



Contents lists available at UGC-CARE

International Journal of Pharmaceutical Sciences and Drug Research

[ISSN: 0975-248X; CODEN (USA): IJPSPP]

Available online at www.ijpsronline.com

Research Article

Computer-Assisted Design of Benzoisoxazol derivatives Inhibitors of Bromodomain-containing Protein 4 (BRD4) with Favorable Pharmacokinetic Profile

Rika J. Kouadja¹, Logbo M. Mouss^{1*}, Koffi C. Kouman¹, Mama Nsangou², Eugene Megnassan¹¹Laboratory of Fundamental and Applied Physics, Nangui Abrogoua University, Abidjan, Côte d'Ivoire²Faculty of Science, University of Maroua, Maroua, Cameroon

ARTICLE INFO

Article history:

Received: 05 August, 2023

Revised: 08 September, 2023

Accepted: 13 September, 2023

Published: 30 September, 2023

Keywords:

ADME properties prediction, Bromodomain-containing protein 4, *In-silico* screening, QSAR models; combinatorial library, Castration-Resistant Prostate Cancer, inhibitors, Benzo[d]isoxazole, pharmacophore model.

DOI:

10.25004/IJPSDR.2023.150514

ABSTRACT

We performed a relation computed-aided design based on the structure of benzo[d]isoxazol derivatives inhibitors (BDIO) derivatives, new potent inhibitors of the BRD4 protein. By using *in-situ* modifications of the three dimensional (3D) models of BRD4-BDIOx complex (Protein Data Bank (PDB) entry code: 5Y8Z) were prepared for the training and validation sets compounds of 29 BDIOx with observed inhibitory potencies (IC_{50}^{exp}). We first built a quantitative structure activity relationship (QSAR) model in the gas phase, linearly correlating the calculated enthalpies of the BRD4-BDIOx complex formation with IC_{50}^{exp} ($pIC_{50}^{exp} = -0.1147 \times \Delta\Delta H_{MM} + 6.6022$; $R^2 = 0.80$) first and then a superior QSAR model was brought forth, correlating computed relative Gibbs' free energies of complexation and IC_{50}^{exp} ($pIC_{50}^{exp} = -0.1205 \times \Delta\Delta G_{com} + 6.9374$; $R^2 = 0.96$) which was then validated by a 3D-QSAR pharmacophore generation model (PH4) ($pIC_{50}^{exp} = 0.996 \times pIC_{50}^{pre} + 0.0554$; $R^2 = 0.95$). The structural information of the active conformation of the training set BDIOs from the models guided us in the design of a virtual combinatorial library (VCL) of 99 225 analogs. We then filtered the VCL by applying Lipinski's rule-of-five, in order to identify new BDIOs drug likely analogs. The pharmacophore (PH4)-based screening retained 106 new and potent BDIOs with predicted inhibitory potencies pIC_{50}^{pre} up to 158 times more active than the most active training set BDIO1 ($IC_{50}^{exp} = 130$ nM). Finally, the predicted pharmacokinetic profiles of the best potent of these new analogs (pIC_{50}^{pre}) were compared to current orally administered anticancer drugs. This computational approach, which combines molecular mechanics and the Poisson-Boltzmann (PB) implicit solvation theory, the pharmacophore model, the analysis of BRD4-BDIOs interaction energies, the *in-silico* screening of VCL compounds, and the inference of ADME properties resulted in a set of new suggested BRD4 inhibitors for the fight against CRPC.

INTRODUCTION

Cancer is a potentially dangerous disease that can cause an imbalance in the body, a cellular malformation or even death. This malformation is caused by transformation of regular cells into aberrant ones becoming immortal and acquiring the capability to multiply without control to a tumor. This can occur anywhere in the body, giving rise to several forms of localized cancers, such as prostate cancer, which is a real public health concern today. Prostate cancer (PCa) is the most common malignancy

and the fifth leading cause of death from malignancy in men^[1-3] evolving to Castration-Resistant Prostate Cancer (CRPC) and in most cases metastatic and incurable form of the disease.^[4,5] Despite technological progress such as androgen deprivation therapy (ADT) with chemical or surgical castration, the cancer cells escape treatment and unavoidably worsen to reach CRPC.^[4-7]

The progress in scientific research sheds light on recurrent prostate cancer after local therapy and the central role of androgen receptor (AR) signaling on cell growth.^[8]

*Corresponding Author: Mr. Logbo M. Mouss

Address: Laboratory of Fundamental and Applied Physics, Nangui Abrogoua University, Abidjan, Côte d'Ivoire

Email ✉: welakoundan@gmail.com

Tel.: +225 07-47-49-96-70

Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2023 Justin R. Kouadja *et al.* This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

Prostate tumor cells acquire resistance to ADT through multiple mechanisms that are essential for cancer progression, including aberrant androgen synthesis, AR gene amplification, AR mutations, production of constitutively active AR splice variants and alternative steroid receptor.^[9,10] Bromodomain and extraterminal domain (BET) proteins (BRD2, BRD3, BRD4, and testis-specific BRDT) are a family of chromatin-associated proteins that regulate gene expression by acting as epigenetic readers through their ability to detect and bind to acetylated lysine residues on histone tails.^[11,12] Recent studies have shown that BET are promising therapeutic targets due to their ability to regulate tumor progression. According to this approach, small-molecule pan-BET inhibitors, such as JQ1^[13] and I-BET15,^[14] have shown great promise as therapeutic agents across a diverse array of human malignancies,^[15,16] including CRPC.^[17,18] However, even pan-BET inhibitors and degraders become more widespread in their investigational and clinical use.^[19,20] Accumulating evidence shows that pan-BET therapies can produce off-target effects, including the reactivation of latent HIV in infected T cells,^[21] and obscure the biology of each independently acting BET protein.^[22] This caution underscores the need for more BET family member-selective chemical intervention.^[23] Therefore, targeting BET proteins is of major interest and an alternative strategy for treating CRPC.^[24] Our goal in this work, is to design new analogs of benzo[d]isoxazole derivatives potentially more active than those currently reported by Zhang *et al.* against BRD4 (BDIO1, $IC_{50}^{exp} \geq 130$ nM).^[25-28] With the help of a structure-based molecular design approach starting from the x-rays crystal structure of BRD4 in complex with BDIO9 inhibitor.^[28] We first elaborate a QSAR model correlating the relative Gibbs' free energy (GFE) of the BRD4-BDIO1-23 complexes formation with their respective experimental activities (IC_{50}^{exp}), and then generate a 3D QSAR pharmacophore (PH4) model of BRD4 inhibition based on the active conformation of bound BDIO1-23. The generated PH4 will serve to screen a virtual library of BDIO drug like analogs to identify the best mapping to the PH4. The predicted activity of these last best fit hits will be evaluated with the initial QSAR correlation equation and their ADME profile subsequently be computed.

