Quantitative chemical proteomics identifies novel targets of the anti-cancer multi-kinase inhibitor E-3810

Colzani et al.

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure S1. Synthesis of solid supported E-3810. Reagents and conditions: a) 2, DIPEA, DMF, 48h, 55°C; b) LiOH, CH_3OH/H_20, 4h, 40°C; c) 9-Fluorenylmethyl *N*-succinimidyl carbonate, $NaHCO_3$, CH_3CN/H_20, 3h, 0°C; d) N,N'-disuccinimidyl carbonate, DIPEA, CH_3CN, 16h, 21°C; e) Affi-Gel 102, CH_3CN, 16h, 21°C; f) piperidine, CH_3CN, 16h, 21°C. Abbreviations: Bz: benzoyl; DIPEA: *N*,*N*-diisopropylethylamine; DMF: *N*,*N*-dimethylformamide; FMOC: 9-Fluorenylmethyloxycarbonyl.

Supplementary Figure S2. Evaluation of a panel of cell lines as model system **A**) Table reporting FGFR2 gene copy number, FGFR2 mRNA level (expressed as fold change relative to HUVE cells) and IC_{50} values calculated using a proliferation assay upon treatment with E-3810 for the indicated cell lines. **B**) Assessment of FGFR2 protein expression by immuno-blot in A2780 whole cell extracts relative to MDA-MB-134, MDA-MB-231, OVCA-432, OVCAR-8 cell lines. Vinculin was used as loading control.

Supplementary Figure S3. Assessment of the ability of immobilized E-3810 to bind FGFR2 from cell lysates. **A)** Schematic view of the protocol for label-free affinity capture by E-3810-derivatized agarose beads, using non-derivatized beads as mock control. The protocol consists in the following steps: preparation of whole cell extracts from A2780 cells; incubation with either E-3810-derivatized or non-derivatized resin; washes and elution of binding proteins from the beads. **B)** A2780 cell lysates used as input and eluates from six different batches of affinity chromatography (A-F), loaded on SDS-PAGE and visualized by Coomassie Blue staining. **C)** Reproducibility of the

amount of FGFR2 captured by E-3810-derivatized resin in six parallel incubation batches (A-F), visualized by immuno-blot using an anti-FGFR2 antibody. Fifty μg of A2780 lysate were used as input in each experiment.

Supplementary Figure S4. Optimization of SILAC-based affinity chromatography competition assay. **A)** Schematic view of the chemical proteomic strategy based on competition, in a label-free setup: 1 mg of A2780 cellular extract was incubated with 40 μl of E-3810-derivatized resin (corresponding to 120 nmol of immobilized drug), in the presence of increasing amounts of E-3810 (0, 15, 60 and 240 nmol). Upon incubation and washes, proteins retained on the beads were eluted from resin and loaded on SDS-PAGE for separation. **B)** Eluates obtained from each affinity chromatography at increasing doses of free E-3810 were analyzed by immuno-blot using an anti-FGFR2 antibody to estimate the amount of captured FGFR2: the decrease of the signal indicates efficient and progressive competition.

Supplementary figure S5. Protein ratios distribution of binders retained by the E-3810 resin in experiment A (no competition with soluble E-3810), in forward **(A)** and reverse **(B)** setups. In the absence of competitor, no outliers are expected, thus the protein ratio distribution is normal and used to set the cutoff to select the non-outlier proteins, which are comprised within the 5^{th} - 95^{th} percentiles of the distribution (blue dotted lines in both panels).

Supplementary Figure S6. Protein ratio trends from A to D samples for the putative E-3810 interactors detected by the SILAC competition assay. The proteins are grouped as: tyrosine kinases (**A**), serine/threonine kinases (**B**), non-kinase proteins (**C**). H/L ratio values measured in the forward experimental setup are coloured in red, whereas L/H ratios obtained from the reverse experiment are in black.

Supplementary Figure S7. Scatterplot of the complete set of protein ratios calculated in experiments A, B, C and D in the forward and reverse experiment. Putative E-3810 targets selected

upon statistical filtering are highlighted in red (kinase proteins) and green (non-kinase proteins), while proteins classified as background binders based on the filtering criteria are in grey. Dotted lines indicate the 5th percentile cutoff computed for sample A and used for the definition of outliers at the highest competition (sample D).

