In silico studies on recreational drugs: 3D-QSAR prediction of classified and de novo designer benzodiazepines

Authors:

Valeria Catalani[1], Giuseppe Floresta[2], Michelle Botha[1], John Martin Corkery[1], Amira Guirguis[1,3], Alessandro Vento[4], Vincenzo Abbate[2], Fabrizio Schifano[1]

Affiliations:
1. Psychopharmacology, Drug Misuse & Novel Psychoactive Substances Research Unit, School of Life & Medical Sciences, University of Hertfordshire, College Lane Campus, Hatfield, United Kingdom
2. Department of Analytical, Environmental and Forensic Sciences, King's College London, London, United Kingdom
3. Swansea University Medical School, The Grove, Swansea University, Singleton Park, Swansea, United Kingdom
4. Guglielmo Marconi University, Department of Psychology, 00193, Rome, Italy

Corresponding author:

Valeria Catalani
v.catalani@herts.ac.uk

Contact emails:

giuseppe.floresta@kcl.ac.uk
vincenzo.abbate@kcl.ac.uk
m.botha@herts.ac.uk
j.corkery@herts.ac.uk
amira.guirguis@swansea.ac.uk
alessandrovento@gmail.com
f.schifano@herts.ac.uk

Running title: Recreational drugs and in silico studies

Abstract

Currently, increasing availability and popularity of designer benzodiazepines (DBZDs) constitutes a primary threat to public health. To assess this threat, the biological activity/potency of DBZDs was investigated using in silico studies.

Specific Quantitative Structure Activity Relationship (QSAR) models were developed in Forge™ for the prediction of biological activity (IC_{50}) on the γ-aminobutyric acid A receptor (GABA-AR) of previously identified classified and unclassified DBDZs. A set of new potential ligands resulting from scaffold hopping studies conducted with MOE® was also evaluated.

Two generated QSAR models (i.e., 3D-field QSAR and RVM) returned very good performance statistics (r^2= 0.98 (both) and q^2 0.75 and 0.72 respectively). The DBZDs predicted to be the most active were flubrotizolam, clonazolam, pynazolam and flucotizolam, consistently with what reported in literature and/or drug discussion fora. The scaffold hopping studies strongly suggests that replacement of the pendant phenyl moiety with a five-membered ring could increase biological activity and highlight the existence of a still unexplored chemical space for DBZDs.

QSAR could be of use as a preliminary risk assessment model for (newly) identified DBZDs, as well as scaffold hopping for the creation of computational libraries that could be used by regulatory bodies as support tools for scheduling procedures.

Introduction

In November 2021, the United Nations Office on Drugs and Crime (UNDOC) reported the persistence of benzodiazepines-type new psychoactive substances (NPS) on illegal markets, as observed in the 2020 trends (UNODC, 2020b, 2021a, 2021c). In fact, an increase in the availability and popularity of DBDZs has been reported across Europe (EMCDDA, 2021a), with 1240 seizures (5% of the NPS total) reported by Member States (EMCDDA, 2021c).These benzodiazepines, also known as designer benzodiazepines (DBZDs), were identified in 68% of toxicological cases related to NPS (clinical admission, drug-facilitated sexual assault, driving under the influence (DUI)), and accounted for 49% of all instances of NPS within a post-mortem setting, qualifying as 'a current primary threat' (UNODC, 2021a).

Designer benzodiazepines started to appear on the market in 2007 when phenazepam was first identified (Greenblatt & Greenblatt, 2019), but it was only in 2012 with the identification of pyrazolam, that the DBZDs phenomenon started (EMCDDA, 2021d; Orsolini et al., 2020). Since then, a total of 31 DBZDs have been reported to the UNODC Early Warning Advisory (UNODC, 2021b) and the EMCDDA Early Warning System (EMCDDA, 2021a), more than 80% of which was detected for the first time between 2014 and 2020.

DBDZs are structural analogues of common benzodiazepines (e.g., alprazolam, diazepam, lorazepam) which act as central nervous system (CNS) depressants through the positive allosteric modulation of the γ-aminobutyric acid (GABA)-A receptor (GABA-AR) (Griffin, Kaye, Rivera Bueno, & Kaye, 2013). Structural modifications of the benzodiazepine scaffold, i.e., addition of a third pentameric fused ring (triazole, imidazole benzodiazepines); replacement of the benzo moiety (thiazole and oxazole); interchanging of nitrogen position on the diazepine ring (1,4-; 1,5-; 2,3-diazepine core), defined six different substructural DBZD groups (Supplementary material Figure S1). Of these, only three core structures have recently been reported among the DBZDs identified (1,4-benzodiazepines (45%), thienotriazolodiazepines (17%), and triazolobenzodiazepines (38%) (UNODC, 2021a) showing less chemical variability compared to other classes of NPS classes (e.g., synthetic cannabinoids and synthetic opioids).

While classical benzodiazepines are used in medicine as hypnotics, sedatives, anxiolytics, anticonvulsant, and muscle relaxants, DBZDs are usually ingested for self-medication; sedative/hypnotic recreational use (Orsolini et al., 2020); potentiation of other sedatives, primarily opioids, effects; promotion of 'come down' after stimulant use; or unintentionally, as counterfeits of prescription benzodiazepines (ACMD, 2020b). Low prices, ease of purchase (e.g., online vendors; without prescription), ease of use, and high availability can be considered as the main reasons for their popularity and continued emergence (EMCDDA, 2021b; Orsolini et al., 2020).

The health concerns associated with these molecules have recently increased (ACMD, 2020a; EMCDDA, 2021a, 2021b). Although some DBZDs show the toxicological profiles of classic BZDs (i.e., increased reaction time, poor motor coordination, anterograde amnesia, restlessness, delirium, aggression, depression, hallucinations, paranoia, and fatalities), for most of them, safety/toxicity profiles are not yet described (Greenblatt & Greenblatt, 2019; Orsolini et al., 2020). This lack of pharmacological data represents a serious health threat, with unforeseeable risks, particularly so in relation to DBZDs use in polydrug consumption scenarios (Carpenter, Murray, Dunkley, Kazzi, & Gittinger, 2019; UNODC, 2017, 2020a, 2020b, 2020c, 2021a).

