Supporting information for

Saturation mutagenesis by efficient free-energy calculation

Zuzana Jandova[1], Daniel Fast[1], Martina Setz[1], Maria Pechlaner[1,‡], Chris Oostenbrink[1*]

[1] Institute of Molecular Modeling and Simulation, University of Natural Resources and Life Sciences, Vienna

[‡] Current address: Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, ETH, 8093 Zürich, Switzerland

*chris.oostenbrink@boku.ac.at

Table S1: end-point energies ($\langle E \rangle$) and fluctuations ($\langle E^2 \rangle - \langle E \rangle^2$) uncharging calculations of sidechains in water for the charged state (λ=0) and neutral state (λ=1) in kJ/mol in Table 1.

Table S2: end-point energies ($\langle E \rangle$) and fluctuations ($\langle E^2 \rangle - \langle E \rangle^2$) of the uncharging calculations of sidechains in tripeptides for the charged state (λ=0) and neutral state (λ=1) in kJ/mol in Table 4.

Table S3: the percentage of contributing configurations in the OSP calculations for all amino acid sidechain analogues (reference states R1 and R2) and for all amino acids (reference states R3, R4, R5 and R6).

Table S4: comparison of the computational efforts of TI and the TPF+OSP approach.

Figure S1: Curves of convergence of fluctuations ($\langle \mathrm{E}^2 \rangle - \langle \mathrm{E} \rangle^2$) at the end points of methione sidechains uncharging. (charged state (λ=0) and neutral state (λ=1).

Data S1: Molecular topology building blocks for the reference states in Figure 2 of the main document.

	charged(Q) (λ=0)		neutral(N) (λ=1)	
	$\langle E \rangle$	$\langle E^2 \rangle - \langle E \rangle^2$	$\langle E \rangle$	$\langle E^2 \rangle - \langle E \rangle^2$
arg	338.69	35.37	-33.41	31.51
asn	129.62	22.94	-0.11	14.25
asp	684.48	41.92	33.87	34.18
cys	25.33	10.70	0.07	6.70
gln	129.85	22.76	-0.10	14.19
glu	680.25	41.23	33.50	34.21
hisa	124.11	20.37	-0.32	13.64
hisb	152.69	21.65	-0.71	15.76
lys	417.29	36.61	-33.21	34.60
met	34.95	11.44	-0.10	7.40
phe	22.10	8.91	-0.62	6.00
ser	80.67	17.96	0.06	9.33
thr	78.49	17.60	0.11	9.11
trp	72.76	15.50	-0.99	12.00
tyr	86.42	17.95	-0.37	10.38

Table S1: End-point energies ($\langle E \rangle$) and fluctuations ($\langle E^2 \rangle - \langle E \rangle^2$) of uncharging calculations of sidechains in water for the charged state (λ=0) and neutral state (λ=1) in Table 1 (in kJ/mol).

	charged(Q) (λ=0)		neutral(N) (λ=1)	
	$\langle E \rangle$	$\langle E^2 \rangle - \langle E \rangle^2$	$\langle E \rangle$	$\langle E^2 \rangle - \langle E \rangle^2$
arg	415.90	35.10	45.60	31.40
asn	328.80	22.50	198.80	14.50
asp	756.70	41.60	118.00	33.40
cys	33.70	10.60	8.30	6.60
gln	372.10	23.70	197.10	16.00
glu	747.10	41.20	104.50	33.70
hisa	101.90	20.40	-22.00	13.40
hisb	146.30	21.80	-2.80	15.60
lys	487.20	36.50	41.50	34.00
met	32.30	11.30	-2.00	7.40
phe	12.70	8.70	-7.90	5.80
ser	91.20	17.50	15.70	9.30
thr	86.50	17.10	17.40	9.80
trp	92.00	15.40	23.20	11.60
tyr	151.80	18.00	64.10	10.40

Table S2: End-point energies ($\langle E \rangle$) and fluctuations ($\langle E^2 \rangle - \langle E \rangle^2$) of uncharging calculations of sidechains in tripeptides for the charged state (λ=0) and neutral state (λ=1) in Table 4 (in kJ/mol).

