerate kinase 1 (PGK1) as a potential target of CBR-470-PAP (Extended Data Fig. 3c). In vitro binding experiments with recombinant protein revealed that CBR-470-PAP selectively labelled PGK1, which was blocked with the soluble competitor CBR-470-1, or shRNA depletion of PGK1 protein levels (Extended Data Fig. 3d–f). Thermal stability assays in the presence of CBR-470-1 resulted in a consistent shift in PGK1 stability, and isothermal dose response profiling[19] against PGK1 and GAPDH also confirmed the selective, dose-dependent alteration of PGK1 stability in the presence of CBR-470-1 (Extended Data Fig. 3g–i). Furthermore, transient and viral shRNA knockdown of *PGK1* in IMR32 cells activated the ARE-LUC reporter signal and expression of *NQO1* (Fig. 2b, c). Knockdown or overexpression of PGK1 protein modulated the NRF2 reporter, with decreased or increased observed CBR-470-1 EC_{50} values, respectively (Fig. 2d, e). Finally, depletion of enolase 1, an enzyme downstream of PGK1, was also found to induce the ARE-LUC signal in IMR32 cells (Extended Data Fig. 3j). These results suggest that CBR-470-1 modulation of PGK1 activity, and therefore glycolysis, regulates NRF2 activation.

Consistent with the PGK1 inhibitory activity of CBR-470-1 (Extended Data Fig. 4a, b), targeted metabolomic profiling[4,20] of IMR32 cells treated with compound revealed a rapid increase in metabolite levels upstream of PGK1 (1,3- and 2,3-bisphosphoglycerate (BPG), and D-glyceraldehyde-3-phosphate (GAP)), and depletion of downstream metabolites such as 3-phosphoglycerate and lactate, which mirrored the profile observed upon viral knockdown of *PGK1* in IMR32 cells (Fig. 2f, Extended Data Fig. 4c, d, Extended Data Table 2). In addition, the co-treatment of ARE-LUC-expressing reporter cells with CBR-470-1 and an inhibitor of glucose entry into glycolysis, 2-deoxyglucose, significantly reduced reporter activation in a dose-dependent manner (Fig. 2g). Together, these data suggested that glycolytic intermediates may serve as a signal to the NRF2 signalling axis.

We first investigated whether 1,3-BPG, which is directly metabolized by PGK1, could be involved in signalling to the KEAP1–NRF2 pathway via phosphoglyceryl-lysine modification of KEAP1. However, CBR-470-1 treatment of IMR32 cells for 30 min, a time at which 1,3-BPG levels are increased, did not result in altered KEAP1 levels, or any anti-phosphoglyceryl-lysine immunoreactive bands using polyclonal antibodies raised against the phosphoglyceryl-lysine epitope (Extended Data Fig. 5a–c). These western blots did, however, reveal the appearance of a CBR-470-1 dose-dependent, high-molecular mass KEAP1 (HMM-KEAP1) band at roughly twice the molecular mass of monomeric KEAP1 (Fig. 3a). The HMM-KEAP1 band was stable to reduction (Extended Data Fig. 5d) and exhibited kinetics and dose-dependent formation consistent with CBR-470-1-dependent NRF2 stabilization and *NQO1* induction (Fig. 1b), but distinct from the direct KEAP1 alkylator TBHQ (Extended Data Fig. 5e–j). Co-treatment of cells with CBR-470-1 and either reduced glutathione (GSH) or *N*-acetylcysteine inhibited HMM-KEAP1 band formation (Extended Data Fig. 5k). Knockdown of *PGK1*, which activates NRF2 target gene expression, also formed HMM-KEAP1, and this band was competed by co-treatment with GSH (Extended Data Fig. 5l). Together, these data indicated that modulation of glycolysis by CBR-470-1 results in the formation of HMM-KEAP1 that is consistent with a covalent KEAP1 dimer, which has been previously observed[21–23], but remained uncharacterized at the molecular level.

Several central glycolytic metabolites other than 1,3-BPG contain reactive functionalities, including the triosephosphate isomers GAP and dihydroxyacetone phosphate, as well as their non-enzymatic elimination product methylglyoxal (MGO), an electrophilic dicarbonyl compound that has been found to form numerous modifications on nucleophilic residues in proteins[24,25]. Among these candidates, only treatment of cell lysates or live cells with MGO resulted in the selective formation of HMM-KEAP1 (Fig. 3b, c). Treatment of Flag–KEAP1-containing lysates or purified KEAP1 with freshly distilled MGO induced dose-dependent formation of HMM-KEAP1 at mid-micromolar concentrations (Extended Data Fig. 5m, n), which is consistent with the range of MGO concentrations previously reported in living cells[26,27]. MGO treatment in cells functionally activated expression of the downstream NRF2 target genes *NQO1* and *HMOX1* (Fig. 3d). Targeted LC–MS measurement of derivatized MGO confirmed that CBR-470-1 treatment resulted in a significant increase in cellular MGO levels in the first few hours of treatment (Fig. 3e, Extended Data Fig. 6a–c), which, along with pathway activation, was sensitive to GSH treatment (Extended Data Fig. 6d–f). To test the involvement of MGO in KEAP1–NRF2 signalling further, we perturbed its degradation, which is mediated by GSH and glyoxylase 1 (GLO1). Knockdown of *GLO1* by shRNA resulted in ARE-LUC reporter activation (Fig. 3f), and also sensitized cells to CBR-470-1 activation of the ARE-LUC reporter (Extended Data Fig. 6g). Direct modulation of GLO1 enzymatic activity with a cell-permeable inhibitor (GLOi) also amplified reporter activation by CBR-470-1 (Extended Data Fig. 6h). Collectively, these metabolomic, proteomic and transcriptomic data established shared kinetics between MGO accumulation, HMM-KEAP1 formation and NRF2 pathway activation, suggesting the existence of a direct link between glycolysis and the KEAP1–NRF2 signalling pathway that is mediated by the modification of KEAP1 by MGO.

A stable isotope labelling with amino acids in cell culture (SILAC)-based quantitative proteomic approach (Extended Data Fig. 7a) suggested the N-terminal region (amino acids 1–50) and BTB domains (amino acids 150–169) as candidate domains and residues that could be involved in HMM-KEAP1 formation in response to the CBR-470-1-induced increase in MGO (Extended Data Fig. 7b–d). We therefore examined more than a dozen C-to-S, K-to-M/R and R-to-A mutations

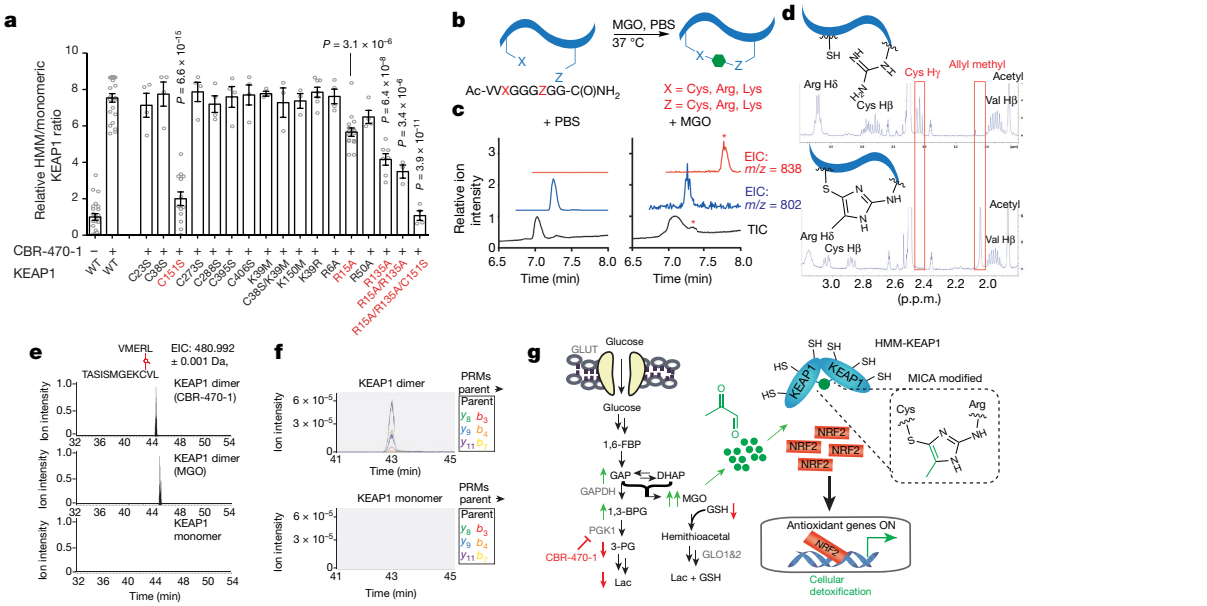


Fig. 4 | Methylglyoxal forms a posttranslational modification between proximal cysteine and arginine residues in KEAP1. a, Quantified HMM-KEAP1 formation of wild-type (WT) or mutant Flag–KEAP1 from HEK293T cells treated with DMSO or CBR-470-1 for 8 h ($n = 23$ for WT; $n = 16$ for R15A; $n = 13$ for C151S; $n = 7$ for K39R, R135A; $n = 4$ for R6A, R50A, all other C-to-S mutations, and R15/135A and C151S triple-mutant; $n = 3$ for R15/135A, and all K-to-M mutations). **b**, Schematic of the model peptide screen for intramolecular modifications formed by MGO and nucleophilic residues. **c**, Total ion and extracted ion chromatograms (TICs and EICs, respectively) from MGO- and mock-treated peptide, with a new peak in the former condition marked by an asterisk. EICs are specific to the indicated m/z ($n = 3$ independent biological replicates). **d**, ^{1}H-NMR spectra of the unmodified (top) and MICA-modified (bottom) model peptide, with pertinent protons highlighted in each. Notable changes in the MICA-modified spectrum include the appearance of a singlet at 2.04 p.p.m., loss of the thiol proton at 2.43 p.p.m., and changes in chemical shift and splitting pattern of the cysteine beta protons and the arginine delta and epsilon protons. Full spectra and additional multidimensional NMR spectra can be found in Extended Data Fig. 7. **e**, EIC from LC–MS/MS analyses of gel-isolated and digested HMM-KEAP1 (CBR-470-1 and MGO-induced) and monomeric KEAP1 for the C151–R135 crosslinked peptide. Slight retention time variation was observed on commercial columns ($n = 3$ independent biological replicates). **f**, PRM chromatograms for the parent and six parent-to-daughter transitions in representative targeted proteomic runs from HMM-KEAP1 and monomeric digests ($n = 6$). **g**, Schematic depicting the direct communication between glucose metabolism and KEAP1–NRF2 signalling mediated by MGO modification of KEAP1 and subsequent activation of the NRF2 transcriptional program. Statistical analyses are by univariate two-sided t-test (**a**). Data are mean ± s.e.m. of biologically independent samples.

within these domains, as well as other known functional residues in KEAP1, for their effect on HMM-KEAP1 formation. Two arginine residues (R15 of the N-terminal region domain and R135 of the BTB domain) significantly, but incompletely, reduced the formation of HMM-KEAP1 (Fig. 4a). More notable was the near complete inhibition of HMM-KEAP1 formation of the C151S mutant in the BTB domain (Fig. 4a). Consistent with this effect, levels of C151-containing tryptic peptides were reduced by MGO treatment, and the pre-treatment of cells with BARD, which alkylates C151, inhibited HMM-KEAP1 formation (Extended Data Fig. 7d, e). C151 lies in an exposed region of the BTB domain that is predicted to mediate the homodimeric interface between two KEAP1 monomers, which is necessary for proper NRF2 binding and ubiquitination[8,23]. Therefore, the strong abrogation of HMM-KEAP1 formation by the mutation of C151 and proximal arginine residues suggested that MGO may be mediating an uncharacterized modification between these residues.

To identify this modification, we synthesized a model peptide containing cysteine and arginine separated by a glycine linker, which was intended to mimic high inter- or intramolecular cysteine and arginine proximity, and treated it with MGO at physiological temperature and pH overnight (Fig. 4b). LC–MS analysis revealed a new peak, which corresponded to a mass increase of 36 Da, consistent with a mercaptomethylimidazole crosslink (Fig. 4c, d) formed by nucleophilic attack of the dicarbonyl by the side chains of cysteine and arginine, followed by dehydration-mediated formation of a methylimidazole crosslink between cysteine and arginine (MICA) posttranslational modification. We purified this product and confirmed its structure by a series of one- and two-dimensional NMR experiments (Fig. 4d, Extended Data Fig. 7f–i). To determine whether MICA modification occurs within KEAP1 protein, we treated cells with CBR-470-1 or MGO, isolated HMM-KEAP1 and monomeric KEAP1 by gel electrophoresis, and then digested these discrete populations for LC–MS/MS analyses. A peptide bearing a MICA crosslink between C151 and R135 was identified in isolated HMM-KEAP1 after both CBR-470-1 and MGO treatment, but not in the isolated monomeric KEAP1 (Fig. 4e). Furthermore, parallel-reaction monitoring (PRM) confirmed the presence and co-elution of more than a dozen parent-to-daughter ion transitions that were uniquely present in HMM-KEAP1 (Fig. 4f, Extended Data Fig. 8a, b). These studies suggest a model in which glycolytic metabolic status is coupled to NRF2-dependent gene expression through the direct interaction of a reactive glycolytic metabolite, MGO, and the sentinel protein KEAP1 via the formation of a stable and mechanistically novel protein posttranslational modification (Fig. 4g).

Although MGO has been previously shown to form covalent modifications on diverse proteins, the compositions, sites and functions of these modifications have remained largely unknown. Similarly, several recent reports have implicated MGO in the pathogenesis of diseases such as diabetes[28] and ageing[29], but the discrete molecular targets of MGO in these contexts are unidentified. Here we found that inhibition of PGK1 increases levels of triosephosphates, which results in increased levels of cellular MGO and the formation of a HMM-KEAP1 species that leads to NRF2-dependent gene expression. Formation of the HMM-KEAP1 species involves the posttranslational modification MICA that is dependent on MGO and forms a covalent linkage between proximal cysteine and arginine residues. These results raise intriguing questions about the general reactivity of MGO, its potential role as a signalling metabolite in other cellular processes, and the specific modifications involved in often-cited advanced glycated end products as biomarkers of disease pathology. We have shown that both cellular and lysate treatment with MGO results in the selective modification of

C151 in KEAP1, probably owing to the intrinsic hyperreactivity of this residue, and the presence of properly oriented arginine(s) that enables formation of the MICA modification. Additional factors such as local metabolite concentration gradients may also contribute to MICA formation in KEAP1. Future studies to determine the full target profile of MGO, and specifically other inter- and intramolecular MICA modifications, are expected to shed light on this model and provide a global view of MGO modification sites in the proteome.

