THE EMBO JOURNAL

Manuscript EMBO-2017-96750

Protease cleavage site fingerprinting by label-free in-gel degradomics reveals novel pH dependant specificity switch of legumain

Robert Vidmar, Matej Vizovišek, Dušan Turk, Boris Turk, and Marko Fonović

Corresponding authors: Boris Turk and Marko Fonovic, Jozef Stefan Institute

Review timeline:

Submission date:	16 February 2017
Editorial Decision:	22 March 2017
Revision received:	12 June 2017
Accepted:	03 July 2017

Editor: Hartmut Vodermaier

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 22 March 2017

Thank you for submitting your manuscript on protease cleavage site fingerprinting for consideration as a Resource Article to The EMBO Journal. We have now received the comments of two expert referees, who both acknowledge the elegance and potential usefulness of your new DIPPS approach. Pending satisfactory revision of a number of specific concerns, we shall therefore be happy to consider the study further for publication.

As you will see, these few issues are well taken and clearly explained in the reports copied below, and should hopefully all be straightforward to address. With regard to point 1 of referee 2, we would consider it important to test the DIPPS approach at high temperature also using a cell lysate from a thermostable organism for comparison, as suggested by the reviewer.

REFEREE REPORTS

Referee #1:

Vidmar, Vizovisek and colleagues have developed an elegant protease substrate profiling method that uses an SDS-denatured human protein extract as a substrate library. They initially separate a protein extract on an SDS-PAGE gel, excise gel bands and digest the embedded protein into peptides with a protease of interest. The protocol used for trypsin substrate specificity profiling is not dissimilar from previously published in-gel digestion protocols such as those published by the Mann group [Nature Protocols 1, - 2856 - 2860 (2007)]. However, instead of using mass spectrometry to catalog the proteins in the gel bands, these researchers use the peptide sequencing information to evaluate protease cleavage specificity. Using this technique, several proteases were characterized, some of which were assayed under conditions of decreasing pH and increasing

temperature. Below are 3 minor comments and 1 major comment that should be addressed:

Minor Comment 1:

Page 5: "With 85 % Asp and 8 % Glu P1 residues identified, the P1 position remains the most important for the caspase-7 substrate recognition."

Please add a sentence to describe what the other 7% of non-Glu and non-Asp cleavage sites are. Are these false positives, or are they simply cleavage events that occurred in the cell lysate prior to treatment with caspase-7?

Minor Comment 2:

Throughout the manuscript the authors use an apostrophe (') to indicate prime side residues. This should be changed to a prime (′) symbol

Minor Comment 3:

"The specificities of cathepsins L, S and V were highly similar and differed slightly from cathepsin K. The most important difference was probably the acceptance of Pro in the P2 position (12%) in the latter, which is linked with the collagenolytic activity of the enzyme".
Since cathepsin L, S and V are so similar, can the author comment in the discussion on whether these enzymes are likely to have distinct functional roles within lysosomes.

Major Comment 1:

In Figure 7A, it is intriguing the substrate specificity of legumain broadens so dramatically as pH decreases and readily accepts Asp residues in the S1 pocket at pH 5.0 and pH 4.0. The authors mention that "the increased specificity for hydrophobic residues at P2 can be attributed to the increase in ionic strength introduced by the acidic medium that facilitates hydrophobic interactions between the substrate and legumain."

This comment is related to Val and Leu and to a lesser extent Ile, Phe and Tyr that are frequently found in the P2 position in the pH 5.0 and 4.0 assays. The authors do not show the amino acids that are below the X-axis (negatively selected) in the iceLogo plots however, Ala is not a preferred amino acid in the P2 position at both pH 5.0 and 4.0 [determined by reviewer using iceLogo software (Version 1.2), Homo sapien as negative set and Datasets EV13 and EV14]. Therefore, it is unclear why AAN and AAD substrates were used in Figure 7E since both have P2-Ala. This assay should be repeated with substrates consisting of Leu or Val in the P2 position and Asn and Asp in the P1 position in order to validate the interesting substrate profiling data generated in Figure 7A. In addition, the amino acids below the X-axis should be shown in Figure 7A as this information is clearly important.

