Scientific African 15 (2022) e01088

Contents lists available at ScienceDirect

Scientific African

journal homepage: www.elsevier.com/locate/sciaf

Structure-based simulated scanning of rheumatoid arthritis inhibitors: 2D-QSAR, 3D-QSAR, docking, molecular dynamics simulation, and lipophilicity indices calculation ☆

Emmanuel Israel Edache [a,b,*], Adamu Uzairu [b], Paul Andrew Mamza [b], Gideon Adamu Shallangwa [b]

[a] Department of Pure and Applied Chemistry, University of Maiduguri, Borno, Nigeria
[b] Department of Chemistry, Ahmadu Bello University, P.M.B. 1044, Zaria, Nigeria

ARTICLE INFO

Article history:
Received 16 November 2020
Revised 30 October 2021
Accepted 16 December 2021

Editor: DR B Gyampoh

Keywords:
Rheumatoid arthritis
QSAR
CoMFA
Docking
MD simulations
ADME
Lipophilicity indices
Golden triangle

ABSTRACT

Rheumatoid arthritis (RA) is an autoimmune condition in the world, affecting about 1% of the population. It is characterized by a cartilage attack unique to the tissue in the peripheral joints. Different types of drugs are used to treat the disease are aspirin, ibuprofen, naproxen and celecoxib, but their use is less effective due to increased drug resistance and the side effects of these drugs. The need for new anti-rheumatoid arthritis drugs with improved activities and a new mechanism of action along diverse mechanisms is therefore necessary. In this study, 30 compounds from PubChem database with accession number 435,024 were carefully chosen as innovative RA inhibitors to battle autoimmune deficiency and were laid open to quantitative structure-activity relationship (QSAR), comparative molecular field analysis (CoMFA), molecular docking, molecular dynamic simulations, ADMET, Frontier Molecular Orbitals, and golden triangle studies. In the QSAR study, multiple linear regression was used for descriptor selection. The results showed that QSAR ($N_{training}$ = 18; R^2 = 0.796; $R^2_{adjusted}$ = 0.7523; F = 18.209; Q^2 = 0.67999) take good stability and predictability. In the CoMFA fractional factorial design (FFD) and CoMFA smart region definition (SRD) study, based on the Open3DALIGN-based arrangement used, results have shown that the CoMFA (FFD) model with reliable predictive ability was chosen as the final model to perform modeling study. The CoMFA (FFD) model, which shows comparatively good execution, was applied to search the necessary structural areas where the adjustment was essential to design a novel compound with bettered bioactivity. The molecular docking study showed that compound 4 has the least binding energy of -8.0 kcal/mol using PyRx software and compound 17 has the least binding energy of -8.55 kcal/mol using AutoDock4.2 software, respectively. Based on docking results on compounds 4 and 17 have good predictive behavior and relatively good ADMET exposure. MD simulation analysis of 1000 ps shows that one hydrogen bond with the residues was formed by compound 17 (Lys11) and binds distantly to the protein. Compound 4 has a smaller energy gap (-

☆ Edited by: B. Gyampoh

* Corresponding author at: Department of Pure and Applied Chemistry, University of Maiduguri, Borno, Nigeria.
E-mail address: edacheson2004@gmail.com (E.I. Edache).

https://doi.org/10.1016/j.sciaf.2021.e01088
2468-2276/© 2021 The Author(s). Published by Elsevier B.V. on behalf of African Institute of Mathematical Sciences / Next Einstein Initiative. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

6.5 eV) have great penetrating power, optimized clearance and oral absorption compared to compound 17. These studies offer valuable insights into the discovery and design of RA inhibitors in a modern age.

© 2021 The Author(s). Published by Elsevier B.V. on behalf of African Institute of Mathematical Sciences / Next Einstein Initiative.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

Introduction

Rheumatoid arthritis (RA) is an autoimmune condition. This means that healthy tissues are mistakenly destroyed by the body's immune system [1]. RA is an infectious disorder and one of the inflammatory diseases contributing to inflammation of the joints and underlying tissues [2,3]. It's a long-term condition. Multiple organs can also be affected [4]. The trigger of RA is unclear. RA can happen at all ages, but in middle age, it is more likely [5]. Women get RA more often than men do. The condition may be linked to inflammation, DNA and hormone differences [6]. The RA can also be linked to smoking. It is less pervasive than osteoarthritis (OA). Osteoarthritis (OA) is a condition that develops in many individuals due to wear and tear in their joints as they age [7]. RA also affects joints on both parts of the body most of the time (https:/medlineplus.gov/ency/article/000,431). The most frequently affected are fingers, hands, ankles, back, elbows, calves, hips and shoulders [6,8]. Pleurisy, dry eyes and mouth (Sjögren Syndrome), burning eyes, itching and discharge, under-skin nodules (most likely a sign of more severe disease), numbness, tingling or swelling of the hands and feet, issues with sleeping [5,9]. Aspirin and non-steroidal anti-inflammatory drugs (NSAIDs) such as ibuprofen, naproxen and celecoxib are the medications used in the treatment of RA [10,11]. To reduce joint stiffness and inflammation, these drugs perform very well, although they can have long-term side effects (https//: www.versusarthritis.org/about-arthritis/conditions/rheumatoid-arthritis). Therefore, they can only be used for a short period and, if possible, at reduced doses. Since they do not avoid typical harm when used on their own, Disease-modifying antirheumatic drugs (DMARDS) should also be used [12]. A delicate combination of biological and physicochemical properties must be demonstrated by a promising drug that crosses the obstacles of clinical trials to achieve acceptance and a secure market place [13]. Theoretical experiments focused on either statistics or statistical models can also play a key role in explaining the disease features of the outbreak, in addition to medical and biological science [14]. The present research was dedicated to defining the quantitative structure-activity relationship (QSAR) of the compounds and to determining their binding modes and structural stabilities as RA inhibitors to further establish novel RA inhibitor compounds based on synthesized compounds. The quantitative structure-activity relationship (2D-QSAR and 3D-QSAR) research combined with molecular docking and molecular dynamics simulation is the recognized approach for predicting novel and potent lead compounds in computer-aided drug design methods [14]. Though QSAR establishes a model for new potent compound prediction [15], molecular docking predicts a ligand's binding mode in its protein target's active site [16]. On the other hand, the simulation of molecular dynamics coupled with the measurement of ADMET, electrostatic potential (EP) and boundary molecular orbital (FMO) helps determine the stability and bind free energy of a ligand to its protein target.

Materials and methods

We scanned the PubChem database (https:/pubchem.ncbi.nlm.gov/bioassay/435,024) of 140 compounds for rheumatoid arthritis and found 30 which had been checked for tyrosine phosphatase, non-receptor type 22 (lymphoid) isoform 1. The balanced geometry of all 30 compounds was calibrated using the semi-empirical PM3 using the Spantan'14 version 1.1.4 program. No molecular symmetry restriction was applied, more complete optimization of all bond lengths and angles was done. The collection of the training set and the test set of molecules are rendered in such a manner that the test set of molecules reflects a spectrum of biological behavior close to that of the training set; therefore, the test set is genuinely representative of the training set. Chemical Structure of compounds and their IC50 values are listed in Table SM1. The IC_{50} values are translated to the logarithmic scale (-log(IC50 $\times$ 10^{-6})) used as dependent variables to create the 2D and 3D QSAR models.

Model development

Separate physicochemical descriptors for each of the compounds in the dataset are calculated using PaDEL v2.20 [17] for the implementation of 2D-QSAR models. Different sets of physicochemical descriptors are estimated with the software. The 2D microstructure was used in the study. 2D-QSAR models were generated using DTC-Lab software (http://dtclab.webs.com/software-tools). The cross-validated determination coefficient $Q^2_{(LOO)}$ which is used as a criterion of both robustness and predictive ability of the model were calculated [18]. The model has also been validated through external validation, including the method of Golbraikh-Tropsha and the method of Roy [19–21]. Galbraith–Tropsha's method and Roy's statistical parameters are shown in Table 1.

Table 1
Internal and external validation parameters of the model.

