

1 **Small-molecule targeting of MUSASHI RNA-binding activity in acute myeloid leukemia**

2

3 Gerard Minuesa¹, Steven K. Albanese^{2,3}, Arthur Chow¹, Alexandra Schurer¹, Sun-Mi Park¹,
4 Christina Z. Rotsides⁴, James Taggart¹, Andrea Rizzi^{3,5}, Levi N. Naden³, Timothy Chou¹, Saroj
5 Gourkanti¹, Daniel Cappel⁶, Maria C. Passarelli^{3,7}, Lauren Fairchild^{3,8}, Carolina Adura⁹, J Fraser
6 Glickman⁹, Jessica Schulman¹⁰, Christopher Famulare¹¹, Minal Patel¹¹, Joseph K. Eibl¹², Gregory
7 M. Ross¹², Derek Tan⁴, Christina Leslie³, Thijs Beuming¹³, Yehuda Goldgur¹⁴, John D. Chodera³,
8 Michael G. Kharas^{1,*}

9

10 1. Molecular Pharmacology Program, Experimental Therapeutics Center and Center for Stem
11 Cell Biology, Memorial Sloan-Kettering Cancer Center, New York, NY, 10065, USA

12 2. Louis V. Gerstner Jr. Graduate School of Biomedical Sciences, Memorial Sloan Kettering
13 Cancer Center, New York, NY, 10065, USA

14 3. Computational and Systems Biology Program, Sloan Kettering Institute, Memorial Sloan
15 Kettering Cancer Center, New York, NY, 10065, USA

16 4. Chemical Biology Program and Tri-Institutional Research Program, Memorial Sloan Kettering
17 Cancer Center, 1275 York Avenue, Box 422, New York, New York, NY, 10065, USA

18 5. Tri-Institutional Training Program in Computational Biology and Medicine, New York, NY,
19 10065, USA

20 6. Schrödinger GmbH, Q7, 23, 68161 Mannheim, Germany

21 7. Weill Cornell Medical College, Rockefeller University & Sloan Kettering Institute Tri-Institutional
22 MD-PhD Program, New York, NY, 10065, USA

23 8. Tri-Institutional Training Program in Computational Biology and Medicine, Weill Cornell Medical
24 College, New York, New York 10065, USA.

25 9. High-Throughput Screening Resource Center, The Rockefeller University, New York, NY,
26 10065, USA

27 10. Hematologic Oncology Tissue Bank, Department of Medicine, Memorial Sloan Kettering
28 Cancer Center, New York, New York, USA

29 11. Center for Hematologic Malignancies, Memorial Sloan Kettering Cancer Center, New York,
30 NY, 10065, USA

31 12. Northern Ontario School of Medicine, Sudbury, ON, Canada

32 13. Schrödinger, Inc., 120 West 45th Street, New York, New York, NY, 10036, USA

33 14. Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY, 10065,
34 USA

35

36 *To whom correspondence should be addressed (kharasm@mskcc.org).

37 **SUMMARY**

38 The MUSASHI family of RNA binding proteins (MSI1 and MSI2) contribute to a wide spectrum of
39 cancers including acute myeloid leukemia. We found that the small molecule Ro 08-2750 (Ro)
40 directly binds to MSI2 and competes for its RNA binding in biochemical assays. Ro treatment in
41 mouse and human myeloid leukemia cells resulted in an increase in differentiation and apoptosis,
42 inhibition of known MSI-targets, and a shared global gene expression signature similar to shRNA
43 depletion of MSI2. Ro demonstrated in vivo inhibition of c-MYC and reduced disease burden in a
44 murine AML leukemia model. Thus, we have identified a small molecule that targets MSI's
45 oncogenic activity. Our study provides a framework for targeting RNA binding proteins in cancer.

46 INTRODUCTION

47 RNA-binding proteins (RBPs) play critical roles in cell homeostasis by controlling gene expression
48 post-transcriptionally. Ribonucleoprotein complexes are essential for all steps of mRNA
49 processing including splicing, polyadenylation, localization, stability, export and translation¹. The
50 contribution of RBPs to tumorigenesis (e.g. SRSF2, SF3B1, MSI and SYNCRIP), through genetic
51 perturbation or epigenetic dysregulation, has been found in a variety of human cancers²⁻⁹.
52 Deregulation of the MSI family of RBPs was initially reported in gliomas¹⁰, medulloblastomas¹¹
53 and hepatomas¹². Since then, studies in a diverse range of neoplasms have involved MSI family,
54 including aggressive forms of colorectal^{13,14}, breast^{15,16}, lung¹⁷, glioblastoma¹⁸ and pancreatic
55 cancers^{19,20} and hematological malignancies. Among these tissues, the hematopoietic system
56 has been the most well characterized to dissect MSI function. The *MSI2* gene was initially
57 reported as a translocation partner with HOXA9 in patients progressing from chronic
58 myelogenous leukemia to blast crisis (CML-BC)²¹. More recently, other rare genetic alterations
59 (involving *MSI2*, *EVI1*, *TTC40* and *PAX5* genes) have been identified in leukemia patients^{22,23,24}.
60 *MSI2* expression is detected in 70% of AML patients and it correlates with a poor clinical
61 prognosis in multiple hematological malignancies²⁵⁻²⁹. Thus, *MSI2* has been proposed as a
62 putative biomarker for diagnosis as well as a potential therapeutic target for AML²⁹.

63
64 The relevance and requirement of *MSI2* function in leukemia was demonstrated by deletion or
65 depletion of *MSI2* with a germline gene-trap knockout or shRNAs resulting in reduced
66 leukemogenesis in a CML-BC model^{25,26}, whereas forced overexpression of *MSI2* and BCR-ABL
67 or NUP98-HOXA13 leads to a more aggressive form of CML²⁶ or myelodysplastic syndromes²⁸,
68 respectively. *MSI2* is upregulated 10-fold as CML progresses to blast crisis state in patients and
69 shRNA-mediated *MSI2* silencing blocks propagation of both CML-BC and AML cell lines^{25,26}.
70 Additionally, *Msi2* was shown to be required for leukemic stem cells (LSC) in a retroviral
71 transplantation MLL-AF9 model of AML^{8,30}. We and others have found that MSI mediates its
72 function as an RNA binding protein controlling translation of its target RNAs^{8,25,30-32}.

73
74 Overall, MSI's requirement in myeloid leukemia makes it an attractive therapeutic target in
75 leukemia and in other malignancies³³. RNA-binding proteins are often considered "undrugable"
76 targets due to their lack of well-defined binding pockets for RNA and their absence of enzymatic
77 activity. Structurally, the MSI family of RBPs -comprising the MSI1 and MSI2- contain two highly
78 conserved RNA-recognition motifs (RRMs) in the N-terminal region and a Poly-A Binding Domain
79 (PABP) at the C-terminal region³⁴. It is known that RRM1 is the determinant for RNA binding
80 specificity whereas RRM2, mainly adds affinity³⁵. The minimal binding consensus described for
81 RRM1 mouse MSI1 is r(GUAG)³⁶ and it is known that MSI it also preferentially binds UAG-
82 containing sequences in human and *Drosophila*^{35,37}. Here, we describe the identification and

83 characterization of Ro 08-2750. Using biochemical and structural approaches, we find that Ro
84 binds to the MSI2 RRM1 RNA-binding site and inhibits MSI RNA-binding activity and regulation of
85 downstream oncogenic targets. Furthermore, we demonstrate that Ro 08-2750 has efficacy in
86 inhibiting leukemogenesis by *in vitro* and *in vivo* models of myeloid leukemia.

87 **RESULTS**

88

89 **Ro 08-2750 (Ro) binds to MSI2 and inhibits its RNA-binding activity**

90 In order to identify a putative MSI inhibitor, we previously performed a fluorescence polarization
91 (FP)-based screen using recombinant MSI1 and MSI2 and a consensus target RNA with a library
92 of 6,208 compounds³⁸. We selected Ro 08-2750 (Ro) based on its RNA-binding inhibition of both
93 MSI1 and MSI2³⁸. MSI2 RNA-binding inhibition was confirmed by FP (IC_{50} of $2.7 \pm 0.4 \mu\text{M}$) (**Fig.**
94 **1a**) We then used a chemiluminescent Electrophoresis Mobility Shift Assay (EMSA) to quantify
95 MSI2-RNA complexes in vitro. GST-MSI2 bound a MSI RNA, which was competed with the
96 addition of unlabeled RNA and by increasing concentrations of Ro (**Fig. 1b and 1c**). To confirm
97 the direct interaction of Ro with MSI2 protein, we performed Microscale Thermophoresis (MST)
98 assays with GST-MSI2 and found that the small-molecule interacted with a K_D of $12.3 \pm 0.5 \mu\text{M}$
99 (**Fig. 1d**). RNA-recognition motif 1 (RRM1) of MSI2 also interacted with a similar affinity
100 (**Extended Data Fig. 1a**) suggesting that the binding was localized to this domain. Similarly,
101 when incubated in the presence of GST-MSI2 and a MSI RNA oligo, Ro could still compete with
102 RNA and bind to MSI2 with a K_D of $27.5 \pm 2.6 \mu\text{M}$. We recently found that SYNCRIP, another
103 RNA binding protein, shares MSI2 target RNAs and is also required in leukemia⁴. SYNCRIP has
104 RRMs that are evolutionarily related to MSI's (with RRM1 and RRM2 sharing 33% and 57% of the
105 residues involved in RNA-binding with MSI2's RRM1 and 2, respectively) (**Extended Data Fig.**
106 **1b**) Ro showed a 19.2-fold lower K_D for SYNCRIP than for MSI2 ($236.0 \pm 167.1 \mu\text{M}$, **Fig.1d**),
107 indicating selectivity toward MSI2.

108

109 **Ro interacts with the RNA recognition site of MSI2 RRM1**

110 To study how Ro interacts with the MSI2 protein, we obtained the crystal structure of *apo* human
111 MSI2 RRM1 at 1.7Å resolution (**Extended Data Table 1**, RCSB PDB accession code 6DBP).
112 This structure allowed us to perform docking analysis to identify a putative binding mode (**Fig. 2a,**
113 **b** and **Extended Data Fig. 2a**). Based on Ro's ability to compete for MSI-RNA complexes, we
114 hypothesized that the binding site is likely to be shared with the RNA binding site. A closer look at
115 the residues involved in putative Ro binding interactions revealed F66 and R100 as crucial amino
116 acids participating in a stacking interaction with the planar tricyclic structure of the small-molecule
117 (**Fig. 2b**). Also, the NH backbone group from F97 formed a stabilizing H-bonding with the oxygen
118 from the aldehyde moiety (**Fig. 2b, c**). A 2D representation of the interacting partners showed
119 R100 forming a π -cation interaction and K22 as a putative amino acid forming an H-bonding with
120 the opposite ring of Ro structure (**Fig. 2c**). To confirm these putative interactions, we performed
121 site-directed mutagenesis on the full-length MSI2 protein by mutating the F97 or the main three
122 potential residues involved in Ro binding (F66, F97 and R100) to alanine. MST interaction assays
123 showed a nearly 7-fold decrease in affinity (measured K_D 69.5 ± 14.7 for F97A versus 10.5 ± 0.3
124 μM for wild-type) for the single mutant. More dramatically, the triple mutant (F66/F97/R100) was

125 incapable of binding Ro, confirming our hypothesis that Ro binds at the RNA-interacting site and
126 can compete for it. We further validated if the mutations of these three residues also disrupt RNA
127 binding by MST: wild-type and F97A possess equivalent RNA binding affinities, whereas the triple
128 mutant F66/F97/R100 only partially ablated RNA binding ($K_D > 50 \mu\text{M}$), (**Extended Data Fig. 2b**).

129

130 To further test structure activity relationships, we obtained two Ro related molecules (Ro-NGF
131 and Ro-OH). The first analog, Ro-NGF, was selected to determine if Ro's activity was related to
132 its anti-NGF activity, as previously described³⁹ because this compound showed the highest
133 affinity ($K_{D[\text{NGF}]} = 1.7 \mu\text{M}$) for NGF in its compound series (**Extended Data Table 2**). The second
134 analog, Ro-OH, a reduced form of the aldehyde to an alcohol, contained a single alteration to the
135 Ro aldehyde moiety (**Fig. 2e** and **Extended Data Figs. 2c, 3a, b**). Alchemical free energy
136 calculations showed computed binding free energies (ΔG_{bind}) for the three ligands (Ro, Ro-OH
137 and Ro-NGF) in a similar range, with a slightly higher affinity predicted binding for Ro and Ro-OH
138 (-5.5 and -6.1 vs -5.1 kcal/mol for Ro-NGF) (**Extended Data Fig. 4a**). Both MSI2 protein and
139 ligands adopted a conformationally heterogeneous ensemble of binding poses, with the protein-
140 ligand complex predicted to undergo a slight conformational change for Ro and Ro-OH upon
141 binding (**Extended Data Fig. 4b**). Free energy calculations for all three small-molecules suggest
142 that Ro-NGF adopts a much more diverse set of conformations (as measured by conformational
143 clustering of the fully-interacting alchemical state) than Ro-OH or Ro (**Fig. 2f**). Ro showed the
144 fewest clusters, with the top three clusters accounting for 92.7% of the sampled configurations
145 (**Extended Data Fig. 4c**). Ro-OH showed a larger number of clusters, with the four clusters
146 accounting for 49.1% of sampled configurations, indicating a greater degree of heterogeneity than
147 Ro (**Extended Data Fig. 4d**). Ro-NGF displayed an even greater degree of heterogeneity,
148 showing a large number of low population of clusters (data not shown). Further structural analysis
149 of our docked model suggests Ro-OH lacking the R100 π -cation interaction and Ro-NGF in a
150 displaced position from the RNA-binding core (**Extended Data Fig. 2e, f, g, h**) as compared to
151 Ro, despite similar interacting residues. To experimentally validate these predictions, we
152 performed EMSA of GST-MSI2 competing Ro-OH and Ro-NGF with RNA, comparing potency
153 with Ro and unlabeled RNA as positive controls. Accordingly, whereas Ro-OH showed partial
154 (~30-40%) but significantly poorer inhibition than Ro (65-75%, $p < 0.05$), Ro-NGF showed no
155 ability to displace MSI2-RNA complexes (**Figure 2g and 2h**). These results were further
156 confirmed by FP assay with Ro-OH inhibiting with 12.5-fold less potency than Ro, and Ro-NGF
157 failing to inhibit of RNA-binding activity (**Supplemental Figure 2d**). Furthermore, in MST assays,
158 Ro-OH showed a 27-fold lower affinity than Ro ($K_D 302.0 \pm 119 \mu\text{M}$ for Ro-OH versus 11.2 ± 0.6
159 μM for Ro) for GST-MSI2, whereas Ro-NGF failed to demonstrate any interaction (**Figure 2i**).

160 Thus, our structural and biochemical experimental data support the conclusion that Ro and MSI2

161 interact via the RRM/RNA binding site and that the drug can displace RNA from its binding site,
162 thus likely inhibiting MSI-related translational regulation.

163

164 **Ro 08-2750 demonstrates therapeutic efficacy in murine MLL-AF9 leukemic cells**

165 To test the MSI-inhibitory effect of Ro in a murine AML of leukemia, we used MLL-AF9
166 expressing leukemic BM cells from secondary transplants previously established in the lab⁸. We
167 first assessed the cytotoxicity effects of Ro and the two analogs against these leukemia cells.
168 Consistent with an on-target effect on MSI inhibition and in agreement with the RNA-binding
169 activity inhibition assays, Ro effectively inhibited leukemia cell proliferation (half-effective
170 concentration, $EC_{50} = 2.6 \pm 0.1 \mu\text{M}$). By comparison, the analogues that failed to interact with
171 MSI2 had a diminished effect (Ro-OH $EC_{50} = 21.5 \pm 0.8 \mu\text{M}$; Ro-NGF $> 50 \mu\text{M}$), suggesting that
172 the antiproliferative effect is due to the ability of Ro to inhibit MSI2 RNA binding-activity (**Fig. 3a**).
173 Treatment of cells with Ro resulted in an increase in the myeloid and granulocyte markers (Mac1
174 and Gr1, respectively) at 5 μM dose and 48h treatment as seen by both flow cytometry (**Fig.3b**)
175 and morphologically by Eosin Y and Methylene Blue/ Azure A staining (**Fig. 3c**). When we
176 assessed apoptosis at different time points, we found a significant increase in the Annexin V+
177 population as early as 8h (both at 5 and 10 μM) with the highest increase at 48h and 10 μM Ro
178 (**Fig. 3d** and **Extended Data Fig. 5**).

179

180 We then assessed how MSI2 overexpression affected the plating capacity of MLL-AF9 BM cells
181 in culture in the absence or presence of Ro. MSI2 overexpressing cells formed 50% more
182 colonies than control cells transduced with an empty vector (MIB). Treatment of cells with Ro
183 resulted in reduced colony formation in control cells by $>50\%$ and $\sim 75\%$ at 1 μM and 5 μM
184 concentrations, respectively. MSI2-overexpressing leukemia cells however showed increased
185 resistance to these doses (**Fig. 3e**). Of note, we assessed MSI2 translational targets^{8,31} in these
186 cells by immunoblotting and we found that Ro treatment reduced protein abundance of SMAD3,
187 c-MYC and HOXA9 in control cells, whereas the levels of these proteins remained unaffected in
188 cells that overexpressed MSI2 (**Fig. 3f**). Indicating a potential therapeutic window between
189 normal and malignant cells, Ro abolished MLL-AF9+ BM colony formation at concentrations that
190 did not affect the plating efficiency of normal Lin-Sca+cKit+ (LSK) cells (**Fig. 3g**).

191

192 **Ro 08-2750 treatment inhibits survival of human AML cell lines and patient cells**

193 To determine if Ro also has activity against human myeloid leukemia, we first tested cytotoxicity
194 effects of the small-molecule and the two analogs in MOLM13 (AML, MLL-AF9+) and K562
195 (CML-BC, BCR-ABL+) cell lines, both known to require MSI2 function^{4,26}. Consistent with our
196 previous data in MLL-AF9 cells, we observed that in both these leukemia cell lines, Ro
197 demonstrated anti-proliferative effect ($EC_{50} \sim 8 \mu\text{M}$), whereas the two analogs (Ro-OH and Ro-

198 NGF) revealed a >4.5-fold weaker potency. Ro induced myeloid differentiation and apoptosis in
199 both K562 and MOLM13 cells based on flow cytometry and by morphology (**Fig. 4b-d** and
200 **Extended Data Fig. 6a-c** and **6b**). Plating activity was >80% inhibited at the 20 μ M Ro dose in
201 the human AML cell lines (**Fig. 4e**). Additionally, Ro demonstrated differential sensitivity in three
202 AML patient samples (**Extended Data Table 3**) colony plating assays compared to normal
203 human CD34+ cord blood cell (>50% inhibition in colony numbers at 5 μ M compared to only a
204 modest reduction at 20 μ M Ro, **Fig. 4f**). These results indicate that Ro can induce differentiation
205 and apoptosis in primary human AML cells and spare normal stem cells up to 2 x EC_{50} Ro.
206

207 **Ro 08-2750 inhibits binding of MSI2 to its RNA targets and exhibits gene signature from** 208 **MSI2 depleted cells**

209 To further investigate the effect and mechanism of action of Ro, we initially performed RNA
210 immunoprecipitation (RNA-IP with FLAG) experiments on K562-MIG (empty vector) and K562-
211 FLAG-MSI2 (MSI2 overexpressing) cells (**Fig. 5a**). After incubating the drug at 10 μ M ($\sim EC_{50}$) for
212 1 hour with the cells, we could detect a significant decrease in MSI2 mRNA binding targets
213 (*TGFBR1*, *cMYC*, *SMAD3*, *CDKN1A*) (**Fig. 5b**). These data suggest that Ro can block MSI2
214 binding to target mRNAs in a cellular context at a short time-point.
215

216 To globally assess the proximal effect of Ro treatment on the transcriptional program, we then
217 performed RNA-sequencing on MOLM13 and K562 cells after 4 hours of treatment. Ro incubation
218 resulted in modest but significant gene expression changes in both the MOLM13 and K562 AML
219 cells (59 upregulated, 221 downregulated and 111 upregulated, 164 downregulated, respectively;
220 FDR<0.05), (**Extended Data Tables 4-5**). Most importantly, this Ro signature enriched for the
221 gene expression profiling after shRNA mediated depletion of MSI2 in CML-BC (AR-230 and
222 LAMA84) and AML cell lines (THP1 and NOMO-1) (**Fig. 5c**)²⁶. To annotate the functional
223 pathway overlap with Ro treatment in both cell lines and MSI2 shRNA depletion, we performed
224 gene-set enrichment analysis (GSEA)⁴⁰ on all 4,733 curated gene sets in the Molecular
225 Signatures Database (MSigDB, <http://www.broadinstitute.org/msigdb>) combined with 92
226 additional relevant gene sets from our experimentally derived or published hematopoietic self-
227 renewal and differentiation signatures^{31,40}. Interestingly, we observed an overlap of MSI-
228 associated signatures from our previous dataset and an enrichment with MSI1 direct mRNA
229 targets from the intestine (**Extended Data Tables 7-12** and **Extended Data Fig. 7a**)⁴. Moreover,
230 we observed a ~70% overlap of the functional pathways between each individual cell line and the
231 pathways altered after shRNA depletion of MSI2 (**Fig. 5d**). Among these shared pathways, 76%
232 (543 out of 717) overlapped in MOLM13 compared to K562 cells treated with Ro, which included
233 c-MYC, mRNA related and leukemia associated gene sets (**Fig. 5d** and **Extended Data Table**

234 **12)**. Thus, Ro treatment after a short administration recapitulated a large portion of the MSI2-
235 associated gene expression program.

