

Q1

1.)

- a. Protein - Cellular tumor antigen p53
Gene - TP53
- b. Organism - Human
Tax. ID - 9606 [NCBI]
- c. Reviewed - Annotation score:(5/5) - Experimental evidence at protein level
 - Reviewed (UniProtKB/Swiss-Prot)
- d.
 - Mutated or turned off in 60% of cancers
 - TP53 defects are found in Barrett metaplasia
 - Esophageal cancer
 - Li-Fraumeni syndrome (LFS)
- e. Ubiquitous. Isoforms are expressed in a wide range of normal tissues but in a tissue-dependent manner.
 - Isoform 2 is expressed in most normal tissues but is not detected in brain, lung, prostate, muscle, fetal brain, spinal cord and fetal liver.
 - Isoform 3 is expressed in most normal tissues but is not detected in lung, spleen, testis, fetal brain, spinal cord and fetal liver.
 - Isoform 7 is expressed in most normal tissues but is not detected in prostate, uterus, skeletal muscle and breast.
 - Isoform 8 is detected only in colon, bone marrow, testis, fetal brain and intestine.
 - Isoform 9 is expressed in most normal tissues but is not detected in brain, heart, lung, fetal liver, salivary gland, breast or intestine.
- f. Integrated into UniProtKB/Swiss-Prot: August 13, 1987
Last sequence update: November 24, 2009
Last modified: March 15, 2017

2.) Acts as a tumor suppressor in many tumor types; induces growth arrest or apoptosis depending on the physiological circumstances and cell type. Involved in cell cycle regulation as a trans-activator that acts to negatively regulate cell division by controlling a set of genes required for this process. One of the activated genes is an inhibitor of cyclin-dependent kinases. Apoptosis induction seems to be mediated either by stimulation of BAX and FAS antigen expression, or by repression of Bcl-2 expression. In cooperation with mitochondrial PPIF is involved in activating oxidative stress-induced necrosis; the function is largely independent of transcription. Induces the transcription of long intergenic non-coding RNA p21 (lincRNA-p21) and lincRNA-Mkln1. LincRNA-p21 participates in TP53-dependent transcriptional repression leading to apoptosis and seem to have to effect on cell-cycle regulation. Implicated in Notch signaling cross-over. Prevents CDK7 kinase activity when associated to CAK complex in response to DNA damage, thus stopping cell cycle progression. Isoform 2 enhances the transactivation activity of isoform 1 from some but not all TP53-inducible promoters. Isoform 4 suppresses transactivation activity and impairs growth suppression mediated by isoform 1.

Isoform 7 inhibits isoform 1-mediated apoptosis. Regulates the circadian clock by repressing CLOCK-ARNTL/BMAL1-mediated transcriptional activation of PER2

a. 574 experimentally validated GO annotations

Cellular component 52

Molecular function 435

Biological process 87

b. Homo Sapien total protein count : 99287

Homo Sapien total annotation count : 1059892

Homo Sapien proteins w/ experimental notation: 19952, 20.1% of all proteins

Protein counts obtained from <http://www.ebi.ac.uk/QuickGO/GProtein?ac=P04637>

Where I removed the protein name filter and applied a taxonomy filter for human to get the human protein and annotation count. Then I filtered by manual experimental and by sub-ontology to get the other counts.

| | |
|-----------------------------|---------|
| Number of annotations | 1060643 |
| Number of distinct proteins | 99287 |

