Electronic Supplementary Material (ESI) for Analyst.

Supplementary information

Blind source separation of molecular components of the human skin in vivo: non-negative matrix factorization of Raman microspectroscopy data

Authors

B.P. Yakimov[a,b†], A.V. Venets[a†], J Schleusener[c] ,V.V. Fadeev[a], J. Lademann[c], E.A. Shirshin[*a, b, d] & M.E. Darvin[*c]

[a.]M.V. Lomonosov Moscow State University, Faculty of physics, 1-2 Leninskie Gory, Moscow, 119991, Russia.

[b.]Medical Research and Education Center, M.V. Lomonosov Moscow State University, Lomonosovsky Prospect 27/10, Moscow, Russia, 119991

[c.]Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Department of Dermatology, Venerology and Allergology, Center of Experimental and Applied Cutaneous Physiology, Charitéplatz 1, Berlin, 10117, Germany

[d.]Institute for Regenerative Medicine, Sechenov First Moscow State Medical University, Trubetskaya 8-2, Moscow, 119991, Russia

*E.S.: shirshin@lid.phys.msu.ru, M.D.: maxim.darvin@charite.de

†Authors contributed equally

*E.S.: shirshin@lid.phys.msu.ru , M. D.: maxim.darvin@charite.de

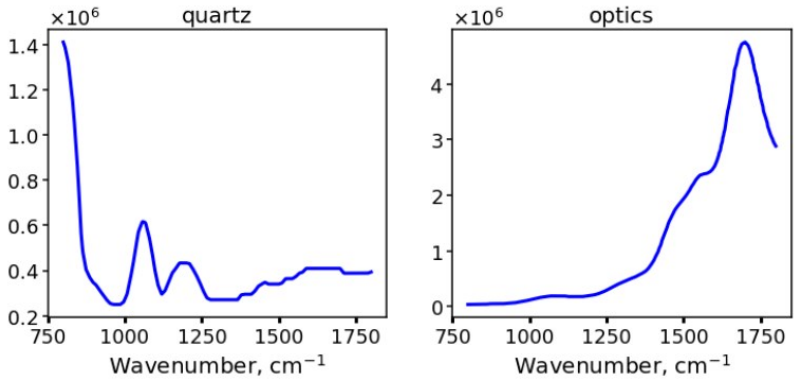

Figure S1. Spectra of residual instrument optics used in the SkinTools 2.0 software.

