The evaluation of exogenous ligands cross-reactivity to a 7TM receptor based on online artificial intelligence engines

Bálint Biró (ABC), László Pitlik (MY-X team, KJU)

Abstract

For every organism from a cell to human it is a general concept that, to survive you have to adapt to the changing environment. To fulfil this requirement, receiving signals coming from the surroundings which are going to generate a response is needed. So, the basic model of signalling consists of a producer, a signal, a target and a transduction pathway which is going to be activated when the special signal has been identified. In cellular signalling processes the acceptor is a molecule called a receptor.

In this study we worked with motilin receptor (MLNR) which is a membrane integrated protein and the natural ligand of it is a peptide hormone called motilin (MLN). This MLN-MLNR system has a pivotal function in smooth muscle contraction and the exact mechanisms in its signalling pathway is relatively well described. MLNR is the member of the 7TM receptor superfamily. There is a huge interest in 7TMs since they are coded by around 1000 genes in mammalian genomes and around 30% of FDA approved drugs act on them.

In life science research it is often necessary to evaluate whether there are other molecules - beside its natural ligand(s) - which can interact with a given receptor. These compounds can be endogens as well as exogens.

Regarding animal welfare and financial considerations, the only possible way to execute such a huge screen is in silico modelling. There are several methods for this topic including virtual screening, machine learning, QSAR analysis etc. Performing these techniques are often a challenging task since they are very computational and time consuming.

Our goal was to try to establish a relatively easy-to-perform-method which can facilitate to explore novel compounds for a given target.

This goal could be approximated based on similarity analyses, where the arbitrary detailed descriptions (attributes) of the potential objects (molecules) could be interpreted in an aggregated way. The optimized analyses proved whether each potential molecule can be seen as the same concerning the complex (context free) similarity index. The multi-layered validated result is atilmotin – which means this molecule can be evaluated as the most similar compared to the molecule of motilin – where the validation based on symmetry of staircase functions like in case of artificial intelligence-oriented term generation processes.

Keywords: similarity, robot-biologist, robot-chemist, Turing-test

# Introduction

## Biological aspects

To adapt to the changing environment a cell has to receive signals and has to send some kind of feedback also. This process is called signalling. In biological signal transduction there are some necessary requirements that have to be fulfilled. The general signal transduction is made up of six steps (Berki et al., 2011). 1 - The production and the release of the extracellular signal. 2 - Signal transport to the target. 3 - Recognition of the signal by a receptor. 4 - Signal transduction. 5 - Activation of the responding mechanism. 6 – Removal of the signal.

Three types of signalling mechanisms can be distinguished by the relationship between the cell which produces the signal and the cell which receives it. In autocrine signalling the producer and receiver cell is the same. In paracrine signalling the producer and the receiver share the same microenvironment while in endocrine signaling the two cells are far away from each other (Iliodromiti et al, 2012; Przybyla & Voldman, 2012; Tse & Wong, 2019) The produced signal can vary from chemical stimuli (molecules such as proteins, lipids or ions, pH etc) to physical ones (radiation, thermal or mechanical stress etc). These wide range of signals make it possible for the cell to maintain its homeostasis (Bernabó et al., 2014). When the interaction between signal and receptor occurs, the inactive receptor is being activated through conformational changes. These changes can lead downstream components of a particular signalling pathway to be activated. The respond mechanism is generally some kind of change in gene expression and/or posttranslational modification of a protein (Hansen & Benenson, 2016).

7 transmembrane receptors (7TM) play a pivotal role in the maintenance of cellular homeostasis. These receptors have 7 transmembrane α-helical domains. Most of them, but not all of them are coupled intracellularly to the heterotrimeric G-proteins (Rajagopal et al., 2005).

Different studies show that in vertebrate animals nearly half of the receptors for transmembrane signalling are rely on 7TMs. 7TMs mediate signalling in neurological, cardiovascular, endocrine, reproductive functions etc (Katritch et al., 2013). Moreover, hundreds of genes have been identified that codes for 7TMs in the human genome and several 7TMs are target of different drugs (Gloriam et al., 2007; Semyonov et al., 2008; Hauser et al., 2018).

The classical way of drug discovery takes years and is very expensive. In the twentieth century with the development of informatics and synthetic chemistry cheminformatics has emerged. This new field uses a great technique, virtual screening which can facilitate the process of drug design. Different types of virtual screening exist but nearly all the methods utilize somehow the structural information about the target (receptor protein) itself or the signal molecule (Kitsuchart, 2012). Generating and working with structural information is very computational. Another problem with these types of information is that only a small portion of known proteins are structurally characterised. This phenomenon is called the „Protein Structure Gap”. This is partially due to the challenging methodology since X-ray crystallography and NMR spectroscopy require experience and different qualifications in protein purification, crystallisation and/or solubilization. In addition, high throughput techniques made it possible to characterize huge amount of protein sequences, but structural techniques could not keep up the pace (Schwede, 2013).

Our research focuses on the motilin receptor (MLNR) and its ligands. MLNR endogen ligand is a polypeptide hormone called motilin (MLN). This MLN-MLNR system has a complex signal transduction pathway. After the formation of receptor/ligand complex, the α subunit of the heterotrimeric G-proteins will activate PLC-β. It catalyses the hydrolysis of PIP2 into IP3 and DAG. The release of IP3 will result a transient Ca2+ efflux. Ca2+/calmodulin complex will be formed and calmodulin will activate the MLC kinase. This enzyme will phosphorylate MLC20 which is the regulatory subunit of myosin muscle protein. Hence smooth muscle contraction can take place (Huang et al., 2005).

There is a great interest around MLNR since it plays important roles in diseases connected to gastro-intestinal motility, food intake and hunger regulation (Ter Beek et al., 2008; Deloose, 2019).

The high-resolution 3D structure of the MLNR is unknown yet.

## AI-aspects

From the point of view of the IT (incl. artificial intelligence) and based on KNUTH’s principle[[1]](#footnote-1), an innovative robot expert (robot-biologist, robot-chemist) is needed being capable of deriving similarity levels of molecules based on their structural data where the basic concept of similarity may not be finetuned even the general problem solving capacity should be involved into the analytical process. The easiness/simplicity should be derivable from the general problem-solving characteristics where the robot expert’s opinion about the aggregated/anti-discriminative (not-learned) similarity ranks of molecules can be compared/evaluated through involving Turing-tests.

The similarity as such is not the ratio of sameness like in case of genetic/genomic analyses. Similarity can be derived based on raw attributes of objects. In this case, objects are the molecules (22 objects – see further figures). Attributes are the physical/chemical/biological/other descriptors (831 attributes – see annex). The aggregated similarities will be derived from the ranking values of the objects in case of each attribute (like Olympic comparison of countries where the descriptors could be the amount of different medallions, etc.). The naive way of aggregation of quasi unlimited impacts of attributes is the subjective scoring where each attribute level can have a kind of score value. In the education, scoring is very simple: the average of the marks will be calculated – it means: each negative change of a mark in case of a subject can be substituted through the same change in the positive direction in case of an other subject. The equivalences are therefore unit-driven although quasi everybody knows that a lower mark in music will be compensated with a better mark in mathematics not so easily as in the reverse modus.

Therefore, AI-based similarity analyses optimize the scoring system in form of a staircase function where each scoring value for a better level in case of every attributes can never be lower.

Based on AI-driven similarities in form of anti-discriminative models, it is possible to create artificial attributes/terms like the term of the similarity as such. The artificial term is a new scale where the objects with normative characteristics will be set into the middle of the scale. It is possible, that each object will have the same similarity value – it means: each object can be evaluated as the same. This is not really possible in the educational evaluation although we all know that everyone in a group could have the same social evaluation grade. In societies where this principle was forgotten a part of the population was simply killed by the others who had just a single attribute where they had better ranking values compared to their victims.

Fortunately, molecules can not be interpreted in a genocide way – but the risk management behind the research activities is also relevant: based on a given amount of financial force fields, it will not lead to the same efficiency if the molecules will be analysed in a randomized way or they will be ranked after the calculated/estimated similarities. In this case, a Turing-test means: the ranking values of a human expert concerning the potential utility of the molecules should be quasi the same as the ranking positions of the molecules generated by a robot-biologist/robot-chemist.

Knuth said – reformulated: Knowledge is, what can be transformed into source code!

# Data assets

The number of the raw descriptors (attributes): 831 (because these data could be identified at all).

The number of the molecules (objects): 22 (because these could be involved into the database at all).

OAMs (6+1):

* <https://miau.my-x.hu/miau/257/motilin.xlsx>,
* <https://miau.my-x.hu/miau/257/motilin_v2.xlsx>

More details can be found in the annexes below. This study is a pilot study for a CRISPR/Cas9 KO animal. We would like to modify the MLN and MLNR genes in rabbit to abolish their biological activity. In this case we are going to be able to further characterize the physiological relevance of the MLN-MLNR system. To study the role of this system we need different exogen substances, drugs that can activate MLNR. Several MLNR agonists and antagonists are commercially available but a global comparison has not been published yet.

Our goal was to perform a complete ranking of the known agonists and antagonists of the MLNR without using structural data, relying only on 2D chemical descriptors.

# Methods

## Biotechnological aspects

22 MLNR ligands were chosen via literature curation. The ligands were alemcinal, ANQ-1125, atilmotin, azithromycin, camicinal, clarithromycin, dirithromycin, EM-523, erythromycin, GHRP, GM-109, idremcinal, KOS-2187, KW-5139, MA-2029, mitemcinal, motilin, oleandomycin, rokitamycin, roxythromycin, RWJ-68023 and telithromycin.

The corresponding SMILES files were exported from ChEMBL, PubChem and Mol-Instinct public databases. ChemDes (Dong et al., 2015) integrated web-based platform was used to generate molecular RDKit and Chemopy 2D chemical descriptors. RDKit descriptors were the following: connectivity, constitutional, MOE-type, CPSA, molecular property and topological. Chemopy descitpros were the following: constitutional, connectivity, Basak, topology, kappa, Burden, E-state, Moran, geary autocorrelation, molecular property, Moreau-Broto autocorrelation, charge, MOE-type.

## Modelling steps

The modelling could be realized based on the online similarity analysis tool (MY-X FREE – COCO/Y0): (for more details: see <https://miau.my-x.hu/myx-free/>, <https://miau.my-x.hu/myx-free/coco/index.html>, <https://miau.my-x.hu/miau/196/My-X%20Team_A5%20fuzet_EN_jav.pdf>)The amount of the raw attributes is 831.

Real involved attribute-positions (583): Here will not be delivered a list with real attribute-names because the 666 attributes having real values could be grouped into 583 different impact-groups where a lot of groups had just one single attribute in the background (545), but a lot of attributes had the same ranking value set in case of the given 22 objects. The maximal number of attributes in a group is 18.

The correlation could not be derived, where the 22 objects had the same raw values. It means: 666 attributes delivered real raw values.

6 attribute-groups were defined (5\*100 attributes and 83 attributes in the last model) because the online tool does let involving more attributes for the general users needing reproducible results.

The 6 partial models got integrated into a closing model.

The modelling philosophy is the anti-discriminative principle where staircase functions are optimized in order to detect non-normative behaviour patterns of the given objects. If each object can be seen as the same, then each object has the same similarity level. If one or more objects are above the similarity norm (hypothetical constant value) then they are the winners and in this case one or more objects should be exiting in the shadow with a similarity potential below the norm.