To assess the possible threat associated with DBZDs, prediction and evaluation of their biological activity/potency should be considered. Computational models, and in particular the quantitative structure-activity relationship (QSAR), have been proven helpful in providing a fast and cost-effective first evaluation of the activity of unknown molecules on a known biological target. In fact,

these models are well established in the pharmaceutical field (Bajorath, 2015; Leelananda & Lindert, 2016; Tian et al., 2015) as very successful tools for drug discovery and development (Valerio & Choudhuri, 2012). In addition, they have been extensively applied to NPS studies as important resources for the preliminary (risk) assessment of unknown molecules (Alam & Khan, 2017; Artemenko et al., 2009; Catalani et al., 2021; Durdagi et al., 2007; Ellis, Kruhlak, Kim, Hawkins, & Stavitskaya, 2018; Floresta & Abbate, 2021; Floresta, Rescifina, & Abbate, 2019; Waters, Manchester, Maskell, Haegeman, & Haider, 2018; Zhang, An, Hu, & Xiang, 2007).

Here, we report three QSAR models (3D field QSAR and two machine-learning) developed in Forge™ (Cresset, 2021) for the prediction of previously identified classified and unclassified DBZDs (Catalani et al., 2021) and of a new set of potential ligands resulting from scaffold hopping studies conducted with MOE® (Chemical Computing Group ULC, 2021).

Methods

Computational models

Two software were used for the computational analysis, Forge™ (Cresset, 2021) developed in England by Cresset and MOE® 2020.00901 developed in Canada by the Chemical Computing Group ULC (Chemical Computing Group ULC, 2021).

Biological activity data

The relevant biological data were retrieved from the literature to create the data-set for the QSAR studies (Hadjipavlou-Litina & Hansch, 1994; Waters et al., 2018). Information used as biological activity was the logarithm of the reciprocal concentration (log 1/c), c being the molar inhibitory concentration (IC_{50}) required to displace 50% of [3H]-diazepam from the rat cerebral cortex. BZDs with provisional log1/c values or atypical atoms or substituents (Tanimoto coefficient (Bajusz, Rácz, & Héberger, 2015)) were not taken into consideration. A total of 77 molecules was identified .These are reported in the Supplementary material (Table S1).

QSAR

The two-dimensional structures of the data-set were built using ChemDraw 20.1.1 (PerkinElmer Inc.). These were prepared with the wash function in MOE®; their predominant protonation state was calculated at neutral pH, and the molecular force field Amber10:EHT was used to rebuild and optimise their 3D geometry.

The optimised 3D structures were uploaded to Forge™ (v10.4.2, Cresset, New Cambridge House, Hertfordshire, UK) for the 3D-QSAR studies. The 77 structures were split according to their log1/c value into a training set (65) and a test set (12) (Table S1). The aim was to obtain two sets, both representative of the activity values space in analysis. The log1/c values (space) ranged from a maximum of 8.92 to a minimum of 6.05. The splitting process was carried out multiple times and all the resulting combinations were used to assess reproducibility.

A set of field points identifying the complex three-dimensional electrostatic/van der Waals properties was calculated in Forge™ with the application of the extended electron distribution force field (XED). In particular, electrostatically positive and negative, van der Waals attractive and hydrophobic features of the molecules were explored. Each feature was identified with a sphere, the size of which was proportional to how much that feature energetically influence the biological activity.

All 77 3D structures were aligned in Forge™ on the previously reported active conformations of diazepam and alprazolam in the crystallised protein (RCSB PDB, 2018b, 2018a), and then submitted to the Forge™ processing application (i.e. conformation search, alignment, and model building calculations).

Three different model building calculations were used and a 3D-field QSAR, a Random Forest and a Relevant Vector Machine models (RVM) were created. Detailed information on the methodology is presented in the Supplementary material (Figures S2-S4).

Scaffold hopping

To investigate the structural/chemical space of benzodiazepine-like NPS, the Fragment application tools in MOE® (v 2020.0901, Chem Comp Group, Montreal, Canada) were used. In particular the MedChem, Scaffold Replacement, and Add Group to Ligand transformation functions were used for the purpose. MedChem transformations applies a set of transformation rules, or bio-isosteric replacements, to existing ligands with the aim of generating/discovering novel chemical structures. Typical transformations tend to exchange functional groups or alter all or part of individual rings while preserving the rest of ligand. These transformations can be applied iteratively and generate cumulative changes (Langdon, Ertl, & Brown, 2010). Scaffold Replacement consists in replacing a portion (the scaffold) of a known compound, while preserving the remaining chemical groups. This aims at obtaining improved or more potent ligands. Add Group to Ligand has the same scope but instead of exchanging scaffold moiety, extends the ligand with the use of a linker database. For the scope of this exercise, the crystallized structure of GABA-AR (PDB6HUP (RCSB PDB, 2018b)) was retrieved from the Protein Data Bank (PDB ("RCSB PDB: Homepage," 2021)) and used. The transformations were performed on the 3D structure of the co-crystallized diazepam. Diazepam was chosen as it is the starting DZP, as already available as co-crystallised ligand, and because in 6HUP, diazepam is bound in its active conformation (RCSB PDB, 2018b) – see Figure 1. This enabled the steric hindrance of the binding pocket (van der Waals interaction surface) and its electrostatic properties to be taken into consideration and included in the scaffold hopping studies.

Three major moieties were identified in the diazepam scaffold and used for the MedChem and Scaffold Replacement studies, as shown in Figure 2. The Add Group to Ligand function was used only subsequently to the Med Chem and Scaffold Replacement studies, as the purpose of the study was to explore diverse chemical structures rather than growing existing ones. The MOE® proprietary linker database containing 46000 linkers, was used. MOE® default descriptor parameters were used to constrain the search, and generated structures were energy-minimised.

Biological activity prediction

The three models generated in Forge™ were used to predict the activity of two sets of molecules. The first set was a total of 102 classified/unclassified DBZDs identified on line (the majority of which was taken from www.isomerdesign.com) and described in previous publications by our research group (Catalani et al., 2021, 2022). These are presented in Table S2. The second set was the total of the DBZDs resulting from the MedChem and Scaffold Replacement transformations. All these molecules were imported and aligned in Forge™ and their putative biological activity predicted.

Results and Discussion

QSAR models statistical analysis

The data-set obtained from the literature (77 molecules) included 1,4-benzodiazepines, thienotriazolo-benzodiazepines, triazolo-benzodiazepines, and imidazo-benzodiazepines (Hadjipavlou-Litina & Hansch, 1994). These were split into a training (65) and a test set (12) (Golbraikh & Tropsha, 2000) – see Table S1 in Supplementary material.

The 3D Field QSAR was generated via a partial least squares (PLS) analysis (Wold, Trygg, Berglund, & Antti, 2001), specifically with the use of the SIMPLS algorithm (de Jong, 1993). The number of

PLS components defined as optimal (8) was identified across 20 models automatically generated, with a reported r^2 (coefficient of determination) of 0.98 and q2 (cross-validated coefficient of determination) of 0.75 as seen in Table 1. The statistics for the 20 methods are reported in the Supplementary material (Table S2).

