

(10) International Publication Number
WO 2017/070611 A1

Fig. 1

WO 2017/070611 A1

SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, — *with sequence listing part of description (Rule 5.2(a))*
GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

CYSTEINE REACTIVE PROBES AND USES THEREOF

CROSS-REFERENCE

[0001] This application claims the benefit of U.S. Provisional Application No. 62/345,710, filed on June 3, 2016, and U.S. Provisional Application No. 62/244,881, filed on October 22, 2015, each of which are incorporated herein by reference in their entireties.

STATEMENT AS TO FEDERALLY SPONSORED RESEARCH

[0002] The invention disclosed herein was made, at least in part, with U.S. government support under Grant Nos. CA087660, GM090294, GM108208, and GM069832 by the National Institutes of Health. Accordingly, the U.S. Government has certain rights in this invention.

SEQUENCE LISTING

[0003] The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on October 19, 2016, is named 48054-702_601_SL.txt and is 372,838 bytes in size.

BACKGROUND OF THE INVENTION

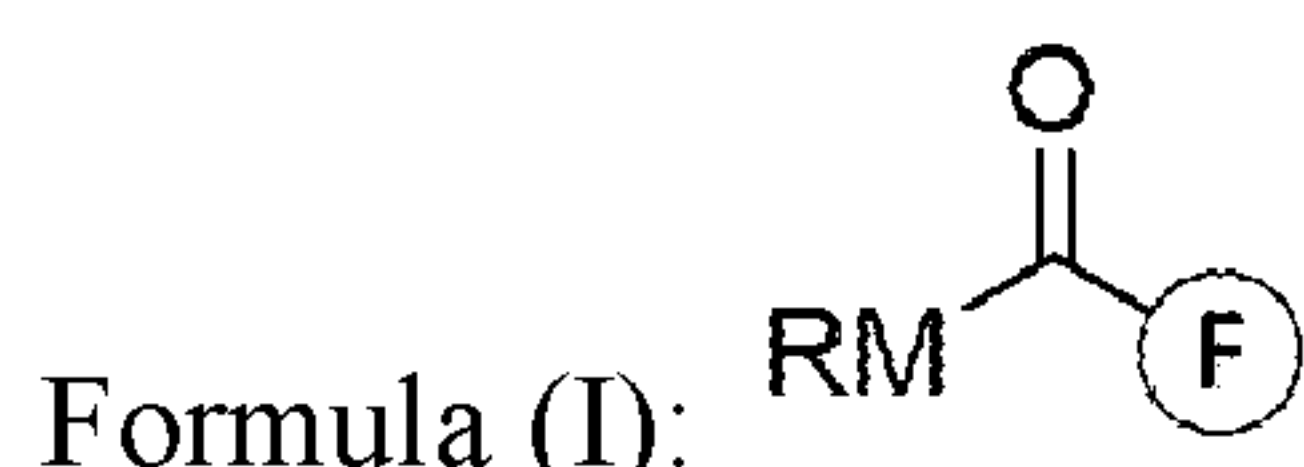
[0004] Protein function assignment has been benefited from genetic methods, such as target gene disruption, RNA interference, and genome editing technologies, which selectively disrupt the expression of proteins in native biological systems. Chemical probes offer a complementary way to perturb proteins that have the advantages of producing graded (dose-dependent) gain- (agonism) or loss- (antagonism) of-function effects that are introduced acutely and reversibly in cells and organisms. Small molecules present an alternative method to selectively modulate proteins and to serve as leads for the development of novel therapeutics.

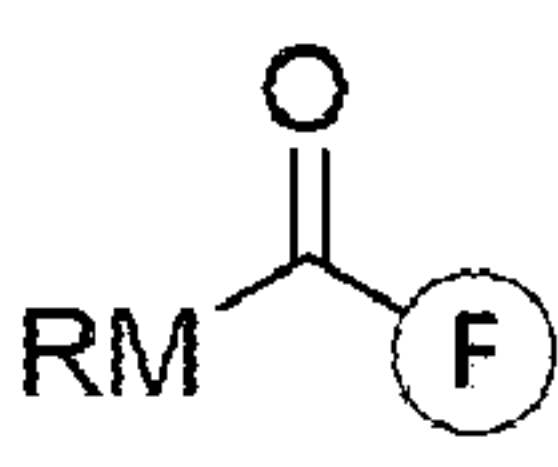
SUMMARY OF THE INVENTION

[0005] Disclosed herein, in certain embodiments, is a method of identifying a cysteine containing protein as a binding target for a small molecule fragment, comprising: (a) obtaining a set of cysteine-reactive probe-protein complexes from a sample treated with a cysteine-reactive probe wherein the cysteine-reactive probe comprises a reactive moiety capable of forming a covalent bond with a cysteine residue located on the cysteine containing protein; (b) analyzing the set of cysteine-reactive probe-protein complexes by a proteomic analysis means; (c) based on step b), identifying a cysteine containing protein as the binding target for the small molecule fragment. In some embodiments, the method further comprises assigning a value to each of the cysteine containing protein from the set of cysteine-reactive probe-protein complexes for identifying a cysteine containing protein as the binding target for the small molecule fragment,

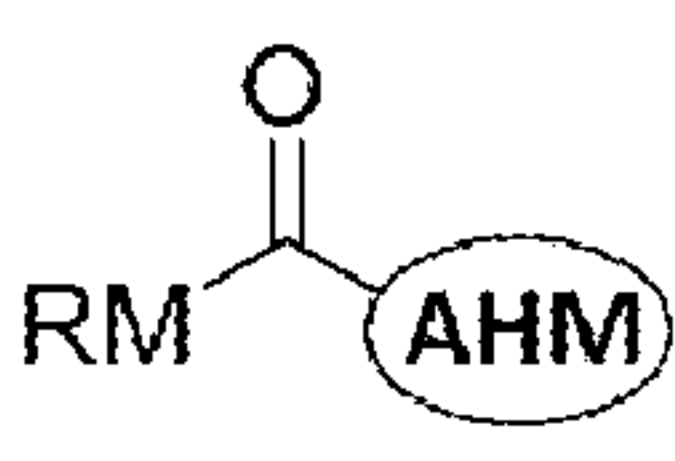
wherein the value is determined based on the proteomic analysis means of step b). In some embodiments, the sample comprises a first cell solution and a second cell solution. In some embodiments, the method further comprises contacting the first cell solution with a small molecule fragment for an extended period of time prior to incubating the first cell solution with a first cysteine-reactive probe to generate a first group of cysteine-reactive probe-protein complexes. In some embodiments, the extended period of time is about 5, 10, 15, 20, 30, 60, 90, 120 minutes or longer. In some embodiments, the method further comprises contacting the second cell solution with a second cysteine-reactive probe to generate a second group of cysteine-reactive probe-protein complexes. In some embodiments, the first cysteine-reactive probe and the second cysteine-reactive probe are the same. In some embodiments, the first group and the second group of cysteine-reactive probe-protein complexes comprise the set of cysteine-reactive probe-protein complexes. In some embodiments, cells from the second cell solution are grown in a media (e.g., an isotopically enriched media). In some embodiments, cells from the first cell solution are grown in a media (e.g., an isotopically enriched media). In some embodiments, cells from both the first cell solution and the second cell solution are grown in two different isotopically enriched media so that cells from the first cell solution is distinguishable from cells obtained from the second cell solution. In other embodiments, cells from only one of the cell solutions (e.g., either the first cell solution or the second cell solution) are grown in an isotopically enriched media. In some embodiments, the method further comprises contacting the first cell solution with a first set of small molecule fragments and a complementing set of cysteine-reactive probes wherein each small molecule fragment competes with its complementing cysteine-reactive probe for binding with a cysteine residue, and wherein each small molecule fragment and each complementing cysteine-reactive probe are different within each respective set. In some embodiments, the method further comprises contacting the second cell solution with a second set of cysteine-reactive probes wherein the second set of cysteine-reactive probes is the same as the complementing set of cysteine-reactive probes, and wherein each cysteine-reactive probe is different within the set. In some embodiments, the first set of cysteine-reactive probes generates a third group of cysteine-reactive probe-protein complexes and the second set of cysteine-reactive probes generates a fourth group of cysteine-reactive probe-protein complexes. In some embodiments, the cysteine containing protein comprises a biologically active cysteine residue. In some embodiments, the biologically active cysteine site is a cysteine residue that is located about 10Å or less to an active-site ligand or residue. In some embodiments, the cysteine residue that is located about 10Å or less to the active-site ligand or residue is an active site cysteine. In some embodiments, the biologically active cysteine site is an active site cysteine. In some embodiments, the biologically active

cysteine site is a cysteine residue that is located greater than 10Å from an active-site ligand or residue. In some embodiments, the cysteine residue that is located greater than 10Å from the active-site ligand or residue is a non-active site cysteine. In some embodiments, the biologically active cysteine site is a non-active site cysteine. In some embodiments, the small molecule fragment that covalently interacts with the biologically active cysteine impairs and/or inhibits activity of the cysteine containing protein. In some embodiments, the cysteine containing protein exists in an active form. In some embodiments, the small molecule fragment and/or the cysteine-reactive probe interact with the active form of the cysteine containing protein. In some embodiments, the cysteine containing protein exists in a pro-active form. In some embodiments, the small molecule fragment and/or the cysteine-reactive probe interact with the pro-active form of the cysteine containing protein. In some embodiments, the structural environment of the biologically active cysteine residue modulates the reactivity of the cysteine residue. In some embodiments, the structural environment is a hydrophobic environment or a hydrophilic environment. In some embodiments, the structural environment is a charged environment. In some embodiments, the structural environment is a nucleophilic environment. In some embodiments, the cysteine containing protein is an enzyme, a transporter, a receptor, a channel protein, an adaptor protein, a chaperone, a signaling protein, a plasma protein, transcription related protein, translation related protein, mitochondrial protein, or cytoskeleton related protein. In some embodiments, the cysteine containing protein is an enzyme, a transporter, a receptor, a channel protein, an adaptor protein, a chaperone, a signaling protein, transcription related protein, or translation related protein. In some embodiments, the enzyme comprises kinases, proteases, or deubiquitinating enzymes. In some embodiments, the protease is a cysteine protease. In some embodiments, the cysteine protease comprises caspases. In some embodiments, the signaling protein comprises vascular endothelial growth factor. In some embodiments, the signaling protein comprises a redox signaling protein. In some embodiments, the cysteine containing protein is a protein illustrated in Table 1. In some embodiments, the cysteine containing protein is a protein illustrated in Table 2. In some embodiments, the cysteine containing protein is a protein illustrated in Table 3. In some embodiments, the cysteine containing protein comprises a cysteine residue denoted in Table 3. In some embodiments, the cysteine containing protein is a protein illustrated in Table 8. In some embodiments, the cysteine containing protein is a protein illustrated in Table 9. In some embodiments, the cysteine containing protein is a protein illustrated in Table 10A, Table 10B, Table 10C, Table 10D or Table 10E. In some embodiments, the small molecule fragment is a small molecule fragment of



Formula (I):  Formula (I), wherein: RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and F is a small molecule fragment moiety. In some embodiments, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some embodiments, F is obtained from a compound library. In some embodiments, the compound library comprises ChemBridge fragment library, Pyramid Platform Fragment-Based Drug Discovery, Maybridge fragment library, FRGx from AnalytiCon, TCI-Frag from AnCoreX, Bio Building Blocks from ASINEX, BioFocus 3D from Charles River, Fragments of Life (FOL) from Emerald Bio, Enamine Fragment Library, IOTA Diverse 1500, BIONET fragments library, Life Chemicals Fragments Collection, OTAVA fragment library, Prestwick fragment library, Selcia fragment library, TimTec fragment-based library, Allium from Vitas-M Laboratory, or Zenobia fragment library. In some embodiments, F is a small molecule fragment moiety illustrated in Fig. 3. In some embodiments, F further comprises a linker moiety that connects F to the carbonyl moiety. In some embodiments, the small molecule fragment is a small molecule fragment illustrated in Fig. 3. In some embodiments, the small molecule fragment is a specific inhibitor or a pan inhibitor. In some embodiments, the cysteine-reactive probe is a cysteine-



reactive probe of Formula (II):  Formula (II), wherein: RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond to the thiol group of a cysteine residue; and AHM is an affinity handle moiety. In some embodiments, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some embodiments, the affinity handle moiety comprises an affinity handle and a binding moiety that facilitates covalent interaction of the cysteine-reactive probe to a cysteine residue of a cysteine-containing protein. In some embodiments, the binding moiety is a small molecule fragment obtained from a compound library. In some embodiments, the compound library comprises ChemBridge fragment library, Pyramid Platform Fragment-Based Drug Discovery, Maybridge fragment library, FRGx from AnalytiCon, TCI-Frag from AnCoreX, Bio Building Blocks from ASINEX, BioFocus 3D from Charles River, Fragments of Life (FOL) from Emerald Bio, Enamine Fragment Library, IOTA Diverse 1500, BIONET fragments library, Life Chemicals Fragments Collection, OTAVA fragment library, Prestwick fragment library, Selcia fragment library, TimTec fragment-based library, Allium from Vitas-M Laboratory, or Zenobia fragment library. In some embodiments, the affinity handle is a bioorthogonal affinity handle. In some embodiments, the affinity handle comprises a carbodiimide, N-

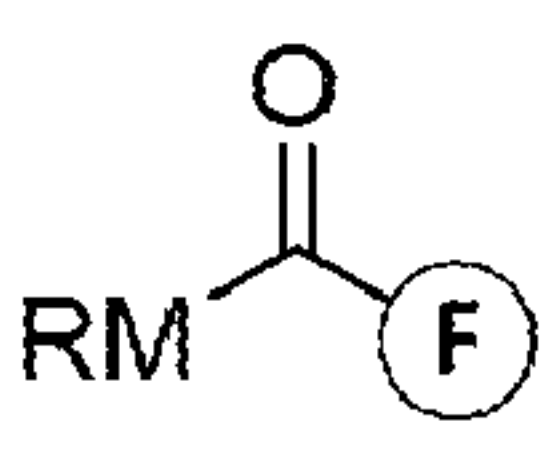
hydroxysuccinimide (NHS) ester, imidoester, pentafluorophenyl ester, hydroxymethyl phosphine, maleimide, haloacetyl, pyridyl disulfide, thiosulfonate, vinylsulfone, hydrazide, alkoxyamine, alkyne, azide, or isocyanate group. In some embodiments, the affinity handle comprises an alkyne or an azide group. In some embodiments, the affinity handle is further conjugated to an affinity ligand. In some embodiments, the affinity ligand comprises a chromophore, a labeling group, or a combination thereof. In some embodiments, the chromophore comprises fluorochrome, non-fluorochrome chromophore, quencher, an absorption chromophore, fluorophore, organic dye, inorganic dye, metal chelate, or a fluorescent enzyme substrate. In some embodiments, the fluorophore comprises rhodamine, rhodol, fluorescein, thiofluorescein, aminofluorescein, carboxyfluorescein, chlorofluorescein, methylfluorescein, sulfofluorescein, aminorhodol, carboxyrhodol, chlororhodol, methylrhodol, sulforhodol, aminorhodamine, carboxyrhodamine, chlororhodamine, methylrhodamine, sulforhodamine, thiorhodamine, cyanine, indocarbocyanine, oxacarbocyanine, thiocarbocyanine, merocyanine, cyanine 2, cyanine 3, cyanine 3.5, cyanine 5, cyanine 5.5, cyanine 7, oxadiazole derivatives, pyridyloxazole, nitrobenzoxadiazole, benzoxadiazole, pyren derivatives, cascade blue, oxazine derivatives, Nile red, Nile blue, cresyl violet, oxazine 170, acridine derivatives, proflavin, acridine orange, acridine yellow, arylmethine derivatives, auramine, crystal violet, malachite green, tetrapyrrole derivatives, porphin, phtalocyanine, bilirubin 1-dimethylaminonaphthyl-5-sulfonate, 1-anilino-8-naphthalene sulfonate, 2-p-touidiny-6-naphthalene sulfonate, 3-phenyl-7-isocyanatocoumarin, N-(p-(2-benzoxazolyl)phenyl)maleimide, stilbenes, pyrenes, 6-FAM (Fluorescein), 6-FAM (NHS Ester), 5(6)-FAM, 5-FAM, Fluorescein dT, 5-TAMRA-cadavarine, 2-aminoacridone, HEX, JOE (NHS Ester), MAX, TET, ROX, TAMRA, TARMA™ (NHS Ester), TEX 615, ATTO™ 488, ATTO™ 532, ATTO™ 550, ATTO™ 565, ATTO™ Rho101, ATTO™ 590, ATTO™ 633, ATTO™ 647N, TYE™ 563, TYE™ 665, or TYE™ 705. In some embodiments, the labeling group is biotin moiety, streptavidin moiety, bead, resin, a solid support, or a combination thereof. In some embodiments, the affinity handle moiety further comprises a chromophore. In some embodiments, the cysteine-reactive probe is a cysteine-reactive probe illustrated in Fig. 3. In some embodiments, the second cell solution further comprises a control. In some embodiments, the control is dimethyl sulfoxide (DMSO). In some embodiments, the proteomic analysis means comprises a mass spectroscopy method. In some embodiments, the mass spectroscopy method is a liquid-chromatography-mass spectrometry (LC-MS) method. In some embodiments, the method further comprises analyzing the results from the mass spectroscopy method by an algorithm for protein identification. In some embodiments, the algorithm combines the results from the mass spectroscopy method with a protein sequence database for protein identification. In some embodiments, the algorithm

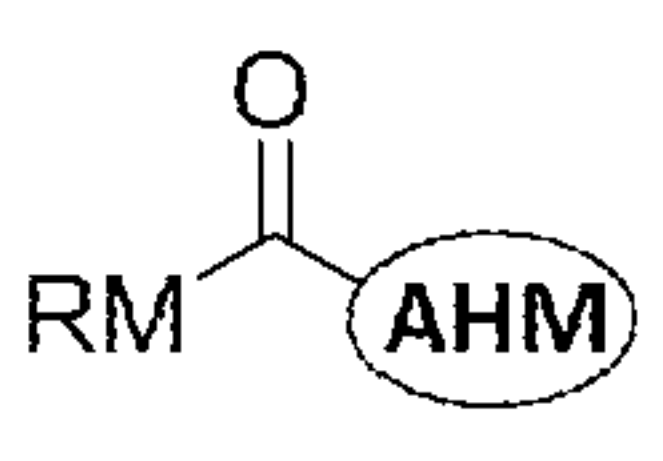
comprises ProLuCID algorithm, Probity, Scaffold, SEQUEST, or Mascot. In some embodiments, the mass spectroscopy method is a MALDI-TOF based method. In some embodiments, the value assigned to each of the cysteine containing protein is obtained from the mass spectroscopy analysis. In some embodiments, the value assigned to each of the cysteine containing protein is the area-under-the curve from a plot of signal intensity as a function of mass-to-charge ratio. In some embodiments, the identifying in step c) further comprises (i) locating a first value assigned to a cysteine containing protein from the first group of cysteine-reactive probe-protein complex and a second value of the same cysteine containing protein from the second group of cysteine-reactive probe-protein complex; and (ii) calculating a ratio between the two values assigned to the same cysteine containing protein. In some embodiments, the ratio of greater than 2 indicates that the cysteine containing protein is a candidate for interacting with the small molecule fragment. In some embodiments, the ratio of greater than 3 indicates that the cysteine containing protein is a candidate for interacting with the small molecule fragment. In some embodiments, the identifying in step c) further comprises calculating a percentage of inhibition of the cysteine-reactive probe to the cysteine containing protein. In some embodiments, the percentage of inhibition of greater than 50%, 60%, 70%, 80%, 90%, or at 100% indicates that the cysteine containing protein is a candidate for interacting with the small molecule fragment. In some embodiments, the cell is obtained from a tumor cell line. In some embodiments, the cell is obtained from a MDA-MB-231, Ramos, or Jurkat cell line. In some embodiments, the cell is obtained from a tumor sample. In some embodiments, the sample is a tissue sample. In some embodiments, the method is an *in situ* method. In some embodiments, the cysteine-reactive probe is not 4-hydroxynonenal or 15-deoxy- $\Delta^{12,14}$ -prostaglandin J2.

[0006] Disclosed herein, in certain embodiments, is a method of screening a small molecule fragment for interaction with a cysteine containing protein, comprising: (a) harvesting a set of cysteine-reactive probe-protein complexes from a sample treated with a cysteine-reactive probe wherein the cysteine-reactive probe comprises a reactive moiety capable of forming a covalent bond with a cysteine residue located on the cysteine containing protein; (b) analyzing the set of cysteine-reactive probe-protein complexes by a proteomic analysis means; and (c) based on step b), identifying the small molecule fragment as interacting with the cysteine containing protein. In some embodiments, the method further comprises assigning a value to each of the cysteine containing protein from the set of cysteine-reactive probe-protein complexes prior to identifying the small molecule fragment as interacting with the cysteine containing protein, wherein the value is determined based on the proteomic analysis means of step b). In some embodiments, the sample comprises a first cell solution and a second cell solution. In some embodiments, the method further comprises contacting the first cell solution with a small molecule fragment for an

extended period of time prior to incubating the first cell solution with a first cysteine-reactive probe to generate a first group of cysteine-reactive probe-protein complexes. In some embodiments, the extended period of time is about 5, 10, 15, 20, 30, 60, 90, 120 minutes or longer. In some embodiments, the method further comprises contacting the second cell solution with a second cysteine-reactive probe to generate a second group of cysteine-reactive probe-protein complexes. In some embodiments, the first cysteine-reactive probe and the second cysteine-reactive probe are the same. In some embodiments, the first group and the second group of cysteine-reactive probe-protein complexes comprise the set of cysteine-reactive probe-protein complexes. In some embodiments, cells from the second cell solution are grown in a media (e.g., an isotopically enriched media). In some embodiments, cells from the first cell solution are grown in a media (e.g., an isotopically enriched media). In some embodiments, cells from both the first cell solution and the second cell solution are grown in two different isotopically enriched media so that cells from the first cell solution is distinguishable from cells obtained from the second cell solution. In other embodiments, cells from only one of the cell solutions (e.g., either the first cell solution or the second cell solution) are grown in an isotopically enriched media. In some embodiments, the method further comprises contacting the first cell solution with a first set of small molecule fragments and a complementing set of cysteine-reactive probes wherein each small molecule fragment competes with its complementing cysteine-reactive probe for binding with a cysteine residue, and wherein each small molecule fragment and each complementing cysteine-reactive probe are different within each respective set. In some embodiments, the method further comprises contacting the second cell solution with a second set of cysteine-reactive probes wherein the second set of cysteine-reactive probes is the same as the complementing set of cysteine-reactive probes, and wherein each cysteine-reactive probe is different within the set. In some embodiments, the first set of cysteine-reactive probes generates a third group of cysteine-reactive probe-protein complexes and the second set of cysteine-reactive probes generates a fourth group of cysteine-reactive probe-protein complexes. In some embodiments, the cysteine containing protein comprises a biologically active cysteine residue. In some embodiments, the biologically active cysteine site is a cysteine residue that is located about 10Å or less to an active-site ligand or residue. In some embodiments, the cysteine residue that is located about 10Å or less to the active-site ligand or residue is an active site cysteine. In some embodiments, the biologically active cysteine site is an active site cysteine. In some embodiments, the biologically active cysteine site is a cysteine residue that is located greater than 10Å from an active-site ligand or residue. In some embodiments, the cysteine residue that is located greater than 10Å from the active-site ligand or residue is a non-active site cysteine. In some embodiments, the biologically active cysteine site

is a non-active site cysteine. In some embodiments, the small molecule fragment that covalently interacts with the biologically active cysteine impairs and/or inhibits activity of the cysteine containing protein. In some embodiments, the cysteine containing protein exists in an active form. In some embodiments, the small molecule fragment and/or the cysteine-reactive probe interact with the active form of the cysteine containing protein. In some embodiments, the cysteine containing protein exists in a pro-active form. In some embodiments, the small molecule fragment and/or the cysteine-reactive probe interact with the pro-active form of the cysteine containing protein. In some embodiments, the structural environment of the biologically active cysteine residue modulates the reactivity of the cysteine residue. In some embodiments, the structural environment is a hydrophobic environment or a hydrophilic environment. In some embodiments, the structural environment is a charged environment. In some embodiments, the structural environment is a nucleophilic environment. In some embodiments, the cysteine containing protein is selected from an enzyme, a transporter, a receptor, a channel protein, an adaptor protein, a chaperone, a signaling protein, a plasma protein, transcription related protein, translation related protein, mitochondrial protein, or cytoskeleton related protein. In some embodiments, the cysteine containing protein is selected from an enzyme, a transporter, a receptor, a channel protein, an adaptor protein, a chaperone, a signaling protein, transcription related protein, or translation related protein. In some embodiments, the enzyme comprises kinases, proteases, or deubiquitinating enzymes. In some embodiments, the protease is a cysteine protease. In some embodiments, the cysteine protease comprises caspases. In some embodiments, the signaling protein comprises vascular endothelial growth factor. In some embodiments, the signaling protein comprises a redox signaling protein. In some embodiments, the cysteine containing protein is selected from Table 1. In some embodiments, the cysteine containing protein is a protein illustrated in Table 2. In some embodiments, the cysteine containing protein is a protein illustrated in Table 3. In some embodiments, the cysteine containing protein comprises a cysteine residue denoted in Table 3. In some embodiments, the cysteine containing protein is a protein illustrated in Table 8. In some embodiments, the cysteine containing protein is a protein illustrated in Table 9. In some embodiments, the cysteine containing protein is a protein illustrated in Table 10A, Table 10B, Table 10C, Table 10D or Table 10E. In some embodiments, the cysteine containing protein is TIGAR, IMPDH2, IDH1, IDH2, BTK, ZAK, TGM2, Map2k7, XPO1, Casp5, Casp8, ERCC3, Park 7 (Toxoplasma DJ-1), GSTO1, ALDH2, CTSZ, STAT1, STAT3, SMAD2, RBPJ, FOXK1, IRF4, IRF8, GTF3C1, or TCERG1. In some embodiments, the small molecule fragment is a small molecule fragment of

Formula (I):  Formula (I), wherein: RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and F is a small molecule fragment moiety. In some embodiments, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some embodiments, F is obtained from a compound library. In some embodiments, the compound library comprises ChemBridge fragment library, Pyramid Platform Fragment-Based Drug Discovery, Maybridge fragment library, FRGx from AnalytiCon, TCI-Frag from AnCoreX, Bio Building Blocks from ASINEX, BioFocus 3D from Charles River, Fragments of Life (FOL) from Emerald Bio, Enamine Fragment Library, IOTA Diverse 1500, BIONET fragments library, Life Chemicals Fragments Collection, OTAVA fragment library, Prestwick fragment library, Selcia fragment library, TimTec fragment-based library, Allium from Vitas-M Laboratory, or Zenobia fragment library. In some embodiments, F is a small molecule fragment moiety illustrated in Fig. 3. In some embodiments, F further comprises a linker moiety that connects F to the carbonyl moiety. In some embodiments, the small molecule fragment is a small molecule fragment illustrated in Fig. 3. In some embodiments, the small molecule fragment is a specific inhibitor or a pan inhibitor. In some embodiments, the cysteine-reactive probe is a cysteine-

reactive probe of Formula (II):  Formula (II), wherein: RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond to the thiol group of a cysteine residue; and AHM is an affinity handle moiety. In some embodiments, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some embodiments, the affinity handle moiety comprises an affinity handle and a binding moiety that facilitates covalent interaction of the cysteine-reactive probe to a cysteine residue of a cysteine-containing protein. In some embodiments, the binding moiety is a small molecule fragment obtained from a compound library. In some embodiments, the compound library comprises ChemBridge fragment library, Pyramid Platform Fragment-Based Drug Discovery, Maybridge fragment library, FRGx from AnalytiCon, TCI-Frag from AnCoreX, Bio Building Blocks from ASINEX, BioFocus 3D from Charles River, Fragments of Life (FOL) from Emerald Bio, Enamine Fragment Library, IOTA Diverse 1500, BIONET fragments library, Life Chemicals Fragments Collection, OTAVA fragment library, Prestwick fragment library, Selcia fragment library, TimTec fragment-based library, Allium from Vitas-M Laboratory, or Zenobia fragment library. In some embodiments, the affinity handle is a bioorthogonal affinity handle. In some embodiments, the affinity handle comprises a carbodiimide, N-

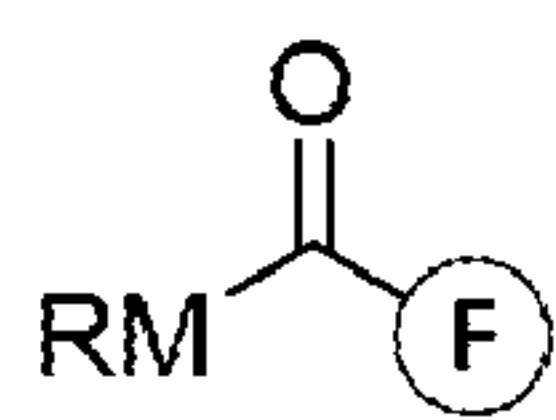
hydroxysuccinimide (NHS) ester, imidoester, pentafluorophenyl ester, hydroxymethyl phosphine, maleimide, haloacetyl, pyridyl disulfide, thiosulfonate, vinylsulfone, hydrazide, alkoxyamine, alkyne, azide, or isocyanate group. In some embodiments, the affinity handle comprises an alkyne or an azide group. In some embodiments, the affinity handle is further conjugated to an affinity ligand. In some embodiments, the affinity ligand comprises a chromophore, a labeling group, or a combination thereof. In some embodiments, the chromophore comprises fluorochrome, non-fluorochrome chromophore, quencher, an absorption chromophore, fluorophore, organic dye, inorganic dye, metal chelate, or a fluorescent enzyme substrate. In some embodiments, the fluorophore comprises rhodamine, rhodol, fluorescein, thiofluorescein, aminofluorescein, carboxyfluorescein, chlorofluorescein, methylfluorescein, sulfofluorescein, aminorhodol, carboxyrhodol, chlororhodol, methylrhodol, sulforhodol; aminorhodamine, carboxyrhodamine, chlororhodamine, methylrhodamine, sulforhodamine, thiorhodamine, cyanine, indocarbocyanine, oxacarbocyanine, thiocarbocyanine, merocyanine, cyanine 2, cyanine 3, cyanine 3.5, cyanine 5, cyanine 5.5, cyanine 7, oxadiazole derivatives, pyridyloxazole, nitrobenzoxadiazole, benzoxadiazole, pyren derivatives, cascade blue, oxazine derivatives, Nile red, Nile blue, cresyl violet, oxazine 170, acridine derivatives, proflavin, acridine orange, acridine yellow, arylmethine derivatives, auramine, crystal violet, malachite green, tetrapyrrole derivatives, porphin, phtalocyanine, bilirubin 1-dimethylaminonaphthyl-5-sulfonate, 1-anilino-8-naphthalene sulfonate, 2-p-touidiny-6-naphthalene sulfonate, 3-phenyl-7-isocyanatocoumarin, N-(p-(2-benzoxazolyl)phenyl)maleimide, stilbenes, pyrenes, 6-FAM (Fluorescein), 6-FAM (NHS Ester), 5(6)-FAM, 5-FAM, Fluorescein dT, 5-TAMRA-cadavarine, 2-aminoacridone, HEX, JOE (NHS Ester), MAX, TET, ROX, TAMRA, TARMA™ (NHS Ester), TEX 615, ATTO™ 488, ATTO™ 532, ATTO™ 550, ATTO™ 565, ATTO™ Rho101, ATTO™ 590, ATTO™ 633, ATTO™ 647N, TYE™ 563, TYE™ 665, or TYE™ 705. In some embodiments, the labeling group is biotin moiety, streptavidin moiety, bead, resin, a solid support, or a combination thereof. In some embodiments, the affinity handle moiety further comprises a chromophore. In some embodiments, the cysteine-reactive probe is a cysteine-reactive probe illustrated in Fig. 3. In some embodiments, the second cell solution further comprises a control. In some embodiments, the control is dimethyl sulfoxide (DMSO). In some embodiments, the proteomic analysis means comprises a mass spectroscopy method. In some embodiments, the mass spectroscopy method is a MALDI-TOF based method. In some embodiments, the mass spectroscopy method is a liquid-chromatography-mass spectrometry (LC-MS) method. In some embodiments, the method further comprises analyzing the results from the mass spectroscopy method by an algorithm for protein identification. In some embodiments, the algorithm combines the results from the mass spectroscopy method with a

protein sequence database for protein identification. In some embodiments, the algorithm comprises ProLuCID algorithm, Probit, Scaffold, SEQUEST, or Mascot. In some embodiments, the value assigned to each of the cysteine containing protein is obtained from the mass spectroscopy analysis. In some embodiments, the value assigned to each of the cysteine containing protein is the area-under-the curve from a plot of signal intensity as a function of mass-to-charge ratio. In some embodiments, the identifying in step c) further comprises (i) locating a first value assigned to a cysteine containing protein from the first group of cysteine-reactive probe-protein complex and a second value of the same cysteine containing protein from the second group of cysteine-reactive probe-protein complex; and (ii) calculating a ratio between the two values assigned to the same cysteine containing protein. In some embodiments, the ratio of greater than 2 indicates that the cysteine containing protein is a candidate for interacting with the small molecule fragment. In some embodiments, the ratio of greater than 3 indicates that the cysteine containing protein is a candidate for interacting with the small molecule fragment. In some embodiments, the identifying in step c) further comprises calculating a percentage of inhibition of the cysteine-reactive probe to the cysteine containing protein. In some embodiments, the percentage of inhibition of greater than 50%, 60%, 70%, 80%, 90%, or at 100% indicates that the cysteine containing protein is a candidate for interacting with the small molecule fragment. In some embodiments, the cell is obtained from a tumor cell line. In some embodiments, the cell is obtained from a MDA-MB-231, Ramos, or Jurkat cell line. In some embodiments, the cell is obtained from a tumor sample. In some embodiments, the sample is a tissue sample. In some embodiments, the method is an *in situ* method.

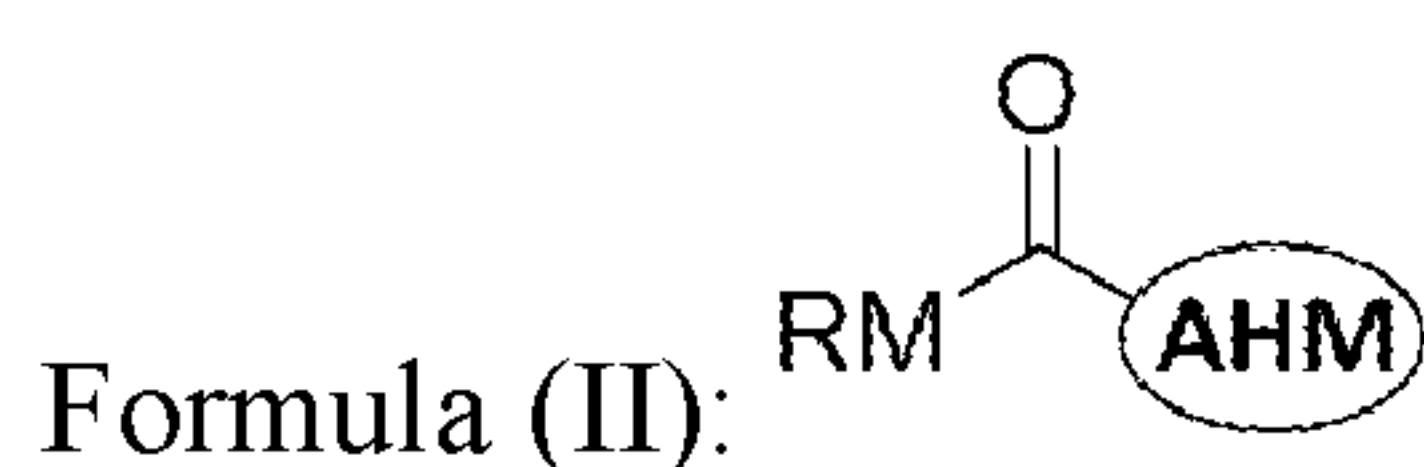
[0007] Disclosed herein, in certain embodiments, is a method of mapping a biologically active cysteine site on a protein, comprising (a) harvesting a set of cysteine-reactive probe-protein complexes from a sample treated with a cysteine-reactive probe wherein the cysteine-reactive probe comprises a reactive moiety capable of forming a covalent bond with a cysteine residue located on the cysteine containing protein; (b) analyzing the set of cysteine-reactive probe-protein complexes by a proteomic analysis means; and (c) based on step b), mapping the biologically active cysteine site on the protein. In some embodiments, the sample comprises a first cell solution and a second cell solution. In some embodiments, the method further comprises contacting the first cell solution with a small molecule fragment for an extended period of time prior to incubating the first cell solution with a first cysteine-reactive probe to generate a first group of cysteine-reactive probe-protein complexes. In some embodiments, the extended period of time is about 5, 10, 15, 20, 30, 60, 90, 120 minutes or longer. In some embodiments, the method further comprises contacting the second cell solution with a second cysteine-reactive probe to generate a second group of cysteine-reactive probe-protein

complexes. In some embodiments, the first cysteine-reactive probe and the second cysteine-reactive probe are the same. In some embodiments, the biologically active cysteine site is a cysteine residue that is located about 10Å or less to an active-site ligand or residue. In some embodiments, the cysteine residue that is located about 10Å or less to the active-site ligand or residue is an active site cysteine. In some embodiments, the biologically active cysteine site is an active site cysteine. In some embodiments, the biologically active cysteine site is a cysteine residue that is located greater than 10Å from an active-site ligand or residue. In some embodiments, the cysteine residue that is located greater than 10Å from the active-site ligand or residue is a non-active site cysteine. In some embodiments, the biologically active cysteine site is a non-active site cysteine. In some embodiments, the small molecule fragment that covalently interacts with the biologically active cysteine impairs and/or inhibits activity of the cysteine containing protein. In some embodiments, the cysteine containing protein exists in an active form. In some embodiments, the small molecule fragment and/or the cysteine-reactive probe interact with the active form of the cysteine containing protein. In some embodiments, the cysteine containing protein exists in a pro-active form. In some embodiments, the small molecule fragment and/or the cysteine-reactive probe interact with the pro-active form of the cysteine containing protein. In some embodiments, the structural environment of the biologically active cysteine residue modulates the reactivity of the cysteine residue. In some embodiments, the structural environment is a hydrophobic environment or a hydrophilic environment. In some embodiments, the structural environment is a charged environment. In some embodiments, the structural environment is a nucleophilic environment. In some embodiments, the protein is an enzyme, a transporter, a receptor, a channel protein, an adaptor protein, a chaperone, a signaling protein, a plasma protein, transcription related protein, translation related protein, mitochondrial protein, or cytoskeleton related protein. In some embodiments, the protein is an enzyme, a transporter, a receptor, a channel protein, an adaptor protein, a chaperone, a signaling protein, transcription related protein, or translation related protein. In some embodiments, the enzyme comprises kinases, proteases, or deubiquitinating enzymes. In some embodiments, the protease is a cysteine protease. In some embodiments, the cysteine protease comprises caspases. In some embodiments, the signaling protein comprises vascular endothelial growth factor. In some embodiments, the signaling protein comprises a redox signaling protein. In some embodiments, the protein is a protein illustrated in Table 1. In some embodiments, the cysteine containing protein is a protein illustrated in Table 2. In some embodiments, the cysteine containing protein is a protein illustrated in Table 3. In some embodiments, the cysteine containing protein comprises a cysteine residue denoted in Table 3. In some embodiments, the cysteine containing protein is a protein illustrated in Table 8. In some embodiments, the cysteine containing protein

is a protein illustrated in Table 9. In some embodiments, the cysteine containing protein is a protein illustrated in Table 10A, Table 10B, Table 10C, Table 10D or Table 10E. In some embodiments, the small molecule fragment is a small molecule fragment of Formula (I):



Formula (I), wherein: RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and F is a small molecule fragment moiety. In some embodiments, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some embodiments, F is obtained from a compound library. In some embodiments, the compound library comprises ChemBridge fragment library, Pyramid Platform Fragment-Based Drug Discovery, Maybridge fragment library, FRGx from AnalytiCon, TCI-Frag from AnCoreX, Bio Building Blocks from ASINEX, BioFocus 3D from Charles River, Fragments of Life (FOL) from Emerald Bio, Enamine Fragment Library, IOTA Diverse 1500, BIONET fragments library, Life Chemicals Fragments Collection, OTAVA fragment library, Prestwick fragment library, Selcia fragment library, TimTec fragment-based library, Allium from Vitas-M Laboratory, or Zenobia fragment library. In some embodiments, F is a small molecule fragment moiety illustrated in Fig. 3. In some embodiments, F further comprises a linker moiety that connects F to the carbonyl moiety. In some embodiments, the small molecule fragment is a small molecule fragment illustrated in Fig. 3. In some embodiments, the small molecule fragment is a specific inhibitor or a pan inhibitor. In some embodiments, the cysteine-reactive probe is a cysteine-reactive probe of

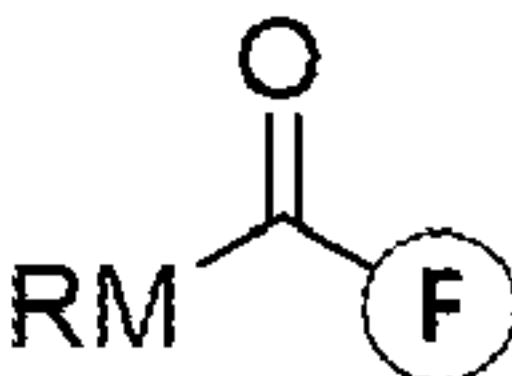


Formula (II), wherein: RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond to the thiol group of a cysteine residue; and AHM is an affinity handle moiety. In some embodiments, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some embodiments, the affinity handle moiety comprises an affinity handle and a binding moiety that facilitates covalent interaction of the cysteine-reactive probe to a cysteine residue of a cysteine-containing protein. In some embodiments, the binding moiety is a small molecule fragment obtained from a compound library. In some embodiments, the compound library comprises ChemBridge fragment library, Pyramid Platform Fragment-Based Drug Discovery, Maybridge fragment library, FRGx from AnalytiCon, TCI-Frag from AnCoreX, Bio Building Blocks from ASINEX, BioFocus 3D from Charles River, Fragments of Life (FOL) from Emerald Bio, Enamine Fragment Library, IOTA Diverse 1500, BIONET fragments library, Life Chemicals Fragments Collection, OTAVA fragment library, Prestwick fragment library, Selcia fragment

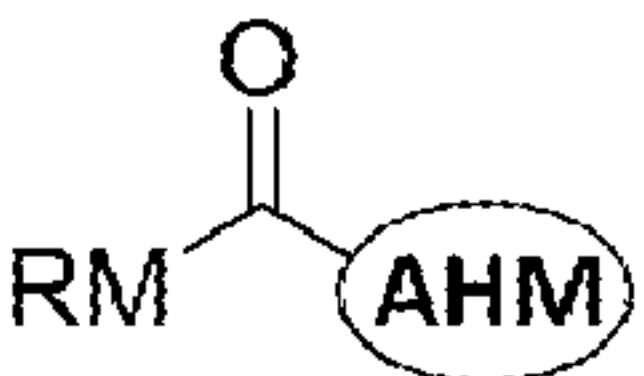
library, TimTec fragment-based library, Allium from Vitas-M Laboratory, or Zenobia fragment library. In some embodiments, the affinity handle is a bioorthogonal affinity handle. In some embodiments, the affinity handle comprises a carbodiimide, N-hydroxysuccinimide (NHS) ester, imidoester, pentafluorophenyl ester, hydroxymethyl phosphine, maleimide, haloacetyl, pyridyl disulfide, thiosulfonate, vinylsulfone, hydrazide, alkoxyamine, alkyne, azide, or isocyanate group. In some embodiments, the affinity handle comprises an alkyne or an azide group. In some embodiments, the affinity handle is further conjugated to an affinity ligand. In some embodiments, the affinity ligand comprises a chromophore, a labeling group, or a combination thereof. In some embodiments, the chromophore comprises fluorochrome, non-fluorochrome chromophore, quencher, an absorption chromophore, fluorophore, organic dye, inorganic dye, metal chelate, or a fluorescent enzyme substrate. In some embodiments, the fluorophore comprises rhodamine, rhodol, fluorescein, thiofluorescein, aminofluorescein, carboxyfluorescein, chlorofluorescein, methylfluorescein, sulfofluorescein, aminorhodol, carboxyrhodol, chlororhodol, methylrhodol, sulforhodol; aminorhodamine, carboxyrhodamine, chlororhodamine, methylrhodamine, sulforhodamine, thiorhodamine, cyanine, indocarbocyanine, oxacarbocyanine, thiocarbocyanine, merocyanine, cyanine 2, cyanine 3, cyanine 3.5, cyanine 5, cyanine 5.5, cyanine 7, oxadiazole derivatives, pyridyloxazole, nitrobenzoxadiazole, benzoxadiazole, pyren derivatives, cascade blue, oxazine derivatives, Nile red, Nile blue, cresyl violet, oxazine 170, acridine derivatives, proflavin, acridine orange, acridine yellow, arylmethine derivatives, auramine, crystal violet, malachite green, tetrapyrrole derivatives, porphin, phtalocyanine, bilirubin 1-dimethylaminonaphthyl-5-sulfonate, 1-anilino-8-naphthalene sulfonate, 2-p-touidiny-6-naphthalene sulfonate, 3-phenyl-7-isocyanatocoumarin, N-(p-(2-benzoxazolyl)phenyl)maleimide, stilbenes, pyrenes, 6-FAM (Fluorescein), 6-FAM (NHS Ester), 5(6)-FAM, 5-FAM, Fluorescein dT, 5-TAMRA-cadavarine, 2-aminoacridone, HEX, JOE (NHS Ester), MAX, TET, ROX, TAMRA, TARMA™ (NHS Ester), TEX 615, ATTO™ 488, ATTO™ 532, ATTO™ 550, ATTO™ 565, ATTO™ Rho101, ATTO™ 590, ATTO™ 633, ATTO™ 647N, TYE™ 563, TYE™ 665, or TYE™ 705. In some embodiments, the labeling group is biotin moiety, streptavidin moiety, bead, resin, a solid support, or a combination thereof. In some embodiments, the affinity handle moiety further comprises a chromophore. In some embodiments, the cysteine-reactive probe is a cysteine-reactive probe illustrated in Fig. 3. In some embodiments, the second cell solution further comprises a control. In some embodiments, the control is dimethyl sulfoxide (DMSO). In some embodiments, the proteomic analysis means comprises a mass spectroscopy method. In some embodiments, the mass spectroscopy method is a liquid-chromatography-mass spectrometry (LC-MS) method. In some embodiments, the method further comprises analyzing the results from the mass

spectroscopy method by an algorithm for protein identification. In some embodiments, the algorithm combines the results from the mass spectroscopy method with a protein sequence database for protein identification. In some embodiments, the algorithm comprises ProLuCID algorithm, Probit, Scaffold, SEQUEST, or Mascot. In some embodiments, the mass spectroscopy method is a MALDI-TOF based method. In some embodiments, the cell is obtained from a tumor cell line. In some embodiments, the cell is obtained from a MDA-MB-231, Ramos, or Jurkat cell line. In some embodiments, the cell is obtained from a tumor sample. In some embodiments, the sample is a tissue sample. In some embodiments, the method is an *in situ* method.

[0008] Disclosed herein, in certain embodiments, is a composition comprising: a small

molecule fragment of Formula (I):  Formula (I), wherein: RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and F is a small molecule fragment moiety; and a cysteine containing protein wherein the cysteine containing protein is covalently bond to the small molecule fragment. In some embodiments, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some embodiments, F is obtained from a compound library. In some embodiments, F is a small molecule fragment moiety illustrated in Fig. 3. In some embodiments, F further comprises a linker moiety that connects F to the carbonyl moiety.

[0009] Disclosed herein, in certain embodiments, is a composition comprising: a cysteine-

reactive probe of Formula (II):  Formula (II), wherein: RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond to the thiol group of a cysteine residue; and AHM is an affinity handle moiety; and a cysteine containing protein wherein the cysteine containing protein is covalently bond to the cysteine-reactive probe. In some embodiments, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some embodiments, the affinity handle moiety comprises an affinity handle and a binding moiety that facilitates covalent interaction of the cysteine-reactive probe to a cysteine residue of a cysteine-containing protein. In some embodiments, the binding moiety is a small molecule fragment obtained from a compound library. In some embodiments, the affinity handle is a bioorthogonal affinity handle. In some embodiments, the affinity handle comprises a carbodiimide, N-hydroxysuccinimide (NHS) ester, imidoester, pentafluorophenyl ester, hydroxymethyl phosphine, maleimide, haloacetyl, pyridyl

disulfide, thiosulfonate, vinylsulfone, hydrazide, alkoxyamine, alkyne, azide, or isocyanate group. In some embodiments, the affinity handle is further conjugated to an affinity ligand. In some embodiments, the affinity handle moiety further comprises a chromophore. In some embodiments, the cysteine-reactive probe is a cysteine-reactive probe illustrated in Fig. 3.

[0010] Disclosed herein, in certain embodiments, is a composition comprising: an isolated sample wherein the isolated sample is an isolated cell or a tissue sample; and a cysteine-reactive probe to be assayed for its ability to interact with a cysteine containing protein expressed in the isolated sample. In some embodiments, the composition further comprises contacting the isolated sample with a small molecule fragment for an extended period of time prior to incubating the isolated sample with the cysteine-reactive probe to generate a cysteine-reactive probe-protein complex. In some embodiments, the extended period of time is about 5, 10, 15, 20, 30, 60, 90, 120 minutes or longer.

[0011] Disclosed herein, in certain embodiments, is an isolated treated cell comprising a cysteine-reactive probe covalently attached to a cysteine containing protein. In some embodiments, the isolated treated cell further comprises a set of cysteine-reactive probes wherein each of the cysteine-reactive probes is covalently attached to a cysteine containing protein.

[0012] Disclosed herein, in certain embodiments, is an isolated treated cell comprising a small molecule fragment covalently attached to a cysteine containing protein. In some embodiments, the isolated treated cell further comprises a set of small molecule fragments wherein each of the small molecule fragments is covalently attached to a cysteine containing protein. In some embodiments, the isolated treated cell further comprises a cysteine-reactive probe. In some embodiments, the isolated treated cell further comprises a set of cysteine-reactive probes.

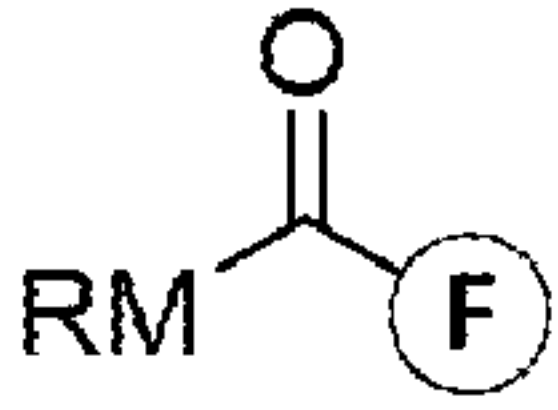
[0013] Disclosed herein, in certain embodiments, is an isolated treated population of cells comprising a set of cysteine-reactive probes covalently attached to cysteine containing proteins. Also disclosed herein, in certain embodiments, is an isolated treated population of cells comprising a set of small molecule fragments covalently attached to cysteine containing proteins. In some embodiments, the isolated treated population of cells further comprises a set of cysteine-reactive probes.

[0014] Disclosed herein, in certain embodiments, is an isolated and purified polypeptide comprising at least 90% sequence identity to at least seven contiguous amino acids of an amino acid sequence selected from Tables 1-3 or 8-9. In some embodiments, the isolated and purified polypeptide comprising at least 95% sequence identity to at least seven contiguous amino acids of an amino acid sequence selected from Tables 1-3 or 8-9. In some embodiments, the isolated and purified polypeptide comprising 100% sequence identity to at least seven contiguous amino

acids of an amino acid sequence selected from Tables 1-3 or 8-9. In some embodiments, the isolated and purified polypeptide consisting 100% sequence identity to the full length of an amino acid sequence selected from Tables 1-3 or 8-9. In some embodiments, the isolated and purified polypeptide is at most 50 amino acids in length. A polypeptide probe for screening a small molecule fragment comprising an isolated and purified polypeptide described herein.

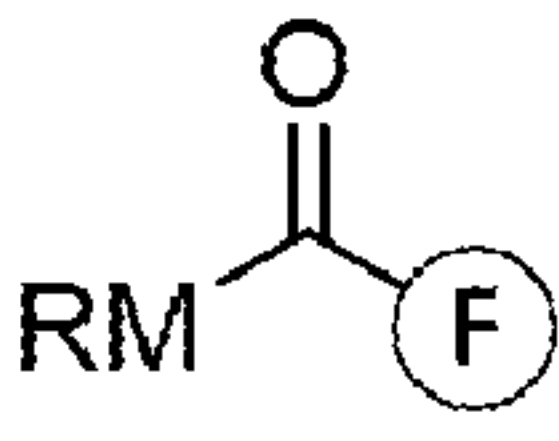
[0015] Further disclosed herein, in certain embodiments, is a nucleic acid encoding a polypeptide comprising at least 90% sequence identity at least seven contiguous amino acids of an amino acid sequence selected from Tables 1-3 or 8-9. In some embodiments, the nucleic acid encoding a polypeptide comprising at least 95% sequence identity at least seven contiguous amino acids of an amino acid sequence selected from Tables 1-3 or 8-9. In some embodiments, the nucleic acid encoding a polypeptide comprising 100% sequence identity at least seven contiguous amino acids of an amino acid sequence selected from Tables 1-3 or 8-9. In some embodiments, the nucleic acid encoding a polypeptide consisting 100% sequence identity to the full length of an amino acid sequence selected from Tables 1-3 or 8-9.

[0016] Disclosed herein, in certain embodiments, is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, wherein the small molecule fragment is a small molecule fragment of

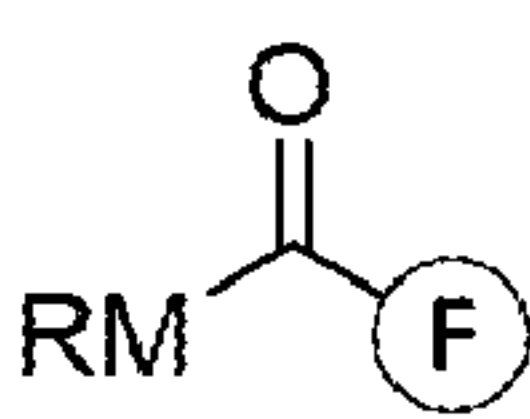
Formula (I):  Formula (I), wherein: RM is a reactive moiety selected from a Michael

acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and F is a small molecule fragment moiety. In some embodiments, the cysteine containing protein is a protein illustrated in Table 1. In some embodiments, the cysteine containing protein is a protein illustrated in Table 2. In some embodiments, the cysteine containing protein is a protein illustrated in Table 3. In some embodiments, the cysteine containing protein comprises a cysteine residue denoted in Table 3. In some embodiments, the cysteine containing protein is a protein illustrated in Table 8. In some embodiments, the cysteine containing protein is a protein illustrated in Table 9. In some embodiments, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some embodiments, F is obtained from a compound library. In some embodiments, F is a small molecule fragment moiety illustrated in Fig. 3. In some embodiments, F further comprises a linker moiety that connects F to the carbonyl moiety. In some embodiments, the small molecule fragment binds irreversibly to the cysteine containing protein. In some embodiments, the small molecule fragment binds reversibly to the cysteine containing protein.

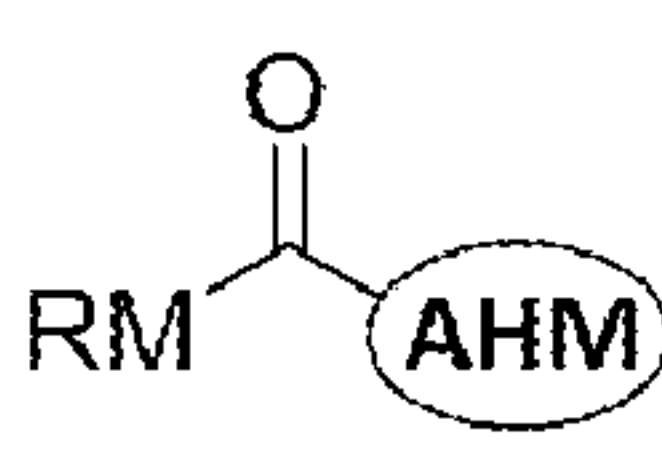
[0017] Disclosed herein, in certain embodiments, is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, wherein the small molecule fragment has a molecular weight of about 150 Dalton or higher. In some embodiments, the small molecule fragment has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some embodiments, the molecular weight of the small molecule fragment is prior to enrichment with a halogen, a nonmetal, or a transition metal. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur atoms. In some embodiments, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some embodiments, the cysteine containing protein is about 20, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000 amino acid residues in length or more. In some embodiments, the cysteine containing protein is a protein illustrated in Table 1. In some embodiments, the cysteine containing protein is a protein illustrated in Table 2. In some embodiments, the cysteine containing protein is a protein illustrated in Table 3. In some embodiments, the cysteine containing protein comprises a cysteine residue denoted in Table 3. In some embodiments, the cysteine containing protein is a protein illustrated in Table 8. In some embodiments, the cysteine containing protein is a protein illustrated in Table 9. In some embodiments, the small molecule fragment is a small molecule

fragment of Formula (I):  Formula (I), wherein: RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and F is a small molecule fragment moiety. In some embodiments, the small molecule fragment of Formula (I) has a molecular weight of about 150, 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some embodiments, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some embodiments, F is obtained from a compound library. In some embodiments, F is a small molecule fragment moiety illustrated in Fig. 3. In some embodiments, F further comprises a linker moiety that connects F to the carbonyl moiety. In some embodiments, the small molecule fragment bond irreversibly to the cysteine containing protein. In some embodiments, the small molecule fragment bond reversibly to the cysteine containing protein.

[0018] Disclosed herein, in certain embodiments, is a cysteine containing protein-small molecule fragment complex produced by a process comprising contacting a cell solution with a

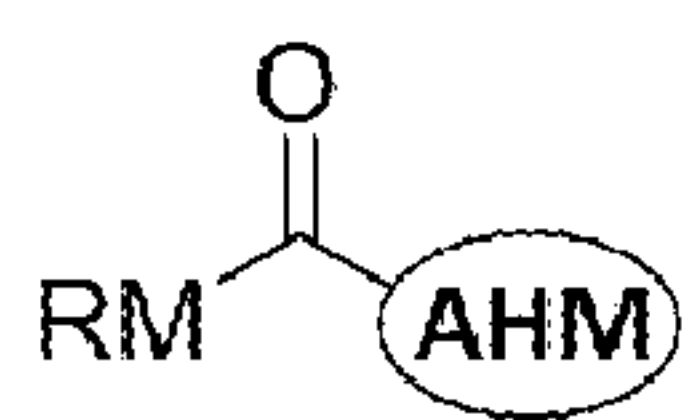
small molecule fragment of Formula (I):  Formula (I), wherein: RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and F is a small molecule fragment moiety; and wherein the contacting time is between about 5 minutes and about 2 hours. In some embodiments, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some embodiments, F is obtained from a compound library. In some embodiments, F is a small molecule fragment moiety illustrated in Fig. 3. In some embodiments, F further comprises a linker moiety that connects F to the carbonyl moiety. In some embodiments, the small molecule fragment of Formula (I) has a molecular weight of about 150, 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some embodiments, the cysteine containing protein is a protein illustrated in Table 1. In some embodiments, the cysteine containing protein is a protein illustrated in Table 2. In some embodiments, the cysteine containing protein is a protein illustrated in Table 3. In some embodiments, the cysteine containing protein comprises a cysteine residue denoted in Table 3. In some embodiments, the cysteine containing protein is a protein illustrated in Table 8. In some embodiments, the cysteine containing protein is a protein illustrated in Table 9. In some embodiments, the cysteine containing protein is a protein illustrated in Table 10A, Table 10B, Table 10C, Table 10D or Table 10E. In some embodiments, the small molecule fragment binds irreversibly to the cysteine containing protein. In some embodiments, the small molecule fragment binds reversibly to the cysteine containing protein.

[0019] Disclosed herein, in certain embodiments, is a modified cysteine containing protein comprising a cysteine-reactive probe having a covalent bond to a cysteine residue of a cysteine containing protein, wherein the cysteine-reactive probe is a cysteine-reactive probe of Formula

(II):  Formula (II), wherein: RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond to the thiol group of a cysteine residue; and AHM is an affinity handle moiety. In some embodiments, the cysteine containing protein is a protein illustrated in Table 1. In some embodiments, the cysteine containing protein is a protein illustrated in Table 2. In some embodiments, the cysteine containing protein is a protein illustrated in Table 8. In some embodiments, the cysteine containing protein is a protein illustrated in Table 9. In some embodiments, the

cysteine containing protein is a protein illustrated in Table 10A, Table 10B, Table 10C, Table 10D or Table 10E. In some embodiments, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some embodiments, the affinity handle moiety comprises an affinity handle and a binding moiety that facilitates covalent interaction of the cysteine-reactive probe to a cysteine residue of a cysteine-containing protein. In some embodiments, the binding moiety is a small molecule fragment obtained from a compound library. In some embodiments, the affinity handle is a bioorthogonal affinity handle. In some embodiments, the affinity handle comprises a carbodiimide, N-hydroxysuccinimide (NHS) ester, imidoester, pentafluorophenyl ester, hydroxymethyl phosphine, maleimide, haloacetyl, pyridyl disulfide, thiosulfonate, vinylsulfone, hydrazide, alkoxyamine, alkyne, azide, or isocyanate group. In some embodiments, the affinity handle is further conjugated to an affinity ligand. In some embodiments, the affinity handle moiety further comprises a chromophore. In some embodiments, the cysteine-reactive probe is a cysteine-reactive probe illustrated in Fig. 3. In some embodiments, the cysteine-reactive probe binds irreversibly to the cysteine containing protein. In some embodiments, the cysteine-reactive probe binds reversibly to the cysteine containing protein.

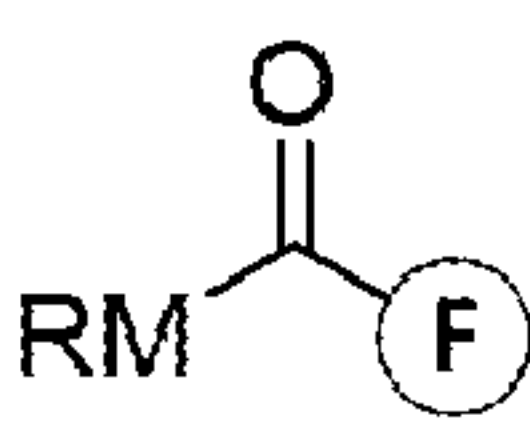
[0020] Disclosed herein, in certain embodiments, is a cysteine-reactive probe of Formula (II):



Formula (II), wherein: RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond to the thiol group of a cysteine residue; and AHM is an affinity handle moiety. In some embodiments, the cysteine-reactive probe covalently binds to a cysteine residue on a cysteine containing protein. In some embodiments, cysteine containing protein is a protein illustrated in Table 1. In some embodiments, the cysteine containing protein is a protein illustrated in Table 2. In some embodiments, the cysteine containing protein is a protein illustrated in Table 3. In some embodiments, the cysteine containing protein comprises a cysteine residue denoted in Table 3. In some embodiments, the cysteine containing protein is a protein illustrated in Table 8. In some embodiments, the cysteine containing protein is a protein illustrated in Table 9. In some embodiments, the cysteine containing protein is a protein illustrated in Table 10A, Table 10B, Table 10C, Table 10D or Table 10E. In some embodiments, the cysteine-reactive probe binds irreversibly to the cysteine containing protein. In some embodiments, the cysteine-reactive probe binds reversibly to the cysteine containing protein.

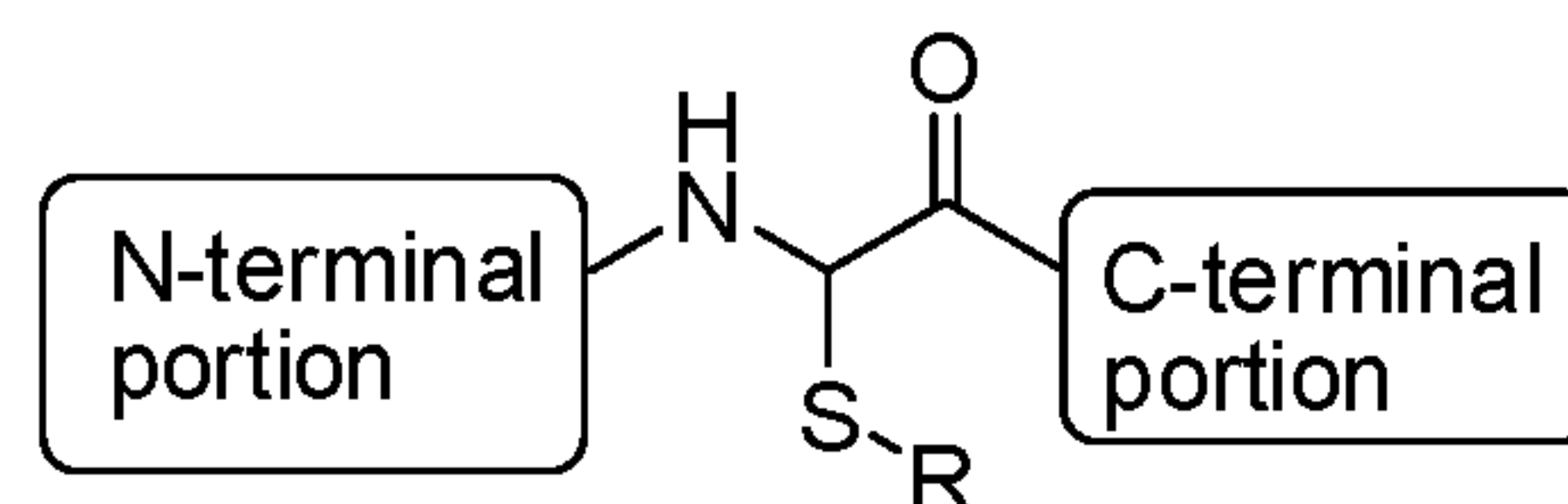
[0021] Disclosed herein, in certain embodiments, is a compound capable of covalently binding to a cysteine containing protein identified, using the method comprising: (a) obtaining a set of

cysteine-reactive probe-protein complexes from a sample wherein the cysteine-reactive probe comprises a reactive moiety capable of forming a covalent bond with a cysteine residue located on the cysteine containing protein; (b) analyzing the set of cysteine-reactive probe-protein complexes by a proteomic analysis means; (c) based on step b), identifying a cysteine containing protein as the binding target for the compound. In some embodiments, the compound is a small molecule fragment. In some embodiments, the small molecule fragment is a small molecule

fragment of Formula (I):  Formula (I), wherein: RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and F is a small molecule fragment moiety. In some embodiments, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some embodiments, F is obtained from a compound library. In some embodiments, the compound library comprises ChemBridge fragment library, Pyramid Platform Fragment-Based Drug Discovery, Maybridge fragment library, FRGx from AnalytiCon, TCI-Frag from AnCoreX, Bio Building Blocks from ASINEX, BioFocus 3D from Charles River, Fragments of Life (FOL) from Emerald Bio, Enamine Fragment Library, IOTA Diverse 1500, BIONET fragments library, Life Chemicals Fragments Collection, OTAVA fragment library, Prestwick fragment library, Selcia fragment library, TimTec fragment-based library, Allium from Vitas-M Laboratory, or Zenobia fragment library. In some embodiments, F is a small molecule fragment moiety illustrated in Fig. 3. In some embodiments, F further comprises a linker moiety that connects F to the carbonyl moiety. In some embodiments, the small molecule fragment is a small molecule fragment illustrated in Fig. 3. In some embodiments, the small molecule fragment is a specific inhibitor or a pan inhibitor. In some embodiments, the cysteine containing protein comprises a biologically active cysteine residue. In some embodiments, the biologically active cysteine site is a cysteine residue that is located about 10Å or less to an active-site ligand or residue. In some embodiments, the cysteine residue that is located about 10Å or less to the active-site ligand or residue is an active site cysteine. In some embodiments, the biologically active cysteine site is an active site cysteine. In some embodiments, the biologically active cysteine site is a cysteine residue that is located greater than 10Å from an active-site ligand or residue. In some embodiments, the cysteine residue that is located greater than 10Å from the active-site ligand or residue is a non-active site cysteine. In some embodiments, the biologically active cysteine site is a non-active site cysteine. In some embodiments, the small molecule fragment that covalently interacts with the biologically active cysteine impairs and/or inhibits activity of the cysteine containing protein. In some embodiments, the cysteine containing protein

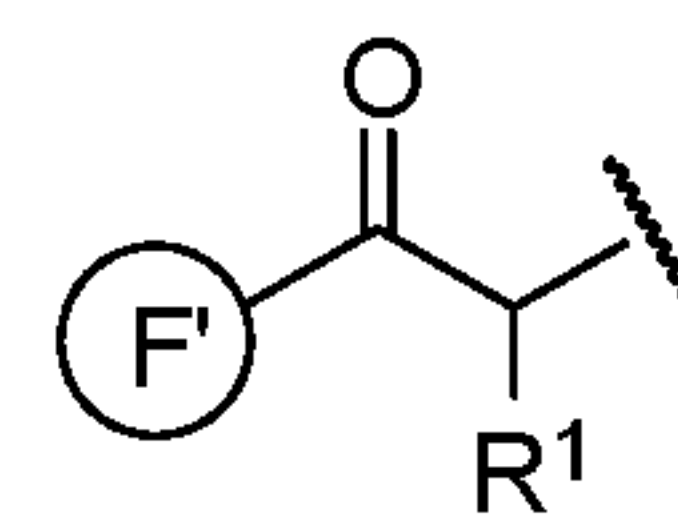
exists in an active form. In some embodiments, the small molecule fragment and/or the cysteine-reactive probe interact with the active form of the cysteine containing protein. In some embodiments, the cysteine containing protein exists in a pro-active form. In some embodiments, the small molecule fragment and/or the cysteine-reactive probe interact with the pro-active form of the cysteine containing protein. In some embodiments, the structural environment of the biologically active cysteine residue modulates the reactivity of the cysteine residue. In some embodiments, the structural environment is a hydrophobic environment or a hydrophilic environment. In some embodiments, the structural environment is a charged environment. In some embodiments, the structural environment is a nucleophilic environment. In some embodiments, the cysteine containing protein is an enzyme, a transporter, a receptor, a channel protein, an adaptor protein, a chaperone, a signaling protein, a plasma protein, transcription related protein, translation related protein, mitochondrial protein, or cytoskeleton related protein. In some embodiments, the cysteine containing protein is an enzyme, a transporter, a receptor, a channel protein, an adaptor protein, a chaperone, a signaling protein, transcription related protein, or translation related protein. In some embodiments, the enzyme comprises kinases, proteases, or deubiquitinating enzymes. In some embodiments, the protease is a cysteine protease. In some embodiments, the cysteine protease comprises caspases. In some embodiments, the signaling protein comprises vascular endothelial growth factor. In some embodiments, the signaling protein comprises a redox signaling protein. In some embodiments, the cysteine containing protein is a protein illustrated in Table 1. In some embodiments, the cysteine containing protein is a protein illustrated in Table 2. In some embodiments, the cysteine containing protein is a protein illustrated in Table 3. In some embodiments, the cysteine containing protein comprises a cysteine residue denoted in Table 3. In some embodiments, the cysteine containing protein is a protein illustrated in Table 8. In some embodiments, the cysteine containing protein is a protein illustrated in Table 9. In some embodiments, the cysteine containing protein is a protein illustrated in Table 10A, Table 10B, Table 10C, Table 10D or Table 10E.

[0022] Disclosed herein, in certain embodiments, is a derivative of a cysteine-containing

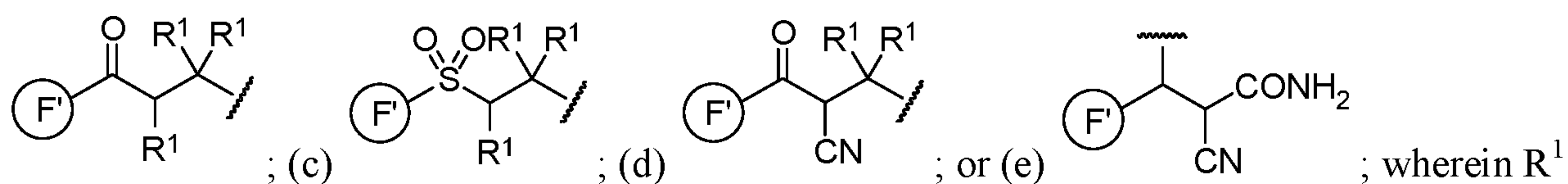


protein having the structure of Formula (I),

(I), wherein,

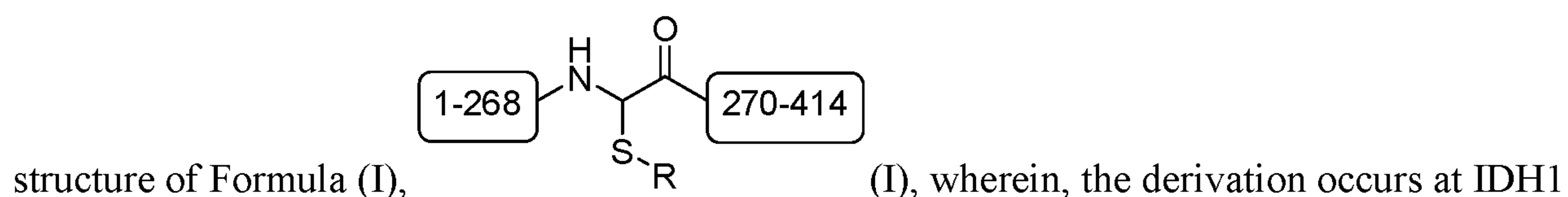


the derivation occurs at a cysteine residue; R is selected from: (a) ; (b)

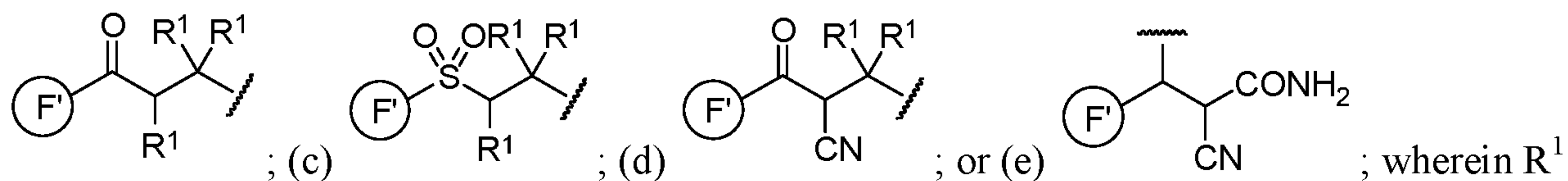


is H, C1-C3 alkyl, or aryl; and F' is a small molecule fragment moiety. In some embodiments, F' has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some embodiments, the molecular weight of F' is prior to enrichment with a halogen, a nonmetal, or a transition metal. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur atoms. In some embodiments, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some embodiments, F' is a small molecule fragment moiety illustrated in Fig. 3. In some embodiments, the cysteine containing protein is a cysteine containing protein described herein. In some embodiments, the cysteine containing protein is a protein illustrated in Tables 1, 2, 3, 8 or 9. In some embodiments, the cysteine containing protein is a protein illustrated in Table 1. In some embodiments, the cysteine containing protein is a protein illustrated in Table 2. In some embodiments, the cysteine containing protein is a protein illustrated in Table 3. In some embodiments, the cysteine containing protein is a protein illustrated in Table 8. In some embodiments, the cysteine containing protein is a protein illustrated in Table 9.

[0023] Disclosed herein, in certain embodiments, is a derivative of IDH1 protein having the



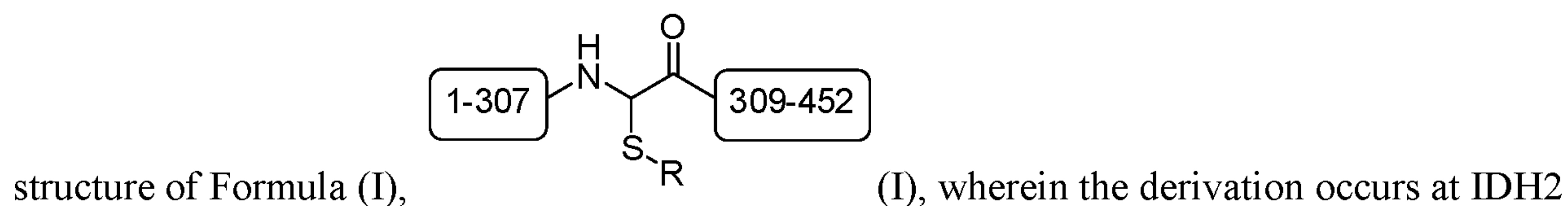
cysteine residue position 269 based on SEQ ID NO: 1; R is selected from: (a) ; (b)

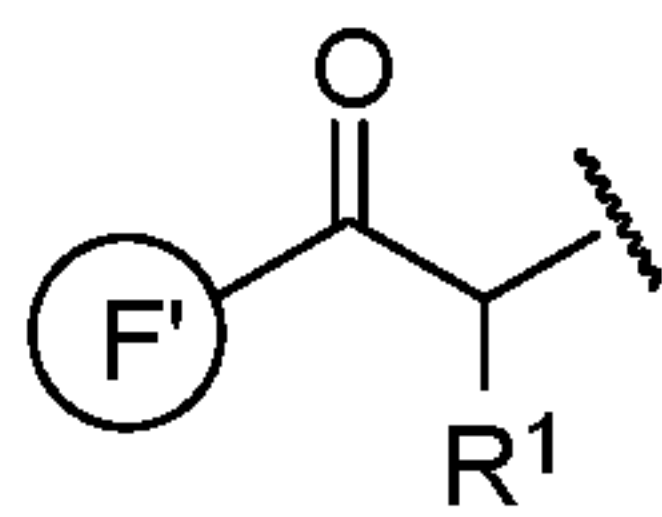


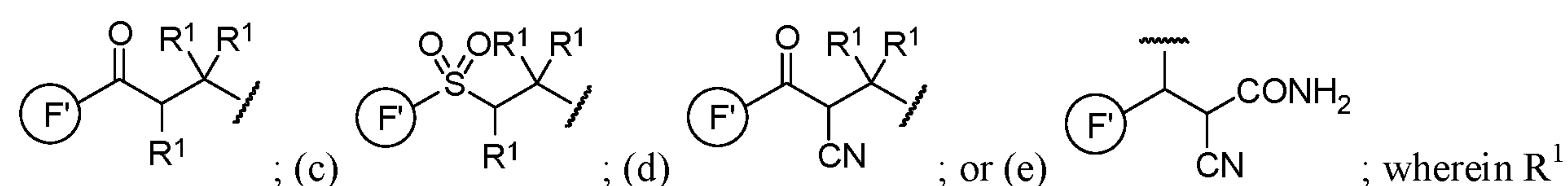
is H, C1-C3 alkyl, or aryl; and F' is a small molecule fragment moiety. In some embodiments, F' has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some embodiments, the molecular weight of F' is prior to enrichment with a halogen, a nonmetal, or a transition metal. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur

atoms. In some embodiments, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some embodiments, F' is a small molecule fragment moiety illustrated in Fig. 3.

[0024] Disclosed herein, in certain embodiments, is a derivative of IDH2 protein having the

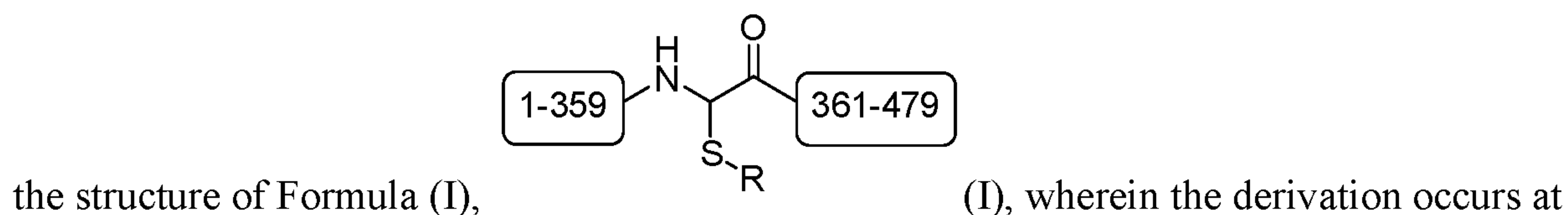


cysteine residue position 308 based on SEQ ID NO: 2; R is selected from: (a) ; (b)

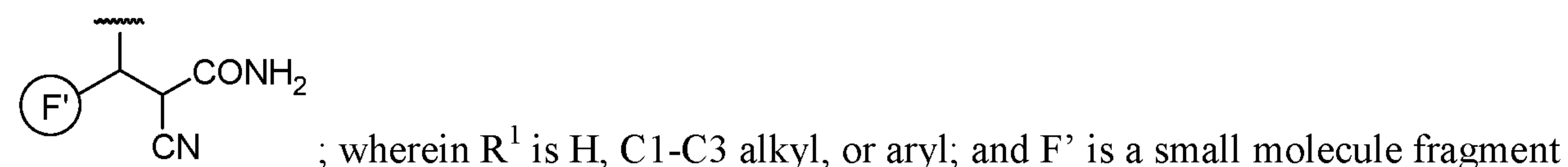
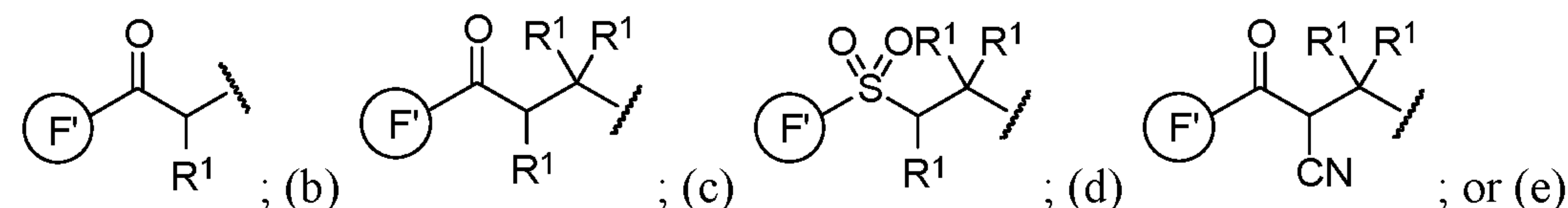


is H, C1-C3 alkyl, or aryl; and F' is a small molecule fragment moiety. In some embodiments, F' has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some embodiments, the molecular weight of F' is prior to enrichment with a halogen, a nonmetal, or a transition metal. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur atoms. In some embodiments, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some embodiments, F' is a small molecule fragment moiety illustrated in Fig. 3.

[0025] Disclosed herein, in certain embodiments, is a derivative of caspase-8 protein having



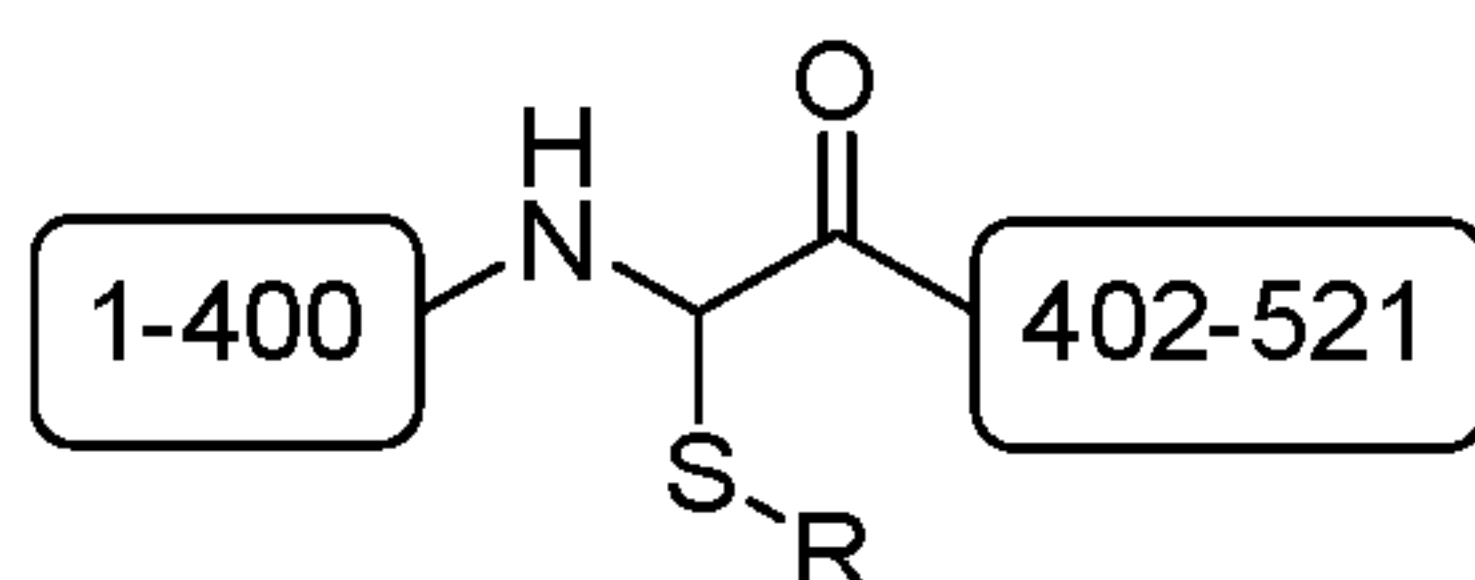
caspase-8 cysteine residue position 360 based on SEQ ID NO: 3; R is selected from: (a)



moiety. In some embodiments, F' has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some

embodiments, the molecular weight of F' is prior to enrichment with a halogen, a nonmetal, or a transition metal. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur atoms. In some embodiments, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some embodiments, F' is a small molecule fragment moiety illustrated in Fig. 3.

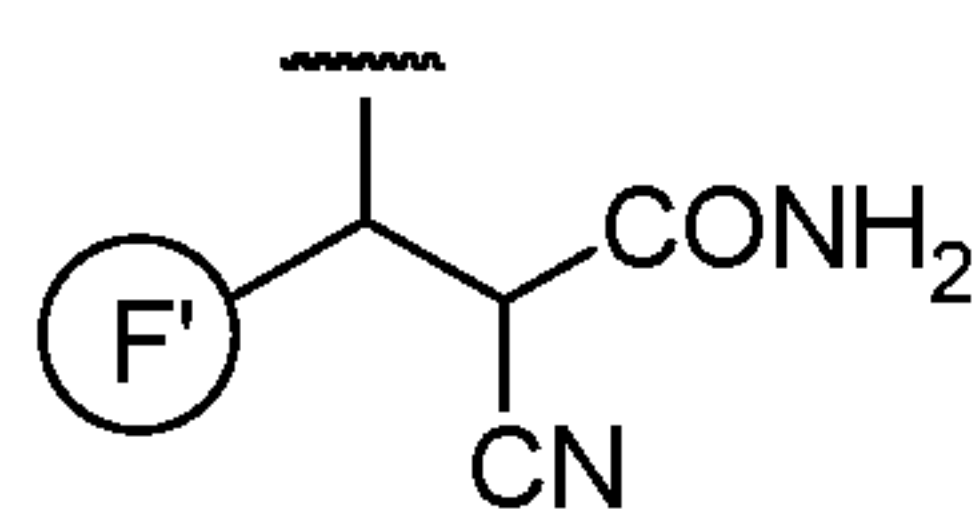
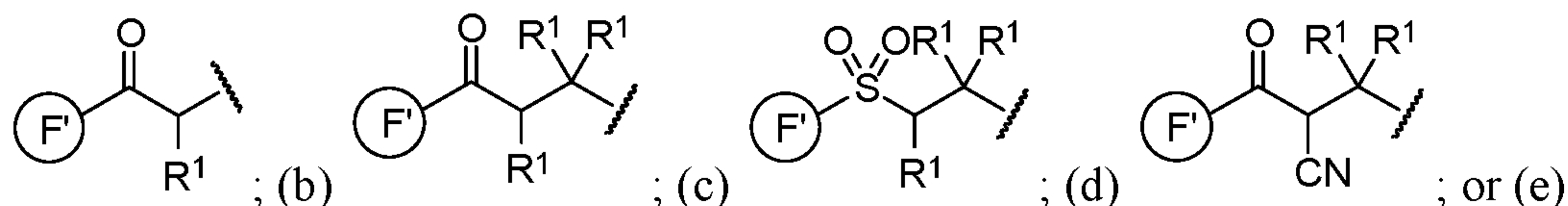
[0026] Disclosed herein, in certain embodiments, is a derivative of caspase-10 protein having



the structure of Formula (I),

(I), wherein the derivation occurs at

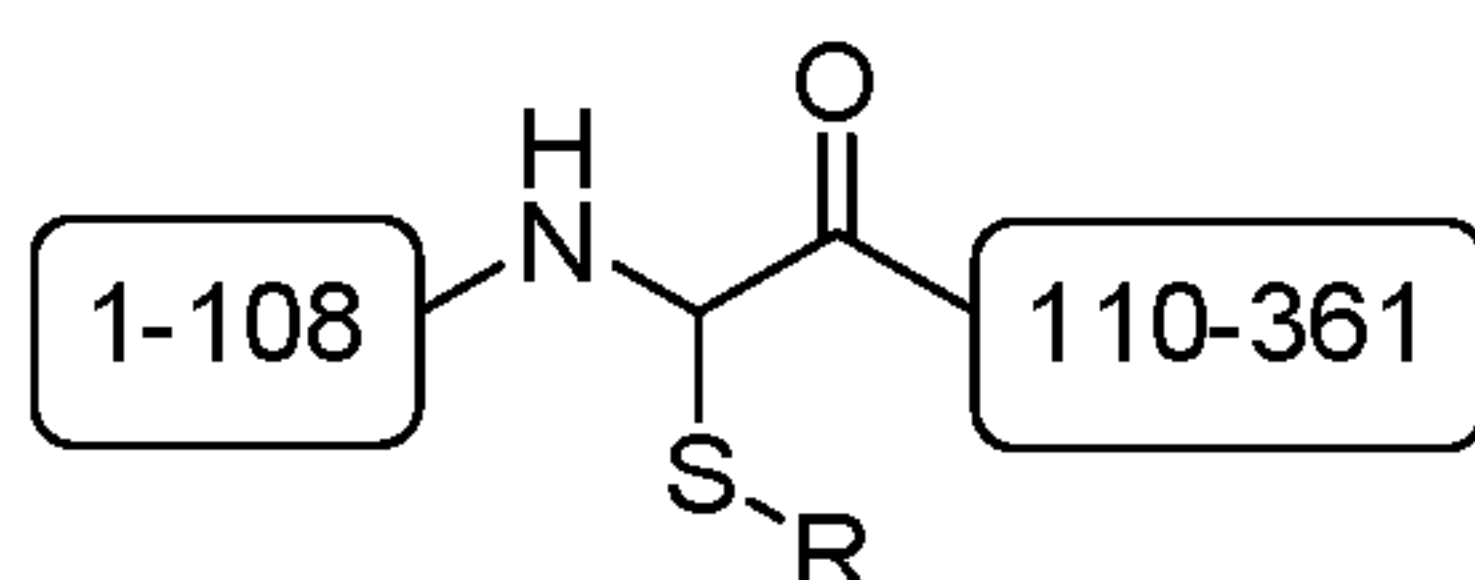
caspase-10 cysteine residue position 401 based on SEQ ID NO: 4; R is selected from: (a)



; wherein R¹ is H, C1-C3 alkyl, or aryl; and F' is a small molecule fragment

moiety. In some embodiments, F' has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some embodiments, the molecular weight of F' is prior to enrichment with a halogen, a nonmetal, or a transition metal. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur atoms. In some embodiments, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some embodiments, F' is a small molecule fragment moiety illustrated in Fig. 3.

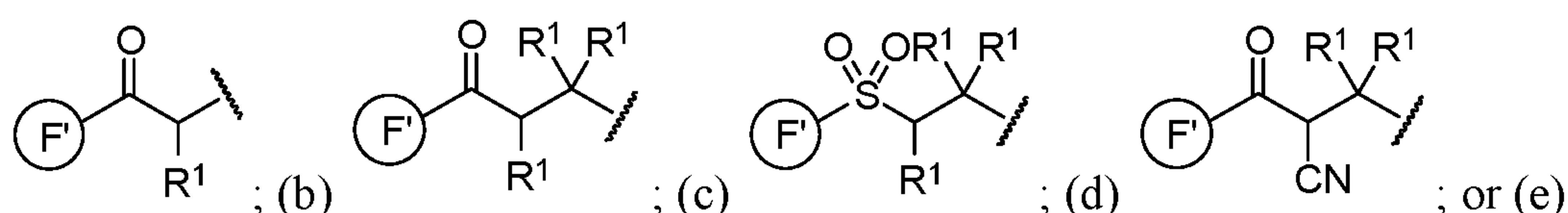
[0027] Disclosed herein, in certain embodiments, is a derivative of PRMT-1 protein having

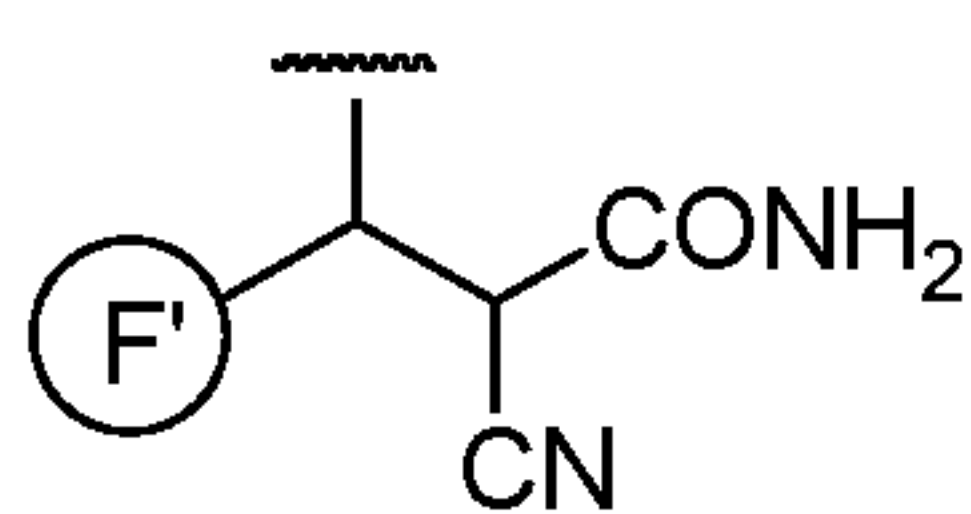


the structure of Formula (I),

(I), wherein the derivation occurs at

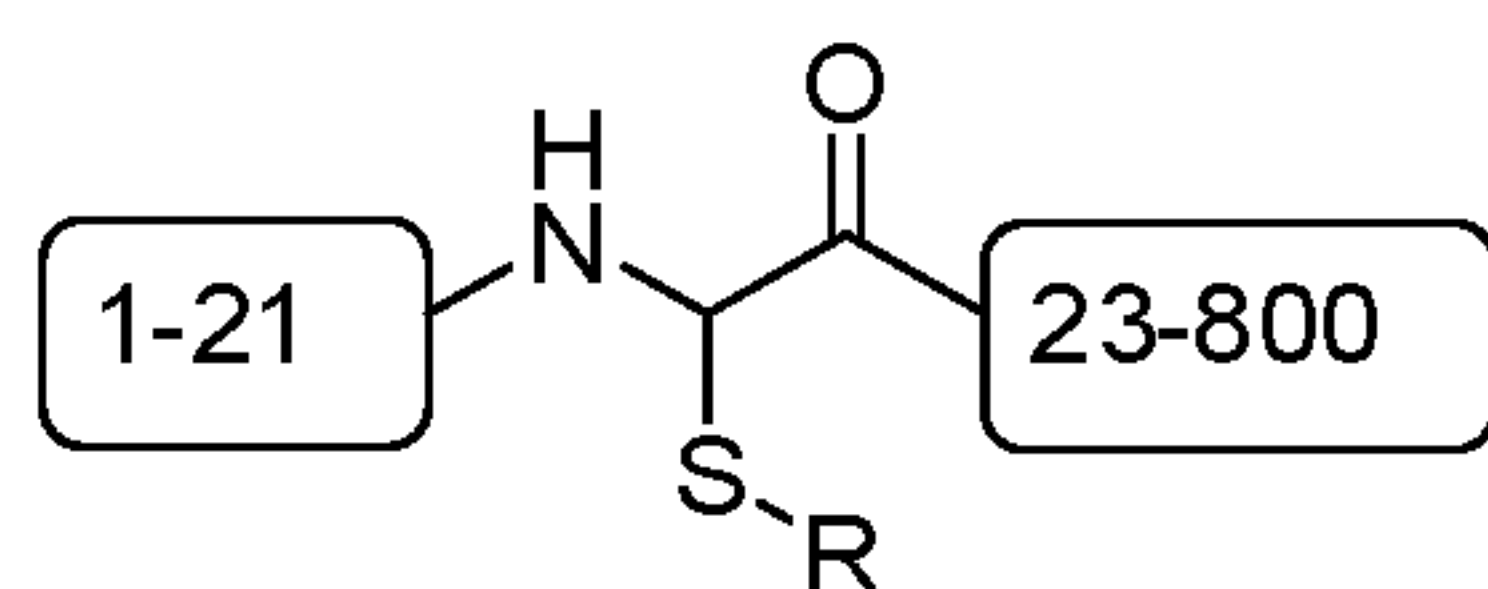
PRMT-1 cysteine residue position 109 based on SEQ ID NO: 5; R is selected from: (a)





; wherein R^1 is H, C1-C3 alkyl, or aryl; and F' is a small molecule fragment moiety. In some embodiments, F' has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some embodiments, the molecular weight of F' is prior to enrichment with a halogen, a nonmetal, or a transition metal. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur atoms. In some embodiments, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some embodiments, F' is a small molecule fragment moiety illustrated in Fig. 3.

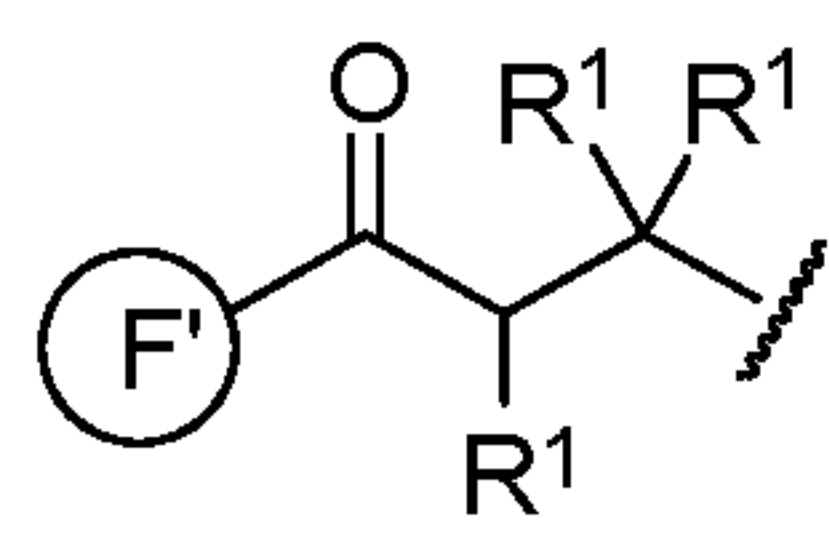
[0028] Disclosed herein, in certain embodiments, is a derivative of ZAK protein having the



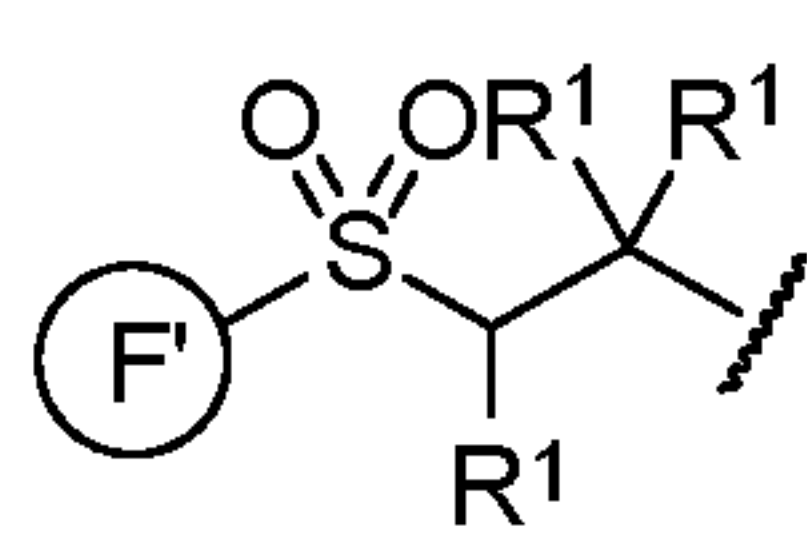
structure of Formula (I),

(I), wherein the derivation occurs at ZAK

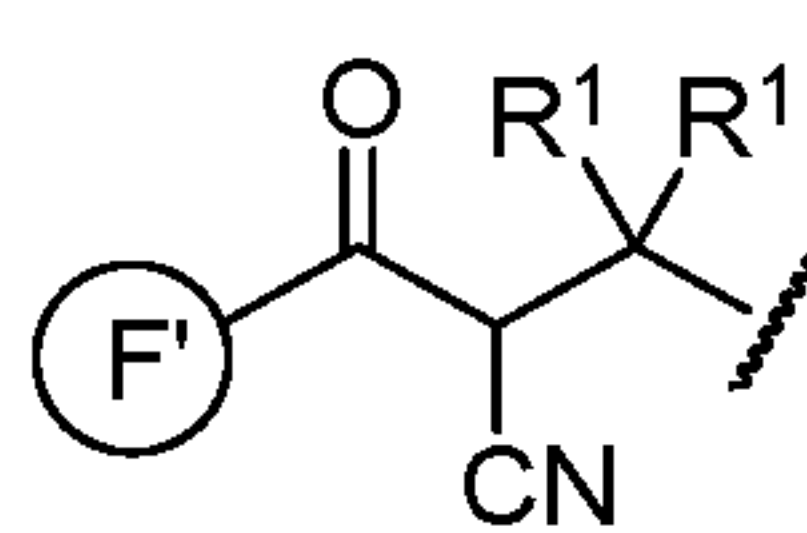
cysteine residue position 22 based on SEQ ID NO: 6; R is selected from: (a) ; (b)



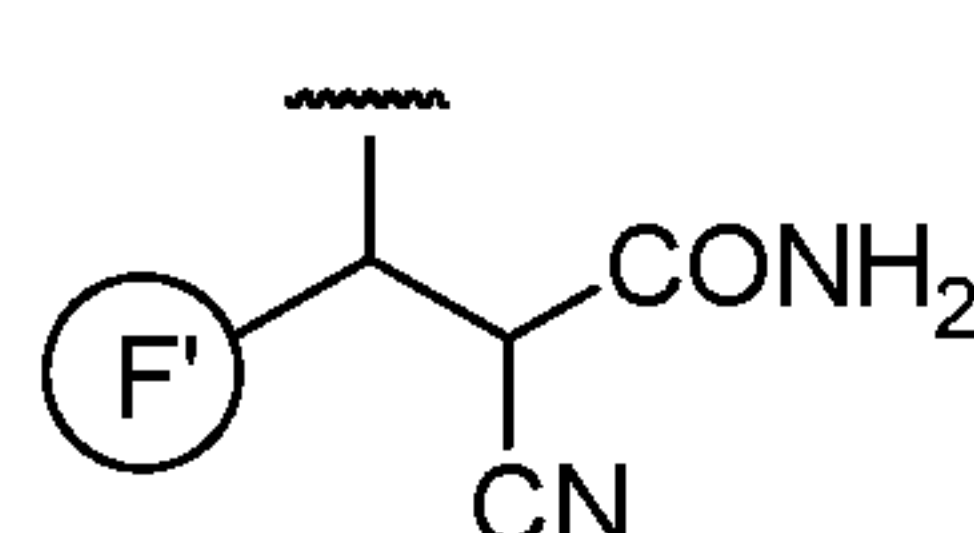
; (c)



; (d)



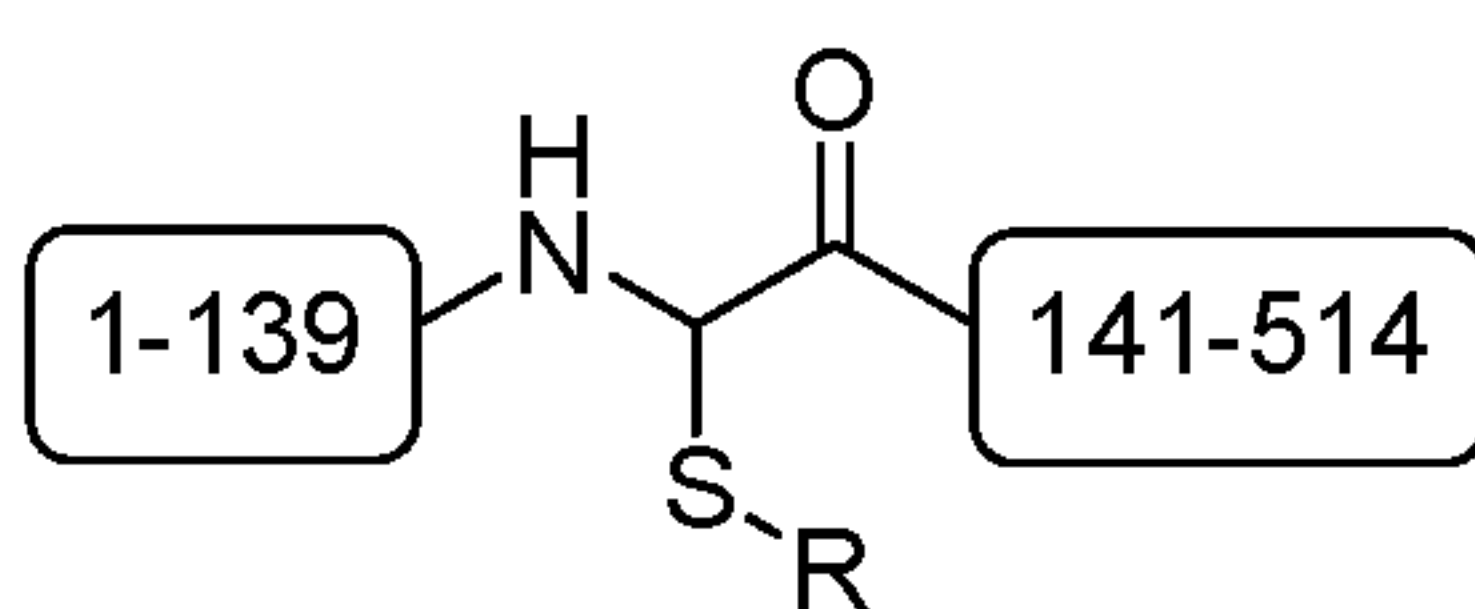
; or (e)



; wherein R^1

is H, C1-C3 alkyl, or aryl; and F' is a small molecule fragment moiety. In some embodiments, F' has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some embodiments, the molecular weight of F' is prior to enrichment with a halogen, a nonmetal, or a transition metal. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur atoms. In some embodiments, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some embodiments, F' is a small molecule fragment moiety illustrated in Fig. 3.

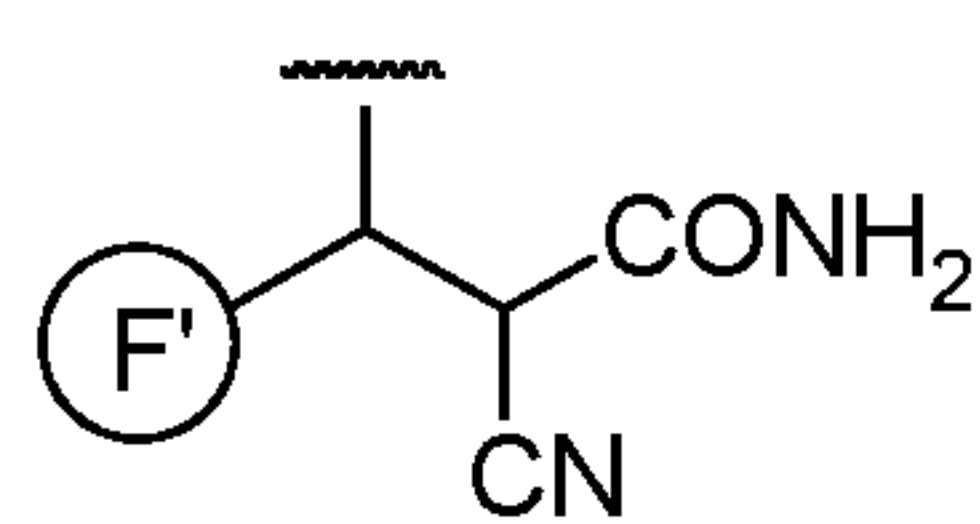
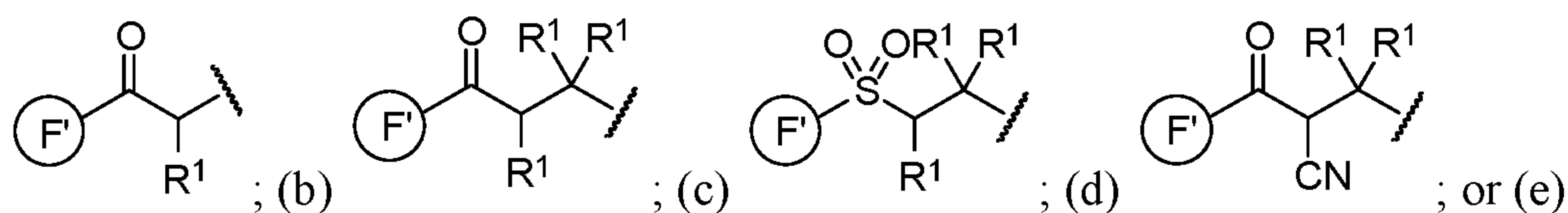
[0029] Disclosed herein, in certain embodiments, is a derivative of IMPDH2 protein having



the structure of Formula (I),

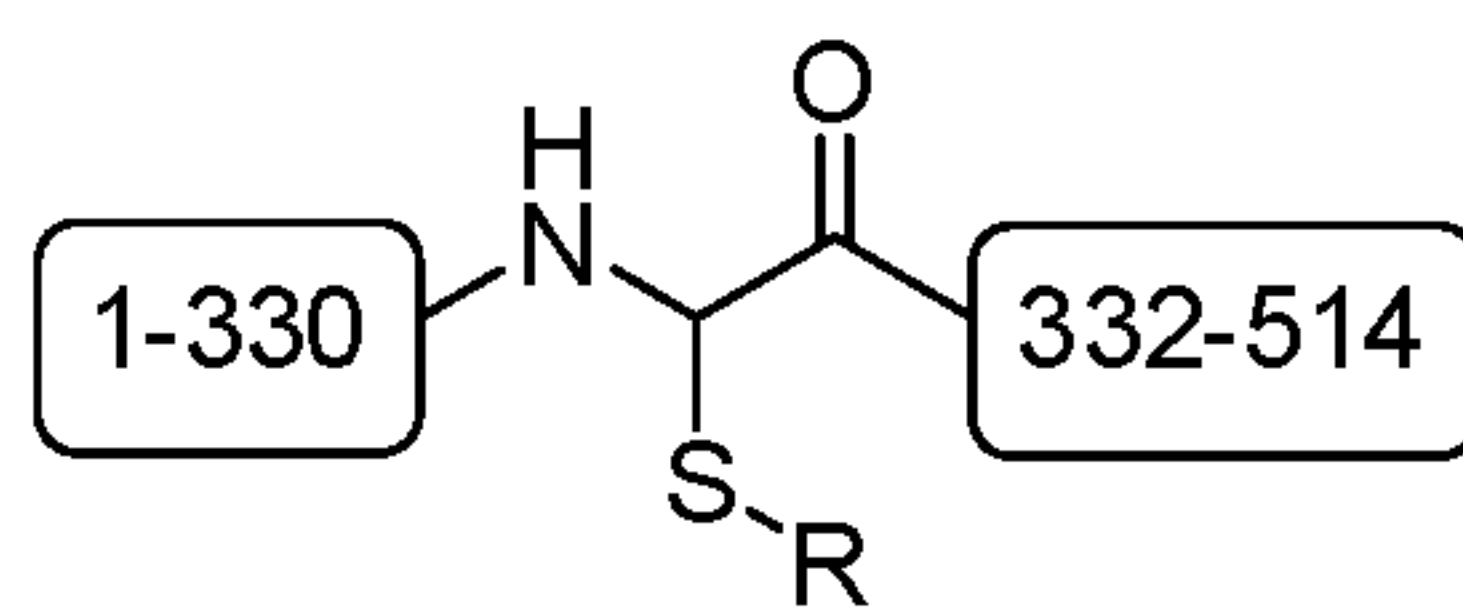
(I), wherein the derivation occurs at

IMPDH2 cysteine residue position 140 based on SEQ ID NO: 7; R is selected from: (a)

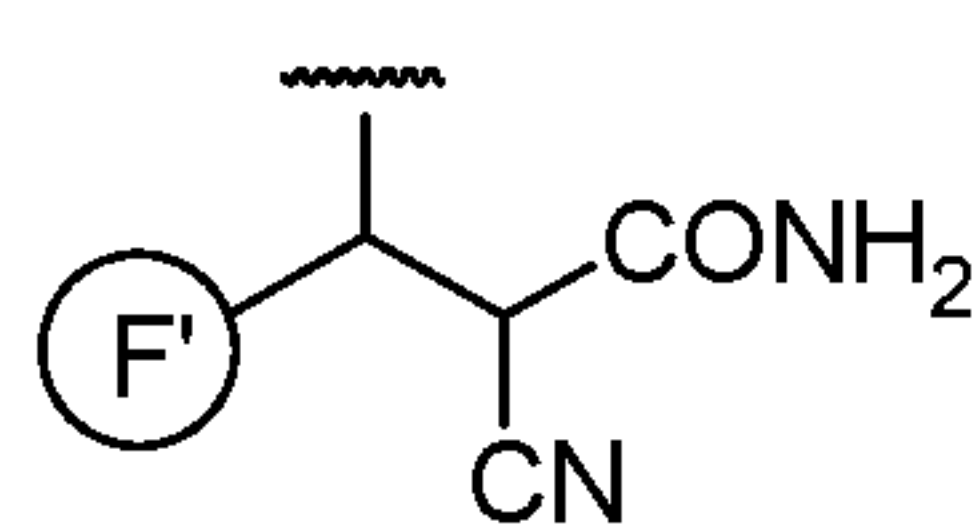
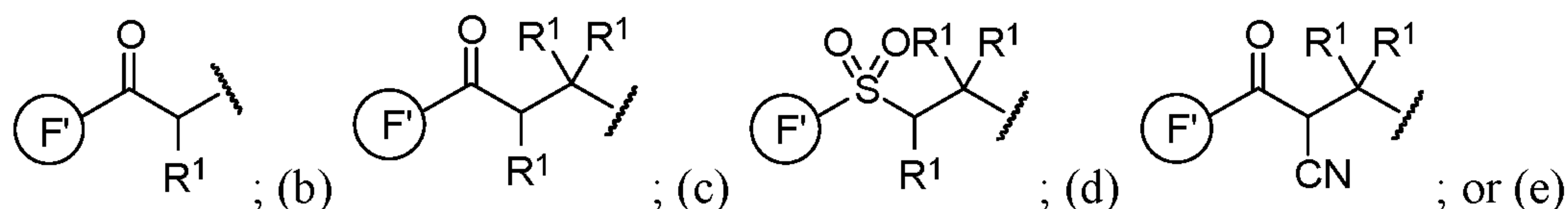


; wherein R^1 is H, C1-C3 alkyl, or aryl; and F' is a small molecule fragment moiety. In some embodiments, F' has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some embodiments, the molecular weight of F' is prior to enrichment with a halogen, a nonmetal, or a transition metal. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur atoms. In some embodiments, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some embodiments, F' is a small molecule fragment moiety illustrated in Fig. 3.

[0030] Disclosed herein, in certain embodiments, is a derivative of IMPDH2 protein having

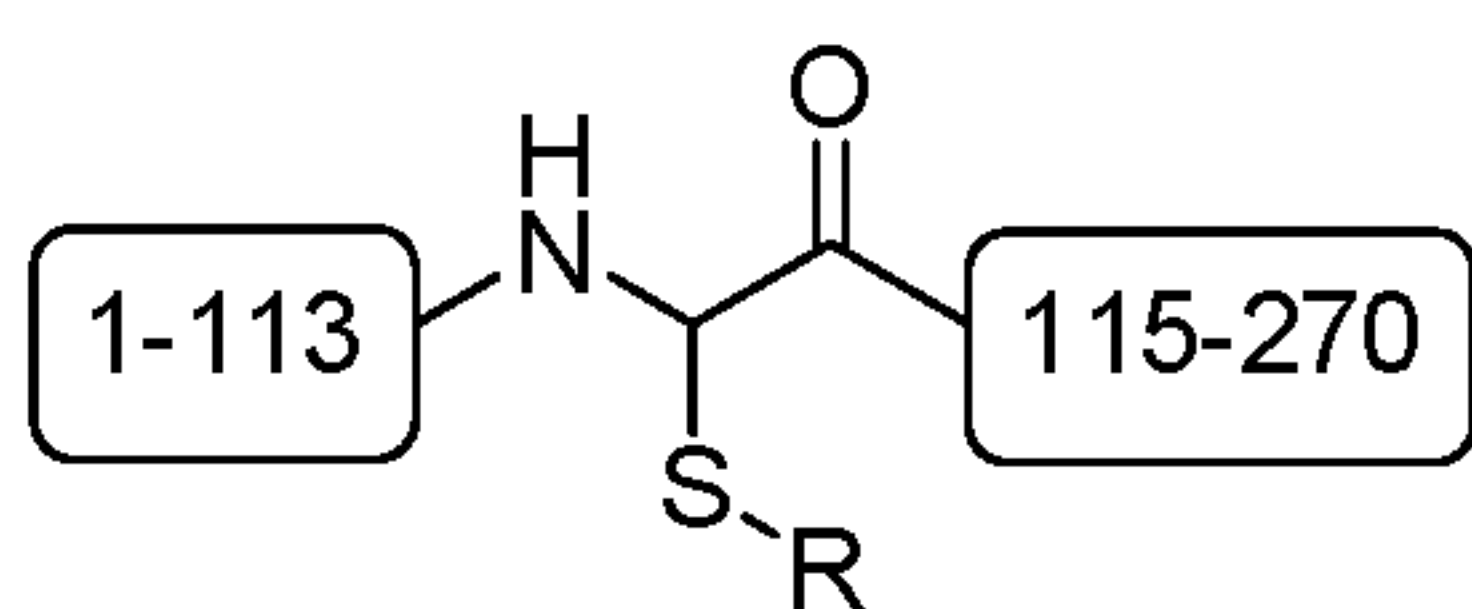


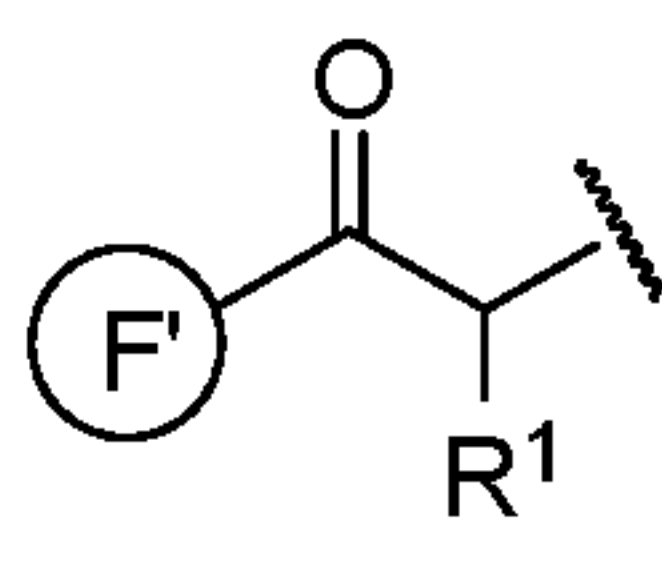
the structure of Formula (I), (I), wherein the derivation occurs at IMPDH2 cysteine residue position 331 based on SEQ ID NO: 7; R is selected from: (a)

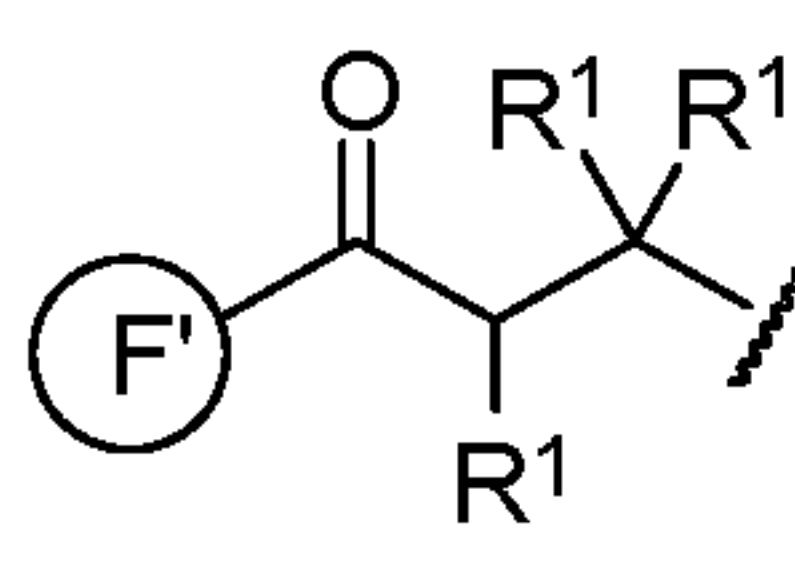
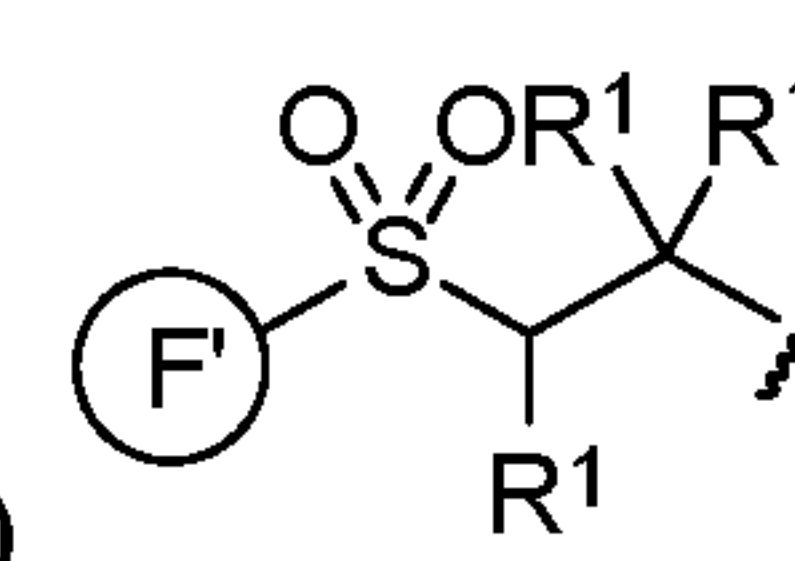
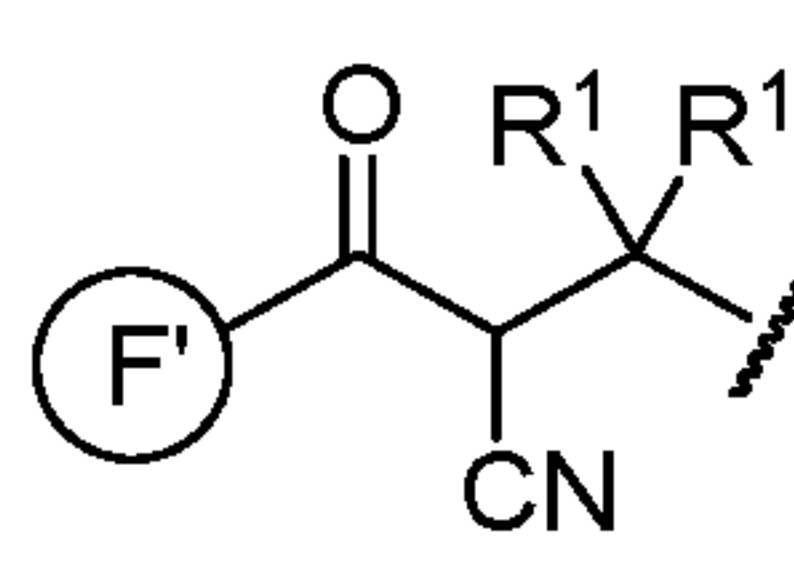
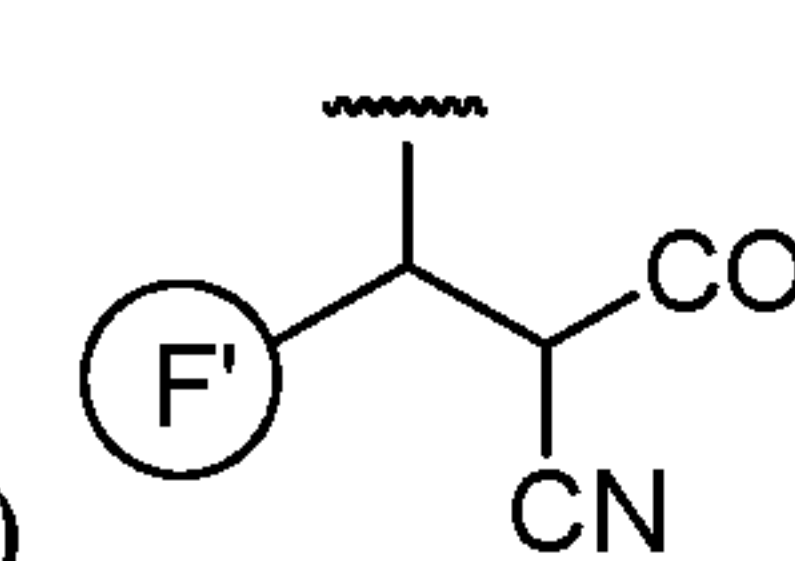


; wherein R^1 is H, C1-C3 alkyl, or aryl; and F' is a small molecule fragment moiety. In some embodiments, F' has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some embodiments, the molecular weight of F' is prior to enrichment with a halogen, a nonmetal, or a transition metal. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur atoms. In some embodiments, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some embodiments, F' is a small molecule fragment moiety illustrated in Fig. 3.

[0031] Disclosed herein, in certain embodiments, is a derivative of TIGAR protein having the

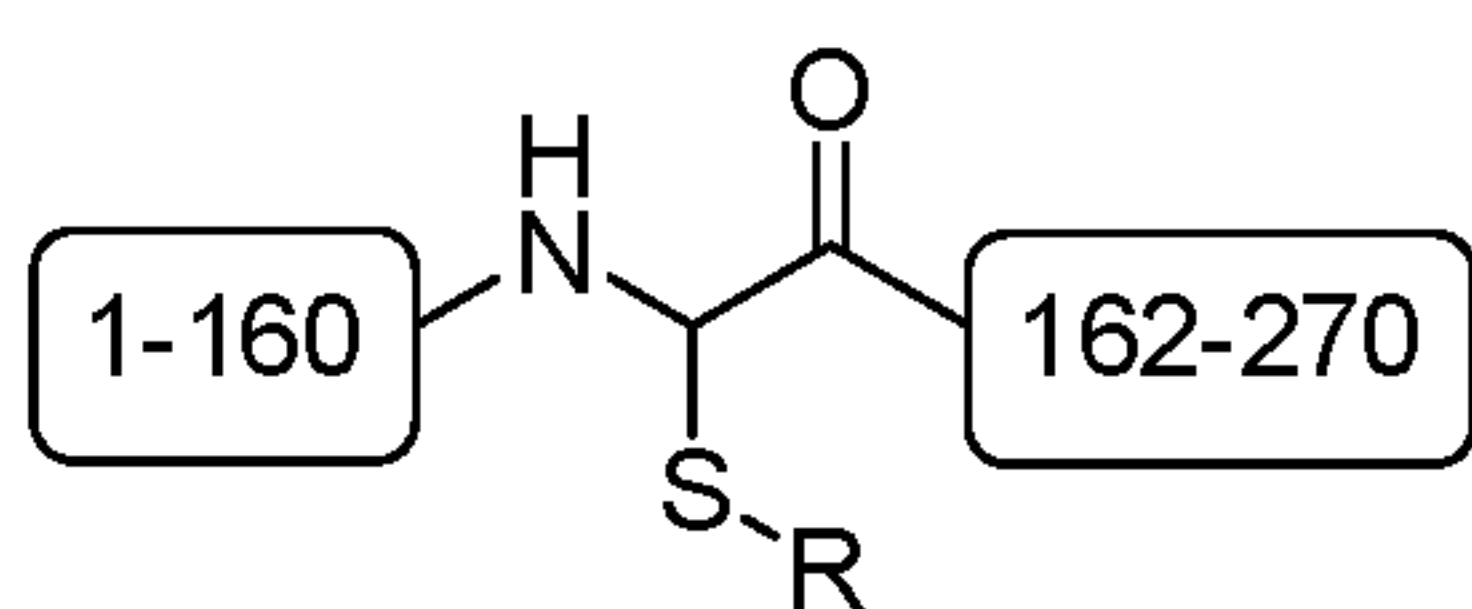
structure of Formula (I),  (I), wherein the derivation occurs at TIGAR

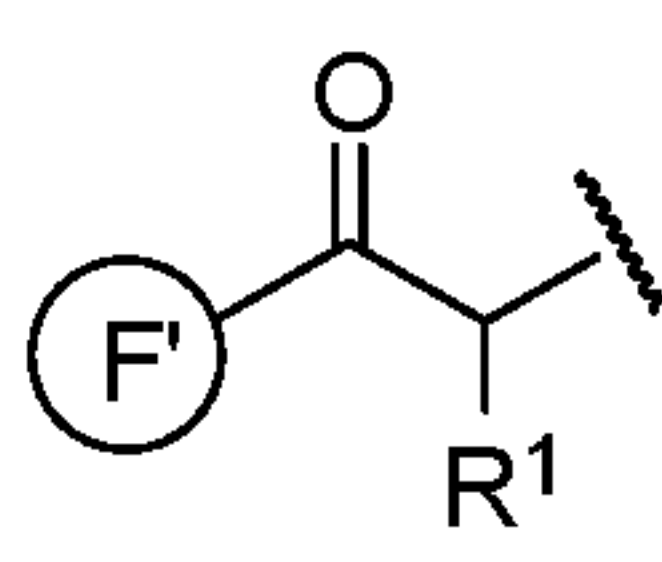
cysteine residue position 114 based on SEQ ID NO: 8; R is selected from: (a) ; (b)

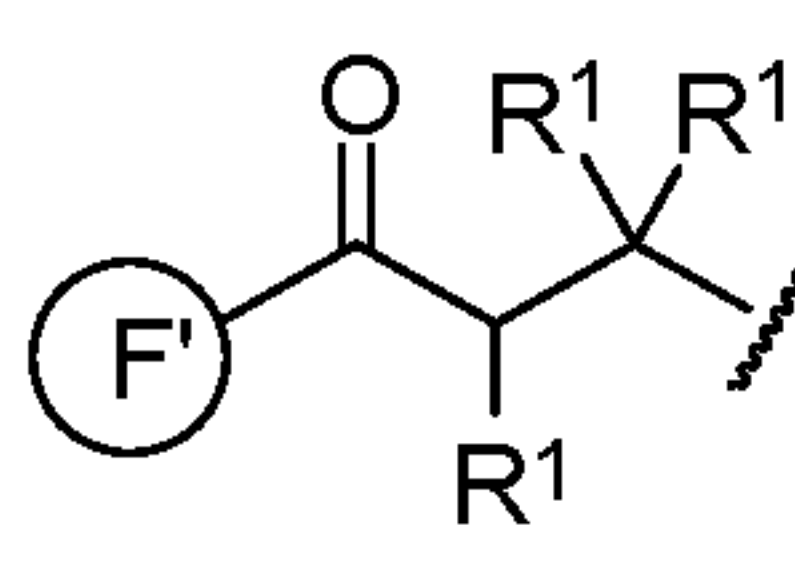
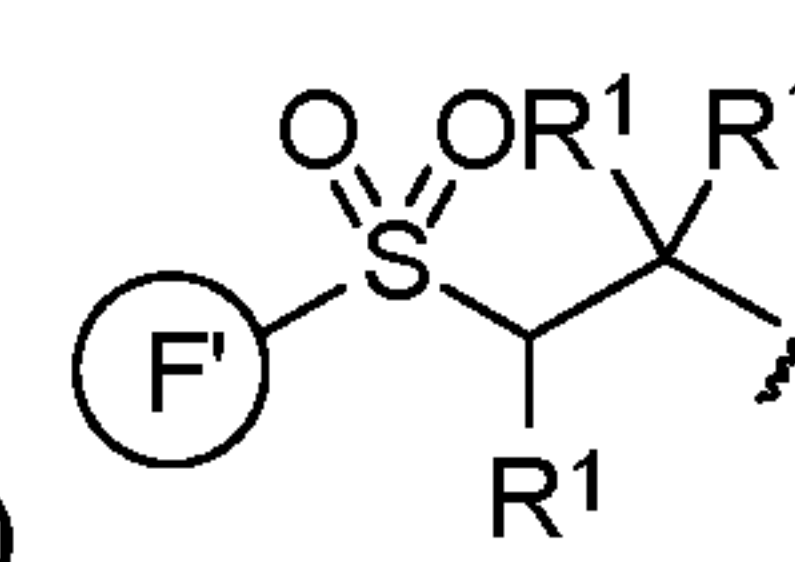
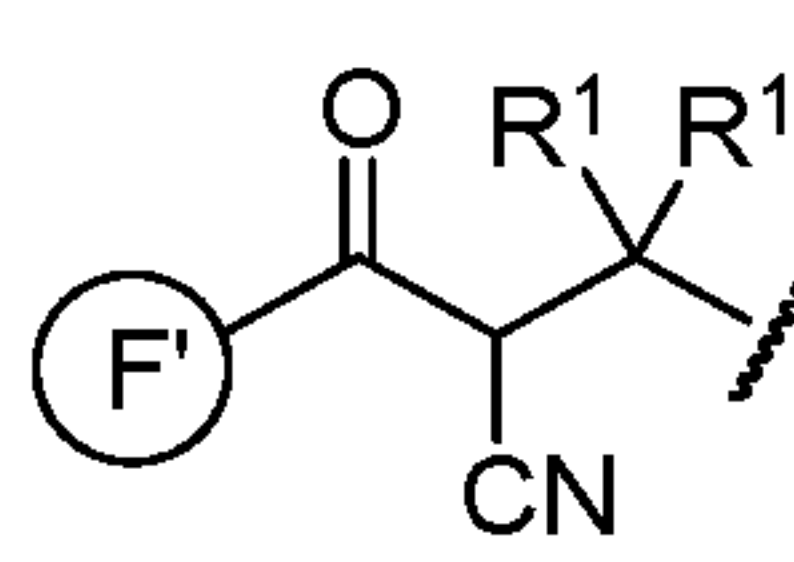
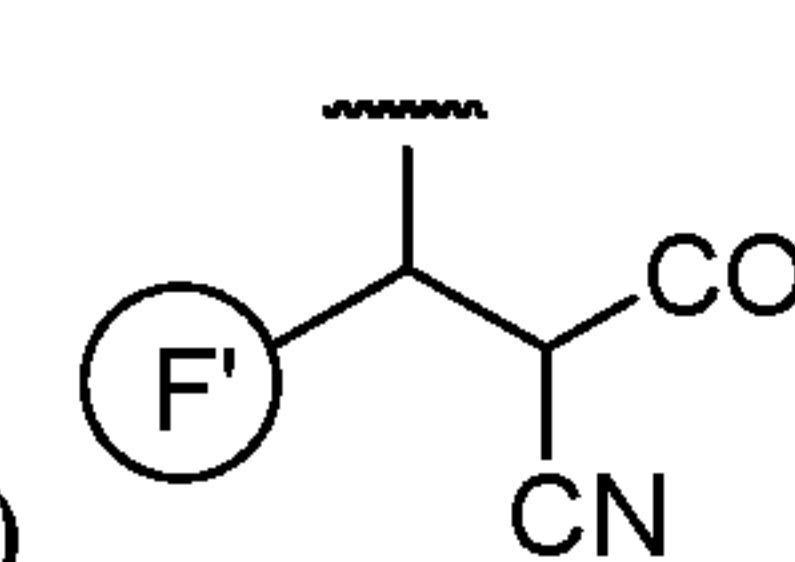
; (c) ; (d) ; or (e) ; wherein R¹

is H, C1-C3 alkyl, or aryl; and F' is a small molecule fragment moiety. In some embodiments, F' has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some embodiments, the molecular weight of F' is prior to enrichment with a halogen, a nonmetal, or a transition metal. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur atoms. In some embodiments, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some embodiments, F' is a small molecule fragment moiety illustrated in Fig. 3.

[0032] Disclosed herein, in certain embodiments, is a derivative of TIGAR protein having the

structure of Formula (I),  (I), wherein the derivation occurs at TIGAR

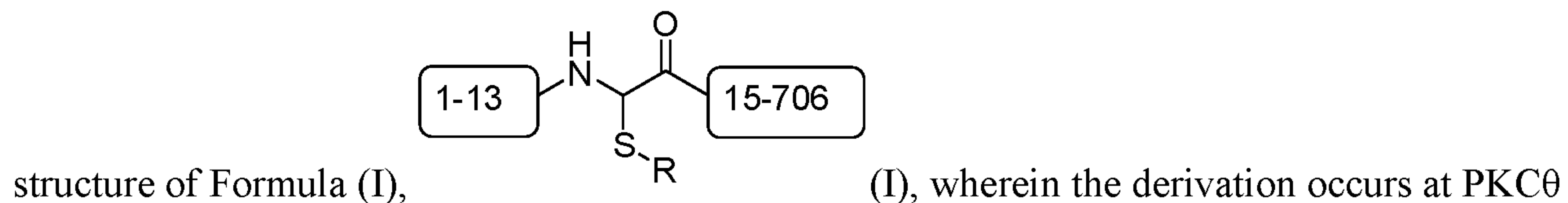
cysteine residue position 161 based on SEQ ID NO: 8; R is selected from: (a) ; (b)

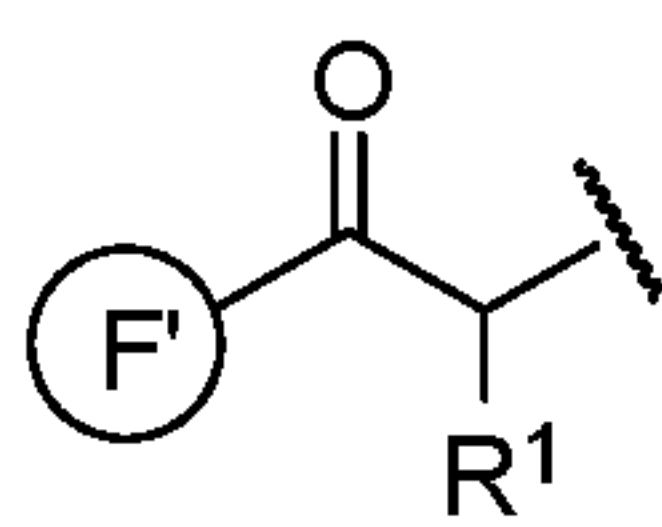
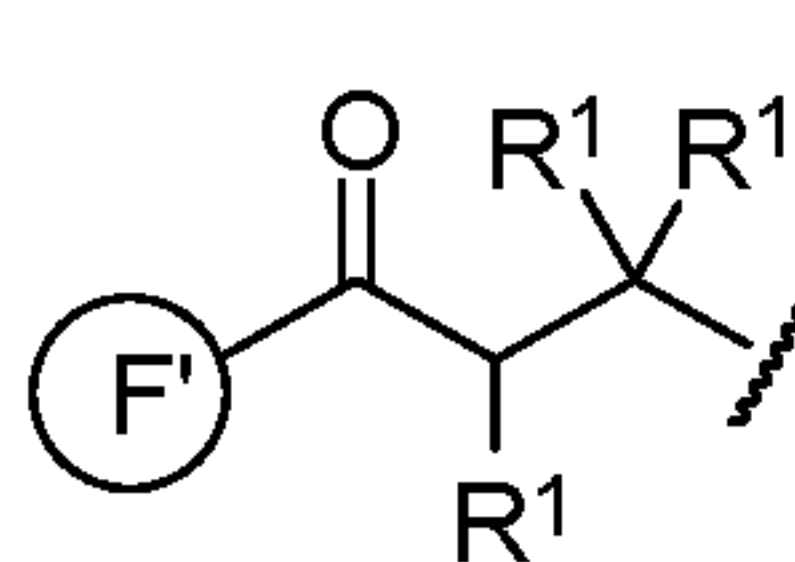
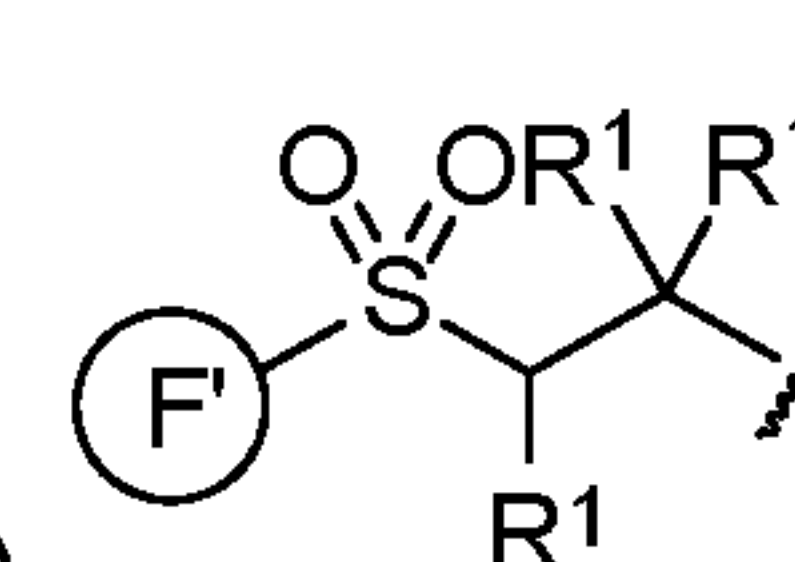
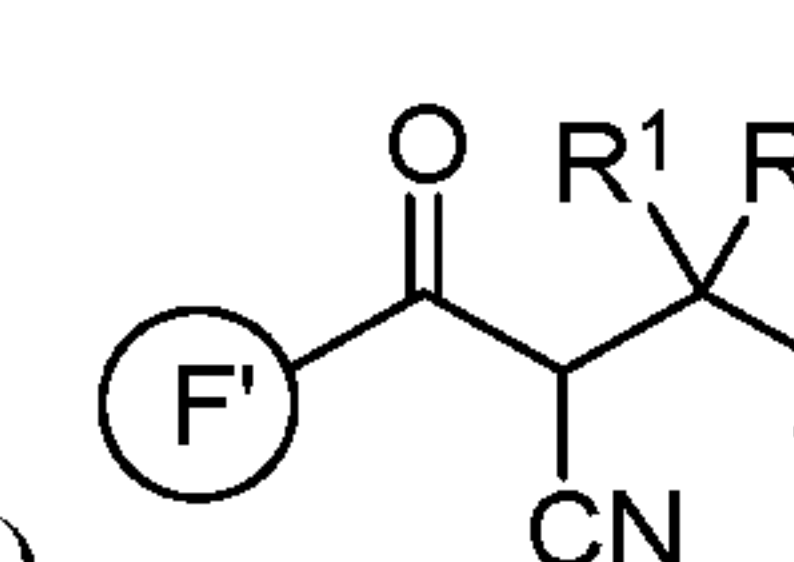
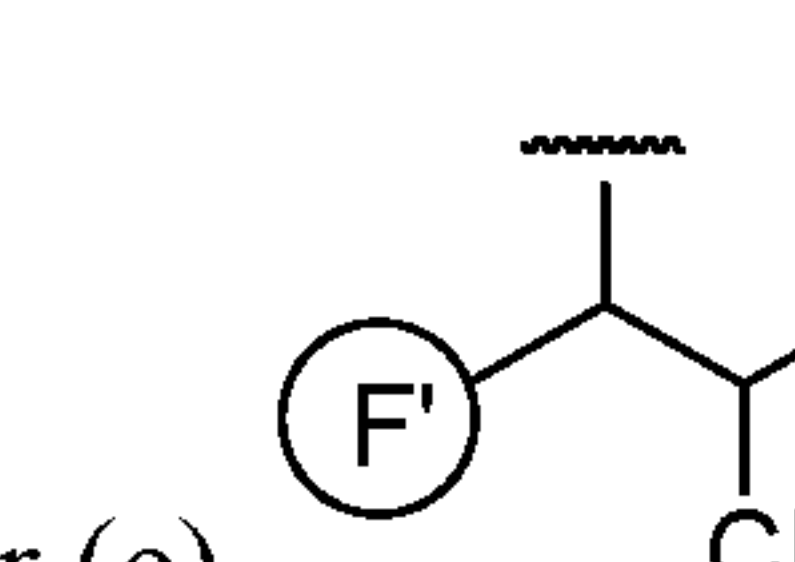
; (c) ; (d) ; or (e) ; wherein R¹

is H, C1-C3 alkyl, or aryl; and F' is a small molecule fragment moiety. In some embodiments, F' has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some embodiments, the molecular weight of F' is prior to enrichment with a halogen, a nonmetal, or a transition metal. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur

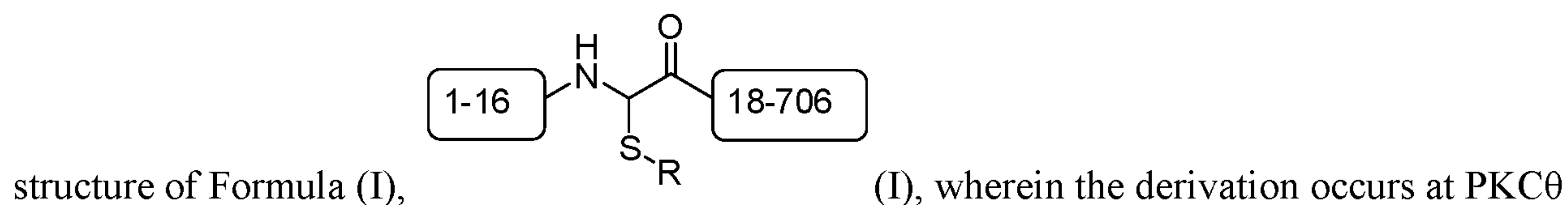
atoms. In some embodiments, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some embodiments, F' is a small molecule fragment moiety illustrated in Fig. 3.

