Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supporting Information

for

Impact of Electronic Polarizability on Protein-Functional Group Interactions

Himanshu Goel[a], Wenbo Yu[a], Vincent D. Ustach[a], Asaminew H. Aytenfisu,[a] Delin Sun[a] and Alexander D. MacKerell Jr.[a*]

[a] Computer Aided Drug Design Center, Department of Pharmaceutical Sciences, University of Maryland School of Pharmacy, 20 Penn St. Baltimore, Maryland 21201, United States

Table S1: Overlap coefficients for the different fragments in consideration to the 1-5 and 6-10 independent runs with Drude and additive model for each protein target.

	P38 kinase		FactorXa		HDM2		TRMD		TYK2	
FragMap	Drude	Additive	Drude	Additive	Drude	Additive	Drude	Additive	Drude	Additive
meoo	0.82	0.84	0.83	0.84	0.83	0.85	0.84	0.84	0.82	0.84
imin	0.80	0.84	0.80	0.84	0.82	0.85	0.83	0.84	0.80	0.84
iminh	0.81	0.84	0.81	0.84	0.83	0.85	0.84	0.84	0.80	0.84
foro	0.82	0.84	0.82	0.83	0.83	0.85	0.84	0.85	0.81	0.84
forn	0.82	0.84	0.82	0.83	0.83	0.85	0.84	0.85	0.82	0.84
mamn	0.82	0.84	0.81	0.82	0.82	0.85	0.82	0.85	0.81	0.84
aceo	0.84	0.85	0.85	0.86	0.86	0.88	0.86	0.86	0.84	0.83
benc	0.85	0.89	0.86	0.90	0.85	0.90	0.86	0.89	0.83	0.88
prpc	0.83	0.87	0.85	0.87	0.84	0.88	0.82	0.86	0.79	0.85
aalo	0.81	0.84	0.81	0.84	0.83	0.85	0.82	0.85	0.80	0.84
tipo	0.97	0.98	0.97	0.98	0.98	0.98	0.98	0.99	0.97	0.99
aalc	0.81	0.85	0.81	0.84	0.83	0.85	0.82	0.85	0.80	0.84
gehc	0.86	0.89	0.86	0.90	0.87	0.90	0.89	0.90	0.85	0.91
acec	0.80	0.81	0.80	0.82	0.82	0.84	0.81	0.82	0.80	0.78

Meoo: methanol O, imin: imidazole acceptor N, imih, imidazole donor N, foro: formamide O, forn, formamide N, mamn: methylammonium N, aceo, acetate O, benc, benzene Cs, prpc, propane Cs, aalo, Acetaldehyde O, tipo, Tip3P O, aalc, Acetaldehyde carbonyl C, gehc: imidazole Cs and acec: Acetate carbonyl C.

Table S2: Overlap coefficient for the different FragMaps for the entire simulation box between the Drude and additive simulations for each protein target.

	P38	FactorXa	HDM2	TRMD	TYK2
FragMap					
Methanol oxygen (meoo)	0.84	0.85	0.86	0.86	0.89
Imidazole acceptor nitrogen (imin)	0.82	0.83	0.84	0.83	0.88
Imidazole hydrogen on nitrogen (iminh)	0.83	0.83	0.84	0.83	0.88
Formamide oxygen (foro)	0.83	0.83	0.85	0.85	0.89
Formamide nitrogen (forn)	0.84	0.83	0.85	0.85	0.89
Methylammonium nitrogen (mamn)	0.75	0.74	0.79	0.77	0.81
Acetate oxygen (aceo)	0.78	0.75	0.79	0.80	0.78
Benzene carbon (benc)	0.79	0.82	0.80	0.83	0.82
Propane carbon (prpc)	0.76	0.80	0.77	0.79	0.82
Acetaldehyde oxygen (aalo)	0.80	0.82	0.81	0.82	0.86
Water (tipo)	0.95	0.95	0.96	0.96	0.97
Acetaldehyde carbon (aalc)	0.80	0.82	0.82	0.81	0.86
Imidazole carbon (gehc)	0.86	0.87	0.87	0.87	0.92
Acetate carbon (acec)	0.75	0.71	0.75	0.77	0.76