MATERIAL AND METHODS

Training and Validation Sets

In this work, the chemical structures and biological activities (IC_{50}^{exp}) of the benzo[d]isoxazole family compounds were reported by Zhang *et al.*^[28,29] These 29 BDIOs' IC_{50}^{exp} span almost two orders of magnitude (101.92), a range of half-maximal inhibitory concentrations ($130 \leq IC_{50}^{exp} \leq 10860$) nM that allows the design of a QSAR model. This dataset was split into a training set (TS) containing

23 BDIO inhibitors and a validation set (VS) including 6 BDIOs inhibitors by a protocol called "Generate training and test data", within the Discovery Studio software.^[30]

Model Building

Three-dimensional (3D) molecular models of protein-inhibitor (P-I) complexes BRD4-BDIOx, free protein BRD4, and free inhibitors BDIO were prepared from the high-resolution (1.84 Å) crystal structure of a reference complex containing the compound BDIO9 inhibitor (PDB entry code: 5Y8Z^[28]) using the Insight-II molecular modeling program.^[31] The structures of BRD4 and the P-I complexes were considered to be at a pH of 7 with neutral N- and C-terminal residues and all protonizable and ionizable residues charged. No crystallographic water molecules are included in the model. The inhibitors were built into the reference structure of BRD4^[28] by *in-situ* replacing of derivatized groups in the molecular scaffold of the template inhibitor BDIO. An exhaustive conformational search over all rotatable bonds (dihedral angles) of the replacing function groups coupled with a careful gradual energy-minimization of the modified inhibitor and active site residues of the BRD4 located in the vicinity of the inhibitor (within 5 Å distance) was employed to identify low-energy bound conformations of the modified inhibitor. The resulting low energy structures of the P-I complexes were carefully refined by minimization of the whole complex. This procedure has been successfully used for model building of viral, bacterial, and protozoal enzyme-inhibitor complexes and design of peptide-mimetic, hydroxynaphthoic, thymidine, triclosan, pyrrolidine carboxamide, nitriles, chalcone and aryl-based inhibitors.^[32-43]

Molecular Mechanics and Conformational Search

The modeling of inhibitors, BRD4 and P-I complexes was carried out by molecular mechanics as described earlier.^[43] As mentioned above, free inhibitor conformations were derived from their bound conformations in the P-I complexes through gradual relaxation to the nearest local energy minimum.^[43]

Solvation Gibbs Free Energies

The electrostatic component of solvation Gibbs free energy (GFE) that includes also the effects of ionic strength via solving nonlinear Poisson-Boltzmann equation^[42,44,45] was computed by the Delphi module in Discovery Studio^[30] as described earlier.^[43]

Calculation of Binding Affinity and QSAR Model

The calculation of binding affinity expressed as complexation GFE has been fully described earlier.^[43]

Interaction Energy

The computation of MM interaction energy (ΔE_{int}) between enzyme residues and the inhibitor CFF force field was performed as mentioned earlier.^[43]



Pharmacophore Generation

Bound conformations of inhibitors from P-I complexes models were used to build a 3D-QSAR pharmacophore (PH4) model via the Discovery Studio Catalyst HypoGen algorithm^[43,46] as given in details earlier.^[43]

ADME Properties

The pharmacokinetic profile of BDIOs were computed by the QikProp program^[47] as described formerly.^[43]

Virtual Library Generation

The virtual library generation was performed as mentioned in our recent work.^[43]

ADME-based Library Searching

The drug-likeness selection criterion served to focus the initial virtual library as fully presented earlier.^[43]

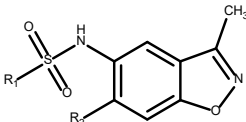
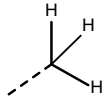

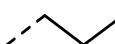
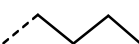
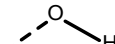
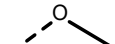
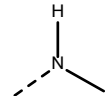
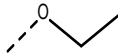
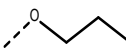
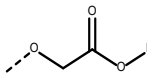
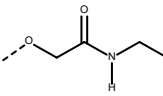
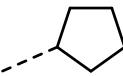
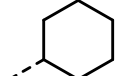
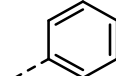
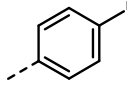
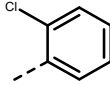
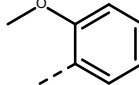
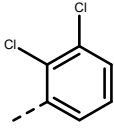
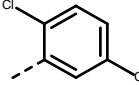
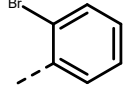
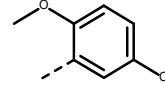
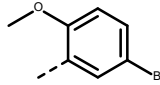
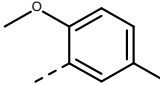
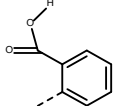
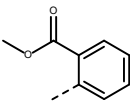
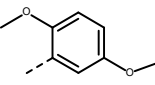
Pharmacophore-based Library Searching

The pharmacophore model (PH4) described in Section 2.8 and derived from the bound conformations of BDIOs at the active site of BRD4 served as a library searching tool as described in a recent study.^[43]

Inhibitory Potency Prediction

The conformer with the best mapping on the PH4 pharmacophore in each cluster of the focused library

Table 1: Training Set (TS) and Validation Set (VS) of BDIO inhibitors^[28] of BRD4 used in the preparation of quantitative structure-activity relationships (QSAR) model of inhibitor binding.

								
#R	1	2	3	4	5	6	7	
R-group								
#R	8	9	10	11	12	13	14	
R-group								
#R	15	16	17	18	19	20	21	
R-group								
#R	22	23	24	25	26			
R-group								
Training set	BDIO1	BDIO2	BDIO3	BDIO4	BDIO5	BDIO6	BDIO7	BDIO8
#R ₁ -#R ₂	22-6	22-7	22-8	21-6	22-5	22-9	19-6	23-6
(nM)	130	150	150	160	230	420	440	470
Training set	BDIO9	BDIO10	BDIO11	BDIO12	BDIO13	BDIO14	BDIO15	BDIO16
#R ₁ -#R ₂	22-1	17-1	21-1	16-1	20-1	13-1	19-1	3-1
(nM)	620	830	1050	1590	1840	2240	2350	2470
Training set	BDIO17	BDIO18	BDIO19	BDIO20	BDIO21	BDIO22	BDIO23	
#R ₁ -#R ₂	4-1	12-1	18-1	14-1	2-1	15-1	26-1	
(nM)	2550	3050	3550	4960	5300	9820	10680	
Validation set	BDIO24	BDIO25	BDIO26	BDIO27	BDIO28	BDIO29		
#R ₁ -#R ₂	20-6	22-10	22-11	23-1	24-1	25-1		
(nM)	220	580	1600	2130	9820	10680		

subset was used for $\Delta\Delta G_{\text{com}}$ calculation and $\text{IC}_{50}^{\text{exp}}$ estimation (virtual screening) by the complexation QSAR model as given in details earlier.^[43]

RESULTS AND DISCUSSION

Training and Validation Sets

The training set of 23 BDIOs and the validation set of 6 BDIOs (Table 1) were selected from the homogeneous series of BRD4 inhibitors for which experimentally determined inhibitory activities were available from the same laboratory.^[28] Substitutions made at two positions R1 and R2 of the benzo[d]isoxazole scaffold and R - group, as shown in (Table 1), made up the entire series. The experimental half-maximal inhibitory concentrations ($130 \leq \text{IC}_{50}^{\text{exp}} \leq 10860$ nM)^[28] cover a sufficiently wide concentration range to build a reliable QSAR model.

