MCP Papers in Press. Published on January 5, 2017 as Manuscript M116.064246

Running title: Eg5 mechanisms by HX-MS

Novel allosteric pathway of Eg5 regulation identified through multivariate statistical analysis of HX-MS ligand screening data

Joey G. Sheff[1], Farshad Farshidfar[2,3], Oliver F. Bathe[2,3], Karen Kopciuk[4,5], Francesco Gentile[6], Jack Tuszynski[6,7], Khaled Barakat[8], David C. Schriemer[1,9]*

[1]Department of Chemistry, University of Calgary, Calgary, Alberta, Canada.
[2]Department of Surgery, University of Calgary, Calgary, Alberta, Canada.
[3]Department of Oncology, University of Calgary, Calgary, Alberta, Canada.
[4]Department of Mathematics and Statistics, University of Calgary, Calgary, Alberta, Canada.
[5]Cancer Epidemiology and Prevention Research, Alberta Health Services, Calgary, AB, Canada.
[6]Department of Physics, University of Alberta, Edmonton, Alberta, Canada.
[7]Department of Oncology, University of Alberta, Edmonton, Alberta, Canada.
[8]Faculty of Pharmacy and Pharmaceutical Sciences, University of Alberta, Edmonton, Canada
[9]Department of Biochemistry and Molecular Biology, University of Calgary, Alberta Canada

*correspondence to: dschriem@ucalgary.ca

Abbreviations

HCA	Hierarchical clustering analysis
HX-MS	Hydrogen-deuterium exchange mass spectrometry
IC50	Half maximal inhibitory concentration
IPTG	Isopropyl β-D-1-thiogalactopyranoside
MT	Microtubule
MESG	2-amino-6-mercapto-7-methylpurine ribonucleoside
NL	Neck linker
OPLS-DA	Orthogonal-projection to latent structures discriminant analysis
PCA	Principal component analysis
PMSF	Phenylmethylsulfonyl fluoride
R^2	Goodness of fit
Q^2	Predictive ability
STLC	S-Trityl-L-Cysteine
VIP	Variable importance to projection

Summary

The mitotic kinesin Eg5 is an important target in cancer chemotherapy. A structurally diverse collection of canonical loop L5 inhibitors engage an allosteric pathway that includes elements of its microtubule binding region. However, recent evidence suggests that Eg5 may permit alternative allosteric mechanisms. Terpendole E, a natural-product Eg5 inhibitor, is active against mutants resistant to canonical loop L5 inhibitors and appears to offer a unique mode of inhibition. To investigate the variety of inhibitor responses, the structure-function properties of eighteen kinesin inhibitors were quantified with hydrogen-exchange mass spectrometry (HX-MS), functional analysis and molecular modeling. A unique strategy for high-density data analysis was implemented, based on a scalable multivariate statistical method, as current HX-MS routines have a limited capacity to guide a characterization of ligands when additional functional data is available. Inhibitor evaluation was achieved using orthogonal partial least squares projection to latent structures discriminant analysis (OPLS-DA). The strategy generated a model that identified functionally-significant conformational elements involved in kinesin inhibition, confirming the canonical allosteric pathway and identifying a novel response pathway. Terpendole E is demonstrated to be an atypical L5 site inhibitor, where binding induces an allosteric effect mediated by a destabilization in the β-sheet core of the molecular motor, an element involved in mechanochemical coupling for structurally-related kinesins. The analysis suggests that a different approach to inhibitor development may be fruitful.

Introduction

The kinesin family of molecular nanomotors convert the free energy of nucleotide hydrolysis into coordinated mechanical movement on microtubules (1, 2). In cell division, the kinesin Eg5 is an essential element for maintaining proper spindle dynamics and preserving spindle bipolarity. Given its role in dividing cells, Eg5 is an important target for anticancer drugs with the potential to overcome the clinical deficiencies presented by conventional mitotic poisons (3, 4). Drug candidates like ispinesib inhibit Eg5 and induce mitotic arrest followed by apoptosis. Continued ligand development is driven in part by the observation of deactivating mutations in the drug binding region, and a dearth of successful monotherapies based on Eg5 inhibition.

Ligands target a surface comprised of a dynamic loop (L5) that exerts allosteric control over motor function (5). All ligands for which mechanistic information is available engage some element of an integrated allosteric circuit that regulates Eg5's motor function. ATP hydrolysis in the "switch I" domain influences the "switch II" microtubule binding region 10's of angstroms away, and controls the orientation of the neck-linker (NL) domain that is involved in the power stroke (5). All known L5 inhibitors slow ADP release and preserve an ATP-like state. They induce a conformational change in the switch II region and their study suggests that motor control necessarily involves switch II (6, 7). Support for this idea is found in the recent discovery of inhibitors that directly engage switch II (8).

Terpendole E, the first natural product inhibitor of kinesin-5 to be discovered (9), appears to be a different class of ligand. While it generates the classical monoastral spindle seen with L5-site ligands, recent work suggests that functional inhibition proceeds through a different mechanism, and perhaps through a novel site (10). Interestingly, Terpendole E inhibits Eg5

mutants that are resistant to loop L5 inhibitors and switch II-directed inhibitors (9, 10). To explore this mechanism in further detail and to determine if an alternative regulatory pathway exists, we implemented a large-scale hydrogen deuterium exchange mass spectrometry (HX-MS) method for conformational analysis. HX-MS is a pseudo-structural method that can provide a peptide-level resolution of conformational response data for ligand binding events, in the form of variable mass shifts (11, 12). The approach is useful for mechanism-of-action studies (13), and is potentially very powerful for driving the development of pharmacophore models.

However, current strategies for HX-MS data analysis cannot easily mine large sets of shift data for mechanistically-significant information. Most protein-ligand analyses involve binary comparisons, where one ligand is related to a control state and *p*-values are measured for each peptide. Application of a Tukey test partially addresses the issue of multiple comparisons in HX-MS data (14), and hierarchical clustering offers a means of peptide classification (15). However, we need methods that scale to any size of screening activity, while retaining an ability to use all the data available from advanced HX-MS technology and other methods. Altered deuteration at a given location in protein can be represented by sets of overlapping peptides of variable lengths, often with multiple charge states, and each with non-equivalent back-exchange properties. Available interpretation strategies do not make use of the rich redundancy of information (and the variable sensitivity of peptides for detecting mass shifts) that such analyses offer. There are similarities between the statistical challenges presented by HX-MS and comparative ‘omics, where multivariate methods have demonstrated utility (16, 17). In this study, we demonstrate how such methods can be used to mine shift data, with a view to classifying and visualizing HX-MS data sets arising from ligand screening exercises. The new methodology was combined with potency measures of nucleotide hydrolysis and computational

modeling to classify a breadth of Eg5 inhibitors. We show that Terpendole E invokes a non-canonical mechanism of regulation, which appears to derive from altered occupancy of the L5 binding site region.

Experimental Procedures

Eg5 motor domain expression and purification

The motor domain of Eg5 (1-386) was expressed and purified following a previously described procedure (18), with some modifications. Briefly, pEt28a(+) plasmid DNA (Genscript) was transformed into competent BL-21 *E. coli*. Cells were grown overnight at 37°C supplemented with kanamycin (50 µg/mL). Next, 19 mL of the culture was transferred to 1L of 2YT supplemented with kanamycin (50 µg/mL) at 37 °C for approximately 7h until an OD600 reading of 0.72 was observed. Expression was induced with the addition of IPTG (0.5 mM). Cells were then left on a shaker overnight at room temperature. Cells were pelleted, and the pellet resuspended in cold buffer A (20 mM piperazine, 200 mM NaCL, 1 mM MgCl2, and 1 mM EGTA), supplemented with 1mM PMSF. Cells were lysed with a high pressure homogenizer, and centrifuged for 30 min at 37000RPM and 4°C. Cell lysate was collected and loaded onto Ni-NTA agarose gel, and washed. Eg5 was eluted with an imidazole gradient (10-500 mM) in a gravity flow column. Collected fractions were analyzed with SDS-PAGE, and purified protein concentrations were determined to be ~3.3 mg/mL (75 µM) by a BCA assay.