Currently selected additional filtering parameters:

| | |
|---------|---------------|
| tax | 9606 |
| protein | (No proteins) |

Cellular Component annotations 386902

Molecular Function annotations 322812

Biological Process annotations 350929

Cellular Component proteins 76439

Molecular Function proteins 70595

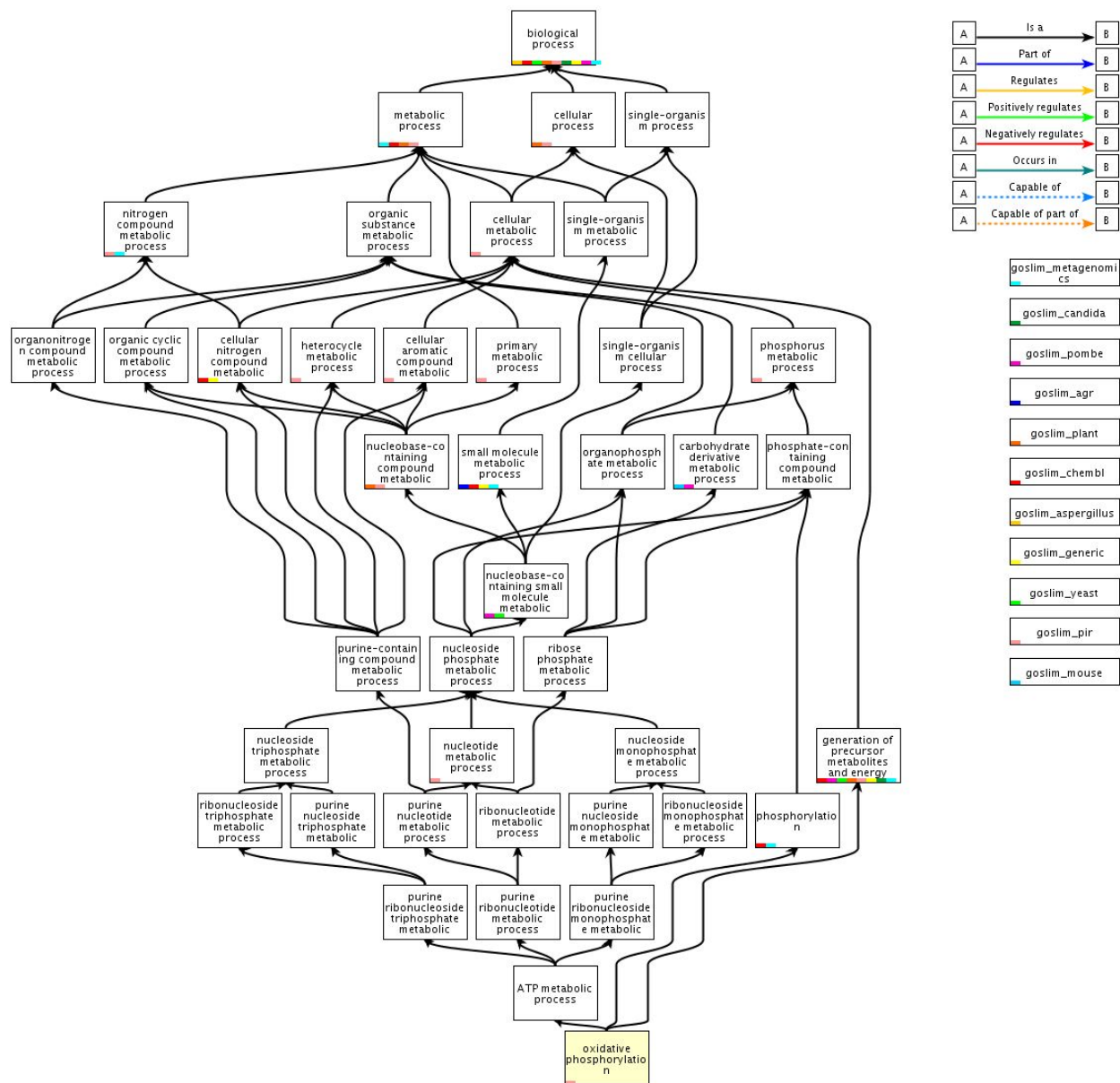
Biological Process proteins 77183

Cellular Component proteins w/ experimental notation 16101, 21.0%

