Supplementary information

Accurate determination of protein:ligand standard binding free energies from molecular dynamics simulations

Haohao Fu,[1] Haochuan Chen,[1] Marharyta Blazhynska,[2] Emma Goulard Coderc De Lacam,[2] Florence Szczepaniak,[2,3] Anna Pavlova,[4] Xueguang Shao,[1] James C. Gumbart,[4] François Dehez,[2] Benoît Roux, [3,5,6] Wensheng Cai,[1,*] and Christophe Chipot[2,7,8,*]

[1]Research Center for Analytical Sciences, Frontiers Science Center for New Organic Matter, College of Chemistry, Nankai University, Tianjin Key Laboratory of Biosensing and Molecular Recognition, State Key Laboratory of Medicinal Chemical Biology, Tianjin 300071, China

[2]Laboratoire International Associé CNRS and University of Illinois at Urbana−Champaign, UMR n°7019, Université de Lorraine, BP 70239, F-54506 Vandœuvre-lès-Nancy, France

[3]Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637-1454, United States

[4]School of Physics, Georgia Institute of Technology, Atlanta, GA 30332, United States

[5]Department of Chemistry, University of Chicago, Chicago, Illinois 60637-1454, United States

[6]Center for Nanoscale Materials, Argonne National Laboratory, Argonne, Illinois 60439-8643, United States

[7]Department of Physics, University of Illinois at Urbana−Champaign, 1110 West Green Street, Urbana, Illinois 61801, United States

[8]Theoretical and Computational Biophysics Group, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801-2325, United States

Supplementary Table 1 | Experimental and estimated binding free energies of protein:ligand complexes shown in Table 2 (main text).

Complex	**PDB ID**[a]	**Δ*G*(exp) (kcal/mol)**	**Δ*G*(calc) (kcal/mol)**
Abl kinase-SH3:p41	1BBZ	-7.99	-7.6 (Geometrical route, this study) -7.9 ± 0.2 (Geometrical route, combination of refs 9, 14 and 16 of the main text) -7.5 ± 0.4 (Alchemical route)
Abl kinase-SH3:p5	—	-5.52	-5.4 ± 0.3
Abl kinase-SH3:p24	—	-5.97	-5.7 ± 0.5
Human $p56^{lck}$-SH2:PYEEI	1LKK	-7.1~-9.5	-8.8[a]
FKBP12:ligand3	—	-8.4	-9.0 ± 1.1
FKBP12:ligand5	—	-9.5	-11.6 ± 2.1
FKBP12:ligand6	—	-10.8	-9.7 ± 2.8
FKBP12:ligand8	1FKG	-10.9	-10.3 ± 0.4
FKBP12:ligand9	1FKH	-11.1	-11.6 ± 1.0
DIAP1-BIR1:grim peptide	1SE0	-9.5	-8.7 ± 0.7
Trypsin:benzamidine	3ATL	-6.4~-7.3	-7.8 ± 0.6
HBV Cp:NVR-010-001-E2	5E0I	-9.9	-11.5 ± 1.1
T4 lysozyme L99A:benzene	4W52	-5.2	-6.0 ± 1.0
T4 lysozyme L99A:ethyl benzene	1NHB	-5.76	-5.2 ± 0.2
T4 lysozyme L99A:paraxylene	187L	-4.67	-4.3 ± 0.4
T4 lysozyme L99A:n-butyl benzene	186L	-6.7	-5.7 ± 0.4
human CREBBP-Bromodomain:dihydroquinoxalinone	4NYX	-7.6	-7.3 ± 0.7
human CREBBP-Bromodomain: isoxazolyl-benzimidazole	4NR7	-10.5	-10.1 ± 0.5
Factor Xa:cation inhibitor	2JKH	-11.0	-10.1 ± 0.6
Factor Xa:cation inhibitor	2Y5H	-7.9	-6.8 ± 0.4
Factor Xa:cation inhibitor	2Y5G	-9.3	-7.2 ± 0.6
Factor Xa:quaternary ammonium	2BOK	-9.0	-8.7 ± 0.4
DNA:netropsin	—	−12.7	-13.2 ± 2.0
MDM2-p53: NVP-CGM097	4ZYF	-11.8	-11.3 ± 0.9
MUP-I:2-methoxy-3-isopropylpyrazine	1QY2	-7.8	-7.8 ± 1.0
MUP-I:6-hydroxy-6-methyl-3-heptanone	1I05	-6.0	-5.5 ± 0.7
β1-adrenergic receptor:4-methyl-2-(piperazin-1-yl) quinoline	3ZPR	-9.07	-8.1 ± 1.0
V_1-ATPase:ATP (tightly bound)	-	-13.24	-11.6 ± 0.8[b]

V_1-ATPase:ATP (bound)	-	-9.96	-8.9 ± 1.1[b]
V_1-ATPase:ATP (empty)	-	-6.08	-4.1 ± 1.1[b]
V_1-ATPase:ADP+P_i (tightly bound)	-	-6.19	-8.3 ± 0.9[b]
V_1-ATPase:ADP+P_i (empty)	-	-6.49	-4.3 ± 0.8[b]

[a]Error not accurately estimated. [b]The basis of comparison is F_1-ATPase.

Supplementary Protocols

Installation of software

BFEE2

We suggest to install BFEE2 through conda by

```
conda create --name bfee   (optional)
conda activate bfee        (optional)
conda install -c conda-forge BFEE2
```

BFEE2 can be also install through pip through

```
pip install BFEE2
```

However, the pip version of BFEE2 may not work well sometimes, as some of the pip versions of PySide2 are not very stable on Windows PCs.

NAMD

The binary of NAMD 3.0 can be downloaded from https://www.ks.uiuc.edu/Development/Download/download.cgi?PackageName=NAMD.

If the end-user wants to use the new definition of Euler and spherical-coordinate angles ("*Compatibility* → *Use quaternion-based CVs*" in BFEE2), compiling NAMD patched by the git version of Colvars is required. We show the steps of compiling a single node + GPU version of NAMD patched by the git version of Colvars below:

(1) Download the latest version of Colvars at http://colvars.github.io/

(2) Download the latest version of NAMD (devel branch) at https://gitlab.com/tcbgUIUC/namd

(3) Uncompress Colvars and NAMD:

```
tar xvf NAMD-devel.tar.gz
tar xvf colvars-master.tar.gz
```

(4) Patch NAMD using Colvars:

```
sh ./colvars-master/update-colvars-code.sh ./NAMD-devel
```

(5) Move to the folder of NAMD source code and untar charm++:

```
tar xvf charm-**********.tar
```

If charm-**********.tar is not found, copy it from the tarball of NAMD source code downloaded at https://www.ks.uiuc.edu/Development/Download/download.cgi?PackageName=NAMD.

(6) Then download TCL and FFTW libraries required by NAMD:

```
wget http://www.ks.uiuc.edu/Research/namd/libraries/fftw-linux-x86_64.tar.gz
tar xzf fftw-linux-x86_64.tar.gz
mv linux-x86_64 fftw
wget http://www.ks.uiuc.edu/Research/namd/libraries/tcl8.5.9-linux-x86_64.tar.gz
wget http://www.ks.uiuc.edu/Research/namd/libraries/tcl8.5.9-linux-x86_64-threaded.tar.gz
tar xzf tcl8.5.9-linux-x86_64.tar.gz
tar xzf tcl8.5.9-linux-x86_64-threaded.tar.gz
mv tcl8.5.9-linux-x86_64 tcl
mv tcl8.5.9-linux-x86_64-threaded tcl-threaded
```

(7) Move to the folder of charm++ source codes, and type

```
./build
```

Then answer some questions before starting building charm++. An example of how to answer these questions is shown below:

```
Are you building to run just on the local machine, and not across multiple nodes? (Y)
Do you want to build Charm++ with GPU Manager support for CUDA? (N)
Do you want to specify a compiler? (N)
Do you want to specify any Charm++ build options, such as Fortran compilers? (N)
Choose a set of compiler flags (3)
        1) none
        2) debug mode                       -g -O0
        3) production build [default]        --with-production
        4) production build w/ projections    --with-production --enable-tracing
        5) custom
What do you want to build? (1)
        1) Charm++ [default] (choose this if you are building NAMD)
        2) Charm++ and AMPI
        3) Charm++, AMPI, ParFUM, FEM and other libraries
Do you want to compile in parallel? (5)
        1) No
        2) Build with -j2
        3) Build with -j4
```

```
4) Build with -j8
5) Build with -j16 [default]
6) Build with -j32
7) Build with -j
Do you want to start the build now? (Y)
```

(8) (Ubuntu users) Add the option `-no-pie` to `CXXOPTS` in the relevant arch file (eg. Linux-x86_64-g++.arch).