Equation	Formular	Equation	Formular
1	$Q^2 = \frac{\sum^{(Y_{obs}-Y_{pred})^2}}{\sum^{(Y_{obs}-\bar{Y})^2}}$	5	$SDEP = \sqrt{\frac{\sum^{(Y_{pred}-Y_{obs})^2}}{n}}$
2	${R_{pred}}^2 = 1 - \frac{\sum^{(Y_{pred(test)}-Y_{test})^2}}{\sum^{(Y_{test}-\bar{Y}_{training})^2}}$	6	$r^2_{adjusted} = \frac{(n-1)\times r^2 - P}{n-1-p}$
3	$r^2_{m(test)} = r^2 \times (1 - \sqrt{r^2 - r_o^2})$	7	$Q^2_{(f2)} = 1 - \frac{\sum^{(Y_{obs(test)}-Y_{pred(test)})^2}}{\sum^{(Y_{obs(test)}-\bar{Y}_{(test)})^2}}$
4	$K = \frac{\sum Y_{obs} Y_{pred(test)}}{\sum Y^2_{pred(test)}}$	8	$K' = \frac{\sum Y_{obs} Y_{pred(test)}}{\sum Y^2_{obs}}$

Yobs, Y(Test), and Ypred(Test) indicate observed activity, observed activity for the test set, and predicted activity, respectively. Ytraining indicates the mean activity value of the training set. The value corresponding to 'n' and 'p' is the number of samples in the training set and the number of components in the MLR model.

According to Galbraith and Tropsha [19,20], the suggested QSAR model is predictive (Table 1) since it meets the following criteria: cross-validation correlation coefficient ($Q^2 \geq 0.6$), squared correlation coefficient ($R^2 \geq 0.6$), predictive correlation coefficient ($R^2_{pred} \geq 0.5$), a modified $r^2_{m(test)} \geq 0.5$, $\frac{r^2 - r_o^2}{r^2} \leq 1$, and Degree of Lack of Relationship ($0.85 \leq K \leq 1.15$, $0.85 \leq K' \leq 1.15$).

Model Y-scrambling

The model was also subjected to a Y-scrambling test. Y-scrambling is useful to avoid any chance-comer correlation between a dependent variable and independent variables. In the simulation, the operation data is arbitrarily permuted to keep the descriptor unchanged, followed by an MLR sprint. Every randomization and subsequent MLR analysis produce a new set of values for R^2 and Q^2 [22]. If the new QSAR models have lower R^2 and Q^2 values for several trials (10 times in this study), then the given QSAR model is thought to be robust.

Applicability domain (AD)

The applicability domain of the QSAR is the knowledge from which training on the model has been developed and for which predictions for new compounds are valid [23]. The applicability domain aims to decide if the assumptions of the model are fulfilled [24]. Besides, this is the case of interpolation.

Molecular alignment and 3D-QSAR analysis

Structural alignment to obtain an effective model of molecular field interaction is one of the most critical phases in 3D-QSAR research. The veracity of the CoMFA theoretical account calculation and the reliability of the contour models relied heavily on the structural orientation of the molecules [25]. In Open3DALIGN v2.3 software [26], Compound 17 obtained the highest O3A score and was used as a proposed superimposition template, assuming that its configuration at the active site stage of the receptor reflects another very bioactive structure. In Fig. SM1, the aligned compounds are shown. This work used Open3DQSAR v2.3 [27] for the review of the CoMFA model. The best-aligned compounds are put in a 3D cubic lattice with a grid size of 1.5 Å and a 5.0 Å out gap. A thorough overview of the input parameters is given by Ahamed and his coworker [28].

Protein-ligand interaction with AutoDock

Docking experiments were conducted using PyRx (AutoDock Vina) [29] to obtain the energy affinity and mode of binding of the protein to the compounds. The height of the grid box has been optimized to 37.7775 × 110.1943 × −19.4534 Å with the base at 59.4273895049, 48.6285671234, 60.8047401428 for x, y, and z, respectively. A thorough overview of the method is given by Shafiu and his coworker [16]. AutoDock software [30] (ADT) version v1.5.4 and AutoDock applications v.4.2 [31] were also used. The search grid expanded beyond the chosen target proteins; polar hydrogen was applied to the protein and Kollman charges were allocated. The target protein complex (PBD I.D: 2AXJ) was docked with 30 anti-rheumatoid arthritis compounds, the protein is called a rigid body, and the compounds being versatile. The protein has two chains (A and B), because chain A and chain B are similar, the hit drugs can attach to either of them. In this research, chain B was used. The quest was expanded over the entire receptor protein with a 60 × 60 × 60 docking range. Affinity maps for all types of atoms current, as well as an electrostatic map, were computed with a grid spacing of 0.375 Å. The Lamarckian Genetic Algorithm was investigated; samples of 150 individuals with a mutation rate of 0.02 have developed over ten existences. The evaluation of the outcomes was carried out by mapping out the different complexes about the necessary binding energy. Cluster analysis focused on the root mean square variance values were subsequently performed to starting geometry, and the most reliable solution was found to be the lowest energy conformation of the more populated cluster.

Molecular dynamics simulations

Calculations of molecular dynamics (MDs) simulations are conducted using VMD [32] and NAMD [33] and CHARMM force field [34]. Interaction parameters have been calculated using the CHARMM27 force field (par-all27-prot-lipid.inp). The protein was dissolved with explicit water at a concentration of 0.1 M NaCl salt for neutralization. Minimization was performed to simplify the original protein structure. To confirm the structure to the force field, the application of solvents and, in particular, to minimize steric clashes that may occur in the device. To perform the simulation, this procedure supplied the computer with the lowest energy. After the minimization, it was characterized by the achievement of energy convergence. After that, the device temperature was steadily heated from 0 K to 310 K in 100 ps. At the end of the day, the system was calibrated at 310 K for 100 ps with the NVT ensemble. On the DELL INSPIRON; Pentium ® Dual-Core CPU T4500 @ 2.30 GHz and 3 GB RAM, 64-bit-Operating System, x64-based CPU, the MD simulation, and analysis results were conducted. In the NAMD documentation (www.ks.uiuc.edu /Research/namd/), a full summary of the input parameters are given. See (www.ks.uiuc.edu/Research/namd /current /ug/) for more detail on running MD simulations with NAMD.

Results and discussion

With a training set of 18 molecules, a n/p ratio of 6, a deviation fraction of 0.796, the QSAR model was developed, representing an accuracy of approximately 80%, leaving one out cross-validation with a value of 0.68, R2adjusted = 0.7523 consistency factor Q = 0.63 and an *F*-value of 18.20905 (DF:3, 14). The compounds have negative ATS3e (Broto-Moreau autocorrelation of lag 3 (log function) weighted by Sanderson electronegativity) and ATSC1c (Average based Broto-Moreau autocorrelation-lag 3 / weighted by charges) values for improved inhibition activity, according to the established QSAR model. The developed QSAR model indicates that a positive nH (number of hydrogen atoms) would be beneficial for the behavior of the compounds. The developed QSAR model also indicates that replacements result in improved action or better bioactivity values at various locations in the rings. The replacements suggested are -CH_3, -NH_2, isopropyl group, -OH, in the ortho-, meta- and para- positions. Along with the model interpretation, a detailed description of the derivatives often acts to go for the two substitutes simultaneously at meta- and para-positions for enhanced and potent bioactivity values. For compounds used for the development of the QSAR equation, an evaluation of observed values and expected values of pIC_{50} is shown in Fig. SM2.

Multiple linear regression (MLR) equation

$$pIC50 = 5.07032(\pm 0.27582) + 0.12145(\pm 0.02428)nH - 0.00404(\pm 0.0006)ATS3e - 2.28527(\pm 0.45238)ATSC1c \quad \text{(Model 1)}$$

SEE = 0.22476; R^2 = 0.796; $R^2_{adjusted}$ = 0.75228; F =18.20905 (DF:3, 14); Q^2 = 0.67999; PRESS = 1.10942; SDEP = 0.24826; $N_{training}$ = 18

External Validation Parameters: R^2= 0.82093; $R_0{}^2$= 0.81351; reverse $R_0{}^2$= 0.72325; $R^2_{m(test)}$= 0.75024; reverse $R^2_{m(test)}$ = 0.56436; average $R^2_{m(test)}$= 0.6573; delta $R^2_{m(test)}$= 0.18588; RMSEP = 0.20407; R^2_{pred}= 0.81482; Q^2_{f1}= 0.81482; Q^2_{f2}= 0.8128; N_{test}= 6

Golbraikh and Tropsha [19] and Golbraikh et al. [20] acceptable model criteria's

1. Q^2 = 0.67999 Passed (Threshold value Q^2 > 0.5)
2. r^2 = 0.82093 Passed (Threshold value r^2 > 0.6)
3. |r0^2-r'0^2| = 0.09026 Passed (Threshold value |r0^2-r'0^2| < 0.3)
4. k = 1.00252; [(r^2-r0^2)/r^2] = 0.00903
5. [0.85<k<1.15 and ((r^2-r0^2)/r^2) < 0.1]
6. k' = 0.99582; [(r^2-r'0^2)/r^2] 0.11898 Passed (Threshold value: [0.85<k'<1.15 and ((r^2-r'0^2)/r^2) < 0.1])

The feasibility of the model was confirmed by the significance of the adjusted R^2 ($R^2_{m(test)}$ = 0.7502) and the novel parameter R^2_p= 0.7079, which was similar to the standard R^2(0.796). As we can see from the internal and external validation parameters, by comparing the related parameters of the method of Golbraikh and Tropsha [19,20] and Roy's method [21], we find that they both follow the requirements, which means that our 2D-QSAR models is accurate and has good statistical significance. Both the compounds in the training set and the evaluation set are closely correlated with their respective experimental anti-rheumatoid arthritis activities (Fig. SM2).