Molecular Function proteins w/ experimental notation 12042, 17.0%

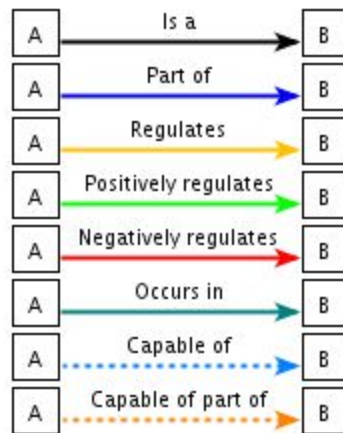
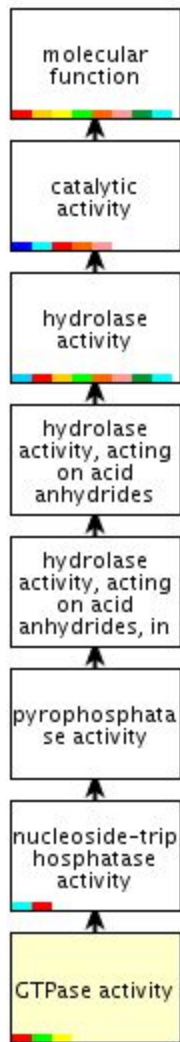
Biological Process proteins w/ experimental notation 8066, 10.4%

c. Biological Process hierarchy picture GO: 0006119:



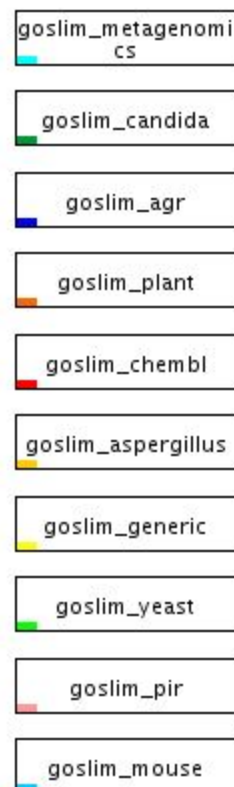
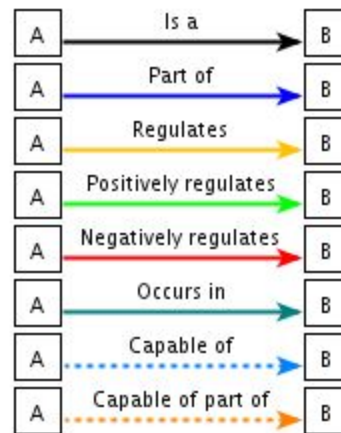
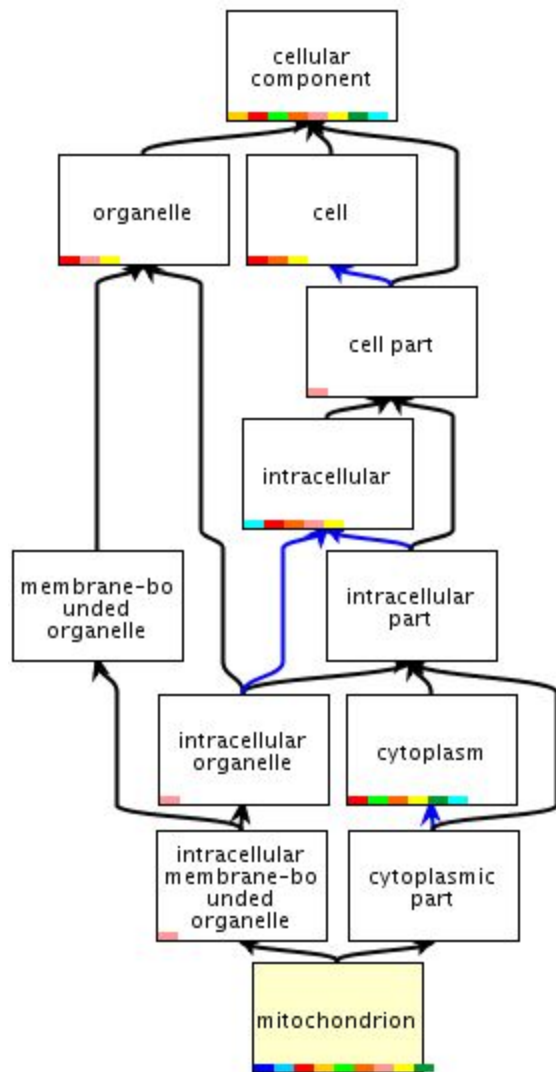
QuickGO - <http://www.ebi.ac.uk/QuickGO>

