Supporting Information for

InterSpin: Integrated Supportive Webtools for Low- and High-Field NMR Analyses Toward Molecular Complexity

Shunji Yamada[†‡], Kengo Ito[‡§], Atsushi Kurotani[‡], Yutaka Yamada[‡],

Eisuke Chikayama[‡∥] and Jun Kikuchi[*†‡§]

[†] Graduate School of Bioagricultural Sciences, Nagoya University, 1 Furo-cho, Chikusa-ku, Nagoya, Aichi 464-0810, Japan

[‡] RIKEN Center for Sustainable Resource Science, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

[§] Graduate School of Medical Life Science, Yokohama City University, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

[∥] Department of Information Systems, Niigata University of International and Information Studies, 3-1-1 Mizukino, Nishi-ku, Niigata-shi, Niigata 950-2292, Japan

Supplementary material

In this document, we provide the following information:

・Supplementary figures illustrating the main functionalities of Interspin.

InterSpin is available at http://dmar.riken.jp/interspin/

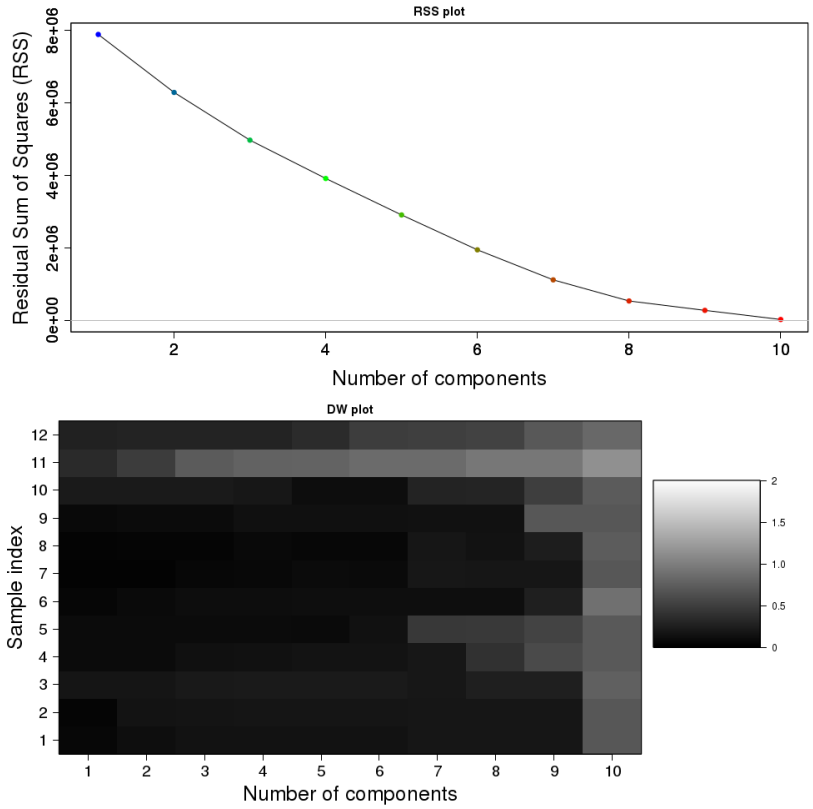
Corresponding Author

*RIKEN Center for Sustainable Resource Science, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan.

Tel: +81455039439; Fax: +81455039489

Email:jun.kikuchi@riken.jp

a) Fast ICA



b) MCR-ALS

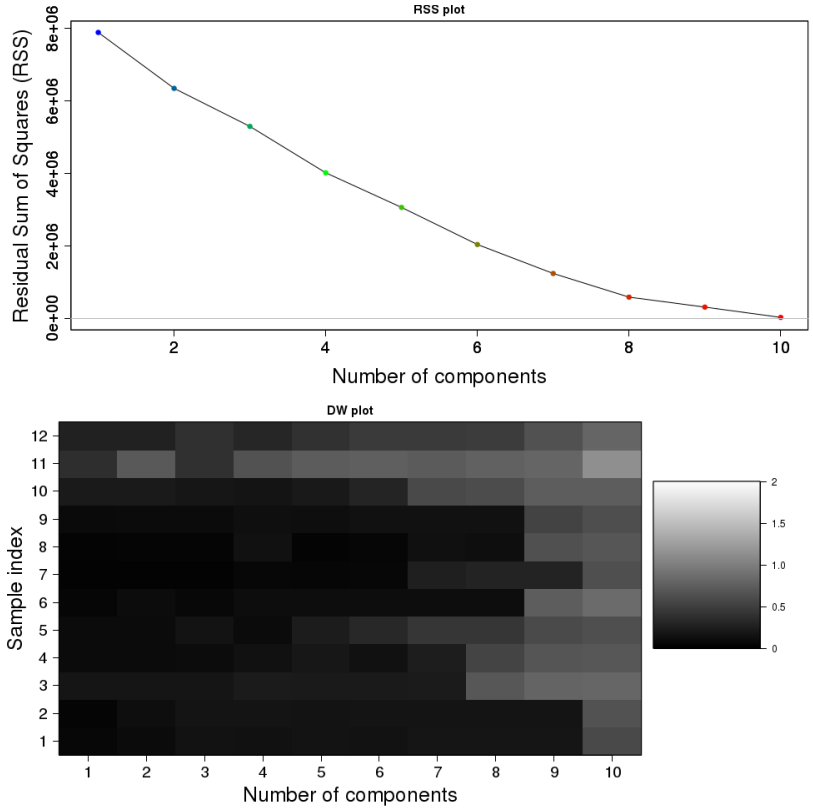
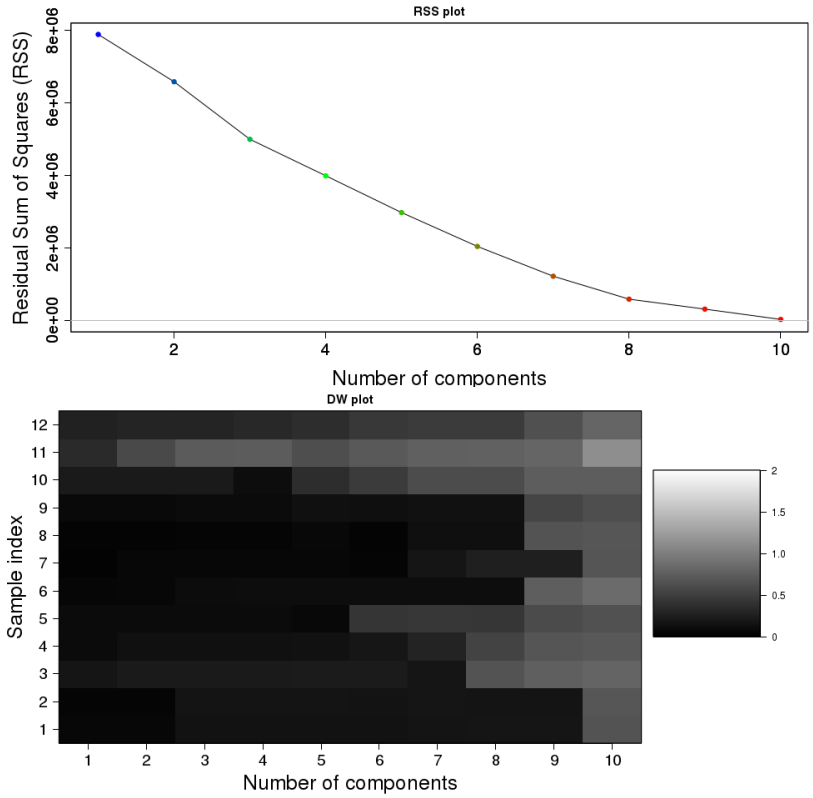


Figure S1: Determining the number of components by RSS and DW plots. (**a**) Fast ICA, and (**b**) MCR-ALS.