(9) Move to the folder of NAMD source code, type:

```
./config Linux-x86_64-g++ --with-cuda –with-single-node-cuda --cuda-prefix /usr/local/cuda
```

The location of CUDA should be changed correspondingly if it is not at `/usr/local/cuda`.

(10) `cd Linux-x86_64-g++/`

Open `Make.config`, change `CHARM = $(CHARMBASE)/$(CHARMARCH)` to the absolute path of charm++ (e.g. `CHARM = /home/********/NAMD-devel/charm-6.10.2`). Then, type

```
make depends
make -j16
```

in terminal to compile NAMD.

VMD

The binary of VMD can be download from https://www.ks.uiuc.edu/Development/Download/download.cgi?PackageName=VMD and its installation guide can be found at https://www.ks.uiuc.edu/Research/vmd/current/ig/node6.html. There is no specific reason to compile VMD from its source code for the end-user.

Gromacs

A version of Gromacs patched with the git version of Colvars is required to finish this protocol. The process of compiling such version of Gromacs is shown below:

(1) Download Gromacs and git version of Colvars (one should check which version of Gromacs is supported by Colvars) at
https://manual.gromacs.org/documentation/2020.6/download.html
and
http://colvars.github.io/, respectively.

(2) Uncompress Gromacs and Colvars by

```
tar xvf gromacs-2020.6.tar.gz
tar xvf colvars-master.tar.gz
```

(3) Patch Gromacs with Colvars by

```
sh ./colvars-master/update-colvars-code.sh ./gromacs-2020.6
```

(4) Compile Gromacs by

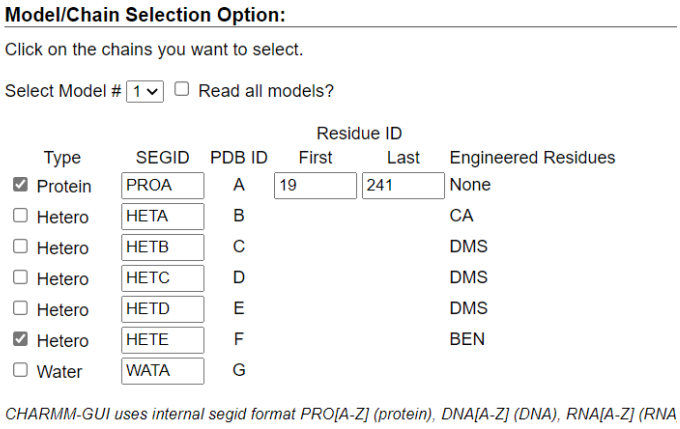
```
cd gromacs-2020.6
mkdir build
cd build
cmake .. -DGMX_BUILD_OWN_FFTW=ON -DGMX_GPU=on -
DCMAKE_INSTALL_PREFIX=****/*****
make -j12
make install
```

Generating structure and topology files using CHARMM-GUI

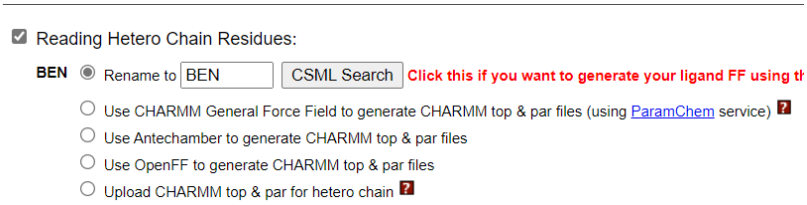
(1) Visit https://charmm-gui.org/. Registration and login may be required.

(2) Click “Input Generator→Solution Builder”. Either type PDB ID (e.g. 3ATL) in “Download PDB File” or upload a PDB file. Then click “Next Step”

(3) Select the protein and ligand that will be included in the molecular assembly. Deselect unnecessary water and other small molecules and then click “Next Step”.

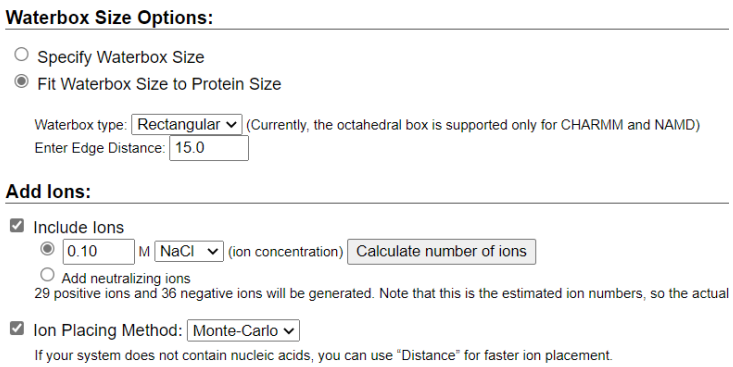


(4) If drug-like molecule is included, click “CSML Search” to search (or generate) the force field of the molecule.



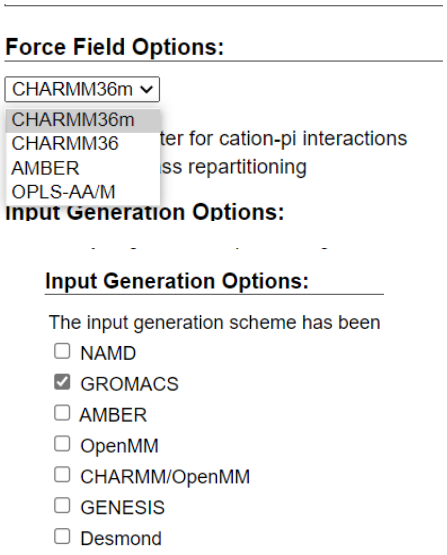
(5) Set disulfide bonds and protonation states if necessary. Then click “Next Step”.

(6) Set the size of water box and the concentration of ions. Then click “Next Step”.



Click “Next Step” again in the “Periodic Boundary Condition” page.

(7) Select the desired force field. Check “GROMACS” if the end-user wants to generate Gromacs-formatted files. Then click “Next Step” and download the generated files.



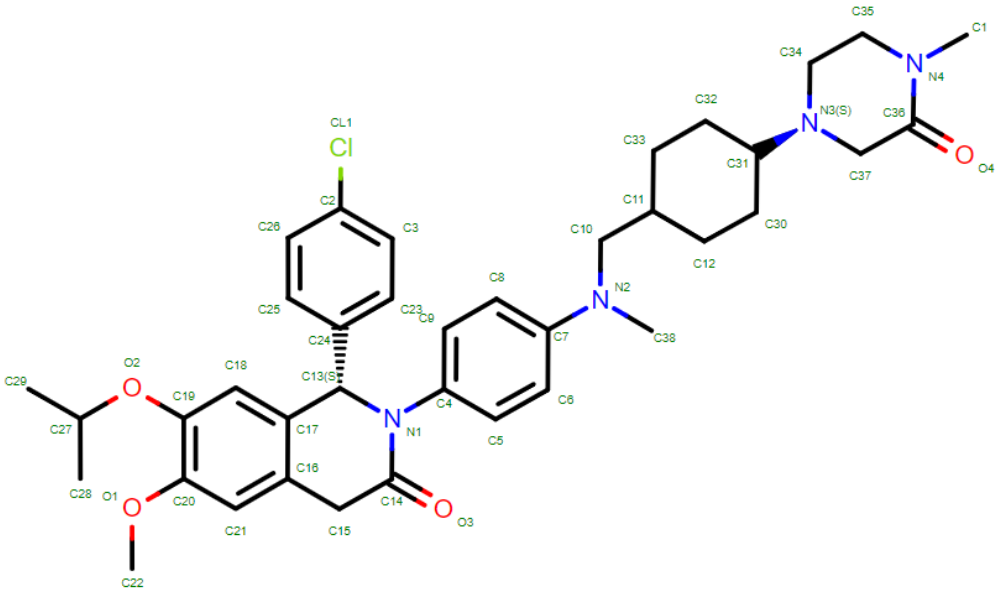
(8) If the ligand-only molecular system is needed (required if Gromacs is used as the MD engine), one can deselect the protein in step (2) and generate the structure and topology files again. If the protein:ligand complex solvated in a large water box is needed (required if the Amber force field is used), one can set a large “Edge Distance” in step (6) and generate the structure and topology files again.

