Unveiling the hidden interactome of CRBN molecular glues with chemoproteomics

Kheewoong Baek[1,2,†], Rebecca J. Metivier[1,2,†], Shourya S. Roy Burman[1,2], Jonathan W. Bushman[1,2], Ryan J. Lumpkin[1,2], Dinah M. Abeja[1], Megha Lakshminarayan[1], Hong Yue[1,2], Samuel Ojeda[1,2], Alyssa L. Verano[1,2], Nathanael S. Gray[3], Katherine A. Donovan[1,2,*], Eric S. Fischer[1,2,4,**]

1. Department of Cancer Biology, Dana–Farber Cancer Institute, Boston, Massachusetts 02215, USA

2. Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA

3. Department of Chemical and Systems Biology, ChEM-H and Stanford Cancer Institute, Stanford Medical School, Stanford University, Stanford, CA, 94305, USA 4. Lead Contact

[†] These authors contributed equally

[*] Correspondence: kdonovan@crystal.harvard.edu (KAD); Eric_Fischer@dfci.harvard.edu (ESF)

SUMMARY

Targeted protein degradation and induced proximity refer to strategies that leverage the recruitment of proteins to facilitate their modification, regulation or degradation. As prospective design of glues remains challenging, unbiased discovery methods are needed to unveil hidden chemical targets. Here we establish a high throughput affinity purification mass spectrometry workflow in cell lysates for the unbiased identification of molecular glue targets. By mapping the targets of 20 CRBN-binding molecular glues, we identify 298 protein targets and demonstrate the utility of enrichment methods for identifying novel targets overlooked using established methods. We use a computational workflow to estimate target confidence and perform a biochemical screen to identify a lead compound for the new non-ZF target PPIL4. Our study provides a comprehensive inventory of targets chemically recruited to CRBN and delivers a robust and scalable workflow for identifying new drug-induced protein interactions in cell lysates.

Keywords: IMiD, molecular glue, E3 ligase, cereblon, targeted protein degradation, proteomics, degrader, ubiquitin, PROTAC, PPIL4

32 **INTRODUCTION**

33 Targeted Protein Degradation (TPD) represents a promising therapeutic approach to
34 remove disease-associated proteins from cells[1,2]. The process of TPD involves hijacking the
35 ubiquitylation machinery for the covalent attachment of ubiquitin molecules to a desired protein
36 of interest, which in turn leads to degradation by the proteasome[3,4]. Ubiquitin-mediated TPD
37 utilizes two types of small molecules, molecular glues[5] and heterobifunctional degraders (also
38 known as PROteolysis Targeting Chimeras, or PROTACs)[3], both of which chemically induce
39 ternary complex formation between a protein target and a ubiquitin E3 ligase, followed by
40 proximity-driven ubiquitylation and subsequent degradation.

41 Despite the rapid growth of TPD as a therapeutic strategy, the discovery and
42 development of effective degraders remains challenging. Heterobifunctional degraders rely on
43 a linker to connect two binding warheads: one for the ligase and one for the protein of
44 interest[3,6,7]. Although this modular design offers the flexibility to target any protein with a known
45 binder, the resulting molecules often possess poor drug-like properties due to their large size.
46 Molecular glues present a possible alternative due to their small size and improved drug-like
47 properties. However, as they lack binding to their target protein and instead enhance protein-
48 protein interactions (PPI) between a ligase and substrate, rational design of molecular glues is
49 far more challenging[8]. Over the last decade, the discovery of new molecular glue degraders has
50 largely relied on serendipity through phenotypic screening of large libraries of molecules and
51 retrospective identification of their degradation targets. Although FDA-approved molecular glue
52 degraders exist, including the immunomodulatory drugs (IMiDs) thalidomide, lenalidomide, and
53 pomalidomide, they have all been characterized as molecular glues in retrospect after their
54 serendipitous discovery. IMiD molecules bind to the CUL4-RBX1-DDB1-CRBN ($CRL4^{CRBN}$) E3
55 ligase[5,9-13] creating a favorable surface for new proteins (neo-substrates) to bind for induced
56 degradation. Since this discovery, significant efforts into the design and screening of new IMiD
57 analogs have revealed up to 50 neo-substrates in the public domain, all carrying a glycine

58 containing β-hairpin structural degron[5,11,14-22]. Remarkably, computational modeling of the
59 AlphaFold2 (AF2) structures available in the Protein Data Bank (PDB) suggest that we are just
60 scraping the surface of what is targetable by these molecules[14].

61 Given the mechanism of action (MoA) of degraders, global chemoproteomics has proven
62 to be an effective tool for the identification of protein degradation targets[19,23-25]. Using this
63 approach, the target space of degraders for multiple therapeutic target classes have been
64 extensively mapped for tractable targets, including kinases[23,26,27], bromo-domains[28-30], HDACs[31]
65 and zinc finger (ZF) transcription factors[19]. Although this method has greatly expanded the
66 repertoire of known targets, limited sensitivity has restricted the ability to identify proteins with
67 low expression levels without screening libraries of cell lines or using target enrichment
68 methods. This approach also remains blind to a key aspect of these molecules: the identification
69 of proteins that are recruited to the ligase but do not ultimately get degraded. Such "non-
70 degrading glue" targets may be subject to poor lysine accessibility, lack of degradative ubiquitin
71 chain formation[32], high deubiquitinase activity, poor proteasome access, or other resistance
72 mechanisms. However, these targets still represent important therapeutic targets if these factors
73 can be overcome to convert silent molecular glues into molecular glue degraders or functional
74 modulators of the target.

75 Methods to identify chemically induced protein-protein interactions include
76 immunoprecipitation mass spectrometry (IP-MS)[33] and proximity labeling approaches coupled to
77 mass spectrometry[34,35]. IP-MS approaches have been employed for the identification of direct
78 protein interactions, whereas proximity labeling approaches are commonly employed for the
79 mapping of proximity interactomes in cells and in vivo[34,36-38], enabling the identification of protein
80 interactions within a 10-30 nm radius of the epitope tagged protein of interest[34,36,37,39]. Although
81 these in-cell methods have demonstrated successful identification of chemically induced
82 interactions, they often require extensive fine tuning of various factors including noise,
83 sensitivity, variability and scalability.

84 In this study, we establish a simple, robust and sensitive workflow to facilitate high
85 throughput discovery of degrader-induced protein-protein interactions and develop them into
86 selective tools and therapeutic candidates. We use this method to build a comprehensive
87 inventory of 298 distinct protein targets recruited to CRBN including many new zinc finger (ZF)
88 transcription factors and proteins from new target classes, including RNA-recognition motif
89 (RRM) domain proteins. We evaluate the binding potential of these targets through structural
90 alignment with IMiD-bound CRBN and performed biochemical and structural validation studies
91 on a series of non-ZF targets. We then screened a library of ~6000 IMiD analogs against a
92 novel non-ZF target, PPIL4, identifying a selective lead degrader molecule, thereby presenting a
93 blueprint for the effective discovery of novel molecular glue degraders.
94

95 **RESULTS**

96 *Unbiased identification of degrader-induced interactors in-lysate*

97 To establish a workflow for the identification of chemically induced protein-protein
98 interactions, we set out to simplify traditional IP-MS methods. We hypothesized that we could
99 create a controlled environment with reduced biological variability and enhanced scalability by
100 establishing a workflow in cell lysate using spiked in recombinant protein as the bait. Our
101 workflow harnesses small molecule degrader-induced ternary complex formation in cell lysate
102 using recombinant FLAG-tagged CRBN in complex with DDB1 excluding the BPB domain (ΔB),
103 which prevents CUL4 interaction to inhibit ubiquitylation of the recruited target, with a small
104 molecule degrader. After incubation, we enrich with a highly selective antibody for the FLAG
105 epitope tag followed by label free quantitative proteomic assessment to identify interactors
106 (**Figure 1A**). To benchmark and explore the viability of this approach for identification of protein-
107 protein interactions, we selected two representative degrader molecules that have been
108 thoroughly profiled in published reports – pomalidomide, a IMiD molecular glue[19,40] and SB1-G-
109 187, a kinase-targeted heterobifunctional degrader[26] (**Figure 1B**). We profiled these two
110 molecules across two different cell lines, MOLT4 and Kelly, selected for their orthogonal
111 expression profiles and broad coverage of known CRBN neo-substrates including IKZF1/3
112 (MOLT4) and SALL4 (Kelly)[19]. The pomalidomide screen revealed 11 different enriched proteins
113 across these two cell lines (9 in MOLT4 and 4 in Kelly cells), which revealed three novel targets,
114 ASS1, ZBED3 and ZNF219 (**Figure 1C, Figure S1A-B, Table S1**). We then validated
115 recruitment of these three novel targets to CRBN using dose response immunoblot or TR-FRET
116 analysis (**Figure S1C-D**).

117 To assess the overlap of these enriched targets with published degradation data for
118 pomalidomide, we performed a hit comparison with publicly available global proteomics data
119 (http://proteomics.fischerlab.org), which includes ten independent pomalidomide treatments
120 spanning HEK293T, Kelly and MOLT4 cell lines (**Figure 1D**). Like our enrichment data, the

121 global degradation data also maps 11 targets as degradable, however only 4 of these targets
122 (IKZF1, IKZF3, ZFP91 and SALL4) overlap with those that we see enriched in this dataset. The
123 SB1-G-187 kinase degrader screen identified 18 enriched targets across the two cell lines (16 in
124 MOLT4 and 7 in Kelly cells) including multiple non-protein kinase targets which raised the
125 question of how these proteins are being recruited to CRBN by a kinase degrader (**Figure 1E,**
126 **Figure S1E-F, Table S1**). Assessment of the non-kinase targets revealed that several are
127 known to form complexes with different kinases, such as TAB1 and TAB2 which form a
128 functional kinase complex with MAP3K7[41,42], and UNC119 which binds to myristoylated SRC to
129 regulate cellular localization[43]. This data suggests that these non-kinases are being recruited to
130 CRBN through piggybacking on their kinase binding partners. Of the other recruited non-kinase
131 targets, ZBED3 is also identified in the pomalidomide treatment suggesting recruitment through
132 the IMiD handle of the degrader and SDR39U1 was reported as a non-kinase target in a
133 chemoproteomics profiling study of kinase inhibitor probes[44]. Next, we assessed the differences
134 and overlap in hits between publicly available degradation data in MOLT4 and Kelly cells for
135 kinase-targeted heterobifunctional degrader, SB1-G-187, and enrichment data (**Figure 1E**). We
136 found 6 overlapping hits - all protein kinases - including CDK1, IRAK1, LCK, LYN, MAP3K7 and
137 SRC. Like the pomalidomide data, we observe similar numbers of proteins identified in either
138 degradation data (15) or enrichment data (12), demonstrating that these two methods
139 complement each other to expand the target scope of these molecules. Together, the data
140 collected for these two degrader molecules demonstrate the value of our workflow for identifying
141 chemically induced protein-protein interactions invisible to degradation assays, while also
142 highlighting opportunities for improving sensitivity.
143
144 *Mapping the interactomes of IMiD molecular glue degraders*
145 Next, using the functional enrichment method as a basis, we set out to optimize and
146 address the critical need for sensitivity and high throughput. IP-MS experiments typically require

147 labor-intensive sample preparation steps which create a significant source of variability and lead
148 to high background and false positive rates while also placing limits on the number of samples
149 that can be prepared in parallel. To address these limitations, we automated the enrichment and
150 sample preparation procedures to enable effective mapping of interaction targets across
151 libraries of molecules at scale. We incorporated a cost effective Opentrons OT-2 liquid handling
152 platform to automate the sample preparation process from addition of all immunoprecipitation
153 components to tryptic digestion for 96 samples in parallel (**Figure 2A**). To address throughput
154 and depth of the proteomics workflow, we took advantage of recent updates in instrumentation
155 (timsTOF Pro2, Bruker) and acquisition methods (diaPASEF)[45] that allow for significant
156 improvements in sensitivity (**Figure 2A**). In contrast to the data-dependent acquisition (DDA)
157 data collected in our proof-of-concept analysis (**Figure 1**), diaPASEF measures peptides by
158 systematically sampling all precursor ions within a specified m/z range regardless of their
159 abundance which enhances the reproducibility and depth of peptide coverage to allow for more
160 accurate and robust quantification of peptides in complex samples. Comparison of the average
161 numbers of proteins and peptides quantified between the DDA collection (**Figure 1A**) and
162 diaPASEF collection (**Figure 2A**) revealed a >5-fold increase in proteins and a ~9-fold increase
163 in peptides quantified (**Figure S2A-B**) confirming a significant improvement in depth and
164 sensitivity.

165 Work over the last several years has led to the identification of a growing list of ~50
166 neo-substrates that are recruited to CRBN by IMiD analogs for chemically induced
167 degradation[14]. Validated targets include a large number of C_2H_2 zinc finger (ZF) transcription
168 factors such as IKZF1/3[5], ZFP91[18], or SALL4[15,19], but only a few non-ZF proteins such as G1 to
169 S phase transition protein 1 (GSPT1)[11] and casein kinase 1 alpha (CK1α)[12,17]. These targets do
170 not possess any similarity, but instead all share a common structural CRBN binding motif
171 consisting of an 8-residue loop that connects the two strands of a β-hairpin and has a glycine at

172 the sixth position (G-loop)[11,12,18]. Remarkably, a recently reported analysis of available
173 AlphaFold2 (AF2) predicted structures for proteins in the human proteome uncovered over
174 2,500 proteins that harbor a G-loop potentially compatible with IMiD-recruitment to CRBN, with
175 C_2H_2 ZF proteins revealing themselves as the most prevalent domain class, aligning with the
176 dominance that this class has amongst the experimentally confirmed targets[14]. Due to the
177 extensive range of proteins that are predicted to be chemically recruitable to CRBN, we asked
178 how many of these proteins are already targeted by existing chemistry, but not yet identified due
179 to lack of sensitivity of existing methods. To explore the range of proteins chemically recruited to
180 CRBN, we screened a curated library of 20 different IMiD analogs through our automated
181 lysate-based IP workflow. We assembled this library to incorporate a broad range of IMiD-based
182 scaffolds including the parental FDA-approved IMiDs (thalidomide, lenalidomide,
183 pomalidomide)[46,47], where there is a high value to identifying new targets for drug repurposing
184 efforts. We included a series of IMiD analogs that are undergoing clinical trials (CC-220, CC-
185 92480, CC-90009)[48-50] and molecules that have demonstrated promiscuity (FPFT-2216, CC-
186 122)[51,52]. Finally, we included a series of in-house synthesized scaffolds developed in the
187 context of targeting Helios (IKZF2)[53] or part of an effort to diversify IMiDs with the addition of
188 fragments on an extended linker (**Figure 2B**). We screened this library at 1 μM concentration
189 across MOLT4 and Kelly cell lines (including a second 5 μM concentration for pomalidomide)
190 and identified proteins that were enriched in the degrader compared to DMSO control IP
191 treatment (**Figure 2C, Figures S3-4, Tables S2-3**). Using significance cutoffs of fold change
192 (FC) >1.5, P-value <0.001 and combining the data from both cell lines, we identified a total of
193 298 enriched proteins (**Table S4**). We rationalized that the likelihood of observing the same
194 proteins enriched as false positives across multiple treatments with similar IMiD analog
195 molecules is low and therefore used 'frequency of enrichment' as a measure of confidence. We
196 observed 102 proteins enriched in at least three independent IPs, and each of the top 5 proteins
197 (PATZ1, ZBED3, WIZ, IKZF2 and ASS1) enriched in more than 20 independent IPs across the

198 database (**Figure 2D, Table S4**). Surprisingly, although published reports have confirmed
199 degradation of PATZ1, WIZ and IKZF2, none of these top 5 enriched proteins regularly feature
200 amongst those proteins that we commonly see reported in existing unbiased screens of IMiD-
201 based molecules indicating the orthogonal data generated by this profiling method, identifying
202 targets that might otherwise be overlooked. ZBED3 and ASS1 showed frequent enrichment
203 across our database without any prior reporting of degradation, even at concentrations up to 5
204 μM of pomalidomide (**Figure S2C, Table S4**), suggesting the first reported examples of targets
205 that are chemically glued to CRBN but lack productive degradation, thus emphasizing the
206 benefit of alternative binding focused approaches for target identification. Also, important to note
207 is that the new targets identified in this study are not only targets of new IMiD analogs but are
208 also identified as targets of IMiDs in clinical trials and with FDA approval.