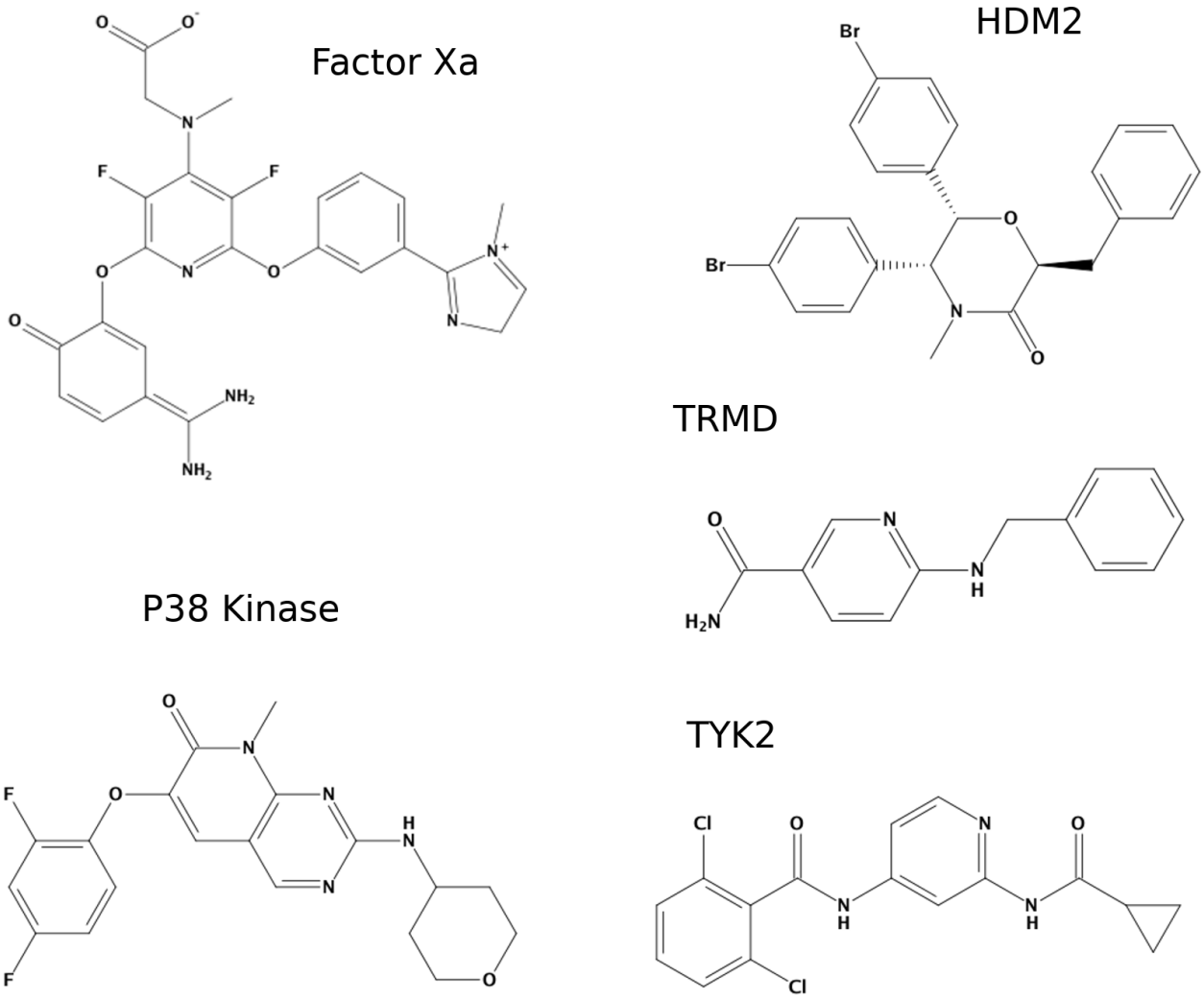

Figure S1: 2D structures for all the ligands present in the crystal structures used in the present study.

HDM2 and TRMD FragMap Analysis

HDM2: The HDM2 FragMaps along with its crystallographic structure are shown in Figure S1A and S1B for additive and Drude simulations, respectively. With the apolar FragMaps, the Drude and additive simulations depict four regions for apolar groups and they match satisfactorily to each other. The ligand structure contains three apolar aromatic rings as shown by arrows 1, 2, and 3 in Figure S1B and both set of FragMaps capture this feature of the crystallographic ligand. There are no FragMaps for the positive or negative charged groups in both sets of figures at the presented contour levels. Hydrogen bond acceptor or donor (GENA/GEND) FragMaps, as shown by arrows 4 and 5, are adjacent to the bromine atoms of the ligand. The Drude simulations show the presence of hydrogen bond donor maps adjacent to one of the bromine atoms as indicated by arrow 4. The presence of the donor FragMap may indicate the potential of this class of functional groups to participate in a favorable interaction with the bromine perpendicular to the C-Br bond as recently reported based on quantum mechanical calculations and a survey of the protein database.[1]

TRMD: The TRMD FragMaps are shown in Figure S1C and S1D for the additive and Drude simulations, respectively, along with the crystallographic structure of the protein and ligand. The apolar groups perfectly capture both the aromatic rings of the ligand as shown by the arrows 1 and 2 in Figure S1D. In addition, hydrogen bond acceptor (red) and donor (blue) FragMaps are observed on both the plots as shown by arrows 3 and 4 in Figure S1D. These recapitulate the positions of the corresponding functional groups on the ligand. Moreover, on the left side of the figure (arrow 5 in Figure S1C) donor and acceptor FragMaps are present which do not correlate with functional groups on the ligand. However, TRMD ligands are known that have polar groups, including hydroxyls and amines, in this region corresponding to the polar FragMaps.[2]