Correlations between the three descriptors are shown in Table 2 as a correlation matrix. The three parameters nH, ATS3e, and ATSC1c, are all correlative to a substituent, moreover, nH belongs to substituent's property, and the difference of substituent must induce the difference of ATS3e and ATSC1c too. It suggests that substituent should be a very important active zone; therefore, we can design some new compounds with high bioactivity by modifying the substituent. ATS3e has the highest mean effect of −1.456 than the other properties which means that these properties have a large effect on the pIC_{50} of the studied compounds. 50.4% of the ATS3e descriptor will be removed to increase bioactivity while 32.7% of the nH

Table 2
Correlation matrix, mean effect, and variance inflation factor.

Descriptors	nH	ATS3e	ATSC1c	MF	VIF	Contribution (%)
nH	1			0.9456	3.6604	32.71
ATS3e	0.8541	1		−1.4560	3.8023	−50.36
ATSC1c	−0.2959	−0.4110	1	−0.4896	1.1271	−16.93

Table 3
Values after several Y-scrambling tests.

Model	R	R^2	Q^2	Random Models Parameters	
Original	0.892188	0.795999	0.679985	Average r	0.407879
Y-scrambling 1	0.519484	0.269863	−0.22599	Average r^2	0.178969
Y-scrambling 2	0.53791	0.289347	−0.05168	Average Q^2	−0.39182
Y-scrambling 3	0.421296	0.17749	−0.18113	cRp^2	0.707946
Y-scrambling 4	0.296705	0.088034	−1.07501		
Y-scrambling 5	0.153083	0.023435	−0.5916		
Y-scrambling 6	0.501631	0.251633	−0.06981		
Y-scrambling 7	0.346289	0.119916	−0.57911		
Y-scrambling 8	0.435616	0.189761	−0.38573		
Y-scrambling 9	0.481143	0.231499	−0.47615		
Y-scrambling 10	0.385632	0.148712	−0.28202		

descriptors should be added to increase the bioactivity. The utmost VIF value of the selected descriptors was 3.8023 indicate that the obtained model has statistical significance.

The applicability domain was generated by determining the leverage values for each compound. The Williams diagram, i.e., the map against leverages (x-axis) of standardized residuals (y-axis) for each of the total set compounds (Fig. SM2). The applicability domain was established from this plot within a squared region within ± 3 standard deviations and a leverage threshold h * = 0.67 (h* = 3(K + 1)/n, where K = the number of model parameters and n = the number of compounds) If seen from Fig. SM3, all the compounds were found inside the square zone. For future research, the predicted evidence on anti-rheumatoid arthritis behavior should only be considered valid for those molecules that fall within the applicability domain for which the model was created.

The Y-scrambling test is designed to test the MLR model's robustness. Ten models were produced; as seen in Table 3, all of them were expected to have lower R^2 and Q^2 values than the initial model [35]. The average R^2 and Q^2, respectively, were 0.181 and −1.018. The findings of Table 2 demonstrate that the MLR approach obtains an acceptable model and that the model produced is statistically valid and stable.

3D-QSAR analysis

Base on the Open3DALIGN-based alignment, a model with the highest predictive ability is considered reliable which meets the 3D-QSAR criterion. We know that one of the better parameters for determining whether a successful 3D-QSAR model is appropriate is Q^2(> 0.5). The CoMFA (FFD) model combined with steric and electrostatic fields was obtained by the study of the obtained 3D-QSAR models. The steric and electrostatic contributions are 79% and 21%, respectively in CoMFA (FFD). In steric contribution, the red polyhedron contributes 80% and also indicate a region where increased steric bulk (negative charge) are predicted to enhance activity. The blue polyhedrons which contribute 20%, indicate regions where decreased steric bulk are predicted to enhance activity. In electrostatic contribution, the dark green polyhedrons which contribute 61% regions where higher electron density (negative charge) are predicted to enhance the activity, while the yellow polyhedrons (partial positive charge) which contribute 39%, lower electron density are predicted to enhance activity as shown in Fig. 1. The statistical parameters derived from the study of CoMFA are: R^2 = 0.9982; SDEC = 0.0182; F-test = 1305.3705; Q^2 = 0.105; $SDEP_{int}$= 0.4152; R2pred = 0.0791; $SDEP_{ext}$ = 0.4551. However, the Q^2 values of these models do not satisfy the criteria (Q2 > 0.5), suggesting that the 3D-QSAR model is undesirable.

Steric and electrostatic inputs were found to be 67% and 33%, respectively, in the CoMFA (SRD) model. In steric contribution, the red polyhedron contributes 80% and also indicate a region where increased steric bulk (negative charge) are predicted to enhance activity. The blue polyhedrons which contribute 20%, indicate regions where decreased steric bulk are predicted to enhance activity. In electrostatic contribution, the dark green polyhedrons which contribute 48% regions where higher electron density (negative charge) are predicted to enhance the activity, while the yellow polyhedrons (partial positive charge) which contribute 52%, lower electron density are predicted to enhance activity as shown in Fig. 2. The statistical parameters obtained from CoMFA (SRD) analysis are: R^2 = 0.9978; SDEC = 0.0207; F-test = 1074.2993; Q^2 = 0.3325; $SDEP_{int}$= 0.3585; R^2_{pred}= 0.0983; $SDEP_{ext}$= 0.4503. Thus, the CoMFA (SRD) models were developed to produce the optimum 3D-QSAR model with the combination of two descriptor fields. The Q^2 values of these models do not follow the criteria (Q^2 > 0.5), which also indicates that the 3D-QSAR model is inappropriate but better than the CoMFA (FFD).

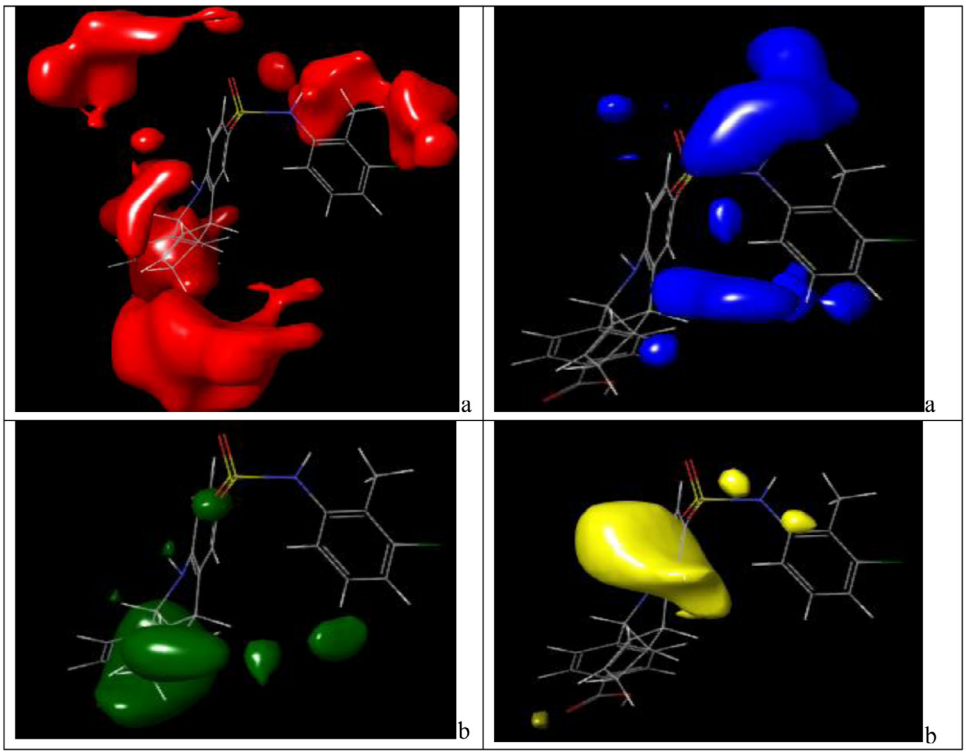

Fig. 1. 3D equipotential maps of CoMFA (FFD) (a) steric and (b) electrostatic field contour maps with compound 17.