Details of practical examples

MDM2-p53:NVP-CGM097 (4ZYF)

Description of the molecular assembly

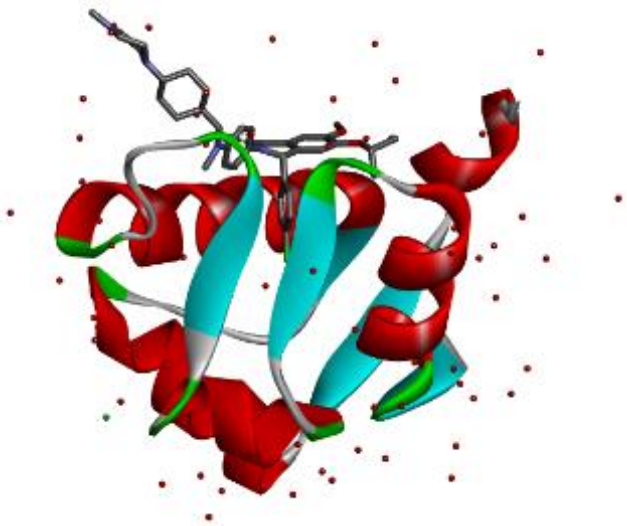
The protein of this protein-ligand molecular assembly is a protein−protein hybrid, MDM2-p53 (the former also known as HDM2, mouse double minute 2 homolog, which is an E3 ubiquitin ligase known to target p53 for degradation by the proteasome). The ligand of the molecular assembly is the p53:MDM2 inhibitor scaffold based on a dihydroisoquinolinone derivative ((S)-1-(4-chlorophenyl)-7-isopropoxy-6-methoxy-2-(4-(methyl(((1r,4S)-4-(4-methyl-3-oxopiperazin-1-yl)cyclohexyl)methyl)amino)phenyl)-1,2-dihydroisoquinolin-3(4H)-one), as shown in Supplementary Fig.1. This compound (known as well as NVP-CGM097) is a promising undergoing phase 1 clinical trials in p53wt tumors.



Supplementary Fig. 1 | The chemical structure of the ligand of the studied complex.

The ligand, namely NVP-CGM097, is rigid and semi-buried inside the protein (Supplementary Fig.2). The dihydroisoquinolinone scaffold of the ligand occupies the middle of the binding site and reaches the three critical binding pockets of the p53 residues Leu26, Trp23, and Phe19. The isopropyl ether at O2-C27-C28-C29 and the methyl ether (O1-C22) fill the Leu26 pocket, while the 4-chlorophenyl group reaches deep into the Trp23 binding cavity (Supplementary Fig.1). The N-methyl piperazinone motif binds toward the exit of the Phe19 (p53) binding cavity.

The experimental value of the binding free energy was obtained by ITC experiment and corresponds to -11.8 kcal/mol. The experimental snapshot was obtained within X-ray diffraction method. The resolution is 1.80 Å, which represents a good reproduction of the binding pose[1].



Supplementary Fig. 2 | Structure of MDM2-p53:NVP-CGM097 (4ZYF).

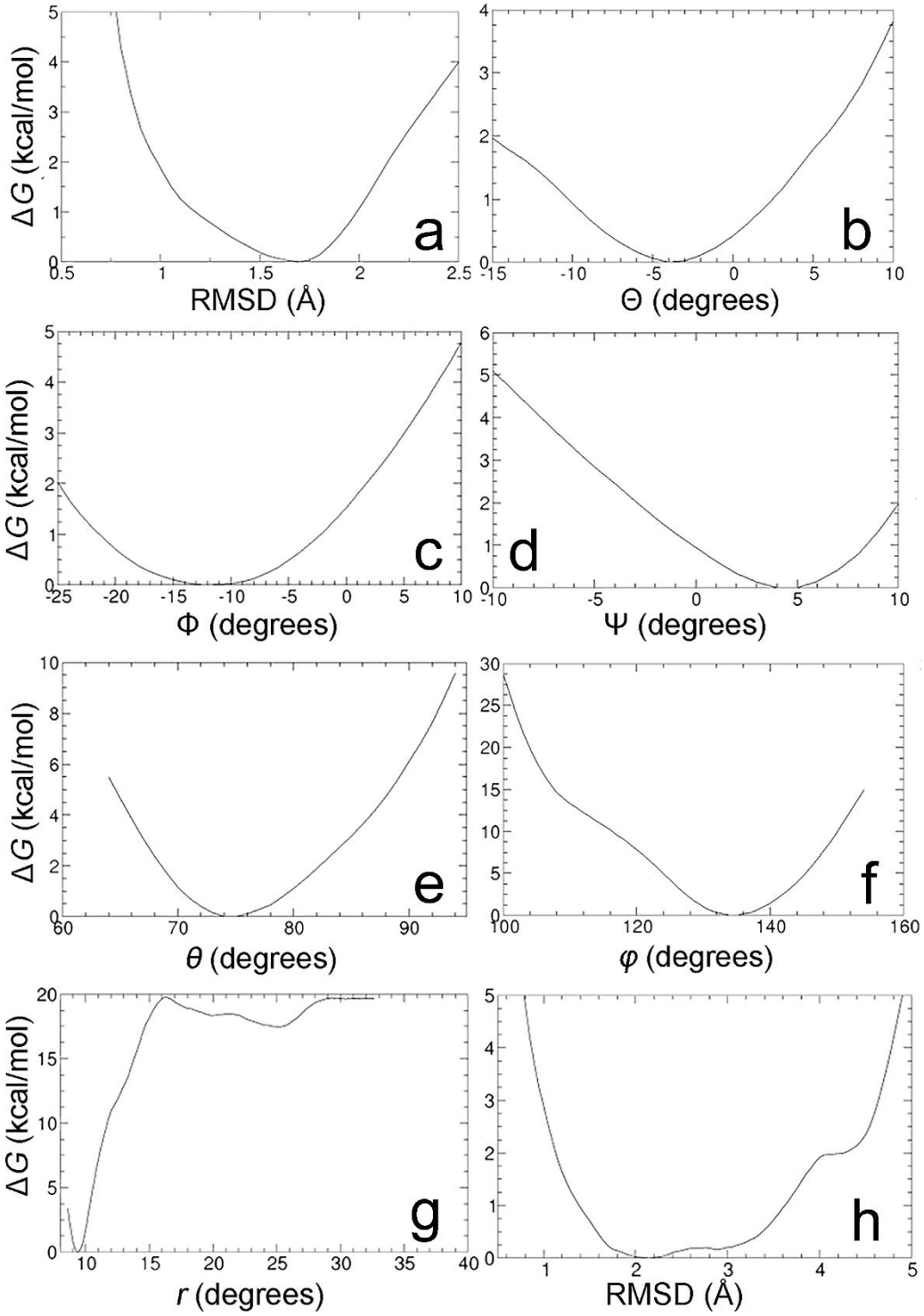
Computational details

The protein and water were described by all-atom CHARMM force field[2] with the TIP3P water model[3], the parameter files for the ligand were automatically generated through CGenFF program v2.4.0[4], with parameter and charge penalties of 44 and 36, respectively. The lone pair of the Chlorine atom were not included for the following simulation procedures. The crystallographic water was kept in the reconstruction of the complex. The obtained system consists of 45192 atoms in total, whereas the ligand has 95 atoms and the protein, 1570 atoms. The rest of the system was represented by counterions of sodium and chloride. The dimensions of the periodic cells for bound and unbound states were $87 \times 85 \times 84$ Å^3 .

