ACADEMIA ROMÂNĂ
Revue Roumaine de Chimie
http://web.icf.ro/rrch/

Rev. Roum. Chim.,
2017, *62*(4-5), 425-432

Dedicated to Professor Ionel Haiduc
on the occasion of his 80th anniversary

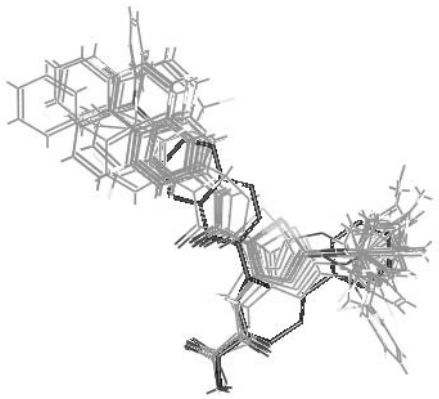
MOLECULAR MODELING STUDIES OF THIAZOLE DERIVATIVES AS PIN1 INHIBITORS**

Daniela VARGA, Luminita CRISAN and Liliana PACUREANU*

Department of Computational Chemistry, Institute of Chemistry of Roumanian Academy, Timisoara, MihaiViteazul Avenue, 24, 300223 Timisoara, Roumania

Received Decembrie 8, 2016

Peptidyl-prolyl cis-trans isomerase NIMA - interacting 1 (PIN1) appeared as a potential therapeutic target in Alzheimer disease, cardiovascular disease and various types of cancer (breast, colon, prostate, esophagus, cervical, etc.). We performed 3D shape similarity search and pharmacophore modeling studies for a series of thiazole derivatives PIN1 inhibitors. A five-point pharmacophore (ADRRR.2) displaying one hydrogen bond acceptor (A), one hydrogen bond donor (D), and three aromatic rings (R) was obtained. Atom-based 3D-QSAR model associated to this hypothesis show significant statistical parameters (training set: R-squared = 0.839, the standard error of estimates, SD = 0.132; test set: correlation coefficient Q-squared = 0.569 and Pearson-R = 0.825. 3D Shape similarity alignment provide information about the possible orientation of these compounds into PIN1 binding site. The pharmacophore hypothesis ADRRR.2, atom-based 3D QSAR and shape similarity search is expected to guide the rational design of novel thiazole derivatives with enhanced affinity towards PIN1.



INTRODUCTION

Peptidyl-prolyl cis-trans isomerase (PIN1) is a human protein that interacts with NIMA (never in mitosis A) kinase. PIN1 belongs to the peptidylprolyl cis-trans isomerase (PPIase) family which catalyze the prolyl isomerization of specific pSer/Thr-Pro motifs that is crucial for cell division in vivo, cell cycle progression, cell signaling, protein folding, assembly and transport, [1-3] cell proliferation.[4,5]

Overexpression of PIN1 was observed in different types of cancer including breast, prostate, colon, cervical, liver and esophagus.[2,6] The previously discovered phosphate-containing small molecules PIN1 inhibitors, have some drawbacks, therefore the rational drug design strategy was repositioned to replace the phosphate group with an bioisostere to improve pharmacokinetic and pharmacodynamic properties.[7] Pharmacophore modeling is a ligand-based approach in fact is one of the best method of choice when the structure of several ligands are known.[8] Hereby, we report a multiple computational ligand-based approach including pharmacophore modeling and 3D shape overly on thiazole PIN1 inhibitors. Pharmacophore modeling was performed via Phase module from Schrödinger, [9-11] based on the 3D structures of 22 thiazole derivatives.[12]

* Corresponding author: pacureanu@acad–icht.tm.edu.ro
** Supporting information on: http://web.icf.ro/rrch or http://revroum.lew.ro

Table 1

The acronym of the thiazole derivatives (Tz-dv) and their biological activity (experimental and predicted) in logarithmic units (pIC_{50})*