HX-MS-based ligand screening assays

Stock solutions (5-10 mM) of 18 kinesin inhibitors were prepared in DMSO, and frozen at -80 °C. For each HX-MS experiment, fresh dilutions (100 µM inhibitor, in 10 mM PIPES, 1 mM ADP, 1 mM $MgCl_2$,) and matched labelling solutions (50 µM inhibitor, in 74% D_2O, 10 mM PIPES, 1 mM ADP, 1 mM $MgCl_2$,) were prepared. Control solutions contained an equivalent 2% DMSO in the appropriate buffer. Eg5-inhibitor solutions were prepared by mixing equal volumes of diluted inhibitor solution with 2.5 µM Eg5, and equilibrated on ice. HX was initiated by adding labelling solution to a final level of 45% D_2O for labeling at 25 °C. Replicate kinetics

data were obtained for the inhibitor S-Trityl-L-Cysteine (STLC), where labelling times ranged from 0.5 to 60 min. For all other inhibitors, labelling was conducted for 5 min at 25 °C. Samples were quenched and digested with cold concentrated nepenthesin extract (~2.3 pmol of active protease) in 7 µL 100 mM Gly-HCl (pH 2.5) (19). Sample (5 pmol) was injected into a LEAP-PAL HTX configuration, outfitted with a self-packed preconcentration cartridge (25 mm x 250 µm i.d. capillary, 200Å, 5 µm Magic C18 beads, Michrom BioResources) and a self-packed analytical column (70 mm x 150 µm i.d. capillary, 200Å, 5 µm Magic C18 beads). Data for all 18 drug states and DMSO control were collected in quadruplicate with a TripleTOF 5600 (SCIEX) coupled to an Eksigent nanoLC-ultra-2D pump. The complete dataset was collected in two separate batches. A single replicate of the control was collected first, followed each ligand state in the batch. This was repeated twice, following the same order of analysis for each replicate.

Mascot v2.4 was used to identify peptides for analysis. Briefly, a search was performed against a database containing all proteins present in this study (Eg5, contaminants and proteases: 13 sequences) with a mass tolerance for precursor ions of 20 ppm, 0.05 Da for fragment ions, and a probability cutoff of p=0.05. No enzyme specificity was selected (thus missed cleavages are irrelevant), and no modifications were considered. The resulting peptide list was next imported into our in-house software package MS Studio (20) for deuteration analysis. Peptide quality was assessed based on intensity, signal to noise ratio and spectral overlap, and only those with reliable isotopic profiles were selected for analysis. The percent relative fitted deuteration for each high-quality peptide was exported. Woods plots for each protein state were created using a statistical module in MS Studio. All 76 raw HX-MS data sets were uploaded on Chorus (https://chorusproject.org/anonymous/download/experiment/4da6e5757fa2408194967ce8e705cf

9f).

Statistical analysis

A matrix of shift data from 76 observations (4 replicates from 19 protein states) and 183 variables (unique peptide features) was built with the HX-MS inhibitor screen data, and trimmed of missing data. The matrix was imported into SIMCA 14 (Umetrics AB, Umeå, Sweden) for subsequent analysis. Data was mean-centered and used in principal component analysis (PCA), generating an initial model of components, which was tested for goodness of fit (R^2) and predictive ability (Q^2) based on 7-fold cross validation. R^2 and Q^2 range between 0 and 1, with 1 representing the ideal performance. An orthogonal-projection to latent structures discriminant analysis (OPLS-DA) model was built using binding responses (Table 1) as class assignments. The OPLS-DA model was validated by applying a similar cross-validation strategy, combined with a permutation test (50 iterations) (21, 22).

IC50 determinations

A Kinesin ELIPA Biochem kit (BK060, Cytoskeleton Inc, Denver, CO) was used to verify the activity of the purified protein, and determine basal, in which no microtubules (MT) are present, IC50 values for all inhibitors. We measured the absorbance shift resulting from the catalytic conversion of 2-amino-6-mercapto-7-methylpurine ribonucleoside (MESG) to 2-amino-6-mercapto-7-methyl purine by purine nucleoside phosphorylase, in the presence of inorganic phosphate generated by the kinesin ATPase activity. Each inhibitor (0.01 to 200 μM) was mixed with Eg5 and ATP in a 384-well plate and absorbance was measured at 370 nm (10 nm bandwidth) with a Filtermax F5 spectrophotometer, for 120 min at 25 °C. IC50 values were calculated from the plate-blank corrected slope of each drug concentration. Dose response curves (Hill slope -1.0).were fit to the data in Graphpad Prism 6.07 (La Jolla, CA) using a least squares

fitting routine.

Molecular docking simulations

Previous Eg5-inhibitor structure determinations were used to prepare a docking strategy for Terpendole-E. We selected nine different structures of Eg5 (in bound form) from the PDB database, out of the thirty-five available, to present the widest diversity in known binding-site conformations. Structure selection was based on an RMSD analysis of the binding sites among the structures and visual inspection, where 1X88 (PDB ID) was taken as reference, and only structures with an RMSD higher than 0.80 Å were selected. Visual Molecular Dynamics (VMD) was employed for this step (23). The binding site was broadly defined as constituted by 15 residues (E116, G117, E118, R119, W127, D130, L132, A133, I136, P137, Y211, L214, E215, A218 and R221). In addition, we included two different mutants, D130A and L214A. Protonation states were assigned to the proteins using the Protonate3D tool in Molecular Operating Environment (MOE, Chemical Computing Group Inc, Montreal, Canada); the protonated structures were then energetically relaxed in MOE. Heteroatoms were removed from the targets. Non-polar hydrogens were merged to the respective carbons and Gasteiger charges were assigned using AutodockTools (23). The docking boxes were designed as 60X60X60 points, centered to the binding site, with 0.375 Å of spacing. The ligand structures were downloaded from PubChem, and an ensemble of representative 3D conformations was generated using LigPrep (Schrödinger, New York, USA). Docking was performed using Autodock4 and the built-in scoring function was used to score the poses (23, 24). Parameters for the simulations were 25,000,000 energetic evaluations, a population of 300 individuals, 27,000 maximal generations and 100 runs of Lamarckian Genetic Algorithm (25). RMSD-based clustering has been performed using a cutoff of 2Å. For each simulation, the final pose has been selected as the

lowest binding energetic one within the most populated cluster. For hydrogen bond detection, a distance cutoff of 3.5 Å and a maximal deviation from linearity of 60^{o} were used.

Chemicals and reagents

HPLC-grade H_2O and acetonitrile (ACN), D_2O (99.9% D), glycine, hydrochloric acid, trifluoroacetic acid (TFA), formic acid (FA), K858 (>98%), S-Trityl-L-cysteine (>97%), cyclobenzaprine hydrochloride and Terpendole E (>95%) were purchased from Sigma Aldrich (St Louis, MO). Monastrol (>98%), dimethylenenastron (>99%), BTB 1 (>99%), paprotrain (>99%) and ARRY 520 trifluroacetate (>98%) were purchased from TOCRIS (Chicago, IL). Ispinesib, SB743921, AZ 3146, GSK923295, ARQ 621 and MPI-0479605 were purchased Selleckchem (Houston, TX). EMD 534085 (>99.9%) was purchased from Chemical solutions corporation (Burnaby, CA). Trans-24 (>95%) and trans-HR22C16 (>95%) were purchased from Enzo Life Sciences, Inc. (Farmingdale, NY).