The anti-discriminative principle needs a raw OAM and the direction for each attribute. The direction means: whether the-more-the-more or the-less-the-more tendency is valid for an attribute concerning the focused similarity potential. In this model-series, the directions are trivial given. The differences in case of each object and each attribute lead to the so called raw data where the direction is monotonous the same for each attribute because each difference to motilin should be minimized in an aggregated similarity model-chain in order to explore the most similar molecules so that motilin as one of the objects should always win the competition compared to itself. Therefore, the real result will be the second position among the winner-molecules.

# Results

The estimation of similarities led to the following details:



Figure Nr.1: Estimations (source: own calculations)

Based on Figure Nr.1, it can be declared, that:

* the partial models (1-2-3-4-5-6) delivered the same winner (“motilin”) and a rel. monotonous ranking value set for the second-best object “atilmotin” (which means: motilin should be the winner because motilin vs motilin is the highest level of similarity and atilmotin can be seen as the most similar molecule compared to the molecule of motilin.)
* the naïve solution (average of the partial model results) delivered the same winners
* the difference between naïve and optimized solutions is moderate (+/- 3 ranking position in case of 22 objects)
* therefore, the randomized partial models can be accepted (c.f. sensitivity analyses)
* the less trivial results (potential winners) are GHRP, dirithromycin, roxithromycin compared to the differences between models (diff<>0)
* erythromycin has a set of attributes leading to a norm-like evaluation
* (the basic models from 1 to 6 used the constant value of 1000000 in order to ensure a higher sensitivity during the optimizing processes concerning to high number of attributes – but in case of the closing model, the constant value is just 1000 because the number of the partial models is low)



Figure Nr.2.: Validation through model-symmetries (source: own calculations)

Figure Nr. 2. demonstrates the specific validation step existing just in similarity analyses where the attributes should have direct and inverse directions – and these two modelling aspects make possible to check whether the expected symmetry of the results (object by object) is given or not?

The values in the figure Nr.2 represents the impact of the partial models level by level (it means in case of each ranking value).

The estimated similarities and the raw attributes have the following (highlighted) correlations (the full table can be read in the annex below):



Figure Nr.3.: The most extreme correlations between raw data and estimated similarities (source: own calculations)

As it can be seen in Figure Nr.3: Just the members of the attribute-group “Chemopy” presented extreme correlations (the group of “RDKit” did not have any attribute with similar characteristics) where the minus sign means: the lower is the raw value the higher will be the similarity.

The weight of the molecules could not lead to extreme correlations – as expected.



Figure Nr.4. Sum of the ranking values of the raw data vs. optimized estimations (source: own calculations)

Figure Nr.4 shows that the sum of the ranking values of the raw data and the optimized estimations deliver a higher difference level (-11 vs +6) than in case of the partial models and the closing model. This kind of verification validates the expected and the explored winners (motilin and atilmotin).

A special sort of the validation is when the anti-winners are interpreted: the anti-winners based on the figure Nr.4 are not clarified enough (see chained ranking values from ranking position 22 towards ranking position 1). Fortunately, the same analysis based on the partial models and the closing model let derive the anti-winners (with the same ranking positions 22:22, or 21:21): MA2029, RWJ\*



Figure Nr.5. The validation of the descriptors using molecular weight correlation.

To test the reliability of the molecular descriptors molecular weight for each molecular descriptor group was compared (Figure Nr.5.). The R2 value of the y=ax+b line equation was 0,89. For four molecules (ANQ-1125, EM-523, KOS-2187, KW-5139) the Chemopy descriptors in the case of Weight described zero values while RDKit descriptors successfully described all the molecules in terms of weight. For RWJ-68023 Chemopy generated negative value for the Weight descriptor. All the other molecules shared the same values in terms of the chosen descriptors in Chemopy and RDKit.



Figure Nr.6. The relaxation of the mean of the ranking values in terms of MA-2029, atilmotin and motilin.

To evaluate the minimum number of descriptors needed to distinguish the similarity, the most likely (atilmotin) and the most unlikely (MA-2029) rankings were plot together with motilin (the standard/bechmark molecule) that served as the fund for the similarity search instead of the number of descriptors (Figure Nr.6.). Around 102 descriptors the average rankings seem to be stabilised. It is obvious that the most likely molecule is much closer to motilin the standard molecule.



Figure Nr.7. The distribution of different ranking values in terms of atilmotin and MA-2029.

The visualization of the different ranking values shows that atilmotin has the same values as motilin in nearly 20% of the calculated descriptors (Figure Nr.7.). More than 30% of atilmotin’s descriptors have the second ranking. MA-2029 has nearly 20% participation in the 17th ranking section. The segregation of the two molecule is straightforward.

# Conclusions

The core logic of the similarity analysis could be used in a successful way to explore similarities between molecules. The raw data were the differences between the focused molecule of motilin and each other competitors concerning each attribute. Therefore, the directions could be defined in a extremely trivial way: similarities need low differences (c.f. the less differences lead to more similarities). The searched molecule is atilmotin. This result could be verified in parallel (mathematical) ways. The robot-biologist and/or robot-chemist can be realized in case of the focused functionality. The human experts could probably not derive a winner with stronger logic…

The validation process with the molecular weight (Figure Nr.5.) indicates some abnormalities. Zero values can be originated from some calculation failure during descriptor generation. This hypothesis can be verified if we check the pipeline to generate the descriptors. For all the descriptor related calculation the inputs were the SMILES files. Even though ANQ-1125, EM-523, KOS-2187, KW-5139 has no Weight values in terms of Chemopy, other values of the mentioned molecules were similar to another ones. Zero values can be observed sometimes through the datasets. There was one negative Weight value (RWJ-68023) which is a physical impossibility. This case brings up some doubt against the reliability of some point of the applied pipeline. Further investigation is needed to go deeper into the causes of the mentioned failures. The impacts of the missing data were tested, and it turned up that the overall results of the present study were not affected by them. Figure Nr.6. shows that approximately 102 descriptors would have been enough to characterise the ranking order of the molecules tested in this study. This quantity can not be extrapolated to other studies since regarding to the Figure Nr.7. adjacent ranking values can have closely related frequency so the more descriptors the better ranking to some extent.

In conclusion this type of overall ranking can be applied to establish the degree of similarity between several molecules. This approach is not as computational as virtual screening or other methods that utilizes structural information of the target or ligand. Careful application of the described pipeline is proposed due to the abnormalities described above. Further investigation is going to be carried out with in vitro muscle contractility or in silico docking studies to cross validate our findings.

# References

Bernabò, N., Barboni, B., & Maccarrone, M. (2014). The biological networks in studying cell signal transduction complexity: The examples of sperm capacitation and of endocannabinoid system. Computational and structural biotechnology journal, 11(18), 11-21.

Hansen, J., & Benenson, Y. (2016). Synthetic biology of cell signaling. Natural Computing, 15(1), 5-13.

Berki, T., Boldizsár, F., Szabó, M., Talabér, G., & Varecza, Z. (2011). Jelátvitel (Orvosi Biotechnológia). Pécsi Tudományegyetem, Pécs.

Tse, L. H., & Wong, Y. H. (2019). GPCRs in Autocrine and Paracrine Regulations. Frontiers in endocrinology, 10.

Przybyla, L., & Voldman, J. (2012). Probing embryonic stem cell autocrine and paracrine signaling using microfluidics. Annual review of analytical chemistry, 5, 293-315.

Kitsuchart, P. (2012). The Review of Virtual Screening Techniques. KMITL Information Technology Journal.

Iliodromiti, Z., Antonakopoulos, N., Sifakis, S., Tsikouras, P., Daniilidis, A., Dafopoulos, K., ... & Vrachnis, N. (2012). Endocrine, paracrine, and autocrine placental mediators in labor. Hormones, 11(4), 397-409.

Semyonov, J., Park, J. I., Chang, C. L., & Hsu, S. Y. T. (2008). GPCR genes are preferentially retained after whole genome duplication. *PLoS One*, *3*(4), e1903.

Gloriam, D. E., Fredriksson, R., & Schiöth, H. B. (2007). The G protein-coupled receptor subset of the rat genome. *BMC genomics*, *8*(1), 338.

Hauser, A. S., Chavali, S., Masuho, I., Jahn, L. J., Martemyanov, K. A., Gloriam, D. E., & Babu, M. M. (2018). Pharmacogenomics of GPCR drug targets. *Cell*, *172*(1-2), 41-54.

Rajagopal, K., Lefkowitz, R. J., & Rockman, H. A. (2005). When 7 transmembrane receptors are not G protein–coupled receptors. *The Journal of clinical investigation*, *115*(11), 2971-2974.

Katritch, V., Cherezov, V., & Stevens, R. C. (2013). Structure-function of the G protein–coupled receptor superfamily. *Annual review of pharmacology and toxicology*, *53*, 531-556.

Schwede, T. (2013). Protein modeling: what happened to the “protein structure gap”?. *Structure*, *21*(9), 1531-1540.

Ter Beek, P. W., Muller, E. S., Van Den Berg, M., Meijer, M. J., Biemond, I., & Lamers, C. B. (2008). Motilin receptor expression in smooth muscle, myenteric plexus, and mucosa of human inflamed and noninflamed intestine. *Inflammatory bowel diseases*, *14*(5), 612-619.

Deloose, E., Verbeure, W., Depoortere, I., & Tack, J. (2019). Motilin: from gastric motility stimulation to hunger signalling. *Nature Reviews Endocrinology*, *15*(4), 238-250.

Huang, J., Zhou, H., Mahavadi, S., Sriwai, W., Lyall, V., & Murthy, K. S. (2005). Signaling pathways mediating gastrointestinal smooth muscle contraction and MLC20 phosphorylation by motilin receptors. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, *288*(1), G23-G31.

Dong, J., Cao, D. S., Miao, H. Y., Liu, S., Deng, B. C., Yun, Y. H., ... & Chen, A. F. (2015). ChemDes: an integrated web-based platform for molecular descriptor and fingerprint computation. *Journal of cheminformatics*, *7*(1), 60.

