

Aminopyrazoles as privileged structures in anticancer drug design - an *in silico* study

G.M. Nitulescu¹, G. Nedelcu¹, A. Buzescu*¹, O.T. Olaru¹

¹“Carol Davila” University of Medicine and Pharmacy, Traian Vuia 6, Bucharest 020956, Romania

Received December 1, 2014, Revised August 7, 2015

Kinases are enzymes with an essential role in cancer progression. Several kinase inhibitors are already used for cancer treatment and extensive efforts are made to develop selective inhibitors for other kinases. Therefore, the assessment of the affinity of some structures for specific molecular targets is mandatory. Our study was focused on aminopyrazoles, as drug-like scaffolds and privileged structures for protein kinases. Molecular descriptors distributions (molecular weight, octanol/water partition coefficient, number of hydrogen bond donors and acceptors, and number of rotatable bonds) were used for characterizing three structural sets containing derivatives of 3-, 4- and 5-aminopyrazole. The analysis of the interaction profiles between protein kinases and specific inhibitors demonstrated their class-selectivity towards protein kinases, suggesting potential antitumor activity. We also showed the importance of the amino group position on the pyrazole ring, indicating a clear difference between aminopyrazole isomers in the drug design process.

Keywords: privileged scaffolds, target selectivity, target affinity, database mining

INTRODUCTION

“Privileged structure” is a concept introduced by Evans in the late 1980s to define the molecular frameworks which are able to provide ligands for more than one type of target, through modification of functional groups [1]. The benzodiazepine scaffold was the first privileged structure cited [2], and thereafter additional similar molecular fragments were revealed. Examples of privileged structures include biphenyls, 1,4-dihydropyridines, bicyclic 6-6 compounds, such as chromones, quinazolines, 2-benzoxazolones, and fused 5-6 ring systems, such as indoles or benzimidazoles [3-5]. Based on Evans definition, the target-family privileged structure concept emerged to describe chemical frameworks which are specific for a single target family and off-target affinities are thus avoided [6].

The aminopyrazole systems prompted enormous research, as they represent valuable templates in drug design. The first aminopyrazole derivative used in antibacterial therapy was sulfaphenazole [7]. MK-0557, a 3-aminopyrazole derivative, was supposed to act by suppressing the appetite-stimulating effects of neuropeptide Y, but it failed to produce clinically meaningful weight loss in humans [8]. Teneligliptin, a 1*H*-pyrazol-5-yl-1-piperazinyl derivative, a dipeptidyl peptidase-4 inhibitor, proved to be useful in the treatment of

type 2 diabetes mellitus [9].

A large panel of aminopyrazole derivatives proved to inhibit various protein kinases with a central role in malignant pathologies. Tozasertib, Doramapimod, Barasertib, AZD1152 and Rebastinib are just some examples of the aminopyrazoles used in anticancer design [10-11].

The aminopyrazole moiety is also used in fused bicyclic compounds, like pyrazolopyrimidine. Zaleplon is a pyrazolopyrimidine that is marketed as a sedative-hypnotic drug in the management of insomnia [12]. Etazolate, Cartazolate and Tracazolate are pyrazolopyridines, structurally related to Zaleplon, exhibiting anxiolytic and anticonvulsant effects [13].

The pharmaceutical impact of the aminopyrazole derivatives has prompted a wide research for developing specific synthetic routes to these compounds [14, 15].

Based on our previous research in the field of antitumor pyrazole-derived compounds [16-18], this study was focused on investigating the target-selectivity patterns of aminopyrazole derivatives by structural and biological *in silico* analysis. The focus of our research was to establish if aminopyrazoles, as drug-like scaffolds, are privileged structures for protein kinases or are promiscuous compounds targeting a plethora of biologic structures.

EXPERIMENTAL

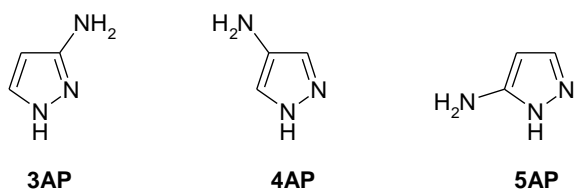
The virtual screening and data mining studies of compounds with certain characteristic substructures

* To whom all correspondence should be sent:
E-mail: zanfirescuanca@yahoo.com

from large chemical databases is an important step in assessing structure-activity relationships [19]. The main resource for obtaining freely-available bioassay data is the PubChem repository provided by the National Center for Biotechnology Information, but the data are not curated and are potentially erroneous [20]. Reaxys is a web-based chemistry database and its Medicinal Chemistry section contains over 5 400 000 substances and more than 26 000 000 bioactivity data points compiled from 320 000 medicinal chemistry publications and patents, fully indexed and normalized [21].

