



(51) International Patent Classification:

C12N 9/02 (2006.01) C07K 14/195 (2006.01)
C07H 21/00 (2006.01) C12N 9/16 (2006.01)
C12N 9/14 (2006.01) C12N 9/10 (2006.01)
C12N 9/00 (2006.01) C12N 9/12 (2006.01)
C12N 9/52 (2006.01) C12N 9/90 (2006.01)

(21) International Application Number:

PCT/US2011/034852

(22) International Filing Date:

2 May 2011 (02.05.2011)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

12/800,045 5 May 2010 (05.05.2010) US

(71) Applicant (for all designated States except US): **BATTELLE ENERGY ALLIANCE, LLC** [US/US]; P.O. Box 1625, Idaho Falls, ID 83415-3899 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **LEE, Brady, D.** [US/US]; 6093 Gleneagles Dr., Idaho Falls, ID 83401 (US). **NEWBY, Deborah, T.** [US/US]; 3785 Summer Circle, Idaho Falls, ID 83404 (US). **LACEY, Jeffrey, A.** [US/US]; 3963 Willow Ridge Drive, Idaho Falls, ID 83406 (US). **THOMPSON, David, N.** [US/US]; 2362 Hoopes Avenue, Idaho Falls, ID 83404 (US). **THOMPSON, Vicki, S.** [US/US]; 2362 Hoopes Avenue, Idaho Falls, ID 83404 (US). **APEL, William, A.** [US/US]; P.O.

Box 9729, Jackson, WY 83002 (US). **ROBERTO, Francisco, F.** [US/US]; 395 Gustafson Drive, Idaho Falls, ID 83402 (US). **REED, David, W.** [US/US]; 1160 Homer Avenue, Idaho Falls, ID 83040 (US).

(74) Agents: **MORATH, Daniel, J.** et al.; Traskbirt, 230 South 500 East, Suite 300, P.O. Box 2550, Salt Lake City, UT 84110-2550 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: GENETIC ELEMENTS, PROTEINS, AND ASSOCIATED METHODS INCLUDING APPLICATION OF ADDITIONAL GENETIC INFORMATION TO GRAM (+) THERMOACIDOPHILES

FIG. 1

```

gb|AA21825.1|          -----SVSTTVSSTSS-----PST
ref|XP_001743680.1|    -----TTSTTSTTSTTSTNSTNT
ref|XP_001525241.1|    -----
RAAC03697              MLGCVKVGLETGKEAIQFVLNWSCTHFIQRETFMRKTIAAIVTLGTVSSMLAGTVYAAT
ref|YP_308280.1|       -----TGSSFIEGNYVADDGIDL
ref|ZP_01916690.1|     -----SGPLQGSQGGQNTVTA

gb|AA21825.1|          TLSTSTTNQTTNTTASSTPANTTLTSTSPNTQTTTNNSSG---TQTSNTSSTQPPTS
ref|XP_001743680.1|    STTSTTSTTSTTSTTSTTSTTSTTSTTSTTSTTSTSTGNTSS---SAISNTTANSTDT
ref|XP_001525241.1|    --TSTTSTTSTTNTNVPANSTASTTNTST--ANSTTSTANNST---QLAQKAAKYE---
RAAC03697              NITNATSTGNTANTTASATSNATSNITGSAPTNPNVLTQKSAVYQLIEVFTIPSWPTGL
ref|YP_308280.1|       WSTFITGDAVTPADINRPGNISGTVTTGAAPQLPNVLTDAEIA---AIVTDIQNETAFAA
ref|ZP_01916690.1|     PNSSGALTNSAPQNTSTSTSSNSPGSAPANPGIATQKSAPSEPVRSRPTTPAPSGS
                        : : . . . . . :
                        *

gb|AA21825.1|          QSLDTST-----
ref|XP_001743680.1|    TSMSPSRT-NGSDTPSSPAPPA-----
ref|XP_001525241.1|    -----
RAAC03697              QSIDPATQNAQWDWKYKPIPPSQWVGLACVFPDPESKYAADLGIQMTDPNGPVTAQQLA
ref|YP_308280.1|       FSPGAI TRSSDWNIRYYPVPAE-----
ref|ZP_01916690.1|     SGPQPAAPVPSQ-----

gb|AA21825.1|          -----
ref|XP_001743680.1|    -----
ref|XP_001525241.1|    -----
RAAC03697              QWIVDNEVKAR
ref|YP_308280.1|       -----
ref|ZP_01916690.1|     -----

```

(57) Abstract: Isolated and/or purified polypeptides and nucleic acid sequences encoding polypeptides from *Alicyclobacillus acidocaldarius* are provided. Further provided are methods for modulating or altering recombination inside or outside of a cell using isolated and/or purified polypeptides and/or nucleic acid sequences from *Alicyclobacillus acidocaldarius*.



Published:

- without international search report and to be republished upon receipt of that report (Rule 48.2(g))
- with sequence listing part of description (Rule 5.2(a))

- 1 -

**GENETIC ELEMENTS, PROTEINS, AND ASSOCIATED METHODS
INCLUDING APPLICATION OF ADDITIONAL GENETIC
INFORMATION TO GRAM (+) THERMOACIDOPHILES**

5

PRIORITY CLAIM

This application claims the benefit of the filing date of United States Patent Application Serial Number 12/800,045, filed May 5, 2010, for “Genetic Elements, Proteins, and Associated Methods Including Application of Additional Genetic
10 Information to Gram (+) Thermoacidophiles.”

GOVERNMENT RIGHTS

The United States Government has certain rights in this invention pursuant to Contract No. DE-AC07-99ID13727 and Contract No. DE-AC07-05ID14517 between the
15 United States Department of Energy and Battelle Energy Alliance, LLC.

TECHNICAL FIELD

The present invention relates generally to biotechnology. More specifically, embodiments of the present invention relate to isolated and/or purified polypeptides and
20 nucleic acid sequences encoding polypeptides from *Alicyclobacillus acidocaldarius* and methods for their use.

BACKGROUND

Thermophilic and/or acidophilic bacteria have great potential for production of
25 useful chemicals in industrial processes. However, most systems for promoting genetic recombination for the purposes of introducing nucleic acids of interest are not well suited for such thermophilic and/or acidophilic bacteria.