# Annex

## Used attributes (831)

Smax38 Smax39 Geto Chi5ch Smax35 Smax36 Smax37 Smax30 Smax31 Smax32 Smax33 bcutv8 bcutv9 bcutv6 bcutv7 bcutv4 bcutv5 bcutv2 bcutv3 QNmin bcutv1 SPP Platt Smax34 smile S44 S45 S46 S47 S40 S41 S42 S43 S48 S49 GATSp5 GATSp4 GATSp7 GATSp6 GATSp1 bcute13 GATSp3 GATSp2 GATSp8 Smin78 Chiv6ch Tac Smin72 Smin73 Smin70 Smin71 Smin76 Smin77 Smin74 Smin75 Chiv9 Chiv8 Chiv5 Chiv4 Chiv7 Chiv6 Chiv1 Chiv0 Chiv3 Chiv2 Smin bcutm9 bcutm8 bcutm7 bcutm6 bcutm5 bcutm4 bcutm3 bcutm2 bcutm1 ATSp3 ATSp2 ATSp1 ATSp7 ATSp6 ATSp5 ATSp4 ATSp8 MATSv1 ndonr TPSA1 bcutv15 bcutv16 QNss bcutv10 nhal bcutv12 bcutv13 MATSv2 MATSv5 SIC0 SIC3 SIC2 PC2 PC3 PC1 PC6 PC4 PC5 Shev Shet Smax15 MR nphos S9 S8 GMTIV S3 S2 S1 S7 S6 S5 S4 nta MATSe4 MATSe5 MATSe6 MATSe7 MATSe1 MATSe2 MATSe3 MATSe8 GATSe7 ZM2 ZM1 Smax67 Smax66 Smax65 Smax64 Smax63 Smax62 Smax61 Smax60 Thara Smax69 Smax68 bcutp16 Smax0 bcutp14 bcutp15 bcutp12 bcutp13 bcutp10 bcutp11 Qmin Smin51 QNmax S39 S38 bcute6 S35 S34 S37 S36 S31 S30 S33 S32 Smin25 Smin24 naro Smin26 Smin21 Smin20 Smin23 Smin22 Smin29 Smin28 QCmax Tpc Mnc ATSm8 ncocl ATSm1 ATSm2 ATSm3 ATSm4 ATSm5 ATSm6 ATSm7 Chi3ch Chiv3c S77 Smin27 S76 Smax29 Smax28 Smax23 Smax22 Smax21 Smax20 Smax27 Smax26 Smax25 Smax24 kappa1 kappa3 kappa2 Hato TPSA S79 IDET Smax58 Smax59 Smax56 Smax57 Smax54 Smax55 S75 Smax53 Smax50 Smax51 Smax73 Smin69 Smin68 Shal Smin61 Smin60 Smin63 Smin62 Smin65 Smin64 Smin67 Smin66 MATSv8 ntb SIC5 SIC4 MATSv3 SIC6 SIC1 MATSv4 MATSv7 MATSv6 GATSe8 W GATSe2 GATSe3 GATSe1 GATSe6 VSAEstate10 GATSe4 GATSe5 Ipc Smin45 Rpc Chiv3ch Smin44 bcute7 LogP2 bcute5 bcute4 bcute3 bcute2 bcute1 bcute9 bcute8 knotp ncobr MATSm8 MATSm4 MATSm5 MATSm6 MATSm7 MATSm1 MATSm2 MATSm3 nnitro Chi3c IVDE MRVSA9 MRVSA8 MRVSA5 MRVSA4 MRVSA7 MRVSA6 MRVSA1 MRVSA0 MRVSA3 MRVSA2 Qmax Smax Xu Smax12 Smax13 Smax10 Smax11 Smax16 Smax17 Smax14 ndb Smax18 Smax19 DZ PEOEVSA12 PEOEVSA11 PEOEVSA10 nsulph DS QCmin IDE Chi4pc S22 S23 S20 S21 S26 S27 S24 S25 S28 S29 Smin10 Smin11 Pol Smin13 Smin14 Smin15 Smin16 Smin17 Smin18 Smin19 Chi0 Chi6ch Save Chi4c Rnc Smax8 Smax9 naccr Smax1 Smax2 Smax3 Smax4 Smax5 Smax6 Smax7 Qindex bcutv14 slogPVSA10 slogPVSA11 Smin57 bcutv11 S50 Chi8 Chi9 Chi2 Chi3 LDI Chi1 Chi6 Chi7 Chi4 Chi5 Smin32 PEOEVSA13 Chiv4c LogP Smax49 Smax48 S68 S69 Smax45 Smax44 Smax47 S65 Smax41 Smax40 Smax43 Smax42 Smin54 Smin55 Smin56 bcute14 Smin50 bcute12 Smin52 Smin53 Smin58 Smin59 dchi4 dchi3 dchi2 dchi1 dchi0 EstateVSA4 EstateVSA5 EstateVSA6 EstateVSA7 EstateVSA0 EstateVSA1 EstateVSA2 EstateVSA3 S66 S67 S64 ATSv1 VSAEstate0 ATSv8 S19 S18 VSAEstate4 VSAEstate5 VSAEstate6 BertzCT S13 S12 S11 S10 S17 S16 S15 S14 GATSm2 GATSm3 GATSm1 GATSm6 S60 GATSm4 GATSm5 GATSm8 S61 bcute16 bcute15 Chiv5ch GMTI nring QHmax bcute11 bcute10 CIC5 CIC4 CIC6 CIC1 CIC0 CIC3 CIC2 bcutp8 bcutp9 bcutp4 bcutp5 bcutp6 bcutp7 bcutp1 bcutp2 bcutp3 ncarb kappam1 TIAC kappam3 kappam2 noxy Chiv4pc PEOEVSA3 PEOEVSA2 PEOEVSA1 PEOEVSA0 PEOEVSA7 PEOEVSA6 PEOEVSA5 PEOEVSA4 Qass PEOEVSA9 PEOEVSA8 phi QOmin GATSv3 GATSv2 GATSv1 GATSv7 GATSv6 GATSv5 GATSv4 GATSv8 QCss Chiv10 slogPVSA2 slogPVSA3 slogPVSA0 slogPVSA1 slogPVSA6 slogPVSA7 slogPVSA4 slogPVSA5 Chi4ch slogPVSA8 slogPVSA9 S57 S56 S55 S54 S53 S52 S51 Gravto EstateVSA8 Sitov S59 S58 Mpc LabuteASA radiust MATSp8 MATSp7 MATSp6 MATSp5 MATSp4 MATSp3 MATSp2 MATSp1 nhet nhev AWeight Chiv4ch ncoi QOmax ncof Hy Smin0 QHss Getov VSAEstate1 VSAEstate2 VSAEstate3 knotpv Weight VSAEstate7 EstateVSA10 VSAEstate8 UI VSAEstate9 mChi1 ATSv3 ATSv2 ATSv5 nsb petitjeant QOss ATSv4 Smax78 ATSv7 Smax70 Smax71 Smax72 ATSv6 Smax74 Smax75 Smax76 Smax77 Smin43 Smin42 Smin41 Smin40 Smin47 Smin46 diametert Scar Smin49 Smin48 ISIZ GATSm7 EstateVSA9 J Tnc QHmin S78 Tigdi Smin36 Smin37 Smin34 Smin35 S62 Smin33 Smin30 Smin31 nhyd Smin38 Smin39 ATSe1 ATSe2 ATSe3 ATSe4 ATSe5 ATSe6 ATSe7 ATSe8 S71 S63 S70 S73 S72 Hatov Smax52 Smin12 S74 Mac Arto IC3 IC2 IC1 IC0 IC6 IC5 IC4 Smin6 Smin7 Smin4 Smin5 Smin2 Smin3 Tsch Smin1 Smin8 Smin9 AW MZM2 MZM1 nrot Sito Smax46 Chi10 bcutm16 bcutm15 bcutm14 bcutm13 bcutm12 bcutm11 bcutm10 BalabanJ BertzCT Chi0 Chi0n Chi0v Chi1 Chi1n Chi1v Chi2n Chi2v Chi3n Chi3v Chi4n Chi4v EState\_VSA1 EState\_VSA10 EState\_VSA11 EState\_VSA2 EState\_VSA3 EState\_VSA4 EState\_VSA5 EState\_VSA6 EState\_VSA7 EState\_VSA8 EState\_VSA9 ExactMolWt FractionCSP3 HallKierAlpha HeavyAtomCount HeavyAtomMolWt Ipc Kappa1 Kappa2 Kappa3 LabuteASA MaxAbsEStateIndex MaxAbsPartialCharge MaxEStateIndex MaxPartialCharge MinAbsEStateIndex MinAbsPartialCharge MinEStateIndex MinPartialCharge MolLogP MolMR MolWt NHOHCount NOCount NumAliphaticCarbocycles NumAliphaticHeterocycles NumAliphaticRings NumAromaticCarbocycles NumAromaticHeterocycles NumAromaticRings NumHAcceptors NumHDonors NumHeteroatoms NumRadicalElectrons NumRotatableBonds NumSaturatedCarbocycles NumSaturatedHeterocycles NumSaturatedRings NumValenceElectrons PEOE\_VSA1 PEOE\_VSA10 PEOE\_VSA11 PEOE\_VSA12 PEOE\_VSA13 PEOE\_VSA14 PEOE\_VSA2 PEOE\_VSA3 PEOE\_VSA4 PEOE\_VSA5 PEOE\_VSA6 PEOE\_VSA7 PEOE\_VSA8 PEOE\_VSA9 RingCount SMR\_VSA1 SMR\_VSA10 SMR\_VSA2 SMR\_VSA3 SMR\_VSA4 SMR\_VSA5 SMR\_VSA6 SMR\_VSA7 SMR\_VSA8 SMR\_VSA9 SlogP\_VSA1 SlogP\_VSA10 SlogP\_VSA11 SlogP\_VSA12 SlogP\_VSA2 SlogP\_VSA3 SlogP\_VSA4 SlogP\_VSA5 SlogP\_VSA6 SlogP\_VSA7 SlogP\_VSA8 SlogP\_VSA9 TPSA VSA\_EState1 VSA\_EState10 VSA\_EState2 VSA\_EState3 VSA\_EState4 VSA\_EState5 VSA\_EState6 VSA\_EState7 VSA\_EState8 VSA\_EState9 fr\_Al\_COO fr\_Al\_OH fr\_Al\_OH\_noTert fr\_ArN fr\_Ar\_COO fr\_Ar\_N fr\_Ar\_NH fr\_Ar\_OH fr\_COO fr\_COO2 fr\_C\_O fr\_C\_O\_noCOO fr\_C\_S fr\_HOCCN fr\_Imine fr\_NH0 fr\_NH1 fr\_NH2 fr\_N\_O fr\_Ndealkylation1 fr\_Ndealkylation2 fr\_Nhpyrrole fr\_SH fr\_aldehyde fr\_alkyl\_carbamate fr\_alkyl\_halide fr\_allylic\_oxid fr\_amide fr\_amidine fr\_aniline fr\_aryl\_methyl fr\_azide fr\_azo fr\_barbitur fr\_benzene fr\_benzodiazepine fr\_bicyclic fr\_diazo fr\_dihydropyridine fr\_epoxide fr\_ester fr\_ether fr\_furan fr\_guanido fr\_halogen fr\_hdrzine fr\_hdrzone fr\_imidazole fr\_imide fr\_isocyan fr\_isothiocyan fr\_ketone fr\_ketone\_Topliss fr\_lactam fr\_lactone fr\_methoxy fr\_morpholine fr\_nitrile fr\_nitro fr\_nitro\_arom fr\_nitro\_arom\_nonortho fr\_nitroso fr\_oxazole fr\_oxime fr\_para\_hydroxylation fr\_phenol fr\_phenol\_noOrthoHbond fr\_phos\_acid fr\_phos\_ester fr\_piperdine fr\_piperzine fr\_priamide fr\_prisulfonamd fr\_pyridine fr\_quatN fr\_sulfide fr\_sulfonamd fr\_sulfone fr\_term\_acetylene fr\_tetrazole fr\_thiazole fr\_thiocyan fr\_thiophene fr\_unbrch\_alkane fr\_urea mol\_smile

## Used objects (20)

alemcinal, ANQ-1125, atilmotin, azithromycin, camicinal, clarithromycin, dirithromycin, EM-523, erythromycin, GHRP, GM-109, idremcinal, KOS-2187, KW-5139, MA-2029, mitemcinal, motilin, oleandomycin, rokitamycin, roxythromycin, RWJ-68023, telithromycin,