The direct connection between glucose metabolism and the KEAP1–NRF2 axis by MGO adds a further layer of regulation to both pathways and global metabolic status. First, this connection highlights the role of glycolysis in regulating cellular redox status beyond the contribution to NADPH and glutathione production. These reducing equivalents are crucial for the regulation of a wide range of reactive species in the cell, and when these levels are deregulated, the KEAP1–NRF2 pathway is poised to respond and limit cellular damage. Recent studies have also implicated the output of the NRF2 transcriptional program in the direct detoxification of MGO through increased glutathione synthesis[30], *GLO1* transcription, as well as the redirection of glucose carbon away from central metabolites (for example, MGO) and into the pentose phosphate pathway[31]. Therefore, the direct coupling of glucose metabolism with KEAP1 function through MGO creates an intrinsic feedback loop to sense and respond to changing metabolic demands in the cell. A final aspect of this study is the notion that modulation of endogenous reactive metabolite levels using small molecules may represent an alternative approach towards activating the ARE pathway for the treatment of several diseases that involve metabolic stress.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41586-018-0622-0.

Received: 22 October 2017; Accepted: 21 August 2018;
Published online 15 October 2018.

1. Lin, H., Su, X. & He, B. Protein lysine acylation and cysteine succination by intermediates of energy metabolism. *ACS Chem. Biol*. **7**, 947–960 (2012).
2. Wagner, G.R., et al. A class of reactive acyl-CoA species reveals the non-enzymatic origins of protein acylation. *Cell Metab*. **25**, 823–837 (2017).
3. Zhang, Z. et al. Identification of lysine succinylation as a new post-translational modification. *Nat. Chem. Biol*. **7**, 58–63 (2011).
4. Moellering, R. E. & Cravatt, B. F. Functional lysine modification by an intrinsically reactive primary glycolytic metabolite. *Science* **341**, 549–553 (2013).
5. Weinert, B. T. et al. Acetyl-phosphate is a critical determinant of lysine acetylation in *E. coli*. *Mol. Cell* **51**, 265–272 (2013).
6. Sabari, B. R., Zhang, D., Allis, C. D. & Zhao, Y. Metabolic regulation of gene expression through histone acylations. *Nat. Rev. Mol. Cell Biol*. **18**, 90–101 (2016).
7. Kobayashi, A. et al. Oxidative and electrophilic stresses activate Nrf2 through inhibition of ubiquitination activity of Keap1. *Mol. Cell. Biol*. **26**, 221–229 (2006).
8. Lo, S. C., Li, X., Henzl, M. T., Beamer, L. J. & Hannink, M. Structure of the Keap1:Nrf2 interface provides mechanistic insight into Nrf2 signaling. *EMBO J*. **25**, 3605–3617 (2006).
9. Jaramillo, M. C. & Zhang, D. D. The emerging role of the Nrf2-Keap1 signaling pathway in cancer. *Genes Dev*. **27**, 2179–2191 (2013).
10. Scannevin, R. H. et al. Fumarates promote cytoprotection of central nervous system cells against oxidative stress via the nuclear factor (erythroid-derived 2)-like 2 pathway. *J. Pharmacol. Exp. Ther*. **341**, 274–284 (2012).
11. Khor, T. O. et al. Nrf2-deficient mice have an increased susceptibility to dextran sulfate sodium-induced colitis. *Cancer Res*. **66**, 11580–11584 (2006).
12. Uruno, A. et al. The Keap1-Nrf2 system prevents onset of diabetes mellitus. *Mol. Cell. Biol*. **33**, 2996–3010 (2013).
13. Sykiotis, G. P. & Bohmann, D. Keap1/Nrf2 signaling regulates oxidative stress tolerance and lifespan in *Drosophila*. *Dev. Cell* **14**, 76–85 (2008).
14. Cleasby, A. et al. Structure of the BTB domain of Keap1 and its interaction with the triterpenoid antagonist CDDO. *PLoS One* **9**, e98896 (2014).
15. Hur, W. et al. A small-molecule inducer of the antioxidant response element. *Chem. Biol*. **17**, 537–547 (2010).
16. Saw, C. L. et al. Impact of Nrf2 on UVB-induced skin inflammation/photoprotection and photoprotective effect of sulforaphane. *Mol. Carcinog*. **50**, 479–486 (2011).
17. Tao, S., Justiniano, R., Zhang, D. D. & Wondrak, G. T. The Nrf2-inducers tanshinone I and dihydrotanshinone protect human skin cells and reconstructed human skin against solar simulated UV. *Redox Biol*. **1**, 532–541 (2013).
18. El-Abaseri, T. B., Putta, S. & Hansen, L. A. Ultraviolet irradiation induces keratinocyte proliferation and epidermal hyperplasia through the activation of the epidermal growth factor receptor. *Carcinogenesis* **27**, 225–231 (2006).
19. Martinez Molina, D. et al. Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. *Science* **341**, 84–87 (2013).
20. Chang, J. W., Lee, G., Coukos, J. S. & Moellering, R. E. Profiling reactive metabolites via chemical trapping and targeted mass spectrometry. *Anal. Chem*. **88**, 6658–6661 (2016).
21. Zhang, D. D. & Hannink, M. Distinct cysteine residues in Keap1 are required for Keap1-dependent ubiquitination of Nrf2 and for stabilization of Nrf2 by chemopreventive agents and oxidative stress. *Mol. Cell. Biol*. **23**, 8137–8151 (2003).
22. Wakabayashi, N. et al. Protection against electrophile and oxidant stress by induction of the phase 2 response: fate of cysteines of the Keap1 sensor modified by inducers. *Proc. Natl Acad. Sci. USA* **101**, 2040–2045 (2004).
23. Ogura, T. et al. Keap1 is a forked-stem dimer structure with two large spheres enclosing the intervening, double glycine repeat, and C-terminal domains. *Proc. Natl Acad. Sci. USA* **107**, 2842–2847 (2010).
24. Lo, T. W., Westwood, M. E., McLellan, A. C., Selwood, T. & Thornalley, P. J. Binding and modification of proteins by methylglyoxal under physiological conditions. A kinetic and mechanistic study with Nα-acetylarginine, Nα-acetylcysteine, and Nα-acetyllysine, and bovine serum albumin. *J. Biol. Chem*. **269**, 32299–32305 (1994).
25. Rabbani, N. & Thornalley, P. J. Dicarbonyl proteome and genome damage in metabolic and vascular disease. *Biochem. Soc. Trans*. **42**, 425–432 (2014).
26. Chaplen, F. W., Fahl, W. E. & Cameron, D. C. Evidence of high levels of methylglyoxal in cultured Chinese hamster ovary cells. *Proc. Natl Acad. Sci. USA* **95**, 5533–5538 (1998).
27. Dhar, A., Desai, K., Liu, J. & Wu, L. Methylglyoxal, protein binding and biological samples: are we getting the true measure? *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci*. **877**, 1093–1100 (2009).
28. Moraru, A. et al. Elevated levels of the reactive metabolite methylglyoxal recapitulate progression of type 2 diabetes. *Cell Metab*. **27**, 926–934 (2018).
29. Ravichandran, M. et al. Impairing l-threonine catabolism promotes healthspan through methylglyoxal-mediated proteohormesis. *Cell Metab*. **27**, 914–925 (2018).
30. Nishimoto, S., Koike, S., Inoue, N., Suzuki, T. & Ogasawara, Y. Activation of Nrf2 attenuates carbonyl stress induced by methylglyoxal in human neuroblastoma cells: Increase in GSH levels is a critical event for the detoxification mechanism. *Biochem. Biophys. Res. Commun*. **483**, 874–879 (2017).
31. Mitsuishi, Y. et al. Nrf2 redirects glucose and glutamine into anabolic pathways in metabolic reprogramming. *Cancer Cell* **22**, 66–79 (2012).

Acknowledgements We thank S. Zhu for discussions about target identification experiments. Animal experiments were approved by the Scripps Research Institute Institutional Review Board. We are grateful for financial support of this work from the following: Kwanjeong Educational Fellowship (to G.L.); NIH MSTP Training Grant (T32GM007281 to J.S.C.); NIH R00CA175399, R01CA211916 and DP2GM128199 (R.E.M.); V Foundation for Cancer Research (V2015-020 to R.E.M.); Damon Runyon Cancer Research Foundation (DFS08-14); The Skaggs Institute for Chemical Biology, and The University of Chicago.

Reviewer information *Nature* thanks H. Christofk, A. Dinkova-Kostova, H. Lin and the anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions All authors reviewed the manuscript. M.J.B., G.L., J.S.C., H.Y., C.Z., J.W.C., I.A., L.L.L. and R.E.M. designed and performed biochemical and cell-based biological experiments. H.Y., J.W.C., J.S.C. and A.K.C. synthesized and characterized chemical probes and reagents. E.N.C., M.J.B., L.L.L. and P.G.S. designed, performed and analysed in vivo experiments. G.L., J.S.C., J.W.C. and R.E.M. designed, performed and analysed metabolomic, proteomic and structural characterization experiments. L.L.L., P.G.S. and R.E.M. conceived of the study and supervised research. M.J.B. and R.E.M. wrote the manuscript with considerable input from all authors.

Competing interests Declared none.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-018-0622-0.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-018-0622-0.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to L.L.L., P.G.S. or R.E.M.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

No statistical methods were used to predetermine sample size. *Chemicals.* TBHQ, 2-deoxyglucose, MGO and GSH were obtained from Sigma Aldrich. The synthesis of AI-1 has been described previously[32]. The GLO1 inhibitor (CAS no. 174568-92-4) was from MedChemExpress. CBR-470-0 and CBR-581-9 were from ChemDiv. CBR-470-1 (initially from ChemDiv as D470-2172) and related analogues were synthesized in house according to full methods described in the Supplementary Information. All commercially obtained chemicals were dissolved in DMSO and used without further purification, with the exception of 2-deoxyglucose, MGO and GSH, which were delivered as aqueous solutions.

Cell culture. IMR32, SH-SY5Y, HeLa and HEK293T cell lines were purchased from ATCC. Human lung fibroblasts and mouse dermal fibroblasts (C57BL/6-derived) were obtained from Sciencell and used before passage 3. IMR32, human lung fibroblasts, SH-SY5Y, HeLa and HEK293T cells were propagated in DMEM (Corning) supplemented with 10% fetal bovine serum (FBS, Corning) and 1% penicillin/streptomycin (Gibco). Mouse dermal fibroblasts were propagated in fibroblast medium 2 from Sciencell. Mouse epidermal keratinocytes (MPEK-BL6) were obtained from Zen Bio and propagated in epidermal keratinocyte medium (Zen Bio).

High-throughput screening and ARE-LUC reporter assay. For high-throughput screening, IMR32 cells were plated at 5×10^3 cells per well in white 384-well plates in 40 μl of growth medium. The next day 100 ng of pTI-ARE-LUC reporter plasmid in 10 μl of Optimem medium (Gibco) was transfected into each well using Fugene HD at a dilution of 1 μg plasmid DNA:4 μl of Fugene. After 24 h, compounds were transferred using a 100 nl pintool head affixed to PerkinElmer FX instrument such that the final screening concentration was 2 μM. After 24 h incubation, ARE-LUC luminance values were recorded on an Envision instrument after the addition of 30 μl of Bright Glo reagent solution (Promega, diluted 1:3 in water). Compounds which increased ARE-LUC signal greater than four *Z*-scores from plate mean were deemed hits. For overexpression and knockdown experiments in HEK293T with ARE-LUC reporter readouts, 5×10^5 cells were plated on poly-D-lysine-coated plates and transfected with 1.5 μg of overexpression or shRNA plasmid and 500 ng of pTI-ARE-LUC using Optimem medium and Fugene in the same mode as above. After 24 h, 10^3 transfected cells were plated in 50 μl of growth medium in white 96-well plates. After a 24 h incubation, a further 50 μl of growth medium with compound at the indicated concentrated was added to each well. ARE-LUC luminance values were recorded on an Envision plate reader 24 h later by the addition of 75 μl of Bright Glo reagent solution (1:3 in water).

Peroxide stress model. Approximately 10^4 SH-SY5Y cells were plated in 100 μl of growth medium in white 96-well plates. After 48 h of compound treatment, 20 μl TBHP diluted to the indicated concentrations was added to each well. After an 8 h incubation, cell viability measurements were recorded on an Envision plate reader after the addition of 50 μl of a Cell Titer Glo solution (Promega, diluted 1:6 in water). Relative viabilities are reported as a fraction relative to the same dose of compound treatment without TBHP.

shRNA knockdown studies. PGK1-targeting shRNA vectors sh10 and sh47 refer to Sigma Mission shRNA lentiviral clones NM_000291.2-338s1c1 and NM_000291.2-935s1c1 respectively. GLO1-targeting shRNA vectors sh29 and sh30 refer to Sigma Mission shRNA lentiviral clones NM_006708.1-195s1c1 and NM_006708.1-292s1c1, respectively. The non-targeting scrambled control vector refers to SHC002 (Sigma). Lentiviruses were generated in HEK293T cells by transient expression of the above vectors with pSPAX2 and pMD2.G packaging vectors (Addgene plasmids 11260 and 12259). Viral supernatants were collected after 48 h of expression and passed through a 70-μm syringe filter before exposure to target cells.

qRT–PCR. Cells were collected by trypsinization and subsequent centrifugation at 500*g*. RNA was isolated using RNeasy kits from Qiagen and concentrations obtained using a NanoDrop instrument. Then 500 ng–5 μg of RNA was then reverse transcribed with oligo dT DNA primers using a SuperScript III First-Strand Synthesis kit from Invitrogen. Quantitative RT–PCR reactions were measured on a Viia 7 Real-Time PCR system (Thermo) using a Clontech SYBR green-based master mix. Gene-specific primer sets are provided in Extended Data Table 1. Reactions were normalized to *TUBG1* levels for each biological replicate and relative transcript abundance calculated using the comparative C_t method.