236

237 To determine how Ro affects previously determined MSI targets, we treated both K562 and
238 MOLM13 cells with increasing concentrations of Ro (up to 20 μ M at 4 hours). In previous studies,
239 MSI was demonstrated to maintain the protein levels of TGF β R1, c-MYC, SMAD3 and HOXA9^{8,31}
240 while suppressing P21 abundance^{41,42}. Consistent with this, we observed a significant and dose
241 dependent reduction of TGF β R1, c-MYC, SMAD3, HOXA9 and an increase in the protein
242 abundance of P21, while the non-target control β -ACTIN remained unchanged (**Fig. 5d** and **5e**).
243 Additionally, Ro could inhibit MSI2 targets in a time-dependent manner with c-MYC, a short half-
244 life protein, being reduced in 1 hour of treatment (**Fig. 5f** and **5g**). In support of Ro altering
245 translation of specific MSI2 targets but not generally inhibiting global translation, we found
246 equivalent global protein synthesis after drug treatment as assessed by O-propargyl-puromycin
247 incorporation (**Extended Data Fig. 7b**). As previously noted by RNA-sequencing, there were
248 modest effects on the mRNA expression of MSI2 targets by qPCR (**Extended Data Fig. 7c**)
249 suggesting that Ro mainly influences its direct targets through a post-transcriptional mechanism.
250 Thus, these results support our hypothesis that Ro acts in the MSI-related translational program.

251

252 **Ro 08-2750 inhibits leukemogenesis in an in vivo MLL-AF9 model of myeloid leukemia**

253 Finally, we sought to determine if Ro has activity in vivo using an aggressive murine MLL-AF9
254 murine leukemia model. Acute treatment of Ro (4h and 12hr) reduced c-KIT protein abundance
255 and intracellular c-MYC (**Fig. 6a-c**). To determine if Ro treatment could affect disease burden we
256 next treated a second cohort of animals and monitored them for disease progression for 19 days
257 after transplantation (**Fig. 6d**). Ro administration every 3 days was well tolerated (**Extended Data**
258 **Fig. 7a, b, c**) demonstrating little to no weight loss and equivalent red blood cells and platelets
259 counts compared to control group. When control mice succumbed to disease (day 19 post-
260 transplantation), we assessed the disease in both groups and found a significant reduction in
261 spleen weights (**Fig. 6e**), white blood cell counts (**Fig. 6f**) and c-MYC levels compared to the
262 controls (**Figure 6g**). These data provide the feasibility that targeting MSI in vivo could have
263 therapeutic efficacy in AML.

264 **DISCUSSION**

265 Inhibiting MSI RNA-binding activity could represent a novel therapeutic avenue in both
266 hematological malignancies and solid cancers. Our previous FP-based screen identified
267 compounds that inhibit MSI binding to RNA³⁸. Here, we characterize Ro 08-2750 as a first
268 selective MSI inhibitor with biochemical, structural and cellular validation linking the compound to
269 the inhibition of the MSI program. Ro falls in the low micromolar range of activity, in line with other
270 RBP associated inhibitors⁴³⁻⁴⁷. We validated Ro as a MSI2 RNA-binding inhibitor with biophysical
271 and biochemical assays by utilizing a high-resolution crystal structure of the MSI2 RRM1. Our
272 newly developed computational molecular modeling algorithm and docking analysis, allowed to
273 predict and validate the key MSI2 residues that were critical for the interactions with Ro in the
274 RNA binding site. Both our novel crystal structure and the computational tools will be useful for
275 the discovery and development of small-molecule RBPs inhibitors. We found that a single
276 chemical reduction of Ro decreased its activity in both in biochemical and in vitro cell based
277 assays. Utilizing a related compound with high affinity binding to NGF, we found that it no longer
278 bound MSI2 and poorly inhibited leukemia cell growth. Further studies involving medicinal
279 chemistry with heterocycle isoalloxazines or pteridine-derived compounds could help identifying
280 more selective and potent MSI-inhibitors.

281

282 Other groups have identified agents that have putative MSI1 inhibitory activity. A natural phenol
283 extracted from cottonseed ((-)-gossypol) was shown to reduce MSI1 to bind RNA⁴⁵ but this
284 interaction was not validated by structure-activity relationships. Of note, (-)-gossypol has been
285 considered to be a pan-active compound that has hit in multiple HTS screens⁴⁸⁻⁵⁰ and assigned to
286 have activity against Bcl-2⁵¹. MSI1 activity was also inhibited by ω -9 monounsaturated fatty acids
287 (e.g. oleic acid), allosterically binding and inducing a conformational change that prevents RNA to
288 bind⁵². It remains unclear if (-)-gossypol or oleic acid have a more broad RNA binding protein
289 inhibitor activity as they were not directly tested against any other RBPs^{43,45,52}. We found that Ro
290 could demonstrate differential binding activity to MSI2 compared to SYNCRIP, Ro's effect on
291 colony formation and direct targets could be rescued by MSI2 overexpression. Moreover, we
292 observed a strong enrichment for the MSI2 shRNAs gene expression signature, associated
293 functional pathways, inhibition of MSI2 binding of target mRNAs and reduced abundance of
294 MSI2 direct targets after Ro treatment. In contrast to other general translational inhibitors Ro did
295 not alter global translation^{53,54}. These data suggest that Ro could be used to probe the acute
296 effects of MSI inhibition in a variety of cellular contexts and cancer models.

297

298 It is also important to note that Ro inhibits both MSI1 and MSI2 and although MSI1 is expressed
299 at low levels in myeloid leukemia it could still be blocking residual MSI1 activity. Moreover, in
300 other models such as the intestine where both factors act redundantly¹³, dual inhibition could

301 provide a powerful therapeutic strategy. Of note based on the close conservation of the RRM
302 between the two proteins it might be challenging to design MSI1 or MSI2 selective inhibitors.

303

304 We demonstrated a therapeutic index for Ro in human AML patient samples versus cord-blood
305 derived CD34+ human stem and progenitor cells. Despite the challenges for in vivo
306 administration, we reduced the disease burden in an aggressive MLL-AF9 leukemia model and
307 decreased MYC levels without overt toxicity. Interestingly, it has previously been shown that MSI2
308 can contribute to chemotherapeutic resistance in different cancer models^{42,55,56}. Future studies
309 could examine if combination therapies could provide additional clinical benefit.

310

311 This study identifies and characterizes Ro 08-2750 as the first compound selectively inhibiting the
312 oncogenic RNA-binding activity of MSI in myeloid leukemia. It will be important to use this
313 compound (or other chemical derivatives) to test their efficacy in other cancer models and on MSI
314 function related to normal physiology. We suggest that Ro provides the rationale for developing
315 more potent compounds with improved clinical utility for the treatment of cancers that are
316 dependent on the MSI family. Additionally, as there are hundreds of RRM containing RNA
317 binding proteins, Ro targeting an RRM motif to block RNA activity represents a valuable proof of
318 concept for the general inhibition of these class of RNA regulators. Thus, we provide a framework
319 to identify and test novel RNA binding protein inhibitors in cancer.

320

321 **Methods**

322 **Purification and culture of cord blood derived HSPC-CD34+ cells**

323 Mononuclear cells were isolated from cord blood using HetaStach solution (6% Hetastarch in
324 0.9% NaCl) and Ficoll-Hypaque Plus density centrifugation. CD34+ Hematopoietic Stem and
325 Progenitor Cells (HSPCs) were subsequently purified by positive selection using the Auto MACS
326 Pro Separator and isolation kit (Miltenyi) and were cultured in Iscove's modified Dulbecco's
327 medium (IMDM, Cellgro), 20% BIT 9500 medium (Stem Cell Technologies) supplemented with
328 SCF (100 ng/ml), FLT-3 ligand (10 ng/ml), IL-6 (20 ng/ml) and TPO (100 ng/ml) as the basic
329 culture. All cytokines were purchased from Peprotech, NJ.

330

331 **Isolation and viral transduction of murine MLL-AF9 leukemia and normal cells**

332 Tibia, femurs, pelvis, and arm bones from leukemia or C57BL/6 wild type mice (10-12 weeks old)
333 were harvested, crushed, filtered, and subjected to red blood cell lysis (Qiagen). To isolate c-Kit⁺
334 cells, bone marrow cells were incubated with anti-CD117 microbeads (Miltenyi Biotec), according
335 to manufacturer's instructions, and then subjected to positive selection using autoMACS Pro
336 Separator. For MLL-AF9 BM cells, previously thawed vials from secondary transplants (Park et al.
337 2015) were used. All murine cells were cultured and transduced in RPMI with 10% FBS and
338 cytokines SCF (10 ng/ml), IL-3 (10 ng/ml), and IL-6 (10 ng/ml) and GM-CSF (10 ng/ml). For MSI2
339 overexpression, cells were spinfected with viral supernatant containing MSCV-IRES-BFP or
340 MSI2-IRES-BFP constructs (see Cloning section).

341

342 **Colony forming unit (CFU) assays**

343 10,000 leukemic MLL-AF9 BM cells or c-Kit enriched normal stem cells (Lin-Sca-Kit⁺) were
344 plated on methylcellulose-based culture media (methocult) GFM3434 (Stem Cell Technologies).
345 Colonies were scored every five days for leukemia cells and every seven to nine days for normal
346 c-kit-enriched bone marrow cells. For human cells, 5,000 of the leukemia cell lines K562 (CML-
347 BC) or MOLM13 (AML) and 10,000 of HSPCs CD34+ or AML patient cells were plated (in
348 duplicate) in methylcellulose (MethoCult™ H4434 Classic, Stem Cell Technologies). CFU
349 colonies in HSPCs CD34+ were scored 14 days after seeding. AML patient cells characteristics
350 are shown in **Extended Data Table 3**.

351

352 **Flow cytometry**

353 To monitor the differentiation status, 200K MLL-AF9 BM cells DMSO or Ro treated (during 8, 16,
354 24, 48h) were stained with the following antibodies: anti-CD11b (Mac1)-PE (clone M1/70,
355 #101208, BioLegend), anti-Ly-6G (Gr1)-APC (clone RB6-8C5, #17-5931-82, eBioscience), and
356 anti-CD117 (c-Kit)-APC-Cy7 (clone 2B8, #105826, BioLegend). For the human cell lines
357 differentiation, we used two panels: (1) anti-CD14-PE (clone M5E2, #555398, BD Pharmingen),

358 anti-CD13-APC (clone TUK1, #MHCD1305, Life Technologies); (2) anti-CD71-APC (clone
359 CY1G4, #334104, BioLegend), anti-CD235a (Glycophorin A)-PE (clone YTH89.1, #MA5-17700,
360 Invitrogen). All samples were stained for 20min in the dark, washed once with PBS 1X and re-
361 suspended in RPMI + 2% FBS for analysis. For intracellular flow cytometry detection of cMYC, 1-
362 2×10^6 cells were fixed in 2% paraformaldehyde for 15 minutes, washed 2 times with 1X PBS and
363 permeabilized with cold methanol and kept at -80 until use. For the staining, cells were washed
364 twice in 1X PBS and stained in 100 μ l final volume. c-MYC (5605, Cell Signaling Technology non-
365 labelled primary antibody was incubated at 1/200 dilution for 1h and labelled donkey anti-rabbit
366 Alexa Fluor 568 (#A10042, Invitrogen) or goat anti-rabbit Alexa Fluor 647 (#A21245, Invitrogen)
367 were used at 1/400 for 20-30 minutes. Cells were washed once with PBS 1X and re-suspended in
368 RPMI + 2% FBS for analysis. All flow cytometry analysis was performed in a LSRII or LSR
369 Fortessa (BD Biosciences) and data was graphed by using FlowJo™ version 10.4.

370

371 **Morphological analysis**

372 After the appropriate time of Ro treatment (or DMSO in controls) in culture, 1.5×10^5 MLL-AF9 BM
373 or human leukemia cells (K562 and MOLM13) were washed once with 1X PBS, counted and
374 centrifuged onto slides for 5 minutes at 500 rpm and air-dried for 24h prior to Richard-Allan
375 Scientific Three-Step Stain Staining Set (Thermo Scientific) based on Eosin Y and Methylene
376 Blue/ Azure A and mounted with Permount solution (Fisher). Cell morphology was evaluated by
377 light microscopy at 400X magnification (Zeiss Imager M-2, equipped with AxioCam ERc 5s).

378

379 **Apoptosis measurements**

380 Apoptosis measurements were taken by MUSE™ Cell Analyzer (Millipore) using the MUSE™
381 Annexin V and Dead Cell Assay Kit (Millipore) as recommended by the instructions from the
382 manufacturer. Dot plots showing viability versus Annexin V+ cells are shown in **Extended Data**
383 **Figures 4 and 5.**

384

385 **In vivo transplantation of leukemia cells and Ro 08-2750 administration**

386 10,000 of MLL-AF9 BM secondary mouse leukemia cells previously obtained¹ were injected retro-
387 orbitally into female C57BL/6 (10-12 weeks old) recipient mice that had been sublethally
388 irradiated at 475 cGy. Drug administration (Ro 08-2750, 13.75 mg/Kg, DMSO) was performed by
389 intraperitoneal injections (50 μ L, top tolerated DMSO volume) 3 weeks after BM transplants
390 (when showing signs of disease) for pharmacodynamic experiments (see **Fig. 6a**), and 3 days
391 after BM transplant for in vivo long-term studies (see **Figure 6d**). Mice weight were monitored
392 every day to check for toxicity. All animal studies were performed on animal protocols approved
393 by the Institutional Animal Care and Use Committee (IACUC) at Memorial Sloan Kettering Cancer
394 Center.

395

396 **Fluorescence Polarization (FP) to assess RNA-binding activity inhibition**

397 To validate RNA-binding activity inhibition by Ro 08-2750 and derivatives (Ro-OH, Ro-NGF) we
398 used Fluorescence Polarization (FP) based assay in as previously described in 384-well format
399 for dose-response curve studies². As previously, the RNA oligo used (Cy3-C₉-[spacer]-
400 rGUAGUAGU, Integrated IDT Technologies) contained 2 MSI motifs (GUAGU) and was 8-
401 nucleotides long, optimal to minimize background and unspecific interactions. Differently, here
402 manual pipetting was used to plate the reagents and the FP reading was performed in a BioTek
403 Synergy Neon Plate Reader (High-Throughput Screening Resource Center, Rockefeller
404 University).

405

406 **Binding affinity quantifications by MicroScale Thermophoresis (MST)**

407 For binding affinity studies of RNA and small-molecules to proteins of interest, purified
408 recombinant GST-MSI2 WT, F97A and F66A/F97A/R100A mutants and GST-SYNCRIP were
409 NT647-labeled using an amine-coupling kit (NanoTemper Technologies). Runs were performed
410 at a concentration range of 50-125 nM (MSI2 and mutants) and 60 nM (SYNCRIP) to get optimal
411 fluorescence signal using an LED power of 40-50% in a red laser equipped Monolith NT.115
412 (NanoTemper Technologies) (High-Throughput Screening Resource Center, Rockefeller
413 University). Prior to each run, protein preparations were diluted in MST buffer (50 mM HEPES,
414 100 mM NaCl, 0.05% Tween-20, pH 7.4) and aggregation was minimized by centrifuging the
415 solutions at 15,000 rpm for 10 minutes. GST-proteins or GST-protein/RNA complexes (15 min
416 pre-incubation) were mixed with increasing concentrations of small-molecules (0.015 to 500 μ M)
417 or RNA (0.0015 to 50 μ M) and loaded onto 16 Premium Coated capillaries. The RNA oligo used
418 (rGUAGUAGUAGUAGUA, Integrated IDT Technologies) contained 4 MSI motifs (GUAGU) and
419 was 15-nucleotides long. The MST measurements were taken at RT and a fixed IR-laser power
420 of 40% for 20 seconds per capillary. GraphPad Prism was used to fit the normalized data and
421 determine apparent K_D values, represented as percent of fraction bound.
422

423

424 **Chemiluminescent Electrophoresis Mobility Shift Assays (EMSA)**

425 An EMSA approach to assess MSI2-RNA complexes and the inhibitory effect of small-molecules
426 was set up by using LightShift Chemiluminescent RNA EMSA kit (Thermo Scientific). In brief,
427 GST-MSI2 (125-250 ng) was preincubated with DMSO or the small-molecule (typically 20 μ M
428 final concentration) during 1h at RT in EMSA buffer 1X RNA EMSA binding buffer (10 mM
429 HEPES, 20 mM KCl, 1 mM MgCl₂, 1 mM DTT, Thermo Scientific) supplemented with 5% glycerol,
430 100 μ g/mL tRNA and additional 10 mM KCl. After this period, 40 nM of biotinylated-RNA (biotin-
431 rGUAGUAGUAGUAGUA, Integrated IDT Technologies -same as for MST-) was added to the
432 mixture (20 μ L final volume) and incubated another 1h at RT. During this second incubation

433 period, a 4-20% TBE polyacrylamide gel (BioRad) was pre-run at 100V for 30-45min in cold 0.5X
434 TBE (RNase free). 5 μ L of 5X loading buffer was added to the 20 μ L reaction and loaded into the
435 pre-run TBE gel and voltage set at 100V. Samples were electrophoresed until 3/4 of the length of
436 the gel. Samples were then transferred in 0.5X TBE at 350-400 mA for 40 min. Membranes were
437 then crosslinked with UV-light crosslinking instrument (UV Stratagene 1800) using Auto-Cross
438 Link function. Membranes were either stored dry for development next day or developed using the
439 detection biotin-labeled RNA chemiluminescence kit (as indicated by the manufacturer) (Thermo
440 Fisher) and Hyperfilm ECL (GE Healthcare).

