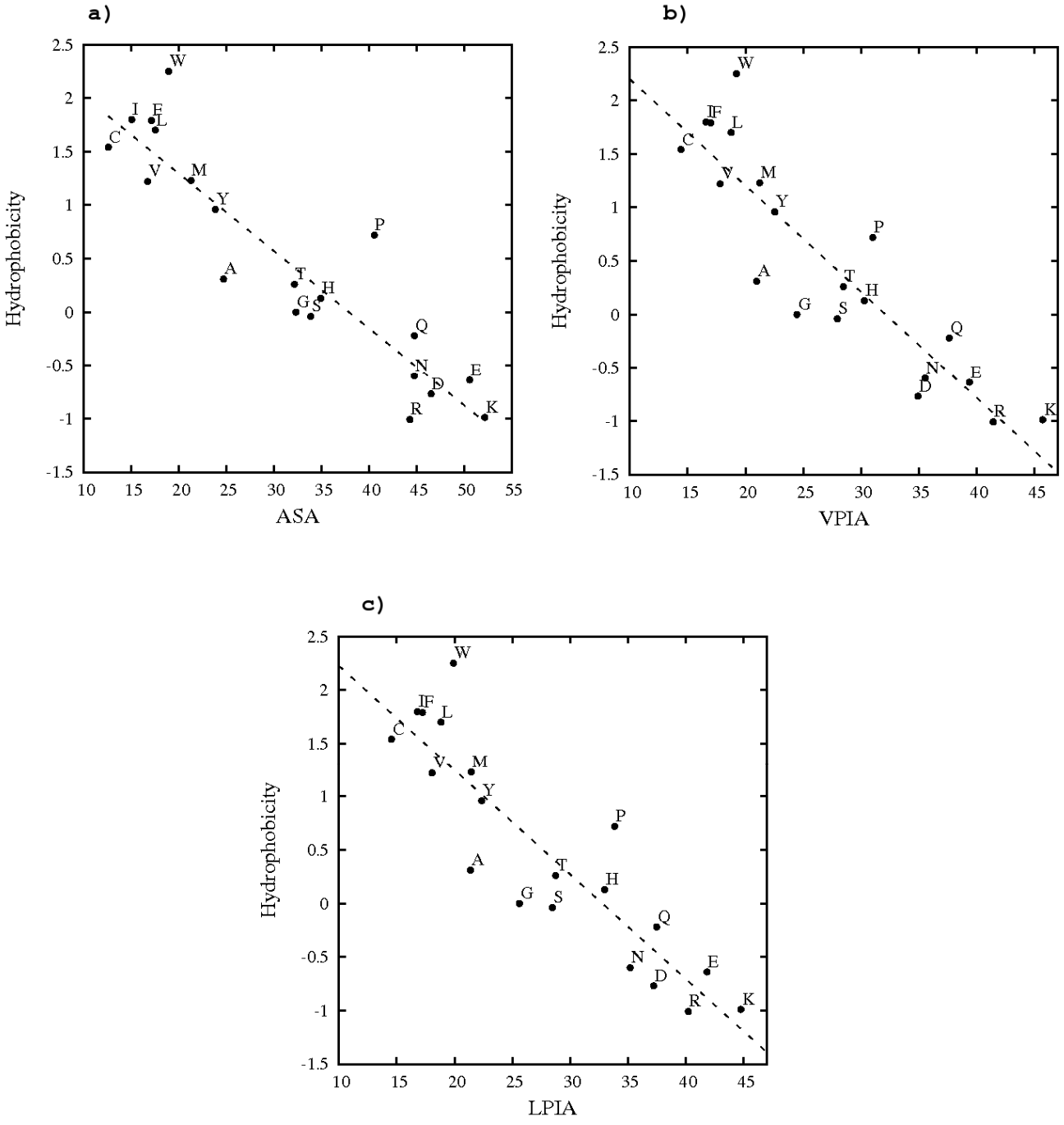
Supplementary data

Supplementary data 1. *Tessellations for proteins.* Given a set of sites X, positions and weights, the Delaunay tessellation divides space into tetrahedra whose vertices belong to X. More precisely, a tetrahedron based on 4 sites of X is Delaunay when its weighted circumscribed sphere does not enclose any other site of X. The Laguerre tessellation consists of polyhedra, often named Laguerre cells, one around each site of X. The Delaunay and Laguerre tessellations are dual of each other, as detailed below. Each Voronoi cell encloses exactly one site of X. In the Laguerre tessellation, depending on the weights, some sites may fall out of their cells. As an alternative definition, the Laguerre cell of a site x is the domain containing all the space points p closer to x than to any other site of X. Here, distance is the "radical distance" depending on the site weight w: $d_w(x, p) = \|x-p\|^2-w$. Radical distance is the square of length of the segment linking point p tangentially to the sphere centered at x and of radius $w^{1/2}$. In the Voronoi case, where all the weights are equal, distance reduces to the standard Euclidean distance. Thus, a Laguerre face marks a boundary equidistant between two sites (or spheres), a Laguerre edge is the locus of points equidistant to three nearby sites, and any Laguerre vertex is the center of the circumscribed sphere of a Delaunay tetrahedron with 4 sites as vertices. Algorithmically, the Laguerre tessellation is constructed as the dual of the Delaunay partition, which is built first. As it can be shown, the three sites around a Laguerre edge form a triangle in the Delaunay partition, which is perpendicular to the edge; there is a Laguerre face between two sites if and only if the two sites are linked by a Delaunay edge, which is then perpendicularly bisected by the plane containing the Laguerre face. In particular, when the atomic spheres intersect (on a circle), the Laguerre face lies in the plane which contains that circle.

