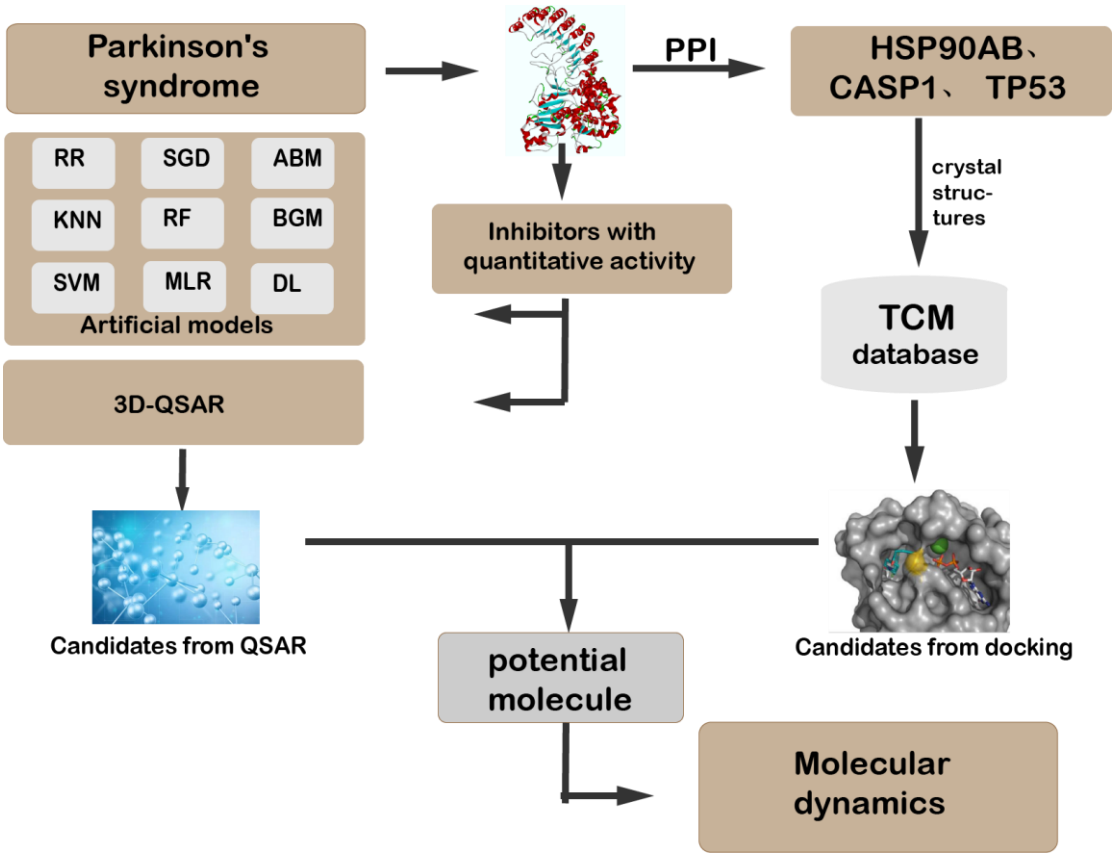


Figure 1. Experimental design flow chart

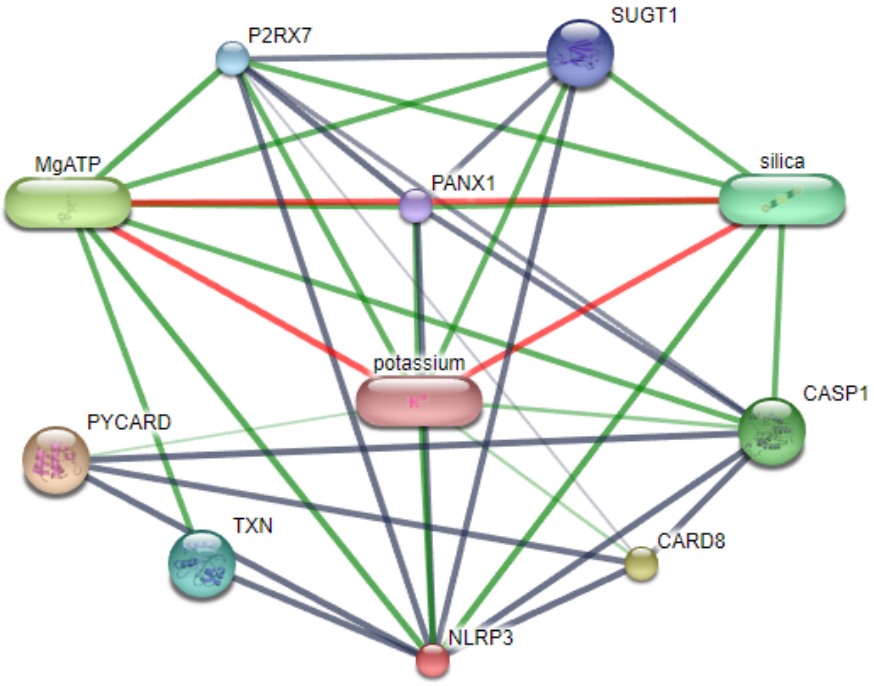


Figure 2. protein protein interaction Schematic diagram

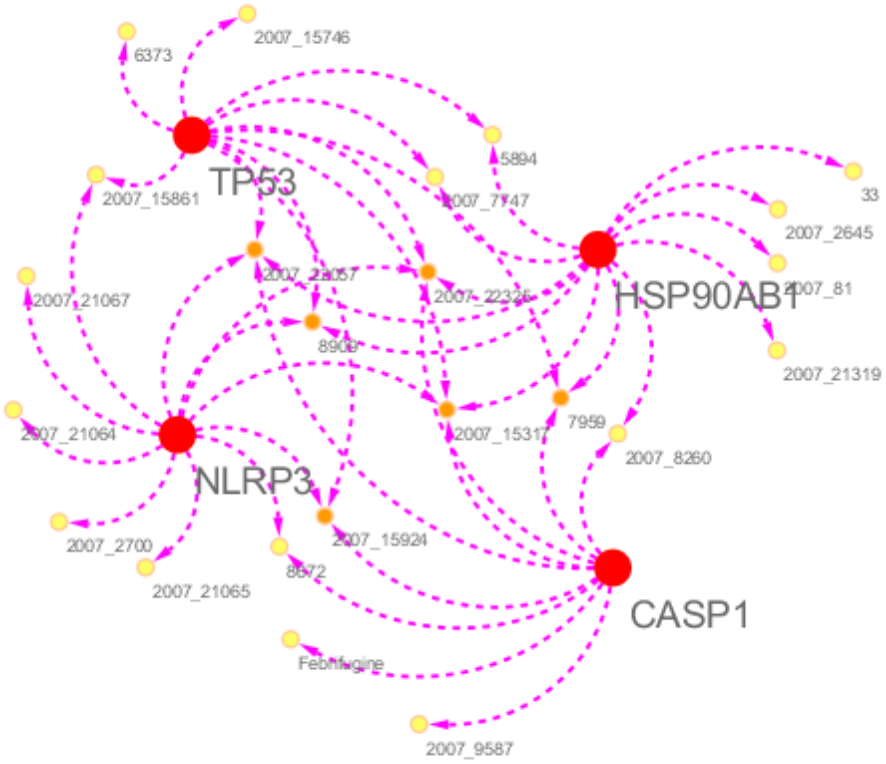
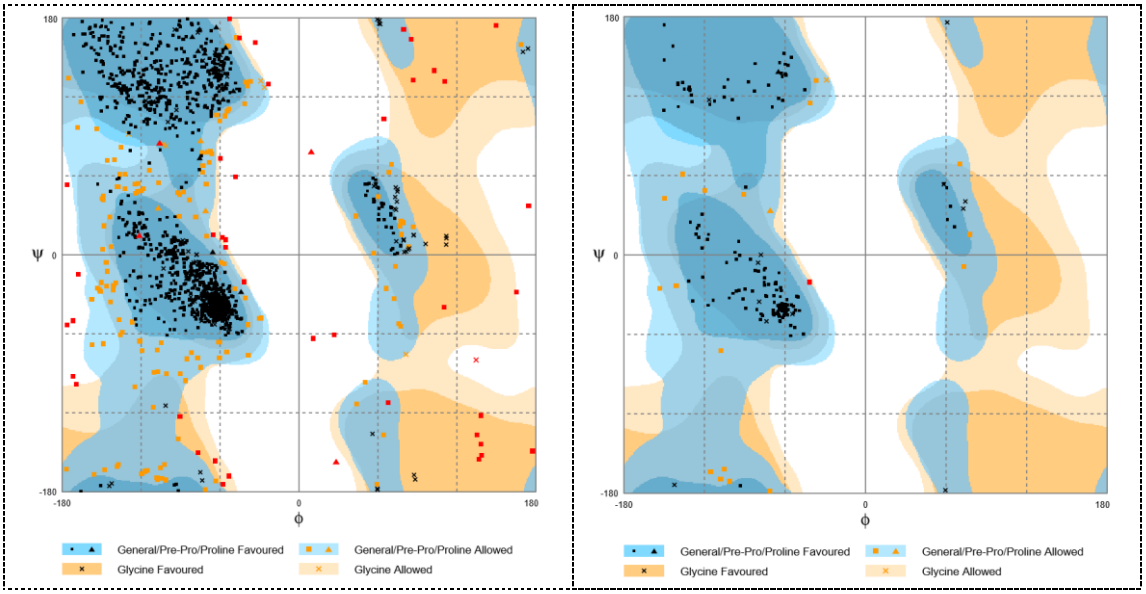


Figure 3. Compound-target network generated by top 10 TCM candidates (ranked by Dock Score) of each protein from database screening.



(a) **(b)**

Figure 4. (a)Ramachandran plot validation of our modeling structure. It can help us to find weather the structure is reliable (b) All the plots mean the amino acids of binding sites, in order to find if these amino acids which is important to our study stood in best or allowed fold region.

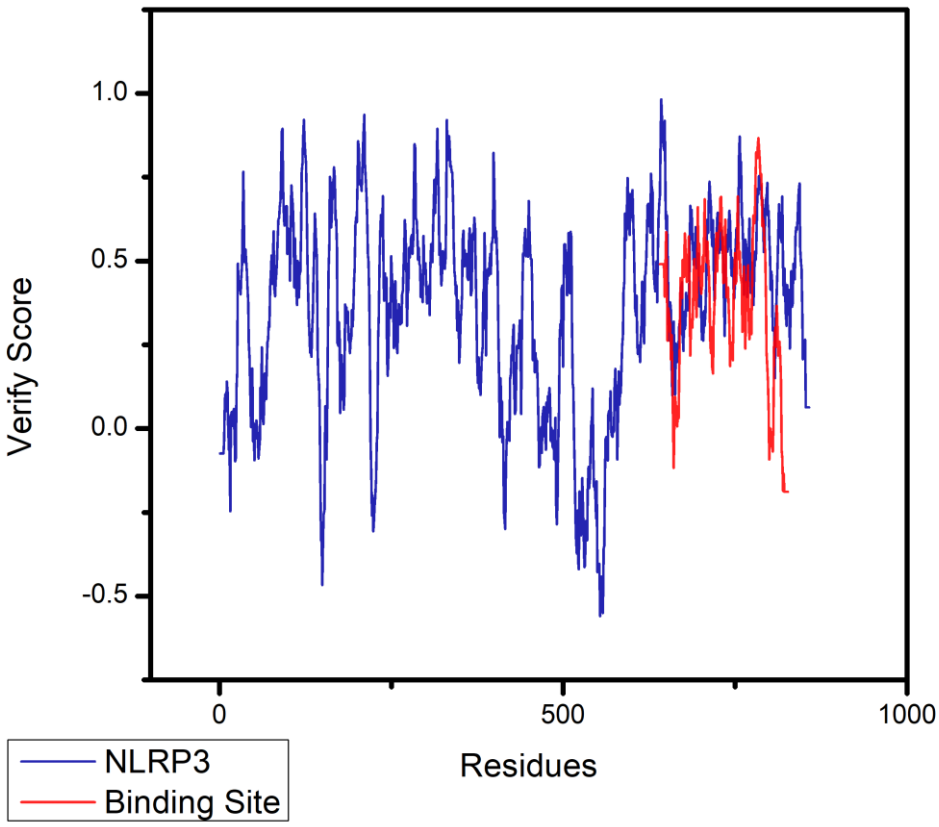


Figure 5. The 3D-profilevalidation of modeling structure. The Verify Score which is higher than 0 signifies the trusted simulation of amino acids.