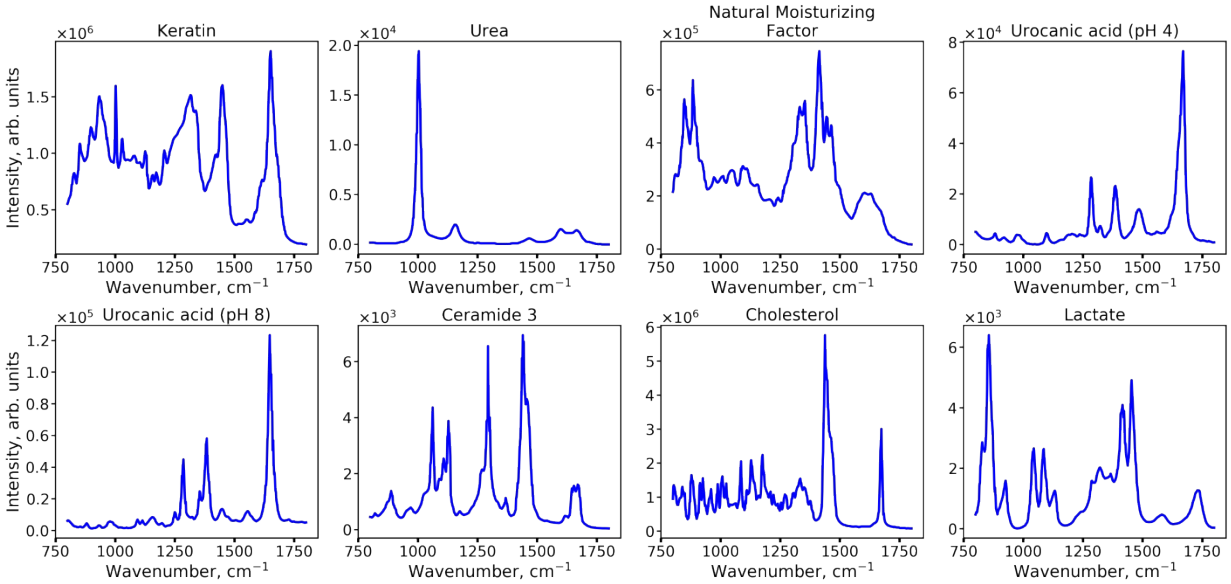

Figure S2. Spectra of individual basic components used for the Raman spectra modeling.

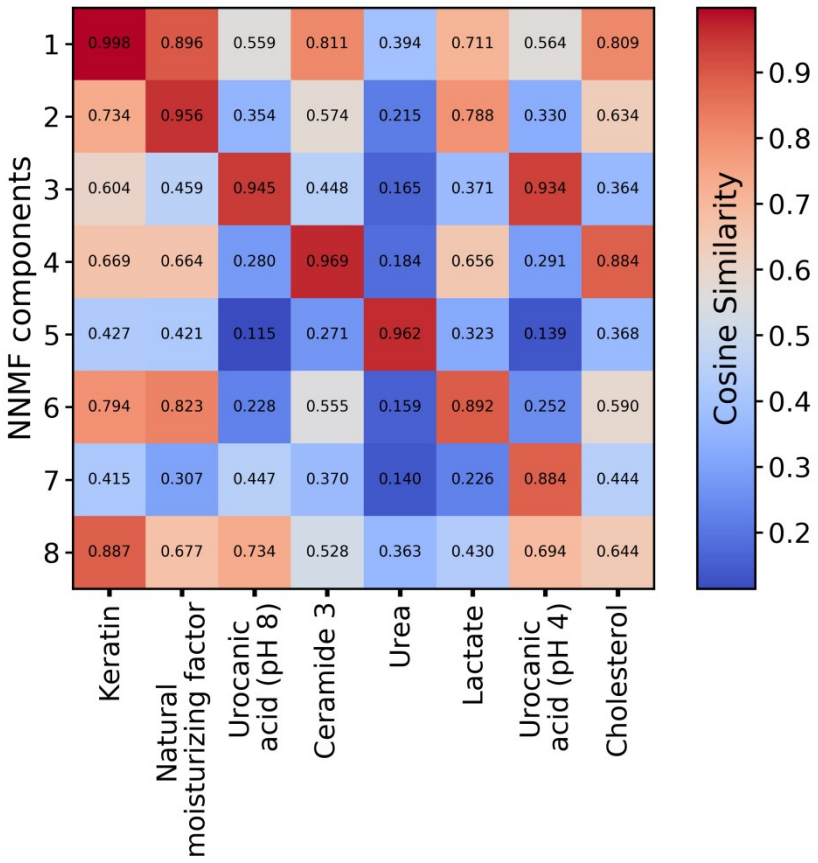

Figure S3. The matrix of pairwise cosine similarities between NNMF-restored components and input basic spectra