441

442 **Cloning, expression, and purification of GST tagged proteins**

443 Human full-length MSI2 was cloned into the retroviral backbone pMSCV-IRES-BFP (MIB) vector
444 (a gift from Dario Vignali; Addgene plasmid # 52115) by Custom DNA Constructs (University
445 Heights, Ohio) introducing a 5'Flag tag and using BamHI and EcoRI restriction sites. Human full-
446 length MSI2 was previously cloned into pGEX6P3 as described². RNA-recognition motif 1
447 (RRM1) from human MSI2 (nucleotides #64-270, NM_138962.2) was subcloned into empty
448 pGEX6P3 using EcoRI and NotI restriction sites. Human SYNCRIP (hnRNP-Q variant 3,
449 NM_001159674.1) was subcloned into empty pGEX6P3 (GE Healthcare) by introducing a 5'Flag
450 sequence (5'-ATGGATTACAAGGATGACGACGATAAG-3') and using Sall and NotI sites. GST-
451 Flag-MSI2 wild-type (WT), Flag-MSI2 mutants (F97A, F66A/F97A/R100A), GST-RRM1 and GST-
452 Flag-SYNCRIP recombinant proteins were produced in BL21 (DE3) competent cells (Agilent
453 Technologies, Santa Clara, CA) as previously reported for MSI2 WT². Here, GST-SYNCRIP
454 protein needed higher content of NaCl (250 mM) in the 1X PBS dialysis step and final buffer for
455 optimal storage and performance in the biochemical and biophysical assays performed.

456

457 **Site-directed mutagenesis**

458 To perform site-directed mutagenesis into pGEX6P3-Flag-MSI2 construct and express the
459 corresponding recombinant GST-MSI2 mutants, we used QuikChange Lightning and Multi Site-
460 Directed Mutagenesis Kit (#210513 and #210518, Agilent Technologies). The primers were
461 designed using QuickChange Primer Design
462 (<https://www.genomics.agilent.com/primerDesignProgram.jsp>) and were the following: F66A
463 (Fwd: 5'-GCTCCAGAGGCTTCGGTGCCGTCACGTTCCGAG-3', Rev: 5'-
464 CTGCGAACGTGACGGCACCGAAGCCTCTGGAGC-3'; F97A/ R100A (Fwd: 5'-
465 AGACGATTGACCCCAAAGTTGCAGCTCCTCGTGCAGCGCAACCCAA-3', Rev: 5'-
466 TTGGGTTGCGCTGCACGAGGAGCTGCAACTTTGGGGTCAATCGTCT-3') and R100A (using
467 F97A mutant construct as template) (Fwd: 5'-
468 CCAAAGTTGCAGCTCCTCGTGCAGCGCAACCCA-3', Rev: 5'-

469 TGGGTTGCGCTGCACGAGGAGCTGCAACTTTGG-3'). PCR reactions and cloning were
470 performed as indicated by the manufacturer (Agilent Technologies).

471

472 **Human MSI2 RRM1 recombinant protein production**

473 GST-RRM1 protein was initially produced in BL21 (DE3) competent cells (Agilent Technologies,
474 Santa Clara, CA) as previously reported for MSI2 WT². Here, the cell lysate of 4L initial culture
475 was centrifuged at 15,000 rpm for 1h and the resulting volume applied to a XK16/20 column pre-
476 packed with Glutathione Sepharose 4 Fast Flow connected to an AKTA Prime FPLC (GE
477 Healthcare). To obtain the RRM1 optimal prep for the crystal preparation, the collected fractions
478 containing GST-RRM1 (in 50 mM Tris-HCl, 20 mM reduced L-Glutathione) were pooled and
479 dialyzed against PreScission Protease Buffer (50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1mM
480 DTT, pH 7.5). GST tag was then cleaved with PreScission Protease overnight at 4°C. Pure RRM1
481 fractions were obtained through size exclusion chromatography (HiLoad Superdex 75, GE
482 Healthcare) and concentrated with a 3K Amicon Ultra Centricon (Millipore).

483

484 **Crystallization and structure determination**

485 A final concentrated MSI2 RRM1 pure protein preparation (>98% by coomassie) at 2 mg/mL in 50
486 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1mM DTT, pH 7.5 was crystallized by sitting drop
487 vapor diffusion. A 1 μ L of protein solution was mixed with an equal volume of precipitant solution
488 containing 100 mM Tris, 200 mM Li₂SO₄, 25% PEG (pH 8.5). Crystals appeared after two weeks.
489 They were cryoprotected by mother liquor containing 25% glycerol and flash frozen in liquid
490 nitrogen. X-ray diffraction data were collected from single crystals at the Advanced Photon
491 Source beamline 24ID-C at 100 K. Indexing and merging of the diffraction data were performed in
492 HKL2000³. The phases were obtained by molecular replacement by PHENIX⁴ using PDB entry
493 1UAW as the search model. Interactive model building was performed using O⁵. Refinement was
494 accomplished with PHENIX. Data collection and refinement statistics are summarized in
495 **Extended Data Table 1**. The crystal structure has been deposited in RCSB PDB under the
496 accession code 6DBP.

497

498 **RNA purification and quantitative real-time PCR**

499 Total RNA was isolated from 1-2x10⁶ cells dry pellets kept at -80C for less than a week using
500 Qiagen RNeasy Plus Mini kit. cDNA was generated from RNA using iScript cDNA Synthesis
501 (#1708891, BioRad) with random hexamers according to the manufacturer's instructions. Real-
502 time PCR reactions were performed using a Vii7 sequence detection system. β -ACTIN was
503 commonly used to normalize for cDNA loading. Relative quantification of the genes was
504 performed using Power SYBR Mix (2X) and specific primers for *c-MYC*, *TGF β 1*, *SMAD3*,
505 *HOXA9* and *CDKN1A* and the $2^{-\Delta\Delta Ct}$ method as described by the manufacturer.

506

507 **Immunoblot analysis**

508 For immunoblot analysis, Ro treated and DMSO control MOLM13 or K562 cells (routinely at
509 0.5×10^6 cells/ mL) were counted and washed twice with cold PBS before collection. $1-5 \times 10^6$ cells
510 were resuspended and lysed in 250 μ l of 1X RIPA Buffer supplemented with Protease Inhibitor
511 Tablets (Sigma-Aldrich) buffer for 30min on ice. After centrifugation at 14,000rpm on a top-bench
512 centrifuge, lysate (supernatant) was collected and total protein quantified by BCA (Thermo
513 Scientific). Cell lysates were separated by 4–15% SDS–PAGE and transferred to 0.45 μ m
514 nitrocellulose membrane. Membranes were blocked and were blotted overnight (4C) for TG β R1
515 (ab31013, Abcam, 1:750 dilution), SMAD3 (9523S, Cell Signaling Technology, 1:750 dilution),
516 HOXA9 (07-178, Millipore, for drug dose-dependent experiments and ab140631, Abcam; 1:1,000
517 dilution for time-course experiments), c-MYC (5605, Cell Signaling Technology; 1:1,000 dilution),
518 P21 (2947S, Cell Signaling Technology, 1:750 dilution), MSI2 (ab76148, Abcam; 1:2,000 dilution)
519 and β -ACTIN-HRP conjugated (A3854, Sigma-Aldrich; 1:20,000 dilution) and developed by
520 Hyperfilm ECL (GE Healthcare) with ECL and pico-ECL reagents (Thermo Scientific).

521

522 **Luminescence-based cytotoxicity assays (EC_{50})**

523 10,000 cells (MLL-AF9 BM from secondary transplants or human leukemic cell lines -K562 or
524 MOLM13-) were plated into U-bottom 96-well plates in the presence of increasing concentration
525 of small-molecules (Ro, Ro-OH or Ro-NGF) up to 100 μ M (in 1:2 serial dilutions). Cells were
526 cultured for 72h at 37C in a 5% CO₂ incubator. To read cell viability, Cell-Titer Glo™ kit
527 (Promega) was used. After cooling down cells to RT for 20-30min, 100 μ L of the cultured cells
528 were transferred to opaque-white bottom 96-well plates and mixed with 100 μ L of Cell-Titer Glo™
529 Reagent (previously prepared by mixing buffer and powdered substrate). The mixture was
530 incubated for 15min at RT and read using a Synergy H1 Hybrid reader (BioTek) for
531 luminescence. Data was normalized as percentage viability and graphed by non-linear regression
532 curves in Graph Pad PRISM 7.0. K562 and MOLM13 lines were purchased from ATCC,
533 authenticated by Genetica, and tested negative for mycoplasma contamination.

534

535 **RNA immunoprecipitation (RNA-IP)**

536 To assess mRNA enrichment and blocking of protein-binding to mRNA by the small-molecules
537 we performed RNA immunoprecipitation (RNA IP) experiments using Magna RIP RNA-binding
538 protein immunoprecipitation kit (#03-115, Millipore). 25×10^6 K562-MIG or MSI2 overexpressing
539 cells 1h treated with DMSO (control) or Ro μ M were used. First, cells were washed with cold PBS
540 and lysed. Five micrograms of mouse anti-Flag (clone M2, #F1804, Sigma-Aldrich) antibody
541 incubated with magnetic beads were used to immunoprecipitate Flag-MSI2 K562 cells. After
542 washing the immunoprecipitated, they were treated with proteinase K. RNA extraction was

543 performed by the phenol–chloroform method, and 200-500 ng of purified RNA was converted to
544 cDNA using the Verso cDNA kit (Thermo Scientific). qPCR was used to validate target mRNAs
545 bound by MSI2 and control cells.

546

547 **O-Propargyl-Puromycin incorporation by flow cytometry**

548 Cells were plated at a density of 200,000 cells/ml and pre-treated with DMSO or Ro up to 4h.
549 Then, 50 μ M O-propargyl-puromycin (OP-Puro; NU-931-05, Jena Bioscience) was added. Control
550 cells were co-incubated with DMSO or Ro and treated with 150 μ g/ml cycloheximide for 15 min.
551 Non-OP-Puro treated cells were also used as negative controls for flow cytometry. Cells were
552 washed twice before collection and subjected to processing using the Click-iT Flow Cytometry
553 Assay kit (C10418, Invitrogen) following the manufacturer's instructions. Labeled cells were
554 analyzed using a BD LSR Fortessa instrument and graphed as Alexa Fluor 647 (AF647) Mean
555 Fluorescence Intensity (normalized to DMSO control treated with OP-Puro).

556

557 **RNA sequencing**

558 Total RNA was isolated from 1×10^6 dry pellets of K562 and MOLM13 4h treated with DMSO
559 (control) or Ro 20 μ M ($n = 4$ for each group) using Qiagen RNeasy Plus Mini kit and the quality
560 assessed on a TapeStation 2200 (Agilent technologies). QuantSeq 3' mRNA-Seq Library Prep Kit
561 FWD (Lexogen, Vienna Austria), supplemented with a common set of external RNA controls,
562 according to manufacturer's recommendations (ERCC RNA Spike-In mix, ThermoFisher
563 Scientific, #4456740). An in-house pipeline was used for read mapping and alignment, transcript
564 construction and quantification of data generated by sequencing (HiSeq 2000, NYGC, NY, USA).
565 This procedure was done in the Epigenetics Core from MSKCC. RNA-seq data has been
566 deposited to GSE114320 and can be viewed for reviewers only:
567 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE114320>.

568

569 **Synthesis of Ro-OH by reduction of Ro 08-2750 aldehyde**

570 To a cooled (0 °C) slurry of Ro 08-2750 (19 mg, 0.070 mmol) in anhydrous MeOH (1.9 mL) was
571 added LiBH₄ (32 mg, 1.5 mmol) in portions over 5 min. The slurry turned from bright orange to
572 dark brown, then dark green within 10 min. The reaction mixture was removed from the ice bath
573 and allowed to warm to rt (22 °C) over 2 h. Reaction progress was monitored by LC-MS (5–95%
574 MeCN in H₂O). Four portions of LiBH₄ (10 mg, 0.04 mmol) were added every 12 h until the
575 reaction was complete. The reaction was quenched with AcOH (10 mL) and filtered. The solids
576 were washed with water (5 mL), MeOH (5 mL), and Et₂O (5 mL). The solid was collected and
577 dried under vacuum to provide a pale orange solid (7 mg, 26%). Purification by HPLC (5–95%
578 MeCN in H₂O) afforded the product as an orange solid (3 mg, 16%). The synthesis was adapted
579 from Salach *et al.*⁶

580 ¹H-NMR (600 MHz, DMSO) δ 11.34 (s, 1H), 7.91 (s, 1H), 7.87 (s, 1H), 5.69 (t, *J* = 4.5, 1H), 4.74
581 (d, *J* = 4.4, 2 H), 3.99 (s, 3H), 2.38 (s, 3H). ¹³C-NMR (150 MHz, DMSO) 159.8 (C), 155.4 (C),
582 150.5 (C), 149.7 (C), 137.4 (C), 133.57 (C), 133.56 (C), 131.5 (CH), 131.0 (C), 112.3 (CH), 60.8
583 (CH₂), 31.7 (CH₃), 17.2 (CH₃); IR (ATR): 2361, 2341, 1717. **ESI-MS** *m/z* (rel int): (pos) 273.1
584 ([M+H]⁺, 100).

585

586 **Statistical analysis**

587 Student's *t* test was used for significance testing in the bar graphs, except where stated
588 otherwise. A two-sample equal-variance model assuming normal distribution was used. The
589 investigators were not blinded to the sample groups for all experiments. *P* values less than 0.05
590 were considered to be significant. Graphs and error bars reflect means + standard error of the
591 mean except stated otherwise. All statistical analyses were carried out using GraphPad Prism 7.0
592 and the R statistical environment.

593

594 **Modeling and System preparation for Computational Modeling**

595 System preparation, modeling, and initial docking calculations were performed using the
596 Schrödinger Suite molecular modeling package (version 2015-4), using default parameters
597 unless otherwise noted. The MSI2 RRM1 protein structure (PDB ID: 6DBP) was prepared using
598 the Protein Preparation Wizard⁷. In this step, force field atom types and bond orders were
599 assigned, missing atoms were added, tautomer/ionization states were assigned, water
600 orientations were sampled, and ionizable residues (Asn, Gln, and His residues) have their
601 tautomers adjusted to optimize the hydrogen bond network. A constrained energy minimization
602 was then performed. All crystallographically resolved water molecules were retained.

603 Potential binding sites were explored and characterized using the SiteMap^{8,9} tool. Ligands with
604 experimental activity and known inactives were docked into putative binding sites using Glide
605 SP^{10,11} to evaluate enrichment of known actives. Best docking scores were for the 'Ro' series for
606 the '(-)-gossypol' binding site described by Lan *et al.*¹² compared to other putative pockets.

607 Since the receptor may not be in an optimal conformation to bind small molecule inhibitors,
608 induced fit docking¹³ of ligand Ro 08-2750 was performed to this binding pocket. Induced fit
609 docking results were validated with the metadynamics protocol described by Clark *et al.*¹⁴ In
610 these metadynamics simulations a biasing potential is applied to the ligand RMSD as collective
611 variable. The resulting potential energy surface is evaluated towards how easy a ligand can move
612 away from the initial binding mode. The underlying assumption is that a ligand pose which is
613 closer to the real one has a higher energetic barrier to leave the pose than an incorrect pose. The
614 pose ranked second using the induced fit docking score retrieved the best score from the
615 metadynamics ranking protocol compared to the other induced fit docking poses. This receptor
616 configuration was furthermore tested towards its suitability for a virtual screening by a Glide SP

617 docking of known actives into this pocket. The docking scores using this receptor conformations
618 were better (down to -6.2) compared to the initial protein conformation in the crystal structure.
619 Furthermore, a WaterMap^{15,16} calculation was done for this receptor.

620

621 **Induced Fit Docking of Ro-NGF and Ro-OH compounds**

622 Induced Fit Docking (IFD) was performed against the receptor pose from the selected Ro 08-
623 2750 pose, using Schödinger molecular modeling suite (version 2017-4). Poses for Ro-NGF and
624 Ro-OH, the top and second scored poses respectively, were selected to most closely match the
625 Ro 08-2750 pose.

626

627 **Alchemical Free Energy Calculations**

628 Absolute alchemical free energy calculations were carried out to validate the putative binding
629 poses in a fully-flexible explicitly solvated system. The YANK GPU-accelerated free energy
630 calculation code with the Amber family of forcefields was used for this purpose. Details follow:

631 *System preparation and modeling.* The top poses generated by induced fit docking, as described
632 above, were selected as input protein and ligand poses. Because proteins and ligands were
633 already prepared, they were simply run through the pdbfixer 1.4 command line tool with add-
634 atoms and add-residues set to None to convert residue and atom names to be compatible with
635 Amber tleap.

636 *Parameterization.* tleap (from the minimal conda-installable AmberTools 16 suite ambermini
637 16.16.0) was used to solvate the complex in a cubic box with a 12Å buffer of TIP3P water
638 molecules around the protein¹⁷. The system was parameterized using AMBER's forcefield
639 ff14sb¹⁸ and GAFF 1.8¹⁹. Missing ligand parameters were determined using antechamber²⁰. The
640 ligand was assigned charges using the AM1-BCC^{21,22} implementation in OpenEye (OEtoolkit
641 2017.6.1²³ through openmoltools 0.8.1).

642 *Minimization.* Minimization was performed using the implementation of the L-BFGS²⁴ algorithm in
643 OpenMM 7.1.1²⁵ with a tolerance of 1kJ/mol/nm.

644 *Production Simulation.* Production simulation was run using YANK 0.19.4²⁶ using OpenMMTools
645 0.13.4. In order to keep the ligand from diffusing away from the protein while in a weakly coupled
646 state, it was confined to the binding site using a Harmonic restraint with an automatically-
647 determined force constant ($K = 0.33 \text{ kcal/mol/Å}^2$). The restraint was centered on the following
648 receptor residues using all-atom selection: 2, 4, 46, 76, 78, and 80. The ligand atoms were
649 automatically determined. The calculation was performed using particle mesh Ewald (PME)²⁷
650 electrostatics with default YANK settings with a real-space cutoff of 9Å. A long-range isotropic
651 dispersion correction was applied to correct for truncation of the Lennard-Jones potential at 9Å.
652 The system was automatically solvated with TIP3P²⁸ solvent and four neutralizing Cl⁻ ions,
653 parameterized using the Joung and Cheatham parameters²⁹. Production alchemical Hamiltonian

654 exchange free energy calculations were carried out at 300 K and 1 atm using a Langevin
655 integrator (VRORV splitting)³⁰ with a 2 fs timestep, 5.0 ps⁻¹ collision rate, and a molecular-scaling
656 Monte Carlo barostat. Ro 08-2750 and Ro-NGF were run for 10000 iterations (50 ns/replica) with
657 2500 timesteps (5 ps) per iteration, while Ro-OH was run for 15000 iterations (75 ns/replica) with
658 2500 timesteps (5 ps) per iteration. Complex configurations were stored for each replica once per
659 iteration. Replica exchange steps were performed each iteration to mix replicas using the Gibbs
660 sampling scheme described previously³¹. The alchemical pathway was automatically determined
661 for each compound using the YANK autoprotocol protocol trailblazing feature.

662 *Absolute binding free energy estimates.* Absolute free energies (ΔG) of binding for each
663 compound was estimated using MBAR³². Samples were reweighted to a cutoff of 16Å to correct
664 the isotropic dispersion correction to a nonisotropic long-range dispersion. This correction is
665 important to account for the heterogeneous density of protein. To remove the harmonic restraint
666 bias, samples were reweighted to substitute a squared well restraint of radius 10Å.

667 *Clustering analysis.* The fully interacting trajectory from YANK was extracted to a PDB file,
668 discarding the following number of initial iterations, which came prior to equilibration³³: 1500 for
669 Ro 08-2750, 1600 for Ro-OH, and 1600 for Ro-NGF. These trajectories were aligned in MDTraj³⁴
670 using only protein backbone atoms. The small molecules were then sliced out and clustered on
671 Cartesian coordinates using the MSMBuild³⁵ implementation of RegularSpatial clustering using
672 a 1Å RMSD cutoff. For the most populated clusters for Ro 08-2750 and Ro-OH, cluster centers
673 were selected and shown with 10 randomly sampled cluster members. Ro-NGF produced a large
674 number of lowly populated clusters with highly heterogeneous binding poses, and were therefore
675 not shown.

