

Brenda Miner

Annotations for ['1f4l', '1mkh'] (6.1.1.10)

October 4, 2005

1 Downloads

1.1 Structure-files

- 1f4l.pdb from <http://projects.villa-bosch.de/dbase/pdba/mmol/1f4l.pdb.gz>
- 1mkh.pdb from <http://projects.villa-bosch.de/dbase/pdba/mmol/1mkh.pdb.gz>

1.2 Engineering

- | Name | Organism | Impact |
|-------|--------------------------|--------------|
| A355C | Saccharomyces cerevisiae | ['activity'] |

Commentary: site-directed mutagenesis, 115

- | Name | Organism | Impact |
|-------------|--------------------------|-------------------|
| C337A/C340A | Saccharomyces cerevisiae | ['notclassified'] |

Commentary: site-directed mutagenesis, inactive

- | Name | Organism | Impact |
|-------|--------------------------|--------------|
| C350A | Saccharomyces cerevisiae | ['activity'] |

Commentary: site-directed mutagenesis, 1.5

- | Name | Organism | Impact |
|-------------|--------------------------|-------------------|
| C350A/C353A | Saccharomyces cerevisiae | ['notclassified'] |

Commentary: site-directed mutagenesis, inactive

- | Name | Organism | Impact |
|-------|--------------------------|-------------------|
| C353A | Saccharomyces cerevisiae | ['notclassified'] |

Commentary: site-directed mutagenesis, catalytically inactive, no in vivo complementation of a deficient yeast strain

- | Name | Organism | Impact |
|-------|--------------------------|-------------------|
| C367A | Saccharomyces cerevisiae | ['notclassified'] |

Commentary: site-directed mutagenesis, catalytically inactive, no in vivo complementation of a deficient yeast strain, mutant shows a second zinc-binding knuckle structure

- | Name | Organism | Impact |
|-------|--------------------------|--------------|
| D348G | Saccharomyces cerevisiae | ['activity'] |

Commentary: site-directed mutagenesis, 4.7

- | Name | Organism | Impact |
|-------|--------------------------|--------------|
| D370A | Saccharomyces cerevisiae | ['activity'] |

Commentary: site-directed mutagenesis, 8.7

- | Name | Organism | Impact |
|-------|------------------|--------------|
| D666A | Escherichia coli | ['activity'] |

Commentary: activity is similar to the wild-type enzyme

- | Name | Organism | Impact |
|------|------------------|-------------------|
| G23A | Escherichia coli | ['notclassified'] |

Commentary: mutant enzymes: L22A variant, G23A variant, G23P variant, H21N variant, H21Q variant, H24N variant, and H24Q variant, with reduced catalytic efficiency and lowered maximal rate

- | Name | Organism | Impact |
|------|------------------|-------------------|
| G23P | Escherichia coli | ['notclassified'] |

Commentary: mutant enzymes: L22A variant, G23A variant, G23P variant, H21N variant, H21Q variant, H24N variant, and H24Q variant, with reduced catalytic efficiency and lowered maximal rate

- | Name | Organism | Impact |
|-------|--------------------------|-------------------|
| G347R | Saccharomyces cerevisiae | ['notclassified'] |

Commentary: site-directed mutagenesis, catalytically inactive, no in vivo complementation of a deficient yeast strain

- | Name | Organism | Impact |
|------|------------------|-------------------|
| H21N | Escherichia coli | ['notclassified'] |

Commentary: mutant enzymes: L22A variant, G23A variant, G23P variant, H21N variant, H21Q variant, H24N variant, and H24Q variant, with reduced catalytic efficiency and lowered maximal rate

- | Name | Organism | Impact |
|------|------------------|-------------------|
| H21Q | Escherichia coli | ['notclassified'] |

Commentary: mutant enzymes: L22A variant, G23A variant, G23P variant, H21N variant, H21Q variant, H24N variant, and H24Q variant, with reduced catalytic efficiency and lowered maximal rate

- | Name | Organism | Impact |
|------|------------------|-------------------|
| H24N | Escherichia coli | ['notclassified'] |

Commentary: mutant enzymes: L22A variant, G23A variant, G23P variant, H21N variant, H21Q variant, H24N variant, and H24Q variant, with reduced catalytic efficiency and lowered maximal rate

- | Name | Organism | Impact |
|------|------------------|-------------------|
| H24Q | Escherichia coli | ['notclassified'] |

Commentary: mutant enzymes: L22A variant, G23A variant, G23P variant, H21N variant, H21Q variant, H24N variant, and H24Q variant, with reduced catalytic efficiency and lowered maximal rate

- | Name | Organism | Impact |
|-------|--------------------------|--------------|
| I363N | Saccharomyces cerevisiae | ['activity'] |

Commentary: site-directed mutagenesis, 84

- | Name | Organism | Impact |
|-------|------------------|--------------|
| K335Q | Escherichia coli | ['activity'] |

Commentary: mutants produced by site-directed mutagenesis, Lys335-Gln substitution results in a complete loss of activity, similar loss of activity is observed when Lys335 is changed into alanine, glutamic acid, or arginine

- | Name | Organism | Impact |
|-------|--------------|--------------|
| K860A | Homo sapiens | ['activity'] |

Commentary: site-directed mutagenesis, reduced activity

- | Name | Organism | Impact |
|-------|--------------|--------------|
| K863A | Homo sapiens | ['activity'] |

Commentary: site-directed mutagenesis, reduced activity

- | Name | Organism | Impact |
|-------|--------------|--------------|
| K866A | Homo sapiens | ['activity'] |

Commentary: site-directed mutagenesis, reduced activity

- | Name | Organism | Impact |
|-------|--------------|---------------|
| K880A | Homo sapiens | ['mechanism'] |

Commentary: site-directed mutagenesis, altered kinetics

- | Name | Organism | Impact |
|------|------------------|-------------------|
| L22A | Escherichia coli | ['notclassified'] |

Commentary: mutant enzymes: L22A variant, G23A variant, G23P variant, H21N variant, H21Q variant, H24N variant, and H24Q variant, with reduced catalytic efficiency and lowered maximal rate

- | Name | Organism | Impact |
|------|--------------------------|-------------------|
| more | Saccharomyces cerevisiae | ['notclassified'] |

Commentary: a strain carrying the MES1 structure gene on a high copy number plasmid, pFL1

- | Name | Organism | Impact |
|------|------------------|--------------|
| more | Escherichia coli | ['kinetics'] |

Commentary: mutant strains with an increased Km for methionine

- | Name | Organism | Impact |
|------|-----------------------------|--------------|
| more | Bacillus stearothermophilus | ['activity'] |

Commentary: construction of a truncated enzyme form with 25

- | Name | Organism | Impact |
|------|--------------------------|-------------------|
| more | Saccharomyces cerevisiae | ['notclassified'] |

Commentary: construction of an inactive strain by gene disruption

- | Name | Organism | Impact |
|------|------------------|-------------------|
| more | Escherichia coli | ['notclassified'] |

Commentary: construction of C-terminal truncated mutant, removal of beta10 strand and insertion of a stop codon at position 666, M665

- | Name | Organism | Impact |
|-------|--------------------------|--------------|
| P338I | Saccharomyces cerevisiae | ['activity'] |

Commentary: site-directed mutagenesis, 74

- | Name | Organism | Impact |
|-------|--------------|--------------|
| R857A | Homo sapiens | ['activity'] |

Commentary: site-directed mutagenesis, reduced activity

- | Name | Organism | Impact |
|------|------------------|--------------|
| T10M | Escherichia coli | ['activity'] |

Commentary: natural mutant, 5

- | Name | Organism | Impact |
|------|------------------|--------------|
| Y15A | Escherichia coli | ['activity'] |

Commentary: natural mutant, very low residual activity, complementation of an enzyme-deficient Escherichia coli strain

- | Name | Organism | Impact |
|------|------------------|-------------------|
| Y94H | Escherichia coli | ['notclassified'] |

Commentary: natural mutant, unstable, no complementation of an enzyme-deficient Escherichia coli strain

2 Alignment of Saccharomyces cerevisiae's Mutations

align to reference sequence:

Compare sequences of length 349 and 750

1	VQQGGYKARWEDIKKEDIRQFLYEAKNINHGELIALHFRDDIKLAPANLMKNIKEG

59	IEPMYPTIISVAYIINLGVAVVADSKEPSQSFLTNDLKNEQLFQNGRASLSMAIE

117	LGRREHGLEMEKVYNSLIENIDKVLGDYNSVKKPDFKPVVGNYSKANVFKILRNVFN

175	GLNALLESNRAVFDDWSFHSDSSEPRVSALYYRWVSPSIGSDQANNGFVGVGRSK

233	SFKGNEYQLYETTNLHHLMTWNEETGLQSGPFVVTHFPVNDKGMFQYLSVHEPNW

291	WQKWEKTYNSTISVYGITADFWVYLVKDKYKELPVPTGWVLDRTICRPKLGDKLWS

```

QTITKSNKSWNGEESAKEVWESIQSELKDLSLFIHDSYKPEPSADDLKCRPNILEF
-----D??C???I???
349 PDLLAGCKDCQDGRADDYHCKPCEGEVYRDALYSNHVPCYLQKMSQEELYGNSNLK
???A?C?C?DG?????????PC????????????????????????????????????
407 TFIHQAIETQKDTTTRGFYDFGIQFWKYVDSHIKHYKDCLQRPTVGEELAKTETAT
????????????????????????????????????????????????????????????
465 GYEDTGCIFLANYNRGKCYRAFIDASLVSGIINGLHPVNNVYPLASTILINRENPK
????????????????????????????????????????????????????????????
523 PLIESDKPKVTLDAQIKVAGINKFSSSDRPVHKKALAVANHVKSPLFPKSHVLS
????????????????????????????????????????????????????????????
581 NLAYVNAFLILDTTTTLPEKLEVLYNEIAKNTLVEVHQOPLEKHYYLLNQLSALAFQY
????????????????????????????????????????????????????????????
639 KDSTQGEFDDMVYRLIANADFLLFPEKTNRLEMAANDNDVEPKLNKSAYELALAIK
????????????????????????????????????????????????????????????
697 L-NNALQLHAPHKKSDFSILFS
?L-----
755

```

minimal editing distance: 4045 Alignment: -22

2.1 Align To Sequence In 1f4l.pdb

Compare sequences of length 750 and 549

```

E-----KSA--EVLA-EVQRMDIRNYL--AKFPNVKHGLLPQQIGDWTLETNLFAE
1 VQQGGYK-ARWEDIKKE-----DIRQFLYEAKNIN--HGEL---IA---LH---F--
ARETLKPLVPK-LYTMLVRFLNI--G---MSC-IAQLDADR---GEQKAVVW----
59 -RDDIK-LAPANL--MK---NIKEGIEPMYPTIISSVAYIINLGV--AVVADSKE
PAQE---DVYRNA-LD---A-LAM-IE--RVAKGFERSEWAEG--IVEAADTFT
117 PSQSFLTNDLK-NEQLFQNGRASLSMAIELGRREHGLEM-EKVYNSLI-ENID---
KYLQPDALLESALVG--DFRKNIFGANRSALNVVK---NVID---A---NVRQVFD
175 KVLG-DYN-S--VKKPDF-KPVVG-NYKA-NVFKILRNVFENGLNALLESNRAVFD
ELNLDIDDI--RSSLKATYYRRL-SDADFHNLWTS--AKI-FTG--RSKSMKAGNV
233 DWSFHSDSSEPRVS--ALYY-RWVSPSIG-----SDQANNGFVGVGRSKSFK-GNE
TV-YGHVFLNSPKRFNSGE--LMA-PWFLS-HFYVIDKGIHFHYL---EATSDK-KW
291 YQLYETTNLHHLMTWNE-ETGLQSGP-FVVTHFPVNDKGMFQYLSVHEPNNWWQKW
YEDFSVSDGRKDCLNKF-SGMYGIPADLWVYF----YKG-PANPIEFGFYPADRSI
349 -E-----KTY-NSTISV-YGITADFWVYLVKDKYKELPV-PT--GWVL-DRTI
DWQQLGSEFWEQM--K-NAVQ--EQLAGSRTWAQLMESF-SPL-DFE-F-HESDRM
407 CRPKLGDKLWSQTITKSNKSWNGEESAKEV-W----ESIQSELKDLSLFIHDSYK
VPTAGSVVSK--PEILETPSYTAGCVECNDGYQDPKCKPCTGKVFDRDPLFMGKEP
465 EPSADDL--KCRPNILEFPDLLAGCKDCQDGRADDYHCKPCEGEVYRDALYSNHVP
DYLQ--SITR---N---KIFGN---E--KL---RS-Y---ILE-----S-----
523 CYLQKMSQEELYGNSNLKTFIHQAIETQKDTTTRGFYDFGIQFWKYVDSHIKHYKD
-LQRN---EESH-TSH---YNDYS-I---NF--G----AF-DTQHEQSMEGIMQE-
581 CLQRPTVGEELAKTETATGYEDTGCIFLANYNRGKCYRAFIDASLV-S--GIINGL
-P-----TIGLQQA---K-LMI----P--TGHADD--ACI--F-----
639 HPVNNVYPLASTI-LINRENPKPL-IESDKPKVTLDAQIKVAGINKFSSSDRPVHK
-----N-V---E---HGRMRQ--YRVWVDAQ-IH-----E-LM-----
697 KALAVANHVKSPLFPKSHVLSNLAY-VN--AFLILDTTTTLPEKLEVLYNEIAKN
----HG--L--HI-----S-----G-----NAY---P-----