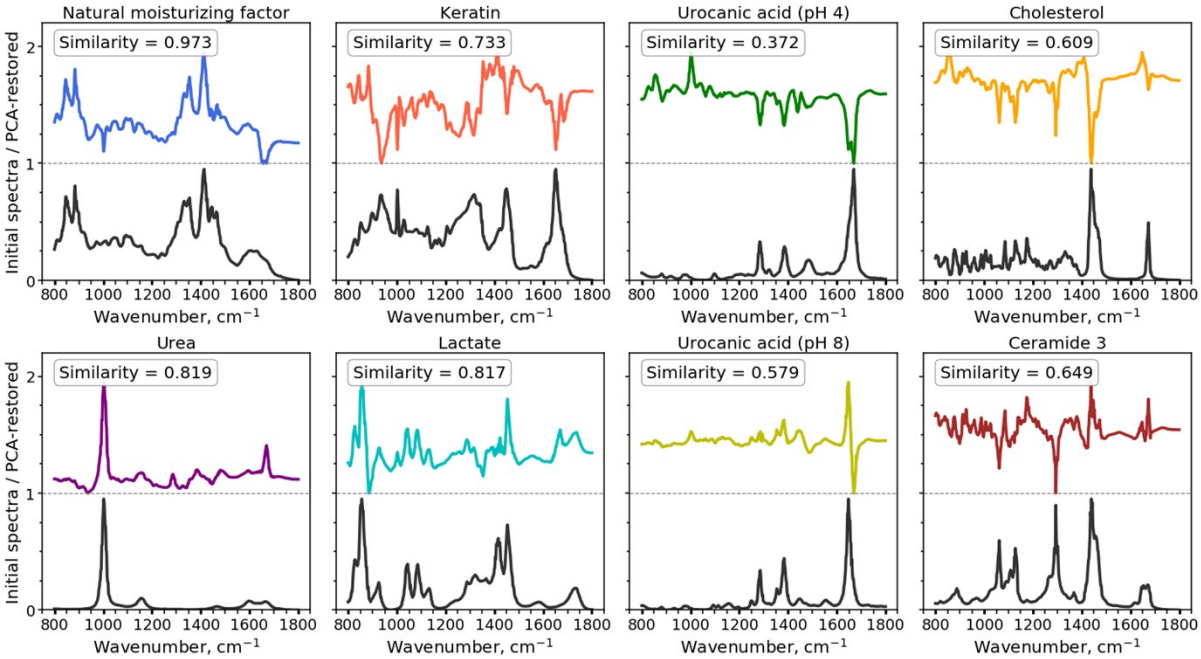

Figure S4. Spectra of the components used for the modeling of the dataset (black) and reconstructed principal components (coloured). The cosine measure of the similarity between the input and restored principal components is shown in the inset.