## Correlations between estimated similarities and raw attributes

The correlation could not be derived, where the 22 objects had the same raw values. It means: 666 attributes delivered real raw values.

|  |  |  |
| --- | --- | --- |
| correlations | groups of attributes | attributes |
| #ZÉRÓOSZTÓ! | Chemopy | Smax38 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax39 |
| -0.29492369 | Chemopy | Geto |
| -0.368206689 | Chemopy | Chi5ch |
| -0.296798573 | Chemopy | Smax35 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax36 |
| -0.282237793 | Chemopy | Smax37 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax30 |
| 0.04959462 | Chemopy | Smax31 |
| 0.238389758 | Chemopy | Smax32 |
| -0.219411012 | Chemopy | Smax33 |
| -0.672543524 | Chemopy | bcutv8 |
| -0.275942462 | Chemopy | bcutv9 |
| -0.605729525 | Chemopy | bcutv6 |
| -0.64463655 | Chemopy | bcutv7 |
| -0.592601065 | Chemopy | bcutv4 |
| -0.559511954 | Chemopy | bcutv5 |
| -0.224250206 | Chemopy | bcutv2 |
| -0.578492604 | Chemopy | bcutv3 |
| -0.34958233 | Chemopy | QNmin |
| -0.404751202 | Chemopy | bcutv1 |
| -0.26938293 | Chemopy | SPP |
| -0.739709042 | Chemopy | Platt |
| -0.575635184 | Chemopy | Smax34 |
| #ÉRTÉK! | Chemopy | smile |
| #ZÉRÓOSZTÓ! | Chemopy | S44 |
| #ZÉRÓOSZTÓ! | Chemopy | S45 |
| #ZÉRÓOSZTÓ! | Chemopy | S46 |
| #ZÉRÓOSZTÓ! | Chemopy | S47 |
| #ZÉRÓOSZTÓ! | Chemopy | S40 |
| #ZÉRÓOSZTÓ! | Chemopy | S41 |
| #ZÉRÓOSZTÓ! | Chemopy | S42 |
| #ZÉRÓOSZTÓ! | Chemopy | S43 |
| #ZÉRÓOSZTÓ! | Chemopy | S48 |
| -0.248089476 | Chemopy | S49 |
| -0.384397985 | Chemopy | GATSp5 |
| -0.439741959 | Chemopy | GATSp4 |
| -0.285455295 | Chemopy | GATSp7 |
| -0.19633536 | Chemopy | GATSp6 |
| -0.290997474 | Chemopy | GATSp1 |
| -0.348641767 | Chemopy | bcute13 |
| -0.409975362 | Chemopy | GATSp3 |
| -0.442745064 | Chemopy | GATSp2 |
| -0.306992184 | Chemopy | GATSp8 |
| -0.199945626 | Chemopy | Smin78 |
| -0.069868462 | Chemopy | Chiv6ch |
| -0.892291138 | Chemopy | Tac |
| #ZÉRÓOSZTÓ! | Chemopy | Smin72 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin73 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin70 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin71 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin76 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin77 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin74 |
| -0.14252876 | Chemopy | Smin75 |
| -0.58422942 | Chemopy | Chiv9 |
| -0.175083805 | Chemopy | Chiv8 |
| -0.651075805 | Chemopy | Chiv5 |
| -0.820886105 | Chemopy | Chiv4 |
| -0.632725703 | Chemopy | Chiv7 |
| -0.114996618 | Chemopy | Chiv6 |
| -0.708827691 | Chemopy | Chiv1 |
| -0.927449833 | Chemopy | Chiv0 |
| -0.717954307 | Chemopy | Chiv3 |
| -0.766477677 | Chemopy | Chiv2 |
| -0.360207208 | Chemopy | Smin |
| -0.211749325 | Chemopy | bcutm9 |
| -0.634390289 | Chemopy | bcutm8 |
| -0.542535056 | Chemopy | bcutm7 |
| -0.502434273 | Chemopy | bcutm6 |
| -0.496140475 | Chemopy | bcutm5 |
| -0.443787516 | Chemopy | bcutm4 |
| -0.521856173 | Chemopy | bcutm3 |
| -0.26565207 | Chemopy | bcutm2 |
| -0.096779744 | Chemopy | bcutm1 |
| -0.756653245 | Chemopy | ATSp3 |
| -0.793471519 | Chemopy | ATSp2 |
| -0.573137774 | Chemopy | ATSp1 |
| -0.688390316 | Chemopy | ATSp7 |
| -0.716748821 | Chemopy | ATSp6 |
| -0.736563428 | Chemopy | ATSp5 |
| -0.687315846 | Chemopy | ATSp4 |
| -0.406268653 | Chemopy | ATSp8 |
| -0.032135161 | Chemopy | MATSv1 |
| -0.038410333 | Chemopy | ndonr |
| -0.930255696 | Chemopy | TPSA1 |
| -0.670997111 | Chemopy | bcutv15 |
| -0.394095167 | Chemopy | bcutv16 |
| -0.581927203 | Chemopy | QNss |
| -0.198040368 | Chemopy | bcutv10 |
| -0.292853411 | Chemopy | nhal |
| -0.567440776 | Chemopy | bcutv12 |
| -0.307322063 | Chemopy | bcutv13 |
| -0.44868928 | Chemopy | MATSv2 |
| -0.502759868 | Chemopy | MATSv5 |
| -0.279135718 | Chemopy | SIC0 |
| -0.385292491 | Chemopy | SIC3 |
| -0.109846559 | Chemopy | SIC2 |
| -0.742580847 | Chemopy | PC2 |
| -0.902276527 | Chemopy | PC3 |
| -0.379699027 | Chemopy | PC1 |
| -0.857701409 | Chemopy | PC6 |
| -0.766864179 | Chemopy | PC4 |
| -0.826911129 | Chemopy | PC5 |
| -0.765224506 | Chemopy | Shev |
| -0.913338212 | Chemopy | Shet |
| -0.102946148 | Chemopy | Smax15 |
| -0.84292342 | Chemopy | MR |
| -0.082389426 | Chemopy | nphos |
| -0.851304885 | Chemopy | S9 |
| -0.215596277 | Chemopy | S8 |
| -0.415409592 | Chemopy | GMTIV |
| #ZÉRÓOSZTÓ! | Chemopy | S3 |
| #ZÉRÓOSZTÓ! | Chemopy | S2 |
| -0.204064265 | Chemopy | S1 |
| -0.610766433 | Chemopy | S7 |
| #ZÉRÓOSZTÓ! | Chemopy | S6 |
| #ZÉRÓOSZTÓ! | Chemopy | S5 |
| -0.102768642 | Chemopy | S4 |
| -0.818996184 | Chemopy | nta |
| -0.473753551 | Chemopy | MATSe4 |
| -0.382748168 | Chemopy | MATSe5 |
| -0.372324948 | Chemopy | MATSe6 |
| -0.136411449 | Chemopy | MATSe7 |
| -0.286259062 | Chemopy | MATSe1 |
| -0.43002895 | Chemopy | MATSe2 |
| -0.351835727 | Chemopy | MATSe3 |
| -0.277143545 | Chemopy | MATSe8 |
| -0.111633449 | Chemopy | GATSe7 |
| -0.800212223 | Chemopy | ZM2 |
| -0.827227592 | Chemopy | ZM1 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax67 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax66 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax65 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax64 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax63 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax62 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax61 |
| -0.0614746 | Chemopy | Smax60 |
| -0.915034593 | Chemopy | Thara |
| #ZÉRÓOSZTÓ! | Chemopy | Smax69 |
| -0.221002702 | Chemopy | Smax68 |
| -0.384202026 | Chemopy | bcutp16 |
| -0.232921451 | Chemopy | Smax0 |
| -0.671001241 | Chemopy | bcutp14 |
| -0.398535358 | Chemopy | bcutp15 |
| -0.5911138 | Chemopy | bcutp12 |
| -0.384538156 | Chemopy | bcutp13 |
| -0.38426587 | Chemopy | bcutp10 |
| -0.257474312 | Chemopy | bcutp11 |
| -0.290408623 | Chemopy | Qmin |
| -0.257520915 | Chemopy | Smin51 |
| -0.349526798 | Chemopy | QNmax |
| -0.184195469 | Chemopy | S39 |
| -0.315879419 | Chemopy | S38 |
| -0.010065596 | Chemopy | bcute6 |
| -0.923497802 | Chemopy | S35 |
| -0.406756002 | Chemopy | S34 |
| -0.19765799 | Chemopy | S37 |
| -0.119201137 | Chemopy | S36 |
| -0.266764515 | Chemopy | S31 |
| -0.469945448 | Chemopy | S30 |
| 0.238389758 | Chemopy | S33 |
| 0.04959462 | Chemopy | S32 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin25 |
| -0.131652656 | Chemopy | Smin24 |
| -0.585405745 | Chemopy | naro |
| #ZÉRÓOSZTÓ! | Chemopy | Smin26 |
| 0.022403274 | Chemopy | Smin21 |
| -0.511200301 | Chemopy | Smin20 |
| -0.320574874 | Chemopy | Smin23 |
| 0.104687814 | Chemopy | Smin22 |
| -0.54130425 | Chemopy | Smin29 |
| 0.048181986 | Chemopy | Smin28 |
| -0.06577074 | Chemopy | QCmax |
| -0.888177902 | Chemopy | Tpc |
| -0.221303106 | Chemopy | Mnc |
| -0.389039003 | Chemopy | ATSm8 |
| -0.219403105 | Chemopy | ncocl |
| -0.627800363 | Chemopy | ATSm1 |
| -0.703603159 | Chemopy | ATSm2 |
| -0.704137876 | Chemopy | ATSm3 |
| -0.687431099 | Chemopy | ATSm4 |
| -0.687828554 | Chemopy | ATSm5 |
| -0.683433775 | Chemopy | ATSm6 |
| -0.388325815 | Chemopy | ATSm7 |
| -0.187333147 | Chemopy | Chi3ch |
| -0.564623139 | Chemopy | Chiv3c |
| #ZÉRÓOSZTÓ! | Chemopy | S77 |
| -0.704979351 | Chemopy | Smin27 |
| -0.279473231 | Chemopy | S76 |
| -0.600132138 | Chemopy | Smax29 |
| -0.063288129 | Chemopy | Smax28 |
| -0.380349322 | Chemopy | Smax23 |
| 0.238389758 | Chemopy | Smax22 |
| 0.022291931 | Chemopy | Smax21 |
| -0.576591774 | Chemopy | Smax20 |
| -0.697945425 | Chemopy | Smax27 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax26 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax25 |
| -0.085217592 | Chemopy | Smax24 |
| -0.938258905 | Chemopy | kappa1 |
| -0.720050884 | Chemopy | kappa3 |
| -0.910596021 | Chemopy | kappa2 |
| -0.037595228 | Chemopy | Hato |
| -0.929639842 | Chemopy | TPSA |
| -0.004978946 | Chemopy | S79 |
| -0.954068839 | Chemopy | IDET |
