Electronic Supplementary Material (ESI) for Chemical Communications

Supporting Information

for

Unmixing the NMR spectra of similar species - vive la différence

Adam A. Colbourne[a], Sebastian Meier[b], Gareth A. Morris[a] and Mathias Nilsson[a,c]*

[a]*School of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, U.K.*
[b]*Carlsberg Laboratory, Gamle Carlsberg Vej 10, 1799 Copenhagen V, Denmark*
[c]*Dept. of Food Science, University of Copenhagen, Rolighedsvej 30, DK-1958 Frederiksberg, Denmark*

**Corresponding Author: mathias.nilsson@manchester.ac.uk*

Figure S1: Processing results for the full spectral width of a progesterone (100 mM) and estradiol (100 mM) sample in DMSO-d_6 for HRDOSY, two-component SCORE and two-component OUTSCORE fitting. The original data were acquired on a Bruker Avance II+ 500 MHz spectrometer in ~ 2 h with 32 gradient increments and 32 scans at 298 K. The raw experimental data were subjected to phasing, baseline correction and reference deconvolution with a 2 Hz Lorenztian target lineshape before processing was applied. A monoexponential fit of the data was used as a starting guess for the multivariate processing. Reference spectra are included in red for comparison with calculated spectra. Slight differences in the chemical shifts of a few signals are due to the concentration difference between reference and mixture samples.

