An integrative "omics" approach, for identification of bona fides PLK1 associated biomarker in Oesophgeal adenocarcinoma

An integrative "omics" approach, for identification of bona fides PLK1 associated biomarker in Oesophgeal adenocarcinoma

Nousheen Bibi[1, 2 *], Sajid Rashid[2], Judith Nicholson[3], Mark Malloy[4], Rob O'Neill[5], David Blake[6] andTed Hupp[5]

[1]*Department of Bioinformatics, Shaheed Benazir Bhutto Women University, Peshawar, Pakistan*

[2]*National Center for Bioinformatics, Quaid-i-Azam University, Islamabad, Pakistan*

[3] *Department of Oncology, University of Oxford, United Kingdom*

[4]*Australian Proteome Analysis Facility, Macquarie University, Sydney, New South Wales 2109, Australia*

[5]*Edinburgh Cancer Research Center, University of Edinburgh, United Kingdom*

[6]*Cyclacel Ltd., Dundee, United Kingdom*

Corresponding author: Nousheen Bibi*

Mailing Address: *Nousheen Bibi, *Departments of Bioinformatics, Shaheed Benazir Bhutto Women University Peshawar, Pakistan*

***Tel**: +92-346-7868789*

e-mail: nousheenimran99@yahoo.com, biogomal@gmail.com

Abstract

Background: The rapid expansion of genome wide profiling techniques offers the opportunity to utilize various types of information collectively in the study of human health and disease. Overexpression of Polo like kinase 1 (PLK1) is associated with oesophageal adenocarcinoma (OAC), however biological functions and molecular targets of PLK1 in OAC are still unknown.

Objectives: Here we performed integrative analysis of two "omics" data sources to reveal high level interactions of PLK1 associated with OAC.

Methods: Initially, quantitative gene expression (RPKM) was measured from transcriptomics data set of four OAC patients. In parallel alteration in phosphorylation levels was evaluated in the proteomics data set (mass spectrometry) in OAC cell line (PLK1 inhibited). Next two "omics" data sets were integrated and through comprehensive analysis possible true PLK1 targets that may serve as OAC biomarkers were assembled.

Results: Through experimental validation small ubiquitin-related modifier 1 (SUMO1) and heat shock protein beta-1 (HSPB1) were identified as novel phosphorylation targets of PLK1. Consequently in vivo, in situ and *in silico* experiments clearly demonstrated the interaction of PLK1 with putative novel targets (SUMO1 and HSPB1).

Conclusion: Identification of a PLK1 dependent biosignature in OAC with high confidence in two omics levels proven the robustness and efficacy of our integrative approach.

Key words: Omics, PLK1, SUMO1, HSPB1, kinase assay, ELISA, co-IP, PLA, *insilico.*

Introduction

Polo-like kinase 1 (PLK1) is a profound member of a serine-threonine kinase family with a kinase domain at the N-terminus and two polo-box domains involved in phosphopeptide-binding at the C-terminus [1, 2]. The expression level of PLK1 changes during cell-cycle progression with a peak level at M phase. PLK1 is an imperative regulator of mitosis through its diverse cellular localization and its protein-binding and phosphorylation activity to several different targets [3-5]. Moreover, overexpression of PLK1 is associated with oncogenesis. Overexpression of PLK1 overrides mitotic checkpoints, and leads to immature cell division without proper chromosome alignment and segregation, resulting in chromosomal instability and aneuploidy, a hallmark of cancer [6]. Indeed, consistent with the role of PLK1 in proliferation, overexpression of PLK1 has been observed in various types of cancers, such as colon [7], breast [8], stomach [9], pancreas [10], head and neck [11], and ovarian cancers [12]. PLK1 is overexpressed in oesophageal cancer (EC) relative to normal organs and exhibited higher promoter activity in EC cells than in normal epithelial cells. Moreover overexpression of PLK1 is associated with apoptosis resistance and proliferation in oesophgeal cancer cells in vitro [13, 14].

Oesophageal cancer is ranked as the sixth most common cause of cancer death with poor prognosis and aggressive behavior as it tends to recur after surgery. The majority of oesophageal cancers are of the adenocarcinoma histological sub-type [15]. Survival of EC patients remains poor, due to late stage presentation of the disease, being approximately 10% for squamous cell carcinoma and 20% for adenocarcinoma [16, 17]. Therefore, to improve the survival of patients with this unmanageable cancer targeting of genes associated with progression of this cancer is being investigated. Recent attempts at identifying new prognostic markers for Oesophageal adenocarcinoma (OAC) have focused on using microarray analysis of mRNA expression patterns, and have led to the identification of two [18] and four [19] gene signatures which are of prognostic value. Thus, further analysis of gene expression

signatures, or the coordinated up regulation of components of molecular pathways might provide further advances in this area. Additionally proteins can be regulated on the proteome level both post-transcriptionally and post-translationally, meaning proteomic analysis can add significant value to biomarker studies in OAC, especially considering many biomarker end-points are looking at protein levels rather than transcript levels, for example immunohistochemistry based methods.

Cancer researchers have made considerable progress in identifying new cancer markers through genome-wide profiling technologies. Comprehensively identifying gene expression on both transcriptomic and proteomic levels in one tissue is a prerequisite for a deeper understanding of its biological functions and to characterize disease associated changes at various levels of genome function. Genomics, transcriptomics and proteomics data provide basically one snapshot from one angle that characterizes different aspects of cancer relevant genome regulation and function [20, 21]. Integrative transcriptomics and proteomics measurements characterize key players and biological processes, a fundamental step in the mechanistic characterization of disease and revealing promising molecular targets for therapeutic intervention of the disease [22, 23]. However, the lack of algorithmic implementations forms a bottleneck hampering integrative approaches.

Here we demonstrated an integrative multi-omics approach to identify PLK1 regulated biomarkers in oesophgeal cancer. RNA-sequencing data from four oesophgeal patients and Mass spectrometry data from PLK1 inhibited oesophageal cell lines were integrated and through extensive downstream bioinformatics analysis promising hits were identified. Putative novel hits were experimentally validated through classical and robust molecular biology and biochemistry techniques. Hence in an effort to identify promising PLK1 targets in oesophageal adenocarcinoma we by-pass the bottleneck of integrative approach by integrating RNA-sequencing and mass spectrometry data set.

Materials and Methods

Identification of PLK1 dependent cancer biomarker through integrative strategy

For uncovering novel PLK1 regulated cancer biomarkers multiple biological information sources were assimilated and functioning premise was measured. We collected RNA sequencing data from four OAC patients (Patient ID; 23T, 26T, 27T and 45T) and similarly Mass Spectrometry (MS) data from PLK1 inhibited oesophageal cell line (OE33). RNA sequencing data was processed and analyzed in CLC genomics and cancer workbench (http://www.clcbio.com/products/clc-genomics-workbench/) resulting in upregulated and downregulated clusters of 6080 and 2810 genes respectively.

Tryptic Digest and Phosphoenrichment

Three replicates of each condition, untreated, arrested, Cyc1 treated and arrested with Cyc1 treatment were digested with trypsin then enriched for phosphopeptides using titanium dioxide. For each sample 500 µg of cell lysate (Sodium deoxycholate lysis buffer) was reduced with 10 mM DTT at 50°C then alkylated in the dark at room temperature with 25 mM iodoacetamide before digesting overnight with Trypsin:Protein 1µg:20µg (Promega) at 37°C. After digestion lysates were acidified to a final concentration of 1% TFA, 80mg/ml glycolic acid, 5% acetonitrile to precipitate deoxycholate and undigested proteins and centrifuged at maximum speed for 2 min. Supernatant was transferred to a lo-bind eppendorf tube. Titanium dioxide beads (10µm, Titansphere, GL Sciences) were washed once with 5% acetonitrile, 1% TFA, 80mg/ml glycolic acid then added to sample (~5mg beads per 500µg protein). Sample was incubated with titanium dioxide beads with shaking for 30min-45mins. Using the batch method beads were centrifuged briefly and supernatant collected as non-phospho fraction. Beads were then washed three times with 200 µl 70% acetonitrile, 80mg/ml glycolic acid, 0.1% TFA, three times

with 80% acetonitrile, 0.1% TFA and phosphopeptides eluted with 80 µl 4% ammonia. Samples were acidified with 2µl 10% TFA, dried in a vacuum centrifuge and stored until LC-MS/MS analysis. Samples were resuspended in MS loading buffer (2% acetonitrile, 0.1% formic acid) and 10 µl injected to a C18 column and eluted over a 180min gradient (Eskigent NanoLC) to a 5600 TripleTOF (ABSciex) and spectra collected via a data dependent acquisition method.

Data analysis

MS data generated was processed with ProteinPilot v4.2 (ABSciex) to generate MGF peaklists which were used to search Mascot (Swissprot 2012 Homo sapiens database, 20232 entries). Mascot .DAT files were subsequently processed in Scaffold and Scaffold PTM (Proteome Software). Results were filtered for peptide identifications with a phosphorylation modification and identification confidence >95% to obtain a false discovery rate of >1%. Proteins were then assessed in Scaffold PTM generating a list of confidently (Ascore >95%) assigned phosphosites for each biological category (DMSO, Cyc1, Arrested, Arrested and Cyc1). To filter for peptides with changes in phosphorylation after drug phosphopeptides were only used which were identified with a confidence of 95%, had an Ascore value of >95% for phosphosite assignment and a spectral count >3 in either biological category.

Plasmids and constructs

Gateway cloning technology was used to make construct for list of selected candidates as follow: His-Tagged full length constructs of human SUMO1 (301 bp), HSPB1 (601 bp), CDK1 (841 bp), SRRM1 (270 bp), DKC1 (1501 bp), CAV2 (481 bp), MCM2 (2701 bp), SRRP1 (2100 bp), RBM39 (1561 bp), SRSF6 (1021 bp), TRA2B (841 bp), EIF3C (2701 bp), EFHD2 (721 bp) and PLK1-PBD (330-603) were

prepared in bacterial expression gateway vector pDEST17 and pDEST14. N-terminal and C-terminal tagged mammalian expression constructs of SUMO1 and HSPB1 were also prepared in Gateway mammalian expression vector pDEST53 and pDEST47. The mutant of SUMO1 S2A/S2D/S2E and HSPB1 S82A/S82D/S82E were prepared using QuickChange® II XL Site-Directed Mutagenesis Kit from Agilent technology.

