**Structure *17***

## Supplemental Data

# Rhomboid Protease Dynamics and Lipid Interactions

**Ana-Nicoleta Bondar, Coral del Val, and Stephen H. White**

## SUPPLEMENTAL EXPERIMENTAL PROCEDURES

### Protocol for Sequence Analysis

To investigate the conservation of specific amino acids in rhomboid sequences from various organisms, we analyzed three datasets that were extracted from the PF01694 rhomboid family from the PFAM domain database (version 23.0; Finn et al., 2008). PFAM is a database of multiple alignments of protein domains or conserved protein regions. The PFAM database can be used to identify key residues by examining multiple alignments, to gain insight into the possible the architectures of protein domains, and their distribution across species.

There are two main sets of PFAM sequence alignments. Pfam-A are accurately crafted, hand-curated multiple alignments; the Pfam-B alignments are automatically generated by taking alignments from the ADDA database (Heger and Holme, 2003), and removing any Pfam-A residues from them. The only available data for rhomboid proteins proceed from PFAM-B_1399 (release 4.1). The PF01694 family from the PFAM domain database describes the rhomboid family. The average length of the rhomboid domain is 149.0 aa, the average identity of the full alignment is 22%, and the average coverage of the sequence by the rhomboid domain is 48.19%. We note that the PFAM alignment does not include the first transmembrane helical segment (TM1). Only TM2-TM6 are common to all rhomboid families and constitute the so-called rhomboid domain.

The PF01694 family provides a PFAM seed alignment of 65 sequences, and a PFAM full alignment with 1582 sequences. The PFAM seed alignment is a set of representative sequences, in this case 65, from Archea, Bacteria, Plants, and Eukaryotes. We analyzed separately the seed and the full sequence datasets. In addition to analyzing the seed and the full sequence datasets, we analyzed the Enterobacteriacea sequences from the PFAM database. The Enterobacteriaceae dataset is a subset of 46 sequences containing only the Enterobacteriaceae sequences, which are the closest relatives to the E. coli rhomboid proteins. Because some Pfam-B families are composed of low complexity regions and may not reflect true relationships between sequences, for the seed and for the Enterobacteriaceae alignments we also used BLAST and manual inspection to check the relationship between sequences.

The PFAM seed alignment is used in the PFAM database to generate a profile Hidden Markov Model (profile HMM) (Eddy, 2001). The HMM is used to generate a full alignment, which are all related sequences with score higher than the manually set threshold values for the HMMs of a particular PFAM entry, in this case the rhomboid family. The curation, details, and the threshold values for the HMM of the PF01694 family are described on the PFAM website (http://pfam.sanger.ac.uk/family?acc=PF01694). The full alignment contains 1582 sequences from all Phyla, all of them containing the rhomboid domain.

Conservation analyses, paying special attention to the conservation of hydrophobicity, were generated following the color scheme from (Kyte and Doolittle, 1982). According to this scheme, the most hydrophobic residues are colored in red, and the most hydrophilic residues in blue. From the alignments, we extracted the conservation of specific amino acids discussed in the main text. The data on the conservation of specific amino acids is summarized in Tables S1-S3. To illustrate the conservation of the functional properties of the amino acids (e.g., H-bonding), we use for the sequence analysis the color scheme from (Kyte and Doolittle, 1982) with the following color codes: the hydrophilic residues D, E, K, R, H, N, and Q, are colored in blue; Y and P - violet; G, S, W, and T - pink; A and M - violet; L and V – red; I – deep red.
The sequence alignment tables were prepared using Jalview (Clamp et al., 2004), a Java multiple alignment editor and analysis tool. The amino acid histogram representation for each position in the alignment was performed using a java implementation of LogoBar (Perez-Bercoff et al., 2006).

## SUPPLEMENTAL REFERENCES

Clamp, M., Cuff, J., Searle, S. M., and Barton, G. J. (2004). The Jalview Java Alignment Editor. Bioinformatics, 20, 426-427.

Eddy, S. R. (2001). HMMER User's Guide: Biological sequence analysis using profile hidden Markov models, version 2.2. Washington University School of Medicine.

Finn, R. D., Tate, J., Mistry, J., Coggill, P. C., Sammut, J. S., Hotz, H. R., Ceric, G., Forslund, K., Eddy, S. R., Sonnhammer, E. L., and Bateman, A. (2008). The Pfam protein families database. Nucleic Acids Res. 36, D281-D288.

Heger A., and Holm, L.U. (2003). Exhaustive enumeration of protein domain families. J Mol Biol 328, 749-767.

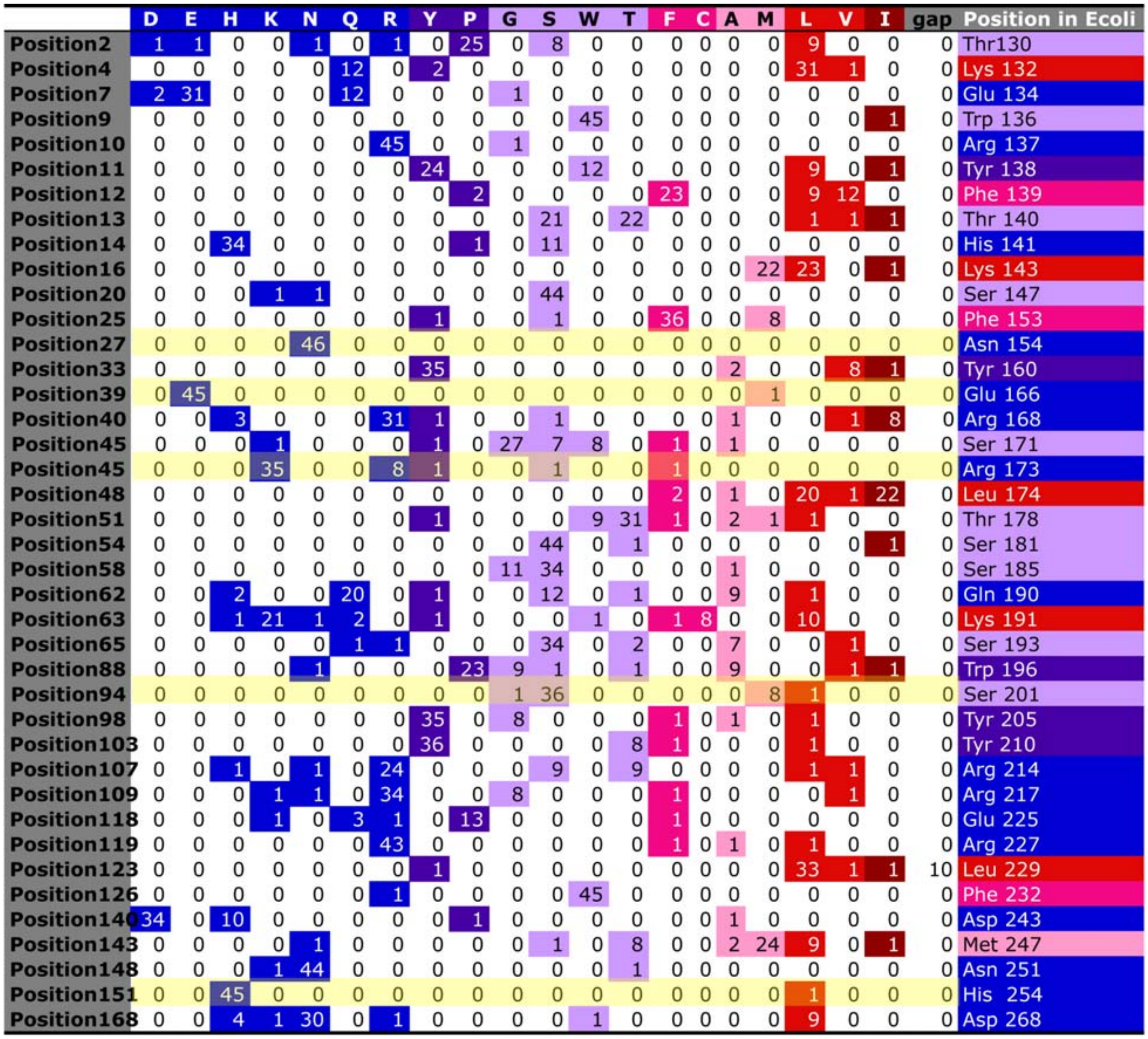
Kyte, J., and Doolittle, R.F. (1982). A simple model for displaying the hydropatic character of a protein. J. Mol. Biol. 157, 105-132.

Pérez-Bercoff, Å., Koch, J., and Bürglin, T.R. (2006). LogoBar: bar graph visualization of protein logos with gaps. Bioinformatics 22, 112-114.

## SUMMARY OF SUPPLEMENTARY INFORMATION TABLES

*Tables S1-S3*.  Conservation of specific amino acids analyzed from three datasets. The conservation of specific amino acids was extracted from the sequence analysis of Enterobacteriaceae (46 sequences; Table S1), the seed sequences of the PFAM database (65 sequences; Table S2), and from the full PFAM database (1582 sequences; Table S3). In each table, 'position' gives the position of the amino acid in the PFAM database sequence, and 'position in E. Coli' indicates the position of the amino acid from the GlpG rhomboid sequence. For example, the GlpG Thr130 amino acid is the second amino acid in the PFAM sequence database. The amino acids are colored according the color scheme of (Kyte and Doolittle, 1982) (see Protocol for Sequence Alignments). The 'gap' gives the insertion/deletion mutations.
*Tables S4-S5*.  Full names and accession codes of the sequences from the Enterobacteriaceae (Table S4) and PFAM seed datasets (Table S5).
*Tables S6-S8*. Sequence alignments for the Enterobacteriaceae (Table S6), PFAM seed (Table S7), and for the complete rhomboid domain from the PFAM database (Table S8).  Table S8 is available from our web site: http://blanco.biomol.uci.edu/download/GlpG_Table_S8.zip.

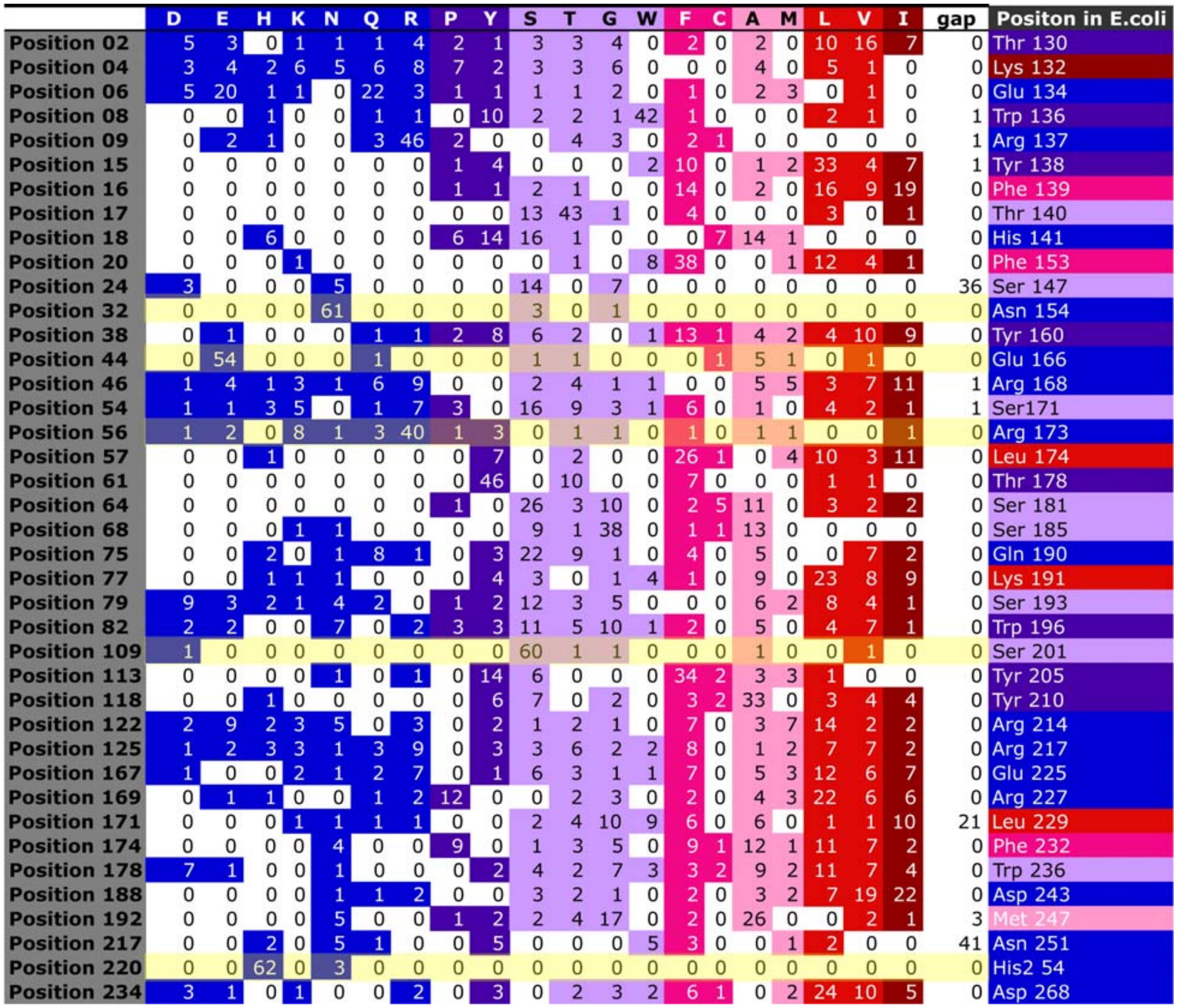
**Table S1**. **The Conservation of Specific Amino Acids Was Extracted from the Sequence Analysis of Enterobacteriaceae (46 Sequences)**



| | D | E | H | K | N | Q | R | Y | P | G | S | W | T | F | C | A | M | L | V | I | gap | Position in Ecoli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Position2 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 25 | 0 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 9 | 0 | 0 | 0 | Thr130 |
| Position4 | 0 | 0 | 0 | 0 | 0 | 12 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 31 | 1 | 0 | 0 | Lys 132 |
| Position7 | 2 | 31 | 0 | 0 | 0 | 12 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | Glu 134 |
| Position9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 45 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | Trp 136 |
| Position10 | 0 | 0 | 0 | 0 | 0 | 0 | 45 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | Arg 137 |
| Position11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 24 | 0 | 0 | 0 | 12 | 0 | 0 | 0 | 0 | 0 | 9 | 0 | 1 | 0 | Tyr 138 |
| Position12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 23 | 0 | 0 | 0 | 9 | 12 | 0 | 0 | Phe 139 |
| Position13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 21 | 0 | 22 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | Thr 140 |
| Position14 | 0 | 0 | 34 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | His 141 |
| Position16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 22 | 23 | 0 | 1 | 0 | Lys 143 |
| Position20 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 44 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | Ser 147 |
| Position25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 36 | 0 | 0 | 8 | 0 | 0 | 0 | 0 | Phe 153 |
| Position27 | 0 | 0 | 0 | 0 | 46 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | Asn 154 |
| Position33 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 35 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 8 | 1 | 0 | Tyr 160 |
| Position39 | 0 | 45 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | Glu 166 |
| Position40 | 0 | 0 | 3 | 0 | 0 | 0 | 31 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 8 | 0 | Arg 168 |
| Position45 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 27 | 7 | 8 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | Ser 171 |
| Position45 | 0 | 0 | 0 | 35 | 0 | 0 | 8 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | Arg 173 |
| Position48 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 1 | 0 | 20 | 1 | 22 | 0 | Leu 174 |
| Position51 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 9 | 31 | 1 | 0 | 2 | 1 | 1 | 0 | 0 | 0 | Thr 178 |
| Position54 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 44 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | Ser 181 |
| Position58 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 11 | 34 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | Ser 185 |
| Position62 | 0 | 0 | 2 | 0 | 0 | 20 | 0 | 1 | 0 | 0 | 12 | 0 | 1 | 0 | 0 | 9 | 0 | 1 | 0 | 0 | 0 | Gln 190 |
| Position63 | 0 | 0 | 1 | 21 | 1 | 2 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 8 | 0 | 0 | 10 | 0 | 0 | 0 | Lys 191 |
| Position65 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 34 | 0 | 2 | 0 | 0 | 7 | 0 | 0 | 1 | 0 | 0 | Ser 193 |
| Position88 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 23 | 9 | 1 | 0 | 1 | 0 | 0 | 9 | 0 | 0 | 1 | 1 | 0 | Trp 196 |
| Position94 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 36 | 0 | 0 | 0 | 0 | 0 | 8 | 1 | 0 | 0 | 0 | Ser 201 |
| Position98 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 35 | 0 | 8 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | Tyr 205 |
| Position103 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 36 | 0 | 0 | 0 | 0 | 8 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | Tyr 210 |
| Position107 | 0 | 0 | 1 | 0 | 1 | 0 | 24 | 0 | 0 | 0 | 9 | 0 | 9 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | Arg 214 |
| Position109 | 0 | 0 | 0 | 1 | 1 | 0 | 34 | 0 | 0 | 8 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | Arg 217 |
| Position118 | 0 | 0 | 0 | 1 | 0 | 3 | 1 | 0 | 13 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | Glu 225 |
| Position119 | 0 | 0 | 0 | 0 | 0 | 0 | 43 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | Arg 227 |
| Position123 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 33 | 1 | 1 | 10 | Leu 229 |
| Position126 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 45 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | Phe 232 |
| Position140 | 34 | 0 | 10 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | Asp 243 |
| Position143 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 8 | 0 | 0 | 2 | 24 | 9 | 0 | 1 | 0 | Met 247 |
| Position148 | 0 | 0 | 0 | 1 | 44 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | Asn 251 |
| Position151 | 0 | 0 | 45 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | His 254 |
| Position168 | 0 | 0 | 4 | 1 | 30 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 9 | 0 | 0 | 0 | Asp 268 |

In the table, ‘position’ gives the position of the amino acid in the PFAM database sequence, and ‘position in E. Coli’ indicates the position of the amino acid from the GlpG rhomboid sequence.

