

# SUPPLEMENTARY MATERIAL

Expanding molecular modeling and design tools to non-natural sidechains

David Gfeller, Olivier Michielin, Vincent Zoete

## 1 Supplementary Figures

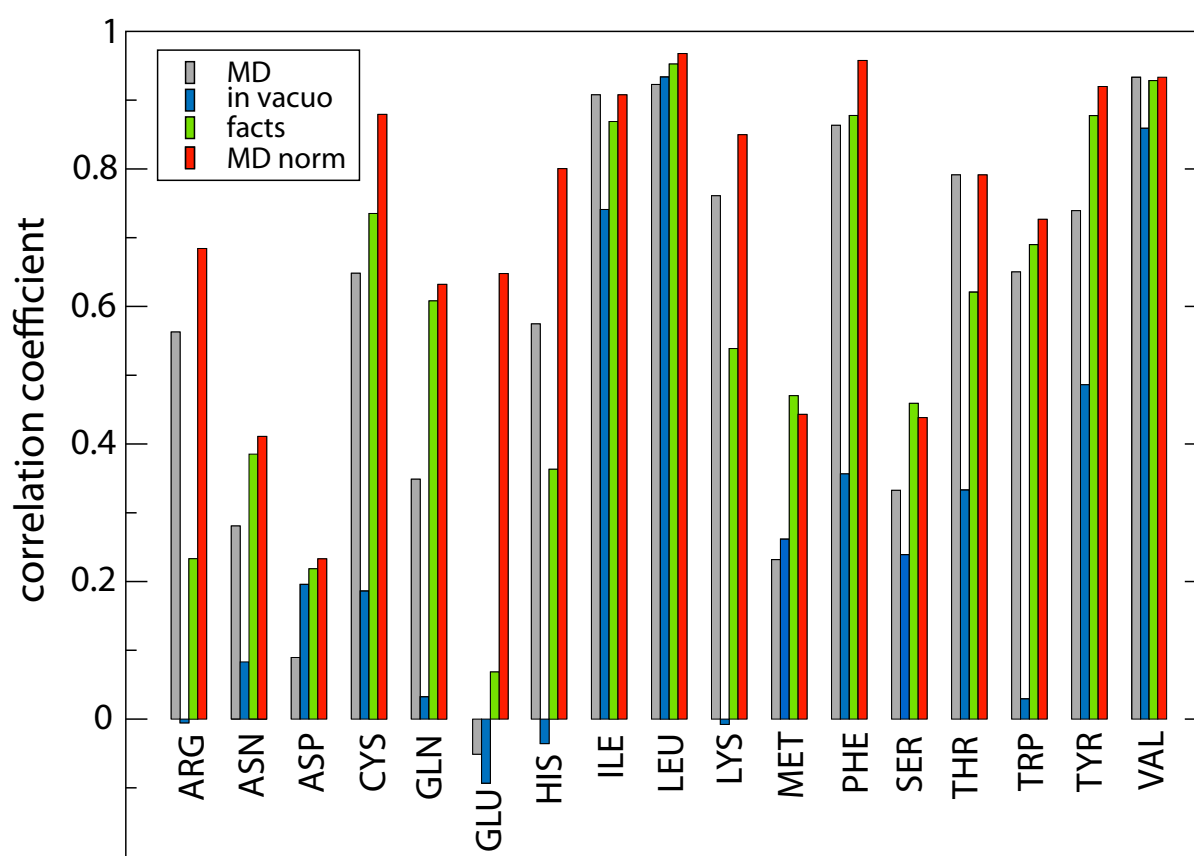


Figure S1: Comparison of rotamer probability predictions using probabilities inferred from frequencies along a MD trajectory (grey), computed from energy calculations in vacuo (blue) and with the FACTS solvent model (green). The blue bars show the cross-validation results after normalization (same data as in Figure 1).

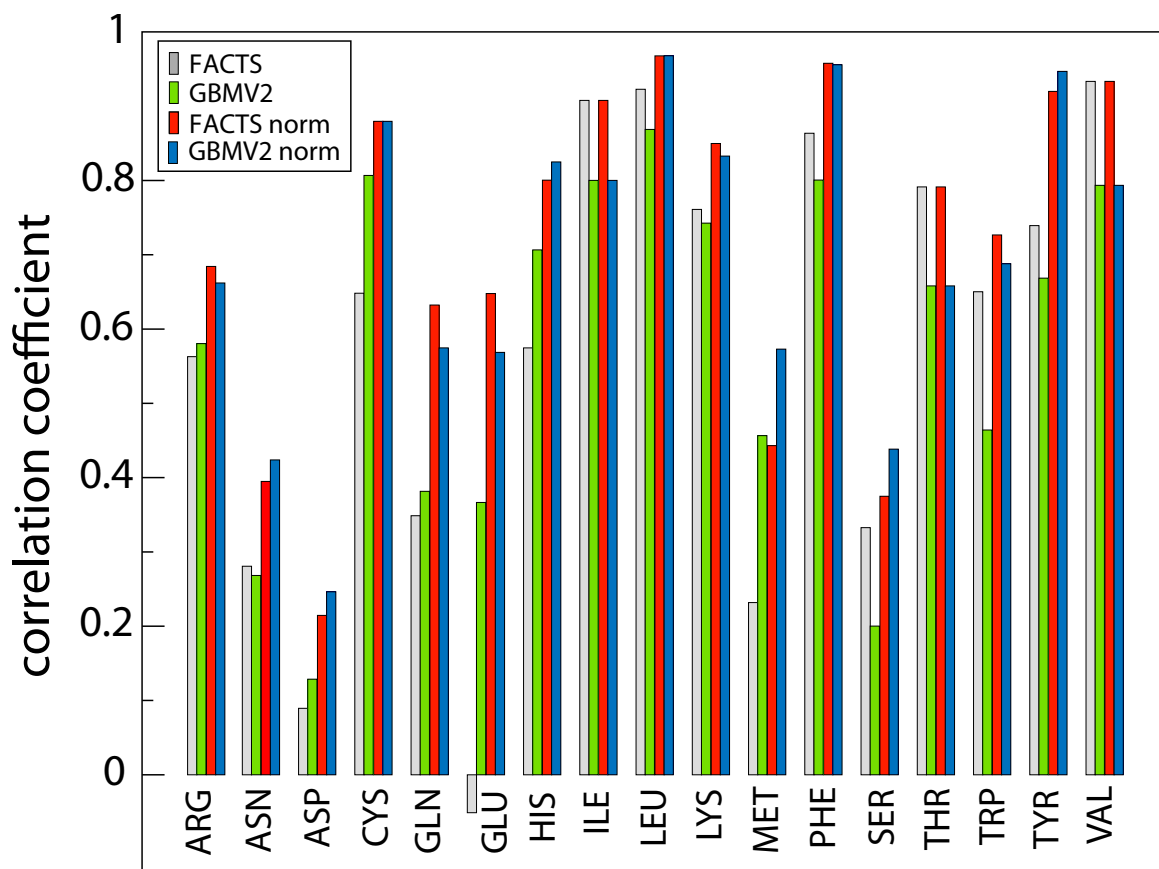


Figure S2: Comparison of rotamer probability predictions using different solvent models. Grey and red bars show the correlation coefficients for rotamers computed with the FACTS solvent both before and after renormalization (same data as in Figure 1). Green, respectively blue, bars show the correlation coefficients for rotamers computed with GBMV2 before, respectively after, renormalization.

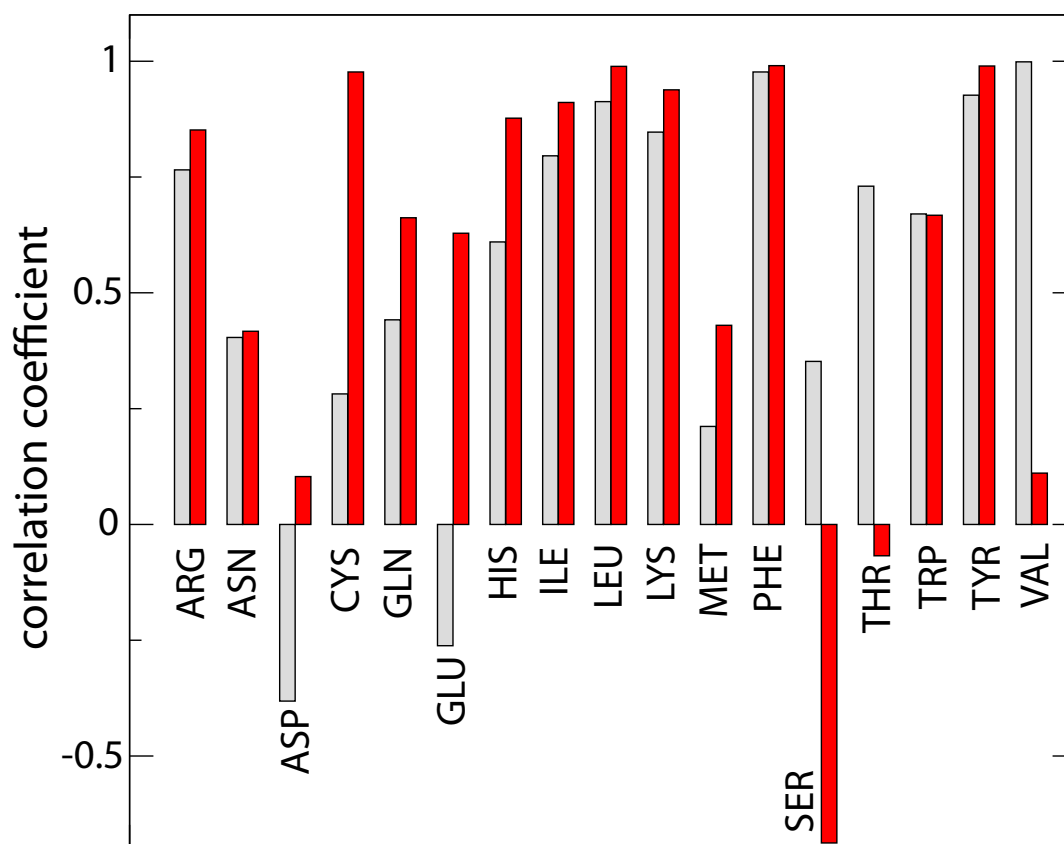


Figure S3: Cross-validation analysis for backbone independent rotamers. All sidechain satisfying the renormalization criteria display better correlations with experimental rotamer probabilities after renormalization. Grey bars show the correlation coefficients for probabilities inferred from MD trajectories. Red bars indicate the cross-validation results after renormalizing the rotamer probabilities by the average of experimental backbone independent first dihedral angle(s) rotamer probabilities.