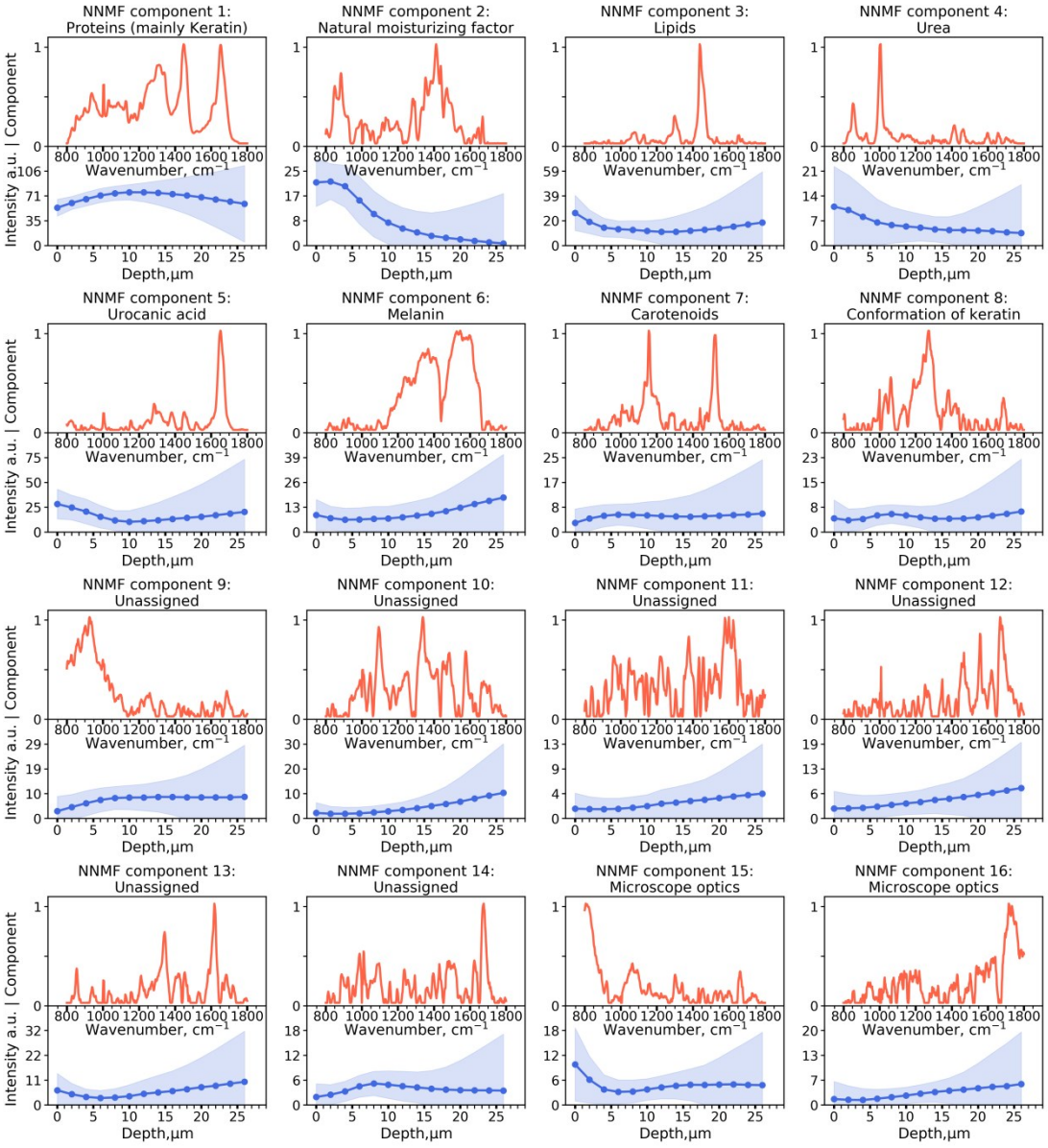

Figure S5. Spectra (top panels) and depth profiles of components (bottom panels) extracted using the NNMF algorithm applied to the experimentally observed depth-resolved skin spectra. Component profiles are averaged over measurements of 19 volunteers. The light-blue corridor represents the standard deviation of the values from the mean (solid blue line).