Molecular Function hierarchy pictureGO:0003924:



QuickGO - <http://www.ebi.ac.uk/QuickGO>

Cellular Component hierarchy pictureGO:0005739:



QuickGO - <http://www.ebi.ac.uk/QuickGO>

3.) SSEARCH against Shuffled SwissProt (550k)

The best scores are:

| | s-w bits | E(553474) | %_id | %_sim | alen | |
|-------------------------------|-----------------------------------|-----------|------|--------|-------|---------------------------------|
| sp shO60393 NOBOX_HUMAN_shuff | Homeobox protein NOBOX OS (691) | 158 | 44.1 | 0.0078 | 0.238 | 0.458 345 align |
| sp shQ0VG99 MESP2_HUMAN_shuff | Mesoderm posterior protei (397) | 144 | 40.9 | 0.041 | 0.266 | 0.463 188 align |
| sp shQ5G268 NETR_NOMLE_shuff | Neurotrypsin OS=Nomascus I (875) | 123 | 35.7 | 3.3 | 0.257 | 0.537 136 align |
| sp shQ8CHH5 GSC1L_MOUSE_shuff | GLTSCR1-like protein OS=M (1074) | 124 | 35.9 | 3.6 | 0.241 | 0.473 319 align |
| sp shQ9ERZ0 HEMGN_MOUSE_shuff | Hemogen OS=Mus musculus G (503) | 118 | 34.7 | 4 | 0.283 | 0.543 138 align |

sp|shP9WKV8|Y477_MYCTO_shuff Uncharacterized protein MT (148) 109 32.9 4.1 0.341 0.553 85 [align](#)
 sp|shQ7TNF8|RIMB1_MOUSE_shuff Peripheral-type benzodiaz (1846) 126 36.3 4.9 0.294 0.516 126 [align](#)
 (these were the only hits given)

BLAST against SwissProt

Select seq sp|P04637.4|P53_HUMAN
 Select seq sp|P13481.1|P53_CHLAE RecName: Full=Cellular tumor antigen p53; AltName: Full=Tumor suppressor p53
 746 746 100% 0.0 96% P13481.1
 Select seq sp|P56424.1|P53_MACMU RecName: Full=Cellular tumor antigen p53; AltName: Full=Tumor suppressor p53
 739 739 100% 0.0 96% P56424.1
 Select seq sp|Q9TTA1.1|P53_TUPBE RecName: Full=Cellular tumor antigen p53; AltName: Full=Tumor suppressor p53
 707 707 100% 0.0 93% Q9TTA1.1
 Select seq sp|O36006.1|P53_MARMO RecName: Full=Cellular tumor antigen p53; AltName: Full=Tumor suppressor p53
 672 672 100% 0.0 86% O36006.1
 Select seq sp|Q95330.1|P53_RABIT RecName: Full=Cellular tumor antigen p53; AltName: Full=Tumor suppressor p53
 651 651 100% 0.0 86% Q95330.1
 Select seq sp|Q8SPZ3.1|P53_DELLE RecName: Full=Cellular tumor antigen p53; AltName: Full=Tumor suppressor p53
 632 632 100% 0.0 85% Q8SPZ3.1
 Select seq sp|Q9WUR6.1|P53_CAVPO RecName: Full=Cellular tumor antigen p53; AltName: Full=Tumor suppressor p53
 618 618 100% 0.0 78% Q9WUR6.1
 Select seq sp|Q00366.1|P53_MESAU RecName: Full=Cellular tumor antigen p53; AltName: Full=Tumor suppressor p53
 613 613 100% 0.0 78% Q00366.1
 Select seq sp|Q9TUB2.1|P53_PIG RecName: Full=Cellular tumor antigen p53; AltName: Full=Tumor suppressor p53
 612 612 100% 0.0 82% Q9TUB2.1
 Select seq sp|O09185.1|P53_CRIGR RecName: Full=Cellular tumor antigen p53; AltName: Full=Tumor suppressor p53
 612 612 100% 0.0 77% O09185.1