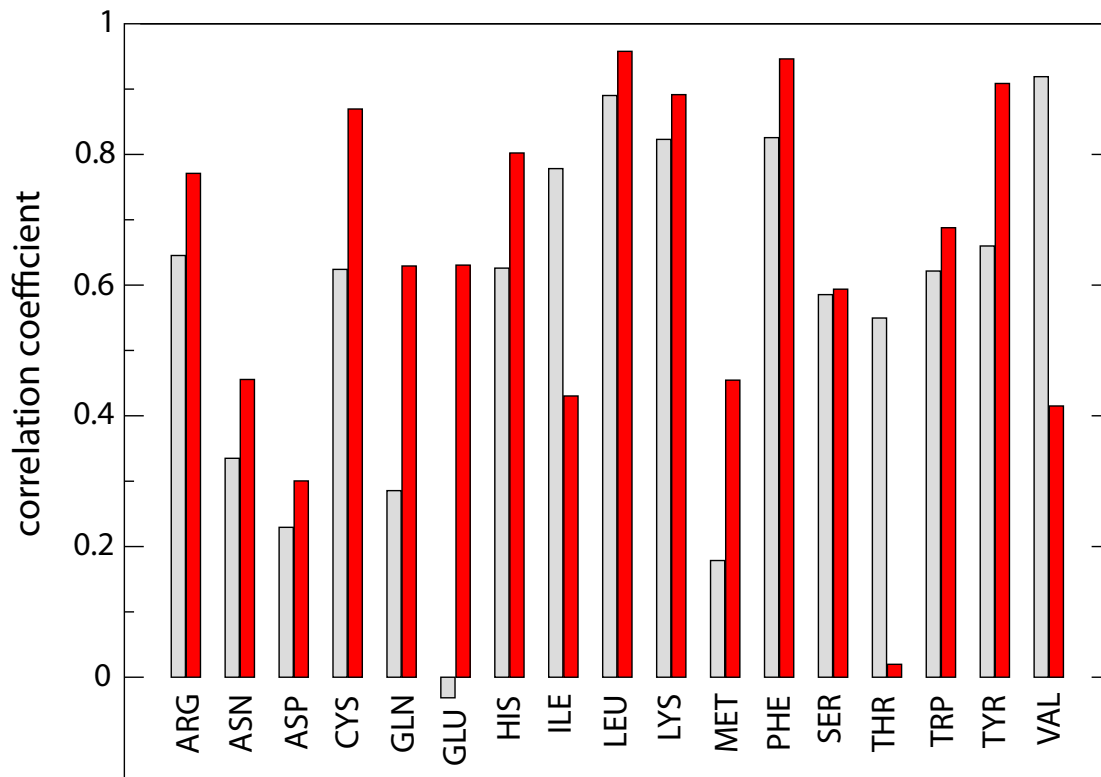


Figure S4: Cross-validation results for bins restricted to the alpha helix and beta sheet regions of the Ramachandran plot. Grey bars show the correlation coefficients for probabilities inferred from MD trajectories. Red bars indicate the cross-validation results after renormalizing the rotamer probabilities.

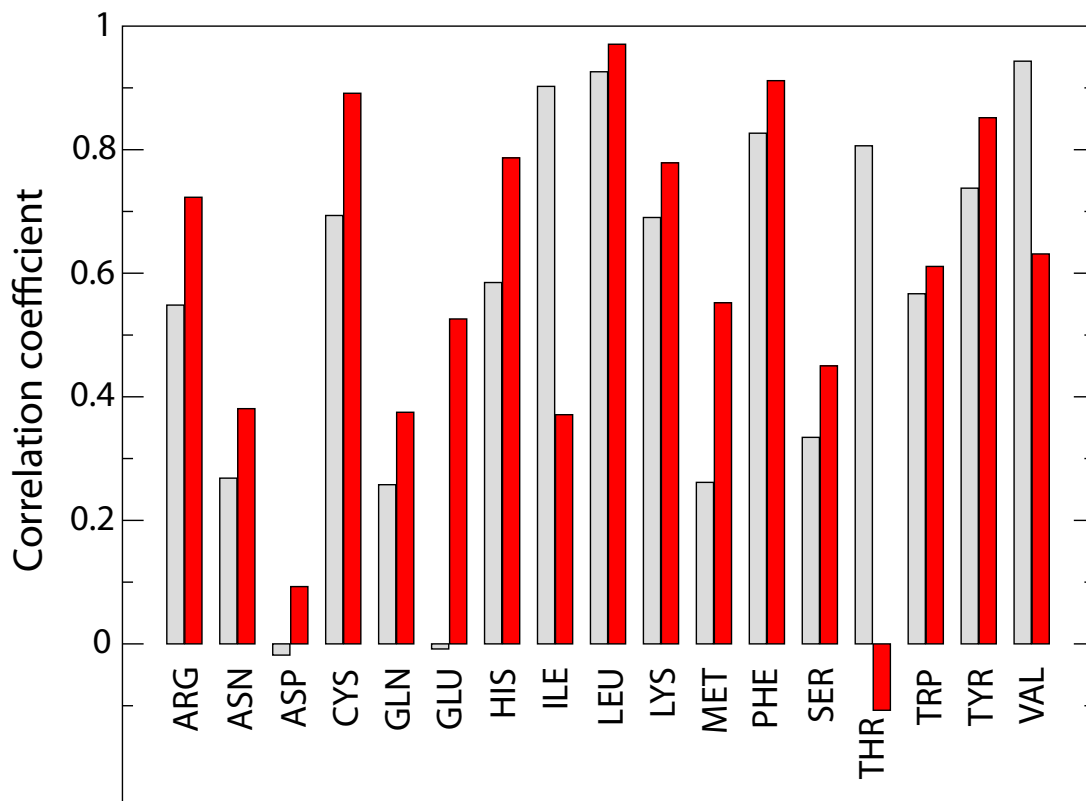
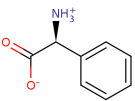
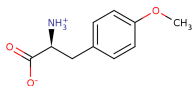
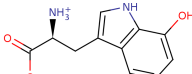
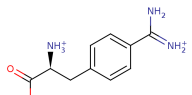
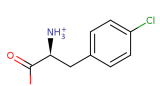
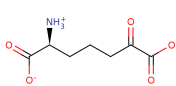
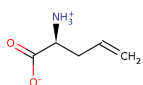


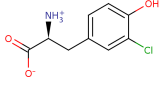
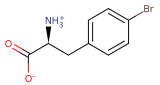
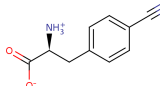
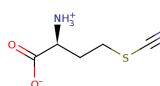
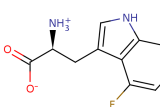
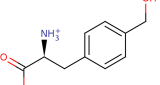
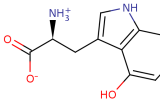
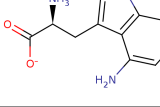
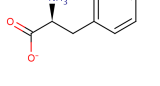
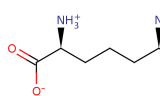
Figure S5: Comparison between the predicted rotamer probabilities and the Dunbrack 2010 rotamer library. Grey bars show the correlation coefficients for probabilities inferred from MD trajectories. Red bars indicate the cross-validation results after renormalizing the rotamer probabilities with the experimental backbone dependent first dihedral angles rotamer probabilities of the Dunbrack 2010 rotamer library. Predicted rotamer probabilities are computed with the same bin definition for sidechain dihedral angles as in Dunbrack 2010 rotamer library.

## 2 Supplementary Tables

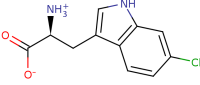
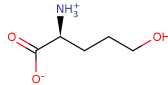
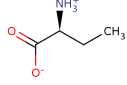
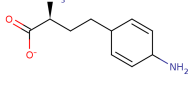
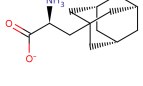
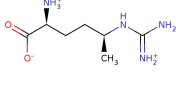
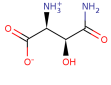
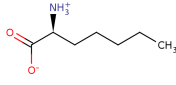
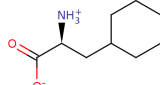
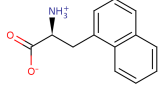
Table S1: List of all non-natural sidechains. The three- or four-letter code is shown in the first column. Column 2 displays the sidechain fullname. The 2D-structure is shown in column 3. The exact definition of freely rotating dihedral angles used to define rotamer is given in column 4 (see data on <http://www.swissnns.ch> for the atom names in the pdb files). Column 5 indicates the corresponding kind of dihedral angles (used to define the bins, see Table S2). The last column shows the type of renormalization that was applied on the rotamer probabilities (see Table S3).

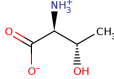
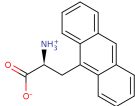
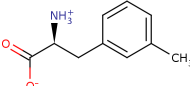
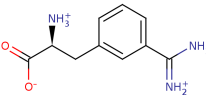
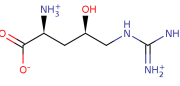
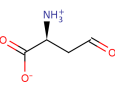
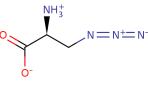
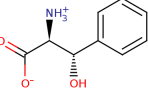
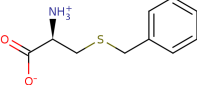
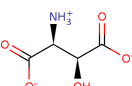
Code	Name	2D structure	Dihedral angles	Angle type	Renormalization scheme
004	Phenylglycine		N-CA-CB-CG1	13	
0A1	4-methoxy-beta-Phe		N-CA-CB-CG, CA-CB-CG-CD1, CE1-CZ-OH-CH	1, 3, 3	PHE
0AF	7-hydroxy-l-tryptophan		N-CA-CB-CG, CA-CB-CG-CD1	1, 4	TRP
0BN	4-carbamimidoyl-l-phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 3	PHE
200	4-chloro-Phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 3	PHE
26P	2-amino-6-oxopimelic acid		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-CE, CG-CD-CE-CZ	1, 1, 1, 9	Chi1,Chi2
2AG	2-Allyl-glycine		N-CA-CB-CG, CA-CB-CG-CD	1, 1	Chi1

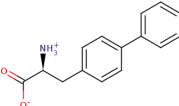
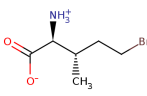
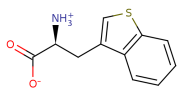
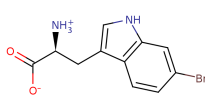
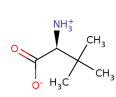
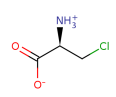
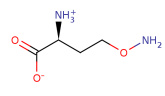
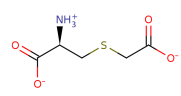
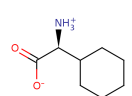
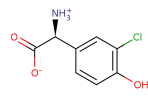