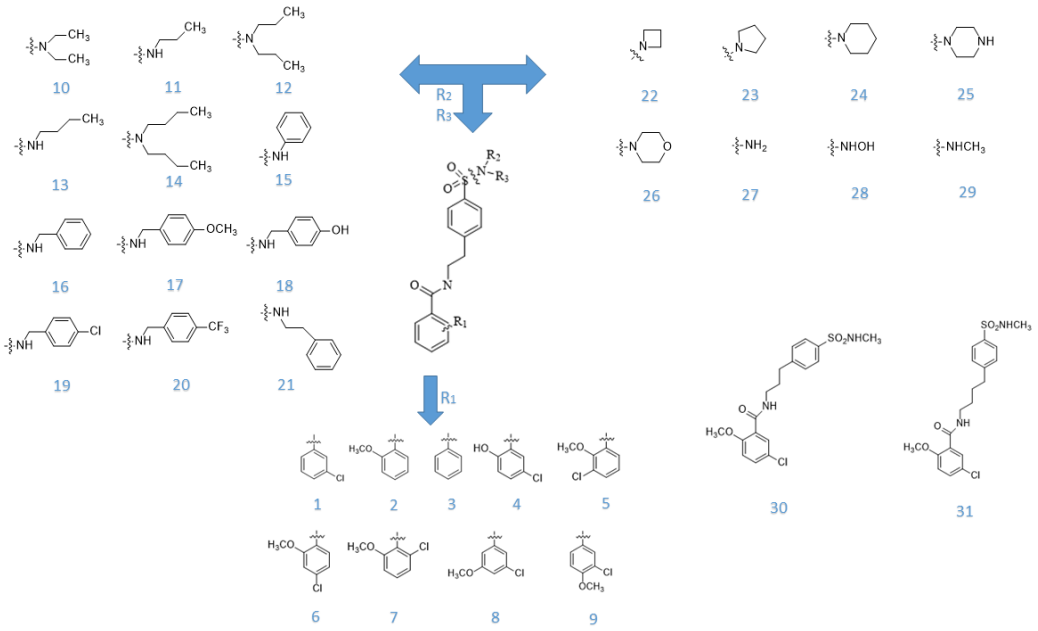


Figure 6. The known small molecular structure of NLRP3 inhibitors from previou study

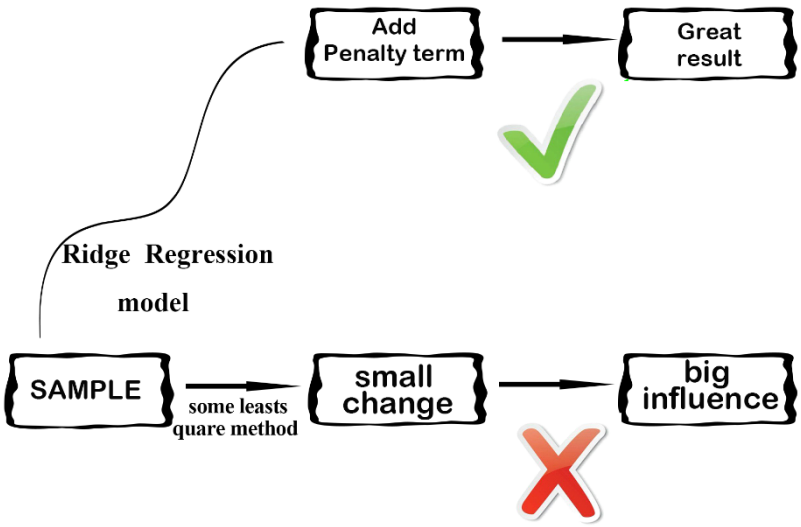


Figure 7. A simplified diagram of the Ridge regression model

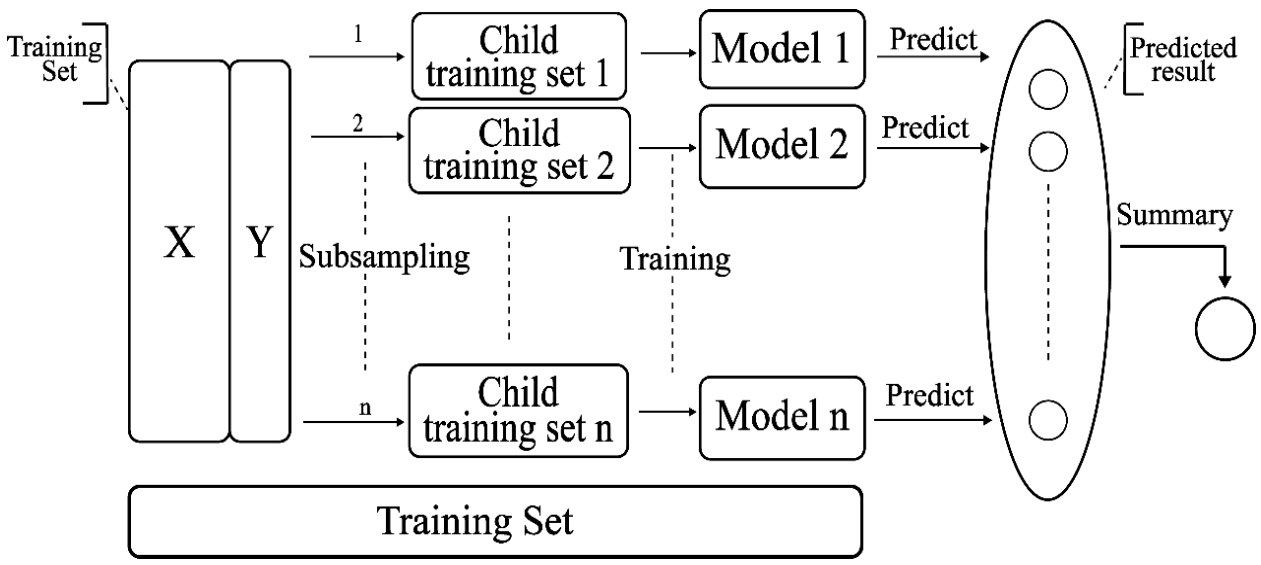


Figure 8. A simplified diagram of the Bagging model

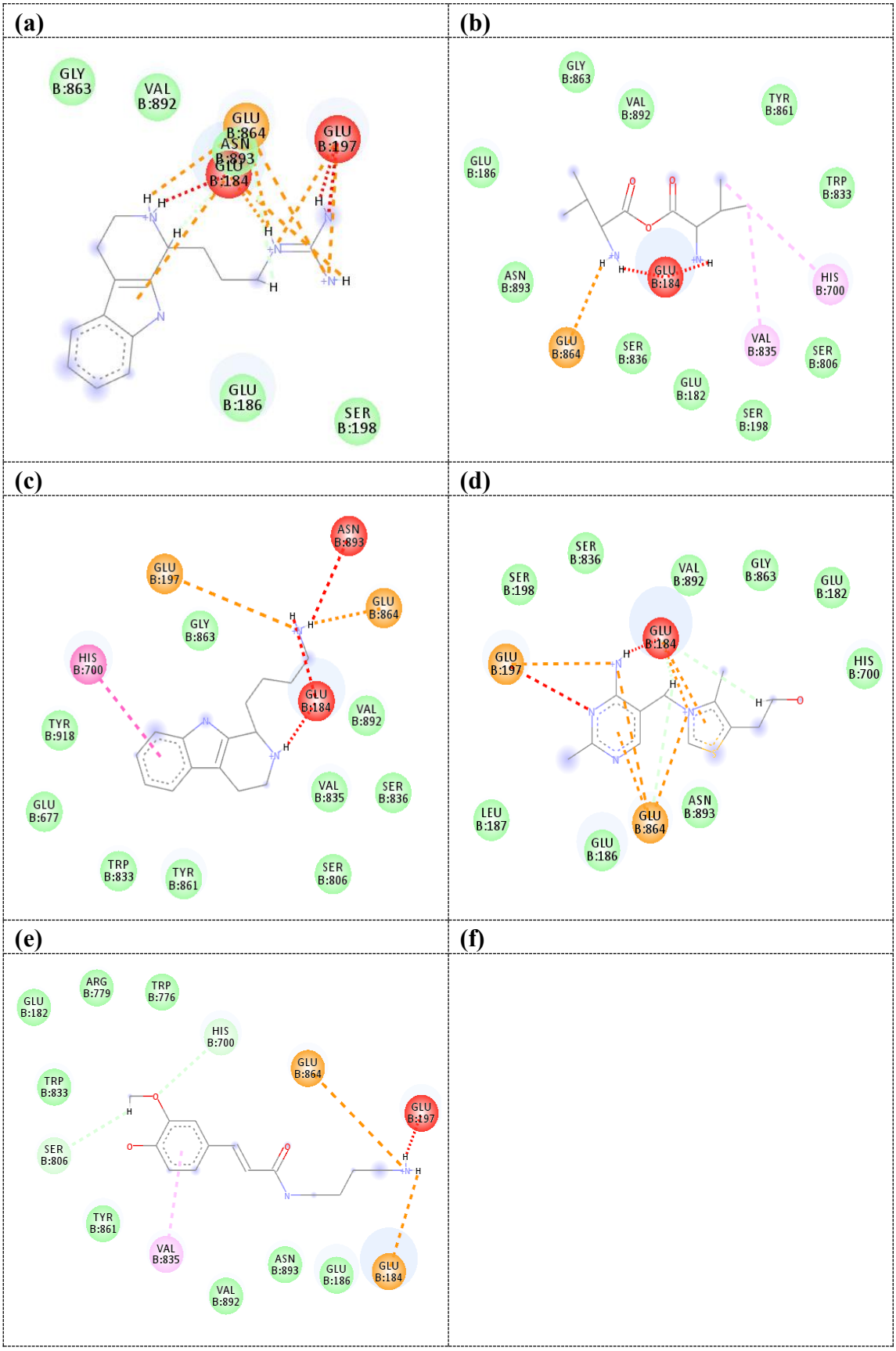
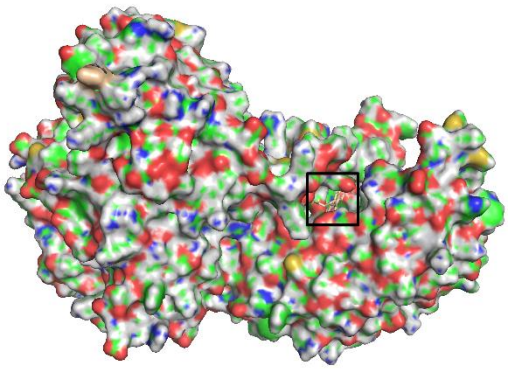
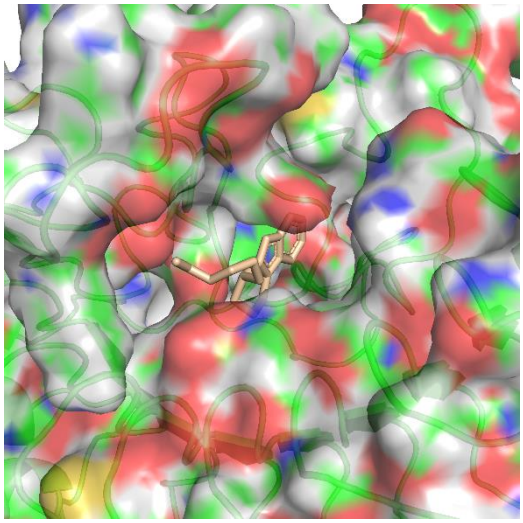