DISCLOSURE OF THE INVENTION

30 Embodiments of the invention relate to purified and/or isolated nucleotide sequences of the genome of *Alicyclobacillus acidocaldarius*, or a homologue or fragment thereof. In one embodiment of the invention, the nucleotide sequence is selected from at least one of SEQ ID Nos. 2, 19, 36, 53, 70, 87, 104, 121, 138, 155, 172, 189, 206, 223, 240,

- 2 -

257, 274, 291, 308, 325, 359, 376, 410, 427, 444, 512, 529, 546, 597,614, 648, 665, 682, 699, 716, 733, 750, 767, 784, 818, 835, 852, 886, 903, 920, 937, 954, 971, 988, 1005, 1039, 1056, 1073, 1090, 1107, 1124, 1141, 1158, 1175, 1192, 1209, 1226, 1243, and 1260, 1277, 1294, 1311, 1328, 1345, 1362, 1379, 1396, 1413, 1430, 1447, 1464, 1481, 1498, 5 1515, 1532, 1549, 1566, 1583, 1600, 1634, 1651, 1685, 1702, 1719, 1787, 1804, 1821, 1872,1889, 1923, 1940, 1957, 1974, 1991, 2008, 2025, 2042, 2059, 2093, 2110, 2127, 2161, 2178, 2195, 2212, 2229, 2246, 2263, 2280, 2314, 2331, 2348, 2365, 2382, 2399, 2416, 2433, 2450, 2467, 2484, 2501, 2518, 2535, 2552, 2569, 2586, 2603, 2620, 2637, 2654, 2671, 2688, 2705, 2722, 2739, 2773, 2790, 2824, 2841, 2858, 2926, 2943, 2960, 10 3011,3028, 3062, 3079, 3096, 3113, 3130, 3147, 3164, 3181, 3198, 3232, 3249, 3266, and 3283 or a homologue or fragment thereof. In another embodiment of the invention, the homologue is selected from the group consisting of a nucleotide sequence having at least 80% sequence identity to at least one of SEQ ID Nos. 2, 19, 36, 53, 70, 87, 104, 121, 138, 155, 172, 189, 206, 223, 240, 257, 274, 291, 308, 325, 359, 376, 410, 427, 444, 512, 529, 15 546, 597,614, 648, 665, 682, 699, 716, 733, 750, 767, 784, 818, 835, 852, 886, 903, 920, 937, 954, 971, 988, 1005, 1039, 1056, 1073, 1090, 1107, 1124, 1141, 1158, 1175, 1192, 1209, 1226, 1243, 1260,1277, 1294, 1311, 1328, 1345, 1362, 1379, 1396, 1413, 1430, 1447, 1464, 1481, 1498, 1515, 1532, 1549, 1566, 1583, 1600, 1634, 1651, 1685, 1702, 1719, 1787, 1804, 1821, 1872,1889, 1923, 1940, 1957, 1974, 1991, 2008, 2025, 2042, 20 2059, 2093, 2110, 2127, 2161, 2178, 2195, 2212, 2229, 2246, 2263, 2280, 2314, 2331, 2348, 2365, 2382, 2399, 2416, 2433, 2450, 2467, 2484, 2501, 2518, 2535, 2552, 2569, 2586, 2603, 2620, 2637, 2654, 2671, 2688, 2705, 2722, 2739, 2773, 2790, 2824, 2841, 2858, 2926, 2943, 2960, 3011,3028, 3062, 3079, 3096, 3113, 3130, 3147, 3164, 3181, 3198, 3232, 3249, 3266, and 3283.

25 Embodiments of the invention may further relate to an isolated and/or purified nucleic acid sequence comprising a nucleic acid sequence encoding a polypeptide selected from the group consisting of a polypeptide having at least 90% sequence identity to at least one of SEQ ID Nos. 1, 18, 35, 52, 69, 86, 103, 120, 137, 154, 171, 188, 205, 222, 239, 256, 273, 290, 307, 324, 358, 375, 409, 426, 443, 511, 528, 545, 596, 613, 647, 664, 681, 698, 30 715, 732, 749, 766, 783, 817, 834, 851, 885, 902, 819, 936, 953, 970, 987, 1004, 1038, 1055, 1072, 1089, 1106, 1123, 1140, 1157, 1174, 1191, 1208, 1225, 1242, 1259, 1276, 1293, 1310, 1327, 1344, 1361, 1378, 1395, 1412, 1429, 1446, 1463, 1480, 1497, 1514,

- 3 -

1531, 1548, 1565, 1582, 1599, 1633, 1650, 1684, 1701, 1718, 1786, 1803, 1820, 1871,
1888, 1922, 1939, 1956, 1973, 1990, 2007, 2024, 2041, 2058, 2092, 2109, 2126, 2160,
2177, 2094, 2211, 2228, 2245, 2262, 2279, 2313, 2330, 2347, 2364, 2381, 2398, 2415,
2432, 2449, 2466, 2483, 2500, 2517, 2534, 2551, 2568, 2585, 2602, 2619, 2636, 2653,
5 2670, 2687, 2704, 2721, 2738, 2772, 2789, 2823, 2840, 2857, 2925, 2942, 2959, 3010,
3027, 3061, 3078, 3095, 3112, 3129, 3146, 3163, 3180, 3197, 3231, 3248, 3265, and 3282.

Embodiments of the invention also relate to isolated and/or purified polypeptides
coded for by a nucleotide sequence comprising a nucleotide sequence of the genome of
Alicyclobacillus acidocaldarius, or a homologue or fragment thereof. In one embodiment,
10 the nucleotide sequence comprises a nucleotide sequence selected from the group
consisting of a nucleotide sequence having at least 80% sequence identity to at least one of
SEQ ID Nos. 2, 19, 36, 53, 70, 87, 104, 121, 138, 155, 172, 189, 206, 223, 240, 257, 274,
291, 308, 325, 359, 376, 410, 427, 444, 512, 529, 546, 597,614, 648, 665, 682, 699, 716,
733, 750, 767, 784, 818, 835, 852, 886, 903, 920, 937, 954, 971, 988, 1005, 1039, 1056,
15 1073, 1090, 1107, 1124, 1141, 1158, 1175, 1192, 1209, 1226, 1243, 1260,1277, 1294,
1311, 1328, 1345, 1362, 1379, 1396, 1413, 1430, 1447, 1464, 1481, 1498, 1515, 1532,
1549, 1566, 1583, 1600, 1634, 1651, 1685, 1702, 1719, 1787, 1804, 1821, 1872,1889,
1923, 1940, 1957, 1974, 1991, 2008, 2025, 2042, 2059, 2093, 2110, 2127, 2161, 2178,
2195, 2212, 2229, 2246, 2263, 2280, 2314, 2331, 2348, 2365, 2382, 2399, 2416, 2433,
20 2450, 2467, 2484, 2501, 2518, 2535, 2552, 2569, 2586, 2603, 2620, 2637, 2654, 2671,
2688, 2705, 2722, 2739, 2773, 2790, 2824, 2841, 2858, 2926, 2943, 2960, 3011,3028,
3062, 3079, 3096, 3113, 3130, 3147, 3164, 3181, 3198, 3232, 3249, 3266, and 3283.

In another embodiment of the invention, the nucleotide sequence comprises a
nucleotide sequence selected from at least one of SEQ ID Nos. 2, 19, 36, 53, 70, 87, 104,
25 121, 138, 155, 172, 189, 206, 223, 240, 257, 274, 291, 308, 325, 359, 376, 410, 427, 444,
512, 529, 546, 597,614, 648, 665, 682, 699, 716, 733, 750, 767, 784, 818, 835, 852, 886,
903, 920, 937, 954, 971, 988, 1005, 1039, 1056, 1073, 1090, 1107, 1124, 1141, 1158,
1175, 1192, 1209, 1226, 1243, and 1260, 1277, 1294, 1311, 1328, 1345, 1362, 1379, 1396,
1413, 1430, 1447, 1464, 1481, 1498, 1515, 1532, 1549, 1566, 1583, 1600, 1634, 1651,
30 1685, 1702, 1719, 1787, 1804, 1821, 1872,1889, 1923, 1940, 1957, 1974, 1991, 2008,
2025, 2042, 2059, 2093, 2110, 2127, 2161, 2178, 2195, 2212, 2229, 2246, 2263, 2280,
2314, 2331, 2348, 2365, 2382, 2399, 2416, 2433, 2450, 2467, 2484, 2501, 2518, 2535,

- 4 -

2552, 2569, 2586, 2603, 2620, 2637, 2654, 2671, 2688, 2705, 2722, 2739, 2773, 2790, 2824, 2841, 2858, 2926, 2943, 2960, 3011, 3028, 3062, 3079, 3096, 3113, 3130, 3147, 3164, 3181, 3198, 3232, 3249, 3266, and 3283 or a homologue or fragment thereof. In still another embodiment, the polypeptide comprises an amino acid sequence of SEQ ID Nos.