Purification of recombinant proteins

Different His tagged wild type and mutant substrates candidates were expressed in *E. coli* strain BCL21 DE3 and purified through Ni-NTA Agarose beads (#30230, QIAGEN technologies). First Ni-NTA agarose beads were washed with buffer I (20mM Tris-Hcl PH 8.0, 400mNaCl, 10mM Mgcl2*6H2O, 0.1% Np40, 5mM Imidazol) and then clear lysates were loaded on to the beads and incubated with shaking at 4°C for 1 hour. Hexa-Histidine tail in the His tag proteins bind to the immobilized nickel ions with high specificity and affinity. The protein suspension was run through column and five washes were performed with different molar concentrations (6M, 4M, 2M, 1M) of buffer I and buffer II (20mM Tris-Hcl PH 8.0, 400 mNaCl, 10mMMgcl2*6H2O, 0.1% Np40, 5mM Imidazol, 8M urea). After washing, His-tagged proteins were eluted in buffer III (20mM Tris-Hcl PH 8.0, 400mNaCl, 10mM Mgcl2*6H2O, 0.1% Np40, 150mM Imidazol). Imidazole in elution buffer (buffer III bind strongly to the Ni-NTA beads and eluted the proteins. The eluted proteins were dialyzed using Zeba™ Desalt Spin Columns (# 89891, Thermo Scientific) by standard protocol.

Antibodies

Following antibodies were used; anti PLK1 (ab 35-206 Abcam, ab Phospho T210 Abcam, ab 3G251 Abcam), anti SUMO1 (ab 21C7 invitrogen) , antiHSPB1 (ab G3.1 abcam) and anti-6X His tag® (ab HIS

H8 abcam). Different concentrations were used for western blotting, ELISA, co-immunoprecipitation and immunoflorescence assays.

Co-Immunoprecipitation

Cells from A375 cell line were grown until confluency was reached. Cells were harvested and lysed in the co-immunoprecipitation buffer (50 mM HEPES-NaOH pH 7.2, 150 mM NaCl, 0.5% (v/v) Tween-20, 1X protease inhibitor cocktail). Initially 1 mg of protein was precleared with protein G sepharose beads for 1 hour at 4°C. After quick centrifugation, supernatant was transferred to new tube and then incubated with appropriate antibody (2 µg/ sample) for 2 hours. The immune complexes were collected with protein A beads. Beads along with attached proteins were washed 3x with lysis buffer. Samples were fractionated by SDS-polyacrylamide gel electrophoresis for immunoblot analysis.

Proximaty ligation assay

Proximity Ligation Assay was performed using Duolink® in situe PLA kit (# DUO92101, Sigma) following the manufacturer's protocol. Duolink anti-rabbit plus probe, anti-mouse minus probe were used, while antibodies against SUMO1, HSPB1 and PLK1 were used from different species. In each experiment a negative control using only one antibody of each pair was included. For each antibody pair, exponentially growing HTC116 and OE33 cells were seeded on cover slips and fixed for immunostaining as described in manufacturer's protocol. After overnight incubation with primary antibody at 4°C, samples were processed for the PLA assay. Imaging was performed using Olympus BX51 fluorescent wide field microscope equipped with a 100-W mercury lamp. we collected images using software embed in microscope and translated them into red, blue and green channel using ImageJ software (http://imagej.nih.gov/ij/) and further processed and analyzed.

Direct protein binding ELlSA

ELISA 96-well plates (Beckman–Coulter) were coated with 20 to 50 ng of proteins 1 (His-tagged SUMO1 WT/S2A/S2D/S2E and His-tagged HSPB1 WT/S82A/S82/D/S82E)) made up to a final volume of 50 µL in coating buffer (0.1 M NaHCO3 pH 8.6). To block the unoccupied sites, wells were washed with PBST and then incubated with 200 µL of PBS plus 3% BSA for 1 hour. A titration of protein 2 (His tagged PLK1-PBD (330-603) was added to each well and incubated for 1hour at room temperature. To terminate the reaction, ELISA plates were washed 4 times with PBST. For detection of bound PLK1-PBD, plates were incubated for 2 hours with 50 µL per well of anti-PLK1 antibody at a concentration of 0.5 µg/ml. After washing the plates 5 times, 100 µL per well of HRP-conjugated secondary antibody (diluted 1:1,000 in blocking buffer) was added and plates were further incubated for 1 hour. After final washing, binding was measured by adding to each well 50 µL of a 1:1 mix of ECL solutions 1 (2.5 mM luminol, 0.4 mM p-coumaric acid, 0.1 M Tris pH 8.5) and ECL solution 2 (0.02% (v/v) H2O2, 0.1 M Tris pH 8.5). The luminescence produced was immediately detected with a illuminometer and analysed with Ascent software version 2.4.1.

In Vitro kinase assay

All *in vitro* kinase assays were performed in kinase buffer (10 mM HEPES pH 7.4, 4 mM MgCl2, 1 mM DTT, 2.5 mM EGTA, 10 mM β-glycerol phosphate) supplemented with 1 µL ATP mix (1/50 dilution of γ32P ATP 3,000 Ci/mmol 10 mCi/ml EasyTide Lead (#NEG502A100UC, Perkin Elmer) in 2 mM ATP (#A7699-5G, Sigma) in 25 mM HEPES pH 7.8) at 30°C for 30 min in the presence of 1.5 µg

of the indicated bacterially expressed His-tagged potential substrates proteins. Reactions were stopped by adding gel sample buffer and analyzed by SDS-PAGE and autoradiography.

Western blotting

After separation by SDS-PAGE, proteins were transferred to Hybond-C Extra 0.45 μm nitrocellulose membrane (#RPN203E, Amersham Biosciences, GE Healthcare) using transfer buffer (24 mM Tris, 191 mM glycine, 20% (v/v) methanol) at 300 mA for 90 min. Membranes were blocked in 1X PBS (made from tablets, #BR0014G , Fisher) -0.1% (v/v) Tween-20 (PBST) containing 5% (w/v) non-fat dried milk (Marvel) for 1 hour at room temperature then incubated overnight at 4°C or several hours at room temperature with primary antibody. Then incubated with appropriate horseradish peroxidise (HRP)-conjugated secondary antibody for 1 hour at room temperature, followed by washes with PBST. Proteins were finally detected by incubation with enhanced chemiluminescence (ECL) reagent: membranes were overlaid with ECL solution 1and ECL solution 2 mixed at a ratio of 1:1, for 1 min.

Molecular structural modeling and docking

Experimentally determined structure of PLK1 kinase domain, polo box domain (PBD) and SUMO1 were retrieved from protein data bank (http://www.rcsb.org/pdb/) with PDB IDs; 3THB, 4LK1 and 1a5r respectively. While structure of HSPB1 was predicted through homology modeling and ab initio method using SWISSMODEL [25] and Muster [26] servers.

Three dimensional model of HSPB1 and SUMO1 encompassing putative kinase phosphorylation and PBD docking motif were subjected to docking against experimental 3D structure of PLK1 kinase domain and PBD. AutoDock 4.0 [27], HADDOCK [28] flexible protein-protein docking server

SWARM dock [29] and GRAMXX [30] were used to cross over the biasness of docking algorithm and cross validation of docking conformation. Best docking cluster from each docking method was selected and thoroughly evaluated to monitor best docking poses and interactions of studied macromolecules using Chimera V1.7.1 [31] and DIMPLOT [32]

Results

Identification of novel PLK1 targets through Integrative omics approach

By focusing the power and need of integrative analysis we used an integrative strategy of transcriptomics-proteomics analysis (Figure 1) to investigate the novel PLK1 targets and biomarkers in OAC. Initially, we assessed alterations in each dataset and then compared the results from both data sources to grasp common biosignatures of OAC targeted by PLK1. RNA-sequencing data from four patients with OAC phenotype was thoroughly analysed through CLC genomics and CLC cancer workbench (http://www.clcbio.com/products/clc-genomics-workbench/). In the first step sequencing reads were mapped to reference genome and after expression level quantification of samples, genes were grouped based on their calculated RPKM (read per kilo meter) [33] values. For both RNA-sequencing data and microarray data, we first log transformed the preprocessed expression data, next we did quantile normalization and transformed the sequence of real numbers into frequency representation. Sub groups of differentially expressed genes with greater than two fold of change in expression were selected and clustered as upregulated (6850 genes) and downregulated (2810 genes).

PLK1 was observed upregulated in four patients with significant fold change. RNA sequencing reads mapped to PLK1 revealed higher expression of PLK1 in tumor sample in comparison with to their normal counterpart (Figure 2). Through *in vitro* studies elevated level of PLK1 in OAC has been suggested as a cause of cell proliferation and inhibition of apoptosis [13, 14].

As a robust strategy, we integrated the information from RNA sequencing or NGS and MS to identify PLK1 associated phosphorylation targets and possible OAC biomarkers. Consistent with the role of PLK1 in OAC and as attractive next generation antimitotic target, our aim was to identify novel molecular targets with related biological function of PLK1 in OAC

Through synchronization of two omics data sources, reasonable number of hits were extracted that were common in both datasets (Table 1) and subjected to downstream gene ontology (GO), interaction network and pathway analysis though multiple Bioinformatics tools and information sources.

All possible direct and indirect interaction networks and pathway links between PLK1 and identified hits were extensively investigated to expose novel PLK1 associated targets. Interestingly, apart from some novel candidates, we observed many previously known PLK1 partners (CDK1, MCM2, TPX2 and TOP2A) in the shortlisted putative targets of PLK1. These findings further strengthened our predictions and cross validated the effectiveness of used strategy.