Code	Name	2D structure	Dihedral angles	Angle type	Renormalization scheme
3MY	3-Chloro-tyrosine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asym- metric
4BF	4-Bromo-phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 3	PHE
4CF	4-cyano-phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 3	PHE
4CY	nitrilo-l-methionine		N-CA-CB-CG, CA-CB-CG-SD, CB-CG-SD-CE	1, 1, 1	MET
4FW	4-fluoro-tryptophan		N-CA-CB-CG, CA-CB-CG-CD1	1, 4	TRP
4HMP	4-hydroxymethyl-phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1, CE1-CZ-CH-OJ	1, 3, 9	PHE
4HT	4-hydroxy-tryptophan		N-CA-CB-CG, CA-CB-CG-CD1	1, 4	TRP
4IN	4-amino-l-tryptophan		N-CA-CB-CG, CA-CB-CG-CD1	1, 4	TRP
4PH	4-methyl-phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 3	PHE
6CL	6-carboxyllysine		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-CE, CG-CD-CE-CZ, CD-CE-CZ-OH1	1, 1, 1, 1, 2	Chi1,Chi2

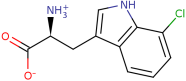
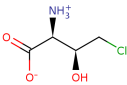
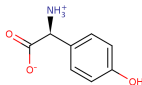
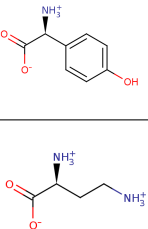
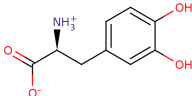
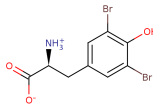
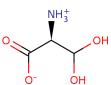
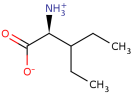
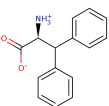
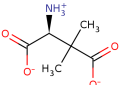


Code	Name	2D structure	Dihedral angles	Angle type	Renormalization scheme
6CW	6-chloro-l-tryptophan		N-CA-CB-CG, CA-CB-CG-CD1	1, 4	TRP
AA4	2-amino-5-hydroxypentanoic acid		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-OE	1, 1, 1	Chi1,Chi2
ABA	2-Aminobutyric acid		N-CA-CB-CG	1	Chi1
ACZ	cis-amclenomycin		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-CE1	1, 1, 1	Chi1,Chi2
ADAM	Adamanthane		N-CA-CB-CG, CA-CB-CG-CD1	1, 1	Chi1
AGM	5-methyl-arginine		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-NE1, CG-CD-NE1-CZ	1, 1, 1, 1	Chi1,Chi2
AHB	beta-hydroxyasparagine		N-CA-CB-CG, CA-CB-CG-OD1	1, 6	
AHP	2-Aminoheptanoic acid		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-CE, CG-CD-CE-CZ	1, 1, 1, 1	Chi1,Chi2
ALC	3-cyclohexyl-alanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 1	Chi1
ALN	1-Naphthalene		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asym- metric

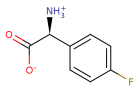
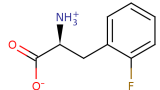
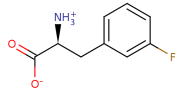
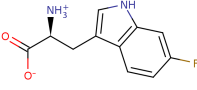
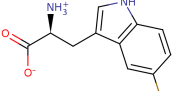
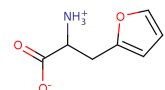

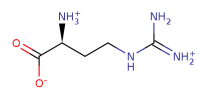
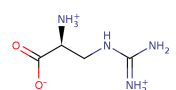
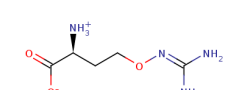
Code	Name	2D structure	Dihedral angles	Angle type	Renormalization scheme
ALO	Allo-threonine		N-CA-CB-OG1	1	
ANTH	3-(9-anthryl)-alanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 3	PHE
APD	3-Methyl-phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asymmetric
APM	m-amidinophenyl-3-alanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asymmetric
ARO	c-gamma-hydroxy arginine		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-NE, CG-CD-NE-CZ	1, 1, 1, 1	Chi1
AS2	(2r)-2-amino-4-oxobutanoic acid		N-CA-CB-CG, CA-CB-CG-OG	1, 8	Chi1
AZDA	azido-alanine		N-CA-CB-NG, CA-CB-NG-ND	1, 1	Chi1
BB8	Phenylserine		N-CA-CB-CG1, CA-CB-CG1-CD1	1, 3	
BCS	benzylcysteine		N-CA-CB-SG, CA-CB-SG-CD, CB-SG-CD-CE, SG-CD-CE-CZ1	1, 1, 1, 3	Chi1
BHD	beta-hydroxyaspartic acid		N-CA-CB-CG, CA-CB-CG-OD1	1, 2	

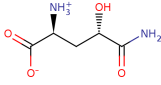
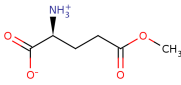
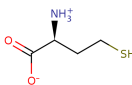
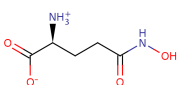
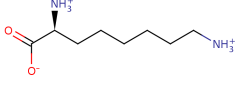
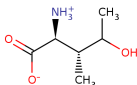
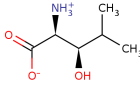
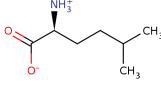
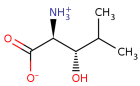
Code	Name	2D structure	Dihedral angles	Angle type	Renormalization scheme
BIF	biphenylalanine		N-CA-CB-CG, CA-CB-CG-CD1, CE1-CZ-C12-C13	1, 3, 10	PHE asym- metric
BIU	5-bromo-l-isoleucine		N-CA-CB-CG1, CA-CB-CG1-CD, CB-CG1-CD-BR	1, 1, 1	
BTH3	3-(3-benzothienyl)-alanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 4	TRP
BTR	6-bromo-tryptophan		N-CA-CB-CG, CA-CB-CG-CD1	1, 4	TRP
BUG	Tertleucine		N-CA-CB-CG1	15	
C2N	3-chloro-d-alanine		N-CA-CB-CL	1	Chi1
CAN	canaline		N-CA-CB-CG, CA-CB-CG-OD, CB-CG-OD-NE	1, 1, 1	Chi1
CCS	carboxymethylated cysteine		N-CA-CB-SG, CA-CB-SG-CD, CB-SG-CD-CE, SG-CD-CE-OZ1	1, 1, 1, 2	Chi1
CHG	Cyclohexylglycine		N-CA-CB-CG1	1	
CHP	3-chloro-4-hydroxy-phenylglycine		N-CA-CB-CG1	14	

Code	Name	2D structure	Dihedral angles	Angle type	Renormalization scheme
CIR	Citrulline		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-NE, CG-CD-NE-CZ	1, 1, 1, 1	Chi1,Chi2
CNP2	2-cyano-phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	Chi1
CP24	2,4-chloro-phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	Chi1
CP34	3,4-chloro-phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asymmetric
CPA3	3-Cyclopentyl-alanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 1	Chi1
CPG2	2-Chloro-phenylglycine		N-CA-CB-CG1	14	
CPG3	3-Chloro-phenylglycine		N-CA-CB-CG1	14	
CPG4	4-Chloro-phenylglycine		N-CA-CB-CG1	13	
CPH2	2-chloro-Phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	Chi1
CSA	s-acetyl cysteine		N-CA-CB-SG, CA-CB-SG-CD, CB-SG-CD-CE, SG-CD-CE-CZ	1, 1, 1, 1	Chi1

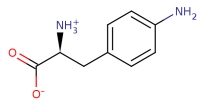
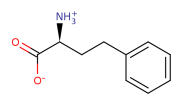
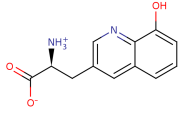
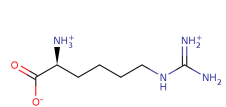
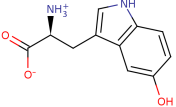
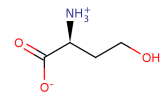
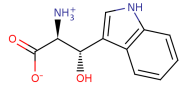
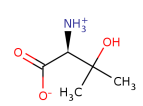
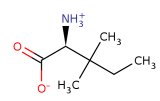
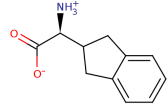
Code	Name	2D structure	Dihedral angles	Angle type	Renormalization scheme
CTE	7-chloro-tryptophan		N-CA-CB-CG, CA-CB-CG-CD1	1, 4	TRP
CTH	4-chloro-threonine		N-CA-CB-CG2, CA-CB-CG2-CL2	1, 1	
D4P	4-Hydroxy-phenylglycine		N-CA-CB-CG1	13	
DAB	Diaminobutyric acid		N-CA-CB-CG, CA-CB-CG-ND	1, 1	Chi1
DAH	3,4-Dihydroxy-phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asym- metric
DBY	3,5 dibromotyrosine		N-CA-CB-CG, CA-CB-CG-CD1	1, 3	PHE
DDZ	3,3-dihydroxy l-alanine		N-CA-CB-OG1	1	
DILE	Diethylalanine		N-CA-CB-CG1, CA-CB-CG1-CD1, CA-CB-CG2-CD2	1, 1, 1	
DIPH	di-phenylalanine		N-CA-CB-CG1, CA-CB-CG1-CD1, CA-CB-CG3-CD3	1, 3, 3	
DMK	3,3-dimethyl aspartic acid		N-CA-CB-CG1, CA-CB-CG1-OD1	1, 2	

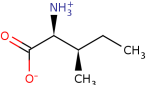
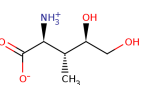
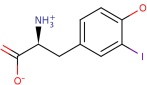
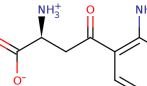
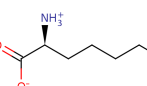
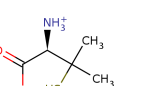
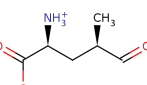
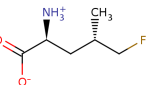
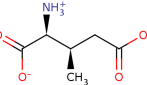
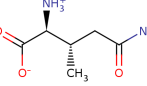
Code	Name	2D structure	Dihedral angles	Angle type	Renormalization scheme
DMP3	Dimethyl-phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1, CD2-CE2-CH-CJ	1, 12, 12	PHE asym- metric
DPP	2,3-Diaminopropanoic acid		N-CA-CB-NG	1	
ESC	2-amino-4-ethyl sulfanyl butyric acid		N-CA-CB-CG, CA-CB-CG-SD, CB-CG-SD-CE, CG-SD-CE-CZ	1, 1, 1, 1	MET
F2F	3,4-Difluoro-phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asym- metric
FCL	3-chloro-Phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asym- metric
FGA4	4-Fluoro-glutamic acid		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-OE1	1, 1, 2	Chi1
FGL	2-amino-propanedioic acid		N-CA-CB-OG1	2	
FLA	Trifluoro-alanine		N-CA-CB-FG1	15	
FPG2	2-Fluoro-phenylglycine		N-CA-CB-CG1	14	
FPG3	3-Fluoro-phenylglycine		N-CA-CB-CG1	14	