```

```

755 TLVEVHQQPLEKHHYLLNQLSSALAFQYKDSTQGEFDDMVYRLIANADFLLFPEKTNR
L--ACT--V--LIK-----K--AXX-----
813 LEMAANDNDVEPKLNKSAYELALAIKLNALQLHAPHKKSKDFSILFS

```

Matches with 36 **A355C** : 333 G \mapsto 143 G

C337A/C340A : 315 K \mapsto 0 -

C350A : 328 D \mapsto 138 D

C350A/C353A : 328 D \mapsto 138 D

C353A : 331 V \mapsto 141 V

C367A : 345 G \mapsto 155 G

D348G : 326 L \mapsto 136 L

D370A : 348 C \mapsto 158 C

G347R : 325 Y \mapsto 135 F

I363N : 341 D \mapsto 151 P

L22A : 0 S \mapsto 0 -

P338I : 316 Q \mapsto 126 Q

2.2 Align To Sequence In 1mkh.pdb

Compare sequences of length 750 and 108

```

-----RV---K-----A-----G-L-----KVE-----K-----
1 VQQGGYKARWEDIKKEDIRQFLYEAKNINHGELIALHFRDDIKLAPANLMKNIKEG
-----D---P-----M---
59 IEPMYPTIISSVAYIINLGVAVVADSKEPSQSFLTNDLKNEQLFQNGRASLSMAIE
L-----L-----A-V---RE---
117 LGRREHGLEMEKVNLSLIENIDKVLGDYNSVKKPDFKPVVGNYKANVFKILRNVFN
G-----DD-----A-----A-----
175 GLNALLESNNRAVFDDDWSFHSDSSEPRVSALYRWVSPSIGSDQANNGFVGVGRSK
-----L---L---M---G-QSGI-----GX---LKK---P---
233 SFGKNEYQLYETTNLHHLMTWNEETGLQSGPFVVTHFPVNDKGMFQYLSVHEPNNW
---EL--N-----A--V--V-----V---VY-R-----
291 WQKWEKTYNSTISVYGITADFWVYLVKDKYKELPVPTGWVLDRTICRPKLGDKLWS
-----N-----L--L-----E-----E---
349 QTITKSNKSWNGEESAKEVWESIQSELKDLSLFIHDSYKPEPSADDLKCRPNILEF
P-----KY_---_Y--K---_---_---K-----L-G-----
407 PDLLAGCKDCQDGRADDYHCKPCEGEVYRDALYSNHVPCYLQKMSQEELYGNSNLK
-----A-----V-----L-TR--V-E-----
465 TFIHQAIETQKDTTTRGFYDFGIQFWKYVDSHIKHYKDCLQRPTVGEELAKTETAT
---D-G---L-----D---VK-----V--VY-L-----K
523 GYEDTGCIFLANYNRGKCYRAFIDASLVSGIINGLHPVNNVYPLASTILINRENPK
---D-----A-----N-----P-HDK---V-----E-----
581 PLIESDKPKVTLDAQIKVAGINKFSSSDRPVHKKALAVANHVKSPLEPFKSHVLSS
-----I-----I-K-----
639 NLAYVNAFLILDTTTTLPEKLEVLYNEIAKNTLVEVHQQPLEKHHYLLNQLSSALAFQY
---G---V-RL---D-L---K-----AF-D-DF--KV---YM-----
697 KDSTQGEFDDMVYRLIANADFLLFPEKTNRLEMAANDNDVEPKLNKSAYELALAIK
-----_-----
755 LNNALQLHAPHKKSKDFSILFS

```

Matches with 20 **A355C** : 333 G \mapsto 0 -
C337A/C340A : 315 K \mapsto 49 K
C350A : 328 D \mapsto 0 -
C350A/C353A : 328 D \mapsto 0 -
C353A : 331 V \mapsto 0 -
C367A : 345 G \mapsto 0 -
D348G : 326 L \mapsto 0 -
D370A : 348 C \mapsto 0 -
G347R : 325 Y \mapsto 0 -
I363N : 341 D \mapsto 0 -
L22A : 0 S \mapsto 0 -
P338I : 316 Q \mapsto 0 -
Compare sequences of length 750 and 108

```

-----RV---K-----A----G-L-----KVE-----K-----
1 VQQGGYKARWEDIKKEDIRQFLYEAKNINHGELIALHFRDDIKLAPANLMKNIKEG
-----D---P-----M---
59 IEPMYPTIISSVAYIINLGVAVVADSKEPSQSFLTNDLKNEQLFQNGRASLSMAIE
L-----L-----A-V---RE---
117 LGRREHGLEMEKVYNSLIENIDKVLGDYNSVKKPDFKPVVGNKYANVFKILRNVFN
G-----DD-----A-----A-----
175 GLNALLESNRAVFDWDFHSDSSEPRVSAlyYRWVSPSIGSDQANNGFVGVGRSK
-----L---L---M---G-QSGI-----GX---LKK---P---
233 SFKGNeyQLyETTnlHhLMTWNEETGLQSGPFVvTHFPVNDKGMFQYLSVHEPNNW
---EL--N-----A--V--V-----V---VY-R-----
291 WQKWEKTYNSTISVYGITADFWVYLvkDKYKELPVPTGwVLDRTICRPKLGDKLWS
-----N-----L--L-----E-----E-----
349 QTITKSNKSWNGEESAKEVWESIQSELkdLSLFIHDSYKPEPSADDLKCRPNILEF
P-----KY-_-_-Y-K-_-_-_-_-K-----L-G-----
407 PDLLAGCKDCQDGRADDYHCKPCEGEVYRDALYSNHVPCYLQKMSQEELYGNSNLK
-----A-----V-----L-TR--V-E-----
465 TFIHQAIETQKDTTTRGFYDFGIQFWKYVDShIKHYKDClQRPTVGEELAKTETAT
---D-G---L-----D---VK-----V--VY-L-----K
523 GYEDTGCIFLANYNRGKCYRAFIDASLVSGIINGLHPVNNVYPLASTILINRENPK
---D---A-----N---P-HDK---V-----E-----
581 PLIESDKPKVTLDAQIKVAGINKFSSSDRPVHKKALAVANHVKSPLKSPKSHVLSS
-----I-----I-K-----
639 NLAYVNAFLILDTTTLPEKLEVLYNEIAKNTLVEVHQQPLEKHyllNQLSALAFQY
---G---V-RL---D-L---K---AF-D-DF--KV---YM-----
697 KDSTQGEFFDDMVYRLIANADFLLFPEKTRLEMAANDNDVEPKLNKSAYELALAIK
-----_-
755 LNNALQLHAPHKKSDFSILFS

```

Matches with 20 **A355C** : 333 G \mapsto 0 -
C337A/C340A : 315 K \mapsto 49 K
C350A : 328 D \mapsto 0 -
C350A/C353A : 328 D \mapsto 0 -
C353A : 331 V \mapsto 0 -
C367A : 345 G \mapsto 0 -
D348G : 326 L \mapsto 0 -

D370A : 348 C \mapsto 0 -
G347R : 325 Y \mapsto 0 -
I363N : 341 D \mapsto 0 -
L22A : 0 S \mapsto 0 -
P338I : 316 Q \mapsto 0 -

3 Alignment of *Saccharomyces cerevisiae*'s Mutations

align to reference sequence:

Compare sequences of length 1 and 750

1	VQQGGYKARWEDIKKEDIRQFLYEAKNINHGELIALHFRDDIKLAPANLMKNIKEG

	IEPMYPTIISSVAYIINLGVAVVADSKEPSQSFLTNDLKNEQLFQNGRASLSMAIE
59	-----
	LGRREHGLEMEKVNSLIENIDKVLGDYNSVKKPDFKPVVGNKYKANVFKILRNVFN
117	-----
	GLNALLESNRAVFDWWSFHSDSSEPRVSALYRWWSPSIGSDQANNFGVGVGRSK
175	-----
	SFKGNEYQLYETTNLHHLMTWNEETGLQSGPFVVTHTFPVNDKGMFQYLSVHEPNW
233	-----
	WQKWEKTYNSTISVYGITADFWVYLKDKYKELPVPTGWVLDRTICRPKLGDKLWS
291	-----
	QTITKSNKSWNGEESAKEVWESIQSELKDLSLFIHDSYKPEPSADDLKCRPNILEF
349	-----
	PDLLAGCKDCQDGRADDYHCKPCEGEVYRDALYSNHVPCYLQKMSQEELYGNSNLK
407	-----
	TFIHQAIETQKDTTTRGFYDFGIQFWKYVDSEHIKHYKDCLQRPTVGEELAKTETAT
465	-----
	GYEDTGCIFLANYNRGKCYRAFIDASLVSGIINGLHPVNNVYPLASTILINRENPK
523	-----
	PLIESDKPKVTLDAQIKVAGINKFSSSDRPVHKKALAVANHVKSPLPEPKSHVLSS
581	-----
	NLAYVNAFLIILDTTTLPEKLEVLYNEIAKNTLVEVHQQPLEKHYYLLNQLSALAFQY
639	-----
	KDSTQGEFDDMVYRLIANADFLLFPEKTNRLEMAANDNDVEPKLNKSAYELALAIk
697	-----L-----
	LNNALQLHAPHKKSDFSILFS
755	-----