c) NMF



d) NNSC

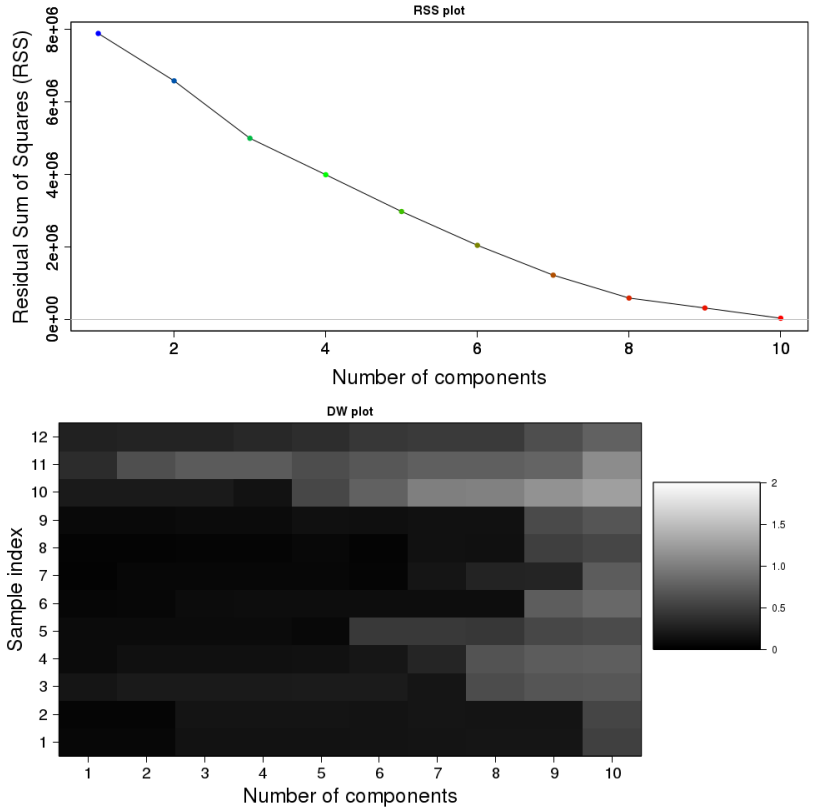


Figure S1: Determining the number of components by RSS and DW plots (continuation). (**c**) NMF, and (**d**) NNSC.

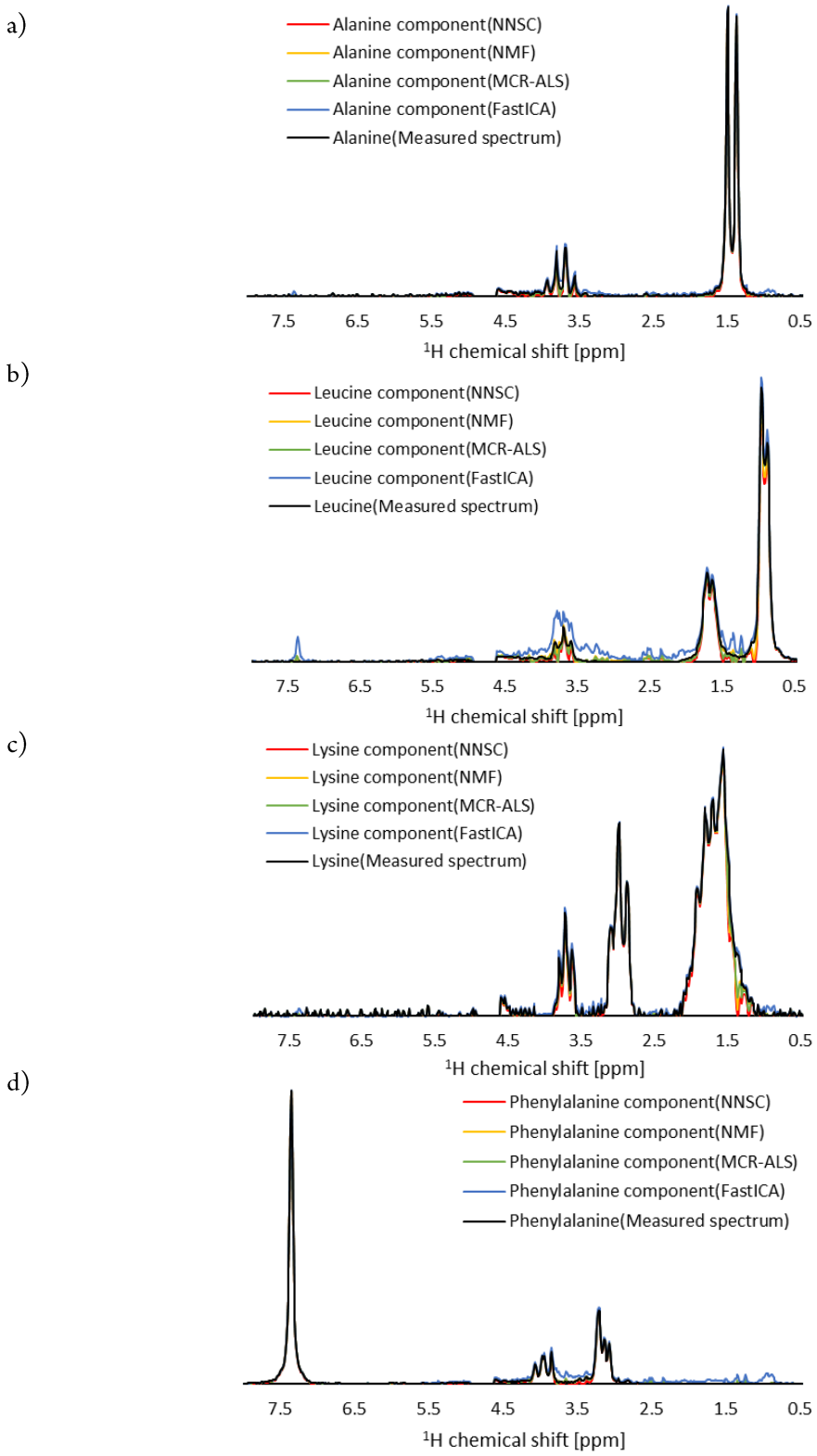
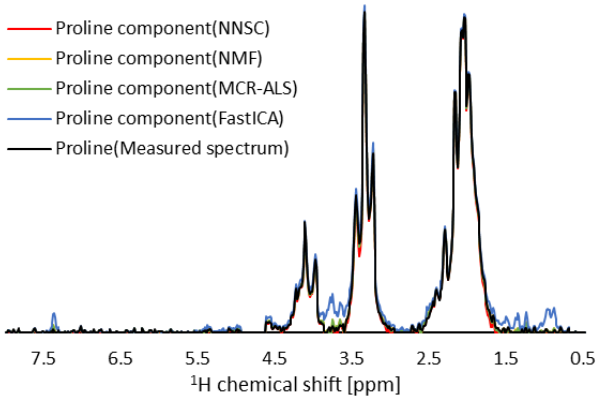
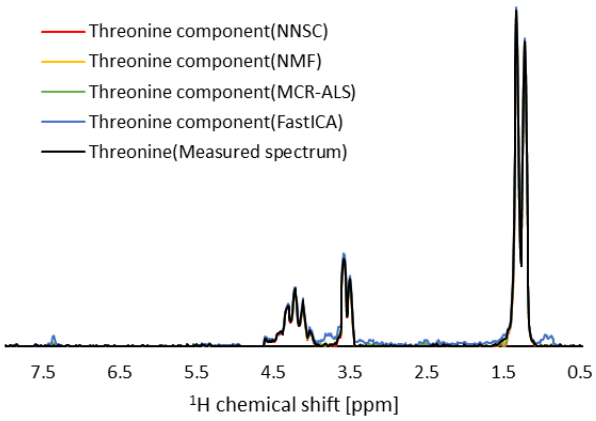


Figure S2: Comparison of the component spectrum separated by different algorithms and the standard spectrum. (**a**) Alanine, (**b**) leucine, (**c**) lysine, and (**d**) phenylalanine.

