Supplementary Figure S1

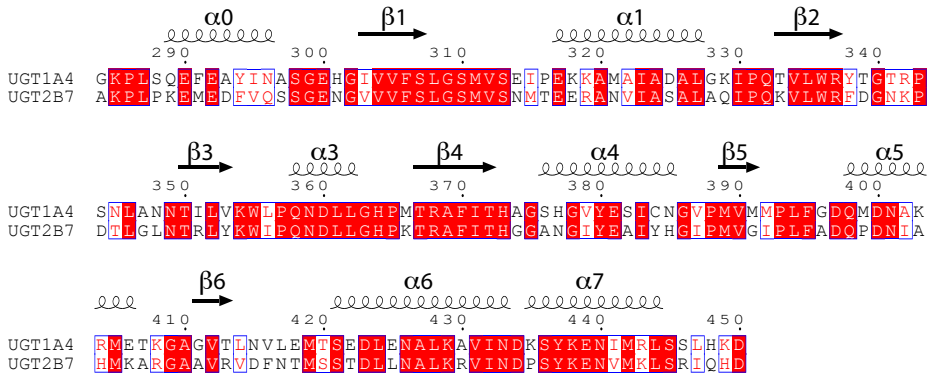

Fig. S1. Alignment of UGT1A4 and UGT2B7 C-domains. Secondary structural elements derived from the UGT2B7 X-ray structure are shown above the sequence (arrows indicate β -strands and loops, α-helices) [23, 24]. The chemical shift index (CSI) for UGT1A-C and these elements were in good agreement [31] providing a strong basis for modelling UGT1A-C on the UGT2B7 structure. The alignment was done using Clustal W2 program (http://www.ebi.ac.uk). Conserved residues are in red and boxed with a white background, while identical residues are boxed on a red background.

