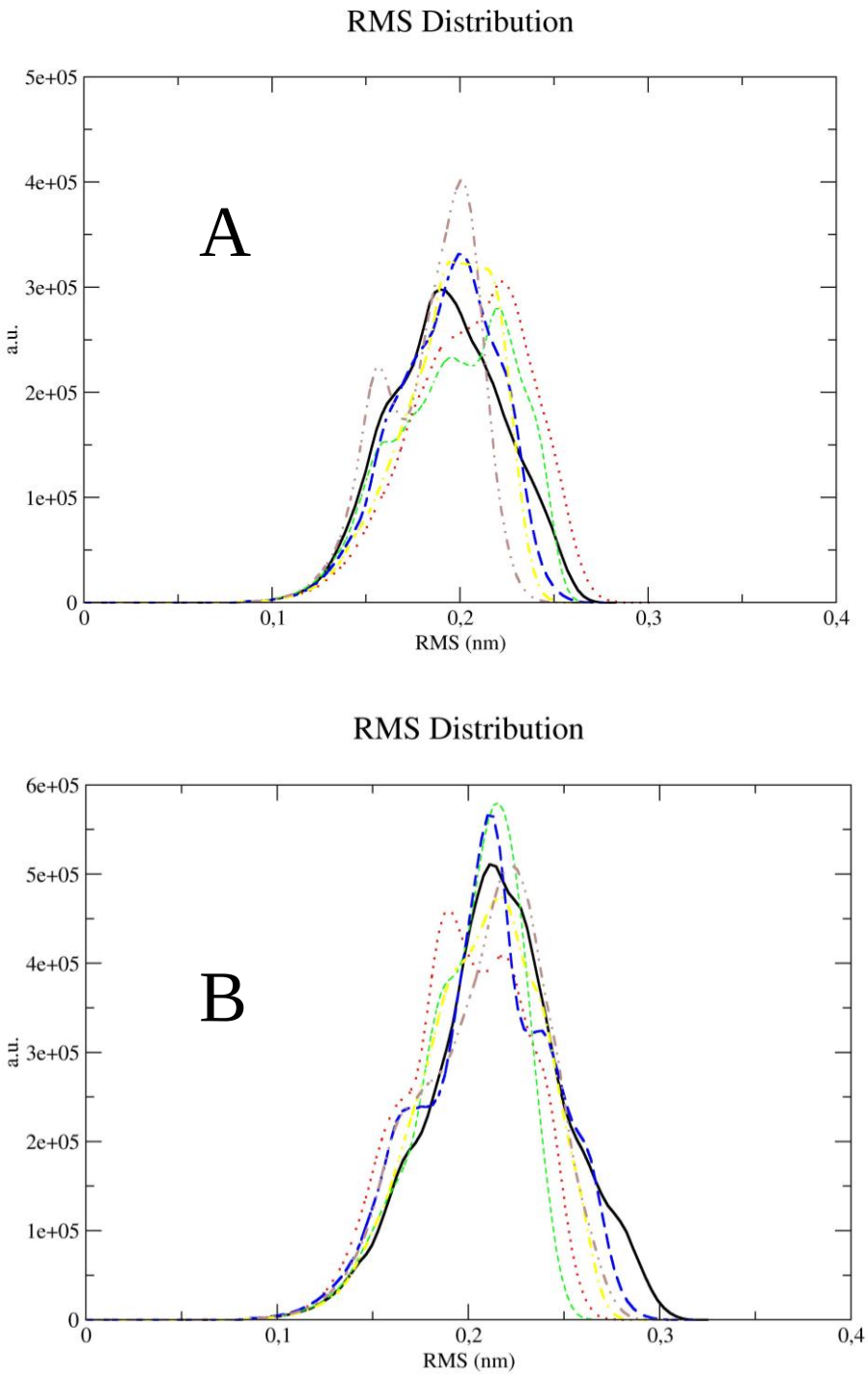