GSEA. Total RNA was extracted from IMR32 cells treated for 24 h with either DMSO or 5 μM CBR-470-1 (3 biological replicates per condition) using an RNeasy kit (Qiagen). RNA sequencing (RNA-seq) experiments were performed by the Scripps Next Generation Sequencing Core according to established in house methods. GSEAs and leading edge heat maps were generated with TPM values from the above experiment using the java GSEA package. 'NFE2L2 targets' gene set refers to Molecular Signature Database (http://software.broadinstitute.org/gsea/msigdb) gene set ID M2662.

Quantitative metabolomic profiling. For polar metabolite profiling experiments, cells were grown in 15 cm plates and cultured in RPMI supplemented with 10% FBS, 2 mM L-glutamine and 1% penicillin/streptomycin before media replacement containing either vehicle (DMSO) or the indicated dose of CBR-470-1. After incubation for the appropriate time, cells were scraped into ice-cold PBS and isolated by centrifugation at 1,400*g* at 4 °C. Cell pellets were resuspended in 300 μl of an 80:20 mixture of cold methanol/water, an internal standard was added (10 nmol d_3-serine; Sigma Aldrich), and the suspension was sonicated (Fisher Scientific, FB-505) for 5 s followed by a 10 min centrifugation at 16,000*g*. The supernatant was collected, dried under N_2 gas and resulting dried metabolites resuspended in 30 μl of 40% methanol/water for analysis on an Agilent triple quadrupole LC–MS/MS (Agilent Technologies, 6460 QQQ). For negative mode operation, metabolites were separated by hydrophilic interaction chromatography with a Luna-NH_2 column (Phenomenex) running mobile phase A (CH_3CN supplemented with 0.2% NH_4OH) and B (95:5 (v/v) H_2O:CH_3CN supplemented with 50 mM ammonium acetate and 0.2% NH_4OH) and the following gradient: 0% B for 3 min; linear increase to 100% B for 27 min at a flow rate of 0.4 ml min^{-1}, followed by an isocratic flow of 100% B for 3 min. The spectrometer settings were: capillary voltage = 4.0 kV, drying gas temperature = 350 °C at 10 l min^{-1}, and the nebulizer pressure was 45 p.s.i. Metabolite peak transitions and retention times are listed in Extended Data Table 2 and were confirmed by running standards for measured glycolytic, pentose phosphate pathway, tricarboxylic acid cycle and co-factor metabolites. 2-phosphoglycerate and 3-phosphoglycerate isomers were quantified in aggregate. Relative metabolite abundance was quantified by integrated peak area for the given multiple reaction monitoring transition, and all metabolite levels were normalized to internal standard extracted ion intensity values for d_3-serine. These parameters were used to quantify all metabolites, with the exception of 1,3-BPG and MGO, which required chemical derivitization to stable intermediates before LC–MS/MS quantification, as previously reported[20,33]. MGO deviated from all other metabolites, as it was separated on a Gemini reverse-phase C18 column (5 mm, 4.6 mm × 50 mm; Phenomenex) together with a pre-column (C18, 3.5 mm, 2 mm × 20 mm) and detected in positive mode analysis, with mobile phase A (H_2O) and B (50:50 (v/v) H_2O:CH_3CN) supplemented with 0.1% trifluoroacetic acid. The gradient started with 0% B for 2 min and increased linearly to 100% B over 10 min with a flow rate of 0.4 ml min^{-1}, followed by an isocratic gradient of 100% B for 5 min at 0.4 ml min^{-1}. The QQQ settings were the same as above.

Flag-tagged protein expression and western blotting. Full-length, human PGK1 (NM_000291, Origene) transiently expressed from a pCMV6 entry vector with a C-terminal Myc-DDK tag; full-length, human KEAP1 (28023, Addgene) was transiently expressed from a pcDNA/FRT/TO plasmid with a C-terminal 3×Flag tag. All references to Flag–PGK1 or Flag–KEAP1 represent the proteins in the aforementioned vectors, respectively. Transient protein expression was performed in confluent 10 cm plates of HEK293T cells by transfection of 1 μg plasmid with Lipofectamine 2000 (Invitrogen) according to manufacturer's protocol. For in situ compound or metabolite treatment experiments, compounds were added approximately 24 h after transfection, and incubated for the indicated duration. For Flag–KEAP1 western blotting and immunoprecipitation experiments, cells were collected by scraping, pelleted by centrifugation, washed twice with PBS and lysed in 8 M urea, 50 mM NH_4HCO_3, phosphatase inhibitor cocktail (Sigma Aldrich), and EDTA-free complete protease inhibitor (Roche), pH 8.0, at 4 °C. Lysate was sonicated (Fisher Scientific, FB-505), insoluble debris cleared by centrifugation, and the supernatant was diluted into 4× Laemmli buffer containing 50 mM dithiothreitol (DTT) as a reducing agent. Samples were prepared for SDS–PAGE by heating to 95 °C for 5 min, cooled to room temperature, resolved on NuPAGE Novex 4–12% Bis-Tris Protein Gels (Invitrogen), and transferred onto nitrocellulose membranes by standard western blotting methods. Membranes were blocked in 2% BSA in TBS containing 0.1% tween-20 (TBST) and probed with primary and secondary antibodies. Primary antibodies used in this study include: anti-Flag-M2 (1:1,000, F1804, Sigma Aldrich), anti-KEAP1 (1:500, SC-15246, Santa Cruz), anti-HSPA1A (1:1,000; 4872, Cell Signaling), anti-ACTB (1:1,000, 4790, Cell Signaling), anti-GAPDH (1:1,000; 2118S, Cell Signaling) and TUBG (1:1,000; 5886, Cell Signaling). Rabbit polyclonal anti-pgK antibody was generated using pgK-modified KLH and affinity purification as described[4] at a 1:400 dilution of a 0.33 mg ml^{-1} stock in 10 mM sodium HEPES (pH 7.5), 150 mM NaCl, 30% glycerol and 0.02% sodium azide. Secondary donkey anti-rabbit, donkey anti-goat, and donkey anti-mouse (Licor), were used at 1:10,000 dilution in 2% BSA-containing TBST and incubated for 1 h before washing and imaging on a Licor infrared scanner. Densitometry measurements were performed with ImageJ software.

Time- and dose-dependent CBR-470-1 treatment studies were performed in HEK293T cells 24 h after transient transfection of Flag–KEAP1, or in IMR32 cells for endogenous KEAP1. Fresh RPMI media with 10% FBS, 2 mM L-glutamine, 1% penicillin/streptomycin and the indicated concentration of CBR470-1 (20 μM for time-dependent experiments) or equivalent DMSO was added to cells in 10 cm dishes. After the indicated incubation time, cells were lysed in lysis buffer (50 mM Tris, 150 mM NaCl, 1% Triton-X 100, phosphatase inhibitor cocktail (Sigma

Aldrich), and EDTA-free complete protease inhibitor (Roche), pH 7.4) and processed for western blot as indicated above.

Target identification studies with CBR-470-PAP. Confluent IMR32 cells in 10 cm dishes were exposed to 5 μM CBR-470-PAP with either DMSO or a 50-fold molar excess of CBR-470-1 (250 μM) for 1 h at 37 °C. Samples were then UV-crosslinked using a Stratalinker 2400 instrument for 10 min. RIPA-extracted lysates were then fractionated with ammonium sulfate with 20% increments. These fractions were then separated via SDS–PAGE and relevant probe labelling was determined by anti-biotin (1:500, ab1227, Abcam) western blotting as above. A parallel gel was silver stained using the Pierce silver stain kit. Relevant gel slices from the 80% fraction were excised and PGK1 identity was determined by LC–MS/MS by the Scripps Center for Metabolomics and Mass Spectrometry. Follow-up shRNA knockdown studies confirmed PGK1 as the target within this fraction.

Dye-based thermal denaturation assay. Thermal denaturation experiments were performed using a Protein Thermal Shift Dye Kit (ThermoFisher, 4461146). Reactions contained 2 μM recombinant PGK1 with the indicated dose of aqueously delivered CBR-470-1 with 1× supplied thermal shift dye and reaction buffer in 20 μl reaction volumes. Fluorescence values were recorded using a Viia7 Real-Time PCR instrument according to supplied instructions.

Recombinant PGK1 assay. PGK1 enzymatic activity in the forward direction was measured with a coupled enzymatic assay[34]. Three PGK1 conditions were prepared by dissolving recombinant PGK1 in potassium phosphate buffer (10 mM KH_2PO_4, 10 mM $MgSO_4$, pH 7.0), and transferring the aliquots of PGK1 solution to the microtubes being treated with same amount of DMSO and indicated concentrations of CBR-470-1. Final concentration of PGK1 is 20 ng ml^{-1} and DMSO is 1% for each sample. Two blank conditions, 0 μM and 100 μM of CBR-470-1 without PGK1, were also prepared for the control measurements. All PGK1 samples and blank samples were pre-incubated for 20 min and then transferred to the UV-transparent 96 well plate (Corning). The assay solution (10 mM KH_2PO_4, 2 mM G3P, 0.6 mM NAD^+, 200 mM glycine, 0.4 mM ADP, pH 7.0) was activated by adding GAPDH with 10 μg ml^{-1} final concentration, and then the assay solution was added to the wells containing PGK1 samples and blank samples. The change in absorbance at 340 nm at room temperature was measured every 20 s for 45 min, by Tecan Infinite M200 plate reader. Each condition was performed with three independent replications.

Isothermal dose response profiling of PGK1. In vitro thermal profiling assay for recombinant proteins was performed by dissolving pure recombinant PGK1 and GAPDH into PBS, and dividing equal amount of mixture into 9 aliquots. Each aliquot was transferred to 0.2 ml PCR microtubes being treated with different amounts of CBR-470-1 added from DMSO stock, and equal amount of DMSO for the control. Each microtube contains 50 μl of mixture with final concentration of 45 μg ml^{-1} for each protein and DMSO concentration 1% with following final concentrations of CBR-470-1; 0 μM, 0.1 μM, 0.3 μM, 1 μM, 3 μM, 10 μM, 33 μM, 100 μM and 333 μM. After a 30 min incubation at 25 °C, samples were heated at 57 °C for 3 min followed by cooling at 25 °C for 3 min using Thermal Cycler. The heated samples were centrifuged at 17,000*g* for 20 min at 4 °C, and the supernatants were transferred to new Eppendorf tubes. Control experiments were performed with heating at 25 °C for 3 min, instead of 57 °C. Samples were analysed by SDS–PAGE and western blot.

Metabolite treatments and HMM-KEAP1 screening. For in vitro screening of glycolytic metabolites, HEK293T cells expressing Flag–KEAP1 were lysed by snap-freeze-thaw cycles (3×) in PBS, pH 7.4, containing EDTA-free complete protease inhibitor (Roche). Lysates were cleared by centrifugation and the supernatants normalized for concentration by Bradford reagent (2 mg ml^{-1}). Concentrated stocks of each metabolite were made in PBS, which were added to the lysate samples for the final indicated concentrations and incubated at 37 °C for 2.5 h with shaking. After incubation, samples were denatured with 6 M urea and processed for SDS–PAGE and western blotting. Methylglyoxal (40% (v/v) with H_2O), GAP, dihydroxyacetone phosphate, and 2,3-BPG were all obtained from Sigma Aldrich and used as PBS stocks. In situ metabolite treatments were performed in HEK293T cells 24 h after transfection of Flag–KEAP1, treated with MGO (1 or 5 mM) in H_2O (Sigma) or equivalent vehicle alone for 8 h. Cells were collected by scraping, washed in PBS and centrifuged, and lysed in urea lysis buffer and analysis by SDS–PAGE and western blot. Dose–response experiments were performed with high purity MGO was prepared by acidic hydrolysis of MG-1,1-dimethylacetal (Sigma Aldrich) followed by fractional distillation under reduced pressure and colorimetric calibration of the distillates, as previously reported[33]. For in vitro MGO dose–response dimerization of KEAP1, HEK293T cells expressing Flag–KEAP1 were lysed in PBS as indicated above, then serial dilutions of high purity MGO in 50 mM sodium phosphate, pH 7.4, were added to the equal volume of lysate aliquots with final protein concentration of 1 mg ml^{-1}. Each mixture was incubated at 37 °C for 8 h with rotating, HMM-KEAP1 formation was analysed by SDS–PAGE and western blot.

For studies with recombinant KEAP1, Flag–KEAP1 was expressed in HEK293T cells from transient transfection of the Flag–KEAP1 plasmid (Addgene plasmid 28023). Flag–KEAP1 protein was immunopurified after overnight incubation at 4 °C with anti-Flag M2 magnetic beads (Sigma) in RIPA buffer in the presence of protease inhibitors, eluted with 3×Flag peptide (150 ng ml^{-1}) in PBS, and desalted completely into PBS. 500 ng of purified Flag–KEAP1 protein was then subjected to reducing conditions with the addition of either TCEP (0.1 mM) or DTT (1 mM) for 10 min at 37 °C. MGO was then added to a final concentration of 5 mM and incubated for 2 h at 37 °C. Reactions were quenched by the addition of 50 μl of 4× sample buffer and subsequent boiling for 10 min. 12 μl of this reaction was then separated by SDS–PAGE and the presence of HMM-KEAP1 evaluated by anti-Flag western blotting as described or by silver staining using the Pierce Silver Stain Kit (ThermoFisher Scientific).

Site-directed mutagenesis of KEAP1. KEAP1 mutants were generated with PCR primers in Extended Data Table 1 according to the Phusion site-directed mutagenesis kit protocol (F-541, Thermo Scientific) and the QuikChange site-directed mutagenesis kit protocol (200523, Agilent). Mutant KEAP1 plasmids were verified by sequencing (CMV (forward), wild-type primers in the middle of the *KEAP1* sequence (forward) and BGH (reverse)), and were transiently expressed in HEK293T cells in the same manner as wild-type KEAP1. Screening of CBR-470-1-induced HMM-KEAP1 formation with mutant constructs was performed just as with wild-type KEAP1, after 8 h CBR-470-1 treatment (20 μM). After treatment, cells were collected and prepared for SDS–PAGE and western blotting as indicated above.

