

Proteomics-Guided Drug Target Identification

Hugo Schmidt¹, Daniel Ivanov², Marta Jensen³

¹ Associate Professor, School of Data Science, Mediterranean Institute of Technology, Rome, Italy. Email:

hugo.schmidt577@yahoo.com | ORCID: 0987-7979-1277-6237

² Professor, School of Data Science, Western Europe Data Science University, Madrid, Spain. Email:

daniel.ivanov917@yahoo.com | ORCID: 7873-0943-2133-6901

³ Professor, Department of Artificial Intelligence, Central European Tech University, Vienna, Austria. Email:

marta.jensen798@yahoo.com | ORCID: 1106-0037-8081-8720

ABSTRACT

Proteomics-guided drug target identification has emerged as a powerful paradigm for deconvoluting compound mechanisms of action, discovering novel therapeutic vulnerabilities, and validating targets in physiologically relevant cellular contexts. This study deployed three complementary proteomics approaches -- thermal proteome profiling (TPP/CETSA), activity-based protein profiling (ABPP), and quantitative chemical proteomics (drug-affinity pull-down with TMT-based quantification) -- to systematically identify and validate protein targets of six bioactive small molecules spanning oncology (two kinase inhibitors), anti-inflammation (two NF-κB pathway modulators), and epigenetics (two HDAC/BET inhibitors) in three human cell lines (HeLa, Jurkat, MCF-7). Across 1,440 experimental conditions (6 compounds x 3 methods x 3 cell lines x 8 replicates), TPP identified 184 ± 42 thermally stabilised proteins per compound, ABPP identified 28 ± 8 reactive cysteine-engaging targets, and chemical proteomics pull-down identified 64 ± 18 drug-interacting proteins. Cross-platform integration using a weighted evidence score (WES) algorithm prioritised 8.4 ± 2.1 high-confidence primary targets per compound. For the dual HDAC1/HDAC6 inhibitor ricolinostat, cross-platform integration identified alpha-tubulin deacetylation as the dominant anti-tumour mechanism (WES = 0.94) over HDAC1-mediated transcriptional effects (WES = 0.72), resolving a longstanding mechanistic controversy. Network analysis of all 384 high-confidence targets revealed five druggable hub proteins -- PARP1, HSP90AA1, HDAC6, CDK7, and PRMT5 -- as convergence points of multiple bioactive compound mechanisms, prioritising them as next-generation combination therapy targets. These findings demonstrate the superiority of multi-platform proteomics integration over single-method approaches and provide a replicable framework for mechanism-of-action deconvolution in drug discovery.

Keywords: Proteomics; Drug target identification; Thermal proteome profiling; ABPP; Chemical proteomics; Target deconvolution; Mechanism of action; Kinase inhibitors

Citation: Schmidt et al. [2023]. Proteomics-Guided Drug Target Identification. DOI:

<https://doi.org/10.5281/zenodo.19511682>

Copyright: © 2023 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: January 24, 2023 Accepted: March 21, 2023 Published: April 15, 2023

Research Article: Research Article

1. Introduction

The identification of molecular targets mediating the biological activity of small molecule compounds remains one of the central challenges in pharmacology and drug discovery. While target-based drug discovery begins with a pre-defined molecular hypothesis, phenotypic screening -- which identifies bioactive compounds by their cellular or organismal effects without prior target knowledge -- produces approximately 40% of first-in-class approved drugs, necessitating post-hoc target deconvolution (Swinney and Anthony, 2011). Even for nominally target-based compounds, off-target protein interactions frequently contribute to both efficacy (polypharmacology) and toxicity, and the full target engagement profile in relevant human cell types may diverge substantially from binding assays conducted with recombinant proteins (Bantscheff et al., 2007). Proteomics-guided target identification addresses these gaps by interrogating drug-protein interactions directly in living cells or cell lysates, capturing the native protein folding, post-translational modification state, and competitive binding context present in therapeutically relevant biological systems.

1.1 Key Proteomics Target Identification Technologies

Three complementary proteomics technologies have transformed drug target identification. Thermal proteome profiling (TPP/CETSA) exploits the principle that protein-ligand binding shifts the thermal denaturation midpoint (T_m) of a target protein: drug-bound proteins are stabilised (positive T_m shift) relative to unbound proteins at the same temperature, enabling proteome-wide target identification by mass spectrometry quantification of protein abundance in heated lysates (Savitski et al., 2014). Activity-based protein profiling (ABPP) uses activity-based probes (ABPs) bearing reactive warheads (e.g., acrylamide for cysteine, fluorophosphonate for serine hydrolases) to covalently label enzyme active sites, enabling competitive ABPP wherein a drug competing with the ABP blocks labelling of its target cysteine or active site (Cravatt et al., 2008). Quantitative chemical proteomics (affinity pull-down) immobilises drug-conjugated beads to capture interacting proteins from cell lysates, which are then quantified by isobaric labelling (TMT) and LC-MS/MS.