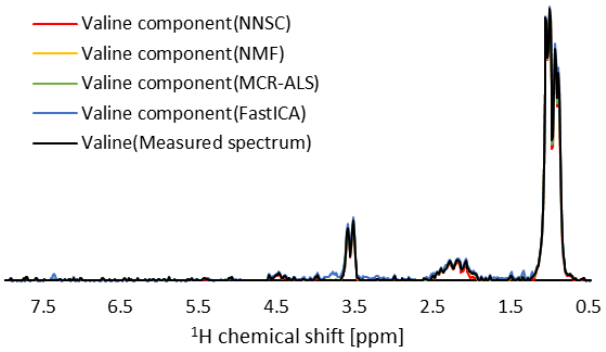
e)



f)



g)



h)

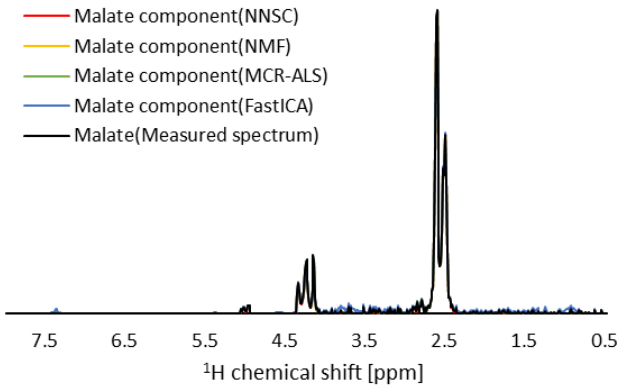
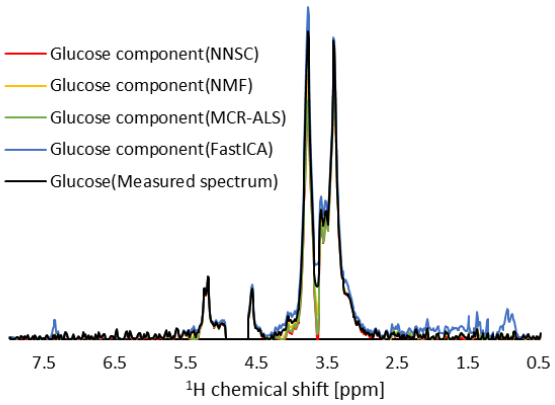


Figure S2: Comparison of the component spectrum separated by different algorithms and the standard spectrum (continuation). (**e**) Proline, (**f**) threonine, (**g**) valine, and (**h**) malate.

i)



j)

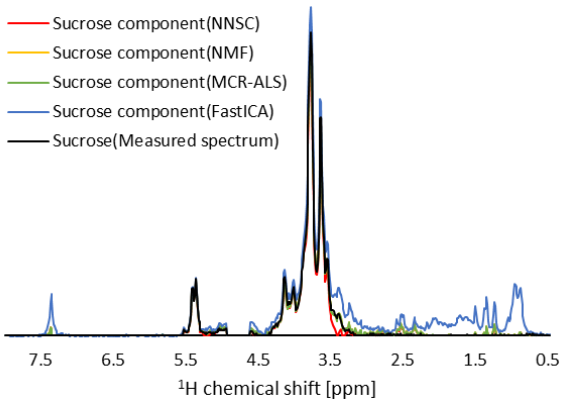


Figure S2: Comparison of the component spectrum separated by different algorithms and the standard spectrum (continuation). (**i**) Glucose, and (**j**) sucrose.

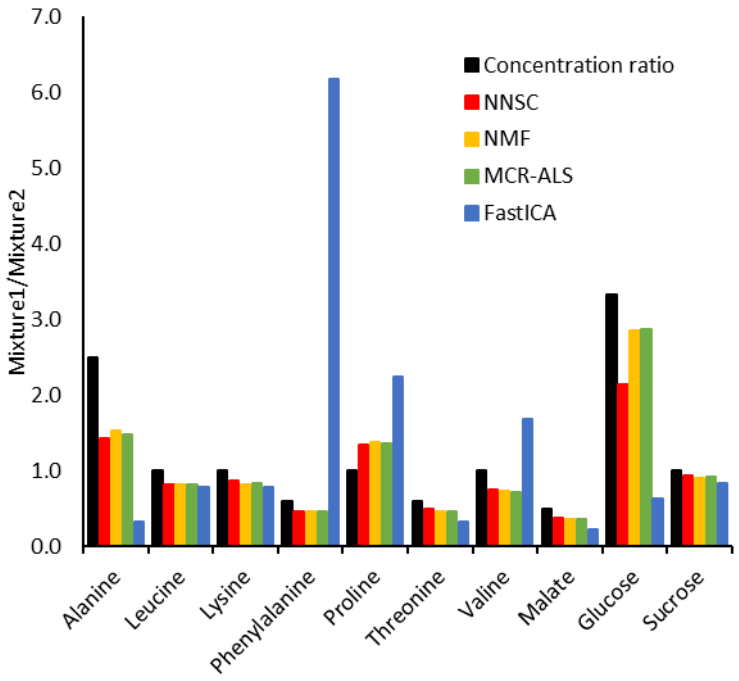


Figure S3: Comparison of concentration ratios and separated spectral score ratios of mixtures 1 and 2. Shown are the concentration ratio of mixtures 1 and 2, and the score ratio of the spectrum separated by each algorithm (NNSC, NMF, MCR-ALS, and Fast ICA).

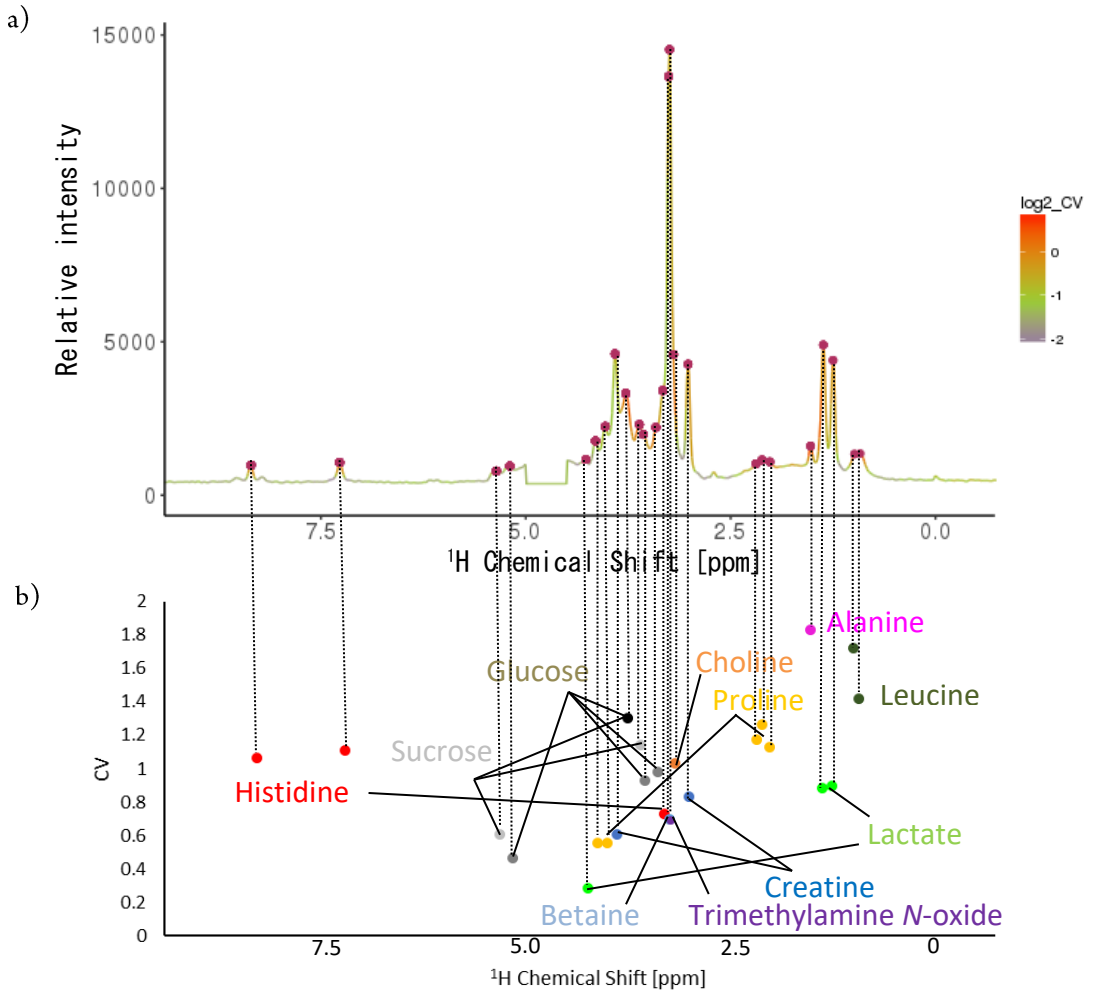
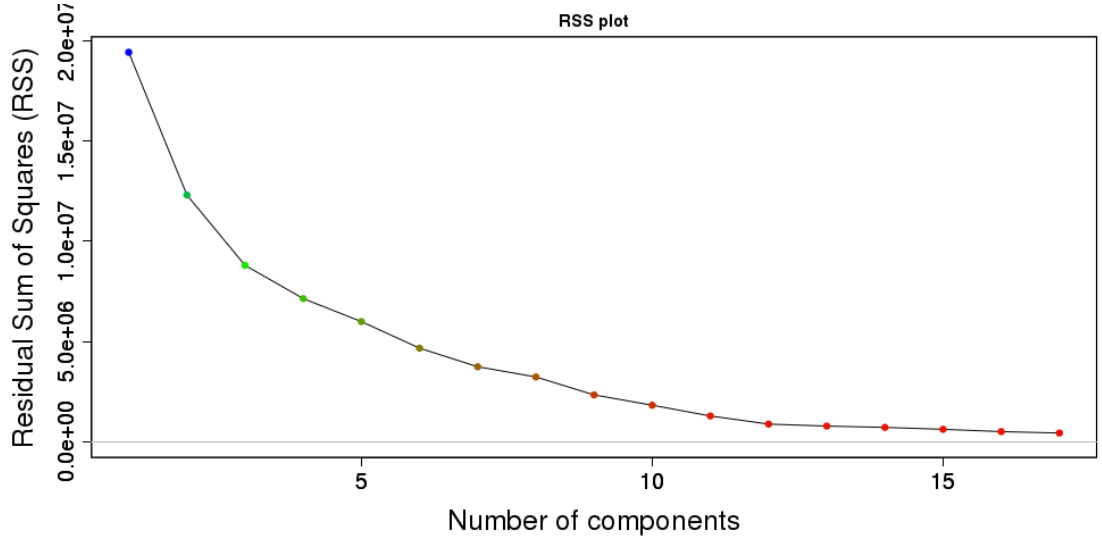