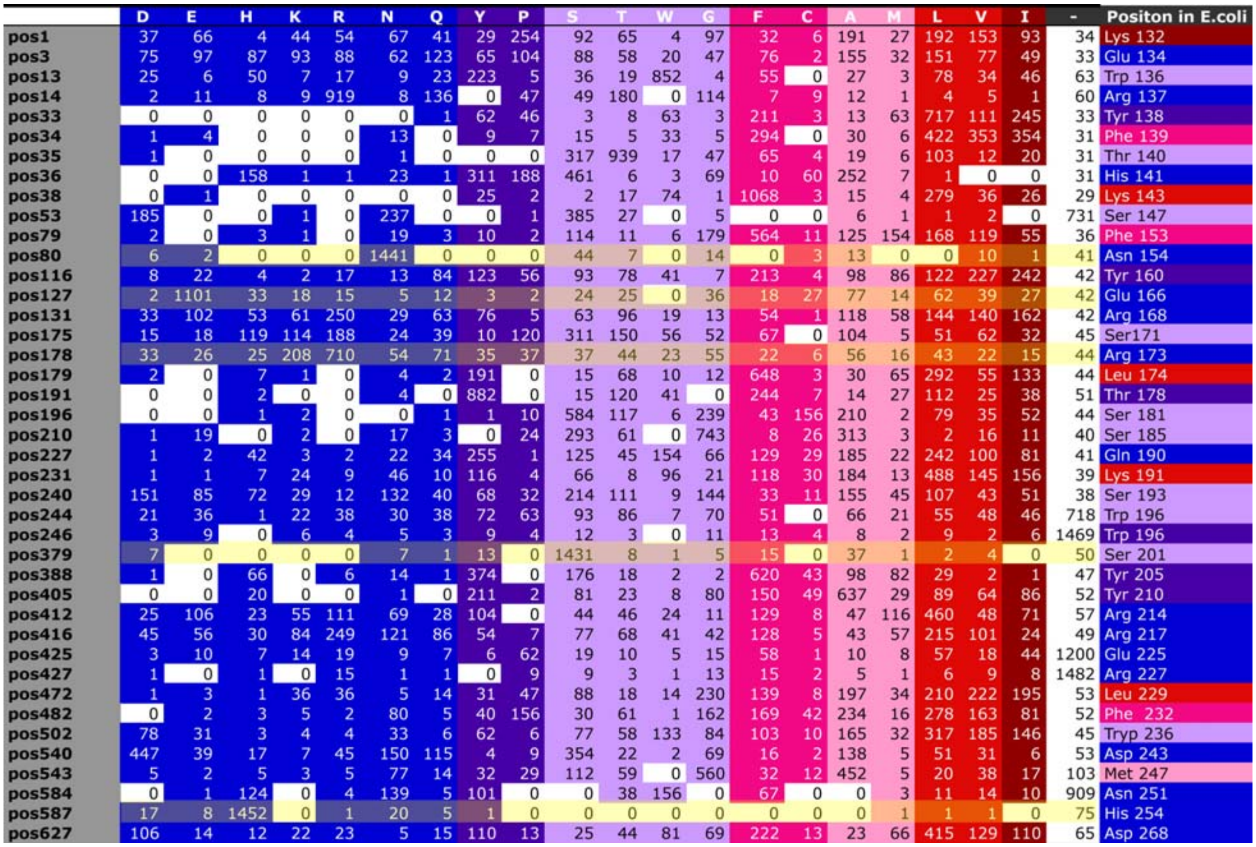
**Table S2. The Seed Sequences of the PFAM Database (65 Sequences)**



| | D | E | H | K | N | Q | R | P | Y | S | T | G | W | F | C | A | M | L | V | I | gap | Positon in E.coli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Position 02 | 5 | 3 | 0 | 1 | 1 | 1 | 4 | 2 | 1 | 3 | 3 | 4 | 0 | 2 | 0 | 2 | 0 | 10 | 16 | 7 | 0 | Thr 130 |
| Position 04 | 3 | 4 | 2 | 6 | 5 | 6 | 8 | 7 | 2 | 3 | 3 | 6 | 0 | 0 | 0 | 4 | 0 | 5 | 1 | 0 | 0 | Lys 132 |
| Position 06 | 5 | 20 | 1 | 1 | 0 | 22 | 3 | 1 | 1 | 1 | 1 | 2 | 0 | 1 | 0 | 2 | 3 | 0 | 1 | 0 | 0 | Glu 134 |
| Position 08 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 10 | 2 | 2 | 1 | 42 | 1 | 0 | 0 | 0 | 2 | 1 | 0 | 1 | Trp 136 |
| Position 09 | 0 | 2 | 1 | 0 | 0 | 3 | 46 | 2 | 0 | 0 | 4 | 3 | 0 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | Arg 137 |
| Position 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 4 | 0 | 0 | 0 | 2 | 10 | 0 | 1 | 2 | 33 | 4 | 7 | 1 | Tyr 138 |
| Position 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 1 | 0 | 0 | 14 | 0 | 2 | 0 | 16 | 9 | 19 | 0 | Phe 139 |
| Position 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 13 | 43 | 1 | 0 | 4 | 0 | 0 | 0 | 3 | 0 | 1 | 0 | Thr 140 |
| Position 18 | 0 | 0 | 6 | 0 | 0 | 0 | 0 | 6 | 14 | 16 | 1 | 0 | 0 | 0 | 7 | 14 | 1 | 0 | 0 | 0 | 0 | His 141 |
| Position 20 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 8 | 38 | 0 | 0 | 1 | 12 | 4 | 1 | 0 | Phe 153 |
| Position 24 | 3 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 14 | 0 | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 36 | Ser 147 |
| Position 32 | 0 | 0 | 0 | 0 | 61 | 0 | 0 | 0 | 0 | 3 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | Asn 154 |
| Position 38 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 2 | 8 | 6 | 2 | 0 | 1 | 13 | 1 | 4 | 2 | 4 | 10 | 9 | 0 | Tyr 160 |
| Position 44 | 0 | 54 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 5 | 1 | 0 | 1 | 0 | 0 | Glu 166 |
| Position 46 | 1 | 4 | 1 | 3 | 1 | 6 | 9 | 0 | 0 | 2 | 4 | 1 | 1 | 0 | 0 | 5 | 5 | 3 | 7 | 11 | 1 | Arg 168 |
| Position 54 | 1 | 1 | 3 | 5 | 0 | 1 | 7 | 3 | 0 | 16 | 9 | 3 | 1 | 6 | 0 | 1 | 0 | 4 | 2 | 1 | 1 | Ser171 |
| Position 56 | 1 | 2 | 0 | 8 | 1 | 3 | 40 | 1 | 3 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | Arg 173 |
| Position 57 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 7 | 0 | 2 | 0 | 0 | 26 | 1 | 0 | 4 | 10 | 3 | 11 | 0 | Leu 174 |
| Position 61 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 46 | 0 | 10 | 0 | 0 | 7 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | Thr 178 |
| Position 64 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 26 | 3 | 10 | 0 | 2 | 5 | 11 | 0 | 3 | 2 | 2 | 0 | Ser 181 |
| Position 68 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 9 | 1 | 38 | 0 | 1 | 1 | 13 | 0 | 0 | 0 | 0 | 0 | Ser 185 |
| Position 75 | 0 | 0 | 2 | 0 | 1 | 8 | 1 | 0 | 3 | 22 | 9 | 1 | 0 | 4 | 0 | 5 | 0 | 0 | 7 | 2 | 0 | Gln 190 |
| Position 77 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 4 | 3 | 0 | 1 | 4 | 1 | 0 | 9 | 0 | 23 | 8 | 9 | 0 | Lys 191 |
| Position 79 | 9 | 3 | 2 | 1 | 4 | 2 | 0 | 1 | 2 | 12 | 3 | 5 | 0 | 0 | 0 | 6 | 2 | 8 | 4 | 1 | 0 | Ser 193 |
| Position 82 | 2 | 2 | 0 | 0 | 7 | 0 | 2 | 3 | 3 | 11 | 5 | 10 | 1 | 2 | 0 | 5 | 0 | 4 | 7 | 1 | 0 | Trp 196 |
| Position 109 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 60 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | Ser 201 |
| Position 113 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 14 | 6 | 0 | 0 | 0 | 34 | 2 | 3 | 3 | 1 | 0 | 0 | 0 | Tyr 205 |
| Position 118 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 6 | 7 | 0 | 2 | 0 | 3 | 2 | 33 | 0 | 3 | 4 | 4 | 0 | Tyr 210 |
| Position 122 | 2 | 9 | 2 | 3 | 5 | 0 | 3 | 0 | 2 | 1 | 2 | 1 | 0 | 7 | 0 | 3 | 7 | 14 | 2 | 2 | 0 | Arg 214 |
| Position 125 | 1 | 2 | 3 | 3 | 1 | 3 | 9 | 0 | 3 | 3 | 6 | 2 | 2 | 8 | 0 | 1 | 2 | 7 | 7 | 2 | 0 | Arg 217 |
| Position 167 | 1 | 0 | 0 | 2 | 1 | 2 | 7 | 0 | 1 | 6 | 3 | 1 | 1 | 7 | 0 | 5 | 3 | 12 | 6 | 7 | 0 | Glu 225 |
| Position 169 | 0 | 1 | 1 | 0 | 0 | 1 | 2 | 12 | 0 | 0 | 2 | 3 | 0 | 2 | 0 | 4 | 3 | 22 | 6 | 6 | 0 | Arg 227 |
| Position 171 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 2 | 4 | 10 | 9 | 6 | 0 | 6 | 0 | 1 | 1 | 10 | 21 | Leu 229 |
| Position 174 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 9 | 0 | 1 | 3 | 5 | 0 | 9 | 1 | 12 | 1 | 11 | 7 | 2 | 0 | Phe 232 |
| Position 178 | 7 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 2 | 4 | 2 | 7 | 3 | 3 | 2 | 9 | 2 | 11 | 7 | 4 | 0 | Trp 236 |
| Position 188 | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 0 | 0 | 3 | 2 | 1 | 0 | 2 | 0 | 3 | 2 | 7 | 19 | 22 | 0 | Asp 243 |
| Position 192 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 1 | 2 | 2 | 4 | 17 | 0 | 2 | 0 | 26 | 0 | 0 | 2 | 1 | 3 | Met 247 |
| Position 217 | 0 | 0 | 2 | 0 | 5 | 1 | 0 | 0 | 5 | 0 | 0 | 0 | 5 | 3 | 0 | 0 | 1 | 2 | 0 | 0 | 41 | Asn 251 |
| Position 220 | 0 | 0 | 62 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | His2 54 |
| Position 234 | 3 | 1 | 0 | 1 | 0 | 0 | 2 | 0 | 3 | 0 | 2 | 3 | 2 | 6 | 1 | 0 | 2 | 24 | 10 | 5 | 0 | Asp 268 |

In the table, ‘position’ gives the position of the amino acid in the PFAM database sequence, and ‘position in E. Coli’ indicates the position of the amino acid from the GlpG rhomboid sequence.

**Table S3. Seed Sequences from the Full PFAM Database (1582 Sequences)**



| | D | E | H | K | R | N | Q | Y | P | S | T | W | G | F | C | A | M | L | V | I | - | Positon in E.coli |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| pos1 | 37 | 66 | 4 | 44 | 54 | 67 | 41 | 29 | 254 | 92 | 65 | 4 | 97 | 32 | 6 | 191 | 27 | 192 | 153 | 93 | 34 | Lys 132 |
| pos3 | 75 | 97 | 87 | 93 | 88 | 62 | 123 | 65 | 104 | 88 | 58 | 20 | 47 | 76 | 2 | 155 | 32 | 151 | 77 | 49 | 33 | Glu 134 |
| pos13 | 25 | 6 | 50 | 7 | 17 | 9 | 23 | 223 | 5 | 36 | 19 | 852 | 4 | 55 | 0 | 27 | 3 | 78 | 34 | 46 | 63 | Trp 136 |
| pos14 | 2 | 11 | 8 | 9 | 919 | 8 | 136 | 0 | 47 | 49 | 180 | 0 | 114 | 7 | 9 | 12 | 1 | 4 | 5 | 1 | 60 | Arg 137 |
| pos33 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 62 | 46 | 3 | 8 | 63 | 3 | 211 | 3 | 13 | 63 | 717 | 111 | 245 | 33 | Tyr 138 |
| pos34 | 1 | 4 | 0 | 0 | 0 | 13 | 0 | 9 | 7 | 15 | 5 | 33 | 5 | 294 | 0 | 30 | 6 | 422 | 353 | 354 | 31 | Phe 139 |
| pos35 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 317 | 939 | 17 | 47 | 65 | 4 | 19 | 6 | 103 | 12 | 20 | 31 | Thr 140 |
| pos36 | 0 | 0 | 158 | 1 | 1 | 23 | 1 | 311 | 188 | 461 | 6 | 3 | 69 | 10 | 60 | 252 | 7 | 1 | 0 | 0 | 31 | His 141 |
| pos38 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 25 | 2 | 2 | 17 | 74 | 1 | 1068 | 3 | 15 | 4 | 279 | 36 | 26 | 29 | Lys 143 |
| pos53 | 185 | 0 | 0 | 1 | 0 | 237 | 0 | 0 | 1 | 385 | 27 | 0 | 5 | 0 | 0 | 6 | 1 | 1 | 2 | 0 | 731 | Ser 147 |
| pos79 | 2 | 0 | 3 | 1 | 0 | 19 | 3 | 10 | 2 | 114 | 11 | 6 | 179 | 564 | 11 | 125 | 154 | 168 | 119 | 55 | 36 | Phe 153 |
| pos80 | 6 | 2 | 0 | 0 | 0 | 1441 | 0 | 0 | 0 | 44 | 7 | 0 | 14 | 0 | 3 | 13 | 0 | 0 | 10 | 1 | 41 | Asn 154 |
| pos116 | 8 | 22 | 4 | 2 | 17 | 13 | 84 | 123 | 56 | 93 | 78 | 41 | 7 | 213 | 4 | 98 | 86 | 122 | 227 | 242 | 42 | Tyr 160 |
| pos127 | 2 | 1101 | 33 | 18 | 15 | 5 | 12 | 3 | 2 | 24 | 25 | 0 | 36 | 18 | 27 | 77 | 14 | 62 | 39 | 27 | 42 | Glu 166 |
| pos131 | 33 | 102 | 53 | 61 | 250 | 29 | 63 | 76 | 5 | 63 | 96 | 19 | 13 | 54 | 1 | 118 | 58 | 144 | 140 | 162 | 42 | Arg 168 |
| pos175 | 15 | 18 | 119 | 114 | 188 | 24 | 39 | 10 | 120 | 311 | 150 | 56 | 52 | 67 | 0 | 104 | 5 | 51 | 62 | 32 | 45 | Ser171 |
| pos178 | 33 | 26 | 25 | 208 | 710 | 54 | 71 | 35 | 37 | 37 | 44 | 23 | 55 | 22 | 6 | 56 | 16 | 43 | 22 | 15 | 44 | Arg 173 |
| pos179 | 2 | 0 | 7 | 1 | 0 | 4 | 2 | 191 | 0 | 15 | 68 | 10 | 12 | 648 | 3 | 30 | 65 | 292 | 55 | 133 | 44 | Leu 174 |
| pos191 | 0 | 0 | 2 | 0 | 0 | 4 | 0 | 882 | 0 | 15 | 120 | 41 | 0 | 244 | 7 | 14 | 27 | 112 | 25 | 38 | 51 | Thr 178 |
| pos196 | 0 | 0 | 1 | 2 | 0 | 0 | 1 | 1 | 10 | 584 | 117 | 6 | 239 | 43 | 156 | 210 | 2 | 79 | 35 | 52 | 44 | Ser 181 |
| pos210 | 1 | 19 | 0 | 2 | 0 | 17 | 3 | 0 | 24 | 293 | 61 | 0 | 743 | 8 | 26 | 313 | 3 | 2 | 16 | 11 | 40 | Ser 185 |
| pos227 | 1 | 2 | 42 | 3 | 2 | 22 | 34 | 255 | 1 | 125 | 45 | 154 | 66 | 129 | 29 | 185 | 22 | 242 | 100 | 81 | 41 | Gln 190 |
| pos231 | 1 | 1 | 7 | 24 | 9 | 46 | 10 | 116 | 4 | 66 | 8 | 96 | 21 | 118 | 30 | 184 | 13 | 488 | 145 | 156 | 39 | Lys 191 |
| pos240 | 151 | 85 | 72 | 29 | 12 | 132 | 40 | 68 | 32 | 214 | 111 | 9 | 144 | 33 | 11 | 155 | 45 | 107 | 43 | 51 | 38 | Ser 193 |
| pos244 | 21 | 36 | 1 | 22 | 38 | 30 | 38 | 72 | 63 | 93 | 86 | 7 | 70 | 51 | 0 | 66 | 21 | 55 | 48 | 46 | 718 | Trp 196 |
| pos246 | 3 | 9 | 0 | 6 | 4 | 5 | 3 | 9 | 4 | 12 | 3 | 0 | 11 | 13 | 4 | 8 | 2 | 9 | 2 | 6 | 1469 | Trp 196 |
| pos379 | 7 | 0 | 0 | 0 | 0 | 7 | 1 | 13 | 0 | 1431 | 8 | 1 | 5 | 15 | 0 | 37 | 1 | 2 | 4 | 0 | 50 | Ser 201 |
| pos388 | 1 | 0 | 66 | 0 | 6 | 14 | 1 | 374 | 0 | 176 | 18 | 2 | 2 | 620 | 43 | 98 | 82 | 29 | 2 | 1 | 47 | Tyr 205 |
| pos405 | 0 | 0 | 20 | 0 | 0 | 1 | 0 | 211 | 2 | 81 | 23 | 8 | 80 | 150 | 49 | 637 | 29 | 89 | 64 | 86 | 52 | Tyr 210 |
| pos412 | 25 | 106 | 23 | 55 | 111 | 69 | 28 | 104 | 0 | 44 | 46 | 24 | 11 | 129 | 8 | 47 | 116 | 460 | 48 | 71 | 57 | Arg 214 |
| pos416 | 45 | 56 | 30 | 84 | 249 | 121 | 86 | 54 | 7 | 77 | 68 | 41 | 42 | 128 | 5 | 43 | 57 | 215 | 101 | 24 | 49 | Arg 217 |
| pos425 | 3 | 10 | 7 | 14 | 19 | 9 | 7 | 6 | 62 | 19 | 10 | 5 | 15 | 58 | 1 | 10 | 8 | 57 | 18 | 44 | 1200 | Glu 225 |
| pos427 | 1 | 0 | 1 | 0 | 15 | 1 | 1 | 0 | 9 | 9 | 3 | 1 | 13 | 15 | 2 | 5 | 1 | 6 | 9 | 8 | 1482 | Arg 227 |
| pos472 | 1 | 3 | 1 | 36 | 36 | 5 | 14 | 31 | 47 | 88 | 18 | 14 | 230 | 139 | 8 | 197 | 34 | 210 | 222 | 195 | 53 | Leu 229 |
| pos482 | 0 | 2 | 3 | 5 | 2 | 80 | 5 | 40 | 156 | 30 | 61 | 1 | 162 | 169 | 42 | 234 | 16 | 278 | 163 | 81 | 52 | Phe 232 |
| pos502 | 78 | 31 | 3 | 4 | 4 | 33 | 6 | 62 | 6 | 77 | 58 | 133 | 84 | 103 | 10 | 165 | 32 | 317 | 185 | 146 | 45 | Tryp 236 |
| pos540 | 447 | 39 | 17 | 7 | 45 | 150 | 115 | 4 | 9 | 354 | 22 | 2 | 69 | 16 | 2 | 138 | 5 | 51 | 31 | 6 | 53 | Asp 243 |
| pos543 | 5 | 2 | 5 | 3 | 5 | 77 | 14 | 32 | 29 | 112 | 59 | 0 | 560 | 32 | 12 | 452 | 5 | 20 | 38 | 17 | 103 | Met 247 |
| pos584 | 0 | 1 | 124 | 0 | 4 | 139 | 5 | 101 | 0 | 0 | 38 | 156 | 0 | 67 | 0 | 0 | 3 | 11 | 14 | 10 | 909 | Asn 251 |
| pos587 | 17 | 8 | 1452 | 0 | 1 | 20 | 5 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 75 | His 254 |
| pos627 | 106 | 14 | 12 | 22 | 23 | 5 | 15 | 110 | 13 | 25 | 44 | 81 | 69 | 222 | 13 | 23 | 66 | 415 | 129 | 110 | 65 | Asp 268 |

In the table, ‘position’ gives the position of the amino acid in the PFAM database sequence, and ‘position in E. Coli’ indicates the position of the amino acid from the GlpG rhomboid sequence.