The cross-validated coefficient of determination is the validation parameter obtained with the leave one out cross-validation (LOO CV) used in Forge™ as internal method validation, and evaluates to which degree the prediction of a model is better compared to a null one (Golbraikh & Tropsha, 2002). To assess further the reliability of the method, the measurements of the root mean squared error (RMSE) forecast are reported, and in particular a value of 0.34 was identified for the eight components final QSAR model. The external validation was conducted on the test set (12 molecules), obtaining an r^2 of 0.82. Of the two machine-learning models, the RVM performed better than the Random Forest as per statistics in Table 1. The performance of these methods is further described by Kendall's Tau, which is a variable that predicts the ability of a model to rank molecules (Kendall, 1938). Tau values range from zero to one, where values closer to one identified more predictive models (McLellan, Ryan, & Breneman, 2011).

A visual representation of the predicted vs. experimental values for training and test sets is reported in Figure 3. The 3D Field QSAR analysis returned a linear relationship between the descriptors and the activity and provided a visual interpretation of the QSAR model (Figure 4).

In the 3D map (Figure 4) the electrostatic (blue and red), and the hydrophobic (green and violet) features identified as important to the biological activity are represented. In particular the red and the blue shapes indicate the space around the molecule in which more positive electrostatic interaction (red) or more negative electrostatic interaction (blue), will be beneficial (i.e. increase) for the activity. More positive interactions (red) could mean that putting strong H-bond donors in that region improves the activity or could mean as well that putting strong H-bond acceptors will worsen the activity, and vice versa with the blue. This means that a triazolo ring fused to the core scaffold (i.e strong H-bond acceptor) in correspondence of the big blue negative patch as seen with alprazolam, will see a positive contribution to the biological activity. The same applies for those DBZDs similar to clonazepam showing a carbonyl substitution on the C2. Substitution with a NO_2 group in position C7 (e.g. meclonazepam) instead match both the negative and positive electrostatic interactions (Figure 4).

The green and the violet areas instead indicate how the presence of a hydrophobic interaction in that region would increase (green) or decrease (violet) the activity. It is interesting to note how the hydrophobic features identified by the 3D Field QSAR model are in line with the electrostatic surface derived via the receptor study with MOE® (Figure 1). Relevant hydrophobic interaction areas are identified around the pendant phenyl ring and in correspondence of a meta substituent. DBZDs showing halogenated substitution in meta (e.g. brotizolam and etizolam) indeed show a higher biological activity when compare with less hydrophobic substituted molecules. Another hydrophobic interaction area is identified between the acceptor/donor features and the molecule scaffold in correspondence of C7. This may suggest an increased activity when a strong H-bond donor is connected to the core structure via a short aliphatic chain, instead of a direct bond with the carbon atoms of the scaffold.

Due to the poor statistics obtained for the Random Forest model (i.e. r^2 Test value < 0.5, Table 1), this was not used to predict the biological activity of both unclassified and de novo DBZDs. This aspect is still under investigation; however , great discordance between the training and test r^2 values when using machine-learning algorithms have been reported before, due to the tendency of the latter to overfitting the training model, or when working with small data-sets (Brownlee, 2018).

Classified / unclassified DBZDs prediction

The two QSAR methods identified above were used to predict the biological activity of a set of DBZDs identified online (the majority of which were retrieved from www.isomerdesign.com) (DAMICON, 2017). Their putative activity was previously assessed with the use of a 2D-QSAR model developed with MOE®, whose predictive statistic were indeed good but inferior if compared to the 3D models described here (Catalani et al., 2021). This may suggest the lack of a strong correlation between the 2D properties of the benzodiazepines scaffold and their activity, making the evaluation of the 3D properties mandatory to obtain more reliable predictions (higher values of r^2, q^2 and test r^2). The predicted biological activities are presented in the Supplementary material (Table S3). The DBZDs showing the highest biological activity were: flubrotizolam (log $1/c$ =9.6), clonazolam (9.5), pynazolam (9.4) and flucotizolam (9.1). Flunitrazolam (8.8) and flubromazolam (8.7) followed with slightly lower values, together with other DBZDs. These results seem to be in line with what is reported in the literature and some drug discussion fora (El Balkhi, Monchaud, Herault, Géniaux, & Saint-Marcoux, 2020; reddit, 2021). The QSAR methods predicted a strong biological activity for some DBZDs that have recently been recommended for placing under Schedule 1 of the Misuse of Drugs Regulations 2001 (ACMD, 2020a), and in particular flunitrazolam (8.8) and flualprazolam (8.3) . It is interesting to note that the presence of a triazole ring seems to be a structural feature that consistently increases the activity of the index molecule.

Scaffold hopping

To assess the possibility of enlarging the chemical diversity of designer benzodiazepines, MedChem and Scaffold Replacement studies were conducted. Using the sole steric hindrance of the receptor's binding pocket for each of the transformations (Figure 2), roughly 4000 results were generated. To generate more informed results, the electrostatic properties of both the receptor and the ligand were taken into consideration using the pharmacophore editor function in MOE®. In particular, for the pendant phenyl an aromatic/hydrophobic pharmacophore feature was created at the centre of the ring, as well as a feature to highlight the hydrophobic portion identified in the receptor pocket in proximity of Tyr160 and Tyr210 (Figure 5a) (Richter et al., 2012). In fact, the aromatic feature of the pendant phenyl engages in an arene-hydrogen bond with His102, an interaction proven essential for receptor activation (Amundarain, Caffarena, & Costabel, 2021; Wieland, Luddens, & Seeburg, 1992). For the benzene ring fused to the diazepine, an aromatic/hydrophobic feature at the centre of the ring was maintained, the importance of which was previously reported (Catalani et al., 2021; Davies, Bateson, & Dunn, 2002; Sigel & P. Luscher, 2012; Sigel, Schaerer, Buhr, & Baur, 1998) (Figure 5b). The hydrophobic feature in proximity of the chlorine atom, the presence of which has been proven to increase the activity, was also kept for the scaffold hopping studies. The acceptor feature on the C=O of the diazepine ring was retained together with the feature identified by the oxygen lone pair projection on the receptor in proximity of Ser205and Ser206 (Figure 5C).

The transformation studies returned a total of 477 entries (364 Scaffold replacement and 113 MedChem, respectively). For these structures, inclusion in the applicability domain (AD) was automatically assessed by Forge via the "distance to model" function. Those structures outside the AD (57% of the Scaffold Replacement) were discarded. Only the top scoring modifications are reported in Figure 6 .