(a)
GLY B:863
VAL B:892
GLU B:864
GLU B:197
ASN B:893
GLU B:184
GLU B:186
SER B:198
(b)
GLY B:863
VAL B:892
TYR B:861
GLU B:186
TRP B:833
ASN B:893
GLU B:184
HIS B:700
GLU B:864
SER B:836
VAL B:835
SER B:806
GLU B:182
SER B:198
(c)
ASN B:893
GLU B:197
GLU B:864
GLY B:863
HIS B:700
GLU B:184
VAL B:892
TYR B:918
VAL B:835
SER B:836
GLU B:677
TRP B:833
TYR B:861
SER B:806
(d)
SER B:836
SER B:198
VAL B:892
GLY B:863
GLU B:182
GLU B:184
GLU B:197
HIS B:700
LEU B:187
ASN B:893
GLU B:864
GLU B:186
(e)
ARG B:779
TRP B:776
GLU B:182
HIS B:700
GLU B:864
TRP B:833
GLU B:197
SER B:806
TYR B:861
VAL B:835
ASN B:893
GLU B:186
GLU B:184
VAL B:892
(f)

Figure 9. 2D diagram of combined pattern in five targets complexes. (a)2007_22057,(b) 2007_22325,(c) 2007_15317 ,(d) 8909 and(e) 7959

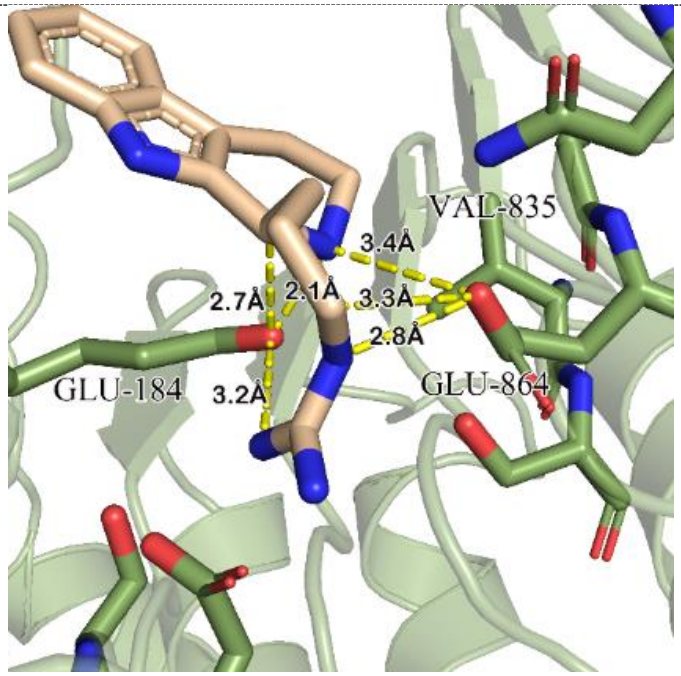
(a1)



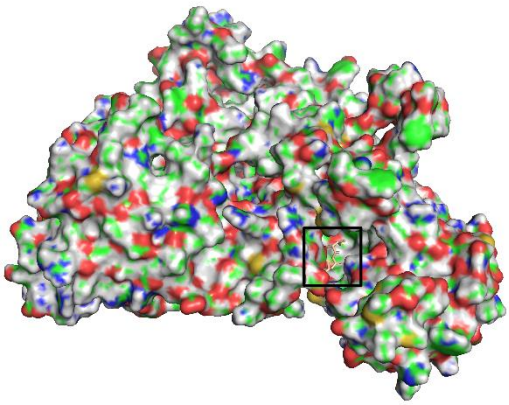
(a2)



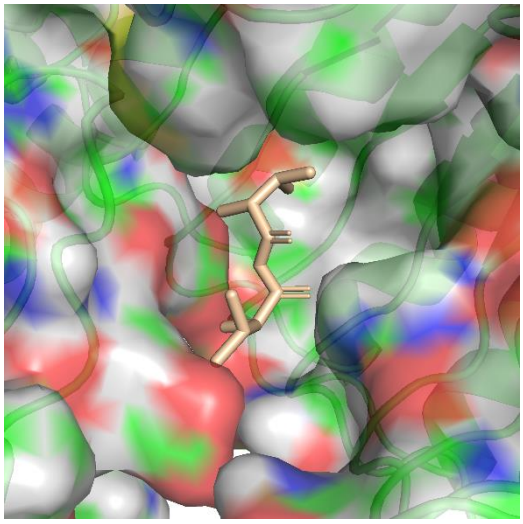
(a3)



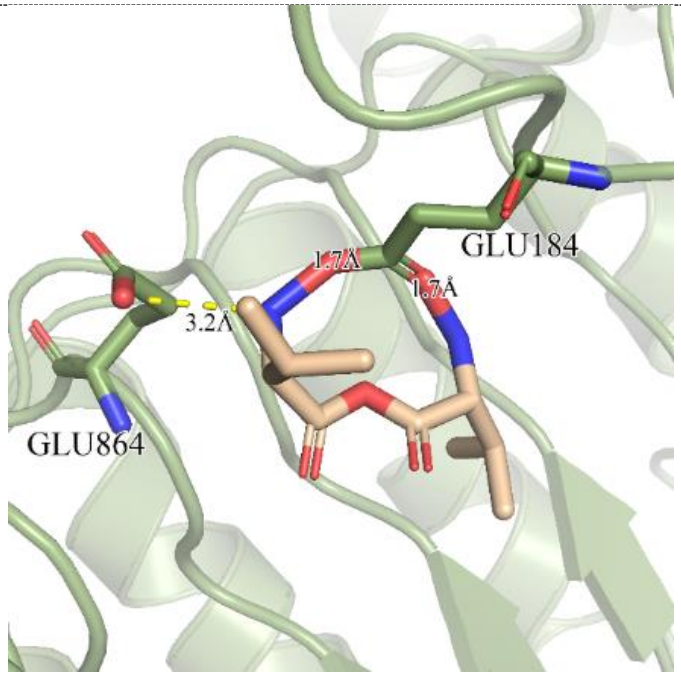
(b1)



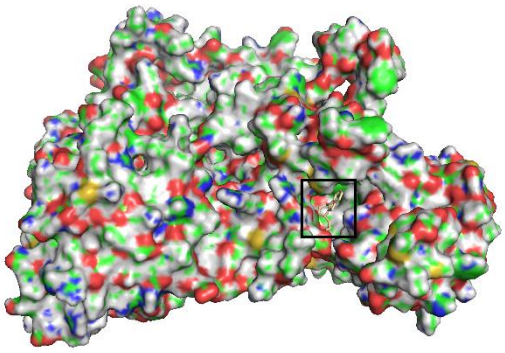
(b2)



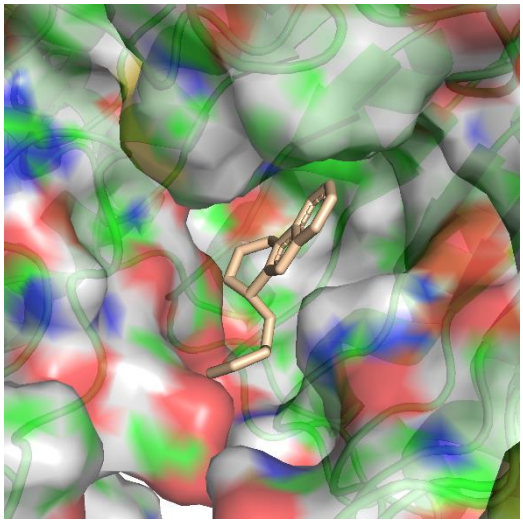
(b3)



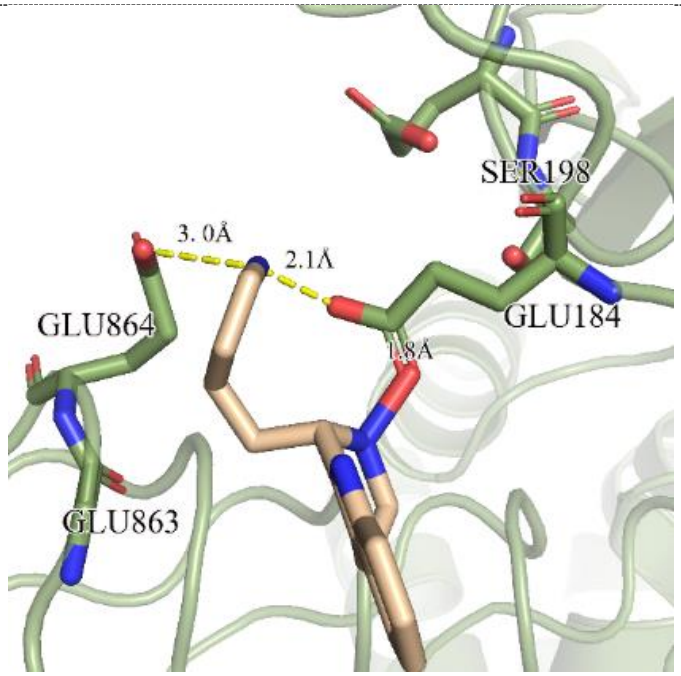
(c1)



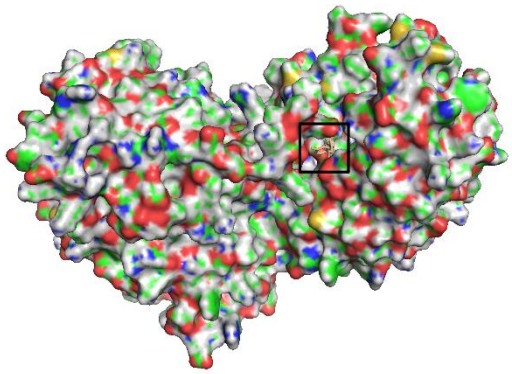
(c2)



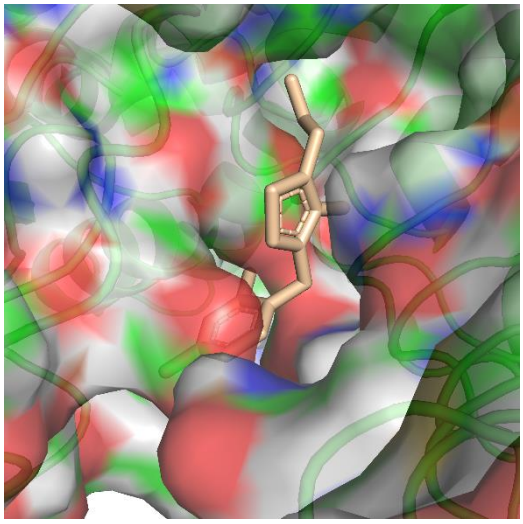
(c3)



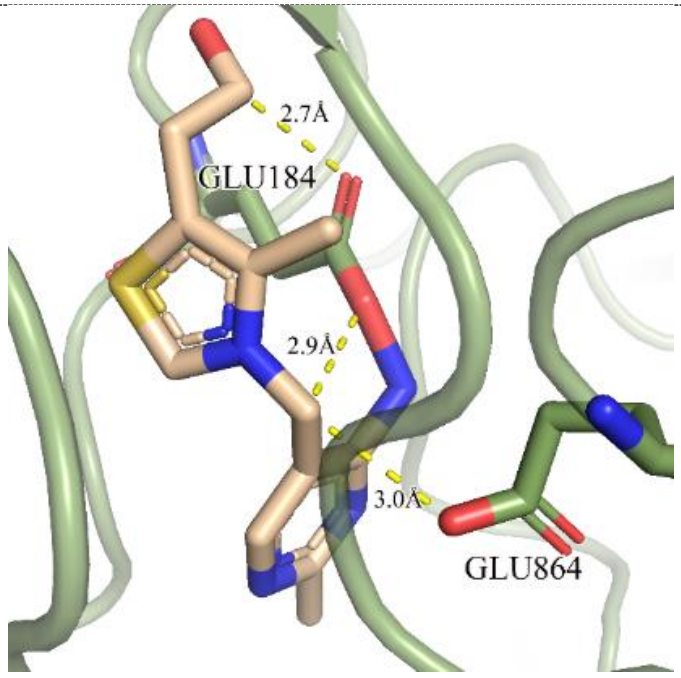
(d1)



(d2)



(d3)