| #ZÉRÓOSZTÓ! | Chemopy | Smax58 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax59 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax56 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax57 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax54 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax55 |
| -0.144159146 | Chemopy | S75 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax53 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax50 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax51 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax73 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin69 |
| -0.23102959 | Chemopy | Smin68 |
| -0.205315548 | Chemopy | Shal |
| #ZÉRÓOSZTÓ! | Chemopy | Smin61 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin60 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin63 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin62 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin65 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin64 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin67 |
| -0.249285073 | Chemopy | Smin66 |
| -0.467809034 | Chemopy | MATSv8 |
| -0.256776289 | Chemopy | ntb |
| -0.552688421 | Chemopy | SIC5 |
| -0.326690064 | Chemopy | SIC4 |
| -0.308883364 | Chemopy | MATSv3 |
| -0.356976977 | Chemopy | SIC6 |
| -0.225981232 | Chemopy | SIC1 |
| -0.551615995 | Chemopy | MATSv4 |
| -0.269483871 | Chemopy | MATSv7 |
| -0.288542951 | Chemopy | MATSv6 |
| 0.003934203 | Chemopy | GATSe8 |
| -0.948717888 | Chemopy | W |
| -0.411138839 | Chemopy | GATSe2 |
| -0.346986197 | Chemopy | GATSe3 |
| -0.341156732 | Chemopy | GATSe1 |
| -0.291289726 | Chemopy | GATSe6 |
| -0.230308847 | Chemopy | VSAEstate10 |
| -0.41758078 | Chemopy | GATSe4 |
| -0.055620022 | Chemopy | GATSe5 |
| -0.919269655 | Chemopy | Ipc |
| -0.302988322 | Chemopy | Smin45 |
| -0.43308008 | Chemopy | Rpc |
| -0.128548653 | Chemopy | Chiv3ch |
| -0.235557525 | Chemopy | Smin44 |
| -0.083119221 | Chemopy | bcute7 |
| -0.87745104 | Chemopy | LogP2 |
| -0.515854843 | Chemopy | bcute5 |
| -0.449169805 | Chemopy | bcute4 |
| -0.448296624 | Chemopy | bcute3 |
| -0.207140681 | Chemopy | bcute2 |
| -0.279650869 | Chemopy | bcute1 |
| -0.231939419 | Chemopy | bcute9 |
| -0.148092028 | Chemopy | bcute8 |
| -0.555795948 | Chemopy | knotp |
| -0.210053488 | Chemopy | ncobr |
| -0.170564921 | Chemopy | MATSm8 |
| -0.323574285 | Chemopy | MATSm4 |
| -0.350342811 | Chemopy | MATSm5 |
| -0.447911368 | Chemopy | MATSm6 |
| -0.201026458 | Chemopy | MATSm7 |
| -0.227778473 | Chemopy | MATSm1 |
| -0.336784241 | Chemopy | MATSm2 |
| -0.023368618 | Chemopy | MATSm3 |
| -0.847167542 | Chemopy | nnitro |
| -0.108169138 | Chemopy | Chi3c |
| -0.87957556 | Chemopy | IVDE |
| -0.923871275 | Chemopy | MRVSA9 |
| -0.264658327 | Chemopy | MRVSA8 |
| -0.135752234 | Chemopy | MRVSA5 |
| -0.680948525 | Chemopy | MRVSA4 |
| -0.142033819 | Chemopy | MRVSA7 |
| -0.597831778 | Chemopy | MRVSA6 |
| -0.080883127 | Chemopy | MRVSA1 |
| -0.868635734 | Chemopy | MRVSA0 |
| -0.757886224 | Chemopy | MRVSA3 |
| -0.873018825 | Chemopy | MRVSA2 |
| -0.240190036 | Chemopy | Qmax |
| -0.100343444 | Chemopy | Smax |
| -0.762042284 | Chemopy | Xu |
| -0.487829236 | Chemopy | Smax12 |
| -0.144159146 | Chemopy | Smax13 |
| -0.207039244 | Chemopy | Smax10 |
| -0.394317405 | Chemopy | Smax11 |
| -0.558674285 | Chemopy | Smax16 |
| 0.053726809 | Chemopy | Smax17 |
| -0.020649812 | Chemopy | Smax14 |
| -0.923738903 | Chemopy | ndb |
| -0.133723526 | Chemopy | Smax18 |
| -0.093362544 | Chemopy | Smax19 |
| -0.855858637 | Chemopy | DZ |
| 0.007448395 | Chemopy | PEOEVSA12 |
| -0.893084478 | Chemopy | PEOEVSA11 |
| -0.435171241 | Chemopy | PEOEVSA10 |
| -0.263639972 | Chemopy | nsulph |
| -0.313713267 | Chemopy | DS |
| -0.208630338 | Chemopy | QCmin |
| -0.13650818 | Chemopy | IDE |
| -0.669414956 | Chemopy | Chi4pc |
| 0.016806485 | Chemopy | S22 |
| 0.238389758 | Chemopy | S23 |
| 0.016027232 | Chemopy | S20 |
| -0.923484891 | Chemopy | S21 |
| #ZÉRÓOSZTÓ! | Chemopy | S26 |
| -0.02001608 | Chemopy | S27 |
| -0.850279619 | Chemopy | S24 |
| 0.053726809 | Chemopy | S25 |
| -0.804131129 | Chemopy | S28 |
| 0.037388084 | Chemopy | S29 |
| -0.218834386 | Chemopy | Smin10 |
| -0.11489375 | Chemopy | Smin11 |
| -0.791736378 | Chemopy | Pol |
| #ZÉRÓOSZTÓ! | Chemopy | Smin13 |
| -0.128157181 | Chemopy | Smin14 |
| -0.686920688 | Chemopy | Smin15 |
| -0.134143763 | Chemopy | Smin16 |
| -0.229635582 | Chemopy | Smin17 |
| -0.123411692 | Chemopy | Smin18 |
| -0.09501452 | Chemopy | Smin19 |
| -0.852683466 | Chemopy | Chi0 |
| -0.261794264 | Chemopy | Chi6ch |
| -0.332130721 | Chemopy | Save |
| -0.133532538 | Chemopy | Chi4c |
| -0.35368262 | Chemopy | Rnc |
| -0.624125244 | Chemopy | Smax8 |
| -0.093380171 | Chemopy | Smax9 |
| -0.779154734 | Chemopy | naccr |
| #ZÉRÓOSZTÓ! | Chemopy | Smax1 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax2 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax3 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax4 |
| -0.20321937 | Chemopy | Smax5 |
| -0.326164071 | Chemopy | Smax6 |
| -0.15699528 | Chemopy | Smax7 |
| -0.664906742 | Chemopy | Qindex |
| -0.264840305 | Chemopy | bcutv14 |
| -0.351651241 | Chemopy | slogPVSA10 |
| -0.895224206 | Chemopy | slogPVSA11 |
| -0.244623007 | Chemopy | Smin57 |
| -0.261252225 | Chemopy | bcutv11 |
| -0.141813224 | Chemopy | S50 |
| -0.918903401 | Chemopy | Chi8 |
| -0.108758896 | Chemopy | Chi9 |
| -0.879647466 | Chemopy | Chi2 |
| -0.804417497 | Chemopy | Chi3 |
| -0.095030881 | Chemopy | LDI |
| -0.938157898 | Chemopy | Chi1 |
| -0.910272817 | Chemopy | Chi6 |
| -0.257094408 | Chemopy | Chi7 |
| -0.850632515 | Chemopy | Chi4 |
| -0.814280829 | Chemopy | Chi5 |
| -0.065069049 | Chemopy | Smin32 |
| -0.65101594 | Chemopy | PEOEVSA13 |
| -0.185953733 | Chemopy | Chiv4c |
| -0.771289013 | Chemopy | LogP |
| -0.895224206 | Chemopy | Smax49 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax48 |
| #ZÉRÓOSZTÓ! | Chemopy | S68 |
| #ZÉRÓOSZTÓ! | Chemopy | S69 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax45 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax44 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax47 |
| #ZÉRÓOSZTÓ! | Chemopy | S65 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax41 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax40 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax43 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax42 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin54 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin55 |
| -0.232037903 | Chemopy | Smin56 |
| -0.354102818 | Chemopy | bcute14 |
| -0.236698182 | Chemopy | Smin50 |
| -0.267499678 | Chemopy | bcute12 |
| -0.144159146 | Chemopy | Smin52 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin53 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin58 |
| -0.105061036 | Chemopy | Smin59 |
| -0.753076245 | Chemopy | dchi4 |
| -0.790653932 | Chemopy | dchi3 |
| -0.820509937 | Chemopy | dchi2 |
| -0.771619406 | Chemopy | dchi1 |
| -0.641698701 | Chemopy | dchi0 |
| -0.166593366 | Chemopy | EstateVSA4 |
| -0.574459163 | Chemopy | EstateVSA5 |
| -0.041891735 | Chemopy | EstateVSA6 |
| -0.252884075 | Chemopy | EstateVSA7 |
| -0.952379439 | Chemopy | EstateVSA0 |
| -0.871176866 | Chemopy | EstateVSA1 |
| -0.026207741 | Chemopy | EstateVSA2 |
| -0.622027016 | Chemopy | EstateVSA3 |
| #ZÉRÓOSZTÓ! | Chemopy | S66 |
| #ZÉRÓOSZTÓ! | Chemopy | S67 |
| -0.217773146 | Chemopy | S64 |
| -0.407200809 | Chemopy | ATSv1 |
| -0.216599983 | Chemopy | VSAEstate0 |
| -0.370222433 | Chemopy | ATSv8 |
| -0.081804868 | Chemopy | S19 |
| 0.053726809 | Chemopy | S18 |
| #ZÉRÓOSZTÓ! | Chemopy | VSAEstate4 |
| #ZÉRÓOSZTÓ! | Chemopy | VSAEstate5 |
| -0.228074645 | Chemopy | VSAEstate6 |
| -0.101032064 | Chemopy | BertzCT |
| -0.539547625 | Chemopy | S13 |
| -0.58065899 | Chemopy | S12 |
