

**Supplementary material**

**Could the FDA-approved anti-HIV PR inhibitors be promising anticancer agents? An answer from enhanced docking and molecular dynamics analyses**

Olayide A. Arodola<sup>a</sup> and Mahmoud E. S. Soliman<sup>a\*</sup>

**Figure S1: The multiple sequence alignment result from CLUSTAW**

CLUSTAL 2.1 multiple sequence alignment

```

3PRY_A|PDBID|CHAIN|SEQUENCE      -----
sp|P08238|HS90B_HUMAN            MPEEVHGGEEVETFAFQAEIAQLMSLIINTFYYSNKEIFLRELISNASDA 50
2CG9_B|PDBID|CHAIN|SEQUENCE      -----MASETFEFQAEITQLMSLIINTVYSNKEIFLRELISNASDA 41
3HJC_A|PDBID|CHAIN|SEQUENCE      -----

3PRY_A|PDBID|CHAIN|SEQUENCE      -----
sp|P08238|HS90B_HUMAN            LDKIRYESLTDPSKLDGKELKIDIIIPNQERTLTLVDTGIGMTKADLIN 100
2CG9_B|PDBID|CHAIN|SEQUENCE      LDKIRYKSLSDPKQLETEPDLFIRITPKPEQKVLEIRDSGIGMTKAELIN 91
3HJC_A|PDBID|CHAIN|SEQUENCE      -----

3PRY_A|PDBID|CHAIN|SEQUENCE      -----
sp|P08238|HS90B_HUMAN            NLGTTAKSGTKAFMEALQAGADISMIGQFVGVFYSAYLVAEKVVVITKHN 150
2CG9_B|PDBID|CHAIN|SEQUENCE      NLGTTAKSGTKAFMEALSAGADVSMIGQFVGVFYSLFLVADRVQVISKSN 141
3HJC_A|PDBID|CHAIN|SEQUENCE      -----

3PRY_A|PDBID|CHAIN|SEQUENCE      -----
sp|P08238|HS90B_HUMAN            DDEQYAWESSAGGSFTVRADHG-EPIGRGTVILHLKEDQTEYLEERRVK 199
2CG9_B|PDBID|CHAIN|SEQUENCE      DDEQYIWESNAGGSFTVTLDEVNERIGRTILRLFLKDDQLEYLEEKRIK 191
3HJC_A|PDBID|CHAIN|SEQUENCE      -----

3PRY_A|PDBID|CHAIN|SEQUENCE      -----
sp|P08238|HS90B_HUMAN            EVVKKHSQFIGYPIITLYLEKEREKESDDEAEEKGEKEEEDKDDEE--K 247
2CG9_B|PDBID|CHAIN|SEQUENCE      EVIKRHSFVAYPIQLVVTKEVEKEVPIPEEKKDEEKKDEEKKDEDDKK 241
3HJC_A|PDBID|CHAIN|SEQUENCE      -----

3PRY_A|PDBID|CHAIN|SEQUENCE      -----MKTKPIWTRNPDDIT 15
sp|P08238|HS90B_HUMAN            PKIEDVGSDEEDDSGKDKKKKTKKIKEYIDQEELNKTPIWTRNPDDIT 297
2CG9_B|PDBID|CHAIN|SEQUENCE      PKLEEVDEEEE-----KKPKTKKVKEEVQEIEELNKTPLWTRNPSDIT 285
3HJC_A|PDBID|CHAIN|SEQUENCE      -----MHHHHHSSGRENLYFQGHKPLWTRDPKDVIT 31
                                     **:***:*.**:

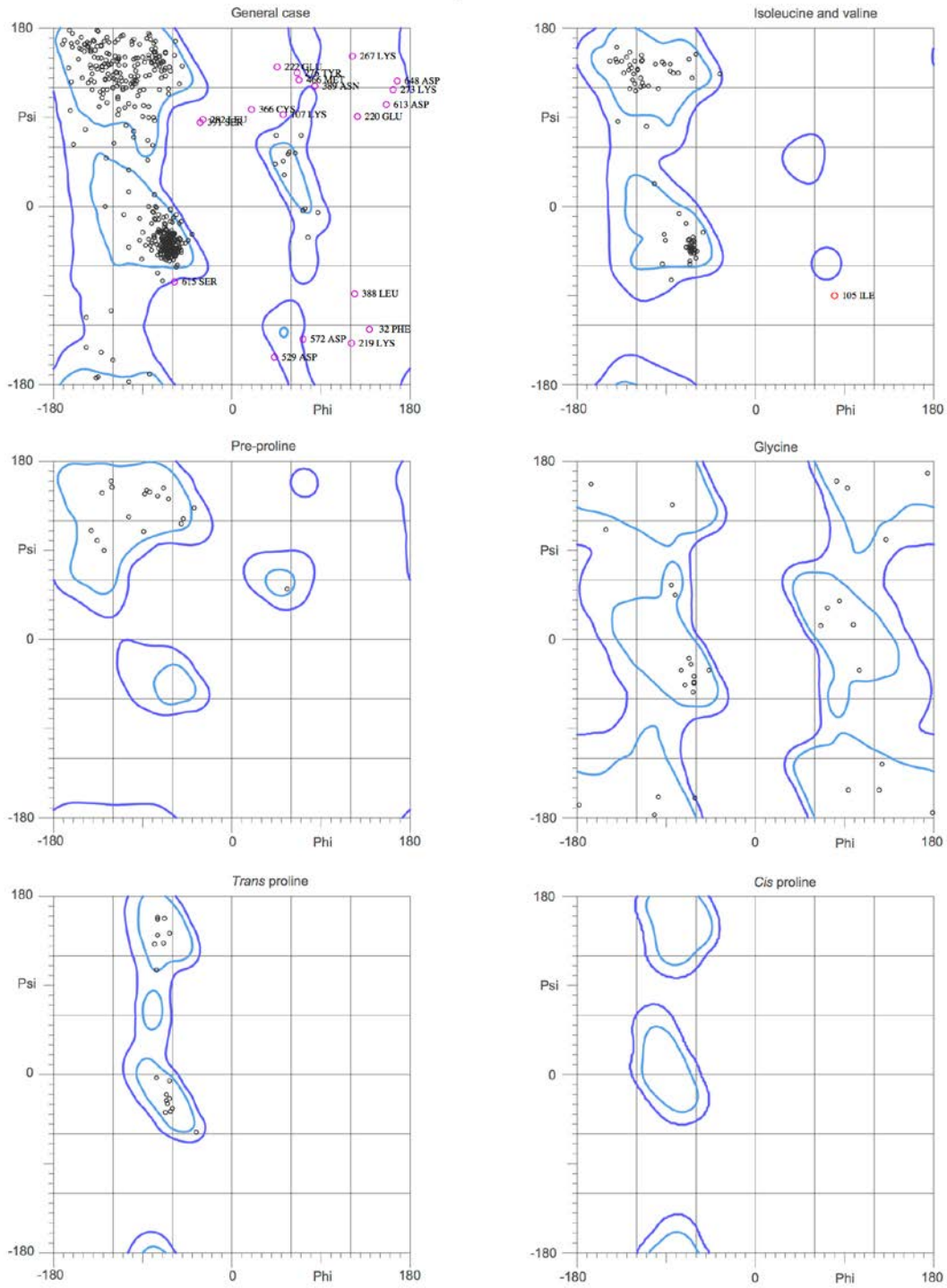
3PRY_A|PDBID|CHAIN|SEQUENCE      QEEYGEFYKSLTNDWEDHLAVKHFVVEGQLEFRALLFIPRRAPFDLFENK 65
sp|P08238|HS90B_HUMAN            QEEYGEFYKSLTNDWEDHLAVKHFVVEGQLEFRALLFIPRRAPFDLFENK 347
2CG9_B|PDBID|CHAIN|SEQUENCE      QEEYNAFYKSI SNDWEDPLVVKHFVVEGQLEFRALLFIPRRAPFDLFESK 335
3HJC_A|PDBID|CHAIN|SEQUENCE      KEEYAAFYKAI SNDWEDPAATKHFVVEGQLEFRSIFVVPKRAPDFMFEPN 81
:***  ***:::***** .*****:::***:*****:

```

3PRY_A   PDBID   CHAIN   SEQUENCE	KKKNNIKLYVRRVFMDSCELIPEYLNFRGVVDS	EDLPLNISREMLQQ	115
sp   P08238   HS90B_HUMAN	KKKNNIKLYVRRVFMDSCELIPEYLNFRGVVDS	EDLPLNISREMLQQ	397
2CG9_B   PDBID   CHAIN   SEQUENCE	KKKNNIKLYVRRVITDEAEDLIPWLSFVKGVVDS	EDLPLNLSREMLQQ	385
3HJC_A   PDBID   CHAIN   SEQUENCE	KKRNNIKLYVRRVIMDNCELDLCPDWLGFKVGVVDS	EDLPLNISRENLQQ	131
	***:***** * . . . . . * : : . * : : ***** : ** *		
3PRY_A   PDBID   CHAIN   SEQUENCE	SKILKVIKKNIVKKCLELFS	ELAEKKNYKFFYEAF	SKNLKLGIHEDSTN 165
sp   P08238   HS90B_HUMAN	SKILKVIKKNIVKKCLELFS	ELAEKKNYKFFYEAF	SKNLKLGIHEDSTN 447
2CG9_B   PDBID   CHAIN   SEQUENCE	NKIMKVIKKNIVKKLIEAFNEIAEDSEQFEKFF	YAFSKNKLGVHEDTQN 435	
3HJC_A   PDBID   CHAIN   SEQUENCE	NKILKVIKKNIVKKCLEMPDEVAENKEDYKQFYEQ	FGKNIKLGIHEDTAN 181	
	. ***:***** : * * . * : : . * : : : : * . * : : ***** : *		
3PRY_A   PDBID   CHAIN   SEQUENCE	RRRLSELLRYHTSQSGDEMTSLSEYVSRMKETQKSI	IYYITGESKEQVANS 215	
sp   P08238   HS90B_HUMAN	RRRLSELLRYHTSQSGDEMTSLSEYVSRMKETQKSI	IYYITGESKEQVANS 497	
2CG9_B   PDBID   CHAIN   SEQUENCE	RAALAKLLRYNSTKSVDELTS	LTDYVTRMPEHQKNIYYITGESLKAVERK 485	
3HJC_A   PDBID   CHAIN   SEQUENCE	RKKLMELLRFYSTESGEEMTTLKDYVTRMKAGQKSI	IYYITGDSKKKLETS 231	
	* * : ***: : : * : * : * : * . * : : * : : * : * : : : *		
3PRY_A   PDBID   CHAIN   SEQUENCE	AFVERVRKRGFEVVMTEPIDEYCVQQLKEFDGKSLVSVT	KEGLELAEN-	264
sp   P08238   HS90B_HUMAN	AFVERVRKRGFEVVMTEPIDEYCVQQLKEFDGKSLVSVT	KEGLELPEDE	547
2CG9_B   PDBID   CHAIN   SEQUENCE	PFLDALKAKNFEVLFLTDPIDEYAF	TQLKEFEGKTLVDITKD-FELEETD 534	
3HJC_A   PDBID   CHAIN   SEQUENCE	PFIEQARRRGLEVLFTTEPIDEYVMQVQKDFEDK	FFACLTKEGVHFESEE 281	
	. * : : : : : : * : : * : : * . * : : * : : * : : * : : *		
3PRY_A   PDBID   CHAIN   SEQUENCE	-----LYFQ-----		268
sp   P08238   HS90B_HUMAN	EEKKMEESKAKFNLC	KLMKEILDKKVEKVTISNRLVSSPCCIVTSTYG 597	
2CG9_B   PDBID   CHAIN   SEQUENCE	EKAEREKEIKEYEPLTKALKEILGDQVEKVVVSYKLLD	DAPAAIRTGQFG 584	
3HJC_A   PDBID   CHAIN   SEQUENCE	EKKQREEEKAACEKLCKTMKEVLGDKVEKVVVSE	RLSTSPCIIVTSEFG 331	
	:		
3PRY_A   PDBID   CHAIN   SEQUENCE	-----		
sp   P08238   HS90B_HUMAN	WTANMERIMKAQALRDNSTMGYMAKKHLEINPDHPIVET	LRQKA-EADK 646	
2CG9_B   PDBID   CHAIN   SEQUENCE	WSANMERIMKAQALRDSSMSYSSKKTFFISPKSPIIKEL	KKRVDGGA 634	
3HJC_A   PDBID   CHAIN   SEQUENCE	WSAHMEQIMRNQALRDSSMAQYMSKKTMELNPRHP	IIKELRRRV-GADE 380	
3PRY_A   PDBID   CHAIN   SEQUENCE	-----		
sp   P08238   HS90B_HUMAN	NDKAVKDLVLLFFETALLSSGFSLEDPQTHSNRIYRMIK	LGIGIDEDEVA 696	
2CG9_B   PDBID   CHAIN   SEQUENCE	QDKTVKDLTKLLYETALLTSGFSLDEPTSFASRINRLISL	GLN----- 677	
3HJC_A   PDBID   CHAIN   SEQUENCE	NDKAVKDLVLLFDTSLLTSGFQLEDPTGYAERINRMIK	LGLSLDEEEEE 430	
3PRY_A   PDBID   CHAIN   SEQUENCE	-----		
sp   P08238   HS90B_HUMAN	AEEPNAAVPEIPPLEGDEDA	SRMEEVD 724	
2CG9_B   PDBID   CHAIN   SEQUENCE	-----		
3HJC_A   PDBID   CHAIN   SEQUENCE	AAEAPVAETAPAEV-----	444	