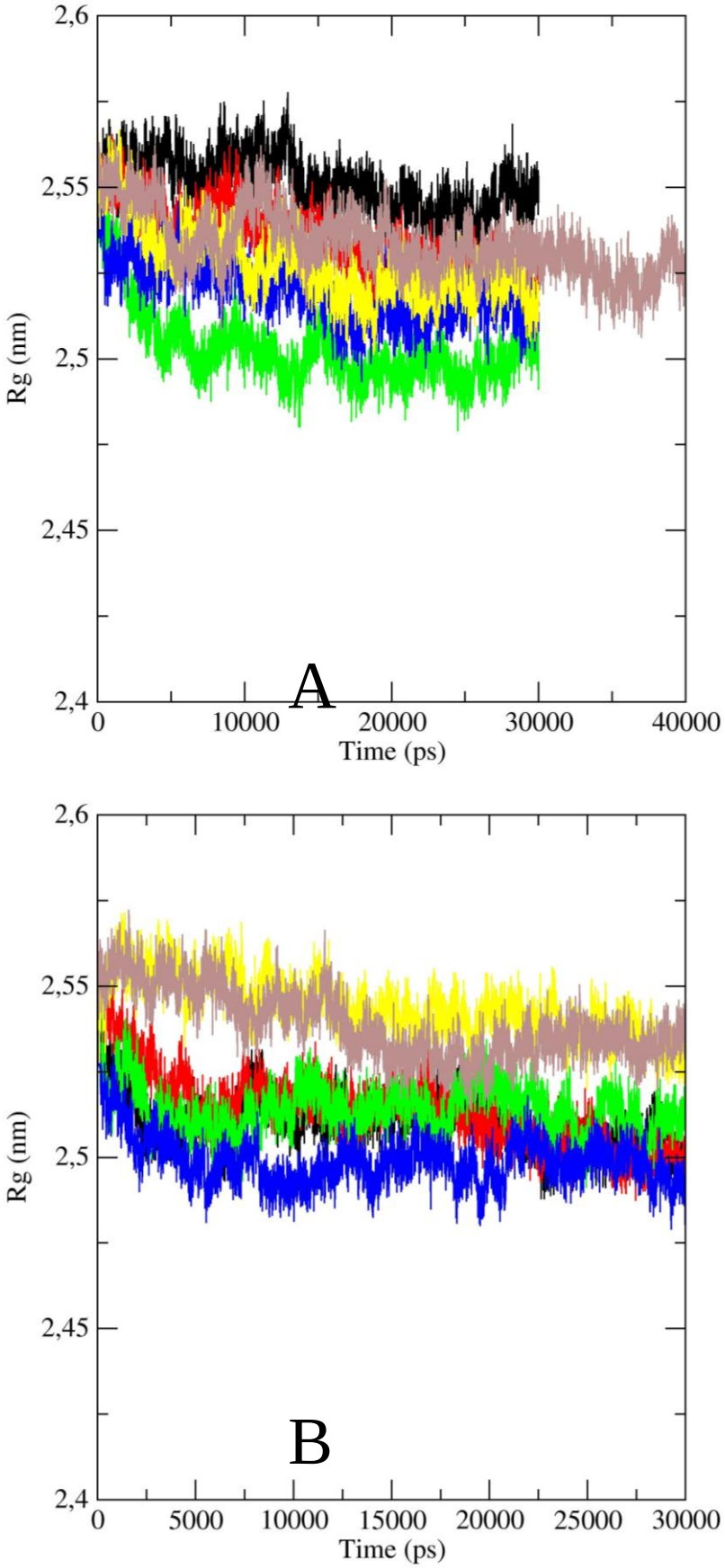