Results and Discussion

Building and validating an HX-MS assay

A full kinetics analysis of the Eg5 motor domain, with and without a known inhibitor, STLC, was used to select a sensitive time-point for screening purposes (Fig. 1). Previous studies have shown that a single HX-MS time-point is suitable for analysis of protein-ligand interactions (26). We determined that a 5 min labeling time in 45% D_2O provided sufficient shift discrimination for the full set of 183 Eg5 peptides, with deuterium uptake averaging 30%, and ranging from 6 to 70% (Fig. S1). A lower D_2O composition of the labelling solution was selected to minimize isotopic peak expansion and increase peak capacity, which we have shown retains sensitivity (27). At this labelling time point, STLC-induced stabilization was observed in Thr112-Ile143 (loop L5-α2 helix), Val158-Leu168 (β4) and Glu200-Ala230 (α3 helix-β5 strand), which encompass the known binding site (Fig. 2a,b). Crystallographic studies have shown that loop L5 reorients by 7Å to form an induced-fit hydrophobic binding site (6). The shifts we see in this region are supported by an earlier HX map of the Eg5-STLC interaction (28). We also observed stabilization in Leu293-Thr300 and Leu316-Asp322, the α4 and α5 helices that comprise the microtubule-binding region for the motor (28). The region constitutes part of the switch II cluster (Fig. 2a), which experiences conformational changes upon nucleotide (2) and L5-based ligand binding (6). Shifts in this region were not identified in previous HX-MS studies with STLC, likely due to lower sequence coverage (28, 29). Together, our analysis confirms that this single time-point HX-MS assay should provide the sensitivity required to detect local perturbations of protein conformation arising from inhibitor binding, whether from the binding event itself or its distal effects.

Functional screening and HX-MS of library

The study was expanded to include a set of 18 kinesin inhibitors, 13 directly targeting Eg5 and 5 selective for other kinesins. The Eg5 inhibitors present a wide range of basal IC50 values, as confirmed in an ATP assay using the Eg5 motor domain (Table 1). We then collected HX-MS assay data for all protein-ligand interactions. For each inhibitor, deuteration was measured in quadruplicate using 183 peptides, which represents a 5.6X sequencing depth and 92% coverage of the Eg5 motor domain (Fig. 2c). The shift data for each ligand relative to the vehicle was displayed using Woods plots (Fig. S2) (20), which provide a typical pairwise comparison of peptide deuteration against the protein sequence. However, the approach is impractical for sustaining a quantitative comparison of the full set of HX-MS profiles. For this purpose, we developed a multivariate data analysis strategy, which is illustrated in Fig. 3. HX-MS data is combined with functional evidence to inform the development of a multivariate statistical model, which groups all ligands according to how they drive a functional response. It returns a set of variables (peptides) that support the classifications, and helps sustain a more robust functional interpretation.

Translating HX data into a multivariate statistical model

In order to interrogate the entire set of assay data, we used OPLS-DA modelling, but began with PCA. PCA is an unsupervised method that allows us to inspect data structure prior to engaging methods like OPLS-DA (16). OPLS-DA is a supervised approach that has the ability to make use of known class information input as an additional Y-matrix (Fig. 3) (30, 31). This information is used to structure the variation present in the X-matrix into components that are predictive (correlated) or non-predictive (orthogonal) of Y-matrix class differentiation (31). In this regard, it seemed to us the most appropriate statistical method to apply towards HX-MS screening data. First, compound identity and/or the functional data associated with the compounds naturally

classify the observations to form a Y-matrix. Second, PLS-DA methods in general have been used very successfully in classifying the underlying data structures in a functionally interpretable manner (16, 32). They drive towards maximum class separation, which works well when the number of classes is limited (31). This is the natural condition found in the HX-MS experiment: the number of possible responses (i.e. classes) a protein can mount to a ligand-binding event will be restricted by its structure/function properties. As a result, the Y-matrix can be considerably smaller than the number of ligands. Third, the orthogonal signal correction filtering in OPLS-DA allows for noise reduction, which in the context of HX-MS applies the rich information in the Y-matrix (i.e. compound type, functional grouping, and/or HX-MS replicates) to highlight and remove non-predictive variation.

To treat HX-MS screening data by PCA and OPLS-DA, peptides missing greater than 50% of deuteration measurements across all 76 inhibitor replicates were removed from the 76 x 183 data matrix, leaving 138 peptides (92% sequence overage, 4.3X redundancy). Rather than implement a strategy of data imputation often used in comparative 'omics, we reasoned that the strong redundancy in sequence coverage tolerates a trimming of variables, and that imputing deuteration values to missing data cannot be done with accuracy. The only additional preprocessing of the data involved mean-centering (Fig. 3a). There was no need for scaling, a common preprocessing step in multivariate methods, as peptide HX-MS data is inherently scaled according to the percent of maximum deuteration (33). The OPLS-DA model output consisted of a single predictive component and three orthogonal components. A goodness-of-fit (R^2) value, which describes the fraction of total variation explained by the predictive components of the model, and a goodness-of-prediction (Q^2) value were determined using cross validation for both PCA (R^2X 0.927, Q^2X 0.883) and OPLS-DA (R^2Y 0.971, Q^2Y 0.958) models. Permutation

confirmed that the final OPLS-DA model was optimal (Fig. S3). Scaled and centered correlation coefficients for the separation of classes in OPLS-DA is shown in Table S1 (34).

Functional interpretation

Our preliminary PCA model demonstrated clear separation between binding and non-binding compounds (Fig. S4), showing a structure that warranted deeper analysis with a supervised method. Table 1 correlates this coarse grouping with the Eg5 inhibition data. OPLS-DA was next employed to evaluate the groups' underlying structure in greater detail, presented as a 3D scores plot (Fig. 4a). The analysis was based on a Y-matrix employing a simple categorization of the IC-50 data as ligands and non-ligands. As anticipated, all compounds with no influence on Eg5 activity (including the solvent vehicle used as a control) are grouped together and collectively define the profile of an insignificant change in protein conformation. In the main cluster of ligands, the strength of inhibition roughly increases from top to bottom along the second orthogonal component.

The loadings plot (Fig. 4b), which illustrates how each peptide feature contributes to each supervised component, can be used for further interpretation of the scores plots. For example, peptides in the L5 region contribute strongly to the model. The classification has strong representation from the magnitude of the deuteration shifts for peptides that span the binding site (6, 7): Thr112-Ile143 (the L5-α2 helix region), Val158-Phe169 (β4) and Glu200- Glu215 (the α3 helix-β5 region). Of course, these conclusions might be gleaned from an individual assessments of the Woods plots for each compound (Fig. S2). However, the Woods plots are difficult to use when comparing large numbers of ligands. Conclusions drawn from an individual plot are hard to extrapolate to the full ligand set, and the statistical basis for identifying important changes over the set is not supplied by such an approach. It is more robust to use the loadings plots in

order to identify variables that contribute strongly to the model. In addition, a comparative scores contribution plot can be generated from the model (e.g. Fig. S4), which specifically identifies variables that maximally discriminate one group from all others. With either approach, once a significant variable is identified in the model, the individual peptides can be validated by generating ligand-wide maps of the shift data, referenced to the control states (Fig. 5). These maps offer an evaluation of the significance of the shift, on a per-peptide level. For example, L5 inhibitors like STLC (Fig. 5a) reduce deuteration strongly in all of L5, as well as in the α2 and α3 helices at the base of the binding site (Fig. 5c and d). L5 ligands generate a conformational response at the switch II site (7), which is measured in our HX-MS experiments as reduced deuteration in the α4 and α5 helices (Fig. 5e and f). OPLS-DA also identified this region as a strong contributor to the separation of binding and non-binding compounds along the first predictive component (PC1, Fig. 4b).