Reaxys database was used to link the screening results to chemical structures in order to identify structure-bioactivity relationships and to study their target promiscuity properties. The database was screened from 10 to 13 November 2014. The access to Reaxys was granted by the UMF Carol Davila's Library.

Pyrazole was the first structure used in the query, and the search filters were "no ring closure" and "no mixtures". The use of the "no ring closure" option removed all fused rings like pyrazolopyrimidine or pyrazolopyridines. The results were again filtered by sub-structure, using 3-aminopyrazole. Next, the compounds containing 3-nitropyrazole were excluded and the non-drug structures were removed using the effect filter. The compounds with insecticidal, pesticidal or herbicidal effects were filtered out and the final set (3AP) was obtained. The same procedure was used in the case of the 4-aminopyrazole set (4AP) and 5-aminopyrazole (5AP).



The resulting structural sets were analyzed regarding their molecular descriptors distribution: the molecular weight (MW), the calculated logarithm of the octanol/water partition coefficient (CLogP), the number of hydrogen bond donors (HBD), the number of hydrogen bond acceptors (HBA), and the number of rotatable bonds (RTB).

The pX querylet was used to filter a particular range of affinities between the compounds and the targets. It represents the logarithmic inverse value of any affinity measure, like inhibitory concentration 50% (IC₅₀), efficacy concentration 50% (EC₅₀), inhibition constant (K_i) or dissociation constant (K_d).

RESULTS AND DISCUSSION

Molecular descriptors distribution

Using the search method described in the "Experimental" section, we identified 3 sets of aminopyrazole compounds, classified by the amino group position as 3AP, 4AP and 5AP. The 3AP set contained 19611 compounds, the 4AP set 13129 and the 5AP set 27058 compounds.

The average value of MWs was 421 g/mol for the 3AP set, 434 g/mol for 4AP, and 466 g/mol for 5AP. The 3AP and 4AP sets have standard deviation close to 80, and for the 5AP set it is around 100 (Figure 1). The high average MW values in comparison with that of the aminopyrazole scaffold provide a higher probability for selective development.

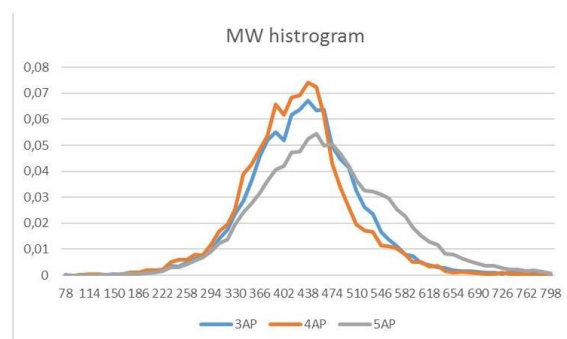


Fig. 1. Histograms of MW distribution in the aminopyrazole derivatives.

The distribution of ClogP values across the three sets of aminopyrazole derivatives took in all cases a bell-shaped curve, but the means differed significantly. In the 3AP group the average ClogP is close to 3, in the 4AP set it is 3.7, the highest being 4.3 in the 5AP set (Figure 2).

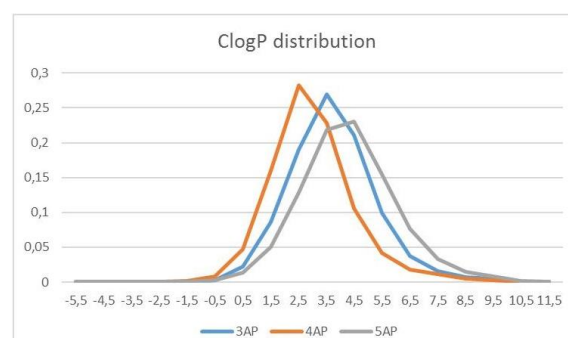


Fig. 2. Histograms of ClogP distribution in the aminopyrazole derivatives.