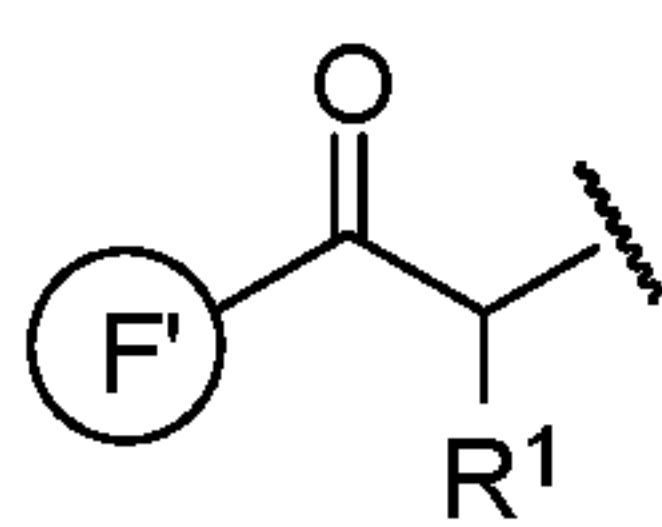
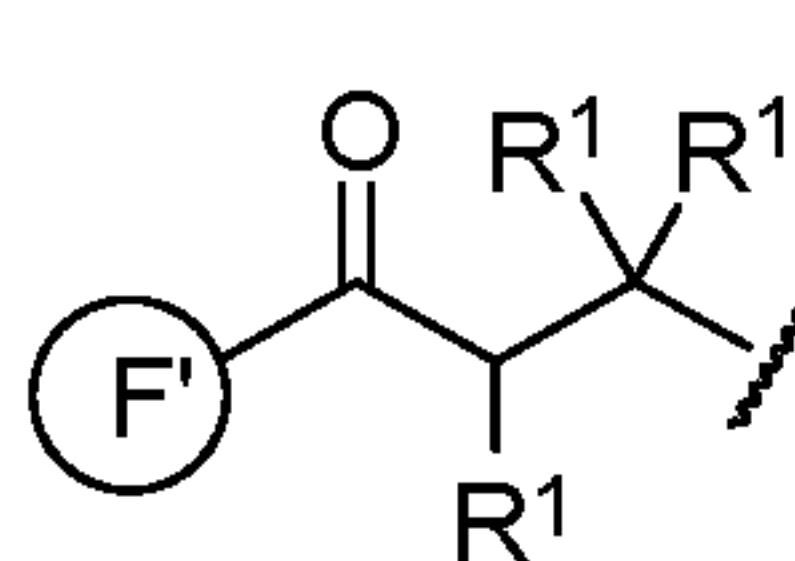
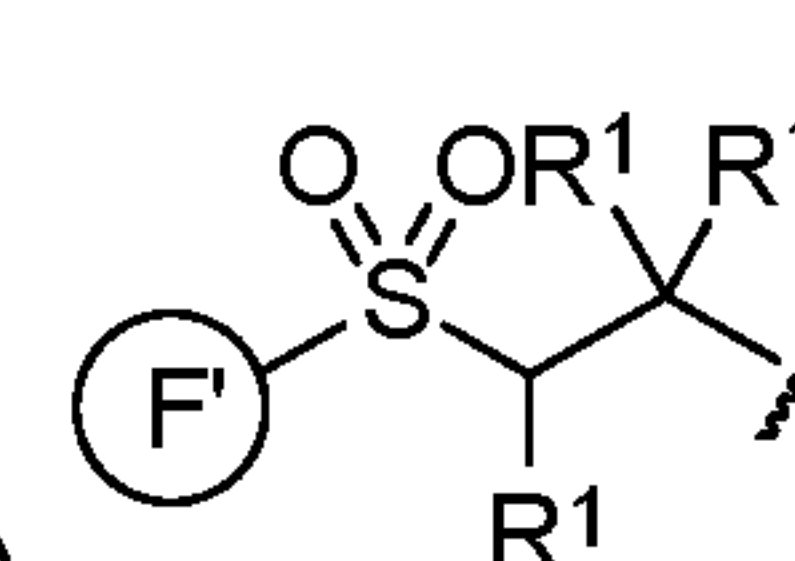
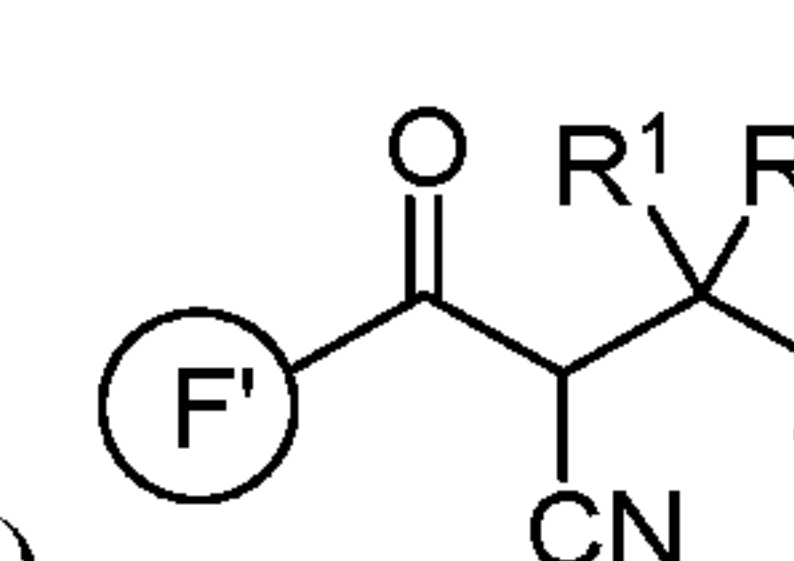
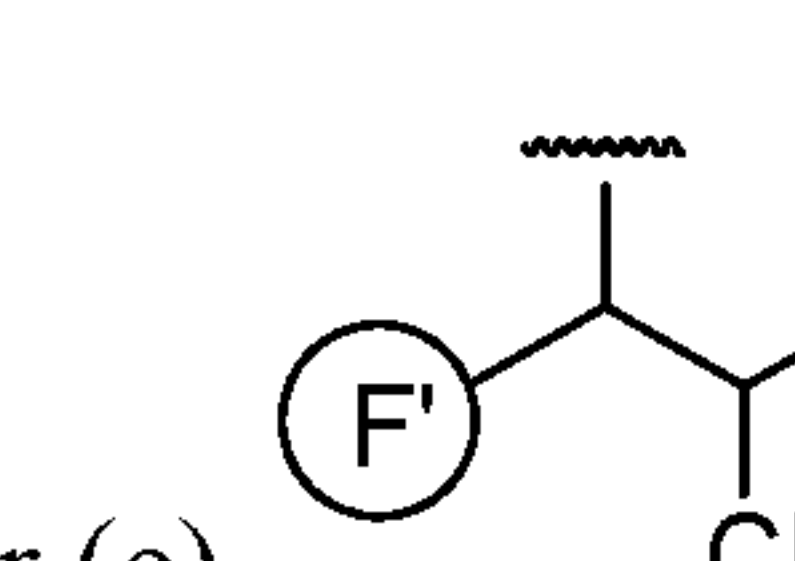
[0033] Disclosed herein, in certain embodiments, is a derivative of PKC θ protein having the



cysteine residue position 14 based on SEQ ID NO: 9; R is selected from: (a) ; (b) ; (c) ; (d) ; or (e) ; wherein R¹ is H, C1-C3 alkyl, or aryl; and F' is a small molecule fragment moiety. In some embodiments, F' has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some embodiments, the molecular weight of F' is prior to enrichment with a halogen, a nonmetal, or a transition metal. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur atoms. In some embodiments, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some embodiments, F' is a small molecule fragment moiety illustrated in Fig. 3.

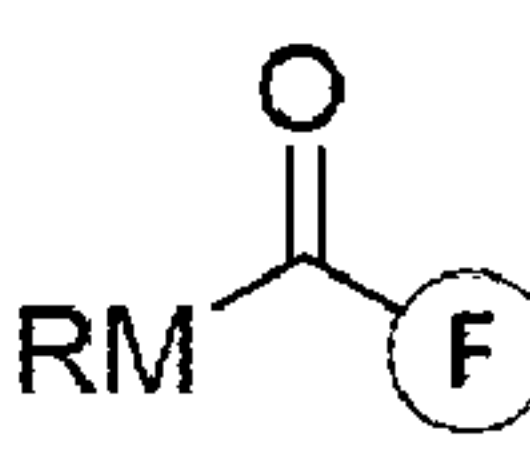
[0034] Disclosed herein, in certain embodiments, is a derivative of PKC θ protein having the

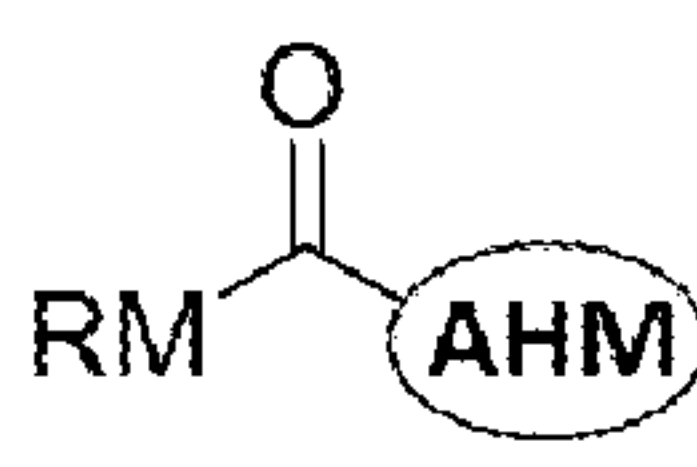


cysteine residue position 17 based on SEQ ID NO: 9; R is selected from: (a) ; (b) ; (c) ; (d) ; or (e) ; wherein R¹ is H, C1-C3 alkyl, or aryl; and F' is a small molecule fragment moiety. In some embodiments, F' has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some embodiments, the molecular

weight of F' is prior to enrichment with a halogen, a nonmetal, or a transition metal. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur atoms. In some embodiments, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some embodiments, F' is a small molecule fragment moiety illustrated in Fig. 3.

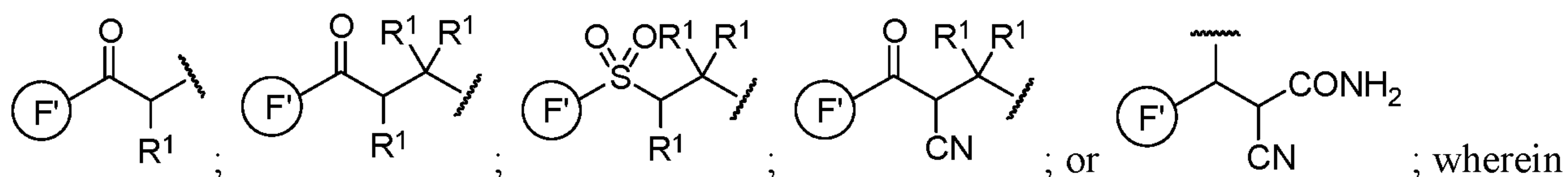
[0035] Disclosed herein, in certain embodiments, is a method of identifying a cysteine containing protein as a binding target for a small molecule fragment, comprising: (a) obtaining a set of cysteine-reactive probe-protein complexes from a sample comprising a first cell solution treated with a small molecule fragment and a cysteine reactive probe wherein the cysteine-reactive probe comprises a reactive moiety capable of forming a covalent bond with a cysteine residue located on the cysteine containing protein; (b) analyzing the set of cysteine-reactive probe-protein complexes by a proteomic analysis means; and (c) based on step b), identifying a cysteine containing protein as the binding target for the small molecule fragment. In some embodiments, the method further comprises determining a value of each of the cysteine containing protein from the set of cysteine-reactive probe-protein complexes for identifying a cysteine containing protein as the binding target for the small molecule fragment, wherein the value is determined based on the proteomic analysis means of step b). In some embodiments, the sample further comprises a second cell solution. In some embodiments, the method further comprises contacting the first cell solution with a small molecule fragment for an extended period of time prior to incubating the first cell solution with a first cysteine-reactive probe to generate a first group of cysteine-reactive probe-protein complexes. In some embodiments, the extended period of time is about 5, 10, 15, 20, 30, 60, 90, 120 minutes or longer. In some embodiments, the method further comprises contacting the second cell solution with a second cysteine-reactive probe to generate a second group of cysteine-reactive probe-protein complexes. In some embodiments, the first cysteine-reactive probe and the second cysteine-reactive probe are the same. In some embodiments, the first group and the second group of cysteine-reactive probe-protein complexes comprise the set of cysteine-reactive probe-protein complexes. In some embodiments, the cysteine containing protein is an enzyme, a transporter, a receptor, a channel protein, an adaptor protein, a chaperone, a signaling protein, a plasma protein, transcription related protein, translation related protein, mitochondrial protein, or cytoskeleton related protein. In some embodiments, the cysteine containing protein is a protein illustrated in Table 3. In some embodiments, the cysteine containing protein is a protein illustrated in Table 1, Table 2, Table 8, Table 9, Table 10A, Table 10B, Table 10C, Table 10D or Table 10E. In some embodiments, the small molecule fragment is a small molecule fragment

of Formula (I):  Formula (I), wherein: RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and F is a small molecule fragment moiety. In some embodiments, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some embodiments, F is obtained from a compound library. In some embodiments, the compound library comprises ChemBridge fragment library, Pyramid Platform Fragment-Based Drug Discovery, Maybridge fragment library, FRGx from AnalytiCon, TCI-Frag from AnCoreX, Bio Building Blocks from ASINEX, BioFocus 3D from Charles River, Fragments of Life (FOL) from Emerald Bio, Enamine Fragment Library, IOTA Diverse 1500, BIONET fragments library, Life Chemicals Fragments Collection, OTAVA fragment library, Prestwick fragment library, Selcia fragment library, TimTec fragment-based library, Allium from Vitas-M Laboratory, or Zenobia fragment library. In some embodiments, F is a small molecule fragment moiety illustrated in Fig. 3. In some embodiments, the cysteine-reactive probe is a cysteine-

reactive probe of Formula (II):  Formula (II), wherein: RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond to the thiol group of a cysteine residue; and AHM is an affinity handle moiety. In some embodiments, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some embodiments, the affinity handle moiety comprises an affinity handle and a binding moiety that facilitates covalent interaction of the cysteine-reactive probe to a cysteine residue of a cysteine-containing protein. In some embodiments, the binding moiety is a small molecule fragment obtained from a compound library. In some embodiments, the affinity handle comprises a carbodiimide, N-hydroxysuccinimide (NHS) ester, imidoester, pentafluorophenyl ester, hydroxymethyl phosphine, maleimide, haloacetyl, pyridyl disulfide, thiosulfonate, vinylsulfone, hydrazide, alkoxyamine, alkyne, azide, or isocyanate group. In some embodiments, the affinity handle is further conjugated to an affinity ligand. In some embodiments, the affinity ligand comprises a chromophore, a labeling group, or a combination thereof. In some embodiments, the chromophore comprises non-fluorochrome chromophore, quencher, an absorption chromophore, fluorophore, organic dye, inorganic dye, metal chelate, or a fluorescent enzyme substrate. In some embodiments, the labeling group is a biotin moiety, a streptavidin moiety, bead, resin, a solid support, or a combination thereof. In some embodiments, the cysteine-reactive probe is a cysteine-reactive probe illustrated in Fig. 3. In some embodiments, the proteomic analysis means comprises a mass spectroscopy method. In

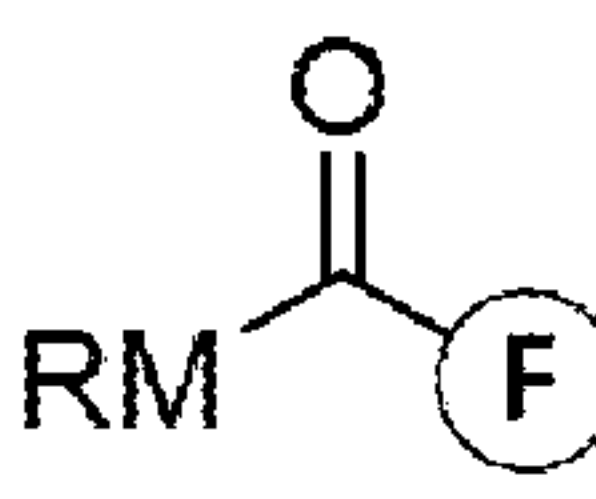
some embodiments, the identifying in step c) further comprises (i) locating a first value assigned to a cysteine containing protein from the first group of cysteine-reactive probe-protein complex and a second value of the same cysteine containing protein from the second group of cysteine-reactive probe-protein complex; and (ii) calculating a ratio between the two values assigned to the same cysteine containing protein. In some embodiments, the ratio of greater than 2 indicates that the cysteine containing protein is a candidate for interacting with the small molecule fragment. In some embodiments, the identifying in step c) further comprises calculating a percentage of inhibition of the cysteine-reactive probe to the cysteine containing protein. In some embodiments, the percentage of inhibition of greater than 50%, 60%, 70%, 80%, 90%, or at 100% indicates that the cysteine containing protein is a candidate for interacting with the small molecule fragment. In some embodiments, the method is an *in situ* method. In some embodiments, the cysteine-reactive probe is not 4-hydroxynonenal or 15-deoxy- $\Delta^{12,14}$ -prostaglandin J2.

[0036] Disclosed herein, modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, wherein the small molecule fragment has a molecular weight of about 150 Dalton or higher. In some embodiments, the cysteine containing protein comprises a cysteine residue site denoted in Table 3. In some embodiments, the cysteine containing protein comprises a protein sequence illustrated in Table 1, Table 2, Table 8, Table 9, Table 10A, Table 10B, Table 10C, Table 10D or Table 10E. In some embodiments, the cysteine containing protein is about 20, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000 amino acid residues in length or more. In some embodiments, the cysteine residue of the modified cysteine containing protein has the structure SR, wherein R is selected from:



R^1 is H, C1-C3 alkyl, or aryl; and F' is the small molecule fragment moiety. In some embodiments, the small molecule fragment has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur atoms. In some embodiments, the modified cysteine containing protein is selected from IDH2, caspase-8, caspase-10 or PRMT1. In some embodiments, IDH2 is modified at cysteine position 308. In some embodiments, caspase-8 is modified at cysteine position 360. In some embodiments, caspase-10 exist in the proform and is modified at cysteine position 401. In some

embodiments, PRMT1 is modified at cysteine position 109. In some embodiments, the small

molecule fragment is a small molecule fragment of Formula (I):  Formula (I), wherein: RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and F is a small molecule fragment moiety. In some embodiments, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some embodiments, F is obtained from a compound library. In some embodiments, F is a small molecule fragment moiety illustrated in Fig. 3. In some embodiments, F further comprises a linker moiety that connects F to the carbonyl moiety. In some embodiments, the small molecule fragment binds irreversibly to the cysteine containing protein. In some embodiments, the small molecule fragment binds reversibly to the cysteine containing protein.

[0037] Disclosed herein, in certain embodiments, is a method of screening a small molecule fragment for interaction with a cysteine containing protein, comprising: (a) harvesting a set of cysteine-reactive probe-protein complexes from a sample comprising a first cell solution treated with a small molecule fragment and a cysteine reactive probe wherein the cysteine-reactive probe comprises a reactive moiety capable of forming a covalent bond with a cysteine residue located on the cysteine containing protein; (b) analyzing the set of cysteine-reactive probe-protein complexes by a proteomic analysis means; and (c) based on step b), identifying the small molecule fragment as interacting with the cysteine containing protein. In some embodiments, the method further comprises determining a value of each of the cysteine containing protein from the set of cysteine-reactive probe-protein complexes prior to identifying the small molecule fragment as interacting with the cysteine containing protein, wherein the value is determined based on the proteomic analysis means of step b). In some embodiments, the cysteine containing protein is a protein illustrated in Table 3. In some embodiments, the cysteine containing protein is a protein illustrated in Table 1, Table 2, Table 8, Table 9, Table 10A, Table 10B, Table 10C, Table 10D or Table 10E.

BRIEF DESCRIPTION OF THE DRAWINGS

[0038] Various aspects of the invention are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present invention will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the invention are utilized, and the accompanying drawings of which:

[0039] **Fig. 1** illustrates proteome-wide screening of covalent fragments. **A**, General protocol for competitive isoTOP-ABPP. Cell lysate or intact cells are pre-treated with a fragment

electrophile or DMSO and then reacted with an IA-alkyne probe **1**. The fragment- and DMSO-treated samples are then conjugated to isotopically-differentiated TEV protease-cleavable biotin tags [light (red) and heavy (blue), respectively] by copper-mediated azide-alkyne cycloaddition (CuAAC or click) chemistry, mixed, and IA-labeled proteins enriched by streptavidin-conjugated beads and digested stepwise on-bead with trypsin and TEV to yield IA-labeled peptides for MS analysis. Competition ratios, or *R* values, are measured by dividing the MS1 ion peaks for IA-labeled peptides in DMSO-treated (heavy or blue) versus fragment-treated (light or red) samples. **B**, Representative members of the electrophilic fragment library, where the reactive (electrophilic) and binding groups are colored green and black, respectively. **C**, Initial analysis of the proteomic reactivity of fragments using an IA-rhodamine probe **16**. Soluble proteome from Ramos cells was treated with the indicated fragments (500 μ M each) for 1 h, followed by labeling with IA-rhodamine (1 μ M, 1 h) and analysis by SDS-PAGE and in-gel fluorescence scanning. Several proteins were identified that show impaired reactivity with IA-rhodamine in the presence of one or more fragments (asterisks). Fluorescent gel shown in grayscale. **D**, Competitive isoTOP-ABPP analysis of fragment-cysteines interactions in the soluble proteome of MDA-MB-231 cells pre-treated with the following fragments (500 μ M each): 3,5-di(trifluoromethyl)aniline chloroacetamide **3**, acrylamide **14**, and acetamide **17**. Proteomic reactivity values, or liganded cysteine rates, for fragments were calculated as the percentage of total cysteines with *R* values ≥ 4 in DMSO/fragment (heavy/light) comparisons. **E**, Concentration-dependent labeling of MDA-MB-231 soluble proteomes with acrylamide **18** and chloroacetamide **19** click probes detected by CuACC with a rhodamine-azide tag and analysis by SDS-PAGE and in-gel fluorescence scanning. **F**, Representative MS1 peptide ion chromatograms from competitive isoTOP-ABPP experiments performed with fragments **3**, **4**, and **23** marking liganded cysteines selectively targeted by one of three fragments (or, in the case of PHGDH C369, by all three fragments).

[0040] Fig. 2 illustrates a conceptual schematic of an exemplary computer server to be used for processing a method described herein.

[0041] Fig. 3 shows composition of fragment electrophile library and structures of additional tool compounds, click probes, and fragments.

[0042] Fig. 4 illustrates analysis of proteomic reactivities of fragment electrophiles determined by competitive isoTOP-ABPP in human cell lysates. **A**, Frequency of quantification of all cysteines across the complete set of competitive isoTOP-ABPP experiments performed with fragment electrophiles. Note that cysteines were required to have been quantified in at least three isoTOP-ABPP data sets for interpretation. **B**, Rank order of proteomic reactivity values (or liganded cysteine rates) of fragments calculated as the percentage of all quantified cysteines with

R values ≥ 4 for each fragment. The majority of fragments were evaluated in 2-4 replicate experiments in MDA-MB-231 and/or Ramos cell lysates, and their proteomic reactivity values are reported as mean $\pm$ SEM values for the replicates. **C**, Comparison of the proteomic reactivities of representative fragments screened at 500 versus 25 μ M in cell lysates. **D**, Comparison of proteomic reactivity values for fragments tested in both Ramos and MDA-MB-231 lysates. **E**, Mean $\pm$ SEM data for proteomic reactivity values of representative fragments tested in at least three independent replicates. **F**, Relative GSH reactivity for representative fragment electrophiles. Consumption of GSH (125 μ M) was measured using Ellman's reagent (5 mM) after 1 h incubation with the indicated fragments (500 μ M). **G**, Proteomic reactivity values for fragments electrophiles (500 μ M) possessing different electrophilic groups attached to a common binding element.

[0043] **Fig. 5** illustrates analysis of cysteines and proteins liganded by fragment electrophiles. **A**, Fraction of total quantified cysteines and proteins that were liganded by fragment electrophiles in competitive isoTOP-ABPP experiments. **B**, Fraction of liganded proteins found in DrugBank. **C**, Functional classes of DrugBank and non-DrugBank proteins containing liganded cysteines. **D**, Functional categorization of liganded and unliganded cysteines based on residue annotations from the Uniprot database. **E**, Comparison of the ligandability of cysteines as a function of their intrinsic reactivity with the IA-alkyne probe. Cysteine reactivity values were taken from Weerapana, et al. *Nature* **468**, 790-795 (2010), where lower ratios correspond to higher cysteine reactivity. Individual cysteines are plotted on the x-axis and were sorted by reactivity, which is shown on the left y-axis. A moving average with a step-size of 50 is shown in blue for the percentage of liganded cysteines within each reactivity bin (percent values shown on the right y-axis). **F**, Number of liganded and quantified cysteines per protein measured by isoTOP-ABPP. Respective average values of one and three for liganded and quantified cysteines per protein were measured by isoTOP-ABPP. **G**, R values for six cysteines in XPO1 quantified by isoTOP-ABPP, identifying C528 as the most liganded cysteine in this protein. Each point represents a distinct fragment-cysteine interaction quantified by isoTOP-ABPP.

[0044] **Fig. 6** illustrates analysis of fragment-cysteine interactions. **A**, Heatmap showing R values for representative cysteines and fragments organized by proteomic reactivity values (high to low, left to right) and percentage of fragment hits for individual cysteines (high to low, top to bottom). R values ≥ 4 designate fragment hits (colored medium and dark blue). White color designates fragment-cysteine interactions that were not detected (ND). **B**, **C**, Histograms depicting the percentage of fragments that are hits ($R \geq 4$) for all 768 liganded cysteines (**B**) or for liganded cysteines found in enzymes for which X-ray and/or NMR structures have been reported (or reported for a close homologue of the enzyme) (**C**). **D**, Percentage of liganded

cysteines targeted only by group A (red) or B (blue) fragments or both group A and B fragments (black). Shown for all liganded cysteines, liganded cysteines in enzyme active and non-active sites, and liganded cysteines in transcription factors/ regulators. For **C**, **D**, active-site cysteines were defined as those that reside $< 10 \text{ \AA}$ from established active-site residues and/or bound substrates/inhibitors in enzyme structures. **E**, Representative example of reactive docking predictions shown for XPO1 (PDB ID: 3GB8). All accessible cysteines were identified and reactive docking was conducted with all fragments from the library within a $25 \text{ \AA}$ docking cube centered on each accessible cysteine. Categories of XPO1 cysteines based on combined docking and isoTOP-ABPP results are shown. **F**, Success rate of reactive docking predictions for liganded cysteines identified by isoTOP-ABPP in 29 representative proteins.

[0045] **Fig. 7** illustrates analysis of cysteines liganded by fragment electrophiles in competitive isoTOP-ABPP experiments. **A**, Representative MS1 ion chromatograms for peptides containing C481 of BTK and C131 of MAP2K7, two cysteines known to be targeted by the anti-cancer drug ibrutinib. Ramos cells were treated with ibrutinib ($1 \mu\text{M}$, 1 h, red trace) or DMSO (blue trace) and evaluated by isoTOP-ABPP. **C**, Total number of liganded cysteines found in the active sites and non-active sites of enzymes for which X-ray and/or NMR structures have been reported (or reported for a close homologue of the enzyme). **C**, *R* values for eight cysteines in PHGDH quantified by isoTOP-ABPP, identifying a single liganded cysteine C369 that is targeted by several fragment electrophiles. Each point represents a distinct fragment-cysteine interaction quantified by isoTOP-ABPP. **D**, Heatmap showing representative fragment interactions for liganded cysteines found in the active sites and non-active sites of kinases. **E**, Histogram showing the fragment hit rate for active- and non-active site cysteines in kinases. **F**, The percentage of liganded cysteines in kinases that were targeted by only group A, only group B, or both group A and B compounds. **G**, Heatmap showing representative fragment interactions for liganded cysteines found in transcription factors/regulators. **H**, The fraction of cysteines predicted to be ligandable or not ligandable by reactive docking that were quantified in isoTOP-ABPP experiments.

[0046] **Fig. 8** illustrates confirmation and functional analysis of fragment-cysteine interactions. **A**, Representative MS1 chromatograms for the indicated Cys-containing peptides from PRMT1 quantified in competitive isoTOP-ABPP experiments of MDA-MB-231 cell lysates, showing blockade of IA-alkyne **1** labeling of C109 by fragment **11**, but not control fragment **3**. **B**, **11**, but not **3** blocked IA-rhodamine ($2 \mu\text{M}$) labeling of recombinant, purified WT-PRMT1 ($1 \mu\text{M}$ protein doped into HEK293T cell lysates). Note that a C109S-PRMT1 mutant did not react with IA-rhodamine. **C**, IC_{50} curve for blockade of **16** labeling of PRMT1 by **11**. CI, 95% confidence intervals. **D**, Effect of **11** and control fragment **3** on methylation of

recombinant histone 4 by recombinant PRMT1. Shown is one representative experiment of three independent experiments that yielded similar results. **E**, **60**, but not control fragment **3** (50 μ M of each fragment) blocked labeling of recombinant MLTK (or ZAK) kinase by a previously reported ibrutinib-derived activity probe **59** (upper panel). A C22A-MLTK mutant did not react with the ibrutinib probe. Anti-FLAG blotting confirmed similar expression of WT- and C22A-MLTK proteins, which were expressed as FLAG-fusion proteins in HEK293T cells (lower panel). **F**, IC₅₀ curve for blockade of ibrutinib probe-labeling of MLTK by **60**. **G**, **60**, but not control fragment **3** (100 μ M of each fragment) inhibited the kinase activity of WT-, but not C22A-MLTK. **H**, Click probe **18** (25 μ M) labeled WT-IMPDH2 and C331S-IMPDH2, but not C140S-IMPDH2 (or C140S/C331S-IMPDH2). Labeling was detected by CuAAC conjugation to a rhodamine-azide reporter tag and analysis by SDS-PAGE and in-gel fluorescence scanning. Recombinant IMPDH2 WT and mutants were expressed and purified from *E. coli* and added to Jurkat lysates to a final concentration of 1 μ M protein. **I**, Nucleotide competition profile for **18**-labeling of recombinant WT-IMPDH2 (500 μ M of each nucleotide). **J**, IC₅₀ curve for blockade of **18** labeling of IMPDH2 by ATP. **K**, **5**, but not control fragment **3** blocked IA-rhodamine (2 μ M) labeling of recombinant, purified C161S-TIGAR (2 μ M protein doped into Ramos cell lysates). **L**, IC₅₀ curve for blockade of IA-rhodamine labeling of C161S-TIGAR by **5**. **M**, **5**, but not control fragment **3** (100 μ M of each fragment) inhibited the catalytic activity of WT-TIGAR, C161S-TIGAR, but not C114S-TIGAR or C114S/C161S-TIGAR. For panels **C**, **F**, **G**, **I**, **J**, **L**, and **M**, data represent mean values $\pm$ SEM for at least three independent experiments. Statistical significance was calculated with unpaired students t-tests comparing DMSO- to fragment-treated samples; **, $p < 0.01$, ***, $p < 0.0001$.

[0047] **Fig. 9** illustrates confirmation and functional analysis of fragment-cysteine interactions. **A**, Representative MS1 ion chromatograms for the MLTK tryptic peptide containing liganded cysteine C22 quantified by isoTOP-ABPP in MDA-MB-231 lysates treated with fragment **4** or control fragment **3** (500 μ M each). **B**, Lysates from HEK293T cells expressing WT- or C22A-MLTK treated with the indicated fragments and then an ibrutinib-derived activity probe **59** at 10 μ M. MLTK labeling by **59** was detected by CuAAC conjugation to a rhodamine-azide tag and analysis by SDS-PAGE and in-gel fluorescence scanning. **C**, Representative MS1 ion chromatograms for IMPDH2 tryptic peptides containing the catalytic cysteine, C331, and Bateman domain cysteine, C140, quantified by isoTOP-ABPP in cell lysates treated with the indicated fragments (500 μ M each). **D**, Structure of human IMPDH2 (PDB ID: 1NF7) (light grey) and its structurally unresolved Bateman domain modeled by ITASSER (dark grey) showing the positions of C331 (red spheres), Ribavirin Monophosphate and C2-Mycophenolic Adenine Dinucleotide (blue), and C140 (yellow spheres). **E**, Fragment reactivity

with recombinant, purified IMPDH2 added to Jurkat lysates to a final concentration of 1 μ M protein, where reactivity was detected in competition assays using the click probe **18** (25 μ M; see **Fig. 8H** for structure of **18**). Note that **18** reacted with WT- and C331S-IMPDH2, but not C140S or C140S/C331S-IMPDH2. **F**, Nucleotide competition of **18** (25 μ M) labeling of WT-IMPDH2 added to cell lysates to a final concentration of 1 μ M protein. **G**, Representative MS1 chromatograms for TIGAR tryptic peptides containing C114 and C161 quantified by isoTOP-ABPP in cell lysates treated with the indicated fragments (500 μ M each). **H**, Crystal structure of TIGAR (PDB ID: 3DCY) showing C114 (red spheres), C161 (yellow spheres), and inorganic phosphate (blue). **I**, Labeling of recombinant, purified TIGAR and mutant proteins by the IA-rhodamine (2 μ M) probe. TIGAR proteins were added to cell lysates, to a final concentration of 2 μ M protein. **J**, Concentration-dependent inhibition of WT-TIGAR by **5**. Note that the C140S-TIGAR mutant was not inhibited by **5**. Data represent mean values $\pm$ SEM for 4 replicate experiments at each concentration.

[0048] **Fig. 10** illustrates *in situ* activity of fragment electrophiles. **A**, X-ray crystal structure of IDH1 (PDB ID: 3MAS) showing the position of C269 and the frequently mutated residue in cancer, R132. **B**, **C**, Reactivity of **20** and control fragment **2** with recombinant, purified WT-IDH1 (**B**) or R132H-IDH1 (**C**) added to cell lysates to a final concentration of 2 or 4 μ M protein, respectively. Fragment reactivity was detected in competition assays using the IA-rhodamine probe (2 μ M); note that the C269S-IDH1 mutant did not react with IA-rhodamine. **D**, Representative MS1 ion chromatograms for the IDH1 tryptic peptides containing liganded cysteine C269 and an unliganded cysteine C379 quantified by isoTOP-ABPP in MDA-MB-231 lysates treated with fragment **20** (25 μ M). **E**, Western blot of MUM2C cells stably overexpressing GFP (mock) or R132H-IDH1 proteins. **F**, Representative MS1 chromatograms for the IDH1 tryptic peptides containing liganded cysteine C269 and an unliganded cysteine C379 quantified by isoTOP-ABPP in R132H-IDH-expressing MUM2C lysates treated with **20** or control fragment **2** (50 μ M, 2 h, *in situ*).

[0049] **Fig. 11** illustrates *in situ* activity of fragment electrophiles. **A**, Blockade of **16** labeling of WT-IDH1 by representative fragment electrophiles. Recombinant, purified WT-IDH1 was added to MDA-MB-231 lysates at a final concentration of 2 μ M, treated with fragments at the indicated concentrations, followed by IA-rhodamine probe **16** (2 μ M) and analysis by SDS-PAGE and in-gel fluorescence scanning. Note that a C269S mutant of IDH1 did not label with IA-rhodamine **16**. **B**, IC₅₀ curve for blockade of IA-rhodamine-labeling of IDH1 by **20**. Note that the control fragment **2** showed much lower activity. **C**, **20**, but not **2**, inhibited IDH1-catalyzed oxidation of isocitrate to α -ketoglutarate (α -KG) as measured by an increase in NADPH production (340 nm absorbance). **20** did not inhibit the C269S-IDH1 mutant. **D**, **20**

inhibited oncometabolite 2-hydroxyglutarate (2-HG) production by R132H-IDH1. MUM2C cells stably overexpressing the oncogenic R132H-IDH1 mutant or control GFP-expressing MUM2C cells were treated with the indicated fragments (2 h, *in situ*). Cells were harvested, lysed and IDH1-dependent production of 2-HG from α -KG and NADPH was measured by LC-MS and from which 2-HG production of GFP-expressing MUM2C cells was subtracted (GFP-expressing MUM2C cells produced < 10% of the 2-HG generated by R132H-IDH1-expressing MUM2C cells). **E**, Proteomic reactivity values for individual fragments are comparable *in vitro* and *in situ*. One fragment (**11**) marked in red showed notably lower reactivity *in situ* versus *in vitro*. Reactivity values were calculated as in **Fig. 1D**. Dashed line mark 90% prediction intervals for the comparison of *in vitro* and *in situ* proteomic reactivity values for fragment electrophiles. Blue and red circles mark fragments that fall above (or just at) or below these prediction intervals, respectively. **F**, Fraction of cysteines liganded *in vitro* that is also liganded *in situ*. Shown are liganded cysteine numbers for individual fragments determined *in vitro* and the fraction of these cysteines that were liganded by the corresponding fragments *in situ*. **G**, Representative cysteines that were selectively targeted by fragments *in situ*, but not *in vitro*. For *in situ*-restricted fragment-cysteine interactions, a second cysteine in the parent protein was detected with an unchanging ratio ($R \sim 1$), thus controlling for potential fragment-induced changes in protein expression. For panels **B-D**, data represent mean values $\pm$ SEM for at least three independent experiments. Statistical significance was calculated with unpaired students t-tests comparing DMSO- to fragment-treated samples; •••, $p < 0.0001$.

[0050] Fig. 12 illustrates fragment electrophiles that target pro-CASP8. **A**, Representative MS1 chromatograms for CASP8 tryptic peptide containing the catalytic cysteine C360 quantified by isoTOP-ABPP in cell lysates or cells treated with fragment **4** (250 μ M, *in vitro*; 100 μ M, *in situ*) and control fragment **21** (500 μ M, *in vitro*; 200 μ M, *in situ*). **B**, Fragment reactivity with recombinant, purified active CASP8 added to cell lysates, where reactivity was detected in competition assays using the caspase activity probe Rho-DEVD-AOMK probe ("DEVD" disclosed as SEQ ID NO: 857) (2 μ M, 1 h). **C**, Western blot of proteomes from MDA-MB-231, Jurkat, and CASP8-null Jurkat proteomes showing that CASP8 was only found in the pro-enzyme form in these cells. **D**, Fragment reactivity with recombinant, purified pro-CASP8 (D374A, D384A, C409S) added to cell lysates to a final concentration of 1 μ M protein, where reactivity was detected in competition assays with the IA-rhodamine probe (2 μ M). Note that mutation of both cysteine-360 and cysteine-409 to serine prevented labeling of pro-CASP8 by IA-rhodamine. **E**, Concentration-dependent reactivity of click probe **61**, with recombinant, purified pro-CASP8 (D374A, D384A) added to cell lysates to a final concentration of 1 μ M protein. Note that pre-treatment with **7** blocked **61** reactivity with pro-CASP8 and mutation of

C360 to serine prevented labeling of pro-CASP8 by **61** (25 μ M). **F**, Fragments **7** and **62** did not block labeling by Rho-DEVD-AOMK ("DEVD" disclosed as SEQ ID NO: 857) (2 μ M) of recombinant, purified active-CASP8 and active-CASP3 added to MDA-MB-231 cell lysates to a final concentration of 1 μ M protein. **G**, Representative MS1 chromatograms for tryptic peptides containing the catalytic cysteines of CASP8 (C360), CASP2 (C320), and CASP7 (C186) quantified by isoTOP-ABPP in Jurkat cell lysates treated with **7** or **62** (50 μ M, 1 h). **H**, Representative MS1 chromatograms for CASP8 tryptic peptide containing C360 quantified by isoTOP-ABPP in cell lysates treated with 10 versus 100 μ M of **61**. Structure of CASP8 C360 tryptic peptide adduct (blue) modified by **61** (black) and conjugated to TEV cleavable tag (red), where underline indicates site of isotopic modification. Figure discloses SEQ ID NO: 864.

[0051] Fig. 13 illustrates fragment electrophiles that target pro-CASP8. **A**, **7** blocked IA-rhodamine **16** labeling of pro-CASP8. Experiments were performed with recombinant, purified pro-CASP8 (bearing a C409S mutation to eliminate IA-rhodamine labeling at this site) added to Ramos cell lysate at a final concentration of 1 μ M and treated with the indicated concentrations of **7** followed by IA-rhodamine (2 μ M). Note that a C360S/C409S-mutant of pro-CASP8 did not label with IA-rhodamine. **B**, IC₅₀ curve for blockade of IA-rhodamine labeling of pro-CASP8 (C409S) by **7**. **C**, **7** (50 μ M) fully competed IA-alkyne-labeling of C360 of endogenous CASP8 in cell lysates as measured by isoTOP-ABPP. Representative MS1 chromatograms are shown for the C360-containing peptide of CASP8. **D**, **7** selectively blocked probe labeling of pro-CASP8 compared to active CASP8. Recombinant pro- and active- CASP8 (added to Ramos cell lysates at a final concentration of 1 μ M each) were treated with **7** (50 μ M) or the established caspase inhibitor, Ac-DEVD-CHO ("DEVD" disclosed as SEQ ID NO: 857) (20 μ M), for 1 h followed by labeling with the click probe **61** (25 μ M) for pro-CASP8 and the Rho-DEVD-AOMK probe ("DEVD" disclosed as SEQ ID NO: 857) (2 μ M) for active-CASP8. SDS-PAGE and in-gel fluorescence scanning revealed that **7** competes **61**-labeling of pro-CASP8, but not Rho-DEVD-AOMK ("DEVD" disclosed as SEQ ID NO: 857) of active-CASP8, and, conversely, DEVD-CHO ("DEVD" disclosed as SEQ ID NO: 857) competed Rho-DEVD-AOMK ("DEVD" disclosed as SEQ ID NO: 857) labeling of active-CASP8, but not **61**-labeling of pro-CASP8. **E**, Neither **7** nor control fragment **62** (100 μ M each) inhibited the activity of recombinant, purified active CASP8 and CASP3, which were assayed following addition to Ramos cell lysate using DEVD-AMC and IETD-AFC substrates, respectively. DEVD-CHO ("DEVD" disclosed as SEQ ID NO: 857) (20 μ M) inhibited the activity of both CASP8 and CASP3. **F**, **7** (30 μ M) blocked IA-alkyne labeling of C360 of pro-CASP8, but not active-CASP8 as measured by isoTOP-ABPP. Recombinant pro- and active-CASP8 were added to Ramos lysates at 1 μ M and then treated with **7** (30 μ M) followed by isoTOP-ABPP. **G**, Substitution of a

naphthylamine for the aniline portion of **7** furnishes a control fragment **62** that did not compete with IA-rhodamine labeling of C360 of pro-CASP8. **H**, **7**, but not control fragment **62**, blocked extrinsic, but not intrinsic apoptosis. Jurkat cells (1.5 million cells/mL) were incubated with **7** or **62** (30 μ M) or the pan-caspase inhibitor VAD-FMK (100 μ M) for 30 min prior to addition of staurosporine (2 μ M) or *SuperFasLigand*TM (100 ng/mL). Cells were incubated for 6 hours and viability was quantified with CellTiter-Glo®. RLU- relative light unit. **I**, For cells treated as described in **H**, cleavage of PARP (89 kDa), CASP8 (p43/p41), and CASP3 (p19/p17) was visualized by western blot. For panels **B**, **E**, and **H**, data represent mean values $\pm$ SEM for at least three independent experiments.

[0052] **Fig. 14** shows electrophile compounds that target pro-CASP8 and pro-CASP10.

Heatmap showing *R* values for caspases measured by quantitative proteomics in Jurkat cells treated with **7**, **63-R**, or **62** followed by probe **61** (10 μ M, 1 h) (**A**). Comparison of effects of **7** and **63-R** on FasL-induced apoptosis in Jurkat cells or anti-CD3, anti-CD28-activated primary human T cells (**B**). For **B**, data represent mean values $\pm$ SEM for at least three independent experiments, and results are representative of multiple experiments performed with T cells from different human subjects. Statistical significance was calculated with unpaired students t-tests comparing DMSO- to fragment treated samples; ****, $p < 0.0001$ and comparing Jurkat to T cells #####, $p < 0.0001$.

[0053] **Fig. 15** illustrates a fraction of liganded (62%; 341 of 553 quantified cysteines) and unliganded (20%; 561 of 2870 quantified cysteines) cysteines that are sensitive to heat denaturation measured by IA-alkyne labeling ($R > 3$ native/heat denatured).

[0054] **Fig. 16** shows a percentage of proteins identified by isoTOP-ABPP as liganded by fragments **3** and **14** and enriched by their corresponding click probes **19** and **18** that are sensitive to heat denaturation (64% (85 of 133 quantified protein targets) and 73% (19 of 26 quantified protein targets), respectively). Protein enrichment by **18** and **19** was measured by whole protein capture of isotopically-SILAC labeled MDA-MB-231 cells using quantitative (SILAC) proteomics.

[0055] **Fig. 17A-B** illustrate exemplary fractions of cysteines predicted based on isoTOP-ABPP method or IA-alkyne probe. **Fig. 17A** shows the fraction of cysteines predicted to be ligandable or unligandable by reactive docking that were quantified in isoTOP-ABPP experiments. **Fig. 17B** shows the fraction of cysteines predicted to be ligandable or unligandable by reactive docking that show heat-sensitive labeling by the IA-alkyne probe ($R > 3$ native/heat denatured).

[0056] **Fig. 18** shows lysates from HEK293T cells expressing WT or C22A-MLTK treated with the indicated fragments and then an ibrutinib-derived activity probe **59** at 10 μ M. MLTK

labeling by **59** was detected by CuAAC conjugation to a rhodamine-azide tag and analysis by SDS-PAGE and in-gel fluorescence scanning.

[0057] **Fig. 19** shows click probe **18** (25 μ M) labeled WT-IMPDH2 and C331S-IMPDH2, but not C140S-IMPDH2 (or C140S/C331S-IMPDH2). Labeling was detected by CuAAC conjugation to a rhodamine-azide reporter tag and analysis by SDS-PAGE and in-gel fluorescence scanning. Recombinant IMPDH2 WT and mutants were expressed and purified from *E. coli* and added to Jurkat lysates to a final concentration of 1 μ M protein.

[0058] **Fig. 20** shows the apparent IC₅₀ curve for blockade of IA rhodamine-labeling of R132H-IDH1 by **20**.

[0059] **Fig. 21A-C** show the activity of compounds **7** and **62** with respect to different recombinant caspases. **Fig. 21A** shows that **7** does not inhibit active caspases. Recombinant, active caspases were added to MDA-MD-231 lysate to a final concentration of 200 nM (CASP2, 3, 6, 7) or 1 μ M (CASP8, 10), treated with z- VAD-FMK (25 μ M) or **7** (50 μ M), followed by labeling with the Rho-DEVD-AOMK probe ("DEVD" disclosed as SEQ ID NO: 857) (2 μ M). **Fig. 21B** shows a western blot of the cleavage of PARP (96 kDa), CASP8 (p43/p41, p18), and CASP3 (p17). **Fig. 21C** shows that **7** protects Jurkat cells from extrinsic, but not intrinsic apoptosis. Cleavage of PARP, CASP8, and CASP3 detected by western blotting as shown in **Fig. 21B** was quantified for three (STS) or two (FasL) independent experiments. Cleavage products (PARP (96 kDa), CASP8 (p43/p41), CASP3 (p17)) were quantified for compound treatment and the % cleavage relative to DMSO treated samples was calculated. For **Fig. 21C**, STS data represent mean values $\pm$ SEM for three independent experiments, and FasL data represent mean values $\pm$ SD for two independent experiments. Statistical significance was calculated with unpaired students t-tests comparing active compounds (VAD-FMK and **7**) to control compound **62**; **, $p < 0.01$, ***, $p < 0.001$, ****, $p < 0.0001$.

[0060] **Fig. 22** shows that CASP10 is involved in intrinsic apoptosis in primary human T cells. **A**, Representative MS1 peptide signals showing *R* values for caspases detected by quantitative proteomics using probe **61**. ABPP-SILAC experiments. Jurkat cells (10 million cells) were treated with either DMSO (heavy cells) or the indicated compounds (light cells) for 2 h followed by probe **61** (10 μ M, 1 h). **B**, **7** competed **61**-labeling of pro-CASP8 and CASP10, whereas **63-R** selectively blocked probe labeling of pro-CASP8. **C**, **7**, but not **63-R** block probe labeling of pro-CASP10. Recombinant pro-CASP10 was added to MDA-MB-231 lysates to a final concentration of 300 nM, treated with the indicated compounds, and labeled with probe **61**. Mutation of the catalytic cysteine C401A fully prevented labeling by **61**. **D**, Apparent IC₅₀ curve for blockade of **61**-labeling of pro-CASP10 by **7**, **63-R** or **63-S**. **E**, Neither **7** nor **63** (25 μ M each) inhibited the activity of recombinant, purified active CASP10 (500 nM), which was

assayed following addition of the protein to MDA-MB-231 lysate using fluorometric AEVD-AMC substrate ("AEVD" disclosed as SEQ ID NO: 859). DEVD-CHO ("DEVD" disclosed as SEQ ID NO: 857) (20 μ M) inhibited the activity of CASP10. **F**, Apparent IC₅₀ curve for blockade of **61** labeling of pro-CASP8 and pro-CASP10 by **63-R**. **G**, **63-R** shows increased potency against pro- CASP8. Recombinant pro-CASP8 was added to MDA-MB-231 lysates to a final concentration of 300 nM, treated with the indicated compounds and labeled with probe **61**. **H**, Apparent IC₅₀ curve for blockade of **61** labeling of pro-CASP8 by **63-R** compared with **63-S**. The structure of 63-S is shown. **I**, CASP10 is more highly expressed in primary human T cells compared to Jurkat cells. Western blot analysis of full-length CASP10, CASP8 and GAPDH expression levels in Jurkat and T-cell lysates (2 mg/mL). **J**, Jurkat cells (150,000 cells/well) were incubated with **7** or **63-R** at the indicated concentrations for 30 min prior to addition of staurosporine (2 μ M) or *SuperFasLigand*TM (100 ng /mL). Cells were incubated for 4 h and viability was quantified with CellTiter-Glo®. **K**, Jurkat cells treated as in **J**, but with **63-R** or **63-S**. **L**, HeLa cells (20,000 cells/well) were seeded and 24 h later treated with the indicated compounds for 30 minutes prior to addition of *SuperFasLigand*TM (100 ng /mL) and cycloheximide (CHX, 2.5 ng/mL). Cells were incubated for 6 h and viability quantified with CTG. **M**, For T cells treated as in **Fig. 14B** cleavage of CASP10 (p22), CASP8 (p18), CASP3 (p17) and RIPK (33 kDa) was visualized by western blotting. For panels **D-F**, **H**, and **J-K**, data represent mean values $\pm$ SEM for at least three independent experiments. Statistical significance was calculated with unpaired students t-tests comparing DMSO- to fragment-treated samples; **, $p < 0.01$, ****, $p < 0.0001$.

[0061] **Fig. 23A–F** exemplify DMF inhibits the activation of primary human T cells. **Fig. 23A** illustrates the chemical structures of DMF, MMF, and DMS. **Fig. 23B – Fig. 23E** illustrate bar graphs that exemplify IL-2 release (**Fig. 23B**), CD25 expression (**Fig. 23C** and **Fig. 23D**), and CD69 expression (**Fig. 23E**) in primary human T cells, either unstimulated (Unstim) or stimulated (Stim) with anti-CD3 + anti-CD28 in the presence of DMSO or the indicated concentrations of DMF, MMF, and DMS for 8 hours. **Fig. 23F** illustrates a bar graph that exemplifies time course of DMF effects. T cells were stimulated with anti-CD3 + anti-CD28 for the indicated periods of time before beginning DMF treatment. Cells were harvested 24 h after beginning T cell stimulation. Shown are data gated on CD4⁺ cells. Data represent mean $\pm$ SE; n = 4-6 experiments/group. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ in comparison to DMSO group.

[0062] **Fig. 24** illustrates a bar graph that exemplifies DMF does not affect T cell viability. Primary human T cells were stimulated with anti-CD3 and anti-CD28 antibodies as indicated and treated concomitantly with compound for 8 h. Cells were then stained with LIVE/DEAD

fixable blue stain, and analyzed by flow cytometry. Shown are data gated on CD4⁺ cells. Data represent mean $\pm$ SE for four experiments per group.

[0063] **Fig. 25A-B** illustrate bar graphs that exemplify DMF, but not MMF, inhibits the activation of primary mouse T cells. Splenic T cells were harvested from C57BL/6 mice and left either unstimulated (Unstim) or stimulated (Stim) with anti-CD3 + anti-CD28 in the presence of DMSO or the indicated concentrations of DMF, MMF, and DMS for 8 h. Activation was assessed by measuring CD25 (Fig. 25A) and CD69 (Fig. 25B) expression. Data represent mean $\pm$ SE for four experiments per group. *** $p < 0.001$ in comparison to DMSO group.

[0064] **Fig. 26A-D** illustrate bar graphs that exemplify inhibitory effects of DMF are equivalent in Nrf2(+/+) and (-/-) T cells and not caused by reductions in cellular GSH. Fig. 26A exemplifies CD25 expression in anti-CD3 + anti-CD28-stimulated Nrf2(+/+) and (-/-) T cells. Splenic T cells were harvested from Nrf2(+/+) and (-/-) mice, then stimulated in the presence of indicated compounds for 24h. Fig. 26B and Fig. 26C exemplify treatment with DMF or BSO causes significant reductions in GSH content of human T cells. Primary human T cells were stimulated with anti-CD3 + anti-CD28 antibodies and treated with DMF (50 μ M, 2 hours) or BSO (2.5 mM, 4 hours), after which intracellular GSH levels were measured. Fig. 26D exemplifies that BSO does not alter T cell activation. Primary human T cells were treated with DMSO, DMF (50 μ M), or BSO (2.5 mM) and stimulated as indicated for 8 h, after which CD25 expression was measured. Data represent mean $\pm$ SE for two biological replicates, with 3-4 technical replicates per biological replicate. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ in comparison to DMSO groups.

[0065] **Fig. 27A-F** exemplify isoTOP-ABPP of DMF-treated primary human T cells. Fig. 27A illustrates a graph that exemplifies isoTOP-ABPP ratios, or R values, for > 2400 Cys residues in primary human T cells treated with DMSO or DMF or MMF (50 μ M, 4 h). Fig. 27B illustrates a graph that exemplifies expanded profile for DMF-sensitive Cys residues (R values > 4 for DMSO/DMF). For Fig. 27A and Fig. 27B, data represent aggregate quantified Cys residues from five biological replicates. For Cys residues quantified in more than one replicate, average ratios are reported. Dashed line designates R values > 4 , which was used to define DMF-sensitive Cys residues (> 4 -fold reductions in IA-alkyne reactivity in DMF-treated T cells). Fig. 27C and Fig. 27D illustrate graphs that exemplify concentration- and time- dependent profiles for DMF-sensitive Cys residues in T cells, respectively. For additional concentrations (10 and 25 μ M) and time points (1 and 2 h), data represent aggregate quantified Cys residues from one-three isoTOP-ABPP experiments per group. Fig. 27E illustrates a chart which exemplifies fraction of proteins for which both a DMF-sensitive Cys residue and at least one additional Cys residue was quantified (Left) and, fraction of these proteins where additional Cys residue was

clearly unchanged (Right) (R value < 2.0 for DMSO/DMF). Unclear calls mark proteins with DMF-sensitive Cys residues where the R value for second Cys showed marginal evidence of potential change (R values between 2.0 and 3.9). Fig. 27F illustrates representative MS1 profiles for quantified Cys residues in PRKDC, one of which (C4045) shows sensitivity to DMF.

[0066] Fig. 28A-B illustrate bar graphs that exemplify the total number of unique quantified peptides (Fig. 28A) and proteins (Fig. 28B) begin to plateau after five biological replicates of the isoTOP-ABPP experiment in primary human T cells (treated with 50 uM DMF for 4 h).

[0067] Fig. 29 illustrates a graph that exemplifies isoTOP-ABPP of BSO-treated primary human T cells. Cells were treated with 2.5 mM BSO for 4 hours. Data represent aggregate quantified Cys residues from two isoTOP-ABPP experiments per group.

[0068] Fig. 30A-C exemplify conservation and functional analysis of DMF-sensitive cysteines. Fig. 30A exemplifies fraction of DMF-sensitive cysteines in the human T cell proteome that are conserved in mice. Fig. 30B exemplifies fraction of conserved DMF-sensitive Cys residues in human T cells that were quantified and also sensitive to DMF in mouse T cells. Fig. 30C exemplifies distribution of proteins harboring DMF-sensitive Cys residues by functional class.

[0069] Fig. 31A-C exemplify DMF inhibits p65 translocation to the nucleus in primary human T cells. Fig. 31A exemplify Human T cells were either left unstimulated or stimulated with anti-CD3 and anti-CD28 antibodies and treated with DMSO or DMF (50 uM) for 1 h. Fig. 31B illustrates a bar graph that exemplifies ratio of nuclear to cytoplasmic localization of p65 for samples shown in Fig. 31A, as well as samples treated with MMF (50 uM) or DMS (50 uM). Fig. 31C exemplifies p65 levels in whole cell lysate.

[0070] Fig. 32A-G exemplify DMF-sensitive C14/C17 residues in PKC θ are important for CD28 interactions and T cell activation. Fig. 32A illustrates representative MS1 profiles for DMF-sensitive (C14/C17) and -insensitive (C322) Cys residues in PKC θ . Fig. 32B exemplifies sequence conservation analysis of human and mouse PKC θ , human PKC δ , and human PKC ϵ (SEQ ID NOS 865-868, respectively, in order of appearance). Shown in red are C14 and C17. Fig. 32C illustrates location of DMF-sensitive C14 and C17 residues in the C2 domain of PKC θ (PDB accession number 2ENJ). Fig. 32D exemplifies DMF, but not MMF, treatment blocks the association of PKC θ with CD28. Peripheral CD4⁺ T cells from C57BL/6 mice were pre-incubated with DMSO, DMF (50 μ M), or MMF (50 μ M), either left unstimulated or stimulated with anti-CD3 + anti-CD28 for 5 min, then washed and lysed. Immunoprecipitations (IPs) were performed in the cell lysates with anti-CD28 or control IgG antibodies and IPs blotted for CD28 or PKC θ . Fig. 32E illustrates Co-IP of WT PKC θ and the C14S/C17S (2CS) PKC θ mutant with CD28. PKC θ (-/-) T cells were reconstituted with empty vector (EV), WT PKC θ , or the 2CS PKC θ mutant. Fig. 32F and Fig. 32 G illustrate PKC θ (-/-) T cells reconstituted with WT or 2CS

PKC θ were assayed for activation potential by measuring CD25 expression (Fig. 32F) and IL-2 (Fig. 32G). For Fig. 32E - Fig. 32G, PKC θ (---) T cell cultures were pre-activated with plate-coated anti-CD3 + anti-CD28 for 24 h before retroviral transduction with empty vector, WT PKC θ , or the 2CS PKC θ mutant. Cells were rested in culture medium without stimulation for 48 h, then re-stimulated with or without 1 μ g/mL plate-coated anti-CD3(+28) overnight (Fig. 32F), for 48 h (Fig. 32G), or with soluble 10 μ g/mL anti-CD3 + anti-CD28 for 5 min prior to IP (Fig. 32D). For Fig. 32D and Fig. 32E, data are from a single experiment representative of three biological replicates. For Fig. 32F and Fig. 32G, data represent mean $\pm$ SE for three biological replicates. ***p < 0.001 in comparison to WT PKC θ group.

[0071] Fig. 33A-D exemplify DMF sensitivity of C14/C17 in PKC θ . Fig. 33A illustrates representative MS1 profile of C14/C17 of mouse PKC θ shows sensitivity to DMF (50 μ M, 4 h) in isoTOP-ABPP experiments. Fig. 33B and Fig. 33C exemplify Time- and concentration-dependence of DMF sensitivity of C14/C17 in human PKC θ , respectively, as determined by isoTOP-ABPP experiments. Fig. 33D exemplifies C14/C17 of human PKC θ are insensitive to MMF treatment (50 μ M MMF, 4 h).

[0072] Fig. 34A-B exemplify DMF-sensitive Cys residue in ADA. Fig. 34A illustrates the DMF-sensitive Cys, C75 (magenta), is $\sim$ 25 angstroms from the ADA active site (orange). Fig. 34B illustrates mutations in both residues neighboring C75 (G74 and R76 (blue)) have been associated with the severe combined immunodeficiency known as ADASID (OMIM: 608958). PDB accession number: 3IAR.

DETAILED DESCRIPTION OF THE INVENTION

[0073] Cysteine containing proteins encompass a large repertoire of proteins that participate in numerous cellular functions such as mitogenesis, proliferation, apoptosis, gene regulation, and proteolysis. These proteins include enzymes, transporters, receptors, channel proteins, adaptor proteins, chaperones, signaling proteins, plasma proteins, transcription related proteins, translation related proteins, mitochondrial proteins, or cytoskeleton related proteins. Dysregulated expression of a cysteine containing protein, in many cases, is associated with or modulates a disease, such as an inflammatory related disease, a neurodegenerative disease, or cancer. As such, identification of a potential agonist/antagonist to a cysteine containing protein aids in improving the disease condition in a patient.

[0074] In some instances, potential constraints exist in drug screening due to the structurally complex compound and the inability of some of the structurally complexed compound to interact with the protein. As such, small molecule fragments are employed in some instances to serve as launching point for structure-guided elaboration of an initial interaction into a high-affinity drug. In some instances, one method of identifying a small molecule fragment that

interacts with a cysteine containing protein is through monitoring their interaction under an *in vitro* environment. However in some cases, the *in vitro* environment does not mimic the native condition of the cysteine containing protein. In other cases, the *in vitro* environment lacks additional helper proteins to facilitate interaction with the small molecule fragment. Further still, in some instances, difficulties arise during the expression and/or purification stage of the cysteine-containing protein.

[0075] Described herein is another method of identifying small molecule fragments for interaction with a cysteine containing protein. In some instances, this method allows for mapping of small molecule fragments for interaction with a cysteine containing protein under native conditions, thereby allows for an accurate mapping of interaction with potential small molecule fragments. In some instances, this method also allows for identification of novel cysteine containing protein targets as this method eliminates the need of recombinant expression and purification.

[0076] In some embodiments, also described herein are compositions, cells, cell populations, assays, probes, and service related to the method of identifying a small molecule fragment for interaction with a cysteine containing protein.

[0077] *General Methodology*

[0078] In some embodiments, the methods described herein utilize a small molecule fragment and a cysteine-reactive probe for competitive interaction with a cysteine-containing protein. In some embodiments, the method is as described in Fig. 1A. Fig. 1A illustrates contacting a first cell solution with a small molecule fragment for an extended period of time prior to incubating the first cell solution with a first cysteine-reactive probe to generate a first group of cysteine-reactive probe-protein complexes. In some embodiments, the extended period of time is about 5, 10, 15, 20, 30, 60, 90, 120 minutes or longer. In some instances, the small molecule fragment competes with the first cysteine-reactive probe for interaction with a protein target. In some instances, the small molecule fragment or the cysteine-reactive probe form a covalent bond via a Michael's reaction with a cysteine residue of the cysteine containing protein. Fig. 1A further illustrates contacting a second cell solution with a second cysteine-reactive probe to generate a second group of cysteine-reactive probe-protein complexes. In some instances, the first cysteine-reactive probe and the second cysteine-reactive probe are the same.

[0079] In some embodiments, cells from the second cell solution are grown in an enriched media (e.g., an isotopically enriched media). In some cases, cells from the first cell solution are grown in an enriched media (e.g., an isotopically enriched media). In some instances, cells from both the first cell solution and the second cell solution are grown in two different enriched media (e.g., two different isotopically enriched media) so that a protein obtained from cells grown in

the first cell solution is distinguishable from a protein obtained from cells grown in the second cell solution. In other embodiments, cells from only one of the cell solutions (e.g., either the first cell solution or the second cell solution) are grown in an enriched media (e.g., isotopically enriched media). In such cases, a protein obtained from the enriched cells (e.g., isotopically enriched cells) is distinguishable from a protein obtained from cells that have not been enriched (e.g., isotopically enriched).

[0080] As illustrated in Fig. 1A, in some instances the second cell solution is not treated with a small molecule fragment. In such cases, the second cell solution acts as a control.

[0081] In some instances, cells from the second cell solution are further treated with a buffer. In some cases, the buffer is DMSO. In some cases, cells from the second cell solution are not treated with a small molecule fragment and the second cell solution acts as a control.

[0082] In some instances, a first group of cysteine-reactive probe-protein complexes and a second group of cysteine-reactive probe-protein complexes are harvested separately and combined to generate a set of cysteine-reactive probe-protein complexes which is further processed by a proteomic analysis means. In some cases, either the first group of cysteine-reactive probe-protein complexes or the second group of cysteine-reactive probe-protein complexes contain labeled proteins obtained from cells grown in an enriched media (e.g., isotopically enriched media). In some cases, both groups of cysteine-reactive probe-protein complexes contain labeled proteins obtained from cells grown in two different enriched media (e.g., two different isotopically enriched media). In other cases, either the first group of cysteine-reactive probe-protein complexes, the second group of cysteine-reactive probe-protein complexes, or both groups of cysteine-reactive probe-protein complexes contain labeled proteins in which the proteins have been labeled after harvesting from a cell.

[0083] In some instances, a first group of cysteine-reactive probe-protein complexes and a second group of cysteine-reactive probe-protein complexes are harvested separately and the proteins from one of the two groups of cysteine-reactive probe-protein complexes are subsequently labeled (e.g., by methylation). In some cases, first group of cysteine-reactive probe-protein complexes and a second group of cysteine-reactive probe-protein complexes are then combined and subjected to proteomic analysis means.

[0084] In other instances, a first group of cysteine-reactive probe-protein complexes and a second group of cysteine-reactive probe-protein complexes are harvested separately and both groups are subjected to proteomic analysis means. In some cases, data obtained from a proteomic analysis means is then combined for further analysis.

[0085] In some embodiments, the proteomic analysis means comprises a mass spectroscopy method. In some instances, the mass spectroscopy method is a liquid-chromatography-mass

spectrometry (LC-MS) method. In some cases, the proteomic analysis means further comprise analyzing the results from the mass spectroscopy method by an algorithm for protein identification. In some cases, the algorithm combines the results from the mass spectroscopy method with a protein sequence database for protein identification. In some cases, the algorithm comprises ProLuCID algorithm, Probity, Scaffold, SEQUEST, or Mascot. In some cases, the mass spectroscopy method is a MALDI-TOF based method.

[0086] In some embodiments, a value is assigned to each of the cysteine binding protein from the cysteine-reactive probe-protein complexes after proteomic analysis, in which the value is determined from the proteomic analysis. In some cases, the value assigned to each of the cysteine containing protein is obtained from a mass spectroscopy analysis. In some instances, the value is an area-under-the curve from a plot of signal intensity as a function of mass-to-charge ratio. In some embodiments, a first value is assigned to a cysteine binding protein from the first group of cysteine-reactive probe-protein complex of the first cell solution and a second value of the same cysteine binding protein from the second group of cysteine-reactive probe-protein complex of the second cell solution. In some instances, a ratio is then calculated between the two values, the first value and the second value, and assigned to the same cysteine binding protein. In some instances, a ratio of greater than 2 indicates that the cysteine binding protein is a candidate for interacting with the small molecule fragment. In some instances, the ratio is greater than 2.5, 3, 3.5, 4, 4.5, 5, 6, 7, 8, 9, or 10. In some cases, the ratio is at most 20. In some instances, the same small molecule fragment interacts with a number of cysteine binding proteins in the presence of a cysteine-reactive probe. In some instances, the small molecule modulates the interaction of a cysteine-reactive probe with its cysteine binding protein partners. In some instances, the spectrum of ratios for a small molecule fragment with its interacting protein partners in the presence of a cysteine-reactive probe indicates the specificity of the small molecule fragment toward the protein. In some instances, the spectrum of ratio indicates whether the small molecule fragment is a specific inhibitor to a protein or a pan inhibitor.

[0087] In some embodiments, the cysteine containing protein identified by the above method comprises a biologically active cysteine residue. In some instances, the biologically active cysteine site is a cysteine residue that is located about 10Å or less to an active-site ligand or residue. In some cases, the cysteine residue that is located about 10Å or less to the active-site ligand or residue is an active site cysteine. In some cases, the biologically active cysteine site is an active site cysteine. In some embodiments, the biologically active cysteine site is a cysteine residue that is located greater than 10Å from an active-site ligand or residue. In some cases, the cysteine residue that is located greater than 10Å from the active-site ligand or residue is a non-

active site cysteine. In some instances, the biologically active cysteine site is a non-active site cysteine.

[0088] In some embodiments, the small molecule fragment that covalently interacts with the biologically active cysteine impairs and/or inhibits activity of the cysteine containing protein. In some instances, the cysteine containing protein exists in an active form. In some embodiments, the small molecule fragment and/or the cysteine-reactive probe interact with the active form of the cysteine containing protein. In some instances, the cysteine containing protein exists in a pro-active form. In some embodiments, the small molecule fragment and/or the cysteine-reactive probe interact with the pro-active form of the cysteine containing protein.

[0089] In some embodiments, the structural environment of the biologically active cysteine residue modulates the reactivity of the cysteine residue. In some embodiments, the structural environment is a hydrophobic environment or a hydrophilic environment. In some embodiments, the structural environment is a charged environment. In some embodiments, the structural environment is a nucleophilic environment.

[0090] In some embodiments, the cysteine containing protein is an enzyme, a transporter, a receptor, a channel protein, an adaptor protein, a chaperone, a signaling protein, a plasma protein, transcription related protein, translation related protein, mitochondrial protein, or cytoskeleton related protein. In some instances, the cysteine containing protein is an enzyme, a transporter, a receptor, a channel protein, an adaptor protein, a chaperone, a signaling protein, transcription related protein, or translation related protein. In some embodiments, the cysteine containing protein is a protein illustrated in Tables 1, 2, 3, 8 or 9. In some instances, the cysteine residue of the cysteine-containing proteins illustrated in Tables 1, 2, 3, 8 or 9 is denoted by (*) in Tables 1, 2, 3, 8 or 9.

[0091] In some instances, a set of cysteine-reactive probes are added to the cell solutions. For example, a first set of cysteine-reactive probes are added to the first cell solution and a second set of cysteine-reactive probes are added to the second cell solution. In some cases, each cysteine-reactive probe is different within the set. In some instances, the first set of cysteine-reactive probes is the same as the second set of cysteine-reactive probes. In some cases, the first set of cysteine-reactive probes generate a third group of cysteine-reactive probe-protein complexes and the second set of cysteine-reactive probes generate a fourth group of cysteine-reactive probe-protein complexes. In some instances, the set of cysteine-reactive probes further facilitates identification of cysteine containing proteins.

[0092] In some embodiments, the sample is a cell sample. In other instances, the sample is a tissue sample.

[0093] In some instances, the method is an *in-situ* method.

Small Molecule Fragments

[0094] In some embodiments, the small molecule fragments described herein comprise non-naturally occurring molecules. In some instances, the non-naturally occurring molecules do not include natural and/or non-natural peptide fragments, or small molecules that are produced naturally within the body of a mammal.

[0095] In some embodiments, the small molecule fragments described herein comprise a molecule weight of about 100 Dalton or higher. In some embodiments, the small molecule fragments comprise a molecule weight of about 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some instances, the molecule weight of the small molecule fragments are between about 150 and about 500, about 150 and about 450, about 150 and about 440, about 150 and about 430, about 150 and about 400, about 150 and about 350, about 150 and about 300, about 150 and about 250, about 170 and about 500, about 180 and about 450, about 190 and about 400, about 200 and about 350, about 130 and about 300, or about 120 and about 250 Dalton.

[0096] In some embodiments, the molecule weight of the small molecule fragments described herein is the molecule weight prior to enrichment with one or more elements selected from a halogen, a nonmetal, a transition metal, or a combination thereof. In some embodiments, the molecule weight of the small molecule fragments described herein is the molecule weight prior to enrichment with a halogen. In some embodiments, the molecule weight of the small molecule fragments described herein is the molecule weight prior to enrichment with a nonmetal. In some embodiments, the molecule weight of the small molecule fragments described herein is the molecule weight prior to enrichment with a transition metal. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur atoms.

[0097] In some embodiments, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some cases, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen. the molecular weight of the small molecule fragment does not include the molecular weight of a transition metal.

[0098] In some embodiments, the small molecule fragments described herein comprise micromolar or millimolar binding affinity. In some instances, the small molecule fragments comprise a binding affinity of about 1 μ M, 10 μ M, 100 μ M, 500 μ M, 1mM, 10mM, or higher.

[0099] In some embodiments, the small molecule fragments described herein has a high ligand efficiency (LE). Ligand efficiency is the measurement of the binding energy per atom of a ligand to its binding partner. In some instances, the ligand efficiency is defined as the ratio of the Gibbs free energy (ΔG) to the number of non-hydrogen atoms of the compound (N):

$$LE = (\Delta G)/N.$$

[00100] In some cases, LE is also arranged as:

$$LE = 1.4 (-\log IC_{50})/N.$$

[00101] In some instances, the LE score is about $0.3 \text{ kcal mol}^{-1} \text{HA}^{-1}$, about $0.35 \text{ kcal mol}^{-1} \text{HA}^{-1}$, about $0.4 \text{ kcal mol}^{-1} \text{HA}^{-1}$, or higher.

[00102] In some embodiments, the small molecule fragments described herein are designed based on the Rule of 3. In some embodiments, the Rule of 3 comprises a non-polar solvent-polar solvent (e.g. octanol-water) partition coefficient $\log P$ of about 3 or less, a molecular mass of about 300 Daltons or less, about 3 hydrogen bond donors or less, about 3 hydrogen bond acceptors or less, and about 3 rotatable bonds or less.

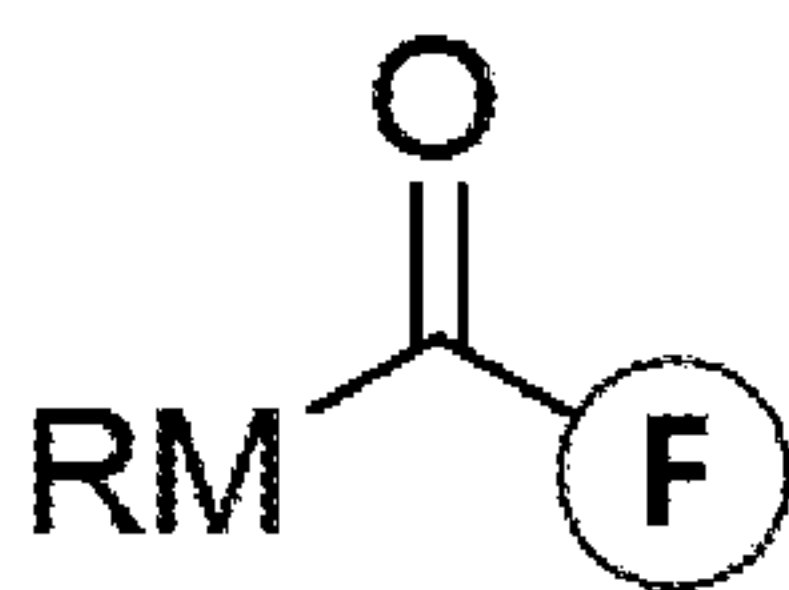
[00103] In some embodiments, the small molecule fragments described herein comprises three cyclic rings or less.

[00104] In some embodiments, the small molecule fragments described herein binds to a cysteine residue of a polypeptide that is about 20 amino acid residues in length or more. In some instances, the small molecule fragments described herein binds to a cysteine residue of a polypeptide that is about 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000 amino acid residues in length or more.

[00105] In some embodiments, the small molecule fragments described herein further comprise pharmacokinetic parameters that are unsuitable as a therapeutic agent for administration without further optimization of the small molecule fragments. In some instances, the pharmacokinetic parameters that are suitable as a therapeutic agent comprise parameters in accordance with FDA guideline, or in accordance with a guideline from an equivalent Food and Drug Administration outside of the United States. In some instances, the pharmacokinetic parameters comprise the peak plasma concentration (C_{max}), the lowest concentration of a therapeutic agent (C_{min}), volume of distribution, time to reach C_{max} , elimination half-life, clearance, and the life. In some embodiments, the pharmacokinetic parameters of the small molecule fragments are outside of the parameters set by the FDA guideline, or by an equivalent Food and Drug Administration outside of the United States. In some instances, a skilled artisan understands in view of the pharmacokinetic parameters of the small molecule fragments described herein that these small molecule fragments are unsuited as therapeutic agents without further optimization.

[00106] In some embodiments, the small molecule fragments described herein comprise a reactive moiety which forms a covalent interaction with the thiol group of a cysteine residue of a cysteine containing protein, and an affinity handle moiety.

[00107] In some instances, a small molecule fragment described herein is a small molecule fragment of Formula (I):



Formula (I)

wherein:

RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and F is a small molecule fragment moiety.

[00108] In some instances, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some cases, F is obtained from a compound library. In some cases, the compound library comprises ChemBridge fragment library, Pyramid Platform Fragment-Based Drug Discovery, Maybridge fragment library, FRGx from AnalytiCon, TCI-Frag from AnCoreX, Bio Building Blocks from ASINEX, BioFocus 3D from Charles River, Fragments of Life (FOL) from Emerald Bio, Enamine Fragment Library, IOTA Diverse 1500, BIONET fragments library, Life Chemicals Fragments Collection, OTAVA fragment library, Prestwick fragment library, Selcia fragment library, TimTec fragment-based library, Allium from Vitas-M Laboratory, or Zenobia fragment library.

[00109] In some embodiments, the small molecule fragment of Formula (I) does not contain a second binding site. In some instances, the small molecule fragment moiety does not bind to the protein. In some cases, the small molecule fragment moiety does not covalently bind to the protein. In some instances, the small molecule fragment moiety does not interact with a secondary binding site on the protein. In some instances, the secondary binding site is an active site such as an ATP binding site. In some cases, the active site is at least about 10, 15, 20, 25, 35, 40Å, or more away from the biologically active cysteine residue. In some instances, the small molecule fragment moiety does not interact with an active site such as an ATP binding site.

[00110] In some instances, F is a small molecule fragment moiety illustrated in Fig. 3. In some cases, F further comprises a linker moiety that connects F to the carbonyl moiety. In some cases, the small molecule fragment is a small molecule fragment illustrated in Fig. 3.

[00111] In some instances, F is a small molecule fragment moiety selected from: *N*-(4-bromophenyl)-*N*-phenylacrylamide, *N*-(1-benzoylpiperidin-4-yl)-2-chloro-*N*-phenylacetamide, 1-(4-benzylpiperidin-1-yl)-2-chloroethan-1-one, *N*-(2-(1H-indol-3-yl)ethyl)-2-chloroacetamide, *N*-(3,5-bis(trifluoromethyl)phenyl)acrylamide, *N*-(4-phenoxy-3-(trifluoromethyl)phenyl)-*N*-(pyridin-3-ylmethyl)acrylamide, *N*-(3,5-bis(trifluoromethyl)phenyl)acetamide, 2-chloro-1-(4-(hydroxydiphenylmethyl)piperidin-1-yl)ethan-1-one, (*E*)-3-(3,5-bis(trifluoromethyl)phenyl)-2-cyanoacrylamide, *N*-(3,5-bis(trifluoromethyl)phenyl)-2-bromopropanamide, *N*-(3,5-bis(trifluoromethyl)phenyl)-2-chloropropanamide, *N*-(3,5-bis(trifluoromethyl)phenyl)-*N*-(pyridin-3-ylmethyl)acrylamide, 3-(2-chloroacetamido)-5-(trifluoromethyl)benzoic acid, 1-(4-(5-fluorobenzisoxazol-3-yl)piperidin-1-yl)prop-2-en-1-one, *tert*-butyl 4-(4-acrylamido-2,6-difluorophenyl)piperazine-1-carboxylate, *N*-(4-bromo-2,5-dimethylphenyl)acrylamide, 2-Chloroacetamido-2-deoxy- α/β -D-glucopyranose, 2-chloro-1-(2-methyl-3,4-dihydroquinolin-1(2*H*)-yl)ethan-1-one, *N*-cyclohexyl-*N*-phenylacrylamide, 1-(5-bromoindolin-1-yl)prop-2-en-1-one, *N*-(1-benzylpiperidin-4-yl)-*N*-phenylacrylamide, 2-chloro-*N*-(2-methyl-5-(trifluoromethyl)phenyl)acetamide, 1-(5-bromoindolin-1-yl)-2-chloroethan-1-one, 2-chloro-*N*-(quinolin-5-yl)acetamide, 1-(4-benzylpiperidin-1-yl)prop-2-en-1-one, 2-chloro-*N*-((3-hydroxy-5-(hydroxymethyl)-2-methylpyridin-4-yl)methyl)acetamide, or 1-(6,7-dimethoxy-3,4-dihydroisoquinolin-2(1*H*)-yl)prop-2-en-1-one.

[00112] In some embodiments, the small molecule fragment of Formula (I) comprise a molecule weight of about 100, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some instances, the molecule weight of the small molecule fragment of Formula (I) is between about 150 and about 500, about 150 and about 450, about 150 and about 440, about 150 and about 430, about 150 and about 400, about 150 and about 350, about 150 and about 300, about 150 and about 250, about 170 and about 500, about 180 and about 450, about 190 and about 400, about 200 and about 350, about 130 and about 300, or about 120 and about 250 Dalton.

[00113] In some embodiments, the molecule weight of the small molecule fragment of Formula (I) is the molecule weight prior to enrichment with one or more elements selected from a halogen, a nonmetal, a transition metal, or a combination thereof. In some embodiments, the molecule weight of the small molecule fragment of Formula (I) is the molecule weight prior to enrichment with a halogen. In some embodiments, the molecule weight of the small molecule fragment of Formula (I) is the molecule weight prior to enrichment with a nonmetal. In some embodiments, the molecule weight of the small molecule fragment of Formula (I) is the molecule weight prior to enrichment with a transition metal.

[00114] In some embodiments, the molecular weight of the small molecule fragment of Formula (I) does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some embodiments, the molecular weight of the small molecule fragment of Formula (I) does not include the molecular weight of a halogen. In some embodiments, the molecular weight of the small molecule fragment of Formula (I) does not include the molecular weight of a transition metal.

[00115] In some instances, the small molecule fragment of Formula (I) comprises micromolar or millimolar binding affinity. In some instances, the small molecule fragment of Formula (I) comprises a binding affinity of about 1 μ M, 10 μ M, 100 μ M, 500 μ M, 1mM, 10mM, or higher.

[00116] In some cases, the small molecule fragment of Formula (I) has a LE score about 0.3 kcal mol⁻¹HA⁻¹, about 0.35 kcal mol⁻¹HA⁻¹, about 0.4 kcal mol⁻¹HA⁻¹, or higher

[00117] In some embodiments, the small molecule fragment of Formula (I) follows the design parameters of Rule of 3. In some instances, the small molecule fragment of Formula (I) has a non-polar solvent-polar solvent (e.g. octanol-water) partition coefficient log P of about 3 or less, a molecular mass of about 300 Daltons or less, about 3 hydrogen bond donors or less, about 3 hydrogen bond acceptors or less, and about 3 rotatable bonds or less.

[00118] In some embodiments, the small molecule fragment of Formula (I) comprises three cyclic rings or less.

[00119] In some embodiments, the small molecule fragment of Formula (I) binds to a cysteine residue of a polypeptide (e.g., a cysteine containing protein) that is about 20 amino acid residues in length or more. In some instances, the small molecule fragments described herein binds to a cysteine residue of a polypeptide (e.g., a cysteine containing protein) that is about 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000 amino acid residues in length or more.

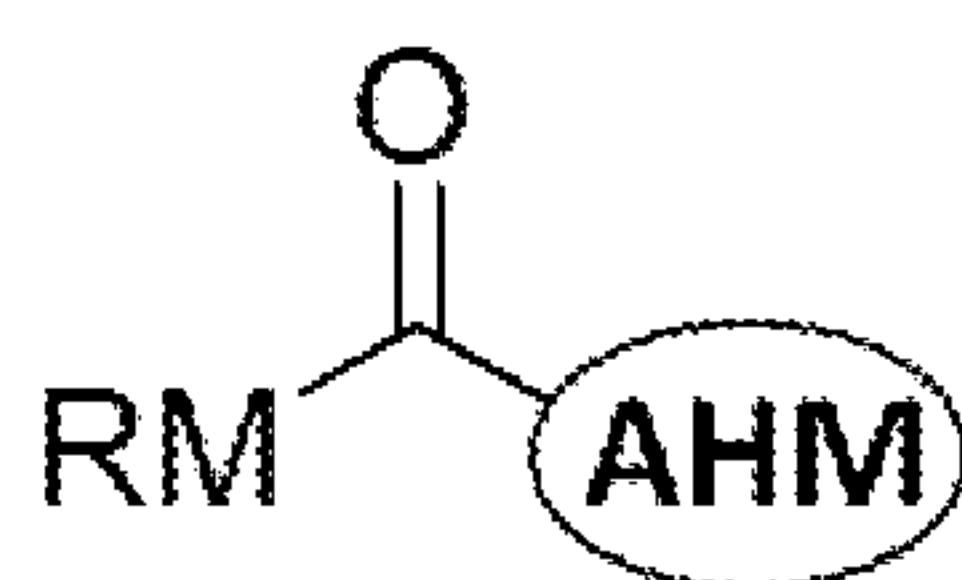
[00120] In some instances, the small molecule fragment of Formula (I) has pharmacokinetic parameters outside of the parameters set by the FDA guideline, or by an equivalent Food and Drug Administration outside of the United States. In some instances, a skilled artisan understands in view of the pharmacokinetic parameters of the small molecule fragment of Formula (I) described herein that these small molecule fragment is unsuited as a therapeutic agent without further optimization.

[00121] In some embodiments, the small molecule fragment is a specific inhibitor or a pan inhibitor.

Cysteine-reactive Probes

[00122] In some embodiments, a cysteine-reactive probe comprises a reactive moiety which forms a covalent interaction with the thiol group of a cysteine residue of a cysteine containing protein, and an affinity handle moiety.

[00123] In some embodiments, a cysteine-reactive probe is a cysteine-reactive probe of Formula (II):



Formula (II)

wherein:

RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond to the thiol group of a cysteine residue; and AHM is an affinity handle moiety.

[00124] In some instances, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some cases, the affinity handle moiety comprises an affinity handle and a binding moiety that facilitates covalent interaction of the cysteine-reactive probe to a cysteine residue of a cysteine-containing protein. In some cases, the binding moiety is a small molecule fragment obtained from a compound library. In some instances, the compound library comprises ChemBridge fragment library, Pyramid Platform Fragment-Based Drug Discovery, Maybridge fragment library, FRGx from AnalytiCon, TCI-Frag from AnCoreX, Bio Building Blocks from ASINEX, BioFocus 3D from Charles River, Fragments of Life (FOL) from Emerald Bio, Enamine Fragment Library, IOTA Diverse 1500, BIONET fragments library, Life Chemicals Fragments Collection, OTAVA fragment library, Prestwick fragment library, Selcia fragment library, TimTec fragment-based library, Allium from Vitas-M Laboratory, or Zenobia fragment library.

[00125] In some embodiments, the affinity handle is a bioorthogonal affinity handle. In some embodiments, the affinity handle utilizes bioorthogonal chemistry. As used herein, bioorthogonal chemistry refers to any chemical reaction that occurs inside of a living system (e.g. a cell) without interfering with native biochemical processes.

[00126] In some cases, the affinity handle comprises a carbodiimide, N-hydroxysuccinimide (NHS) ester, imidoester, pentafluorophenyl ester, hydroxymethyl phosphine, maleimide, haloacetyl, pyridyl disulfide, thiosulfonate, vinylsulfone, hydrazide, alkoxyamine, alkyne, azide, or isocyanate group. In some cases, the affinity handle comprises an alkyne or an azide group.

[00127] In some instances, the affinity handle is an alkyne group. The term “alkyne group” as used in the context of an affinity handle refers to a group with a chemical formula of $\text{H-C}\equiv\text{C-R}$, HC_2R , $\text{R}_1\text{-C}\equiv\text{C-R}_2$, or $\text{R}_1\text{C}_2\text{R}_2$. In the context of the present chemical formula, R, R_1 , and R_2 are independently a cysteine-reactive probe portion described herein, a linker, or a combination thereof. In some cases, the alkyne group is capable of being covalently linked in a chemical reaction with a molecule containing an azide. In some instances, the affinity handle is an azide group.

[00128] In some instances, the affinity handle (e.g. alkyne group or azide group) serve as nonnative and non-perturbed bioorthogonal chemical handles. In some instances, the affinity handle (e.g. alkyne group or azide group) is further derivatized through chemical reactions such as click chemistry. In some instances, the click chemistry is a copper(I)-catalyzed [3+2]-Huisgen 1,3-dipolar cyclo-addition of alkynes and azides leading to 1,2,3-triazoles. In other instances, the click chemistry is a copper free variant of the above reaction.

[00129] In some instances, the affinity handle further comprises a linker. In some instances, the linker bridges the affinity handle to the reactive moiety.

[00130] In some instances, the affinity handle is further conjugated to an affinity ligand. In some cases, the affinity ligand comprises a chromophore, a labeling group, or a combination thereof. In some embodiments, the chromophore comprises fluorochrome, non-fluorochrome chromophore, quencher, an absorption chromophore, fluorophore, organic dye, inorganic dye, metal chelate, or a fluorescent enzyme substrate. In some cases, the chromophore comprises non-fluorochrome chromophore, quencher, an absorption chromophore, fluorophore, organic dye, inorganic dye, metal chelate, or a fluorescent enzyme substrate. In other cases, the chromophore comprises a fluorophore.

[00131] In some embodiments, the fluorophore comprises rhodamine, rhodol, fluorescein, thiofluorescein, aminofluorescein, carboxyfluorescein, chlorofluorescein, methylfluorescein, sulfofluorescein, aminorhodol, carboxyrhodol, chlororhodol, methylrhodol, sulforhodol, aminorhodamine, carboxyrhodamine, chlororhodamine, methylrhodamine, sulforhodamine, thiorhodamine, cyanine, indocarbocyanine, oxacarbocyanine, thiocarbocyanine, merocyanine, cyanine 2, cyanine 3, cyanine 3.5, cyanine 5, cyanine 5.5, cyanine 7, oxadiazole derivatives, pyridyloxazole, nitrobenzoxadiazole, benzoxadiazole, pyren derivatives, cascade blue, oxazine derivatives, Nile red, Nile blue, cresyl violet, oxazine 170, acridine derivatives, proflavin, acridine orange, acridine yellow, arylmethine derivatives, auramine, crystal violet, malachite green, tetrapyrrole derivatives, porphyrin, phthalocyanine, bilirubin 1-dimethylaminonaphthyl-5-sulfonate, 1-anilino-8-naphthalene sulfonate, 2-p-touidiny-6-naphthalene sulfonate, 3-phenyl-7-isocyanatocoumarin, N-(p-(2-benzoxazolyl)phenyl)maleimide, stilbenes, pyrenes, 6-FAM

(Fluorescein), 6-FAM (NHS Ester), 5(6)-FAM, 5-FAM, Fluorescein dT, 5-TAMRA-cadavarine, 2-aminoacridone, HEX, JOE (NHS Ester), MAX, TET, ROX, TAMRA, TARMATM (NHS Ester), TEX 615, ATTOTM 488, ATTOTM 532, ATTOTM 550, ATTOTM 565, ATTOTM Rho101, ATTOTM 590, ATTOTM 633, ATTOTM 647N, TYETM 563, TYETM 665, or TYETM 705.

[00132] In some embodiments, the labeling group is a biotin moiety, a streptavidin moiety, bead, resin, a solid support, or a combination thereof. As used herein, a biotin moiety described herein comprises biotin and biotin derivatives. Exemplary biotin derivatives include, but are not limited by, desthiobiotin, biotin alkyne or biotin azide. In some instances, a biotin moiety described herein is desthiobiotin. In some cases, a biotin moiety described herein is d-Desthiobiotin.

[00133] In some instances, the labeling group is a biotin moiety. In some instances, the biotin moiety further comprises a linker such as a 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or more residues in length. In some instances, the linker further comprises a cleavage site, such as a protease cleavage site. In some cases, the biotin moiety interacts with a streptavidin moiety. In some instances, the biotin moiety is further attached to a bead, such as a streptavidin-coupled bead. In some instances, the biotin moiety is further attached to a resin or a solid support, such as a streptavidin-coupled resin or a streptavidin-coupled solid support. In some instances, the solid support is a plate, a platform, a cover slide, a microfluidic channel, and the like.

[00134] In some embodiments, the affinity handle moiety further comprises a chromophore.

[00135] In some embodiments, the cysteine-reactive probe is a cysteine-reactive probe illustrated in Fig. 3. In some embodiments, the cysteine-reactive probe is a cysteine-reactive probe selected from: *N*-(hex-5-yn-1-yl)-2-iodoacetamide, Iodoacetamide-rhodamine, 3-acrylamido-*N*-(hex-5-yn-1-yl)-5-(trifluoromethyl)benzamide, 3-acrylamido-*N*-(hex-5-yn-1-yl)-5-(trifluoromethyl)benzamide, or 2-chloro-*N*-(1-(3-ethynylbenzoyl)piperidin-4-yl)-*N*-phenylacetamide.

Cysteine Containing Proteins

[00136] In some instances, the cysteine containing protein is a soluble protein or a membrane protein. In some instances, the cysteine containing protein is involved in one or more of a biological process such as protein transport, lipid metabolism, apoptosis, transcription, electron transport, mRNA processing, or host-virus interaction. In some instances, the cysteine containing protein is associated with one or more of diseases such as cancer or one or more disorders or conditions such as immune, metabolic, developmental, reproductive, neurological, psychiatric, renal, cardiovascular, or hematological disorders or conditions.

[00137] In some embodiments, the cysteine containing protein comprises a biologically active cysteine residue. In some embodiments, the cysteine containing protein comprises one or more

cysteines in which at least one cysteine is a biologically active cysteine residue. In some cases, the biologically active cysteine site is a cysteine residue that is located about 10Å or less to an active-site ligand or residue. In some cases, the cysteine residue that is located about 10Å or less to the active-site ligand or residue is an active site cysteine. In other cases, the biologically active cysteine site is a cysteine residue that is located greater than 10Å from an active-site ligand or residue. In some instances, the cysteine residue is located greater than 12Å, 15Å, 20Å, 25Å, 30Å, 35Å, 40Å, 45Å, or greater than 50Å from an active-site ligand or residue. In some cases, the cysteine residue that is located greater than 10Å from the active-site ligand or residue is a non-active site cysteine. In additional cases, the cysteine containing protein exists in an active form, or in a pro-active form.

[00138] In some embodiments, the cysteine containing protein comprises one or more functions of an enzyme, a transporter, a receptor, a channel protein, an adaptor protein, a chaperone, a signaling protein, a plasma protein, transcription related protein, translation related protein, mitochondrial protein, or cytoskeleton related protein. In some embodiments, the cysteine containing protein is an enzyme, a transporter, a receptor, a channel protein, an adaptor protein, a chaperone, a signaling protein, a plasma protein, transcription related protein, translation related protein, mitochondrial protein, or cytoskeleton related protein. In some instances, the cysteine containing protein has an uncategorized function.

[00139] In some embodiments, the cysteine containing protein is an enzyme. An enzyme is a protein molecule that accelerates or catalyzes chemical reaction. In some embodiments, non-limiting examples of enzymes include kinases, proteases, or deubiquitinating enzymes.

[00140] In some instances, exemplary kinases include tyrosine kinases such as the TEC family of kinases such as Tec, Bruton's tyrosine kinase (Btk), interleukin-2-inducible T-cell kinase (Itk) (or Emt/Tsk), Bmx, and Txk/Rlk; spleen tyrosine kinase (Syk) family such as SYK and Zeta-chain-associated protein kinase 70 (ZAP-70); Src kinases such as Src, Yes, Fyn, Fgr, Lck, Hck, Blk, Lyn, and Frk; JAK kinases such as Janus kinase 1 (JAK1), Janus kinase 2 (JAK2), Janus kinase 3 (JAK3), and Tyrosine kinase 2 (TYK2); or ErbB family of kinases such as Her1 (EGFR, ErbB1), Her2 (Neu, ErbB2), Her3 (ErbB3), and Her4 (ErbB4).

[00141] In some embodiments, the cysteine containing protein is a protease. In some embodiments, the protease is a cysteine protease. In some cases, the cysteine protease is a caspase. In some instances, the caspase is an initiator (apical) caspase. In some instances, the caspase is an effector (executioner) caspase. Exemplary caspase includes CASP2, CASP8, CASP9, CASP10, CASP3, CASP6, CASP7, CASP4, and CASP5. In some instances, the cysteine protease is a cathepsin. Exemplary cathepsin includes Cathepsin B, Cathepsin C,

Cathepsin F, Cathepsin H, Cathepsin K, Cathepsin L1, Cathepsin L2, Cathepsin O, Cathepsin S, Cathepsin W, or Cathepsin Z.

[00142] In some embodiments, the cysteine containing protein is a deubiquitinating enzyme (DUB). In some embodiments, exemplary deubiquitinating enzymes include cysteine proteases DUBs or metalloproteases. Exemplary cysteine protease DUBs include ubiquitin-specific protease (USP/UBP) such as USP1, USP2, USP3, USP4, USP5, USP6, USP7, USP8, USP9X, USP9Y, USP10, USP11, USP12, USP13, USP14, USP15, USP16, USP17, USP17L2, USP17L3, USP17L4, USP17L5, USP17L7, USP17L8, USP18, USP19, USP20, USP21, USP22, USP23, USP24, USP25, USP26, USP27X, USP28, USP29, USP30, USP31, USP32, USP33, USP34, USP35, USP36, USP37, USP38, USP39, USP40, USP41, USP42, USP43, USP44, USP45, or USP46; ovarian tumor (OTU) proteases such as OTUB1 and OTUB2; Machado-Josephin domain (MJD) proteases such as ATXN3 and ATXN3L; and ubiquitin C-terminal hydrolase (UCH) proteases such as BAP1, UCHL1, UCHL3, and UCHL5. Exemplary metalloproteases include the Jab1/Mov34/Mpr1 Pad1 N-terminal+ (MPN+) (JAMM) domain proteases.

[00143] In some embodiments, exemplary cysteine containing proteins as enzymes include, but are not limited to, Glyceraldehyde-3-phosphate dehydrogenase (GAPDH), Protein arginine N-methyltransferase 1 (PRMT1), Peptidyl-prolyl cis-trans isomerase NIMA-interaction (PIN1), Acetyl-CoA acetyltransferase (mitochondrial) (ACAT1), Glutathione S-transferase P (GSTP1), Elongation factor 2 (EEF2), Glutathione S-transferase omega-1 (GSTO1), Acetyl-CoA acetyltransferase (mitochondrial) (ACAT1), Protein disulfide-isomerase A4 (PDIA4), Prostaglandin E synthase 3 (PTGES3), Adenosine kinase (ADK), Elongation factor 2 (EEF2), Isoamyl acetate-hydrolyzing esterase 1 homolog (IAH1), Peroxiredoxin-5 (mitochondrial) (PRDX5), Inosine-5-monophosphate dehydrogenase 2 (IMPDH2), 3-hydroxyacyl-CoA dehydrogenase type-2 (HSD17B10), Omega-amidase NIT2 (NIT2), Aldose reductase (AKR1B1), Monofunctional C1-tetrahydrofolate synthase (mitochondrial) (MTHFD1L), Protein disulfide-isomerase A6 (PDIA6), Pyruvate kinase isozymes M1/M2 (PKM), 6-phosphogluconolactonase (PGLS), Acetyl-CoA acetyltransferase (mitochondrial) (ACAT1), ERO1-like protein alpha (ERO1L), Thioredoxin domain-containing protein 17 (TXNDC17), Protein disulfide-isomerase A4 (PDIA4), Protein disulfide-isomerase A3 (PDIA3), 3-ketoacyl-CoA thiolase (mitochondrial) (ACAA2), Dynamin-2 (DNM2), DNA replication licensing factor MCM3 (MCM3), Serine--tRNA ligase (cytoplasmic) (SARS), Fatty acid synthase (FASN), Acetyl-CoA acetyltransferase (mitochondrial) (ACAT1), Protein disulfide-isomerase (P4HB), Deoxycytidine kinase (DCK), Eukaryotic translation initiation factor 3 subunit (EIF3F), Protein disulfide-isomerase A6 (PDIA6), UDP-N-acetylglucosamine-peptide N-acetylglucosamine

(OGT), Ketosamine-3-kinase (FN3KRP), Protein DJ-1 (PARK7), Phosphoglycolate phosphatase (PGP), DNA replication licensing factor MCM6 (MCM6), Fructose-2,6-bisphosphatase TIGAR (TIGAR), Cleavage and polyadenylation specificity factor subunit (CPSF3), Ubiquitin-conjugating enzyme E2 L3 (UBE2L3), Alanine--tRNA ligase, cytoplasmic (AARS), Mannose-1-phosphate guanyltrtransferase alpha (GMPPA), C-1-tetrahydrofolate synthase (cytoplasmic) (MTHFD1), Dynamin-1-like protein (DNM1L), Protein disulfide-isomerase A3 (PDIA3), Aspartyl aminopeptidase (DNPEP), Acetyl-CoA acetyltransferase (cytosolic) (ACAT2), Thioredoxin domain-containing protein 5 (TXNDC5), Thymidine kinase (cytosolic) (TK1), Inosine-5-monophosphate dehydrogenase 2 (IMPDH2), Ubiquitin carboxyl-terminal hydrolase isozyme L3 (UCHL3), Integrin-linked protein kinase (ILK), Cyclin-dependent kinase 2 (CDK2), Histone acetyltransferase type B catalytic subunit (HAT1), Enoyl-CoA delta isomerase 2 (mitochondrial) (ECI2), C-1-tetrahydrofolate synthase (cytoplasmic) (MTHFD1), Deoxycytidine kinase (DCK), Ubiquitin-like modifier-activating enzyme 6 (UBA6), Protein-L-isoaspartate(D-aspartate) O-methyltransferase (PCMT1), Monofunctional C1-tetrahydrofolate synthase (mitochondrial) (MTHFD1L), Thymidylate kinase (DTYMK), Protein ETHE1 (mitochondrial) (ETHE1), Arginine--tRNA ligase (cytoplasmic) (RARS), NEDD8-activating enzyme E1 catalytic subunit (UBA3), Dual specificity mitogen-activated protein kinase (MAP2K3), Ubiquitin-conjugating enzyme E2S (UBE2S), Amidophosphoribosyltransferase (PPAT), Succinate-semialdehyde dehydrogenase (mitochondrial) (ALDH5A1), CAD, Phosphoenolpyruvate carboxykinase (PCK2), 6-phosphofructokinase type C (PFKP), Acyl-CoA synthetase family member 2 (mitochondrial) (ACSF2), Multifunctional protein ADE2 (PAICS), Desumoylating isopeptidase 1 (DESI1), 6-phosphofructokinase type C (PFKP), V-type proton ATPase catalytic subunit A (ATP6V1A), 3-ketoacyl-CoA thiolase (peroxisomal) (ACAA1), Galactokinase (GALK1), Thymidine kinase (cytosolic) (TK1), ATPase WRNIP1 (WRNIP1), Phosphoribosylformylglycinamide synthase (PFAS), V-type proton ATPase catalytic subunit A (ATP6V1A), Thioredoxin domain-containing protein 5 (TXNDC5), 4-trimethylaminobutyraldehyde dehydrogenase (ALDH9A1), Dual specificity mitogen-activated protein kinase (MAP2K4), Calcineurin-like phosphoesterase domain-containing (CPPED1), Dual specificity protein phosphatase 12 (DUSP12), Phosphoribosylformylglycinamide synthase (PFAS), Diphosphomevalonate decarboxylase (MVD), D-3-phosphoglycerate dehydrogenase (PHGDH), Cell cycle checkpoint control protein RAD9A (RAD9A), Peroxiredoxin-1 (PRDX1), Sorbitol dehydrogenase (SORD), Peroxiredoxin-4 (PRDX4), AMP deaminase 2 (AMPD2), Isocitrate dehydrogenase (IDH1), Pyruvate carboxylase (mitochondrial) (PC), Integrin-linked kinase-associated serine/threonine (ILKAP), Methylmalonate-semialdehyde dehydrogenase (ALDH6A1), 26S proteasome non-ATPase regulatory subunit 14

(PSMD14), Thymidylate kinase (DTYMK), 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (PFKFB2), Peroxiredoxin-5 (mitochondrial) (PRDX5), PDP1, Cathepsin B (CTSB), Transmembrane protease serine 12 (TMPRSS12), UDP-glucose 6-dehydrogenase (UGDH), Histidine triad nucleotide-binding protein 1 (HINT1), E3 ubiquitin-protein ligase UBR5 (UBR5), SAM domain and HD domain-containing protein 1 (SAMHD1), Probable tRNA threonylcarbamoyladenine biosynthesis (OSGEP), Methylated-DNA--protein-cysteine methyltransferase (MGMT), Fatty acid synthase (FASN), Adenosine deaminase (ADA), Cyclin-dependent kinase 19 (CDK19), Serine/threonine-protein kinase 38 (STK38), Mitogen-activated protein kinase 9 (MAPK9), tRNA (adenine(58)-N(1))-methyltransferase catalytic (TRMT61A), Glyoxylate reductase/hydroxypyruvate reductase (GRHPR), Aldehyde dehydrogenase (mitochondrial) (ALDH2), Mitochondrial-processing peptidase subunit beta (PMPCB), 3-ketoacyl-CoA thiolase, peroxisomal (ACAA1), Lysophosphatidic acid phosphatase type 6 (ACP6), Ubiquitin/ISG15-conjugating enzyme E2 L6 (UBE2L6), Caspase-8 (CASP8), 2,5-phosphodiesterase 12 (PDE12), Thioredoxin domain-containing protein 12 (TXNDC12), Nitrilase homolog 1 (NIT1), ERO1-like protein alpha (ERO1L), SUMO-activating enzyme subunit 1 (SAE1), Leucine--tRNA ligase (cytoplasmic) (LARS), Protein-glutamine gamma-glutamyltransferase 2 (TGM2), Probable DNA dC- dU-editing enzyme APOBEC-3C (APOBEC3C), Double-stranded RNA-specific adenosine deaminase (ADAR), Isocitrate dehydrogenase (IDH2), Methylcrotonoyl-CoA carboxylase beta chain (mitochondrial) (MCCC2), Uridine phosphorylase 1 (UPP1), Glycogen phosphorylase (brain form) (PYGB), E3 ubiquitin-protein ligase UBR5 (UBR5), Procollagen-lysine,2-oxoglutarate 5-dioxygenase 1 (PLOD1), Ubiquitin carboxyl-terminal hydrolase 48 (USP48), Aconitate hydratase (mitochondrial) (ACO2), GMP reductase 2 (GMPR2), Pyrroline-5-carboxylate reductase 1 (mitochondrial) (PYCR1), Cathepsin Z (CTSZ), E3 ubiquitin-protein ligase UBR2 (UBR2), Cysteine protease ATG4B (ATG4B), Serine/threonine-protein kinase Nek9 (NEK9), Lysine-specific demethylase 4B (KDM4B), Insulin-degrading enzyme (IDE), Dipeptidyl peptidase 9 (DPP9), Decaprenyl-diphosphate synthase subunit 2 (PDSS2), TFIIH basal transcription factor complex helicase (ERCC3), Methionine-R-sulfoxide reductase B2 (mitochondrial) (MSRB2), E3 ubiquitin-protein ligase BRE1B (RNF40), Thymidylate synthase (TYMS), Cyclin-dependent kinase 5 (CDK5), Bifunctional 3-phosphoadenosine 5-phosphosulfate (PAPSS2), Short/branched chain specific acyl-CoA dehydrogenase (ACADSB), Cathepsin D (CTSD), E3 ubiquitin-protein ligase HUWE1 (HUWE1), Calpain-2 catalytic subunit (CAPN2), Dual specificity mitogen-activated protein kinase (MAP2K7), Mitogen-activated protein kinase kinase kinase MLT (MLTK), Bleomycin hydrolase (BLMH), Probable ATP-dependent RNA helicase DDX59 (DDX59), Cystathionine gamma-lyase (CTH), S-

adenosylmethionine synthase isoform type-2 (MAT2A), 6-phosphofructokinase type C (PFKP), Cytidine deaminase (CDA), DNA-directed RNA polymerase II subunit RPB2 (POLR2B), Protein disulfide-isomerase (P4HB), Procollagen-lysine,2-oxoglutarate 5-dioxygenase 3 (PLOD3), Nucleoside diphosphate-linked moiety X motif 8 (mitochondrial) (NUDT8), E3 ubiquitin-protein ligase HUWE1 (HUWE1), Methylated-DNA--protein-cysteine methyltransferase (MGMT), Nitrilase homolog 1 (NIT1), Interferon regulatory factor 2-binding protein 1 (IRF2BP1), Ubiquitin carboxyl-terminal hydrolase 16 (USP16), Glycylpeptide N-tetradecanoyltransferase 2 (NMT2), Cyclin-dependent kinase inhibitor 3 (CDKN3), Hydroxysteroid dehydrogenase-like protein 2 (HSDL2), Serine/threonine-protein kinase VRK1 (VRK1), Serine/threonine-protein kinase A-Raf (ARAF), ATP-citrate synthase (ACLY), Probable ribonuclease ZC3H12D (ZC3H12D), Peripheral plasma membrane protein CASK (CASK), DNA polymerase epsilon subunit 3 (POLE3), Aldehyde dehydrogenase X (mitochondrial) (ALDH1B1), UDP-N-acetylglucosamine transferase subunit ALG13 (ALG13), Protein disulfide-isomerase A4 (PDIA4), DNA polymerase alpha catalytic subunit (POLA1), Ethylmalonyl-CoA decarboxylase (ECHDC1), Protein-tyrosine kinase 2-beta (PTK2B), E3 SUMO-protein ligase RanBP2 (RANBP2), Legumain (LGMN), Non-specific lipid-transfer protein (SCP2), Long-chain-fatty-acid--CoA ligase 4 (ACSL4), Dual specificity protein phosphatase 12 (DUSP12), Oxidoreductase HTATIP2 (HTATIP2), Serine/threonine-protein kinase MRCK beta (CDC42BPB), Histone-lysine N-methyltransferase EZH2 (EZH2), Non-specific lipid-transfer protein (SCP2), Dual specificity mitogen-activated protein kinase (MAP2K7), Ubiquitin carboxyl-terminal hydrolase 28 (USP28), 6-phosphofructokinase (liver type) (PFKL), SWI/SNF-related matrix-associated actin-dependent (SMARCA1), Protein phosphatase methylesterase 1 (PPME1), DNA replication licensing factor MCM5 (MCM5), 6-phosphofructo-2-kinase/fructose-2,6-bisphosphata (PFKFB4), Dehydrogenase/reductase SDR family member 11 (DHRS11), Pyroglutamyl-peptidase 1 (PGPEP1), Probable E3 ubiquitin-protein ligase (MYCBP2), DNA fragmentation factor subunit beta (DFFB), Deubiquitinating protein VCIP135 (VCPIP1), Putative transferase CAF17 (mitochondrial) (IBA57), Calpain-7 (CAPN7), GDP-L-fucose synthase (TSTA3), Protein disulfide-isomerase A4 (PDIA4, Probable ATP-dependent RNA helicase (DDX59), RNA exonuclease 4 (REXO4), PDK1, E3 SUMO-protein ligase (PIAS4), DNA (cytosine-5)-methyltransferase 1 (DNMT1), Alpha-amino adipic semialdehyde dehydrogenase (ALDH7A1), Hydroxymethylglutaryl-CoA synthase (cytoplasmic) (HMGCS1), E3 ubiquitin-protein ligase (SMURF2), Aldehyde dehydrogenase X (mitochondrial) (ALDH1B1), Tyrosine-protein kinase (BTK), DNA repair protein RAD50 (RAD50), ATP-binding domain-containing protein 4 (ATPBD4), Nucleoside diphosphate kinase 3 (NME3), Interleukin-1 receptor-associated kinase 1 (IRAK1), Ribonuclease P/MRP protein

subunit POP5 (POP5), Peptide-N(4)-(N-acetyl-beta-glucosaminyl)asparagin (NGLY1), Caspase-2 (CASP2), Ribosomal protein S6 kinase alpha-3 (RPS6KA3), E3 ubiquitin-protein ligase UBR1 (UBR1), Serine/threonine-protein kinase Chk2 (CHEK2), Phosphatidylinositol 3,4,5-trisphosphate 5-phosphatase (INPPL1), Histone acetyltransferase p300 (EP300), Creatine kinase U-type (mitochondrial) (CKMT1B), E3 ubiquitin-protein ligase TRIM33 (TRIM33), Cancer-related nucleoside-triphosphatase (NTPCR), Aconitate hydratase (mitochondrial) (ACO2), Ubiquitin carboxyl-terminal hydrolase 34 (USP34), Probable E3 ubiquitin-protein ligase HERC4 (HERC4), E3 ubiquitin-protein ligase HECTD1 (HECTD1), Peroxisomal 2,4-dienoyl-CoA reductase (DECR2), Helicase ARIP4 (RAD54L2), Ubiquitin-like modifier-activating enzyme 7 (UBA7), ER degradation-enhancing alpha-mannosidase-like 3 (EDEM3), Ubiquitin-conjugating enzyme E2O (UBE2O), Dual specificity mitogen-activated protein kinase (MAP2K7), Myotubularin-related protein 1 (MTMR1), Calcium-dependent phospholipase A2 (PLA2G5), Mitotic checkpoint serine/threonine-protein kinase (BUB1B), Putative transferase CAF17 (mitochondrial) (IBA57), Tyrosine-protein kinase ZAP-70 (ZAP70), E3 ubiquitin-protein ligase pellino homolog 1 (PELI1), Neuropathy target esterase (PNPLA6), Ribosomal protein S6 kinase alpha-3 (RPS6KA3), N6-adenosine-methyltransferase 70 kDa subunit (METTL3), Fructosamine-3-kinase (FN3K), Ubiquitin carboxyl-terminal hydrolase 22 (USP22), Rab3 GTPase-activating protein catalytic subunit (RAB3GAP1), Caspase-5 (CASP5), L-2-hydroxyglutarate dehydrogenase (mitochondrial) (L2HGDH), Saccharopine dehydrogenase-like oxidoreductase (SCCPDH), FLAD1 FAD synthase, Lysine-specific demethylase 3A (KDM3A), or Ubiquitin carboxyl-terminal hydrolase 34 (USP34).

[00144] In some embodiments, the cysteine containing protein is a signaling protein. In some instances, exemplary signaling protein includes vascular endothelial growth factor (VEGF) proteins or proteins involved in redox signaling. Exemplary VEGF proteins include VEGF-A, VEGF-B, VEGF-C, VEGF-D, and PGF. Exemplary proteins involved in redox signaling include redox-regulatory protein FAM213A.

[00145] In some embodiments, the cysteine containing protein is a transcription factor or regulator. Exemplary cysteine containing proteins as transcription factors and regulators include, but are not limited to, 40S ribosomal protein S3 (RPS3), Basic leucine zipper and W2 domain-containing protein (BZW1), Poly(rC)-binding protein 1 (PCBP1), 40S ribosomal protein S11 (RPS11), 40S ribosomal protein S4, X isoform (RPS4X), Signal recognition particle 9 kDa protein (SRP9), Non-POU domain-containing octamer-binding protein (NONO), N-alpha-acetyltransferase 15, NatA auxiliary subunit (NAA15), Cleavage stimulation factor subunit 2 (CSTF2), Lamina-associated polypeptide 2, isoform alpha (TMPO), Heterogeneous nuclear ribonucleoprotein R (HNRNPR), MMS19 nucleotide excision repair protein homolog (MMS19),

SWI/SNF complex subunit SMARCC2 (SMARCC2), Enhancer of mRNA-decapping protein 3 (EDC3), H/ACA ribonucleoprotein complex subunit 2 (NHP2), WW domain-containing adapter protein with coiled-c (WAC), N-alpha-acetyltransferase 15 NatA auxiliary subunit (NAA15), 40S ribosomal protein S11 (RPS11), Signal transducer and activator of transcription 1 (STAT1), Mediator of RNA polymerase II transcription subunit (MED15), Lamina-associated polypeptide 2 (isoform alpha) (TMPO), MMS19 nucleotide excision repair protein homolog (MMS19), DNA mismatch repair protein Msh2 (MSH2), Recombining binding protein suppressor of hairless (RBPJ), Mediator of RNA polymerase II transcription subunit (MED17), Heterogeneous nuclear ribonucleoprotein U (HNRNPU), Transcription initiation factor IIA subunit 2 (GTF2A2), Chromatin accessibility complex protein 1 (CHRAC1), CDKN2A-interacting protein (CDKN2AIP), Zinc finger protein 217 (ZNF217), Signal transducer and activator of transcription 3 (STAT3), WD repeat and HMG-box DNA-binding protein 1 (WDHD1), Lamina-associated polypeptide 2 (isoform alpha) (TMPO), Lamina-associated polypeptide 2 (isoforms beta/gam) (TMPO), Interferon regulatory factor 4 (IRF4), Protein flightless-1 homolog (FLII), Heterogeneous nuclear ribonucleoprotein F (HNRNPF), Nucleus accumbens-associated protein 1 (NACC1), Transcription elongation regulator 1 (TCERG1), Protein HEXIM1 (HEXIM1), Enhancer of mRNA-decapping protein (EDC3), Zinc finger protein Aiolos (IKZF3), Transcription elongation factor SPT5 (SUPT5H), Forkhead box protein K1 (FOXK1), LIM domain-containing protein 1 (LIMD1), MMS19 nucleotide excision repair protein homolog (MMS19), Elongator complex protein 4 (ELP4), Ankyrin repeat and KH domain-containing protein 1 (ANKHD1), PML, Nuclear factor NF-kappa-B p100 subunit (NFKB2), Heterogeneous nuclear ribonucleoprotein L-like (HNRPLL), CCR4-NOT transcription complex subunit 3 (CNOT3), Constitutive coactivator of PPAR-gamma-like protein (FAM120A), Mediator of RNA polymerase II transcription subunit (MED15), 60S ribosomal protein L7 (RPL7), Interferon regulatory factor 8 (IRF8), COUP transcription factor 2 (NR2F2), Mediator of RNA polymerase II transcription subunit (MED1), tRNA (uracil-5-)-methyltransferase homolog A (TRMT2A), Transcription factor p65 (RELA), Exosome complex component RRP42 (EXOSC7), General transcription factor 3C polypeptide 1 (GTF3C1), Mothers against decapentaplegic homolog 2 (SMAD2), Ankyrin repeat domain-containing protein 17 (ANKRD17), MMS19 nucleotide excision repair protein homolog (MMS19), Death domain-associated protein 6 (DAXX), Zinc finger protein 318 (ZNF318), Thioredoxin-interacting protein (TXNIP), Glucocorticoid receptor (NR3C1), Iron-responsive element-binding protein 2 (IREB2), Zinc finger protein 295 (ZNF295), Polycomb protein SUZ12 (SUZ12), Cleavage stimulation factor subunit 2 tau variant (CSTF2T), C-myc promoter-binding protein (DENND4A), Pinin (PNN), Mediator of RNA polymerase II transcription subunit (MED9),

POU domain, class 2, transcription factor 2 (POU2F2), Enhancer of mRNA-decapping protein 3 (EDC3), A-kinase anchor protein 1 (mitochondrial) (AKAP1), Transcription factor RelB (RELB), RNA polymerase II-associated protein 1 (RPAP1), Zinc finger protein 346 (ZNF346), Chromosome-associated kinesin KIF4A (KIF4A), Mediator of RNA polymerase II transcription subunit (MED12), Protein NPAT (NPAT), Leucine-rich PPR motif-containing protein (mitochondrial) (LRPPRC), AT-hook DNA-binding motif-containing protein 1 (AHDC1), Mediator of RNA polymerase II transcription subunit (MED12), Bromodomain-containing protein 8 (BRD8), Trinucleotide repeat-containing gene 6B protein (TNRC6B), Aryl hydrocarbon receptor nuclear translocator (ARNT), Activating transcription factor 7-interacting protein (ATF7IP), Glucocorticoid receptor (NR3C1), Chromosome transmission fidelity protein 18 homolog (CHTF18), or C-myc promoter-binding protein (DENND4A).

[00146] In some embodiments, the cysteine containing protein is a channel, transporter or receptor. Exemplary cysteine containing proteins as channels, transporters, or receptors include, but are not limited to, Chloride intracellular channel protein 4 (CLIC4), Exportin-1 (XPO1), Thioredoxin (TXN), Protein SEC13 homolog (SEC13), Chloride intracellular channel protein 1 (CLIC1), Guanine nucleotide-binding protein subunit beta-2 (GNB2L1), Sorting nexin-6 (SNX6), Conserved oligomeric Golgi complex subunit 3 (COG3), Nuclear cap-binding protein subunit 1 (NCBP1), Cytoplasmic dynein 1 light intermediate chain 1 (DYNC1LI1), MOB-like protein phocein (MOB4), Programmed cell death 6-interacting protein (PDCD6IP), Glutaredoxin-1 (GLRX), ATP synthase subunit alpha (mitochondrial) (ATP5A1), Treacle protein (TCOF1), Dynactin subunit 1 (DCTN1), Importin-7 (IPO7), Exportin-2 (CSE1L), ATP synthase subunit gamma (mitochondrial) (ATP5C1), Trafficking protein particle complex subunit 5 (TRAPPC5), Thioredoxin mitochondrial (TXN2), THO complex subunit 6 homolog (THOC6), Exportin-1 (XPO1), Nuclear pore complex protein Nup50 (NUP50), Treacle protein (TCOF1), Nuclear pore complex protein Nup93 (NUP93), Nuclear pore glycoprotein p62 (NUP62), Cytoplasmic dynein 1 heavy chain 1 (DYNC1H1), Thioredoxin-like protein 1 (TXNL1), Nuclear pore complex protein Nup214 (NUP214), Protein lin-7 homolog C (LIN7C), ADP-ribosylation factor-binding protein GGA2 (GGA2), Trafficking protein particle complex subunit 4 (TRAPPC4), Protein quaking (QKI), Perilipin-3 (PLIN3), Copper transport protein ATOX1 (ATOX1), Unconventional myosin-Ic (MYO1C), Nucleoporin NUP53 (NUP35), Vacuolar protein sorting-associated protein 18 homolog (VPS18), Dedicator of cytokinesis protein 7 (DOCK7), Nucleoporin p54 (NUP54), Ras-related GTP-binding protein C (RRAGC), Arf-GAP with Rho-GAP domain (ANK repeat and PH domain) (ARAP1), Exportin-5 (XPO5), Kinectin (KTN1), Chloride intracellular channel protein 6 (CLIC6), Voltage-gated potassium channel subunit beta-2 (KCNAB2), Exportin-5 (XPO5), Ras-related GTP-binding protein C

(RRAGC), Ribosome-binding protein 1 (RRBP1), Acyl-CoA-binding domain-containing protein 6 (ACBD6), Chloride intracellular channel protein 5 (CLIC5), Pleckstrin homology domain-containing family A member (PLEKHA2), ADP-ribosylation factor-like protein 3 (ARL3), Protein transport protein Sec24C (SEC24C), Voltage-dependent anion-selective channel protein (VDAC3), Programmed cell death 6-interacting protein (PDCD6IP), Chloride intracellular channel protein 3 (CLIC3), Multivesicular body subunit 12A (FAM125A), Eukaryotic translation initiation factor 4E transporter (EIF4ENIF1), NmrA-like family domain-containing protein 1 (NMRAL1), Nuclear pore complex protein Nup98-Nup96 (NUP98), Conserved oligomeric Golgi complex subunit 1 (COG1), Importin-4 (IPO4), Pleckstrin homology domain-containing family A member (PLEKHA2), Cytoplasmic dynein 1 heavy chain 1 (DYNC1H1), DENN domain-containing protein 1C (DENND1C), Cytoplasmic dynein 1 heavy chain 1 (DYNC1H1), Protein ELYS (AHCTF1), Trafficking protein particle complex subunit 1 (TRAPPC1), Guanine nucleotide-binding protein-like 3 (GNL3), or Importin-13 (IPO13).

[00147] In some embodiments, the cysteine containing protein is a chaperone. Exemplary cysteine containing proteins as chaperones include, but are not limited to, 60 kDa heat shock protein (mitochondrial) (HSPD1), T-complex protein 1 subunit eta (CCT7), T-complex protein 1 subunit epsilon (CCT5), Heat shock 70 kDa protein 4 (HSPA4), GrpE protein homolog 1 (mitochondrial) (GRPEL1), Tubulin-specific chaperone E (TBCE), Protein unc-45 homolog A (UNC45A), Serpin H1 (SERPINH1), Tubulin-specific chaperone D (TBCD), Peroxisomal biogenesis factor 19 (PEX19), BAG family molecular chaperone regulator 5 (BAG5), T-complex protein 1 subunit theta (CCT8), Protein canopy homolog 3 (CNPY3), DnaJ homolog subfamily C member 10 (DNAJC10), ATP-dependent Clp protease ATP-binding subunit clp (CLPX), or Midasin (MDN1).

[00148] In some embodiments, the cysteine containing protein is an adapter, scaffolding or modulator protein. Exemplary cysteine containing proteins as adapter, scaffolding, or modulator proteins include, but are not limited to, Proteasome activator complex subunit 1 (PSME1), TIP41-like protein (TIPRL), Crk-like protein (CRKL), Cofilin-1 (CFL1), Condensin complex subunit 1 (NCAPD2), Translational activator GCN1 (GCN1L1), Serine/threonine-protein phosphatase 2A 56 kDa regulatory (PPP2R5D), UPF0539 protein C7orf59 (C7orf59), Protein diaphanous homolog 1 (DIAPH1), Protein asunder homolog (Asun), Ras GTPase-activating-like protein IQGAP1 (IQGAP1), Sister chromatid cohesion protein PDS5 homolog A (PDS5A), Reticulon-4 (RTN4), Proteasome activator complex subunit 4 (PSME4), Condensin complex subunit 2 (NCAPH), Sister chromatid cohesion protein PDS5 homolog A (PDS5A), cAMP-dependent protein kinase type I-alpha regulatory (PRKAR1A), Host cell factor 1 (HCFC1), Serine/threonine-protein phosphatase 4 regulatory (PPP4R2), Apoptotic chromatin condensation

inducer in the nucleus (ACIN1), BRISC and BRCA1-A complex member 1 (BABAM1), Interferon-induced protein with tetratricopeptide (IFIT3), Ras association domain-containing protein 2 (RASSF2), Hsp70-binding protein 1 (HSPBP1), TBC1 domain family member 15 (TBC1D15), Dynamin-binding protein (DNMBP), Condensin complex subunit 1 (NCAPD2), Beta-2-syntrophin (SNTB2), Disks large homolog 1 (DLG1), TBC1 domain family member 13 (TBC1D13), Formin-binding protein 1-like (FNBP1L), Translational activator GCN1 (GCN1L1), GRB2-related adapter protein (GRAP), G2/mitotic-specific cyclin-B1 (CCNB1), Myotubularin-related protein 12 (MTMR12), Protein FADD (FADD), Translational activator GCN1 (GCN1L1), Wings apart-like protein homolog (WAPAL), cAMP-dependent protein kinase type II-beta regulatory (PRKAR2B), Malcavernin (CCM2), MPP1 55 kDa erythrocyte membrane protein, Actin filament-associated protein 1 (AFAP1), Tensin-3 (TNS3), tRNA methyltransferase 112 homolog (TRMT112), Symplekin (SYMPK), TBC1 domain family member 2A (TBC1D2), ATR-interacting protein (ATRIP), Ataxin-10 (ATXN10), Succinate dehydrogenase assembly factor 2 (mitochondrial) (SDHAF2), Formin-binding protein 1 (FNBP1), Myotubularin-related protein 12 (MTMR12), Interferon-induced protein with tetratricopeptide (IFIT3), Protein CBFA2T2 (CBFA2T2), Neutrophil cytosol factor 1 (NCF1), or Protein syndesmos (NUDT16L1).

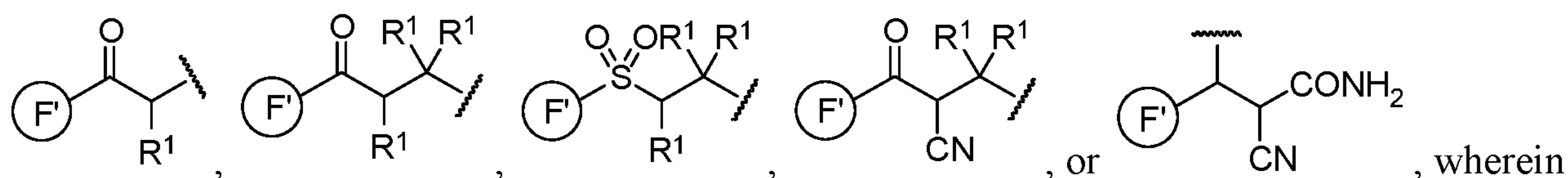
[00149] In some embodiments, a cysteine containing protein comprises a protein illustrated in Tables 1-5 or Tables 7-9. In some instances, a cysteine containing protein comprises a protein illustrated in Table 1. In some embodiments, the cysteine containing protein comprises a cysteine residue denoted in Table 1. In some instances, a cysteine containing protein comprises a protein illustrated in Table 2. In some embodiments, the cysteine containing protein comprises a cysteine residue denoted in Table 2. In some instances, a cysteine containing protein comprises a protein illustrated in Table 3. In some embodiments, the cysteine containing protein comprises a cysteine residue denoted in Table 3. In some instances, a cysteine containing protein comprises a protein illustrated in Table 4. In some embodiments, the cysteine containing protein comprises a cysteine residue denoted in Table 4. In some instances, a cysteine containing protein comprises a protein illustrated in Table 5. In some embodiments, the cysteine containing protein comprises a cysteine residue denoted in Table 5. In some instances, a cysteine containing protein comprises a protein illustrated in Table 7. In some embodiments, the cysteine containing protein comprises a cysteine residue denoted in Table 7. In some instances, a cysteine containing protein comprises a protein illustrated in Table 8. In some embodiments, the cysteine containing protein comprises a cysteine residue denoted in Table 8. In some instances, a cysteine containing protein comprises a protein illustrated in Table 9. In some embodiments, the cysteine containing protein comprises a cysteine residue denoted in Table 9. In some instances, the cysteine containing protein is a

modified protein, in which the protein is modified at a cysteine residue site by a small molecule fragment described herein, such as for example, by a small molecule fragment of Formula (I) described herein, a cysteine-reactive probe of Formula (II) described herein, or by a small molecule fragment illustrated in Fig. 3.

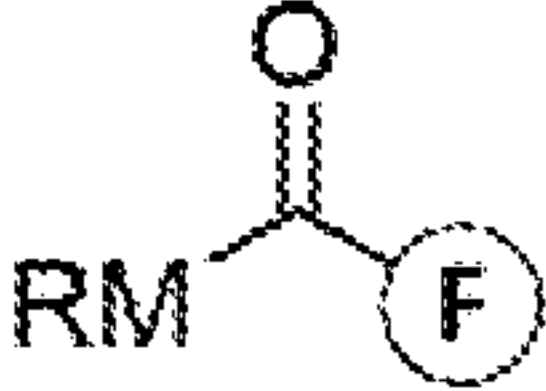
[00150] In some embodiments, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein. In some instances, the cysteine containing protein is selected from Table 3. In some cases, one or more cysteine residues of each respective cysteine containing protein are denoted in Table 3. In some cases, a cysteine containing protein selected from Table 3 is modified by a small molecule fragment at at least one cysteine site denoted in Table 3 to generate a modified cysteine containing protein. In some cases, the cysteine containing protein is selected from AIP, PES1, IKBKB, XPO1, KDM4B, NR3C1, GSTP1, TNFAIP3, ACAT1, IRAK1, GNB2L1, IRF4, USP34, ZC3HAV1, USP7, PELI1, DCUN1D1, USP28, UBE2O, RRAGC, MLTK, USP22, KDM3A, or USP16. In some cases, the cysteine containing protein is selected from AIP, PES1, IKBKB, XPO1, GSTP1, ACAT1, IRAK1, IRF4, ZC3HAV1, USP7, PELI1, USP28, UBE2O, RRAGC, MLTK, USP22, KDM3A, or USP16. In some cases, the cysteine containing protein is selected from KDM4B, NR3C1, TNFAIP3, USP7 or USP22. In some cases, the cysteine containing protein is selected from GNB2L1 or USP34. In some cases, the cysteine containing protein is DCUN1D1. In some cases, the cysteine containing protein is selected from PES1, IKBKB, GSTP1, ACAT1, IRAK1, ZC3HAV1 or RRAGC. In some cases, the cysteine containing protein is selected from XPO1, GNB2L1, USP34, UBE2O, MLTK or USP22. In some cases, the cysteine containing protein is selected from KDM4B or NR3C1. In some cases, the cysteine containing protein is selected from TNFAIP3, USP7, USP28, KDM3A or USP16. In some cases, the cysteine containing protein is selected from IRF4, PELI1, DCUN1D1 or USP22. In some cases, the cysteine containing protein is AIP. In some cases, the cysteine containing protein is an enzyme and the enzyme is selected from IKBKB, KDM4B, GSTP1, TNFAIP3, ACAT1, IRAK1, USP34, USP7, PELI1, USP28, UBE2O, MLTK, USP22, KDM3A, or USP16. In some cases, the cysteine containing protein is a transcription factor or regulator and the transcription factor or regulator is selected from NR3C1, IRF4 or ZC3HAV1. In some cases, the cysteine containing protein is a channel, a transporter, or a receptor and the channel, transporter, or receptor is selected from GNB2L1 or RRAGC. In some cases, the cysteine containing protein is selected from AIP, PES1, XPO1 or DCUN1D1. In some cases, the cysteine containing protein is selected from PES1, CYR61, UBE2L6, XPO1, ADA, NR3C1, POU2F2, UCHL3, MGMT, ERCC3, ACAT1, STAT3, UBA7, CASP2, IDH2, LRBA, UBE2L3, RELB, IRF8, CASP8, PDIA6, PCK2, PFKFB4, PDE12, USP34, USP48, SMARCC2 or

SAMHD1. In some cases, the cysteine containing protein is selected from PES1, CYR61, NR3C1, UCHL3, ERCC3, ACAT1, STAT3, CASP2, LRBA, UBE2L3, RELB, PDIA6, PCK2, PFKFB4, USP48 or SMARCC2. In some cases, the cysteine containing protein is selected from UBE2L6, POU2F2, MGMT, ACAT1, UBA7, CASP8, PDE12 or USP34. In some cases, the cysteine containing protein is selected from CYR61 or XPO1. In some cases, the cysteine containing protein is selected from ADA, MGMT, IDH2, IRF8 or SAMHD1. In some cases, the cysteine containing protein is selected from PES1, CYR61, XPO1, NR3C1 or SMARCC2. In some cases, the cysteine containing protein is selected from CYR61, UBE2L6, MGMT, ERCC3, ACAT1 or USP48. In some cases, the cysteine containing protein is selected from ADA, RELB or USP34. In some cases, the cysteine containing protein is selected from UCHL3, CASP2, IDH2, LRBA, CASP8, PCK2 or PDE12. In some cases, the cysteine containing protein is selected from MGMT, ACAT1, UBA7, UBE2L3 or IRF8. In some cases, the cysteine containing protein is selected from PFKFB4, ACAT1 or STAT3. In some cases, the cysteine containing protein is selected from POU2F2, PDIA6 or SAMHD1. In some cases, the cysteine containing protein is an enzyme and the enzyme is selected from UBE2L6, ADA, UCHL3, MGMT, ERCC3, ACAT1, UBA7, CASP2, IDH2, UBE2L3, CASP8, PDIA6, PCK2, PFKFB4, PDE12, USP34, USP48 or SAMHD1. In some cases, the cysteine containing protein is a transcription factor or a regulator and the transcription factor or regulator is selected from NR3C1, POU2F2, STAT3, RELB, IRF8 or SMARCC2. In some cases, the cysteine containing protein is selected from ZAP70, PRKCQ or PRMT1. In some cases, the cysteine containing protein is selected from ZAP70 or PRKCQ. In some cases, the cysteine containing protein is selected from CYR61, ZNF217, NCF1, IREB2, LRBA, CDK5, EP300, EZH2, UBE2S, VCPIP1, RRAGC or IRAK4. In some cases, the cysteine containing protein is selected from CYR61, ZNF217, IREB2, EP300, UBE2S, VCPIP1, RRAGC or IRAK4. In some cases, the cysteine containing protein is selected from NCF1, LRBA or CDK5. In some cases, the cysteine containing protein is EZH2. In some cases, the cysteine containing protein is selected from ZNF217, NCF1, CDK5, EP300 or IRAK4. In some cases, the cysteine containing protein is selected from CYR61, IREB2, LRBA or UBE2S. In some cases, the cysteine containing protein is selected from EZH2, VCPIP1 or RRAGC. In some cases, the cysteine containing protein is an enzyme and the enzyme is selected from CDK5, EP300, EZH2, UBE2S, VCPIP1 or IRAK4. In some cases, the cysteine containing protein is a transcription factor or a regulator and the transcription factor or regulator is selected from ZNF217 or IREB2. In some cases, the cysteine containing protein is an adapter, a scaffolding protein or a modulator protein and the adapter, scaffolding protein or the modulator protein is selected from NCF1. In some cases, the cysteine containing protein is a channel, a transporter or a receptor and the channel, transporter, or

receptor is selected from RRAGC. In some cases, the cysteine containing protein is selected from CYR61 or LRBA. In some cases, the cysteine containing protein is about 20, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000 amino acid residues in length or more. In some cases, the cysteine residue of the modified cysteine containing protein has the structure SR, wherein R is selected from:



R^1 is H, C1-C3 alkyl, or aryl; and F' is the small molecule fragment moiety. In some cases, the small molecule fragment has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some cases, the molecular weight of the small molecule fragment is prior to enrichment with a halogen, a nonmetal, or a transition metal. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur atoms. In some embodiments, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some cases, the small molecule fragment is a small

molecule fragment of Formula (I):  wherein RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and F is a small molecule fragment moiety. In some cases, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some cases, F is obtained from a compound library. In some cases, F is a small molecule fragment moiety illustrated in Fig. 3. In some cases, F further comprises a linker moiety that connects F to the carbonyl moiety. In some cases, the small molecule fragment binds irreversibly to the cysteine containing protein. In some cases, the small molecule fragment binds reversibly to the cysteine containing protein.

