

## Rapid covalent-probe discovery by electrophile-fragment screening

Efrat Resnick<sup>1</sup>, Anthony Bradley<sup>2,3,4</sup>, Jinrui Gan<sup>1</sup>, Alice Douangamath<sup>4</sup>, Tobias Krojer<sup>2</sup>, Ritika Sethi<sup>2,5,6</sup>, Paul P. Geurink<sup>7</sup>, Anthony Aimon<sup>3,4</sup>, Gabriel Amitai<sup>8</sup>, Dom Bellini<sup>9</sup>, James Bennett<sup>2,10</sup>, Michael Fairhead<sup>2</sup>, Oleg Fedorov<sup>2,10</sup>, Ronen Gabizon<sup>1</sup>, Jin Gan<sup>7</sup>, Jingxu Guo<sup>11</sup>, Alexander Plotnikov<sup>8</sup>, Nava Reznik<sup>1</sup>, Gian Filippo Ruda<sup>2,10</sup>, Laura Díaz-Sáez<sup>2,10</sup>, Verena M. Straub<sup>2,10</sup>, Tamas Szommer<sup>2,10</sup>, Srikannathasan Velupillai<sup>2,10</sup>, Daniel Zaidman<sup>1</sup>, Yanling Zhang<sup>12</sup>, Alun R. Coker<sup>11</sup>, Christopher G. Dowson<sup>9</sup>, Haim M. Barr<sup>8</sup>, Chu Wang<sup>12,13</sup>, Kilian V.M. Huber<sup>2,10</sup>, Paul E. Brennan<sup>2,9,14</sup>, Huib Ovaa<sup>7</sup>, Frank von Delft<sup>2,4,15</sup>, Nir London<sup>1,#</sup>

<sup>1</sup> Department of Organic Chemistry, Weizmann Institute of Science, Rehovot 7610001, Israel.

<sup>2</sup> Structural Genomics Consortium, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7DQ, UK.

<sup>3</sup> Department of Chemistry, Chemistry Research Laboratory, 12 Mansfield Road, Oxford OX1 3TA, UK.

<sup>4</sup> Diamond Light Source Ltd., Harwell Science and Innovation Campus, Didcot OX11 0QX, UK.

<sup>5</sup> Structural Biology Research Center, VIB, Brussels, Belgium.

<sup>6</sup> Structural Biology Brussels, Vrije Universiteit Brussel, Brussels, Belgium.

<sup>7</sup> Oncode Institute and Department of Cell and Chemical Biology, Leiden University Medical Center, Einthovenweg 20, 2333 ZC Leiden, The Netherlands.

<sup>8</sup> Wohl Institute for Drug Discovery of the Nancy and Stephen Grand Israel National Center for Personalized Medicine, Weizmann Institute of Science, Rehovot 76100 Israel.

<sup>9</sup> School of Life Sciences, University of Warwick, Coventry, UK.

<sup>10</sup> Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7FZ, UK.

<sup>11</sup> Division of Medicine, University College London, Gower Street, London WC1E 6BT, UK.

<sup>12</sup> Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, China.

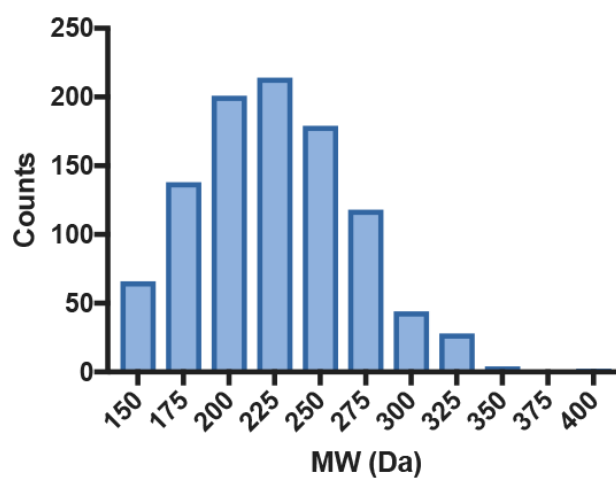
<sup>13</sup> College of Chemistry and Molecular Engineering, Peking University, Beijing, China.

<sup>14</sup> Alzheimer's Research UK Oxford Drug Discovery Institute, University of Oxford, NDMRB, Roosevelt Drive, Oxford OX3 7FZ, UK.

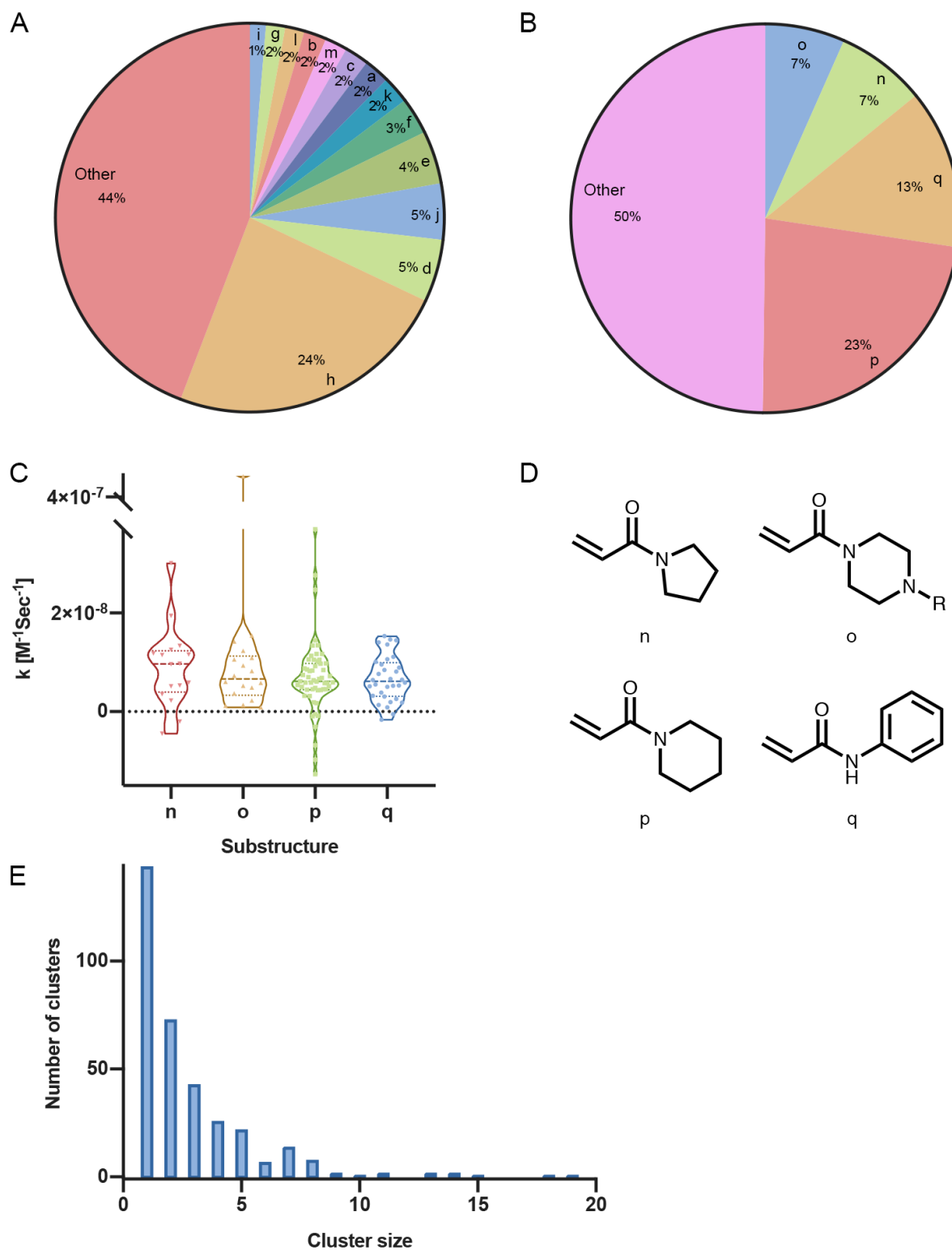
<sup>15</sup> Department of Biochemistry, University of Johannesburg, Auckland Park 2006, South Africa.

# e-mail: nir.london@weizmann.ac.il

## Supplementary Figures

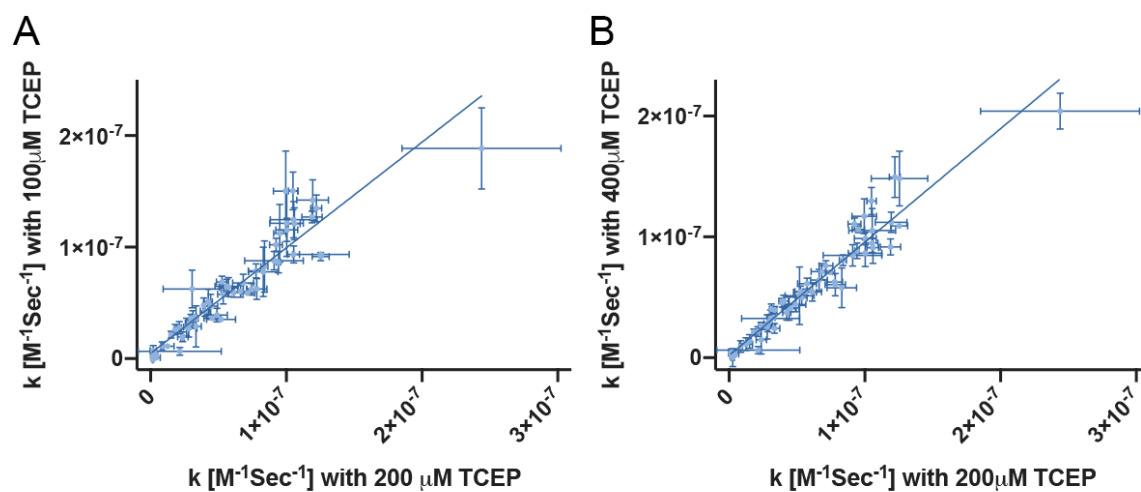


**Supplementary Figure 1. Distribution of molecular weights of the adduct forms of the fragments in the electrophile library**



### Supplementary Figure 2. Diversity of substructures in the library

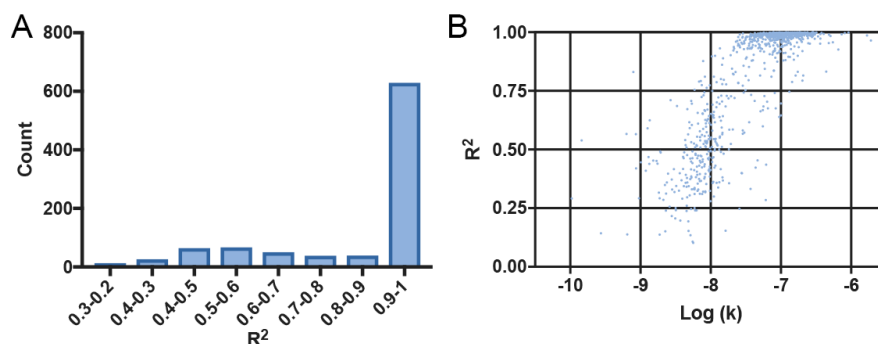
**A.** Proportion of chloroacetamide containing substructures. Named groups (see structures in Fig. 2 of the main text) represent substructures with more than ten instances in the library. **B.** Proportion of acrylamide containing substructures in the library containing more than ten instances. **C.** Distribution of the rate constants of the different acrylamide substructure groups. **D.** Acrylamide substructures with more than ten instances in the library. **E.** Distribution of cluster sizes of the electrophile library, only ten clusters contain ten or more members.



**Supplementary Figure 3. Concentration of TCEP does not affect the measured rate constant in the thiol reactivity assay.**

Calculated rate constant of 64 electrophiles with 100  $\mu\text{M}$ , 200  $\mu\text{M}$  and 400  $\mu\text{M}$  TCEP. **A.** Rate constant correlation between values measured with 100  $\mu\text{M}$  and 200  $\mu\text{M}$  TCEP.  $R^2=0.88$ . **B.** Rate constant correlation between values measured with 400  $\mu\text{M}$  and 200  $\mu\text{M}$  TCEP.  $R^2=0.94$ .