	sidechain analogues		tripeptides			
	R1	**R2**	**R3**	**R4**	**R5**	**R5**
ala	49.85	-	79.14	82.23	83.68	83.23
arg	0.00	6.29	0.02	1.06	0.01	0.01
asn	5.00	57.96	17.14	38.26	15.97	15.84
asp	6.69	69.19	16.62	37.75	15.38	14.95
cys	18.23	80.92	45.11	55.24	45.17	43.98
gln	0.00	41.26	4.06	24.50	4.10	4.18
glu	1.63	57.01	2.62	19.65	2.40	2.44
hisa	0.53	43.37	1.19	16.20	1.12	1.15
hisb	0.32	33.44	1.14	14.21	1.13	1.08
ile	0.95	72.36	1.74	5.74	3.04	3.13
leu	1.65	53.16	4.38	19.57	4.32	4.34
lys	0.11	11.96	0.51	9.87	0.42	0.46
met	0.00	67.64	1.87	18.40	2.03	2.05
phe	0.03	18.95	0.08	4.01	0.05	0.05
ser	27.07	84.09	53.60	51.60	54.50	53.30
thr	9.46	93.44	21.73	24.80	25.64	26.33
trp	0.00	2.76	0.00	0.36	0.00	0.00
tyr	0.00	12.12	0.04	2.61	0.03	0.03
val	7.38	82.34	9.38	11.98	11.96	12.12

Table S3: Percentage of configurations significantly contributing to the value ($\Delta G_{R>N}^{OSP}$) in OSP. The number of configurations is counted for which ($H_N - H_R$) in equation (8) is less than $\Delta G_{N>R}^{OSP} + k_B T$.

ns	TI			TPF+OSP		
	Q>N	**N>A**	**total**	**Q>N**	**N>A**	**total**
1 AA	29.2	11.2	40.4	20.0	50.0	70.0
17 AA	408.5	190.7	**599.3**	280.6	50.0	**330.6**

Table S4: Computational resources spent on calculation of the free energy of mutation into alanine. The value for TI is the result of 14 uncharging TI calculations and 17 annihilation TI calculations of mutations to Ala in the tripeptides. In OSP only one 50 ns simulation of R5 reference in tripeptide was included, while for TPF 14 uncharging calculations (two simulations each) are taken into account. The total amount of ns for every process is labelled by Q>N and N>A for the calculation of uncharging and annihilation free energies, respectively.