Supplementary data 2. *Average volumes of buried residues*. The selection of buried residues obeys both ASA and V/LPIA equal to zero. For each method, Voronoi and Laguerre, the tessellation has been computed with (wH) or without water hydrogen (no H). Column *N* gives the number of each amino acid, <V> the average volumes and σ the standard deviations. The *global* line contains the total number of residues Sum(*N*), the overall average residue volume and the rms deviation (square root of weighted average of variances) . The *%buried* line gives the proportion of buried residues compared with all residues of databank (80,360 residues).

	Voronoi						Laguerre					
	wH			no H			wH			no H		
AA	***N***	***<V>***	**σ**	***N***	***<V>***	**σ**	***N***	***<V>***	**σ**	***N***	***<V>***	**σ**
A	495	87.7	4.9	576	87.7	5.0	557	87.1	4.7	591	87.2	4.7
C	73	111.1	6.0	83	111.4	5.9	81	113.7	6.0	84	113.6	6.1
D	20	128.2	5.9	26	128.5	6.6	25	122.5	5.2	27	122.7	5.5
E	4	144.7	4.1	7	147.7	6.8	7	143.2	6.0	8	143.0	5.6
F	126	190.5	7.3	154	190.4	7.3	145	192.0	6.8	156	192.1	6.8
G	292	66.0	4.1	341	66.1	4.2	323	64.4	3.8	345	64.5	3.9
H	13	157.7	5.4	15	157.9	5.1	16	157.8	4.8	17	157.9	4.7
I	300	156.5	6.9	361	156.4	6.8	344	157.5	6.4	359	157.5	6.3
K	1	157.8	0.0	1	157.8	0.0	1	159.8	0.0	1	159.8	0.0
L	370	157.9	6.8	449	157.7	6.9	432	158.5	6.7	461	158.3	6.6
M	45	161.2	7.0	52	161.3	6.7	54	163.4	6.4	55	163.6	6.3
N	20	130.2	4.8	24	130.3	6.8	24	125.4	4.9	25	124.9	5.4
P	30	118.4	5.6	40	118.2	5.8	36	120.6	5.0	42	120.4	5.0
Q	6	151.2	5.7	8	150.0	5.4	9	146.1	5.1	9	145.4	4.6
R	4	189.4	9.1	4	189.4	9.1	4	187.5	9.3	4	187.5	9.3
S	121	100.6	5.0	144	100.3	5.1	139	96.4	4.5	146	96.5	4.5
T	91	122.4	5.9	109	122.3	5.9	103	119.6	5.4	111	119.7	5.1
V	368	134.2	6.0	438	134.1	6.2	415	134.8	5.7	437	134.8	5.7
W	13	229.1	10.0	18	226.6	9.5	16	228.5	9.3	17	228.2	9.1
Y	38	200.1	6.6	45	199.7	7.3	42	198.7	7.0	44	198.5	7.1
global	2430	125.1	5.9	2895	125.6	6.0	2773	125.4	5.6	2939	125.3	5.6
%buried	3.0			3.6			3.5			3.7		