minimal editing distance: 30 Alignment: 3

3.1 Align To Sequence In 1f4l.pdb

Compare sequences of length 750 and 549

1	E-----KSA--EVLA-EVQRMDIRNYL--AKFPNVKHGLLPQQIGDWTLETNLFAE
	VQQGGYK-ARWEDIKKE-----DIRQFLYEAKNIN--HGEL---IA---LH---F--

```

ARETLKPLVPK-LYTMLVRFLNI--G---MSC-IAQLDADR---GEQKAVVW----
59 -RDDIK-LAPANL--MK----NIKEGIEPMYPTIISSVAYIINLGV--AVVADSKE
PAQE----DVRNA-LD----A-LAM-IE--RVAKGFERSEWAEG--IVEAADTFT
117 PSQSFLTNDLK-NEQLFQNGRASLSMAIELGRREHGLEM-EKVYNSLI-ENID---
KYLQPDALASALVG--DFRKNIFGANRSALNVVK---NVID---A----NVRQVFD
175 KVLG-DYN-S--VKKPDF-KPVVG-NYKA-NVFKILRNVFNGLNALLESNRAVFD
ELNLDIDDI--RSSLKATYYRRL-SDADFHNLWTS--AKI-FTG--RSKSMKAGNV
233 DWSFHSDSSEPRVS--ALYY-RWVSPSIG-----SDQANNGFVGVGRSKSFK-GNE
TV-YGHVFLNSPKRFNSGE--LMA-PWFLS-HFYVIDKGI FHYL---EATSDK-KW
291 YQLYETTNLHHLMTWNE-ETGLQSGP-FVVTHFPVNDKGMFQYLSVHEPNWQKW
YEDFSVSDGRKDCLNKF-SGMYGIPADLWVYF----YKG-PANPIEFGFYPADR SI
349 -E-----KTY-NSTISV-YGITADFWVYLKDKYKELPV-PT--GWVL-DRTI
DWQQLGSEFWEQM--K-NAVQ--EQLAGSRTWAQLMESF-SPL-DFE-F-HESDRM
407 CRPKLGDKLWSQTITKSNKSWNGEESAKEV-W----ESIQSELKDL S LFIHDSYKP
VPTAGSVVSK--PEILETPSYTAGCVCNDGYQDPSKCKPCTGKVFRDPLFMGKEP
465 EPSADDL--KCRPNILEFPDLLAGCKDCQDGRADDYHCKPCEGEVYRDALYSNHVP
DYLQ--SITR---N---KIFGN---E--KL---RS-Y---ILE-----S-----
523 CYLQKMSQEELYGNSNLKTFIHQAIEQKDTTTRGFYDFGIQFWKYVD SHIKHYKD
-LQRN---EESH-TSH---YNDYS-I---NF--G---AF-DTQHEQSMEGIMQE-
581 CLQRPTVGEELAKTETATGYEDTGCIFLANYNRGKCYRAFIDASLV-S--GIINGL
-P-----TIGLQQA---K-LMI-----P--TGHADD--ACT--F-----
639 HPVNNVYPLASTI-LINRENPKPL-IESDKPKVTLDAQIKVAGINKFSSSDRPVHK
-----N-V---E---HGRMRQ--YRVWVDAQ-IH-----E-LM-----
697 KALAVANHVKSPLLEPKSHVLSNLAY-VN--AFLILDTTTLEKLEVLYNEIAKN
---HG--L--HI-----S-----G-----NAY-----P-----
755 TLVEVHQPLEKHYYLLNQLSALAFQYKDYDSTQGEFDDMVYRLIANADFLFP EKTNR
L--ACT---V---LIK-----K---AXX-----
813 LEMAANDNDVEPKLNKSAYELALAIKLNALQLHAPHKSKDFSILFS

```

Matches with 36 L22A : 25 L → 0 -

3.2 Align To Sequence In 1mkh.pdb

Compare sequences of length 750 and 108

```

-----RV---K-----A-----G-L-----KVE-----K-----
1 VQGGGYKARWEDIKKEDIRQFLYEAKNINHGELIALHFRDDIKLAPANLMKNIKEG
-----D---P-----M---
59 IEPMYPTIISSVAYIINLGVAVVADSKEPSQSFLTNDLKNEQLFQNGRASLSMAIE
L-----L-----A-V---RE---
117 LGRREHGLEMEKVYNSLIENIDKVLGDYNSVKKPDFKPVVGNKYANVFKILRNVFN
G-----DD-----A-----A-----
175 GLNALLESNRAVFD DWSFHSDSSEPRVSALYYRWVSPSIGSDQANNGFVGVGRSK
-----L---L---M---G-QSGI-----GX---LKK--P---
233 SFGKNEYQLYETTNLHHLMTWNEETGLQSGPFVVTHFPVNDKGMFQYLSVHEPNW
---EL--N-----A--V--V-----V---VY-R-----
291 WQKWEKTYNSTISVYGITADFWVYLKDKYKELPVPTGWVLDRTICRPKLGDKLWS
-----N-----L--L-----E-----E-----
349 QTITKSNKSWNGEESAKEVWESIQSELKDL S LFIHDSYKPEPSADDLKCRPNILEF

```

```

P-----KY-----Y--K-----K-----L-G-----
407 PDLLAGCKKDCQDGRADDYHCKPCEGEVYRDALYSNHVPCYLQKMSQEELYGNSNLK
-----A-----V-----LTR--V-E-----
465 TFIHQAIETQKDTTTRGFYDFGIQFWKYVDSHIKHYKDCCLRPTVGEELAKTETAT
---D-G---L-----D--VK-----V--VY-L-----K
523 GYEDTGCIFLANYNRGKCYRAFIDASLVSGIINGLHPVNNVYPLASTILINRENPK
---D---A-----N-----P-HDK---V-----E---
581 PLIESDKPKVTLDAQIKVAGINKFSSSDRPVHKKALAVANHVKSPLPEPKSHVLSS
-----I-----I-K-----
639 NLAYVNAFLILDTTTTLPEKLEVLVLYNEIAKNTLVEVHQPPLEKHYYLLNQLSALAFQY
-----G-----V-RL---D-L---K-----AF-D-DF--KV----YM--_
697 KDSTQGEFDDMVYRLIANADFLLFPEK'TNRLEMAANDNDVEPKLNKSAYELALAIK
-----
755 LNNALQLHAPHKKSDFSILFS

```

Matches with 20 L22A : 25 L → 0 -
Compare sequences of length 750 and 108

```

-----RV---K-----A-----G-L-----KVE-----K-----
1 VQQGGYKARWEDIKKEDIRQFLYEAKNINHGELIALHFRDDIKLAPANLMMKNIKEG
-----D---P-----M---
59 IEPMYPTIISVAYIINLGVAVVADSKEPSQSFLTNDLKNEQLFQNGRASLSMAIE
L-----L-----A-V---RE---
117 LGRREHGLEMEKVNSLIENIDKVLGDYNSVKKPDFKPVVGNKYKANVFKILRNVFN
G-----DD-----A-----A-----
175 GLNALLESNRAVFDWWSFHSDSSEPRVSALYRWVSPSIGSDQANNFVGVGRSK
-----L---L---M---G-QSGI-----GX---LKK--P---
233 SFKGNEYQLYETTNLHHLMTWNEETGLQSGPFVVTHTFPVNDKGMFQYLSVHEPNNW
---EL--N---A--V--V-----V--R-----
291 WQKWEKTYNSTISVYGITADFVWVYLKDKYKELPVPTGWVLDRTICRPKLGDKLWS
-----N-----L--L-----E-----E---
349 QTITKSNKSWNGEESAKEVWESIQSELKDLSLFIHDSYKPEPSADDLKCPRNILEF
P-----KY-----Y--K-----K-----L-G-----
407 PDLLAGCKKDCQDGRADDYHCKPCEGEVYRDALYSNHVPCYLQKMSQEELYGNSNLK
-----A-----V-----LTR--V-E-----
465 TFIHQAIETQKDTTTRGFYDFGIQFWKYVDSHIKHYKDCCLRPTVGEELAKTETAT
---D-G---L-----D--VK-----V--VY-L-----K
523 GYEDTGCIFLANYNRGKCYRAFIDASLVSGIINGLHPVNNVYPLASTILINRENPK
---D---A-----N-----P-HDK---V-----E---
581 PLIESDKPKVTLDAQIKVAGINKFSSSDRPVHKKALAVANHVKSPLPEPKSHVLSS
-----I-----I-K-----
639 NLAYVNAFLILDTTTTLPEKLEVLVLYNEIAKNTLVEVHQPPLEKHYYLLNQLSALAFQY
-----G-----V-RL---D-L---K-----AF-D-DF--KV----YM--_
697 KDSTQGEFDDMVYRLIANADFLLFPEK'TNRLEMAANDNDVEPKLNKSAYELALAIK
-----
755 LNNALQLHAPHKKSDFSILFS

```

Matches with 20 L22A : 25 L → 0 -

4 Alignment of Bacillus stearothermophilus's Mutations

align to reference sequence:

Compare sequences of length 1 and 649

```

1 KIKTGNPVDVTALSFGGGSGGALIMGESWEGRLKAPKLNVCIVKKGILEEPKY
-----
59 FEAIGSIVQRKEGGLDLQLKLLKDANKMREPQVVEAVRLDVKAFDDISIEEAAAQ
-----
117 KEEKAAEAPKGGQMHAKIYEVEVGIDLRPFLPEGKQVNTGEPILGFDYLSDWEKLS
-----
175 RDSIGLQTFIREPTRLFPQLLVATYRLSEALHAMVSALEERKSEEKALVWPQTED
-----
233 IYKNTRGILQWVASLAVSFEMREMAEYQVVERATESLDRDFPTKPGRYPIAGG
-----
291 FYKEIMAVTRHLLNGLDNALDYNIREIFGEPTFVGDSGFPVERLLYYRLADLGYRD
-----
349 IIMVPDVVNGKSKSMKGDKMLLWGHGFVKKPLPLGLAMLMIWPYITHFRVIEKGVL
-----
407 HVDAPWYKRFKEDNDTGYGLATIYNALADIWVYIVHKPDGPVKIGWDFTTRSVALD
-----
465 ELGPKIFNNIMENKRSEPQIFDPNEEYYQLLRDVYKSMRFFYSEEKVEVPRGCDP
-----
523 CNGDVLQRETYFSECPTCYWGEYEGLYIDGQEVLRAFIKEVIKKHREQTTRIFDDY
-----
581 SIDLKRWLEQIGAVIDDVYQQPTVGKEQAKRQIKQGHEDTGTLYMVDYGRRLRYRA
-----
639 MADGAVTTYAhGIHLKDSPYYIPTTLYFTKKEM
-----L-----

```

minimal editing distance: 30 Alignment: -4

4.1 Align To Sequence In 1f4l.pdb

Compare sequences of length 649 and 549

```

1 E-K-----S-----A---EV---L-A-----E----
KIKTGNPVDVTALSFGGGSGGALIMGESWEGRLKAPKLNVCIVKKGILEEPKY
59 -----VQRM---DIRNYL---A-KF--PNV-----K-----
FEAIGSIVQRKEGGLDLQLKLLKDANKMREPQVVEAVRLDVKAFDDISIEEAAAQ
-----H-----G--L---LPQ--Q-----I-G-D---WT-LE
117 KEEKAAEAPKGGQMHAKIYEVEVGIDLRPFLPEGKQVNTGEPILGFDYLSDWEKLS
TN---L--FA-EA-RETLKP-L-VPKLYTMLVRFLNIGM-SCIAQLDADRG-EQKA
175 RDSIGLQTFIREPTRLFPQLLVAT-YR-LSEALHA-MVS--A-LEE-RKSEEKA
VWVPAQ-EDVYRNA---LD--A-LAMIERVAKGFE-RSEWAEGL--V-EAA----D
233 LVWP-QTEDIYKNTRGILQWVASLA----VS--FEMR-EMAEYQVVERATESLD
T-F-TK---YLQPDALLESALVGDFRKNIFGAN-RSALNVVKNVIDANVRQVFDELN
291 RDFPTKPGRYP-PIAG-----G-FYKEIM-AVTRHLLNGLDNALDYNIREIFGEPT
LDIDDI----RSSLKATYYYRLSDADFHNLWTSAKIFT-----GRSKSMKAGNVTV