In vitro PLK1 phosphorylation assay

As a consequence of reduced PLK1 activity, phosphorylation sites on direct PLK1 substrates are expected to be downregulated in PLK1 inhibited cells. To corroborate a direct role of PLK1 in phosphorylation of identified putative direct targets, we performed PLK1 *in vitro* phosphorylation assay on a set of 8 candidates, selected throughout the ranked list on the basis of their common cellular localization and biological process association with PLK1 and availability of full length bacterially expressed proteins. ATP alone was used as negative control. The kinase assay revealed that 2 out of 8 candidate substrates incorporated ^{32}P on incubation with PLK1. In order to quantify the results and compare the amount of ATP incorporation between substrates, we normalized autoradiography band signal to exclude differences due to time of film exposure and amount of protein, which suggested that

SUMO1 and HSPB1 could be phosphorylation targets of PLK1. Overall, our experimental validation assay showed that 2 out of 8 candidate proteins were PLK1 substrates *in vitro*, however accuracy with individual phosphorylation site remains to be assessed.

Since our *in vitro* assay showed that Ser2 of SUMO1 and Ser82 of HSPB1 were phosphorylated by PLK1, to further verify the phosphorylation sites we generated full length His-SUMO1 phosphomimic and phospho-dead mutants (Ser2A, Ser2D, Ser2E) and His-HSPB1 phosphomimic and phospho-dead mutants (Ser82A, Ser82D and Ser82E) and performed *in vitro* kinase assays with PLK1. There was a significant reduction in the phosphorylation signal of SUMO1 and HSPB1 mutants as compared to their wild type counterparts (Figure 3e). Hence these results confirmed that Ser2 of SUMO1 and Ser82 of HSPB1 are the true phosphorylation sites targeted by PLK1.

Protein-protein interaction experiments

To determine *in vivo* interaction of SUMO1 and HSPB1 with PLK1, we performed immunoprecipitation assay in A375 cell lines. In reverse immunoprecipitation experiments, we observed that endogenous SUMO1 and HSPB1 were co-precipitated with PLK1 (Figures 4 f and g), thereby suggesting *in vivo* interaction among them. We also performed reverse co-immunoprecipitation assay to further validate our *in vivo* interaction. Furthermore, we looked for possible protein-protein interactions of PLK1 with SUMO1 and HSPB1 to demonstrate their involvements in common signaling network. Therefore we performed an *in situ* proximity ligation assay (PLA) in HCT116 and OE33 cells. The immunofluorescence signal against PLK1-SUMO1 and PLK1-HSPB1 in two different PLA experiments showed well defined protein-protein interactions (Figures 4a and b). Images were collected and analysed in ImageJ software. We multiplied the readings in green and blue channels with inverse value of red channel. Similarly, we multiplied values in blue channel with values in green channel. Prominent

fluorescence spots detected through *in situ* PLA experiments revealed that SUMO1 and HSPB1 are novel interacting partners of PLK1.

In agreement to the previous studies which suggest that PLK1 stably interacts with its targets via its PBD, we performed direct binding ELISA to substantiate the direct protein-protein interaction between PLK1-PBD and identified putative hits (SUMO1 and HSPB1). Against a buffer control, SUMO1 and HSPB1 were shown to bind significantly to PLK1-PBD in a dose dependent manner (Figure 5.4d). ELISA test was also carried for phospho-mutant forms of SUMO1 (Ser2Ala, Ser2Glu and Ser2Asp) and HSPB1 (S82Ala, S82Glu and S82Asp) proteins in comparison to wild type proteins against PLK1-PBD. Results indicated binding of WT and mutant proteins (SUMO1 and HSPB1) with equal affinity to PLK1-PBD (Figures 4 d and e) hence suggesting that Ser2 of SUMO1 and Ser82 of HSPB1 were phosphorylation sites of PLK1, however the identified phosphorylation sites were not involved in binding to PBD of PLK1 as mutant and wild type proteins bound with equal affinity

Next we asked whether SUMO1 and HSPB1 carry PBD docking sites along with putative phosphorylation site identified in the present study. Interestingly, we found that SUMO1 and HSPB1 held previously identified PBD docking consensus **S[pS/pT]X** [1, 34], which is thought to be critical for binding with PBD prior to phosphorylation of putative site by kinase domain. Furthermore, to map the PBD binding motif and to characterize the binding of identified targets with PLK1, we performed detailed structural analysis through structural bioinformatics approaches.

Structural bioinformatics approach to investigate interaction of SUMO1 and HSPB1 with PLK1

Putative novel PLK1 targets (SUMO1 and HSPB1) were subjected to molecular modeling and docking assays to characterize their mode of interaction and site of interaction with both domains of PLK1. The 3D structure of small ubiquitin like modifier (SUMO1) was retrieved through protein data bank having PDB ID: 1a5r.A. Due to lack of experimentally determined full length HSPB1/HSP27 structures, we used homology modeling and *ab initio* structure prediction approach to determine the 3D structures for ongoing interaction mapping experiments. SWISSMODEL [25] and Muster [26] servers were used to predict HSPB1 three dimensional structure, followed by model refinement and geometry optimization through Wincoot [35] and UCSF chimera 1.7.1 [31] tools.

Ramachandran plot indicated that approximately 95% residues of predicted models lie in the allowed regions. Moreover, parameters like peptide bond planarity, non-bonded interactions, Ca tetrahedral distortion, main chain H-bond energy and overall G-factor for the modeled structures lie within favorable range.

PDB determined three dimensional structure of SUMO1 (PDB ID: 1a5r) and predicted 3D model of HSPB1 were docked against the experimentally known structure of PLK1 kinase domain (PDB ID: 3THB) and PBD (PDB ID: 4LK1) through multiple docking algorithm (AutoDock 4.0 [27], HADDOCK [28] SWARM dock [29] and GRAMXX [30]). Deep structural analysis was performed on set of most promising docked complexes (SUMO1-PLK1 and HSPB1-PLK1) selected from each methods to configure mode and site of bindings between studies macromolecules. Structure based interaction measurement and binding site characterization of SUMO1 with PLK1 revealed that N-terminal part of SUMO1 a loop region encompassing putative phosphorylation site (S2) was in fine docking pose to be phosphorylated by kinase domain. While the C-terminal part of SUMO1 encompassing PBD binding consensus S[pS/pT]X [1, 34] was observed in tight interaction with critical substrate binding residues of PBD (TRP414, ASN416, HIS489, PHE535, HIS538 and LYS540) [1] (Figure 5). Surprisingly the

docking pose adopted by SUMO1 and PLK1 (both domain) was consistent in all docking algorithms used in the present study.

Likewise, binding analysis of most consistent and significant docked complex of HSPB1 and PLK1 exposed important structural twist. N-terminus of HSPB1 carrying putative PLK1 phosphorylation site was observed to be pointed towards the kinase domain of PLK1, while HSPB1 region encompassing β3 and small loop (Pro106-Gly116) was observed having strong interactions with critical residues of PBD involved in substrate binding (Figure 6).

Discussion

The integration of transcriptomics (data from next generation sequencing) and proteomics (data from mass spectrometry) or multi-omics provides a comprehensive overview of the biological features that drive cancer than analysis of individual dataset alone. Such analysis helps in the identification of the most important targets for cancer detection and intervention.

In recent years, despite increased understanding of PLK1 pleiotropic functions [36-39], many of its identified substrates do not alone explain all its physiological functions. In line with this, we used an integrative omics approach to uncover the unexplored PLK1 targets and to extend our knowledge of PLK1 signaling. Our transcriptomics and proteomics profiling led to the identification of SUMO1 and HSPB1 as novel phosphorylation targets of PLK1. Integration of RNA sequencing data from four OAC patients (23T, 26T, 27T, 45T) with PLK1 dependent mass spectrometry data from OE33 cell line treated with PLK1 inhibitor allowed us to quickly eliminate a number of candidate genes which were significant in RNA sequencing data, but were not present in phosphoproteomic data (possible direct PLK1 targets).

Even though it is tough and challenging to integrate and amalgamate two different technologies due to different sensitivity and accuracy, multiple bioinformatics approaches may make it possible. Here through integration of information from two data sources our focus was to perform a high throughput analysis to have a good coverage and promising results that would not have been reached with individual data type. Post-translational modification data which is highly relevant when studying kinases was also only available via proteomics methods, as this information is not present in transcriptomic data. Combination of the two datasets leads to especially robust targets for future development.

Our integrative approach resulted in the identification of many proteins that have previously been associated with PLK1. It has been known that PLK1 co-immunoprecipitates with members of the minichromosome maintenance MCM2-7 protein complex, having important roles in DNA replication and contributes in initiation and elongation of replication [40]. In the present study, MCM2 was observed to be have a phosphosite downregulated with >2 fold reduction in phosphorylation signal upon PLK1 inhibition. Similarly, Van Vagurt in 2010 reported that mitotic phosphorylation feedback network connects Cdk1, PLK1, 53BP1 and Chk2 to inactivate the G(2)/M DNA damage checkpoint [41]. In our results, we identified reduction in Cdk1 phosphorylation level upon PLK1 inhibition apart from up-regulation in four OAC patients. TPX2 and TOP2A were previously known to interplay with PLK1 in cell cycle [42-43] and interestingly filtered out in our shortlisted candidates set. These results further strengthened the prediction accuracy and soundness of our strategy. In addition to the proteins discussed above that have previously been associated with PLK1, several novel PLK1 interacting proteins that share ontology with PLK1 were also identified in present study.