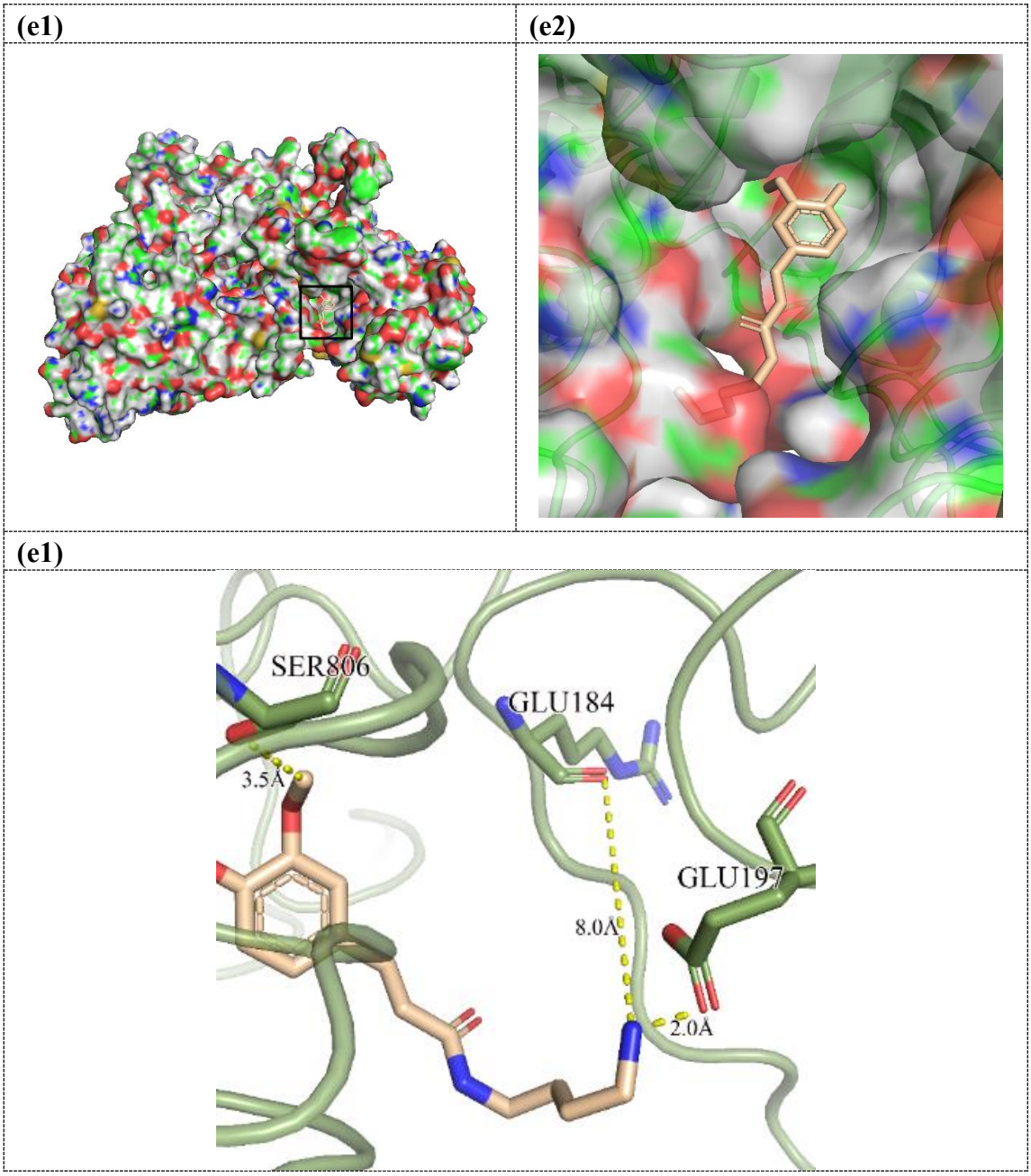


Figure 10. Docking pose of (a)2007_22057, (b)2007_22325, (c)2007_15317, (d) 8909 and (e)7959 with NLRP3. In (3), the yellow dash lines stand for H-bonds.

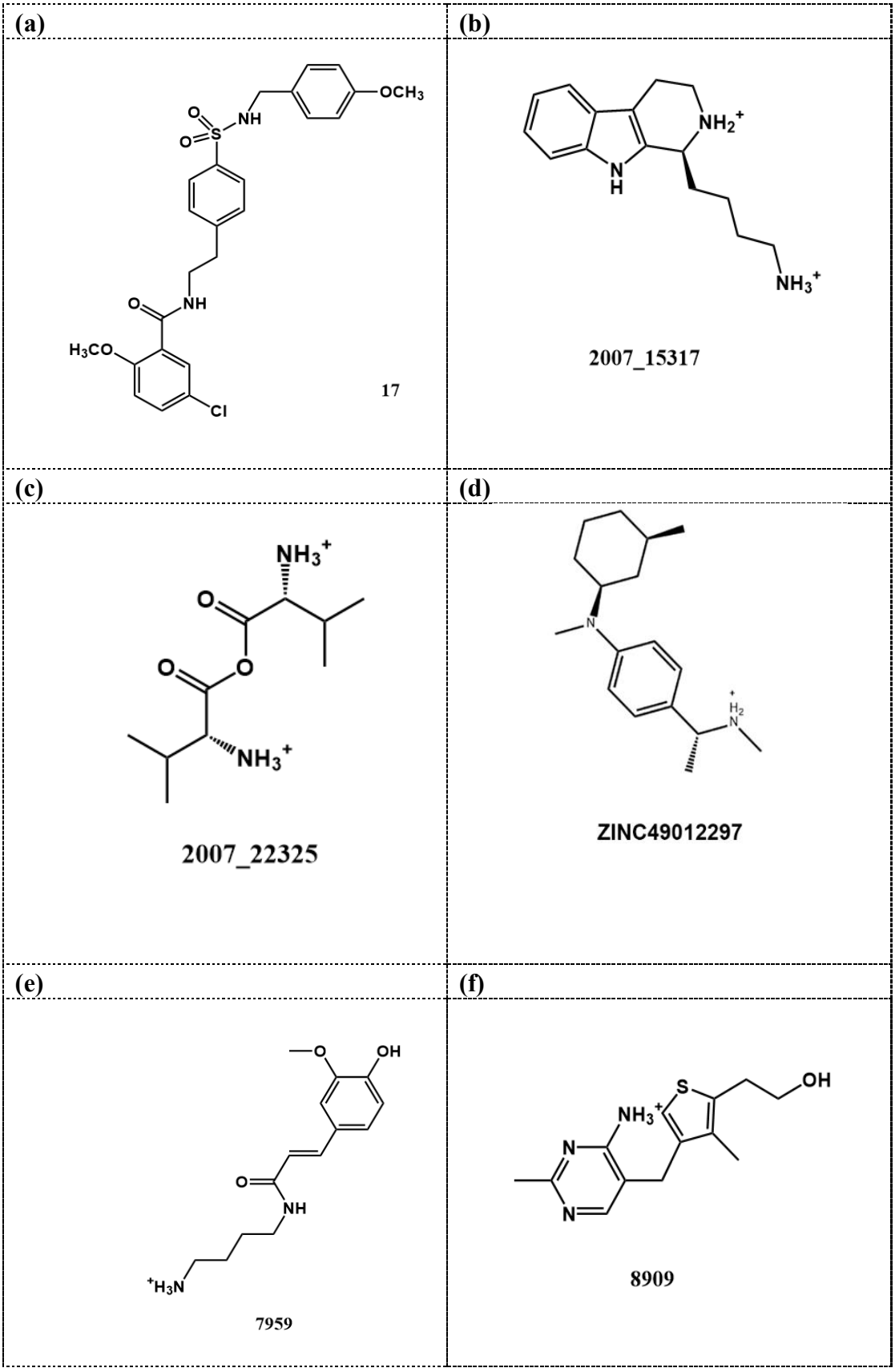


Figure 11. The chemical scaffold of TCM candidates and controls.

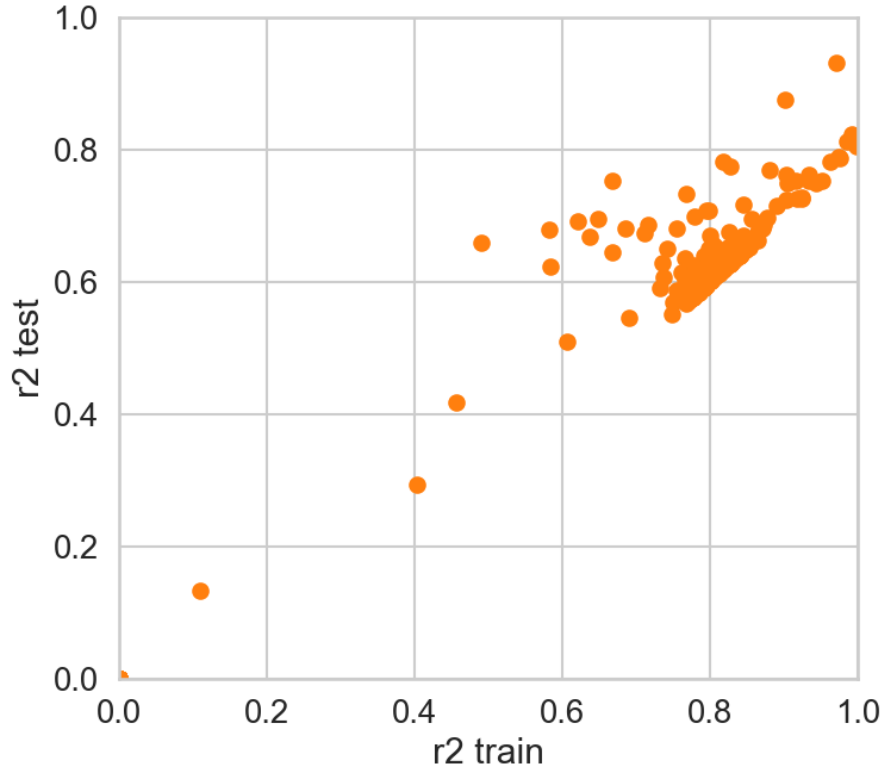
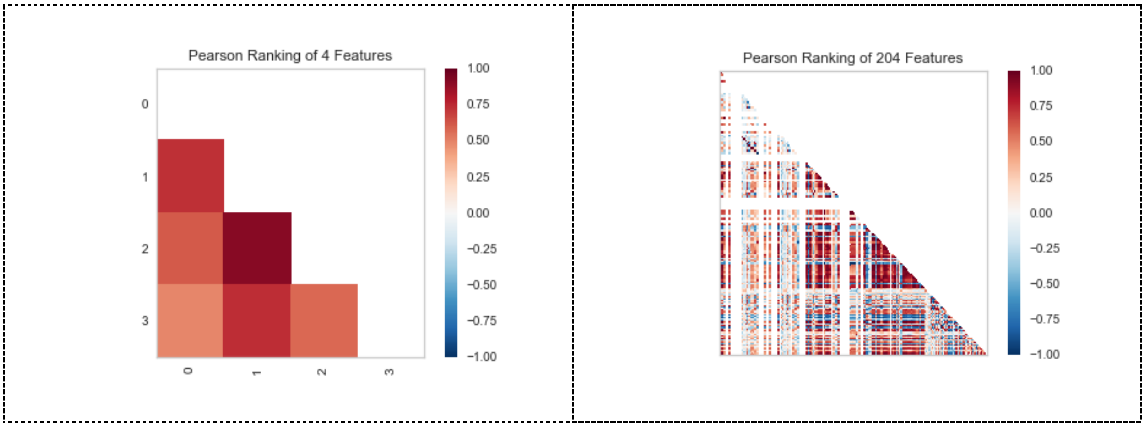


Figure 12. Scatter plot presenting the results of experiments.



(a) **(b)**

Figure 13. (a) Pearson correlation coefficient matrix heat map of three selected features.(b) Relation of 204 features ranked by Pearson correlation coefficient.

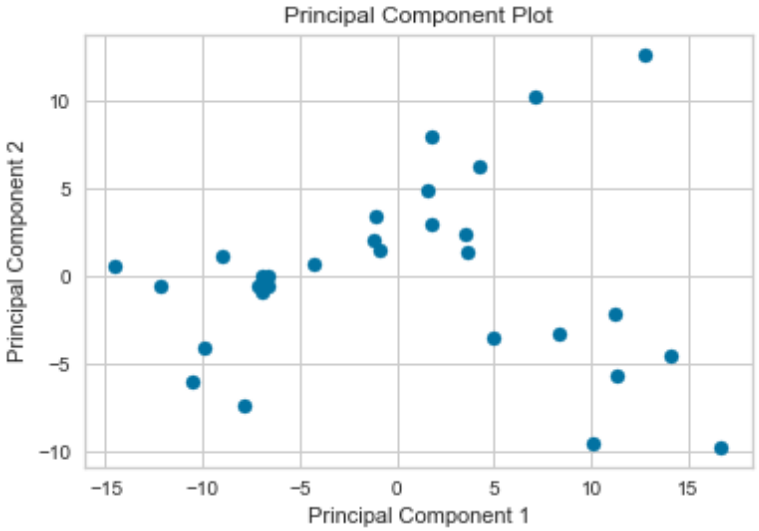


Figure 14. 2D PCA visualizations obtained using Yellowbrick.