SILAC cell culture methods and proteomic sample preparation. SILAC labelling was performed by growing cells for at least five passages in lysine- and arginine-free SILAC medium (RPMI, Invitrogen) supplemented with 10% dialysed fetal calf serum, 2 mM L-glutamine and 1% penicillin/streptomycin. 'Light' and 'heavy' media were supplemented with natural lysine and arginine (0.1 mg ml^{-1}), and ^{13}C-, ^{15}N-labelled lysine and arginine (0.1 mg ml^{-1}), respectively.

General protein digestion for LC–MS/MS analysis was performed by dissolving protein (for example, whole lysate or enriched proteins) in digestion buffer (8 M urea, 50 mM NH_4HCO_3, pH 8.0), followed by disulfide reduction with DTT (10 mM, 40 min, 50 °C), alkylation (iodoacetamide, 15 mM, 30 min, room temperature, protected from light) and quenching (DTT, 5mM, 10 min, room temperature). The proteome solution was diluted fourfold with ammonium bicarbonate solution (50 mM, pH 8.0), $CaCl_2$ added (1 mM) and digested with sequencing grade trypsin (~1:100 enzyme/protein ratio; Promega) at 37 °C while rotating overnight. Peptide digestion reactions were stopped by acidification to pH 2–3 with 1% formic acid, and peptides were then desalted on ZipTip C18 tips (100 μl, Millipore), dried under vacuum, resuspended with LC–MS grade water (Sigma Aldrich), and then lyophilized. Lyophilized peptides were dissolved in LC–MS/MS buffer A (H_2O with 0.1% formic acid, LC–MS grade, Sigma Aldrich) for proteomic analysis.

Proteomic LC–MS/MS and data analysis. LC–MS/MS experiments were performed with an Easy-nLC 1000 ultra-high pressure LC system (ThermoFisher) using a PepMap RSLC C18 column heated to 40 °C (column: 75 μm × 15 cm; 3 μm, 100 Å) coupled to a Q Exactive HF orbitrap and Easy-Spray nanosource (ThermoFisher). Digested peptides (500 ng) in MS/MS buffer A were injected onto the column and separated using the following gradient of buffer B (0.1% formic acid acetonitrile) at 300 nl min^{-1}: 0–2% buffer B over 10 min, 2–40% buffer B over 120 min, 40–70% buffer B over 10 min, and 70–100% buffer B over 5 min. MS/MS spectra were collected from 0 to 150 min using a data-dependent, top-20 ion setting with the following settings: full MS scans were acquired at a resolution of 120,000, scan range of 400–1,600 *m*/*z*, maximum IT of 50 ms, AGC target of 1×10^6, and data collection in profile mode. MS/MS scans were performed by HCD fragmentation with a resolution of 15,000, AGC target of 1×10^5, maximum IT of 30 ms, NCE of 26, and data type in centroid mode. Isolation window for precursor ions was set to 1.5 *m*/*z* with an underfill ratio of 0.5%. Peptides with charge state >5, 1 and undefined were excluded and dynamic exclusion was set to eight seconds. Furthermore, S-lens RF level was set to 60 with a spray voltage value of 2.60 kV and ionization chamber temperature of 300 °C.

MS/MS files were generated and searched using the ProLuCID algorithm in the Integrated Proteomics Pipeline (IP2) software platform. Human proteome data were searched using a concatenated target/decoy UniProt database (UniProt_Human_reviewed_04-10-2017.fasta). Basic searches were performed with the following search parameters: HCD fragmentation method; monoisotopic precursor ions; high resolution mode (3 isotopic peaks); precursor mass range 600–6,000 and initial fragment tolerance at 600 p.p.m.; enzyme cleavage specificity at C-terminal lysine and arginine residues with three missed cleavage sites permitted; static modification of +57.02146 on cysteine (carboxyamidomethylation); two total differential modification sites per peptide, including oxidized methionine (+15.9949); primary scoring type by XCorr and secondary by Zscore; minimum peptide length of six residues with a candidate peptide threshold of 500. A minimum of one peptide per protein and half-tryptic peptide specificity were required. Starting statistics were performed with a Δmass cutoff = 15 p.p.m. with modstat, and trypstat settings. False discovery rates of peptide (sfp) were set to 1%, peptide

modification requirement (-m) was set to 1, and spectra display mode (-t) was set to 1. SILAC searchers were performed as above with light and heavy database searches of MS1 and MS2 files by including static modification of +8.014168 for lysine and +10.0083 for arginine in a parallel heavy search. SILAC quantification was performed using the QuantCompare algorithm, with a mass tolerance of 10 p.p.m. or less in cases in which co-eluting peptide interferes. In general, all quantified peptides have a mass error within 3 p.p.m.

Quantitative proteomic detection of potential KEAP1 modification sites. Quantitative surface mapping with SILAC quantitative proteomics was performed with heavy- and light-labelled HEK293T cells expressing Flag–KEAP1. Cells were incubated with DMSO alone (light cells) or CBR-470-1 (20 μM, heavy cells) for 8 h. After incubation cells were scraped, washed with PBS (3×) and combined before lysis in urea lysis buffer (8 M urea, 50 mM NH_4HCO_3, nicotinamide (1 mM), phosphatase inhibitor cocktail (Sigma Aldrich), and EDTA-free complete protease inhibitor (Roche), pH 8.0) by sonication at 4 °C. After sonication insoluble debris was cleared by centrifugation (17,000*g*, 10 min), diluted with Milli-Q water to give 1 M urea, and lysate was incubated with anti-Flag M2 resin (100 μl slurry, A2220, Sigma Aldrich) at 4 °C overnight while rotating. For SILAC label-swap experiments, light HEK293T cells were incubated with CBR-470-1 and heavy cells were incubated with DMSO and processed as above. Flag resin was washed with PBS (7 × 1 ml), Flag–KEAP1 protein eluted with glycine-HCl buffer (0.1 M glycine, pH 3.5, 2 × 500 μl), followed by 8 M urea (2 × 100 μl). The combined eluent was brought up to 8 M urea total concentration and processed for trypsin digestion and LC–MS/MS analysis as indicated above.

The SILAC maps were generated by comparing SILAC ratios for each peptide, relative to the median value for all KEAP1 peptides. SILAC ratios were converted to Log_2 values and plotted to visualize peptides that are significantly perturbed, for example by modification, relative to the rest of the protein. A minimum of three SILAC ratios for each peptide was required for inclusion in KEAP1 surface maps, which allowed for approximately 85–90% coverage of the KEAP1 protein. Missing sequences were caused by the lack or close spacing of tryptic sites, resulting in inadequate peptides for MS/MS detection.

In vitro MGO peptide reactions. 'CR' peptide was synthesized using standard solid phase peptide synthesis with FMOC-protected amino acids on MBHA rink amide resin. Peptides were cleaved in a solution of 94% trifluoroacetic acid, 2.5% triisopropyl silane, 2.5% H_2O, 1% β-mercaptoethanol and precipitated with ether. Peptide identity was confirmed using an Agilent 1100 series LC–MS. Peptides were purified via reverse phase HPLC on an Agilent Zorbax SB-C18 250 mm column and dried via lyophilization. For MGO reactions CR peptide (1 mM) was incubated with 12.5 mM MGO (diluted from 40% solution in water; Sigma Aldrich) or equivalent amount of water (mock) in 1× PBS pH 7.4 at 37 °C overnight. Reactions were diluted 1:25 in 95%/5% H_2O/acetonitrile and 0.1% trifluoroacetic acid and analysed by LC–MS.

For NMR experiments, approximately 1.5 mg of the CR or CR-MGO crosslinked peptide was purified by reverse phase HPLC, lyophilized and dissolved in 700 μl d_6-DMSO. The peptides were dried via lyophilization. All NMR experiments were performed on a Bruker Avance II+ 500 MHz 11.7 Tesla NMR. Data were processed and plotted in Bruker Topspin 3.5. CR peptide NMR experiments were run with a spectral width of 8.5 for 2D experiments (in both dimensions) and 15 for 1D proton NMR with a pulse width of 13.5 μs and an interscan delay of 3 s. For the proton NMR, 256 scans were taken. For the COSY-DQF experiment, 128 and 2,048 complex points were acquired in the F_1 and F_2 dimensions respectively, with 8 scans per point. For the TOCSY experiment, a mixing time of 60 μs was used, and 256 and 1,024 complex points were acquired with 8 scans per point. All CR-MGO peptide NMR experiments were run with a spectral width of 13 (in both dimensions) with a pulse width of 11.5 μs and an interscan delay of 2.2 s. For the proton NMR, 256 scans were taken. For the COSY-DQF experiment, 128 and 2,048 complex points were acquired in the F_1 and F_2 dimensions, respectively, with 8 scans per point. For the TOCSY experiment, a mixing time of 80 μs was used, and 256 and 1,024 complex points were acquired with 8 scans per point.

In-gel digestion of KEAP1. Targeted proteomic analyses of KEAP1 protein were performed by running anti-Flag enriched HMM-KEAP1 and low-molecular mass (LMM)-KEAP1 (from both CBR-470-1 or MGO treatments as above) on SDS–PAGE gels, and isolated gel pieces were digested in-gel with sequencing grade trypsin (Promega), as previously reported[35]. Tryptic peptides from in-gel tryptic digestions were dissolved in 100 mM Tris-HCl, pH 8.0, with 2 mM of $CaCl_2$, and further digested with mass spectrometry-grade chymotrypsin (Thermo Scientific) according to manufacturer's protocol. Chymotryptic digestion reactions were stopped by acidification, and desalted on Ziptip C18 tips.

Targeted proteomic analysis of crosslinked KEAP1 peptides. Double-digested KEAP1 peptides from isolated HMM-KEAP1 and monomeric KEAP1 were analysed by LC–MS/MS on an Easy-nLC 1000 ultra-high pressure LC system coupled to a Q Exactive HF orbitrap and Easy-Spray nanosource as indicated above. Candidate peptides were initially searched by manual inspection of chromatograms and MS1 spectra for *m/z* values of peptide candidates from predicted digestion sites, crosslink sites and differential presence in HMM- and monomeric KEAP1 from both CBR-470-1-and MGO-treated samples. Extracted MS1 ions of the candidates were present in HMM-KEAP1 digests but not in LMM-KEAP1 digests. MS/MS spectra and PRM experiments were collected on the same instrument using the following settings: global and general settings included lock masses of off, chromatography peak width of 15 s, polarity of positive, in-source CID of 0.0 eV, inclusion list set to 'on,' and an *m/z* value of the target parent ion with its charge state in the inclusion list. MS2 scans were performed by HCD fragmentation with micro-scans of 1, resolution of 120,000, AGC target of 5×10^5, maximum IT of 200 ms, loop count of 1, MSX count of 1, isolation window of 2.0 *m/z*, isolation offset of 0.0 *m/z*, NCE of 16, and spectrum data type in profile mode. Furthermore, S-lens RF level was set to 60 with a spray voltage value of 2.20 kV and ionization chamber temperature of 275 °C. Targeted PRM experiments were performed on CBR-470-1-, MGO-induced HMM-KEAP1 and monomeric KEAP1 samples.

UVB skin damage model. Thirty-two 5-week-old BALB/c male mice were randomized into 4 groups of 8 animals such that each group had similar body weight means. Mice were prepared for removal of hair from their entire back two days before UVB exposure (day 3) by using an electric shaver and depilatory cream. On day 5, mice received exposure to UVB (200 mJ cm^{-2}) produced by a broad band UVB lamp (Dermapal UVB Rev 2) powered by a Kernel UV Phototherapy system. UVB exposure was confined to a rectangular area of ~8 cm^2 by a lead shielding mask. UVB doses were confirmed by dosimeter measurements (Daavlin X96). Sham animals were shaved but received no UVB treatment. Mice were dosed from day 0 to study end at day 10 via oral gavage twice daily (CBR-470-2, 50 mg kg^{-1} twice daily, orally; BARD, 3 mg kg^{-1} orally; vehicle, 0.5% methyl cellulose/0.5% Tween80). Mice were monitored daily for body weight changes and erythema scoring from days 5 to 10. Mice were euthanized at day 10 and specimens collected for histological analysis from the wounded area. These studies were performed at Biomodels, LLC. Blinded erythema scores were recorded by a blinded, trained investigator according to established in house scale. In short, a scale of 0 to 4 was generated with a score of 0 referring to normal skin and a score of 4 indicating severe ulceration.

Percentage wounded area measurements. Photographs of animals on day 10 of the study were taken such that the distances from camera, aperture and exposure settings were identical. Images were then cropped such that only the shaved, wounded area encompassed the imaging field. These images were then processed with a custom ImageJ macro which first performed a three colour image deconvolution to separate the red content of the image[36]. The thresholding function within ImageJ software was then used to separate clear sites of wounding from red background present in normal skin. Red content corresponding to wounds was then quantified as a fraction of the whole imaging field and reported as the percentage wounded area.