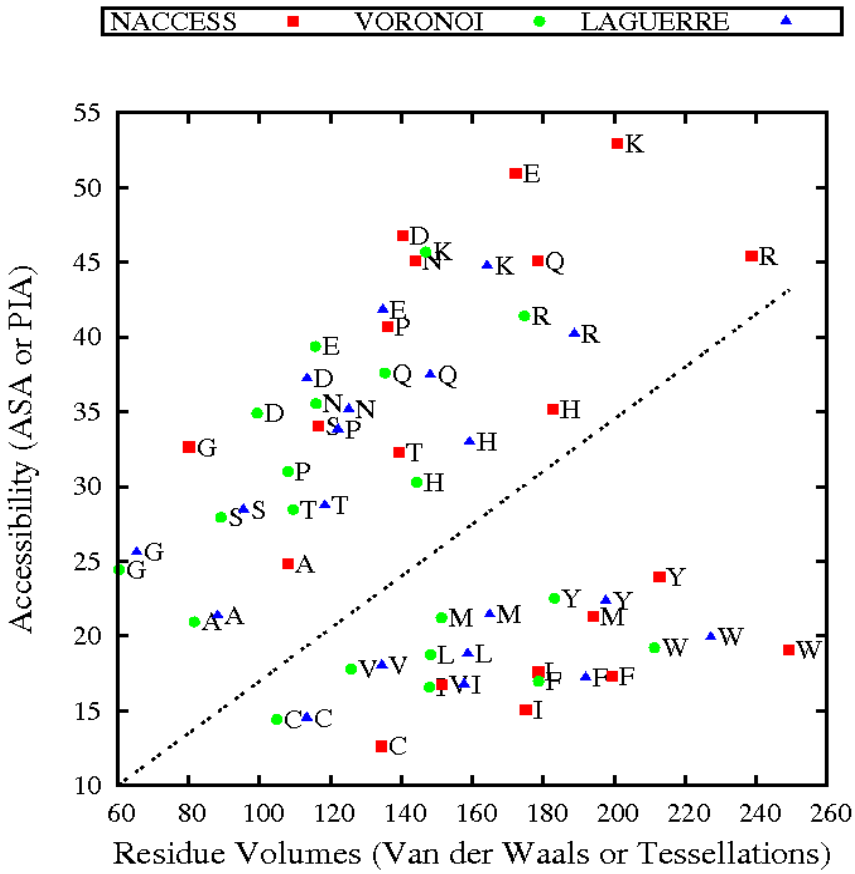
Supplementary data 3. *Hydrophobicity and accessibility*. Plot of hydrophobicity scale (kcal/mol) against (a) ASA, (b) VPIA and (c) LPIA.



Supplementary data 4. *Accessibility as indicator of hydrophobicity*. Starting from the hydrophobicity and accessibility data given in supplementary data 3, the equation $h = a\ x + b$ with the following coefficients was obtained by least squares regressions. Notice the very good Pearson correlation coefficients (PCC).

h = ax+b			
Accessibility	a	b	PCC
ASA	-0.072	2.74	-0.92
VPIA	-0.099	3.19	-0.90
LPIA	-0.098	3.20	-0.90

Supplementary data 5. *Mean accessibility versus residue volume*. For NACCESS, the volumes are taken from tables provided with the software (sum of VdW volumes) and accessibility is ASA. For Voronoi and Laguerre, the mean volumes and PIA are calculated from the tessellations. The diagonal line of equation $y = 0.175\ x - 0.50$ separates two clusters with different properties.