[00151] In some embodiments, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein is selected from Table 10A, enzymes. In some cases, one or more cysteine residues of each respective cysteine containing protein are denoted in Table 10A. In some cases, a cysteine containing protein selected from Table 10A is modified by a small molecule fragment at at least one cysteine site denoted in Table 10A to generate a modified cysteine containing protein. In some cases, the cysteine

containing protein is about 20, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000 amino acid residues in length or more. In some cases, the cysteine residue of the modified cysteine containing protein has the

structure SR, wherein R is selected from:

, wherein R^1 is H, C1-C3 alkyl, or aryl; and F' is the small molecule fragment moiety. In some cases, the small molecule fragment has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some cases, the molecular weight of the small molecule fragment is prior to enrichment with a halogen, a nonmetal, or a transition metal. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur atoms. In some embodiments, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some

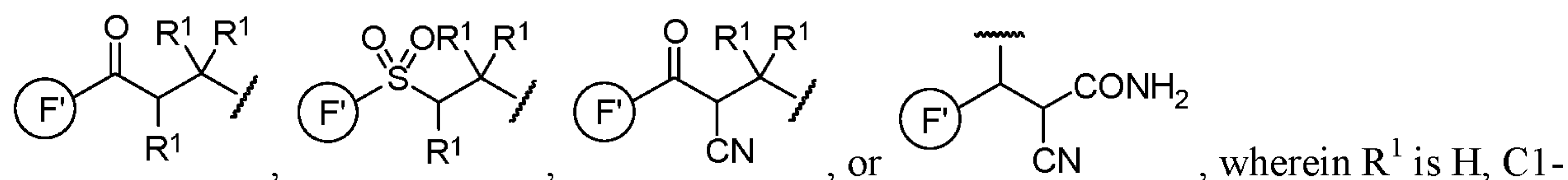
cases, the small molecule fragment is a small molecule fragment of Formula (I):

wherein RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and F is a small molecule fragment moiety. In some cases, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some cases, F is obtained from a compound library. In some cases, F is a small molecule fragment moiety illustrated in Fig. 3. In some cases, F further comprises a linker moiety that connects F to the carbonyl moiety. In some cases, the small molecule fragment binds irreversibly to the cysteine containing protein. In some cases, the small molecule fragment binds reversibly to the cysteine containing protein.

[00152] In some embodiments, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein is selected from Table 10B, transcription factors and regulators. In some cases, one or more cysteine residues of each respective cysteine containing protein are denoted in Table 10B. In some cases, a cysteine containing protein selected from Table 10B is modified by a small molecule fragment at at least one cysteine site denoted in Table 10B to generate a modified cysteine containing protein. In

some cases, the cysteine containing protein is about 20, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000 amino acid residues in length or more. In some cases, the cysteine residue of the modified cysteine

containing protein has the structure SR, wherein R is selected from:



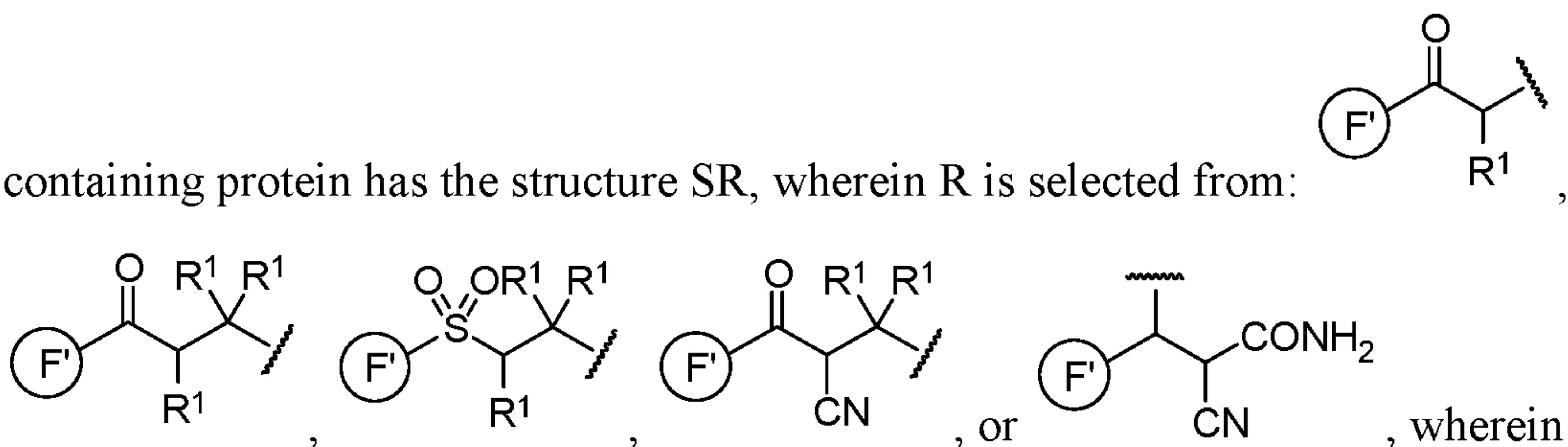
wherein R^1 is H, C1-C3 alkyl, or aryl; and F' is the small molecule fragment moiety. In some cases, the small molecule fragment has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some cases, the molecular weight of the small molecule fragment is prior to enrichment with a halogen, a nonmetal, or a transition metal. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur atoms. In some embodiments, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some cases, the small molecule fragment is a small molecule

fragment of Formula (I): wherein RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and F is a small molecule fragment moiety. In some cases, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some cases, F is obtained from a compound library. In some cases, F is a small molecule fragment moiety illustrated in Fig. 3. In some cases, F further comprises a linker moiety that connects F to the carbonyl moiety. In some cases, the small molecule fragment binds irreversibly to the cysteine containing protein. In some cases, the small molecule fragment binds reversibly to the cysteine containing protein.

[00153] In some embodiments, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein is selected from Table 10C, channels, transporters or receptors. In some cases, one or more cysteine residues of each respective cysteine containing protein are denoted in Table 10C. In some cases, a cysteine containing protein selected from Table 10C is modified by a small molecule fragment at at least

one cysteine site denoted in Table 10C to generate a modified cysteine containing protein. In some cases, the cysteine containing protein is about 20, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000 amino acid residues in length or more. In some cases, the cysteine residue of the modified cysteine

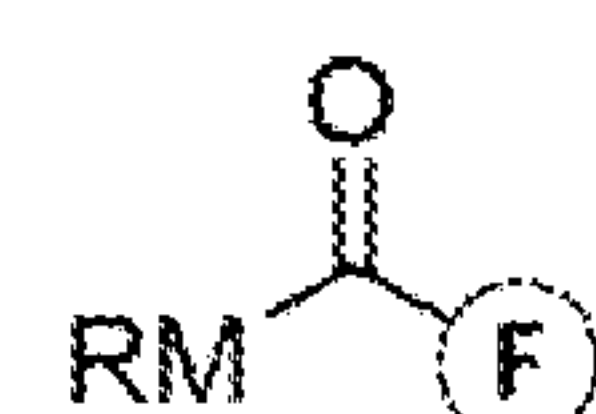
containing protein has the structure SR, wherein R is selected from:



, wherein R¹ is H, C1-C3 alkyl, or aryl; and F' is the small molecule fragment moiety. In some cases, the small

molecule fragment has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some cases, the molecular weight of the small molecule fragment is prior to enrichment with a halogen, a nonmetal, or a transition metal. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur atoms. In some embodiments, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some cases, the small molecule fragment is a small molecule

fragment of Formula (I):

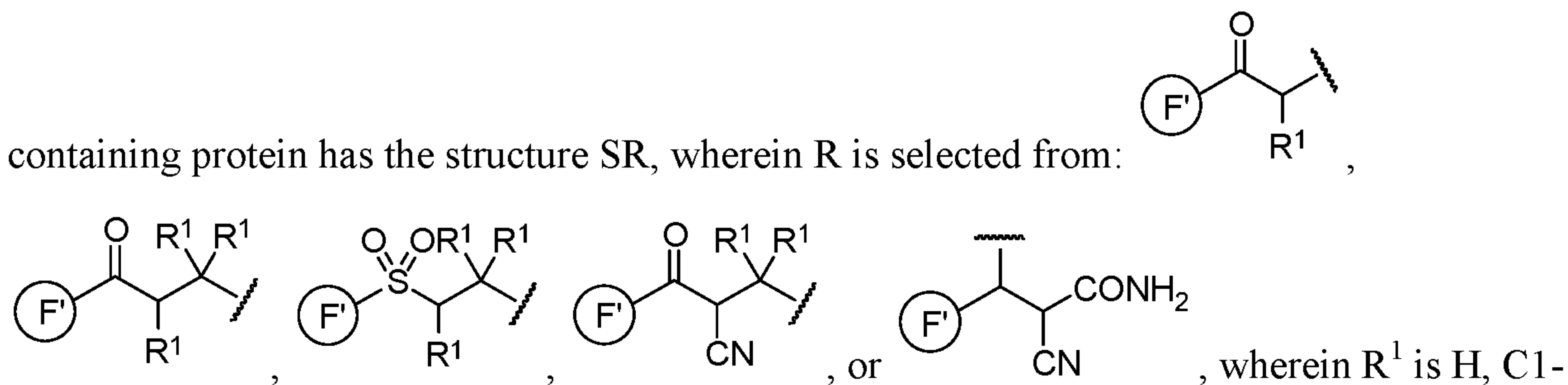


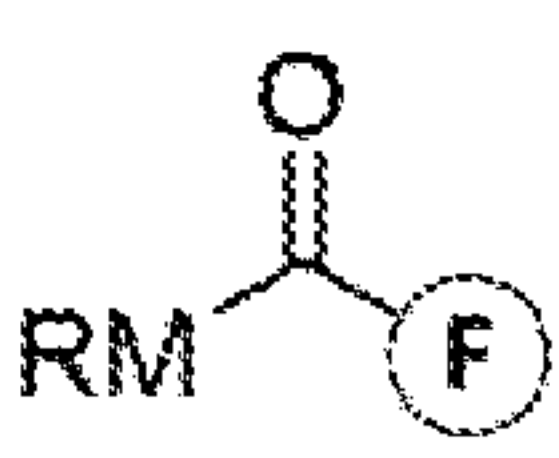
wherein RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and F is a small molecule fragment moiety. In some cases, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some cases, F is obtained from a compound library. In some cases, F is a small molecule fragment moiety illustrated in Fig. 3. In some cases, F further comprises a linker moiety that connects F to the carbonyl moiety. In some cases, the small molecule fragment binds irreversibly to the cysteine containing protein. In some cases, the small molecule fragment binds reversibly to the cysteine containing protein.

[00154] In some embodiments, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein is selected from Table 10D, adapter, scaffolding, or modulator proteins. In some cases, one or more cysteine residues of each respective cysteine containing protein are denoted in Table 10D. In some cases, a cysteine

containing protein selected from Table 10D is modified by a small molecule fragment at at least one cysteine site denoted in Table 10D to generate a modified cysteine containing protein. In some cases, the cysteine containing protein is about 20, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000 amino acid residues in length or more. In some cases, the cysteine residue of the modified cysteine

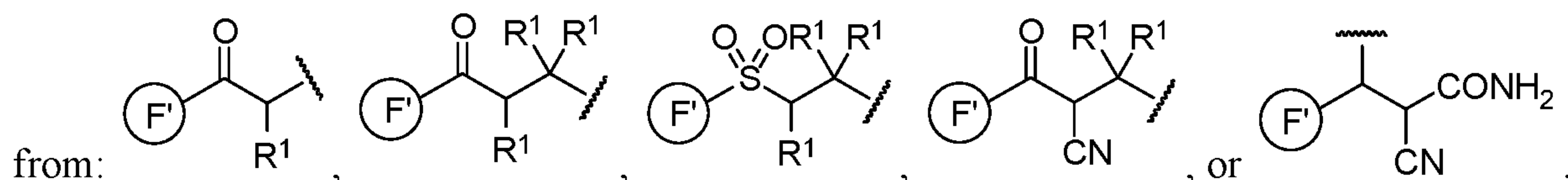
containing protein has the structure SR, wherein R is selected from:



fragment of Formula (I):  wherein RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and F is a small molecule fragment moiety. In some cases, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some cases, F is obtained from a compound library. In some cases, F is a small molecule fragment moiety illustrated in Fig. 3. In some cases, F further comprises a linker moiety that connects F to the carbonyl moiety. In some cases, the small molecule fragment binds irreversibly to the cysteine containing protein. In some cases, the small molecule fragment binds reversibly to the cysteine containing protein.

[00155] In some embodiments, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein is selected from Table 10E. In some cases, one or more cysteine residues of each respective cysteine containing protein are denoted

in Table 10E. In some cases, a cysteine containing protein selected from Table 10E is modified by a small molecule fragment at at least one cysteine site denoted in Table 10E to generate a modified cysteine containing protein. In some cases, the cysteine containing protein is about 20, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000 amino acid residues in length or more. In some cases, the cysteine residue of the modified cysteine containing protein has the structure SR, wherein R is selected



wherein R^1 is H, C1-C3 alkyl, or aryl; and F' is the small molecule fragment moiety. In some cases, the small molecule fragment has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some cases, the molecular weight of the small molecule fragment is prior to enrichment with a halogen, a nonmetal, or a transition metal. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur atoms. In some embodiments, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some cases, the small molecule fragment is a small

molecule fragment of Formula (I): wherein RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and F is a small molecule fragment moiety. In some cases, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some cases, F is obtained from a compound library. In some cases, F is a small molecule fragment moiety illustrated in Fig. 3. In some cases, F further comprises a linker moiety that connects F to the carbonyl moiety. In some cases, the small molecule fragment binds irreversibly to the cysteine containing protein. In some cases, the small molecule fragment binds reversibly to the cysteine containing protein.

[00156] In some embodiments, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif X_pC^*Z , wherein X_p is a polar residue, C^* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from AIP, PES1, IKBKB, XPO1, KDM4B, NR3C1, GSTP1,

TNFAIP3, ACAT1, IRAK1, GNB2L1, IRF4, USP34, ZC3HAV1, USP7, PELI1, DCUN1D1, USP28, UBE2O, RRAGC, MLTK, USP22, KDM3A, or USP16.

[00157] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif $X_pC^*X_n$, wherein X_p is a polar residue, C^* denotes the site of modification, and X_n is a nonpolar residue. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from AIP, PES1, IKBKB, XPO1, GSTP1, ACAT1, IRAK1, IRF4, ZC3HAV1, USP7, PELI1, USP28, UBE2O, RRAGC, MLTK, USP22, KDM3A, or USP16.

[00158] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif $X_pC^*X_p$, wherein X_p is a polar residue and C^* denotes the site of modification. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from KDM4B, NR3C1, TNFAIP3, USP7 or USP22.

[00159] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif $X_pC^*X_b$, wherein X_p is a polar residue, C^* denotes the site of modification, and X_b is a basic residue. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from GNB2L1 or USP34.

[00160] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif $X_pC^*X_a$, wherein X_p is a polar residue, C^* denotes the site of modification, and X_a is an acidic residue. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is DCUN1D1.

[00161] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif SC^*Z , wherein C^* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from PES1, IKBKB, GSTP1, ACAT1, IRAK1, ZC3HAV1 or RRAGC.

[00162] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif NC*Z, wherein C* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from XPO1, GNB2L1, USP34, UBE2O, MLTK or USP22.

[00163] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif YC*Z, wherein C* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from KDM4B or NR3C1.

[00164] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif TC*Z, wherein C* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from TNFAIP3, USP7, USP28, KDM3A or USP16.

[00165] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif QC*Z, wherein C* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from IRF4, PELI1, DCUN1D1 or USP22.

[00166] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif CC*Z, wherein C* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is AIP.

[00167] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein is an enzyme and the enzyme comprises the motif X_pC*Z, wherein X_p is a polar residue, C* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3.

In some cases, the enzyme is selected from IKBKB, KDM4B, GSTP1, TNFAIP3, ACAT1, IRAK1, USP34, USP7, PELI1, USP28, UBE2O, MLTK, USP22, KDM3A, or USP16.

[00168] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein is a transcription factor or a regulator and the transcription factor or regulator comprises the motif X_pC^*Z , wherein X_p is a polar residue, C^* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the transcription factor or regulator is selected from NR3C1, IRF4 or ZC3HAV1.

[00169] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein is a channel, transporter or a receptor and the channel, transporter or receptor comprises the motif X_pC^*Z , wherein X_p is a polar residue, C^* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the channel, transporter, or receptor is selected from GNB2L1 or RRAGC.

[00170] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif X_pC^*Z , wherein X_p is a polar residue, C^* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from AIP, PES1, XPO1 or DCUN1D1.

[00171] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif X_nC^*Z , wherein X_n is a nonpolar residue, C^* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from PES1, CYR61, UBE2L6, XPO1, ADA, NR3C1, POU2F2, UCHL3, MGMT, ERCC3, ACAT1, STAT3, UBA7, CASP2, IDH2, LRBA, UBE2L3, RELB, IRF8, CASP8, PDIA6, PCK2, PFKFB4, PDE12, USP34, USP48, SMARCC2 or SAMHD1.

[00172] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif $X_nC^*X_n$, wherein X_n is a nonpolar residue and C^* denotes the site of modification. In some cases, the

cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from PES1, CYR61, NR3C1, UCHL3, ERCC3, ACAT1, STAT3, CASP2, LRBA, UBE2L3, RELB, PDIA6, PCK2, PFKFB4, USP48 or SMARCC2.

[00173] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif $X_nC^*X_p$, wherein X_n is a nonpolar residue, C^* denotes the site of modification, and X_p is a polar residue. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from UBE2L6, POU2F2, MGMT, ACAT1, UBA7, CASP8, PDE12 or USP34.

[00174] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif $X_nC^*X_a$, wherein X_n is a nonpolar residue, C^* denotes the site of modification, and X_a is an acidic residue. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from CYR61 or XPO1.

[00175] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif $X_nC^*X_b$, wherein X_n is a nonpolar residue, C^* denotes the site of modification, and X_b is a basic residue. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from ADA, MGMT, IDH2, IRF8 or SAMHD1.

[00176] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif LC^*Z , wherein C^* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from PES1, CYR61, XPO1, NR3C1 or SMARCC2.

[00177] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif PC^*Z , wherein C^* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from CYR61, UBE2L6, MGMT, ERCC3, ACAT1 or USP48.

[00178] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif GC*Z, wherein C* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from ADA, RELB or USP34.

[00179] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif AC*Z, wherein C* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from UCHL3, CASP2, IDH2, LRBA, CASP8, PCK2 or PDE12.

[00180] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif VC*Z, wherein C* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from MGMT, ACAT1, UBA7, UBE2L3 or IRF8.

[00181] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif IC*Z, wherein C* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from PFKFB4, ACAT1 or STAT3.

[00182] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif X_rC*Z, wherein X_r denotes an aromatic residue, C* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from POU2F2, PDIA6 or SAMHD1.

[00183] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein is an enzyme and the enzyme comprises the motif X_nC*Z, wherein X_n is a nonpolar residue, C* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected

from Table 3. In some cases, the enzyme is selected from UBE2L6, ADA, UCHL3, MGMT, ERCC3, ACAT1, UBA7, CASP2, IDH2, UBE2L3, CASP8, PDIA6, PCK2, PFKFB4, PDE12, USP34, USP48 or SAMHD1.

[00184] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein is a transcription factor or a regulator and the transcription factor or regulator comprises the motif X_nC^*Z , wherein X_n is a nonpolar residue, C^* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the transcription factor or regulator is selected from NR3C1, POU2F2, STAT3, RELB, IRF8 or SMARCC2.

[00185] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif X_nC^*Z , wherein X_n is a nonpolar residue, C^* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from PES1, CYR61, XPO1 or LRBA.

[00186] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif X_aC^*Z , wherein X_a is an acidic residue, C^* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from ZAP70, PRKCQ or PRMT1.

[00187] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif EC^*Z , wherein C^* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from ZAP70 or PRKCQ.

[00188] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif X_bC^*Z , wherein X_b is a basic residue, C^* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from CYR61, ZNF217, NCF1, IREB2, LRBA, CDK5, EP300, EZH2, UBE2S, VCPIP1, RRAGC or IRAK4.

[00189] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif $X_bC^*X_n$, wherein X_b is a basic residue, C^* denotes the site of modification, and X_n is a nonpolar residue. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from CYR61, ZNF217, IREB2, EP300, UBE2S, VCPIP1, RRAGC or IRAK4.

[00190] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif $X_bC^*X_p$, wherein X_b is a basic residue, C^* denotes the site of modification, and X_p is a polar residue. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from NCF1, LRBA or CDK5.

[00191] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif $X_bC^*X_b$, wherein X_b is a basic residue and C^* denotes the site of modification. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is EZH2.

[00192] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif RC^*Z , wherein C^* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from ZNF217, NCF1, CDK5, EP300 or IRAK4.

[00193] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif KC^*Z , wherein C^* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from CYR61, IREB2, LRBA or UBE2S.

[00194] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif HC^*Z , wherein C^* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine

containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from EZH2, VCPIP1 or RRAGC.

[00195] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein is an enzyme and the enzyme comprises the motif X_bC^*Z , wherein X_b is a basic residue, C^* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the enzyme is selected from CDK5, EP300, EZH2, UBE2S, VCPIP1 or IRAK4.

[00196] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein is a transcription factor or a regulator and the transcription factor or regulator comprises the motif X_bC^*Z , wherein X_b is a basic residue, C^* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the transcription factor or regulator is selected from ZNF217 or IREB2.

[00197] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein is an adapter, a scaffolding protein, or a modulator protein and the adapter, scaffolding protein or the modulator protein comprises the motif X_bC^*Z , wherein X_b is a basic residue, C^* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the adapter, scaffolding protein or the modulator protein is selected from NCF1.

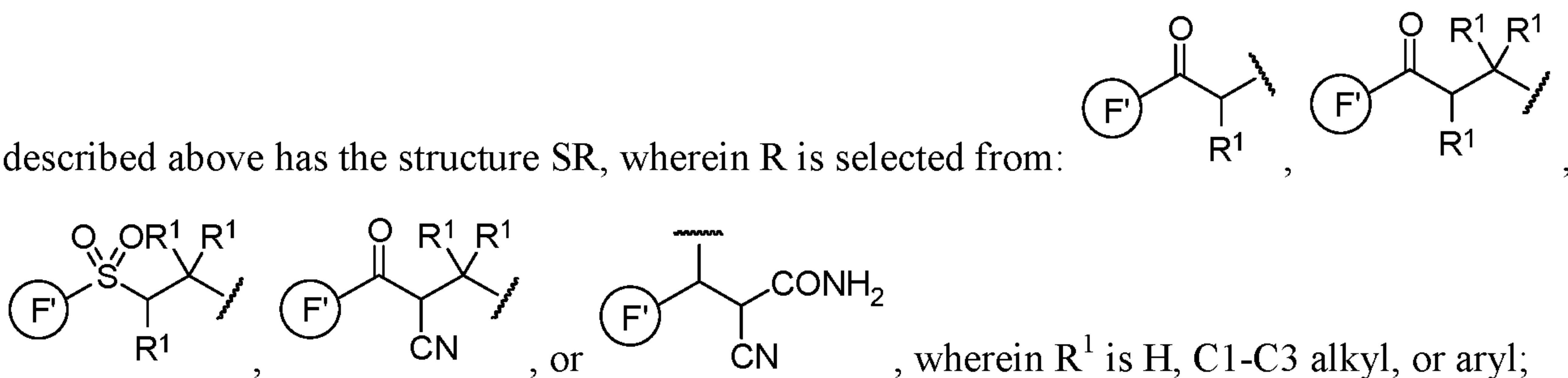
[00198] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein is a channel, a transporter, or a receptor and the channel, transporter, or receptor comprises the motif X_bC^*Z , wherein X_b is a basic residue, C^* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the channel, transporter, or receptor is selected from RRAGC.

[00199] In some instances, described herein is a modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, in which the cysteine containing protein comprises the motif X_bC^*Z , wherein X_b is a basic residue, C^* denotes the site of modification, and Z is any amino acid. In some cases, the cysteine containing protein is selected from Table 3. In some cases, the cysteine containing protein is selected from CYR61 or LRBA.

[00200] In some cases, a cysteine containing protein described above comprises about 20, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000 amino acid residues in length or more.

[00201] In some cases, the cysteine residue of a modified cysteine containing protein

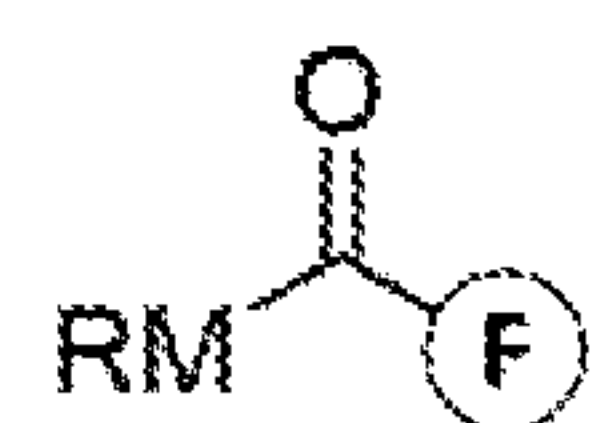
described above has the structure SR, wherein R is selected from:



The structures shown are: $\text{F}'-\text{C}(=\text{O})-\text{CH}(\text{R}^1)-$, $\text{F}'-\text{C}(=\text{O})-\text{CH}(\text{R}^1)-\text{C}(\text{R}^1)_2-$, $\text{F}'-\text{S}(=\text{O})(\text{OR}^1)-\text{CH}(\text{R}^1)-\text{C}(\text{R}^1)_2-$, $\text{F}'-\text{C}(=\text{O})-\text{CH}(\text{CN})-\text{C}(\text{R}^1)_2-$, or $\text{F}'-\text{CH}(\text{CN})-\text{CONH}_2$, wherein R^1 is H, C1-C3 alkyl, or aryl; and F' is the small molecule fragment moiety.

In some cases, the small molecule fragment has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher. In some cases, the molecular weight of the small molecule fragment is prior to enrichment with a halogen, a nonmetal, or a transition metal. In some embodiments, the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur atoms. In some embodiments, the molecular weight of the small molecule fragment does not include the molecular weight of a halogen, a transition metal or a combination thereof. In some

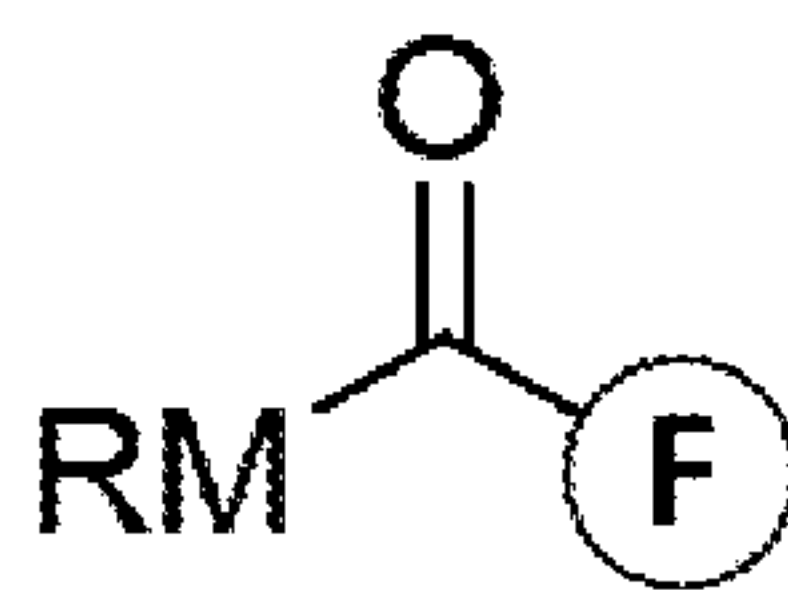
cases, the small molecule fragment is a small molecule fragment of Formula (I):



wherein RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and F is a small molecule fragment moiety. In some cases, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some cases, F is obtained from a compound library. In some cases, F is a small molecule fragment moiety illustrated in Fig. 3. In some cases, F further comprises a linker moiety that connects F to the carbonyl moiety. In some cases, the small molecule fragment binds irreversibly to a cysteine containing protein described above. In some cases, the small molecule fragment binds reversibly to a cysteine containing protein described above.

Compositions, Cells, and Cell Populations

[00202] Disclosed herein also include compositions of a small molecule fragment conjugated to a cysteine containing protein, a cysteine-reactive probe conjugated to a cysteine containing protein, and treated sample compositions. In some embodiments, a composition described herein comprises a small molecule fragment of Formula (I):



Formula (I)

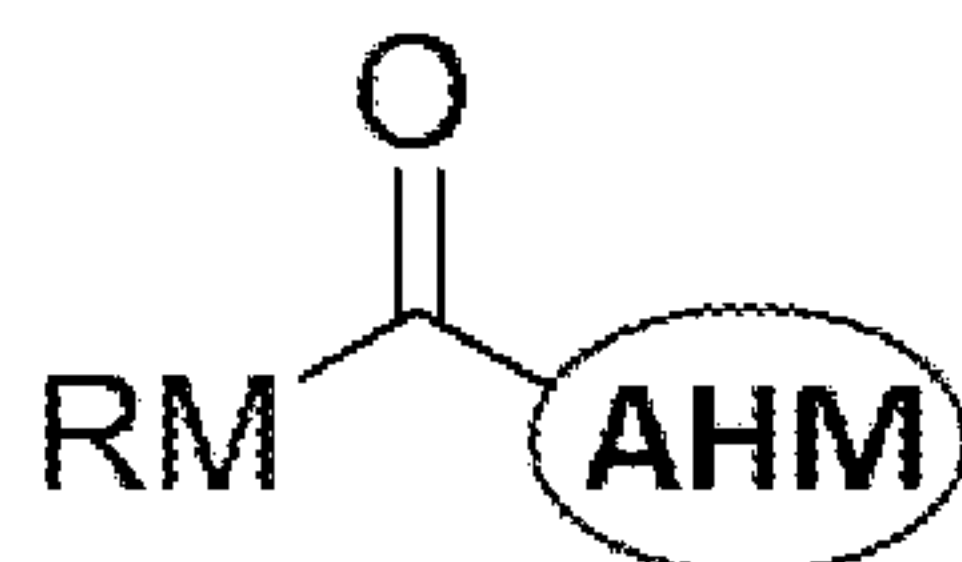
wherein:

RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and

F is a small molecule fragment moiety; and

a cysteine containing protein wherein the cysteine containing protein is covalently bond to the small molecule fragment.

[00203] In some embodiments, also described herein is a composition that comprises a cysteine-reactive probe of Formula (II):



Formula (II)

wherein:

RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond to the thiol group of a cysteine residue; and

AHM is an affinity handle moiety; and

a cysteine containing protein wherein the cysteine containing protein is covalently bond to the cysteine-reactive probe.

[00204] In some embodiments, also described herein is a composition that comprises an isolated sample wherein the isolated sample is an isolated cell or a tissue sample; and a cysteine-reactive probe to be assayed for its ability to interact with a cysteine containing protein expressed in the isolated sample.

[00205] Disclosed herein further include isolated treated cell and cell populations. In some embodiments, described herein is an isolated treated cell that comprises a cysteine-reactive probe covalently attached to a cysteine containing protein. In some instances, the isolated treated cell further comprises a set of cysteine-reactive probes wherein each of the cysteine-reactive probes is covalently attached to a cysteine containing protein.

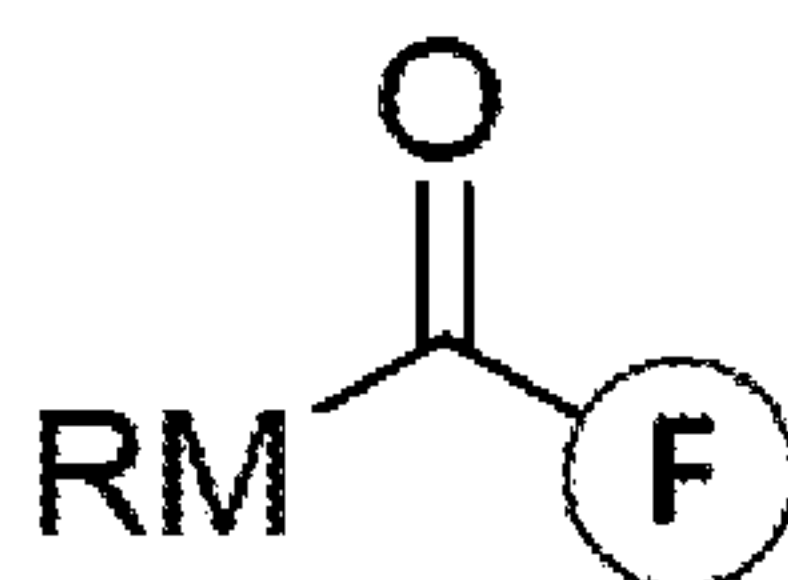
[00206] In some embodiments, described herein is an isolated treated cell that comprises a small molecule fragment covalently attached to a cysteine containing protein. In some instances, the isolated treated cell further comprises a set of small molecule fragments wherein each of the

small molecule fragment is covalently attached to a cysteine containing protein. In some instances, the isolated treated cell further comprises a cysteine-reactive probe. In some instances, the isolated treated cell further comprises a set of cysteine-reactive probes.

[00207] In some embodiments, also described herein is an isolated treated population of cells that comprises a set of cysteine-reactive probes covalently attached to cysteine containing proteins.

[00208] In some embodiments, further described herein is an isolated treated population of cells that comprises a set of small molecule fragments covalently attached to cysteine containing proteins. In some instances, the isolated treated population of cells further comprises a set of cysteine-reactive probes.

[00209] As disclosed elsewhere herein, the small molecule fragment is a small molecule fragment of Formula (I):



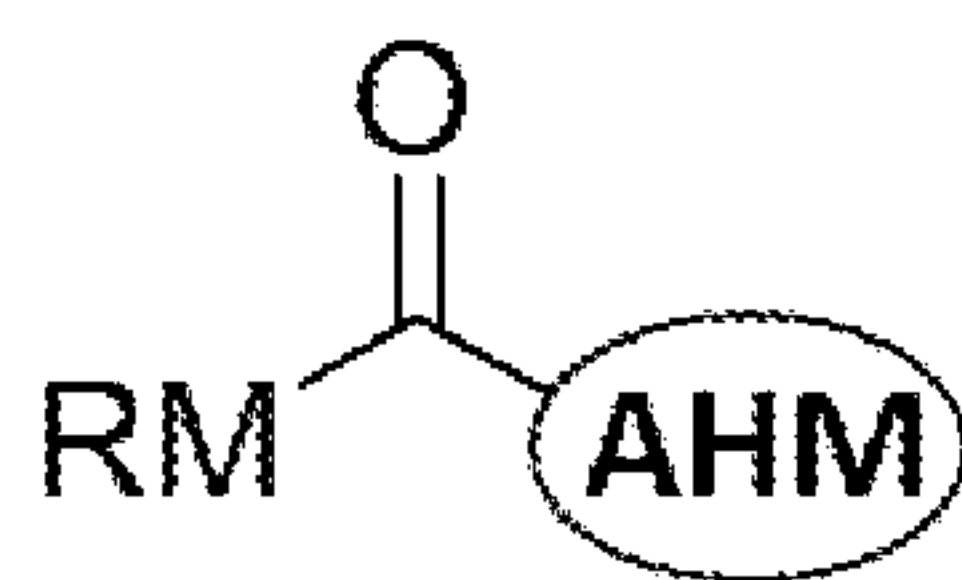
Formula (I)

wherein:

RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and F is a small molecule fragment moiety.

[00210] In some instances, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some cases, F is obtained from a compound library. In some embodiments, the compound library comprises ChemBridge fragment library, Pyramid Platform Fragment-Based Drug Discovery, Maybridge fragment library, FRGx from AnalytiCon, TCI-Frag from AnCoreX, Bio Building Blocks from ASINEX, BioFocus 3D from Charles River, Fragments of Life (FOL) from Emerald Bio, Enamine Fragment Library, IOTA Diverse 1500, BIONET fragments library, Life Chemicals Fragments Collection, OTAVA fragment library, Prestwick fragment library, Selcia fragment library, TimTec fragment-based library, Allium from Vitas-M Laboratory, or Zenobia fragment library. In some cases, F is a small molecule fragment moiety illustrated in Fig. 3. In some cases, F further comprises a linker moiety that connects F to the carbonyl moiety. In some embodiments, the small molecule fragment is a small molecule fragment illustrated in Fig. 3.

[00211] Also described elsewhere herein, the cysteine-reactive probe is a cysteine-reactive probe of Formula (II):



Formula (II)

wherein:

RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond to the thiol group of a cysteine residue; and AHM is an affinity handle moiety.

[00212] In some embodiments, the Michael acceptor moiety comprises an alkene or an alkyne moiety. In some instances, the affinity handle moiety comprises an affinity handle and a binding moiety that facilitates covalent interaction of the cysteine-reactive probe to a cysteine residue of a cysteine-containing protein. In some cases, the binding moiety is a small molecule fragment obtained from a compound library. In some cases, the compound library comprises ChemBridge fragment library, Pyramid Platform Fragment-Based Drug Discovery, Maybridge fragment library, FRGx from AnalytiCon, TCI-Frag from AnCoreX, Bio Building Blocks from ASINEX, BioFocus 3D from Charles River, Fragments of Life (FOL) from Emerald Bio, Enamine Fragment Library, IOTA Diverse 1500, BIONET fragments library, Life Chemicals Fragments Collection, OTAVA fragment library, Prestwick fragment library, Selcia fragment library, TimTec fragment-based library, Allium from Vitas-M Laboratory, or Zenobia fragment library.

[00213] In some instances, the affinity handle is a bioorthogonal affinity handle. In some cases, the affinity handle comprises a carbodiimide, N-hydroxysuccinimide (NHS) ester, imidoester, pentafluorophenyl ester, hydroxymethyl phosphine, maleimide, haloacetyl, pyridyl disulfide, thiosulfonate, vinylsulfone, hydrazide, alkoxyamine, alkyne, azide, or isocyanate group. In some cases, the affinity handle comprises an alkyne or an azide group. In some instances, the affinity handle is further conjugated to an affinity ligand. In some instances, the affinity ligand comprises a chromophore, a labeling group, or a combination thereof. In some cases, the chromophore comprises fluorochrome, non-fluorochrome chromophore, quencher, an absorption chromophore, fluorophore, organic dye, inorganic dye, metal chelate, or a fluorescent enzyme substrate. In some cases, the labeling group is biotin moiety, streptavidin moiety, bead, resin, a solid support, or a combination thereof. In some instances, the affinity handle moiety further comprises a chromophore. In some embodiments, the cysteine-reactive probe is a cysteine-reactive probe illustrated in Fig. 3.

[00214] Further described elsewhere herein, the cell or cell population is obtained from any mammal, such as human or non-human primates. In some embodiments, the cell or cell

population is an epithelial cell, connective tissue cell, hormone secreting cell, a nerve cell, a skeletal muscle cell, a blood cell, or an immune system cell. In additional embodiments, the cell or cell population is cancerous or is obtained from a tumor site.

Polypeptides comprising a cysteine interacting site

[00215] Further disclosed herein are polypeptides that comprise one or more of the cysteine interacting sites identified by a method described herein. In some embodiments, described herein is an isolated and purified polypeptide that comprises at least 90% sequence identity to at least seven contiguous amino acids of an amino acid sequence selected from Tables 1-3 or 8-9. In some embodiments, the isolated and purified polypeptide comprises at least 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity to at least seven contiguous amino acids of an amino acid sequence selected from Tables 1-3 or 8-9. In some embodiments, the isolated and purified polypeptide comprises 100% sequence identity to at least seven contiguous amino acids of an amino acid sequence selected from Tables 1-3 or 8-9. In some instances, the isolated and purified polypeptide consists 100% sequence identity to the full length of an amino acid sequence selected from Tables 1-3 or 8-9. In some instances, the isolated and purified polypeptide is at most 50 amino acids in length.

[00216] In some embodiments, additionally described herein include nucleic acid encoding a polypeptide that comprises at least 90% sequence identity at least seven contiguous amino acids of an amino acid sequence selected from Tables 1-3 or 8-9. In some embodiments, the nucleic acid encoding a polypeptide comprises at least 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence identity at least seven contiguous amino acids of an amino acid sequence selected from Tables 1-3 or 8-9. In some embodiments, the nucleic acid encoding a polypeptide comprises 100% sequence identity at least seven contiguous amino acids of an amino acid sequence selected from Tables 1-3 or 8-9. In some embodiments, the nucleic acid encoding a polypeptide consists 100% sequence identity to the full length of an amino acid sequence selected from Tables 1-3 or 8-9.

[00217] In some embodiments, further disclosed herein include a method of mapping a biologically active cysteine site on a protein, which comprises harvesting a set of cysteine-reactive probe-protein complexes from a sample wherein the cysteine-reactive probe comprises a reactive moiety capable of forming a covalent bond with a cysteine residue located on the cysteine containing protein; analyzing the set of cysteine-reactive probe-protein complexes by a proteomic analysis means; and based on the previous step, mapping the biologically active cysteine site on the protein.

[00218] In some embodiments, the analyzing further comprises treating the set of cysteine-reactive probe-protein complexes with a protease to generate a set of protein fragments. The

protease is a serine protease, a threonine protease, a cysteine protease, an aspartate protease, a glutamic acid protease, or a metalloprotease. In some instances, the protease is a serine protease. In some instances, the protease is trypsin. In some instances, cysteine-reactive probe-protein complex is further attached to a labeling group such as a biotin moiety. In some instances, the labeling group such as a biotin moiety further comprises a linker. In some instances, the linker is a peptide. In some instances, the peptide linker is about 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more amino acid residues in length. In some instances, the peptide linker contains a cleavage site. A non-limiting list of cleavage sites includes Tobacco Etch Virus (TEV), thrombin (Thr), enterokinase (EKT), activated Factor X (Xa), or human Rhinovirus 3C protease (3C/PreScission). In some instances, the peptide linker contains a TEV protease cleavage site. In some instances, the TEV protease cleavage site comprises the following sequence Gly-Gln-Phe-Tyr-Leu-Asn-Glu (SEQ ID NO: 860). In some instances, the biotin moiety is further coupled to a bead (e.g. a streptavidin-coupled bead).

[00219] In some instances, the protein from the cysteine-reactive probe-protein complex attached to the bead (via a biotin moiety comprising a linker and attached to a streptavidin-coupled bead) is digested with trypsin, and the immobilized peptide or protein fragment is further separated and collected. In some instances, the collected peptide or protein fragment is then digested by a protease (e.g. TEV protease), and the treated protein fragment is then separated, and collected for analysis. In some instances, the analysis is a proteomic analysis as described above and elsewhere herein. In some instances, the sequence of the protein fragment is further determined. In some instances, the protein fragment correlates to a small molecule fragment binding site on the cysteine containing protein.

[00220] In some embodiments, the sequence of the protein fragment correlates to a sequence as illustrated in Tables 1-3 or 8-9. In some instances, the sequence as shown in Tables 1-3 or 8-9 correlate to a site on the full length protein as a drug binding site. In some instances, the sequence as shown in Tables 1-3 or 8-9 correlate to a drug binding site. In some instances, polypeptides comprising one or more of the sequences as shown in Tables 1-3 or 8-9 serve as probes for small molecule fragment screening.

[00221] In some instances after the generation of a polypeptide, the polypeptide is subjected to one or more rounds of purification steps to remove impurities. In some instances, the purification step is a chromatographic step utilizing separation methods such as affinity-based, size-exclusion based, ion-exchange based, or the like. In some cases, the polypeptide is at most 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 99.9%, or 100% pure or without the presence of impurities. In some cases, the polypeptide is at least 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 99.9%, or 100% pure or without the presence of impurities.

[00222] As described above, nucleic acid encoding a polypeptide that is derived from a cysteine containing protein is subjected to one or more rounds of purification steps to remove impurities. In some instances, the purification step is a chromatographic step utilizing separation methods such as affinity-based, size-exclusion based, ion-exchange based, or the like. In some cases, the nucleic acid is at most 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 99.9%, or 100% pure or without the presence of impurities. In some cases, the nucleic acid is at least 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 99.9%, or 100% pure or without the presence of impurities.

[00223] As used herein, a polypeptide includes natural amino acids, unnatural amino acids, or a combination thereof. In some instances, an amino acid residue refers to a molecule containing both an amino group and a carboxyl group. Suitable amino acids include, without limitation, both the D- and L-isomers of the naturally-occurring amino acids, as well as non-naturally occurring amino acids prepared by organic synthesis or other metabolic routes. The term amino acid, as used herein, includes, without limitation, α -amino acids, natural amino acids, non-natural amino acids, and amino acid analogs.

[00224] The term “ α -amino acid” refers to a molecule containing both an amino group and a carboxyl group bound to a carbon which is designated the α -carbon.

[00225] The term “ β -amino acid” refers to a molecule containing both an amino group and a carboxyl group in a β configuration.

[00226] “Naturally occurring amino acid” refers to any one of the twenty amino acids commonly found in peptides synthesized in nature, and known by the one letter abbreviations A, R, N, C, D, Q, E, G, H, I, L, K, M, F, P, S, T, W, Y and V.

[00227] The following table shows a summary of the properties of natural amino acids:

Amino Acid	3-Letter Code	1-Letter Code	Side-chain Polarity	Side-chain charge (pH 7.4)	Hydropathy Index
Alanine	Ala	A	nonpolar	neutral	1.8
Arginine	Arg	R	polar	positive	-4.5
Asparagine	Asn	N	polar	neutral	-3.5
Aspartic acid	Asp	D	polar	negative	-3.5
Cysteine	Cys	C	polar	neutral	2.5
Glutamic acid	Glu	E	polar	negative	-3.5
Glutamine	Gln	Q	polar	neutral	-3.5
Glycine	Gly	G	nonpolar	neutral	-0.4
Histidine	His	H	polar	positive(10%) neutral(90%)	-3.2

Amino Acid	3-Letter Code	1-Letter Code	Side-chain Polarity	Side-chain charge (pH 7.4)	Hydropathy Index
Isoleucine	Ile	I	nonpolar	neutral	4.5
Leucine	Leu	L	nonpolar	neutral	3.8
Lysine	Lys	K	polar	positive	-3.9
Methionine	Met	M	nonpolar	neutral	1.9
Phenylalanine	Phe	F	nonpolar	neutral	2.8
Proline	Pro	P	nonpolar	neutral	-1.6
Serine	Ser	S	polar	neutral	-0.8
Threonine	Thr	T	polar	neutral	-0.7
Tryptophan	Trp	W	nonpolar	neutral	-0.9
Tyrosine	Tyr	Y	polar	neutral	-1.3
Valine	Val	V	nonpolar	neutral	4.2

[00228] “Hydrophobic amino acids” include small hydrophobic amino acids and large hydrophobic amino acids. “Small hydrophobic amino acid” are glycine, alanine, proline, and analogs thereof. “Large hydrophobic amino acids” are valine, leucine, isoleucine, phenylalanine, methionine, tryptophan, and analogs thereof. “Polar amino acids” are serine, threonine, asparagine, glutamine, cysteine, tyrosine, and analogs thereof. “Charged amino acids” are lysine, arginine, histidine, aspartate, glutamate, and analogs thereof. In some cases, aspartic acid and glutamic acid are referred to as acidic amino acids. In other cases, lysine, arginine and histidine are referred to as basic amino acids.

[00229] The term “amino acid analog” refers to a molecule which is structurally similar to an amino acid and which is substituted for an amino acid in the formation of a peptidomimetic macrocycle. Amino acid analogs include, without limitation, β -amino acids and amino acids where the amino or carboxy group is substituted by a similarly reactive group (e.g., substitution of the primary amine with a secondary or tertiary amine, or substitution of the carboxy group with an ester).

[00230] The term “non-natural amino acid” refers to an amino acid which is not one of the twenty amino acids commonly found in peptides synthesized in nature, and known by the one letter abbreviations A, R, N, C, D, Q, E, G, H, I, L, K, M, F, P, S, T, W, Y and V.

[00231] In some instances, amino acid analogs include β -amino acid analogs. Examples of β -amino acid analogs include, but are not limited to, the following: cyclic β -amino acid analogs; β -alanine; (R)- β -phenylalanine; (R)-1,2,3,4-tetrahydro-isoquinoline-3-acetic acid; (R)-3-amino-4-(1-naphthyl)-butyric acid; (R)-3-amino-4-(2,4-dichlorophenyl)butyric acid; (R)-3-amino-4-(2-chlorophenyl)-butyric acid; (R)-3-amino-4-(2-cyanophenyl)-butyric acid; (R)-3-amino-4-(2-fluorophenyl)-butyric acid; (R)-3-amino-4-(2-furyl)-butyric acid; (R)-3-amino-4-(2-

methylphenyl)-butyric acid; (R)-3-amino-4-(2-naphthyl)-butyric acid; (R)-3-amino-4-(2-thienyl)-butyric acid; (R)-3-amino-4-(2-trifluoromethylphenyl)-butyric acid; (R)-3-amino-4-(3,4-dichlorophenyl)butyric acid; (R)-3-amino-4-(3,4-difluorophenyl)butyric acid; (R)-3-amino-4-(3-benzothienyl)-butyric acid; (R)-3-amino-4-(3-chlorophenyl)-butyric acid; (R)-3-amino-4-(3-cyanophenyl)-butyric acid; (R)-3-amino-4-(3-fluorophenyl)-butyric acid; (R)-3-amino-4-(3-methylphenyl)-butyric acid; (R)-3-amino-4-(3-pyridyl)-butyric acid; (R)-3-amino-4-(3-thienyl)-butyric acid; (R)-3-amino-4-(3-trifluoromethylphenyl)-butyric acid; (R)-3-amino-4-(4-bromophenyl)-butyric acid; (R)-3-amino-4-(4-chlorophenyl)-butyric acid; (R)-3-amino-4-(4-cyanophenyl)-butyric acid; (R)-3-amino-4-(4-fluorophenyl)-butyric acid; (R)-3-amino-4-(4-iodophenyl)-butyric acid; (R)-3-amino-4-(4-methylphenyl)-butyric acid; (R)-3-amino-4-(4-nitrophenyl)-butyric acid; (R)-3-amino-4-(4-pyridyl)-butyric acid; (R)-3-amino-4-(4-trifluoromethylphenyl)-butyric acid; (R)-3-amino-4-pentafluoro-phenylbutyric acid; (R)-3-amino-5-hexenoic acid; (R)-3-amino-5-hexynoic acid; (R)-3-amino-5-phenylpentanoic acid; (R)-3-amino-6-phenyl-5-hexenoic acid; (S)-1,2,3,4-tetrahydro-isoquinoline-3-acetic acid; (S)-3-amino-4-(1-naphthyl)-butyric acid; (S)-3-amino-4-(2,4-dichlorophenyl)butyric acid; (S)-3-amino-4-(2-chlorophenyl)-butyric acid; (S)-3-amino-4-(2-cyanophenyl)-butyric acid; (S)-3-amino-4-(2-fluorophenyl)-butyric acid; (S)-3-amino-4-(2-furyl)-butyric acid; (S)-3-amino-4-(2-methylphenyl)-butyric acid; (S)-3-amino-4-(2-naphthyl)-butyric acid; (S)-3-amino-4-(2-thienyl)-butyric acid; (S)-3-amino-4-(2-trifluoromethylphenyl)-butyric acid; (S)-3-amino-4-(3,4-dichlorophenyl)butyric acid; (S)-3-amino-4-(3,4-difluorophenyl)butyric acid; (S)-3-amino-4-(3-benzothienyl)-butyric acid; (S)-3-amino-4-(3-chlorophenyl)-butyric acid; (S)-3-amino-4-(3-cyanophenyl)-butyric acid; (S)-3-amino-4-(3-fluorophenyl)-butyric acid; (S)-3-amino-4-(3-methylphenyl)-butyric acid; (S)-3-amino-4-(3-pyridyl)-butyric acid; (S)-3-amino-4-(3-thienyl)-butyric acid; (S)-3-amino-4-(3-trifluoromethylphenyl)-butyric acid; (S)-3-amino-4-(4-bromophenyl)-butyric acid; (S)-3-amino-4-(4-chlorophenyl) butyric acid; (S)-3-amino-4-(4-cyanophenyl)-butyric acid; (S)-3-amino-4-(4-fluorophenyl) butyric acid; (S)-3-amino-4-(4-iodophenyl)-butyric acid; (S)-3-amino-4-(4-methylphenyl)-butyric acid; (S)-3-amino-4-(4-nitrophenyl)-butyric acid; (S)-3-amino-4-(4-pyridyl)-butyric acid; (S)-3-amino-4-(4-trifluoromethylphenyl)-butyric acid; (S)-3-amino-4-pentafluoro-phenylbutyric acid; (S)-3-amino-5-hexenoic acid; (S)-3-amino-5-hexynoic acid; (S)-3-amino-5-phenylpentanoic acid; (S)-3-amino-6-phenyl-5-hexenoic acid; 1,2,5,6-tetrahydropyridine-3-carboxylic acid; 1,2,5,6-tetrahydropyridine-4-carboxylic acid; 3-amino-3-(2-chlorophenyl)-propionic acid; 3-amino-3-(2-thienyl)-propionic acid; 3-amino-3-(3-bromophenyl)-propionic acid; 3-amino-3-(4-chlorophenyl)-propionic acid; 3-amino-3-(4-methoxyphenyl)-propionic acid; 3-amino-4,4,4-trifluoro-butyric acid; 3-aminoadipic acid; D- β -phenylalanine; β -leucine; L- β -homoalanine; L- β -

homoaspartic acid γ -benzyl ester; L- β -homoglutamic acid δ -benzyl ester; L- β -homoisoleucine; L- β -homoleucine; L- β -homomethionine; L- β -homophenylalanine; L- β -homoproline; L- β -homotryptophan; L- β -homovaline; L-N ω -benzyloxycarbonyl- β -homolysine; N ω -L- β -homoarginine; O-benzyl-L- β -homohydroxyproline; O-benzyl-L- β -homoserine; O-benzyl-L- β -homothreonine; O-benzyl-L- β -homotyrosine; γ -trityl-L- β -homoasparagine; (R)- β -phenylalanine; L- β -homoaspartic acid γ -t-butyl ester; L- β -homoglutamic acid δ -t-butyl ester; L-N ω - β -homolysine; N δ -trityl-L- β -homoglutamine; N ω -2,2,4,6,7-pentamethyl-dihydrobenzofuran-5-sulfonyl-L- β -homoarginine; O-t-butyl-L- β -homohydroxy-proline; O-t-butyl-L- β -homoserine; O-t-butyl-L- β -homothreonine; O-t-butyl-L- β -homotyrosine; 2-aminocyclopentane carboxylic acid; and 2-aminocyclohexane carboxylic acid.

[00232] In some instances, amino acid analogs include analogs of alanine, valine, glycine or leucine. Examples of amino acid analogs of alanine, valine, glycine, and leucine include, but are not limited to, the following: α -methoxyglycine; α -allyl-L-alanine; α -aminoisobutyric acid; α -methyl-leucine; β -(1-naphthyl)-D-alanine; β -(1-naphthyl)-L-alanine; β -(2-naphthyl)-D-alanine; β -(2-naphthyl)-L-alanine; β -(2-pyridyl)-D-alanine; β -(2-pyridyl)-L-alanine; β -(2-thienyl)-D-alanine; β -(2-thienyl)-L-alanine; β -(3-benzothienyl)-D-alanine; β -(3-benzothienyl)-L-alanine; β -(3-pyridyl)-D-alanine; β -(3-pyridyl)-L-alanine; β -(4-pyridyl)-D-alanine; β -(4-pyridyl)-L-alanine; β -chloro-L-alanine; β -cyano-L-alanin; β -cyclohexyl-D-alanine; β -cyclohexyl-L-alanine; β -cyclopenten-1-yl-alanine; β -cyclopentyl-alanine; β -cyclopropyl-L-Ala-OH.dicyclohexylammonium salt; β -t-butyl-D-alanine; β -t-butyl-L-alanine; γ -aminobutyric acid; L- α,β -diaminopropionic acid; 2,4-dinitro-phenylglycine; 2,5-dihydro-D-phenylglycine; 2-amino-4,4,4-trifluorobutyric acid; 2-fluoro-phenylglycine; 3-amino-4,4,4-trifluoro-butyric acid; 3-fluoro-valine; 4,4,4-trifluoro-valine; 4,5-dehydro-L-leu-OH.dicyclohexylammonium salt; 4-fluoro-D-phenylglycine; 4-fluoro-L-phenylglycine; 4-hydroxy-D-phenylglycine; 5,5,5-trifluoro-leucine; 6-aminohexanoic acid; cyclopentyl-D-Gly-OH.dicyclohexylammonium salt; cyclopentyl-Gly-OH.dicyclohexylammonium salt; D- α,β -diaminopropionic acid; D- α -aminobutyric acid; D- α -t-butylglycine; D-(2-thienyl)glycine; D-(3-thienyl)glycine; D-2-aminocaproic acid; D-2-indanylglycine; D-allylglycine-dicyclohexylammonium salt; D-cyclohexylglycine; D-norvaline; D-phenylglycine; β -aminobutyric acid; β -aminoisobutyric acid; (2-bromophenyl)glycine; (2-methoxyphenyl)glycine; (2-methylphenyl)glycine; (2-thiazoyl)glycine; (2-thienyl)glycine; 2-amino-3-(dimethylamino)-propionic acid; L- α,β -diaminopropionic acid; L- α -aminobutyric acid; L- α -t-butylglycine; L-(3-thienyl)glycine; L-2-amino-3-(dimethylamino)-propionic acid; L-2-aminocaproic acid dicyclohexyl-ammonium salt; L-2-indanylglycine; L-allylglycine.dicyclohexyl ammonium salt; L-cyclohexylglycine; L-phenylglycine; L-propargylglycine; L-norvaline; N- α -aminomethyl-L-alanine; D- α,γ -

diaminobutyric acid; L- α,γ -diaminobutyric acid; β -cyclopropyl-L-alanine; (N- β -(2,4-dinitrophenyl))-L- α,β -diaminopropionic acid; (N- β -1-(4,4-dimethyl-2,6-dioxocyclohex-1-ylidene)ethyl)-D- α,β -diaminopropionic acid; (N- β -1-(4,4-dimethyl-2,6-dioxocyclohex-1-ylidene)ethyl)-L- α,β -diaminopropionic acid; (N- β -4-methyltrityl)-L- α,β -diaminopropionic acid; (N- β -allyloxycarbonyl)-L- α,β -diaminopropionic acid; (N- γ -1-(4,4-dimethyl-2,6-dioxocyclohex-1-ylidene)ethyl)-D- α,γ -diaminobutyric acid; (N- γ -1-(4,4-dimethyl-2,6-dioxocyclohex-1-ylidene)ethyl)-L- α,γ -diaminobutyric acid; (N- γ -4-methyltrityl)-D- α,γ -diaminobutyric acid; (N- γ -4-methyltrityl)-L- α,γ -diaminobutyric acid; (N- γ -allyloxycarbonyl)-L- α,γ -diaminobutyric acid; D- α,γ -diaminobutyric acid; 4,5-dehydro-L-leucine; cyclopentyl-D-Gly-OH; cyclopentyl-Gly-OH; D-allylglycine; D-homocyclohexylalanine; L-1-pyrenylalanine; L-2-aminocaproic acid; L-allylglycine; L-homocyclohexylalanine; and N-(2-hydroxy-4-methoxy-Bzl)-Gly-OH.

[00233] In some instances, amino acid analogs include analogs of arginine or lysine. Examples of amino acid analogs of arginine and lysine include, but are not limited to, the following: citrulline; L-2-amino-3-guanidinopropionic acid; L-2-amino-3-ureidopropionic acid; L-citrulline; Lys(Me)₂-OH; Lys(N₃)—OH; N δ -benzyloxycarbonyl-L-ornithine; N ω -nitro-D-arginine; N ω -nitro-L-arginine; α -methyl-ornithine; 2,6-diaminoheptanedioic acid; L-ornithine; (N δ -1-(4,4-dimethyl-2,6-dioxo-cyclohex-1-ylidene)ethyl)-D-ornithine; (N δ -1-(4,4-dimethyl-2,6-dioxo-cyclohex-1-ylidene)ethyl)-L-ornithine; (N δ -4-methyltrityl)-D-ornithine; (N δ -4-methyltrityl)-L-ornithine; D-ornithine; L-ornithine; Arg(Me)(Pbf)-OH; Arg(Me)₂-OH (asymmetrical); Arg(Me)₂-OH (symmetrical); Lys(ivDde)-OH; Lys(Me)₂-OH.HCl; Lys(Me)₃-OH chloride; N ω -nitro-D-arginine; and N ω -nitro-L-arginine.

[00234] In some instances, amino acid analogs include analogs of aspartic or glutamic acids. Examples of amino acid analogs of aspartic and glutamic acids include, but are not limited to, the following: α -methyl-D-aspartic acid; α -methyl-glutamic acid; α -methyl-L-aspartic acid; γ -methylene-glutamic acid; (N- γ -ethyl)-L-glutamine; [N- α -(4-aminobenzoyl)]-L-glutamic acid; 2,6-diaminopimelic acid; L- α -aminosuberic acid; D-2-aminoadipic acid; D- α -aminosuberic acid; α -aminopimelic acid; iminodiacetic acid; L-2-aminoadipic acid; threo- β -methyl-aspartic acid; γ -carboxy-D-glutamic acid γ,γ -di-*t*-butyl ester; γ -carboxy-L-glutamic acid γ,γ -di-*t*-butyl ester; Glu(OAll)-OH; L-Asu(OtBu)—OH; and pyroglutamic acid.

[00235] In some instances, amino acid analogs include analogs of cysteine and methionine. Examples of amino acid analogs of cysteine and methionine include, but are not limited to, Cys(farnesyl)-OH, Cys(farnesyl)-OMe, α -methyl-methionine, Cys(2-hydroxyethyl)-OH, Cys(3-aminopropyl)-OH, 2-amino-4-(ethylthio)butyric acid, buthionine, buthioninesulfoximine, ethionine, methionine methylsulfonium chloride, selenomethionine, cysteic acid, [2-(4-pyridyl)ethyl]-DL-penicillamine, [2-(4-pyridyl)ethyl]-L-cysteine, 4-methoxybenzyl-D-

penicillamine, 4-methoxybenzyl-L-penicillamine, 4-methylbenzyl-D-penicillamine, 4-methylbenzyl-L-penicillamine, benzyl-D-cysteine, benzyl-L-cysteine, benzyl-DL-homocysteine, carbamoyl-L-cysteine, carboxyethyl-L-cysteine, carboxymethyl-L-cysteine, diphenylmethyl-L-cysteine, ethyl-L-cysteine, methyl-L-cysteine, t-butyl-D-cysteine, trityl-L-homocysteine, trityl-D-penicillamine, cystathionine, homocystine, L-homocystine, (2-aminoethyl)-L-cysteine, seleno-L-cystine, cystathionine, Cys(StBu)—OH, and acetamidomethyl-D-penicillamine.

[00236] In some instances, amino acid analogs include analogs of phenylalanine and tyrosine. Examples of amino acid analogs of phenylalanine and tyrosine include β -methyl-phenylalanine, β -hydroxyphenylalanine, α -methyl-3-methoxy-DL-phenylalanine, α -methyl-D-phenylalanine, α -methyl-L-phenylalanine, 1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid, 2,4-dichloro-phenylalanine, 2-(trifluoromethyl)-D-phenylalanine, 2-(trifluoromethyl)-L-phenylalanine, 2-bromo-D-phenylalanine, 2-bromo-L-phenylalanine, 2-chloro-D-phenylalanine, 2-chloro-L-phenylalanine, 2-cyano-D-phenylalanine, 2-cyano-L-phenylalanine, 2-fluoro-D-phenylalanine, 2-fluoro-L-phenylalanine, 2-methyl-D-phenylalanine, 2-methyl-L-phenylalanine, 2-nitro-D-phenylalanine, 2-nitro-L-phenylalanine, 2,4,5-trihydroxy-phenylalanine, 3,4,5-trifluoro-D-phenylalanine, 3,4,5-trifluoro-L-phenylalanine, 3,4-dichloro-D-phenylalanine, 3,4-dichloro-L-phenylalanine, 3,4-difluoro-D-phenylalanine, 3,4-difluoro-L-phenylalanine, 3,4-dihydroxy-L-phenylalanine, 3,4-dimethoxy-L-phenylalanine, 3,5,3'-triiodo-L-thyronine, 3,5-diiodo-D-tyrosine, 3,5-diiodo-L-tyrosine, 3,5-diiodo-L-thyronine, 3-(trifluoromethyl)-D-phenylalanine, 3-(trifluoromethyl)-L-phenylalanine, 3-amino-L-tyrosine, 3-bromo-D-phenylalanine, 3-bromo-L-phenylalanine, 3-chloro-D-phenylalanine, 3-chloro-L-phenylalanine, 3-chloro-L-tyrosine, 3-cyano-D-phenylalanine, 3-cyano-L-phenylalanine, 3-fluoro-D-phenylalanine, 3-fluoro-L-phenylalanine, 3-fluoro-tyrosine, 3-iodo-D-phenylalanine, 3-iodo-L-phenylalanine, 3-iodo-L-tyrosine, 3-methoxy-L-tyrosine, 3-methyl-D-phenylalanine, 3-methyl-L-phenylalanine, 3-nitro-D-phenylalanine, 3-nitro-L-phenylalanine, 3-nitro-L-tyrosine, 4-(trifluoromethyl)-D-phenylalanine, 4-(trifluoromethyl)-L-phenylalanine, 4-amino-D-phenylalanine, 4-amino-L-phenylalanine, 4-benzoyl-D-phenylalanine, 4-benzoyl-L-phenylalanine, 4-bis(2-chloroethyl)amino-L-phenylalanine, 4-bromo-D-phenylalanine, 4-bromo-L-phenylalanine, 4-chloro-D-phenylalanine, 4-chloro-L-phenylalanine, 4-cyano-D-phenylalanine, 4-cyano-L-phenylalanine, 4-fluoro-D-phenylalanine, 4-fluoro-L-phenylalanine, 4-iodo-D-phenylalanine, 4-iodo-L-phenylalanine, homophenylalanine, thyroxine, 3,3-diphenylalanine, thyronine, ethyl-tyrosine, and methyl-tyrosine.

[00237] In some instances, amino acid analogs include analogs of proline. Examples of amino acid analogs of proline include, but are not limited to, 3,4-dehydro-proline, 4-fluoro-proline, cis-4-hydroxy-proline, thiazolidine-2-carboxylic acid, and trans-4-fluoro-proline.

[00238] In some instances, amino acid analogs include analogs of serine and threonine. Examples of amino acid analogs of serine and threonine include, but are not limited to, 3-amino-2-hydroxy-5-methylhexanoic acid, 2-amino-3-hydroxy-4-methylpentanoic acid, 2-amino-3-ethoxybutanoic acid, 2-amino-3-methoxybutanoic acid, 4-amino-3-hydroxy-6-methylheptanoic acid, 2-amino-3-benzyloxypropionic acid, 2-amino-3-benzyloxypropionic acid, 2-amino-3-ethoxypropionic acid, 4-amino-3-hydroxybutanoic acid, and α -methylserine.

[00239] In some instances, amino acid analogs include analogs of tryptophan. Examples of amino acid analogs of tryptophan include, but are not limited to, the following: α -methyl-tryptophan; β -(3-benzothienyl)-D-alanine; β -(3-benzothienyl)-L-alanine; 1-methyl-tryptophan; 4-methyl-tryptophan; 5-benzyloxy-tryptophan; 5-bromo-tryptophan; 5-chloro-tryptophan; 5-fluoro-tryptophan; 5-hydroxy-tryptophan; 5-hydroxy-L-tryptophan; 5-methoxy-tryptophan; 5-methoxy-L-tryptophan; 5-methyl-tryptophan; 6-bromo-tryptophan; 6-chloro-D-tryptophan; 6-chloro-tryptophan; 6-fluoro-tryptophan; 6-methyl-tryptophan; 7-benzyloxy-tryptophan; 7-bromo-tryptophan; 7-methyl-tryptophan; D-1,2,3,4-tetrahydro-norharman-3-carboxylic acid; 6-methoxy-1,2,3,4-tetrahydronorharman-1-carboxylic acid; 7-azatryptophan; L-1,2,3,4-tetrahydro-norharman-3-carboxylic acid; 5-methoxy-2-methyl-tryptophan; and 6-chloro-L-tryptophan.

[00240] In some instances, amino acid analogs are racemic. In some instances, the D isomer of the amino acid analog is used. In some cases, the L isomer of the amino acid analog is used. In some instances, the amino acid analog comprises chiral centers that are in the R or S configuration. Sometimes, the amino group(s) of a β -amino acid analog is substituted with a protecting group, e.g., tert-butyloxycarbonyl (BOC group), 9-fluorenylmethyloxycarbonyl (FMOC), tosyl, and the like. Sometimes, the carboxylic acid functional group of a β -amino acid analog is protected, e.g., as its ester derivative. In some cases, the salt of the amino acid analog is used.

[00241] In some embodiments, nucleic acid molecules refer to at least two nucleotides covalently linked together. In some instances, a nucleic acid described herein contains phosphodiester bonds, although in some cases, as outlined below (for example in the construction of primers and probes such as label probes), nucleic acid analogs are included that have alternate backbones, comprising, for example, phosphoramidate (Beaucage et al., Tetrahedron 49(10):1925 (1993) and references therein; Letsinger, J. Org. Chem. 35:3800 (1970); Sprinzl et al., Eur. J. Biochem. 81:579 (1977); Letsinger et al., Nucl. Acids Res. 14:3487 (1986); Sawai et al, Chem. Lett. 805 (1984), Letsinger et al., J. Am. Chem. Soc. 110:4470 (1988); and Pauwels et al., Chemica Scripta 26:141 91986)), phosphorothioate (Mag et al., Nucleic Acids Res. 19:1437 (1991); and U.S. Pat. No. 5,644,048), phosphorodithioate (Briu et al., J. Am. Chem. Soc. 111:2321 (1989), O-methylphosphoroamidite linkages (see Eckstein,

Oligonucleotides and Analogues: A Practical Approach, Oxford University Press), and peptide nucleic acid (also referred to herein as “PNA”) backbones and linkages (see Egholm, J. Am. Chem. Soc. 114:1895 (1992); Meier et al., Chem. Int. Ed. Engl. 31:1008 (1992); Nielsen, Nature, 365:566 (1993); Carlsson et al., Nature 380:207 (1996), all of which are incorporated by reference). Other analog nucleic acids include those with bicyclic structures including locked nucleic acids (also referred to herein as “LNA”), Koshkin et al., J. Am. Chem. Soc. 120:13252 3 (1998); positive backbones (Denpcy et al., Proc. Natl. Acad. Sci. USA 92:6097 (1995); non-ionic backbones (U.S. Pat. Nos. 5,386,023, 5,637,684, 5,602,240, 5,216,141 and 4,469,863; Kiedrowshi et al., Angew. Chem. Intl. Ed. English 30:423 (1991); Letsinger et al., J. Am. Chem. Soc. 110:4470 (1988); Letsinger et al., Nucleoside & Nucleotide 13:1597 (1994); Chapters 2 and 3, ASC Symposium Series 580, “Carbohydrate Modifications in Antisense Research”, Ed. Y. S. Sanghui and P. Dan Cook; Mesmaeker et al., Bioorganic & Medicinal Chem. Lett. 4:395 (1994); Jeffs et al., J. Biomolecular NMR 34:17 (1994); Tetrahedron Lett. 37:743 (1996)) and non-ribose backbones, including those described in U.S. Pat. Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7, ASC Symposium Series 580, “Carbohydrate Modifications in Antisense Research”, Ed. Y. S. Sanghui and P. Dan Cook. Nucleic acids containing one or more carbocyclic sugars are also included within the definition of nucleic acids (see Jenkins et al., Chem. Soc. Rev. (1995) pp 169 176). Several nucleic acid analogs are described in Rawls, C & E News Jun. 2, 1997 page 35. “Locked nucleic acids” are also included within the definition of nucleic acid analogs. LNAs are a class of nucleic acid analogues in which the ribose ring is “locked” by a methylene bridge connecting the 2'-O atom with the 4'-C atom. All of these references are hereby expressly incorporated by reference. In some instances, these modifications of the ribose-phosphate backbone are done to increase the stability and half-life of such molecules in physiological environments. For example, PNA:DNA and LNA-DNA hybrids exhibit higher stability and thus are used in some embodiments. The target nucleic acids are single stranded or double stranded, as specified, or contain portions of both double stranded or single stranded sequence. Depending on the application, the nucleic acids are DNA (including, e.g., genomic DNA, mitochondrial DNA, and cDNA), RNA (including, e.g., mRNA and rRNA) or a hybrid, where the nucleic acid contains any combination of deoxyribo- and ribo-nucleotides, and any combination of bases, including uracil, adenine, thymine, cytosine, guanine, inosine, xanthine hypoxanthine, isocytosine, isoguanine, etc.

Samples, Analytical Techniques, and Instrumentation

[00242] In certain embodiments, one or more of the methods disclosed herein comprise a sample. In some embodiments, the sample is a cell sample or a tissue sample. In some instances, the sample is a cell sample. In some embodiments, the sample for use with the methods

described herein is obtained from cells of an animal. In some instances, the animal cell includes a cell from a marine invertebrate, fish, insects, amphibian, reptile, or mammal. In some instances, the mammalian cell is a primate, ape, equine, bovine, porcine, canine, feline, or rodent. In some instances, the mammal is a primate, ape, dog, cat, rabbit, ferret, or the like. In some cases, the rodent is a mouse, rat, hamster, gerbil, hamster, chinchilla, or guinea pig. In some embodiments, the bird cell is from a canary, parakeet or parrots. In some embodiments, the reptile cell is from a turtles, lizard or snake. In some cases, the fish cell is from a tropical fish. In some cases, the fish cell is from a zebrafish (e.g. Danino rerio). In some cases, the worm cell is from a nematode (e.g. C. elegans). In some cases, the amphibian cell is from a frog. In some embodiments, the arthropod cell is from a tarantula or hermit crab.

[00243] In some embodiments, the sample for use with the methods described herein is obtained from a mammalian cell. In some instances, the mammalian cell is an epithelial cell, connective tissue cell, hormone secreting cell, a nerve cell, a skeletal muscle cell, a blood cell, or an immune system cell.

[00244] Exemplary mammalian cells include, but are not limited to, 293A cell line, 293FT cell line, 293F cells, 293 H cells, HEK 293 cells, CHO DG44 cells, CHO-S cells, CHO-K1 cells, Expi293F™ cells, Flp-In™ T-REx™ 293 cell line, Flp-In™-293 cell line, Flp-In™-3T3 cell line, Flp-In™-BHK cell line, Flp-In™-CHO cell line, Flp-In™-CV-1 cell line, Flp-In™-Jurkat cell line, FreeStyle™ 293-F cells, FreeStyle™ CHO-S cells, GripTite™ 293 MSR cell line, GS-CHO cell line, HepaRG™ cells, T-REx™ Jurkat cell line, Per.C6 cells, T-REx™-293 cell line, T-REx™-CHO cell line, T-REx™-HeLa cell line, NC-HIMT cell line, and PC12 cell line.

[00245] In some instances, the sample for use with the methods described herein is obtained from cells of a tumor cell line. In some instances, the sample is obtained from cells of a solid tumor cell line. In some instances, the solid tumor cell line is a sarcoma cell line. In some instances, the solid tumor cell line is a carcinoma cell line. In some embodiments, the sarcoma cell line is obtained from a cell line of alveolar rhabdomyosarcoma, alveolar soft part sarcoma, ameloblastoma, angiosarcoma, chondrosarcoma, chordoma, clear cell sarcoma of soft tissue, dedifferentiated liposarcoma, desmoid, desmoplastic small round cell tumor, embryonal rhabdomyosarcoma, epithelioid fibrosarcoma, epithelioid hemangioendothelioma, epithelioid sarcoma, esthesioneuroblastoma, Ewing sarcoma, extrarenal rhabdoid tumor, extraskeletal myxoid chondrosarcoma, extraskeletal osteosarcoma, fibrosarcoma, giant cell tumor, hemangiopericytoma, infantile fibrosarcoma, inflammatory myofibroblastic tumor, Kaposi sarcoma, leiomyosarcoma of bone, liposarcoma, liposarcoma of bone, malignant fibrous histiocytoma (MFH), malignant fibrous histiocytoma (MFH) of bone, malignant

mesenchymoma, malignant peripheral nerve sheath tumor, mesenchymal chondrosarcoma, myxofibrosarcoma, myxoid liposarcoma, myxoinflammatory fibroblastic sarcoma, neoplasms with perivascular epithelioid cell differentiation, osteosarcoma, parosteal osteosarcoma, neoplasm with perivascular epithelioid cell differentiation, periosteal osteosarcoma, pleomorphic liposarcoma, pleomorphic rhabdomyosarcoma, PNET/extraskeletal Ewing tumor, rhabdomyosarcoma, round cell liposarcoma, small cell osteosarcoma, solitary fibrous tumor, synovial sarcoma, telangiectatic osteosarcoma.

[00246] In some embodiments, the carcinoma cell line is obtained from a cell line of adenocarcinoma, squamous cell carcinoma, adenosquamous carcinoma, anaplastic carcinoma, large cell carcinoma, small cell carcinoma, anal cancer, appendix cancer, bile duct cancer (i.e., cholangiocarcinoma), bladder cancer, brain tumor, breast cancer, cervical cancer, colon cancer, cancer of Unknown Primary (CUP), esophageal cancer, eye cancer, fallopian tube cancer, gastroenterological cancer, kidney cancer, liver cancer, lung cancer, medulloblastoma, melanoma, oral cancer, ovarian cancer, pancreatic cancer, parathyroid disease, penile cancer, pituitary tumor, prostate cancer, rectal cancer, skin cancer, stomach cancer, testicular cancer, throat cancer, thyroid cancer, uterine cancer, vaginal cancer, or vulvar cancer.

[00247] In some instances, the sample is obtained from cells of a hematologic malignant cell line. In some instances, the hematologic malignant cell line is a T-cell cell line. In some instances, B-cell cell line. In some instances, the hematologic malignant cell line is obtained from a T-cell cell line of: peripheral T-cell lymphoma not otherwise specified (PTCL-NOS), anaplastic large cell lymphoma, angioimmunoblastic lymphoma, cutaneous T-cell lymphoma, adult T-cell leukemia/lymphoma (ATLL), blastic NK-cell lymphoma, enteropathy-type T-cell lymphoma, hematosplenic gamma-delta T-cell lymphoma, lymphoblastic lymphoma, nasal NK/T-cell lymphomas, or treatment-related T-cell lymphomas.

[00248] In some instances, the hematologic malignant cell line is obtained from a B-cell cell line of: acute lymphoblastic leukemia (ALL), acute myelogenous leukemia (AML), chronic myelogenous leukemia (CML), acute monocytic leukemia (AMoL), chronic lymphocytic leukemia (CLL), high-risk chronic lymphocytic leukemia (CLL), small lymphocytic lymphoma (SLL), high-risk small lymphocytic lymphoma (SLL), follicular lymphoma (FL), mantle cell lymphoma (MCL), Waldenstrom's macroglobulinemia, multiple myeloma, extranodal marginal zone B cell lymphoma, nodal marginal zone B cell lymphoma, Burkitt's lymphoma, non-Burkitt high grade B cell lymphoma, primary mediastinal B-cell lymphoma (PMBL), immunoblastic large cell lymphoma, precursor B-lymphoblastic lymphoma, B cell prolymphocytic leukemia, lymphoplasmacytic lymphoma, splenic marginal zone lymphoma, plasma cell myeloma,

plasmacytoma, mediastinal (thymic) large B cell lymphoma, intravascular large B cell lymphoma, primary effusion lymphoma, or lymphomatoid granulomatosis.

[00249] In some embodiments, the sample for use with the methods described herein is obtained from a tumor cell line. Exemplary tumor cell line includes, but is not limited to, 600MPE, AU565, BT-20, BT-474, BT-483, BT-549, Evsa-T, Hs578T, MCF-7, MDA-MB-231, SkBr3, T-47D, HeLa, DU145, PC3, LNCaP, A549, H1299, NCI-H460, A2780, SKOV-3/Luc, Neuro2a, RKO, RKO-AS45-1, HT-29, SW1417, SW948, DLD-1, SW480, Capan-1, MC/9, B72.3, B25.2, B6.2, B38.1, DMS 153, SU.86.86, SNU-182, SNU-423, SNU-449, SNU-475, SNU-387, Hs 817.T, LMH, LMH/2A, SNU-398, PLHC-1, HepG2/SF, OCI-Ly1, OCI-Ly2, OCI-Ly3, OCI-Ly4, OCI-Ly6, OCI-Ly7, OCI-Ly10, OCI-Ly18, OCI-Ly19, U2932, DB, HBL-1, RIVA, SUDHL2, TMD8, MEC1, MEC2, 8E5, CCRF-CEM, MOLT-3, TALL-104, AML-193, THP-1, BDCM, HL-60, Jurkat, RPMI 8226, MOLT-4, RS4, K-562, KASUMI-1, Daudi, GA-10, Raji, JeKo-1, NK-92, and Mino.

[00250] In some embodiments, the sample for use in the methods is from any tissue or fluid from an individual. Samples include, but are not limited to, tissue (e.g. connective tissue, muscle tissue, nervous tissue, or epithelial tissue), whole blood, dissociated bone marrow, bone marrow aspirate, pleural fluid, peritoneal fluid, central spinal fluid, abdominal fluid, pancreatic fluid, cerebrospinal fluid, brain fluid, ascites, pericardial fluid, urine, saliva, bronchial lavage, sweat, tears, ear flow, sputum, hydrocele fluid, semen, vaginal flow, milk, amniotic fluid, and secretions of respiratory, intestinal or genitourinary tract. In some embodiments, the sample is a tissue sample, such as a sample obtained from a biopsy or a tumor tissue sample. In some embodiments, the sample is a blood serum sample. In some embodiments, the sample is a blood cell sample containing one or more peripheral blood mononuclear cells (PBMCs). In some embodiments, the sample contains one or more circulating tumor cells (CTCs). In some embodiments, the sample contains one or more disseminated tumor cells (DTC, e.g., in a bone marrow aspirate sample).

[00251] In some embodiments, the samples are obtained from the individual by any suitable means of obtaining the sample using well-known and routine clinical methods. Procedures for obtaining tissue samples from an individual are well known. For example, procedures for drawing and processing tissue sample such as from a needle aspiration biopsy is well-known and is employed to obtain a sample for use in the methods provided. Typically, for collection of such a tissue sample, a thin hollow needle is inserted into a mass such as a tumor mass for sampling of cells that, after being stained, will be examined under a microscope.

Sample Preparation and Analysis

[00252] In some embodiments, the sample is a sample solution. In some instances, the sample solution comprises a solution such as a buffer (e.g. phosphate buffered saline) or a media. In some embodiments, the media is an isotopically labeled media. In some instances, the sample solution is a cell solution.

[00253] In some embodiments, the sample (e.g., cells or a cell solution) is incubated with a cysteine-reactive probe for analysis of protein cysteine-reactive probe interactions. In some instances, the sample (e.g., cells or a cell solution) is further incubated in the presence of a small molecule fragment prior to addition of the cysteine-reactive probe. In some instances, the sample is compared with a control. In some instances, the control comprises the cysteine-reactive probe but not the small molecule fragment. In some instances, a difference is observed between a set of cysteine-reactive probe protein interactions between the sample and the control. In some instances, the difference correlates to the interaction between the small molecule fragment and the cysteine containing proteins.

[00254] In some embodiments, the sample (e.g. cells or a cell solution) is further labeled for analysis of cysteine-reactive probe protein interactions. In some instances, the sample (e.g. cells or a cell solution) is labeled with an enriched media. In some cases, the sample (e.g. cells or a cell solution) is labeled with isotope-labeled amino acids, such as ^{13}C or ^{15}N -labeled amino acids. In some cases, the labeled sample is further compared with a non-labeled sample to detect differences in cysteine-reactive probe protein interactions between the two samples. In some instances, this difference is a difference of a cysteine containing protein and its interaction with a small molecule fragment in the labeled sample versus the non-labeled sample. In some instances, the difference is an increase, decrease or a lack of protein cysteine-reactive probe interaction in the two samples. In some instances, the isotope-labeled method is termed SILAC, stable isotope labeling using amino acids in cell culture.

[00255] In some instances, the sample is divided into a first cell solution and a second cell solution. In some cases, the first cell solution is incubated with a small molecule fragment for an extended period of time prior to incubating the first cell solution with a first cysteine-reactive probe to generate a first group of cysteine-reactive probe-protein complexes. In some instances, the extended period of time is about 5, 10, 15, 20, 30, 60, 90, 120 minutes or longer. In some instances, the second cell solution comprises a second cysteine-reactive probe to generate a second group of cysteine-reactive probe-protein complexes. In some instances, the first cysteine-reactive probe and the second cysteine-reactive probe are the same. In some embodiments, cells from the second cell solution are further treated with a buffer, such as a control buffer, in which

the buffer does not contain a small molecule fragment. In some embodiments, the control buffer comprises dimethyl sulfoxide (DMSO).

[00256] In some embodiments, the cysteine-reactive probe-protein complex is further conjugated to a chromophore, such as a fluorophore. In some instances, the cysteine-reactive probe-protein complex is separated and visualized utilizing an electrophoresis system, such as through a gel electrophoresis, or a capillary electrophoresis. Exemplary gel electrophoresis includes agarose based gels, polyacrylamide based gels, or starch based gels. In some instances, the cysteine-reactive probe-protein is subjected to a native electrophoresis condition. In some instances, the cysteine-reactive probe-protein is subjected to a denaturing electrophoresis condition.

[00257] In some instances, the cysteine-reactive probe-protein after harvesting is further fragmentized to generate protein fragments. In some instances, fragmentation is generated through mechanical stress, pressure, or chemical means. In some instances, the protein from the cysteine-reactive probe-protein complexes is fragmented by a chemical means. In some embodiments, the chemical means is a protease. Exemplary proteases include, but are not limited to, serine proteases such as chymotrypsin A, penicillin G acylase precursor, dipeptidase E, DmpA aminopeptidase, subtilisin, prolyl oligopeptidase, D-Ala-D-Ala peptidase C, signal peptidase I, cytomegalovirus assemblin, Lon-A peptidase, peptidase Clp, Escherichia coli phage K1F endosialidase CIMCD self-cleaving protein, nucleoporin 145, lactoferrin, murein tetrapeptidase LD-carboxypeptidase, or rhomboid-1; threonine proteases such as ornithine acetyltransferase; cysteine proteases such as TEV protease, amidophosphoribosyltransferase precursor, gamma-glutamyl hydrolase (*Rattus norvegicus*), hedgehog protein, DmpA aminopeptidase, papain, bromelain, cathepsin K, calpain, caspase-1, separase, adenain, pyroglutamyl-peptidase I, sortase A, hepatitis C virus peptidase 2, sindbis virus-type nsP2 peptidase, dipeptidyl-peptidase VI, or DeSI-1 peptidase; aspartate proteases such as beta-secretase 1 (BACE1), beta-secretase 2 (BACE2), cathepsin D, cathepsin E, chymosin, napsin-A, nepenthesin, pepsin, plasmepsin, presenilin, or renin; glutamic acid proteases such as AfuGprA; and metalloproteases such as peptidase_M48.

[00258] In some instances, the fragmentation is a random fragmentation. In some instances, the fragmentation generates specific lengths of protein fragments, or the shearing occurs at particular sequence of amino acid regions.

[00259] In some instances, the protein fragments are further analyzed by a proteomic method such as by liquid chromatography (LC) (e.g. high performance liquid chromatography), liquid chromatography-mass spectrometry (LC-MS), matrix-assisted laser desorption/ionization

(MALDI-TOF), gas chromatography-mass spectrometry (GC-MS), capillary electrophoresis-mass spectrometry (CE-MS), or nuclear magnetic resonance imaging (NMR).

[00260] In some embodiments, the LC method is any suitable LC methods well known in the art, for separation of a sample into its individual parts. This separation occurs based on the interaction of the sample with the mobile and stationary phases. Since there are many stationary/mobile phase combinations that are employed when separating a mixture, there are several different types of chromatography that are classified based on the physical states of those phases. In some embodiments, the LC is further classified as normal-phase chromatography, reverse-phase chromatography, size-exclusion chromatography, ion-exchange chromatography, affinity chromatography, displacement chromatography, partition chromatography, flash chromatography, chiral chromatography, and aqueous normal-phase chromatography.

[00261] In some embodiments, the LC method is a high performance liquid chromatography (HPLC) method. In some embodiments, the HPLC method is further categorized as normal-phase chromatography, reverse-phase chromatography, size-exclusion chromatography, ion-exchange chromatography, affinity chromatography, displacement chromatography, partition chromatography, chiral chromatography, and aqueous normal-phase chromatography.

[00262] In some embodiments, the HPLC method of the present disclosure is performed by any standard techniques well known in the art. Exemplary HPLC methods include hydrophilic interaction liquid chromatography (HILIC), electrostatic repulsion-hydrophilic interaction liquid chromatography (ERLIC) and reverse phase liquid chromatography (RPLC).

[00263] In some embodiments, the LC is coupled to a mass spectroscopy as a LC-MS method. In some embodiments, the LC-MS method includes ultra-performance liquid chromatography-electrospray ionization quadrupole time-of-flight mass spectrometry (UPLC-ESI-QTOF-MS), ultra-performance liquid chromatography-electrospray ionization tandem mass spectrometry (UPLC-ESI-MS/MS), reverse phase liquid chromatography-mass spectrometry (RPLC-MS), hydrophilic interaction liquid chromatography-mass spectrometry (HILIC-MS), hydrophilic interaction liquid chromatography-triple quadrupole tandem mass spectrometry (HILIC-QQQ), electrostatic repulsion-hydrophilic interaction liquid chromatography-mass spectrometry (ERLIC-MS), liquid chromatography time-of-flight mass spectrometry (LC-QTOF-MS), liquid chromatography-tandem mass spectrometry (LC-MS/MS), multidimensional liquid chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS). In some instances, the LC-MS method is LC/LC-MS/MS. In some embodiments, the LC-MS methods of the present disclosure are performed by standard techniques well known in the art.

[00264] In some embodiments, the GC is coupled to a mass spectroscopy as a GC-MS method. In some embodiments, the GC-MS method includes two-dimensional gas

chromatography time-of-flight mass spectrometry (GC*GC-TOFMS), gas chromatography time-of-flight mass spectrometry (GC-QTOF-MS) and gas chromatography-tandem mass spectrometry (GC-MS/MS).

[00265] In some embodiments, CE is coupled to a mass spectroscopy as a CE-MS method. In some embodiments, the CE-MS method includes capillary electrophoresis- negative electrospray ionization-mass spectrometry (CE-ESI-MS), capillary electrophoresis-negative electrospray ionization-quadrupole time of flight-mass spectrometry (CE-ESI-QTOF-MS) and capillary electrophoresis-quadrupole time of flight-mass spectrometry (CE-QTOF-MS).

[00266] In some embodiments, the nuclear magnetic resonance (NMR) method is any suitable method well known in the art for the detection of one or more cysteine binding proteins or protein fragments disclosed herein. In some embodiments, the NMR method includes one dimensional (1D) NMR methods, two dimensional (2D) NMR methods, solid state NMR methods and NMR chromatography. Exemplary 1D NMR methods include ¹Hydrogen, ¹³Carbon, ¹⁵Nitrogen, ¹⁷Oxygen, ¹⁹Fluorine, ³¹Phosphorus, ³⁹Potassium, ²³Sodium, ³³Sulfur, ⁸⁷Strontium, ²⁷Aluminium, ⁴³Calcium, ³⁵Chlorine, ³⁷Chlorine, ⁶³Copper, ⁶⁵Copper, ⁵⁷Iron, ²⁵Magnesium, ¹⁹⁹Mercury or ⁶⁷Zinc NMR method, distortionless enhancement by polarization transfer (DEPT) method, attached proton test (APT) method and 1D-incredible natural abundance double quantum transition experiment (INADEQUATE) method. Exemplary 2D NMR methods include correlation spectroscopy (COSY), total correlation spectroscopy (TOCSY), 2D-INADEQUATE, 2D-adequate double quantum transfer experiment (ADEQUATE), nuclear overhauser effect spectroscopy (NOSEY), rotating-frame NOE spectroscopy (ROESY), heteronuclear multiple-quantum correlation spectroscopy (HMQC), heteronuclear single quantum coherence spectroscopy (HSQC), short range coupling and long range coupling methods. Exemplary solid state NMR method include solid state ¹³Carbon NMR, high resolution magic angle spinning (HR-MAS) and cross polarization magic angle spinning (CP-MAS) NMR methods. Exemplary NMR techniques include diffusion ordered spectroscopy (DOSY), DOSY-TOCSY and DOSY-HSQC.

[00267] In some embodiments, the protein fragments are analyzed by method as described in Weerapana et al., “Quantitative reactivity profiling predicts functional cysteines in proteomes,” *Nature*, 468:790-795 (2010).

[00268] In some embodiments, the results from the mass spectroscopy method are analyzed by an algorithm for protein identification. In some embodiments, the algorithm combines the results from the mass spectroscopy method with a protein sequence database for protein identification. In some embodiments, the algorithm comprises ProLuCID algorithm, Probit, Scaffold, SEQUEST, or Mascot.

[00269] In some embodiments, a value is assigned to each of the protein from the cysteine-reactive probe-protein complex. In some embodiments, the value assigned to each of the protein from the cysteine-reactive probe-protein complex is obtained from the mass spectroscopy analysis. In some instances, the value is the area-under-the curve from a plot of signal intensity as a function of mass-to-charge ratio. In some embodiments, a first value is assigned to the protein obtained from the first cell solution and a second value is assigned to the same protein obtained from the second cell solution. In some instances, a ratio is calculated between the two values. In some instances, a ratio of greater than 2 indicates that the protein is a candidate for interacting with a drug or that the protein is a cysteine binding protein. In some instances, the ratio is greater than 2.5, 3, 3.5, 4, 4.5, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20. In some cases, the ratio is at most 20.

[00270] In some instances, the ratio is calculated based on averaged values. In some instances, the averaged value is an average of at least two, three, or four values of the protein from each cell solution, or that the protein is observed at least two, three, or four times in each cell solution and a value is assigned to each observed time. In some instances, the ratio further has a standard deviation of less than 12, 10, or 8.

[00271] In some instances, a value is not an averaged value. In some instances, the ratio is calculated based on value of a protein observed only once in a cell population. In some instances, the ratio is assigned with a value of 20.

[00272] In some embodiments, in the context of identifying a cysteine containing protein as a small fragment molecule binding target, a first ratio is obtained from two cell solutions in which both cell solutions have been incubated with a cysteine-reactive probe and the first cell solution is further incubated with a small molecule fragment. In some instances, the first ratio is further compared to a second ratio in which both cell solutions have been treated by cysteine-reactive probes in the absence of a small molecule fragment. In some instances, the first ratio is greater than 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20. In some instances, the second ratio is greater than 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20. In some instances, if the first ratio is greater than 2, 2.5, 3, 3.5, 4, 4.5, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 and the second ratio is from about 0.5 to about 2, the two ratios indicate that a protein is a drug binding target.

[00273] In some embodiments, the value further enables calculating a percentage of inhibition of the cysteine-reactive probe to the cysteine containing protein. In some embodiments, the percentage of inhibition of greater than 50%, 60%, 70%, 80%, 90%, or at 100% indicates that the cysteine containing protein is a candidate for interacting with the small molecule fragment.

Kits/Article of Manufacture

[00274] Disclosed herein, in certain embodiments, are kits and articles of manufacture for use with one or more methods described herein. In some embodiments, described herein is a kit for identifying a cysteine containing protein as a small molecule fragment binding target. In some instances, also described herein is a kit for mapping binding sites on a cysteine containing protein. In some cases, described herein is a kit for identifying cysteine binding proteins. In some embodiments, also described herein is a kit for a high throughput screening of a small molecule fragment for interaction with a cysteine containing protein.

[00275] In some embodiments, such kit includes cysteine-reactive probes such as the cysteine-reactive probes described herein, test compounds such as small molecule fragments or libraries and/or controls, and reagents suitable for carrying out one or more of the methods described herein. In some instances, the kit further comprises samples, such as a cell sample, and suitable solutions such as buffers or media. In some embodiments, the kit further comprises recombinant proteins for use in one or more of the methods described herein. In some embodiments, additional components of the kit comprises a carrier, package, or container that is compartmentalized to receive one or more containers such as vials, tubes, and the like, each of the container(s) comprising one of the separate elements to be used in a method described herein. Suitable containers include, for example, bottles, vials, plates, syringes, and test tubes. In one embodiment, the containers are formed from a variety of materials such as glass or plastic.

[00276] The articles of manufacture provided herein contain packaging materials. Examples of pharmaceutical packaging materials include, but are not limited to, bottles, tubes, bags, containers, and any packaging material suitable for a selected formulation and intended mode of use.

[00277] For example, the container(s) include cysteine-reactive probes, test compounds, and one or more reagents for use in a method disclosed herein. Such kits optionally include an identifying description or label or instructions relating to its use in the methods described herein.

[00278] A kit typically includes labels listing contents and/or instructions for use, and package inserts with instructions for use. A set of instructions will also typically be included.

[00279] In one embodiment, a label is on or associated with the container. In one embodiment, a label is on a container when letters, numbers or other characters forming the label are attached, molded or etched into the container itself; a label is associated with a container when it is present within a receptacle or carrier that also holds the container, e.g., as a package insert. In one embodiment, a label is used to indicate that the contents are to be used for a specific therapeutic application. The label also indicates directions for use of the contents, such as in the methods described herein.

Services

[00280] In some embodiments, the methods provided herein also perform as a service. In some instances, a service provider obtain from the customer a plurality of small molecule fragment candidates for analysis with one or more of the cysteine-reactive probes for screening. In some embodiments, the service provider screens the small molecule fragment candidates using one or more of the methods described herein, and then provide the results to the customer. In some instances, the service provider provides the appropriate reagents to the customer for analysis utilizing one or more of the cysteine-reactive probes and one or more of the methods described herein. In some cases, the customer performs one or more of the methods described herein and then provide the results to the service provider for analysis. In some embodiments, the service provider then analyzes the results and provides the results to the costumer. In some cases, the customer further analyze the results by interacting with software installed locally (at the customer's location) or remotely (e.g., on a server reachable through a network). Exemplary customers include pharmaceutical companies, clinical laboratories, physicians, patients, and the like. In some instances, a customer is any suitable customer or party with a need or desire to use the methods, systems, compositions, and kits described herein.

Digital Processing Device

[00281] In some embodiments, the methods described herein include a digital processing device, or use of the same. In further embodiments, the digital processing device includes one or more hardware central processing units (CPU) that carry out the device's functions. In still further embodiments, the digital processing device further comprises an operating system configured to perform executable instructions. In some embodiments, the digital processing device is optionally connected to a computer network. In further embodiments, the digital processing device is optionally connected to the Internet such that it accesses the World Wide Web. In still further embodiments, the digital processing device is optionally connected to a cloud computing infrastructure. In other embodiments, the digital processing device is optionally connected to an intranet. In other embodiments, the digital processing device is optionally connected to a data storage device.

[00282] In accordance with the description herein, suitable digital processing devices include, by are not limited to, server computers, desktop computers, laptop computers, notebook computers, sub-notebook computers, netbook computers, netpad computers, set-top computers, media streaming devices, handheld computers, Internet appliances, mobile smartphones, tablet computers, personal digital assistants, video game consoles, and vehicles. Suitable tablet computers include those with booklet, slate, or convertible configurations.

[00283] In some embodiments, the digital processing device includes an operating system configured to perform executable instructions. The operating system is, for example, software, including programs and data, which manages the device's hardware and provides services for execution of applications. Suitable server operating systems include, by way of non-limiting examples, FreeBSD, OpenBSD, NetBSD[®], Linux, Apple[®] Mac OS X Server[®], Oracle[®] Solaris[®], Windows Server[®], and Novell[®] NetWare[®]. Suitable personal computer operating systems include, by way of non-limiting examples, Microsoft[®] Windows[®], Apple[®] Mac OS X[®], UNIX[®], and UNIX-like operating systems such as GNU/Linux[®]. In some embodiments, the operating system is provided by cloud computing. Suitable mobile smart phone operating systems include, by way of non-limiting examples, Nokia[®] Symbian[®] OS, Apple[®] iOS[®], Research In Motion[®] BlackBerry OS[®], Google[®] Android[®], Microsoft[®] Windows Phone[®] OS, Microsoft[®] Windows Mobile[®] OS, Linux[®], and Palm[®] WebOS[®]. Suitable media streaming device operating systems include, by way of non-limiting examples, Apple TV[®], Roku[®], Boxee[®], Google TV[®], Google Chromecast[®], Amazon Fire[®], and Samsung[®] HomeSync[®]. Suitable video game console operating systems include, by way of non-limiting examples, Sony[®] PS3[®], Sony[®] PS4[®], Microsoft[®] Xbox 360[®], Microsoft Xbox One, Nintendo[®] Wii[®], Nintendo[®] Wii U[®], and Ouya[®].

[00284] In some embodiments, the device includes a storage and/or memory device. The storage and/or memory device is one or more physical apparatuses used to store data or programs on a temporary or permanent basis. In some embodiments, the device is volatile memory and requires power to maintain stored information. In some embodiments, the device is non-volatile memory and retains stored information when the digital processing device is not powered. In further embodiments, the non-volatile memory comprises flash memory. In some embodiments, the non-volatile memory comprises dynamic random-access memory (DRAM). In some embodiments, the non-volatile memory comprises ferroelectric random access memory (FRAM). In some embodiments, the non-volatile memory comprises phase-change random access memory (PRAM). In other embodiments, the device is a storage device including, by way of non-limiting examples, CD-ROMs, DVDs, flash memory devices, magnetic disk drives, magnetic tapes drives, optical disk drives, and cloud computing based storage. In further embodiments, the storage and/or memory device is a combination of devices such as those disclosed herein.

[00285] In some embodiments, the digital processing device includes a display to send visual information to a user. In some embodiments, the display includes a cathode ray tube (CRT), a liquid crystal display (LCD), a thin film transistor liquid crystal display (TFT-LCD), an organic light emitting diode (OLED) display, a plasma display, a video projector, or a combination thereof.

[00286] In some embodiments, the digital processing device includes an input device to receive information from a user. In some embodiments, the input device is a keyboard. In some embodiments, the input device is a pointing device including, by way of non-limiting examples, a mouse, trackball, track pad, joystick, game controller, or stylus. In some embodiments, the input device is a touch screen or a multi-touch screen. In other embodiments, the input device is a microphone to capture voice or other sound input. In other embodiments, the input device is a video camera or other sensor to capture motion or visual input. In further embodiments, the input device is a Kinect™, Leap Motion™, or the like. In still further embodiments, the input device is a combination of devices such as those disclosed herein.

[00287] In some embodiments, the systems and methods disclosed herein include one or more non-transitory computer readable storage media encoded with a program including instructions executable by the operating system of an optionally networked digital processing device. In further embodiments, a computer readable storage medium is a tangible component of a digital processing device. In still further embodiments, a computer readable storage medium is optionally removable from a digital processing device. In some embodiments, a computer readable storage medium includes, by way of non-limiting examples, CD-ROMs, DVDs, flash memory devices, solid state memory, magnetic disk drives, magnetic tape drives, optical disk drives, cloud computing systems and services, and the like. In some cases, the program and instructions are permanently, substantially permanently, semi-permanently, or non-transitorily encoded on the media.

[00288] In some embodiments, the systems and methods disclosed herein include at least one computer program, or use of the same. A computer program includes a sequence of instructions, executable in the digital processing device's CPU, written to perform a specified task. In some embodiments, computer readable instructions are implemented as program modules, such as functions, objects, Application Programming Interfaces (APIs), data structures, and the like, that perform particular tasks or implement particular abstract data types.

[00289] In some embodiments, the functionality of the computer readable instructions are combined or distributed as desired in various environments. In some embodiments, a computer program comprises one sequence of instructions. In some embodiments, a computer program comprises a plurality of sequences of instructions. In some embodiments, a computer program is provided from one location. In other embodiments, a computer program is provided from a plurality of locations. In various embodiments, a computer program includes one or more software modules. In various embodiments, a computer program includes, in part or in whole, one or more web applications, one or more mobile applications, one or more standalone

applications, one or more web browser plug-ins, extensions, add-ins, or add-ons, or combinations thereof.

[00290] In some embodiments, a computer program includes a web application. A web application, in various embodiments, utilizes one or more software frameworks and one or more database systems. In some embodiments, a web application is created upon a software framework such as Microsoft[®] .NET or Ruby on Rails (RoR). In some embodiments, a web application utilizes one or more database systems including, by way of non-limiting examples, relational, non-relational, object oriented, associative, and XML database systems. In further embodiments, suitable relational database systems include, by way of non-limiting examples, Microsoft[®] SQL Server, MySQL[™], and Oracle[®]. A web application, in various embodiments, is written in one or more versions of one or more languages. In some embodiments, a web application is written in one or more markup languages, presentation definition languages, client-side scripting languages, server-side coding languages, database query languages, or combinations thereof. In some embodiments, a web application is written to some extent in a markup language such as Hypertext Markup Language (HTML), Extensible Hypertext Markup Language (XHTML), or eXtensible Markup Language (XML). In some embodiments, a web application is written to some extent in a presentation definition language such as Cascading Style Sheets (CSS). In some embodiments, a web application is written to some extent in a client-side scripting language such as Asynchronous Javascript and XML (AJAX), Flash[®] Actionscript, Javascript, or Silverlight[®]. In some embodiments, a web application is written to some extent in a server-side coding language such as Active Server Pages (ASP), ColdFusion[®], Perl, Java[™], JavaServer Pages (JSP), Hypertext Preprocessor (PHP), Python[™], Ruby, Tcl, Smalltalk, WebDNA[®], or Groovy. In some embodiments, a web application is written to some extent in a database query language such as Structured Query Language (SQL). In some embodiments, a web application integrates enterprise server products such as IBM[®] Lotus Domino[®]. In some embodiments, a web application includes a media player element. In various further embodiments, a media player element utilizes one or more of many suitable multimedia technologies including, by way of non-limiting examples, Adobe[®] Flash[®], HTML 5, Apple[®] QuickTime[®], Microsoft[®] Silverlight[®], Java[™], and Unity[®].

[00291] In some embodiments, a computer program includes a mobile application provided to a mobile digital processing device. In some embodiments, the mobile application is provided to a mobile digital processing device at the time it is manufactured. In other embodiments, the mobile application is provided to a mobile digital processing device via the computer network described herein.

[00292] In view of the disclosure provided herein, a mobile application is created by techniques using hardware, languages, and development environments. Suitable programming languages include, by way of non-limiting examples, C, C++, C#, Objective-C, Java™, Javascript, Pascal, Object Pascal, Python™, Ruby, VB.NET, WML, and XHTML/HTML with or without CSS, or combinations thereof.

[00293] Suitable mobile application development environments are available from several sources. Commercially available development environments include, by way of non-limiting examples, AirplaySDK, alcheMo, Appcelerator®, Celsius, Bedrock, Flash Lite, .NET Compact Framework, Rhomobile, and WorkLight Mobile Platform. Other development environments are available without cost including, by way of non-limiting examples, Lazarus, MobiFlex, MoSync, and Phonegap. Also, mobile device manufacturers distribute software developer kits including, by way of non-limiting examples, iPhone and iPad (iOS) SDK, Android™ SDK, BlackBerry® SDK, BREW SDK, Palm® OS SDK, Symbian SDK, webOS SDK, and Windows® Mobile SDK.

[00294] In some embodiments, commercial forums for distribution of mobile applications include, by way of non-limiting examples, Apple® App Store, Android™ Market, BlackBerry® App World, App Store for Palm devices, App Catalog for webOS, Windows® Marketplace for Mobile, Ovi Store for Nokia® devices, Samsung® Apps, and Nintendo® DSi Shop.

[00295] In some embodiments, a computer program includes a standalone application, which is a program that is run as an independent computer process, not an add-on to an existing process, e.g., not a plug-in. In some instances, standalone applications are compiled. A compiler is a computer program(s) that transforms source code written in a programming language into binary object code such as assembly language or machine code. Suitable compiled programming languages include, by way of non-limiting examples, C, C++, Objective-C, COBOL, Delphi, Eiffel, Java™, Lisp, Python™, Visual Basic, and VB .NET, or combinations thereof.

Compilation is often performed, at least in part, to create an executable program. In some embodiments, a computer program includes one or more executable compiled applications.

[00296] In some embodiments, the computer program includes a web browser plug-in. In computing, a plug-in is one or more software components that add specific functionality to a larger software application. Makers of software applications support plug-ins to enable third-party developers to create abilities which extend an application, to support easily adding new features, and to reduce the size of an application. When supported, plug-ins enable customizing the functionality of a software application. For example, plug-ins are commonly used in web browsers to play video, generate interactivity, scan for viruses, and display particular file types. In some instances, web browser plug-ins include Adobe® Flash® Player, Microsoft® Silverlight®,

and Apple[®] QuickTime[®]. In some embodiments, the toolbar comprises one or more web browser extensions, add-ins, or add-ons. In some embodiments, the toolbar comprises one or more explorer bars, tool bands, or desk bands.

[00297] In view of the disclosure provided herein, plug-in frameworks are available that enable development of plug-ins in various programming languages, including, by way of non-limiting examples, C++, Delphi, Java[™], PHP, Python[™], and VB .NET, or combinations thereof.

[00298] Web browsers (also called Internet browsers) are software applications, designed for use with network-connected digital processing devices, for retrieving, presenting, and traversing information resources on the World Wide Web. Suitable web browsers include, by way of non-limiting examples, Microsoft[®] Internet Explorer[®], Mozilla[®] Firefox[®], Google[®] Chrome, Apple[®] Safari[®], Opera Software[®] Opera[®], and KDE Konqueror. In some embodiments, the web browser is a mobile web browser. Mobile web browsers (also called microbrowsers, mini-browsers, and wireless browsers) are designed for use on mobile digital processing devices including, by way of non-limiting examples, handheld computers, tablet computers, netbook computers, subnotebook computers, smartphones, music players, personal digital assistants (PDAs), and handheld video game systems. Suitable mobile web browsers include, by way of non-limiting examples, Google[®] Android[®] browser, RIM BlackBerry[®] Browser, Apple[®] Safari[®], Palm[®] Blazer, Palm[®] WebOS[®] Browser, Mozilla[®] Firefox[®] for mobile, Microsoft[®] Internet Explorer[®] Mobile, Amazon[®] Kindle[®] Basic Web, Nokia[®] Browser, Opera Software[®] Opera[®] Mobile, and Sony[®] PSP[™] browser.

[00299] In some embodiments, the systems and methods disclosed herein include software, server, and/or database modules, or use of the same. In view of the disclosure provided herein, software modules are created and implemented in a multitude of ways. In various embodiments, a software module comprises a file, a section of code, a programming object, a programming structure, or combinations thereof. In further various embodiments, a software module comprises a plurality of files, a plurality of sections of code, a plurality of programming objects, a plurality of programming structures, or combinations thereof. In various embodiments, the one or more software modules comprise, by way of non-limiting examples, a web application, a mobile application, and a standalone application. In some embodiments, software modules are in one computer program or application. In other embodiments, software modules are in more than one computer program or application. In some embodiments, software modules are hosted on one machine. In other embodiments, software modules are hosted on more than one machine. In further embodiments, software modules are hosted on cloud computing platforms. In some

embodiments, software modules are hosted on one or more machines in one location. In other embodiments, software modules are hosted on one or more machines in more than one location.

[00300] In some embodiments, the methods and systems disclosed herein include one or more databases, or use of the same. In view of the disclosure provided herein, databases are suitable for storage and retrieval of analytical information described elsewhere herein. In various embodiments, suitable databases include, by way of non-limiting examples, relational databases, non-relational databases, object oriented databases, object databases, entity-relationship model databases, associative databases, and XML databases. In some embodiments, a database is internet-based. In further embodiments, a database is web-based. In still further embodiments, a database is cloud computing-based. In other embodiments, a database is based on one or more local computer storage devices.

Server

[00301] In some embodiments, the methods provided herein are processed on a server or a computer server (Fig. 2). In some embodiments, the server **401** includes a central processing unit (CPU, also “processor”) **405** which is a single core processor, a multi core processor, or plurality of processors for parallel processing. In some embodiments, a processor used as part of a control assembly is a microprocessor. In some embodiments, the server **401** also includes memory **410** (e.g. random access memory, read-only memory, flash memory); electronic storage unit **415** (e.g. hard disk); communications interface **420** (e.g. network adaptor) for communicating with one or more other systems; and peripheral devices **425** which includes cache, other memory, data storage, and/or electronic display adaptors. The memory **410**, storage unit **415**, interface **420**, and peripheral devices **425** are in communication with the processor **405** through a communications bus (solid lines), such as a motherboard. In some embodiments, the storage unit **415** is a data storage unit for storing data. The server **401** is operatively coupled to a computer network (“network”) **430** with the aid of the communications interface **420**. In some embodiments, a processor with the aid of additional hardware is also operatively coupled to a network. In some embodiments, the network **430** is the Internet, an intranet and/or an extranet, an intranet and/or extranet that is in communication with the Internet, a telecommunication or data network. In some embodiments, the network **430** with the aid of the server **401**, implements a peer-to-peer network, which enables devices coupled to the server **401** to behave as a client or a server. In some embodiments, the server is capable of transmitting and receiving computer-readable instructions (e.g., device/system operation protocols or parameters) or data (e.g., sensor measurements, raw data obtained from detecting metabolites, analysis of raw data obtained from detecting metabolites, interpretation of raw data obtained from detecting metabolites, etc.) via electronic signals transported through the network **430**. Moreover, in some

embodiments, a network is used, for example, to transmit or receive data across an international border.

[00302] In some embodiments, the server **401** is in communication with one or more output devices **435** such as a display or printer, and/or with one or more input devices **440** such as, for example, a keyboard, mouse, or joystick. In some embodiments, the display is a touch screen display, in which case it functions as both a display device and an input device. In some embodiments, different and/or additional input devices are present such as an enunciator, a speaker, or a microphone. In some embodiments, the server uses any one of a variety of operating systems, such as for example, any one of several versions of Windows®, or of MacOS®, or of Unix®, or of Linux®.

[00303] In some embodiments, the storage unit **415** stores files or data associated with the operation of a device, systems or methods described herein.

[00304] In some embodiments, the server communicates with one or more remote computer systems through the network **430**. In some embodiments, the one or more remote computer systems include, for example, personal computers, laptops, tablets, telephones, Smart phones, or personal digital assistants.

[00305] In some embodiments, a control assembly includes a single server **401**. In other situations, the system includes multiple servers in communication with one another through an intranet, extranet and/or the Internet.

[00306] In some embodiments, the server **401** is adapted to store device operation parameters, protocols, methods described herein, and other information of potential relevance. In some embodiments, such information is stored on the storage unit **415** or the server **401** and such data is transmitted through a network.

Certain Terminology

[00307] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of skill in the art to which the claimed subject matter belongs. It is to be understood that the foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of any subject matter claimed. In this application, the use of the singular includes the plural unless specifically stated otherwise. It must be noted that, as used in the specification and the appended claims, the singular forms “a,” “an” and “the” include plural referents unless the context clearly dictates otherwise. In this application, the use of “or” means “and/or” unless stated otherwise. Furthermore, use of the term “including” as well as other forms, such as “include,” “includes,” and “included,” is not limiting.

[00308] As used herein, ranges and amounts can be expressed as “about” a particular value or range. About also includes the exact amount. Hence “about 5 μ L” means “about 5 μ L” and also “5 μ L.” Generally, the term “about” includes an amount that would be expected to be within experimental error.

[00309] The section headings used herein are for organizational purposes only and are not to be construed as limiting the subject matter described.

[00310] The term “protein”, as used herein, encompasses a full-length cysteine containing protein, a full-length functional cysteine containing protein, a cysteine containing protein fragment, or a functionally active cysteine containing protein fragment. In some instances, a protein described herein is also referred to as an “isolated protein”, or a protein that by virtue of its origin or source of derivation is not associated with naturally associated components that accompany it in its native state; is substantially free of other proteins from the same species; is expressed by a cell from a different species; or does not occur in nature.

[00311] The term “polypeptide”, as used herein, refers to any polymeric chain of amino acids. The term “polypeptide” encompasses native or modified cysteine containing protein, cysteine containing protein fragments, or polypeptide analogs comprising non-native amino acid residues. In some instances, a polypeptide is monomeric. In other instances, a polypeptide is polymeric. In some instances, a polypeptide described herein is also referred to as an “isolated polypeptide”, or a polypeptide that by virtue of its origin or source of derivation is not associated with naturally associated components that accompany it in its native state; is substantially free of other proteins from the same species; is expressed by a cell from a different species; or does not occur in nature.

[00312] As used herein, the terms “individual(s)”, “subject(s)” and “patient(s)” mean any mammal. In some embodiments, the mammal is a human. In some embodiments, the mammal is a non-human. None of the terms require or are limited to situations characterized by the supervision (e.g. constant or intermittent) of a health care worker (e.g. a doctor, a registered nurse, a nurse practitioner, a physician’s assistant, an orderly or a hospice worker).

[00313] The term “alkyl” as used herein is a branched or unbranched saturated hydrocarbon group of 1 to 24 carbon atoms, such as methyl, ethyl, n-propyl, isopropyl, n-butyl, isobutyl, s-butyl, t-butyl, n-pentyl, isopentyl, s-pentyl, neopentyl, hexyl, heptyl, octyl, nonyl, decyl, dodecyl, tetradecyl, hexadecyl, eicosyl, tetracosyl, and the like. It is understood that the alkyl group is acyclic. In some instances, the alkyl group is branched or unbranched. In some instances, the alkyl group is also substituted or unsubstituted. For example, the alkyl group is substituted with one or more groups including, but not limited to, alkyl, cycloalkyl, alkoxy, amino, ether, halide, hydroxy, nitro, silyl, sulfo-oxo, or thiol. A “lower alkyl” group is an alkyl group containing from

one to six (e.g., from one to four) carbon atoms. In some instances, the term alkyl group is also a C1 alkyl, C1-C2 alkyl, C1-C3 alkyl, C1-C4 alkyl, C1-C5 alkyl, C1-C6 alkyl, C1-C7 alkyl, C1-C8 alkyl, C1-C9 alkyl, C1-C10 alkyl, and the like up to and including a C1-C24 alkyl.

[00314] The term “aryl” as used herein is a group that contains any carbon-based aromatic group including, but not limited to, benzene, naphthalene, phenyl, biphenyl, anthracene, and the like. The aryl group can be substituted or unsubstituted. In some instances, the aryl group is substituted with one or more groups including, but not limited to, alkyl, cycloalkyl, alkoxy, alkenyl, cycloalkenyl, alkynyl, cycloalkynyl, aryl, heteroaryl, aldehyde, —NH₂, carboxylic acid, ester, ether, halide, hydroxy, ketone, azide, nitro, silyl, sulfo-oxo, or thiol. The term “biaryl” is a specific type of aryl group and is included in the definition of “aryl.” In addition, the aryl group is optionally a single ring structure or comprises multiple ring structures that are either fused ring structures or attached via one or more bridging groups such as a carbon-carbon bond. For example, biaryl refers to two aryl groups that are bound together via a fused ring structure, as in naphthalene, or are attached via one or more carbon-carbon bonds, as in biphenyl.

EXAMPLES

[00315] These examples are provided for illustrative purposes only and not to limit the scope of the claims provided herein.

Example 1

Biological Methods

Preparation of human cancer cell line proteomes

[00316] All cell lines were obtained from ATCC, were used with a low passage number and were grown at 37 °C with 5% CO₂. MDA-MB-231 cells and HEK-293T cells were grown in DMEM supplemented with 10% fetal bovine serum, penicillin, streptomycin and glutamine. Jurkat, Ramos and MUM2C cells were grown in RPMI-1640 medium supplemented with 10% fetal bovine serum, penicillin and streptomycin. For *in vitro* labeling, cells were grown to 100% confluence for MDA-MB-231 cells or until cell density reached 1.5 million cells/mL for Ramos and Jurkat cells. Cells were washed with cold PBS, scraped with cold PBS and cell pellets were isolated by centrifugation (1,400 g, 3 min, 4 °C), and stored at -80 °C until use. Cell pellets were lysed by sonication and fractionated (100,000 g, 45 min) to yield soluble and membrane fractions, which were then adjusted to a final protein concentration of 1.5 mg/mL for proteomics experiments and 1 mg/mL for gel-based ABPP experiments. The soluble lysate was prepared fresh from frozen pellets directly before each experiment. Protein concentration was determined using the Bio-Rad DC™ protein assay kit.

Screening of fragment electrophile library by gel-based ABPP with IA-rhodamine and Ac-Rho-DEVD-AMK ("DEVD" disclosed as SEQ ID NO: 857)

[00317] 25 μ L of soluble proteome (1 mg/mL) was treated with fragment electrophiles (1 μ L of 25 $\times$ stock solution in DMSO) at ambient temperature for 1 h. IA-rhodamine (1 μ L of 25 μ M, final concentration = 1 μ M) was then added and allowed to react for an additional 1 h. The reactions were quenched with 8 μ L of 4 $\times$ SDS-PAGE loading buffer and the quenched samples analyzed by SDS-PAGE (10% polyacrylamide; 15 μ L of sample/lane) and visualized by in-gel fluorescence using a flatbed fluorescent scanner (BioRad ChemiDoc™ MP or Hitachi FMBio IIe). To measure labeling of recombinant proteins expressed in *E. coli*, purified protein was added to soluble proteome to a final concentration of 1 μ M (CASP8, PRMT1, IMPDH2), 2 μ M (TIGAR, IDH1) or 4 μ M (IDH1 R132H) and the proteomes were treated as detailed above. IDH1 labeling by IA-rhodamine is relatively better in MDA-MB-231 soluble proteome when compared with Ramos and Jurkat soluble proteome. Recombinant, active CASP8 in soluble proteome was labeled with Rho-DEVD-AOMK ("DEVD" disclosed as SEQ ID NO: 857) (1 μ L of 50 μ M, final concentration = 2 μ M), quenched and analyzed by SDS-PAGE on 14% polyacrylamide gels.

Gel-based ABPP with alkyne-containing click probes

[00318] 25 μ L of soluble proteome (1 mg/mL) was labeled with the indicated concentration of **18** or **19** (1 μ L of 25 $\times$ stock solution in DMSO) for 1 h at ambient temperature followed by copper-mediated azide-alkyne cycloaddition (CuAAC) conjugation to rhodamine-azide. CuAAC was performed with 20 μ M rhodamine-azide (50 $\times$ stock in DMSO), 1 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP; fresh 50 $\times$ stock in water, final concentration = 1 mM), ligand (17 $\times$ stock in DMSO:t-butanol 1:4, final concentration = 100 μ M) and 1 mM CuSO₄ (50 $\times$ stock in water, final concentration = 1 mM). Samples were allowed to react for 1 h at ambient temperature before quenching with 8 μ L 4 $\times$ SDS-PAGE loading buffer. Quenched reactions were analyzed by SDS-PAGE and visualized by in-gel fluorescence. For CASP8 and IMPDH2 25 μ L of soluble proteomes containing IMPDH2 or Pro-CASP8 (1 μ M each respectively) were treated with the indicated fragment for 1 h prior to incubation for 1 h with **18** (1 μ L, of 625 μ M, final concentration = 25 μ M) for IMPDH2 or **61** (1 μ L, of 625 μ M, final concentration = 25 μ M) for CASP8. For MLTK, HEK 293T cells stably overexpressing MTLK² were treated with the indicated fragment electrophiles for 1 h, followed by labeling with **59** (1 μ L, of 125 μ M, final concentration = 5 μ M) for 1 h. These were followed by CuAAC conjugation to rhodamine-azide and evaluation by SDS-PAGE as described above.

Determination of in vitro IC₅₀ values

[00319] 25 μ L of proteomes containing the indicated protein were treated with fragment electrophiles for 1 h at ambient temperature, labeled with the probes detailed above for 1 h, quenched, and analyzed by SDS-PAGE and in-gel fluorescence visualization ($n = 3$). IA-

rhodamine was used as the probe for C161S-TIGAR, C409S-CASP8 and PRMT1. **59** was used as a probe for MLTK. The soluble proteome containing IMPDH2 was treated with ATP for 15 min prior to incubation with **18** (1 μ l, of 625 μ M, final concentration = 25 μ M) for 1 h. MLTK and IMPDH2 were subjected to CuAAC conjugation to rhodamine-azide as detailed above. The percentage of labeling was determined by quantifying the integrated optical intensity of the bands, using ImageJ software. Nonlinear regression analysis was used to determine the IC50 values from a dose-response curve generated using GraphPad Prism 6.

isoTOP-ABPP sample preparation

[00320] For *in situ* labeling, MDA-MB-231 cells were grown to 95% confluence and Ramos cells were grown to 1 million cells/mL. The media in all samples was replaced with fresh media, containing 200 μ M of the indicated fragments and the cells were incubated at 37 °C for 2 h, washed with cold PBS, scraped into cold PBS and harvested by centrifugation (see prior section on “Preparation of human cancer cell line proteomes”).

[00321] Fragments **2, 3, 8, 9, 10, 12, 13, 14, 21, 27, 28, 29, 31, 33, 38, 45, 51** and **56** were screened at 200 μ M *in situ*. Fragments **4** and **11** were screened at 100 μ M *in situ*. Fragments **2, 3, 8,** and **20** were tested at 50 μ M *in situ*.

[00322] After *in vitro* or *in situ* fragment treatment, the samples were labeled for 1 h at ambient temperature with 100 μ M iodoacetamide alkyne (IA-alkyne, 5 μ L of 10 mM stock in DMSO). For direct labeling with **61, 61** (5 μ L of 1 or 10 mM stocks in DMSO, final concentration = 10 or 100 μ M) was substituted for IA-alkyne. Samples were conjugated by CuAAC to either the light (fragment treated) or heavy (DMSO treated) TEV tags (10 μ L of 5 mM stocks in DMSO, final concentration = 100 μ M), TCEP, TBTA ligand and CuSO₄ as detailed above. The samples were allowed to react for 1 h at which point the samples were centrifuged (16,000 g, 5 min, 4 °C). The resulting pellets were sonicated in ice-cold methanol (500 μ L) and the resuspended light- and heavy-labeled samples were then combined and centrifuged (16,000 g, 5 min, 4 °C). The pellets were solubilized in PBS containing 1.2% SDS (1 mL) with sonication and heating (5 min, 95 °C) and any insoluble material was removed by an additional centrifugation step at ambient temperature (14,000 g, 1 min).

[00323] For each sample, 100 μ L of streptavidin-agarose beads slurry (Pierce) was washed in 10 mL PBS and then resuspended in 5 mL PBS. The SDS-solubilized proteins were added to the suspension of streptavidin-agarose beads and the bead mixture was rotated for 3 h at ambient temperature. After incubation, the beads were pelleted by centrifugation (1,400 g, 3 min) and were washed (2 $\times$ 10 mL PBS and 2 $\times$ 10 mL water).

[00324] The beads were transferred to eppendorf tubes with 1 mL PBS, centrifuged (1,400 g, 3 min), and resuspended in PBS containing 6 M urea (500 μ L). To this was added 10 mM DTT

(25 μ L of a 200 mM stock in water) and the beads were incubated at 65 °C for 15 mins. 20 mM iodoacetamide (25 μ L of a 400 mM stock in water) was then added and allowed to react at 37 °C for 30 mins with shaking. The bead mixture was diluted with 900 μ L PBS, pelleted by centrifugation (1,400 g, 3 min), and resuspended in 200 μ L PBS. To this was added 1 mM CaCl_2 (2 μ L of a 200 mM stock in water) and trypsin (2 μ g, Promega, sequencing grade) and the digestion was allowed to proceed overnight at 37 °C with shaking. The beads were separated from the digest with Micro Bio-Spin columns (Bio-Rad) by centrifugation (1,000 g, 1 min), washed (2×1 mL PBS and 2×1 mL water) and then transferred to fresh eppendorfs with 1 mL water. The washed beads were washed once further in 140 μ L TEV buffer (50 mM *Tris*, pH 8, 0.5 mM EDTA, 1 mM DTT) and then resuspended in 140 μ L TEV buffer. 5 μ L TEV protease (80 μ M) was added and the reactions were rotated overnight at 29 °C. The TEV digest was separated from the beads with Micro Bio-Spin columns by centrifugation (1,400 g, 3 min) and the beads were washed once with water (100 μ L). The samples were then acidified to a final concentration of 5% (v/v) formic acid and stored at -80 °C prior to analysis.

Liquid-chromatography-mass-spectrometry (LC-MS) analysis of isoTOP-ABPP samples

[00325] TEV digests were pressure loaded onto a 250 μ m (inner diameter) fused silica capillary column packed with C18 resin (Aqua 5 μ m, Phenomenex). The samples were analyzed by multidimensional liquid chromatography tandem mass spectrometry (MudPIT), using an LTQ-Velos Orbitrap mass spectrometer (Thermo Scientific) coupled to an Agilent 1200- series quaternary pump. The peptides were eluted onto a biphasic column with a 5 μ m tip (100 μ m fused silica, packed with C18 (10 cm) and bulk strong cation exchange resin (3 cm , SCX, Phenomenex,)) in a 5-step MudPIT experiment, using 0%, 30%, 60%, 90%, and 100% salt bumps of 500 mM aqueous ammonium acetate and using a gradient of 5–100% buffer B in buffer A (buffer A: 95% water, 5% acetonitrile, 0.1% formic acid; buffer B: 5% water, 95% acetonitrile, 0.1% formic acid) as has been described in Weerapana *et al.* Nat Protoc 2:1414-1425 (2007). Data was collected in data-dependent acquisition mode with dynamic exclusion enabled (20 s, repeat of 2). One full MS (MS1) scan (400-1800 m/z) was followed by 30 MS2 scans (ITMS) of the nth most abundant ions.

Peptide and protein identification

[00326] The MS2 spectra data were extracted from the raw file using RAW Xtractor (version 1.9.9.2; available at <http://fields.scripps.edu/downloads.php>). MS2 spectra data were searched using the ProLuCID algorithm (publicly available at <http://fields.scripps.edu/downloads.php>) using a reverse concatenated, nonredundant variant of the Human UniProt database (release-2012_11). Cysteine residues were searched with a static modification for carboxyamidomethylation (+57.02146) and up to one differential modification for either the

light or heavy TEV tags (+464.28595 or +470.29976 respectively). Peptides were required to have at least one tryptic terminus and to contain the TEV modification. ProLuCID data was filtered through DTASelect (version 2.0) to achieve a peptide false-positive rate below 1%.

R value calculation and processing

[00327] The ratios of heavy/light for each unique peptide (DMSO/compound treated; isoTOP-ABPP ratios, *R* values) were quantified with in-house CIMAGE software, using default parameters (3 ms1s per peak and signal to noise threshold 2.5). Site-specific engagement of electrophilic fragments was assessed by blockade of IA-alkyne probe labeling. For peptides that showed a $\geq 95\%$ reduction in MS1 peak area from the fragment treated proteome (light TEV tag) when compared to the DMSO treated proteome (heavy TEV tag), a maximal ratio of 20 was assigned. Ratios for unique peptide entries were calculated for each experiment; overlapping peptides with the same modified cysteine (e.g. different charge states, MudPIT chromatographic steps or tryptic termini) were grouped together and the median ratio was reported as the final ratio (*R*). The peptide ratios reported by CIMAGE were further filtered to ensure the removal or correction of low quality ratios in each individual dataset. The quality filters applied were the following: removal of half tryptic peptides; for ratios with high standard deviations from the median (90% of the median or above) the lowest ratio was taken instead of the median; removal of peptides with *R*=20 and only a single ms2 event triggered during the elution of the parent ion; manual annotation of all the peptides with ratios of 20, removing any peptides with low-quality elution profiles that remained after the previous curation steps. Proteome reactivity values for individual fragments were computed as the percentage of the total quantified cysteine-containing peptides with *R* values ≥ 4 (defined as liganded cysteines) for each replicate experiment and the final proteome reactivity value was calculated as the mean for all replicate experiments for each fragment from both MDA-MB-231 and Ramos cellular proteomes.

Cross-data processing

[00328] Biological replicates of the same compound and cell-line were averaged if the standard deviation was below 60% of the mean; otherwise the lowest value of the ratio set was taken. For peptides with multiple modified cysteines, the cysteine with the highest number of quantification events was kept and the remaining, redundant peptides were discarded. Peptides included in the aggregate dataset (those used for further bioinformatics and statistical analyses) were required to have been quantified in 3 experiments. Cysteines were categorized as liganded if they had at least two ratios $R \geq 4$ (hit fragments) and one ratio between 0.5 and 2 (control fragments). Although the majority ($> 75\%$ of fragments) were profiled in at least two biological replicates, some data from single replicate MS experiments were included. Averaged filtered data for all fragments and representative individual filtered datasets are found in **Tables 1-3**.

In situ data processing

[00329] R values were calculated and individual datasets were filtered as described above (*R* value calculation and processing). Two categories of hits *in situ* were defined: 1) cysteines liganded *in situ* that were also observed as hits *in vitro* and 2) cysteines that detected *in vitro*, but were only liganded *in situ*. For the first category, *R* values for the same cysteine containing peptide from *in vitro* and *in situ* experiments were compared and if both had ratios $R \geq 4$, the cysteine was considered ligandable *in situ*. To qualify for the second category, two ratios $R \geq 4$ for replicates of two different fragments were required to be detected *in situ* and at least one of these fragments must be quantified as a non-hit with $R \leq 2$ *in vitro*. Additionally, another cysteine from the same protein was required to be unliganded *in situ* ($R \leq 2$) by the same fragment to control for the possibility that changes in *R* values from changes in protein expression upon fragment treatment rather than from fragment competition.

Functional annotation of liganded cysteines

[00330] Custom python scripts were used to compile functional annotations available in the UniProtKB/Swiss-Prot Protein Knowledge database (release-2012_11). Relevant Uniprot entries were mined for available functional annotations at the residue level, specifically for annotations regarding enzyme catalytic residues (active sites), disulfides (redox active and structural) and metal binding sites. Liganded proteins were queried against the Drugbank database (Version 4.2) and fractionated into DrugBank and non-Drugbank proteins. Functional keywords assigned at the protein level were collected from the Uniprot database and the Drugbank and non-drugbank categories were further classified into protein functional classes. Cysteine reactivity data was re-processed using ProLuCID as detailed above (Peptide and protein identification). Cysteines found in both the reactivity and ligandability datasets were sorted based on their reactivity values (lower ratio indicates higher reactivity). The moving average of the percentage of total liganded cysteines within each reactivity bin (step-size 50) was taken. Custom python scripts were developed to collect relevant NMR and X-ray structures from the RCSB Protein Data Bank (PDB). For proteins without available PDB structures, sequence alignments, performed with BLAST to proteins deposited in the PDB, were used to identify structural homologues. For annotation of active-site and non-active cysteines, enzymes with structures in the PDB were manually inspected to evaluate the location of the cysteine. Cysteines were considered to reside in enzyme active sites if they were within 10 Å of active-site ligand or residue(s). Cysteines outside of the 10 Å range were deemed non-active-site residues. Histograms of fragment hit-rates across high-coverage, ligandable cysteines, active-site and non-active site cysteines were calculated from the subset of ligandable cysteines quantified in 10 or more separate experiments. The fragment hit rate is reported as the percentage of the total

quantification events with $R \geq 4$. For analyses of trends within the whole data, including histograms and heatmaps, a cell-line merged dataset was used where data from the MDA-MB-231 experiments was taken first and the Ramos data was used if there was no data from MDA-MB-231 experiments for a particular fragment and cysteine. Heatmaps were generated in R (version 3.1.3) using the heatmap.2 algorithm. Protein structures were rendered using Pymol.

GSH reactivity

[00331] Glutathione (GSH) was diluted to a final concentration of 125 μ M in assay buffer (100 mM *Tris*, pH 8.8, 10% ethanol as co solvent). In triplicate, to 100 μ L of the GSH mixture in a clear 96 well plate (Costar® Corning®), the indicated electrophile (2 μ L of a 50 mM stock solution in DMSO, final concentration = 500 μ M) was added and the reaction mixture was incubated at room temperature for 1 h. 5 μ L of Ellman's reagent (100 mM stock in 1M NaOH, final concentration = 5 mM) was added and the absorbance was measure at 440 nm on a plate reader (Tecan Infinite F500). The concentration of GSH remaining was calculated from a standard curve.

Reactive cysteine docking

[00332] *In silico* fragment library containing all chloroacetamide and acrylamide fragments from **Fig. 3** was prepared using Open Babel library with custom Python scripts. Fragments were modeled in their reactive form (i.e., with explicit chloroacetamide and acrylamide warheads). 3D coordinates were generated from SMILES strings, calculating their protonation state at pH 7.4, and then minimizing them using MMF94s forcefield (50K iterations steepest descent; 90K conjugate gradient); for chiral molecules with undefined configuration, all enantiomers were generated, resulting in 53 total fragments

[00333] For each protein, the UniProtKB ID was used to filter the PDB. Structures determined by X-ray crystallography were selected, privileging higher sequence coverage and structure resolution (See Table 5 for selected PDB IDs). When no human structures were available, the closest homologous organism available was selected (e.g. PRMT1: *R. norvegicus*). Protein structures were prepared following the standard AutoDock protocol. Waters, salts, and crystallographic additives were removed; AutoDockTools was used to add hydrogens, calculate Gasteiger-Marsili charges and generate PDBQT files.

[00334] MSMS reduced surface method was used to identify accessible cysteines. The protein volume was scanned using a probe radius of 1.5 Å; residues were considered accessible if they had at least one atom in contact with either external surfaces or internal cavities.

[00335] The fragment library was docked independently on each accessible cysteine using AutoDock 4.2. A grid box of 24.4×24.4×24.4 Å was centered on the geometric center of the residue; thiol hydrogen was removed from the side-chain, which was modeled as flexible during

the docking; the rest of the structure was kept rigid. A custom 13-7 interaction potential was defined between the nucleophile sulfur and the reactive carbon in the ligands. The equilibrium distance (r_{eq}) was set to the length of the C-S covalent bond (1.8 Å); the potential well depth (ϵ_{eq}) varied between 1.0 and 0.175 to model to the reactivity of the different ligands. For each fragment, potential well depth was determined by dividing its proteomic reactivity percentage by 20, and the value for iodoacetamide was approximated as the maximum (2.5) for reference. The potential was implemented by modifying the force field table of AutoDock. Fragments were docked with no constraints, generating 100 poses using the default GA settings. For each fragment, the best docking score pose was analyzed: if the distance between the nucleophilic sulfur and the reactive carbon was ≤ 2.0 Å, the cysteine was considered covalently modified. If a residue was alkylated by at least one ligand, it was considered labeled. The docking score (i.e., negative binding energy) was calculated based on the estimated interaction energy of each fragment in its docked pose. The docking score of the best alkylating fragment defined the labeling score. The residue with the best labeling score was considered the most probable to be labeled.

Structural modeling

[00336] IMPDH2 structure, including the Bateman domain, was modeled using I-TASSER.

Subcloning and mutagenesis

[00337] Full length cDNAs encoding for IDH1 (Open Biosystems, Clone ID: 3880331) and IMPDH2 (Open Biosystems, Clone ID: 3447994) were subcloned into pET22b (+) (Novagen) with C-terminal His₆-affinity tag (SEQ ID NO: 861). Full length cDNA encoding for TIGAR (Origene, Sc320794) was subcloned into pET28a (+) (Novagen) with N-terminal His₆-affinity tag (SEQ ID NO: 861). Full length PRMT1 subcloned into pET45b (+) (Novagen) was previously generated by the Cravatt lab. Full-length human CASP3 (residues 1-277) and a truncated CASP8 (residues 217-479) without the CARD domain was subcloned into pET23b (Novagen) with C-terminal His₆-affinity tags (SEQ ID NO: 861). Cysteine mutants were generated using QuikChange site-directed mutagenesis, using primers containing the desired mutations and their respective compliments.

Recombinant overexpression of TIGAR, IDH1, PRMT1 and IMPDH2

[00338] TIGAR, IDH1, PRMT1 and IMPDH2 were expressed in BL21(DE3) Chemically Competent Cells (NEB), grown on Terrific Broth supplemented with the desired antibiotic (50 µg/mL Kanamycin or 50 µg/mL Carbenicillin) to OD₆₀₀ of 0.8 and induced with 0.5 mM IPTG for 16 h at 18 °C. Cells were immediately harvested and resuspended in 30 mL cold buffer A (25 mM Tris, pH 7.4, 200 mM NaCl, 10% glycerol, 1 mM BME), supplemented with lysozyme (Sigma), DNAase (NEB) and cOmplete protease inhibitor tablets (Roche), sonicated, and

centrifuged (45,000 g, 30 min, 4 °C). The soluble fractions were collected and rotated for 1 h with 1 mL Ni-NTA slurry (Qiagen) at 4 °C. The slurry was then transferred to a 50 mL volume, fritted column and collected by gravity flow. The resin was then washed with 100 mL buffer A containing 20 mM imidazole and eluted with 10 mL buffer A containing 200 mM imidazole. The eluant was concentrated to 2.5 mL (Amicon-Ultra-15, 10 kDa MW cutoff), buffer exchanged using PD10 columns (GE Amersham) into the storage buffer (50 mM HEPES, pH 7.4, 150 mM NaCl, 10% glycerol, 1 mM BME) and further concentrated (Amicon-Ultra-4, 10 kDa MW cutoff) to a final concentration of approximately 100 µM protein. Protein concentration was determined using the Bio-Rad DC™ protein assay kit. Protein purity was assayed by SDS-PAGE under reducing conditions and were >95% pure.

Recombinant CASP3, CASP8 and TEV protease expression

[00339] CASP3, CASP8, pro-CASP8 (D374A, D384A) and an N-terminal MBP fusion-His₆-TEV-Arg₆ protease construct pRK793 ("His₆" disclosed as SEQ ID NO: 861 and "Arg₆" disclosed as SEQ ID NO: 862) were expressed in *E. coli* BL21(DE3)pLysS cells (Stratagene). Cells were grown in 2xYT medium supplemented with 200 µg/ml ampicillin and 50 µg/ml chloramphenicol at 37 °C to an OD₆₀₀ of 0.8-1.0. Overexpression of caspase was induced with 0.2 mM IPTG at 30 °C for 4 h (CASP3) or at 12 °C overnight (CASP8) or with 0.5 mM IPTG at 30 °C for 4 h (TEV protease). Cells were immediately harvested and resuspended in ice cold buffer A (caspases: 100 mM Tris, pH 8.0, 100 mM NaCl; TEV protease: PBS) and subjected to 3 cycles of lysis by microfluidization (Microfluidics). The cell lysate was clarified by centrifugation (45,000 g, 30 min, 4 °C) and soluble fractions were loaded onto a 1 mL HisTrap HP Ni-NTA affinity column (GE Amersham) pre-equilibrated with buffer A and eluted with buffer A containing 200 mM imidazole. The eluted protein was immediately diluted two-fold with buffer B (20 mM Tris, pH 8.0) and purified by anion-exchange chromatography (HiTrap Q HP, GE Amersham) with a 30-column volume gradient to 50% of buffer B containing 1 M NaCl. The caspases were injected over a Superdex 200 16/60 gel filtration column (GE Amersham) and TEV protease over a Superdex 75 gel filtration column (GE Amersham) in buffer C (caspases: 20 mM Tris, pH 8.0, 50 mM NaCl; TEV protease: PBS, 10 mM DTT) to buffer exchange and to remove any remaining contaminants. Fractions containing the desired protein were pooled and concentrated to approximately 1 mg/mL (Millipore Ultrafree-15, 10 kDa MW cutoff). The purified proteins were immediately frozen and stored at -80 °C. Protein concentrations were measured using both Bio-Rad colorimetric assay and A₂₈₀ absorbance in denaturing conditions. Protein purity was assayed by SDS-PAGE under reducing conditions and were >98% pure.