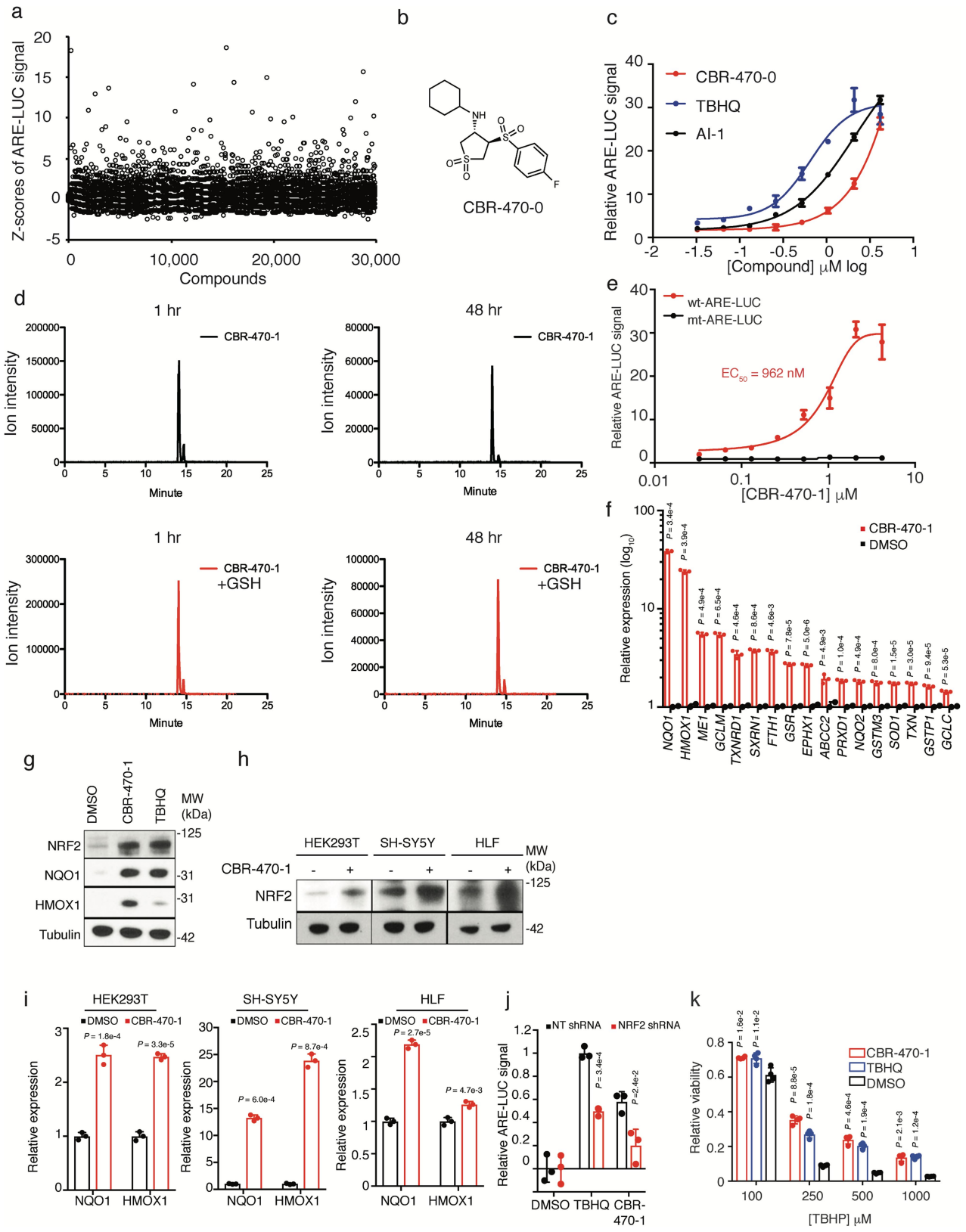
Epidermal thickness measurements. H&E-stained skin sections corresponding to the wounded area were generated by Histotox Labs and accessed via pathxl software. Twenty-four individual measurements of epidermal thickness from 8 sections spanning a 400-μm step distance were recorded per animal by a non-blinded, trained investigator. These measurements were then averaged to generate a mean epidermal thickness measurement per animal.

Reporting summary. Information on research design is available the Nature Research Reporting Summary.

Data availability

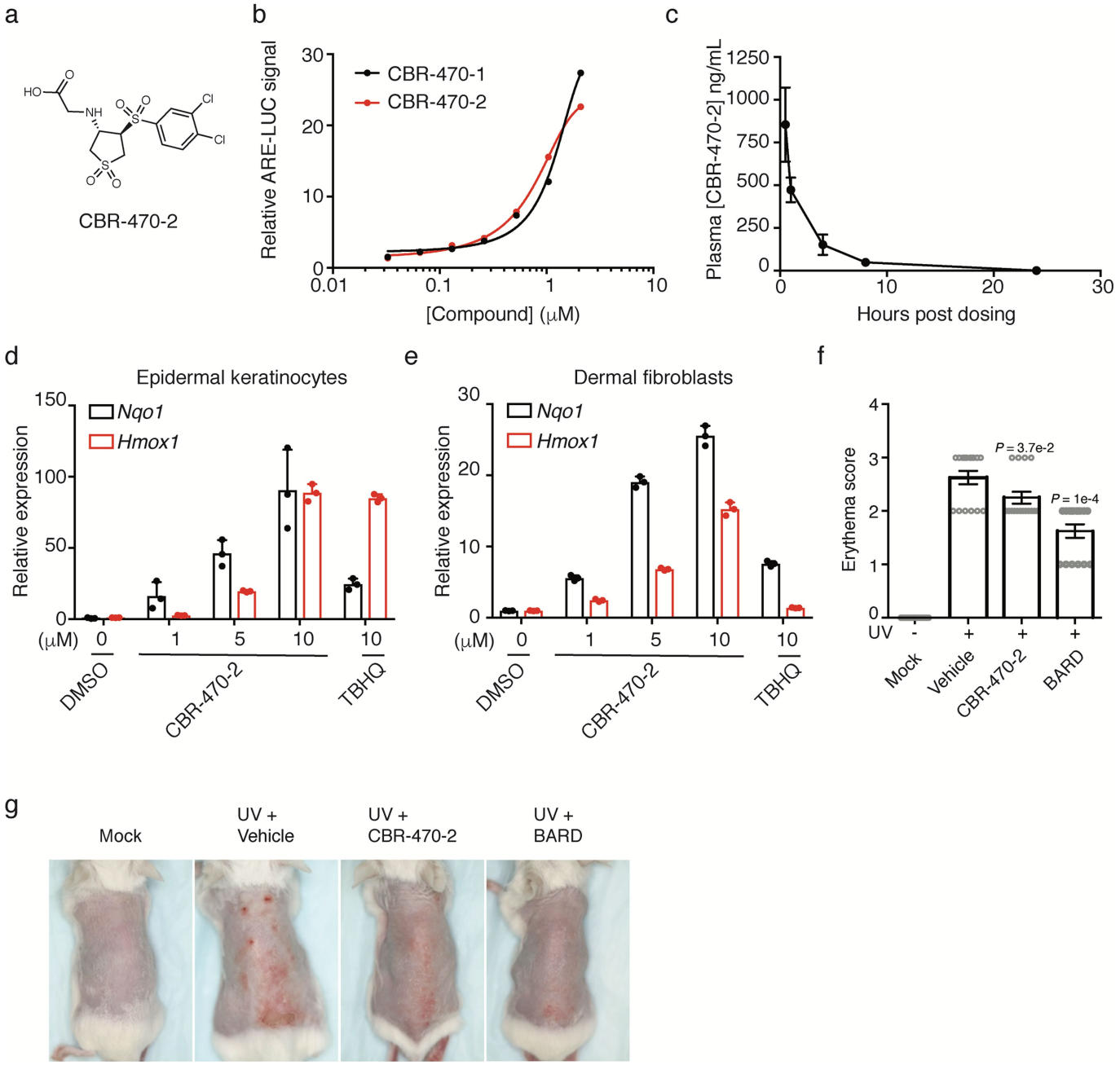
RNA-seq primary data are deposited in the Gene Expression Omnibus (GEO) under accession number GSE116642. Source data for all mouse experiments have been provided. Full scans for western blots and gels are available in the Supplementary Information. All other data are available on reasonable request.

32. Hur, W. & Gray, N. S. Small molecule modulators of antioxidant response pathway. *Curr. Opin. Chem. Biol.* **15**, 162–173 (2011).
33. Rabbani, N. & Thornalley, P. J. Measurement of methylglyoxal by stable isotopic dilution analysis LC–MS/MS with corroborative prediction in physiological samples. *Nat. Protoc.* **9**, 1969–1979 (2014).
34. Bücher, T. Phosphoglycerate kinase from Brewer's yeast. *Methods Enzymol.* **1**, 415–422 (1955).
35. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1**, 2856–2860 (2007).
36. Ruifrok, A. C. & Johnston, D. A. Quantification of histochemical staining by color deconvolution. *Anal. Quant. Cytol. Histol.* **23**, 291–299 (2001).

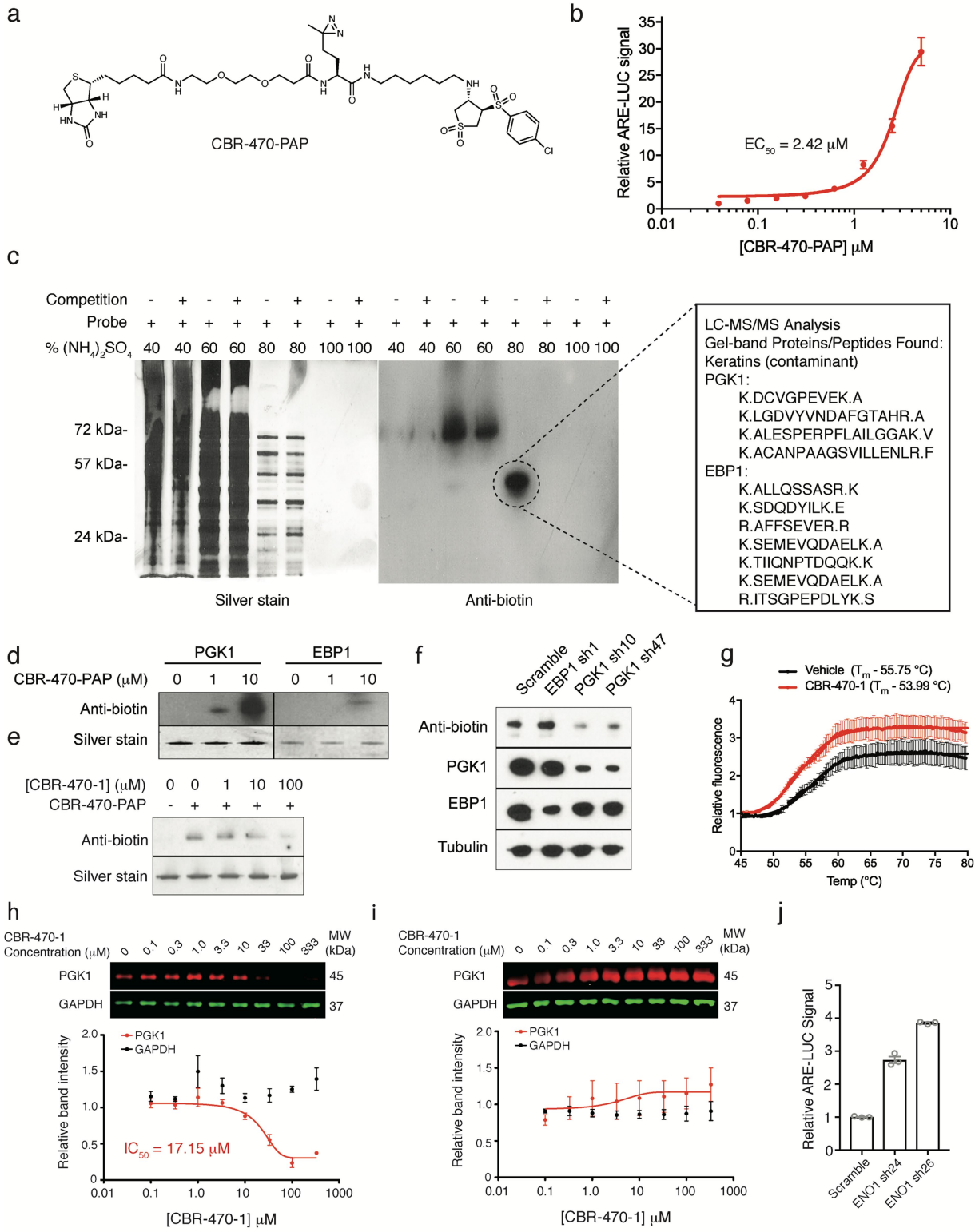


Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | A high-throughput screen identifies a non-covalent NRF2 activator chemical series that activate a robust NRF2 transcriptional program in multiple cell types. a, Plate-based *Z*-scores of ARE-LUC luminance measurements of all test compounds from a 30,000 compound screen in IMR32 cells. **b**, Structure of screening hit CBR-470-0. **c**, Relative ARE-LUC luminance measurements from IMR32 cells treated for 24 h with a concentration response of CBR-470-0 and reported NRF2 activators TBHQ and AI-1 ($n = 3$ biologically independent samples, mean and s.e.m.). **d**, LC–MS quantification of CBR-470-1 (50 μM) incubated in the presence or absence of GSH (1 mM) in PBS for 1 h (left) and 48 h (right). Relative ion intensities within each time point were compared with representative chromatograms shown ($n = 2$). **e**, Relative ARE-LUC luminance values from IMR32 cells transfected with wild-type (wt) or mutant (mt; two core nucleotides necessary for NRF2 binding were changed from GC to AT) ARE-LUC reporter constructs and treated with the indicated doses of CBR-470-1 for 24 h ($n = 3$, mean and s.e.m.). **f**, Relative abundance of NRF2-dependent transcripts as determined by qRT–PCR from IMR32 cells treated for 24 h with 5 μM CBR-470-1 ($n = 3$). **g**, Western blot analyses of total NRF2 protein content or NRF2-controlled genes (*NQO1* and *HMOX1*) from IMR32 cells treated for 24 h with 5 μM CBR-470-1 ($n = 5$). **h**, Western blot analyses of total NRF2 protein content from the indicated cell types treated for 4 h with 5 μM CBR-470-1 ($n = 3$). **i**, Relative expression levels of NQO1 and HMOX1 from the indicated cell types treated for 24 h with 5 μM CBR-470-1 ($n = 3$, mean and s.d.). **j**, Relative ARE-LUC luminescence values from HEK293T cells transfected with the indicated shRNA constructs and pTI-ARE-LUC and then treated with TBHQ (10 μM) or CBR-470-1 (5 μM) for 24 h ($n = 3$). **k**, Relative viability measurements of SH-SY5Y cells treated with either CBR-470-1 (5 μM) or TBHQ (10 μM) for 48 h and then challenged with the indicated doses of TBHP for 8 h ($n = 4$). Data are mean and s.d. or s.e.m. of biologically independent samples. $^{*}P < 0.05$, $^{**}P < 0.005$, $^{***}P < 0.001$, univariate two-sided *t*-test.