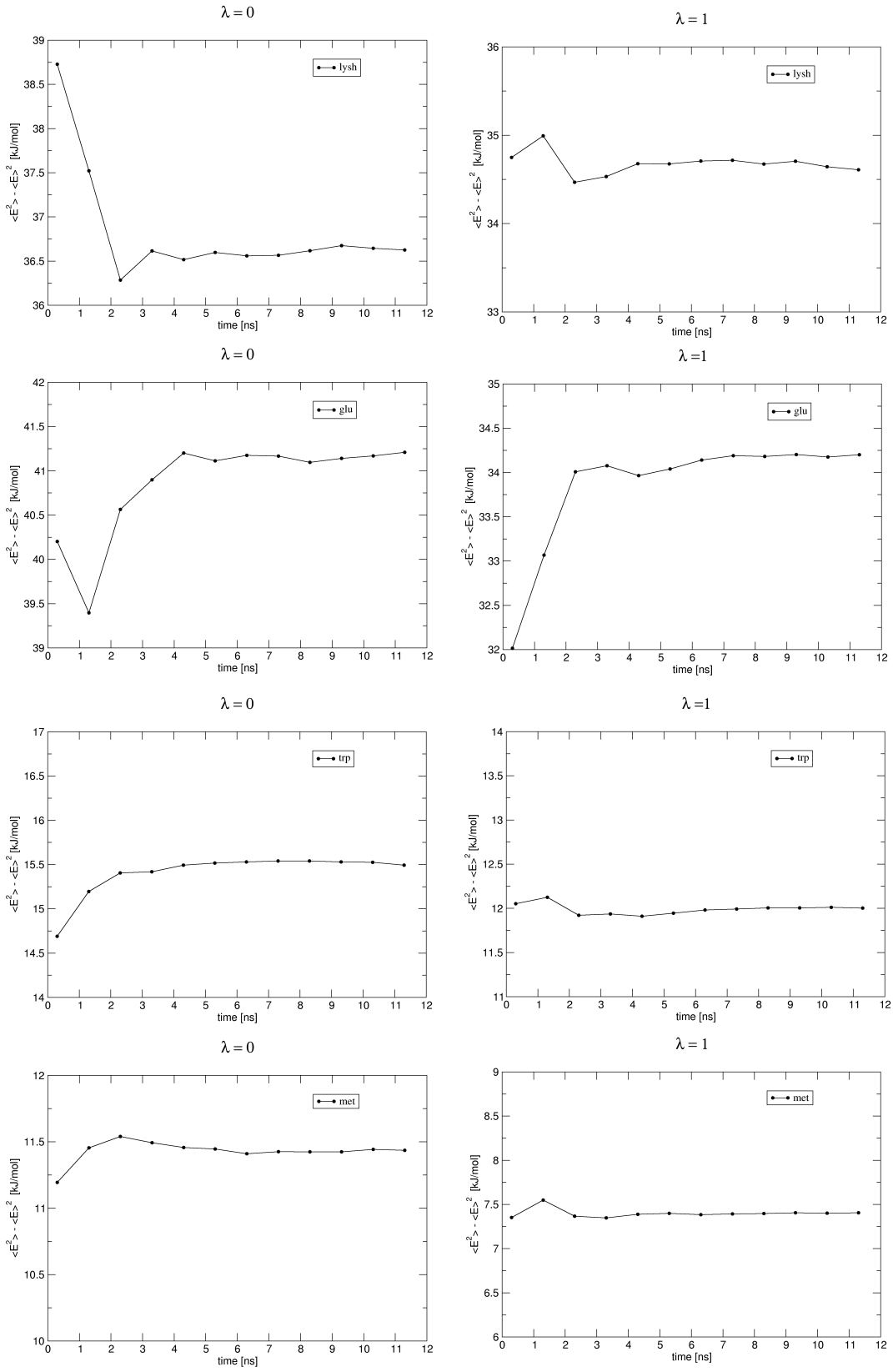

Figure S1: Cumulative estimates of the fluctuations ($\langle E^2 \rangle - \langle E \rangle^2$) at the end points of uncharging (charged state (λ=0) and neutral state (λ=1)) of selected sidechain analogues as a function of simulation time.

