Biophysical Journal, Volume *116*

Supplemental Information

Molecular Mechanisms of Macular Degeneration Associated with the Complement Factor H Y402H Mutation

Reed E.S. Harrison and Dimitrios Morikis

Molecular Mechanisms of Macular Degeneration Associated with the Complement Factor H Y402H Mutation

Reed E. S. Harrison, Dimitrios Morikis
Department of Bioengineering, University of California, Riverside, California, United States of America

SUPPLEMENTARY TABLES AND FIGURES

Table S1: Significant coevolved pairs in SCR7 with minimum distances between heavy atoms of residues measured from $SCR7^{Y402}$ (PDB: 2jgx) and $SCR7^{H402}$ (PDB: 2uwn).

Position 1	Position 2	DCA Score	Residue 1	Residue 2	p-value	Minimum distance (Å)
24	31	1.01	429	438	4.0E-08	5.53
2	4	0.95	394	397	2.5E-07	4.53
5	9	0.91	402	412	8.8E-07	9.31
27	28	0.75	432	434	4.8E-05	6.43
2	3	0.64	394	395	6.0E-04	5.40
20	21	0.64	424	425	6.7E-04	5.40
1	4	0.62	389	397	8.4E-04	17.14
16	17	0.59	419	420	1.6E-03	4.48
19	21	0.57	422	425	2.2E-03	5.74
8	23	0.57	411	427	2.3E-03	9.60
13	26	0.57	416	431	2.5E-03	17.91
8	25	0.57	411	430	2.6E-03	5.29
10	23	0.53	413	427	5.0E-03	5.36
32	34	0.52	440	442	6.2E-03	7.08
3	4	0.51	395	397	7.3E-03	5.89
13	30	0.50	416	436	8.7E-03	14.47
11	32	0.49	414	440	9.8E-03	7.45

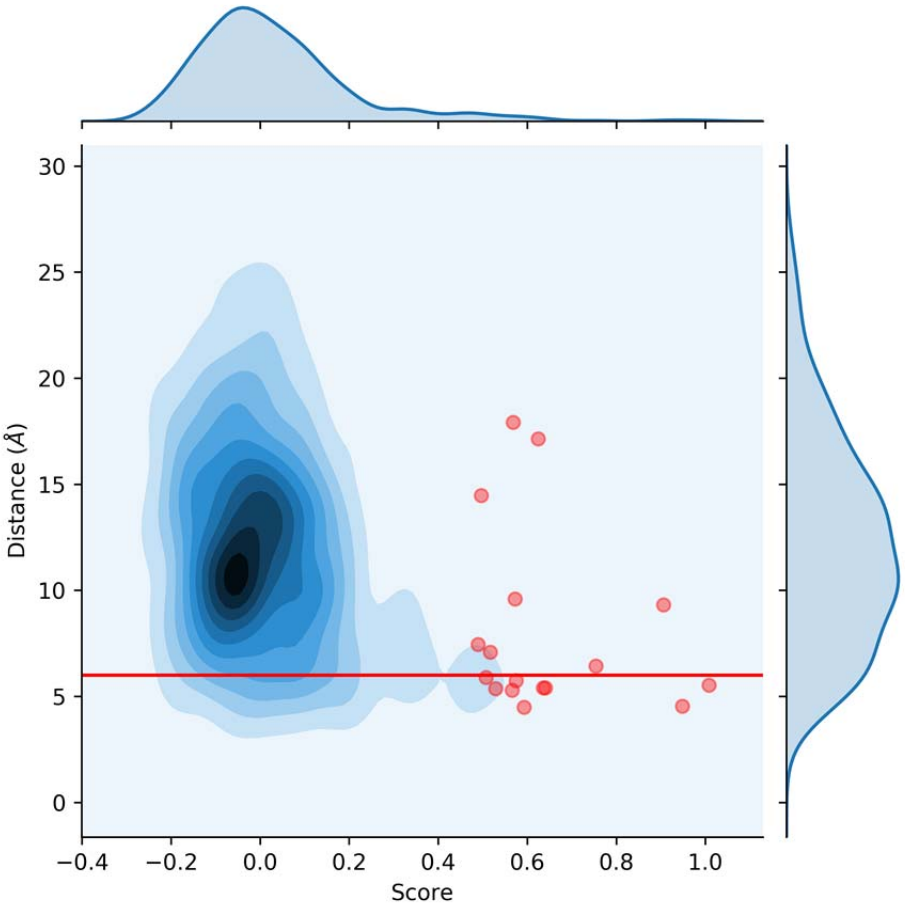

Figure S1: Joint distribution for DCA scores and minimum distance of separation for heavy atoms between residue pairs. The cutoff distance of 6 Å is shown by a red line, and coevolutionary couplings with scores in the top 1% from a Gaussian distribution are shown with red circles. As shown in the above image, significant coevolved pairs (DCA scores in the top 1%) are more likely to be true contacts in FH SCR7.

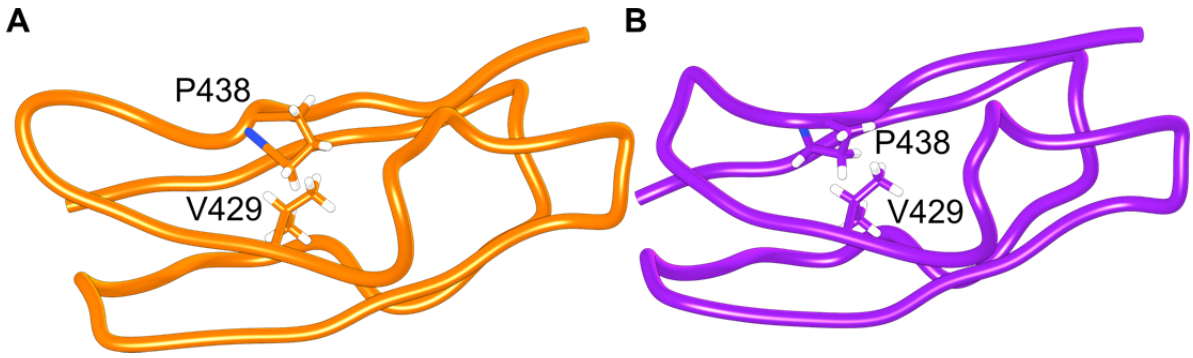

Figure S2: Structural representations of (A) SCR7^{H402} from NMR (PDB:2jgx) and (B) SCR7^{Y402} from X-ray crystallography (PDB: 2uwn) are shown with residues from the V429-P438 coevolved pair displayed. In both structures, a contact is observed between members of this coevolved pair