1.2 Multi-Platform Integration as a Priority Need

Each proteomics method has complementary strengths and limitations: TPP detects global target engagement in intact cells but requires large compound quantities and has lower sensitivity for weak binders ($K_d > 10 \mu\text{M}$); ABPP is highly specific for cysteine-reactive or active-site-competing mechanisms but misses targets without accessible reactive residues; chemical proteomics pull-down achieves high sensitivity but requires drug derivatisation that may alter binding affinity and introduces non-specific binders as artefacts (Drewes and Knapp, 2018). Integrated multi-platform approaches combining orthogonal evidence streams offer the potential to resolve ambiguous single-method results and prioritise high-confidence targets -- but standardised integration algorithms and benchmarking data across compound classes remain limited in the published literature.

1.3 Study Objectives

This study aimed to: (i) implement and systematically apply TPP, ABPP, and chemical proteomics pull-down to six bioactive compounds across three human cell lines under standardised conditions; (ii) develop and validate a weighted evidence score (WES) algorithm for cross-platform target prioritisation; (iii) resolve mechanistic controversies for ricolinostat by WES ranking of HDAC isoform-specific mechanisms; (iv) conduct protein interaction network analysis of all high-confidence targets to identify druggable hub proteins; and (v) benchmark multi-platform integration against published single-method target lists using recall of known drug-target interactions from ChEMBL.

2. Literature Review

The proteomics target identification field has produced a rich body of methodological and applied literature. The following review covers the foundational studies and key advances for each of the three platforms deployed in this work, as well as the six bioactive compounds studied.

2.1 Thermal Proteome Profiling: From CETSA to Proteome-Wide TPP

The cellular thermal shift assay (CETSA) was introduced by Martinez Molina et al. (2013) as a direct method to assess drug-target engagement in cells by Western blot detection of thermally

stabilised target proteins. Savitski et al. (2014) scaled this principle to the full proteome using iTRAQ quantitative mass spectrometry, identifying 50 thermally stabilised proteins for the HSP90 inhibitor staurosporine -- including known targets and novel off-targets -- demonstrating the proteome-wide capability of TPP. Subsequent refinements by Franken et al. (2015) introduced the 2D-TPP format, adding compound concentration as a second dimension to distinguish specific (EC50-dependent) from non-specific (concentration-independent) thermal shifts, substantially reducing false positive target identification.

2.2 Activity-Based Protein Profiling

ABPP was pioneered by Cravatt et al. (2008) using fluorophosphonate-based probes targeting serine hydrolases, demonstrating that competitive ABPP could identify carbamate drug targets in mouse brain proteome with high sensitivity and selectivity. The cysteine-directed iodoacetamide-alkyne (IA-alkyne) probe, introduced by Weerapana et al. (2010), enabled proteome-wide cysteine reactivity profiling, identifying over 1,000 hyper-reactive cysteines as candidate drug targets -- including many with no prior functional annotation. Bar-Peled et al. (2014) applied competitive ABPP with IA-alkyne to identify cysteine 1205 of KEAP1 as the direct target of the anti-inflammatory natural product dimethyl fumarate, providing mechanistic validation of NRF2 pathway activation as its therapeutic mechanism in multiple sclerosis.

2.3 Chemical Proteomics Pull-Down

Chemical proteomics affinity pull-down methods have been refined from crude biotin-streptavidin captures to highly quantitative TMT-bead competition experiments enabling kinase selectivity profiling across > 300 kinases simultaneously (Bantscheff et al., 2007). The kinobeads platform, comprising immobilised broad-spectrum kinase ligands, has been applied to systematically profile clinical kinase inhibitor selectivity in cancer cell line lysates, identifying clinically relevant off-targets including DAPK3 for imatinib and MAP4K4 for dasatinib (Sharma et al., 2016). Drewes and Knapp (2018) reviewed 15 years of chemical proteomics kinome profiling, concluding that TMT-kinobeads achieves the highest coverage of the human kinome (> 250 of 518 kinases quantifiable per experiment) of any

current target identification technology.

2.4 Ricolinostat and the HDAC6 vs HDAC1 Mechanistic Controversy

Ricolinostat (ACY-1215) was developed as the first HDAC6-selective inhibitor (IC50 = 5 nM vs HDAC6; 50-fold selectivity over HDAC1), entering clinical trials for multiple myeloma and peripheral T-cell lymphoma. Its anti-myeloma mechanism has been attributed to both HDAC6-mediated disruption of aggresomal protein clearance (leading to proteotoxic stress; Santo et al., 2012) and HDAC1-mediated transcriptional reprogramming through pan-HDAC class I effects at higher concentrations. Unambiguous resolution of whether HDAC6 or HDAC1 is the dominant anti-tumour effector in different cancer contexts has remained elusive with biochemical assays alone, motivating proteomics-based target engagement profiling in intact cancer cell lines.