CoMFA model, contour plots

The contour plots of the CoMFA (FFD) steric and electrostatic fields are presented in Fig. 1 for the model of RA operations. For chasteness, only the most active compound (compound 17) contour map is seen. Red and blue contours reflect regions in this figure (Fig. 1) where steric bulk classes preferred and hindered the activity, respectively. The red outline, seen in Fig. 1, shows that for high inhibitory action, broad substituents -COOH and phenyl group of the compound 17 are important. The red contours at the location of -NH and -CNH show regions where the activity is decreased by an increase in the positive charge. Thus, the substituents are ideally broad and charged negatively. Blue contours are also found near the -NH, O = S = O, and -CNH, respectively, so broad substituents are also favored. Green contours mean that functional groups that have an electronegative character, such as carboxylic acid (-COOH) and hydroxyl (-OH), are useful, as seen in the figure. Positively charged groups, such as bases and OH groups, are also beneficial for having high activity, as seen by the yellow contours (Fig. 1).

To define the essential regions in the 3D space surrounding the molecules, CoMFA (SRD) contour maps have been developed such that changes can be made in certain areas to enhance the inhibitory effect, which can be used to enhance the inhibitory action of RA. The green color indicates a favorable component, which leads to an increase in biological activity by maintaining the bulkier population, while the yellow color indicates a decrease in biological activity due to the bulkier area. In contrast, colors of red and blue are seen in the steric contour diagram. The colors red and blue signify a favorable and unfavorable area. Here, the red color region reveals that a negative charge increases biological activity, while the blue color region reveals a favorable charge for biological activity to increase. A broad green contour cloud in an area at -NH group suggests that the introduction of the bulky group is preferred for activity, whereas bulky substituent is not preferred at the para location of the phenyl group. This is apparent from compound 9, 11,13,21,29, and 30 experimental activity benefits, of which the bulky group results in the least RA inhibitory activity. Fig. 2(b), uses CoMFA (SRD) to show the electrostatic contour map. Green and yellow contour maps are used to represent electrostatic fields. A green contour near the carboxylic ring, SO_2, and the group at toluene suggests the need for a negatively charged electrostatic interaction substitution with a receptor to illustrate strong inhibitory function. Although the yellow contour observed at the phenyl ring in the SO_2 area suggests that the incorporation of the positively charged substituent favored the activity at the phenyl ring.

Molecular docking interaction using PyRx (AutoDock Vina)

Docking tests of compounds with enzyme structures of rheumatoid arthritis-related *T* cell receptor beta chains produced a docking score ranging from −8.0 to −5.1 (Table 4) suggesting stable associations of enzyme-substrate. The high-

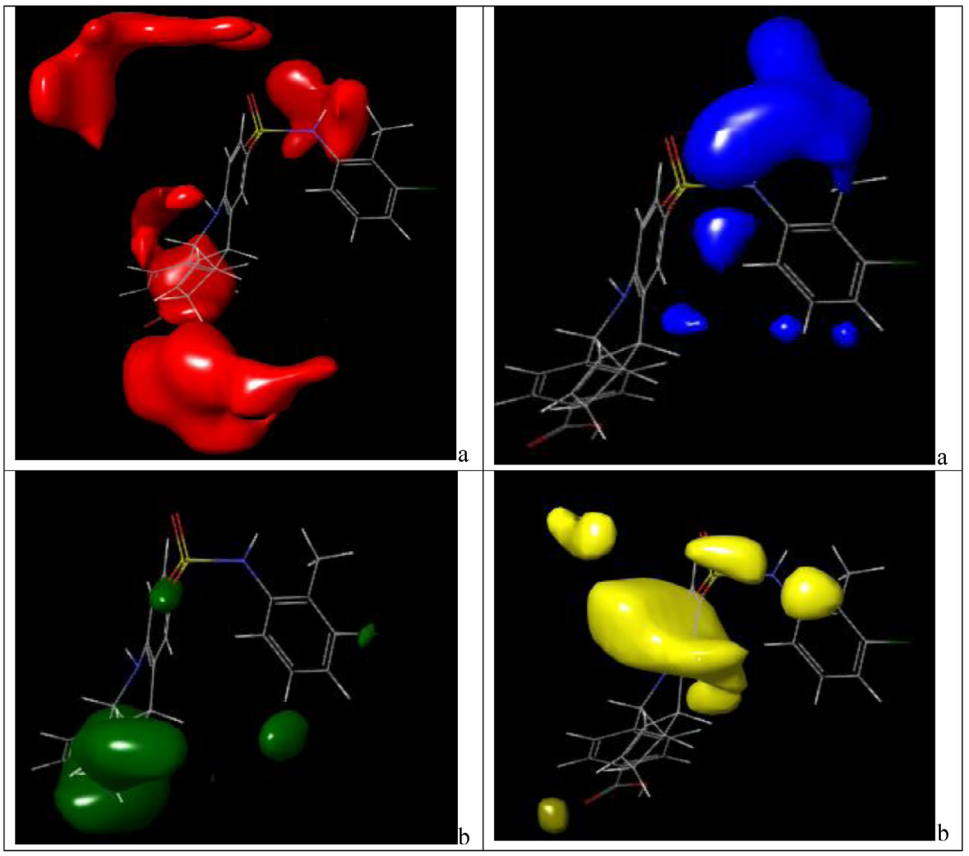

Fig. 2. 3D equipotential maps of CoMFA (SRD) (a) steric and (b) electrostatic field contour maps with compound 17.

est inhibition of the enzyme structures of *T* cell receptors in the dataset was shown by compound 4 (4-((6-oxo-3-(1,4-dioxa-8-azaspiro[4.5]decan-8-yl)−6H-anthra[1,9-cd]isoxazol-5-yl)amino)butanoic acid) with the best docking performance (−8.0 kcal/mol). The compounds interact with beta-chain *T* cell receptors by binding to the amino acid residues of Pro176, Gln175, Asp173, Pro174, Val155, Thr172, Glu156, Leu157, Gln211, Ser158, Val170, His167, Trp16, and Glu165, as seen in Fig. 3. The *T* cell receptor inhibitory assay and molecular docking analysis show that *T* cell receptor inhibits compound 4 may be due to its hydrogen bond association with amino acid residues His176 (4.64 Å), Glu165 (5.52 Å), Pro174 (4.99 Å), and Pi-Alkyl association with Val170 (5.54 Å). Compound 17, which is the next highest binding energy close to compound 4 has a docking score of −7.8 kcal/mol by binding with Pro176, Asp152, Tyr12, Val155, Glu156, Asp173, Pro174, Thr172, Leu157, Ser171, Val170, Trp159, His167, Val166, and Glu165 amino acid residues as shown in Fig. 4. The docking studying reveals that compound 17 has hydrogen bond interactions with Thr172 (3.64 Å), Pro174 (4.71 Å), Tyr12 (5.26 Å), and Pi-Alkyl interaction with Val170 (5.44 Å) and Pro176 (6.49 Å) amino acid residues, respectively.

Molecular docking interaction using AutoDock4.2

Docking study with AutoDock 4.2 software [36,37]. All the 30 molecules were docked to the 3D structure of *T* cell receptor beta chains (PDB code: 2AXY). In the active site, identified by a grid box, the structures were docked using a Lamarckian genetic algorithm ($x = 60$, $y = 60$, $z = 60$). In Table 4, the docking effects of the compounds are shown. Compound 17 has the most binding energy of −8.55 kcal / mol using AutoDock4.1 and is the highest performing data collection molecule with a binding energy of −6.87 kcal/mol using PyRx (AutoDock-Vina) software. Compound 17 (3-((2-(ethoxycarbonyl)−7-(2–hydroxy-5-methoxybenzoyl)−3-methyl-4-oxo-4H-pyrido[1,2 a]thieno[2,3-d]pyrimidin-9-yl)thio)propanoic acid) using AutoDock4.2 has hydrogen bond interaction between Tyr215 (5.81 Å) and Ser107 (4.90 Å) and Alkyl-pi-Alkyl groups of Pro10, Tyr215, and Lys11, respectively, while compound 4 has hydrogen bonding interaction with Leu157 (5.00 Å), Glu156 (5.40 Å), Lys11 (4.12 Å and 4.84 Å), pi-Anion with Glu165 and pi-Sigma with Ser158, respectively. mpound 4 and 17 closure contacts linked to residues at the active site as shown in Figs. (5 and 6). Compound 17 had the highest Ligand efficiency (LE) values than compound 4 and was estimated to be the best compound.

Table 4
The predicted binding affinity of the molecules and their inhibition constant.