676 *Conformational heterogeneity analysis.* To investigate the conformational heterogeneity in the
677 presence or absence of the ligand, the fully interacting thermodynamic state (corresponding to
678 the holo protein bound to the ligand) and fully non-interacting state (corresponding to the apo
679 protein free of ligand interactions) for all three ligands were extracted using a 4-frame skip,
680 discarding the initial frames as above.

681

682 *Code availability.* All Schrödinger project files, YANK simulation inputs, and analysis scripts have
683 been made publicly available (<https://github.com/choderalab/musashi>).

684

685 **Methods. Supplemental References.**

- 686 1 Park, S. M. *et al.* Musashi2 sustains the mixed-lineage leukemia-driven stem cell
687 regulatory program. *J Clin Invest* **125**, 1286-1298 (2015).
- 688 2 Minuesa, G. *et al.* A 1536-well fluorescence polarization assay to screen for modulators
689 of the MUSASHI family of RNA-binding proteins. *Comb Chem High Throughput Screen*
690 **17**, 596-609 (2014).
- 691 3 Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation
692 mode. *Methods Enzymol* **276**, 307-326 (1997).

- 693 4 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular
694 structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010).
- 695 5 Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building
696 protein models in electron density maps and the location of errors in these models. *Acta*
697 *Crystallogr A* **47 (Pt 2)**, 110-119 (1991).
- 698 6 Salach, J. *et al.* Studies on succinate dehydrogenase. Site of attachment of the
699 covalently-bound flavin to the peptide chain. *Eur J Biochem* **26**, 267-278 (1972).
- 700 7 Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and
701 ligand preparation: parameters, protocols, and influence on virtual screening
702 enrichments. *J Comput Aided Mol Des* **27**, 221-234 (2013).
- 703 8 Halgren, T. A. Identifying and characterizing binding sites and assessing druggability. *J*
704 *Chem Inf Model* **49**, 377-389 (2009).
- 705 9 Halgren, T. New method for fast and accurate binding-site identification and analysis.
706 *Chem Biol Drug Des* **69**, 146-148 (2007).
- 707 10 Halgren, T. A. *et al.* Glide: a new approach for rapid, accurate docking and scoring. 2.
708 Enrichment factors in database screening. *J Med Chem* **47**, 1750-1759 (2004).
- 709 11 Friesner, R. A. *et al.* Glide: a new approach for rapid, accurate docking and scoring. 1.
710 Method and assessment of docking accuracy. *J Med Chem* **47**, 1739-1749 (2004).
- 711 12 Lan, L. *et al.* Natural product (-)-gossypol inhibits colon cancer cell growth by targeting
712 RNA-binding protein Musashi-1. *Mol Oncol* **9**, 1406-1420 (2015).
- 713 13 Sherman, W., Day, T., Jacobson, M. P., Friesner, R. A. & Farid, R. Novel procedure for
714 modeling ligand/receptor induced fit effects. *J Med Chem* **49**, 534-553 (2006).
- 715 14 Clark, A. J. *et al.* Prediction of Protein-Ligand Binding Poses via a Combination of
716 Induced Fit Docking and Metadynamics Simulations. *J Chem Theory Comput* **12**, 2990-
717 2998 (2016).
- 718 15 Young, T., Abel, R., Kim, B., Berne, B. J. & Friesner, R. A. Motifs for molecular
719 recognition exploiting hydrophobic enclosure in protein-ligand binding. *Proc Natl Acad Sci*
720 *U S A* **104**, 808-813 (2007).
- 721 16 Abel, R., Young, T., Farid, R., Berne, B. J. & Friesner, R. A. Role of the active-site
722 solvent in the thermodynamics of factor Xa ligand binding. *J Am Chem Soc* **130**, 2817-
723 2831, (2008).
- 724 17 Case, D. A., Betz, R.M., Cerutti, D.S., Cheatham, III, T.E., Darden, T.A., Duke, R.E.,
725 Giese, T.J., Gohlke, H., Goetz, A.W., Homeyer, N., Izadi, S., Janowski, P., Kaus, J.,
726 Kovalenko, A., Lee, T.S., LeGrand, S., Li, P., Lin, C., Luchko, T., Luo, R., Madej, R.,
727 Mermelstein, D., Merz, K.M., Monard, G., Nguyen, H., Nguyen, H.T., Omelyan, I.,
728 Onufriev, A., Roe, D.R., Roitberg, A., Sagui, C., Simmerling, C.L., Botello-Smith, W.M.,
729 Swails, J., Walker, R.C., Wang, J., Wolf, R.M., Wu, X., Xiao, L. and Kollman, P.A.
730 AMBER 2016, University of California, San Francisco (2016).
- 731 18 Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone
732 Parameters from ff99SB. *J Chem Theory Comput* **11**, 3696-3713 (2015).
- 733 19 Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and
734 testing of a general amber force field. *J Comput Chem* **25**, 1157-1174 (2004).
- 735 20 Wang, J., Wang, W., Kollman, P. A. & Case, D. A. Automatic atom type and bond type
736 perception in molecular mechanical calculations. *J Mol Graph Model* **25**, 247-260 (2006).
- 737 21 Jakalian, A., Jack, D. B. & Bayly, C. I. Fast, efficient generation of high-quality atomic
738 charges. AM1-BCC model: II. Parameterization and validation. *J Comput Chem* **23**, 1623-
739 1641 (2002).
- 740 22 Jakalian, A., Bush, B. L., Jack, D. B. & Bayly, C. I. Fast, efficient generation of high-
741 quality atomic charges AM1-BCC model: I. Method. *J Comput Chem* **21**, 132-146 (2000).
- 742 23 Toolkits, O. Oct.1 OpenEye Scientific Software, Santa Fe, NM.
743 <http://www.eyesopen.com/>. (2017).
- 744 24 Nocedal, J. American Mathematical Society. *Math. Comp.* **35**, 773-782 (1980).
- 745 25 Eastman, P. *et al.* OpenMM 7: Rapid development of high performance algorithms for
746 molecular dynamics. *PLoS Comput Biol* **13**, e1005659 (2017).

- 747 26 Wang, K., Chodera, J. D., Yang, Y. & Shirts, M. R. Identifying ligand binding sites and
748 poses using GPU-accelerated Hamiltonian replica exchange molecular dynamics. *J*
749 *Comput Aided Mol Des* **27**, 989-1007 (2013).
- 750 27 Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An $N \log(N)$ method for Ewald
751 sums in large systems *J Chem Phys* **98**, 10089-10092 (1998).
- 752 28 Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L.
753 Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **79**,
754 926-935 (1998).
- 755 29 Joung, I. S. & Cheatham, T. E., 3rd. Determination of alkali and halide monovalent ion
756 parameters for use in explicitly solvated biomolecular simulations. *J Phys Chem B* **112**,
757 9020-9041 (2008).
- 758 30 Leimkuhler, B. & Matthews, C. Efficient molecular dynamics using geodesic integration
759 and solvent-solute splitting. *Proc Math Phys Eng Sci* **472**, 20160138 (2016).
- 760 31 Chodera, J. D. & Shirts, M. R. Replica exchange and expanded ensemble simulations as
761 Gibbs sampling: simple improvements for enhanced mixing. *J Chem Phys* **135**, 194110
762 (2011).
- 763 32 Shirts, M. R. & Chodera, J. D. Statistically optimal analysis of samples from multiple
764 equilibrium states. *J Chem Phys* **129**, 124105 (2008).
- 765 33 Chodera, J. D. A Simple Method for Automated Equilibration Detection in Molecular
766 Simulations. *J Chem Theory Comput* **12**, 1799-1805 (2016).
- 767 34 McGibbon, R. T. *et al.* MDTraj: A Modern Open Library for the Analysis of Molecular
768 Dynamics Trajectories. *Biophys J* **109**, 1528-1532 (2015).
- 769 35 Beauchamp, K. A. *et al.* MSMBuilder2: Modeling Conformational Dynamics at the
770 Picosecond to Millisecond Scale. *J Chem Theory Comput* **7**, 3412-3419 (2011).
- 771
- 772

773 **References**

- 774 1 Hentze, M. W., Castello, A., Schwarzl, T. & Preiss, T. A brave new world of RNA-binding
775 proteins. *Nat Rev Mol Cell Biol*, doi:10.1038/nrm.2017.130 (2018).
- 776 2 Kharas, M. G. & Lengner, C. J. Stem Cells, Cancer, and MUSASHI in Blood and Guts.
777 *Trends Cancer* **3**, 347-356, doi:10.1016/j.trecan.2017.03.007 (2017).
- 778 3 Pereira, B., Billaud, M. & Almeida, R. RNA-Binding Proteins in Cancer: Old Players and
779 New Actors. *Trends Cancer* **3**, 506-528, doi:10.1016/j.trecan.2017.05.003 (2017).
- 780 4 Vu, L. P. *et al.* Functional screen of MSI2 interactors identifies an essential role for
781 SYNCRIP in myeloid leukemia stem cells. *Nat Genet* **49**, 866-875, doi:10.1038/ng.3854
782 (2017).
- 783 5 Han, T. *et al.* Anticancer sulfonamides target splicing by inducing RBM39 degradation via
784 recruitment to DCAF15. *Science* **356**, doi:10.1126/science.aal3755 (2017).
- 785 6 Ghosh, M. *et al.* Essential role of the RNA-binding protein HuR in progenitor cell survival
786 in mice. *J Clin Invest* **119**, 3530-3543, doi:10.1172/JCI38263 (2009).
- 787 7 Palanichamy, J. K. *et al.* RNA-binding protein IGF2BP3 targeting of oncogenic transcripts
788 promotes hematopoietic progenitor proliferation. *J Clin Invest* **126**, 1495-1511,
789 doi:10.1172/JCI80046 (2016).
- 790 8 Park, S. M. *et al.* Musashi2 sustains the mixed-lineage leukemia-driven stem cell
791 regulatory program. *J Clin Invest* **125**, 1286-1298, doi:10.1172/JCI78440 (2015).
- 792 9 Lee, S. C. & Abdel-Wahab, O. Therapeutic targeting of splicing in cancer. *Nat Med* **22**,
793 976-986, doi:10.1038/nm.4165 (2016).
- 794 10 Kanemura, Y. *et al.* Musashi1, an evolutionarily conserved neural RNA-binding protein, is
795 a versatile marker of human glioma cells in determining their cellular origin, malignancy,
796 and proliferative activity. *Differentiation* **68**, 141-152 (2001).
- 797 11 Hemmati, H. D. *et al.* Cancerous stem cells can arise from pediatric brain tumors. *Proc*
798 *Natl Acad Sci U S A* **100**, 15178-15183, doi:10.1073/pnas.2036535100 (2003).
- 799 12 Shu, H. J. *et al.* Expression of the Musashi1 gene encoding the RNA-binding protein in
800 human hepatoma cell lines. *Biochem Biophys Res Commun* **293**, 150-154,
801 doi:10.1016/S0006-291X(02)00175-4 (2002).
- 802 13 Li, N. *et al.* The Msi Family of RNA-Binding Proteins Function Redundantly as Intestinal
803 Oncoproteins. *Cell Rep* **13**, 2440-2455, doi:10.1016/j.celrep.2015.11.022 (2015).
- 804 14 Wang, S. *et al.* Transformation of the intestinal epithelium by the MSI2 RNA-binding
805 protein. *Nat Commun* **6**, 6517, doi:10.1038/ncomms7517 (2015).
- 806 15 Oskarsson, T. *et al.* Breast cancer cells produce tenascin C as a metastatic niche
807 component to colonize the lungs. *Nat Med* **17**, 867-874, doi:10.1038/nm.2379 (2011).
- 808 16 Kang, M. H. *et al.* Musashi RNA-binding protein 2 regulates estrogen receptor 1 function
809 in breast cancer. *Oncogene* **36**, 1745-1752, doi:10.1038/onc.2016.327 (2017).
- 810 17 Wang, X. Y. *et al.* Musashi1 as a potential therapeutic target and diagnostic marker for
811 lung cancer. *Oncotarget* **4**, 739-750, doi:10.18632/oncotarget.1034 (2013).
- 812 18 Vo, D. T. *et al.* The oncogenic RNA-binding protein Musashi1 is regulated by HuR via
813 mRNA translation and stability in glioblastoma cells. *Mol Cancer Res* **10**, 143-155,
814 doi:10.1158/1541-7786.MCR-11-0208 (2012).
- 815 19 Guo, K. *et al.* The Novel KLF4/MSI2 Signaling Pathway Regulates Growth and
816 Metastasis of Pancreatic Cancer. *Clin Cancer Res* **23**, 687-696, doi:10.1158/1078-
817 0432.CCR-16-1064 (2017).
- 818 20 Fox, R. G. *et al.* Image-based detection and targeting of therapy resistance in pancreatic
819 adenocarcinoma. *Nature* **534**, 407-411, doi:10.1038/nature17988 (2016).
- 820 21 Barbouti, A. *et al.* A novel gene, MSI2, encoding a putative RNA-binding protein is
821 recurrently rearranged at disease progression of chronic myeloid leukemia and forms a
822 fusion gene with HOXA9 as a result of the cryptic t(7;17)(p15;q23). *Cancer Res* **63**, 1202-
823 1206 (2003).
- 824 22 De Weer, A. *et al.* EVI1 overexpression in t(3;17) positive myeloid malignancies results
825 from juxtaposition of EVI1 to the MSI2 locus at 17q22. *Haematologica* **93**, 1903-1907,
826 doi:10.3324/haematol.13192 (2008).

- 827 23 Saleki, R. *et al.* A novel TTC40-MSI2 fusion in de novo acute myeloid leukemia with an
828 unbalanced 10;17 translocation. *Leuk Lymphoma* **56**, 1137-1139,
829 doi:10.3109/10428194.2014.947611 (2015).
- 830 24 Wang, K. *et al.* Patient-derived xenotransplants can recapitulate the genetic driver
831 landscape of acute leukemias. *Leukemia* **31**, 151-158, doi:10.1038/leu.2016.166 (2017).
- 832 25 Ito, T. *et al.* Regulation of myeloid leukaemia by the cell-fate determinant Musashi.
833 *Nature* **466**, 765-768, doi:10.1038/nature09171 (2010).
- 834 26 Kharas, M. G. *et al.* Musashi-2 regulates normal hematopoiesis and promotes aggressive
835 myeloid leukemia. *Nat Med* **16**, 903-908, doi:10.1038/nm.2187 (2010).
- 836 27 Thol, F. *et al.* Prognostic significance of expression levels of stem cell regulators MSI2
837 and NUMB in acute myeloid leukemia. *Ann Hematol* **92**, 315-323, doi:10.1007/s00277-
838 012-1637-5 (2013).
- 839 28 Taggart, J. *et al.* MSI2 is required for maintaining activated myelodysplastic syndrome
840 stem cells. *Nat Commun* **7**, 10739, doi:10.1038/ncomms10739 (2016).
- 841 29 Byers, R. J., Currie, T., Tholouli, E., Rodig, S. J. & Kutok, J. L. MSI2 protein expression
842 predicts unfavorable outcome in acute myeloid leukemia. *Blood* **118**, 2857-2867,
843 doi:10.1182/blood-2011-04-346767 (2011).
- 844 30 Kwon, H. Y. *et al.* Tetraspanin 3 Is Required for the Development and Propagation of
845 Acute Myelogenous Leukemia. *Cell Stem Cell* **17**, 152-164,
846 doi:10.1016/j.stem.2015.06.006 (2015).
- 847 31 Park, S. M. *et al.* Musashi-2 controls cell fate, lineage bias, and TGF-beta signaling in
848 HSCs. *J Exp Med* **211**, 71-87, doi:10.1084/jem.20130736 (2014).
- 849 32 Rentas, S. *et al.* Musashi-2 attenuates AHR signalling to expand human haematopoietic
850 stem cells. *Nature* **532**, 508-511, doi:10.1038/nature17665 (2016).
- 851 33 Kudinov, A. E., Karanicolas, J., Golemis, E. A. & Bumber, Y. Musashi RNA-Binding
852 Proteins as Cancer Drivers and Novel Therapeutic Targets. *Clin Cancer Res* **23**, 2143-
853 2153, doi:10.1158/1078-0432.CCR-16-2728 (2017).
- 854 34 Sakakibara, S., Nakamura, Y., Satoh, H. & Okano, H. Rna-binding protein Musashi2:
855 developmentally regulated expression in neural precursor cells and subpopulations of
856 neurons in mammalian CNS. *J Neurosci* **21**, 8091-8107 (2001).
- 857 35 Zearfoss, N. R. *et al.* A conserved three-nucleotide core motif defines Musashi RNA
858 binding specificity. *J Biol Chem* **289**, 35530-35541, doi:10.1074/jbc.M114.597112 (2014).
- 859 36 Ohyama, T. *et al.* Structure of Musashi1 in a complex with target RNA: the role of
860 aromatic stacking interactions. *Nucleic Acids Res* **40**, 3218-3231,
861 doi:10.1093/nar/gkr1139 (2012).
- 862 37 Katz, Y. *et al.* Musashi proteins are post-transcriptional regulators of the epithelial-luminal
863 cell state. *Elife* **3**, e03915, doi:10.7554/eLife.03915 (2014).
- 864 38 Minuesa, G. *et al.* A 1536-well fluorescence polarization assay to screen for modulators
865 of the MUSASHI family of RNA-binding proteins. *Comb Chem High Throughput Screen*
866 **17**, 596-609 (2014).
- 867 39 Eibl, J. K., Strasser, B. C. & Ross, G. M. Identification of novel
868 pyrazoloquinazolinecarboxylate analogues to inhibit nerve growth factor in vitro. *Eur J*
869 *Pharmacol* **708**, 30-37, doi:10.1016/j.ejphar.2013.03.029 (2013).
- 870 40 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for
871 interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-
872 15550, doi:10.1073/pnas.0506580102 (2005).
- 873 41 Zhang, H. *et al.* Musashi2 modulates K562 leukemic cell proliferation and apoptosis
874 involving the MAPK pathway. *Exp Cell Res* **320**, 119-127,
875 doi:10.1016/j.yexcr.2013.09.009 (2014).
- 876 42 Han, Y. *et al.* Musashi-2 Silencing Exerts Potent Activity against Acute Myeloid Leukemia
877 and Enhances Chemosensitivity to Daunorubicin. *PLoS One* **10**, e0136484,
878 doi:10.1371/journal.pone.0136484 (2015).
- 879 43 Meisner, N. C. *et al.* Identification and mechanistic characterization of low-molecular-
880 weight inhibitors for HuR. *Nat Chem Biol* **3**, 508-515, doi:10.1038/nchembio.2007.14
881 (2007).

