

SUPPLEMENTARY INFORMATION

The structure of a thermophilic kinase shapes fitness upon random circular permutation

Alicia M. Jones¹, Manan M. Mehta², Emily E. Thomas¹, Joshua T. Atkinson³, Thomas H. Segall-Shapiro⁴, Shirley Liu¹, and Jonathan J. Silberg^{1,*}

Author affiliations:

1. Department of Biosciences, Rice University, MS-140, 6100 Main Street, Houston, TX, 77005
2. Medical Scientist Training Program, Northwestern University, 303 East Chicago Avenue, Morton 1-670, Chicago, IL, 60611
3. Systems, Synthetic, and Physical Biology Graduate Program, Rice University, 6100 Main MS-180, Houston, Texas 77005
4. Department of Biological Engineering, Synthetic Biology Center, Massachusetts Institute of Technology, 500 Technology Square, NE47-257, Cambridge, MA, 02139

*To whom correspondence should be addressed. Jonathan J. Silberg, Biosciences Department, 6100 Main Street GBH W200-I, Houston TX 77005; Tel: 713-348-3849; Fax: 713-348-5154; Email: joff@rice.edu

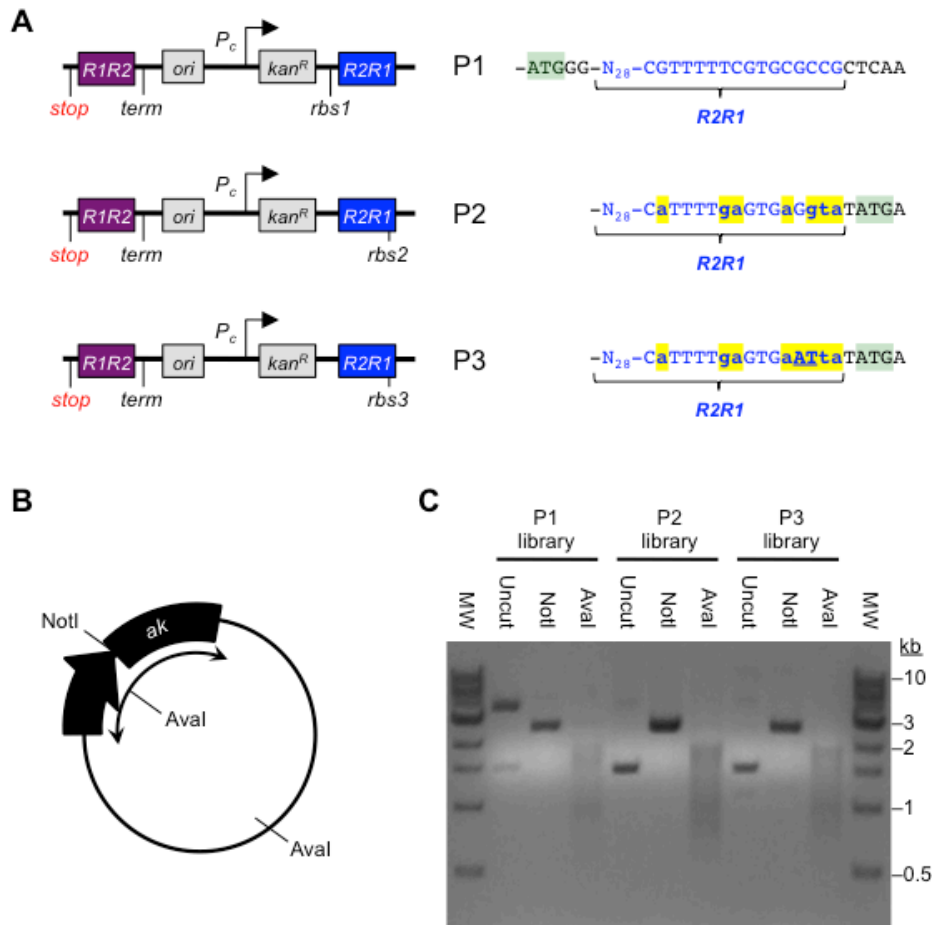


Figure S1. Permuteposons used to create libraries. **(A)** The architecture of each permuteposon. The start codon designed to initiate translation of open reading frames encoding permuted proteins is shown in green, the transposase recognition sequences (R1R2 and R2R1) are shown in purple and blue, and the mutations introduced into R2R1 are shown in yellow. Mutations in P3 that differ from P2 are underlined. **(B)** The different vectors created by PERMUTE contain a single NotI site within the sequence that connects the permuted *TnAK* gene fragments and two Aval sites. One of the Aval sites is within the permuteposon, while the other is within the permuted *TnAK* genes. **(C)** When treated with NotI, all three libraries yield a single band at the expected molecular weight of linear permuteposons fused to the *TnAK* gene (2.5 kb). When treated with Aval, the three libraries yield smears whose length varies from 600 to 1900 base pairs. This smear is expected to occur when libraries contain many of the possible variants because the distance between the two Aval sites within different vectors depends on the location where the *TnAK* gene is permuted.

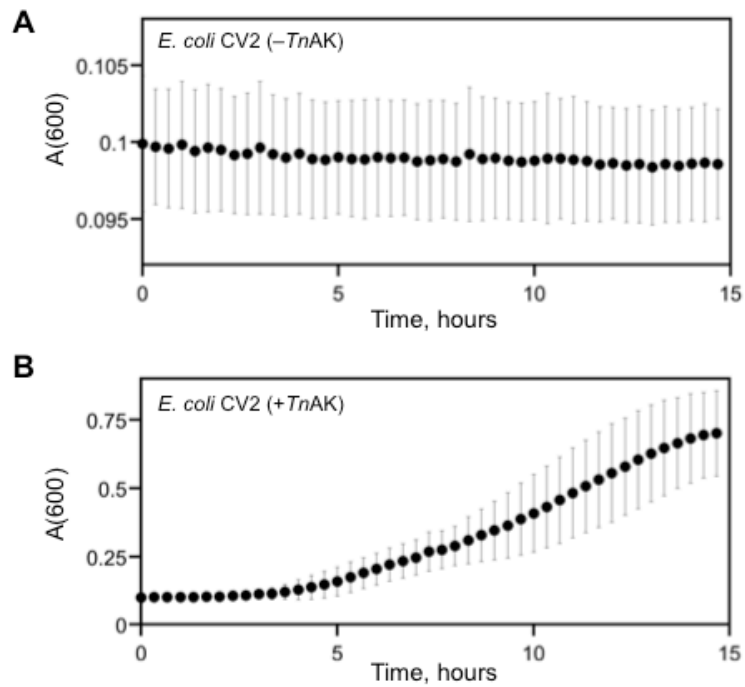


Figure S2. Assessing AK function using bacterial complementation. The absorbance at 600 nm of *E. coli* CV2 cells transformed with (A) circularized P1 is compared with cells transformed with (B) P1 that expresses *TnAK*. Cell growth was performed at 42°C, and data are reported as the mean $\pm 1\sigma$ ($n \geq 28$).