The conformational change we observed in the switch II region does not appear to require a full stabilization of L5. Cyclobenzaprine, an inhibitor with a relatively weak IC50, shows some separation from the other inhibitors. It has a diminished effect on the conformational state of Phe113-Glu124, the N-terminal end of L5 (data not shown). It appears that partial L5 stabilization, provided it is accompanied by a strong conformational change in α2 and α3, is sufficient to stabilize the switch II region (Fig. 5e and f). However, overall, this allosteric pathway is invoked by all ligands with the exception of Terpendole E.

Alternative allosteric pathway of Eg5 regulation

Terpendole E is very distinctly separated in the OPLS-DA model from all other L5 ligands (Fig. 4a). It demonstrates stabilization of L5 and α2 (Fig. 5c), but very weak stabilization of α3 (Fig. 5d). We also do not observe a measurable effect on the β4 region of the binding site (data not

shown). Since this reduced labeling is still broadly within the known binding site for canonical inhibitors, we centered a molecular docking analysis of Terpendole E on the region, to determine if altered site occupancy could provide an explanation for the different stabilization pattern. As validation for our docking routine, we generated a docking analysis of all the Eg5 inhibitors used in this study, against a diverse set of Eg5 structures, and confirmed the known binding orientations (Fig. S5). Using this approach, Terpendole E was observed to orient in the binding site in a manner that partially overlaps the other inhibitors (Fig. 6). The docked model presents a different binding mode, most noticeable through a nonequivalent hydrogen bonding pattern (Table S2) and an α3 helix orientation that is farther away from both L5 and α2. This analysis suggests why the conformational state of the α3 helix was mainly unchanged in the HX-MS data, but the structure will require confirmation via a crystallographic analysis. We note that the model does not fully explain a recent observation that Terpendole E remains active against a Glu130Ala mutation in L5 (10). This residue is only slightly removed from Terpendole E relative to other ligands, and it is difficult to see how mutation would impact binding.

Nevertheless, inhibition by Terpendole E does not appear to exert allosteric control over the switch II region. No conformational changes are observed in either α4 or α5 (Fig. 5 e,f). These effects cannot be attributed to low potency, as Terpendole E has an IC50 in the range of group 2 ligands (Table 1). We dug deeper into the separation of Terpendole E from the rest of the inhibitors using a comparative contribution scores plot (Fig S4), to identify additional variables that significantly inform the discrimination of this compound from the rest. While the largest contributors span the inhibitor and microtubule binding regions, peptides covering the β-sheet core are also featured. A unique conformational response to this inhibitor is observed in the core β-sheet of the motor domain (Fig. 5b and g), where an increase in deuteration indicates

destabilization. Two overlapping peptides illustrate the change: Glu247-Leu261 and Lys255-Leu261. The strongest change is observed in Lys255-Leu261, located in β7, the middle strand of the β-sheet core. The transition appears conformationally significant, as β-sheets are normally very slow exchangers in the HX experiment.

It would seem that Terpendole E achieves motor inhibition using an alternative structural transition involving the central β-sheet in some fashion. To generate force on the microtubule lattice, a ratcheting motion centered on the neck-linker (NL) is a central element of kinesin motor function. Canonical inhibitors "dock" the NL onto the motor domain, facilitated by a conformational change in the switch II region (e.g. PDB 4AP0) (35, 36). These changes are accompanied by, and likely contribute to, the slowed release of ADP. The switch II region, and particularly the α4 helix within it, is a central element of the microtubule binding domain in Eg5. These inhibitors weaken the microtubule-motor interaction, which may also contribute to a nonproductive motor (37, 38). The HX-MS data we collected for these inhibitors correlate strongly with this mechanism (e.g. Fig. 5a). Conversely, with Terpendole E we observe no conformational effect on the switch II region, and detect no influence of Terpendole E on microtubule binding in a pulldown assay (data not shown). A role for the core β-sheet in regulating motor activity is more suggestive of force generation mechanisms in myosin (39, 40). In myosin, hydrolysis product release leads to a substantial twist of the core β-sheet. Recent structural data emerging from tubulin-kinesin-1 interactions do not show such a dramatic effect (1), however other studies of Eg5 suggest a role for hydrolysis- and inhibitor-regulated β-sheet conformational change. For example, nucleotide-dependent docking of the NL to the β-sheet core is essential for forward motility of the second motor head in the kinesin dimer, towards the MT plus-end (41). It was shown in this study that Eg5 sensitivity to inhibitors can be abolished if

there is a twist in β6- β7, as observed in a resistant Ala133Asp mutant. Further work will be necessary to determine if the terpendole/β-sheet pathway influences motor at the level of force generation, and independent of its behavior in ADP release.

Conclusions and perspective

It is possible to rank and differentiate ligands in large scale binding assays based on their full HX-MS profiles and combine this with activity data to correlate the conformation responses that align with the functional responses. Although traditional pairwise comparisons of drug states are suitable in many instances, multivariate methods of the type we have described are needed to identify and rank patterns of response in large sets of data. They scale better than existing approaches, and offer a useful way to integrate other data into the evaluation. These tools are needed as HX-MS technology continues to develop in robustness and throughput. Whether comparing large sets of ligands or different biologics, robust methods built on accepted statistical routines will ease the difficulty that we currently experience in the interpretation of HX-MS output. For example, the magnitude of a significant mass shift has little meaning in a straightforward pairwise comparison (e.g. drug vs control), but when it is combined with other data, and studied over large sets of ligands, it can become a potent discriminating feature. Multivariate methods therefore use the full richness of the data better than our existing tools. In addition, OPLS-DA models of the sort described in this study can be mined to define a simpler list of the most discriminating peptide features. These features can be tracked as peptide "biomarkers" for accurate predictive classification of inhibitors, in subsequent higher throughput HX-MS assays (42-44).

Our analysis shows that Terpendole E invokes an alternative allosteric pathway of motor regulation, and one that may support a novel pharmacophore. The strategy we present for a supervised multivariate analysis of screening data demonstrates that this classification can be a predictor of mechanism and potency, and can be useful for categorizing protein/ligand interactions prior to more involved structural characterizations.