Among the top scoring moieties, some well-known DBZD scaffolds were identified/suggested by MOE®, in particular the triazole (Figure 6h), and the thieno (6d) moieties. According to UNODC reports, these can be found among the majority of DBZDs recently identified (e.g. clonazolam, etizolam, flualprazolam, flubromazepam, flubromazolam) (UNODC, 2020b, 2021a). This finding is extremely important because it suggests that computational studies are reliable and could be used to

predict future NPS scaffolds together with their biological activity. The MedChem/Scaffold Replacement transformations that returned molecules with the highest values of predicted log1/c were those on the pendant phenyl ring. In particular, the latter was substituted by another aromatic five-membered ring (matching the pharmacophore feature previously identified) (Catalani et al., 2021). The five-membered rings showing the highest predicted activity values were the 1,3,4-triazole (Figure 6a, 6b), and the imidazole (6c). Interestingly, the binding pocket cavity that accommodates the pendant phenyl is very narrow and leaves very little room for chemical modification/growth. Indeed, SAR studies in the literature report how meta- and para- substitutions (small groups) on the ring are not beneficial for the activity (Davies et al., 2002; Hadjipavlou-Litina & Hansch, 1994). The suggestion/prediction that the use of a smaller aromatic ring could increase the activity is indeed very interesting for further chemical space investigation. As per Figure 7a, the five-membered ring seems to still be able to engage in the hydrophobic interaction with His102. Moreover, the reduced size of the pendant moiety gives opportunity for a greater number of ring substitutions, either in position two or five of the triazole. It should be noted that the substitution with bulky chain (hydrophobic more favourable) is preferred only at position 2 due to the steric hindrance of the receptor creating a very tight pocket near position 5 (Figure 7b).

Further substitution in position 2 was explored with the Add Group to Ligand function in MOE®. A very high number of entries (> 2000) was generated and analysed. Among them, some high (log 1/c =>9.5) biological activities were predicted. Examples are shown in Figure 7c.

A set of halogenated substituents (that is, Br, Cl, F, CF_3, etc.) in R_1 (Figure 7c) was also investigated but returned lower predicted biological activities (log 1/c =<8.0). The results obtained suggest that changes in the benzo or diazepine moieties of the molecule may have a smaller impact on biological activity when compared to modification of the pendant phenyl. Indeed, the new scaffolds for series 1-2 (Figure 2) show log1/c values lower than 9, suggesting still a potent activity profile but in line with the DBZDs currently on the market. A new scaffold for series 3 instead suggests the possibility of creating very potent and potentially more harmful DBZDs.

Limitations

The major limitation is represented by the size of the data-sets (training and test) used for the computational studies. Indeed, for a robust QSAR model, the data-set should consist of roughly 100 entries (Fourches, Muratov, & Tropsha, 2010; Golbraikh, Muratov, Fourches, & Tropsha, 2014). However, to the best of our knowledge, no other experimental data comparable to the ones used here are available in the literature. That is, no other IC_{50} values were found for benzodiazepines-like structures at the GABA-AR, calculated as displacement of 50% of [3H]-diazepam.

Other limitations include the use of only one crystallised structure of the GABA-AR; and the use of force field methodology only for the energy minimisation of the molecules analysed, which could give slightly less accurate conformations if compared to semi empirical calculations.

Conclusions

In line with the latest international reports (EMCDDA, 2021a; UNODC, 2021a), DBZDs represent a chemical class of strong interest for drug misusers (UNODC, 2020b). In fact, DBDZs are increasingly identified in polydrug consumption scenarios, either with stimulants or other CNS depressants (e.g. synthetic or classic opioids). The consumption of two substances (e.g. DBZDs mixed with novel

synthetic opioids (NSOs) or traditional opioids, and vice versa), displaying similar effect on the CNS (i.e. depressant), can cause strong enhancement of both their sought after and adverse effects, and could lead to very severe side-effects as respiratory depression, coma and death. Examples could include diazepam, alprazolam and clonazepam used with fentanyl analogues or nitazenes, or heroin in combination with clonazolam, etizolam. The threat associated with polydrug consumption (in particular opioids and benzodiazepines, whether novel or not) is actual and is even more worrisome if one considers that for the majority of NPS, which are constantly introduced on the market, very little data on their safety/toxicity profile is available (El Balkhi et al., 2020).

Hence, it is extremely important to assess these new molecules and more so with regard to their pharmacology. The use of QSAR methodology has proven very helpful in doing so (Catalani et al., 2021; Floresta & Abbate, 2021; Floresta, Apirakkan, Rescifina, & Abbate, 2018; Floresta et al., 2019; Waters et al., 2018), especially for NPS. The latter, despite their structural similarity, can indeed display very different biological activity profiles. The 3D-QSAR models identified here seem to be very reliable in their predictive power. They identified as most potent, DBZDs such as flubromazolam, clonazolam, pynazolam and flucotizolam, which were indeed reported as such both in the scientific literature and by users. Moreover, the molecule predicted as the most potent, flubrotizolam, is a new DBZDS for which no data is available in literature (to the best of our knowledge). This finding in particular underscores the importance and the need for these computational models to be used as preventive and informative tools.

Indeed, these models could be used to assess, in a rapid and cost-effective way, the biological activity profile of a new DBZD, as soon as the latter is identified on the illegal market. Moreover, they could be of use to define activity differences, of different DBZDs, on the GABA-AR receptor. The results obtained with the scaffold hopping exercise are also very promising because, despite suggesting the existence of a wider chemical landscape for this NPS class, they could be used as an effective tool in the prediction of the latter. It should be noted that among the new structures generated are some very well-known and potent scaffolds. According to the predicted biological activity values, further modifications to the classical core structure could significantly increase the biological activity of an index molecule, and hence they need to be carefully investigated.

Finally, this exercise could be considered as the first step towards the creation of computational libraries that could be used by regulatory bodies as support tools for risk assessment and scheduling procedures.

Authors' contributions

VC carried out the planning of the manuscript and drafted the paper. GF, VA, MB, AG, JC, AV and FS critically reviewed the final draft prior to submission. All authors have read and agreed to the published version of the manuscript.

Data and Software Availability

All data are available in a machine readable format on request. MOE® and Forge™ are commercial software for which a trial free version is available upon request.

Funding

Conflicts of Interest

The authors declare no conflict of interest.