Figure S3. Complementation strength of variants selected from each library. For the (A) P1, (B) P2, and (C) P3 libraries, the absorbance at 600 nm of *E. coli* CV2 cultures after 15 hours of growth at 42°C is reported for cells transformed with circularized P1 (-*ctrl*), P1 that expresses *TnAK* (+*TnAK*), and P1, P2, and P3 permuteposons expressing permuted proteins selected from each library. Permuted AK are named based on the *TnAK* codon that occurs at the beginning of each permuted gene, and names are colored red (AMP binding), blue (core), and green (lid) based on the domain location of the residue at their N-terminus. Data are reported as the mean $\pm 1\sigma$ ($n \geq 4$). A(600) represents the absorbance obtained at 600 nm after 15 hours of growth. All variants displayed significant growth compared to the negative control (t-test; $p < 0.05$) except 144 in the P1 library and 44 and 47 in the P3 library. Because the growth conditions (42°C for 15 hours) were more stringent than the original selections (40°C for 48 hours), we performed measurements with these three variants at 40°C. Analysis of *E. coli* CV2 complementation by these three variants at 40°C revealed significant growth after 15 hours ($p < 0.001$).

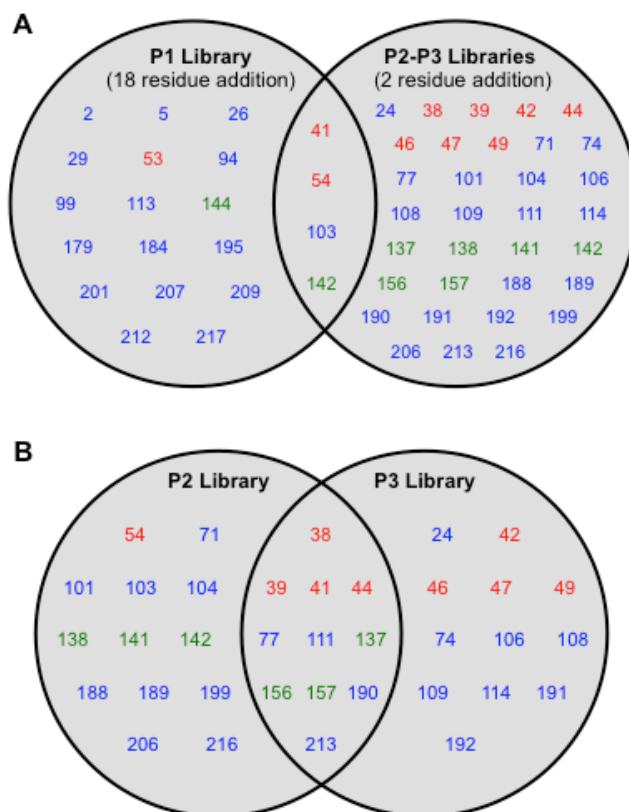


Figure S4. Comparison of permuted *TnAK* discovered in each library. **(A)** Comparison of the functional permuted *TnAK* selected from libraries that amend small versus large peptides to the N-terminus of permuted proteins. **(B)** A comparison of variants selected from the P2 and P3 libraries, which express identical variants using different RBSs. Variants are named based on the *TnAK* codon that occurs at the beginning of each permuted gene, and names are colored red (AMP binding), blue (core), and green (lid) based on the domain location of the residue at their N-terminus.

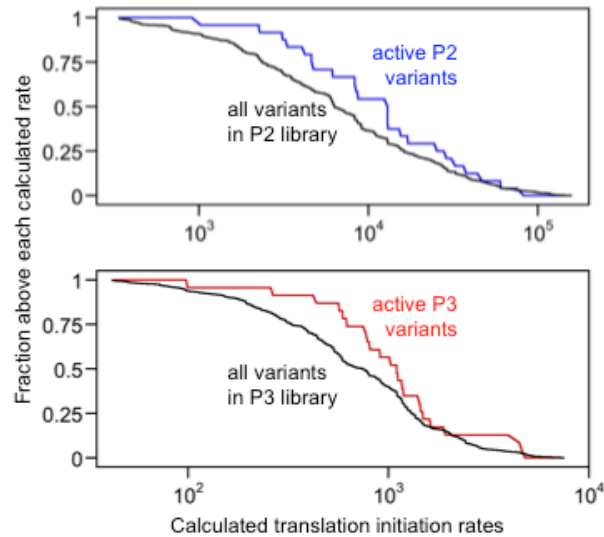


Figure S5. Translation initiation rates calculated using intended and alternative start codons. The relative expression of each permuted *TnAK* was estimated by summing the translation initiation rates at the intended start codon and alternative in frame start codons. Previous studies have shown that *TnAK* retains activity upon removal of more than twenty residues at its termini, suggesting that some permuted *TnAK* variants can be expressed as truncations from alternative translation initiation sites without loss of protein function. The fraction of active P2 and P3 variants above each calculated threshold for translation initiation was compared with the fraction of total possible variants to visualize the enrichment of active permuted proteins over that expected in the absence of enrichment. Application of a one-tailed Mann-Whitney-Wilcoxon test yielded a >98% probability that the median expression for active permuted *TnAK* in the P2 library (and P3 library) is greater than the median expression for all theoretically possible variants in each library.

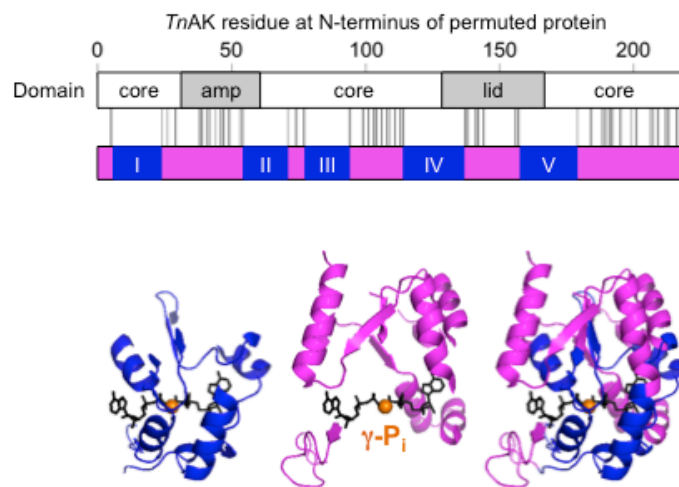


Figure S6. Mapping peptides I-V onto AK structure. The AK structure shows the proximity of residues within peptides I-V (blue) to the γ -P_i (orange) within Ap5A, an inhibitor that has been used to model the catalytic transition state.