- 882 44 Wu, X. *et al.* Identification and validation of novel small molecule disruptors of HuR-
883 mRNA interaction. *ACS Chem Biol* **10**, 1476-1484, doi:10.1021/cb500851u (2015).
- 884 45 Lan, L. *et al.* Natural product (-)-gossypol inhibits colon cancer cell growth by targeting
885 RNA-binding protein Musashi-1. *Mol Oncol* **9**, 1406-1420,
886 doi:10.1016/j.molonc.2015.03.014 (2015).
- 887 46 Lim, D., Byun, W. G., Koo, J. Y., Park, H. & Park, S. B. Discovery of a Small-Molecule
888 Inhibitor of Protein-MicroRNA Interaction Using Binding Assay with a Site-Specifically
889 Labeled Lin28. *J Am Chem Soc*, doi:10.1021/jacs.6b06965 (2016).
- 890 47 Roos, M. *et al.* A Small-Molecule Inhibitor of Lin28. *ACS Chem Biol* **11**, 2773-2781,
891 doi:10.1021/acscchembio.6b00232 (2016).
- 892 48 Jarvis, W. D., Turner, A. J., Povirk, L. F., Traylor, R. S. & Grant, S. Induction of apoptotic
893 DNA fragmentation and cell death in HL-60 human promyelocytic leukemia cells by
894 pharmacological inhibitors of protein kinase C. *Cancer Res* **54**, 1707-1714 (1994).
- 895 49 Zhu, J. *et al.* Niemann-Pick C2 Proteins: A New Function for an Old Family. *Front Physiol*
896 **9**, doi:10.3389/fphys.2018.00052 (2018).
- 897 50 Judge, J. L. *et al.* The Lactate Dehydrogenase Inhibitor Gossypol Inhibits Radiation-
898 Induced Pulmonary Fibrosis. *Radiat Res* **188**, 35-43, doi:10.1667/RR14620.1 (2017).
- 899 51 Zeng, Y., Ma, J., Xu, L. & Wu, D. Natural Product Gossypol and Its Derivatives in
900 Precision Cancer Medicine. *Curr Med Chem*, doi:10.2174/0929867324666170523123655
901 (2017).
- 902 52 Clingman, C. C. *et al.* Allosteric inhibition of a stem cell RNA-binding protein by an
903 intermediary metabolite. *Elife* **3**, doi:10.7554/eLife.02848 (2014).
- 904 53 Sadlish, H. *et al.* Evidence for a functionally relevant rocaglamide binding site on the
905 eIF4A-RNA complex. *ACS Chem Biol* **8**, 1519-1527, doi:10.1021/cb400158t (2013).
- 906 54 Choo, A. Y., Yoon, S. O., Kim, S. G., Roux, P. P. & Blenis, J. Rapamycin differentially
907 inhibits S6Ks and 4E-BP1 to mediate cell-type-specific repression of mRNA translation.
908 *Proc Natl Acad Sci U S A* **105**, 17414-17419, doi:10.1073/pnas.0809136105 (2008).
- 909 55 Fang, T. *et al.* Musashi 2 contributes to the stemness and chemoresistance of liver
910 cancer stem cells via LIN28A activation. *Cancer Lett* **384**, 50-59,
911 doi:10.1016/j.canlet.2016.10.007 (2017).
- 912 56 Sheng, W. *et al.* Cooperation of Musashi-2, Numb, MDM2, and P53 in drug resistance
913 and malignant biology of pancreatic cancer. *FASEB J* **31**, 2429-2438,
914 doi:10.1096/fj.201601240R (2017).
- 915

916 **Figure Legends**

917 **Figure 1. Ro 08-2750 (Ro) is a novel selective MSI RNA-binding activity inhibitor.**

918 (a) Fluorescence polarization secondary validation of Ro 08-2750 (Ro) IC_{50} for MSI-RNA binding
919 inhibition in 384-well format. Seven independent experiments performed in duplicate \pm standard
920 error mean (s.e.m.) are shown; (b) Representative Electrophoresis Mobility Shift Assays (EMSA)
921 for GST- and GST-MSI2 proteins (125 and 250ng) using biotinylated-RNA oligo in the absence or
922 presence of unlabeled RNA (*left*); quantification of MSI2-RNA complexes of at five independent
923 experiments \pm s.e.m. is shown in bar graph (*right*); (c) EMSA for GST-MSI2 (125ng) in the
924 presence of increasing concentrations of Ro (5 to 40 μ M); quantification of RNA-protein
925 complexes of at least four independent experiments \pm s.e.m. is shown in bar graph (*right*); (d)
926 Microscale Thermophoresis (MST) assay showing interaction of Ro with GST-MSI2, GST-
927 MSI2/RNA complexes or the RRM-RBP control GST-SYNCRIP. Ro concentrations ranged from
928 0.0153 to 500 μ M. Affinity (K_D) values \pm s.e.m. (μ M) of three independent experiments are shown
929 as percentage of fraction bound. For (b) and (c): two-tailed Paired *t*-test; * p <0.05; ** p <0.01,
930 *** p <0.005, **** p <0.001.

931

932 **Figure 2. Ro 08-2750 interacts with the RNA-recognition motif 1 (RRM1) of MSI2 and its**
933 **analogs show minimal or null residual activity.**

934 (a) Global front view of the docked Ro 08-2750 (Ro) molecule in the RNA-binding site of human
935 MSI2 RRM1 based on the X-ray diffraction crystal structure obtained at 1.7 \AA resolution (RCSB
936 PDB 6DBP); (b) Lateral and close up (*inset*) view of Ro showing the most relevant interaction
937 residues (F66, F97 and R100) and the distances (\AA) between them and Ro closest atoms; (c) 2D
938 representation of residues involved in Ro binding showing F66 (hydrophobic stacking), K22 (H-
939 bonding), F97 (H-bonding with the backbone) and R100 (π -cation interaction) from RRM1 as
940 main interaction partners; (d) Microscale Thermophoresis (MST) assay showing affinity of
941 interaction of Ro with full-length GST-MSI2 WT (*red*), GST-MSI2 F97A (*cyan*) and GST-
942 F66A/F97A/R100A (*orange*). K_D values \pm standard deviation (μ M) of at least three independent
943 experiments are shown as percentage of fraction bound; (e) Chemical structures of Ro analogues
944 used in (f), (g), (h) and (i) panels. Ro-NGF (high affinity Neural Growth Factor -NGF- inhibitor, K_D
945 $(_{\text{NGF}}) = 1.7 \times 10^{-6}$ M) and Ro-OH (reduced form of Ro); (f) The cluster centers for Ro (*left*), RoOH,
946 (*center*) and Ro-NGF (*right*), derived using regular spatial clustering with a ligand RMSD cutoff of
947 1 \AA . Ro-NGF (*right*) showing a much larger number of clusters than Ro 08-2750 (*left*) or RoOH
948 (*center*). (g) Representative EMSA for GST-MSI2 (125ng) in the absence (DMSO) or presence of
949 Ro (20 μ M), Ro-OH (20 μ M) or unlabeled RNA oligo (1 μ M) and quantification of RNA-protein
950 complexes of at least three independent experiments (bar graph, *below*); (h) Representative

951 EMSA for GST-MSI2 (125ng) in the absence (DMSO) or presence of Ro (20 μ M), Ro-NGF (20
952 μ M) or unlabeled RNA oligo (1 μ M) and quantification of RNA-protein complexes of at least three
953 independent experiments (bar graph, *below*) \pm s.e.m.; (i) MST assays showing interaction of Ro,
954 Ro-OH and Ro-NGF with GST-MSI2 WT. Drug concentrations ranged from 0.0153 to 500 μ M. K_D
955 values \pm standard deviation (μ M) of at least three experiments are shown as percentage of
956 fraction bound; For (g) and (h), two-tailed Paired *t*-test; *ns*, not significant, * p <0.05, ** p <0.01,
957 *** p <0.005.

958

959 **Figure 3. Ro 08-2750 treatment leads to preferential increase in differentiation and**
960 **apoptosis in murine MLL-AF9 leukemic cells compared to Lin⁻Sca⁺cKit⁺ (LSK) cells.**

961 (a) Cytotoxicity assay (Cell-Titer Glo[®]) of Ro (*red*), Ro-OH (*cyan*) and Ro-NGF (*orange*) in MLL-
962 AF9+ BM cells. 50% Effective Concentration (EC_{50}) values, average of at least three independent
963 experiments \pm standard deviation are shown. (b) Flow cytometry representative histograms of
964 DMSO (*grey*) and 5 μ M Ro (*red*) treated MLL-AF9+ BM cells showing myeloid differentiation
965 markers (Mac1 and Gr1); bar graphs (*below*) show average (fold change increase) \pm standard
966 error mean of three independent experiments, performed in triplicate. Paired *t*-test, * p <0.05;
967 ** p <0.01. (c) Representative immunocytochemistry images of cytospun MLL-AF9+ BM cells
968 control (DMSO) or Ro treated (5 and 10 μ M) and stained by Eosin Y and Methylene Blue/ Azure
969 A. Scale, 50 μ m. (d) Apoptosis analysis by Annexin V+ (% population) for MLL-AF9+ BM cells
970 cultured in absence (DMSO, *black*) or presence of Ro 5 μ M (*light red*) or 10 μ M (*red*). Results
971 represent at least three independent experiments \pm s.e.m.. (e) Colony Formation Unit (CFU)
972 assay of MLL-AF9+ BM cells transduced with MSCV-IRES-BFP (MIB, control) or MSCV-IRES-
973 MSI2-BFP (MSI2-BFP) retroviral vectors. Results represent the average \pm s.e.m. of colony
974 numbers of at least five experiments performed in duplicate. (f) Representative immunoblot of
975 MLL-AF9+ BM MIB (*black bars*) and MSI2-BFP (*red bars*) cells (used in panel e) after DMSO or
976 10 μ M Ro treatment for 4h. β -ACTIN, loading control. (g) CFU assay of Lin⁻Sca⁺cKit⁺ (LSK)
977 versus MLL-AF9+ BM cells demonstrates Ro 08-2750 therapeutic window. Results represent the
978 average \pm s.e.m. of colony numbers of three experiments performed in duplicate. Two tailed
979 Paired *t*-test (b, d, e and g), * p <0.05, ** p <0.01, *** p <0.005.

980

981 **Figure 4. Ro 08-2750 treatment inhibits survival of human AML cell lines and patient cells.**

982 (a) Cytotoxicity assay (Cell-Titer Glo[®]) of Ro, Ro-OH and Ro-NGF in MOLM13 and K562 cells.
983 EC_{50} values average of three independent experiments \pm standard deviation is shown. (b) Mean
984 Fluorescence Intensity (MFI) fold changes of CD14 (myeloid marker, MOLM13) and CD235a
985 (Glycophorin-A; erythroid marker, K562) after 48h treatment with DMSO (control, *black bars*) or

986 Ro 20 μM (*red bars*). Data is normalized to DMSO control cells. Representative histograms are
987 shown in *Extended Data Figure 6a*. (c) Representative immunocytochemistry images of cytospun
988 MOLM13 and K562 cells treated for 48h with DMSO (control) or Ro 20 μM and stained with Eosin
989 Y and Methylene Blue/ Azure A. Scale, 20 μm . (d) Apoptosis analysis by Annexin V+ (%
990 population). MOLM13 and K562 were cultured in DMSO (*black bars*) or in the presence of Ro 20
991 μM (*red bars*) for the indicated times and Annexin V positivity and 7AAD was measured. Results
992 represent three independent experiments \pm standard deviation. (e) CFU assay of MOLM13 and
993 K562 in the presence of Ro 08-2750 at different concentrations (1, 5, 10 and 20 μM). Data is
994 shown as average colony numbers (normalized to DMSO control) \pm s.e.m. of at least three
995 independent experiments. (f) CFU assay of cord-blood derived CD34+ HSPCs and AML patient
996 BM cells. Data is shown as average colony numbers (normalized to DMSO) \pm s.e.m. of three
997 different blood donors for CD34+ and three independent AML patients. Two tailed Paired *t*-test
998 (DMSO vs Ro treated, unless indicated with lines); * $p < 0.05$; ** $p < 0.01$; *** $p < 0.005$, **** $p < 0.001$.

999

1000 **Figure 5. Ro 08-2750 treatment resembles gene signature from MSI2 depleted cells and**
1001 **demonstrates inhibition of MSI2 target translation.**

1002 (a) Scheme of RNA-immunoprecipitation (IP) protocol followed with K562-MIG (MSCV-IRES-
1003 GFP) or FLAG-MSI2 overexpressing cells. (b) Ro 08-2750 inhibitory effect in the RNA-IP
1004 enrichment of MSI2 mRNA targets in K562-FLAG-MSI2 versus K562-MIG after 1h treatment at
1005 10 μM . Data is shown as average of inhibition effect (normalized to DMSO cells) \pm s.e.m. of four
1006 independent experiment. (c) Up-regulated and down-regulated gene sets obtained by RNA-seq
1007 analysis after 20 μM Ro 4h treatment in K562 and MOLM13 cells showing identical signature as
1008 previously obtained using shRNA against MSI2 in CML-BC and AML lines²⁶. (d) Venn diagram
1009 showing gene Set Enrichment Analysis (GSEA) overlap between MOLM13 (*red*), K562 (*blue*)
1010 (after 20 μM Ro 4h treatment) and AML/CML-BC cell lines MSI2 depleted with shRNAs (*yellow*)
1011 from²⁶. Bold values inside brackets below each group are total gene sets numbers. (e)
1012 Representative immunoblot for K562 treated with Ro at different concentrations (1, 5, 10 and 20
1013 μM) for 4h showing expression of MSI2 targets. HOXA9 is not expressed in this BCR-ABL+
1014 (CML-BC) leukemia cell line. (f) Representative immunoblot for MOLM13 treated with Ro at
1015 different concentrations for 4h showing expression of MSI2 targets. (g) Representative
1016 immunoblot for K562 treated with Ro 20 μM at different time points (1, 4, 12 and 24h) showing
1017 expression of same MSI2 targets as in panel (e). P21 and β -ACTIN from a different
1018 representative gel are shown. (h) Representative immunoblot for MOLM13 treated with Ro 20 μM
1019 at different time points showing effect on MSI2 targets.

1020

1021 **Figure 6. Ro 08-2750 demonstrates efficacy in inhibiting leukemogenesis in short-time and**
1022 **long-term treatment in a MLL-AF9 *in vivo* model.**

1023 (a) Scheme of pharmacodynamics marker experiments with Ro short-time points performed with
1024 MLL-AF9+ secondary BM cells. 10,000 MLL-AF9 GFP+ cells were transplanted and, after 3
1025 weeks, mice were injected with DMSO or Ro (13.75 mg/kg) and were sacrificed for analysis after
1026 4h and 12h (b) Surface flow analysis of c-Kit receptor in spleen cells of Ro at 4h and 12h *versus*
1027 DMSO treated mice. Results are represented as MFI of cKit-PE-Cy7 normalized to DMSO group.
1028 Each data point is an independent treated mouse. Mean \pm s.e.m. is shown. (c) Intracellular (IC)
1029 flow analysis of c-MYC expression in spleen cells of Ro at 4h and 12h *versus* DMSO treated
1030 mice. Results are represented as MFI of c-MYC normalized to DMSO group. Each data point is
1031 an independent treated mouse. Mean \pm s.e.m. is shown; (a-c, DMSO and Ro 4h, $n=9$; Ro 12h,
1032 $n=6$). (d) Scheme of *in vivo* Ro treatment in MLL-AF9+ model of myeloid leukemia. 10,000 MLL-
1033 AF9 GFP+ cells were transplanted and after 3 days, mice were injected with DMSO or Ro 13.75
1034 mg/kg (in DMSO) intraperitoneally (IP) at days 1, 4, 7, 10 and 13 (one day *on*, two days *off* drug).
1035 At day 19 of treatment, mice were sacrificed for organ weight and flow cytometry analysis of
1036 disease burden and MSI2 target, c-MYC. (e) Spleen weights at time of sacrifice. Results are
1037 represented in weight (g) and each data point represents an individual DMSO or Ro treated
1038 mouse. (f) White blood cell (WBC) counts ($K/\mu L$) at time of sacrifice. Each data point represents
1039 an individually treated mouse. (g) Intracellular (IC) flow analysis of c-MYC expression in spleen
1040 cells of Ro vs DMSO treated mice. Results are represented as % frequency (% freq) of c-MYC+
1041 cells. Each data point is an independent treated mouse. Mean \pm s.e.m. is shown. (d-g, DMSO,
1042 $n=9$; Ro, $n=8$). For all graphs, Unpaired *t*-test; * $p<0.05$, ** $p<0.005$.

1043

1044 **Extended Data Figure Legends**

1045 **Extended Data Figure 1. Ro 08-2750 (Ro) binds to RRM1 and SYNCRIP RRM identities.**

1046 (a) MicroScale Thermophoresis (MST) assay showing interaction of Ro with GST-RRM1 (hMSI2).
1047 Ro concentrations ranged from 0.0153 to 500 μM . K_D values \pm s.e.m. (μM) of at least three
1048 experiments are shown as percentage of fraction bound. (b) Sequence alignment of RRM1
1049 (*above*) and RRM2 (*below*) of human MSI2, MSI1 and SYNCRIP. Numbers indicate crucial RNA-
1050 binding conserved residues (in bold red) in hMSI2 (e.g. F24, corresponding to F23 in hMSI1,
1051 F165 in SYNCRIP). Grey highlights indicate conserved amino acids.

1052

1053 **Extended Data Figure 2. Ro 08-2750 docking and interacting residues in comparison with**
1054 **Ro-OH and Ro-NGF in the RNA-binding site of RRM1.**

1055 (a) Ro docked in the RNA-binding site of MSI2 RRM1 with interacting residues. Distances shown
1056 in \AA ; (b) MST experiments showing GST-MSI2 WT (*red*), F97A (*cyan*) and Triple
1057 (F66A/F97A/R100, *orange*) mutants interaction to MSI2 RNA oligo (4 MSI motifs; 15-nt). K_D
1058 values \pm s.e.m. of at least three experiments are shown (μM); (c) Chemical synthesis scheme of
1059 Ro-OH from Ro 08-2750 compound (see *Methods*); (d) FP confirmation of Ro, Ro-OH and Ro-
1060 NGF MSI2-RNA binding inhibition in 384-well format. IC_{50} values of two independent experiments
1061 performed in triplicate with s.e.m., $2.0 \pm 0.3 \mu\text{M}$ (Ro, *red*) and $25.0 \pm 8.0 \mu\text{M}$ (RoOH, *cyan*). Ro-NGF
1062 (*orange*) showed null inhibition of RNA-binding activity; (e) Docked pose of Ro-OH in the RNA
1063 binding site of MSI2 RRM1. Distances in \AA ; (f) Docked pose of Ro-NGF in the RNA binding site of
1064 MSI2 RRM1 showing a displaced center of the small-molecule from the binding site; (g) 2D
1065 representation of Ro-OH docked pose in the RRM1 of MSI2; (h) 2D representation of Ro-NGF
1066 docked pose in the RRM1 of MSI2 showing H-bonding of K22 changing from the O to the N in the
1067 middle ring, and π -cation interacting with R100 displaced with respect to Ro (see **Figure 2b**).

1068

1069 **Extended Data Figure 3. ^1H NMR (a) and NMR ^{13}C spectrum (b) of Ro-OH, the synthesized**
1070 **reduced form of Ro.**

1071

1072 **Extended Data Figure 4. Alchemical free energy calculations show that both protein and**
1073 **ligands adopt a conformationally heterogeneous ensemble of binding poses.**

1074 (a) Computed binding free energy (ΔG_{bind} , kcal/mol) estimates from alchemical free energy
1075 calculations (*y-axis*) for Ro, Ro-OH, and Ro-NGF for different definitions of the “bound” complex
1076 as a function of distance cutoff (*x-axis*, in \AA). Reported statistical errors and error bars correspond
1077 one standard error. The inset ΔG_{bind} was calculated for a cutoff of 20\AA . (b) In the alchemical

1078 Hamiltonian replica exchange simulations, a conformational change is induced when MSI2 is
1079 bound (“Complex”; *green*) to Ro (*right*) or Ro-OH (*center*), as compared to apo MSI2 (“Apo”;
1080 *gray*). Ro-NGF (*left*) does not induce the same conformational change. (c) The top three most
1081 populous clusters for Ro 08-2750. The protein structure and solid-color ligand pose depict cluster
1082 centers, while transparent ligand poses depict 10 randomly sampled frames assigned to that
1083 cluster. Sidechains within 4Å of any of the ligands are shown as lines. (d) The top four most
1084 populous clusters for Ro-OH, using the same depiction scheme as (c).

1085

1086 **Extended Data Figure 5. Flow cytometry plots showing apoptosis in MLL-AF9 leukemic**
1087 **cells after treatment with Ro 08-2750.**

1088 Apoptosis plots (graphs in **Figure 3d**) showing Annexin V+ and 7AAD (live/dead staining) by
1089 Apoptosis MUSE® Cell kit and MUSE® Cell Analyzer (Millipore-Sigma) in MLL-AF9+ BM cells at
1090 8, 16, 24 and 48 hours post treatment with Ro 5 and 10 µM.