5 1, 18, 35, 52, 69, 86, 103, 120, 137, 154, 171, 188, 205, 222, 239, 256, 273, 290, 307, 324, 358, 375, 409, 426, 443, 511, 528, 545, 596, 613, 647, 664, 681, 698, 715, 732, 749, 766, 783, 817, 834, 851, 885, 902, 819, 936, 953, 970, 987, 1004, 1038, 1055, 1072, 1089, 1106, 1123, 1140, 1157, 1174, 1191, 1208, 1225, 1242, 1259, 1276, 1293, 1310, 1327, 1344, 1361, 1378, 1395, 1412, 1429, 1446, 1463, 1480, 1497, 1514, 1531, 1548, 1565,

10 1582, 1599, 1633, 1650, 1684, 1701, 1718, 1786, 1803, 1820, 1871, 1888, 1922, 1939, 1956, 1973, 1990, 2007, 2024, 2041, 2058, 2092, 2109, 2126, 2160, 2177, 2094, 2211, 2228, 2245, 2262, 2279, 2313, 2330, 2347, 2364, 2381, 2398, 2415, 2432, 2449, 2466, 2483, 2500, 2517, 2534, 2551, 2568, 2585, 2602, 2619, 2636, 2653, 2670, 2687, 2704, 2721, 2738, 2772, 2789, 2823, 2840, 2857, 2925, 2942, 2959, 3010, 3027, 3061, 3078,

15 3095, 3112, 3129, 3146, 3163, 3180, 3197, 3231, 3248, 3265, and 3282. In yet another embodiment, the polypeptide comprises an amino acid sequence selected from the group consisting of a polypeptide having at least 90% sequence identity to at least one of SEQ ID Nos. 1, 18, 35, 52, 69, 86, 103, 120, 137, 154, 171, 188, 205, 222, 239, 256, 273, 290, 307, 324, 358, 375, 409, 426, 443, 511, 528, 545, 596, 613, 647, 664, 681, 698, 715, 732, 749,

20 766, 783, 817, 834, 851, 885, 902, 819, 936, 953, 970, 987, 1004, 1038, 1055, 1072, 1089, 1106, 1123, 1140, 1157, 1174, 1191, 1208, 1225, 1242, 1259, 1276, 1293, 1310, 1327, 1344, 1361, 1378, 1395, 1412, 1429, 1446, 1463, 1480, 1497, 1514, 1531, 1548, 1565, 1582, 1599, 1633, 1650, 1684, 1701, 1718, 1786, 1803, 1820, 1871, 1888, 1922, 1939, 1956, 1973, 1990, 2007, 2024, 2041, 2058, 2092, 2109, 2126, 2160, 2177, 2094, 2211,

25 2228, 2245, 2262, 2279, 2313, 2330, 2347, 2364, 2381, 2398, 2415, 2432, 2449, 2466, 2483, 2500, 2517, 2534, 2551, 2568, 2585, 2602, 2619, 2636, 2653, 2670, 2687, 2704, 2721, 2738, 2772, 2789, 2823, 2840, 2857, 2925, 2942, 2959, 3010, 3027, 3061, 3078, 3095, 3112, 3129, 3146, 3163, 3180, 3197, 3231, 3248, 3265, and 3282.

In embodiments of the invention, the polypeptides may be acidophilic and/or

30 thermophilic. In further embodiments, the polypeptides may be glycosylated, pegylated, and/or otherwise post-translationally modified.

- 5 -

Embodiments of methods include methods of altering genetic recombination inside or outside of a cell, the methods comprising providing a recombinant, purified, and/or isolated nucleotide sequence comprising a nucleotide sequence selected from the group consisting of a nucleotide sequences having at least 90% sequence identity to at least one

5 of the sequences of SEQ ID Nos. 2, 19, 36, 53, 70, 87, 104, 121, 138, 155, 172, 189, 206, 223, 240, 257, 274, 291, 308, 325, 359, 376, 410, 427, 444, 512, 529, 546, 597, 614, 648, 665, 682, 699, 716, 733, 750, 767, 784, 818, 835, 852, 886, 903, 920, 937, 954, 971, 988, 1005, 1039, 1056, 1073, 1090, 1107, 1124, 1141, 1158, 1175, 1192, 1209, 1226, 1243, 1260, 1277, 1294, 1311, 1328, 1345, 1362, 1379, 1396, 1413, 1430, 1447, 1464, 1481,

10 1498, 1515, 1532, 1549, 1566, 1583, 1600, 1634, 1651, 1685, 1702, 1719, 1787, 1804, 1821, 1872, 1889, 1923, 1940, 1957, 1974, 1991, 2008, 2025, 2042, 2059, 2093, 2110, 2127, 2161, 2178, 2195, 2212, 2229, 2246, 2263, 2280, 2314, 2331, 2348, 2365, 2382, 2399, 2416, 2433, 2450, 2467, 2484, 2501, 2518, 2535, 2552, 2569, 2586, 2603, 2620, 2637, 2654, 2671, 2688, 2705, 2722, 2739, 2773, 2790, 2824, 2841, 2858, 2926, 2943,

15 2960, 3011, 3028, 3062, 3079, 3096, 3113, 3130, 3147, 3164, 3181, 3198, 3232, 3249, 3266, and 3283 and/or a recombinant, purified, and/or isolated polypeptide selected from the group consisting of a polypeptide having at least 90% sequence identity to at least one of the sequences of SEQ ID Nos. 1, 18, 35, 52, 69, 86, 103, 120, 137, 154, 171, 188, 205, 222, 239, 256, 273, 290, 307, 324, 358, 375, 409, 426, 443, 511, 528, 545, 596, 613, 647,

20 664, 681, 698, 715, 732, 749, 766, 783, 817, 834, 851, 885, 902, 819, 936, 953, 970, 987, 1004, 1038, 1055, 1072, 1089, 1106, 1123, 1140, 1157, 1174, 1191, 1208, 1225, 1242, 1259, 1276, 1293, 1310, 1327, 1344, 1361, 1378, 1395, 1412, 1429, 1446, 1463, 1480, 1497, 1514, 1531, 1548, 1565, 1582, 1599, 1633, 1650, 1684, 1701, 1718, 1786, 1803, 1820, 1871, 1888, 1922, 1939, 1956, 1973, 1990, 2007, 2024, 2041, 2058, 2092, 2109,

25 2126, 2160, 2177, 2094, 2211, 2228, 2245, 2262, 2279, 2313, 2330, 2347, 2364, 2381, 2398, 2415, 2432, 2449, 2466, 2483, 2500, 2517, 2534, 2551, 2568, 2585, 2602, 2619, 2636, 2653, 2670, 2687, 2704, 2721, 2738, 2772, 2789, 2823, 2840, 2857, 2925, 2942, 2959, 3010, 3027, 3061, 3078, 3095, 3112, 3129, 3146, 3163, 3180, 3197, 3231, 3248, 3265, and 3282 to a nucleotide sequence with which recombination is desired.

