NPR

Integrated profiling methods for identifying the targets of bioactive compounds: MorphoBase and ChemProteoBase

Journal:	*Natural Product Reports*
Manuscript ID	NP-HIG-08-2015-000106.R1
Article Type:	Highlight
Date Submitted by the Author:	17-Feb-2016
Complete List of Authors:	Muroi, Makoto; RIKEN CSRS, Chemical Biology Research Group Futamura, Yushi; RIKEN CSRS, Chemical Biology Research Group Osada, Hiroyuki; RIKEN CSRS, Chemical Biology Research Group

Highlight:

Integrated profiling methods for identifying the targets of bioactive compounds: MorphoBase and ChemProteoBase

Makoto MUROI, Yushi FUTAMURA and Hiroyuki OSADA*

Chemical Biology Research Group, RIKEN CSRS, 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan

*To whom correspondence should be addressed:

Hiroyuki OSADA

Tel: +81-48-4679541; Fax: +81-48-4679542

E-mail: hisyo@riken.jp

Abstract

Many useful compounds from natural products have been discovered through phenotype-based screening. However, the target identification process for compounds is laborious and time-consuming. With the development of new equipment and methodologies for biological analyses, a variety of profiling methods that utilize large sets of experimental data have been established. Here, we highlight the utility of our identification approaches, MorphoBase and ChemProteoBase.

Natural products have yielded compounds with novel biological activity and novel structures that would be difficult to create using synthetic chemistry. Identifying such useful compounds from natural products using phenotype-based screenings is important, as understanding the molecular targets of these compounds can help elucidate their biological functions and potential pharmaceutical applications. For instance, studies on compounds isolated from microbial products, such as lactacystin (a proteasome inhibitor)[1], trichostatin A[2] (a histone deacetylase inhibitor) and rapamycin (a mammalian target of rapamycin inhibitor)[3], have revealed some of the mechanisms involved in cell survival such as protein degradation, epigenetic gene regulation, and signal transduction, respectively. These studies also demonstrated that the targets of these compounds are applicable to cancer chemotherapy, thus leading to the development of anticancer drugs such as bortezomib, vorinostat and temsirolimus, respectively.

Therefore, target identification is one of the most important steps in biological research and drug development. With the development of new equipment and techniques, many approaches for identifying the target of compounds have been discovered. However, it is still a laborious and time-consuming process because there are a wide variety of target candidates in cells, and the strength of the effect and mechanism of action depend on the compound. Furthermore, DNA or lipids can be targets for compounds. Therefore, using multiple approaches is essential for the rapid and accurate identification of targets.

Generally, two fundamentally different approaches exist for identifying molecular targets after novel biologically active compounds have been obtained. One is the direct approach and the other is the indirect approach. The direct approach is based on the detection of a direct interaction between the compound and its target molecule. Traditionally, affinity beads conjugated with compounds via a suitable functional group have been used for obtaining binding proteins[4], and this method is still widely used. Improvements to the technology utilized in this process have been made, such as new methodologies for producing the matrix, linker, and conjugation of compounds[5, 6]. Our laboratory also developed an easy-to-use and universal coupling method that employs a photo-activated crosslinking reaction in order to immobilize of compounds to affinity beads[7, 8]. Moreover, we adapted ribosome display technology for use with pull-down assays to overcome the limitation of low target protein expression in cell lysates[9].

Recently, target identification methods, such as the drug affinity responsive target stability (DARTS)[10] method and the cellular thermal shift assay (CETSA)[11, 12], which use intact ligands without any additive chemical modifications have been developed; these methods utilize the principle of protein stability changes, which occur during interactions with the compound. The molecular targets of many compounds have been identified using these direct approaches.

Although the direct approach may be the most powerful method, it is associated with a few problems such as nonspecific binding. Therefore, indirect approaches can be used to complement the direct approaches. With the development of new equipment and methodologies for biological analyses, a variety of methods that use large experimental data sets have been established, such as cancer cell panel profiling[13, 14], chemical genomic profiling[15, 16], transcriptomic profiling[17], proteomic profiling[18], metabolomic profiling[19], and cell morphology profiling[20-24]. Many of the details of these methods have been described in previous reviews[5, 6]; thus, in this paper, we highlight the utility of our phenotypic profiling approaches, namely MorphoBase[25] and ChemProteoBase[26].