References

ACMD. (2020a). Addendum to ACMD's report on novel benzodiazepines. Retrieved January 31, 2022, from https://www.gov.uk/government/publications/novel-benzodiazepines-prevalence-and-harms-in-the-uk/addendum-to-acmds-report-on-novel-benzodiazepines

ACMD. (2020b). *Novel Benzodiazepines A review of the evidence of use and harms of Novel Benzodiazepines*. Retrieved from https://www.gov.uk

Alam, S., & Khan, F. (2017). 3D-QSAR studies on Maslinic acid analogs for Anticancer activity against Breast Cancer cell line MCF-7. *Scientific Reports*, 7(1). https://doi.org/10.1038/S41598-017-06131-0

Amundarain, M. J., Caffarena, E. R., & Costabel, M. D. (2021). How does α1Histidine102 affect the binding of modulators to α1β2γ2 GABAA receptors? molecular insights from in silico experiments. *Undefined*, 23(6), 3993–4006. https://doi.org/10.1039/D0CP05081D

Artemenko, A. G., Kuz'Min, V. E., Muratov, E. N., Polishchuk, P. G., Borisyuk, I. Y., & Golovenko, N. Y. (2009). Influence of the structure of substituted benzodiazepines on their pharmacokinetic properties. *Pharmaceutical Chemistry Journal*, 43(8), 454–462. https://doi.org/10.1007/s11094-009-0332-x

Bajorath, J. (2015). Computer-aided drug discovery. *F1000Research*, Vol. 4. https://doi.org/10.12688/f1000research.6653.1

Bajusz, D., Rácz, A., & Héberger, K. (2015). *Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? 7*, 20. https://doi.org/10.1186/s13321-015-0069-3

Brownlee, J. (2018). The Model Performance Mismatch Problem (and what to do about it). Retrieved February 24, 2022, from https://machinelearningmastery.com/the-model-performance-mismatch-problem/

Carpenter, J. E., Murray, B. P., Dunkley, C., Kazzi, Z. N., & Gittinger, M. H. (2019). Designer benzodiazepines: a report of exposures recorded in the National Poison Data System, 2014–2017. *Clinical Toxicology*, 57(4), 282–286. https://doi.org/10.1080/15563650.2018.1510502

Catalani, V., Botha, M., Corkery, J. M., Guirguis, A., Vento, A., Scherbaum, N., & Schifano, F. (2021). The psychonauts' benzodiazepines; quantitative structure-activity relationship (QSAR) analysis and docking prediction of their biological activity. *Pharmaceuticals 2021, Vol. 14, Page 720*, 14(8), 720. https://doi.org/10.3390/PH14080720

Catalani, V., Botha, M., Corkery, J. M., Guirguis, A., Vento, A., & Schifano, F. (2022). Designer benzodiazepines' activity on opioid receptors: a docking study. *Current Pharmaceutical Design*, 28. https://doi.org/10.2174/1381612828666220510153319

Chemical Computing Group ULC. (2021). *Molecular Operating Enviroment (MOE), 2020.0901*. Retrieved from https://www.chemcomp.com/index.htm

Cresset. (2021). *Forge for ligand-based molecular design and structure-activity relationships*. Retrieved from https://www.cresset-group.com/software/forge/

DAMICON. (2017). NPSfinder. Retrieved September 4, 2020, from https://www.npsfinder.com/

Davies, M., Bateson, A. N., & Dunn, S. M. J. (2002). Structural Requirements for Ligand Interactions at the Benzodiazepine Recognition Site of the GABAA Receptor. *Journal of Neurochemistry*, 70(5), 2188–2194. https://doi.org/10.1046/j.1471-4159.1998.70052188.x

de Jong, S. (1993). SIMPLS: An alternative approach to partial least squares regression. *Chemometrics and Intelligent Laboratory Systems*, 18(3), 251–263. https://doi.org/10.1016/0169-7439(93)85002-X

Durdagi, S., Kapou, A., Kourouli, T., Andreou, T., Nikas, S. P., Nahmias, V. R., … Mavromoustakos, T. (2007). The application of 3D-QSAR studies for novel cannabinoid ligands substituted at the C1′ position of the alkyl side chain on the structural requirements for binding to cannabinoid receptors CB1 and CB2. *Journal of Medicinal Chemistry*, 50(12), 2875–2885.

https://doi.org/10.1021/jm0610705

El Balkhi, S., Monchaud, C., Herault, F., Géniaux, H., & Saint-Marcoux, F. (2020). Designer benzodiazepines' pharmacological effects and potencies: How to find the information. *Journal of Psychopharmacology*, 269881119901096. https://doi.org/10.1177/0269881119901096

Ellis, C. R., Kruhlak, N. L., Kim, M. T., Hawkins, E. G., & Stavitskaya, L. (2018). Predicting opioid receptor binding affinity of pharmacologically unclassified designer substances using molecular docking. *PLoS ONE*, *13*(5). https://doi.org/10.1371/journal.pone.0197734

EMCDDA. (2021a). *New benzodiazepines in Europe – a review*. Retrieved from https://www.emcdda.europa.eu/publications/rapid-communications/new-benzodiazepines-europe-review_en

EMCDDA. (2021b). New report highlights public health risks and increasing availability of 'designer benzodiazepines' in Europe. Retrieved January 31, 2022, from https://www.emcdda.europa.eu/news/2021/6/report-highlights-public-health-risks-and-increasing-availability-designer-benzodiazepines-europe_en

EMCDDA. (2021c). The EU Early Warning System on new psychoactive substances (NPS). Retrieved October 23, 2019, from https://www.emcdda.europa.eu/publications/topic-overviews/eu-early-warning-system_en

EMCDDA. (2021d). Timeline of benzodiazepines formally notified to the EU Early Warning System for the first time, 2007–2020. Retrieved January 31, 2022, from https://www.emcdda.europa.eu/media-library/timeline-benzodiazepines-formally-notified-eu-early-warning-system-first-time-2007–2020_en

Floresta, G., & Abbate, V. (2021). Machine learning vs. field 3D-QSAR models for serotonin 2A receptor psychoactive substances identification. *RSC Advances*, *11*(24), 14587–14595. https://doi.org/10.1039/D1RA01335A

Floresta, G., Apirakkan, O., Rescifina, A., & Abbate, V. (2018). Discovery of high-affinity cannabinoid receptors ligands through a 3D-QSAR ushered by scaffold-hopping analysis †. *Molecules*, *23*(9), 2183. https://doi.org/10.3390/molecules23092183

Floresta, G., Rescifina, A., & Abbate, V. (2019). Structure-based approach for the prediction of mu-opioid binding affnity of unclassified designer fentanyl-like molecules. In *International Journal of Molecular Sciences* (Vol. 20, p. 2311). https://doi.org/10.3390/ijms20092311

Fourches, D., Muratov, E., & Tropsha, A. (2010). Trust, but verify: On the importance of chemical structure curation in cheminformatics and QSAR modeling research. *Journal of Chemical Information and Modeling*, *50*(7), 1189–1204. https://doi.org/10.1021/ci100176x