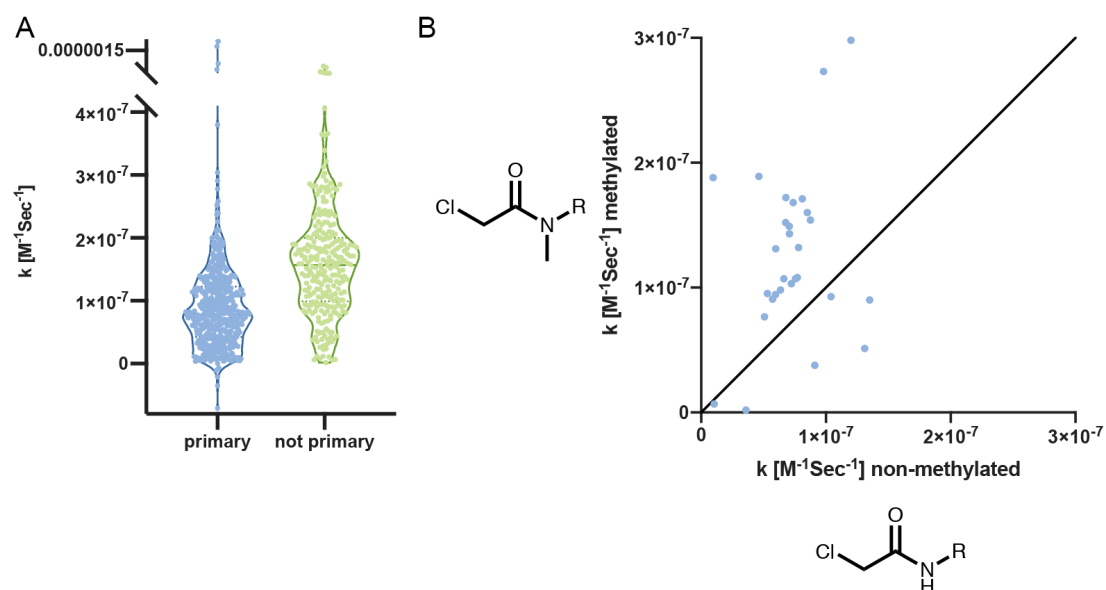




**Supplementary Figure 4. Most kinetic rates fit well to a second order reaction equation**

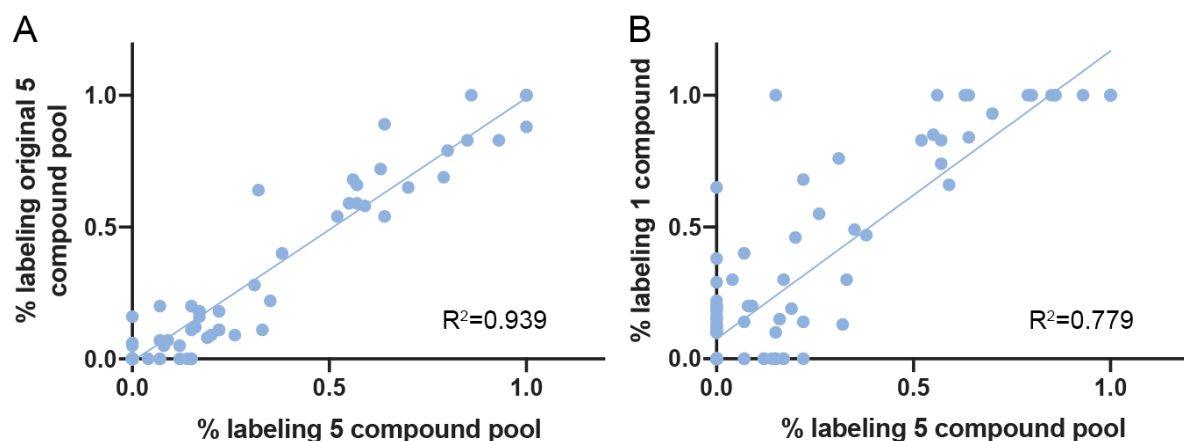
**A.** Distribution of  $R^2$  values for rate calculations. 67% of the compounds show  $R^2 > 0.8$ .

**B.** Goodness of fit ( $R^2$ ) for the linear regression of the kinetic rate calculations (see methods) plotted against the log of the rate constant. Very unreactive compounds show a poor fit. However, most compounds display a very good fit as shown in B.



**Supplementary Figure 5. N,N-disubstituted chloroacetamides are more reactive than singly substituted chloroacetamides.**

**A.** Rate constant distributions of chloroacetamides based on primary amines/anilines and chloroacetamides based on secondary amines/anilines. The latter are significantly more reactive ( $p < 0.0001$ ). **B.** Rate constants of methylated chloroacetamides plotted against the rate constant of their non-methylated counterparts. The line represents  $X=Y$

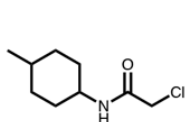


**Supplementary Figure 6. Good correlation between labeling of NUDT7 in pools of five compounds per well compared to one compound per well.**

**A.** Comparison between two independent labeling experiments, for 100 random compounds, performed with the same pooling (five compounds per well).

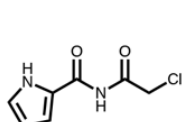
**B.** Comparison between % labeling of NUDT7 by 100 individual compounds plotted against same compounds in pool of five compounds per well.

## NV3CP



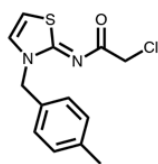
PCM-0102913

Labeling 85%  
k [M<sup>-1</sup>Sec<sup>-1</sup>] 6.63×10<sup>-8</sup>



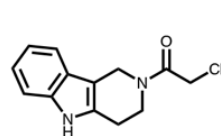
PCM-0102795

79%  
1.01×10<sup>-8</sup>



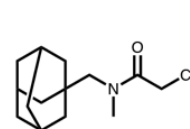
PCM-0102708

56%  
1.45×10<sup>-7</sup>



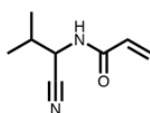
PCM-0102591

56%  
8.13×10<sup>-7</sup>



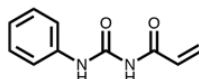
PCM-0102942

53%  
1.52×10<sup>-7</sup>



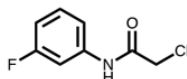
PCM-0102529

Labeling 53%  
k [M<sup>-1</sup>Sec<sup>-1</sup>] 8.65×10<sup>-10</sup>



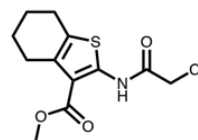
PCM-0102883

52%  
6.06×10<sup>-8</sup>



PCM-0102364

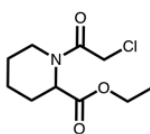
52%  
1.17×10<sup>-7</sup>



PCM-0102143

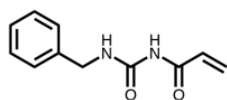
50%  
1.04×10<sup>-10</sup>

## K-Ras<sup>G12C</sup>



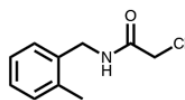
PCM-0102937

Labeling 81%  
k [M<sup>-1</sup>Sec<sup>-1</sup>] 1.09×10<sup>-7</sup>



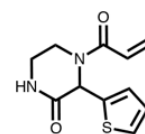
PCM-0102834

79%  
3.52×10<sup>-8</sup>



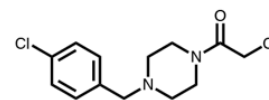
PCM-0102253

70%  
7.05×10<sup>-8</sup>



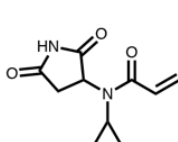
PCM-0102597

70%  
1.54×10<sup>-8</sup>



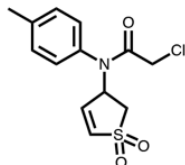
PCM-0102818

63%  
1.94×10<sup>-7</sup>



PCM-0102646

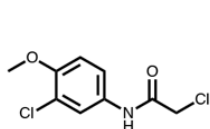
Labeling 62%  
k [M<sup>-1</sup>Sec<sup>-1</sup>] 6.86×10<sup>-9</sup>



PCM-0102539

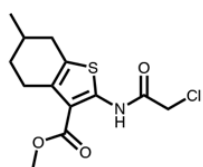
54%  
2.36×10<sup>-7</sup>

# NNMT



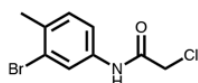
PCM-0102652

Labeling 83%  
k [M<sup>-1</sup>Sec<sup>-1</sup>] 1.55×10<sup>-7</sup>



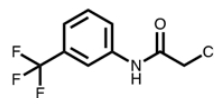
PCM-0102418

80%  
-1.07×10<sup>-8</sup>



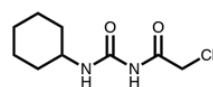
PCM-0102955

78%  
8.04×10<sup>-8</sup>



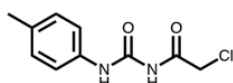
PCM-0102290

77%  
1.83×10<sup>-7</sup>



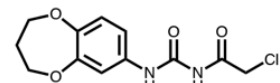
PCM-0102692

76%  
1.16×10<sup>-7</sup>



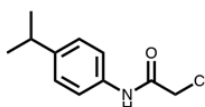
PCM-0102623

Labeling 71%  
k [M<sup>-1</sup>Sec<sup>-1</sup>] 8.56×10<sup>-8</sup>



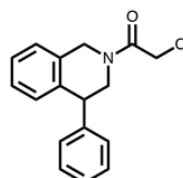
PCM-0102910

69%  
1.02×10<sup>-7</sup>



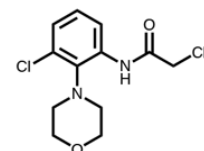
PCM-0102532

68%  
1.14×10<sup>-7</sup>



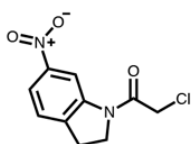
PCM-0102432

65%  
8.46×10<sup>-8</sup>



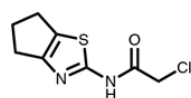
PCM-0102893

62%  
5.64×10<sup>-8</sup>



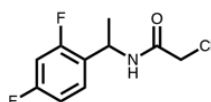
PCM-0102777

Labeling 61%  
k [M<sup>-1</sup>Sec<sup>-1</sup>] 1.00×10<sup>-7</sup>



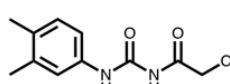
PCM-0102457

59%  
2.14×10<sup>-7</sup>



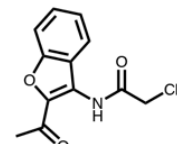
PCM-0103035

58%  
1.27×10<sup>-7</sup>



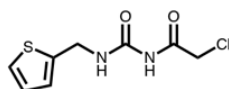
PCM-0102119

57%  
1.09×10<sup>-7</sup>



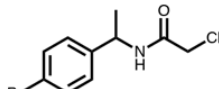
PCM-0102918

56%  
1.32×10<sup>-7</sup>



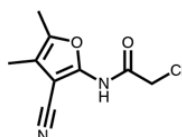
PCM-0102753

Labeling 55%  
k [M<sup>-1</sup>Sec<sup>-1</sup>] 1.14×10<sup>-7</sup>



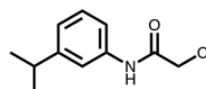
PCM-0102374

55%  
8.13×10<sup>-8</sup>



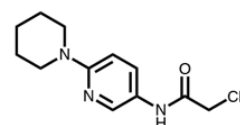
PCM-0102337

53%  
1.19×10<sup>-7</sup>



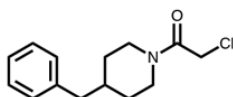
PCM-0102429

52%  
1.13×10<sup>-7</sup>



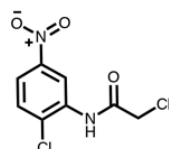
PCM-0102812

53%  
1.38×10<sup>-7</sup>



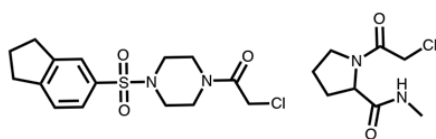
PCM-0102492

Labeling 51%  
k [M<sup>-1</sup>Sec<sup>-1</sup>] 1.04×10<sup>-7</sup>

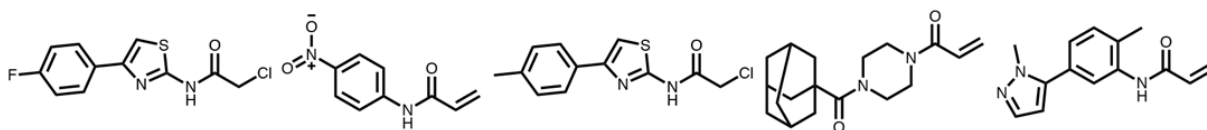


PCM-0102366

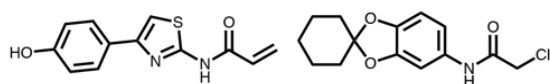
51%  
1.26×10<sup>-7</sup>

**PBP3<sup>R504C</sup>**

	PCM-0102193	PCM-0102518
Labeling	100%	78%
k [M <sup>-1</sup> Sec <sup>-1</sup> ]	2.58×10 <sup>-7</sup>	1.01×10 <sup>-7</sup>

**USP8**

	PCM-0102809	PCM-0102740	PCM-0102168	PCM-0102682	PCM-0102543
Labeling	100%	69%	58%	56%	54%
k [M <sup>-1</sup> Sec <sup>-1</sup> ]	1.75×10 <sup>-7</sup>	1.45×10 <sup>-8</sup>	1.43×10 <sup>-7</sup>	1.24×10 <sup>-8</sup>	5.90×10 <sup>-9</sup>



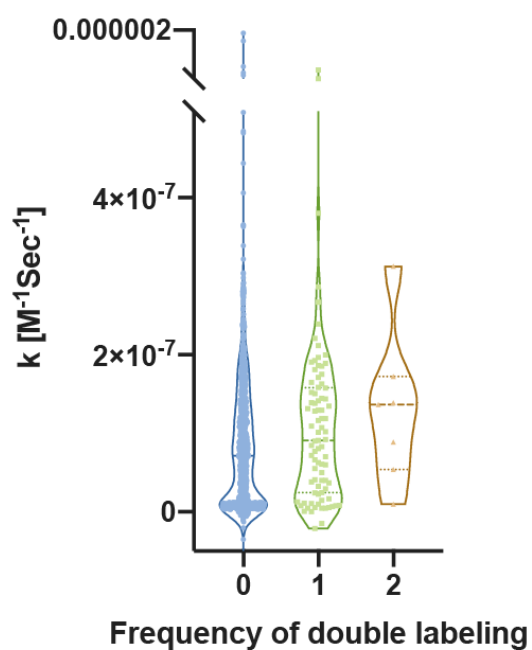
	PCM-0102884	PCM-0102104
Labeling	52%	50%
k [M <sup>-1</sup> Sec <sup>-1</sup> ]	4.09×10 <sup>-8</sup>	1.58×10 <sup>-7</sup>

**Supplementary Figure 7. Chemical structures of non-promiscuous hits from primary screens against NV3CP, K-Ras<sup>G12C</sup>, NNMT, PBP3<sup>R504C</sup> and USP8.**

Indicated labeling refers to the primary screen (24 h, 4 °C, 200 μM compound). The thiol reactivity rate constant for each compound is also indicated.

