

Preparation for Variant Analysis (4A & 4B)

Preliminaries for 4A:

Protein of interest: GTP Cyclohydrolase I, AA sequence in FASTA format:

```
>EAW80648.1 GTP cyclohydrolase 1 (dopa-responsive dystonia)[Homo sapiens]
MEKGPVVRAPAEKPRGARCSNGFPERDPPRPGPSRPAEKPPRPEAKSAQPADGWKGERPRSEEDNELNLPN
LAAAYSSILSSLGENPQRQGLLKTPWRAASAMQFFTKGYQETISDVLNDAIFDEDHDEMIVKIDIDMFSM
CEHHLVPPFVGKVHIGYLPNKQVLGLSKLARIVEIYSRRLQVQERLTKQIAVAITEALRPAGVGVVVEATH
MCMVMRGVQKMNSKTVTSTMLGVFREDPKTREEFLTLIRS
```

List of variants to be analyzed (note, only 6 required):

E65Q, A73V, A98V, S100L, I154M, P199A, M211I, M215T

Output of ProtParam analysis (numbered protein sequence):

```

      10           20           30           40           50           60
MEKGPVVRAPAEKPRGARCSNGFPERDPPRPGPSRPAEKPPRPEAKSAQPADGWKGERPRSEEDNELNLPN
      70           80           90           100          110          120
EEDNELNLPN LAAAYSSILSSLGENPQRQGLLKTPWRAASAMQFFTKGYQETISDVLNDAIFDEDHDEMIVKIDIDMFSM
      130          140          150          160          170          180
IFDEDHDEMIVKIDIDMFSMCEHHLVPPFVGKVHIGYLPNKQVLGLSKLARIVEIYSRRLQVQERLTKQIAVAITEALRPAGVGVVVEATH
      190          200          210          220          230          240
VQERLTKQIAVAITEALRPAGVGVVVEATHMCMVMRGVQKMNSKTVTSTMLGVFREDPKTREEFLTLIRS
      250
REEFLTLIRS
```

Completed main item (only the marked version must be submitted):

4A. Your POI AA sequence with variant locations marked.

GTPCH1 - AA variant locations marked

```

      10           20           30           40           50           60
MEKGPVVRAPAEKPRGARCSNGFPERDPPRPGPSRPAEKPPRPEAKSAQPADGWKGERPRSEEDNELNLPN
      70           80           90           100          110          120
EEDNEELNLPN LAAAYSSILSSLGENPQRQGLLKTPWRAASAMQFFTKGYQETISDVLNDAIFDEDHDEMIVKIDIDMFSM
      130          140          150          160          170          180
IFDEDHDEMIVKIDIDMFSMCEHHLVPPFVGKVHIGYLPNKQVLGLSKLARIVEIYSRRLQVQERLTKQIAVAITEALRPAGVGVVVEATH
      190          200          210          220          230          240
VQERLTKQIAVAITEALRPAGVGVVVEATHMCMVMRGVQKMNSKTVTSTMLGVFREDPKTREEFLTLIRS
      250
REEFLTLIRS
```

List of variants to be analyzed:

E65Q, A73V, A98V, S100L, I154M, P199A, M211I, M215T

Completed main item (examples):

4C. Variant Analysis based on AA type change only (a few examples)

(Note: these are not actual GTPCH1 variants, but other hypothetical changes. You may put all of your variants into some kind of table, or simple lists like below. The descriptions refer to the side chain, of course.)

1. L86G - Leu (L) to Gly (G)
 - i. Change from a large aliphatic, hydrophobic AA to a tiny, slightly hydrophobic amino acid.
 - ii. BLOSUM score: -4
 - iii. Summary: Radical change

2. F141E - Phe (F) to Glu (E)
 - i. Change from an aromatic, hydrophobic AA to a polar, acidic amino acid.
 - ii. BLOSUM score: -3
 - iii. Summary: Radical change

2. S165A - Serine (S) to Alanine (A)
 - i. Change from a tiny, polar AA to a tiny, hydrophobic amino acid.
 - ii. BLOSUM score: 1
 - iii. Summary: minimal to moderate change

4D. Variant Analysis including MSA conservation information

Summary variant analysis

E65Q - likely benign
A73V - likely benign
A98V - likely pathogenic
S100L - likely benign
I154M - uncertain, benign?
P199A - likely pathogenic
M211I - likely pathogenic
M215T - likely pathogenic

Additional explanation of your reasoning for each designation of your variants is appropriate.

Example: A98V - likely pathogenic. Although A & V are similar hydrophobic AAs, they're not highly similar (e.g., A is small, V larger; BLOSUM score A->V is 0); but most importantly, based on the MSA, **A at this position is 100% conserved.**