Small ubiquitin-like modifier (SUMO1) is found in all eukaryotes and becomes covalently conjugated to other cellular proteins [44-47]. SUMO, pathways play critical roles in several aspects of mitosis including cell cycle progression, chromosome structure, kinetochore function, cytokinesis and cell

division [48]. SUMOylation also controls transcription, DNA repair, DNA recombination and mitotic chromosome segregation [49-54]. Biological knowledge between PLK1 and SUMO1 suggests some common functionalities and dynamic localization of these key players throughout in mitosis. In current study, we identified phosphorylation dependent interaction of SUMO1 and PLK1. MS data revealed phosphorylation of SUMO1 by PLK1 at Ser2 which was validated through our *in vitro* PLK1 phosphorylation assay. Phosphorylation of Ser2 of SUMO1 was detected in yeast, drosophila and human through mass spectrophotometry, signifying evolutionary conserved function for this modification [55]. Binding of SUMO1 with PLK1 was further validated through protein-protein interaction mapping assays. Through our *in silico* deep structural analysis, we mapped docking sites of SUMO1 with kinase domain and PBD of PLK1 which clearly revealed interaction of SUMO1-PBD binding consensus with PBD (PLK1) and putative phosphorylation motif with kinase domain (PLK1). We are tempting to speculate that our findings open a room for cancer research through the functional interplay between PLK1, SUMO1 and tumor suppressor protein p53. The p53 is modified by SUMO1 at single phosphorylation site (Lys386) and acts as a potential regulator of p53 response [56]. Recently it has been reported that p53 physically interacts with PLK1 and is necessarily and specifically binds to PLK1 promoter and suppresses the endogenous level of PLK1 expression [57-58]. Once the intertwined relationship of PLK1, SUMO1 and p53 is established and evaluated, this could be a novel and interesting area for cancer therapeutic development. This is also a broader mechanism of regulation of SUMO, which is increasingly identified as a post-translational modification of important proteins such as DNA repair proteins. The post-translational modification of a post-translational modification leads to intriguing new levels of complexity in cellular signaling pathways.

In our study, we also identified HSPB1, a heat shock protein as a novel phosphorylation target of PLK1. HSPB1 is ubiquitously expressed in all human tissues [59-60] and increased expression level of HSPB1

was detected in many types of tumors [61-64], while its decreased expression level was measured in oesophageal adenocarcinoma and adrenal adenoma [63, 65]. Interestingly, HSPB1 was also downregulated in four oesophageal patients in present study. In response to growth factor and heat shock, HSPB1 is known to be phosphorylated *in vivo* at multiple sites like Ser15, Ser78 and Ser82. Clerk *et al* in 1998 identified MAPAK2 as major enzyme involved in phosphrylation HSPB1 at Ser82 [66]. Through our mass spec screen we observed 2.6 fold reductions in phosphorylation signal of HSPB1 upon PLK1 inhibition. Through extensive investigation by using *in vivo, in vitro, in situ* and *in silico* approaches we demonstrated that Ser82 of HSPB1 is phosphorylated by PLK1. It has been known that upon phosphorylation, HSPB1 forms different oligomers and phosphorylation affects the chaperon activity of HSPB1 [67-68]. HSPB1 is involved in signal transduction and prevent apoptosis by inhibiting capases [69]. Role of HSPB1 and PLK1 is needed to be investigated to have a clear picture of mechanism regulated by these two proteins. PLK1 has also been shown to phosphorylate HSp70 another heat shock protein [70]. HSP90 is also a heat shock protein known to regulate mitosis through its association with PLK1 [71]. Thus identification of HSPB1/HSP27 as novel PLK1 substrate suggests additional relation of PLK1 with family of heat shock proteins.

In summary our high throughput analysis using two omics information extends the knowledge of PLK1 interactions and novel signaling cascade in OAC. *In vitro* PLK1 phosphorylation assay revealed SUMO1 and HSPB1 as phosphorylation targets of PLK1. Consequently, the interaction of SUMO1 and HSPB1 with PLK1 was characterized by conducting comprehensive set of *in vivo, in situ* and *in silico* assays that clearly demonstrated the interaction of PLK1 with putative novel targets (SUMO1 and HSPB1). The clinical relevance of PLK1 dependent modification in OAC remains to be explored. Owning to the emerging role of PLK1 in OAC proliferation, identification of PLK1 dependent phosphorylation targets suggested that PLK1 is a promising molecular target for the treatment of OAC.

Author Contributions

Conceived and designed the experiments: NB, TH. Mass spectrometry experiment and data analysis JN, MM, RN, DB. Wrote paper: NB. Prepared figures: NB. Manuscript editing and review: TH, SR.

Acknowledgments

We thank Miss.Taiba Anwar for critical reading of the manuscript.

Competing interests

The authors of the manuscript declare no conflict of interest.

References

1. Elia, A. E.; Rellos, P.; Haire, L. F.; Chao, J. W.; Ivins, F. J.; Hoepker, K.; Mohammad, D.; Cantley, L. C.; Smerdon, S. J.; Yaffe, M. B. The molecular basis for phosphodependent substrate targeting and regulation of Plks by the Polo-box domain. *Cell*, **2003**, 115 (1), 83-95.
2. Strebhardt, K.; Ullrich, A., Targeting polo-like kinase 1 for cancer therapy. *Nat Rev Cancer* , **2006**, 6 (4), 321-330.
3. Golsteyn, R. M.; Mundt, K. E.; Fry, A. M.; Nigg, E. A.Cell cycle regulation of the activity and subcellular localization of Plk1, a human protein kinase implicated in mitotic spindle function. *The J Cell Biol*, **1995**, 129 (6), 1617-1628.
4. Toyoshima-Morimoti, F.; Taniguchi, E.; Shinya, N.; Iwamatsu, A.; Nishida, E. erratum: Polo-like kinase 1 phosphorylates cyclin B1 and targets it to the nucleus during prophase. *Nature* , **2001**, 410 (6830), 847-847.
5. Lane, H. A.; Nigg, E. A. Antibody microinjection reveals an essential role for human polo-like kinase 1 (Plk1) in the functional maturation of mitotic centrosomes. *The J Cell Biol*, **1996**, 135 (6), 1701-1713.
6. Kops, G. J.; Weaver, B. A.; Cleveland, D. W. On the road to cancer: aneuploidy and the mitotic checkpoint. *Nat Rev Cancer*, **2005**, 5 (10), 773-785.
7. Takahashi, T.; Sano, B.; Nagata, T.; Kato, H.; Sugiyama, Y.; Kunieda, K.; Kimura, M.; Okano, Y.; Saji, S. Polo-like kinase 1 (PLK1) is overexpressed in primary colorectal cancers. *Cancer Sci*, **2003**, 94 (2), 148-152.
8. Wolf, G.; Hildenbrand, R.; Schwar, C.; Grobholz, R.; Kaufmann, M.; Stutte, H.-J.; Strebhardt, K.; Bleyl, U. Polo-like kinase: a novel marker of proliferation: correlation with estrogen-receptor expression in human breast cancer. *Pathol Res Pract***, 2000**, 196 (11), 753-759.

9. Tokumitsu, Y.; Mori, M.; Tanaka, S.; Akazawa, K.; Nakano, S.; Niho, Y., Prognostic significance of polo-like kinase expression in esophageal carcinoma. *Int J Oncol*, **1999**, 15 (4), 687-779.
10. Gray, P. J.; Bearss, D. J.; Han, H.; Nagle, R.; Tsao, M.-S.; Dean, N.; Von Hoff, D. D. Identification of human polo-like kinase 1 as a potential therapeutic target in pancreatic cancer. *Cancer Res*, **2004**, 3 (5), 641-646.
11. Knecht, R.; Elez, R.; Oechler, M.; Solbach, C.; von Ilberg, C.; Strebhardt, K. Prognostic significance of polo-like kinase (PLK) expression in squamous cell carcinomas of the head and neck. *Cancer Res*, **1999**, 59 (12), 2794-2797.
12. Takai, N.; Miyazaki, T.; Fujisawa, K.; Nasu, K.; Hamanaka, R.; Miyakawa, I. Expression of polo-like kinase in ovarian cancer is associated with histological grade and clinical stage. *Cancer Lett*, **2001**, 164 (1), 41-49.
13. Sato, F.; Abraham, J. M.; Yin, J.; Kan, T.; Ito, T.; Mori, Y.; Hamilton, J. P.; Jin, Z.; Cheng, Y.; Paun, B. Polo-like kinase and survivin are esophageal tumor-specific promoters. *Biochem Biophys Res Commun*, **2006**, 342 (2), 465-471.
14. Feng, Y. B.; Lin, D. C.; Shi, Z. Z.; Wang, X. C.; Shen, X. M.; Zhang, Y.; Du, X. L.; Luo, M. L.; Xu, X.; Han, Y. L. Overexpression of PLK1 is associated with poor survival by inhibiting apoptosis via enhancement of survivin level in esophageal squamous cell carcinoma. *Int J Cancer*, **2009**, 124 (3), 578-588.
15. Coupland, V. H.; Allum, W.; Blazeby, J. M.; Mendall, M. A.; Hardwick, R. H.; Linklater, K. M.; Møller, H.; Davies, E. A. Incidence and survival of oesophageal and gastric cancer in England between 1998 and 2007, a population-based study. *BMC cance*r, **2012**, 12 (1)
16. Adams, R.; Morgan, M.; Mukherjee, S.; Brewster, A.; Maughan, T.; Morrey, D.; Havard, T.; Lewis, W.; Clark, G.; Roberts, S. A prospective comparison of multidisciplinary treatment of oesophageal cancer with curative intent in a UK cancer network. *Eur J Surg Oncol*, **2007**, 33 (3), 307-313.
17. WHO. World Cancer Report. Lyon: IARC Press; **2003**
18. Kim, S. M.; Park, Y.-Y.; Park, E. S.; Cho, J. Y.; Izzo, J. G.; Zhang, D.; Kim, S.-B.; Lee, J. H.; Bhutani, M. S.; Swisher, S. G. Prognostic biomarkers for esophageal adenocarcinoma identified by analysis of tumor transcriptome. *PloS on*e, **2010**, 5 (11), e15074
19. Peters, C. J.; Rees, J. R.; Hardwick, R. H.; Hardwick, J. S.; Vowler, S. L.; Ong, C. A. J.; Zhang, C.; Save, V.; O'Donovan, M.; Rassl, D. A 4-gene signature predicts survival of patients with resected adenocarcinoma of the esophagus, junction, and gastric cardia. *Gastroenterology*, **2010**, 139 (6), 1995-2004. e15.
20. Chin, L.; Gray, J. W. Translating insights from the cancer genome into clinical practice. *Nature*, **2008**, 452 (7187), 553-563.
21. Hawkins, R. D.; Hon, G. C.; Ren, B. Next-generation genomics: an integrative approach. *Nat RevGenet*, **2010**, 11 (7), 476-486.
22. Berger, J. A.; Hautaniemi, S.; Mitra, S. K.; Astola, J. Jointly analyzing gene expression and copy number data in breast cancer using data reduction models. IEEE/ACM Tr *Comp Biol Bioinf*, **2006**, 3 (1), 2.
23. Shen, R.; Olshen, A. B.; Ladanyi, M. Integrative clustering of multiple genomic data types using a joint latent variable model with application to breast and lung cancer subtype analysis. *Bioinformatics*, **2009**, 25 (22), 2906-2912.