| -0.044068338 | Chemopy | S11 |
| -0.198393663 | Chemopy | S10 |
| -0.025053029 | Chemopy | S17 |
| -0.939120726 | Chemopy | S16 |
| #ZÉRÓOSZTÓ! | Chemopy | S15 |
| -0.231867324 | Chemopy | S14 |
| -0.400137256 | Chemopy | GATSm2 |
| -0.389442904 | Chemopy | GATSm3 |
| -0.300760655 | Chemopy | GATSm1 |
| -0.283278435 | Chemopy | GATSm6 |
| -0.195351299 | Chemopy | S60 |
| -0.098678501 | Chemopy | GATSm4 |
| -0.261137491 | Chemopy | GATSm5 |
| -0.306678211 | Chemopy | GATSm8 |
| -0.211951184 | Chemopy | S61 |
| -0.637608624 | Chemopy | bcute16 |
| -0.389547784 | Chemopy | bcute15 |
| -0.222675969 | Chemopy | Chiv5ch |
| -0.67562858 | Chemopy | GMTI |
| -0.300571796 | Chemopy | nring |
| -0.267334344 | Chemopy | QHmax |
| -0.411061471 | Chemopy | bcute11 |
| -0.334251193 | Chemopy | bcute10 |
| -0.645283432 | Chemopy | CIC5 |
| -0.664106611 | Chemopy | CIC4 |
| -0.375298722 | Chemopy | CIC6 |
| -0.408319565 | Chemopy | CIC1 |
| -0.452614393 | Chemopy | CIC0 |
| -0.369221592 | Chemopy | CIC3 |
| -0.40821918 | Chemopy | CIC2 |
| -0.442008548 | Chemopy | bcutp8 |
| -0.237512368 | Chemopy | bcutp9 |
| -0.679932499 | Chemopy | bcutp4 |
| -0.606779534 | Chemopy | bcutp5 |
| -0.653592336 | Chemopy | bcutp6 |
| -0.675965615 | Chemopy | bcutp7 |
| -0.578689623 | Chemopy | bcutp1 |
| -0.388284011 | Chemopy | bcutp2 |
| -0.104881118 | Chemopy | bcutp3 |
| -0.788474553 | Chemopy | ncarb |
| -0.152767752 | Chemopy | kappam1 |
| -0.893432579 | Chemopy | TIAC |
| -0.721950329 | Chemopy | kappam3 |
| -0.950629321 | Chemopy | kappam2 |
| -0.883536215 | Chemopy | noxy |
| -0.489722798 | Chemopy | Chiv4pc |
| -0.660638048 | Chemopy | PEOEVSA3 |
| -0.012773407 | Chemopy | PEOEVSA2 |
| -0.676156421 | Chemopy | PEOEVSA1 |
| -0.951225137 | Chemopy | PEOEVSA0 |
| -0.144136147 | Chemopy | PEOEVSA7 |
| -0.799459593 | Chemopy | PEOEVSA6 |
| -0.731938978 | Chemopy | PEOEVSA5 |
| -0.06204647 | Chemopy | PEOEVSA4 |
| -0.083075242 | Chemopy | Qass |
| -0.890825051 | Chemopy | PEOEVSA9 |
| -0.376211104 | Chemopy | PEOEVSA8 |
| -0.895853389 | Chemopy | phi |
| -0.278522149 | Chemopy | QOmin |
| -0.432604541 | Chemopy | GATSv3 |
| -0.365481825 | Chemopy | GATSv2 |
| -0.507405663 | Chemopy | GATSv1 |
| -0.2856956 | Chemopy | GATSv7 |
| -0.298428514 | Chemopy | GATSv6 |
| -0.399642641 | Chemopy | GATSv5 |
| -0.585811549 | Chemopy | GATSv4 |
| -0.252305176 | Chemopy | GATSv8 |
| -0.920855641 | Chemopy | QCss |
| -0.069152746 | Chemopy | Chiv10 |
| -0.937836862 | Chemopy | slogPVSA2 |
| -0.033486463 | Chemopy | slogPVSA3 |
| -0.426006722 | Chemopy | slogPVSA0 |
| -0.787762375 | Chemopy | slogPVSA1 |
| #ZÉRÓOSZTÓ! | Chemopy | slogPVSA6 |
| -0.097010895 | Chemopy | slogPVSA7 |
| -0.849633633 | Chemopy | slogPVSA4 |
| -0.629500753 | Chemopy | slogPVSA5 |
| #ZÉRÓOSZTÓ! | Chemopy | Chi4ch |
| -0.189683617 | Chemopy | slogPVSA8 |
| -0.167210668 | Chemopy | slogPVSA9 |
| #ZÉRÓOSZTÓ! | Chemopy | S57 |
| #ZÉRÓOSZTÓ! | Chemopy | S56 |
| -0.144159146 | Chemopy | S55 |
| #ZÉRÓOSZTÓ! | Chemopy | S54 |
| #ZÉRÓOSZTÓ! | Chemopy | S53 |
| #ZÉRÓOSZTÓ! | Chemopy | S52 |
| -0.09801575 | Chemopy | S51 |
| -0.891308514 | Chemopy | Gravto |
| -0.163754313 | Chemopy | EstateVSA8 |
| -0.877036608 | Chemopy | Sitov |
| #ZÉRÓOSZTÓ! | Chemopy | S59 |
| -0.20713776 | Chemopy | S58 |
| -0.10020067 | Chemopy | Mpc |
| -0.857042759 | Chemopy | LabuteASA |
| -0.83647446 | Chemopy | radiust |
| -0.367886165 | Chemopy | MATSp8 |
| -0.257466972 | Chemopy | MATSp7 |
| -0.404000401 | Chemopy | MATSp6 |
| -0.329867025 | Chemopy | MATSp5 |
| -0.542894407 | Chemopy | MATSp4 |
| -0.424916317 | Chemopy | MATSp3 |
| -0.553336429 | Chemopy | MATSp2 |
| -0.064148359 | Chemopy | MATSp1 |
| -0.360974547 | Chemopy | nhet |
| -0.892879266 | Chemopy | nhev |
| -0.257583637 | Chemopy | AWeight |
| #ZÉRÓOSZTÓ! | Chemopy | Chiv4ch |
| -0.251265822 | Chemopy | ncoi |
| -0.249021555 | Chemopy | QOmax |
| -0.256859277 | Chemopy | ncof |
| -0.323218493 | Chemopy | Hy |
| -0.050350882 | Chemopy | Smin0 |
| -0.722604293 | Chemopy | QHss |
| -0.318791963 | Chemopy | Getov |
| #ZÉRÓOSZTÓ! | Chemopy | VSAEstate1 |
| #ZÉRÓOSZTÓ! | Chemopy | VSAEstate2 |
| -0.159154684 | Chemopy | VSAEstate3 |
| -0.097138549 | Chemopy | knotpv |
| -0.855462591 | Chemopy | Weight |
| -0.182624383 | Chemopy | VSAEstate7 |
| -0.072784629 | Chemopy | EstateVSA10 |
| -0.890629364 | Chemopy | VSAEstate8 |
| -0.473054924 | Chemopy | UI |
| -0.555872698 | Chemopy | VSAEstate9 |
| -0.222365556 | Chemopy | mChi1 |
| -0.756770737 | Chemopy | ATSv3 |
| -0.694694952 | Chemopy | ATSv2 |
| -0.09811547 | Chemopy | ATSv5 |
| -0.788146847 | Chemopy | nsb |
| -0.076653106 | Chemopy | petitjeant |
| -0.416952137 | Chemopy | QOss |
| -0.381413726 | Chemopy | ATSv4 |
| -0.219312151 | Chemopy | Smax78 |
| -0.393290219 | Chemopy | ATSv7 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax70 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax71 |
| -0.220649153 | Chemopy | Smax72 |
| -0.385396338 | Chemopy | ATSv6 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax74 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax75 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax76 |
| #ZÉRÓOSZTÓ! | Chemopy | Smax77 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin43 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin42 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin41 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin40 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin47 |
| -0.093896819 | Chemopy | Smin46 |
| -0.553631185 | Chemopy | diametert |
| -0.704314561 | Chemopy | Scar |
| -0.895224206 | Chemopy | Smin49 |
| -0.079907648 | Chemopy | Smin48 |
| -0.870931985 | Chemopy | ISIZ |
| -0.045075532 | Chemopy | GATSm7 |
| -0.916239987 | Chemopy | EstateVSA9 |
| -0.09919508 | Chemopy | J |
| -0.891757932 | Chemopy | Tnc |
| -0.258738677 | Chemopy | QHmin |
| -0.214890078 | Chemopy | S78 |
| -0.415067423 | Chemopy | Tigdi |
| -0.184195469 | Chemopy | Smin36 |
| -0.341208542 | Chemopy | Smin37 |
| -0.296644181 | Chemopy | Smin34 |
| -0.135058099 | Chemopy | Smin35 |
| -0.198730862 | Chemopy | S62 |
| -0.355468961 | Chemopy | Smin33 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin30 |
| -0.108724813 | Chemopy | Smin31 |
| -0.782775297 | Chemopy | nhyd |
| #ZÉRÓOSZTÓ! | Chemopy | Smin38 |
| -0.218661766 | Chemopy | Smin39 |
| -0.652297623 | Chemopy | ATSe1 |
| -0.706100502 | Chemopy | ATSe2 |
| -0.702488811 | Chemopy | ATSe3 |
| -0.688589205 | Chemopy | ATSe4 |
| -0.688990905 | Chemopy | ATSe5 |
| -0.682992606 | Chemopy | ATSe6 |
| -0.67092931 | Chemopy | ATSe7 |
| -0.404171863 | Chemopy | ATSe8 |
| #ZÉRÓOSZTÓ! | Chemopy | S71 |
| #ZÉRÓOSZTÓ! | Chemopy | S63 |
| #ZÉRÓOSZTÓ! | Chemopy | S70 |
| #ZÉRÓOSZTÓ! | Chemopy | S73 |
| -0.223909914 | Chemopy | S72 |
| -0.342897618 | Chemopy | Hatov |
| -0.133105457 | Chemopy | Smax52 |
| -0.554610979 | Chemopy | Smin12 |
| -0.209663372 | Chemopy | S74 |
| -0.260356916 | Chemopy | Mac |
| -0.241859131 | Chemopy | Arto |
| -0.497574147 | Chemopy | IC3 |
| -0.364512205 | Chemopy | IC2 |
| -0.435126631 | Chemopy | IC1 |
| -0.254677699 | Chemopy | IC0 |
| -0.684453388 | Chemopy | IC6 |
| -0.656002508 | Chemopy | IC5 |
| -0.323773882 | Chemopy | IC4 |
| -0.536956862 | Chemopy | Smin6 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin7 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin4 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin5 |
| #ZÉRÓOSZTÓ! | Chemopy | Smin2 |
| 0.003725832 | Chemopy | Smin3 |
| -0.948962945 | Chemopy | Tsch |
| -0.056051438 | Chemopy | Smin1 |
| -0.759899148 | Chemopy | Smin8 |
| -0.113058611 | Chemopy | Smin9 |
| -0.911141075 | Chemopy | AW |
| -0.109525866 | Chemopy | MZM2 |
| -0.795252374 | Chemopy | MZM1 |
| -0.648256562 | Chemopy | nrot |
| -0.842187936 | Chemopy | Sito |
| -0.126107878 | Chemopy | Smax46 |
| -0.965829357 | Chemopy | Chi10 |
| -0.614084058 | Chemopy | bcutm16 |