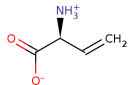
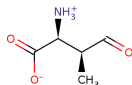
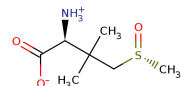
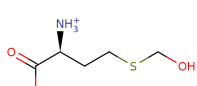
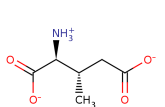
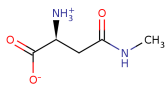
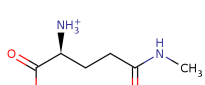
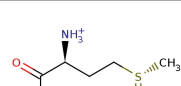
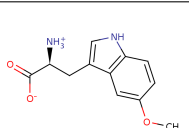
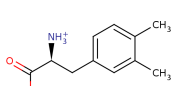
Code	Name	2D structure	Dihedral angles	Angle type	Renormalization scheme
FPG4	4-Fluoro-phenylglycine		N-CA-CB-CG1	13	
FPH2	2-Fluoro-Phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	Chi1
FPH3	3-Fluoro-Phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asym- metric
FT6	6-fluoro-1-tryptophan		N-CA-CB-CG, CA-CB-CG-CD1	1, 4	TRP
FTR	5-Fluoro-tryptophan		N-CA-CB-CG, CA-CB-CG-CD1	1, 4	TRP
FUA2	(2-furyl)-alanine		N-CA-CB-CG, CA-CB-CG-OD1	1, 5	Chi1
FVAL	3-Fluoro-valine		N-CA-CB-FG3	1	
GBUT	Guanidinobutyric		N-CA-CB-CG, CA-CB-CG-ND, CB-CG-ND-CE	1, 1, 1	Chi1
GDPR	2-Amino-3-guanidinopropionic acid		N-CA-CB-NG, CA-CB-NG-CD	1, 1	
GGB	Canavanine		N-CA-CB-CG, CA-CB-CG-OD, CB-CG-OD-NE, CG-OD-NE-CZ	1, 1, 1, 1	Chi1

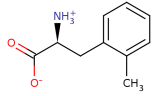
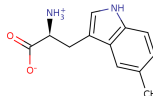
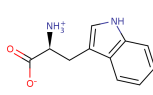
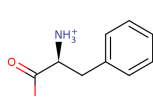
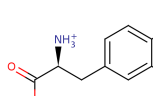
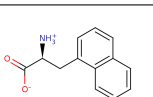
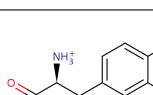
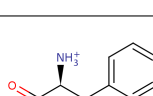
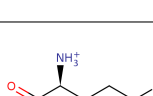
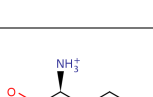
Code	Name	2D structure	Dihedral angles	Angle type	Renormalization scheme
GHG	(2s,4s)-2,5-diamino-4-hydroxy-5-oxopentanoic acid		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-OE1	1, 1, 7	Chi1
GME	5-o-methylglutamic acid		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-OE2, CG-CD-OE2-CZ	1, 1, 1, 1	Chi1,Chi2
HCS	homocysteine		N-CA-CB-CG, CA-CB-CG-SD	1, 1	Chi1,Chi2
HGA	glutamine hydroxamate		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-NE2, CG-CD-NE2-OZ	1, 1, 1, 11	Chi1,Chi2
HHK	(2s)-2,8-diaminooctanoic acid		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-CE, CG-CD-CE-CZ, CD-CE-CZ-CH, CE-CZ-CH-NJ	1, 1, 1, 1, 1, 1	Chi1,Chi2
HIL4	4-Hydroxy-L-isoleucine		N-CA-CB-CG1, CA-CB-CG1-CD2	1, 1	
HL2	(2s,3r)-2-amino-3-hydroxy-4-methylpentanoic acid		N-CA-CB-CG, CA-CB-CG-CD1	1, 1	
HLEU	Homoleucine		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-CE1	1, 1, 1	Chi1,Chi2
HLU	2s,3s)-2-amino-3-hydroxy-4-methylpentanoic acid		N-CA-CB-CG, CA-CB-CG-CD1	1, 1	

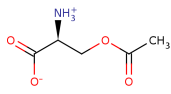
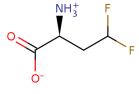
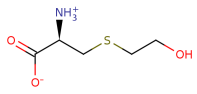
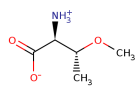
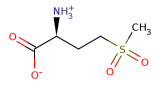
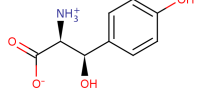
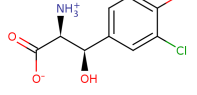
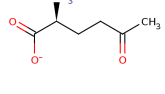
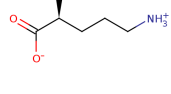
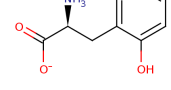


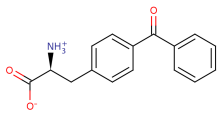
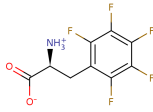
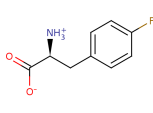
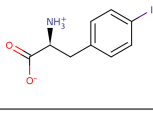
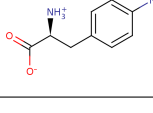
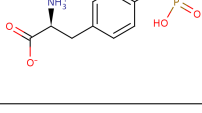
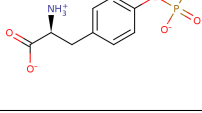
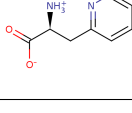
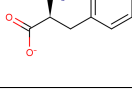
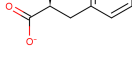
Code	Name	2D structure	Dihedral angles	Angle type	Renormalization scheme
HOX	4-amino-L-phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 3	PHE
HPE	Homophenylalanine		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-CE1	1, 1, 3	Chi1,Chi2
HQA	3-(8-hydroxyquinolin-3-yl)-l-alanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asymmetric
HRG	homoarginine		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-CE, CG-CD-CE-NZ, CD-CE-NZ-CH	1, 1, 1, 1, 1	Chi1,Chi2
HRP	5-Hydroxy-tryptophan		N-CA-CB-CG, CA-CB-CG-CD1	1, 4	TRP
HSER	homoserine		N-CA-CB-CG, CA-CB-CG-OD	1, 1	Chi1
HTR	beta-hydroxy-tryptophane		N-CA-CB-CG, CA-CB-CG-CD1	1, 4	
HVA	3-hydroxy-l-valine		N-CA-CB-OG3	1	
I2M	3-methyl-l-alloisoleucine		N-CA-CB-CG1, CA-CB-CG1-CD1	1, 1	
IGL	alpha-amino-2-indanacetic acid		N-CA-CB-CG1	1	

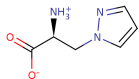
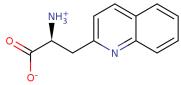
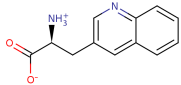
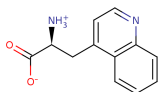
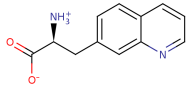
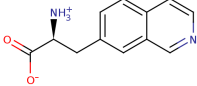
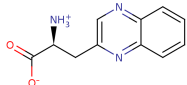
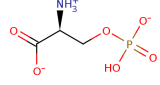
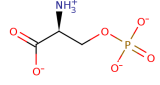
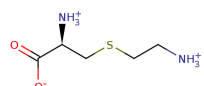
Code	Name	2D structure	Dihedral angles	Angle type	Renormalization scheme
IIL	Allo-Isoleucine		N-CA-CB-CG1, CA-CB-CG1-CD1	1, 1	
ILX	4,5-dihydroxy-isoleucine		N-CA-CB-CG1, CA-CB-CG1-CD1, CB-CG1-CD1- OD1	1, 1, 1	
IYR	3-iodo-tyrosine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asym- metric
KYN	kynurenine		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-CE1	1, 1, 16	Chi1
LDO	6-hydroxy-l-norleucine		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-CE, CG-CD-CE-OZ	1, 1, 1, 1	Chi1,Chi2
LE1	Penicillamine		N-CA-CB-SG3	1	
LED	(4r)-5-oxo-l-leucine		N-CA-CB-CG, CA-CB-CG-CD2, CB-CG-CD2-OE	1, 1, 8	Chi1
LEF	(4s)-5-fluoro-l-leucine		N-CA-CB-CG, CA-CB-CG-CD1, CB-CG-CD1-F1	1, 1, 1	Chi1
LME	(3r)-3-methyl-l-glutamic acid		N-CA-CB-CG1, CA-CB-CG1-CD, CB-CG1-CD-OE1	1, 1, 2	
LMQ	3-methyl-l-glutamine		N-CA-CB-CG1, CA-CB-CG1-CD, CB-CG1-CD-OE1	1, 1, 7	

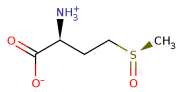
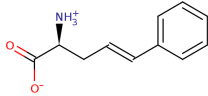
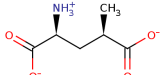
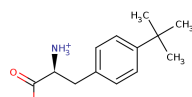
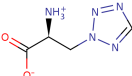
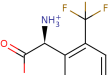
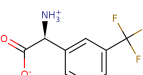
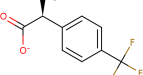
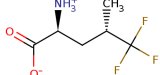
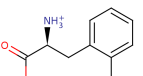
Code	Name	2D structure	Dihedral angles	Angle type	Renormalization scheme
LVG	vinylglycine		N-CA-CB-CG	1	
LVN	4-oxo-l-valine		N-CA-CB-CG1, CA-CB-CG1-OG1	1, 8	
M2S	3-{{(r)-methylsulfinyl}-methyl}-l-valine		N-CA-CB-CG, CA-CB-CG-SD, CB-CG-SD-CE	1, 1, 1	
ME0	hydroxy-l-methionine		N-CA-CB-CG, CA-CB-CG-SD, CB-CG-SD-CE, CG-SD-CE-OZ	1, 1, 1, 1	MET
MEG	(3s)-3-methyl-l-glutamic acid		N-CA-CB-CG1, CA-CB-CG1-CD, CB-CG1-CD-OE1	1, 1, 2	
MEN	n-methyl-asparagine		N-CA-CB-CG, CA-CB-CG-ND, CB-CG-ND-CE	1, 1, 11	Chi1
MEQ	n5-methyl-glutamine		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-NE2, CG-CD-NE2-CZ	1, 1, 1, 11	Chi1,Chi2
MHO	s-oxymethionine		N-CA-CB-CG, CA-CB-CG-SD, CB-CG-SD-CE	1, 1, 1	Chi1,Chi2
MOT5	5-Methoxy-tryptophan		N-CA-CB-CG, CA-CB-CG-CD1, CE3-CZ3-OZ3- CJ3	1, 4, 3	TRP
MP34	3,4-Dimethyl-phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asym- metric