Figure S4: SENSI (SENsitivity improvement with Spectral Integration) result for integrated 51 fish spectra of fish extracts obtained by benchtop 60 MHz NMR. (**a**) Picked peaks, and (**b**) CV.

a)



b)

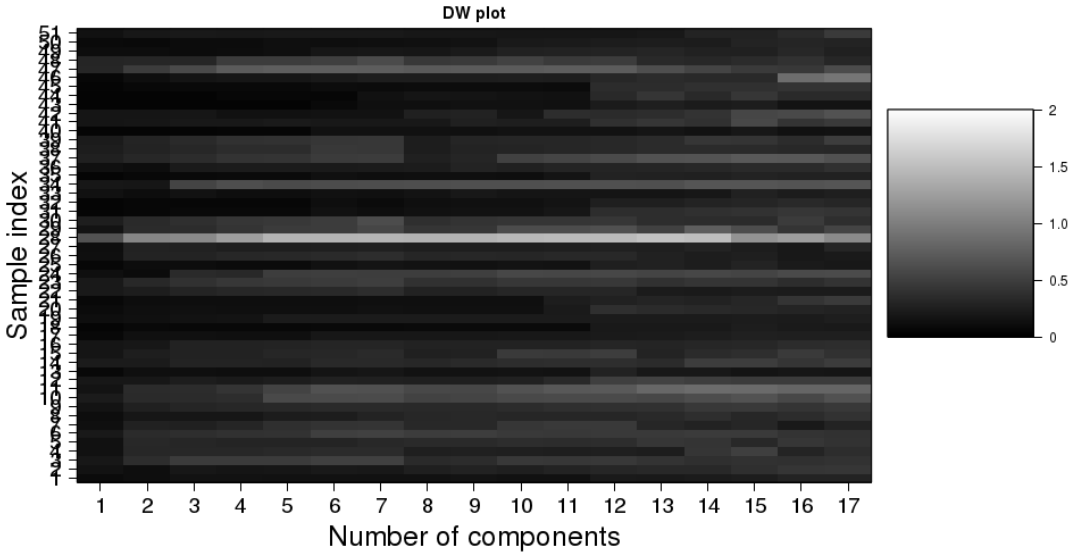


Figure S5: PKSP results of spectrum from Fig. S5a using the NNSC method. (**a**) RSS plot, and (**b**) DW plot.

c)

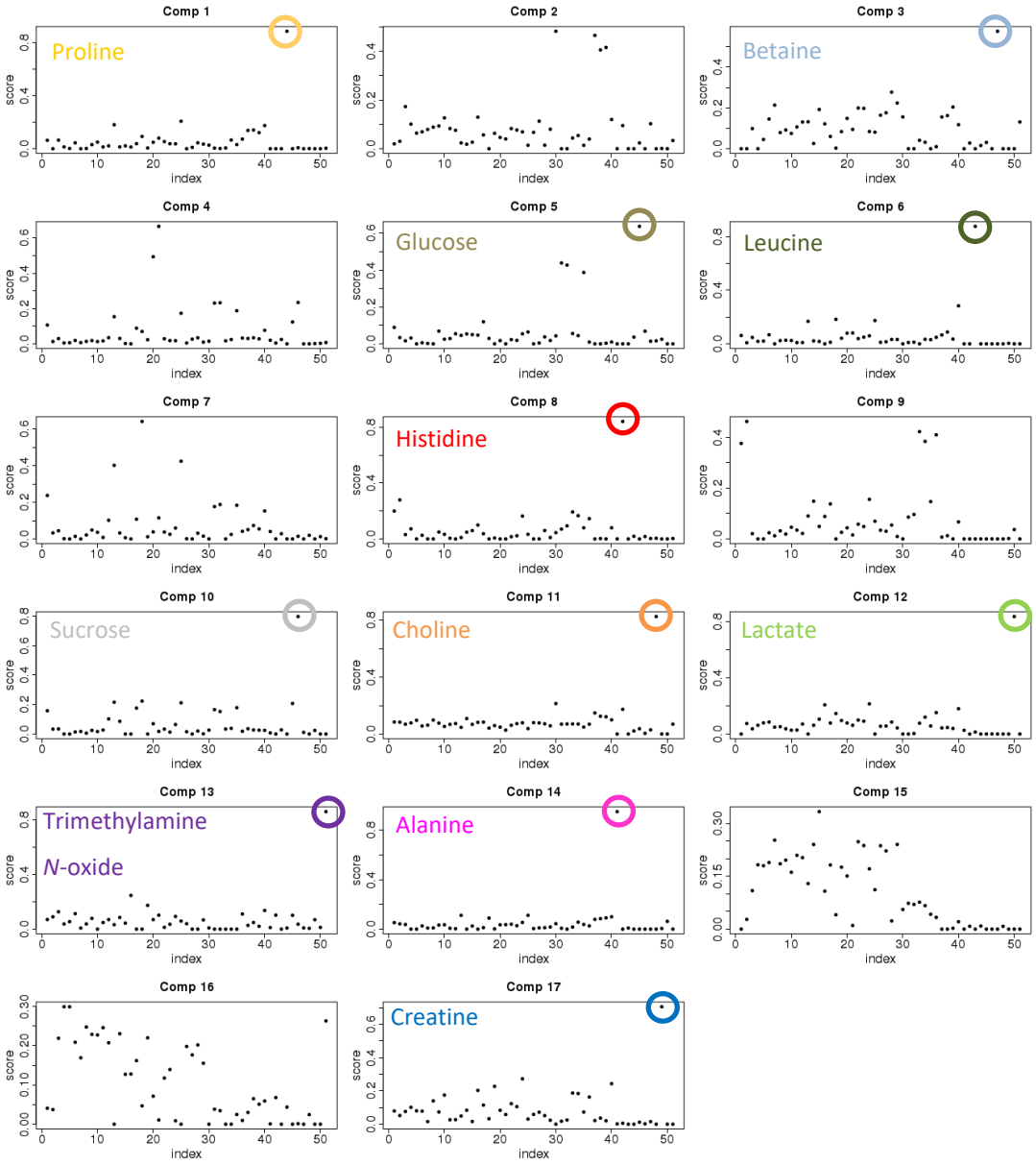


Figure S5: PKSP results of spectrum from Fig. S5a using the NNSC method (continuation). (**c**) Plot of scores for different components.

d)