209 To assess the fraction of newly identified IMiD targets, we compared the 298 enriched
210 proteins to a list of literature reported targets and discovered an overlap of only 28 targets. We
211 identified 270 novel targets and found only 22 targets were reported in the literature but not
212 identified as hits in our study (**Figure 2E, Table S4**)[14]. Considering the prevalence of C_2H_2 ZF
213 transcription factors amongst reported IMiD targets, we asked whether this dominance holds
214 true across our extended list of targets. To assess this, we extracted superfamily, family and
215 domain information from curated databases including InterPro[54,55], Uniprot[56] and Superfamily[57]
216 to categorize the targets based on studied features (**Figure 2F-G, Table S4**). Of the 298 targets
217 identified, after C-terminal domain classification, the C_2H_2 ZF superfamily represents the largest
218 segment, accounting for >14% of the targets in the top 10 enriched superfamilies. This is
219 followed by RNA-recognition motif domain proteins (RRM, >13%) and nucleotide-binding alpha-
220 beta plait domain superfamilies (α-β plait domain, >12%). Notably, protein kinase-like domain
221 proteins also features on this top 10 list of superfamilies (kinase-like domain, >6.4%) which
222 aligns with our knowledge that kinases can be targets of IMiD molecular glues (eg, CSNK1A1 or
223 WEE1)[12,17,58] and suggests that molecular glues may be a viable alternative to PROTACs, which

224 are currently widely explored for kinase targeting. Exploration of the top 10 domain
225 classifications across the dataset shows a similar trend with C_2H_2 ZF, RRM, ZF, protein kinases
226 and BTB/POZ domains showing the highest representation across the targets identified (**Figure**
227 **2G**).
228 This dataset builds upon previous identifications of protein kinases as targets of IMiD
229 molecules[17,22,59], and further extends the kinase list adding CDK7, IRAK1 and TBK1 as novel
230 putative molecular glue targets. It also broadens the scope of tractable targets by introducing
231 multiple new families as targetable by CRBN-based molecular glues, illustrating the extensive
232 potential of these molecules. Through the application of our unbiased target enrichment
233 workflow, we have significantly increased the number of experimentally detected IMiD targets,
234 expanding beyond the C_2H_2 ZF protein family to a wide range of protein families including
235 protein kinases and proteins involved in RNA metabolism.
236
237 *Zinc finger transcription factors enriched among targets*
238 To validate the 270 previously unreported targets, we sought to establish a
239 computational screening pipeline to score the compatibility of targets for recruitment to CRBN.
240 Structural studies on IMiD-mediated CRBN neo-substrates, both natural and designed, have
241 established the common G-loop motif that is recognized by the CRBN-IMiD complex[12,60]. We
242 used MASTER[61] to mine the AF2 database[62] for proteins containing G-loops with similar
243 backbone architecture to the G-loop in known neo-substrate CSNK1A1 (PDB: 5FQD, aa35-42)
244 resulting in a set of 46,040 loops across 10,926 proteins (**Figure 3A**).
245 Due to structural constraints, not all these proteins are compatible with CRBN. To
246 identify CRBN-compatible proteins, we first extracted domains containing the G-loops based on
247 domain definitions from DPAM, a tool that parses domains from AF models based upon
248 predicted aligned errors (PAE) and evolutionary classification[63]. Next, we aligned the domains to
249 our reference CSNK1A1-IMiD-CRBN-DDB1ΔB structure based on the G-loop and calculated a

250 clash score. We used the van der Waals force term for interchain contacts in Rosetta's low-
251 resolution mode[64] to obtain a side-chain independent clash estimate. Out of 16 known neo-
252 substrates with validated G-loops (**Table S5, Figure S5**), 15 had clash scores below 2, while
253 ZNF654[40] had a score of 172, indicating a minor clash. The clash was caused by a low
254 confidence region in the AF2 structure and could be resolved by relaxing the complex with
255 Rosetta (**Figure S6A**)[65]. On the other hand, a protein with no evidence supporting it being a
256 neo-substrate, PAAF1, had a major clash with a score of 1,551 which could not be resolved by
257 relaxation (**Figure S6B**). Based on these examples, and analysis of the clash scores of all hit
258 proteins containing a clear structural G-loop in AF2 (**Figure S6C)**, we filtered out domains with
259 scores greater than 200 resulting in a list of 14,189 loops across 6,018 proteins with
260 nonexistent, or marginal clashes with CRBN (**Figure 3A, Table S5**).
261 Of the 298 total enriched candidates,199 were found to have a clear structural G-loop
262 with 162 having a clash score below 200. Given the high proportion of ZF proteins identified as
263 targets across this enrichment database (**Figure 2F**), we mapped the fold change in enrichment
264 for proteins with an annotated ZF domain across all 20 degraders for both MOLT4 (**Figure 3B**)
265 and Kelly cells (**Figure 3C**). Across these two cell lines, we identified 19 previously reported and
266 28 new neo-substrates as chemically recruited to CRBN. We then used our G-loop database to
267 inform on which of these targets have a tractable G-loop and found that only five of the 57
268 identified targets do not contain a structural G-loop (**Figure 3C, Table S5**). Given what we know
269 about the recruitment and binding of CRBN neo-substrates, targets usually bind through a
270 dominant structural hairpin. Since we do not have validated degron information for all these ZF
271 targets, we assumed that the G-loop with the lowest clash score has the highest likelihood to
272 bind and therefore proceeded with evaluation of a single G-loop for each target. To gauge how
273 the clash scores for these ZF targets compare to all hits in the G-loop database, we compared
274 the clash scores for our ZF targets to those of all hits (**Figure S6C**) demonstrating a pronounced
275 trend towards lower scores for ZF targets suggesting fewer unfavorable interactions (**Figure**

276 **S5A-B**). Notably, when we explored the ZF hits with higher clash scores (>10) and >3 hit
277 frequency, we realized that almost all of these have a reported association with at least one of
278 the validated hits – ZMYND8 (cs 455, binds to ZNF687), and RNF166 (cs 17, binds to
279 ZNF653/ZBTB39/ZNF827) – which also offers the possibility that these proteins could be
280 collateral targets, recruited via piggybacking on their binding partners, the direct binders (**Table**
281 **S4**). Finally, we compared ZF targets across the two cell lines as an additional means for
282 validation, and found 10 overlapping proteins, 6 of which are novel recruited targets (ZBED3,
283 MNAT1, MTA2, ZBTB44, TRIM28) (**Figure 3D**).

284 There are many factors to take into consideration when looking to predict target
285 degradability, such as ternary complex formation[26,31,66] and target ubiquitylation[12,32,67-69], and
286 multiple studies have placed an emphasis on exploring their role in driving productive
287 degradation[70,71]. For degrader-induced degradation to occur, a ternary complex consisting of
288 ligase-degrader-target needs to form for proximity-mediated ubiquitin transfer to the target
289 protein. Because ternary complex formation is necessary for successful protein degradation, we
290 set out to explore the relationship between ternary complex formation and degradation for ZF
291 targets identified in this study. We focused our evaluation on the parental IMiD molecules which
292 have been subjected to degradation target profiling using unbiased global proteomics analysis
293 across a panel of four cell lines (SK-N-DZ, Kelly, MM.1S, hES)[19]. Comparison of the enriched
294 ZF targets to the published degradation data shows a consistent trend across the three IMiDs
295 where only ~30% of the enriched targets that were quantified in global proteomics studies were
296 degraded, with ~ 60% of the targets quantified but not reported as degraded (**Figure 3E, Table**
297 **S5**). The data was then grouped to allow a global comparison of the enriched versus degraded
298 IMiD targets. The comparison revealed that of the 31 ZF targets enriched across these three
299 molecular glues, only 11 of the 29 proteins quantified in global proteomics experiments were
300 found to be degraded (**Figure 3F, Table S5**). 18 proteins were quantified in global proteomics
301 but were not identified as degraded. This prompted us to question whether these targets were

302 resistant to degradation by IMiDs and their analogs, or if they were not identified as degraded
303 due to experimental limitations such as inadequate sensitivity to detect minor changes in protein
304 abundance, rapid protein turnover or suboptimal experimental conditions. We found that
305 although several of the targets (WIZ, PATZ1, ZNF687, ZMYM2 and HIC2) were not reported as
306 degraded in Donovan et al.[19], they have since been reported as degraded in other published
307 studies[40,72,73] confirming that IPs provide a complementary approach able to overcome
308 limitations in sensitivity. The absence of degradation data for the remaining targets could imply
309 that these targets are resistant to degradation, or similar to the above proteins, the appropriate
310 degradation experiment has yet to be performed. These data demonstrate that our IP workflow
311 provides a significant advantage over global proteomics analysis by enabling selective isolation
312 and enrichment of targets that may be below the change in abundance threshold for consistent
313 identification with global proteomics approaches.
314
315 *IMiD derived molecular glues recruit hundreds of non-zinc finger proteins*
316 The largest target class of CRBN neo-substrates today are ZF containing proteins,
317 however, of the ~20,000 proteins in the human proteome, ZF containing proteins only make up
318 a relatively small proportion with about ~1700 ZF proteins reported[74]. So far, only a handful of
319 targets are reported to lack a ZF motif, which includes GSPT1[11], CK1α[12,17], PDE6D[75], and
320 RAB28[75]. With this in mind, we examined our list of targeted proteins with a focus on those that
321 do not contain a reported ZF domain and found 251 non-ZF proteins enriched across the IP
322 dataset (**Figure 4A, Table S4**). These non-ZF proteins include a wide range of families such as
323 protein kinases (IRAK1, TBK1, CDK7), RNA recognition motif proteins (ELAVL1, PPIL4, CSTF2,
324 RBM45), metabolic enzymes (ASS1, PAICS, ACLY, CS, ACADVL), translational proteins
325 (MARS1, ETF1, EEF1E1, EIF4B) and more spanning different biological pathways. To assist in
326 establishing confidence in some of these targets, we performed a comparison of the non-ZF
327 targets enriched in the two tested cell lines and found 39 targets were identified in both MOLT4

328 and Kelly cells, including the four above mentioned targets (**Figure 4A-B**). We then assessed
329 the AF2 structures of each of these 39 proteins and found that almost all of them (33/39) contain
330 a structural G-loop (**Figure 4B, Figure S5, Table S5**).

331 Given the large number of non-ZF targets identified in this study and the lack of
332 emphasis in the public domain with regards to non-ZF CRBN neo-substrates, we selected a
333 series of non-ZF proteins for further experimental validation. Firstly, to demonstrate that these
334 neo-substrates are directly recruited to CRBN, we examined ternary complex formation using
335 recombinant purified proteins. Using two of the more promiscuous molecular glues,
336 pomalidomide and FPFT-2216, we tested previous reported degradation targets PDE6D,
337 RAB28 and DTWD1, along with a newly discovered target PPIL4. Indeed, PDE6D, DTWD1,
338 and PPIL4 formed compound dependent ternary complex with CRBN at varying effective
339 concentrations (**Figure 4C**). However, RAB28, which was previously reported to be degraded
340 by IMiDs[19] and FPFT-2216[75], did not show any evidence for direct binding to CRBN using
341 purified proteins. Since RAB28 has previously been reported as a CRBN neo-substrate and
342 consistently scored across our enrichment study, we explored whether there was any evidence
343 suggesting that RAB28 could be a collateral target. Exploration of protein-protein interaction
344 databases including BioPlex[33] and STRING-DB[76] revealed that RAB28 is known to bind to two
345 validated IMiD-CRBN neo-substrates PDE6D and ZNF653 (**Table S4**), suggesting that RAB28
346 is likely an indirectly recruited target. These data demonstrate that in addition to identifying
347 direct binders, we can also identify indirect binding partners that may be simultaneously
348 recruited together with direct binding neo-substrates. Given that targeted protein degradation
349 requires not only recruitment to CRBN, but also CRBN mediated ubiquitin transfer for
350 degradation, we also monitored whether the recruited proteins can be ubiquitylated by
351 $CRL4^{CRBN}$. In vitro ubiquitylation assays showed robust ubiquitin modification on all 3 recruited
352 non-ZF proteins in the presence of pomalidomide or FPFT-2216 (**Figure 4D**). In addition, all
353 three of these targets were degraded in response to IMiD treatment as observed by global

354 proteomics analysis (**Figure S6D, Table S5**). Using structural G-loop alignments, we then
355 assessed the potential for each of these three proteins to bind to IMiD-CRBN and found that all
356 three proteins had a G-loop with a clash score of <200 (**Figure 4E**). However, the aligned clash
357 score for DTWD1 was relatively high and approaching the upper 200 threshold (cs 198). We
358 performed relaxation with Rosetta and found that this reduced the clash score to 1.58 by
359 allowing minor shifts in the overall conformation while retaining the structural G-loop (**Figure**
360 **S6E**). This process demonstrates that in some cases, clash scores can be relieved through
361 minor structural rearrangements using Rosetta relax.

362 To expand our understanding of the recruitment of non-ZF targets, we determined
363 cryo-EM structures of CRBN-DDB1ΔB-FPFT-2216 bound to PPIL4 and PDE6D, respectively
364 (**Figure 4F, Figure S7, Table 1**). The complex structures were both refined to a global
365 resolution of around 3.4 Å and the quality of the resulting maps were sufficient to dock the
366 complex components, but the flexibly tethered PPIL4 resulted in a lower local resolution. We
367 were able to observe PPIL4 engagement with FPFT-2216-CRBN via its Gly278 harboring G-
368 loop as expected from the G-loop alignment, as well as for PDE6D via its Gly28 G-loop.
369 Furthermore, overall density allowed fitting of FPFT-2216 in bulk although the reduced
370 resolution in that region did not permit exact positioning of the molecule. Nevertheless, we were
371 able to see that the glutarimide ring engages CRBN's binding pocket in a similar manner to
372 other IMiD molecular glues. The triazole interacts with the backbone of the G-loop, and the
373 methoxythiophene moiety potentially contacts both the PPIL4 backbone of the G-loop and
374 Arg273. This suggests that the triazole and the methoxythiophene moieties could provide
375 specificity elements to FPFT-2216 mediated neo-substrate recruitment. The methoxythiophene
376 moiety also engaged Arg23 of PDE6D, indicating that FPFT-2216 might derive specificity in
377 engaging an arginine residue from its neo-substrates. Analysis of the non-ZF targets of FPFT-
378 2216 revealed several other proteins harboring an arginine or a lysine residue at this sequence
379 location (PDE6D, SCYL1, RBM45, PPIL4). Finally, we compared the experimental structure to

380 the AF2 predicted G-loop aligned structure of PPIL4 (**Figure S6F**). The G-loop aligned structure
381 of PPIL4 presented a clash score of 3.38, which showed the C-terminal region of CRBN around
382 residue Arg373 to be clashing with PPIL4's loop harboring residue Val250. Although the low
383 resolution permitted only backbone level fitting of PPIL4, we observed that the cryo-EM
384 structure revealed a minor shift in the RRM domain of PPIL4 to accommodate this minor clash
385 suggested in the G-loop aligned structure while retaining overall conformational similarity of the
386 G-loop (**Figure S6F**). Meanwhile, PDE6D retained overall similar conformation with minor shifts
387 that did not alter the interaction with CRBN (**Figure S6G**).
388 These data demonstrate that RRM domain containing proteins represent a new class
389 of proteins targetable through CRBN dependent molecular glues. Using structural modeling we
390 increase confidence in these new targets while also providing a reminder that structural analysis
391 and AF2 predicted structures are static models and although they provide excellent structural
392 guidance, we need to keep in mind that proteins in solution are flexible and dynamic.
393
394 *Discovery of new and selective molecular glue for PPIL4*
395 While the proteomics-based screening workflow identifies novel putative CRBN targets
396 and provides initial chemical matter, it does not necessarily provide the best starting point for
397 developing a chemical probe or therapeutic due to the limited number of molecules screened.
398 We hypothesized that this limitation could be overcome by following up proteomics screening
399 with a target centric screen of a larger CRBN binder library to identify the optimal chemical
400 starting point. To test this, we set out to identify PPIL4 targeting molecular glues with improved
401 selectivity and lacking the triazole moiety. We employed an IMiD molecular glue library
402 consisting of ~6000 compounds of various IMiD analogs that were either synthesized in-house
403 or purchased externally. We screened this library against PPIL4 using TR-FRET to measure
404 compound-induced PPIL4 recruitment to CRBN (**Figure 5A**). TR-FRET ratios were obtained by
405 incubating the library with GFP fused CRBN-DDB1ΔB, biotinylated PPIL4, and Tb-labeled

406 streptavidin that binds to the biotinylated PPIL4. The library was compared relative to the
407 positive control, whereby the 520/490 ratio of FPFT-2216 at 10 μM was normalized as 1, and
408 compounds were tested at 1.66 μM or 3.33 μM to find hits with equal or improved efficacy in
409 directly recruiting PPIL4 to CRBN-DDB1ΔB. We were able to narrow down the library to two
410 molecules that performed similar or better than FPFT-2216 (**Figure 5B)**. These lead compounds
411 were subject to a full titration to assess recruitment efficacy by TR-FRET. Ultimately, after
412 recognizing one of the two hits was due to autofluorescence, we were able to identify a
413 molecule, Z6466608628, that produced a higher 520/490 ratio, and a better EC50 of 0.34 μM
414 compared to FPFT-2216, measured at 1.05 μM in this experiment (**Figure 5C-D**).
415 To test the efficacy and selectivity of our lead compound, we first performed IP-MS in
416 comparison with FPFT-2216 in Kelly cell lysate. While FPFT-2216 recruited many proteins,
417 Z6466608628 selectively recruited PPIL4, along with its binding partner DHX40 (**Figure 5E,**
418 **Table S5**). We then performed global proteomics in MOLT4 cells to confirm that Z6466608628
419 can induce selective downregulation of PPIL4 (**Figure 5F, Table S5**). These data collectively
420 demonstrate the complete workflow, starting from the identification of a novel non-ZF target
421 PPIL4 in a chemoproteomics screen, to the discovery of a new PPIL4 selective molecular glue
422 that would serve as an excellent lead molecule for structural optimization.
423

424 **DISCUSSION**

425 Targeted protein degradation and induced proximity are part of a rapidly expanding field
426 focused on the development of small molecules that leverage induced neo-protein-protein
427 interactions to drive pharmacology. In this study, we develop and showcase a new workflow for
428 high sensitivity, unbiased target identification of degraders and non-degrading molecular glues,
429 identifying more than 290 targets recruited to CRBN by IMiD-like molecules. We demonstrate
430 that this new approach to target identification can reveal critical insights and new targets that
431 are missed by traditional screening methodologies and provide a blueprint from discovery to
432 optimization and structure guided design of new molecular glue degraders.