Code	Name	2D structure	Dihedral angles	Angle type	Renormalization scheme
MPH2	2-Methyl-phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	Chi1
MTR5	5-Methyl-tryptophan		N-CA-CB-CG, CA-CB-CG-CD1	1, 4	TRP
MTR6	6-Methyl-tryptophan		N-CA-CB-CG, CA-CB-CG-CD1	1, 4	TRP
MTY	m-Tyrosine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asym- metric
NAL	2-Naphthalene		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asym- metric
NAO1	5-hydroxy-1-naphthalene		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	Chi1
NAO2	6-hydroxy-2-naphthalene		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asym- metric
NIY	meta-nitro-tyrosine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	Chi1
NLE	Nor-leucine		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-CE	1, 1, 1	Chi1,Chi2
NVA	Norvaline		N-CA-CB-CG, CA-CB-CG-CD	1, 1	Chi1,Chi2

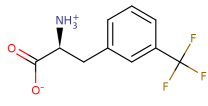
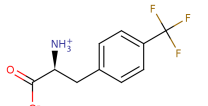
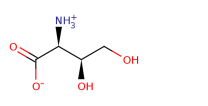
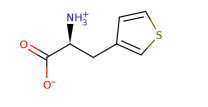
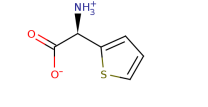
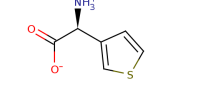
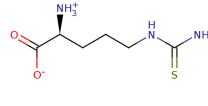
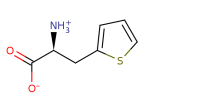
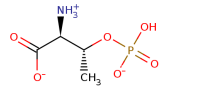
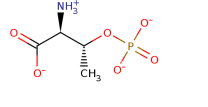
Code	Name	2D structure	Dihedral angles	Angle type	Renormalization scheme
OAS	o-acetylserine		N-CA-CB-OG, CA-CB-OG-CD, CB-OG-CD-CE	1, 1, 1	
OBF	(2s)-2-amino-4,4-difluorobutanoic acid		N-CA-CB-CG, CA-CB-CG-FG1	1, 1	Chi1
OCY	s-(2-hydroxyethyl)-l-cysteine		N-CA-CB-SG, CA-CB-SG-CD, CB-SG-CD-CE, SG-CD-CE-OZ	1, 1, 1, 1	Chi1
OLT	o-methyl-l-threonine		N-CA-CB-OG1, CA-CB-OG1-CD1	1, 1	
OMT	Methionine sulfone		N-CA-CB-CG, CA-CB-CG-SD, CB-CG-SD-CE	1, 1, 1	MET
OMX	(betar)-beta-hydroxy-l-tyrosine		N-CA-CB-CG, CA-CB-CG-CD1	1, 3	PHE asym- metric
OMY	(betar)-3-chloro-beta-hydroxy-l-tyrosine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	
ONL	5-oxo-l-norleucine		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-CE	1, 1, 1	Chi1,Chi2
ORN	Ornithine		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-NE	1, 1, 1	Chi1,Chi2
OTYR	o-Tyrosine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	Chi1

Code	Name	2D structure	Dihedral angles	Angle type	Renormalization scheme
PBF	4-benzoyl-Phe		N-CA-CB-CG, CA-CB-CG-CD1, CE1-CZ-CF-CJ, CZ-CF-CJ-CH1	1, 3, 12, 3	PHE
PF5	pentafluoro-phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 3	Chi1
PFF	4-Fluoro-Phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 3	PHE
PHI	4-Iodo-Phe		N-CA-CB-CG, CA-CB-CG-CD1	1, 3	PHE
PPN	4-Nitro-phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 3	PHE
PTR	phosphonotyrosine		N-CA-CB-CG, CA-CB-CG-CD1, CE1-CZ-OH-P, CZ-OH-P-O3P	1, 3, 9, 1	PHE
PTR2	phosphonotyrosine (charge: -2)		N-CA-CB-CG, CA-CB-CG-CD1, CE1-CZ-OH-P, CZ-OH-P-O1P	1, 3, 9, 1	PHE
PYR2	3-(2-Pyridyl)-alanine		N-CA-CB-CG, CA-CB-CG-ND1	1, 12	PHE asym- metric
PYR3	3-(3-Pyridyl)-alanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asym- metric
PYR4	3-(4-Pyridyl)-alanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 3	PHE

Code	Name	2D structure	Dihedral angles	Angle type	Renormalization scheme
PYZ1	3-(1-Pyrazolyl)-alanine		N-CA-CB-NG, CA-CB-NG-ND1	1, 5	HIS
QU32	3-(2-Quinolyl)-alanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asym- metric
QU33	3-(3-quinolyl)-alanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asym- metric
QU34	3-(4-quinolyl)-alanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	Chi1
QU35	3-(5-Quinolyl)-alanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asym- metric
QU36	3-(6-Quinolyl)-alanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asym- metric
QX32	3-(2-quinoxalyl)-alanine		N-CA-CB-CG, CA-CB-CG-ND1	1, 12	PHE asym- metric
SEP	phosphoserine		N-CA-CB-OG, CA-CB-OG-P, CB-OG-P-O3P	1, 1, 1	
SEP2	phosphoserine (charge: -2)		N-CA-CB-OG, CA-CB-OG-P, CB-OG-P-O1P	1, 1, 1	
SLZ	thialysine		N-CA-CB-SG, CA-CB-SG-CD, CB-SG-CD-CE, SG-CD-CE-NZ	1, 1, 1, 1	Chi1

Code	Name	2D structure	Dihedral angles	Angle type	Renormalization scheme
SME	Methionine sulfoxide		N-CA-CB-CG, CA-CB-CG-SD, CB-CG-SD-CE	1, 1, 1	MET
STYA	styryl-alanine		N-CA-CB-CG, CA-CB-CG-CD, CG-CD-CE-CZ1	1, 12, 3	Chi1
SYM	2s,4r-4-methylglutamate		N-CA-CB-CG, CA-CB-CG-CD1, CB-CG-CD1-OE1	1, 1, 2	Chi1
TBP4	4-tert-butyl-phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1, CE1-CZ-CH-CJ1	1, 3, 17	PHE
TEZA	3-(2-Tetrazolyl)-alanine		N-CA-CB-NG, CA-CB-NG-ND1	1, 5	HIS
TFG2	2-(Trifluoromethyl)-phenylglycine		N-CA-CB-CG1, CB-CG1-CH1-FH1	14, 17	
TFG3	3-(Trifluoromethyl)-phenylglycine		N-CA-CB-CG1, CG1-CD1-CH1-FH1	14, 17	
TFG4	4-(Trifluoromethyl)-phenylglycine		N-CA-CB-CG1, CD1-CE-CH-FH1	13, 17	
TFLE	5,5,5-Trifluoro-DL-leucine		N-CA-CB-CG, CA-CB-CG-CD1	1, 1	Chi1
TFP2	2-(Trifluoromethyl)-phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1, CG-CD1-CH1-FH1	1, 12, 17	Chi1



Code	Name	2D structure	Dihedral angles	Angle type	Renormalization scheme
TFP3	3-(Trifluoromethyl)-phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1, CD1-CE1-CH1- FH1	1, 12, 17	PHE asym- metric
TFP4	4-(Trifluoromethyl)-phenylalanine		N-CA-CB-CG, CA-CB-CG-CD1, CE1-CZ-CH1-FH1	1, 3, 17	PHE
TH6	4-hydroxy-l-threonine		N-CA-CB-CG, CA-CB-CG-OD	1, 1	
THA3	3-(3-thienyl)-alanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 5	HIS
THG2	2-thienylglycine		N-CA-CB-SG1	14	
THG3	3-thienylglycine		N-CA-CB-CG1	14	
THIC	Thio-citrulline		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-NE, CG-CD-NE-CZ	1, 1, 1, 1	Chi1,Chi2
TIH	3-(2-thienyl)-alanine		N-CA-CB-CG, CA-CB-CG-SD1	1, 3	HIS
TPO	phosphothreonine		N-CA-CB-OG1, CA-CB-OG1-P, CB-OG1-P-O3P	1, 1, 1	
TPO2	phosphothreonine (charge: -2)		N-CA-CB-OG1, CA-CB-OG1-P, CB-OG1-P-O1P	1, 1, 1	

Code	Name	2D structure	Dihedral angles	Angle type	Renormalization scheme
TRO	2-hydroxy-tryptophan		N-CA-CB-CG, CA-CB-CG-CD1	1, 4	Chi1
TRX	6-hydroxy-tryptophan		N-CA-CB-CG, CA-CB-CG-CD1	1, 4	TRP
TRZ4	3-(1,2,4-Triazol-1-yl)-alanine		N-CA-CB-NG, CA-CB-NG-ND1	1, 5	HIS
TTQ	6-amino-7-hydroxy-l-tryptophan		N-CA-CB-CG, CA-CB-CG-CD1	1, 4	TRP
TY2	3-Amino-L-tyrosine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asym- metric
TYI	3,5-diiodotyrosine		N-CA-CB-CG, CA-CB-CG-CD1	1, 3	PHE
TYQ	3-amino-6-hydroxy-tyrosine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	Chi1
TZA4	(4-thiazolyl)-alanine		N-CA-CB-CG, CA-CB-CG-ND2	1, 5	HIS
UN1	2-Aminoadipic acid		N-CA-CB-CG, CA-CB-CG-CD, CB-CG-CD-CE, CG-CD-CE-OZ1	1, 1, 1, 2	Chi1,Chi2
VAH	Hydroxynorvaline		N-CA-CB-CG2, CA-CB-CG2-CD	1, 1	

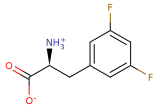
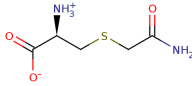
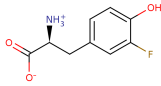
Code	Name	2D structure	Dihedral angles	Angle type	Renormalization scheme
WFP	3,5-Difluorophenylalanine		N-CA-CB-CG, CA-CB-CG-CD1	1, 3	PHE
YCM	cysteine-s-acetamide		N-CA-CB-SG, CA-CB-SG-CD, CB-SG-CD-CE, SG-CD-CE-OZ1	1, 1, 1, 7	Chi1
YOF	3-fluorotyrosine		N-CA-CB-CG, CA-CB-CG-CD1	1, 12	PHE asym- metric

Table S2: Bins used on different types of dihedral angles. Types 1-7 correspond approximately to the bins used in the 2002 Dunbrack rotamer library. Types 8-15 have been introduced to handle dihedral angles along bonds that do not have equivalent in natural sidechains.