The Figure above shows the 2D sequence multi-alignment of Hsp90 from *saccharomyces cerevisiae* (PDB Code: 2CG9), which contained the ATP bound in its active site; Hsp90 middle domain from *homo sapiens* (PDB Code: 3PRY), Hsp90 C- terminal domain from *Leishmania major* (PDB Code: 3HJC) and the human Hsp90 obtained from Uniprot (fasta code P08238).

**Figure S2: Ramachandran plot for the human Hsp90 homologue**



**Below is the molprobit analyses of the Ramachandran plot**

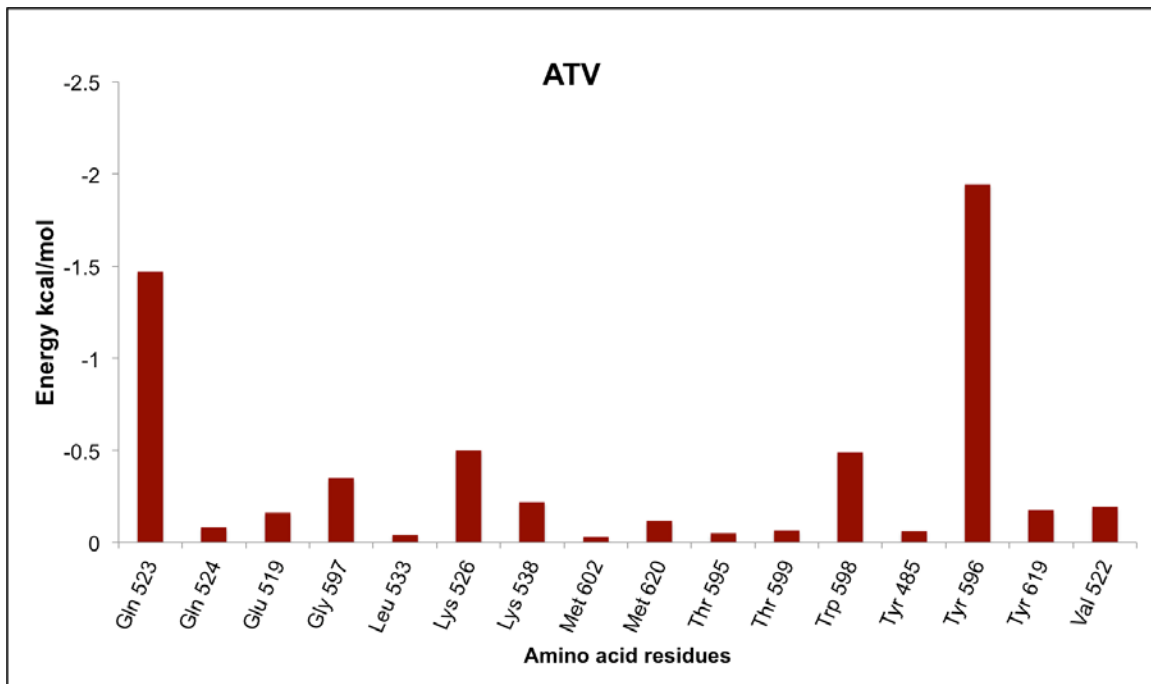
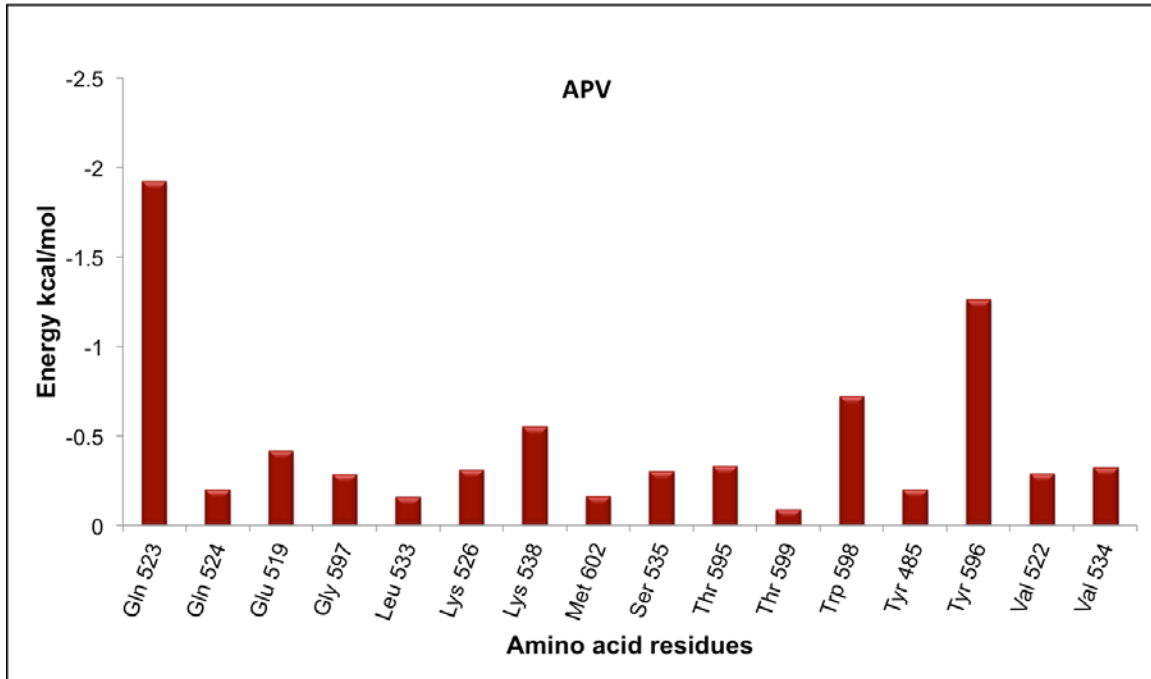
91.7% (620/676) of all residues were in favored (98%) regions.

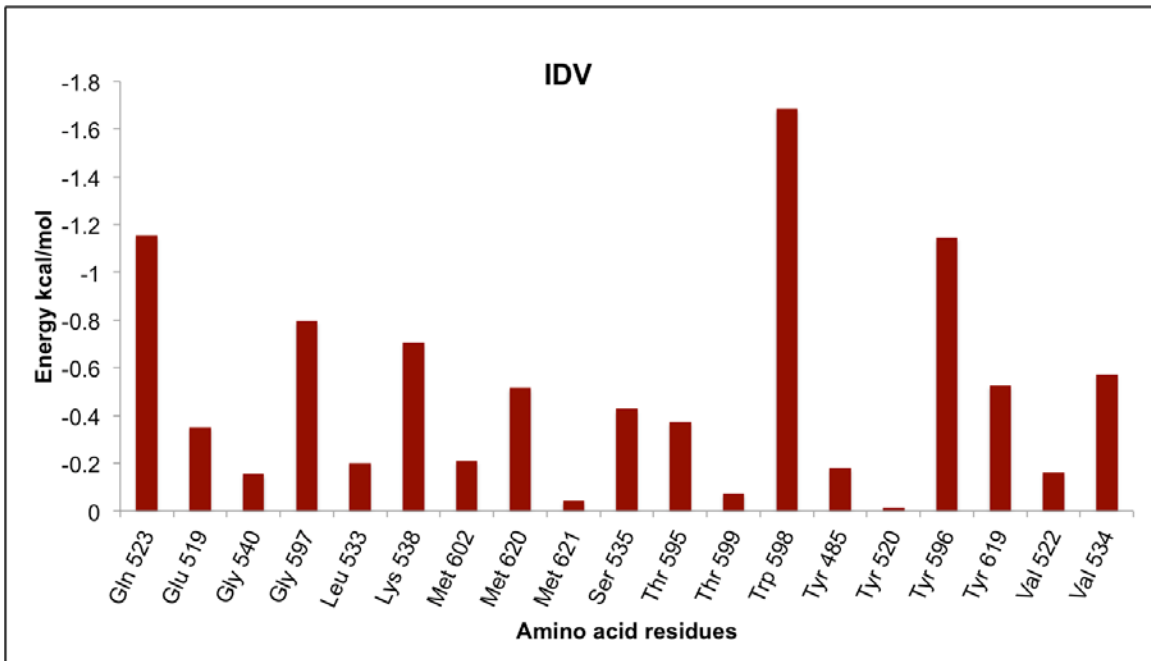
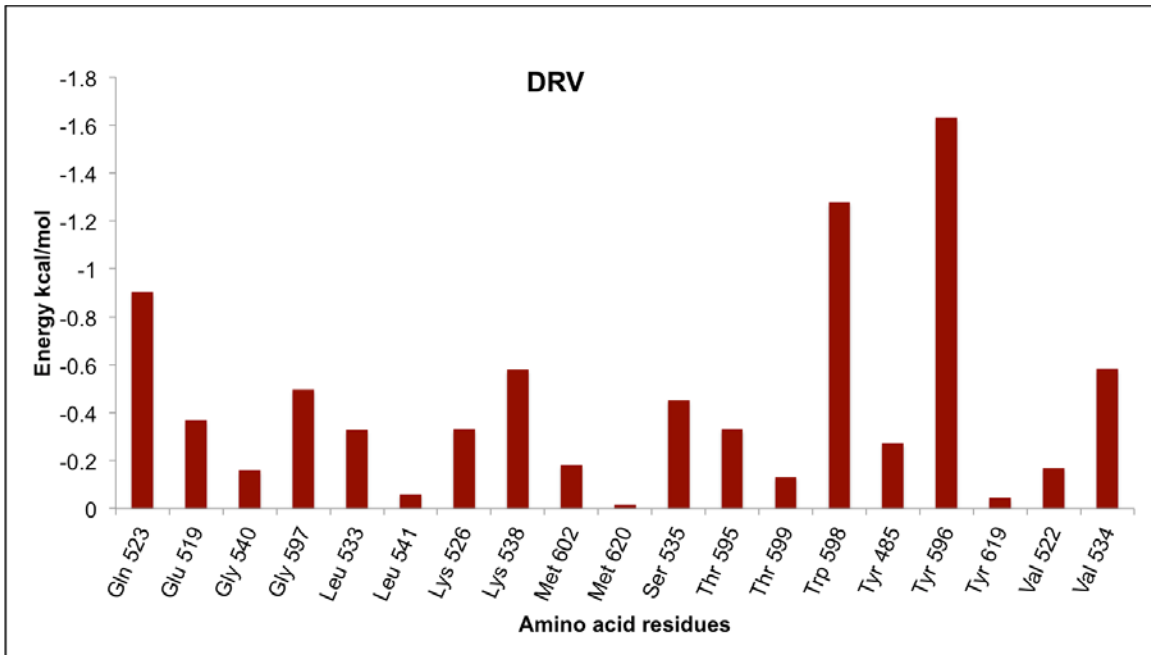
97.0% (656/676) of all residues were in allowed (>99.8%) regions.

