Pharmacophore-Based Virtual Screening versus Docking-Based Virtual Screening: a Benchmark Comparison against Eight Targets

Zhi Chen[a], Honglin Li[b,*], Qijun Zhang[a], Xiaoguang Bao[a], Kunqian Yu[a], Xiaomin Luo[a], Weiliang Zhu[a], Hualiang Jiang[a,b,*]

[a] Center for Drug Discovery and Design, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China

[b] School of Pharmacy, East China University of Science and Technology, Shanghai 200237, China.

Supporting Methods

The detailed procedure of virtual screening using DOCK, GOLD and Glide is described below.

DOCK. For each target, the Connolly molecular surface was calculated with a probe radius of 1.4 Å. The residues around the co-crystallized ligand within 6 Å were defined as binding site with Sybyl 6.8. Within the binding site, spheres were generated therein with the SPHGEN program encoded in DOCK4.0. SPHGEN output the spheres in clusters. Cluster 1 was selected for each target. Compounds in the databases were docked into the binding site, allowing for ligand flexibility, using the grid-based energy scoring option for minimization after initial placement in the site. The box for the scoring grid was defined such that all spheres were enclosed with an extra 7 Å added in each dimension. Scoring grids for contact and energy scores were calculated with a grid spacing of 0.3 Å. The bump check was set such that atoms closer than 75 percent of the van der Waals radii of the respective atoms were rejected. A 6-12 Lennard-Jones van der Waals potential was used along with a Coulomb potential using a distance-dependent dielectric coefficient of 4 to simulate solvation effects. The energy cutoff was set to 10 Å. The radii used were those in the vdw.defn set.

GOLD. The GOLD docking program of Version 3.1 was evaluated in this study. The active site was defined to encompass any protein atom included in a 10-Å-radius sphere centered on the mass of the co-crystallized ligand in the complex structure. 7-8 times speedup setting was selected in virtual screening. All the variable parameters were set as the default values. The compounds were flexibly docked into the binding site of the targets, and the pose with the best GoldScore for each compound was adopted.

Glide. The Glide of version 4.0 was used to screen the databases. Target protein protonation was performed with the pprep script in Glide. Glide constructed a grid that defined the ligand-binding site search region, which was centered on the co-crystallized ligand in the complex structure. The grid was defined as an enclosing box with 15 Å in all three dimensions. The standard precision (SP) docking mode was selected. The ligands in the MOL2 format in databases were transformed into MAE format with mol2convert utility encoded in the program, and then docked flexibly into the binding site. The pose with the best GlideScore for each ligand was saved for further analysis.

Supporting Tables and Figures

Table S1. Chemical Structures of the Actives Used in This Study. The DrugBank Numbers Were Used to Name the Actives.
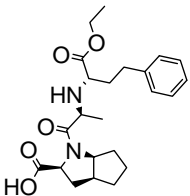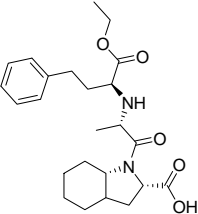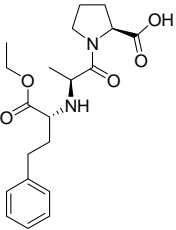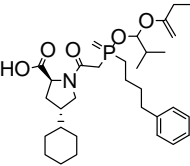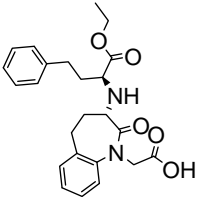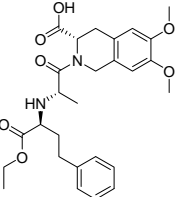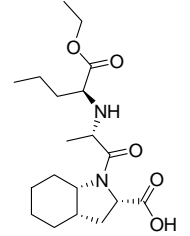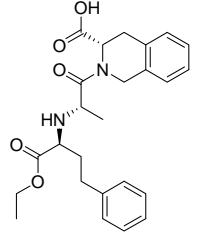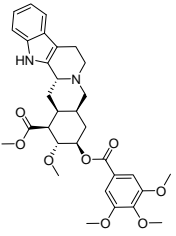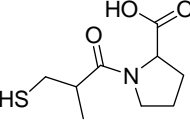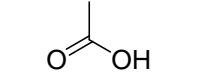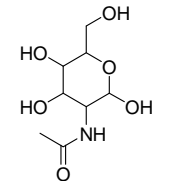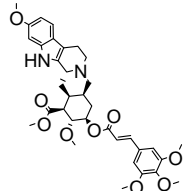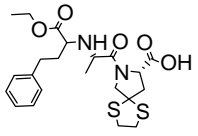
Compound	Structure	Compound	Structure	Compound	Structure

ACE

Compound	Structure	Compound	Structure	Compound	Structure
DB00178		DB00519		DB00584	
DB00492		DB00542		DB00691	
DB00790		DB00881		DB01089	
DB02032		DB04184		DB03740	
DB01180		DB01348			

AChE
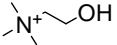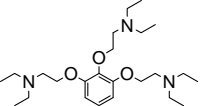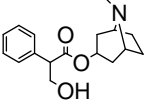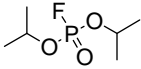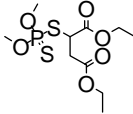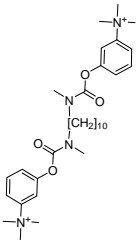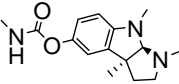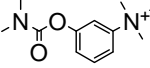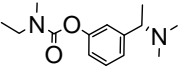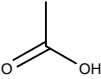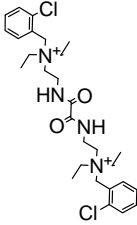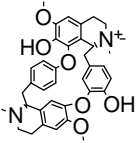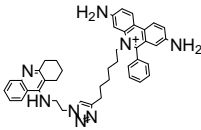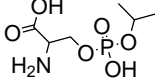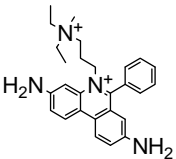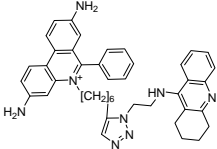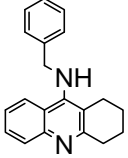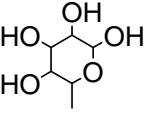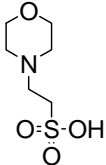
DB00122   DB00483   DB00545

DB00572   DB00677   DB00772

DB00944   DB00981   DB01400

DB00989   DB03166   DB01122

DB01199   DB01364   DB02226

DB01805   DB02166   DB03005
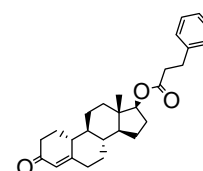
DB03672   DB03283   DB03814
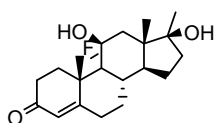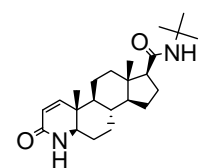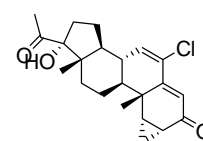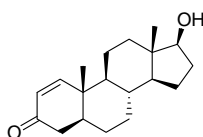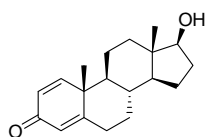
DB03740
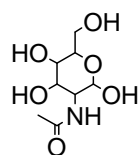
AR

DB00499
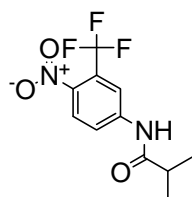
DB00621
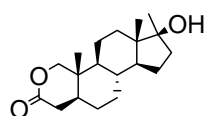
DB02901
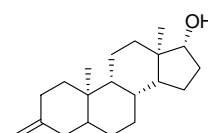
DB00665
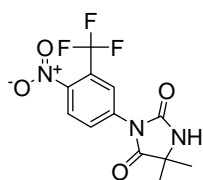
DB00858
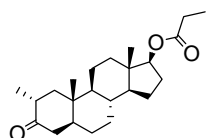
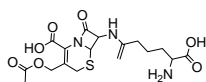DB00984

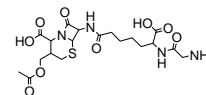DB01185

DB01216

DB04839

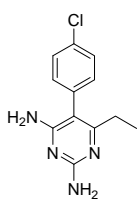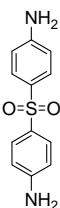DB01481

DB01541

DB02998
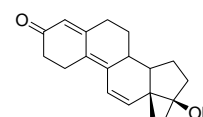
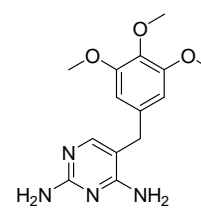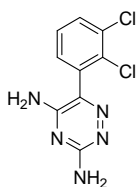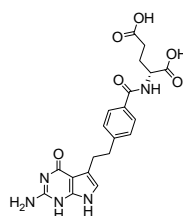DB04709
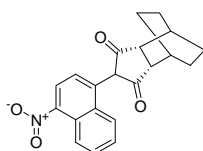
DB01564
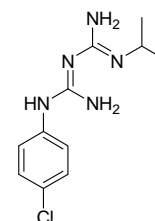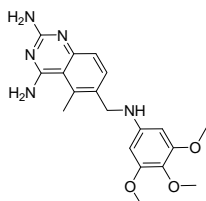
DB02932

DB02478

DacA

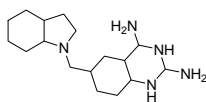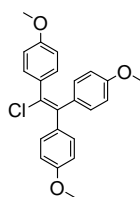DB00456
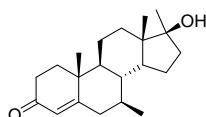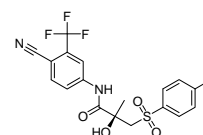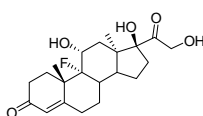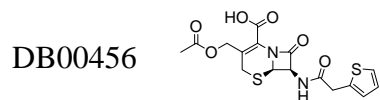
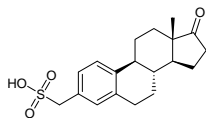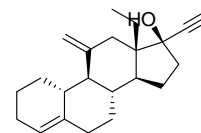DB03313