433 Thalidomide and its derivatives, lenalidomide and pomalidomide (IMiDs), have had a
434 checkered past. These molecules have been in use for a variety of indications, on and off, since
435 the 1950's and have experienced perhaps the greatest turnaround in drug history. From
436 devastating birth defects to effective hematological cancer therapy, and more recently,
437 significant investment in utilizing these molecules for TPD-based therapeutics. While a decade
438 of research has slowly uncovered around 50 reported neo-substrates of IMiD's, thousands of
439 proteins harbor G-loops that have the potential for recruitment to CRBN by IMiD molecules. Our
440 simple, cost effective and highly scalable unbiased screening workflow combines whole cell
441 lysate with recombinant Flag-CRBN and degraders to enrich target binders from the complex
442 proteome. Through an IMiD-analog diversity screen across two cell lines, we mapped a
443 significant expansion of the neo-substrate repertoire by identifying 298 proteins recruited to
444 CRBN, with 270 of these being novel targets. Unlike many current high throughput screening
445 workflows that focus on the end point – degradation, this workflow allows us to explore the
446 fundamental first step of proximity induced degradation – recruitment, where we are now able to
447 identify targets that are directly or indirectly recruited to CRBN. This sensitive workflow sheds
448 light onto a previously unchartered element of the molecular glue mechanism of action and
449 establishes insights into how and why certain molecular glues may exhibit higher efficacy than

450 others. Surprisingly, we discovered targets recruited to CRBN that are resistant to degradation,
451 demonstrating the first examples of targets being glued to CRBN without productive
452 degradation. Exploration of two of these targets, ASS1 and ZBED3, does not offer any clues as
453 to why they are not degraded since both have reported ubiquitylated sites[77]. Numerous
454 possibilities exist, from these targets being tightly preoccupied by other binding partners,
455 geometric constraints leading to inaccessibility of lysines, removal of ubiquitylation by
456 deubiquitinases, or preclusion of the catalytic sites due to size and shape preventing active
457 ubiquitylation. It is important to note that the non-degrading functions of these molecular glues
458 may have interesting degradation-independent pharmacology that have not yet been
459 investigated, thus providing an opportunity for future experimental research.

460 The comprehensive G-loop database provided us with prefiltered insights as to whether
461 these targets have the potential to be recruited to CRBN through the currently established
462 mechanism of G-loop binding. However, although most targets identified in this study do have a
463 structural G-loop, we do have numerous instances of proteins that do not harbor a G-loop.
464 Some of these targets do have a structurally similar hairpin motif but are lacking the 'essential'
465 glycine in position six. Whereas other targets did not have this structural motif at all. These
466 findings indicate the potential for alternative recruitment mechanisms such as proteins
467 piggybacking on a direct G-loop carrying target. This concept of collateral (or bystander)
468 targeting was also demonstrated in a study exploring HDAC degradability, where it was found
469 that both HDACs, and their known complex binding partners can be degraded[31]. Alternatively,
470 and perhaps more intriguingly, the potential for recruitment of proteins through a distinct
471 structural motif suggesting there may be new binding mechanisms that are pending discovery.
472 The potential capacity for IMiDs to yield interfaces favorable for recruitment of various structural
473 motifs would considerably expand the diversity CRBN neo-substrates and broaden therapeutic
474 applications.

475 Amongst the targets identified in this study, we not only discovered many new C_2H_2 ZF
476 transcription factor targets but also extended targets beyond C_2H_2 ZF proteins, into additional
477 classes of proteins such as those containing RNA recognition motif (RRM) domain and kinase
478 domains, confirming that CRBN is an incredibly versatile ligase and very well suited to hijacking
479 for TPD applications. We reveal 251 non-ZF targets, a dramatic increase in the breadth and
480 number of proteins targeted by CRBN from the currently reported targets of less than a dozen.
481 Direct binding data using TR-FRET on a selection of these targets validates their direct binding
482 mechanism, and structural characterization further corroborates this binding while validating the
483 generated G-loop alignment database as a tool to assist prioritization of targets using clash
484 score assessment. Using the accumulative data, we selected a novel non-ZF neo-substrate,
485 PPIL4, for additional screening to illustrate the utility of this workflow for prioritization efforts.
486 After a biochemical ternary complex recruitment screen of around 6,000 IMiD analogs, we
487 selected a single hit compound and used chemoproteomics to confirm selective recruitment of
488 PPIL4 to CRBN. Genomic studies have reported that PPIL4 is essential for brain specific
489 angiogenesis and has implications in intracranial aneurysms[78], and is known to regulate the
490 catalytic activation of the spliceosome[79]. Thus, this new molecular glue could be of great interest
491 to target the splicing pathway, in relation to intracranial aneurysms, or in other contexts.
492 We believe our strategic workflow and comprehensive data package, along with outlining
493 specific applications of these, provides a valuable resource for the chemical biology, drug
494 discovery and induced proximity communities. Importantly, the workflow is neither limited to
495 CRBN nor to TPD, but rather can be applied to any induced proximity application. We expect
496 the enrichment workflow will provide a blueprint for expansion into target identification for
497 induced proximity platforms as well as further expansion of targets for protein degraders beyond
498 molecular glues. Through initial scouting efforts on heterobifunctional degraders and additional
499 ligases we are confident there are many novel discoveries to be made with already existing

500 chemistry and we envision this as an evolving resource where we will continue to release data
501 as it becomes available.
502
503
504 **SIGNIFICANCE**
505 Degraders and molecular glues are small molecules that can target and promote the
506 degradation of specific proteins providing a novel approach for modulating protein function.
507 Currently available unbiased methods to identify targets of degraders, although successful in
508 identifying transient and/or degraded targets, are limited in sensitivity and ability to identify direct
509 binders of these molecules, prohibiting identification of targets that have weak expression
510 changes or are glued and not degraded. Here, we develop an automatable high throughput
511 method for the identification of chemically-induced binders. We demonstrate the ability to
512 comprehensively identify new targets by identifying 298 neo-substrates of CRBN, significantly
513 expanding the repertoire of actionable targets. We then used structural and biochemical
514 characterization alongside a computational structural alignment workflow to validate hit targets
515 and selected one novel target, PPIL4, to perform a focused biochemical screen for the
516 identification of a new lead molecule. CRBN is the most targeted ligase in the TPD field, with
517 molecules FDA approved and more in clinical trials it is important that we understand the
518 complete cellular and molecular impact of targeting this ligase. The findings presented in this
519 study, open a new and complementary avenue for target identification and create a valuable
520 data resource mapping a wide range of neo-substrates of the CRBN ligase. Through expansion
521 of the range of CRBN targets, we not only enhance our knowledge of newly druggable targets
522 and offer new avenues for therapeutic development, but we also enhance our understanding of
523 the molecular mechanisms and cellular pathways that are influenced by existing and future IMiD
524 molecules providing opportunities for improved drug design.
525

526 **ACKNOWLEDGEMENTS**

527 We thank S. Dixon-Clarke, M. Hunkeler, T. Levitz, Y. Xiong, J. Che, T. Zhang and members of
528 the Fischer and Gray labs for helpful discussions, reagents, and support. We thank the Harvard
529 Cryo-EM center for Structural Biology for support on data collection. Financial support for this
530 work was provided by the National Institutes of Health (R01CA214608 and R01CA262188 (both
531 to E.S.F.). K.B. is a Damon Runyon Fellow supported by the Damon Runyon Cancer Research
532 Foundation (DRG-2514-24). Figures 1A, 2A and 5A were created in Biorender.

533

534 **AUTHOR CONTRIBUTIONS**

535 K.B. designed experiments, performed structural work and biochemical assays, analyzed the
536 data, interpreted results, wrote the manuscript. R.J.M. initiated the study, designed proteomics
537 experiments, performed proteomics experiments, analyzed the data. S.S.R.B. performed
538 computational alignment analysis, interpreted results. J.W.B. initiated the study and performed
539 biochemical experiments. R.J.L. wrote proteomics analysis code. D.M.A performed proteomics.
540 M.L. performed immunoblots. H.Y. performed TR-FRET screen. S.O. performed TR-FRET
541 screen. A.L.V. synthesized molecules. N.S.G. supervised experiments. K.A.D. conceived the
542 study, designed experiments, analyzed the data, interpreted the results, wrote the manuscript
543 and supervised the study. E.S.F. conceived the study, interpreted results, supervised and
544 funded the study. All authors read, edited and approved the final manuscript.

545

546 **DECLARATION OF INTERESTS**

547 E.S.F. is a founder, scientific advisory board (SAB) member, and equity holder of Civetta
548 Therapeutics, Proximity Therapeutics, Stelexis Biosciences, and Neomorph, Inc. (also board of
549 directors). He is an equity holder and SAB member for Avilar Therapeutics, Photys
550 Therapeutics, and Ajax Therapeutics and an equity holder in Lighthorse Therapeutics and Anvia
551 Therapeutics. E.S.F. is a consultant to Novartis, EcoR1 capital, Odyssey and Deerfield. The

Fischer lab receives or has received research funding from Deerfield, Novartis, Ajax, Interline, Bayer and Astellas. K.A.D receives or has received consulting fees from Kronos Bio and Neomorph Inc. N.S.G. is a founder, science advisory board member (SAB) and equity holder in Syros, C4, Allorion, Lighthorse, Inception, Matchpoint, Shenandoah (board member), Larkspur (board member) and Soltego (board member). The Gray lab receives or has received research funding from Novartis, Takeda, Astellas, Taiho, Jansen, Kinogen, Arbella, Deerfield, Springworks, Interline and Sanofi.

FIGURE TITLES AND LEGENDS

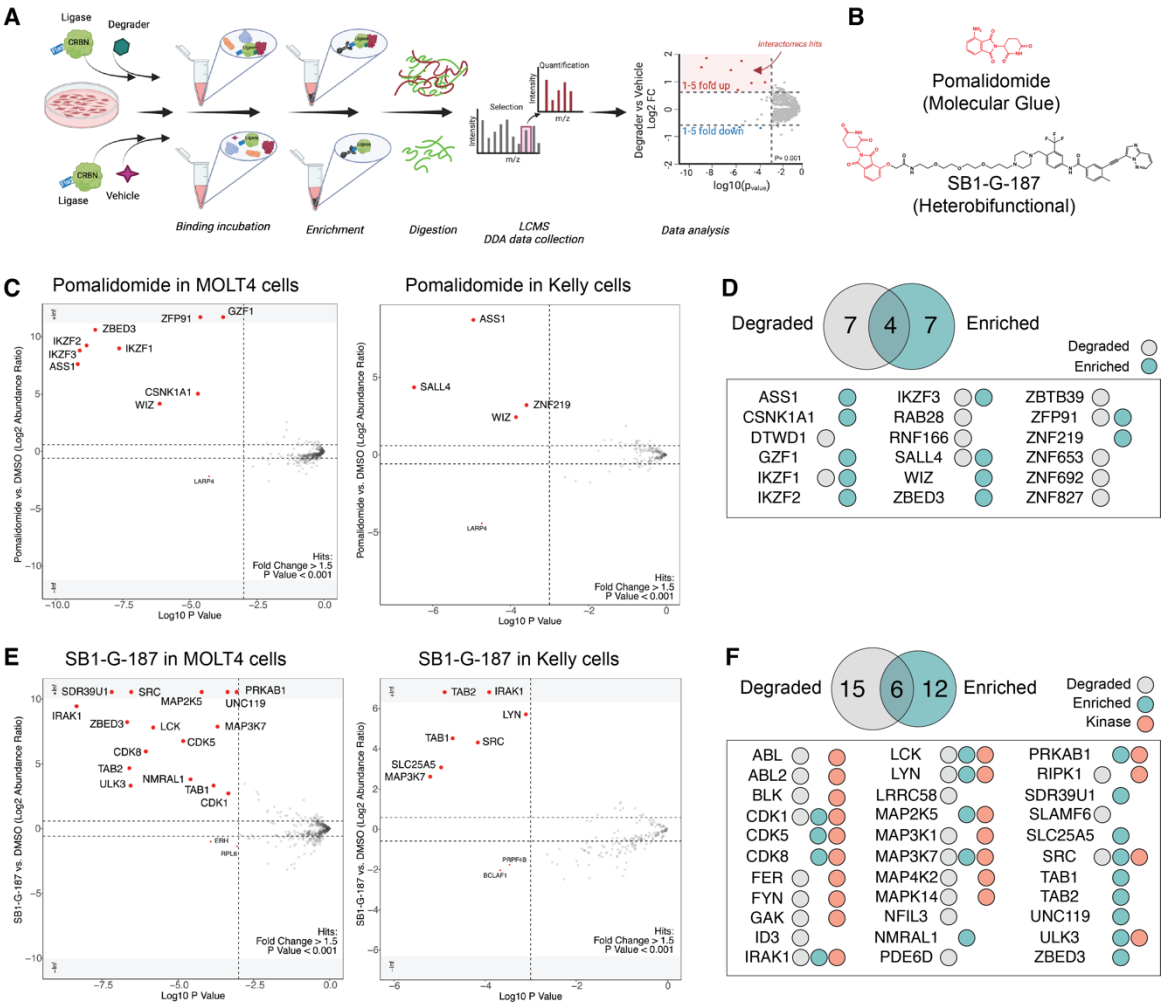


Figure 1 | Proof of concept for target enrichment in-lysate.

(**A**) Schematic representation of the first-generation enrichment-based quantitative proteomics workflow established for target enrichment and identification. (**B**) Chemical structures of degraders – Pomalidomide (molecular glue) and SB1-G-187 (heterobifunctional). (**C**) Scatterplots depicting relative protein abundance following Flag-CRBN-DDB1ΔB enrichment from in-lysate treatment with 1 μM Pomalidomide and recombinant Flag-CRBN-DDB1ΔB spike in. Left: MOLT4 cells and Right: Kelly Cells. Scatterplots display fold change in abundance to DMSO. Significant changes were assessed by moderated t-test as implemented in the limma package[80] with $\log_2$ FC shown on the y-axis and negative $\log_{10}$ P-value on the x-axis. (**D**) Venn diagram showing unique and overlapping hits for Pomalidomide found in our enrichment study and in publicly available whole cell degradation data. (**E**) As in **C**, but with 1 μM SB1-G-187 treatment. (**F**) As in **D**, but with SB1-G-187 treatment.