24. Guex, N.; Peitsch, M. C.; Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis*, **2009**, 30 (S1).
25. Sheelagh, F,; Claire A,; Robert O,; Jonathan H,; Stephen T,; Ted H,; David B,; Daniella Z. Potent and selective small molecule inhibitors of Polo-like kinase 1:Biological characterization AACR 103rd Annual Meeting, , 2012, Chicago, IL
26. Wu, S.; Zhang, Y. MUSTER: improving protein sequence profile–profile alignments by using multiple sources of structure information. *Proteins: Structure, Function, and Bioinformatics*, **2008**, 72 (2), 547-556.
27. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J *Comput Chem*, **2009**, 30 (16), 2785-2791.
28. De Vries, S. J.; Van Dijk, M.; Bonvin, A. M. The HADDOCK web server for data-driven biomolecular docking. *Nat Protoc*, **2010**, 5 (5), 883-897.
29. Torchala, M.; Moal, I. H.; Chaleil, R. A.; Fernandez-Recio, J.; Bates, P. A. SwarmDock: a server for flexible protein–protein docking. *Bioinformatics*, **2013**, 29 (6), 807-809.
30. Tovchigrechko, A.; Vakser, I. A. GRAMM-X public web server for protein–protein docking. *Nucleic Acids Res*, **2006**, 34 (suppl_2), W310-W314.
31. Meng, E. C.; Pettersen, E. F.; Couch, G. S.; Huang, C. C.; Ferrin, T. E. Tools for integrated sequence-structure analysis with UCSF Chimera. B*MC bioinformatics*, **2006**, 7 (1), 339.
32. Wallace, A. C.; Laskowski, R. A.; Thornton, J. M., LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng*, **1995**, 8 (2), 127-134.
33. Mortazavi, A.; Williams, B. A.; McCue, K.; Schaeffer, L.; Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods*,**2008**, 5 (7), 621-628.
34. Yaffe, M. B.; Smerdon, S. J. The use of in vitro peptide-library screens in the analysis of phosphoserine/threonine-binding domain structure and function. *Annu Rev Biophys Biomol Struct*, **2004**, 33, 225-244.
35. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr D*, **2010**, 66 (4), 486-501.
36. Barr, F. A.; Silljé, H. H.; Nigg, E. A. Polo-like kinases and the orchestration of cell division. *Nat. Rev. Mol. Cell Bi*ol, **2004**, 5 (6), 429-441.
37. Petronczki, M.; Lénárt, P.; Peters, J.-M. Polo on the rise—from mitotic entry to cytokinesis with Plk1. *Dev. Cel*l, **2008**, 14 (5), 646-659.
38. Archambault, V.; Glover, D. M. Polo-like kinases: conservation and divergence in their functions and regulation. *Nat. Rev. Mol. Cell Biol*, **2009**, 10 (4), 265-275.
39. Taylor, S.; Peters, J.-M. Polo and Aurora kinases—lessons derived from chemical biology. *Curr. Opin. Cell Bio*l, **2008**, 20 (1), 77-84.
40. Tsvetkov, L.; Stern, D. F. Interaction of chromatin-associated Plk1 and Mcm7. *J Biol Chem*, **2005**, 280 (12), 11943-11947.
41. Van Vugt, M. A.; Gardino, A. K.; Linding, R.; Ostheimer, G. J.; Reinhardt, H. C.; Ong, S.-E.; Tan, C. S.; Miao, H.; Keezer, S. M.; Li, J. A mitotic phosphorylation feedback network connects Cdk1, Plk1, 53BP1, and Chk2 to inactivate the G2/M DNA damage checkpoint. *PLoS Biol*, **2010**, 8 (1), e1000287.
42. Luca, M. D.; Lavia, P.; Guarguaglini, G. A functional interplay between Aurora-A, Plk1 and TPX2 at spindle poles: Plk1 controls centrosomal localization of Aurora-A and TPX2 spindle association. *Cell Cycle*, **2006**, 5 (3), 296-303.

43. Li, H.; Wang, Y.; Liu, X. Plk1-dependent phosphorylation regulates functions of DNA topoisomerase IIα in cell cycle progression. *J Biol Chem*, **2008**, 283 (10), 6209-6221.
44. Denison, C.; Rudner, A. D.; Gerber, S. A.; Bakalarski, C. E.; Moazed, D.; Gygi, S. P. A proteomic strategy for gaining insights into protein sumoylation in yeast. . *Mol Cell Proteomics***, 2005,** 4 (3), 246-254.
45. Hannich, J. T.; Lewis, A.; Kroetz, M. B.; Li, S.-J.; Heide, H.; Emili, A.; Hochstrasser, M. Defining the SUMO-modified proteome by multiple approaches in Saccharomyces cerevisiae *J BiolChem*, **2005**, 280 (6), 4102-4110.
46. Panse, V. G.; Hardeland, U.; Werner, T.; Kuster, B.; Hurt, E. A proteome-wide approach identifies sumoylated substrate proteins in yeast. *J BiolChem*, **2004**, 279 (40), 41346-41351.
47. Wohlschlegel, J. A.; Johnson, E. S.; Reed, S. I.; Yates, J. R. Global analysis of protein sumoylation in Saccharomyces cerevisiae. *J Biol Chem*, **2004**, 279 (44), 45662-45668.
48. Dasso, M. Emerging roles of the SUMO pathway in mitosis Cell Div, **2008**, 3 (1), 5.
49. Gill, G. Something about SUMO inhibits transcription *CurrOpin Genet Dev*, **2005**, 15 (5), 536-541.
50. Hay, R. T. SUMO: a history of modification. *Mol Cell*, **2005**, 18 (1), 1-12.
51. Hay, R. Role of ubiquitin-like proteins in transcriptional regulation. In The Histone Code and Beyond, Springer: 2006; pp 173-192.
52. Moschos, S. J.; Mo, Y.-Y. Role of SUMO/Ubc9 in DNA damage repair and tumorigenesis. J *MolHisto*l, **2006**, 37 (5), 309-319.
53. Pastushok, L.; Xiao, W. DNA postreplication repair modulated by ubiquitination and sumoylation *Adv Protein Chem* **2004**, 69, 279-306.
54. Seeler, J.-S.; Bischof, O.; Nacerddine, K.; Dejean, A. SUMO, the three Rs and cancer. In Acute Promyelocytic Leukemia, Springer: 2007; pp 49-71.
55. Matic, I.; Macek, B.; Hilger, M.; Walther, T. C.; Mann, M. Phosphorylation of SUMO-1 occurs in vivo and is conserved through evolution. . *J Proteome Res*,**2008**, 7 (9), 4050-4057.
56. Rodriguez, M. S.; Desterro, J. M.; Lain, S.; Midgley, C. A.; Lane, D. P.; Hay, R. T. SUMO-1 modification activates the transcriptional response of p53. *EMBO J*, **1999**, 18 (22), 6455-6461.
57. McKenzie, L.; King, S.; Marcar, L.; Nicol, S.; Dias, S. S.; Schumm, K.; Robertson, P.; Bourdon, J.-C.; Perkins, N.; Fuller-Pace, F. p53-dependent repression of polo-like kinase-1 (PLK1). *Cell Cycle*, **2010**, 9 (20), 4200-4212.
58. Zhu, H.; Chang, B.-D.; Uchiumi, T.; Roninson, I. B. Identification of Promoter Elements Responsible for Transcriptional Inhibition of Polo-like Kinase 1 and Topoisomerase IIα Genes by p21WAF1/CIP1/SDI1. *Cell Cycle*, **2002**, 1 (1), 55-62.
59. Stamler, R.; Kappé, G.; Boelens, W.; Slingsby, C. Wrapping the α-crystallin domain fold in a chaperone assembly. *J Mol Biol*, **2005**, 353 (1), 68-79.
60. Mymrikov, E. V.; Seit-Nebi, A. S.; Gusev, N. B. Large potentials of small heat shock proteins. *Physiol Rev*, **2011**, 91 (4), 1123-1159.
61. Assimakopoulou, M.; Sotiropoulou-Bonikou, G.; Maraziotis, T.; Varakis, I. Prognostic significance of Hsp-27 in astrocytic brain tumors: an immunohistochemical study. *Physiol Rev*, **1997**, 17 (4A), 2677-2682.
62. Ciocca, D. R.; Calderwood, S. K. Heat shock proteins in cancer: diagnostic, prognostic, predictive, and treatment implications. *Cell Stress Chaperones*, **2005**, 10 (2), 86-103.
63. Fuqua, SA,; Oesterreich, S,; Hilsenbeck SG,; Von Hoff, D.D,; Eckardt, J,; Osborne, C.K. Heat shock proteins and drug resistance. *Breast Cancer Res Treat*, **1994**, 32, 67-71.

64. Huang, Q.; Ye, J.; Huang, Q.; Chen, W.; Wang, L.; Lin, W.; Lin, J.; Lin, X. Heat shock protein 27 is over-expressed in tumor tissues and increased in sera of patients with gastric adenocarcinoma. *ClinChem Lab Med*, **2010**, 48 (2), 263-269.
65. Muzio, L. L.; Leonardi, R.; Mariggio, M.; Mignogna, M.; Rubini, C.; Vinella, A.; Pannone, G.; Giannetti, L.; Serpico, R.; Testa, N. HSP 27 as possible prognostic factor in patients with oral squamous cell carcinoma. . *HistolHistopathol*, **2004**, 19 (1), 119-128.
66. Kato, K.; Hasegawa, K.; Goto, S.; Inaguma, Y. Dissociation as a result of phosphorylation of an aggregated form of the small stress protein, hsp27. *J Biol Chem*, **1994**, 269 (15), 11274-11278.
67. Hayes, D.; Napoli, V.; Mazurkie, A.; Stafford, W. F.; Graceffa, P. Phosphorylation dependence of hsp27 multimeric size and molecular chaperone function. *J Biol Chem*, **2009**, 284 (28), 18801-18807.
68. Chauhan, D.; Li, G.; Hideshima, T.; Podar, K.; Mitsiades, C.; Mitsiades, N.; Catley, L.; Tai, Y. T.; Hayashi, T.; Shringarpure, R. Hsp27 inhibits release of mitochondrial protein Smac in multiple myeloma cells and confers dexamethasone resistance. *Blood*, **2003**, 102 (9), 3379-3386.
69. Chen, Y. J.; Lin, Y. P.; Chow, L. P.; Lee, T. C. Proteomic identification of Hsp70 as a new Plk1 substrate in arsenic trioxide-induced mitotically arrested cells. *Proteomics*, **2011**, 11 (22), 4331-4345.
70. de Cárcer, G. Heat shock protein 90 regulates the metaphase-anaphase transition in a polo-like kinase-dependent manner. *Cancer Res*,**2004**, 64 (15), 5106-5112.