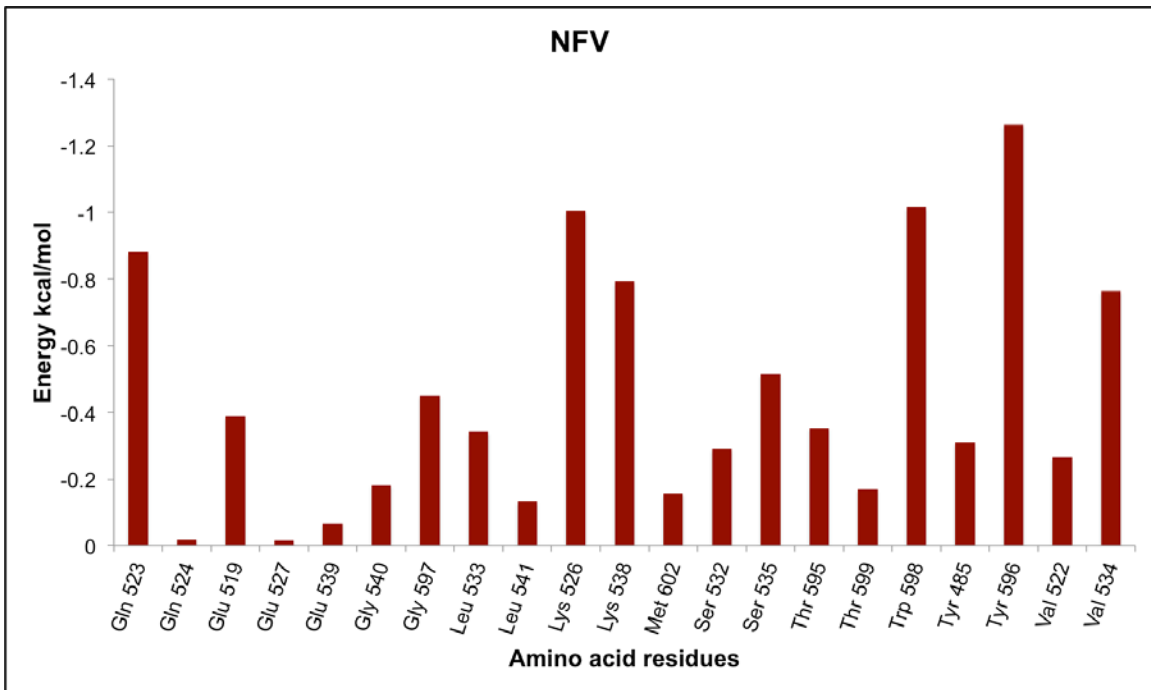
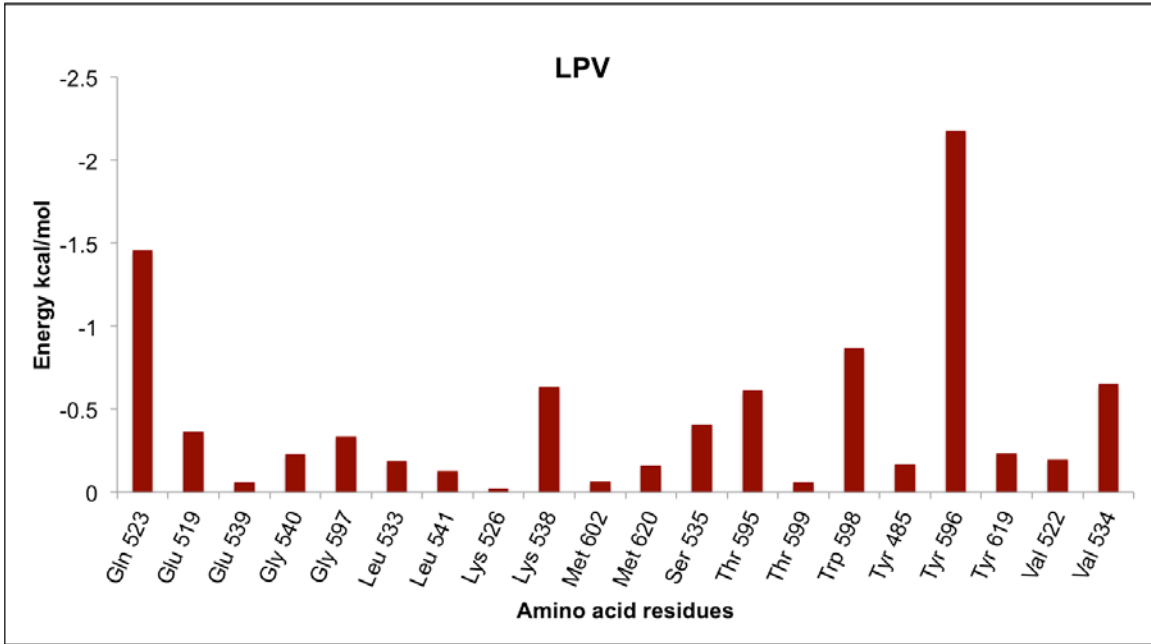
There were 20 outliers (phi, psi):

32 PHE (140.0, -124.6)  
105 ILE (80.7, -90.8)  
107 LYS (53.0, 93.3)  
219 LYS (121.6, -138.6)  
220 GLU (127.9, 92.0)  
222 GLU (46.8, 141.4)  
267 LYS (122.1, 153.0)  
273 LYS (164.0, 118.7)  
276 TYR (66.5, 135.4)  
282 LEU (-29.3, 88.2)  
366 CYS (20.7, 98.3)  
388 LEU (124.5, -88.7)  
389 ASN (84.2, 123.0)  
391 SER (-32.3, 85.1)  
466 MET (68.7, 128.0)  
529 ASP (43.6, -152.6)  
572 ASP (72.8, -134.9)  
613 ASP (156.6, 103.0)  
615 SER (-58.5, -76.8)  
648 ASP (167.7, 127.8)

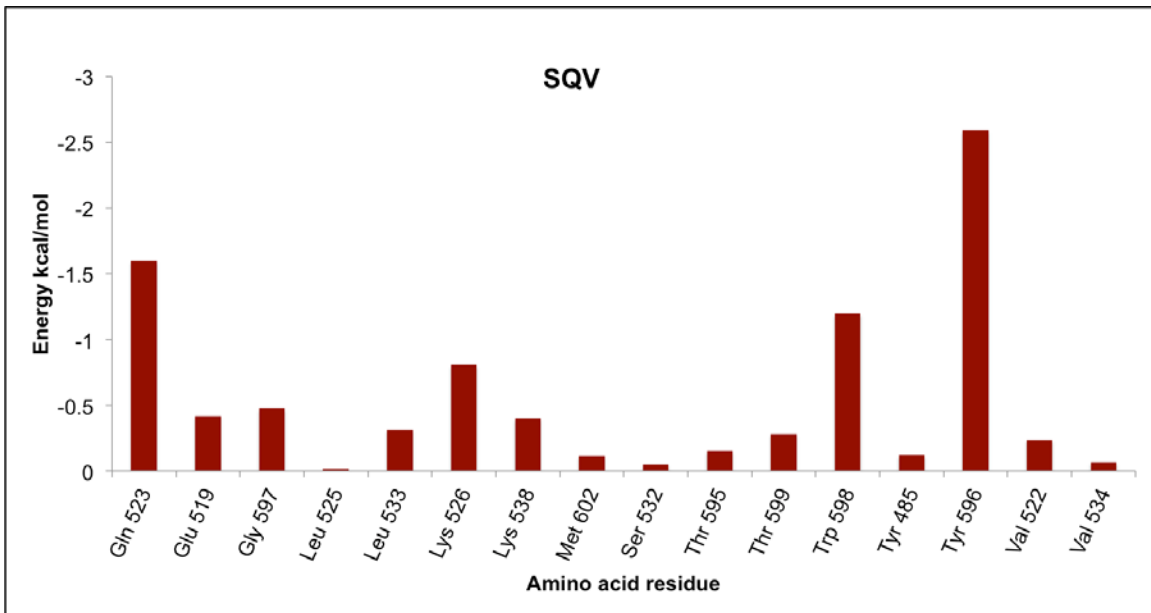
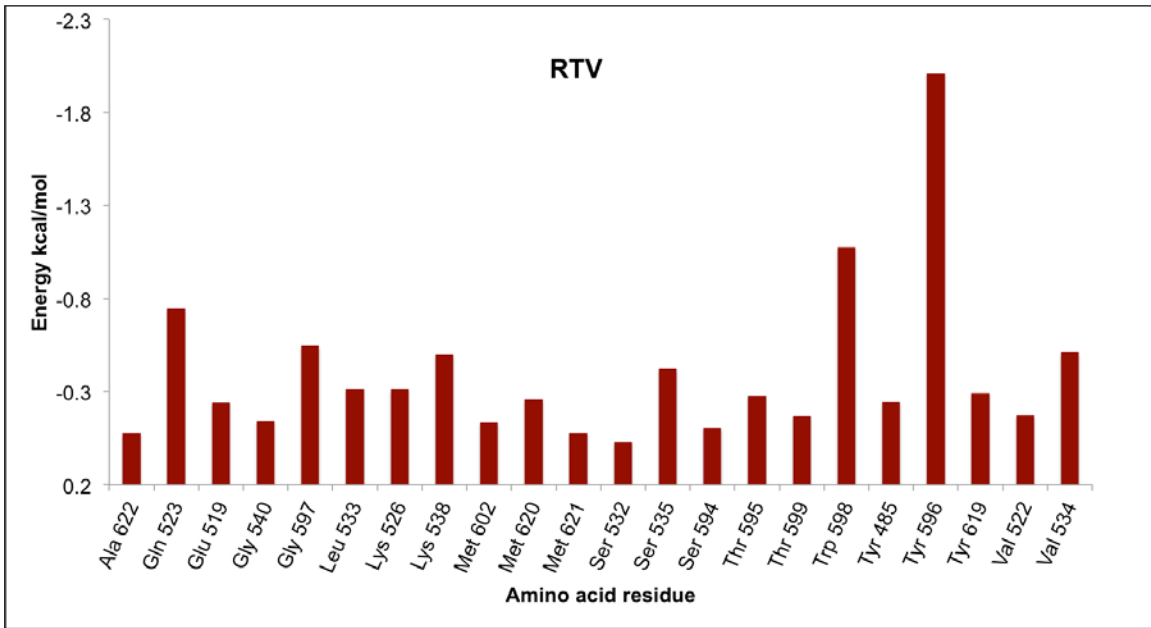
**Figure S3: Per-residue decomposition analyses plot for all inhibitors at the Hsp90 active site pocket**