30 Further embodiments of methods include placing a cell producing or encoding a recombinant, purified, and/or isolated nucleotide sequence comprising a nucleotide sequence selected from the group consisting of a nucleotide sequences having at least 90%

- 6 -

sequence identity to at least one of the sequences of SEQ ID Nos. 2, 19, 36, 53, 70, 87, 104, 121, 138, 155, 172, 189, 206, 223, 240, 257, 274, 291, 308, 325, 359, 376, 410, 427, 444, 512, 529, 546, 597, 614, 648, 665, 682, 699, 716, 733, 750, 767, 784, 818, 835, 852, 886, 903, 920, 937, 954, 971, 988, 1005, 1039, 1056, 1073, 1090, 1107, 1124, 1141, 1158, 5 1175, 1192, 1209, 1226, 1243, 1260, 1277, 1294, 1311, 1328, 1345, 1362, 1379, 1396, 1413, 1430, 1447, 1464, 1481, 1498, 1515, 1532, 1549, 1566, 1583, 1600, 1634, 1651, 1685, 1702, 1719, 1787, 1804, 1821, 1872, 1889, 1923, 1940, 1957, 1974, 1991, 2008, 2025, 2042, 2059, 2093, 2110, 2127, 2161, 2178, 2195, 2212, 2229, 2246, 2263, 2280, 2314, 2331, 2348, 2365, 2382, 2399, 2416, 2433, 2450, 2467, 2484, 2501, 2518, 2535, 10 2552, 2569, 2586, 2603, 2620, 2637, 2654, 2671, 2688, 2705, 2722, 2739, 2773, 2790, 2824, 2841, 2858, 2926, 2943, 2960, 3011, 3028, 3062, 3079, 3096, 3113, 3130, 3147, 3164, 3181, 3198, 3232, 3249, 3266, and 3283 and/or a recombinant, purified, and/or isolated polypeptide selected from the group consisting of a polypeptide having at least 90% sequence identity to at least one of the sequences of SEQ ID Nos. 1, 18, 35, 52, 69, 15 86, 103, 120, 137, 154, 171, 188, 205, 222, 239, 256, 273, 290, 307, 324, 358, 375, 409, 426, 443, 511, 528, 545, 596, 613, 647, 664, 681, 698, 715, 732, 749, 766, 783, 817, 834, 851, 885, 902, 819, 936, 953, 970, 987, 1004, 1038, 1055, 1072, 1089, 1106, 1123, 1140, 1157, 1174, 1191, 1208, 1225, 1242, 1259, 1276, 1293, 1310, 1327, 1344, 1361, 1378, 1395, 1412, 1429, 1446, 1463, 1480, 1497, 1514, 1531, 1548, 1565, 1582, 1599, 1633, 20 1650, 1684, 1701, 1718, 1786, 1803, 1820, 1871, 1888, 1922, 1939, 1956, 1973, 1990, 2007, 2024, 2041, 2058, 2092, 2109, 2126, 2160, 2177, 2094, 2211, 2228, 2245, 2262, 2279, 2313, 2330, 2347, 2364, 2381, 2398, 2415, 2432, 2449, 2466, 2483, 2500, 2517, 2534, 2551, 2568, 2585, 2602, 2619, 2636, 2653, 2670, 2687, 2704, 2721, 2738, 2772, 2789, 2823, 2840, 2857, 2925, 2942, 2959, 3010, 3027, 3061, 3078, 3095, 3112, 3129, 25 3146, 3163, 3180, 3197, 3231, 3248, 3265, and 3282 in a environment comprising temperatures at or above about 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, and/or 95 degrees Celsius and/or a pH at, below, and/or above 8, 7, 6, 5, 4, 3, 2, 1, and/or 0.

These and other aspects of the invention will become apparent to the skilled artisan in view of the teachings contained herein.

- 7 -

DESCRIPTION OF THE DRAWINGS

FIG. 1 depicts a sequence alignment between SEQ ID NO: 1 (RAAC03697) and ref|ZP_01916690.1|, ref|YP_308280.1|, ref|XP_001525241.1|, gb|AAY21825.1|, and ref|XP_001743680.1| (SEQ ID Nos: 3-7 respectively). Amino acids conserved among all
5 sequences are indicated by a "*" and generally conserved amino acids are indicated by a ":".

FIG. 2 depicts a sequence alignment between SEQ ID NO: 18 (RAAC02297) and ref|ZP_01171092.1|, ref|YP_429214.1|, gb|EAZ41188.1|, gb|AAR38445.1|, and ref|ZP_01774730.1| (SEQ ID Nos: 20-24 respectively). Amino acids conserved among all
10 sequences are indicated by a "*" and generally conserved amino acids are indicated by a ":".

FIG. 3 depicts a sequence alignment between SEQ ID NO: 35 (RAAC02298) and ref|ZP_01575699.1|, ref|ZP_01171091.1|, ref|ZP_02598168.1|, ref|ZP_01900573.1|, and ref|ZP_01301851.1| (SEQ ID Nos: 37-41 respectively). Amino acids conserved among all
15 sequences are indicated by a "*" and generally conserved amino acids are indicated by a ":".

FIG. 4 depicts a sequence alignment between SEQ ID NO: 52 (RAAC02299) and ref|ZP_01171090.1|, ref|ZP_02598167.1|, ref|ZP_01575700.1|, ref|ZP_02849387.1|, and ref|NP_627754.1| (SEQ ID Nos: 54-58 respectively). Amino acids conserved among all
20 sequences are indicated by a "*" and generally conserved amino acids are indicated by a ":".

FIG. 5 depicts a sequence alignment between SEQ ID NO: 69 (RAAC02300) and ref|ZP_02598166.1|, ref|ZP_01171089.1|, ref|ZP_02849386.1|, ref|YP_159112.1|, and ref|ZP_02007550.1| (SEQ ID Nos: 71-75 respectively). Amino acids conserved among all
25 sequences are indicated by a "*" and generally conserved amino acids are indicated by a ":".

FIG. 6 depicts a sequence alignment between SEQ ID NO: 86 (RAAC02301) and ref|ZP_02756760.1|, ref|ZP_01171088.1|, emb|CAE47790.1|, emb|CAE47778.1|, and ref|NP_967133.1| (SEQ ID Nos: 88-92 respectively). Amino acids conserved among all
30 sequences are indicated by a "*" and generally conserved amino acids are indicated by a ":".