Golbraikh, A., Muratov, E., Fourches, D., & Tropsha, A. (2014). Data set modelability by QSAR. *Journal of Chemical Information and Modeling*, *54*(1), 1–4. https://doi.org/10.1021/ci400572x

Golbraikh, A., & Tropsha, A. (2000). Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. *Molecular Diversity*, *5*(4), 231–243. https://doi.org/10.1023/A:1021372108686

Golbraikh, A., & Tropsha, A. (2002). Beware of q2! *Journal of Molecular Graphics and Modelling*, *20*(4), 269–276. https://doi.org/10.1016/S1093-3263(01)00123-1

Greenblatt, H. K., & Greenblatt, D. J. (2019). Designer Benzodiazepines: A Review of Published Data and Public Health Significance. *Clinical Pharmacology in Drug Development*, *8*(3), 266–269. https://doi.org/10.1002/cpdd.667

Griffin, C. E., Kaye, A. M., Rivera Bueno, F., & Kaye, A. D. (2013). Benzodiazepine pharmacology and central nervous system-mediated effects. *Ochsner Journal*, *13*(2), 214–223.

Hadjipavlou-Litina, D., & Hansch, C. (1994). Quantitative Structure–Activity Relationships of the Benzodiazepines. A Review and Reevaluation. *Chemical Reviews*, *94*(6), 1483–1505. https://doi.org/10.1021/cr00030a002

Kendall, M. G. (1938). A New Measure of Rank Correlation. *Biometrika*, *30*(1/2), 81. https://doi.org/10.2307/2332226

Langdon, S. R., Ertl, P., & Brown, N. (2010). Bioisosteric Replacement and Scaffold Hopping in

Lead Generation and Optimization. *Molecular Informatics*, *29*(5), 366–385. https://doi.org/10.1002/MINF.201000019

Leelananda, S. P., & Lindert, S. (2016). Computational methods in drug discovery. *Beilstein Journal of Organic Chemistry*, *12*, 2694–2718. https://doi.org/10.3762/bjoc.12.267

McLellan, M. R., Ryan, M. D., & Breneman, C. M. (2011). Rank order entropy: why one metric is not enough. *Journal of Chemical Information and Modeling*, *51*(9), 2302–2319. https://doi.org/10.1021/CI200170K

Orsolini, L., Corkery, J. M., Chiappini, S., Guirguis, A., Vento, A., De Berardis, D., … Schifano, F. (2020). "New/Designer Benzodiazepines": an analysis of the literature and psychonauts' trip reports. *Current Neuropharmacology*, *18(9)*(9), 809–837. https://doi.org/10.2174/1570159X18666200110121333

RCSB PDB: Homepage. (2021). Retrieved February 4, 2021, from https://www.rcsb.org/

RCSB PDB. (2018a). 6HUO: CryoEM structure of human full-length heteromeric alpha1beta3gamma2L GABA(A)R in complex with alprazolam (Xanax), GABA and megabody Mb38. Retrieved March 8, 2021, from https://www.rcsb.org/structure/6HUO

RCSB PDB. (2018b). 6HUP: CryoEM structure of human full-length alpha1beta3gamma2L GABA(A)R in complex with diazepam (Valium), GABA and megabody Mb38. Retrieved March 8, 2021, from https://www.rcsb.org/structure/6HUP

reddit. (2021). reddit: the front page of the internet. Retrieved October 8, 2020, from https://www.reddit.com/

Richter, L., De Graaf, C., Sieghart, W., Varagic, Z., Mörzinger, M., De Esch, I. J. P., … Ernst, M. (2012). Diazepam-bound GABAA receptor models identify new benzodiazepine binding-site ligands. *Nature Chemical Biology*, *8*(5), 455. https://doi.org/10.1038/NCHEMBIO.917

Sigel, E., & P. Luscher, B. (2012). A Closer Look at the High Affinity Benzodiazepine Binding Site on GABAA Receptors. *Current Topics in Medicinal Chemistry*, *11*(2), 241–246. https://doi.org/10.2174/156802611794863562

Sigel, E., Schaerer, M. T., Buhr, A., & Baur, R. (1998). The benzodiazepine binding pocket of recombinant α1β2γ2 γ- aminobutyric acid(A) receptors: Relative orientation of ligands and amino acid side chains. *Molecular Pharmacology*, *54*(6), 1097–1105. https://doi.org/10.1124/mol.54.6.1097

Tian, S., Wang, J., Li, Y., Li, D., Xu, L., & Hou, T. (2015). The application of in silico drug-likeness predictions in pharmaceutical research. *Advanced Drug Delivery Reviews*. https://doi.org/10.1016/j.addr.2015.01.009

UNODC. (2017). Non-medical use of benzodiazepines : a growing threat to public health ? In *Global Smart Update* (Vol. 18). Vienna.

UNODC. (2020a). *Current NPS Threaths Volume II*. Retrieved from https://www.unodc.org/documents/scientific/Current_NPS_Threats_Volume_II_Web.pdf

UNODC. (2020b). *Current NPS Threats Volume III*. Retrieved from https://www.unodc.org/unodc/en/scientists/current-nps

UNODC. (2020c). News: October 2020 - UNODC SMART: Benzodiazepine-type NPS feature strongly in DUID and post-mortem cases. Retrieved January 21, 2022, from https://www.unodc.org/LSS/Announcement/Details/55c68fff-1371-4382-bc18-9acc54fd6588

UNODC. (2021a). *Current NPS Threats Volume IV*. Retrieved from https://www.unodc.org/unodc/en/scientists/current-nps-threats.html

UNODC. (2021b). Early Warning Advisory (EWA) on New Psychoactive Substances (NPS). Retrieved February 4, 2021, from https://www.unodc.org/LSS/Home/NPS

UNODC. (2021c). News: September 2021 – UNODC: Falsified Xanax® containing NPS flualprazolam and flubromazolam. Retrieved January 21, 2022, from https://www.unodc.org/LSS/Announcement/Details/bb1a14f0-75bb-499e-bba7-984c72822370

Valerio, L. G., & Choudhuri, S. (2012). Chemoinformatics and chemical genomics: potential utility of in silico methods. *Journal of Applied Toxicology*, *32*(11), 880–889.

https://doi.org/10.1002/jat.2804

Waters, L., Manchester, K. R., Maskell, P. D., Haegeman, C., & Haider, S. (2018). The use of a quantitative structure-activity relationship (QSAR) model to predict GABA-A receptor binding of newly emerging benzodiazepines. *Science and Justice*, *58*(3), 219–225. https://doi.org/10.1016/j.scijus.2017.12.004