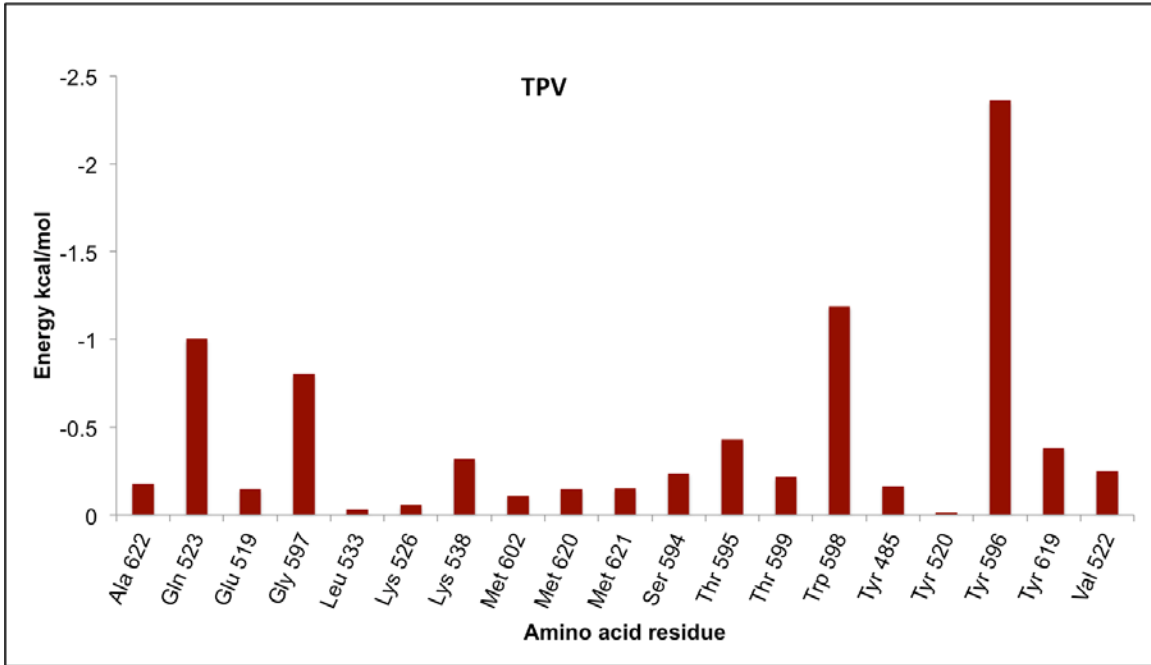
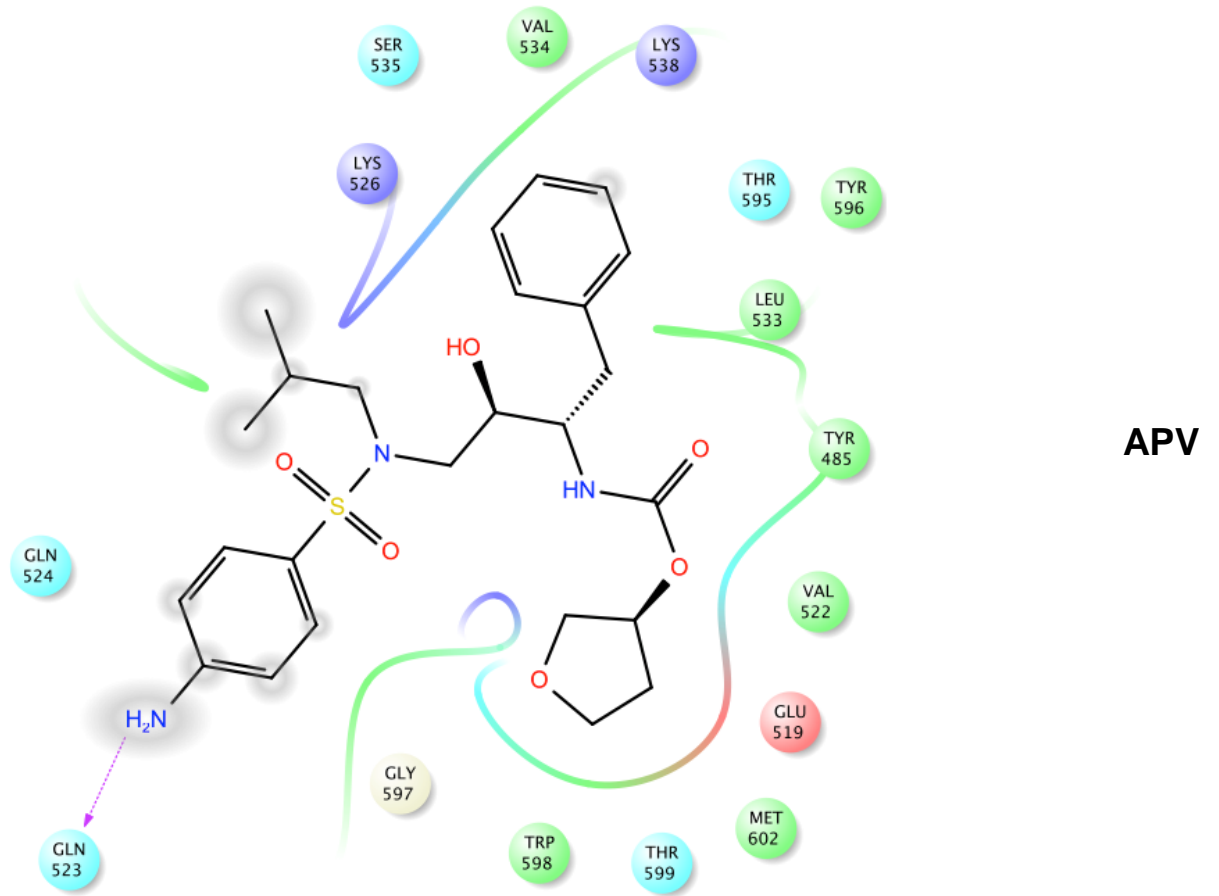
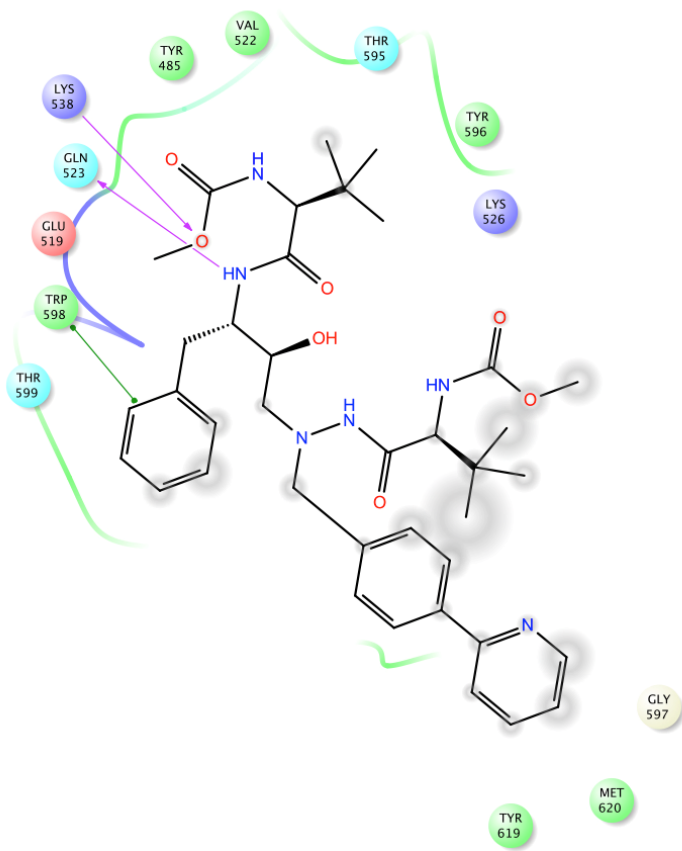
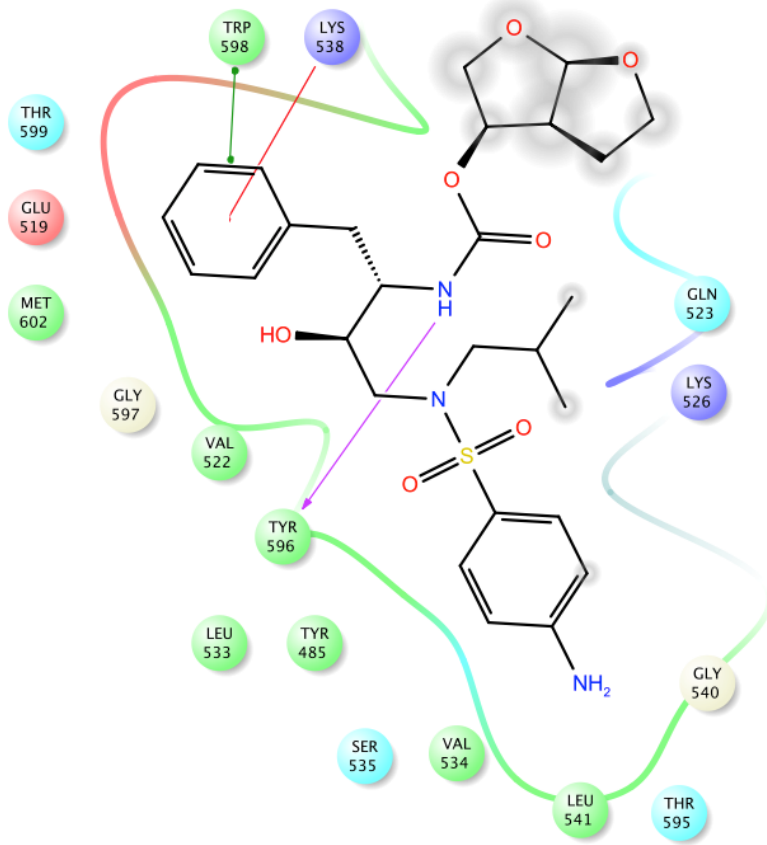