Data S1: Molecular topology building blocks for the reference states in Figure 2 of the main document.

```
MTBUILDBLSOLUTE
# building block (residue, nucleotide, etc.)
# RNME
R1
# number of atoms, number of preceding exclusions
# NMAT,NLIN
    2    0
# preceding exclusions
#ATOM                                MAE MSAE
# atoms
#ATOM ANM  IACM MASS        CGMICGM MAE MSAE
    1 AR     31   39    0.00000   0   1    2
# bonds
#  NB
    0
#  IB   JB  MCB
# bond angles
# NBA
    0
#  IB   JB   KB  MCB
# improper dihedrals
# NIDA
    0
#  IB   JB   KB   LB  MCB
# dihedrals
# NDA
    0
#  IB   JB   KB   LB  MCB
# LJ exceptions
# NEX
    0
#  IB   JB  MCB  NCO  IND  CON
END
MTBUILDBLSOLUTE
# building block (residue, nucleotide, etc.)
# RNME
R2
# number of atoms, number of preceding exclusions
# NMAT,NLIN
    2    0
# preceding exclusions
#ATOM                                MAE MSAE
# atoms
#ATOM ANM  IACM MASS        CGMICGM MAE MSAE
    1 AR     31   39    0.00000   0   1    2
    2 AR     31   39    0.00000   1   0
# bonds
#  NB
    1
#  IB   JB  MCB
    1    2   54
# bond angles
# NBA
    0
#  IB   JB   KB  MCB
# improper dihedrals
# NIDA
```

```
    0
#  IB   JB   KB   LB  MCB
# dihedrals
# NDA
    0
#  IB   JB   KB   LB  MCB
# LJ exceptions
# NEX
    0
#  IB   JB  MCB  NCO  IND  CON
END
MTBUILDBLSOLUTE
# building block (residue, nucleotide, etc.)
# RNME
R3
# number of atoms, number of preceding exclusions
# NMAT,NLIN
    6    2
# preceding exclusions
#ATOM                      MAE MSAE
   -1                                4    0    1    2    3
    0                                1    1

# atoms
#ATOM ANM  IACM MASS         CGMICGM MAE MSAE
    1 N       6   14   -0.31000   0   4    2    3    4    5
    2 H      21    1    0.31000   1   1    3
    3 CA     14    3    0.00000   1   4    4    5    6     7
    4 AR     31   39    0.00000   1   1    5
    5 C      12   12    0.45000   0
    6 O       1   16   -0.45000   1
# bonds
#  NB
    6
#  IB   JB  MCB
    1    2    2
    1    3   21
    3    4   53
    3    5   27
    5    6   5
    5  7   10
# bond angles
# NBA
    9
#  IB   JB   KB  MCB
   -1    1    2   32
   -1    1    3   31
    2    1    3   18
    1    3    4   13
    1    3    5   13
    4    3    5   13
    3    5    6   30
    3    5    7   19
    6    5    7   33
# improper dihedrals
# NIDA
    3
#  IB   JB   KB   LB  MCB
    1   -1    3    2    1
    3    1    4    5    2
```

```
    5    3    7    6    1
# dihedrals
# NDA
    5
#  IB   JB   KB   LB  MCB
   -2   -1    1    3   14
   -1    1    3    5   44
   -1    1    3    5   43
    1    3    5    7   45
    1    3    5    7   42
# LJ exceptions
# NEX
    0
#  IB   JB  MCB  NCO  IND  CON
END

MTBUILDBLSOLUTE
# building block (residue, nucleotide, etc.)
# RNME
R4
# number of atoms, number of preceding exclusions
# NMAT,NLIN
    7    2
# preceding exclusions
#ATOM                                MAE MSAE
# atoms
   -1                                  4    0    1    2    3
    0                                  1    1
#ATOM ANM  IACM MASS         CGMICGM MAE MSAE
    1 N       6   14   -0.31000   0   4    2    3    4    6
    2 H      21    1    0.31000   1   1    3
    3 CA     14    3    0.00000   1   5    4    5    6    7    8
    4 AR     31   39    0.00000   1   2    5    6
    5 AR     31   39    0.00000   1   0
    6 C      12   12    0.45000   0
    7 O       1   16   -0.45000   1
# bonds
#  NB
    7
#  IB   JB  MCB
    1    2    2
    1    3   21
    3    4   53
    3    6   27
    4    5   54
    6    7    5
    6  8   10
# bond angles
# NBA
    10
#  IB   JB   KB  MCB
   -1    1    2   32
   -1    1    3   31
    2    1    3   18
    1    3    4   13
    1    3    6   13
    4    3    6   13
    3    4    5   13
    3    6    7   30
    3    6    8   19
```

```
    7    6    8   33
# improper dihedrals
# NIDA
    3