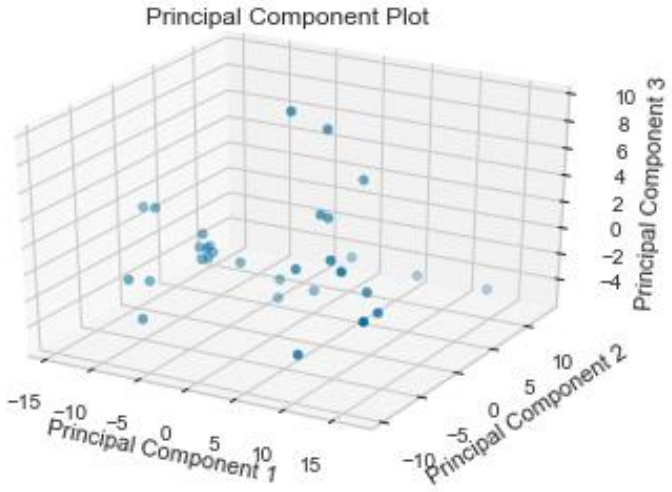
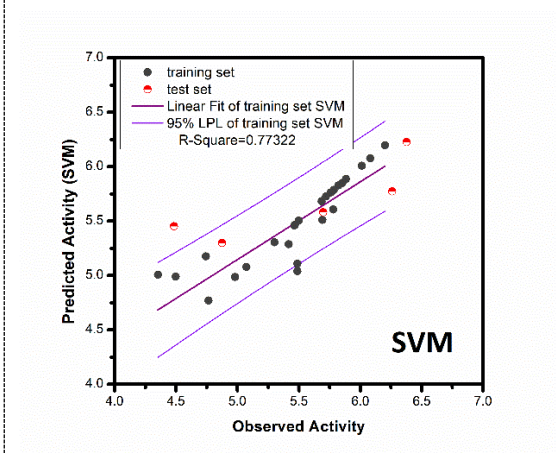
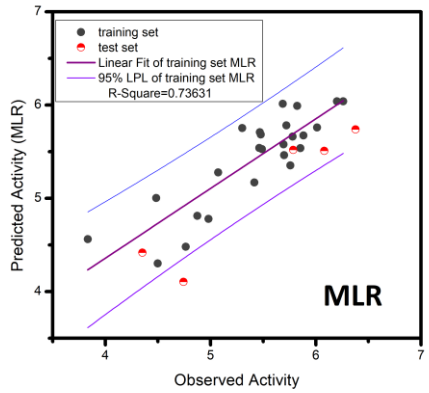


Figure 15. 3D PCA visualizations obtained using Yellowbrick.

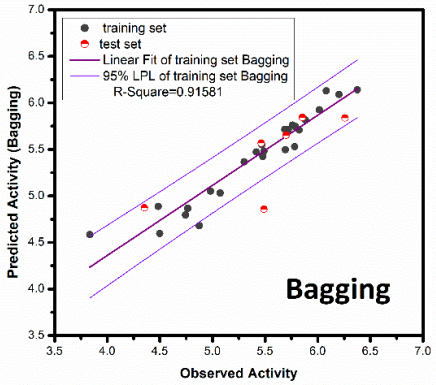
(a)


training set
test set
Linear Fit of training set SVM
95% LPL of training set SVM
R-Square=0.77322
SVM
Predicted Activity (SVM)
Observed Activity

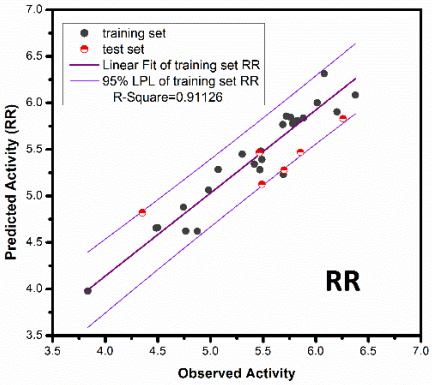
(b)


training set
test set
Linear Fit of training set MLR
95% LPL of training set MLR
R-Square=0.73631
MLR
Predicted Activity (MLR)
Observed Activity

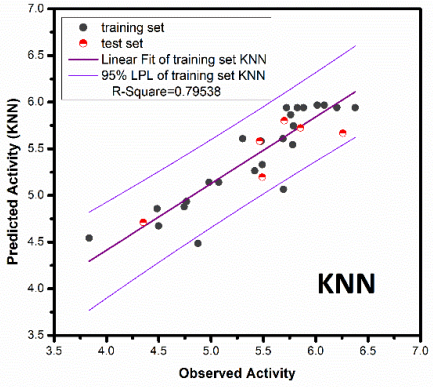
(c)


training set
test set
Linear Fit of training set Bagging
95% LPL of training set Bagging
R-Square=0.91581
Bagging
Predicted Activity (Bagging)
Observed Activity

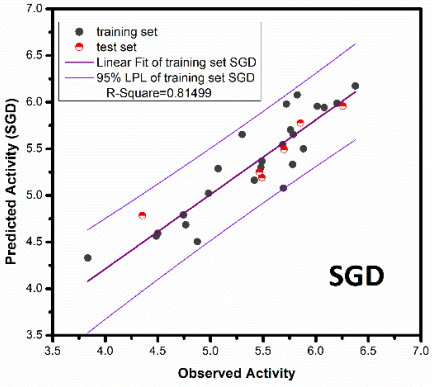
(d)


training set
test set
Linear Fit of training set RR
95% LPL of training set RR
R-Square=0.91126
RR
Predicted Activity (RR)
Observed Activity

(e)


training set
test set
Linear Fit of training set KNN
95% LPL of training set KNN
R-Square=0.79538
KNN
Predicted Activity (KNN)
Observed Activity

(f)


training set
test set
Linear Fit of training set SGD
95% LPL of training set SGD
R-Square=0.81499
SGD
Predicted Activity (SGD)
Observed Activity

(g)

(h)

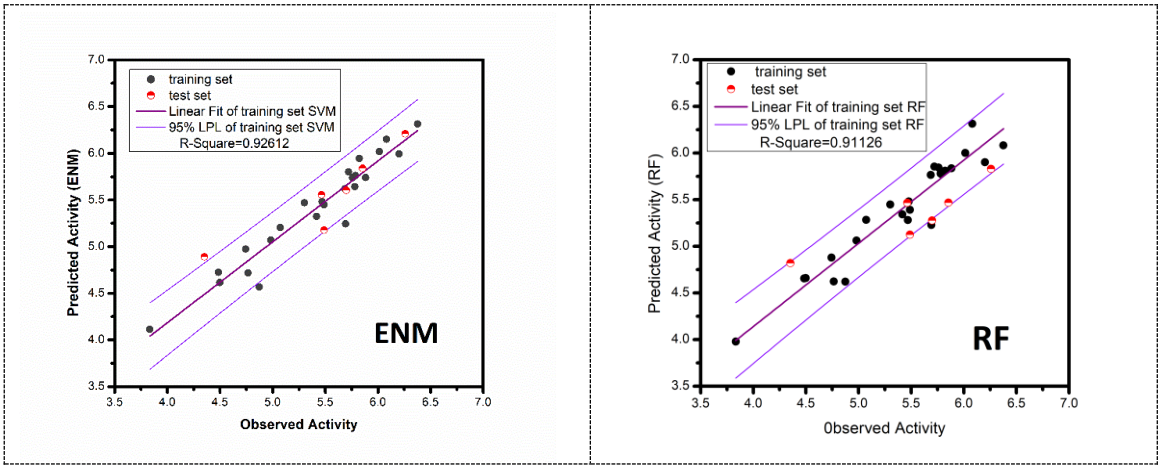
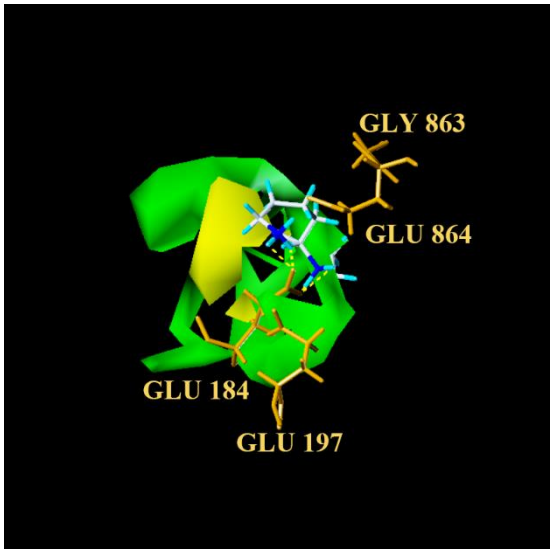
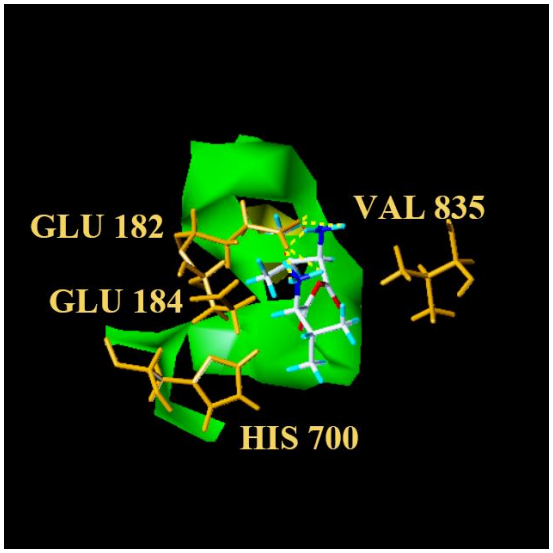
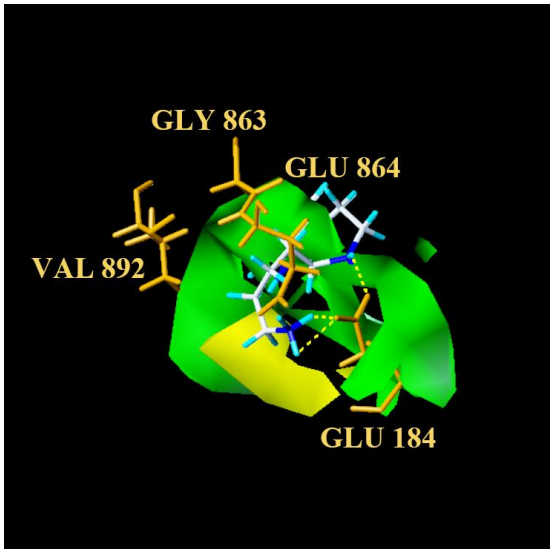
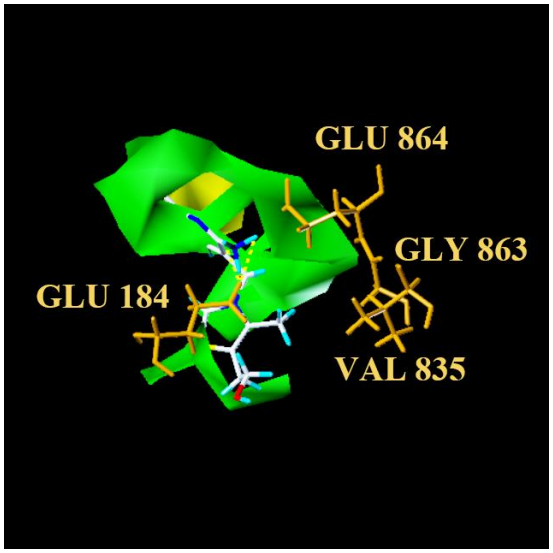


Figure 16. AI models for NLRP3 respectively. The SVM and MLR models identify relationships between observed and predicted activity (pIC_{50}). Correlation trend (purple lines) and 95% prediction confidence regions (enclosed by red lines) were presented. Training set (black dots) and testing set (red dots) were shown. Correlation coefficients (R^2) of the QSAR models were all higher than 0.70.