Retroviral overexpression of flag-tagged IDH1 proteins

[00340] R132H-IDH1, including an additional K345K silent mutation to remove an unwanted restriction site and GFP were subcloned into a modified pCLNCX retroviral vector. Retrovirus was prepared by taking 1.5 µg of each pCLNCX vector and 1.5 µg pCMV-VSV-G and 20 µL of Roche X-tremeGeneHP DNA transfection reagent to transfect HEK-293RTV cells. The medium was replaced after 1 day of transfection and the following day the culture supernatant was collected and filtered through 0.5 µm filter. 10 mL of the filtrate, containing the desired virus, was used to infect MUM2C cells in the presence of polybrene (8 µg/mL) for 48 h, at which point the infected cells were selected for in medium containing 100 µg/mL hygromycin for 7-10 days. Surviving cells were expanded and cultured in complete RPMI-1640 medium containing hygromycin.

IDH1 NADP assay

[00341] Recombinant IDH1 and C269S -IDH1 (100 µM in storage buffer) were diluted 1:200 in MDA-MB-231 cellular proteome (1 mg/mL). To 25 µL of this mixture was added 1 µL of the indicated compound (25 X stock solution in DMSO) and the lysates were incubated for 1 h at room temperature in clear 96 well plates (Corning® Costar®). 75 µL per well of a stock solution of NADP (13.3 mM) and isocitrate (13.3 mM) in IDH1 buffer (40 mM Tris, pH 7.4, 2 mM MgCl₂, 0.01% pluronic) was added immediately before measuring UV absorbance at 340 nm on a 96 well UV absorbance plate reader (TECAN). Absorbance was measured for 45 minutes and the relative activities were calculated from the change in absorbance for the linear portion of the curve.

IDH1 2-hydroxyglutarate (2-HG) formation assay

[00342] MUM2C cells stably overexpressing IDH1 R132H were seeded 1.5×10^6 cells/150 mm dish. The following day the indicated compounds (50 mM stock solutions in DMSO) or DMSO were added to the cells to the final concentrations indicated and were allowed to incubate for 2 h. Control cells overexpressing GFP were treated in parallel. The cells were washed in ice-cold PBS and collected by scraping in ice-cold PBS and centrifugation (1,400 g, 3 min, 4 °C). The cell pellets were then resuspended in 100 µL ice-cold PBS followed by sonication and centrifugation at 16,000 g for 10 min. Lysates were then buffer exchanged into IDH1 buffer (40 mM Tris, pH 7.4, 2 mM MgCl₂) with 0.5 mL ZEBRA spin desalting columns (Thermo Fisher, 89882). The protein concentrations were adjusted to 3.5 mg/mL and 25 µL of the lysate was mixed with 25 µL of the reaction mixture (2.5 mM NADPH and 2.5 mM α-ketoglutarate in IDH1 buffer) and the reaction was allowed to proceed for 4 h at which point the reaction mixtures were quenched with 50 µL cold methanol, followed by a centrifugation (16,000 g, 10 min, 4 °C). Formation of 2-HG was followed by targeted LC/MS analysis. The

reaction mixture was separated with a Luna-NH₂ column (5 μ m, 100 Å, 50 $\times$ 4.6 mm, Phenomenex) with a precolumn (NH₂, 4 $\times$ 3.0 mm) using a gradient of mobile phases A and B (mobile phase A: 100% CH₃CN, 0.1% formic acid; mobile phase B: 95:5 (v/v) H₂O:CH₃CN, 50 mM NH₄OAc, 0.2% NH₄OH). The flow rate started at 0.1 mL/min, and the gradient consisted of 5 min 0% B, a linear increase to 100% B over 20 min at a flow rate of 0.4 mL/min, followed by an isocratic gradient of 100% B for 2 min at 0.5 mL/min before equilibrating for 3 min at 0% B at 0.4 mL/min (30 min total). For each run, the injection volume was 25 μ L. MS analysis was performed on an Agilent G6410B tandem mass spectrometer with ESI source. The dwell time for 2-HG was set to 100 ms. and collision energy for 2-HG was set to 5. The capillary was set to 4 kV, and the fragmentor was set to 100 V. The drying gas temperature was 350 °C, the drying gas flow rate was 11 L/min and the nebulizer pressure was 35 psi. The mass spectrometer was run in MRM mode, monitoring the transition of *m/z* from 146.7 to 129 for 2-HG (negative ionization mode). Treatments were conducted in triplicate. Background 2-HG production, calculated from the 'mock' GFP over expressing cells, was subtracted from the total 2-HG production. *TIGAR activity assay*

[00343] TIGAR activity assay was conducted as described in Gerin *et al. The Biochemical Journal* **458**:439-448 (2014). Formation of 3PG (3-phosphoglycerate) production from 23BPG (2,3-bisphosphoglycerate) was measured spectrophotometrically on a TECAN plate reader, measuring decrease in absorbance at 340 nm in clear, flat-bottom 96 well microplate (Corning® Costar®). 2 μ L of recombinant TIGAR (10 mg/mL) was diluted into 1 mL dilution buffer (25 mM HEPES, pH 7.1, 25 mM KCl, 1 mM MgCl₂). 25 μ L of diluted protein was incubated for 1 h with the indicated concentration of compound (1 μ L, 25 $\times$ stock in DMSO). Then 75 μ L of assay mixture comprised of 25 mM HEPES (pH 7.1), 25 mM KCl, 1 mM MgCl₂, 0.5 mM NADH, 1 mM DTT, 1 mM 23BPG, 1 mM ATP-Mg, the equivalent of 1 μ L each of rabbit muscle GAPDH (4000 units/mL, Sigma, G5537) and yeast PG kinase (6300 units/mL, Sigma, P7634) was added to the protein and decrease in absorbance was monitored at 340 nm. The background, calculated from samples lacking TIGAR, was subtracted from samples containing TIGAR. Experiments were performed in quadruplicate.

PRMT1 in vitro methylation assays

[00344] PRMT1 assays were conducted as described in Weerapana *et al. Nature* **468**:790-795 (2010). Recombinant human PRMT1 (0.85 μ M, wild type or C101S mutant) in 25 μ L methylation buffer (20 mM Tris, pH 8.0, 200 mM NaCl, 0.4 mM EDTA) was pre-incubated with indicated fragments for 1 h and methylation activity was monitored after addition of 1 mg of recombinant histone 4 (NEB, M2504S) and ³H-SAM (2 μ Ci). Reactions were further incubated for 60 min at ambient temperature and stopped with 4 $\times$ SDS sample buffer. SDS-PAGE gels

were fixed with 10% acetic acid/10% methanol (v/v), washed, and incubated with Amplify reagent (Amersham) before exposing to film at -80°C for 3 days.

MLTK in vitro kinase activity assay

[00345] The kinase activity assay protocol was conducted as described in Wang *et al.* ACS Chemical biology 9:2194-2198 (2014). Kinase assay buffers, myelin basic protein (MBP) substrate and ATP stock solution were purchased from SignalChem. Radio-labeled $[\gamma\text{-}^{33}\text{P}]$ ATP was purchased from PerkinElmer. 250 μL of HEK-293T soluble lysates (8 mg/mL), stably overexpressing WT, C22A or K45M MLTK were labeled for 1 h with 100 μM fragment or DMSO. The samples were then individually immunoprecipitated with 20 μL flag resin slurry per sample and then eluted with 15 μL 3 $\times$ Flag-peptide. To each sample was added 5 μL of MBP and then 5 μL of $[\gamma\text{-}^{33}\text{P}]$ ATP assay cocktail (250 μM , 167 $\mu\text{Ci/mL}$) was added to initiate the kinase reaction. Each reaction mixture was incubated at ambient temperature for 30 min, and the reactions were terminated by spotting 25 μL of the reaction mixture onto individual precut phosphocellulose P81 paper. The spotted P81 strips were washed with 10 mL of 1% phosphoric acid (3×10 min). MLTK activity was measured by counting the radioactivity on the P81 paper in the presence of scintillation fluid in a scintillation counter. The background was determined from the K45M- inactive mutant MLTK activity level, which was subtracted from the WT and C22A samples. Relative activities for WT and C22A were normalized to their respective DMSO treated samples. Experiments were performed in triplicate.

CASP3 and CASP8 in vitro activity assays

[00346] Caspase 3 and 8 assays were conducted with CASP8 activity assay kit (BioVision, K112-100) and Caspase 3 activity assay kit (Invitrogen, EnzChek® Caspase-3 Assay Kit), following the manufacturer's instructions. Briefly, recombinant Caspase 3 (10 μM) was added to soluble Ramos lysates (1 mg/mL) to a 100 nM final concentration of protease. Caspase 8 (30 μM) was added to soluble Ramos lysates to a 1 μM final concentration of protease. In triplicate, 50 μL lysate was treated with either DMSO, DEVD-CHO ("DEVD" disclosed as SEQ ID NO: 857) (20 μM) or the indicated compounds (100 μM) for 1 h, following which 50 μL of 2 $\times$ reaction buffer containing 10mM DTT and 5 μL substrate (4 mM stock in DMSO of IETD-AFC ("IETD" disclosed as SEQ ID NO: 858) for CASP8; 10 mM stock in DMSO of DEVD-AMC ("DEVD" disclosed as SEQ ID NO: 857) for CASP3) was added to each well and the samples were incubated at ambient temperature for 2 h. Caspase activity was measured from the increase in fluorescence (excitation 380 nm emission 460 nm). Experiments were performed in triplicate. Background was calculated from samples lacking the recombinant caspase.

Apoptosis assays with Caspase 8 inhibitors

[00347] 4 mL of Jurkat cells in RPMI (1.5 million cells/mL) were treated with the indicated compound at 30 μ M for 30 min (50 mM stock solution in DMSO). Z-VAD-FMK (EMD Millipore Biosciences, 627610) and was used at a final concentration of 100 μ M. After pre-incubation, FASL (4 μ L of 100 μ g/ μ L stock solution of *SuperFasLigand*[™] in water, final concentration =100 ng/mL, Enzo life Sciences) or staurosporine (8 μ L of 1 mM stock solution in DMSO, final concentration =2 μ M, Fisher Scientific, 50664333). After 6 hours, cells were harvested by centrifugation, washed and lysed in cell lysis buffer (BioVision, 1067-100) and 40 μ g of each sample were separated by SDS-Page on 14% polyacrylamide gels. The gels were transferred to nitrocellulose membranes and were immunoblotted overnight with the indicated antibodies. For measurements of cell viability, in quadruplicate for each condition, 150,000 cells (100 μ L of 1.5 million cells/mL) were plated in Nunc[™] MicroWell[™] 96-Well Optical-Bottom Plates with Polymer Base (Fisher Scientific). Compounds, FASL and STS were used at the same concentrations indicated above with a 30 minute pre-incubation with compound, followed by 6 hours with either STS or FASL or DMSO. Cell viability was measured with CellTiter-Glo[®] Luminescent Cell Viability Assay (Promega) and was read on a Biotech Synergy 4 plate reader.

Western blotting

[00348] For CASP8, CASP3 and PARP, cell pellets were resuspended in cell lysis buffer from (BioVision, 1067-100) with 1 $\times$ cOmplete protease inhibitor (Roche) and allowed to incubate on ice for 30 min prior to centrifugation (10 min, 16,000 g). For all other proteins, cell pellets were resuspended in PBS and lysed with sonication prior to centrifugation (10 min, 16,000 g). The proteins were then resolved by SDS-PAGE and transferred to nitrocellulose membranes, blocked with 5% BSA in TBST and probed with the indicated antibodies. The primary antibodies and the dilutions used are as follows: anti-parp (Cell Signaling, 9532, 1:1000), anti-casp3 (Cell Signaling, 9662, 1:500), anti-casp8 (Cell Signaling, 9746, 1:500), anti-IDH1 (Cell Signaling, 1:500, 3997s), anti-actin (Cell Signaling, 3700, 1:3000), anti-gapdh (Santa Cruz, sc-32233, 1:2000) anti-flag (Sigma Aldrich, F1804, 1:3000) . Blots were incubated with primary antibodies overnight at 4 °C with rocking and were then washed (3 $\times$ 5 min, TBST) and incubated with secondary antibodies (LICOR, IRDye[®] 800CW or IRDye[®] 800LT, 1:10,000) for 1 h at ambient temperature. Blots were further washed (3 $\times$ 5 min, TBST) and visualized on a LICOR Odyssey Scanner.

Statistical analysis

[00349] Data are shown as mean $\pm$ SEM. *P* values were calculated using unpaired, two-tailed Student's *t*-test. *P* values of <0.05 were considered significant.

Prediction Failures in Reactive Docking

[00350] Prediction failures were due to the approximations of the rigid model used with highly flexible/solvent exposed loop regions (STAT1:C255, PDB ID:1YVL; HAT1:C101, PDB ID:2P0W; ZAP70:C117, PDB ID:4K2R), or with partially buried residues (SARS:C438, PDB ID:4I87; PAICS:C374, PDB ID:2H31). In some embodiments, the simulation of some degree of flexibility (such as flexible side chains) improves the success rate. In some embodiments, the method was limited by availability and quality of crystallographic structures, when sequences were not fully resolved in available models (XPO1:C34, C1070, PDB ID:3GB8, FNBP1:C511,C555,C609, PDB ID:2EFL; IMPDH2:C140, PDB ID:1NF7), or when only orthologue sequences were available (PRMT1: R. Norvegicus, PDB ID:1ORI).

General Synthetic Methods

[00351] Chemicals and reagents were purchased from a variety of vendors, including Sigma Aldrich, Acros, Fisher, Fluka, Santa Cruz, CombiBlocks, BioBlocks, and Matrix Scientific, and were used without further purification, unless noted otherwise. Anhydrous solvents were obtained as commercially available pre-dried, oxygen-free formulations. Flash chromatography was carried out using 230–400 mesh silica gel. Preparative thin layer chromatography (PTLC) was carried out using glass backed PTLC plates 500-2000 μm thickness (Analtech). All reactions were monitored by thin layer chromatography carried out on 0.25 mm E. Merck silica gel plates (60F-254) and visualized with UV light, or by ninhydrin, ethanolic phosphomolybdic acid, iodine, *p*-anisaldehyde or potassium permanganate stain. NMR spectra were recorded on Varian INOVA-400, Bruker DRX-600 or Bruker DRX-500 spectrometers in the indicated solvent. Multiplicities are reported with the following abbreviations: s singlet; d doublet; t triplet; q quartet; p pentet; m multiplet; br broad. Chemical shifts were reported in ppm relative to TMS and *J* values were reported in Hz. Mass spectrometry data were collected on a HP1100 single-quadrupole instrument (ESI; low resolution) or an Agilent ESI-TOF instrument (HRMS).

[00352] In some embodiments, General Procedure A was used for the synthesis of one or more of the small molecule fragments and/or cysteine-reactive probes described herein. The amine was dissolved in anhydrous CH_2Cl_2 (0.2 M) and cooled to 0 °C. To this, anhydrous pyridine (1.5 equiv.) was added in one portion, then chloroacetyl chloride (1.5 equiv.) dropwise and the reaction was monitored by TLC until complete disappearance of starting material and conversion to product was detected (typically 1 h). If the reaction did not proceed to completion, additional aliquots of pyridine (0.5 equiv.) and chloroacetyl chloride (0.5 equiv.) were added. The reaction was quenched with H_2O (1 mL), diluted with CH_2Cl_2 (20 mL), and washed twice with saturated NaHCO_3 (100 mL). The organic layer was concentrated *in vacuo* and purified by preparatory thin layer or flash column chromatography to afford the desired product. In some

embodiments, General Procedure A1 is similar to General Procedure A except triethylamine (3 equiv.) was used instead of pyridine. In some embodiments, General Procedure A2 is similar to General Procedure A except *N*-methylmorpholine (3 equiv.) was used instead of pyridine.

[00353] In some embodiments, General Procedure B was used for the synthesis of one or more of the small molecule fragments and/or cysteine-reactive probes described herein. The amine was dissolved in anhydrous CH₂Cl₂ (0.2 M) and cooled to 0 °C. To this, triethylamine (TEA, 1.5 equiv.), was added in one portion, then acryloyl chloride (1.5 equiv.) dropwise, and the reaction was monitored by TLC until complete disappearance of starting material and conversion to product was detected (typically 1 h). If the reaction did not proceed to completion, additional aliquots of TEA (0.5 equiv.) and acryloyl chloride (0.5 equiv.) were added. The reaction was quenched with H₂O (1 mL), diluted with CH₂Cl₂ (20 mL), and washed twice with saturated NaHCO₃ (100 mL). The organic layer was passed through a plug of silica, after which, the eluant was concentrated *in vacuo* and purified by preparatory thin layer or flash column chromatography to afford the desired product.

[00354] In some embodiments, General Procedure C was used for the synthesis of one or more of the small molecule fragments and/or cysteine-reactive probes described herein. Acryloyl chloride (80.4 µL, 1.0 mmol, 2 equiv.) was dissolved in anhydrous CH₂Cl₂ (4 mL) and cooled to 0 °C. A solution of the amine (0.5 mmol, 1 equiv.) and *N*-methylmorpholine (0.16 mL, 1.5 mmol, 3 equiv.) in CH₂Cl₂ (2 mL) was then added dropwise. The reaction was stirred for 1 hr at 0 °C then allowed to warm up to room temperature slowly. After TLC analysis showed disappearance of starting material, or 6 h, whichever was sooner, the reaction was quenched with saturated aqueous NaHCO₃ (5 mL) and extracted with CH₂Cl₂ (3 × 10 mL). The combined organic layers were dried over anhydrous Na₂SO₄, concentrated *in vacuo*, and the residue obtained was purified by preparatory thin layer chromatography to afford the desired product.

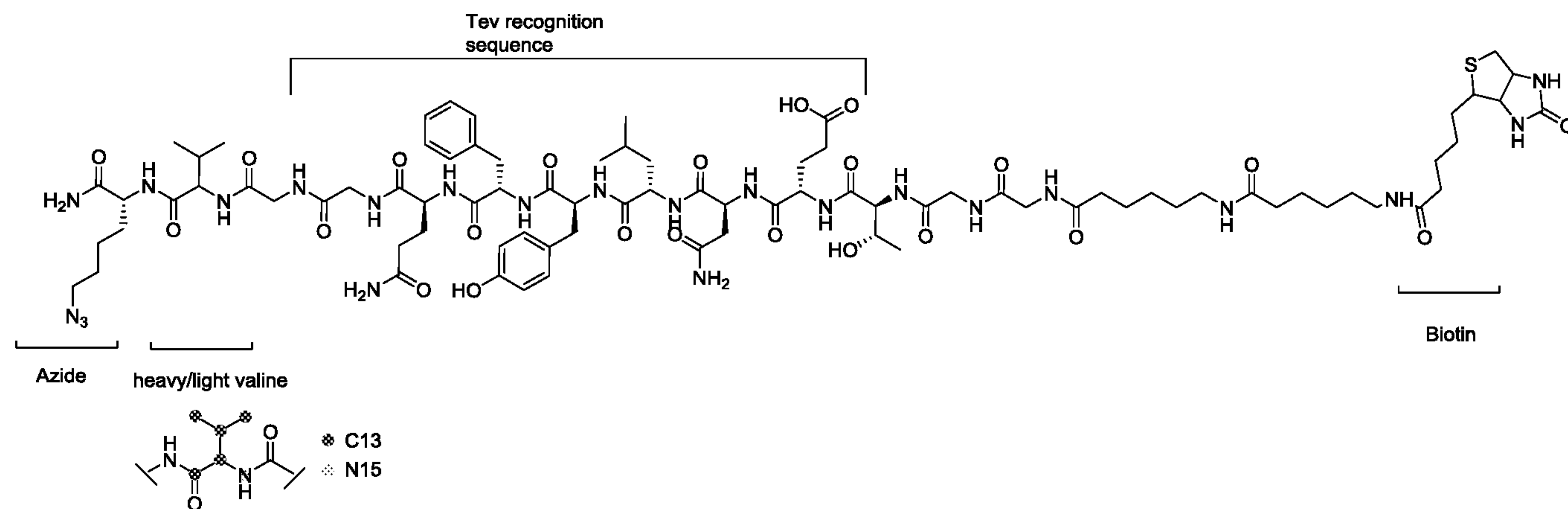
Synthesis of probes and fragments

Purchased fragments

[00355] The following electrophilic fragments were purchased from the indicated vendors. **2** (Santa Cruz Biotechnology sc-345083), **3** (Key Organics JS-092C), **4** (Sigma Aldrich T142433-10mg), **6** (Toronto Research Chemicals M320600), **8** (Alfa Aesar H33763), **10** (Santa Cruz Biotechnology sc-345060), **11** (Santa Cruz Biotechnology sc-354895), **12** (Santa Cruz Biotechnology sc-354966), **21** (Santa Cruz Biotechnology, sc-279681), **22** (Sigma Aldrich 699357-5G), **26** (Sigma Aldrich T109959), **27** (Santa Cruz Biotechnology sc-342184), **28** (Santa Cruz Biotechnology sc-335173), **29** (Santa Cruz Biotechnology sc-348978), **30** (Santa Cruz Biotechnology sc-355362), **32** (Santa Cruz Biotechnology sc-354613), **33** (Sigma Aldrich R996505), **34** (Santa Cruz Biotechnology sc-355477), **35** (Santa Cruz Biotechnology sc-

328985), **41** (Sigma Aldrich L469769), **42** (Sigma Aldrich R901946), **43** (Santa Cruz Biotechnology sc-307626), **52** (Enamine, EN300-08075), **55** (Santa Cruz Biotechnology sc-354880), **57** (VWR 100268-442), **58** (Enzo Life Sciences ALX-430-142-M005), **62** (WuXi Apptec).

Synthesis of isotopically-labeled TEV-tags:

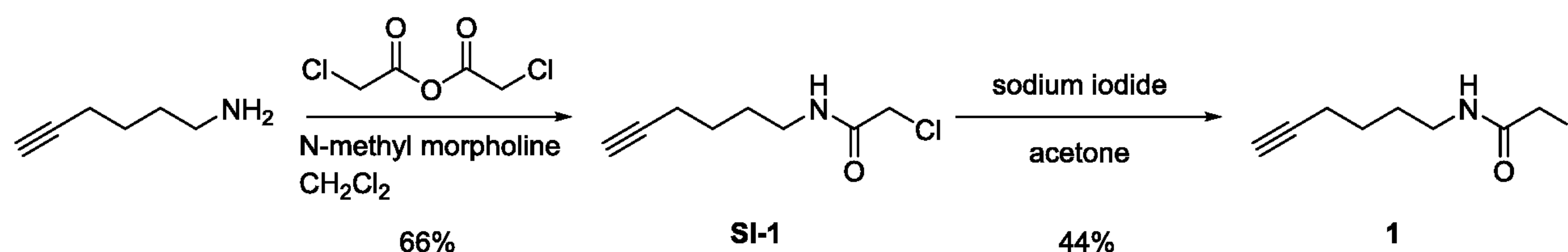


[00356] Isotopically-labeled heavy and light tags were synthesized with minor modifications to the procedure reported in Weerapana *et al. Nat Protoc* **2**:1414-1425 (2007) and Weerapana *et al. Nature* **468**:790-795 (2010). Fmoc-Rink-Amide-MBHA resin (EMD Biosciences; 0.5 M, 830 mg, 0.6 mmol / g loading) was deprotected with 4-methylpiperidine in DMF (50% v/v, 2 × 5 mL, 1 min). Fmoc-Lys(N₃)-OH (Anaspec) (500 mg, 1.26 mmol, 1.26 equiv.) was coupled to the resin overnight at room temperature with DIEA (113 μl) and 2-(6-chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HCTU; 1.3 mL of 0.5 M stock in DMF) followed by a second overnight coupling with Fmoc-Lys(N₃)-OH (500 mg, 1.26 mmol, 1.26 equiv.), DIEA (113 μl), O-(7-azabenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate (HATU; 1.3 mL of 0.5 M stock in DMF). Unmodified resin was then capped (2 × 30 min) with Ac₂O (400 μL) and DIEA (700 μL) in DMF after which the resin was washed with DMF (2 × 1 min). Deprotection with 4-methylpiperidine in DMF (50% v/v, 2 × 5 mL, 1 min) and coupling cycles (4 equiv. Fmoc-protected amino acid (EMD biosciences) in DMF) with HCTU (2 mL, 0.5 M in DMF) and DIEA (347.7 μL) were then repeated for the remaining amino acids. For the heavy TEV-tag, Fmoc-Valine-OH (¹³C₅C₁₅H₂₁¹⁵NO₄, ¹³C₅, 97-99%, ¹⁵N, 97-99%, Cambridge Isotope Laboratories, Inc.) was used. Reactions were monitored by ninhydrin stain and dual couplings were used for all steps that did not go to completion. Biotin (0.24 g, 2 equiv.) was coupled for two days at room temperature with NHS (0.1 g, 2 equiv.), DIC (0.16 g, 2 equiv.) and DIEA (0.175 g, 2 equiv.). The resin was then washed with DMF (5 mL, 2 × 1 min) followed by 1:1 CH₂Cl₂:MeOH (5 mL, 2 × 1 min), dried under a stream of nitrogen and transferred to a round-bottom flask. The peptides were cleaved for 90 minutes

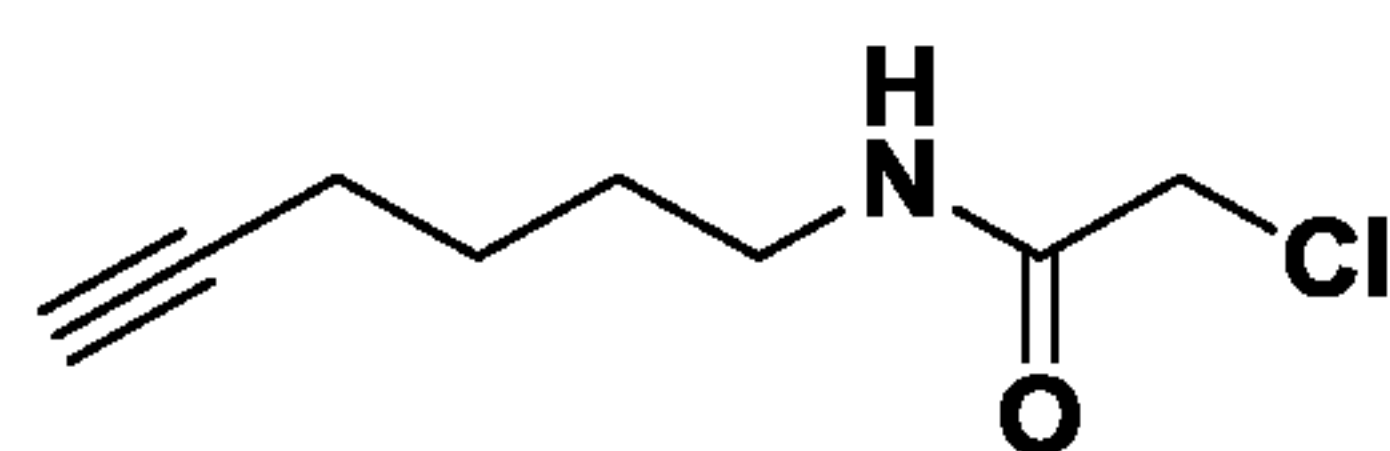
from the resin by treatment with 95:2.5:2.5 trifluoroacetic acid: water:triisopropylsilane. The resin was removed by filtration and the remaining solution was triturated with cold ether to provide either the light or heavy TEV-tag as a white solid. HPLC-MS revealed only minor impurities and the compounds were used without further purification. HRMS-ESI (m/z): calculated for $C_{83}H_{128}N_{23}O_{23}S$ [$M+H$]: (Light-TEV-Tag) 1846.9268; found: 1846.9187; calculated for $C_{78}^{13}C_5H_{128}N_{22}^{15}NO_{23}S$ [$M+H$]: (Heavy-TEV-Tag): 1852.9237; found: 1852.9309.

Synthesis of probes and fragments

Synthesis of 1

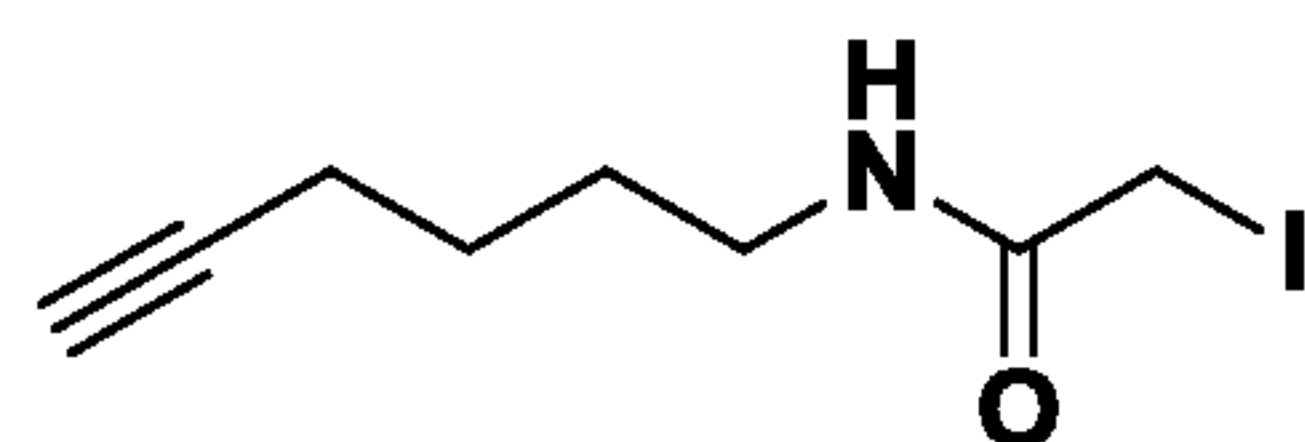


N-(hex-5-yn-1-yl)-2-chloroacetamide (SI-1)



[00357] To a solution of 5-hexynylamine (63 mg, 0.65 mmol, 1.0 equiv.) in CH_2Cl_2 (3.2 mL, 0.2 M) at 0 °C was added *N*-methylmorpholine (215 μL , 3 equiv.) followed by chloroacetic anhydride portionwise (222 mg, 2 equiv.). The reaction was allowed to come to room temperature and then stirred overnight. The reaction was then diluted with ether (50 mL), washed with 1 M HCl, 1 M NaOH, then brine (20 mL each). The combined organic layers were dried over magnesium sulfate and concentrated to yield chloroacetamide **SI-1** (74 mg, 66%). ^1H NMR (400 MHz, Chloroform- d) δ 6.79 (s, 1H), 4.09 (d, $J = 1.1$ Hz, 2H), 3.34 (q, $J = 6.8$ Hz, 2H), 2.23 (td, $J = 6.9, 2.7$ Hz, 2H), 1.98 (t, $J = 2.7$ Hz, 1H), 1.75 – 1.62 (m, 4H), 1.62 – 1.51 (m, 2H).

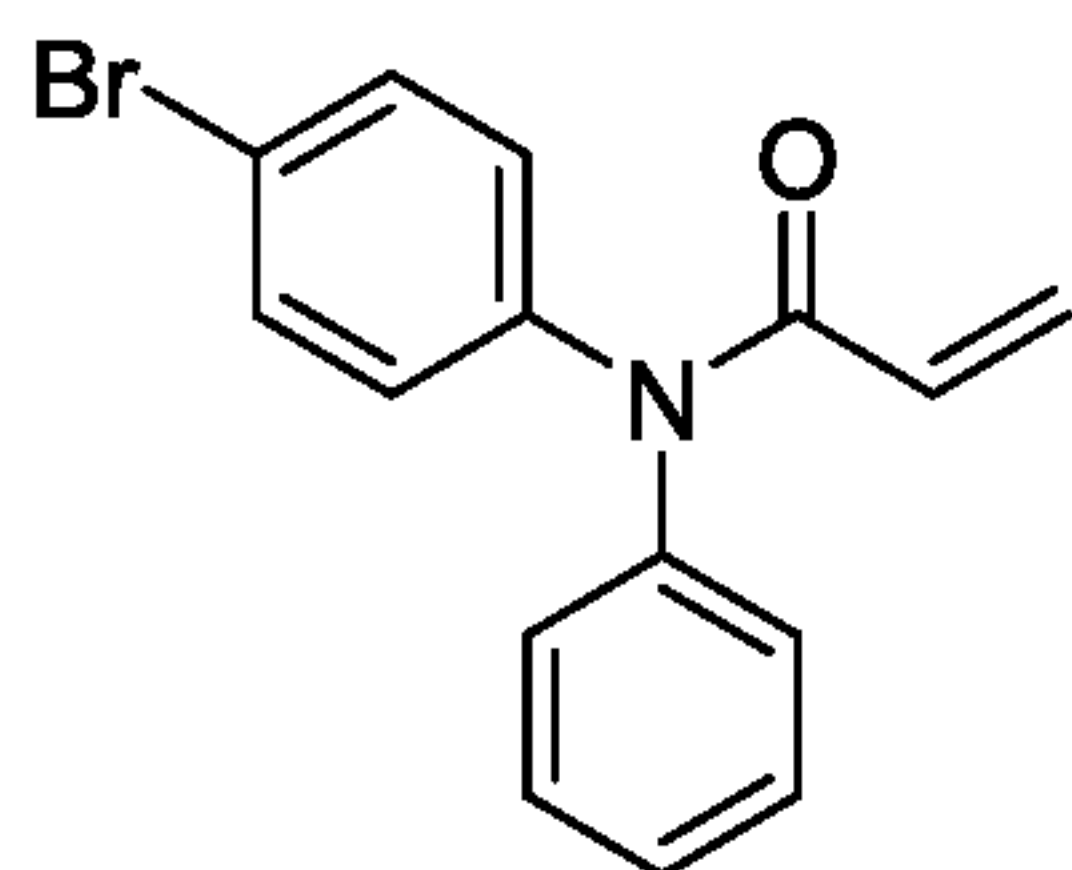
N-(hex-5-yn-1-yl)-2-iodoacetamide (1)



[00358] To a solution of chloroacetamide **SI-1** (36.1 mg, 0.2 mmol) in acetone (1 mL, 0.2 M) was added sodium iodide (47 mg, 1.5 equiv.) and the reaction was stirred overnight. The next day the reaction was filtered through a plug of silica eluting with 20% ethyl acetate in hexanes, and the filtrate was concentrated to yield a 10:1 mixture of the desired iodoacetamide **1** and starting material. This mixture was re-subjected to the reaction conditions for one further day, at which point complete conversion was observed. The product was purified by silica gel

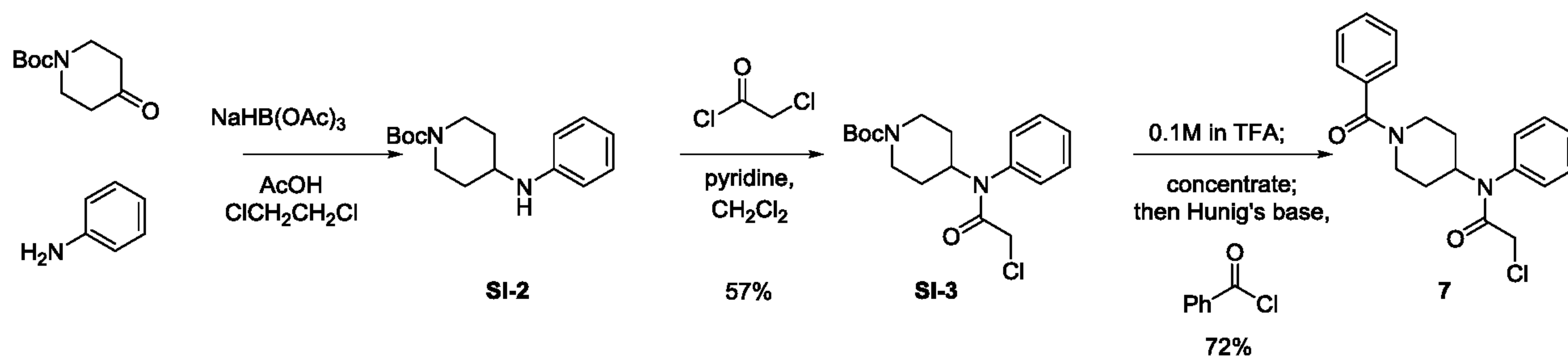
chromatography, utilizing a gradient of 5 to 10 to 15 to 20% ethyl acetate in hexanes to yield the desired product (24 mg, 44%). In some embodiments, the reaction is performed with 2.5 equiv. of sodium iodide, in which case re-subjection is not necessary, and purification by PTLC is accomplished in 30% EtOAc/hexanes as eluent. ^1H NMR (500 MHz, Chloroform-*d*) δ 6.16 (s, 1H), 3.69 (s, 2H), 3.30 (q, $J = 6.8$ Hz, 2H), 2.23 (td, $J = 6.8, 2.6$ Hz, 2H), 1.97 (t, $J = 2.6$ Hz, 1H), 1.75 – 1.61 (m, 2H), 1.61 – 1.52 (m, 2H).

N-(4-bromophenyl)-*N*-phenylacrylamide (5)

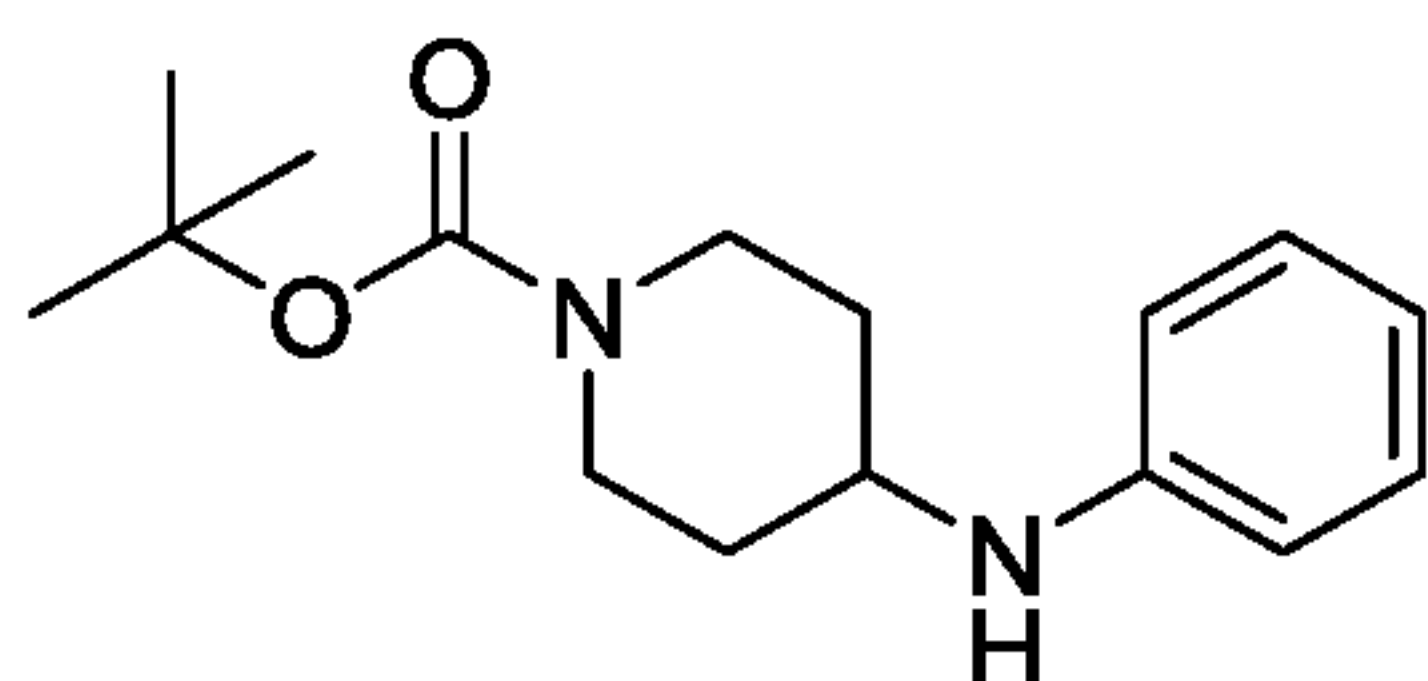


[00359] The title compound was synthesized according to General Procedure C from 4-bromophenylaniline (18.9 mg, 0.0762 mmol, 1 equiv.). Purification of the crude product by prep. TLC (30% EtOAc / hexanes) provided the title compound as a white solid (12.5 mg, 54%). ^1H NMR (500 MHz, Chloroform-*d*) δ 7.47 (d, $J = 8.2$ Hz, 2H), 7.39 (t, $J = 7.6$ Hz, 2H), 7.32 (d, $J = 7.4$ Hz, 1H), 7.21 (d, $J = 7.7$ Hz, 2H), 7.12 (d, $J = 8.2$ Hz, 2H), 6.48 (d, $J = 16.7$ Hz, 1H), 6.17 (dd, $J = 16.8, 10.3$ Hz, 1H), 5.65 (d, $J = 10.3$ Hz, 1H); HRMS-ESI (m/z) calculated for $\text{C}_{15}\text{H}_{13}\text{BrNO}$ [$\text{M}+\text{H}$]: 302.0175; found: 302.0176.

Synthesis of 7

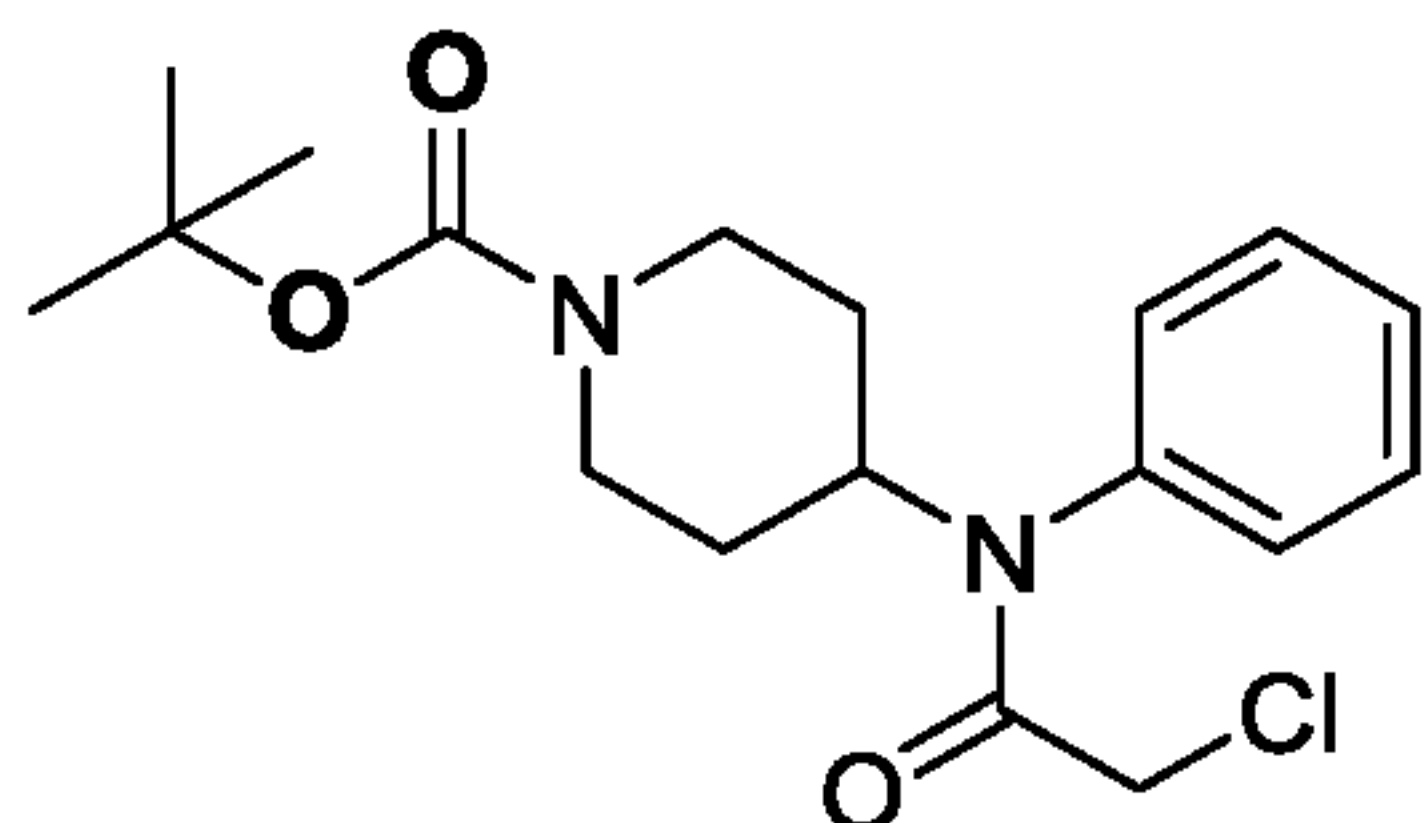


tert-butyl 4-(phenylamino)piperidine-1-carboxylate (SI-2)



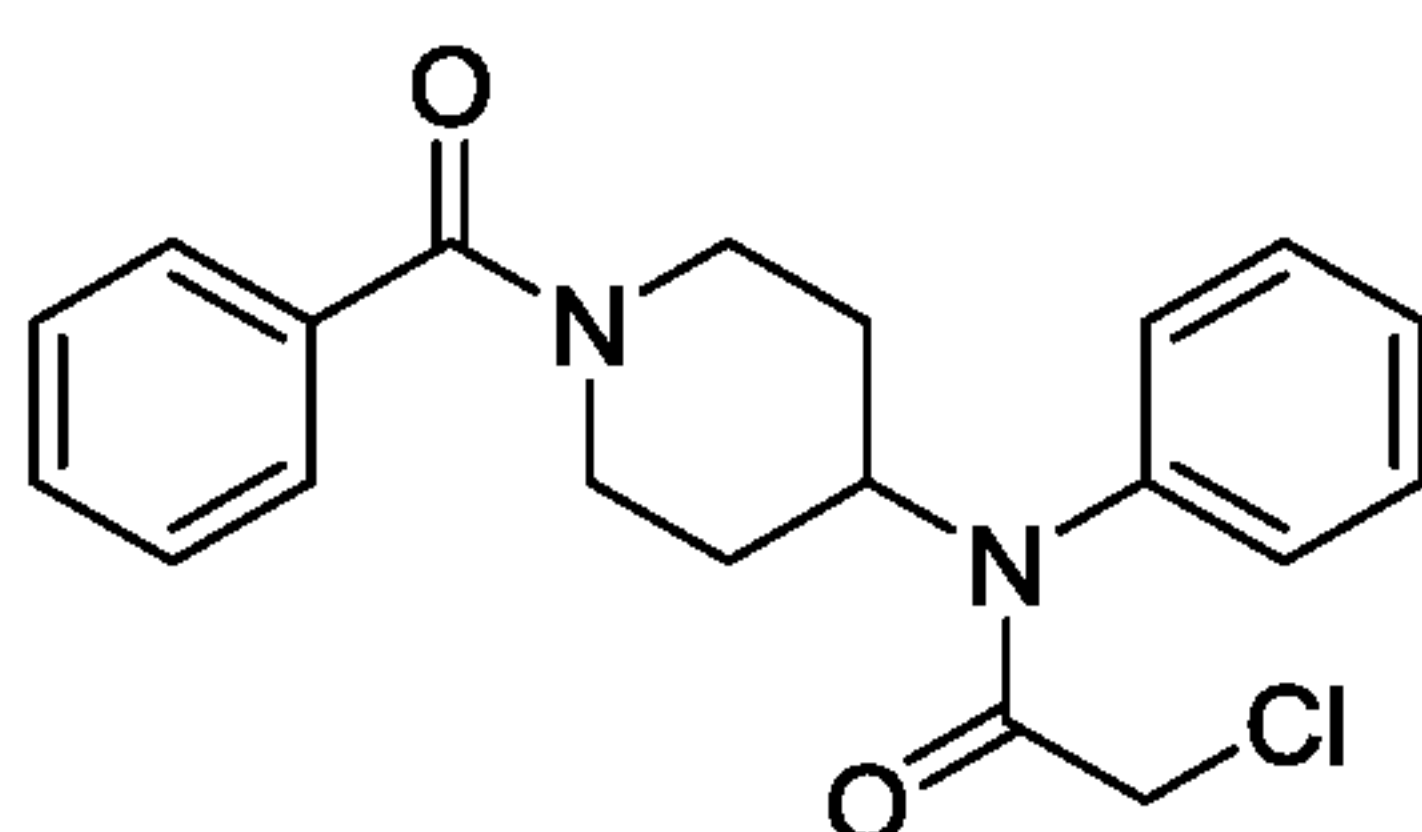
SI-2 was prepared according to Thoma *et al*, *J. Med. Chem.* **47**:1939-1955 (2004). ^1H NMR (400 MHz, Chloroform-*d*) δ 7.24 – 7.12 (m, 2H), 6.75 – 6.68 (m, 1H), 6.66 – 6.58 (m, 2H), 3.88 – 3.81 (m, 1H), 3.44 (tt, $J = 10.4, 3.9$ Hz, 2H), 3.00 – 2.88 (m, 2H), 2.10 – 1.99 (m, 2H), 1.48 (bs 9H), 1.41 – 1.27 (m, 2H).

tert-butyl 4-(2-chloro-*N*-phenylacetamido)piperidine-1-carboxylate (**SI-3**)



[00360] To a solution of aniline **SI-2** (65 mg, 0.24 mmol) at 0 °C in CH₂Cl₂ (0.6 mL) was added pyridine (38 μL, 2 equiv.) followed by chloroacetyl chloride (37.4 μL, 2.0 equiv.) in CH₂Cl₂ (0.6 mL). The resulting solution was allowed to warm to room temperature and stirred overnight. The solution was then quenched with saturated aqueous sodium bicarbonate, extracted with Et₂O (3 × 10 mL). The combined organic layers were dried over magnesium sulfate, filtered and concentrated to give an off-white solid, which was used without further purification (47 mg, 57%). ¹H NMR (400 MHz, Chloroform-*d*) δ 7.47 – 7.38 (m, 3H), 7.18 – 7.03 (m, 2H), 4.75 – 4.63 (m, 1H), 4.07 (s, 2H), 3.68 (s, 2H), 2.76 (s, 2H), 1.84 – 1.69 (m, 2H), 1.35 (s, 9H), 1.27 – 1.12 (m, 2H).

N-(1-benzoylpiperidin-4-yl)-2-chloro-*N*-phenylacetamide (**7**)

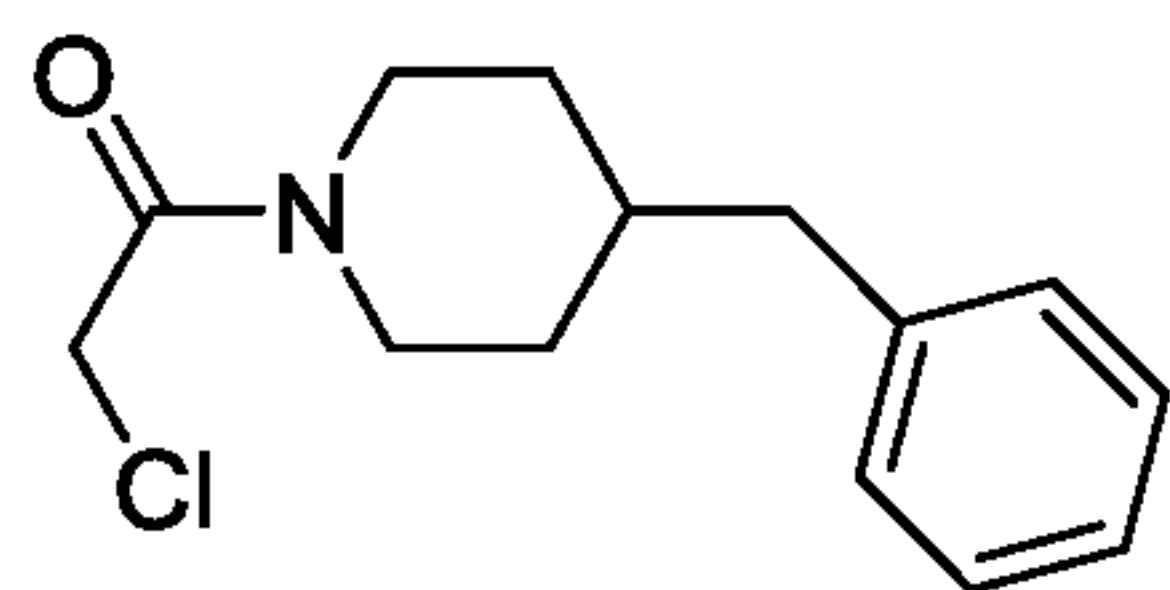


[00361] To neat **SI-3** (47 mg, 0.128 mmol) was added trifluoroacetic acid (0.7 mL, final 0.2 M). The resulting solution was concentrated under a stream of nitrogen until no further evaporation was observed, providing the deprotected amine as its trifluoroacetate salt. This viscous gum was then treated with triethylamine in ethyl acetate (10% v/v, 2 mL; solution smokes upon addition). The resulting solution was concentrated to afford the free base, which contained only triethylammonium trifluoroacetate and the free amine by proton NMR. A stock solution was prepared by dissolving the resulting gum in CH₂Cl₂ (1.2 mL, ~0.1 M final).

[00362] The deprotected amine (0.3 mL of stock solution, 0.0319 mmol) was treated with Hunig's base (17.5 μL, 3 equiv.) and benzoyl chloride (7.6 μL, 2.0 equiv.). This solution was stirred overnight, quenched with saturated aqueous sodium bicarbonate, extracted with Et₂O (3 × 10 mL). The resulting solution was dried over magnesium sulfate, filtered and concentrated. The resulting oil was purified by silica gel chromatography (20% EtOAc/hexanes) to afford chloroacetamide **7** as a white solid (8.6 mg, 75%). ¹H NMR (500 MHz, Chloroform-*d*) δ 7.55 (dd, *J* = 5.5, 3.0 Hz, 3H), 7.50 – 7.32 (m, 5H), 7.21 (s, 2H), 4.92 (tt, *J* = 12.3, 4.0 Hz, 1H), 4.87 (s, 1H), 3.87 (s, 1H), 3.78 (s, 2H), 3.21 (s, 1H), 2.97 – 2.90 (m, 1H), 2.01 (s, 1H), 1.90 (s, 1H),

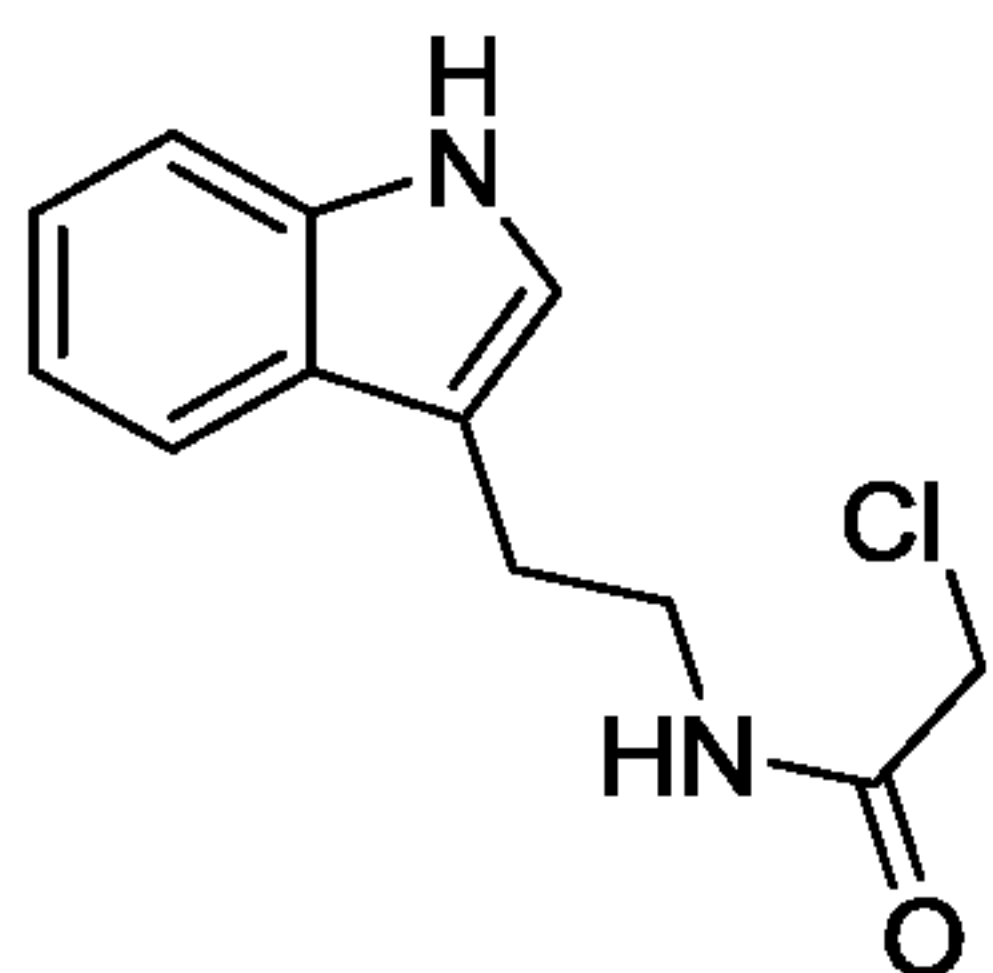
1.45 (s, 1H), 1.36 – 1.26 (m, 1H); HRMS-ESI (m/z) calculated for C₂₀H₂₂ClN₂O₂ [M+H]: 357.1364; found: 357.1362.

1-(4-benzylpiperidin-1-yl)-2-chloroethan-1-one (9)



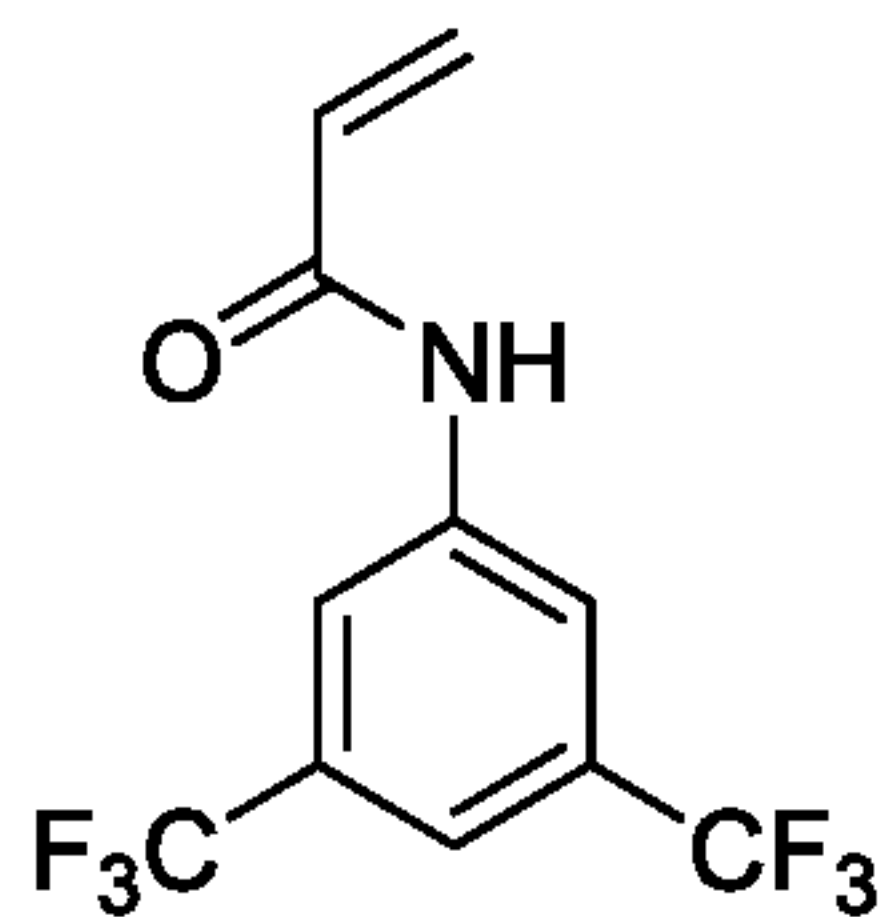
[00363] Following General Procedure A, starting from 4-benzylpiperidine (840 mg, 5.2 mmol, 1 equiv.), the desired compound was obtained after column chromatography as a yellow oil (1 g, 81%). Spectroscopic data matches those reported previously reported in Papadopoulou *et al. J. Med. Chem.* **55**:5554-5565 (2012). ¹H NMR (500 MHz, Chloroform-*d*) δ 7.42 – 7.14 (m, 5H), 4.61 (d, *J* = 13.4 Hz, 1H), 4.14 (q, *J* = 21.9, 11.5 Hz, 2H), 3.89 (d, *J* = 13.5, 1H), 3.11 (td, *J* = 13.1, 2.7 Hz, 1H), 2.69 – 2.57 (m, 3H), 1.92 – 1.75 (m, 3H), 1.40 – 1.21 (m, 2H); HRMS-ESI (m/z) calculated for C₁₄H₁₉ClNO [M+H]: 252.115; found: 252.115.

N-(2-(1H-indol-3-yl)ethyl)-2-chloroacetamide (13)

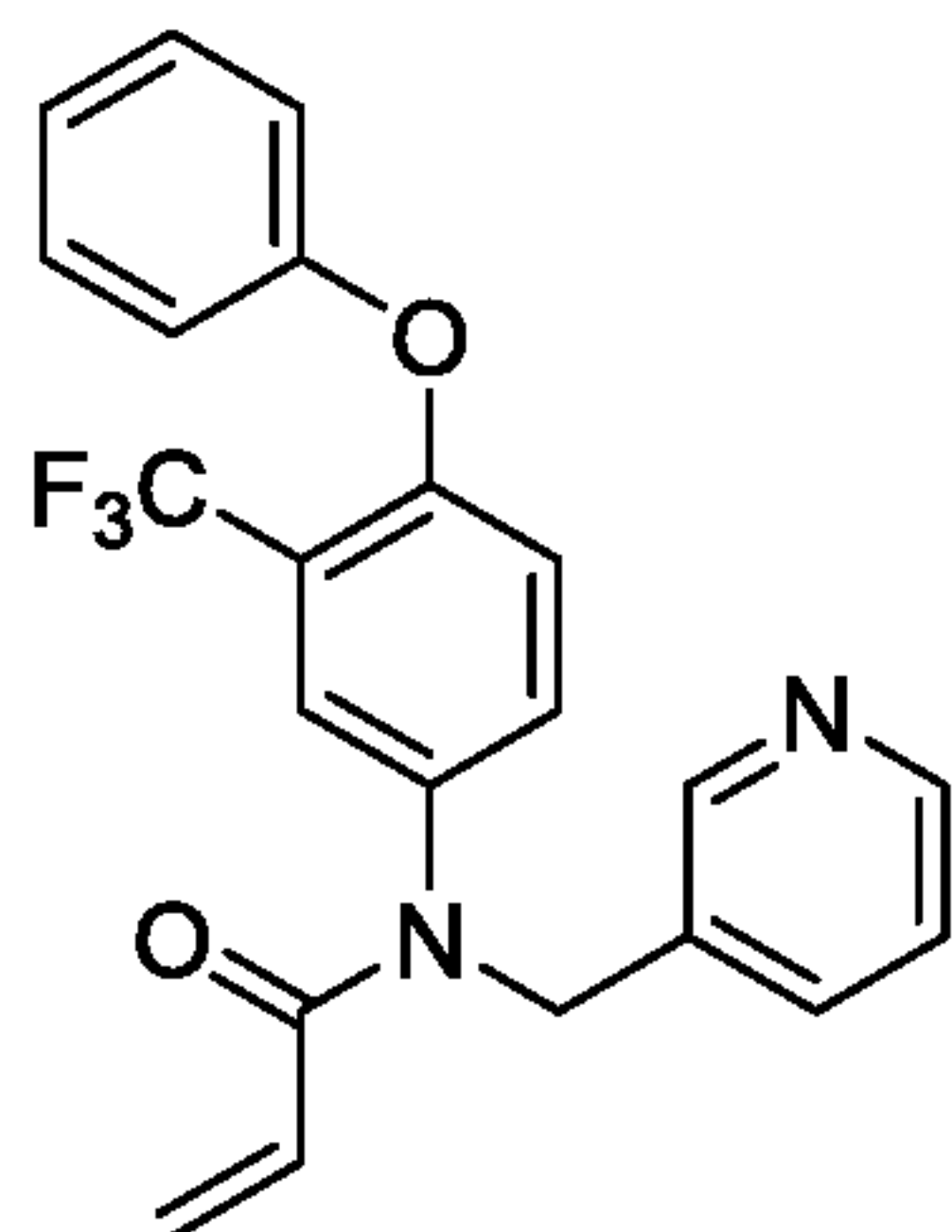


[00364] Following General Procedure A, starting from tryptamine (400 mg, 2.5 mmol, 1 equiv.), the desired compound was obtained after column chromatography as a brownish solid (460 mg, 77%). ¹H NMR (500 MHz, Chloroform-*d*) δ 8.55 (s, 1H), 7.70 (d, *J* = 7.9 Hz, 1H), 7.45 (d, *J* = 8.1 Hz, 1H), 7.30 (t, *J* = 7.5 Hz, 1H), 7.23 (t, *J* = 7.4 Hz, 1H), 7.10 (s, 1H), 6.84 (s, 1H), 4.08 (s, 2H), 3.72 (q, *J* = 6.4 Hz, 2H), 3.10 (t, *J* = 6.8 Hz, 2H); HRMS-ESI (m/z) calculated for C₁₂H₁₄ClN₂O₂ [M+H]: 237.0789; found: 237.0791.

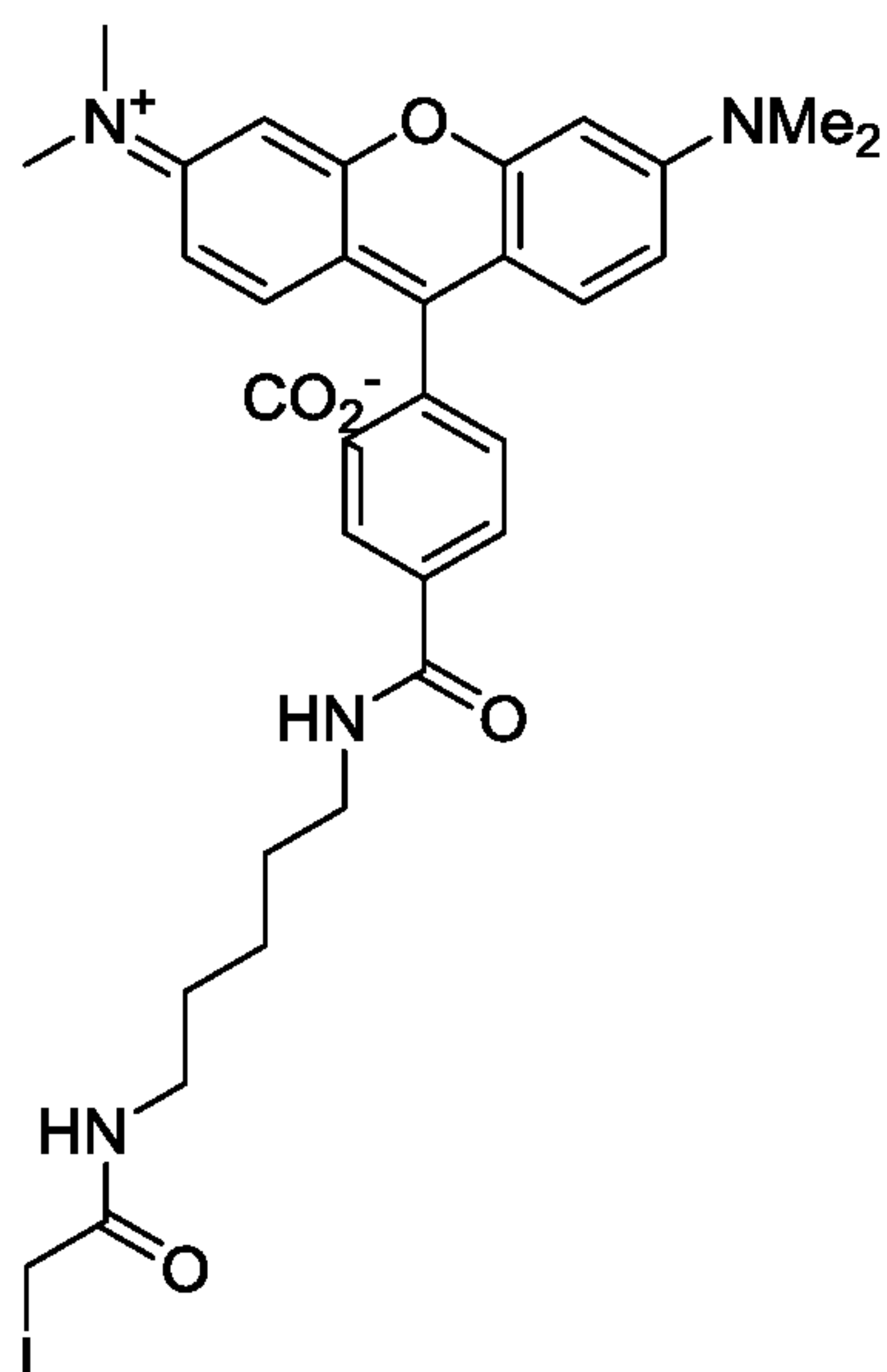
N-(3,5-bis(trifluoromethyl)phenyl)acrylamide (14)



[00365] Following General Procedure B, starting from 3,5-bis(trifluoromethyl)aniline (1.16 g, 5 mmol, 1 equiv.), the desired compound was obtained after column chromatography as a white solid (1.05 g, 74%). ¹H NMR (500 MHz, Chloroform-*d*) δ 8.33 (s, 1H), 8.18 (s, 2H), 7.68 (s, 1H), 6.57 (d, *J* = 17.5 Hz, 1H), 6.38 (dd, *J* = 16.9, 10.3 Hz, 1H), 5.93 (d, *J* = 12.5 Hz, 1H); HRMS-ESI (m/z) calculated for C₁₁H₈F₆NO₂ [M+H]: 284.0505; found: 284.0504.

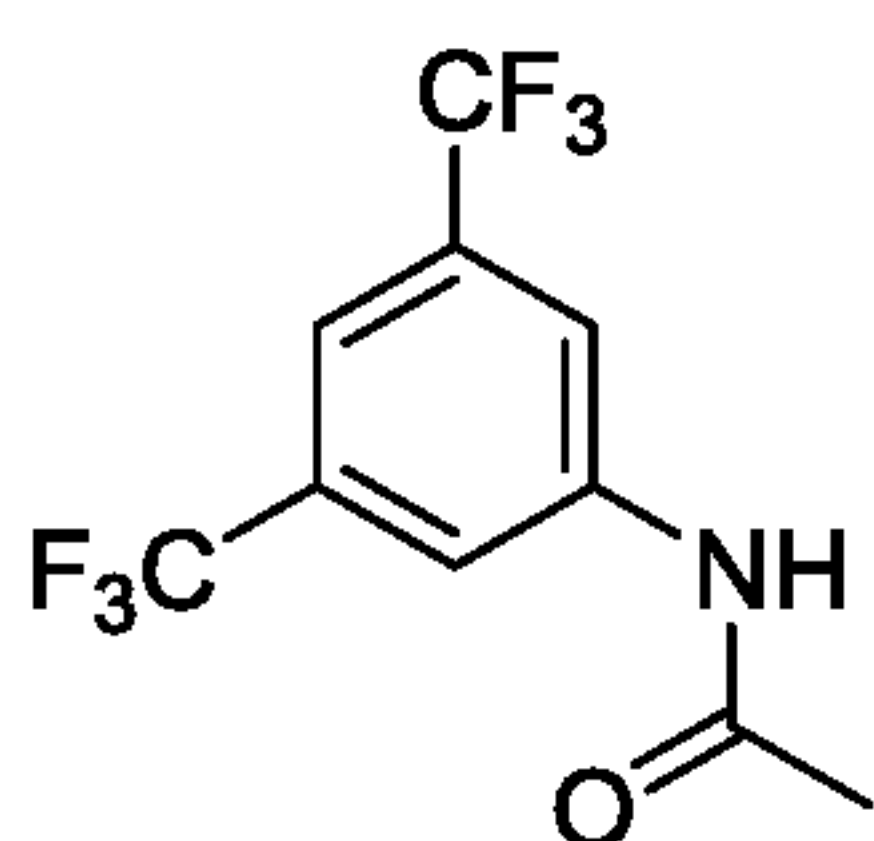
N-(4-phenoxy-3-(trifluoromethyl)phenyl)-*N*-(pyridin-3-ylmethyl)acrylamide (15)

[00366] 4-phenoxy-3-(trifluoromethyl)aniline (260 mg, 1 mmol, 1 equiv.) (Combi-Blocks) was dissolved in TFA (5 mL). Following the reductive amination protocol reported by Boros *et al. J. Org. Chem* **74**:3587-3590 (2009), the reaction mixture was cooled to 0 °C and to this sodium triacetoxyborohydride (STAB) (270 mg, 1.3 mmol, 1.3 equiv.) was added. 3-pyridinecarboxaldehyde (200 mg, 2 mmol, 2 equiv.) was dissolved in CH₂Cl₂ (5 mL) and slowly added to the reaction mixture. Upon complete conversion to product, the reaction was diluted with CH₂Cl₂ (20 mL) and washed with saturated sodium bicarbonate solution (3 × 20 mL) and the organic layer was dried then concentrated under reduced pressure. Without further purification the crude material was dissolved in anhydrous CH₂Cl₂ and subjected to General Procedure B. The resulting crude was purified by prep. TLC to give a white solid (31 mg, 10%). ¹H NMR (500 MHz, Chloroform-*d*) δ 8.52 (d, *J* = 3.5 Hz, 1H), 8.39 (s, 1H), 7.68 (d, *J* = 7.8 Hz, 1H), 7.40 (t, *J* = 7.7 Hz, 2H), 7.34 (s, 1H), 7.28 – 7.18 (m, 2H), 7.07 (d, *J* = 8.2 Hz, 2H), 6.98 (d, *J* = 7.5 Hz, 1H), 6.82 (d, *J* = 8.8 Hz, 1H), 6.46 (d, *J* = 16.8 Hz, 1H), 6.01 (dd, *J* = 16.2, 10.7 Hz, 1H), 5.64 (d, *J* = 10.3 Hz, 1H), 4.96 (s, 2H). HRMS-ESI (*m/z*) calculated for C₂₂H₁₈F₃N₂O₂ [M+H]⁺; 399.1315; found: 399.1315.

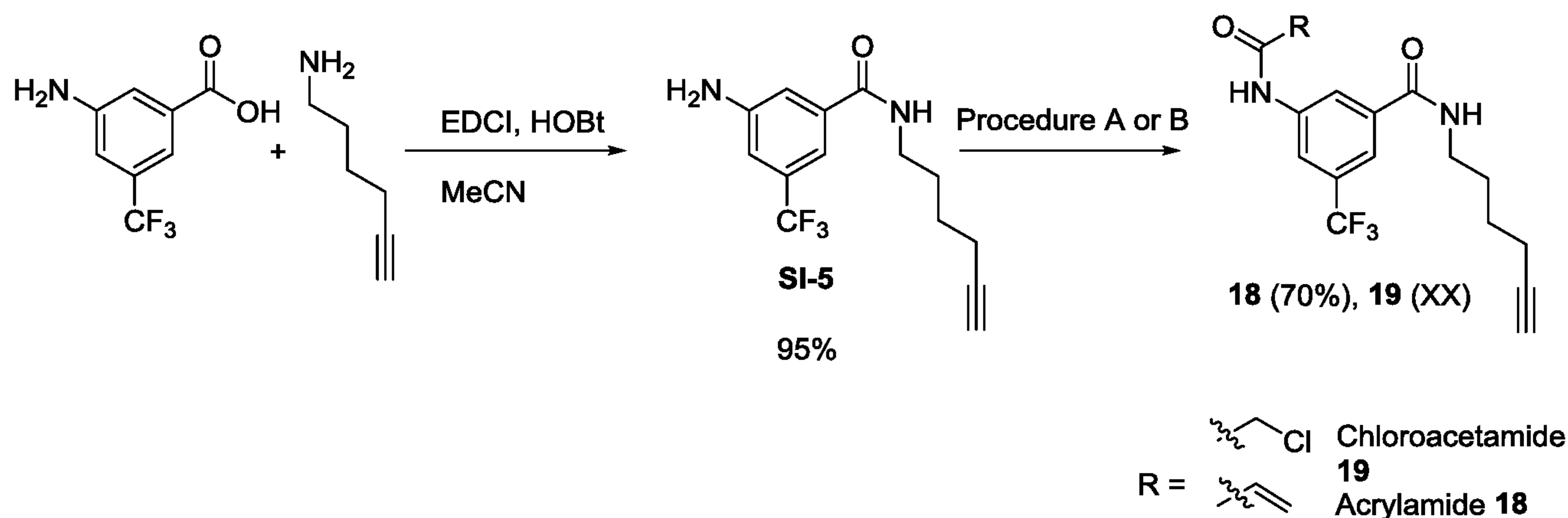
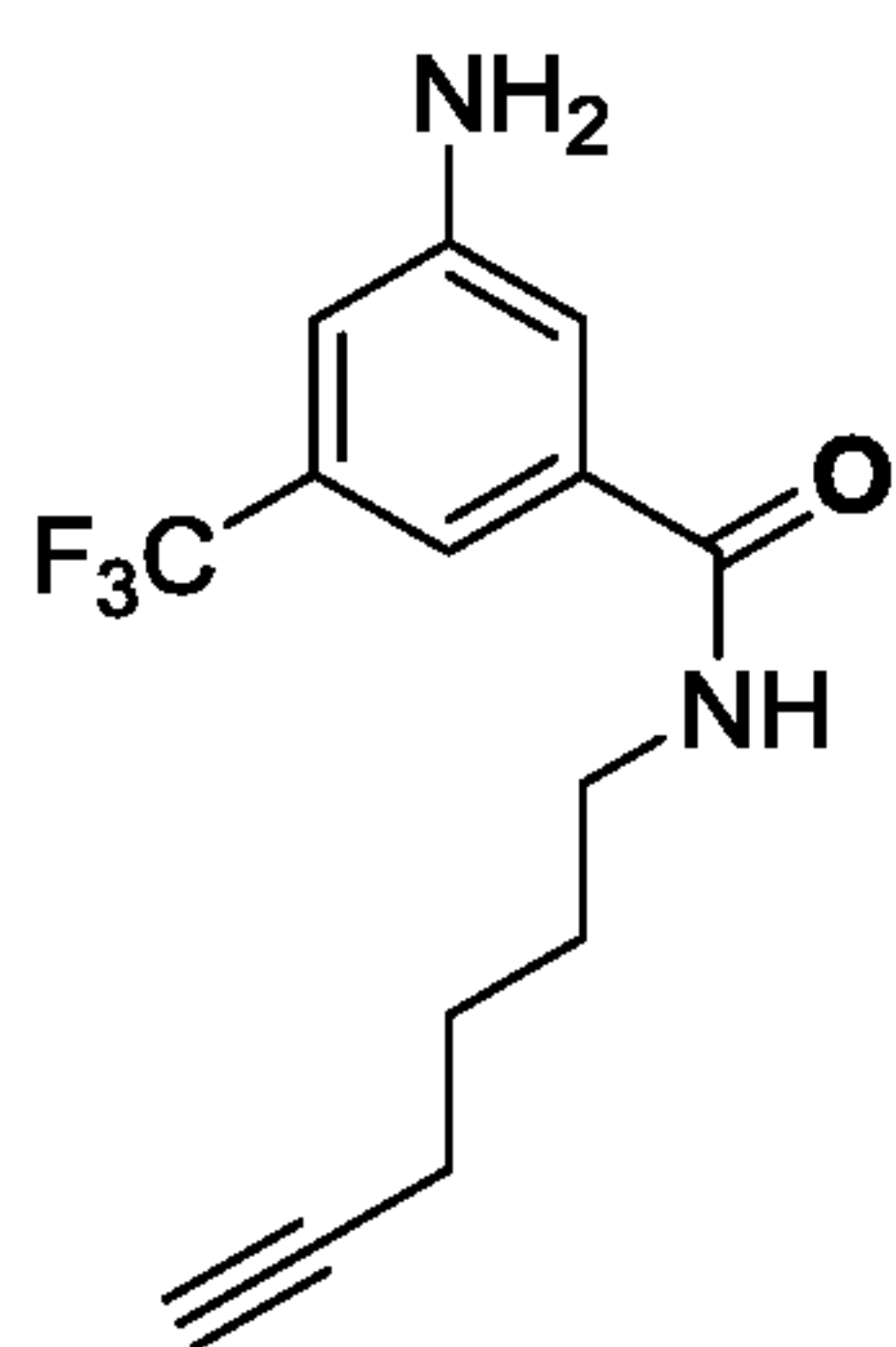
Iodoacetamide-rhodamine (16)

[00367] 5-(and-6)-((*N*-(5-aminopentyl)amino)carbonyl)tetramethylrhodamine (tetramethylrhodamine cadaverine) mixed isomers (60 mg, 0.12 mmol, 1 equiv.) were dissolved in anhydrous DMF (500 μ L) with sonication. To this was added DIPEA (60 μ L, 0.34 mmol, 3 equiv.) and chloroacetyl chloride (10 μ L, 0.13 mmol, 1 equiv., diluted 1:10 in DMF) and the reaction was stirred at room temperature for 20 min until complete conversion to the product was detected by TLC. The DMF was removed under a stream of nitrogen and the reaction mixture was separated by PTLC in MeOH:CH₂Cl₂:TEA (15:85:0.001). The chloroacetamide rhodamine was then eluted in MeOH:CH₂Cl₂ (15:85), concentrated under reduced pressure and redissolved in acetone (500 μ L). NaI (150 mg, 1 mmol, 10 equiv.) was added to this and the reaction was stirred for 20 min at 50 °C until complete conversion to product was detected and the crude reaction mixture was purified by reverse phase HPLC on a C18 column and concentrated to yield the title compound as a purple solid that is a mixture of 5 and 6 carboxamide tetramethylrhodamine isomers (ratio ~ 6:1) (10 mg, 12 %). ¹H NMR (600 MHz, Methanol-*d*₄) δ 8.87 (t, *J* = 4.8 Hz, 0.14 H), 8.80 – 8.71 (m, 1H), 8.41 (dd, *J* = 8.2, 1.1 Hz, 0.86H), 8.35 (br s, 1H), 8.27 (dt, *J* = 7.9, 1.5 Hz, 0.164 H), 8.20 (dt, *J* = 8.2, 1.5 Hz, 0.86H), 7.81 (s, 0.86H), 7.53 (d, *J* = 7.8 Hz, 0.14 H), 7.18 – 7.11 (m, 2H), 7.07 (d, *J* = 9.5 Hz, 2H), 7.00 (s, 2H), 3.68 – 3.62 (m, 2H), 3.46 – 3.37 (m, 2H), 3.31 (s, 12H, obscured by solvent) 3.21 – 3.12 (m, 2H), 1.81 – 1.21 (m, 6H); HRMS-ESI (*m/z*) calculated for C₃₂H₃₆IN₄O₅ [*M*+*H*]: 683.1725; found: 683.1716.

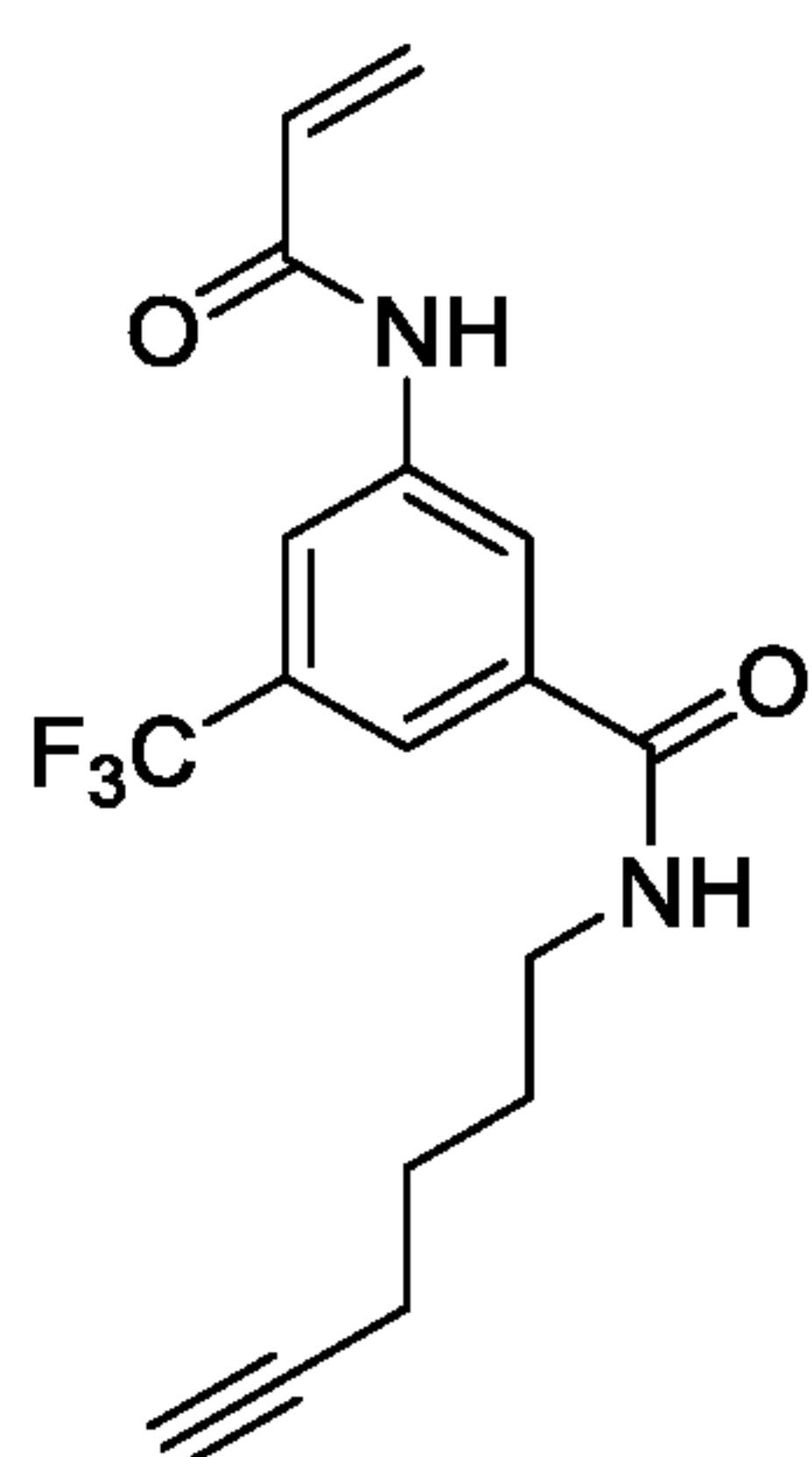
N-(3,5-bis(trifluoromethyl)phenyl)acetamide (17)



[00368] Following General Procedure A, starting with 3,5-bis(trifluoromethyl)aniline (327 mg, 1.42 mmol, 1 equiv.) and acetic anhydride (200 μ L, 3 mmol, 2 equiv.), the title compound was obtained after PTLC as a white solid (302 mg, 78%). ¹H NMR (500 MHz, Chloroform-*d*) δ 8.10 (s, 2H), 7.72 (s, 1H), 7.68 (s, 1H), 2.32 (d, *J* = 0.9 Hz, 3H). HRMS-ESI (*m/z*) calculated for C₁₁H₈F₆NO₂ [*M*+*H*]: 284.0505; found: 284.0504.

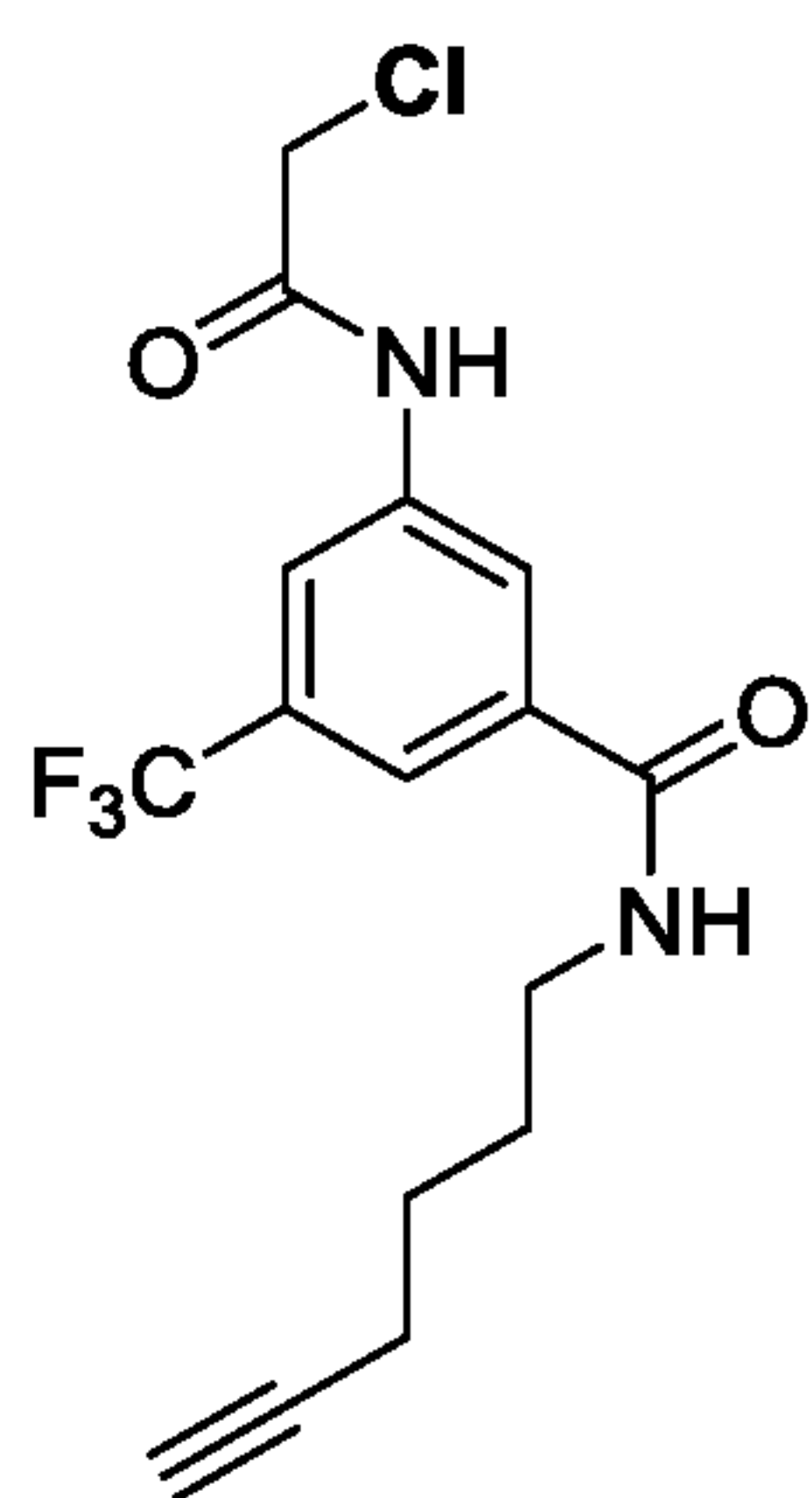
Synthesis of 18 and 19*3-amino-N-(hex-5-yn-1-yl)-5-(trifluoromethyl)benzamide (SI-5)*

[00369] To a solution of 3-amino-5-(trifluoromethyl)benzoic acid (74 mg, 0.36 mmol) in acetonitrile (3.6 mL, 0.1 M final) was added EDCI (83 mg, 1.2 equiv.) followed by hex-5-ynamine (35 mg, 1.0 equiv.) followed by 1-hydroxybenzotriazole hydrate (HOBt, 66.3 mg, 1.2 equiv.) and the resulting solution was stirred overnight. The reaction was diluted with ethyl acetate, washed with 1 M HCl twice and then brine. The organic layer was dried over magnesium sulfate and concentrated to yield aniline **SI-5** (97.4 mg, 95%) as a white solid. ^1H NMR (400 MHz, Chloroform-*d*) δ 7.29 – 7.22 (m, 2H), 6.98 (t, J = 1.8 Hz, 1H), 6.38 (t, J = 5.5 Hz, 1H), 4.08 (s, 2H), 3.46 (td, J = 7.1, 5.7 Hz, 2H), 2.25 (td, J = 6.9, 2.6 Hz, 2H), 1.99 (t, J = 2.7 Hz, 1H), 1.81 – 1.55 (m, 4H).

3-acrylamido-N-(hex-5-yn-1-yl)-5-(trifluoromethyl)benzamide (18)

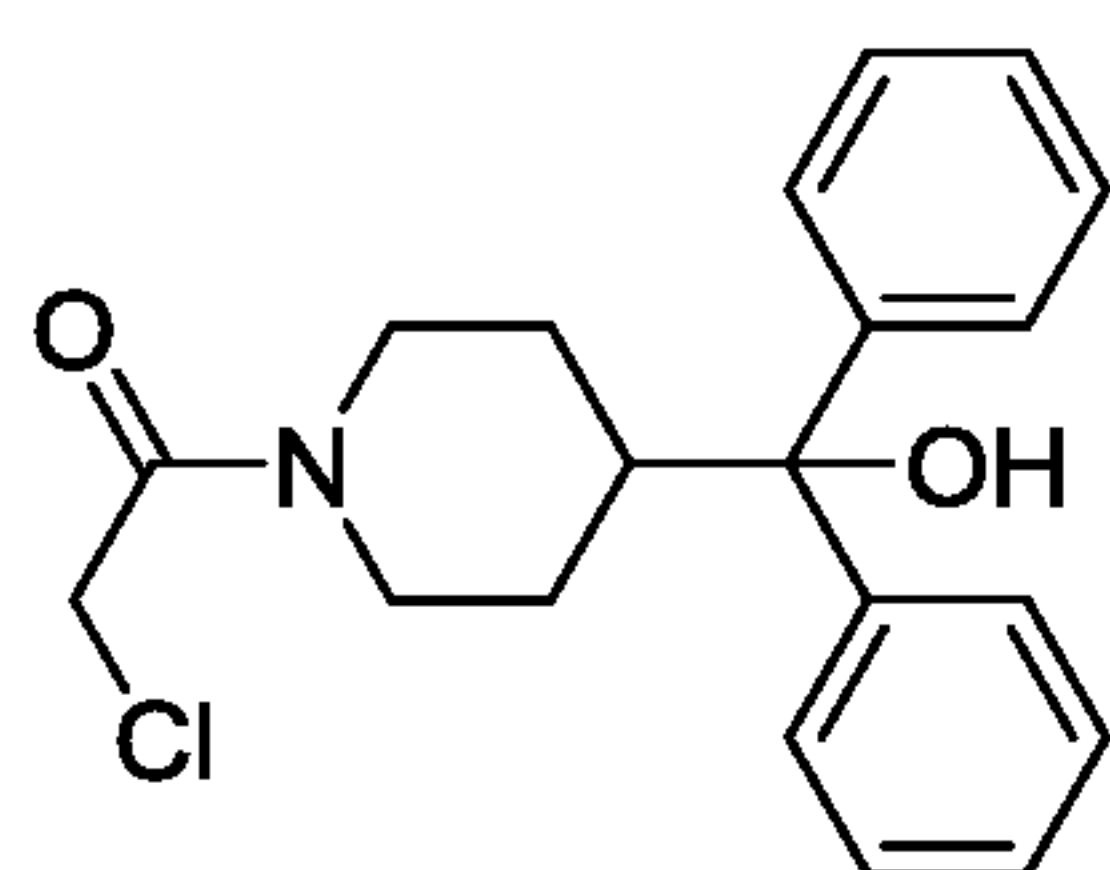
[00370] Following General Procedure B, starting with **SI-5** (42 mg, 0.15 mmol, 1 equiv.), the title compound was obtained after column chromatography as a white solid (34 mg, 70%). ¹H NMR (500 MHz, Chloroform-*d*) δ 8.94 (s, 1H), 8.24 (d, J = 11.9 Hz, 2H), 7.71 (s, 1H), 6.87 (t, J = 5.7 Hz, 1H), 6.55 (dd, J = 17.4, 0.7 Hz, 1H), 6.43 (dd, J = 16.9, 10.1 Hz, 1H), 5.88 (dd, J = 10.1, 1.3 Hz, 1H), 3.56 (q, J = 6.7 Hz, 2H), 2.33 (td, J = 6.9, 2.7 Hz, 2H), 2.06 (t, J = 2.7 Hz, 1H), 1.87 (p, J = 7.3 Hz, 2H), 1.69 (p, J = 7.8 Hz, 2H); HRMS-ESI (m/z) calculated for C₁₇H₁₈F₃N₂O₂ [M+H]: 339.1314; found 339.1313.

3-acrylamido-N-(hex-5-yn-1-yl)-5-(trifluoromethyl)benzamide (19)



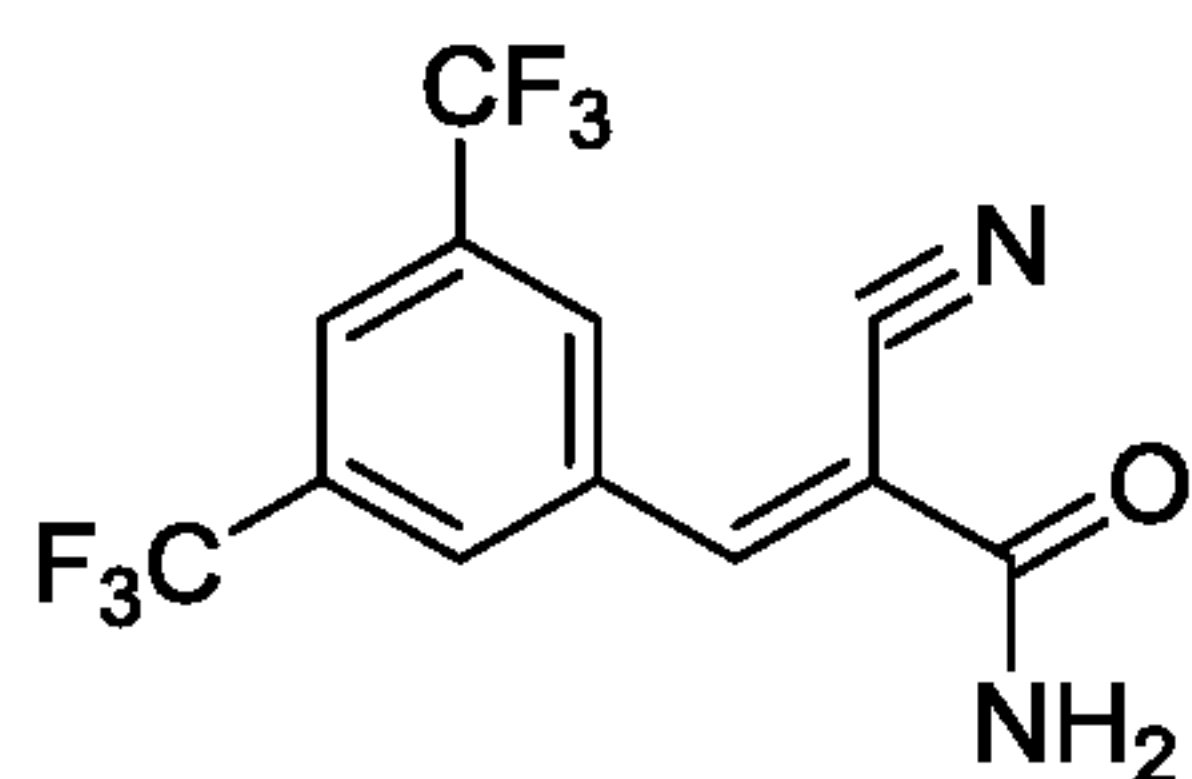
[00371] Synthesized according to General Procedure A2, starting from **SI-5**. ¹H NMR (600 MHz, Chloroform-*d*) δ 8.57 (s, 1H), 8.16 (t, J = 1.8 Hz, 1H), 8.05 (t, J = 1.8 Hz, 1H), 7.79 (d, J = 2.0 Hz, 1H), 6.38 (d, J = 6.1 Hz, 1H), 4.23 (s, 2H), 3.51 (td, J = 7.1, 5.7 Hz, 2H), 2.27 (td, J = 6.9, 2.7 Hz, 2H), 2.00 (t, J = 2.6 Hz, 1H), 1.82 – 1.74 (m, 2H), 1.71 – 1.59 (m, 2H); HRMS-ESI (m/z) calculated for C₁₆H₁₇ClF₃N₂O₂ [M+H]: 361.0925; found: 361.0925.

2-chloro-1-(4-(hydroxydiphenylmethyl)piperidin-1-yl)ethan-1-one (20)



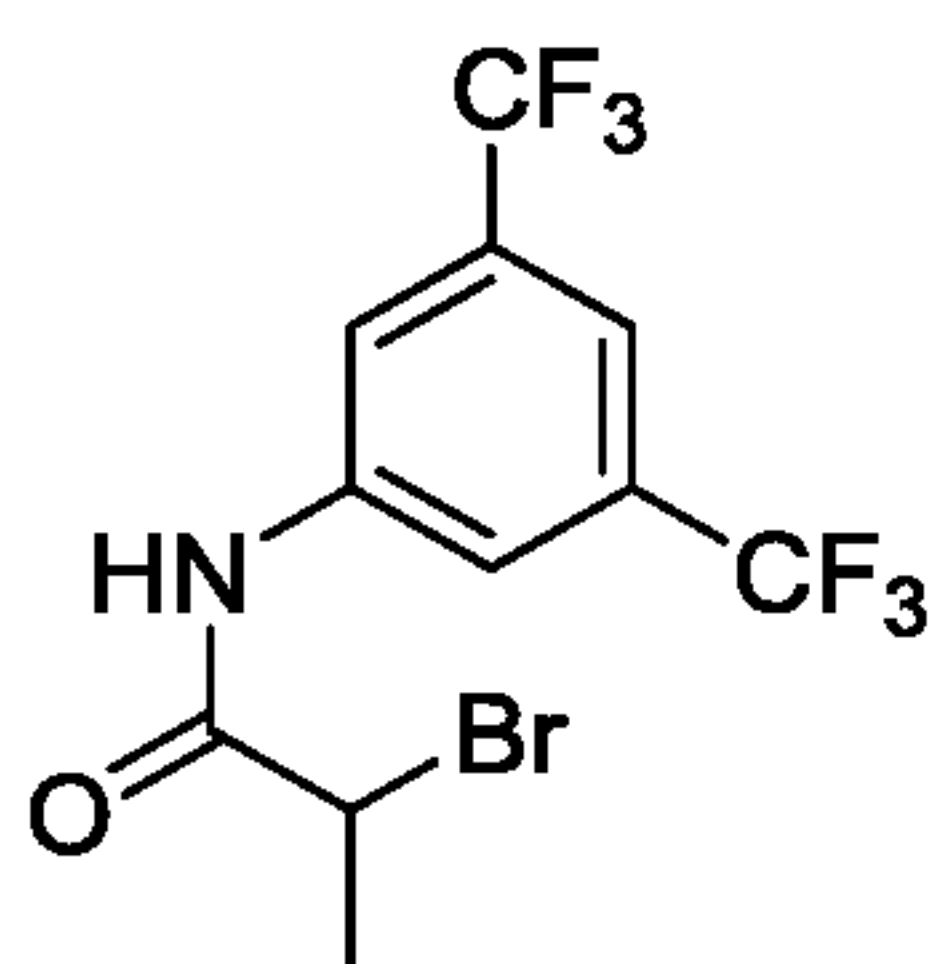
[00372] Following General Procedure A, starting with α,α -diphenyl-4-piperidinomethanol (800 mg, 3 mmol, 1 equiv.), the title compound was obtained after column chromatography as a white solid (637 mg, 61%). ¹H NMR (500 MHz, Chloroform-*d*) δ 7.56 (d, J = 7.6 Hz, 4H), 7.39 (q, J = 7.1 Hz, 4H), 7.28 (q, J = 6.8 Hz, 2H), 4.66 (d, J = 13.3 Hz, 1H), 4.07 (dd, J = 12.2, 4.2 Hz, 2H), 3.91 (d, J = 13.4 Hz, 1H), 3.18 (t, J = 12.9 Hz, 1H), 2.77 – 2.62 (m, 3H), 1.67 (t, J = 12.5 Hz, 2H), 1.56 (q, J = 11.8 Hz, 1H), 1.44 (q, J = 12.4, 11.8 Hz, 1H); HRMS-ESI (m/z) calculated for C₂₀H₂₃ClNO₂ [M+H]: 344.1412; found: 344.1412.

(*E*)-3-(3,5-bis(trifluoromethyl)phenyl)-2-cyanoacrylamide (23)



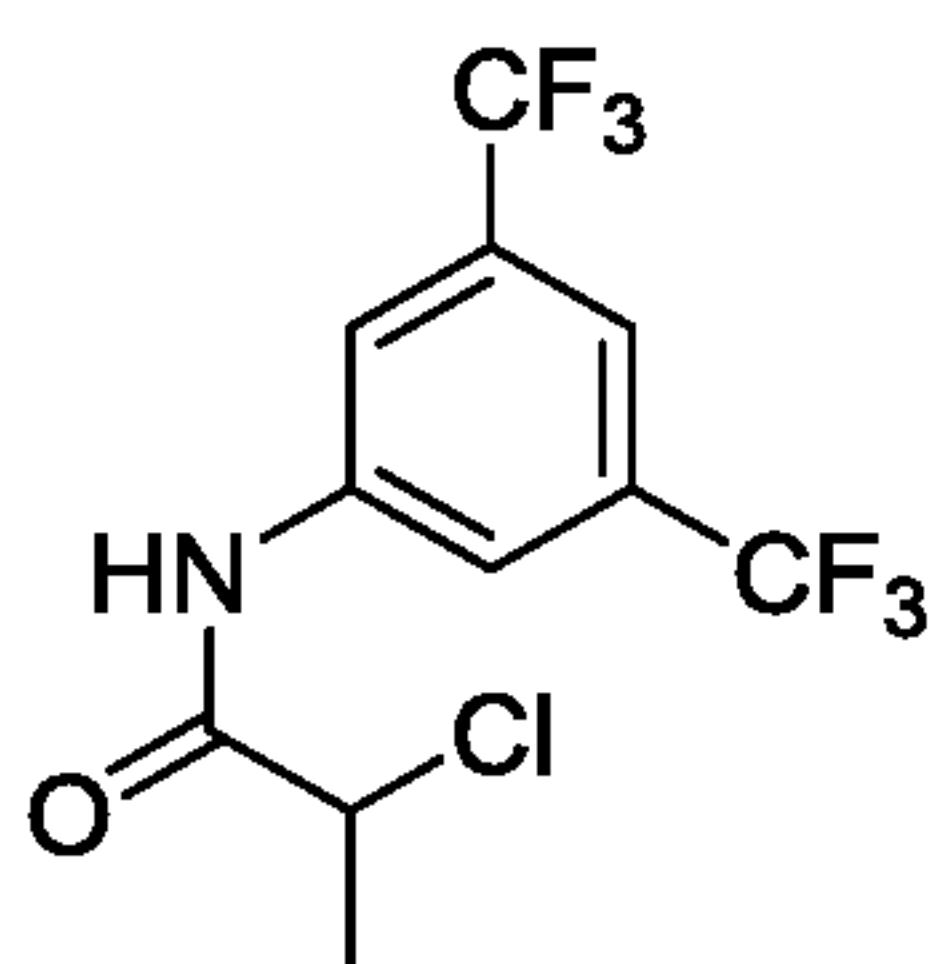
[00373] 3,5-bis(trifluoromethyl)benzaldehyde (880 mg, 3.6 mmol, 1 equiv.) and 2-cyanoacetamide (460 mg, 5.5 mmol, 1.5 equiv.) were dissolved in MeOH (10 mL). To this was added *piperidine* (214 mg, 0.7 equiv.) and the reaction was stirred at room temperature for 30 minutes at which point starting material was consumed. After addition of an equivalent volume of water (10 mL), the precipitate was collected by filtration and washed with water/methanol (1:1) to yield the title compound as a white solid (534 mg, 47%); ^1H NMR (400 MHz, Acetone- d_6) δ 8.78 (s, 2H), 8.61 (s, 1H), 8.41 (s, 1H), 7.57 (s, 1H), 7.42 (s, 1H); HRMS-ESI (m/z) calculated for $\text{C}_{12}\text{H}_7\text{F}_6\text{N}_2\text{O}_2$ [$\text{M}+\text{H}$]: 309.0457; found: 309.0459.

N-(3,5-bis(trifluoromethyl)phenyl)-2-bromopropanamide (24)

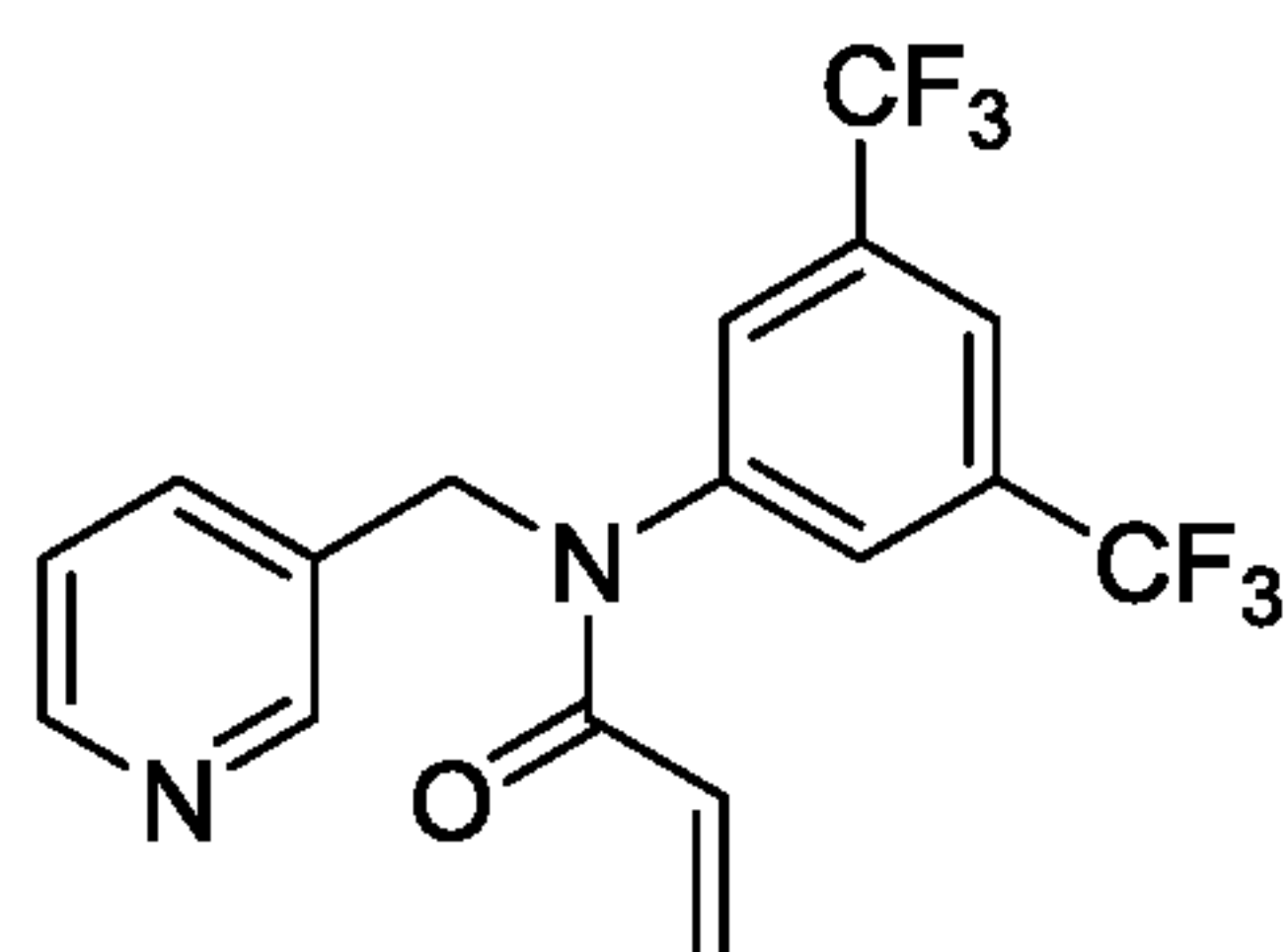


[00374] Following General Procedure A1, starting with 3,5-bis(trifluoromethyl)aniline (250 mg, 1.1 mmol, 1 equiv.) and 2-bromopropionyl chloride (200 μL , 2 mmol, 1.8 equiv.) the title compound was obtained by PTLC as a white solid (130 mg, 35%). ^1H NMR (500 MHz, Chloroform- d) δ 8.34 (s, 1H), 8.06 (s, 2H), 7.66 (s, 1H), 4.58 (q, $J = 7.0$ Hz, 1H), 1.98 (d, $J = 7.0$ Hz, 3H); HRMS-ESI (m/z) calculated for $\text{C}_{11}\text{H}_7\text{BrF}_6\text{NO}$ [$\text{M}-\text{H}$]: 361.9621; found: 361.9623.

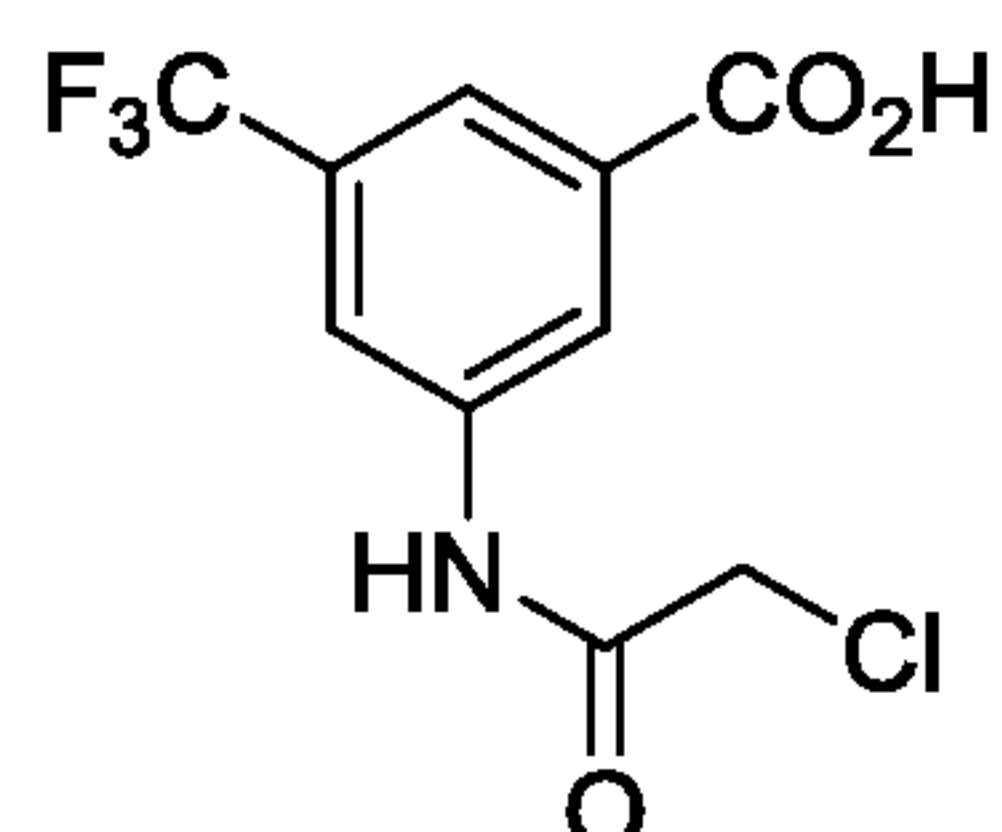
N-(3,5-bis(trifluoromethyl)phenyl)-2-chloropropanamide (25)



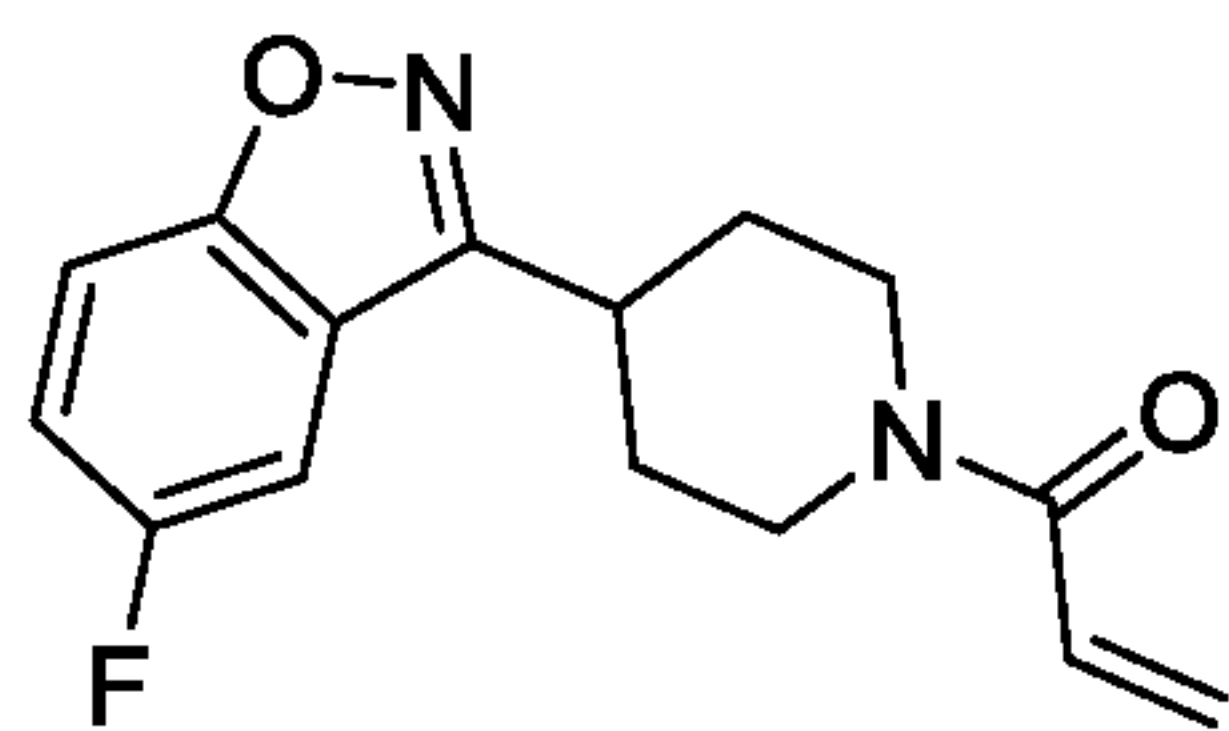
[00375] Following General Procedure A1, starting with 3,5-bis(trifluoromethyl)aniline (327 mg, 1.42 mmol, 1 equiv.) and 2-chloropropionyl chloride (200 μL , 2 mmol, 1.8 equiv.) the title compound was obtained by PTLC as a white solid (250 mg, 55%). ^1H NMR (500 MHz, Chloroform- d) δ 8.61 (s, 1H), 8.16 (s, 2H), 7.75 (s, 1H), 4.67 (q, $J = 7.1$ Hz, 1H), 1.93 (d, $J = 7.1$ Hz, 3H). HRMS-ESI (m/z) calculated for $\text{C}_{11}\text{H}_7\text{ClF}_6\text{NO}$ [$\text{M}-\text{H}$]: 318.0126; found: 318.0126.

N-(3,5-bis(trifluoromethyl)phenyl)-*N*-(pyridin-3-ylmethyl)acrylamide (31)

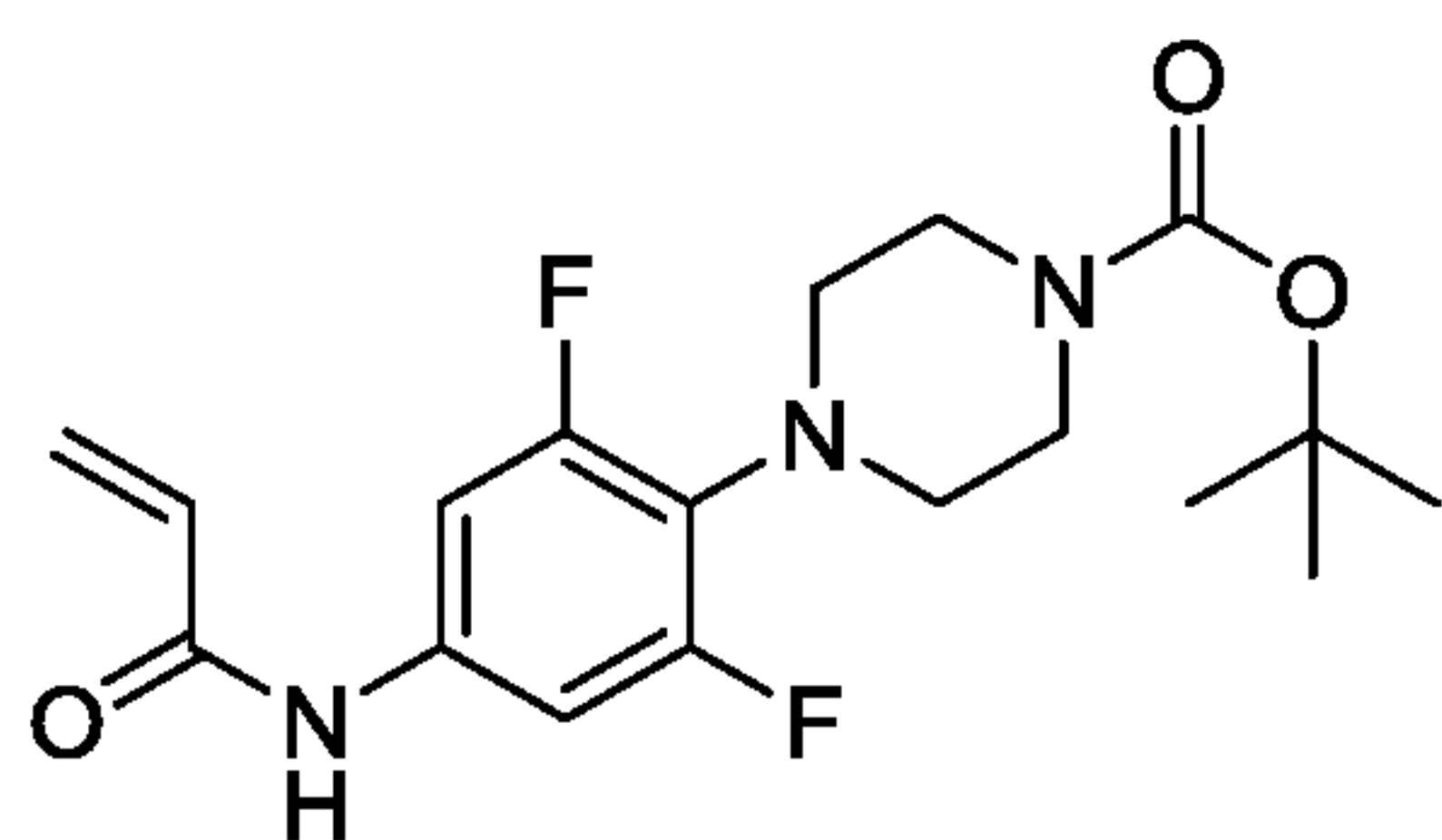
[00376] 3,5-bis(trifluoromethyl)aniline (350 mg, 1.6 mmol, 1 equiv.) was dissolved in TFA (5 mL). The reaction mixture was cooled to 0 °C and to this sodium triacetoxyborohydride (STAB) (400 mg, 2 mmol, 1.3 equiv.) was added. 3-pyridinecarboxaldehyde (244 mg, 1.5 mmol, 1 equiv.) was dissolved in CH₂Cl₂ (5 mL) and slowly added to the reaction mixture dropwise over 10 minutes. Upon complete conversion to product, the reaction mixture was diluted with CH₂Cl₂ (20 mL) and washed with saturated sodium bicarbonate solution (3 × 20 mL) and the organic layer was dried then concentrated under reduced pressure. Without further purification the crude material was dissolved in anhydrous CH₂Cl₂ and subjected to General Procedure B. The resulting crude was purified by PTLC to give a white solid (10 mg, 2%). ¹H NMR (500 MHz, Chloroform-*d*) δ 8.63 (d, *J* = 3.8 Hz, 1H), 8.49 (s, 1H), 7.93 (s, 1H), 7.70 (d, *J* = 7.7 Hz, 1H), 7.55 (s, 2H), 7.35 (dd, *J* = 7.6, 5.3 Hz, 1H), 6.60 (dd, *J* = 16.6, 1.6 Hz, 1H), 6.02 (dd, *J* = 16.9, 10.2 Hz, 1H), 5.79 (dd, *J* = 10.3, 1.6 Hz, 1H), 5.11 (s, 2H). HRMS-ESI (*m/z*) calculated for C₁₇H₁₃F₆N₂O [*M*+*H*]: 375.0927; found: 375.0928.

3-(2-chloroacetamido)-5-(trifluoromethyl)benzoic acid (36)

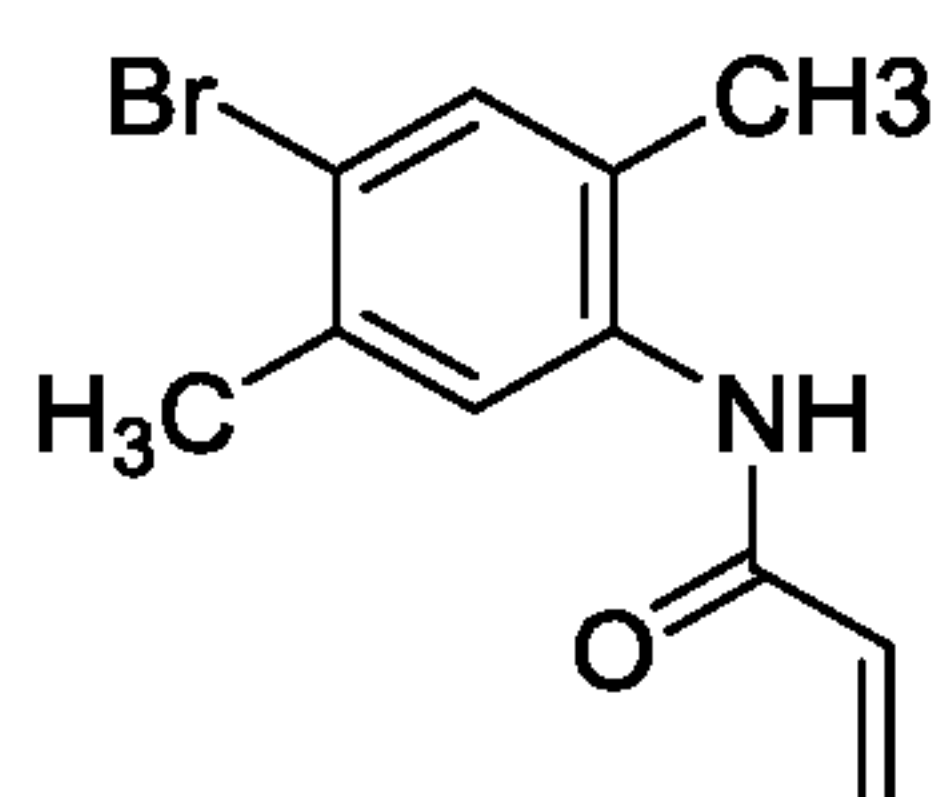
[00377] To a solution of 3-amino-5-(trifluoromethyl)benzoic acid (500 mg, 2.44 mmol) in 1.5 mL of dimethylacetamide (1.6 M) at 0 °C was added chloroacetyl chloride (214 μL, 2.69 mmol, 1.1 equiv.). The resulting solution was warmed to ambient temperature and stirred for 20 minutes, at which point ethyl acetate (40 mL) and water (30 mL) were added. The pH of the aqueous layer was adjusted to pH 10 via addition of 1 N NaOH, and the phases were separated. The aqueous layer was washed with 40 mL of ethyl acetate, then acidified by adding 1 N HCl. The product was extracted with ethyl acetate (40 mL), and the organic layer was washed with 1M HCl (2 × 40 mL), brine (40 mL), dried over magnesium sulfate and concentrated to provide the desired product (456 mg, 66%). ¹H NMR (500 MHz, Chloroform-*d*) δ 8.31 (s, 1H), 8.27 (s, 1H), 8.14 (s, 1H), 4.13 (s, 2H); HRMS-ESI (*m/z*) calculated for C₁₀H₈ClF₃NO₃ [*M*+*H*]: 282.0139; found: 282.0141.

1-(4-(5-fluorobenzisoxazol-3-yl)piperidin-1-yl)prop-2-en-1-one (37)

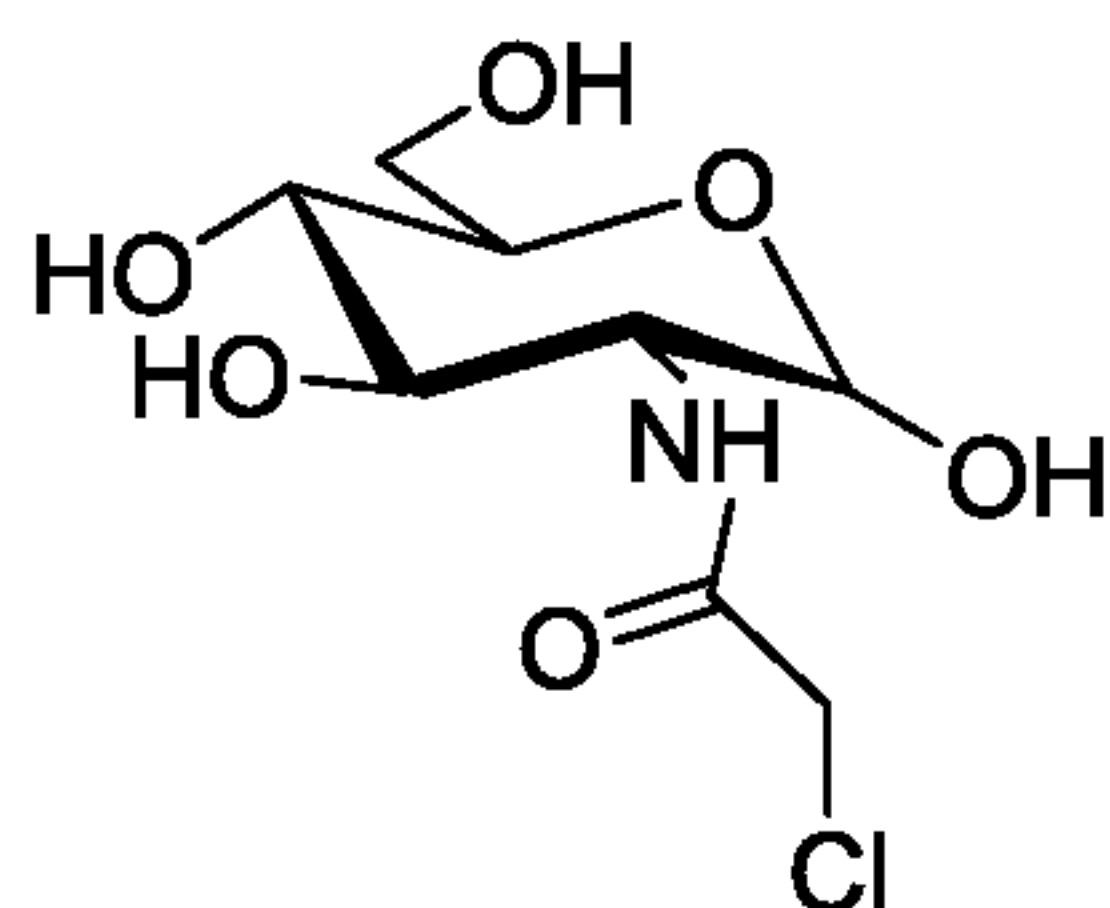
[00378] The title compound was obtained starting from 6-fluoro-3(4-piperidiny)-1,2-benzisoxazole hydrochloride (53 mg, 0.2 mmol, 1 equiv.) according to General Procedure C as a colorless oil (49.1 mg, 87%). ¹H NMR (400 MHz, Chloroform-*d*) δ 7.64 (dd, *J* = 8.7, 5.1 Hz, 1H), 7.27 (dd, *J* = 8.4, 2.3 Hz, 1H), 7.08 (td, *J* = 8.9, 2.1 Hz, 1H), 6.64 (dd, *J* = 16.8, 10.6 Hz, 1H), 6.32 (dd, *J* = 16.9, 1.9 Hz, 1H), 5.73 (dd, *J* = 10.6, 1.9 Hz, 1H), 4.70 (d, *J* = 13.4 Hz, 1H), 4.15 (d, *J* = 12.4 Hz, 1H), 3.53 – 3.13 (m, 2H), 2.99 (t, *J* = 13.1 Hz, 1H), 2.25 – 2.07 (m, 2H), 2.00 (ddd, *J* = 23.1, 14.2, 7.8 Hz, 2H); HRMS-ESI (*m/z*) calculated for C₁₅H₁₆FN₂O [M+H]: 275.119; found: 275.119.

tert-butyl 4-(4-acrylamido-2,6-difluorophenyl)piperazine-1-carboxylate (38)

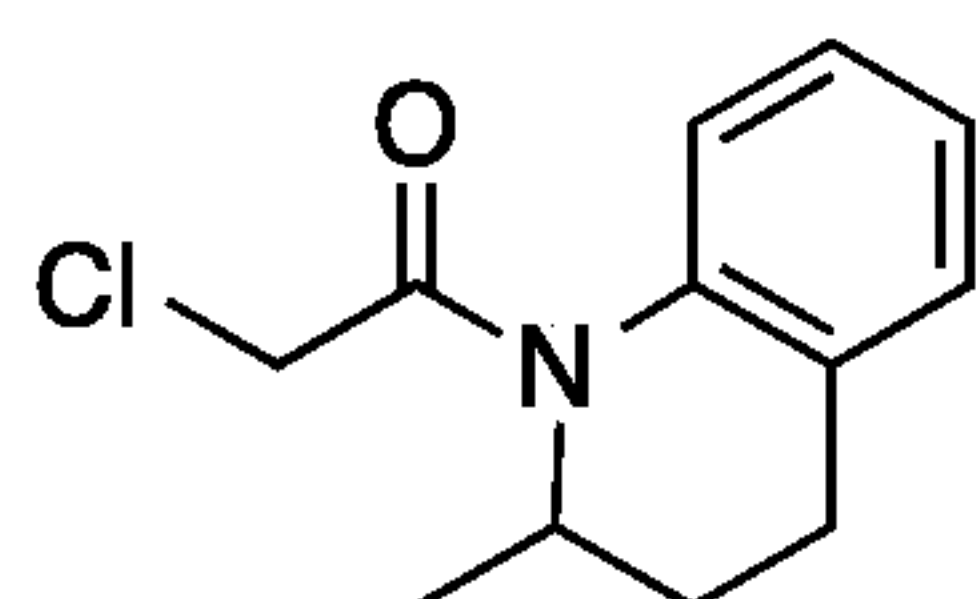
[00379] The title compound was obtained starting from tert-Butyl 4-(4-amino-2,6-difluorophenyl)piperazine-1-carboxylate according to General Procedure B. ¹H NMR (400 MHz, Chloroform-*d*) δ 8.12 (s, 1H), 7.13 (d, *J* = 10.4 Hz, 2H), 6.36 (d, *J* = 16.9 Hz, 1H), 6.19 (dd, *J* = 16.8, 10.2 Hz, 1H), 5.70 (d, *J* = 10.2 Hz, 1H), 3.45 (t, *J* = 4.7 Hz, 4H), 3.00 (t, *J* = 3.7 Hz, 4H), 1.41 (s, 9H); HRMS-ESI (*m/z*) calculated for C₁₈H₂₄F₂N₃O₃ [M+H]: 368.178; found: 368.178.

N-(4-bromo-2,5-dimethylphenyl)acrylamide (40)

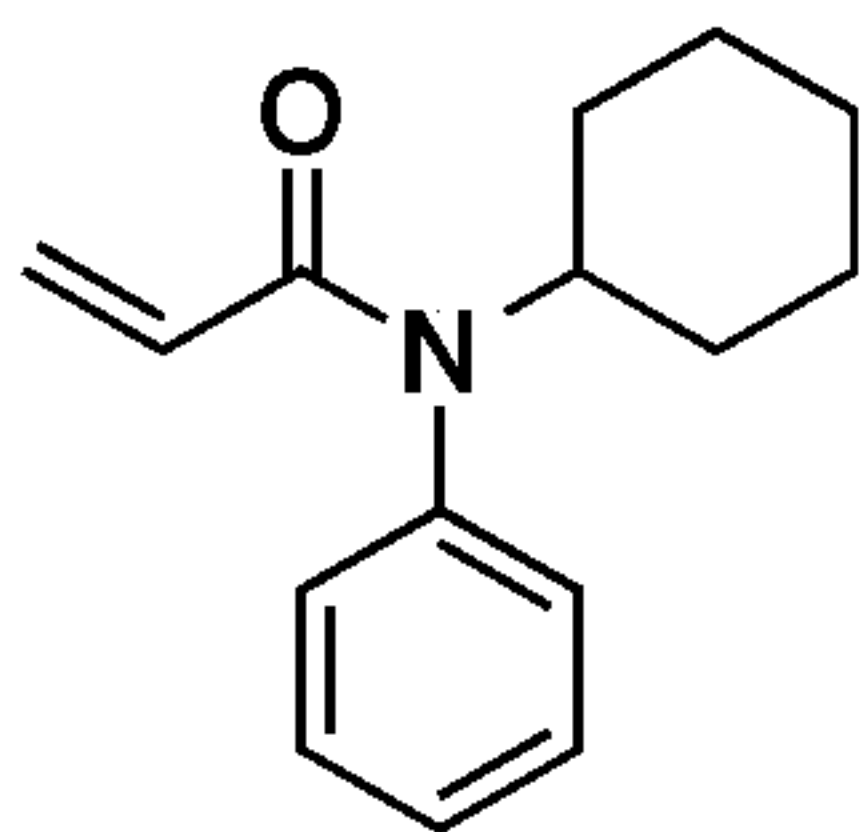
[00380] Following General Procedure B, starting from 4-bromo-2,5-dimethylaniline (900 mg, 4.5 mmol, 1 equiv.), the title compound was obtained after column chromatography and recrystallization from cold CH₂Cl₂ as a white solid (611 mg, 40%). ¹H NMR (500 MHz, Chloroform-*d*) δ 7.87 (s, 1H), 7.43 (s, 1H), 7.16 (s, 1H), 6.50 (d, *J* = 16.7 Hz, 1H), 6.35 (dd, *J* = 16.4, 10.3 Hz, 1H), 5.86 (d, *J* = 10.3 Hz, 1H), 2.42 (s, 3H), 2.28 (s, 3H); HRMS-ESI (*m/z*) calculated for C₁₁H₁₃BrNO [M+H]: 254.0175; found: 254.0175.

2-Chloroacetamido-2-deoxy- α/β -D-glucopyranose (44)

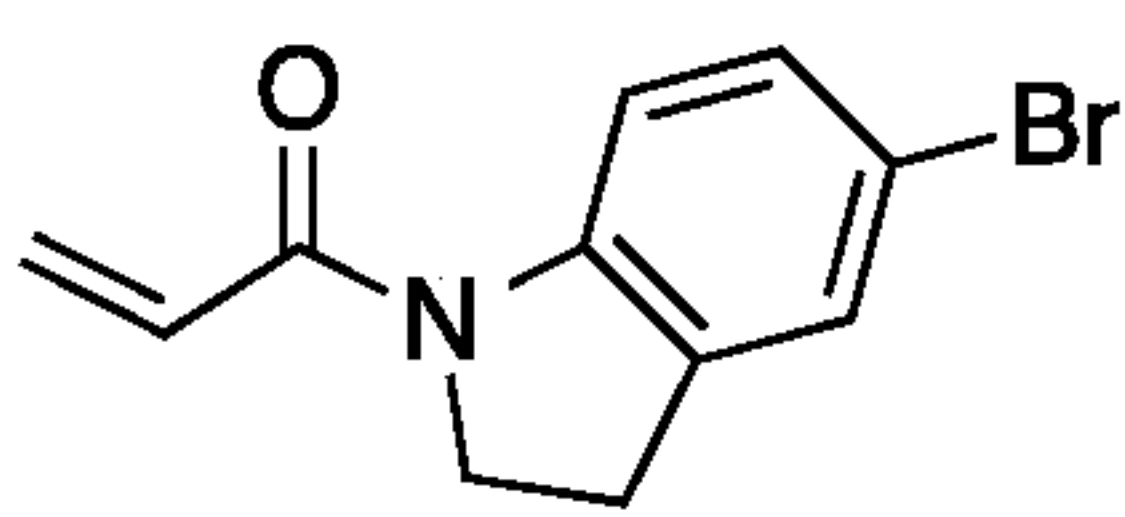
[00381] To a stirred solution of hexosamine hydrochloride (590 mg, 3.39 mmol, 1 equiv.) in anhydrous MeOH (200 mL) at room temperature was added sodium metal (60 mg, 2.6 mmol, 0.78 equiv.), TEA (400 μ L, 5.7 mmol, 1.8 equiv.). Chloroacetic anhydride (1 g, 5.9 mmol, 1 equiv.) was then added and the mixture stirred for 6 h, monitoring for completeness by TLC. After which, the reaction mixture was concentrated *in vacuo*. The crude product then was purified by two rounds of column chromatography to afford the pure title product as a white solid (610 mg, 72%). ^1H NMR (500 MHz, Methanol- d_4) δ 5.20 (d, J = 3.7 Hz, 1H α), 4.75 (d, J = 8.3 Hz, 1H β), 4.19 (dd, J = 20.2, 13.9 Hz, 2H), 4.19 (d, J = 12.6 Hz, 1H), 3.95 (dd, J = 10.6, 3.5 Hz, 1H α), 3.83 (m, 3H α , 3H β), 3.74 (d, J = 5.1 Hz, 1H β), 3.70 (dd, J = 11.4, 8.9 Hz, 1H β), 3.60 (dd, J = 10.7, 9.5 Hz, 1H β), 3.46 (t, J = 9.3 Hz, 1H), 3.42 (t, J = 10.0 Hz, 1H β); HRMS-ESI (m/z) calculated for $\text{C}_8\text{H}_{15}\text{ClNO}_6$ [$\text{M}+\text{H}$]: 256.0582; found: 256.0582.