Dihedral angle kind (Table S1)	Examples	Full interval	Bins
1	$\chi_1$ for all except PRO; $\chi_2$ for ARG, GLN, GLU, ILE, LEU, MET, LYS; $\chi_3$ for ARG, LYS, MET; $\chi_4$ for ARG, LYS	(-120, 240)	(0, 120), (120, 240), (-120, 0)
2	$\chi_2$ for ASP and $\chi_3$ for GLU	(-90, 90)	(30, 90), (-30, 30), (-90, -30)
3	$\chi_2$ for PHE and TYR	(-30, 150)	(30, 150), (-30, 30)
4	$\chi_2$ for TRP	(-180, 180)	(-180, -60), (-60, 60), (60, 180)
5	$\chi_2$ for HIS	(-120, 240)	(0, 120), (120, 240), (-120, 0)
6	$\chi_2$ for ASN	(-165, 195)	(-165, -90), (-90, -42), (-42, 12), (12, 60), (60, 110), (110, 195)
7	$\chi_3$ for GLN	(-210, 150)	(-210, -98), (-98, -18), (-18, 61), (61, 150)
8	AS2	(-180, 180)	(-180, -60), (-60, 60), (60, 180)
9	$\chi_4$ for 2NP	(-180, 180)	(-180, 0), (0, 180)
10	Biphenylalanine	(-90, 90)	(-90, 0), (0, 90)
11	N-methyl-asparagine	(-90, 270)	(-90, 90), (90, 270)
12	$\chi_2$ for ortho or meta substitutes of PHE (e.g., APD)	(-210, 150)	(-210, -150), (-150, -30), (-30, 30), (30, 150)
13	$\chi_1$ for Phenylglycine	(-30, 150)	(-30, 150)
14	$\chi_1$ for ortho or meta substituents of Phenylglycine (e.g., CHP)	(-210, 150)	(-210, -30), (-30, 150)
15	$\chi_1$ for Tertleucine	(0, 120)	(0, 120)
16	$\chi_3$ for kynurenine	(-180, 180)	(-180, -90), (-90, 0), (0, 90), (90, 180)
17	tert-butyl or trifluoromethyl substituents on aromatic rings	(0, 120)	(0, 60), (60, 120)

Table S3: Description of the different renormalization schemes used to predict rotamer probabilities.

<b>Type of renormalization (Table S1)</b>	<b>Description</b>	<b>Application criterion</b>
Chi1	Renormalize the rotamer probabilities of the first dihedral angle	Sidechain starting with linear chains of three carbon or sulfur atoms
Chi1,Chi2	Renormalize the rotamer probabilities of the first two dihedral angles	Sidechain starting with linear chains of four carbon or sulfur atoms
PHE	Renormalize the first two dihedral angles with the experimental rotamer probabilities of PHE	Derivatives of PHE with symmetric aromatic ring
PHE asymmetric	Renormalize the first two dihedral angles with the experimental rotamer probabilities of PHE shared equally between each 180 degree symmetrical bin	Derivatives of PHE with asymmetric aromatic ring (meta-substituents)
TRP	Renormalize the first two dihedral angles with the experimental rotamer probabilities of TRP	Derivatives of TRP
HIS	Renormalize the first two dihedral angles with the experimental rotamer probabilities of HIS	Derivatives of HIS
MET	Renormalize the first three dihedral angles with the experimental rotamer probabilities of MET	Derivatives of MET

Table S4: Comparison between experimental and predicted 3D conformations of non-natural sidechains. **Nd** indicates the number of dihedral angles, **Nr** indicates the total number of instances in all considered structures, **S** stands for the S score, as defined in Materials and Methods.  $N_i$  stands for the fraction of correctly predicted rotameric conformations up the  $i^{th}$  dihedral angle. The last column gives the list of structures used for each non-natural sidechain. The three phosphoryated residues (PTR, SEP, TPO) are listed at the end.

Name	Nd	Nr	S	$N_1$	$N_2$	$N_3$	$N_4$	List of structures
0A1	3	1	1.00	1.00	1.00	1.00		3QTC
0AF	2	12	1.00	1.00	1.00			1MAE, 1MAF, 3L4M, 3ORV, 3PXS, 3PXT, 3PXW
0BN	2	2	1.00	1.00	1.00			7KME, 8KME
200	2	1	1.00	1.00	1.00			2AKW
2AG	2	3	0.00	0.00	0.00			3COG
2AS	2	14	0.54	1.00	0.07			1KKR, 1W3M
2FM	4	1	1.00	1.00	1.00	1.00	1.00	1PFV
32T	2	1	1.00	1.00	1.00			1RMO
3CF	2	2	0.75	1.00	0.50			3OQZ
3GL	3	1	0.33	1.00	0.00	0.00		2ZJP
3MY	2	9	0.44	0.67	0.22			2WDX, 2XAD, 3MG9, 3MGB
4BF	2	22	0.98	1.00	0.95			2AG6, 3NI3, 3Q9G, 3Q9I
4CF	2	4	1.00	1.00	1.00			3OQY, 3QE4
4CY	3	2	0.67	1.00	0.50	0.50		3OR0
4FW	2	1	1.00	1.00	1.00			1RM9
4HT	2	4	0.00	0.00	0.00			1RU9, 1RUA, 1RUL, 1RUM
4IN	2	7	0.79	1.00	0.57			1OXF, 2HXX
4PH	2	1	1.00	1.00	1.00			3BV9
6CW	2	4	1.00	1.00	1.00			2AXI, 2GV2, 3FEA
AA4	3	1	1.00	1.00	1.00	1.00		1MIK
ABA	1	261	0.88	0.88				1B6J, 1B6K, 1B6L, 1B6M, 1B6P, 1C5F, 1CPI, 1CSA, 1CWA, 1CWB, 1CWC, 1CWH, 1CWL, 1CWM, 1CYN, 1D4K, 1D4L, 1IKF, 1M63, 1MF8, 1MTR, 1QNG, 1QNH, 1RBD, 1XQ7, 1YLC, 1YLD, 1Z1H, 1Z1R, 1Z3L, 1ZII, 1ZIJ, 2C5V, 2ESL, 2J9K, 2JUV, 2O40, 2OJU, 2POY, 2RMA, 2RMB, 2RMC, 2WFJ, 2X2C, 2X7K, 2Z6W, 2ZOK, 2ZOL, 2ZSV, 2ZSW, 3A0K, 3BO7, 3BXR, 3BXS, 3DCK, 3DCR, 3EOV, 3FSM, 3GIO, 3HAU, 3HAW, 3HBO, 3HDK, 3HLO, 3HVP, 3HZC, 3I2L, 3IA9, 3IAW, 3KA2, 3LO6, 3ODI, 3ODL, 4HVP, 7HVP, 8HVP, 9JDW
ACZ	3	2	0.67	1.00	0.50	0.50		1MLY
AGM	4	22	0.53	1.00	0.95	0.09	0.09	1E6Y, 1HBM, 1HBN, 1HBO, 1HBU, 1MRO, 3M1V, 3M2R, 3M2U, 3M2V, 3M30, 3M32, 3POT
AHB	2	5	0.50	0.80	0.20			1NT0, 2QC9, 2ZGD
AHP	4	3	0.58	0.67	0.67	0.67	0.33	1J4X, 3LO9

Name	Nd	Nr	S	$N_1$	$N_2$	$N_3$	$N_4$	List of structures
ALC	2	32	0.55	0.84	0.25			1B3H, 1D5M, 1D5X, 1D5Z, 1D6E, 1HBT, 1NZQ, 1O0D, 1QUR, 1THS, 1YWH, 2A2X, 2ANK, 2FEQ, 2FES, 3DPO, 3DPP, 3DPQ, 4THN, 5GDS
ALN	2	10	0.65	0.70	0.60			1B0H, 1FIV, 2FIV, 3FIV, 3OE0
ALO	1	1	0.00	0.00				2JUJ
ARO	4	4	0.81	1.00	0.75	0.75	0.75	1FFU, 1FFV
AS2	2	4	0.75	1.00	0.50			2GZ3
BB8	2	3	0.67	0.67	0.67			1D8T, 2C77
BCS	4	2	0.75	1.00	1.00	1.00	0.00	1EH8, 3L00
BHD	2	6	0.25	0.33	0.17			1AUT, 1ERM, 1PFX, 1XKA, 1XKB
BIF	3	4	0.33	1.00	0.00	0.00		1YYL, 2I5Y
C2N	1	2	1.00	1.00				1TZM
CAN	3	3	0.00	0.00	0.00	0.00		2CAN
CCS	4	17	0.31	0.59	0.35	0.18	0.12	1DSS, 1ERR, 1GTI, 1L2I, 1STF, 2BJ4, 2JFA, 3DMT, 3PQZ
CHG	1	4	0.25	0.25				4THN, 5GDS, 7KME, 8KME
CSA	4	4	0.56	1.00	0.50	0.50	0.25	1DWQ, 3NPF
CTE	2	3	0.83	1.00	0.67			2AR8, 2V7L, 2X68
CTH	2	2	0.50	1.00	0.00			1A7Z
DAB	2	13	0.50	0.62	0.38			1B4H, 1J73, 2PM1, 2W76, 3D3S, 3DS9, 3FQ9, 3KED
DAH	2	10	0.65	0.90	0.40			1IVV, 1RNR, 2VH3, 2ZWE, 2ZWF, 2ZWG, 6PAH
DBY	2	2	1.00	1.00	1.00			1EBA
DDZ	1	4	0.50	0.50				1E33, 1HDH, 2VQR
DMK	2	8	0.12	0.12	0.12			1NJT, 1NKK
DPP	1	7	0.00	0.00				1B5H, 3KNH, 3KNJ, 3KNL, 3KNN
ESC	4	1	0.50	1.00	1.00	0.00	0.00	1KBG
F2F	2	1	1.00	1.00	1.00			3D3V
FCL	2	2	0.50	1.00	0.00			1OKW
FGL	1	5	0.40	0.40				2VH3, 2W8S
FT6	2	1	1.00	1.00	1.00			2NW9
FTR	2	26	0.92	1.00	0.85			1I45, 1NEY, 1NF0, 2ZNX, 5FWG
GHG	3	3	0.56	0.67	0.67	0.33		1RU9, 1RUA, 1RUL
HCS	2	14	0.79	1.00	0.57			1Q8A, 1Q8J, 1U22, 1XDJ, 2CI5, 3BOF, 3BOL, 3BQ5
HGA	4	2	0.50	1.00	1.00	0.00	0.00	1XFG
HL2	2	2	0.50	1.00	0.00			3AH8
HLU	2	1	0.50	1.00	0.00			1ROV
HOX	2	1	0.50	1.00	0.00			3IDN
HPE	3	1	1.00	1.00	1.00	1.00		1B1H
HQA	2	2	1.00	1.00	1.00			3FCA
HRP	2	1	0.50	1.00	0.00			1YIA
IIL	2	25	0.84	0.88	0.80			1Q4V, 1RH4, 1TGG, 3I2L