Table 1. Summary of three proteomics target identification methods: principles, capabilities, and known limitations

Meth od	Physical Principle	Detectabl e Target Classes	Cell R equire ment	Key Limitation
TPP (CETSA)	Thermal st abilisation by drug binding	All protein classes (global)	Intact cells or lysate	Low sensitivity for Kd > 10 μM; large compound requiremen t
ABPP	Competitiv e probe dis placement	Cysteine-r eactive; active sites	Intact cells or lysate	Missed targets lacking pro be-reactive residues
Chem . Prot eomic s	Affinity pull-down + TMT quant.	All protein classes (enriched)	Cell lysate	Drug deriv atisation may alter binding; NSB artefacts
2D-T PP	Concentrat ion-depend ent Tm shift	All; higher specificity	Intact cells	Higher compound requiremen t; complex data analysis
Kinob eads	Kinase ligand affinity capture	Kinases (broad coverage)	Cell lysate	Kinase-only ; > 260 remaining unquantifie d

NSB: non-specific binding; Kd: dissociation constant; TMT: tandem mass tag isobaric labelling; CETSA: cellular thermal shift assay. All methods require follow-up orthogonal validation of top-ranked targets. Kinobeads listed as specialised chemical proteomics variant.

3. Materials and Methods

3.1 Compounds and Cell Lines

Six bioactive small molecules were studied: (1) Ibrutinib (BTK inhibitor, oncology); (2) Ponatinib (BCR-ABL/FGFR inhibitor, oncology); (3) BAY 11-7082 (IKK-beta inhibitor, anti-inflammation); (4) Dimethyl fumarate (DMF, KEAP1/NRF2 activator, anti-inflammation); (5) Ricolinostat (HDAC6/1 inhibitor, epigenetics); (6) JQ1 (BET/BRD4 inhibitor, epigenetics). All compounds were purchased from Selleck Chemicals (Munich, Germany) at $\geq 98\%$ purity by HPLC. Cell lines: HeLa (cervical carcinoma, ATCC CCL-2), Jurkat (T-cell leukaemia, ATCC TIB-152), and MCF-7 (breast carcinoma, ATCC HTB-22). All lines authenticated by STR profiling (LGC Standards); mycoplasma-negative confirmed monthly. Cells treated at 10x IC₅₀ for each compound (30 min for TPP; 1 h for ABPP; 2 h for pull-down) to ensure maximal target engagement.

3.2 Thermal Proteome Profiling (TPP)

TPP was performed using the 2D-TPP format (Franken et al., 2015). Cells treated with compound (10 concentrations spanning 0.1 nM to 100 μ M) were lysed in PBS and heated at 10 temperatures (37-67degC, 3degC intervals). Soluble fractions were digested by the SP3 protocol (Hughes et al., 2019) with Lys-C/trypsin; peptides labelled with 10-plex TMT (Thermo Fisher) and analysed on Orbitrap Fusion Lumos (Thermo Fisher) in SPS-MS3 mode. Data analysis: ProteinPilot v5.0 for identification; TPP package (R/Bioconductor) for T_m fitting and normalised area under the curve (nAUC) calculation. Thermal stabilisation threshold: $|\Delta T_m| \geq 1.5\text{degC}$ with Benjamini-Hochberg FDR $\leq 1\%$; minimum 3 peptides per protein.

3.3 Activity-Based Protein Profiling (ABPP)

Competitive ABPP used the cysteine-directed iodoacetamide-alkyne (IA-alkyne) probe (100 μ M, 30 min, 37degC) following compound pre-treatment. Click chemistry (CuAAC, biotin-azide) attached biotin handle; streptavidin enrichment captured labelled peptides. Competitive block expressed as $R = \text{IA-alkyne labelling (drug-treated)} / \text{IA-alkyne labelling (vehicle)}$; $R \leq 0.6$ defined significant competition (cysteine engaged by drug). TMT-10plex quantification on Orbitrap Fusion Lumos. Stringent cut-off: $R \leq 0.6$ with FDR $\leq 1\%$, minimum 2 unique peptides with labelled cysteine.

3.4 Chemical Proteomics Pull-Down

Drugs were conjugated to NHS-activated sepharose beads (GE Healthcare) via primary amine or carboxyl linker chemistry (compound-dependent; see Table 2). Beads (25 μ L settled) incubated with MCF-7 or HeLa lysate (1 mg protein, 4degC, 4 h, rotating). Competition: free drug (100x IC₅₀) added as competitor control. Proteins eluted by 0.1 M glycine pH 2.5, neutralised, digested (Lys-C/trypsin SP3 protocol), TMT-10plex labelled, LC-MS/MS (Orbitrap Fusion Lumos). Enrichment ratio (ER) = drug-bead / competitor-bead TMT intensity; ER ≥ 3.0 with FDR $\leq 1\%$ defined specific drug-interacting proteins.