There were no sequences common to both the BLAST search and the SSEARCH

Topmost hit:

Protein: Cellular tumor antigen p53
 Organism: Chlorocebus aethiops (Green monkey)

>P13481.1 RecName: Full=Cellular tumor antigen p53; AltName: Full=Tumor suppressor p53

| Score | Expect | Method | Identities | Positives | Gaps |
|--|--------|------------------------------|---------------|---------------|-----------|
| 813 bits(2101) | 0.0 | Compositional matrix adjust. | 393/393(100%) | 393/393(100%) | 0/393(0%) |
| Query 1 MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPPLSQAMDDLMLSPDDIEQWFTEDPGP 60 | | | | | |
| MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPPLSQAMDDLMLSPDDIEQWFTEDPGP | | | | | |
| Sbjct 1 MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPPLSQAMDDLMLSPDDIEQWFTEDPGP 60 | | | | | |
| Query 61 DEAPRMPEAAPPVAPAPAAPTPAAPAPAPSWPLSSSVPSQKTYQGSYGFRGLHSGTAK 120 | | | | | |
| DEAPRMPEAAPPVAPAPAAPTPAAPAPAPSWPLSSSVPSQKTYQGSYGFRGLHSGTAK | | | | | |
| Sbjct 61 DEAPRMPEAAPPVAPAPAAPTPAAPAPAPSWPLSSSVPSQKTYQGSYGFRGLHSGTAK 120 | | | | | |
| Query 121 SVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGRVVRAMAIYKQSQHMTEVVRRCPHHE 180 | | | | | |
| SVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGRVVRAMAIYKQSQHMTEVVRRCPHHE | | | | | |
| Sbjct 121 SVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGRVVRAMAIYKQSQHMTEVVRRCPHHE 180 | | | | | |

Query 181 RCSDSDGLAPPQHILIRVEGNLRVEYLDDRNTFRHSVVVPYEPPEVGSDCTTIHYNMCNS 240
RCSDSDGLAPPQHILIRVEGNLRVEYLDDRNTFRHSVVVPYEPPEVGSDCTTIHYNMCNS
Sbjct 181 RCSDSDGLAPPQHILIRVEGNLRVEYLDDRNTFRHSVVVPYEPPEVGSDCTTIHYNMCNS 240

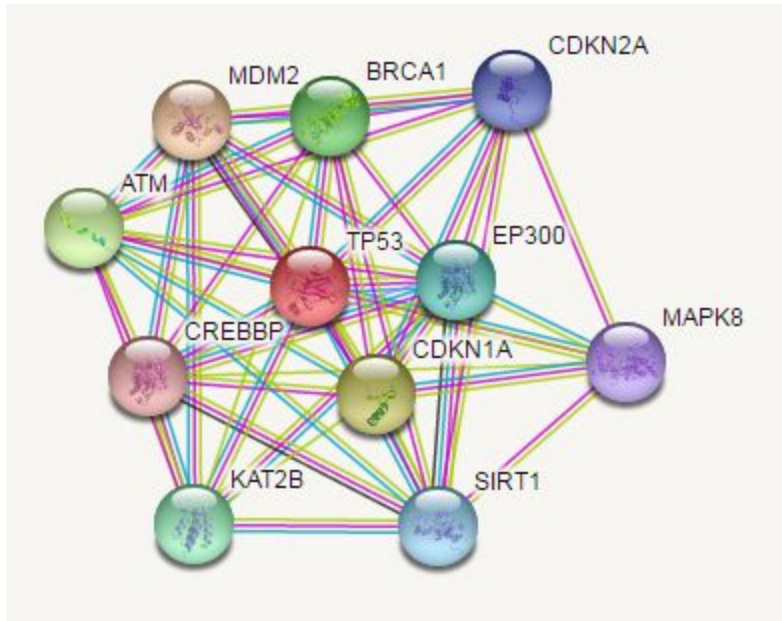
Query 241 SCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCACPGRRRTEENLRKKGEPHHELP 300
SCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCACPGRRRTEENLRKKGEPHHELP
Sbjct 241 SCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCACPGRRRTEENLRKKGEPHHELP 300