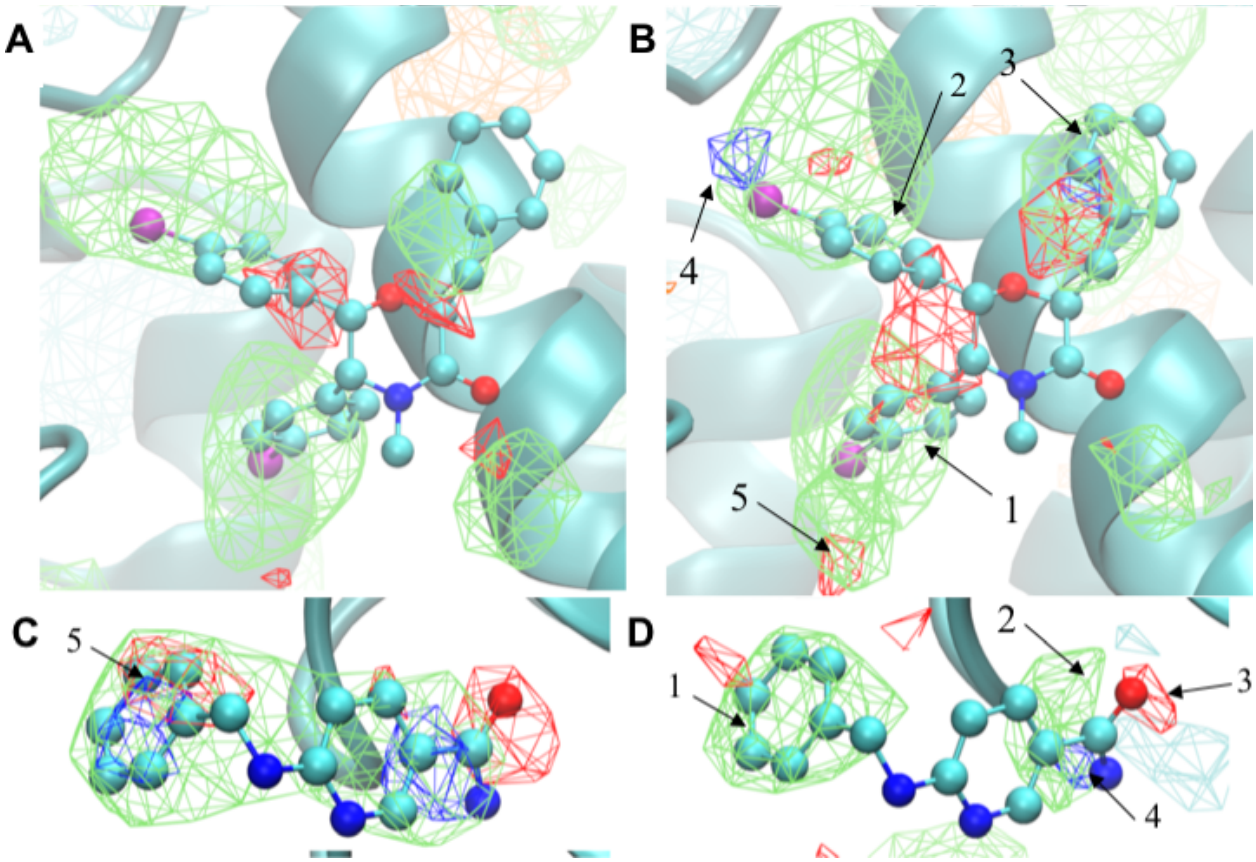

Figure S2: FragMaps overlaid on the proteins HDM2 and TRMD. Cartoon representation are shown based on the crystal conformations (PDB 4JV7 and PDB 4YPW) with portions of the protein occluding the binding pocket view omitted. The ligands from the respective crystal structures are shown in CPK representation with atom type color. Figures A and C show the FragMaps for HDM2 and TRMD obtained from the additive simulations while figures B and D show the FragMaps for HDM2 and TRMD, respectively obtained from the Drude simulations. The FragMaps color are GENN (green), GENA (red), GEND (blue), MAMN (cyan), and ACEO (orange). All FragMaps isocontour surfaces are displayed at a cutoff of -1.2 kcal/mol. 2D images of the ligands are shown in Figure S1.