No	Name	pIC_{50} exp.	pIC_{50} pred.	No	structure	pIC_{50} exp.	pIC_{50} pred.
1	Tz-dv 1	5.340	5.170	12	Tz-dv 12	5.266	5.180
2*	Tz-dv 2	5.296	5.230	13	Tz-dv 13	4.873	4.850
3	Tz-dv 3	5.260	5.160	14	Tz-dv 14	5.331	5.380
4	Tz-dv 4	5.330	5.330	15*	Tz-dv 15	5.029	5.140
5	Tz-dv 5	5.025	5.070	16*	Tz-dv 16	5.091	5.140
6*	Tz-dv 6	5.286	5.240	17*	Tz-dv 17	4.955	5.140
7	Tz-dv 7	5.502	5.460	18	Tz-dv 18	4.631	4.830
8	Tz-dv 8	5.533	5.550	19	Tz-dv 19	4.83	5.090
9	Tz-dv 9	5.492	5.520	20*	Tz-dv 20	4.676	4.870
10	Tz-dv 10	5.219	5.190	21	Tz-dv 21	4.854	4.810
11*	Tz-dv 11	5.299	5.100	22	Tz-dv 22	4.742	4.640

* test set compounds

To the best of our knowledge this series of thiazole derivatives[12] have not been studied in terms of pharmacophore and shape similarity search. Our goal is to identify the pharmacophore points which further can be used for screening large databases of chemical compounds, to investigate the shape pattern, and identify common outcomes provided by both methodologies.

METHODS

Dataset

A series of novel thiazole derivatives as PIN1 inhibitors synthesized by Zhao *et al.*[12] were used to generate pharmacophore hypotheses with Phase software and 3D shape similarity search using ROCS software.[13-20] Twenty-two compounds whose experimental activity was determined were considered in the present study. The skeleton of the compounds was drawn with the help of SymyxDraw[21] software together with IUPAC names and acronyms are included in Supplementary material.

Pharmacophore modeling was performed with Phase module from Schrödinger package,[22] which includes LigPrep[23] and ConfGen[24,25] to generate: (i) ionization states and tautomers (pH range of 7.2±0.2) and (ii) explore the conformational space for each ligand followed by energy minimization based on the OPLS-2005 force field. The pharmacophore hypotheses were obtained using the most active five compounds and the least potent four compounds as inactives (see Table 1).

For 3D shape and the so called 'color' similarity search the conformational space for each ligand was generated with Omega version 2.3.2 from OpenEye[26-31] package using by default settings RMSD = 0.8 Å, and an energy window of 10 kcal.

Pharmacophore modeling and validation

The default parameters implemented in Develop Pharmacophore Model module of Phase software[9-11] were used in order to generate all possible pharmacophore hypothesis with a root mean squared deviation (RMSD) for distinct conformations higher than 2Å. During pharmacophore hypothesis search we observed no significant increase of the statistics if the number of PLS factors would be higher than four. Phase recognize the following pharmacophore features: acceptor (A), donor (D), hydrophobic (H) and ring R (R).[9-11] A RMSD cut-off of 1.2 Å was used to rank pharmacophore hypotheses in order to detect the best alignment of the active compounds. On the basis of statistical parameters such as: (i) SD - standard deviation of the regression, (ii) R-squared - the coefficient of determination, (iii) F - the ratio of the model variance to the observed activity variance, (iv) P - the significance level of variance

ratio, (v) RMSE - the root-mean-square error in the test set predictions, (vi) Q-squared - the coefficient of determination for the test set predictions, (vii) Pearson-R - the correlation coefficient between predicted and observed activities for the test set, the resulted pharmacophore models and their associated 3D atom-based QSAR models were ranked.

3D Atom-based quantitative structure-activity relationships (QSAR) models corresponding to each pharmacophore hypothesis developed were built in order to validate the pharmacophore hypotheses. In 3D atom-based QSAR a molecule is described as overlapping van der Waals spheres which have the following features: hydrogen-bond donor (D), electron withdrawing (i.e. including hydrogen bond acceptors, W), hydrophobic (H), negative ionic (N), positive ionic (P), and miscellaneous (X).[11] The training set includes 15 randomly selected compounds, whereas the remaining 7 compounds were used to validate the model (test set).