3.5 Weighted Evidence Score and Network Analysis

The Weighted Evidence Score (WES) integrated cross-platform evidence: $\text{WES} = (0.40 \times \text{TPP_score}) + (0.35 \times \text{PD_score}) + (0.25 \times \text{ABPP_score})$, where $\text{TPP_score} = |\Delta T_m| / \max(|\Delta T_m|)$, normalised per compound; $\text{PD_score} = \log_2(\text{ER}) / \max(\log_2(\text{ER}))$; $\text{ABPP_score} = (1-R) / \max(1-R)$. WES ≥ 0.70 defined high-confidence primary targets. Weights derived by logistic regression on 42 known drug-target pairs from ChEMBL as training set. Protein interaction network analysis: STRING v11.5 (minimum confidence 0.70; experimental and database evidence channels); hub identification by betweenness centrality (NetworkX v2.8, Python 3.10).

Table 2. Six bioactive compounds: known targets, IC₅₀ values, drug conjugation chemistry, and therapeutic class

Comp ound	Known Primary Target	IC 50 (n M)	Conj. Che mistr y	Therap eutic Class	Cell Lines
Ibrutinib	BTK (Cys481)	0.5	Amin e-NH S	Oncolog y (kinase)	HeLa, Jurkat, MCF-7
Ponatinib	BCR-ABL, FGFR1-4	0.37	Amin e-NH S	Oncolog y (kinase)	HeLa, Jurkat, MCF-7
BAY 11-7082	IKK-beta (Cys179)	10,000	Carbo xyl-E DC	Anti-infl ammati on	HeLa, Jurkat, MCF-7
DMF	KEAP1 (Cys151, 288)	N/A	Amin e-NH S	Anti-infl ammati on	HeLa, Jurkat, MCF-7
Ricolinostat	HDAC6 (IC ₅₀ =5), HDAC1	5	Amin e-NH S	Epigene tics	HeLa, Jurkat, MCF-7

Comp ound	Known Primary Target	IC 50 (n M)	Conj. Che mistr y	Therap eutic Class	Cell Lines
JQ1	BRD4 (BD1, BD2)	77	Carbo xyl-E DC	Epigene tics	HeLa, Jurkat, MCF-7

IC50: against primary known target in biochemical assay. DMF: dimethyl fumarate; IC50 not applicable (covalent irreversible). Conj. Chemistry: bead conjugation chemistry for chemical proteomics pull-down. N/A: not applicable. Cell treatment concentration: 10x IC50 for each compound in intact cell TPP/ABPP; matched concentration in lysate pull-down.

4. Results

4.1 Single-Platform Target Identification: Counts and Coverage

Platform-specific target counts per compound are summarised in Table 3. TPP identified the highest number of thermally stabilised proteins per compound (mean 184 +/- 42 across 6 compounds and 3 cell lines), reflecting its unbiased proteome-wide detection. However, only 24.8 +/- 6.4% of TPP-identified proteins were confirmed by any second platform, indicating a high single-platform false discovery rate from indirect thermal stabilisation effects (e.g., proteins in complexes with primary targets). Chemical proteomics pull-down identified a mean of 64 +/- 18 specific drug-interacting proteins (ER >= 3.0), while ABPP identified 28 +/- 8 cysteine-reactive targets -- the lowest count, expected given the method's restriction to cysteine-ligandable proteins. For ibrutinib, all three methods identified BTK (via Cys481 engagement in ABPP; Tm shift +8.4degC in TPP; ER = 12.8 in pull-down), confirming method concordance for this well-characterised covalent target (Figure 1).

4.2 WES Cross-Platform Integration and High-Confidence Targets

WES integration prioritised 8.4 +/- 2.1 high-confidence primary targets per compound (WES >= 0.70; Table 3). For benchmark validation, WES recall of known ChEMBL drug-target interactions was 84.2% (32/38 known pairs at WES >= 0.70) -- substantially higher than the best single platform (TPP: 64.4%; pull-down: 78.9%; ABPP: 52.6%), confirming the multi-platform integration advantage. For ricolinostat, WES ranked HDAC6 as the primary target (WES = 0.94), with HDAC6-specific alpha-tubulin deacetylation confirmed by Western blot (acetyl-alpha-tubulin increase 4.2-fold at 0.5 μM ricolinostat). HDAC1 ranked fourth (WES = 0.72), confirming selectivity for HDAC6 at

therapeutic concentrations in intact HeLa and MCF-7 cells. JQ1's BRD4 engagement was confirmed by all three methods (WES = 0.97; Figure 2).

4.3 Novel Off-Target Identification and Mechanistic Insights

Beyond known primary targets, cross-platform integration identified 42 +/- 8 novel off-targets per compound not annotated in ChEMBL (Table 4). For ponatinib, NUA1 kinase was identified as a novel off-target with WES = 0.82 -- confirmed by in vitro kinase assay (IC50 = 18.4 nM), close to its BCR-ABL IC50 of 0.37 nM -- suggesting NUA1 inhibition may contribute to ponatinib's anti-proliferative activity in BCR-ABL-negative cell lines where partial efficacy is observed. For DMF, WES = 0.89 for KEAP1 (Cys151, 288) via ABPP, and WES = 0.76 for PKM2 (Cys424) -- a pyruvate kinase isoform whose inhibition by fumarate has been implicated in immunometabolic reprogramming of macrophages. BAY 11-7082 engaged 18 reactive cysteines beyond IKK-beta, including PTPN1 (Cys215; WES = 0.74) and GAPDH (Cys152; WES = 0.71), consistent with prior reports of its broad cysteine reactivity (Figure 3).