Figure S4: Ligand-enzyme interactions for all inhibitors

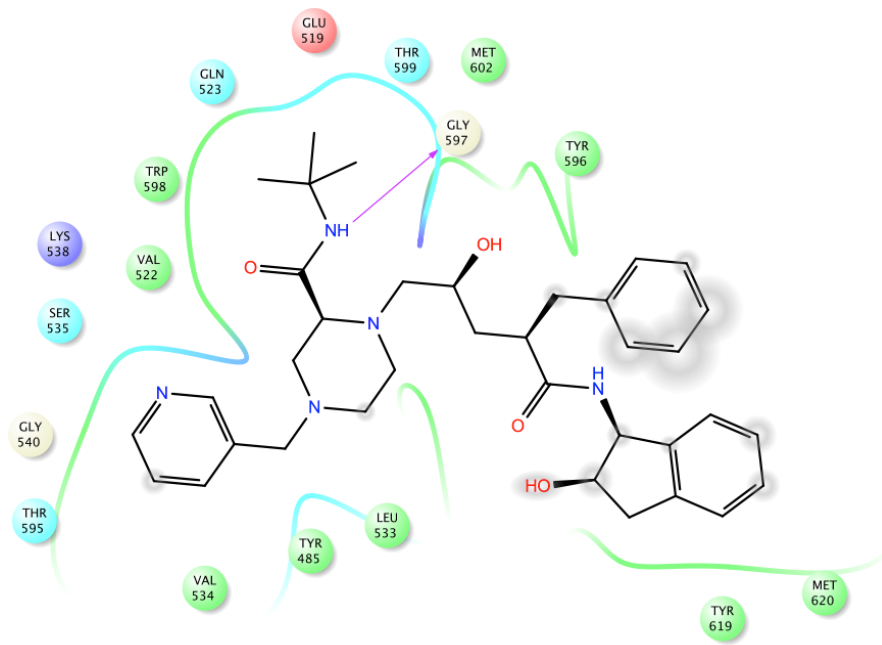




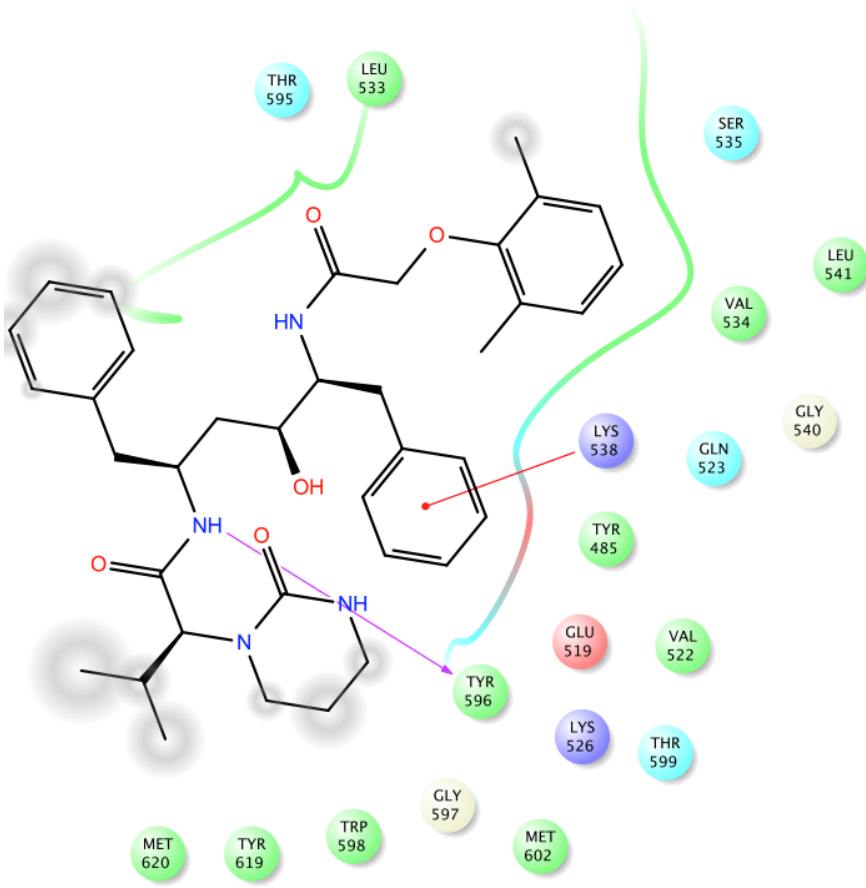
ATV



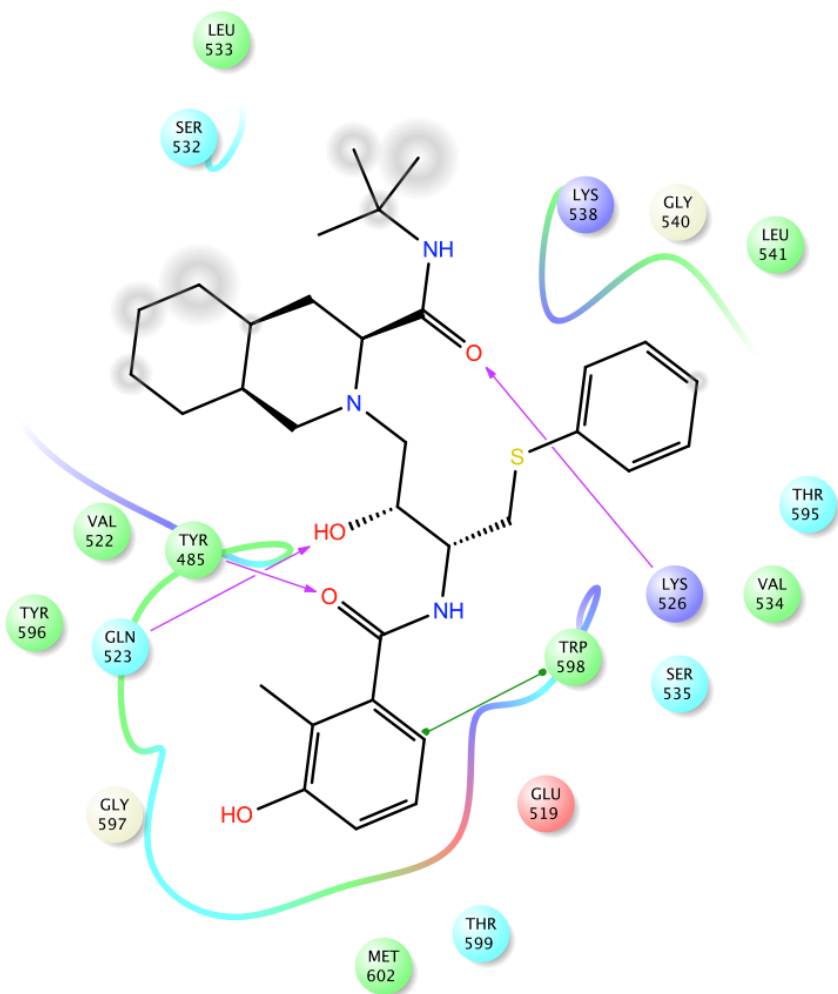
DRV



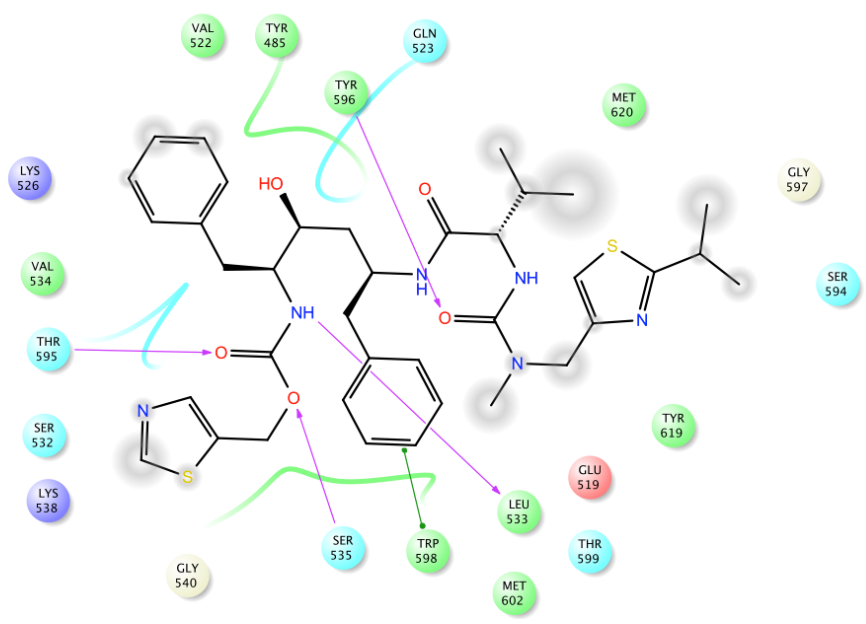
IDV



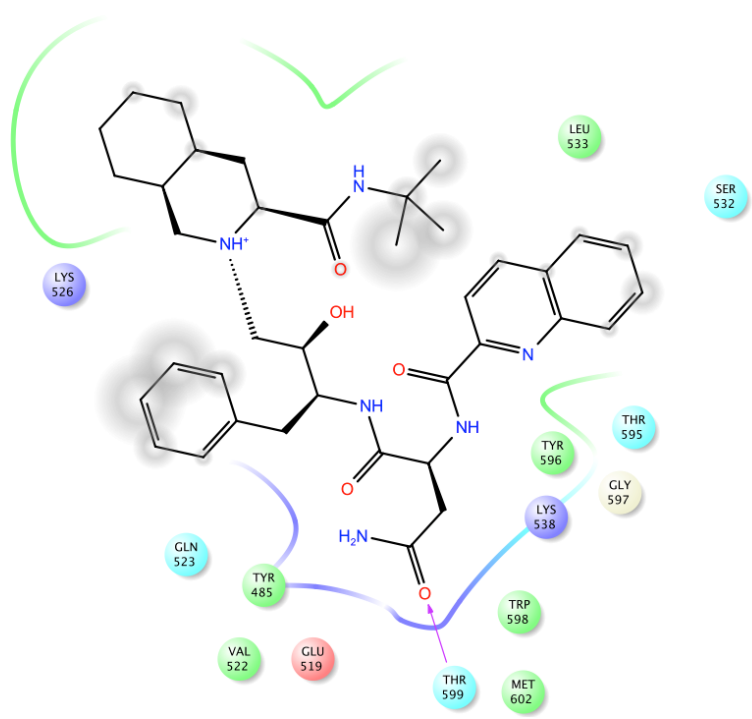
LPV