The analysis of the HBD values distribution (Figure 3) showed a good similarity between the 3AP and 5AP sets, with unimodal bell-shaped curves. Meanwhile, the 4AP had a bimodal distribution. This may indicate different

implications of the hydrogen bonds donors, depending on the type of target.

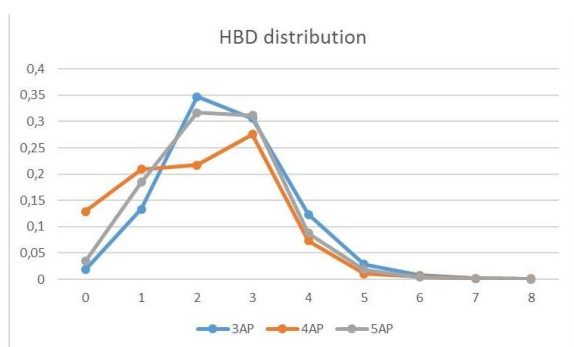


Fig. 3. HBD distribution in the aminopyrazole derivatives.

HBA values were similarly distributed in all aminopyrazole derivatives, having a mean value between 7 and 8 HBD.

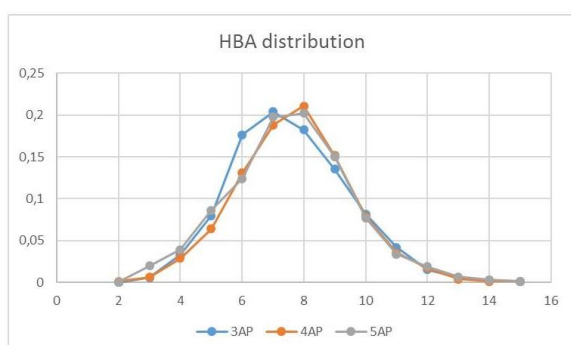


Fig. 4. HBA distribution in the aminopyrazole derivatives.

The RTB descriptor distribution in the three sets indicated a close similarity between the 3AP and 4AP sets (Figure 5). The distribution of the RTB values in the 5AP group differs significantly from those in the 3AP and 4AP sets. The RTB distribution curves resemble those of the MW distribution, with positively skewed data.

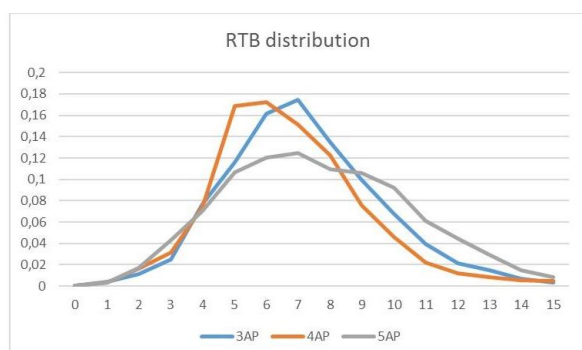


Fig. 5. The distribution of RTB values in the aminopyrazole derivatives.

The number of compounds in the three sets, containing fluorine (F), chlorine (Cl), bromine (Br), iodine (I) or any of these four halogen elements was computed. In each set, the number of compounds containing at least one halogen atom in their structure was close to 54%. The fluorine was found in 34% to 39% of the compounds, and the chlorine atom in 20% to 25% of the compounds, depending on the set.

The analysis of the descriptors frequency in each group of derivatives indicated significant differences between the 3 sets, emphasizing the importance of the amino position on the pyrazole ring. The study also showed that the aminopyrazole scaffold needs a larger framework to become drug-like.

Target-selectivity patterns

The set of 3-aminopyrazole derivatives contained 19611 compounds which can interact with 1858 biological targets; the 4AP set was formed of 13129 compounds which are active against 1343 targets, the 5AP set contained 27058 compounds which can interact with 2237 targets. The Analysis View tool was used to compute the number of compounds acting on each biologic target (Table 1). The results are presented in Table 1 as percentage of the number of compounds in each set.

The occurrence frequency for the top 20 biologic targets interacting with aminopyrazole derivatives clearly indicated a class-selectivity of these compounds for kinases, especially for protein kinases. Exceptions are: cytochrome P450 3A4 (cyp3A4) which interacts with 621 4-aminopyrazoles, (4.3% of the 4AP set); sodium-glucose linked transporter 1 (splt1) interacting with almost 3% of the 3AP group.