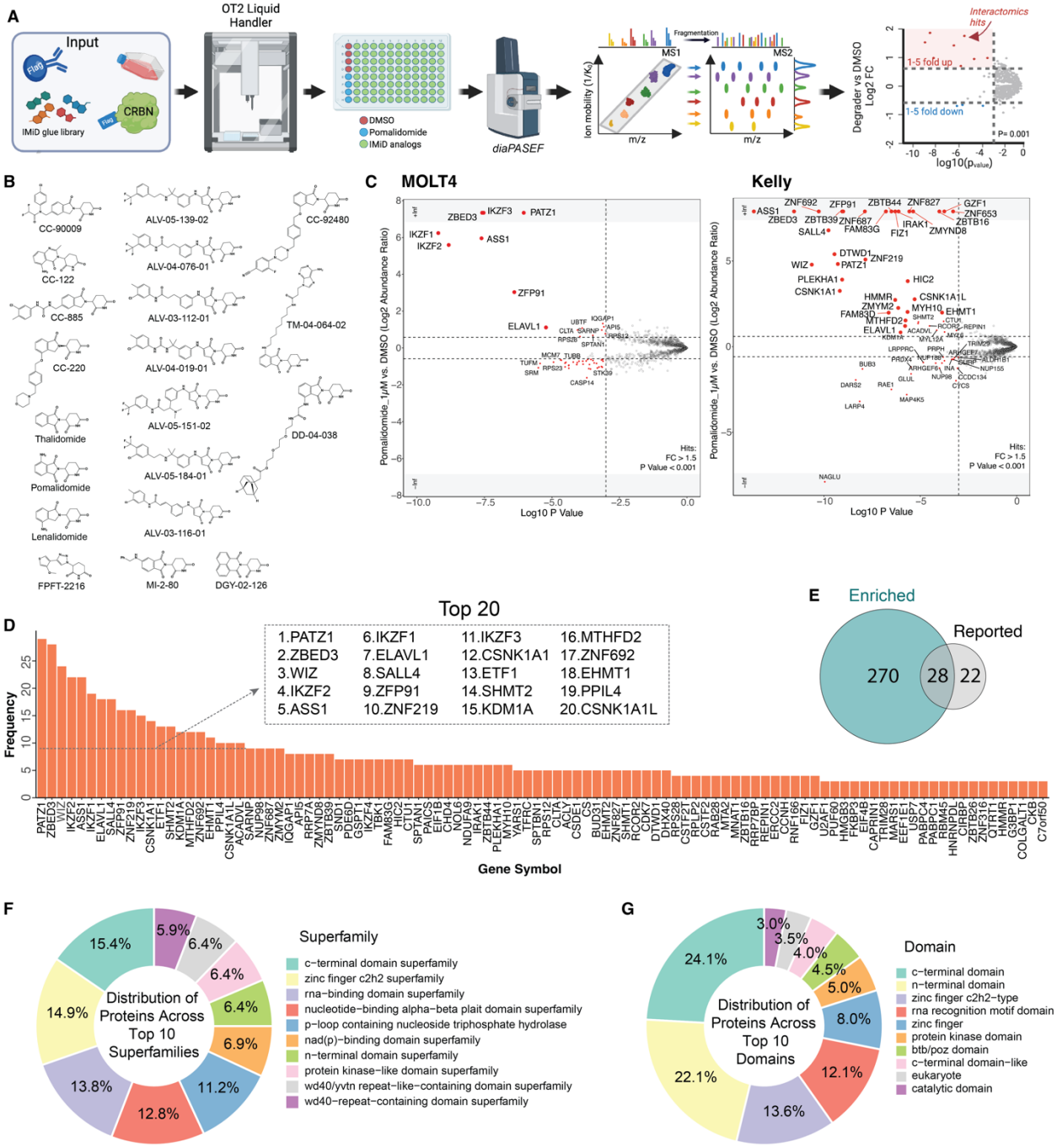


Figure 2 | Unveiling and mapping CRBN recruited neo-substrates.

(**A**) Schematic representation of the second-generation enrichment-based quantitative proteomics workflow established for target enrichment and identification. (**B**) Chemical structures of the 20 CRBN-based degraders profiled in this study. (**C**) Scatterplot depicting relative protein abundance following Flag-CRBN-DDB1ΔB enrichment from in-lysate treatment

582 with degrader and recombinant Flag-CRBN-DDB1ΔB spike in. Scatterplot displays fold change
583 in abundance to DMSO. Significant changes were assessed by moderated t-test as
584 implemented in the limma package[80] with $\log_2$ FC shown on the y-axis and negative $\log_{10}$ P-
585 value on the x-axis. Scatterplots for all 21 treatments across MOLT4 and Kelly cells can be
586 found in separate PDF's "**Figures S3-4**", representative example for a single treatment
587 (Pomalidomide, 1 μM) is displayed here. (**D**) The number of independent IPs for which
588 enrichment was observed for each target. Inset, the top 20 frequently enriched target proteins.
589 (**E**) Venn diagram showing unique and overlapping hits found in our enrichment study and in
590 published literature. (**F**) Donut chart representing the proportions of enriched proteins contained
591 within the Top 10 different superfamily categories. (**G**) Donut chart representing the proportions
592 of enriched proteins contained within the Top 10 different domain categories.
593

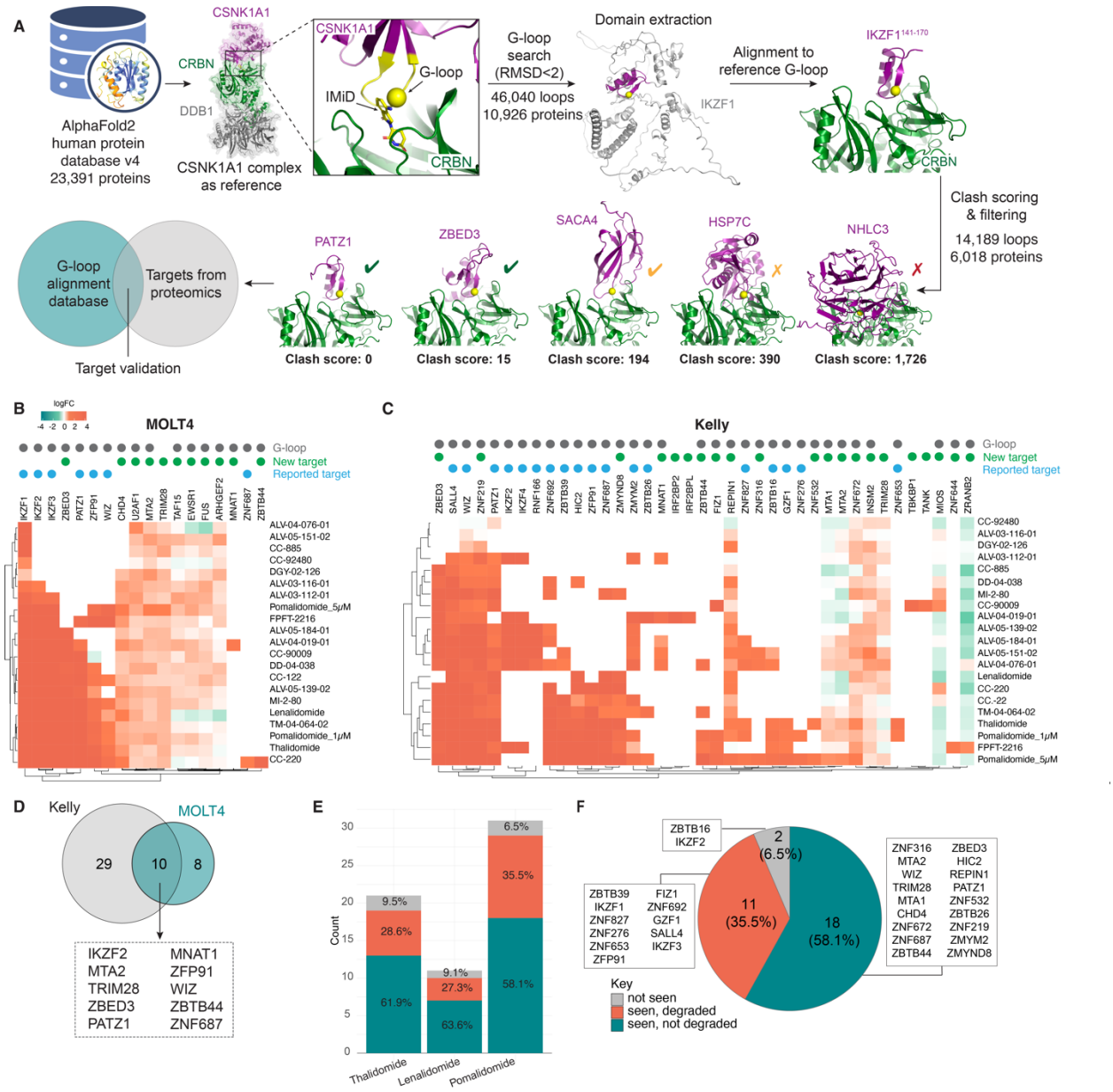
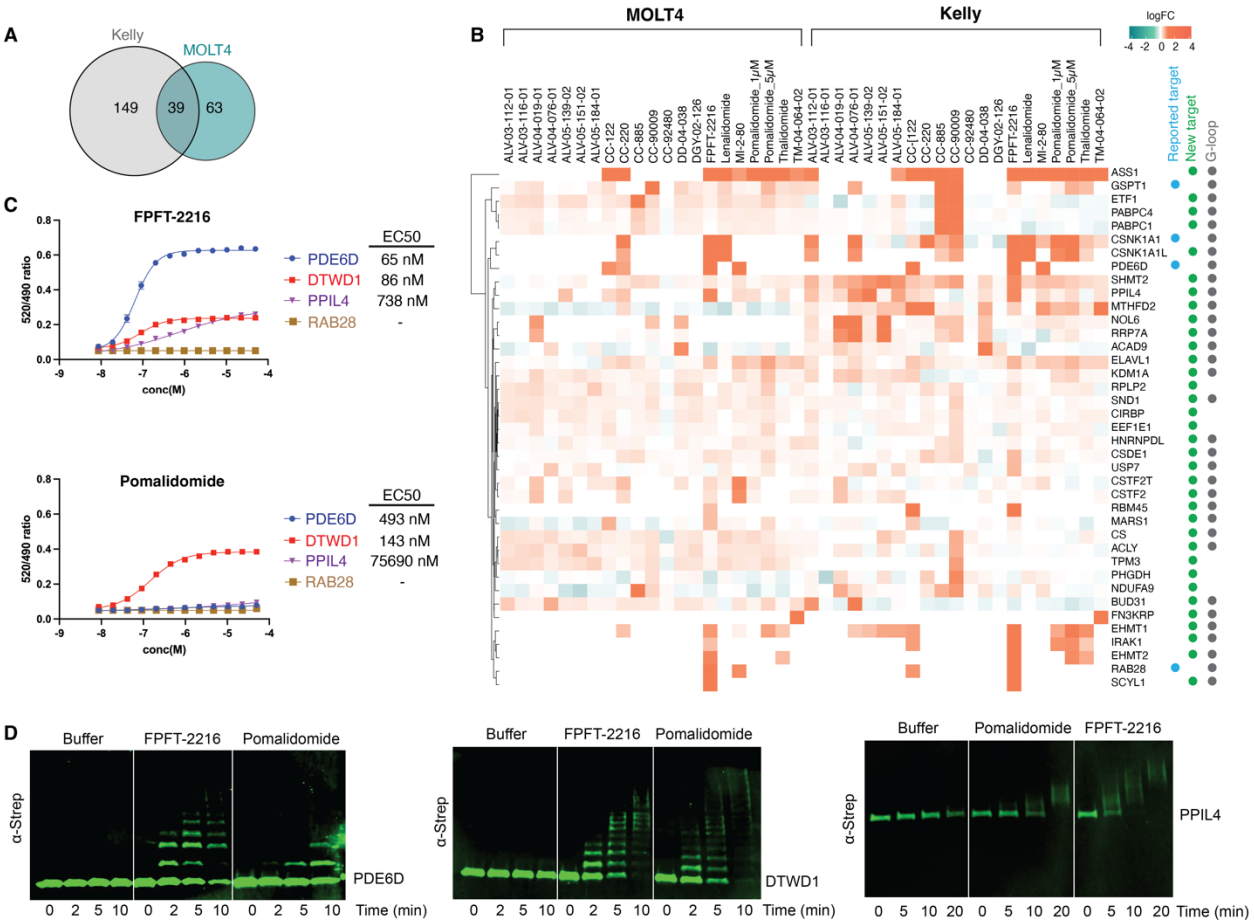
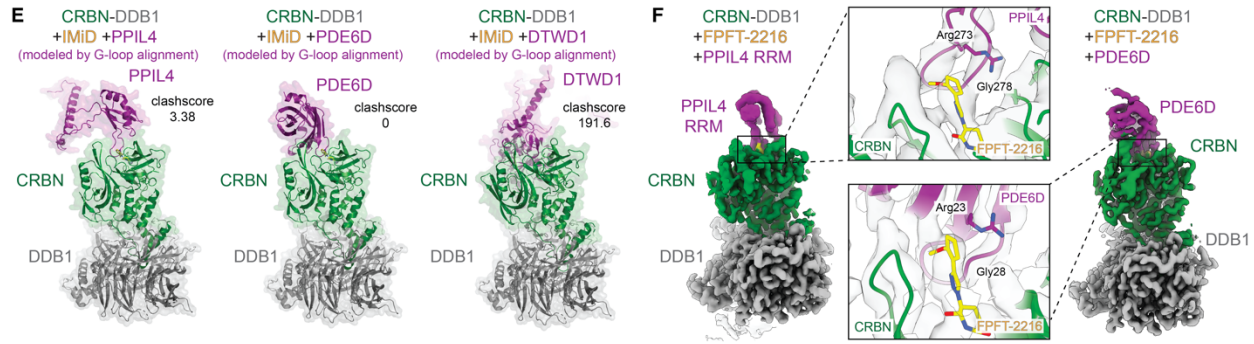


Figure 3 | Structural alignment and assessment of ZF CRBN neo-substrates. (**A**) Schematic representation of the computational workflow established for AF2 G-loop binding compatibility with CRBN-IMiD. (**B**) Heatmap displaying the log2 fold change (log2 FC) of significant (P-value <0.001) molecular glue dependent ZF targets in MOLT4 cells. White space in the heatmap corresponds to log2FC = 0 or no quantification. Previously reported targets are marked with a blue dot, newly reported targets are marked with a green dot and targets with a structural G-loop are marked with a gray dot. Significant changes were assessed by moderated t-test as implemented in the limma package[80]. (**C**) As in **B**, but with Kelly cells. (**D**) Venn diagram showing unique and overlapping ZF hits comparing MOLT4 and Kelly cell targets. (**E**)

604 Stacked bar plot showing the proportion of targets complexed and degraded by each of the
605 indicated IMiD molecules. “not seen” indicates enriched targets were not quantified in global
606 proteomics studies. “seen, degraded” indicates enriched targets quantified and reported as
607 degraded in global proteomics. “seen, not degraded” indicates enriched targets were quantified
608 but not degraded in global proteomics [19]. (**F**) Pie chart displays the IMiD-grouped data from **E**.
609

610


A
Kelly
MOLT4
149
39
63
B
MOLT4
Kelly
logFC
-4 -2 0 2 4
Reported target
New target
G-loop
ASS1
GSPT1
ETF1
PABPC4
PABPC1
CSNK1A1
CSNK1A1L
PDE6D
SHMT2
PPIL4
MTHFD2
NOL6
RRP7A
ACAD9
ELAVL1
KDM1A
RPLP2
SND1
CIRBP
EEF1E1
HNRNPDL
CSDE1
USP7
CSTF2T
CSTF2
RBM45
MARS1
CS
ACLY
TPM3
PHGDH
NDUFA9
BUD31
FN3KRP
EHMT1
IRAK1
EHMT2
RAB28
SCYL1
C
FPFT-2216
520/490 ratio
conc(M)
EC50
PDE6D 65 nM
DTWD1 86 nM
PPIL4 738 nM
RAB28 -
Pomalidomide
EC50
PDE6D 493 nM
DTWD1 143 nM
PPIL4 75690 nM
RAB28 -
D
Buffer
FPFT-2216
Pomalidomide
α-Strep
PDE6D
DTWD1
PPIL4
Time (min)