Supplementary Figure S2

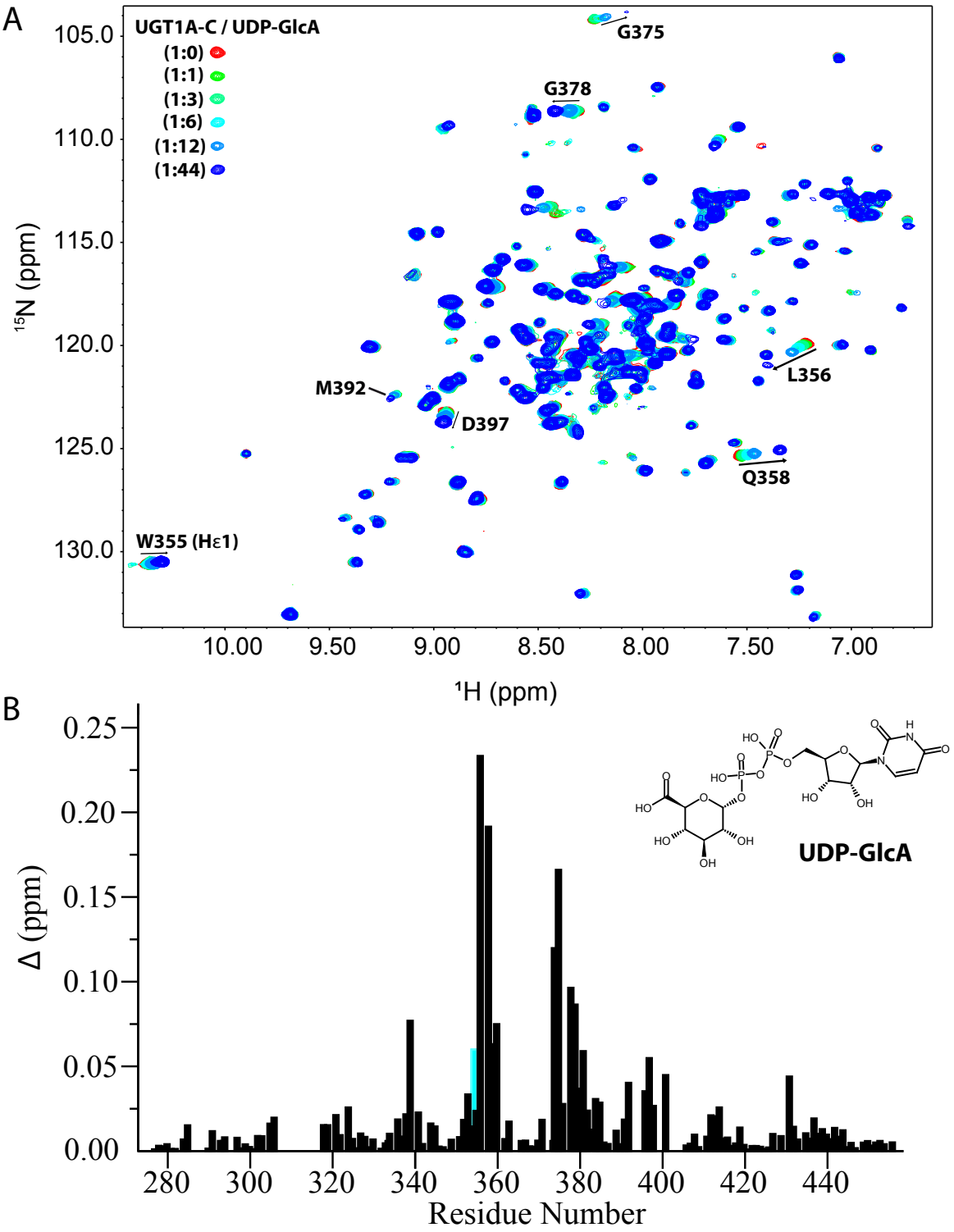

Fig. S2. UDP-GlcA-binding site on UGT1A-C from NMR studies. (A) Chemical shift perturbation (CSP) observed upon UDP-GlcA addition in the ^{1}H-^{15}N HSQC using ^{15}N-labelled UGT1A-C. Residues located at the UDP-GlcA binding site are highlighted. (B) Quantitation of CSPs derived from ^{1}H-^{15}N HSQC experiments using ^{15}N-labelled UGT1A-C in the presence of UDP-GlcA. The blue bar indicates the CSP for the indole nitrogen from W355. The chemical structure of UDP-GlcA is also shown