Referee #2:

In this manuscript, Vidmar and co-workers present a simple, but clever way of characterizing protease specificity using in-gel digestion of cell lysate followed by mass spectrometry. They called their method Direct in-gel profiling of protease specificity (DIPPS). Considering the complexity of many of N-terminomics and degradomics approaches available to study proteases, this technique could be appealing to the scientific community, as it requires no particular homemade bioinformatics, chemical or enzymatic reagents. The authors applied their methodology to study the specificity of several proteases, such as trypsin, GluC, caspase-7 and legumain, as well as matrixmetalloproteinase-3, thermolysin and cathepsins K, L, S and V. Biologically, the authored didn't find anything new, other than perhaps a small change in the specificity of legumain at pH 4.0 compared to 5.0 and 6.0. Nonetheless, this study is well executed and brings a new and simple method to study protease specificity that warrants publication.

However, although acknowledged by the authors in the discussion section, this methods fall short in many aspects compared to current methods used to identify biological substrate of proteases, and

this should be emphasized a little more in the text - this method should probably not be used to identify new biological substrates, as all substrates are denatured before proteolysis. Moreover, because of the limiting size of the fragments extracted from the gel, any inter-domain cleavage will most likely not be observed; and often these are the most important ones.

Comments:

1. The protease specificity study at high temperature is interesting, but it's unclear how useful this is. The argument that most proteins denature at 75C is correct, for human proteins. But wouldn't it be more relevant to use a cell lysate originating from the same thermostable organism (e.g. Bacillus thermoproteolyticus), which I assume would be thermostable?

2. The specificity of legumain changed as a function of pH, but what about it's activity (kcat/Km)? Any changes by comparing the z-AAN-AMC and z-AAD-AMC?

3. Page 10. "The S1 binding pocket is formed by amino acid side chains that are more flexible than the main chain atoms. The absence of significant differences in the crystal structures determined at different pH values suggests that the changes in pH may enable induced fit into S1 pocket caused by the residues other than Asn and Asp." Unless the authors want to perform dynamics studies to support their claim, I would suggest taking this statement out.

Minor comments:

1. The authors should probably define DIPPS in the abstract.
2. The authors failed to cite original N-terminomics papers: COFRADIC (Gevaert 2003), PROTOMAP (Dix 2008), subtiligase (Mahrus 2008), TAILS (Kleifeld 2010). O'Donoghue et al. 2012 should also be appropriate to acknowledge, describing the use of peptide libraries to characterize protease specificity.
3. Please reference correctly the original papers reporting caspase-7 substrates (page 9.)

1st Revision - authors' response 12 June 2017

We would like to really thank the reviewers for their very positive and valuable comments on our manuscript, which we have positively addressed as described below. We therefore hope that you are now able to accept the manuscript for publication in the EMBO Journal.

Response to Referee #1:

Major comment 1:

In Figure 7A, it is intriguing the substrate specificity of legumain broadens so dramatically as pH decreases and readily accepts Asp residues in the S1 pocket at pH 5.0 and pH 4.0. The authors mention that "the increased specificity for hydrophobic residues at P2 can be attributed to the increase in ionic strength introduced by the acidic medium that facilitates hydrophobic interactions between the substrate and legumain."

This comment is relatied to Val and Leu and to a lesser extent Ile, Phe and Tyr that are frequently found in the P2 position in the pH 5.0 and 4.0 assays. The authors do not show the amino acids that are below the X-axis (negatively selected) in the iceLogo plots however, Ala is not a preferred amino acid in the P2 position at both pH 5.0 and 4.0 [determined by reviewer using iceLogo software (Version 1.2), Homo sapien as negative set and Datasets EV13 and EV14]. Therefore, it is unclear why AAN and AAD substrates were used in Figure 7E since both have P2-Ala. This assay should be repeated with substrates consisting of Leu or Val in the P2 position and Asn and Asp in the P1 position in order to validate the interesting substrate profiling data generated in Figure 7A. In addition, the amino acids below the X-axis should be shown in Figure 7A as this information is clearly important.