DB04488

DR

DB00205
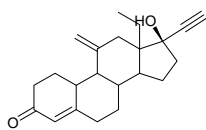
DB00250

DB00440

DB00555

DB00642

DB01131

DB01157

DB02559

ER

DB00269

DB00286

DB00304

DB00294

DB00351

DB00367

DB00396

DB00506

DB00539

DB00603

DB00655

DB00675

DB00823

DB00858

DB00860

DB00890

DB00947

DB00957

DB00977

DB01006

DB01065

DB01183

DB01357

DB01185

DB01196

DB01395

DB01406
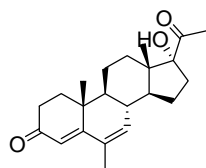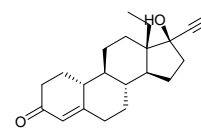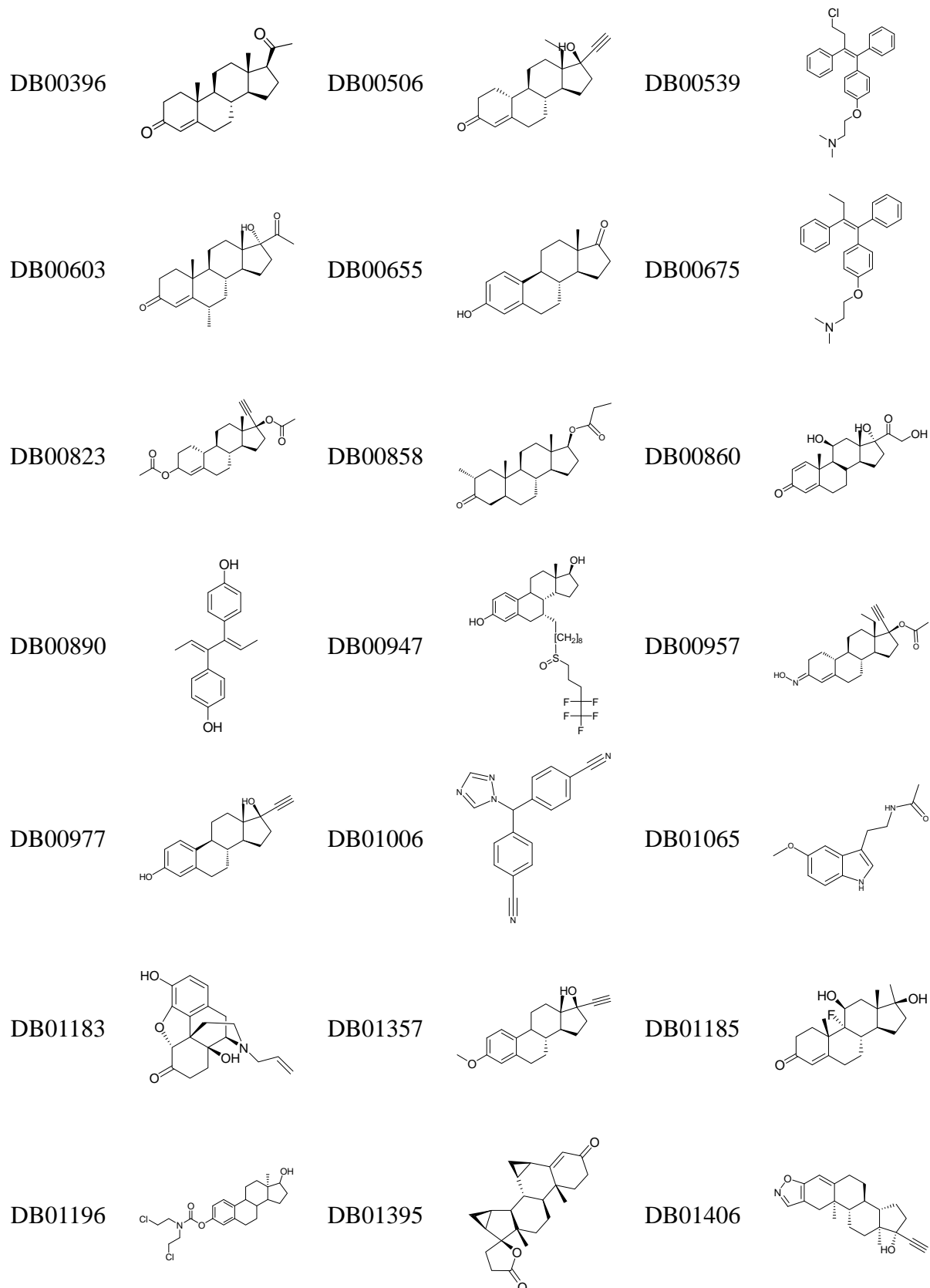
Table S1 (*continued*)

DB03742

DB03802

DB04573

DB04575

DB00882

HIV-pr

DB00220

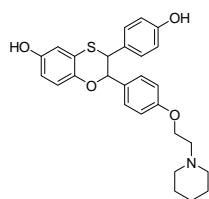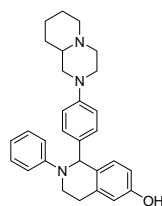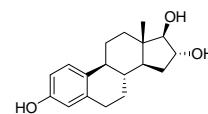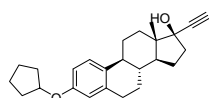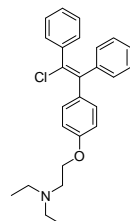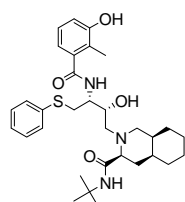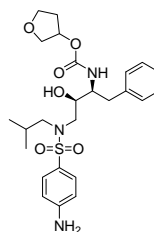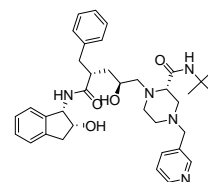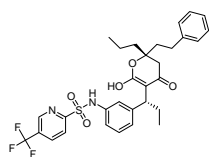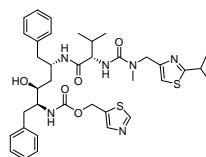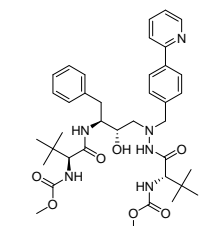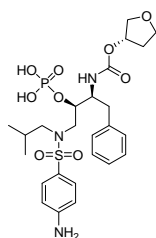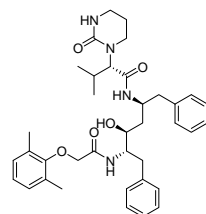DB00701

DB00224

DB00932

DB00503

DB01072
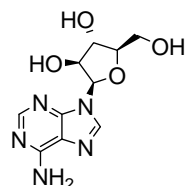
DB01319

DB01601

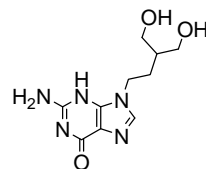DB02785
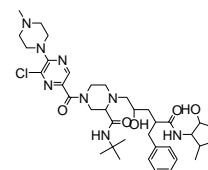
TK