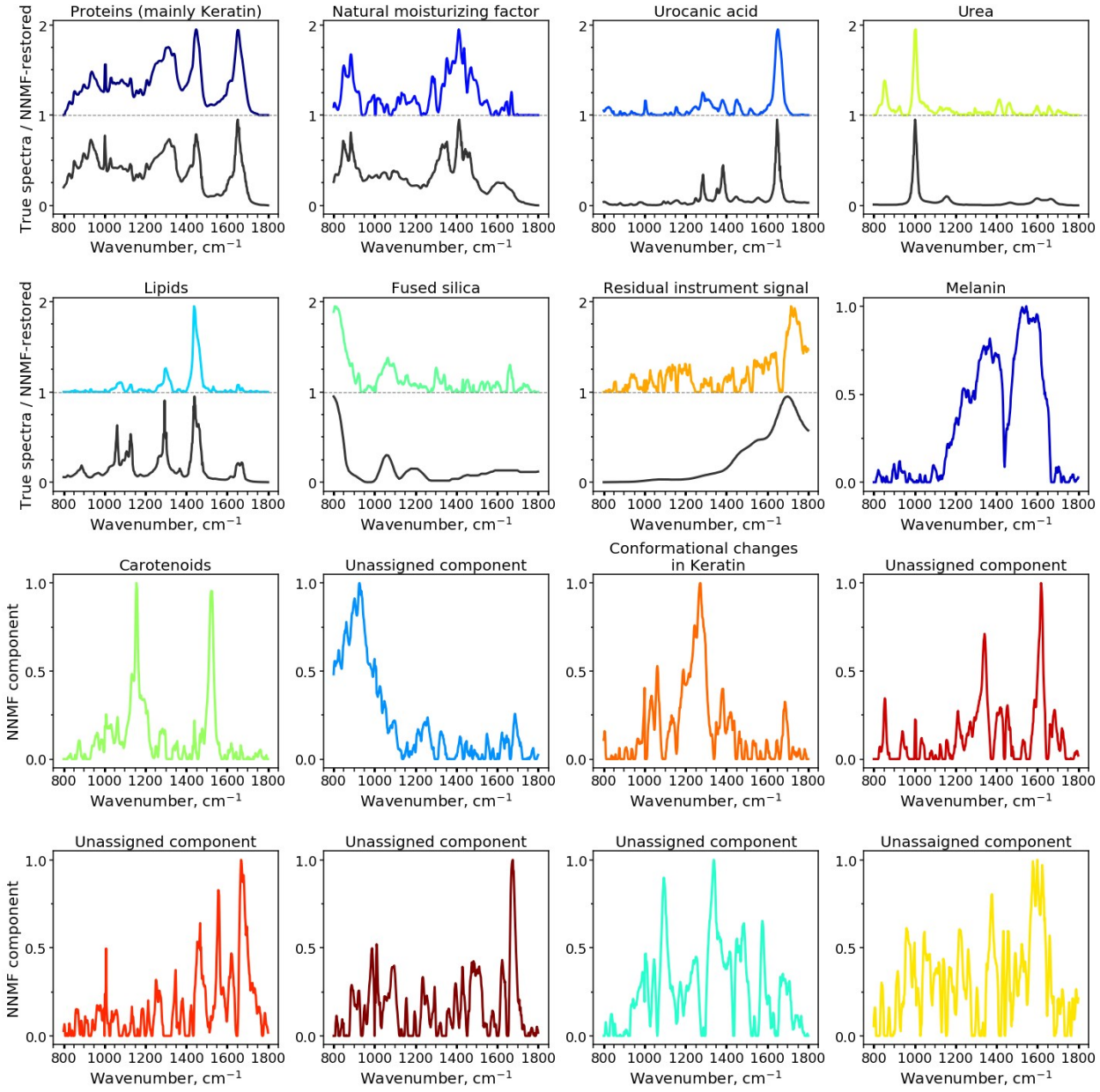

Figure S6. The spectra of all NNMF-restored components for the experimental Raman spectra of the skin.