Supplementary Figure S8. Immobilized E-3810 is present in molar excess over its putative targets in the chemoproteomics assays. **A)** Schematic view of the experimental design: 40 µl of E-3810 resin were incubated with either 1 mg of heavy A2780 lysate (the amount used in the chemoproteomics assays) or a 3-fold excess of light lysate (3 mg). After incubation and washing, the resins were mixed to produce a H/L SILAC sample and the eluted proteins were analyzed by geLC-MS for identification and quantification. H/L ratio depends on protein abundance: eluted proteins, if present at non-saturating conditions, are expected to have H/L ratio values close to 0.33, since they are captured 3 times more in the light (3 mg) than in the heavy lysate (1 mg). Abundant proteins, already present in saturating conditions in 1 mg of lysate, bind to the resin to a similar extent in both heavy and light channels, returning a H/L ratio close to 1.0. **B)** The H/L protein ratio distribution of the 673 proteins that were identified with at least 2 unique peptides and quantified with at least 2 ratio counts has a mean value of 0.66 and standard deviation 0.20, suggesting that the majority of the proteins (i.e., the background binders) was captured by the resin about 1.5-times more in the light than in the heavy lysate. **C)** H/L ratio values for the putative E-3810 targets identified in the K_d assay plotted against their calculated $K_{d\ imm}$ values. The 0.33 ratio value, expected for non-saturating interactors, is indicated by the red line.

Supplementary Figure S9. A) Distribution of iBAQ scores for the 5967 proteins (black dots) identified with at least 2 unique peptides by shotgun proteomics in A2780 cell lysate. After geLC-MS analysis, acquired spectra were processed by MaxQuant to compute the iBAQ score, which estimates protein abundance based on spectral counting: high scores indicate highly abundant proteins. Kinase proteins selected as putative targets by the K_d assay are highlighted in red, while

non kinase proteins are highlighted in blue. **B)** Distribution of IBAQ scores for all kinase proteins identified in the A2780 proteome (black dots). Within this group, kinase proteins selected as putative targets through the chemo-proteomics K_d assay are highlighted in red. Putative target kinases cover the whole range of the kinome iBAQ scores, ruling out biases towards the most abundant kinases. Non kinase proteins and kinase proteins were labelled in panels A and B, respectively.

Supplementary Figure S10. Fragmentation spectra for putative E3810 interactors from the competition assay that were identified with 2 peptides, of which 1 unique (LYN and CSPB(1)).

Supplementary Figure S11. Fragmentation spectra for putative E3810 interactors from the K_d assay that were identified with 2 peptides, of which 1 unique (FGFR2, ACTBL2 and DPDE3).

Supplemental Table S1. Half-maximal inhibitory concentration (IC_{50}) obtained by *in vitro* kinase assay for both non-derivatized and derivatized E-3810 (corresponding to compound 1 and 4 in Figure S1). Compounds were tested against the kinases FGFR-1, -2, -3 and VEGFR-1 and -2, already characterized as targets of E-3810. c-kit and VEGFR-3 were included as negative controls. E-3810 derivatization leads to a slight increase of IC_{50} values, which was considered acceptable for the chemoproteomics screening.

Supplementary Table S2. Unsupervised clustering of putative targets. Putative E-3810 targets selected by SILAC competition assay were grouped in 3 clusters based on the trends of ratio modulation from sample A to D. The three clusters, automatically generated by GproX, group the proteins in the following categories: proteins eluted at low (cluster 1, high affinity), intermediate (cluster 2, medium affinity) and high concentrations (cluster 3, low affinity) of E-3810. For each protein, the number of the cluster is indicated for both the forward and reverse experiments.

Supplementary dataset 1. MaxQuant protein output for the chemoproteomics competition assay, showing protein lists after progressive filtering. Red: summary of results; yellow: list of E3810 interactors; orange: total proteins identified; purple: column name description.

Supplementary dataset 2. MaxQuant protein output for the chemoproteomics Kd assay, showing protein lists after progressive filtering. Red: summary of results; yellow: list of E3810 interactors; green: protein list without reverse and contaminants, identified with at least two peptides, of which at least one unique; orange: total proteins identified; purple: column name description.

Supplementary dataset 3. MaxQuant protein output for the saturation assay, showing protein lists after progressive filtering. Red: summary of results; yellow: proteins considered for the analysis; green: protein list without reverse and contaminants, identified with at least two unique peptides; orange: total proteins identified; purple: column name description.

Supplementary dataset 4. MaxQuant protein output for the iBAQ score calculation, showing protein lists after progressive filtering. Red: summary of results; yellow: proteins considered for the analysis; orange: total proteins identified; purple: column name description.