**NFV**

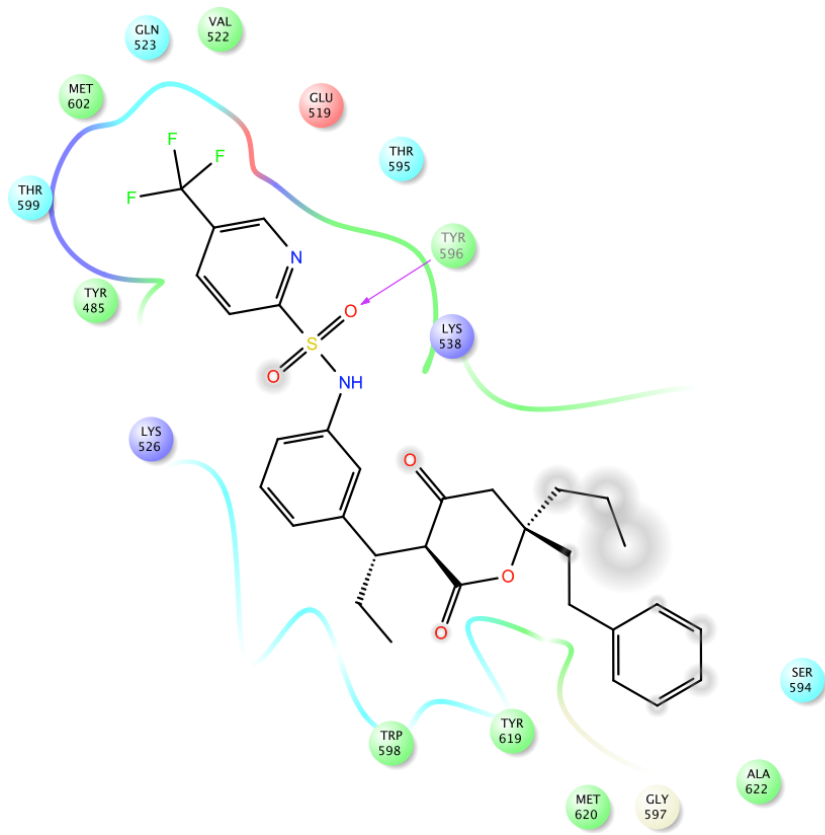


RTV



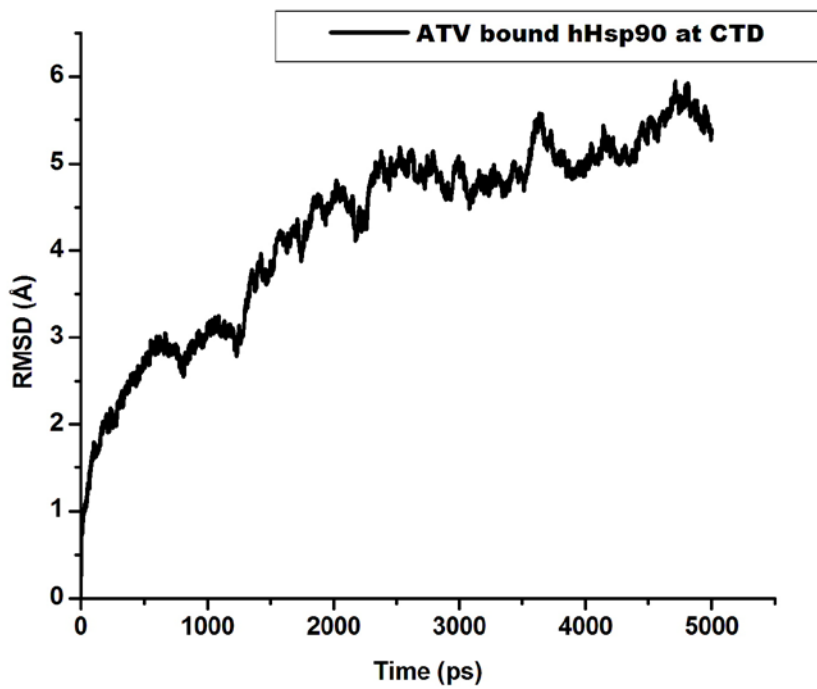
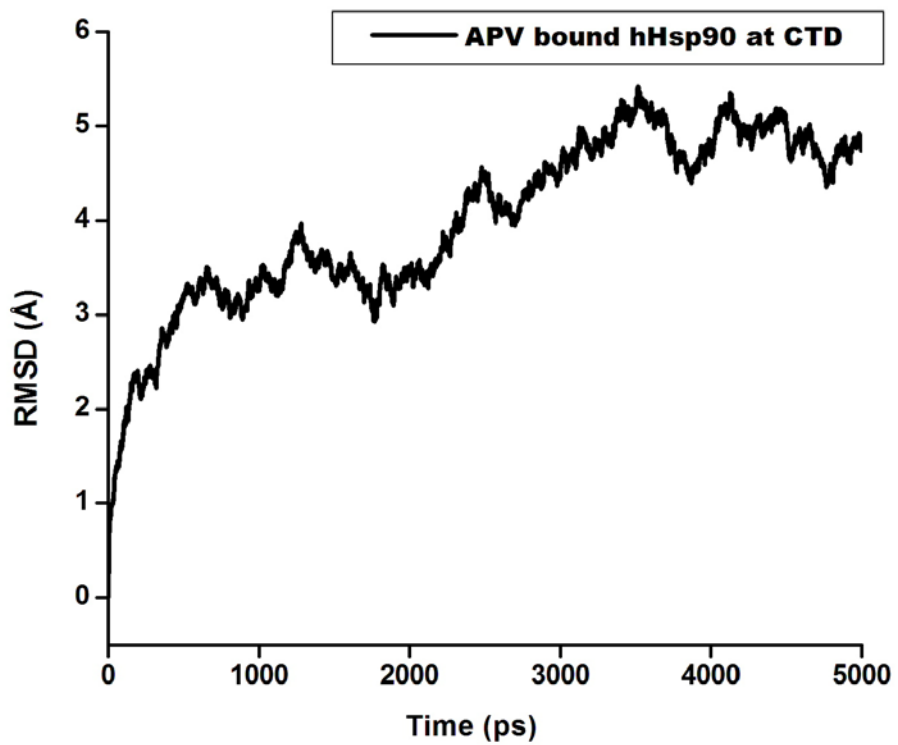
SQV

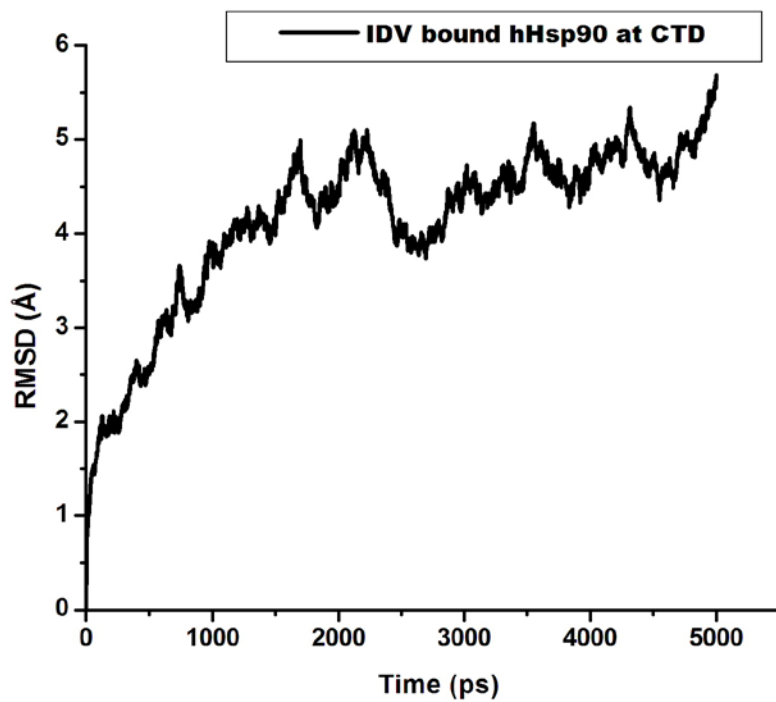
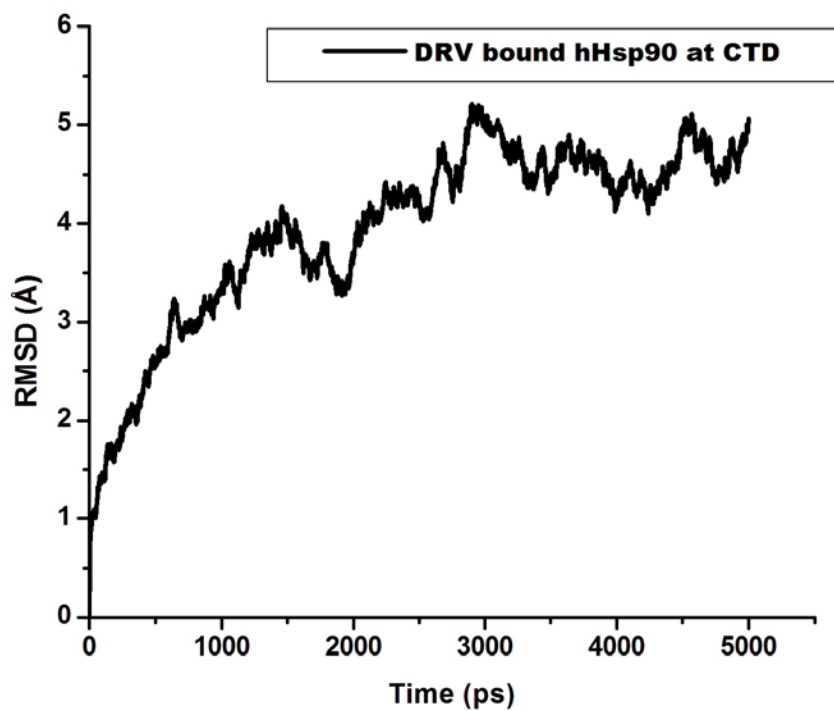


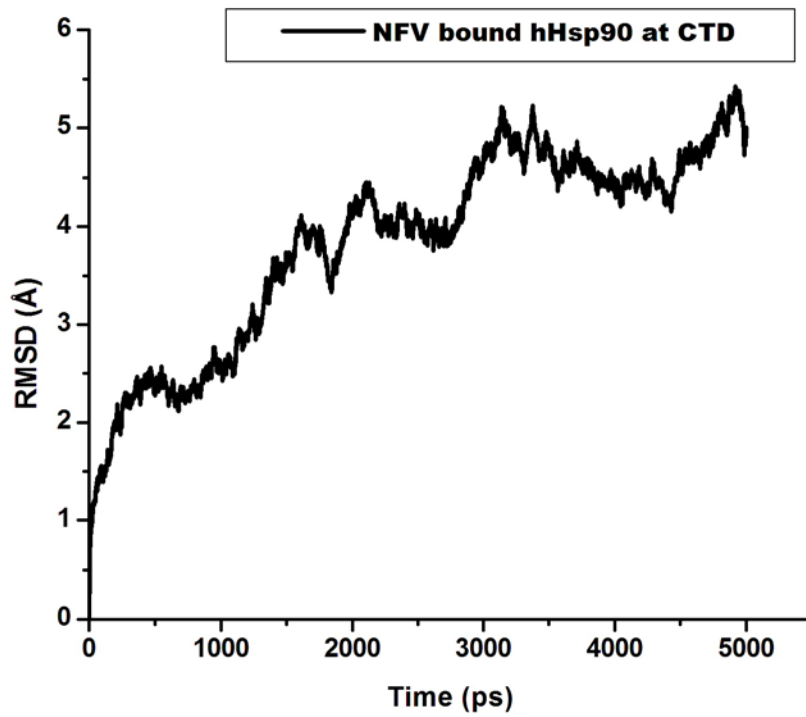
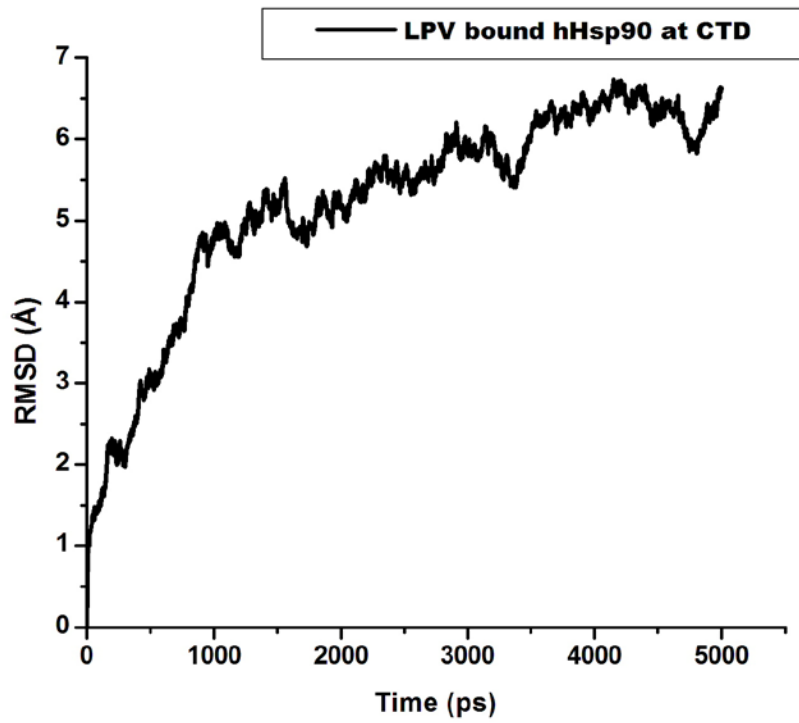