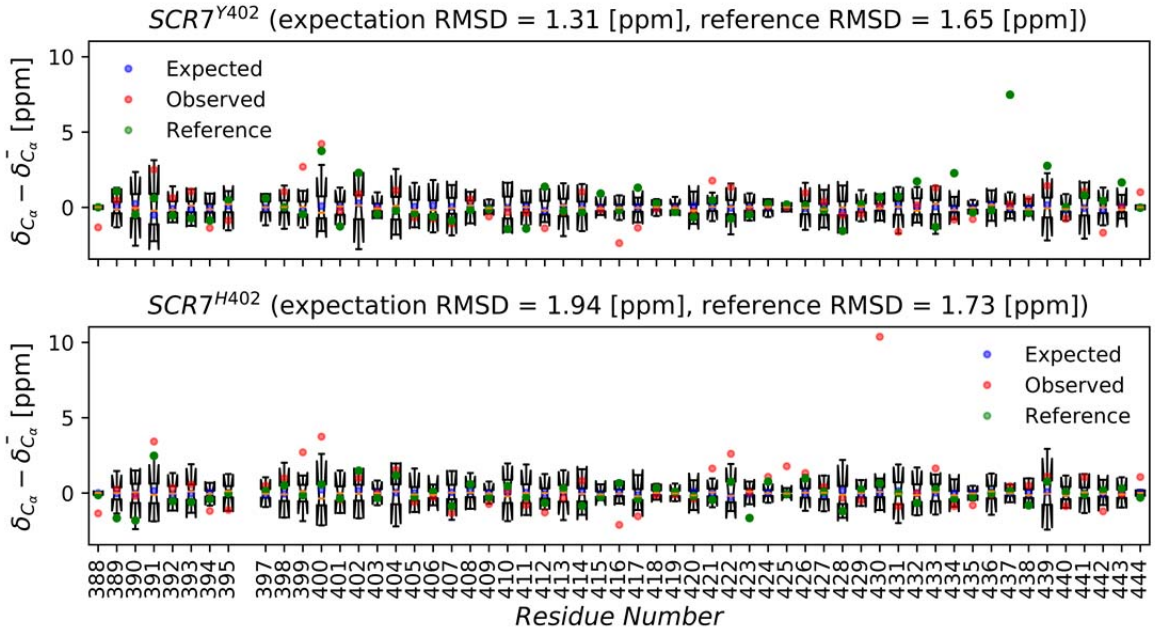

Figure S3: Boxplots for predicted C_α chemical shifts from representative structures from Markov chains for $SCR7^{Y402}$ (top) and $SCR7^{H402}$ (bottom) are displayed above. Expected chemical shifts from the Markov chains for each SCR7 isoform are marked by blue circles, while predicted shifts for reference structures 2jgx and 2uwn from the PDB are marked by green circles. Known chemical shifts from experiments are annotated by red circles.

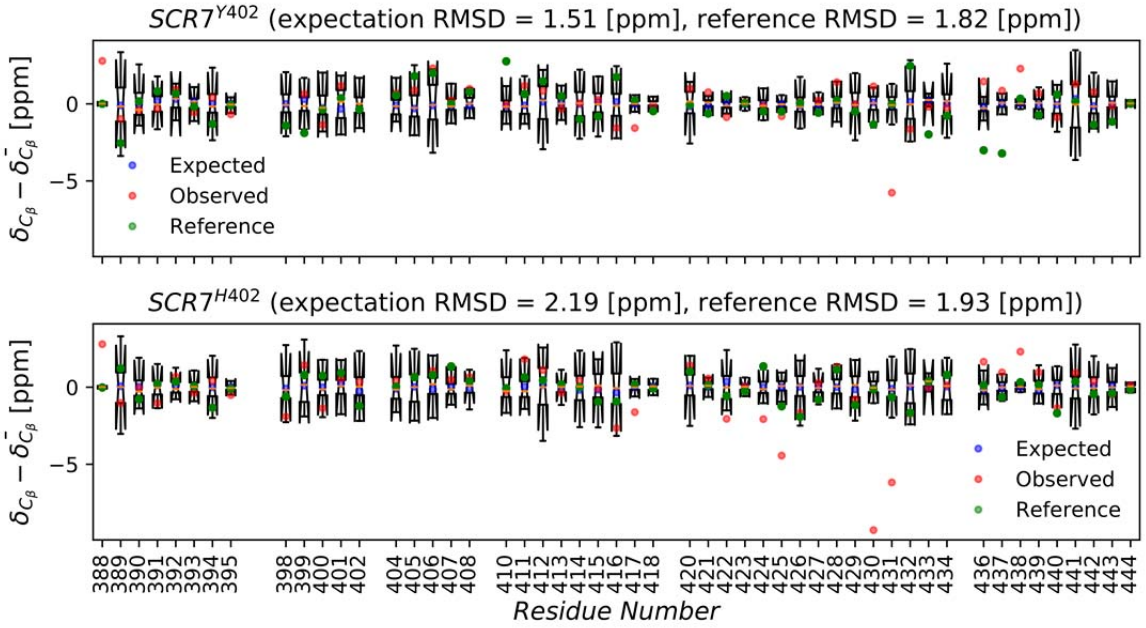

Figure S4: Boxplots for predicted C_β chemical shifts from representative structures from Markov chains for $SCR7^{Y402}$ (top) and $SCR7^{H402}$ (bottom) are displayed above. Expected chemical shifts from the Markov chains for each SCR7 isoform are marked by blue circles, while predicted shifts for reference structures 2jgx and 2uwn from the PDB are marked by green circles. Known chemical shifts from experiments are annotated by red circles.