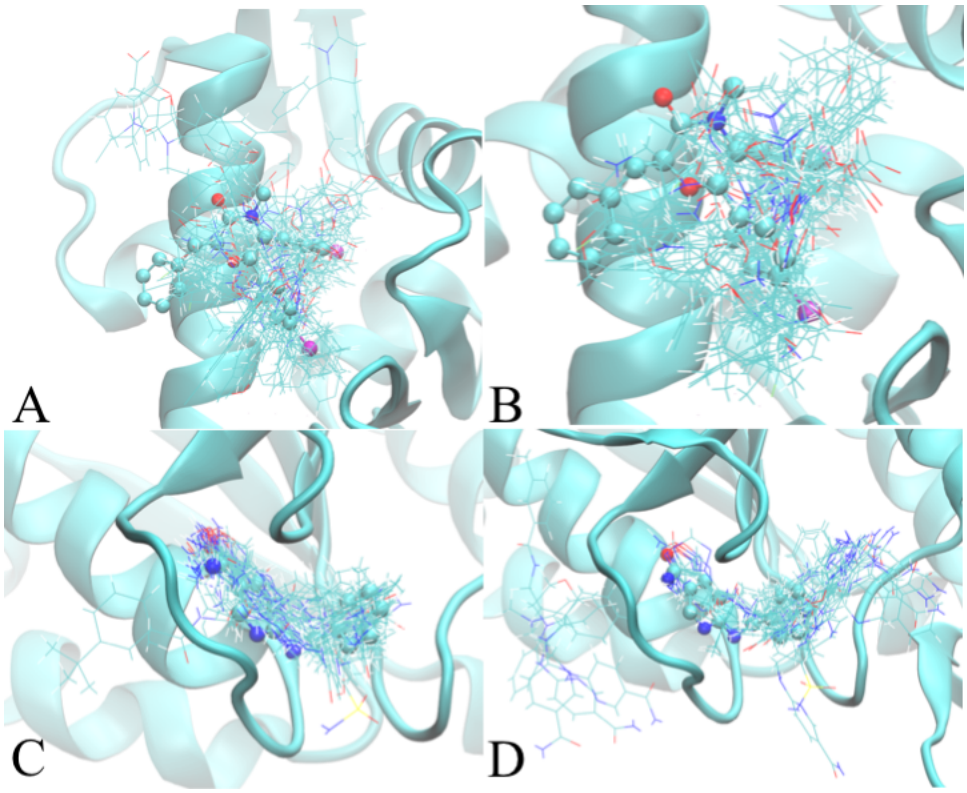

Figure S3: Ligand minimum LGFE conformations from the Exhaustive SILCS-MC protocol for HDM2 (A/B) and TRMD (C/D) . The left (A and C) and right (B and D) panels shows the ligand conformations obtained from the SILCS-MC additive and Drude, respectively. The crystallographic position of the ligand from the crystal structures used to initiate the SILCS simulations are displayed in CPK model to show the binding pocket.

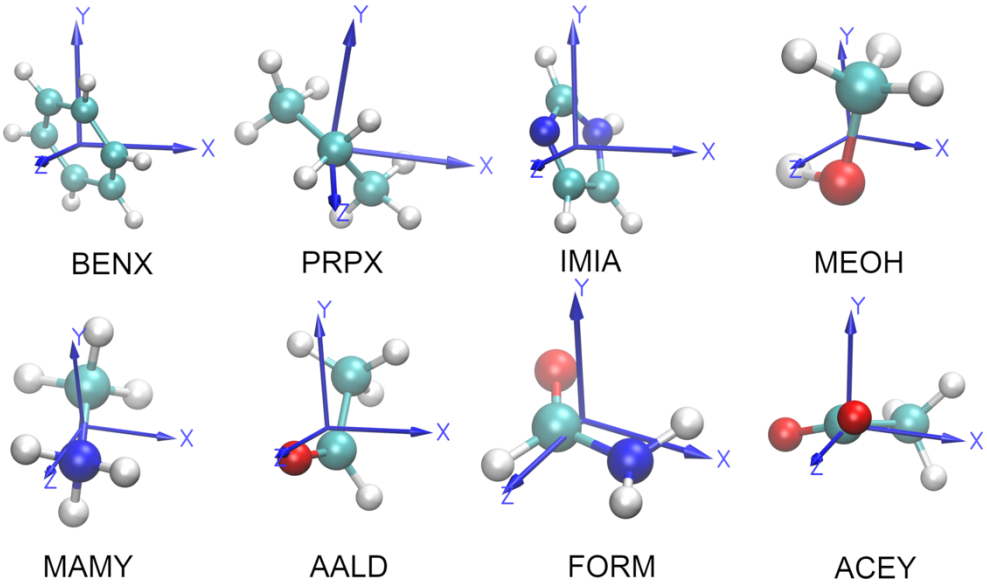

Figure S4: The x, y and z dipole moment vectors drawn on the eight solutes to show their orientations on the molecules.

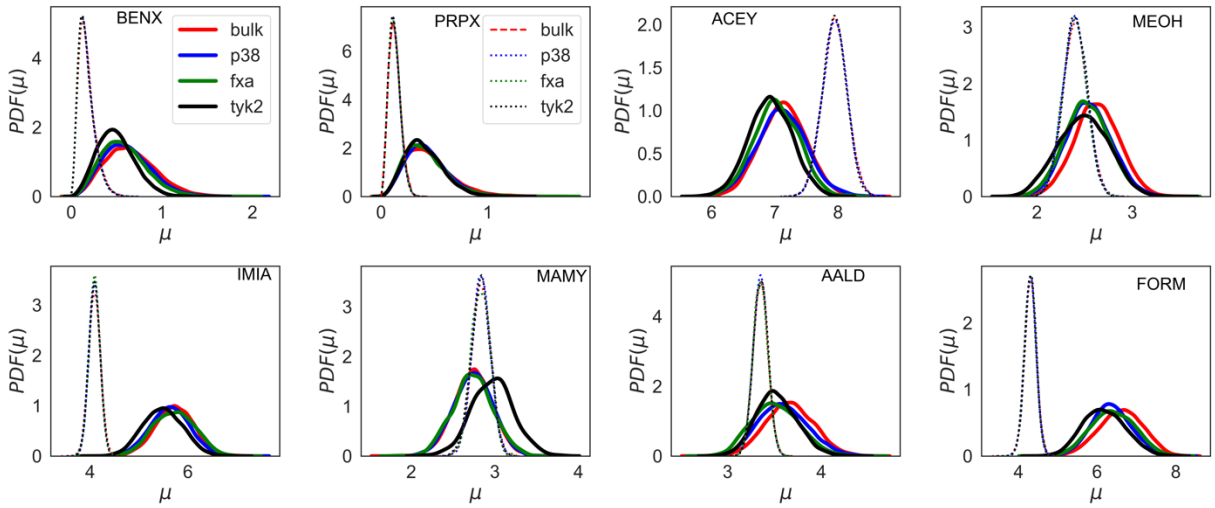

Figure S5: Total dipole moment distribution for all the eight solutes from the P38, Factor Xa, and TYK2 protein simulations. Solid and dotted/dashed lines represent the dipole distribution from Drude and additive simulations, respectively. The red color shows the dipole moment for the solutes present in the bulk phase and the blue, green, and black color represent dipole moments in the ligand binding pockets of P38, Factor Xa, and TYK2, respectively.