| -0.586720569 | Chemopy | bcutm15 |
| -0.549582349 | Chemopy | bcutm14 |
| -0.455600473 | Chemopy | bcutm13 |
| -0.348472414 | Chemopy | bcutm12 |
| -0.421985894 | Chemopy | bcutm11 |
| -0.260840749 | Chemopy | bcutm10 |
| -0.39848162 | RDKit | BalabanJ |
| -0.715691083 | RDKit | BertzCT |
| -0.731575347 | RDKit | Chi0 |
| -0.737226971 | RDKit | Chi0n |
| -0.736794538 | RDKit | Chi0v |
| -0.730794589 | RDKit | Chi1 |
| -0.737617054 | RDKit | Chi1n |
| -0.741060523 | RDKit | Chi1v |
| -0.743441273 | RDKit | Chi2n |
| -0.736117541 | RDKit | Chi2v |
| -0.742714542 | RDKit | Chi3n |
| -0.747724325 | RDKit | Chi3v |
| -0.745175712 | RDKit | Chi4n |
| -0.7461083 | RDKit | Chi4v |
| -0.72535036 | RDKit | EState\_VSA1 |
| -0.703077194 | RDKit | EState\_VSA10 |
| -0.653346111 | RDKit | EState\_VSA11 |
| -0.738843151 | RDKit | EState\_VSA2 |
| -0.003489035 | RDKit | EState\_VSA3 |
| -0.483302065 | RDKit | EState\_VSA4 |
| -0.478617309 | RDKit | EState\_VSA5 |
| -0.639548624 | RDKit | EState\_VSA6 |
| -0.302760856 | RDKit | EState\_VSA7 |
| -0.631537396 | RDKit | EState\_VSA8 |
| -0.543159652 | RDKit | EState\_VSA9 |
| -0.734164927 | RDKit | ExactMolWt |
| -0.39746839 | RDKit | FractionCSP3 |
| -0.676906698 | RDKit | HallKierAlpha |
| -0.730842057 | RDKit | HeavyAtomCount |
| -0.733437012 | RDKit | HeavyAtomMolWt |
| -0.741692408 | RDKit | Ipc |
| -0.73180784 | RDKit | Kappa1 |
| -0.726104096 | RDKit | Kappa2 |
| -0.708569652 | RDKit | Kappa3 |
| -0.732486112 | RDKit | LabuteASA |
| -0.587769342 | RDKit | MaxAbsEStateIndex |
| -0.178881791 | RDKit | MaxAbsPartialCharge |
| -0.587769342 | RDKit | MaxEStateIndex |
| -0.234131215 | RDKit | MaxPartialCharge |
| -0.521713888 | RDKit | MinAbsEStateIndex |
| -0.234131215 | RDKit | MinAbsPartialCharge |
| -0.388720246 | RDKit | MinEStateIndex |
| -0.178881791 | RDKit | MinPartialCharge |
| -0.70363364 | RDKit | MolLogP |
| -0.731625532 | RDKit | MolMR |
| -0.734146917 | RDKit | MolWt |
| -0.699087892 | RDKit | NHOHCount |
| -0.734833472 | RDKit | NOCount |
| -0.144159146 | RDKit | NumAliphaticCarbocycles |
| -0.321356401 | RDKit | NumAliphaticHeterocycles |
| -0.364354 | RDKit | NumAliphaticRings |
| -0.336208201 | RDKit | NumAromaticCarbocycles |
| 0.073854979 | RDKit | NumAromaticHeterocycles |
| -0.33996117 | RDKit | NumAromaticRings |
| -0.736505205 | RDKit | NumHAcceptors |
| -0.709507658 | RDKit | NumHDonors |
| -0.736957832 | RDKit | NumHeteroatoms |
| #ZÉRÓOSZTÓ! | RDKit | NumRadicalElectrons |
| -0.703547191 | RDKit | NumRotatableBonds |
| #ZÉRÓOSZTÓ! | RDKit | NumSaturatedCarbocycles |
| -0.320449721 | RDKit | NumSaturatedHeterocycles |
| -0.320449721 | RDKit | NumSaturatedRings |
| -0.733844716 | RDKit | NumValenceElectrons |
| -0.736152671 | RDKit | PEOE\_VSA1 |
| -0.713329174 | RDKit | PEOE\_VSA10 |
| -0.383939826 | RDKit | PEOE\_VSA11 |
| -0.677219115 | RDKit | PEOE\_VSA12 |
| 0.000887673 | RDKit | PEOE\_VSA13 |
| -0.540430692 | RDKit | PEOE\_VSA14 |
| -0.7015406 | RDKit | PEOE\_VSA2 |
| -0.218036994 | RDKit | PEOE\_VSA3 |
| -0.895224206 | RDKit | PEOE\_VSA4 |
| #ZÉRÓOSZTÓ! | RDKit | PEOE\_VSA5 |
| -0.532833414 | RDKit | PEOE\_VSA6 |
| -0.746274752 | RDKit | PEOE\_VSA7 |
| -0.532117414 | RDKit | PEOE\_VSA8 |
| -0.37126265 | RDKit | PEOE\_VSA9 |
| -0.222066072 | RDKit | RingCount |
| -0.703958197 | RDKit | SMR\_VSA1 |
| -0.697017432 | RDKit | SMR\_VSA10 |
| 0.117260247 | RDKit | SMR\_VSA2 |
| -0.699156804 | RDKit | SMR\_VSA3 |
| -0.782592345 | RDKit | SMR\_VSA4 |
| -0.679309619 | RDKit | SMR\_VSA5 |
| -0.446900799 | RDKit | SMR\_VSA6 |
| -0.368999873 | RDKit | SMR\_VSA7 |
| #ZÉRÓOSZTÓ! | RDKit | SMR\_VSA8 |
| -0.315938921 | RDKit | SMR\_VSA9 |
| -0.675391169 | RDKit | SlogP\_VSA1 |
| -0.260267836 | RDKit | SlogP\_VSA10 |
| -0.314166717 | RDKit | SlogP\_VSA11 |
| -0.728792658 | RDKit | SlogP\_VSA12 |
| -0.744516472 | RDKit | SlogP\_VSA2 |
| -0.726785439 | RDKit | SlogP\_VSA3 |
| -0.256307164 | RDKit | SlogP\_VSA4 |
| -0.717361148 | RDKit | SlogP\_VSA5 |
| -0.424731487 | RDKit | SlogP\_VSA6 |
| #ZÉRÓOSZTÓ! | RDKit | SlogP\_VSA7 |
| 0.071899968 | RDKit | SlogP\_VSA8 |
| #ZÉRÓOSZTÓ! | RDKit | SlogP\_VSA9 |
| -0.725500929 | RDKit | TPSA |
| #ZÉRÓOSZTÓ! | RDKit | VSA\_EState1 |
| -0.206737829 | RDKit | VSA\_EState10 |
| #ZÉRÓOSZTÓ! | RDKit | VSA\_EState2 |
| #ZÉRÓOSZTÓ! | RDKit | VSA\_EState3 |
| #ZÉRÓOSZTÓ! | RDKit | VSA\_EState4 |
| #ZÉRÓOSZTÓ! | RDKit | VSA\_EState5 |
| #ZÉRÓOSZTÓ! | RDKit | VSA\_EState6 |
| #ZÉRÓOSZTÓ! | RDKit | VSA\_EState7 |
| -0.291444321 | RDKit | VSA\_EState8 |
| -0.729574028 | RDKit | VSA\_EState9 |
| -0.677808968 | RDKit | fr\_Al\_COO |
| -0.325675641 | RDKit | fr\_Al\_OH |
| -0.461301582 | RDKit | fr\_Al\_OH\_noTert |
| #ZÉRÓOSZTÓ! | RDKit | fr\_ArN |
| #ZÉRÓOSZTÓ! | RDKit | fr\_Ar\_COO |
| 0.074460072 | RDKit | fr\_Ar\_N |
| 0.053726809 | RDKit | fr\_Ar\_NH |
| -0.395959929 | RDKit | fr\_Ar\_OH |
| -0.677808968 | RDKit | fr\_COO |
| -0.677808968 | RDKit | fr\_COO2 |
| -0.69651419 | RDKit | fr\_C\_O |
| -0.696822198 | RDKit | fr\_C\_O\_noCOO |
| #ZÉRÓOSZTÓ! | RDKit | fr\_C\_S |
| #ZÉRÓOSZTÓ! | RDKit | fr\_HOCCN |
| -0.895224206 | RDKit | fr\_Imine |
| -0.542850239 | RDKit | fr\_NH0 |
| -0.685336178 | RDKit | fr\_NH1 |
| -0.763697521 | RDKit | fr\_NH2 |
| #ZÉRÓOSZTÓ! | RDKit | fr\_N\_O |
| -0.268955836 | RDKit | fr\_Ndealkylation1 |
| -0.504814388 | RDKit | fr\_Ndealkylation2 |
| 0.053726809 | RDKit | fr\_Nhpyrrole |
| #ZÉRÓOSZTÓ! | RDKit | fr\_SH |
| -0.044068338 | RDKit | fr\_aldehyde |
| #ZÉRÓOSZTÓ! | RDKit | fr\_alkyl\_carbamate |
| -0.144159146 | RDKit | fr\_alkyl\_halide |
| -0.038928909 | RDKit | fr\_allylic\_oxid |
| -0.671847277 | RDKit | fr\_amide |
| #ZÉRÓOSZTÓ! | RDKit | fr\_amidine |
| -0.225088799 | RDKit | fr\_aniline |
| 0.04959462 | RDKit | fr\_aryl\_methyl |
| #ZÉRÓOSZTÓ! | RDKit | fr\_azide |
| #ZÉRÓOSZTÓ! | RDKit | fr\_azo |
| #ZÉRÓOSZTÓ! | RDKit | fr\_barbitur |
| -0.336208201 | RDKit | fr\_benzene |
| #ZÉRÓOSZTÓ! | RDKit | fr\_benzodiazepine |
| -0.043091156 | RDKit | fr\_bicyclic |
| #ZÉRÓOSZTÓ! | RDKit | fr\_diazo |
| #ZÉRÓOSZTÓ! | RDKit | fr\_dihydropyridine |
| -0.128548653 | RDKit | fr\_epoxide |
| -0.190207358 | RDKit | fr\_ester |
| -0.236989677 | RDKit | fr\_ether |
| #ZÉRÓOSZTÓ! | RDKit | fr\_furan |
| -0.697978712 | RDKit | fr\_guanido |
| -0.255708696 | RDKit | fr\_halogen |
| #ZÉRÓOSZTÓ! | RDKit | fr\_hdrzine |
| #ZÉRÓOSZTÓ! | RDKit | fr\_hdrzone |
| 0.074863487 | RDKit | fr\_imidazole |
| #ZÉRÓOSZTÓ! | RDKit | fr\_imide |
| #ZÉRÓOSZTÓ! | RDKit | fr\_isocyan |
| #ZÉRÓOSZTÓ! | RDKit | fr\_isothiocyan |
| -0.057594643 | RDKit | fr\_ketone |
| -0.057594643 | RDKit | fr\_ketone\_Topliss |
| #ZÉRÓOSZTÓ! | RDKit | fr\_lactam |
| -0.233372733 | RDKit | fr\_lactone |
| -0.146864616 | RDKit | fr\_methoxy |
| -0.144159146 | RDKit | fr\_morpholine |
| #ZÉRÓOSZTÓ! | RDKit | fr\_nitrile |
| #ZÉRÓOSZTÓ! | RDKit | fr\_nitro |
| #ZÉRÓOSZTÓ! | RDKit | fr\_nitro\_arom |
| #ZÉRÓOSZTÓ! | RDKit | fr\_nitro\_arom\_nonortho |
| #ZÉRÓOSZTÓ! | RDKit | fr\_nitroso |
| #ZÉRÓOSZTÓ! | RDKit | fr\_oxazole |
| 0.015618841 | RDKit | fr\_oxime |
| -0.016737109 | RDKit | fr\_para\_hydroxylation |
| -0.395959929 | RDKit | fr\_phenol |
| -0.395959929 | RDKit | fr\_phenol\_noOrthoHbond |
| #ZÉRÓOSZTÓ! | RDKit | fr\_phos\_acid |
| #ZÉRÓOSZTÓ! | RDKit | fr\_phos\_ester |
| -0.205315548 | RDKit | fr\_piperdine |
| -0.205315548 | RDKit | fr\_piperzine |
| -0.708499456 | RDKit | fr\_priamide |
| #ZÉRÓOSZTÓ! | RDKit | fr\_prisulfonamd |
| 0.04959462 | RDKit | fr\_pyridine |
| 0.238389758 | RDKit | fr\_quatN |
| -0.895224206 | RDKit | fr\_sulfide |
| #ZÉRÓOSZTÓ! | RDKit | fr\_sulfonamd |
| #ZÉRÓOSZTÓ! | RDKit | fr\_sulfone |
| #ZÉRÓOSZTÓ! | RDKit | fr\_term\_acetylene |
| #ZÉRÓOSZTÓ! | RDKit | fr\_tetrazole |
| #ZÉRÓOSZTÓ! | RDKit | fr\_thiazole |
| #ZÉRÓOSZTÓ! | RDKit | fr\_thiocyan |
| #ZÉRÓOSZTÓ! | RDKit | fr\_thiophene |
| -0.65655239 | RDKit | fr\_unbrch\_alkane |
| -0.144159146 | RDKit | fr\_urea |
| #ÉRTÉK! | RDKit | mol\_smile |