```

```

349 FVGDSGFPVERL-L---YY-RL--AD---LGYRDIIMVPDVVNGKSKSMK-GDKML
Y-GHVFLNSPKRFNSGE--LMAPWFLSHFYVIDKGIHYLEATSDKKWYEDFSVSD
407 LWGHGFVKKPLPL--GLAMLMIPWYITHFRVIEKGV LHVD-AP----WYKRF----
GRK-DCLNKFSGMYG---IP-A--DLWVYFY-KGPANP--IEFG--FYPADRSI--
465 --KED--NDT-G-YGLATIYNALADIWVYIVHK-PDGPVKI--GWDFTT--RSVAL
DWQQLGSE-FWEQMKNVQEQ-QLAGSRTWA-QLMESFSPL-DFFFHESDRMVPTAG
523 DE--LGPKIFNNIMENKRSEPQIFDPNEEYYQL-----LRDVYK--SMRFFY---
SV--VSKPEILETPSYTAGCVCN-DGYQDP---SKCKPCTGKVFDRDPLFMGKEPD
581 SEEKV-K-EV---PR---GCDPCNGDVLQRETYFSEC-P-TCYW-----G-E--
Y--LQSITRNKIFGNEKLSYILESLO--RNEESHTSHYNDYSINFGAFDTQH--E
639 YEGLY-ID-----GQEVLRAFI KEVIKKHR-EQT-TRIFDDYSI-----DLKRWLE
QSMEG--IM---QEPTIGLQOAKLMIPTGHADDACIFNVEHGRMRQYRVWVD-AQI
697 QI--GAVIDDVYQQPTVKGKEQAKRQIKQGHEDTGTLYMVDYGR LRKYRAMADGAVT
HELMHGLHI--SGNAY-PLACTVLI--KKAXXX-
755 TYA-HGIHLKDSPPY-YIPT--T-LYFTKKEM---

```

Matches with 38 L22A : 18 L → 20 I

4.2 Align To Sequence In 1mkh.pdb

Compare sequences of length 649 and 108

```

RVK-----A--G-----LKV-E-----KDPML-----L-----
1 KIKTGNPVHQDVTALSFGGGSGGALIMGESWEGRLKAPKLNVAIVKKGILEEPKY
--A--V-R-EG--D-----DA-----AL-L-----
59 FEAGSIVQRKEGGLDLQKLLKDKANKMREPQVVEAVRLDVKAFDDISIEEARAAQ
-----M-----G-----QS--G--I-GX--LK--K--
117 KEEKAAEAPKGGOMHAKIYEVEVGIDLRPFLPEGKQVNTGEPILGFDYLSDWEKLS
-----P-----E-LNA-V-----V-----
175 RDSIGLQTFIREPTRTLFPQLLVATYRLSEALHAMVSALEERKSEEKALVWPQTED
-----V--V-----Y-R-----
233 IYKNTRGILQWVASLAVSFEMREMAEYQRVVERATESLDRDFPTKPGRYPIIAGG
-----N-L--L--E--EPK-----YY-----
291 FYKEIMAVTRHLLNGLDNALDYNIREIFGEPFVGD SGFPVERLLYYRLADLGYRD
-----K-K-----L-GA--V--L-----T--RV-E--
349 IIMVPDVVNGKSKSMKGDKMLLWGHGFVKKPLPLGLAMLMIPWYITHFRVIEKGV
--D-----G--L-----D--V--K--V-----V--
407 HVDAPWYKRFKEDNDTGYGLATIYNALADIWVYIVHKPDGPVKIGWDFTT RSVALD
-----Y--L--K-----DA-----
465 ELGPKIFNNIMENKRSEPQIFDPNEEYYQLLRDVYKSMRFFYSEEKVKEVPRGCDP
--N-----PH-----D-----K-V--E--I--
523 CNGDVLQRETYFSECPTCYWGEYEGLYIDGQEVLRAFI KEVIKKHREQTTRIFDDY
--I--K-----G-V-----R-----L--D--L-K--A
581 SIDLKRWLEQIGAVIDDVYQQPTVKGKEQAKRQIKQGHEDTGTLYMVDYGR LRKYRA
F--D-----D-----F-KVYM-
639 MADGAVTTYAHGIHLKDSPPYYIPTTLYFTKKEM-

```

Matches with 23 L22A : 18 L → 0 -
Compare sequences of length 649 and 108

```

RVK-----A--G-----LKV-E-----KDPML-----L-----
1  KIKTGNPVHQDVTALSFGGGSGGALIMGESWEGRLKAPKLNVCIVKKGILEEPKY
--A---V-R-EG--D-----DA-----AL-L-----
59  FEAIGSIVQRKEGGDLQKLLKDKANKMREPQVVEAVRLDVKAFDDISIEEARAAQ
-----M-----G-----QS--G--I-GX--LK--K--
117  KEEKAAEAPKGGQMHAKIYEVEVVGIDLRPFLPEGKQVNTGEPILGFDYLSDWEKLS
-----P-----E-LNA-V-----V-----
175  RDSIGLQTFIREPTRLFPQLLVATYRLSEALHAMVSALEERKSEEKALVWPQTED
-----V---V---Y-R-----
233  IYKNTRGILQWVASLAVSFEMREMAEYQRVVERATESLDRDFPTKPGRYPIAGG
-----N-L---L---E---EPK-----YY-----
291  FYKEIMAVTRHLLNGLDNALDYNIREIFGEP'FVGDSGFPVERLLYYRLADLGYRD
-----K-K---L---GA--V---L-----T--RV-E---
349  IIMVPDVVNGKSKSMGDKMLLWGHGFVKKPLPLGLAMLMI PWYITHFRVIEKGVL
--D-----G-V---L---D--V---K---V-----V-----
407  HVDAPWYKRFKEDNDTGYGLATIYNALADIWVYIVHKPDGPVKIGWDF'TTRSVALD
-----Y--L---K-----DA
465  ELGPKIFNNIMENKRSEPQIFDPNEEYQLLRDVYKSMRFFYSEEKVKVPRGCDP
--N-----PH-----D-----K-V-----E---I---
523  CNGDVLQRETYFSECP'CYWGEYEGLYIDGQEVLRAFIKEVIKKHREQTTRIFDDY
--I--K-----G-V-----R-----L--D--L-K--A
581  SIDLKRWLEQIGAVIDDVYQQPTVGKEQAKRQIKQGHEDTGTLYMVDYGRLRKYRA
F-D-----D-----F-KVYM-
639  MADGAVTTYAHGIHLKDSPYYIPTTLYFTKKEM-

```

Matches with 23 **L22A** : 18 L → 0 -

5 Alignment of Escherichia coli's Mutations

align to reference sequence:
 Compare sequences of length 85 and 676

```

KVQHGPAGADPSLLFIDKGGPGAAMVMGESIGFRMKRPALNAVMI THRGILAQPD
1  -----
PYASRIGSFVNRKEGGDLTLRLLKDSGEVFEANEILAVRLDVKAFDDFTITEQIP
59  -----
DDALPGTVPAAAAKVEEKSAEVLAEVQRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
117  -----
TNLFAEARETLKPLVPKLYTMLVRFNLNIGMSCIAQLDADRGEQKAVVWPAQEDVYR
175  -----
NALDALAMI ERVAKGFERSEWAE GIVEAADTFTKYLQPDAL ESALVGDFRKNIFGA
233  -----
NRSALNVVKNVIDANVRQVFDELNLDIDDIRSSLKATYYYRLSDADFHNLWTSAKI
291  -----
F'TGRSKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGI FHYLEATS
349  -----
DKKWYEDFSVSDGRKCLNKFSGMYGIPADLWVYFYKGPANPI EFGFYPADRSIDW

```

```

407 -----
      QQLGSEFWEQMKNVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
465 -----
      EILETPSYTAGCVECNDGYQDPSKCKPCTGKVFRDPLFMGKEPDYLSITRNKIFG
523 -----
      NEKLRSYILESQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQOAK
581 -----Y????????????????????????????????????????????
      LMIPTGHADDACIFNVEHGRMRQYRVWVDAQIHELMHGLHISGNAYPLACTVLIKK
639 -----Y?????T-----
      AVQT
697 -----

```

minimal editing distance: 30 Alignment: -1

5.1 Align To Sequence In 1f4l.pdb

Compare sequences of length 676 and 549

```

-----E-----K-----
1  KVQHGPKAGADPSLLFIDKGGPGAAMVMGESIGFRMKR PALNAVMI THRGILAQPD
   -S-----A-E---V-L---A-----E---
59  PYASRIGSFVNRKEGGLDLTLRLLKDSGEVFEANEILAVRLDVKAFFDDFTITEQIP
   -----V-----QRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
117 DDALPGTVPAAAAKVEEKSAEVLAEVQRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
     TNLFAEARETLKPLVPKLYTMLVRFLNIGMSCIAQLDADRGEQKAVVWPAQEDVYR
175 TNLFAEARETLKPLVPKLYTMLVRFLNIGMSCIAQLDADRGEQKAVVWPAQEDVYR
     NALDALAMI ERVAKGFERSEWAE GIVEAADTF TKYLQPDAL ESALVGD FRKNIFGA
233 NALDALAMI ERVAKGFERSEWAE GIVEAADTF TKYLQPDAL ESALVGD FRKNIFGA
     NRSALNVVKNVIDANVRQVFDELNLDIDDIRS SLKATYYYRLSDAD FHNLWTS AKI
291 NRSALNVVKNVIDANVRQVFDELNLDIDDIRS SLKATYYYRLSDAD FHNLWTS AKI
     FTGRSKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGI FHYLEATS
349 FTGRSKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGI FHYLEATS
     DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPI EFGFYPADRS IDW
407 DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPI EFGFYPADRS IDW
     QQLGSEFWEQMKNVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
465 QQLGSEFWEQMKNVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
     EILETPSYTAGCVECNDGYQDPSKCKPCTGKVFRDPLFMGKEPDYLSITRNKIFG
523 EILETPSYTAGCVECNDGYQDPSKCKPCTGKVFRDPLFMGKEPDYLSITRNKIFG
     NEKLRSYILESQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQOAK
581 NEKLRSYILESQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQOAK
     LMIPTGHADDACIFNVEHGRMRQYRVWVDAQIHELMHGLHISGNAYPLACTVLIKK
639 LMIPTGHADDACIFNVEHGRMRQYRVWVDAQIHELMHGLHISGNAYPLACTVLIKK
     AXXX-
697 AVQT-

```

Matches with 88 T10M : 9 T → 10 T
Y15A : 14 Y → 15 Y
Y94H : 93 Y → 94 Y

5.2 Align To Sequence In 1mkh.pdb

Compare sequences of length 676 and 108

```

RV---KAG---L---K-----V--E-----KDPML-----LA---
1 KVQHGPKAGADPSLLFIDKGGPGAAMVMGESIGFRMKRPALNAVMITHRGILAQPD
-----V-R-EG--D-----D-----A---A--L-----
59 PYASRIGSFVNRKEGGLDLTLRLLKDSGEVFEANEILAVRLDVKAFDDFTTITEQIP
---LMGQ-----SG-----IGX-L-K---K---PE-----L-
117 DDALPGTVPAAAAKVEEKSAEVLAEVQRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
-N--A-----V-----V-----VV-----YR
175 TNLFAEARETLKPLVPKLYTMLVRFLNIGMSCIAQLDADRGEQKAVVWPAQEDVYR
N-L--L--E-----E-----P-----K-----
233 NALDALAMIERVAKGFERSEWAEGIVEAADTFTKYLQPDALESALVGDFRKNIFGA
-----YY-----K-
291 NRSALNVVKNVIDANVRQVFDELNLDIDDIRSSLKATYYYRLSDADFHNLWTSAKI
---K-----L-----G--A-----V-----L--TR
349 FTGRSKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGIFHYLEATS
V---ED---G---L-----D--V---KVV-----Y-----
407 DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPIEFGFYPADRSIDW
--L---KDAN-----P---H--DK-V-----
465 QQLGSEFWEQMKNAVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
EII-----K---G-V-R-L-----D-L-----K---
523 EIIETPSYTAGCVECNDGYQDPSKCKPCTGKVFRDPLFMGKEPDYLQSITRNKIFG
-----AFD-----
581 NEKLRSYILESLQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQQAK
---D---FKV---Y---M-----
639 LMIPTGHADDACIFNVEHGRMRQYRVWVDAQIHELMHGLHISGNAYPLACTVLIKK
----
697 AVQT