- 8 -

FIG. 7 depicts a sequence alignment between SEQ ID NO: 103 (RAAC02302) and ref[ZP_01171087.1], ref[YP_061819.1], emb[CAJ49597.1], ref[YP_158155.1], and ref[NP_779769.1] (SEQ ID Nos: 105-109 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 8 depicts a sequence alignment between SEQ ID NO: 120 (RAAC02303) and ref[YP_001371728.1], ref[YP_001235767.1], ref[YP_674884.1], ref[NP_046584.1], and ref[ZP_00630666.1] (SEQ ID Nos: 122-126 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 9 depicts a sequence alignment between SEQ ID NO: 137 (RAAC02304) and ref[YP_793245.1], ref[YP_386759.1], ref[YP_868126.1], ref[NP_253469.1], and ref[ZP_01591801.1] (SEQ ID Nos: 139-143 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 10 depicts a sequence alignment between SEQ ID NO: 154 (RAAC02305) and ref[YP_645800.1], ref[ZP_01697403.1], ref[NP_111721.1], dbj[BAB60367.1], and ref[YP_950098.1] (SEQ ID Nos: 156-160 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 11 depicts a sequence alignment between SEQ ID NO: 171 (RAAC02306) and gb[EAU91762.1], ref[ZP_01035289.1], ref[ZP_01076306.1], ref[YP_173223.1], and ref[ZP_01500882.1] (SEQ ID Nos: 173-177 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 12 RAAC02306 RAAC02306 RAAC02306 depicts a sequence alignment between SEQ ID NO: 188 (RAAC02289) and ref[YP_517477.1], ref[YP_001568284.1], ref[YP_503850.1], ref[NP_783815.1], and ref[YP_149134.1] (SEQ ID Nos: 190-194 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 13 depicts a sequence alignment between SEQ ID NO: 205 (RAAC02307) and ref[XP_001317319.1], ref[YP_303751.1], ref[ZP_01222568.1], ref[XP_001191064.1],

- 9 -

and ref|ZP_01697132.1| (SEQ ID Nos: 207-211 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 14 depicts a sequence alignment between SEQ ID NO: 222 (RAAC02309)
5 and ref|NP_623664.1|, ref|ZP_02755290.1|, ref|YP_001662357.1|, ref|YP_001666189.1|, and ref|YP_001181426.1| (SEQ ID Nos: 224-228 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 15 depicts a sequence alignment between SEQ ID NO: 239 (RAAC02310)
10 and ref|NP_735797.1|, ref|XP_001247966.1|, gb|AAG38042.1|AF295925_7, ref|ZP_00874806.1|, and gb|ABV55445.1| (SEQ ID Nos: 241-245 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 16 depicts a sequence alignment between SEQ ID NO: 256 (RAAC02311)
15 and ref|YP_001205829.1|, ref|YP_001240062.1|, ref|YP_359336.1|, ref|YP_001451893.1|, and ref|YP_466026.1| (SEQ ID Nos: 258-262 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 17 depicts a sequence alignment between SEQ ID NO: 273 (RAAC02312)
20 and ref|YP_001506532.1|, ref|NP_279998.1|, ref|ZP_02850831.1|, ref|ZP_02248080.1|, and ref|XP_001637270.1| (SEQ ID Nos: 275-279 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 18 depicts a sequence alignment between SEQ ID NO: 290 (RAAC02313)
25 and ref|YP_245669.1|, ref|ZP_02369868.1|, ref|YP_438666.1|, ref|NP_695275.1|, and emb|CAB06069.2| (SEQ ID Nos: 292-296 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 19 depicts a sequence alignment between SEQ ID NO: 307 (RAAC02314)
30 and ref|XP_503244.1|, ref|NP_218466.1|, ref|ZP_01058859.1|, ref|NP_126488.1|, and ref|YP_754274.1| (SEQ ID Nos: 309-313 respectively). Amino acids conserved among all

- 10 -

sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 20 depicts a sequence alignment between SEQ ID NO: 324 (RAAC02315) and ref|ZP_02854145.1|, ref|YP_145847.1|, ref|YP_536482.1|, ref|YP_799230.1|, and
5 ref|NP_714527.1| (SEQ ID Nos: 326-330 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 21 depicts a sequence alignment between SEQ ID NO: 341 (RAAC02316) and ref|ZP_02180762.1|, ref|ZP_02077766.1|, ref|ZP_01893908.1|, ref|XP_001444409.1|,
10 and ref|XP_624126.2| (SEQ ID Nos: 343-347 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 22 depicts a sequence alignment between SEQ ID NO: 358 (RAAC02290) and emb|CAJ50746.1|, gb|AAM28266.1|, ref|XP_816394.1|, ref|XP_001585185.1|, and
15 gb|EAU92316.1| (SEQ ID Nos: 360-364 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 23 depicts a sequence alignment between SEQ ID NO: 375 (RAAC02317) and ref|YP_001210712.1|, ref|ZP_02727046.1|, ref|ZP_02758954.1|, ref|ZP_01593342.1|,
20 and ref|ZP_00235902.1| (SEQ ID Nos: 377-381 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 24 depicts a sequence alignment between SEQ ID NO: 409 (RAAC02319) and ref|YP_431168.1|, ref|YP_001212944.1|, ref|YP_754944.1|, ref|YP_754864.1|, and
25 ref|ZP_02171383.1| (SEQ ID Nos: 411-415 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 25 depicts a sequence alignment between SEQ ID NO: 426 (RAAC02320) and ref|YP_431169.1|, ref|YP_754945.1|, ref|YP_754863.1|, ref|ZP_01287154.1|, and
30 ref|ZP_01287577.1| (SEQ ID Nos: 428-432 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

- 11 -

FIG. 26 depicts a sequence alignment between SEQ ID NO: 443 (RAAC02321) and ref|NP_982177.1|, ref|ZP_02595431.1|, ref|NP_927486.1|, ref|YP_001108426.1|, and emb|CAN89659.1| (SEQ ID Nos: 445-449 respectively). Amino acids conserved among all sequences are indicated by a "*" and generally conserved amino acids are indicated by a ":".

FIG. 27 depicts a sequence alignment between SEQ ID NO: 460 (RAAC02322) and ref|YP_504284.1|, ref|YP_001046337.1|, ref|ZP_02131576.1|, ref|ZP_01288161.1|, and ref|ZP_02132246.1| (SEQ ID Nos: 462-466 respectively). Amino acids conserved among all sequences are indicated by a "*" and generally conserved amino acids are indicated by a ":".

FIG. 28 depicts a sequence alignment between SEQ ID NO: 477 (RAAC02323) and ref|ZP_01287831.1|, ref|YP_076198.1|, ref|YP_446560.1|, ref|NP_634267.1|, and ref|ZP_01106621.1| (SEQ ID Nos: 479-483 respectively). Amino acids conserved among all sequences are indicated by a "*" and generally conserved amino acids are indicated by a ":".

FIG. 29 depicts a sequence alignment between SEQ ID NO: 494 (RAAC02324) and ref|ZP_02595423.1|, ref|NP_982173.1|, ref|ZP_02367476.1|, ref|YP_001616264.1|, and ref|YP_827637.1| (SEQ ID Nos: 496-500 respectively). Amino acids conserved among all sequences are indicated by a "*" and generally conserved amino acids are indicated by a ":".

FIG. 30 depicts a sequence alignment between SEQ ID NO: 511 (RAAC02326) and ref|NP_982172.1|, ref|ZP_01872101.1|, ref|NP_922949.1|, emb|CAO48005.1|, and sp|P08995|NO26_SOYBN (SEQ ID Nos: 513-517 respectively). Amino acids conserved among all sequences are indicated by a "*" and generally conserved amino acids are indicated by a ":".