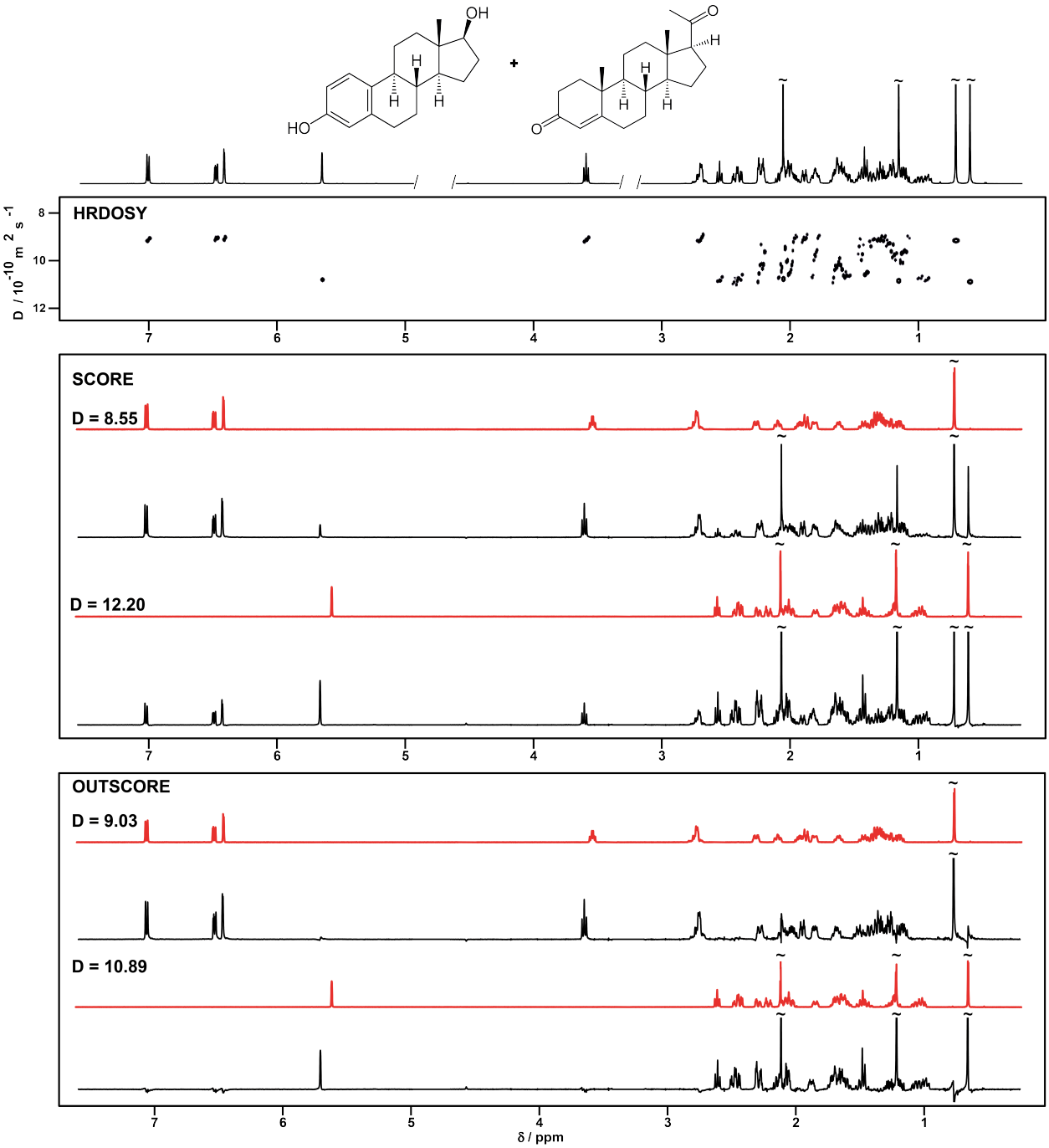

Figure S2: Processing results for the full spectral width of a concentrated ribose sample in D_2O for HRDOSY and two-component OUTSCORE fitting. The original data were acquired on a Bruker Avance II+ 500 MHz spectrometer in ~ 2 h 20 min with 20 gradient increments and 64 scans at 298 K. The raw experimental data were subjected to phasing, baseline correction and reference deconvolution with a 2 Hz Gaussian target lineshape before solvent signals were removed and processing carried out. A monoexponential fit of the data was used as a starting guess for the OUTSCORE processing.

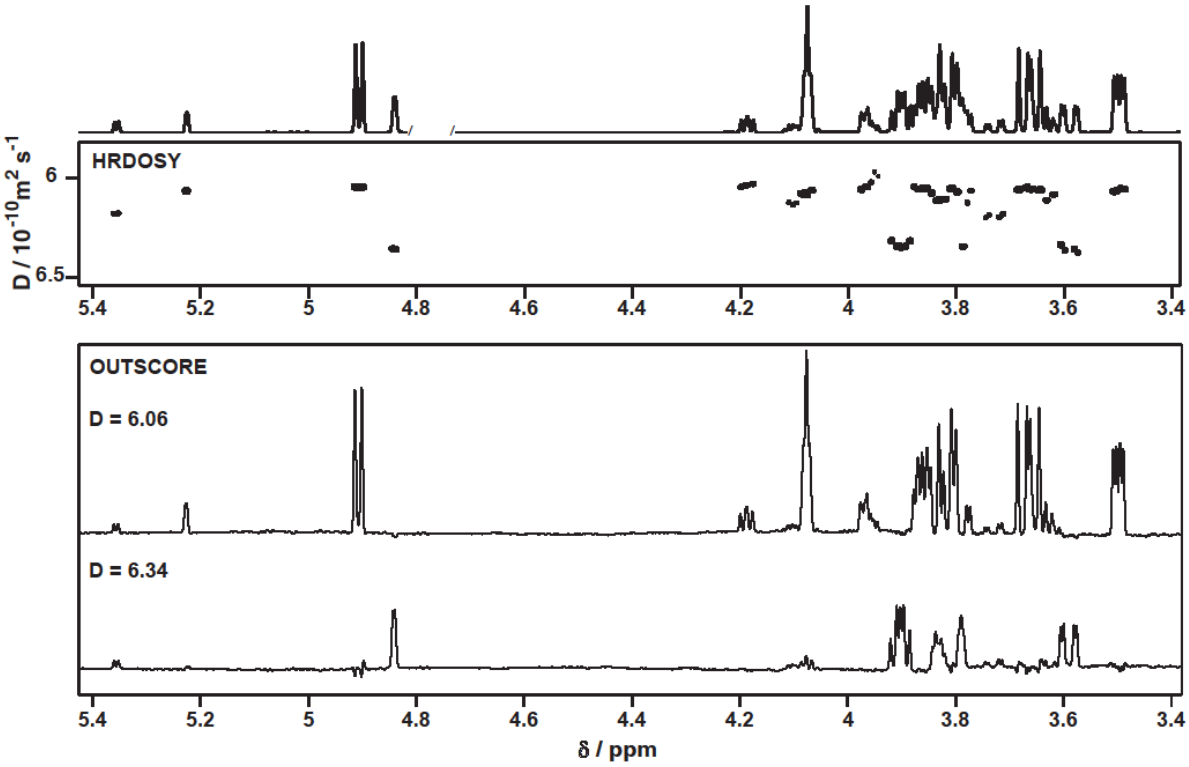

Figure S3: A comparison of OUTSCORE processing results for data from a sample of progesterone and estradiol with different signal-to-noise ratios (SNRs). The original data were collected on a Bruker Avance II+ 500 MHz spectrometer in18.5 min using 16 gradient increments and 16 scans, giving a SNR ~30000:1. Two other datasets at 1000:1 and 100:1 SNR (relative to the tallest signal) were generated by adding Gaussian noise to the original data. Solvent signals were excised and OUTSCORE processing was applied to each dataset and a monoexponential fit of the data was used as a starting guess for the OUTSCORE processing. Even at 100:1 SNR, the two component spectra are separated, though only the methyl signals are visible above the noise in the resultant spectra. The increased noise level in the component spectra is caused by uncertainty in the matrix decomposition; the additional noise is anti-correlated between the two component spectra such that if they are summed, the noise cancels and the original spectrum is recovered.