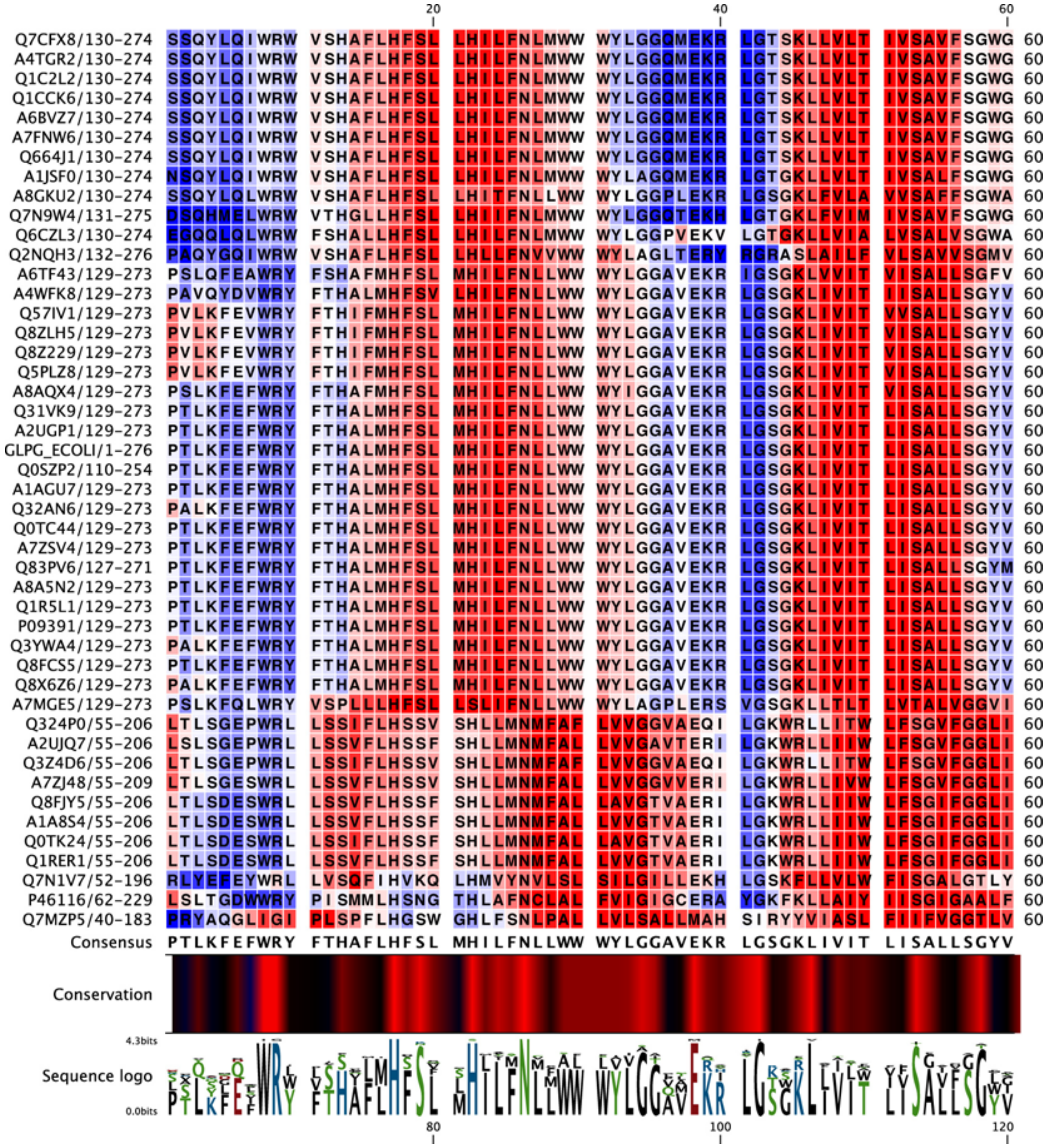
## Table S4. Full Names and Accession Codes of the Sequences from the Enterobacteriaceae Dataset

| Alignment position | Accession number | Organism | | Begin Alignment | end Alignment |
|---|---|---|---|---|---|
| **1** | Q7CFX8 | Yersinia pestis | 278 | 130 | 274 |
| **2** | A4TGR2 | Yersinia pestis Pestoides F | 278 | 130 | 274 |
| **3** | Q1C2L2 | Yersinia pestis Antiqua | 278 | 130 | 274 |
| **4** | Q1CCK6 | Yersinia pestis Nepal516 | 278 | 130 | 274 |
| **5** | A6BVZ7 | Yersinia pestis CA88-4125. | 278 | 130 | 274 |
| **6** | A7FNW6 | Yersinia pseudotuberculosis IP 31758 | 278 | 130 | 274 |
| **7** | Q664J1 | Yersinia pseudotuberculosis | 278 | 130 | 274 |
| **8** | A1JSF0 | Yersinia enterocolitica subsp. enterocolitica 8081 | 278 | 130 | 274 |
| **9** | A8GKU2 | Serratia proteamaculans 568 | 278 | 130 | 274 |
| **10** | Q7N9W4 | Photorhabdus luminescens subsp. laumondii | 282 | 131 | 275 |
| **11** | Q6CZL3 | Sodalis glossinidius str. 'morsitans' | 282 | 130 | 274 |
| **12** | Q2NQH3 | Sodalis glossinidius (strain morsitans) | 276 | 132 | 276 |
| **13** | A6TF43 | Klebsiella pneumoniae subsp | 276 | 129 | 273 |
| **14** | A4WFK8 | Enterobacter sp. (strain 638). | 276 | 129 | 273 |
| **15** | Q57IV1 | Salmonella choleraesuis. | 276 | 129 | 273 |
| **16** | Q8ZLH5 | Salmonella typhimurium. | 276 | 129 | 273 |
| **17** | Q8Z229 | Salmonella typhi | 276 | 129 | 273 |
| **18** | Q5PLZ8 | Salmonella paratyphi A. | 276 | 129 | 273 |
| **19** | A8AQX4 | Citrobacter koser | 276 | 129 | 273 |
| **20** | Q31VK9 | Shigella boydii serotype 4 (strain Sb227). | 276 | 129 | 273 |
| **21** | A2UGP1 | Escherichia coli O9:H4 (strain HS). | 276 | 129 | 273 |
| **22** | GLPG_ECOLI | Escherichia coli (strain K12). | 276 | 1 | 276 |
| **23** | Q0SZP2 | Shigella flexneri serotype 5b (strain 8401). | 276 | 110 | 254 |
| **24** | A1AGU7 | Escherichia coli O157:H7 | 276 | 129 | 273 |
| **25** | Q32AN6 | Shigella dysenteriae Sd197 | 276 | 129 | 273 |
| **26** | Q0TC44 | Shigella sonnei Ss046 | 276 | 129 | 273 |
| **27** | A7ZSV4 | Salmonella typhimurium | 276 | 129 | 273 |
| **28** | Q83PV6 | Salmonella typhi | 276 | 127 | 271 |
| **29** | A8A5N2 | Salmonella enterica subsp. enterica serovar Paratyphi | 276 | 129 | 273 |
| **30** | Q1R5L1 | Salmonella choleraesuis | 276 | 129 | 273 |
| **31** | P09391 | Escherichia coli O1:K1 / APEC. | 276 | 129 | 273 |
| **32** | Q3YWA4 | Shigella sonnei (strain Ss046). | 276 | 129 | 273 |
| **33** | Q8FCS5 | Escherichia coli O6. | 276 | 129 | 273 |
| **34** | Q8X6Z6 | Escherichia coli O157:H7. | 276 | 129 | 273 |
| **35** | A7MGE5 | Enterobacter sakazakii ATCC BAA-894 | 276 | 129 | 273 |
| **36** | Q324P0 | Shigella boydii serotype 4 (strain Sb227). | 588 | 55 | 206 |
| **37** | A2UJQ7 | Escherichia coli (ATCC 8739 / DSM 1576 / Crooks) | 625 | 55 | 206 |
| **38** | Q3Z4D6 | Shigella sonnei (strain Ss046). | 625 | 55 | 206 |
| **39** | A7ZJ48 | Escherichia coli O139:H28 (strain E24377A / ETEC). | 625 | 55 | 209 |
| **40** | Q8FJY5 | Escherichia coli O6 | 625 | 55 | 206 |
| **41** | A1A8S4 | Escherichia coli O1:K1 / APEC. | 588 | 55 | 206 |
| **42** | Q0TK24 | Escherichia coli O6:K15:H31 | 625 | 55 | 206 |
| **43** | Q1RER1 | Providencia stuartii | 281 | 55 | 206 |
| **44** | Q7N1V7 | Photorhabdus luminescens subsp. laumondii | 204 | 52 | 196 |
| **45** | P46116 | Photorhabdus luminescens subsp. laumondii | 186 | 62 | 229 |
| **46** | Q7MZP5 | Photorhabdus luminescens subsp. laumondii | 186 | 40 | 183 |

# Table S5. Full Names and Accession Codes of the Sequences from PFAM Seed Dataset

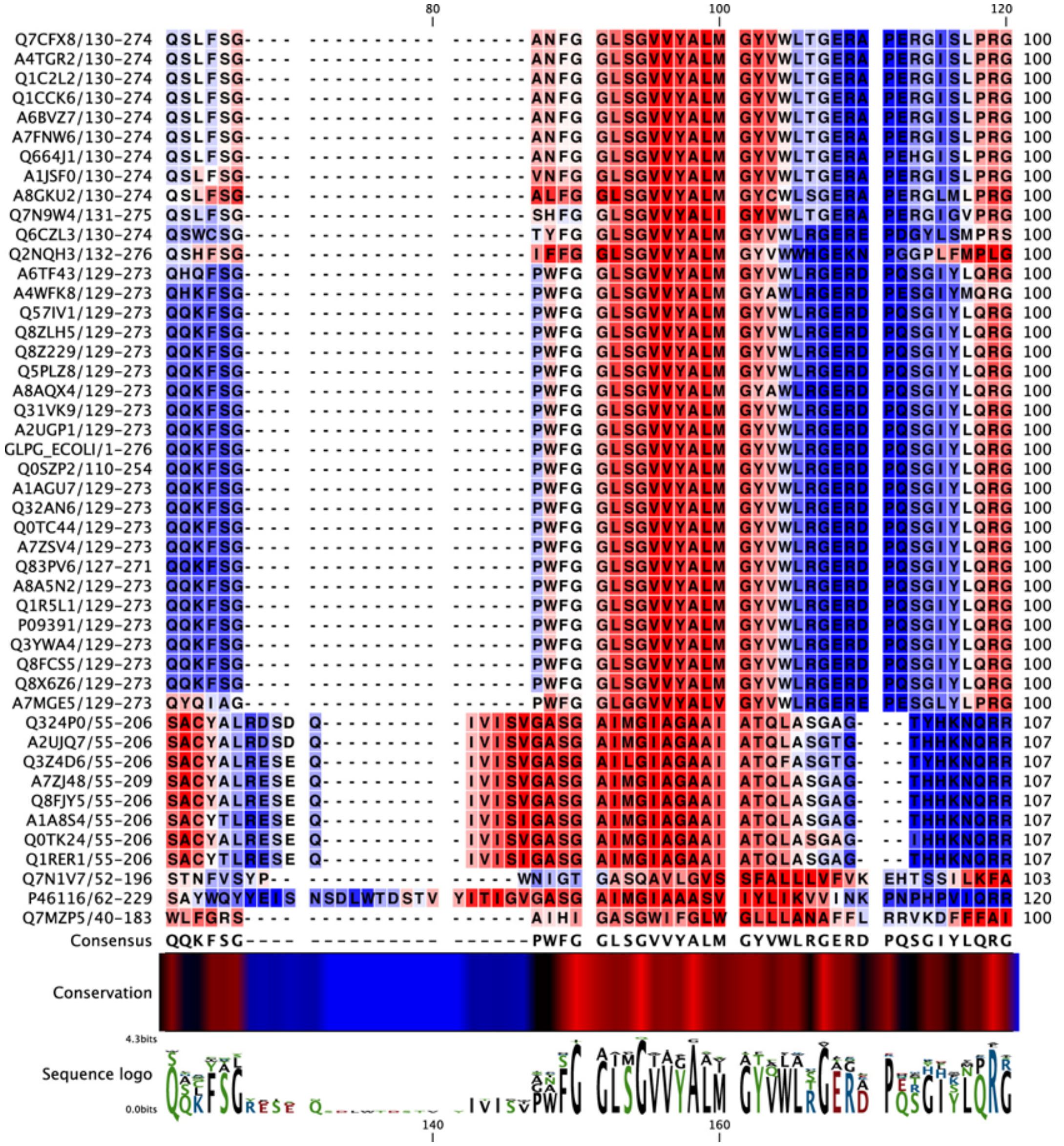
| Alignment position | Accession number | Specie | Sequence length | Begin Alignment | end Alignment |
|---|---|---|---|---|---|
| **1** | Q9FRH8 | Arabidopsis thaliana | 369 | 187 | 338 |
| **2** | Q97KG5 | Clostridium acetobutylicum | 328 | 183 | 326 |
| **3** | Q9FFX0 | Arabidopsis thaliana | 434 | 225 | 370 |
| **4** | YQGP | Bacillus subtilis | 507 | 216 | 358 |
| **5** | Q9KCZ8 | Bacillus halodurans | 514 | 225 | 367 |
| **6** | AARA | Providencia stuartii | 281 | 62 | 229 |
| **7** | Q9HEF8 | Neurospora crassa | 548 | 275 | 420 |
| **8** | Q9SH15 | Arabidopsis thaliana | 735 | 109 | 274 |
| **9** | Q9SSR0 | Arabidopsis thaliana | 309 | 100 | 246 |
| **10** | Q9M348 | Arabidopsis thaliana | 361 | 105 | 244 |
| **11** | O82756 | Arabidopsis thaliana | 313 | 98 | 242 |
| **12** | Q9LN72 | Arabidopsis thaliana | 302 | 86 | 230 |
| **13** | O81073 | Arabidopsis thaliana | 372 | 149 | 321 |
| **14** | Q43323 | Saccharum hybrid cultivar H65-7052 | 325 | 108 | 253 |
| **15** | Q9CAN1 | Arabidopsis thaliana | 317 | 104 | 248 |
| **16** | Q9HBK7 | Homo sapiens | 434 | 99 | 250 |
| **17** | RBD2 | Saccharomyces cerevisiae | 262 | 47 | 198 |
| **18** | RBD2 | Saccharomices pombe | 262 | 47 | 198 |
| **19** | Q9M1B5 | Arabidopsis thaliana | 269 | 45 | 204 |
| **20** | Y3846 | Yersinia pestis KIM | 289 | 58 | 216 |
| **21** | Q9AQU7 | Oryza sativa (japonica cultivar-group) | 350 | 85 | 239 |
| **22** | Q9HZC2 | Pseudomonas aeruginosa | 286 | 134 | 283 |
| **23** | Q9CL10 | Pasteurella multocida | 291 | 144 | 289 |
| **24** | GLPG_HAIN | Haemophilus influenzae | 192 | 44 | 188 |
| **25** | Q9KVP2 | Vibrio cholerae | 277 | 132 | 275 |
| **26** | GLPG_ECOLI | Escherichia coli K12 | 276 | 129 | 273 |
| **27** | Q9PFK8 | Xylella fastidiosa | 224 | 53 | 207 |
| **28** | O83947 | Treponema pallidum | 208 | 47 | 200 |
| **29** | PCP1 | Schizosaccharomyces pombe | 346 | 173 | 332 |
| **30** | Q9FZ81 | Arabidopsis thaliana | 336 | 168 | 334 |
| **31** | PARL | Human | 379 | 198 | 354 |
| **32** | Q9A538 | Caulobacter crescentus | 218 | 43 | 209 |
| **33** | Q98ND6 | Mesorhizobium loti | 237 | 54 | 226 |
| **34** | Q9HS82 | Halobacterium salinarum | 333 | 132 | 304 |
| **35** | O29251 | Archaeoglobus fulgidus | 330 | 166 | 324 |
| **36** | Q9PEH4 | Xylella fastidiosa | 232 | 68 | 227 |
| **37** | Q9X7Z8 | Streptomyces coelicolor | 256 | 83 | 240 |
| **38** | O69868 | Streptomyces coelicolor | 383 | 202 | 365 |
| **39** | Q9X0H3 | Thermotoga maritima | 222 | 56 | 217 |
| **40** | O67346 | Aquifex aeolicus | 227 | 62 | 226 |
| **41** | Q9YAR6 | Aeropyrum pernix | 253 | 61 | 239 |
| **42** | Q9RSX7 | Deinococcus radiodurans | 232 | 86 | 230 |
| **43** | P74553 | Synechocystis sp. PCC 6803 | 198 | 49 | 193 |
| **44** | Y1372 | Mycobacterium bovis | 240 | 71 | 223 |
| **45** | Y1171 | Mycobacterium leprae | 251 | 83 | 234 |
| **46** | Q9VKA6 | Drosophila melanogaster | 171 | 9 | 162 |
| **47** | Q9BML4 | Drosophila melanogaster | 431 | 208 | 377 |
| **48** | RHOM | Drosophila melanogaster | 355 | 144 | 300 |
| **49** | Q9W0F8 | Drosophila melanogaster | 341 | 129 | 285 |
| **50** | ROM1 | Caenorhabditis elegans | 1348 | 160 | 312 |
| **51** | RHBL2 | Toxoplasma gondii | 283 | 114 | 269 |
| **52** | RHBL1 | Human | 865 | 239 | 396 |
| **53** | ROM2 | Caenorhabditis elegans | 435 | 163 | 314 |
| **54** | Q9U2V5 | Caenorhabditis elegans | 727 | 446 | 590 |
| **55** | Q9U2S3 | Caenorhabditis elegans | 861 | 579 | 725 |
| **56** | Q97NE9 | Streptococcus pneumoniae | 225 | 52 | 195 |
| **57** | Q9CJ36 | Lactococcus lactis subsp. lactis | 230 | 56 | 200 |
| **58** | O59166 | Pyrococcus horikoshii | 197 | 45 | 184 |
| **59** | Q9WZ53 | Thermotoga maritima | 235 | 48 | 194 |
| **60** | Q99TU5 | Staphylococcus aureus subsp. aureus Mu50 | 487 | 196 | 338 |
| **61** | O24850 | Acinetobacter sp. ADP1 | 276 | 60 | 226 |
| **62** | Q9KFG2 | Bacillus halodurans | 248 | 51 | 196 |
| **63** | P96617 | Bacillus subtilis | 199 | 52 | 196 |
| **64** | Q9XA09 | Streptomyces coelicolor | 297 | 128 | 267 |
| **65** | O53632 | Mycobacterium tuberculosis | 249 | 70 | 209 |