Supplementary Figure S3

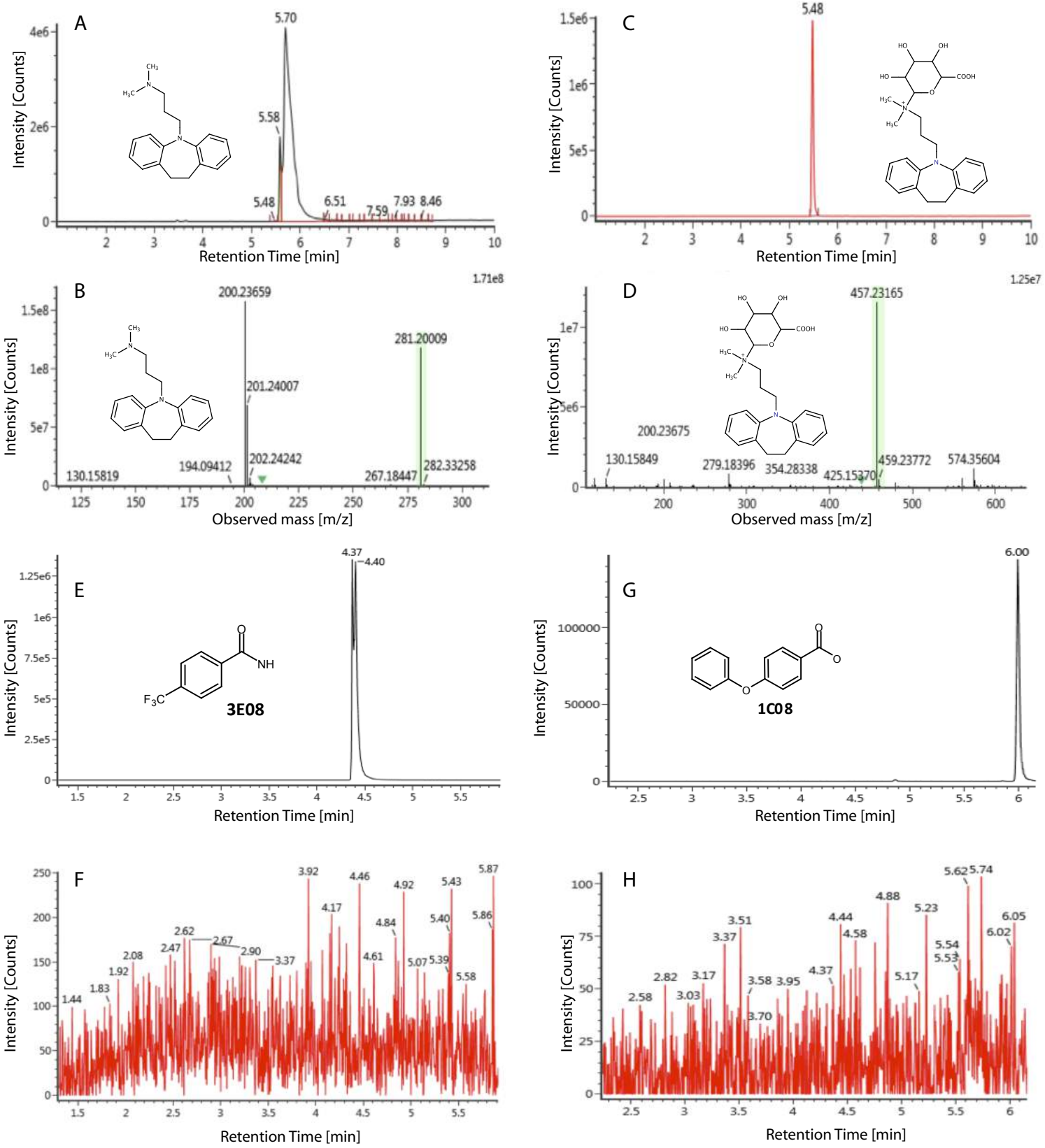

Fig. S3. Summary of mass spectrometry experiments searching for glucuronides of selected compounds after incubation with UGT1A4 or UGT1A1 microsomes. Top panel; UPLC extracted ion chromatogram (EIC) traces for (A) imipramine at the m/z of 281.2 for the parent compound and (C) for the glucuronide at the mass of the m/z of the parent ion +176.0321 Da. B) ESI mass spectra showing the parent compound and D) the glucronide at m/z 457.231565 Da E) EIC for compound 3E08 at m/z 190.0479 (+/-5 ppm) for the parent compound and (F) its glucuronide at 366.08 Da (+/- 5ppm) after incubation in UGT1A4 microsomes. G) EIC for compound 1C08 at m/z of the parent compound (215.0708 Da) and H) for the glucuronide (391.1029 +/- 5 ppm) after incubation in UGT1A1 microsomes. No glucuronides were detected for 3E08 and 1C08.