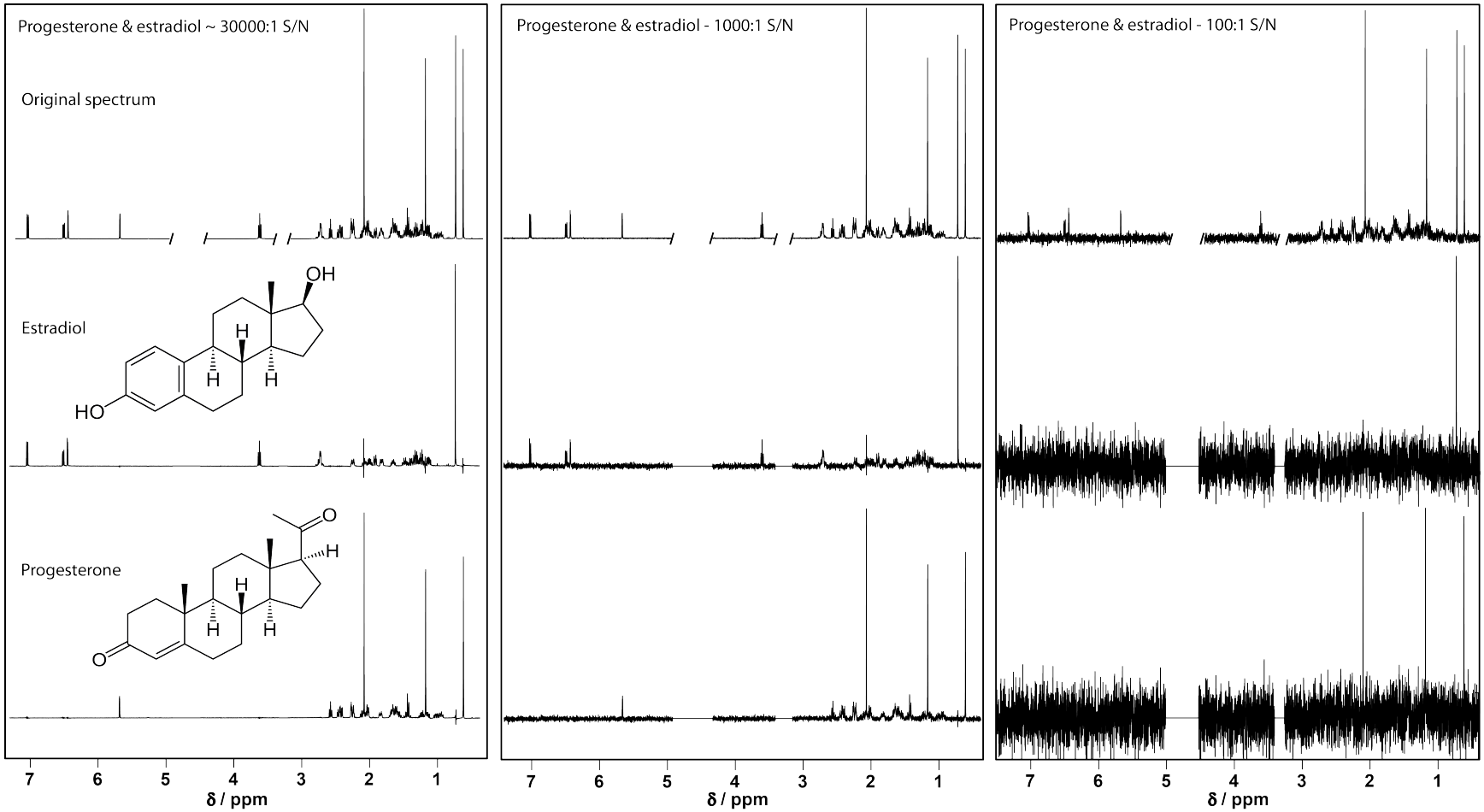

Figure S4: A comparison of SCORE processing results for data from a sample of progesterone and estradiol with different signal-to-noise ratios (SNRs). The original data were collected on a Bruker Avance II+ 500 MHz spectrometer in18.5 min using 16 gradient increments and 16 scans, giving a SNR ~30000:1. Two other datasets at 1000:1 and 100:1 SNR (relative to the tallest signal) were generated by adding Gaussian noise to the original data. Solvent signals were excised and SCORE processing applied to each dataset; a monoexponential fit of the data was used as a starting guess for the SCORE processing. The SCORE processing fails to resolve the component spectra, with cross-talk observable in all calculated component spectra.