1091

1092 **Extended Data Figure 6. Differentiation and apoptosis are induced in MOLM13 and K562**
1093 **cells after Ro 08-2750 treatment.**

1094 (a) Representative histograms showing CD14 and CD13 myeloid markers in MOLM13 and
1095 erythroid differentiation markers CD235a (Glycophorin-A) and CD71 in K562 after 48h of 20 µM
1096 Ro treatment. (b) Mean Fluorescence Intensity (MFI) fold changes of CD13 (myeloid marker,
1097 MOLM13) and CD71 (erythroid marker, K562) after 48h treatment of leukemia cell lines with
1098 DMSO (control, *black bars*) or Ro 20 µM (*red bars*). Data is shown as average (normalized to
1099 DMSO control cells) ± standard error mean of three independent experiments performed in
1100 triplicate. Paired *t*-test (DMSO vs Ro treated); **p*<0.05. (c) Apoptosis plots (from graphs in **Figure**
1101 **4d**) showing Annexin V+ and 7AAD (live/dead staining) in MOLM13 and K562 by MUSE® Cell
1102 Analyzer (Millipore-Sigma) in DMSO and Ro 20 µM treatments at 48, 72 and 96h.

1103

1104 **Extended Data Figure 7. Ro effects on global translation and mRNA of MSI2 targets**

1105 (a) OP-Puromycin incorporation to assess global translation rates in MOLM13 leukemia cells.
1106 Results are represented as average of Alexa Fluor 647 (AF647) Mean Fluorescence Intensity
1107 (MFI) normalized to DMSO control cells ± standard error mean of four independent experiments
1108 performed in duplicate. Paired *t*-test (DMSO vs Ro treated); *ns*, non-significant, ***p*<0.005. (b)
1109 Expression levels of mRNA targets of MSI2 by qPCR in K562 and (c) MOLM13. Cells were
1110 treated for 4h at 20 µM Ro. Results represent the average of ten independent experiments ±
1111 standard error mean. Paired *t*-test (DMSO vs Ro treated); **p*<0.05.

1112

1113 **Extended Data Figure 8. No toxicity of Ro 08-2750 after in vivo treatment of MLL-AF9 mice.**

1114 (a) Mice weight in DMSO (*cyan lines*, left panel) and Ro 13.75 mg/kg (*orange lines*, right panel)
1115 groups during the duration of the *in vivo* experiment. (b) Red Blood Cell (RBC) counts (M/ μ L) at
1116 time of sacrifice. Each data point represents an individually treated mouse. Unpaired *t*-test; *ns*,
1117 non-significant. (c) Platelets counts (PLT) counts (K/ μ L) at time of sacrifice. Each data point
1118 represents an individually treated mouse. Unpaired *t*-test; *ns*, non-significant. DMSO, *n*=9; Ro,
1119 *n*=8.

1120

1121 **Acknowledgements**

1122 We would like to thank the current members of the M. Kharas and R. Levine labs, R. Levine, P.
1123 Lito, J. Vidigal, P. Rocha, C. Lengner, A. Kentsis and K. Keshari for their critical advice and
1124 helpful suggestions. H. Djaballah, R. Garippa and the members of the former High-Throughput
1125 Screening Core and current RNAi Core (MSKCC) for their technical and project support
1126 throughout the screen and secondary validation. We would like to thank R. Sridharan, Y.
1127 Shamay, D. Heller (MSKCC) and Rui Liang (Tri-TDI) for their scientific advice and technical
1128 support. We would also like to thank A. Gruet and the Epigenetics core for his technical support
1129 with RNA-sequencing library preparation and J. Charavalli (HTSRC, The Rockefeller University)
1130 for her technical support with initial FP measurements. M.G.K. was supported by US National
1131 Institutes of Health National Institute of Diabetes, Digestive and Kidney Diseases Career
1132 Development Award, NIDDK NIH R01-DK101989-01A1, NCI 1R01CA193842-01, Louis V.
1133 Gerstner Young Investigator Award, American Society of Hematology Junior Scholar Award,
1134 Kimmel Scholar Award, V-Scholar Award, Geoffrey Beene Award and Alex's Lemonade Stand A
1135 Award and the Starr Foundation. The research was funded by Mr. William H. Goodwin and Mrs.
1136 Alice Goodwin and the Commonwealth Foundation for Cancer Research and The Center for
1137 Experimental Therapeutics at Memorial Sloan Kettering Cancer Center. The research was funded
1138 in part through NIH/NCI Cancer Support Core Grant P30 CA08748 to M.G.K. J.D.C.
1139 acknowledges support from the Sloan Kettering Institute, a Louis V. Gerstner Young Investigator
1140 Award, NIH grant P30 CA008748, and NIH grant R01 GM121505. L.N.N. acknowledges support
1141 from Merck KGaA to support the development of open source tools for GPU-accelerated
1142 alchemical free energy calculations. M.C.P was supported by a Medical Scientist Training
1143 Program grant from the National Institute of General Medical Sciences of the National Institutes of
1144 Health under award number T32GM007739 to the Weill Cornell/Rockefeller/Sloan Kettering Tri-
1145 Institutional MD-PhD Program. The MSKCC structural biology core laboratory is supported by
1146 National Cancer Institute grant P30-CA008748. X-ray diffraction data were collected at
1147 synchrotron facilities supported by grants and contracts from the National Institutes of Health
1148 (P41GM103403, HEI-S10RR029205) and the Department of Energy (DE-AC02-06CH11357).

1149

1150 **Author Contributions**

1151 G.M. led the project, performed experiments, analyzed data and wrote the manuscript. M.G.K.
1152 directed the project, analyzed data and wrote the manuscript. S.A., D.C., T.B., A.R., L.N., J.C.
1153 performed experiments, analyzed data and provided project support. A.C., A.S., S.M.P., T.C., J.T.
1154 performed experiments and analyzed data. M.C.P., L.F., C.L. analyzed data. J.S., C.F., M.P.
1155 provided clinical data and analysis. C.Z.R. and D.T. performed experiments and provided critical

1156 reagents. J.E. and G.M.R. provided critical reagents. Y.G. performed experiments and analyzed
1157 data. C.A. and J.F.G. provided suggestions, project support and assisted analyzing data.

1158

1159 **Author Information**

1160 J.D.C. is a member of the Scientific Advisory Board for Schrödinger.

1161

1162

1163

Figure 1

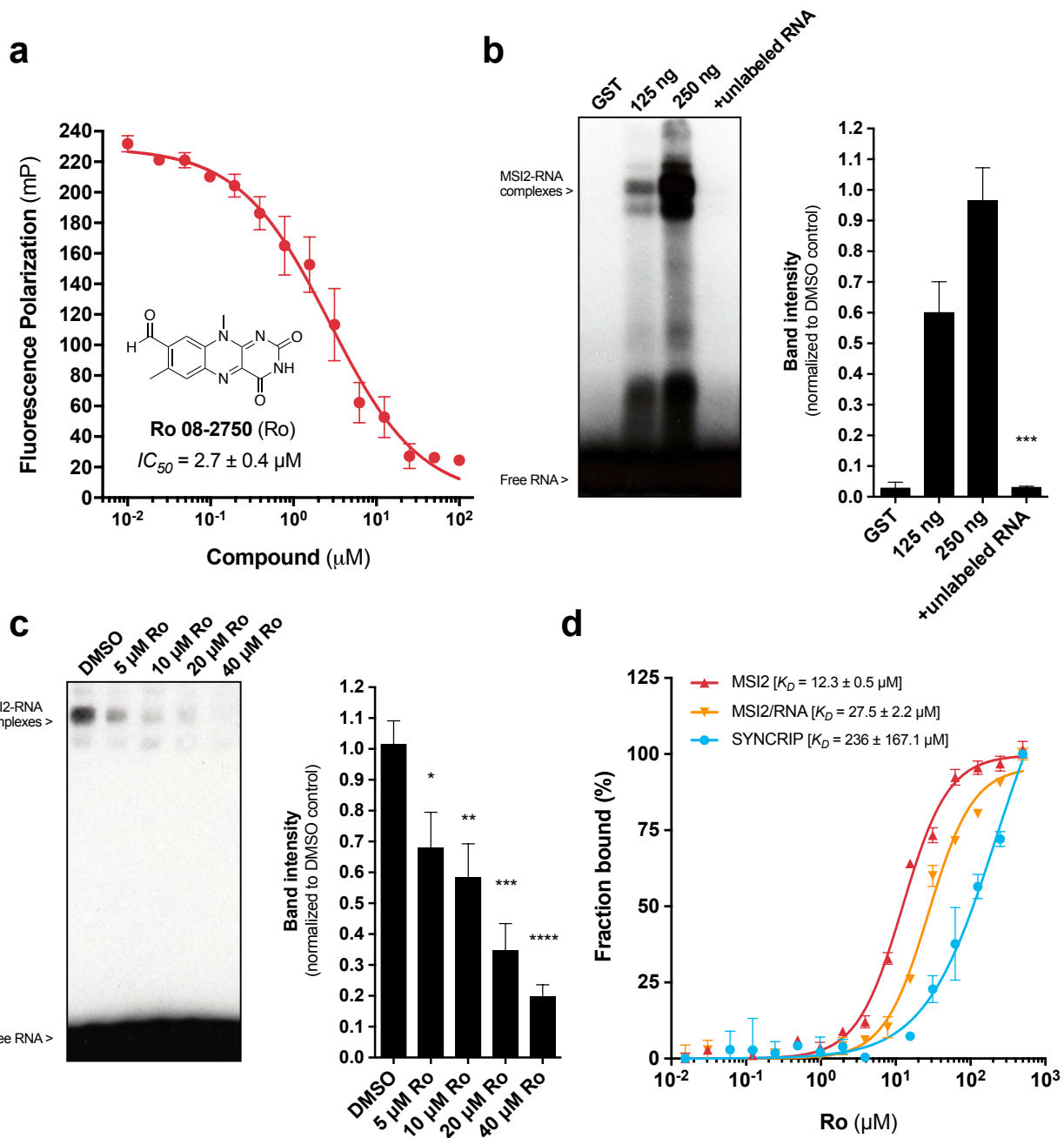


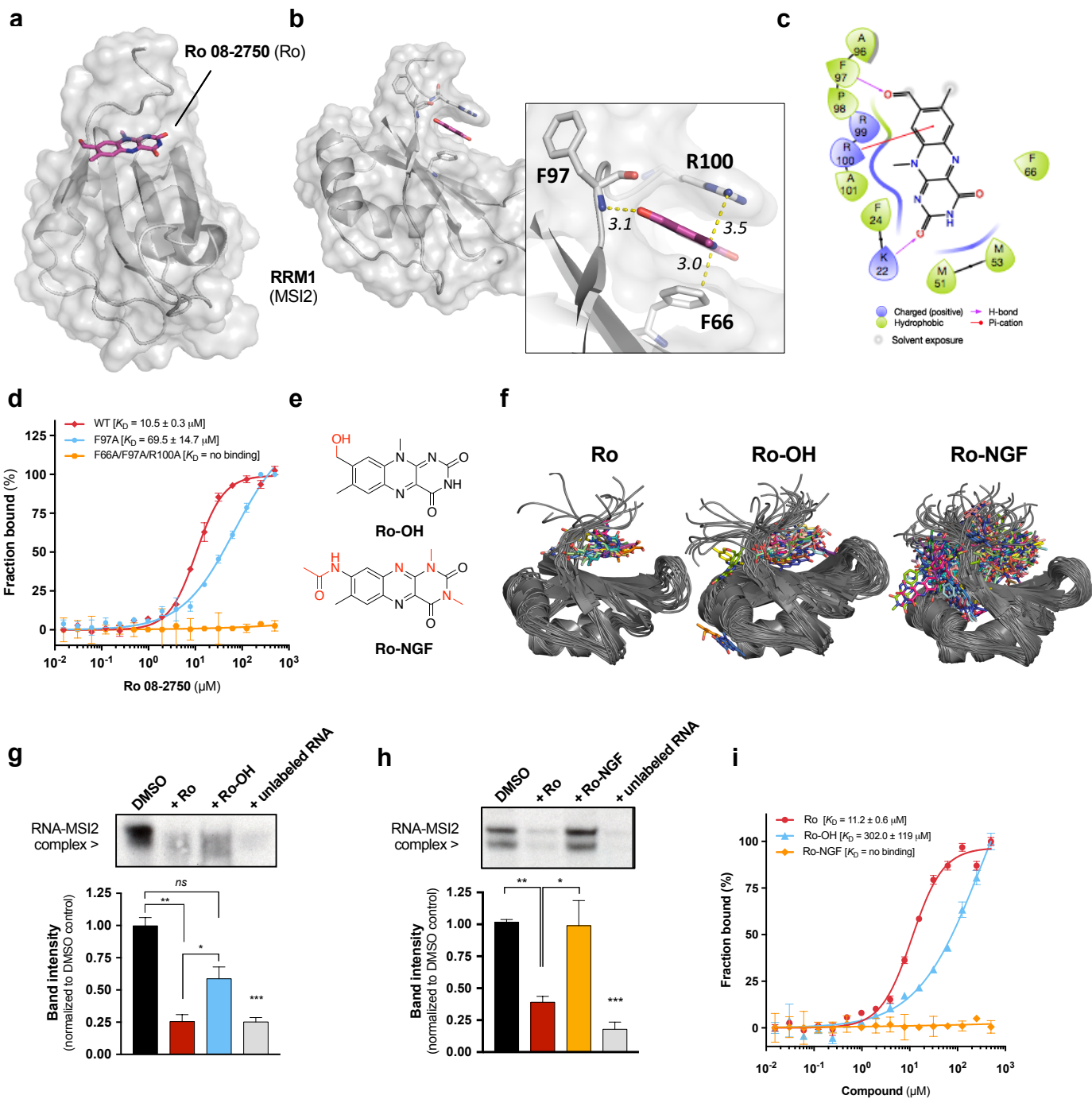
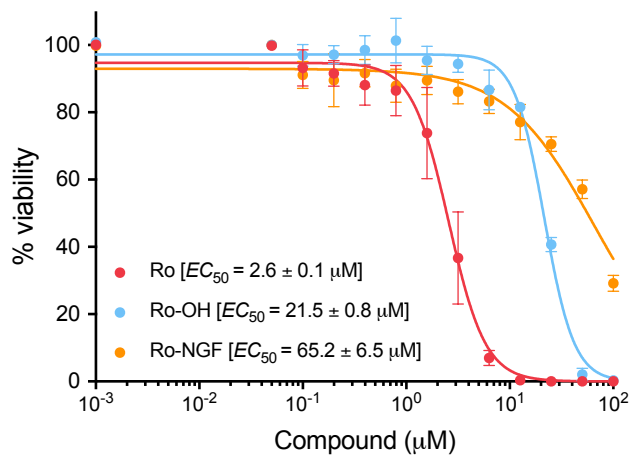
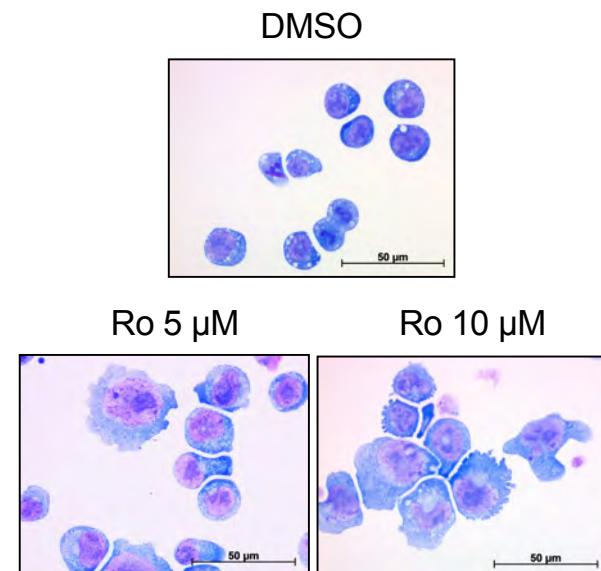
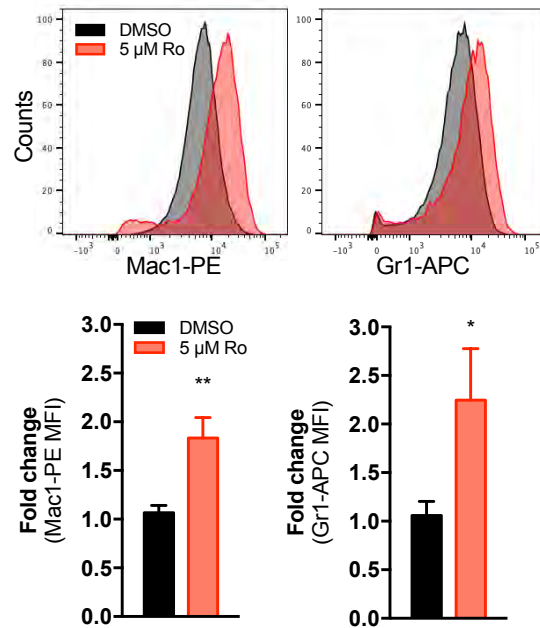
Figure 2

Figure 3

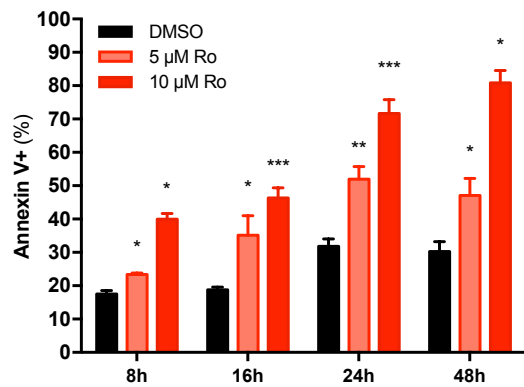
a



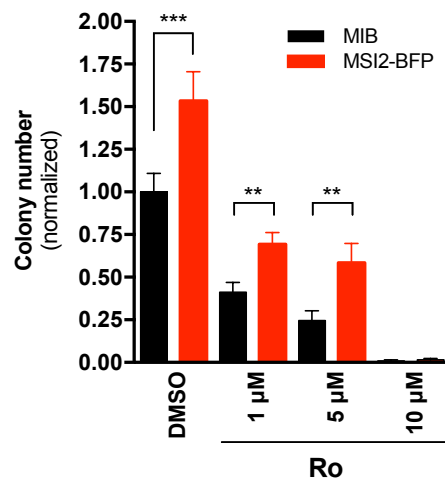
c



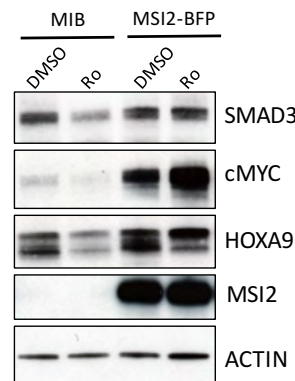
d



e



f



g

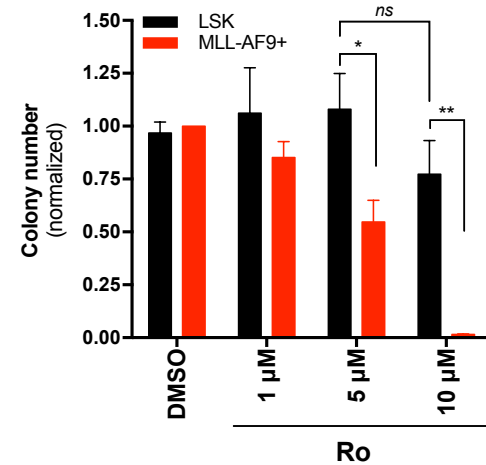
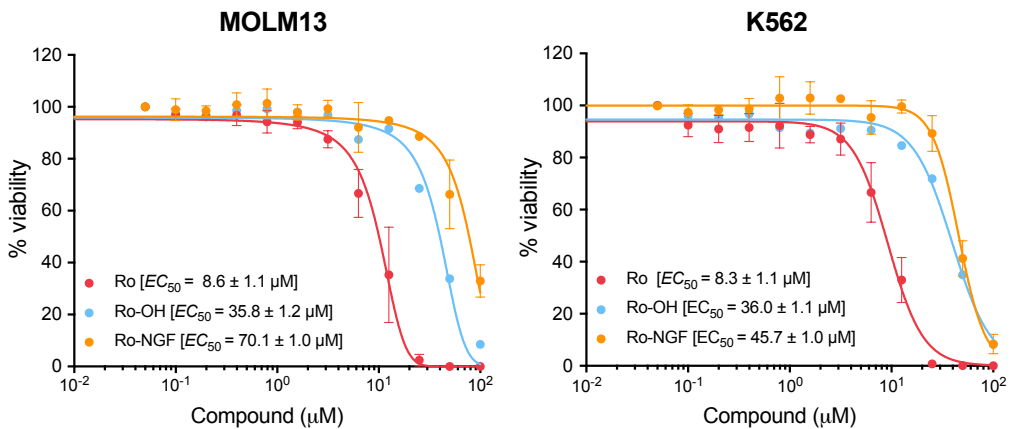
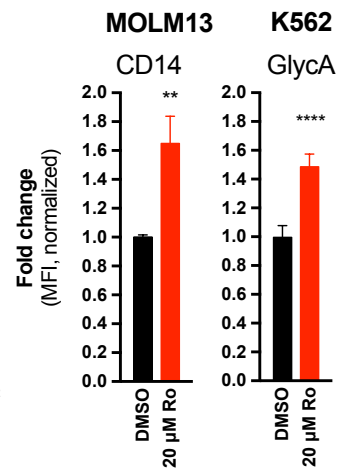