**Table 6A. Sequence Alignments for the Enterobacteria**



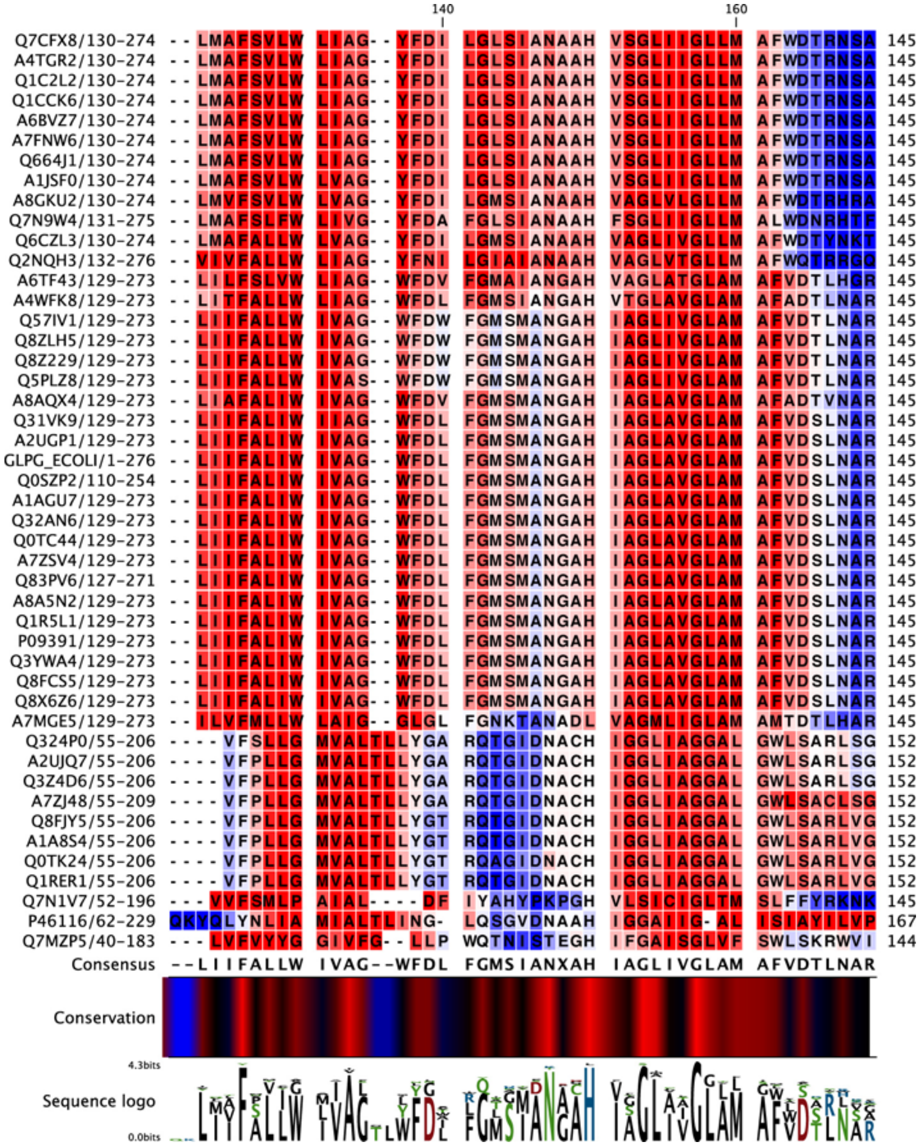
```
Q7CFX8/130-274   SSQYLQIWRW VSHAFLHFSL LHILFNLMWW WYLGGQMEKR LGTSKLLVLT IVSAVFSGWG 60
A4TGR2/130-274   SSQYLQIWRW VSHAFLHFSL LHILFNLMWW WYLGGQMEKR LGTSKLLVLT IVSAVFSGWG 60
Q1C2L2/130-274   SSQYLQIWRW VSHAFLHFSL LHILFNLMWW WYLGGQMEKR LGTSKLLVLT IVSAVFSGWG 60
Q1CCK6/130-274   SSQYLQIWRW VSHAFLHFSL LHILFNLMWW WYLGGQMEKR LGTSKLLVLT IVSAVFSGWG 60
A6BVZ7/130-274   SSQYLQIWRW VSHAFLHFSL LHILFNLMWW WYLGGQMEKR LGTSKLLVLT IVSAVFSGWG 60
A7FNW6/130-274   SSQYLQIWRW VSHAFLHFSL LHILFNLMWW WYLGGQMEKR LGTSKLLVLT IVSAVFSGWG 60
Q664J1/130-274   SSQYLQIWRW VSHAFLHFSL LHILFNLMWW WYLGGQMEKR LGTSKLLVLT IVSAVFSGWG 60
A1JSF0/130-274   NSQYLQIWRW VSHAFLHFSL LHILFNLMWW WYLAGQMEKR LGTGKLLVLT IVSALFSGWG 60
A8GKU2/130-274   SSQYLQLWRW VSHAFLHFSL LHITFNLLWW WYLGGPLEKR LGSGKLFVLA VVSAFFSGWA 60
Q7N9W4/131-275   DSQHMELWRW VTHGLLHFSL LHIIFNLMWW WYLGGQTEKH LGTGKLFVIM IVSAVFSGWG 60
Q6CZL3/130-274   EGQQLQLWRW FSHALLHFSL LHILFNLMWW WYLGGPVEKV LGTGKLLVIA LVSALVSGWA 60
Q2NQH3/132-276   PAQYGQIWRW VSHAFLHFSL LHLLFNVVWW WYLAGLTERY RGRASLAILF VLSAVVSGMV 60
A6TF43/129-273   PSLQFEAWRY FSHAFMHFSL MHILFNLLWW WYLGGAVEKR IGSGKLVVIT VISALLSGFV 60
A4WFK8/129-273   PAVQYDVWRY FTHALMHFSV LHILFNLLWW WYLGGAVEKR LGSGKLIVIT IISALLSGYV 60
Q57IV1/129-273   PVLKFEVWRY FTHIFMHFSL MHILFNLLWW WYLGGAVEKR LGSGKLIVIT VVSALLSGYV 60
Q8ZLH5/129-273   PVLKFEVWRY FTHIFMHFSL MHILFNLLWW WYLGGAVEKR LGSGKLIVIT VISALLSGYV 60
Q8Z229/129-273   PVLKFEVWRY FTHIFMHFSL MHILFNLLWW WYLGGAVEKR LGSGKLIVIT VISALLSGYV 60
Q5PLZ8/129-273   PVLKFEVWRY FTHIFMHFSL MHILFNLLWW WYLGGAVEKR LGSGKLIVIT VISALLSGYV 60
A8AQX4/129-273   PSLKFEFWRY FTHAFMHFSL MHILFNLLWW WYIGGAVEKR LGSGKLIVIT VISALLSGYV 60
Q31VK9/129-273   PTLKFEFWRY FTHALMHFSL MHILFNLLWW WYLGGAVEKR LGSGKLIVIT LISALLSGYV 60
A2UGP1/129-273   PTLKFEFWRY FTHALMHFSL MHILFNLLWW WYLGGAVEKR LGSGKLIVIT LISALLSGYV 60
GLPG_ECOLI/1-276 PTLKFEFWRY FTHALMHFSL MHILFNLLWW WYLGGAVEKR LGSGKLIVIT LISALLSGYV 60
Q0SZP2/110-254   PTLKFEFWRY FTHALMHFSL MHILFNLLWW WYLGGAVEKR LGSGKLIVIT LISALLSGYV 60
A1AGU7/129-273   PTLKFEFWRY FTHALMHFSL MHILFNLLWW WYLGGAVEKR LGSGKLIVIT LISALLSGYV 60
Q32AN6/129-273   PALKFEFWRY FTHALMHFSL MHILFNLLWW WYLGGAVEKR LGSGKLIVIT LISALLSGYV 60
Q0TC44/129-273   PTLKFEFWRY FTHALMHFSL MHILFNLLWW WYLGGAVEKR LGSGKLIVIT LISALLSGYV 60
A7ZSV4/129-273   PTLKFEFWRY FTHALMHFSL MHILFNLLWW WYLGGAVEKR LGSGKLIVIT LISALLSGYV 60
Q83PV6/127-271   PTLKFEFWRY FTHALMHFSL MHILFNLLWW WYLGGAVEKR LGSGKLIVIT LISALLSGYM 60
A8A5N2/129-273   PTLKFEFWRY FTHALMHFSL MHILFNLLWW WYLGGAVEKR LGSGKLIVIT LISALLSGYV 60
Q1R5L1/129-273   PTLKFEFWRY FTHALMHFSL MHILFNLLWW WYLGGAVEKR LGSGKLIVIT LISALLSGYV 60
P09391/129-273   PTLKFEFWRY FTHALMHFSL MHILFNLLWW WYLGGAVEKR LGSGKLIVIT LISALLSGYV 60
Q3YWA4/129-273   PALKFEFWRY FTHALMHFSL MHILFNLLWW WYLGGAVEKR LGSGKLIVIT LISALLSGYV 60
Q8FCS5/129-273   PTLKFEFWRY FTHALMHFSL MHILFNLLWW WYLGGAVEKR LGSGKLIVIT LISALLSGYV 60
Q8X6Z6/129-273   PALKFEFWRY FTHALMHFSL MHILFNLLWW WYLGGAVEKR LGSGKLIVIT LISALLSGYV 60
A7MGE5/129-273   PSLKFQLWRY VSPLLLHFSL LSLIFNLLWW WYLAGPLERS VGSGKLLTLT LVTALVGGVI 60
Q324P0/55-206    LTLSGEPWRL LSSIFLHSSV SHLLMNMFAF LVVGGVAEQI LGKWRLLITW LFSGVFGGLI 60
A2UJQ7/55-206    LSLSGEPWRL LSSVFLHSSF SHLLMNMFAL LVVGAVTERI LGKWRLLIIW LFSGVFGGLI 60
Q3Z4D6/55-206    LTLSGEPWRL LSSIFLHSSV SHLLMNMFAF LVVGGVAEQI LGKWRLLITW LFSGVFGGLI 60
A7ZJ48/55-209    LTLSGESWRL LSSVFLHSSV SHLLMNMFAL LVVGGVVERI LGKWRLLIVW LFSGVFGGLI 60
Q8FJY5/55-206    LTLSDESWRL LSSVFLHSSF SHLLMNMFAL LAVGTVAERI LGKWRLLIIW LFSGIFGGLI 60
A1A8S4/55-206    LTLSDESWRL LSSVFLHSSF SHLLMNMFAL LVVGTVAERI LGKWRLLIIW LFSGIFGGLI 60
Q0TK24/55-206    LTLSDESWRL LSSIFLHSSF SHLLMNMFAL LAVGTVAERI LGKWRLLIIW LFSGIFGGLI 60
Q1RER1/55-206    LTLSDESWRL LSSVFLHSSF SHLLMNMFAL LVVGTVAERI LGKWRLLIIW LFSGIFGGLI 60
Q7N1V7/52-196    RLYEFEYWRL LVSQFIHVKQ LHMVYNVLSL SILGILLEKH LGSKFLLVLW FISGALGTLY 60
P46116/62-229    LSLTGDWWRY PISMMLHSNG THLAFNCLAL FVIGIGCERA YGKFKLLAIY IISGIGAALF 60
Q7MZP5/40-183    PRYAQGLIGI PLSPFLHGSW GHLFSNLPAL LVLSALLMAH SIRYYVIASL FIIFVGGTLV 60
Consensus        PTLKFEFWRY FTHAFLHFSL MHILFNLLWW WYLGGAVEKR LGSGKLIVIT LISALLSGYV
```

## Table 6B. Sequence Alignments for the Enterobacteria



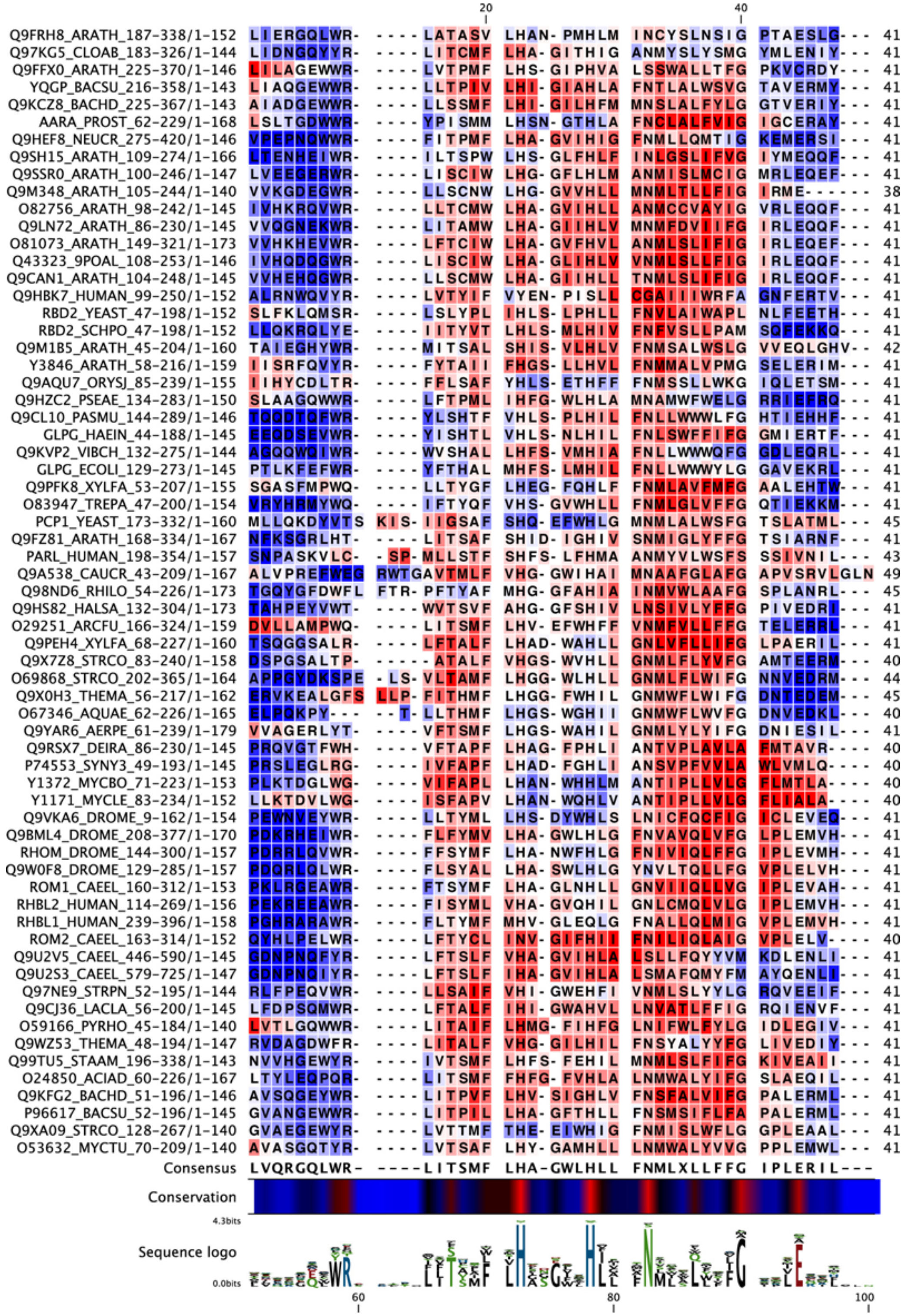
```
Q7CFX8/130-274  QSLFSG---- ---------- ------ANFG GLSGVVYALM GYVWLTGERA PERGISLPRG 100
A4TGR2/130-274  QSLFSG---- ---------- ------ANFG GLSGVVYALM GYVWLTGERA PERGISLPRG 100
Q1C2L2/130-274  QSLFSG---- ---------- ------ANFG GLSGVVYALM GYVWLTGERA PERGISLPRG 100
Q1CCK6/130-274  QSLFSG---- ---------- ------ANFG GLSGVVYALM GYVWLTGERA PERGISLPRG 100
A6BVZ7/130-274  QSLFSG---- ---------- ------ANFG GLSGVVYALM GYVWLTGERA PERGISLPRG 100
A7FNW6/130-274  QSLFSG---- ---------- ------ANFG GLSGVVYALM GYVWLTGERA PERGISLPRG 100
Q664J1/130-274  QSLFSG---- ---------- ------ANFG GLSGVVYALM GYVWLTGERA PEHGISLPRG 100
A1JSF0/130-274  QSLFSG---- ---------- ------VNFG GLSGVVYALM GYVWLTGERA PERGISLPRG 100
A8GKU2/130-274  QSLFSG---- ---------- ------ALFG GLSGVVYALM GYCWLSGERA PERGLMLPRG 100
Q7N9W4/131-275  QSLFSG---- ---------- ------SHFG GLSGVVYALI GYVWLTGERA PERGIGVPRG 100
Q6CZL3/130-274  QSWCSG---- ---------- ------TYFG GLSGVVYALM GYVWLRGERE PDGYLSMPRS 100
Q2NQH3/132-276  QSHFSG---- ---------- ------IFFG GLSGVVYALM GYVWWHGEKN PGGPLFMPLG 100
A6TF43/129-273  QHQFSG---- ---------- ------PWFG GLSGVVYALM GYVWLRGERD PQSGIYLQRG 100
A4WFK8/129-273  QHKFSG---- ---------- ------PWFG GLSGVVYALM GYAWLRGERD PESGIYMQRG 100
Q57IV1/129-273  QQKFSG---- ---------- ------PWFG GLSGVVYALM GYVWLRGERD PQSGIYLQRG 100
Q8ZLH5/129-273  QQKFSG---- ---------- ------PWFG GLSGVVYALM GYVWLRGERD PQSGIYLQRG 100
Q8Z229/129-273  QQKFSG---- ---------- ------PWFG GLSGVVYALM GYVWLRGERD PQSGIYLQRG 100
Q5PLZ8/129-273  QQKFSG---- ---------- ------PWFG GLSGVVYALM GYVWLRGERD PQSGIYLQRG 100
A8AQX4/129-273  QQKFSG---- ---------- ------PWFG GLSGVVYALM GYAWLRGERD PQSGIYLQRG 100
Q31VK9/129-273  QQKFSG---- ---------- ------PWFG GLSGVVYALM GYVWLRGERD PQSGIYLQRG 100
A2UGP1/129-273  QQKFSG---- ---------- ------PWFG GLSGVVYALM GYVWLRGERD PQSGIYLQRG 100
GLPG_ECOLI/1-276  QQKFSG---- ---------- ------PWFG GLSGVVYALM GYVWLRGERD PQSGIYLQRG 100
Q0SZP2/110-254  QQKFSG---- ---------- ------PWFG GLSGVVYALM GYVWLRGERD PQSGIYLQRG 100
A1AGU7/129-273  QQKFSG---- ---------- ------PWFG GLSGVVYALM GYVWLRGERD PQSGIYLQRG 100
Q32AN6/129-273  QQKFSG---- ---------- ------PWFG GLSGVVYALM GYVWLRGERD PQSGIYLQRG 100
Q0TC44/129-273  QQKFSG---- ---------- ------PWFG GLSGVVYALM GYVWLRGERD PQSGIYLQRG 100
A7ZSV4/129-273  QQKFSG---- ---------- ------PWFG GLSGVVYALM GYVWLRGERD PQSGIYLQRG 100
Q83PV6/127-271  QQKFSG---- ---------- ------PWFG GLSGVVYALM GYVWLRGERD PQSGIYLQRG 100
A8A5N2/129-273  QQKFSG---- ---------- ------PWFG GLSGVVYALM GYVWLRGERD PQSGIYLQRG 100
Q1R5L1/129-273  QQKFSG---- ---------- ------PWFG GLSGVVYALM GYVWLRGERD PQSGIYLQRG 100
P09391/129-273  QQKFSG---- ---------- ------PWFG GLSGVVYALM GYVWLRGERD PQSGIYLQRG 100
Q3YWA4/129-273  QQKFSG---- ---------- ------PWFG GLSGVVYALM GYVWLRGERD PQSGIYLQRG 100
Q8FCS5/129-273  QQKFSG---- ---------- ------PWFG GLSGVVYALM GYVWLRGERD PQSGIYLQRG 100
Q8X6Z6/129-273  QQKFSG---- ---------- ------PWFG GLSGVVYALM GYVWLRGERD PQSGIYLQRG 100
A7MGE5/129-273  QYQIAG---- ---------- ------PWFG GLGGVVYALV GYVWLRGERE PESGLYLPRG 100
Q324P0/55-206   SACYALRDSD Q--------- -IVISVGASG AIMGIAGAAI ATQLASGAG- --TYHKNQRR 107
A2UJQ7/55-206   SACYALRDSD Q--------- -IVISVGASG AIMGIAGAAI ATQLASGTG- --THHKNQRR 107
Q3Z4D6/55-206   SACYALRESE Q--------- -IVISVGASG AILGIAGAAI ATQFASGTG- --TYHKNQRR 107
A7ZJ48/55-209   SACYALRESE Q--------- -IVISVGASG AIMGIAGAAI ATQLASGAG- --THHKNQRR 107
Q8FJY5/55-206   SACYALRESE Q--------- -IVISVGASG AIMGIAGAAI ATQLASGAG- --THHKNQRR 107
A1A8S4/55-206   SACYTLRESE Q--------- -IVISIGASG AIMGIAGAAI ATQLASGAG- --THHKNQRR 107
Q0TK24/55-206   SACYALRESE Q--------- -IVISVGASG AIMGIAGAAI ATQLASGAG- --IHHKNQRR 107
Q1RER1/55-206   SACYTLRESE Q--------- -IVISIGASG AIMGIAGAAI ATQLASGAG- --THHKNQRR 107
Q7N1V7/52-196   STNFVSYP-- ---------- -----WNIGT GASQAVLGVS SFALLLVFVK EHTSSILKFA 103
P46116/62-229   SAYWQYYEIS NSDLWTDSTV YITIGVGASG AIMGIAAASV IYLIKVVINK PNPHPVIQRR 120
Q7MZP5/40-183   WLFGRS---- ---------- ------AIHI GASGWIFGLW GLLLANAFFL RRVKDFFFAI 100
Consensus       QQKFSG---- ---------- ------PWFG GLSGVVYALM GYVWLRGERD PQSGIYLQRG
```