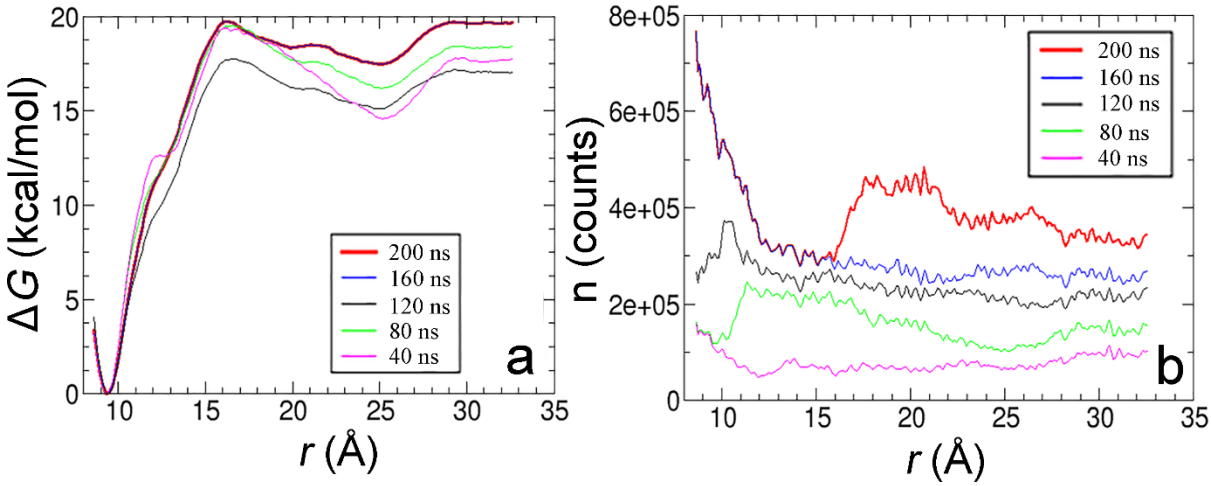
Before using the BFEE2 pre-treatment tool, the system was pre-equilibrated for 100 ns in NVT ensemble with NAMD 3.0 program[5], keeping the temperature at 298.15K (in accordance with the experimental conditions of the ITC experiment for the studied complex) and the pressure of 1 atm, using damped Langevin dynamics[6] and Langevin piston[7]. A time step of 2 fs was used for the integration of the equation of motion. A smoothed 12 Å spherical cutoff was applied to truncate van der Waals and short-range electrostatic interactions. The distance between pairs for inclusion in pair lists was 14 Å. For the long-range electrostatic interactions, the particle-mesh Ewald (PME) algorithm[8] was used. The obtained coordinates after the pre-equilibration were saved in a new PDB file and were used in the BFEE2 tool.

For the free energy calculations, the geometrical route was chosen at the BFEE2 pre-treatment section. By default, the PMF approach is based on well-tempered meta-eABF method. For the standard binding free energy calculations through the geometrical route, the procedure detailed in the main text was followed. It is important to check the convergence of the PMF calculations through the time-evolution of PMF RMSD with respect to a zero vector. The checkup of the continuous and asymptotic behavior of the curves convince that the sampling of the system was enough. Another convergence test is checking the distribution of sampling along the collective variable (`*.hist.count file`).

Results



Supplementary Fig. 3 | Individual PMFs for all components. The PMF calculations using RMSD of the ligand with respect to its bound-state conformation (a), Θ (b), Φ (c), Ψ (d), θ (e), φ (f), centers-of-mass distance between the ligand and protein (g), and RMSD of the ligand with respect to its bound-state conformation with the ligand in the unbound state (h), as the collective variable, respectively.



Supplementary Fig. 4 | Convergence of PMF calculation characterizing the separation of the ligand and the protein. Time-evolution of PMF (a) and of the distribution of sampling along collective variable (b).

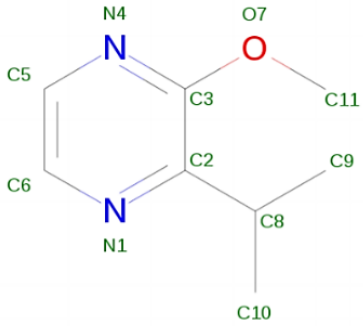
Supplementary Table 2 | Results for each contribution to the binding free energy of MDM2-p53:NVP-CGM097. The number of nanoseconds enough for the reasonable convergence of the components and computer time used to perform the simulation on 32 CPU cores and 2 GTX 2080Ti.

Contribution	Free energy (kcal/mol)	Simulation time (ns)	Speed (ns/day)
$\Delta G_{\mathrm{c}}^{\mathrm{site}}$	-7.2 ± 0.1	100	60
$\Delta G_{\Theta}^{\mathrm{site}}$	-0.3 ± 0.0	20	61
$\Delta G_{\Phi}^{\mathrm{site}}$	-0.5 ± 0.0	20	61
$\Delta G_{\Psi}^{\mathrm{site}}$	-0.2 ± 0.0	20	61
$\Delta G_{\theta}^{\mathrm{site}}$	-0.2 ± 0.0	20	44
$\Delta G_{\varphi}^{\mathrm{site}}$	-0.2 ± 0.0	20	42
$-\frac{1}{\beta}\ln(S^{*}I^{*}C^{\circ})$	-17.9 ± 0.6	180	36
$\Delta G_{\mathrm{c}}^{\mathrm{bulk}}$	8.6 ± 0.3	100	96
$\Delta G_{\mathrm{o}}^{\mathrm{bulk}}$	6.6		
ΔG_{bind}	-11.3 ± 0.9 (calculation) -11.8 (experiment[1])		

MUP-I:2-methoxy-3-isopropylpyrazine (1QY2)

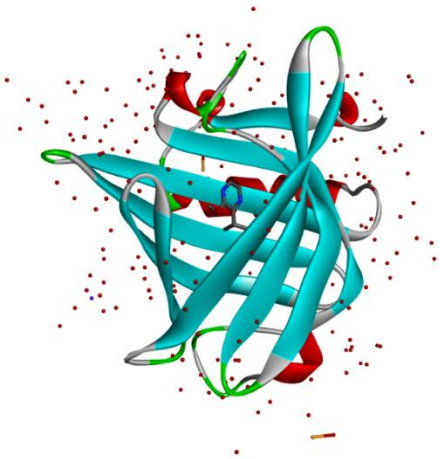
Description of the molecular assembly

The protein, MUP-I, is an abundant pheromone binding protein found in male mouse urine. The protein has a typical lipocalin fold that consists of an eight-stranded β-barrel and a single α-helix, and the interior of the barrel forms a hydrophobic cavity. The ligand of the system is 2-methoxy-3-isopropylpyrazine (Supplementary Fig. 5), a small hydrophobic molecule, which is located inside the cavity of the MUP-I protein.



Supplementary Fig. 5 | The chemical structure of 2-methoxy-3-isopropylpyrazine.

This ligand is buried inside the globular protein system (Supplementary Fig. 6). The pyrazine moiety is located inside the β-barrel, within a hydrophobic environment formed by the side chains of Phe38, Leu40, Leu42, Ile45, Leu54, Phe56, Met69, Val82, Tyr84, Phe90, Ala103, Leu105, Leu116, and Tyr120. The hydroxyl group of Tyr120 forms hydrogen bonds directly to one of the nitrogen atoms of the ring of the ligand. The experimental value of the binding free energy was obtained by ITC experiment and corresponds to -7.8 kcal/mol. The experimental snapshot was obtained with X-ray diffraction method, with the resolution of 1.75 Å.