## Not monotonous attributes (666 of 831)

Geto, Chi5ch, Smax35, Smax37, Smax31, Smax32, Smax33, bcutv8, bcutv9, bcutv6, bcutv7, bcutv4, bcutv5, bcutv2, bcutv3, QNmin, bcutv1, SPP, Platt, Smax34, S49, GATSp5, GATSp4, GATSp7, GATSp6, GATSp1, bcute13, GATSp3, GATSp2, GATSp8, Smin78, Chiv6ch, Tac, Smin75, Chiv9, Chiv8, Chiv5, Chiv4, Chiv7, Chiv6, Chiv1, Chiv0, Chiv3, Chiv2, Smin, bcutm9, bcutm8, bcutm7, bcutm6, bcutm5, bcutm4, bcutm3, bcutm2, bcutm1, ATSp3, ATSp2, ATSp1, ATSp7, ATSp6, ATSp5, ATSp4, ATSp8, MATSv1, ndonr, TPSA1, bcutv15, bcutv16, QNss, bcutv10, nhal, bcutv12, bcutv13, MATSv2, MATSv5, SIC0, SIC3, SIC2, PC2, PC3, PC1, PC6, PC4, PC5, Shev, Shet, Smax15, MR, nphos, S9, S8, GMTIV, S1, S7, S4, nta, MATSe4, MATSe5, MATSe6, MATSe7, MATSe1, MATSe2, MATSe3, MATSe8, GATSe7, ZM2, ZM1, Smax60, Thara, Smax68, bcutp16, Smax0, bcutp14, bcutp15, bcutp12, bcutp13, bcutp10, bcutp11, Qmin, Smin51, QNmax, S39, S38, bcute6, S35, S34, S37, S36, S31, S30, S33, S32, Smin24, naro, Smin21, Smin20, Smin23, Smin22, Smin29, Smin28, QCmax, Tpc, Mnc, ATSm8, ncocl, ATSm1, ATSm2, ATSm3, ATSm4, ATSm5, ATSm6, ATSm7, Chi3ch, Chiv3c, Smin27, S76, Smax29, Smax28, Smax23, Smax22, Smax21, Smax20, Smax27, Smax24, kappa1, kappa3, kappa2, Hato, TPSA, S79, IDET, S75, Smin68, Shal, Smin66, MATSv8, ntb, SIC5, SIC4, MATSv3, SIC6, SIC1, MATSv4, MATSv7, MATSv6, GATSe8, W, GATSe2, GATSe3, GATSe1, GATSe6, VSAEstate10, GATSe4, GATSe5, Ipc, Smin45, Rpc, Chiv3ch, Smin44, bcute7, LogP2, bcute5, bcute4, bcute3, bcute2, bcute1, bcute9, bcute8, knotp, ncobr, MATSm8, MATSm4, MATSm5, MATSm6, MATSm7, MATSm1, MATSm2, MATSm3, nnitro, Chi3c, IVDE, MRVSA9, MRVSA8, MRVSA5, MRVSA4, MRVSA7, MRVSA6, MRVSA1, MRVSA0, MRVSA3, MRVSA2, Qmax, Smax, Xu, Smax12, Smax13, Smax10, Smax11, Smax16, Smax17, Smax14, ndb, Smax18, Smax19, DZ, PEOEVSA12, PEOEVSA11, PEOEVSA10, nsulph, DS, QCmin, IDE, Chi4pc, S22, S23, S20, S21, S27, S24, S25, S28, S29, Smin10, Smin11, Pol, Smin14, Smin15, Smin16, Smin17, Smin18, Smin19, Chi0, Chi6ch, Save, Chi4c, Rnc, Smax8, Smax9, naccr, Smax5, Smax6, Smax7, Qindex, bcutv14, slogPVSA10, slogPVSA11, Smin57, bcutv11, S50, Chi8, Chi9, Chi2, Chi3, LDI, Chi1, Chi6, Chi7, Chi4, Chi5, Smin32, PEOEVSA13, Chiv4c, LogP, Smax49, Smin56, bcute14, Smin50, bcute12, Smin52, Smin59, dchi4, dchi3, dchi2, dchi1, dchi0, EstateVSA4, EstateVSA5, EstateVSA6, EstateVSA7, EstateVSA0, EstateVSA1, EstateVSA2, EstateVSA3, S64, ATSv1, VSAEstate0, ATSv8, S19, S18, VSAEstate6, BertzCT, S13, S12, S11, S10, S17, S16, S14, GATSm2, GATSm3, GATSm1, GATSm6, S60, GATSm4, GATSm5, GATSm8, S61, bcute16, bcute15, Chiv5ch, GMTI, nring, QHmax, bcute11, bcute10, CIC5, CIC4, CIC6, CIC1, CIC0, CIC3, CIC2, bcutp8, bcutp9, bcutp4, bcutp5, bcutp6, bcutp7, bcutp1, bcutp2, bcutp3, ncarb, kappam1, TIAC, kappam3, kappam2, noxy, Chiv4pc, PEOEVSA3, PEOEVSA2, PEOEVSA1, PEOEVSA0, PEOEVSA7, PEOEVSA6, PEOEVSA5, PEOEVSA4, Qass, PEOEVSA9, PEOEVSA8, phi, QOmin, GATSv3, GATSv2, GATSv1, GATSv7, GATSv6, GATSv5, GATSv4, GATSv8, QCss, Chiv10, slogPVSA2, slogPVSA3, slogPVSA0, slogPVSA1, slogPVSA7, slogPVSA4, slogPVSA5, slogPVSA8, slogPVSA9, S55, S51, Gravto, EstateVSA8, Sitov, S58, Mpc, LabuteASA, radiust, MATSp8, MATSp7, MATSp6, MATSp5, MATSp4, MATSp3, MATSp2, MATSp1, nhet, nhev, AWeight, ncoi, QOmax, ncof, Hy, Smin0, QHss, Getov, VSAEstate3, knotpv, Weight, VSAEstate7, EstateVSA10, VSAEstate8, UI, VSAEstate9, mChi1, ATSv3, ATSv2, ATSv5, nsb, petitjeant, QOss, ATSv4, Smax78, ATSv7, Smax72, ATSv6, Smin46, diametert, Scar, Smin49, Smin48, ISIZ, GATSm7, EstateVSA9, J, Tnc, QHmin, S78, Tigdi, Smin36, Smin37, Smin34, Smin35, S62, Smin33, Smin31, nhyd, Smin39, ATSe1, ATSe2, ATSe3, ATSe4, ATSe5, ATSe6, ATSe7, ATSe8, S72, Hatov, Smax52, Smin12, S74, Mac, Arto, IC3, IC2, IC1, IC0, IC6, IC5, IC4, Smin6, Smin3, Tsch, Smin1, Smin8, Smin9, AW, MZM2, MZM1, nrot, Sito, Smax46, Chi10, bcutm16, bcutm15, bcutm14, bcutm13, bcutm12, bcutm11, bcutm10, BalabanJ, BertzCT, Chi0, Chi0n, Chi0v, Chi1, Chi1n, Chi1v, Chi2n, Chi2v, Chi3n, Chi3v, Chi4n, Chi4v, EState\_VSA1, EState\_VSA10, EState\_VSA11, EState\_VSA2, EState\_VSA3, EState\_VSA4, EState\_VSA5, EState\_VSA6, EState\_VSA7, EState\_VSA8, EState\_VSA9, ExactMolWt, FractionCSP3, HallKierAlpha, HeavyAtomCount, HeavyAtomMolWt, Ipc, Kappa1, Kappa2, Kappa3, LabuteASA, MaxAbsEStateIndex, MaxAbsPartialCharge, MaxEStateIndex, MaxPartialCharge, MinAbsEStateIndex, MinAbsPartialCharge, MinEStateIndex, MinPartialCharge, MolLogP, MolMR, MolWt, NHOHCount, NOCount, NumAliphaticCarbocycles, NumAliphaticHeterocycles, NumAliphaticRings, NumAromaticCarbocycles, NumAromaticHeterocycles, NumAromaticRings, NumHAcceptors, NumHDonors, NumHeteroatoms, NumRotatableBonds, NumSaturatedHeterocycles, NumSaturatedRings, NumValenceElectrons, PEOE\_VSA1, PEOE\_VSA10, PEOE\_VSA11, PEOE\_VSA12, PEOE\_VSA13, PEOE\_VSA14, PEOE\_VSA2, PEOE\_VSA3, PEOE\_VSA4, PEOE\_VSA6, PEOE\_VSA7, PEOE\_VSA8, PEOE\_VSA9, RingCount, SMR\_VSA1, SMR\_VSA10, SMR\_VSA2, SMR\_VSA3, SMR\_VSA4, SMR\_VSA5, SMR\_VSA6, SMR\_VSA7, SMR\_VSA9, SlogP\_VSA1, SlogP\_VSA10, SlogP\_VSA11, SlogP\_VSA12, SlogP\_VSA2, SlogP\_VSA3, SlogP\_VSA4, SlogP\_VSA5, SlogP\_VSA6, SlogP\_VSA8, TPSA, VSA\_EState10, VSA\_EState8, VSA\_EState9, fr\_Al\_COO, fr\_Al\_OH, fr\_Al\_OH\_noTert, fr\_Ar\_N, fr\_Ar\_NH, fr\_Ar\_OH, fr\_COO, fr\_COO2, fr\_C\_O, fr\_C\_O\_noCOO, fr\_Imine, fr\_NH0, fr\_NH1, fr\_NH2, fr\_Ndealkylation1, fr\_Ndealkylation2, fr\_Nhpyrrole, fr\_aldehyde, fr\_alkyl\_halide, fr\_allylic\_oxid, fr\_amide, fr\_aniline, fr\_aryl\_methyl, fr\_benzene, fr\_bicyclic, fr\_epoxide, fr\_ester, fr\_ether, fr\_guanido, fr\_halogen, fr\_imidazole, fr\_ketone, fr\_ketone\_Topliss, fr\_lactone, fr\_methoxy, fr\_morpholine, fr\_oxime, fr\_para\_hydroxylation, fr\_phenol, fr\_phenol\_noOrthoHbond, fr\_piperdine, fr\_piperzine, fr\_priamide, fr\_pyridine, fr\_quatN, fr\_sulfide, fr\_unbrch\_alkane, fr\_urea,

1. Science is what we understand well enough to explain to a computer. Art is everything else we do. [↑](#footnote-ref-1)