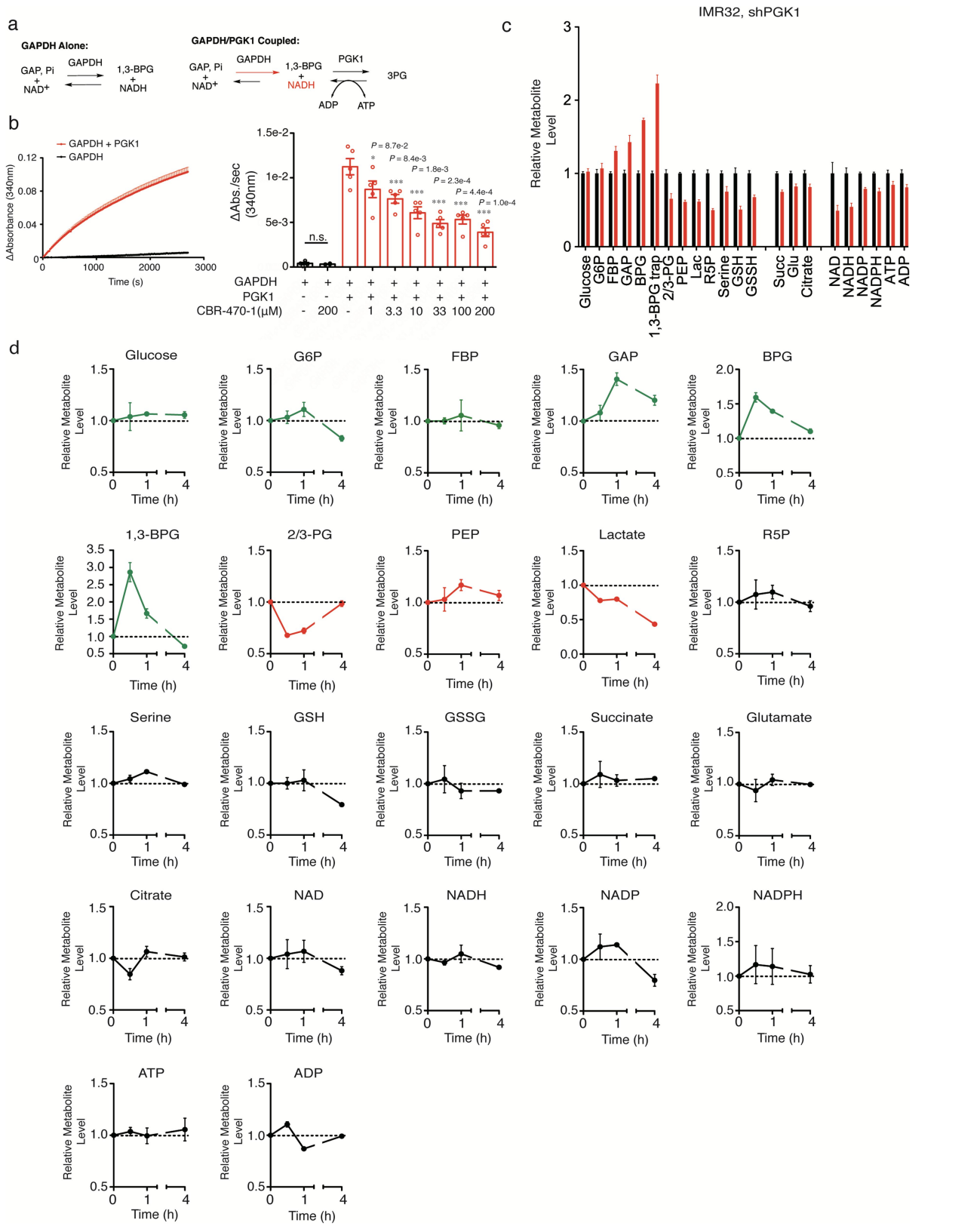


Extended Data Fig. 2 | CBR-470-2 pharmacokinetics and in vivo activity. a, Structure of CBR-470-2. **b**, Relative ARE-LUC luminance values from IMR32 cells transfected with pTI-ARE-LUC and treated with the indicated doses of CBR-470-1 and CBR-470-2 for 24 h ($n = 3$ biologically independent samples). **c**, Plasma concentrations of CBR-470-2 from mice treated with a single 20 mg kg^{-1} dose of compound ($n = 3$ animals, mean and s.e.m.). **d**, **e**, Relative transcript levels of *Nqo1* and *Hmox1* from mouse epidermal keratinocytes (**d**) and mouse dermal fibroblasts (**e**) treated for 24 h with the indicated doses of compound ($n = 3$ biologically independent samples, mean and s.d.). **f**, Blinded erythema scores from mice treated with vehicle, CBR-470-2 or bardoxolone after exposure to UV ($n = 8$ animals, mean and s.e.m.). *$P < 0.05$, *** $P < 0.005$, one-way ANOVA with Dunnett's correction. **g**, Representative images of UV-exposed dorsal regions of animals at day 10 of the study.

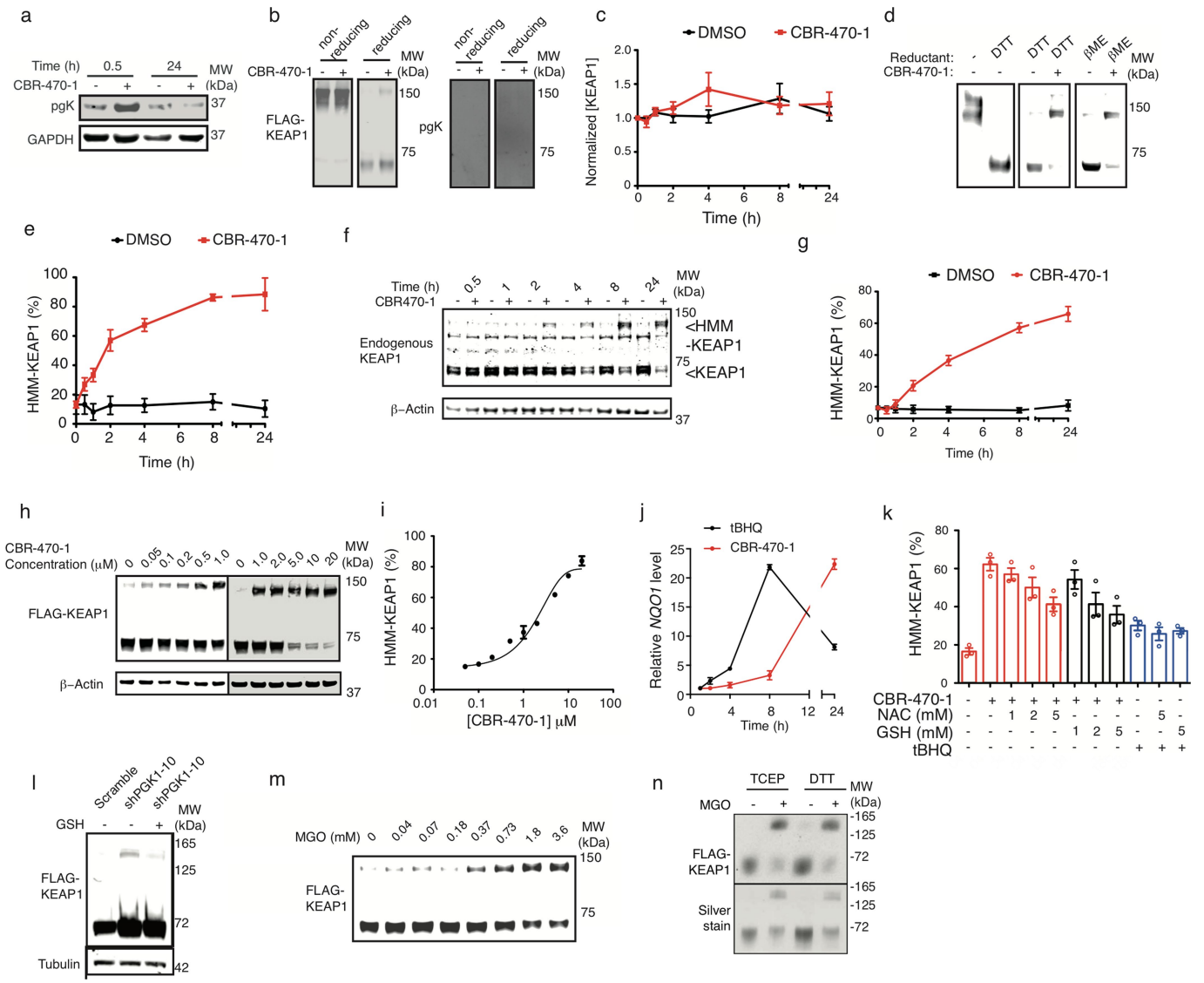


Extended Data Fig. 3 | See next page for caption.

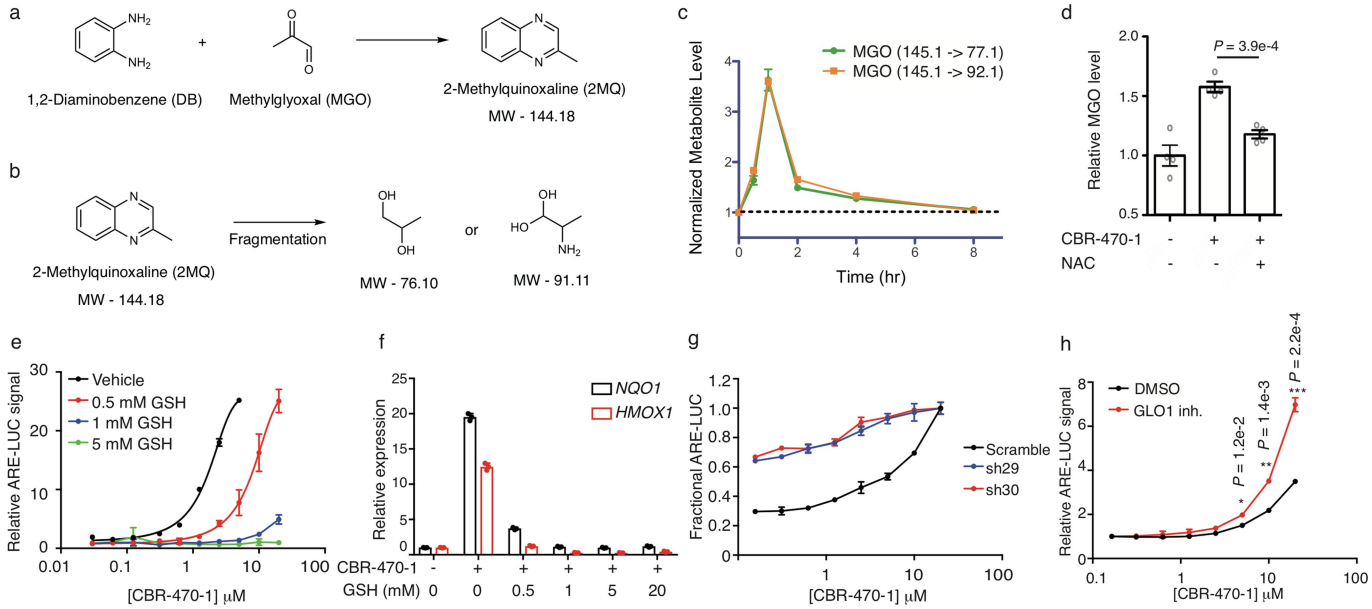
Extended Data Fig. 3 | A photoactivatable affinity probe-based approach identifies PGK1 as the relevant cellular target of CBR-470-1. **a**, Structure of CBR-470-PAP. **b**, Relative ARE-LUC luminance values from IMR32 cells transfected with pTI-ARE-LUC and then treated with the indicated doses of CBR-470-PAP for 24 h ($n = 3$). **c**, Silver staining and anti-biotin western blots of ammonium sulfate fractionated lysates from UV-irradiated IMR32 cells treated with 5 μM for 1 h with or without CBR-470-1 competition (250 μM) ($n = 3$). Shown on the right are initial proteomic target results from gel-band digestion and LC–MS/MS analysis. **d**, Anti-biotin western blots from in vitro crosslinking assays with recombinant PGK1 and EBP1 in the presence of the indicated doses of CBR-470-PAP ($n = 2$). **e**, Anti-biotin western blot analyses from an in vitro crosslinking assay with recombinant PGK1 in the presence of CBR-470-PAP (1 μM) and indicated concentration of soluble CBR-470-1 competitor ($n = 2$). **f**, Anti-biotin western blot analyses of cells treated with 5 μM CBR-470-PAP after transduction with anti-PGK1 and anti-EBP1 shRNA for 48 h. Depletion of PGK1 protein selectively reduces CBR-470-PAP-dependent labelling ($n = 2$). **g**, Dye-based thermal denaturation assay with recombinant PGK1 in the presence CBR-470-1 (20 μM) or vehicle alone ($n = 3$). Calculated melting temperature (T_m) values are listed. **h**, **i**, Dose-dependent thermal stability assay of recombinant PGK1 and GAPDH in the presence of increasing doses of CBR-470-1 near the T_m of both proteins (57 °C) ($n = 5$) (**h**) or room temperature ($n = 3$) (**i**). Western blot of sample supernatants after centrifugation (13,000 r.p.m.) detected total PGK1 and GAPDH protein, which were plotted in Prism (below). **j**, ARE-LUC reporter activity in HEK293T cells with transient shRNA knockdown of *ENO1* ($n = 3$). Data are mean ± s.e.m. of biologically independent samples.