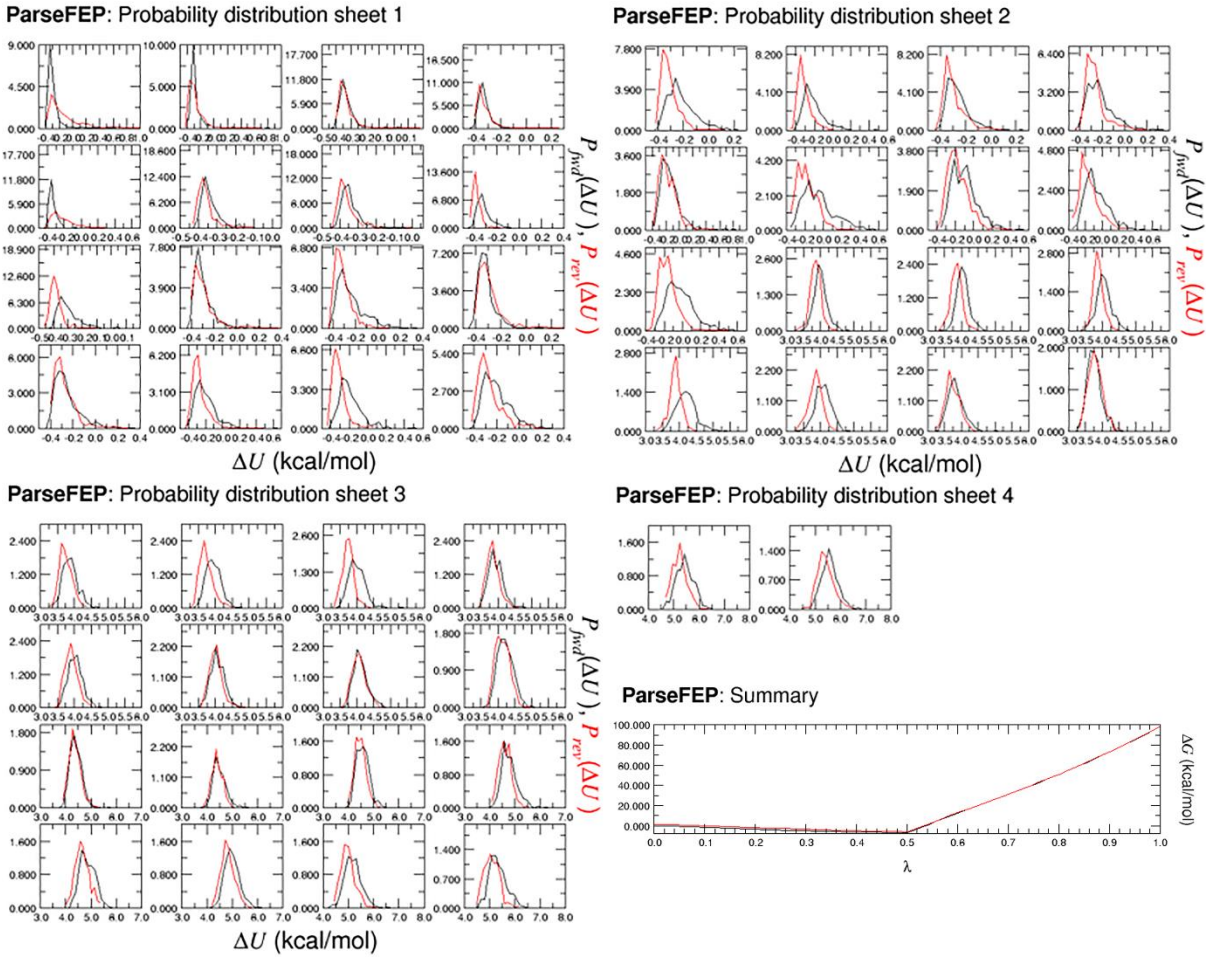
Supplementary Fig. 6 | Structure of MUP-I:2-methoxy-3-isopropylpyrazine (1QY2).

Computational details

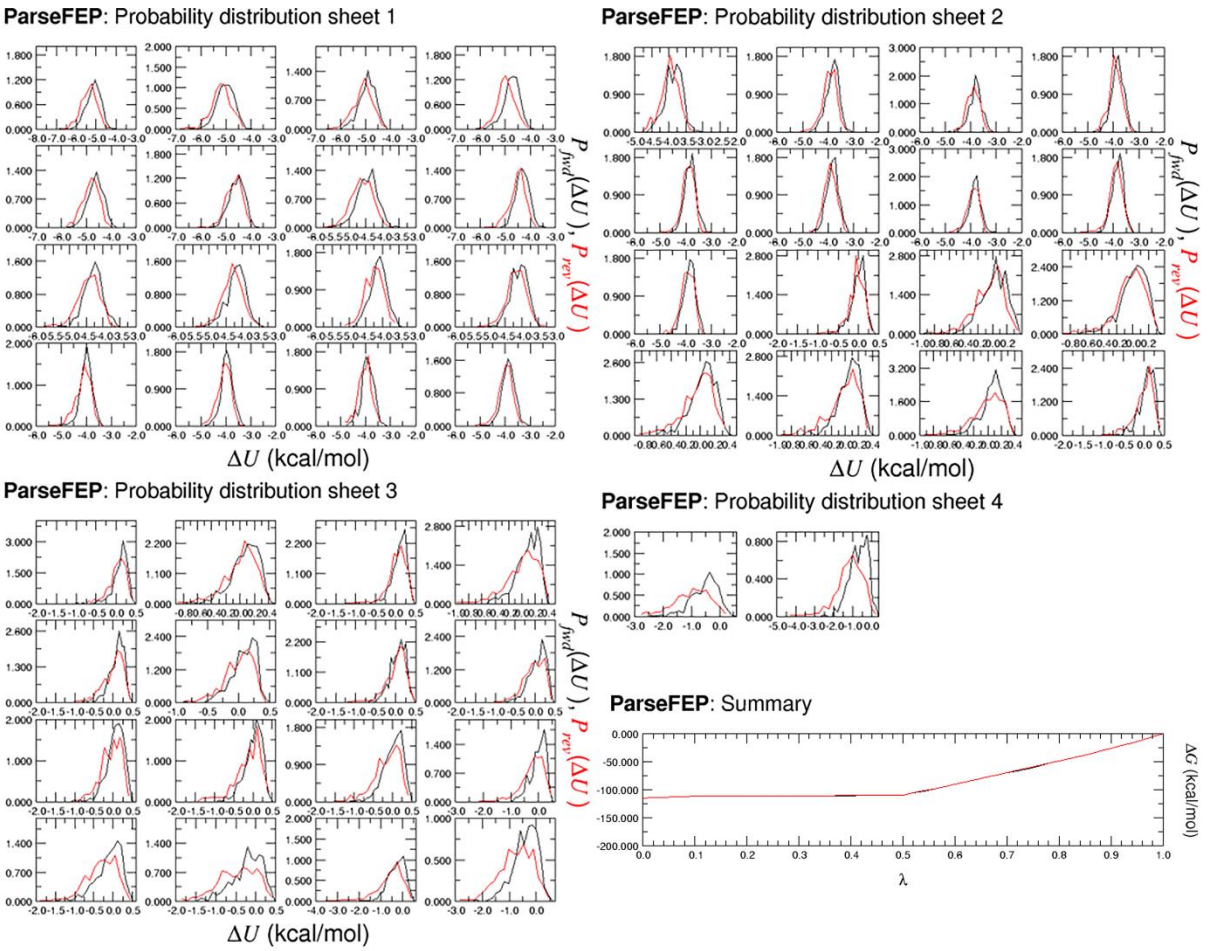
The procedure of modeling the molecular assembly is the same as that of the previous example (4ZYF). The molecular assembly consists of 55837 atoms in total and the dimensions of the water box was 87 × 85 × 84 $Å^3$.

For the free energy calculations, the alchemical route was chosen in the BFEE2 pre-treatment tool. In the advanced settings, a lambda schedule of 50 stratification windows for FEP calculations and 15 for TI calculations were chosen for both bound and unbound states. For FEP calculations, sampling time of 1 ns per window and equilibration prior to data collection of 0.2 ns per window were adopted. The electrostatic interaction was completely decoupled at lambda equal to 0.5 and the Van der Waals interaction, 1.0. For the TI calculations, the total time of sampling was chosen as 2 ns per window, while other parameters were kept as the default.