Supplementary data 6. *Mean residue volumes and protein size*. <*V*> is the mean Voronoi or Laguerre volume and σ the standard deviation. *N* is the number of amino acids. Four classes of protein size (*L*) have been defined: <150, 150-250, 251-400 and >400 amino acids. *%s* is the proportion of residues near surface (ASA>25%) .

Voronoi																
L	<150				150-250				250-400				>400			
AA	***N***	**<V>**	**σ**	**%s**	***N***	**<V>**	**σ**	**%s**	***N***	**<V>**	**σ**	**%s**	***N***	**<V>**	**σ**	**%s**
A	968	80.2	7.6	55	1741	81.7	7.9	41	2641	81.7	7.9	35	1192	82.4	8.1	33
C	187	102.0	9.0	22	245	104.5	8.9	18	353	106.3	9.4	16	149	105.9	9.3	15
D	784	96.6	9.1	86	1287	98.6	10.2	76	1733	100.2	11.3	71	884	100.9	11.3	68
E	1124	113.4	9.6	89	1633	115.0	10.4	84	2132	116.5	11.3	79	1003	117.8	12.3	76
F	533	176.1	14.0	36	953	178.6	14.0	27	1231	179.3	13.8	22	631	179.9	12.8	18
G	807	59.1	5.7	66	1434	60.4	6.2	53	2235	60.3	6.1	47	1094	61.2	6.3	42
H	290	141.1	9.8	78	495	143.8	10.6	55	751	144.5	11.2	54	329	147.5	11.9	40
I	839	145.7	11.7	33	1410	148.0	11.4	24	1814	148.5	11.0	19	793	149.3	10.9	14
K	832	145.7	8.9	94	1302	146.9	9.1	87	1742	147.3	8.8	84	860	147.0	9.3	82
L	1398	145.8	12.1	41	2259	148.4	11.8	27	3062	148.9	11.4	24	1424	149.1	10.9	19
M	157	148.1	12.0	52	220	150.9	12.6	39	428	152.0	11.9	30	186	152.8	10.8	20
N	559	114.4	8.8	84	919	115.4	9.4	76	1248	116.2	9.5	68	671	117.6	9.6	59
P	537	106.4	8.8	77	953	107.5	9.2	70	1453	108.4	9.1	62	716	109.1	9.7	53
Q	657	133.2	9.6	85	849	135.2	10.5	77	1061	135.8	10.7	70	542	137.2	11.4	65
R	714	172.4	9.3	89	1104	174.1	9.7	79	1540	175.4	10.1	73	747	176.5	10.4	60
S	796	86.9	7.5	71	1311	89.0	8.5	57	1758	89.5	8.7	52	829	90.6	9.0	44
T	707	107.0	8.9	69	1085	109.4	9.6	56	1549	110.1	10.0	49	808	110.6	10.0	47
V	1000	123.6	10.2	39	1612	126.3	10.0	28	2127	126.2	10.1	22	1042	126.5	9.8	21
W	174	207.4	14.4	43	271	211.0	14.3	28	413	212.4	13.0	23	210	213.2	13.8	22
Y	471	180.1	13.0	53	747	182.4	13.3	40	1087	184.2	13.8	35	528	185.2	12.9	25
glob	13534	123.7	9.9		21830	124.9	10.2		30358	124.8	10.3		14638	125.6	10.3	
Laguerre																
AA	***N***	**<V>**	**σ**		***N***	**<V>**	**σ**		***N***	**<V>**	**σ**		***N***	**<V>**	**σ**	
A	968	88.3	5.7		1741	88.4	5.3		2641	88.0	5.4		1192	88.1	5.6	
C	187	111.0	7.4		245	113.1	6.7		353	114.6	7.4		149	113.5	7.4	
D	784	112.6	5.3		1287	113.3	5.6		1733	113.6	6.0		884	113.9	6.0	
E	1124	133.9	5.8		1633	134.7	6.1		2132	135.0	6.2		1003	135.6	6.9	
F	533	191.6	8.2		953	192.6	9.0		1231	191.7	8.8		631	191.6	8.6	
G	807	65.4	4.3		1434	65.5	4.5		2235	65.1	4.4		1094	65.5	4.5	
H	290	159.7	7.5		495	159.1	7.1		751	158.9	7.5		329	159.7	7.6	
I	839	157.6	7.6		1410	157.9	7.7		1814	157.6	7.6		793	157.5	7.9	
K	832	164.0	7.1		1302	164.4	7.0		1742	164.2	6.7		860	163.9	7.0	
L	1398	158.7	8.1		2259	158.9	7.9		3062	158.8	7.9		1424	158.3	7.5	
M	157	165.3	6.8		220	165.3	7.9		428	164.9	8.5		186	164.0	7.7	
N	559	125.5	6.6		919	125.2	6.6		1248	124.9	6.4		671	125.3	6.2	
P	537	122.1	7.1		953	122.3	7.0		1453	122.0	6.7		716	122.1	6.9	
Q	657	147.8	7.0		849	148.5	7.6		1061	148.0	7.4		542	148.2	7.7	
R	714	188.8	7.6		1104	188.8	7.6		1540	188.9	7.6		747	188.7	7.9	
S	796	94.9	5.4		1311	95.5	5.5		1758	95.5	5.8		829	95.9	6.0	
T	707	118.0	6.2		1085	118.3	6.0		1549	118.5	6.4		808	118.6	6.6	
V	1000	134.3	6.9		1612	134.9	6.7		2127	134.3	7.0		1042	134.4	6.9	
W	174	226.6	9.1		271	227.9	9.4		413	226.8	8.2		210	228.0	9.4	
Y	471	197.4	8.1		747	197.6	8.0		1087	197.8	8.6		528	197.5	8.5	
glob	13534	137.1	6.8		21830	136.6	6.8		30358	135.5	6.9		14638	135.7	6.9	

