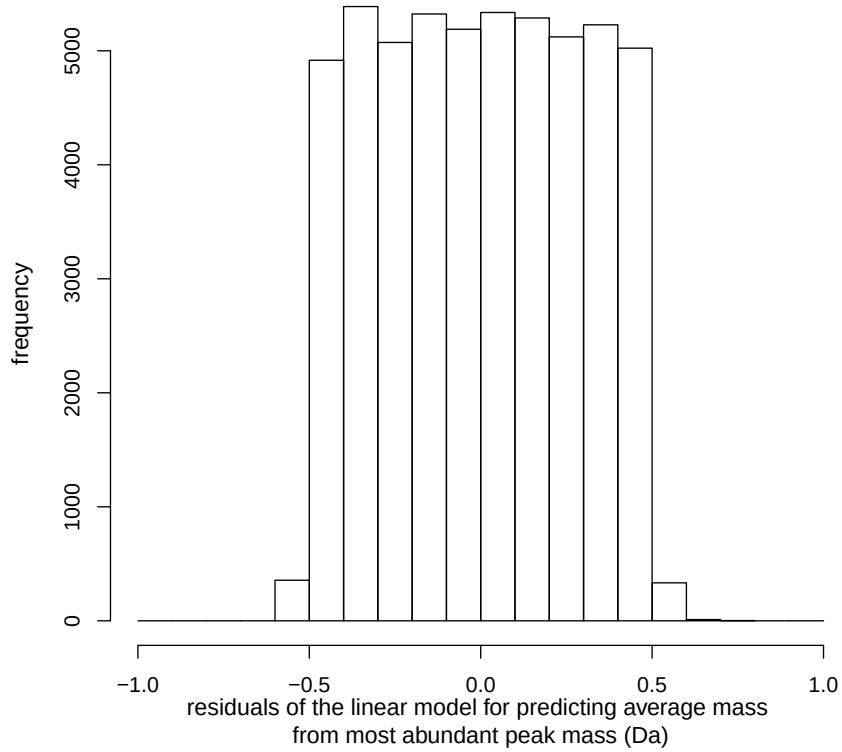

Suppl. Figure 1: Residuals of the linear model predicting average mass from the most abundant peak mass with standard deviation of 0.29.

Suppl. Table 1: Computational time (in seconds) for selected biomolecules calculated by R implementation of BRAIN algorithm. For each molecules there are provided two values (in seconds, average from 20 runs): (a) time1 - time for calculating peaks up to 10th peak after obtaining local maximum (in total noVaraints1 peaks calculated), (b) time2 - time for calculating requested number of variants (noVariants2) defined for each molecules separately. The computations were made on PC with two Intel(R) Core(TM)2 2.40GHz CPUs. The results for simple MATLAB implementation (for noVariants2 peaks) were taken from (Claesen et al. 2012). Please, note the computations for R and MATLAB were performed on different machines.

	name	formula	noVariants1	time1	noVariants2	time2	MATLABtime
1	Angiotensin II	$C_{50}H_{71}N_{13}O_{12}$	11	0.01	50	0.02	0.04
2	Bovine insulin	$C_{254}H_{377}N_{65}O_{75}S_6$	14	0.01	50	0.02	0.04
3	Human insulin	$C_{520}H_{817}N_{139}O_{147}S_8$	18	0.01	50	0.02	0.04
4	Human myoglobin	$C_{744}H_{1224}N_{210}O_{222}S_5$	21	0.01	100	0.04	0.04
5	Human intrinsic factor	$C_{2023}H_{3208}N_{524}O_{619}S_{20}$	39	0.02	322	0.12	0.07
6	Bovine serum albumin	$C_{2934}H_{4615}N_{781}O_{897}S_{39}$	53	0.02	400	0.15	0.07
7	Human Na/K ATPase	$C_{5047}H_{8014}N_{1338}O_{1495}S_{48}$	81	0.03	643	0.28	0.16
8	Renal isoform, subunit Human ATP binding cassette protein	$C_{8574}H_{13378}N_{2092}O_{2392}S_{77}$	129	0.05	807	0.40	0.22
9	Human intrinsic factor – hydroxocobalamin receptor	$C_{17600}H_{26474}N_{4752}O_{5486}S_{197}$	262	0.10	1163	0.67	0.36
10	Human dynein heavy chain	$C_{23832}H_{37816}N_{6528}O_{7031}S_{170}$	341	0.14	1325	0.79	0.41