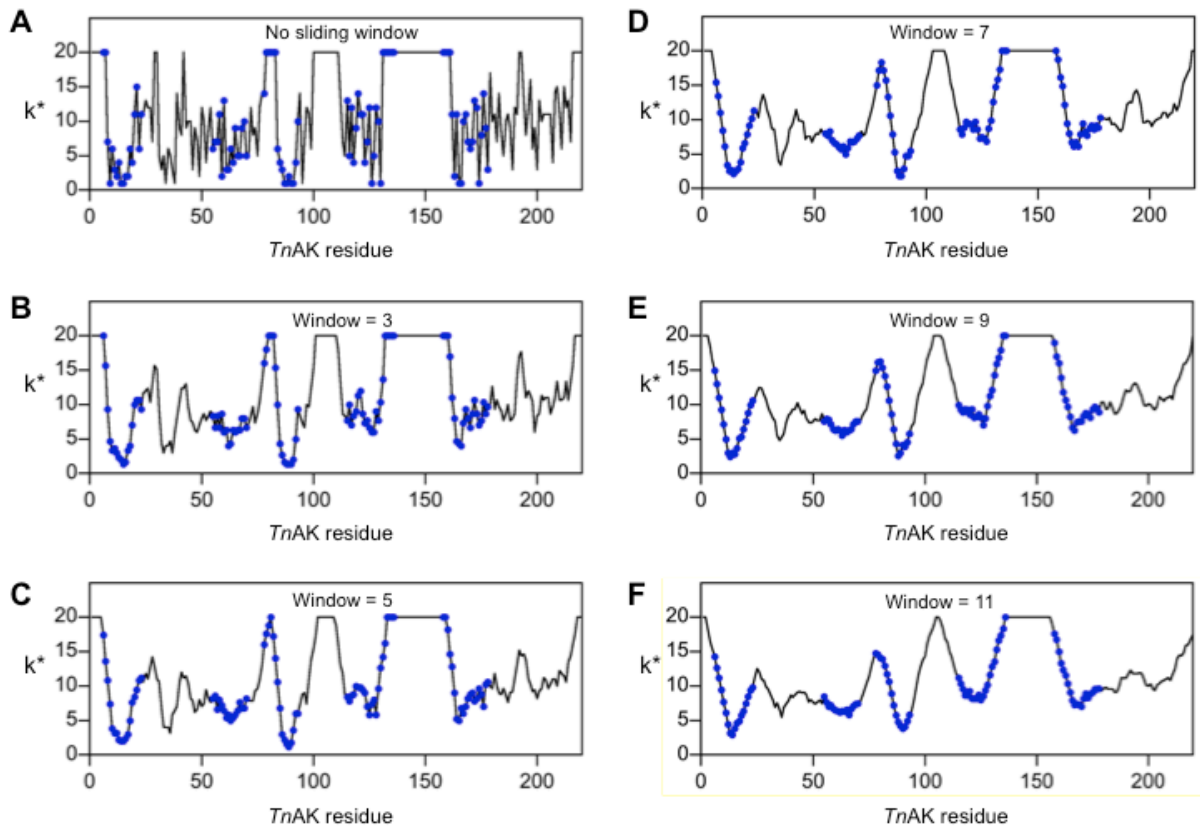


Figure S7. Mutational tolerance (k^*) calculated using a multiple sequence alignment. k^* is shown for sliding windows of (A) 1, (B) 3, (C) 5, (D) 7, (E) 9, and (F) 11 residues. When windows were applied, k^* for a residue site was calculated as the mean k^* from a window of residues centered on each native position. The *TnAK* residues that correspond to peptides I-V, which lacked backbone fission sites in combinatorial experiments, are mapped in blue on each profile.

AK with Ap5A structures

Organism	PDB	NH2	COOH
<i>Geobacillus stearothermophilus</i>	1ZIN	1	217
<i>Bacillus subtilis</i>	1P3J	1	212
<i>Bacillus globisporus</i>	1S3G	1	217
<i>Escherichia coli</i>	1AKE	1	214
<i>Saccharomyces cerevisiae</i>	1AKY	3	220
<i>Streptococcus pneumoniae</i>	4NU0	1	212
<i>Aquifex aeolicus</i>	2RGX	1	203
<i>Marinabacillus marinus</i>	3FB4	1	215
<i>Zea mays</i>	1ZAK	3	222
<i>Homo sapiens</i>	2C95	0	195

Pairwise sequence identity

TnAK	1ZIN	1P3J	1S3G	1AKE	1AKY	4NU0	2RGX	3FB4	1ZAK	2C95	
46%	1ZIN	---	75%	68%	45%	42%	53%	45%	73%	40%	32%
50%	1P3J	---	---	69%	47%	43%	57%	45%	71%	39%	32%
43%	1S3G	---	---	---	50%	44%	53%	46%	74%	38%	31%
43%	1AKE	---	---	---	---	46%	39%	44%	48%	42%	35%
41%	1AKY	---	---	---	---	---	41%	41%	43%	36%	30%
39%	4NU0	---	---	---	---	---	---	39%	55%	38%	29%
47%	2RGX	---	---	---	---	---	---	---	43%	37%	34%
48%	3FB4	---	---	---	---	---	---	---	---	39%	30%
38%	1ZAK	---	---	---	---	---	---	---	---	---	37%
35%	2C95	---	---	---	---	---	---	---	---	---	---

Figure S8. Sequence variability for AK orthologs used in RMSD calculations. The organism, PDB code, and residues with defined coordinates are shown for the ten Ap5A-bound AK structures used in calculations. These include AK from bacteria, fungi, plants, and animals. The pairwise sequence identities are shown for all ten AK as well as the sequence identities between *TnAK* and each ortholog having a crystal structure. This analysis yielded four pairs of sequences with identities between 70-79% (yellow), two pairs of sequences with identities between 60-69% (green), five pairs of sequences with identities between 50-59% (orange), seventeen pairs of sequences with identities between 40-49% (gray), sixteen pairs of sequences with identities between 30-39% (magenta), and one sequence pair with identity below 30% (blue).

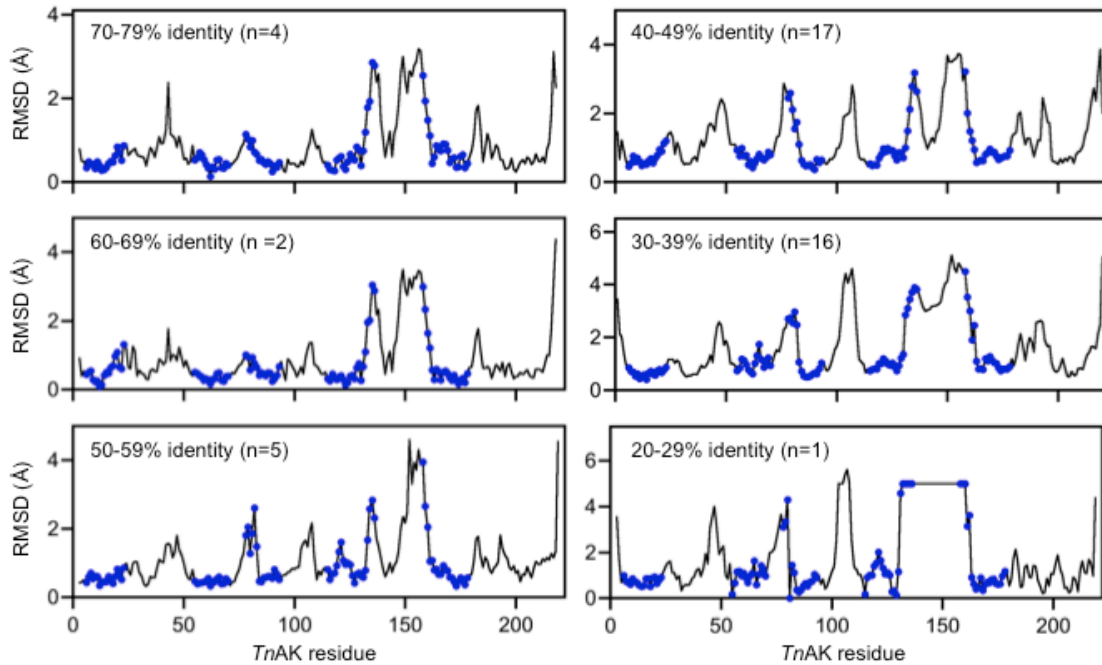


Figure S9. Positional RMSD calculated using AK structures of varying identity. The positional RMSD is shown for AK ortholog pairs having sequences with 70-79%, 60-69%, 50-59%, 40-49%, 30-39%, and 20-29% identity. The *TnAK* residues that correspond to peptides I-V, which lacked backbone fission sites in laboratory evolution experiments, are mapped in blue on each profile.