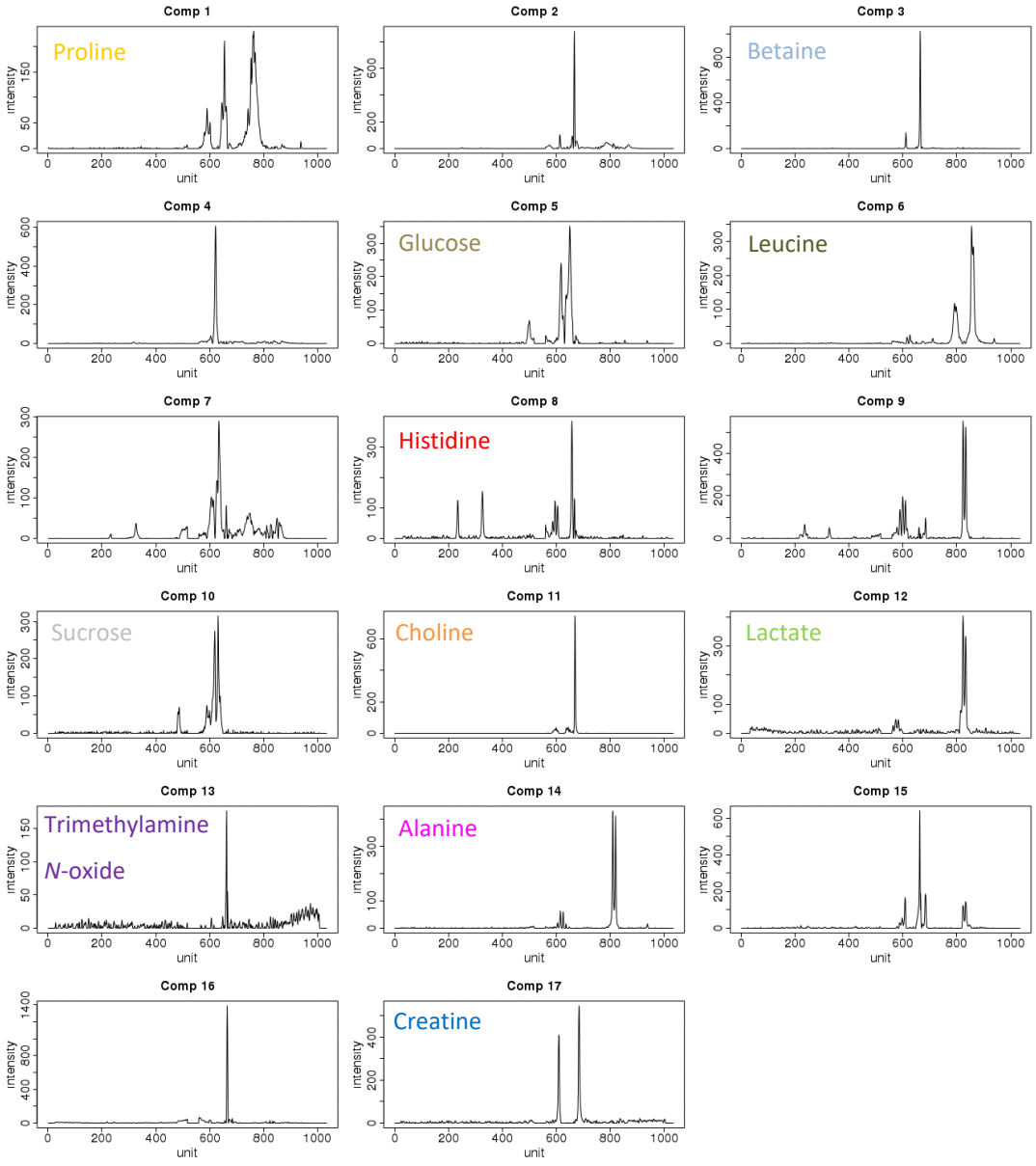


Figure S5: PKSP results of spectrum from Fig. S5a using the NNSC method (continuation). (**d**) Plot of loadings for different components.

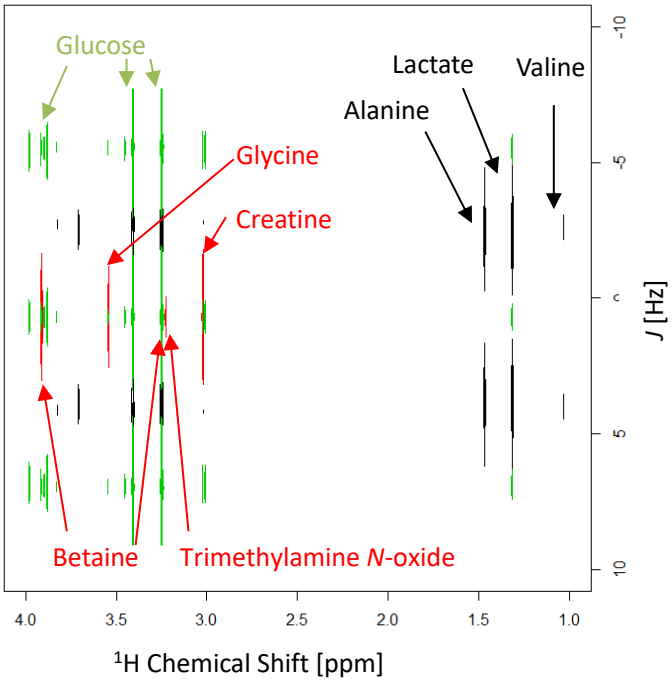


Figure S6: The result of separating one spectrum of 2D-*J*res from *Acanthogobius flavimanus* (Yellowfin goby) body muscle extracts in D_2O by PKSP. As a result of peak separation using PKSP's MCR-ALS algorithm, it was separated into three components of singlet (red), doublet (black) and triplet (multiplet, green). As shown in figure S6, 8 compounds (Valine, Lactate, Alanine, Creatine, Trimethylamine *N*-oxide, Betaine, Glycine and Glucose) were assigned.

Summary

QUERY PEAK	ANNOTATED PEAK
35	31 (88.571428571429%)

QUERY			ANNOTATION	
PEAK	^{1}H PPM	^{13}C PPM	COMPOUND & ATOM	COUNTS
1		180.7	Lignin(Nakarai) ; Xylan(Birchwood) ; Xantangum ; Lipid ;	4
2		172.7	Chitin ; Peptide ;	2
3		156.8	Peptide ;	1
4		148.3	Lignin(Nakarai) ; Lignin(Nakarai) ;	2
5		136.5	Peptide ;	1
6		129.3	Lignin(Nakarai) ; HumicAcid ; Peptide ; Peptide ; Lipid ;	5
7		128.3	Lignin(Nakarai) ; HumicAcid ; HumicAcid ; Peptide ;	4
8		118.2	Lipid ;	1
9		105.3	AlphaCellulose ; Rhamnogalacturonan ; Inulin ; Carrageenan ; Chitosan ; Galactan(Potate) ; Curdran ; Agarose ; Agarose ; Cellulose ; Paramylon ;	11
10		103.5	Laminarin ; Galactan ; Rhamnogalacturonan ; Arabinoxylan(Wheat) ; Arabinoxylan(Wheat) ; Inulin ; Inulin ; Chitin ; Xyloglucan ; Starch(Wheat) ; Curdran ; Pullulan ; Lichenan ; Lichenan ; Laminaran ; Laminaran ; Agarose ; Agarose ; Peptide ;	19
11		100.9	AlginicAcid ; AlginicAcid ; AlginicAcid ; Fucoidan ; Pectin(Apple) ; Rhamnogalacturonan ; Arabinoxylan(Wheat) ; Xylan(Birchwood) ; Starch(Wheat) ; Starch(Wheat) ; PolygalacturonicAcid ; PolygalacturonicAcid ; Galactomannan(Guar) ; Agarose ; Agarose ;	15
12		92.7	Galactan ; Fucoidan ; Lichenan ; Lichenan ; Dextran ; Dextran ;	6
13		90.4	Laminarin ; Laminarin ; Carrageenan ; Curdran ; Curdran ; Lichenan ; Lichenan ; Laminaran ;	8
14		83.5	Laminarin ; AlphaCellulose ; AlphaCellulose ; Rhamnogalacturonan ; Arabinoxylan(Wheat) ; Inulin ; Chitin ; Xyloglucan ; Chitosan ; Chitosan ; Arabinan ; Pullulan ; Lichenan ; Agarose ; Xantangum ; Cellulose ; Paramylon ;	17
15		79.7	AlginicAcid ; Fucoidan ; Pectin(Apple) ; Pectin(Apple) ; Pectin(Apple) ; Rhamnogalacturonan ; Arabinoxylan(Wheat) ; Galactomannan(Carob) ; Glucomannan ; PolygalacturonicAcid ; PolygalacturonicAcid ; PolygalacturonicAcid ; Galactan(Potate) ; Arabinan ; Arabinan ; Lichenan ; Agarose ; Xantangum ;	18
16		75.9	Laminarin ; AlphaCellulose ; AlginicAcid ; Pectin(Apple) ; Inulin ; Chitin ; Starch(Wheat) ; Chitosan ; Curdran ; Lichenan ; Agarose ; Paramylon ;	12
17		73.1	AlphaCellulose ; Galactan ; Fucoidan ; Chitin ; Lignin(Nakarai) ; Xyloglucan ; Glucomannan ; Galactan(Potate) ; Arabinan ; Arabinan ; Arabinan ; Curdran ; Curdran ; Pullulan ; Xantangum ; Xantangum ; Cellulose ; Peptide ;	18
18		68.8	Laminarin ; AlginicAcid ; Fucoidan ; Pectin(Apple) ; Lignin(Nakarai) ; Carrageenan ; PolygalacturonicAcid ; Galactan(Potate) ; Arabinan ; Xantangum ; Paramylon ;	11
19		62.6	Laminarin ; Galactan ; Galactan ; Xylan(Beechwood) ; Arabinoxylan(Wheat) ; Inulin ; Xylan(Birchwood) ; Galactomannan(Carob) ; Galactomannan(Carob) ; Xyloglucan ; Glucomannan ; Carrageenan ; Starch(Wheat) ; Galactan(Potate) ; Arabinan ; Curdran ; Pullulan ; Pullulan ; Pullulan ; Galactomannan(Guar) ; Laminaran ; Laminaran ; Agarose ; Cellulose ; Paramylon ;	25
20		59.4	Galactan ; Galactan ; Lignin(Nakarai) ; Galactan(Potate) ; Arabinan ; Agarose ;	6
21		53.1	Pectin(Apple) ; Peptide ;	2
22		48.3		
23		42.8	Xylan(Beechwood) ;	1
24		39.6	Peptide ;	1
25		37.4		
26		33.7	Peptide ; Lipid ;	2
27		32.3	Galactan ; Fucoidan ; Peptide ; Lipid ;	4
28		30.1	Lignin(Nakarai) ; HumicAcid ; Peptide ; Lipid ;	4
29		27.3		
30		24.8	Arabinoxylan(Wheat) ; Lignin(Nakarai) ; Xylan(Birchwood) ; Xantangum ; Xantangum ; Peptide ; Lipid ;	7
31		24.3	Arabinoxylan(Wheat) ; Lignin(Nakarai) ; Xylan(Birchwood) ; Xantangum ; Peptide ; Peptide ; Lipid ;	7
32		22.9	Chitin ; Peptide ;	2
33		20.5	Fucoidan ; Lignin(Nakarai) ; HumicAcid ; Xantangum ; Peptide ;	5
34		14.3	Lignin(Nakarai) ; Peptide ; Lipid ;	3
35		12.3		