The biological targets with the highest frequency of occurrence are classified based on their type and function, as presented in Table 2.

The analysis of the interaction profiles of protein kinase – inhibitors indicated a clear difference between the utility of aminopyrazoles isomers in the drug design process. For example, 3-aminopyrazole derivatives had the highest probability to interact with the Janus kinase family (jak1, jak2, jak3 and trk2), whereas 4-aminopyrazole derivatives had a higher affinity for the cyclin-dependent kinase family. The 5-aminopyrazole group showed selectivity mostly for the mitogen-activated protein kinase 14 (p38a).

Table 1. Biological targets and number (%) of compounds to act on each target. The targets are mentioned in descending order of the compounds number.

No	3AP	%	4AP	%	5AP	%
1	jak2	14.93	jak2	11.76	p38a	17.03
2	aura	14.83	tyk2	10.39	kdr	8.59
3	jak3	12.04	aura	9.51	trka	6.74
4	igf1r	9.39	gsk3beta	8.42	abl	6.16
5	tyk2	9.16	aurb	8.40	c-src	5.43
6	jak1	8.48	kdr	8.26	b-raf	4.94
7	aurb	7.96	pi3k	8.24	aura	4.49
8	c-src	7.56	cdk2/cyclin	7.90	c-raf	3.76
9	erk2	7.29	cdk1/cyclin	7.26	cdk2/cyclin	3.46
10	hk4	6.91	jak3	6.92	tie2	3.33
11	cdk2/cyclin	6.70	pim1	6.30	flt3	3.22
12	trka	6.36	met	5.59	fgfr1	2.71
13	gsk3beta	5.78	jak1	5.03	lck	2.59
14	akt1	5.49	cyp3A4	4.73	erk2	2.55
15	flt3	4.36	cdk5/cyclin	4.67	jak2	2.40
16	trkb	3.96	cdk6/cyclin	4.57	aurb	2.36
17	kdr	3.89	cdk3/cyclin	4.20	igf1r	2.31
18	splt1	2.99	pim2	4.04	jnk2	2.31
19	mapkapk2	2.68	pim3	4.03	jak3	2.29
20	syk	2.64	frap	3.96	jnk3	2.26

Table 2. Biological target classification.

Function	Biological target
Receptor tyrosine kinases	igf1r (<i>insulin-like growth factor 1 receptor</i>), trka (<i>tropomyosin-receptor-kinase A</i>), trkb (<i>tropomyosin-receptor-kinase B</i>), flt3 (<i>Fms-like tyrosine kinase 3</i>), kdr (<i>vascular endothelial growth factor receptor 2</i>), met (<i>hepatocyte growth factor receptor</i>), tie2 (<i>tyrosine kinase with immunoglobulin-like and EGF-like domains</i>), fgfr1 (<i>fibroblast growth factor receptor 1</i>)
Non-receptor tyrosine kinases	jak1 (<i>Janus kinase 1</i>), jak2 (<i>Janus kinase 2</i>), jak3 (<i>Janus kinase 3</i>), tyk2 (<i>tyrosine kinase 2</i>), c-src (<i>c-src kinase</i>), lck (<i>lymphocyte-specific protein tyrosine kinase</i>), syk (<i>spleen tyrosine kinase</i>), abl (<i>Bcr-Abl tyrosine-kinase</i>)
Serine/threonine protein kinases	aura (<i>Aurora A kinase</i>), aurb (<i>Aurora B kinase</i>), cdk1/cyclin (<i>cyclin-dependent kinase 1</i>), cdk2/cyclin (<i>cyclin-dependent kinase 2</i>), cdk3/cyclin (<i>cyclin-dependent kinase 3</i>), cdk5/cyclin (<i>cyclin-dependent kinase 5</i>), cdk6/cyclin (<i>cyclin-dependent kinase 6</i>), erk2 (<i>mitogen-activated protein kinase 1</i>), mapkapk2 (<i>MAP kinase-activated protein kinase 2</i>), jnk2 (<i>mitogen-activated protein kinase 9</i>), p38a (<i>mitogen-activated protein kinase 14</i>), gsk3beta (<i>glycogen synthase kinase 3 beta</i>), akt1 (<i>v-akt murine thymoma viral oncogene homolog 1</i>), pim1 (<i>proto-oncogene serine/threonine-protein kinase 1</i>), pim2 (<i>proto-oncogene serine/threonine-protein kinase 2</i>), pim3 (<i>proto-oncogene serine/threonine-protein kinase 3</i>), frap (<i>mechanistic target of rapamycin</i>), b-raf (<i>serine/threonine-protein kinase B-Raf</i>), c-raf (<i>proto-oncogene c-RAF</i>)
Carbohydrate kinases	hk4 (<i>glucokinase</i>)
Phosphatidylinositol kinases	pi3k (<i>phosphatidylinositol-3-kinases</i>)