Figure 4

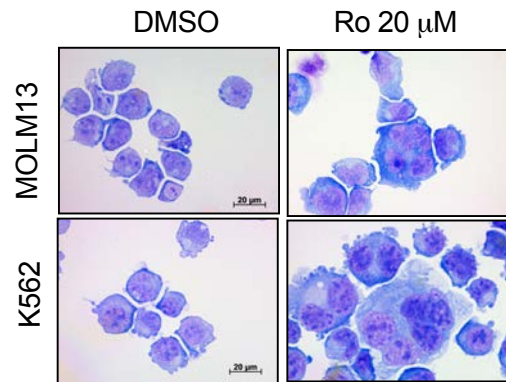
a



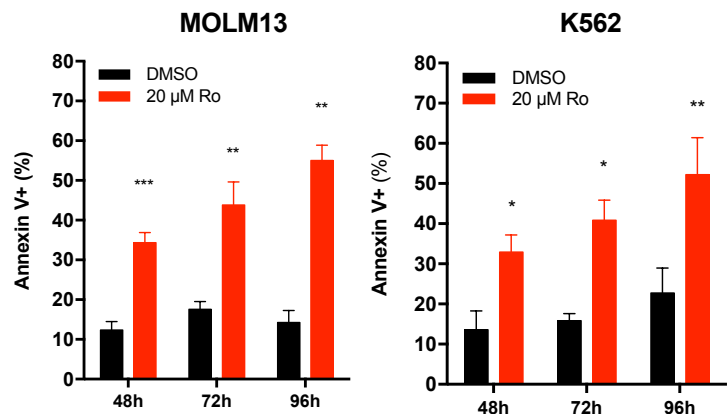
b



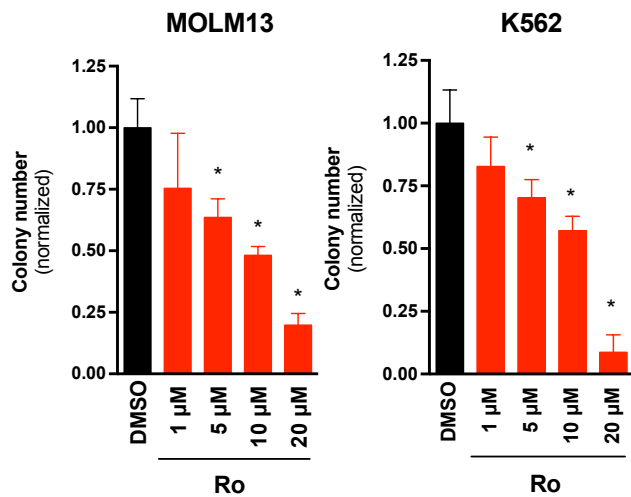
c



d



e



f

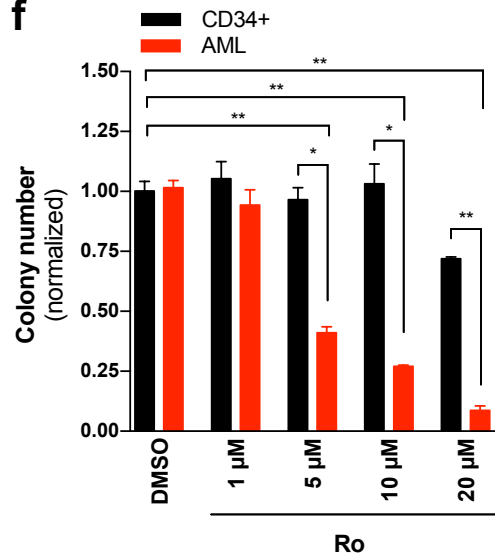


Figure 5

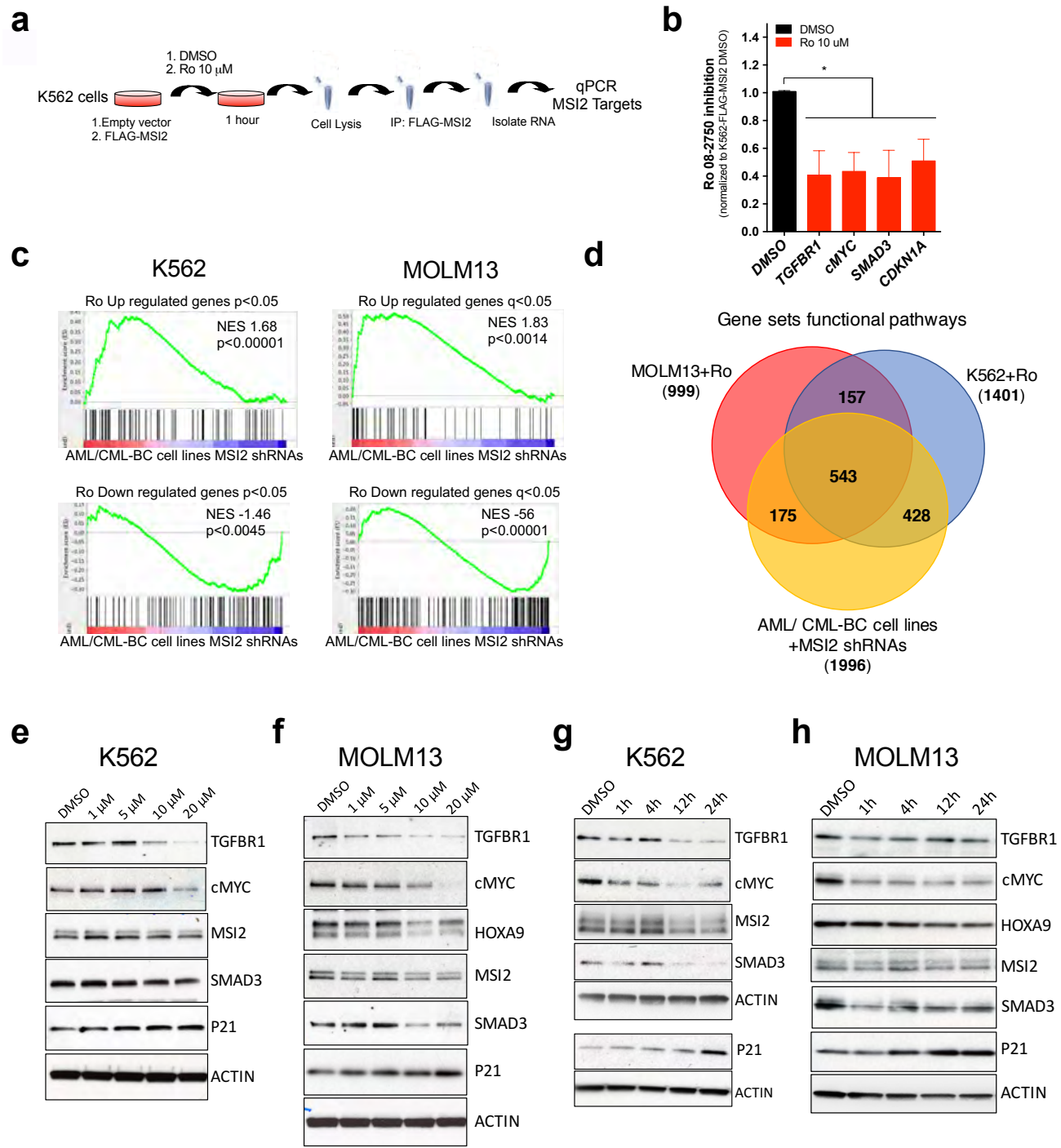
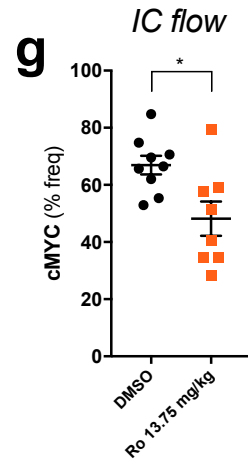
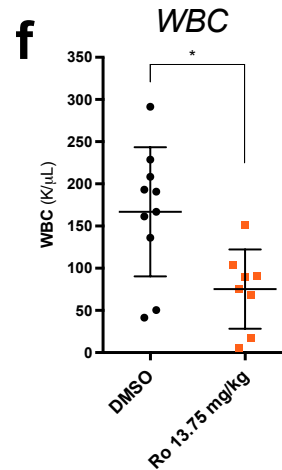
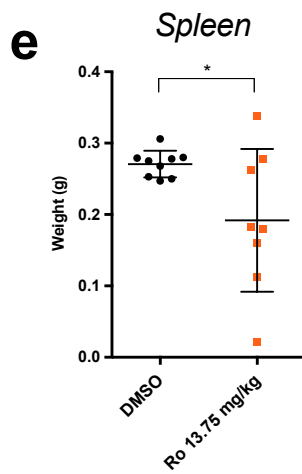
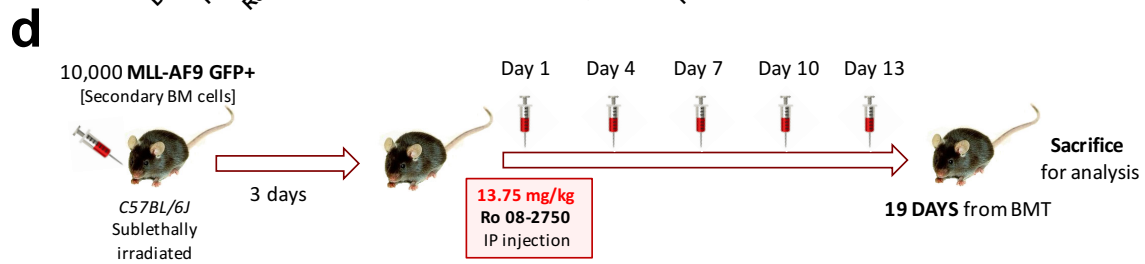
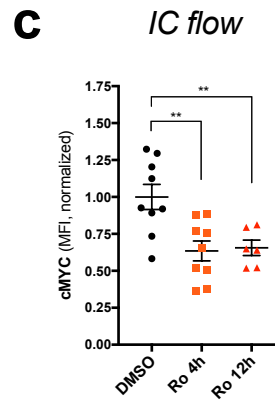
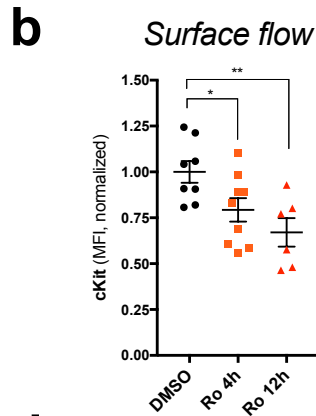
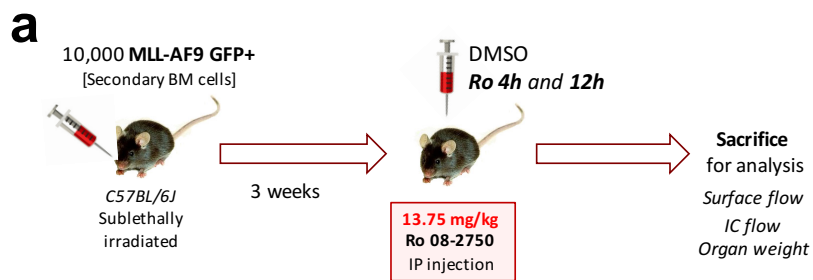
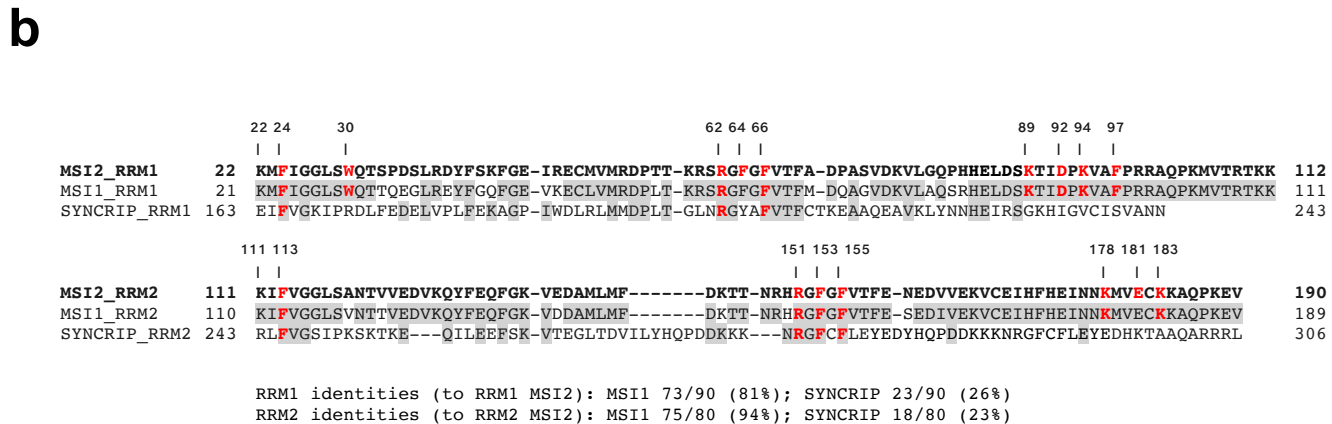
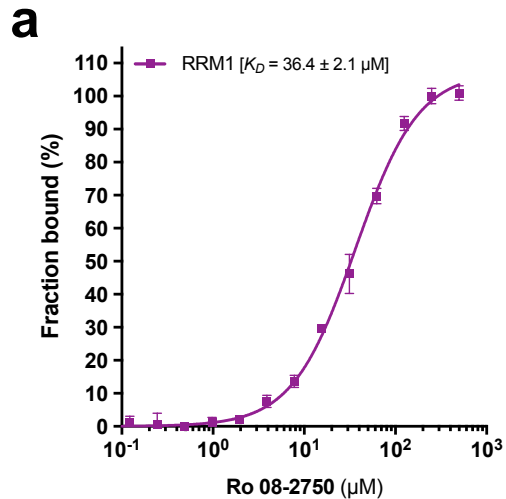