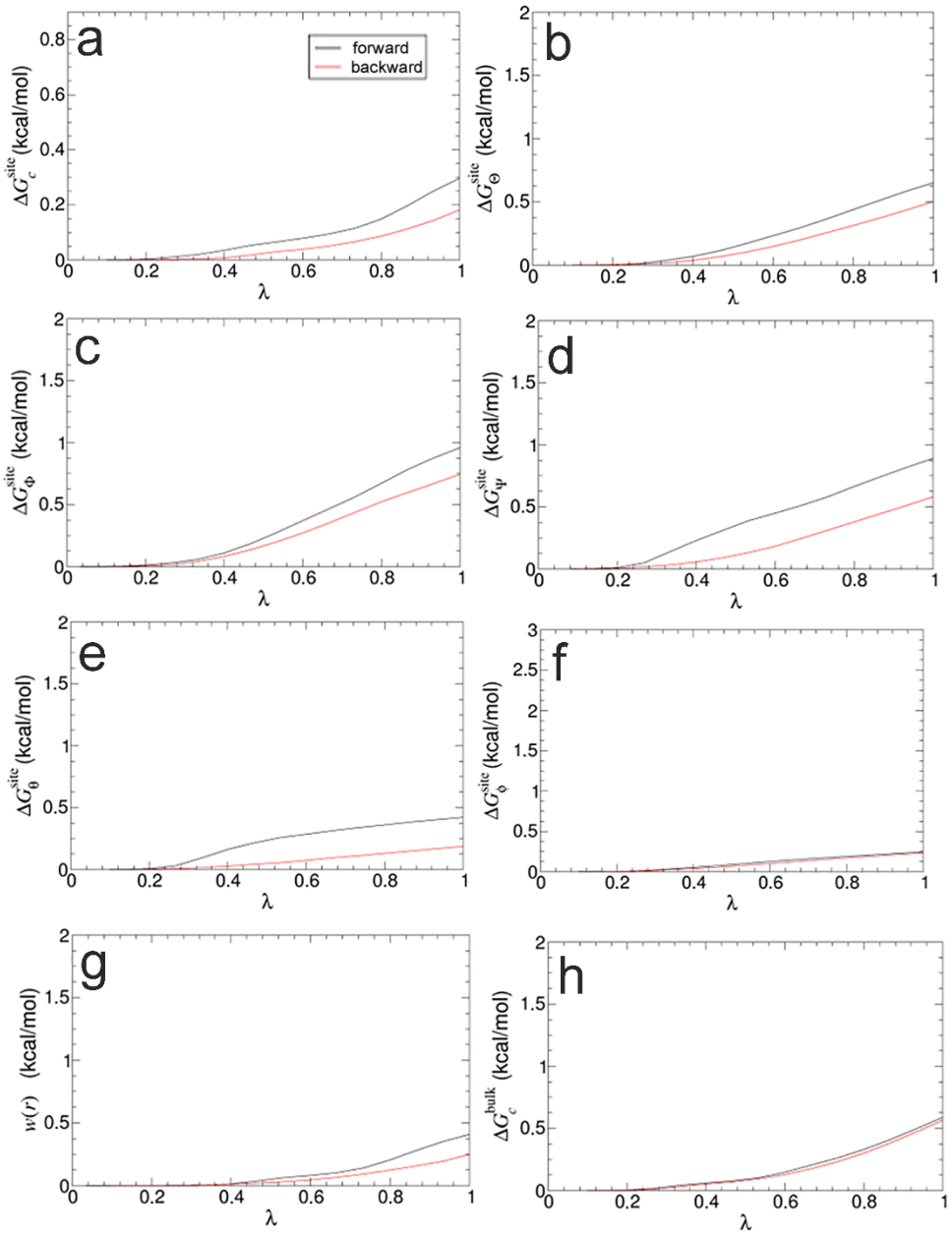
Results



Supplementary Fig. 7 | Distribution of sampling in each window for reversible decoupling the ligand in the bound state for MUP-I:2-methoxy-3-isopropylpyrazine and the hysteresis between the forward and backward transformation, mirrored by the free-energy change with respect to *λ*.



Supplementary Fig. 8 | Distribution of sampling in each window for reversible decoupling the ligand in the unbound state for MUP-I:2-methoxy-3-isopropylpyrazine and the hysteresis between the forward and backward transformation, mirrored by the free-energy change with respect to λ.



Supplementary Fig. 9 | Free-energy change accounting for reversible introduction of restraints to the ligand in the bound state for MUP-I:2-methoxy-3-isopropylpyrazine (a-g), and the ligand in the unbound state (h), respectively.

Supplementary Table 3 | Results for each contribution to the binding free energy of MUP-I:2-methoxy-3-isopropylpyrazine. The number of nanoseconds enough for the reasonable convergence of the components

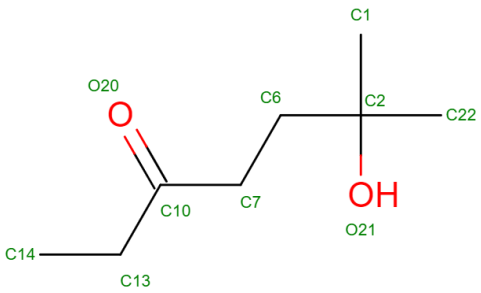
Contribution	Free energy (kcal/mol)	Simulation time (ns)	Speed (ns/day)
$\Delta G^{\text{site}}_{\text{decouple}}$	97.6 ± 0.9	100	22.0
$\Delta G^{\text{site}}_{\text{c}}$	-0.3 ± 0.1		
$\Delta G^{\text{site}}_{\Theta}$	-0.6 ± 0.1		
$\Delta G^{\text{site}}_{\Phi}$	-0.9 ± 0.2		
$\Delta G^{\text{site}}_{\Psi}$	-0.7 ± 0.2	64	25.8
$\Delta G^{\text{site}}_{\theta}$	-0.4 ± 0.1		
$\Delta G^{\text{site}}_{\varphi}$	-0.2 ± 0.0		
$\Delta G^{\text{site}}_{r}$	-0.3 ± 0.1		
$\Delta G^{\text{bulk}}_{\text{decouple}}$	-114.9 ± 0.1	100	54.4
$\Delta G^{\text{bulk}}_{\text{c}}$	0.6 ± 0.0	64	95.1
$\Delta G^{\text{bulk}}_{\text{o+r+a}}$	12.3		
ΔG_{bind}	-7.8 ± 1.0 (calculation) -7.8 kcal/mol (experiment[9])		

and computer time used to perform the simulation on 32 CPU cores and 2 GTX 2080Ti.

MUP-I:6-hydroxy-6-methyl-3-heptanone (1I05)

Description of the molecular assembly

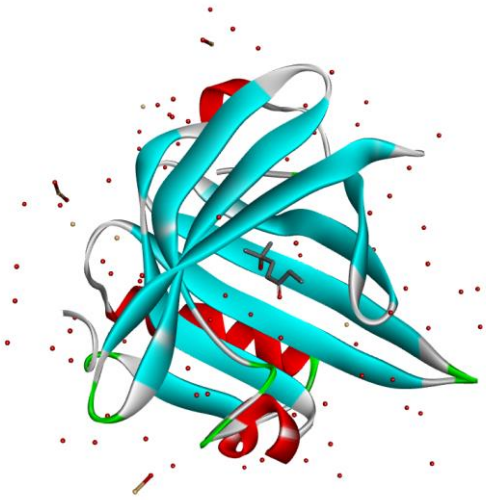
The protein of the molecular assembly, MUP-I, is an abundant pheromone binding protein found in male mouse urine. The protein has a typical lipocalin fold that consists of an eight-stranded β-barrel and a single α-helix, and the interior of the barrel forms a hydrophobic cavity. The corresponding ligand is 6-hydroxy-6-methyl-3-heptanone (Supplementary Fig. 10), a small hydrophobic molecule, which binds the MUP-I protein within the hydrophobic environment at one end of the β-barrel formed by the side chains of Phe56, Leu58, Leu60, Ile63, Leu72, Phe74, Met87, Val100, Tyr102, Phe108, Ala121, Leu123, Leu134, and Tyr138.



Supplementary Fig. 10 | Chemical structure of 6-hydroxy-6-methyl-3-heptanone.

A snapshot of the globular protein:buried ligand complex is provided in Supplementary Fig. 11. The outermost side chains of the ligand binding site (Met87, Phe56, and Leu58) are in van der Waals contact with the side chain of Tyr102, which caps the entrance to the interior of the β-barrel. Two water molecules are also present in the ligand binding site, interacting with the hydroxyl group of Tyr138 and the carbonyl

oxygen atoms of Phe56 and Leu58.