E
CRBN-DDB1
+IMiD +PPIL4
(modeled by G-loop alignment)
PPIL4
clashscore
3.38
CRBN
DDB1
CRBN-DDB1
+IMiD +PDE6D
(modeled by G-loop alignment)
PDE6D
clashscore
0
CRBN-DDB1
+IMiD +DTWD1
(modeled by G-loop alignment)
DTWD1
clashscore
191.6
F
CRBN-DDB1
+FPFT-2216
+PPIL4 RRM
PPIL4
RRM
Arg273
Gly278
FPFT-2216
Arg23
Gly28
CRBN-DDB1
+FPFT-2216
+PDE6D
PDE6D

611

612 **Figure 4 | Assessment and validation of CRBN non-ZF neo-substrates.**
613 (**A**) Venn diagram showing unique and overlapping non-ZF hits comparing MOLT4 and Kelly
614 cell targets. (**B**) Heatmap displaying the log2 fold change (log2 FC) for the 39 overlapping hits
615 from **A**. White space in the heatmap corresponds to log2FC = 0 or no quantification. Previously
616 reported targets are marked with a blue dot, newly reported targets are marked with a green dot
617 and targets with a structural G-loop are marked with a gray dot. Significant changes were
618 assessed by moderated t-test as implemented in the limma package[80]. (**C**) TR-FRET with
619 titration of FPFT-2216 or pomalidomide to N-terminally biotinylated FL PDE6D, DTWD1, PPIL4
620 or RAB28 at 20 nM, incubated with terbium-streptavidin at 2 nM to monitor binding to GFP-
621 CRBN-DDB1ΔB at 200 nM. Values were determined by technical replicates of n=2. **(D)**
622 Immunoblots of ubiquitylation assay establishing PDE6D, DTWD1 and PPIL4 as FPFT-2216
623 and pomalidomide-induced neo-substrates of CRBN. (**E**) Structural G-loop alignment of AF2
624 PPIL4, PDE6D, and DTWD1 with CRBN-DDB1ΔB (PDB ID 5QFD, 6UML). Corresponding clash
625 score is displayed. (**F**) Cryo-EM 3D reconstruction of PPIL4-RRM bound in ternary complex with
626 FPFT-2216-CRBN-DDB1 FL, and PDE6D bound with FPFT-2216-CRBN-DDB1ΔB. Maps are
627 postprocessed with DeepEMhancer[81]. Inset of each shows the potential binding mode of action
628 of FPFT-2216 engaging PPIL4 or PDE6D via neo-substrate G-loop and its interacting residues.
629

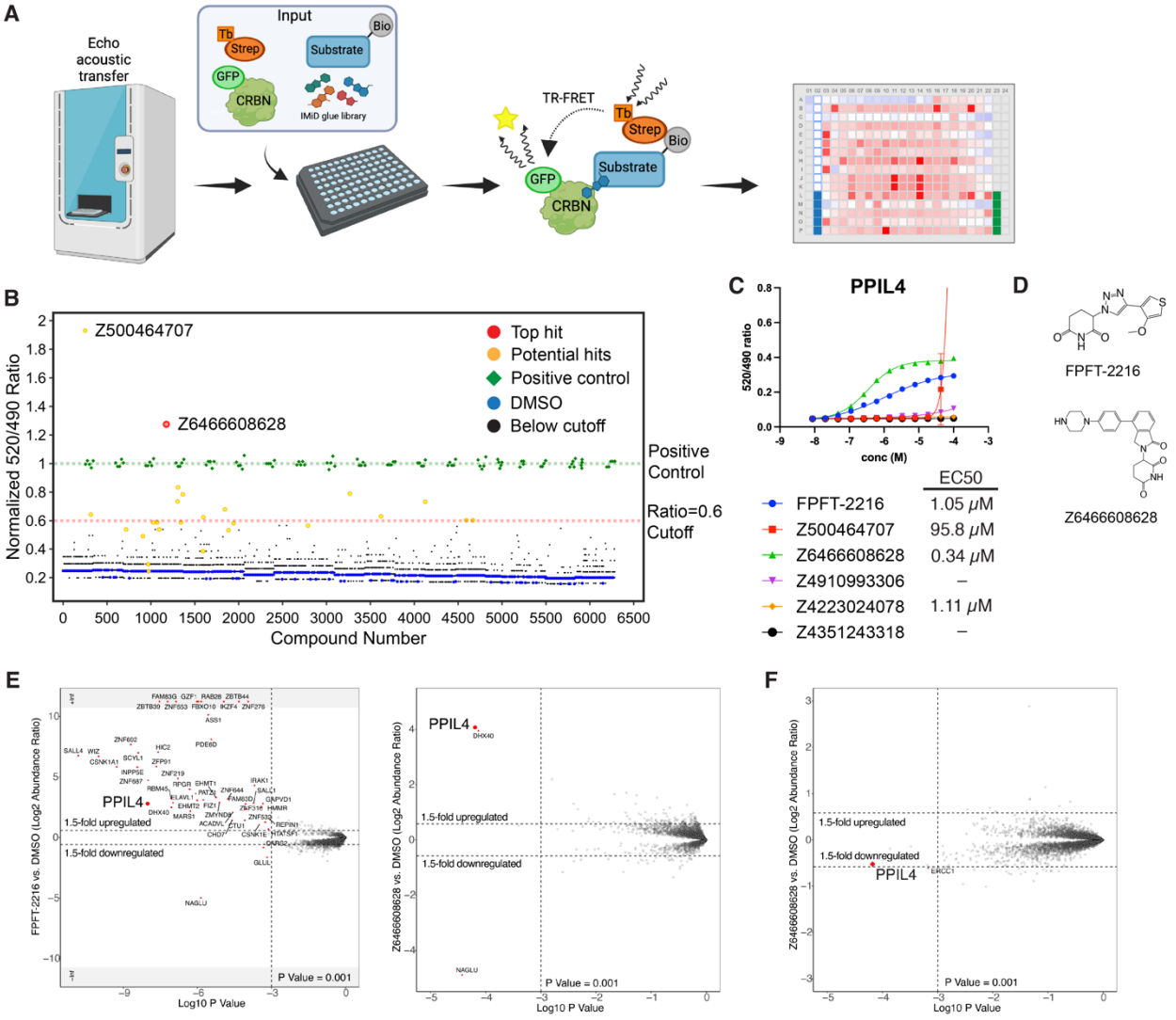


Figure 5 | High throughput IMiD analog library screen for improved hit molecules for PPIL4. (**A**) Schematic of the high throughput TR-FRET screening workflow used to screen >6,000 IMiD analogs. (**B**) TR-FRET: Normalized 520/490 ratio for each of the >6,000 compounds derived from IMiD molecules combined from the Gray/Fischer laboratories and those purchased from Enamine with GFP-CRBN-DDB1ΔB at 50 nM, biotinylated PPIL4 at 20 nM, and terbium-streptavidin at 2 nM. (**C**) TR-FRET: titration of FPFT-2216 and lead compounds to GFP-CRBN-DDB1ΔB at (200 nM), biotinylated PPIL4 at 20 nM, and terbium-streptavidin at 2 nM. Values were determined by technical replicates of n=2. (**D**) Chemical structures of FPFT-2216 alongside new lead compound from **B** and **C**. (**E**) Scatterplot depicting relative protein abundance following Flag-CRBN-DDB1ΔB enrichment from Kelly cell in-lysate

642 treatment with FPFT-2216 (left) and Z6466608626 (right) and recombinant Flag-CRBN-
643 DDB1ΔB spike in. Scatterplot displays fold change in abundance to DMSO. Significant changes
644 were assessed by moderated t-test as implemented in the limma package[80] with $\log_2$ FC shown
645 on the y-axis and negative $\log_{10}$ P-value on the x-axis. (**F**) Scatterplot depicting relative protein
646 abundance following Z6466608626 treatment in MOLT4 cells. Significant changes were
647 assessed by moderated t-test as implemented in the limma package[80] with $\log_2$ FC shown on
648 the y-axis and negative $\log_{10}$ P-value on the x-axis.
649
650

651 **Table 1. Data collection and refinement statistics for cryo-EM datasets, related to Figures**
652 **4 and S7**

	CRBN-DDB1 FPFT-2216 PPIL4 RRM	**CRBN-DDB1ΔB FPFT-2216 PDE6D Consensus refine**	**CRBN-DDB1ΔB FPFT-2216 PDE6D Local refine**
Microscope	Talos Arctica	Talos Arctica	
Voltage (kV)	200	200	
Camera	Gatan K3	Gatan K3	
Magnification (×)	36,000	36,000	
Pixel size (Å)	1.1	1.1	
Total electron exposure (e^-/Å^2)	54	51.2	
Number of frames	40	50	
Defocus range (μm)	−2.0 to −0.8	−2.0 to −0.8	
Data collection software	SerialEM4.1b		
Micrographs collected	4,170	4,474	
Total extracted particles	3,198,055	3,393,589	
EMDB accession code	EMD-XXXXX	EMD-XXXXX	EMD-XXXXX
PDB accession code	PDB-XXXX	PDB-XXXX	
Final particles used	219,802	515,636	515,636
Map resolution (Å, FSC 0.143)	3.5	3.3	3.4
FSC threshold	0.143	0.143	0.143
Model composition			
Protein residues	1211	1260	
Ligands	2	2	
Refinement package	phenix.real_space_refine		
Model-to-map CC	0.72	0.78	
Model-to-map FSC (Å, FSC 0.5)	3.8	3.5	
Mean *B* factors (Å^2)			
Protein	63.69	98.29	
Ligand	69.37	118.85	
Water	–	–	
Bond root-mean-square deviation (RMSD)			
Lengths (Å)	0.005	0.005	

	CRBN-DDB1 FPFT-2216 PPIL4 RRM	CRBN-DDB1ΔB FPFT-2216 PDE6D Consensus refine	CRBN-DDB1ΔB FPFT-2216 PDE6D Local refine
Angles (°)	0.615	0.797	
Validation			
MolProbity score	2.05	1.53	
Clash score	11.78	6.37	
Rotamer outliers (%)	1.27	0.11	
CaBLAM outliers (%)	2.76	1.54	
Ramachandran plot			
Favored (%)	94.26	96.96	
Allowed (%)	5.74	3.04	
Disallowed (%)	0.00	0.00	

656 **RESOURCES AVAILABILITY**

657 **Lead Contact**

658 Further information and requests for resources and reagents should be directed to and will be

659 fulfilled by the Lead Contact, Eric Fischer (Eric_Fischer@DFCI.HARVARD.EDU).

660

661 **Materials Availability**

662 Small molecules described in this study will be made available on request, upon completion of a

663 Materials Transfer Agreement.

664

665 **Supplemental information.**

666 **Table S1.** Table reporting Log2 Fold Change and P-value for all proteins quantified in CRBN-
667 DDB1ΔB IP-MS experiments in the presence of pomalidomide (1 μM) or SB1-G-187 (1 μM) in
668 MOLT4 or Kelly cell lysate. Statistics generated through moderated t-test in limma package[80].
669 Table related to Figures 1 and S1.

670 **Table S2.** Table reporting Log2 Fold Change and P-value for all proteins quantified in CRBN-
671 DDB1ΔB IP-MS experiments in the presence of IMiD analogs in MOLT4 cell lysate. Statistics
672 generated through moderated t-test in limma package[80]. Table related to Figure 2-4 and S3.

673 **Table S3.** Table reporting Log2 Fold Change and P-value for all proteins quantified in CRBN-
674 DDB1ΔB IP-MS experiments in the presence of IMiD analogs in Kelly cell lysate. Statistics
675 generated through moderated t-test in limma package[80]. Table related to Figure 2-4 and S4.

676 **Table S4.** Table reporting frequency of "hit" for all proteins identified as a hit across the IMiD
677 library IP-MS screens in MOLT4 and Kelly cell lysates. Table reports target
678 domain/family/superfamily classifications and interactome information. Table related to Figures
679 2, 4 and S2.

680 **Table S5.** Table reporting the G-loop details for all proteins in the human proteome that were
681 identified as having a domain carrying a G-loop. Table reports IP and global proteomics data for
682 PPIL4 studies. Table related to Figures 3-5, S5-6.

683
684

SUPPLEMENTAL FIGURES AND LEGENDS

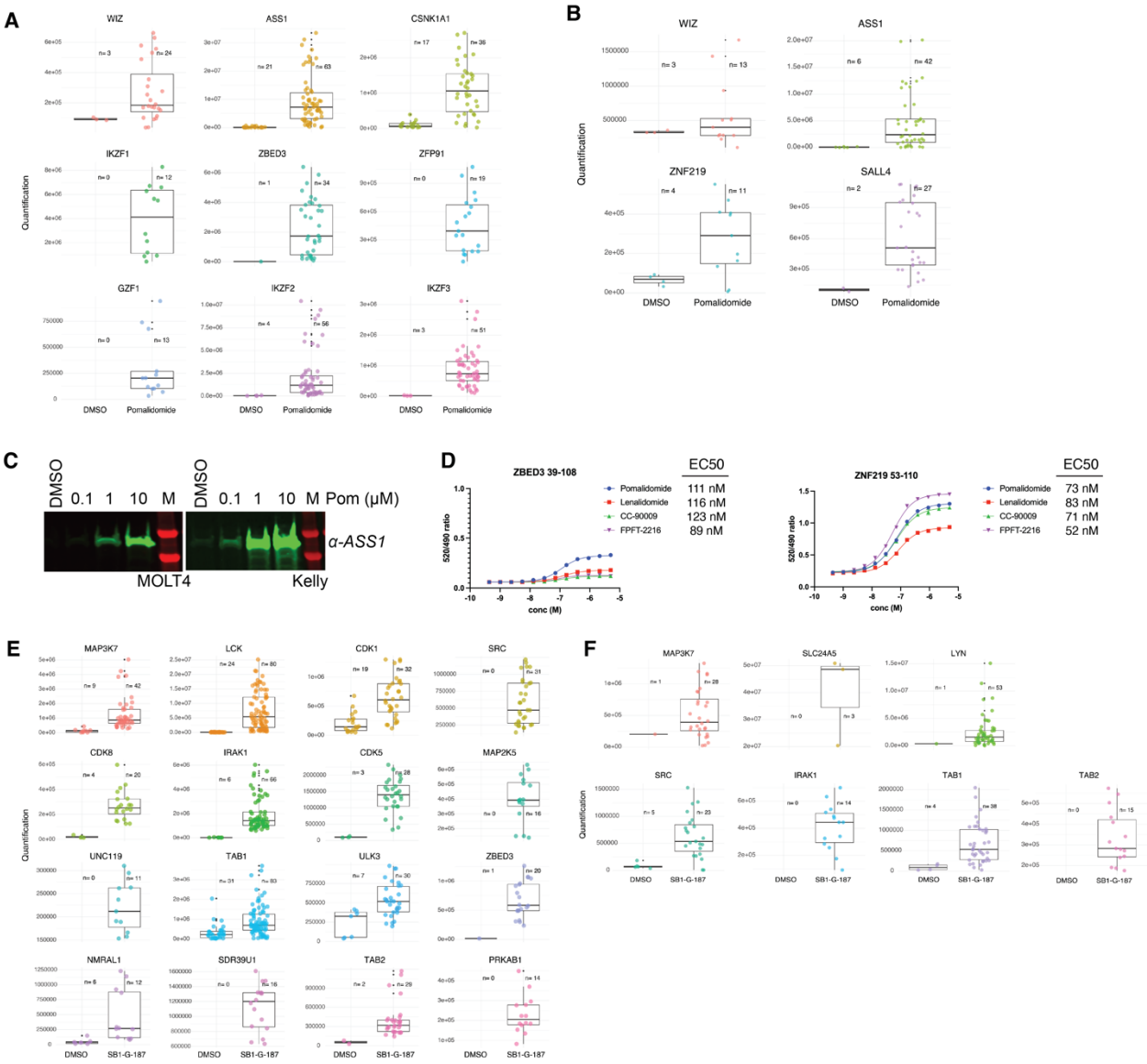


Figure S1. Validation of enriched targets from lysate IPs, related to Figure 1, Table S1

(**A**) Box plot depicting all quantified peptides for each of the enriched targets from MOLT4 cells comparing DMSO control and Pomalidomide treatments. (**B**) As in **A** but for Kelly cells. (**C**) Flag-CRBN IP experiments were performed in the presence of increasing concentration of pomalidomide in both MOLT4 and Kelly cells. Following elution, ASS1 protein levels were assessed by immunoblot. (**D**) TR-FRET: titration of IMiD analogs to GFP-CRBN-DDB1ΔB at 200 nM, $ZBED3_{39\text{-}108}$ or $ZNF219_{53\text{-}110}$ at 20 nM, and terbium-streptavidin at 2 nM. Values were determined by technical replicates of n=2. (**E**) Box plot depicting all quantified peptides for each