(a)	(b)
 GLY 863 GLU 864 GLU 184 GLU 197	 GLU 182 VAL 835 GLU 184 HIS 700
(c)	**(d)**
 GLY 863 GLU 864 VAL 892 GLU 184	 GLU 864 GLY 863 GLU 184 VAL 835
(e)	**(f)**

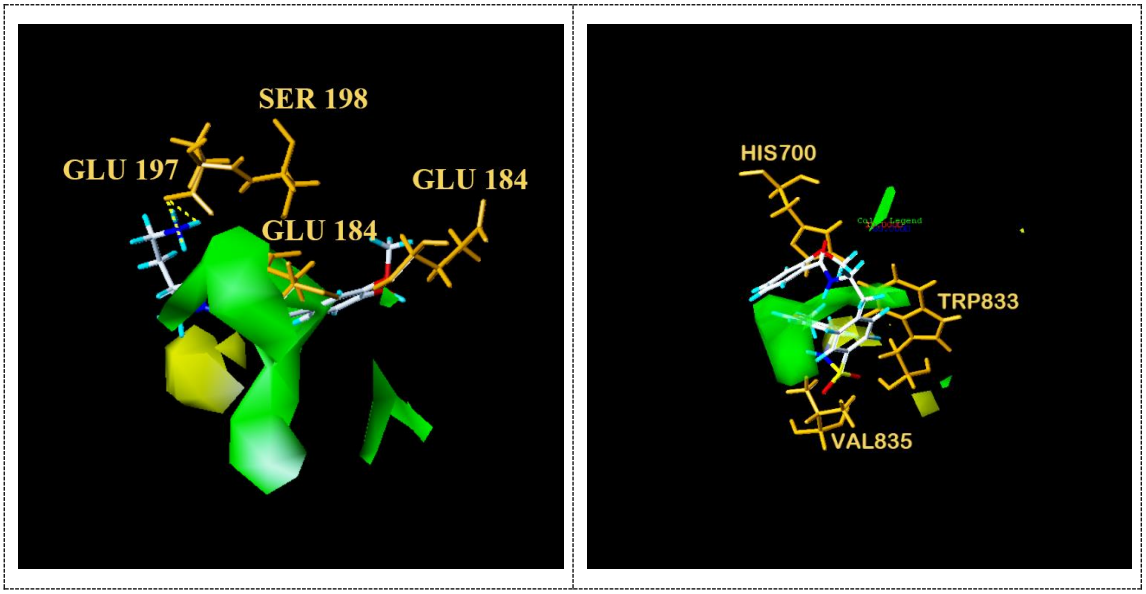
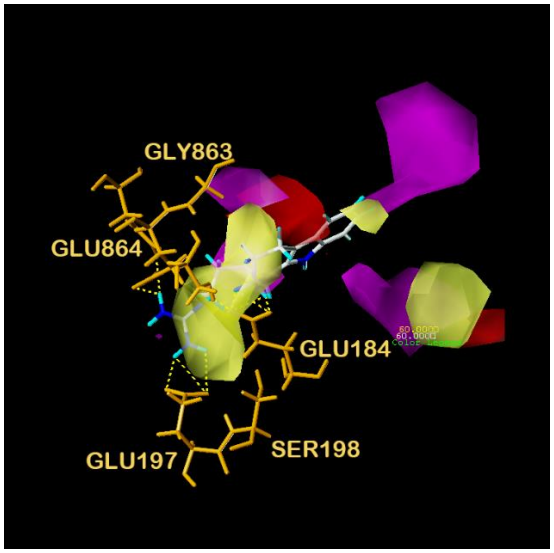
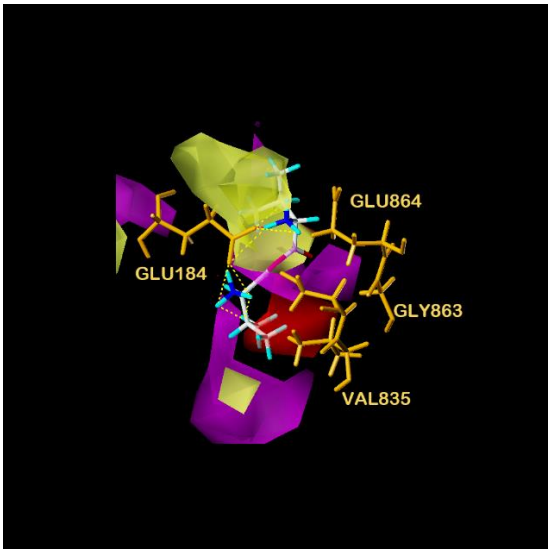
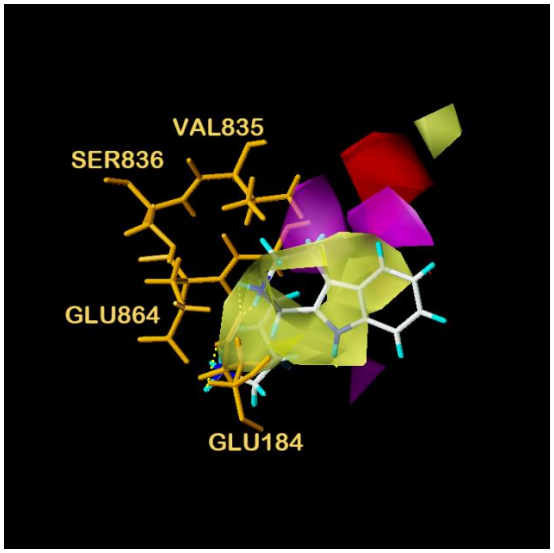
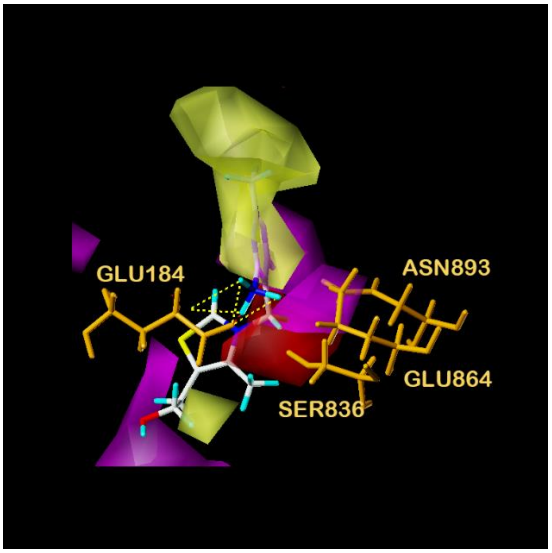


Figure 17. Structural contouring of TCM candidates and NLRP3 control to CoMFA mapping. (a)2007_22057, (b) 2007_22325, (c) 2007_15317, (d) 8909, (e)7959, (f) 17. The yellow dotted line indicates a hydrogen bond

(a)	(b)
 GLY863 GLU864 GLU184 GLU197 SER198	 GLU864 GLU184 GLY863 VAL835
(c)	**(d)**
 VAL835 SER836 GLU864 GLU184	 GLU184 ASN893 GLU864 SER836
(e)	**(f)**

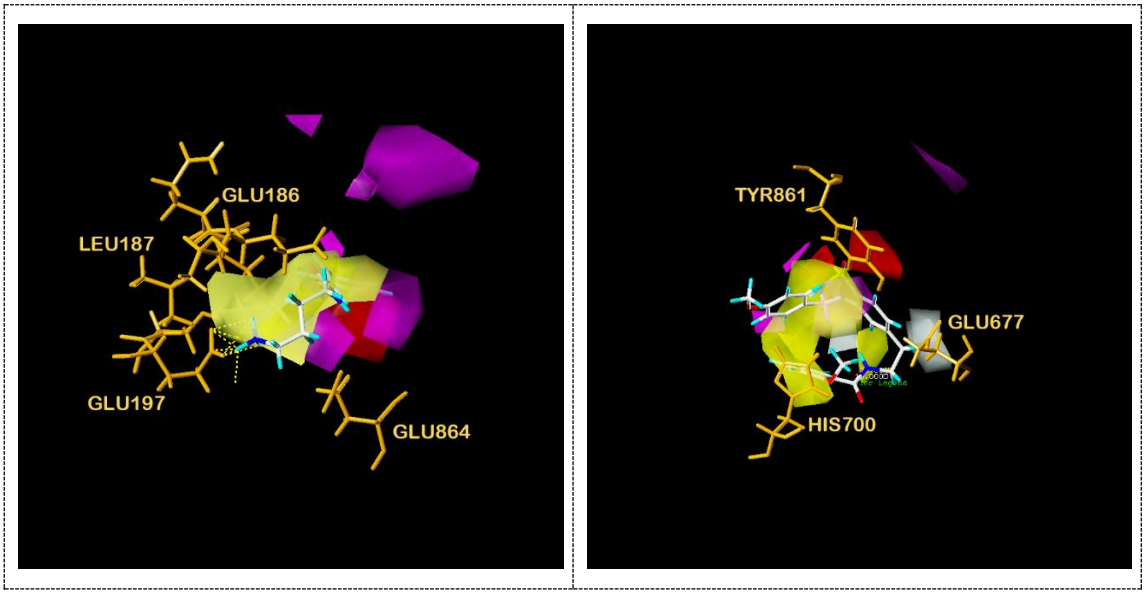


Figure 18. Structural contouring of TCM candidates and NLRP3 control to CoMSIA mapping. (a)2007_22057, (b) 2007_22325, (c) 2007_15317, (d) 8909, (e)7959, (f) 17. The yellow dotted line indicates a hydrogen bond.

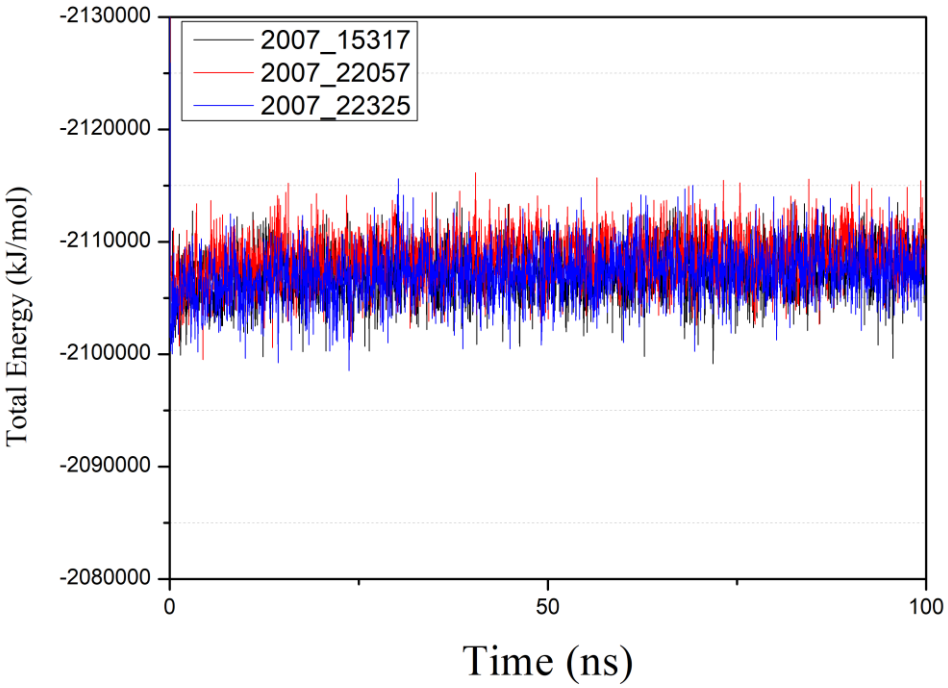


Figure 19. Total energy changes during molecular dynamics simulations between NLRP3 protein with three candidates. Different colors represent different molecular candidates, which could demonstratively reveal the state of complex.