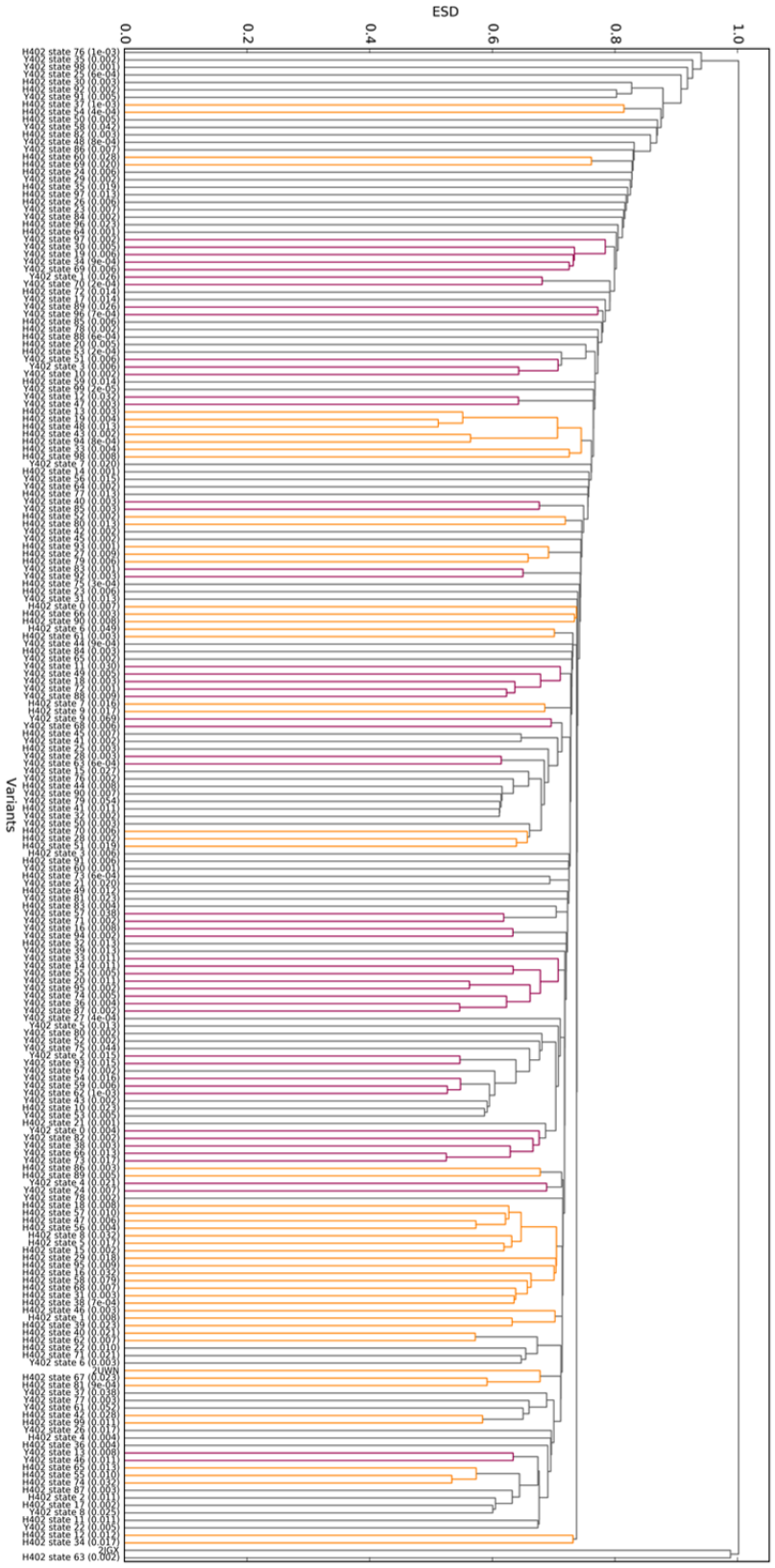

Figure S5: Electrostatic similarity distance (ESD) is compared for each representative structure of SCR7 isoforms and for reference structures 2jgx and 2uwn from the PDB. Most ESD values are below 0.7, indicating a large degree of similarity (an ESD of zero indicates identical potentials). The maximum ESD theoretically possible is 2.0. We expect that large differences between representative structures of SCR7 isoforms are unlikely to arise from a single mutation.

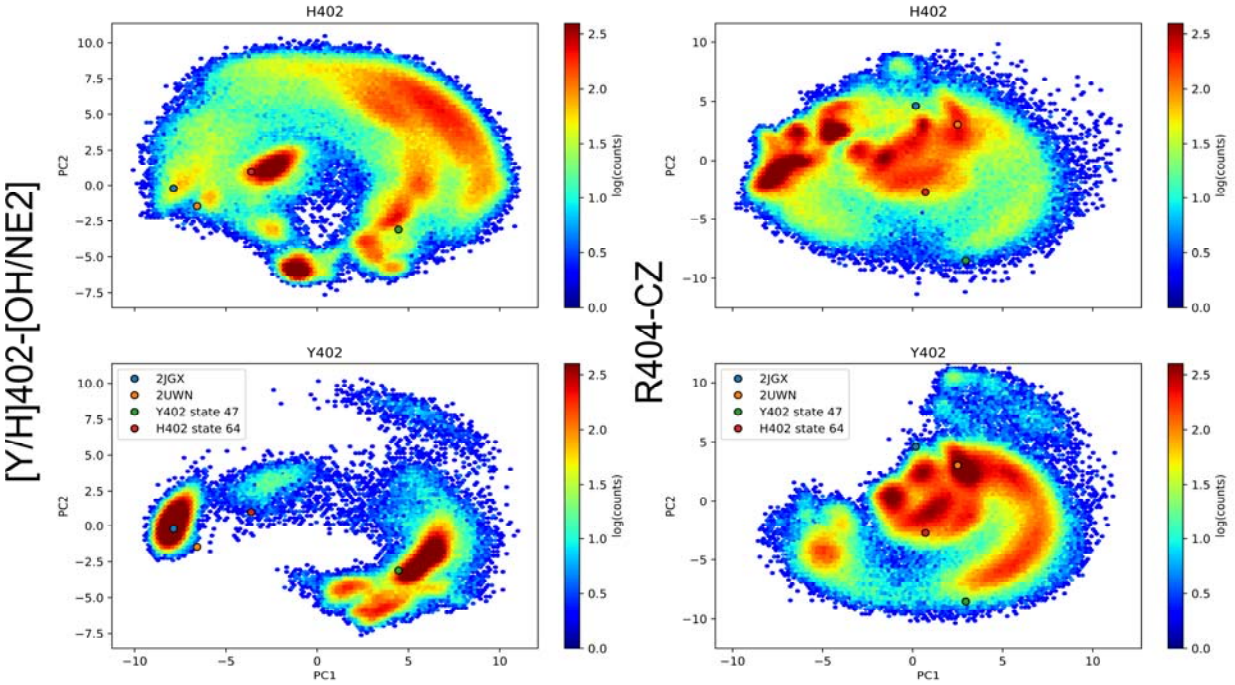

Figure S6: Free energy landscapes for a two-dimensional representation of sidechain orientations for residues (A) [Y/H]402 and (B) R404 are shown for SCR7 isoforms, where a larger value for the logarithm of the number of observations in a bin indicates a lower free energy. To describe the orientation of sidechains, position vectors are found from coordinates of atom NE2 in H402, OH in Y402, and CZ in R404 and decomposed into two dimensions with principle component analysis. Regions of the landscape are annotated with circles to indicate the side chain orientations in references structures for $SCR7^{Y402}$ (2JGX) and $SCR7^{H402}$ (2UWN) and for states from each isoform where the maximum rate-constant for association with heparin is observed (Y402 state 47, H402 state 64).