FIG. 31 depicts a sequence alignment between SEQ ID NO: 528 (RAAC02327) and ref|NP_982171.1|, ref|ZP_01090358.1|, ref|ZP_01856486.1|, ref|ZP_02736297.1|, and ref|ZP_01311632.1| (SEQ ID Nos: 530-534 respectively). Amino acids conserved among all sequences are indicated by a "*" and generally conserved amino acids are indicated by a ":".

FIG. 32 depicts a sequence alignment between SEQ ID NO: 545 (RAAC02328) and ref|NP_982170.1|, ref|XP_955124.1|, ref|XP_763458.1|, ref|XP_845342.1|, and

- 12 -

ref|XP_666904.1| (SEQ ID Nos: 547-551 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 33 depicts a sequence alignment between SEQ ID NO: 562 (RAAC02332) and ref|YP_079109.1|, ref|ZP_01858609.1|, gb|AAV70501.1|, ref|YP_001319533.1|, and ref|ZP_00539168.1| (SEQ ID Nos: 564-568 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 34 depicts a sequence alignment between SEQ ID NO: 596 (RAAC02334) and ref|YP_001485227.1|, ref|YP_001358015.1|, ref|ZP_02013298.1|, ref|ZP_02178354.1|, and ref|YP_001356736.1| (SEQ ID Nos: 598-602 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 35 depicts a sequence alignment between SEQ ID NO: 613 (RAAC02335) and dbj|BAF91394.1|, gb|AAL17690.1|, dbj|BAF91409.1|, ref|NP_001058416.1|, and dbj|BAD45624.1| (SEQ ID Nos: 615-619 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 36 depicts a sequence alignment between SEQ ID NO: 630 (RAAC02336) and ref|YP_517477.1|, ref|YP_517489.1|, ref|ZP_01370335.1|, gb|ACA46983.1|, and ref|YP_001568284.1| (SEQ ID Nos: 632-636 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 37 depicts a sequence alignment between SEQ ID NO: 647 (RAAC02292) and gb|AAB91591.1|, ref|YP_001422657.1|, ref|NP_391247.1|, ref|YP_093160.1|, and ref|NP_391246.1| (SEQ ID Nos: 649-653 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 38 depicts a sequence alignment between SEQ ID NO: 664 (RAAC02337) and gb|ACA41259.1|, ref|ZP_02626811.1|, ref|NP_664934.1|, ref|YP_195796.1|, and ref|NP_817052.1| (SEQ ID Nos: 666-670 respectively). Amino acids conserved among all

- 13 -

sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 39 depicts a sequence alignment between SEQ ID NO: 681 (RAAC02338) and gb|ACA41261.1|, ref|ZP_02626812.1|, ref|NP_664935.1|, ref|XP_001701427.1|, and
5 sp|Q2M3V2|ANR43_HUMAN (SEQ ID Nos: 683-687 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 40 depicts a sequence alignment between SEQ ID NO: 698 (RAAC02339) and gb|AAU83457.1|, emb|CAJ70907.1|, ref|ZP_00514953.1|, ref|YP_322920.1|, and
10 ref|YP_183482.1| (SEQ ID Nos: 700-704 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 41 depicts a sequence alignment between SEQ ID NO: 715 (RAAC02340) and ref|ZP_01731985.1|, ref|ZP_02429891.1|, ref|ZP_02432977.1|, ref|ZP_02085861.1|, and
15 and ref|ZP_02207628.1| (SEQ ID Nos: 717-721 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 42 depicts a sequence alignment between SEQ ID NO: 732 (RAAC02341) and ref|ZP_02596024.1|, ref|NP_150014.1|, ref|ZP_00231288.1|, ref|YP_001213007.1|, and
20 ref|YP_001113884.1| (SEQ ID Nos: 734-738 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 43 depicts a sequence alignment between SEQ ID NO: 749 (RAAC02342) and ref|NP_783868.1|, ref|YP_805310.1|, ref|ZP_01273840.1|, ref|YP_803555.1|, and
25 ref|YP_001270615.1| (SEQ ID Nos: 751-755 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 44 depicts a sequence alignment between SEQ ID NO: 766 (RAAC02293) and ref|ZP_01171099.1|, ref|ZP_02598174.1|, ref|YP_429218.1|, ref|YP_001112194.1|, and
30 ref|YP_001112320.1| (SEQ ID Nos: 768-772 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

- 14 -

FIG. 45 depicts a sequence alignment between SEQ ID NO: 783 (RAAC02346) and ref|NP_623604.1|, ref|ZP_00235040.1|, ref|ZP_02327783.1|, ref|ZP_02082031.1|, and ref|ZP_02846176.1| (SEQ ID Nos: 785-789 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 46 depicts a sequence alignment between SEQ ID NO: 800 (RAAC02347) and ref|XP_001383704.2|, ref|XP_001664270.1|, ref|YP_113896.1|, ref|XP_761114.1|, and ref|XP_001015776.2| (SEQ ID Nos: 802-806 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 47 depicts a sequence alignment between SEQ ID NO: 817 (RAAC03510) and gb|ACA42330.1|, ref|ZP_02605128.1|, ref|ZP_02586769.1|, ref|YP_001049647.1|, and ref|YP_913053.1| (SEQ ID Nos: 819-823 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 48 depicts a sequence alignment between SEQ ID NO: 834 (RAAC02348) and ref|XP_414088.2|, ref|ZP_01171110.1|, gb|AAF98351.1|, ref|NP_607899.1|, and ref|NP_269831.1| (SEQ ID Nos: 836-840 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 49 depicts a sequence alignment between SEQ ID NO: 851 (RAAC02349) and ref|YP_001038860.1|, ref|YP_754926.1|, ref|YP_001396671.1|, ref|YP_001254879.1|, and ref|YP_001396310.1| (SEQ ID Nos: 853-857 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 50 depicts a sequence alignment between SEQ ID NO: 868 (RAAC03270) and ref|YP_001655174.1|, emb|CAJ73677.1|, emb|CAJ73386.1|, emb|CAJ74660.1|, and ref|ZP_00516046.1| (SEQ ID Nos: 870-874 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 51 depicts a sequence alignment between SEQ ID NO: 885 (RAAC03271) and sp|P26545|VE2_HP5B, ref|XP_001371550.1|, sp|P36786|VE2_HP19,