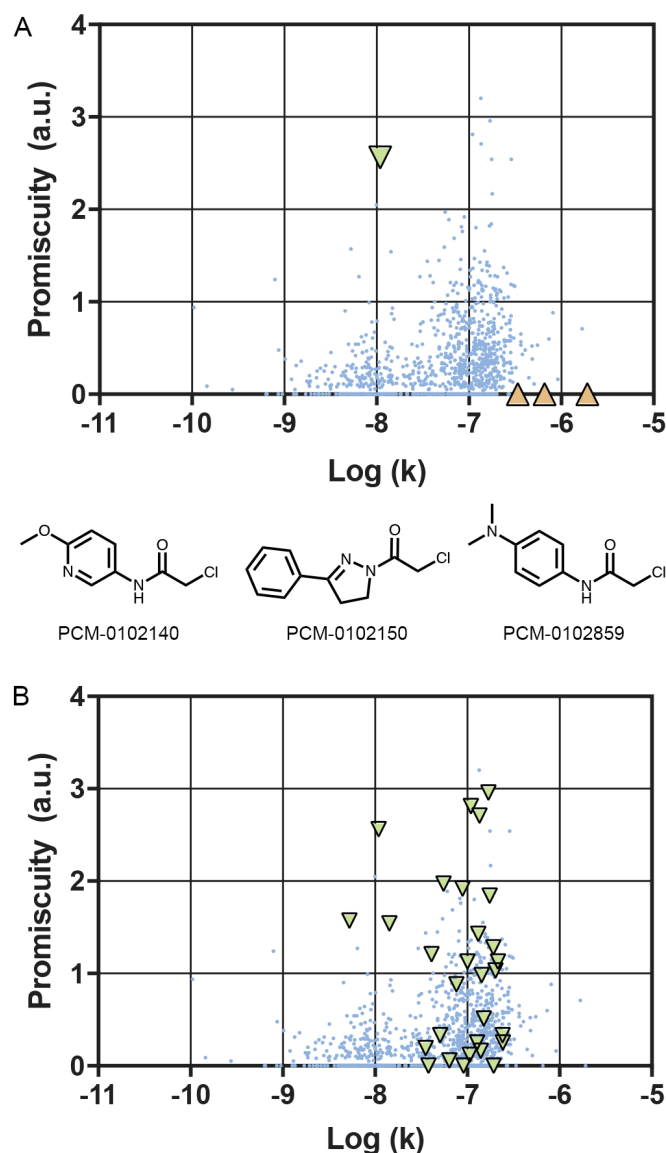
PCM-0102330	PCM-0103067	PCM-0102382	PCM-0102158	PCM-0102296
k [M <sup>-1</sup> Sec <sup>-1</sup> ] 1.69×10 <sup>-7</sup>	1.35×10 <sup>-7</sup>	1.35×10 <sup>-7</sup>	9.72×10 <sup>-8</sup>	1.75×10 <sup>-7</sup>
PCM-0102339	PCM-0102784	PCM-0102260	PCM-0102496	PCM-0102262
k [M <sup>-1</sup> Sec <sup>-1</sup> ] 1.79×10 <sup>-7</sup>	1.09×10 <sup>-7</sup>	-2.11×10 <sup>-8</sup>	1.10×10 <sup>-8</sup>	1.57×10 <sup>-7</sup>
PCM-0102957	PCM-0102401	PCM-0102196	PCM-0102126	PCM-0102746
k [M <sup>-1</sup> Sec <sup>-1</sup> ] 2.87×10 <sup>-7</sup>	1.47×10 <sup>-7</sup>	5.38×10 <sup>-8</sup>	8.90×10 <sup>-8</sup>	6.90×10 <sup>-8</sup>
PCM-0102577	PCM-0102246	PCM-0102539	PCM-0102455	PCM-0102883
k [M <sup>-1</sup> Sec <sup>-1</sup> ] 1.19×10 <sup>-7</sup>	5.53×10 <sup>-8</sup>	2.36×10 <sup>-7</sup>	9.94×10 <sup>-9</sup>	6.06×10 <sup>-8</sup>
PCM-0102244	PCM-0102799	PCM-0102432	PCM-0102695	PCM-0102374
k [M <sup>-1</sup> Sec <sup>-1</sup> ] 5.25×10 <sup>-9</sup>	8.30×10 <sup>-8</sup>	8.46×10 <sup>-8</sup>	1.43×10 <sup>-8</sup>	8.13×10 <sup>-8</sup>
PCM-0102834	PCM-0102252			
k [M <sup>-1</sup> Sec <sup>-1</sup> ] 3.52×10 <sup>-8</sup>	1.67×10 <sup>-7</sup>			

**Supplementary Figure 8. Chemical structures and reactivity rates of promiscuous fragments.**



**Supplementary Figure 9. Distribution of rate constants for compounds with no occurrences of double labeling or with one or two occurrences of double labeling.** 3 compounds that showed double labeling in 3 or 4 occurrences are not shown.



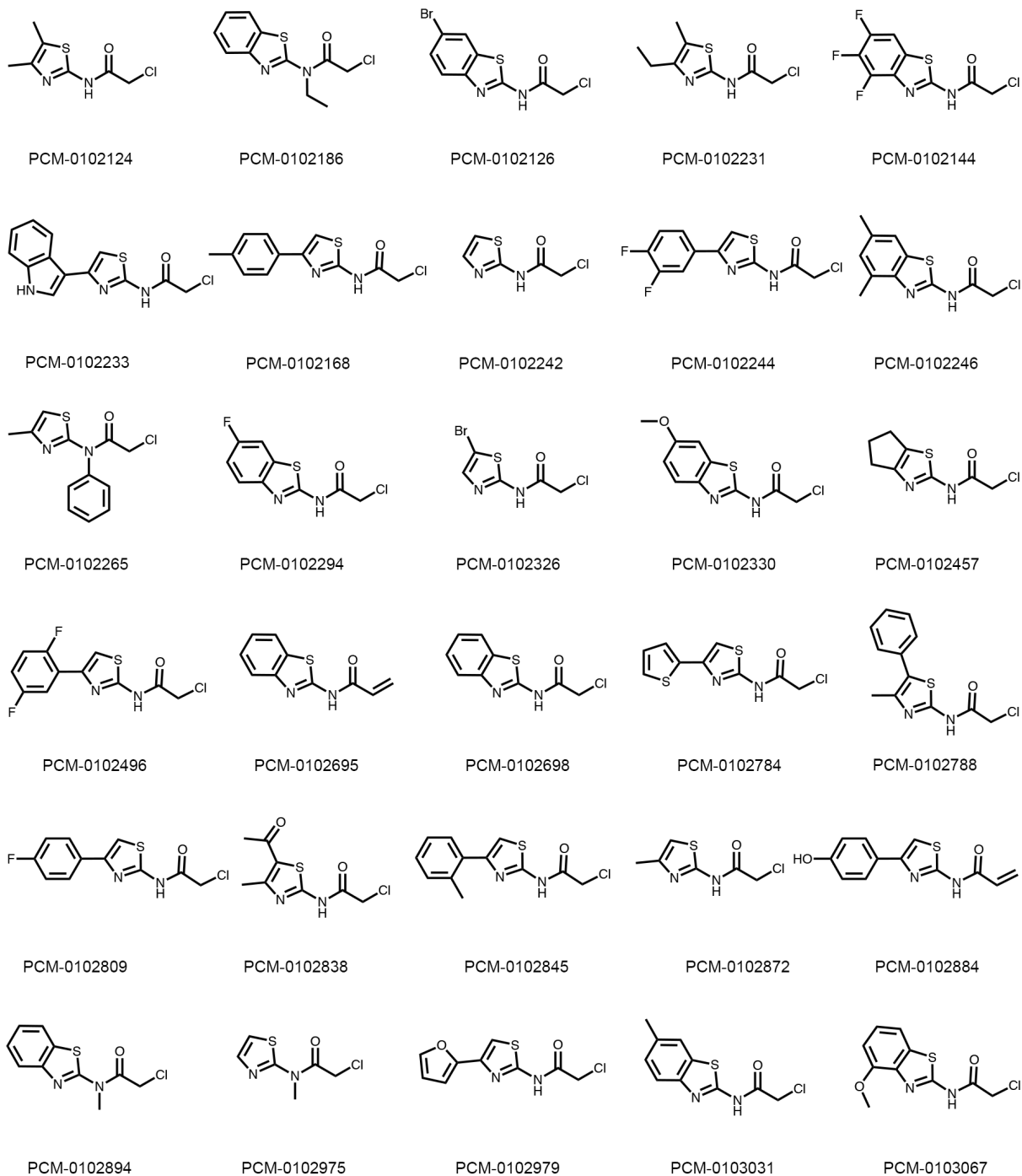


**Supplementary Figure 10. No clear correlation between compound thiol reactivity and overall protein labeling.**

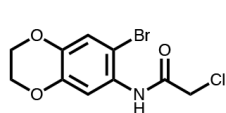
**A.** Promiscuity of compounds is calculated as the total labeling percentage across all ten screened proteins. Reactivity is presented as the log of the rate constant. PCM-0102496 (green triangle) is a compound which is promiscuous but displays very low reactivity. PCM-0102140, PCM-0102859, PCM-0102150 (orange triangles; structures illustrated) are on the contrary, highly reactive but do not label any protein. We should note though that for at least two or three proteins these compounds are found in wells for which we were not able to interpret labeling data.

**B.** The aminothiazole substructure (denoted in green triangles; see Supp. Fig. 12 for structures) confers promiscuity, however, not all aminothiazoles are highly reactive, nor are they all promiscuous.

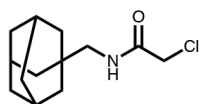




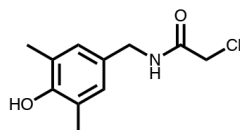
**Supplementary Figure 12. Chemical structures of aminothiazole containing fragments.**