3D-Similarity search

The shape similarity search was carried out with the help of 3D similarity tool Rapid Overly Chemical Structure (ROCS) which is acknowledged as the standard for ligand shape-based virtual screening. ROCS employ 3D Gaussian functions[13-16] to picture the shape of molecules, whereas the chemical force field (color field) measure chemical equivalence. The use of color field and can improve the outcomes provided by shape based superpositions.[13] ImplicitMillsDean color force field which accounts for hydrogen-bond donors, hydrogen-bond acceptors, hydrophobes, anions, cations, and rings is recommended for use by OpenEye.[13] Two similarity metrics are calculated by ROCS Tanimoto and Tversky. Since ROCS performance strongly depends on the selection of query molecule,[32] we choose the bioactive conformation of (2R,4E)-2-[(naphthalen-2-ylcarbonyl)amino]- 5-phenylpent-4-enoic acid extracted from its co-crystal with PIN1 (PDB ID: JZI from 3JYJ co-crystal)[33] as template. Thirteen similarity functions such as TanimotoCombo, ShapeTanimoto, ColorTanimoto, ScaledColor, ComboScore, ColorScore, FitTverskyCombo, FitTversky, FitColorTversky, RefTverskyCombo, RefTversky, RefColorTversky, and Overlap implemented in ROCS were computed.[13-16]

RESULTS AND DISCUSSION

Two hundred twenty-two pharmacophore hypotheses were generated using minimum three and maximum five pharmacophore sites at 4 partial least squares (PLS) factors. The best nine hypotheses internally and externally validated by 3D atom-based QSAR module from Schrödinger package based on statistical parameters R-squared, SD, F, P, RMSE, Q-squared, Pearson-R are presented in Table 2. The pharmacophore hypothesis with the highest values for R-squared, Q-squared and Pearson-R was selected and considered further in the current study (see Figure1 and Table 2). The selected pharmacophore model hypothesis ADRRR.2 (Table 2, in bold) includes three distinct pharmacophore features one hydrogen-bond donor, one hydrogen bond acceptor, and three rings.

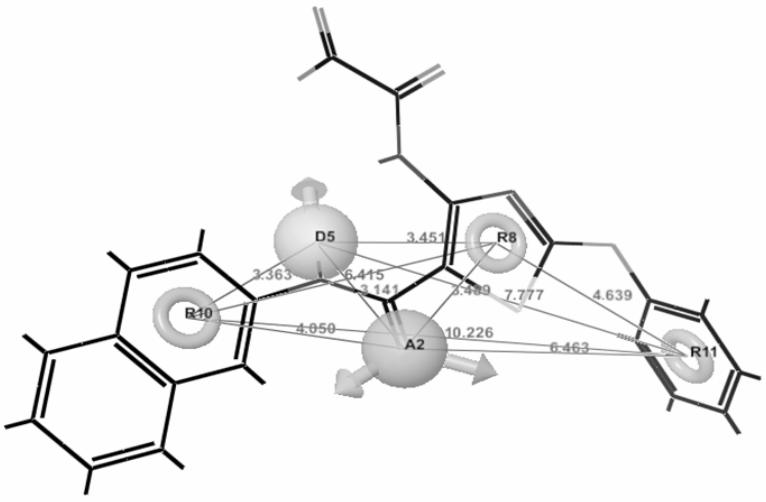


Fig. 1 – The pharmacophore hypothesis ADRRR.2 (acceptor (A2, pink), donor (D5, blue), ring (R8, R10, R11, orange)) aligned to the most active compound **10a**.

Table 2

The statistical parameters obtained for the best QSAR models (the Pearson-R values greater than 0.7)

No.	ID	#	SD	R-squared	F	P	RMSE	Q-squared	Pearson-R
1	AADRR.199	4	0.118	0.887	19.6	0.000101	0.138	0.567	0.781
2	AANRR.34	4	0.112	0.899	22.3	0.000058	0.155	0.453	0.704
3	AARRR.1	4	0.114	0.894	21.2	0.000072	0.151	0.482	0.763
4	AARRR.45	4	0.127	0.871	16.8	0.000195	0.157	0.440	0.772
5	ADNRR.72	4	0.120	0.884	19.0	0.000115	0.156	0.449	0.770
6	**ADRRR.2**	**4**	**0.132**	**0.858**	**15.1**	**0.000302**	**0.138**	**0.569**	**0.825**
7	ANRRR.86	4	0.116	0.891	20.5	0.000082	0.141	0.545	0.760
8	DNRRR.21	4	0.126	0.871	16.9	0.000189	0.147	0.508	0.821
9	DNRRR.25	4	0.124	0.875	17.6	0.000161	0.155	0.456	0.706