QSAR Model

Single-descriptor QSAR models

Each of the 23 training sets (TS) and 6 validation sets (VS) of BRD4-BDIOs complexes (Table 1) was prepared by in situ modification of the refined template crystal structure (pdb entry code 5Y8Z^[28]) of the complex BRD4-BRD9 as described in the Material and Methods section. Further, the relative Gibbs free energy of the BRD4-BDIOx complex formation ($\Delta\Delta G_{\text{com}}$) was computed for each of the 29 optimized protein-inhibitor complexes. (Table 2) lists computed values of $\Delta\Delta G_{\text{com}}$ and its components for the TS and VS of benzo[d]isoxazole derivatives.^[29] The QSAR model explained variation in the BDIOs experimental inhibitory potencies ($\text{pIC}_{50}^{\text{exp}} = -\log_{10} (\text{IC}_{50}^{\text{exp}})$ ^[29]) by correlating it with computed GFE $\Delta\Delta G_{\text{com}}$ through a linear regression. In addition, a significant correlation obtained in this QSAR relationship permitted the identification of the active bound conformation of the BDIOs at the BRD4 binding site and enabled the definition of the PH4. In search for better insight into the binding affinity of BDIOs towards BRD4, we have analyzed the enthalpy of complexation in gas phase $\Delta\Delta H_{\text{MM}}$. The validity of this linear correlation allowed assessment of the significance of inhibitor-protein interactions ($\Delta\Delta H_{\text{MM}}$) when solvent effect and loss of entropy of the inhibitor upon binding to the protein were neglected. For statistical data of the regression, see Table 3, Equation A. This correlation explained about 80% of the variation in $\text{pIC}_{50}^{\text{exp}}$ data and underlined the role of the enthalpic contribution to the binding affinity of the ligand. Similarly, the more advanced descriptor, namely the GFE of the BRD4-BDIOx complex formation including all components: $\Delta\Delta H_{\text{MM}}$, $\Delta\Delta TS_{\text{vib}}$ and $\Delta\Delta G_{\text{sol}}$, have been assessed (for statistical data see Table 3, Equation (B)). Relatively high values of the regression coefficient R^2 , leave-one-out cross-validated regression coefficient R^2_{xv} and Fischer F-test of the correlation suggest a strong

relationship between the 3D model of inhibitor binding and the observed inhibitory potencies of the BDIOs.^[32] Therefore, structural information derived from the 3D models of BRD4-BDIOx complexes can be expected to lead to the reliable prediction of BRD4 inhibitory potencies for new BDIOs analogs based on the QSAR model B, (see Table 3).

The statistical data confirmed the validity of the correlation Equations (A) and (B) plotted on Fig. 1. The ratio $\text{pIC}_{50}^{\text{pre}} / \text{pIC}_{50}^{\text{exp}} \cong 1$ (the values were estimated using correlation Equation (B), Table 3) calculated for the validation set BDIO24-29 documents the substantial predictive power of the complexation QSAR model from Table 2. Thus, the regression Equation (B) (Table 3) and computed $\Delta\Delta G_{\text{com}}$ GFE can be used for the prediction of inhibitory potencies $\text{IC}_{50}^{\text{exp}}$ against BRD4 for novel BDIO analogs, provided that they share the same binding mode as the training set BDIO1-23.

Binding Mode of BDIOs

The structural information from the BRD4-BDIOx complexes help identify key interactions which explain the affinity of benzo[d]isoxazole derivatives with BDIO. In fact, the key interactions which are involved in the shaping of BRD4-BDIOx complexes and which justify the affinity of benzofuran derivatives with BRD4, are hydrogen bonds, van der Waals and hydrophobic contacts, etc. Fig. 2 shows the binding mode of the most active ligand of the test set (BDIO1): hydrogen bonds with residues Tyr97, Gln85, Pro86, Asp88 and Pi-Pi stacking with Tyr97 and Phe83 and Pi-alkyl interactions with Leu92 and Ile146 on the phenyl substituent at the R1 position. Most training set molecules share these interactions enumerated for BDIO1. On the other hand, some of these interactions are lost for less active training set compounds (BDIO4, BDIO5, BDIO7, BDIO9, BDIO10) or the hydrogen interactions with residue Asn140 at the R1 position and as Asp88 at the R2 position. For other less active molecules (BDIO14, BDIO16, BDIO17, BDIO18, BDIO21), the lack of phenyl fragments and the presence of the methyl groups in R1 and R2 position respectively cause to absence of hydrogen and Pi-alkyl interactions with some key residues. In this work, interactions with these residues have been successfully preserved. However, it should be noted that several other interactions with residues around the active site participated in the formation of complex.

Interaction energy

Other key structural information was provided by the Interaction Energy (IE, E_{int}) diagram obtained for each training set inhibitor. IE break-down to contributions from BRD4 active site residues is helpful for the choice of relevant R1-groups and R2-groups, which could improve the binding affinity of BDIO analogs to the BRD4 and subsequently enhance the inhibitory potency. A comparative analysis of computed IE for the training



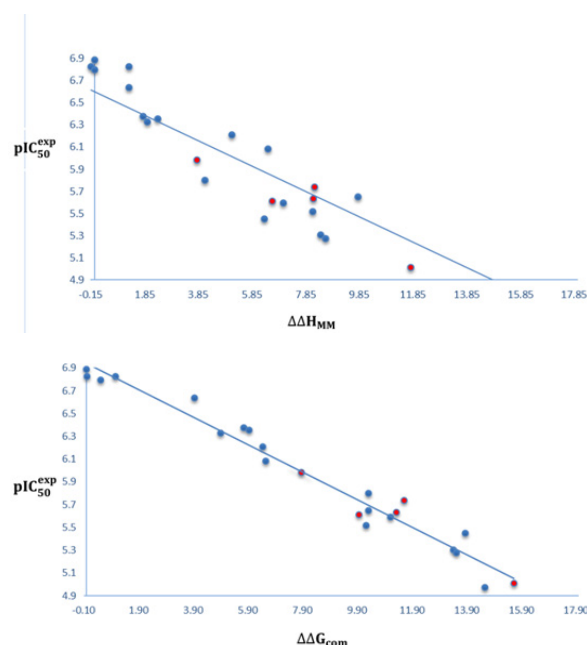
Table 2: Gibbs free energy (binding affinity) and its components for the training set of BRD4 inhibitors BDI01-23 and validation set inhibitors BDI024-29.^[32]