Figure Legends

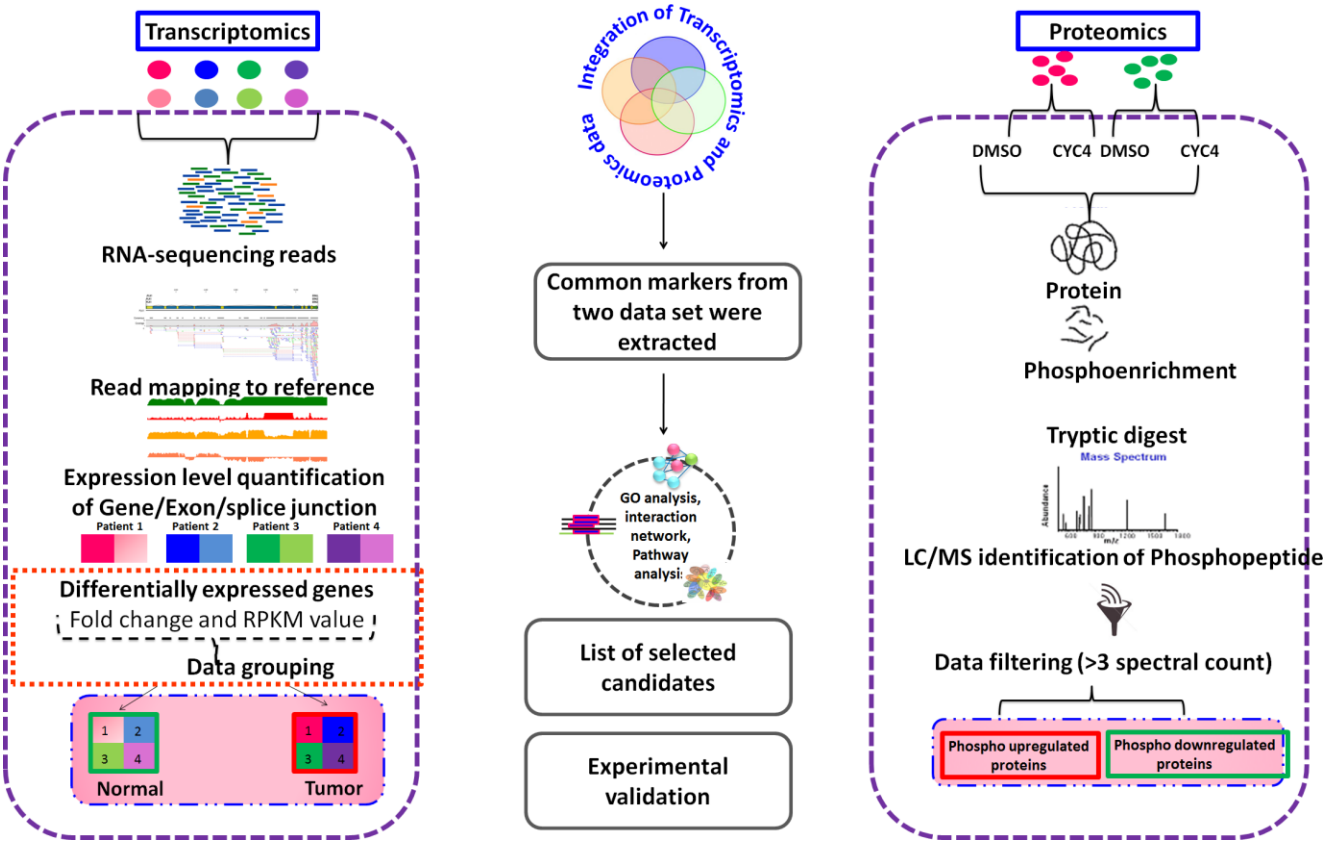

Figure 1: Integrative approach schematics. Transcriptomics and proteomics data types were integrated for high throughput investigation of PLK1 targeted biomarkers of OAC. Each data type was analyzed individually; information from both data type was amalgamated and later on, by downstream Bioinformatics analysis putative direct PLK1 target was selected for experimental validation.

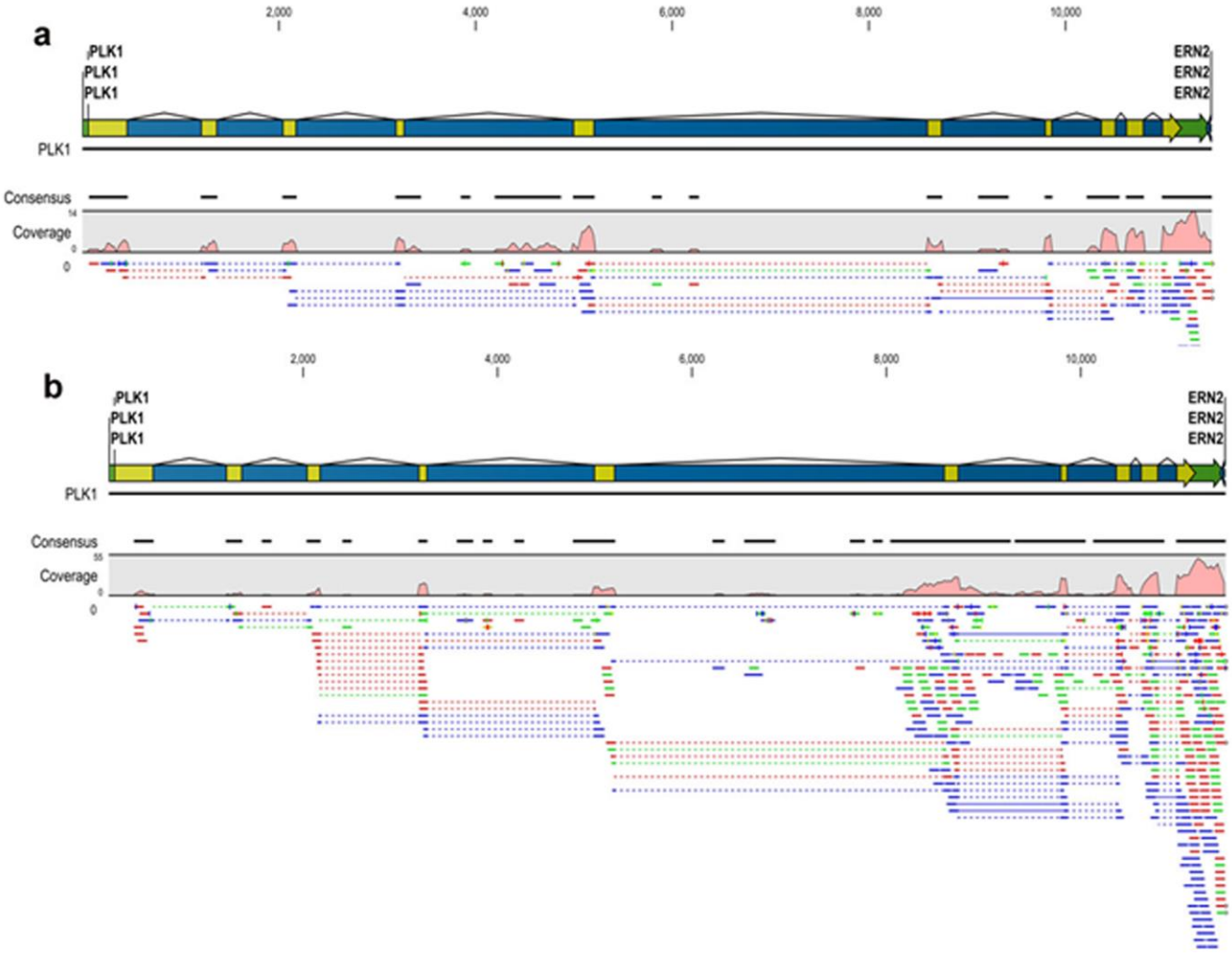

Figure 2: Overexpression of PLK1 in RNA sequencing data from four patients with OAC phenotype. *PLK1* gene (a) normal tissue (b) tumor tissue. Reads are mapped for each exon. Number of mapped reads is higher for tumor sample in comparison to normal counterpart which clearly indicates overexpression of PLK1 in OAC.