Name	Nd	Nr	S	$N_1$	$N_2$	$N_3$	$N_4$	List of structures
ILX	3	3	0.33	1.00	0.00	0.00		1K83, 2VUM, 3CQZ
IYR	2	24	0.54	0.71	0.38			1CF0, 1WQ3, 2D8O, 2D8P, 2D8W, 2D97, 2D98, 2R1Q, 2Z10, 2Z11, 2ZP1, 2ZXV, 3GFD
KYN	3	11	0.30	0.64	0.27	0.00		1T5M, 1T5N, 2R2N, 2VOV, 2VOX, 3E2Z, 3PD6
LDO	4	4	0.00	0.00	0.00	0.00	0.00	1JSR
LEI	1	9	0.22	0.22				3H5F, 3M17, 3M1B
LED	3	1	0.33	1.00	0.00	0.00		2IUW
LEF	3	2	1.00	1.00	1.00	1.00		1OGW
LME	3	2	0.33	0.50	0.50	0.00		1T5M, 1T5N
LVG	1	2	0.50	0.50				1TDK
LVN	2	1	1.00	1.00	1.00			3KT7
ME0	4	12	0.38	1.00	0.50	0.00	0.00	3NBB, 3NBJ
MEN	3	64	0.60	0.78	0.52	0.50		1ALL, 1B33, 1B8D, 1D8T, 1EYX, 1GH0, 1HA7, 1JBO, 1KTP, 1PHN, 1QGW, 1XF6, 1XG0, 2BV8, 2C77, 2V8A, 2VJH, 2VJT, 3BRP, 3DBJ, 3O18, 3O2C
MEQ	4	4	0.19	0.50	0.25	0.00	0.00	1NV8, 3FMY, 3GN5
MHO	3	47	0.51	0.77	0.53	0.23		1EK0, 1F4G, 1NLR, 1SF5, 1ST2, 1UP0, 1UP2, 1XJO, 2A9E, 2IDT, 2IDU, 2IQF, 2IUF, 2QDV, 2QDW, 3D94, 3F00, 3M1X, 9PTI
MTY	2	5	0.50	0.80	0.20			1BIQ, 2TOH, 3HFV, 3HFZ
NAL	2	7	0.29	0.57	0.00			1SKL, 1ZH0, 2ITK, 2Q5A, 4THN, 5GDS
NIY	2	34	0.53	0.91	0.15			1K4Q, 1SDA, 2ADP, 2H5U, 2XAK, 2XAP, 2XAV, 2XAW, 2XAX, 2XAY, 2XAZ, 2XOF, 3DIV
NLE	3	104	0.55	0.73	0.65	0.28		1B7H, 1CFN, 1D5E, 1EOL, 1F3R, 1FGL, 1HD9, 1JB6, 1MMT, 1R1L, 1WY3, 1WY4, 1Z3L, 1Z3M, 2F4K, 2GYP, 2J9J, 2J9K, 2JE4, 2O40, 2RLN, 2RM9, 2RMD, 2RME, 2WUH, 2Z2T, 3DCK, 3DCR, 3FSM, 3G3P, 3GI0, 3HAU, 3HAW, 3HBO, 3HDK, 3HLO, 3HZC, 3I2L, 3IA9, 3IAW, 3JWB, 3KA2
NVA	2	13	0.85	0.92	0.77			1B6H, 1C9Y, 1JDX, 1OBC, 1OBH, 1Z3M, 1Z3P, 2BTE, 2I6U
OAS	3	16	0.31	0.94	0.00	0.00		1EBV, 2C58, 2VAX, 3FYU
OCY	4	24	0.39	0.54	0.54	0.25	0.21	1ASV, 1VSD, 1VSH, 1VSI, 1VSJ, 2R9S, 3FI2, 3FI3, 3FV8, 3KVX, 3O4Q
OMT	3	21	0.52	0.67	0.67	0.24		1EA0, 1H6N, 1H7K, 1M85, 1MQF, 1NM0, 2CAG, 2CAH, 2F1K, 2ISA
OMX	2	16	1.00	1.00	1.00			1HH3, 1HHA, 1HHC, 1HHF
OMY	2	51	0.73	0.86	0.59			1AA5, 1C0Q, 1C0R, 1FVM, 1GHG, 1GO6, 1HHU, 1HHY, 1HHZ, 1PN3, 1PNV, 1QD8, 1RRV, 1SHO, 2WDX, 2XAD, 3MG9, 3MGB



Name	Nd	Nr	S	$N_1$	$N_2$	$N_3$	$N_4$	List of structures
ONL	3	15	0.44	0.87	0.33	0.13		1ECC, 1ECG, 1OFE, 2J6H, 2Q3Z, 3BRM, 3DLA
ORN	3	117	0.48	0.86	0.43	0.15		1C30, 1HQG, 1KEE, 1M6V, 1T36, 1T5M, 1T5N, 1TK2, 1X7D, 2PM5, 2W6T, 2W6U, 3GMZ, 3IDJ, 3IT6, 3NI3, 3Q9G, 3Q9H, 3Q9I, 3Q9J, 3S61, 3THJ
PBF	4	4	0.38	1.00	0.50	0.00	0.00	1EEN, 2HGZ, 3AJI
PFF	2	13	0.85	1.00	0.69			1OL1, 1OL2, 2C5V, 2UUE, 2WHB, 2X1N, 3D39, 3F3C
PHI	2	11	0.68	0.82	0.55			1CZI, 1GA1, 1GA4, 1NLU, 1ORW, 1T6H, 1TF9
PPN	2	4	0.50	0.50	0.50			1YTJ, 2WHH, 3RCE
SLZ	4	6	0.29	0.50	0.50	0.17	0.00	1BM2, 2J9J, 2J9K, 3DIG
SME	3	32	0.49	0.69	0.50	0.28		1GKF, 1GM7, 1GM9, 1GWF, 1WMF, 2ATM, 2OOV, 2OQE, 2XN2, 3H6S, 3KO6, 3KSG, 3MMH, 3O1F
SYM	3	2	1.00	1.00	1.00	1.00		1SD3
TH6	2	5	0.40	0.80	0.00			2WW7
TIH	2	3	1.00	1.00	1.00			1EOJ, 1KW0, 1MMK
TRO	2	4	0.75	0.75	0.75			1G3P, 1KB0, 3NZJ
TRX	2	3	0.83	1.00	0.67			1K83, 2VUM, 3CQZ
TTQ	2	4	1.00	1.00	1.00			2HXC, 2I0S
TY2	2	8	0.44	0.88	0.00			2VH3, 2XO4, 2XO5, 3MBB
TYI	2	30	0.83	0.90	0.77			1CTP, 1L0S, 2AXE, 2C3V, 2D8O, 2D8P, 2D8W, 2D97, 2D98, 2NOO, 2VLW, 3GH8
TYQ	2	4	1.00	1.00	1.00			1D6U, 1D6Y
VAH	2	2	1.00	1.00	1.00			3AJR
WFP	2	21	0.95	1.00	0.90			3KTJ, 3KTK
YCM	4	91	0.30	0.57	0.35	0.21	0.09	1L0Q, 1N1X, 1N3Z, 1PX5, 1TJG, 1TJH, 1TJI, 1TQ9, 2A8P, 2A8Q, 2A8R, 2A8S, 2O40, 2R0W, 2R0Z, 3BCP, 3DCK, 3DCR, 3ESW, 3EYS, 3EYU, 3FKZ, 3FL0, 3FL1, 3FL3, 3FSM, 3GI0, 3HLO, 3HZC, 3I2L, 3IA9, 3IAW, 3JWD, 3KA2
YOF	2	69	0.82	0.97	0.67			1RRX, 1XDC, 1XIL, 3FYG

Name	Nd	Nr	S	$N_1$	$N_2$	$N_3$	$N_4$	List of structures
PTR	4	467	0.53	0.81	0.52	0.25		1A1B, 1A1C, 1A1E, 1A31, 1A81, 1AD5, 1AYA, 1AYB, 1AYC, 1BF5, 1BG1, 1BM2, 1BMB, 1CM8, 1D4W, 1EEN, 1EEO, 1F1W, 1FBV, 1FMK, 1FPR, 1FYR, 1G1F, 1G1G, 1G1H, 1GAG, 1GNG, 1H9O, 1I3Z, 1IR3, 1J4X, 1JYQ, 1JYR, 1K3A, 1K4S, 1K4T, 1KA6, 1KC2, 1KSW, 1LCJ, 1LCK, 1LKK, 1LKL, 1NH3, 1NZL, 1NZV, 1O9U, 1OO4, 1P13, 1PKG, 1PTT, 1PTU, 1PTV, 1PTY, 1QCF, 1QG1, 1QPC, 1QPD, 1QPE, 1QPJ, 1R1P, 1R1Q, 1R1S, 1RQQ, 1RR8, 1RRJ, 1SC7, 1SEU, 1SHA, 1SHB, 1SHD, 1SPS, 1T8I, 1TL8, 1TRN, 1TZE, 1U54, 1UUR, 1UUS, 1X27, 1XXP, 1YGR, 1YGU, 1YI6, 1YRK, 1YVH, 1YVJ, 1YVL, 1YWN, 1ZFP, 2AUH, 2B4S, 2B7A, 2C0I, 2C0O, 2C0T, 2CBL, 2CI9, 2CIA, 2CJM, 2CJZ, 2DQ7, 2DVJ, 2DXP, 2ERK, 2FCI, 2H5K, 2H7F, 2H8H, 2HCK, 2HDX, 2HMH, 2HWL, 2I6O, 2IUH, 2IUI, 2IVT, 2IVU, 2IVV, 2J0L, 2L1C, 2OFU, 2OH4, 2OQ1, 2OW3, 2PLD, 2PLE, 2PTK, 2PVF, 2Q8Y, 2QO7, 2QOB, 2QON, 2QOQ, 2QYQ, 2SRC, 2V7A, 2VIF, 2VX3, 2W1I, 2WO6, 2X2K, 2X2L, 2X2M, 2XA4, 2XKK, 2Z8C, 2Z8P, 2ZM1, 2ZM3, 2ZM4, 2ZOQ, 2ZYB, 3AC1, 3AC2, 3AC3, 3AC4, 3AC5, 3AC8, 3ACJ, 3ACK, 3AD4, 3AD5, 3AD6, 3ANQ, 3ANR, 3BU3, 3BU5, 3BU6, 3BUM, 3BUN, 3BUO, 3BUW, 3BUX, 3BYM, 3BYO, 3C7Q, 3CD3, 3CI5, 3CLY, 3D42, 3D44, 3DA9, 3DK6, 3DK7, 3DQW, 3E62, 3E63, 3E64, 3EB0, 3EYG, 3EYH, 3F5P, 3F7Z, 3F88, 3FUP, 3FXX, 3GB2, 3GQI, 3IO7, 3IOK, 3JY9, 3K2L, 3KCK, 3KMM, 3KRR, 3KUL, 3KVW, 3KXZ, 3L4J, 3L4K, 3LCK, 3LPB, 3LXN, 3LXP, 3MAZ, 3ML4, 3MXC, 3MXY, 3N7Y, 3N84, 3N8M, 3NNX, 3OB1, 3OB2, 3OLL, 3OLR, 3OMH, 3OP0, 3PFV, 3PLF, 3PX7, 3PY3, 3Q32, 3Q6W, 3SRV, 3TL0, 3ZRK, 3ZRL, 3ZRM