695 of the enriched targets from MOLT4 cells comparing DMSO control and SB1-G-187 treatments.
696 (**F**) As in **E** but for Kelly cells.
697

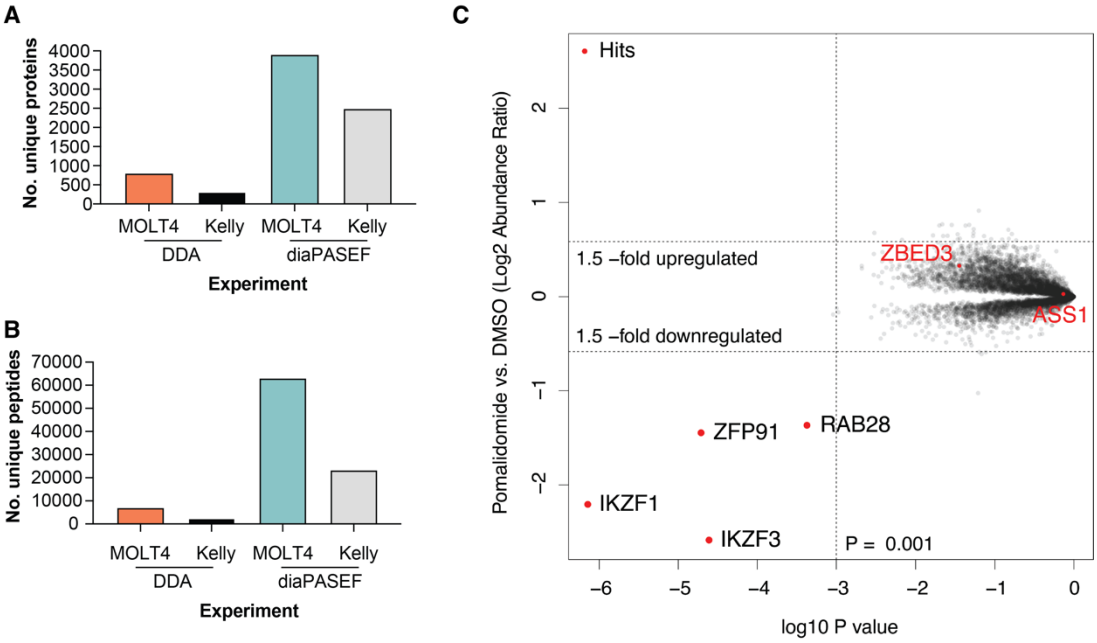


Figure S2. Quantitative proteomics for exploration of targets recruited by IMiD analogs, related to Figure 2

(**A**) Histogram displaying number of unique proteins quantified in DDA and diaPASEF IP-MS experiments for MOLT4 and Kelly cells. (**B**) Histogram displaying number of unique peptides quantified in DDA and diaPASEF IP-MS experiments for MOLT4 and Kelly cells. (**C**) Scatterplot depicting relative protein abundance following 5 μM Pomalidomide treatment in MOLT4 cells. Significant changes were assessed by moderated t-test as implemented in the limma package[80] with $\log_2$ FC shown on the y-axis and negative $\log_{10}$ P-value on the x-axis.

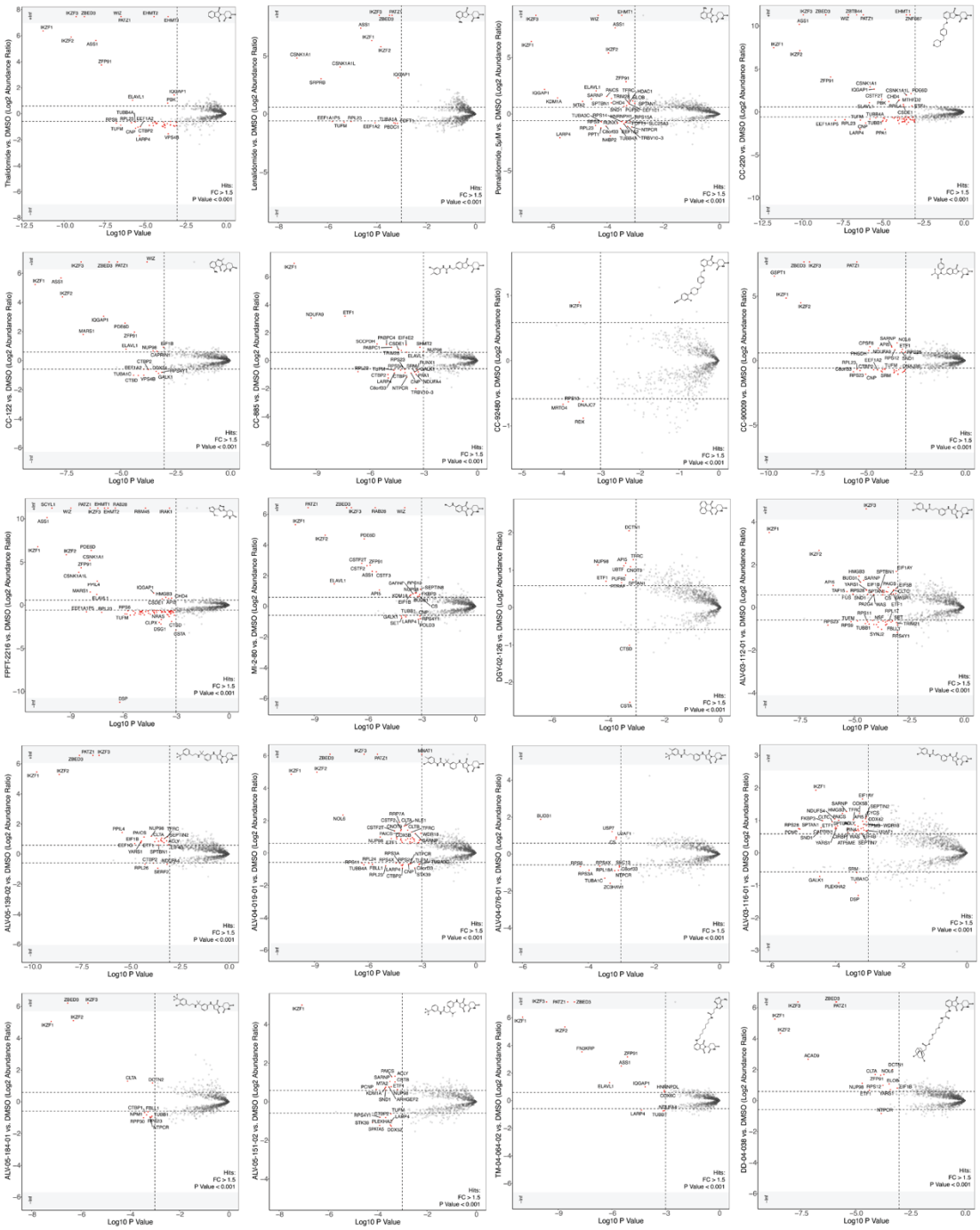


Figure S3. Scatterplots of IP-MS with molecular glues, related to Figure 2, 3, S4, Table S2, S3, S4

Scatterplots depicting relative protein abundance following Flag-CRBN-DDB1 enrichment from in-lysate treatment with degrader and recombinant Flag-CRBN-DDB1 spike in. Scatterplot displays fold change in abundance for each of the 20 molecules relative to DMSO in MOLT4

715 cells. Significant changes were assessed by moderated t-test as implemented in the limma
716 package[80] with $\log_2$ FC shown on the y-axis and negative $\log_{10}$ P-value on the x-axis.
717

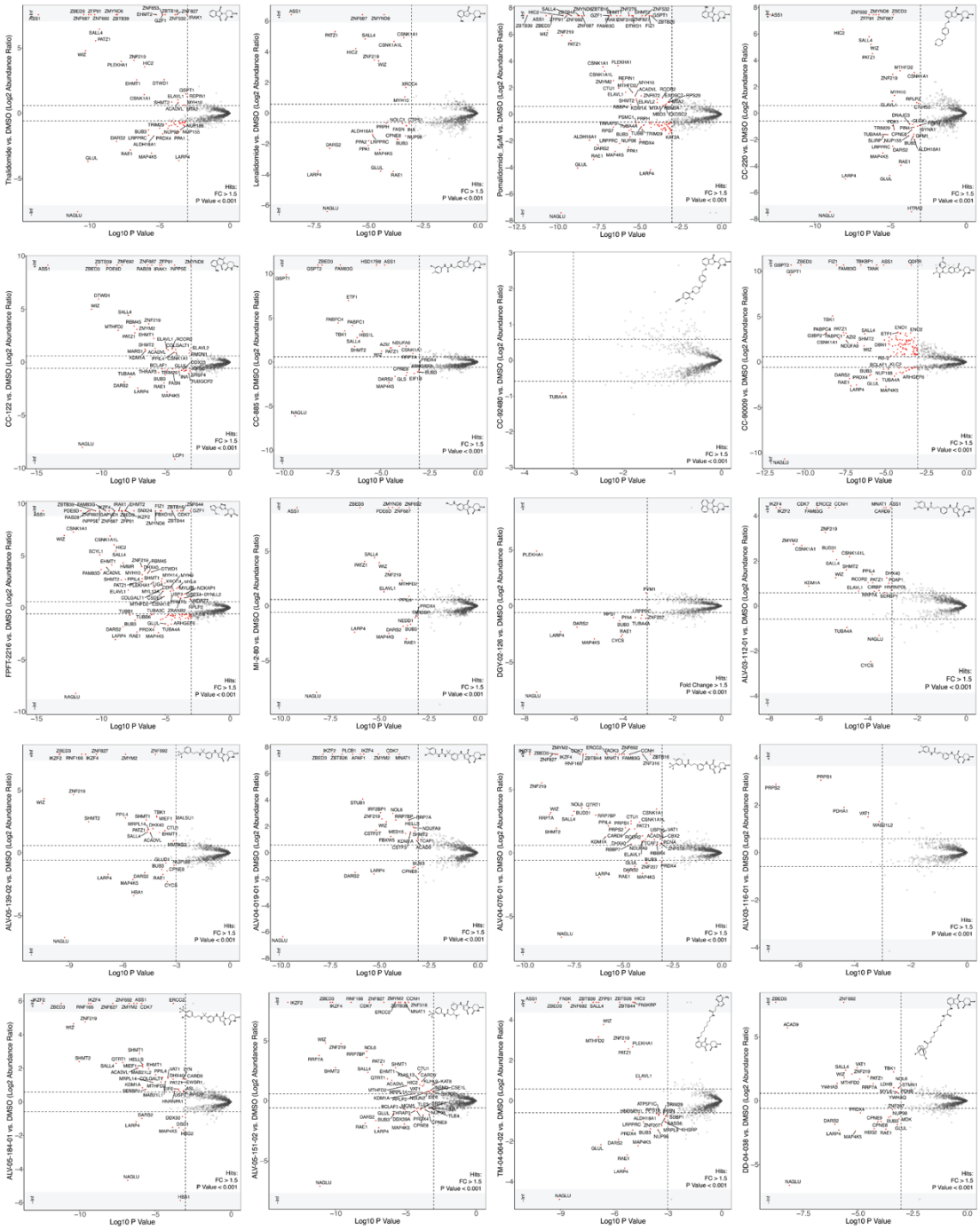


Figure S4. Scatterplots of IP-MS with molecular glues, related to Figure 2, 3, S4, Table S2, S3, S4

Scatterplots depicting relative protein abundance following Flag-CRBN-DDB1 enrichment from in-lysate treatment with degrader and recombinant Flag-CRBN-DDB1 spike in. Scatterplot displays fold change in abundance for each of the 20 molecules relative to DMSO in Kelly cells.

725 Significant changes were assessed by moderated t-test as implemented in the limma package[80]
726 with $\log_2$ FC shown on the y-axis and negative $\log_{10}$ P-value on the x-axis.
727

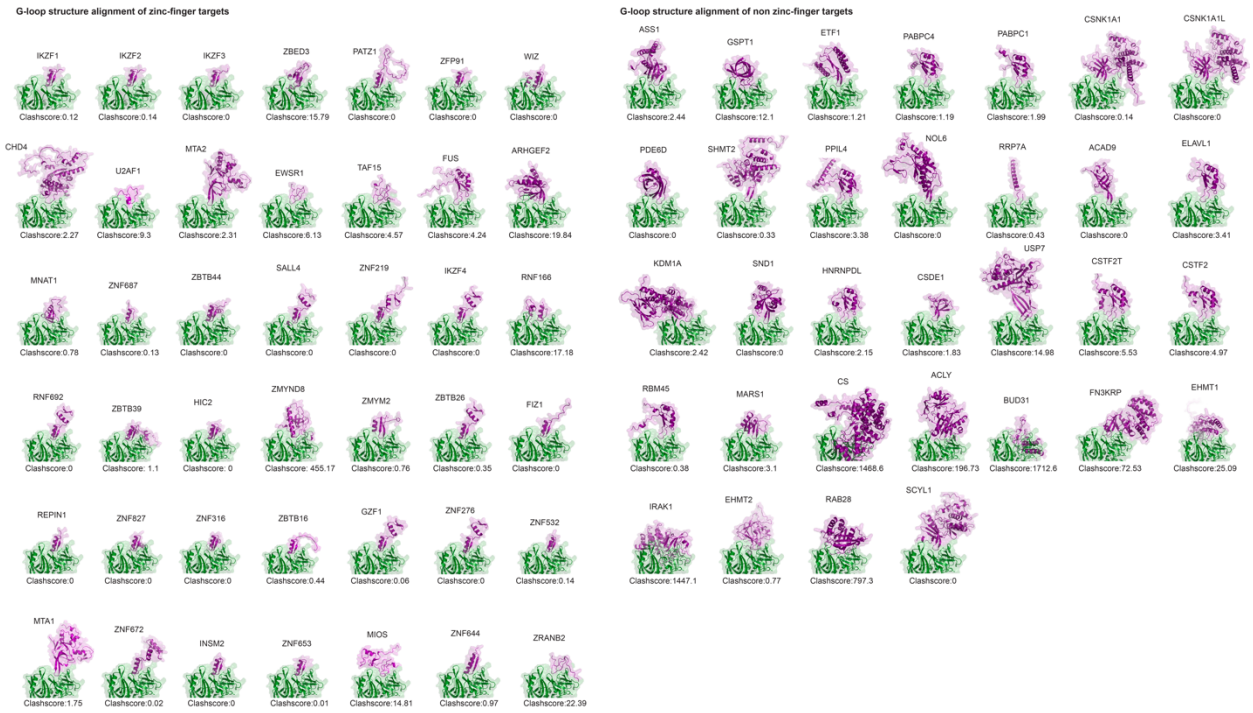


Figure S5. Structural alignments of G-loop neo-substrates, related to Figure 3, 4, S5, Table S5

Structural G-loop alignment showing alignment of AF2 structures by the G-loop for a subset of the neo-substrate targets identified as hits in this study with CRBN (PDB ID: 5FQD). Corresponding clash scores are displayed.