#  IB   JB   KB   LB  MCB
    1   -1    3    2    1
    3    1    4    6    2
    6    3    8    7    1
# dihedrals
# NDA
    6
#  IB   JB   KB   LB  MCB
   -2   -1    1    3   14
   -1    1    3    6   44
   -1    1    3    6   43
    1    3    4    5   34
    1    3    7    8   45
    1    3    7    8   45
# LJ exceptions
# NEX
    0
#  IB   JB  MCB  NCO  IND  CON
END
MTBUILDBLSOLUTE
# building block (residue, nucleotide, etc.)
# RNME
R5
# number of atoms, number of preceding exclusions
# NMAT,NLIN
    7    2
# preceding exclusions
#ATOM                                MAE MSAE
   -1                                  4    0    1    2    3
    0                                  1    1
# atoms
#ATOM ANM  IACM MASS        CGMICGM MAE MSAE
    1 N       6   14   -0.31000   0   4    2    3    4    6
    2 H      21    1    0.31000   1   1    3
    3 CA     14    3    0.00000   1   5    4    5    6    7  8
    4 D      22    4    0.00000   0   2    5    6
    5 AR     31   39    0.00000   1   0
    6 C      12   12    0.45000   0
    7 O       1   16   -0.45000   1
# bonds
#  NB
    7
#  IB   JB  MCB
    1    2    2
    1    3   21
    3    4   27
    3    6   27
    4    5   27
    6    7    5
    6    8    10
# bond angles
# NBA
    10
#  IB   JB   KB  MCB
   -1    1    2   32
   -1    1    3   31
```

```
    2    1    3   18
    1    3    4   13
    1    3    6   13
    4    3    6   13
    3    4    5   15
    3    6    7   30
    3    6    8   19
    7    6    8   33
# improper dihedrals
# NIDA
    3
#  IB   JB   KB   LB  MCB
    1   -1    3    2    1
    3    1    6    4    2
    6    3    8    7    1
# dihedrals
# NDA
    6
#  IB   JB   KB   LB  MCB
   -2   -1    1    3   14
   -1    1    3    6   44
   -1    1    3    6   43
    1    3    4    5   34
    1    3    6    8   42
    1    3    6    8   45
# LJ exceptions
# NEX
    0
#  IB   JB  MCB  NCO  IND  CON
END
MTBUILDBLSOLUTE
# building block (residue, nucleotide, etc.)
# RNME
R6
# number of atoms, number of preceding exclusions
# NMAT,NLIN
   10    2
# preceding exclusions
#ATOM                                MAE MSAE
   -1                                  4    0    1    2    3
    0                                  1    1
# atoms
#ATOM ANM  IACM MASS        CGMICGM MAE MSAE
    1 N       6   14   -0.31000   0   4    2    3    4    9
    2 H      21    1    0.31000   1   1    3
    3 CA     14    3    0.00000   1   6    4    5    8    9   10  11
    4 D      22   14    0.00000   0   5    5    6    7    8    9
    5 AR     31   39    0.00000   0   3    6    7    8
    6 D0     22   16    0.00000   0   1    7
    7 D1     22   16    0.00000   0   0
    8 D2     22   16    0.00000   1   0
    9 C      12   12    0.45000   0
   10 O       1   16   -0.45000   1
# bonds
#  NB
   10
#  IB   JB  MCB
    1    2    2
    1    3   21
    3    4   27
```

```
     3     9    27
     4     5    27
     4     8    27
     5     6    27
     5     7    27
     9    10     5
     9    11    10
# bond angles
# NBA
    15
#  IB   JB   KB  MCB
    -1    1    2   32
    -1    1    3   31
     2    1    3   18
     1    3    4   13
     1    3    9   13
     4    3    9   13
     3    4    5   15
     3    4    8   15
     5    4    8   15
     4    5    6   15
     4    5    7   15
     6    5    7   15
     3    9   10   30
     3    9   11   19
    10    9   11   33
# improper dihedrals
# NIDA
     5
#  IB   JB   KB   LB  MCB
     1   -1    3    2    1
     3    1    9    4    2
     4    8    5    3    2
     5    4    6    7    2
     9    3   11   10    1
# dihedrals
# NDA
     7
#  IB   JB   KB   LB  MCB
    -2   -1    1    3   14
    -1    1    3    9   44
    -1    1    3    9   43
     1    3    4    5   34
     1    3    9   11   42
     1    3    9   11   45
     3    4    5    6   34
# LJ exceptions
# NEX
     0
#  IB   JB  MCB  NCO  IND  CON
END
```