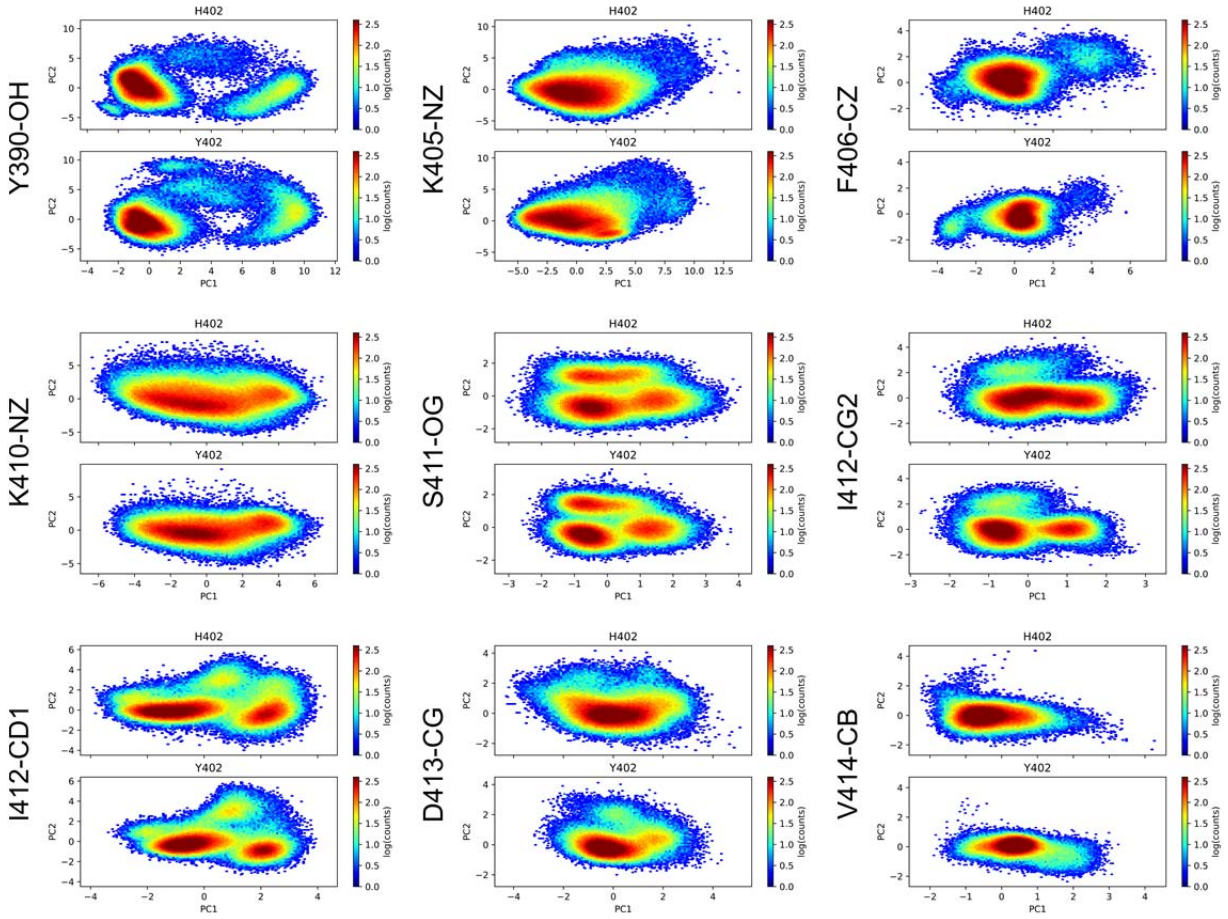

Figure S7: Free energy landscapes for a two-dimensional representation of sidechain orientations for residues Y390, K405, F406, K410, S411, I412, D413, and V414 are shown for SCR7 isoforms, where a larger value for the logarithm of the number of observations in a bin indicates a lower free energy. To describe the orientation of sidechains, position vectors are found from coordinates of panel labels and decomposed into two dimensions with principle component analysis. For these residues, free energy landscapes between SCR7 isoforms are nearly identical.