Supplementary Figure S4

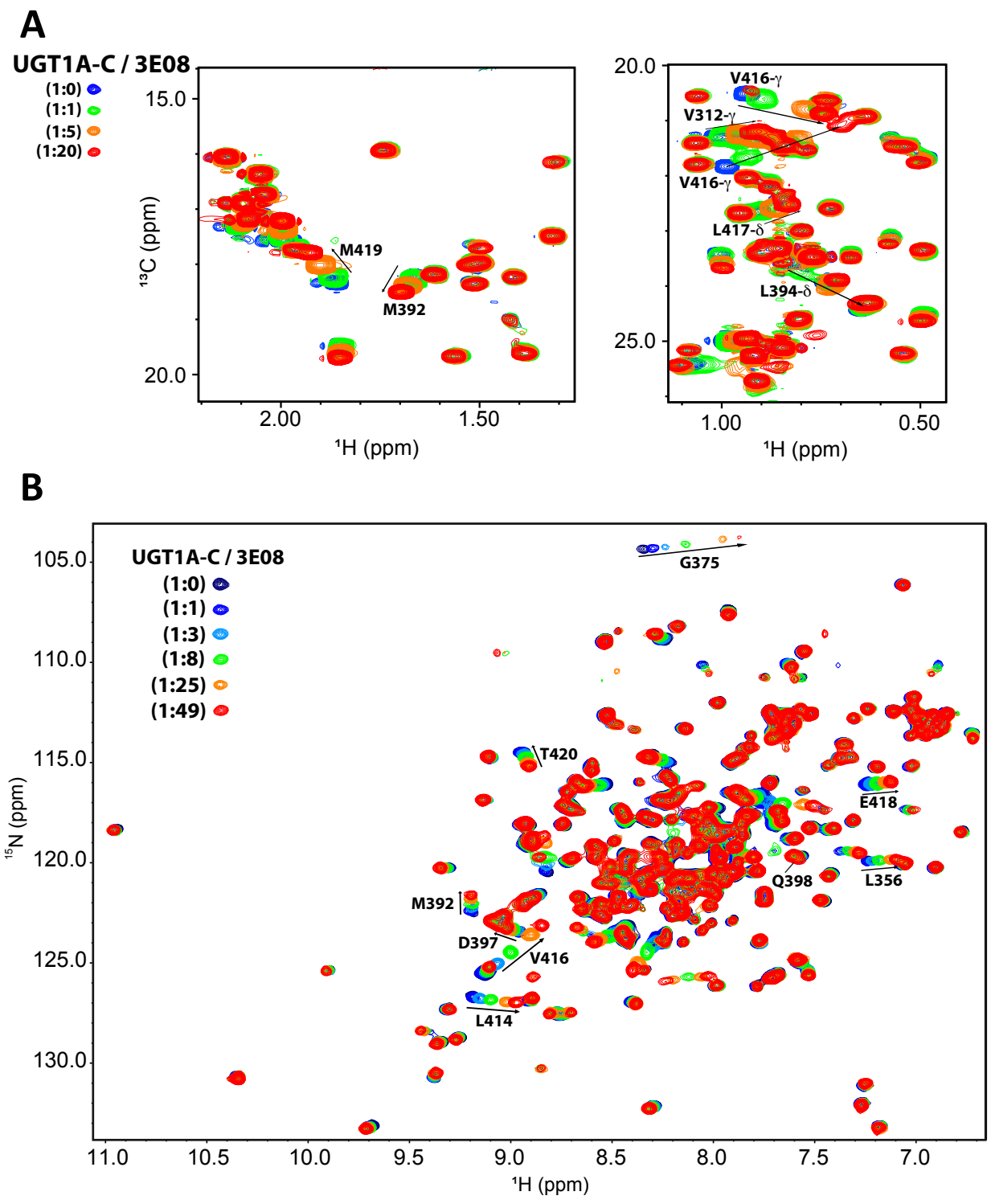

Fig. S4. 3E08 binding to UGT1A-C. (A) Regions of ^{1}H-^{13}C HSQC showing CSPs upon 3E08 addition. Relevant residues are highlighted. (B) ^{1}H-^{15}N HSQC showing CSPs at different ratios of 3E08.