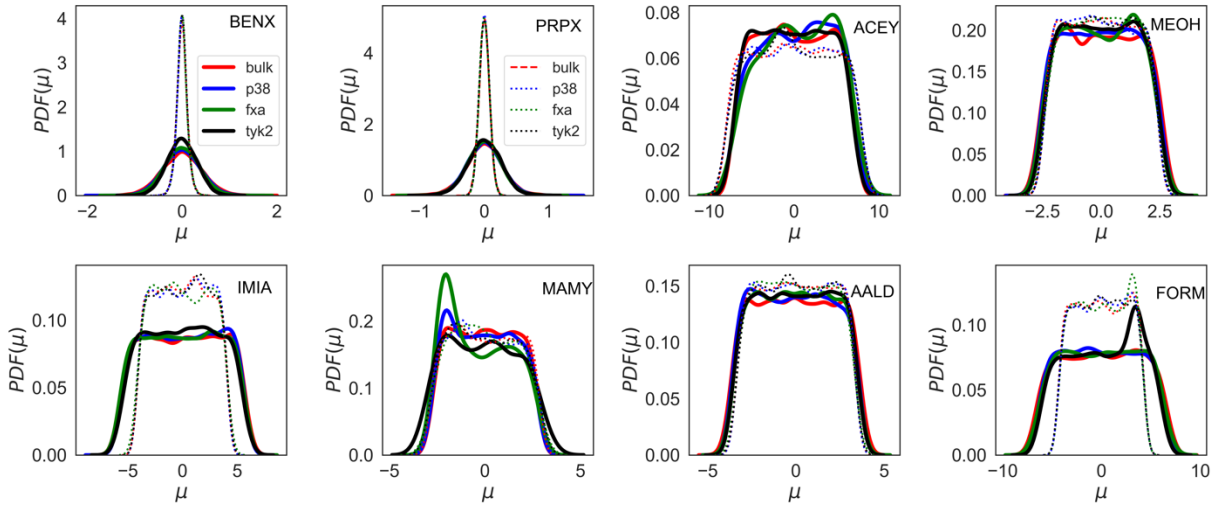

Figure S6: Dipole moment distributions in the x direction for all the eight solutes from the P38, Factor Xa, and TYK2 protein simulations. Solid and dotted/dashed lines represent the dipole distribution from Drude and additive simulations, respectively. The red color shows the dipole moment for the solutes present in the bulk phase and the blue, green, and black color represents dipole moments in the ligand binding pockets of P38, Factor Xa, and TYK2, respectively.

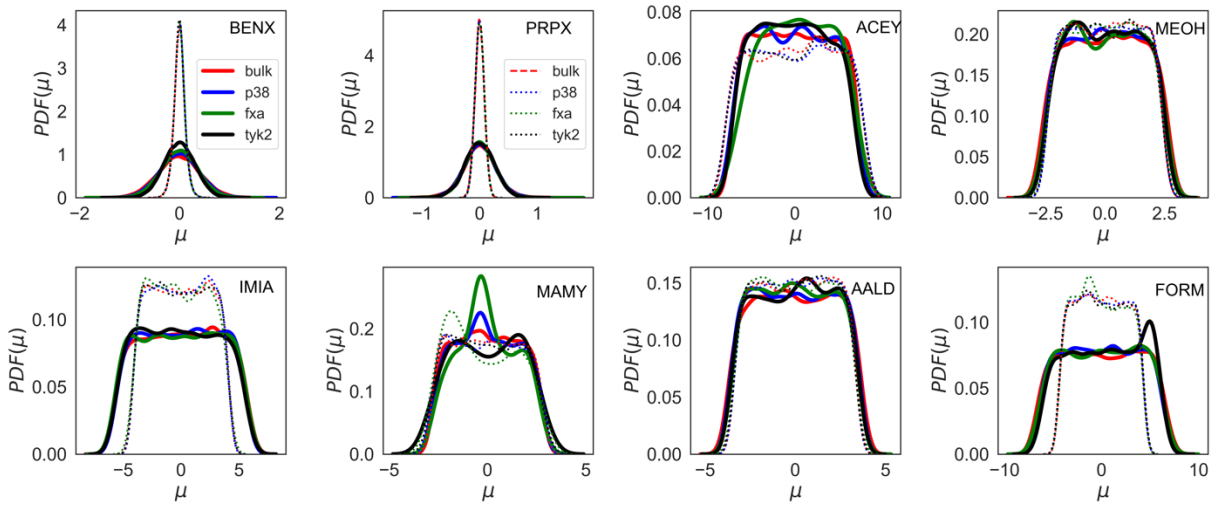

Figure S7: Dipole moment distributions in the y direction for all the eight solutes from the P38, Factor Xa, and TYK2 protein simulations. Solid and dotted/dashed lines represent the dipole distribution from Drude and additive simulations, respectively. The red color shows the dipole moment for the solutes present in the bulk phase and the blue, green, and black color represents dipole moments in the ligand binding pockets of P38, Factor Xa, and TYK2, respectively.