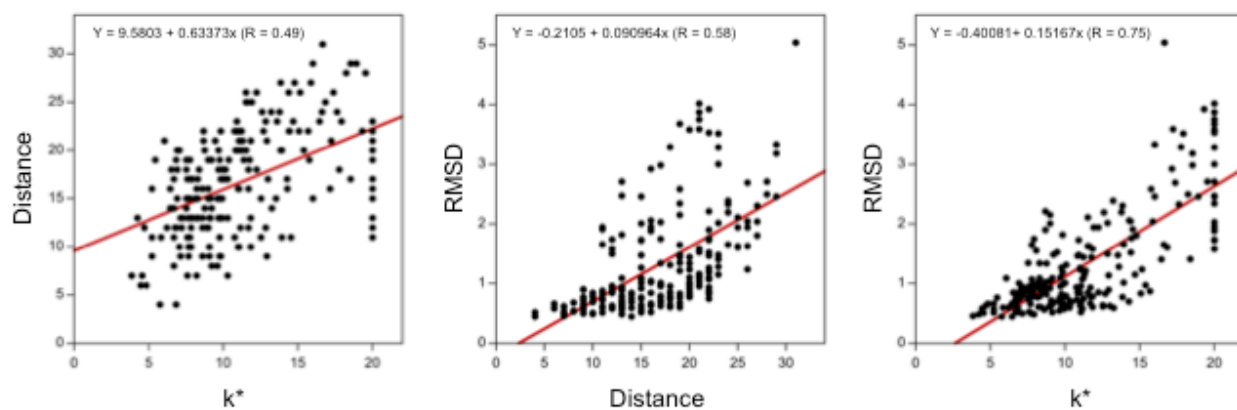


Figure S10. Comparison of distance, k^* , and RMSD for each possible variant. For each possible *TnAK* in our library, we compared the pairwise relationships between each structure-based metric. As expected, k^* and RMSD displayed the strongest correlation, while distance and k^* showed the weakest correlation.

Supplementary Sequences. Sequences of permuted *TnAK* encoded by the open reading frames in permuteposon vectors selected from the P1, P2, and P3 libraries, as well as permuted *TnAK* encoded by rationally designed pET vectors. Variants are named by the *TnAK* residue that is translated first. Peptides amended to the N-terminus are colored red. Blue residues are duplicated by the transposase reaction. One extra residue is added to the C-terminus in P1 variants (green), while P2 and P3 variants lack an extra residue at their C-terminus. The peptide linker AAA connecting the original termini is shown in bold.

P1 LIBRARY

>P1 #2, Residues 2-220/1-2

MGFRIYRETL**SRFS****CAAQ****M**AYLVFLGPPGAGKGTYAKRIQEK**T**GIPHISTGDIFRDIVK**K**ENDELGKK**I**KEIMEK**G**ELVLPDELVNEVVKRRLSEKDCEKGFILDGYPR**T**V**A**Q**A**EFLDSFLESQ**N**Q**L**TA**A**VLFDVPEDVVVQ**R**LT**S**RRIC**P**CKGRIY**N**MISLPPKEDEL**C**DDCKV**K**LVQ**R**DD**D**KEETV**R**HRYKVY**L**E**K**T**Q**PVIDY**G**KK**G**ILK**R**VDGTIGIDNV**V**AEVL**K**IIGWSD**K****AAAM****M**A

>P1 #5, Residues 5-220/1-5

MGFRIYRETL**SRFS****CAAQ****L**VFLGPPGAGKGTYAKRIQEK**T**GIPHISTGDIFRDIVK**K**ENDELGKK**I**KEIMEK**G**ELVLPDELVNEVVKRRLSEKDCEKGFILDGYPR**T**V**A**Q**A**EFLDSFLESQ**N**Q**L**TA**A**VLFDVPEDVVVQ**R**LT**S**RRIC**P**CKGRIY**N**MISLPPKEDEL**C**DDCKV**K**LVQ**R**DD**D**KEETV**R**HRYKVY**L**E**K**T**Q**PVIDY**G**KK**G**ILK**R**VDGTIGIDNV**V**AEVL**K**IIGWSD**K****AAAM****M**AY**L****V**

>P1 #26, Residues 26-220/1-26

MGFRIYRETL**SRFS****CAAQ****T**GIPHISTGDIFRDIVK**K**ENDELGKK**I**KEIMEK**G**ELVLPDELVNEVVKRRLSEKDCEKGFILDGYPR**T**V**A**Q**A**EFLDSFLESQ**N**Q**L**TA**A**VLFDVPEDVVVQ**R**LT**S**RRIC**P**CKGRIY**N**MISLPPKEDEL**C**DDCKV**K**LVQ**R**DD**D**KEETV**R**HRYKVY**L**E**K**T**Q**PVIDY**G**KK**G**ILK**R**VDGTIGIDNV**V**AEVL**K**IIGWSD**K****AAAM****M**AYLVFLGPPGAGKGTYAKRIQEK**T****G**

>P1 #29, Residues 29-220/1-29

MGFRIYRETL**SRFS****CAAQ****P**HISTGDIFRDIVK**K**ENDELGKK**I**KEIMEK**G**ELVLPDELVNEVVKRRLSEKDCEKGFILDGYPR**T**V**A**Q**A**EFLDSFLESQ**N**Q**L**TA**A**VLFDVPEDVVVQ**R**LT**S**RRIC**P**CKGRIY**N**MISLPPKEDEL**C**DDCKV**K**LVQ**R**DD**D**KEETV**R**HRYKVY**L**E**K**T**Q**PVIDY**G**KK**G**ILK**R**VDGTIGIDNV**V**AEVL**K**IIGWSD**K****AAAM****M**AYLVFLGPPGAGKGTYAKRIQEK**T****G****I****P****H**

>P1 #41, Residues 41-220/1-41 (also found in P2 and P3 libraries)