# Table 6C. Sequence Alignments for the Enterobacteria



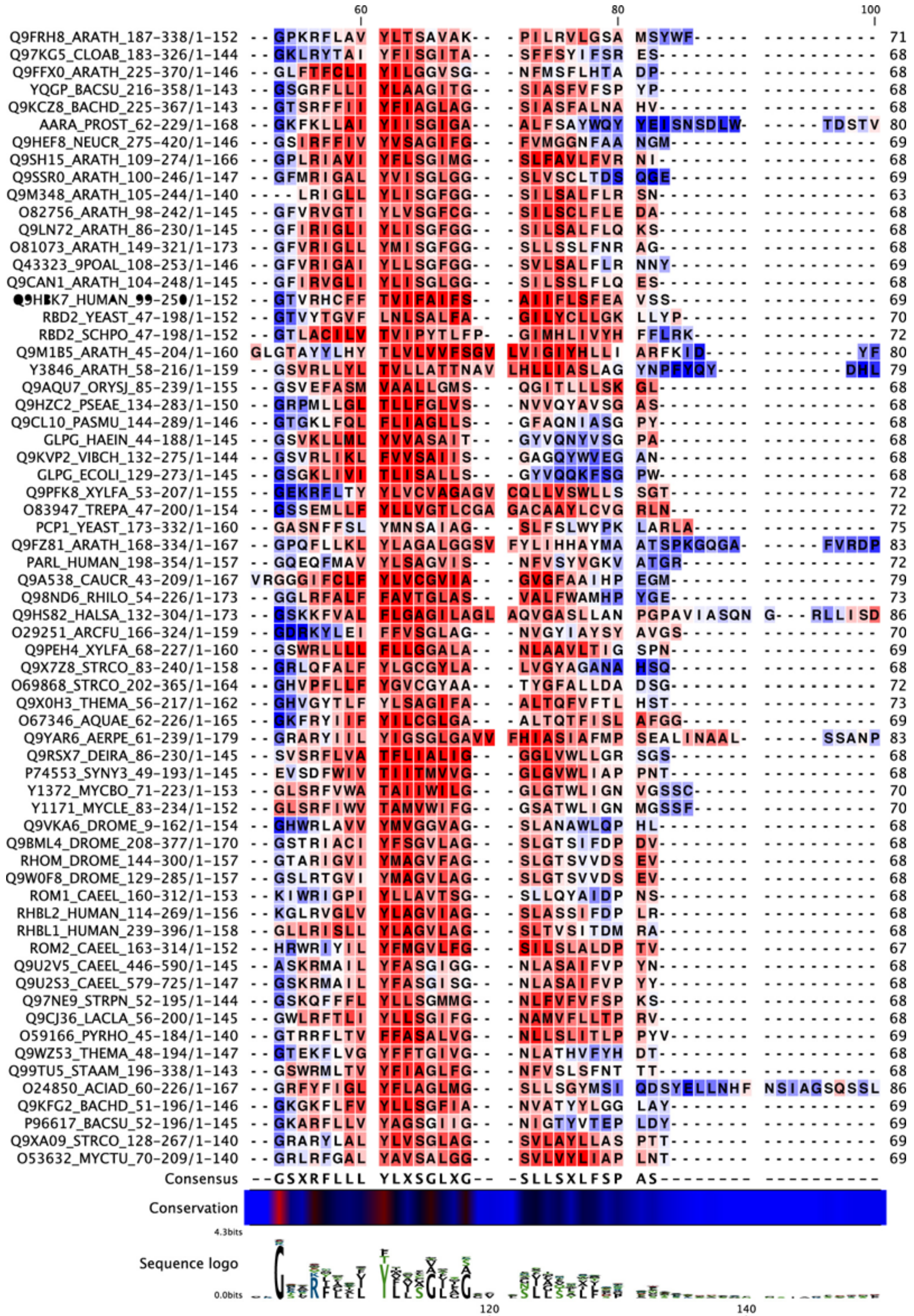
```
                                           140                  160
                                             |                    |
Q7CFX8/130-274     --LMAFSVLW LIAG--YFDI LGLSIANAAH VSGLIIGLLM AFWDTRNSA 145
A4TGR2/130-274     --LMAFSVLW LIAG--YFDI LGLSIANAAH VSGLIIGLLM AFWDTRNSA 145
Q1C2L2/130-274     --LMAFSVLW LIAG--YFDI LGLSIANAAH VSGLIIGLLM AFWDTRNSA 145
Q1CCK6/130-274     --LMAFSVLW LIAG--YFDI LGLSIANAAH VSGLIIGLLM AFWDTRNSA 145
A6BVZ7/130-274     --LMAFSVLW LIAG--YFDI LGLSIANAAH VSGLIIGLLM AFWDTRNSA 145
A7FNW6/130-274     --LMAFSVLW LIAG--YFDI LGLSIANAAH VSGLIIGLLM AFWDTRNSA 145
Q664J1/130-274     --LMAFSVLW LIAG--YFDI LGLSIANAAH VSGLIIGLLM AFWDTRNSA 145
A1JSF0/130-274     --LMAFSVLW LVAG--YFDI LGLSIANAAH VSGLIIGLLM AFWDTRNSA 145
A8GKU2/130-274     --LMVFSVLW LVAG--YFDI LGMSIANAAH VAGLVLGLLM AFWDTRHRA 145
Q7N9W4/131-275     --LMAFSLFW LIVG--YFDA FGLSIANAAH FSGLIIGLLM ALWDNRHTF 145
Q6CZL3/130-274     --LMAFALLW LVAG--YFDI LGMSIANAAH VAGLIVGLLM AFWDTYNKT 145
Q2NQH3/132-276     --VIVFALLW LIAG--YFNI LGIAIANAAH VAGLVTGLLM AFWQTRRGQ 145
A6TF43/129-273     --LILFSLVW LIAG--WFDV FGMAIANGAH VAGLATGLAM AFVDTLHGR 145
A4WFK8/129-273     --LITFALLW LIAG--WFDL FGMSIANGAH VTGLAVGLAM AFADTLNAR 145
Q57IV1/129-273     --LIIFALLW IVAG--WFDW FGMSMANGAH IAGLIVGLAM AFVDTLNAR 145
Q8ZLH5/129-273     --LIIFALLW IVAG--WFDW FGMSMANGAH IAGLIVGLAM AFVDTLNAR 145
Q8Z229/129-273     --LIIFALLW IVAG--WFDW FGMSMANGAH IAGLIVGLAM AFVDTLNAR 145
Q5PLZ8/129-273     --LIIFALLW IVAS--WFDW FGMSMANGAH IAGLIVGLAM AFVDTLNAR 145
A8AQX4/129-273     --LIAFALIW IVAG--WFDV FGMSMANGAH IAGLAVGLAM AFADTVNAR 145
Q31VK9/129-273     --LIIFALIW IIAG--WFDL FGMSMANGAH IAGLAVGLAM AFVDSLNAR 145
A2UGP1/129-273     --LIIFALIW IVAG--WFDL FGMSMANGAH IAGLAVGLAM AFVDSLNAR 145
GLPG_ECOLI/1-276   --LIIFALIW IVAG--WFDL FGMSMANGAH IAGLAVGLAM AFVDSLNAR 145
Q0SZP2/110-254     --LIIFALIW IVAG--WFDL FGMSMANGAH IAGLAVGLAM AFVDSLNAR 145
A1AGU7/129-273     --LIIFALIW IVAG--WFDL FGMSMANGAH IAGLAVGLAM AFVDSLNAR 145
Q32AN6/129-273     --LIIFALIW IVAG--WFDL FGMSMANGAH IAGLAVGLAM AFVDSLNAR 145
Q0TC44/129-273     --LIIFALIW IVAG--WFDL FGMSMANGAH IAGLAVGLAM AFVDSLNAR 145
A7ZSV4/129-273     --LIIFALIW IVAG--WFDL FGMSMANGAH IAGLAVGLAM AFVDSLNAR 145
Q83PV6/127-271     --LIIFALIW IVAG--WFDL FGMSMANGAH IAGLAVGLAM AFVDSLNAR 145
A8A5N2/129-273     --LIIFALIW IVAG--WFDL FGMSMANGAH IAGLAVGLAM AFVDSLNAR 145
Q1R5L1/129-273     --LIIFALIW IVAG--WFDL FGMSMANGAH IAGLAVGLAM AFVDSLNAR 145
P09391/129-273     --LIIFALIW IVAG--WFDL FGMSMANGAH IAGLAVGLAM AFVDSLNAR 145
Q3YWA4/129-273     --LIIFALIW IVAG--WFDL FGMSMANGAH IAGLAVGLAM AFVDSLNAR 145
Q8FCS5/129-273     --LIIFALIW IVAG--WFDL FGMSMANGAH IAGLAVGLAM AFVDSLNAR 145
Q8X6Z6/129-273     --LIIFALIW IVAG--WFDL FGMSMANGAH IAGLAVGLAM AFVDSLNAR 145
A7MGE5/129-273     --ILVFMLLW LAIG--GLGL FGNKTANADL VAGMLIGLAM AMTDTLHAR 145
Q324P0/55-206      ----VFSLLG MVALTLLYGA RQTGIDNACH IGGLIAGGAL GWLSARLSG 152
A2UJQ7/55-206      ----VFPLLG MVALTLLYGA RQTGIDNACH IGGLIAGGAL GWLSARLSG 152
Q3Z4D6/55-206      ----VFPLLG MVALTLLYGA RQTGIDNACH IGGLIAGGAL GWLSARLSG 152
A7ZJ48/55-209      ----VFPLLG MVALTLLYGA RQTGIDNACH IGGLIAGGAL GWLSACLSG 152
Q8FJY5/55-206      ----VFPLLG MVALTLLYGT RQTGIDNACH IGGLIAGGAL GWLSARLVG 152
A1A8S4/55-206      ----VFPLLG MVALTLLYGT RQTGIDNACH IGGLIAGGAL GWLSARLVG 152
Q0TK24/55-206      ----VFPLLG MVALTLLYGT RQAGIDNACH IGGLIAGGAL GWLSARLVG 152
Q1RER1/55-206      ----VFPLLG MVALTLLYGT RQTGIDNACH IGGLIAGGAL GWLSARLVG 152
Q7N1V7/52-196      ---VVFSMLP AIAL----DF IYAHYPKPGH VLSICIGLTM SLFFYRKNK 145
P46116/62-229      QKYQLYNLIA MIALTLING- LQSGVDNAAH IGGAIIG-AL ISIAYILVP 167
Q7MZP5/40-183      ---LVFVYYG GIVFG--LLP WQTNISTEGH IFGAISGLVF SWLSKRWVI 144
Consensus          --LIIFALLW IVAG--WFDL FGMSIANXAH IAGLIVGLAM AFVDTLNAR
```