MorphoBase is a quick profiling system that is based on the morphological changes in two cancer cell lines: *src*^{ts}-NRK and HeLa cells. Specifically, we designed a high-content imaging method that segments the cells and quantifies twelve morphological parameters of the cells following treatment with ~200 well-characterized compounds. Furthermore, we developed a data analysis program that incorporates multivariate statistical tools. Principal component analysis is applied to visualize the phenotypic responses, and target prediction of a compound is performed by using two statistical computations: “similarity ranking” as determined by Euclidean distance metrics and “probability scores”, which are based on the z-scores of a test compound relative to the 14 target classes. The simplicity of MorphoBase, which requires only the bright-field cell images and nuclear staining image, makes it an easy and fast way of obtaining reproducible data in comparison with other methods that require multiple staining images.

Our other approach, ChemProteoBase, is a proteomic profiling system that identifies the targets of compounds with proteome analyses that use two-dimensional difference gel electrophoresis (2-D DIGE)[26]. The expression data that were presented by ~300 spots were reproducibly detected in all images of HeLa cells treated with

well-known inhibitors and were successfully classified by cluster analyses according to their mechanism of action. The mechanism of action of a novel compound can be accurately established if it induces a proteomic profile that is similar to a profile induced by a reference compound in the database. To obtain more information about the effects of compounds linking to the gene ontology or biological pathway, about 90% of the spots in ChemProteoBase were identified by mass spectrometry. Spot information and a subset of the expression data for 19 inhibitors are shown on our web site (http://www.npd.riken.jp/csrs/xtra/ProteomePage/). Currently, the ChemProteoBase dataset contains the spot expression data from various experiments on ~150 compounds. Because MorphoBase is a quick profiling system, it is easy to acquire many data-points. On the other hand, in ChemProteoBase, it is usually difficult to increase the number of experiments because of the low throughput and high cost. However, ChemProteoBase may provide additional information such as the proteins in a pathway involving the compound target and post-transcriptional modification. Recent advancements in the technology for mass-based proteomic analysis have made it possible to measure more proteins, thus expanding accuracy of the profiling system. The 2-D DIGE system used ChemProteoBase system is a conventional system; however, it is easy to use and reproducible data for constructing the database.

We have found new compounds with novel biological activity by using various cell-based screening systems. As mentioned above, a variety of methods have been developed for identifying the targets identification of compounds. However, a "gold standard" has not yet been established. When we obtain new compounds, we routinely use MorphoBase and ChemProteoBase as indirect approaches and analyze the compounds' binding proteins by using photo-crosslinking affinity beads as a direct approach (Fig. 1). Based on the results, a validation study is performed. Using these approaches, we successfully identified the molecular target of several new compounds[25, 27-30] (Fig. 2).

Of the compounds we have analyzed, the analysis of pyrrolizilactone in particular demonstrated the potential power of using both MorphoBase and ChemProteoBase[31]. Pyrrolizilactone was originally isolated from the fraction library of a microbial culture broth in our laboratory and demonstrated potent inhibition of cancer cell growth[32]. Two fungal metabolites that were structurally related to pyrrolizilactone had been reported, previously, but their targets remained unclear. Therefore, the target