Training set ^a	M_W^b	$\Delta\Delta H_{MM}^c$	$\Delta\Delta G_{sol}^d$	$\Delta\Delta TS_{vib}^e$	$\Delta\Delta G_{com}^f$	$IC_{50}^{exp} g$ [nM]
BDI01	427	0.0	0.0	0.0	0.0	130
BDI02	426	1.27	-0.05	1.20	0.02	150
BDI03	441	-0.13	-0.32	-1.50	1.06	150
BDI04	382	0.01	0.72	0.22	0.50	160
BDI05	413	1.28	0.39	-2.28	3.95	230
BDI06	455	1.80	-0.04	-3.98	5.74	420
BDI07	387	2.34	-0.31	-3.91	5.93	440
BDI08	362	1.97	2.13	-0.81	4.91	470
BDI09	411	5.11	0.25	-1.08	6.44	620
BDI010	332	6.45	0.93	0.83	6.55	830
BDI011	366	3.83	2.93	-1.10	7.86	1050
BDI012	336	4.10	4.90	-1.29	10.30	1590
BDI013	381	8.20	-0.53	-3.95	11.62	1840
BDI014	308	9.59	6.38	5.89	10.30	2240
BDI015	371	8.16	-1.75	-4.92	11.34	2350
BDI016	268	6.62	3.62	0.26	9.97	2470
BDI017	282	7.03	6.31	2.21	11.12	2550
BDI018	394	8.13	4.29	2.20	10.22	3050
BDI019	371	6.32	2.72	-4.81	13.86	3550
BDI020	302	8.43	3.47	-1.52	13.42	4960
BDI021	254	8.60	6.39	1.47	13.52	5300
BDI022	320	11.79	2.18	-1.68	15.64	9820
BDI023	362	19.75	-3.90	1.28	14.58	10680
validation set ^a	M_W^b	$\Delta\Delta H_{MM}^c$	$\Delta\Delta G_{sol}^d$	$\Delta\Delta TS_{vib}^e$	$\Delta\Delta G_{com}^f$	$\frac{pIC_{50}^{prg}}{pIC_{50}^{exp} h}$
BDI024	353	3.09	-3.05	-3.63	3.67	0.98
BDI025	471	0.11	1.32	-3.54	4.96	1.02
BDI026	499	5.79	-2.19	-5.96	9.56	1.00
BDI027	411	10.18	1.62	2.17	9.68	1.02
BDI028	360	2.37	5.90	-0.54	8.12	1.14
BDI029	346	10.78	0.82	-4.35	14.34	1.01

^afor the chemical structures of the training set of inhibitors see Table 1; ^b M_W is the molar mass of inhibitors (g/mol); ^c $\Delta\Delta H_{MM}$ (kcal/mol) is the relative enthalpic contribution to the GFE change related to P-I complex formation derived by MM; $\Delta\Delta H_{MM} \cong [E_{MM}\{P; I_x\} - E_{MM}\{I_x\}] - [E_{MM}\{P; I_{ref}\} - E_{MM}\{I_{ref}\}]$; I_{ref} is the reference inhibitor BDI01; ^d $\Delta\Delta G_{sol}$ (kcal/mol) is the relative solvent effect contribution to the GFE change of P-I complex formation: $\Delta\Delta G_{sol} = [G_{sol}\{P; I_x\} - G_{sol}\{I_x\}] - [G_{sol}\{P; I_{ref}\} - G_{sol}\{I_{ref}\}]$; ^e $\Delta\Delta TS_{vib}$ (kcal/mol) is the relative entropic contribution of inhibitor I_x to the GFE upon P-I complex formation $\Delta\Delta TS_{vib} = [\Delta TS_{vib}\{P; I_x\} - \Delta TS_{vib}\{I_x\}] - [\Delta TS_{vib}\{P; I_{ref}\} - \Delta TS_{vib}\{I_{ref}\}]$; ^f $\Delta\Delta G_{com}$ (kcal/mol) is the overall relative GFE change of P-I complex formation: $P:I: \Delta\Delta G_{com} \cong \Delta\Delta H_{MM} + \Delta\Delta G_{sol} - T\Delta\Delta TS_{vib}$; ^g IC_{50}^{exp} is the experimental half-maximal inhibition concentration of BRD4 obtained from ref. ^[31]; ^h ratio of predicted and experimental half-maximal inhibition concentrations $pIC_{50}^{prg}/pIC_{50}^{exp}$ ($pIC_{50}^{prg} = -\log_{10}(IC_{50}^{prg})$) was predicted from computed $\Delta\Delta G_{com}$ using the regression equation for BRD4 shown in Table 3, (B).

Table 3: Analysis of computed binding affinities, its enthalpic component, and experimental half-maximal inhibitory concentrations $pIC_{50}^{exp} = -\log_{10}(IC_{50}^{exp})$ of BDIOs towards BRD4.^[32]

Statistical Data of Linear Regression	(A)	(B)
$pIC_{50}^{exp} = -0.1147 \times \Delta\Delta H_{MM} + 6.6022$ (A)		
$pIC_{50}^{exp} = -0.1205 \times \Delta\Delta G_{com} + 6.9374$ (B)		
Number of compound n	23	23
Squared correlation coefficient of regression R^2	0.80	0.96
LOO cross-validated Squared Correlation coef. R^2_{XV}	0.79	0.96
Standard error of regression σ	0.27	0.12
Statistical significance of regression. Fisher F-test	83.1	552.8
Level of statistical significance α	> 95%	> 95%
Range of activities IC_{50}^{exp} [nM]	130–10680	

**Fig. 1:** (a) Plot of correlation equation between pIC_{50}^{exp} and relative enthalpic contribution to the GFE $\Delta\Delta H_{MM}$; (b) Plot for relative complexation GFE $\Delta\Delta G_{com}$ of the training set of BDIOs, all in $\text{kcal}\cdot\text{mol}^{-1}$. Validation set data is shown in red color.

set BDIOs (Fig. 3) divided into three classes (highest, moderate, and lowest active) was conducted to identify the residues, contribution of which to binding affinity could be increased. However, the comparative analysis of these contributions with respect to their interaction energies facilitates the identification of residues with a more significant contribution to binding affinity. Further comparative analysis of BRD4 active site residues contribution to the interaction energies revealed no difference for the three classes, and consequently, no specific suggestion for relevant substitutions able to