# Table 7A. Sequence Alignments for the PFAM Seed



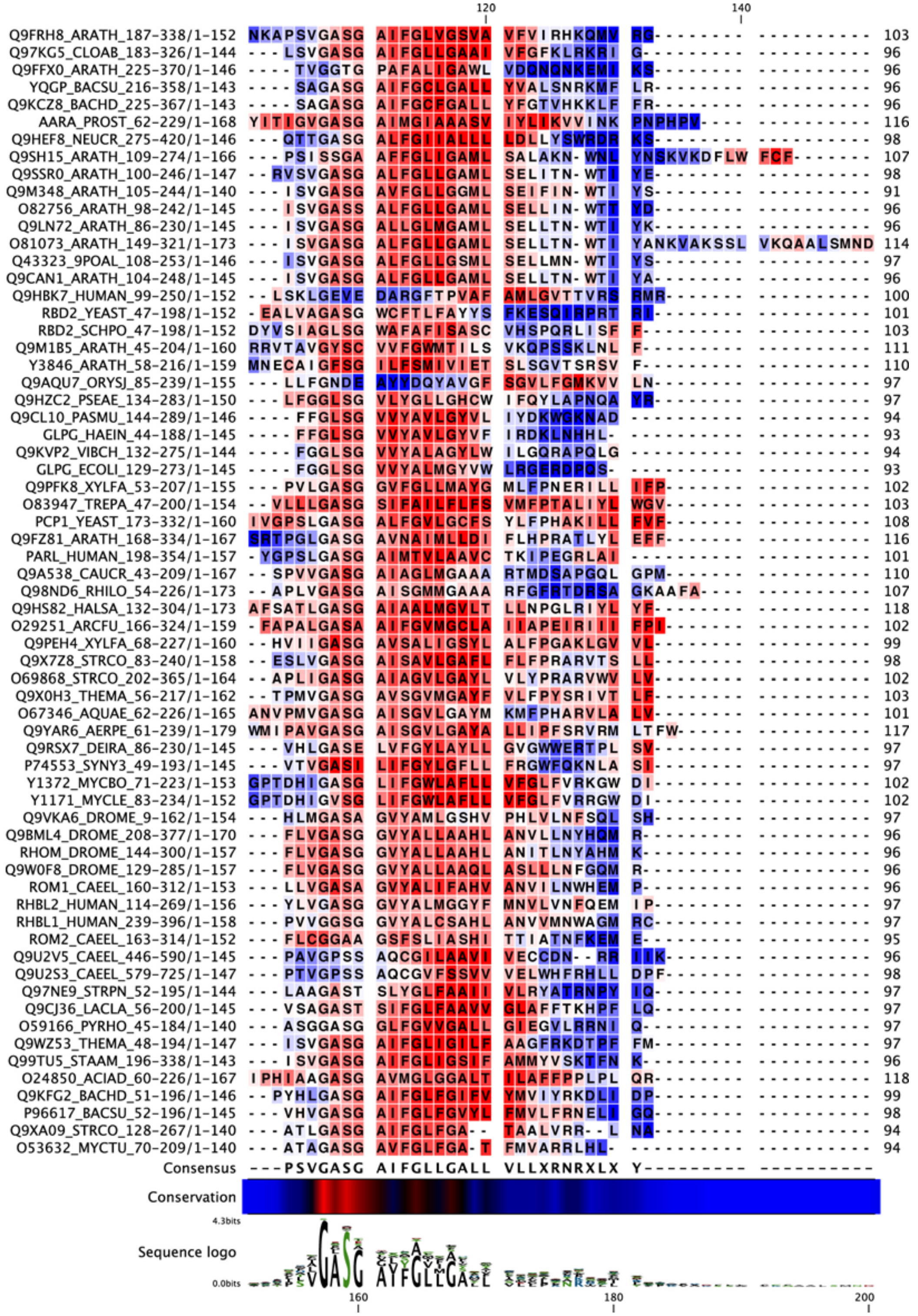
```
                                                   20                   40
                                                    |                    |
Q9FRH8_ARATH_187-338/1-152  LIERGQLWR- ----LATASV LHAN-PMHLM INCYSLNSIG PTAESLG---  41
Q97KG5_CLOAB_183-326/1-144  LIDNGQYYR- ----LITCMF LHA-GITHIG ANMYSLYSMG YMLENIY---  41
Q9FFX0_ARATH_225-370/1-146  LILAGEWWR- ----LVTPMF LHS-GIPHVA LSSWALLTFG PKVCRDY---  41
 YQGP_BACSU_216-358/1-143   LIAQGEWWR- ----LLTPIV LHI-GIAHLA FNTLALWSVG TAVERMY---  41
Q9KCZ8_BACHD_225-367/1-143  AIADGEWWR- ----LLSSMF LHI-GILHFM MNSLALFYLG GTVERIY---  41
  AARA_PROST_62-229/1-168   LSLTGDWWR- ----YPISMM LHSN-GTHLA FNCLALFVIG IGCERAY---  41
Q9HEF8_NEUCR_275-420/1-146  VPEPNQWWR- ----FITPMF LHA-GVIHIG FNMLLQMTIG KEMERSI---  41
Q9SH15_ARATH_109-274/1-166  LTENHEIWR- ----ILTSPW LHS-GLFHLF INLGSLIFVG IYMEQQF---  41
Q9SSR0_ARATH_100-246/1-147  LVEEGERWR- ----LISCIW LHG-GFLHLM ANMISLMCIG MRLEQEF---  41
Q9M348_ARATH_105-244/1-140  VVKGDEGWR- ----LLSCNW LHG-GVVHLL MNMLTLLFIG IRME------  38
 O82756_ARATH_98-242/1-145  IVHKRQVWR- ----LLTCMW LHA-GVIHLL ANMCCVAYIG VRLEQQF---  41
 Q9LN72_ARATH_86-230/1-145  VVQGNEKWR- ----LITAMW LHA-GIIHLV MNMFDVIIFG IRLEQQF---  41
O81073_ARATH_149-321/1-173  VVHKHEVWR- ----LFTCIW LHA-GVFHVL ANMLSLIFIG IRLEQEF---  41
Q43323_9POAL_108-253/1-146  IVHQDQGWR- ----LISCIW LHA-GLIHLV VNMLSLLFIG IRLEQQF---  41
Q9CAN1_ARATH_104-248/1-145  VVHEHQGWR- ----LLSCMW LHA-GIIHLL TNMLSLIFIG IRLEQQF---  41
Q9HBK7_HUMAN_99-250/1-152   ALRNWQVYR- ----LVTYIF VYEN-PISLL CGAIIIWRFA GNFERTV---  41
  RBD2_YEAST_47-198/1-152   SLFKLQMSR- ----LSLYPL IHLS-LPHLL FNVLAIWAPL NLFEETH---  41
  RBD2_SCHPO_47-198/1-152   LLQKRQLYE- ----IITYVT LHLS-MLHIV FNFVSLLPAM SQFEKKQ---  41
 Q9M1B5_ARATH_45-204/1-160  TAIEGHYWR- ----MITSAL SHIS-VLHLV FNMSALWSLG VVEQLGHV--  42
   Y3846_ARATH_58-216/1-159 IISRFQVYR- ----FYTAII FHGS-LLHVL FNMMALVPMG SELERIM---  41
 Q9AQU7_ORYSJ_85-239/1-155  IIHYCDLTR- ----FFLSAF YHLS-ETHFF FNMSSLLWKG IQLETSM---  41
Q9HZC2_PSEAE_134-283/1-150  SLAAGQWWR- ----LFTPML IHFG-WLHLA MNAMWFWELG RRIEFRQ---  41
Q9CL10_PASMU_144-289/1-146  TQQDTQFWR- ----YLSHTF VHLS-PLHIL FNLLWWWLFG HTIEHHF---  41
   GLPG_HAEIN_44-188/1-145  EEQDSEVWR- ----YISHTL VHLS-NLHIL FNLSWFFIFG GMIERTF---  41
Q9KVP2_VIBCH_132-275/1-144  AGQQWQIWR- ----WVSHAL LHFS-VMHIA FNLLWWWQFG GDLEQRL---  41
  GLPG_ECOLI_129-273/1-145  PTLKFEFWR- ----YFTHAL MHFS-LMHIL FNLLWWWYLG GAVEKRL---  41
 Q9PFK8_XYLFA_53-207/1-155  SGASFMPWQ- ----LLTYGF LHEG-FQHLF FNMLAVFMFG AALEHTW---  41
 O83947_TREPA_47-200/1-154  VRYHRMYWQ- ----IFTYQF VHS-GVWHLL FNMLGLVFFG QTIEKKM---  41
  PCP1_YEAST_173-332/1-160  MLLQKDYVTS KIS-IIGSAF SHQ-EFWHLG MNMLALWSFG TSLATML---  45
Q9FZ81_ARATH_168-334/1-167  NFKSGRLHT- ----LITSAF SHID-IGHIV SNMIGLYFFG TSIARNF---  41
  PARL_HUMAN_198-354/1-157  SNPASKVLC- -SP-MLLSTF SHFS-LFHMA ANMYVLWSFS SSIVNIL---  43
 Q9A538_CAUCR_43-209/1-167  ALVPREFWEG RWTGAVTMLF VHG-GWIHAI MNAAFGLAFG APVSRVLGLN  49
 Q98ND6_RHILO_54-226/1-173  TGQYGFDWFL FTR-PFTYAF MHG-GFAHIA INMVWLAAFG SPLANRL---  45
Q9HS82_HALSA_132-304/1-173  TAHPEYVWT- ----WVTSVF AHG-GFSHIV LNSIVLYFFG PIVEDRI---  41
O29251_ARCFU_166-324/1-159  DVLLAMPWQ- ----LITSMF LHV-EFWHFF VNMFVLLFFG TELERRL---  41
 Q9PEH4_XYLFA_68-227/1-160  TSQGGSALR- ----LFTALF LHAD-WAHLL GNLVFLLIFG LPAERIL---  41
 Q9X7Z8_STRCO_83-240/1-158  DSPGSALTP- -----ATALF VHGS-WVHLL GNMLFLYVFG AMTEERM---  40
O69868_STRCO_202-365/1-164  APPGYDKSPE -LS-VLTAMF LHGG-WLHLL GNMLFLWIFG NNVEDRM---  44
 Q9X0H3_THEMA_56-217/1-162  ERVKEALGFS LLP-FITHMF LHGG-FWHIL GNMWFLWIFG DNTEDEM---  45
 O67346_AQUAE_62-226/1-165  ELPQKPY--- --T-LLTHMF LHGS-WGHII GNMWFLWVFG DNVEDKL---  40
 Q9YAR6_AERPE_61-239/1-179  VVAGERLYT- ----VFTSMF LHGS-WAHIL GNMLYLYIFG DNIESIL---  41
 Q9RSX7_DEIRA_86-230/1-145  PRQVGTFWH- ----VFTAPF LHAG-FPHLI ANTVPLAVLA FMTAVR----  40
 P74553_SYNY3_49-193/1-145  PRSLEGLRG- ----IVFAPF LHAD-FGHLI ANSVPFVVLA WLVMLQ----  40
  Y1372_MYCBO_71-223/1-153  PLKTDGLWG- ----VIFAPL LHAN-WHHLM ANTIPLLVLG FLMTLA----  40
  Y1171_MYCLE_83-234/1-152  LLKTDVLWG- ----ISFAPV LHAN-WQHLV ANTIPLLVLG FLIALA----  40
  Q9VKA6_DROME_9-162/1-154  PEWNVEYWR- ----LLTYML LHS-DYWHLS LNICFQCFIG ICLEVEQ---  41
Q9BML4_DROME_208-377/1-170  PDKRHEIWR- ----FLFYMV LHA-GWLHLG FNVAVQLVFG LPLEMVH---  41
  RHOM_DROME_144-300/1-157  PDRRLQVWR- ----FFSYMF LHA-NWFHLG FNIVIQLFFG IPLEVMH---  41
Q9W0F8_DROME_129-285/1-157  PDQRLQLWR- ----FLSYAL LHA-SWLHLG YNVLTQLLFG VPLELVH---  41
  ROM1_CAEEL_160-312/1-153  PKLRGEAWR- ----FTSYMF LHA-GLNHLL GNVIIQLLVG IPLEVAH---  41
 RHBL2_HUMAN_114-269/1-156  PEKREEAWR- ----FISYML VHA-GVQHIL GNLCMQLVLG IPLEMVH---  41
 RHBL1_HUMAN_239-396/1-158  PGHRARAWR- ----FLTYMF MHV-GLEQLG FNALLQLMIG VPLEMVH---  41
  ROM2_CAEEL_163-314/1-152  QYHLPELWR- ----LFTYCL INV-GIFHII FNILIQLAIG VPLELV----  40
 Q9U2V5_CAEEL_446-590/1-145 GDNPNQFYR- ----LFTSLF VHA-GVIHLA LSLLFQYYVM KDLENLI---  41
 Q9U2S3_CAEEL_579-725/1-147 GDNPNQIYR- ----LFTSLF IHA-GVIHLA LSMAFQMYFM AYQENLI---  41
 Q97NE9_STRPN_52-195/1-144  RLFPEQVWR- ----LLSAIF VHI-GWEHFI VNMLSLYYLG RQVEEIF---  41
 Q9CJ36_LACLA_56-200/1-145  LFDPSQMWR- ----LFTALF IHI-GWAHVL LNVATLFFIG RQIENVF---  41
 O59166_PYRHO_45-184/1-140  LVTLGQWWR- ----LITAIF LHMG-FIHFG LNIFWLFYLG IDLEGIV---  41
 Q9WZ53_THEMA_48-194/1-147  RVDAGDWFR- ----LITALF VHG-GILHIL FNSYALYYFG LIVEDIY---  41
 Q99TU5_STAAM_196-338/1-143 NVVHGEWYR- ----IVTSMF LHFS-FEHIL MNMLSLFIFG KIVEAII---  41
 O24850_ACIAD_60-226/1-167  LTYLEQPQR- ----LITSMF FHFG-FVHLA LNMWALYIFG SLAEQIL---  41
 Q9KFG2_BACHD_51-196/1-146  AVSQGEYWR- ----LITPVF LHV-SIGHLV FNSFALVIFG PALERML---  41
 P96617_BACSU_52-196/1-145  GVANGEWWR- ----LITPIL LHA-GFTHLL FNSMSIFLFA PALERML---  41
Q9XA09_STRCO_128-267/1-140  GVAEGEWYR- ----LVTTMF THE-EIWHIG FNMISLWFLG GPLEAAL---  41
 O53632_MYCTU_70-209/1-140  AVASGQTYR- ----LVTSAF LHY-GAMHLL LNMWALYVVG PPLEMWL---  41
                 Consensus  LVQRGQLWR- ----LITSMF LHA-GWLHLL FNMLXLLFFG IPLERIL---
```

# Table 7B. Sequence Alignments for the PFAM Seed



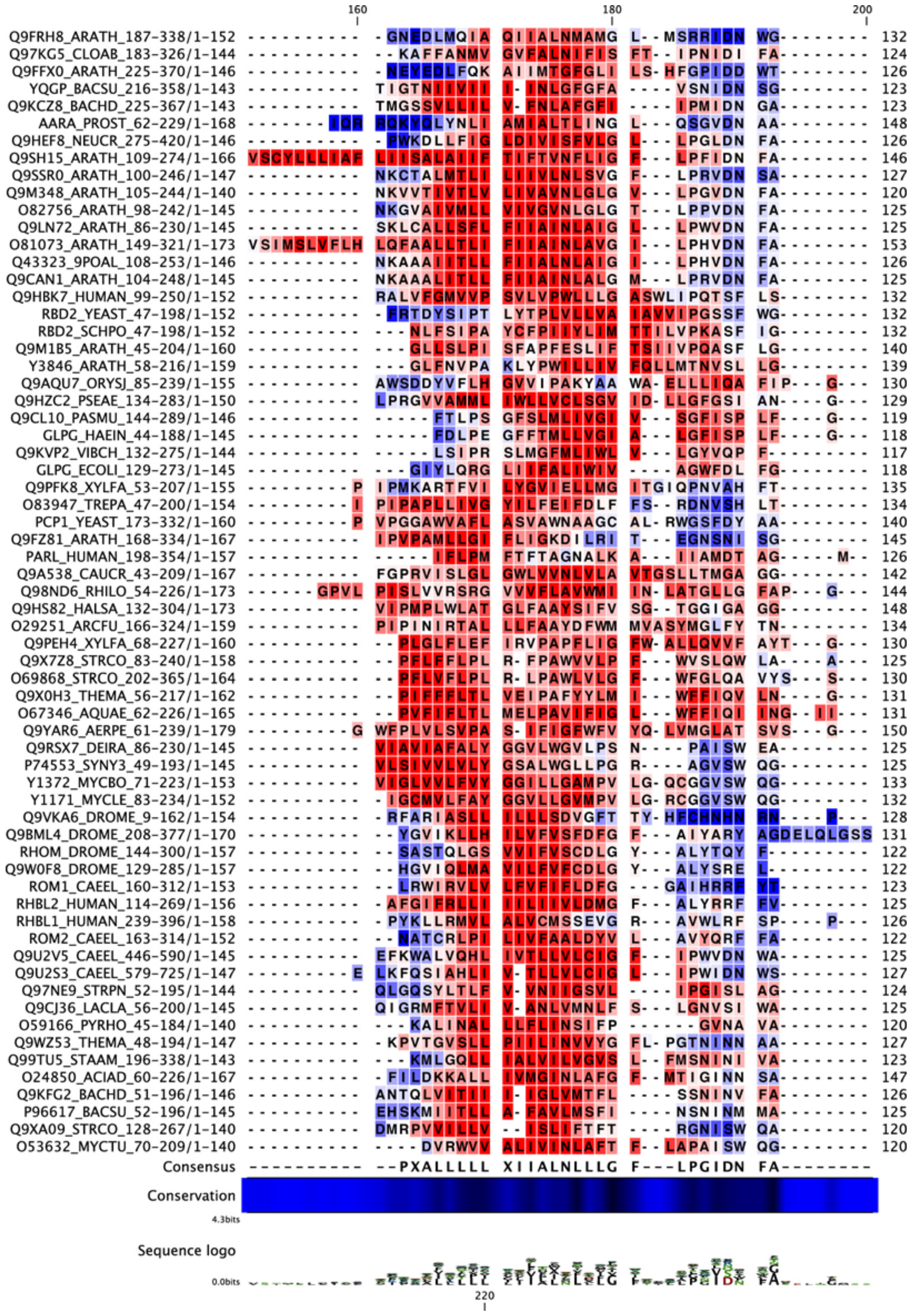
```
Q9FRH8_ARATH_187-338/1-152   --GPKRFLAV YLTSAVAK-- -PILRVLGSA MSYWF----- ---------- 71
Q97KG5_CLOAB_183-326/1-144   --GKLRYTAI YFISGITA-- -SFFSYIFSR ES-------- ---------- 68
Q9FFX0_ARATH_225-370/1-146   --GLFTFCLI YILGGVSG-- -NFMSFLHTA DP-------- ---------- 68
YQGP_BACSU_216-358/1-143     --GSGRFLLI YLAAGITG-- -SIASFVFSP YP-------- ---------- 68
Q9KCZ8_BACHD_225-367/1-143   --GTSRFFII YFIAGLAG-- -SIASFALNA HV-------- ---------- 68
AARA_PROST_62-229/1-168      --GKFKLLAI YIISGIGA-- -ALFSAYWQY YEISNSDLW- -----TDSTV 80
Q9HEF8_NEUCR_275-420/1-146   --GSIRFFIV YVSAGIFG-- -FVMGGNFAA NGM------- ---------- 69
Q9SH15_ARATH_109-274/1-166   --GPLRIAVI YFLSGIMG-- -SLFAVLFVR NI-------- ---------- 68
Q9SSR0_ARATH_100-246/1-147   --GFMRIGAL YVISGLGG-- -SLVSCLTDS QGE------- ---------- 69
Q9M348_ARATH_105-244/1-140   ----LRIGLL YLISGFGG-- -SILSALFLR SN-------- ---------- 63
O82756_ARATH_98-242/1-145    --GFVRVGTI YLVSGFCG-- -SILSCLFLE DA-------- ---------- 68
Q9LN72_ARATH_86-230/1-145    --GFIRIGLI YLISGFGG-- -SILSALFLQ KS-------- ---------- 68
O81073_ARATH_149-321/1-173   --GFVRIGLL YMISGFGG-- -SLLSSLFNR AG-------- ---------- 68
Q43323_9POAL_108-253/1-146   --GFVRIGAI YLLSGFGG-- -SVLSALFLR NNY------- ---------- 69
Q9CAN1_ARATH_104-248/1-145   --GFIRVGLI YLISGLGG-- -SILSSLFLQ ES-------- ---------- 68
Q9HBK7_HUMAN_99-250/1-152    --GTVRHCFF TVIFAIFS-- -AIIFLSFEA VSS------- ---------- 69
RBD2_YEAST_47-198/1-152      --GTVYTGVF LNLSALFA-- -GILYCLLGK LLYP------ ---------- 70
RBD2_SCHPO_47-198/1-152      --GTLACILV TVIPYTLFP- -GIMHLIVYH FFLRK----- ---------- 72
Q9M1B5_ARATH_45-204/1-160    GLGTAYYLHY TLVLVVFSGV LVIGIYHLLI ARFKID---- --------YF 80
Y3846_ARATH_58-216/1-159     --GSVRLLYL TVLLATTNAV LHLLIASLAG YNPFYQY--- -------DHL 79
Q9AQU7_ORYSJ_85-239/1-155    --GSVEFASM VAALLGMS-- -QGITLLLSK GL-------- ---------- 68
Q9HZC2_PSEAE_134-283/1-150   --GRPMLLGL TLLFGLVS-- -NVVQYAVSG AS-------- ---------- 68
Q9CL10_PASMU_144-289/1-146   --GTGKLFQL FLIAGLLS-- -GFAQNIASG PY-------- ---------- 68
GLPG_HAEIN_44-188/1-145      --GSVKLLML YVVASAIT-- -GYVQNYVSG PA-------- ---------- 68
Q9KVP2_VIBCH_132-275/1-144   --GSVRLIKL FVVSAIIS-- -GAGQYWVEG AN-------- ---------- 68
GLPG_ECOLI_129-273/1-145     --GSGKLIVI TLISALLS-- -GYVQQKFSG PW-------- ---------- 68
Q9PFK8_XYLFA_53-207/1-155    --GEKRFLTY YLVCVAGAGV CQLLVSWLLS SGT------- ---------- 72
O83947_TREPA_47-200/1-154    --GSSEMLLF YLLVGTLCGA GACAAYLCVG RLN------- ---------- 72
PCP1_YEAST_173-332/1-160     --GASNFFSL YMNSAIAG-- -SLFSLWYPK LARLA----- ---------- 75
Q9FZ81_ARATH_168-334/1-167   --GPQFLLKL YLAGALGGSV FYLIHHAYMA ATSPKGQGA- -----FVRDP 83
PARL_HUMAN_198-354/1-157     --GQEQFMAV YLSAGVIS-- -NFVSYVGKV ATGR------ ---------- 72
Q9A538_CAUCR_43-209/1-167    VRGGGIFCLF YLVCGVIA-- -GVGFAAIHP EGM------- ---------- 79
Q98ND6_RHILO_54-226/1-173    --GGLRFALF FAVTGLAS-- -VALFWAMHP YGE------- ---------- 73
Q9HS82_HALSA_132-304/1-173   --GSKKFVAL FLGAGILAGL AQVGASLLAN PGPAVIASQN G---RLLISD 86
O29251_ARCFU_166-324/1-159   --GDRKYLEI FFVSGLAG-- -NVGYIAYSY AVGS------ ---------- 70
Q9PEH4_XYLFA_68-227/1-160    --GSWRLLLL FLLGGALA-- -NLAAVLTIG SPN------- ---------- 69
Q9X7Z8_STRCO_83-240/1-158    --GRLQFALF YLGCGYLA-- -LVGYAGANA HSQ------- ---------- 68
O69868_STRCO_202-365/1-164   --GHVPFLLF YGVCGYAA-- -TYGFALLDA DSG------- ---------- 72
Q9X0H3_THEMA_56-217/1-162    --GHVGYTLF YLSAGIFA-- -ALTQFVFTL HST------- ---------- 73
O67346_AQUAE_62-226/1-165    --GKFRYIIF YILCGLGA-- -ALTQTFISL AFGG------ ---------- 69
Q9YAR6_AERPE_61-239/1-179    --GRARYIIL YIGSGLGAVV FHIASIAFMP SEALINAAL- -----SSANP 83
Q9RSX7_DEIRA_86-230/1-145    --SVSRFLVA TFLIALIG-- -GGLVWLLGR SGS------- ---------- 68
P74553_SYNY3_49-193/1-145    --EVSDFWIV TIITMVVG-- -GLGVWLIAP PNT------- ---------- 68
Y1372_MYCBO_71-223/1-153     --GLSRFVWA TAIIWILG-- -GLGTWLIGN VGSSC----- ---------- 70
Y1171_MYCLE_83-234/1-152     --GLSRFIWV TAMVWIFG-- -GSATWLIGN MGSSF----- ---------- 70
Q9VKA6_DROME_9-162/1-154     --GHWRLAVV YMVGGVAG-- -SLANAWLQP HL-------- ---------- 68
Q9BML4_DROME_208-377/1-170   --GSTRIACI YFSGVLAG-- -SLGTSIFDP DV-------- ---------- 68
RHOM_DROME_144-300/1-157     --GTARIGVI YMAGVFAG-- -SLGTSVVDS EV-------- ---------- 68
Q9W0F8_DROME_129-285/1-157   --GSLRTGVI YMAGVLAG-- -SLGTSVVDS EV-------- ---------- 68
ROM1_CAEEL_160-312/1-153     --KIWRIGPI YLLAVTSG-- -SLLQYAIDP NS-------- ---------- 68
RHBL2_HUMAN_114-269/1-156    --KGLRVGLV YLAGVIAG-- -SLASSIFDP LR-------- ---------- 68
RHBL1_HUMAN_239-396/1-158    --GLLRISLL YLAGVLAG-- -SLTVSITDM RA-------- ---------- 68
ROM2_CAEEL_163-314/1-152     --HRWRIYIL YFMGVLFG-- -SILSLALDP TV-------- ---------- 67
Q9U2V5_CAEEL_446-590/1-145   --ASKRMAIL YFASGIGG-- -NLASAIFVP YN-------- ---------- 68
Q9U2S3_CAEEL_579-725/1-147   --GSKRMAIL YFASGISG-- -NLASAIFVP YY-------- ---------- 68
Q97NE9_STRPN_52-195/1-144    --GSKQFFFL YLLSGMMG-- -NLFVFVFSP KS-------- ---------- 68
Q9CJ36_LACLA_56-200/1-145    --GWLRFTLI YLLSGIFG-- -NAMVFLLTP RV-------- ---------- 68
O59166_PYRHO_45-184/1-140    --GTRRFLTV FFASALVG-- -NLLSLITLP PYV------- ---------- 69
Q9WZ53_THEMA_48-194/1-147    --GTEKFLVG YFFTGIVG-- -NLATHVFYH DT-------- ---------- 68
Q99TU5_STAAM_196-338/1-143   --GSWRMLTV YFIAGLFG-- -NFVSLSFNT TT-------- ---------- 68
O24850_ACIAD_60-226/1-167    --GRFYFIGL YFLAGLMG-- -SLLSGYMSI QDSYELLNHF NSIAGSQSSL 86
Q9KFG2_BACHD_51-196/1-146    --GKGKFLFV YLLSGFIA-- -NVATYYLGG LAY------- ---------- 69
P96617_BACSU_52-196/1-145    --GKARFLLV YAGSGIIG-- -NIGTYVTEP LDY------- ---------- 69
Q9XA09_STRCO_128-267/1-140   --GRARYLAL YLVSGLAG-- -SVLAYLLAS PTT------- ---------- 69
O53632_MYCTU_70-209/1-140    --GRLRFGAL YAVSALGG-- -SVLVYLIAP LNT------- ---------- 69
Consensus                    --GSXRFLLL YLXSGLXG-- -SLLSXLFSP AS-------- ----------
```