2-chloro-1-(2-methyl-3,4-dihydroquinolin-1(2H)-yl)ethan-1-one (45)

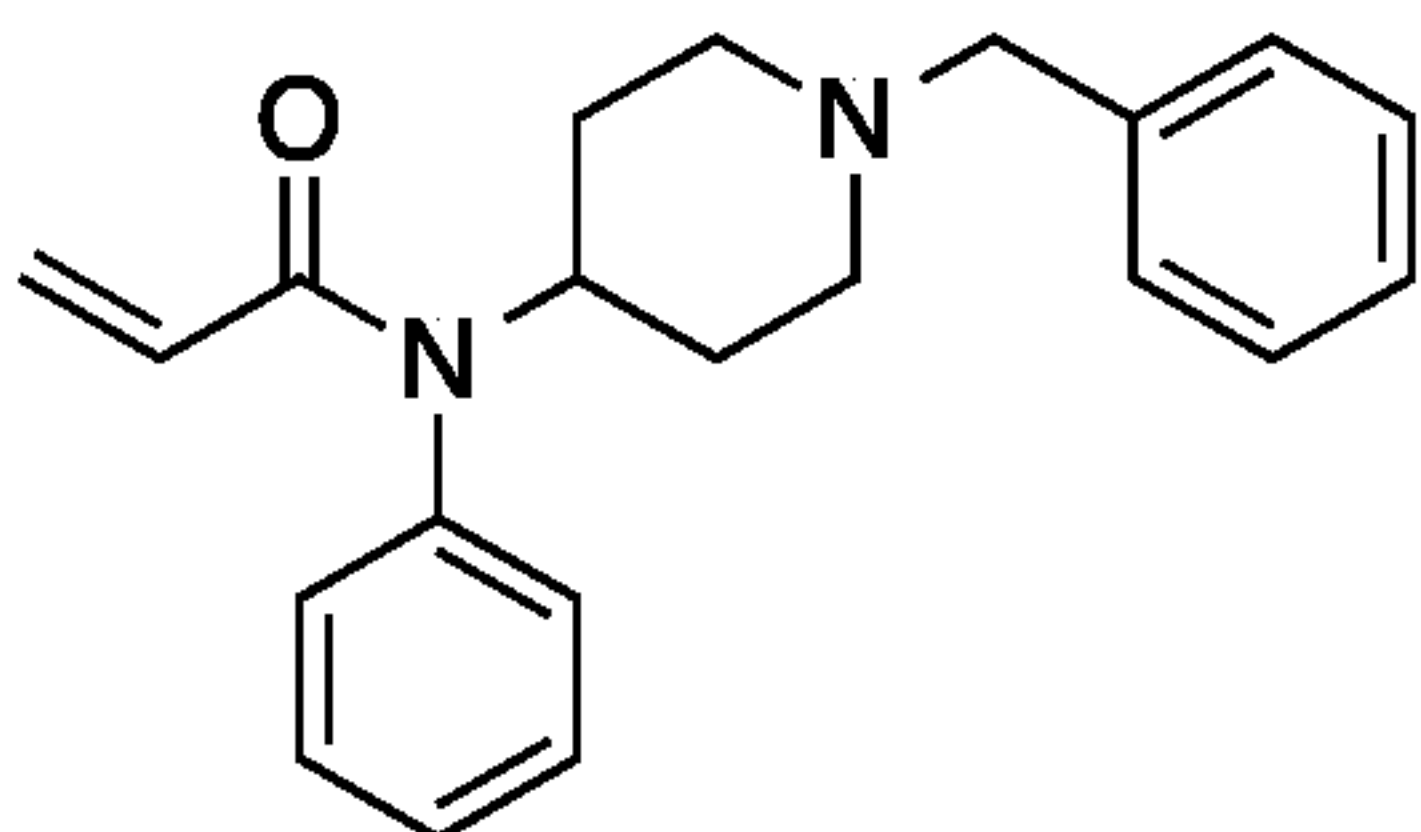
[00382] Chloroacetyl chloride (80.4 μ L, 0.9 mmol, 1.7 equiv.) was dissolved in anhydrous CH_2Cl_2 (3 mL) and cooled to 0 $^\circ\text{C}$. A solution of 2-methyl-1,2,3,4-tetrahydroquinoline (80.1 mg, 0.544 mmol, 1 equiv.) and *N*-methylmorpholine (0.11 mL, 1.0 mmol, 1.8 equiv.) in CH_2Cl_2 (2 mL) was then added dropwise. After 6 h, the reaction was quenched with saturated aqueous NaHCO_3 (5 mL) and extracted with CH_2Cl_2 (3×10 mL). The combined organic layers were dried over anhydrous Na_2SO_4 and concentrated under reduced pressure. The resultant residue was purified by prep. TLC (30% EtOAc / hexanes), providing the title compound as an off-white solid (108.8 mg, 89%). ^1H NMR (400 MHz, chloroform- d) δ 7.30 – 7.13 (m, 4H), 4.86 – 4.75 (m, 1H), 4.20 (d, J = 12.5 Hz, 1H), 4.09 (d, J = 12.5 Hz, 1H), 2.69 – 2.58 (m, 1H), 2.59 – 2.46 (m, 1H), 2.46 – 2.31 (m, 1H), 1.36 – 1.29 (m, 1H), 1.15 (d, J = 6.5 Hz, 3H); HRMS-ESI (m/z) calculated for $\text{C}_{12}\text{H}_{15}\text{ClNO}$ [$\text{M}+\text{H}$]: 224.0837; found: 224.0836.

N-cyclohexyl-*N*-phenylacrylamide (46)

[00383] The title compound was synthesized according to General Procedure C from *N*-cyclohexylaniline (89.5 mg, 0.511 mmol, 1 equiv.). Purification of the crude product by flash column chromatography (10–20% EtOAc / hexanes) then prep. TLC (30% EtOAc / hexanes) provided the title compound as an off-white solid (53.1 mg, 45%). ¹H NMR (400 MHz, chloroform-*d*) δ 7.42 – 7.33 (m, 3H), 7.10 – 7.06 (m, 2H), 6.31 (dd, *J* = 16.7, 2.1 Hz, 1H), 5.77 (dd, *J* = 16.7, 10.3 Hz, 1H), 5.41 (dd, *J* = 10.4, 2.1 Hz, 1H), 4.65 (tt, *J* = 12.2, 3.7 Hz, 1H), 1.85 (dt, *J* = 11.2, 1.8 Hz, 2H), 1.75 – 1.68 (m, 2H), 1.61 – 1.53 (m, 1H), 1.40 (qt, *J* = 13.3, 3.6 Hz, 2H), 1.07 (qd, *J* = 12.4, 3.6 Hz, 2H), 0.91 (qt, *J* = 13.1, 3.8 Hz, 1H); HRMS-ESI (*m/z*) calculated for C₁₅H₂₀NO [*M*+*H*]: 230.1539; found: 230.1539.

1-(5-bromoindolin-1-yl)prop-2-en-1-one (47)

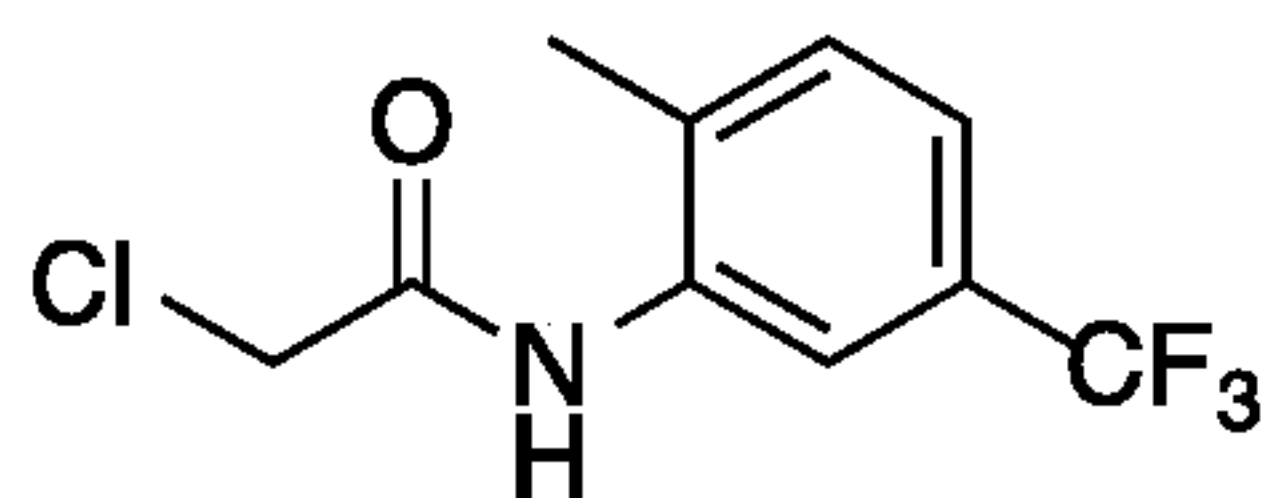
[00384] The title compound was synthesized according to General Procedure C from 5-bromoindoline (41.7 mg, 0.211 mmol, 1 equiv.), acryloyl chloride (32 μL, 0.40 mmol, 1.9 equiv.), and changing the base to pyridine (32 μL, 0.40 mmol, 1.9 equiv.). Purification of the crude product by re-precipitation from EtOAc provided the title compound as a white solid (67.8 mg, 64%). ¹H NMR (400 MHz, chloroform-*d*) δ 8.16 (d, *J* = 8.6 Hz, 1H), 7.33 – 7.25 (m, 2H), 6.60 – 6.42 (m, 2H), 5.84 – 5.76 (m, 1H), 4.15 (t, *J* = 8.6 Hz, 2H), 3.17 (t, *J* = 8.6 Hz, 2H); HRMS-ESI (*m/z*) calculated for C₁₁H₁₁BrNO [*M*+*H*]: 252.0018; found: 252.0017.

N-(1-benzylpiperidin-4-yl)-*N*-phenylacrylamide (48)

[00385] The title compound was synthesized according to General Procedure C from 1-benzyl-*N*-phenylpiperidin-4-amine (30.0 mg, 0.113 mmol, 1 equiv.), acryloyl chloride (17 μL, 0.21 mmol, 1.9 equiv.), and changing the base to pyridine (17 μL, 0.21 mmol, 1.9 equiv.). Purification of the crude product by prep. TLC provided the title compound as a white solid (22.5 mg, 64%). ¹H NMR (400 MHz, chloroform-*d*) δ 7.62 – 7.56 (m, 2H), 7.43 – 7.36 (m, 6H),

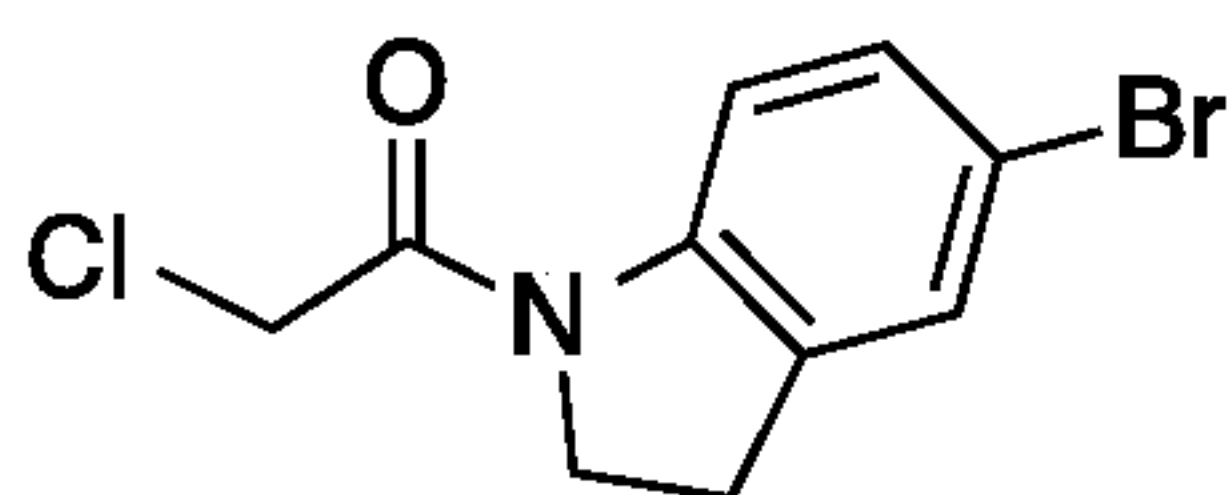
7.05 (d, $J = 6.2$ Hz, 2H), 6.29 (dd, $J = 16.8, 2.1$ Hz, 1H), 5.79 (dd, $J = 16.8, 10.3$ Hz, 1H), 5.46 (dd, $J = 10.3, 2.1$ Hz, 1H), 4.81 – 4.70 (m, 1H), 4.09 (s, 2H), 3.41 (d, $J = 12.0$ Hz, 2H), 2.82 (q, $J = 11.5$ Hz, 2H), 2.21 (q, $J = 11.9$ Hz, 2H), 1.94 (d, $J = 14.2$ Hz, 2H); HRMS-ESI (m/z) calculated for $C_{21}H_{25}N_2O$ [$M+H$]: 321.1961; found: 321.1962.

2-chloro-N-(2-methyl-5-(trifluoromethyl)phenyl)acetamide (49)



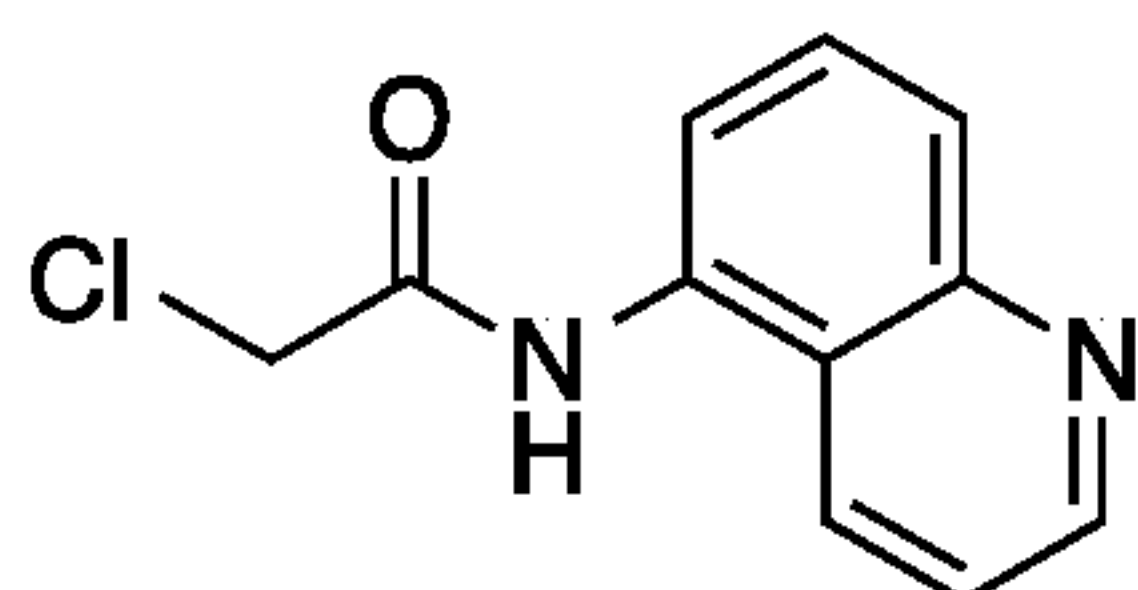
[00386] The title compound was synthesized according to General Procedure A1 from 2-methyl-5-(trifluoromethyl)aniline (35.0 mg, 0.2 mmol, 1 equiv.). Purification of the crude product by prep. TLC (20% EtOAc / hexanes) provided the title compound as a white solid (48.2 mg, 95%). ^1H NMR (600 MHz, chloroform- d) δ 8.31 (s, 1H), 8.25 (d, $J = 1.9$ Hz, 1H), 7.37 (dd, $J = 7.9, 1.8$ Hz, 1H), 7.32 (d, $J = 7.9$ Hz, 1H), 4.25 (s, 2H), 2.36 (s, 3H); HRMS-ESI calculated for $C_{10}H_{10}ClF_3NO$ [$M+H$]: 252.0397; found: 252.0397.

1-(5-bromoindolin-1-yl)-2-chloroethan-1-one (50)



[00387] The title compound was synthesized according to General Procedure A1 from 5-bromoindoline (39.6 mg, 0.2 mmol, 1 equiv.). Purification of the crude product by prep. TLC (25% EtOAc / hexanes) provided the title compound as an off-white solid (48.6 mg, 89%). ^1H NMR (600 MHz, $CDCl_3$) δ 8.07 (d, $J = 8.4$ Hz, 1H), 7.32 (d, $J = 8.8$ Hz, 2H), 4.17 (t, $J = 8.6$ Hz, 2H), 4.14 (s, 2H), 3.22 (t, $J = 8.4$ Hz, 2H); HRMS-ESI (m/z) calculated for $C_{10}H_{10}BrClNO$ [$M+H$]: 273.9629; found: 273.9629.

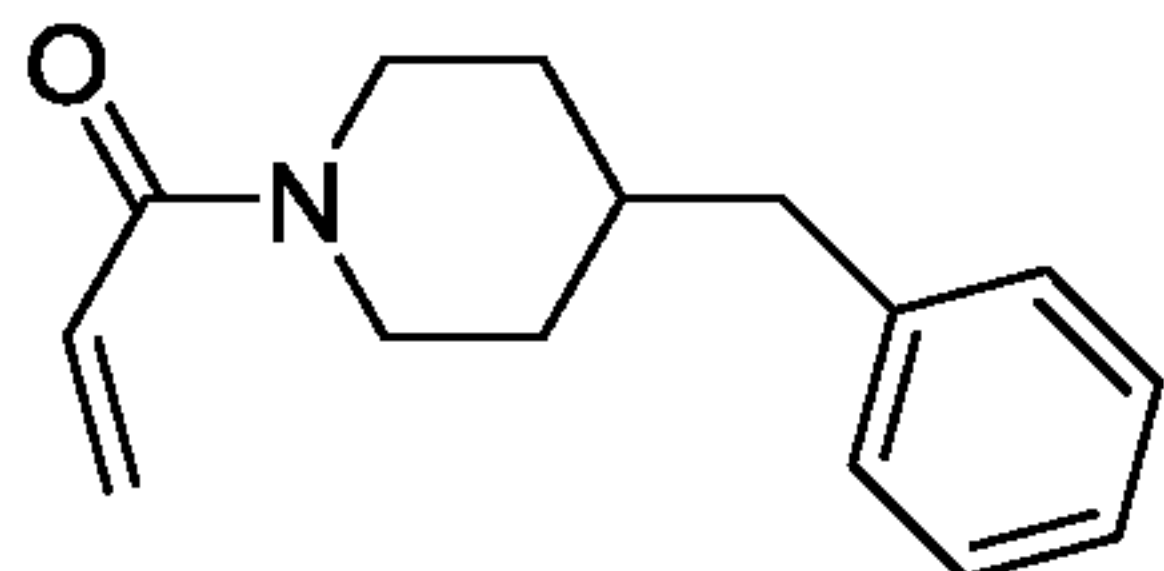
2-chloro-N-(quinolin-5-yl)acetamide (51)



[00388] To a stirring suspension of 5-aminoquinoline (28.8 mg, 0.2 mmol, 1 equiv.) and potassium carbonate (82.9 mg, 0.6 mmol, 3 equiv.) in anhydrous CH_2Cl_2 (3 mL) at 0 °C was added chloroacetyl chloride (24 μ L, 1.5 equiv.). The reaction was allowed to slowly warm up to room temperature. After 3 hours, the mixture was filtered, washed with EtOAc (10 mL) and CH_2Cl_2 (10 mL). The solid cake was then eluted with MeOH (20 mL) and the filtrate concentrated *in vacuo*. The residue was taken up in 10% MeOH / CH_2Cl_2 and passed through a pad of silica to provide the title compound as an off-white solid (42.6 mg, 82%). ^1H NMR (500

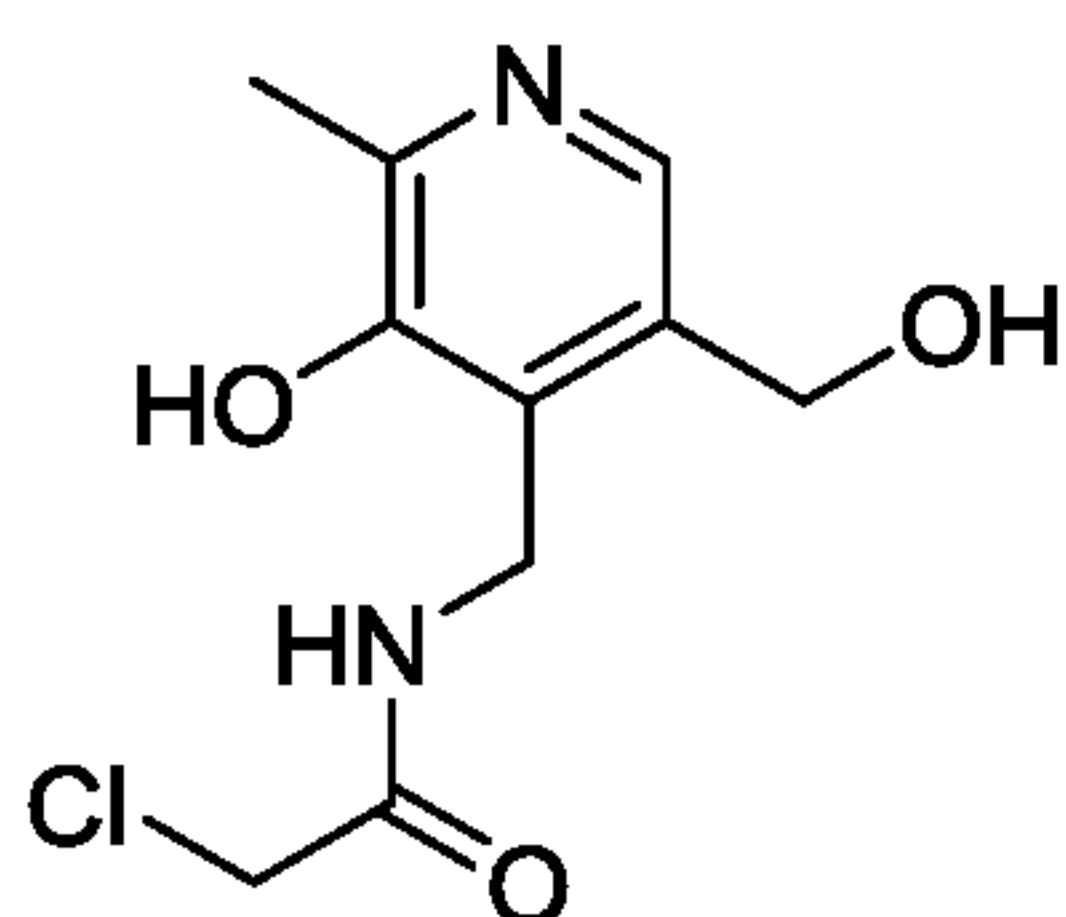
MHz, CDCl_3) δ 8.96 (d, $J = 2.5$ Hz, 1H), 8.71 (s, 1H), 8.20 (d, $J = 8.6$ Hz, 1H), 8.04 (d, $J = 8.5$ Hz, 1H), 7.94 (d, $J = 7.5$ Hz, 1H), 7.74 (t, $J = 8.0$ Hz, 1H), 7.48 (dd, $J = 8.5, 4.2$ Hz, 1H), 4.35 (s, 2H); HRMS-ESI (m/z) calculated for $\text{C}_{11}\text{H}_9\text{ClN}_2\text{O}$ [$\text{M}+\text{H}$]: 221.0476; found: 221.0477.

1-(4-benzylpiperidin-1-yl)prop-2-en-1-one (53)



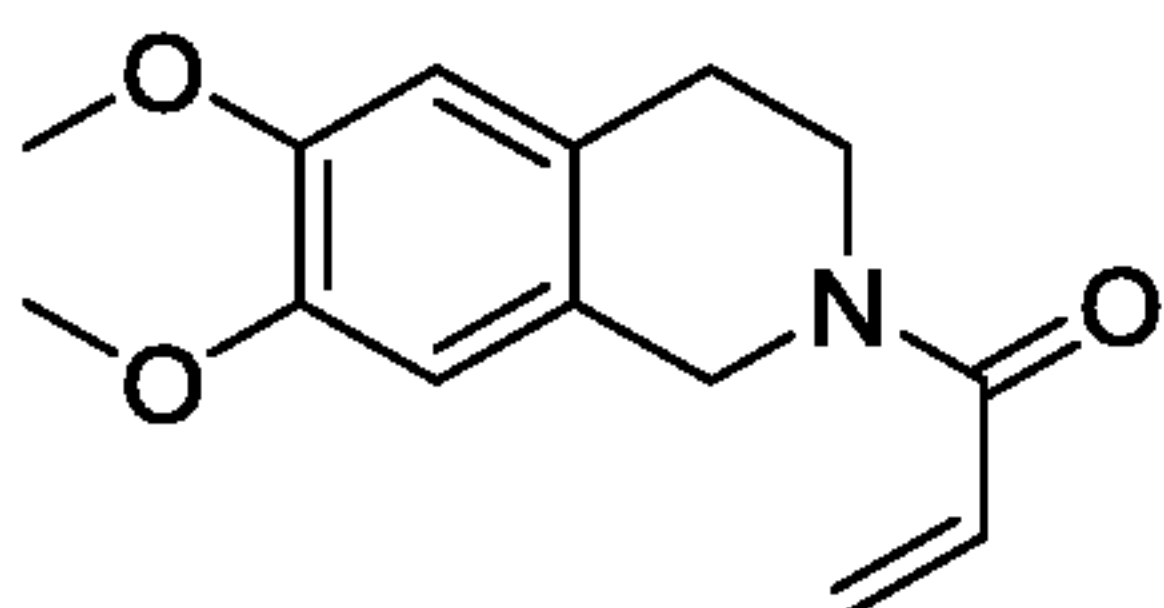
[00389] Following General Procedure B, starting from 4-benzylpiperidine (1 g, 5.7 mmol, 1 equiv.), the title compound was obtained after column chromatography as a yellow oil (748 mg, 57%). ^1H NMR (500 MHz, Chloroform- d) δ 7.36 (t, $J = 7.4$ Hz, 2H), 7.28 (t, $J = 7.4$ Hz, 1H), 7.20 (d, $J = 7.1$ Hz, 2H), 6.64 (dd, $J = 16.8, 10.6$ Hz, 1H), 6.32 (dd, $J = 16.8, 1.9$ Hz, 1H), 5.72 (dd, $J = 10.6, 1.9$ Hz, 1H), 4.72 (d, $J = 12.7$ Hz, 1H), 4.03 (d, $J = 13.0$ Hz, 1H), 3.05 (t, $J = 12.7$ Hz, 1H), 2.70 – 2.59 (m, 3H), 1.86 (ddp, $J = 14.6, 7.2, 3.5$ Hz, 1H), 1.77 (m, 2H), 1.37-1.18 (m, 2H); HRMS-ESI (m/z) calculated for $\text{C}_{15}\text{H}_{20}\text{ClNO}$ [$\text{M}+\text{H}$]: 230.1539; found: 230.1539.

2-chloro-N-((3-hydroxy-5-(hydroxymethyl)-2-methylpyridin-4-yl)methyl)acetamide (54)



[00390] To a stirred solution of pyridoxamine hydrochloride (150 mg, 0.64 mmol, 1 equiv.) in anhydrous MeOH (20 mL) at room temperature was added sodium metal (30 mg, 1.5 mmol, 2.3 equiv.), TEA (100 μL , 1 mmol, 1.6 equiv.). Chloroacetic anhydride (390 mg, 2.29 mmol, 3.5 equiv.) was added and the mixture stirred for 6 h, monitoring for completeness by TLC. After which, the reaction mixture was concentrated *in vacuo*. The crude product then was purified by prep. TLC to afford the title compound as a white solid (46 mg, 30%). ^1H NMR (500 MHz, Methanol- d_4) δ 7.97 (s, 1H), 4.81 (s, 2H), 4.61 (s, 2H), 4.17 (s, 3H), 4.06 (s, 1H), 3.35 (s, 1H), 2.52 (s, 3H); HRMS-ESI (m/z) calculated for $\text{C}_{10}\text{H}_{14}\text{ClN}_2\text{O}_3$ [$\text{M}+\text{H}$]: 245.0687; found: 245.0688.

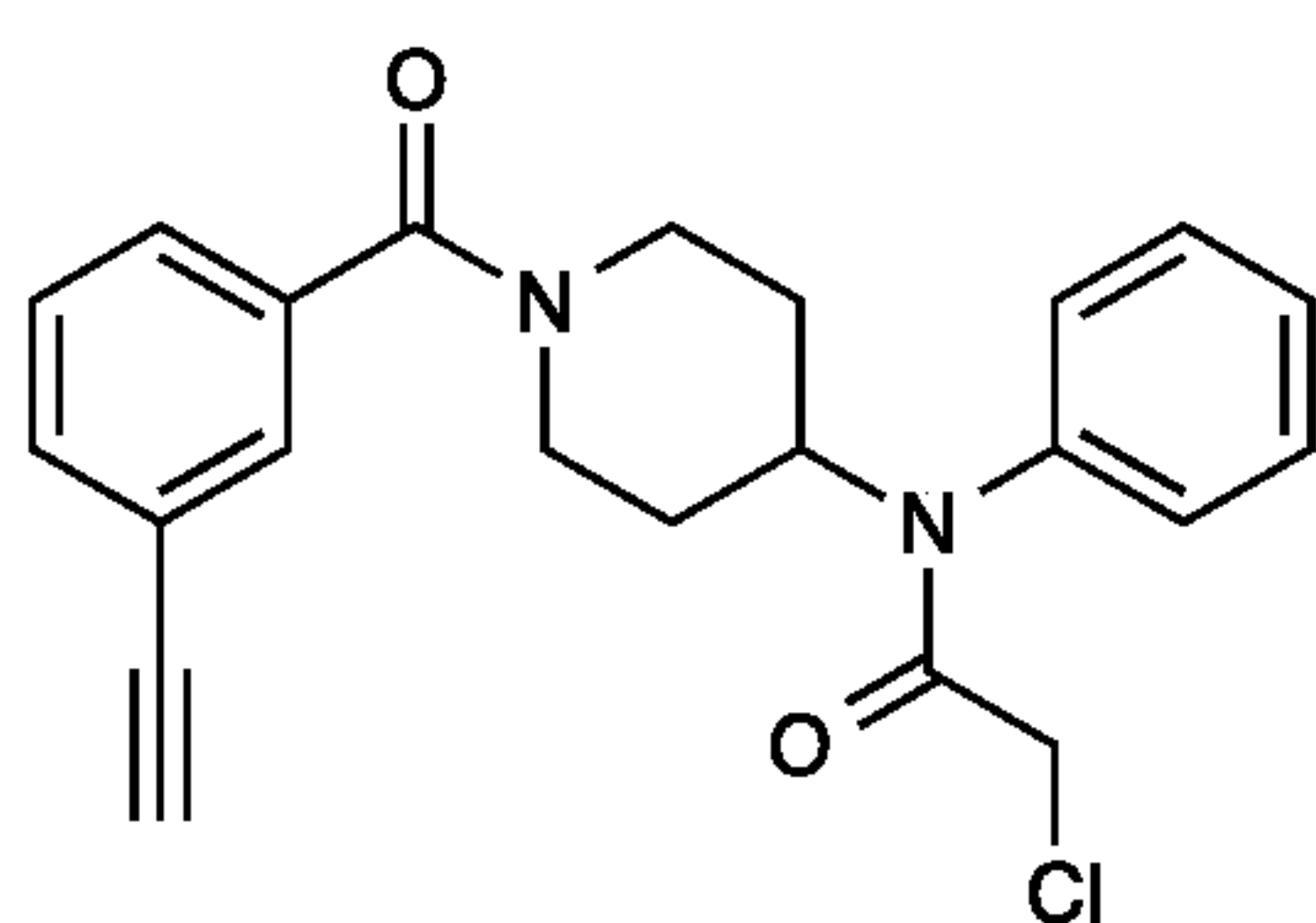
1-(6,7-dimethoxy-3,4-dihydroisoquinolin-2(1H)-yl)prop-2-en-1-one (56)



[00391] To a stirring suspension of the 6,7-dimethoxy-3,4-dihydroisoquinoline (1 g, 5.2 mmol, 1 equiv.) and TEA (1800 μL , 12.6 mmol, 2.5 equiv.) in anhydrous THF (10 mL) at

0 °C was added acryloyl chloride (1320 μ L, 13.2 mmol, 2.6 equiv.) and the reaction was allowed to slowly warm up to room temperature. After 2 hours, the mixture was diluted with CH_2Cl_2 (2×50 mL) and washed with saturated brine (2×50 mL) and the combined organics were concentrated *in vacuo*. The residue was taken up in 10% MeOH / CH_2Cl_2 and purified by column chromatography to afford the title compound as a white solid (700 mg, 54%, mixture of *E/Z* isomers). ^1H NMR (500 MHz, Chloroform-*d*) δ 6.63 (m, 3H), 6.29 (d, $J = 16.8$ Hz, 1H), 5.69 (dd, $J = 10.6, 1.8$ Hz, 1H), 4.69 (s, 1H [major]), 4.63 (s, 0.8H [minor]), 3.82 (s, 7H), 3.73 (t, $J = 5.6$ Hz, 1H), 2.84 – 2.77 (m, 2H); HRMS-ESI (m/z) calculated for $\text{C}_{14}\text{H}_{18}\text{NO}_3$ [$\text{M}+\text{H}$]: 248.128; found: 248.1281.

2-chloro-N-(1-(3-ethynylbenzoyl)piperidin-4-yl)-N-phenylacetamide (61)



[00392] To an excess of neat **SI-3** was added 0.7 mL of trifluoroacetic acid (0.2 M). The resulting solution was concentrated under a stream of nitrogen until no further evaporation was observed, providing the deprotected amine as its trifluoroacetate salt. The trifluoroacetate amine salt (90.6 mg, 0.25 mmol) was taken up in DMF (0.5 mL, 0.5 M) and the resulting solution was cooled to 0 °C. 3-ethynyl benzoic acid (44 mg, 1.2 equiv.), HATU (113 mg, 1.2 equiv.), and Hunig's base (86 μ L, 2 equiv.) were sequentially added. The reaction was stirred for 2 hours at 0 °C, diluted with Et_2O , and then washed with 1 M HCl. The organic layer was dried over magnesium sulfate, concentrated, and purified by flash chromatography (gradient from 40 to 70 % ethyl acetate in hexanes) to provide the title compound (87 mg, 92%). ^1H NMR (400 MHz, Chloroform-*d*) δ 7.51 (dd, $J = 9.5, 5.4$ Hz, 4H), 7.43 (d, $J = 1.9$ Hz, 1H), 7.39 – 7.25 (m, 2H), 7.14 (d, $J = 10.4$ Hz, 2H), 4.86 (tt, $J = 15.1, 5.3$ Hz, 2H), 3.72 (s, 3H), 3.19 (d, $J = 14.0$ Hz, 1H), 3.11 (s, 1H), 2.86 (s, 1H), 1.90 (d, $J = 36.6$ Hz, 2H), 1.38 (s, 1H), 1.24 (d, $J = 19.9$ Hz, 1H); HRMS-ESI (m/z) calculated for $\text{C}_{22}\text{H}_{22}\text{ClN}_2\text{O}_2$ [$\text{M}+\text{H}$]: 381.1364; found: 381.1363.

Global profiling of cysteine-reactive fragments in native populations

[00393] Cysteine is unique among protein-coding amino acids owing to its high nucleophilicity and sensitivity to oxidative modification. Cysteine residues perform catalytic functions in diverse enzyme classes and represent sites for post-translational regulation of proteins through disulfide bonding, iron-sulfur cluster formation, conversion to sulfinic and sulfonic acid, nitrosylation, S-glutathionylation and lipid modification. Using a quantitative chemical proteomic method termed isoTOP-ABPP (isotopic Tandem Orthogonal Proteolysis–

Activity-Based Protein Profiling), global measurements of the intrinsic reactivity of cysteine residues was carried out and their sensitivity to modification by lipid-derived electrophiles was assessed. In order to determine whether isoTOP-ABPP was adapted to perform covalent FBLD in native biological systems, a cell preparation (lysate or intact cells) was pre-treated with DMSO or one member of a library of electrophilic small-molecule fragments and then exposed to a broad-spectrum cysteine-reactive probe iodoacetamide (IA)-alkyne **1** (**Fig. 1A**). Proteins harboring IA-alkyne-labeled cysteine residues from DMSO- and fragment-treated samples were conjugated by copper-mediated azide-alkyne cycloaddition (CuAAC or click) chemistry to isotopically differentiated azide-biotin tags (heavy and light, respectively), combined, enriched by streptavidin, and proteolytically digested on-bead to yield isotopic peptide pairs that were analyzed by LC-MS. Quantification of MS1 chromatographic peak ratios for peptide pairs identified fragment-competed Cys residues as those displaying high competition ratios, or *R* values, in DMSO/fragment comparisons.

[00394] A 50+ member fragment library was constructed with most compounds containing either a chloroacetamide or acrylamide electrophile (**Fig. 1B** and **Fig. 3**), which are well-characterized cysteine-reactive groups found in many chemical probes and some clinically approved drugs. These electrophiles were appended to structurally diverse small-molecule fragments (< 300 Da) intended to serve as recognition elements that promote interactions with different subsets of the human proteome. The library also contained some additional electrophiles, such as cyanoacrylamides and vinylsulfonamides, and known bioactive electrophilic compounds (e.g., the anti-cancer agent piperlongumine and anti-migratory agent locostatin) (**Fig. 1B**, and **Fig. 3**). The electrophile library was screened at a high concentration (500 μ M) comparable to the ligand concentrations used in typical FBLD experiments. A subset of the fragment library was initially assayed by competitive profiling in a human MDA-MB-231 breast cancer cell line proteome using an IA-rhodamine probe **16**, which permitted facile SDS-PAGE detection of cysteine reactivity events. This experiment identified several proteins that showed reductions in IA-rhodamine labeling in the presence of one or more fragments (**Fig. 1C**, asterisks). Interestingly, the proteins exhibited distinct SARs across the test fragment set, indicating that the library recognition elements exert a strong influence over specific fragment-protein reactivity events.

[00395] Competitive isoTOP-ABPP was used to globally map human proteins and the cysteine residues within these proteins that were targeted by fragment electrophiles. Each fragment was tested, in general, against two distinct human cancer cell proteomes (MDA-MB-231 and Ramos cells) and most fragments were screened in duplicate against at least one of these proteomes. On average, 927 cysteines were quantified per data set, and it was required that

individual cysteines were quantified in at least three data sets for interpretation. Based on these criteria, more than 6157 cysteines from 2885 proteins were quantified in aggregate across all data sets with an average quantification frequency of 22 data sets per cysteine (**Fig. 4A**).

Fragment-competed cysteine residues, or “liganded” cysteines, were defined as those showing $\geq 75\%$ reductions in IA-alkyne labeling (R values ≥ 4 for DMSO/fragment). To minimize the potential for false-positives, only cysteines that showed R values ≥ 4 in two or more data sets and met additional criteria for data quality control were considered as targets of the fragment electrophiles. The proteomic reactivity values, or liganded cysteine rates, of individual fragments were then calculated as the percentage of liganded/total quantified cysteines in isoTOP-ABPP experiments performed on that fragment.

[00396] Most fragment electrophiles showed a tempered reactivity across the human proteome, with a median liganded cysteine rate of 3.8% for the library (**Fig. 4B**). Substantial differences in reactivity were, however, observed, with individual electrophiles showing liganded cysteine rates of $< 0.1\%$ and others displaying rates $> 15\%$ (**Fig. 4B**). That piperlongumine and locostatin fell into the latter category indicated the intrinsic proteomic reactivity of the fragment electrophiles did not, in general, exceed that of previously described electrophilic probes. A subset of fragments was also screened at lower concentrations (25-50 μM), which confirmed that their proteomic reactivities were concentration-dependent (**Fig. 4C**). The relative reactivity of fragment electrophiles was similar in MDA-MB-231 and Ramos cell proteomes (**Fig. 4D**), indicating that this parameter is an intrinsic property of the compounds. Fragments also showed consistent reactivity profiles when assayed in biological replicate experiments (**Fig. 4E**). Interestingly, it was found that the proteomic reactivity of fragment electrophiles was only marginally correlated with their glutathione adduction potential, which is a commonly used surrogate assay for measurements of proteinaceous cysteine reactivity (**Fig. 4F**). These differences are attributed to the impact of the recognition element of fragment electrophiles on their interactions and, ultimately, reactivity with proteins.

[00397] A comparison of fragments **3**, **14**, **17**, and **23-26** provided insights into the relative proteomic reactivity of different electrophilic groups coupled to a common recognition element (3,5-di(trifluoromethyl)phenyl group). Chloroacetamide **3** exhibited greater reactivity than acrylamide **14** (15% versus 3.4% liganded cysteines, respectively; **Fig. 1D**), with cyanoacrylamide **23** exhibiting similar reactivity to acrylamide **14** and other, more sterically congested electrophiles (**24-26**) showing reduced proteomic reactivity (**Fig. 4G**). Importantly, the non-electrophilic acetamide control fragment **17** showed negligible activity in competitive isoTOP-ABPP experiments (liganded cysteine rate $< 0.2\%$) (**Fig. 1D**), indicating that the vast majority of detected fragment-cysteine interactions reflected covalent reactions versus non-

covalent binding events. Also in support of this conclusion, “clickable” alkyne analogues of **3** and **14** (compounds **19** and **18**, respectively) exhibited different concentration-dependent proteome labeling profiles (**19** > **18**; **Fig. 1E**) that mirrored the respective liganded cysteine rates displayed by **3** and **14** in competitive isoTOP-ABPP experiments (**3** > **14**; **Fig. 1D**). Despite the greater overall proteomic reactivity of chloroacetamide **3** relative to acrylamide **14** and cyanoacrylamide **23**, clear examples of cysteines were found that were preferentially liganded by the latter fragments (**Fig. 1F**).

[00398] In some instances, these findings demonstrate that the isoTOP-ABPP platform is one method for use to competitively profile fragment electrophiles against thousands of cysteine residues in native proteomes.

Cysteines targeted by fragment electrophiles in native proteomes

[00399] Across all isoTOP-ABPP data sets combined, 758 liganded cysteines were identified on 637 distinct proteins, which corresponded to ~12 and 22% of the total quantified cysteines and proteins, respectively (**Fig. 5A** and **Tables 1-3**). Only a modest fraction of the proteins harboring liganded cysteines were found in the DrugBank database (15%; **Fig. 5B**), indicating the fragment electrophiles targeted many proteins that lack small-molecule probes. Among protein targets with known covalent ligands, the fragment electrophiles frequently targeted the same cysteine residues as these known ligands (Table 4); examples include the protein kinase BTK, in which electrophilic fragments targeted an active-site cysteine that also reacts with the cancer drug ibrutinib, and XPO1 and ERCC3, in which electrophilic fragments targeted conserved cysteines that are modified by bioactive natural products and candidate anti-cancer agents. In the case of BTK, it was confirmed that the interaction of ibrutinib with this kinase was detected by isoTOP-ABPP, which also identified a known ibrutinib off-target – MAP2K7 – in Ramos cell lysates (**Fig. 7A**).

[00400] DrugBank proteins with liganded cysteines mostly originated from classes that are regarded as “druggable”, including enzymes, channels, and transporters (**Fig. 5C**). Non-DrugBank proteins with liganded cysteines, on the other hand, showed a broader class distribution that included proteins, such as transcription factors and adaptor/scaffolding proteins, that are considered challenging to target with small-molecule ligands (**Fig. 5C**). Even among the enzymes targeted by fragment electrophiles, many examples were noted where the liganded cysteine was a non-active site residue (**Fig. 7B**). These data indicated that the cysteines modified by fragment electrophiles were not restricted to classical ligand-binding pockets on proteins. Also consistent with this premise, only ~6% of all of the liganded cysteines were functionally annotated as active-site residues (**Fig. 5D**). Active-site cysteines, as well as redox-active cysteines, were still, however, substantially enriched among the liganded cysteine group

compared to unliganded cysteines quantified by isoTOP-ABPP (**Fig. 5D**). It had been previously found that active-site and redox-active cysteines also show, in general, greater intrinsic reactivity (as measured with the IA-alkyne probe) compared to other cysteines. While this heightened reactivity is a likely contributory factor to the ligandability of cysteines, as reflected in the high proportion of hyperreactive cysteines that were detected as targets of fragment electrophiles (**Fig. 5E**), liganded cysteines were also well-represented across a broad range of intrinsic reactivities (**Fig. 5E**). Finally, most proteins were found to harbor a single liganded cysteine among the several cysteines that were, on average, quantified per protein by isoTOP-ABPP (**Fig. 5F**). The nuclear export factor XPO1 and metabolic enzyme PHGDH provide compelling examples of the selectivity displayed by fragment electrophiles for individual cysteines within proteins (**Fig. 5G** and **Fig. 7C**). Among the six different XPO1 cysteine residues quantified by isoTOP-ABPP, a single cysteine, C528, was frequently targeted by fragment electrophiles (**Fig. 5G**), and this residue is also modified by electrophilic drugs in clinical development for cancer⁴⁰. Similarly, among eight quantified cysteines in PHGDH, only C369, a non-active site residue, was targeted by electrophilic fragments (**Fig. 7C**).

[00401] Liganded cysteines displayed strikingly distinct SARs with the fragment electrophile library (**Fig. 6A** and **Tables 1-3**). While a handful of cysteines were targeted by a large number of fragments (> 50%), most cysteines exhibited more restricted reactivity (**Fig. 6A, B** and **Tables 1-3**). The operational grouping of fragment electrophiles based on their relative proteomic reactivity values (group A, > 10%; group B, < 10%) revealed SAR features that emphasized both the recognition and reactivity components of cysteine-electrophile interactions. Certain cysteines, for instance, preferentially interacted with the less reactive (group B) fragments (e.g., GLRX5; MSTO1; SRP9; UCHL3; **Fig. 6A**), while others were mainly liganded by the most reactive (group A) fragments (e.g., ATXN7L3B; CRKL; C2ORF49; **Fig. 6A**), although, even in these cases, the interactions differed substantially across group A fragments. Liganded cysteines located in the active sites of proteins tended to show broader reactivity with the fragment electrophiles compared to other cysteines (**Fig. 6C**), possibly reflecting their greater ligandability, but clear SARs were observed for many non-active site cysteines and these residues were not disproportionately targeted by group A fragments (**Fig. 6D**). These principles applied across different protein classes and were well-exemplified in kinases, for which > 20 liganded cysteines were identified that distributed near-evenly between active- and non-active-site residues (**Fig. 7D-F**). Even cysteines found in proteins considered challenging to drug, such as transcription factors/regulators, showed distinct SARs indicative of specific interactions involving both binding and reactivity (**Fig. 6D** and **Fig. 9G**). In addition, about greater than 60% of liganded cysteines, electrophile (IA-alkyne or fragment) reactivity was blocked by heat

denaturation of the proteome, while about a fraction of unliganded cysteines (about 20%) showed decreased IA-alkyne labeling following heat denaturation (**Figs. 15 and 16**). In some instances, these results showed that the ligand-cysteine interactions are specific in that they depend on both the binding groups of ligands and structured sites in protein.

[00402] The availability of three-dimensional structures for a subset of proteins with liganded cysteines provided an opportunity to test whether docking predicts sites of fragment electrophile reactivity. Covalent docking programs have recently been introduced to discover ligands that target pre-specified cysteines in proteins; here, however, the aim was to computationally assess the relative ligandability of all cysteines within a protein and match these outputs to the data acquired in isoTOP-ABPP experiments. First, 29 representative protein targets were scanned and 99 solvent-accessible cysteines were identified. Then, the fragment electrophile library was docked on each residue independently using a modified potential to simulate non-covalent interactions preceding the alkylation event. In cases where the fragment electrophile bound favorably near a cysteine and the reactive group was within covalent bond distance of the cysteine, the cysteine was considered to be modified by the fragment. Docking scores were then calculated based on the estimated interaction energy of each fragment in its docked pose, and the ranking of these predictions matched the experimental data in 19 out of the 29 systems (i.e., cases where the top predicted ligandable cysteine matched the liganded cysteine determined by isoTOP-ABPP) (**Fig. 6E, F** and Table 5). In six out of the remaining 10 systems, the liganded cysteines were ranked second by reactive docking. In the remaining four systems, reactive docking failed to predict the liganded cysteine due to limitations in the docking scoring function or structural issues in the models used. Notably, across the entire 29 proteins evaluated by reactive docking, it was found that cysteines predicted to be ligandable were much more likely to have been detected by isoTOP-ABPP compared to cysteines not predicted to be ligandable (**Fig. 6E** and **Fig. 7H**). It was also found that cysteines predicted to be ligandable were more likely to have been detected by isoTOP-ABPP and exhibited heat-sensitive IA-alkyne reactivity (**Fig. 17A** and **Fig. 17B**). These results indicate that reactive docking provides a good overall prediction of the ligandability of proteinaceous cysteines and suggest that IA-alkyne reactivity itself provides an independent experimental parameter useful for designating potentially ligandable cysteines in proteins.

Functional analysis of ligand-cysteines interactions

[00403] The next step was to confirm and determine the functional impact of ligand-cysteine interactions mapped by isoTOP-ABPP using recombinant proteins. Two proteins were selected for which the functional significance of the liganded cysteines had been previously demonstrated. The protein methyltransferase PRMT1 possesses a non-catalytic active-site

cysteine (C109) that, when modified by electrophilic small molecules like 4-hydroxynonenal (HNE), results in the inhibition of PRMT1 activity²⁷. Competitive isoTOP-ABPP revealed a very selective SAR for ligand engagement of C109 of PRMT1, with only three fragments (**2**, **11**, and **51**) blocking IA-alkyne labeling of this residue (**Fig. 6A** and **Fig. 8A** and **Tables 1-3**). Even though several additional cysteines in PRMT1 were quantified in isoTOP-ABPP experiments (none of which showed sensitivity to the tested fragment electrophiles; **Fig. 8A** and **Tables 1-3**), it was found that IA-rhodamine labeling of recombinant PRMT1 was blocked by mutating C109 to serine (**Fig. 8B**). These data are consistent with past studies indicating that C109 is the most reactive cysteine in PRMT1 and is selectively labeled by low concentrations of electrophilic probes. Using a convenient SDS-PAGE readout, it was confirmed that fragment **11** blocked IA-rhodamine labeling of PRMT1 with an IC₅₀ value of 36 μ M, whereas control fragment **3** was inactive (**Fig. 8B, C**), despite displaying similar overall proteome reactivity to **11** (**Fig. 4B**). Pre-treatment with **11**, but not **3**, also inhibited PRMT1-catalyzed methylation of histone 4 in a C109-dependent manner (**Fig. 8D**). These data indicate that electrophilic ligands targeting C109 act as PRMT1 inhibitors.

[00404] MLTK, or ZAK, which is a MAP3 kinase that possesses an active site-proximal cysteine residue C22 that is modified by HNE to feedback-inhibit JNK pathways under conditions of oxidative stress, was then examined. MLTK has also recently been implicated as an oncogenic driver in gastric cancer and is an off-target for ibrutinib, which reacts with C22 of MLTK. Competitive isoTOP-ABPP experiments identified a subset of fragment electrophiles that blocked IA-alkyne labeling of C22 in MLTK (**Fig. 9A** and **Tables 1-3**). The SAR provided by isoTOP-ABPP was verified and extended by testing fragments for blockade of labeling of recombinant MLTK using an ibrutinib-derived activity probe (**Fig. 8E** and **Fig. 9B**), which identified the benzofuran fragment **60** as having good potency for inhibiting MLTK (IC₅₀ value of 2.6 μ M) and **3** as an inactive control probe (**Fig. 8E, F** and **Fig. 9A, B**). Fragment **60**, but not **3**, also blocked the catalytic activity of MLTK using a substrate phosphorylation assay, and this inhibitory effect was not observed with a C22A-MLTK mutant (**Fig. 8G** and **Fig. 18**).

[00405] Next, proteins were evaluated that possessed previously uncharacterized liganded cysteines. IMPDH2, which is the rate-limiting enzyme in *de novo* synthesis of guanine nucleotides and regulates immune cell proliferation and cancer, contained two liganded cysteines – C140 and C331 – that showed overlapping, but distinct SARs in competitive isoTOP-ABPP experiments (**Fig. 9C, D**; **Fig. 19** and **Tables 1-3**). C331 serves as a catalytic nucleophile and active site-directed inhibitors of IMPDH2 have been described. C140, on the other hand, is found in a separate Bateman domain of IMPDH2, which serves as a module for allosteric regulation by sensing nucleotides (**Fig. 9D**) and has not been shown to react with

electrophilic small molecules. Therefore focused was placed on the characterization of C140. It was first confirmed that fragment **14** directly labeled C140 of recombinant IMPDH2 by MS methods (Table 6). An alkyne analogue of **14** (**18**; **Fig. 8H**) was then synthesized, which provided a means to directly monitor ligand interactions at C140 by click chemistry conjugation to a rhodamine-azide tag and SDS-PAGE analysis. Click probe **18** labeled WT-IMPDH2 and a C331S-IMPDH2 mutant, but not the C140S or C140S/C331S mutants of this enzyme (**Fig. 8H**). Using this assay, it was confirmed that **14**, but not control fragment **8**, inhibited the labeling of IMPDH2 by **18** (**Fig. 9E**). IMPDH2 labeling by **18** was also inhibited by nucleotides ATP, AMP, and GTP, but not UTP or IMP (**Fig. 8I** and **Fig. 9F**). ATP blocked **18** labeling of IMPDH2 with an IC_{50} value of 45 μ M (**Fig. 8J**). Thus, covalent ligands targeting the Bateman domain of IMPDH2 serves not only as inhibitors, but also probes of nucleotide binding to this enzyme.

[00406] Two liganded cysteines – C114 and C161 – were also identified in the p53-induced phosphatase TIGAR (**Fig. 9G, H**). In some instances, TIGAR acts as both a fructose-2,6-bisphosphatase and 2,3-bisphosphoglycerate phosphatase to shape the metabolic state of cancer cells and protect them from ROS-induced apoptosis. Inhibitors of TIGAR have not been described. C114 is found on the lid of the TIGAR active site, ~ 15 Å from the phosphate substrate binding site (**Fig. 9H**). C161 resides on the opposite side of the protein. Focus was placed on the characterization of fragment labeling of C114 given its proximity to the TIGAR active site. It was first confirmed that both C114 and C161 of recombinant TIGAR were labeled by the IA-rhodamine probe and this labeling was partly diminished in C114S and C161S single mutants and fully blocked in a C114S/C161S double mutant of TIGAR (**Fig. 9I**). It was also verified interactions of hit fragment **5** with C114 of TIGAR by LC-MS analysis (Table 6) and by showing that the fragment blocked IA-rhodamine labeling of a C161S-TIGAR mutant with an IC_{50} value of 16 μ M (**Fig. 8K, L**); in contrast, the control fragment **3** showed much lower potency (**Fig. 8K, L**). **5** also blocked the catalytic activity of WT- and C161S-, but not C114S- or C114S/C161S-TIGAR using a substrate assay (**Fig. 8M**). Control fragment **3** did not affect TIGAR catalytic activity (**Fig. 8L**). Inhibition of TIGAR substrate turnover by **5** plateaued at 70% (**Fig. 9J**), which indicates that the covalent ligand acts by an allosteric mechanism or does not extend fully into the active site of TIGAR to produce complete inhibition.

Electrophilic ligands that inhibit IDH1 activity in cancer cells

[00407] Isocitrate dehydrogenase 1 (IDH1) and 2 (IDH2) are mutated in a number of human cancers to produce enzyme variants with a neomorphic catalytic activity that converts isocitrate to 2-hydroxyglutarate (2-HG). Increases in 2-HG inhibit α -ketoglutarate-dependent dioxygenases that function as tumor suppressors, in particular, by methylating DNA and proteins. Competitive isoTOP-ABPP experiments identified distinct subsets of ligands that

targeted a conserved cysteine in IDH1 and IDH2 (C269 and C308, respectively; **Tables 1-3**). This cysteine is an active site-proximal residue that is 13 Å from the NADP⁺ molecule in a crystal structure of IDH1 (**Fig. 10A**); glutathionylation of C308 has previously been shown to block IDH2 activity, but, to our knowledge, irreversible inhibitors of IDH enzymes have not been characterized.

[00408] The functional significance of ligand interactions with IDH enzymes by recombinantly expressing wild type (WT) and a C269S mutant of IDH1 was explored. WT-, but not C269S-IDH1 reacted with the IA-rhodamine probe as detected by SDS-PAGE, and fragment electrophiles blocked this reaction with an SAR that mirrored that observed for endogenous IDH1 in competitive isoTOP-ABPP experiments (**Fig. 11A** and **Tables 1-3**). Fragment **20** inhibited IA-rhodamine labeling of WT-IDH1 with an IC₅₀ value of 2.9 μM (**Fig. 11B** and **Fig. 10B**) and showed similar activity with the R132H oncogenic mutant of IDH1 (**Fig. 10C** and **Fig. 20**). It was also confirmed by isoTOP-ABPP that **20** (25 μM) completely blocked IA-alkyne labeling of endogenous IDH1 in MDA-MD-231 proteomes (R value = 20; **Fig. 10D**) and, by MS analysis, that **20** directly modifies C269 of IDH1 (Table 6). Fragment **2** showed much less activity against C269 of IDH1 (IC₅₀ > 50 μM; **Fig. 11B** and **Fig. 10B**) and was therefore selected as a control probe. It was found that **20** blocked in a concentration-dependent manner the catalytic activity of WT-IDH1 (as measured by the reduction of NADP⁺ to NADPH in the presence of isocitrate), but did not inhibit the activity of the C269S-IDH1 mutant (**Fig. 11C**). The *in situ* activity of **20** was also tested by generating a human cancer cell line that stably overexpressed R132H-IDH1 (**Fig. 10E**). The R132H-IDH1 cells were treated with fragments **20** and **2** for 2 h, lysed, and assayed *ex situ* for 2-HG production. **20** (50 μM) near-completely blocked 2-HG production by R132H cell lysates, while **2** (50 μM) only caused a slight decrease in this activity (**Fig. 11D**). Parallel competitive isoTOP-ABPP experiments confirmed that fragment **20**, but not fragment **2** inhibited IA-alkyne labeling of C269 of IDH1 *in situ* (**Fig. 10F**).

Global profiling of cysteine-reactive fragments in cells

[00409] Encouraged by the cellular activity of the IDH1 ligand **20**, the capacity of fragment electrophiles to modify proteinaceous cysteines *in situ* was more broadly assessed. MDA-MB-231 and Ramos cells were treated with representative members of the fragment library (23 compounds tested in total; each compound tested at 200 μM, 2 h *in situ* treatment), and the cells were then harvested, lysed, and analyzed by isoTOP-ABPP. A handful of fragments were cytotoxic to cells and re-tested at lower (50 or 100 μM) concentrations. The tested fragments showed a broad range of *in situ* reactivities that generally matched their respective reactivities *in vitro* (**Fig. 11E** and **Tables 1-3**). Some fragments, however, showed somewhat greater reactivity

in cells, while fragment **11** was notably devoid of activity *in situ* (**Fig. 11E**). These differences reflect the impact of transport and/or metabolic pathways on the cellular concentrations of fragment electrophiles. A substantial fraction (64%) of the liganded cysteines identified in cell lysates were also sensitive to the same electrophilic fragments in cells (**Fig. 11F**). A handful of fragment-cysteine interactions were also observed selectively *in situ*, but not in lysates, including C182 of p53 (TP53), a redox-regulated residue at the dimerization interface of the DNA binding domain⁵⁰ (**Fig. 11G**). In some instances, these liganded cysteines require an intact cellular environment to preserve their interactions with fragment electrophiles. Taken together, these findings indicate that the ligandability of cysteine residues is generally similar in lysates and cells, although exceptional cases underscore the importance of having the capability to perform ligand discovery experiments *in situ*.

Electrophilic ligands that target pro-caspase-8 and block extrinsic apoptosis

[00410] Several fragments targeted the catalytic cysteine nucleophile C360 of the protease caspase-8 (CASP8) in isoTOP-ABPP experiments performed *in vitro* and *in situ* (**Fig. 12A** and **Tables 1-3**). CASP8 plays important roles apoptosis, immune cell proliferation, and embryonic development, but selective, non-peptidic, and cell-active inhibitors for this protease are lacking. Representative fragment hits against recombinant, active CASP8 were screened using substrate and activity-based probe (Rho-DEVD-AOMK probe ("DEVD" disclosed as SEQ ID NO: 857)) assays and observed marginal to no inhibition with most fragments (**Fig. 12B**). Initially puzzled by this outcome, it was hypothesized that fragment labeling of CASP8 in isoTOP-ABPP experiments might reflect reaction with the inactive zymogen (pro-) rather than active form of this protease. Western blots confirmed that most, if not all of the CASP8 in MDA-MB-231 cell lysates existed in the pro-form (**Fig. 12C**). Next a recombinant form of pro-CASP8 was expressed with mutated cleavage sites (D374A and D384A) to prevent processing and activation. A non-catalytic cysteine C409S of pro-CASP8 was also mutated, which enabled detection of C360 labeling with IA-rhodamine by SDS-PAGE analysis (**Fig. 13A**). Several hit fragments detected in isoTOP-ABPP experiments completely blocked IA-rhodamine labeling of pro-CASP8 (**Fig. 12D**). Fragment **7** displayed the highest potency, with an IC₅₀ value of ~5 μ M (**Fig. 13A, B**), which, when combined with the low overall proteome reactivity of this fragment (3%), designated it as suitable tool compound for further studies.

[00411] Fragment **7** (50 μ M) fully blocked IA-alkyne labeling of C360 of CASP8 in isoTOP-ABPP experiments performed in both Ramos and Jurkat cell lysates (**Fig. 13C**). Next, a clickable analogue of **7** (**61**) was synthesized and it was found that this probe (25 μ M) strongly labeled pro-CASP8, but not a C360S-pro-CASP8 mutant (**Fig. 13D** and **Fig. 12E**). **7** (50 μ M) blocked labeling of pro-CASP8 by **61**, but did not inhibit labeling of active CASP8 by the Rho-

DEVD-AOMK probe ("DEVD" disclosed as SEQ ID NO: 857) developed to target active caspases (**Fig. 13D** and **Fig. 12F**). Conversely, the general caspase inhibitor Ac-DEVD-CHO ("DEVD" disclosed as SEQ ID NO: 857) (20 μ M) blocked Rho-DEVD-AOMK ("DEVD" disclosed as SEQ ID NO: 857) labeling of active CASP8, but not **61** labeling of pro-CASP8 (**Fig. 13D**, **Fig. 12F**, and **Fig. 21A**). Similar results were obtained in substrate assays, where DEVD-CHO ("DEVD" disclosed as SEQ ID NO: 857), but not **7**, blocked CASP8 activity (**Fig. 13E**). Cross-reactivity of **7** with other caspases was not observed, including recombinant, active CASP3 assayed with a substrate (**Fig. 13E**) or the Rho-DEVD-AOMK probe ("DEVD" disclosed as SEQ ID NO: 857) (**Fig. 12F**) or CASP2 and CASP7 in cell lysates measured by isoTOP-ABPP (**Fig. 12G**). Finally, to further verify that **7** preferentially reacts with pro-CASP8 over active CASP8 in complex biological systems, recombinant forms of these proteins were doped into MDA-MB-231 cell lysates followed by treatment with **7** (30 μ M, 1 h) or DMSO and analysis by isoTOP-ABPP. **7** produced a near-complete blockade of IA-alkyne labeling of C360 for pro-CASP8 ($R = 10$), but had little effect on IA-alkyne reaction with C360 of active CASP8 ($R = 1.9$) (**Fig. 13F**).

[00412] Treatment of Jurkat cell lysates with 10 or 100 μ M of **61**, followed by analysis of the combined samples by isoTOP-ABPP, confirmed direct labeling of C360 of CASP8 by **61** (**Fig. 12H**). The low R value observed for C360 in this analysis ($R = 2$) indicated near complete labeling of this cysteine by **61** at 10 μ M in cell lysates, consistent with the low μ M IC_{50} value displayed by the parent fragment **7** for inhibiting IA-rhodamine labeling of C360 of CASP8 (**Fig. 13B**). The effect of pro-CASP8 inhibition in cellular apoptosis assays was next to be evaluated. Because C360 is the catalytic nucleophile of CASP8, mutation of this residue was not possible to create a control protein for evaluating the pharmacological effects of **7** in cells. Instead, a structurally related inactive probe was developed for this purpose. It was found that bulky substituents placed on the aniline ring of **7** furnished compounds such as **62** that did not inhibit pro-CASP8 labeling by IA-rhodamine (**Fig 13B, G**). It was confirmed that **62** also did not inhibit active CASP3 or CASP8 using substrate (**Fig. 13E**) and activity-probe (**Fig. 12F**) assays and was inactive against endogenous CASP8, CASP2, or CASP7 in Jurkat lysates as determined by isoTOP-ABPP (**Fig. 12G**). Based on these data, **62** was designed as a suitable inactive control probe for studying the inhibition of pro-CASP8 by **7**. Jurkat cells were treated with **7** or **62** (30 μ M, 30 min) prior to addition of FASL or staurosporine (STS) to induce extrinsic and intrinsic apoptosis, respectively. **7**, but not **62**, completely blocked FASL-induced apoptosis (**Fig. 13H** and **Fig. 21B-C**), as well as the proteolytic processing of CASP3, CASP8, and the apoptosis marker PARP (**Fig. 13I**). In contrast, **7** did not block STS-induced intrinsic apoptosis (**Fig. 13H**) or the cleavage of PARP and CASP3, although the compound did

substantially inhibit cleavage of CASP8 in these cells (**Fig. 13I**). The non-selective caspase inhibitor VAD-FMK prevented both FASL- and STS-induced apoptosis and associated proteolytic processing events (**Fig. 13H, I**). Chemical proteomic experiments revealed that **7** fully inhibited CASP8, as well as the related initiator caspase CASP10 (but not other caspases, including CASP2, 3, 6, and 9) in Jurkat cells (**Fig. 14A** and **Fig. 22A**). It was confirmed that **7** blocked labeling of pro-CASP10 by **61** with an apparent IC₅₀ value of 4.5 μ M (**Fig. 22B-D**), but did not inhibit active CASP10 as measured by labeling with the Rho-DEVD-AOMK probe ("DEVD" disclosed as SEQ ID NO: 857) (**Fig. 21A**) or a substrate assay (**Fig. 22E**). As such, in some instances, **7** blocking CASP8 processing in both FASL- and STS-treated cells supports a model where CASP8 activation mainly occurs through auto-processing in either extrinsic or intrinsic apoptosis, but is only required for the former type of programmed cell death.

[00413] In some instances, the respective functions of CASP8 and CASP10 in extrinsic apoptosis and other cellular processes remain poorly understood in large part due to a lack of selective, non-peptidic, and cell-active inhibitors for these enzymes and the absence of animal models for CASP10 (which is not expressed in rodents). In some cases, the potency and selectivity of **7** was improved to address this issue. Conversion of the 4-piperidino moiety to a 3-piperidino group and addition of a *p*-morpholino substituent to the benzoyl ring of **7** furnished compound **63** that was separated by chiral chromatography into its two purified enantiomers, **63-R** (**Fig. 4c**) and **63-S**, the former of which showed substantially improved activity against CASP8 (apparent IC₅₀ value of 0.7 μ M (95% CI, 0.5 – 0.8); **Fig. 22F-H**) and negligible cross-reactivity with CASP10 (IC₅₀ value > 100 μ M; **Fig. 22C, D, F**). **63-S** was much less active against CASP8 (apparent IC₅₀ value of 15 μ M; **Fig. 22G, H**) and also inactive against CASP10 (**Fig. 14A**). With dual CASP8/10 (**7**) and CASP8-selective (**63-R**) ligands in hand, we next set out to investigate the biological functions of these proteases.

[00414] The effects of caspase ligands in human T cells were evaluated, where both CASP8 and CASP10 are highly expressed (**Fig. 22I**) in Jurkat cells, which are a commonly studied immortalized human T cell line. It was found that **63-R** fully blocked FasL-induced apoptosis in Jurkat cells and did so with greater potency than **7** (**Fig. 14B** and **Fig. 22J**) or **63-S** (**Fig. 22K**). Similar results were obtained in HeLa cells, which express CASP8, but not CASP10 (26) (**Fig. 22L**). In contrast to these cell line results, FasL-induced apoptosis in primary human T cells showed substantial resistance to **63-R** at all tested concentrations and instead was completely inhibited by the dual CASP8/10 ligand **7** (**Fig. 14B**). It was confirmed by chemical proteomics with probe **61** that **7** blocked both CASP8 and CASP10, while **63-R** inhibited CASP8, but not CASP10, in primary human T cells and Jurkat cells (**Fig. 14A**). Consistent with these cell death results, **7**, but not **63-R**, prevented proteolytic processing of CASP3 and CASP10 in primary

human T cells (**Fig. 22M**). In some instances, the processing of both CASP8 and the initiator caspase substrate RIP kinase were also preferentially inhibited by **7** versus **63-R** (**Fig. 22M**, indicating that CASP10 also contribute to these proteolytic events in T cells, as has been suggested by biochemical studies.

Example 2

[00415] Dimethyl fumarate (DMF) is a drug used to treat autoimmune conditions, including multiple sclerosis and psoriasis. In some instances, the mechanism of action of DMF is unclear, but is proposed to involve covalent modification of proteins and/or serving as a pro-drug that is converted to monomethyl fumarate (MMF). Using an isoTOP-ABPP approach, the mechanism of action of DMF is examined.

Chemical reagents

[00416] Assays were performed with the following reagents: dimethyl fumarate (DMF; 242926; Sigma Aldrich), monomethyl fumarate (MMF; 651419; Sigma Aldrich), dimethyl succinate (DMS; W239607; Sigma Aldrich), and buthionine sulfoximine (BSO; 14484; Cayman Chemical).

Isolation of primary human T cells

[00417] All studies using samples from human volunteers follow protocols approved by the TSRI institutional review board. Blood from healthy donors (females aged 30-49) were obtained after informed consent. Peripheral blood mononuclear cells (PBMCs) were purified over Histopaque-1077 gradients (10771; Sigma) following the manufacturer's instructions. Briefly, blood (20 × 25 mL blood aliquots) were layered over Histopaque-1077 (12.5 mL) and the samples were then fractionated by centrifugation (2000 rpm, 20 min, 20 °C, no brake). PBMC's were harvested from the Histopaque-plasma interface and washed twice with PBS. After that time, the T cells were isolated using an EasySep™ Human T Cell Isolation Kit (17951; STEMCELL) per the manufacturer's instructions.

Mice

[00418] C57BL/6J and *Nrf2*^{-/-} mice (Stock No:017009; *Nfe2l2*^{tm1Ywk}; Jackson Labs) were bred and maintained in a closed breeding facility at The Scripps Research Institute and were 6-8 weeks old when used in experiments. All mice were used in accordance with guidelines from the Institutional Animal Care and Use Committee of The Scripps Research Institute.

[00419] For the PKCθ studies, C57BL/6 mice and *Prkcg*^{-/-} mice were housed under specific pathogen-free conditions and used in accordance with a protocol approved by the La Jolla Institute for Allergy and Immunology Animal Care Committee.

Isolation of primary mouse T cells

[00420] Spleens were harvested from female mice, perfused with collagenase, and incubated at 37 °C with 5% CO₂ for 30 min. After this time, the spleens were homogenized. Cells that passed through a 100 µm cell strainer were collected and washed with RPMI. T cells were isolated from the splenocytes using the EasySep™ Mouse T cell Isolation Kit (19851; STEMCELL) according to manufacturer's instructions.

[00421] For the PKCθ studies, CD4⁺ T cells were isolated by anti-mouse CD4 magnetic particles (L3T4; BD IMag) and were cultured in RPMI-1640 medium (Gibco) supplemented with 10% (vol/vol) heat-inactivated FBS, 2 mM glutamine, 1 mM sodium pyruvate, 1 mM MEM nonessential amino acids, 100 U/mL each of penicillin G and streptomycin (Life Technologies) and recombinant IL-2 (100 U/mL, Biolegend).

T cell stimulation

[00422] 96-well plates were coated with anti-CD3 (1:200; BioXcell) and anti- CD28 (1:500; 302933; BioLegend) in PBS (100 µL/well) overnight at 4 °C. The plates were then washed twice with PBS and to each well was added 500,000 primary T cells in 100 µL of RPMI supplemented with 10% FBS, glutamine, and Pen-Strep. Cells were then treated with 100 µL of media containing compound at the indicated concentrations (final well volume of 200 µL). Cells were left at 37 °C in a 5% CO₂ incubator for the indicated periods of time and harvested by centrifugation (500g, 8 min, 4 °C), followed by washing with PBS.

Cellular analysis and sorting by flow cytometry

[00423] Cells were transferred to a round bottom 96-well plate (0720095; Fisher Scientific), harvested by centrifugation (500g, 3 min, 4 °C), washed with PBS, and stained with LIVE/DEAD fixable cell stain (L23105; ThermoFisher) according to the manufacturer's instructions. Briefly, one vial of LIVE/DEAD stain was resuspended in 50 µL of DMSO and added to 20 mL of PBS. To each well of the 96-well plate was added 200 µL of the stain, and the cells were incubated on ice for 30 min in the dark. After this time, cells were pelleted and washed once with PBS, then stained for cell surface antigens.

[00424] Flow cytometry analysis of cell surface antigens was performed with the following antibodies: Pacific Blue-conjugated anti-CD8 (1:25 dilution; clone RPA-T8; BD Biosciences), APC-conjugated anti-CD4 (1:25 dilution; clone RPA-T4; eBioscience), phycoerythrin-conjugated anti-CD25 (1:25 dilution; clone BC96; eBioscience or PC61; BioLegend (PKCθ studies)), FITC-conjugated anti-CD69 (1:25 dilution; clone FN50; eBioscience). All antibodies were diluted in 1% FBS in PBS, and 50 µL of the stain solution was added to each well. Cells were stained for 15 min on ice in the dark, after which cells were harvested by centrifugation (500g, 3 min, 4 °C), washed with 1% FBS in PBS, and resuspended in 200 µL/well of 4% PFA

in PBS. Flow cytometry acquisition was performed with BD FACSDiva™-driven BD™ LSR II flow cytometer (Becton, Dickinson and Company). Data was then analyzed with FlowJo software (Treestar Inc.). Data represent mean $\pm$ SE for four-five experiments per group.

Quantification of secreted cytokines by enzyme-linked immunosorbent assay (ELISA)

[00425] T cells were harvested and stimulated as described above. At the indicated time points, cell culture supernatants were collected and IL-2 levels were measured in clear microplates (991427; R&D Systems) according to the manufacturer's instructions (Human IL-2 DuoSet ELISA; DY202; R&D Systems). Plates were read in a Gemini SpectraMax 250 microplate reader set to 450 nm. Data represent mean $\pm$ SE for four experiments per group.

[00426] For the PKC θ studies, aliquots of transduced *Prkcd*^{-/-} CD4⁺ T cells (1×10^6) were stimulated for 48 h with anti-CD3 alone or anti-CD3 plus anti-CD28, and the concentration of IL-2 in culture supernatants was determined by enzyme-linked immunosorbent assay according to the manufacturer's instructions (BioLegend). Briefly, a 96-well plate (Corning Costar) was coated overnight at 4 °C with mAb to IL-2. Triplicates of IL-2 standards and supernatants from cultured cells were then added to the plate, followed by 2h incubation at room temperature. A biotinylated polyclonal antibody to IL-2 was added to the plate, followed by incubation for 1 h at room temperature, and then Avidin-HRP was added, followed by incubation for 30 min at room temperature. The amount of bound avidin was then assessed with TMB peroxidase that was acidified by 2 N H₂SO₄. The absorbance of each well at 450 nm was then measured with a spectrophotometric plate reader (BioTek).

Quantification of cellular glutathione (GSH) levels

[00427] Primary human T cells (2.5 million cells/mL, 20 mL per condition) were treated as indicated, harvested by centrifugation (500g, 8 min, 4 °C), and washed twice with PBS. To the cell pellet was added 75 μ L of lysis buffer. After vortexing, the samples were incubated on ice for 15 min, then harvested by centrifugation (16,000g, 10 min, 4 °C). Protein concentrations were adjusted to at least 5 mg/mL and the assay performed according to manufacturer's instructions (Sigma-Aldrich, CS1020). Data represent mean $\pm$ SE for two biological replicates.

Protein labeling and click chemistry

[00428] Cells were lysed by sonication and diluted to a concentration of 2 mg protein/mL. Protein concentrations were measured with the Bio-Rad DC™ protein assay reagents A and B (5000113, 5000114; Bio-Rad). 500 μ L of proteome sample was treated with 100 μ M of IA-alkyne probe using 10 μ L of a 10 mM DMSO stock. The labeling reactions were incubated at room temperature for 1 h upon which time the samples were conjugated to isotopically-labeled TEV-cleavable tags (TEV tags) by copper-catalyzed azide-alkyne cycloaddition (CuACC or 'click chemistry'). 60 μ L of heavy click chemistry reaction mixture was added to the DMSO-

treated control sample and 60 μ L of the light reaction mixture was added to the compound-treated sample. The click reaction mixture comprised TEV tags (10 μ L of a 5 mM stock, light (fragment treated) or heavy (DMSO treated)), CuSO_4 (10 μ L of a 50 mM stock in water), and TBTA (30 μ L of a 1.7 mM stock in 4:1 tBuOH:DMSO). To this was added TCEP (10 μ L of a 50 mM stock). The reaction was performed for 1 h at room temperature.

[00429] The light- and heavy-labeled samples were then centrifuged (16,000g, 5 min, 4 °C) to harvest the precipitated proteins. The resulting pellets were resuspended in 500 μ L of cold methanol by sonication and the heavy and light samples combined pairwise. Combined pellets were then washed with cold MeOH, after which the pellet was solubilized in PBS containing 1.2% SDS by sonication. The samples were heated at 90 °C for 5 min and subjected to streptavidin enrichment of probe-labeled proteins, sequential on-bead trypsin and TEV digestion, and liquid chromatography-tandem mass spectrometry (LC-MS/MS) according to the published isoTOP-ABPP protocols.

Peptide and protein identification

[00430] RAW Xtractor (version 1.9.9.2) was used to extract the MS2 spectra data from the raw files. MS2 data were searched against a reverse concatenated, nonredundant variant of the Human UniProt database (release-2012_11) using the ProLuCID algorithm. Cysteine residues were searched with a static modification for carboxyamidomethylation (+57.02146) and up to one differential modification for either the light or heavy TEV tags (+464.28595 or +470.29976, respectively). Peptides were required to have at least one tryptic terminus and to contain the TEV modification. ProLuCID data was filtered through DTASelect (version 2.0) to achieve a peptide false-positive rate below 1%.

R value calculation and processing

[00431] The quantification of heavy/light ratios (isoTOP-ABPP ratios, *R* values) was performed by in-house CIMAGE software using default parameters (3 MS1's per peak and signal to noise threshold of 2.5). Site-specific engagement of electrophilic compounds was assessed by blockade of IA-alkyne probe labeling. For peptides that showed a $\approx$ 95% reduction in MS1 peak area from the compound-treated proteome (light TEV tag) when compared to the DMSO treated proteome (heavy TEV tag), a maximal ratio of 20 was assigned. Overlapping peptides with the same labeled cysteine (for example, same local sequence around the labeled cysteines but different charge states, MudPIT segment numbers, or tryptic termini) were grouped together, and the median ratio from each group was recorded as the *R* value of the peptide for that run.

Analysis of cysteine conservation

[00432] For each human protein containing a DMF-sensitive cysteine, the mouse homolog was identified and the human and mouse sequences aligned using the Align tool on UniProt.

Immunofluorescent analysis of NF-kB translocation

[00433] Primary human T cells were harvested and stimulated as described above (500,000 cells/well), with concomitant treatment with DMSO or DMF for 60 min. Cells were pelleted (500g, 3 min, 4 °C), then each well was resuspended in 50 µL PBS and added to Poly-D-lysine coated coverslips (12mm; 354087; Corning® BioCoat™). Cells were allowed to adhere to the coverslips for 30-60 min at 4 °C. Coverslips were transferred to a 6 well plate and fixed with 4% PFA (157-4-100; Electron Microscopy Sciences) at room temperature for 10 min. After washing three times with PBS, cells were permeabilized with 0.1% Triton X-100 in PBS at room temperature for 10 min. Cells were washed three times with PBS, then placed cell-side-up on Parafilm. To each cover slip was added 150 µL of blocking buffer (2% BSA in PBS), and the slides were blocked for 30 min at room temperature.

[00434] The blocking buffer was aspirated, coverslips placed face down in 40 µL of antibody buffer (anti-human p65; p65Ab; FivePhoton Biochemicals; 1:500 dilution in blocking buffer), and allowed to stain overnight at 4 °C in a wet chamber. Cover slips were washed three times with PBS, then incubated with 150 µL of secondary antibody (anti-rabbit Alexa Fluor 488; A21441; Life Technologies; 1:200 dilution in PBS) for 2 h at room temperature. After washing three times with PBS, 150 µL of Hoechst counter stain was added (5 µg/mL in PBS) and coverslips were left at room temperature for 30-60 min. Cells were again washed with PBS three times, then stained with Alexa Fluor 555 Phalloidin red (8953S; Cell Signaling; 1:20 dilution in PBS). The coverslips were washed with PBS a final three times, then transferred to SuperFrost Plus slides (12-550-15, Fisherbrand) spotted with 10 µL of Prolong® Gold Antifade Mountant (P36934, ThermoFisher). The circumference of each coverslip was sealed with clear nail polish (72180; Electron Microscopy Sciences).

[00435] Images were acquired using a Zeiss 780 laser scanning confocal microscope with a 63x Objective (0.3µm image step size) and the automated stitching module to merged (10% overlap) and create a three dimensional multi-paneled mega image composite. The composite image was gathered as a z-series of at least 9 individual image panels that were auto-merged using zen software. The mega-image composite was projected into a maximum image projection in the zen software then analyzed using the colocalization modual in Zen (Zeiss Inc) and Image Pro Premier (Media Cybernetics). The Mander's Correlation Coefficients (MCC), specifically M1 and M2 between the various combination of fluorescent label (Rhodamine Phalloidin vs NFkB-P65 and Hoechst vs NFkB-p65) are calculated in ZEN (Zeiss inc) per cell and displayed

as a percent. Each cell was outlined using the region of interest module and the software then calculated the M1 and M2 correlation coefficients between the two fluorophores and tabulated the results. The fluorescent signal dynamic range and threshold cutoff of real signal was defined by multiple background and secondary controls. Correlation coefficient values were compared using Image Pro Premier (IPP) (Media Cybernetics), where images were imported as raw calibrated czi files and analyzed using a similar module in IPP. Similar results were obtained with both platforms (not shown). Data represent mean $\pm$ SE for two-three biological replicates.

Subcloning and mutagenesis

[00436] QuikChange site-directed mutagenesis was performed on a pEF4 His A plasmid containing the full length human PKC θ (residues 1-707). The PKC θ insert was excised using BamHI and XhoI, then ligated into a pMIG vector.

PKC θ retroviral transduction and stimulation

[00437] Platinum-E packaging cells were plated in a six-well plate in 2 mL RPMI-1640 medium plus 10% FBS. After 24 h, cells were transfected with empty pMIG vector or the appropriate PKC θ -expressing vector DNA (3 μ g) with TransIT-LT1 transfection reagent (Mirus Bio). After overnight incubation, the medium was replaced and cultures were maintained for another 24 h. Retroviral supernatants were then collected and filtered, supplemented with 8 μ g/mL of polybrene and used to infect CD4⁺ T cells that had been pre-activated for 24 h with plate-bound monoclonal antibody to CD3 (8 μ g/mL) and CD28 (8 μ g/mL). After centrifuging plates for 1.5-2 h at 2,000 r.p.m., cell supernatants were replaced by fresh RPMI-1640 supplemented with 10% FBS and recombinant IL-2 (100 U/mL). Cells were incubated for another 24h at 37 °C. On day 3, cells were washed, moved to new plates and cultured in RPMI-1640 medium containing 10% FBS and recombinant IL-2 (100 U/mL) without stimulation for 2 additional days before restimulation with mAb to CD3 alone or plus mAb to CD28.

PKC θ immunoprecipitation and immunoblot analysis

[00438] Cells were lysed in 1% (wt/vol) digitonin (D141, Sigma) lysis buffer (20mM Tris-HCl, pH7.5, 150mM NaCl, 5mM EDTA) supplemented with protease inhibitors (10 μ g/mL aprotinin, 10 μ g/mL leupeptin and 1mM PMSF) and phosphatase inhibitors (5 mM sodium pyrophosphate and 1 mM Na₃VO₄). Supernatants were incubated 2h with 1 μ g anti-CD28 mAb, and proteins were immunoprecipitated overnight at 4 °C with protein G-Sepharose beads (GE Healthcare). The immunoprecipitated proteins were resolved by SDS-PAGE, transferred onto a PVDF membrane and probed overnight at 4 °C with primary antibodies, followed by incubation for 1 h at room temperature with horseradish peroxidase (HRP)-conjugated secondary antibodies. Signals were visualized by enhanced chemiluminescence (ECL; GE Healthcare) and were exposed to X-ray film. Densitometry analysis was performed with ImageJ software.

Immunoblotting antibodies to CD28 (C-20) and PKC θ (C-19) were obtained from Santa Cruz Biotechnology.

DMF, but not MMF, inhibits T cell activation

[00439] Multiple sclerosis is an autoimmune disease with a prominent T cell component; as such, it was reasoned that DMF in some cases impact primary T cell activation. Consistent with this, previous reports have shown that DMF inhibits cytokine release from mouse splenocytes and promotes a Th2 phenotype via induction of IL-10-producing type II dendritic cells. The effects of DMF and MMF (**Fig. 23A**) were tested on cytokine release from primary human T cells activated with anti-CD3, anti-CD28 antibodies. Secretion of IL-2 was strongly inhibited by DMF, but not MMF (**Fig. 23B**). DMF, but not MMF or the non-electrophilic analogue dimethyl succinate (DMS, **Fig. 23A**) also blocked the expression of the early activation markers CD25 (**Fig. 23C, D**) and CD69 (**Fig. 23E**) in anti-CD3, anti-CD28-stimulated T cells. The blockade of T cell activation by DMF was concentration-dependent, with 10, 25 and 50 μ M of the drug producing marginal/negligible, partial, and near-complete inhibition, respectively (**Fig. 23B, D, E**). In some instances, the effects of DMF on cytokine release and activation markers occurred at concentrations of the drug that did not impair T cell viability (**Fig. 24**). Similar results were obtained with primary splenic T cells from C57BL/6 mice, the activation of which was also suppressed by DMF, but not MMF or DMS (**Fig. 25**). Of note, the inhibitory effects of DMF were reduced if the drug was added two hours after anti-CD3, anti-CD28 stimulation and completely ablated if the drug was added six hours after stimulation (**Fig. 23F**), suggesting that DMF inhibits an early event(s) in the T cell activation pathway

DMF effects on T cell activation are independent of Nrf2 and GSH

[00440] DMF is thought to produce neuroprotective effects through activating the Nrf2-Keap1 pathway, but whether this pathway contributes to the immunomodulatory effects of DMF is unclear. A recent study showed that DMF inhibits pro-inflammatory cytokine release from primary mouse splenocytes and this effect was comparable in wild type and Nrf2(−/−) splenocytes (Gillard, *et al.*, “DMF, but not other fumarates, inhibits NF-kappaB activity in vitro in an Nrf2-independent manner,” *J. Neuroimmunol.* **283**, 74-85 (2015)). Consistent with this, it was found that the activation of Nrf2(+/+) and (−/−) T cells was similarly sensitive to inhibition by DMF (**Fig. 26A**). In some instances, DMF also impair T cell activation through depleting glutathione (GSH), and, indeed, DMF-treated primary human T cells showed a significant decrease in cellular GSH content (**Fig. 26B**). Significant reductions in GSH were, however, also observed with the GSH synthesis inhibitor buthionine sulfoximine (BSO), which had no effect on T cell activation (**Fig. 26C, D**). In some cases, these data indicate that the blockade of T cell activation by DMF involves processes other than Nrf2 activation or GSH depletion.

Chemical proteomic discovery of DMF-sensitive Cys residues in T cells

[00441] The inhibition of T cell activation by DMF, but not the non-electrophilic analogues MMF and DMS, pointed to a mechanism that involves covalent reactivity with one or more proteins important for T cell function. As such, a globally inventory of DMF-sensitive Cys residues in primary human and mouse T cells were examined using the quantitative chemical proteomic platform isoTOP-ABPP. In this method, DMF is evaluated for its ability to block the reactivity of proteinaceous Cys residues with the general electrophilic probe iodoacetamide-alkyne (IA-alkyne). Using isotopically differentiated azide-biotin tags (containing a TEV protease-cleavable linker), Cys residues are identified and comparatively quantified for their IA-reactivity in cells treated with DMF versus DMSO control. Primary advantages of the isoTOP-ABPP platform include: 1) the competing electrophile does not itself need to be chemically altered for target identification, which is particularly beneficial when studying very small compounds like DMF; and 2) isotopic labeling occurs late in the sample processing, which facilitates the quantitative analysis of primary cells and tissues that are not readily amenable to metabolic labeling.

[00442] The isoTOP-ABPP method was performed on primary human T cells treated with DMSO or DMF (50 μ M, 4 h). Five independent replicates were performed, and the total aggregate number of unique quantified peptides and proteins began to plateau by the fourth and fifth replicate (**Fig. 28**), indicating that we approached maximal proteomic coverage of IA-reactive Cys residues in human T cells under the conditions employed. Of the more than 2400 quantified Cys residues, a small fraction ($\sim$ 40) showed substantial reductions ($>$ four-fold; isoTOP-ABPP ratio (R value) $>$ 4) in IA-alkyne labeling in DMF-treated T cells (**Fig. 27A**, and **Tables 7-9**). Similar isoTOP-ABPP analyses revealed that none of the $\sim$ 40 DMF-sensitive Cys residues were altered by MMF (50 μ M, 4 h) or BSO (2.5 mM, 4 h) treatment, which, in general, affected the reactivity of very few Cys residues across the T cell proteome (**Fig. 27A, B** and **Fig. 29**, respectively). The Cys residues targeted by DMF exhibited concentration- (**Fig. 27C** and **Tables 8-9**) and time (**Fig. 27D**) dependent increases in DMF sensitivity, as revealed by isoTOP-ABPP experiments performed with human T cells treated with lower concentrations of DMF (10 and 25 μ M, 4 h) or for shorter periods of time (50 μ M DMF, 1 or 2 h). Of note, very few DMF-sensitive Cys residues were detected in T cells treated with 10 μ M DMF, a concentration of the drug that also had limited impact T cell activation (**Fig. 23B, D, E**). These concentration- and time-dependent studies uncovered another $\sim$ 10 DMF-sensitive Cys residues that were not detected in the original 50 μ M/4 h isoTOP-ABPP experiments, likely reflecting the stochastic nature of peptide discovery in datadependent MS experiments.

[00443] The possibility that some of the alterations in Cys reactivity following DMF treatment could reflect changes in protein expression was considered; however, multiple Cys residues were quantified by isoTOP-ABPP for the majority of proteins harboring DMF-sensitive Cys residues, and, in most of these cases, the additional quantified Cys residues were clearly unaffected by DMF treatment (**Fig. 27E**). The DNAactivated protein kinase PRKDC was shown as one representative example, for which IA-alkyne reactivity was quantified for several Cys residues, only one of which (C4045) was blocked by DMF (**Fig. 27F**). These results indicate that DMF directly impaired the IAalkyne reactivity of specific Cys residues rather than indirectly affecting protein expression in human T cells.

Conservation of DMF-sensitive Cys residues in human and mouse T cells

[00444] Considering that DMF impaired the activation of both human and mouse T cells, it was surmised that at least a subset of Cys residues potentially important for mediating DMF action were conserved in humans and mice. Consistent with this, approximately two-thirds of the DMF-sensitive Cys residues discovered in human T cells are conserved in mice (**Fig. 30A** and **Table 7**). The isoTOP-ABPP experiments were performed on mouse T cells treated with DMF (50 μ M, 4 h) and found that the vast majority (> 80%) of the conserved, quantified Cys residues sensitive to DMF in human T cells were also blocked (R values > 4) by this drug in mouse T cells (**Fig. 30B** and **Tables 8-9**). These results indicate that DMF targets a similar array of Cys residues in human and mouse T cells, pointing to a specific set of proteins as candidate sites of action for this electrophilic drug.

[00445] The proteins containing DMF-sensitive Cys residues, as a whole, originated from several functional classes, including enzymes, channels, transporters, scaffolding proteins, and transcriptional regulators (**Fig. 30C**). Among these proteins were several with important immune functions (**Table 7**). DMF-sensitive Cys residues were found, for instance, in multiple proteins that are either components or regulators of the NF- κ B signaling pathway, including I κ B kinase β (IKK β or IKBKB), protein kinase C- θ (PKC θ or PRKCQ), and TNFAIP3 (**Table 7**). Consistent with these sites of DMF action and potentially others within the NF- κ B pathway, it was found that DMF treatment blocked p65 nuclear translocation (**Fig. 31**), as has been shown in other cell types. DMF-sensitive Cys residues were also found in: 1) the adenosine deaminase enzyme ADA, deleterious mutations in which cause severe combined immunodeficiency in humans, 2) the transcription factors interferon regulatory factors-4 (IRF4) and -8 (IRF8), and 3) the immunomodulatory cytokine IL-16 (**Table 7**).

DMF perturbs a CXXC motif critical for PKC θ -CD28 interactions and T cell Activation

[00446] PKC θ is a key kinase involved in T cell signaling at the immunological synapse where engagement of the T cell receptor and CD28 co-receptor initiates activation of multiple

downstream pathways, including NF- κ B. T cells from PKC θ ($-/-$) mice are defective in early activation. The isoTOP-ABPP analysis identified two DMF-sensitive Cys residues – C14 and C17 – in human (**Fig. 32A**) and mouse (**Fig. 33A**) T cells, and these Cys residues showed time- and concentration-dependent increases in DMF sensitivity (**Fig. 33B, C**), but were not affected by MMF treatment (**Fig. 33D**). Because C14 and C17 are found on the same tryptic peptide, it was difficult to distinguish whether one or both residues was sensitive to DMF treatment, but, in certain isoTOP-ABPP experiments, this tryptic peptide appeared to migrate as two adjacent peaks, both of which showed DMF sensitivity (**Fig. 32A**), suggesting that the IA-alkyne reactivity of both C14 and C17 is blocked by DMF treatment. The isoTOP-ABPP experiments also identified a third Cys in PKC θ (C322) that was unaffected by DMF treatment (**Fig. 32A**), indicating that DMF caused reductions in C14/17 reactivity rather than changes in PKC θ expression. C14 and C17 form a CXXC motif found in the C2 domain of PKC θ , but not other PKC isoforms (**Fig. 32B, C**). The C2 domain of PKC θ was recently shown to bind phosphotyrosine-containing peptides and has been postulated to stabilize plasma membrane association of PKC θ at the immunological synapse. Upon TCR/CD28 stimulation, PKC θ is recruited to the immunological synapse where it interacts with the CD28 co-receptor by associating with the CD28 cytoplasmic tail. It was found that DMF, but not MMF, blocked the interaction between PKC θ and CD28 in mouse T cells (**Fig. 32D**). A retroviral transduction was used to reconstitute PKC θ ($-/-$) T cells with either WT- or a C14S/C17S-PKC θ mutant and found that the mutant protein failed to associate with CD28 (**Fig. 32E**). PKC θ ($-/-$) T cells reconstituted with the C14S/C17SPKC θ mutant also showed impaired expression of CD25 (**Fig. 32F**) and IL-2 release (**Fig. 32G**) compared to cells reconstituted with WT PKC θ following anti-CD3, anti-CD28 treatment. Taken together, these data indicate that the C14/C17 motif within the C2 domain of PKC θ regulates localization of this kinase to the immunological synapse, and disruption of this motif by DMF or genetic mutation impairs T cell activation.

Sensitive cysteine residue sites in DMF toward probe ADA

[00447] The DMF-sensitive Cys residue C75 is located between two amino acids – G74 and R76 – that, when mutated in humans, contribute to an immunosuppressive phenotype. The amino acid 74-76 region of ADA is over 25 angstroms from the active site of the enzyme (**Fig. 34**), suggesting that it performs a non-catalytic function possibly perturbed by DMF reactivity. The DMF-sensitive Cys in IKBKB is located in the leucine-zipper domain and is distinct from another electrophile-sensitive Cys residue C179 found in the active site of this kinase.

[00448] Table 1 illustrates a list of liganded cysteines and their reactivity profiles with the fragment electrophile library from isoTOP-ABPP experiments performed in cell lysates (*in vitro*). Table 1 further shows the accession number (or the protein identifier) of the protein.

Table 1A

Identifier	Protein	SEQ ID NO:	2_500 uM_in vitro_231	2_500 uM_in vitro_ramos	3_500 uM_in vitro_231	3_500 uM_in vitro_ramos	4_250 uM_in vitro_231	4_250 uM_in vitro_ramos
Q99873_C109	PRMT1 Protein arginine N-methyltransferase 1	17	3.1	12.6	1.3	1.8	1.1	0.2
P24752_C119	ACAT1 Acetyl-CoA acetyltransferase, mitochondrial	22	3.9	3.2	2.1	5.3	1.8	0.2
P09211_C48	GSTP1 Glutathione S-transferase P	25	3.1	1.6	2.9	1.7	3.2	0.5
O14980_C34	XPO1 Exportin-1	28	2.8	4.6	1.4	1.6	1.0	0.2
P24752_C196	ACAT1 Acetyl-CoA acetyltransferase, mitochondrial	33	12.4	9.3	1.7	3.1	1.9	0.5
Q15084_C55	PDIA6 Protein disulfide-isomerase A6	51	6.3	7.1	17.9	14.9	1.3	1.2
P24752_C413	ACAT1 Acetyl-CoA acetyltransferase, mitochondrial	56	18.1	--	15.3	20.0	3.3	--
P63244_C182	GNB2L1 Guanine nucleotide-binding protein subunit beta-2-	85	1.2	1.2	20.0	4.7	0.9	--
P24752_C126	ACAT1 Acetyl-CoA acetyltransferase, mitochondrial	89	20.0	--	2.6	2.4	2.1	1.2
Q15084_C190	PDIA6 Protein disulfide-isomerase A6	96	--	15.4	19.9	13.1	1.8	1.2
Q8TAQ2_C145	SMARCC2 SWI/SNF complex subunit SMARCC2	119	8.5	14.0	9.7	6.6	1.3	--
P68036_C86	UBE2L3 Ubiquitin-conjugating enzyme E2 L3	120	2.8	2.5	1.2	3.0	1.2	--
P15374_C95	UCHL3 Ubiquitin carboxyl-terminal hydrolase isozyme L3	146	--	1.8	1.7	1.4	--	0.9
Q16763_C118	UBE2S Ubiquitin-conjugating enzyme E2 S	187	4.2	6.9	1.3	1.5	1.2	--
Q16822_C306	PCK2 Phosphoenolpyruvate carboxykinase	192	--	--	10.6	0.9	2.2	--
O14980_C528	XPO1 Exportin-1 DLLGLCEQK K.DLLGLC*EQKR.G	218	4.9	3.1	20.0	20.0	0.9	1.1
O00170_C122	AIP AH receptor-interacting protein	240	12.5	6.7	7.0	3.1	2.5	0.3
O75874_C269	IDH1 Isocitrate dehydrogenase	260	20.0	--	20.0	1.0	1.6	0.6
O75362_C286	ZNF217 Zinc finger protein 217	268	20.0	20.0	1.4	--	1.7	--
P40763_C259	STAT3 Signal transducer and activator of transcription 3	283	--	--	2.5	2.9	2.0	1.9
Q9Y3Z3_C522	SAMHD1 SAM domain and HD domain-containing protein 1	288	--	4.0	2.6	--	1.5	--
P16455_C150	MGMT Methylated-DNA--protein-cysteine methyltransferase	291	--	20.0	--	17.1	--	6.9
Q96GG9_C115	DCUN1D1 DCN1-like protein 1	293	--	20.0	--	5.5	--	--
P00813_C75	ADA Adenosine deaminase	296	--	7.3	--	1.5	--	0.1
Q14790_	CASP8 Caspase-8	335	9.8	--	3.3	2.3	12.3	--

Identifier	Protein	SEQ ID NO:	2_500 uM_in vitro_231	2_500 uM_in vitro_ramos	3_500 uM_in vitro_231	3_500 uM_in vitro_ramos	4_250 uM_in vitro_231	4_250 uM_in vitro_ramos
C360								
Q15306_C194	IRF4 Interferon regulatory factor 4	338	--	--	--	20.0	--	--
Q6L8Q7_C108	PDE12 2,5-phosphodiesterase 12	339	--	5.1	3.6	1.8	--	--
P48735_C308	IDH2 Isocitrate dehydrogenase	360	1.4	--	2.2	--	0.7	--
Q86UV5_C39	USP48 Ubiquitin carboxyl-terminal hydrolase 48	381	--	7.1	1.8	1.5	1.4	1.4
P50851_C1704	LRBA Lipopolysaccharide-responsive and beige-like ancho	388	--	3.4	--	3.6	--	1.5
O94953_C694	KDM4B Lysine-specific demethylase 4B	395	20.0	5.1	20.0	7.4	20.0	--
P19447_C342	ERCC3 TFIIH basal transcription factor complex helicase	402	20.0	--	12.1	--	1.6	1.1
Q00535_C157	CDK5 Cyclin-dependent kinase 5	407	--	3.2	--	20.0	--	1.3
Q9UPT9_C171	USP22 Ubiquitin carboxyl-terminal hydrolase 22	413	--	7.1	--	--	--	4.0
Q9HB90_C377	RRAGC Ras-related GTP-binding protein C	417	20.0	--	3.7	--	3.5	--
P50851_C2675	LRBA Lipopolysaccharide-responsive and beige-like ancho	426	--	3.0	--	5.1	1.1	--
Q9NYL2_C22	MLTK Mitogen-activated protein kinase kinase kinase MLT	430	20.0	--	1.8	--	20.0	--
Q5T1V6_C414	DDX59 Probable ATP-dependent RNA helicase DDX59	439	20.0	20.0	--	6.0	--	--
Q9HB90_C358	RRAGC Ras-related GTP-binding protein C	452	20.0	20.0	1.2	--	1.5	--
P16455_C145	MGMT Methylated-DNA--protein-cysteine methyltransferase	470	--	20.0	--	20.0	--	20.0
Q9Y5T5_C205	USP16 Ubiquitin carboxyl-terminal hydrolase 16	474	20.0	--	--	8.8	--	20.0
Q02556_C306	IRF8 Interferon regulatory factor 8	513	--	--	--	5.6	--	--
Q15910_C503	EZH2 Histone-lysine N-methyltransferase EZH2	557	--	2.8	--	2.0	--	1.5
Q96RU2_C171	USP28 Ubiquitin carboxyl-terminal hydrolase 28	569	--	1.1	--	1.7	--	--
Q16877_C159	PFKFB4 6-phosphofructo-2-kinase/fructose-2,6-bisphosphata	582	--	--	--	12.9	1.4	--
P04150_C302	NR3C1 Glucocorticoid receptor	600	2.3	--	7.6	--	--	--
Q96JH7_C219	VCPIP1 Deubiquitinating protein VCIP135	601	--	1.1	20.0	15.1	--	--
P48200_C137	IREB2 Iron-responsive element-binding protein 2	603	--	10.5	--	--	--	--
O00622_C39	CYR61 Protein CYR61	612	--	--	4.0	--	2.8	--

Identifier	Protein	SEQ ID NO:	2_500 uM_in vitro_231	2_500 uM_in vitro_amos	3_500 uM_in vitro_231	3_500 uM_in vitro_amos	4_250 uM_in vitro_231	4_250 uM_in vitro_amos
Q5T1V6_C453	DDX59 Probable ATP-dependent RNA helicase DDX59	620	20.0	20.0	20.0	--	--	--
P51617_C608	IRAK1 Interleukin-1 receptor-associated kinase 1	656	--	2.5	--	20.0	--	--
P42575_C370	CASP2 Caspase-2	661	--	--	--	7.4	--	--
P09086_C346	POU2F2 POU domain, class 2, transcription factor 2	663	--	--	1.7	5.5	--	--
Q09472_C1738	EP300 Histone acetyltransferase p300	676	--	12.7	--	--	--	--
Q01201_C109	RELB Transcription factor RelB	681	20.0	20.0	--	--	--	--
Q70CQ2_C741	USP34 Ubiquitin carboxyl-terminal hydrolase 34	688	--	--	--	3.2	--	--
P41226_C599	UBA7 Ubiquitin-like modifier-activating enzyme 7	702	--	--	--	--	--	20.0
P14598_C378	NCF1 Neutrophil cytosol factor 1	705	--	4.4	--	2.1	--	1.4
Q9C0C9_C375	UBE2O Ubiquitin-conjugating enzyme E2O	707	--	--	9.8	4.1	--	--
O00622_C134	CYR61 Protein CYR61	713	--	--	20.0	--	--	--
O00541_C361	PES1 Pescadillo homolog	718	--	--	1.5	--	--	--
P43403_C117	ZAP70 Tyrosine-protein kinase ZAP-70	726	--	--	--	--	--	--
Q96FA3_C282	PELI1 E3 ubiquitin-protein ligase pellino homolog 1	729	--	--	--	--	--	--
Q9UPT9_C44	USP22 Ubiquitin carboxyl-terminal hydrolase 22	737	--	--	--	--	--	--
Q9Y4C1_C251	KDM3A Lysine-specific demethylase 3A	753	--	--	4.4	4.0	--	--
Q70CQ2_C1090	USP34 Ubiquitin carboxyl-terminal hydrolase 34	761	--	--	--	19.7	--	--
O00622_C70	CYR61 Protein CYR61	762	--	--	20.0	--	--	--
P04150_C622	NR3C1 Glucocorticoid receptor	765	--	--	20.0	--	--	--

Table 1B

Identifier	5_500u M_in vitro_231	5_500 uM_in vitro_amos	6_500u M_in vitro_231	7_500u M_in vitro_231	7_500 uM_in vitro_amos	8_500u M_in vitro_231	8_500u M_in vitro_amos	9_500u M_in vitro_231	9_500 uM_in vitro_amos	10_500uM_in vitro_231
Q99873_C109	0.9	1.3	1.1	1.1	1.1	1.2	1.2	1.0	1.5	2.3
P24752_C119	1.8	2.4	1.1	2.3	2.0	1.8	0.7	2.2	3.1	1.5
P09211_C48	2.2	2.4	1.2	1.8	0.9	2.6	2.0	2.3	1.4	1.4
O14980_C34	1.1	1.2	1.2	1.7	1.8	2.4	1.5	1.2	1.1	1.2
P24752_C196	1.5	2.2	1.0	1.9	--	2.7	1.5	3.1	2.1	2.5
Q15084_C55	0.9	1.1	1.0	1.0	--	4.7	2.2	1.4	1.0	20.0
P24752_C413	1.3	--	1.1	4.2	--	1.4	2.1	15.3	--	13.0
P63244_C182	1.5	1.4	1.5	1.5	--	0.9	1.2	0.9	0.8	1.1

WO 2017/070611

PCT/US2016/058308

Identifier	5_500u M_invit ro_231	5_500 uM_in vitro_r amos	6_500u M_invit ro_231	7_500u M_invit ro_231	7_500 uM_in vitro_r amos	8_500u M_invit ro_231	8_500u M_invit ro_ram os	9_500u M_invit ro_231	9_500 uM_in vitro_r amos	10_5 00uM _invit ro_23 1
P24752_C126	0.9	1.3	0.9	2.2	--	2.1	12.6	1.4	--	3.2
Q15084_C190	--	1.7	1.1	1.0	1.2	20.0	2.7	2.0	1.3	--
Q8TAQ2_C1 45	--	2.8	--	1.0	--	3.1	1.7	1.0	1.6	4.2
P68036_C86	--	4.8	1.8	0.9	--	1.4	1.0	0.8	1.5	1.1
P15374_C95	1.5	1.1	1.1	0.8	--	20.0	1.2	1.0	1.2	2.0
Q16763_C118	2.8	--	--	1.5	--	--	1.1	0.8	1.0	1.7
Q16822_C306	1.7	--	--	1.4	--	2.2	1.0	1.3	--	1.9
O14980_C528	20.0	--	0.7	0.7	--	1.8	--	0.8	0.7	0.8
O00170_C122	--	--	--	--	--	--	1.3	--	1.1	--
O75874_C269	--	0.8	--	1.2	--	12.4	--	--	1.4	20.0
O75362_C286	--	--	0.9	1.3	--	1.8	--	1.2	--	1.1
P40763_C259	--	--	0.6	1.3	1.8	1.8	1.7	1.5	1.2	3.4
Q9Y3Z3_C52 2	--	5.4	--	--	--	1.8	1.1	1.0	--	1.8
P16455_C150	--	9.6	--	--	20.0	--	15.8	--	4.0	--
Q96GG9_C11 5	--	--	--	--	1.6	--	1.4	--	1.3	1.3
P00813_C75	--	2.5	--	--	1.4	--	1.3	1.3	1.1	--
Q14790_C360	1.0	1.3	--	--	--	--	1.4	--	--	2.8
Q15306_C194	--	2.2	--	--	--	--	2.1	--	1.1	--
Q6L8Q7_C10 8	--	--	--	--	--	--	--	1.3	--	--
P48735_C308	0.8	--	--	0.9	--	1.3	--	1.4	--	1.1
Q86UV5_C39	--	1.2	--	--	--	--	--	0.8	0.9	2.5
P50851_C170 4	--	3.6	--	1.8	1.6	--	1.2	1.0	1.2	--
O94953_C694	--	--	1.1	--	--	20.0	1.1	1.2	1.0	--
P19447_C342	--	4.1	1.1	1.3	2.0	5.9	--	2.5	--	3.0
Q00535_C157	--	20.0	--	--	--	--	1.3	1.2	0.8	--
Q9UPT9_C17 1	--	20.0	--	--	3.2	--	2.3	--	--	--
Q9HB90_C37 7	--	--	--	--	--	5.1	--	1.9	--	--
P50851_C267 5	--	20.0	--	--	--	--	1.1	1.5	--	--
Q9NYL2_C2 2	20.0	--	--	3.1	--	--	--	20.0	--	--
Q5T1V6_C41 4	--	--	--	--	--	--	2.0	--	--	--
Q9HB90_C35 8	--	--	1.2	--	--	--	--	1.9	--	--
P16455_C145	--	2.0	--	--	--	--	20.0	--	20.0	--
Q9Y5T5_C20 5	--	1.1	--	--	--	--	20.0	--	--	--
O00541_C272	--	12.3	--	--	--	--	1.6	--	--	--
Q02556_C306	--	2.6	--	--	20.0	--	2.2	--	1.0	--
Q15910_C503	--	20.0	1.8	--	--	--	--	1.1	--	--
Q96RU2_C17 1	--	--	0.8	--	--	7.5	--	1.1	--	1.4
Q16877_C159	--	20.0	--	--	--	--	1.0	1.5	--	1.2
P04150_C302	--	--	--	--	--	--	--	--	--	4.1

WO 2017/070611

PCT/US2016/058308

Identifier	5_500u M_invit ro_231	5_500 uM_in vitro_r amos	6_500u M_invit ro_231	7_500u M_invit ro_231	7_500 uM_in vitro_r amos	8_500u M_invit ro_231	8_500u M_invit ro_ram os	9_500u M_invit ro_231	9_500 uM_in vitro_r amos	10_5 00uM _invit ro_23 1
Q96JH7_C219	--	--	--	--	--	--	--	--	--	--
P48200_C137	--	--	--	--	2.2	--	1.6	--	--	--
O00622_C39	--	--	--	--	--	--	--	2.0	--	--
Q5T1V6_C453	--	--	--	--	--	20.0	--	2.1	--	--
P51617_C608	--	10.1	--	--	--	--	--	--	--	--
P42575_C370	--	--	--	--	--	--	2.0	--	1.2	--
P09086_C346	--	4.6	--	--	--	--	--	--	--	--
Q09472_C1738	--	20.0	--	--	--	--	--	1.5	--	--
Q01201_C109	--	--	--	--	--	--	--	--	--	--
Q70CQ2_C741	--	--	--	--	--	--	--	--	--	--
P41226_C599	--	--	--	--	--	--	--	--	--	--
P14598_C378	--	4.0	--	--	--	--	1.3	--	--	--
Q9C0C9_C375	--	--	--	--	--	--	--	--	--	--
O00622_C134	--	--	--	--	--	--	--	1.7	--	20.0
O00541_C361	--	--	--	--	--	--	--	20.0	--	--
P43403_C117	--	--	--	--	--	20.0	--	--	--	20.0
Q96FA3_C282	--	--	--	--	--	--	--	--	--	--
Q9UPT9_C44	--	20.0	--	--	--	--	1.5	--	--	--
Q9Y4C1_C251	--	--	--	--	--	--	--	--	--	--
Q70CQ2_C1090	--	20.0	--	--	--	--	--	--	--	--
O00622_C70	--	--	--	--	--	20.0	--	--	--	--
P04150_C622	--	--	--	--	--	--	--	--	--	--

Table 1C

Identifier	10_500 uM_inv itro_ra mos	11_500 uM_inv itro_23 1	11_50 0uM_i nvitro _ramo s	12_500 uM_inv itro_23 1	12_500 uM_inv itro_ra mos	13_50 0uM_i nvitro _231	13_50 0uM_i nvitro _ramo s	14_500 uM_inv itro_23 1	14_500 uM_inv itro_ra mos	15_500 uM_inv itro_23 1
Q99873_C109	1.3	20.0	20.0	1.4	1.3	1.0	0.9	0.8	1.4	0.9
P24752_C119	1.4	1.2	0.9	1.0	0.7	1.2	0.6	1.2	1.8	1.0
P09211_C48	0.8	1.7	1.1	2.9	1.1	1.5	1.2	3.2	2.1	2.5
O14980_C34	0.7	1.2	1.0	1.2	0.7	0.9	0.7	1.1	1.5	--
P24752_C196	1.7	1.5	1.3	1.2	0.9	1.6	1.7	1.1	1.4	0.9
Q15084_C55	--	20.0	20.0	3.7	3.1	1.9	1.1	1.0	1.4	0.8
P24752_C413	1.6	1.9	1.7	1.5	1.1	5.5	7.3	1.3	1.6	1.2
P63244_C182	1.1	1.0	1.1	0.9	0.8	1.4	0.8	1.2	1.5	--
P24752_C126	1.5	1.5	--	1.0	1.1	9.6	20.0	1.1	1.4	0.8
Q15084_C190	3.3	20.0	20.0	--	20.0	1.5	1.3	1.1	1.8	0.9
Q8TAQ2_C145	--	20.0	--	2.8	1.6	--	0.8	1.1	2.3	--
P68036_C86	0.4	--	1.0	--	--	1.1	0.8	1.0	1.1	--
P15374_C95	--	0.8	0.8	--	0.5	--	0.8	1.2	1.0	--
Q16763_C118	--	--	--	--	0.6	--	0.7	0.7	1.2	--

WO 2017/070611

PCT/US2016/058308

Identifier	10_500 uM_inv itro_ra mos	11_500 uM_inv itro_23 1	11_50 0uM_i nvitro _ramo s	12_500 uM_inv itro_23 1	12_500 uM_inv itro_ra mos	13_50 0uM_i nvitro _231	13_50 0uM_i nvitro _ramo s	14_500 uM_inv itro_23 1	14_500 uM_inv itro_ra mos	15_500 uM_inv itro_23 1
Q16822_C306	1.1	3.2	1.2	1.5	0.7	0.9	1.0	--	20.0	--
O14980_C528	--	1.0	0.7	1.1	--	--	0.5	20.0	20.0	0.8
O00170_C122	1.7	--	2.8	--	1.3	--	0.9	--	2.5	--
O75874_C269	--	1.1	0.7	--	--	2.7	--	0.9	1.3	--
O75362_C286	--	1.2	1.2	1.0	--	0.9	0.9	0.9	--	--
P40763_C259	--	20.0	20.0	--	--	--	1.2	0.8	3.1	--
Q9Y3Z3_C52 2	--	2.1	--	--	--	0.8	--	1.7	1.5	--
P16455_C150	20.0	--	20.0	--	20.0	--	4.2	5.5	5.1	--
Q96GG9_C11 5	1.0	--	0.9	--	0.6	--	--	1.1	1.6	--
P00813_C75	1.0	--	7.5	--	0.7	--	0.9	1.5	1.7	--
O14933_C98	--	--	3.4	--	1.3	--	1.0	--	2.8	--
Q14790_C360	1.5	20.0	20.0	--	--	--	--	1.4	1.9	--
Q15306_C194	2.5	--	4.3	--	3.7	--	1.1	1.3	2.2	--
Q6L8Q7_C10 8	--	4.2	--	--	1.5	--	1.0	0.5	2.2	1.1
P48735_C308	--	1.1	--	1.7	--	1.1	--	4.7	--	--
Q86UV5_C39	--	--	1.0	--	--	0.8	0.6	0.8	--	--
P50851_C170 4	0.8	--	1.4	--	--	0.7	0.8	1.9	1.7	--
O94953_C694	1.4	--	1.9	--	--	1.4	1.1	--	2.0	--
P19447_C342	--	2.9	1.5	3.0	--	1.0	--	4.7	--	--
Q00535_C157	--	--	1.3	--	1.0	--	--	6.1	2.3	--
Q9UPT9_C17 1	3.8	--	8.2	--	4.1	--	0.8	4.0	4.0	--
Q9HB90_C37 7	--	--	1.6	--	--	2.1	--	--	20.0	--
P50851_C267 5	1.3	--	1.3	--	1.5	--	--	1.9	2.9	--
Q9NYL2_C2 2	--	--	--	--	--	5.2	--	1.6	--	--
Q5T1V6_C41 4	--	--	--	--	--	--	1.2	--	4.0	--
Q9HB90_C35 8	--	--	--	--	--	1.5	--	0.9	2.1	1.1
P16455_C145	--	--	20.0	--	--	--	20.0	1.5	2.0	--
Q9Y5T5_C20 5	--	20.0	--	2.2	--	--	3.3	--	--	--
O00541_C272	2.2	--	4.2	--	--	1.7	--	2.6	--	--
Q02556_C306	3.6	--	3.1	--	--	--	0.7	2.4	2.3	--
Q15910_C503	--	--	1.3	--	--	--	0.8	0.6	1.9	0.8
Q96RU2_C17 1	--	20.0	--	--	--	--	2.4	1.0	--	--
Q16877_C159	--	--	0.6	--	--	0.9	--	2.3	--	--
P04150_C302	--	--	1.5	--	1.3	--	--	--	2.2	--
Q96JH7_C21 9	--	--	--	--	--	--	--	0.6	1.5	--
P48200_C137	2.1	--	3.7	--	--	--	--	5.2	--	--
O00622_C39	--	9.6	--	--	--	1.2	--	3.6	--	--
Q5T1V6_C45 3	--	20.0	--	--	--	--	1.4	2.2	--	--

WO 2017/070611

PCT/US2016/058308

Identifier	10_500 uM_inv itro_ra mos	11_500 uM_inv itro_23 1	11_50 0uM_i nvitro _ramo s	12_500 uM_inv itro_23 1	12_500 uM_inv itro_ra mos	13_50 0uM_i nvitro _231	13_50 0uM_i nvitro _ramo s	14_500 uM_inv itro_23 1	14_500 uM_inv itro_ra mos	15_500 uM_inv itro_23 1
P51617_C608	--	--	1.3	--	--	--	--	--	--	--
P42575_C370	1.2	--	6.0	--	--	--	--	--	--	--
P09086_C346	--	--	1.7	--	--	--	0.6	--	--	--
Q09472_C173 8	--	--	--	--	--	--	0.7	2.1	--	--
Q01201_C109	--	--	--	--	--	--	--	1.7	2.4	--
Q70CQ2_C74 1	0.9	--	--	--	--	0.8	--	--	--	--
P41226_C599	--	--	1.9	--	--	--	--	--	--	--
P14598_C378	1.1	--	--	--	--	--	0.7	--	--	--
Q9C0C9_C37 5	0.8	--	--	--	--	--	--	2.8	--	--
O00622_C134	--	--	--	--	--	--	--	0.7	--	--
O00541_C361	--	--	--	--	--	--	--	--	--	--
P43403_C117	20.0	--	--	--	--	--	--	--	1.3	--
Q96FA3_C28 2	--	--	--	--	--	--	--	1.4	--	--
Q9UPT9_C44	--	--	--	--	--	--	--	--	--	--
Q9Y4C1_C25 1	--	--	--	--	--	--	--	--	--	--
Q70CQ2_C10 90	--	--	--	--	--	--	--	--	--	--
O00622_C70	--	--	--	--	--	--	--	--	--	--
P04150_C622	--	1.7	--	--	--	--	--	--	--	--

Table 1D

Identifier	15_500 uM_inv itro_ra mos	27_500 uM_inv itro_23 1	20_50 0uM_i nvitro _231	20_50 0uM_i nvitro _ramo s	21_500 uM_inv itro_23 1	21_50 0uM_i nvitro _ramo s	22_500 uM_inv itro_23 1	22_50 0uM_i nvitro _ramo s	23_50 0uM_i nvitro _231	23_500 uM_inv itro_ra mos
Q99873_C109	1.2	0.7	1.9	2.2	1.1	1.5	1.0	1.0	0.8	1.0
P24752_C119	1.2	0.9	3.6	2.6	0.9	1.1	1.0	0.8	1.7	2.0
P09211_C48	1.3	1.3	2.5	1.1	1.1	1.3	0.9	0.7	0.8	0.9
O14980_C34	1.2	0.9	1.0	1.2	0.9	1.2	1.1	--	3.4	5.3
P24752_C196	1.6	0.9	3.7	2.9	1.0	1.1	1.4	0.7	--	1.8
Q15084_C55	1.0	0.8	1.7	2.0	1.1	2.1	2.1	1.6	0.8	1.0
P24752_C413	0.9	0.9	20.0	14.0	1.4	1.4	1.1	1.1	1.1	1.2
P63244_C182	1.3	1.0	1.4	1.1	0.9	0.9	1.6	--	4.6	--
P24752_C126	--	0.9	10.6	20.0	1.2	1.4	2.4	1.7	0.8	0.8
Q15084_C190	1.2	0.9	2.1	2.1	1.1	2.3	2.2	4.2	0.8	1.1
Q8TAQ2_C1 45	1.2	0.9	2.9	3.9	1.2	--	--	1.2	1.3	--
P68036_C86	1.3	1.2	1.1	1.2	1.2	2.0	--	0.8	7.4	13.2
P15374_C95	0.7	0.4	1.0	0.9	1.1	--	--	0.7	0.8	1.1
Q16763_C118	1.2	0.8	--	2.9	0.9	--	--	0.9	1.6	2.4
Q16822_C306	--	1.4	2.5	1.6	1.2	2.0	--	--	3.0	--
O14980_C528	1.0	1.2	1.4	4.0	0.9	0.8	--	--	--	2.9
O00170_C122	1.4	0.4	--	2.2	1.4	--	1.3	1.1	0.7	1.4
O75874_C269	0.6	0.4	--	0.6	1.5	--	1.0	--	0.7	1.1
O75362_C286	--	20.0	1.3	--	0.9	--	--	--	20.0	--
P40763_C259	1.9	0.7	--	3.0	--	2.9	2.0	--	1.1	0.8

WO 2017/070611

PCT/US2016/058308

Identifier	15_500 uM_inv itro_ra mos	27_500 uM_inv itro_23 1	20_50 0uM_i nvitro _231	20_50 0uM_i nvitro _ramo s	21_500 uM_inv itro_23 1	21_50 0uM_i nvitro _ramo s	22_500 uM_inv itro_23 1	22_50 0uM_i nvitro _ramo s	23_50 0uM_i nvitro _231	23_500 uM_inv itro_ra mos
Q9Y3Z3_C52 2	1.6	0.8	2.1	1.3	--	1.4	--	0.8	--	--
P16455_C150	1.9	1.0	--	18.3	--	17.2	--	2.9	--	1.0
Q96GG9_C11 5	0.8	1.1	1.2	1.2	1.0	1.0	1.0	0.8	--	20.0
P00813_C75	1.1	1.1	--	1.5	--	1.3	--	1.3	--	1.1
O14933_C98	1.3	1.2	--	3.9	--	--	--	2.0	0.7	1.2
Q14790_C360	--	0.8	1.9	1.5	1.9	1.9	--	--	1.4	1.4
Q15306_C194	2.2	--	--	3.0	--	1.7	--	1.1	--	1.9
Q6L8Q7_C10 8	--	0.5	1.6	--	--	--	--	--	1.0	--
P48735_C308	--	1.9	0.5	--	0.6	--	--	--	--	--
Q86UV5_C39	0.7	0.8	1.1	1.8	20.0	1.2	--	--	--	1.3
P50851_C170 4	1.9	--	2.0	1.7	--	1.6	--	--	--	3.4
O94953_C694	1.1	1.4	2.8	2.0	--	1.8	--	--	--	3.2
P19447_C342	--	0.8	2.6	--	6.1	--	--	--	1.6	--
Q00535_C157	--	--	--	2.0	--	--	--	0.9	--	2.8
Q9UPT9_C17 1	--	--	--	20.0	--	--	--	1.1	--	5.1
Q9HB90_C37 7	--	1.0	--	--	1.4	--	1.7	0.9	1.5	--
P50851_C267 5	1.6	--	--	--	0.9	1.5	1.1	1.1	--	2.6
Q9NYL2_C2 2	--	1.1	20.0	--	2.0	--	1.0	--	1.4	--
Q5T1V6_C41 4	1.0	0.8	--	3.1	1.2	--	--	1.3	--	1.4
Q9HB90_C35 8	--	0.3	--	--	1.2	--	1.7	--	--	--
P16455_C145	1.3	--	--	20.0	--	--	--	--	--	1.0
Q9Y5T5_C20 5	--	1.6	--	--	1.6	--	--	1.6	--	1.2
O00541_C272	--	1.1	--	--	--	5.5	1.6	--	--	2.0
Q02556_C306	2.1	--	--	2.1	--	--	--	--	--	--
Q15910_C503	--	--	--	--	--	--	--	--	--	2.0
Q96RU2_C17 1	--	0.9	--	2.4	1.1	--	--	--	--	1.2
Q16877_C159	--	--	--	1.0	--	1.2	--	--	--	1.0
P04150_C302	0.9	1.3	--	--	--	--	--	0.9	--	4.7
Q96JH7_C21 9	1.3	1.7	--	1.2	--	--	--	--	--	--
P48200_C137	2.0	1.4	--	5.2	--	--	--	--	--	1.6
O00622_C39	--	1.0	--	--	--	--	--	--	4.4	--
Q5T1V6_C45 3	--	--	--	--	--	--	--	--	--	--
P51617_C608	--	--	--	--	--	--	--	--	--	1.5
P42575_C370	--	0.9	--	--	--	--	--	--	--	2.1
P09086_C346	0.8	--	--	2.3	0.9	--	--	--	--	1.1
Q09472_C173 8	--	--	--	3.2	--	--	--	--	--	--
Q01201_C109	--	0.8	--	--	--	--	--	--	--	1.3
Q70CQ2_C74	--	--	2.8	1.7	--	1.7	--	--	--	--

WO 2017/070611

PCT/US2016/058308

Identifier	15_500 uM_inv itro_ra mos	27_500 uM_inv itro_23 1	20_50 0uM_i nvitro _231	20_50 0uM_i nvitro _ramo s	21_500 uM_inv itro_23 1	21_50 0uM_i nvitro _ramo s	22_500 uM_inv itro_23 1	22_50 0uM_i nvitro _ramo s	23_50 0uM_i nvitro _231	23_500 uM_inv itro_ra mos
1										
P41226_C599	--	0.7	1.1	12.2	--	2.5	--	--	--	3.2
P14598_C378	--	--	--	--	--	--	--	--	--	--
Q9C0C9_C37 5	--	0.7	--	--	--	2.4	--	--	--	1.4
O00622_C134	--	1.5	20.0	--	--	--	--	--	--	--
O00541_C361	--	0.6	20.0	20.0	--	--	--	--	--	1.1
P43403_C117	20.0	--	--	20.0	--	--	--	--	--	--
Q96FA3_C28 2	--	--	--	--	--	--	--	--	--	--
Q9UPT9_C44	--	20.0	--	--	--	--	--	--	--	--
Q9Y4C1_C25 1	--	--	--	--	0.7	--	--	--	--	--
Q70CQ2_C10 90	--	--	--	2.7	--	0.9	--	--	--	--
O00622_C70	--	1.1	--	--	--	--	--	--	--	--
P04150_C622	--	--	--	--	--	--	--	--	--	--

Table 1E

Identifier	24_500 uM_in vitro_r amos	25_500 uM_in vitro_2 31	26_500 uM_in vitro_r amos	27_500 uM_in vitro_2 31	27_500 uM_in vitro_r amos	28_500 uM_in vitro_2 31	28_500 uM_in vitro_r amos	29_500 uM_in vitro_2 31	29_500 uM_in vitro_r amos	30_500 uM_in vitro_2 31
Q99873_C109	1.0	1.0	0.9	0.9	1.0	0.8	1.1	1.0	0.9	1.1
P24752_C119	1.3	1.0	1.0	1.0	1.0	1.1	1.1	1.3	0.6	0.7
P09211_C48	1.6	1.8	0.9	1.0	1.0	1.3	1.5	1.0	0.4	1.4
O14980_C34	1.1	0.9	0.9	1.0	1.0	0.8	0.9	1.3	0.9	0.7
P24752_C196	1.1	0.9	1.0	1.0	1.1	1.0	1.1	1.3	1.0	0.8
Q15084_C55	1.3	1.0	0.9	1.1	1.2	1.1	1.5	1.0	1.3	2.2
P24752_C413	1.2	1.0	0.9	1.1	1.3	1.2	1.4	1.3	0.7	1.1
P63244_C182	2.6	1.2	1.0	0.8	1.1	1.0	1.1	1.3	1.2	1.2
P24752_C126	1.1	0.9	0.9	1.0	1.3	1.7	1.6	1.1	0.8	1.1
Q15084_C190	1.7	1.0	0.9	--	1.1	1.1	1.6	1.2	--	2.3
Q8TAQ2_C145	5.4	0.9	0.9	--	1.2	1.1	1.2	1.0	--	1.1
P68036_C86	--	1.5	0.7	0.9	--	1.4	--	0.9	1.5	--
P15374_C95	1.1	0.8	0.8	--	1.0	1.1	0.9	0.9	1.0	0.7
Q16763_C118	1.2	1.0	0.9	--	0.9	1.0	0.8	--	--	0.7
Q16822_C306	--	0.8	--	1.5	1.1	1.1	--	1.1	0.5	1.2
O14980_C528	20.0	--	0.9	0.9	0.8	--	--	1.0	--	--
O00170_C122	1.3	--	0.9	--	1.0	0.9	1.2	--	1.3	1.1
O75874_C269	0.9	1.0	--	1.0	--	1.1	--	0.8	--	0.7
O75362_C286	--	20.0	--	0.9	1.0	1.0	1.0	0.9	--	--
P40763_C259	1.7	0.9	--	--	--	1.0	1.1	1.5	--	--
Q9Y3Z3_C522	1.1	--	0.9	1.0	0.9	1.0	--	1.5	--	0.7
P16455_C150	4.0	--	0.9	--	2.3	--	1.7	--	1.3	--
Q96GG9_C115	1.7	--	0.9	--	1.0	--	--	--	--	0.7
P00813_C75	1.2	--	0.8	--	1.0	--	0.9	--	--	--
O14933_C98	--	--	1.0	--	1.0	--	1.2	--	--	1.3
Q14790_C360	1.4	1.0	--	--	1.2	1.1	--	--	--	2.0
Q15306_C194	1.9	--	1.1	--	1.0	--	1.1	--	1.2	--
Q6L8Q7_C108	1.4	0.8	1.1	--	1.3	1.0	1.2	--	--	1.1
P48735_C308	--	14.5	--	1.0	--	0.9	--	1.0	--	0.7

WO 2017/070611

PCT/US2016/058308

Identifier	24_500 uM_in vitro_r amos	25_500 uM_in vitro_2 31	26_500 uM_in vitro_r amos	27_500 uM_in vitro_2 31	27_500 uM_in vitro_r amos	28_500 uM_in vitro_2 31	28_500 uM_in vitro_r amos	29_500 uM_in vitro_2 31	29_500 uM_in vitro_r amos	30_500 uM_in vitro_2 31
Q86UV5_C39	1.0	1.0	--	--	--	--	--	1.2	--	--
P50851_C1704	1.3	0.8	--	1.1	--	--	--	--	--	--
O94953_C694	--	--	--	1.0	--	--	--	1.2	--	--
P19447_C342	--	0.9	--	1.1	--	--	--	0.8	--	1.3
Q00535_C157	--	--	0.9	--	1.0	0.9	1.0	--	--	--
Q9UPT9_C171	3.1	1.1	1.0	--	1.2	--	1.3	--	--	--
Q9HB90_C377	1.5	1.0	--	--	--	1.3	--	1.3	--	1.4
P50851_C2675	1.0	--	0.9	--	--	--	--	1.1	--	0.9
Q9NYL2_C22	--	0.8	--	--	--	1.4	--	1.2	--	1.1
Q5T1V6_C414	1.7	--	--	--	1.0	--	--	--	--	--
Q9HB90_C358	0.9	0.9	--	--	--	1.1	--	--	--	1.0
P16455_C145	20.0	--	--	--	20.0	--	20.0	--	20.0	--
Q9Y5T5_C205	--	--	--	--	0.9	--	1.6	--	--	--
O00541_C272	2.6	--	--	--	--	--	--	0.9	--	--
Q02556_C306	1.2	--	--	--	1.1	--	--	--	--	--
Q15910_C503	1.3	0.8	--	--	--	--	--	1.4	--	--
Q96RU2_C171	--	--	--	1.2	--	--	--	--	--	--
Q16877_C159	--	--	--	--	--	--	--	--	--	--
P04150_C302	--	--	--	1.0	--	--	--	--	--	--
Q96JH7_C219	1.2	--	--	--	--	--	--	--	--	--
P48200_C137	--	--	--	--	--	--	--	--	--	--
O00622_C39	--	1.0	--	--	--	1.0	--	--	--	--
Q5T1V6_C453	1.7	0.9	--	1.0	--	--	--	--	--	--
P51617_C608	1.3	--	--	--	--	--	--	--	--	--
P42575_C370	2.2	--	--	--	--	--	--	--	--	--
P09086_C346	1.1	0.7	--	--	--	--	--	--	--	--
Q09472_C1738	--	1.6	--	1.0	--	--	--	1.3	--	--
Q01201_C109	1.5	--	--	--	--	--	1.1	--	--	--
Q70CQ2_C741	--	--	--	--	--	--	--	--	--	--
P41226_C599	--	--	--	--	1.3	--	--	--	--	--
P14598_C378	--	--	--	--	--	--	--	--	--	--
Q9C0C9_C375	--	--	--	--	--	--	--	--	--	--
O00622_C134	--	1.0	--	--	--	--	--	--	--	--
O00541_C361	--	--	--	2.0	--	--	--	--	--	--
P43403_C117	--	--	--	--	--	--	1.1	--	--	--
Q96FA3_C282	--	--	0.5	--	--	--	0.7	--	--	--
Q9UPT9_C44	--	--	--	--	1.2	--	--	--	--	--
Q9Y4C1_C251	1.3	--	--	--	--	--	--	--	--	--
Q70CQ2_C109 0	--	--	--	--	--	--	--	--	--	--
O00622_C70	--	--	--	--	--	--	--	--	--	--
P04150_C622	--	--	--	--	--	--	--	--	--	--

Table 1F

Identifier	30_500 uM_in vitro_r amos	31_500 uM_in vitro_2 31	31_500 uM_in vitro_r amos	32_500 uM_in vitro_2 31	32_500 uM_in vitro_r amos	33_500 uM_in vitro_2 31	33_500 uM_in vitro_r amos	34_500 uM_in vitro_2 31	34_500 uM_in vitro_r amos	35_500 uM_in vitro_2 31
Q99873_C109	1.8	1.0	1.3	1.7	1.8	1.0	0.9	0.8	1.1	0.8
P24752_C119	0.9	1.3	0.7	1.7	1.3	0.7	0.7	0.8	--	0.9
P09211_C48	1.2	2.0	1.3	1.9	1.4	2.0	0.7	0.8	1.0	0.7
O14980_C34	1.1	1.0	0.8	1.2	1.0	0.8	0.9	0.8	1.1	0.8

WO 2017/070611

PCT/US2016/058308

Identifier	30_500 uM_in vitro_r amos	31_500 uM_in vitro_2 31	31_500 uM_in vitro_r amos	32_500 uM_in vitro_2 31	32_500 uM_in vitro_r amos	33_500 uM_in vitro_2 31	33_500 uM_in vitro_r amos	34_500 uM_in vitro_2 31	34_500 uM_in vitro_r amos	35_500 uM_in vitro_2 31
P24752_C196	1.1	1.2	0.9	2.3	2.0	0.9	1.3	0.8	1.0	0.9
Q15084_C55	--	1.2	1.1	20.0	20.0	3.3	--	0.9	--	0.9
P24752_C413	1.9	1.6	1.0	6.8	4.1	1.3	--	0.9	1.3	--
P63244_C182	1.6	1.3	--	1.9	--	--	1.1	--	--	0.9
P24752_C126	1.6	1.1	--	7.3	8.6	4.3	--	--	1.1	--
Q15084_C190	--	1.3	1.2	20.0	20.0	3.3	--	0.9	--	0.9
Q8TAQ2_C145	1.5	1.1	0.9	20.0	20.0	--	1.5	--	--	0.9
P68036_C86	1.5	1.7	1.3	1.1	0.9	--	0.9	--	--	--
P15374_C95	1.0	0.9	0.7	1.2	1.0	--	--	--	--	--
Q16763_C118	1.1	0.9	0.8	1.1	1.1	--	--	--	--	--
Q16822_C306	1.1	1.9	1.2	3.2	--	--	--	--	--	0.9
O14980_C528	--	5.2	3.2	1.7	1.2	--	--	--	1.1	--
O00170_C122	2.0	--	0.6	3.7	1.9	--	--	1.0	--	--
O75874_C269	--	0.4	0.7	--	--	1.2	--	--	--	--
O75362_C286	2.1	--	--	--	--	--	--	--	--	--
P40763_C259	--	1.7	--	10.2	--	--	--	--	--	--
Q9Y3Z3_C522	--	1.5	--	--	1.8	0.8	--	--	--	--
P16455_C150	5.1	--	2.8	--	20.0	--	1.9	--	1.4	--
Q96GG9_C115	1.1	--	0.9	--	0.9	--	--	--	--	--
P00813_C75	3.1	--	0.8	--	1.7	--	0.8	--	1.1	--
O14933_C98	2.1	--	--	13.8	4.6	--	1.2	0.8	--	--
Q14790_C360	5.2	--	--	--	--	3.1	--	--	--	--
Q15306_C194	1.7	--	1.1	--	3.7	--	1.1	--	--	--
Q6L8Q7_C108	1.6	2.0	--	2.5	1.9	--	--	--	--	0.8
P48735_C308	--	1.3	1.1	--	--	--	--	--	--	--
Q86UV5_C39	--	--	--	--	--	--	--	--	--	--
P50851_C1704	--	--	1.4	--	2.0	--	--	--	--	--
O94953_C694	--	--	--	--	--	--	--	--	--	--
P19447_C342	--	--	--	--	--	--	--	--	--	--
Q00535_C157	--	--	1.1	1.7	1.7	--	--	0.7	--	--
Q9UPT9_C171	2.3	2.1	--	--	9.3	--	--	--	--	--
Q9HB90_C377	--	2.2	--	--	1.5	1.1	--	--	--	--
P50851_C2675	--	--	--	1.8	1.1	--	--	--	1.1	--
Q9NYL2_C22	--	--	--	2.1	--	3.7	--	0.8	--	--
Q5T1V6_C414	2.0	--	--	--	--	--	--	0.6	--	--
Q9HB90_C358	--	2.3	--	1.9	1.2	--	--	--	--	--
P16455_C145	--	--	1.0	--	20.0	--	--	--	--	--
Q9Y5T5_C205	--	--	--	--	--	--	--	--	--	--
O00541_C272	--	--	--	--	--	1.4	--	--	--	0.8
Q02556_C306	2.0	--	--	--	--	--	--	--	--	--
Q15910_C503	--	--	--	--	1.7	--	--	--	--	--
Q96RU2_C171	--	--	--	--	--	--	--	--	--	--
Q16877_C159	--	--	--	--	1.5	--	--	--	1.4	--
P04150_C302	--	--	--	--	--	--	--	--	--	--
Q96JH7_C219	--	--	1.0	--	--	--	--	--	--	--
P48200_C137	--	--	--	--	--	--	--	--	--	--
O00622_C39	--	1.2	--	--	--	--	--	--	--	--
Q5T1V6_C453	--	--	--	--	--	--	--	--	--	--
P51617_C608	--	--	--	5.8	2.9	--	--	--	--	--
P42575_C370	4.1	--	--	--	--	--	--	--	--	--
P09086_C346	--	--	0.6	--	--	--	--	--	--	--
Q09472_C1738	--	--	--	--	--	--	--	--	--	--
Q01201_C109	--	--	--	--	20.0	--	--	--	--	--

WO 2017/070611

PCT/US2016/058308

Identifier	30_500 uM_in vitro_r amos	31_500 uM_in vitro_2 31	31_500 uM_in vitro_r amos	32_500 uM_in vitro_2 31	32_500 uM_in vitro_r amos	33_500 uM_in vitro_2 31	33_500 uM_in vitro_r amos	34_500 uM_in vitro_2 31	34_500 uM_in vitro_r amos	35_500 uM_in vitro_2 31
Q70CQ2_C741	--	4.2	--	--	--	--	--	--	--	--
P41226_C599	--	--	--	--	--	--	--	--	--	--
P14598_C378	--	--	--	--	--	--	--	--	--	--
Q9C0C9_C375	--	--	--	--	--	--	--	--	--	--
O00622_C134	--	--	--	--	--	--	--	--	--	--
O00541_C361	--	--	--	--	--	--	--	--	--	--
P43403_C117	--	--	--	--	--	--	--	--	--	--
Q96FA3_C282	1.5	--	--	--	--	--	--	--	--	--
Q9UPT9_C44	1.7	--	--	--	--	--	--	--	--	--
Q9Y4C1_C251	--	--	--	--	--	--	--	--	--	--
Q70CQ2_C1090	--	--	--	--	--	--	--	--	--	--
O00622_C70	--	--	--	--	--	--	--	--	--	--
P04150_C622	--	4.8	--	--	--	--	--	--	--	--

Table 1G

Identifier	35_500 uM_in vitro_r amos	36_500 uM_in vitro_2 31	37_500 uM_in vitro_r amos	38_500 uM_in vitro_2 31	38_500 uM_in vitro_r amos	39_500 uM_in vitro_2 31	40_500 uM_in vitro_2 31	40_500 uM_in vitro_r amos	41_500 uM_in vitro_2 31	41_500 uM_in vitro_r amos
Q99873_C109	1.0	1.0	1.0	0.9	1.4	1.5	0.8	1.3	0.9	0.9
P24752_C119	1.0	1.2	0.8	1.0	1.7	1.7	0.8	0.9	1.0	0.8
P09211_C48	1.3	1.6	1.2	1.9	--	2.0	1.6	1.6	1.3	0.9
O14980_C34	1.3	1.3	0.9	1.0	1.0	1.3	0.8	1.1	3.1	2.0
P24752_C196	1.4	1.5	1.0	1.0	1.3	3.0	0.8	--	1.3	1.0
Q15084_C55	--	--	1.0	1.0	1.3	3.9	0.8	--	0.9	0.8
P24752_C413	1.3	2.7	0.9	1.1	20.0	2.7	0.9	--	0.8	0.9
P63244_C182	0.9	1.3	0.9	1.1	--	1.0	1.0	0.6	--	4.7
P24752_C126	10.4	--	0.9	1.0	1.4	--	0.9	--	1.0	0.9
Q15084_C190	--	--	1.1	1.1	--	7.0	0.9	1.7	1.0	--
Q8TAQ2_C145	--	--	--	1.1	1.9	1.2	0.9	1.1	1.8	2.8
P68036_C86	--	--	--	--	1.5	1.9	0.8	0.9	9.8	7.9
P15374_C95	--	--	--	2.7	12.5	--	--	0.9	1.0	0.8
Q16763_C118	--	--	--	1.0	1.0	1.4	0.9	0.8	--	3.4
Q16822_C306	1.0	--	1.0	1.0	--	--	--	--	1.2	1.0
O14980_C528	0.8	--	0.9	1.0	--	--	--	--	0.6	--
O00170_C122	--	--	--	1.5	1.5	--	--	--	1.1	1.0
O75874_C269	--	1.7	--	0.7	--	--	--	0.8	0.8	--
O75362_C286	--	--	0.9	0.9	1.2	20.0	--	--	--	0.6
P40763_C259	1.1	--	0.9	--	1.5	--	--	1.4	0.9	--
Q9Y3Z3_C522	--	--	0.7	--	--	1.1	0.8	--	0.9	--
P16455_C150	1.3	--	1.0	--	1.8	--	--	1.9	--	0.9
Q96GG9_C115	0.9	--	0.8	1.2	1.0	--	--	0.9	--	--
P00813_C75	1.1	--	0.9	--	2.3	--	--	0.9	--	1.1
O14933_C98	--	--	--	1.6	1.7	--	--	1.3	1.2	1.1
Q14790_C360	1.3	--	1.0	1.0	4.1	--	--	--	1.0	0.8
Q15306_C194	1.0	--	0.9	--	2.1	--	--	1.1	--	1.2
Q6L8Q7_C108	--	--	--	0.9	6.8	3.2	0.8	--	1.3	--
P48735_C308	--	1.3	--	1.2	--	--	--	--	--	--
Q86UV5_C39	1.0	--	0.7	3.7	--	--	--	1.6	--	--
P50851_C1704	1.1	--	1.0	--	2.5	--	--	--	--	--
O94953_C694	1.1	--	1.1	--	2.4	--	--	--	--	--
P19447_C342	--	--	--	1.3	--	--	--	--	1.1	--

WO 2017/070611

PCT/US2016/058308

Identifier	35_500 uM_in vitro_r amos	36_500 uM_in vitro_2 31	37_500 uM_in vitro_r amos	38_500 uM_in vitro_2 31	38_500 uM_in vitro_r amos	39_500 uM_in vitro_2 31	40_500 uM_in vitro_2 31	40_500 uM_in vitro_r amos	41_500 uM_in vitro_2 31	41_500 uM_in vitro_r amos
Q00535_C157	--	--	--	0.9	--	--	--	--	1.5	--
Q9UPT9_C171	1.5	--	--	1.4	--	--	--	--	3.0	1.4
Q9HB90_C377	--	--	--	1.0	--	--	--	--	1.1	--
P50851_C2675	1.0	--	1.0	--	--	--	--	--	1.9	--
Q9NYL2_C22	--	--	--	1.2	--	--	--	--	1.5	--
Q5T1V6_C414	1.3	--	1.0	0.9	1.8	--	--	1.4	1.2	1.0
Q9HB90_C358	--	--	--	1.1	--	--	--	--	1.3	--
P16455_C145	--	--	--	--	--	--	--	1.0	1.9	--
Q9Y5T5_C205	--	--	--	--	1.9	--	--	0.9	0.8	--
O00541_C272	1.2	--	1.0	1.0	--	--	--	--	--	--
Q02556_C306	1.2	--	--	--	--	--	--	0.9	--	1.4
Q15910_C503	--	--	--	--	--	--	--	--	--	--
Q96RU2_C171	--	--	--	3.3	--	--	--	--	--	--
Q16877_C159	--	--	--	--	--	--	--	1.2	--	--
P04150_C302	0.9	--	0.9	--	--	--	--	--	--	--
Q96JH7_C219	--	--	--	--	1.8	--	--	1.9	--	--
P48200_C137	1.1	--	0.9	--	--	--	--	1.1	1.2	--
O00622_C39	--	--	--	0.8	--	--	--	--	--	--
Q5T1V6_C453	--	--	--	--	--	--	--	--	--	--
P51617_C608	--	--	--	--	2.4	--	--	--	1.1	--
P42575_C370	--	--	0.7	--	--	--	--	2.7	--	1.4
P09086_C346	--	--	--	--	--	--	--	--	--	--
Q09472_C1738	--	--	--	--	--	--	--	--	--	--
Q01201_C109	--	--	--	--	--	--	--	--	--	--
Q70CQ2_C741	--	--	1.1	--	4.1	--	--	--	--	--
P41226_C599	1.2	--	--	--	--	--	--	--	--	--
P14598_C378	--	--	--	--	--	--	--	--	--	--
Q9C0C9_C375	--	--	--	--	--	--	--	--	--	--
O00622_C134	--	--	--	--	--	--	--	--	--	--
O00541_C361	--	--	--	--	--	--	--	--	--	--
P43403_C117	--	--	--	--	--	--	--	--	--	--
Q96FA3_C282	--	--	--	--	7.0	--	--	--	--	--
Q9UPT9_C44	--	--	--	--	--	--	--	--	--	--
Q9Y4C1_C251	--	--	--	--	--	--	--	--	--	--
Q70CQ2_C1090	--	--	--	--	--	--	--	--	--	--
O00622_C70	--	--	--	--	--	--	--	--	--	--
P04150_C622	--	--	--	--	--	--	--	--	--	--

Table 1H

Identifier	42_500 uM_in vitro_r amos	43_500 uM_in vitro_2 31	43_500 uM_in vitro_r amos	44_500 uM_in vitro_2 31	45_500 uM_in vitro_2 31	46_500 uM_in vitro_2 31	47_500 uM_in vitro_2 31	48_500 uM_in vitro_2 31	49_500 uM_in vitro_2 31	50_500 uM_in vitro_2 31
Q99873_C109	0.9	1.4	1.2	0.8	2.2	0.8	0.9	1.0	2.0	1.0
P24752_C119	0.8	1.8	1.1	0.8	12.4	0.8	1.0	1.3	4.9	2.1
P09211_C48	0.7	4.1	2.2	0.9	19.7	1.0	2.5	1.1	--	2.3
O14980_C34	0.9	1.1	--	0.8	1.8	1.0	1.0	1.1	1.9	1.2
P24752_C196	0.8	2.4	2.3	0.8	20.0	0.9	1.1	1.3	4.7	3.9
Q15084_C55	0.9	20.0	20.0	1.0	--	0.8	0.9	1.0	12.8	1.3
P24752_C413	0.7	15.9	--	0.8	20.0	1.1	0.9	1.1	2.6	9.2
P63244_C182	0.8	0.9	0.9	--	0.9	0.8	1.1	1.1	--	0.8
P24752_C126	0.8	--	20.0	0.8	20.0	0.9	1.0	--	20.0	20.0

WO 2017/070611

PCT/US2016/058308

Identifier	42_500 uM_in vitro_r amos	43_500 uM_in vitro_2 31	43_500 uM_in vitro_r amos	44_500 uM_in vitro_2 31	45_500 uM_in vitro_2 31	46_500 uM_in vitro_2 31	47_500 uM_in vitro_2 31	48_500 uM_in vitro_2 31	49_500 uM_in vitro_2 31	50_500 uM_in vitro_2 31
Q15084_C190	--	20.0	--	1.0	--	0.9	1.0	--	7.7	--
Q8TAQ2_C145	1.0	20.0	--	1.0	--	1.0	1.5	1.1	--	2.8
P68036_C86	0.8	1.1	1.7	--	--	1.2	2.0	--	1.9	--
P15374_C95	0.9	--	20.0	0.8	--	1.0	1.0	--	--	1.3
Q16763_C118	0.8	--	1.0	0.8	--	1.0	--	--	--	0.8
Q16822_C306	0.8	3.3	--	--	3.3	0.9	1.0	--	--	2.7
O14980_C528	--	4.6	--	--	--	0.7	2.2	--	--	--
O00170_C122	0.8	--	2.1	--	--	--	--	--	4.8	2.3
O75874_C269	--	0.8	--	0.8	20.0	0.6	0.9	1.0	--	20.0
O75362_C286	1.0	1.9	1.2	--	20.0	0.9	--	--	1.0	--
P40763_C259	--	--	--	--	--	1.2	--	--	--	--
Q9Y3Z3_C522	--	2.2	--	0.8	--	--	--	1.4	3.0	--
P16455_C150	0.8	--	20.0	--	--	--	--	--	--	--
Q96GG9_C115	0.9	--	--	--	--	--	--	--	--	--
P00813_C75	0.8	--	1.2	--	--	--	--	--	--	--
O14933_C98	1.0	--	--	2.0	--	--	--	--	--	1.5
Q14790_C360	--	--	--	--	--	--	--	--	--	--
Q15306_C194	1.2	--	3.0	--	--	--	--	--	--	--
Q6L8Q7_C108	--	--	--	0.8	--	--	--	--	--	1.8
P48735_C308	--	0.8	--	--	1.0	0.2	1.2	1.1	--	2.2
Q86UV5_C39	--	1.5	--	--	--	--	--	--	--	--
P50851_C1704	--	--	--	--	--	--	--	--	4.0	--
O94953_C694	0.9	20.0	--	--	--	--	--	--	--	--
P19447_C342	--	--	--	--	--	1.0	--	--	--	--
Q00535_C157	--	--	1.2	--	--	--	--	--	--	--
Q9UPT9_C171	1.0	--	4.5	--	--	--	--	--	--	--
Q9HB90_C377	--	--	--	0.7	--	--	1.6	1.4	--	--
P50851_C2675	--	--	--	--	--	--	--	--	--	--
Q9NYL2_C22	--	2.9	--	0.8	20.0	--	--	--	--	--
Q5T1V6_C414	--	--	--	0.8	--	--	--	--	--	--
Q9HB90_C358	--	--	--	--	--	1.1	--	--	--	--
P16455_C145	0.9	--	--	--	--	--	--	--	--	--
Q9Y5T5_C205	1.5	--	--	--	--	1.0	--	--	--	--
O00541_C272	0.9	--	--	--	--	--	--	--	--	--
Q02556_C306	0.9	--	2.7	--	--	--	--	--	--	--
Q15910_C503	--	20.0	--	--	--	--	--	--	--	--
Q96RU2_C171	--	--	--	--	--	0.9	1.5	--	--	--
Q16877_C159	--	--	--	--	--	--	--	--	--	--
P04150_C302	--	1.5	--	--	--	--	--	--	--	--
Q96JH7_C219	--	--	--	--	--	--	--	--	--	--
P48200_C137	--	--	--	--	--	--	--	--	--	--
O00622_C39	--	--	--	--	--	1.9	--	--	--	2.3
Q5T1V6_C453	1.0	20.0	--	--	--	--	--	--	--	--
P51617_C608	--	--	--	--	--	--	--	--	--	--
P42575_C370	--	--	--	--	--	--	--	--	--	--
P09086_C346	--	--	--	--	--	--	--	--	--	--
Q09472_C1738	--	2.2	--	--	--	--	--	--	--	--
Q01201_C109	--	--	--	--	--	--	--	--	--	--
Q70CQ2_C741	--	--	--	--	--	--	--	--	--	--
P41226_C599	--	--	--	--	--	--	--	--	--	--
P14598_C378	--	--	1.2	--	--	--	--	--	--	--
Q9C0C9_C375	--	--	--	--	--	--	--	--	--	--
O00622_C134	--	--	--	--	--	--	--	--	--	--

WO 2017/070611

PCT/US2016/058308

Identifier	42_500 uM_in vitro_r amos	43_500 uM_in vitro_2 31	43_500 uM_in vitro_r amos	44_500 uM_in vitro_2 31	45_500 uM_in vitro_2 31	46_500 uM_in vitro_2 31	47_500 uM_in vitro_2 31	48_500 uM_in vitro_2 31	49_500 uM_in vitro_2 31	50_500 uM_in vitro_2 31
O00541_C361	--	--	--	--	--	--	--	--	--	--
P43403_C117	--	--	--	--	--	--	--	--	--	--
Q96FA3_C282	1.2	--	--	--	--	--	--	--	--	--
Q9UPT9_C44	--	--	2.1	--	--	--	--	--	--	--
Q9Y4C1_C251	--	--	--	--	--	1.0	--	--	--	--
Q70CQ2_C1090	--	--	--	--	--	--	--	--	--	--
O00622_C70	--	--	--	--	--	--	--	--	--	--
P04150_C622	--	--	--	--	--	--	--	--	--	--

Table 1I

Identifier	51_500 uM_in vitro_2 31	51_500 uM_in vitro_r amos	52_500 uM_in vitro_2 31	52_500 uM_in vitro_r amos	53_500 uM_in vitro_2 31	53_500 uM_in vitro_r amos	54_500 uM_in vitro_2 31	55_500 uM_in vitro_2 31	56_500 uM_in vitro_2 31	56_500 uM_in vitro_r amos
Q99873_C109	0.9	9.5	0.9	1.3	0.9	1.1	1.2	1.0	1.0	1.1
P24752_C119	0.8	0.9	1.4	0.8	0.9	0.7	1.0	0.9	1.0	1.2
P09211_C48	1.1	1.0	1.0	1.9	2.3	1.2	1.3	1.4	1.9	1.2
O14980_C34	0.9	1.0	1.0	1.1	0.9	1.0	1.1	0.9	0.9	0.9
P24752_C196	1.1	--	1.3	1.5	1.2	0.9	1.1	1.0	1.1	1.0
Q15084_C55	2.5	2.7	1.1	1.8	0.9	1.0	1.2	0.9	1.0	1.1
P24752_C413	1.0	1.0	1.9	1.5	1.0	--	1.1	1.0	1.0	0.9
P63244_C182	0.7	0.9	1.0	--	0.9	0.8	--	1.0	1.1	--
P24752_C126	1.0	1.2	1.3	--	--	--	1.1	0.9	0.9	--
Q15084_C190	--	5.6	1.0	1.9	1.0	1.0	1.2	--	0.9	1.1
Q8TAQ2_C145	--	1.1	0.9	1.8	--	1.2	--	--	1.4	1.2
P68036_C86	--	1.0	0.8	1.3	1.0	1.0	--	1.0	1.5	1.3
P15374_C95	--	0.9	1.2	--	--	0.7	--	--	1.0	1.1
Q16763_C118	--	0.9	1.2	--	1.0	0.8	--	--	1.2	1.1
Q16822_C306	--	0.8	1.3	1.6	1.3	--	--	--	1.4	--
O14980_C528	--	0.9	1.3	--	--	0.8	--	--	20.0	--
O00170_C122	1.4	--	--	1.4	--	0.9	--	--	0.8	0.9
O75874_C269	--	--	--	--	--	--	1.1	--	0.9	--
O75362_C286	0.9	20.0	1.0	--	1.1	--	--	--	1.0	--
P40763_C259	--	--	2.2	--	--	1.9	1.3	--	--	--
Q9Y3Z3_C522	--	--	1.2	--	--	0.7	--	--	1.0	1.0
P16455_C150	--	--	--	2.2	--	1.1	--	--	--	1.4
Q96GG9_C115	--	0.8	--	1.0	1.0	0.9	1.1	--	--	1.0
P00813_C75	--	1.0	--	1.2	--	--	--	--	--	0.9
O14933_C98	--	--	1.4	1.4	--	1.0	--	--	--	1.1
Q14790_C360	--	--	--	--	--	--	1.2	--	1.0	--
Q15306_C194	--	1.4	--	1.4	--	1.0	--	--	--	1.5
Q6L8Q7_C108	--	--	1.4	6.8	--	1.1	--	--	--	--
P48735_C308	--	--	1.1	--	0.5	--	--	--	1.0	--
Q86UV5_C39	--	1.8	0.9	--	--	--	--	--	1.0	--
P50851_C1704	--	1.1	--	1.1	--	--	--	--	--	--
O94953_C694	--	1.4	--	1.4	--	--	--	--	--	--
P19447_C342	1.4	--	1.0	--	--	--	--	--	1.2	--
Q00535_C157	--	1.0	--	--	1.2	0.8	--	--	1.1	--
Q9UPT9_C171	--	--	--	--	--	--	--	--	--	1.9
Q9HB90_C377	0.6	--	--	--	--	1.1	--	--	1.2	--
P50851_C2675	--	0.9	--	--	--	--	--	--	--	--
Q9NYL2_C22	--	--	--	--	--	--	1.1	1.2	1.1	--

Identifier	51_500 uM_in vitro_2 31	51_500 uM_in vitro_r amos	52_500 uM_in vitro_2 31	52_500 uM_in vitro_r amos	53_500 uM_in vitro_2 31	53_500 uM_in vitro_r amos	54_500 uM_in vitro_2 31	55_500 uM_in vitro_2 31	56_500 uM_in vitro_2 31	56_500 uM_in vitro_r amos
Q5T1V6_C414	--	--	--	--	0.8	--	--	--	1.0	1.0
Q9HB90_C358	--	--	--	--	--	0.6	--	--	1.0	1.0
P16455_C145	--	20.0	--	--	--	1.0	--	--	--	--
Q9Y5T5_C205	1.7	3.2	--	--	1.1	--	--	--	1.1	0.9
O00541_C272	--	--	--	--	--	--	--	1.0	1.1	--
Q02556_C306	--	1.5	--	--	--	--	--	--	--	--
Q15910_C503	--	--	0.9	--	--	--	--	--	--	--
Q96RU2_C171	--	--	--	--	--	--	--	--	--	--
Q16877_C159	--	1.1	--	--	--	--	--	--	--	--
P04150_C302	--	--	--	1.8	--	--	--	--	--	--
Q96JH7_C219	--	2.3	--	1.6	--	--	--	--	--	1.0
P48200_C137	--	--	--	1.3	--	--	--	--	--	--
O00622_C39	--	--	--	--	--	--	--	--	--	--
Q5T1V6_C453	--	1.8	--	--	--	--	--	--	--	--
P51617_C608	--	1.1	--	17.6	--	--	--	--	--	--
P42575_C370	--	--	--	--	--	--	--	--	--	--
P09086_C346	--	--	--	--	--	--	--	--	--	--
Q09472_C1738	--	--	1.1	--	--	--	--	--	--	--
Q01201_C109	--	1.6	--	--	--	--	--	--	--	--
Q70CQ2_C741	--	0.9	--	--	--	--	--	--	--	--
P41226_C599	--	--	--	--	--	--	--	--	--	--
P14598_C378	--	--	--	--	--	--	--	--	--	--
Q9C0C9_C375	--	0.8	--	--	--	--	--	--	--	--
O00622_C134	--	--	--	--	--	--	--	--	--	--
O00541_C361	--	--	--	--	--	--	--	--	--	--
P43403_C117	--	--	--	--	--	--	--	--	--	--
Q96FA3_C282	--	--	--	20.0	--	--	--	--	--	--
Q9UPT9_C44	--	--	--	--	--	--	--	--	--	--
Q9Y4C1_C251	--	--	--	--	--	--	--	--	--	--
Q70CQ2_C1090	--	--	--	--	--	--	--	--	--	--
O00622_C70	--	--	--	--	--	--	--	--	--	--
P04150_C622	--	--	--	--	--	--	--	--	--	--

[00449] Table 2 illustrates a list of liganded cysteines and their reactivity profiles with the fragment electrophile library from isoTOP-ABPP experiments performed *in situ*. Table 2 further shows the accession number (or the protein identifier) of the protein.