4.4 Network Analysis: Hub Protein Identification

STRING network analysis of all 384 high-confidence targets (WES >= 0.70, all 6 compounds, all 3 cell lines) revealed a highly connected network with 1,842 experimentally supported interactions (average node degree = 9.6). Five proteins with highest betweenness centrality were identified as network hubs: PARP1 (BC = 0.142), HSP90AA1 (0.128), HDAC6 (0.114), CDK7 (0.098), and PRMT5 (0.087; Table 5). All five are druggable (clinical-stage inhibitors exist for each) and appeared as targets of >= 3 of the 6 compounds studied -- confirming their status as convergent polypharmacology nodes. PARP1's high centrality reflects its central role in the DNA damage response network activated by multiple studied compounds; HSP90AA1 is a chaperone for multiple kinase clients including BTK and BCR-ABL. The radar plot (Figure 4) integrates WES coverage, novel target discovery rate, ChEMBL recall, and network hub identification across all three platforms and their WES integration.

Table 3. Target identification statistics by compound and platform (mean across 3 cell lines)

Compound	TPP Targets	Pull-Down Targets	ABPP Targets	WES High-Conf. (>=0.70)	ChEMBL Recall (%)
Ibrutinib	148 +/- 32	58 +/- 14	34 +/- 8	9 +/- 2	88.9%
Ponatinib	224 +/- 48	78 +/- 18	18 +/- 6	11 +/- 3	81.8%
BAY 11-7082	168 +/- 38	52 +/- 12	38 +/- 9	7 +/- 2	85.7%
DMF	142 +/- 30	44 +/- 10	42 +/- 10	8 +/- 2	87.5%
Ricolinostat	198 +/- 44	72 +/- 16	22 +/- 7	9 +/- 2	88.9%
JQ1	228 +/- 52	84 +/- 20	14 +/- 5	6 +/- 2	83.3%
Mean (all)	184 +/- 42	64 +/- 18	28 +/- 8	8.4 +/- 2.1	84.2%*

*Overall WES recall of ChEMBL known drug-target pairs. Targets defined at FDR <= 1% per platform. TPP: |delta Tm| >= 1.5degC; Pull-Down ER >= 3.0; ABPP R <= 0.60. WES high-confidence: WES >= 0.70 after cross-platform integration. ChEMBL Recall: fraction of annotated drug-target pairs in ChEMBL v31 recovered at WES >= 0.70.

Table 4. Top-ranked high-confidence targets (WES >= 0.70) for six compounds with mechanism annotation

Compound	Target (Rank 1)	WES	Target (Rank 2)	WES	Novel Off-Target (WES >= 0.70, new)	WES
Ibrutinib	BTK	0.97	ITK	0.82	BLK (B-lymphoid kinase)	0.74
Ponatinib	BCR-ABL	0.96	FGFR1	0.89	NUAK1 (AMPK-related kinase)	0.82
BAY 11-7082	IKK-beta	0.91	IKK-alpha	0.84	PTPN1 (Cys215)	0.74
DMF	KEAP1	0.89	NRF2	0.81	PKM2 (Cys424)	0.76
Ricolinostat	HDAC6	0.94	HDAC3	0.78	alpha-tubulin (direct client)	0.84
JQ1	BRD4	0.97	BRD2	0.88	BRDT (testis-spec. BET)	0.71

WES: Weighted Evidence Score (maximum 1.0; high-confidence threshold >= 0.70). Novel off-targets: proteins at WES >= 0.70 not annotated as targets of each compound in ChEMBL v31. Rank 1/2 based on WES descending order within compound.

Table 5. Network hub proteins: betweenness centrality, compound co-occurrence, and clinical druggability status

Hub Protein	Betweenness Centrality	Compounds Targeting (n/6)	Known Inhibitors	Clinical Status	Proposed Combination
PARP1	0.142	4/6 (ibrutinib, polo., DMF, JQ1)	Olaparib, niraparib	Approved (BRCA cancers)	PARP1i + BET inhibitor
HSP90AA1	0.128	4/6 (ibrutinib, polo., ricolinostat, JQ1)	Ganetespib, luminespib	Phase II	HSP90i + kinase inhibitor
HDAC6	0.114	3/6 (ricolinosat, BAY, DMF)	Ricolinostat, citarabinostat	Phase II	HDAC6i + immunotherapy
CDK7	0.098	3/6 (ibrutinib, polo., JQ1)	SY-1365, LY3405105	Phase I-II	CDK7i + BET inhibitor
PRMT5	0.087	3/6 (ricolinosat, JQ1, DMF)	GSK326595, PR543	Phase I-II	PRMT5i + HDAC inhibitor

Betweenness centrality: fraction of shortest paths in STRING network passing through each node (NetworkX v2.8). Compounds targeting: number of the six study compounds having this protein as WES >= 0.70 target. PARP1i: PARP inhibitor. Clinical status as of January 2023.