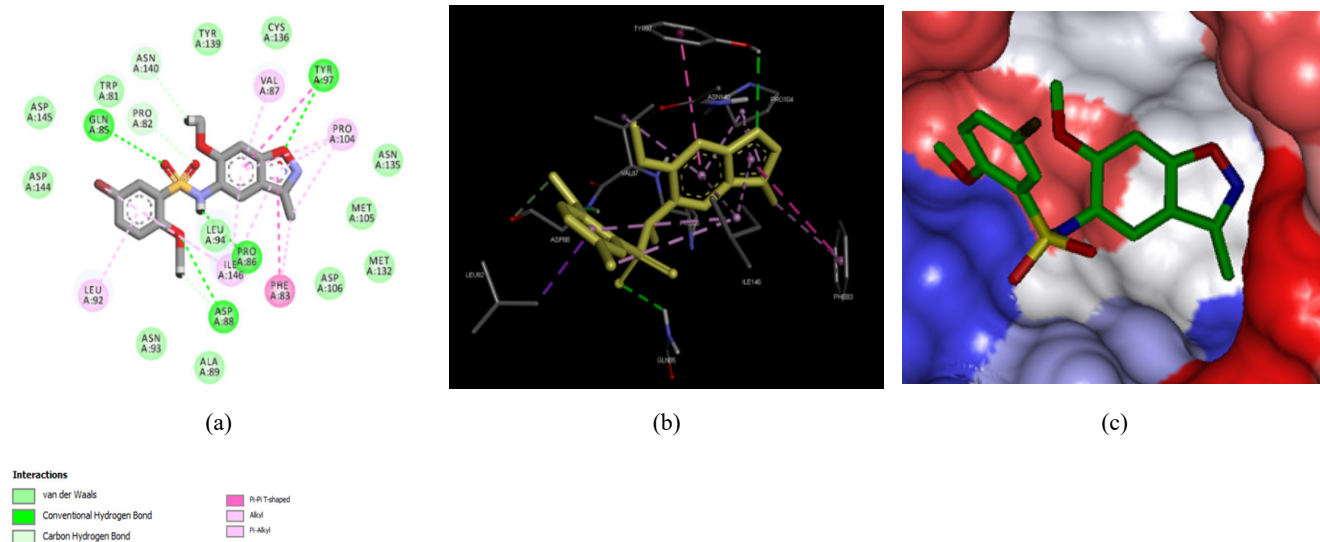


Fig. 2: (a) 2D schematic interaction diagram of the most potent inhibitor BDIO1^{[29][29]} at the active site; (b) 3D structure of the BDR4 active site with bond inhibitor BDIO1; (c) Connolly surface of the BDR4 active site for BDIO1. Surface coloring legend: red = hydrophobic, blue = hydrophilic and white = intermediate

potentially improve binding affinity was available. Therefore, the design of new BDIOs analogs obeys to a (so called) combinatorial approach. Through this approach, we generated an *in-silico library* (VL, virtual library to be screened) of 99 225 benzo[d]isoxazole analogs.

3D-QSAR Pharmacophore Model

BRD4 active site pharmacophore

The Connolly surface generation protocol in Insight-II molecular modeling program^[28] allows for mapping of hydrophobic and hydrophilic character of the active site of a protein. The surface of the active site of BRD4 is both hydrophobic and hydrophilic (Fig. 2, c). Building BRD4 inhibition 3D-QSAR PH4 from BDIOs active conformation will complement the enzyme active site pharmacophore and enhance affinity for those new analogs from VL which map well the inhibition PH4.

Generation and validation of 3D-QSAR pharmacophore

BRD4 inhibition 3D-QSAR PH4 was generated from the active conformation of 23 TS BDIO1-23 and evaluated by 6 VS BDIO24-29 covering a large range of experimental activity (130–10680 nM). The generation process is divided into three main steps: (i) the constructive, (ii) the subtractive, and (iii) the optimization step as described earlier.^[48,43,30] During the constructive phase BDIO1, BDIO2, BDIO3 and BDIO4 were retained as the lead (since their activities fulfilled the threshold criterion, $IC_{50}^{exp} \leq 1.25 \times 130$ nM) and used to generate the starting PH4 features. In the subtractive phase, compounds for which $IC_{50}^{exp} > 130 \times 10^{3.5}$ nM = 411096 nM were considered inactive. Accordingly, none of the training set BDIOs was inactive and no starting PH4 features were removed.

Finally, during the optimization phase, the score of the pharmacophoric hypotheses was improved. Hypotheses were scored according to errors in the estimated activity from regression and complexity via a simulated annealing approach. At the end of the optimization, the top scoring 10 unique PH4 hypotheses were kept, all displaying three features. The cost values, correlation coefficients, root-mean-square deviation (RMSD) values, the PH4 features, and the max-fit value of the top 10 ranked hypotheses (Hypo1–Hypo10) are listed in Table 4. They are selected based on significant statistical parameters, such as high correlation coefficient, low total cost, and low RMSD.

The generated pharmacophore models were assessed for their reliability based on the calculated cost parameters ranging from 101.4 (hypo1) to 122.4 (hypo10). The relatively small gap between the highest and lowest cost parameter corresponds well with the homogeneity of the generated hypotheses and the consistency of the TS of BDIOx. For this PH4 model, the fixed cost (53) is lower than the null cost (457.54) by a difference $\Delta = 404.54$. To be statistically significant, a hypothesis has to be as close as possible to the fixed cost and as far as possible from the null cost. For the set of 10 hypotheses, the difference $\Delta \geq 182.1$ attests to the pharmacophore model's high quality. The standard indicators such as the RMSD between the hypotheses ranged from 1.92 to 2.36, and the squared correlation coefficient (R^2) falls to an interval from 0.95 to 0.92. The first PH4 hypothesis with the closest cost (101.4) to the fixed one (53) and best RMSD and R^2 was retained for further analysis. The statistical data for the set of hypotheses (costs, RMSD, R^2) are listed in Table 4. The link between the 98% significance and the number 49 scrambled runs of each hypothesis is based on the formula: $S = [1 - (1 + X)/Y] \times 100$, with X the total number



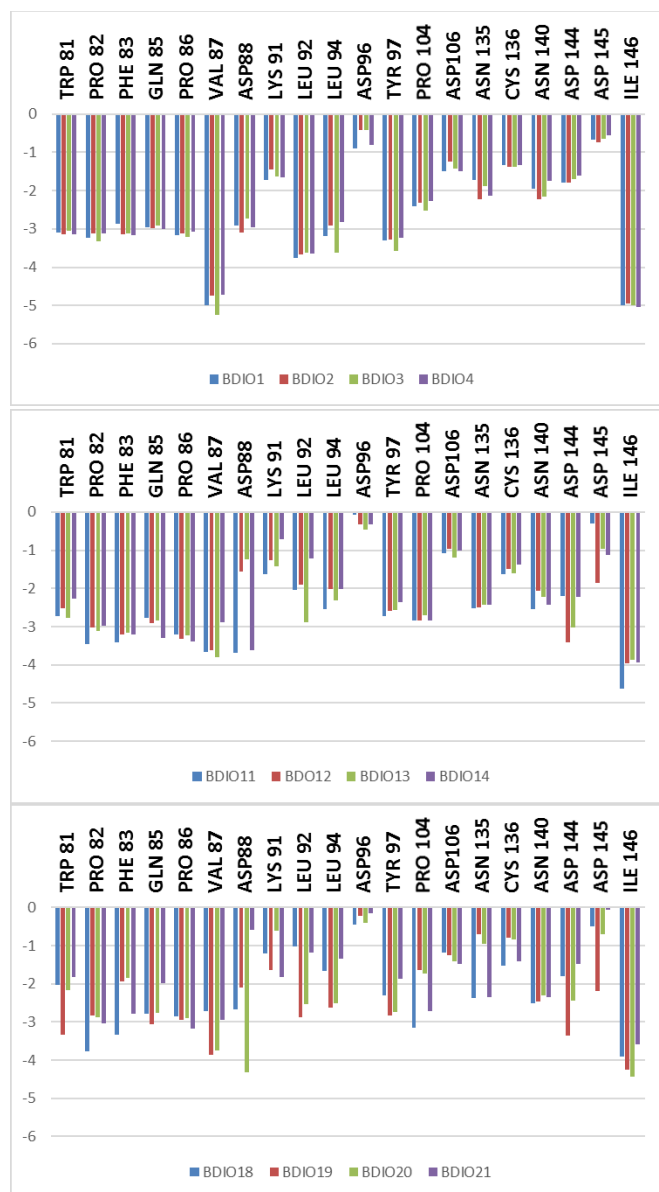


Fig. 3: Molecular mechanics intermolecular interaction energy breakdown to residue contributions in [kcal.mol⁻¹]: (Top) the most active inhibitors; (middle) moderately active inhibitors; (bottom) less active inhibitors BDIO, Table 2.^[28]

of hypotheses having a total cost lower than the original hypothesis (Hypo 1) and Y the total number of HypoGen runs (initial + random runs): $X=0$ and $Y = (1+49)$, hence $98\% = \{1 - [(1+0)/(49+1)]\} \times 100$.