molecule of pyrrolizilactone was difficult to determine based on its structure. Furthermore, we could not obtain enough pyrrolizilactone to prepare its affinity beads, which is a problem that frequently occurs when analyzing natural products. Therefore, we choose to use our profiling methods, i.e., MorphoBase and ChemProteoBase, to estimate its molecular target. In MorphoBase, pyrrolizilactone was classified into the proteasome class with the score 1.71. However, the DNA class had also about the same score (1.78) and the closest neighbor was a HSP60 inhibitor. Meanwhile, in ChemProteoBase, pyrrolizilactone was classified into a cluster that included proteasome inhibitors, HSP90 inhibitors and rapamycin. The two compounds that were the most similar were proteasome inhibitors, and the HSP27 expression induced by pyrrolizilactone increased at a greater rate than the rate noted in HSP90-inhibitor-treated cells. Because profiling analysis can yield multiple candidates targets for a compound, it is sometimes difficult to reduce the number of targets to only one. Collectively, the results obtained using MorphoBase and ChemProteoBase suggested that the target molecules of pyrrolizilactone were the most likely proteasomes. Next, we confirmed the inhibitory activities of the proteasomes *in vitro*. Interestingly, pyrrolizilactone specifically inhibited the trypsin-like activity of proteasome unlike many other proteasome inhibitors such as MG-132 and lactacystin. Additionally, the intracellular inhibition of proteasomes resulted in the accumulation of ubiquitinated proteins in HeLa cells treated with pyrrolizilactone. Thus, we could promptly identify the target of pyrrolizilactone by using the combined analysis (Fig. 3).

Since the two methods are based on the comparisons to reference compounds, appropriate reference compounds are necessary for predicting the compound target, thus analyzing compounds with multiple targets is difficult. However, cell-based profiling systems, such as our analysis systems, generally require a smaller amount of the test compounds than do direct methods, and the non-specific interactions of compounds in cells might not influenced to the results. In the analysis of pyrrolizilactone, HSP90 inhibitors and proteasome inhibitors were classified into the same group by ChemProteoBase, while MorphoBase clearly separated the two groups. Hence, different approaches might be complement to the other methods.

In conclusion, although many methods are available for identifying the target of a compound, for rapid target identification, it is necessary to use a combination of several methods to take the advantage of the benefits of each method.

Acknowledgements

This study was supported in part by Grants-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Science and Technology Research Promotion Program for Agriculture, Forestry, Fisheries and Food Industry; and the Cross-ministerial Strategic Innovation Promotion Program (SIP), "Technologies for creating next-generation agriculture, forestry and fisheries"; research Program on Hepatitis from the Japan Agency for Medical Research and Development, AMED.

Figure legend

Fig. 1. A schematic demonstrating the experimental process of identifying the target of new compounds.

Fig. 2. Representative compounds that were analyzed using MorphoBase and ChemProteoBase.

Fig. 3. Target identification of pyrrolizilactone by using the combined analysis approaches of MorphoBase and ChemProteoBase.