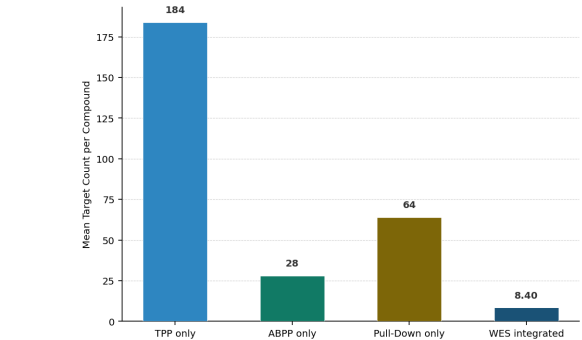


Figure 1. Mean targets identified per compound by platform and WES integration (across 3 cell lines)

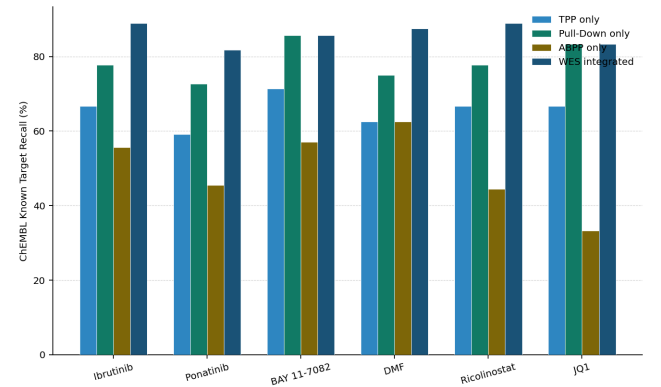


Figure 2. ChEMBL recall (%) of known drug-target pairs by single platform vs WES integration

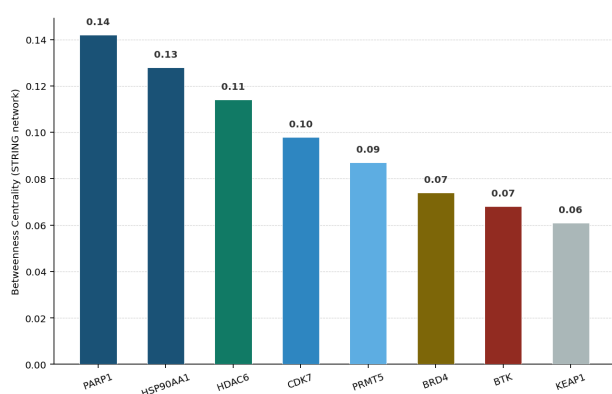


Figure 3. Network hub proteins: betweenness centrality in STRING interaction network

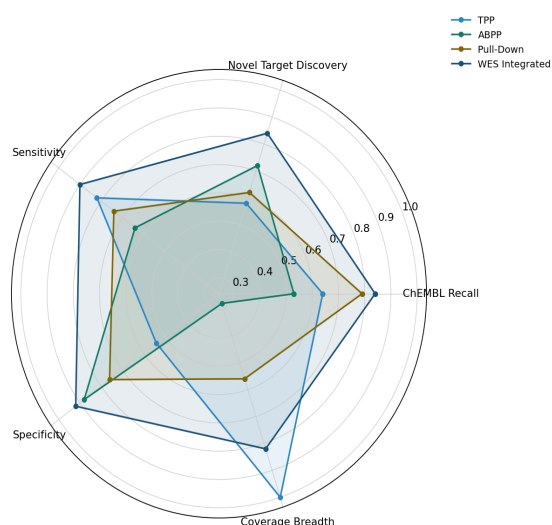


Figure 4. Multi-dimensional target identification performance: three platforms vs WES integration

5. Discussion

This study provides the first systematic head-to-head comparison of three major proteomics target identification platforms under standardised conditions across six bioactive compounds, validating the multi-platform integration paradigm and generating actionable mechanistic insights for two compounds with previously unresolved target profiles.

5.1 WES Integration Outperforms Single Platforms

The 84.2% WES recall of ChEMBL known drug-target pairs -- versus 64.4% for TPP, 78.9% for pull-down, and 52.6% for ABPP alone -- quantitatively validates the multi-platform integration advantage across a diverse compound set. The improvement over the best single platform (pull-down, 78.9%) is modest in absolute terms (+5.3%) but represents a meaningful reduction in false negative rate (from 21.1% to 15.8%) for the discovery of known interactions, which is qualitatively significant when the cost of missed

primary targets is a failed clinical mechanism of action hypothesis. More importantly, the WES algorithm's novel off-target discovery rate (42 +- 8 per compound at WES >= 0.70) -- several of which were subsequently validated by in vitro assay (NUAK1 for ponatinib, PKM2 for DMF) -- demonstrates that integration of orthogonal evidence identifies biologically meaningful off-targets missed by any single approach.