DB00194

DB00299

DB00432

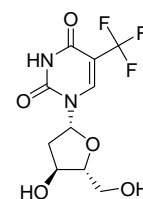**Table S1 (*continued*)**

DB00577
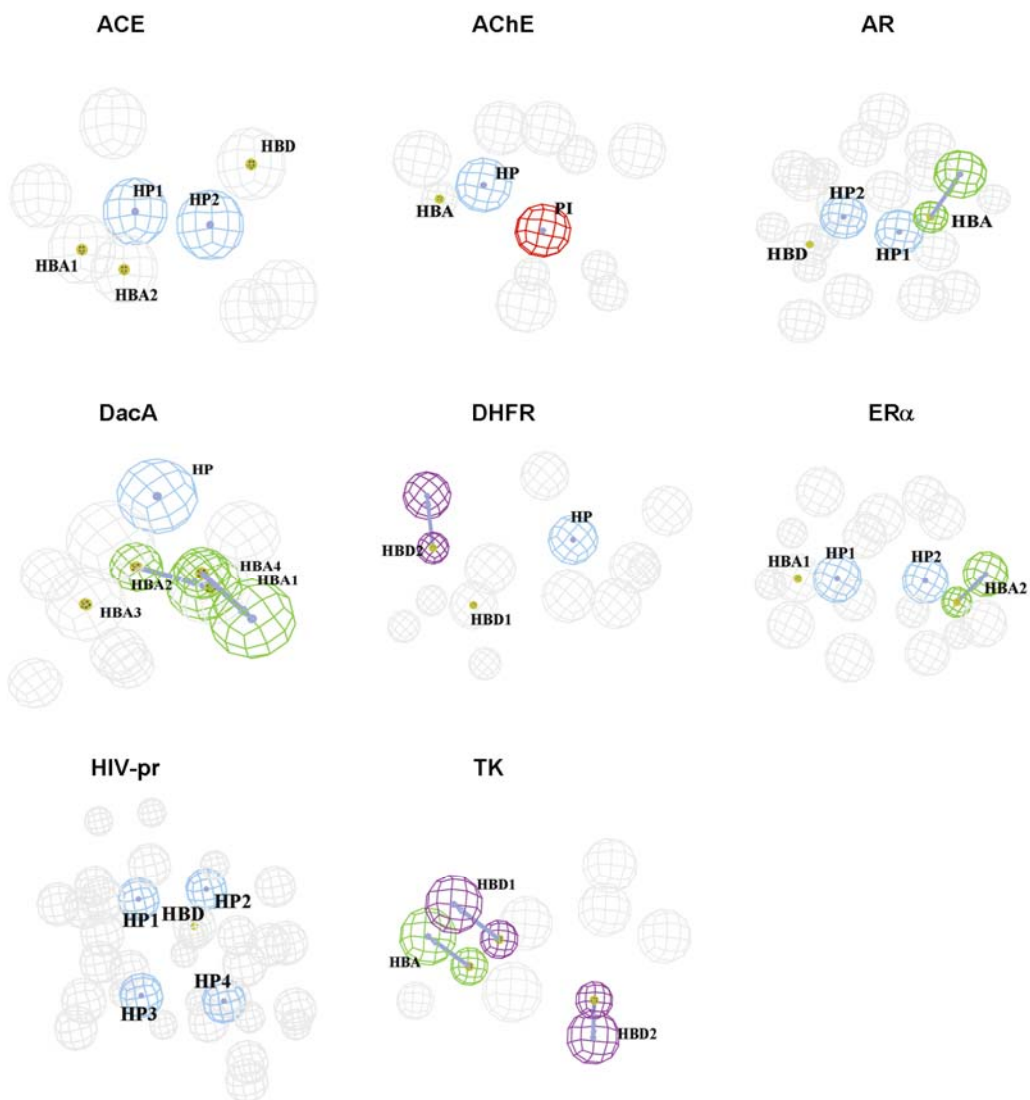
DB00787

DB01004

DB03280

DB03431

Figure S1. Refined pharmacophore models for the eight targets represented in Catalyst form: hydrogen bond acceptor (HBA), hydrogen bond donor (HBD), hydrophobic (HP), and excluded volume (gray sphere).
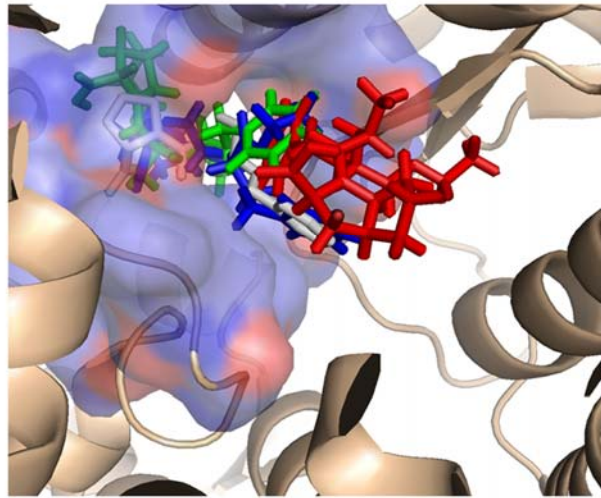
Figure S2. Actives mapping to the pharmacophore model and fitting to the binding pocket of ACE. Actives in the 5% of the highest ranks of the entire database produced by the pharmacophore-based method were used in the mapping and fitting. Pharmacophore model was represented in Catalyst form: hydrogen bond acceptor (HBA, green), hydrogen bond donor (HBD, purple), hydrophobic group (HP, cyan), and excluded volume (EV, gray). Binding conformations of the actives were isolated from the results of virtual screens. Actives in the docking models were represented in stick model. Ligand bound in the crystal structure was used as reference molecule (white), and actives docked by DOCK, GOLD and Glide were displayed in red, green, and blue colors, respectively.
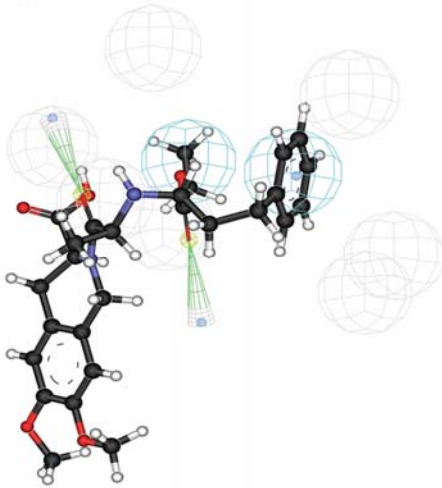
A
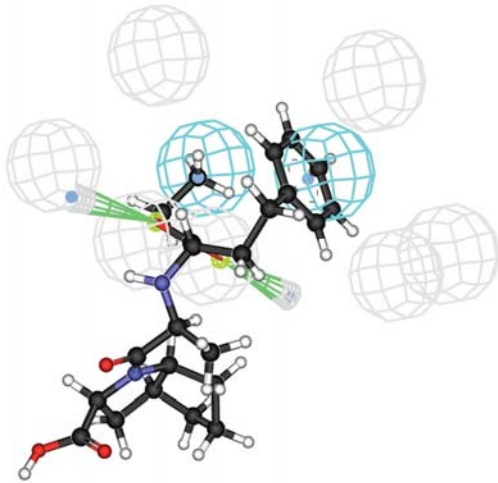
DB00178 Fit=3.38

B

C

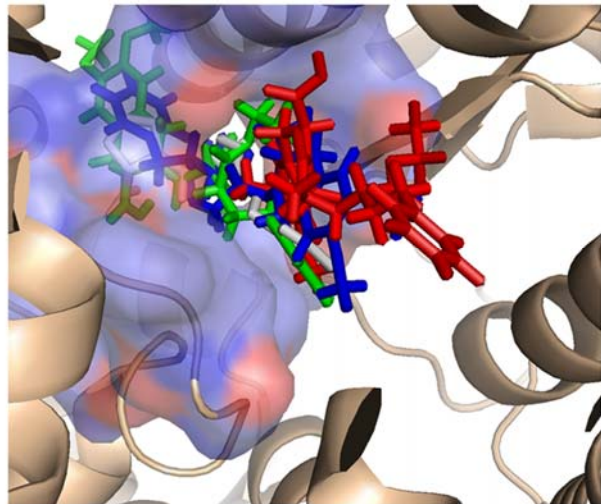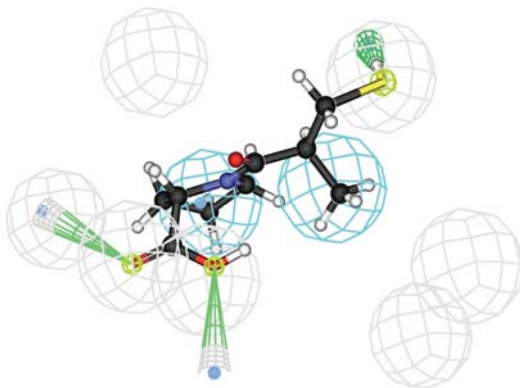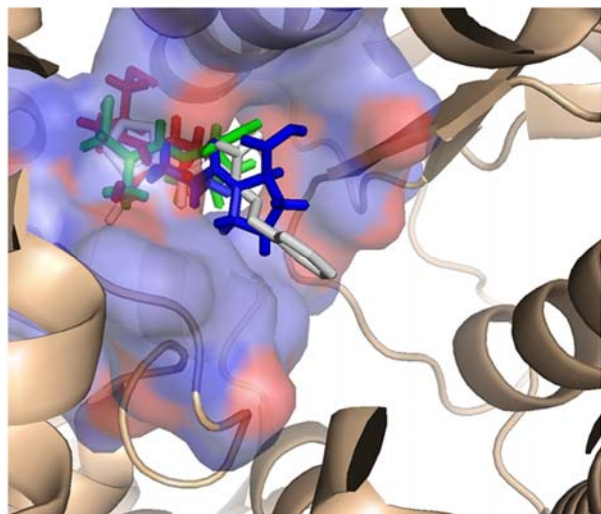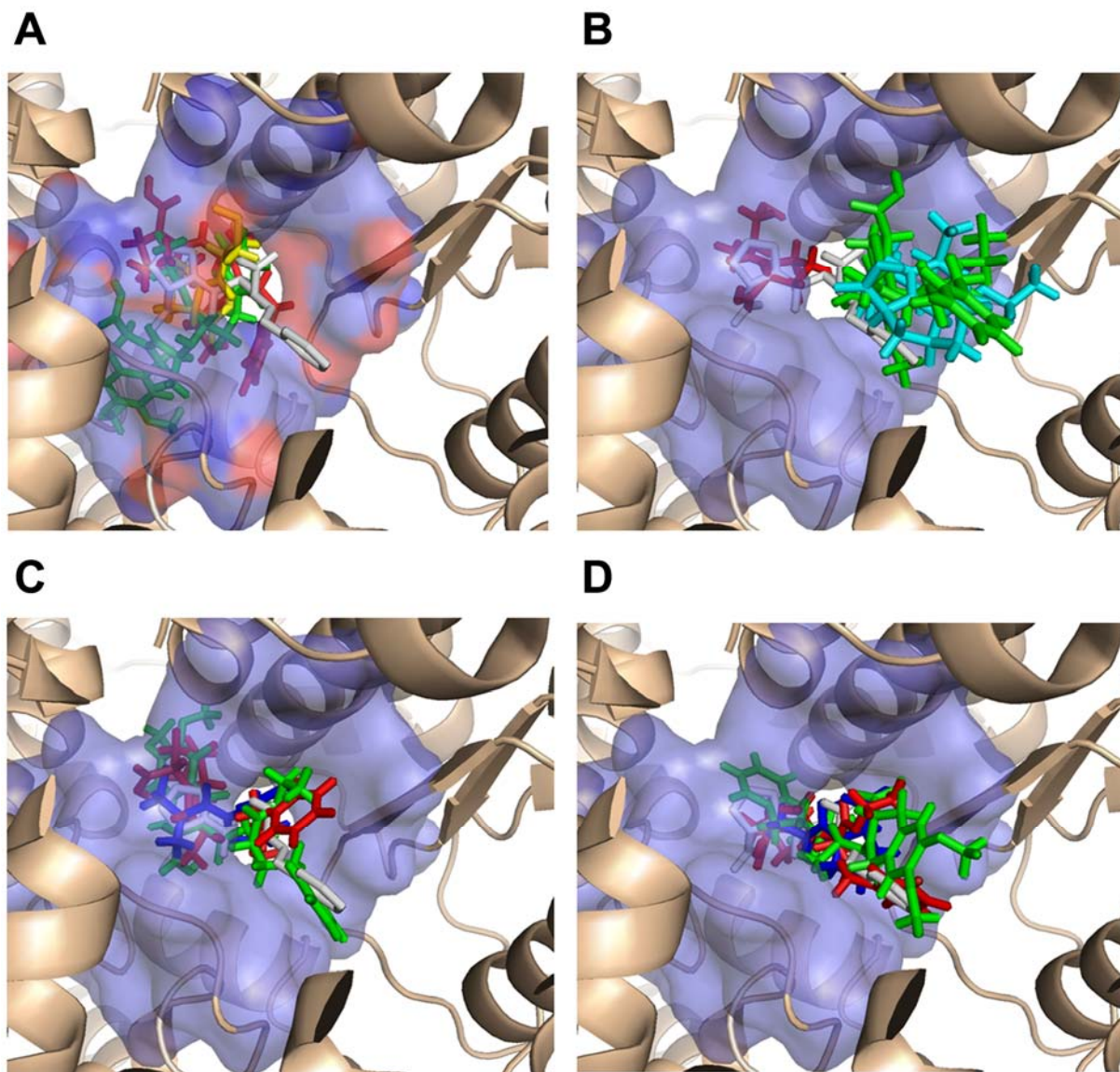DB00691 Fit=3.33

D

E

DB02032 Fit=3.79

F

Figure S3. (A) Structural superposition of ACE actives after pharmacophore mapping and binding pocket fitting. The conformation of the actives predicted by DOCK (B), GOLD (C), and Glide (D) are also displayed.