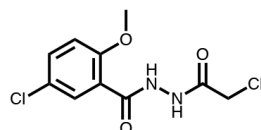
PCM-0102299



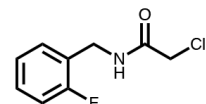
PCM-0102706



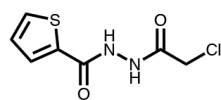
PCM-0102431



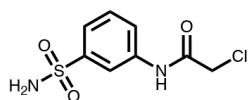
PCM-0102763



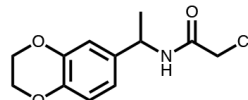
PCM-0102149



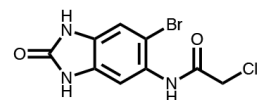
PCM-0102426



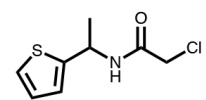
PCM-0102162



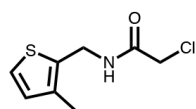
PCM-0102342



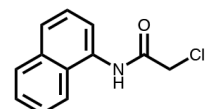
PCM-0102388



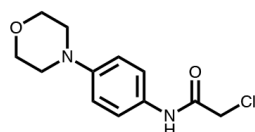
PCM-0102180



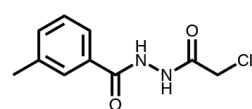
PCM-0103075



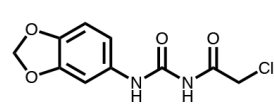
PCM-0102396



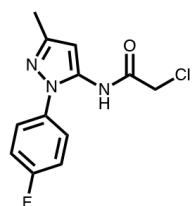
PCM-0102582



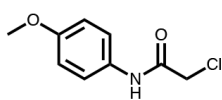
PCM-0102737



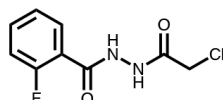
PCM-0102333



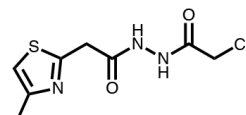
PCM-0102665



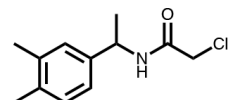
PCM-0102156



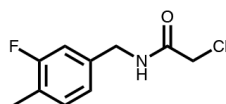
PCM-0102510



PCM-0102095

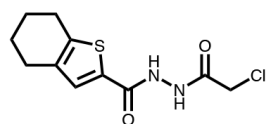


PCM-0102270

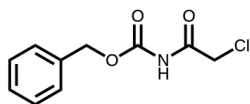


PCM-0102959

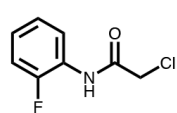
### Fragments for follow-up



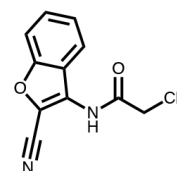
PCM-0102142



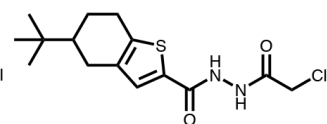
PCM-0102153



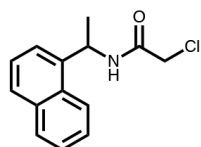
PCM-0102158



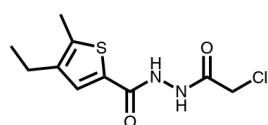
PCM-0102252



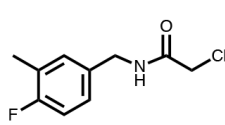
PCM-0102260



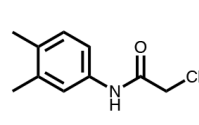
PCM-0102275



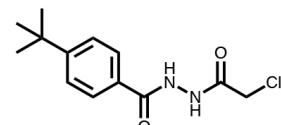
PCM-0102300



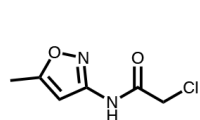
PCM-0102305



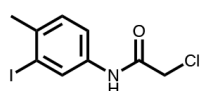
PCM-0102339



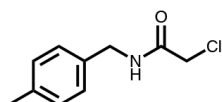
PCM-0102355



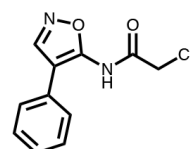
PCM-0102500



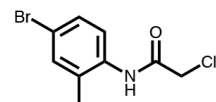
PCM-0102577



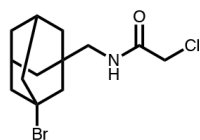
PCM-0102660



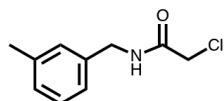
PCM-0102746



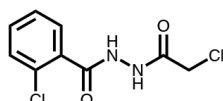
PCM-0102799



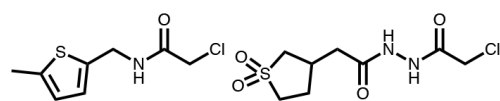
PCM-0102819



PCM-0102821

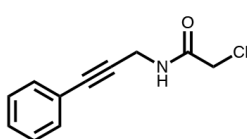


PCM-0102954

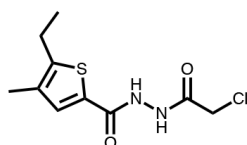


PCM-0102973

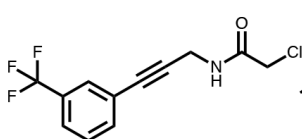
PCM-0102998



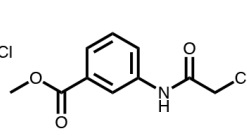
PCM-0103007



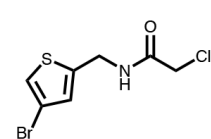
PCM-0103009



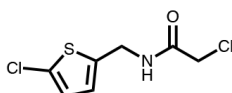
PCM-0103011



PCM-0103048

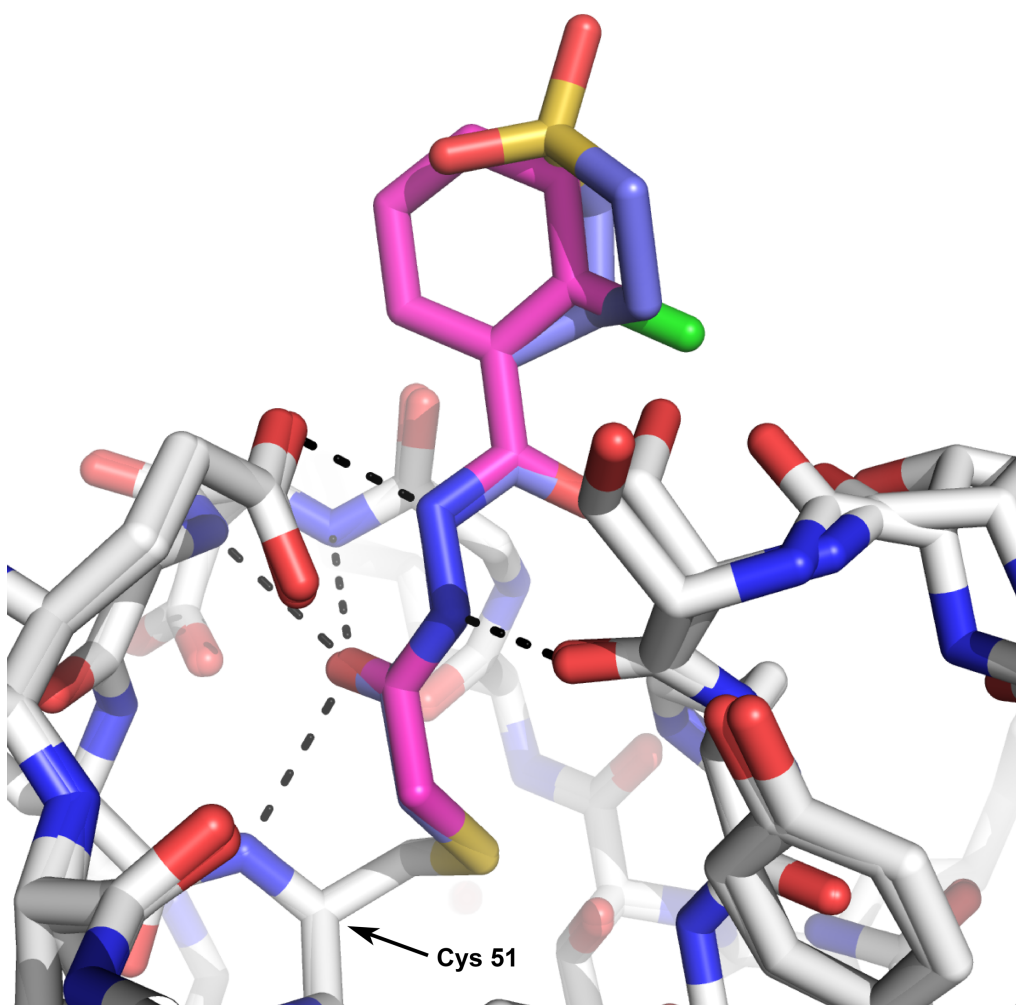


PCM-0103050



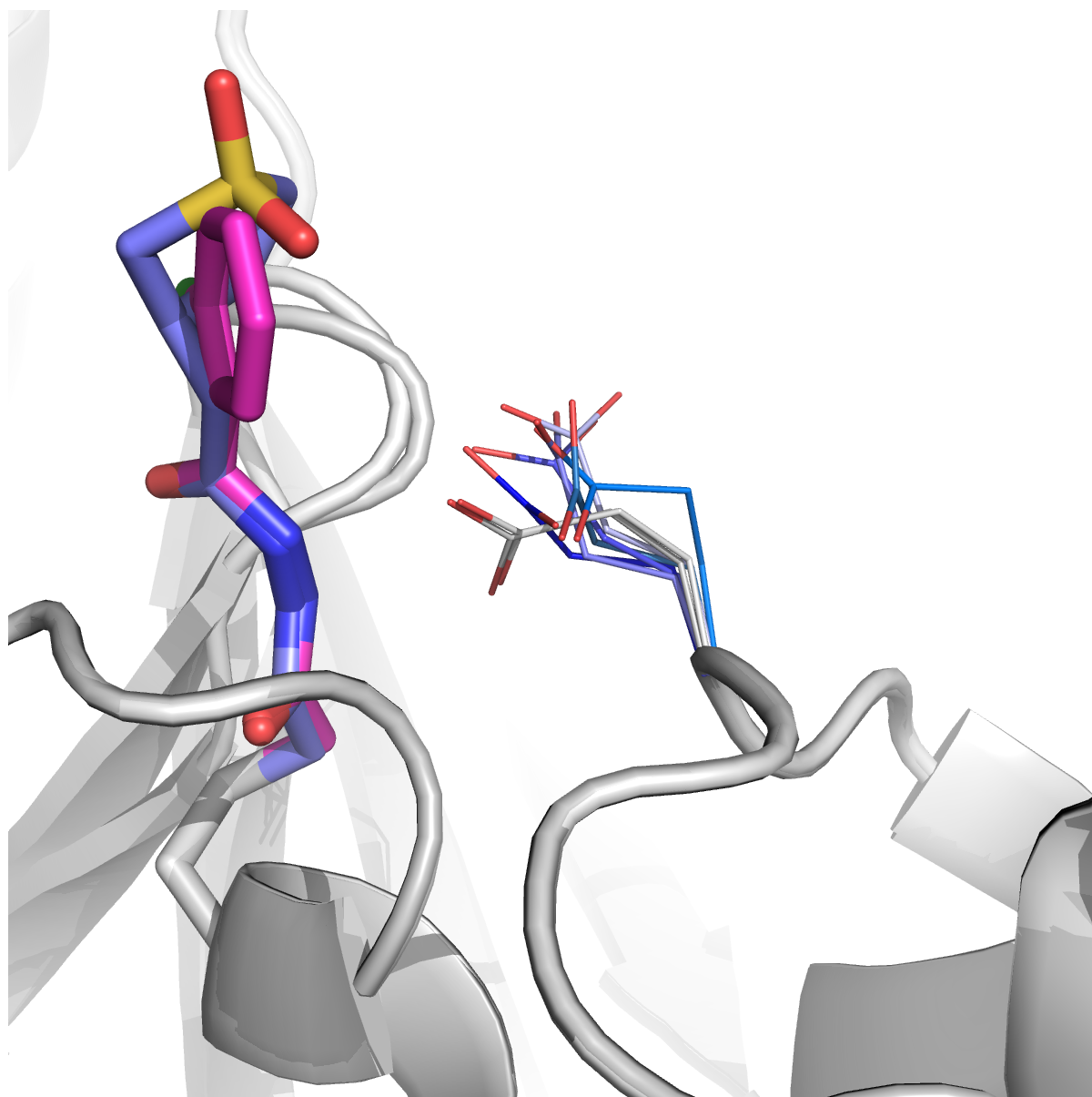
PCM-0103080

**Supplementary Figure 13. Chemical structures of OTUB2 hits**  
Fragments selected for follow-up analysis are indicated.



**Supplementary Figure 14. Identified recognition motif mediates a hydrogen bond network with OTUB2.**

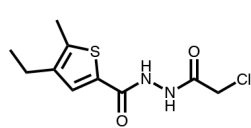
Overlay of the co-crystal structures of OTUB2 with compounds PCM-0102954 (magenta; PDB 5QIY) and PCM-0102998 (pink; PDB 5QIV) respectively. The hydrazide motif that re-occurred in seven of the OTUB2 selective hits mediates three hydrogen bonds with the backbone of residues Ser223, Asp48, Gly49 and Cys51 itself (to which it is also covalently bound) and another hydrogen bond with the sidechain of Glu174.



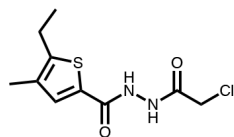
**Supplementary Figure 15. A rotamer flip mediates a new hydrogen bond with hydrazide motif.**

Overlay of 11 co-crystal structures of OTUB2 in complex with various electrophilic fragments (See Fig. 4A for the various fragments). The rotamer of glutamate 174 is shown in white and gray for the two fragments containing the hydrazide motif (see Supp. Fig. 7) and in blue for all other fragments. It clearly shows how this glutamate adopts a different rotamer in order to form a new hydrogen bond with the hydrazide.