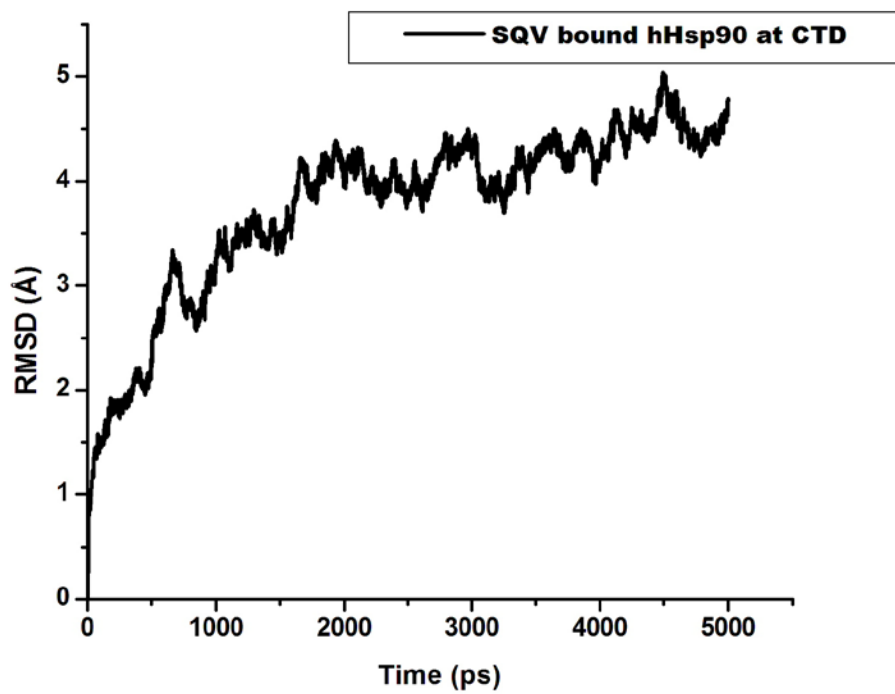
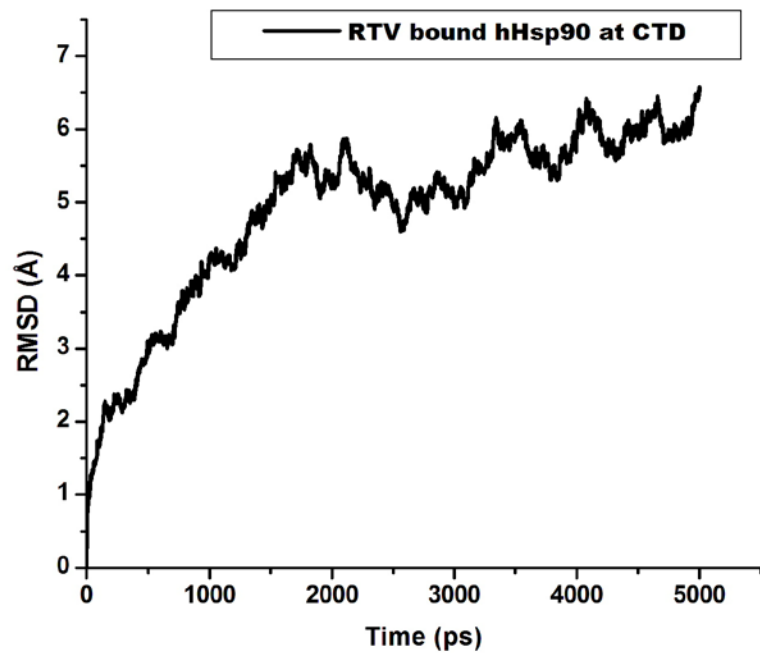
TPV

Figure S5: RMSD plots for all inhibitors









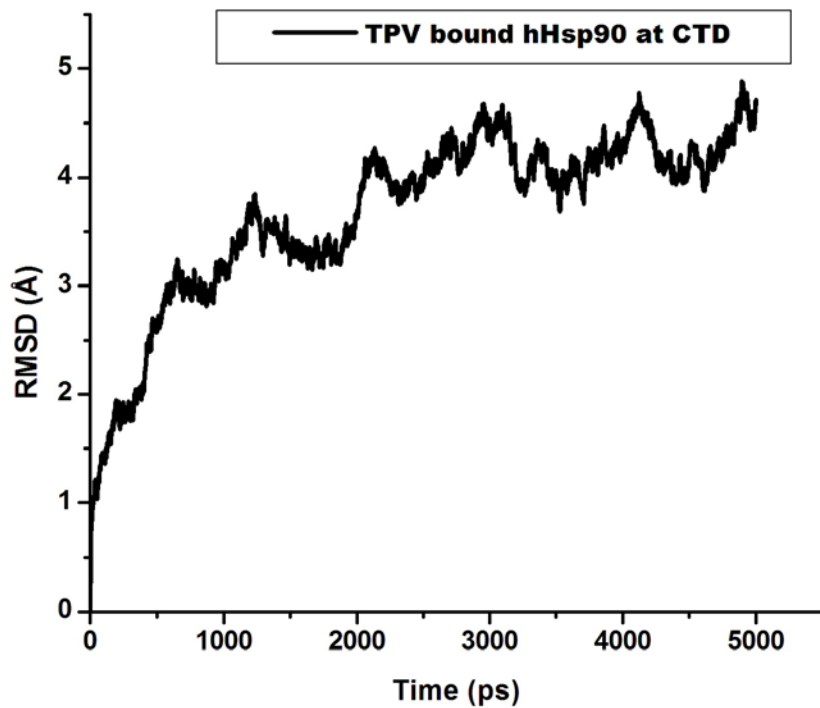
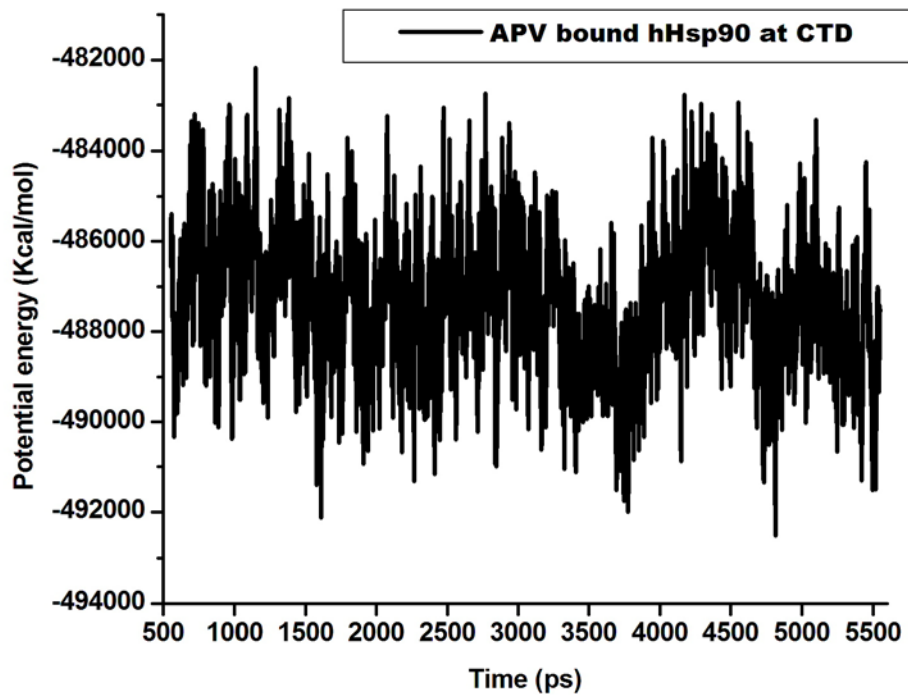
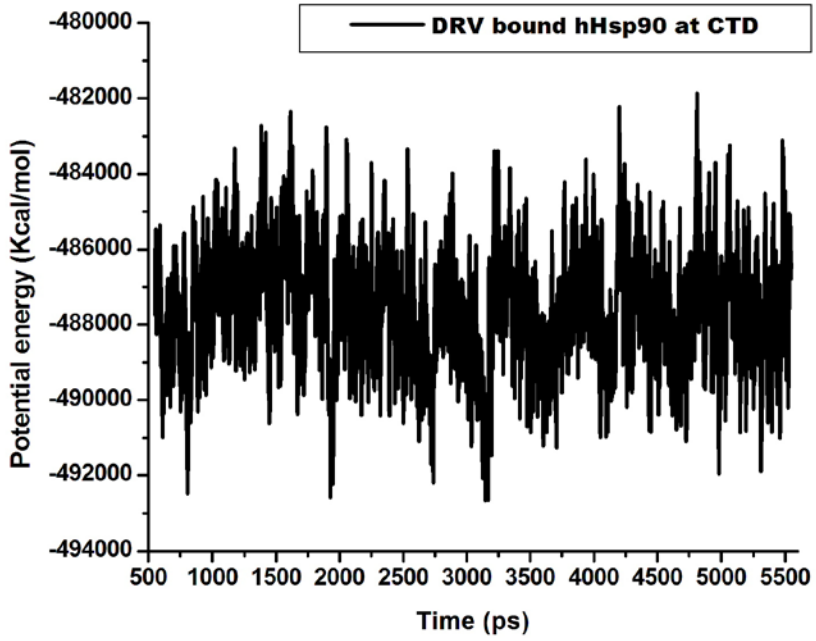
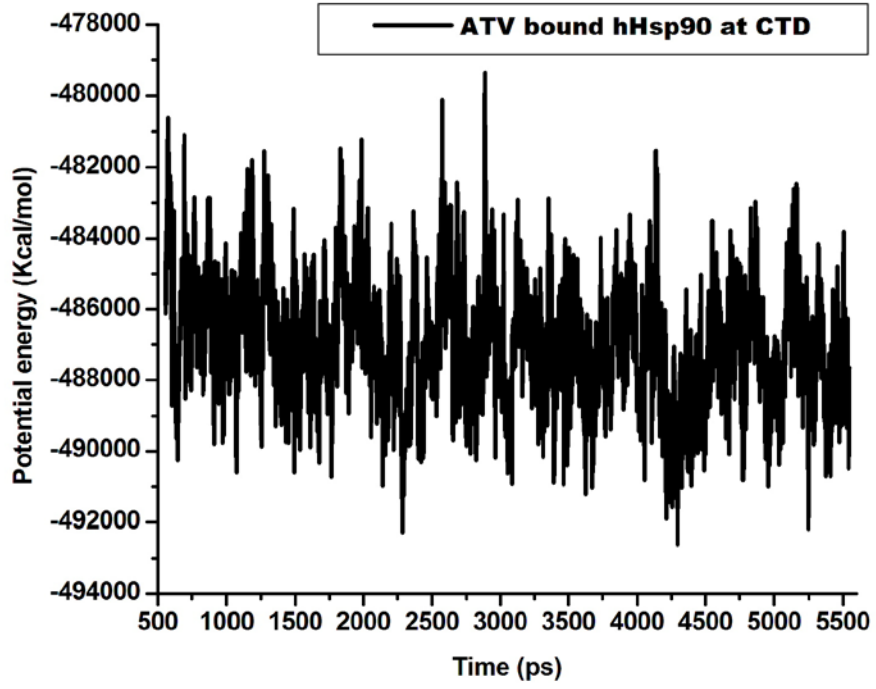
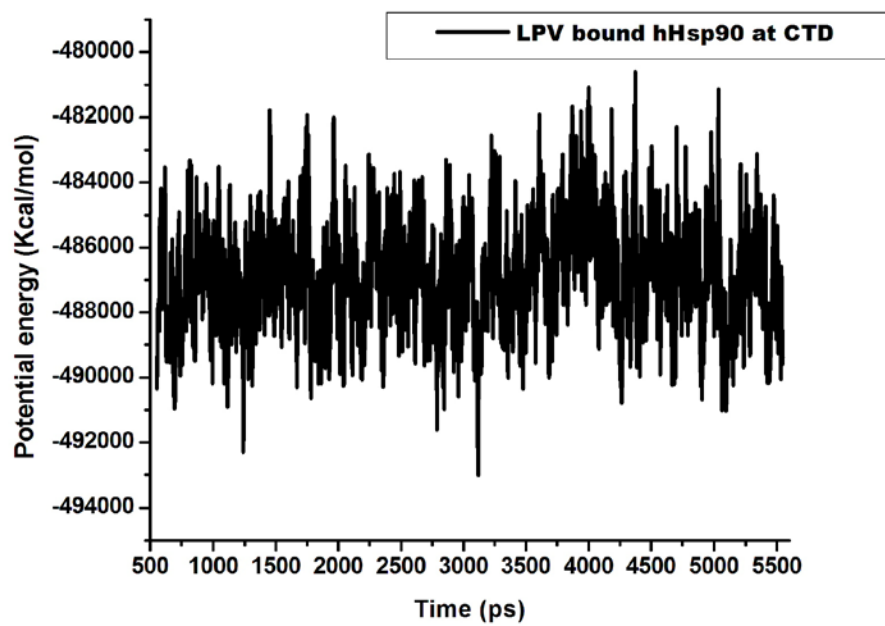
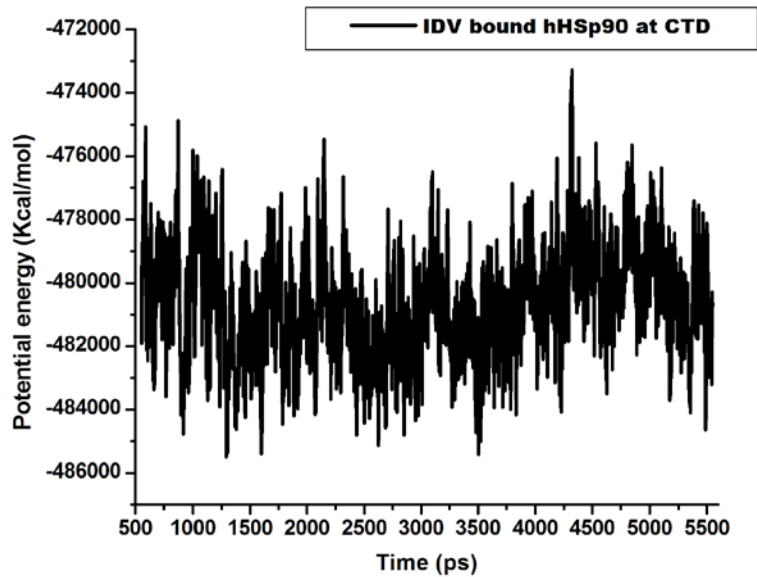