Query 301 PGSTKRALPNNTSSSPQPKKKPLDGEYFTLQIRGRERFEMFRELNEALELKDAQAGKEPG 360
PGSTKRALPNNTSSSPQPKKKPLDGEYFTLQIRGRERFEMFRELNEALELKDAQAGKEPG
Sbjct 301 PGSTKRALPNNTSSSPQPKKKPLDGEYFTLQIRGRERFEMFRELNEALELKDAQAGKEPG 360

Query 361 GSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDSD 393
GSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDSD
Sbjct 361 GSRAHSSHLKSKKGQSTSRHKKLMFKTEGPDSD 393

These two proteins share a single experimentally validated GO category, GO:0005515
This GO category belongs to the sub ontology of Molecular function

| Code | Name | % | Count | |
|------------|--|--------|-------|--|
| GO:0005515 | protein binding | 100.00 | 2 | |
| GO:2001244 | positive regulation of intrinsic apoptotic signaling pathway | 50.00 | 1 | |
| GO:2000379 | positive regulation of reactive oxygen species metabolic process | 50.00 | 1 | |
| GO:1900119 | positive regulation of execution phase of apoptosis | 50.00 | 1 | |
| GO:0097718 | disordered domain specific binding | 50.00 | 1 | |
| GO:0097252 | oligodendrocyte apoptotic process | 50.00 | 1 | |
| GO:0090403 | oxidative stress-induced premature senescence | 50.00 | 1 | |
| GO:0090399 | replicative senescence | 50.00 | 1 | |
| GO:0090200 | positive regulation of release of cytochrome c from mitochondria | 50.00 | 1 | |
| GO:0072717 | cellular response to actinomycin D | 50.00 | 1 | |
| GO:0072332 | intrinsic apoptotic signaling pathway by p53 class mediator | 50.00 | 1 | |
| GO:0072331 | signal transduction by p53 class mediator | 50.00 | 1 | |
| GO:0071480 | cellular response to gamma radiation | 50.00 | 1 | |
| GO:0071479 | cellular response to ionizing radiation | 50.00 | 1 | |
| GO:0071456 | cellular response to hypoxia | 50.00 | 1 | |
| GO:0071158 | positive regulation of cell cycle arrest | 50.00 | 1 | |
| GO:0051974 | negative regulation of telomerase activity | 50.00 | 1 | |
| GO:0051721 | protein phosphatase 2A binding | 50.00 | 1 | |
| GO:0051289 | protein homotetramerization | 50.00 | 1 | |
| GO:0051087 | chaperone binding | 50.00 | 1 | |
| GO:0048147 | negative regulation of fibroblast proliferation | 50.00 | 1 | |
| GO:0047485 | protein N-terminus binding | 50.00 | 1 | |
| GO:0046982 | protein heterodimerization activity | 50.00 | 1 | |
| GO:0046677 | response to antibiotic | 50.00 | 1 | |



4.)

5.)

Cellular tumor antigen p53 (Antigen NY-CO-13) (Phosphoprotein p53) (Tumor suppressor p53)
Homo sapiens (Human).