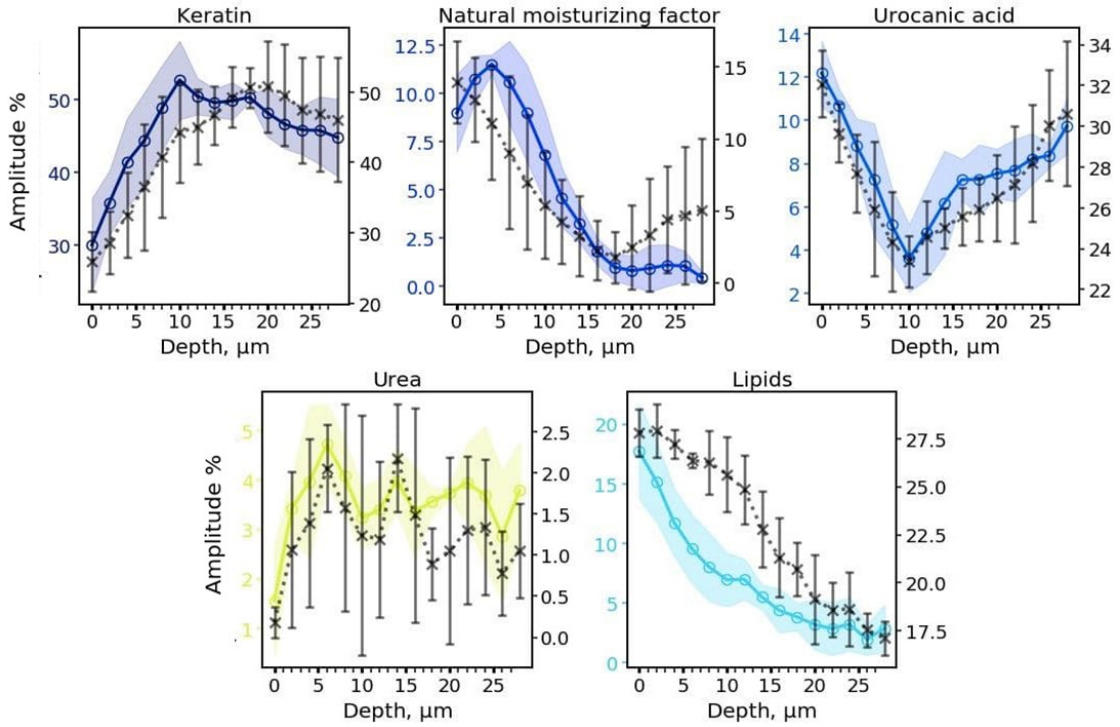

Figure S7. Average depth profiles of individual skin component amplitudes one volunteer (skin type II), obtained by NNMF (solid coloured lines) and by the decomposition with predefined components (dashed black lines) of the experimentally obtained depth-resolved Raman spectra.

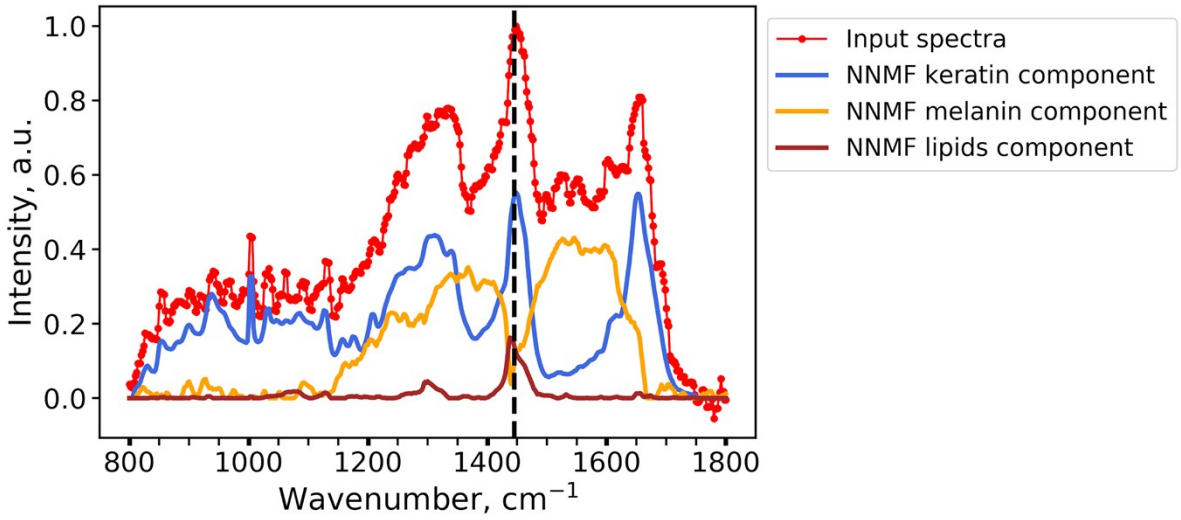

Figure S8. Normalized input Raman spectrum and spectra of the NNMF components related to keratin, melanin and lipids.

Table S1. The positions, full width at the half maximum (FWHM) and relative intensities of the bands of the input basic spectra and best-matched NNMF components obtained for modeled depth-resolved Raman spectra of the skin. For the FWHM calculation linear baseline subtraction was applied in the spectrum region specified in the Table. The bands are stated with relative intensity.