Supplementary Fig. 11 | Structure of MUP-I:6-hydroxy-6-methyl-3-heptanone (1I05)

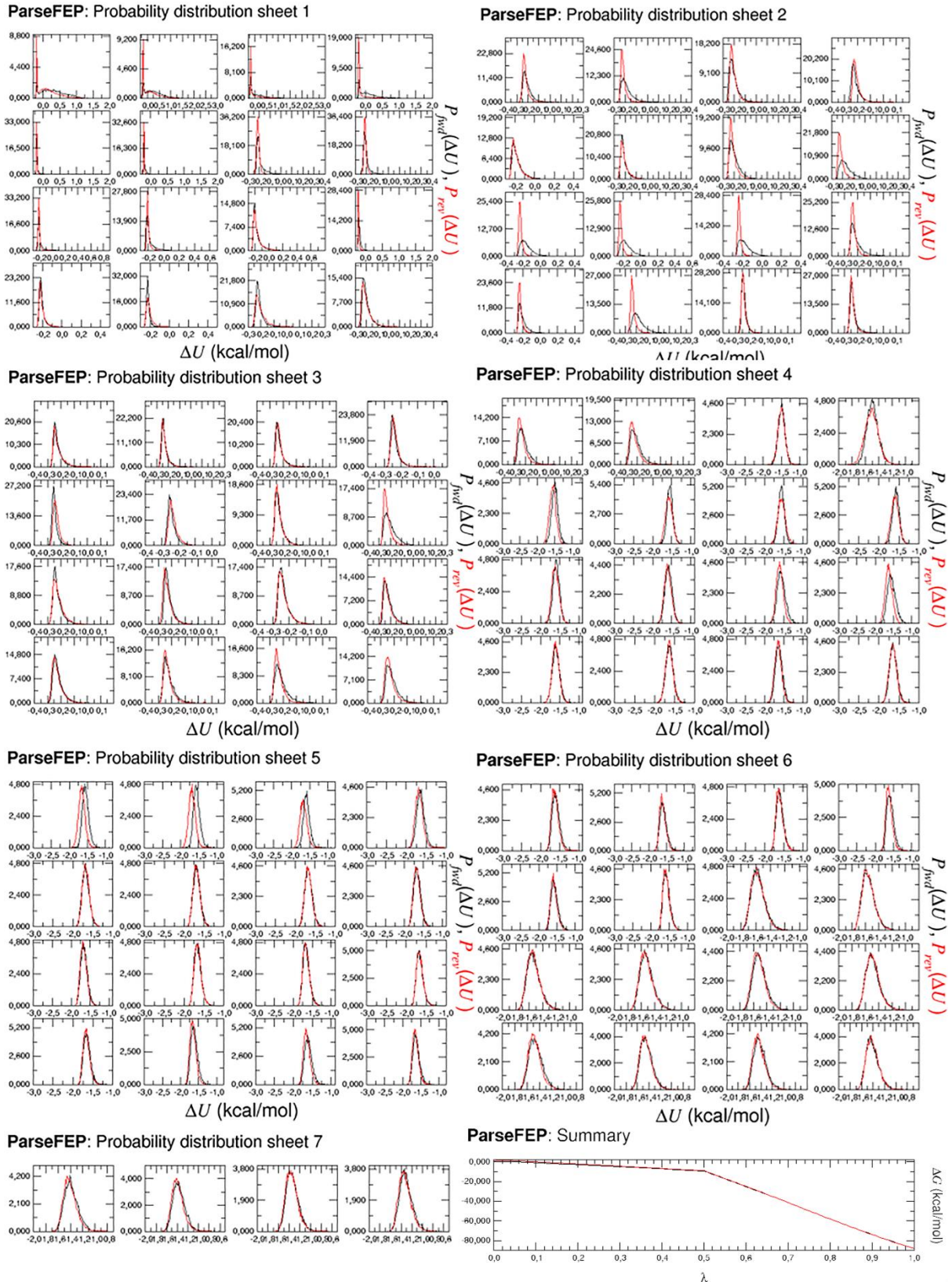
The experimental value of the binding free energy was obtained by ITC experiment at 30°C and corresponds to -6 ± 1 kcal/mol, wherein ΔH = -13 kcal/mol and $-T\Delta S$ = +7 kcal/mol. The PDB structure was obtained within the X-ray diffraction method. The resolution is 2.00 Å, which represents a sufficient reproduction of the binding pose.

Computational details

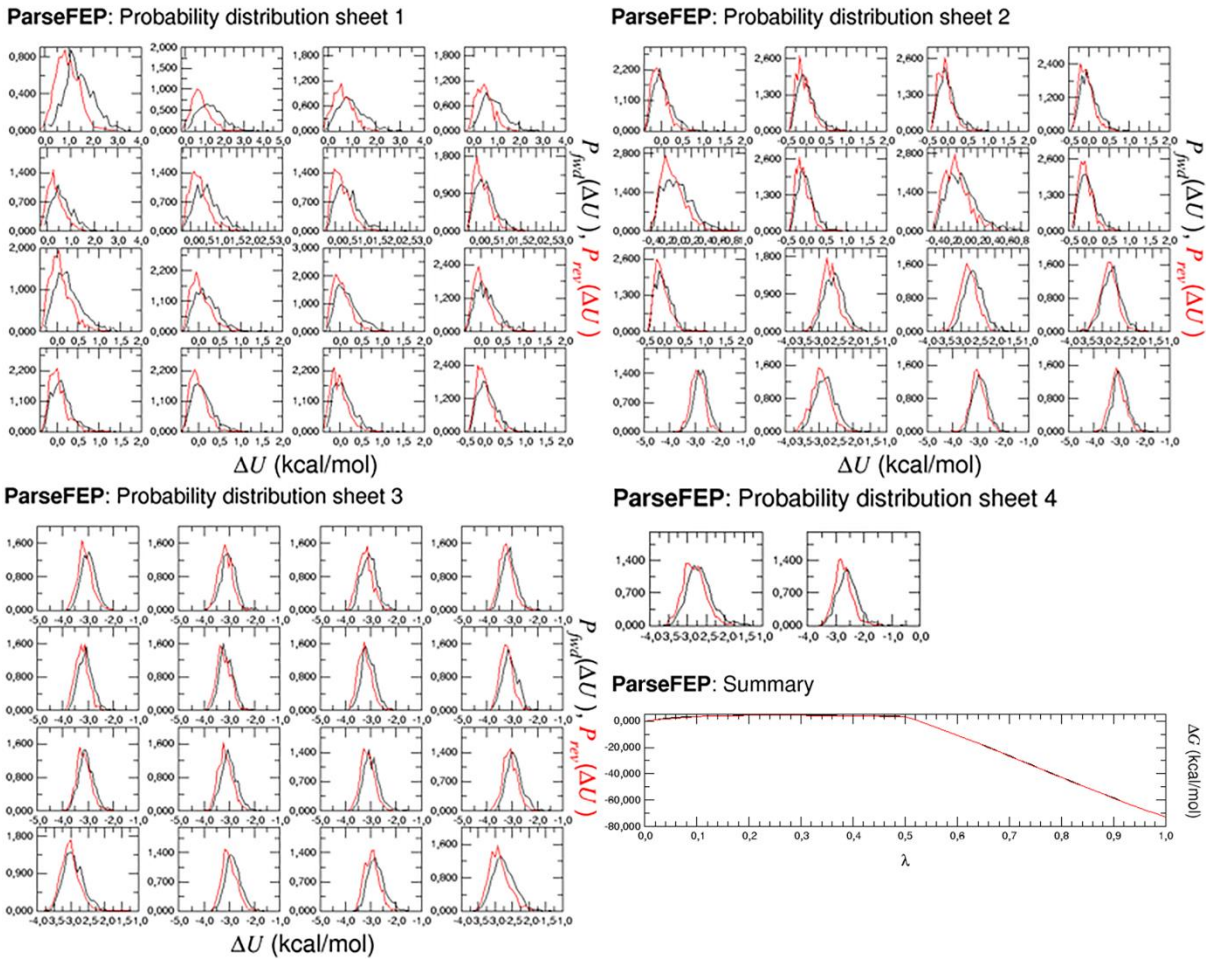
The crystal structure of MUP-I:6-hydroxy-6-methyl-3-heptanone is directly provided to BFEE2 without pre-equilibration, as the binding event is driven by enthalpy and we want to guarantee the most favorable structure before starting the free-energy calculation.