We would like to thank the Referee for this extremely valuable comment. A very thorough analysis of the legumain sample namely revealed presence of trace amounts of the cathepsin L-like insect

cysteine protease from Spodoptera frugiperda (i.e. the expression system), which could be only identified by activity-based probe labelling (DCG-04) and MS analysis with much larger amounts of legumain searching below the threshold values. Inhibition of this protease by the general cathepsin inhibitor E-64 revealed that the P1 preference for Asn and Asp was the result of legumain activity, while the observed preference for Val and Leu in the P2 position was actually an artifact originating from the insect cysteine cathepsin background. Since the protease is extremely unusual having a pH optimum at pH 3.0 and low activity at pH 6.0, the Val/Leu preference at pH 6.0 was hidden due to the low number of cleavages as compared to legumain. We have therefore corrected this problem by repeating DIPPS profiling of legumain in the presence of general cathepsin inhibitor E-64 and we accordingly modified the Figures 7 and EV5, which now show the correct legumain data. The new profiling data was also deposited in the PRIDE depository (PRIDE data accession information is listed at the end of the letter). We have also repeated the degradation profile of BSA with legumain in the presence of E-64 (Fig. 7C) and made appropriate corrections in the Results (Page 8), Discussion (Page 10) and Materials and Methods (Page 13) sections of the text. However, this does not affect the performance and applicability of the DIPPS methodology in any way.

As suggested, we have also included aminoacids below X-axes in Figure 7A. Since the enrichment of Leu and Val residues at P2 position was found to be an artifact, the rationale for kinetic testing of additional substrates was no longer valid. However, such experiment was recently performed by small peptide library scan (Poreba et al., 2016), which showed that replacement of Ala with Leu or Val at the P2 position did not affect the cleavage efficiency of the small peptide substrates, which is now also cited in the manuscript. We have also modified the phrase "novel pH dependant specificity switch" into "pH dependant specificity switch" in the title of the paper.

Minor Comment 1:

Page 5: "With 85 % Asp and 8 % Glu P1 residues identified, the P1 position remains the most important for the caspase-7 substrate recognition."
Please add a sentence to describe what the other 7% of non-Glu and non-Asp cleavage sites are. Are these false positives, or are they simply cleavage events that occurred in the cell lysate prior to treatment with caspase-7?

About 1% of non-Glu and non-Asp cleavage sites can be attributed to false positives, since 1% false discovery rate (FDR) was applied during the database search. Furthermore, endogenous proteolytic degradation and processing of proteins is constantly present inside the living cells. Such degradation products are therefore present also in the cell lysate and detected by DIPPS as the nonspecific proteolytical background. As the reviewer proposed, we described this by including the following sentence on Page 5: "The remaining 7% of the identified random cleavages can be attributed to false positives (1% false discovery rate) and background proteolysis that happened prior to lysate preparation and caspase treatment."

Minor Comment 2:

Throughout the manuscript the authors use an apostrophe (') to indicate prime side residues. This should be changed to a prime (′) symbol.

As suggested, we have corrected the apostrophe to a prime symbol.

Minor Comment 3:

"The specificities of cathepsins L, S and V were highly similar and differed slightly from cathepsin K. The most important difference was probably the acceptance of Pro in the P2 position (12%) in the latter, which is linked with the collagenolytic activity of the enzyme." Since cathepsin L, S and V are so similar, can the author comment in the discussion on whether these enzymes are likely to have distinct functional roles within lysosomes.

Cathepsins L, V and S are generaly known to have similar aminoacid specificity preference. However, that does not mean that they have the same cellular functions since their biochemical properties and tissue localisation differ significantly. For example cathepsins L and V that also

share a high degree of sequence homology (78%), have completely different expression patterns. Contrary to cathepsin L, which is ubiquitously expressed, cathepsin V expression is more limited. In addition, cathepsin V was found to bind DNA, in contrast to cathepsin L. Similarly, cathepsin S is expressed mainly by the antigen presenting cells. Such different cellular functions were seen also from the knockout experiments, where cathepsin S knockout was found to have defective processing of the invariant chain of the MHC II complex and thus defective MHC II antigen presentation in all spleen-derived cells, whereas cathepsin L (in mice, V in human) was responsible for invariant chain processing in thymus. As suggested, we have commented this on page 9: "Although similar specificity of cathepsins L, V and S imply certain degree of functional redundancy, due to their different localization they perform also individual tissue-specific functions (Turk et al, 2012b)."