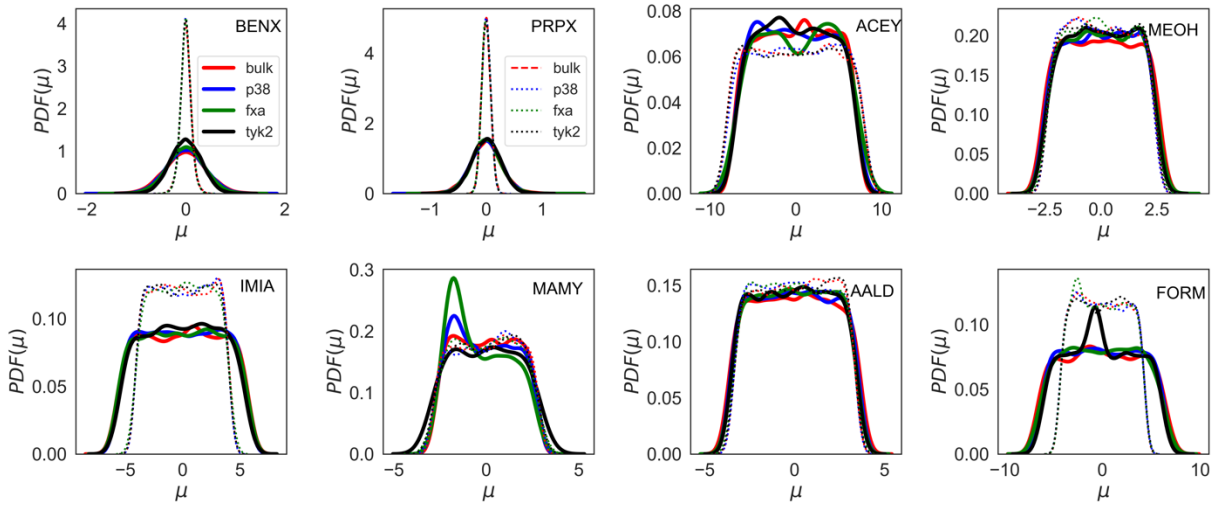

Figure S8: Dipole moment distributions in the z direction for all the eight solutes from the P38, Factor Xa, and TYK2 protein simulations. Solid and dotted/dashed lines represent the dipole distribution from Drude and additive simulations, respectively. The red color shows the dipole moments for the solutes present in the bulk phase and the blue, green, and black color represents dipole moments in the ligand binding pockets of P38, Factor Xa, and TYK2, respectively.