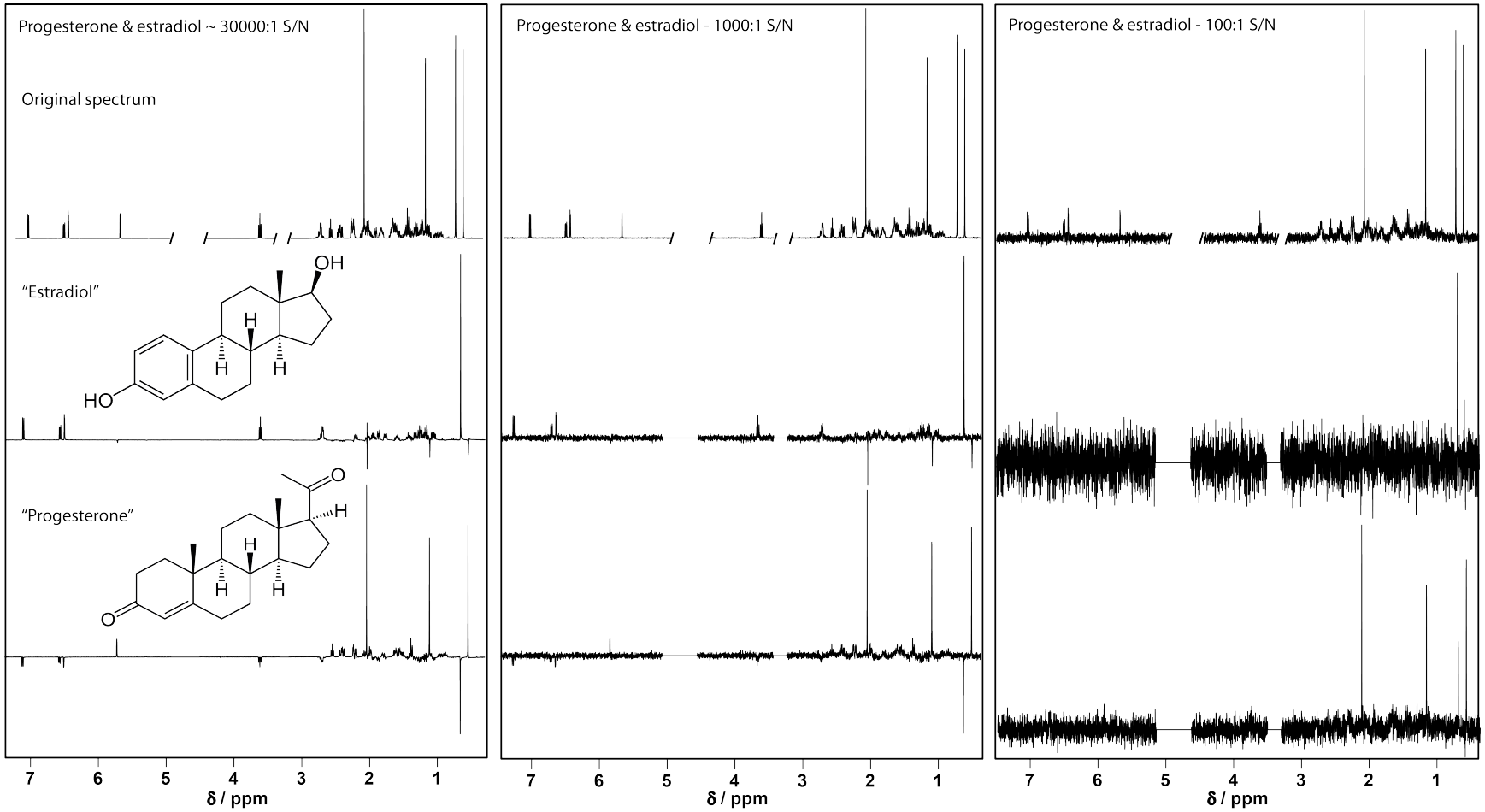

Figure S5: OUTSCORE processing results for a sample containing a mixture of quinine, geraniol and camphene. The original data were acquired on a Varian INOVA 400 MHz spectrometer with 30 gradient increments and 256 transients in 17 h with no temperature regulation; reference deconvolution was applied with a 3 Hz Lorentzian target lineshape. A monoexponential fit of the data was used as a starting guess for the OUTSCORE processing. Very little cross-talk is observable, although the phase instability of the water peak does cause a slight dispersion-mode presence in the geraniol and camphene spectra.