Figure S4. Actives mapping to the pharmacophore model and fitting to the binding pocket of AChE. The figure caption is the same as Figure S2.
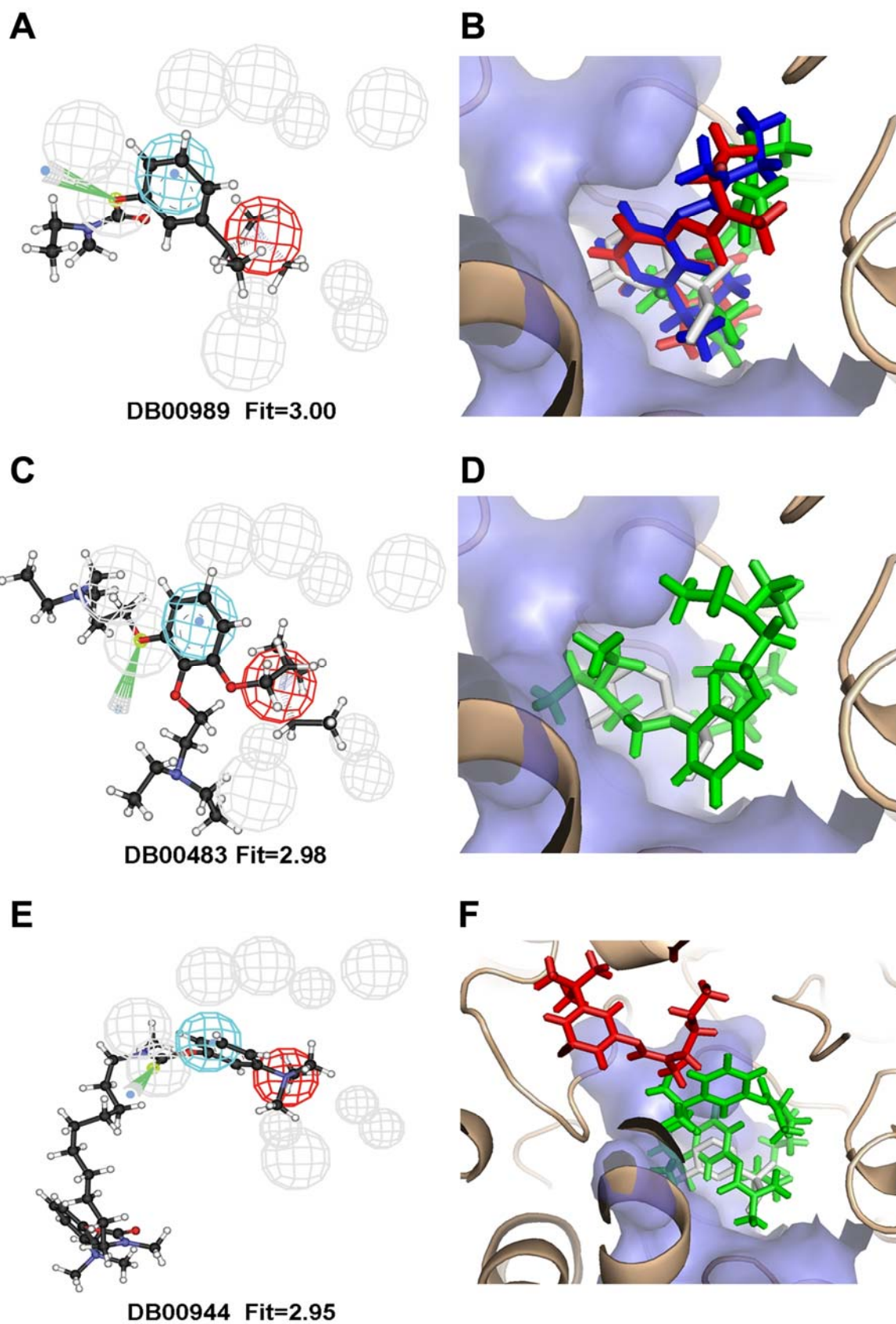

DB00989 Fit=3.00

DB00483 Fit=2.98

DB00944 Fit=2.95

G

DB01400 Fit=2.94

H

I

DB01805 Fit=2.84

J

K

DB00981 Fit=2.13

L

Figure S5. Structural superposition of AChE actives after pharmacophore mapping and binding pocket fitting. The figure caption is the same as Figure S3.

Figure S6. Actives mapping to the pharmacophore model and fitting to the binding pocket of AR. The figure caption is the same as Figure S2.
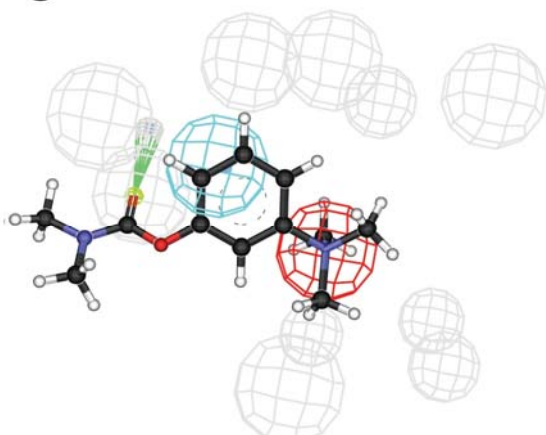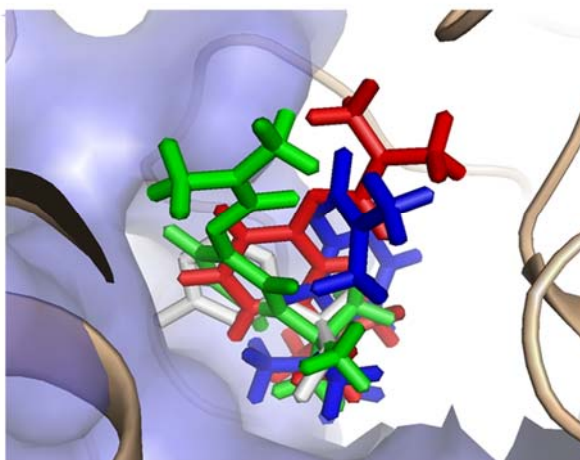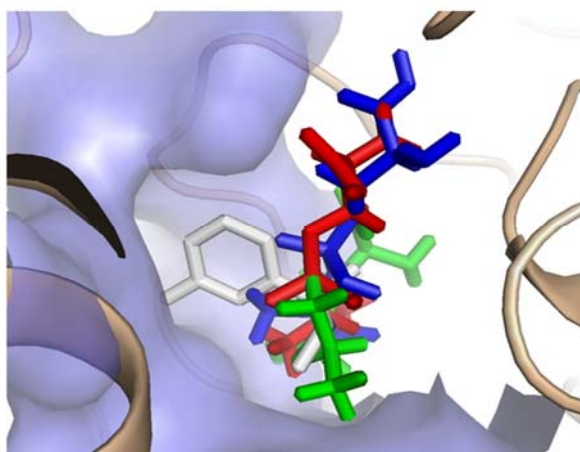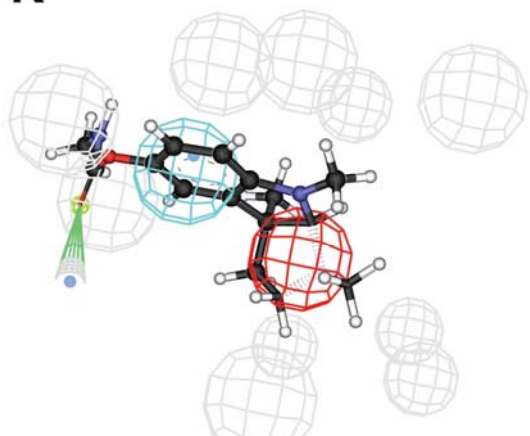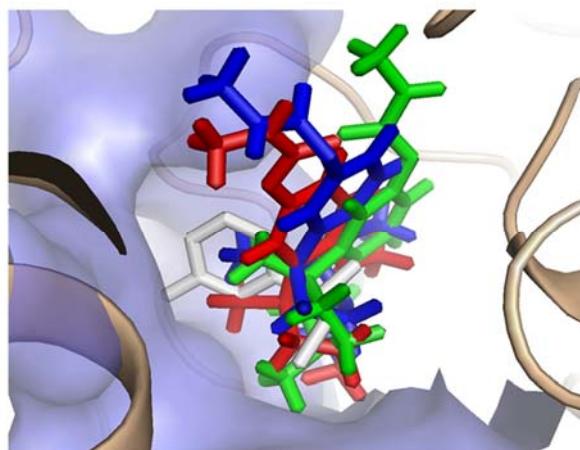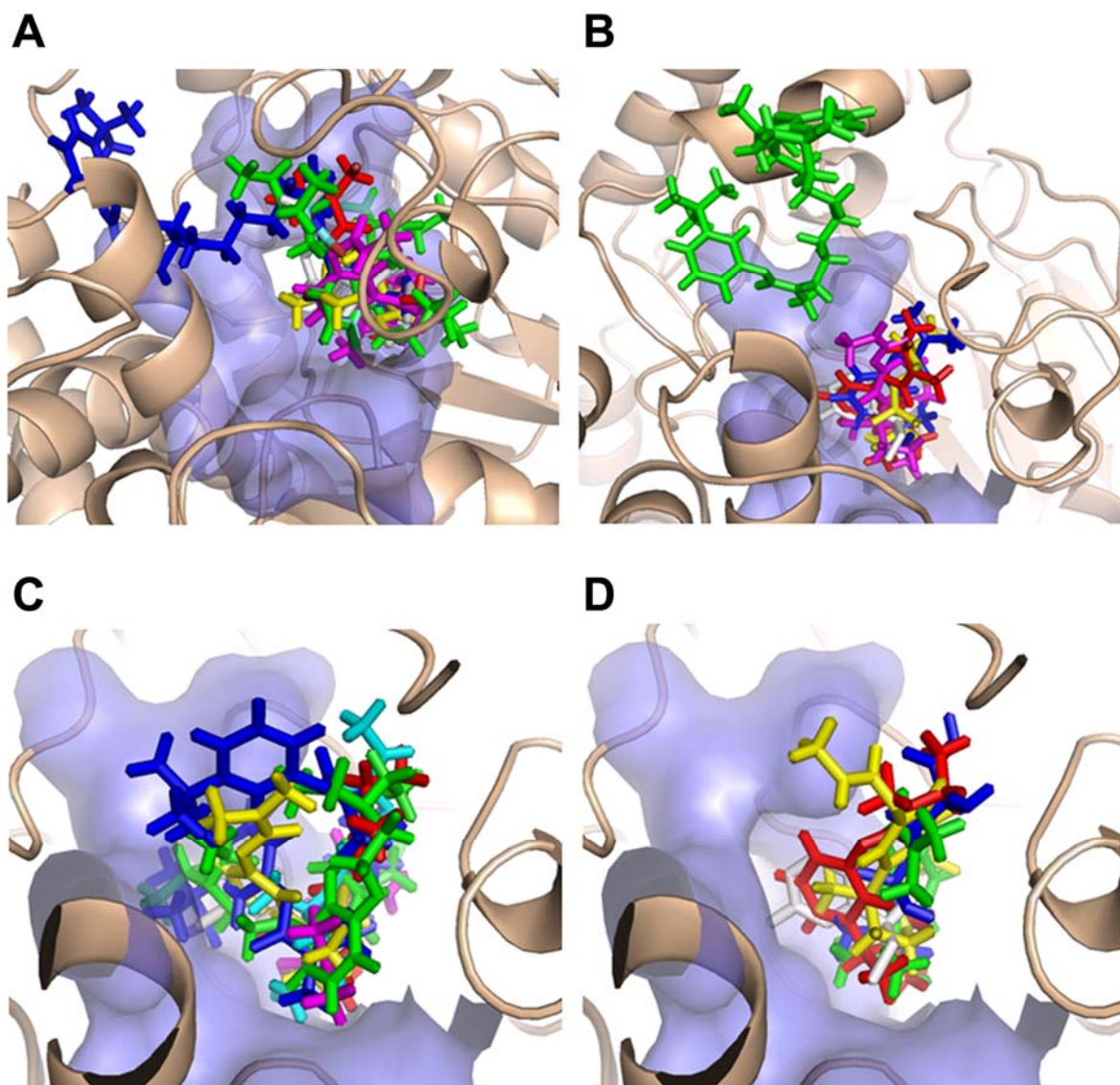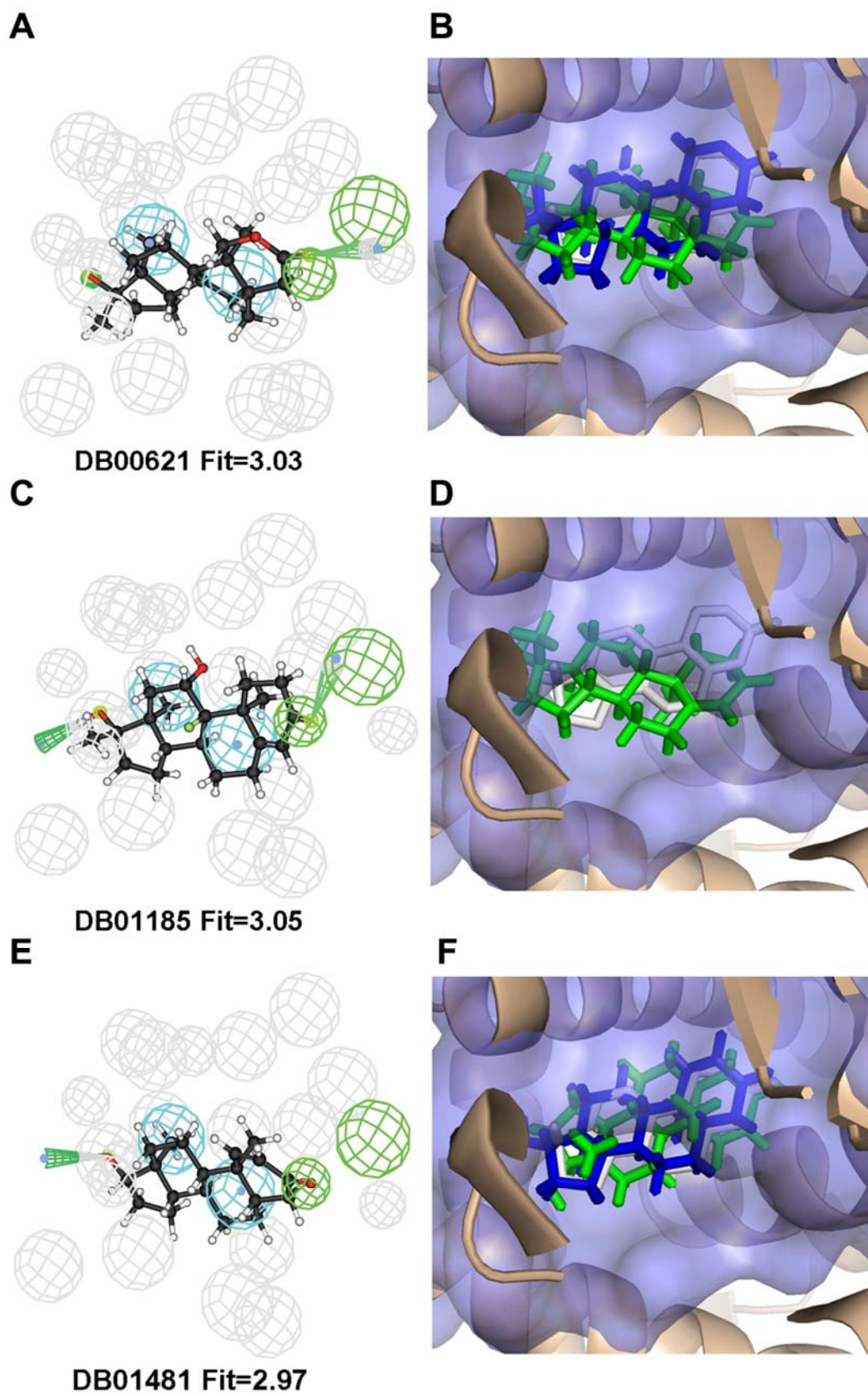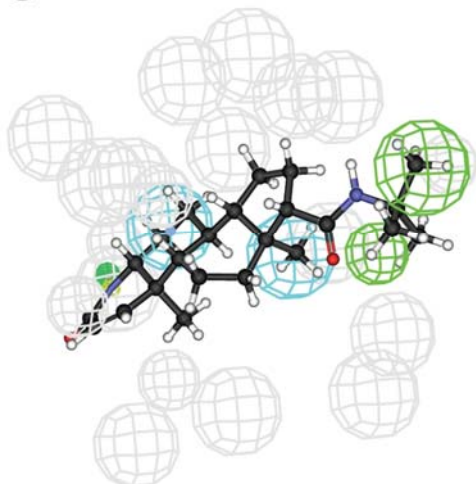