Number of factors in the PLS regression model; SD – standard deviation of the regression; R-squared – the coefficient of determination; F – the ratio of the model variance to the observed activity variance; P – the significance level of variance ratio; RMSE – the root-mean-square error in the test set predictions; Q-squared – the squared correlation coefficient for the test set predictions; Pearson-R – the correlation coefficient between predicted and observed activities for the test set.[9-11]

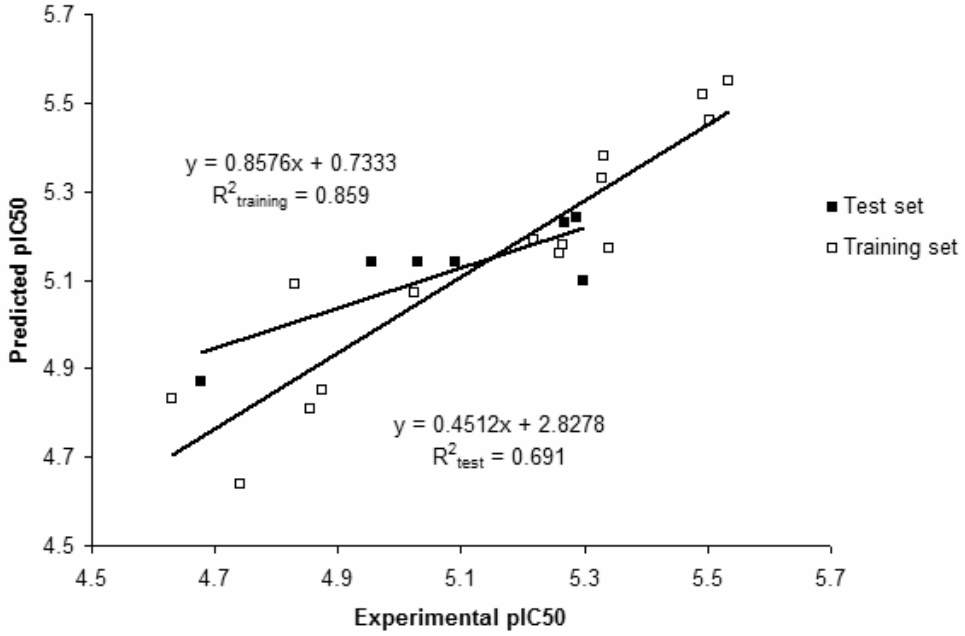


Fig. 2 – The plot of observed versus predicted activities for the ADRRR.2 hypothesis .

The selected atom-based 3D QSAR model at 4 PLS factors ADRRR.2 are statistically reliable, internally and externally (Table 2). The correlation coefficient R-squared higher than 0.85 suggests that the X and Y matrices for the training set are well correlated. The predictive abilities of the model for the test set are acceptable Q-squared (>0.6) and Pearson's R (>0.8). The very low value of the significance level of variance ratio P indicates a statistically significant model (see Table 2). The statistics obtained for the test set illustrates that the selected atom based 3D QSAR model is stable and predictive. The plot of experimental against predicted pIC_{50} for the training and test sets for 3D QSAR model is shown in Figure 2.

In order to examine the atom-based 3D-QSAR model and to investigate the correlation of the most important interactions with biological activity, two property fields (hydrogen bond donor and electron withdrawing) were generated for most active compound (**10a**). Graphical representations for these two properties are displayed in Fig. 3 and Fig. 4. The favorable regions for the biological activity are indicated in blue, while the unfavorable regions are shown in red. The graphical representations suggest that hydrogen bond donor and electron withdrawing fields are situated in the

central part of the molecule resulting that in this region the molecule can make polar interactions. The extreme positions occupied by hydrophobic groups such as naphtyl and phenyl can make hydrophobic interactions with PIN1. Similarly, in the 3JYJ co-crystal of PIN1 with (2R,4E)-2-[(naphthalen-2-ylcarbonyl)amino]- 5-phenylpent-4-enoic acid, the ligand interacts with Ser114 and Lys63 by polar interactions in the central part, whereas naphtyl substituent make hydrophobic interactions with Leu122, Cys113 and phenyl ring undergo pi-pi interactions with His157.[33]