Response to Referee #2:

However, although acknowledged by the authors in the discussion section, this methods fall short in many aspects compared to current methods used to identify biological substrate of proteases, and this should be emphasized a little more in the text - this method should probably not be used to identify new biological substrates, as all substrates are denatured before proteolysis. Moreover, because of the limiting size of the fragments extracted from the gel, any inter-domain cleavage will most likely not be observed; and often these are the most important ones.

We would like to thank the reviewer for this comment and we apologize if we did not explain this more clearly in the Discussion section. DIPPS was never intended for the identification of physiological substrates and cleavage sites in vivo, as we have originally stated on the Page 10: "Although the use of linearized proteins (polypeptides) in DIPPS precludes identification of natural substrates of proteases, the high number of cleavage sites resulting in high accuracy in determining protease specificities more than compensates this drawback". The sole purpose of DIPPS is determination of primary aminoacid specificity preferences, which can be very valuable in assay development. To emphasize this, we have mentioned it again at the end of Page 11, where we modified the sentence "Application of DIPPS is limited only in the case of extremely large proteases (e.g. proteasome), which could not be absorbed into polyacrylamide gel." into "However, DIPPS also has some limitations. Because of the use of denatured proteins it is not suitable for identification of physiological substrates. In addition, its applicability is limited in the case of extremely large proteases (e.g. proteasome), which are not efficiently absorbed into a polyacrylamide gel, or exopeptidases."

Comment 1:

The protease specificity study at high temperature is interesting, but it's unclear how useful this is. The argument that most proteins denature at 75°C is correct, for human proteins. But wouldn't it be more relevant to use a cell lysate originating from the same thermostable organism (e.g. Bacillus thermoproteolyticus), which I assume would be thermostable?

Specificity profilling of thermophylic proteases is problematic since their proteolytic efficiency is highest at their optimal temperature which is usually in the range of 60-90°C. Such conditions are not compatible with N-terminomics approaches which use in solution digestion of protein library (COFRADIC, TAILS, FPPS and N-terminal biotinylation), since complex proteomes (e.g. cellular lysates) tend to denature and precipitate at elevated temperatures. This is one of the reasons that thermophylic proteases have so far only been profiled by small peptide library screens and not by proteomic approaches. The result is that specificity data of thermophylic proteases remains scarce. Even thermolysin which is the most studied thermophylic protease has specificity profile based on solely 280 cleavages (MEROPS database). Using a cell lysate from thermostable organisms might elevate the problem with precipitation in solution, but thermophylic organisms are not readilly avaliable in biochemical laboratories since they require very specific growth conditions. Many of them also do not have reliable protein databases which are required for proteomic identification of peptides. The benefit of DIPPS is that it utilises a library of fully denatured and gel-imbeded proteins, which cannot precipitate under elevated temperature. This means that they can be processed also at conditions that are optimal for thermophylic proteases, regardless of the source of the lysate. As the reviewer suggested, and to show that DIPPS profiling is not dependent on the lysate source, we profiled thermolysin at 75°C using a human cell lysate and a cell lysate prepared from hyperthermophilic archea Aeropyrum pernix. A pernix is an obligatory thermophile with

optimum growth temperature at 90-95°C. We chose to include this organism in our study because it has even higher temperature tolerance than Bacillus thermoproteolyticus and because it has a larger and much better annotated protein database (B thermoproteolyticus has only 163 protein entries in the NCBI database (and which would preclude any meaningful analysis), while A pernix has 2358 entries). A comparison of the profiles obtained by both lysates revealed identical specificity profiles of thermolysin, demonstrating that the protein library obtained from a thermostable organism does not influence the specificity profile of the protease, and that common cell lysates can be used also for the specificity profiling of thermophylic proteases. We included the data obtained from A pernix lysate into Figure 6D, E and Dataset EV11. We have also commented on this in the Results section (Page 7), where we have added the sentence: »For comparison, we have additionally performed DIPPS profiling of thermolysin using the cell lysate of the obligatory thermophile Aeropyrum pernix. Both datasets correlated well and indicated that the most important substrate specificity determinant was the S1' subsite, primarily accommodating aliphatic and aromatic amino acid residues (Fig 6D, 6E and Dataset EV11).«