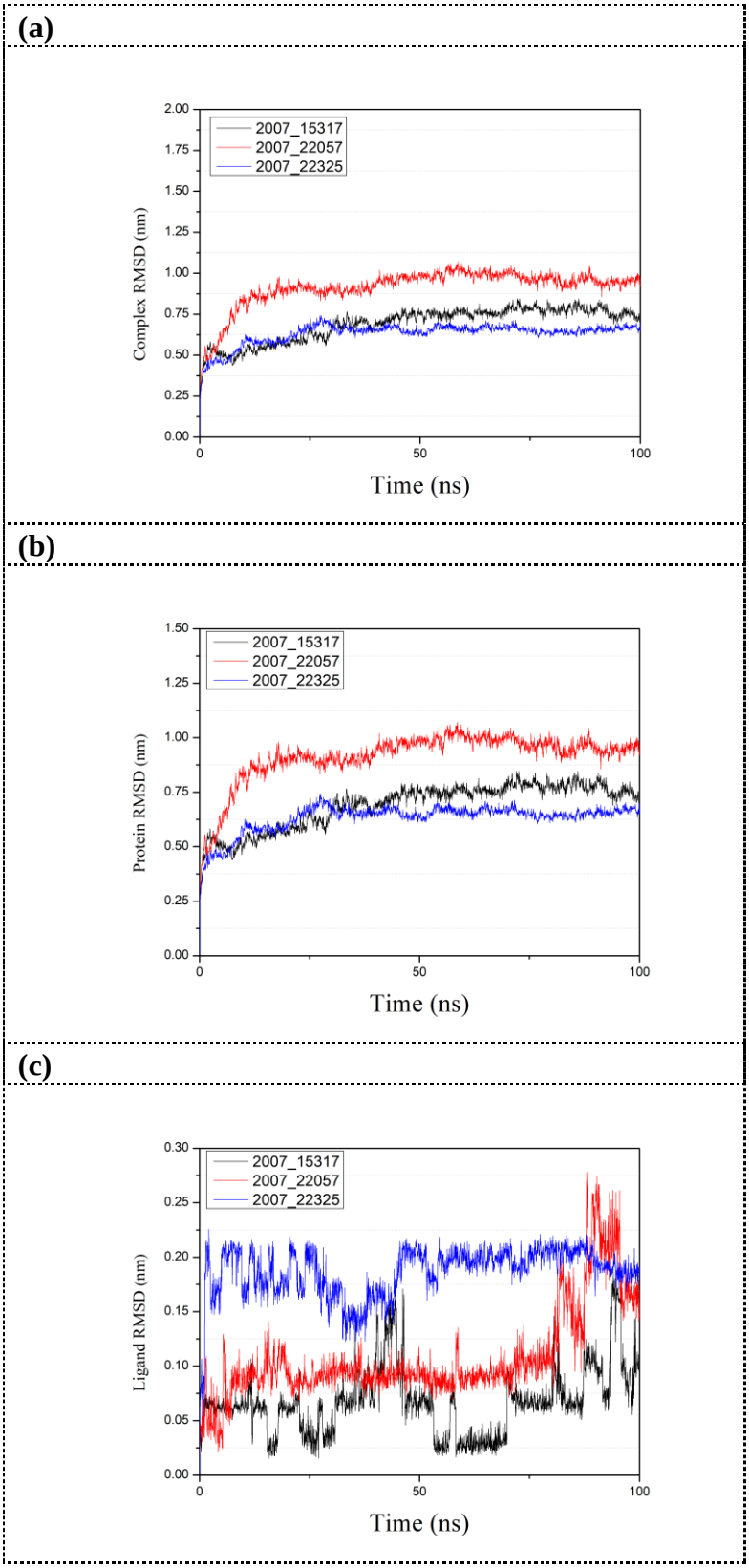


Figure 20. RMSD changes during molecular dynamics simulations between NLRP3 protein with three candidates. Different colors represent different molecular candidates, which could demonstratively reveal the state of complex. (a) RMSD changes of complex (b) RMSD changes of protein (c) RMSD changes of ligand

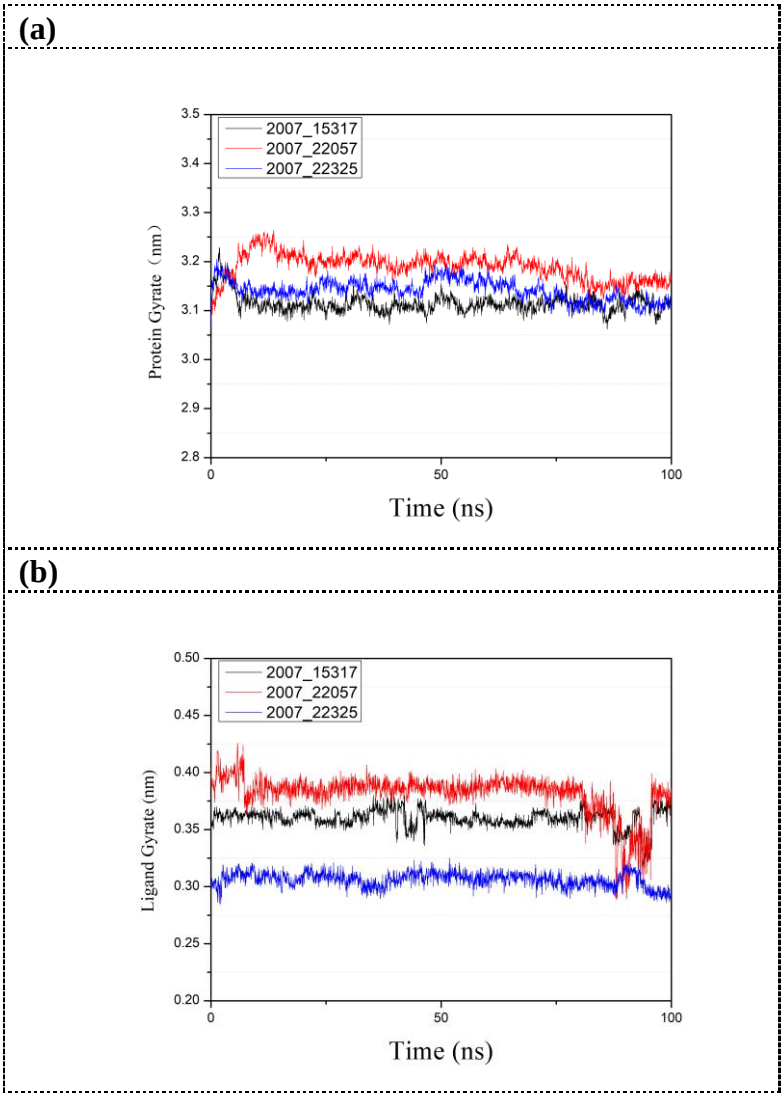


Figure 21. Gyrate result of target complex with three candidates. Different colors represent different molecular candidates, which could demonstratively reveal the state of complex. (a) Gyrate result of protein (b) Gyrate result of ligand

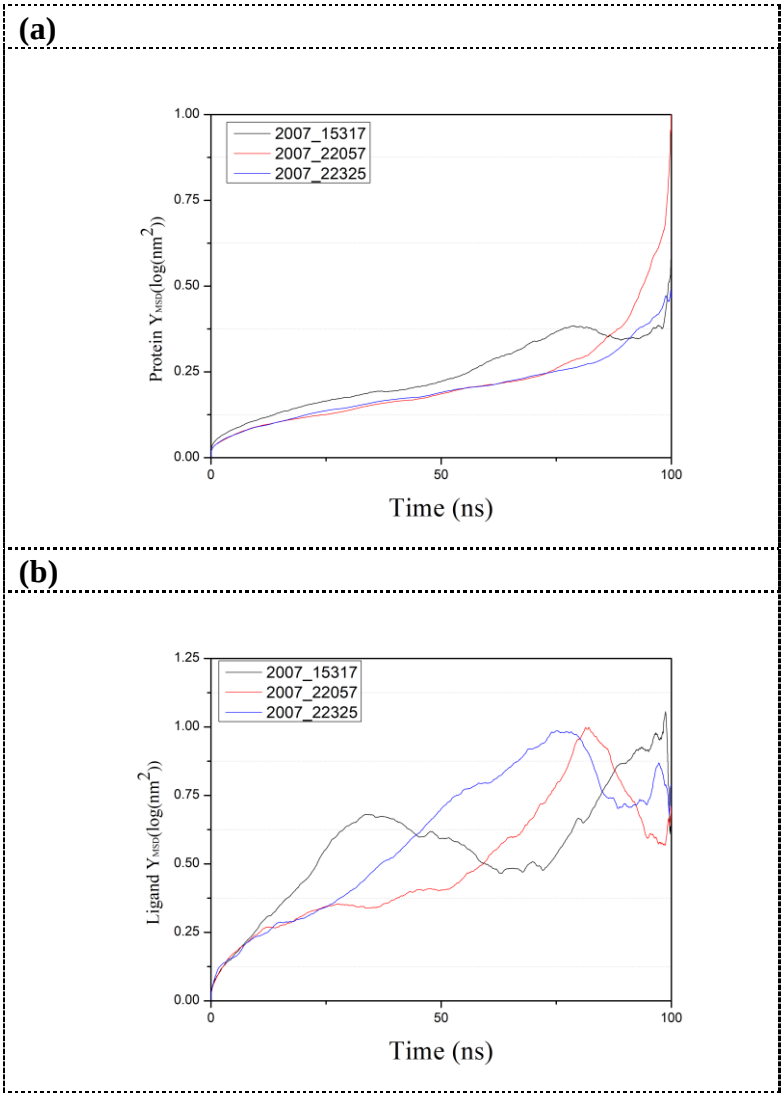


Figure 22. MSD result of target complex with three candidates. Different colors represent different molecular candidates, which could demonstratively reveal the state of complex. (a) MSD result of protein (b) MSD result of ligand

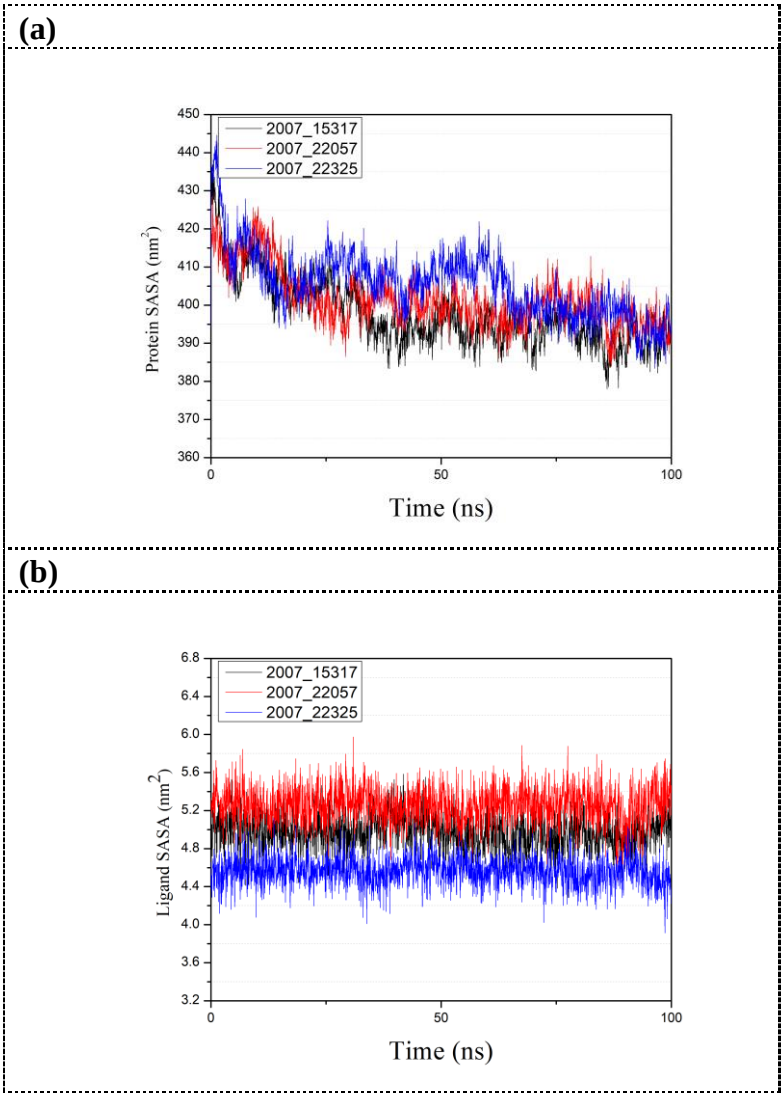


Figure 23. SASA result of target complex with three candidates. Different colors represent different molecular candidates, which could demonstratively reveal the state of complex. (a) SASA result of protein (b) SASA result of ligand