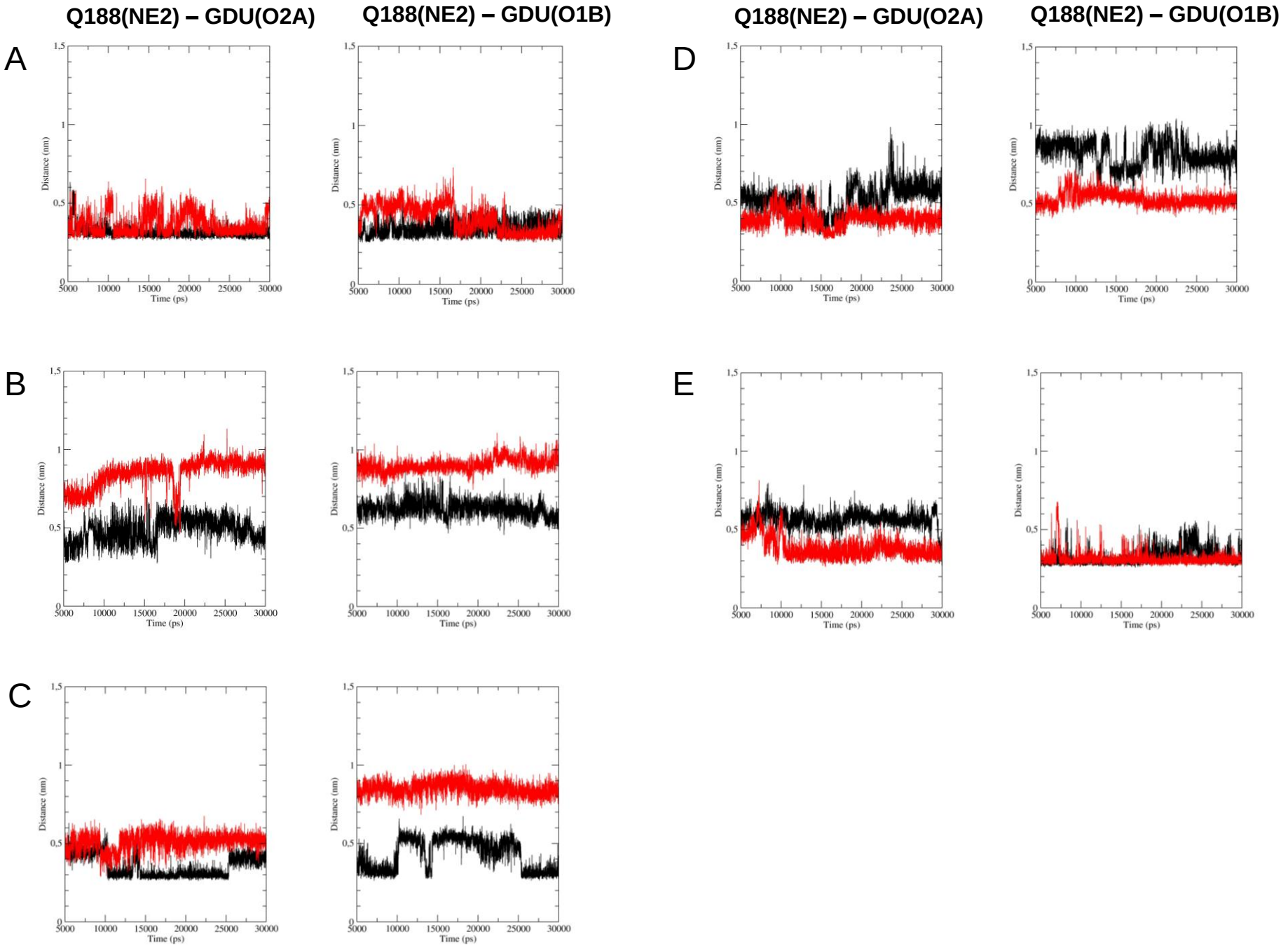

**Supp. Figure S1.** RMS distribution of cluster analysis. In A), RMS distribution was shown for "apo" forms; in B), RMS distribution was reported for "holo" forms, i.e. with UDP-Gal bound. a.u.: arbitrary units. Color code: wild-type GALT is in black, P185H in red, H132Q in green, I278N in blue, V168L in yellow, Q188R in light brown.




**Supp. Figure S2.** Radius of gyration (Rg) of "apo" (A) and "holo" (B) forms of GALT. The color code is the same as in Supp. Figure S1.

**Supp. Figure S3 (next page).** Variation of the distances between UDP-Gal and Gln 188. The variation of the distances between atom O2A and O1B of UDP-Gal and atom NE2 of residue Gln 188 was followed during the trajectory after the equilibration phase ($\geq$ 5 ns) for wild-type GALT (A), mutant P185H (B), mutant H132Q (C), mutant I278N (D) and mutant V168L (E). Black line refers to the distances measured in chain A, red line to the distances measured in chain B.

**Supp. Figure S3**



**Supp. Table S1. 14 *GALT* gene variations in this study**

| # | Location | Type | Nucleotide Change | Protein Change | Comments |
|---|---|---|---|---|---|
| 1 | Exon II | Missense | c.100T>A | p.Y34N | A father had p.Y34N and enzyme activity=16.4 µmol/h/gHb. Mother was a carrier for the Duarte 2 haplotype. Child had enzyme and phenotype consistent with D/G. |
| 2 | Exon V | Missense | c.396C>A | p.H132Q | Patient had p.H132Q and the Duarte 2 haplotype and enzyme activity of 3.9 µmol/h/gHb. |
| 3 | Exon V | Missense | c.502G>T | p.V168L | Pilipino female patient presenting with bilirubinemia was homozygous for pV168L. GALT enzyme activity was 0 µmol/h/gHb. |
| 4 | Exon VI | Missense | c.509T>C | p.I170T | Non-identical twins were heterozygous for p.I170T and p.Q188R and had absent enzyme activity. |
| 5 | Exon VI | Missense | c.554C>A (Calderon et al. 2007) | p.P185H | Patient was heterozygous for p.P185H and p.R201C. Enzyme activity was 1.8 µmol/h/gHb. |
| 6 | Exon VII | Missense | c.601C>T (Calderon et al. 2007) | p.R201C | Patient was heterozygous for p.P185H and p.R201C. Enzyme activity was 1.8 µmol/h/gHb. |
| 7 | Exon VII | Missense | c.658G>A (Calderon et al. 2007) | p.E220K | Patient was heterozygous for p.E220K and enzyme activity=11.2 umol/h/gHb. |
| 8 | Exon VII | Missense | c.667C>A (Calderon et al. 2007) | p.R223S | Patient was heterozygous for p.R223S and p.Q188R. Enzyme activity was 1.7µmol/h/gHb. |