In contrast to 2D similarity methods, 3D similarity methods retrieve compounds that match the shape of the query, although they are topologically different.[35] The superposition of thiazole series of compounds over (2R,4E)-2-[(naphthalen-2-ylcarbonyl)amino]- 5-phenylpent-4-enoic acid suggests a similar orientation into the binding site and analogous hydrogen bond interactions for thiazole derivatives with the receptor.

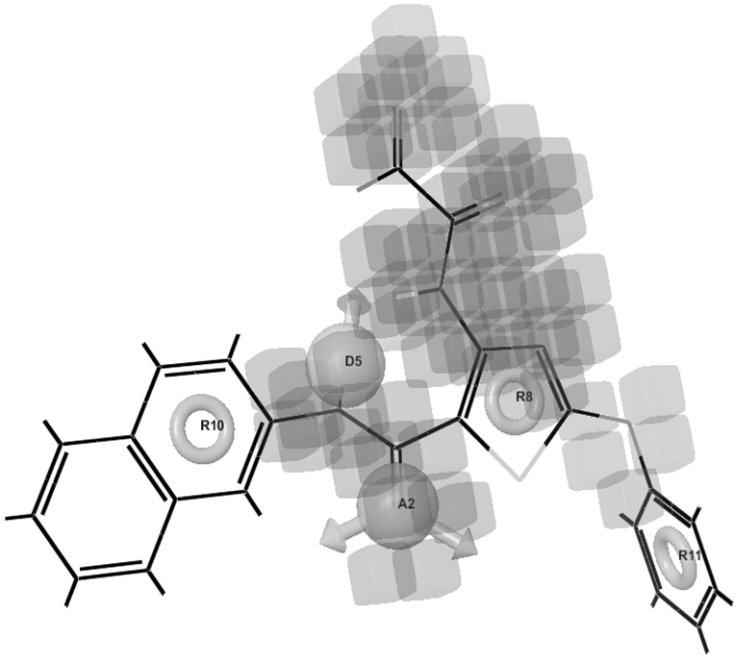


Fig. 3 – Electron withdrawing regions observed on most active compounds **10a**.

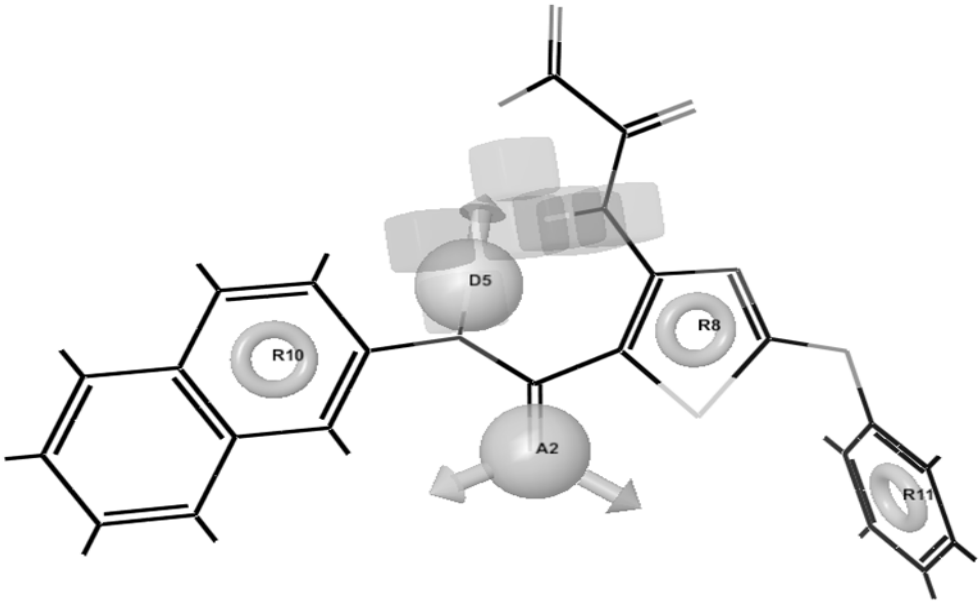


Fig. 4 – Hydrogen bond donor regions observed on most active compound **10**a.