Supplementary Figure S5

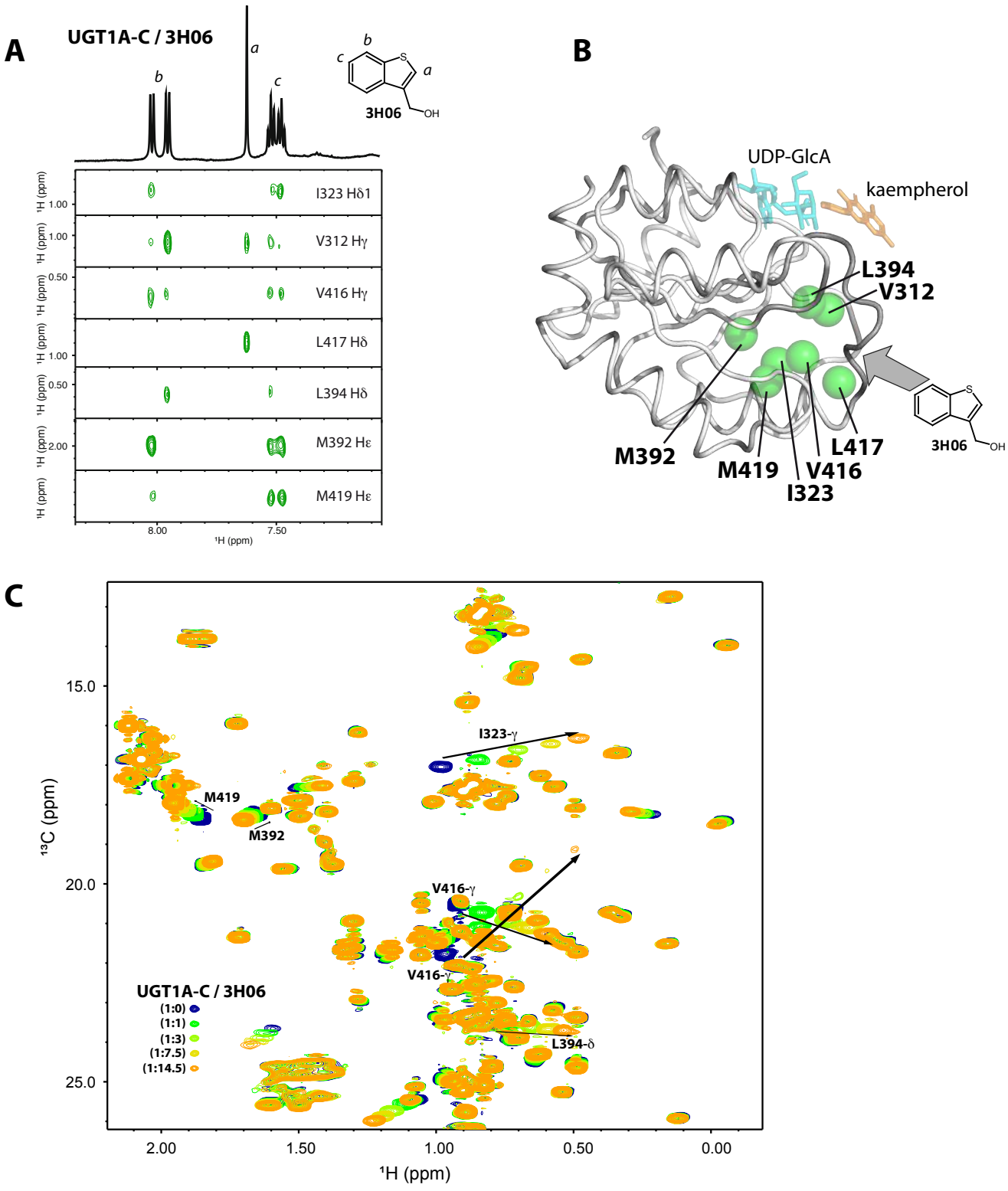

Fig. S5. NMR analysis of 3H06 binding to UGT1A-C. (A) ^{1}H-^{1}H strips from the ^{13}C-edited, ^{13}C/^{15}N filtered NOESY showing intermolecular NOEs between methyl resonances of ^{13}C/^{15}N UGT1A-C and 3H06. 1D ^{1}H NMR spectrum of 3H06 and its chemical structure are also shown. (B) NOEs are mapped onto the UGT1A-C model structure, as shown by green balls. The active site is marked by UDP-GlcA (cyan) and kaempherol (orange) to provide orientation, but these compounds were not present in the mapping experiments. (C) ^{1}H-^{13}C HSQC of UGT1A-C showing CSP upon addition of 3H06. Perturbed shifts agree well with the observed NOEs.