Compound No.	PyRx (Vina)	AutoDock4.2	Ligand_efficiency	Inhib_constant (uM)
1	−6.2	−4.89	−0.23	259.18
2	−6.1	–	–	-
3	−5.5	−6.57	−0.55	15.4
4	−8.0	−6.87	−0.2	9.22
5	−5.2	−3.99	−0.29	1.19
6	−5.1	−4.28	−0.29	727.19
7	−4.5	−4.22	−0.22	802.97
8	−7.1	−5.59	−0.2	79.83
9	−7.1	−7.65	−0.26	2.46
10	−7.4	−6.49	−0.19	17.37
11	−6.4	−6.26	−0.24	25.77
12	−6.2	−5.18	−0.18	160.38
13	−6.4	−6.15	−0.19	31.04
14	−6.5	−5.79	−0.15	56.84
15	−6.9	−7.56	−0.2	2.88
16	−5.9	−5.52	−0.15	89.57
17	−7.8	−8.55	−0.25	538.45
18	−6.0	−5.8	−0.2	56.08
19	−6.6	−5.14	−0.22	169.85
20	−5.1	−4.15	−0.24	902.92
21	−7.5	−5.63	−0.21	74.6
22	−7.3	−7.46	−0.21	3.42
23	−7.5	−5.73	−0.18	63.59
24	−7.2	−6.07	−0.21	35.47
25	−6.2	−4.25	−0.13	770.27
26	−6.8	−5.57	−0.18	82.71
27	−6.9	−5.15	−0.2	116.59
28	−6.7	−4.99	−0.15	218.15
29	−7.6	−5.18	−0.16	159.17
30	−7.1	−4.9	−0.14	254.21

Molecular dynamics simulations

To consider the properties of the molecules' assemblies in terms of their arrangement and the microscopic interactions between them, molecular dynamic simulations were carried out. MDs simulations model the motion of a collection of the chemical system over time, according to the law of classical physics. After 2600 ps, MD simulations reached equilibrium. The enzyme was stabilized at −80,197, 16,837.2, −97,034.1 kcal/mol, total energy, kinetic energy, and potential energy, respectively. Observation of overall energy, kinetic energy, and possible changes in energy during the aggregation process revealed that an increase in the energy of the device was followed by the process. When the simulation was stopped, the energy changes resulted in phases before reaching a limit. This suggests that a metastable system is formed and a stable structure with a minimum energy state is consequently pursued. To the investigation of the effects of the inhibitor on the simulated enzyme, the same ligands (compound 4 and 17) were docked on simulated protein as shown in Figs. (6 and SM4) which resulted in the decline in the docking interaction. Molecular docking interaction of compound 4 Fig. SM4, shows no hydrogen bond interaction, while compound 17 with one long-distance hydrogen bond interaction with Lys11 (5.84 Å) as shown in Fig. SM5.

In the first 900 ps, RMSD values for the protein have emerged and then remain constant for the remainder of the simulation time. This result indicates that our complexes were stable throughout the experiment. The protein RMSD (PDB ID: 2axj) was determined from a trajectory of 1000 ps, where the RMSD points fluctuated by 1.81 Å (Fig. SM6A). This fluctuation can happen because each residue's interactions, such as electrostatic contacts, hydrogen bonds, hydrophobic interactions, and even water molecules, play a role in structural reorganization at the protein interface. After running 500,000 at 1000 ps NVT, the temperature fluctuated near 211.86 K (Fig. SM6B).

ADMET and drug-likeness

The essence of this correspondence was to ascertain the physicochemical domain of the compounds being investigated. The properties involved are polar surface area (PSA), number of rotatable bonds (NumRot), hydrogen bond donors (NumHBA), hydrogen bond acceptors (NumHBD), and molecular weight (MW). The results were calculated using Spartan'14 v1.1.4, PowerMV v0.61 (www.niss.org/PowerMV), and MedChem Designer v3.0 (www.simulations-plus.com) are shown in Table 5. These criteria permit the determination of oral absorption and membrane permeability when the measured molecule is accompanied by (1) Veber's two rules [38] as rotatable bonds $\leq$ 10 and polar surface area $\leq$ 140 $Å^2$ are highly likely to have strong oral bioavailability in the human intestine. (2) Lipinski's rule of five as molecular weight (MW) $\leq$ 500 Da, logP $\leq$ 5 octanol-water partition coefficient, H-bond donors, atoms of nitrogen or oxygen with one or more atoms

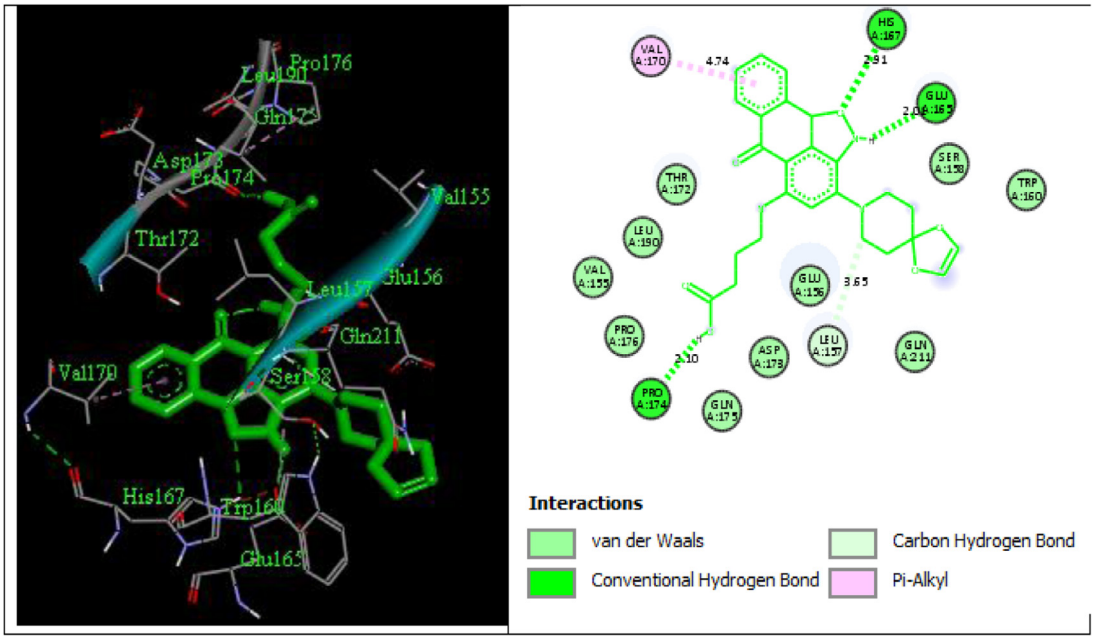

Fig. 3. Docking interactions of compound 4 in the *T* cell receptor binding site.

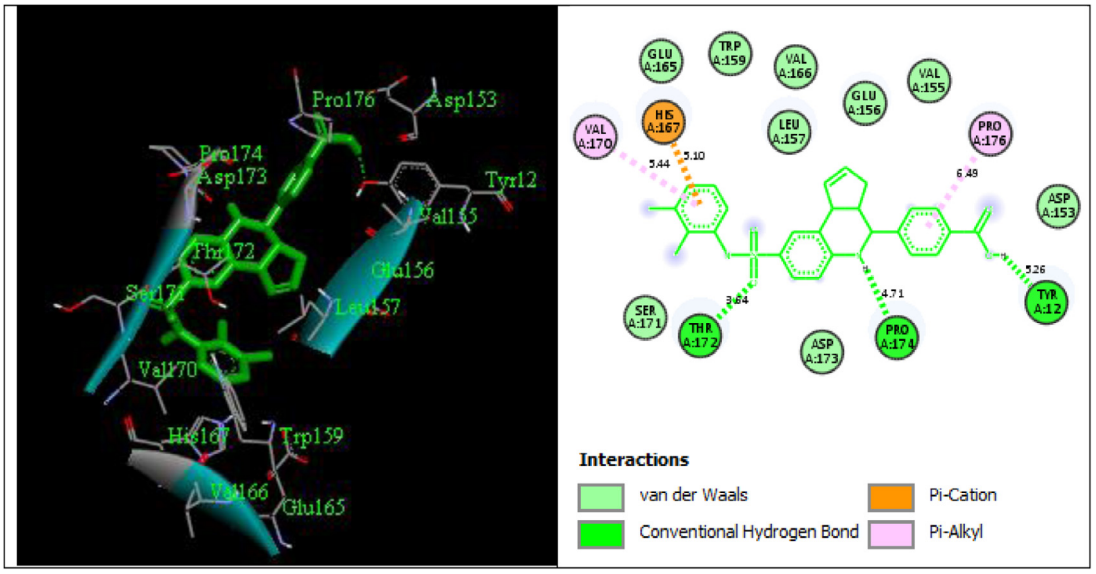

Fig. 4. Docking interactions of compound 17 in the *T* cell receptor binding site.

of hydrogen (HBD) $\leq$ 5 and H-bond acceptors, atoms of nitrogen or oxygen (HBA) $\leq$ 10 [39]. Molecules that lack more than one of these requirements may have bioavailability issues and a high risk of not being drug-like. From the ADMET drug-like analysis, the emphasis is placed on compound 4 and 17 (Fig. SM7). From the findings in Table 5, it was observed that compounds 4 and 17 conform to the law of Veber and Lipinski, which indicates strong oral bioavailability involving a balance between the aqueous solubility of compounds and their capacity to passively disperse through the various biological barriers.