# Table 7C. Sequence Alignments for the PFAM Seed



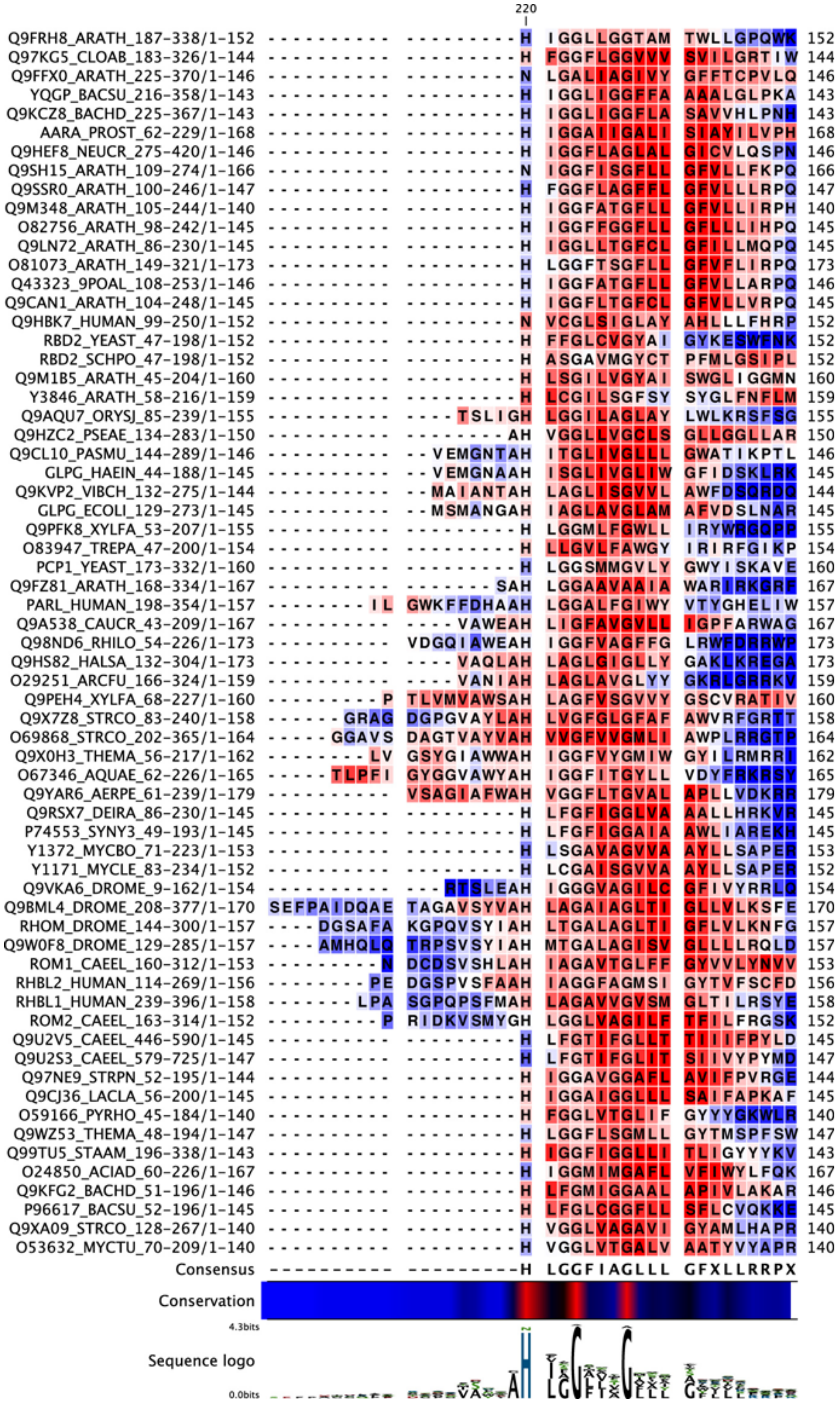
```
Q9FRH8_ARATH_187-338/1-152  NKAPSVGASG AIFGLVGSVA VFVIRHKQMV RG-------- ---------- 103
Q97KG5_CLOAB_183-326/1-144  ---LSVGASG AIFGLLGAAI VFGFKLRKRI G--------- ---------- 96
Q9FFX0_ARATH_225-370/1-146  ----TVGGTG PAFALIGAWL VDQNQNKEMI KS-------- ---------- 96
YQGP_BACSU_216-358/1-143    ----SAGASG AIFGCLGALL YVALSNRKMF LR-------- ---------- 96
Q9KCZ8_BACHD_225-367/1-143  ----SAGASG AIFGCFGALL YFGTVHKKLF FR-------- ---------- 96
AARA_PROST_62-229/1-168     YITIGVGASG AIMGIAAASV IYLIKVVINK PNPHPV---- ---------- 116
Q9HEF8_NEUCR_275-420/1-146  ---QTTGASG ALFGIIALLL LDLLYSWRDR KS-------- ---------- 98
Q9SH15_ARATH_109-274/1-166  ---PSISSGA AFFGLIGAML SALAKN-WNL YNSKVKDFLW FCF------- 107
Q9SSR0_ARATH_100-246/1-147  --RVSVGASG ALFGLLGAML SELITN-WTI YE-------- ---------- 98
Q9M348_ARATH_105-244/1-140  ---ISVGASG AVFGLLGGML SEIFIN-WTI YS-------- ---------- 91
O82756_ARATH_98-242/1-145   ---ISVGASS ALFGLLGAML SELLIN-WTT YD-------- ---------- 96
Q9LN72_ARATH_86-230/1-145   ---ISVGASG ALLGLMGAML SELLTN-WTI YK-------- ---------- 96
O81073_ARATH_149-321/1-173  ---ISVGASG ALFGLLGAML SELLTN-WTI YANKVAKSSL VKQAALSMND 114
Q43323_9POAL_108-253/1-146  ---ISVGASG ALFGLLGSML SELLMN-WTI YS-------- ---------- 97
Q9CAN1_ARATH_104-248/1-145  ---ISVGASG ALFGLLGAML SELLTN-WTI YA-------- ---------- 96
Q9HBK7_HUMAN_99-250/1-152   --LSKLGEVE DARGFTPVAF AMLGVTTVRS RMR------- ---------- 100
RBD2_YEAST_47-198/1-152     -EALVAGASG WCFTLFAYYS FKESQIRPRT RI-------- ---------- 101
RBD2_SCHPO_47-198/1-152     DYVSIAGLSG WAFAFISASC VHSPQRLISF F--------- ---------- 103
Q9M1B5_ARATH_45-204/1-160   RRVTAVGYSC VVFGWMTILS VKQPSSKLNL F--------- ---------- 111
Y3846_ARATH_58-216/1-159    MNECAIGFSG ILFSMIVIET SLSGVTSRSV F--------- ---------- 110
Q9AQU7_ORYSJ_85-239/1-155   ---LLFGNDE AYYDQYAVGF SGVLFGMKVV LN-------- ---------- 97
Q9HZC2_PSEAE_134-283/1-150  ---LFGGLSG VLYGLLGHCW IFQYLAPNQA YR-------- ---------- 97
Q9CL10_PASMU_144-289/1-146  ----FFGLSG VVYAVLGYVL IYDKWGKNAD ---------- ---------- 94
GLPG_HAEIN_44-188/1-145     ----FFGLSG VVYAVLGYVF IRDKLNHHL- ---------- ---------- 93
Q9KVP2_VIBCH_132-275/1-144  ----FGGLSG VVYALAGYLW ILGQRAPQLG ---------- ---------- 94
GLPG_ECOLI_129-273/1-145    ----FGGLSG VVYALMGYVW LRGERDPQS- ---------- ---------- 93
Q9PFK8_XYLFA_53-207/1-155   ---PVLGASG GVFGLLMAYG MLFPNERILL IFP------- ---------- 102
O83947_TREPA_47-200/1-154   --VLLLGASG SIFAILFLFS VMFPTALIYL WGV------- ---------- 103
PCP1_YEAST_173-332/1-160    IVGPSLGASG ALFGVLGCFS YLFPHAKILL FVF------- ---------- 108
Q9FZ81_ARATH_168-334/1-167  SRTPGLGASG AVNAIMLLDI FLHPRATLYL EFF------- ---------- 116
PARL_HUMAN_198-354/1-157    -YGPSLGASG AIMTVLAAVC TKIPEGRLAI ---------- ---------- 101
Q9A538_CAUCR_43-209/1-167   --SPVVGASG AIAGLMGAAA RTMDSAPGQL GPM------- ---------- 110
Q98ND6_RHILO_54-226/1-173   --APLVGASG AISGMMGAAA RFGFRTDRSA GKAAFA---- ---------- 107
Q9HS82_HALSA_132-304/1-173  AFSATLGASG AIAALMGVLT LLNPGLRIYL YF-------- ---------- 118
O29251_ARCFU_166-324/1-159  -FAPALGASA AIFGVMGCLA IIAPEIRIII FPI------- ---------- 102
Q9PEH4_XYLFA_68-227/1-160   --HVIIGASG AVSALIGSYL ALFPGAKLGV VL-------- ---------- 99
Q9X7Z8_STRCO_83-240/1-158   --ESLVGASG AISAVLGAFL FLFPRARVTS LL-------- ---------- 98
O69868_STRCO_202-365/1-164  --APLIGASG AIAGVLGAYL VLYPRARVWV LV-------- ---------- 102
Q9X0H3_THEMA_56-217/1-162   --TPMVGASG AVSGVMGAYF VLFPYSRIVT LF-------- ---------- 103
O67346_AQUAE_62-226/1-165   ANVPMVGASG AISGVLGAYM KMFPHARVLA LV-------- ---------- 101
Q9YAR6_AERPE_61-239/1-179   WMIPAVGASG AISGVLGAYA LLIPFSRVRM LTFW------ ---------- 117
Q9RSX7_DEIRA_86-230/1-145   ---VHLGASE LVFGYLAYLL GVGWWERTPL SV-------- ---------- 97
P74553_SYNY3_49-193/1-145   ---VTVGASI LIFGYLGFLL FRGWFQKNLA SI-------- ---------- 97
Y1372_MYCBO_71-223/1-153    GPTDHIGASG LIFGWLAFLL VFGLFVRKGW DI-------- ---------- 102
Y1171_MYCLE_83-234/1-152    GPTDHIGVSG LIFGWLAFLL VFGLFVRRGW DI-------- ---------- 102
Q9VKA6_DROME_9-162/1-154    ---HLMGASA GVYAMLGSHV PHLVLNFSQL SH-------- ---------- 97
Q9BML4_DROME_208-377/1-170  ---FLVGASG GVYALLAAHL ANVLLNYHQM R--------- ---------- 96
RHOM_DROME_144-300/1-157    ---FLVGASG GVYALLAAHL ANITLNYAHM K--------- ---------- 96
Q9W0F8_DROME_129-285/1-157  ---FLVGASG GVYALLAAQL ASLLLNFGQM R--------- ---------- 96
ROM1_CAEEL_160-312/1-153    ---LLVGASA GVYALIFAHV ANVILNWHEM P--------- ---------- 96
RHBL2_HUMAN_114-269/1-156   ---YLVGASG GVYALMGGYF MNVLVNFQEM IP-------- ---------- 97
RHBL1_HUMAN_239-396/1-158   ---PVVGGSG GVYALCSAHL ANVVMNWAGM RC-------- ---------- 97
ROM2_CAEEL_163-314/1-152    ---FLCGGAA GSFSLIASHI TTIATNFKEM E--------- ---------- 95
Q9U2V5_CAEEL_446-590/1-145  ---PAVGPSS AQCGILAAVI VECCDN--RR IIK------- ---------- 96
Q9U2S3_CAEEL_579-725/1-147  ---PTVGPSS AQCGVFSSVV VELWHFRHLL DPF------- ---------- 98
Q97NE9_STRPN_52-195/1-144   ---LAAGAST SLYGLFAAII VLRYATRNPY IQ-------- ---------- 97
Q9CJ36_LACLA_56-200/1-145   ---VSAGAST SIFGLFAAVV GLAFFTKHPF LQ-------- ---------- 97
O59166_PYRHO_45-184/1-140   ---ASGGASG GLFGVVGALL GIEGVLRRNI Q--------- ---------- 97
Q9WZ53_THEMA_48-194/1-147   ---ISVGASG AIFGLIGILF AAGFRKDTPF FM-------- ---------- 97
Q99TU5_STAAM_196-338/1-143  ---ISVGASG AIFGLIGSIF AMMYVSKTFN K--------- ---------- 96
O24850_ACIAD_60-226/1-167   IPHIAAGASG AVMGLGGALT ILAFFPPLPL QR-------- ---------- 118
Q9KFG2_BACHD_51-196/1-146   --PYHLGASG AIFGLFGIFV YMVIYRKDLI DP-------- ---------- 99
P96617_BACSU_52-196/1-145   ---VHVGASG AIFGLFGVYL FMVLFRNELI GQ-------- ---------- 98
Q9XA09_STRCO_128-267/1-140  ---ATLGASG AIFGLFGA-- TAALVRR--L NA-------- ---------- 94
O53632_MYCTU_70-209/1-140   ---ATAGASG AVFGLFGA-T FMVARRLHL- ---------- ---------- 94
Consensus                   ---PSVGASG AIFGLLGALL VLLXRNRXLX Y--------- ----------
```