```

Matches with 22 **T10M** : 9 T \mapsto 0 -

Y15A : 14 Y \mapsto 0 -

Y94H : 93 Y \mapsto 0 -

Compare sequences of length 676 and 108

```

RV---KAG---L---K-----V--E-----KDPML-----LA---
1 KVQHGPKAGADPSLLFIDKGGPGAAMVMGESIGFRMKRPALNAVMITHRGILAQPD
-----V-R-EG--D-----D-----A---A--L-----
59 PYASRIGSFVNRKEGGLDLTLRLLKDSGEVFEANEILAVRLDVKAFDDFTTITEQIP
---LMGQ-----SG-----IGX-L-K---K---PE-----L-
117 DDALPGTVPAAAAKVEEKSAEVLAEVQRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
-N--A-----V-----V-----VV-----YR
175 TNLFAEARETLKPLVPKLYTMLVRFLNIGMSCIAQLDADRGEQKAVVWPAQEDVYR
N-L--L--E-----E-----P-----K-----
233 NALDALAMIERVAKGFERSEWAEGIVEAADTFTKYLQPDALESALVGDFRKNIFGA
-----YY-----K-
291 NRSALNVVKNVIDANVRQVFDELNLDIDDIRSSLKATYYYRLSDADFHNLWTSAKI
---K-----L-----G--A-----V-----L--TR
349 FTGRSKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGIFHYLEATS
V---ED---G---L-----D--V---KVV-----Y-----

```

```

407 DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPIEFGFYPADRSIDW
--L-----K-DAN-----P-----H--DK-V-----
465 QQLGSEFWEQMKNVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
EII-----K-----G-V-R--L-----D-L-----K---
523 EILETPSYTAGCVECNDGYQDPSKCKPCTGKVFRDPLFMGKEPDYLSITRNKIFG
-----AFD-----
581 NEKLRSYILESQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQOAK
-----D---FKV-----Y-----M-----
639 LMIPTGHADDACIFNVEHGRMRQYRVWVDAQIHELMHGLHISGNAYPLACTVLIKK
-----
697 AVQT

```

Matches with 22 T10M : 9 T → 0 -
Y15A : 14 Y → 0 -
Y94H : 93 Y → 0 -

6 Alignment of Escherichia coli's Mutations

align to reference sequence:

Compare sequences of length 645 and 676

```

KVQHGPAGADPSLLFIDKGGPGAAMVMGESIGFRMKRPALNAVMI THRGILAQPD
1 -----D????????????????????????????????????????????????
PYASRIGSFVNRKEGGLDLTLRLLKDSGEVFEANEILAVRLDVKAFDDFTITEQIP
59 ?????????????????????????????????????????????????????????
DDALPGTVPAAAAKVEEKSAEVLAEVQRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
117 ?????????????????????????????????????????????????????????
TNLFAEARETLKPLVPKLYTMLVRFNIGMSCIAQLDADRGEQKAVVWPAQEDVYR
175 ?????????????????????????????????????????????????????????
NALDALAMI ERVAKGFERSEWAE GIVEAADTFTKYLQPDAL ESALVGDFRKNIFGA
233 ?????????????????????????????????????????????????????????
NRSALNVKKNVIDANVRQVFDELNLDIDDIRSSLKATYYYRLSDADFHNLWTS AKI
291 ?????????????????????????????????????????????????????????
F'TGRSKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGI FHYLEATS
349 ?????????????????????????????????????????????????????????
DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPIEFGFYPADRSIDW
407 ?????????????????????????????????????????????????????????
QQLGSEFWEQMKNVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
465 ?????????????????????????????????????????????????????????
EILETPSYTAGCVECNDGYQDPSKCKPCTGKVFRDPLFMGKEPDYLSITRNKIFG
523 ?????????????????????????????????????????????????????????
NEKLRSYILESQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQOAK
581 ?????????????????????????????????????????????????????????
LMIPTGHADDACIFNVEHGRMRQYRVWVDAQIHELMHGLHISGNAYPLACTVLIKK
639 ?????????????????????????????????????????????????????????L-----
AVQT
697 ----

```

minimal editing distance: 30 Alignment: -1

6.1 Align To Sequence In 1f4l.pdb

Compare sequences of length 676 and 549

```

-----E-----K-----
1 KVQHGPKAGADPSLLFIDKGGPGAAMVMGESIGFRMKRPALNAVMITHRGILAQPD
---S-----A-E---V-L---A-----E---
59 PYASRIGSFVNRKEGGLDLTLRLLKDSGEVFEANEILAVRLDVKAFFDDFTITEQIP
-----V-----QRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
117 DDALPGTVPAAAAKVEEKSAEVLAEVQRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
TNLFAEARETLKPLVPKLYTMLVRFNLNIGMSCIAQLDADRGEQKAVVWPAQEDVYR
175 TNLFAEARETLKPLVPKLYTMLVRFNLNIGMSCIAQLDADRGEQKAVVWPAQEDVYR
NALDALAMIERVAKGFERSEWAEGIVEAADTFTKYLQPDALLESALVGDFRKNIFGA
233 NALDALAMIERVAKGFERSEWAEGIVEAADTFTKYLQPDALLESALVGDFRKNIFGA
NRSALNVVKNVIDANVRQVFDELNLDIDDIRSSLKATYYYRLSDADFHNLWTSAKI
291 NRSALNVVKNVIDANVRQVFDELNLDIDDIRSSLKATYYYRLSDADFHNLWTSAKI
FTGRSKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGIFHYLEATS
349 FTGRSKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGIFHYLEATS
DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPIEFGFYPADRSIDW
407 DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPIEFGFYPADRSIDW
QQLGSEFWEQMKNVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
465 QQLGSEFWEQMKNVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
EILETPSYTAGCVECNDGYQDPSKCKPCTGKVFRDPLFMGKEPDYLSITRNKIFG
523 EILETPSYTAGCVECNDGYQDPSKCKPCTGKVFRDPLFMGKEPDYLSITRNKIFG
NEKLSYILESLQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQQAK
581 NEKLSYILESLQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQQAK
LMIPTGHADDACIFNVEHGRMRQYRVWVDAQIHELMHGLHISGNAYPLACTVLIKK
639 LMIPTGHADDACIFNVEHGRMRQYRVWVDAQIHELMHGLHISGNAYPLACTVLIKK
AXXX-
697 AVQT-

```

Matches with 88 D666A : 665 D \mapsto 0 -
L22A : 21 L \mapsto 22 L

6.2 Align To Sequence In 1mkh.pdb

Compare sequences of length 676 and 108

```

RV---KAG--L---K-----V--E-----KDPML-----LA---
1 KVQHGPKAGADPSLLFIDKGGPGAAMVMGESIGFRMKRPALNAVMITHRGILAQPD
-----V-R-EG--D-----D-----A---A--L-----
59 PYASRIGSFVNRKEGGLDLTLRLLKDSGEVFEANEILAVRLDVKAFFDDFTITEQIP
---LMGQ-----SG-----IGX-L-K---K---PE-----L-
117 DDALPGTVPAAAAKVEEKSAEVLAEVQRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
-N--A-----V-----V-----VV-----YR
175 TNLFAEARETLKPLVPKLYTMLVRFNLNIGMSCIAQLDADRGEQKAVVWPAQEDVYR
N-L--L---E-----E-----P-----K-----
233 NALDALAMIERVAKGFERSEWAEGIVEAADTFTKYLQPDALLESALVGDFRKNIFGA
-----YY-----K-
291 NRSALNVVKNVIDANVRQVFDELNLDIDDIRSSLKATYYYRLSDADFHNLWTSAKI

```

```

-----K-----L-----G--A-----V-----L--TR
349 FTGRSKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGI FHYLEATS
V---ED---G---L---D--V---KVV-----Y-----
407 DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPIEFGFYPADRSIDW
--L-----KDAN-----P-----H--DK-V-----
465 QQLGSEFWEQMKNVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
EII-----K-----G-V-R--L-----D-L-----K---
523 EILETPSYTAGCVECNDGYQDPSKCKPCTGKVFRDPLFMGKEPDYLSITRNKIFG
-----AFD-----
581 NEKLRSYILESLQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQQAK
-----D---FKV-----Y-----M-----
639 LMIPTGHADDACIFNVEHGRMRQYRVVWDAQIHELMHGLHISGNAYPLACTVLIKK
-----
697 AVQT

```

Matches with 22 **D666A** : 665 D \mapsto 0 -
L22A : 21 L \mapsto 0 -
Compare sequences of length 676 and 108

```

RV---KAG---L---K-----V--E-----KDPML-----LA---
1 KVQHGPKAGADPSLLFIDKGGPGAAMVMGESIGFRMKRPALNAVMI THRGILAQPD
---V-R-EG--D---D-----A---A--L-----
59 PYASRIGSFVNRKEGGLDLTLRLLKDSGEVFEANEILAVRLDVKAFDDFTITEQIP
---LMGQ---SG-----IGX-L-K---K---PE---L---
117 DDALPGTVPAAAAKVEEKSAEVLAEVQRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
-N--A-----V-----V-----VV-----YR
175 TNLFAEARETLKPLVPKLYTMLVRFNLNIGMSCIAQLDADRGEQKAVVWPAQEDVYR
N-L--L--E---E-----P-----K-----
233 NALDALAMIERVAKGFERSEWAE GIVEAADTFTKYLQPDALLESALVGDFRKNIFGA
-----YY---K-----
291 NRSALNVVKNVIDANVRQVFDELNLDIDDIRSSLKATYYRLSDADFHNLWTSAKI
-----K-----L-----G--A-----V-----L--TR
349 FTGRSKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGI FHYLEATS
V---ED---G---L---D--V---KVV-----Y-----
407 DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPIEFGFYPADRSIDW
--L-----KDAN-----P-----H--DK-V-----
465 QQLGSEFWEQMKNVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
EII-----K-----G-V-R--L-----D-L-----K---
523 EILETPSYTAGCVECNDGYQDPSKCKPCTGKVFRDPLFMGKEPDYLSITRNKIFG
-----AFD-----
581 NEKLRSYILESLQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQQAK
-----D---FKV-----Y-----M-----
639 LMIPTGHADDACIFNVEHGRMRQYRVVWDAQIHELMHGLHISGNAYPLACTVLIKK
-----
697 AVQT

```

Matches with 22 **D666A** : 665 D \mapsto 0 -
L22A : 21 L \mapsto 0 -

7 Alignment of Escherichia coli's Mutations

align to reference sequence:

Compare sequences of length 1 and 676

1	KVQHGP	KAGADPSLLFIDKGGPGAAMVMGESIGFRMKRPALNAVMITHRGILAQPD
	-----	-----
59	PYASRIGSFVNRKEGGLDLTLRLLKDSGEVFEANEILAVRLDVKAFDDFTITEQIP	
	-----	-----
117	DDALPGTVPAAAAKVEEKSAEVLAEVQRMDIRNYLAKFPNVKHGLLPQQIGDWTLE	
	-----	-----
175	TNLF AEARETLKPLVPKLYTMLVRFLNIGMSCIAQLDADRGEQKAVVWPAQEDVYR	
	-----	-----
233	NALDALAMI ERVAKGFERSEWAE GIVEAADTFTKYLQPDAL ESALVGDFRKNIFGA	
	-----	-----
291	NRSALNVVKNVIDANVRQVFDELNLDIDDIRSSLKATYYYYRLSDADFHNLWTS AKI	
	-----	-----
349	FTGRSKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGI FHYLEATS	
	-----	-----
407	DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPIEFGFYPADRSIDW	
	-----	-----
465	QQLGSEFWEQMKNAVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP	
	-----	-----
523	EILETPSYTAGCVCNDGYQDPSKCKPCTGKVFRDPLFMGKEPDYLSITRNKIFG	
	-----	-----
581	NEKLSYI LESLQRNEESH TSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQQAK	
	-----	-----
639	LMIPTGHADDACIFNVEHGRMRQYRVWVDAQIHELMH _g LHISGNAYPLACTVLIKK	
	-----	-----L-----
697	AVQT	
	-----	-----

minimal editing distance: 30 Alignment: 3

7.1 Align To Sequence In 1f4l.pdb

Compare sequences of length 676 and 549

	-----	-----E-----K-----	-----
1	KVQHGP	KAGADPSLLFIDKGGPGAAMVMGESIGFRMKRPALNAVMITHRGILAQPD	
	-----	-----S-----A-E---V-L---A-----E---	-----
59	PYASRIGSFVNRKEGGLDLTLRLLKDSGEVFEANEILAVRLDVKAFDDFTITEQIP		
	-----	-----V-----QRMDIRNYLAKFPNVKHGLLPQQIGDWTLE	-----
117	DDALPGTVPAAAAKVEEKSAEVLAEVQRMDIRNYLAKFPNVKHGLLPQQIGDWTLE		
	-----	-----	-----
175	TNLF AEARETLKPLVPKLYTMLVRFLNIGMSCIAQLDADRGEQKAVVWPAQEDVYR		
	-----	-----	-----
233	NALDALAMI ERVAKGFERSEWAE GIVEAADTFTKYLQPDAL ESALVGDFRKNIFGA		
	-----	-----	-----
291	NRSALNVVKNVIDANVRQVFDELNLDIDDIRSSLKATYYYYRLSDADFHNLWTS AKI		
	-----	-----	-----

```

349 FTGRSKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGIHFHYLEATS
DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPIIEFGFYPADRSIDW
407 DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPIIEFGFYPADRSIDW
QQLGSEFWEQMKNVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
465 QQLGSEFWEQMKNVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
EILETPSYTAGCVECNDGYQDP SKCKPCTGKVFRDPLFMGKEPDYLOSITRNKIFG
523 EILETPSYTAGCVECNDGYQDP SKCKPCTGKVFRDPLFMGKEPDYLOSITRNKIFG
NEKLSYILES LQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQOAK
581 NEKLSYILES LQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQOAK
LMIPTGHADDACIFNVEHGRMRQYRVVWDAQIHELMHGLHISGNAYPLACTVLIKK
639 LMIPTGHADDACIFNVEHGRMRQYRVVWDAQIHELMHGLHISGNAYPLACTVLIKK
AVXX-
697 AVQT-

```

Matches with 88 L22A : 25 L → 26 L

7.2 Align To Sequence In 1mkh.pdb

Compare sequences of length 676 and 108

```

RV---KAG---L---K-----V--E-----KDPML-----LA---
1 KVQHGPKAGADPSLLFIDKGGPGAAMVMGESIGFRMKRPALNAVMITHRGILAQPD
-----V-R-EG--D-----D-----A---A--L-----
59 PYASRIGSFVNRKEGGLDLTLRLLKDSGEVFEANEILAVRLDVKAFFDFTITEQIP
---LMGQ-----SG-----IGX-L-K---K---PE-----L-
117 DDALPGLTPAAAAKVEEKSAEVLAEVQRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
-N--A-----V-----V-----VV-----YR
175 TNLFAEARETLKPLVPKLYTMLVRFLNIGMSCIAQLDADRGEQKAVVWPAQEDVYR
N-L--L--E-----E-----P-----K-----
233 NALDALAMI ERVAKGF ERSEWAE GIVEAADTFTKYLQPDAL ESALVGDFRKNIFGA
-----YY-----K-
291 NRSALNVVKNVIDANVRQVFDELNLDIDDIRSSLKATYYYRLSDADFHNLWTSAKI
---K-----L-----G--A-----V-----L--TR
349 FTGRSKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGIHFHYLEATS
V---ED---G---L-----D--V---KVV-----Y-----
407 DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPIIEFGFYPADRSIDW
--L-----K DAN-----P-----H--DK-V-----
465 QQLGSEFWEQMKNVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
EII-----K---G-V-R--L-----D-L-----K---
523 EILETPSYTAGCVECNDGYQDP SKCKPCTGKVFRDPLFMGKEPDYLOSITRNKIFG
-----AFD-----
581 NEKLSYILES LQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQOAK
-----D---FKV-----Y-----M-----
639 LMIPTGHADDACIFNVEHGRMRQYRVVWDAQIHELMHGLHISGNAYPLACTVLIKK
----
697 AVQT

```

Matches with 22 **L22A** : 25 L \mapsto 0 -
 Compare sequences of length 676 and 108

```

RV---KAG---L---K-----V--E-----KDPML-----LA---
1 KVQHGPKAGADPSLLFIDKGGPGAAMVMGESIGFRMKRPALNAVMITHRGLAQPD
-----V-R-EG--D-----D-----A---A--L-----
59 PYASRIGSFVNRKEGGLDLTLRLLKDSGEVFEANEILAVRLDVKAFDDFTITEQIP
---LMGQ-----SG-----IGX-L-K---K---PE-----L-
117 DDALPGTVPAAAAKVEEKSAEVLAEVQRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
-N--A-----V-----V-----VV-----YR
175 TNLFAEARETLKPLVPKLYTMLVRFLNIGMSCIAQLDADRGEQKAVVWPAQEDVYR
N-L--L---E-----E-----P-----K-----
233 NALDALAMIERVAKGFERSEWAEGIVEAADTFTKYLQPALESALVGDFRKNIFGA
-----YY-----K-
291 NRSALNVVKNVIDANVRQVFDELNLDIDDIRSSLKATYYYRLSDADFHNLWTSAKI
---K-----L-----G--A-----V-----L--TR
349 FTGRSKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGIFHYLEATS
V---ED---G---L-----D--V---KVV-----Y-----
407 DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPIEFGFYPADRSIDW
--L---KDAN-----P-----H--DK-V-----
465 QQLGSEFWEQMNAVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
EII-----K---G-V-R--L-----D-L-----K---
523 EILETPSYTAGCVECNDGYQDPSKCKPCTGKVFRDPLFMGKEPDYLQSITRNKIFG
-----AFD-----
581 NEKLRSYILESLQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQQAK
-----D---FKV-----Y-----M-----
639 LMIPTGHADDACIFNVEHGRMRQYRVVDAQIHELMHGLHISGNAYPLACTVLIKK
----
697 AVQT

```

Matches with 22 **L22A** : 25 L \mapsto 0 -

8 Alignment of Escherichia coli's Mutations

align to reference sequence:
 Compare sequences of length 4 and 676

```

KVQHGPKAGADPSLLFIDKGGPGAAMVMGESIGFRMKRPALNAVMITHRGLAQPD
1 -----
59 PYASRIGSFVNRKEGGLDLTLRLLKDSGEVFEANEILAVRLDVKAFDDFTITEQIP
-----
117 DDALPGTVPAAAAKVEEKSAEVLAEVQRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
-----
175 TNLFAEARETLKPLVPKLYTMLVRFLNIGMSCIAQLDADRGEQKAVVWPAQEDVYR
-----
233 NALDALAMIERVAKGFERSEWAEGIVEAADTFTKYLQPALESALVGDFRKNIFGA
-----
NRSALNVVKNVIDANVRQVFDELNLDIDDIRSSLKATYYYRLSDADFHNLWTSAKI

```

```

291  -----
    FTGRSKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGIHFHYLEATS
349  -----
    DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPIEFGFYPADRSIDW
407  -----
    QQLGSEFWEQMKNVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
465  -----
    EILETPSYTAGCVECNDGYQDPSKCKPCTGKVFRDPLFMGKEPDYLSITRNKIFG
523  -----
    NEKLRSYILESLQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQOAK
581  -----
    LMIPTGHADDACIFNVEHGRMRQYRVWVDAQIHELMHGLHISGNAYPLACTVLIKK
639  -----HGLH-----
    AVQT
697  -----

```

minimal editing distance: 30 Alignment: -1

8.1 Align To Sequence In 1f4l.pdb

Compare sequences of length 676 and 549

```

-----E-----K-----
1  KVQHGPKAGADPSLLFIDKGGPGAAMVMGESIGFRMKRPALNAVMIHRGILAQPD
    S-----A-E---V-L---A-----E---
59  PYASRIGSFVNRKEGGLDLTLRLLKDSGEVFEANEILAVRLDVKAFFDDFTITEQIP
    -----V-----QRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
117  DDALPGTVPAAAAKVEEKSAEVLAEVQRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
    TNLFAEARETLKPLVPKLYTMLVRFLNIGMSCIAQLDADRGEQKAVVWPAQEDVYR
175  TNLFAEARETLKPLVPKLYTMLVRFLNIGMSCIAQLDADRGEQKAVVWPAQEDVYR
    NALDALAMIERVAKGFERSEWAEGIVEAADTFTKYLQPDALLESALVGDFRKNIFGA
233  NALDALAMIERVAKGFERSEWAEGIVEAADTFTKYLQPDALLESALVGDFRKNIFGA
    NRSALNVVKNVIDANVRQVFDELNLDIDDIRSSLKATYYYRLSDADFNLWTSAKI
291  NRSALNVVKNVIDANVRQVFDELNLDIDDIRSSLKATYYYRLSDADFNLWTSAKI
    FTGRSKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGIHFHYLEATS
349  FTGRSKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGIHFHYLEATS
    DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPIEFGFYPADRSIDW
407  DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPIEFGFYPADRSIDW
    QQLGSEFWEQMKNVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
465  QQLGSEFWEQMKNVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
    EILETPSYTAGCVECNDGYQDPSKCKPCTGKVFRDPLFMGKEPDYLSITRNKIFG
523  EILETPSYTAGCVECNDGYQDPSKCKPCTGKVFRDPLFMGKEPDYLSITRNKIFG
    NEKLRSYILESLQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQOAK
581  NEKLRSYILESLQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQOAK
    LMIPTGHADDACIFNVEHGRMRQYRVWVDAQIHELMHGLHISGNAYPLACTVLIKK
639  LMIPTGHADDACIFNVEHGRMRQYRVWVDAQIHELMHGLHISGNAYPLACTVLIKK
    AXXX-
697  AVQT-

```

Matches with 88 **G23A** : 22 G \mapsto 23 G
G23P : 22 G \mapsto 23 G
H21N : 20 H \mapsto 21 H
H21Q : 20 H \mapsto 21 H
H24N : 23 H \mapsto 24 H
H24Q : 23 H \mapsto 24 H
L22A : 21 L \mapsto 22 L

8.2 Align To Sequence In 1mkh.pdb

Compare sequences of length 676 and 108

```

1  RV---KAG---L---K-----V--E-----KDPML-----LA---
KVQHGPKAGADPSLLFIDKGGPGAAMVMGESIGFRMKRPALNAVMITHRGILAQPD
-----V-R-EG--D-----D-----A---A--L-----
59  PYASRIGSFVNRKEGGLDLTLRLLKDSGEVFEANEILAVRLDVKAFFDFTITEQIP
---LMGQ-----SG-----IGX-L-K---K---PE-----L-
117  DDALPGTVPAAAAKVEEKSAEVLAEVQRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
-N--A-----V-----V-----VV-----YR
175  TNLFAEARETLKPLVPKLYTMLVRFLNIGMSCIAQLDADRGEQKAVVWPAQEDVYR
N-L--L--E-----E-----P-----K-----
233  NALDALAMIERVAKGFERSEWAE GIVEAADTFTKYLQPDALLESALVGDFRKNIFGA
-----YY-----K-
291  NRSALNVVKNVIDANVRQVFDELNLDIDDIRSSLKATYYYRLSDADFNHLWTSAKI
-----K-----L-----G--A-----V-----L--TR
349  FTGRSKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGIFHYLEATS
V---ED---G---L-----D--V---KVV-----Y-----
407  DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPIEFGFYPADRSIDW
--L-----KDAN-----P-----H--DK-V-----
465  QQLGSEFWEQMKNAVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
EII-----K-----G-V-R--L-----D-L-----K---
523  EILETPSYTAGCVECNDGYQDPSKCKPCTGKVFRDPLFMGKEPDYLSITRNKIFG
-----AFD-----
581  NEKLSYILESLQRNEESHSTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQQAK
---D---FKV-----Y-----M-----
639  LMIPTGHADDACIFNVEHGRMRQYRVVWDAQIHELMHGLHISGNAYPLACTVLIKK
----
697  AVQT