Table S3: Topology and parameter information for the SILCS polarizable Drude force field solutes.

```
* DRUDE SILCS solute topology file
*

read rtf card append
* Topology for drude model compounds
*
34

DEFA FIRS NONE LAST NONE
AUTOGENERATE ANGLES DIHEDRALS DRUDE  !note use of DRUDE


RESI BENX        0.00 ! benzene
!
!            HD1  HE1
!             |    |
!            CD1--CE1
!            /      \
!       HG--CG      CZ--HZ
!            \      /
!            CD2--CE2
!             |    |
!            HD2  HE2
!
GROUP
ATOM CG   CD2R6A  -0.1106  ALPHA -1.615  THOLE 1.195
ATOM HG   HDR6A    0.1106
ATOM CD1  CD2R6A  -0.1106  ALPHA -1.615  THOLE 1.195
ATOM HD1  HDR6A    0.1106
ATOM CD2  CD2R6A  -0.1106  ALPHA -1.615  THOLE 1.195
ATOM HD2  HDR6A    0.1106
ATOM CE1  CD2R6A  -0.1106  ALPHA -1.615  THOLE 1.195
ATOM HE1  HDR6A    0.1106
ATOM CE2  CD2R6A  -0.1106  ALPHA -1.615  THOLE 1.195
ATOM HE2  HDR6A    0.1106
ATOM CZ   CD2R6A  -0.1106  ALPHA -1.615  THOLE 1.195
ATOM HZ   HDR6A    0.1106
ATOM LPA  LS       0.00

BOND CD1 CG CD2 CG CE1 CD1
BOND CE2 CD2 CZ CE1 CZ CE2
BOND CG HG CD1 HD1 CD2 HD2 CE1 HE1
BOND CE2 HE2 CZ HZ CG LPA

LONEPAIR bisector LPA CG CD1 CD2 distance  1.375 angle   0.0 dihe 0.0

IC CG   CD1  CE1  CZ    0.0000  0.0000   0.0000  0.0000  0.0000
IC CD1  CE1  CZ   CE2   0.0000  0.0000   0.0000  0.0000  0.0000
IC CE1  CZ   CE2  CD2   0.0000  0.0000   0.0000  0.0000  0.0000
```

```
IC CD1  CD2  *CG  HG    0.0000  0.0000 180.0000  0.0000  0.0000
IC CE1  CG   *CD1 HD1   0.0000  0.0000 180.0000  0.0000  0.0000
IC CE2  CG   *CD2 HD2   0.0000  0.0000 180.0000  0.0000  0.0000
IC CZ   CD1  *CE1 HE1   0.0000  0.0000 180.0000  0.0000  0.0000
IC CZ   CD2  *CE2 HE2   0.0000  0.0000 180.0000  0.0000  0.0000
IC CE1  CE2  *CZ  HZ    0.0000  0.0000 180.0000  0.0000  0.0000
PATCHING FIRST NONE LAST NONE


RESI PRPX         0.000 ! propane
!
!  H11   H21    H31
!    \    |     /
! H12-C1--C2--C3-H32
!    /    |     \
!  H13   H22    H33
!
GROUP
ATOM H11 HDA3A     0.059
ATOM H12 HDA3A     0.059
ATOM H13 HDA3A     0.059
ATOM C1  CD33A    -0.177  ALPHA -2.051  THOLE 1.3
GROUP
ATOM C2  CD32A    -0.156  ALPHA -1.660  THOLE 1.3
ATOM H21 HDA2A     0.078
ATOM H22 HDA2A     0.078
GROUP
ATOM H31 HDA3A     0.059
ATOM H32 HDA3A     0.059
ATOM H33 HDA3A     0.059
ATOM C3  CD33A    -0.177  ALPHA -2.051  THOLE 1.3
ATOM LPA LS        0.00

BOND c1 h11 c1 h12  c1 h13
BOND c1 c2  c2 h21  c2 h22
BOND c2 c3  c3 h31  c3 h32  c3 h33 C2  LPA

LONEPAIR bisector LPA C2 C1 C3  distance  0.1 angle   0.0 dihe 0.0

! IC for PROP staggered conformer MP2/6-31G* optimized
IC   C1    C2   C3    H31   1.5260  112.39   59.77  110.82   1.0946
IC   C1    C2   C3    H32   1.5260  112.39  180.00  111.49   1.0938
IC   C1    C2   C3    H33   1.5260  112.39  -59.77  110.82   1.0946
IC   H11   C1   C2    C3    1.0946  110.82  -59.77  112.39   1.5260
IC   H11   C1   C2    H21   1.0946  110.82   62.14  109.48   1.0960
IC   H11   C1   C2    H22   1.0946  110.82  178.32  109.48   1.0960
IC   C3    C2   C1    H12   1.5260  112.39   59.77  110.82   1.0946
IC   C3    C2   C1    H13   1.5260  112.39  180.00  111.49   1.0938
PATCH FIRST NONE LAST NONE
```

```
RESI MAMY          1.00 ! from MAMM
!             HZ2
!              |
!              |
!    HZ1------NZ------HZ3     (+)
!              |
!              |
!         HE1-CE-HE3
!              |
!            HE2
GROUP
ATOM CE  CD33A     -0.100  ALPHA -1.656 THOLE 0.895
ATOM NZ  ND3P3A    -0.349  ALPHA -1.298 THOLE 0.895
ATOM HE1 HDA3C      0.143
ATOM HE2 HDA3C      0.143
ATOM HE3 HDA3C      0.143
ATOM HZ1 HDP1B      0.340
ATOM HZ2 HDP1B      0.340
ATOM HZ3 HDP1B      0.340
ATOM LPA  LT        0.00
BOND CE HE1  CE HE2  CE HE3
BOND CE NZ  NZ HZ1  NZ HZ2  NZ HZ3 NZ  LPA
LONEPAIR bisector LPA NZ HZ1 HZ3  distance  0.1 angle   0.0 dihe 0.0
IC   NZ   HE3  *CE  HE1  0.0000    0.00  120.00    0.00   0.0000
IC   NZ   HE3  *CE  HE2  0.0000    0.00 -120.00    0.00   0.0000
IC   HE3  CE   NZ   HZ3  0.0000    0.00  180.00    0.00   0.0000
IC   CE   HZ3  *NZ  HZ1  0.0000    0.00  120.00    0.00   0.0000
IC   CE   HZ3  *NZ  HZ2  0.0000    0.00 -120.00    0.00   0.0000


RESI FORM  0.000