A

B

DB00621 Fit=3.03

C

D

DB01185 Fit=3.05

E

F

DB01481 Fit=2.97

G

DB01216 Fit=2.95

H

I

DB01564 Fit=2.99

J

K

DB02478 Fit=2.71

L

M

DB02998 Fit=3.45

N

O
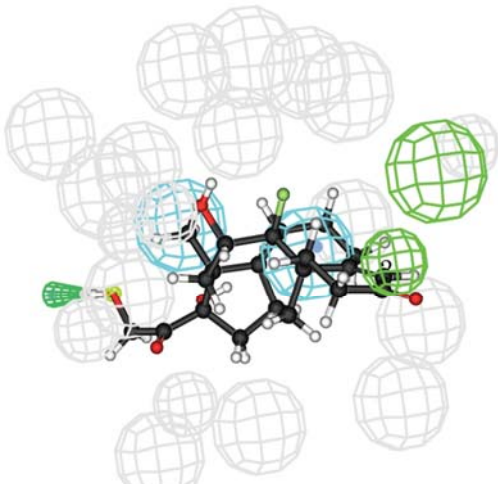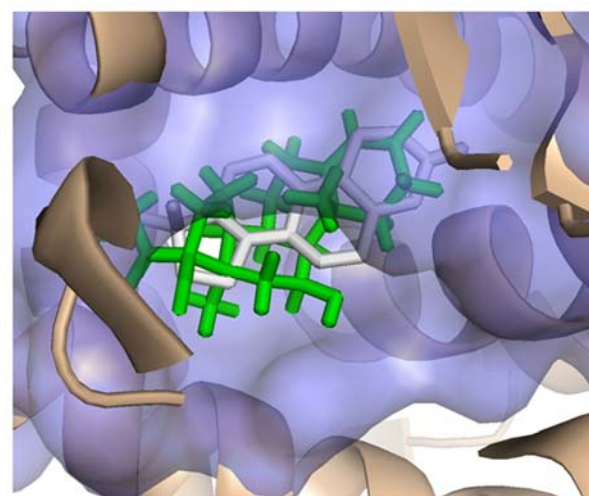
DB02901 Fit=3.06

P

Figure S7. Structural superposition of AR actives after pharmacophore mapping and binding pocket fitting. The figure caption is the same as Figure S3.

Figure S8. Actives mapping to the pharmacophore model and fitting to the binding pocket of DHFR. The figure caption is the same as Figure S2.
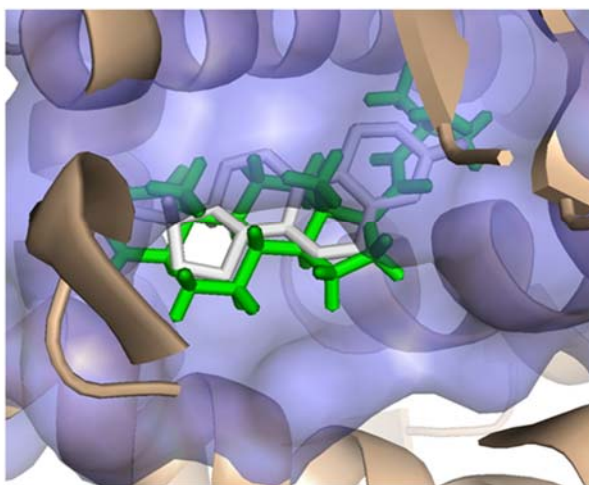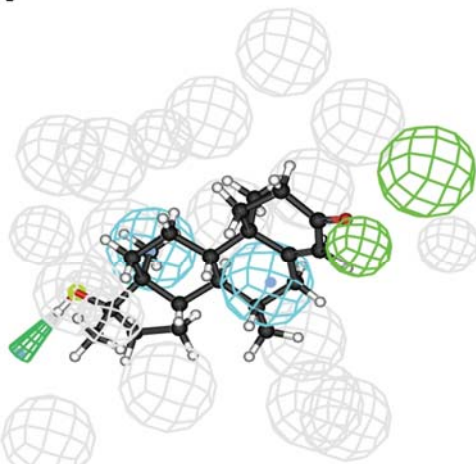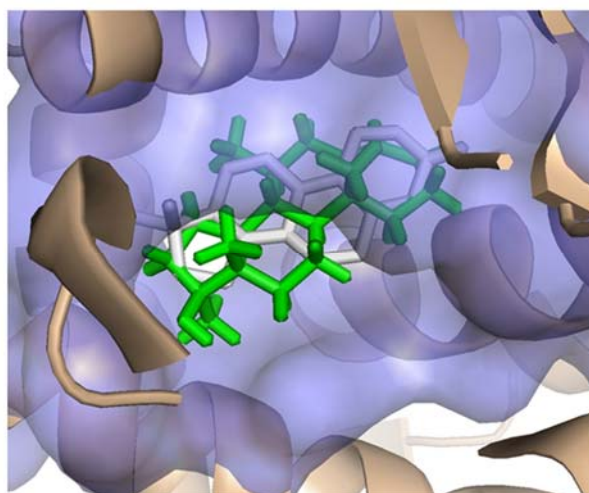