QUERY PEAK=Query peak No., QUERY 1H PPM=1H chemical shift of query, QUERY 13C PPM=13C chemical shift of query, ANNOTATION COMPOUND & ATOM=All annotated compounds with atom names, ANNOTATION COUNTS=No. of annotations

Figure S7: SpinMacro assignment results of peaks picked by SENSI of the previously reported solid-state CP-MAS spectra of *Euglena gracilis* and standards (paramylon, peptide, lipid). Regarding the assigned chemical shift, paramylon is indicated by the red frame, peptides by the black frame, and lipids by the blue frame. The ^{13}C tolerance is 1 ppm.

a)

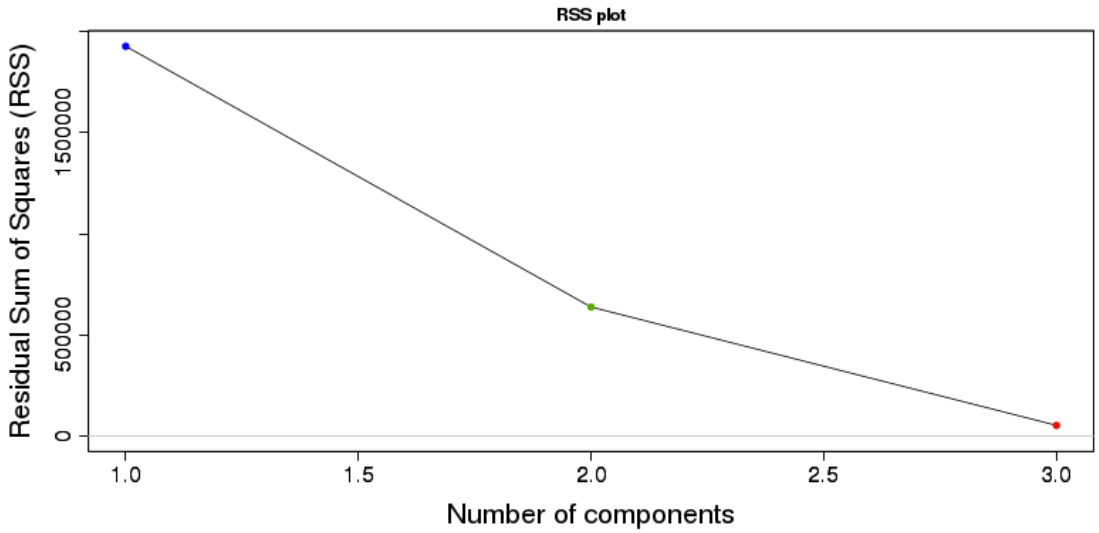


Figure S8: Signal separattion by NNSC of PKSP from *E. gracilis* CP-MAS spectrum. (**a**) RSS plot shows that there are three main components.

b)

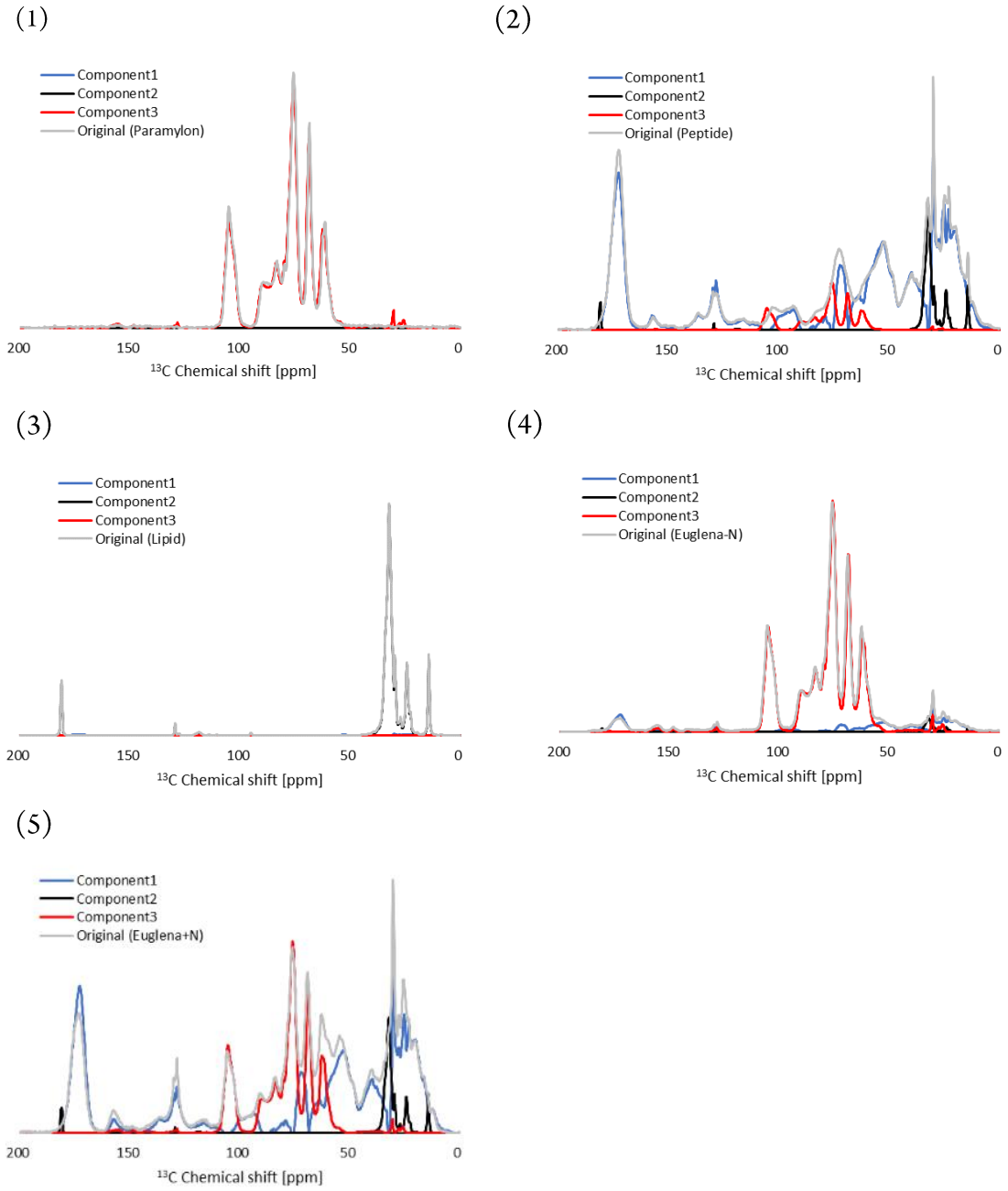


Figure S8: Signal separation by NNSC of PKSP from *E. gracilis* CP-MAS spectrum (continuation). (**b**) Separated spectra of paramylon, peptide, and lipid. Component 1 was identified as peptide (b-1), component 2 was identified as lipid (b-2), and component 3 was identified as paramylon (b-3). The *E. gracilis* sample was separated into 3 components (b-4, b-5). The variation in paramylon and peptide components indicates metabolic fluctuation of *E. gracilis*.