| # | Location | Type | Nucleotide Change | Protein Change | Comments |
|---|---|---|---|---|---|
| 9 | Exon VII | Missense | c.680T>C | p.L227P | Patient was heterozygous for p.L227P and the Duarte 2 haplotype. Enzyme activity was 3.4 µmol/h/gHb. |
| 10 | Exon VIII | Missense | c.776G>A | p.R259Q | Patient was heterozygous for p.R259Q and p.S135L. Enzyme activity was 0 µmol/h/gHb. |
| 11 | Exon IX | Missense | c.833T>A (Calderon et al. 2007) | p.I278N | Two unrelated patients were heterozygous for p.I278N. Enzyme activity was 9.8 and 11.4 µmol/h/gHb respectively. |
| 12 | Exon IX | Missense | c.865C>T (Calderon et al. 2007) | p.L289F | Patient was heterozygous for p.L289F and enzyme activity=11 µmol/h/gHb. |
| 13 | Exon IX | Missense | c.872A>T | p.E291V | Hispanic patient was heterozygous for p.E291V and p.Q188R. Enzyme activity consistent with D/G. |
| 14 | Exon X | Missense | c.980T>C | p.L327P | Patient was heterozygous for p.L327P and p.Q188R. Enzyme activity was 0 µmol/h/gHb. |

**Note:** DNA variation numbering system used here is based on the reference cDNA sequence from GenBank (BC015045.2 GI:34783207). Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the ATG translation initiation codon in the reference sequence.

**Supp. Table S2. Analysis of the effects of selected variations on GALT secondary structures content**

| **Analysis on apo forms** | | | | | | |
|---|---|---|---|---|---|---|
| **Secondary structure[a]** | **wt GALT** | **H132Q** | **V168L** | **P185H** | **I278N** | **Q188R** |
| Helices | 26.4 | 24.8 | 25.4 | 26.2 | 25.0 | 24.5 |
| Strands | 18.7 | 17.6 | 17.7 | 18.6 | 17.1 | 16.2 |
| Other structures | 29.5 | 31.1 | 30.4 | 30.0 | 31.9 | 33.8 |
| Random coil | 25.4 | 26.5 | 26.5 | 25.2 | 26.0 | 25.5 |
| | | | | | | |
| **Analysis on holo forms** | | | | | | |
| **Secondary structure[a]** | **wt GALT** | **H132Q** | **V168L** | **P185H** | **I278N** | **Q188R** |
| Helices | 24.2 | 24.6 | 23.4 | 24.4 | 24.7 | 24.8 |
| Strands | 20.8 | 17.3 | 18.1 | 17.6 | 17.3 | 21.3 |
| Other structures | 28.3 | 32.2 | 31.8 | 32.8 | 30.8 | 28.5 |
| Random coil | 26.7 | 25.9 | 26.7 | 25.2 | 27.2 | 25.4 |

The table shows the % of residues included in different secondary structures, calculated on the representative structures obtained by the cluster analysis.

[a]: secondary structures are classified as following: helices: DSSP codes H+I+G; strands: DSSP code E; other structures: DSSP codes B+S+T; random coil: residues with unassigned DSSP code.

**Supp. Table S3. Analysis of the effects of selected variations on GALT H-bond content**

| **Analysis on apo enzymes** | | | | | | |
|---|---|---|---|---|---|---|
| | **wt GALT** | **H132Q** | **V168L** | **P185H** | **I278N** | **Q188R** |
| Donor groups | 1040 | 1040 | 1040 | 1044 | 1042 | 1044 |
| Acceptor groups | 2082 | 2082 | 2082 | 2086 | 2086 | 2084 |
| Average number of H-bonds per frame (into the whole protein)[a] | 548 | 540 | 539 | 542 | 549 | 544 |
| Average number of H-bonds per frame (chainA vs chainB)[b] | 29 | 30 | 29 | 33 | 34 | 28 |
| **Analysis on holo enzymes** | | | | | | |
| | **wt GALT** | **H132Q** | **V168L** | **P185H** | **I278N** | **Q188R** |
| Donor groups | 1054 | 1054 | 1054 | 1058 | 1056 | 1058 |
| Acceptor groups | 2120 | 2120 | 2120 | 2124 | 2124 | 2122 |
| Average number of H-bonds per frame (into the whole protein)[a] | 542 | 534 | 522 | 538 | 543 | 529 |
| Average number of H-bonds per frame (chainA vs chainB)[b] | 32 | 31 | 29 | 29 | 31 | 28 |
| Average number of H-bonds per frame (protein vs ligand)[c] | 21 | 27 | 23 | 23 | 26 | 19 |

[a]: the average number of H-bonds that are present into the whole structure per timeframe
[b]: the average number of H-bonds that are present between chain A and chain B per timeframe
[c]: the average number of H-bonds that are present between protein and ligand per timeframe