- 15 -

ref|NP_041368.1|, and ref|XP_001131003.1| (SEQ ID Nos: 887-891 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 52 depicts a sequence alignment between SEQ ID NO: 902 (RAAC02294) and ref|YP_941474.2|, gb|AAZ42391.1|, ref|ZP_01171098.1|, prf|2123261AD, and ref|NP_498368.2| (SEQ ID Nos: 904-908 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 53 depicts a sequence alignment between SEQ ID NO: 919 (RAAC02353) and ref|YP_430185.1|, ref|YP_001233893.1|, emb|CAO81523.1|, ref|YP_001603689.1|, and ref|YP_719187.1| (SEQ ID Nos: 921-925 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 54 depicts a sequence alignment between SEQ ID NO: 936 (RAAC02354) and ref|YP_001438903.1|, ref|YP_001251565.1|, ref|YP_126374.1|, ref|YP_123382.1|, and ref|YP_001251158.1| (SEQ ID Nos: 938-942 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 55 depicts a sequence alignment between SEQ ID NO: 953 (RAAC02355) and ref|ZP_02758276.1|, gb|ACA42232.1|, ref|ZP_02602342.1|, ref|ZP_02597242.1|, and ref|ZP_02524501.1| (SEQ ID Nos: 955-959 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 56 depicts a sequence alignment between SEQ ID NO: 970 (RAAC02356) and ref|NP_347717.1|, ref|YP_423535.1|, ref|ZP_01860459.1|, ref|YP_429187.1|, and ref|YP_645289.1| (SEQ ID Nos: 972-976 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 57 depicts a sequence alignment between SEQ ID NO: 987 (RAAC02357) and gb|EAY58379.1|, ref|ZP_01505670.1|, ref|NP_856790.1|, ref|NP_217638.1|, and ref|ZP_00876805.1| (SEQ ID Nos: 989-993 respectively). Amino acids conserved among

- 16 -

all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 58 depicts a sequence alignment between SEQ ID NO: 1004 (RAAC02358) and ref|YP_024839.1|, ref|YP_001456771.1|, ref|YP_655149.1|, ref|XP_748956.1|, and
5 ref|NP_943831.1| (SEQ ID Nos: 1006-1010 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 59 depicts a sequence alignment between SEQ ID NO: 1038 (RAAC02361) and ref|ZP_01966753.1|, ref|YP_001086797.1|, ref|ZP_01805266.1|, ref|YP_001681547.1|,
10 and ref|ZP_01188985.1| (SEQ ID Nos: 1040-1044 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 60 depicts a sequence alignment between SEQ ID NO: 1055 (RAAC02362) and ref|YP_157691.1|, ref|ZP_02509777.1|, ref|ZP_02485527.1|, ref|ZP_02406784.1|, and
15 ref|YP_335221.1| (SEQ ID Nos: 1057-1061 respectively) 1055 in Table 1. Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 61 depicts a sequence alignment between SEQ ID NO: 1072 (RAAC02363) and ref|YP_001681084.1|, ref|YP_177318.1|, ref|ZP_02329650.1|, ref|NP_243607.1|, and
20 ref|YP_001421775.1| (SEQ ID Nos: 1074-1078 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 62 depicts a sequence alignment between SEQ ID NO: 1089 (RAAC02364) and ref|YP_079689.1|, ref|ZP_02329649.1|, ref|ZP_01860132.1|, ref|ZP_01171904.1|, and
25 ref|YP_001487332.1| (SEQ ID Nos: 1091-1095 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 63 depicts a sequence alignment between SEQ ID NO: 1106 (RAAC02366) and ref|ZP_01614696.1|, gb|ABH06559.1|, gb|AAB95339.1|, ref|YP_161675.1|, and
30 ref|YP_001202661.1| (SEQ ID Nos: 1108-1112 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

- 17 -

FIG. 64 depicts a sequence alignment between SEQ ID NO: 1123 (RAAC02367) and ref[ZP_02093159.1], ref[ZP_02026447.1], ref[YP_001127515.1], ref[ZP_02091713.1], and ref[ZP_02423704.1] (SEQ ID Nos: 1125-1129 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 65 depicts a sequence alignment between SEQ ID NO: 1140 (RAAC02370) and ref[ZP_00960984.1], ref[ZP_01035782.1], ref[ZP_01903846.1], ref[ZP_01880414.1], and ref[ZP_01742943.1] (SEQ ID Nos: 1142-1146 respectively). Amino acids conserved among all sequences are indicated by a “*” and gene conserved amino acids are indicated by a “:”.

FIG. 66 depicts a sequence alignment between SEQ ID NO: 1157 (RAAC02371) and ref[ZP_02756730.1], ref[ZP_02833143.1], ref[YP_521772.1], ref[ZP_01551668.1], and ref[YP_887014.1] (SEQ ID Nos: 1159-1163 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 67 depicts a sequence alignment between SEQ ID NO: 1174 (RAAC02372) and sp|Q7ZXB1|MCM7B_XENLA, ref[YP_324842.1], ref[NP_486002.1], ref[XP_660834.1], and ref[ZP_01733540.1] (SEQ ID Nos: 1176-1180 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 68 depicts a sequence alignment between SEQ ID NO: 1191 (RAAC02296) and ref[XP_001563017.1], ref[XP_001615133.1], ref[YP_001236354.1], ref[ZP_01776409.1], and ref[ZP_01565636.1] (SEQ ID Nos: 1193-1197 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 69 depicts a sequence alignment between SEQ ID NO: 1208 (RAAC02373) and ref[YP_001376930.1], gb|AAW81277.1, ref[ZP_02852259.1], ref[YP_001642790.1], and ref[YP_001312077.1] (SEQ ID Nos: 1210-1214 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 70 depicts a sequence alignment between SEQ ID NO: 1225 (RAAC02374) and ref[XP_001467069.1], ref[YP_001376929.1], ref[YP_001208199.1], gb|EAU86007.1,

- 18 -

and ref|YP_946581.1| (SEQ ID Nos: 1227-1231 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 71 depicts a sequence alignment between SEQ ID NO: 1242 (RAAC02375)
5 and ref|XP_975359.1|, ref|XP_001315633.1|, ref|ZP_02840410.1|, ref|ZP_01467536.1|, and
ref|YP_001016790.1| (SEQ ID Nos: 1244-1248 respectively). Amino acids conserved
among all sequences are indicated by a “*” and generally conserved amino acids are
indicated by a “:”.

FIG. 72 depicts a sequence alignment between SEQ ID NO: 1259 (RAAC03273)
10 and ref|YP_502758.1|, gb|EAU81483.1|, ref|ZP_01091610.1|, ref|YP_944003.1|, and
ref|YP_462360.1| (SEQ ID Nos: 1261-1265 respectively). Amino acids conserved among
all sequences are indicated by a “*” and generally conserved amino acids are indicated by
a “:”.

FIG. 73 depicts a sequence alignment between SEQ ID NO: 1276 (RAAC02967)
15 and ref|YP_074959.1|, ref|YP_594046.1|, ref|ZP_01846154.1|, sp|Q45618|TRA6_BACST,
and ref|YP_828009.1| (SEQ ID Nos: 1278-1282 respectively). Amino acids conserved
among all sequences are indicated by a “*” and generally conserved amino acids are
indicated by a “:”.

FIG. 74 depicts a sequence alignment between SEQ ID NO: 1293 (RAAC03589)
20 and ref|YP_146741.1|, ref|YP_148969.1|, ref|YP_001126171.1|, ref|ZP_02130848.1|, and
ref|YP_146154.1| (SEQ ID Nos: 1295-1299 respectively). Amino acids conserved among
all sequences are indicated by a “*” and generally conserved amino acids are indicated by
a “:”.

FIG. 75 depicts a sequence alignment between SEQ ID NO: 1310 (RAAC03695)
25 and ref|ZP_01665148.1|, ref|YP_001111903.1|, ref|YP_752864.1|, ref|YP_753434.1|, and
ref|YP_753226.1| (SEQ ID Nos: 1312-1316 respectively). Amino acids conserved among
all sequences are indicated by a “*” and generally conserved amino acids are indicated by
a “:”.