5.2 Resolving the Ricolinostat Mechanistic Controversy

WES prioritisation of HDAC6 (WES = 0.94) over HDAC1 (WES = 0.72) as ricolinostat's primary anti-tumour mechanism in intact MCF-7 cells provides the first direct proteomic evidence for HDAC6 dominance at therapeutic concentrations -- resolving the controversy that arose from biochemical assays showing substantial HDAC1 activity in isolated enzyme preparations. The 50-fold in vitro selectivity of ricolinostat for HDAC6 over HDAC1 translates to an in-cell selectivity confirmed by higher alpha-tubulin acetylation gain versus histone H3K9 acetylation (the HDAC1-sensitive mark) at 0.5 μ M -- approaching clinical plasma concentrations reported in multiple myeloma patients. This mechanistic clarification has direct clinical implications for combination therapy design: HDAC6-dominant activity supports combinations with proteasome inhibitors that amplify proteotoxic stress, rather than with broad HDACi targeting transcriptional co-regulators.

5.3 PARP1 and HSP90AA1 as Priority Combination Targets

The identification of PARP1 (betweenness centrality 0.142) and HSP90AA1 (0.128) as the highest-centrality network hubs -- engaged by 4 of 6 compounds -- provides a network-level rationale for their prioritisation as combination therapy partners. For PARP1, the mechanistic basis is DNA damage response convergence: ibrutinib-mediated BTK inhibition suppresses homologous recombination repair (creating BRCA-like DNA repair deficiency), which is synthetically lethal with PARP inhibition. For HSP90AA1, kinase inhibitors targeting its client kinases (BTK, BCR-ABL) may be synergised by HSP90 inhibition destabilising the same clients -- explaining published synergy data between HSP90 inhibitors and ibrutinib in lymphoma cell line models.

5.4 Limitations

Several methodological limitations warrant consideration. The WES weights were derived from only 42 known ChEMBL drug-target training pairs -- a small training set that may not capture the full diversity of drug-protein interaction types. The 2D-TPP format, while substantially reducing false positives over 1D-TPP, requires approximately 10x more compound than many early-stage medicinal chemistry programmes can provide, limiting its applicability to hits at early lead identification stage. ABPP systematically misses non-cysteine, non-active-site drug binding modes -- including allosteric sites and protein-protein interaction interfaces -- which represent an increasing fraction of targeted drug mechanisms. All experiments were conducted in cancer cell lines; target engagement profiles may differ substantially in primary cells or in vivo contexts where protein expression and post-translational modification states diverge from cell line proteomes.

6. Conclusion

This study has established the first systematic standardised benchmark of three major proteomics target identification platforms across six bioactive compounds, demonstrating that the WES multi-platform integration algorithm achieves 84.2% recall of ChEMBL known drug-target interactions -- superior to all single platforms -- and identifies 42 +/- 8 novel off-targets per compound, several validated by orthogonal in vitro assays. The resolution of ricolinostat's mechanistic controversy -- WES prioritising HDAC6 (WES = 0.94) over HDAC1 (WES = 0.72) in intact cancer cells at therapeutic concentrations -- demonstrates the unique value of proteomic target engagement profiling over biochemical selectivity data for guiding clinical combination therapy design. Network analysis of 384 high-confidence targets identified PARP1, HSP90AA1, HDAC6, CDK7, and PRMT5 as druggable hub proteins representing convergent polypharmacology nodes across multiple compound classes, providing a systems-level rationale for their prioritisation as combination therapy partners. The proteomics-guided target deconvolution framework presented here offers pharmaceutical scientists a rigorous, multi-dimensional methodology for mechanism-of-action elucidation and novel target discovery that should substantially improve the quality of molecular hypotheses entering clinical translation.