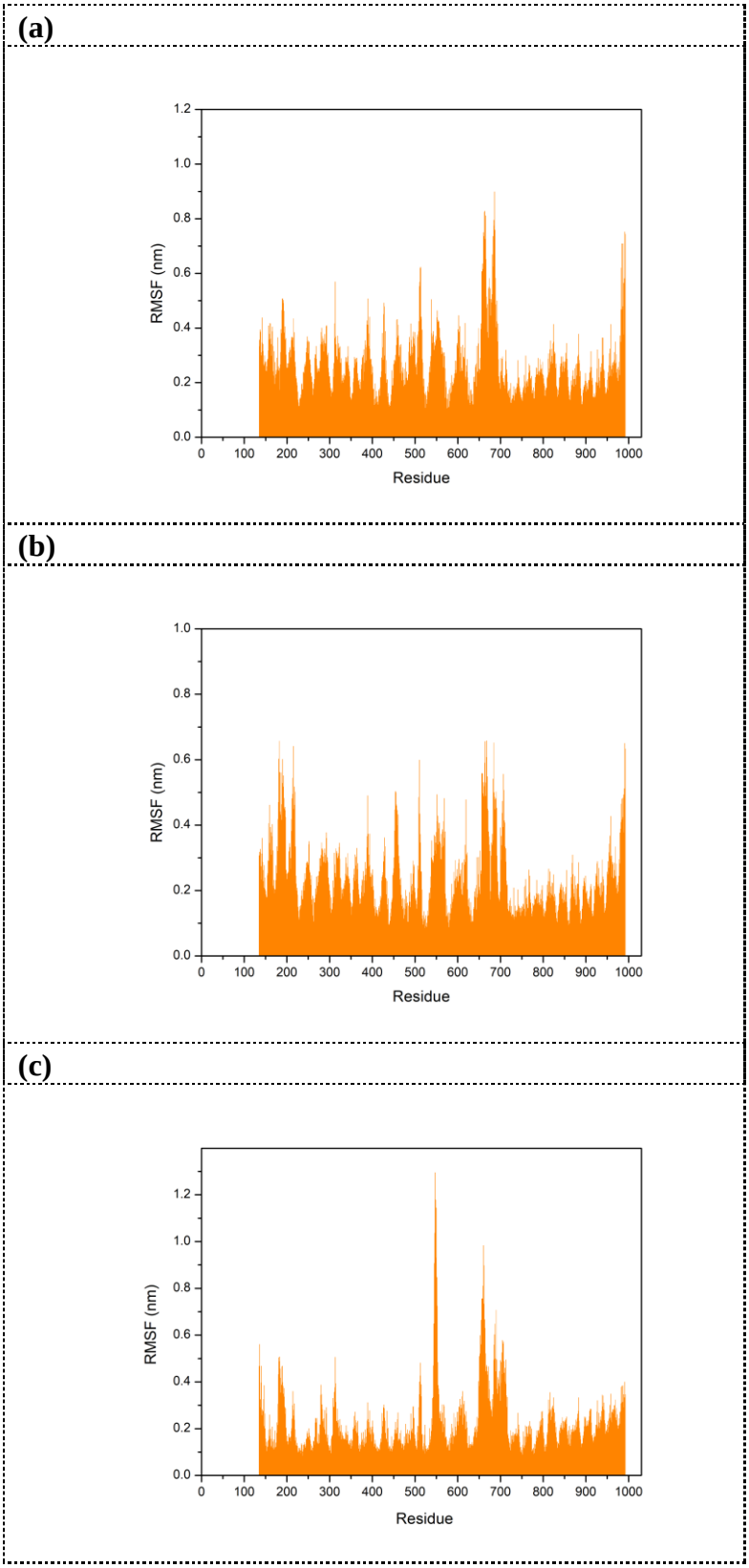


Figure 24. RMSF value of each residue on various protein. The figure can intuitively show the amplitude of each residual and show which residual has a larger range of variation. The abscissa was the number of protein residue sequences. Whether different ligands have similar effects to the target protein can also be judged by the figure. (a) 2007_15317 (b) 2007_22057 (c) 2007_22325.

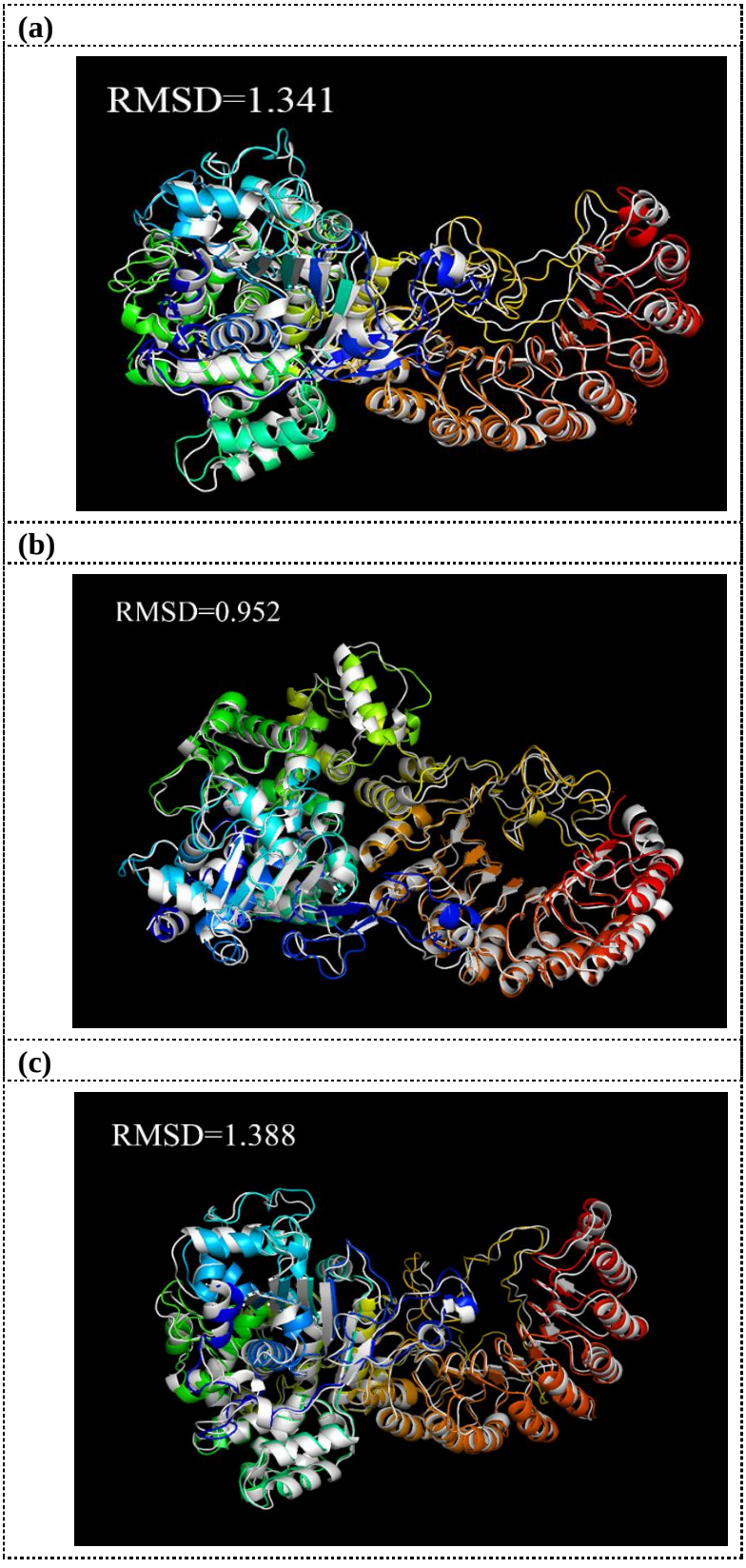


Figure 25. Average structure of each proteins reacted with different ligands. The average structure and the final state structure are superimposed to obtain RMSD values, and these structures were to observe whether their conformational changes were consistent, which indicate that the final state structure has good stability. (a) 2007_15317 (b) 2007_22057 (c) 2007_22325. (The ligand was not shown in figure)

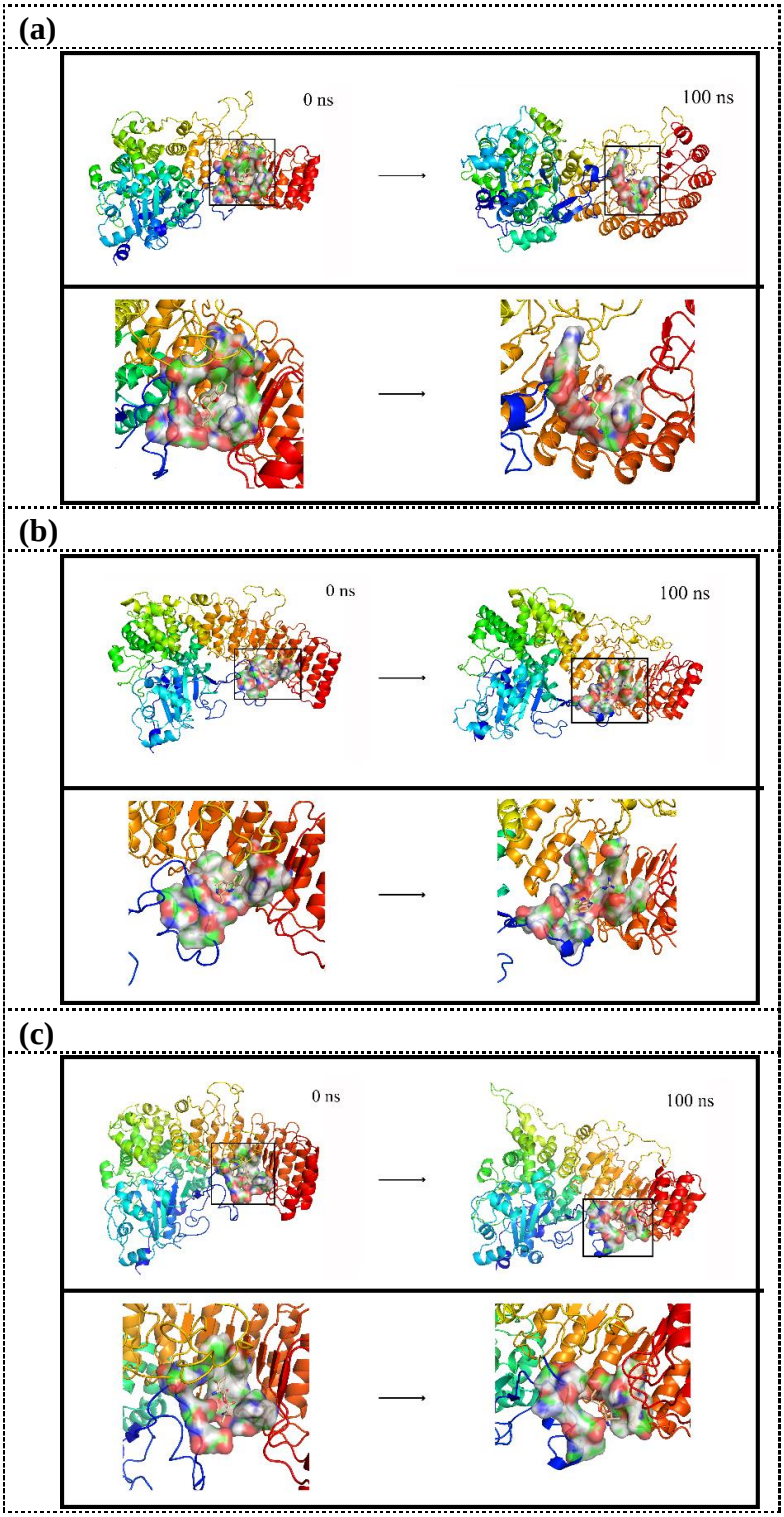


Figure 26. Combining posture changes during MD in microenvironment. Although the state of the ligand changes, the position where the ligand binds to the target protein does not change. (a) 2007_15317 (b) 2007_22057 (c) 2007_22325