References

1. G. Fenteany, R. F. Standaert, W. S. Lane, S. Choi, E. J. Corey and S. L. Schreiber, *Science*, 1995, **268**, 726-731.
2. M. Yoshida and S. Horinouchi, *Ann. N. Y. Acad. Sci.*, 1999, **886**, 23-36.
3. E. J. Brown, M. W. Albers, T. B. Shin, K. Ichikawa, C. T. Keith, W. S. Lane and S. L. Schreiber, *Nature*, 1994, **369**, 756-758.
4. M. W. Harding, A. Galat, D. E. Uehling and S. L. Schreiber, *Nature*, 1989, **341**, 758-760.
5. Y. Futamura, M. Muroi and H. Osada, *Mol. BioSyst.*, 2013, **9**, 897-914.
6. S. Ziegler, V. Pries, C. Hedberg and H. Waldmann, *Angew. Chem., Int. Ed. Engl.*, 2013, **52**, 2744-2792.
7. N. Kanoh, K. Honda, S. Simizu, M. Muroi and H. Osada, *Angew. Chem., Int. Ed. Engl.*, 2005, **44**, 3559-3562.
8. N. Kanoh, H. Takayama, K. Honda, T. Moriya, T. Teruya, S. Simizu, H. Osada and Y. Iwabuchi, *Bioconjugate Chem.*, 2010, **21**, 182-186.
9. A. Wada, S. Hara and H. Osada, *Anal. Chem.*, 2014, **86**, 6768-6773.
10. B. Lomenick, R. Hao, N. Jonai, R. M. Chin, M. Aghajan, S. Warburton, J. Wang, R. P. Wu, F. Gomez, J. A. Loo, J. A. Wohlschlegel, T. M. Vondriska, J. Pelletier, H. R. Herschman, J. Clardy, C. F. Clarke and J. Huang, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 21984-21989.
11. R. Jafari, H. Almqvist, H. Axelsson, M. Ignatushchenko, T. Lundback, P. Nordlund and D. Martinez Molina, *Nat. Protoc.*, 2014, **9**, 2100-2122.
12. T. P. Miettinen and M. Bjorklund, *Mol. Pharmaceutics*, 2014, **11**, 4395-4404.
13. R. H. Shoemaker, *Nat. Rev. Cancer*, 2006, **6**, 813-823.
14. S. Dan, M. Okamura, M. Seki, K. Yamazaki, H. Sugita, M. Okui, Y. Mukai, H. Nishimura, R. Asaka, K. Nomura, Y. Ishikawa and T. Yamori, *Cancer Res.*, 2010, **70**, 4982-4994.
15. S. Nishimura, Y. Arita, M. Honda, K. Iwamoto, A. Matsuyama, A. Shirai, H. Kawasaki, H. Kakeya, T. Kobayashi, S. Matsunaga and M. Yoshida, *Nat. Chem. Biol.*, 2010, **6**, 519-526.
16. M. Costanzo, A. Baryshnikova, J. Bellay, Y. Kim, E. D. Spear, C. S. Sevier, H. Ding, J. L. Koh, K. Toufighi, S. Mostafavi, J. Prinz, R. P. St Onge, B. VanderSluis, T. Makhnevych, F. J. Vizeacoumar, S. Alizadeh, S. Bahr, R. L. Brost, Y. Chen, M. Cokol, R. Deshpande, Z. Li, Z. Y. Lin, W. Liang, M. Marback, J. Paw, B. J. San Luis, E. Shuteriqi, A. H. Tong, N. van Dyk, I. M. Wallace, J. A. Whitney, M. T. Weirauch, G. Zhong, H. Zhu, W. A. Houry, M. Brudno, S. Ragibizadeh, B. Papp, C. Pal, F. P. Roth, G. Giaever, C. Nislow, O. G. Troyanskaya, H. Bussey, G. D. Bader, A. C. Gingras, Q. D. Morris, P. M. Kim, C. A. Kaiser, C. L. Myers, B. J. Andrews and C. Boone, *Science*, 2010, **327**, 425-431.
17. J. Lamb, E. D. Crawford, D. Peck, J. W. Modell, I. C. Blat, M. J. Wrobel, J. Lerner, J. P. Brunet, A. Subramanian, K. N. Ross, M. Reich, H. Hieronymus, G. Wei, S. A. Armstrong, S. J. Haggarty, P. A. Clemons, R. Wei, S. A. Carr, E. S. Lander and T. R. Golub, *Science*, 2006, **313**, 1929-1935.
18. M. Bantscheff, D. Eberhard, Y. Abraham, S. Bastuck, M. Boesche, S. Hobson, T. Mathieson, J. Perrin, M. Raida, C. Rau, V. Reader, G. Sweetman, A. Bauer, T. Bouwmeester, C. Hopf, U.

Kruse, G. Neubauer, N. Ramsden, J. Rick, B. Kuster and G. Drewes, *Nat. Biotechnol.*, 2007, **25**, 1035-1044.