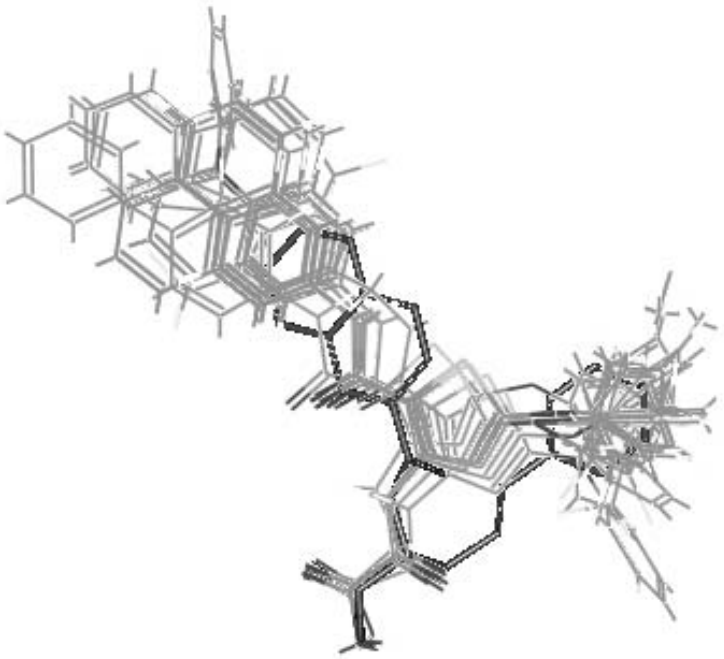


Fig. 5 – Alignment of the 22 hits (line) to the RX ligand (stick).[34]

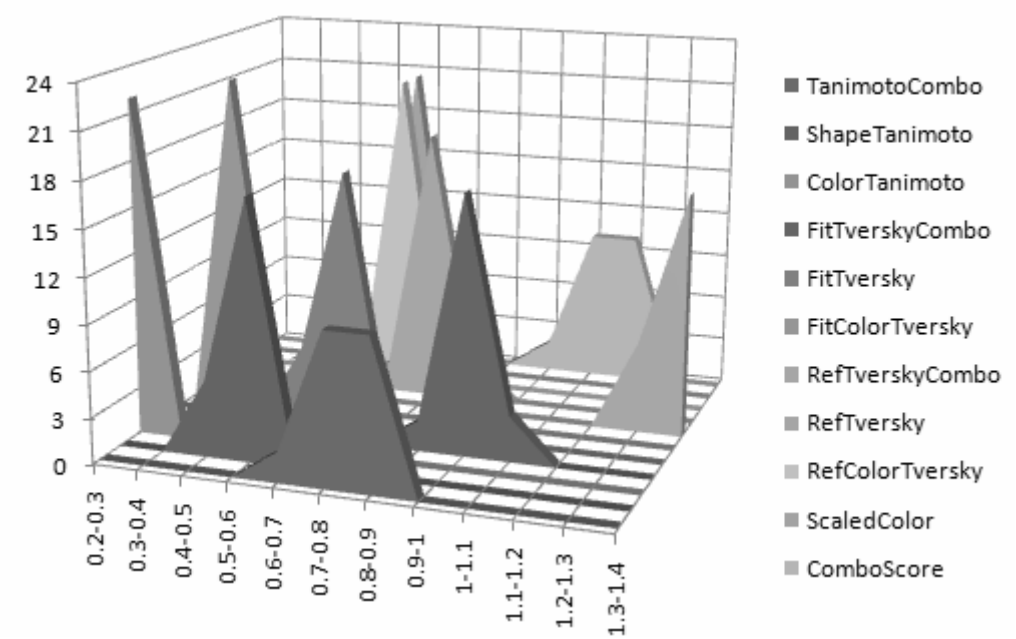


Fig. 6 –The distribution of the ROCS coefficients.