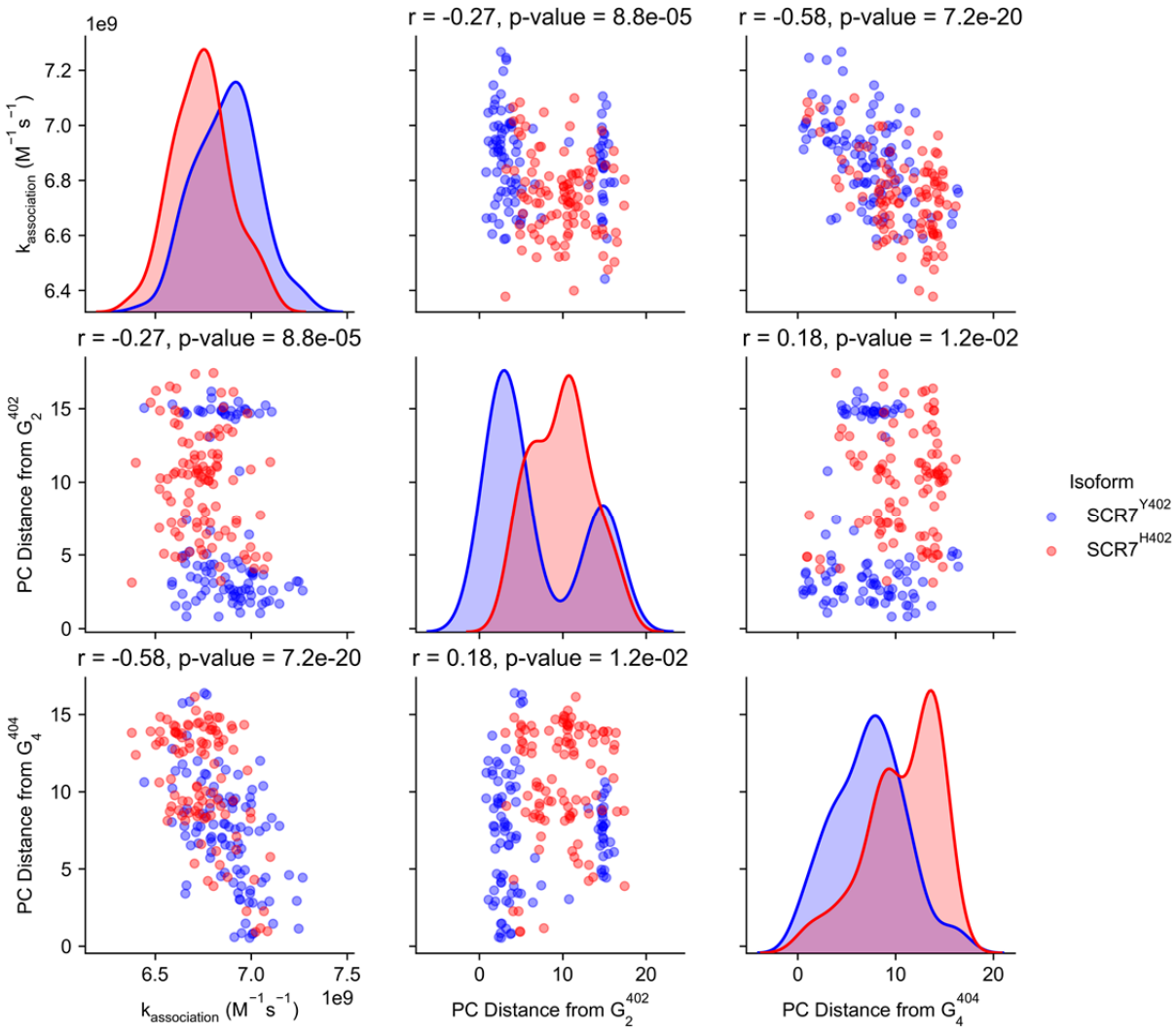

Figure S8: Associations between orientations of side chains for residues [Y/H]402 and R404 and predicted association rate constants. The PC distance from G_2^{402} is the distance (in PC space) from energy minimum 2 for the side-chain orientation of [Y/H]402. The PC distance from G_4^{404} is the distance (in PC space) from energy minimum 4 for the side-chain orientation of R404. Only the orientation of R404 is strongly correlated with the predicted association rate constant.

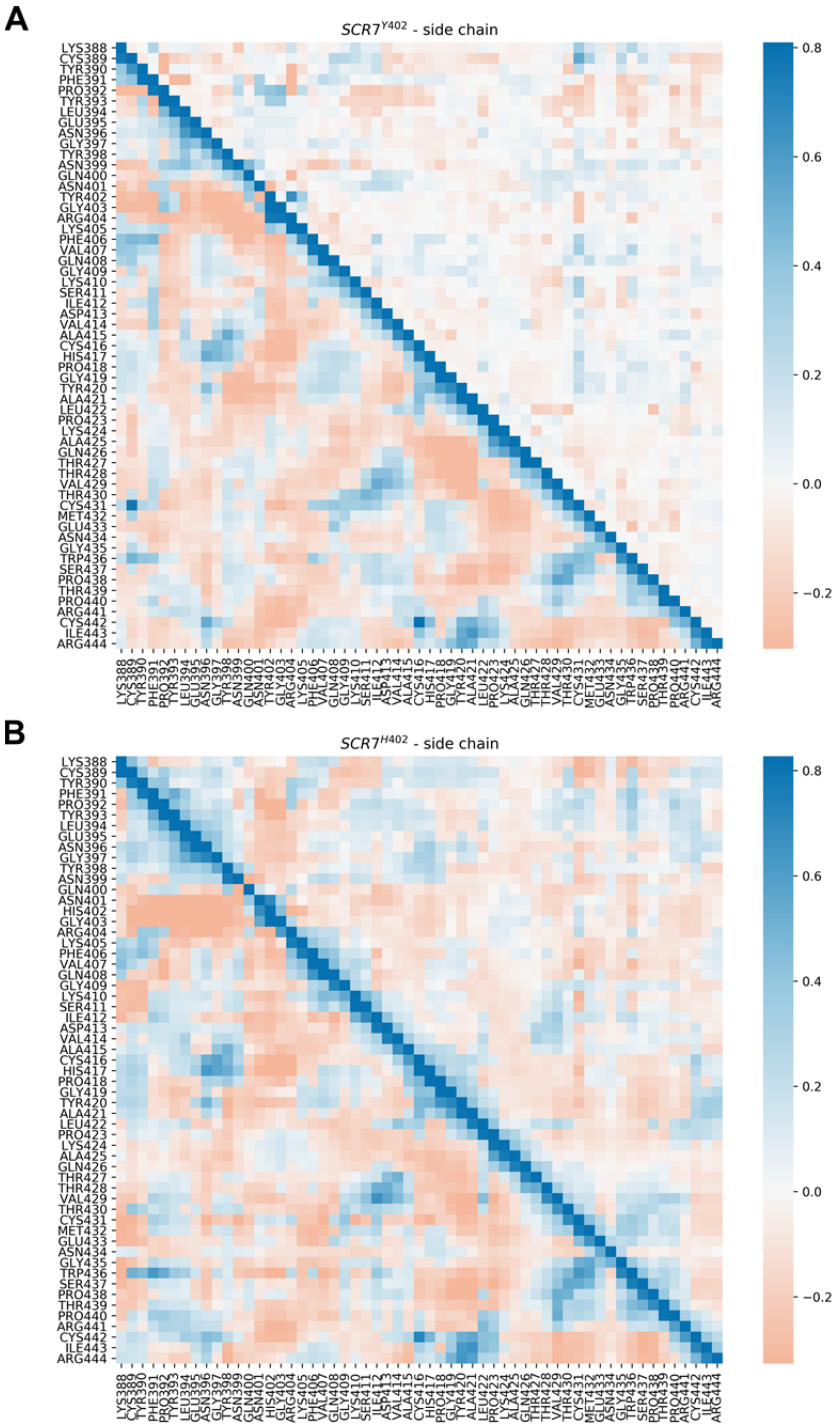

Figure S9: Dynamic cross-correlation matrices (DCCM) for (A) SCR7^{Y402} and (B) SCR7^{H402} side-chains. The lower triangle of the DCCM shows the mean values from leave-one out crossvalidation of all trajectories (100 ps time step) for a single SCR7 isoform, while the upper triangle shows mean DCCM differences between isoforms. In comparing DCCM matrices between isoforms, the other isoform is always subtracted from the current isoform being analyzed. For example, the upper triangle in panel A is calculated by subtracting the DCCM for SCR7^{H402} from the DCCM for SCR7^{Y402}. Note how R404 anti-correlated with H402 but correlated with Y402.