19. M. Kitagawa, S. Ikeda, E. Tashiro, T. Soga and M. Imoto, *Chem. Biol.*, 2010, **17**, 989-998.
20. M. Breinig, F. A. Klein, W. Huber and M. Boutros, *Mol. Syst. Biol.*, 2015, **11**, 846.
21. H. Sasaki, I. Takeuchi, M. Okada, R. Sawada, K. Kanie, Y. Kiyota, H. Honda and R. Kato, *PloS one*, 2014, **9**, e93952.
22. C. J. Schulze, W. M. Bray, M. H. Woerhmann, J. Stuart, R. S. Lokey and R. G. Linington, *Chem. Biol.*, 2013, **20**, 285-295.
23. S. M. Gustafsdottir, V. Ljosa, K. L. Sokolnicki, J. Anthony Wilson, D. Walpita, M. M. Kemp, K. Petri Seiler, H. A. Carrel, T. R. Golub, S. L. Schreiber, P. A. Clemons, A. E. Carpenter and A. F. Shamji, *PloS one*, 2013, **8**, e80999.
24. D. W. Young, A. Bender, J. Hoyt, E. McWhinnie, G. W. Chirn, C. Y. Tao, J. A. Tallarico, M. Labow, J. L. Jenkins, T. J. Mitchison and Y. Feng, *Nat. Chem. Biol.*, 2008, **4**, 59-68.
25. Y. Futamura, M. Kawatani, S. Kazami, K. Tanaka, M. Muroi, T. Shimizu, K. Tomita, N. Watanabe and H. Osada, *Chem. Biol.*, 2012, **19**, 1620-1630.
26. M. Muroi, S. Kazami, K. Noda, H. Kondo, H. Takayama, M. Kawatani, T. Usui and H. Osada, *Chem. Biol.*, 2010, **17**, 460-470.
27. M. Kawatani, H. Takayama, M. Muroi, S. Kimura, T. Maekawa and H. Osada, *Chem. Biol.*, 2011, **18**, 743-751.
28. H. Minegishi, Y. Futamura, S. Fukashiro, M. Muroi, M. Kawatani, H. Osada and H. Nakamura, *J. Med. Chem.*, 2015, **58**, 4230-4241.
29. Y. Tarui, T. Chinen, Y. Nagumo, T. Motoyama, T. Hayashi, H. Hirota, M. Muroi, Y. Ishii, H. Kondo, H. Osada and T. Usui, *ChemBioChem*, 2014, **15**, 934-938.
30. T. Kawamura, Y. Kondoh, M. Muroi, M. Kawatani and H. Osada, *Biochem. J.*, 2014, **463**, 53-63.
31. Y. Futamura, M. Kawatani, M. Muroi, H. Aono, T. Nogawa and H. Osada, *ChemBioChem*, 2013, **14**, 2456-2463.
32. T. Nogawa, M. Kawatani, M. Uramoto, A. Okano, H. Aono, Y. Futamura, H. Koshino, S. Takahashi and H. Osada, *J. Antibiot.*, 2013, **66**, 621-623.