Our *in silico* study demonstrated that aminopyrazoles are privileged structures in the design of protein kinases inhibitors. We further investigated whether the aminopyrazole scaffold is necessary for a compound to interact with a particular protein kinase. Therefore we searched the Reaxys database for all the substances interacting with a certain protein kinase at a pX value over 3. We also calculated the number of compounds which contain an aminopyrazole scaffold in their structure (Table 3).

Table 3. The number (%) of compounds containing an aminopyrazole scaffold and the kinases specifically inactivated by these compounds.

No	Target	3AP	4AP	5AP
1	jak2	7.17	3.62	1.33
2	jak3	3.15	2.33	1.45
3	tyk2	1.84	11.12	1.85
4	c-src	4.01	0.02	3.64
5	abl	0.46	0.14	3.44
6	kdr	1.07	1.48	2.32
7	igf1r	10.99	0.69	3.51
8	trka	7.36	0.63	0.24
9	aura	9.51	4.44	3.10
10	aurb	10.79	7.58	2.75
11	p38a	0.23	0.17	13.24
12	gsk3beta	3.39	1.22	1.15

This analysis revealed the importance of the 5-aminopyrazole scaffold in the development of p38a inhibitors, approximately 13% of them containing this framework in their structure. Only a very small percentage contained the 3-aminopyrazole or the 4-aminopyrazole scaffold.

The 3-aminopyrazole scaffold proved to be important in the design of insulin-like growth factor I receptor inhibitors, and of Aurora A and B kinase inhibitors.

The 4-aminopyrazole structure exhibited affinity for the tyrosine kinase 2, almost 11% of its inhibitors sharing this scaffold.

Correlating these data with the molecular descriptors distribution, we found that the aminopyrazole may be important for a certain target, but it needs a larger framework in order to reach a molecular weight in the range of 300 to 600 g/mol, a logP value between 1 and 5, and the proper number of hydrogen bonds donors and acceptors.

CONCLUSIONS

Using data mining techniques, we demonstrated that the aminopyrazole derivatives represent privileged structures for protein kinases, despite their apparent promiscuity. We also emphasized the importance of the amino group position in the pyrazole ring, which dictates the affinity profile for

particular protein kinases. Protein kinases are key players in cancer progression, being involved in uncontrolled growth, survival, neovascularization, metastasis and invasion [21]. By suppressing the activity of particular kinases, the development of cancer cells might be impaired, whilst normal cells are minimally affected [22]. It is therefore expected that the new aminopyrazole derivatives would possess antitumor effects, if properly targeted.

Acknowledgements: This work received financial support through the project entitled "CERO – Career profile: Romanian Researcher", grant number POSDRU/159/1.5/S/135760, cofinanced by the European Social Fund for Sectoral Operational Programme Human Resources Development 2007-2013.

The authors wish to thank Dr. Gina Manda ("Victor Babeş" National Institute for Pathology and Biomedical Sciences, Bucharest) who assisted in the proof-reading of the manuscript.