Comment 2:

The specificity of legumain changed as a function of pH, but what about it's activity (kcat/Km)? Any changes by comparing the z-AAN-AMC and z-AAD-AMC?

As suggested, we have determined the kcat/Km values for the hydrolysis of tested substrates and included these data in the new Figure 7 and Appendix TableS1. At pH=6 and above the kcat/Km value for z-AAN-AMC substrate hydrolysis is at least 300-fold higher than for z-AAD-AMC. However, at pH 4, the difference in kcat/Km value between both substrates is only about 4-fold. We have commented on this result in the Results section (Page 8) where we stated: »This pH-dependent shift of P1 specificity from Asn to Asp was also seen during the hydrolysis of small substrates z-AAD-AMC and z-AAN-AMC. A comparison of k_{cat}/K_M values thus showed that at pH 6.0 z-AAN-AMC was cleaved over 300-fold more efficiently than z-AAD-AMC, whereas at pH 4.0 this difference was only about 4-fold (Fig 7D, Appendix Table S1).« In addition, the number of cleavages as determined by DIPPS profiling at all pH values was substantially higher than expected based on small substrate hydrolysis, suggesting differential binding of small and large substrates, which is now also added in the Discussion section (p. 10, last paragraph): "However, the total number of cleavages as well as the percentage of cleavages after Asp observed in large substrates at acidic pH were much higher as that in small fluorogenic substrates (z-AAN-AMC and z-AAD-AMC), suggesting differential binding of the substrates. While small substrates bind only to the non-prime side of the active site cleft, large polypeptide substrates occupy the entire active site on prime- and non-prime side (Turk et al., 2012, EMBO),), possibly explaining the difference."

Comment 3:

Page 10. "The S1 binding pocket is formed by amino acid side chains that are more flexible than the main chain atoms. The absence of significant differences in the crystal structures determined at different pH values suggests that the changes in pH may enable induced fit into S1 pocket caused by the residues other than Asn and Asp."
Unless the authors want to perform dynamics studies to support their claim, I would suggest taking this statement out.

We would like to thank the reviewer also fort his comment. However, in light of new results on legumain specificity, which confirmed only the pH-dependent Asn to Asp switch, this statement became redundant and was deleted as suggested.

Minor comment 1:

The authors should probably define DIPPS in the abstract.

We have modified the abstract by including the definition of DIPPS, as suggested.

Minor comment 2:

The authors failed to cite original N-terminomics papers: COFRADIC (Gevaert 2003), PROTOMAP (Dix 2008), subtiligase (Mahrus 2008), TAILS (Kleifeld 2010). O'Donoghue et al. 2012 should also be appropriate to acknowledge, describing the use of peptide libraries to characterize protease specificity.

We completely agree with the reviewers suggestion that original methodological references should also be included in the paper. We have therefore modified the manuscript accordingly and cited the N-terminomics papers in the Introduction section (Page 3) (Gevaert et al., 2003; Mahrus et al., 2008; Kleifeld et al., 2010; O'Donoghue et al., 2012). However, it should be noted that the PROTOMAP methodology (Dix et al., 2008) is not an N-terminomics approach. PROTOMAP detects protein cleavages based on the protein positional shift on SDS-PAGE gel and can not be applied to cleavage specificity profiling since it can not identify the actual cleavage sites. This is the main reason that this reference could not be cited in the paper.

Minor comment 3:

Please reference correctly the original papers reporting caspase-7 substrates (page 9.)

We have corrected the citations regarding the caspase-7 substrates. Reference "Agard and Wells, 2009" was replaced by reference "Agard et al., 2012", which we believe is the original reference suggested by the reviewer.