Electrostatic potential (EP) surface of compound 4 and 17

The electrostatic potential is used to distinguish electron-rich areas on the surface that are subject to electrophilic attack from electron-poor regions that are subject to nucleophilic attack. In terms of the electrostatic surface, the electrostatic potential contour map shows the molecular distance, form, positive, negative, and neutral electrostatic potential areas, which

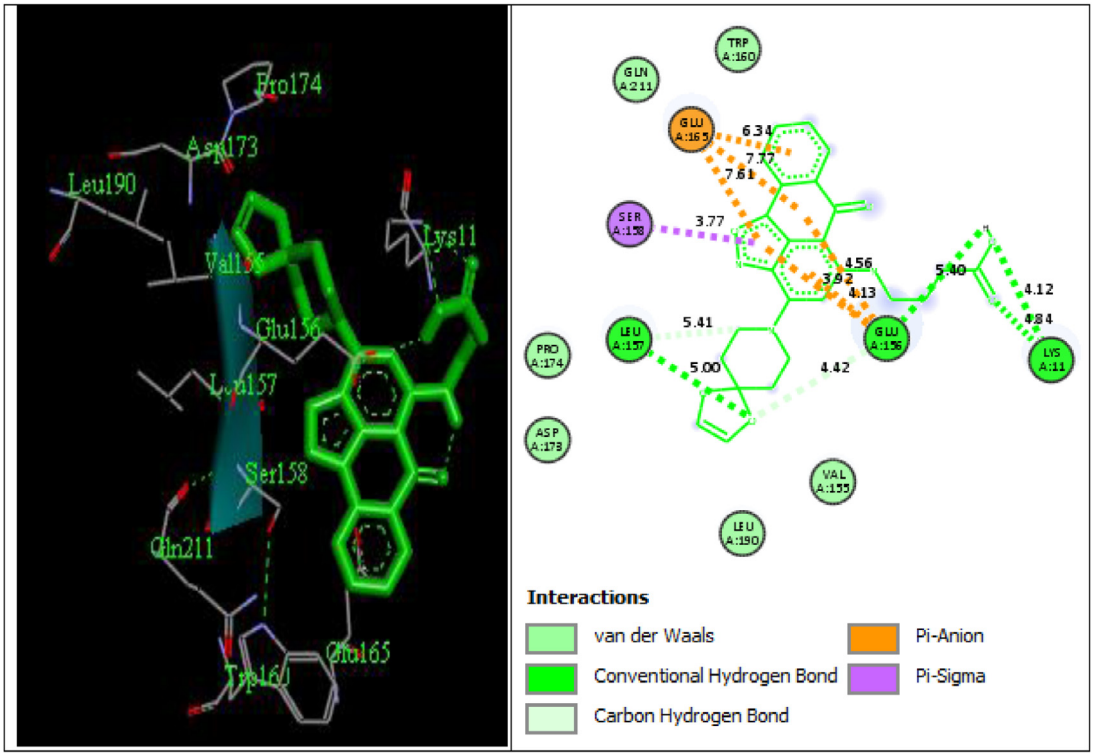

Fig. 5. The mode of binding between compound 4 and the *T* cell receptor binding pocket.

describes the study of the molecular structure and its relationships with physicochemical properties. The EP surface map of compounds 4 and 17 (Fig. SM8) reveals the three regions around the tow cyclic nitrogen atoms marked by red color (negative electrostatic potential) that describe the capacity to the electrophilic assault on these positions, the blue color (positive electrostatic potential) around the two hydrogen atoms defines the susceptibility of these regions to nucleophilic attack. Finally, the green color situated between the red and blue regions clarifies the neutral electrostatic potential surface. The difference in the electrostatic potential generated by a molecule is primarily responsible for binding the drug to its active sites (receptor) since it is usually assumed that the binding site would have opposite electrostatic potential areas.

The heat of formation of compound 4 has the highest negative value of −497.33 Kcal/mol, while compound 17 is −321.09 Kcal/mol. The calculated HOMO and LUMO of compound 4 and 17 are −8.32 eV and −1.82 eV (donor), −9.00 eV and −0.90 eV (acceptor), respectively. Fig. SM9 demonstrates the energies of compound 4 and compound 17 of HOMO (Highest Occupied Molecular Orbital) and LUMO (Lowest Unoccupied Molecular Orbital). The positive phase is red and green is the negative one. For a comparative study between the effects of electron donors and acceptors, we have studied compounds 4 and 17 in the present study and also determined how the molecule interacts with the receptor. The compound 4 E_{HOMO} level demonstrates that both the positive and negative phases are distributed symmetrically. The E_{LUMO} is almost distributed without hydrogen group over the molecule, and both positive and negative phases are symmetrically distributed but compared to E_{HOMO}, the E_{LUMO} reflects a great deal of antibonding π^* character. Compound 4 take a smaller $E_{HOMO} - E_{LUMO}$ energy gap (−6.5 eV) related to compound 17, thus, compound 4 is the most active which promote intra-molecular charge transfer (ICT), and likewise possesses higher chemical potential (μ), which contains information about electron transfer, lower global hardness (η) which contains the stability of the compound, higher electron affinity (A), small ionization energy (I), thus, compound 4 is a strong electrophile compare to compound 17 (Table 6).

Fig. SM10 shows that compound 4 has the probability of success in maximizing potency, stability, and permeability since it falls inside the golden triangle plot [40] with the baseline of S+LogD = −6 to 6 at MW = 500. These show that compound 4 has great penetrating power, optimized clearance, and oral absorption compared to compound 17.

Conclusion

In this work, 2D-QSAR, COMFA, docking, MD simulations, ADMET, and Frontier Molecular Orbital's (FMOs) methods were performed to investigate structural properties and pathways for attachment of 30 inhibitors of *T* cell receptor beta chains related to rheumatoid arthritis (RA). In this analysis, the QSAR model evolved is robust and has a high predictive ability, revealing the primary systemic factors influencing the inhibitory behavior of RA. The compound 4 (4-((6-oxo-3-(1,4-dioxa-8-azaspiro[4.5]decan-8-yl)−6H-anthra[1,9-cd]isoxazol-5-yl)amino)butanoic acid) and compound 17 (3-((2-(ethoxycarbonyl)−7-

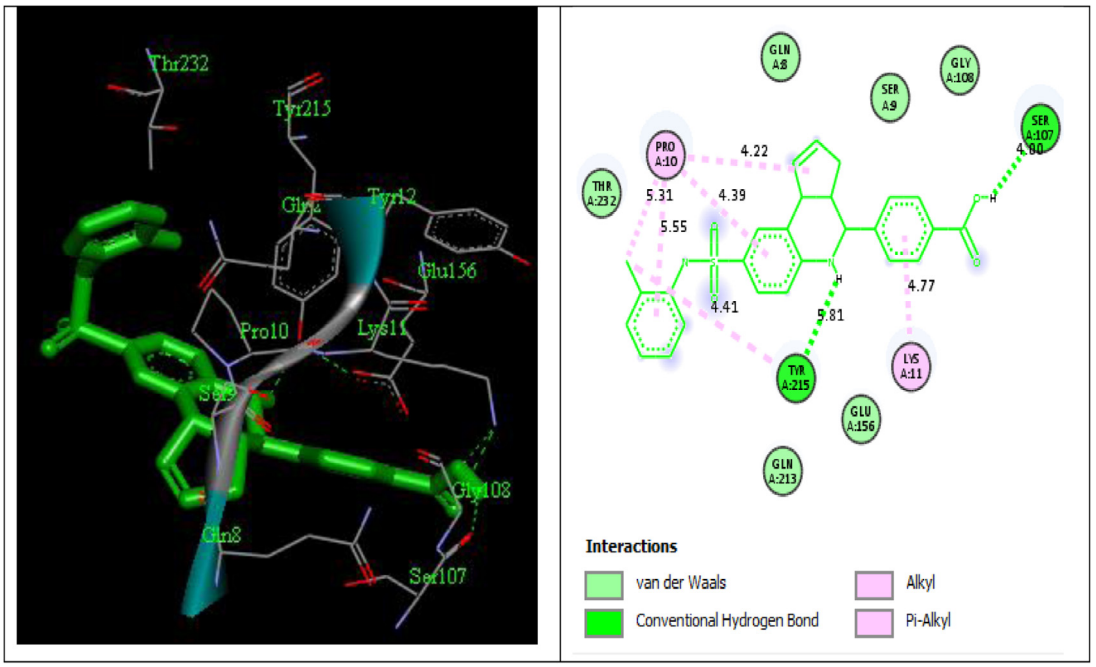

Fig. 6. The mode of binding between compound 17 and the *T* cell receptor binding pocket.