MGFRIYRETL**SRFS****CAAQ****V**KKENDELGKK**I**KEIMEK**G**ELVLPDELVNEVVKRRLSEKDCEKGFILDGYPR**T**V**A**Q**A**EFLDSFLESQ**N**Q**L**TA**A**VLFDVPEDVVVQ**R**LT**S**RRIC**P**CKGRIY**N**MISLPPKEDEL**C**DDCKV**K**LVQ**R**DD**D**KEETV**R**HRYKVY**L**E**K**T**Q**PVIDY**G**KK**G**ILK**R**VDGTIGIDNV**V**AEVL**K**IIGWSD**K****AAAM****M**AYLVFLGPPGAGKGTYAKRIQEK**T****G****I****P****H**ISTGDIFRDIV**N**

>P1 #53, Residues 53-220/1-53

MGFRIYRETL**SRFS****CAAQ****K**EIMEK**G**ELVLPDELVNEVVKRRLSEKDCEKGFILDGYPR**T**V**A**Q**A**EFLDSFLESQ**N**Q**L**TA**A**VLFDVPEDVVVQ**R**LT**S**RRIC**P**CKGRIY**N**MISLPPKEDEL**C**DDCKV**K**LVQ**R**DD**D**KEETV**R**HRYKVY**L**E**K**T**Q**PVIDY**G**KK**G**ILK**R**VDGTIGIDNV**V**AEVL**K**IIGWSD**K****AAAM****M**AYLVFLPPGAGKGTYAKRIQEK**T****G****I****P****H**ISTGDIFRDIVK**K**ENDELGKK**I****K****D**

>P1 #54, Residues 54-220/1-54 (also found in P2 library)

MGFRIYRETL**SRFS****CAAQ****E**IMEK**G**ELVLPDELVNEVVKRRLSEKDCEKGFILDGYPR**T**V**A**Q**A**EFLDSFLESQ**N**Q**L**TA**A**VLFDVPEDVVVQ**R**LT**S**RRIC**P**CKGRIY**N**MISLPPKEDEL**C**DDCKV**K**LVQ**R**DD**D**KEETV**R**HRYKVY**L**E**K**T**Q**PVIDY**G**KK**G**ILK**R**VDGTIGIDNV**V**AEVL**K**IIGWSD**K****AAAM****M**AYLVFLGPPGAGKGTYAKRIQEK**T****G****I****P****H**ISTGDIFRDIVK**K**ENDELGKK**I****K****E****I**

>P1 #94, Residues 94-220/1-94

MGFRIYRETL~~SRFS~~CAAQ~~LA~~AEFLDSFLESQNKQLTAAVLF~~FDV~~PE~~DD~~VVVQRLTSRRICPKCGRINMISLPPKEDELCD
DDCKVKLVQRDDDDKEETV~~RR~~HRYKVYLEKTQPVIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDK~~AA~~AMMAYLVFL
GPPGAGKGT~~YAKRIQ~~EKTGIPHISTGDI~~FR~~DIVK~~KENDEL~~GKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFIL
LDGYPRTVA~~HA~~

>P1 #99, Residues 99-220/1-99

MGFRIYRETL~~SRFS~~CAAQ~~LD~~SFLESQNKQLTAAVLF~~FDV~~PE~~DD~~VVVQRLTSRRICPKCGRIYNMISLPPKEDELCD
DDCKVKLVQRDDDDKEETV~~RR~~HRYKVYLEKTQPVIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDK~~AA~~AMMAYLVFL
GPPGAGKGT~~YAKRIQ~~EKTGIPHISTGDI~~FR~~DIVK~~KENDEL~~GKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFIL
DGYPR~~VAQA~~EFLD

>P1 #103, Residues 103-220/1-103 (also found in P2 library)

MGFRIYRETL~~SRFS~~CAAQ~~LES~~QNKQLTAAVLF~~FDV~~PE~~DD~~VVVQRLTSRRICPKCGRIYNMISLPPKEDELCD
DDCKVKLVQRDDDDKEETV~~RR~~HRYKVYLEKTQPVIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDK~~AA~~AMMAYLVFL
GPPGAGKGT~~YAKRIQ~~EKTGIPHISTGDI~~FR~~DIVK~~KENDEL~~GKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFIL
DGYPR~~VAQA~~EFLDSFLD

>P1 #113, Residues 113-220/1-113

MGFRIYRETL~~SRFS~~CAAQ~~AVL~~FDVPE~~DD~~VVVQRLTSRRICPKCGRIYNMISLPPKEDELCD
DDCKVKLVQRDDDDKEETV~~RR~~HRYKVYLEKTQPVIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDK~~AA~~AMMAYLVFL
GPPGAGKGT~~YAKRIQ~~EKTGIPHISTGDI~~FR~~DIVK~~KENDEL~~GKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFIL
DGYPR~~VAQA~~EFLDSFLESQNKQLTAAV

>P1 #142, Residues 142-220/1-142 (also found in P2 library)

MGFRIYRETL~~SRFS~~CAAQ~~MIS~~LPPKEDELCD
DDCKVKLVQRDDDDKEETV~~RR~~HRYKVYLEKTQPVIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDK~~AA~~AMMAYLVFL
GPPGAGKGT~~YAKRIQ~~EKTGIPHISTGDI~~FR~~DIVK~~KENDEL~~GKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFIL
DGYPR~~VAQA~~EFLDSFLESQNKQLTAAVLF~~FDV~~PE~~DD~~VVVQRLTSRRICPKCGRIYNMI

>P1 #144, Residues 144-220/1-144

MGFRIYRETL~~SRFS~~CAAQ~~SL~~LPPKEDELCD
DDCKVKLVQRDDDDKEETV~~RR~~HRYKVYLEKTQPVIDYGGKILKRVDGTIGIDNVVAEVLKIIIGWSDK~~AA~~AMMAYLVFL
GPPGAGKGT~~YAKRIQ~~EKTGIPHISTGDI~~FR~~DIVK~~KENDEL~~GKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFIL
DGYPR~~VAQA~~EFLSFLDSFLESQNKQLTAAVLF~~FDV~~PE~~DD~~VVVQRLTSRRICPKCGRIYNMISL

>P1 #179, Residues 179-220/1-179

MGFRIYRETL~~SRFS~~CAAQ~~LE~~KTQPVIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDK~~AA~~AMMAYLVFL
GPPGAGKGT~~YAKRIQ~~EKTGIPHISTGDI~~FR~~DIVK~~KENDEL~~GKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFIL
DGYPR~~VAQA~~EFLDSFLESQNKQLTAAVLF~~FDV~~PE~~DD~~VVVQRLTSRRICPKCGRIYNMISLPPKEDELCD
DDCKVKLVQRDDDDKEETV~~RR~~HRYKVYLE

>P1 #184, Residues 184-220/1-184

MGFRIYRETL~~SRFS~~CAAQ~~PV~~VIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDK~~AA~~AMMAYLVFL
GPPGAGKGT~~YAKRIQ~~EKTGIPHISTGDI~~FR~~DIVK~~KENDEL~~GKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFIL
DGYPR~~VAQA~~EFLDSFLESQNKQLTAAVLF~~FDV~~PE~~DD~~VVVQRLTSRRICPKCGRIYNMISLPPKEDELCD
DDCKVKLVQRDDDDKEETV~~RR~~HRYKVYLEKTQP