The singular 3D similarity coefficients Color Tanimoto ShapeTanimoto and FitColorTversky have the lowest values. The lowest values of ColorScore indicate pharmacophoric variability inside this series of compounds. Among the combined similarity coefficients (shape and color) RefTverskyCombo, ComboScore, FitTverskyCombo, and TanimotoCombo display the highest values. The ComboScore presents high frequency in the range 1÷1.1 and 1.1÷1.2 displaying ten compounds for each interval. The ScaledColor score presents one high frequency area (twenty-one compounds) in the range 0.6÷0.7, while in the range 0.5÷0.6 only one compound is included (see Fig. 6).

Table 3

The best values of correlation coefficient between 3D coefficients values with fitness score

3D coefficient	TanimotoCombo	ShapeTanimoto	RefTverskyCombo	ComboScore
R-squared	0.5543	0.4482	0.6582	0.6496

As expected the case of combined 3D similarity coefficients (shape and color) RefTverskyCombo and ComboScore we obtained the highest values for the correlation coefficient between 3D similarity and fitness score which illustrate the convergence of both methods in identifying key molecular features. However, significant correlation was obtained for TanimotoCombo which includes a lower contribution of color score. The simple ShapeTanimoto show also a meaningful correlation with fitness score substantiating the previous findings that identified the shape as the basic feature which contribute to molecular recognition.[36]

CONCLUSIONS

The pharmacophore and shape similarity search was employed for a series of 22 thiazole derivatives with experimental inhibitory activities against PIN1. We developed a 222 pharmacophoric hypotheses from which among which nine models displays good statistical results (R-squared >0.8 and R-Pearson >0.7). The best pharmacophore model ADRRR.2, networks three distinct chemical features which includes acceptor (1), donor (1) and ring (3) sites. 3D shape overly with ROCS provided information about the possible orientation and binding mode of thiazole derivatives. The highest values for the correlation coefficient between 3D similarity and fitness score were obtained for RefTverskyCombo and ComboScore which underline the intersection of both methods in identifying crucial molecular features.

Acknowledgements: This project was financially supported by Project 1.2 of the Institute of Chemistry of the Romanian Academy. The authors thank OpenEye Ltd. for providing academic license and to Dr. Ramona Curpăn (Institute of Chemistry Timişoara of Roumanian Academy), for providing access to Schrödinger software acquired through the PN-II-RU-TE-2014-4-422 projects funded by CNCS-UEFISCDI Roumania.