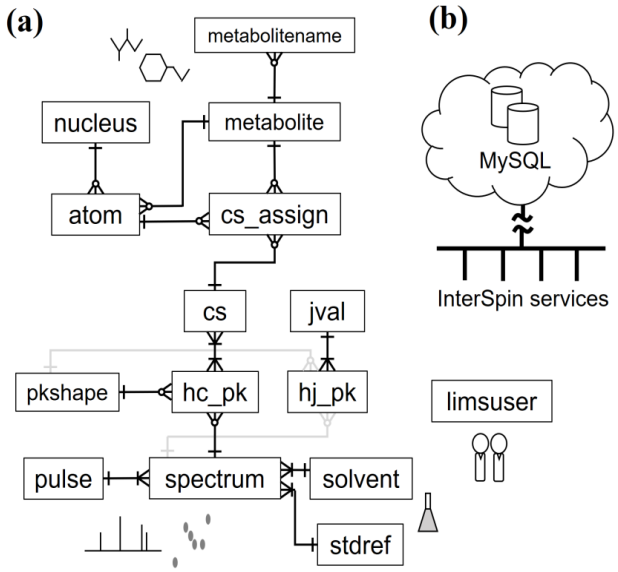


Figure S9: Schematic diagram of SpinLIMS.

(**a**) SpinLIMS is a relational database based on NMR spectra and molecular information. Thirty-four tables store a variety of information: for example, "metabolitename" is compound name, "metabolite" is compound, "atom" is atom, "nucleus" is nuclide, "cs_assign" is assignment of chemical shift, "cs" is chemical shift, "Jval" is *J* values, "hc_pk" is HSQC peaks ("h_pk" for ^{1}H-1D NMR, "c_pk" for ^{13}C-1D NMR), "pkshape" is peak linear type, "hj_pk" is 2D-*J*res peak, "spectrum" is spectrum, "pulse" is NMR pulse sequence type, "solvent" is solvent (register as "none" in case of solid), "stdref" is reference compound. (**b**) SpinLIMS services are provided to users from the MySQL server via InterSpin (http://dmar.riken.jp/interspin/).

a)

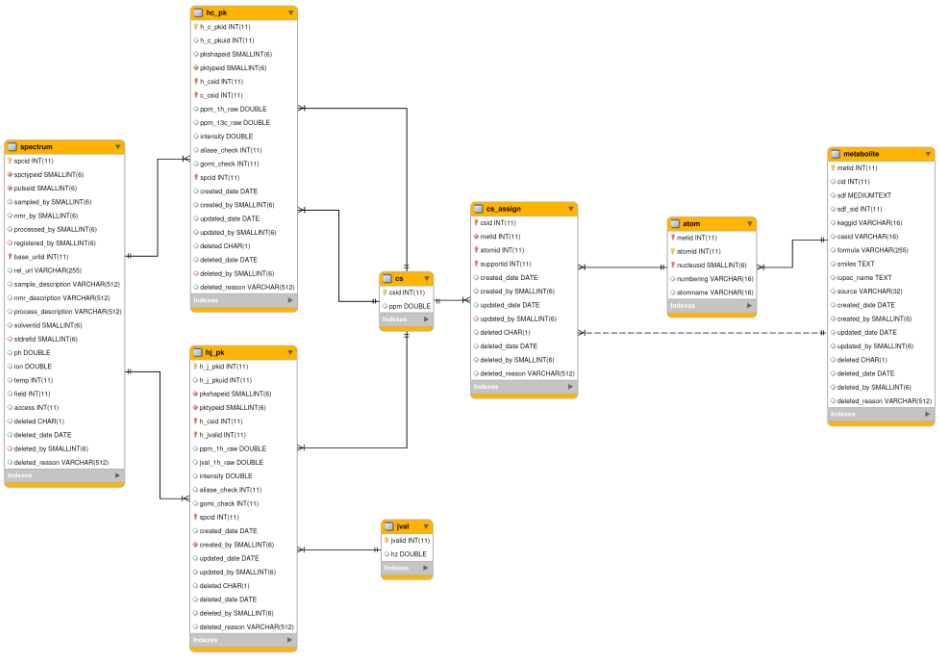


Figure S10: Diagram showing the relationship among different entities on the SpinLIMS database. (**a**) Core entities or "tables" within the relational database.

b)

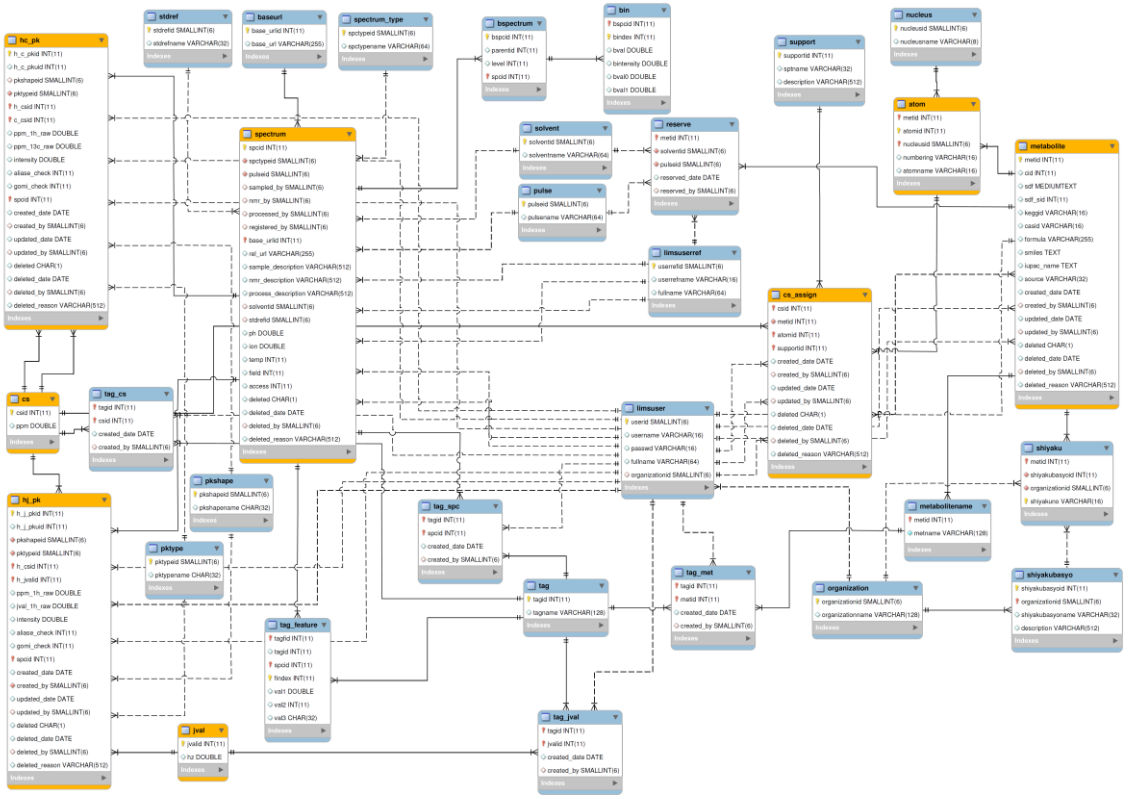
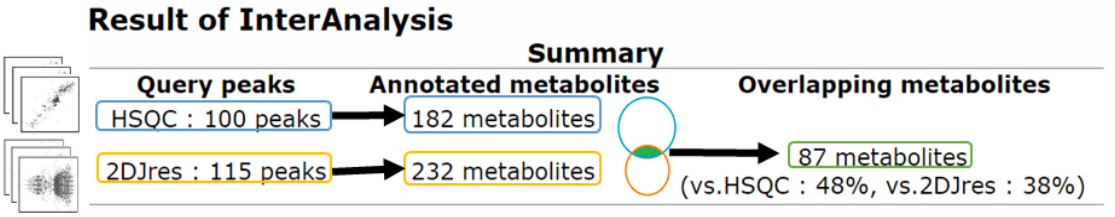