The evaluation of Hypo 1 was performed first through Fischer's randomization cross-validation test. The Cat-Scramble program was used to randomize the experimental activities of the training set. At 98% confidence level, each of the 49 scramble runs created ten valid hypotheses, using the same features and parameters as in the generation of the original 10 PH4 hypotheses. Among them, the cost value of Hypo1 is the lowest compared with those of the 49 randomly generated hypotheses, as we can see in Table 4 where the lowest cost of the 49 random runs

Table 4: Parameters of 10 generated PH4 pharmacophoric hypotheses BRD4 inhibitors^[28] after the Cat Scramble validation procedure (49 scrambled runs for each hypothesis at the selected level of confidence of 98%)

Hypothesis	RMSD ^a	R ² ^b	Total costs ^c	Costs difference ^d	Closest Random ^e
Hypo1	1.915	0.95	101.4	306.4	151.1
Hypo2	1.979	0.94	105.1	297.3	160.2
Hypo3	2.027	0.94	106.3	297.1	160.4
Hypo4	2.172	0.93	111.6	284.2	173.4
Hypo5	2.183	0.93	114.7	218.0	239.5
Hypo6	2.267	0.93	118.7	205.6	252.0
Hypo7	2.265	0.93	120.4	201.7	255.8
Hypo8	2.303	0.92	121.9	192.4	265.2
Hypo9	2.358	0.92	122.2	182.2	275.3
Hypo10	2.358	0.92	122.4	182.1	275.5

^a Root Mean Square Deviation; ^b Squared correlation coefficient; ^c Overall cost parameter of the PH4; ^d Cost difference between Null cost and hypothesis total cost; ^e Lowest cost from 49 scrambled runs at a selected level of confidence of 98%. The Fixed Cost = 53 with RMSD = 0, Null Cost = 457.54 with RMSD = 6.02 and the Configuration cost = 10.38

is listed for each original hypothesis, and none of them was as predictive as the original hypotheses generated shown in Table 4. Thus, there is a 98% probability that the best selected hypothesis Hypo1 represents a PH4 model for inhibitory activity of BDIOs with a similar level of predictive power as the complexation QSAR model, which relies on the benzoisoxazol active conformation from 3D structures of the BRD4-BDIOx complexes and their computed GFE $\Delta\Delta G_{\text{com}}$. Another evaluation of Hypo 1 is the mapping of the best active training set BDIO1 (Fig. 4) displaying the geometry of the Hypo1 PH4 of BRD4 inhibition. The Hypo1-based regression equation for $\text{pIC}_{50}^{\text{exp}}$ vs $\text{pIC}_{50}^{\text{pre}}$: $\text{pIC}_{50}^{\text{exp}} = 0.9909 \times \text{pIC}_{50}^{\text{pre}} + 0.0554$ ($n = 23$, $R^2 = 0.95$, $R_{\text{XV}}^2 = 0.89$, $\sigma = 0.194$, $\alpha > 98\%$, $F\text{-test} = 186.28$) Fig. 4.

Virtual Screening

In-silico screening of a virtual library of ligands can lead to hits identification as it was shown in our previous works on inhibitor design.^[32,43,48]

Virtual library

An initial virtual combinatorial library (VCL) was generated by substitutions at positions R1 and R2 (see Table 5) on the benzo[d]isoxazole derivatives scaffold. During the VCL enumeration, the R-groups listed in Table 6 were attached to positions R1 and R2 of the BDIO scaffold to form a virtual combinatorial library of the size: $R1 \times R2 = 315 \times 315 = 99225$ BDIO analogs.

All analogs are matching the substitution pattern of the best inhibitor BDIO1. This BDIOs analogs library was

generated from fragments (chemicals) listed in databases of available chemicals.^[49] Nowadays, one of the criteria for the design of new anticancer drugs, for the target population, is their oral bioavailability.

To design a more targeted library of reduced size and increased rate of drug-like and orally bioavailable BDIOs analogs, a set of filters and penalties were introduced, such as the Lipinski rule-of-five^[37] facilitating the selection of a smaller number of suitable BDIOs that can be submitted to *in silico* screening.

In-silico Screening of Library of BDIOs

The focused library of 89 779 analogs was further screened for molecular structures matching the 3D-QSAR PH4 pharmacophore model Hypo1 of BRD4 inhibition. 209 BDIOs mapped at least 5 pharmacophoric features and these 106 best fitting analogs (PH4 hits) were selected and subjected to complexation QSAR model screening. The computed relative GFE of BRD4-BDIOx complex formation, their components, and predicted half-maximal inhibitory concentrations calculated from the correlation equation (B) (Table 3) are listed in Table 5.

Analysis of novel BDIO analogs substituents

To identify which substituents on R-positions of BDIO scaffold (Table 5) lead to new inhibitor candidates with the highest predicted potencies towards the BDIO, histograms of the absolute frequency of occurrence of R1- and R2- groups among the 106 best fit PH4 hits were prepared (Fig. 5). From these histograms, it comes out that R1-groups numbered 1(20), 2(4), 4(3), 7(14), 44(5), 50(6) and 62(4) are almost equally represented with the highest occurrence in BDIO subset. The R2-groups contain preferentially 6(4), 11(4), 14(8), 100(6) and 113(8).

ADME Profile of Novel BDIO Analogs

The properties related to ADME such as octanol-water partitioning coefficient, aqueous solubility, blood-brain partition coefficient, Caco-2 cell permeability, serum protein binding, number of likely metabolic reactions, and another eighteen descriptors related to absorption, distribution, metabolism, and excretion (ADME) were calculated by the QikProp program^[50] for the new best BDIO analogs (Table 7). This program is based on the method of Jorgensen.^[51] Experimental data from more than 710 compounds, including about 500 drugs and related heterocycles, were used to produce regression equations correlating experimental and computed descriptors resulting in an accurate prediction of pharmacokinetic properties of molecules. The drug-likeness (number of stars) i.e., the number of property descriptors that fall outside the range of optimal values determined for 95% of known drugs out of 24 selected descriptors computed by the QikProp,^[47] was used as an additional ADME-related compound selection criterion. The values for the best active designed BDIOs are compared with those computed