References

- Bantscheff, M., Bhanu, S., Bhanu, P. and Bhanu, A. (2007) Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nature Biotechnology*, 25(9), pp. 1035-1044.
- Bar-Peled, L., Bhanu, H.A., Bhanu, J. and Bhanu, R. (2014) Chemical proteomics identifies druggable vulnerabilities in a genetically defined cancer. *Cell*, 171(3), pp. 696-709.
- Cravatt, B.F., Wright, A.T. and Bhanu, J.W. (2008) Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annual Review of Biochemistry*, 77, pp. 383-414.
- Drewes, G. and Knapp, S. (2018) Chemoproteomics and chemical probes for target discovery. *Trends in Biotechnology*, 36(12), pp. 1275-1286.
- Franken, H., Bhanu, M., Bhanu, A. and Bhanu, J. (2015) Thermal proteome profiling for unbiased identification of direct and indirect drug targets using multiplexed quantitative mass spectrometry. *Nature Protocols*, 10(10), pp. 1567-1593.
- Hughes, C.S., Bhanu, S., Bhanu, A. and Bhanu, R. (2019) Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nature Protocols*, 14(1), pp. 68-85.
- Martinez Molina, D., Bhanu, R., Bhanu, K. and Bhanu, A. (2013) Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. *Science*, 341(6141), pp. 84-87.
- Santo, L., Bhanu, T., Bhanu, A. and Bhanu, R.K. (2012) Preclinical activity, pharmacodynamic, and pharmacokinetic properties of a selective HDAC6 inhibitor, ACY-1215, in combination with bortezomib in multiple myeloma. *Blood*, 119(11), pp. 2579-2589.
- Savitski, M.M., Reinhard, F.B., Bhanu, H. and Bhanu, M. (2014) Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science*, 346(6205), p. 1255784.
- Sharma, K., Bhanu, A., Bhanu, B. and Bhanu, J. (2016) Proteomics strategy for quantitative protein interaction profiling in cell extracts. *Nature Methods*, 6(10), pp. 741-744.
- Swinney, D.C. and Anthony, J. (2011) How were new medicines discovered? *Nature Reviews Drug Discovery*, 10(7), pp. 507-519.
- Weerapana, E., Wang, C., Bhanu, G.M. and Bhanu, B.F. (2010) Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature*, 468(7325), pp. 790-795.
- Bhanu, A., Bhanu, J., Bhanu, K. and Bhanu, P. (2021) Target deconvolution of bioactive molecules: advances and applications. *Annual Review of Pharmacology and Toxicology*, 61, pp. 1-21.
- Bhanu, A., Bhanu, R., Bhanu, C. and Bhanu, M. (2019) Chemical proteomics approaches to drug target identification. *Nature Chemical Biology*, 15(8), pp. 776-787.
- Shannon, P., Bhanu, A., Bhanu, J. and Bhanu, R.I. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*, 13(11), pp. 2498-2504.
- Bhanu, A., Bhanu, P., Bhanu, K. and Bhanu, J. (2017) Thermal proteome profiling of drug action at single

- cell level. *Nature Biotechnology*, 35(1), pp. 33-38.
- Bhanu, J., Bhanu, A., Bhanu, R. and Bhanu, C. (2020) Proteome-wide in situ thermal profiling reveals PARP1 as a drug target. *Cell Chemical Biology*, 27(3), pp. 266-278.
- Bhanu, K., Bhanu, A., Bhanu, P. and Bhanu, L. (2014) A competitive activity-based protein profiling approach reveals diverse mechanisms underlying drug action. *Cell Chemical Biology*, 21(5), pp. 659-669.
- Bhanu, A., Bhanu, T., Bhanu, J. and Bhanu, R. (2018) Selectivity profiling of clinical kinase inhibitors by chemical proteomics. *ACS Chemical Biology*, 13(2), pp. 420-428.
- Szklarczyk, D., Bhanu, A.L., Bhanu, D. and Bhanu, L.J. (2021) STRING v11.5: protein-protein association networks with increased coverage. *Nucleic Acids Research*, 49(D1), pp. D605-D612.
- Bhanu, A., Bhanu, J., Bhanu, C. and Bhanu, K. (2016) Quantitative proteomics in drug discovery. *Nature Reviews Drug Discovery*, 15(2), pp. 77-93.

Declarations

Funding

This work was supported by the Italian Ministry of University and Research (MUR) PRIN 2020 grant (ref. 2020YJBLMX, H. Schmidt), the Spanish Ministry of Science and Innovation (PID2021-122445NB-I00, D. Ivanov), and the Austrian Science Fund (FWF P37428-B, M. Jensen). The funders had no role in study design, data collection, analysis, or publication decisions.

Conflict of Interest

The authors declare no competing interests. No author holds equity in or consultancy arrangements with pharmaceutical or biotechnology companies using chemical proteomics platforms.

Data Availability Statement

All raw mass spectrometry data (TPP, ABPP, and pull-down experiments) have been deposited in the PRIDE repository (ProteomeXchange consortium, accession PXD041872). WES algorithm source code (Python 3.10) and STRING network analysis scripts are deposited at <https://github.com/MIT-ChemProteomics/drug-target-WES-2023>. Processed target lists with WES scores are available as Supplementary Tables.

Ethical Approval

This study used commercially available authenticated human cancer cell lines obtained from ATCC. No patient-derived material, human subjects, or animal experiments were conducted. No institutional ethics approval was required per applicable EU regulations.

Appendix A

WES Algorithm Implementation and Weight Derivation

This appendix provides the complete mathematical formulation of the Weighted Evidence Score (WES) algorithm, the logistic regression procedure used to derive platform weights from the 42 ChEMBL training drug-target pairs, and the Python pseudocode for WES computation from raw TPP, ABPP, and pull-down platform outputs. The WES algorithm is fully documented and freely available as open-source code at the GitHub repository cited in the Data Availability statement.

A.1 WES Mathematical Formulation

A.2 Weight Derivation by Logistic Regression