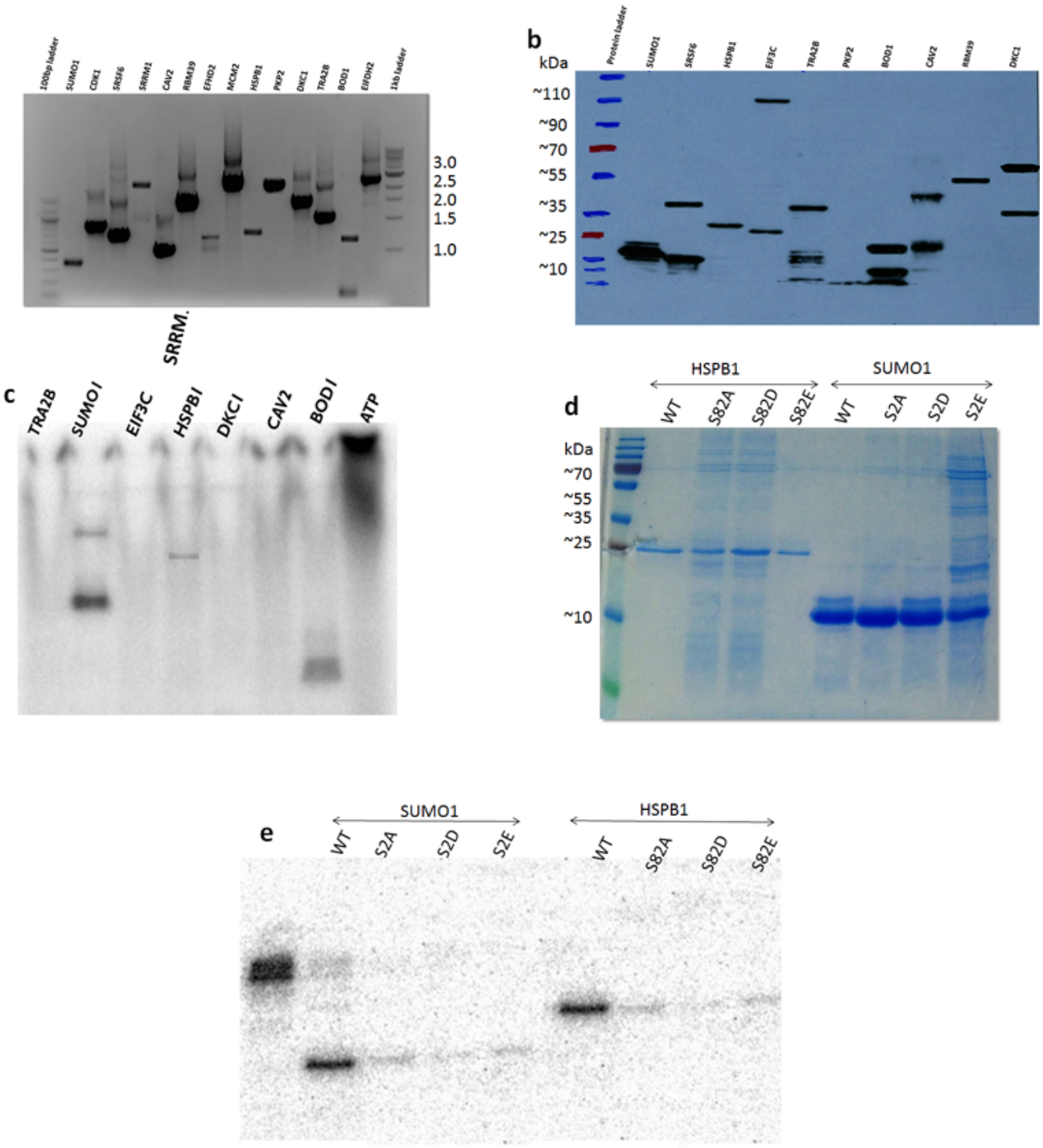

Figure 3: Cloning, protein purification and *in vitro* validation of identified hits. (a) Substrates were cloned in bacterial expression Gateway cloning vector and (b) expressed in E. coli BL21 DE3 cell line and proteins for 10 putative candidate substrates including SUMO1 (11 kDa), SRSF6 (39kDa) HSPB1

(27kDa), EIF3C (105kDa), TRA2B (33kDa), PKP2 (97kDa), BOD1 (19kDa), CAV2 (18kDa), RBM39 (59kDa) and DKC1 (57kDa) were purified and analyzed through immunoblotting. (c) PLK1 phosphorylation assays were performed on potential substrates from the candidate list. Precipitated proteins were mixed with PLK1 or only buffer in the presence of ^{32}P-ATP, then separated by SDS-PAGE and stained with coomassie brilliant blue. Incorporated ^{32}P was visualized by autoradiography. Exposure time of the autoradiographs was adapted for each substrate to allow the visualization of ^{32}P incorporation. (d) To confirm the site of phosphorylation, mutant constructs were generated for SUMO1 (Ser2A/Ser2D/Ser2E) and HSPB1 (Ser82A/Ser82D/Ser82E) with subsequent purification of proteins separated by SDS-PAGE. (e) Upon Phosphorylation with PLK1, reduced signal was observed through autoradiography for mutant proteins in comparison to wild type.

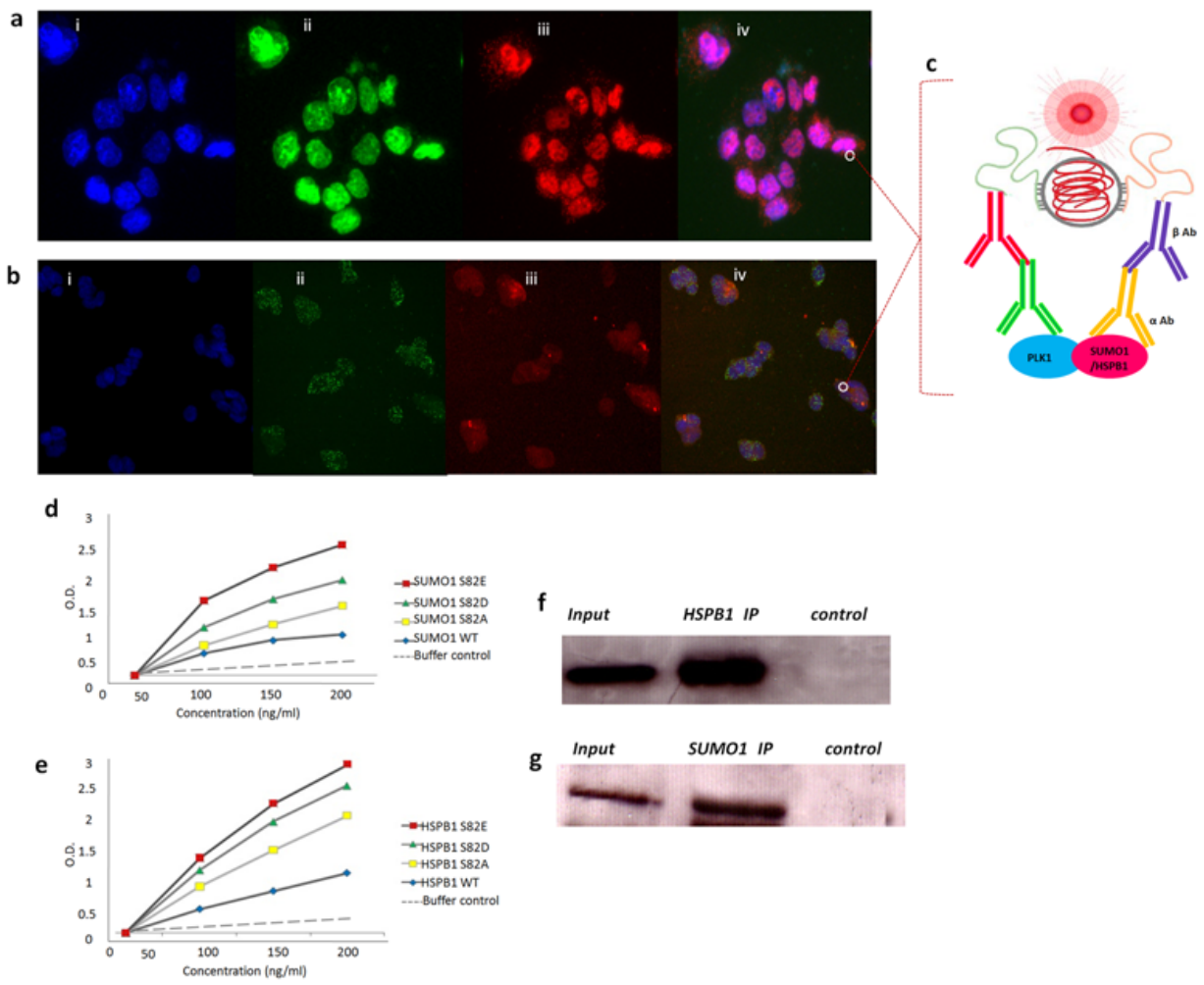

Figure 4: Protein-protein interaction experiments. Proximity ligation assay (PLA) was performed using primary antibodies directed against (a) SUMO1, (b) HSPB1 and PLK1. Red labeled particles [a (iv), b (iv)] represent interaction of PLK1 with SUMO1 and HSPB1. The hybridization probe is labeled red, cytoplasm is labeled green and nuclei blue. (c) If two proximity probes come closer to each other, then subsequently added linear connector oligonucleotide are guided to form a circular structure which is covalently joined by enzymatic DNA ligation. After ligation, hybridization of fluorescent labeled oligonucleotide complementary to a tag sequence can be detected. (d) Direct protein binding ELISA of SUMO1-WT and its mutant proteins with PLK1-PBD. (e) Binding analysis of HSPB1-WT and its mutants proteins with PLK1-PBD through direct protein binding ELISA. Mutant and WT of SUMO1

and HSPB1 showed binding to PLK1-PBD with equal affinity in comparison to buffer control. (f) Co-immunoprecipitation of endogenous HSPB1 with PLK1 in A375 cell line. (g) Co-immunoprecipitation of endogenous SUMO1 with PLK1 in A375 cell line. The lanes are: Input lane represents whole cell extract, control lane represents cell extract without any antibody and finally lane 2 represents immunoprecipitated proteins.