Wieland, H. A., Luddens, H., & Seeburg, P. H. (1992). A single histidine in GABA(A) receptors is essential for benzodiazepine agonist binding. *Journal of Biological Chemistry*, *267*(3), 1426–1429. https://doi.org/10.1016/S0021-9258(18)45961-3

Wold, S., Trygg, J., Berglund, A., & Antti, H. (2001). Some recent developments in PLS modeling. *Chemometrics and Intelligent Laboratory Systems*, *58*(2), 131–150. https://doi.org/10.1016/S0169-7439(01)00156-3

Zhang, Z., An, L., Hu, W., & Xiang, Y. (2007). 3D-QSAR study of hallucinogenic phenylalkylamines by using CoMFA approach. *Journal of Computer-Aided Molecular Design*, *21*(4), 145–153. https://doi.org/10.1007/s10822-006-9090-y

Table 1 Values for the statistics obtained for the three calculated QSAR models

Model	r^2	q^2	r^2 Test	RMSE	Tau
3D Field QSAR	0.98	0.75	0.82	0.34	NA
Random Forest	0.91	0.51	0.46	0.53	0.55
RVM	0.98	0.72	0.86	0.40	0.71

Notes. Here are presented the statistic of the QSAR models generated in the form of: the coefficient of determination (r2) which indicates the goodness of fit; the cross-validated coefficient of determination (q^2) which indicates the robustness; the coefficient of determination for the test set (r^2 test), which indicates the predictive power; the root mean square error (RMSE) as reliability measure; and Tau as a further parameter to assess the predictivity of the model.

Figures legends

Figure 1 Diazepam bound in is active conformation in the GABA-AR (PDB6HUP (RCSB PDB, 2018b)).

The benzodiazepine allosteric binding pocket of the GABA-AR is presented, with the subunit α1 dentified in pink and the subuninit γ2 identified in light blue. On the left side the interactions with the binding pocket are visualised; in the middle the elettrostatic properties of the pocket and on the right the van der Waals interactions surface are presented. The elettrostatic properties are identified with three different colours: green for the hydrophobic portion, red for the H-bond acceptor and blue for the H-bond donor like portion. These were all retrieved from the analysis of the PDB6HUP crystallised structure.

Figure 2 Three major moieties (1,2,and 3) for the MedChem (green) and Scaffold Replacement (red) studies

Notes. The green moieties include the benzene ring (1), diazepine ring (2) and pendant phenyl ring (3). The red moieties include the whole benzodiazepine ring (1), the benzene ring (2) and the pendant phenyl ring (3).

Figure 3 Visual representation of the predicted (x axis) vs. experimental (y axis) log 1/c values for training (blue) and test (yellow)sets

Figure 4 Visual representation of the generated 3D Field QSAR model

Notes. The eletrostatic properties are identified by the red (positive) and blue (negative) colours, while the green and violet identify respectively areas of favorable and unfavorable hydrophobics.In particular the red and the blue shapes indicate the space around the molecule in which more positive electrostatic interaction (red) or more negative electrostatic interaction (blue), will be beneficial (i.e. increase) for the activity. More positive interactions (red) could mean that putting strong H-bond donors in that region improves the activity or could mean as well that putting strong H-bond acceptors will worsen the activity, and vice versa with the blue. The green and the violet areas instead indicate how the presence of a hydrophobic interaction in that region would increase (green) or decrease (violet) the activity Please note the colours for acceptor and donors are inverted when compared to Figure 1.
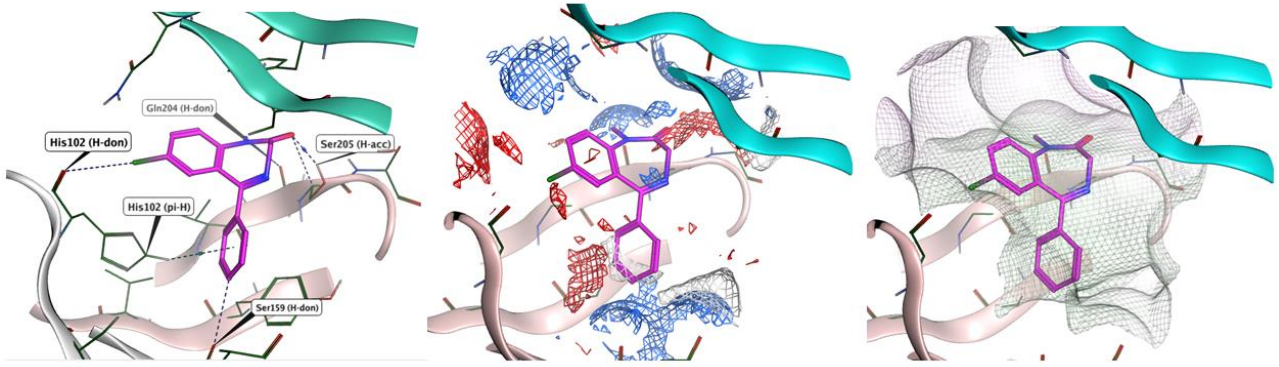
Figure 5 Ligand pharmacophore features taken into consideration for the scaffold hopping exercise

Notes. The pharmacophore features are represented by coloured spheres. The hydrophobic and aromatic features are presented in orange, while the H bond accpetor feauture is presented in light blue. The elettrostatic properties of the binding pocket are also presented: in green the hydrophobic, in red the H-bond acceptor and in blue the H-bond donor.
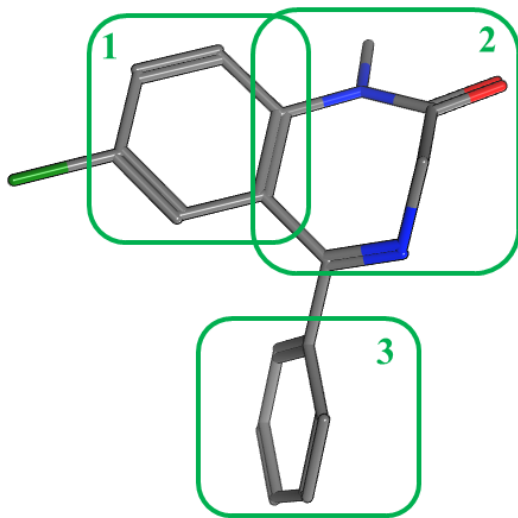
Figure 6 MedChem and Scaffold Replacement top scoring moieties