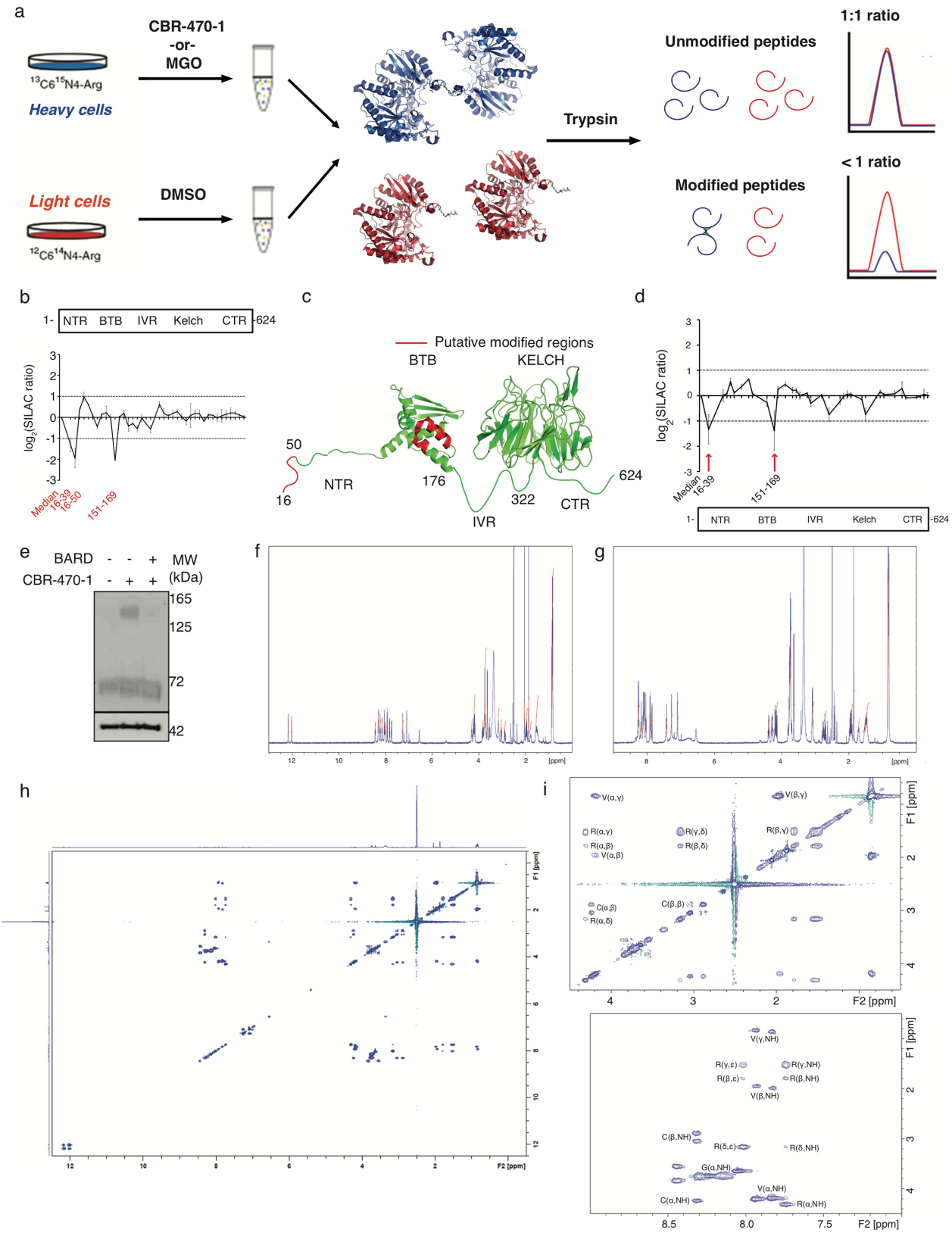
Extended Data Fig. 4 | CBR-470-1 inhibits PGK1 in vitro and in situ. **a**, Schematic of the GAPDH/PGK1 coupled assay. Pre-equilibration of the GAPDH reaction (top left) results in an NAD^+/NADH equilibrium, which after addition of PGK1 and ADP pulls the reaction to the right, producing more NADH. Monitoring NADH absorbance after the addition of PGK1 (bottom right) can be used to monitor PGK1 activity in the forward direction (right). Kinetic monitoring of NADH absorbance (340 nm) after established equilibrium with GAPDH shows little change (black curve), but is significantly increased after the addition of PGK1, pulling the equilibrium to the right (red curve). **b**, CBR-470-1 does not affect the GAPDH equilibrium alone, but significantly inhibits PGK1-dependent activity and accumulation of NADH ($n = 5$). **c**, **d**, Relative level of central metabolites in IMR32 cells treated with viral knockdown of PGK1 for 72 h (**c**) ($n = 4$) and with CBR-470-1 relative DMSO alone for the indicative times (**d**) ($n = 3$). Each metabolite is normalized to the control condition at each time point. Statistical analysis was by univariate two-sided *t*-test (**b**). Data are mean ± s.e.m. of biologically independent samples.



Extended Data Fig. 5 | Modulation of PGK1 induces HMM-KEAP1. **a**, Anti-phosphoglyceryl-lysine (pgK) and anti-GAPDH western blots analysis of CBR-470-1 or DMSO-treated IMR32 cells at early (30 min) and late (24 h) time points ($n = 6$). **b**, Anti-Flag (left) and anti-pgK (right) western blot analysis of affinity purified Flag–KEAP1 from HEK293T cells treated with DMSO or CBR-470-1 for 30 min. Duplicate samples were run under non-reducing (left) and reducing (DTT, right) conditions ($n = 6$). **c**, Densitometry quantification of total endogenous KEAP1 levels (combined bands at approximately 70 and 140 kDa) in IMR32 cells treated with DMSO or CBR-470-1 for the indicated times ($n = 6$). **d**, Western blot detection of Flag–KEAP1 in HEK293T cells comparing non-reducing reagent to DTT (left), and the stability of CBR-470-1-dependent HMM-KEAP1 to the presence of DTT (12.5 mM final concentration, middle) or β-mercaptoethanol (βME; 5% (v/v) final concentration, right) during sample preparation. **e**, Time-dependent CBR-470-1 treatment of HEK293T cells expressing Flag–KEAP1. Time-dependent assays were run with 20 μM CBR-470-1 with western blot analysis at the indicated time-points ($n = 8$). **f**, **g**, Western blot detection (**f**) and quantification (**g**) of endogenous KEAP1 and β-actin in IMR32 cells treated with DMSO or CBR-470-1 for the indicated times ($n = 6$). Arrows indicate monomeric KEAP1 (70 kDa) and HMM-KEAP1 (140 kDa) bands. **h**, **i**, Western blot (**h**) detection and quantification (**i**) of Flag–KEAP1 in HEK293T cells exposed to increasing doses of CBR-470-1 ($n = 3$). **j**, Kinetic qRT–PCR measurement of *NQO1* mRNA levels from IMR32 cells treated with TBHQ (10 μM) or CBR-470-1 (10 μM) for the indicated times ($n = 3$). **k**, Quantification of HMM-KEAP1 formation after treatment with CBR-470-1 or the direct KEAP1 alkylator TBHQ, in the presence or absence of GSH or *N*-acetylcysteine (NAC) ($n = 3$). All measurements were taken after 8 h of treatment in Flag–KEAP1-expressing HEK293T cells. **l**, Transient shRNA knockdown of *PGK1* induced HMM-KEAP1 formation, which was blocked by co-treatment of cells with GSH ($n = 3$). **m**, Anti-Flag western blot analysis of Flag–KEAP1 monomer and the HMM–KEAP1 fraction, with dose-dependent incubation of distilled MGO in lysates from HEK293T cells expressing Flag–KEAP1 ($n = 4$). **n**, SDS–PAGE gel (silver stain) and anti-Flag western blot analysis of purified KEAP1 treated with MGO under the indicated reducing conditions for 2 h at 37 °C ($n = 3$). Purified protein reactions were quenched in 4× SDS loading buffer containing β-mercaptoethanol and processed for gel analysis as in **d**. Data are mean ± s.e.m. of biologically independent samples.

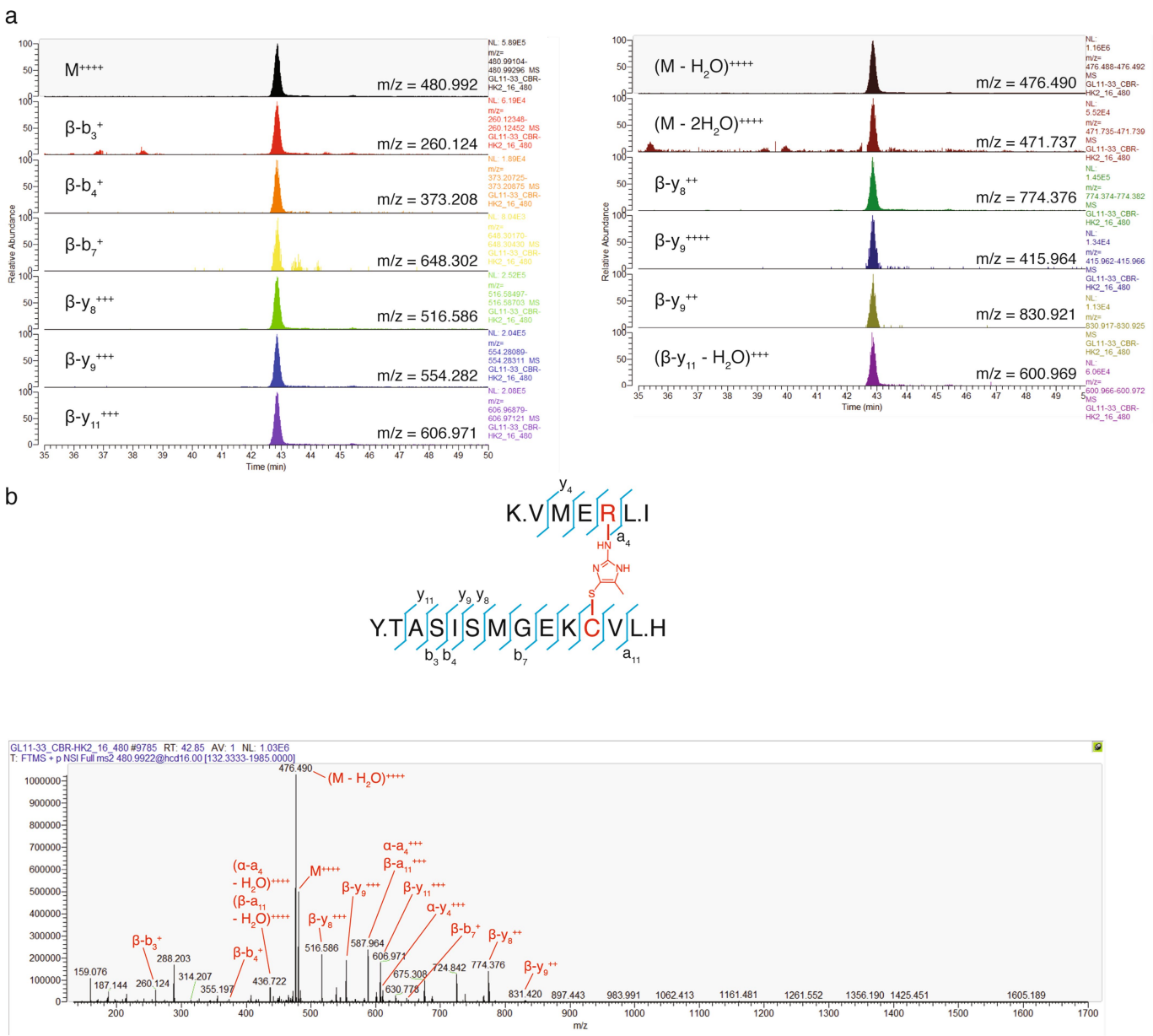


Extended Data Fig. 6 | MGO and glyoxylase activity regulates NRF2 activation. CBR-470-1 causes increased levels of MGO in cells. **a**, Schematic depicting chemical derivatization and trapping of cellular MGO for analysis by targeted metabolomics using two unique fragment ions. **b**, **c**, Daughter ion fragments (**b**) and resulting MS/MS quantification of MGO levels (**c**) in IMR32 cells treated with CBR-470-1, relative to DMSO ($n = 4$). **d**, Quantitative LC–MS/MS measurement of cellular MGO levels in IMR32 cells treated for 2 h with CBR-470-1 or co-treated for 2 h with CBR-470-1 and *N*-acetylcysteine (2 mM) relative to DMSO ($n = 4$). **e**, Relative ARE-LUC luminance values from IMR32 cells transfected with pTI-ARE-LUC and co-treated with the indicated doses of CBR-470-1 and GSH ($n = 3$). **f**, Relative levels of *NQO1* and *HMOX1* transcripts from IMR32 cells co-treated with CBR-470-1 (10 μM) and the indicated concentrations of GSH for 24 h ($n = 3$). **g**, Fractional ARE-LUC values from HEK293T cells transiently co-transfected with pTI-ARE-LUC and the indicated shRNAs and then treated for 24 h with the indicated doses of CBR-470-1 ($n = 3$). **h**, ARE-LUC reporter activity in HEK293T cells treated with CBR-470-1 alone (black) and with a cell-permeable small-molecule GLO1 inhibitor (red) ($n = 3$). Statistical analysis was by univariate two-sided *t*-test (**d**, **h**). Data are mean ± s.e.m. of biologically independent samples.



Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Schematic of SILAC-based proteomic mapping of KEAP1 modifications in response to CBR-470-1 and NMR characterization of CR-MGO peptide. a, SILAC experiments in which Flag-tagged KEAP1 was treated with vehicle ('light') and CBR-470-1 or MGO ('heavy'). Subsequent mixing of the cell lysates, anti-Flag enrichment, tryptic digestion and LC–MS/MS analysis permitted the detection of unmodified portions of KEAP1, which retained approximate 1:1 SILAC ratios relative to the median ratios for all detected KEAP1 peptides. By contrast, peptides that are modified under one condition will no longer match tryptic MS/MS searches, resulting in skewed SILAC ratios that 'drop out' (bottom). **b**, SILAC ratios for individual tryptic peptides from Flag–KEAP1-enriched DMSO-treated light cells and CBR-470-1-treated heavy cells, relative to the median ratio of all KEAP1 peptides. Highlighted tryptic peptides were significantly reduced by three- to fourfold relative to the KEAP1 median, indicative of structural modification ($n = 8$). **c**, Structural depiction of potentially modified stretches of human KEAP1 (red) using published X-ray crystal structure of the BTB (PDB code 4CXI) and KELCH (PDB code 1U6D) domains. Intervening protein stretches are depicted as unstructured loops in green. **d**, SILAC ratios for individual tryptic peptides from Flag–KEAP1-enriched MGO-treated heavy cell lysates and non-treated light cell lysates, relative to the median ratio of all KEAP1 peptides. Highlighted tryptic peptides were significantly reduced by 2- to 2.5-fold relative to the KEAP1 median, indicative of structural modification ($n = 12$). **e**, Representative western blotting analysis of Flag–KEAP1 dimerization from HEK293T cells pre-treated with BARD followed by CBR-470-1 treatment for 4 h ($n = 3$). **f**, ^{1}H-NMR of CR-MGO peptide (isolated product of MGO incubated with Ac-NH-VVCGGGRGG-C(O)NH_2 peptide). ^{1}H-NMR (500 MHz, d_6-DMSO) δ 12.17 (s, 1H), 12.02 (s, 1H), 8.44 (t, J = 5.6 Hz, 1H), 8.32–8.29 (m, 2H), 8.23 (t, J = 5.6 Hz, 1H), 8.14 (t, J = 5.9 Hz, 1H), 8.05 (t, J = 5.9 Hz, 1H), 8.01 (t, J = 5.9 Hz, 1H), 7.93 (d, J = 8.5 Hz, 1H), 7.74 (d, J = 8.0 Hz, 1H), 7.26 (s, 1H), 7.09 (s, 1H), 4.33–4.28 (m, 1H), 4.25-4.16 (m, 3H), 3.83 (dd, J = 6.9 Hz, J = 16.2 Hz, 1H), 3.79-3.67 (m, 6H), 3.63 (d, J = 5.7 Hz, 2H), 3.54 (dd, J = 4.9 Hz, J = 16.2 Hz, 1H), 3.18-3.13 (m, 2H), 3.04 (dd, J = 4.9 Hz, J = 13.9 Hz, 1H), 2.88 (dd, J = 8.6 Hz, J = 13.6 Hz, 1H), 2.04 (s, 3H), 1.96 (sep, J = 6.8 Hz, 2H), 1.87 (s, 3H), 1.80–1.75 (m, 1H), 1.56–1.47 (m, 3H), 0.87–0.82 (m, 12H). **g**, ^{1}H-NMR of CR peptide (Ac-NH-VVCGGGRGG-C(O)NH_2). ^{1}H-NMR (500 MHz, d_6-DMSO) δ 8.27–8.24 (m, 2H), 8.18 (t, J = 5.7 Hz, 1H), 8.13–8.08 (m, 3H), 8.04 (t, J = 5.7 Hz, 1H), 7.91 (d, J = 8.8 Hz), 7.86 (d, J = 8.8 Hz, 1H), 7.43 (t, J = 5.4 Hz, 1H), 7.28 (s, 1H), 7.10 (s, 1H), 4.39 (dt, J = 5.6 Hz, J = 7.4 Hz, 1H), 4.28 (dt, J = 5.7 Hz, J = 7.2 Hz, 1H), 4.21-4.13 (m, 2H), 3.82-3.70 (m, 8H), 3.64 (d, J = 5.8, 2H), 3.08 (dt, J = 6.5 Hz, J = 6.5 Hz, 2H), 2.80-2.67 (m, 2H), 2.43 (t, J = 8.6 Hz, 1H), 1.94 (sep, J = 6.8 Hz, 2H), 1.85 (s, 3H), 1.75–1.68 (m, 1H), 1.54–1.42 (m, 3H), 0.85–0.81 (m, 12H) **h**, ^{1}H–^{1}H total correlation spectroscopy (TOCSY) of CR-MGO peptide. **i**, Peak assignment for CR-MGO peptide TOCSY spectrum. Data are mean ± s.e.m. of biologically independent samples.



Extended Data Fig. 8 | MS2 analysis of CR-MGO-crosslinked KEAP1 peptide. a, Targeted PRM transitions ($n = 6$). **b**, Annotated MS/MS spectrum from the crosslinked C151–R135 KEAP1 peptide.

Extended Data Table 1 | Primer sequences for qPCR and cloning experiments

Gene	Forward Primer Sequence	Reverse Primer Sequence
NQO1	GCCTCCTTCATGGCATAGTT	GGACTGCACCAGAGCCAT
HMOX1	GAGTGTAAGGACCCATCGGA	GCCAGCAACAAAGTGCAAG
ME1	GGAGACGAAATGCATTCACA	ACGAATTCATGGAGGCAGTT
GCLM	GCTTCTTGGAAACTTGCTTCA	CTGTGTGATGCCACCAGATT
TXNRD1	TCAGGGCCGTTCATTTTTAG	GATCTGCCCGTTGTGTTTG
FTH1	GGCAAAGTTCTTCAAAGCCA	CATCAACCGCCAGATCAAC
GSR	TTGGAAAGCCATAATCAGCA	CAAGCTGGGTGGCACTTG
EPHX1	CTTCACGTGGATGAAGTGGA	CTGGCGGAATGAATTTGACT
ABCC2	GGGATCTCTTCCACACTGGAT	CATACAGGCCCTGAAGAGGA
PRDX1	GGGCACACAAAGGTGAAGTC	GCTGTTATGCCAGATGGTCAG
NQO2	TGCGTAGTCTCTCTTCAGCG	GCAACTCCTAGAGCGGTCCT
GSTM3	GGGTGATCTTGTTCTTCCCA	GGGGAAGCTCCTGACTATGA
SOD1	CCACACCTTCACTGGTCCAT	CTAGCGAGTTATGGCGACG
TXNRD1	TCAGGGCCGTTCATTTTTAG	GATCTGCCCGTTGTGTTTG
GSTP1	CTCAAAAGGCTTCAGTTGCC	ACCTCCGCTGCAAATACATC
GCLC	CTTTCTCCCCAGACAGGACC	CAAGGACGTTCTCAAGTGGG
GLO1	TGGATTAGCGTCATTCCAAG	GCGGACCCCAGTACCAAG
PGK1	CTTGGGACAGCAGCCTTAAT	CAAGCTGGACGTTAAAGGGA
TUBG1	ATCTGCCTCCCGGTCTATG	TACCTGTCGGAACATGGAGG

Mutation	Primer (Forward)	Primer (Reverse)
C23S	5'-/5Phos/GCA GGG GAC GCG GTG ATG TAC -3'	5'-/5Phos/CCC CTC AGG AGA CTG TGA CTG CAG GGG C -3'
C38S	5'-/5Phos/GCC CTC CCA GCA TGG CAA -3'	5'-/5Phos/GTC ACC TCC GCC TTG GAC TCA GT -3'
C151S	5'-/5Phos/TGA ACG GTG CTG TCA TGT ACC AGA TC -3'	5'-/5Phos/TGA CGT GGA GGA CAG ACT TCT CGC -3'
C273S	5'-/5Phos/CCG AAC TTC CTG CAG ATG CAG CT -3'	5'-/5Phos/CGT CAA CGA GTG GGA GCG CAC G -3'
C288S	5'-/5Phos/GTC CGA CTC CCG CTG CAA GGA CT -3'	5'-/5Phos/TGC AGG ATC TCG GAC TTC TGC AGC T -3'
C396S	5'-/5Phos/GAC CAA TCA GTG GTC GCC CTG -3'	5'-/5Phos/ATG GGG TTG TAA GAG TCC AGG GC -3'
C405S	5'-/5Phos/CGT GCC CCG TAA CCG CAT CG -3'	5'-/5Phos/CTC ATG GGG GCG CTG GGC G -3'
K39R	5'-/5Phos/GCC CTC CCA GCA TGG CAA -3'	5'-/5Phos/GTC ACC TCC GCC CTG CAC TCA GT -3'
K39M	5'-/5Phos/GCC CTC CCA GCA TGG CAA -3'	5'- GTC ACC TCC GCC ATG CAC TCA GT -3'
C38S/K39M	5'-/5Phos/GCC CTC CCA GCA TGG CAA -3'	5'- GTC ACC TCC GCC ATG GAC TCA GT -3'
K150M	5'-/5Phos/TGA ACG GTG CTG TCA TGT ACC AGA TC -3'	5'- TGA CGT GGA GGA CAC ACA TCT CGC C -3'
R6A	5'- GCA GCC AGA TCC CGC GCC TAG CGG GGC TG -3'	5'- CAG CCC CGC TAG GCG CGG GAT CTG GCT GC -3'
R15A	5'- GGG CCT GCT GCG CAT TCC TGC CCC TGC A -3'	5'- TGC AGG GGC AGG AAT GCG CAG CAG GCC C -3'
R50A	5'- CTC CCA GCA TGG CAA CGC CAC CTT CAG CTA CAC -3'	5'- GTG TAG CTG AAG GTG GCG TTG CCA TGC TGG GAG -3'
R135A	5'- CCC AAG GTC ATG GAG GCC CTC ATT GAA TTC GCC T -3'	5'- AGG CGA ATT CAA TGA GGG CCT CCA TGA CCT TGG G -3'

Extended Data Table 2 | Acquisition parameters used for targeted metabolomic measurements on a triple quadrupole mass spectrometer

Metabolite	Precursor mass	MS1 Resolution	Product ion	MS2 Resolution	Dwell	Fragmentor	Collision Energy	Polarity	Retention time (min)
Glucose	179.05	Wide	89.2	Unit	5	68	12	Neg	12.2
G6P	258.9	Wide	138.9	Unit	100	100	5	Neg	22.3
FBP	339.1	Wide	96.9	Unit	100	100	20	Neg	26.8
GAP	169	Wide	96.9	Unit	100	100	5	Neg	22.1
BPG	264.9	Wide	96.9	Unit	5	86	21	Neg	30.9
2/3-PG	184.98	Wide	78.9	Unit	5	86	21	Neg	24.6
PEP	166.97	Wide	79	Unit	5	78	9	Neg	25.4
Pyruvate	87.1	Wide	43	Unit	100	100	10	Neg	14.8
Lac	89.1	Wide	43	Unit	100	100	20	Neg	13.5
D3-Serine	107.05	Wide	75.1	Unit	5	18	9	Neg	13.9
R5P	228.7	Wide	78.8	Unit	100	100	35	Neg	19.9
Serine	104.2	Wide	73.8	Unit	5	100	5	Neg	13.9
GSH	305.7	Wide	143.0	Unit	100	100	15	Neg	16.7
GSSG	610.7	Wide	305.9	Unit	100	100	15	Neg	20.5
Succ	117	Wide	73.1	Unit	100	100	5	Neg	18.8
Glu	146.1	Wide	102.1	Unit	100	100	5	Neg	15.9
Cit	191	Wide	111	Unit	5	100	5	Neg	24.4
NAD^+	662.1	Wide	540	Unit	100	100	15	Neg	16.1
NADH	663.4	Wide	407.9	Unit	100	100	35	Neg	16.1
$NADP^+$	742	Wide	619.9	Unit	100	100	25	Neg	24.1
NADPH	743.5	Wide	407.8	Unit	100	100	25	Neg	24.1
ATP	506	Wide	159	Unit	100	100	25	Neg	27.5
ADP	425.8	Wide	134	Unit	100	100	15	Neg	26.5
3PGha	199.98	Wide	199.98	Unit	5	116	0	Neg	22.4
3PGha	199.98	Wide	79	Unit	5	116	15	Neg	22.4
2MQ	145.1	Wide	77.1	Unit	5	100	24	Pos	8.5
2MQ	145.1	Wide	92.1	Unit	5	100	20	Pos	8.5
D3-Serine	109.07	Wide	63.1	Unit	5	40	12	Pos	4.3

2/3-PG, 2/3-phosphoglycerate; 2MQ, 2-methylquinoxaline(derivatization product of MGO); 3PGha, 3-phosphoglyceroyl hydroxamic acid (derivatization product of 1,3-BPG). d_3-serine is an isotopically labelled serine standard included in all runs as an internal normalization control.

nature research

Corresponding author(s): Lairson, Schultz & Moellering

☐ Initial submission ☒ Revised version ☐ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

Describe how sample size was determined.

Sample size for cellular and animal experiments were determined according to the minimal number of independent biological replicates powered to significantly identify an effect. In vitro experiments were performed in triplicate independent measurements or more as indicated in the text. Mouse experiments were performed with n = 8 per cohort, which was expected to yield significant results based on previous experience using this mouse model.

2. Data exclusions

Describe any data exclusions.

No data was excluded.

3. Replication

Describe whether the experimental findings were reliably reproduced.

All experiments were independently replicated, with biological and technical replicates listed in the legends of the corresponding figures.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Mouse cohorts were randomized to match mean animal weight between groups.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

In vivo experiments were performed in a single blinded fashion, such that the researchers were not aware of the treatment identity until after completion and unblinding.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	Graphs presented were generated in Prism 5.0, with analysis algorithms indicated in figure legends. Other software used for data analysis or processing are indicated in the corresponding Methods sections.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	All materials are available through commercial vendors or from the authors.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	Commercial Antibodies: Anti-Biotin, Abcam (AB1227), rabbit Anti-FLAG M2, Sigma Aldrich (F1804), mouse Anti-GAPDH, Cell Signaling (2118S), rabbit Anti-PGK1, Santa Cruz (SC-130335), mouse Anti-TUBG, Cell Signaling (5886), rabbit Anti-NQO1, Abcam (ab34173), rabbit Anti-HMOX1, Enzo Life Sciences (ADI-SPA-894), rabbit Anti-ACTB, Cell Signaling (4790), rabbit Anti-KEAP1, Santa Cruz (SC-15246), goat Anti-HSPA1A, Cell Signaling (4872), rabbit Anti-NRF2, Santa Cruz (clone H300), rabbit pan anti-pgK, reference 4, rabbit - generated, affinity purified and validated for pgK specificity as described in reference 4.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	All cell lines were obtained from the ATCC.
b. Describe the method of cell line authentication used.	None of the cell lines were profiled for authentication.
c. Report whether the cell lines were tested for mycoplasma contamination.	Cells lines used in this study tested negative for mycoplasma contamination.
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	No commonly misidentified cell lines were used in this study.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	Animal experiments involved male mice, of the Balb/c strain, aged approximately 3-6 months.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	This study did not involve human subjects.