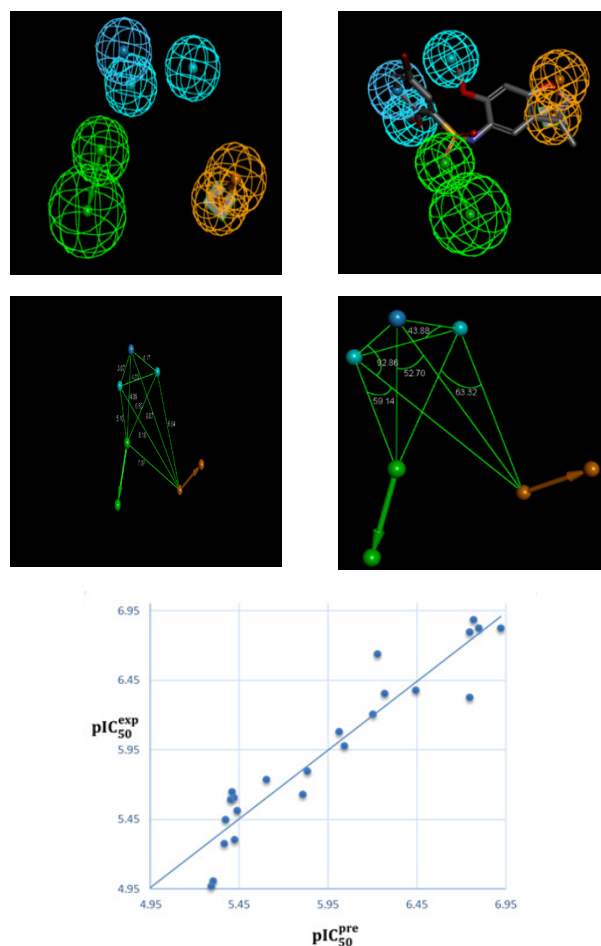


Fig. 4: Top left features of the pharmacophore model of BRD4 inhibition; (Top right) PH4 mapping with the most potent BDIO1 ($IC_{50}^{exp} = 130$ nM); (Middle left) distances between centers; (Middle right) angle between centers of pharmacophoric features. The features are colored green for Hydrogen Bond Acceptor (HBA), orange for aromatic ring (Ar) and blue for hydrophobic; (Bottom) correlation plot of experimental vs. predicted inhibitory activity (open circles correspond to TS).

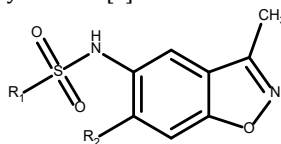
for drugs used for the treatment of CRPC or currently undergoing clinical trials, Table 7.

Predicted BRD4 Inhibition Potency for Current Drugs benzoisoxazole Scaffold

Since the benzoisoxazole scaffold has been analyzed in this study, drugs currently used in the clinical practice sharing this scaffold are worth evaluating with the help of our 3D-QSAR generated PH4 pharmacophore. The list of 10 compounds given in Table 8 is mostly indicated for treatment of neural disorders, dementia, schizophrenia in adults, psychosomatic disorders and cancers.^[52]

As we can see on Fig. 6, the mapping of the three most potent predicted BRD4 inhibitors to PH4 pharmacophore sheds light on their affinity towards the protein. However, the low affinity of these inhibitors towards the protein cannot allow to suggest their experimental evaluation as



Table 5: R1- and R2-groups (fragments, building blocks, substituents) were used in the design of the initial diversity virtual combinatorial library of benzo[d]isoxazole derivatives.

R-groups ^a					
1	5-ethyl-2-methoxybenzenyl	2	2-hydroxybenzenyl	3	1-(2-methoxyphenoxy) methyl
4	5-amino-2-methylbenzenyl	5	3-ethoxybenzenyl	6	2-hydroxyl-3-methylbenzenyl
7	2-methoxy-6-methylbenzenyl	8	4-amino-3-ethylbenzenyl	9	4-amino-3-propylbenzenyl
10	4-(ethylamino)-3-methylbenze	11	4-(ethylamino)-3-methylbenzenyl	12	4-amino-3,5-dimethylbenzenyl
13	3-methyl-4-(propylamino-benzenyl	14	3,5-diMe-4-(propylamino) benzenyl	15	Methoxyl
16	Ethoxyl	17	Propoxyl	18	Hydroxyl
19	Ammoniac	20	Methylamin	21	Ethylamin
22	Propylamin	23	Methyl	24	4-methyl-1H-2-pyrrolyl
25	Chlorure	26	Bromure	27	Hydrogen sulfure
28	Ethyl	29	Propyl	30	Prop-1-yl
31	Ethynyl	32	1-oxoethanyl	33	1-oxomethyl
34	Methylamide	35	1-N-methylmethanamide	36	Methanoic acid
37	1,1-dihydroxymethanyl	38	1-hydroxy-1-methoxymethanyl	39	3-methyl-4-(propylamino) benzenyl
40	2-(2-fluoroethyl)-6-methoxybenzenyl	41	3,4-difluoro-2-hydroxybenzenyl	42	4-(propylamino) benzenyl
43	4-(butylamino)-3,5-dimethylbenzenyl	44	2-methoxybenzenyl	45	2-methoxy-6-methylbenzenyl
46	2-ethoxy-6-methylbenzenyl	47	2-methyl-4-(propylamino) benzenyl	48	4-(ethylamino)-3,5-dimethylbenzenyl
49	4-ethoxy-3,5-dimethylbenzenyl	50	2-hydroxy-6-methoxybenzenyl	51	5-bromo-2-hydroxybenzenyl
52	3,5-dimethyl-4-propoxybenzenyl	53	3-methyl-4-propoxybenzenyl	54	2-methoxy-3-methylbenzenyl
55	3-methylbenzenyl	56	4-((OHMe)amino)-3,5-diMebenzenyl	57	2-(hydroxymethyl)-6-methylbenzenyl
58	4-hydroxybenzenyl	59	4-(ethylamino)-2,3,5-trimethylbenzenyl	60	4-OH-2-(propa-1,2-dien-1-yl) benzenyl
61	2-aminobenzenyl	62	5-bromo-2-methoxybenzenyl	63	4-oxo-4H-2-pyranyl
64	2-hydroxy-6-methylbenzenyl	65	2-(methylamino) benzenesulfonyl	66	2-acetylbenzenyl
67	4-(ethylamino) benzenyl	68	Acrylamide	69	2-methoxy-5-methylbenzenyl
70	3-acetylbenzoic acid	71	3-benzoic acid	72	4-fluorobenzenyl
73	2-(trifluoromethyl) benzenyl	74	5-bromo-2-methoxybenzenyl	75	2-methoxybenzenyl
76	5-bromo-2-methoxybenzenyl	77	2,6-dichloro-benzenyl	78	2-methoxy-6-propylbenzenyl
79	2-hydroxy-4-methylbenzenyl	80	2-phenyl methylcarbamate	81	2-amino-3-methylbenzenyl
82	2-OH-6-Me-5-methylenecyclohexa-1,3-diene	83	3-amino-2-hydroxybenzenyl	84	5-amino-4-hydroxypyridine
85	Fluoride	86	1,1,1-trihydroxymethanyl	87	2-methoxy-5-methylbenzenyl
88	4-methoxy-2-methylpyrimidine	89	Hypofluorous anhydride	90	Hypochlorous anhydride
91	Hypobromous anhydride	92	3-amino-benzenyl	93	5-fluoro-2-hydroxybenzenyl
94	4-amino-3-methylbenzenyl	95	3-amino-2-fluoro-6-methylbenzenyl	96	3-amino-4-methylbenzenyl
97	2-methylcyclohexa-1,3,5-trien-1-amine oxo-4H-pyran	98	4-amino-3-methylbenzenyl	99	4-amino-2-methylbenzenyl
100	2-amino-3-methylbenzenyl	101	2-hydroxybenzenyl	102	4-amino-2-fluoro-6-hydroxybenzenyl
103	3-methyl-4-(methylamino) benzenyl	104	2-fluoro-6-methoxybenzenyl	105	5-fluoro-2-methoxybenzenyl
106	4-fluoro-2-methoxybenzenyl	107	3-fluoro-2-methoxybenzenyl	108	4-amino-3-hydroxybenzenyl
109	2,3-dichlorobenzenyl	110	1-aminomethane	111	4H-1,4-oxazine

Contin..