Table 2A

Identifier	Protein	SEQ ID NO:	2_200 uM_in situ_231	4_100 uM_in situ_231	8_200 uM_in situ_231	9_200 uM_in situ_231
P04406_C152	GAPDH Glyceraldehyde-3-phosphate dehydrogenase	16	20.0	1.6	1.1	1.7
P61978_C132	HNRNPK Heterogeneous nuclear ribonucleoprotein K	768	5.0	3.8	4.2	1.4
Q13526_C113	PIN1 Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1	19	20.0	3.4	2.6	3.3
P24752_C119	ACAT1 Acetyl-CoA acetyltransferase, mitochondrial	22	2.6	3.0	5.5	3.8
P24752_C413	ACAT1 Acetyl-CoA acetyltransferase, mitochondrial	56	20.0	4.5	20.0	19.5

Identifier	Protein	SEQ ID NO:	2_200 uM_in situ_231	4_100 uM_in situ_231	8_200 uM_in situ_231	9_200 uM_in situ_231
Q9NUY8_C283	TBC1D23 TBC1 domain family member 23	101	4.1	2.1	4.9	2.1
P13667_C206	PDIA4 Protein disulfide-isomerase A4	36	2.6	6.8	15.2	1.1
P12268_C140	IMPDH2 Inosine-5-monophosphate dehydrogenase 2	45	1.3	1.0	0.9	1.2
Q15365_C194	PCBP1 Poly(rC)-binding protein 1	29	5.1	1.4	1.4	10.1
Q9NVC6_C649	MED17 Mediator of RNA polymerase II transcription subunit 17	211	7.7	1.7	3.2	2.2
P42166_C561	TMPO Lamina-associated polypeptide 2, isoform alpha	88	8.3	20.0	4.6	3.7
Q9Y696_C35	CLIC4 Chloride intracellular channel protein 4	21	1.7	20.0	4.7	2.6
P10599_C32	TXN Thioredoxin	34	7.9	3.3	5.9	20.0
P31943_C267	HNRNPH1 Heterogeneous nuclear ribonucleoprotein H	769	3.8	6.0	5.3	2.1
Q86SX6_C67	GLRX5 Glutaredoxin-related protein 5, mitochondrial	26	1.1	1.4	1.0	5.1
P15121_C299	AKR1B1 Aldose reductase	48	0.9	1.0	1.7	1.4
P52597_C267	HNRNPF Heterogeneous nuclear ribonucleoprotein F	108	4.2	20.0	--	3.2
Q9ULV4_C420	CORO1C Coronin-1C	770	3.0	1.3	4.6	2.0
P62888_C92	RPL30 60S ribosomal protein L30	100	1.0	2.7	1.5	1.3
Q9NQR4_C153	NIT2 Omega-amidase NIT2	47	13.1	0.7	20.0	20.0
P42765_C92	ACAA2 3-ketoacyl-CoA thiolase, mitochondrial	71	20.0	3.7	20.0	4.3
Q15084_C55	PDIA6 Protein disulfide-isomerase A6	51	3.0	4.9	4.4	2.1
Q96HE7_C241	ERO1L ERO1-like protein alpha	61	2.4	3.0	3.7	--
Q99439_C164	CNN2 Calponin-2	41	--	1.5	1.1	4.7
P25205_C119	MCM3 DNA replication licensing factor MCM3	74	3.5	4.3	3.5	1.9
Q9NS86_C187	LANCL2 LanC-like protein 2	203	--	1.6	6.1	2.5
Q15233_C145	NONO Non-POU domain-containing octamer-binding protein	72	--	1.6	12.0	1.4
Q9BRA2_C43	TXNDC17 Thioredoxin domain-containing protein 17	62	15.5	3.5	20.0	20.0
P35611_C68	ADD1 Alpha-adducin	771	5.8	4.3	2.4	4.3
O75521_C380	ECI2 Enoyl-CoA delta isomerase 2, mitochondrial	155	0.6	1.6	0.8	0.7
Q9BXW7_C392	CECR5 Cat eye syndrome critical region protein 5	772	3.6	2.2	3.2	1.5
P30101_C406	PDIA3 Protein disulfide-isomerase A3	773	3.2	9.0	20.0	1.4

Identifier	Protein	SEQ ID NO:	2_200 uM_in situ_231	4_100 uM_in situ_231	8_200 uM_in situ_231	9_200 uM_in situ_231
Q96AB3_C114	ISOC2 Isochorismatase domain-containing protein 2, mitochondria	159	1.6	0.8	4.7	--
P13667_C555	PDIA4 Protein disulfide-isomerase A4	774	3.2	9.0	20.0	1.4
Q09161_C44	NCBP1 Nuclear cap-binding protein subunit 1	102	2.1	1.8	2.0	--
P78417_C32	GSTO1 Glutathione S-transferase omega-1	32	--	19.9	20.0	--
Q9ULW0_C536	TPX2 Targeting protein for Xklp2	437	20.0	20.0	20.0	5.7
Q9NRG0_C55	CHRAC1 Chromatin accessibility complex protein 1	252	20.0	20.0	7.4	2.9
Q96T76_C848	MMS19 MMS19 nucleotide excision repair protein homolog	114	3.9	20.0	1.7	20.0
Q8TAQ2_C145	SMARCC2 SWI/SNF complex subunit SMARCC2	775	4.6	3.1	--	--
Q9BVC5_C10	C2orf49 Ashwin	168	2.7	3.3	4.5	--
Q7Z2W4_C645	ZC3HAV1 Zinc finger CCCH-type antiviral protein 1	776	--	4.4	5.1	3.0
Q9BQ69_C186	MACROD1 O-acetyl-ADP-ribose deacetylase MACROD1	777	4.8	1.8	5.0	1.2
Q16831_C162	UPP1 Uridine phosphorylase 1	364	1.2	1.0	20.0	7.7
P30101_C57	PDIA3 Protein disulfide-isomerase A3	133	2.2	20.0	20.0	--
P12268_C331	IMPDH2 Inosine-5-monophosphate dehydrogenase 2	144	12.1	4.6	--	1.3
O95571_C170	ETHE1 Protein ETHE1, mitochondrial	176	3.8	2.3	6.8	2.0
O00299_C24	CLIC1 Chloride intracellular channel protein 1	69	9.5	20.0	10.3	3.0
O14879_C343	IFIT3 Interferon-induced protein with tetratricopeptide	308	7.6	6.6	2.1	17.2
Q96CM8_C64	ACSF2 Acyl-CoA synthetase family member 2, mitochondrial	194	20.0	20.0	20.0	--
P51946_C244	CCNH Cyclin-H	778	1.9	5.3	--	--
P49588_C773	AARS Alanine--tRNA ligase, cytoplasmic	122	--	1.3	1.5	1.5
Q96RN5_C618	MED15 Mediator of RNA polymerase II transcription subunit	171	6.4	4.5	20.0	--
O15294_C758	OGT UDP-N-acetylglucosamine--peptide N-acetylglucosamine	104	2.6	6.4	2.4	2.1
P46734_C207	MAP2K3 Dual specificity mitogen-activated protein kinase	184	0.8	0.6	0.8	3.3
Q96S55_C272	WRNIP1 ATPase WRNIP1	215	2.2	3.7	1.2	4.5
O95229_C54	ZWINT ZW10 interactor	779	3.4	12.4	3.1	3.7
O60610_C1227	DIAPH1 Protein diaphanous homolog 1	780	2.2	1.6	--	1.5
Q13428_C38	TCOF1 Treacle protein	150	3.3	2.8	2.0	--

Identifier	Protein	SEQ ID NO:	2_200 uM_in situ_231	4_100 uM_in situ_231	8_200 uM_in situ_231	9_200 uM_in situ_231
Q9Y277_C65	VDAC3 Voltage-dependent anion-selective channel protein	512	3.3	4.7	6.2	--
P57764_C268	GSDMD Gasdermin-D	110	20.0	20.0	8.4	--
Q9Y3A3_C134	MOB4 MOB-like protein phocein	121	--	1.2	1.6	20.0
Q02252_C317	ALDH6A1 Methylmalonate-semialdehyde dehydrogenase	265	10.3	14.4	--	--
Q9NYL9_C132	TMOD3 Tropomodulin-3	130	2.2	0.8	0.9	7.6
P83731_C6	RPL24 60S ribosomal protein L24	147	2.4	0.7	--	4.3
O95336_C32	PGLS 6-phosphogluconolactonase	55	14.9	--	20.0	6.7
Q13155_C291	AIMP2 Aminoacyl tRNA synthase complex-interacting multif	83	20.0	20.0	2.4	1.9
Q13418_C346	ILK Integrin-linked protein kinase	149	1.4	0.8	--	2.7
A6NDU8_C179	C5orf51 UPF0600 protein C5orf51	298	4.2	3.0	--	20.0
Q9UKF6_C498	CPSF3 Cleavage and polyadenylation specificity factor su	118	6.2	3.5	20.0	--
Q96F86_C413	EDC3 Enhancer of mRNA-decapping protein 3	141	20.0	20.0	--	--
P42224_C492	STAT1 Signal transducer and activator of transcription 1	228	16.9	--	--	3.1
P11216_C326	PYGB Glycogen phosphorylase, brain form	366	--	1.7	20.0	2.6
P21980_C277	TGM2 Protein-glutamine gamma-glutamyltransferase 2	356	0.7	0.5	0.7	2.3
Q9HAV7_C124	GRPEL1 GrpE protein homolog 1, mitochondrial	206	3.2	1.7	2.2	4.7
P24752_C126	ACAT1 Acetyl-CoA acetyltransferase, mitochondrial	89	9.1	2.8	4.5	14.1
Q9NQ88_C161	TIGAR Fructose-2,6-bisphosphatase TIGAR	117	6.0	--	--	2.6
Q13155_C23	AIMP2 Aminoacyl tRNA synthase complex-interacting multifunctional protein 2	132	--	2.8	2.7	--
Q9NQW6_C712	ANLN Actin-binding protein anillin	281	--	4.1	14.1	3.1
P51649_C340	ALDH5A1 Succinate-semialdehyde dehydrogenase, mitochondria	189	1.9	1.0	2.7	--
Q15021_C439	NCAPD2 Condensin complex subunit 1	139	--	2.2	0.6	6.1
Q5T0N5_C69	FNBP1L Formin-binding protein 1-like	418	20.0	20.0	3.6	--
P38606_C138	ATP6V1A V-type proton ATPase catalytic subunit A	201	--	20.0	20.0	20.0
Q9HCC0_C216	MCCC2 Methylcrotonoyl-CoA carboxylase beta chain, mitoch	363	4.6	5.9	3.9	--
Q9NQC3_C1101	RTN4 Reticulon-4	247	12.0	20.0	--	20.0
P35754_C23	GLRX Glutaredoxin-1	142	--	20.0	--	--

Identifier	Protein	SEQ ID NO:	2_200 uM_in situ_231	4_100 uM_in situ_231	8_200 uM_in situ_231	9_200 uM_in situ_231
Q99757_C90	TXN2 Thioredoxin, mitochondrial	208	10.1	3.6	4.5	20.0
Q9Y3D0_C93	FAM96B Mitotic spindle-associated MMXD complex subunit MIP18	179	20.0	20.0	--	--
Q9UMS0_C213	NFU1 NFU1 iron-sulfur cluster scaffold homolog, mitochondrial	394	7.4	4.4	--	--
Q9NXV6_C516	CDKN2AIP CDKN2A-interacting protein	255	20.0	--	--	6.6
Q96RS6_C376	NUDCD1 NudC domain-containing protein 1	79	--	1.6	--	2.2
Q14997_C1840	PSME4 Proteasome activator complex subunit 4	257	2.8	20.0	20.0	--
P50570_C27	DNM2 Dynamin-2	73	4.6	2.2	--	--
Q86YH6_C71	PDSS2 Decaprenyl-diphosphate synthase subunit 2	401	20.0	20.0	--	--
Q99497_C106	PARK7 Protein DJ-1	109	--	0.9	--	--
Q9UJW0_C258	DCTN4 Dynactin subunit 4	103	4.0	--	20.0	--
Q9BUH6_C180	C9orf142 Uncharacterized protein C9orf142	348	--	1.2	6.6	2.8
P24752_C196	ACAT1 Acetyl-CoA acetyltransferase, mitochondrial	33	20.0	3.5	--	5.8
Q13162_C51	PRDX4 Peroxiredoxin-4	781	4.6	1.2	1.4	3.3
Q9BTA9_C553	WAC WW domain-containing adapter protein with coiled-coil	153	17.6	17.8	9.6	--
P48643_C253	CCT5 T-complex protein 1 subunit epsilon	126	--	0.8	0.8	--
O75362_C286	ZNF217 Zinc finger protein 217	268	8.9	--	--	--
O60825_C158	PFKFB2 6-phosphofructo-2-kinase/fructose-2,6-bisphosphata	272	0.7	--	--	--
Q8NBS9_C350	TXNDC5 Thioredoxin domain-containing protein 5	136	--	2.4	--	1.7
Q9NYL2_C22	MLTK Mitogen-activated protein kinase kinase kinase MLTK	430	5.5	20.0	1.4	10.3
P27707_C9	DCK Deoxycytidine kinase	93	--	1.4	--	4.9
Q93009_C223	USP7 Ubiquitin carboxyl-terminal hydrolase 7	782	--	6.8	20.0	--
O14929_C101	HAT1 Histone acetyltransferase type B catalytic subunit	154	--	--	--	9.0
Q9UPQ0_C140	LIMCH1 LIM and calponin homology domains-containing protein	783	20.0	--	20.0	--
Q96NY7_C487	CLIC6 Chloride intracellular channel protein 6	447	--	20.0	5.1	3.6
Q9NQ88_C114	TIGAR Fructose-2,6-bisphosphatase TIGAR	143	2.8	--	--	1.9
Q14790_C360	CASP8 Caspase-8	335	20.0	20.0	--	--
P04183_C230	TK1 Thymidine kinase, cytosolic	784	1.9	0.8	--	10.0

Identifier	Protein	SEQ ID NO:	2_200 uM_in situ_231	4_100 uM_in situ_231	8_200 uM_in situ_231	9_200 uM_in situ_231
P68366_C54	TUBA4A Tubulin alpha-4A chain	137	9.7	0.8	0.9	--
Q13428_C1298	TCOF1 Treacle protein	785	4.7	5.6	20.0	--
Q5MNZ6_C63	WDR45L WD repeat domain phosphoinositide-interacting protein	434	--	2.0	20.0	--
O14980_C528	XPO1 Exportin-1	786	20.0	1.9	--	--
Q86W42_C35	THOC6 THO complex subunit 6 homolog	217	--	--	--	20.0
Q9Y6G9_C51	DYNC1LI1 Cytoplasmic dynein 1 light intermediate chain 1	107	--	3.5	4.1	--
Q9NY27_C22	PPP4R2 Serine/threonine-protein phosphatase 4 regulatory	282	20.0	20.0	--	--
Q8NFH5_C255	NUP35 Nucleoporin NUP53	386	5.1	3.9	--	--
Q9Y676_C128	MRPS18B 28S ribosomal protein S18b, mitochondrial	495	10.2	2.3	1.0	--
P35658_C728	NUP214 Nuclear pore complex protein Nup214	311	3.9	1.7	--	--
Q9NTX5_C133	ECHDC1 Ethylmalonyl-CoA decarboxylase	532	1.5	1.7	2.0	--
Q15118_C71	PDK1	623	20.0	20.0	20.0	--
Q00765_C18	REEP5 Receptor expression-enhancing protein 5	344	--	20.0	2.7	--
P22307_C71	SCP2 Non-specific lipid-transfer protein	787	3.2	2.7	20.0	--
O75521_C312	ECI2 Enoyl-CoA delta isomerase 2, mitochondrial	788	0.7	1.0	0.5	0.5
P49189_C288	ALDH9A1 4-trimethylaminobutyraldehyde dehydrogenase	236	--	--	--	20.0
Q5T440_C170	IBA57 Putative transferase CAF17, mitochondrial	608	14.0	1.5	--	--
Q15084_C190	PDIA6 Protein disulfide-isomerase A6	96	--	20.0	5.2	--
Q96C19_C172	EFHD2 EF-hand domain-containing protein D2	246	--	5.3	--	2.8
P22061_C102	PCMT1 Protein-L-isoaspartate(D-aspartate) O-methyltransferase	163	--	2.8	4.5	1.2
Q9NP73_C86	ALG13 UDP-N-acetylglucosamine transferase subunit ALG13	528	20.0	20.0	--	--
Q9BRF8_C54	CPPED1 Calcineurin-like phosphoesterase domain-containing	239	--	20.0	--	2.4
Q6ICB0_C108	DESI1 Desumoylating isopeptidase 1	197	20.0	--	20.0	--
P29590_C389	PML Protein PML	304	15.6	--	--	--
P07858_C211	CTSB Cathepsin B	323	--	2.3	--	5.0
Q9NX18_C83	SDHAF2 Succinate dehydrogenase assembly factor 2, mitochondrial	658	3.1	3.0	--	--
P46109_C249	CRKL Crk-like protein VPCAYDK K.RVPC*AYDK.T	99	20.0	--	--	6.7

Identifier	Protein	SEQ ID NO:	2_200 uM_in situ_231	4_100 uM_in situ_231	8_200 uM_in situ_231	9_200 uM_in situ_231
P45984_C177	MAPK9 Mitogen-activated protein kinase 9	302	--	12.0	20.0	--
P19447_C342	ERCC3 TFIIH basal transcription factor complex helicase	402	--	--	12.4	--
P42166_C341	TMPO Lamina-associated polypeptide 2, isoform alpha	175	10.4	--	4.4	--
Q8N1F7_C522	NUP93 Nuclear pore complex protein Nup93	275	--	--	20.0	--
Q86UY8_C276	NT5DC3 5-nucleotidase domain-containing protein 3	789	3.8	--	--	--
Q8WWI1_C228	LMO7 LIM domain only protein 7	504	--	--	--	--
Q9NWA0_C139	MED9 Mediator of RNA polymerase II transcription subunit	654	4.3	3.4	--	--
P09110_C381	ACAA1 3-ketoacyl-CoA thiolase, peroxisomal	205	--	--	0.8	--
Q2NL82_C126	TSR1 Pre-rRNA-processing protein TSR1 homolog	233	20.0	13.8	--	--
Q5JPI3_C308	C3orf38 Uncharacterized protein C3orf38	544	2.9	--	--	--
P23919_C163	DTYMK Thymidylate kinase	173	--	0.4	0.6	--
Q96EB1_C218	ELP4 Elongator complex protein 4	466	--	20.0	--	8.4
Q96FX7_C209	TRMT61A tRNA (adenine(58)-N(1))-methyltransferase catalytic	313	20.0	2.4	--	--
O14933_C98	UBE2L6 Ubiquitin/ISG15-conjugating enzyme E2 L6	331	--	3.0	--	--
Q29RF7_C242	PDS5A Sister chromatid cohesion protein PDS5 homolog A	232	--	3.0	1.0	--
Q96T76_C819	MMS19 MMS19 nucleotide excision repair protein homolog	181	--	--	--	--
P23919_C117	DTYMK Thymidylate kinase	267	--	--	0.4	7.3
Q15149_C4574	PLEC Plectin	336	--	4.2	--	0.8
Q96RP9_C153	GFM1 Elongation factor G, mitochondrial	524	--	1.0	--	--
P04818_C199	TYMS Thymidylate synthase	406	20.0	20.0	--	--
P27708_C73	CAD CAD protein	191	3.1	--	20.0	--
P55265_C1224	ADAR Double-stranded RNA-specific adenosine deaminase	359	9.1	--	--	--
Q9Y3D2_C105	MSRB2 Methionine-R-sulfoxide reductase B2, mitochondrial	403	--	20.0	20.0	--
O00244_C12	ATOX1 Copper transport protein ATOX1	365	--	--	--	4.7
Q8WV74_C207	NUDT8 Nucleoside diphosphate-linked moiety X motif 8, mitochondrial	468	20.0	20.0	20.0	--
Q9NRW3_C130	APOBEC3C Probable DNA dC- dU-editing enzyme APOBEC-3C	358	20.0	20.0	--	--
P24468_C326	NR2F2 COUP transcription factor 2	790	20.0	20.0	20.0	--

Identifier	Protein	SEQ ID NO:	2_200 uM_in situ_231	4_100 uM_in situ_231	8_200 uM_in situ_231	9_200 uM_in situ_231
P42166_ C684	TMPO Lamina-associated polypeptide 2, isoform alpha	312	--	--	4.6	--
Q96EY5_ C231	FAM125A Multivesicular body subunit 12A	540	4.1	1.7	--	1.9
P14635_ C238	CCNB1 G2/mitotic-specific cyclin-B1	448	--	--	--	3.7
Q8NDH3_ C81	NPEPL1 Probable aminopeptidase NPEPL1	791	20.0	20.0	20.0	--
Q9P0J1_ C149	PDP1	276	20.0	--	--	--
Q96P48_ C900	ARAP1 Arf-GAP with Rho-GAP domain, ANK repeat and PH domain	433	5.5	--	--	--
Q96HE7_ C37	ERO1L ERO1-like protein alpha	347	--	5.4	--	20.0
Q07065_ C100	CKAP4 Cytoskeleton-associated protein 4	733	--	--	--	15.4
Q9BRJ7_ C88	NUDT16L1 Protein syndesmos	432	20.0	20.0	--	--
O75439_ C265	PMPCB Mitochondrial-processing peptidase subunit beta	320	--	--	--	--
O43175_ C369	PHGDH D-3-phosphoglycerate dehydrogenase	248	20.0	--	--	--
Q9UNI6_ C265	DUSP12 Dual specificity protein phosphatase 12	241	--	0.8	--	--
Q06203_ C100	PPAT Amidophosphoribosyltransferase	188	--	1.7	--	--
A0AVT1_ C347	UBA6 Ubiquitin-like modifier-activating enzyme 6	158	--	20.0	--	3.3
Q86X76_ C203	NIT1 Nitrilase homolog 1	471	--	0.7	--	--
Q6XZF7_ C691	DNMBP Dynamin-binding protein	353	--	1.2	--	3.4
Q15398_ C129	DLGAP5 Disks large-associated protein 5	167	20.0	--	--	--
O75717_ C773	WDHD1 WD repeat and HMG-box DNA-binding protein 1	289	--	--	4.2	--
Q01433_ C107	AMPD2 AMP deaminase 2	259	4.4	2.3	6.2	3.1
Q8WVV9_ C464	HNRPLL Heterogeneous nuclear ribonucleoprotein L-like	487	--	--	--	--
O14733_ C131	MAP2K7 Dual specificity mitogen-activated protein kinase	427	--	--	--	--
Q14137_ C404	BOP1 Ribosome biogenesis protein BOP1	535	20.0	1.2	--	--
Q96RU2_ C171	USP28 Ubiquitin carboxyl-terminal hydrolase 28	569	--	20.0	1.2	--
Q9Y679_ C391	AUP1 Ancient ubiquitous protein 1	564	--	20.0	20.0	--
P51610_ C1872	HCFC1 Host cell factor 1	270	4.1	--	--	--
P22307_ C307	SCP2 Non-specific lipid-transfer protein	541	20.0	20.0	20.0	--
Q9BTE3_ C325	MCMBP Mini-chromosome maintenance complex-binding protein	792	--	--	--	--

Identifier	Protein	SEQ ID NO:	2_200 uM_in situ_231	4_100 uM_in situ_231	8_200 uM_in situ_231	9_200 uM_in situ_231
Q9HA64_C24	FN3KRP Ketosamine-3-kinase	106	6.3	20.0	--	20.0
Q5TFE4_C119	NT5DC1 5-nucleotidase domain-containing protein 1	75	--	--	--	20.0
Q96N67_C2125	DOCK7 Dedicator of cytokinesis protein 7	408	--	4.1	5.4	--
P52948_C1312	NUP98 Nuclear pore complex protein Nup98-Nup96	575	--	--	--	--
Q5UIP0_C2298	RIF1 Telomere-associated protein RIF1	607	--	20.0	--	--
P51812_C436	RPS6KA3 Ribosomal protein S6 kinase alpha-3	662	--	--	--	2.1
Q92616_C1692	GCN1L1 Translational activator GCN1	174	1.4	--	--	--
Q15345_C297	LRRC41 Leucine-rich repeat-containing protein 41	536	--	--	--	--
Q9NPH0_C267	ACP6 Lysophosphatidic acid phosphatase type 6	329	--	20.0	20.0	--
P04183_C66	TK1 Thymidine kinase, cytosolic	138	--	--	2.3	7.0
P42166_C629	TMPO Lamina-associated polypeptide 2, isoform alpha	793	4.4	--	1.7	--
Q15013_C124	MAD2L1BP MAD2L1-binding protein	794	--	13.0	5.4	--
Q9Y5Y2_C72	NUBP2 Cytosolic Fe-S cluster assembly factor NUBP2	795	5.3	--	--	4.0
O15446_C86	CD3EAP DNA-directed RNA polymerase I subunit RPA34	796	3.9	--	--	--
Q13630_C116	TSTA3 GDP-L-fucose synthase	797	20.0	--	--	1.9
Q8IYQ7_C324	THNSL1 Threonine synthase-like 1	628	--	--	20.0	1.6
P05091_C319	ALDH2 Aldehyde dehydrogenase, mitochondrial	318	--	20.0	--	--
Q29RF7_C532	PDS5A Sister chromatid cohesion protein PDS5 homolog A	261	--	20.0	--	--
Q9Y570_C381	PPME1 Protein phosphatase methylesterase 1	578	12.6	1.1	--	20.0
Q14980_C961	NUMA1 Nuclear mitotic apparatus protein 1	429	--	--	--	--
P53384_C235	NUBP1 Cytosolic Fe-S cluster assembly factor NUBP1	552	--	0.6	0.6	6.8
Q15003_C418	NCAPH Condensin complex subunit 2	258	20.0	20.0	--	--
P53634_C258	CTSC Dipeptidyl peptidase 1	798	0.9	1.2	--	--
Q8NFF5_C499	FLAD1 FAD synthase	750	6.2	--	--	--
Q9ULA0_C144	DNPEP Aspartyl aminopeptidase	134	--	--	--	2.1
P22307_C94	SCP2 Non-specific lipid-transfer protein	559	--	20.0	--	--
O15294_C620	OGT UDP-N-acetylglucosamine--peptide N-acetylglucosamine	279	--	20.0	--	--

Identifier	Protein	SEQ ID NO:	2_200 uM_in situ_231	4_100 uM_in situ_231	8_200 uM_in situ_231	9_200 uM_in situ_231
Q9Y5S2_C1517	CDC42BPB Serine/threonine-protein kinase MRCK beta	554	20.0	20.0	--	--
Q8TD19_C623	NEK9 Serine/threonine-protein kinase Nek9	393	--	--	--	--
Q8N2W9_C326	PIAS4 E3 SUMO-protein ligase PIAS4	625	--	--	--	--
Q13158_C98	FADD Protein FADD	799	20.0	--	--	--
Q9UKX7_C151	NUP50 Nuclear pore complex protein Nup50	234	6.1	--	--	--
Q6PCB5_C280	RSBN1L Round spermatid basic protein 1-like protein	519	--	--	--	--
P10398_C597	ARAF Serine/threonine-protein kinase A-Raf	505	--	--	--	5.0
Q9UL40_C68	ZNF346 Zinc finger protein 346	706	--	--	--	--
P46013_C903	MKI67 Antigen KI-67	595	20.0	--	--	--
Q16667_C39	CDKN3 Cyclin-dependent kinase inhibitor 3	482	--	0.6	--	--
O75150_C890	RNF40 E3 ubiquitin-protein ligase BRE1B	405	--	--	--	--
Q00610_C870	CLTC Clathrin heavy chain 1	113	--	2.1	--	2.8
Q9Y5T5_C205	USP16 Ubiquitin carboxyl-terminal hydrolase 16	474	7.9	20.0	--	--
O95881_C66	TXNDC12 Thioredoxin domain-containing protein 12	342	--	--	--	--
Q7Z5K2_C160	WAPAL Wings apart-like protein homolog	800	20.0	--	--	--
P42166_C518	TMPO Lamina-associated polypeptide 2, isoform alpha	801	--	--	--	3.4
Q9Y2S7_C143	POLDIP2 Polymerase delta-interacting protein 2	574	--	--	--	2.3
E2QRD5_C183	C15orf38-AP3S2 Protein C15orf38-AP3S2	581	--	--	--	4.9
O95833_C22	CLIC3 Chloride intracellular channel protein 3	531	--	--	--	20.0
O94953_C694	KDM4B Lysine-specific demethylase 4B	395	20.0	--	--	--
O00541_C272	PES1 Pescadillo homolog	511	--	--	--	5.0
Q9NXJ5_C149	PGPEP1 Pyroglutamyl-peptidase 1	587	--	20.0	--	--
Q8N5L8_C131	RPP25L Ribonuclease P protein subunit p25-like protein	670	--	--	--	--
Q8IZ73_C246	RPUSD2 RNA pseudouridylate synthase domain-containing protein	441	--	--	--	--
Q99798_C385	ACO2 Aconitate hydratase, mitochondrial	685	20.0	1.0	--	--
Q9GZR2_C382	REXO4 RNA exonuclease 4	621	--	--	--	--
Q13613_C117	MTMR1 Myotubularin-related protein 1	717	--	--	--	--

Identifier	Protein	SEQ ID NO:	2_200 uM_in situ_231	4_100 uM_in situ_231	8_200 uM_in situ_231	9_200 uM_in situ_231
Q9NUI1_C22	DECR2 Peroxisomal 2,4-dienoyl-CoA reductase	698	--	--	--	--
Q02556_C306	IRF8 Interferon regulatory factor 8	513	--	--	--	--
Q9UPT9_C171	USP22 Ubiquitin carboxyl-terminal hydrolase 22	802	--	--	--	--
Q8N999_C302	C12orf29 Uncharacterized protein C12orf29	484	--	--	--	--
Q8IU81_C363	IRF2BP1 Interferon regulatory factor 2-binding protein 1	803	--	--	--	--
Q9C0I1_C152	MTMR12 Myotubularin-related protein 12	671	--	--	--	--
Q9P2X3_C195	IMPACT Protein IMPACT	678	--	20.0	--	--
Q6QNY0_C168	BLOC1S3 Biogenesis of lysosome-related organelles complex	411	--	--	--	--
Q15796_C81	SMAD2 Mothers against decapentaplegic homolog 2	561	20.0	--	--	--
Q9NZB2_C531	FAM120A Constitutive coactivator of PPAR-gamma-like protei	492	--	--	--	--
Q9HB90_C377	RRAGC Ras-related GTP-binding protein C	417	3.3	--	--	4.3
Q9BR61_C267	ACBD6 Acyl-CoA-binding domain-containing protein 6	472	--	--	--	--
P16455_C145	MGMT Methylated-DNA--protein-cysteine methyltransferase	470	--	--	--	--
Q86UV5_C39	USP48 Ubiquitin carboxyl-terminal hydrolase 48	381	20.0	--	--	--
A2A288_C367	ZC3H12D Probable ribonuclease ZC3H12D	515	--	--	--	--
Q8NEC7_C140	GSTCD Glutathione S-transferase C-terminal domain-containing protein	602	--	--	--	--
Q6PJG6_C673	BRAT1 BRCA1-associated ATM activator 1	695	--	--	--	--
Q13232_C158	NME3 Nucleoside diphosphate kinase 3	653	--	--	--	2.7
Q86X76_C165	NIT1 Nitrilase homolog 1	345	--	0.9	--	--
P42695_C541	NCAPD3 Condensin-2 complex subunit D3	573	--	--	--	--
P41226_C599	UBA7 Ubiquitin-like modifier-activating enzyme 7	702	--	--	--	--
Q99986_C50	VRK1 Serine/threonine-protein kinase VRK1	497	--	--	--	--
Q8WU_M4_C90	PDCD6IP Programmed cell death 6-interacting protein	527	--	--	--	--
P29590_C213	PML Protein PML	477	--	--	--	--
Q9P0K7_C973	RAI14 Ankycorbin	638	--	8.7	--	--
P53992_C78	SEC24C Protein transport protein Sec24C	498	--	--	5.4	--
Q13867_C73	BLMH Bleomycin hydrolase	431	--	--	--	3.9

WO 2017/070611

PCT/US2016/058308

Identifier	Protein	SEQ ID NO:	2_200 uM_in situ_231	4_100 uM_in situ_231	8_200 uM_in situ_231	9_200 uM_in situ_231
Q8ND24_C655	RNF214 RING finger protein 214	451	6.5	--	--	4.0
Q96EK4_C48	THAP11 THAP domain-containing protein 11	538	7.9	--	--	--
Q96IV0_C309	NGLY1 Peptide-N(4)-(N-acetyl-beta-glucosaminy)asparagin	660	--	--	--	--
Q5T1V6_C414	DDX59 Probable ATP-dependent RNA helicase DDX59	439	--	--	--	--
Q9UHQ1_C99	NARF Nuclear prelamin A recognition factor	740	--	--	--	--
O43396_C34	TXNL1 Thioredoxin-like protein 1	310	--	--	--	--
Q8IV53_C174	DENND1C DENN domain-containing protein 1C	804	--	--	--	--
Q8N9T8_C673	KRI1 Protein KRI1 homolog	563	--	--	5.3	--

Table 2B

Identifier	9_200uM_in situ_231	10_200 uM_in situ_231	10_200 uM_in situ_231	11_100 uM_in situ_231	12_200 uM_in situ_231	13_200 uM_in situ_231	13_200 uM_in situ_231	14_200 uM_in situ_231	14_200 uM_in situ_231	21_200 uM_in situ_231
P04406_C152	1.1	1.0	0.9	0.8	0.9	0.9	0.9	0.9	0.6	1.3
P61978_C132	1.4	5.6	1.6	0.9	7.5	1.1	1.1	1.3	0.9	1.4
Q13526_C113	1.3	2.0	1.3	0.7	0.6	0.9	1.0	0.8	1.0	0.7
P24752_C119	4.3	4.5	1.5	1.2	4.1	2.1	2.8	1.0	--	1.6
P24752_C413	20.0	7.8	2.2	1.2	9.9	3.5	20.0	0.9	0.8	1.5
Q9NUY8_C283	1.5	2.0	--	1.1	2.9	1.1	1.1	0.8	0.6	0.8
P13667_C206	1.4	3.8	--	4.1	16.6	1.1	--	1.0	0.7	0.8
P12268_C140	1.1	0.7	0.9	0.6	1.2	0.7	--	10.3	--	0.6
Q15365_C194	1.5	1.6	2.0	0.4	1.3	0.9	1.4	1.0	--	1.1
Q9NVC6_C649	1.3	2.3	2.1	0.8	4.0	1.1	1.0	1.0	1.4	0.8
P42166_C561	2.1	17.8	--	0.6	16.1	1.2	1.0	2.9	--	1.4
Q9Y696_C35	1.6	2.7	1.6	1.9	20.0	2.3	3.2	0.7	--	1.1
P10599_C32	2.1	3.8	--	13.0	20.0	7.9	4.1	3.1	--	3.9
P31943_C267	1.5	4.8	2.5	0.9	5.0	1.1	1.2	1.3	--	1.5
Q86SX6_C67	14.0	1.3	1.1	1.4	1.6	10.3	12.4	3.1	--	1.4
P15121_C299	1.5	20.0	--	0.7	0.7	--	0.8	0.9	--	2.7
P52597_C267	1.6	2.5	2.4	1.0	6.8	1.2	1.3	1.5	--	1.5
Q9ULV4_C420	1.4	3.4	--	0.9	2.2	1.1	1.1	1.4	--	1.3
P62888_C92	2.4	1.3	4.5	0.8	1.1	0.8	2.0	1.0	--	0.9
Q9NQR4_C153	20.0	20.0	--	3.7	0.9	6.2	12.2	0.8	--	0.9
P42765_C92	1.2	3.7	1.1	--	20.0	--	2.2	1.2	0.8	1.3
Q15084_C55	1.4	3.1	--	4.3	15.4	--	1.5	1.0	--	0.9
Q96HE7_C241	1.8	2.0	--	1.8	16.4	1.5	1.3	1.1	--	1.6
Q99439_C164	1.3	1.2	1.3	0.6	0.9	0.9	1.3	0.9	--	0.9
P25205_C119	1.8	2.3	--	0.7	1.0	1.1	1.3	1.2	--	1.1
Q9NS86_C187	1.7	6.3	--	0.8	2.0	1.0	1.1	1.0	3.3	0.9
Q15233_C145	1.4	--	2.2	1.2	2.4	1.0	1.1	1.0	--	--
Q9BRA2_C43	5.4	20.0	--	17.7	20.0	20.0	5.2	1.4	--	16.0
P35611_C68	2.3	1.9	--	0.9	3.6	1.1	1.0	1.6	--	0.8
O75521_C380	1.1	4.8	--	0.8	1.9	1.0	--	1.1	--	1.8
Q9BXW7_C392	--	20.0	--	1.1	5.1	2.0	--	1.6	1.1	1.6

WO 2017/070611

PCT/US2016/058308

P30101_C406	1.6	3.3	1.4	--	20.0	--	2.0	1.1	--	--
Q96AB3_C114	3.7	20.0	--	0.9	1.6	2.4	--	1.0	--	4.9
P13667_C555	1.6	3.3	1.4	--	20.0	--	2.0	1.1	--	--
Q09161_C44	1.4	12.4	--	1.1	4.7	1.0	--	2.0	--	1.3
P78417_C32	20.0	20.0	20.0	--	20.0	20.0	20.0	1.7	--	--
Q9ULW0_C536	1.6	17.6	--	1.4	20.0	1.5	--	1.5	--	0.9
Q9NRG0_C55	2.5	2.7	--	1.1	20.0	1.2	--	1.7	--	--
Q96T76_C848	1.8	2.5	--	--	20.0	0.6	1.5	5.0	20.0	0.7
Q8TAQ2_C145	2.4	--	20.0	7.6	--	1.1	--	1.3	--	1.6
Q9BVC5_C10	1.4	3.1	--	1.1	2.9	1.6	--	1.2	--	1.5
Q7Z2W4_C645	1.3	3.8	--	1.0	3.3	1.3	0.9	1.4	--	0.9
Q9BQ69_C186	--	2.4	--	1.0	2.7	1.4	--	1.1	--	1.4
Q16831_C162	--	7.3	--	0.8	0.9	1.0	--	0.6	--	1.6
P30101_C57	1.5	3.3	--	2.2	20.0	--	1.6	1.0	--	1.0
P12268_C331	2.6	--	2.1	1.0	--	--	1.4	1.7	--	0.8
O95571_C170	1.8	9.0	--	1.2	6.3	1.2	1.9	1.0	--	1.7
O00299_C24	1.4	5.0	--	2.5	20.0	--	1.8	0.8	--	0.7
O14879_C343	--	4.9	--	0.4	3.5	1.0	--	1.5	--	0.8
Q96CM8_C64	20.0	20.0	--	1.5	17.9	1.4	--	1.5	--	2.0
P51946_C244	2.0	1.7	--	1.3	1.3	1.0	1.2	1.6	--	1.4
P49588_C773	2.0	2.1	1.7	0.9	0.9	0.8	1.1	1.0	--	1.1
Q96RN5_C618	--	--	--	1.0	20.0	1.0	1.4	1.6	--	1.0
O15294_C758	--	2.3	--	1.0	20.0	1.0	--	1.6	2.9	1.1
P46734_C207	1.8	0.8	--	0.6	0.7	0.8	--	13.8	--	0.9
Q96S55_C272	2.5	1.2	--	--	2.4	0.7	--	1.4	--	1.3
O95229_C54	--	2.3	--	1.2	20.0	0.8	1.4	5.0	20.0	0.8
O60610_C1227	--	1.5	1.4	0.9	0.8	0.7	0.9	--	--	20.0
Q13428_C38	1.7	4.2	--	1.1	4.7	1.6	--	3.6	--	1.5
Q9Y277_C65	0.8	3.3	2.5	1.4	3.9	1.0	--	1.7	--	2.9
P57764_C268	4.9	--	2.2	0.7	--	--	1.5	1.6	--	0.7
Q9Y3A3_C134	20.0	1.9	--	1.8	1.4	--	1.9	1.1	--	--
Q02252_C317	2.6	1.5	0.7	0.9	3.2	1.3	--	1.1	--	1.5
Q9NYL9_C132	--	--	--	0.5	0.6	--	--	0.5	--	1.4
P83731_C6	1.3	0.4	2.0	0.5	0.3	0.7	1.0	--	--	1.0
O95336_C32	2.6	--	--	0.9	20.0	1.6	2.2	3.2	--	0.9
Q13155_C291	--	1.7	--	0.9	1.1	--	--	1.6	--	0.8
Q13418_C346	--	1.1	8.3	0.6	0.6	0.7	2.1	0.8	--	0.5
A6NDU8_C179	3.7	1.3	--	0.7	--	0.8	--	1.9	--	0.9
Q9UKF6_C498	1.7	20.0	2.8	1.3	4.1	--	1.4	1.9	--	1.4
Q96F86_C413	4.0	20.0	20.0	20.0	--	1.6	2.0	1.0	--	0.8
P42224_C492	--	20.0	--	0.7	1.0	1.0	--	20.0	--	0.7
P11216_C326	--	8.6	--	0.8	3.9	1.3	--	0.8	--	0.9
P21980_C277	--	0.6	--	0.6	0.3	0.8	--	20.0	--	0.7
Q9HAV7_C124	3.8	1.0	--	0.9	1.1	--	--	1.3	--	1.7
P24752_C126	20.0	5.9	--	1.2	5.8	--	--	0.9	--	1.4
Q9NQ88_C161	2.0	4.6	2.1	0.9	--	1.1	1.5	20.0	--	1.0
Q13155_C23	1.6	2.3	--	1.2	1.6	0.9	--	1.1	--	0.9
Q9NQW6_C712	--	11.2	--	0.9	16.7	1.3	--	2.1	--	--
P51649_C340	1.5	13.4	20.0	--	20.0	1.1	1.3	0.7	--	1.2
Q15021_C439	--	5.2	--	0.9	6.7	--	--	5.9	--	0.7
Q5T0N5_C69	--	2.0	--	1.3	20.0	1.0	--	1.5	--	0.8
P38606_C138	--	20.0	--	9.2	20.0	20.0	--	1.9	--	20.0
Q9HCC0_C216	3.3	2.3	--	1.9	20.0	--	1.8	1.0	--	2.2
Q9NQC3_C1101	20.0	20.0	--	1.1	20.0	--	--	20.0	--	2.1
P35754_C23	5.7	--	13.0	20.0	20.0	--	--	0.7	--	--
Q99757_C90	--	--	--	2.7	5.0	--	--	3.2	--	4.3
Q9Y3D0_C93	2.1	2.1	--	0.9	--	--	1.4	2.7	--	0.5

WO 2017/070611

PCT/US2016/058308

Q9UMS0_C213	--	20.0	--	--	20.0	1.6	--	3.5	--	--
Q9NXV6_C516	13.8	7.7	--	1.0	3.3	2.4	1.6	1.4	--	0.8
Q96RS6_C376	1.7	5.7	--	0.8	1.1	0.7	1.2	0.9	1.8	1.4
Q14997_C1840	--	20.0	--	--	20.0	0.9	1.6	--	20.0	1.0
P50570_C27	1.4	3.1	--	0.7	0.7	--	1.0	1.0	--	0.8
Q86YH6_C71	6.6	20.0	--	1.0	20.0	1.2	1.8	20.0	--	2.1
Q99497_C106	2.6	1.7	--	1.3	3.0	--	3.3	0.8	--	--
Q9UJW0_C258	20.0	20.0	20.0	20.0	20.0	--	1.9	3.0	--	--
Q9BUH6_C180	--	3.8	--	--	1.4	1.1	--	1.3	--	--
P24752_C196	6.8	5.1	2.3	1.1	4.6	--	--	1.0	--	1.5
Q13162_C51	--	2.0	--	0.9	1.0	--	--	0.9	--	1.7
Q9BTA9_C553	2.1	20.0	--	1.7	--	1.4	--	1.7	--	2.8
P48643_C253	1.1	--	0.9	0.6	5.1	--	--	0.8	--	--
O75362_C286	1.8	2.6	--	1.0	20.0	1.1	20.0	1.4	--	1.1
O60825_C158	1.8	0.7	1.2	--	0.8	--	0.9	1.5	14.2	1.4
Q8NBS9_C350	1.8	--	1.4	3.4	9.4	1.3	--	1.0	--	--
Q9NYL2_C22	--	1.2	--	0.7	0.7	1.3	--	0.9	--	--
P27707_C9	1.5	1.9	1.7	0.5	1.4	1.0	1.3	--	--	1.3
Q93009_C223	20.0	2.5	--	1.0	3.1	--	1.3	1.1	--	1.8
O14929_C101	--	20.0	--	0.8	20.0	0.9	1.2	1.6	1.9	1.2
Q9UPQ0_C140	--	3.9	--	1.0	1.9	1.1	--	4.6	--	0.8
Q96NY7_C487	--	3.2	--	3.1	20.0	1.9	--	0.8	--	1.2
Q9NQ88_C114	1.4	2.1	--	0.8	--	1.1	1.2	2.2	5.1	0.9
Q14790_C360	3.2	--	5.8	--	--	1.2	--	--	--	0.9
P04183_C230	2.0	1.3	--	--	--	0.7	--	--	--	0.7
P68366_C54	1.8	7.9	2.8	0.3	--	--	1.3	--	5.1	0.8
Q13428_C1298	2.5	5.9	--	0.8	20.0	--	--	3.0	--	--
Q5MNZ6_C63	11.2	3.6	--	0.9	20.0	1.1	--	1.2	--	0.8
O14980_C528	1.8	1.3	0.7	0.7	--	1.1	1.1	20.0	--	0.8
Q86W42_C35	2.7	--	--	1.3	--	--	--	0.9	--	--
Q9Y6G9_C51	--	4.9	--	--	6.5	--	1.5	2.1	4.3	0.6
Q9NY27_C22	1.8	--	2.1	2.3	20.0	--	1.5	1.3	--	--
Q8NFB5_C255	1.9	3.5	--	1.3	12.4	1.5	--	1.0	--	1.8
Q9Y676_C128	1.3	2.9	--	0.8	1.4	--	--	1.2	--	1.5
P35658_C728	4.2	20.0	--	0.8	--	1.0	--	1.2	--	1.3
Q9NTX5_C133	--	1.6	--	1.0	1.2	1.1	--	1.0	--	1.1
Q15118_C71	--	18.0	2.7	0.6	--	1.6	--	4.4	--	1.7
Q00765_C18	--	4.2	--	0.8	20.0	1.2	--	20.0	--	0.7
P22307_C71	--	8.3	--	2.0	3.7	4.9	--	1.1	--	8.3
O75521_C312	--	5.0	--	--	4.2	0.7	--	1.4	--	1.6
P49189_C288	20.0	20.0	--	--	1.8	--	12.0	0.9	--	1.0
Q5T440_C170	2.4	5.7	--	--	--	1.2	--	1.2	--	1.6
Q15084_C190	1.6	3.7	--	--	20.0	--	--	1.1	--	--
Q96C19_C172	--	2.4	--	0.9	0.9	0.7	--	1.0	--	--
P22061_C102	1.3	1.6	--	0.8	5.8	0.9	--	1.0	--	--
Q9NP73_C86	--	--	--	0.7	--	0.9	1.4	1.1	--	--
Q9BRF8_C54	1.9	3.8	--	1.0	1.2	1.4	--	0.9	--	--
Q6ICB0_C108	--	--	--	0.7	--	--	4.0	--	--	0.6
P29590_C389	--	7.9	--	0.5	--	1.3	--	1.2	--	20.0
P07858_C211	--	7.4	--	1.1	--	1.6	--	1.0	--	2.6
Q9NX18_C83	--	4.4	--	1.1	20.0	--	--	2.2	--	1.3
P46109_C249	4.0	20.0	--	--	20.0	--	1.8	1.3	--	0.9
P45984_C177	3.8	20.0	--	0.9	2.1	--	--	1.5	--	--
P19447_C342	2.7	20.0	--	0.7	6.7	1.1	--	2.5	--	2.0
P42166_C341	1.5	--	--	0.8	--	1.3	1.4	2.3	--	1.7
Q8N1F7_C522	3.7	--	--	--	20.0	--	2.2	3.2	20.0	--
Q86UY8_C276	--	5.8	--	1.4	20.0	1.5	--	2.8	--	1.8

WO 2017/070611

PCT/US2016/058308

Q8WWI1_C228	1.4	17.2	--	1.3	20.0	2.0	1.0	--	--	1.1
Q9NWA0_C139	--	2.9	--	--	20.0	1.1	--	--	--	1.5
P09110_C381	0.9	--	--	0.8	4.5	--	--	1.0	--	--
Q2NL82_C126	--	4.9	--	--	20.0	--	--	0.8	--	1.5
Q5JPI3_C308	4.0	--	--	1.0	--	1.1	1.7	3.2	--	0.9
P23919_C163	2.4	--	--	0.2	--	--	--	0.9	--	--
Q96EB1_C218	3.7	1.1	--	--	--	1.2	1.7	--	--	0.7
Q96FX7_C209	16.1	--	--	--	20.0	2.1	6.8	2.7	--	2.0
O14933_C98	2.8	20.0	3.2	--	--	--	--	2.5	--	0.7
Q29RF7_C242	--	--	--	0.9	12.9	--	--	1.8	--	0.9
Q96T76_C819	4.0	--	3.0	0.8	--	--	--	3.4	--	--
P23919_C117	1.4	0.4	--	--	0.2	0.6	--	1.2	--	0.9
Q15149_C4574	--	2.2	--	1.1	--	1.4	--	1.2	--	1.0
Q96RP9_C153	2.8	14.8	--	--	20.0	--	0.8	1.2	--	1.4
P04818_C199	11.7	1.5	--	--	--	--	--	--	--	--
P27708_C73	2.2	--	--	--	--	--	5.0	20.0	3.7	0.6
P55265_C1224	--	10.7	--	--	--	1.9	--	4.2	--	2.2
Q9Y3D2_C105	--	20.0	--	1.3	20.0	--	--	2.7	--	--
O00244_C12	1.6	2.1	--	0.7	--	1.5	--	0.7	--	0.7
Q8WV74_C207	20.0	--	20.0	--	--	13.1	--	20.0	--	--
Q9NRW3_C130	6.2	--	--	--	--	--	--	--	1.4	1.3
P24468_C326	--	20.0	--	0.9	--	--	--	4.9	--	1.1
P42166_C684	--	5.2	--	0.7	--	1.2	--	3.1	--	--
Q96EY5_C231	2.1	--	--	0.7	--	--	--	--	--	0.7
P14635_C238	2.1	6.7	--	0.7	--	1.3	1.5	2.6	--	--
Q8NDH3_C81	--	5.7	--	--	20.0	--	--	1.4	--	1.6
Q9P0J1_C149	--	18.0	--	1.4	18.5	--	--	1.3	--	--
Q96P48_C900	--	1.8	--	--	1.9	1.2	--	--	--	0.7
Q96HE7_C37	--	--	--	--	--	--	1.4	1.4	--	--
Q07065_C100	--	--	--	1.0	20.0	1.6	--	11.2	--	2.2
Q9BRJ7_C88	4.0	20.0	--	--	--	--	--	--	--	--
O75439_C265	1.7	--	2.1	1.0	--	--	--	20.0	--	--
O43175_C369	2.0	--	2.4	--	--	--	2.0	--	20.0	1.0
Q9UNI6_C265	--	1.2	--	--	1.2	--	--	0.7	--	--
Q06203_C100	2.9	--	20.0	--	--	--	2.0	--	--	--
A0AVT1_C347	2.6	--	20.0	1.6	--	1.3	--	--	--	--
Q86X76_C203	20.0	--	--	0.8	3.9	--	--	0.7	--	0.8
Q6XZF7_C691	10.0	20.0	20.0	--	--	1.6	--	--	--	0.8
Q15398_C129	--	--	--	0.6	--	--	--	--	--	--
O75717_C773	--	2.2	--	1.2	--	1.1	--	1.7	--	--
Q01433_C107	1.4	--	--	0.6	--	--	--	--	--	--
Q8WVV9_C464	--	--	--	0.9	20.0	--	--	2.2	--	1.4
O14733_C131	--	1.8	--	--	1.7	--	1.6	--	20.0	0.7
Q14137_C404	2.1	20.0	--	--	--	1.1	--	--	--	1.1
Q96RU2_C171	--	--	--	2.2	20.0	--	--	1.3	--	--
Q9Y679_C391	--	--	--	3.0	--	--	--	--	--	1.5
P51610_C1872	1.0	1.5	--	0.5	--	1.0	--	1.1	--	--
P22307_C307	--	--	--	2.9	20.0	--	--	0.9	--	20.0
Q9BTE3_C325	5.4	--	--	--	--	--	--	5.7	--	--
Q9HA64_C24	1.6	--	--	--	--	--	--	1.7	--	0.6
Q5TFE4_C119	20.0	--	20.0	--	--	--	1.7	3.8	--	1.2
Q96N67_C2125	--	--	--	--	--	--	--	1.8	--	--
P52948_C1312	--	20.0	--	--	20.0	1.2	--	1.7	--	1.1
Q5UIP0_C2298	--	20.0	--	1.3	--	--	--	20.0	--	0.9
P51812_C436	1.8	--	--	0.5	1.4	1.0	--	0.6	--	--
Q92616_C1692	--	--	--	--	--	--	0.9	20.0	--	--
Q15345_C297	--	--	--	1.1	--	1.6	--	5.8	--	1.1

WO 2017/070611

PCT/US2016/058308

Q9NPH0_C267	2.6	13.4	--	--	--	--	--	1.2	--	--
P04183_C66	1.8	--	--	0.6	--	1.0	--	--	--	--
P42166_C629	--	2.5	--	0.6	--	1.0	--	20.0	--	1.6
Q15013_C124	--	--	--	1.0	--	--	--	1.4	--	--
Q9Y5Y2_C72	1.3	1.3	--	0.5	--	--	1.5	--	--	--
O15446_C86	--	4.7	--	--	3.3	1.3	--	1.0	--	1.7
Q13630_C116	1.6	20.0	--	--	--	--	1.2	--	--	0.8
Q8IYQ7_C324	--	5.3	--	--	--	1.3	--	1.6	--	2.0
P05091_C319	20.0	--	--	--	20.0	--	--	3.3	--	--
Q29RF7_C532	--	20.0	--	--	--	--	2.0	2.5	--	--
Q9Y570_C381	--	4.7	--	--	--	0.9	--	--	--	--
Q14980_C961	--	--	--	--	20.0	2.2	--	--	--	3.7
P53384_C235	--	--	--	0.5	--	--	--	--	--	--
Q15003_C418	2.5	--	--	1.6	--	--	--	--	--	--
P53634_C258	--	5.5	--	--	13.4	1.6	--	1.9	--	2.8
Q8NFF5_C499	12.1	--	--	1.6	--	--	--	20.0	--	--
Q9ULA0_C144	1.3	5.2	--	--	--	--	--	--	--	--
P22307_C94	--	--	--	3.3	20.0	--	--	1.0	--	--
O15294_C620	--	--	12.0	--	--	--	--	1.0	--	--
Q9Y5S2_C1517	--	--	--	1.2	--	--	--	0.9	--	--
Q8TD19_C623	20.0	--	--	--	--	--	20.0	--	--	0.8
Q8N2W9_C326	2.6	4.5	--	--	--	--	20.0	--	--	1.0
Q13158_C98	--	--	--	--	--	1.6	--	--	--	--
Q9UKX7_C151	1.8	--	--	1.0	6.7	--	--	1.8	--	--
Q6PCB5_C280	2.5	--	--	--	20.0	--	0.8	20.0	--	--
P10398_C597	1.3	1.4	--	--	--	--	1.0	--	--	0.7
Q9UL40_C68	20.0	--	--	--	20.0	--	20.0	20.0	--	--
P46013_C903	--	--	--	--	--	--	--	--	--	--
Q16667_C39	4.5	2.3	--	--	--	--	1.2	--	--	0.8
O75150_C890	--	4.2	--	1.1	--	--	--	1.1	--	--
Q00610_C870	--	--	--	0.9	1.7	--	--	20.0	--	--
Q9Y5T5_C205	--	--	--	--	--	--	20.0	--	--	0.9
O95881_C66	20.0	--	--	--	20.0	20.0	--	1.3	--	--
Q7Z5K2_C160	--	20.0	--	1.0	--	--	--	2.4	--	--
P42166_C518	1.3	4.2	--	0.8	--	--	--	--	--	1.4
Q9Y2S7_C143	--	13.9	--	--	--	--	--	1.3	--	--
E2QRD5_C183	--	--	--	0.9	--	--	--	--	--	0.8
O95833_C22	--	--	--	--	--	--	--	0.8	--	1.0
O94953_C694	4.2	--	3.9	--	--	--	1.5	--	--	--
O00541_C272	--	--	--	1.0	--	1.4	--	3.8	--	--
Q9NXJ5_C149	20.0	--	11.0	--	3.9	--	20.0	--	--	--
Q8N5L8_C131	--	--	--	--	--	1.2	--	1.3	--	0.8
Q8IZ73_C246	6.5	--	--	--	--	--	--	--	--	1.8
Q99798_C385	--	0.7	--	--	--	--	--	--	--	2.0
Q9GZR2_C382	1.0	--	--	--	20.0	--	0.9	1.2	--	--
Q13613_C117	4.3	--	--	--	--	--	--	1.0	1.7	0.9
Q9NUI1_C22	--	--	--	2.1	--	--	--	1.1	--	--
Q02556_C306	5.5	--	--	--	--	--	1.5	--	--	--
Q9UPT9_C171	4.4	--	9.3	0.8	--	--	3.2	--	--	--
Q8N999_C302	20.0	--	--	--	--	--	--	--	--	--
Q8IU81_C363	--	5.9	--	--	--	--	--	--	--	--
Q9C0I1_C152	--	--	--	--	20.0	--	6.5	--	--	0.7
Q9P2X3_C195	--	--	--	--	--	--	--	1.6	--	--
Q6QNY0_C168	4.6	--	--	--	--	--	1.2	--	--	--
Q15796_C81	5.6	--	--	--	--	--	--	--	--	1.3
Q9NZB2_C531	--	--	--	--	20.0	--	--	--	--	1.6
Q9HB90_C377	20.0	--	--	--	--	--	--	1.1	--	--

WO 2017/070611

PCT/US2016/058308

Q9BR61_C267	--	--	--	--	--	1.1	--	--	6.9	--
P16455_C145	20.0	--	20.0	--	--	--	--	--	--	--
Q86UV5_C39	--	--	--	--	--	--	--	--	--	1.4
A2A288_C367	5.2	--	--	--	--	--	--	--	--	--
Q8NEC7_C140	--	--	--	--	--	1.8	20.0	--	--	--
Q6PJG6_C673	--	--	--	--	--	--	--	--	--	--
Q13232_C158	--	--	--	--	--	--	--	--	--	--
Q86X76_C165	20.0	--	--	--	--	--	--	0.8	--	--
P42695_C541	--	20.0	6.5	--	--	--	--	--	--	2.3
P41226_C599	20.0	8.0	--	--	--	--	--	--	--	1.0
Q99986_C50	--	--	0.7	--	--	--	--	1.6	--	--
Q8WUM4_C90	5.0	--	--	--	--	--	--	1.0	--	--
P29590_C213	4.5	--	--	--	--	--	--	--	--	--
Q9P0K7_C973	--	--	--	0.8	--	--	--	--	--	--
P53992_C78	--	--	--	--	--	0.7	--	--	--	--
Q13867_C73	--	4.2	--	--	--	--	--	--	--	--
Q8ND24_C655	--	--	--	--	--	--	--	--	--	--
Q96EK4_C48	2.3	--	--	--	--	--	--	--	--	--
Q96IV0_C309	6.8	--	--	--	--	--	--	--	--	--
Q5T1V6_C414	3.7	11.3	--	--	--	--	--	--	--	--
Q9UHQ1_C99	20.0	--	--	--	--	--	20.0	--	--	--
O43396_C34	--	--	--	--	--	--	--	7.3	--	--
Q8IV53_C174	--	--	5.0	--	--	--	--	--	--	--
Q8N9T8_C673	--	--	--	--	--	--	--	--	--	--

Table 2C

Identifier	27_20 0uM_i nsitu_ 231	28_20 0uM_i nsitu_ 231	29_20 0uM_i nsitu_ ramos	31_20 0uM_i nsitu_ 231	31_20 0uM_i nsitu_ ramos	33_20 0uM_i nsitu_ 231	38_20 0uM_i nsitu_ 231	41_20 0uM_i nsitu_ 231	45_20 0uM_i nsitu_ 231	51_20 0uM_i nsitu_ 231	56_20 0uM_i nsitu_ 231
P04406_C152	0.5	0.8	1.2	0.8	1.0	1.1	0.9	0.6	20.0	1.0	0.9
P61978_C132	0.9	1.0	1.2	0.9	1.1	1.3	1.0	0.9	9.3	1.2	0.8
Q13526_C113	0.5	0.7	1.0	0.6	1.1	1.2	1.0	0.6	10.0	1.4	--
P24752_C119	1.6	1.0	1.2	0.6	0.7	2.3	1.2	1.9	12.4	0.7	0.8
P24752_C413	1.9	1.0	1.2	1.0	--	20.0	1.1	0.9	20.0	1.1	0.7
Q9NUY8_C283	0.6	0.7	1.3	0.6	--	1.6	1.0	0.7	2.7	2.1	0.9
P13667_C206	1.0	1.3	1.4	0.8	1.1	2.7	0.9	1.0	1.9	5.3	0.8
P12268_C140	0.5	0.6	1.6	0.6	1.0	1.4	1.2	0.7	1.9	1.0	1.1
Q15365_C194	0.5	--	1.0	0.7	1.6	2.2	2.5	0.3	1.5	0.6	1.2
Q9NVC6_C649	1.3	1.1	1.6	0.8	--	1.2	1.2	0.8	20.0	--	0.8
P42166_C561	0.8	1.1	0.9	1.3	2.5	2.9	20.0	0.7	20.0	1.3	1.1
Q9Y696_C35	1.0	2.7	1.0	0.6	--	8.9	1.0	0.6	1.4	2.6	0.9
P10599_C32	2.5	20.0	1.9	1.2	1.7	20.0	2.1	0.4	5.7	13.9	1.5
P31943_C267	1.1	--	1.4	0.7	0.9	1.3	1.2	1.0	3.6	1.3	0.8
Q86SX6_C67	1.5	--	1.2	--	0.7	1.0	1.0	1.0	9.1	1.0	0.8
P15121_C299	0.5	0.8	1.1	0.8	1.3	0.7	1.0	0.9	1.0	0.9	0.9
P52597_C267	1.0	--	1.2	1.2	1.3	1.6	1.8	0.8	4.1	1.3	0.9
Q9ULV4_C420	0.8	0.7	--	0.8	1.1	1.6	1.4	0.8	3.2	1.1	1.0
P62888_C92	0.6	0.9	2.0	--	--	1.1	3.2	0.8	2.2	0.9	0.9
Q9NQR4_C153	1.3	--	--	0.6	1.4	1.0	1.1	0.7	2.0	0.6	0.8
P42765_C92	1.2	--	0.9	0.8	1.2	1.3	0.9	1.1	15.1	1.1	--
Q15084_C55	1.2	1.4	1.0	0.7	--	3.7	0.8	0.9	2.0	5.0	0.8
Q96HE7_C241	0.9	1.1	1.3	0.7	--	2.3	1.6	0.7	2.1	1.7	0.9
Q99439_C164	--	0.9	1.0	0.7	1.3	1.5	1.5	0.4	1.8	--	1.2
P25205_C119	1.0	--	1.3	0.5	--	1.2	0.9	0.9	6.2	2.6	1.0
Q9NS86_C187	0.4	--	1.7	0.5	--	1.5	2.0	0.6	3.1	2.0	1.2

WO 2017/070611

PCT/US2016/058308

Identifier	27_20 0uM_i nsitu_ 231	28_20 0uM_i nsitu_ 231	29_20 0uM_i nsitu_ ramos	31_20 0uM_i nsitu_ 231	31_20 0uM_i nsitu_ ramos	33_20 0uM_i nsitu_ 231	38_20 0uM_i nsitu_ 231	41_20 0uM_i nsitu_ 231	45_20 0uM_i nsitu_ 231	51_20 0uM_i nsitu_ 231	56_20 0uM_i nsitu_ 231
Q15233_C145	1.2	1.1	--	0.7	0.8	1.3	0.9	1.1	3.7	1.6	0.8
Q9BRA2_C43	20.0	--	2.3	0.6	--	20.0	1.1	0.6	8.0	--	3.9
P35611_C68	0.9	1.7	--	--	--	1.6	1.4	0.9	20.0	1.0	0.8
O75521_C380	1.6	0.8	1.2	--	--	0.7	0.9	1.1	1.2	1.0	1.1
Q9BXW7_C392	1.3	1.0	1.5	--	--	1.1	0.8	1.0	4.9	1.1	0.8
P30101_C406	1.2	--	1.5	0.8	1.1	--	0.8	1.0	2.6	6.0	0.7
Q96AB3_C114	1.0	0.8	1.1	0.7	--	0.9	1.2	0.7	3.1	0.8	--
P13667_C555	1.2	--	1.5	0.8	1.1	--	0.8	1.0	2.6	6.0	0.7
Q09161_C44	0.8	--	1.1	0.6	--	1.1	1.2	0.8	2.5	1.2	1.0
P78417_C32	20.0	20.0	20.0	0.8	2.6	20.0	1.0	0.3	1.9	--	1.3
Q9ULW0_C536	1.1	1.0	1.8	--	--	2.2	1.2	1.2	20.0	1.4	--
Q9NRG0_C55	0.9	--	1.5	--	1.6	1.1	1.0	0.9	20.0	1.1	1.0
Q96T76_C848	0.5	0.8	1.0	20.0	--	--	--	0.7	20.0	--	1.4
Q8TAQ2_C145	2.1	0.9	2.7	0.8	1.1	1.3	1.3	1.2	2.9	1.2	0.8
Q9BVC5_C10	1.1	0.9	1.8	--	--	1.3	0.9	1.0	3.4	1.0	0.8
Q7Z2W4_C645	0.8	--	1.6	--	--	2.1	0.9	0.9	4.6	1.6	0.9
Q9BQ69_C186	2.4	--	1.5	0.5	--	0.9	3.6	1.2	3.3	1.2	0.8
Q16831_C162	0.5	0.7	--	--	--	1.3	1.4	0.8	1.3	1.1	1.1
P30101_C57	1.1	1.6	1.2	0.9	--	--	0.9	1.3	2.6	--	0.7
P12268_C331	0.7	0.8	1.3	1.0	1.1	1.5	1.3	0.8	4.1	--	--
O95571_C170	1.1	--	1.5	--	--	1.2	--	1.1	3.1	--	0.8
O00299_C24	0.7	--	1.1	0.6	--	8.4	1.0	0.6	--	--	0.9
O14879_C343	0.4	0.7	--	0.9	--	2.4	--	0.3	4.0	1.3	1.2
Q96CM8_C64	1.8	--	2.3	0.8	--	2.3	1.1	0.8	20.0	--	0.8
P51946_C244	1.1	--	1.1	--	--	0.9	1.0	1.2	2.4	0.7	1.0
P49588_C773	0.6	--	1.1	--	1.2	--	0.7	0.7	10.1	--	--
Q96RN5_C618	0.8	0.7	1.3	0.9	--	2.0	1.4	--	20.0	2.1	0.8
O15294_C758	1.0	1.0	--	--	--	1.2	--	0.9	2.9	1.1	1.0
P46734_C207	0.8	--	--	--	--	1.2	1.5	0.7	1.5	0.7	1.3
Q96S55_C272	0.9	0.5	--	--	--	1.2	1.3	0.8	2.5	0.9	0.9
O95229_C54	0.5	0.7	1.2	--	--	--	--	0.6	--	1.0	1.0
O60610_C1227	0.5	--	1.6	0.6	1.2	--	0.3	0.6	3.0	--	1.2
Q13428_C38	1.3	1.0	1.5	--	--	--	1.1	1.1	2.8	1.1	0.8
Q9Y277_C65	1.6	1.2	--	--	1.0	--	20.0	1.4	--	--	0.9
P57764_C268	0.6	--	1.1	--	3.1	--	1.8	0.6	20.0	1.8	0.8
Q9Y3A3_C134	--	--	1.2	1.0	1.8	1.2	1.0	0.9	2.3	--	0.9
Q02252_C317	2.0	1.1	1.2	--	--	2.3	2.4	0.8	2.6	--	--
Q9NYL9_C132	0.6	0.7	0.9	0.7	0.7	--	2.6	0.3	0.9	0.8	--
P83731_C6	0.4	--	1.6	--	--	1.6	2.1	0.3	1.2	--	--
O95336_C32	0.5	--	3.0	--	1.7	2.3	2.8	0.6	--	--	0.9
Q13155_C291	0.7	0.9	--	0.8	--	--	1.7	0.7	7.3	1.1	1.1
Q13418_C346	0.6	--	--	0.6	--	1.4	1.0	0.6	--	0.8	--
A6NDU8_C179	0.7	0.8	1.0	--	--	7.4	1.4	0.6	20.0	--	1.2
Q9UKF6_C498	1.5	--	2.3	--	--	1.6	--	1.0	20.0	--	1.0
Q96F86_C413	0.7	--	1.9	--	1.3	3.0	--	--	20.0	20.0	1.2
P42224_C492	--	0.8	3.8	0.8	1.9	1.3	--	0.6	20.0	2.4	1.2
P11216_C326	0.7	--	--	0.5	--	1.6	1.2	0.7	5.3	0.8	--
P21980_C277	0.6	0.4	--	--	--	--	1.4	0.6	1.1	1.0	--
Q9HAV7_C124	0.6	--	1.0	--	--	--	1.0	0.6	2.1	--	0.7
P24752_C126	--	1.2	1.2	0.7	--	20.0	--	0.8	--	--	0.8
Q9NQ88_C161	0.6	--	1.3	1.0	2.6	--	--	--	20.0	--	1.4
Q13155_C23	0.7	--	1.6	--	--	1.4	1.4	--	4.4	1.2	0.9
Q9NQW6_C712	1.0	0.9	--	0.7	--	1.5	1.5	0.7	5.9	--	1.0

WO 2017/070611

PCT/US2016/058308

Identifier	27_20 0uM_i nsitu_ 231	28_20 0uM_i nsitu_ 231	29_20 0uM_i nsitu_ ramos	31_20 0uM_i nsitu_ 231	31_20 0uM_i nsitu_ ramos	33_20 0uM_i nsitu_ 231	38_20 0uM_i nsitu_ 231	41_20 0uM_i nsitu_ 231	45_20 0uM_i nsitu_ 231	51_20 0uM_i nsitu_ 231	56_20 0uM_i nsitu_ 231
P51649_C340	0.9	--	1.0	--	1.0	0.9	--	--	2.4	--	--
Q15021_C439	0.6	0.9	1.4	20.0	--	--	1.1	0.8	--	1.1	1.2
Q5T0N5_C69	0.6	--	--	--	--	2.6	1.5	0.6	5.4	1.7	1.0
P38606_C138	4.6	--	--	--	1.7	20.0	2.1	0.8	--	10.6	1.1
Q9HCC0_C216	1.0	--	--	--	--	1.0	--	1.0	7.0	0.9	0.7
Q9NQC3_C1101	1.4	--	--	12.8	--	2.1	20.0	1.0	--	1.2	0.8
P35754_C23	0.7	1.4	--	0.6	1.1	20.0	0.9	0.5	16.3	--	0.9
Q99757_C90	6.3	--	1.6	1.4	--	--	3.6	0.5	6.7	5.0	--
Q9Y3D0_C93	0.7	0.8	1.0	--	--	--	2.4	0.6	20.0	20.0	--
Q9UMS0_C213	1.3	--	2.0	1.5	--	1.9	3.5	0.9	7.1	1.8	0.7
Q9NXV6_C516	1.4	--	1.1	--	--	1.4	1.8	0.9	--	--	--
Q96RS6_C376	--	--	1.3	0.8	--	--	1.3	0.7	--	--	--
Q14997_C1840	20.0	0.9	--	--	1.2	--	20.0	0.9	--	20.0	--
P50570_C27	0.5	--	1.0	--	--	--	0.9	0.5	15.4	1.1	--
Q86YH6_C71	1.3	--	1.1	--	--	--	--	0.9	20.0	--	1.0
Q99497_C106	0.9	1.0	--	0.8	1.0	--	1.1	0.6	20.0	1.4	--
Q9UJW0_C258	0.8	--	20.0	--	--	--	--	0.7	2.7	1.1	1.3
Q9BUH6_C180	--	1.1	1.5	0.8	--	1.4	1.4	0.6	5.5	--	1.3
P24752_C196	--	--	1.1	--	--	3.4	1.0	0.9	--	--	0.8
Q13162_C51	0.9	--	--	--	1.3	1.7	2.4	--	2.3	--	1.0
Q9BTA9_C553	--	--	1.1	--	--	--	1.1	1.0	19.1	1.2	0.9
P48643_C253	--	0.8	--	0.5	1.0	--	1.1	0.8	1.2	0.8	0.9
O75362_C286	1.0	0.9	1.0	--	--	--	--	1.0	20.0	--	0.9
O60825_C158	--	--	0.9	--	1.3	--	--	0.7	1.3	0.9	1.2
Q8NBS9_C350	1.1	1.1	--	0.8	--	2.6	--	0.9	2.1	--	0.8
Q9NYL2_C22	0.7	--	1.2	--	--	1.5	--	0.6	2.3	--	1.0
P27707_C9	--	--	1.2	--	1.2	1.9	1.8	0.4	--	--	--
Q93009_C223	1.1	1.0	1.0	--	--	--	0.8	--	20.0	1.0	--
O14929_C101	--	--	1.0	--	--	1.3	1.1	0.6	20.0	--	--
Q9UPQ0_C140	0.7	--	--	--	--	2.5	2.5	0.5	20.0	--	1.1
Q96NY7_C487	--	--	--	0.5	--	14.9	--	0.4	1.8	--	1.3
Q9NQ88_C114	0.4	--	0.9	--	--	--	--	0.6	4.6	--	--
Q14790_C360	0.4	--	1.6	0.5	--	3.2	1.0	0.5	20.0	--	1.2
P04183_C230	--	--	1.2	--	2.1	1.6	2.8	0.6	2.0	0.7	--
P68366_C54	0.3	--	1.0	--	1.1	--	2.0	--	--	--	--
Q13428_C1298	1.0	0.9	--	--	2.2	--	2.0	1.0	4.7	--	--
Q5MNZ6_C63	--	1.7	--	--	--	1.6	--	0.8	2.6	--	0.7
O14980_C528	--	--	1.1	--	--	--	2.1	0.9	--	1.1	--
Q86W42_C35	1.3	1.3	--	0.5	--	1.6	0.8	1.1	20.0	1.4	0.7
Q9Y6G9_C51	0.6	0.7	--	--	1.6	1.8	--	0.7	--	--	--
Q9NY27_C22	--	--	1.7	--	--	--	1.0	0.8	20.0	2.2	--
Q8NFH5_C255	--	--	1.6	--	--	--	1.1	0.8	5.5	--	--
Q9Y676_C128	1.0	--	--	--	--	--	1.0	--	2.0	1.0	--
P35658_C728	--	--	1.7	0.7	--	6.4	--	1.1	3.1	--	--
Q9NTX5_C133	1.5	--	--	--	--	1.0	--	--	4.2	--	0.9
Q15118_C71	1.5	--	1.8	--	--	--	1.3	--	20.0	--	--
Q00765_C18	1.1	--	1.4	--	--	1.1	--	--	20.0	--	0.9
P22307_C71	20.0	--	--	--	--	--	--	1.2	16.8	5.6	--
O75521_C312	1.1	--	--	--	--	--	1.1	--	0.7	--	0.9
P49189_C288	--	--	--	--	--	20.0	0.7	0.8	20.0	--	0.9
Q5T440_C170	--	--	1.3	--	--	2.5	--	0.7	3.5	--	1.0
Q15084_C190	1.0	--	1.0	1.0	--	--	1.0	0.9	2.8	--	--
Q96C19_C172	--	--	1.2	1.6	--	--	0.7	0.5	--	1.2	--

WO 2017/070611

PCT/US2016/058308

Identifier	27_20 0uM_i nsitu_ 231	28_20 0uM_i nsitu_ 231	29_20 0uM_i nsitu_ ramos	31_20 0uM_i nsitu_ 231	31_20 0uM_i nsitu_ ramos	33_20 0uM_i nsitu_ 231	38_20 0uM_i nsitu_ 231	41_20 0uM_i nsitu_ 231	45_20 0uM_i nsitu_ 231	51_20 0uM_i nsitu_ 231	56_20 0uM_i nsitu_ 231
P22061_C102	2.2	--	1.2	--	--	--	--	0.7	--	--	--
Q9NP73_C86	0.4	--	2.0	--	--	--	1.3	0.5	20.0	1.3	--
Q9BRF8_C54	--	--	1.5	--	--	1.8	--	0.7	20.0	--	--
Q6ICB0_C108	1.9	--	1.0	1.2	1.3	--	1.1	--	3.1	1.1	--
P29590_C389	0.5	--	--	1.1	--	--	2.2	0.4	3.4	0.9	--
P07858_C211	0.8	--	--	--	--	2.2	1.8	0.4	2.4	--	--
Q9NX18_C83	1.5	--	1.9	--	--	--	--	--	3.7	1.0	0.7
P46109_C249	--	--	4.2	--	--	--	1.6	--	20.0	--	0.8
P45984_C177	0.5	--	1.7	--	--	2.0	--	0.5	--	--	1.1
P19447_C342	--	--	--	--	--	--	1.0	0.7	5.9	--	0.7
P42166_C341	0.6	--	--	--	2.0	--	--	0.5	--	1.2	--
Q8N1F7_C522	--	--	1.7	--	3.6	1.7	1.7	0.8	20.0	--	--
Q86UY8_C276	1.7	--	--	--	--	0.8	--	1.7	2.1	1.2	--
Q8WWI1_C228	0.8	--	1.0	--	--	--	--	0.8	20.0	--	--
Q9NWA0_C139	1.3	--	0.7	--	--	1.6	--	0.8	--	--	1.1
P09110_C381	--	0.8	--	0.7	--	--	0.8	0.9	1.2	--	0.8
Q2NL82_C126	0.4	--	--	--	--	--	0.9	0.8	20.0	1.9	--
Q5JPI3_C308	0.7	--	1.3	--	--	--	--	0.7	--	--	0.7
P23919_C163	0.2	0.4	--	--	1.0	--	5.7	0.2	--	--	0.9
Q96EB1_C218	0.6	--	1.1	--	--	--	1.3	--	20.0	--	--
Q96FX7_C209	--	--	--	--	--	--	1.0	--	20.0	--	0.9
O14933_C98	--	--	2.4	--	1.9	--	1.6	0.5	20.0	--	--
Q29RF7_C242	0.9	1.0	--	--	--	--	1.0	--	--	1.8	1.0
Q96T76_C819	0.5	--	--	20.0	20.0	--	4.6	0.6	20.0	--	0.9
P23919_C117	--	--	1.6	--	--	2.5	4.4	--	--	--	--
Q15149_C4574	1.3	--	--	--	--	1.4	--	0.9	--	--	1.0
Q96RP9_C153	1.6	--	1.2	--	--	--	--	1.1	--	--	--
P04818_C199	0.5	--	2.5	--	--	--	20.0	0.7	20.0	20.0	--
P27708_C73	0.5	--	1.8	--	--	--	--	0.7	--	--	--
P55265_C1224	1.4	--	1.8	--	--	1.6	--	--	20.0	1.3	--
Q9Y3D2_C105	1.5	--	--	--	--	--	1.6	1.2	20.0	--	--
O00244_C12	0.4	--	1.4	--	--	1.8	--	--	--	--	--
Q8WV74_C207	2.2	--	--	--	--	--	--	0.9	20.0	--	--
Q9NRW3_C130	1.3	--	13.3	--	--	--	--	0.9	20.0	--	0.9
P24468_C326	1.2	--	--	--	--	--	--	0.8	20.0	--	--
P42166_C684	--	--	--	--	3.3	--	3.0	--	6.8	1.2	1.1
Q96EY5_C231	0.6	--	1.1	--	--	1.2	1.8	--	--	--	--
P14635_C238	--	--	--	--	--	1.2	1.8	0.5	--	--	--
Q8NDH3_C81	--	--	--	1.0	--	--	2.0	--	--	--	1.0
Q9P0J1_C149	--	--	--	1.0	--	--	--	0.9	20.0	--	0.8
Q96P48_C900	--	--	1.1	1.1	--	1.7	--	0.5	--	--	--
Q96HE7_C37	--	--	2.1	0.6	--	--	3.4	0.4	20.0	--	--
Q07065_C100	--	--	--	--	--	1.7	--	1.0	20.0	--	--
Q9BRJ7_C88	--	--	1.7	--	2.9	--	--	0.4	--	1.6	1.0
O75439_C265	1.6	0.6	--	1.0	1.4	--	--	1.3	--	--	--
O43175_C369	--	--	1.1	20.0	20.0	--	--	--	--	--	--
Q9UNI6_C265	0.6	0.8	--	--	--	--	1.0	1.0	7.5	--	--
Q06203_C100	0.8	--	1.6	0.6	1.2	--	1.2	--	--	--	--
A0AVT1_C347	0.8	--	--	--	1.9	--	--	0.6	--	--	--
Q86X76_C203	--	--	--	--	--	--	--	1.3	20.0	--	0.8
Q6XZF7_C691	--	--	--	2.3	--	2.5	--	--	--	--	--
Q15398_C129	0.5	--	1.2	--	2.4	1.3	1.4	--	--	7.4	1.3
O75717_C773	--	--	--	--	--	--	2.3	1.0	3.6	--	0.9

WO 2017/070611

PCT/US2016/058308

Identifier	27_20 0uM_i nsitu_ 231	28_20 0uM_i nsitu_ 231	29_20 0uM_i nsitu_ ramos	31_20 0uM_i nsitu_ 231	31_20 0uM_i nsitu_ ramos	33_20 0uM_i nsitu_ 231	38_20 0uM_i nsitu_ 231	41_20 0uM_i nsitu_ 231	45_20 0uM_i nsitu_ 231	51_20 0uM_i nsitu_ 231	56_20 0uM_i nsitu_ 231
Q01433_C107	0.4	--	--	0.7	--	--	--	0.5	--	--	--
Q8WVV9_C464	--	--	--	--	--	2.2	1.5	--	20.0	3.2	--
O14733_C131	--	--	1.0	--	--	--	20.0	0.8	--	--	--
Q14137_C404	0.6	--	1.5	--	--	--	--	--	--	--	--
Q96RU2_C171	1.1	1.1	--	--	--	--	0.8	--	--	--	--
Q9Y679_C391	1.2	--	--	--	--	--	--	--	20.0	20.0	0.7
P51610_C1872	--	--	--	--	--	--	1.3	0.5	--	--	--
P22307_C307	--	--	--	--	--	--	--	1.1	--	--	--
Q9BTE3_C325	--	1.0	--	1.9	3.5	--	2.2	0.6	20.0	--	--
Q9HA64_C24	--	--	--	--	--	--	1.7	--	--	--	1.0
Q5TFE4_C119	--	--	--	3.1	--	--	--	0.5	--	--	--
Q96N67_C2125	0.9	--	--	0.8	--	--	1.5	0.7	--	--	0.9
P52948_C1312	1.3	--	--	--	--	1.8	1.3	--	--	--	--
Q5UIP0_C2298	--	--	--	--	--	3.4	--	1.2	20.0	--	--
P51812_C436	--	--	1.3	--	--	--	--	4.5	--	--	--
Q92616_C1692	0.5	--	--	0.6	2.0	1.2	--	0.7	--	--	--
Q15345_C297	--	--	--	0.9	--	1.9	1.4	0.8	--	--	--
Q9NPH0_C267	1.3	--	2.1	--	--	--	--	--	20.0	--	--
P04183_C66	--	--	1.6	0.7	--	--	--	0.4	--	--	--
P42166_C629	--	--	1.6	--	--	--	--	--	--	--	--
Q15013_C124	1.0	1.0	--	--	--	--	--	--	20.0	1.2	--
Q9Y5Y2_C72	--	--	1.1	--	--	--	1.7	--	--	--	--
O15446_C86	--	--	--	--	--	1.6	--	--	4.6	--	--
Q13630_C116	--	--	--	--	--	--	--	--	20.0	--	--
Q8IYQ7_C324	--	--	--	--	--	1.4	--	--	--	--	--
P05091_C319	--	--	10.5	--	--	--	--	0.9	20.0	--	--
Q29RF7_C532	--	--	--	5.1	8.1	--	--	0.8	--	--	--
Q9Y570_C381	--	--	1.2	--	--	--	--	0.2	--	--	--
Q14980_C961	2.4	--	--	1.1	--	4.8	2.6	--	--	--	--
P53384_C235	--	--	--	0.7	--	--	2.0	0.2	--	--	--
Q15003_C418	--	--	1.4	--	--	--	--	--	20.0	--	1.3
P53634_C258	--	--	--	--	--	--	--	--	--	--	--
Q8NFF5_C499	--	--	3.4	--	--	--	--	--	--	--	1.3
Q9ULA0_C144	--	--	1.3	--	--	--	--	0.8	--	--	0.9
P22307_C94	--	20.0	--	--	--	--	--	--	20.0	--	--
O15294_C620	--	--	2.6	1.1	--	--	--	1.0	--	--	--
Q9Y5S2_C1517	--	--	--	--	--	--	--	0.4	20.0	--	--
Q8TD19_C623	--	--	--	0.7	1.4	--	--	0.6	--	--	--
Q8N2W9_C326	--	--	--	--	--	--	0.9	0.8	--	--	--
Q13158_C98	0.6	--	1.3	0.9	--	--	1.5	--	--	--	--
Q9UKX7_C151	--	--	--	--	--	--	--	0.8	--	--	--
Q6PCB5_C280	1.0	--	1.4	--	--	--	--	--	--	--	--
P10398_C597	--	--	1.2	--	--	--	--	--	--	--	--
Q9UL40_C68	1.9	--	2.4	--	--	--	--	--	--	--	--
P46013_C903	--	--	1.6	--	--	--	1.1	1.0	20.0	1.1	--
Q16667_C39	--	--	1.7	--	--	--	--	--	--	--	--
O75150_C890	1.2	--	--	0.7	--	--	--	1.3	--	--	--
Q00610_C870	--	--	--	--	--	--	20.0	--	--	--	--
Q9Y5T5_C205	1.2	--	--	--	--	--	1.5	--	--	--	--
O95881_C66	--	--	--	--	--	--	1.0	1.0	--	--	--
Q7Z5K2_C160	--	--	0.7	--	--	--	--	1.0	--	--	--
P42166_C518	--	--	--	--	--	--	2.2	--	--	--	--
Q9Y2S7_C143	--	--	--	--	--	--	--	1.2	--	--	0.8

Identifier	27_20 0uM_i nsitu_ 231	28_20 0uM_i nsitu_ 231	29_20 0uM_i nsitu_ ramos	31_20 0uM_i nsitu_ 231	31_20 0uM_i nsitu_ ramos	33_20 0uM_i nsitu_ 231	38_20 0uM_i nsitu_ 231	41_20 0uM_i nsitu_ 231	45_20 0uM_i nsitu_ 231	51_20 0uM_i nsitu_ 231	56_20 0uM_i nsitu_ 231
E2QRD5_C183	--	--	--	0.8	--	--	--	0.5	--	--	--
O95833_C22	--	--	--	0.5	--	--	--	--	13.8	--	--
O94953_C694	--	--	1.1	--	--	--	--	--	--	--	--
O00541_C272	--	--	--	--	--	--	--	0.9	--	--	--
Q9NXJ5_C149	--	--	--	--	--	--	--	--	--	--	--
Q8N5L8_C131	--	--	--	--	--	1.7	--	--	7.3	--	--
Q8IZ73_C246	--	--	--	0.4	1.0	--	1.7	--	--	--	--
Q99798_C385	--	--	--	--	--	--	--	0.1	--	--	--
Q9GZR2_C382	--	--	--	--	--	--	--	--	20.0	--	--
Q13613_C117	--	--	1.5	--	--	--	--	--	--	--	--
Q9NUI1_C22	--	--	--	0.7	--	--	--	--	17.0	--	--
Q02556_C306	--	--	1.4	--	2.0	--	--	--	--	--	--
Q9UPT9_C171	--	--	--	--	--	--	--	--	--	--	--
Q8N999_C302	--	0.6	--	--	--	--	--	0.3	20.0	--	--
Q8IU81_C363	--	--	--	0.7	--	--	1.3	--	1.7	--	--
Q9C0I1_C152	--	--	4.6	--	--	--	--	--	--	--	--
Q9P2X3_C195	--	--	--	0.5	--	--	--	--	20.0	--	--
Q6QNY0_C168	--	--	1.6	--	--	--	1.0	--	--	--	--
Q15796_C81	--	--	1.3	--	--	--	--	--	--	--	--
Q9NZZB2_C531	--	--	--	--	--	2.4	--	--	--	--	0.6
Q9HB90_C377	--	--	--	--	--	--	--	--	--	--	--
Q9BR61_C267	--	--	2.5	--	--	--	--	0.5	--	--	--
P16455_C145	--	--	--	--	1.2	--	--	--	--	--	--
Q86UV5_C39	--	--	--	--	1.3	--	--	--	--	--	--
A2A288_C367	--	--	1.3	--	1.9	--	--	--	--	--	--
Q8NEC7_C140	--	--	--	--	--	--	--	--	20.0	--	--
Q6PJG6_C673	--	--	4.6	--	20.0	--	--	--	--	--	0.9
Q13232_C158	--	--	1.5	--	--	--	--	--	8.8	--	--
Q86X76_C165	--	--	--	--	--	--	--	--	--	--	--
P42695_C541	--	--	--	--	--	--	--	--	--	--	--
P41226_C599	--	--	--	--	--	--	--	--	--	--	--
Q99986_C50	--	--	--	--	--	--	--	--	4.1	--	--
Q8WUM4_C90	0.6	--	--	--	--	--	--	--	--	--	--
P29590_C213	0.8	--	--	--	--	--	--	--	--	--	--
Q9P0K7_C973	--	--	--	--	--	--	--	--	--	--	--
P53992_C78	--	--	--	--	--	--	--	--	--	--	--
Q13867_C73	--	--	--	--	--	--	--	--	--	--	--
Q8ND24_C655	--	--	--	--	--	--	--	--	--	--	--
Q96EK4_C48	--	--	--	--	--	--	--	--	--	--	--
Q96IV0_C309	--	--	3.0	--	--	--	--	--	--	--	--
Q5T1V6_C414	--	--	--	--	--	--	--	--	--	--	--
Q9UHQ1_C99	--	--	--	--	--	--	--	--	--	--	--
O43396_C34	--	--	--	--	--	--	--	--	--	--	--
Q8IV53_C174	--	--	--	--	--	--	--	--	--	--	--
Q8N9T8_C673	--	--	--	--	--	--	--	--	--	--	--

[00450] Table 3 illustrates a list of cysteine containing proteins and potential cysteine site of conjugation.

Identifier	Protein Name	Cysteine Location	Protein Class
O00170	AIP AH receptor-interacting protein	C122	Uncategorized

Identifier	Protein Name	Cysteine Location	Protein Class
O00541	PES1 Pescadillo homolog	C272; C361	Uncategorized
O00622	CYR61 Protein CYR61	C39; C70; C134	Uncategorized
O14920	IKBKB Inhibitor of nuclear factor kappa-B kinase subunit	C464	Enzyme
O14933	UBE2L6 Ubiquitin/ISG15-conjugating enzyme E2 L6 PCTK	C98	Enzyme
O14980	XPO1 Exportin-1	C34; C528; C1070	Uncategorized
O75362	ZNF217 Zinc finger protein 217	C286	Transcription factors and regulators
O94953	KDM4B Lysine-specific demethylase 4B	C694	Enzyme
P00813	ADA Adenosine deaminase	C75	Enzyme
P04150	NR3C1 Glucocorticoid receptor	C302; C622	Transcription factors and regulators
P09086	POU2F2 POU domain, class 2, transcription factor 2	C346	Transcription factors and regulators
P09211	GSTP1 Glutathione S-transferase P	C48	Enzyme
P14598	NCF1 Neutrophil cytosol factor 1	C378	Adapter, scaffolding, modulator proteins
P15374	UCHL3 Ubiquitin carboxyl-terminal hydrolase isozyme L3	C95	Enzyme
P16455	MGMT Methylated-DNA--protein-cysteine methyltransferase	C145; C150	Enzyme
P17812	CTP synthase 1	C491	Enzyme
P19447	ERCC3 TFIIH basal transcription factor complex helicase	C342	Enzyme
P21580	TNFAIP3 Tumor necrosis factor alpha-induced protein 3	C54	Enzyme
P24752	ACAT1 Acetyl-CoA acetyltransferase, mitochondrial	C119; C126; C196; C413	Enzyme
P40261	Nicotinamide N-methyltransferase	C165	Enzyme
P40763	STAT3 Signal transducer and activator of transcription 3	C259	Transcription factors and regulators
P41226	UBA7 Ubiquitin-like modifier-activating enzyme 7	C599	Enzyme
P42575	CASP2 Caspase-2	C370	Enzyme
P43403	ZAP70 Tyrosine-protein kinase ZAP-70	C117	Enzyme
P48200	IREB2 Iron-responsive element-binding protein 2	C137	Transcription factors and regulators
P48735	IDH2 Isocitrate dehydrogenase	C308	Enzyme
P50851	LRBA Lipopolysaccharide-responsive and beige-like anchor protein	C1704; C2675	Uncategorized
P51617	IRAK1 Interleukin-1 receptor-associated kinase 1	C608	Enzyme
P61081	NEDD8-conjugating enzyme Ubc12	C47	Enzyme
P61088	Ubiquitin-conjugating enzyme E2 N	C87	Enzyme
P63244	GNB2L1 Guanine nucleotide-binding protein subunit beta-2-like 1	C182	Channels, Transporters, Receptors

Identifier	Protein Name	Cysteine Location	Protein Class
P68036	UBE2L3 Ubiquitin-conjugating enzyme E2 L3	C86	Enzyme
Q00535	CDK5 Cyclin-dependent kinase 5	C157	Enzyme
Q01201	RELB Transcription factor RelB	C109	Transcription factors and regulators
Q02556	IRF8 Interferon regulatory factor 8	C306	Transcription factors and regulators
Q04759	PRKCQ Protein kinase C theta type	C14; C17	Enzyme
Q06124	Tyrosine-protein phosphatase non-receptor type 11	C573	Enzyme
Q09472	EP300 Histone acetyltransferase p300	C1738	Enzyme
Q14790	CASP8 Caspase-8	C360	Enzyme
Q15084	PDIA6 Protein disulfide-isomerase A6	C55; C58; C190; C193	Enzyme
Q15306	IRF4 Interferon regulatory factor 4	C194	Transcription factors and regulators
Q15910	EZH2 Histone-lysine N-methyltransferase EZH2	C503	Enzyme
Q16186	Proteasomal ubiquitin receptor ADRM1	C88	Channels, Transporters, Receptors
Q16763	UBE2S Ubiquitin-conjugating enzyme E2 S	C118	Enzyme
Q16822	PCK2 Phosphoenolpyruvate carboxykinase	C306	Enzyme
Q16875	6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 3	C155	Enzyme
Q16877	PFKFB4 6-phosphofructo-2-kinase/fructose-2,6-bisphosphata	C159	Enzyme
Q6L8Q7	PDE12 2,5-phosphodiesterase 12	C108	Enzyme
Q70CQ2	USP34 Ubiquitin carboxyl-terminal hydrolase 34	C741; C1090	Enzyme
Q7Z2W4	ZC3HAV1 Zinc finger CCCH-type antiviral protein 1	C645	Transcription factors and regulators
Q86UV5	USP48 Ubiquitin carboxyl-terminal hydrolase 48	C39	Enzyme
Q8TAQ2	SMARCC2 SWI/SNF complex subunit SMARCC2	C145	Transcription factors and regulators
Q92851	Caspase-10	C401	Enzyme
Q93009	USP7 Ubiquitin carboxyl-terminal hydrolase 7	C223; C315	Enzyme
Q96FA3	PELI1 E3 ubiquitin-protein ligase pellino homolog 1	C282	Enzyme
Q96GG9	DCUN1D1 DCN1-like protein 1	C115	Uncategorized
Q96JH7	VCPIP1 Deubiquitinating protein VCIP135	C219	Enzyme
Q96RU2	USP28 Ubiquitin carboxyl-terminal hydrolase 28	C171; C733	Enzyme
Q99873	PRMT1 Protein arginine N-methyltransferase 1	C109	Enzyme
Q9C0C9	UBE2O Ubiquitin-conjugating enzyme E2 O	C375	Enzyme
Q9HB90	RRAGC Ras-related GTP-binding protein C	C358; C377	Channels, transporters, and receptors
Q9NRW4	Dual specificity protein phosphatase 22	C124	Enzyme
Q9NWZ3	IRAK4 Interleukin-1 receptor-associated kinase 4	C13	Enzyme
Q9NYL2	MLTK Mitogen-activated protein kinase kinase	C22	Enzyme

Identifier	Protein Name	Cysteine Location	Protein Class
	kinase MLT		
Q9UPT9	USP22 Ubiquitin carboxyl-terminal hydrolase 22	C44; C171	Enzyme
Q9Y3Z3	SAMHD1 SAM domain and HD domain-containing protein 1	C522	Enzyme
Q9Y4C1	KDM3A Lysine-specific demethylase 3A	C251	Enzyme
Q9Y5T5	USP16 Ubiquitin carboxyl-terminal hydrolase 16	C205	Enzyme

[00451] Table 4 shows representative cysteines with known covalent ligands targeted by fragment electrophiles in isoTOP-ABPP experiments.

Protein	Liganded cysteine	Fragment(s)	Other cysteines quantified by isoTOP-ABPP	Previous covalent inhibitor(s)	Cysteine location
BTK	C481	2, 3, 14, 31	C145, C337	Ibrutinib	Active site
TGM2	C277	12, 14, 32	C10, C27, C230, C269, C290, C336, C370, C524, C545, C620	18d	Active Site
Map2k7	C131	2, 3, 11, 14, 20, 21, 38	C260, C280	Ibrutinib	Active Site
XPO1	C528	2, 3, 5, 14, 24, 31, 43, 56	C34, C119, C164, C199, C327, C498, C723, C1070	KPT-330	Non-active site
Casp5	C315	3, 50	—	Z-WEHD-CHO/FMK ("WEHD" disclosed as SEQ ID NO: 863)	Active Site
Casp8	C360	2, 4, 11	C236, C409	Z-VAD-FMK, CV8/9-AOMK	Active Site
ERCC3	C342	2, 3, 5, 8, 14, 21	—	Triptolide	Active Site
Park 7 (Toxoplasma DJ-1)	C106	2, 9, 8, 11, 13, 43, 45, 50, 52	, C46, C53	WRR-086	Active Site
GSTO1	C32	2-13, 16, 18-22, 33, 27-30, 32-34, 36, 39, 43, 49, 50, 52, 54, 55	C90, C192, C237	KT53	Active Site
ALDH2	C319	3, 8-10, 12, 27, 28, 32, 39, 40, 43, 49, 50	C66, C179, C386, C472	Disulfiram	Active Site
CTSZ	C92	4, 11, 20, 28, 32	C89, C126, C132, C154, C170, C173, C179, C214	Cy5DCG04	Active Site

[00452] Table 5 shows Reactive docking results for liganded cysteines.

Protein	PDB ID:	Most ligandable cysteine by docking	Cysteine location	Most ligandable cysteine by isoTOP-ABPP	Match
Aldh2	1O05	C319	Active site	C319	Yes
BTK	1K2P	C481	Active site	C481	Yes
CASP8	1QTN	C360	Active Site	C360	Yes
CCNB1	2JGZ	C238	Non-active site	C238	Yes
CDKN3	1FQ1	C39	Non-active site	C39	Yes
CLIC4	2AEH	C35	Non-active site	C35	Yes
DTYMK	1E2G	C163	Non-active site	C163	Yes
IDH1	3MAP	C269	Non-active site	C269	Yes
IMPDH2	1NF7	C331	Active site	C331, C140	Yes
GLRX5	2WUL	C67	Active site	C67	Yes
GSTO1	1EEM	C32	Active site	C32	Yes
NME3	1ZS6	C158	Non-active site	C158	Yes
PKM	4JPG	C423	Non-active site	C423	Yes
SRC	2SRC	C277	Active Site	C277	Yes
TIGAR	3DCY	C114	Non-active site	C114,C161	Yes
TXNDC	1WOU	C43	Active site	C43	Yes
UGDH	3ITK	C276	Active site	C276	Yes
UPP1	3EUF	C162	Non-active site	C162	Yes
XPO1	3GB8	C528	Non-active site	C528	Yes
CDK5	1UNG	C157	Non-active site	C269	Second
EDC3	3D3K	C311	Non-active site	C137, C413, C499	Second
NR2F2	3CJW	C213	Non-active site	C326, C213 (in situ)	Second
PDCD6IP	2R02	C231	Non-active site	C90	Second
PRMT1	1ORI	C285	Active site	C109	Second
UBE2S	1ZDN	C118	Non-active site	C95	Second
FNBP1	2EFL	C145	Non-active site	C70	No
HAT1	2P0W	C120	Non-active site	C101	No
MAPK9	3NPC	C163	Active site	C177	No
STAT1	1YVL	C543	Non-active site	C492, C255	No

[00453] Table 6 shows site of fragment labeling for recombinant proteins. The underlines portion indicates the fragment-modified cysteines.

Protein	Cysteine	Fragment #	Peptide	SEQ ID NO:	M+H calculated (m/z)	M+H observed (m/z)	Charge
IMPDH2	C140	14	R.HGFC <u>G</u> IPITDTGR.M	45	715.86	715.86	2
TIGAR	C114	5	R.EEC <u>P</u> VFTPPGGETLDQVK.M	143	1123.97	1123.97	2
CASP8	C360	7	K.VFFIQAC <u>Q</u> GDNYQK.G	335	660.98	660.98	3
IDH1	C269	20	K.SEGGFIWAC <u>K</u> .N	260	702.84	702.84	2

[00454] Table 7 illustrates a list of DMF-sensitive Cys residues in human T cells, defined as Cys residues that showed R values (DMSO/DMF) > 4 in isoTOP-ABPP experiments comparing DMSO- versus DMF-treated T cells.

Name	Full name	Protein function	Residue	Conserved in mice	Role in immunology
ADA	Adenosine deaminase	Adenosine deaminase	C75	yes	Positive regulator of T cell co-activation
AGFG2	Arf-GAP domain and FG repeat-containing protein 2	GTPase activator	C39	yes	Unknown
AIP	AH receptor-interacting protein	Transcription factor binding	C122	yes	Unknown
CRKL	Crk-like protein	Poly(A) RNA binding	C249	yes	Unknown
FLII	Protein flightless-1 homolog	Actin binding	C46	yes	Unknown
GAK	Cyclin-G-associated kinase	Serine/threonine protein kinase	C87	yes	Unknown
HUWE1	E3 ubiquitin-protein ligase HUWE1	E3 ubiquitin-protein ligase	C3372	yes	Unknown
IKBKB	Inhibitor of nuclear factor kappa-B kinase subunit	Serine kinase	C464	yes	Phosphorylates IκB-α in NF-κB pathway
IL16	Pro-interleukin-16	Cytokine	C1004	yes	Influences migration of CD4+ lymphocytes
IRF4	Interferon regulatory factor 4	DNA binding	C194	yes	Regulates dendritic cell and B cell development, as well as T/B cell differentiation
IRF8	Interferon regulatory factor 8	DNA binding	C306	yes	Plays a negative regulatory role in immune cells. Binds to upstream regulatory region of MHC class I genes. Regulates the development and differentiation of myeloid cells.
KIAA0528	Uncharacterized protein	Calcium-dependent phospholipid binding	C993	yes	Unknown
LAS1L	Ribosomal biogenesis protein	Poly(A) RNA binding	C456	yes	Unknown
MARS2	Methionine--tRNA ligase, mitochondrial	Methionine-tRNA ligase	C425	yes	Unknown
MAT2A	S-adenosylmethionine synthase isoform type-2	Methionine adenosyltransferase	C56	yes	Unknown
MAT2A	S-adenosylmethionine synthase isoform type-2	Methionine adenosyltransferase	C104	yes	Unknown

MTCH2	Mitochondrial carrier homolog 2	Induces mitochondrial depolarization	C296	yes	Unknown
PGP	Phosphoglycolate phosphatase	Phosphatase	C297	yes	Unknown
PML	Protein Promyelocytic leukemia	RNA/DNA binding	C479	yes	Modulates TGF-beta signaling, induced by interferon to promote antiviral responses
PRKCQ	Protein kinase C theta type	Serine/threonine protein kinase	C14	yes	Promotes TCR signaling through activation of NF- κ B and other transcription factors
PYGB	Glycogen phosphorylase, brain form	Phosphorylase	C326	yes	Unknown
RARS	Arginine--tRNA ligase, cytoplasmic	tRNA binding	C32	yes	Unknown
SON	Protein SON	RNA/DNA binding	C92	yes	Unknown
SYNE2	Nesprin-2	Actin binding	C553	yes	Unknown
TDRKH	Tudor and KH domain-containing protein	RNA binding	C109	yes	Unknown
THNSL1	Threonine synthase-like 1	Threonine synthase	C324	yes	Unknown
THOC1	THO complex subunit 1	RNA/DNA binding	C49	yes	Unknown
TNFAIP3	Tumor necrosis factor alpha-induced protein 3	Ubiquitin-specific protease	C54	yes	Inhibits NF- κ B signaling upon TCR-mediated T cell activation
UBR4	E3 ubiquitin-protein ligase	Ubiquitin ligase	C2554	yes	Unknown
USP7	Ubiquitin carboxyl-terminal hydrolase 7	Ubiquitin-specific protease	C315	yes	Deubiquitinates FOXP3, increasing Treg suppressive capacity
VDAC3	Voltage-dependent anion-selective channel protein	Mitochondrial outer membrane channel	C65	yes	Unknown
VDAC3	Voltage-dependent anion-selective channel protein	Voltage-gated anion channel	C36	yes	Unknown
ZC3HAV1	Zinc finger CCCH-type antiviral protein 1	Poly(A) RNA binding	C645	yes	Inhibits viral replication
ZNF346	Zinc finger protein 346	RNA binding	C68	yes	Unknown
AARS	Alanine--tRNA ligase, cytoplasmic	Alanine-tRNA ligase	C773	no	Unknown
APOBEC3C	Probable DNA dC-dU- editing enzyme	Cytidine deaminase	C130	no	Inhibits retrovirus replication

BCL2A1	Bcl-2-related protein A1	Scaffolding protein	C55	no	Expression induced by inflammatory cytokines
BCL2A1	Bcl-2-related protein A1	Scaffolding protein	C19	no	Unknown
CHRAC1	Chromatin accessibility complex protein 1	Chromatin remodeling	C55	no	Unknown
DCXR	L-xylulose reductase	Xylulose reductase	C244	no	Unknown
GHDC	GH3 domain-containing protein	Uncharacterized	C502	no	Unknown
IRAK4	Interleukin-1 receptor-associated kinase 4	Serine/threonine protein kinase	C13	no	Helps initiate innate immune response by promoting ubiquitination of IRAK1 upon TLR activation. Also implicated in T cell activation
NADSYN1	Glutamine-dependent NAD(+) synthetase	NAD(+) synthase	C428	no	Unknown
PGLS	6-phospho-gluconolactonase	Hydrolysis of 6-phosphogluconolactone	C32	no	Unknown
PRKDC	DNA-dependent protein kinase catalytic subunit	Serine/threonine protein kinase	C4045	no	Regulates DNA damage response, involved in V(D)J recombination
PUSL1	tRNA pseudouridine synthase-like 1	Pseudouridine synthase	C292	no	Unknown
RIN3	Ras and Rab interactor 3	GTPase activator	C942	no	Unknown
SCLY	Selenocysteine lyase	Selenocysteine lyase	C22	no	Unknown
SPCS2	Signal peptidase complex subunit 2	Peptidase	C17	no	Unknown
TRNT1	CCA tRNA nucleotidyltransferase 1, mitochondrial	tRNA binding	C373	no	Mutations lead to B-cell immunodeficiency as well as progressive reductions in T and NK cells (OMIM number 616084)
TUBGCP3	Gamma-tubulin complex component 3	Gamma-tubulin binding	C194	no	Unknown
UBE2L6	Ubiquitin/ISG 15-conjugating enzyme E2 L6	Ubiquitin-conjugating enzyme	C98	no	Acts as an E2 enzyme for an IFN-induced ubiquitin-like protein

[00455] Table 8 illustrates an exemplary list of DMF sensitive cysteine-containing proteins in human T cell targets. Table 8 further shows the accession number (or the protein identifier) of the protein.

Identifier	Protein Name	SEQ ID	DMF 50u	DMF 50u	DMF 50u	DMF 25u	DMF 10u	MMF 50u
------------	--------------	--------	---------	---------	---------	---------	---------	---------

		NO:	M_4h	M_2h	M_1h	M_4h	M_4h	M_4h
Q9NRW3_C130	APOBEC3C Probable DNA dC-dU-editing enzyme APOBEC-3C	805	20	20	8.6	--	1.92	1.48
Q9NWZ3_C13	IRAK4 Interleukin-1 receptor-associated kinase 4	806	20	--	8.3	--	--	1.48
Q9Y2W6_C109	TDRKH Tudor and KH domain-containing protein	807	20	20	4	--	2.34	1.36
Q6IA69_C428	NADSYN1 Glutamine-dependent NAD(+) synthetase	808	20	2.31	1.81	--	1.43	1.33
O14920_C464	IKBKB Inhibitor of nuclear factor kappa-B kinase subunit	809	20	10.12	3.96	--	2.59	--
P00813_C75	ADA Adenosine deaminase	810	20	5.08	2.51	--	2.29	--
Q9Y277_C65	VDAC3 Voltage-dependent anion-selective channel protein	811	15.94	7.53	3.35	5.64	1.73	1.39
P49588_C773	AARS Alanine--tRNA ligase, cytoplasmic	812	12.75	10.16	9.34	--	2.84	1.24
O14933_C98	UBE2L6 Ubiquitin/ISG15-conjugating enzyme E2 L6	813	12.55	2.92	2.44	--	1.49	1.7
O95336_C32	PGLS 6-phosphogluconolactonase	814	11.51	9.49	3.42	5.32	1.9	1.26
A6NDG6_C297	PGP Phosphoglycolate phosphatase	815	10.77	4.21	3.06	--	--	1.52
Q7Z6Z7_C3372	HUWE1 E3 ubiquitin-protein ligase HUWE1	816	10.48	4.43	2.28	--	1.58	1.2
Q16548_C55	BCL2A1 Bcl-2-related protein A1	817	7.18	--	--	--	--	0.97
P11216_C326	PYGB Glycogen phosphorylase, brain form	818	6.76	3.73	2.47	3.53	1.65	1.29
O95081_C39	AGFG2 Arf-GAP domain and FG repeat-containing protein 2	819	6.39	3.85	--	--	1.42	1.24
Q7Z2W4_C645	ZC3HAV1 Zinc finger CCCH-type antiviral protein 1	820	6.28	3.13	2.36	2.52	1.46	1.3
O00170_C122	AIP AH receptor-interacting protein	821	6.14	3.05	--	--	--	1.24
Q96Q11_C373	TRNT1 CCA tRNA nucleotidyltransferase 1, mitochondrial	822	5.83	2.66	1.97	--	--	1.29
Q8TB24_C942	RIN3 Ras and Rab interactor 3	823	5.7	3	--	--	--	1.23
Q9Y4W2_C456	LAS1L Ribosomal biogenesis protein LAS1L	824	5.61	3.42	1.8	--	1.29	1.14
Q02556_C306	IRF8 Interferon regulatory factor 8	825	5.32	--	1.66	--	1.9	--
Q96GW9_C425	MARS2 Methionine--tRNA ligase, mitochondrial .A	826	5.3	4.16	2.23	2.86	1.84	1.3
Q15306_C194	IRF4 Interferon regulatory factor 4	827	5.25	3.13	1.32	1.78	1.69	1.33
Q15005_C17	SPCS2 Signal peptidase complex subunit 2	828	5.09	3.86	2.25	2.41	1.42	1.32
P54136_C32	RARS Arginine--tRNA ligase, cytoplasmic	829	5.02	3.58	2.58	4.03	0.62	1.78
Q96CW5_C194	TUBGCP3 Gamma-tubulin complex component 3	830	4.94	2.44	--	--	--	--
P46109_C249	CRKL Crk-like protein	831	4.86	3.21	2.21	--	1.38	1.27
Q8N0Z8_C292	PUSL1 tRNA pseudouridine synthase-like 1	832	4.68	--	--	--	--	1.36

Q5T4S7_ C2554	UBR4 E3 ubiquitin-protein ligase UBR4	833	4.63	2.1	1.6	--	1.52	1.2
Q9UL40_ C68	ZNF346 Zinc finger protein 346	834	4.6	3.91	2.5	--	1.98	1.26
Q13045_ C46	FLII Protein flightless-1 homolog	835	4.5	3.57	2.05	--	1.55	1.27
Q86YS7_ C993	KIAA0528 Uncharacterized protein KIAA0528	836	4.38	--	--	--	1.4	1.4
Q9Y6C9_ C296	MTCH2 Mitochondrial carrier homolog 2	837	4.3	2.44	1.81	--	1.68	1.35
Q7Z4W1_ C244	DCXR L-xylulose reductase	838	4.24	2.76	1.29	--	2.3	--
Q04759_ C14	PRKCQ Protein kinase C theta type	839	4.21	2.92	--	3.29	1.62	1.14
P18583_ C92	SON Protein SON	840	4.17	6.31	2.5	--	--	1.31
P31153_ C56	MAT2A S-adenosylmethionine synthase isoform type-2	841	4.17	--	--	--	--	--
Q16548_ C19	BCL2A1 Bcl-2-related protein A1	842	4.16	2.09	--	2.19	1.15	1.28
Q14005_ C1004	IL16 Pro-interleukin-16	843	4.13	3.32	1.95	--	1.37	1.31
P31153_ C104	MAT2A S-adenosylmethionine synthase isoform type-2	844	4.11	--	--	--	1.5	1.31
Q9Y277_ C36	VDAC3 Voltage-dependent anion- selective channel protein	845	4.11	3.98	3.21	--	--	1.18
Q8WXH0_ C553	SYNE2 Nesprin-2	846	4.05	3.29	--	--	1.61	--
Q96I15_ C22	SCLY Selenocysteine lyase	847	4.04	2.16	1.9	2.16	1.31	1.27
P29590_ C479	PML Protein PML	848	--	4.57	--	2.17	2.1	1.27
Q8IYQ7_ C324	THNSL1 Threonine synthase-like 1	849	--	19.36	15.93	--	--	1.4
Q93009_ C315	USP7 Ubiquitin carboxyl-terminal hydrolase 7	850	--	14.06	5.33	--	1.9	1.4
P21580_ C54	TNFAIP3 Tumor necrosis factor alpha-induced protein 3	851	--	5.34	--	--	1.58	--
O14976_ C87	GAK Cyclin-G-associated kinase	852	--	4.79	--	--	1.36	--
Q96FV9_ C49	THOC1 THO complex subunit 1	853	--	5.7	3.93	--	--	0.97
P78527_ C4045	PRKDC DNA-dependent protein kinase catalytic subunit	854	--	10.53	4.14	--	--	1.23
Q9NRG0_ C55	CHRAC1 Chromatin accessibility complex protein 1	855	--	11.72	12.59	--	5.07	1.27
Q8N2G8_ C502	GHDC GH3 domain-containing protein	856	--	20	4.23	--	--	--

[00456] Table 9 illustrates the full protein sequence of exemplary cysteine-containing proteins described herein. The cysteine residue of interest is denoted with (*).

Protein Identifier (Accession No.)	Protein Name	Cysteine Residue Number	Sequence	SEQ ID NO:
O75874	Isocitrate dehydrogenase 1 (IDH1)	C269	MSKKISGGSV VEMQGDDEMTR IIWELIKEKL IFPYVELDLH SYDLGIENRD ATNDQVTKDA AEAIIKHNVG VKCATITPDE KRVEEFKLKQ MWKSPNGTIR NILGGTVFRE AIICKNIPRL VSGWVKPIII GRHAYGDQYR ATDFVVPGP KVEITYTPSD GTQKVTYLVH NFEEGGGVAM GMYNQDKSIE DFAHSSFQMA LSKGWPLYLS TKNTILKKYD GRFKDIFQEI YDKQYKSQFE AQKIWYEHRL IDDMVAQAMK SEGGFIWAC*K NYDGDVQSDS VAQGYGSLGM MTSVLVCPDG KTVEAEAAHG TVTRHYRMYQ KGQETSTNPI ASIFAWTRGL AHRAKLDNNK ELAFFANALE EVSIIETIEAG FMTKDLAACI KGLPNVQRSD YLNTFEFMDK LGENLKI KLA QAKL	1
P48735	Isocitrate dehydrogenase 2 (IDH2)	C308	MAGYLRVVRS LCRASGSRPA WAPAALTAPT SSEQPRRHYA DKRIKVAKPV VEMDGDDEMTR IIWQFIKEKL ILPHVDIQLK YFDLGLPNRD QTDDQVTIDS ALATQKYSVA VKCATITPDE ARVEEFKLKK MWKSPNGTIR NILGGTVFRE PIICKNIPRL VPGWTKPITI GRHAHGDQYK ATDFVADRAG TFKMVFTPKD GSGVKEWEVY NFPAGGVGMG MYNTDESIG FAHSCFQYAI QKKWPLYMST KNTILKAYDG RFKDIFQEIF DKHYKTDFDK NKIWYEHRLI DDMVAQVLKS SGGFVWAC*KN YDGDVQSDIL AQGFGLGLM TSVLVCPDGK TIEAEAAHGT VTRHYREHQK GRPTSTNPIA SIFAWTRGLE HRGKLDGNQD LIRFAQMLEK VCVETVESGA MTKDLAGCIH GLSNVKLNEH FLNTTDFLDT	2

			IKSNLDRALG RQ	
Q14790	CASP8	C360	MDFSRNLYDI GEQLDSEDLA SLKFLSLDYI PQRKQEPIKD ALMLFQRLQE KRMLEESNLS FLKELLFRIN RLDLLITYLN TRKEEMEREL QTPGRAQISA YRVMLYQISE EVSRSELRSF KFLQEEISK CKLDDDMNLL DIFIEMEKRV ILGEGKLDIL KRVCAQINKS LLKIINDYEE FSKERSSSLE GSPDEFSNGE ELCGVMTISD SPREQDSESQ TLDKVYQMKS KPRGYCLIIN NHNFAKAREK VPKLHSIRDR NGTHLDAGAL TTTFEELHFE IKPHDDCTVE QIYEILKIYQ LMDHSNMDCF ICCILSHGDK GIIYGTGQE APIYELTSQF TGLKCPSLAG KPKVFFIQAC* QGDNYQKGIP VETDSEEQPY LEMDLSSPQT RYIPDEADFL LGMATVNNCV SYRNPAEGTW YIQSLCQSLR ERCPRGDDIL TILTEVNYEV SNKDDKKNMG KQMPQPTFTL RKKLVFPSD	3
Q92851	CASP10	C401	MKSQGQHWYS SSDKNCKVSF REKLLIIDSN LGVQDVENLK FLCIGLVPNKKLEKSSSASD VFEHLLAEDL LSEEDPFFLA ELLYIIRQKK LLQHLNCTKE EVERLLPTRQ RVSLFRNLLY ELSEGIDSEN LKDMIFLLKD SLPKTEMTSL SFLAFLEKQG KIDEDNLTCL EDLCKTVVPK LLRNIEKYKR EKAIQIVTPP VDKEAESYQG EEELVSQTDV KTFLEALPQE SWQNKHAGSN GNRATNGAPS LVSRGMQGAS ANTLNSETST KRAAVYRMNR NHRGLCVIVN NHSFTSLKDR QGTHKDAEIL SHVFQWLGFT VHIHNNVTKV EMEMVLQKQK CNPAHADGDC FVFCILTHGR FGAVYSSDEA LIPIREIMSH FTALQCPRLA EKPKLFFIQA C* QGEEIQPSV SIEADALNPE QAPTSLQDSI PAEADFLGL ATPPGYVSFR HVEEGSWYIQ SLCNHLKKLV	4

			PRMLKFLEKT MEIRGRKRTV WGAKQISATS LPTAISAQTP RPPMRRWSSV S	
Q99873	PRMT1	C109	MENFVATLAN GMSLQPPLEE VSCGQAESSE KPNAEDMTSK DYYFDSYAHF GIHEEMLKDE VRTLTyrNSM FHNRHLFKDK VVLdVGSgtG ILCMFAAKAG ARKVIGIEC* S SISDYAVKIV KANKLDHVVT IIKGKVEEVE LPVEKVDIII SEWMGYCLFY ESMLNTVLYA RDKWLAPDGL IFPD RATLYV TAIEDRQYKD YKIHWWENVY GFDMSCIKDV AIKEPLVDVV DPKQLVTNAC LIKEVDIYTV KVEDLTFTSP FCLQVKRNDY VHALVAYFNI EFTRCHKRTG FSTSPESPYT HWKQTVFYME DYLT VKTGEE IFGTIGMRPN AKNNRDLDF IDLDFKGQLC ELSCSTDYRM R	5
Q9NYL2	MAP3 kinase MLTK (or ZAK)	C22	MSSLGASFVQ IKFDDLQFFE NC*GGGSFGSV YRAKWISQDK EVAVKKLLKI EKEAEILSVL SHRNIIQFYG VILEPPNYGI VTEYASLGSL YDYINSNRSE EMDMDHIMTW ATDVAKGMHY LHMEAPVKVI HRDLKSRNVV IAADGV LKIC DFGASRFH NH TTHMSLVGTF PWMAPEVIQS LPVSETCDTY SYGVVLW EML TREVPFKGLE GLQVAWL VVE KNERLTIPSS CPRSFAELLH QCWEADAKKR PSFKQIISIL ESMSNDTSLP DKCNSFLHNK AEWRCIEAT LERLKKLERD LSFKEQELKE RERRLKMWEQ KLTEQSN TPL LPSFEIGAWT EDDVYCWVQQ LVRKGDSSAE MSVYASLFKE NNITGKRLLL LEEEDLKDMG IVSKGHIIHF KSAIEKLTHD YINLFHFPPL IKDSGGEPEE NEEKIVNLEL VFGFHLKPGT GPQDCKWKMY MEMDGDEIAI TYIKDVT FNT NLPDAEILKM TKPPFVMEKW IVGIAKSQTV ECTVTYESDV RTPKSTKHVH	6

			SIQWSRTKPQ DEVKAVQLAI QTLFTNSDGN PGSRSDSSAD CQWLDTLRMR QIASNTSLQR SQSNPILGSP FFSHFDGQDS YAAAVRRPQV PIKYQQITPV NQSRSSSPTQ YGLTKNFSSL HLNSRDSGFS SGNTDTSSER GRYSDRSRKN YGRGSISLNS SPRGRYSGKS QHSTPSRGRY PGKFYRVSQS ALNPHQSPDF KRSPRDLHQP NTIPGMPLHP ETDSRASEED SKVSEGGWTK VEYRKKPHRP SPAKTNKERA RGDHRGWRF	
P12268	IMPDH2	C140, C331	MADYLISGGT SYVPDDGLTA QQLFNCGDGL TYNDLILPG YIDFTADQVD LTSALTKKIT LKTPLVSSPM DTVTEAGMAI AMALTGGIGF IHHNCTPEFQ ANEVRKVKKY EQGFITDPVV LSPKDRVRDV FEAKARHGFC* GIPITDTGRM GSRLVGISS RDIDFLKEEE HDCFLEEIMT KREDLVVAPA GITLKEANEI LQRSKKGKLP IVNEDDELVA IIARTDLKKN RDYPLASKDA KKQLLCGAAI GTHEDDKYRL DLLAQAGVDV VVLDSSQGNS IFQINMIKYI KDKYPNLQVI GGNVVTAQA KNLIDAGVDA LRVGMGSGSI C*ITQEVLAG RPQATAVYKV SEYARRFGVP VIADGGIQNV GHIAKALALG ASTVMMGSL AATTEAPGEY FFSDGIRLKK YRGMGSLDAM DKHLSSQNRV FSEADKIKVA QGVSGAVQDK GSIHKFVPYL IAGIQHSCQD IGAKSLTQVR AMMYSGELKF EKRTSSAQVE GGVHSLHSYE KRLF	7
Q9NQ88	TIGAR	C114, C161	MARFALTVVR HGETRFNKEK IIQGQGVDEP LSETGFKQAA AAGIFLNNVK FTHAFSSDLM RTKQTMHGIL ERSKFCKDMT VKYDSRLRER KYGVVEGKAL SELRAMAKAA REEC*PVFTPP GGETLDQVKM RGIDFFFLC QLILKEADQK EQFSQGSPSN C*LETSLAEIF	8

			PLGKNHSSKV NSDSGIPGLA ASVLVVSHGA YMRS�FDYFL TDLKCSLPAT LSRSELMSVT PNTGMSLFII NFEEGREVKP TVQCICMNLQ DHLNGLTETR	
Q04759	PKCθ	C14, C17	MSPFLRIGLS NFDC*GSC*QSC QGEAVNPYCA VLVKEYVESE NGQMYIQKKP TMYPPWDSTF DAHINKGRVM QIIVKGKNVD LISSETVELY SLAERCCKNN GKTEIWLELK PQGRMLMNAR YFLEMSDTKD MNEFETEGFF ALHQRRGAIK QAKVHHVKCH EFTATFFPQP TFCSVCHEFV WGLNKQGYQC RQCNAAIHKK CIDKVIKCT GSAINSRETM FHKERFKIDM PHRFKVYNYK SPTFCEHCGT LLWGLARQGL KCDACGMNVH HRCQTKVANL CGINQKLMAE ALAMIESTQQ ARCLRDTEQI FREGPVEIGL PCSIKNEARP PCLPTPGKRE PQGISWESPL DEVDKMCHLP EPELNKERPS LQIKLKIEDF ILHKMLGKGS FGKVFLAEFK KTNQFFAIKA LKKDVVLMDD DVECTMVEKR VLSLAWEHF LTHMFCTFQT KENLFFVMEY LNGGDLMYHI QSCHKFDLSR ATFYAAEIL GLQFLHSKGI VYRDLKLDNI LLDKDGHKI ADFGMCKENM LGDAKTNTFC GTPDYIAPEI LLGQKYNHSV DWWSFGVLLY EMLIGQSPFH GQDEEELFHS IRMDNPFYPR WLEKEAKDLL VKLFVREPEK RLGVRGDIRQ HPLFREINWE ELERKEIDPP FRPKVKSPFD CSNFDKEFLN EKPRLSFADR ALINSMDQNM FRNFSFMNPG MERLIS	9

[00457] Table 10A – Table 10E illustrate a list of cysteine containing proteins and potential cysteine site of conjugation separated by protein class. Table 10A illustrates cysteine containing enzymes and potential cysteine conjugation site. Table 10B shows a list of cysteine containing transcription factors and regulators. Table 10C shows an exemplary list of cysteine containing

channels, transporters and receptors. Table 10D illustrates an exemplary cysteine containing adapter, scaffolding, and modulator protein. Table 10E provides an exemplary list of uncategorized cysteine containing proteins.