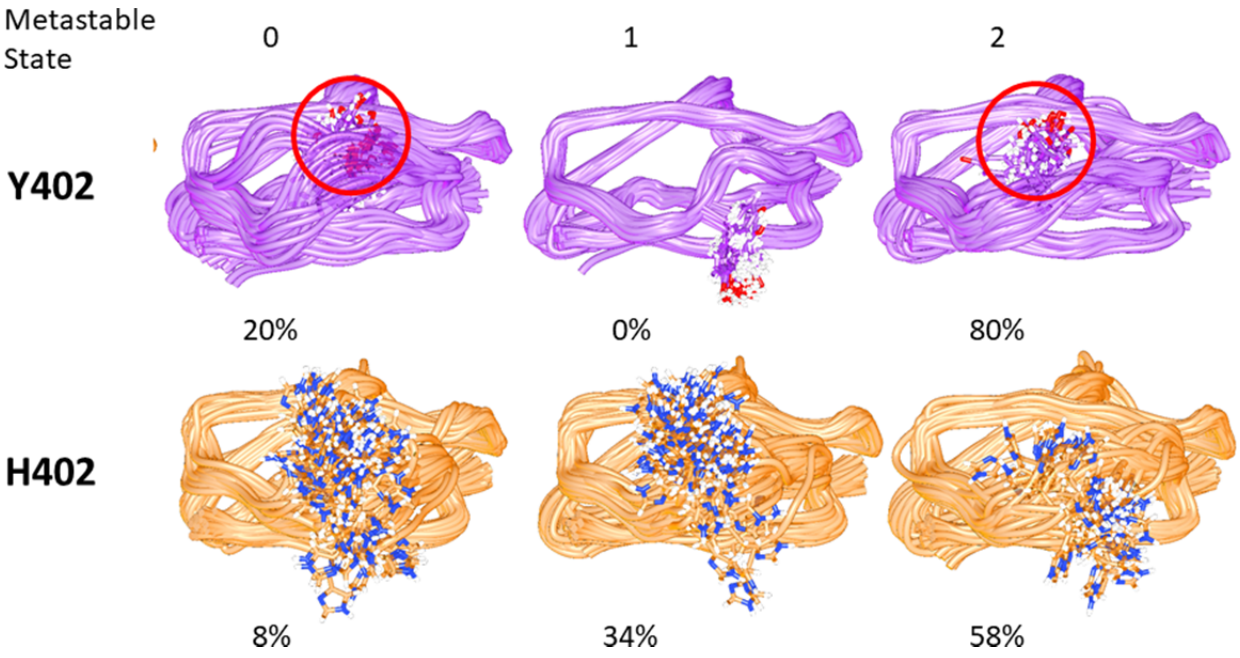

Figure S10: Coarse grain (metastable) states with associated probabilities from preliminary Markov models of SCR6-8 conformational dynamics. SCR7^{Y402} is colored purple, and SCR7^{H402} is color orange. Only the side chain for position 402 is displayed for multiple samples from each metastable state. Models were based on three 100 ns simulations for each SCR6-8 isoform and constructed similarly to the SCR7 Markov chain. These results suggest that we can recapitulate the behavior of Y402 forming of a coevolved contact with I412 in SCR6-8^{Y402} as shown in metastable states 0 and 2 with the red circle. These data were not used in the SCR7 study since we required more data to construct a high quality model.

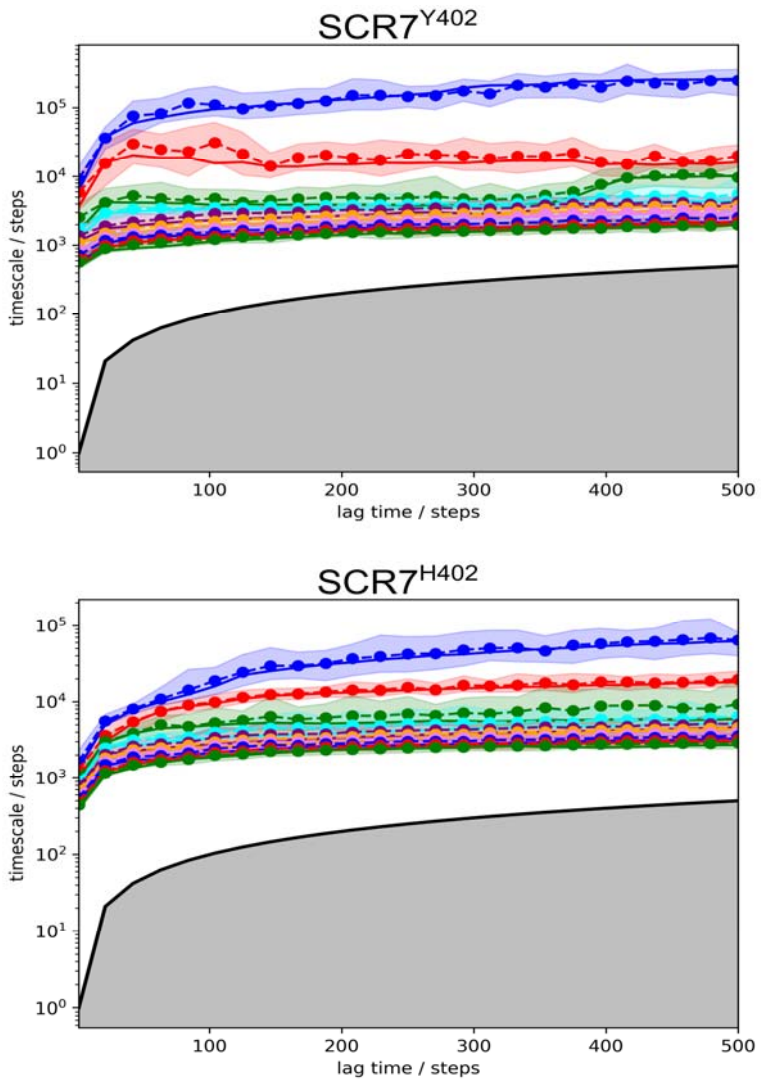

Figure S11: Timescales calculated from the probability transition matrices from Markov chains for SCR7^{Y402} (top) and SCR7^{H402} (bottom) are plotted versus lag time used to construct the Markov chain. To satisfy the Markov property, a lag time should be selected where the timescale does not change with increasing values of the timescale. Different timescale responses are colored uniquely, and one step corresponds to 100 ps. We selected a lag time of 200 steps (20 ns).