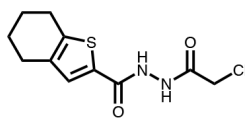
### First generation fragments



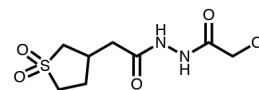
PCM-0102300



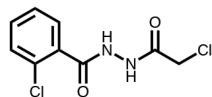
PCM-0103009



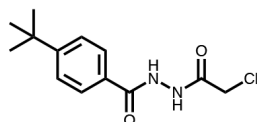
PCM-0102142



PCM-0102998

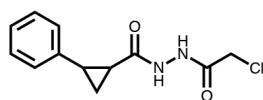


PCM-0102954

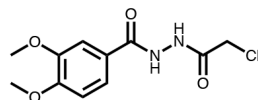


PCM-0102355

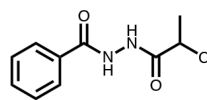
### Second generation fragments



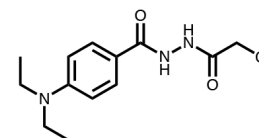
OTUB2-COV-1



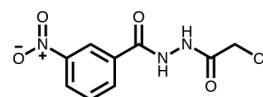
OTUB2-COV-2



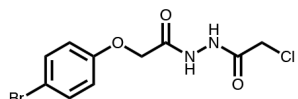
OTUB2-COV-3



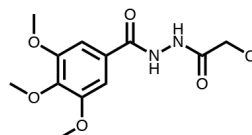
OTUB2-COV-4



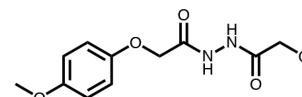
OTUB2-COV-5



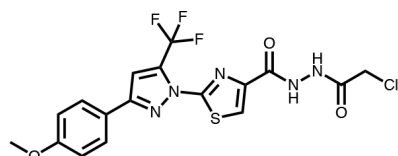
OTUB2-COV-6



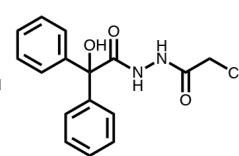
OTUB2-COV-7



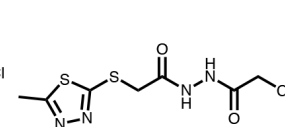
OTUB2-COV-8



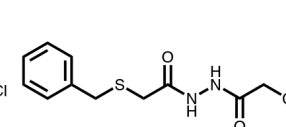
OTUB2-COV-9



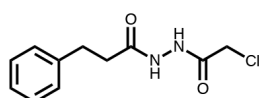
OTUB2-COV-10



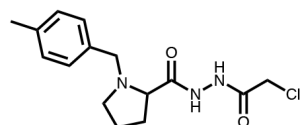
OTUB2-COV-11



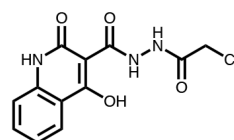
OTUB2-COV-12



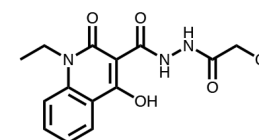
OTUB2-COV-13



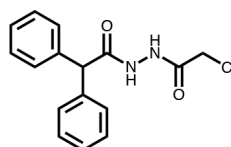
OTUB2-COV-14



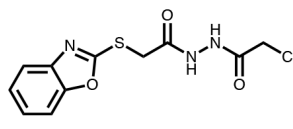
OTUB2-COV-15



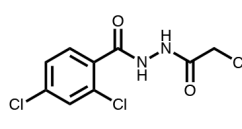
OTUB2-COV-16



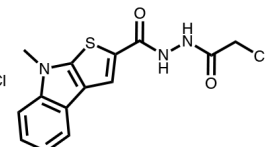
OTUB2-COV-17



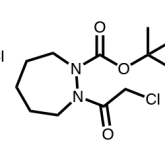
OTUB2-COV-18



OTUB2-COV-19



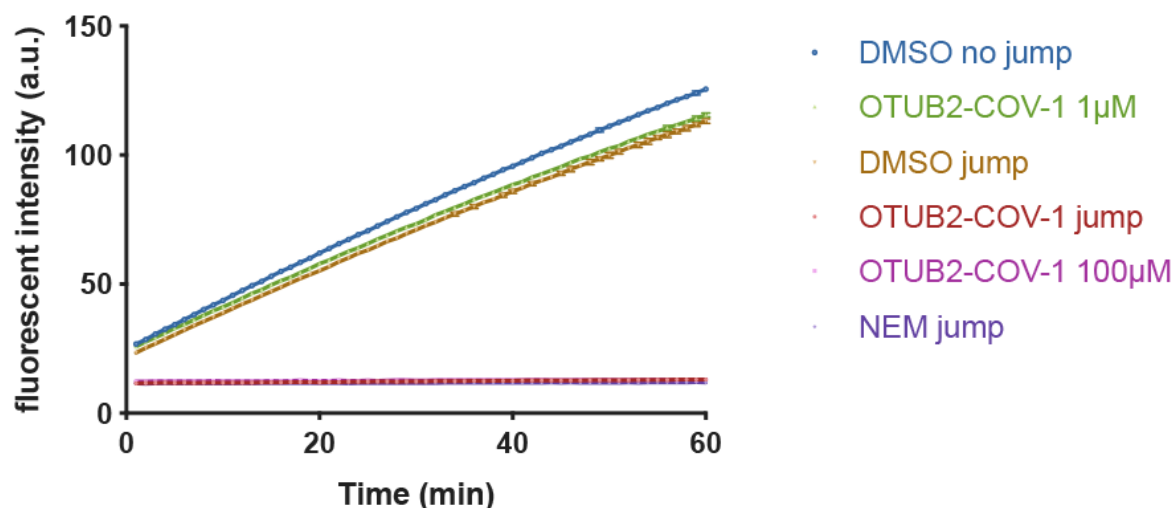
OTUB2-COV-20



OTUB2-COV-21

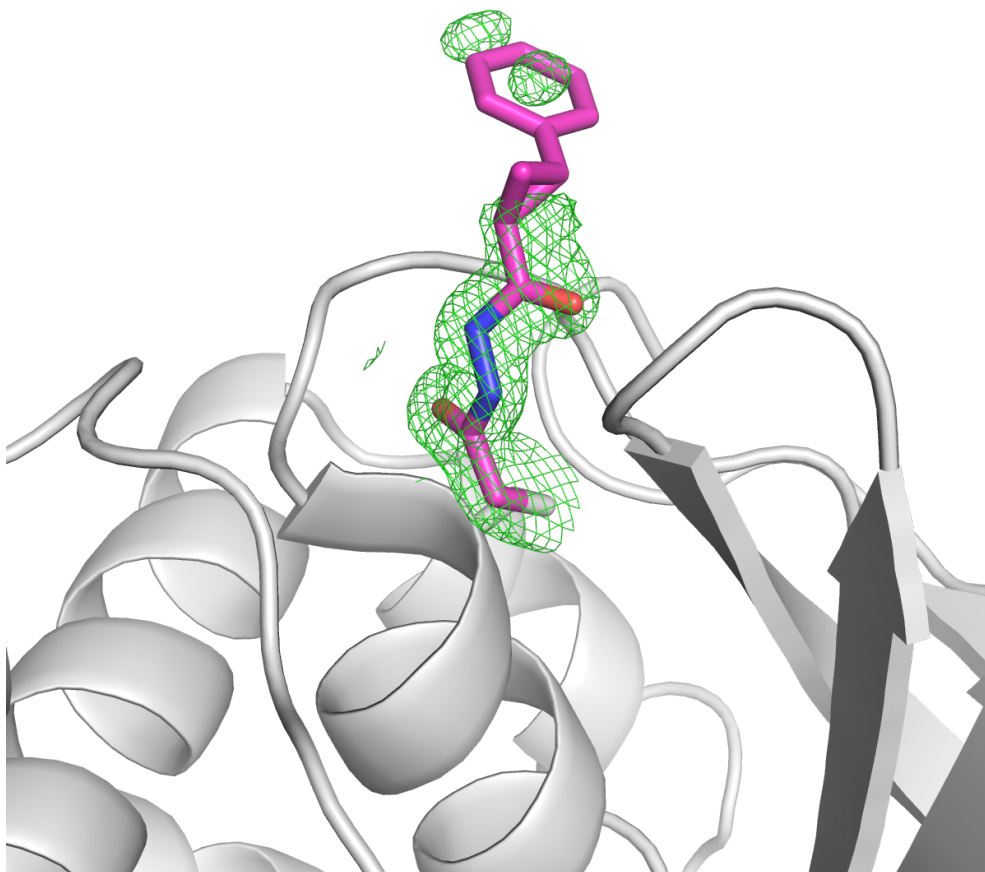
**Supplementary Figure 16. Chemical structures of second generation OTUB2 binders sharing the chloroacetylhydrazide motif.**





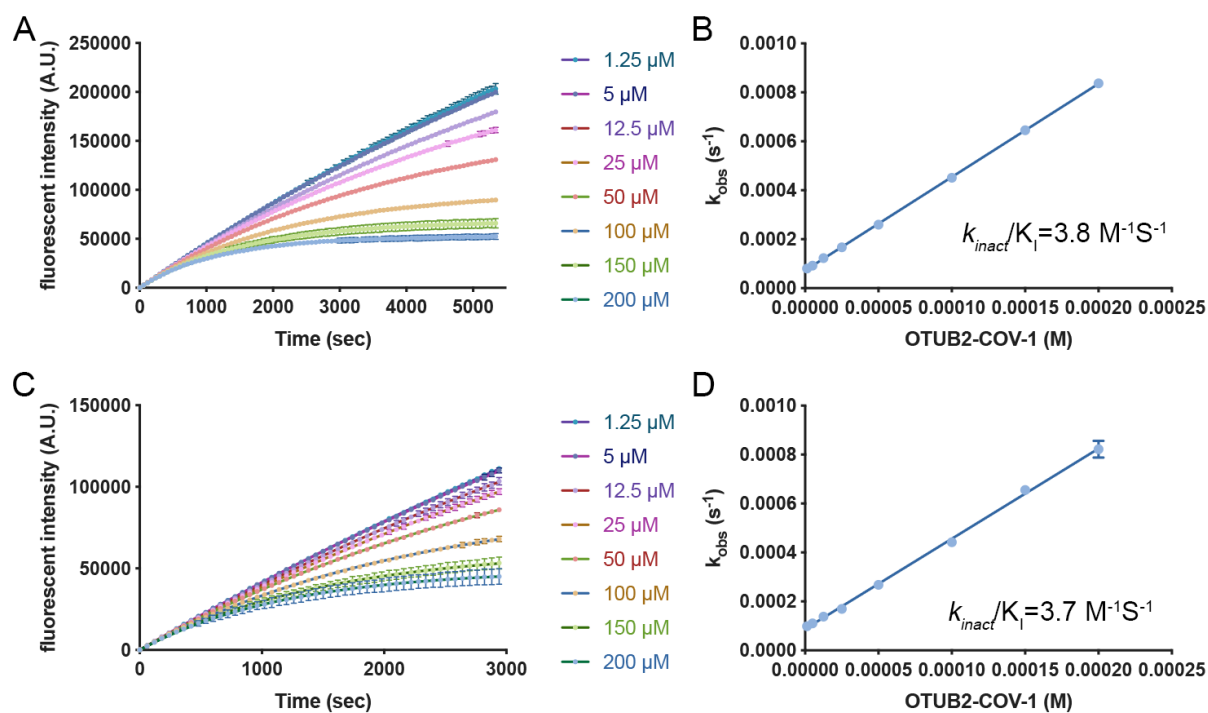
**Supplementary Figure 17. OTUB2 inhibition is irreversible and not affected by dilution of the sample.**

OTUB2 enzymatic rate was measured with DMSO; when diluted 100x (DMSO jump) there was very little effect of the rate. The rate was then measured in the presence of inhibitors. 1  $\mu$ M OTUB2-COV-1 had little inhibition, whereas both OTUB2-COV-1 at 100  $\mu$ M and NEM completely inhibited the enzyme. Diluting OTUB2-COV-1 x100 to 1  $\mu$ M (OTUB2-COV-1 jump), or N-Ethylmaleimide (NEM jump) did not rescue the enzymatic activity - suggesting irreversible inhibition.



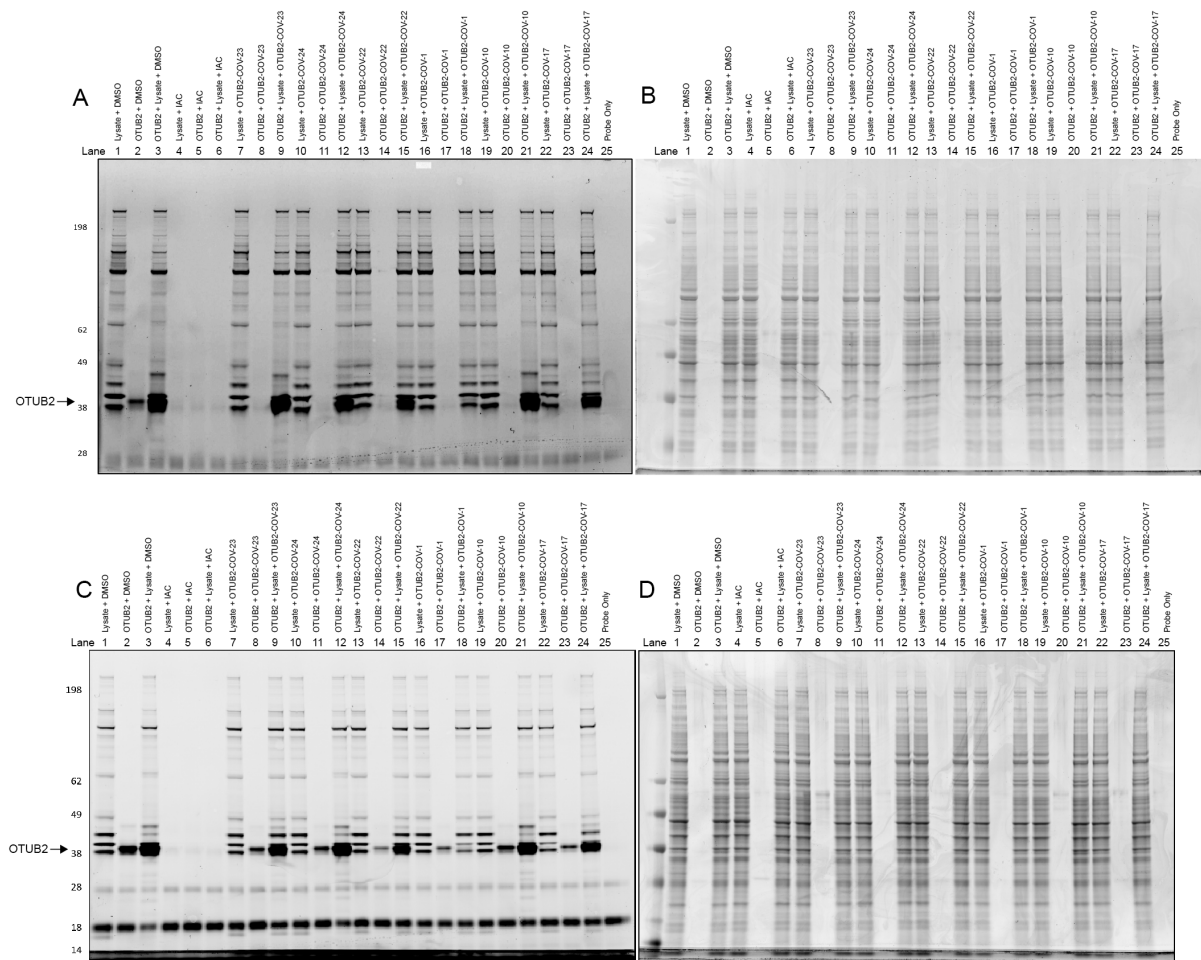
**Supplementary Figure 18. Partial electron density for lead OTUB2 inhibitor.**

Co-crystal structure of OTUB2 (white) in complex with OTUB2-COV-1 (magenta; PDB: 5QIO).  $2mF_o-DF_c$  map is contoured at  $1\sigma$ , showing continuous density from the cysteine thiol to the compound, indicating the formation of the covalent bond. Density is also clear for the hydrazide motif, however, there is barely any electron density for the phenyl-cyclopropyl moiety, suggesting it may be flexible.