Table 5
Drug properties of the studied derivatives.

Row	XLogP	PSA	NumRot	NumHBA	NumHBD	MW	BBB	BadGroup	MLogP	*S*+LogP	*S*+LogD	M_NO	HBDH
1	0.65	82.67	2	2	0	283.291	0	0	1.904	1.017	1.017	7	0
2	1.575	69.56	1	4	3	231.251	1	0	−0.456	0.659	−1.81	4	3
3	2.276	123.32	4	6	1	307.332	0	0	1.658	2.506	2.506	6	0
4	3.168	114.13	6	7	2	463.49	0	0	2.094	3.214	0.249	9	2
5	−1.636	82.67	0	2	0	193.166	0	0	−0.199	−1.195	−1.195	7	0
6	−1.321	82.67	0	2	0	207.193	0	0	0.135	−0.957	−0.957	7	0
7	−1.958	112.08	9	7	3	297.397	0	1	−4.472	−4.224	−4.224	7	3
8	2.845	137.18	6	4	1	430.532	0	1	1.106	3.039	0.172	6	1
9	5.231	108.77	5	7	2	446.912	0	1	0.983	3.369	0.769	7	2
10	4.907	167.88	9	6	2	496.57	0	1	2.475	4.352	2.724	9	2
11	4.582	116.23	6	8	2	371.374	0	2	2.752	3.878	3.784	8	2
12	3.181	113.96	7	8	3	461.268	0	1	1.985	3.486	0.767	8	3
13	4.863	117.09	10	10	3	434.408	0	2	2.564	3.482	3.424	10	3
14	5.901	150.38	12	8	3	530.606	0	1	2.31	4.534	1.122	9	3
15	5.118	129.52	9	7	1	527.602	0	0	2.096	2.896	0.64	8	2
16	4.011	198.04	10	8	2	542.591	0	0	1.08	2.982	0.233	10	2
17	5.317	103.88	5	6	3	495	0	0	3.615	4.641	1.539	6	3
18	6.074	70.32	6	4	1	452.279	0	1	4.419	5.841	2.534	4	1
19	0.922	82.67	3	2	0	309.329	0	0	0.756	−0.07	−0.07	7	0
20	−0.499	82.67	2	2	0	235.247	0	0	0.756	−0.07	−0.07	7	0
21	4.436	81.41	5	3	1	379.807	0	0	3.58	4.568	4.567	7	1
22	6.016	141.15	10	7	3	506.971	0	1	2.649	4.591	1.409	8	1
23	2.39	156.82	6	10	3	441.445	0	1	1.433	1.543	−0.198	10	3
24	4.426	91.64	4	6	2	407.813	0	0	2.699	3.976	1.403	7	2
25	1.9	130.47	9	8	2	472.544	0	1	1.209	2.976	0.27	8	2
26	5.348	99.42	8	6	2	432.863	0	0	3.195	5.284	2.091	6	2
27	3.252	128.36	4	6	1	371.374	0	1	1.9	2.261	−1.131	8	2
28	5.594	108.65	10	7	2	521.367	0	1	2.941	5.548	2.076	7	2
29	4.23	120.85	6	10	2	445.387	0	1	2.317	3.929	0.311	10	3
30	1.029	141.55	8	9	2	497.554	0	1	0.752	2.519	0.18	9	2

Table 6
Thermodynamic proprieties for compound 4 and compound 17.

No.	Area (Å^2)	H^o(au)	S^o(J/mol)	G^0(au)	A (eV)	μ (eV)	I (eV)	η (eV)	Dipole moment (debye)
Cpd 4	455.51	0.304	653.94	0.230	1.82	−5.07	8.32	3.25	2.70
Cpd 17	477.	0.343	667.11	0.267	0.90	−4.95	9.00	4.05	4.46

(2–hydroxy-5-methoxybenzoyl)−3-methyl-4-oxo-4H-pyrido[1,2-a]thieno[2,3-d]pyrimidin-9-yl)thio)propanoic acid) were absolutely placed within the *T* cell receptor binding pocket of the protein and forms relevant interactions with the amino acid, which lead to the competitive binding suppression of the protein cell. The generated COMFA (FFD) is rich and extremely validated. The ADMET and FMOs model that accurately reveals the features of RA inhibitors. Lastly, the MD simulation was conducted in a water environment to create the 3D structure of a protein. The protein with the ligands was used to realize the conformation shifts of the complex in MD simulation. From the above methods, compound 4 (4-((6-oxo-3-(1,4-dioxa-8-azaspiro[4.5]decan-8-yl)−6H-anthra[1,9-cd]isoxazol-5-yl)amino)butanoic acid) and compound 17 (3-((2-(ethoxycarbonyl)−7-(2–hydroxy-5-methoxybenzoyl)−3-methyl-4-oxo-4H-pyrido[1,2-a]thieno[2,3-d]pyrimidin-9-yl)thio)propanoic acid) have stronger BBB permeation, strong intestinal permeability and oral bioavailability, they have the optimal *in vitro* ADME and protection characteristics. Through these trials, we have gained useful insight into the conditions regulating the efficacy of rheumatoid arthritis-related inhibitors for the *T* cell receptor beta chains. Compound 4 and compound 17 are potential inhibitors of *T* cell receptor and are drug-like compounds. These analyses might offer worthful perception for further development of novel competent rheumatoid arthritis inhibitors.

Financial support and sponsorship

Nil.

Declaration of Competing Interest

There are no conflicts of interest.

Acknowledgment

The authors gratefully acknowledge the Department of Chemistry, Ahmadu Bello University, Zaria (Samaru, Zaria-Nigeria); for computational studies, and as part of the Ph.D. thesis.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.sciaf.2021.e01088.