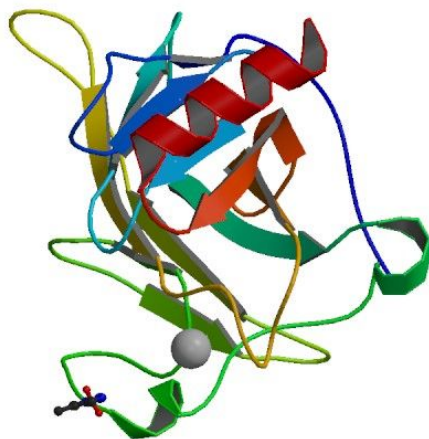
The parameters have been computed for the following feature:

| | | | | |
|----|-------|---|-----|-----------------------------|
| FT | CHAIN | 1 | 393 | Cellular tumor antigen p53. |
|----|-------|---|-----|-----------------------------|

The computation has been carried out on the complete sequence (**393** amino acids).

Molecular weight (Da): 43653.18 (average mass), 43625.38 (monoisotopic mass)

Theoretical pI: 6.33



6.)

7.) First I would go to <http://www.uniprot.org/uniprot/P04637#structure> to find what gene expresses it. That gene is TP53. I would then search for TP53 at <https://ghr.nlm.nih.gov/> which would lead me here: <https://ghr.nlm.nih.gov/gene/TP53> where I can find good information on diseases associated with this gene as well as other links to more material on this gene. After looking through that site and the associated links I would then go to <https://www.ncbi.nlm.nih.gov/projects/genome/guide/human/> and search for TP53 which would lead me here <https://www.ncbi.nlm.nih.gov/gene/7157>. From here I can get large amounts of more detailed, more in depth knowledge of the gene including statistical data on what tissues are most likely to express the gene, where this gene sits on the chromosome in relation to other genes, research papers written about this gene and much more info. I would probably be able to find everything i need to know about this gene here.

Q2

1.) Humans: 367 receptors / 99287 total proteins= .003696% of total proteome

Mice : 392 receptors / 66017 total proteins = .005937% of total proteome

Total protein counts were obtained as described in Q1) 2 b

Receptor counts were taken from this paper linked to by InterPro

<http://europepmc.org/abstract/MED/12679517>

2.) Top 15 most common protein domains for Escherichia coli (strain K12)

| Pfam | Description | Number of sequences | Number of occurrences |
|--------------------------------|---|---------------------|-----------------------|
| ABC_tran | ABC transporter | 78 | 95 |
| MFS_1 | Major Facilitator Superfamily | 57 | 72 |
| BPD_transp_1 | Binding-protein-dependent transport system inner membrane component | 51 | 52 |
| LysR_substrate | LysR substrate binding domain | 47 | 47 |
| HTH_1 | Bacterial regulatory helix-turn-helix protein, lysR family | 46 | 46 |
| Response_reg | Response regulator receiver domain | 38 | 38 |
| HATPase_c | Histidine kinase-, DNA gyrase B-, and HSP90-like ATPase | 32 | 32 |
| RHS_repeat | RHS Repeat | 5 | 28 |
| Fimbrial | Fimbrial protein | 28 | 28 |
| Fer4 | 4Fe-4S binding domain | 20 | 26 |
| Hexapep | Bacterial transferase hexapeptide (six repeats) | 13 | 24 |
| DDE_Tnp_1 | Transposase DDE domain | 21 | 21 |
| GntR | Bacterial regulatory proteins, gntR family | 21 | 21 |
| HTH_AraC | Bacterial regulatory helix-turn-helix proteins, AraC family | 12 | 20 |
| GerE | Bacterial regulatory proteins, luxR family | 20 | 20 |

Yes the following domains are present in humans as well:

| | | | |
|--------------------------|-------------------------------|-----|-----|
| ABC_tran | ABC transporter | 107 | 155 |
| MFS_1 | Major Facilitator Superfamily | 133 | 148 |