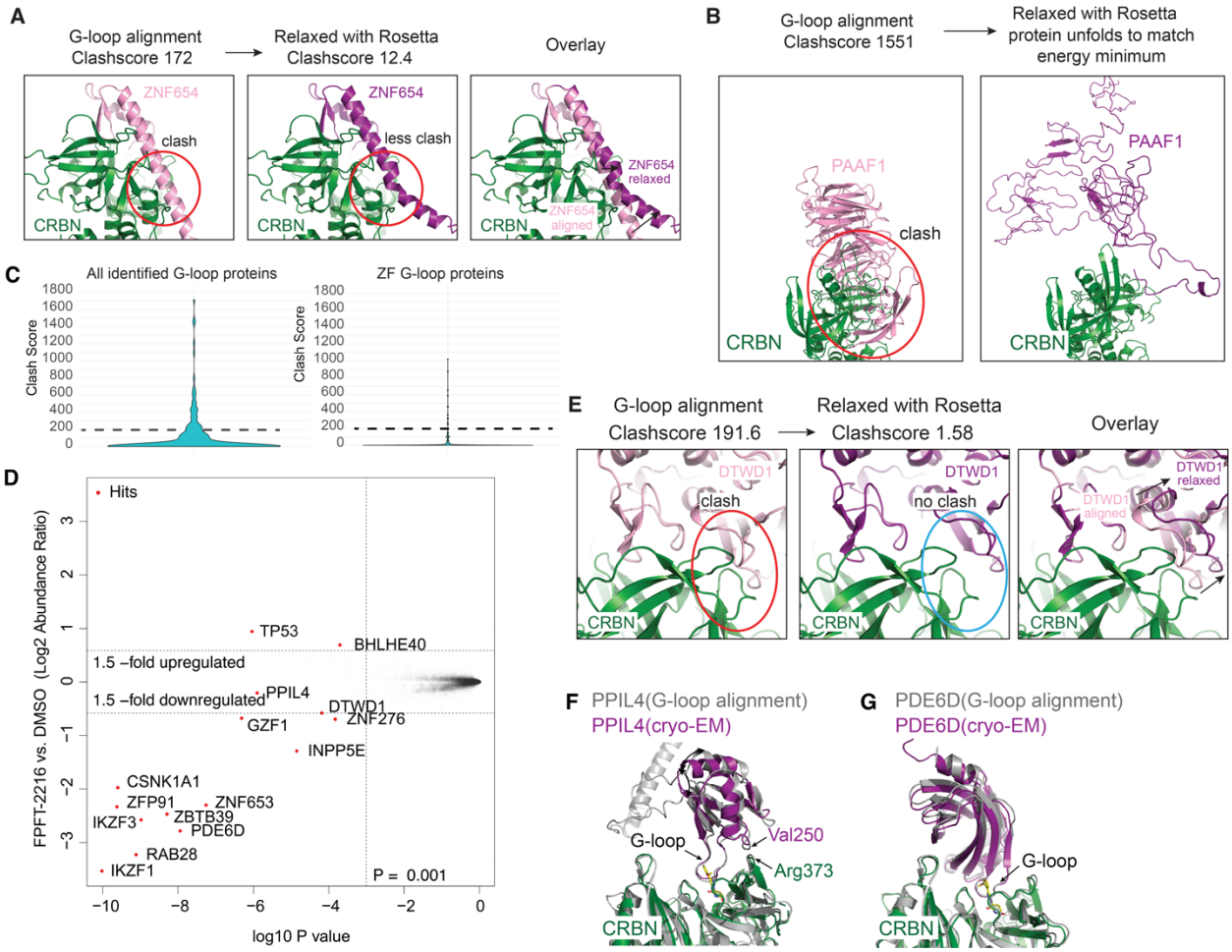


Figure S6. Resolving clashes from G-loop alignments, related to Figure 3, 4, S5, Table S5 (**A**) G-loop aligned AF2 structures for ZNF654 demonstrating the clash score and structural shift before and after relaxation with Rosetta. (**B**) As in **A**, but for PAAF1. (**C**) Violin plot depicting the distribution of class scores for a single G-loop from each of the identified targets in this dataset (left), and those for ZF proteins only (right). (**D**) Scatterplot depicting relative protein abundance following FPFT-2216 treatment in MOLT4 cells. Significant changes were assessed by moderated t-test as implemented in the limma package[80] with $\log_2$ FC shown on the y-axis and negative $\log_{10}$ P-value on the x-axis. (**E**) As in **A**, but for DTWD1. (**F**) Overlay of G-loop aligned AF2 structure of PPIL4 and structure obtained by cryo-EM. (**G**) Same as in (**F**) but with PDE6D.

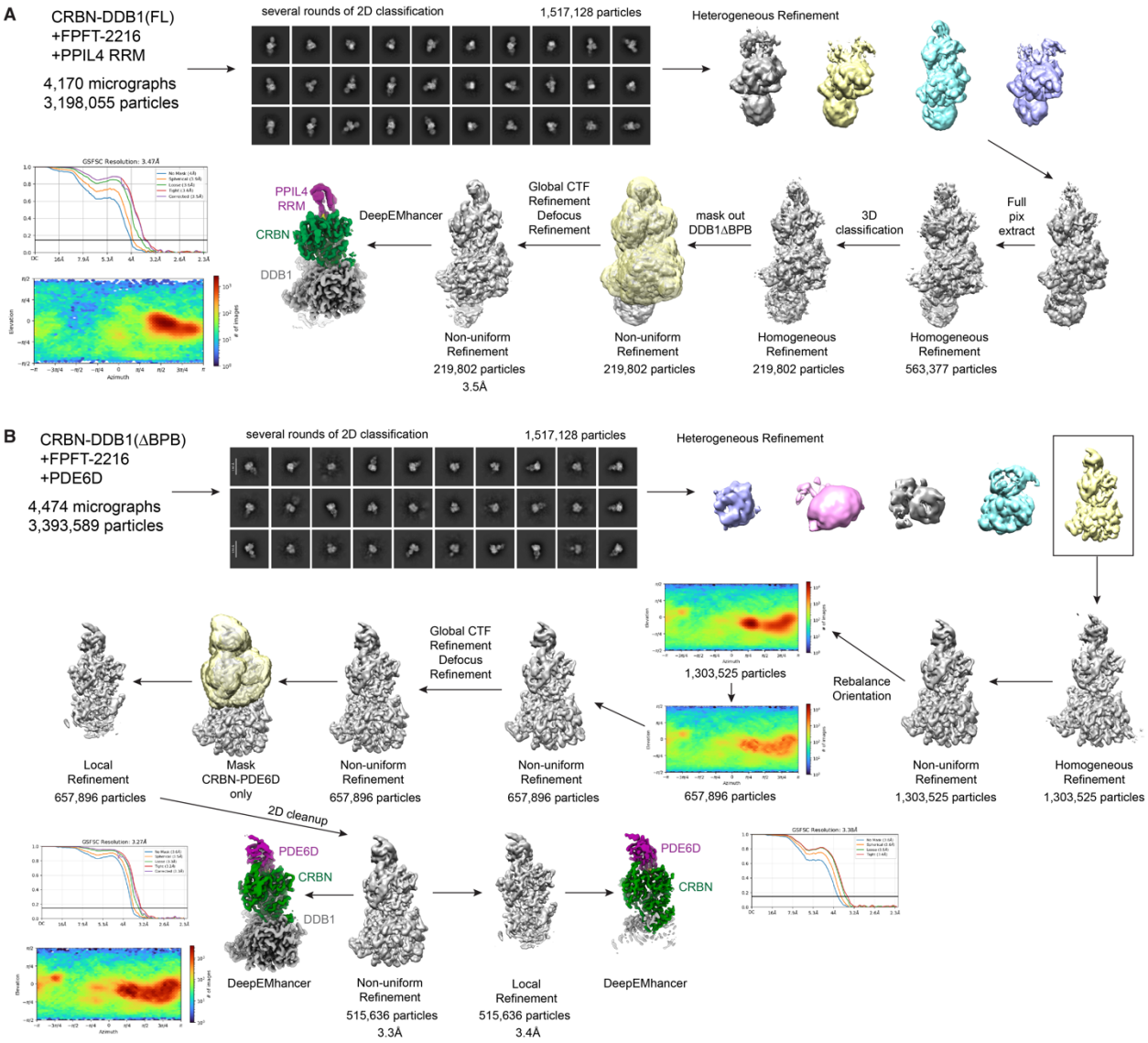


Figure S7. Cryo-EM processing schematic of neosubstrate-molecular glue-CRBN-DDB1, related to Figure 4, Table 1

(**A**) Cryo-EM processing schematic of CRBN-DDB1 (FL) with PPIL4 RRM domain and FPFT-2216. **(B)** Cryo-EM processing schematic of CRBN-DDB1ΔB with PDE6D and FPFT-2216.

754
755 **REFERENCES**
756

757 1. Bondeson, D.P., and Crews, C.M. (2017). Targeted protein degradation by small
758 molecules. Annual review of pharmacology and toxicology *57*, 107-123.
759 2. Neklesa, T.K., Winkler, J.D., and Crews, C.M. (2017). Targeted protein degradation by
760 PROTACs. Pharmacology & therapeutics *174*, 138-144.
761 10.1016/j.pharmthera.2017.02.027.
762 3. Sakamoto, K.M., Kim, K.B., Kumagai, A., Mercurio, F., Crews, C.M., and Deshaies, R.J.
763 (2001). Protacs: chimeric molecules that target proteins to the Skp1–Cullin–F box
764 complex for ubiquitination and degradation. Proceedings of the National Academy of
765 Sciences *98*, 8554-8559.
766 4. Komander, D., and Rape, M. (2012). The ubiquitin code. Annual review of biochemistry
767 *81*, 203-229.
768 5. Kronke, J., Udeshi, N.D., Narla, A., Grauman, P., Hurst, S.N., McConkey, M., Svinkina, T.,
769 Heckl, D., Comer, E., Li, X., et al. (2014). Lenalidomide causes selective degradation of
770 IKZF1 and IKZF3 in multiple myeloma cells. Science *343*, 301-305.
771 10.1126/science.1244851.
772 6. Winter, G.E., Buckley, D.L., Paulk, J., Roberts, J.M., Souza, A., Dhe-Paganon, S., and
773 Bradner, J.E. (2015). Phthalimide conjugation as a strategy for in vivo target protein
774 degradation. Science *348*, 1376-1381.
775 7. Bondeson, D.P., Mares, A., Smith, I.E., Ko, E., Campos, S., Miah, A.H., Mulholland, K.E.,
776 Routly, N., Buckley, D.L., and Gustafson, J.L. (2015). Catalytic in vivo protein knockdown
777 by small-molecule PROTACs. Nature chemical biology *11*, 611.
778 8. Wang, B., Cao, S., and Zheng, N. (2024). Emerging strategies for prospective discovery of
779 molecular glue degraders. Current Opinion in Structural Biology *86*, 102811.
780 9. Ito, T., Ando, H., Suzuki, T., Ogura, T., Hotta, K., Imamura, Y., Yamaguchi, Y., and Handa,
781 H. (2010). Identification of a primary target of thalidomide teratogenicity. Science *327*,
782 1345-1350. 10.1126/science.1177319.
783 10. Fischer, E.S., Bohm, K., Lydeard, J.R., Yang, H., Stadler, M.B., Cavadini, S., Nagel, J.,
784 Serluca, F., Acker, V., Lingaraju, G.M., et al. (2014). Structure of the DDB1-CRBN E3
785 ubiquitin ligase in complex with thalidomide. Nature *512*, 49-53. 10.1038/nature13527.
786 11. Matyskiela, M.E., Lu, G., Ito, T., Pagarigan, B., Lu, C.C., Miller, K., Fang, W., Wang, N.Y.,
787 Nguyen, D., Houston, J., et al. (2016). A novel cereblon modulator recruits GSPT1 to the
788 CRL4(CRBN) ubiquitin ligase. Nature *535*, 252-257. 10.1038/nature18611.
789 12. Petzold, G., Fischer, E.S., and Thoma, N.H. (2016). Structural basis of lenalidomide-
790 induced CK1alpha degradation by the CRL4(CRBN) ubiquitin ligase. Nature *532*, 127-130.
791 10.1038/nature16979.
792 13. Lu, G., Middleton, R.E., Sun, H., Naniong, M., Ott, C.J., Mitsiades, C.S., Wong, K.K.,
793 Bradner, J.E., and Kaelin, W.G., Jr. (2014). The myeloma drug lenalidomide promotes the
794 cereblon-dependent destruction of Ikaros proteins. Science *343*, 305-309.
795 10.1126/science.1244917.

796 14. Oleinikovas, V., Gainza, P., Ryckmans, T., Fasching, B., and Thomä, N.H. (2024). From
797 thalidomide to rational molecular glue design for targeted protein degradation. Annual
798 Review of Pharmacology and Toxicology *64*, 291-312.
799 15. Matyskiela, M.E., Couto, S., Zheng, X., Lu, G., Hui, J., Stamp, K., Drew, C., Ren, Y., Wang,
800 M., Carpenter, A., et al. (2018). SALL4 mediates teratogenicity as a thalidomide-
801 dependent cereblon substrate. Nature chemical biology *14*, 981-987. 10.1038/s41589-
802 018-0129-x.
803 16. Sievers, Q.L., Gasser, J.A., Cowley, G.S., Fischer, E.S., and Ebert, B.L. (2018). Genome-
804 wide screen identifies cullin-RING ligase machinery required for lenalidomide-
805 dependent CRL4(CRBN) activity. Blood *132*, 1293-1303. 10.1182/blood-2018-01-821769.
806 17. Kronke, J., Fink, E.C., Hollenbach, P.W., MacBeth, K.J., Hurst, S.N., Udeshi, N.D.,
807 Chamberlain, P.P., Mani, D.R., Man, H.W., Gandhi, A.K., et al. (2015). Lenalidomide
808 induces ubiquitination and degradation of CK1alpha in del(5q) MDS. Nature *523*, 183-
809 188. 10.1038/nature14610.
810 18. An, J., Ponthier, C.M., Sack, R., Seebacher, J., Stadler, M.B., Donovan, K.A., and Fischer,
811 E.S. (2017). pSILAC mass spectrometry reveals ZFP91 as IMiD-dependent substrate of
812 the CRL4(CRBN) ubiquitin ligase. Nature communications *8*, 15398.
813 10.1038/ncomms15398.
814 19. Donovan, K.A., An, J., Nowak, R.P., Yuan, J.C., Fink, E.C., Berry, B.C., Ebert, B.L., and
815 Fischer, E.S. (2018). Thalidomide promotes degradation of SALL4, a transcription factor
816 implicated in Duane Radial Ray syndrome. Elife *7*. 10.7554/eLife.38430.
817 20. Yamamoto, J., Suwa, T., Murase, Y., Tateno, S., Mizutome, H., Asatsuma-Okumura, T.,
818 Shimizu, N., Kishi, T., Momose, S., and Kizaki, M. (2020). ARID2 is a pomalidomide-
819 dependent CRL4CRBN substrate in multiple myeloma cells. Nature chemical biology *16*,
820 1208-1217.
821 21. Renneville, A., Gasser, J.A., Grinshpun, D.E., Jean Beltran, P.M., Udeshi, N.D., Matyskiela,
822 M.E., Clayton, T., McConkey, M., Viswanathan, K., and Tepper, A. (2021). Avadomide
823 induces degradation of ZMYM2 fusion oncoproteins in hematologic malignancies. Blood
824 cancer discovery *2*, 250-265.
825 22. Li, L., Xue, W., Shen, Z., Liu, J., Hu, M., Cheng, Z., Wang, Y., Chen, Y., Chang, H., and Liu,
826 Y. (2020). A cereblon modulator CC-885 induces CRBN-and p97-dependent PLK1
827 degradation and synergizes with volasertib to suppress lung cancer. Molecular Therapy-
828 Oncolytics *18*, 215-225.
829 23. Huang, H.T., Dobrovolsky, D., Paulk, J., Yang, G., Weisberg, E.L., Doctor, Z.M., Buckley,
830 D.L., Cho, J.H., Ko, E., Jang, J., et al. (2018). A Chemoproteomic Approach to Query the
831 Degradable Kinome Using a Multi-kinase Degrader. Cell chemical biology *25*, 88-99.e86.
832 10.1016/j.chembiol.2017.10.005.
833 24. Sathe, G., and Sapkota, G.P. (2023). Proteomic approaches advancing targeted protein
834 degradation. Trends in pharmacological sciences *44*, 786-801.
835 25. King, E.A., Cho, Y., Hsu, N.S., Dovala, D., McKenna, J.M., Tallarico, J.A., Schirle, M., and
836 Nomura, D.K. (2023). Chemoproteomics-enabled discovery of a covalent molecular glue
837 degrader targeting NF-κB. Cell chemical biology *30*, 394-402. e399.