Supplementary data 7. *Average residue Voronoi volumes according to secondary structure*. The secondary structures are divided in three classes: α-helix, β-sheet and coil. <*V*> is the mean Voronoi volume and σ the standard deviation. *N* gives the number of amino acids, and *%surf* is the proportion of residues near surface (ASA>25%) for each kind of amino acid.

	helix				sheet				coil			
AA	***N***	**<V>**	**σ**	**%surf**	***N***	**<V>**	**σ**	**%surf**	***N***	**<V>**	**σ**	**%surf**
A	3106	81.1	7.6	38	923	87.0	7.1	18	1836	79.9	7.6	50
C	270	106.3	8.8	15	229	107.8	8.3	7	324	102.1	9.4	74
D	1394	98.4	10.5	75	455	106.7	11.7	51	2367	98.8	10.3	22
E	2756	115.5	10.6	82	684	123.1	11.2	65	1781	113.7	10.8	14
F	1183	179.9	13.7	23	896	182.1	12.4	16	933	174.0	13.7	66
G	939	61.2	5.8	35	664	64.3	5.3	18	3424	59.4	6.1	40
H	567	145.1	11.2	54	334	148.3	9.9	38	778	142.6	11.0	38
I	1761	147.6	11.0	19	1560	151.9	10.5	15	1075	143.6	10.6	66
K	1894	146.4	8.7	85	674	151.8	8.7	76	1693	145.6	8.8	9
L	3751	148.7	11.3	24	1581	152.3	10.5	17	1999	144.9	11.6	63
M	469	151.9	11.7	30	198	156.2	11.2	20	299	147.2	11.4	53
N	930	116.0	9.4	69	376	121.1	9.8	50	1776	114.9	9.1	24
P	710	106.4	8.6	66	278	111.5	9.3	56	2293	108.1	9.2	36
Q	1399	135.1	10.6	74	414	140.0	10.3	60	962	133.8	10.3	21
R	1652	174.3	9.8	74	706	179.1	9.7	64	1310	173.1	9.8	20
S	1454	88.3	8.8	56	741	93.9	8.1	33	2046	88.1	8.1	38
T	1140	109.9	10.5	51	922	112.7	8.9	40	1681	107.6	9.1	37
V	1754	125.4	9.9	24	2086	128.8	9.3	18	1386	122.4	10.1	62
W	410	214.5	13.4	19	241	212.3	12.6	24	337	207.3	13.7	63
Y	1011	183.8	13.3	35	740	187.0	12.5	26	825	180.0	13.0	52
global	28550	128.3	10.3	49	14702	135.8	9.9	31	29125	116.2	9.7	38