References

1. Gigant, B., Wang, W., Dreier, B., Jiang, Q., Pecqueur, L., Pluckthun, A., Wang, C., and Knossow, M. (2013) Structure of a kinesin-tubulin complex and implications for kinesin motility. *Nature Structural & Molecular Biology* 20, 1001-1007
2. Kikkawa, M., Sablin, E. P., Okada, Y., Yajima, H., Fletterick, R. J., and Hirokawa, N. (2001) Switch-based mechanism of kinesin motors. *Nature* 411, 439-445
3. Waitzman, J. S., and Rice, S. E. (2014) Mechanism and regulation of kinesin-5, an essential motor for the mitotic spindle. *Biol Cell* 106, 1-12
4. Perez-Melero, C. (2014) KSP Inhibitors as Antimitotic Agents. *Curr Top Med Chem* 14, 2286-2311
5. Wojcik, E. J., Buckley, R. S., Richard, J., Liu, L., Huckaba, T. M., and Kim, S. (2013) Kinesin-5: cross-bridging mechanism to targeted clinical therapy. *Gene* 531, 133-149
6. Yan, Y., Sardana, V., Xu, B., Homnick, C., Halczenko, W., Buser, C. A., Schaber, M., Hartman, G. D., Huber, H. E., and Kuo, L. C. (2004) Inhibition of a mitotic motor protein: where, how, and conformational consequences. *Journal of Molecular Biology* 335, 547-554
7. Kaan, H. Y., Major, J., Tkocz, K., Kozielski, F., and Rosenfeld, S. S. (2013) "Snapshots" of ispinesib-induced conformational changes in the mitotic kinesin Eg5. *Journal of Biological Chemistry* 288, 18588-18598
8. Ulaganathan, V., Talapatra, S. K., Rath, O., Pannifer, A., Hackney, D. D., and Kozielski, F. (2013) Structural Insights into a Unique Inhibitor Binding Pocket in Kinesin Spindle Protein. *J Am Chem Soc* 135, 2263-2272
9. Nakazawa, J., Yajima, J., Usui, T., Ueki, M., Takatsuki, A., Imoto, M., Toyoshima, Y. Y., and Osada, H. (2003) A novel action of terpendole E on the motor activity of mitotic Kinesin Eg5. *Chemistry & Biology* 10, 131-137
10. Tarui, Y., Chinen, T., Nagumo, Y., Motoyama, T., Hayashi, T., Hirota, H., Muroi, M., Ishii, Y., Kondo, H., Osada, H., and Usui, T. (2014) Terpendole E and its derivative inhibit STLC- and GSK-1-resistant Eg5. *Chembiochem* 15, 934-938
11. Chalmers, M. J., Busby, S. A., Pascal, B. D., West, G. M., and Griffin, P. R. (2011) Differential hydrogen/deuterium exchange mass spectrometry analysis of protein-ligand interactions. *Expert Rev Proteomics* 8, 43-59
12. Bennett, M. J., Barakat, K., Huzil, J. T., Tuszynski, J., and Schriemer, D. C. (2010) Discovery and characterization of the laulimalide-microtubule binding mode by mass shift perturbation mapping. *Chem Biol* 17, 725-734
13. Kojetin, D. J., Matta-Camacho, E., Hughes, T. S., Srinivasan, S., Nwachukwu, J. C., Cavett, V., Nowak, J., Chalmers, M. J., Marciano, D. P., Kamenecka, T. M., Shulman, A. I., Rance, M., Griffin, P. R., Bruning, J. B., and Nettles, K. W. (2015) Structural mechanism for signal transduction in RXR nuclear receptor heterodimers. *Nat Commun* 6, 8013
14. West, G. M., Chien, E. Y., Katritch, V., Gatchalian, J., Chalmers, M. J., Stevens, R. C., and Griffin, P. R. (2011) Ligand-dependent perturbation of the conformational ensemble for the GPCR beta2 adrenergic receptor revealed by HDX. *Structure* 19, 1424-1432
15. Chalmers, M. J., Pascal, B. D., Willis, S., Zhang, J., Iturria, S. J., Dodge, J. A., and Griffin, P. R. (2011) Methods for the Analysis of High Precision Differential Hydrogen Deuterium Exchange Data. *Int J Mass Spectrom* 302, 59-68
16. Worley, B., and Powers, R. (2013) Multivariate Analysis in Metabolomics. *Curr Metabolomics* 1, 92-107

17. Farshidfar, F., Weljie, A. M., Kopciuk, K., Buie, W. D., MacLean, A., Dixon, E., Sutherland, F. R., Molckovsky, A., Vogel, H. J., and Bathe, O. F. (2012) Serum metabolomic profile as a means to distinguish stage of colorectal cancer. *Genome Med* 4
18. DeBonis, S., Simorre, J. P., Crevel, I., Lebeau, L., Skoufias, D. A., Blangy, A., Ebel, C., Gans, P., Cross, R., Hackney, D. D., Wade, R. H., and Kozielski, F. (2003) Interaction of the mitotic inhibitor monastrol with human kinesin Eg5. *Biochemistry* 42, 338-349
19. Rey, M., Yang, M., Burns, K. M., Yu, Y., Lees-Miller, S. P., and Schriemer, D. C. (2013) Nepenthesin from monkey cups for hydrogen/deuterium exchange mass spectrometry. *Molecular & Cellular Proteomics* 12, 464-472
20. Rey, M., Sarpe, V., Burns, K. M., Buse, J., Baker, C. A., van Dijk, M., Wordeman, L., Bonvin, A. M., and Schriemer, D. C. (2014) Mass spec studio for integrative structural biology. *Structure* 22, 1538-1548
21. Edgington, E. S., and Onghena, P. (2007) *Randomization Tests, 4th Edition*, 4th Ed., Chapman & Hall/CRC
22. UmetricsAB (2015) User guide to SIMCA, Version 14. Umetrics AB
23. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., and Olson, A. J. (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem* 30, 2785-2791
24. Huey, R., Morris, G. M., Olson, A. J., and Goodsell, D. S. (2007) A semiempirical free energy force field with charge-based desolvation. *J Comput Chem* 28, 1145-1152
25. Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., and Olson, A. J. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry* 19, 1639-1662
26. Chalmers, M. J., Busby, S. A., Pascal, B. D., Southern, M. R., and Griffin, P. R. (2007) A two-stage differential hydrogen deuterium exchange method for the rapid characterization of protein/ligand interactions. *J Biomol Tech* 18, 194-204
27. Slysz, G. W., Percy, A. J., and Schriemer, D. C. (2008) Restraining expansion of the peak envelope in H/D exchange-MS and its application in detecting perturbations of protein structure/dynamics. *Anal Chem* 80, 7004-7011
28. Brier, S., Lemaire, D., Debonis, S., Forest, E., and Kozielski, F. (2004) Identification of the protein binding region of S-trityl-L-cysteine, a new potent inhibitor of the mitotic kinesin Eg5. *Biochemistry* 43, 13072-13082
29. Brier, S., Lemaire, D., DeBonis, S., Kozielski, F., and Forest, E. (2006) Use of hydrogen/deuterium exchange mass spectrometry and mutagenesis as a tool to identify the binding region of inhibitors targeting the human mitotic kinesin Eg5. *Rapid Commun Mass Spectrom* 20, 456-462
30. Ward, J. H. (1963) Hierarchical Grouping to Optimize an Objective Function. *J Am Stat Assoc* 58, 236-244
31. Bylesjo, M., Rantalainen, M., Cloarec, O., Nicholson, J. K., Holmes, E., and Trygg, J. (2006) OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *J Chemometr* 20, 341-351
32. Boulesteix, A. L., and Strimmer, K. (2007) Partial least squares: a versatile tool for the analysis of high-dimensional genomic data. *Brief Bioinform* 8, 32-44
33. van den Berg, R. A., Hoefsloot, H. C., Westerhuis, J. A., Smilde, A. K., and van der Werf, M. J. (2006) Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics* 7, 142
34. Vandervoet, H. (1994) Comparing the Predictive Accuracy of Models Using a Simple Randomization Test. *Chemometr Intell Lab* 25, 313-323
35. Talapatra, S. K., Schuttelkopf, A. W., and Kozielski, F. (2012) The structure of the ternary Eg5-ADP-ispinesib complex. *Acta Crystallogr D Biol Crystallogr* 68, 1311-1319

36. Garcia-Saez, I., Yen, T., Wade, R. H., and Kozielski, F. (2004) Crystal structure of the motor domain of the human kinetochore protein CENP-E. *J Mol Biol* 340, 1107-1116
37. Cochran, J. C., Gatial, J. E., 3rd, Kapoor, T. M., and Gilbert, S. P. (2005) Monastrol inhibition of the mitotic kinesin Eg5. *J Biol Chem* 280, 12658-12667
38. Crevel, I. M., Alonso, M. C., and Cross, R. A. (2004) Monastrol stabilises an attached low-friction mode of Eg5. *Curr Biol* 14, R411-412
39. Coureux, P. D., Sweeney, H. L., and Houdusse, A. (2004) Three myosin V structures delineate essential features of chemo-mechanical transduction. *EMBO J* 23, 4527-4537
40. Kull, F. J., and Endow, S. A. (2013) Force generation by kinesin and myosin cytoskeletal motor proteins. *J Cell Sci* 126, 9-19
41. Kim, E. D., Buckley, R., Learman, S., Richard, J., Parke, C., Worthylake, D. K., Wojcik, E. J., Walker, R. A., and Kim, S. (2010) Allosteric drug discrimination is coupled to mechanochemical changes in the kinesin-5 motor core. *Journal of Biological Chemistry* 285, 18650-18661
42. Gorrochategui, E., Casas, J., Porte, C., Lacorte, S., and Tauler, R. (2015) Chemometric strategy for untargeted lipidomics: biomarker detection and identification in stressed human placental cells. *Anal Chim Acta* 854, 20-33
43. Wang, X., Zhang, A., Han, Y., Wang, P., Sun, H., Song, G., Dong, T., Yuan, Y., Yuan, X., Zhang, M., Xie, N., Zhang, H., Dong, H., and Dong, W. (2012) Urine metabolomics analysis for biomarker discovery and detection of jaundice syndrome in patients with liver disease. *Mol Cell Proteomics* 11, 370-380
44. Percy, A. J., and Schriemer, D. C. (2011) MRM methods for high precision shift measurements in H/DX-MS. *International Journal of Mass Spectrometry* 302, 26-35
45. Marx, A., Muller, J., Mandelkow, E. M., Hoenger, A., and Mandelkow, E. (2006) Interaction of kinesin motors, microtubules, and MAPs. *J Muscle Res Cell Motil* 27, 125-137