### Supplementary Figure 19. Determination of $K_{\text{inact}}/K_i$ for OTUB2-COV-1

**A.** and **C.** Rate measurements of the enzymatic activity in the presence of various OTUB2-COV-1 concentrations to determine the observed first-order rate constant ( $k_{\text{obs}}$ ). **B.** and **D.** Linear fit of  $k_{\text{obs}}$  vs. the concentration of OTUB2-COV-1 to calculate  $k_{\text{inact}}/K_i$  which is the slope of the fit.

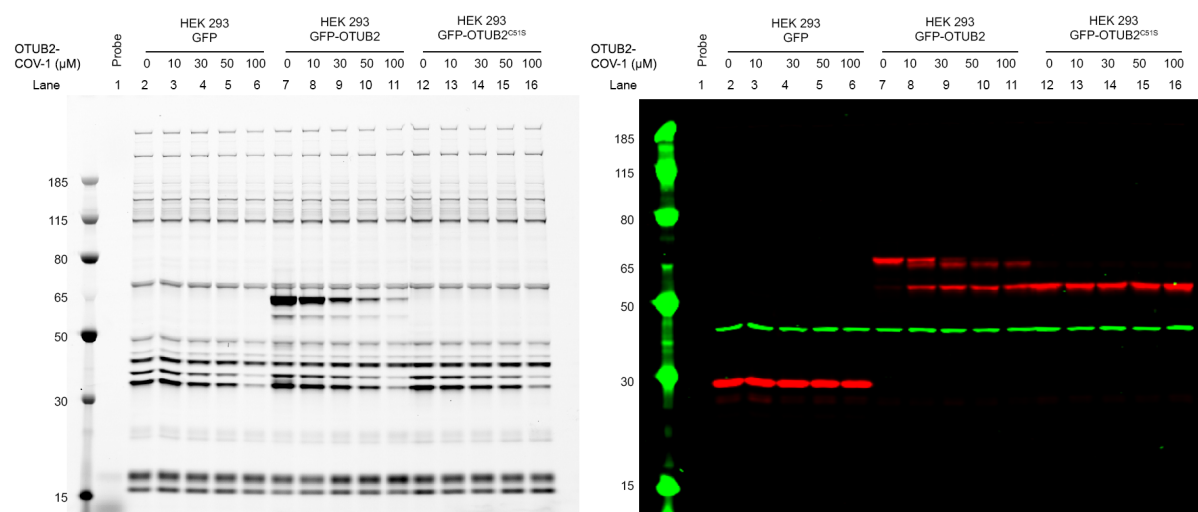


### Supplementary Figure 20. OTUB2 inhibitors are selective in lysates.

Gel-based ABPP selectivity assessment. HEK293T cell lysate with or without addition of purified recombinant OTUB2 was treated with 50  $\mu$ M of the inhibitors. All residual cysteine DUB activity was labelled with the fluorescent Rho-Ub-PRG DUB probe. Inhibition of OTUB2 becomes apparent by disappearance of the corresponding band just above 38 kDa. **A.** Fluorescence scan and **B.** InstantBlue stain of samples spiked with 0.05  $\mu$ g/ $\mu$ L OTUB2. **C.** Fluorescence scan and **D.** InstantBlue stain of samples spiked with 0.1  $\mu$ g/ $\mu$ L OTUB2.

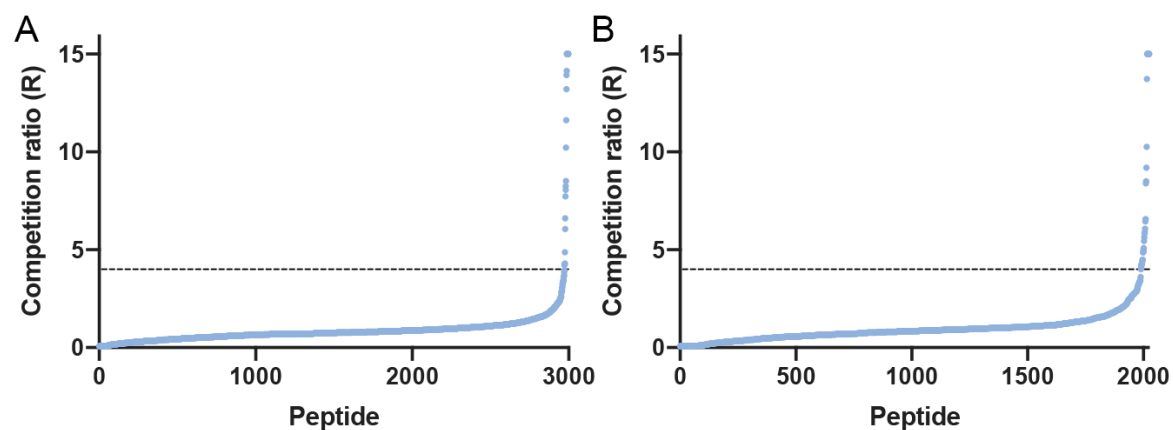
Lanes 1-3: DMSO negative controls showing probe labeling of DUBs in lysate, of purified OTUB2 and lysate spiked with OTUB2. Lanes 4-6: Iodoacetamide (10 mM) as positive control eliminates all probe labeling. Lanes 16-18: OTUB2-COV-1 specifically compete with probe only for OTUB2. Lanes 19-21: Negative control compound OTUB2-COV-10 can only compete with probe for purified OTUB2 but not for OTUB2 in lysate. Note, it does have other DUB off-targets (compare lanes 19 and 21)

When incubated with OTUB2 alone, the fragments outcompeted the DUB probe, completely blocking any labeling at 0.05  $\mu$ g and significantly diminished labeling at 0.1  $\mu$ g. This effect is much more pronounced in spiked lysates. A very pronounced band appears for OTUB2 in the DMSO control as well as the inactive compound control OTUB2-COV-10 (likely due to merging with bands of close molecular-weight DUBs). However, for the most potent compound OTUB2-COV-1, this band completely disappears. For the other analogs, while all diminish the probe labeling against recombinant OTUB2, none are able to compete with it as well in the spiked lysate.



**Supplementary Figure 21. Gel-based ABPP shows selectivity using a fluorescent activity-based DUB probe in HEK293 cells overexpressing OTUB2-GFP.**

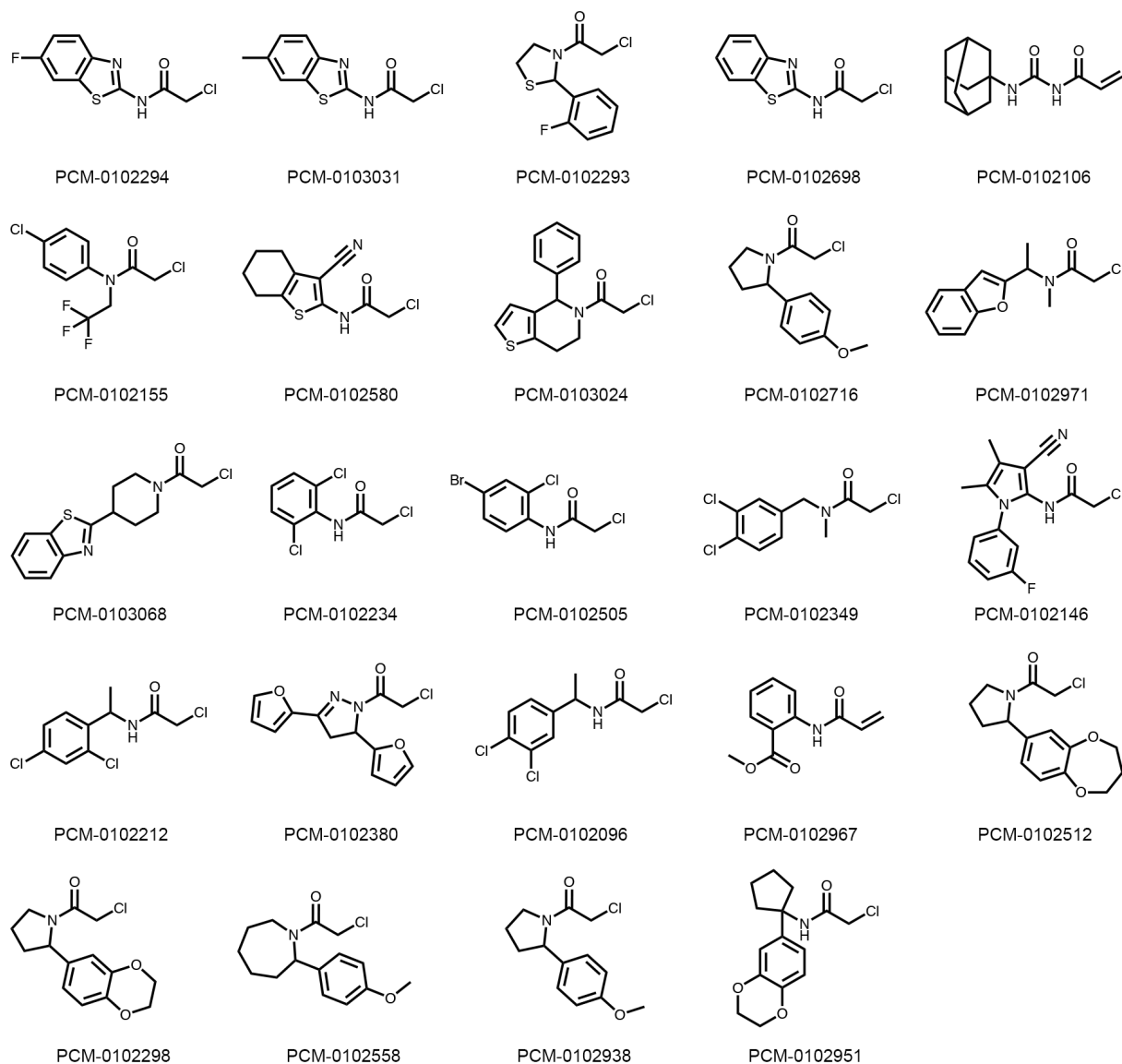
HEK293 transfected with GFP (lanes 2-6), GFP-OTUB2 (lanes 7-11) or GFP-OTUB2-C51S (lanes 12-16). **A.** Fluorescence scan of DUB probe. Each band corresponds to a DUB. The absence of OTUB2 bands (~65kDa and ~55kDa) at lanes 2-7 indicates the low endogenous expression levels of OTUB2. The band at ~55kDa likely corresponds to an OTUB2 degradation product since it is not detected in both the endogenous cells (lanes 2-7) or in cells overexpressing the OTUB2-C51S mutant - not able to be labeled by the DUB probe. **B.** Western blot: anti-GFP in red, anti- $\beta$ -actin in green. The upper red band is the OTUB2-GFP bound by the DUB probe and is diminishing in a dose-dependent manner with increasing OTUB2-COV-1. The lower red band is OTUB2-GFP without the DUB probe and is only found in cells overexpressing OTUB2-GFP or the catalytically dead C51S OTUB2-GFP.



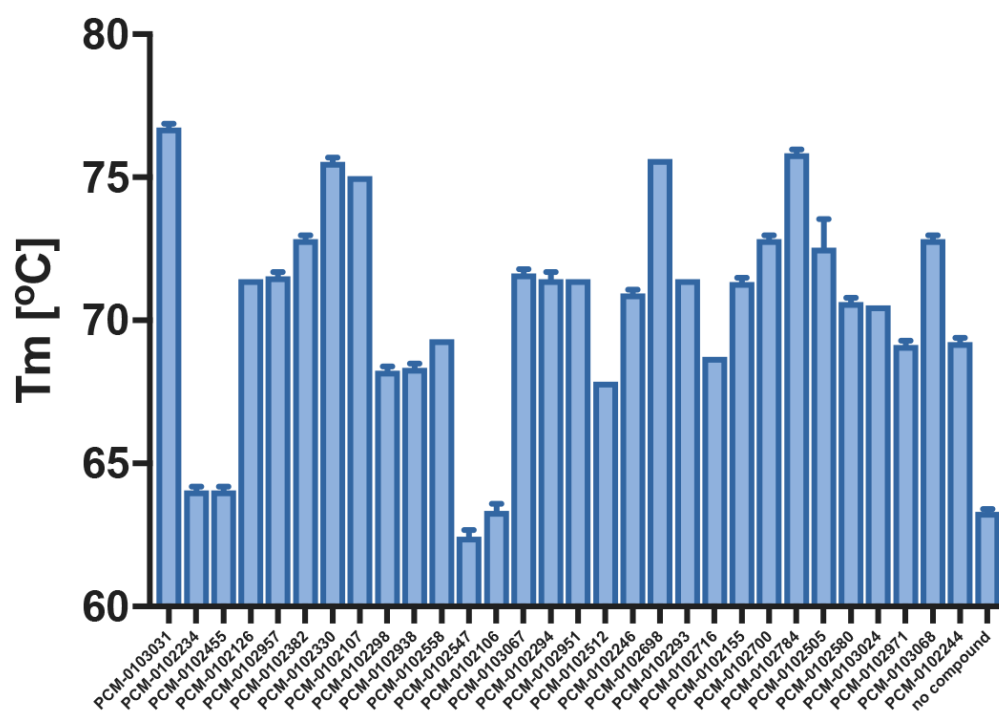
**Supplementary Figure 22. Competitive isoTOP-ABPP in HEK293 cells.**

Heavy/Light ratio of peptides detected after in a competitive isoTOP-ABPP experiment in cells.