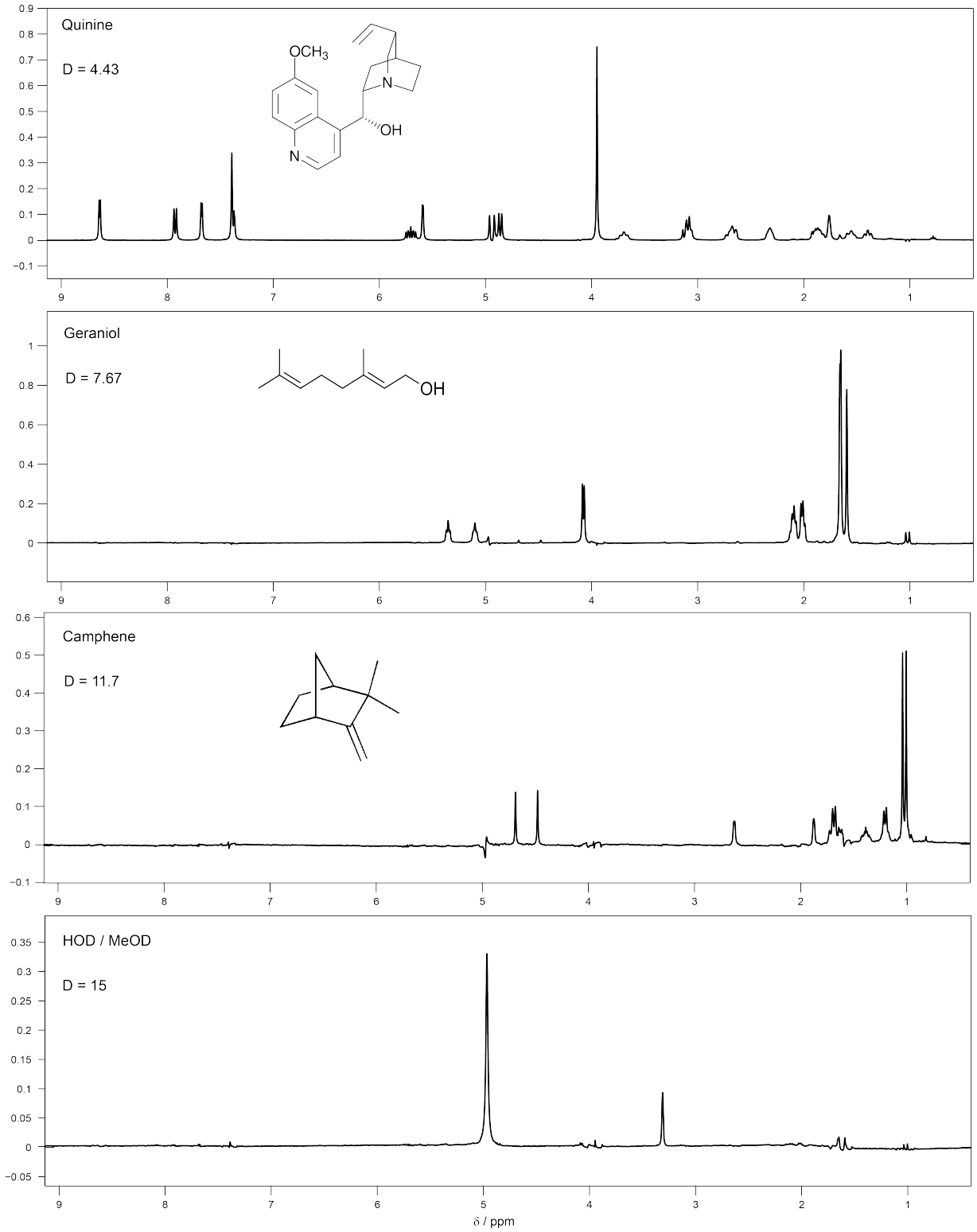

Figure S6: SCORE processing results for the same data as Figure S5. A monoexponential fit of the data was used as a starting guess for the SCORE processing. There is minor but significant cross-talk to be observed throughout, especially between the quinine and the other spectra, as well as between the HOD / MeOD spectrum and the other calculated components; this results from the uncertainty in the diffusion coefficients found for these components.