Figure legends

Figure 1. HX-MS kinetics of STLC binding. Time-point deuteration kinetics of Eg5 incubated with STLC (open circles) and DMSO vehicle (crosses), showing (a) residues L34-L48 which demonstrate no shift upon ligand binding, (b) residues Y125-L140 (α2/L5) and (c) residues D208-K216 (α3) from the induced-fit inhibitor binding pocket, and (d) residues L316-D322 (α5) from the microtubule binding region. Data were collected in quadruplicate, and deuteration data are normalized to the maximum D_2O (45%) used in the labelling experiment (+/- 1 standard deviation).

Figure 2. Structural properties and domain organization for Eg5. (a) Schematic of key structural elements of the Eg5 motor domain (1-386) involved in nucleotide binding (blue), the induced-fit binding pocket (green), and the microtubule binding region (orange) (2, 45). (b) Cartoon representation of Eg5 (PDB 3KEN) bound to STLC (cyan spheres), ADP (magenta spheres) and Mg^{2+} (green sphere). Key structural domains are highlighted according to the colouring scheme in (a). (c) Peptide sequence coverage map of Eg5 digested with *nepenthes* proteases, summarizing 183 deuterated peptides with an average sequence redundancy of 5.6. Residues are coloured based on localized redundancy, according the scale bar shown.

Figure 3. Statistical method for combining HX-MS and functional data. High volume HX-MS screening data is mined by MS Studio to generate deuteration data from all identifiable features, and arrayed with other evidence for an OPLS-DA multivariate statistical analysis performed in Simca v14.1. The deuteration data are input in the X-matrix, and class assignments (here based on IC50 measurements) are input in the Y-matrix. The analysis returns a model that groups ligands according to conformational changes that drive a functional response. The variables (peptides) used to build the statistical model are scored for their contribution to the model, validated, then used in structure-function interpretation of the distinct groupings.

Figure 4. OPLS-DA visualization of composite screening data (IC50 and HX-MS). (a) Scores plot projected for the first principal components (PC1) and first two orthogonal components (PO1 and PO2) of the OPLS-Da model. Non-binding compounds are shown in purple, and binding compounds are shown in yellow. (b) Loadings plots for the corresponding principal components, highlighting the X-variables (peptide features) contributing to the model (labeled with residue numbers and charge state). Peptides of interest covering the L5-binding site are shown in red, peptides covering the microtubule binding region in blue, with the remaining peptides shown in green. The predictive Y-variables shown in grey.

Figure 5. Functional interpretation of class separation. Mapping of significant HX-MS changes induced by binding (a) STLC or (b) Terpendole E. Ligand-induced stabilization is shown in blue, destabilization in red, and no measurable effect in grey. Mass shifts identified as significant if >2% altered deuteration, with $p<0.05$ when compared to all non-ligands. Ligands in green spheres. Models use PDB 4BBG (STLC) and 4BXN (Terpendole E). Relative deuteration for select peptides identified as significant OPLS-DA model contributors in (c,d) the L5 binding pocket, (e,f) the switch II microtubule binding region, and (g) the β-sheet core. Plots show deuteration across all replicates, organized with replicates for each state grouped together, ordered with non-ligands on the left and ligands on the right. Grey dashed lines represent the mean deuteration of non-ligands and control samples. Colored dashed boxes mark ligands with statistically-significant shifts from the controls ($p=0.05$).

Figure 6. Comparative molecular docking. Molecular docking of STLC (cyan) and Terpendole E (magenta) on Eg5 structures shown in green and orange, respectively. Residues Glu116 and Gly117 are highlighted as green spheres.

Tables

Table 1. Classification of Eg5 inhibitors

Ligand	IC50(μM)[1]	HX class[2]	Ligand	IC50(μM)[1]	HX class[2]
Paprotrain (PAP)	>200	Non-binding	S-Trityl-L-Cysteine	0.7	Binding
BTB1 (BTB)	>200	Non-binding	trans-HR22C16 (THR)	0.5	Binding
AZ 3146 (AZ)	>200	Non-binding	EMD 534085 (EMD)	0.6	Binding
GSK923295 (GSK)	>100	Non-binding	Ispinesib (ISP)	0.3	Binding
MPI-0479605 (MPI)	>100	Non-binding	SB743921 (SB7)	0.3	Binding
Monastrol (Mona)	27.3	Binding	ARRY 520 (ARRY)	0.2	Binding
Dimethylenastron	1.7	Binding	Terpendole E (TER)	6.0	Binding
Trans-24 (T24)	0.8	Binding	Cyclobenzaprine (Cyclo)	15.8	Binding
K 858	2.3	Binding	ARQ 621 (ARQ)	1.0	Binding

[1]Inhibition of basal ATPase rate (see supporting information for methodology).
[2]HX classification according to strength of inhibition