A
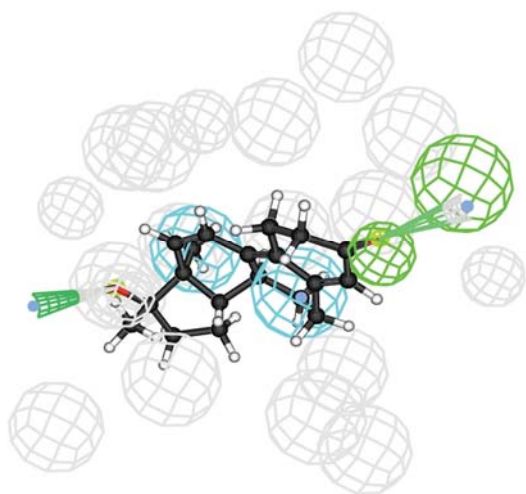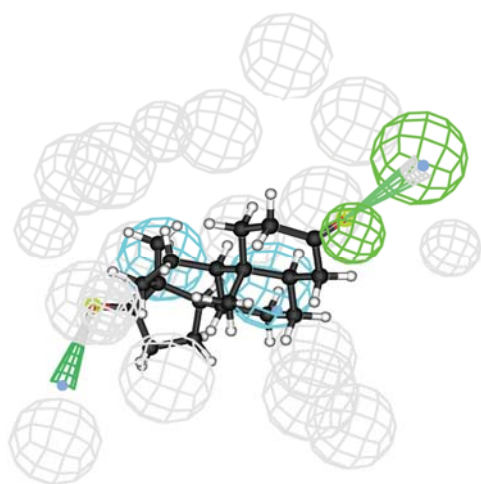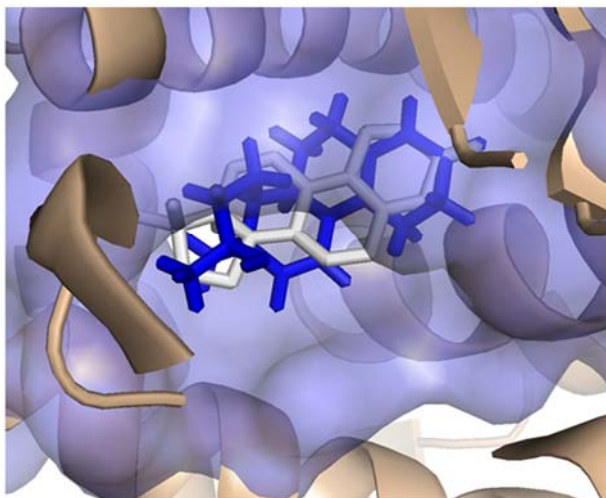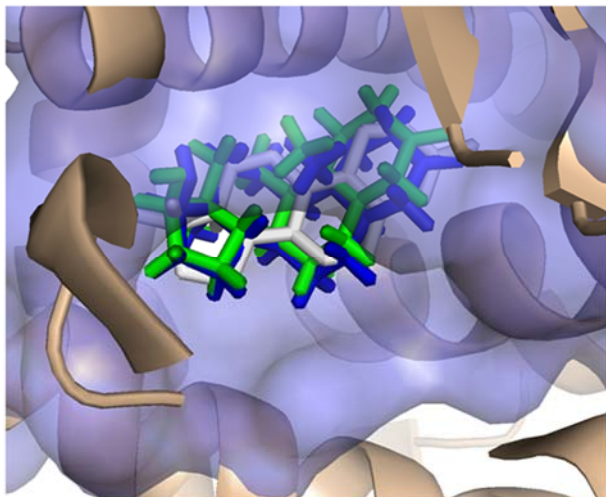
DB02559 Fit=2.71

B

C

DB00440 Fit=2.96

D

E

DB01157 Fit=2.90
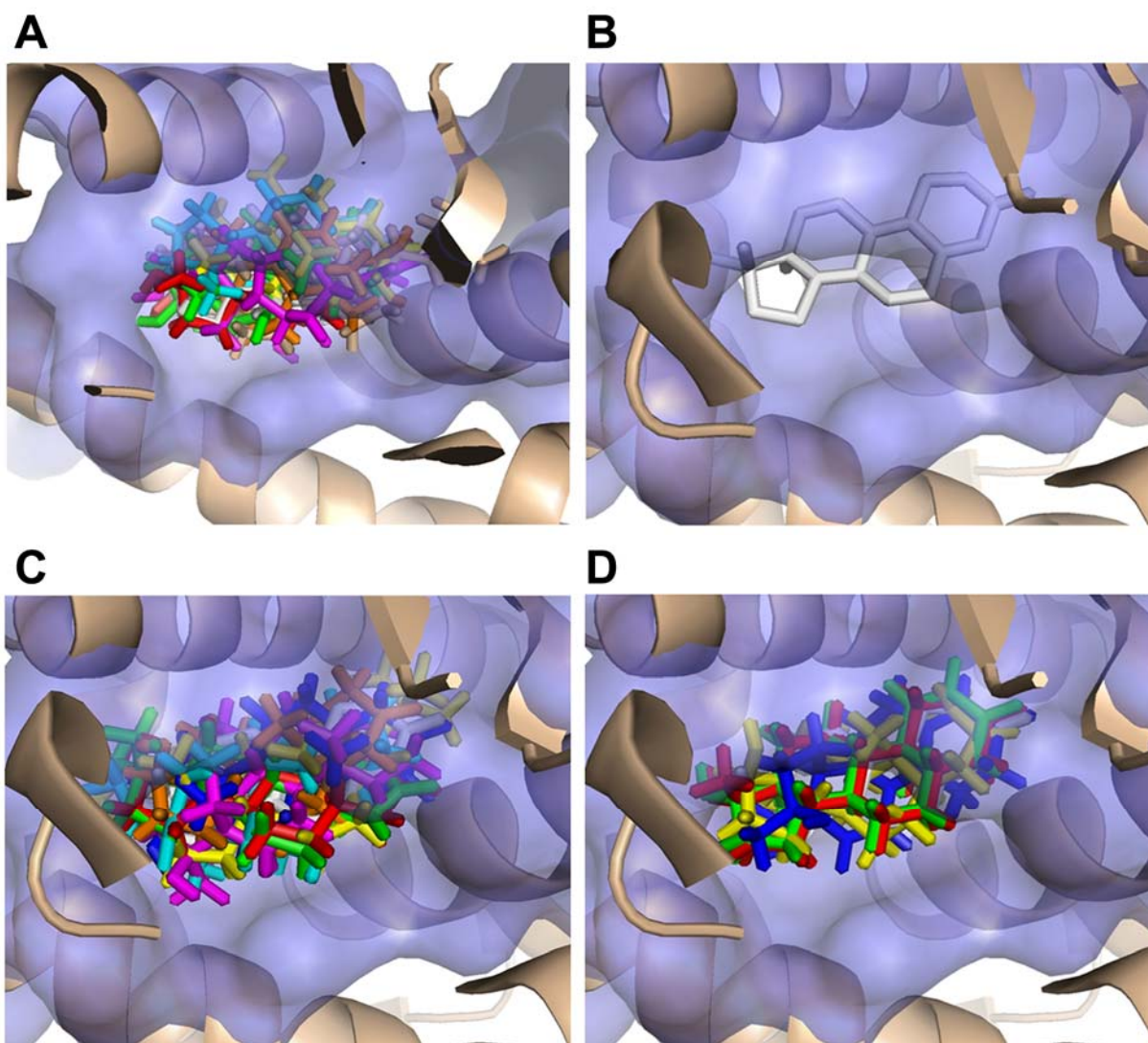
F

Figure S8 *(continued)*

G

DB00642 Fit=2.76

H

Figure S9. Structural superposition of DHFR actives after pharmacophore mapping and binding pocket fitting. The figure caption is the same as Figure S3.

Figure S10. Actives mapping to the pharmacophore model and fitting to the binding pocket of DacA. The figure caption is the same as Figure S2.
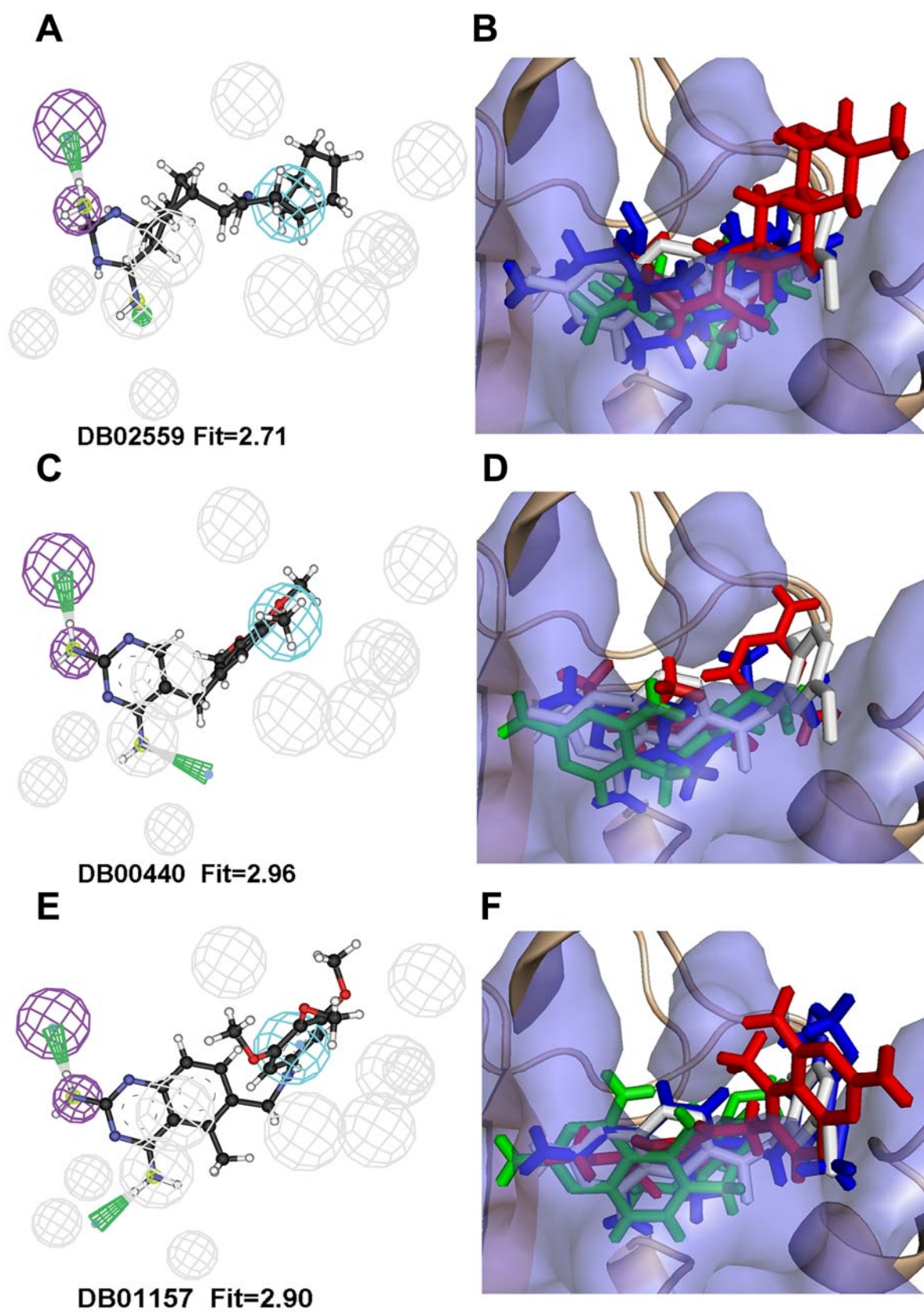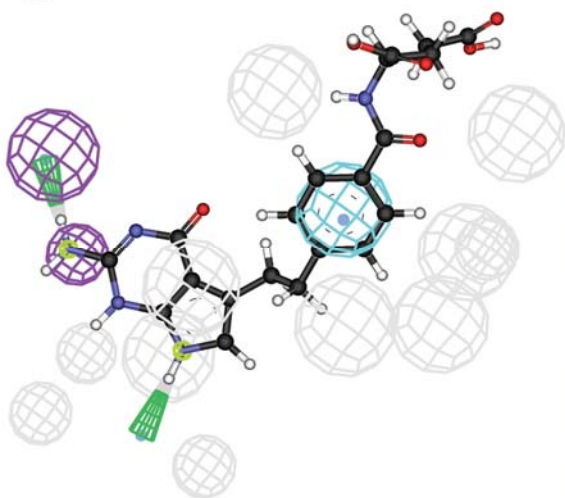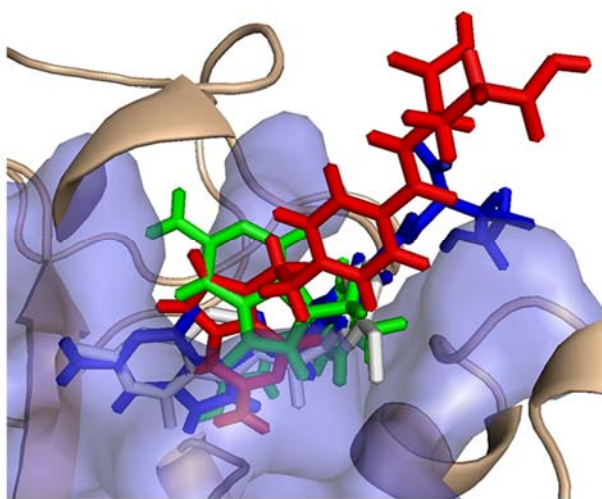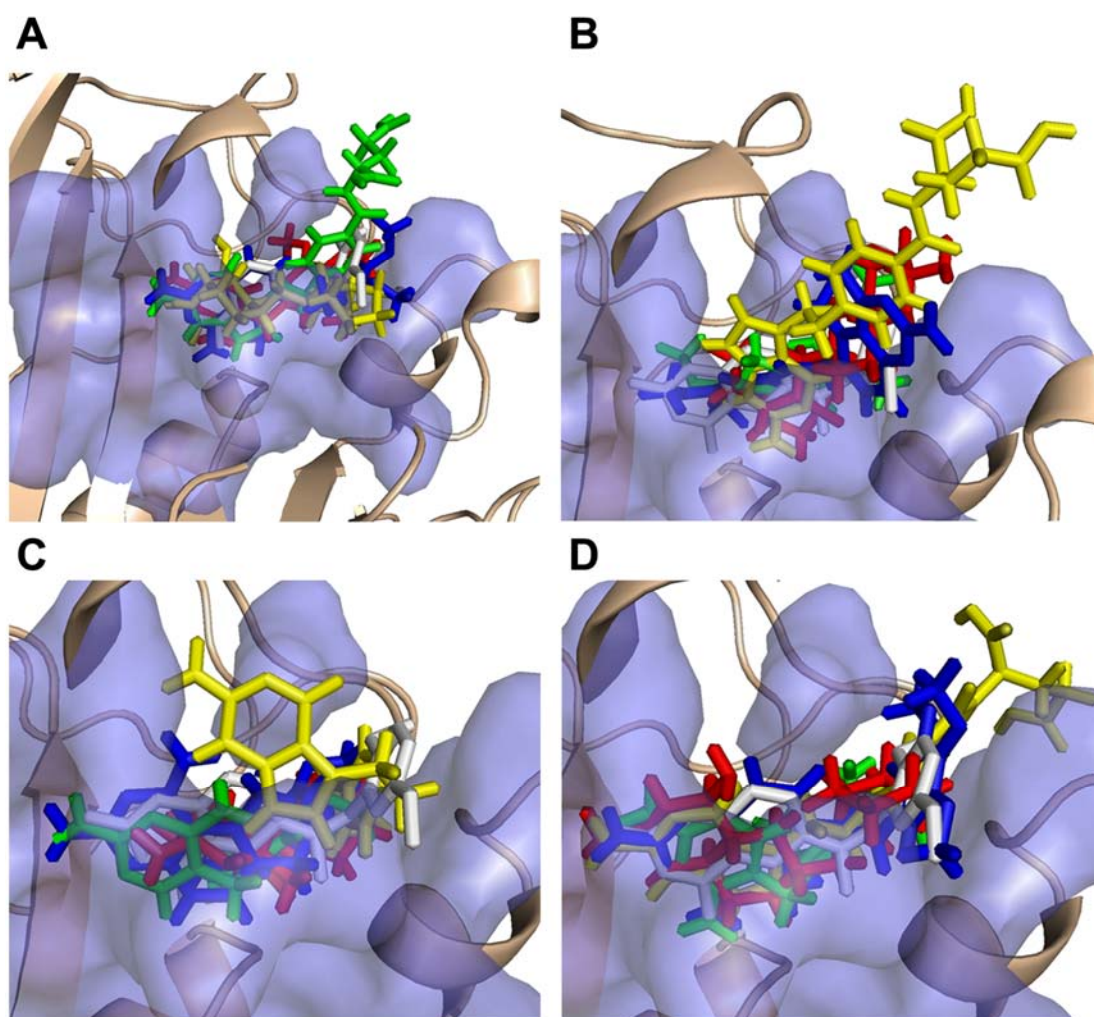