Name	Nd	Nr	S	$N_1$	$N_2$	$N_3$	List of structures
SEP	3	781	0.49	0.58	0.40		1APM, 1ATP, 1BKX, 1BX6, 1C8L, 1CMK, 1F34, 1F8A, 1FA9, 1FMO, 1FU0, 1GPA, 1GZ2, 1H1W, 1H4X, 1HJK, 1I7W, 1J3H, 1JBP, 1JDY, 1JLU, 1K35, 1KHX, 1KKM, 1L3R, 1L7P, 1LF8, 1LWN, 1LWO, 1MKI, 1OHE, 1OKY, 1OKZ, 1P16, 1P22, 1P5D, 1P5G, 1PCJ, 1PCM, 1PJQ, 1PJS, 1PUA, 1PY1, 1Q24, 1Q61, 1Q62, 1Q8T, 1Q8U, 1Q8W, 1QJA, 1QJB, 1R0Z, 1RDQ, 1RE8, 1REJ, 1REK, 1RZR, 1SMH, 1STC, 1SVE, 1SVG, 1SVH, 1SYK, 1SZA, 1SZM, 1T15, 1T29, 1T2V, 1TH1, 1U5Q, 1U5R, 1U7F, 1U7V, 1U9I, 1UHG, 1UJK, 1UU3, 1UU7, 1UU8, 1UU9, 1UVR, 1V18, 1VEB, 1VKL, 1VRV, 1XH4, 1XH5, 1XH6, 1XH7, 1XH8, 1XH9, 1XHA, 1XJD, 1Y98, 1YDR, 1YDS, 1YDT, 1YHS, 1YI3, 1YI4, 1YWT, 1Z5M, 1Z8D, 1ZEB, 1ZEF, 1ZVV, 2AK7, 2AZM, 2B05, 2BIK, 2BR9, 2BTP, 2BVA, 2BZI, 2C1A, 2C1B, 2C1J, 2C1N, 2C30, 2C63, 2C74, 2CDZ, 2CEF, 2CEZ, 2CFJ, 2CPK, 2DU3, 2DU5, 2DU6, 2ERZ, 2F57, 2FEP, 2FKF, 2FO0, 2FWN, 2G57, 2GBL, 2GCD, 2GFC, 2GHQ, 2GHT, 2GLQ, 2GNF, 2GNG, 2GNH, 2GNI, 2GNJ, 2GNL, 2GPA, 2GU8, 2H4L, 2H5A, 2I0E, 2J0I, 2J90, 2JDS, 2JDT, 2JDV, 2JED, 2JFL, 2JG8, 2KMD, 2L5J, 2LAJ, 2LAY, 2LAZ, 2LB0, 2NPM, 2NRU, 2NZU, 2NZV, 2OBJ, 2OEN, 2OIB, 2OIC, 2OID, 2OVQ, 2PE0, 2PE1, 2PE2, 2PIL, 2PSG, 2PT3, 2PUM, 2Q0N, 2QCS, 2QG5, 2QKW, 2QLL, 2QPK, 2QQT, 2QRB, 2QUR, 2QVS, 2R5L, 2UVX, 2UVY, 2UVZ, 2UW0, 2UW3, 2UW4, 2UW5, 2UW6, 2UW7, 2UW8, 2V7O, 2VAG, 2VNW, 2VNY, 2VO0, 2VO3, 2VO6, 2VO7, 2VXC, 2W5V, 2W5W, 2W5X, 2WH0, 2X4Z, 2XCH, 2XCK, 2XIX, 2XIY, 2XIZ, 2XJ1, 2XJ2, 2XZ0, 2XZ1, 2Y1K, 2Z5Z, 3A77, 3AG9, 3AGL, 3AGM, 3AJ4, 3AMA, 3AMB, 3AMV, 3AXY, 3BEG, 3BGM, 3BGP, 3BGQ, 3BGZ, 3BH8, 3BH9, 3BHB, 3BKQ, 3BWJ, 3BXI, 3C04, 3COJ, 3CU8, 3CY3, 3D9K, 3D9L, 3D9M, 3D9N, 3D9P, 3DCV, 3DD1, 3DDS, 3DDW, 3DND, 3DNE, 3DVL, 3E8C, 3E8E, 3EFZ, 3EQU, 3EQV, 3ERH, 3ERI, 3EXH, 3F2A, 3F3Z, 3FAQ, 3FBV, 3FHI, 3FJQ, 3FNL, 3FQR, 3FQU, 3FQX, 3G2T, 3G2V, 3GA7, 3GC1, 3GCJ, 3GCK, 3GCL, 3GP3, 3H9F, 3H9O, 3HGK, 3HL2, 3HRC, 3HRF, 3I3W, 3I6N, 3IAF, 3IDB, 3IDC, 3IFQ, 3IOP, 3IQJ, 3IQU, 3IQV, 3IW4, 3JRW, 3JZM, 3K05, 3K0E, 3K0F, 3K0H, 3K0K, 3K15, 3K16, 3K2L, 3KKV, 3KRQ, 3KVW, 3L41, 3L6F, 3L9L, 3L9M, 3L9N, 3LJ0, 3LJ1, 3LJ2, 3MA3, 3MHR, 3MK0, 3MK1, 3MK2, 3MVJ, 3N8F, 3NAK, 3NAY, 3NIU, 3NKX, 3NUN, 3NUS, 3NUU, 3NUY, 3NX8, 3NYH, 3O2Q, 3O7L, 3O8I, 3OB2, 3OGW, 3OOG, 3OQM, 3OQN, 3ORX, 3ORZ, 3OTU, 3OVV, 3OW3, 3OWP, 3OXT, 3P0M, 3P35, 3P9Y, 3PFQ, 3PMG, 3PVB, 3PXE, 3PY4, 3Q1I, 3Q4A, 3Q9K, 3QC4, 3QCQ, 3QCS, 3QCX, 3QD3, 3QF1, 3QIC, 3QL6, 3R4X, 3R55, 3R5O, 3R5Q, 3RCJ, 3RKE, 3RTX, 3S1A, 3S4F, 3SDJ, 3SDM, 3SXV, 3TGY, 4ICD, 4PEP

Name	Nd	Nr	S	$N_1$	$N_2$	$N_3$	List of structures
TPO	3	589	0.69	0.81	0.57		1APM, 1ATP, 1BKX, 1BX6, 1CDK, 1CM8, 1CMK, 1CTP, 1E9H, 1FMO, 1FOT, 1FQ1, 1G6G, 1GXC, 1GY3, 1H1P, 1H1Q, 1H1R, 1H1S, 1H24, 1H25, 1H26, 1H27, 1H28, 1IB1, 1J3H, 1J4X, 1JBP, 1JLU, 1JST, 1JSU, 1L3R, 1LC7, 1NEX, 1O6K, 1O6L, 1O9D, 1O9F, 1OGU, 1OI9, 1OIU, 1OIY, 1OL5, 1OL7, 1P5E, 1PKD, 1Q24, 1Q4K, 1Q61, 1Q62, 1Q8T, 1Q8U, 1Q8W, 1QMZ, 1RDQ, 1RE8, 1REJ, 1REK, 1SMH, 1STC, 1SVE, 1SVG, 1SVH, 1SYK, 1SZM, 1TH1, 1U9I, 1UA2, 1UMW, 1V54, 1V55, 1VEB, 1W98, 1XH4, 1XH5, 1XH6, 1XH7, 1XH8, 1XH9, 1XHA, 1XJD, 1YDR, 1YDS, 1YDT, 1YJM, 1YRP, 1ZRZ, 2A19, 2A1A, 2AST, 2B2T, 2BFX, 2BFY, 2C1A, 2C1B, 2C6T, 2CCH, 2CCI, 2CJM, 2CPK, 2DYR, 2DYS, 2EIJ, 2EIK, 2EIL, 2EIM, 2EIN, 2ERK, 2ERZ, 2FF4, 2G9X, 2GA3, 2GBL, 2GFC, 2GNF, 2GNG, 2GNH, 2GNI, 2GNJ, 2GNL, 2GU8, 2IOE, 2IW6, 2IW8, 2IW9, 2J90, 2JDO, 2JDR, 2JDS, 2JDT, 2JDV, 2JFL, 2JFM, 2JGZ, 2JOC, 2KFU, 2KMD, 2LAX, 2LB2, 2LB3, 2NRU, 2NRY, 2O8Y, 2OIB, 2OIC, 2OID, 2OVQ, 2OVR, 2PIE, 2Q5A, 2Q8Y, 2QCS, 2QKW, 2QUR, 2QVS, 2RLT, 2UVX, 2UVY, 2UVZ, 2UW0, 2UW3, 2UW4, 2UW5, 2UW6, 2UW7, 2UW8, 2UW9, 2UZB, 2UZD, 2UZE, 2UZL, 2V7D, 2VAG, 2VGO, 2VGP, 2VNW, 2VNY, 2VO0, 2VO3, 2VO6, 2VO7, 2VRX, 2W08, 2W1C, 2W3O, 2W8D, 2WMA, 2WMB, 2WTV, 2X39, 2XH5, 2XIK, 2Y69, 2Y94, 2Z8P, 2ZXW, 3A62, 3A7F, 3A7G, 3A7H, 3A7I, 3A7J, 3A8W, 3A8X, 3ABK, 3ABL, 3ABM, 3AG1, 3AG2, 3AG3, 3AG4, 3AG9, 3AGL, 3AGM, 3AL3, 3ALO, 3AMA, 3AMB, 3ASN, 3ASO, 3BHT, 3BHU, 3BHV, 3BLH, 3BLQ, 3BLR, 3BWJ, 3BZI, 3C4W, 3C5L, 3CKX, 3COM, 3CQU, 3CQW, 3D0E, 3D5W, 3DDP, 3DDQ, 3DND, 3DNE, 3DOG, 3DPC, 3DVL, 3E5A, 3E6Y, 3E87, 3E88, 3E8C, 3E8D, 3E8E, 3F69, 3FBV, 3FHI, 3FJQ, 3FVH, 3FXZ, 3FY0, 3H9F, 3HA6, 3HGK, 3HIK, 3HUF, 3IDB, 3IDC, 3IW4, 3K09, 3K0A, 3K0C, 3K0E, 3K2L, 3KK8, 3KK9, 3KKV, 3L9L, 3L9M, 3L9N, 3LJ0, 3LJ1, 3LJ2, 3LQ5, 3LW1, 3M8W, 3M8Y, 3M8Z, 3MI9, 3MIA, 3MV5, 3MVH, 3MVJ, 3MY1, 3MY5, 3NX8, 3O7L, 3OB1, 3OCB, 3OJY, 3OOG, 3OT9, 3OVV, 3OW3, 3OW4, 3OWP, 3OXT, 3P0M, 3P2Z, 3P34, 3P36, 3P37, 3PFQ, 3POA, 3PVB, 3PY3, 3QD2, 3QHR, 3QHW, 3QKK, 3QKL, 3QKM, 3RQ7, 3SDJ, 3SDM