<i>R-groups^a</i>			
112	2-bromo-6-methoxybenzenyl	113	4-chloro-2-methoxyphenyl
115	3,4-diaminobenzenyl	116	2,4-diaminobenzenyl
118	2,6-diamino-3-methylbenzenyl	119	2,6-diaminobenzenyl
121	5-methoxy-2-methylbenzenyl	122	2-ethoxybenzenyl
124	2-ethoxy-5-(methylamino) benzenyl	125	6-Et-2-MeO-4-methylenecyclohexa-1,5-diene
127	4-amino-2-MeO-5-methylbenzenyl	128	2-chloro-4-methylbenzenyl
130	3-hydroxy-2-methylpyridine	131	3-bromo-2-formylbenzenyl
133	3-carbonylbenzoic acid	134	3-bromo-2-methoxybenzenyl
136	5-(aminomethyl)-2-hydroxybenzenyl	137	5-(aminomethyl)-2-methoxybenzenyl
139	4-amino-2-(methylamino) benzenyl	140	2-amino-5-hydroxy-3-methylbenzenyl
142	4-(ethylamino)-2-methylbenzenyl	143	3-formylbenzenyl
145	5-acetyl-2-aminobenzenyl	146	3-acetyl-5-aminobenzenyl
148	3-acetyl-2-aminobenzenyl	149	2-amino-3-benzenyl
151	3-amino-5-formylbenzenyl	152	2-amino-5-formylbenzenyl
154	2-acetyl-6-aminobenzenyl	155	2-amino-6-formylbenzenyl
157	2-acetyl-4-aminobenzenyl	158	4-amino-2-formylbenzenyl
160	2-acetyl-3-aminobenzenyl	161	2-acetyl-5-bromobenzenyl
163	2-acetyl-4-bromobenzenyl	164	3-bromo-2-formylbenzenyl
166	5-bromo-2-formylbenzenyl	167	2-chloro-6-formylbenzenyl
169	2-acetyl-6-chlorobenzenyl	170	4-fluorocyclohexanyl
172	Methyl 4-cyclohexanecarboxylate	173	4-formylcyclohexyl
175	4-methylcyclohexyl	176	2-cyclohexanecarboxylic acid
178	2-formylcyclohexyl	179	2-acetylcyclohexyl
181	2-(1-methoxyethyl) cyclohexyl	182	2-(methoxymethyl) cyclohexyl
184	2-methylcyclohexyl	185	3-cyclohexyl carboxylic acid
187	3-formylcyclohexyl	188	3-(hydroxymethyl) cyclohexyl
190	3-methylcyclohexyl	191	3-methylpiperidinyl
193	3-methylpiperidinyl	194	2-methylpiperidinyl
196	1,3-oxazinanyl	197	6-methyl-1,3-oxazinanyl
199	4-methyl-1,3-oxazinanyl	200	3-methylmorpholinyl
202	2,6-dimethylmorpholyl	203	2,3-dimethylmorpholinyl
205	4H-1,4-oxaziny	206	3-amino-4H-1,4-oxaziny
208	2-amino-5-fluoro-4H-1,4-oxaziny	209	3-(fluoroamino)-4H-1,4-oxaziny
211	2-amino-5-methyl-4H-1,4-oxaziny	212	2-amino-3-methyl-4H-1,4-oxaziny
214	2-amino-6-methyl-4H-1,4-oxaziny	215	3-chloro-4-fluorobenzenyl
217	3-chloro-4-fluoro-N-methylaniliny	218	2-chloro-4-fluoro-N-methylaniliny
220	5-F-N-methylbenzene-1,3-diaminy	221	N-methylpyridin-3-aminy
223	5-Cl-3-F-2-(methylamino) pyridiny	224	5-(aminomethyl)-2-pyridiny
226	5-amino-2-ethylpyridiny	227	5-amino-2-fluoropyridiny
229	5-amino-2-ethoxy-3-fluoropyridiny	230	3-amino-2-fluoropyridiny
232	3-amino-6-formyl-2-MeO-pyridiny	233	3-amino-4-methoxypyridiny
235	3-methylaminopyridiny	236	6-hydroxy-3-methylaminopyridiny
114	2-methyl-1,4-dihydropyridine	117	2,4-diaminobenzenyl
120	5-hydroxy-2-methylbenyl	123	2-ethoxy-3,6-dimethylbenzenyl
126	2-methyl-6-(methylamino) benzenyl	129	4,6-dichloro-5-methylpyridine
132	(E)-3-hydroxyacrylamide	135	2-amino-3-ethylbenzenyl
138	2,5-bis(methylamino)benzenyl	141	2-amino-5-methoxy-3-methylbenzenyl
144	3-acetylbenzenyl	147	3-acetyl-4-aminobenzenyl
150	4-amino-3-formylbenzenyl	153	2-amino-6-formylbenzenyl
156	2-acetyl-5-amino-benzenyl	159	3-amino-2-formylbenzenyl
162	2-acetyl-6-bromobenzenyl	165	4-bromo-2-formylbenzenyl
168	2-bromo-6-formylbenzenyl	171	4-cyclohexanecarboxylic acid
174	4-(hydroxymethyl) cyclohexyl	177	Methyl 2-cyclohexanecarboxylate
180	2-(1-hydroxyethyl) cyclohexyl	183	2-(hydroxymethyl) cyclohexyl
186	Methyl 3-cyclohexanecarboxylate	189	3-(methoxymethyl) cyclohexyl
192	4-methylcyclohexayl	195	2-methyl-1,3-oxazinanyl
198	5-methyl-1,3-oxazinanyl	201	2-methylmorpholinyl
204	1H-1,4-oxaziny	207	2-amino-4H-1,4-oxazine
210	3-amino-5-methyl-4H-1,4-oxazine	213	2-amino-5-methyl-4H-1,4-oxazine
216	3-amino-5-methylpyrazine	219	4-fluoro-N-methylbenzene-1,3-diamine
222	2-chloro-5-fluoropyridine	225	2-amino-5-chloropyridine
228	5-amino-2-methoxypyridine	231	3-amino-2-methoxypyridine
234	3-amino-4-hydroxypyridine	237	6-methoxy-3-methylaminopyridine

Contin..