```

Matches with 22 **G23A** : 22 G \mapsto 0 -
G23P : 22 G \mapsto 0 -
H21N : 20 H \mapsto 0 -
H21Q : 20 H \mapsto 0 -
H24N : 23 H \mapsto 0 -
H24Q : 23 H \mapsto 0 -
L22A : 21 L \mapsto 0 -

Compare sequences of length 676 and 108

```

1  RV---KAG---L---K-----V--E-----KDPML-----LA---
KVQHGPKAGADPSLLFIDKGGPGAAMVMGESIGFRMKRPALNAVMITHRGILAQPD

```

```

-----V-R-EG--D-----D-----A--A--L-----
59 PYASRIGSFVNRKEGGLDLTLRLLKDSGEVFEANEILAVRLDVKAFDDFTITEQIP
---LMGQ-----SG-----IGX-L-K---K---PE---L-
117 DDALPGTVPAAAAKVEEKSAEVLAEVQRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
-N--A-----V-----V-----VV-----YR
175 TNLFAEARETLKPLVPKLYTMLVRFNLNIGMSCIAQLDADRGEQKAVVWPAQEDVYR
N-L--L--E-----E-----P-----K-----
233 NALDALAMIERVAKGFERSEWAE GIVEAADTF'TKYLQPDALLESALVGDFRKNIFGA
-----YY-----K-
291 NRSALNVVKNVIDANVRQVFDELNLDIDDIRSSLKATYYRLSADDFHNLWTSAKI
-----K-----L-----G--A-----V-----L--TR
349 F'TGRSKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGI FHYLEATS
V---ED---G---L-----D--V---KV-----Y-
407 DKKWYEDFVSVDGRKDCLNKFSGMYGIPADLWVYFYKGPANPIEFGFYPADRSIDW
--L-----KDN-----P-----H--DK--V-
465 QQLGSEFWEQMKNVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
EII-----K-----G-V-R--L-----D-L--K---
523 EILETPSYTAGCVECNDGYQDPSKCKPCTGKVFRDPLFMGKEPDYLSITRNKIFG
-----AFD-----
581 NEKLRSYILESLQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQQAQ
--D---FKV-----Y-----M-----
639 LMIPTGHADDACIFNVEHGRMRQYRVWVDAQIHELMHGLHISGNAYPLACTVLIKK
-----
697 AVQT

```

Matches with 22 **G23A** : 22 G \mapsto 0 -
G23P : 22 G \mapsto 0 -
H21N : 20 H \mapsto 0 -
H21Q : 20 H \mapsto 0 -
H24N : 23 H \mapsto 0 -
H24Q : 23 H \mapsto 0 -
L22A : 21 L \mapsto 0 -

9 Alignment of Escherichia coli's Mutations

align to reference sequence:

Compare sequences of length 1 and 676

```

KVQHGPAGADPSLLFIDKGGPGAAMVMGESIGFRMKRPALNAVMI THRGI LAQPD
1 -----
59 PYASRIGSFVNRKEGGLDLTLRLLKDSGEVFEANEILAVRLDVKAFDDFTITEQIP
-----
117 DDALPGTVPAAAAKVEEKSAEVLAEVQRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
-----
175 TNLFAEARETLKPLVPKLYTMLVRFNLNIGMSCIAQLDADRGEQKAVVWPAQEDVYR
-----
233 NALDALAMIERVAKGFERSEWAE GIVEAADTF'TKYLQPDALLESALVGDFRKNIFGA
-----

```

```

291 NRSALNVVKNVIDANVRQVFDELNLDIDDIRSSLKATYYYYRLSDADFHNLWTSAKI
-----
FTGRsKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGIFHYLEATS
349 -----K-----
DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPIEFGFYPADRSIDW
407 -----
QQLGSEFWEQMKNVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
465 -----
EILETPSYTAGCVECNDGYQDPKCKPCTGKVFRDPLFMGKEPDYLSITRNKIFG
523 -----
NEKLRSYILESLQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQQAK
581 -----
LMIPTGHADDACIFNVEHGRMRQYRVVWDAQIHELMHGLHISGNAYPLACTVLIKK
639 -----
AVQT
697 -----

```

minimal editing distance: 30 Alignment: -1

9.1 Align To Sequence In 1f4l.pdb

Compare sequences of length 676 and 549

```

-----E-----K-----
1 KVQHGPKAGADPSLLFIDKGGPGAAMVMGESIGFRMKRPALNAVMIHRGILAQPD
---S-----AE---VL---A-----E---
59 PYASRIGSFVNRKEGGLDLTLRLLKDSGEVFEANEILAVRLDVKAFFDDFTITEQIP
-----V-----QRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
117 DDALPGTVPAAAAKVEEKSAEVLAEVQRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
TNLFAEARETLKPLVPKLYTMLVRFLNIGMSCIAQLDADRGEQKAVVWPAQEDVYR
175 TNLFAEARETLKPLVPKLYTMLVRFLNIGMSCIAQLDADRGEQKAVVWPAQEDVYR
NALDALAMIERVAKGFERSEWAEGIVEAADTFTKYLQPDALESALVGDFRKNIFGA
233 NALDALAMIERVAKGFERSEWAEGIVEAADTFTKYLQPDALESALVGDFRKNIFGA
NRSALNVVKNVIDANVRQVFDELNLDIDDIRSSLKATYYYYRLSDADFHNLWTSAKI
291 NRSALNVVKNVIDANVRQVFDELNLDIDDIRSSLKATYYYYRLSDADFHNLWTSAKI
FTGRsKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGIFHYLEATS
349 FTGRsKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGIFHYLEATS
DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPIEFGFYPADRSIDW
407 DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPIEFGFYPADRSIDW
QQLGSEFWEQMKNVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
465 QQLGSEFWEQMKNVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
EILETPSYTAGCVECNDGYQDPKCKPCTGKVFRDPLFMGKEPDYLSITRNKIFG
523 EILETPSYTAGCVECNDGYQDPKCKPCTGKVFRDPLFMGKEPDYLSITRNKIFG
NEKLRSYILESLQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQQAK
581 NEKLRSYILESLQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQQAK
LMIPTGHADDACIFNVEHGRMRQYRVVWDAQIHELMHGLHISGNAYPLACTVLIKK
639 LMIPTGHADDACIFNVEHGRMRQYRVVWDAQIHELMHGLHISGNAYPLACTVLIKK
AXXX-
697 AVQT-

```

Matches with 88 **K335Q** : 334 K \mapsto 335 K

9.2 Align To Sequence In 1mkh.pdb

Compare sequences of length 676 and 108

```
RV---KAG---L---K-----V--E-----KDPML-----LA---
1 KVQHGPKAGADPSLLFIDKGGPGAAMVMGESIGFRMKRPALNAVMITHRGILAQPD
-----V-R-EG-D-----D-----A--A-L-----
59 PYASRIGSFVNRKEGGLDLTLRLLKDSGEVFEANEILAVRLDVKAFDDFTITEQIP
---LMGQ-----SG-----IGX-L-K---K---PE-----L-
117 DDALPGTVPAAAAKVEEKSAEVLAEVQRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
-N-A-----V-----V-----VV-----YR
175 TNLFAEARETLKPLVPKLYTMLVRFLNIGMSCIAQLDADRGEQKAVVWPAQEDYR
N-L--L--E-----E-----P-----K-----
233 NALDALAMIERVAKGFERSEWAEGIVEAADTFTKYLQPDALESALVGDFRKNIFGA
-----YY-----K-
291 NRSALNVVKNVIDANVRQVFDELNLDIDDIRSSLKATYYYRLSDADFHNLWTSSAKI
---K-----L-----G--A-----V-----L--TR
349 FTGRSKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGIFHYLEATS
V---ED---G---L-----D-V--KVV-----Y-----
407 DKKWYEDFSVSDGRKCLNKFSGMYGIPADLWVYFYKGPANPIEFGFYPADRSIDW
--L-----KDAN-----P-----H--DK-V-----
465 QQLGSEFWEQMKNAVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
EII-----K-----G-V-R-L-----D-L-----K---
523 EIILETPSYTAGCVECNDGYQDPSKCKPCTGKVFRDPLFMGKEPDYLQSITRNKIFG
-----AFD-----
581 NEKLRSYILESLQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQQAK
-----D--FKV-----Y-----M-----
639 LMIPTGHADDACIFNVEHGRMRQYRVWVDAQIHELMHGLHISGNAYPLACTVLIKK
----
697 AVQT
```

Matches with 22 **K335Q** : 334 K \mapsto 49 K

Compare sequences of length 676 and 108

```
RV---KAG---L---K-----V--E-----KDPML-----LA---
1 KVQHGPKAGADPSLLFIDKGGPGAAMVMGESIGFRMKRPALNAVMITHRGILAQPD
-----V-R-EG-D-----D-----A--A-L-----
59 PYASRIGSFVNRKEGGLDLTLRLLKDSGEVFEANEILAVRLDVKAFDDFTITEQIP
---LMGQ-----SG-----IGX-L-K---K---PE-----L-
117 DDALPGTVPAAAAKVEEKSAEVLAEVQRMDIRNYLAKFPNVKHGLLPQQIGDWTLE
-N-A-----V-----V-----VV-----YR
175 TNLFAEARETLKPLVPKLYTMLVRFLNIGMSCIAQLDADRGEQKAVVWPAQEDYR
N-L--L--E-----E-----P-----K-----
233 NALDALAMIERVAKGFERSEWAEGIVEAADTFTKYLQPDALESALVGDFRKNIFGA
-----YY-----K-
291 NRSALNVVKNVIDANVRQVFDELNLDIDDIRSSLKATYYYRLSDADFHNLWTSSAKI
---K-----L-----G--A-----V-----L--TR
```

```

349 FTGRSKSMKAGNVTVYGHVFLNSPKRFNSGELMAPWFLSHFYVIDKGIFHYLEATS
V---ED---G---L-----D--V--KVV-----Y-----
407 DKKWYEDFSVSDGRKDCLNKFSGMYGIPADLWVYFYKGPANPIEFGFYPADRSIDW
--L-----KDAN-----P-----H--DK-V-----
465 QQLGSEFWEQMNAVQEQLAGSRTWAQLMESFSPLDFFFHESDRMVPTAGSVVSKP
EII-----K-----G-V-R--L-----D-L-----K---
523 EILETPSYTAGCVECNDGYQDPSKCKPCTGKVFRDPLFMGKEPDYLQSITRNKIFG
-----AFD-----
581 NEKLRSYILESLQRNEESHTSHYNDYSINFGAFDTQHEQSMEGIMQEPTIGLQQAK
-----D---FKV---Y-----M-----
639 LMIPTGHADDACIFNVEHGRMRQYRVVDAQIHELMHGLHISGNAYPLACTVLIKK
-----
697 AVQT

```

Matches with 22 **K335Q** : 334 K \mapsto 49 K

10 Alignment of Homo sapiens's Mutations

align to reference sequence:

Compare sequences of length 24 and 900

```

KKKKGKPAEPPKGEAVALQKLDLLKAVEAAVENKDAKQAKLeRVINGQKTVEDML
1 -----K???????????????K??K??K??R-----
AQIQQPKATTVTEVVAPKPSTKAQGGGFRQRLSEIQDNELKQFLPSVTGIQHGAPL
59 -----
TCLFNTLLISCAPPLQLQAITASVTPMYPQLMVSLLAINVALGTVTGARQRDA
117 -----
ESGKIRKWPENVQIYQNGHRSITLISRLADRIRVKELLQHYHQLELTVHALLRQDD
175 -----
PTLVMEPVYGGFFKSVFMGARNIFNGLNNLESNNKLLLDTWSFASDQGEPRIYLL
233 -----
YFRWIDAPIGTDQAMDGFVGRSKSFKGDEYNLYETAILHSVLTYNDEAGLASCP
291 -----
FVLSHFPVNDKAMFQYLDVQEPNKWREWQDTYNATISLYGITADFWVYFVKDEFG
349 -----
ELPVPTGWKLDRTICRPKLGDRLWSRTIFQANPTWDSGPLTRGLWEELRKELKPLD
407 -----
LFLHQSSQVPCSRCVKCQPKLEVANILKGCKDCQDGRAEEYGCFPCVGEVFRDA
465 -----
LFRACHECRLQEVTDQLVFGRKLLQQFIDQTIKTQQPTTRGFIDFSINFWRYIDA
523 -----
HIIHYKDCIEQPTLGEELAKTETATGYEDTGCLYLTNWQRLRSYRAFVDASLVCGI
581 -----
INGLHPVNNVYPLASTILVNREGAVPLVPNQQPRLPPLSELGKEWATVAMAIEES
639 -----
LTALEEEEEPENTVARGEAPSPQQPQKQLYPRLALVGQQKLVTEAARQCPEQTSLTQF
697 -----
WSHLASLEEPLYADQLLPYLAGWLVIDALSETEGALFPCNQRSLSHDIHTLARRV