2nd Editorial Decision 3 July 2017

Thank you for submitting your revised manuscript for our consideration. Referee 1 has now assessed it once more, and I am happy to say that in light of his/her feedback, we consider all scientific issues satisfactorily addressed and the study suitable for publication in The EMBO Journal.

REFEREE REPORT

Referee #1:

The authors have fully addressed all of my comments, both major and minor. They have performed additional experiments and have updated the methods and results section appropriately. I have determined that the findings and conclusions are of "high" general significance because the technique described here can be performed in any biochemical laboratory that has access to an LC-MS/MS instrument.

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name:	Prof. Dr. Marko Fonovic
Journal Submitted to:	EMBO Journal
Manuscript Number:	EMBOJ-2017-96750

Reporting Checklist For Life Sciences Articles (Rev. July 2015)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- If n< 5, the individual data points from each experiment should be plotted and any statistical test employed should be justified
- Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.

- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple χ2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

Please ensure that the answers to the following questions are reported in the manuscript itself. We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

In the pink boxes below, provide the page number(s) of the manuscript draft or figure legend(s) where the information can be located. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable).

USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com	Antibodypedia
http://1degreebio.org	1DegreeBio
http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repo	ARRIVE Guidelines
http://grants.nih.gov/grants/olaw/olaw.htm	NIH Guidelines in animal use
http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm	MRC Guidelines on animal use
http://ClinicalTrials.gov	Clinical Trial registration
http://www.consort-statement.org	CONSORT Flow Diagram
http://www.consort-statement.org/checklists/view/32-consort/66-title	CONSORT Check List
http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tum	REMARK Reporting Guidelines (marker prognostic studies)
http://datadryad.org	Dryad
http://figshare.com	Figshare
http://www.ncbi.nlm.nih.gov/gap	dbGAP
http://www.ebi.ac.uk/ega	EGA
http://biomodels.net/	Biomodels Database
http://biomodels.net/miriam/	MIRIAM Guidelines
http://jjj.biochem.sun.ac.za	JWS Online
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html	Biosecurity Documents from NIH
http://www.selectagents.gov/	List of Select Agents

B- Statistics and general methods

Please fill out these boxes (Do not worry if you cannot see all your text once you press return)

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	N/A
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	N/A
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	N/A
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	N/A
For animal studies, include a statement about randomization even if no randomization was used.	N/A
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	N/A
4.b. For animal studies, include a statement about blinding even if no blinding was done	N/A
5. For every figure, are statistical tests justified as appropriate?	YES
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	N/A
Is there an estimate of variation within each group of data?	N/A
Is the variance similar between the groups that are being statistically compared?	N/A

C- Reagents

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	N/A
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	MDA-MB-231 cell line was obtained from ATCC (HBT-26) and was tested for mycoplasma contamination. Aeropyrum pernix cell lysate was obtained from the laboratory of prof. dr. Natasa Poklar Ulrih (University of Ljubljana)

* for all hyperlinks, please see the table at the top right of the document

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	N/A
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	N/A
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	N/A

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	N/A
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	N/A
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	N/A
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	N/A
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	N/A
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	N/A
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	N/A

F- Data Accessibility

18. Provide accession codes for deposited data. See author guidelines, under 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	Mass spectrometry raw data and database search results have been deposited tothe ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD004218.
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	See point 18
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	N/A
21. As far as possible, primary and referenced data should be formally cited in a Data Availability section. Please state whether you have included this section. Examples: **Primary Data** Wetmore KM, Deutschbauer AM, Price MN, Arkin AP (2012). Comparison of gene expression and mutant fitness in Shewanella oneidensis MR-1. Gene Expression Omnibus GSE39462 **Referenced Data** Huang J, Brown AF, Lei M (2012). Crystal structure of the TRBD domain of TERT and the CR4/5 of TR. Protein Data Bank 4O26 AP-MS analysis of human histone deacetylase interactions in CEM-T cells (2013). PRIDE PXD000208	N/A
22. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	N/A

G- Dual use research of concern

23. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	N/A