Table 10A

Identifier	Protein Name	Cysteine Location	Protein Class
O14920	IKBKB Inhibitor of nuclear factor kappa-B kinase subunit	C464	Enzyme
O14933	UBE2L6 Ubiquitin/ISG15-conjugating enzyme E2 L6 PCTK	C98	Enzyme
O94953	KDM4B Lysine-specific demethylase 4B	C694	Enzyme
P00813	ADA Adenosine deaminase	C75	Enzyme
P09211	GSTP1 Glutathione S-transferase P	C48	Enzyme
P15374	UCHL3 Ubiquitin carboxyl-terminal hydrolase isozyme L3	C95	Enzyme
P16455	MGMT Methylated-DNA--protein-cysteine methyltransferase	C145; C150	Enzyme
P17812	CTP synthase 1	C491	Enzyme
P19447	ERCC3 TFIIH basal transcription factor complex helicase	C342	Enzyme
P21580	TNFAIP3 Tumor necrosis factor alpha-induced protein 3	C54	Enzyme
P24752	ACAT1 Acetyl-CoA acetyltransferase, mitochondrial	C119; C126; C196; C413	Enzyme
P40261	Nicotinamide N-methyltransferase	C165	Enzyme
P41226	UBA7 Ubiquitin-like modifier-activating enzyme 7	C599	Enzyme
P42575	CASP2 Caspase-2	C370	Enzyme
P43403	ZAP70 Tyrosine-protein kinase ZAP-70	C117	Enzyme
P48735	IDH2 Isocitrate dehydrogenase	C308	Enzyme
P51617	IRAK1 Interleukin-1 receptor-associated kinase 1	C608	Enzyme
P61081	NEDD8-conjugating enzyme Ubc12	C47	Enzyme
P61088	Ubiquitin-conjugating enzyme E2 N	C87	Enzyme
P68036	UBE2L3 Ubiquitin-conjugating enzyme E2 L3	C86	Enzyme
Q00535	CDK5 Cyclin-dependent kinase 5	C157	Enzyme
Q04759	PRKCQ Protein kinase C theta type	C14; C17	Enzyme
Q06124	Tyrosine-protein phosphatase non-receptor type 11	C573	Enzyme
Q09472	EP300 Histone acetyltransferase p300	C1738	Enzyme
Q14790	CASP8 Caspase-8	C360	Enzyme
Q15084	PDIA6 Protein disulfide-isomerase A6	C55; C58; C190; C193	Enzyme
Q15910	EZH2 Histone-lysine N-methyltransferase EZH2	C503	Enzyme
Q16763	UBE2S Ubiquitin-conjugating enzyme E2 S	C118	Enzyme
Q16822	PCK2 Phosphoenolpyruvate carboxykinase	C306	Enzyme
Q16875	6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 3	C155	Enzyme
Q16877	PFKFB4 6-phosphofructo-2-kinase/fructose-2,6-bisphosphata	C159	Enzyme
Q6L8Q7	PDE12 2,5-phosphodiesterase 12	C108	Enzyme

Identifier	Protein Name	Cysteine Location	Protein Class
Q70CQ2	USP34 Ubiquitin carboxyl-terminal hydrolase 34	C741; C1090	Enzyme
Q86UV5	USP48 Ubiquitin carboxyl-terminal hydrolase 48	C39	Enzyme
Q92851	Caspase-10	C401	Enzyme
Q93009	USP7 Ubiquitin carboxyl-terminal hydrolase 7	C223; C315	Enzyme
Q96FA3	PELI1 E3 ubiquitin-protein ligase pellino homolog 1	C282	Enzyme
Q96JH7	VCPIP1 Deubiquitinating protein VCIP135	C219	Enzyme
Q96RU2	USP28 Ubiquitin carboxyl-terminal hydrolase 28	C171; C733	Enzyme
Q99873	PRMT1 Protein arginine N-methyltransferase 1	C109	Enzyme
Q9C0C9	UBE2O Ubiquitin-conjugating enzyme E2 O	C375	Enzyme
Q9NRW4	Dual specificity protein phosphatase 22	C124	Enzyme
Q9NWZ3	IRAK4 Interleukin-1 receptor-associated kinase 4	C13	Enzyme
Q9NYL2	MLTK Mitogen-activated protein kinase kinase kinase MLT	C22	Enzyme
Q9UPT9	USP22 Ubiquitin carboxyl-terminal hydrolase 22	C44; C171	Enzyme
Q9Y3Z3	SAMHD1 SAM domain and HD domain-containing protein 1	C522	Enzyme
Q9Y4C1	KDM3A Lysine-specific demethylase 3A	C251	Enzyme
Q9Y5T5	USP16 Ubiquitin carboxyl-terminal hydrolase 16	C205	Enzyme

Table 10B

Identifier	Protein Name	Cysteine Location	Protein Class
O75362	ZNF217 Zinc finger protein 217	C286	Transcription factors and regulators
P04150	NR3C1 Glucocorticoid receptor	C302; C622	Transcription factors and regulators
P09086	POU2F2 POU domain, class 2, transcription factor 2	C346	Transcription factors and regulators
P40763	STAT3 Signal transducer and activator of transcription 3	C259	Transcription factors and regulators
P48200	IREB2 Iron-responsive element-binding protein 2	C137	Transcription factors and regulators
Q01201	RELB Transcription factor RelB	C109	Transcription factors and regulators
Q02556	IRF8 Interferon regulatory factor 8	C306	Transcription factors and regulators
Q15306	IRF4 Interferon regulatory factor 4	C194	Transcription factors and regulators

Identifier	Protein Name	Cysteine Location	Protein Class
Q7Z2W4	ZC3HAV1 Zinc finger CCCH-type antiviral protein 1	C645	Transcription factors and regulators
Q8TAQ2	SMARCC2 SWI/SNF complex subunit SMARCC2	C145	Transcription factors and regulators

Table 10C

Identifier	Protein Name	Cysteine Location	Protein Class
P63244	GNB2L1 Guanine nucleotide-binding protein subunit beta-2-like 1	C182	Channels, Transporters, Receptors
Q16186	Proteasomal ubiquitin receptor ADRM1	C88	Channels, Transporters, Receptors
Q9HB90	RRAGC Ras-related GTP-binding protein C	C358; C377	Channels, transporters, and receptors

Table 10D

Identifier	Protein Name	Cysteine Location	Protein Class
P14598	NCF1 Neutrophil cytosol factor 1	C378	Adapter, scaffolding, modulator proteins

Table 10E

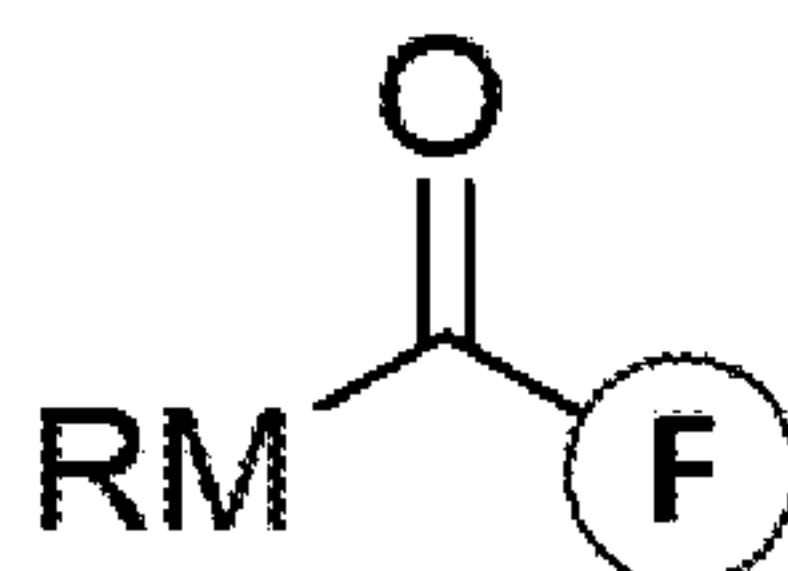
Identifier	Protein Name	Cysteine Location	Protein Class
O00170	AIP AH receptor-interacting protein	C122	Uncategorized
O00541	PES1 Pescadillo homolog	C272; C361	Uncategorized
O00622	CYR61 Protein CYR61	C39; C70; C134	Uncategorized
O14980	XPO1 Exportin-1	C34; C528; C1070	Uncategorized
P50851	LRBA Lipopolysaccharide-responsive and beige-like anchor protein	C1704; C2675	Uncategorized
Q96GG9	DCUN1D1 DCN1-like protein 1	C115	Uncategorized

[00458] The examples and embodiments described herein are for illustrative purposes only and various modifications or changes suggested to persons skilled in the art are to be included within the spirit and purview of this application and scope of the appended claims.

CLAIMS**WHAT IS CLAIMED IS:**

1. A method of identifying a cysteine containing protein as a binding target for a small molecule fragment, comprising:
 - a) obtaining a set of cysteine-reactive probe-protein complexes from a sample comprising a first cell solution treated with a small molecule fragment and a cysteine reactive probe wherein the cysteine-reactive probe comprises a reactive moiety capable of forming a covalent bond with a cysteine residue located on the cysteine containing protein;
 - b) analyzing the set of cysteine-reactive probe-protein complexes by a proteomic analysis means; and
 - c) based on step b), identifying a cysteine containing protein as the binding target for the small molecule fragment.
2. The method of claim 1, further comprising determining a value of each of the cysteine containing protein from the set of cysteine-reactive probe-protein complexes for identifying a cysteine containing protein as the binding target for the small molecule fragment, wherein the value is determined based on the proteomic analysis means of step b).
3. The method of claim 1, wherein the sample further comprises a second cell solution.
4. The method of claim 1, further comprising contacting the first cell solution with a small molecule fragment for an extended period of time prior to incubating the first cell solution with a first cysteine-reactive probe to generate a first group of cysteine-reactive probe-protein complexes.
5. The method of claim 4, wherein the extended period of time is about 5, 10, 15, 20, 30, 60, 90, 120 minutes or longer.
6. The method of claim 3, further comprising contacting the second cell solution with a second cysteine-reactive probe to generate a second group of cysteine-reactive probe-protein complexes.
7. The method of any one of the claims 3-6, wherein the first cysteine-reactive probe and the second cysteine-reactive probe are the same.
8. The method of any one of the claims 3-7, wherein the first group and the second group of cysteine-reactive probe-protein complexes comprise the set of cysteine-reactive probe-protein complexes.

9. The method of claim 1, wherein the cysteine containing protein is an enzyme, a transporter, a receptor, a channel protein, an adaptor protein, a chaperone, a signaling protein, a plasma protein, transcription related protein, translation related protein, mitochondrial protein, or cytoskeleton related protein.
10. The method of claim 1, wherein the cysteine containing protein is a protein illustrated in Table 3.
11. The method of claim 1, wherein the cysteine containing protein is a protein illustrated in Table 1, Table 2, Table 8, Table 9, Table 10A, Table 10B, Table 10C, Table 10D or Table 10E.
12. The method of claim 1, wherein the small molecule fragment is a small molecule fragment of Formula (I):



Formula (I)

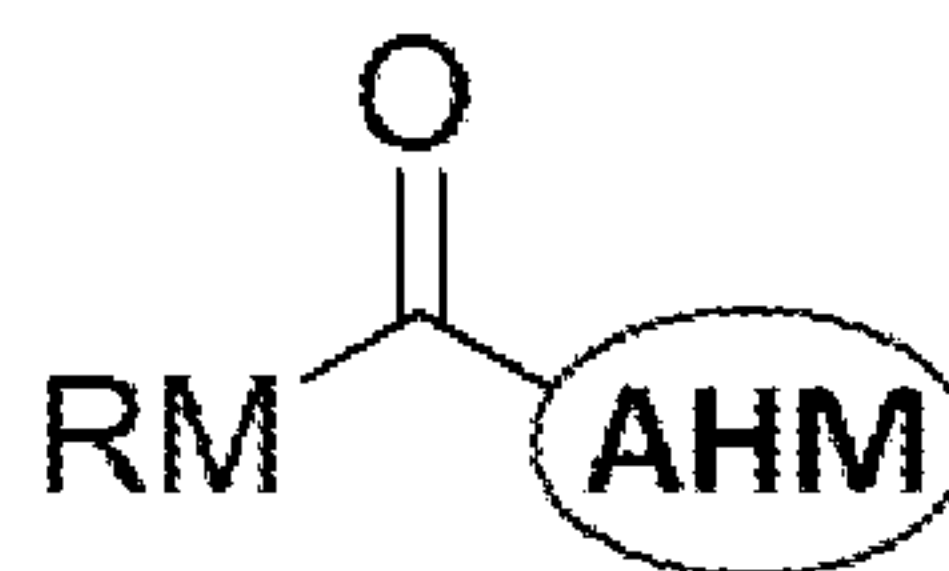
wherein:

RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and

F is a small molecule fragment moiety.

13. The method of claim 12, wherein the Michael acceptor moiety comprises an alkene or an alkyne moiety.
14. The method of claim 12, wherein F is obtained from a compound library.
15. The method of claim 14, wherein the compound library comprises ChemBridge fragment library, Pyramid Platform Fragment-Based Drug Discovery, Maybridge fragment library, FRGx from AnalytiCon, TCI-Frag from AnCoreX, Bio Building Blocks from ASINEX, BioFocus 3D from Charles River, Fragments of Life (FOL) from Emerald Bio, Enamine Fragment Library, IOTA Diverse 1500, BIONET fragments library, Life Chemicals Fragments Collection, OTAVA fragment library, Prestwick fragment library, Selcia fragment library, TimTec fragment-based library, Allium from Vitas-M Laboratory, or Zenobia fragment library.
16. The method of any one of the claims 12-15, wherein F is a small molecule fragment moiety illustrated in Fig. 3.

17. The method of claim 1, wherein the cysteine-reactive probe is a cysteine-reactive probe of Formula (II):



Formula (II)

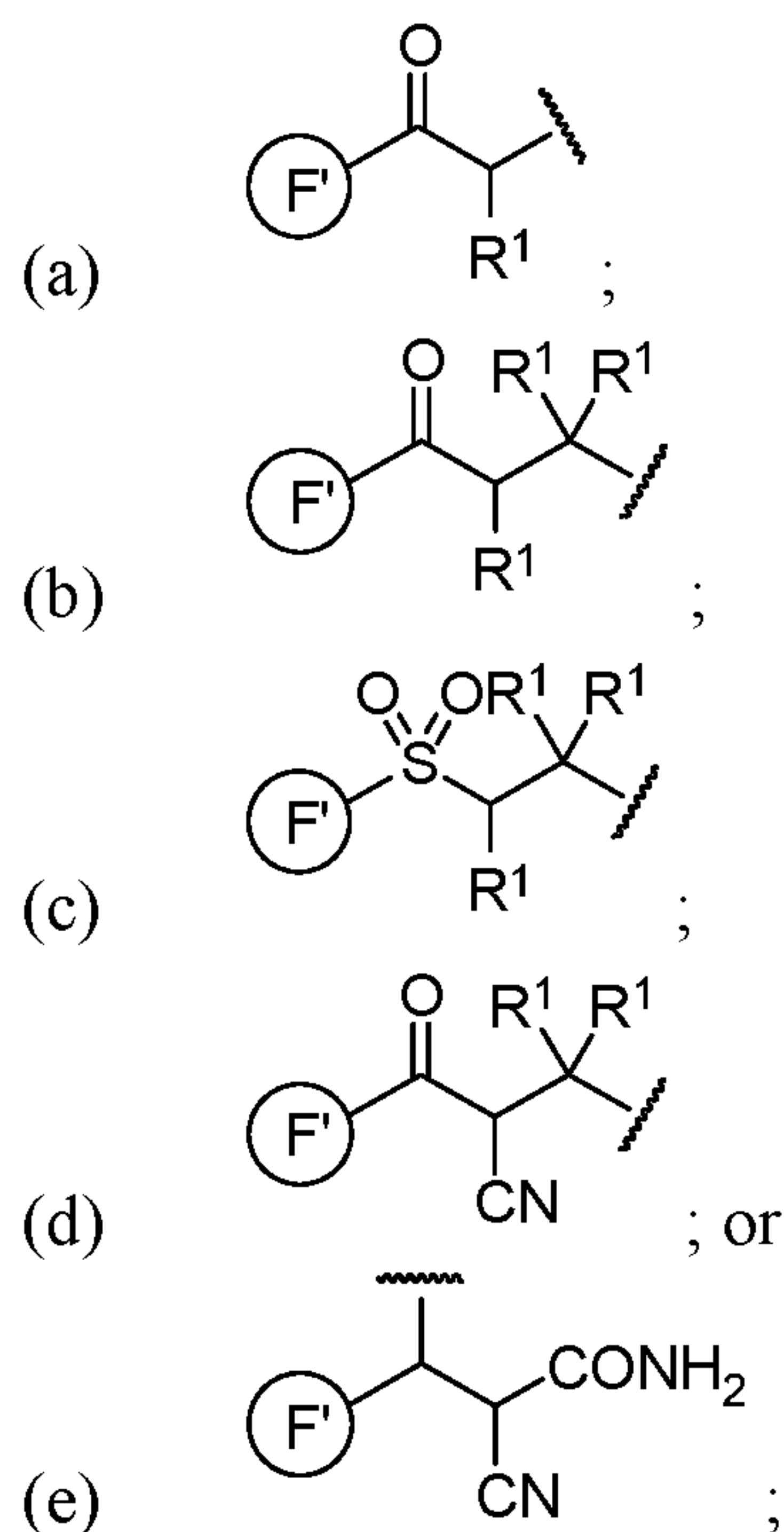
wherein:

RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond to the thiol group of a cysteine residue; and

AHM is an affinity handle moiety.

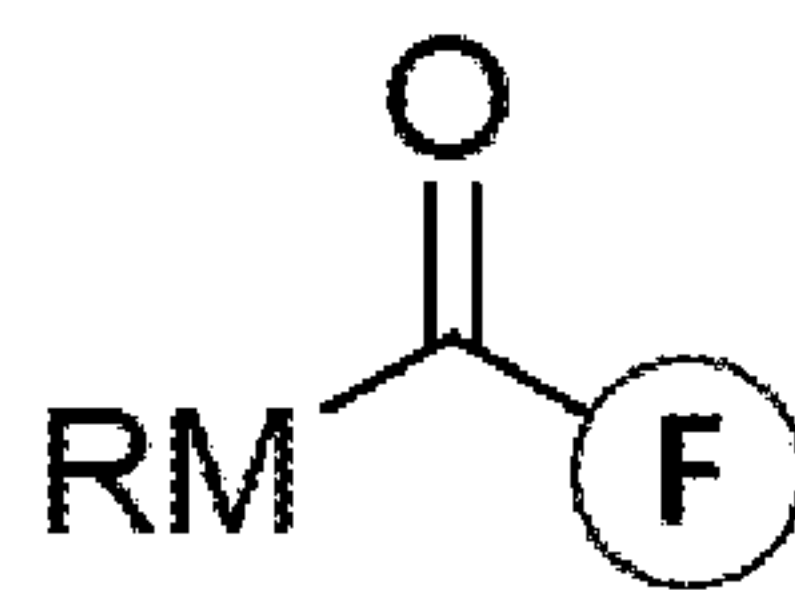
18. The method of claim 17, wherein the Michael acceptor moiety comprises an alkene or an alkyne moiety.
19. The method of claim 17, wherein the affinity handle moiety comprises an affinity handle and a binding moiety that facilitates covalent interaction of the cysteine-reactive probe to a cysteine residue of a cysteine-containing protein.
20. The method of claim 19, wherein the binding moiety is a small molecule fragment obtained from a compound library.
21. The method of claim 17 or 19, wherein the affinity handle comprises a carbodiimide, N-hydroxysuccinimide (NHS) ester, imidoester, pentafluorophenyl ester, hydroxymethyl phosphine, maleimide, haloacetyl, pyridyl disulfide, thiosulfonate, vinylsulfone, hydrazide, alkoxyamine, alkyne, azide, or isocyanate group.
22. The method of any one of the claims 17, or 19-21, wherein the affinity handle is further conjugated to an affinity ligand.
23. The method of claim 22, wherein the affinity ligand comprises a chromophore, a labeling group, or a combination thereof.
24. The method of claim 23, wherein the chromophore comprises non-fluorochrome chromophore, quencher, an absorption chromophore, fluorophore, organic dye, inorganic dye, metal chelate, or a fluorescent enzyme substrate.
25. The method of claim 23, wherein the labeling group is a biotin moiety, a streptavidin moiety, bead, resin, a solid support, or a combination thereof.
26. The method of claim 17, wherein the cysteine-reactive probe is a cysteine-reactive probe illustrated in Fig. 3.

27. The method of claim 1, wherein the proteomic analysis means comprises a mass spectroscopy method.
28. The method of claim 1, wherein the identifying in step c) further comprises
 - i. locating a first value assigned to a cysteine containing protein from the first group of cysteine-reactive probe-protein complex and a second value of the same cysteine containing protein from the second group of cysteine-reactive probe-protein complex; and
 - ii. calculating a ratio between the two values assigned to the same cysteine containing protein.
29. The method of claim 28, wherein the ratio of greater than 2 indicates that the cysteine containing protein is a candidate for interacting with the small molecule fragment.
30. The method of claim 28, wherein the identifying in step c) further comprises calculating a percentage of inhibition of the cysteine-reactive probe to the cysteine containing protein.
31. The method of claim 30, wherein the percentage of inhibition of greater than 50%, 60%, 70%, 80%, 90%, or at 100% indicates that the cysteine containing protein is a candidate for interacting with the small molecule fragment.
32. The method of any one of the claims 1-31, wherein the method is an *in situ* method.
33. The method of any one of the claims 1-32, wherein the cysteine-reactive probe is not 4-hydroxynonenal or 15-deoxy- $\Delta^{12,14}$ -prostaglandin J2.
34. A modified cysteine containing protein comprising a small molecule fragment having a covalent bond to a cysteine residue of a cysteine containing protein, wherein the small molecule fragment has a molecular weight of about 150 Dalton or higher.
35. The modified cysteine containing protein of claim 34, wherein the cysteine containing protein comprises a cysteine residue site denoted in Table 3.
36. The modified cysteine containing protein of claim 34, wherein the cysteine containing protein comprises a protein sequence illustrated in Table 1, Table 2, Table 8, Table 9, Table 10A, Table 10B, Table 10C, Table 10D or Table 10E.
37. The modified cysteine containing protein of claim 34, wherein the cysteine containing protein is about 20, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000 amino acid residues in length or more.
38. The modified cysteine containing protein of claim 34, wherein the cysteine residue of the modified cysteine containing protein has the structure SR, wherein R is selected from:



wherein R^1 is H, C1-C3 alkyl, or aryl; and F' is the small molecule fragment moiety.

39. The modified cysteine containing protein of claim 34 or 38, wherein the small molecule fragment has a molecular weight of about 175, 200, 225, 250, 275, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 Dalton, or higher.
40. The modified cysteine containing protein of any one of the claims 34, 38 or 39, wherein the molecular weight of the small molecule fragment is calculated based on carbon and hydrogen atoms and optionally further based on nitrogen, oxygen and/or sulfur atoms.
41. The modified cysteine containing protein of claim 34, wherein the modified cysteine containing protein is selected from IDH2, caspase-8, caspase-10 or PRMT1.
42. The modified cysteine containing protein of claim 34, wherein IDH2 is modified at cysteine position 308.
43. The modified cysteine containing protein of claim 34, wherein caspase-8 is modified at cysteine position 360.
44. The modified cysteine containing protein of claim 34, wherein caspase-10 exist in the proform and is modified at cysteine position 401.
45. The modified cysteine containing protein of claim 34, wherein PRMT1 is modified at cysteine position 109.
46. The modified cysteine containing protein of claim 34, wherein the small molecule fragment is a small molecule fragment of Formula (I):



Formula (I)

wherein:

RM is a reactive moiety selected from a Michael acceptor moiety, a leaving group moiety, or a moiety capable of forming a covalent bond with the thiol group of a cysteine residue; and

F is a small molecule fragment moiety.

47. The modified cysteine containing protein of claim 46, wherein the Michael acceptor moiety comprises an alkene or an alkyne moiety.
48. The modified cysteine containing protein of claim 46, wherein F is obtained from a compound library.
49. The modified cysteine containing protein of claim 46, wherein F is a small molecule fragment moiety illustrated in Fig. 3.
50. The modified cysteine containing protein of claim 46, wherein F further comprises a linker moiety that connects F to the carbonyl moiety.
51. The modified cysteine containing protein of any one of the claims 34-50, wherein the small molecule fragment binds irreversibly to the cysteine containing protein.
52. The modified cysteine containing protein of any one of the claims 34-50, wherein the small molecule fragment binds reversibly to the cysteine containing protein.
53. A method of screening a small molecule fragment for interaction with a cysteine containing protein, comprising:
 - a) harvesting a set of cysteine-reactive probe-protein complexes from a sample comprising a first cell solution treated with a small molecule fragment and a cysteine reactive probe wherein the cysteine-reactive probe comprises a reactive moiety capable of forming a covalent bond with a cysteine residue located on the cysteine containing protein;
 - b) analyzing the set of cysteine-reactive probe-protein complexes by a proteomic analysis means; and
 - c) based on step b), identifying the small molecule fragment as interacting with the cysteine containing protein.
54. The method of claim 53, further comprising determining a value of each of the cysteine containing protein from the set of cysteine-reactive probe-protein complexes prior to

identifying the small molecule fragment as interacting with the cysteine containing protein, wherein the value is determined based on the proteomic analysis means of step b).

55. The method of claim 53, wherein the cysteine containing protein is a protein illustrated in Table 3.
56. The method of claim 53, wherein the cysteine containing protein is a protein illustrated in Table 1, Table 2, Table 8, Table 9, Table 10A, Table 10B, Table 10C, Table 10D or Table 10E.

Fig. 1

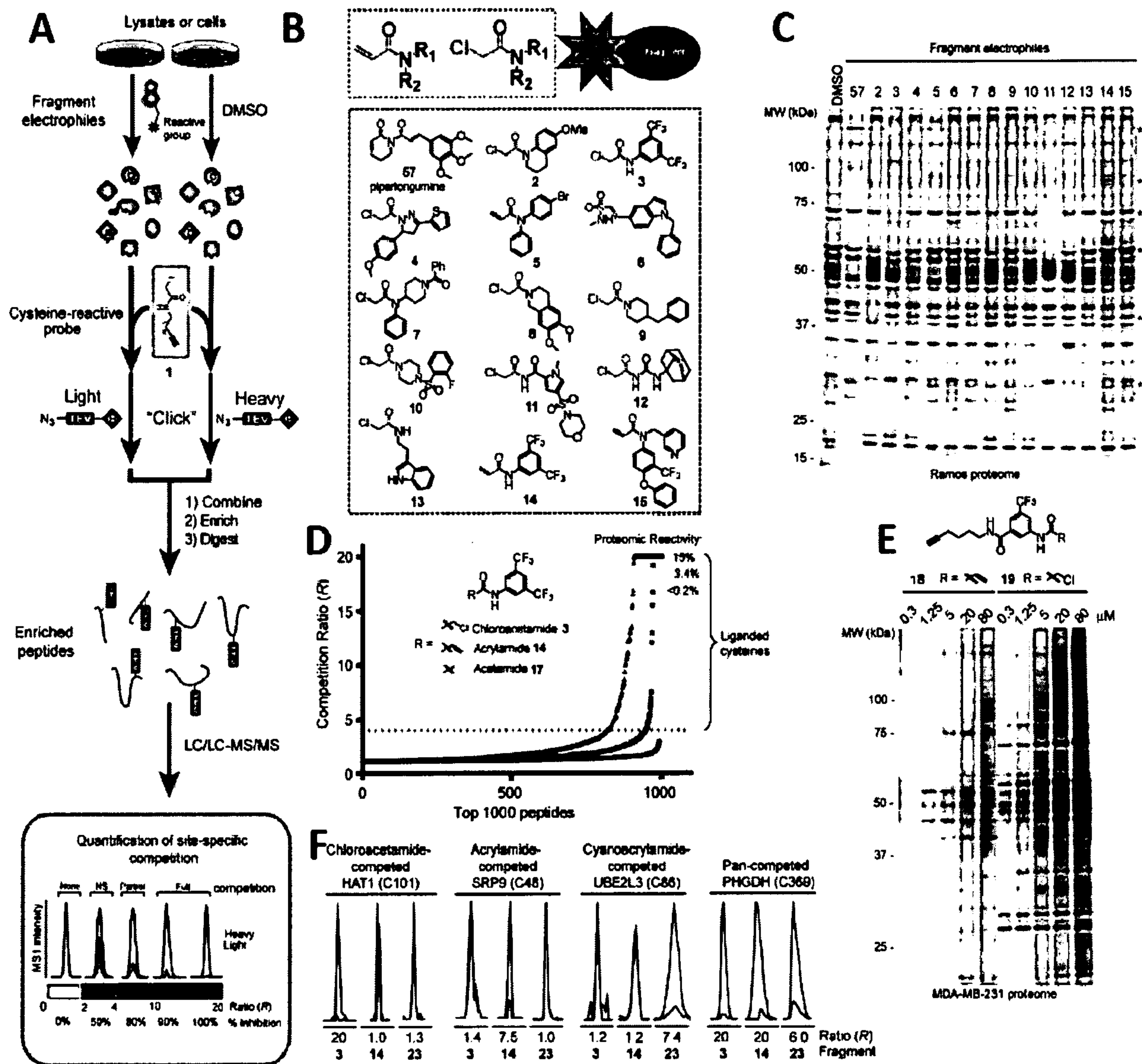


Fig. 2

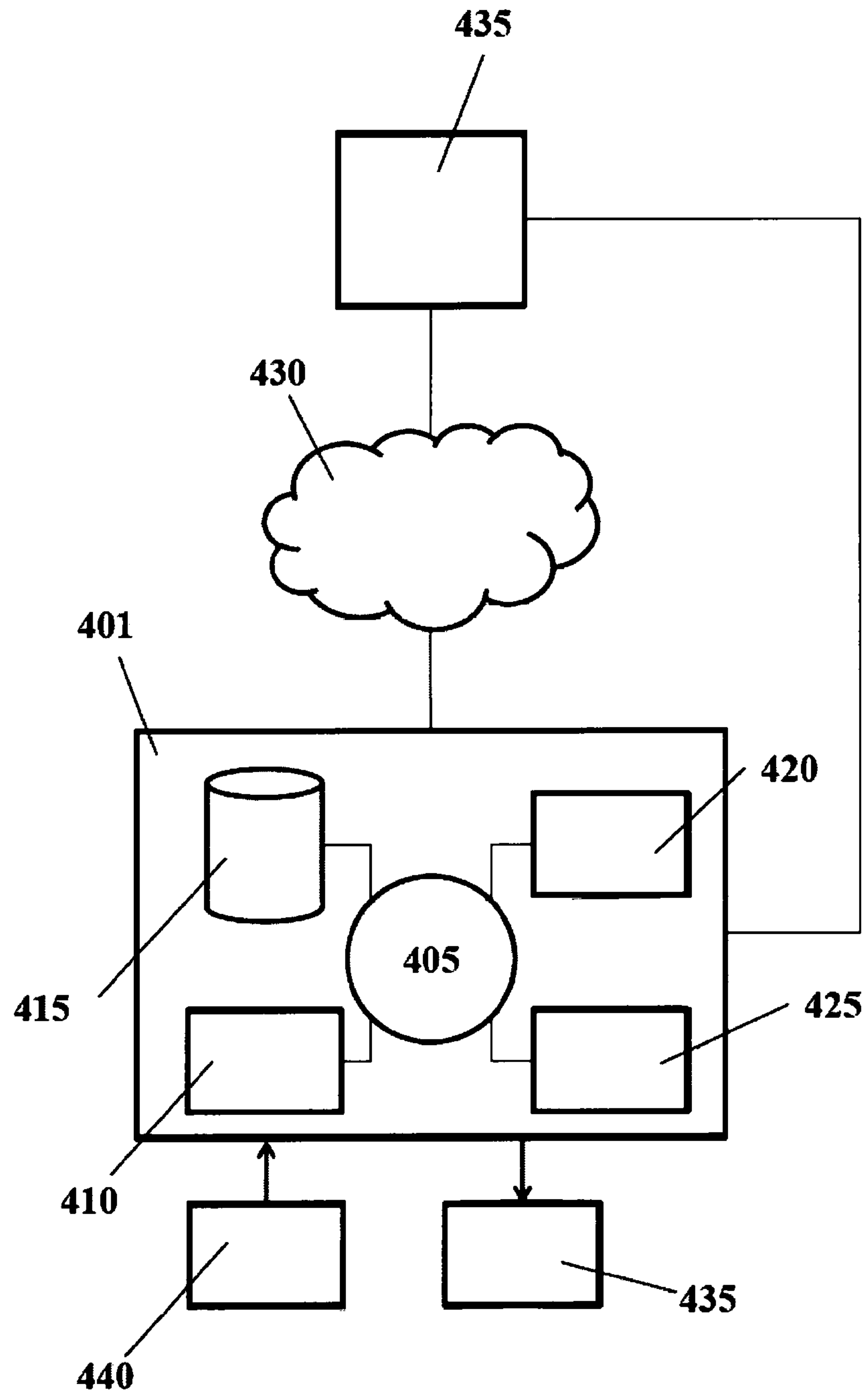


Fig. 3

Fragment electrophiles screened by isoTOP-ABPP

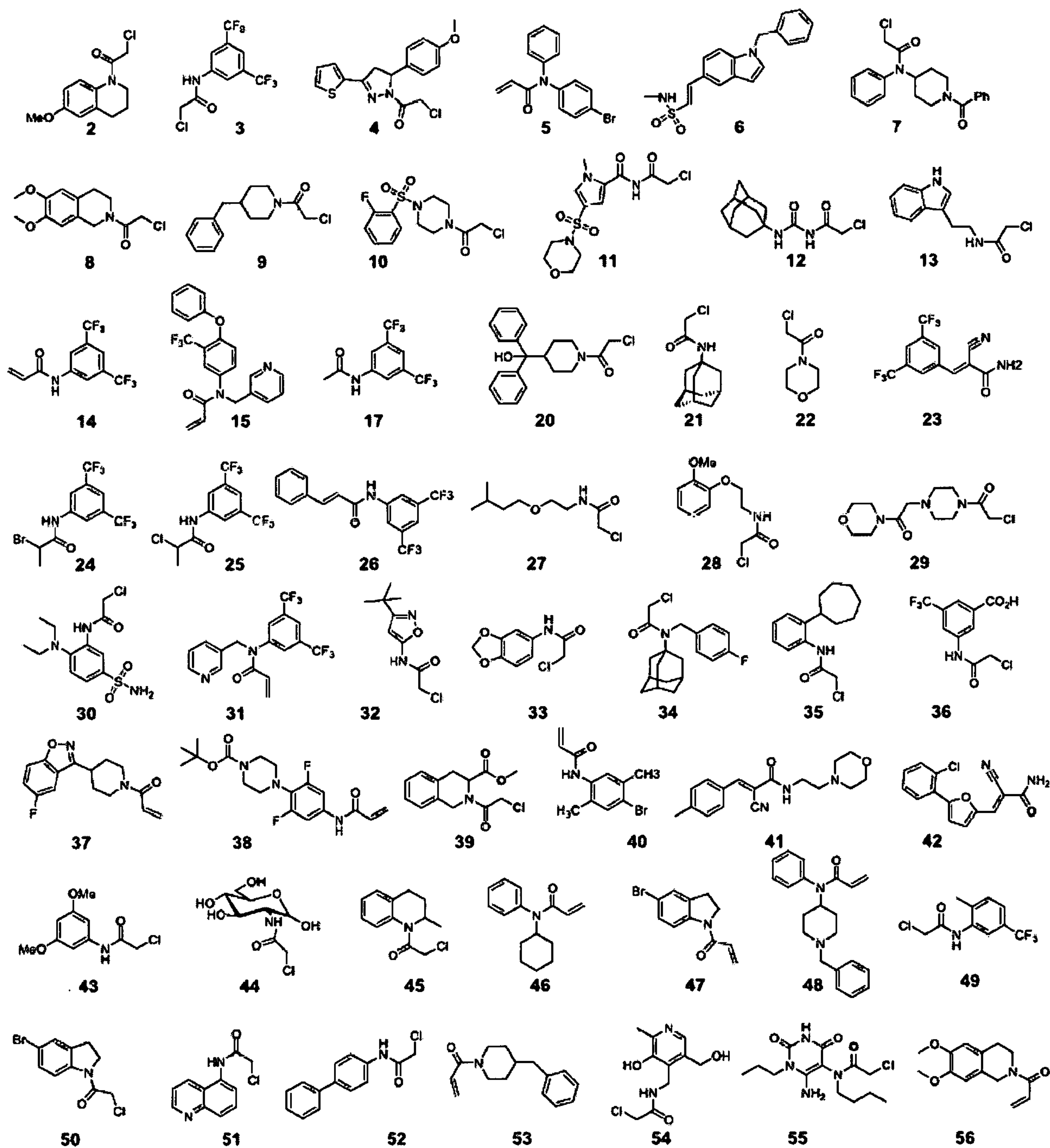


FIG. 3
Continued

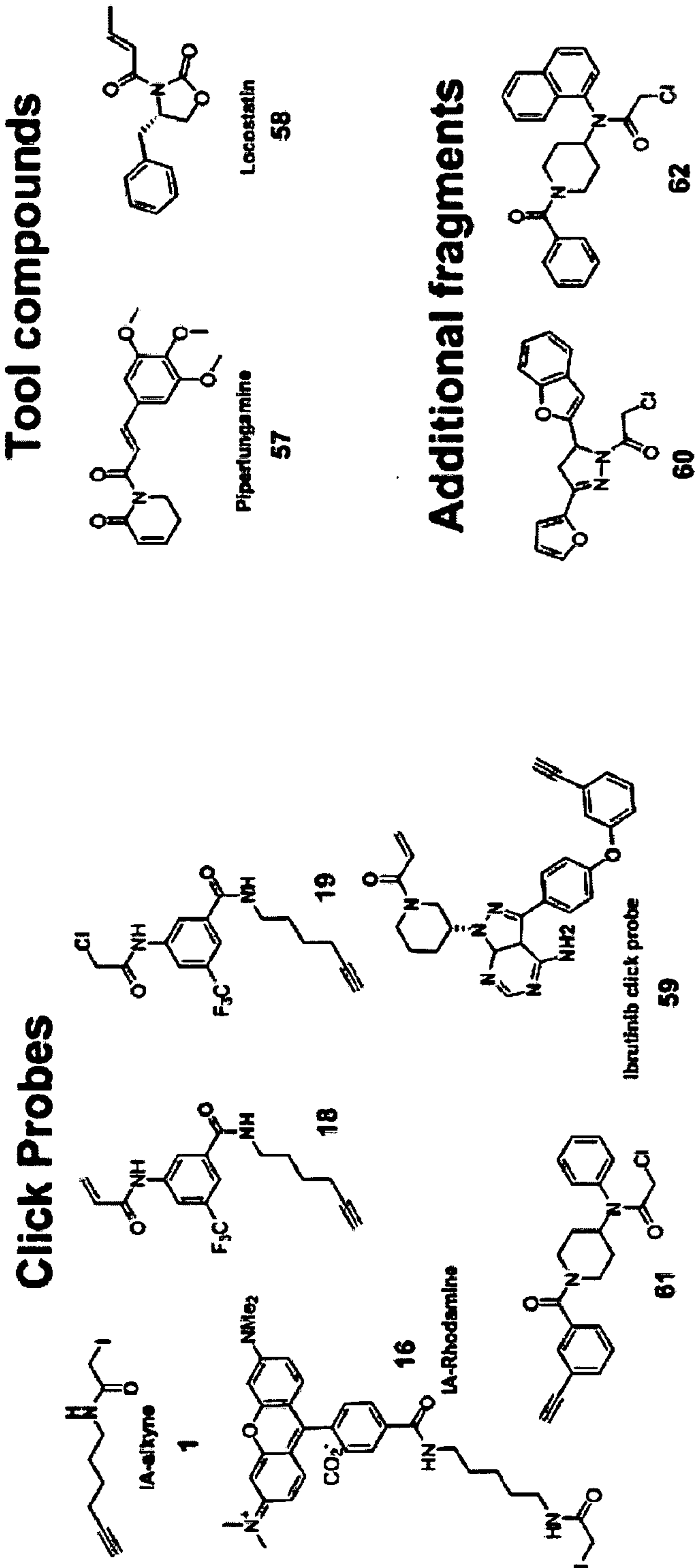
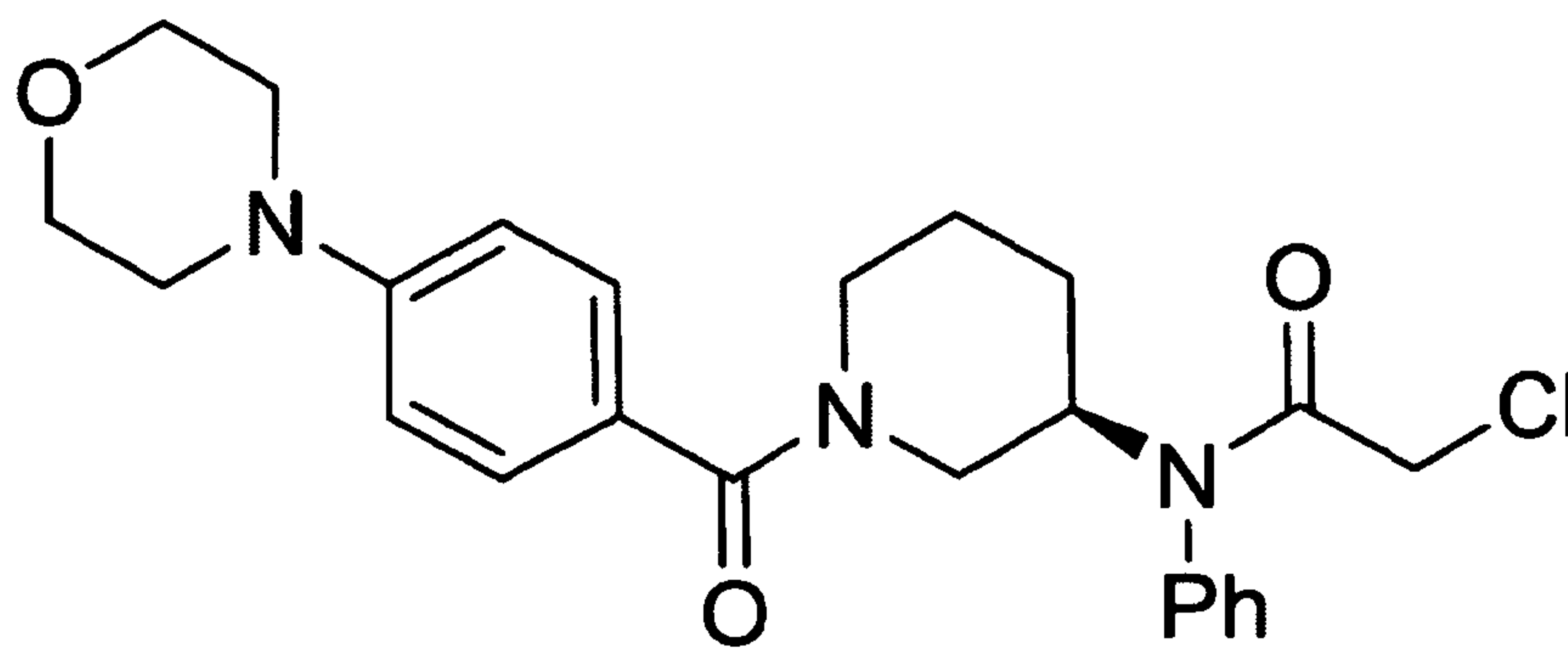
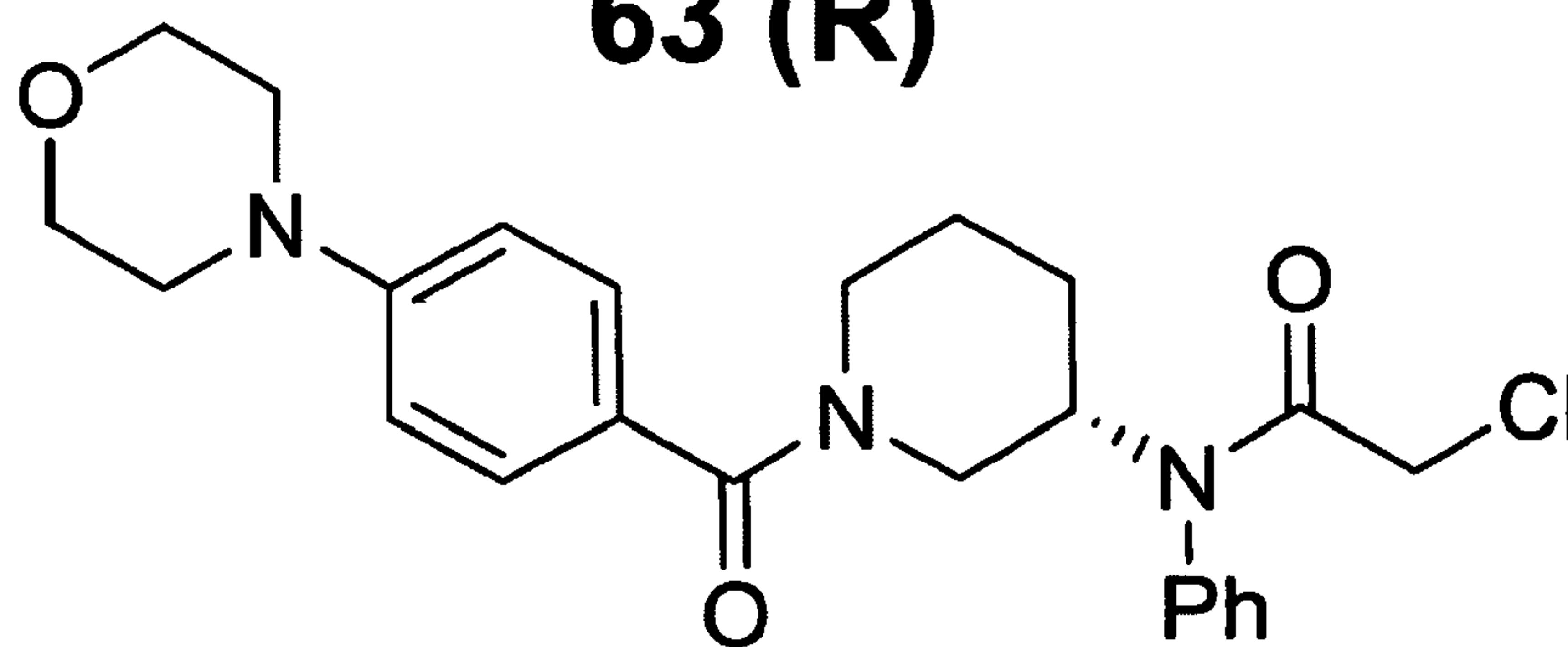


Fig. 3
Continued



63 (R)



63 (S)

A

of cysteines

of quantification events

B

Proteomic reactivity (%)

locostatin
piperlongumine

Fragment

C

Proteomic reactivity (%)

500 μ M
25 μ M

Fragment

D

Proteomic reactivity (Ramos %)

Proteomic reactivity (MDA-MB-231 %)

$R^2 = 0.86$

E

Proteomic reactivity (%)

Fragment

F

GSH consumed (%)

Fragment

Increasing proteomic reactivity values

G

Proteomic reactivity (%)

Fragment

$R =$

3, 23, 14, 24, 25, 26, 17

FIG. 5

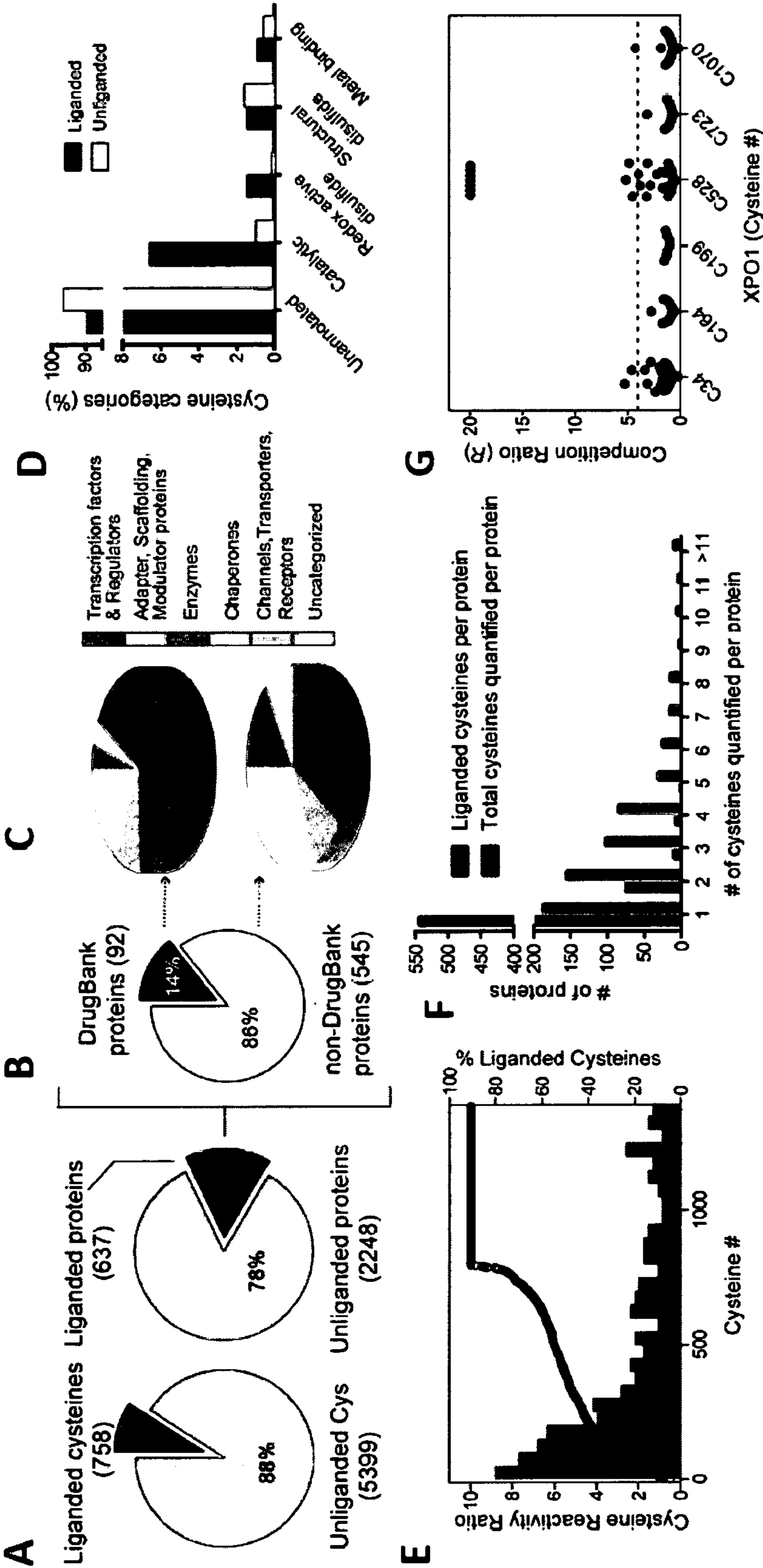


Fig. 6

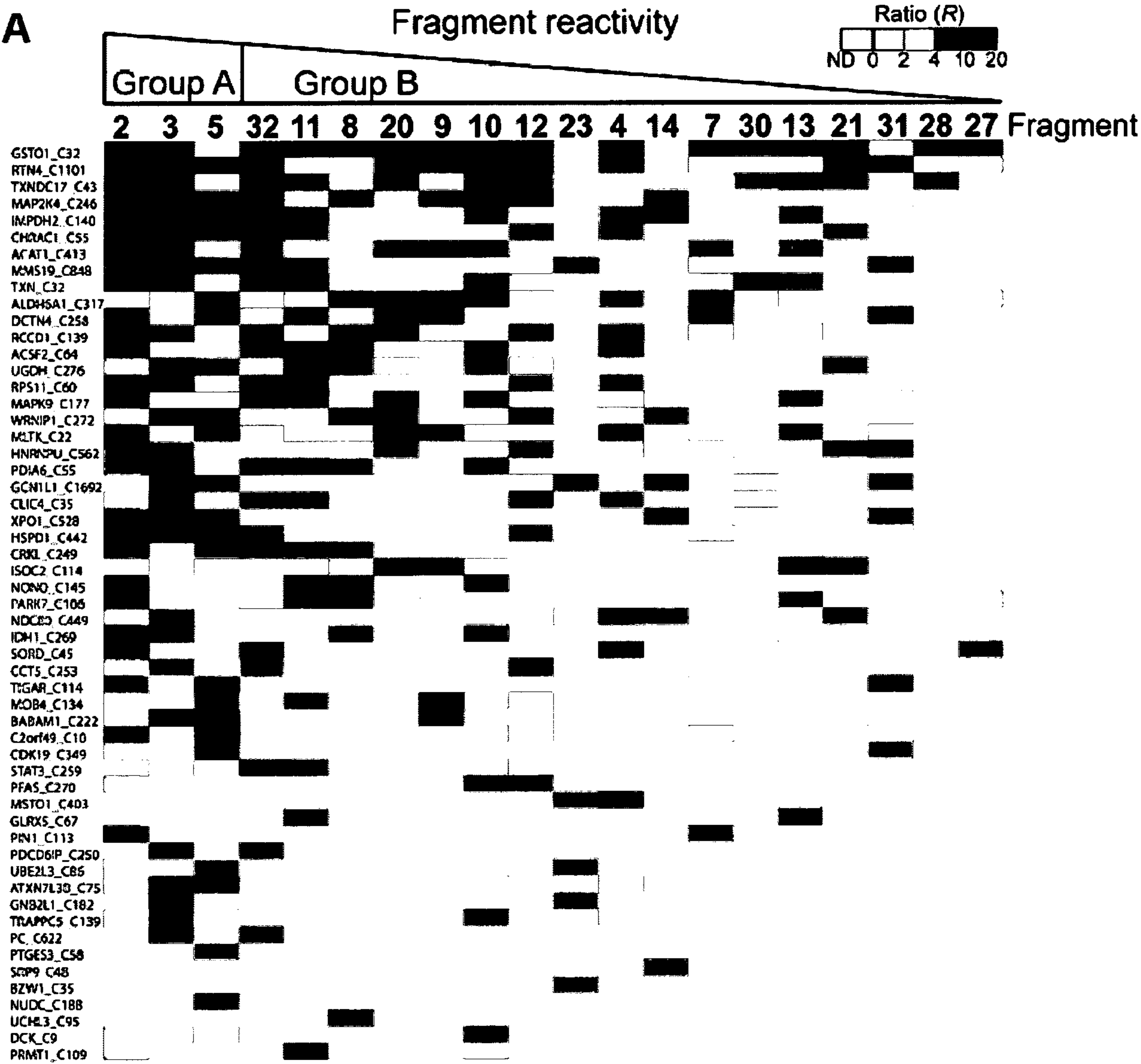
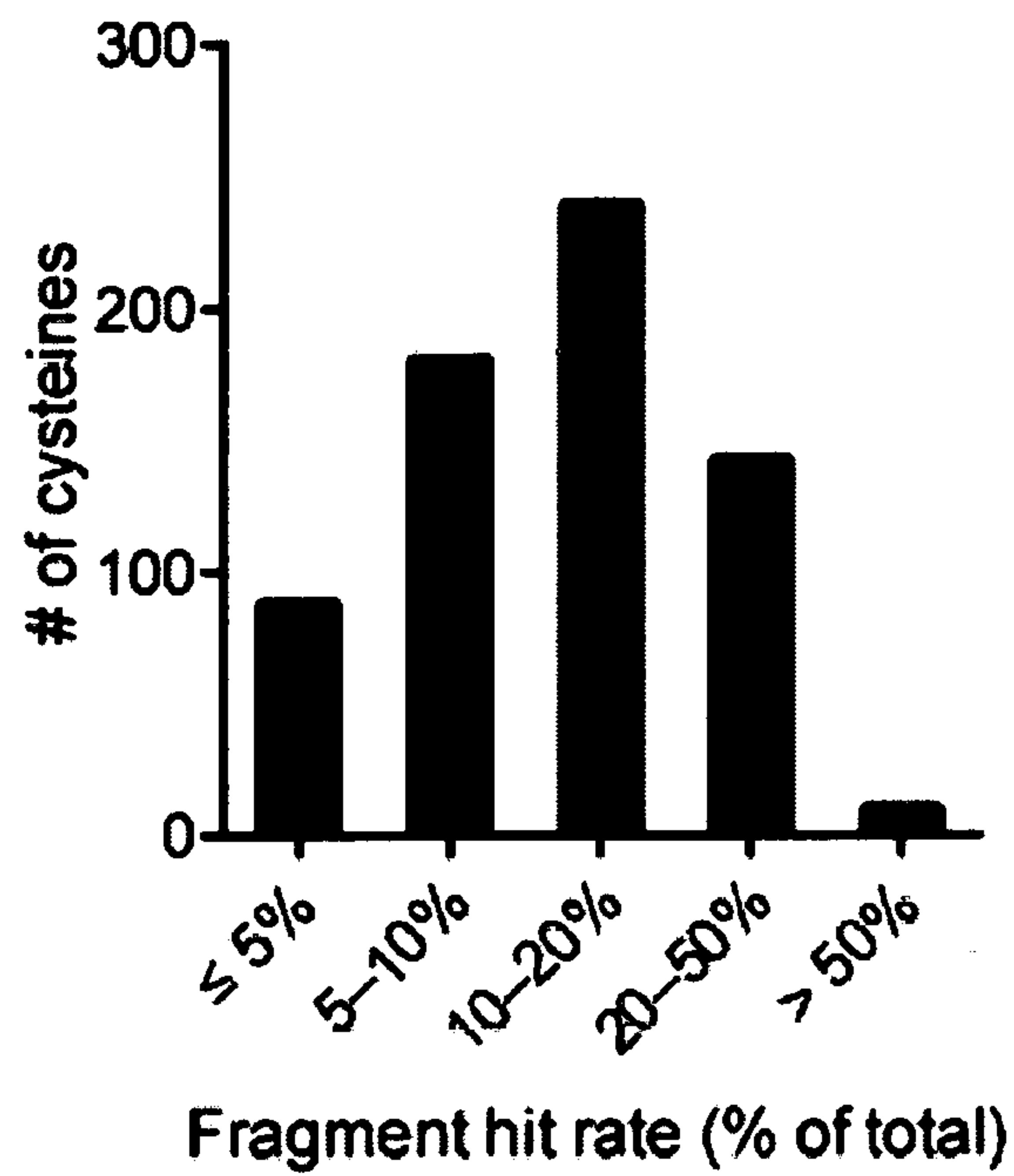


Fig. 6
Continued

B



C

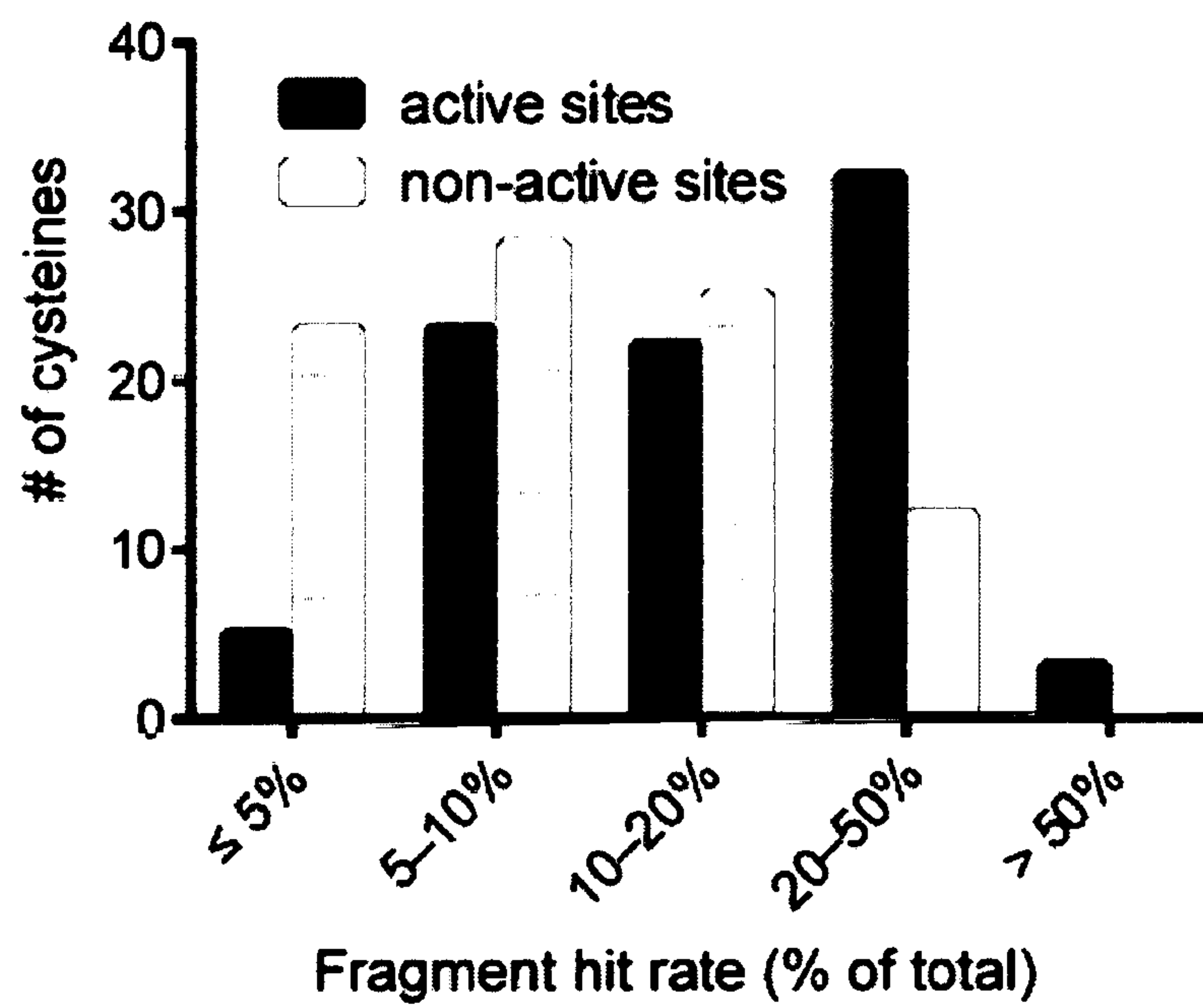
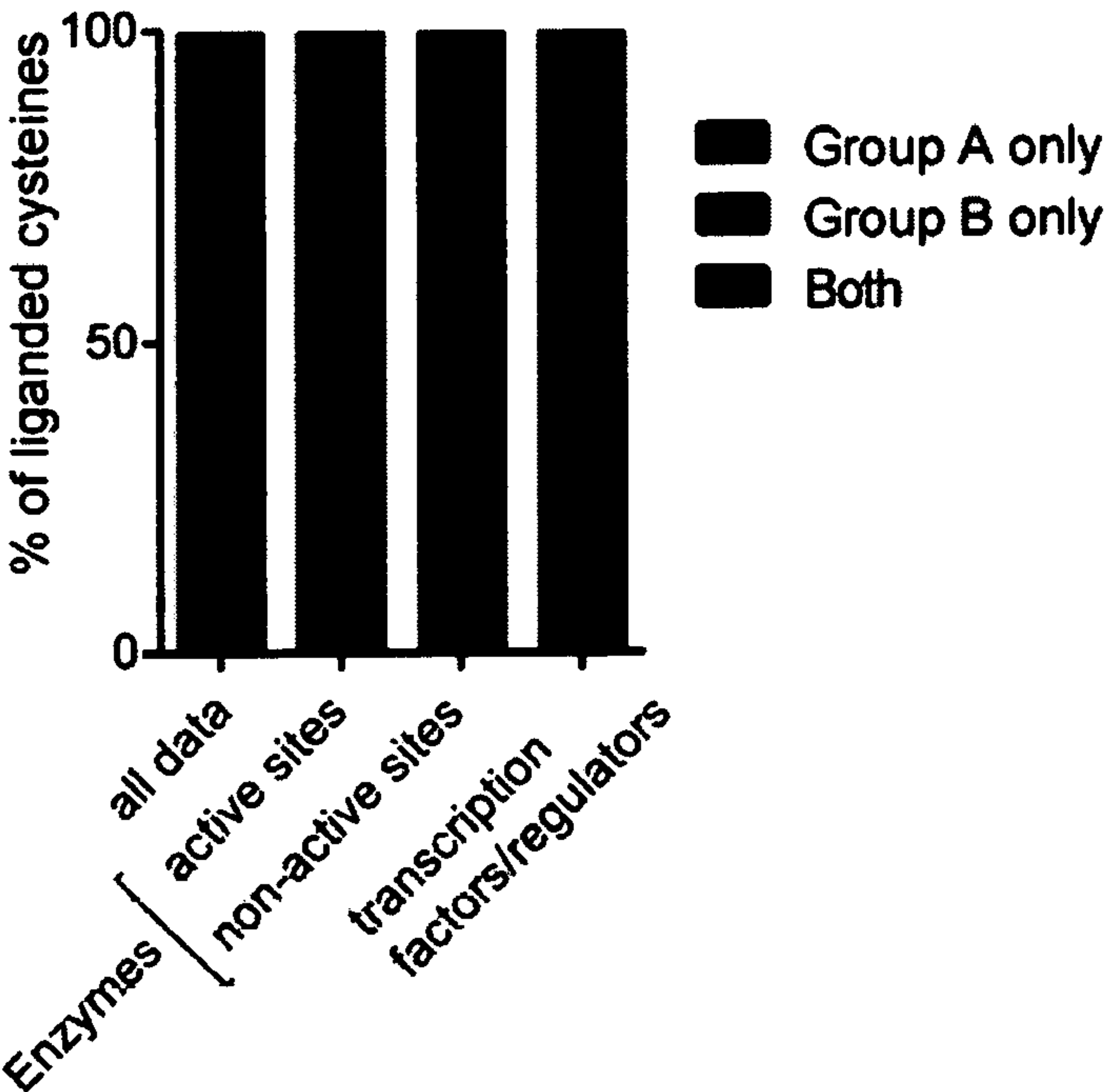
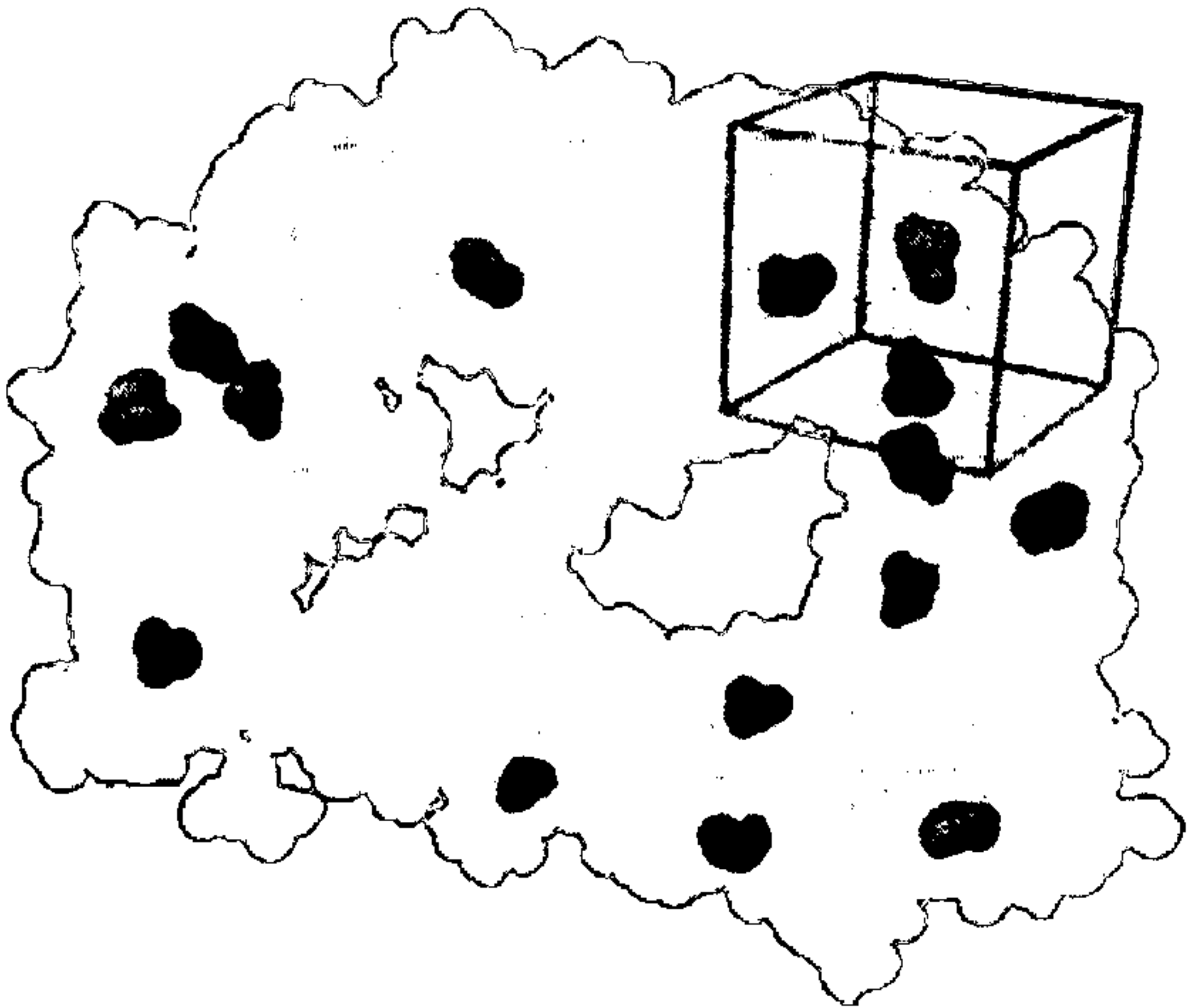


Fig. 6
Continued

D



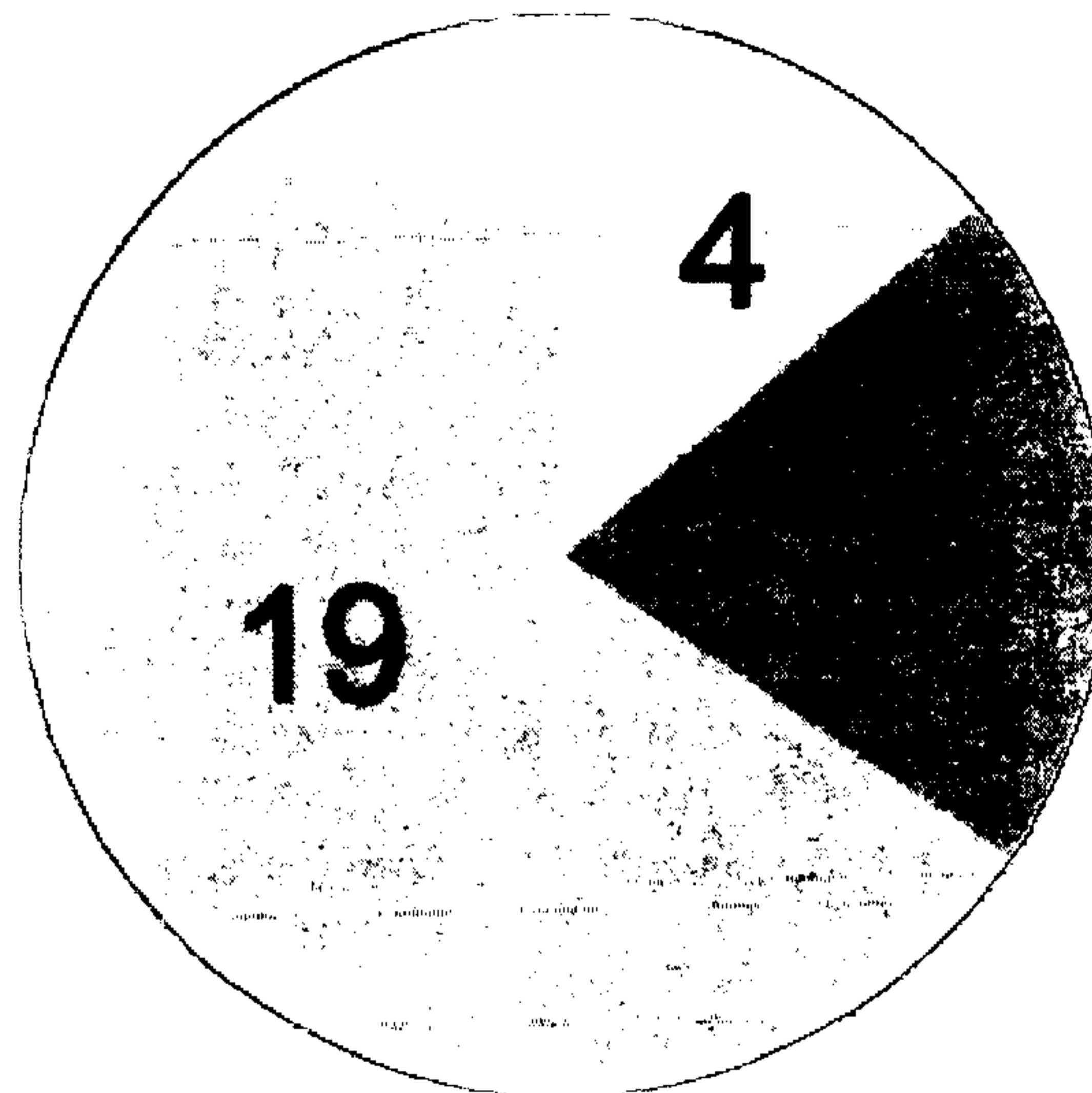
E



XPO1 cysteine (events)	isoTOP-ABPP		Docking prediction
	Liganded	Quantified	
Green (1)	+	+	Top result
Orange (4)	+ (1)	+ (4)	Ligandable
Blue (1)	-	+	Not ligandable
Red (2)	-	-	Ligandable
Gray (6)	-	-	Not ligandable

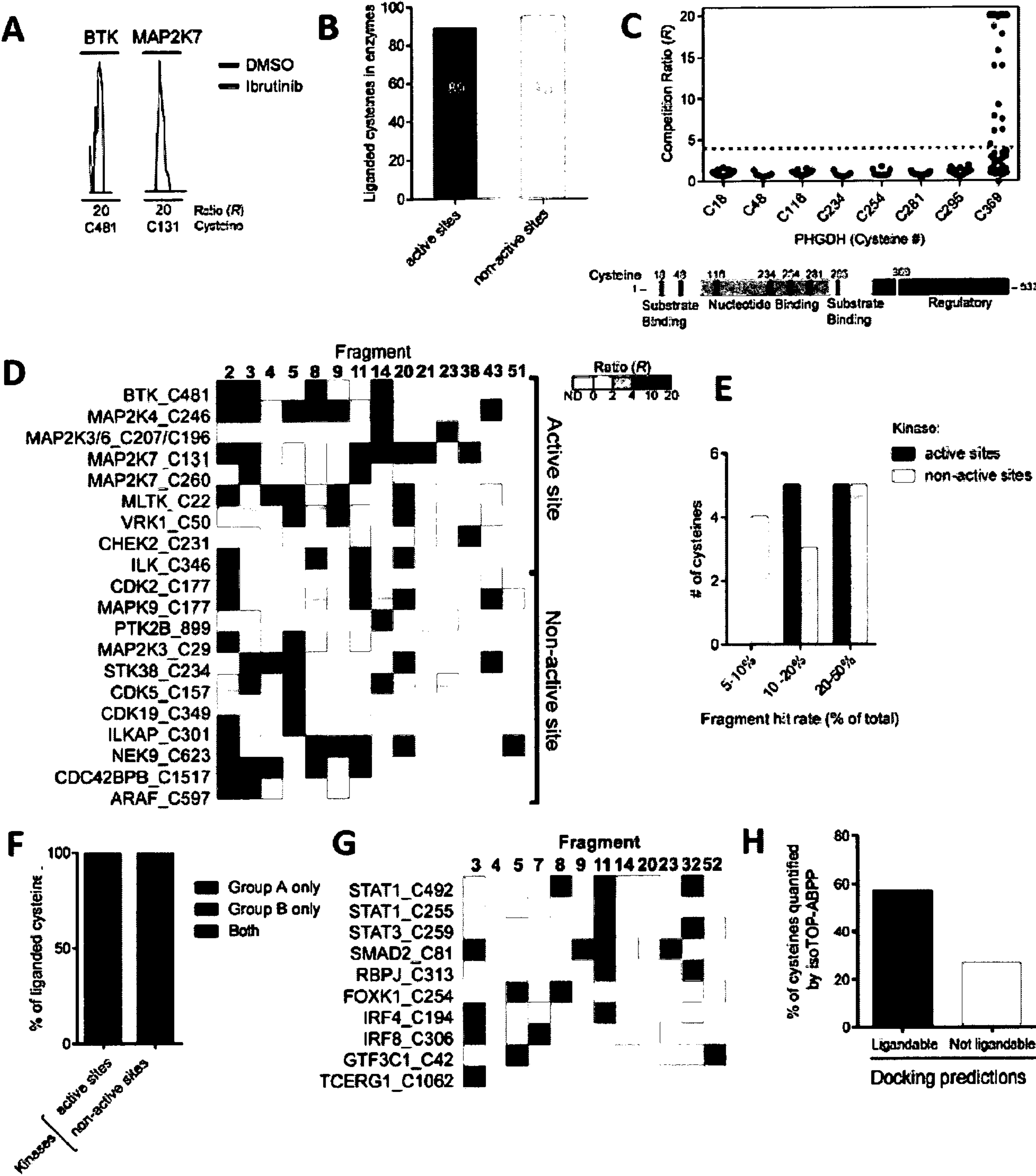
Fig. 6
Continued

F



☐ Top result
☒ Second result
☐ Incorrect prediction

Fig. 7



[illegible]

Fig. 9

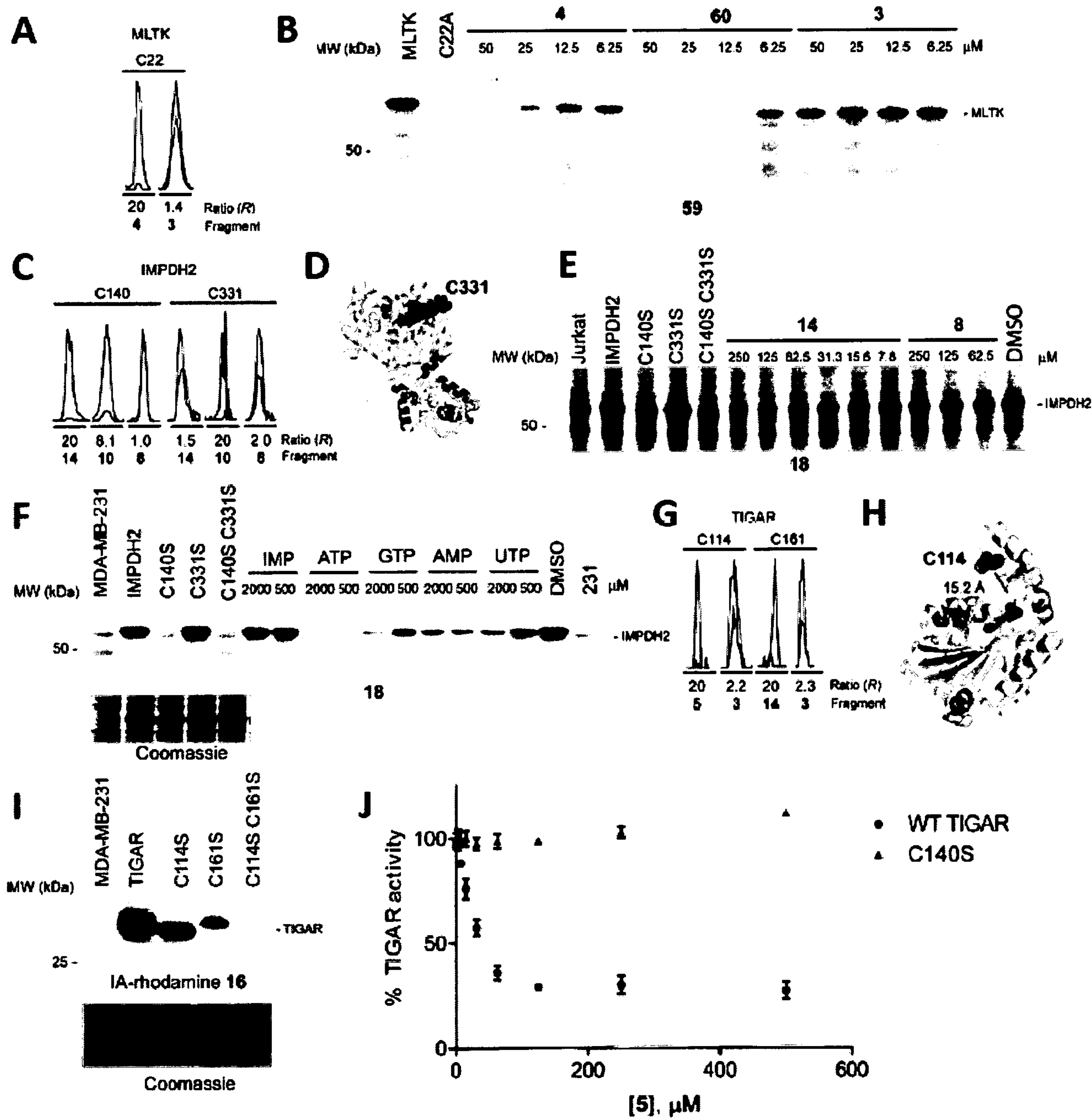


FIG. 10

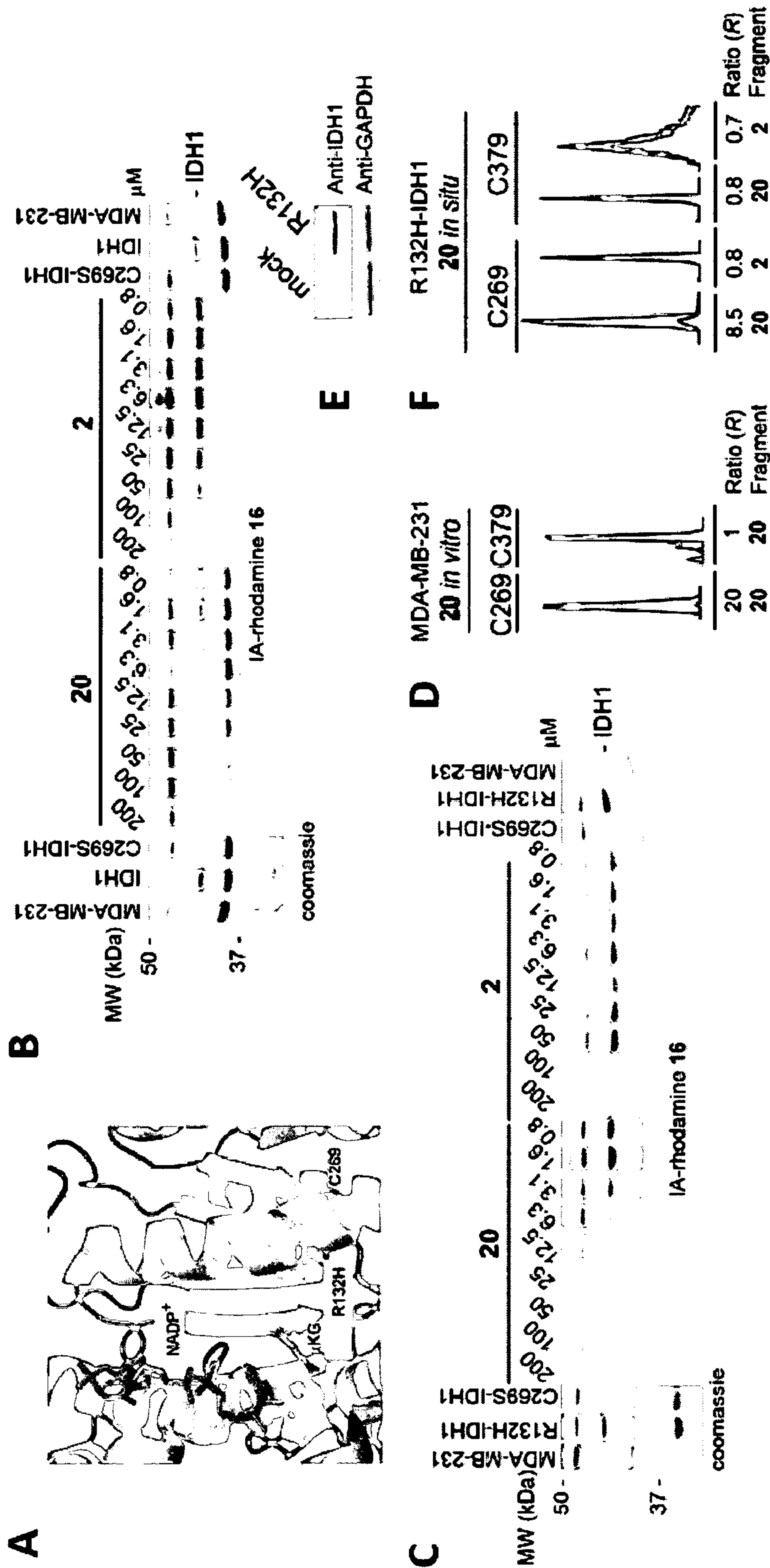


FIG. 11

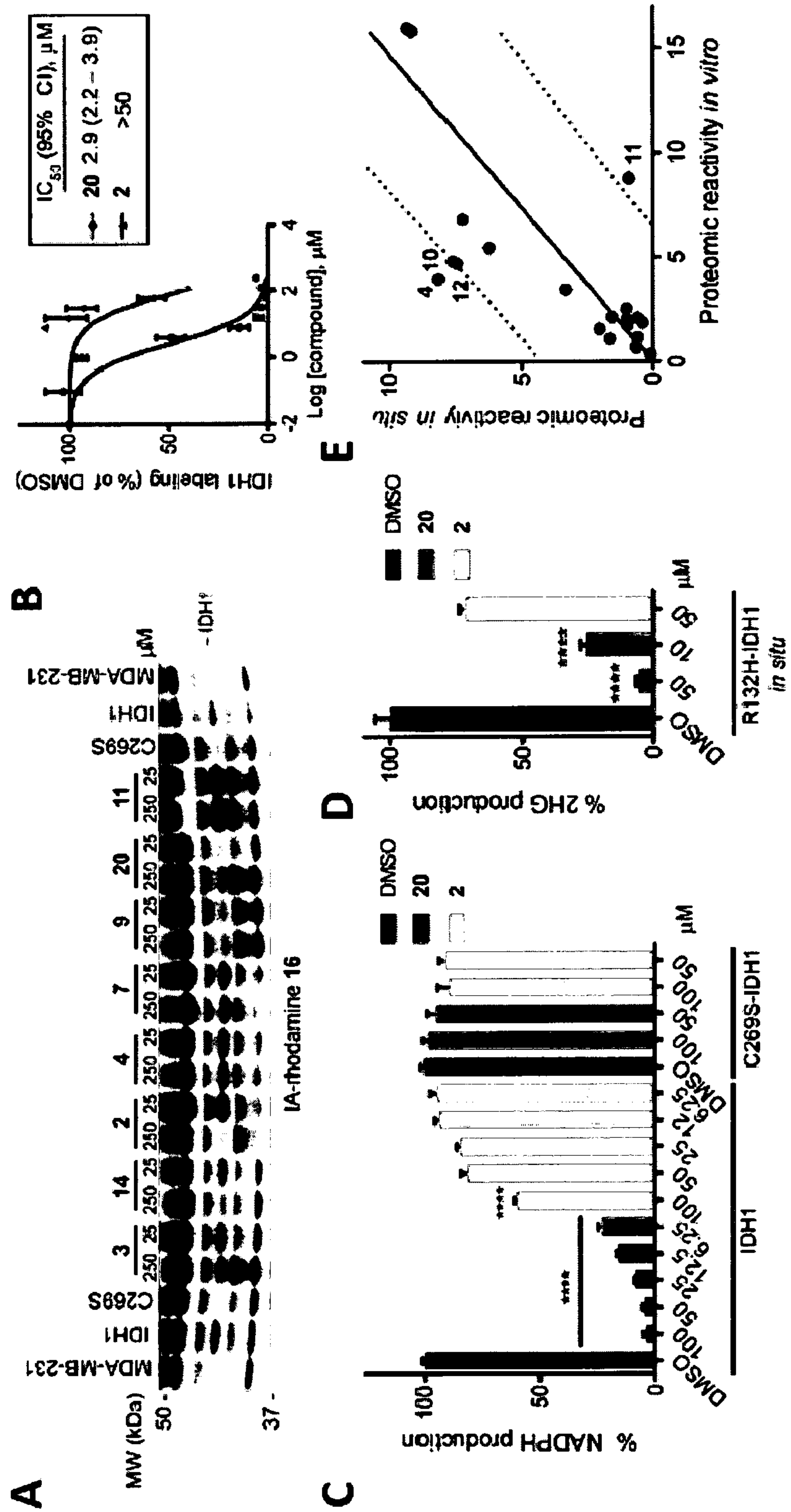


FIG. 11
Continued

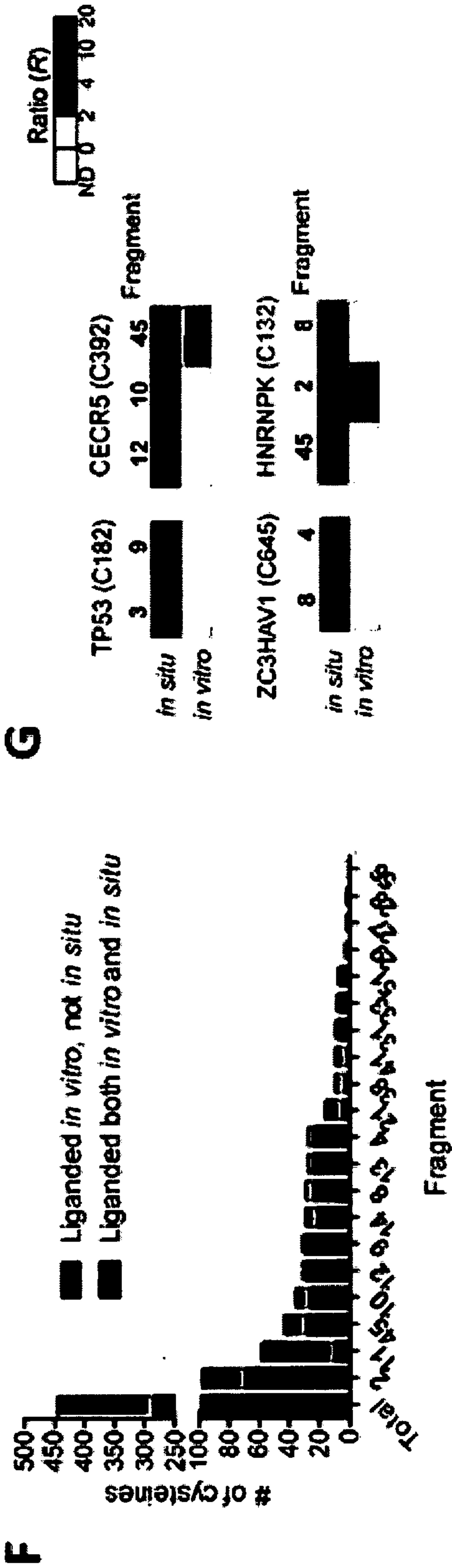


Fig. 12

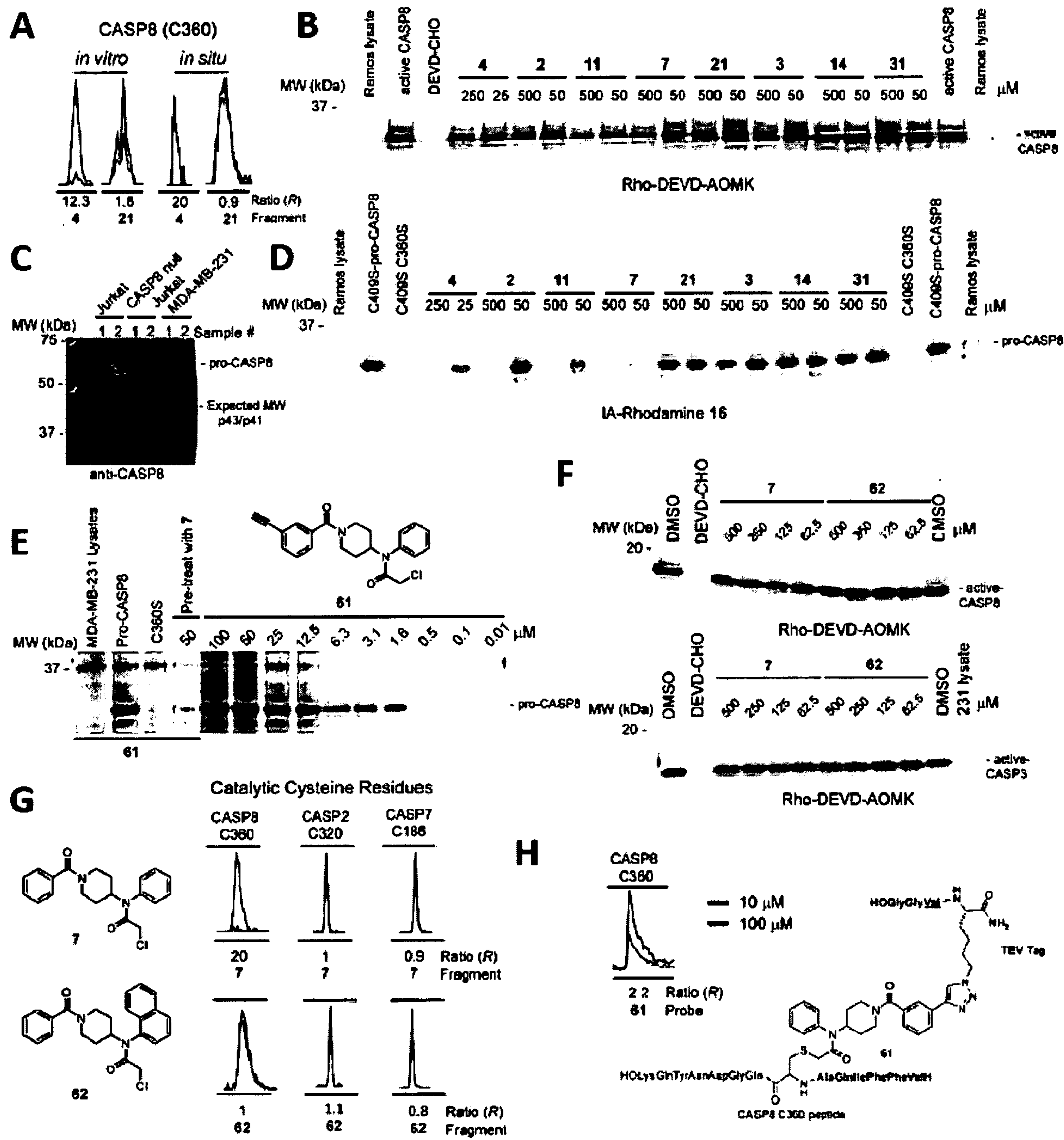


Fig. 13

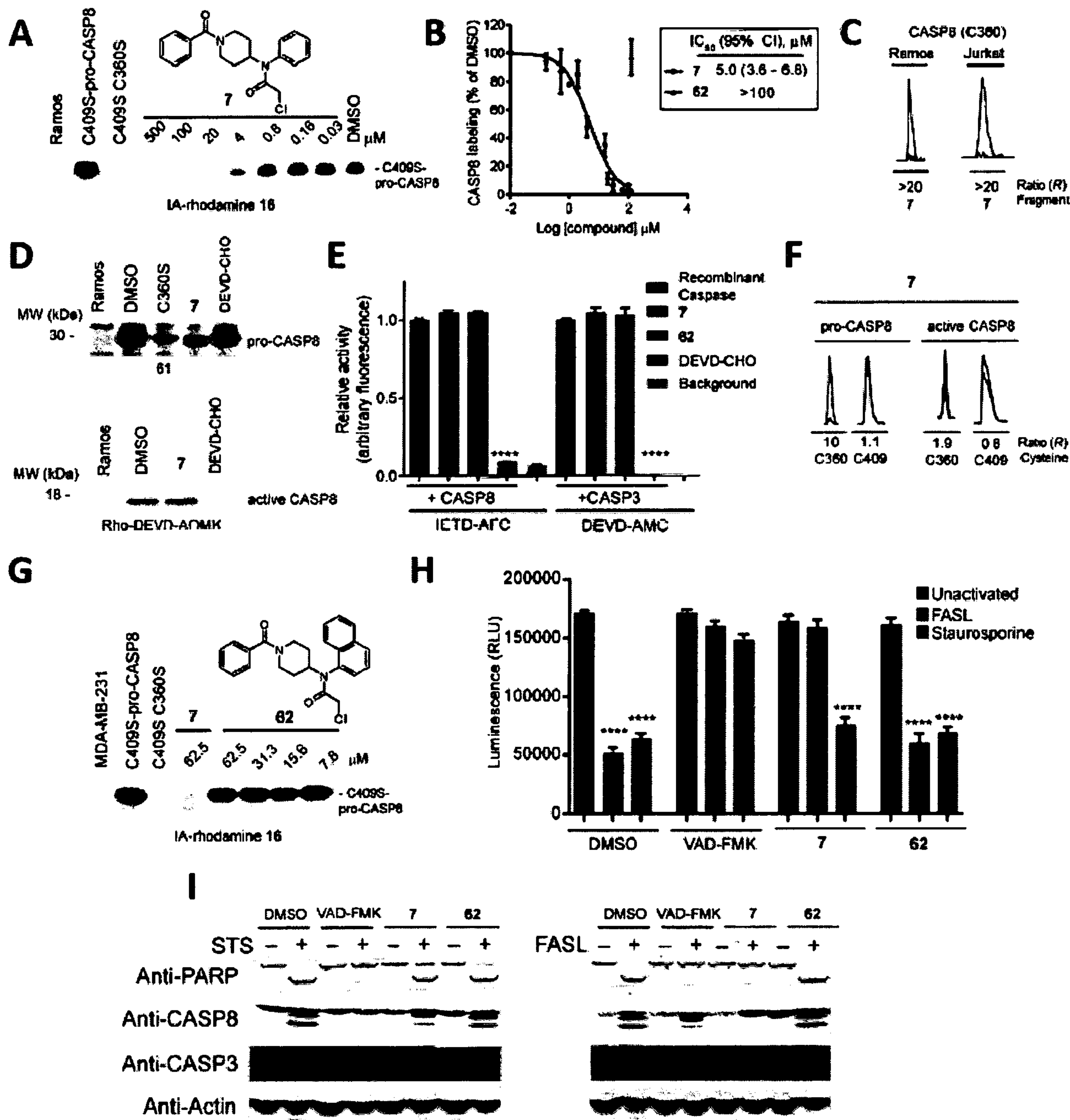
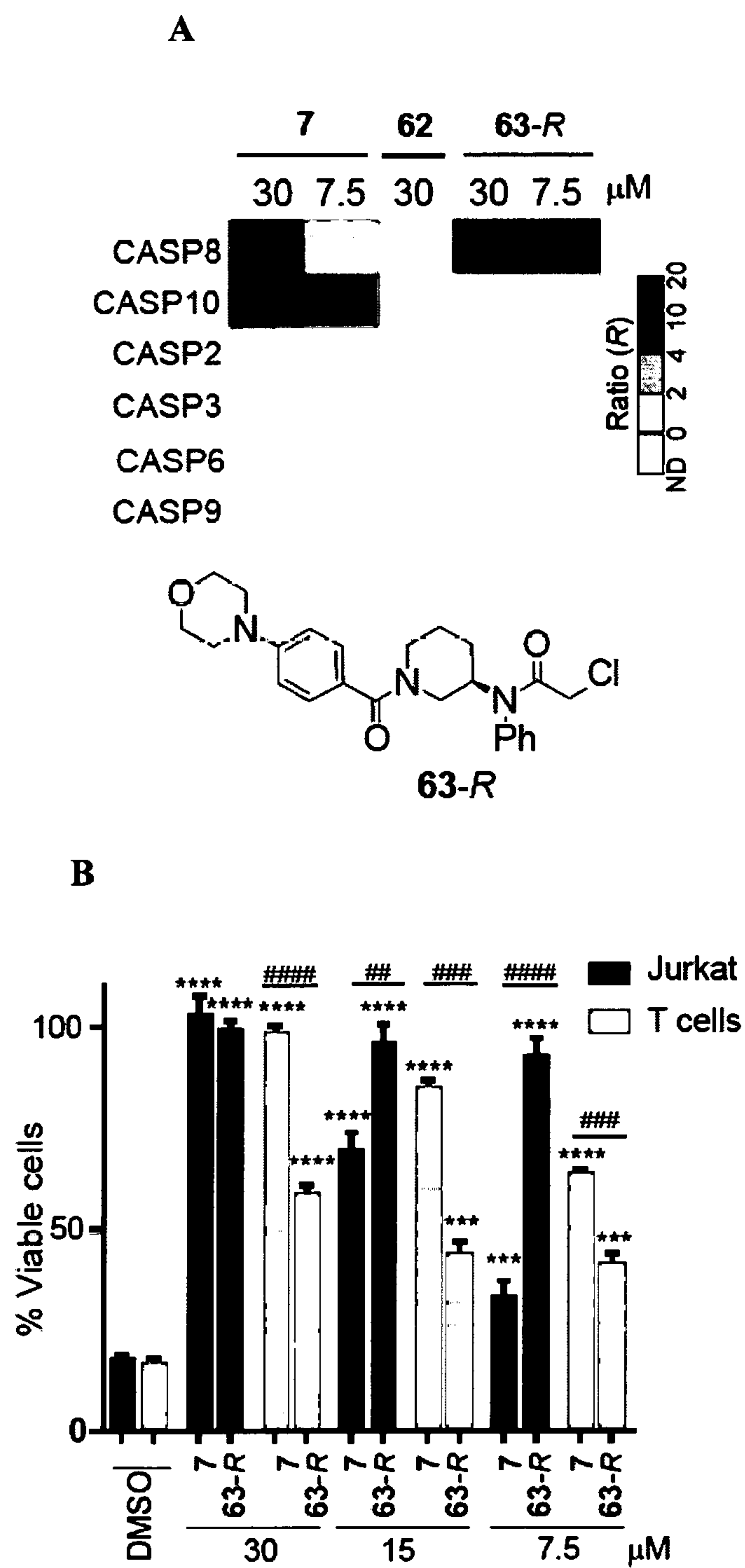
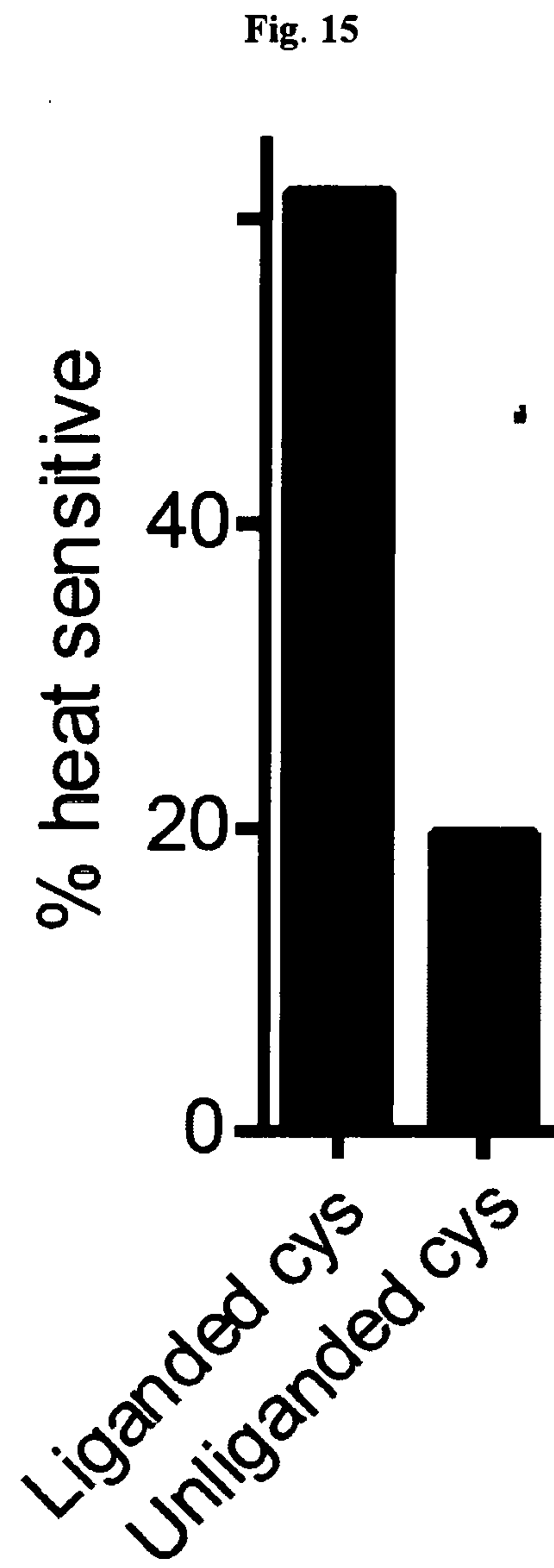