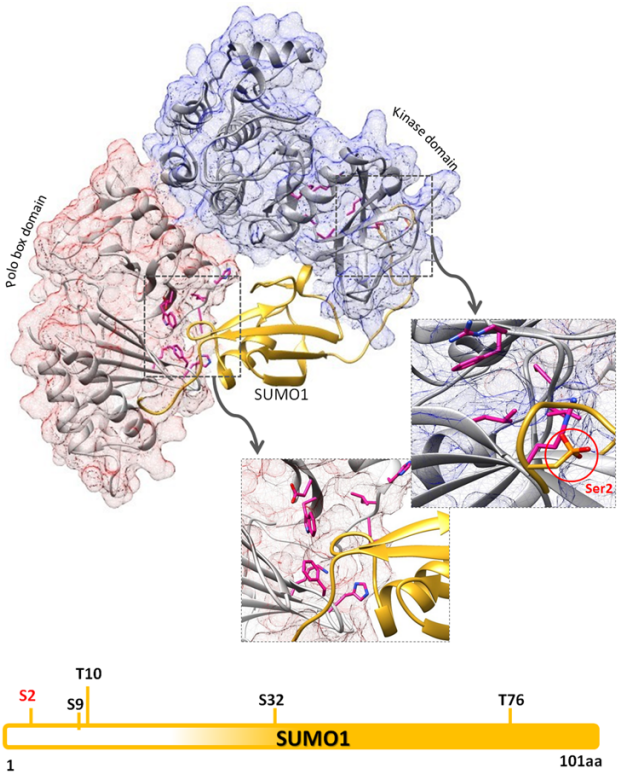

Figure 5. *In silico* interaction mapping of SUMO1 and PLK1. Interaction of SUMO1 was mapped with PBD (light gray ribbon) and kinase domain (dim gray ribbon) of PLK1. The interacting residues of both PLK1 domains are shown in pink sticks. SUMO1 is represented in golden ribbons, its N-terminal region with putative phopshorylation site (Ser2) which is pointed towards kinase domain and showed interaction of its C-terminal part with key residues (Trp414, His538, Lys540, Phe535 and His489) of PBD. Architecture of SUMO1 is shown along with phosphorylation sites mapped on it. Information about phosphorylation site was gained from phosphositeplus (*http://www.phosphosite.org/*) and

phosphoELM (*http://phospho.elm.eu.org/*). Phosphorylation site targeted by PLK1 (identified in present study) is highlighted in red color.

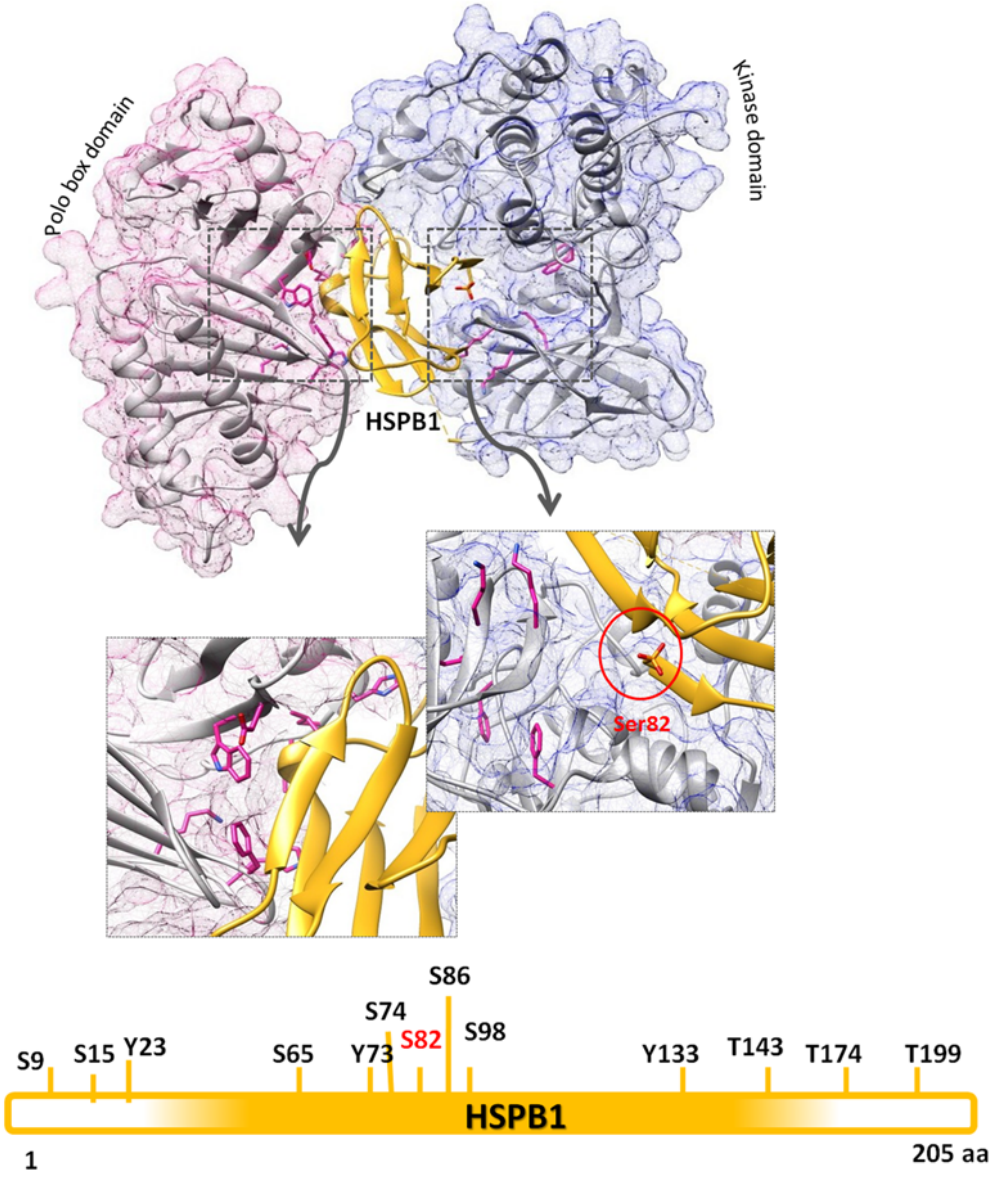

Figure 6: *In silico* interaction mapping of HSPB1 and PLK1. Interaction of HSPB1was mapped with PBD (light gray ribbon) and kinase domain (dim gray ribbon) of PLK1. The interacting residues of both PLK1 domains are shown in pink sticks. HSPB1 is represented in golden ribbons; its N-terminal

region with putative phosphorylation site (Ser2) is shown to bind with kinase domain while its (HSPB1) C-terminal part showed interaction with key residues (Trp414, His538, Lys540, Phe535 and His489) of PBD. Architecture of HSPB1 is shown along with predicted phosphorylation sites mapped on it.

Table 1: List of candidates shortlisted through integration of two omics data sources**.**

OVEREXPRESSED/PHOSPHO-UP				UNDEREXPRESSED/PHOSPHO-DOWN			
23T	**26T**	**27T**	**45T**	**23T**	**26T**	**27T**	**45T**
CAV2	**CAV2**	**CAV2**	**CAV2**	**ANXA2**	**ANXA2**	**ANXA2**	**ANXA2**
DDX21	**CCDC6**	**BAG3**	**CDK1**	**BAG3**	**BAG3**	**BIN1**	**BAG3**
DKC1	**CDK1**	**CDK1**	**COPA**	**CCDC6**	**DKC1**	**CDC37**	**CCDC6**
DOCK5	**COPA**	**COPA**	**DDX21**	**HSPB1**	**HSPB1**	**HSPB1**	**HSPB1**
DSG2	**DDX46**	**CTR9**	**DDX46**	**CDK1**	**EIF3C**	**LAD1**	**CDC37**
EFHD2	**DOCK5**	**DDX21**	**DKC1**	**COPA**	**FLNA**	**MYL9**	**EIF3C**
FLNB	**DSG2**	**DDX46**	**DOCK5**	**CTR9**	**LAD1**	**NAF1**	**HMGA1**
HDGF	**EFHD2**	**DKC1**	**DSG2**	**DDX46**	**LMNA**	**NEMF**	**LAD1**
CDK1	**FLNB**	**DSG2**	**EFHD2**	**EIF3C**	**MFAP1**	**PTRF**	**MAP4**
MCM2	**MCM2**	**MCM2**	**MCM2**	**G3BP1**	**MYL9**	**RBM14**	**MFAP1**
MYH9	**HDGF**	**EFHD2**	**FLNB**	**HN1**	**NEMF**	**RTN4**	**NEMF**
NOLC1	**HN1**	**EIF3C**	**HDGF**	**LAD1**	**NOC2L**	**SCRIB**	**PA2G4**
NOP2	**LIMA1**	**HDGF**	**LIMA1**	**LIMA1**	**NOP2**	**STMN1**	**PCDH1**
NOP58	**MCM2**	**HN1**	**MYH9**	**LMNA**	**NOP56**		**RBM14**
NSUN2	**MYH9**	**LIMA1**	**MYL9**	**MFAP1**	**PA2G4**		**RTN4**
PA2G4	**NOP58**	**MYH9**	**NOLC1**	**MYL9**	**PCDH1**		**SF3A1**
PKP2	**PAK2**	**NOLC1**	**NOP56**	**NEMF**	**PLEC**		**SLTM**
PTRF	**PKP2**	**NOP2**	**NOP58**	**NOP56**	**RTN4**		**SRRM2**
RBM39	**PTRF**	**NOP56**	**PKP2**	**PAK2**	**SRRM2**		**SRSF9**

SRRT	**RBM39**	**NOP58**	**PTRF**	**PCDH1**	**STRN**	**STRN**
SRSF6	**SF3A1**	**PA2G4**	**RBM39**	**RABL6**		**TMX1**
SRSF9	**SLTM**	**PAK2**	**SNW1**	**RTN4**		**TOP2B**
SSFA2	**SNW1**	**PKP2**	**SRRM1**	**SLTM**		
SSRP1	**SRRM1**	**RBM39**	**SRSF6**	**SNW1**		
SUMO1	**SRRT**	**SLTM**	**SSFA2**	**SRRM1**		
TOP1	**SRSF6**	**SNW1**	**SSRP1**	**SRRM2**		
TOP2A	**SRSF9**	**SRRM1**	**STMN1**	**STMN1**		
TPX2	**SSFA2**	**SRRM2**	**SUMO1**	**STRN**		
TRA2A	**SSRP1**	**SRSF6**	**TOP1**	**TMX1**		
	STMN1	**SSFA2**	**TOP2A**	**TOP2B**		
	SUMO1	**SSRP1**	**TPX2**	**TRA2B**		
	TMX1	**TMX1**	**TRA2A**			
	TOP1	**TOP1**				
	TOP2A	**TOP2A**				
	TOP2B	**TOP2B**				
	TPX2	**TPX2**				
	TRA2A	**TRA2B**				