REFERENCES

1. K.P. Lu, S.D. Hanes and T. Hunter, *Nature*, **1996**, *380*, 544-547.
2. E. S. Yeh and A. R. Means, *Nat. Rev. Cancer*, **2007**, *7*, 381-388.
3. K.P. Lu, G. Finn, T. H. Lee and L. K. Nicholson, *Nat. Chem. Biol.*, **2007**, *3*, 619-629.
4. Y. Tong, H. Ying, R. Liu, L. Li, J. Bergholz and Z. Xiao, *Cell Death Dis.*, **2015**, 6, e1640.
5. T. J. Barberi, A. Dunkle, W. Y. He, L. Racioppi and A. R. Means, *PLoS One*, **2012**, 7, e29808.
6. T. Wang, Z. Liu, F. Shi and J. Wang, *Mol Cell. Biochem.*, **2016**, *413*, 179-187.
7. L. Dong, J. Marakovits, X. Hou, C. Guo, S. Greasley, E. Dagostino, R. Ferre, M.C. Johnson, E. Kraynov, J. Thomson, V. Pathak and B.W. Murray, *Bioorg. Med. Chem. Lett*., **2010**, *20*, 2210-2214.
8. P. Kaur, V. Sharma and V. Kumar, *Int. J. Med. Chem.,* **2012**, 2012, id:452325.
9. S.L. Dixon, A.M. Smondyrev, E.H. Knoll, S.N. Rao, D.E. Shaw and R.A. Friesner, *J. Comput. Aided. Mol. Des.,* **2006**, *20*, 647-671.
10. S.L. Dixon, A.M. Smondyrev and S.N. Rao, *Chem. Biol. Drug. Des.*, **2006**, *67*, 370-372.
11. Schrödinger Release 2016-1**:** Phase, Schrödinger, LLC, New York, NY, **2016**.
12. H. Zhao, G. Cui, J. Jin, X. Chen and B. Xu, *Bioorg. Med. Chem*., **2016**, *24*, 5911-5920.
13. OpenEye Scientific Software; ROCS version 3.2.0.4 Santa Fe NM 87508; http://www.eyesopen.com, **2013**.
14. P.C.D. Hawkins, A.G. Skillman and A. Nicholls, *J. Med. Chem.,* **2007**, *50*, 74-82.
15. T.S. Rush, J.A. Grant, L. Mosyak and A. Nicholls, *J. Med. Chem.,* **2005**, *48*, 1489-1495.
16. T. Tuccinardi, G. Ortore, M. Amelia Santos, S.M. Marques, E. Nuti, A. Rosello and A. Martinelli, *J. Chem. Inf. Model.,* **2009**, *49*, 1715-1724.
17. J. Venhorst, S. Nunez, J.W. Terpstra, and C.G. Kruse, *J. Med. Chem.* **2008**, *51*, 3222-3229.
18. R.P. Sheridan, G.B. McGaughey and W.D. Cornell, *J. Comput. Aided Mol. Des.*, **2008**, *22*, 257-265.
19. J.A. Grant, M.A. Gallardo and B. Pickup, *J. Comp. Chem.*, **1996**, *17*, 1653-1666.
20. J.J. Sutherland, R.K. Nandigam, J.A. Erickson and M. Vieth, *J. Chem. Inf. Model.*, **2007**, *47*, 2293-2302.
21. SymyxDraw version 3.2, https://symyx-draw.en.softonic.com, **2009**.
22. Small-Molecule Drug Discovery Suite 2016-1, Schrödinger, LLC, New York, NY, **2016**.
23. Schrödinger Release 2016-1:LigPrep, Schrödinger, LLC, New York, NY, **2016**.
24. K.S. Watts, P. Dalal, R.B. Murphy, W. Sherman, R.A. Friesner and J.C. Shelley, *J.Chem. Inf. Model.*, **2010**, *50*, 534-546.
25. Schrödinger Release 2016-1:ConfGen, Schrödinger, LLC, New York, NY, **2016**.
26. OpenEye Scientific Software; Omega version 2.5.1.4 Santa Fe NM 87508; http://www.eyesopen.com, **2013**.
27. P.C.D. Hawkins, A.G. Skillman, G. L.Warren, B.A. Ellingson and M.T. Stahl, *J. Chem. Inf. Model.*, **2010**, *50*, 572-584.
28. P.C.D. Hawkins and A. Nicholls, *J. Chem. Inf. Model*., **2012**, *52*, 2919-2936.

29. E. Perola and P.S. Charifson., *J. Med. Chem*., **2004**, *47*, 2499-2510.
30. R. Kristam, V.J. Gillet, R.A. Lewis and D. Thorner, *J. Chem. Inf. Model.*, **2005**, *45*, 461-476.
31. J. Boström, J.R. Greenwood and J.J. Gottfries, *J. Mol. Graph. Model*. **2003**, *21*, 449-462.
32. J. Kirchmair, S. Distinto, P. Markt, D. Schuster, G.M. Spitzer, K.R. Liedl and G. Wolber, *J. Chem. Inf. Model*., **2009**, *49*, 678-692.
33. RCSBProteinDataBank. https://www3.rcsb.org/ligand/JZI.
34. Discovery Studio Visualizer-Accelrys, version 2.5, **2009**, San Diego, CA.
35. G.L. Warren, C.W. Andrews, A.M. Capelli A.M., B. Clarke, J. LaLonde, M.H. Lambert, M. Lindvall, N. Nevins, S.F. Semus, S. Senger, G. Tedesco, I.D. Wall, J.M. Woolven, C.E. Peishoff and M.S. Head, *J. Med. Chem*., **2006**, *49*, 5912-5931.
36. S. Kortagere, M.D. Krasowski and S. Ekins, *Trends Pharmacol. Sci.,* **2009**, *30*, 138–147.