Supplementary data 8. *Average volumes of buried residues in each secondary structure state*. The secondary structures are divided in three classes: α-helix, β-sheet and coil. <*V*> is the mean Voronoi or Laguerre volume and σ the standard deviation. *N* is the number of buried amino acid. The buried residues have been selected by ASA equal to zero.

Voronoi									
	helix			sheet			coil		
AA	***N***	**<V>**	**σ**	***N***	**<V>**	**σ**	***N***	**<V>**	**σ**
A	311	86.5	4.4	141	90.4	4.6	27	87.8	5.3
C	36	111.5	5.2	24	109.3	5.9	8	112.7	7.5
D	11	127.4	3.8	6	128.4	5.5	3	130.5	10.4
E	3	145.0	4.7	1	143.8	0.0	0	0.0	0.0
F	63	190.9	7.3	53	190.7	7.6	6	185.6	3.4
G	108	64.5	3.6	114	67.2	3.7	56	66.1	4.4
H	6	155.8	5.3	3	161.5	4.7	3	159.5	3.8
I	137	155.1	6.7	138	157.8	6.7	16	156.7	6.7
K	0	0.0	0.0	1	157.8	0.0	0	0.0	0.0
L	216	156.9	6.6	115	160.2	6.7	25	157.5	6.8
M	34	160.3	7.0	10	164.4	6.6	0	0.0	0.0
N	12	130.2	4.8	2	131.8	3.7	5	129.6	5.6
P	8	114.2	5.1	7	124.0	4.6	14	118.3	3.9
Q	3	148.0	3.3	1	158.4	0.0	2	152.4	6.1
R	2	197.8	4.9	0	0.0	0.0	2	181.0	0.8
S	62	100.5	4.8	37	101.9	5.3	17	98.7	4.4
T	43	122.4	6.6	32	123.7	5.6	14	119.6	3.3
V	131	133.0	6.6	209	135.2	5.4	13	131.5	5.9
W	9	230.3	11.6	3	225.8	3.7	1	227.9	0.0
Y	20	201.1	6.5	14	197.8	3.9	2	206.3	9.7
global	1215	124.8	5.8	911	128.7	5.7	214	112.5	5.4
Laguerre									
	helix			sheet			coil		
AA	***N***	**<V>**	**σ**	***N***	**<V>**	**σ**	***N***	**<V>**	**σ**
A	356	86.0	4.2	150	89.8	4.5	32	87.8	4.5
C	39	114.0	5.0	27	112.4	6.3	8	115.3	7.3
D	11	121.9	3.3	6	122.9	4.7	7	123.5	7.7
E	4	140.1	2.4	2	143.8	5.2	1	154.9	0.0
F	76	192.3	6.8	58	192.4	7.2	6	188.0	2.0
G	121	63.1	3.3	124	65.4	3.5	64	64.7	4.0
H	7	155.7	4.9	4	160.3	4.1	4	160.1	3.2
I	154	156.1	6.1	160	159.0	6.1	19	157.8	6.5
K	0	0.0	0.0	1	159.8	0.0	0	0.0	0.0
L	251	157.6	6.8	138	160.5	6.5	29	158.2	6.0
M	38	162.6	6.5	14	165.5	5.9	1	162.0	0.0
N	14	124.0	4.5	4	130.5	5.4	5	124.8	3.2
P	10	116.9	4.3	7	125.7	5.1	18	120.7	3.4
Q	3	142.3	4.1	3	147.8	4.4	2	146.1	4.1
R	2	196.3	2.9	0	0.0	0.0	2	178.7	3.2
S	73	95.9	4.3	43	98.0	4.7	17	95.2	3.8
T	50	118.6	5.7	33	121.4	4.8	16	118.1	3.8
V	151	133.4	6.2	231	135.8	5.2	17	133.1	6.1
W	11	228.2	10.7	4	228.4	5.3	1	231.6	0.0
Y	23	198.5	7.6	14	197.4	3.2	2	205.5	8.7
global	1394	124.8	5.6	1023	129.6	5.4	251	112.6	4.9