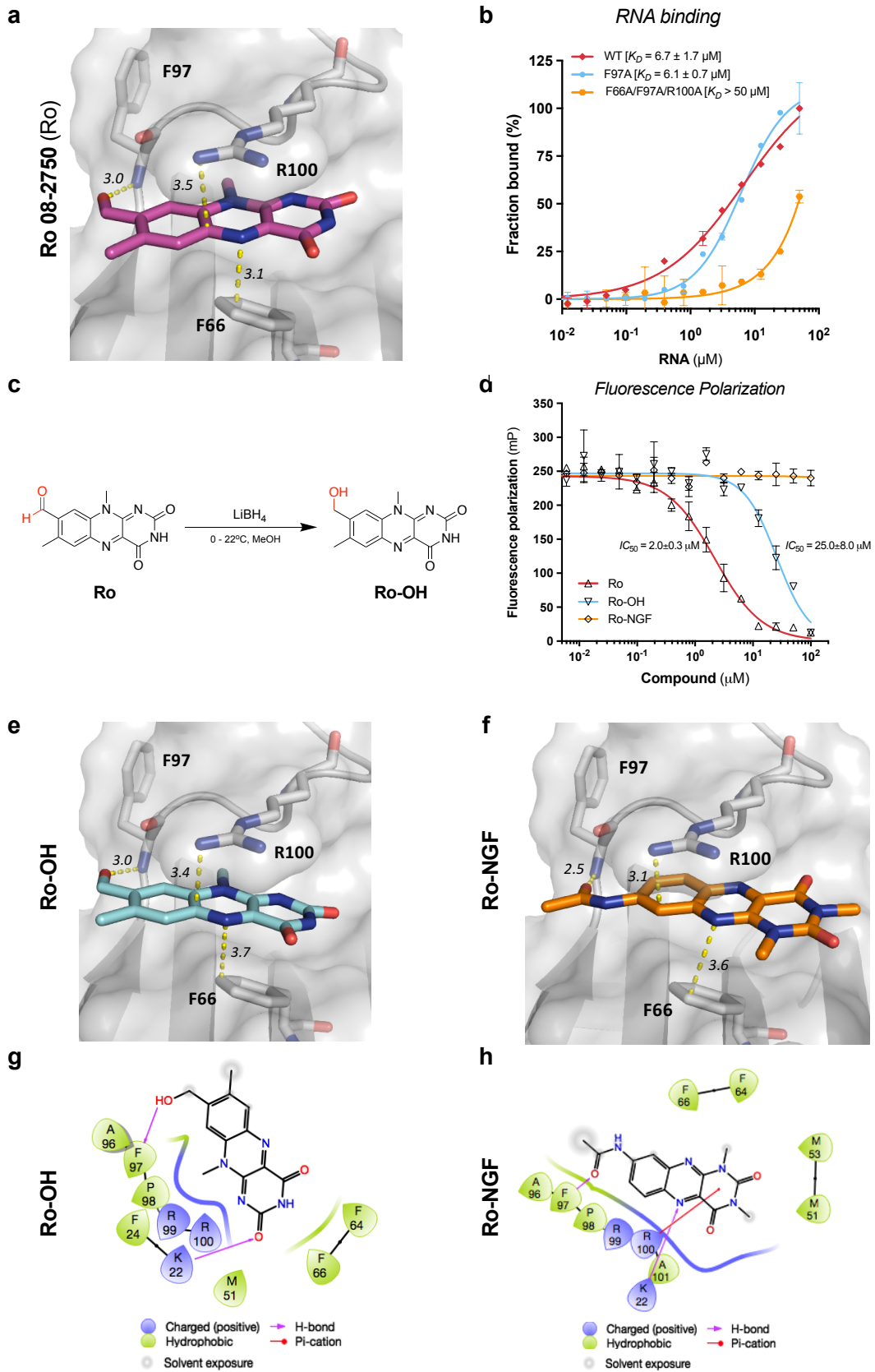
Figure 6



Extended Data Figure 1



Extended Data Figure 2

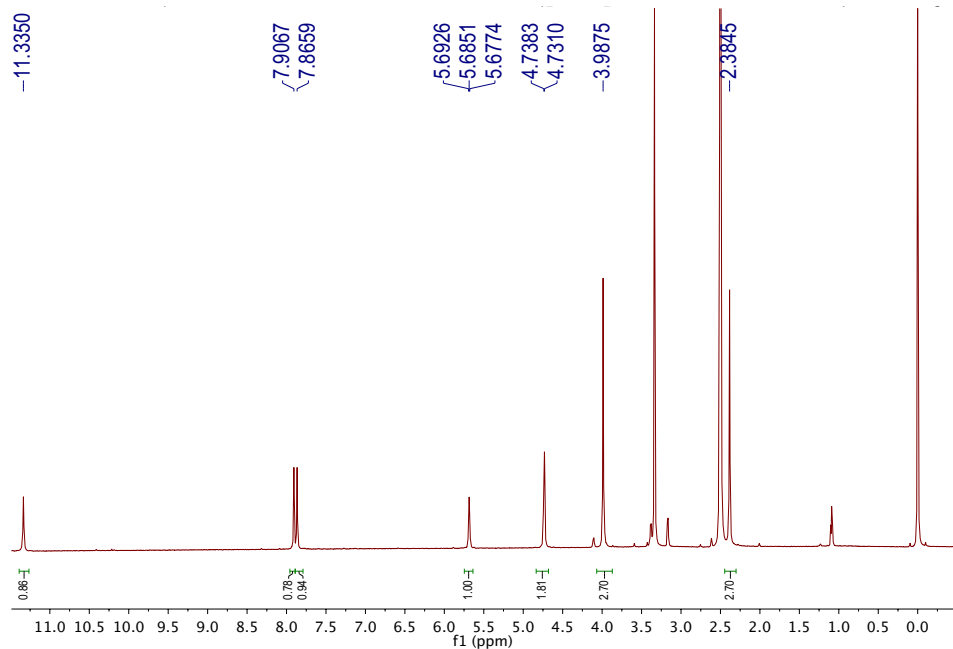


Extended Data Figure 3

Ro-OH

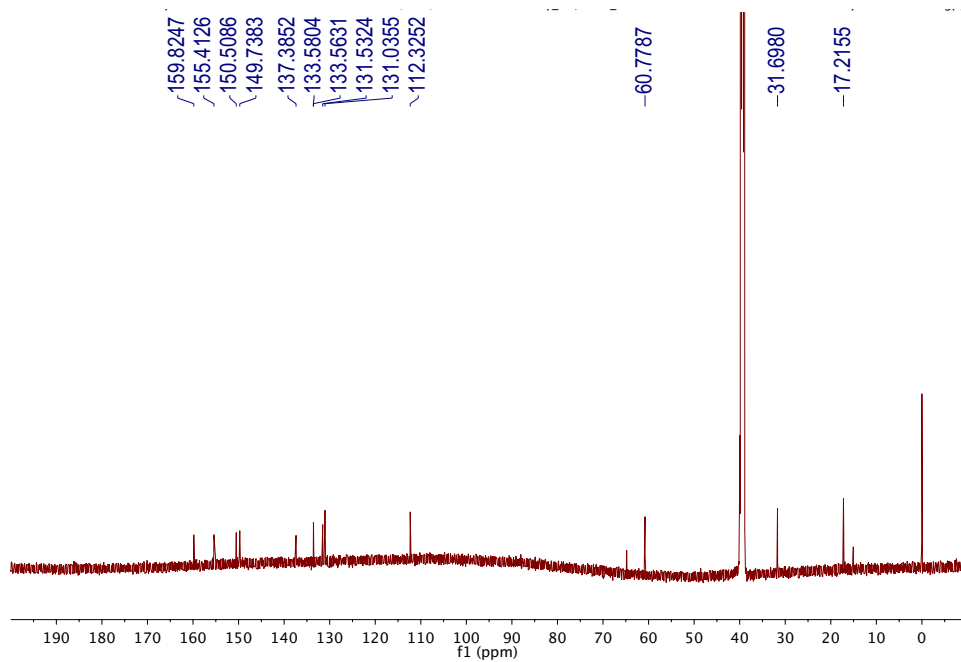
a

NMR (^1H) spectrum

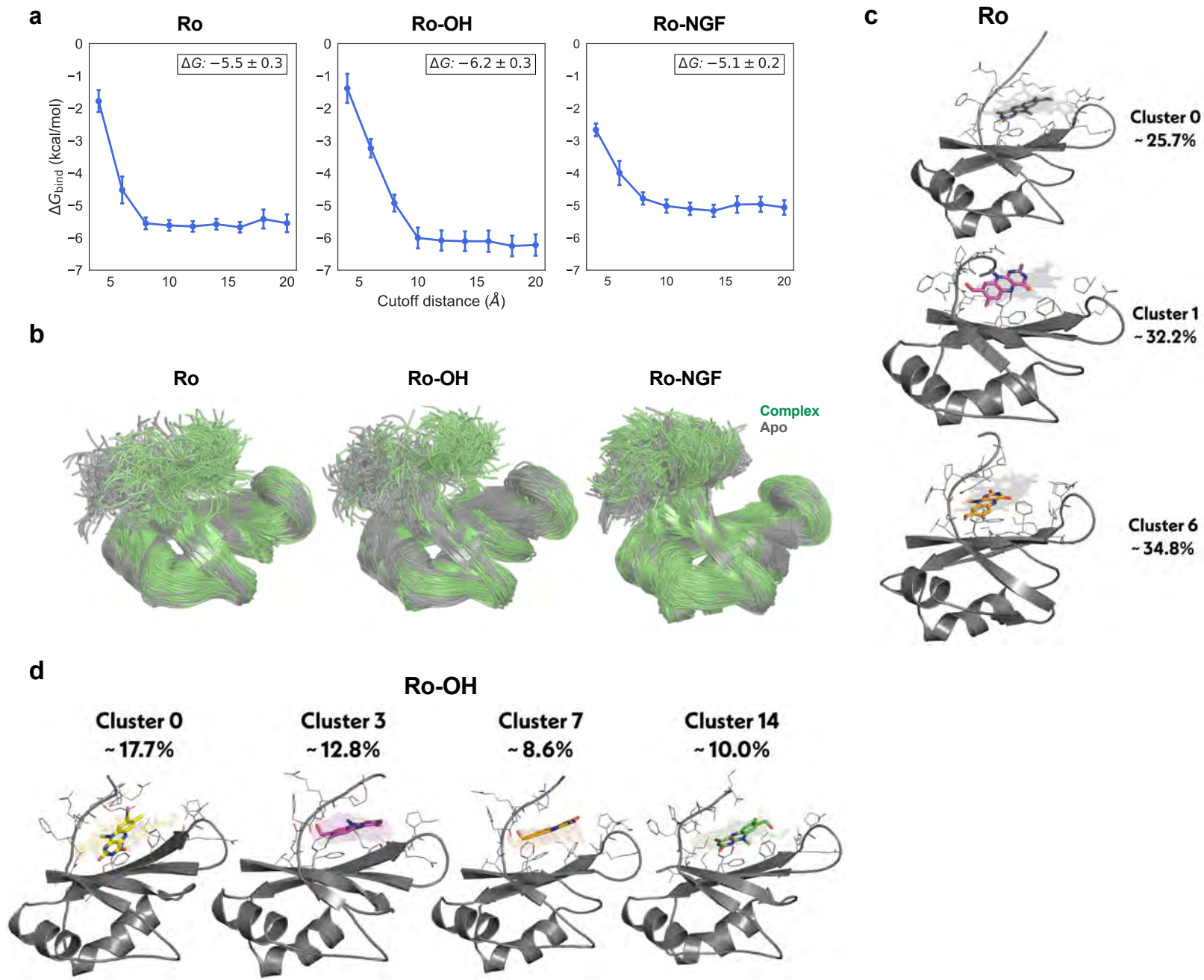


b

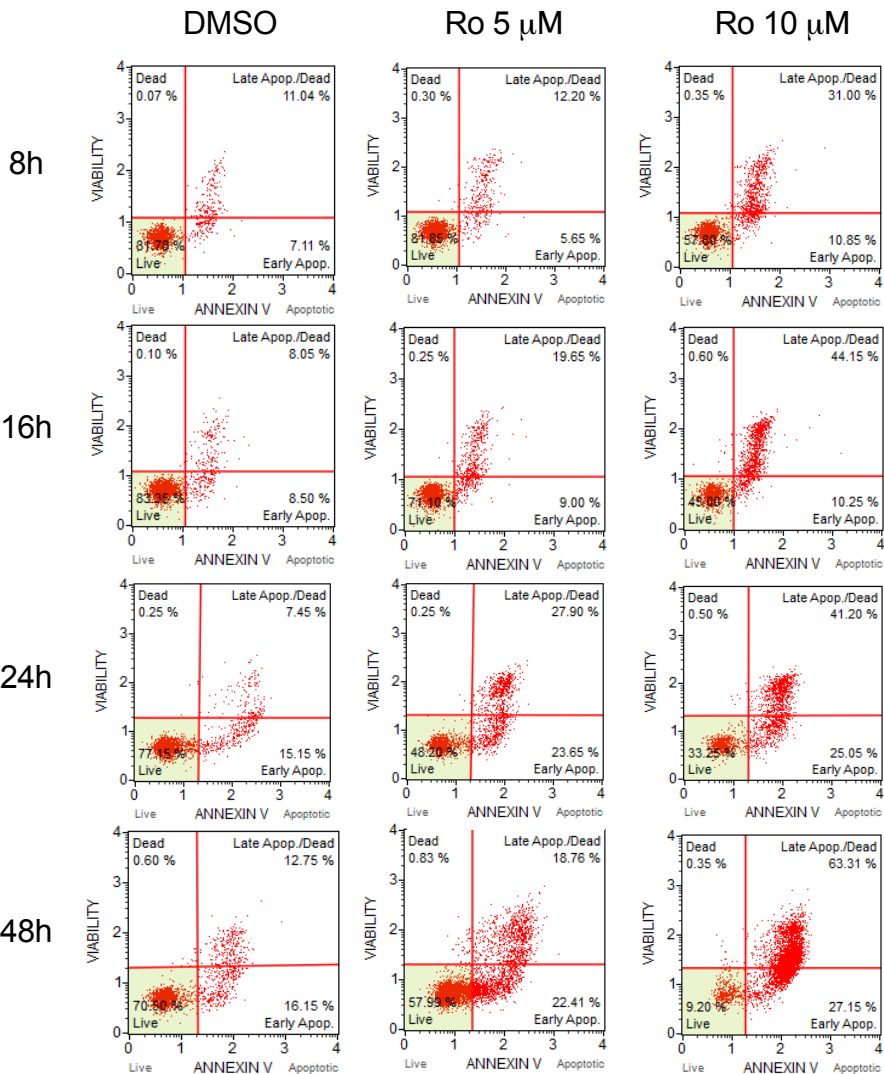
NMR (^{13}C) spectrum



Extended Data Figure 4

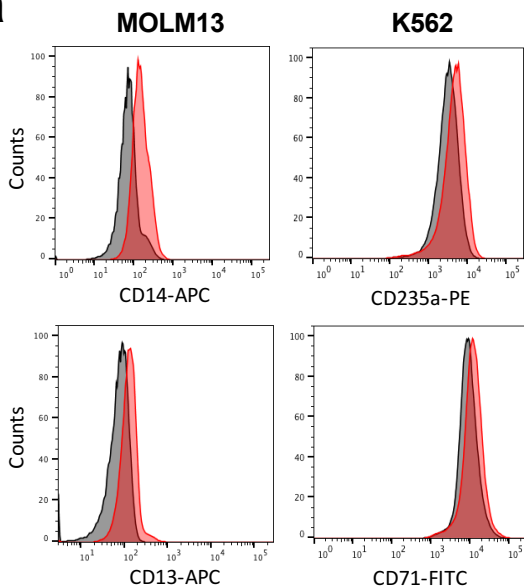


Extended Data Figure 5

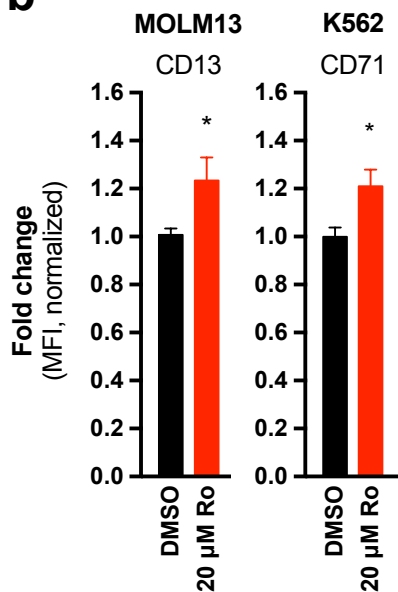


Extended Data Figure 6

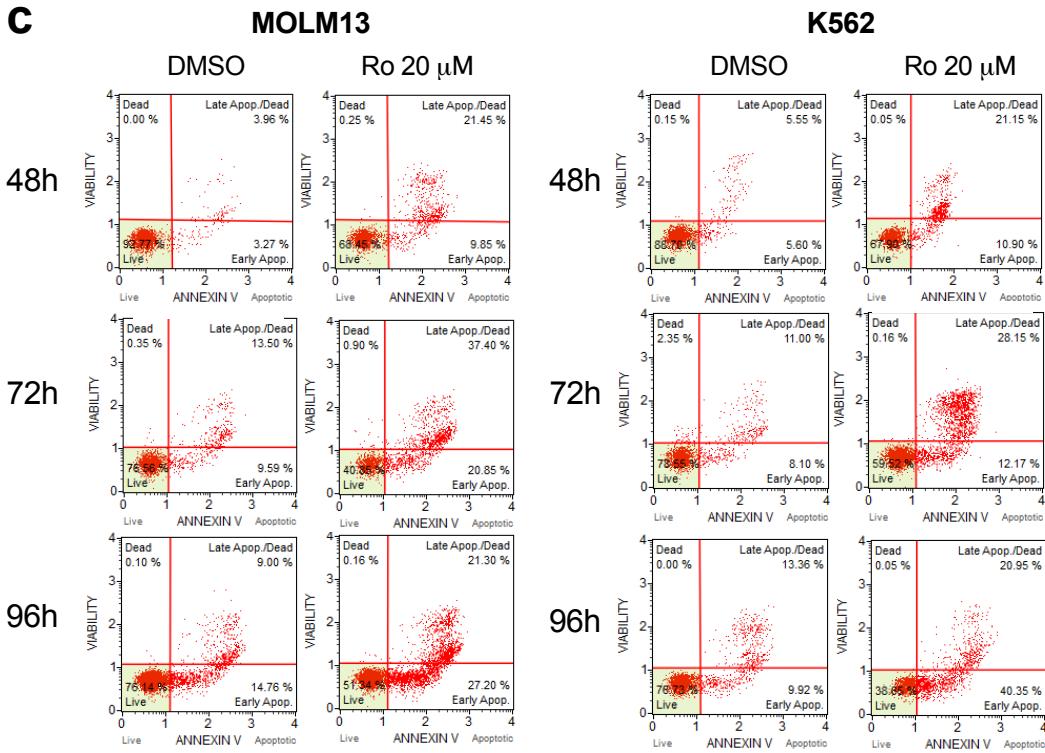
a



b

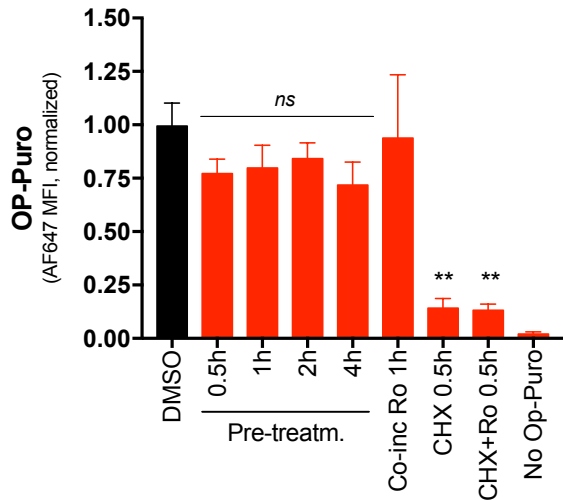


c

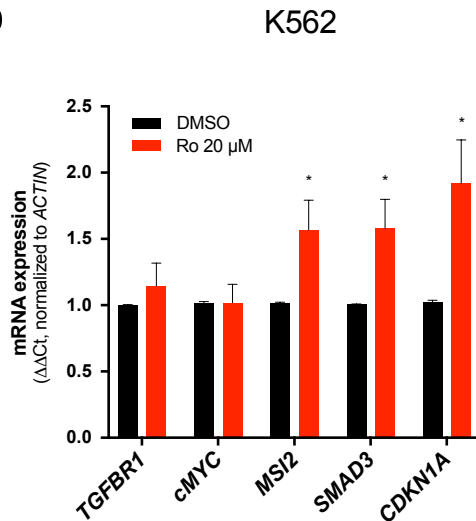


Extended Data Figure 7

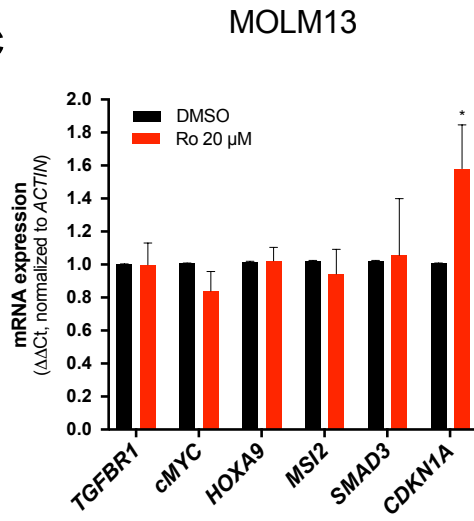
a



b

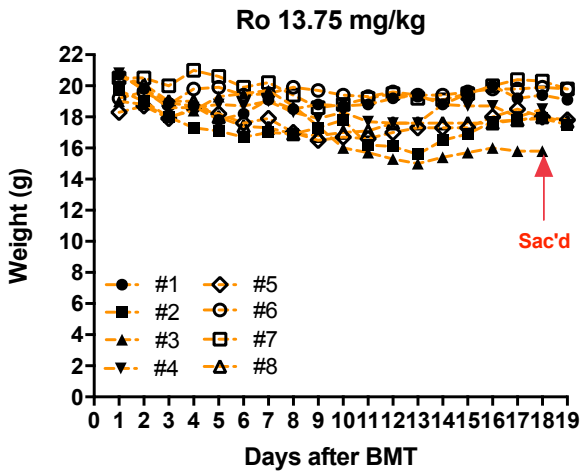
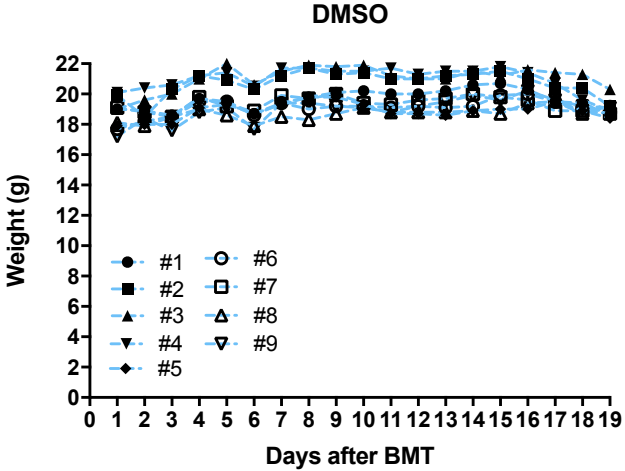


c

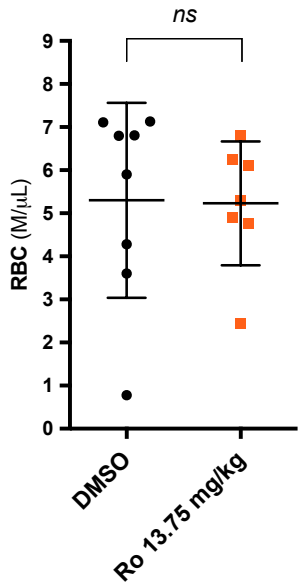


Extended Data Figure 8

a



b



c