Figures

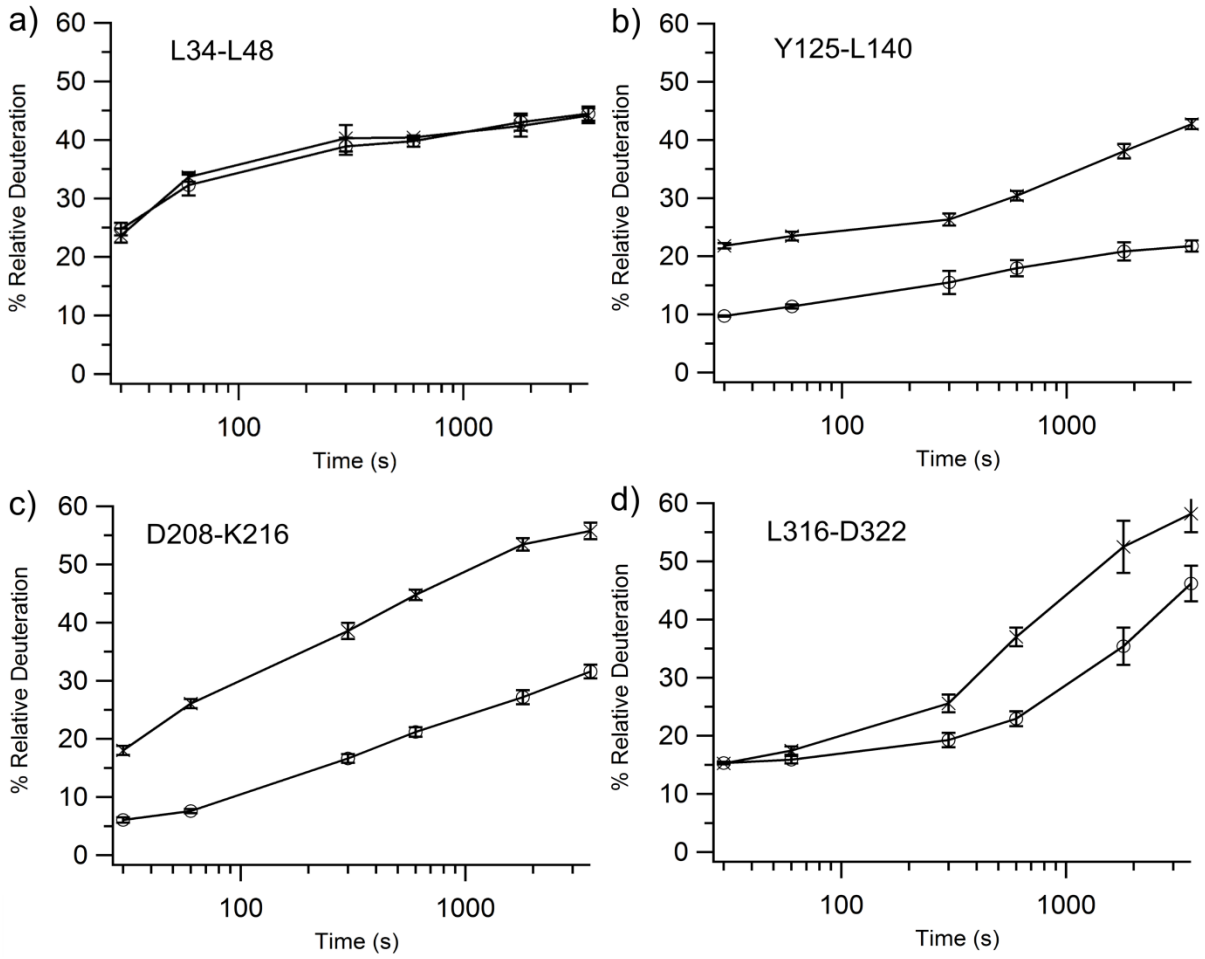

Figure 1

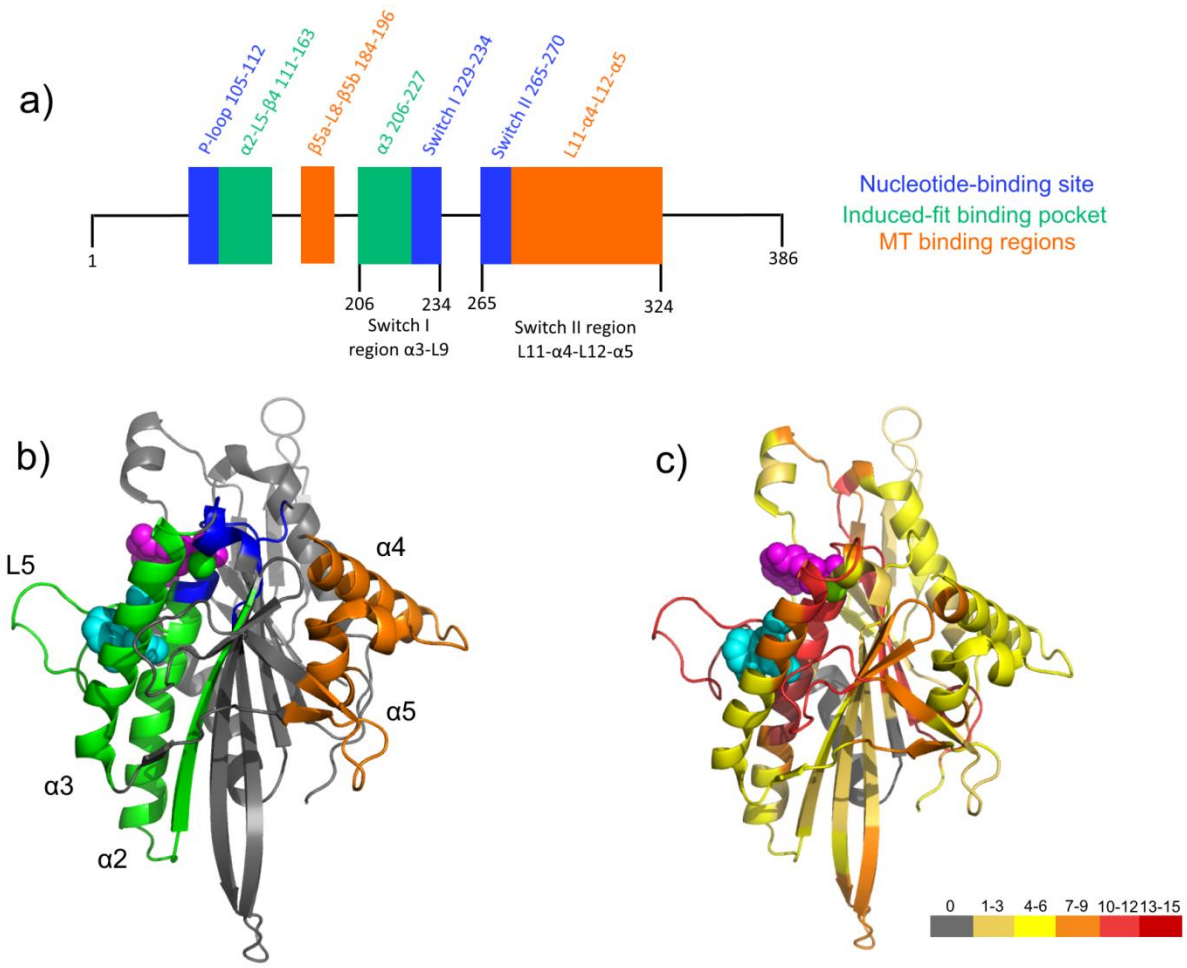

Figure 2

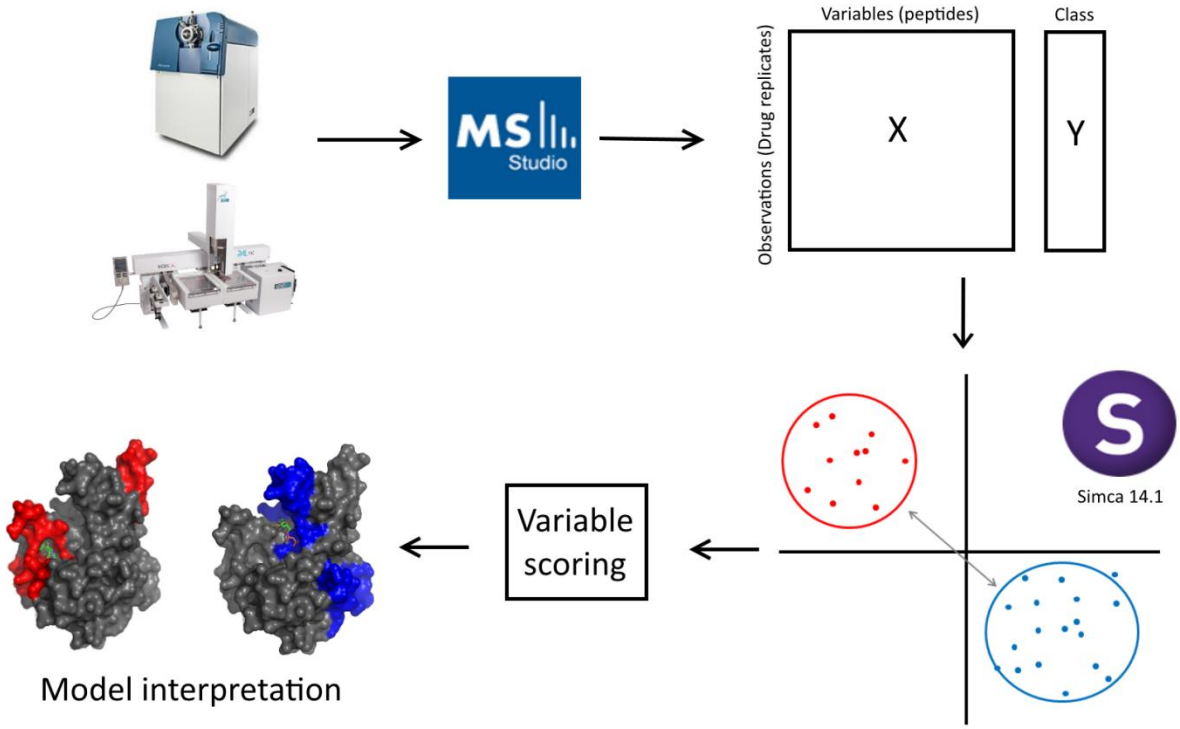

Figure 3

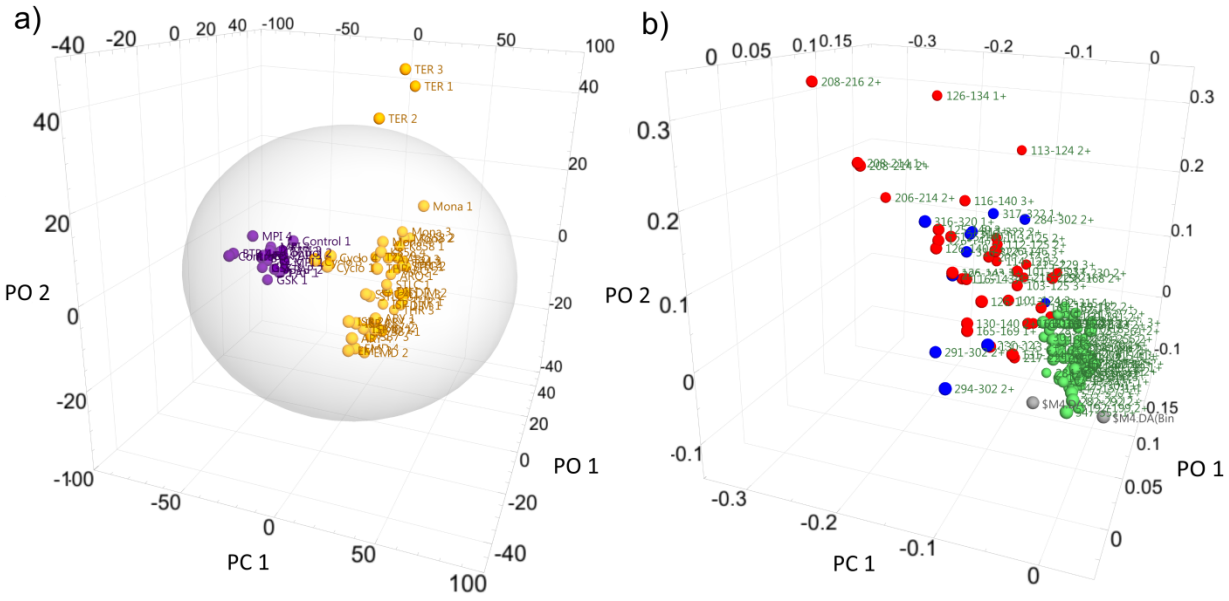

Figure 4

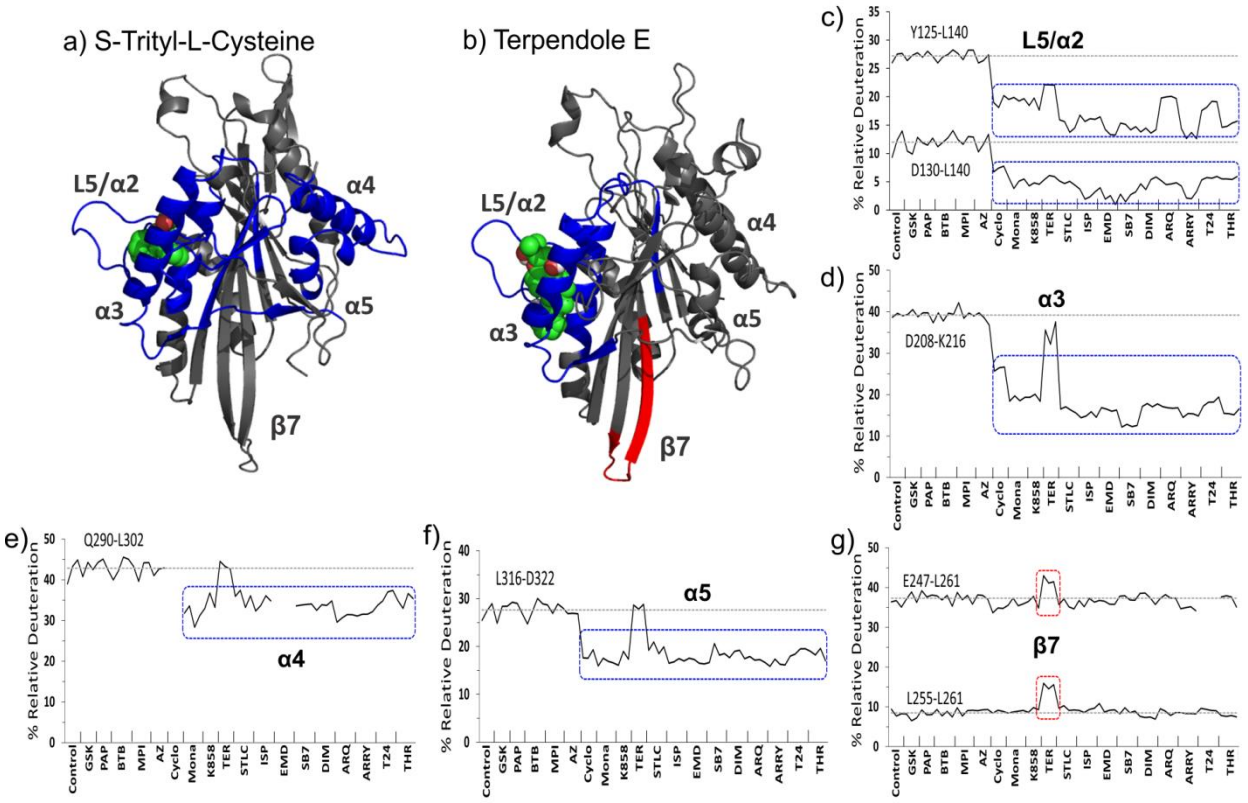
a) S-Trityl-L-Cysteine
b) Terpendole E