```

```

755 -----
SGLVDEGKKGQVVLYLAASLAPQLETAEWELWQNTLDDQEWGSLFFYRCIASTS
813 -----
FLYNGSDLQLVPVKPRTLFPVVCDEPGVTSILVEARGRARGAAALVPLCGPVGDSV
871 -----
FLRM
929 -----

```

minimal editing distance: 30 Alignment: -1

10.1 Align To Sequence In 1f4l.pdb

Compare sequences of length 900 and 549

```

1 EKS----AE-----V-L--_-----A-E--VQRMDIRNYLAKFPNVKHGL-----
KKKKGKPAEPPKGEAVALQKKLDLLKAVEAAVENKDAKQ--AKLERVINGQKTVED
-LPQ-QIGDWTLETNLFAE-----A-----R-----ET-----LKP-L--V-----
59 MLAQIQPKATTVTEVVAPKPKSTKAQGGGFRQRLSEIQDNELKQFLPSVTGIQHGA
PKLYTMLVRF--LNIGMSCIA----QLDADRGEQK-AVVWPA--QEDVYRNALD--
117 P-L-TCL--FNTLLI--SC-APPPLQLQA----QITASVTPMYPQLMVSLLAAINV
AL-----AMIERVAK-GFERSEWAE-----G---IVEA---ADTFT-KYL-QPD
175 ALGTVTGARQ-RDAESGKIRK-WPENVQIYQNGHRSITLISRLADRIRVKELLQHY
A-LES---AL-----V-----GDFRKNIF-GANRSALNVVKNVIDANVRQVFD
233 HQLELTVHALLRQDDPTLVMEPVYGGFFKSVFMGA-RNIFNGL-N----NLL----
ELN--LDIDDIRS--SLKATY--Y--Y-RLSDADFHNLWTS-AKI-FTG--RSKSM
291 ESNNKLLLDDTW-SFASDQGEPRIYLLYFRWIDAPIG---TDQAMDGFVGVGRSKSF
KAG---NVTVYGHVFLNSPKRFNSGE--LMA--PWF-LSHFYVIDKGIFHYL---E
349 K-GDEYNL--YETALHSVLTYND-EAGL-ASCP-FVLSHFVNDKAMFQYLDVQE
ATSDKKWYEDFSVSDGRKDCLNKF-SGMYGIPADLWVYFYK---G--PANPIEFGF
407 PN--KWWREWQ---DTY----NATISL-YGITADFWVYFVKDEFGELPV-PT--GW
YPADRSIDWQQLG----SE--FWEQMKNAVQEQLAGSRT---WAQLM-ESFSPLDF
465 KL-DRTICRPKLGDRLWSRTIF--QA-NPTWDS--GPLTRGLWEELRKELK-PLDL
FFHESDRMVPTAGSV-VSK--PEILETPSYTAGCVECNDGYQDPSKCKPCTGKVFR
523 FLHQSSQVVPC--SRCV-KCQPKKLEVANILKGCKDCQDGRAEEYGCPCVGEVFR
DPLFMGK-EP-----DYL-----QS-I--T-----RNKI-FG-NEKLRS
581 DALFRACHECRLQEVTDQLVFGRKLLQQFIDQTIKTQQPTTTRGFIDFSINFW-RY
-----Y---I---L-ESLQRNE-----ESH-----T-----S-----
639 IDAHIIHYKDCIEQPTLGEELAKTETATGYEDTGCLYLTNWQRLRSYRAFVDASLV
-----HY-ND-Y---S-I--NF-GAFDT---QHEQ-----SME-GI-----M
697 CGIINGLHPVNNVYPLASTILVNREGAVPLVPNQ--QPRLPPLS-ELGKEWATVAM
Q-E-----P--TI--GL---Q-Q-----A-----KLMI-----P--
755 AIEEESLTALEEEEEPENTVARGEAPSPQPQKQLYPRLALVGQQKLVTEAARQCPEQ
TG-----H-AD-----D-----AC---IFNV-EH-G-----RM-----
813 TSLTQFWSHLASLEEPLYAPDQLLPYLAGWLVIDALSETEGALFPCNQRSLSHDIH
---RQYRVW--VD-----A-----QIH--E--LMH-----G-LH--
871 TLAR--RVSGLVDEGKKGQVVLYLAASLAPQLETAEWELWQNTLDDQEWGSLFF
---I-SGNA-Y-----P-----LA---CT---VL-IKK-AXXX-----
929 YRCIASTSFLYNGSDLQLVPVKPRTLFPVVCDEPGVTSILVEARGRARGAAALVPL
-----
987 CGPVGDSVFLRM

```

Matches with 35 **K860A** : 859 K \mapsto 528 K
K863A : 862 K \mapsto 533 R
K866A : 865 K \mapsto 536 M
K880A : 879 K \mapsto 0 -
R857A : 856 R \mapsto 525 N

10.2 Align To Sequence In 1mkh.pdb

Compare sequences of length 900 and 108

```

R-----V--K--A-----_--_--_--G-----L
1  KKKKGKPAEPPKGEAVLQKKLDLLKAVEAAVENKDAKQAKLERVINGQKTVEDML
-----K--V-E--K-----D-----P-----
59  AQIQQPKATTVEVVAPKPSTKAQGGGFRQRLSEIQDNELKQFLPSVTGIQHGAPL
-----M--L--L-A--V-----R-----
117  TCLFNTLLISCAPPPLQLQAQITASVTPMYPQLMVSLLAAINVALGTVTGARQRDA
E-GDD-----A-----ALL-----
175  ESGKIRKWPENVQIYQNGHRSITLISRLADRIRVKELLQHYHQLELTVHALLRQDD
---M--GQ--S--G--I--GX--LK--K-----P-----
233  PTLVMEPVYGGFFKSVFMGARNIFNGLNLLLESNNKLLLDTWSFASDQGEPRIYLL
-----E-L-----N-A-----
291  YFRWIDAPIGTDQAMDGFVGVGRSKSFKGDEYNLYETAILHSVLTYNDEAGLASCP
-V--V--V-----VY-----
349  FVLSHFVNDKAMFQYLDVQEPNKWWREWQDTYNATISLYGITADFWVYFVKDEFG
-----R-----N-----L--L-EE-----P--
407  ELPVPTGWKLDRTICRPKLGDRLWSRTIFQANPTWDSGPLTRGLWEELRKELKPLD
-----KYY-KKL-----G-----A-----V-----
465  LFLHQSSQVPCSRVCVKCQPKKLEVANILKGCKDCQDGRAEEYGCFCVGEVFRDA
LTR-----V-----
523  LFRACHECRLQEVTDLVFGRLKLLQQFIDQTIKTQQPTTTRGFIDFSINFWRYIDA
-----E-----D-G-L-----D--VK--
581  HI IHYKDCIEQPTLGEELAKTETATGYEDTGCLYLTNWQRLRSYRAFVDASLVCGI
-----V--VY-L-----KD-A-----
639  INGLHPVNNVYPLASTILVNREGAVPLVPNQPRLPPLSELGKEWATVAMAIEEES
-----N-----PHD--K-----V-----E-----
697  LTALEEEEPENTVARGEAPSPQPQKQLYPRALVGQQKLVTEAARQCPEQTSLTQF
-----I-----I-----
755  WSHLASLEEPLYAPDQLLPYLAGWLVIDALSETEGALFPCNQRSLSHDIHTLARRV
-----K-G-V-----R-----
813  SGLVDEGKKGQVVLYYLAASLAPQLETAEWELWQNTLDDQEWGSLFFYRCIASTS
-L---DL---KA--F---D-----DFK
871  FLYNGSDLQLVPVKPRTLFPVVCDEPGVTSILVEARGRARGAAALVPLCGPVGDS-
VY--M-
929  VFLRM-

```

Matches with 17 **K860A** : 859 K \mapsto 0 -
K863A : 862 K \mapsto 0 -
K866A : 865 K \mapsto 0 -
K880A : 879 K \mapsto 0 -

R857A : 856 R → 0 -

Compare sequences of length 900 and 108

```
R-----V---K-----A-----_--_--_--_--G-----L
1 KKKKGKPAEPPKGEAVALQKKLDLLKAVEAAVENKDAKQAKLERVINGQKTVEDML
-----K---V-E---K-----D-----P-----
59 AQIQQPKATTVTEVVAPKPSTKAQGGGFRQRLSEIQDNELKQFLPSVTGIQHGAPL
-----M---L---L-A---V-----R-----
117 TCLFNTLLISCAPPLQLQAQITASVTPMYPQLMVSLLAAIVALGTVTGARQRDA
E-GDD-----A-----ALL-----
175 ESGKIRKWPENVQIYQNHRSITLISRLADRIRVKELLQHYHQLELTVHALLRQDD
---M---GQ---S---G---I---GX---LK---K-----P-----
233 PTLVMEPVYGGFFKSVFMGARNIFNGLNNLESNNKLLLDTWSFASDQEPRIYLL
-----E---L-----N---A-----
291 YFRWIDAPIGTDQAMDGFVGVRSKSFKGDEYNLYETAILHSVLTYNDEAGLASCP
-V-----V-----V-----VY-----
349 FVLSHFPVNDKAMFQYLDVQEPNKWWREWQDTYNATISLYGITADFWVYFVKDEFG
-----R-----N-----L---L-EE-----P--
407 ELPVPTGWKLDRTICRPKLGDRLWSRTIFQANPTWDSGPLTRGLWEELRKELKPLD
-----KYY-KKL-----G-----A-----V-----
465 LFLHQSSQVVPCSRCVKCQPKKLEVANILKGCKDCDGRAEEYGCFPCVGEVFRDA
LTR-----V-----
523 LFRACHECRLQEVTDQLVFGRKLLQQFIDQTIKTQQPTTTRGFIDFSINFWRYIDA
-----E-----D-G-L-----D---VK--
581 HIIHYKDCIEQPTLGEELAKTETATGYEDTGCLYLTNWQRLRSYRAFVDASLVCGI
-----V---VY-L-----KD-A-----
639 INGLHPVNNVYPLASTILVNREGAVPLVPNQQPRLPPLSELGKEWATVAMAIEES
-----N-----PHD---K-----V-----E-----
697 LTALEEEEEPENTVARGEAPSPQPQKQLYPRLALVGQQKLVTEAARQCPEQTSLTQF
-----I-----I-----
755 WSHLASLEEPLYAPDQLLPYLAGWLVIDALSETEGALFPCNQRSLSHDIHTLARRV
-----K-G-V-----R-----
813 SGLVDEGKKGQVVLYYLAASLAPQLETAEWELWQNTLDDQEWGSLLFFYRCIASTS
-L---DL---KA---F---D-----DFK
871 FLYNGSDLQLVPVKPRTLFPVVCDEPGVTSILVEARGRAGAALVPLCGPVGDS-
VY--M-
929 VFLRM-
```

Matches with 17 K860A : 859 K → 0 -

K863A : 862 K → 0 -

K866A : 865 K → 0 -

K880A : 879 K → 0 -

R857A : 856 R → 0 -