A

DB00456 Fit=4.30

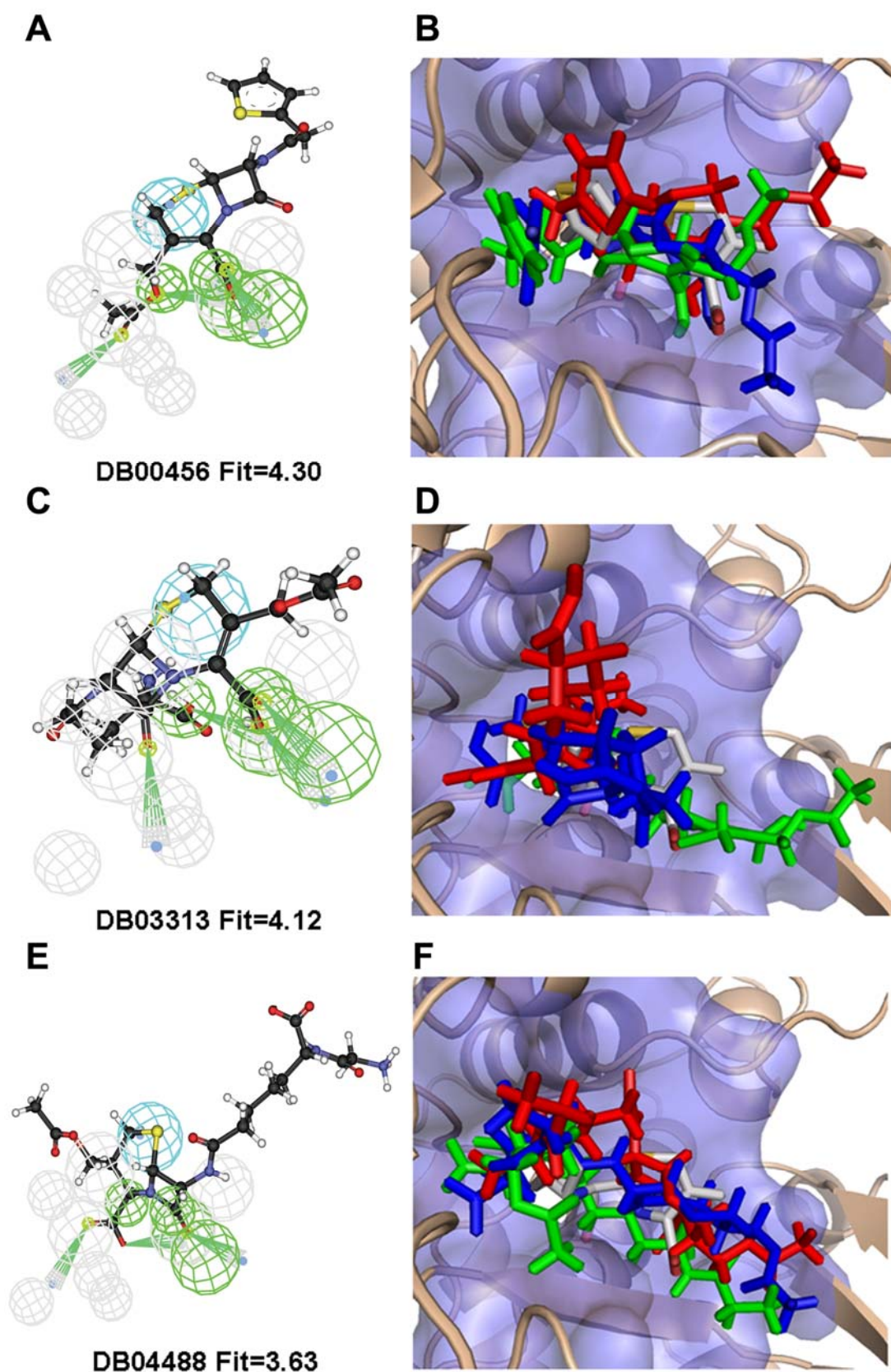B

C

DB03313 Fit=4.12

D

E

DB04488 Fit=3.63

F

Figure S11. Structural superposition of DacA actives after pharmacophore mapping and binding pocket fitting. The figure caption is the same as Figure S3.
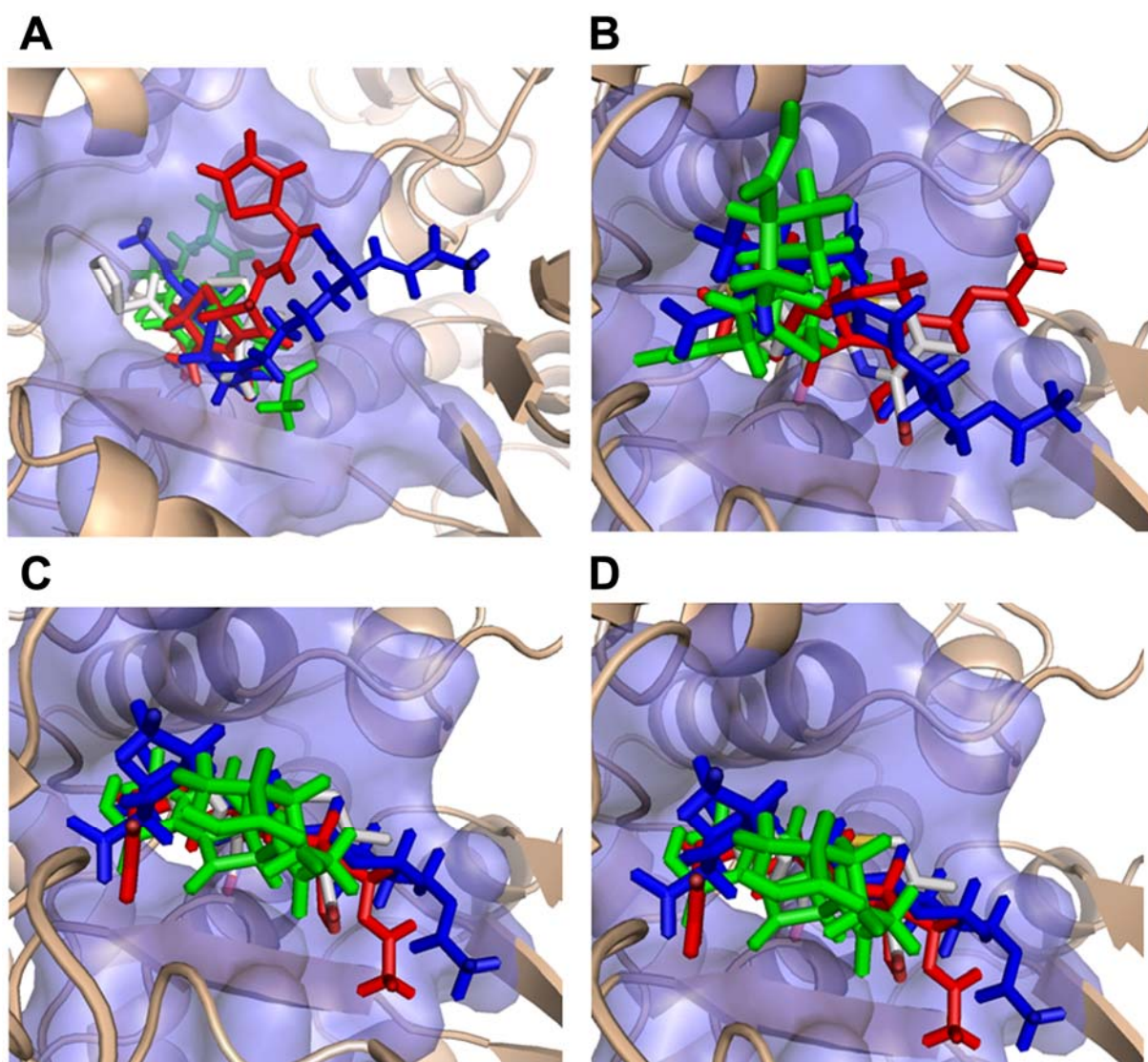
Figure S12. Actives mapping to the pharmacophore model and fitting to the binding pocket of HIV-pr. The figure caption is the same as Figure S2.