Fig. 14





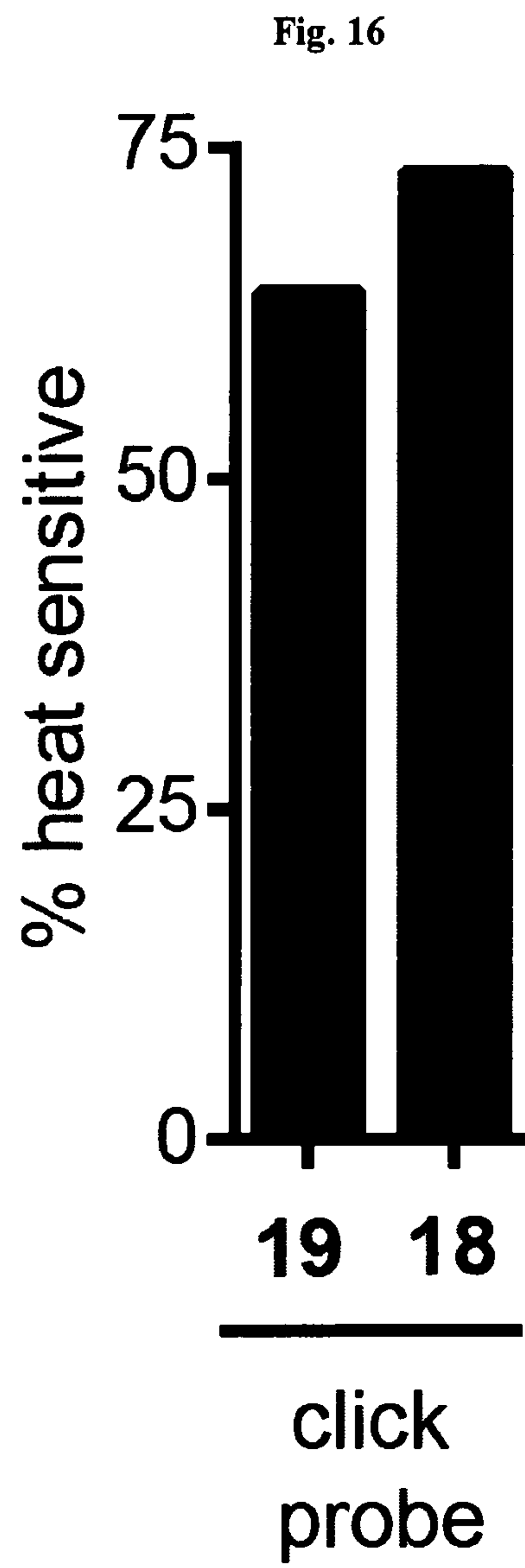


Fig. 17A

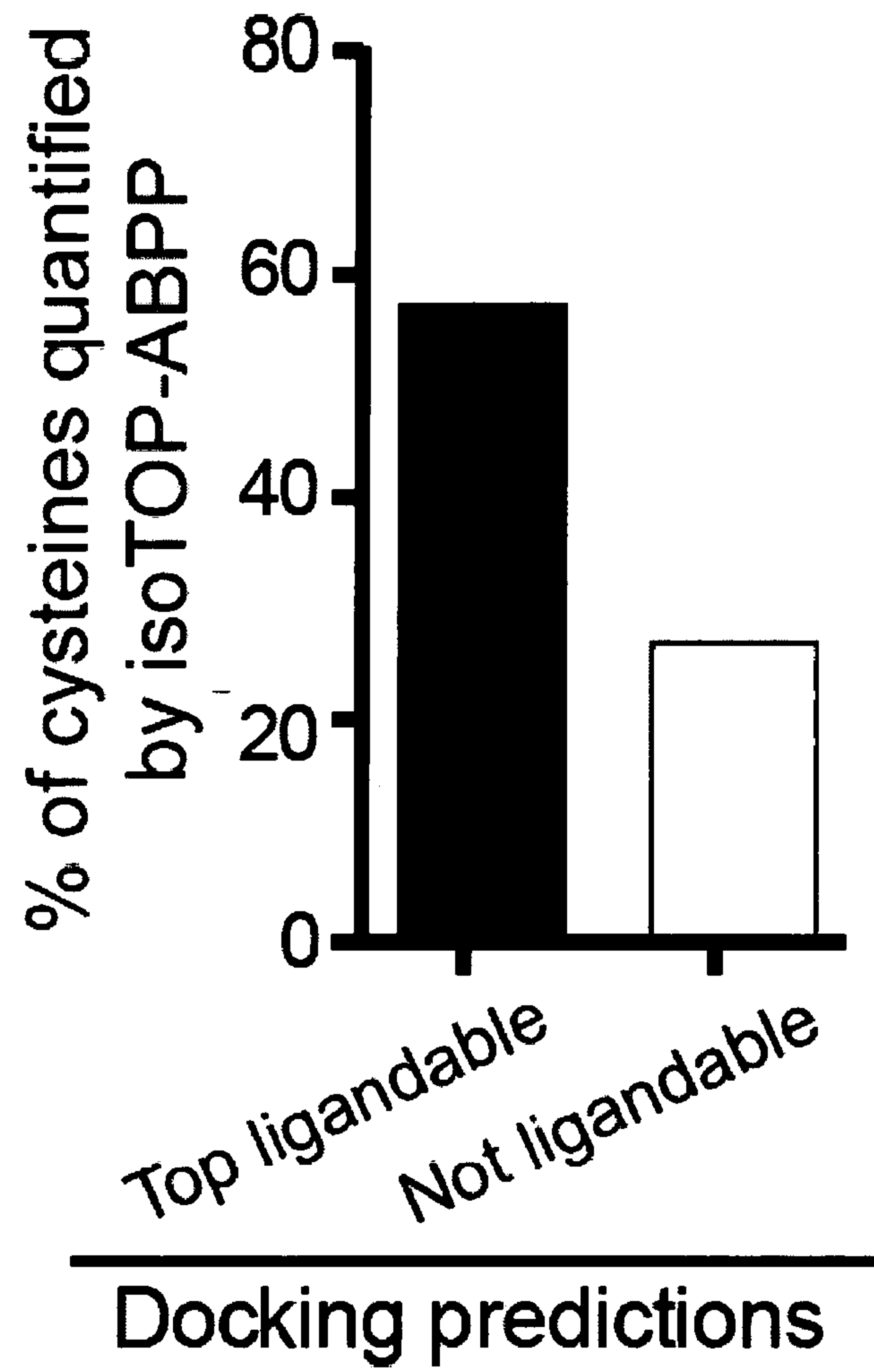


Fig. 17B

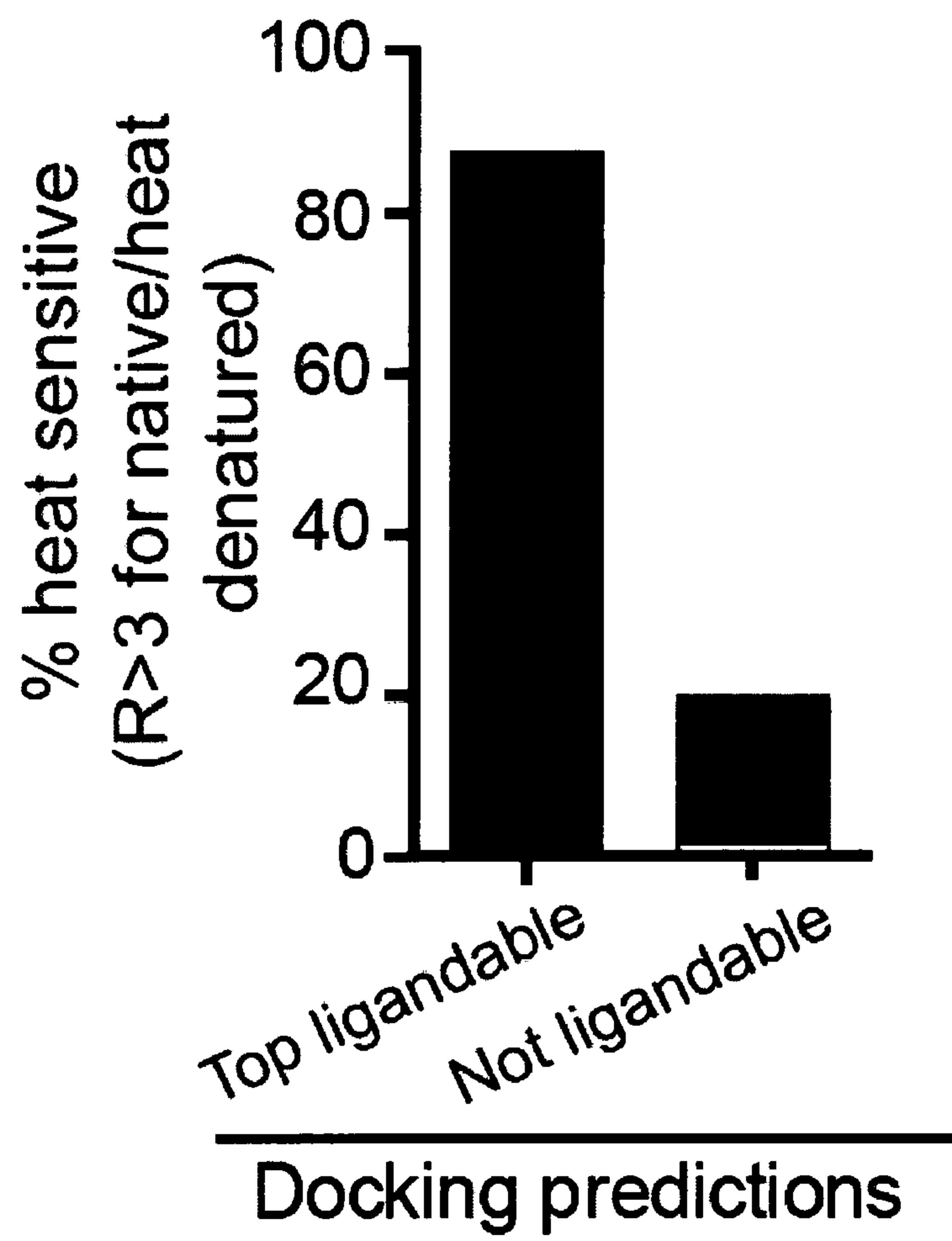


Fig. 18

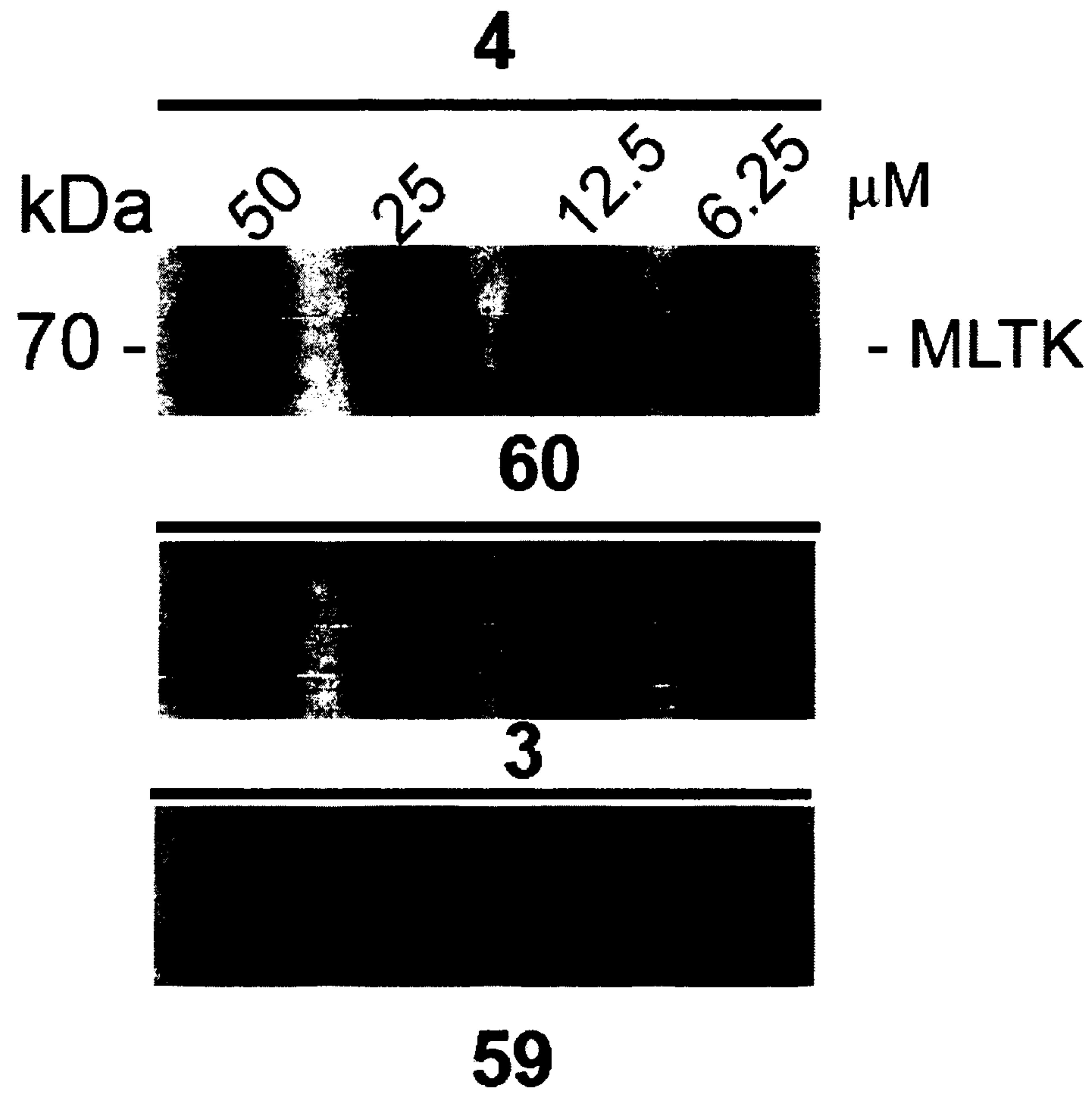


Fig. 19

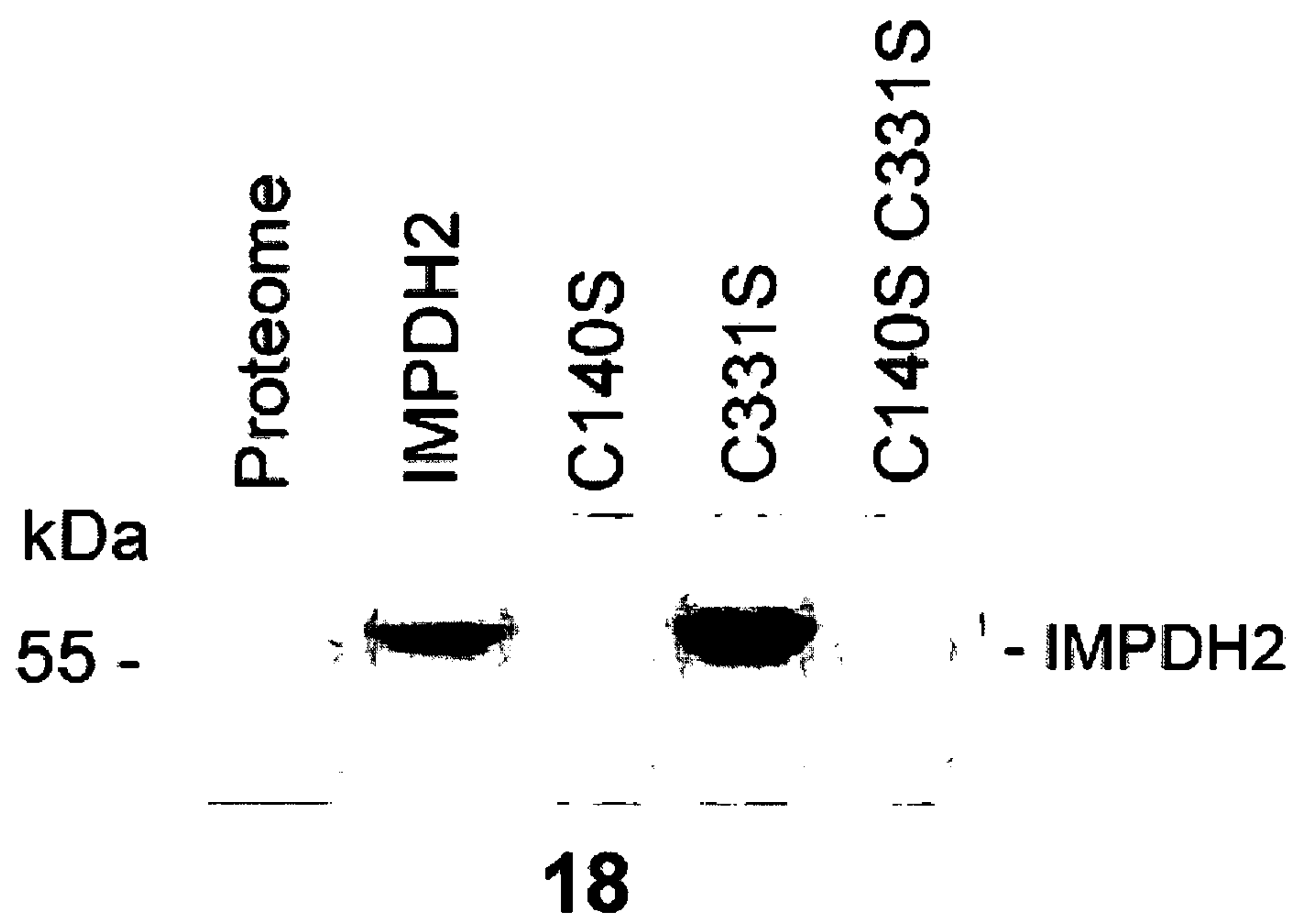


Fig. 20

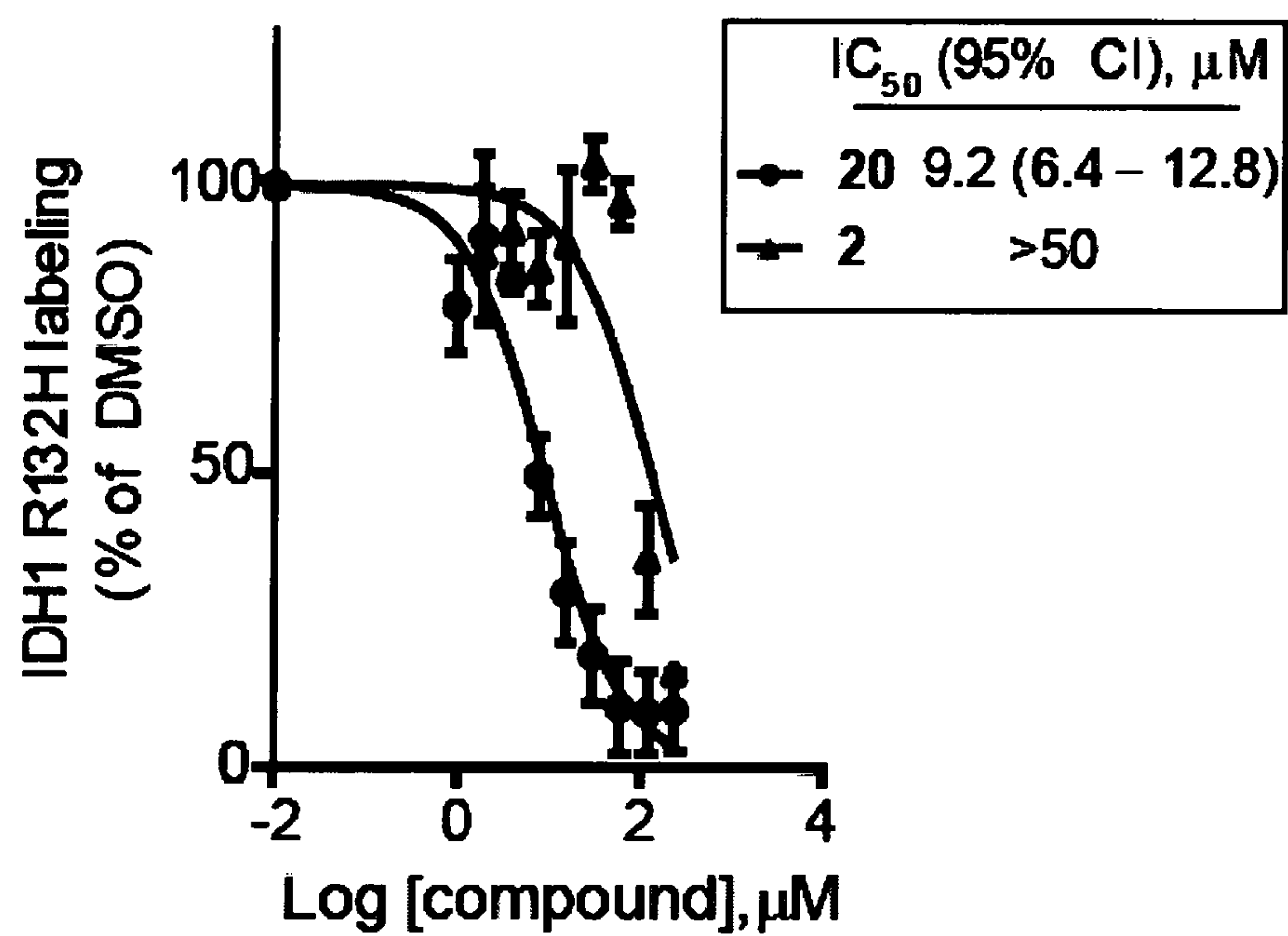


Fig. 21A

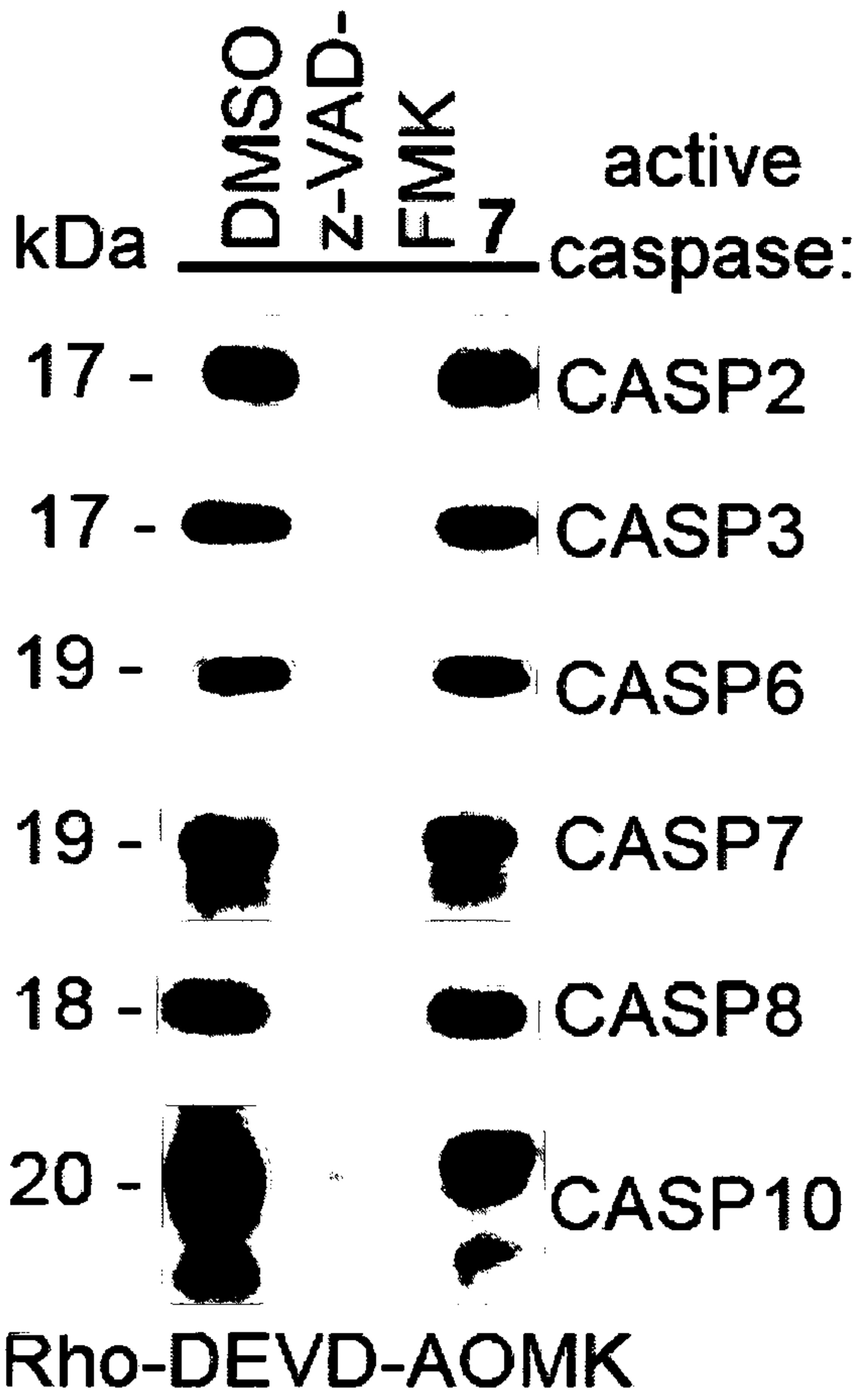


Fig. 21B

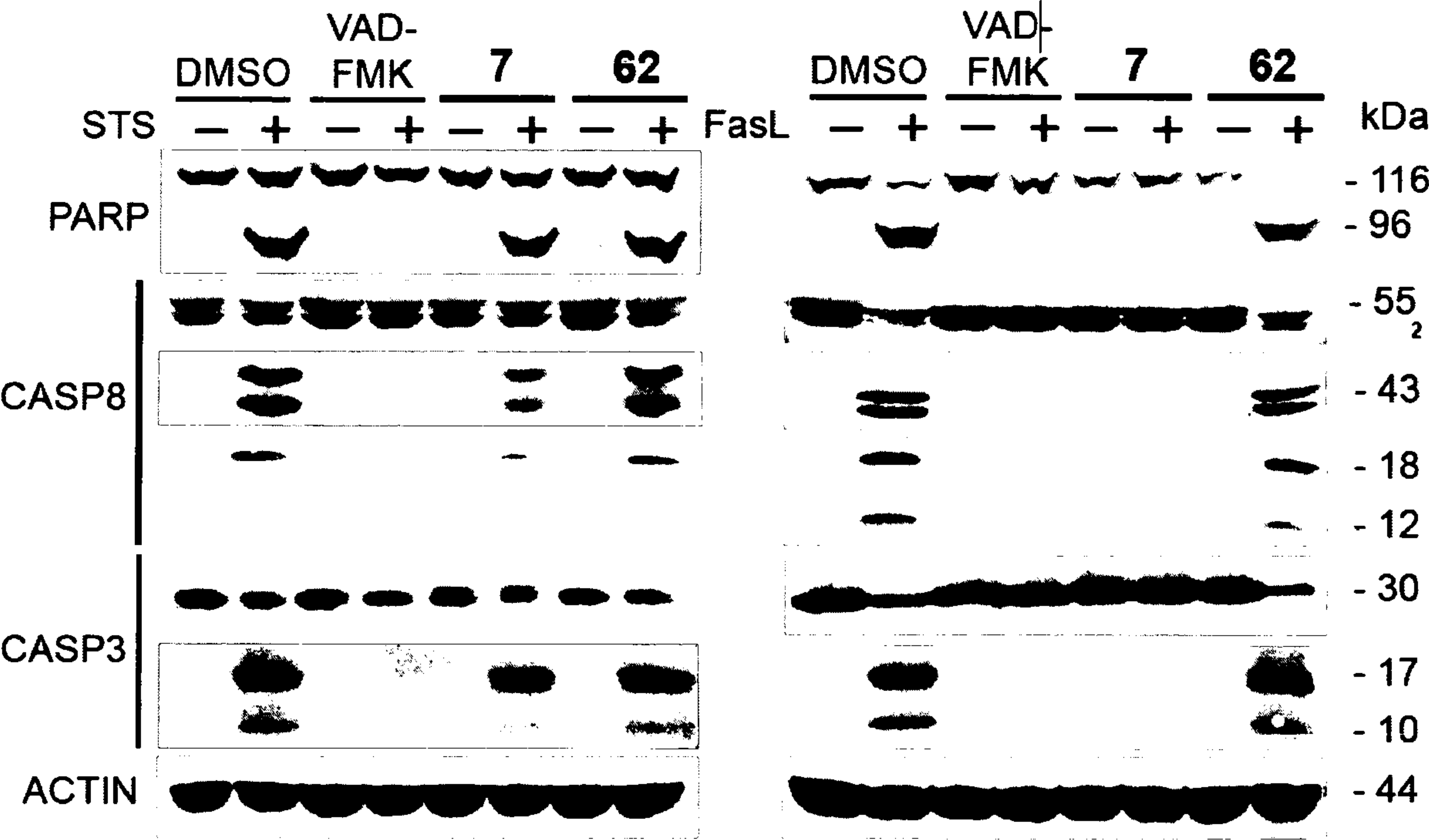


Fig. 21C

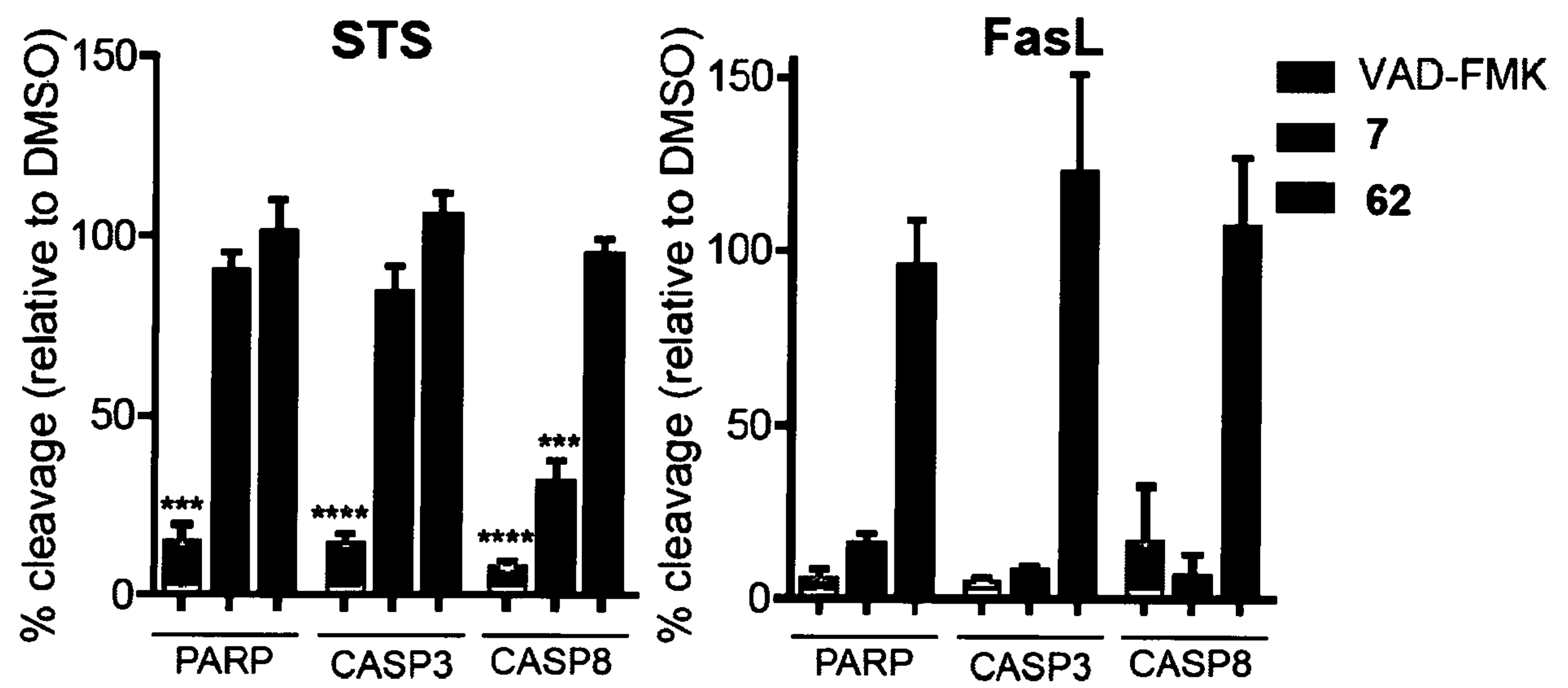


Fig. 22

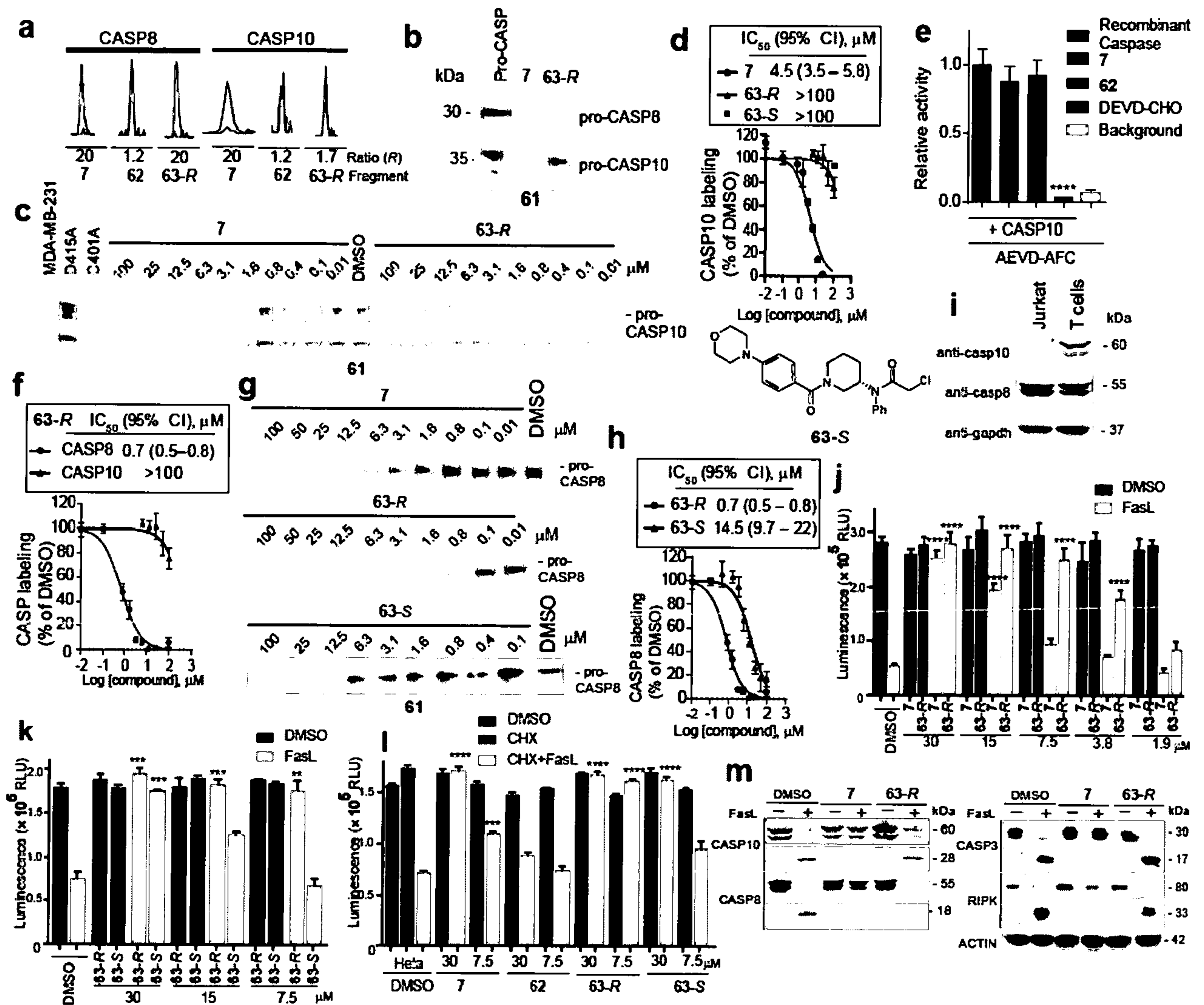


Fig. 23A

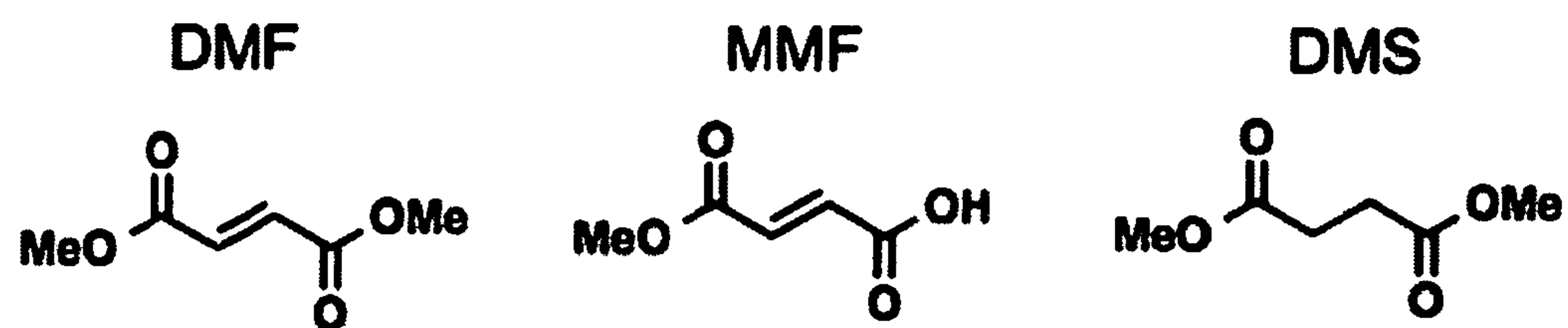


Fig. 23B

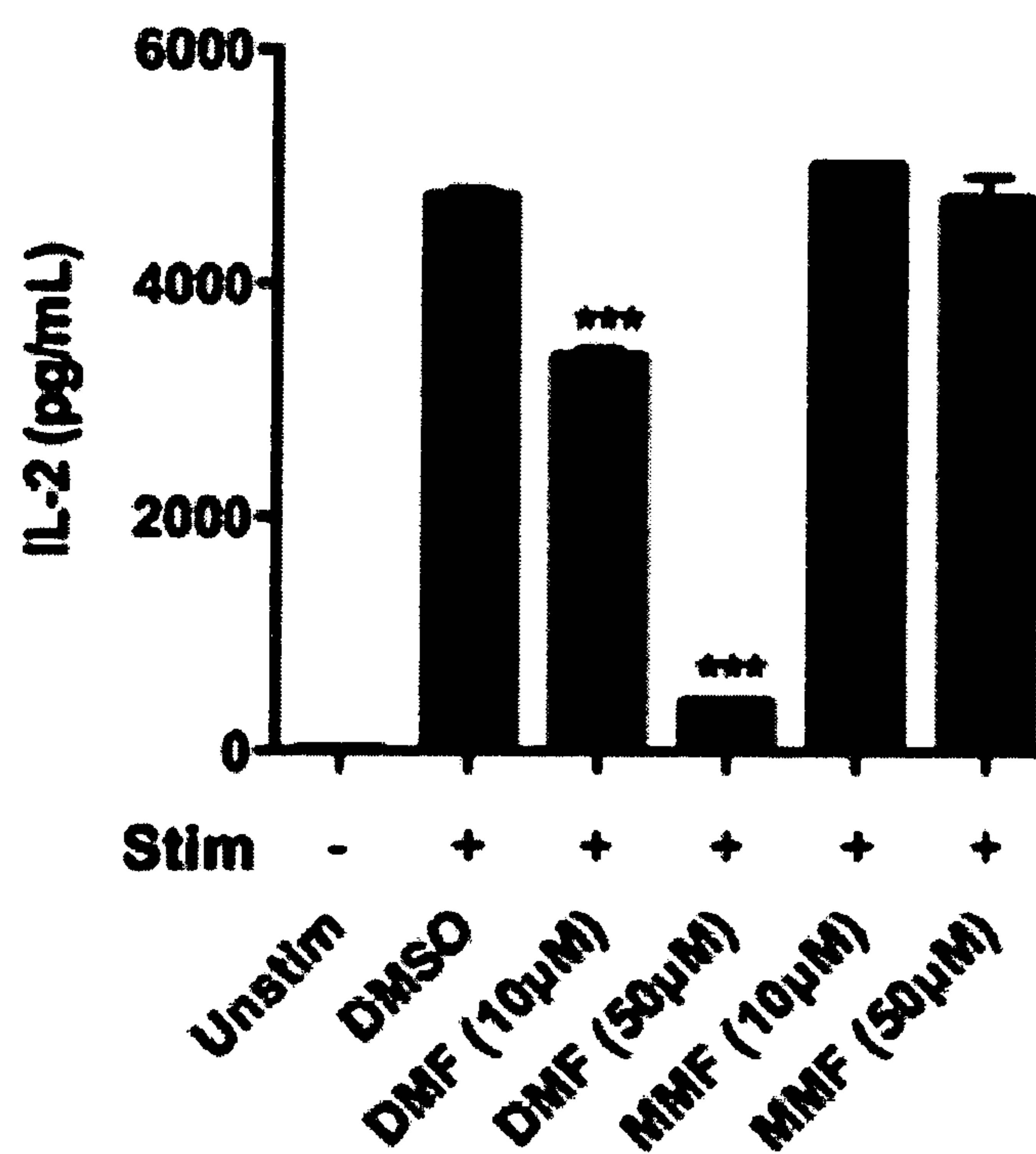


Fig. 23C

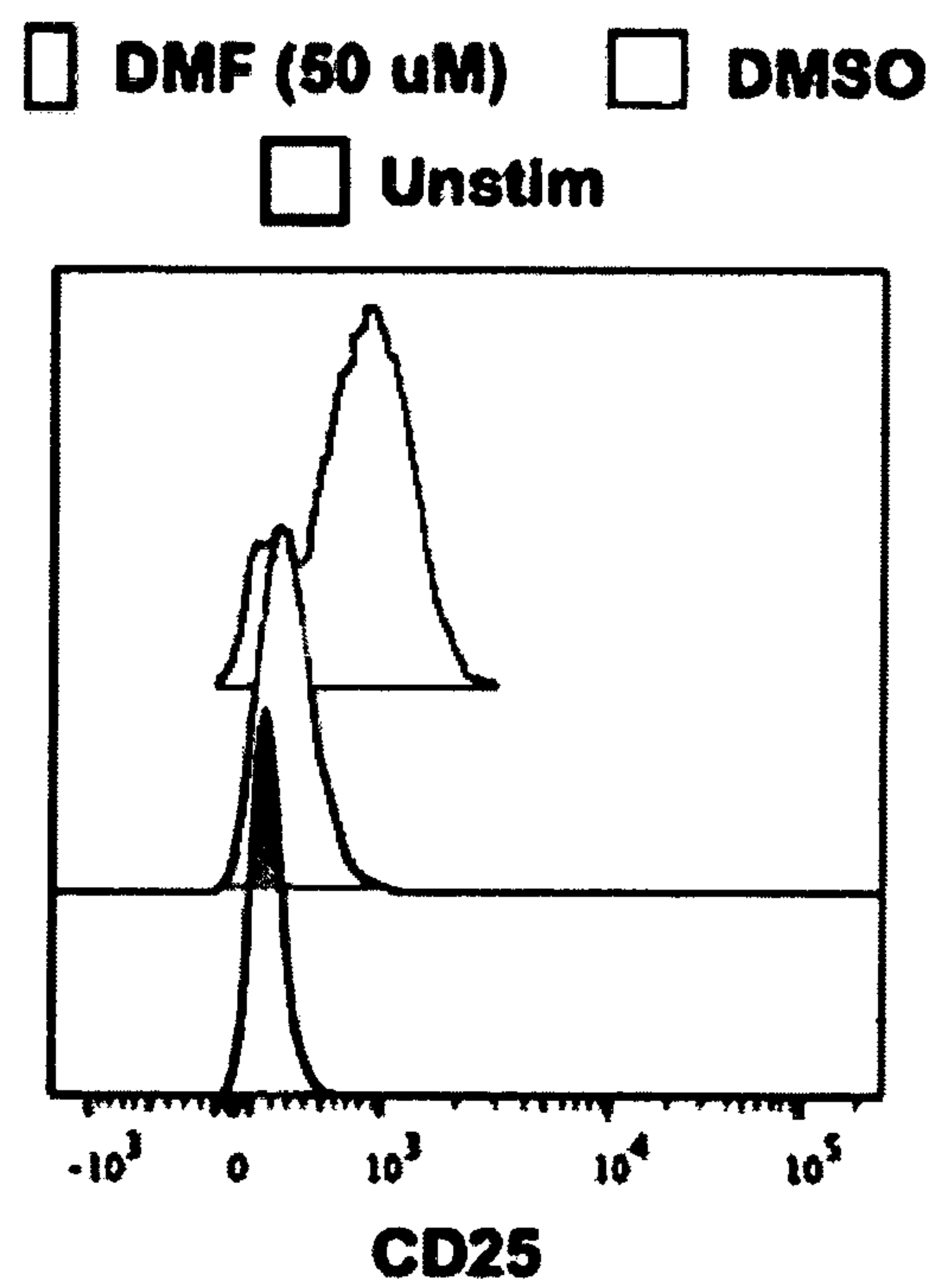


Fig. 23D

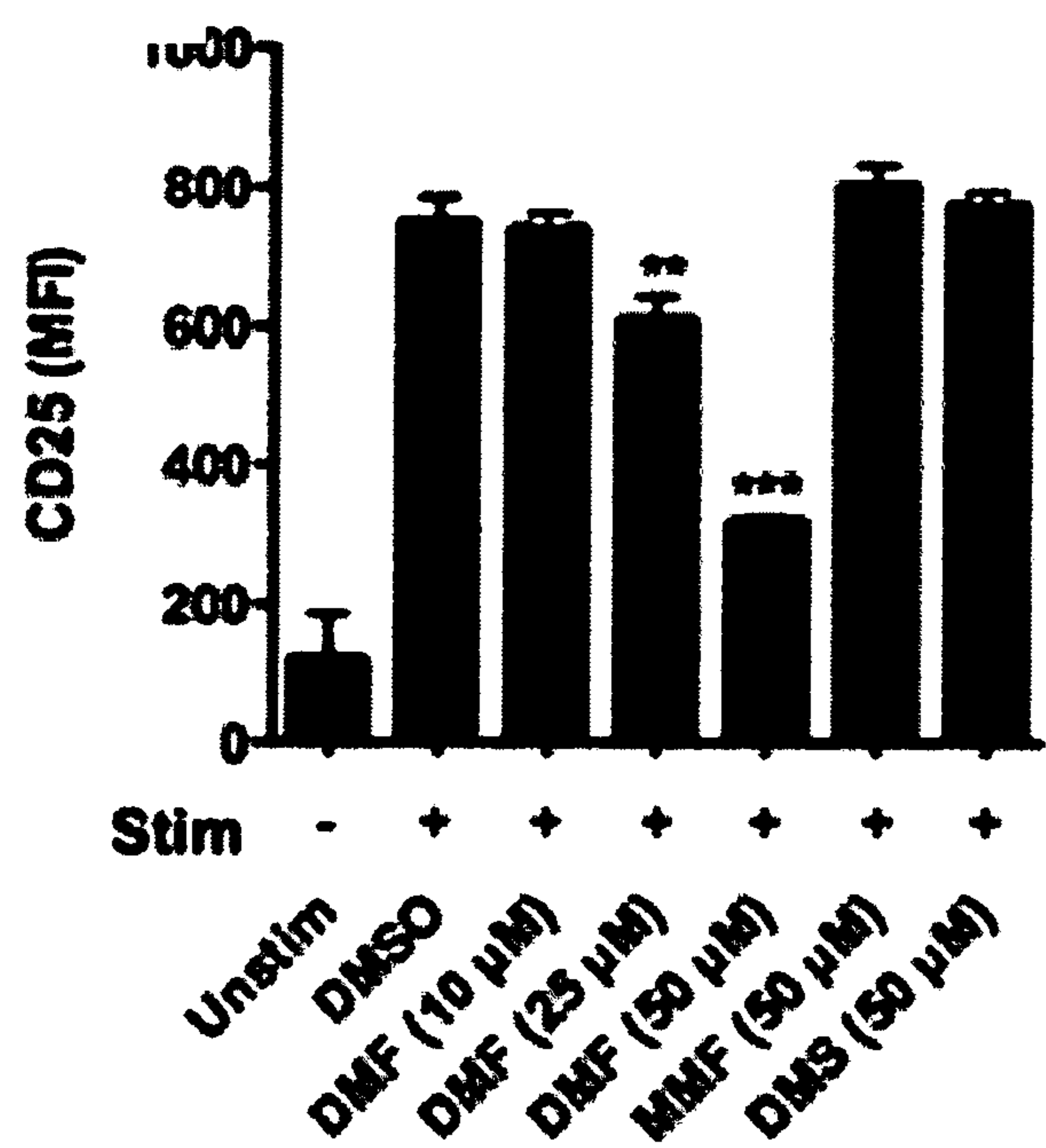


Fig. 23E

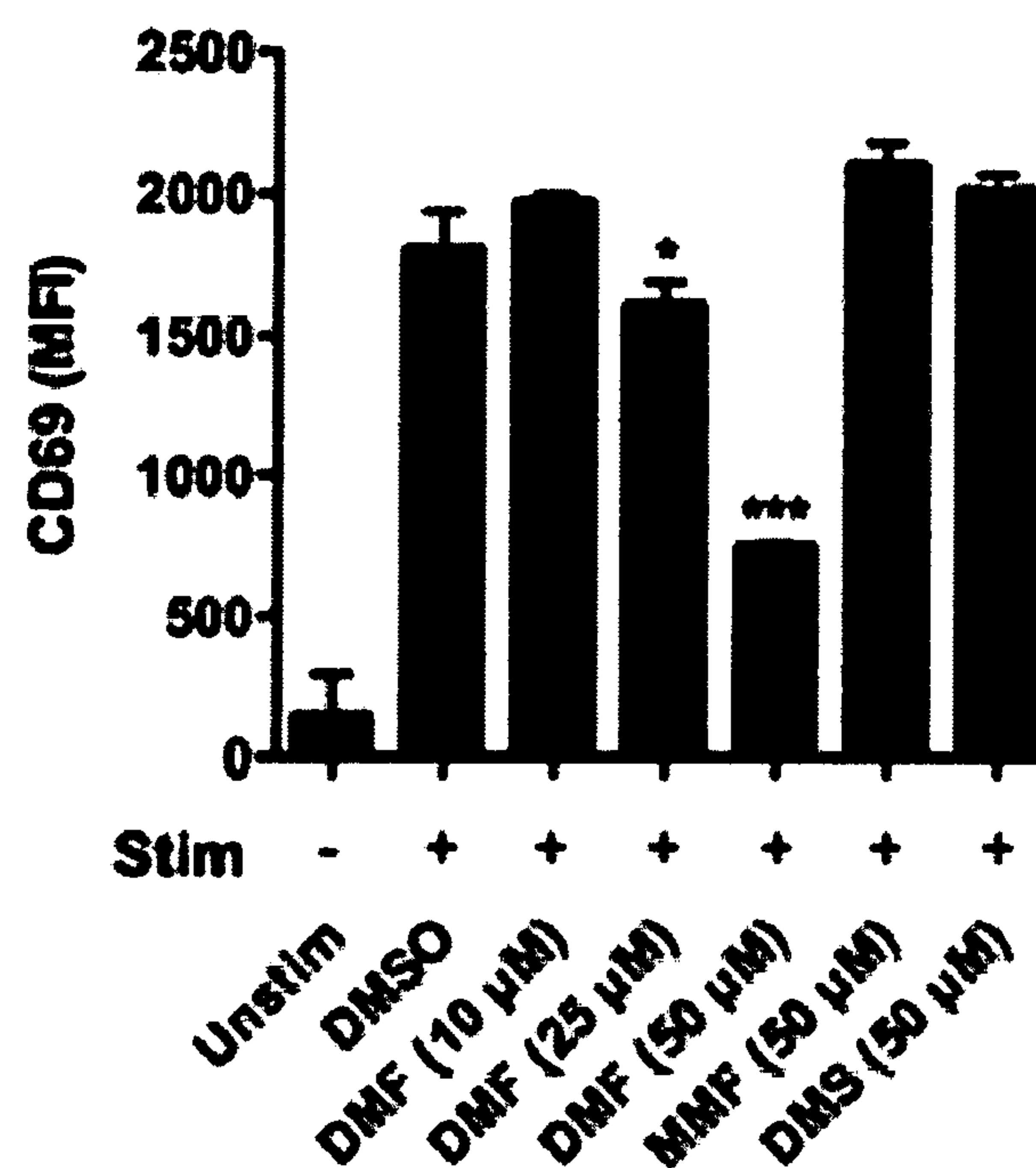


Fig. 23F

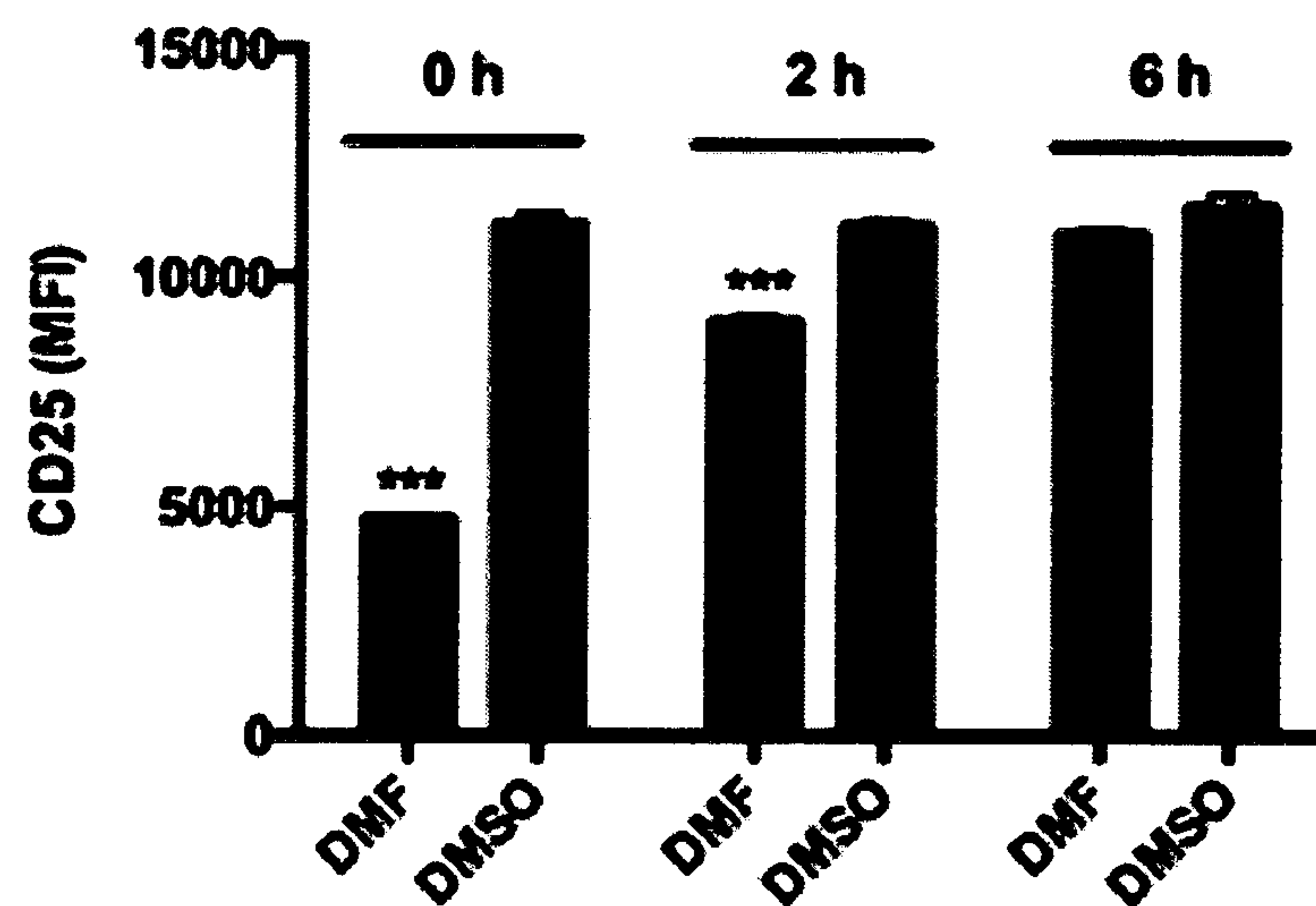


Fig. 24

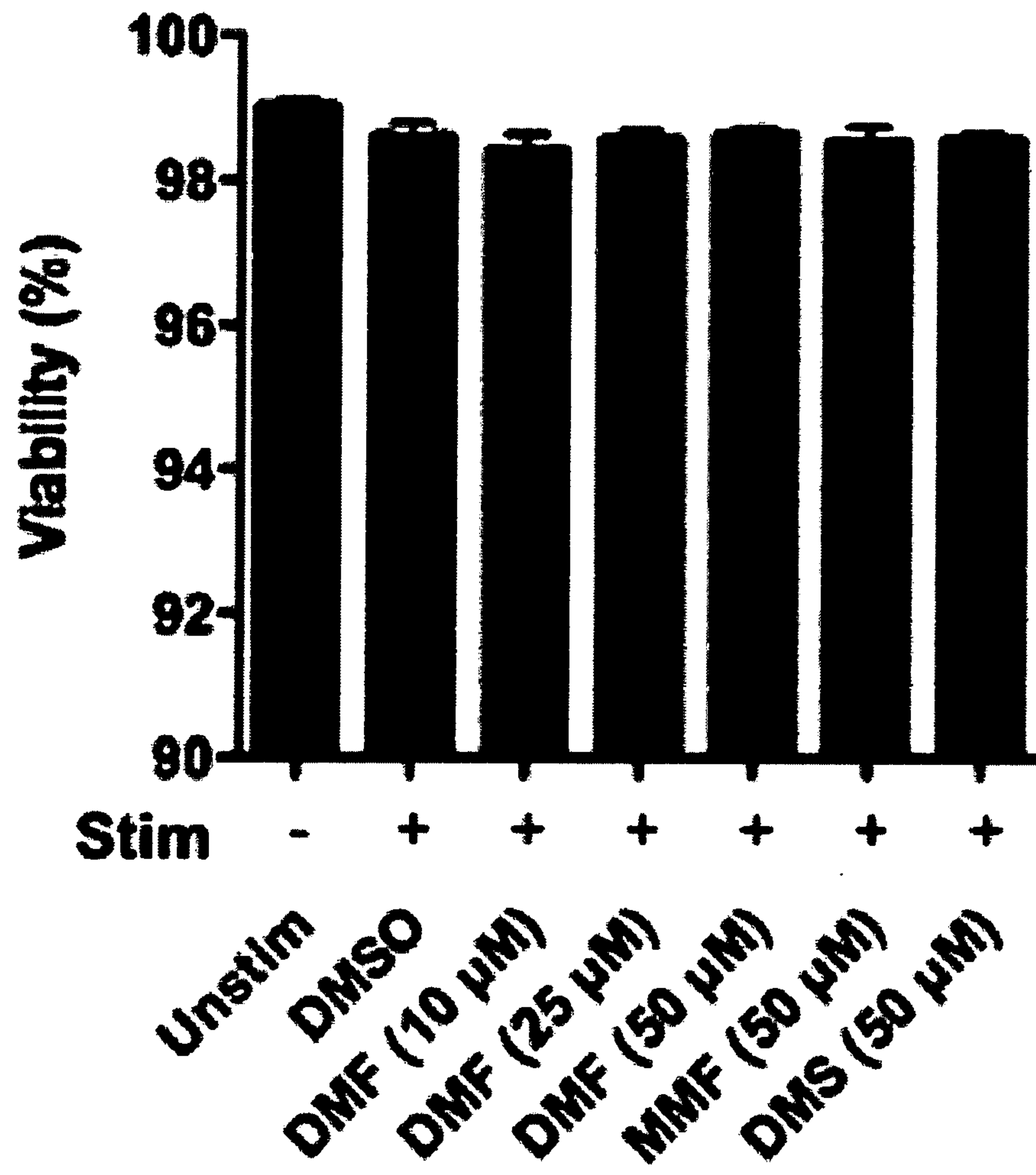


Fig. 25A

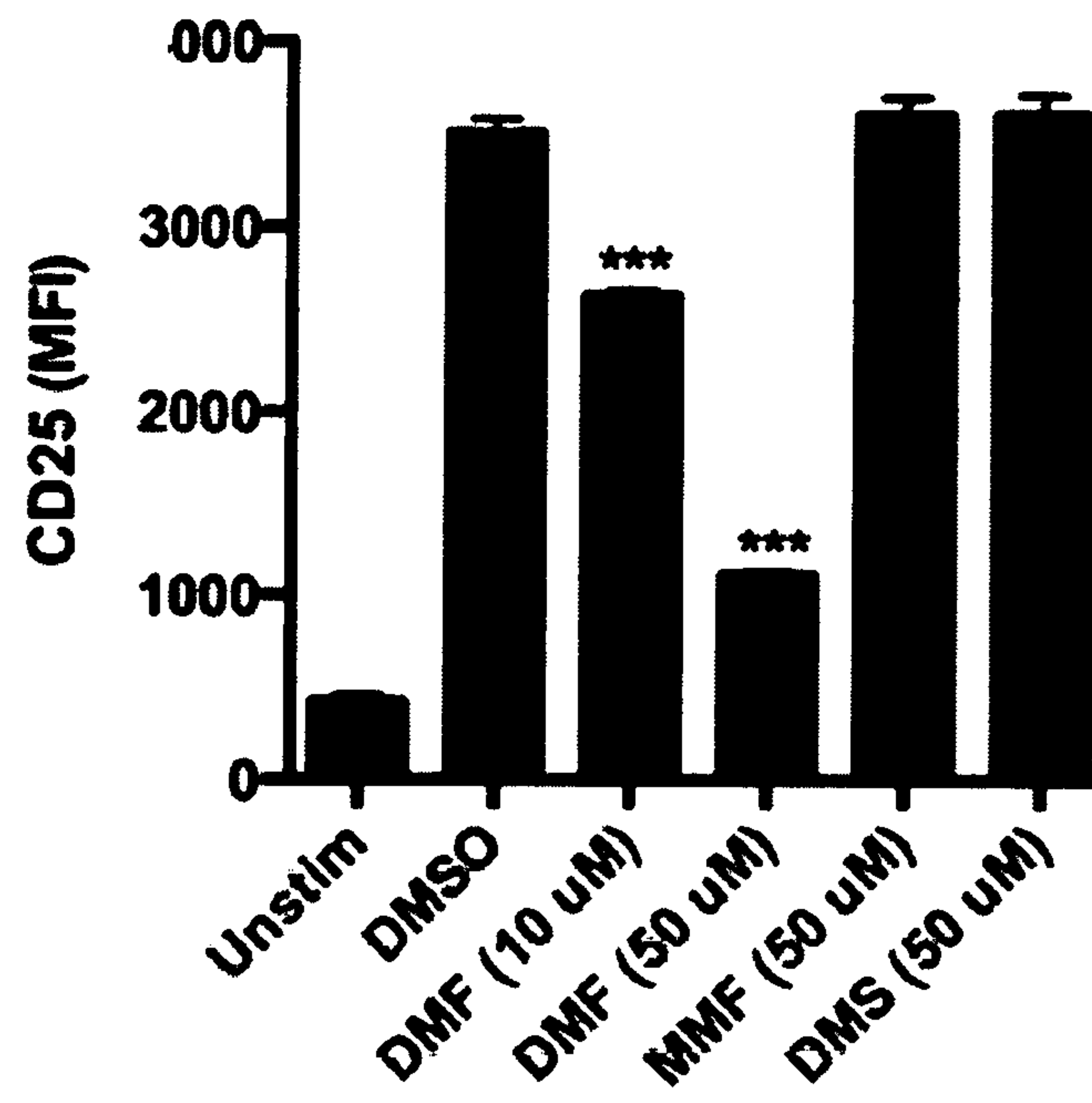


Fig. 25B

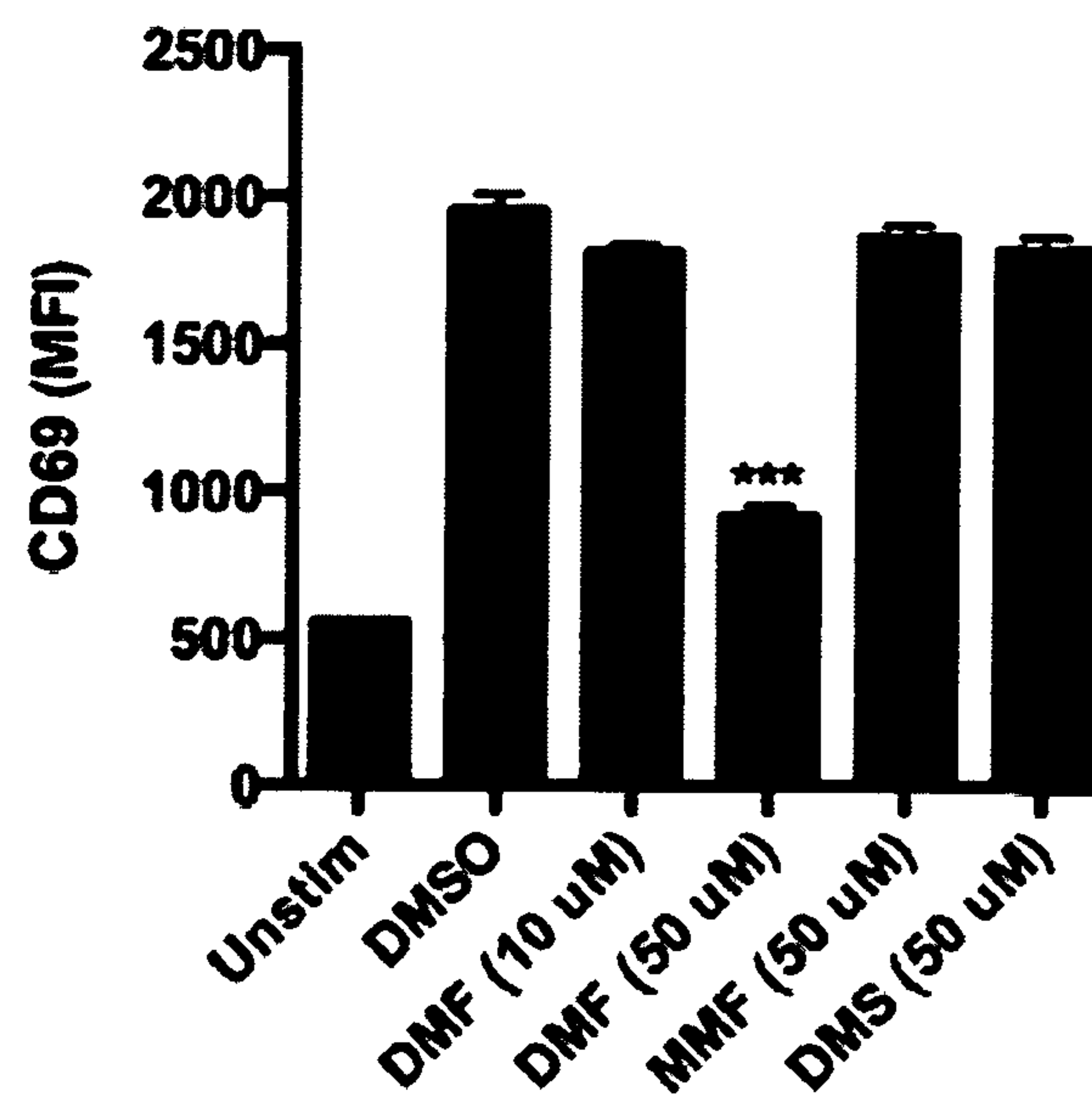


Fig. 26A

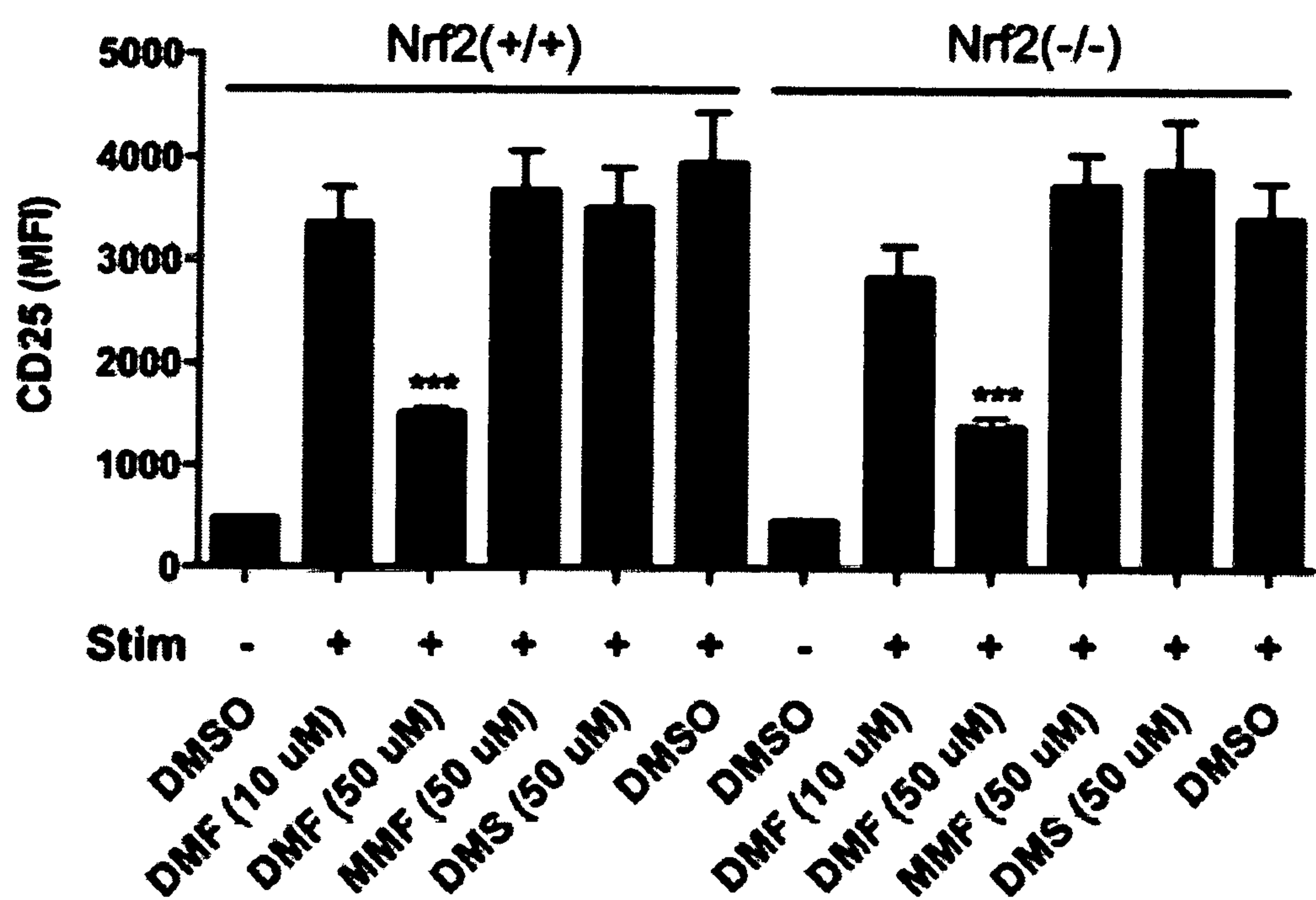


Fig. 26B

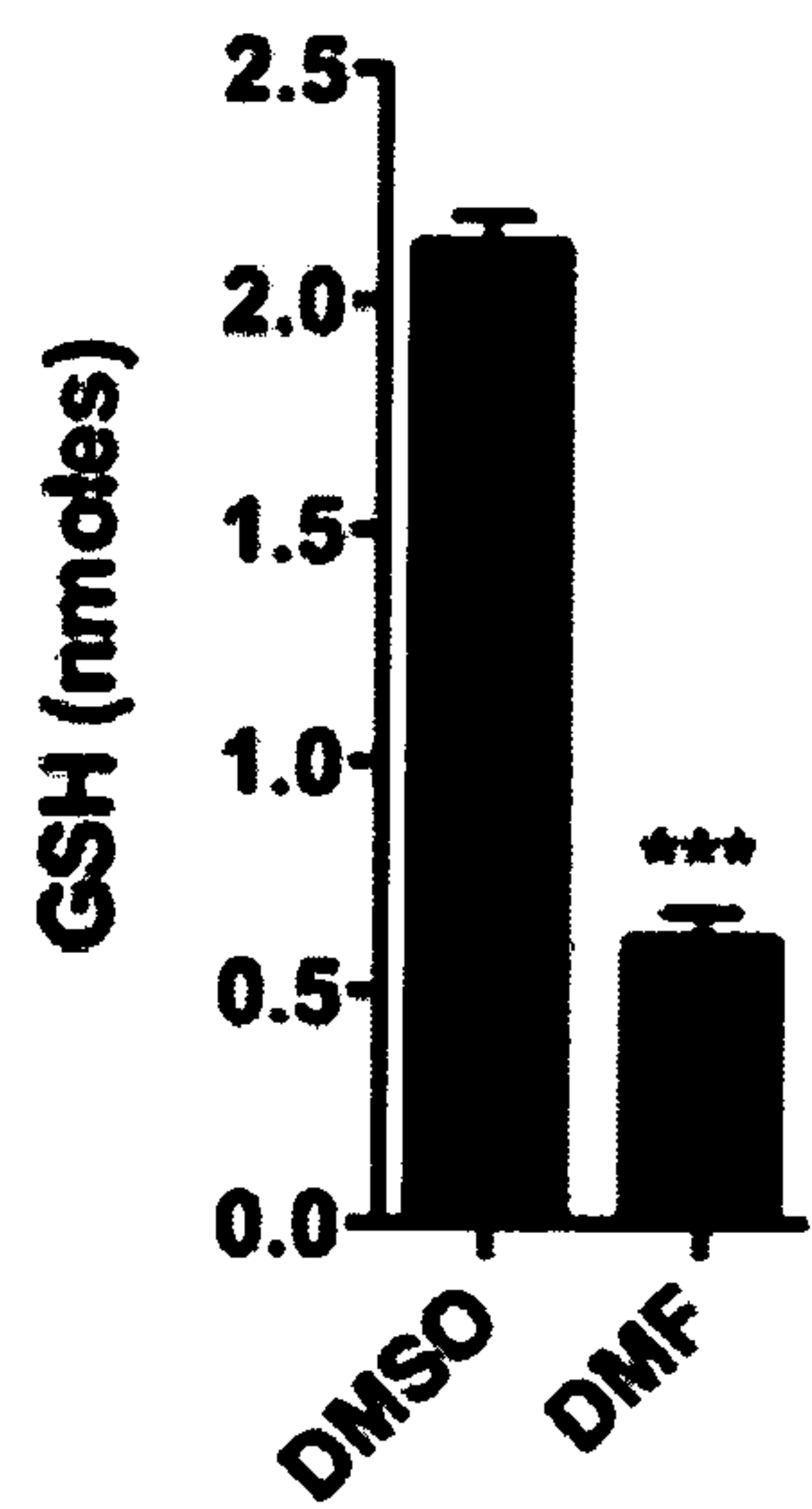


Fig. 26C

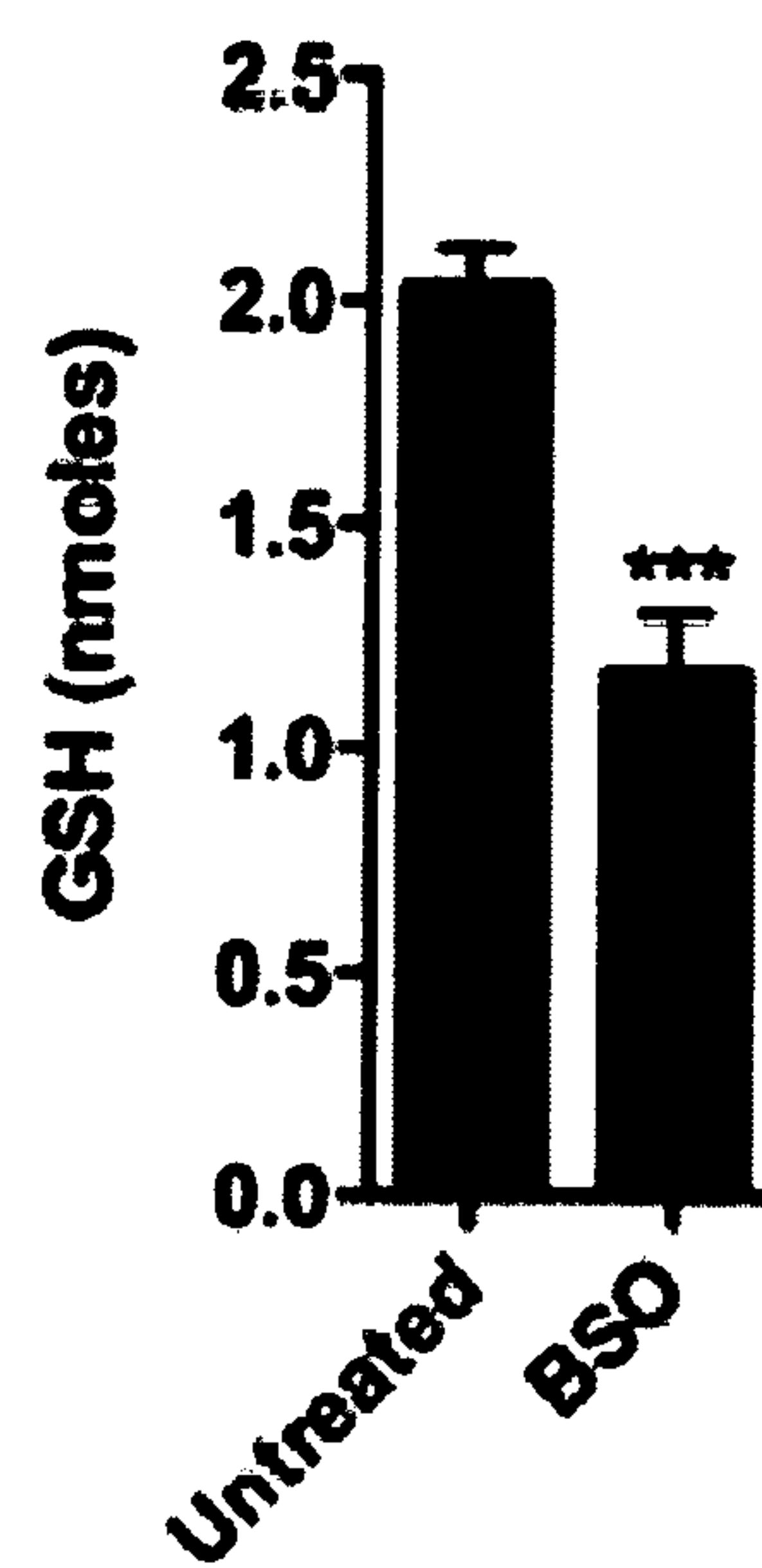


Fig. 26D

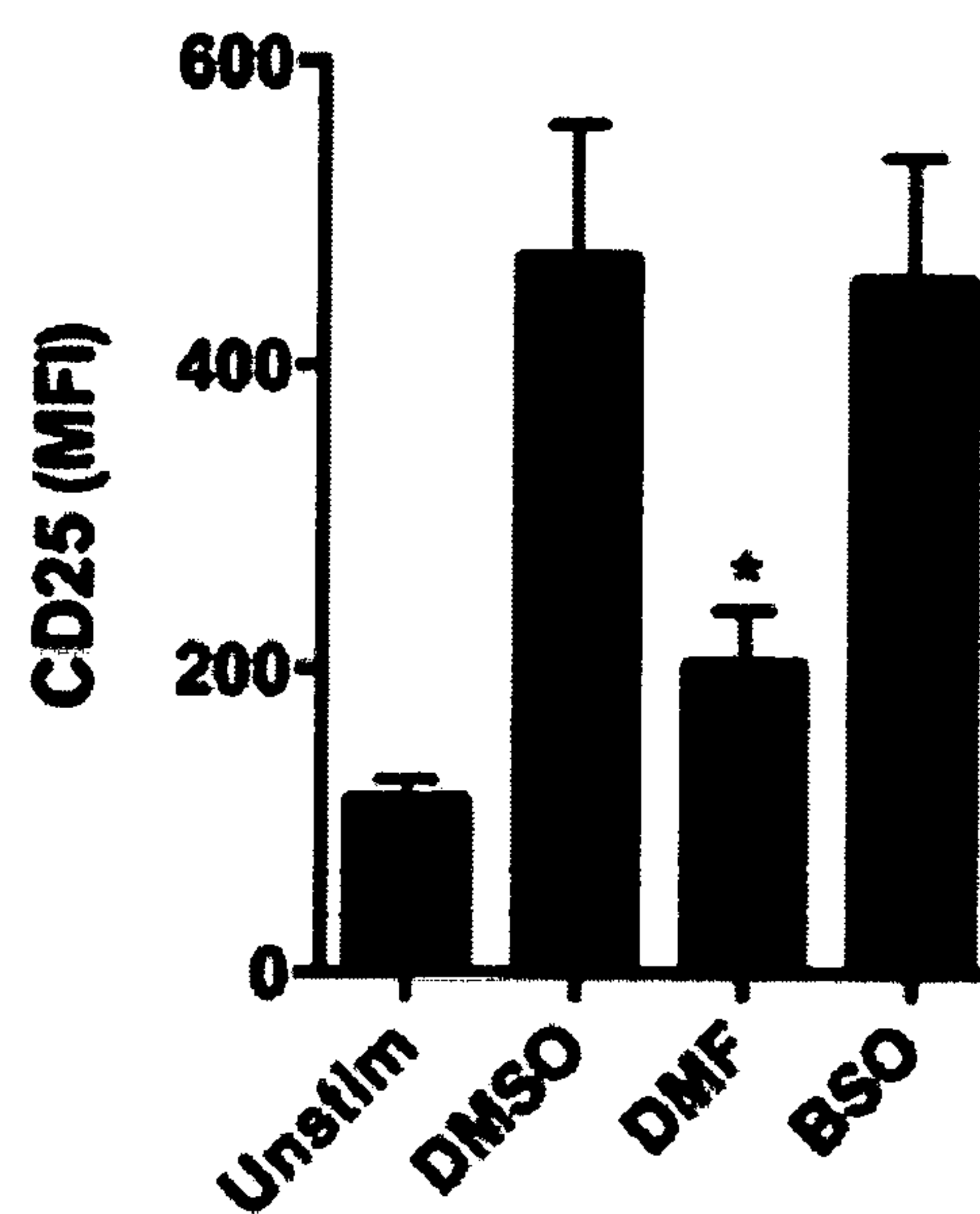


Fig. 27A

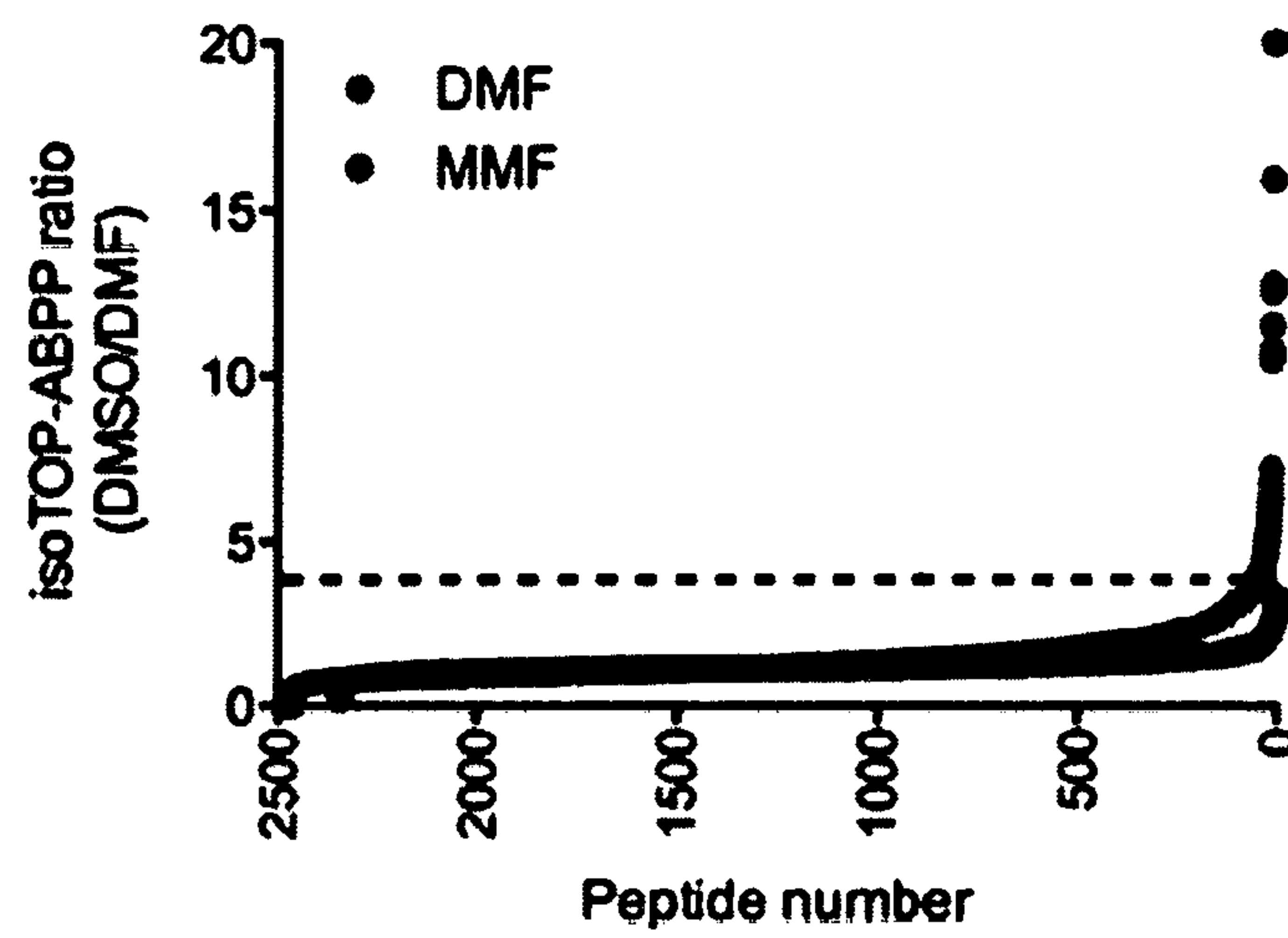


Fig. 27B

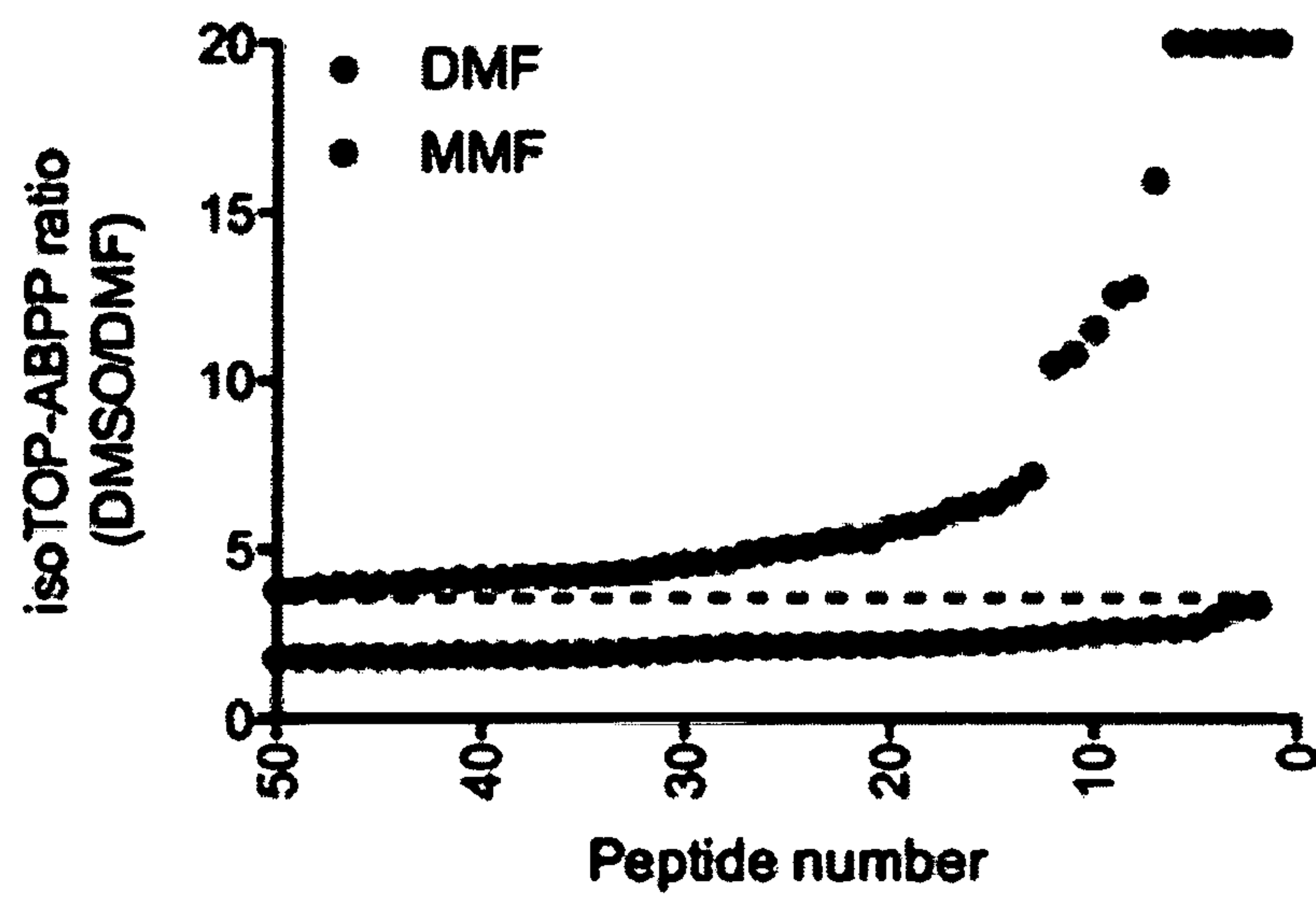


Fig. 27C

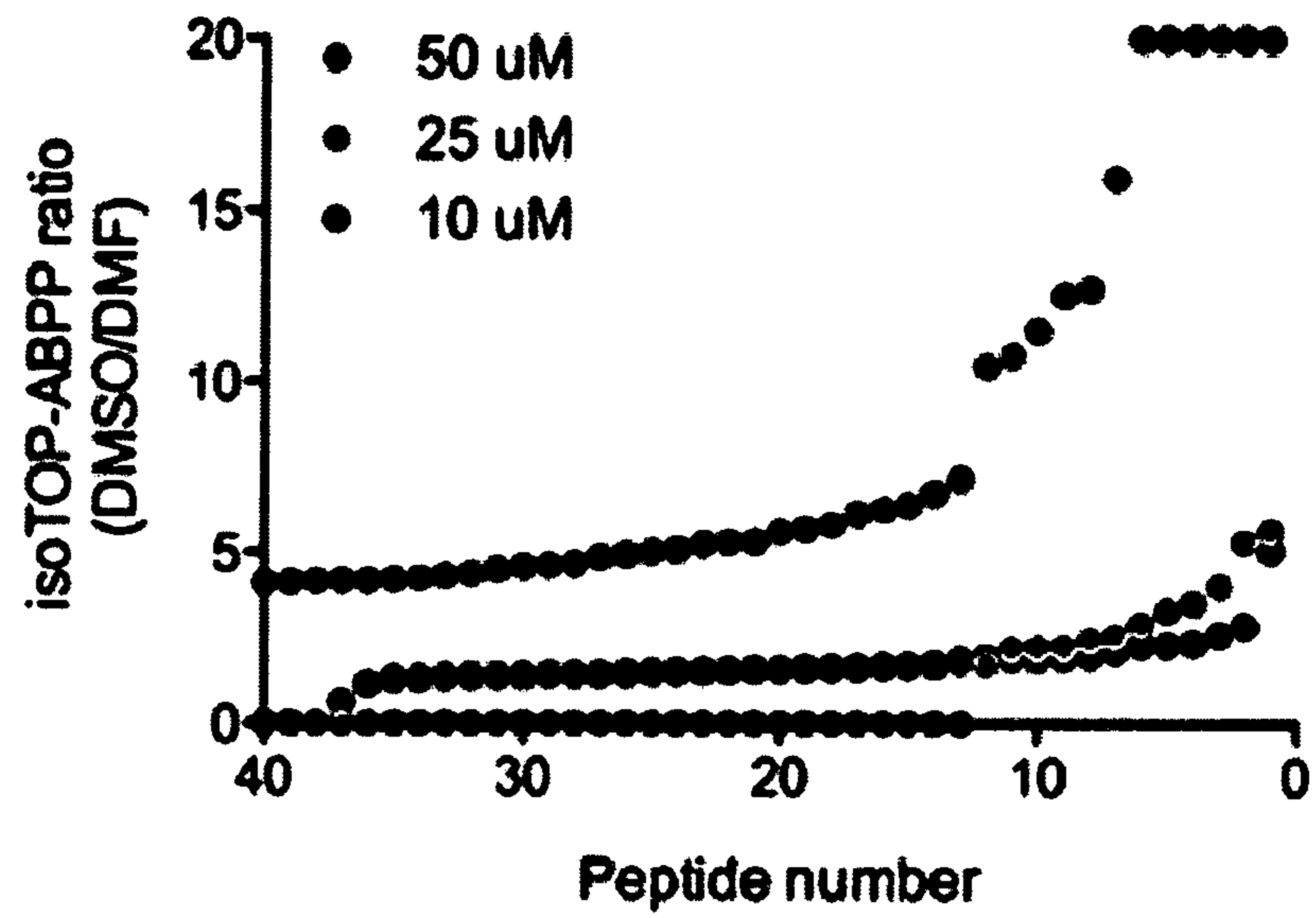


Fig. 27D

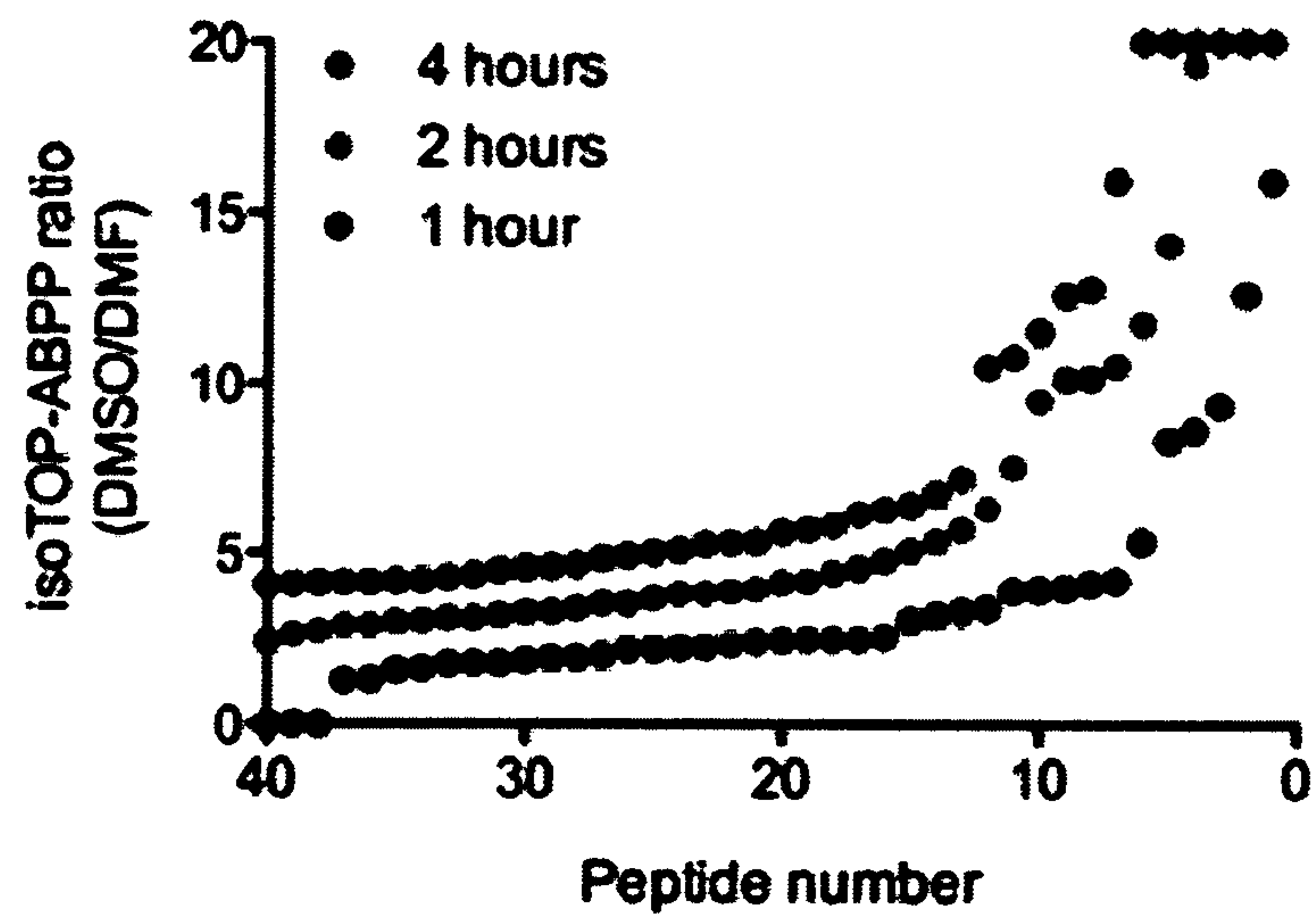


Fig. 27E

E

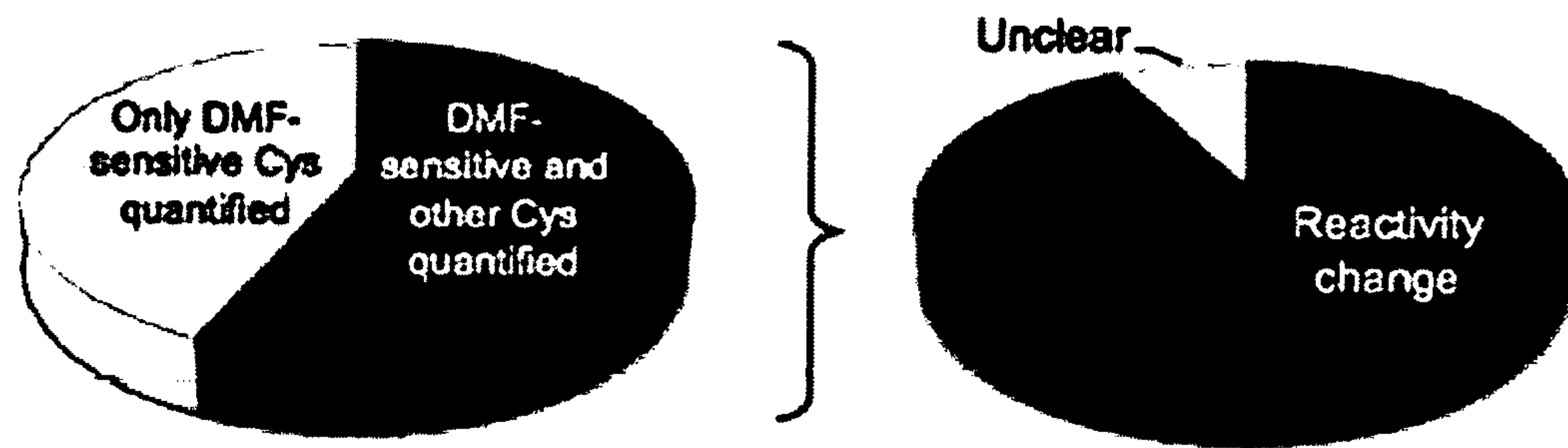


Fig. 27F

F

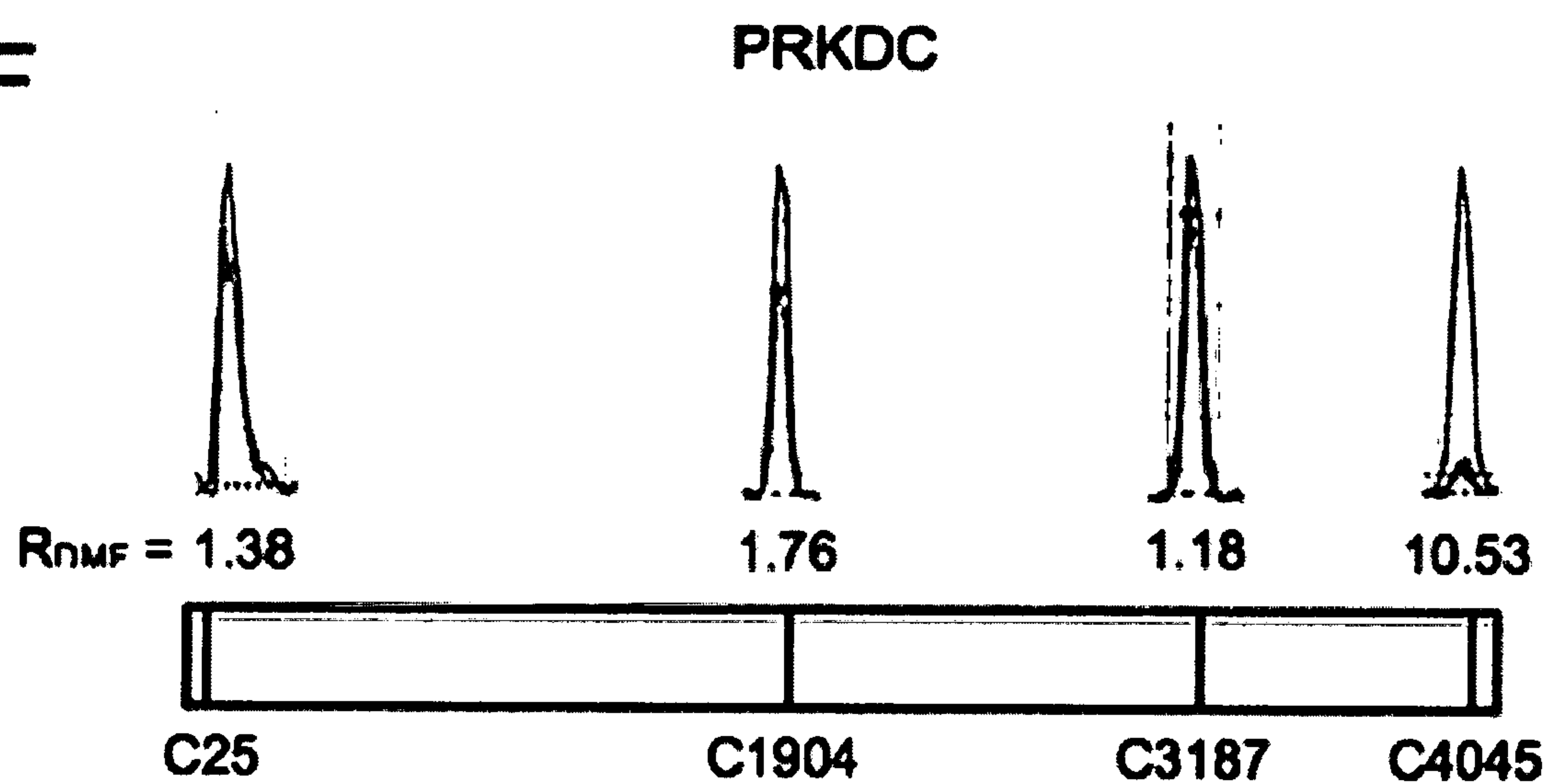


Fig. 28A

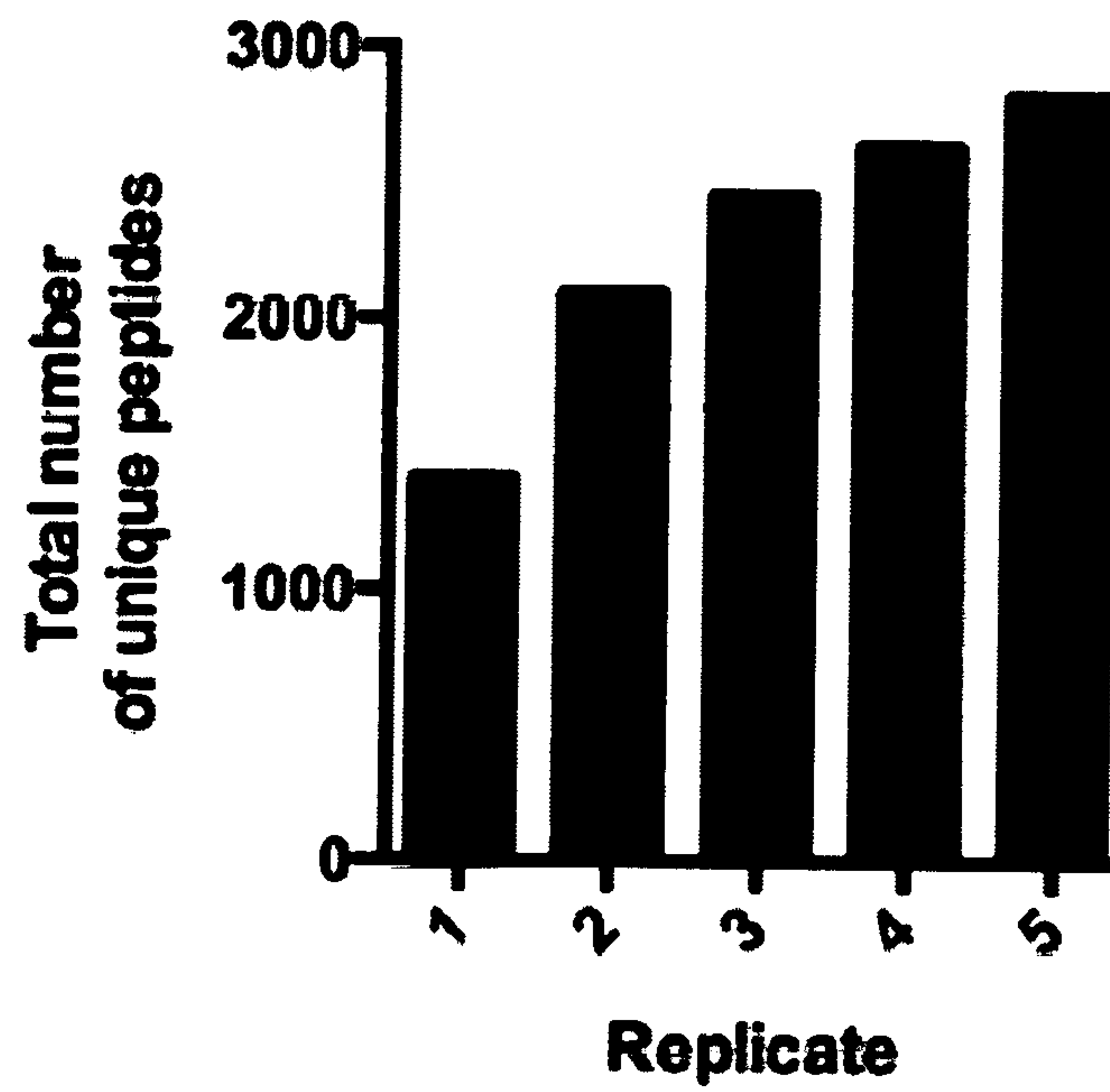


Fig. 28B

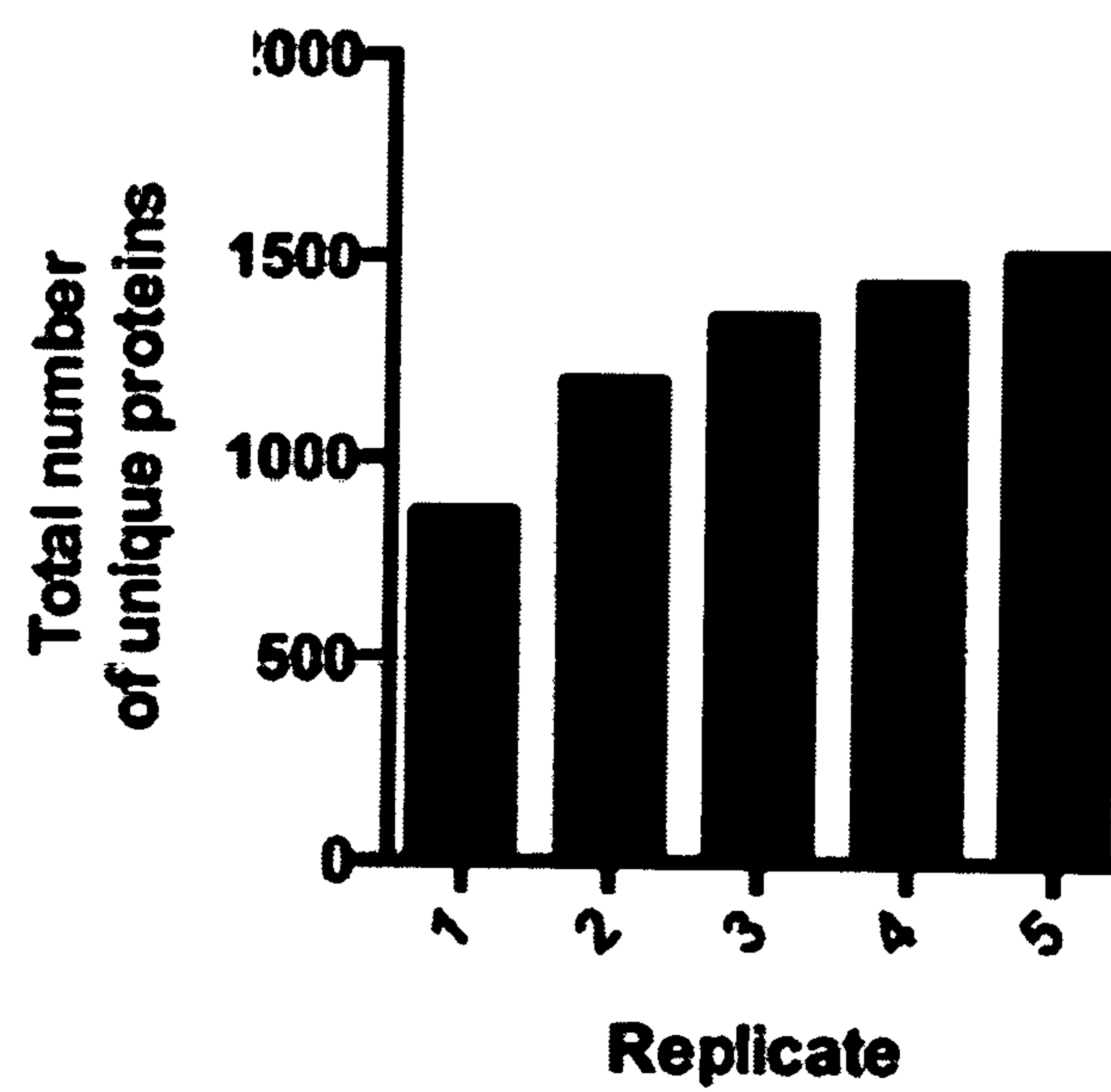


Fig. 29

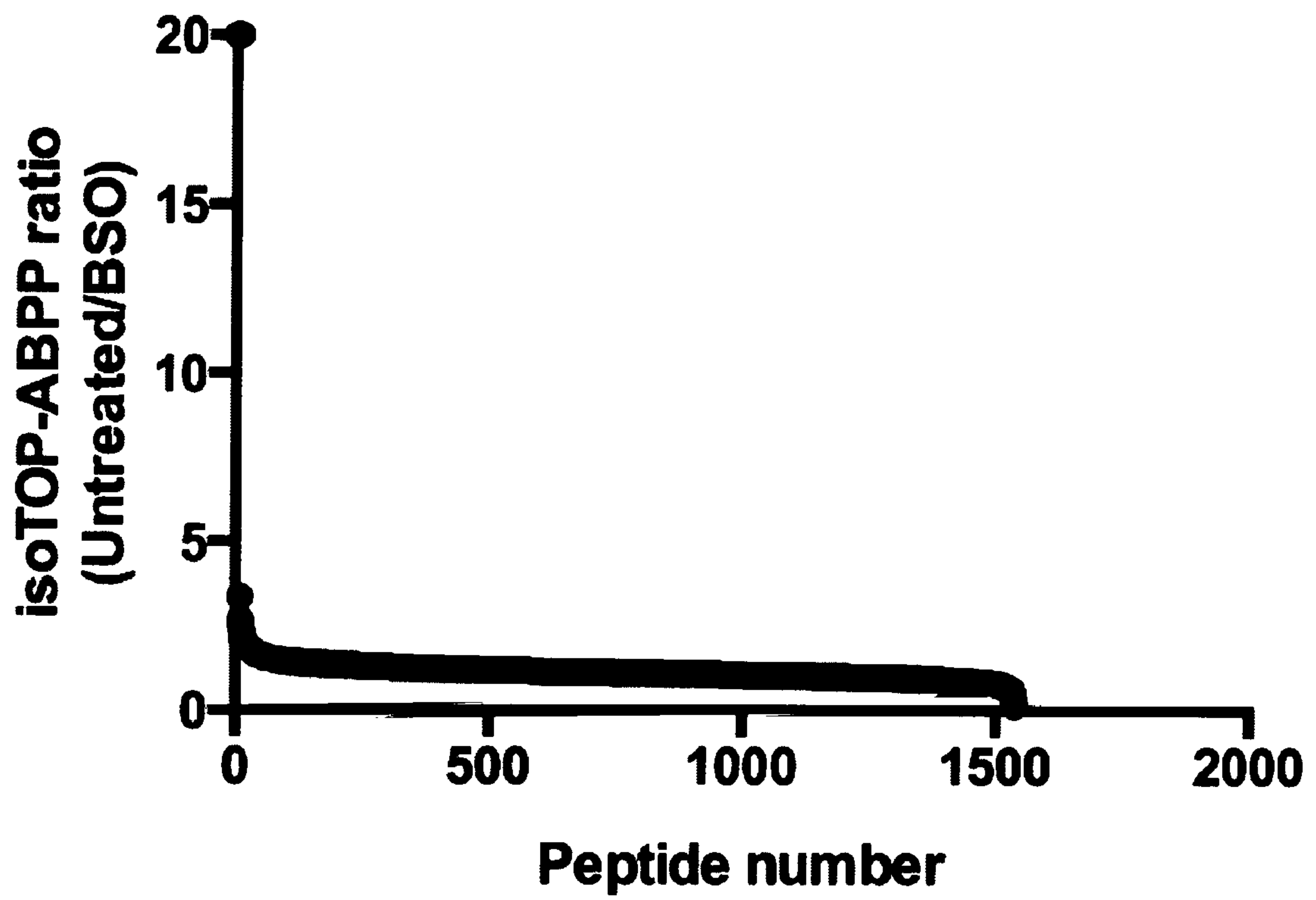


Fig. 30A

DMF-sensitive human cysteines
conserved in mice

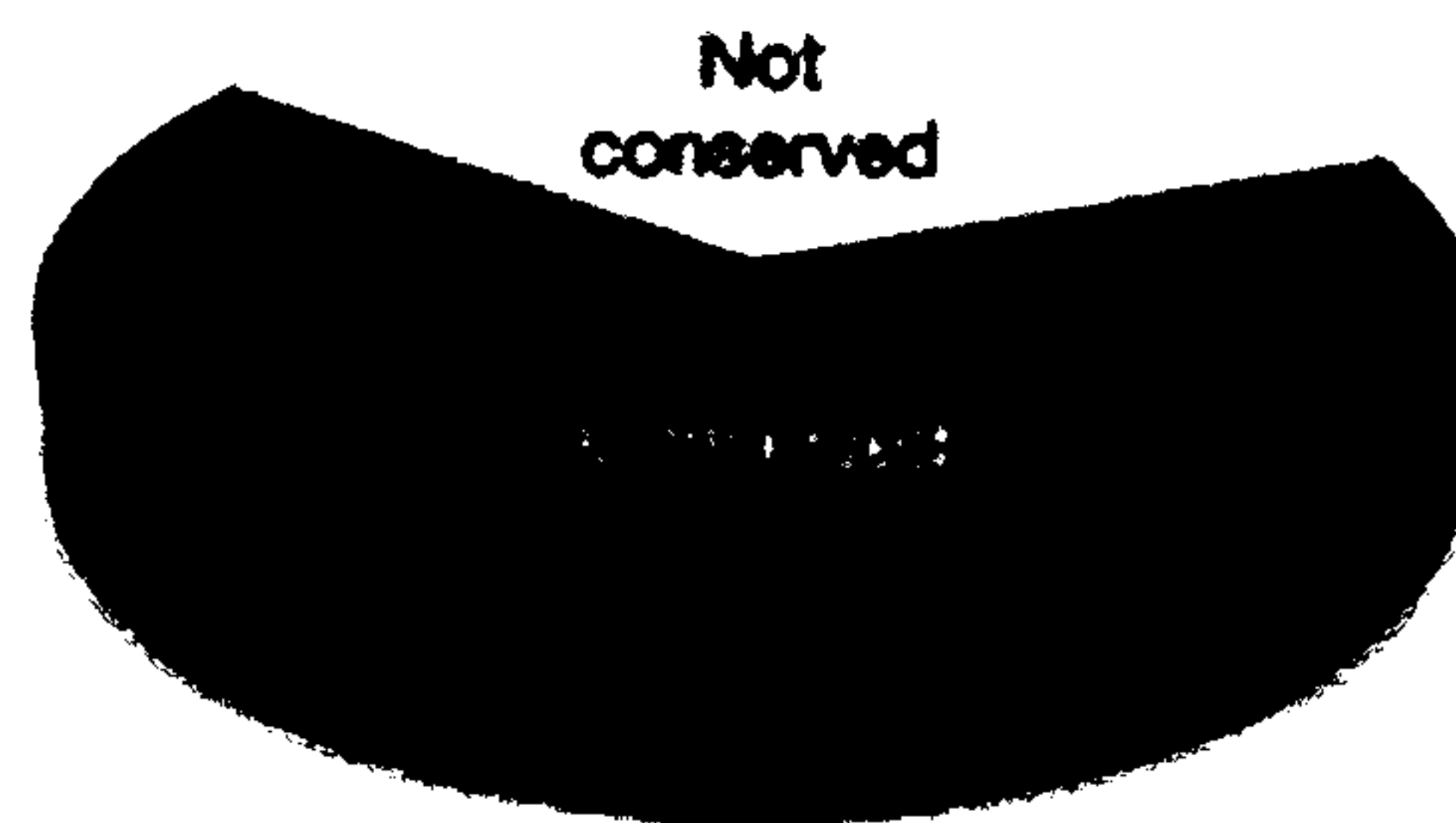


Fig. 30B

Conserved DMF-sensitive human
cysteines also quantified and
sensitive to DMF in mice

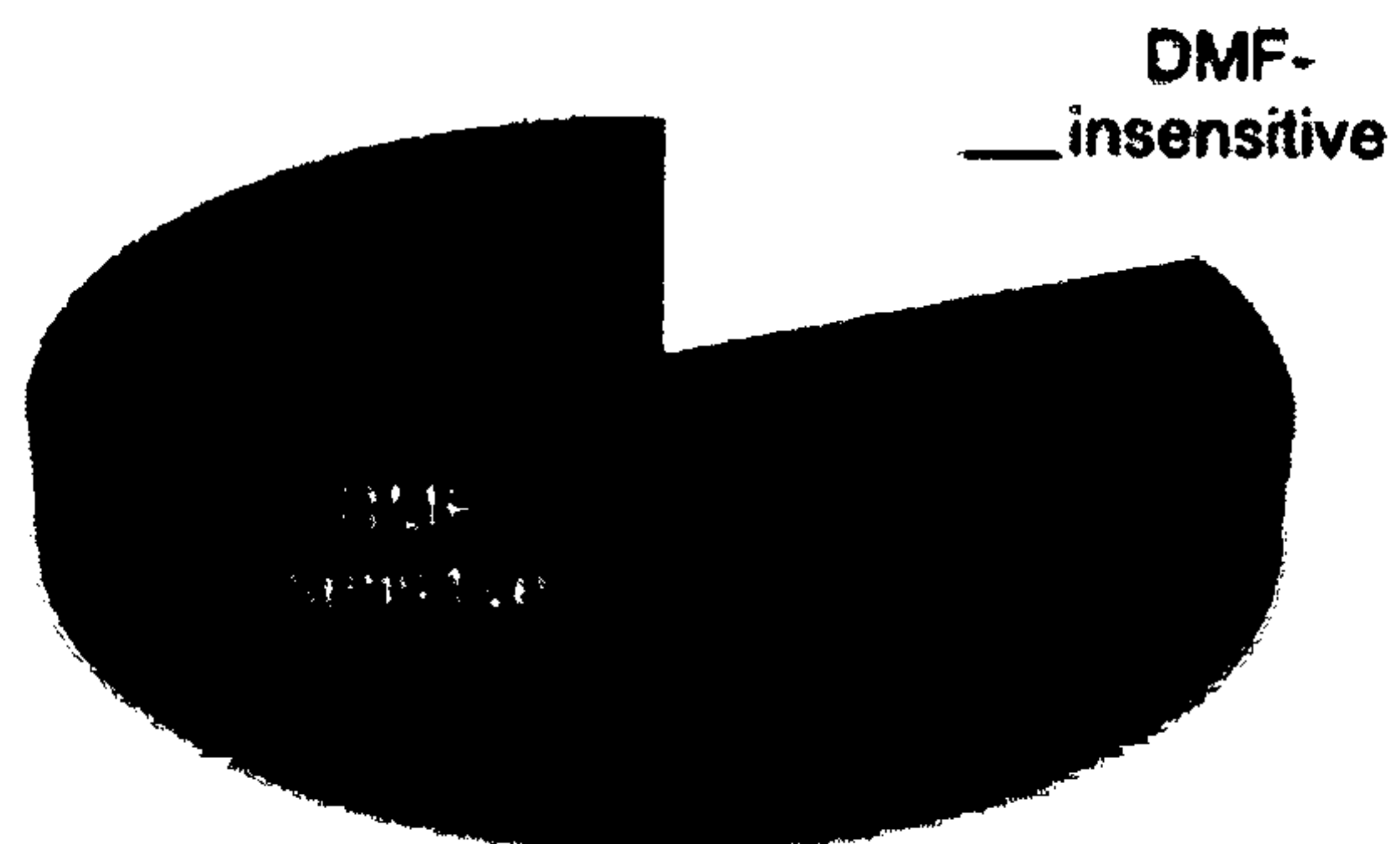


Fig. 30C

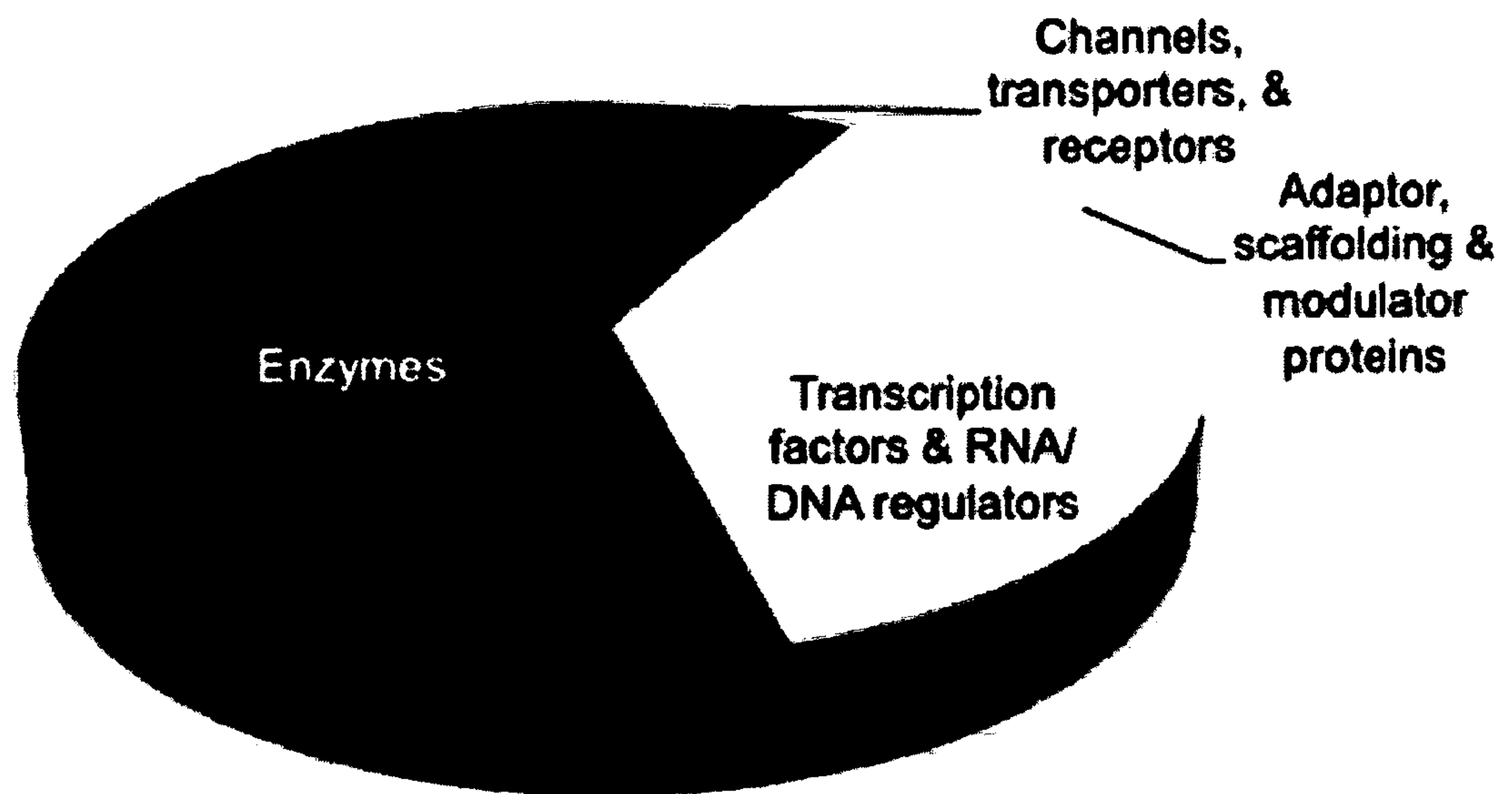


Fig. 31A

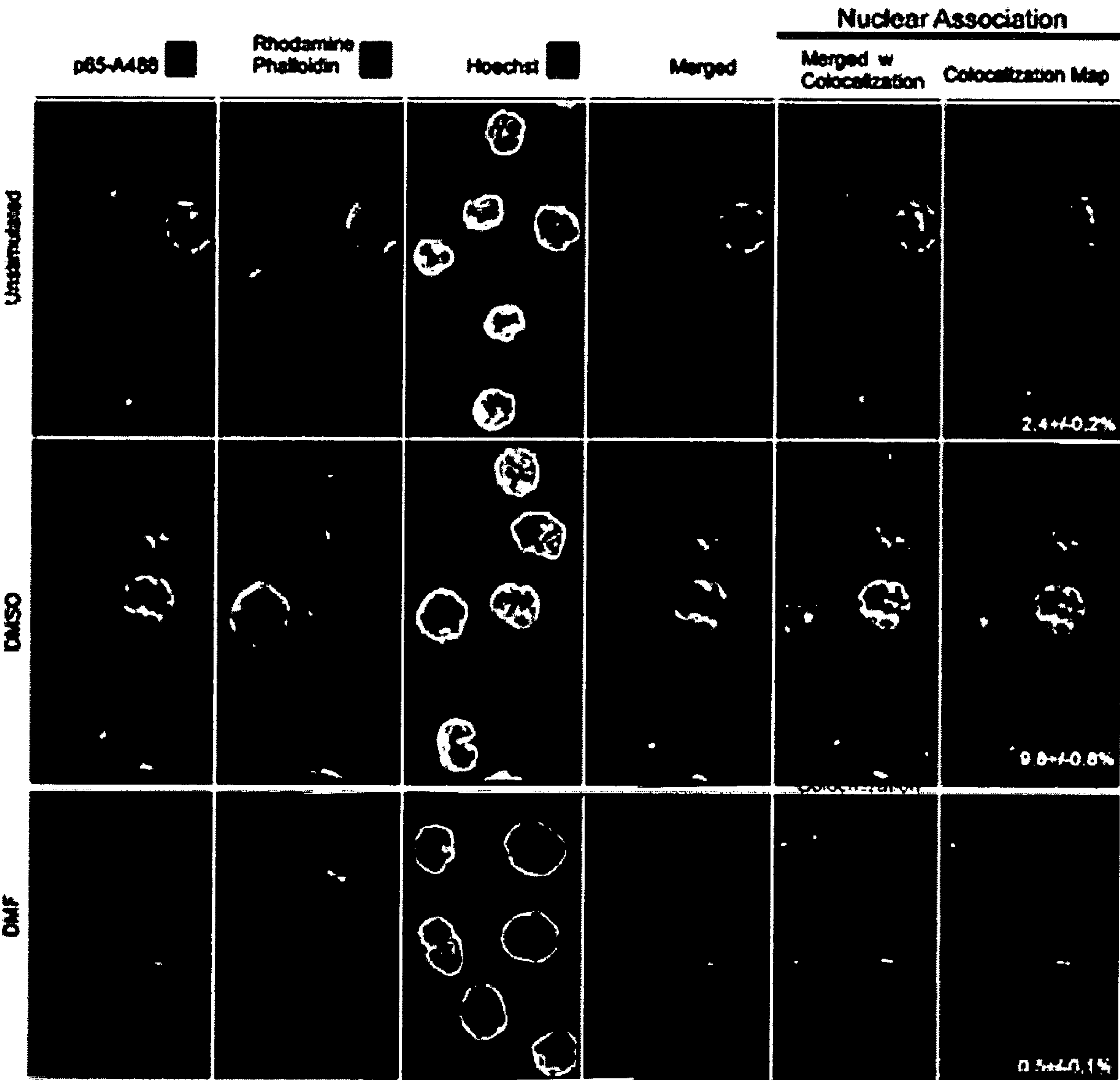


Fig. 31B

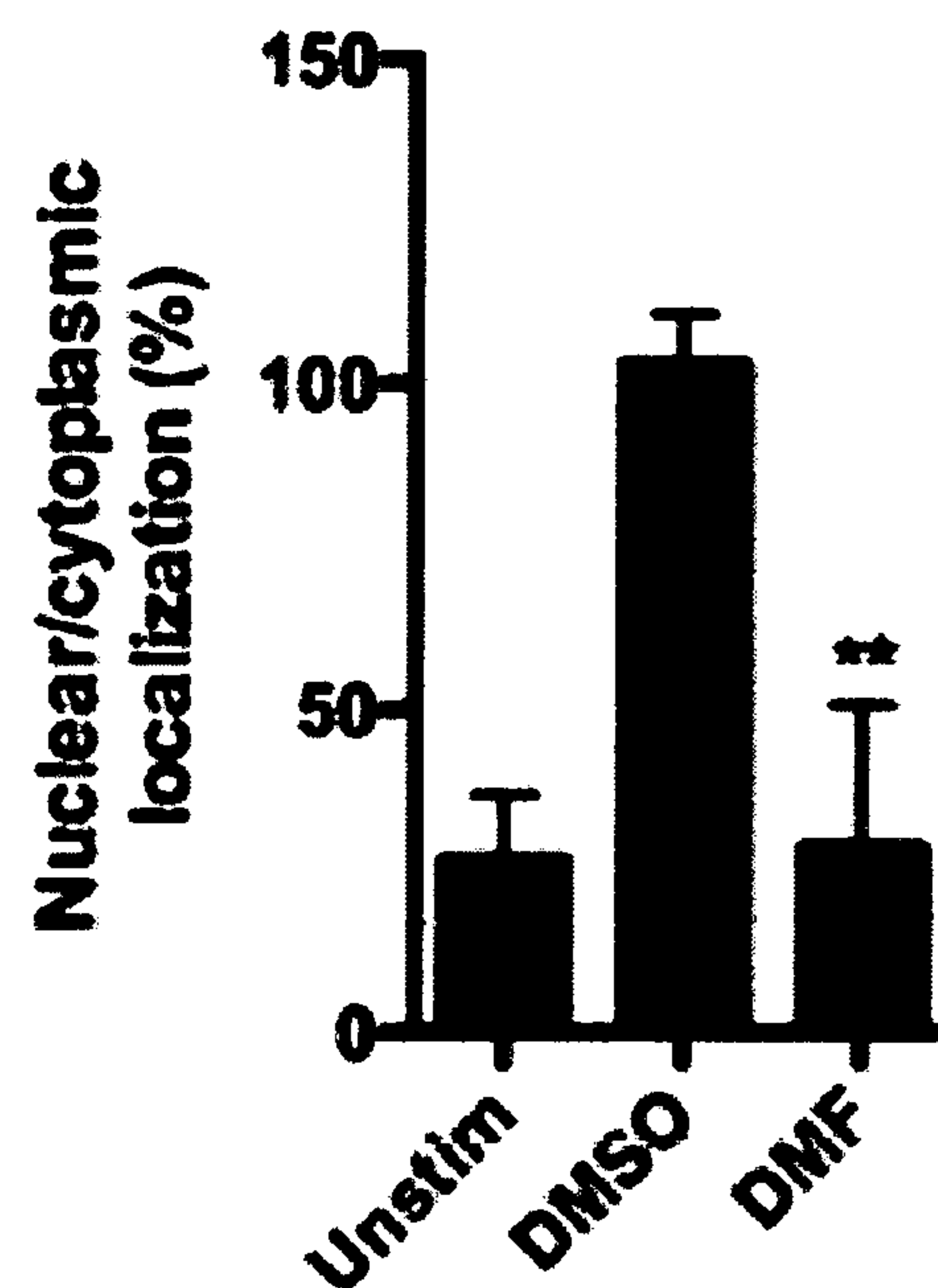
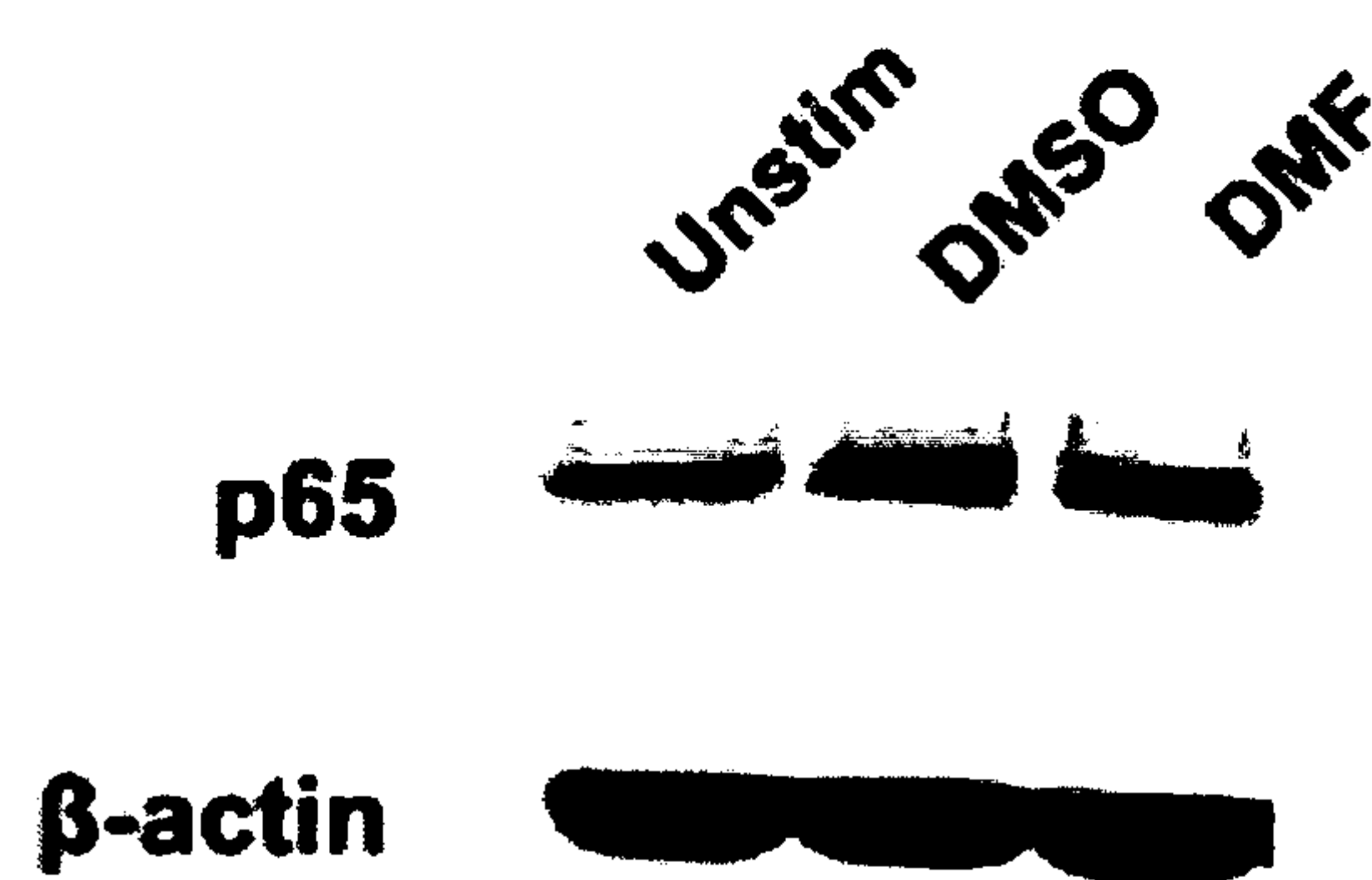
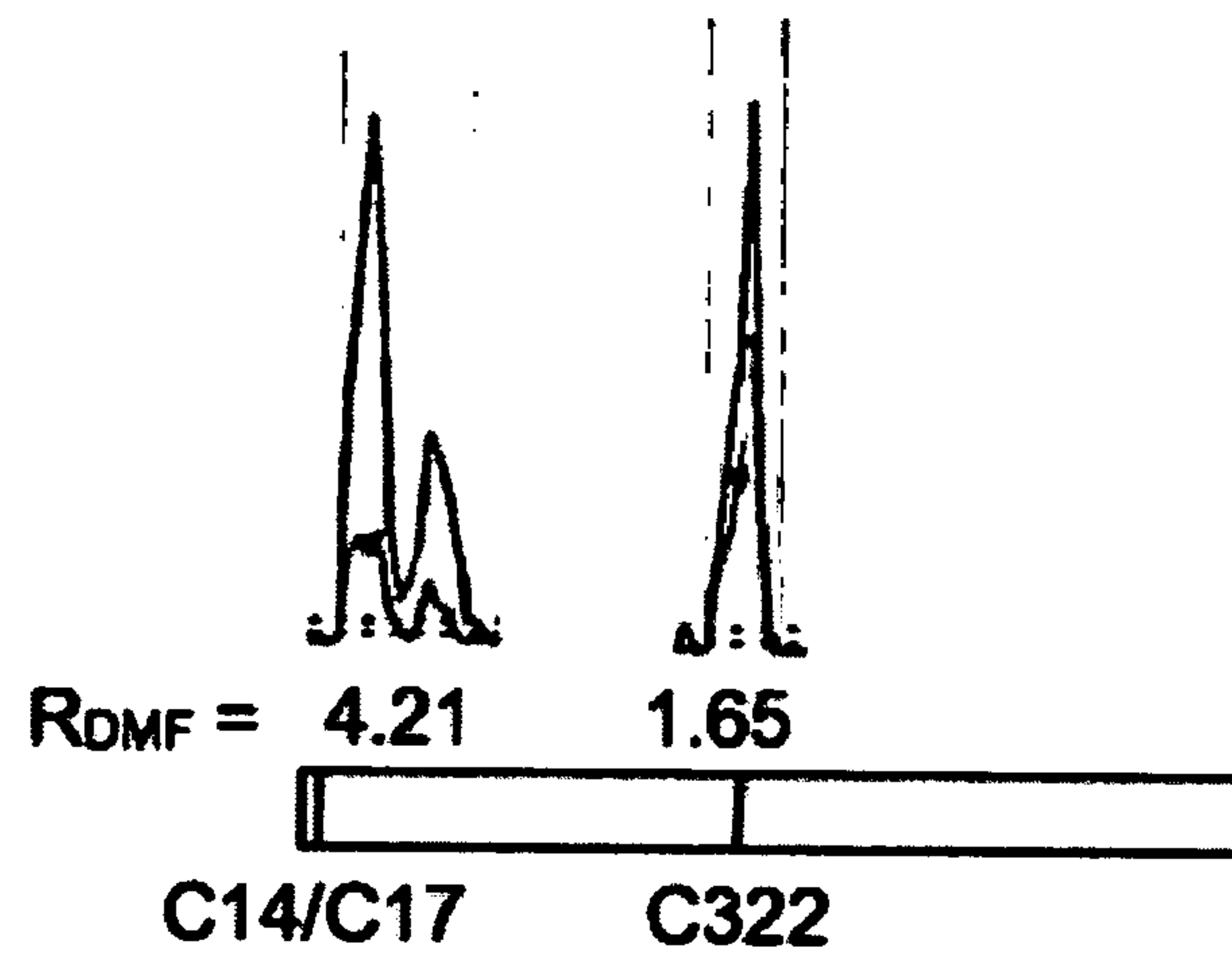


Fig. 31C





PKCθ (human) -----**SNFDCG--SCQSCQGEAVNP**

PKCθ (mouse) -----**SNFDCG--TCQACQGEAVNP**

PKCδ (human) -----**NSYELG--SLQA-EDEANQP**

PKCε (human) **AVSLKPTAWSLRHVGPRPOTFLLDPL**

· · : : *

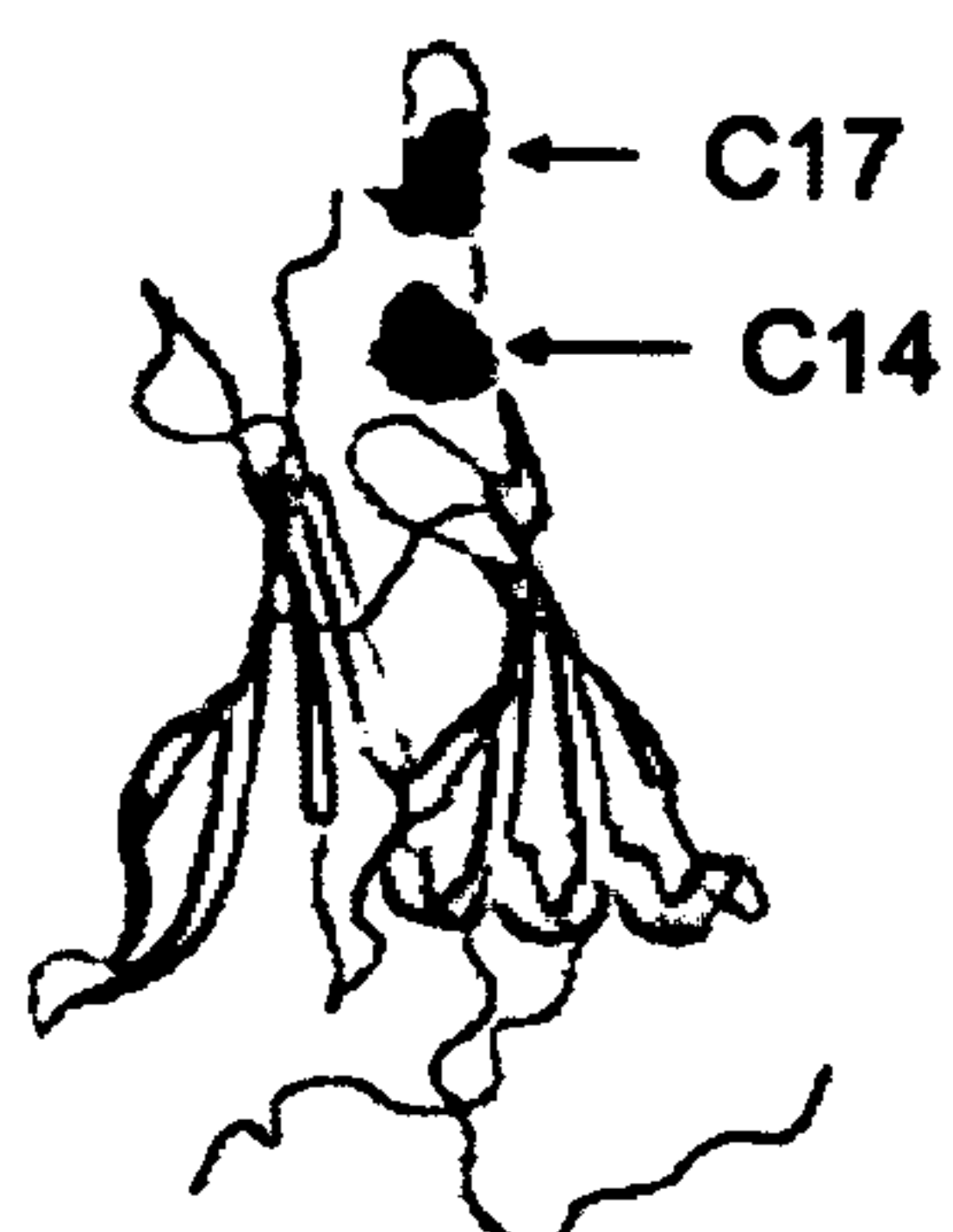


Fig. 32D

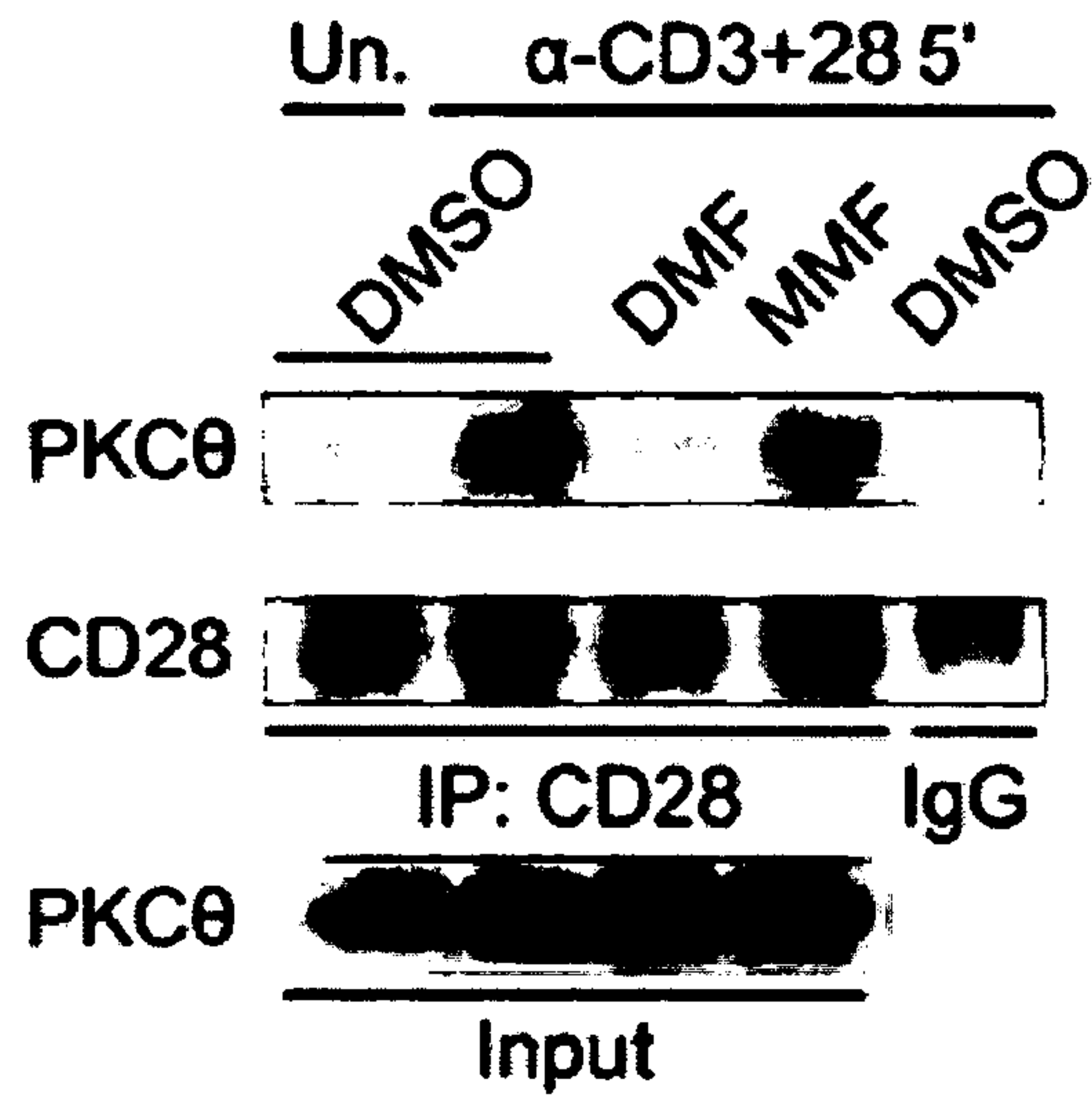


Fig. 32E

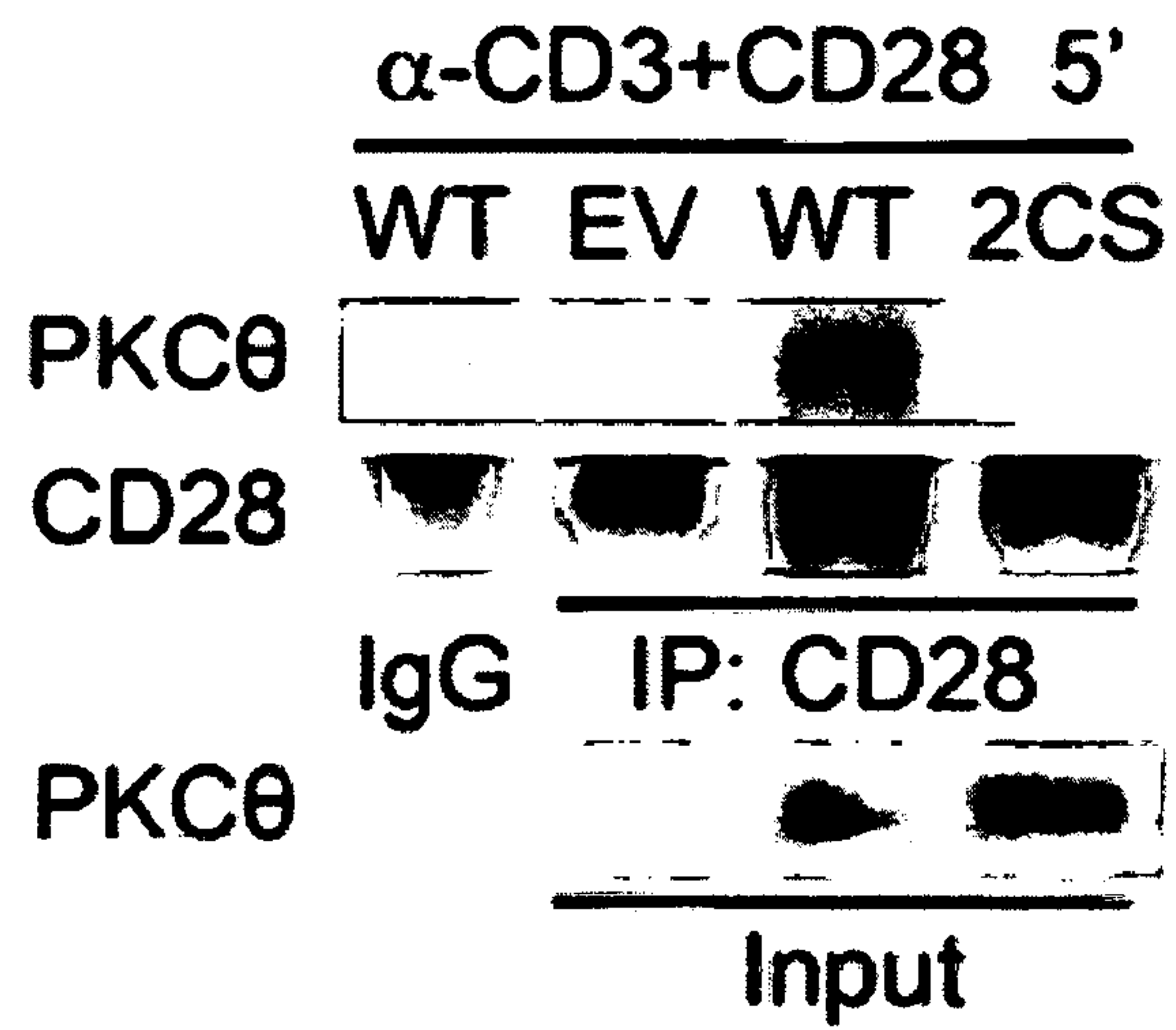


Fig. 32F

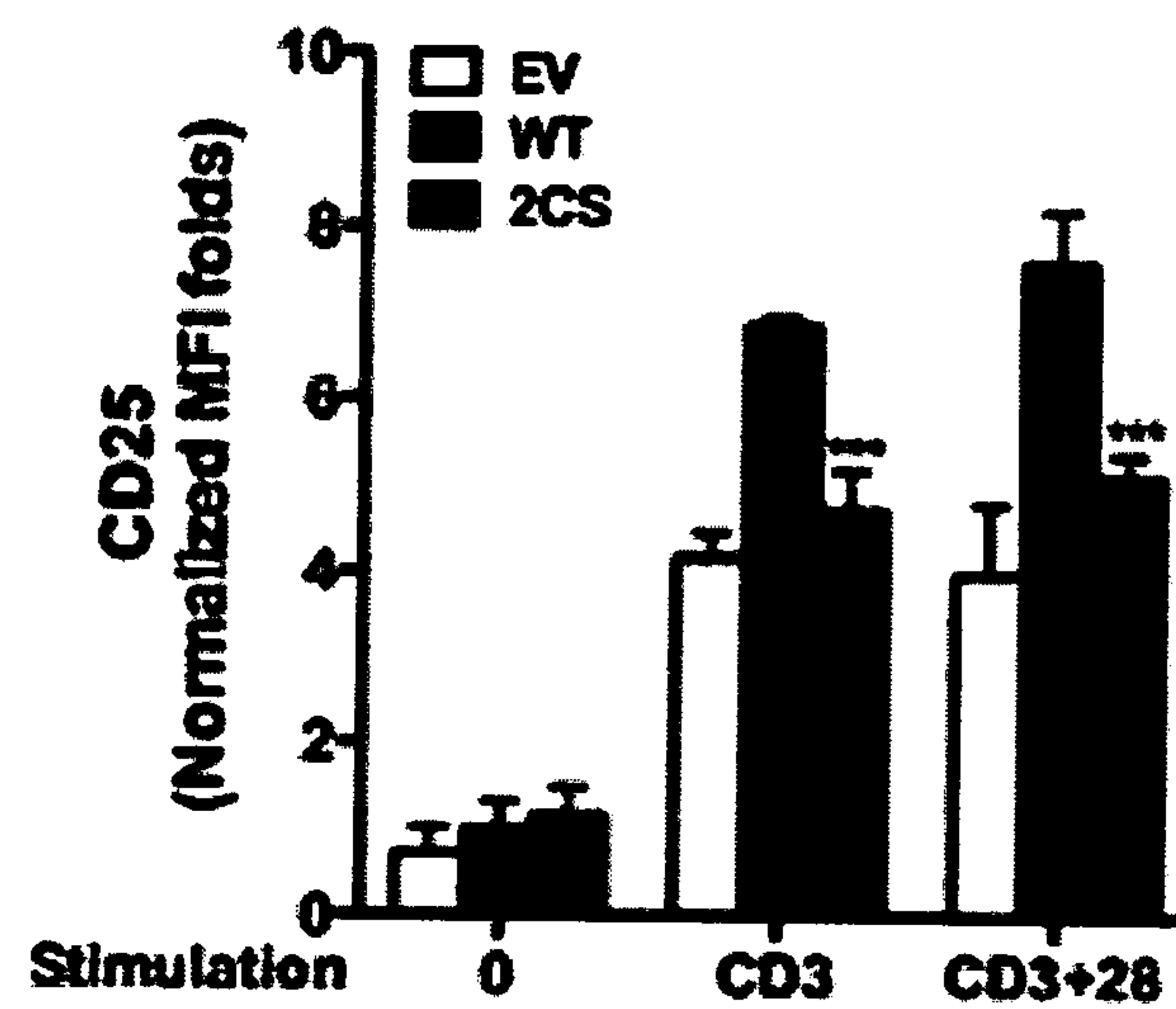


Fig. 32G

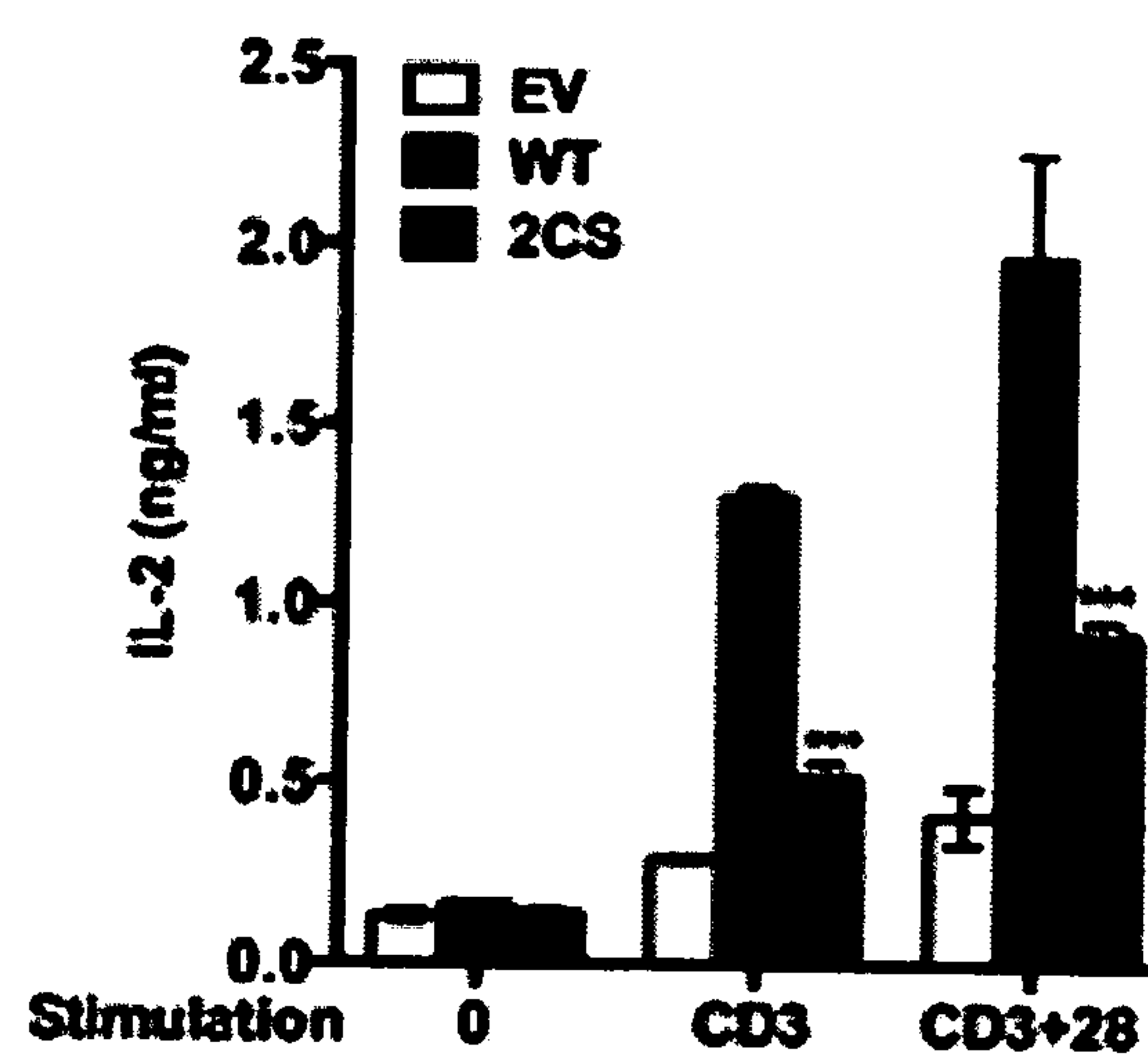
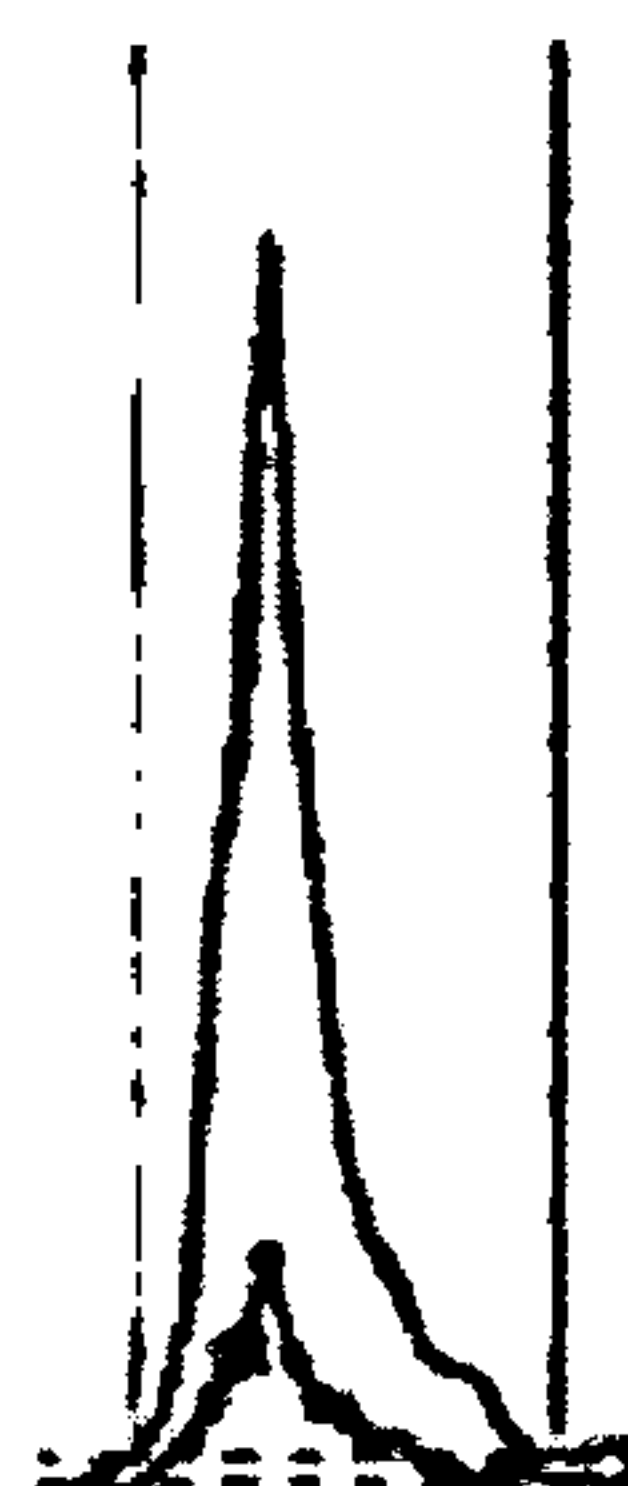
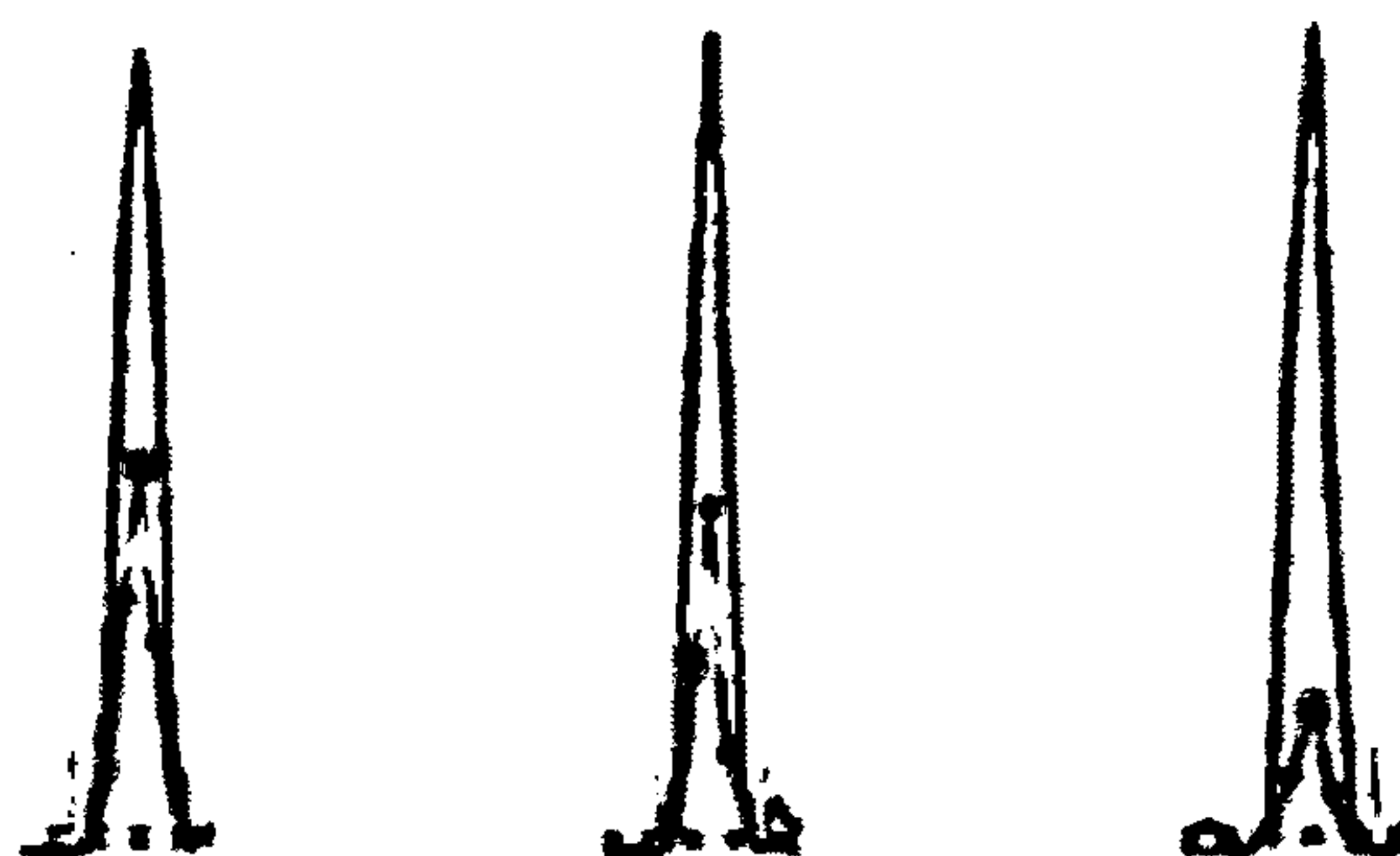


Fig. 33A



R_{DMF} 12.77

Fig. 33B



Time (hours)

1

2

4

R_{DMF}

2.24

2.92

4.21

Fig. 33C

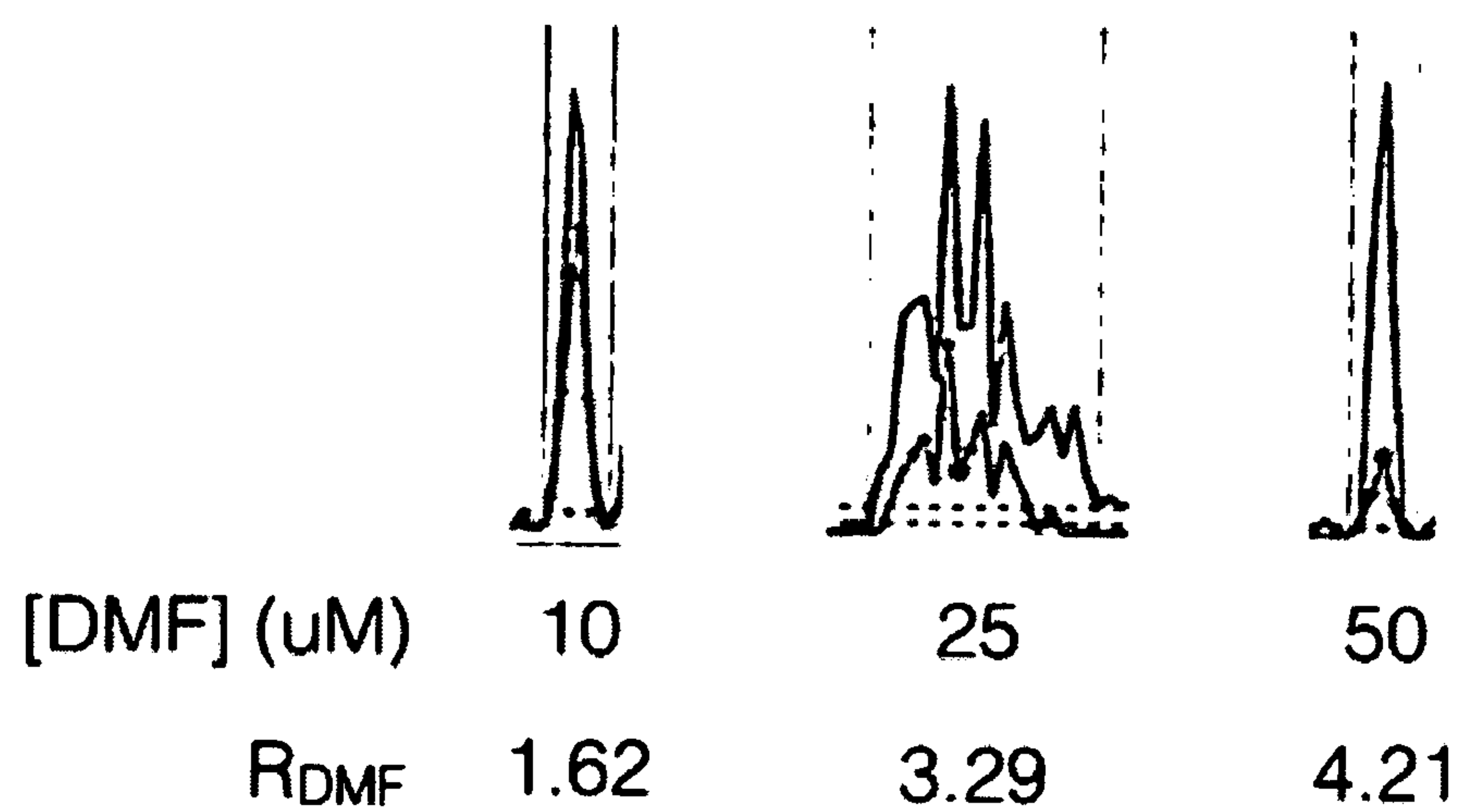


Fig. 33D

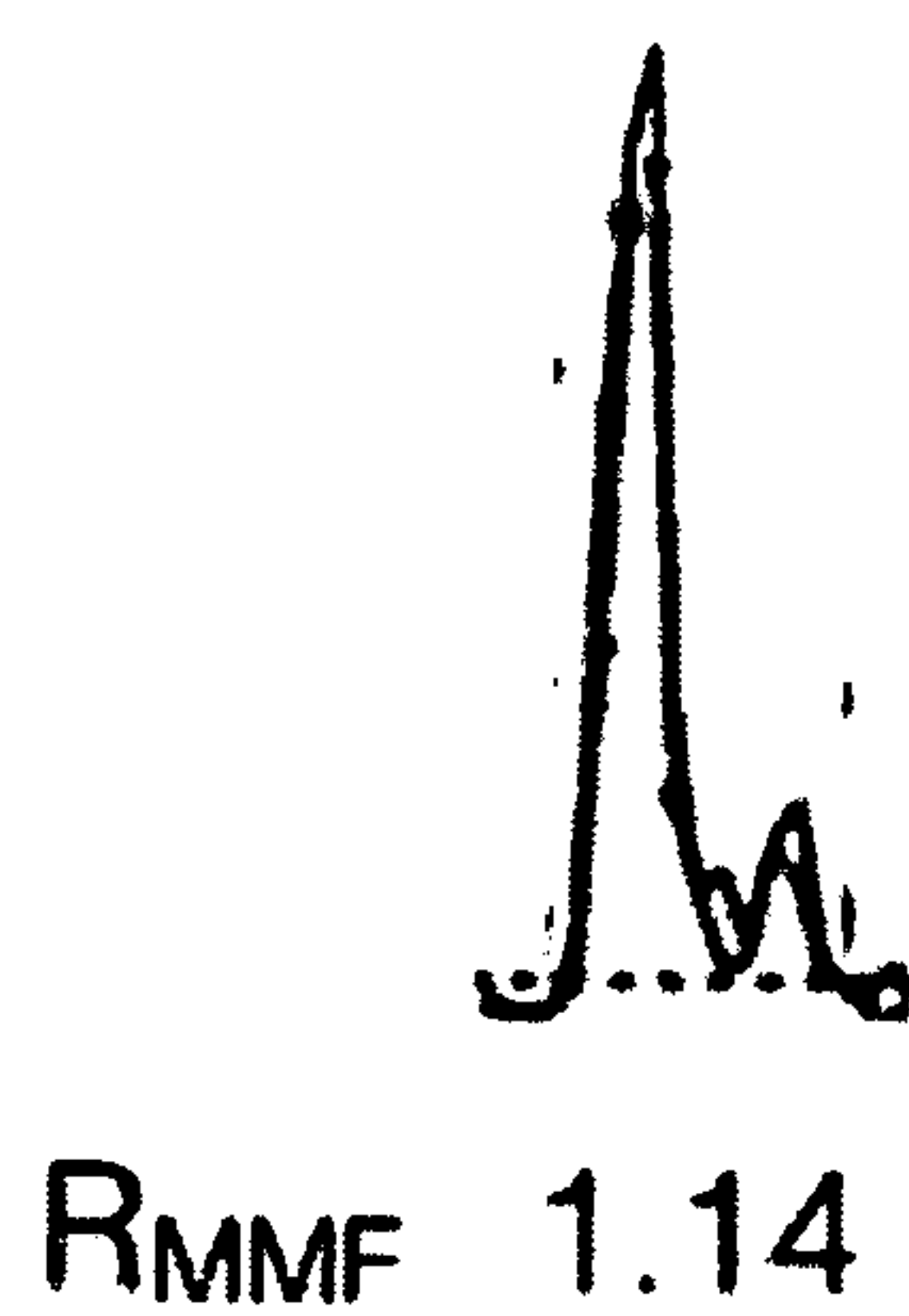


Fig. 34A**Fig. 34B**