# Table 7D. Sequence Alignments for the PFAM Seed



```
Q9FRH8_ARATH_187-338/1-152   ---------- -GNEDLMQIA QIIALNMAMG L--MSRRIDN WG-------- 132
Q97KG5_CLOAB_183-326/1-144   ---------- --KAFFANMV GVFALNIFIS FT--IPNIDI FA-------- 124
Q9FFX0_ARATH_225-370/1-146   ---------- -NEYEDLFQK AIIMTGFGLI LS-HFGPIDD WT-------- 126
YQGP_BACSU_216-358/1-143     ---------- TIGTNIIVII I-INLGFGFA ----VSNIDN SG-------- 123
Q9KCZ8_BACHD_225-367/1-143   ---------- TMGSSVLLIL V-FNLAFGFI ----IPMIDN GA-------- 123
AARA_PROST_62-229/1-168      -------IQR RQKYQLYNLI AMIALTLING L---QSGVDN AA-------- 148
Q9HEF8_NEUCR_275-420/1-146   ---------- -PWKDLLFIG LDIVISFVLG L---LPGLDN FA-------- 126
Q9SH15_ARATH_109-274/1-166   VSCYLLLIAF LIISALAIIF TIFTVNFLIG F---LPFIDN FA-------- 146
Q9SSR0_ARATH_100-246/1-147   ---------- NKCTALMTLI LIIVLNLSVG F---LPRVDN SA-------- 127
Q9M348_ARATH_105-244/1-140   ---------- NKVVTIVTLV LIVAVNLGLG V---LPGVDN FA-------- 120
O82756_ARATH_98-242/1-145    ---------- NKGVAIVMLL VIVGVNLGLG T---LPPVDN FA-------- 125
Q9LN72_ARATH_86-230/1-145    ---------- SKLCALLSFL FIIAINLAIG L---LPWVDN FA-------- 125
O81073_ARATH_149-321/1-173   VSIMSLVFLH LQFAALLTLI FIIAINLAVG I---LPHVDN FA-------- 153
Q43323_9POAL_108-253/1-146   ---------- NKAAAIITLL FIIALNLAIG I---LPHVDN FA-------- 126
Q9CAN1_ARATH_104-248/1-145   ---------- NKAAALITLL FIIAINLALG M---LPRVDN FA-------- 125
Q9HBK7_HUMAN_99-250/1-152    ---------- RALVFGMVVP SVLVPWLLLG ASWLIPQTSF LS-------- 132
RBD2_YEAST_47-198/1-152      ---------- -FRTDYSIPT LYTPLVLLVA IAVVIPGSSF WG-------- 132
RBD2_SCHPO_47-198/1-152      ---------- ---NLFSIPA YCFPIIYLIM TTILVPKASF IG-------- 132
Q9M1B5_ARATH_45-204/1-160    ---------- ---GLLSLPI SFAPFESLIF TSIIVPQASF LG-------- 140
Y3846_ARATH_58-216/1-159     ---------- ---GLFNVPA KLYPWILLIV FQLLMTNVSL LG-------- 139
Q9AQU7_ORYSJ_85-239/1-155    ---------- AWSDDYVFLH GVVIPAKYAA WA-ELLLIQA FIP---G--- 130
Q9HZC2_PSEAE_134-283/1-150   ---------- LPRGVVAMML IWLLVCLSGV ID-LLGFGSI AN----G--- 129
Q9CL10_PASMU_144-289/1-146   ---------- -----FTLPS GFSLMLIVGI V---SGFISP LF----G--- 119
GLPG_HAEIN_44-188/1-145      ---------- -----FDLPE GFFTMLLVGI A---LGFISP LF----G--- 118
Q9KVP2_VIBCH_132-275/1-144   ---------- -----LSIPR SLMGFMLIWL V---LGYVQP F--------- 117
GLPG_ECOLI_129-273/1-145     ---------- ---GIYLQRG LIIFALIWIV ----AGWFDL FG-------- 118
Q9PFK8_XYLFA_53-207/1-155    ---------P IPMKARTFVI LYGVIELLMG ITGIQPNVAH FT-------- 135
O83947_TREPA_47-200/1-154    ---------I PIPAPLLIVG YILFEIFDLF FS--RDNVSH LT-------- 134
PCP1_YEAST_173-332/1-160     ---------P VPGGAWVAFL ASVAWNAAGC AL-RWGSFDY AA-------- 140
Q9FZ81_ARATH_168-334/1-167   ---------- IPVPAMLLGI FLIGKDILRI T---EGNSNI SG-------- 145
PARL_HUMAN_198-354/1-157     ---------- -----IFLPM FTFTAGNALK A---IIAMDT AG-----M-- 126
Q9A538_CAUCR_43-209/1-167    ---------- FGPRVISLGL GWLVVNLVLA VTGSLLTMGA GG-------- 142
Q98ND6_RHILO_54-226/1-173    ------GPVL PISLVVRSRG VVVFLAVWMI IN-LATGLLG FAP---G--- 144
Q9HS82_HALSA_132-304/1-173   ---------- VIPMPLWLAT GLFAAYSIFV SG--TGGIGA GG-------- 148
O29251_ARCFU_166-324/1-159   ---------- PIPINIRTAL LLFAAYDFWM MVASYMGLFY TN-------- 134
Q9PEH4_XYLFA_68-227/1-160    ---------- --PLGLFLEF IRVPAPFLIG FW-ALLQVVF AYT---G--- 130
Q9X7Z8_STRCO_83-240/1-158    ---------- --PFLFFLPL R-FPAWVVLP F---WVSLQW LA----A--- 125
O69868_STRCO_202-365/1-164   ---------- --PFLVFLPL R-LPAWLVLG F---WFGLQA VYS---S--- 130
Q9X0H3_THEMA_56-217/1-162    ---------- --PIFFFLTL VEIPAFYYLM I---WFFIQV LN----G--- 131
O67346_AQUAE_62-226/1-165    ---------- --PVFIFLTL MELPAVIFIG L---WFFIQI ING--II--- 131
Q9YAR6_AERPE_61-239/1-179    ---------G WFPLVLSVPA S-IFIGFWFV YQ-LVMGLAT SVS---G--- 150
Q9RSX7_DEIRA_86-230/1-145    ---------- VIAVIAFALY GGVLWGVLPS N----PAISW EA-------- 125
P74553_SYNY3_49-193/1-145    ---------- VLSIVVLVLY GSALWGLLPG R----AGVSW QG-------- 125
Y1372_MYCBO_71-223/1-153     ---------- VIGLVVLFVY GGILLGAMPV LG-QCGGVSW QG-------- 133
Y1171_MYCLE_83-234/1-152     ---------- -IGCMVLFAY GGVLLGVMPV LG-RCGGVSW QG-------- 132
Q9VKA6_DROME_9-162/1-154     ---------- -RFARIASLL ILLLSDVGFT TY-HFCHNHN RN----P--- 128
Q9BML4_DROME_208-377/1-170   ---------- --YGVIKLLH ILVFVSFDFG F---AIYARY AGDELQLGSS 131
RHOM_DROME_144-300/1-157     ---------- --SASTQLGS VVIFVSCDLG Y---ALYTQY F--------- 122
Q9W0F8_DROME_129-285/1-157   ---------- --HGVIQLMA VILFVFCDLG Y---ALYSRE L--------- 122
ROM1_CAEEL_160-312/1-153     ---------- --LRWIRVLV LFVFIFLDFG ---GAIHRRF YT-------- 123
RHBL2_HUMAN_114-269/1-156    ---------- -AFGIFRLLI IILIIVLDMG F---ALYRRF FV-------- 125
RHBL1_HUMAN_239-396/1-158    ---------- -PYKLLRMVL ALVCMSSEVG R---AVWLRF SP----P--- 126
ROM2_CAEEL_163-314/1-152     ---------- --NATCRLPI LIVFAALDYV L---AVYQRF FA-------- 122
Q9U2V5_CAEEL_446-590/1-145   ---------- EFKWALVQHL IVTLLVLCIG F---IPWVDN WA-------- 125
Q9U2S3_CAEEL_579-725/1-147   ---------E LKFQSIAHLI V-TLLVLCIG L---IPWIDN WS-------- 127
Q97NE9_STRPN_52-195/1-144    ---------- QLGQSYLTLF V-VNIIGSVL ----IPGISL AG-------- 124
Q9CJ36_LACLA_56-200/1-145    ---------- QIGRMFTVLI V-ANLVMNLF S---LGNVSI WA-------- 125
O59166_PYRHO_45-184/1-140    ---------- ---KALINAL LLFLINSIFP ------GVNA VA-------- 120
Q9WZ53_THEMA_48-194/1-147    ---------- -KPVTGVSLL PIILINVVYG FL-PGTNINN AA-------- 127
Q99TU5_STAAM_196-338/1-143   ---------- ---KMLGQLL IALVILVGVS L--FMSNINI VA-------- 123
O24850_ACIAD_60-226/1-167    ---------- -FILDKKALL IVMGINLAFG F--MTIGINN SA-------- 147
Q9KFG2_BACHD_51-196/1-146    ---------- ANTQLVITII I-IGLVMTFL ----SSNINV FA-------- 126
P96617_BACSU_52-196/1-145    ---------- EHSKMIITLL A-FAVLMSFI ----NSNINM MA-------- 125
Q9XA09_STRCO_128-267/1-140   ---------- DMRPVVILLV --ISLIFTFT ----RGNISW QA-------- 120
O53632_MYCTU_70-209/1-140    ---------- ----DVRWVV ALIVINLAFT F--LAPAISW QG-------- 120
Consensus                    ---------- --PXALLLLL XIIALNLLLG F---LPGIDN FA--------
```

## Table 7E. Sequence Alignments for the PFAM Seed



```
                                                    220
                                                      |
Q9FRH8_ARATH_187-338/1-152   ---------- ---------H IGGLLGGTAM TWLLGPQWK 152
Q97KG5_CLOAB_183-326/1-144   ---------- ---------H FGGFLGGVVV SVILGRTIW 144
Q9FFX0_ARATH_225-370/1-146   ---------- ---------N LGALIAGIVY GFFTCPVLQ 146
YQGP_BACSU_216-358/1-143     ---------- ---------H IGGLIGGFFA AAALGLPKA 143
Q9KCZ8_BACHD_225-367/1-143   ---------- ---------H IGGLIGGFLA SAVVHLPNH 143
AARA_PROST_62-229/1-168      ---------- ---------H IGGAIIGALI SIAYILVPH 168
Q9HEF8_NEUCR_275-420/1-146   ---------- ---------H IGGFLAGLAL GICVLQSPN 146
Q9SH15_ARATH_109-274/1-166   ---------- ---------N IGGFISGFLL GFVLLFKPQ 166
Q9SSR0_ARATH_100-246/1-147   ---------- ---------H FGGFLAGFFL GFVLLLRPQ 147
Q9M348_ARATH_105-244/1-140   ---------- ---------H IGGFATGFLL GFVLLIRPH 140
O82756_ARATH_98-242/1-145    ---------- ---------H IGGFFGGFLL GFLLLIHPQ 145
Q9LN72_ARATH_86-230/1-145    ---------- ---------H IGGLLTGFCL GFILLMQPQ 145
O81073_ARATH_149-321/1-173   ---------- ---------H LGGFTSGFLL GFVFLIRPQ 173
Q43323_9POAL_108-253/1-146   ---------- ---------H IGGFATGFLL GFVLLARPQ 146
Q9CAN1_ARATH_104-248/1-145   ---------- ---------H IGGFLTGFCL GFVLLVRPQ 145
Q9HBK7_HUMAN_99-250/1-152    ---------- ---------N VCGLSIGLAY AHLLLFHRP 152
RBD2_YEAST_47-198/1-152      ---------- ---------H FFGLCVGYAI GYKESWFNK 152
RBD2_SCHPO_47-198/1-152      ---------- ---------H ASGAVMGYCT PFMLGSIPL 152
Q9M1B5_ARATH_45-204/1-160    ---------- ---------H LSGILVGYAI SWGLIGGMN 160
Y3846_ARATH_58-216/1-159     ---------- ---------H LCGILSGFSY SYGLFNFLM 159
Q9AQU7_ORYSJ_85-239/1-155    ---------- ----TSLIGH LGGILAGLAY LWLKRSFSG 155
Q9HZC2_PSEAE_134-283/1-150   ---------- --------AH VGGLLVGCLS GLLGGLLAR 150
Q9CL10_PASMU_144-289/1-146   ---------- --VEMGNTAH ITGLIVGLLL GWATIKPTL 146
GLPG_HAEIN_44-188/1-145      ---------- --VEMGNAAH ISGLIVGLIW GFIDSKLRK 145
Q9KVP2_VIBCH_132-275/1-144   ---------- --MAIANTAH LAGLISGVVL AWFDSQRDQ 144
GLPG_ECOLI_129-273/1-145     ---------- --MSMANGAH IAGLAVGLAM AFVDSLNAR 145
Q9PFK8_XYLFA_53-207/1-155    ---------- ---------H LGGMLFGWLL IRYWRGQPP 155
O83947_TREPA_47-200/1-154    ---------- ---------H LLGVLFAWGY IRIRFGIKP 154
PCP1_YEAST_173-332/1-160     ---------- ---------H LGGSMMGVLY GWYISKAVE 160
Q9FZ81_ARATH_168-334/1-167   ---------- -------SAH LGGAAVAAIA WARIRKGRF 167
PARL_HUMAN_198-354/1-157     --------IL GWKFFDHAAH LGGALFGIWY VTYGHELIW 157
Q9A538_CAUCR_43-209/1-167    ---------- ----VAWEAH LIGFAVGVLL IGPFARWAG 167
Q98ND6_RHILO_54-226/1-173    ---------- VDGQIAWEAH IGGFVAGFFG LRWFDRRWP 173
Q9HS82_HALSA_132-304/1-173   ---------- ----VAQLAH LAGLGIGLLY GAKLKREGA 173
O29251_ARCFU_166-324/1-159   ---------- ----VANIAH LAGLAVGLYY GKRLGRRKV 159
Q9PEH4_XYLFA_68-227/1-160    ---------P TLVMVAWSAH LAGFVSGVVY GSCVRATIV 160
Q9X7Z8_STRCO_83-240/1-158    ------GRAG DGPGVAYLAH LVGFGLGFAF AWVRFGRTT 158
O69868_STRCO_202-365/1-164   -----GGAVS DAGTVAYVAH VVGFVVGMLI AWPLRRGTP 164
Q9X0H3_THEMA_56-217/1-162    --------LV GSYGIAWWAH IGGFVYGMIW GYILRMRRI 162
O67346_AQUAE_62-226/1-165    -----TLPFI GYGGVAWYAH IGGFITGYLL VDYFRKRSY 165
Q9YAR6_AERPE_61-239/1-179    ---------- VSAGIAFWAH VGGFLTGVAL APLLVDKRR 179
Q9RSX7_DEIRA_86-230/1-145    ---------- ---------H LFGFIGGLVA AALLHRKVR 145
P74553_SYNY3_49-193/1-145    ---------- ---------H LFGFIGGAIA AWLIAREKH 145
Y1372_MYCBO_71-223/1-153     ---------- ---------H LSGAVAGVVA AYLLSAPER 153
Y1171_MYCLE_83-234/1-152     ---------- ---------H LCGAISGVVA AYLLSAPER 152
Q9VKA6_DROME_9-162/1-154     ---------- ---RTSLEAH IGGGVAGILC GFIVYRRLQ 154
Q9BML4_DROME_208-377/1-170   SEFPAIDQAE TAGAVSYVAH LAGAIAGLTI GLLVLKSFE 170
RHOM_DROME_144-300/1-157     ----DGSAFA KGPQVSYIAH LTGALAGLTI GFLVLKNFG 157
Q9W0F8_DROME_129-285/1-157   ----AMHQLQ TRPSVSYIAH MTGALAGISV GLLLLRQLD 157
ROM1_CAEEL_160-312/1-153     ---------N DCDSVSHLAH IAGAVTGLFF GYVVLYNVV 153
RHBL2_HUMAN_114-269/1-156    --------PE DGSPVSFAAH IAGGFAGMSI GYTVFSCFD 156
RHBL1_HUMAN_239-396/1-158    -------LPA SGPQPSFMAH LAGAVVGVSM GLTILRSYE 158
ROM2_CAEEL_163-314/1-152     ---------P RIDKVSMYGH LGGLVAGILF TFILFRGSK 152
Q9U2V5_CAEEL_446-590/1-145   ---------- ---------H LFGTIFGLLT TIIIFPYLD 145
Q9U2S3_CAEEL_579-725/1-147   ---------- ---------H LFGTIFGLIT SIIVYPYMD 147
Q97NE9_STRPN_52-195/1-144    ---------- ---------H IGGAVGGAFL AVIFPVRGE 144
Q9CJ36_LACLA_56-200/1-145    ---------- ---------H IGGAIGGLLL SAIFAPKAF 145
O59166_PYRHO_45-184/1-140    ---------- ---------H FGGLVTGLIF GYYYGKWLR 140
Q9WZ53_THEMA_48-194/1-147    ---------- ---------H LGGFLSGMLL GYTMSPFSW 147
Q99TU5_STAAM_196-338/1-143   ---------- ---------H IGGFIGGLLI TLIGYYYKV 143
O24850_ACIAD_60-226/1-167    ---------- ---------H IGGMIMGAFL VFIWYLFQK 167
Q9KFG2_BACHD_51-196/1-146    ---------- ---------H LFGMIGGAAL APIVLAKAR 146
P96617_BACSU_52-196/1-145    ---------- ---------H LFGLCGGFLL SFLCVQKKE 145
Q9XA09_STRCO_128-267/1-140   ---------- ---------H VGGLVAGAVI GYAMLHAPR 140
O53632_MYCTU_70-209/1-140    ---------- ---------H VGGLVTGALV AATYVYAPR 140
Consensus                    ---------- ---------H LGGFIAGLLL GFXLLRRPX
```