838 26. Donovan, K.A., Ferguson, F.M., Bushman, J.W., Eleuteri, N.A., Bhunia, D., Ryu, S., Tan, L.,
839 Shi, K., Yue, H., and Liu, X. (2020). Mapping the Degradable Kinome Provides a Resource
840 for Expedited Degrader Development. Cell.
841 27. Bondeson, D.P., Smith, B.E., Burslem, G.M., Buhimschi, A.D., Hines, J., Jaime-Figueroa, S.,
842 Wang, J., Hamman, B.D., Ishchenko, A., and Crews, C.M. (2018). Lessons in PROTAC
843 Design from Selective Degradation with a Promiscuous Warhead. Cell chemical biology
844 *25*, 78-87.e75. 10.1016/j.chembiol.2017.09.010.
845 28. Nowak, R.P., DeAngelo, S.L., Buckley, D., He, Z., Donovan, K.A., An, J., Safaee, N.,
846 Jedrychowski, M.P., Ponthier, C.M., Ishoey, M., et al. (2018). Plasticity in binding confers
847 selectivity in ligand-induced protein degradation. Nature chemical biology *14*, 706-714.
848 10.1038/s41589-018-0055-y.
849 29. Li, Y.-D., Ma, M.W., Hassan, M.M., Hunkeler, M., Teng, M., Puvar, K., Lumpkin, R.,
850 Sandoval, B., Jin, C.Y., and Ficarro, S.B. (2023). Template-assisted covalent modification
851 of DCAF16 underlies activity of BRD4 molecular glue degraders. bioRxiv, 2023.2002.
852 2014.528208.
853 30. Hsia, O., Hinterndorfer, M., Cowan, A.D., Iso, K., Ishida, T., Sundaramoorthy, R.,
854 Nakasone, M.A., Imrichova, H., Schätz, C., and Rukavina, A. (2024). Targeted protein
855 degradation via intramolecular bivalent glues. Nature *627*, 204-211.
856 31. Xiong, Y., Donovan, K.A., Eleuteri, N.A., Kirmani, N., Yue, H., Razov, A., Krupnick, N.M.,
857 Nowak, R.P., and Fischer, E.S. (2021). Chemo-proteomics exploration of HDAC
858 degradability by small molecule degraders. Cell chemical biology *28*, 1514-1527. e1514.
859 32. Huang, H.-T., Lumpkin, R.J., Tsai, R.W., Su, S., Zhao, X., Xiong, Y., Chen, J., Mageed, N.,
860 Donovan, K.A., and Fischer, E.S. (2024). Ubiquitin-specific proximity labeling for the
861 identification of E3 ligase substrates. Nature Chemical Biology, 1-10.
862 33. Huttlin, E.L., Ting, L., Bruckner, R.J., Gebreab, F., Gygi, M.P., Szpyt, J., Tam, S., Zarraga,
863 G., Colby, G., and Baltier, K. (2015). The BioPlex network: a systematic exploration of the
864 human interactome. Cell *162*, 425-440.
865 34. Cho, K.F., Branon, T.C., Udeshi, N.D., Myers, S.A., Carr, S.A., and Ting, A.Y. (2020).
866 Proximity labeling in mammalian cells with TurboID and split-TurboID. Nature Protocols
867 *15*, 3971-3999.
868 35. Roux, K.J., Kim, D.I., Burke, B., and May, D.G. (2018). BioID: a screen for protein-protein
869 interactions. Current protocols in protein science *91*, 19.23. 11-19.23. 15.
870 36. Kim, D.I., KC, B., Zhu, W., Motamedchaboki, K., Doye, V., and Roux, K.J. (2014). Probing
871 nuclear pore complex architecture with proximity-dependent biotinylation. Proceedings
872 of the National Academy of Sciences *111*, E2453-E2461.
873 37. Kim, D.I., Jensen, S.C., Noble, K.A., Kc, B., Roux, K.H., Motamedchaboki, K., and Roux, K.J.
874 (2016). An improved smaller biotin ligase for BioID proximity labeling. Molecular biology
875 of the cell *27*, 1188-1196.
876 38. Kido, K., Yamanaka, S., Nakano, S., Motani, K., Shinohara, S., Nozawa, A., Kosako, H., Ito,
877 S., and Sawasaki, T. (2020). AirID, a novel proximity biotinylation enzyme, for analysis of
878 protein–protein interactions. Elife *9*, e54983.
879 39. Martell, J.D., Deerinck, T.J., Sancak, Y., Poulos, T.L., Mootha, V.K., Sosinsky, G.E.,
880 Ellisman, M.H., and Ting, A.Y. (2012). Engineered ascorbate peroxidase as a genetically
881 encoded reporter for electron microscopy. Nature biotechnology *30*, 1143-1148.

882 40. Sievers, Q.L., Petzold, G., Bunker, R.D., Renneville, A., Slabicki, M., Liddicoat, B.J.,
883 Abdulrahman, W., Mikkelsen, T., Ebert, B.L., and Thoma, N.H. (2018). Defining the
884 human C2H2 zinc finger degrome targeted by thalidomide analogs through CRBN.
885 Science *362*. 10.1126/science.aat0572.
886 41. Roh, Y.S., Song, J., and Seki, E. (2014). TAK1 regulates hepatic cell survival and
887 carcinogenesis. Journal of gastroenterology *49*, 185-194.
888 42. Shibuya, H., Yamaguchi, K., Shirakabe, K., Tonegawa, A., Gotoh, Y., Ueno, N., Irie, K.,
889 Nishida, E., and Matsumoto, K. (1996). TAB1: an activator of the TAK1 MAPKKK in TGF-β
890 signal transduction. Science *272*, 1179-1182.
891 43. Garivet, G., Hofer, W., Konitsiotis, A., Klein, C., Kaiser, N., Mejuch, T., Fansa, E., Alsaabi,
892 R., Wittinghofer, A., and Bastiaens, P.I. (2019). Small-molecule inhibition of the UNC-Src
893 interaction impairs dynamic Src localization in cells. Cell Chemical Biology *26*, 842-851.
894 e847.
895 44. Golkowski, M., Brigham, J.L., Perera, B.G.K., Romano, G.S., Maly, D.J., and Ong, S.-E.
896 (2014). Rapid profiling of protein kinase inhibitors by quantitative proteomics.
897 Medchemcomm *5*, 363-369.
898 45. Meier, F., Brunner, A.-D., Frank, M., Ha, A., Bludau, I., Voytik, E., Kaspar-Schoenefeld, S.,
899 Lubeck, M., Raether, O., and Bache, N. (2020). diaPASEF: parallel accumulation–serial
900 fragmentation combined with data-independent acquisition. Nature methods *17*, 1229-
901 1236.
902 46. Knight, R. (2005). IMiDs: a novel class of immunomodulators. (Elsevier), pp. 24-30.
903 47. D'Amato, R.J., Loughnan, M.S., Flynn, E., and Folkman, J. (1994). Thalidomide is an
904 inhibitor of angiogenesis. Proc Natl Acad Sci U S A *91*, 4082-4085.
905 48. Lonial, S., Popat, R., Hulin, C., Jagannath, S., Oriol, A., Richardson, P.G., Facon, T., Weisel,
906 K., Larsen, J.T., and Minnema, M.C. (2022). Iberdomide plus dexamethasone in heavily
907 pretreated late-line relapsed or refractory multiple myeloma (CC-220-MM-001): a
908 multicentre, multicohort, open-label, phase 1/2 trial. The Lancet Haematology *9*, e822-
909 e832.
910 49. Richardson, P.G., Trudel, S., Popat, R., Mateos, M.-V., Vangsted, A.J., Ramasamy, K.,
911 Martinez-Lopez, J., Quach, H., Orlowski, R.Z., and Arnao, M. (2023). Mezigdomide plus
912 dexamethasone in relapsed and refractory multiple myeloma. New England Journal of
913 Medicine *389*, 1009-1022.
914 50. Surka, C., Jin, L., Mbong, N., Lu, C.-C., Jang, I.S., Rychak, E., Mendy, D., Clayton, T.,
915 Tindall, E., and Hsu, C. (2021). CC-90009, a novel cereblon E3 ligase modulator, targets
916 acute myeloid leukemia blasts and leukemia stem cells. Blood, The Journal of the
917 American Society of Hematology *137*, 661-677.
918 51. Gemechu, Y., Millrine, D., Hashimoto, S., Prakash, J., Sanchenkova, K., Metwally, H.,
919 Gyanu, P., Kang, S., and Kishimoto, T. (2018). Humanized cereblon mice revealed two
920 distinct therapeutic pathways of immunomodulatory drugs. Proc Natl Acad Sci U S A
921 *115*, 11802-11807. 10.1073/pnas.1814446115.
922 52. Hagner, P.R., Man, H.-W., Fontanillo, C., Wang, M., Couto, S., Breider, M., Bjorklund, C.,
923 Havens, C.G., Lu, G., and Rychak, E. (2015). CC-122, a pleiotropic pathway modifier,
924 mimics an interferon response and has antitumor activity in DLBCL. Blood, The Journal
925 of the American Society of Hematology *126*, 779-789.

926 53. Wang, E.S., Verano, A.L., Nowak, R.P., Yuan, J.C., Donovan, K.A., Eleuteri, N.A., Yue, H.,
927 Ngo, K.H., Lizotte, P.H., and Gokhale, P.C. (2021). Acute pharmacological degradation of
928 Helios destabilizes regulatory T cells. Nature chemical biology *17*, 711-717.
929 54. Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen,
930 J., Mitchell, A., and Nuka, G. (2014). InterProScan 5: genome-scale protein function
931 classification. Bioinformatics *30*, 1236-1240.
932 55. Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B.L., Salazar, G.A.,
933 Bileschi, M.L., Bork, P., Bridge, A., and Colwell, L. (2023). InterPro in 2022. Nucleic acids
934 research *51*, D418-D427.
935 56. UniProt: the universal protein knowledgebase in 2023. (2023). Nucleic acids research
936 *51*, D523-D531.
937 57. Gough, J., Karplus, K., Hughey, R., and Chothia, C. (2001). Assignment of homology to
938 genome sequences using a library of hidden Markov models that represent all proteins
939 of known structure. Journal of molecular biology *313*, 903-919.
940 58. Razumkov, H., Jiang, Z., Baek, K., You, I., Geng, Q., Donovan, K., Tang, M.T., Metivier,
941 R.J., Mageed, N., and Seo, P. (2024). Discovery of CRBN-dependent WEE1 Molecular
942 Glue Degraders from a Multicomponent Combinatorial Library. bioRxiv, 2024.2005.
943 2004.592550.
944 59. Zhao, M., Hu, M., Chen, Y., Liu, H., Chen, Y., Liu, B., and Fang, B. (2021). Cereblon
945 modulator CC-885 induces CRBN-dependent ubiquitination and degradation of CDK4 in
946 multiple myeloma. Biochemical and biophysical research communications *549*, 150-156.
947 60. Furihata, H., Yamanaka, S., Honda, T., Miyauchi, Y., Asano, A., Shibata, N., Tanokura, M.,
948 Sawasaki, T., and Miyakawa, T. (2020). Structural bases of IMiD selectivity that emerges
949 by 5-hydroxythalidomide. Nature communications *11*, 4578.
950 61. Zhou, J., and Grigoryan, G. (2015). Rapid search for tertiary fragments reveals protein
951 sequence–structure relationships. Protein Science *24*, 508-524.
952 62. Varadi, M., Bertoni, D., Magana, P., Paramval, U., Pidruchna, I., Radhakrishnan, M.,
953 Tsenkov, M., Nair, S., Mirdita, M., and Yeo, J. (2024). AlphaFold Protein Structure
954 Database in 2024: providing structure coverage for over 214 million protein sequences.
955 Nucleic acids research *52*, D368-D375.
956 63. Zhang, J., Schaeffer, R.D., Durham, J., Cong, Q., and Grishin, N.V. (2023). DPAM: A
957 domain parser for AlphaFold models. Protein Science *32*, e4548.
958 64. Leman, J.K., Weitzner, B.D., Lewis, S.M., Adolf-Bryfogle, J., Alam, N., Alford, R.F.,
959 Aprahamian, M., Baker, D., Barlow, K.A., and Barth, P. (2020). Macromolecular modeling
960 and design in Rosetta: recent methods and frameworks. Nature methods *17*, 665-680.
961 65. Maguire, J.B., Haddox, H.K., Strickland, D., Halabiya, S.F., Coventry, B., Griffin, J.R.,
962 Pulavarti, S.V.K., Cummins, M., Thieker, D.F., and Klavins, E. (2021). Perturbing the
963 energy landscape for improved packing during computational protein design. Proteins:
964 Structure, Function, and Bioinformatics *89*, 436-449.
965 66. Rodriguez-Rivera, F.P., and Levi, S.M. (2021). Unifying catalysis framework to dissect
966 proteasomal degradation paradigms. ACS Central Science *7*, 1117-1125.
967 67. Barroso-Gomila, O., Merino-Cacho, L., Muratore, V., Perez, C., Taibi, V., Maspero, E.,
968 Azkargorta, M., Iloro, I., Trulsson, F., and Vertegaal, A.C. (2023). BioE3 identifies specific
969 substrates of ubiquitin E3 ligases. Nature Communications *14*, 7656.

970 68. Mukhopadhyay, U., Levantovsky, S., Carusone, T.M., Gharbi, S., Stein, F., Behrends, C.,
971 and Bhogaraju, S. (2024). A ubiquitin-specific, proximity-based labeling approach for the
972 identification of ubiquitin ligase substrates. Science Advances *10*, eadp3000.
973 69. Liwocha, J., Li, J., Purser, N., Rattanasopa, C., Maiwald, S., Krist, D.T., Scott, D.C.,
974 Steigenberger, B., Prabu, J.R., and Schulman, B.A. (2024). Mechanism of millisecond
975 Lys48-linked poly-ubiquitin chain formation by cullin-RING ligases. Nature Structural &
976 Molecular Biology *31*, 378-389.
977 70. Bai, N., Riching, K.M., Makaju, A., Wu, H., Acker, T.M., Ou, S.-C., Zhang, Y., Shen, X.,
978 Bulloch, D.N., and Rui, H. (2022). Modeling the CRL4A ligase complex to predict target
979 protein ubiquitination induced by cereblon-recruiting PROTACs. Journal of Biological
980 Chemistry *298*.
981 71. Békés, M., Langley, D.R., and Crews, C.M. (2022). PROTAC targeted protein degraders:
982 the past is prologue. Nature Reviews Drug Discovery *21*, 181-200.
983 72. Yamanaka, S., Horiuchi, Y., Matsuoka, S., Kido, K., Nishino, K., Maeno, M., Shibata, N.,
984 Kosako, H., and Sawasaki, T. (2022). A proximity biotinylation-based approach to identify
985 protein-E3 ligase interactions induced by PROTACs and molecular glues. Nature
986 Communications *13*, 183.
987 73. Yu, H.H., Reitsma, J.M., Sweredoski, M.J., Moradian, A., Hess, S., and Deshaies, R.J.
988 (2019). Single subunit degradation of WIZ, a lenalidomide-and pomalidomide-
989 dependent substrate of E3 ubiquitin ligase CRL4CRBN. BioRxiv, 595389.
990 74. Cassandri, M., Smirnov, A., Novelli, F., Pitolli, C., Agostini, M., Malewicz, M., Melino, G.,
991 and Raschellà, G. (2017). Zinc-finger proteins in health and disease. Cell death discovery
992 *3*, 1-12.
993 75. Teng, M., Lu, W., Donovan, K.A., Sun, J., Krupnick, N.M., Nowak, R.P., Li, Y.-D., Sperling,
994 A.S., Zhang, T., and Ebert, B.L. (2021). Development of PDE6D and CK1α degraders
995 through chemical derivatization of FPFT-2216. Journal of medicinal chemistry *65*, 747-
996 756.
997 76. Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A.L.,
998 Fang, T., Doncheva, N.T., and Pyysalo, S. (2023). The STRING database in 2023: protein–
999 protein association networks and functional enrichment analyses for any sequenced
1000 genome of interest. Nucleic acids research *51*, D638-D646.
1001 77. Hornbeck, P.V., Zhang, B., Murray, B., Kornhauser, J.M., Latham, V., and Skrzypek, E.
1002 (2015). PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. Nucleic acids
1003 research *43*, D512-D520.
1004 78. Barak, T., Ristori, E., Ercan-Sencicek, A.G., Miyagishima, D.F., Nelson-Williams, C., Dong,
1005 W., Jin, S.C., Prendergast, A., Armero, W., and Henegariu, O. (2021). PPIL4 is essential for
1006 brain angiogenesis and implicated in intracranial aneurysms in humans. Nature
1007 medicine *27*, 2165-2175.
1008 79. Schmitzová, J., Cretu, C., Dienemann, C., Urlaub, H., and Pena, V. (2023). Structural basis
1009 of catalytic activation in human splicing. Nature *617*, 842-850.
1010 80. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015).
1011 limma powers differential expression analyses for RNA-sequencing and microarray
1012 studies. Nucleic Acids Res *43*, e47. 10.1093/nar/gkv007.

1013 81. Sanchez-Garcia, R., Gomez-Blanco, J., Cuervo, A., Carazo, J.M., Sorzano, C.O.S., and
1014 Vargas, J. (2021). DeepEMhancer: a deep learning solution for cryo-EM volume post-
1015 processing. Communications biology *4*, 874.
1016