Fig. 1

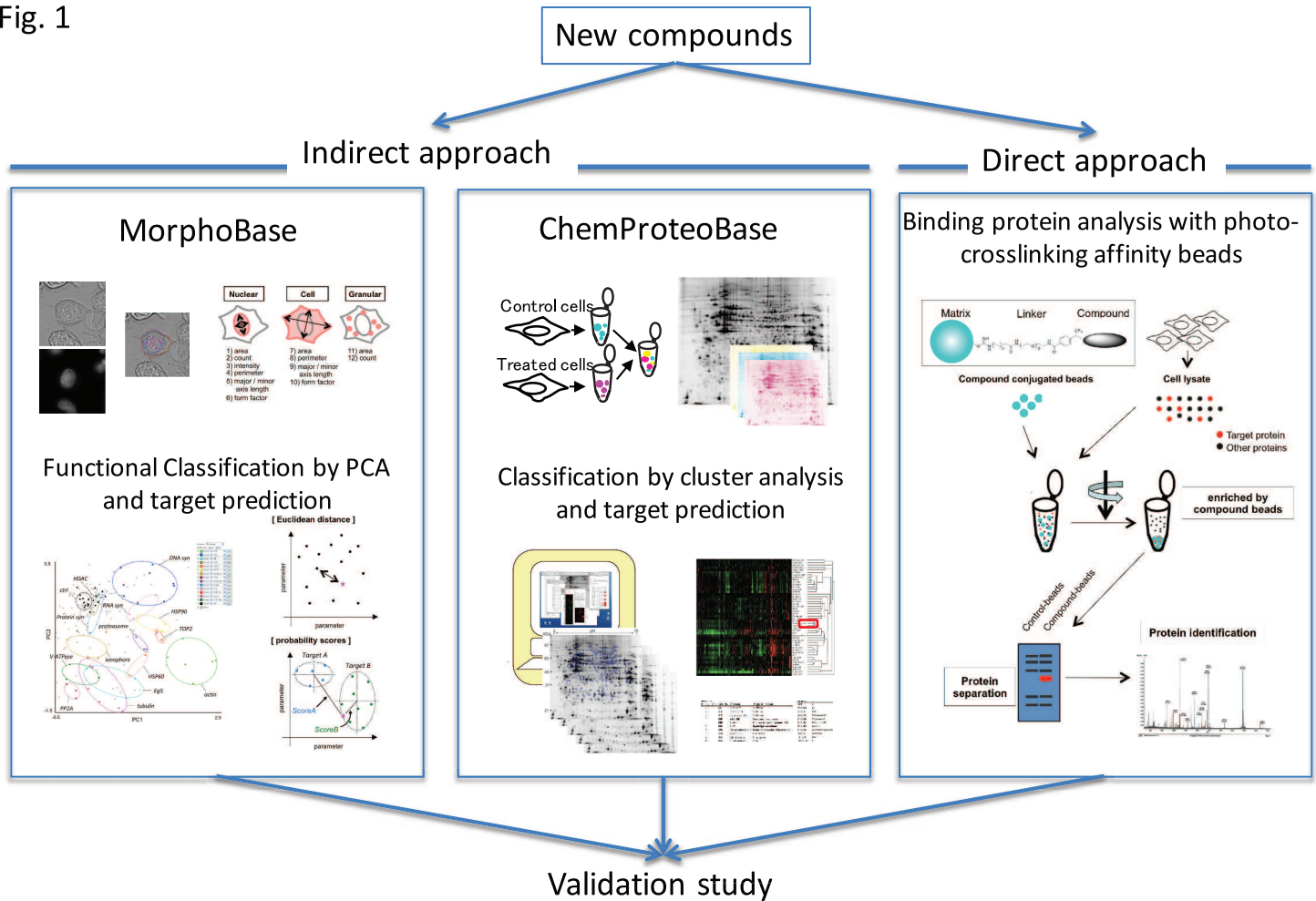
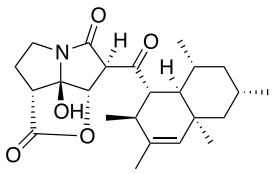
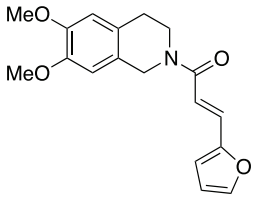

New compounds
Indirect approach
Direct approach
MorphoBase
Nuclear
Cell
Granular
Functional Classification by PCA and target prediction
[Euclidean distance]
[probability scores]
ChemProteoBase
Control cells
Treated cells
Classification by cluster analysis and target prediction
Binding protein analysis with photo-crosslinking affinity beads
Matrix
Linker
Compound
Compound conjugated beads
Cell lysate
Target protein
Other proteins
enriched by compound beads
Control-beads
Compound-beads
Protein identification
Protein separation
Validation study

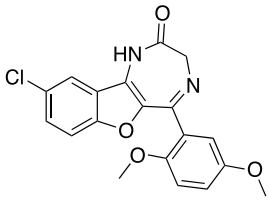
Fig. 2



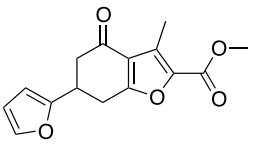
Pyrrolizilactone
(Proteasome inhibitor)



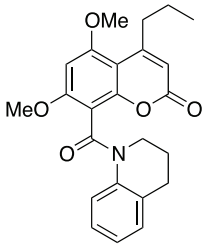
NPD6689
(Tubulin inhibitor)