To prevent the molecular assembly from reformation of a buried water-mediated hydrogen bond network between the protein and ligand, two water molecules were treated as "part of the protein" by restraining them from leaving their hydrogen-bonding partners (F56, L58, T39, and Y138, force constant: 0.3 kcal/(mol·Å²)). The soft CV force constant of 0.3 kcal/(mol·Å²). To avert the isomerization of the ligand inside the binding pocket, we increased the force constant restraining the RMSD of the ligand with respect to its bound-state conformation to 100 kcal/(mol·Å²).

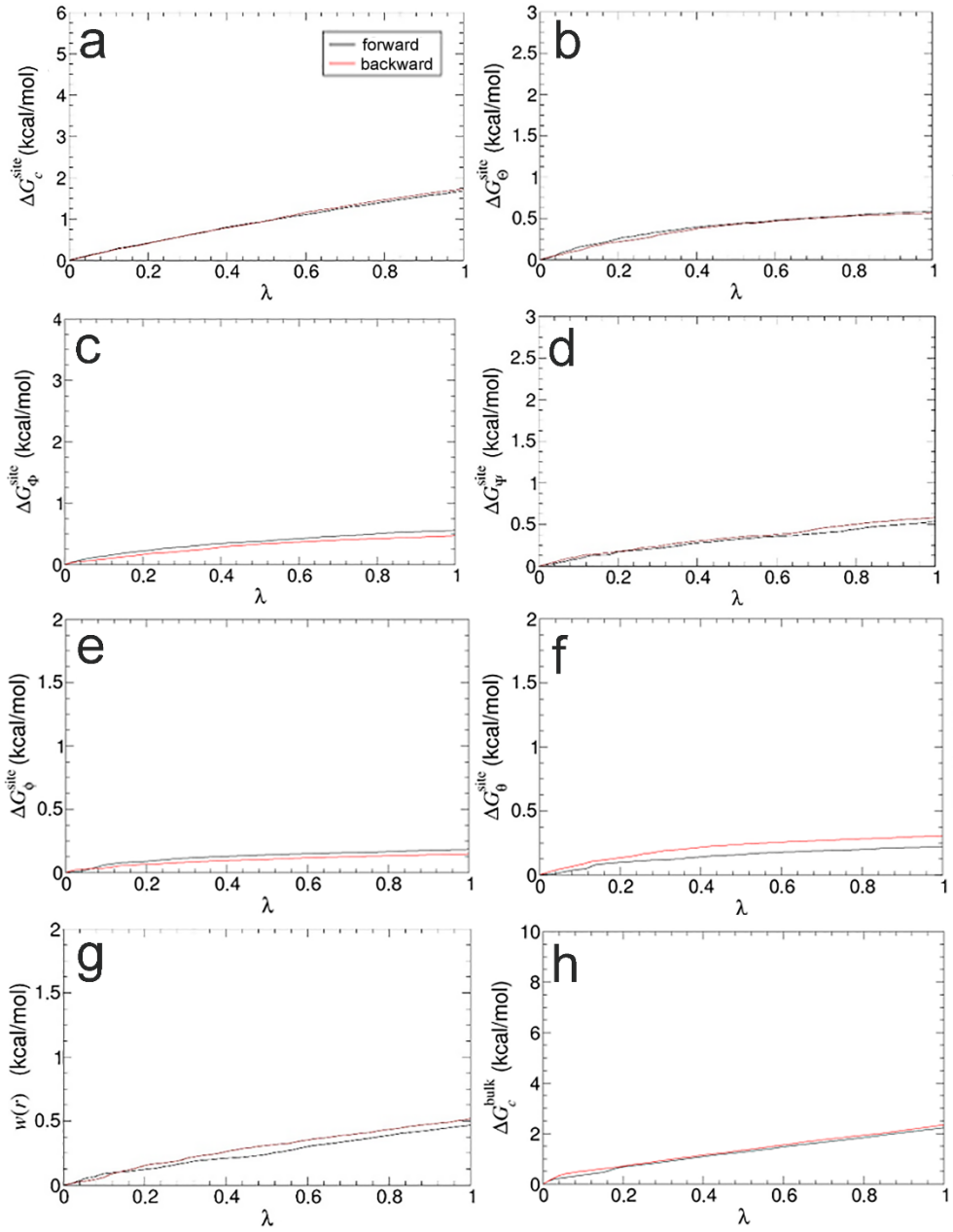
100, 100, 50 and 100 strata were adopted for simulations characterizing the reversible decoupling the ligand from the protein, adding restraints on the bound-state ligand, decoupling the ligand from the bulk water and adding restraints on the unbound-state ligand, respectively. The simulation times are 2, 1, 1, 1 ns per window for the aforementioned four simulations, respectively. Other parameters are the same as those of the previous example (1QY2).



Supplementary Fig. 12 | Distribution of sampling in each window for reversible decoupling the ligand in the bound state for MUP-I:6-hydroxy-6-methyl-3-heptanone and the hysteresis between the forward and backward transformation, mirrored by the free-energy change with respect to λ.



Supplementary Fig. 13 | Distribution of sampling in each window for reversible decoupling the ligand in the unbound state for MUP-I:6-hydroxy-6-methyl-3-heptanone and the hysteresis between the forward and backward transformation, mirrored by the free-energy change with respect to λ.



Supplementary Fig. 14 | Free-energy change accounting for reversible introduction of restraints to the ligand in the bound state for MUP-I:6-hydroxy-6-methyl-3-heptanone (a-g), and the ligand in the unbound state (h), respectively.

Supplementary Table 4 | Results for each contribution to the binding free energy of MUP-I:6-hydroxy-6-methyl-3-heptanone. The number of nanoseconds enough for the reasonable convergence of the components and computer time used to perform the simulation on 32 CPU cores and 2 GTX 2080Ti.

Contribution	Free energy (kcal/mol)	Simulation time (ns)	Speed (ns/day)
$\Delta G^{\text{site}}_{\text{decouple}}$	-88.7 ± 0.7	200	27.9
$\Delta G^{\text{site}}_{\text{c}}$	-1.8 ± 0.0		
$\Delta G^{\text{site}}_{\Theta}$	-0.5 ± 0.0	0.1	24.0
$\Delta G^{\text{site}}_{\Phi}$	-0.5 ± 0.0		

$\Delta G_{\Psi}^{\text{site}}$	-0.5 ± 0.0		
$\Delta G_{\theta}^{\text{site}}$	-0.3 ± 0.0		
$\Delta G_{\varphi}^{\text{site}}$	-0.2 ± 0.0		
$\Delta G_{r}^{\text{site}}$	-0.5 ± 0.0		
$\Delta G_{\text{decouple}}^{\text{bulk}}$	72.8 ± 0.1	100	49.0
$\Delta G_{\text{c}}^{\text{bulk}}$	2.3 ± 0.0	0.1	40.0
$\Delta G_{\text{o+r+a}}^{\text{bulk}}$	12.4		
ΔG_{bind}	-5.5 ± 0.7 (calculation) -6.0 ± 1.0 (experiment[10])		

Supplementary References

1. Holzer, P. *et al.* Discovery of a dihydroisoquinolinone derivative (NVP-CGM097): A highly potent and selective MDM2 inhibitor undergoing phase 1 clinical trials in p53wt tumors. *J. Med. Chem.* **58**, 6348–6358 (2015).
2. Huang, J. *et al.* CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71 (2016).
3. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
4. Vanommeslaeghe, K. *et al.* CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31**, 671–690 (2010).
5. Phillips, J. C. *et al.* Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* **153**, 44130 (2020).
6. Uhlenbeck, G. E. & Ornstein, L. S. On the theory of the brownian motion. *Phys. Rev.* **36**, 823–841 (1930).
7. Feller, S. E., Zhang, Y., Pastor, R. W. & Brooks, B. R. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* **103**, 4613–4621 (1995).
8. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N · log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
9. Bingham, R. J. *et al.* Thermodynamics of binding of 2-methoxy-3-isopropylpyrazine and 2-methoxy-3-isobutylpyrazine to the major urinary protein. *J. Am. Chem. Soc.* **126**, 1675–1681 (2004).
10. Timm, D. E., Baker, L. J., Mueller, H., Zidek, L. & Novotny, M. V. Structural basis of pheromone binding to mouse major urinary protein (MUP-I). *Protein Sci.* **10**, 997–1004 (2001).