!
!       O     HC
!       ||   /
!  HA---C---N
!            \
!             HT
!
GROUP
ATOM C      CD2O1C     0.302   ALPHA -1.757  THOLE 1.482
ATOM O      OD2C1C     0.000   ALPHA -0.901  THOLE 1.311
ATOM LPA    LPD       -0.233
ATOM LPB    LPD       -0.226
ATOM HA     HDP1C      0.060
ATOM N      ND2A1     -0.545   ALPHA -1.314  THOLE  1.187
ATOM HC     HDP1A      0.346
ATOM HT     HDP1A      0.296
BOND C    O
BOND C    HA
```

```
BOND C    N
BOND N    HC
BOND N    HT
BOND        O    LPA         O    LPB
LONEPAIR relative LPA   O      C      HA      distance 0.30 angle 91.00
dihe   0.00
LONEPAIR relative LPB   O      C      HA      distance 0.30 angle 91.00
dihe 180.00
ANISOTROPY       O       C    LPA    LPB  A11  0.583  A22 0.711
ANISOTROPY       N       C    HC     HT   A11  0.970 A22  1.347
IMPR C      N      O      HA
IMPR N      HC    HT     C
IC O   C   N   HT         1.2250  125.00  180.00  123.50   1.0250
IC O   N   *C  HA         1.2250  125.00  180.00  114.00   0.0000
IC C   HT  *N  HC         1.3500  123.50  180.00  113.00   1.0250
IC HA  C    N  HC         1.3500  123.50    0.00  113.00   1.0250


RESI ACEY         -1.00 ! from ACET
GROUP
!
!   H1      O1
!     \    //
! H2--C1--C2 (-1)
!     /    \
!   H3      O2
!
GROUP
ATOM C1   CD33A   -0.194  ALPHA -2.528  THOLE 1.414
ATOM H1   HDA3A    0.004
ATOM H2   HDA3A    0.004
ATOM H3   HDA3A    0.004
ATOM C2   CD2O2A   0.708  ALPHA -1.016  THOLE 0.899
ATOM O1   OD2C2A   0.003  ALPHA -0.699  THOLE 2.399
ATOM O2   OD2C2A   0.003  ALPHA -0.699  THOLE 2.399
ATOM LP1A LP      -0.383
ATOM LP1B LP      -0.383
ATOM LP2A LP      -0.383
ATOM LP2B LP      -0.383
ATOM LPA  LT       0.00

BOND C1 C2  C2 O1  C2 O2
BOND C1 H1  C1 H2  C1 H3
BOND O1 LP1A O1 LP1B
BOND O2 LP2A O2 LP2B C2 LPA
IMPH O1 C1 O2 C2

LONEPAIR relative LP1A O1 C2 C1  distance 0.35 angle 110.0 dihe   0.0
LONEPAIR relative LP1B O1 C2 C1  distance 0.35 angle 110.0 dihe 180.0
ANISOTROPY O1 C2 LP1A LP1B A11 0.7229 A22 1.265
LONEPAIR relative LP2A O2 C2 C1  distance 0.35 angle 110.0 dihe   0.0
LONEPAIR relative LP2B O2 C2 C1  distance 0.35 angle 110.0 dihe 180.0
ANISOTROPY O2 C2 LP2A LP2B A11 0.7229 A22 1.265
LONEPAIR bisector LPA C2 O1 O2  distance  0.1 angle   0.0 dihe 0.0
```

```
IC  O1  C2 C1  H1   1.2543  118.04  180.00  111.57   1.1108
IC  O1  O2 *C2 C1   1.2543  122.88  180.00  119.08   1.5229
IC  H1  C2 *C1 H2   1.1108  111.57  120.15  110.12   1.1112
IC  H1  C2 *C1 H3   1.1108  111.57 -120.15  110.12   1.1112
IC  H1  C1 C2  O2   1.1108  111.57    0.00  119.08   1.2536


PATCHING FIRST NONE LAST NONE


END

read para card append
* SILCS LP VdW and NBFIX
*

BONDS
!acey
OD2C2A   LP          0.00      0.000 !
CD2R6A   LS          0.00      0.000 ! BENZENE LONEPAIR
CD2O2A   LT          0.00      0.000
CD32A    LS          0.00      0.000
ND3P3A   LT          0.00      0.000
!meox
OD31A    LP          0.00      0.000 !

! aald and meagan's
CD2O1C   LPDO1       0.00      0.000 !
! form
CD2O1C   ND2A1   420.000     1.3700

ANGLES

! aald and meagan's

OD2C1D   CD2O1C   CD33C     63.60    123.81  ! acetone 63.6 121.78; 2-
butanone 63.6 121.78; 2-pentanone 63.6 121.88;
! form
ND2A1    CD2O1C   OD2C1C     67.000   127.00   50.000    2.37000  75
ND2A1    CD2O1C  HDP1C      30.000   115.00    50.000    1.98000
CD2O1C   ND2A1    HDP1A      52.000   117.00


DIHEDRALS

! aald and meagan's
OD2C1D   CD2O1C   CD33C    HDA3A       0.015    3     0.00 !
! form
HDP1C    CD2O1C   ND2A1    HDP1A     1.400    2   180.0
OD2C1C   CD2O1C   ND2A1    HDP1A     3.000    2   180.0

IMPROPERS
```

```
! form
CD2O1C   ND2A1   OD2C1C   HDP1C     61.00     0     0.0
ND2A1    HDP1A    HDP1A   CD2O1C     3.00     0     0.0


NONBONDED nbxmod  5 atom cdiel fshift vatom vdistance vswitch -
cutnb 14.0 ctofnb 12.0 ctonnb 10.0 eps 1.0 e14fac 1.0 wmin 1.5


NBFIX
ODW     CD2O1C -0.16560  3.31690 ! original 3.3669, eps -0.16560
ODW     OD2C1D -0.20540  3.50690 ! NMA, PEML 3.60690, normally 3.5869, eps
-0.20540

END
```

References

1. F. Y. Lin and A. D. MacKerell, Jr., *J. Phys. Chem. B*, 2017, **121**, 6813-6821.
2. GSK Trmd Dataset, (https://drugdesigndata.org/about/datasets/226).