**A.** Incubation with OTUB2-COV-1 **B.** Incubation with NUDT7-COV-1. Black line represents Heavy/Light ratio = 4. The identity of the peptides can be found in Supplementary Datasets 6 & 7.



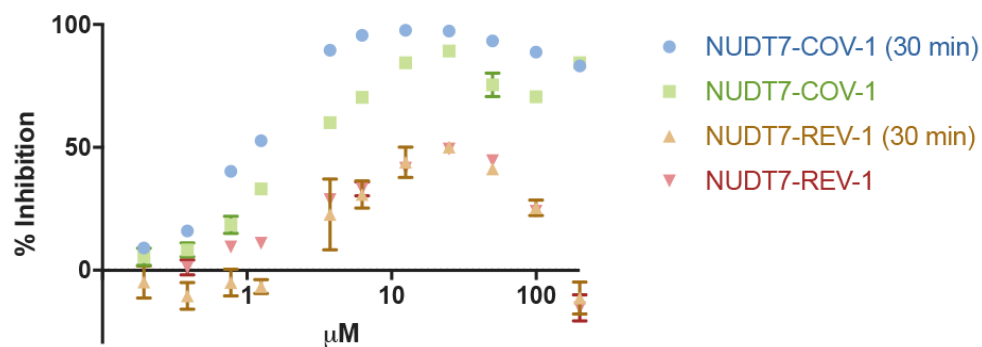
**Supplementary Figure 23. Chemical structures of non-promiscuous NUDT7 hits.**



#### Supplementary Figure 24. Covalent fragments stabilize NUDT7.

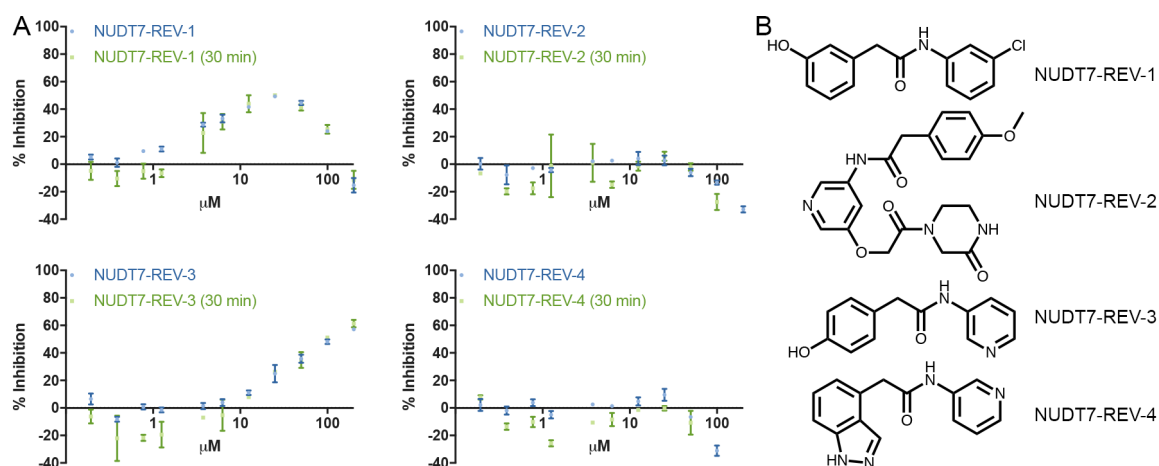
Hits against NUDT7 (see Supp. Fig. 10) were evaluated for thermal stabilization of NUDT7 by Differential Scanning Fluorimetry (DSF). Thermal denaturation curves were measured for the untreated protein as well as for protein incubated with 200  $\mu$ M of the indicated compounds for 24 hours; 4 °C. 26 of the compounds showed substantial stabilization. Experiments were performed in triplicates and are represented as mean  $\pm$  standard deviation.





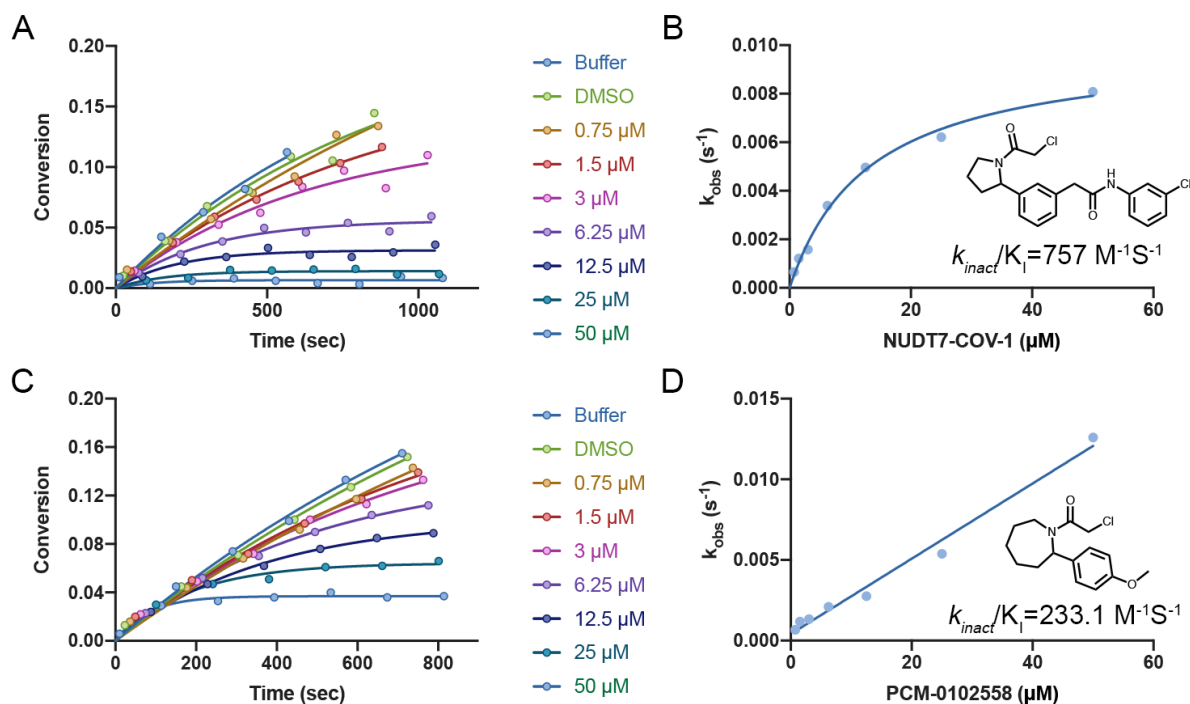
**Supplementary Figure 25. Optimized compound potently inhibits NUDT7 in an enzymatic assay.**

Full dose response curve for enzymatic inhibition of NUDT7 by the covalent optimized compound NUDT7-COV-1 and the parent non-covalent compound NUDT7-REV-1, with either no pre-incubation or 30 minutes pre-incubation.



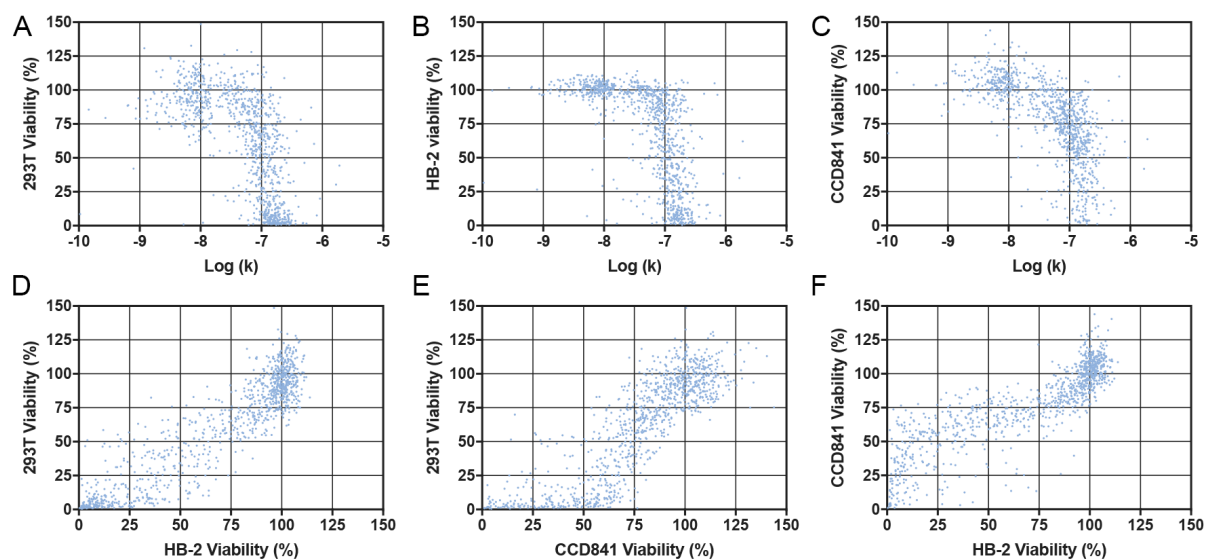
**Supplementary Figure 26. Reversible NUDT7 binders do not show enzyme inhibition.**

**A.** Dose response curves for enzymatic inhibition of NUDT7 by non-covalent compounds, with either no pre-incubation or 30 minutes pre-incubation. **B.** Structures of non-covalent compounds.



**Supplementary Figure 27. Determination of  $k_{\text{inact}}/K_i$  of primary hit and optimized probe against NUDT7.**

**A.** Enzymatic activity of NUDT7 as a function of increasing concentrations of NUDT7-COV-1 to determine the observed first-order rate constant ( $k_{\text{obs}}$ ). **B.** Non-linear fit of  $k_{\text{obs}}$  vs. NUDT7-COV-1 concentration using  $Y = k_{\text{inact}} \cdot X / (K_i + X)$  results in a goodness of fit  $R^2 = 0.99$  with  $k_{\text{inact}} = 0.01 \text{ s}^{-1}$ ,  $K_i = 13.21 \mu\text{M}$  and  $k_{\text{inact}}/K_i = 757 \text{ M}^{-1}\text{s}^{-1}$ . **C.** Enzymatic activity of NUDT7 as a function of increasing concentrations of PCM-0102558 to determine the observed first-order rate constant ( $k_{\text{obs}}$ ). **D.** The data could not be fit to a Michaelis-Menten curve; thus we used a linear fit to calculate  $k_{\text{inact}}/K_i = 233.1 \text{ M}^{-1}\text{s}^{-1}$  ( $R^2 = 0.984$ ).



**Supplementary Figure 28. Correlation between cellular toxicity and thiol reactivity at high compound concentrations.**

**A.** % viability of HEK293 cells after 48 hours incubation with the electrophile library at a concentration of 10  $\mu$ M as a function of reactivity (second order kinetic rate).

**B.** % viability of HB-2 cells after 48 hours incubation with the electrophile library at a concentration of 10  $\mu$ M as a function of reactivity (second order kinetic rate).

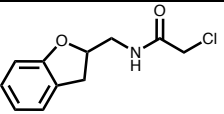
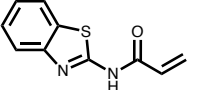
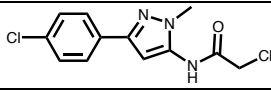
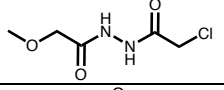
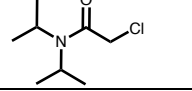
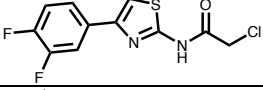
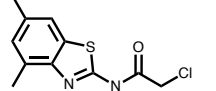
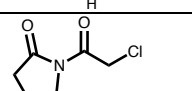
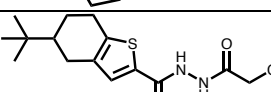
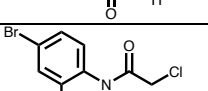
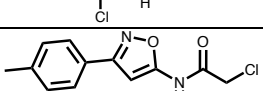
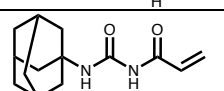
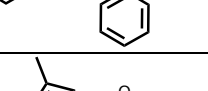
**C.** % viability of CCD841 cells after 48 hours incubation with the electrophile library at a concentration of 10  $\mu$ M as a function of reactivity (second order kinetic rate).

**D.** % viability correlation between HEK293 and HB-2 cells.

**E.** % viability correlation between HEK293 and CCD841 cells.

**F.** % viability correlation between CCD841 and HB-2 cells.

# Supplementary Tables

compound	Avg. k (M <sup>-1</sup> S <sup>-1</sup> )	St. Dev.	re-measured k (M <sup>-1</sup> S <sup>-1</sup> )	St. Dev.	GSH t <sub>1/2</sub> (hours)	structure
PCM-0103020	3.85E-08	1.10E-08	6.01E-08	6.54E-09	>>6	
PCM-0102695	1.43E-08	3.68E-08	2.11E-08	5.41E-09	0.3	
PCM-0102296	1.75E-07	7.41E-08	1.27E-07	8.19E-09	1.6	
PCM-0103070	2.57E-08	2.49E-09	3.68E-08	3.82E-09		
PCM-0103023	3.55E-08	1.70E-08	1.57E-07	2.37E-08	3.2	
PCM-0102244	5.25E-09	2.69E-08	6.88E-08	5.05E-08	2.6	
PCM-0102246	5.53E-08	2.39E-08	6.53E-08	8.55E-09	2.0	
PCM-0102982	6.43E-09	6.14E-09	2.29E-07	1.15E-08		
PCM-0102260	-2.11E-08	2.79E-08	-5.53E-09	1.21E-08	>>6	
PCM-0102505	6.69E-08	6.77E-09	5.90E-08	2.32E-08	2.8	
PCM-0102382	1.35E-07	6.10E-08	2.32E-07	4.95E-08	1.1	
PCM-0102106	7.90E-10	1.55E-08	4.84E-09	3.32E-09		
PCM-0102957	2.87E-07	1.04E-07	3.75E-07	1.34E-08	0.6	
PCM-0102665	1.24E-07	1.95E-08	2.15E-07	1.57E-08	1.8	

Supplementary Table 1. Reactivity comparison for 14 re-sourced compounds.