Figure S10: Diagram showing the relationship among different entities on the SpinLIMS database (continuation). (**b**) The complete relationship among all entities (tables). For a given query, each table contains the following data: "hc_pk" is HSQC peaks; "hj_pk" is 2D-*J*res peaks; "h_pk" is ^{1}H-1D NMR peaks; "c_pk" is ^{13}C-1D NMR peaks; "spectrum" is the spectrum; "spectrum_type" is the type of spectrum; "pulse" is the NMR pulse sequence; "solvent" is the solvent; "stdref" is the standard compound; "baseurl" is the base URL for raw spectrum; "pkshape" is the peak shape; "pktype" is the type of peak; "cs" is chemical shifts; "jval" is *J* values; "cs_assign" is one-atom assignment for one chemical shift; "support" is evidence of assignment; "metabolite" is metabolite; "atom" is atom; "nucleus" is nucleus; "metabolitename" is the name of metabolite; "shiyaku" is the compound number; "Shiyakubasyo" is the storage location of compound; "organization" is the organization who measured the data; "limsuser" is the user who registered the data; "reserve" is the measuring reservation; "tag" is the tag for clustering compound, metabolite, and spectrum; "tag_met" is the tag for metabolite; "tag_cs" is the tag for chemical shift; "tag_jval" is the tag for *J* value; "limsuserref" is the reference user for sampled_by, nmr_by, processed_by of spectrum table, and reserved_by of reserve table; "tag_feature" is a feature to tag related spectra, for example, for tasting data, etc.; "bin" is the bin for spectrum; "bspectrum" is the binning spectrum; and "tag_spc" is the tag for spectrum. The complete entity relationship diagram is published at http://dmar.riken.jp/interspin/



Show 10 entries

Input:HSQC peak No.	Input:HSQC 1h ppm	Input:HSQC 13c ppm	Input:2DJ peak No.	Input:2DJ 1h ppm	Input:2DJ J val hz	Annotated metabolite name
2	3.2861	38.8347	42	3.9867	7.14152999626929	L-phenylalanine
2	3.2861	38.8347	42	3.9867	7.14152999626929	L-phenylalanine
2	3.2861	38.8347	47	3.9867	1.89040499901246	L-phenylalanine
2	3.2861	38.8347	47	3.9867	1.89040499901246	L-phenylalanine
2	3.2861	38.8347	48	3.0166	1.89040499901246	L-phenylalanine
2	3.2861	38.8347	48	3.0166	1.89040499901246	L-phenylalanine
2	3.2861	38.8347	49	3.9884	0.350074999817122	L-phenylalanine
2	3.2861	38.8347	57	3.9858	0.350074999817122	L-phenylalanine
5	0.9477	23.2475	11	3.4209	-3.36071999824437	pantothenic acid
5	0.9477	23.2475	11	3.4209	-3.36071999824437	pantothenic acid

Showing 1 to 10 of 6,216 entries

First Previous 1 2 3 4 5 ... 622 Next Last

Figure S11: Result of InterAnalysis for HSQC and 2D *J*res peaks from *Acanthogobius flavimanus* body muscle extracted in deuterated potassium phosphate (KPi). The summary shows the number of query peaks, the number of assigned molecules, and the narrowed down set of molecules. The table shows some of the molecular assignment results for each query peak. For data acquired in KPi extract, the original webtools SpinAssign and SpinCouple assigned 182 and 232 molecules, respectively. By contrast, InterAnalysis assigned 87 molecules, thereby narrowing the down molecules to 48% and 38%, respectively.

Table S1: List of compounds in mixtures 1 and 2.

No.	Compound	Mixture 1 (mM)	Mixture 2 (mM)
1	Alanine	5	2
2	Leucine	5	5
3	Lysine	2	2
4	Phenylalanine	3	5
5	Proline	5	5
6	Threonine	3	5
7	Valine	2	2
8	Malate	5	10
9	Glucose	10	3
10	Sucrose	5	5

Table S2: List of 40 fish-based food mixture samples.

No.	Scientific name	Recipe	Storage time
1	*Auxis rochei rochei*	Soy marinated, Fried	□
2	*Auxis rochei rochei*	Grilled	□
3	*Branchiostegus japonicus*	Fried	□
4	*Epinephelus areolatus*	Fried	1day
5	*Epinephelus areolatus*	Fried	2day
6	*Epinephelus septemfasciatus*	Raw	□
7	*Lateolabrax japonicus*	Raw	□
8	*Lateolabrax japonicus*	Raw	□
9	*Lateolabrax japonicus*	Steamed rice wine	□
10	*Lateolabrax japonicus*	Steaming	□
11	*Nemipterus virgatus*	Fried	□
12	*Nemipterus virgatus*	Mayonnaise	□
13	*Pagrus major*	Soy marinated	1day
14	*Pagrus major*	Fried	□
15	*Pagrus major*	Grilled	□
16	*Pagrus major*	Raw	□
17	*Pagrus major*	Mayonnaise	□
18	*Paralichthys olivaceus*	Soy marinated, Fried	1day
19	*Paralichthys olivaceus*	Raw	□
20	*Paralichthys olivaceus*	Marinated kelp	1hour
21	*Paralichthys olivaceus*	Kelp, Salt, Citrus depressa	4day
22	*Paralichthys olivaceus*	Raw	□
23	*Paralichthys olivaceus*	Yuzu, Salt, Grilled	□
24	*Sardinops melanostictus*	Raw	□
25	*Scombrops gilberti*	Soy marinated	□
26	*Scombrops gilberti*	Raw	□
27	*Scombrops gilberti*	Steaming	2day
28	*Scorpaenopsis cirrhosa*	Fried	□
29	*Scorpaenopsis cirrosa*	Raw	□
30	*Sepia esculenta*	Raw	□
31	*Seriola quinqueradiata*	Grilled miso	1hour
32	*Seriola quinqueradiata*	Grilled miso	2hour
33	*Thunnus*	Raw	□
34	*Thunnus*	Raw	1day
35	*Thunnus orientalis*	Soy marinated	□
36	*Thunnus orientalis*	Grilled	□
37	*Todarodes pacificus*	Dried	□
38	*Todarodes pacificus*	Dried	1day
39	*Todarodes pacificus*	Dried	3day
40	*Zeus faber*	Raw	□

Table S3: List of standard compounds.

No.	Compound
41	Alanine
42	Histidine
43	Leucine
44	Proline
45	Glucose
46	Sucrose
47	Betaine
48	Choline
49	Creatine
50	Lactate
51	Trimethylamine *N*-oxide

Table S4: Improvement of signal-to-noise ratio by SENSI.

^{1}H chemical shift of picked peaks [ppm]	Intensity of picked peaks by SENSI	Average intensity of original spectrum	S/N after SENSI	S/N before SENSI	Sensitivity improvement rate	Peaks enhancement rate
8.35	980.1	12.3	16.0	2.4	6.7	79.9
7.27	1063.5	13.9	17.3	2.7	6.5	76.5
5.36	788.8	8.5	12.9	1.6	7.8	92.6
5.19	947.8	11.6	15.4	2.2	6.9	81.4
4.27	1162.9	15.9	19.0	3.1	6.2	73.3
4.15	1770.3	27.8	28.8	5.4	5.4	63.8
4.03	2241.1	37.0	36.5	7.1	5.1	60.6
3.91	4606.0	83.4	75.1	16.1	4.7	55.2
3.78	3318.3	58.1	54.1	11.2	4.8	57.1
3.61	2302.3	38.2	37.5	7.4	5.1	60.3
3.57	1979.2	31.9	32.3	6.1	5.3	62.1
3.41	2211.9	36.4	36.0	7.0	5.1	60.7
3.33	3409.6	59.9	55.6	11.5	4.8	56.9
3.26	13649.8	260.7	222.4	50.3	4.4	52.4
3.25	14525.5	277.9	236.7	53.6	4.4	52.3
3.20	4585.9	83.0	74.7	16.0	4.7	55.3
3.02	4268.6	76.8	69.6	14.8	4.7	55.6
2.19	1028.2	13.2	16.8	2.5	6.6	77.8
2.12	1149.3	15.6	18.7	3.0	6.2	73.7
2.02	1087.3	14.4	17.7	2.8	6.4	75.6
1.52	1602.3	24.5	26.1	4.7	5.5	65.5
1.37	4897.7	89.1	79.8	17.2	4.6	55.0
1.25	4397.1	79.3	71.7	15.3	4.7	55.5
0.99	1340.2	19.3	21.8	3.7	5.9	69.3
0.92	1357.9	19.7	22.1	3.8	5.8	69.0
Average	3226.9	56.3	52.6	10.9	5.5	65.5

Signal-to-noise (max–min, 9–10 ppm) is 5.2 before SENSI and 61.4 after SENSI.