>P1 #195, Residues 195-220/1-195

MGFRIYRETL~~SRFS~~CAAQ~~LK~~RV~~DD~~GTIGIDNVVAEVLKIIIGWSDK~~AA~~AMMAYLVFL
GPPGAGKGT~~YAKRIQ~~EKTGIPHISTGDI~~FR~~DIVK~~KENDEL~~GKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFIL
DGYPR~~VAQA~~EFLDSFLESQNKQLTAAVLF~~FDV~~PE~~DD~~VVVQRLTSRRICPKCGRIYNMISLPPKEDELCD
DDCKVKLVQRDDDDKEETV~~RR~~HRYKVYLEKTQP
VIDYYGKKGILN

>P1 #201, Residues 201-220/1-201

MGFRIYRETLSRFSCAAQTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI
FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVVAQAEFLDSFLESQNKQLTAA
VLFVDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKKEETVRHRYKVYLEKTQPVIDYYG
KKGILKRVDTI

>P1 #207, Residues 207-220/1-207

MGFRIYRETLSRFSCAAQVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI
FRDIVK
KENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVVAQAEFLDSFLESQNKQLTAAVLFVDP
EDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKKEETVRHRYKVYLEKTQPVIDYYGKKGILK
RVDGTIGIDNVV

>P1 #209, Residues 209-220/1-209

MGFRIYRETLSRFSCAAQVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI
FRDIVKKE
NDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVVAQAEFLDSFLESQNKQLTAAVLFVDPED
VVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKKEETVRHRYKVYLEKTQPVIDYYGKKGILKRV
DGTIGIDNVVAD

>P1 #212, Residues 212-220/1-212

MGFRIYRETLSRFSCAAQLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI
FRDIVK
KENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVVAQAEFLDSFLESQNKQLTAAVLFVDPEDVVV
QRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTI
GIDNVVAEVLN

>P1 #217, Residues 217-220/1-217

MGFRIYRETLSRFSCAAQWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI
FRDIVK
KENDELGKKI
KEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVVAQAEFLDSFLESQNKQLTAAVLFVDPEDVVVQRLTS
RRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDN
VVAEVLKIIIGWS

P2 LIBRARY

>P2 #38, Residues 38-220/1-38 (also found P3 library)

MIRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVVAQAEFLDSFLESQNKQLTA
AVLFVDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKKEETVRHRYKVYLEKTQPVIDYY
GKKGILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI
FR

>P2 #39, Residues 39-220/1-39 (also found P3 library)

MRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVVAQAEFLDSFLESQNKQLTAA
VLFVDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKKEETVRHRYKVYLEKTQPVIDYYG
KKGILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQKTGIPHISTGDI
FRD

>P2 #41, Residues 41-220/1-41 (also found in P1 and P3 libraries)

MIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVVAQAEFLDSFLESQNKQLTAAV
LFDVDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKKEETVRHRYKVYLEKTQPVIDYYGK
GILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI
FRDIV

>P2 #44, Residues 44-220/1-44 (also found P3 library)

MKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVVAQAEFLDSFLESQNKQLTAAVLFV
DPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKKEETVRHRYKVYLEKTQPVIDYYGKKGIL
KRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI
FRDIVKKE

>P2 #54, Residues 54-220/1-54 (also found P1 library)

MKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVVAQAEFLDSFLESQNKQLTAAVLFVDPEDVVVQRLT
SRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGID
NVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI
FRDIVK
KENDELGKKIKE

>P2 #71, Residues 71-220/1-71

MIVKRRRLSEKDCEKGFILDGYPRVVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDK**AA**AMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIRDIVKKENDELGKKIKEIMEKGELVPDELVNEV**V**

>P2 #77, Residues 77-220/1-77 (also found P3 library)

MTEKDCEKGFILDGYPRVVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDK**AA**AMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIRDIVKKENDELGKKIKEIMEKGELVPDELVNEV**VKRR**LSE

>P2 #101, Residues 101-220/1-101

MNSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDK**AA**AMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIRDIVKKENDELGKKIKEIMEKGELVPDELVNEV**VKRR**LSEKDCEKGFILDGYPRVVAQAEFLDS**S**

>P2 #103, Residues 103-220/1-103 (also found P1 library)

MILESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDK**AA**AMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIRDIVKKENDELGKKIKEIMEKGELVPDELVNEV**VKRR**LSEKDCEKGFILDGYPRVVAQAEFLDS**FL**

>P2 #104, Residues 104-220/1-104

MMESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDK**AA**AMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIRDIVKKENDELGKKIKEIMEKGELVPDELVNEV**VKRR**LSEKDCEKGFILDGYPRVVAQAEFLDS**FL**

>P2 #111, Residues 111-220/1-111 (also found P3 library)

MITAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDK**AA**AMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIRDIVKKENDELGKKIKEIMEKGELVPDEEV**VKRR**LSEKDCEKGFILDGYPRVVAQAEFLDSFLESQNKQL**T**

>P2 #137, Residues 137-220/1-137 (also found P3 library)

MSGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDK**AA**AMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIRDIVKKENDELGKKIKEIMEKGELVPDELVNEV**VKRR**LSEKDCEKGFILDGYPRVVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCG**R**

>P2 #138, Residues 138-220/1-138

MSRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDK**AA**AMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIRDIVKKENDELGKKIKEIMEKGELVPDELVNEV**VKRR**LSEKDCEKGFILDGYPRVVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCG**R**

>P2 #141, Residues 141-220/1-141

MNNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDK**AA**AMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIRDIVKKENDELGKKIKEIMEKGELVPDELVNEV**VKRR**LSEKDCEKGFILDGYPRVVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIY**N**

>P2 #142, Residues 142-220/1-142 (also found P1 library)

MNMMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDK**AA**AMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIRDIVKKENDELGKKIKEIMEKGELVPDELVNEV**VKRR**LSEKDCEKGFILDGYPRVVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIY**N**

>P2 #156, Residues 156-220/1-156 (also found P3 library)

MNCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDK**AA**AMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIRDIVKKENDELGKKIKEIMEKGELVPDELVNEV**VKRR**LSEKDCEKGFILDGYPRVVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCD**CC**

>P2 #157, Residues 157-220/1-157 (also found P3 library)
MSKVKLVRDDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIFRDIVKKNENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCK

>P2 #188, Residues 188-220/1-188
MNYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIFRDIVKKNENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDDKEETVRHRYKVYLEKTQPVIDY

>P2 #189, Residues 189-220/1-189
MNYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIFRDIVKKNENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDDKEETVRHRYKVYLEKTQPVIDY

>P2 #190, Residues 190-220/1-190 (also found P3 library)
MNGKKGILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIFRDIVKKNENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDDKEETVRHRYKVYLEKTQPVIDYYG

>P2 #199, Residues 199-220/1-199
MMDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIFRDIVKKNENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRV

>P2 #206, Residues 206-220/1-206
MNNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIFRDIVKKNENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDN

>P2 #213, Residues 213-220/1-213 (also found P3 library)
MIKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIFRDIVKKNENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLK

>P2 #216, Residues 216-220/1-216
MIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIFRDIVKKNENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKII

P3 LIBRARY

>P3 #24, Residues 24-220/1-24
MKEKTGIPHISTGDIFRDIVKKNENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQ

>P3 #38, Residues 38-220/1-38 (also found in P2 library)
MIRDIVKKNENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIFR

>P3 #39, Residues 39-220/1-39 (also found in P2 library)
MRDIVKKNENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIFR

>P3 #41, Residues 41-220/1-41 (also found in P1 and P2 libraries)
MIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVAQAEFLDSFLESQNKQLTAAVL
FDVPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKK
GILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIFRDIV

>P3 #42, Residues 42-220/1-42
MMKKNDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVAQAEFLDSFLESQNKQLTAAVL
DVPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKK
ILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIFRDIVK

>P3 #44, Residues 44-220/1-44 (also found in P2 library)
MKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVAQAEFLDSFLESQNKQLTAAVLFDV
PEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGIL
KRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIFRDIVKKE

>P3 #46, Residues 46-220/1-46
MNDDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVAQAEFLDSFLESQNKQLTAAVLFDVPE
DVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILK
VDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIFRDIVKKN

>P3 #47, Residues 47-220/1-47
MNELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVAQAEFLDSFLESQNKQLTAAVLFDVPE
VVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILK
DGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIFRDIVKKN

>P3 #49, Residues 49-220/1-49
MIGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVAQAEFLDSFLESQNKQLTAAVLFDVPE
VQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILK
TIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIFRDIVKKN

>P3 #74, Residues 74-220/1-74
MRRLSEKDCEKGFILDGYPRVAQAEFLDSFLESQNKQLTAAVLFDVPEDVVVQRLTSRRICPKCGRIYNMISLPP
EDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDKAAAM
AYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIFRDIVKKNDELGKKIKEIMEKGELVPDELVNEVVKR

>P3 #77, Residues 77-220/1-77 (also found in P2 library)
MTEKDCEKGFILDGYPRVAQAEFLDSFLESQNKQLTAAVLFDVPEDVVVQRLTSRRICPKCGRIYNMISLPPKE
LCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAY
LVFLGPPGAGKGTYAKRIQEKTGIPHISTGDIFRDIVKKNDELGKKIKEIMEKGELVPDELVNEVVKRRLSE

>P3 #106, Residues 106-220/1-106
MTQNKQLTAAVLFDVPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLE
KTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHIST
DIFRDIVKKNDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVAQAEFLDSFLESQ

>P3 #108, Residues 108-220/1-108
MNKQLTAAVLFDVPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKT
QPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTG
IFRDIVKKNDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVAQAEFLDSFLESQNK

>P3 #109, Residues 109-220/1-109
MKQLTAAVLFDVPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKT
PVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTG
RDIVKKNDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVAQAEFLDSFLESQNKQ

>P3 #111, Residues 111-220/1-111 (also found in P2 library)
MITAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPV
IDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDIFRD
IVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLT

>P3 #114, Residues 114-220/1-114
MTVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPV
IDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDIFRD
IVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAV

>P3 #137, Residues 137-220/1-137 (also found in P2 library)
MSGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPV
IDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDIFRD
IVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCG

>P3 #156, Residues 156-220/1-156 (also found in P2 library)
MNCVKLVQRDDDKEETVRHRYKVYLEKTQPV
IDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDIFRD
IVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCK

>P3 #157, Residues 157-220/1-157 (also found in P2 library)
MSKVKLVQRDDDKEETVRHRYKVYLEKTQPV
IDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDIFRD
IVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCK

>P3 #190, Residues 190-220/1-190 (also found in P2 library)
MNGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDIFRD
IVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPV
IDYYG

>P3 #191, Residues 191-220/1-191
MRKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDIFRD
IVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPV
IDYYG

>P3 #192, Residues 192-220/1-192
MKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDIFRD
IVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPV
IDYYG

>P3 #213, Residues 213-220/1-213 (also found in P2 library)
MIKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDIFRD
IVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPV
IDYYGKKGILKRVDGTIGIDNVVAEVLK

Rationally designed pET vectors (discovered in a single library)

>pET #26, Residues 26-220/1-26
MKTGIPHISTGDIFRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLD
SFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKV
YLEKTQPV
IDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEK

>pET #26, Residues 26-220/1-26
MGFRIYRETLSRFSCAAQTGIPHISTGDIFRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGF
ILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLV
QRDDDKEETVRHRYKVYLEKTQPV
IDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEK

>pET #44, Residues 44-220/1-44

MKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDV
PEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGIL
KRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDI FRDIVKKE

>pET #44, Residues 44-220/1-44

MGFRIYRETLSRFSCAAQENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSF
LESQNKQLTAAVLFDVDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYL
EKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHIST
GDIFRDIVKKEN

>pET #53, Residues 53-220/1-53

MIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDVDPEDVVVQRL
TSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGI
DNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIK

>pET #53, Residues 53-220/1-53

MGFRIYRETLSRFSCAAQKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLT
AAVLFDVDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDY
YGKKGILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDI FRDIVK
KENDELGKKIKD

>pET #111, Residues 111-220/1-111

MITAAVLFDVDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQP
VIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDI FR
DIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLT

>pET #111, Residues 111-220/1-111

MGFRIYRETLSRFSCAAQTAAVLFDVDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEE
TVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQ
EKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLD
SFLESQNKQLTS

>pET #137, Residues 137-220/1-137

MSGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAE
VLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVP
DELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDVDPEDVVVQRLTSRRICPKCG

>pET #137, Residues 137-220/1-137

MGFRIYRETLSRFSCAAQGRINMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGIL
KRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDI FRDIVKKENDELG
KKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDVDPEDVVVQ
RLTSRRICPKCGS

>pET #157, Residues 157-220/1-157

MSKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFL
GPPGAGKGTAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGF
ILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDVDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCK

>pET #157, Residues 157-220/1-157

MGFRIYRETLSRFSCAAQVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKII
GWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNE
VVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFDVDPEDVVVQRLTSRRICPKCGRIYNMISL
PPKEDELCDCKV

>pET #179, Residues 179-220/1-179

MNLEKTQPVIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPH
ISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQN
KQLTAAVLFVDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDEL CDDCKVKLVQRDDDKEETVRHRYKVYL

>pET #179, Residues 179-220/1-179

MGFRIYRETL SRFSCAAQLEKTQPVIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAG
KGTYAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYP
RTVAQAEFLDSFLESQNKQLTAAVLFVDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDEL CDDCKVKLVQRDDDKE
ETVRHRYKVYLD

>pET #190, Residues 190-220/1-190

MNGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI FRDIV
KKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFV
DPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDEL CDDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYG

>pET #190, Residues 190-220/1-190

MGFRIYRETL SRFSCAAQGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEK
TGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSF
LESQNKQLTAAVLFVDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDEL CDDCKVKLVQRDDDKEETVRHRYKVYL
EKTQPVIDYYGT

>pET #209, Residues 209-220/1-209

MIAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGE
LVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFVDPEDVVVQRLTSRRICPKCG
RIYNMISLPPKEDEL CDDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDGTIGIDNVVA

>pET #209, Residues 209-220/1-209

MGFRIYRETL SRFSCAAQAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI FRDIVKKE
NDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFVDPED
VVVQRLTSRRICPKCGRIYNMISLPPKEDEL CDDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKR
VDGTIGIDNVVAD

>pET #217, Residues 217-220/1-217

MRWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVN
EVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFVDPEDVVVQRLTSRRICPKCGRIYNMISL
PPKEDEL CDDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGW

>pET #217, Residues 217-220/1-217

MGFRIYRETL SRFSCAAQWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKI
KEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFVDPEDVVVQRLTS
RRICPKCGRIYNMISLPPKEDEL CDDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDGTIGIDN
VVAEVLKIIIGWS

Rationally designed pET vectors (fission within Peptides I-V)

>pET #12, Residues 12-220/1-12

MTGAGKGTYAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFIL
DGYPRTVAQAEFLDSFLESQNKQLTAAVLFVDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDEL CDDCKVKLVQR
DDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPG

>pET #12, Residues 12-220/1-12

MGFRIYRETL SRFSCAAQGAGKGTYAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNEV
VKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFVDPEDVVVQRLTSRRICPKCGRIYNMISLPP
KEDEL CDDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAM
MAYLVFLGPPGA

>pET #15, Residues 15-220/1-15

MRKGTYAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRRLSEKDCEKGFILDGY
PRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDEL CDDCKVKLVQRDDD
KEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGK

>pET #15, Residues 15-220/1-15

MGFRIYRETLSRFSCAAQKGTYAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRR
RLSEKDCEKGFILDGY PRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKED
ELCDDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLKIIIGWSDKAAAMMAY
LVFLGPPGAGKG

>pET #59, Residues 59-220/1-59

MKGELVPDELVNEVVKRRRLSEKDCEKGFILDGY PRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRIC
PKCGRIYNMISLPPKEDEL CDDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAAE
VLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKG

>pET #59, Residues 59-220/1-59

MGFRIYRETLSRFSCAAQGELVPDELVNEVVKRRRLSEKDCEKGFILDGY PRTVAQAEFLDSFLESQNKQLTAAVLFDP
VEDVVVQRLTSRRICPKCGRIYNMISLPPKEDEL CDDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGI
LKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI FRDIVKKENDEL
GKKIKEIMEKGD

>pET #62, Residues 62-220/1-62

MMVPDELVNEVVKRRRLSEKDCEKGFILDGY PRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKC
GRIYNMISLPPKEDEL CDDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAEVLK
IIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELV

>pET #62, Residues 62-220/1-62

MGFRIYRETLSRFSCAAQVPDELVNEVVKRRRLSEKDCEKGFILDGY PRTVAQAEFLDSFLESQNKQLTAAVLFDPVE
DVVVQRLTSRRICPKCGRIYNMISLPPKEDEL CDDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKR
VDGTIGIDNVVAAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI FRDIVKKENDELGKK
IKEIMEKGELVP

>pET #85, Residues 85-220/1-85

MIILDGY PRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDEL CDDCKVK
LVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAAEVLKIIIGWSDKAAAMMAYLVFLGPPGA
GKGTYAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRRLSEKDCEKGF I

>pET #85, Residues 85-220/1-85

MGFRIYRETLSRFSCAAQILDGY PRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMI
SLPPKEDEL CDDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAAEVLKIIIGWSDK
AAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRR
LSEKDCEKGFIL

>pET #88, Residues 88-220/1-88

MNGY PRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDEL CDDCKVKLVQ
RDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGK
TYAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRRLSEKDCEKGFILDG

>pET #88, Residues 88-220/1-88

MGFRIYRETLSRFSCAAQGY PRTVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLP
PKEDEL CDDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDTIGIDNVVAAEVLKIIIGWSDKAAA
MMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRRLSE
KDCEKGFILDGY

>pET #125, Residues 125-220/1-125
MIQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVD
GTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIK
EIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFVDPEDVVVQ

>pET #125, Residues 125-220/1-125
MGFRIYRETLRFSCAAQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQ
PVIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI
FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAV
LFDVPEDVVVQS

>pET #127, Residues 127-220/1-127
MRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDGT
IGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIK
EIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFVDPEDVVVQRL

>pET #127, Residues 127-220/1-127
MGFRIYRETLRFSCAAQLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQP
VIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI
FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAV
LFDVPEDVVVQRLT

>pET #165, Residues 165-220/1-165
MNDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGK
GTYAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRT
VAQAEFLDSFLESQNKQLTAAVLFVDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDD

>pET #165, Residues 165-220/1-165
MGFRIYRETLRFSCAAQDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAA
MAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSE
KDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFVDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCD
CKVKLVQRDDD

>pET #168, Residues 168-220/1-168
MKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGT
YAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVA
QAEFLDSFLESQNKQLTAAVLFVDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKE

>pET #168, Residues 168-220/1-168
MGFRIYRETLRFSCAAQETVRHRYKVYLEKTQPVIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAM
MAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSE
KDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAVLFVDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCD
CKVKLVQRDDDKE

Rationally designed pET vectors (fission outside Peptides I-V)

>pET #28, Residues 28-220/1-28
MSIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDS
FLESQNKQLTAAVLFVDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYL
EKTQPVIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTI

>pET #40, Residues 40-220/1-40
MNIKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRTVAQAEFLDSFLESQNKQLTAAV
LFDVPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYG
KILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTYAKRIQEKTGIPHISTGDI FRDI

>pET #50, Residues 50-220/1-50

MRKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVAQAEFLDSFLESQNKQLTAAVLFDPEDVVV
QRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDGT
IGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDI FRDIVKKENDELGK

>pET #73, Residues 73-220/1-73

MKRRLSEKDCEKGFILDGYPRVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPP
KEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAM
MAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKR

>pET #98, Residues 98-220/1-98

MKFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVR
HRYKVYLEKTQPVIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEK
TGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVAQAEF

>pET #110, Residues 110-220/1-110

MKLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQP
VIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDI FR
DIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVAQAEFLDSFLESQNKQL

>pET #154, Residues 154-220/1-154

MSDDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYL
VFLGPPGAGKGTAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEK
GFILDGYPRVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCD

>pET #185, Residues 185-220/1-185

MTVIDYYGKKGILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDI
FRDIVKKENDELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVAQAEFLDSFLESQNKQLTAA
VLFDPEDVVVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQP

>pET #194, Residues 194-220/1-194

MSILKRVDGTIGIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDI FRDIVKKEN
DELGKKIKEIMEKGELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVAQAEFLDSFLESQNKQLTAAVLFDPEDV
VVQRLTSRRICPKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGI

>pET #205, Residues 205-220/1-205

MIDNVVAEVLKIIIGWSDKAAAMMAYLVFLGPPGAGKGTAKRIQEKTGIPHISTGDI FRDIVKKENDELGKKIKEIM
KELVPDELVNEVVKRRLSEKDCEKGFILDGYPRVAQAEFLDSFLESQNKQLTAAVLFDPEDVVVQRLTSRRIC
PKCGRIYNMISLPPKEDELCDCKVKLVQRDDDKEETVRHRYKVYLEKTQPVIDYYGKKGILKRVDGTIGID