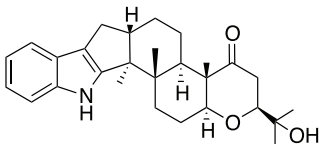
NPD8617
(Tubulin inhibitor)



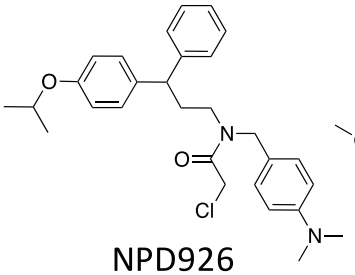
NPD8969
(Tubulin inhibitor)



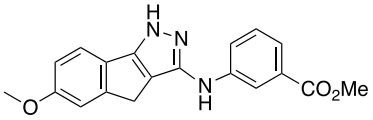
BNS-22
(Top 2 inhibitor)



11-ketopaspaline
(Eg5 inhibitor)

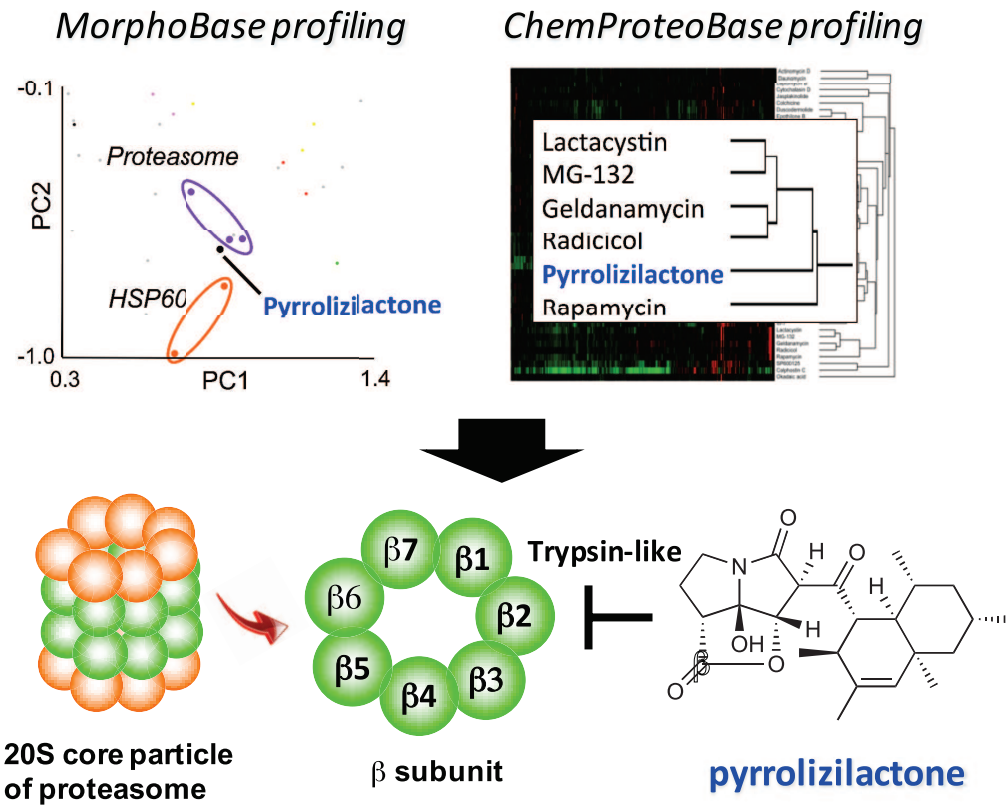


NPD926
(Glutathione inhibitor)



GN39482
(Tubulin inhibitor)

Fig. 3


MorphoBase profiling
ChemProteoBase profiling
-0.1
-1.0
PC2
0.3
PC1
1.4
Proteasome
HSP60
Pyrrolizilactone
Lactacystin
MG-132
Geldanamycin
Radicicol
Pyrrolizilactone
Rapamycin
β7
β1
β6
β2
β5
β4
β3
Trypsin-like
20S core particle of proteasome
β subunit
pyrrolizilactone