Compound	200 $\mu$ M <sup>a</sup>	100 $\mu$ M
PCM-0102142	90%	80%
PCM-0102153	75%	N/A <sup>b</sup>
PCM-0102158	13%	8%
PCM-0102252	90%	76%
PCM-0102260	7%	18%
PCM-0102275	97%	N/A
PCM-0102300	97%	84%
PCM-0102305	71%	48%
PCM-0102339	58%	35%
PCM-0102355	28%	16%
PCM-0102500	30%	16%
PCM-0102577	72%	53%
PCM-0102660	57%	30%
PCM-0102746	85%	70%
PCM-0102799	73%	49%
PCM-0102819	51%	29%
PCM-0102821	57%	29%
PCM-0102954	70%	54%
PCM-0102973	54%	31%
PCM-0102998	76%	55%
PCM-0103007	67%	45%
PCM-0103009	93%	78%
PCM-0103011	97%	84%
PCM-0103048	69%	37%
PCM-0103050	43%	22%
PCM-0103080	52%	30%

**Supplementary Table 2. Labeling of OTUB2 with selected hits**

<sup>a</sup> Compounds were incubated at the indicated concentration for 24 hours with 2  $\mu$ M protein in 4 °C. % labeling was assigned via intact protein LC/MS. See Supp. Fig. 6 for chemical structures of compounds.

<sup>b</sup> N/A - % labeling could not be assigned due to noisy spectrum.

PDB ID	5QIO	5QIP	5QIQ	5QIR	5QIS	5QIT	5QIU	5QIV	5QIW	5QIX	5QIY	5QIZ
Compound	OTUB-COV-1	PCM-0102153	PCM-0103050	PCM-0102305	PCM-0102500	PCM-0102821	PCM-0103011	PCM-0102998	PCM-0102660	PCM-0103007	PCM-0102954	PCM-0103080
Beamline	DLS I04-1	DLS I04-1	DLS I04-1	DLS I04-1	DLS I04-1	DLS I04-1	DLS I04-1	DLS I04-1	DLS I04-1	DLS I04-1	DLS I04-1	DLS I04-1
Wavelength (Å)	0.91587	0.92819	0.92819	0.92819	0.92819	0.92819	0.92819	0.92819	0.92819	0.92819	0.92819	0.92819
Space group	P 2 <sub>1</sub>	P 2 <sub>1</sub>	P 2 <sub>1</sub>	P 2 <sub>1</sub>	P 2 <sub>1</sub>	P 2 <sub>1</sub>	P 2 <sub>1</sub>	P 2 <sub>1</sub>	P 2 <sub>1</sub>	P 2 <sub>1</sub>	P 2 <sub>1</sub>	P 2 <sub>1</sub>
Cell dimensions:												
a, b, c (Å)	46.78 58.15 49.30	47.58 57.98 50.30	47.61 58.79 50.20	47.30 58.46 49.79	47.42 58.75 50.10	47.29 58.44 49.71	47.44 58.43 50.14	47.24 58.64 49.79	47.37 58.59 49.86	47.55 58.33 50.33	47.09 58.17 49.45	47.21 58.41 49.58
α, β, γ (deg)	90.00 116.25 90.00	90.00 116.68 90.00	90.00 116.15 90.00	90.00 115.94 90.00	90.00 116.36 90.00	90.00 115.88 90.00	90.00 116.28 90.00	90.00 116.18 90.00	90.00 116.07 90.00	90.00 116.38 90.00	90.00 115.90 90.00	90.00 115.85 90.00
Resolution range	29.07 - 1.46 (1.51 - 1.46)	22.00 - 1.63 (1.69 - 1.63)	45.06 - 1.44 (1.49 - 1.44)	29.23 - 1.43 (1.48 - 1.43)	29.37 - 1.53 (1.59 - 1.53)	29.22 - 1.46 (1.51 - 1.46)	29.21 - 1.56 (1.62 - 1.56)	29.32 - 1.39 (1.44 - 1.39)	23.87 - 1.71 (1.77 - 1.71)	29.17 - 1.39 (1.44 - 1.39)	29.08 - 1.58 (1.64 - 1.58)	44.62 - 1.63 (1.69 - 1.63)
Unique reflections	40592 (3873)	30418 (3025)	44937 (4478)	44585 (4383)	37219 (3695)	41298 (4036)	34928 (3475)	47746 (4613)	25921 (2573)	48711 (4750)	32418 (3189)	30234 (2993)
Multiplicity	3.0 (2.1)	3.3 (3.4)	3.3 (2.9)	3.3 (2.9)	3.4 (3.4)	3.4 (3.2)	3.4 (3.4)	3.4 (2.8)	3.4 (3.2)	3.3 (2.8)	3.4 (3.5)	3.3 (3.4)
Completeness (%)	98.43 (94.62)	99.34 (99.51)	99.75 (99.78)	98.76 (97.42)	99.84 (99.92)	97.41 (95.68)	99.76 (99.48)	97.10 (94.92)	97.41 (96.77)	98.04 (95.89)	98.18 (97.31)	99.57 (99.83)
Mean I/sigma(I)	14.3 (1.7)	10.8 (1.2)	9.6 (1.2)	14.3 (1.7)	14.5 (1.7)	10.2 (1.5)	9.9 (1.5)	11.1 (1.6)	8.8 (1.1)	9.1 (1.5)	12.6 (1.6)	8.5 (1.2)
Wilson B-factor	16.51	21.38	12.62	14.51	17.5	14.16	16.2	13.41	17.49	7.53	18.46	18.44
Rmerge	0.040 (0.357)	0.070 (0.991)	0.067 (0.927)	0.047 (0.626)	0.048 (0.657)	0.071 (0.677)	0.078 (0.702)	0.059 (0.603)	0.107 (1.064)	0.074 (0.621)	0.059 (0.702)	0.094 (0.948)
CC1/2	0.999 (0.850)	0.998 (0.484)	0.998 (0.461)	0.999 (0.636)	0.999 (0.700)	0.998 (0.650)	0.998 (0.646)	0.998 (0.643)	0.996 (0.456)	0.997 (0.629)	0.999 (0.684)	0.997 (0.525)
<b>Refinement</b>												
Rwork / Rfree (%)	17.7 / 20.0	18.3 / 20.2	15.9 / 20.2	14.5 / 18.3	15.1 / 19.5	15.2 / 19.6	15.5 / 20.8	14.8 / 18.9	19.1 / 21.6	15.8 / 19.7	18.6 / 20.0	18.5 / 20.4
Number of non-hydrogen atoms												
(protein/ other/ water)	1865 / 24 / 191	1825 / 41 / 197	1847 / 22 / 228	1869 / 40 / 233	1872 / 29 / 239	1871 / 31 / 207	1858 / 26 / 234	1858 / 40 / 235	1859 / 20 / 194	1846 / 44 / 256	1859 / 22 / 177	1865 / 19 / 206
Average B-factor												
(protein/ other/ water)	18.18 / 40.7 / 29.66	24.09 / 49.64 / 37.82	17.76 / 27.29 / 31.58	18.6 / 32.76 / 32.54	22.06 / 32.78 / 34.63	17.78 / 30.3 / 31.6	20.89 / 38.17 / 34.72	19.64 / 34.57 / 35.04	19.18 / 27.85 / 28.75	13.69 / 30.66 / 27.59	21.95 / 26.35 / 32.7	19.98 / 26.14 / 31.51
Protein residues	225	223	225	225	225	225	225	225	225	223	225	225
r.m.s.d. bonds (Å)	0.014	0.016	0.019	0.019	0.019	0.018	0.016	0.02	0.017	0.02	0.015	0.015
r.m.s.d. angles (deg)	1.64	1.74	1.88	1.93	1.86	1.79	1.73	1.94	1.68	2.06	1.63	1.62
Ramachandran												
favored (%)	99.1	99.09	98.21	98.65	98.65	99.55	99.1	98.21	99.1	100	97.76	99.1
allowed (%)	0.9	0.91	1.79	1.35	1.35	0.45	0.9	1.79	0.9	0	1.79	0.9
outliers (%)	0	0	0	0	0	0	0	0	0	0	0.45	0

**Supplementary Table 3. Crystallography Statistics for OTUB2 complex crystal structures.**

compound	% labeling <sup>a</sup>
Cov-Frag-OTUB2-1	100%
Cov-Frag-OTUB2-2	9%
Cov-Frag-OTUB2-3	0%
Cov-Frag-OTUB2-4	0%
Cov-Frag-OTUB2-5	9%
Cov-Frag-OTUB2-6	53%
Cov-Frag-OTUB2-7	64%
Cov-Frag-OTUB2-8	37%
Cov-Frag-OTUB2-9	0%
Cov-Frag-OTUB2-10	0%
Cov-Frag-OTUB2-11	16%
Cov-Frag-OTUB2-12	56%
Cov-Frag-OTUB2-13	32%
Cov-Frag-OTUB2-14	0%
Cov-Frag-OTUB2-15	0%
Cov-Frag-OTUB2-16	0%
Cov-Frag-OTUB2-17	100%
Cov-Frag-OTUB2-18	0%
Cov-Frag-OTUB2-19	33%
Cov-Frag-OTUB2-20	72%
Cov-Frag-OTUB2-21	0%

**Supplementary Table 4. Labeling of OTUB2 with second generation chloroacethydrazide analogs.**

<sup>a</sup> Compounds were incubated at a concentration of 100  $\mu$ M for 24 hours with 2  $\mu$ M protein in 4 °C. % labeling was assigned via intact protein LC/MS. See Supp. Fig. 8 for chemical structures of compounds.



PDB ID	5QH1	5QH8	5QH9	5QHA	5QHH
Compound	NUDT7- REV-1	PCM- 0102558	PCM- 0102716	PCM- 0102951	NUDT7- COV-1
Beamline	DLS I04-1	DLS I04-1	DLS I04-1	DLS I04-1	DLS I03
Wavelength (Å)	0.9159	0.9159	0.9159	0.9159	0.9762
Space group	P 3 2 1	P 3 2 1	P 3 2 1	P 3 2 1	P 3 2 1
Cell dimensions:					
a, b, c (Å)	123.75 123.75 40.87	124.88 124.88 41.28	124.67 124.67 41.17	123.80 123.80 41.06	124.40 124.40 41.10
α, β, γ (deg)	90.00 90.00 120.00	90.00 90.00 120.00	90.00 90.00 120.00	90.00 90.00 120.00	90.00 90.00 120.00
Resolution range	29.72 - 1.65 (1.71-1.65)	41.28 - 1.75 (1.813 - 1.75)	28.98 - 1.72 (1.78 - 1.72)	53.61 - 1.57 (1.63 - 1.57)	34.29 - 1.52 (1.57 - 1.52)
Unique reflections	43383 (4303)	37402 (3688)	39176 (3874)	49681 (4751)	56248 (5573)
Multiplicity	11.2 (11.0)	8.9 (7.9)	10.1 (9.8)	7.7 (5.8)	10.1 (10.1)
Completeness (%)	99.96 (99.95)	99.76 (99.06)	99.91 (99.95)	97.41 (93.41)	99.95 (100.00)
Mean I/sigma(I)	31.4 (2.9)	18.7 (2.8)	31.1 (2.6)	8.2 (0.5)	16.0 (1.2)
Wilson B-factor	27.8	31.98	28.14	16.91	26.12
Rmerge	0.040 (0.837)	0.093 (0.699)	0.042 (0.796)	0.122 (2.865)	0.063 (1.859)
CC1/2	1.000 (0.860)	0.998 (0.921)	1.000 (0.830)	0.997 (0.534)	0.999 (0.574)
<b>Refinement</b>					
Rwork / Rfree (%)	19.8 / 21.2	20.7 / 22.4	18.9 / 21.2	20.3 / 24.1	19.5 / 21.2
Number of non- hydrogen atoms					
(protein/ other/ water)	1489 / 34 /162	1517 / 26 / 159	1487 / 24 / 196	1484 / 27 / 175	1497 / 41 / 179
Average B-factor					
(protein/ other/ water)	32.53 / 39.56 / 43.66	36.38 / 53.74 / 46.34	29.89 / 49.01 / 42.67	30.97 / 49.05 / 47.08	30.25 / 38.72 / 42.56
Protein residues	186	186	185	185	186
r.m.s.d. bonds (Å)	0.013	0.025	0.012	0.019	0.014
r.m.s.d. angles (deg)	1.62	1.58	1.61	1.82	1.69
Ramachandran					
favoured (%)	98.3	97.73	98.29	97.71	98.3
allowed (%)	1.7	2.27	1.71	2.29	1.7
outliers (%)	0	0	0	0	0

**Supplementary Table 5. Crystallography Statistics for NUDT7 complex crystal structures.**