The top scoring moieties are: (a, b) triazole; (c) imidazole; (d) triazole; (e) pyridine; (f, g, h) triazolobenzodiazepines; (i) pyrimidine.
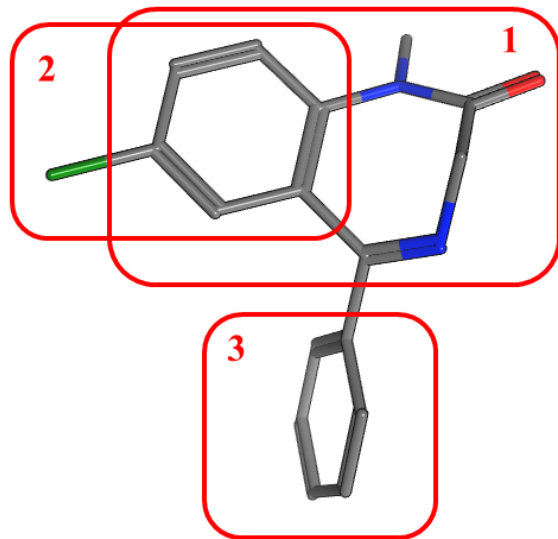
Figure 7 (a) 3D representation of the new scaffold created and its adopted conformation inside the binding pockets which maintains the hydrophobic interaction with His102; b) same representation with the addition of the vanderWaal interaction surface which shows how narrow is the portion of the pocket which host the pendant aromatic ring; c)the predicted activities retrieved from "Add Group to Ligand" exercise on the C2 of the triazole pendant ring (R_1)
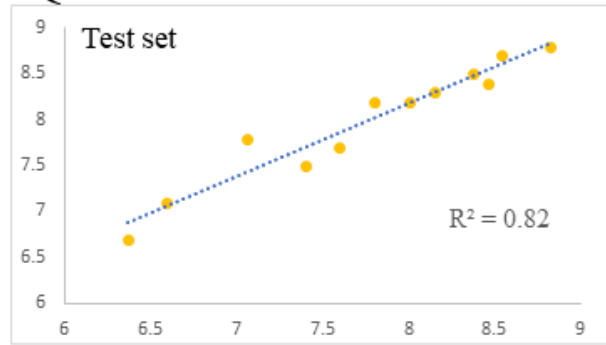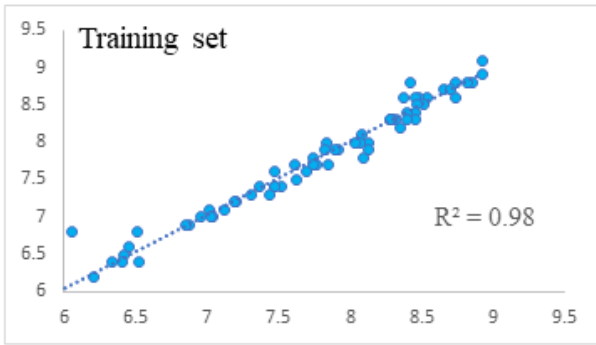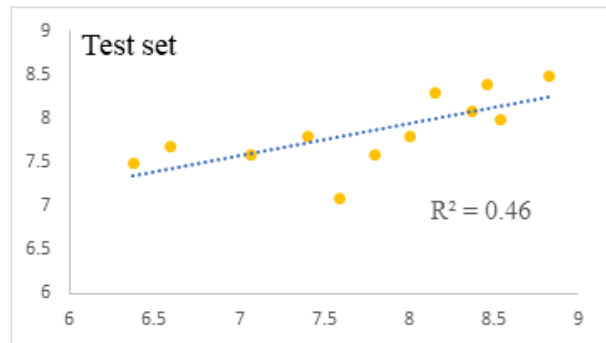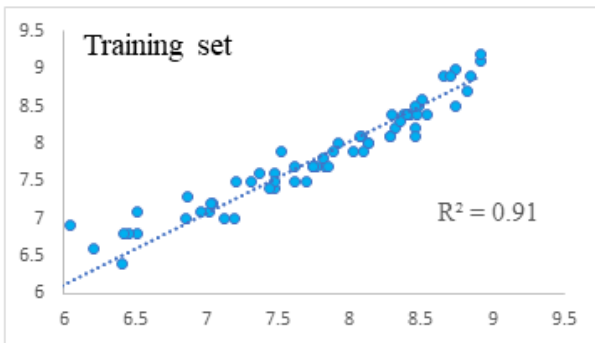
Med Chem

Scaffold Replacement

3D-Field QSAR

Training set

R² = 0.98

Test set

R² = 0.82

Random Forest

Training set

R² = 0.91

Test set

R² = 0.46

Relevance Vector Machine

Training set

R² = 0.98

Test set

R² = 0.86

Negative electrostatics
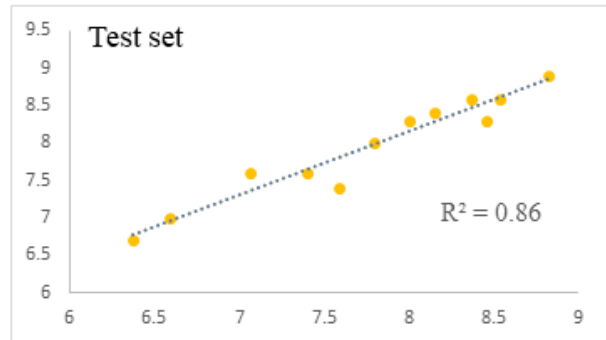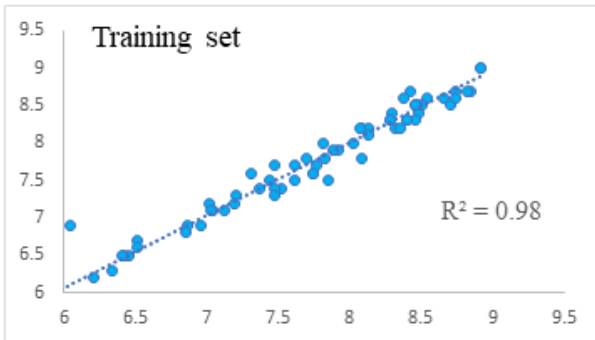
Positive electrostatics

Favorable Hydrophobics

Unfavorable Hydrophobics

a
Log 1/c= 9.5

b
Log 1/c= 9.0

c
Log 1/c= 8.7

d
Log 1/c= 8.7

e
Log 1/c= 8.7

f
Log 1/c= 8.8

g
Log 1/c= 8.7

h
Log 1/c= 8.7

i
Log 1/c= 8.6

a

b

His102

c

Log 1/c= 10.1

Log 1/c= 10.0

Log 1/c= 10.0

Log 1/c= 9.7

Log 1/c= 9.5