References

[1] S. Raman, Computer-aided drug design (CADD) for Rheumatoid arthritis using QSAR approach, Int. J. 1 (6) (2013) 608–613.
[2] R.V. Antre, R.J. Oswal, S.S. Kshirsagar, P.P. Kore, M.M. Mutha, 2D-QSAR studies of substituted pyrazolone derivatives as anti-inflammatory agents, Med. Chem. 2 (2012) 126–130, doi:10.4172/2161-0444.1000127.
[3] R.K. Prasad, R. Sharma, In silico 2D-QSAR analysis of 2-ArylPyridine inhibitors of mitogen-activated protein kinase-2 as anti-rheumatoid arthritis agents, AJPCT 1 (1) (2013) 001.
[4] M. Zaka, B.H. Abbasi, S. Durdagi, Novel tumor necrosis factor-α (TNF-α) inhibitors from small molecule library screening for their therapeutic activity profiles against rheumatoid arthritis using target-driven approaches and binary QSAR models, J. Biomol. Struct. Dyn. (2018), doi:10.1080/07391102.2018.1491423.
[5] D.S.V.G.K. Kaladhar, K.V.V.V. Satyanarayana, A.K. Chaitanya, S.A.K.Z. Hussain, Clinical analysis, drug designing and QSAR studies on rheumatoid arthritis, Int. J. Pharm. Bio Sci. 1 (4) (2010) 132–138.
[6] J.R. O'Dell, T.R. Mikuls, T.H. Taylor, V. Ahluwalia, M. Brophy, S.R. Warren, R.A. Lew, A.C. Cannella, G. Kunkel, C.S. Phibbs, A.H. Anis, S. Leatherman, E. Keystone, Therapies for active rheumatoid arthritis after methotrexate failure, N. Engl. J. Med. 369 (4) (2013) 307–318, doi:10.1056/NEJMoa1303006.
[7] R. Fleischmann, A.L. Pangan, I.H. Song, E. Mysler, L. Bessette, C. Peterfy, P. Durez, A.J. Ostor, Y. Li, Y. Zhou, A.A. Othman, M.C. Genovese, Upadacitinib versus placebo or adalimumab in patients with rheumatoid arthritis and an inadequate response to methotrexate: results of a phase III, double-blind, randomized controlled trial, Arthritis Rheumatol. 71 (11) (2019) 1788–1800, doi:10.1002/art.41032.
[8] J.M. Kremer, W. Rigby, N.G. Singer, C. Birchwood, D. Gill, W. Reiss, J. Pei, M. Michalska, Sustained response following discontinuation of methotrexate in patients with rheumatoid arthritis treated with subcutaneous tocilizumab: results from a randomized, Controll. Trial. Arthritis Rheumatol. 70 (8) (2018) 1200–1208, doi:10.1002/art.40493.
[9] J.A. Singh, K.G. Saag, S.L. Bridges, E.A. Akl, R.R. Bannuru, M.C. Sullivan, E. Vaysbrot, C. McNaughton, M. Osani, R.H. Shmerling, J.R. Curtis, D.E. Furst, D. Parks, A. Kavanaugh, J. O'Dell, C. King, A. Leong, E.L. Matteson, J.T. Schousboe, B. Drevlow, S. Ginsberg, J. Grober, E.W. St Clair, E. Tindall, A.S. Miller, T. McAlindon, 2015 American college of rheumatology guideline for the treatment of rheumatoid arthritis, Arthritis Rheumatol. 68 (1) (2016) 1–26, doi:10.1002/art.39480.
[10] B. Mondal, A.K. Manna, P. Talukdar, S.N.Talapatra I.Ghosh, Phytochemicals of avicennia species: prediction of toxicity through QSAR modeling and lead compound identification on TNF-α through molecular docking, RJLBPCS 5 (1) (2019) 853–868, doi:10.26479/2019.0501.71.
[11] Y.S. Sherif, M.A. Gouda, Quantitative structure structure-activity relationship (QSAR) for 2-amino-5-selenothiazole derivatives as anti-inflammatory and analgesic agents, Der Pharm. Chem. 6 (2) (2014) 1–6.
[12] A. Sepriano, A. Kerschbaumer, J.S. Smolen, D. Heijde, M. Dougados, I.B.McInnes R.Vollenhoven, J.W. Bijlsma, G.R. Burmester, M. Wit, L. Falzon, R. Landewé, Safety of synthetic and biological DMARDs: a systematic literature review informing the 2019 update of the EULAR recommendations for the management of rheumatoid arthritis, Ann. Rheum. Dis. 0 (2020) 1–11, doi:10.1136/annrheumdis-2019-216653.
[13] K. Zitouni, S. Belaidi, A. Kerassa, Conformational analysis and qsar modeling of 14-membered macrolide analogues against mycobacterium tuberculosis, J. Fundam. Appl. Sci. 12 (3) (2020) 1035–1066.
[14] E.I. Edache, A. Uzairu, P.A. Mamza, G.A. Shallangwa, A comparative QSAR analysis, 3D-QSAR, molecular docking and molecular design of iminoguanidine-based inhibitors of HemO: a rational approach to antibacterial drug design, J. Drugs Pharm. Sci. 4 (3) (2020) 21–36, doi:10.31248/JDPS2020.036.
[15] E.I. Edache, A. Uzairu, S.E. Abechi, Development and estimation of an in-silico model for Anti-HIV-1 integrase inhibitor using genetic function approximation, JAMPS 5 (2) (2016) 1–18, doi:10.9734/JAMPS/2016/22227.

[16] S. Shafiu, E.I. Edache, U. Sani, M. Abatyough, Docking and virtual screening studies of tetraketone derivatives as tyrosine kinase (EGFR) inhibitors: a rational approach to anti-fungi drug design, J. Pharm. Med. Res. 3 (1) (2017) 78–80.
[17] C.W. Yap, PaDEL-descriptor: an open-source software to calculate molecular descriptors and fingerprints, J. Comput. Chem. 32 (7) (2011) 1466–1474.
[18] E.I. Edache, D.E. Arthur, U. Abdulfatai, Quantitative structure-activity relationship analysis of the anti-tyrosine activity of some tetraketone and benzyl-benzoate derivatives based on genetic algorithm-multiple linear regression, JCMR 6 (2016) 3–13.
[19] A. Golbraikh, A. Tropsha, Beware of q^2!, J. Mol. Graph. Model. 20 (2002) 269–276, doi:10.1016/S1093-3263(01)00123-1.
[20] A. Golbraikh, M. Shen, Z. Xiao, Y.D. Xiao, K.H. Lee, A. Tropsha, Rational selection of training and test sets for the development of validated QSAR models, J. Comput. Aided Mol. Des. 17 (2003) 241–253.
[21] K. Roy, P.P. Roy, Comparative chemometric modeling of cytochrome 3A4 inhibitory activity of structurally diverse compounds using stepwise MLR, FA-MLR, PLS, GFA, G/PLS and ANN techniques, Eur. J. Med. Chem. 44 (2009) 2913–2922.
[22] E.I. Edache, A. Uzairu, S.E. Abechi, Investigation of 5,6-dihydro-2-pyrones derivatives as potent anti-HIV agents inhibitors, J. Comput. Methods Mol. Des. 5 (3) (2015) 135–149.
[23] S. Kar, K. Roy, Development and validation of a robust QSAR model for prediction of carcinogenicity of drugs, India J. Biochem. Biophys. 48 (2011) 111–122.
[24] E.I. Edache, H.U. Hambali, D.E. Arthur, A. Oluwaseye, O.C. Chinweuba, In-silico discovery and simulated selection of multi-target anti-HIV-1 inhibitors, IRJPAC 11 (1) (2016) 1–15, doi:10.9734/IRJPAC/2016/22863.
[25] Y.D. Aher, A. Agrawal, P.V. Bharatam, P. Garg, 3D-QSAR studies of substituted 1-(3, 3-diphenylpropyl)-piperidinyl amides and ureas as CCR5 receptor antagonists, J. Mol. Model. 13 (2007) 519–529, doi:10.1007/s00894-007-0173-z.
[26] P. Tosco, T. Balle, F. Shiri, Open3DALIGN: an open-source software aimed at unsupervised ligand alignment, J. Comput. Aided. Mol. Des. 25 (2011) 777–783, doi:10.1007/s10822-011-9462-9.
[27] P. Tosco, T. Balle, Open3DQSAR: a new open-source software aimed at high throughput chemometric analysis of molecular interaction fields, J. Mol. Model. 17 (2011) 201–208, doi:10.1007/s00894-010-0684-x.
[28] T.K.S. Ahamed, V.K. Rajan, K. Muraleedharan, QSAR modeling of benzoquinone derivatives as 5-lipoxygenase inhibitors, Food Sci. Hum. Wellness 8 (2019) 53–62, doi:10.1016/j.fshw.2019.02.001.
[29] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J. Comput. Chem. 31 (2) (2010) 455–461, doi:10.1002/jcc.21334.
[30] M.F. Sanner, Python: a programming language for software integration and development, J. Mol. Graph. Model. 17 (1999) 57–61.
[31] G. Morris, R. Huey, AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility, J. Comput. Chem. 30 (2009) 2785–2791, doi:10.1002/jcc.21256.AutoDock4.
[32] W. Humphrey, A. Dalke, K. Schulten, Visual molecular dynamics, J. Mol. Graph. 7855 (1996) 33–38.
[33] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, K. Schulten, Scalable molecular dynamics with NAMD, J. Comput. Chem. 26 (2005) 1781–1802, doi:10.1002/jcc.20289.
[34] A.D. MacKerell, D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins, J. Phys. Chem. B 102 (18) (1998) 3586–3616, doi:10.1021/jp973084f.
[35] M. Abbasi, F. Ramezani, M. Elyasi, H. Sadeghi-Aliabadi, M. Amanlou, A study on quantitative structure-activity relationship and molecular docking of metalloproteinase inhibitors based on L-tyrosine scaffold, DARU J. Pharm. Sci. 23 (2015) 29, doi:10.1186/s40199-015-0111-z.
[36] E. Qudjani, M. Iman, M.F.D. Ramandi, A. Shafiee, Design and synthesis of curcumin-like diarylheptanoid analogues as potential anticancer agents, Recent Pat. Anticancer Drug Discov. 11 (2016) 342–351, doi:10.2174/1574892811666160420141.
[37] M. Iman, Z. Khansefid, A. Davood, Modeling and proposed molecular mechanism of hydroxyurea through docking and molecular dynamic simulation to curtail the action of ribonucleotide reductase, Recent Pat. Anticancer Drug Discov. 11 (4) (2016) 461–468.
[38] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, Adv. Drug Deliv. Rev. 23 (2001) 3–25.
[39] D.F. Veber, S.R. Johnson, H. Cheng, B.R. Smith, K.W. Ward, K.D. Kopple, Molecular properties that influence the oral bioavailability of drug candidates, J. Med. Chem. 45 (2002) 2615–2623, doi:10.1021/jm020017n.
[40] T.W. Johnson, K.R. Dress, M. Edwards, Using the golden triangle to optimize clearance and oral absorption, Bioorganic Med. Chem. Lett. 19 (2009) 5560–5564, doi:10.1016/j.bmcl.2009.08.045.