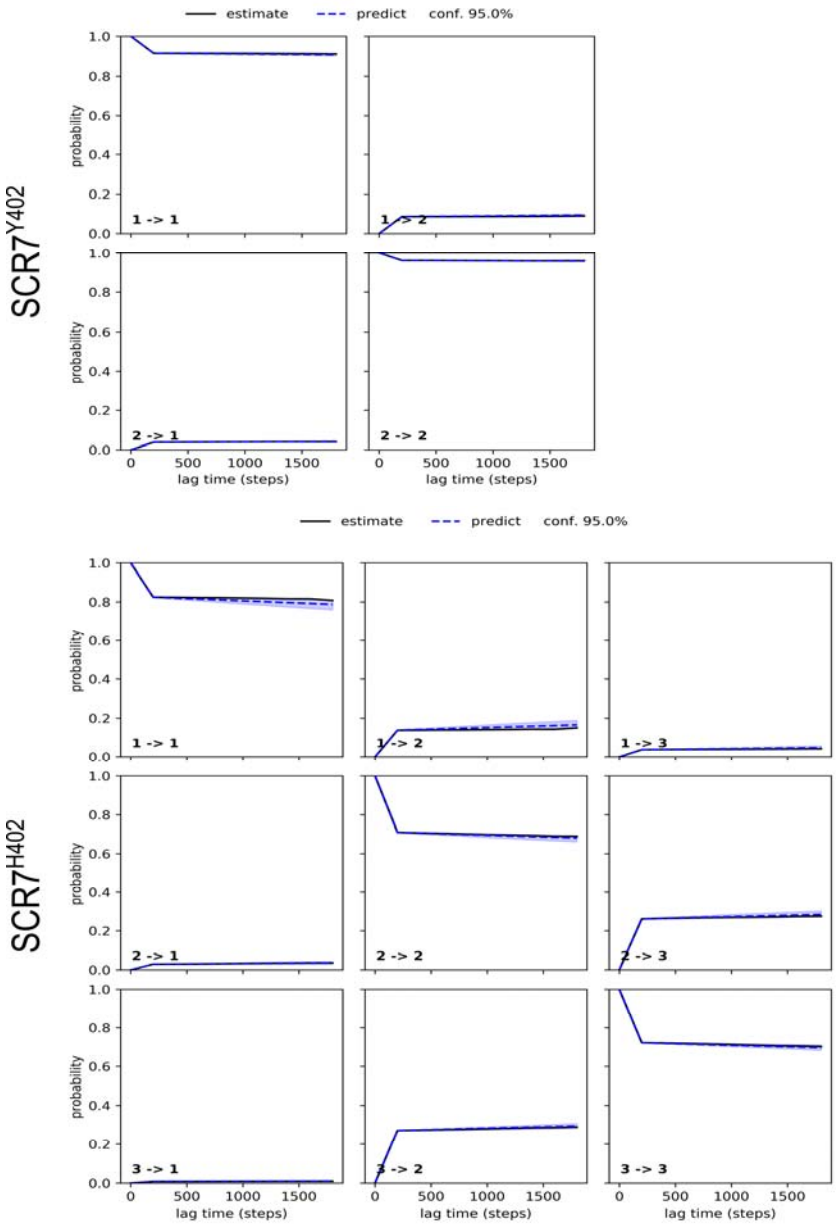

Figure S12: Chapman-Kolmogorov validation of Markov chains for SCR7^{Y402} (top) and SCR7^{H402} (bottom) are show above. These panels suggest the probabilities of transitioning between metastable states by propagation of the Markov chain reproduces (within a 95% confidence interval) probabilities directly calculated from observed data. One step corresponds to 100 ps, and each subpanel describes the probabilities of a particular transition calculated at multiple lag times.

SIMULATION INPUT FILES (Gromacs)

ion.mdp – solvation

```
constraints = h-bonds
cutoff-scheme = Verlet
vdwtype = cutoff
vdw-modifier = force-switch
rlist = 1.2
rvdw = 1.2
rvdw-switch = 1.0
coulombtype = PME
rcoulomb = 1.2
DispCorr = no
```

minim.mdp – minimization

```
; minim.mdp - used as input into grompp to generate em.tpr
integrator  = steep        ; Algorithm (steep = steepest descent minimization)
emtol       = 1000.0       ; Stop minimization when the maximum force < 1000.0 kJ/mol/nm
emstep      = 0.01         ; Energy step size
nsteps      = 50000        ; Maximum number of (minimization) steps to perform

; Parameters describing how to find the neighbors of each atom and how to calculate the interactions
constraints = h-bonds
cutoff-scheme = Verlet
vdwtype = cutoff
vdw-modifier = force-switch
rlist = 1.2
rvdw = 1.2
rvdw-switch = 1.0
coulombtype = PME
rcoulomb = 1.2
DispCorr = no
```

nvt.mdp – NVT equilibration

```
title           = fh ccp7 h402 equilibration
define          = -DPOSRES    ; position restrain the protein
; Run parameters
integrator    = md              ; leap-frog integrator
nsteps          = 50000                 ; 2 * 50000 = 100 ps
dt              = 0.002         ; 2 fs
; Output control
nstxout         = 500           ; save coordinates every 1.0 ps
nstvout         = 500           ; save velocities every 1.0 ps
nstenergy     = 500           ; save energies every 1.0 ps
nstlog          = 500           ; update log file every 1.0 ps
; Bond parameters
continuation            = no            ; first dynamics run
constraint_algorithm    = lincs       ; holonomic constraints
constraints             = all-bonds     ; all bonds (even heavy atom-H bonds) constrained
lincs_iter              = 1             ; accuracy of LINCS
lincs_order             = 4             ; also related to accuracy
; Neighborsearching
cutoff-scheme   = Verlet
ns_type                 = grid          ; search neighboring grid cells
nstlist                 = 10          ; 20 fs, largely irrelevant with Verlet
rlist           = 1.2
rcoulomb        = 1.2           ; short-range electrostatic cutoff (in nm)
rvdw                    = 1.2         ; short-range van der Waals cutoff (in nm)
rvdw-switch      = 1.0
vdwtype          = cutoff
vdw-modifier      = force-switch
; Electrostatics
coulombtype             = PME       ; Particle Mesh Ewald for long-range electrostatics
```

```
pme_order      = 4             ; cubic interpolation
fourierspacing          = 0.16          ; grid spacing for FFT
; Temperature coupling is on
tcoupl                  = V-rescale             ; modified Berendsen thermostat
tc-grps                 = Protein Non-Protein   ; two coupling groups - more accurate
tau_t                   = 0.1      0.1          ; time constant, in ps
ref_t                   = 300      300          ; reference temperature, one for each group, in K
; Pressure coupling is off
pcoupl                  = no            ; no pressure coupling in NVT
; Periodic boundary conditions
pbc             = xyz               ; 3-D PBC
; Dispersion correction
DispCorr    = no            ; account for cut-off vdW scheme
; Velocity generation
gen_vel                 = yes           ; assign velocities from Maxwell distribution
gen_temp    = 300           ; temperature for Maxwell distribution
gen_seed    = -1            ; generate a random seed
```