L5/α2
α4
α3
α5
β7
c)
Y125-L140
L5/α2
D130-L140
% Relative Deuteration
d)
α3
D208-K216
e)
Q290-L302
α4
f)
L316-D322
α5
g)
E247-L261
β7
L255-L261
Control GSK PAP BTB MPI AZ Cyclo Mona K858 TER STLC ISP EMD SB7 DIM ARQ ARRY T24 THR

Figure 5

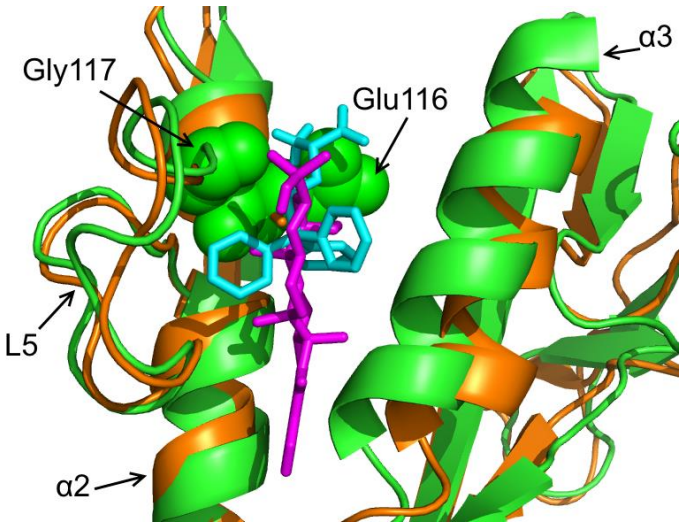

Figure 6