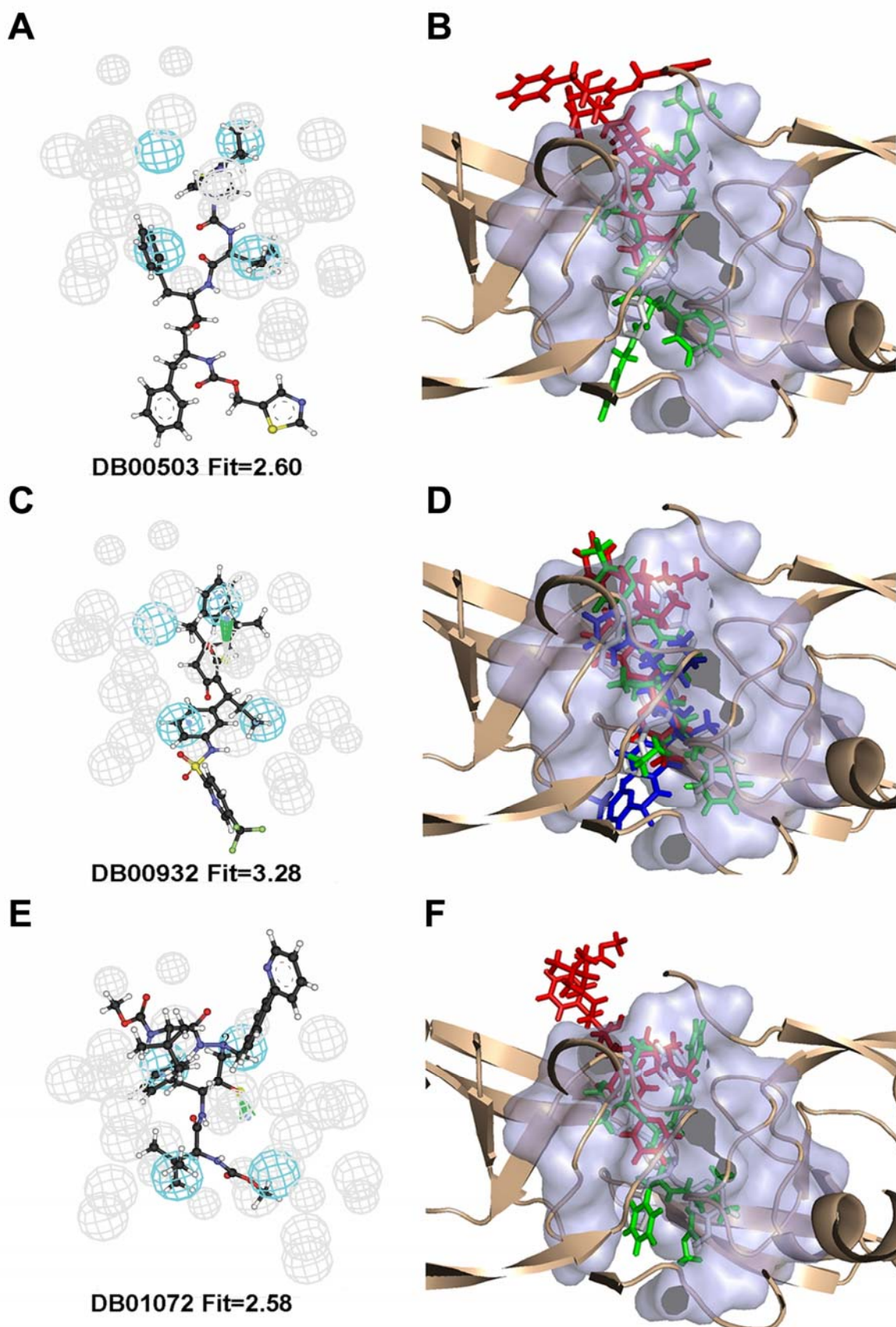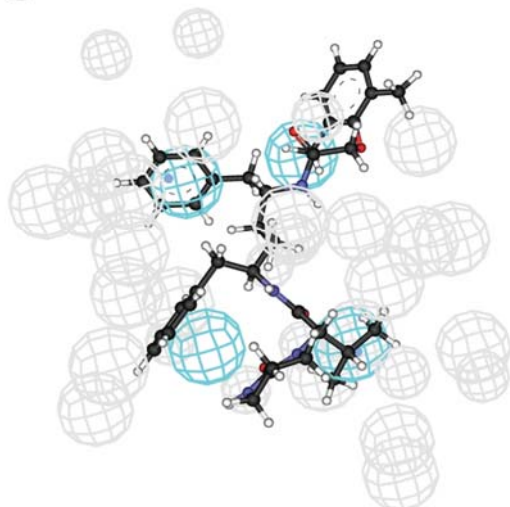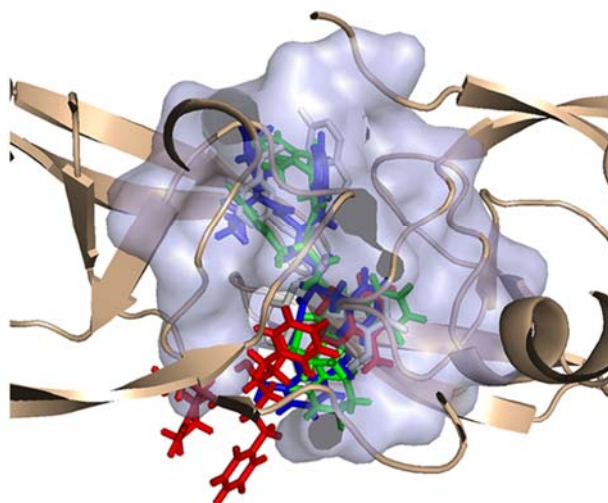

DB00503 Fit=2.60

DB00932 Fit=3.28

DB01072 Fit=2.58

G

DB01601 Fit=2.59

H
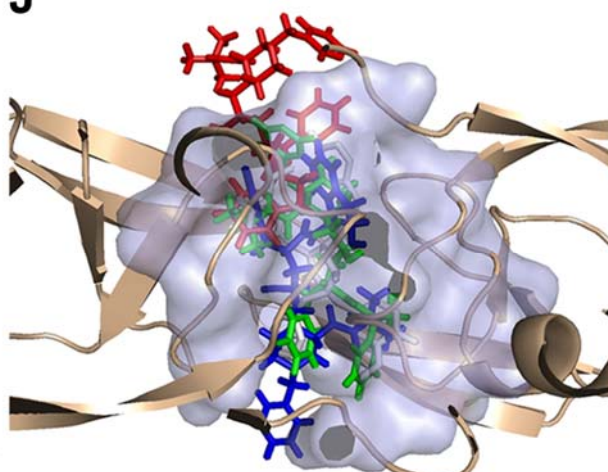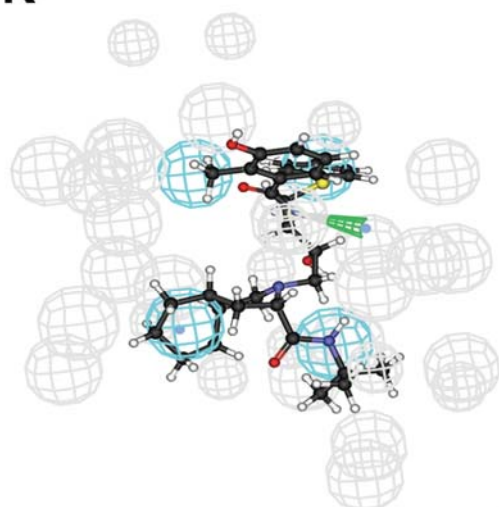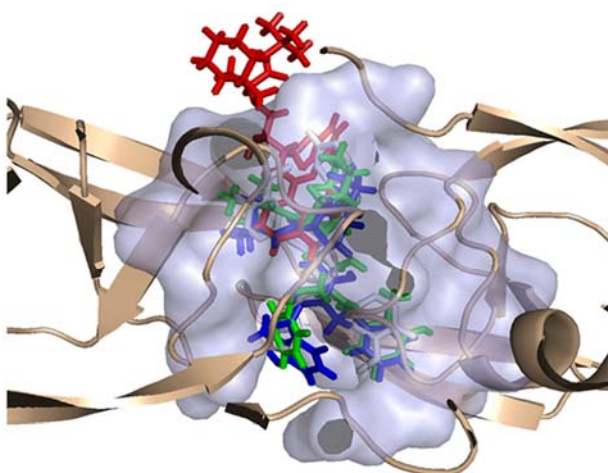
I
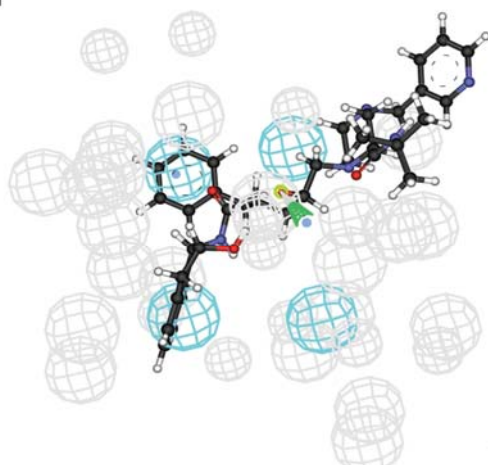
DB00224 Fit=2.40

J

K

DB00220 Fit=2.87

L

Figure S13. Structural superposition of HIV-pr actives after pharmacophore mapping and binding pocket fitting. The figure caption is the same as Figure S3.