Component name	**Band centerline, cm^{-1}**		**FWHM, cm^{-1}**		**Relative intensity**		**Baseline subtraction region, cm^{-1}**	
	Input	NNMF	Input	NNMF	Component	NNMF	Component	NNMF
Keratin / NNMF -1	852		14		0.164	0.171	838–872	838–872
	934		34	36	0.27	0.246	914–994	914–984
	1002		6		0.422	0.457	994–1014	984–1016
	1126	1128	12		**0.138**	**0.18**	1114–1142	1114–1144
	1316	1294	80	86	0.465	0.463	1218–1376	1218–1378
	1450	1444	36		**0.68**	**0.773**	1376–1508	1378–1510
	1652	1652	34	36	1	1	1508–1792	1510–1792
Natural moisturizing factor / NNMF-2	846		20	14	**0.602**	**0.498**	820–872	824–862
	884		12		**0.68**	**0.752**	872–912	862–912
	1354		60	38	**0.742**	**0.65**	1224–1380	1294–1372
	1414		24		1	1	1380–1434	1372–1454
Urocanic acid (pH 8) / NNMF-3	1286		12		**0.323**	**0.44**	1262–1330	1262–1334
	1384		18	20	0.437	0.472	1330–1420	1331–1438
	1648		22	36	1	1	1592–1708	1592–1726
Ceramide 3 / NNMF-4	1062		10		**0.474**	**0.331**	990–1076	1012–1074
	1128		12		0.332	0.314	1118–1162	1118–1158
	1294		10	6	**0.732**	**0.506**	1276–1352	1278–1314
	1440		36		1	1	1390–1530	1396–1530
Urea /	1002		16		1	1	868–1100	940–1064

N	1156		38	36	**0.104**	**0.091**	1100–1226	1108–1206
Lactate (pH4) / NNMF-6	856		24		**1.866**	**1.507**	806–892	802–886
	926	930	20	44	**0.413**	**0.769**	892–974	886–1010
	1044		18	28	0.658	0.703	974–1064	1010–1062
	1086	1082	18	28	0.598	0.602	1064–1110	1062–1110
	1416		20	28	0.712	0.77	1380–1438	1380–1438
	1454		18	16	1	1	1438–1526	1438–1500
Urocanic acid (pH 4) / NNMF-7	1386	1384	22	18	**0.277**	**0.335**	1342–1424	1364–1412
	1670	1648	28	30	1	1	1578–1744	1576–1728
Cholesterol / NNMF- 8	1438	1436	24	16	**1.88**	**0.176**	1398–1582	1416–1456
	1674	1670	12	18	1	1	1582–1780	1646–1730

Table S2. The positions and relative intensities of several bands of the spectra of multiple skin constituents (keratin, natural moisturizing factor, ceramide 3 and urocanic acid) and best-matched NNMF components obtained as a result of the decomposition of the experimental depth-resolved Raman spectra of skin.

Component name	**Band centerline, cm^{-1}**	**Relative intensity**	
		Component	**NNMF**
Keratin & NNMF-1	852	0.519	0.278
	934	0.768	0.507
	1002	0.813	0.591
	1126	0.482	0.427
	1316	0.772	0.793
	1450	0.824	1
	1652	1	0.993
Natural moisturizing factor & NNMF-2	846	0.752	0.579
	884	0.85	0.708
	1354	0.743	0.664
	1414	1	1
Ceramide 3 & NNMF-3	1062	0.626	0.103
	1128	0.556	0.086
	1294	0.943	0.27
	1440	1	1
Urocanic acid & NNMF-4	1286	0.357	0.263
	1384	0.465	0.176
	1648	1	1

Supplementary Note 1. Assessment of conformational changes of keratin by Amide I Raman band decomposition