Supplementary data 9. *Sensitivity of the volumes to the replacement of residue side chains*. A subset of 18 proteins was taken from our databank as a test case. The side chains were replaced using four different softwares: SCWRL 4 (S4) [72], SCWRL 3 (S3) [73], SCCOMP (SO) [74] and SCATD (SC) [75]. Then our protocol was applied to all the structures: generation of solvent and minimization by GROMACS, Laguerre tessellation. The following Table gives the residue mean volumes (V) and standard deviations (σ) for each software used (S4, S3, SO and SC). For comparison, the values obtained for the original, reference, structures (no replacement) are provided in columns *V real* and *σ real*. The averages are reported for the whole set of residues in the 18 proteins with the amino acids number in column *N*. Finally, Rmsd with the reference structure (used for *V real*) was calculated using GROMACS dropping out all hydrogens. *Rmsd all* includes all atoms, except hydrogen; *rmsd sc* is restricted to side chains, averaged over the 18 proteins in each case. The row, named *score*, gives the weighted average of absolute differences between references volumes (*V real*) and the others (*V S4*, *S3*, *SO* and *SC*) on all residues.

	AA	*N*	V real	σ real	V S4	σ S4	V S3	σ S3	V SO	σ SO	V SC	σ SC
	A	393	88.0	5.18	89.7	5.51	90.1	5.20	90.2	5,90	90,2	5,41
	C	74	115.2	7.35	118.4	5.96	118.5	5.58	117.1	5,76	118,4	6,33
	D	298	113.2	5.79	114.1	5.97	113.9	5.50	113.8	6,09	114,4	5,60
	E	343	135.0	6.42	137.1	6.08	137.4	6.73	136.9	6,91	136,7	7,22
	F	203	191.5	9.31	196.9	8.20	197.7	8.72	198.4	9,25	199,3	9,30
	G	336	65.1	4.22	66.6	4.36	67.1	4.50	66.8	4,54	67,3	4,96
	H	129	159.2	7.59	164.4	9.19	165.2	7.72	165.3	8,43	166,9	8,68
	I	317	158.2	8.13	161.2	7.63	161.4	7.55	161.7	7,62	162,4	7,95
	K	284	164.0	6.60	167.8	7.45	168.2	6.73	168.2	7,67	168,5	7,16
	L	490	158.7	7.72	161.8	7.37	162.2	7.42	162.4	7,54	163,3	8,21
	M	68	15.2	9.10	169.7	9.57	170.5	9.13	168.5	7,68	172,2	10,07
	N	235	125.3	6.32	128.1	6.78	129.0	7.11	129.4	6,71	129,1	6,18
	P	241	122.3	6.53	124.6	7.05	125.5	7.00	125.2	6,29	125,5	7,05
	Q	203	148.0	7.43	151.9	7.11	152.0	6.74	153.3	7,36	152,4	7,54
	R	270	189.2	7.06	193.2	8.10	194.1	7.96	193.3	8,33	194,3	8,08
	S	299	95.4	5.90	97.7	5.81	97.8	6.00	97.7	5,71	98,8	5,91
	T	278	117.9	6.20	120.6	6.17	121.1	6.34	120.8	6,46	121,1	6,56
	V	396	134.7	7.06	137.8	7.04	137.4	6.54	137.3	6,40	138,2	7,06
	W	58	227.0	7.23	237.7	9.06	239.0	8.38	239.1	9,85	237,7	9,54
	Y	152	197.2	8.43	205.0	9.00	205.7	9.28	205.3	9,38	206,3	9,17
rmsd all					0.13	0.02	0.13	0.01	0.14	0,03	0,13	0,01
rmsd sc					0.17	0.02	0.18	0.02	0.19	0,03	0,18	0,02
score					3.07		3.46		3.39		3.92	