npt.mdp – NPT equilibration

```
title           = fh ccp7 h402 equilibration
define                  = -DPOSRES  ; position restrain the protein
; Run parameters
integrator  = md            ; leap-frog integrator
nsteps                  = 50000                 ; 2 * 50000 = 100 ps
dt              = 0.002         ; 2 fs
; Output control
nstxout                 = 500           ; save coordinates every 1.0 ps
nstvout                 = 500           ; save velocities every 1.0 ps
nstenergy   = 500           ; save energies every 1.0 ps
nstlog                  = 500           ; update log file every 1.0 ps
; Bond parameters
```

```
continuation                = yes               ; Restarting after NVT
constraint_algorithm        = lincs          ; holonomic constraints
constraints             = all-bonds        ; all bonds (even heavy atom-H bonds) constrained
lincs_iter              = 1                 ; accuracy of LINCS
lincs_order             = 4                 ; also related to accuracy
; Neighborsearching
cutoff-scheme       = Verlet
ns_type                 = grid              ; search neighboring grid cells
nstlist                 = 10     ; 20 fs, largely irrelevant with Verlet scheme
rlist            = 1.2
rcoulomb          = 1.2             ; short-range electrostatic cutoff (in nm)
rvdw                    = 1.2           ; short-range van der Waals cutoff (in nm)
rvdw-switch         = 1.0
vdwtype           = cutoff
vdw-modifier        = force-switch
; Electrostatics
coulombtype             = PME               ; Particle Mesh Ewald for long-range electrostatics
pme_order      = 4              ; cubic interpolation
fourierspacing          = 0.16              ; grid spacing for FFT
; Temperature coupling is on
tcoupl                  = V-rescale             ; modified Berendsen thermostat
tc-grps                 = Protein Non-Protein    ; two coupling groups - more accurate
tau_t                   = 0.1       0.1         ; time constant, in ps
ref_t                   = 300       300         ; reference temperature, one for each group, in K
; Pressure coupling is on
pcoupl                  = Parrinello-Rahman         ; Pressure coupling on in NPT
pcoupltype          = isotropic          ; uniform scaling of box vectors
tau_p                   = 2.0                       ; time constant, in ps
ref_p                   = 1.0                       ; reference pressure, in bar
compressibility         = 4.5e-5                ; isothermal compressibility of water, bar^-1
refcoord_scaling        = com
```

```
; Periodic boundary conditions
pbc          = xyz              ; 3-D PBC
; Dispersion correction
DispCorr     = no               ; account for cut-off vdW scheme
; Velocity generation
gen_vel              = no              ;          Velocity          generation          is          off
```

md.mdp – NPT production run

```
title          = fh ccp7 h402 production run 1us
; Run parameters
integrator     = md               ; leap-frog integrator
nsteps                 = 500000000     ; 2 * 500000 = 1000 ps (1 ns)
dt               = 0.002                ; 2 fs
; Output control
nstxout                      = 5000              ; save coordinates every 10.0 ps
nstvout                      = 5000              ; save velocities every 10.0 ps
nstenergy          = 5000              ; save energies every 10.0 ps
nstlog                       = 5000              ; update log file every 10.0 ps
nstxout-compressed      = 5000        ; save compressed coordinates every 10.0 ps
                          ; nstxout-compressed replaces nstxtcout
compressed-x-grps       = System      ; replaces xtc-grps
; Bond parameters
continuation                 = yes               ; Restarting after NPT
constraint_algorithm    = lincs                 ; holonomic constraints
constraints          = h-bonds  ; all bonds (even heavy atom-H bonds) constrained
lincs_iter           = 1                     ; accuracy of LINCS
lincs_order          = 4                     ; also related to accuracy
; Neighborsearching
cutoff-scheme       = Verlet
ns_type                  = grid              ; search neighboring grid cells
nstlist                  = 10           ; 20 fs, largely irrelevant with Verlet scheme
rlist            = 1.2
```

```
rcoulomb        = 1.2                   ; short-range electrostatic cutoff (in nm)
rvdw                    = 1.2           ; short-range van der Waals cutoff (in nm)
rvdw-switch     = 1.0
vdwtype         = cutoff
vdw-modifier    = force-switch
; Electrostatics
coulombtype             = PME                   ; Particle Mesh Ewald for long-range electrostatics
pme_order       = 4                 ; cubic interpolation
fourierspacing          = 0.16                  ; grid spacing for FFT
; Temperature coupling is on
tcoupl                  = V-rescale             ; modified Berendsen thermostat
tc-grps                 = Protein Non-Protein   ; two coupling groups - more accurate
tau_t                   = 0.1   0.1             ; time constant, in ps
ref_t                   = 300   300             ; reference temperature, one for each group, in K
; Pressure coupling is on
pcoupl                  = Parrinello-Rahman             ; Pressure coupling on in NPT
pcoupltype      = isotropic             ; uniform scaling of box vectors
tau_p                   = 2.0                           ; time constant, in ps
ref_p                   = 1.0                           ; reference pressure, in bar
compressibility         = 4.5e-5                ; isothermal compressibility of water, bar^-1
; Periodic boundary conditions
pbc             = xyz               ; 3-D PBC
; Dispersion correction
DispCorr        = No                ; account for cut-off vdW scheme
; Velocity generation
gen_vel                 = no                ; Velocity generation is off
```