To estimate the prevailing secondary structure of keratin (*α*-helices or *β*-sheets) in the skin, Raman spectra deconvolution of the Amide I band located near 1650 cm^{-1} was performed, as suggested in [1]. First, linear baseline corrections estimated by the 1530–1540 and 1770–1780 cm^{-1} regions were applied to the raw Raman spectra. After that, the Raman signal in the 1570–1720 cm^{-1} range was fitted using 4 Gaussian functions centered at 1617±7, 1655±5, 1670±5 and 1685±5 cm^{-1} (the second number denotes the range in which position centers varied). The full width at half maximum (FWHM) for each peak varied in the 13–33, 24–36, 8–22 and 30–44 cm^{-1} ranges, respectively. The Gaussian centered near 1617 cm^{-1} corresponds to the signal from the tryptophan aromatic ring, 1655 cm^{-1} to *α*-helices, 1670 cm^{-1} to *β*-sheets, and the band at 1685 cm^{-1} corresponds to turns and random coils of skin proteins [1]. The depth profiles of the *β*-sheet/*α*-helix keratin were calculated as the ratio between the sum of intensities of bands centered near 1670 and 1685 cm^{-1} to the intensity of the Gaussian centered at 1655 cm^{-1}. SI Figure 9 demonstrates the example deconvolution of Amide I band by the described procedure.

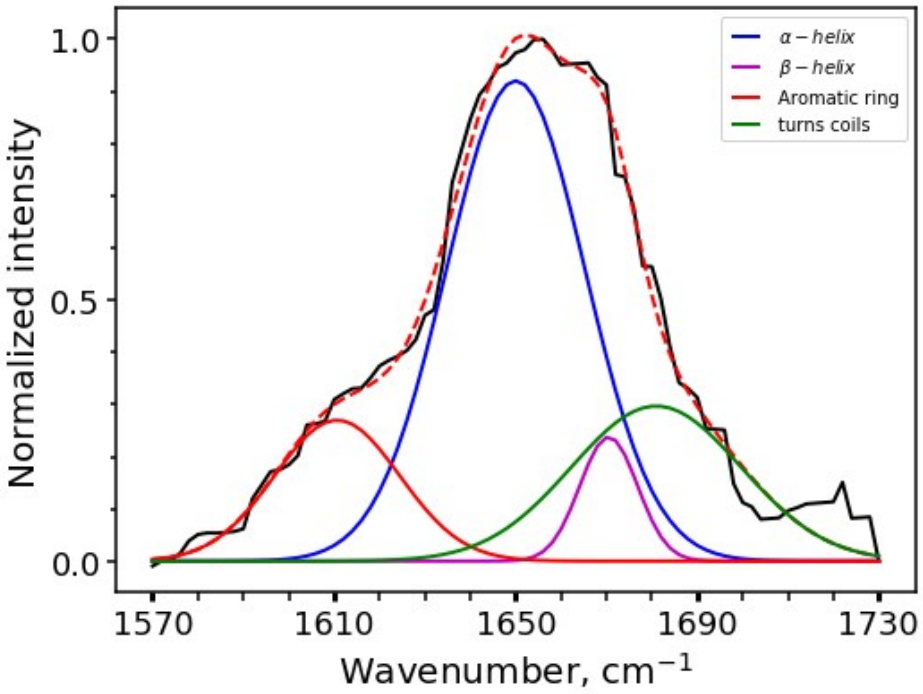

Figure S9. Example of Amide I band deconvolution by 4 Gaussians centered at 1617, 1655, 1670 and 1685 cm^{-1}. The *β*-sheet/*α*-helix ratio was estimated as sum intensities of the 1670 and 1685 cm^{-1} bands to the intensity of the 1655 cm^{-1} band.

References

[1] Choe, ChunSik, et al. "Keratin-water-NMF interaction as a three layer model in the human stratum corneum using in vivo confocal Raman microscopy." *Scientific reports* 7.1 (2017): 1-13.