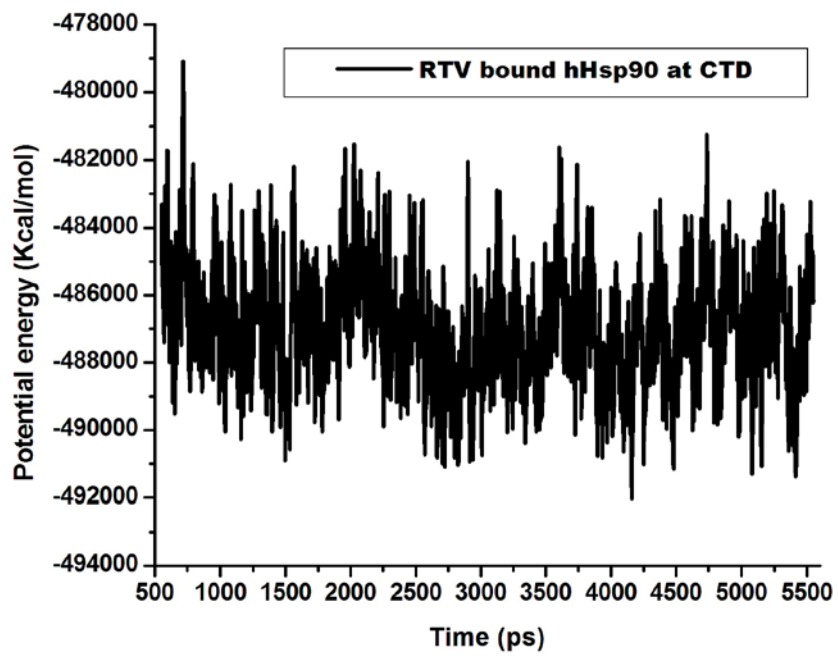
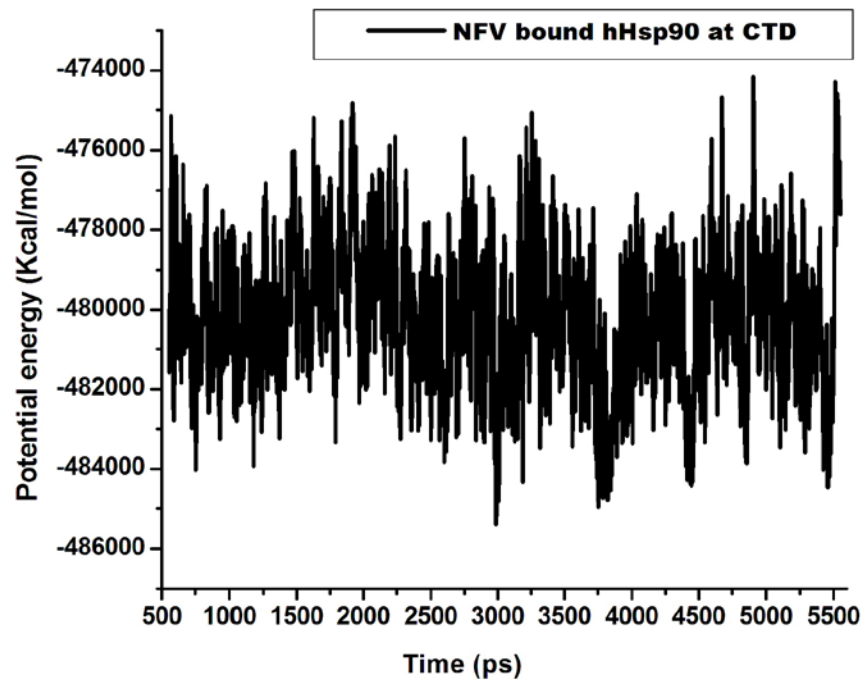
Figure S6: Potential energy plots for all inhibitors

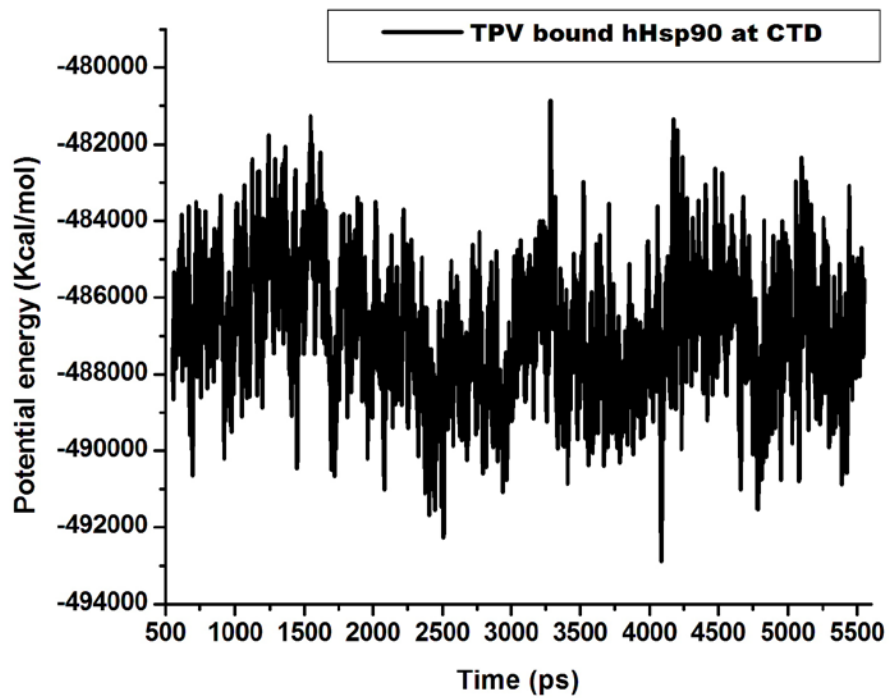
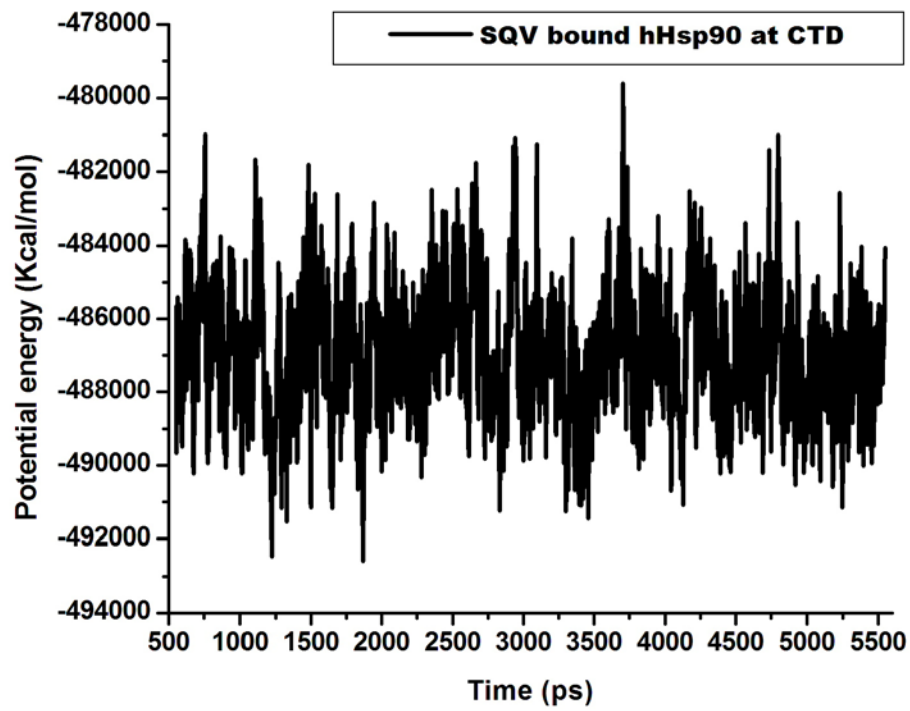












## Figure S7. Validation of SiteHound-web predictions

**Data 1:** The HIV-1 protease ( $\alpha$  and  $\beta$  chains) active site residues predicted by SiteHound-web:

Arg8, Leu23, Asp25, Thr26, Gly27, Ile50, Pro81, Val82, Ile84,  
Asp25', Thr26', Gly27', Ala28, Asp29, Asp30, Val32, Ile47, Gly48, Gly49,  
Ile50', Ile84'

**Data 2:** The Hsp90 N-terminal domain active site residues predicted by SiteHound-web:

Glu33, Leu34, Asn37, Ala38, Asp40, Ala41, Lys44, Ile77, Asp79, Ile82, Gly83,  
Met84, Glu88, Asn92, Leu93, Gly94, Thr95, Ile96, Ala97, Lys98, Ser99, Gly100,  
Lys102, Ala103, Gly118, Gln119, Phe120, Gly121, Val122, Gly123, Phe124,  
Tyr125, Val136, Thr171, Leu173, Arg380