Supplementary Figure S6

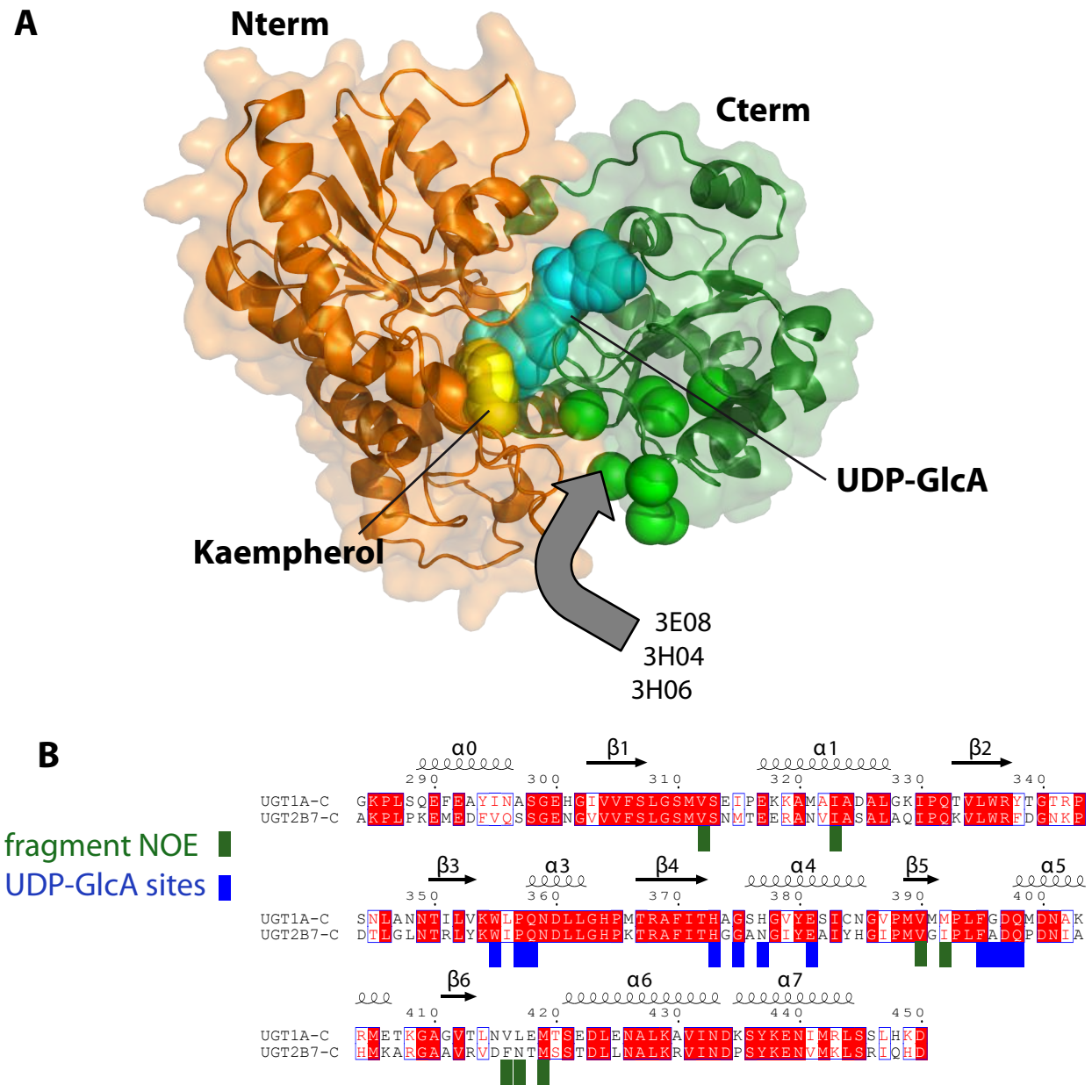

Fig. S6. (A) Model based on Fig. 1B which shows position of the fragment binding site as well as kaempherol and UDP-GlcA. Colouring is the same as in Fig 1B. Green balls represent methyl groups which make NOEs to E308 as a marker of the fragment binding site for 3E08, 3H04 and 3H06. This model shows that the fragments would make contacts with the N-domain and thus provide a basis for the observed allostery. (B) Summary of the NMR defined UDP-GlcA and fragment sites are represented on the sequence of UGT1A4, for comparisonthe UGT2B7 aligmnent is shown. Blue bars represent binding sites for UGP-GlcA from CSP and the UGT1A-C model. Green bars represent NOE's from the fragments (3E08 and 3H06) to UGT1A-C identified.