REFERENCES

1. R.W. DeSimone, K.S. Currie, S.A. Mitchell, J.W. Darrow, D.A. Pippin, *Comb Chem High Throughput Screen.*, **7**, 473 (2004).
2. B.E. Evans, K.E. Rittle, M.G. Bock, R.M. Dipardo, R.M. Freidinger, W.L. Whitter, G.F. Lundell, D.F. Veber, P.S. Anderson, R.S. Chang, V.J. Lotti, D.J. Cerino, T.B. Chen, P.J. Kling, K.A. Kunkel, J.P. Springer, J. Hirshfield, *J Med Chem.*, **31**, 2235 (1988).
3. L. Costantino, D. Barlocco. *Curr Med Chem.*, **13**, 65 (2006).
4. C.D. Duarte, E.J. Barreiro, C.A. Fraga, *Mini Rev Med Chem.*, **7**, 1108 (2007).
5. M.E. Welsch, S.A. Snyder, B.R. Stockwell. *Curr Opin Chem Biol.*, **14**, 347 (2010).
6. D.M. Schnur, M.A. Hermsmeier, A.J. Tebben, *J Med Chem.*, **49**, 2000 (2006).
7. J. G. Susset, *Can Med Assoc J.*, **79**, 992 (1958).
8. D.J. MacNeil, *Curr Top Med Chem.* **7**, 1721 (2007).
9. R. Baetta, A. Corsini. *Drugs*, **71**, 1441 (2011).
10. H. Kumar, D. Saini, S. Jain, N. Jain, *Eur. J. Med. Chem.*, **70**, 248 (2013)
11. D. Pal, S. Saha, S. Singh, *Int J Pharm Pharm Sci.*, **4**, 98 (2012).
12. K. Vanover, R.M. Mangano, J.E. Barrett, *Drug Dev Res.*, **33**, 39 (1994).
13. J.B. Patel, J.B. Malick, A.I. Salama, M.E. Goldberg. *Pharmacol Biochem Behav.*, **23**, 675 (1985).
14. K.A. Kumar, P. Jayaroopa, *Int J PharmTech Res.*, **5**, 1473 (2013).
15. H.F. Anwar, M.H. Elnagdi. *Arkivoc.*, **1**, 198 (2009).
16. V. Anuta, G.M. Nitulescu, C.E. Dinu-Pîrvu, O.T. Olaru, *Molecules*, **19**, 16381 (2014).
17. G.M. Nitulescu, C. Draghici, O.T. Olaru, *Int J Mol Sci.*, **14**, 21805 (2013).

18. G.M. Nitulescu, C. Draghici, A.V. Missir. *Eur J Med Chem.*, **45**, 4914 (2010)
19. Y. Wang, J. Xiao, T.O. Suzek, J. Zhang, J. Wang, S.H. Bryant, *Nucleic Acids Res.*, **37**, 623 (2009).
20. A.C. Schierz. *J. Cheminform.*, **1**, 21 (2009).
21. Reaxys, version 2.18608; Elsevier; 2014; RRN 969209 (accessed, 2014).
22. C.J.Tsai, R. Nussinov, *Semin. Cancer Biol.* **23**, 235 (2013).
23. 21. J. Downward, *Nat. Rev. Cancer*, **3**, 11 (2003).

АМИНОПИРАЗОЛИТЕ КАТО ПРЕДПОЧЕТЕНИ СТРУКТУРИ ПРИ ДИЗАЙНА НА ПРОТИВО-РАКОВИ ЛЕКАРСТВА - *in silico* ИЗСЛЕДВАНЕ

Г.М. Нитулеску¹, Г. Неделку¹, А. Бузеску*¹, О.Т. Олару¹

¹Университет по медицина и фармация „Карол Давила“, Букурещ 020956, Румъния

Постъпила на 1 декември, 2014 г., коригирана на 7 август, 2015 г.

(Резюме)

Киназите са ензими с съществена роля за развитието на раковите заболявания. Някои инхибитори на киназите вече се използват за лечението на рака, като са правени много опити за разработването на селективни инхибитори за други кинази. За тази цел е задължително да се оцени афинитета на структурите на някои целеви молекули. Нашето изследване е фокусирано върху аминокпиразолите с лекарство-подобна структура, предпочетена за протеин-киназа. Разпределението на молекулните дескриптори (молекулно тегло, коефициент на разпределение октанол/вода, броя донори и акцептори на водородни връзки и броя на ротиращите връзки) е използвано за охарактеризирането на три структурни групи, съдържащи производни на 3-, 4- и 5-аминопипразоли. Анализът на профилите на взаимодействие между протеин-киназите и специфичните инхибитори показва клас-селективността спрямо протеин-киназите, внушавайки антитуморно действие. Ние също показахме значението на положението на аминок-групата към пипразоловия пръстен, показвайки ялната разлика между изомерите на аминокпиразолите при дизайна на лекарствените препарати.