FIG. 76 depicts a sequence alignment between SEQ ID NO: 1327 (RAAC02318)
30 and ref|YP_754943.1|, ref|YP_754865.1|, ref|YP_431166.1|, ref|ZP_02171171.1|, and
ref|YP_519650.1| (SEQ ID Nos: 1329-1333 respectively). Amino acids conserved among

- 19 -

all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 77 depicts a sequence alignment between SEQ ID NO: 1344 (RAAC02319) and ref|YP_431168.1|, ref|YP_001212944.1|, ref|YP_754944.1|, ref|YP_754864.1|, and
5 ref|ZP_02171383.1| (SEQ ID Nos: 1346-1350 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 78 depicts a sequence alignment between SEQ ID NO: 1361 (RAAC02333) and ref|ZP_02734990.1|, ref|ZP_01265219.1|, ref|YP_266430.1|, ref|YP_008142.1|, and
10 ref|YP_713924.1| (SEQ ID Nos: 1363-1367 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 79 depicts a sequence alignment between SEQ ID NO: 1378 (RAAC03703) and ref|YP_001036724.1|, ref|YP_001039064.1|, ref|YP_001039349.1|, ref|YP_076118.1|, and
15 ref|YP_074958.1|X (SEQ ID Nos: 1380-1384 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 80 depicts a sequence alignment between SEQ ID NO: 1395 (RAAC03568) and ref|YP_001039349.1|, ref|YP_001036724.1|, ref|YP_001039064.1|, ref|YP_076118.1|, and
20 ref|YP_074958.1|X (SEQ ID Nos: 1397-1401 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 81 depicts a sequence alignment between SEQ ID NO: 1412 (RAAC03707) and ref|YP_519534.1|, ref|ZP_01370818.1|, ref|ZP_01372264.1|, ref|YP_516922.1|, and
25 ref|YP_519084.1|X (SEQ ID Nos: 1414-1418 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 82 depicts a sequence alignment between SEQ ID NO: 1429 (RAAC03173) and ref|YP_431168.1|, ref|YP_001212944.1|, ref|YP_754944.1|, ref|YP_754864.1|, and
30 ref|ZP_02171383.1|X (SEQ ID Nos: 1431-1435 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

- 20 -

FIG. 83 depicts a sequence alignment between SEQ ID NO: 1446 (RAAC02966) and ref|YP_074959.1|, ref|NP_634718.1|, ref|NP_616807.1|, gb|AAR99616.1|, and sp|Q45618|TRA6_BACSTX (SEQ ID Nos: 1448-1452 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 84 depicts a sequence alignment between SEQ ID NO: 1463 (RAAC00757) and ref|YP_430545.1|, ref|YP_148444.1|, ref|NP_980798.1|, ref|YP_001211577.1|, and sp|Q9KDI8|RUVB_BACHDX (SEQ ID Nos: 1465-1469 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 85 depicts a sequence alignment between SEQ ID NO: 1480 (RAAC00756) and ref|YP_234498.1|, ref|NP_793742.1|, ref|YP_275913.1|, ref|ZP_00991066.1|, and ref|YP_001186768.1|X (SEQ ID Nos: 1482-1486 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 86 depicts a sequence alignment between SEQ ID NO: 1497 (RAAC00755) and ref|YP_644098.1|, ref|ZP_02848139.1|, ref|YP_518701.1|, ref|YP_074988.1|, and ref|YP_001180347.1|X (SEQ ID Nos: 1499-1503 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 87 depicts a sequence alignment between SEQ ID NO: 1514 (RAAC01468) and ref|YP_146341.1|, ref|YP_001513188.1|, ref|ZP_02261478.1|, ref|YP_001373830.1|, and ref|ZP_02257063.1|X (SEQ ID Nos: 1516-1520 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 88 depicts a sequence alignment between SEQ ID NO: 1531 (RAAC03178) and ref|ZP_01665148.1|, ref|YP_752864.1|, ref|YP_001111903.1|, ref|YP_753434.1|, and ref|YP_753226.1|X (SEQ ID Nos: 1533-1537 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 89 depicts a sequence alignment between SEQ ID NO: 1548 (RAAC01937) and dbj|BAF33373.1|, ref|YP_001126744.1|, gb|AAB52611.1|, gb|ABM97416.1|, and

sp|P52026|DPO1_BACSTX (SEQ ID Nos: 1550-1554 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 90 depicts a sequence alignment between SEQ ID NO: 1565 (RAAC01372) and ref|YP_001681573.1|, ref|ZP_00651175.1|, ref|YP_473713.1|, gb|ACA11657.1|, and ref|YP_475384.1|X (SEQ ID Nos: 1567-1571 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 91 depicts a sequence alignment between SEQ ID NO: 1582 (RAAC00062) and ref|ZP_01695687.1|, ref|ZP_01695971.1|, ref|ZP_01695982.1|, ref|ZP_01695655.1|, and ref|YP_430569.1|X (SEQ ID Nos: 1584-1588 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 92 depicts a sequence alignment between SEQ ID NO: 1599 (RAAC02377) and ref|YP_752864.1|, ref|ZP_01665148.1|, ref|YP_001111903.1|, ref|YP_754667.1|, and ref|YP_753226.1|X (SEQ ID Nos: 1601-1605 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 93 depicts a sequence alignment between SEQ ID NO: 1633 (RAAC03117) and ref|YP_754865.1|, ref|YP_754943.1|, ref|YP_431166.1|, ref|ZP_02171171.1|, and ref|YP_001318017.1|X (SEQ ID Nos: 1635-1639 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 94 depicts a sequence alignment between SEQ ID NO: 1650 (RAAC00037) and ref|YP_001514390.1|, ref|NP_780819.1|, ref|YP_001663996.1|, ref|YP_699899.1|, and ref|YP_077153.1|X (SEQ ID Nos: 1652-1656 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 95 depicts a sequence alignment between SEQ ID NO: 1667 (RAAC00054) and ref|NP_387885.1|, ref|YP_803557.1|, ref|YP_077286.1|, ref|ZP_02326643.1|, and ref|YP_001419683.1|X (SEQ ID Nos: 1669-1673 respectively). Amino acids conserved

- 22 -

among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 96 depicts a sequence alignment between SEQ ID NO: 1684 (RAAC03102) and ref|YP_074959.1|, gb|AAR99616.1|, sp|Q45618|TRA6_BACST, ref|YP_075129.1|, and ref|YP_359963.1|X (SEQ ID Nos: 1686-1690 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 97 depicts a sequence alignment between SEQ ID NO: 1701 (RAAC03103) and ref|YP_074959.1|, ref|YP_594046.1|, ref|ZP_02563000.1|, ref|ZP_02516401.1|, and ref|ZP_02335796.1|X (SEQ ID Nos: 1703-1707 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 98 depicts a sequence alignment between SEQ ID NO: 1718 (RAAC03341) and ref|YP_076075.1|, ref|YP_074958.1|, ref|YP_076118.1|, ref|YP_001039349.1|, and ref|YP_001036724.1|X (SEQ ID Nos: 1720-1724 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 99 depicts a sequence alignment between SEQ ID NO: 1786 (RAAC02852) and ref|YP_001125185.1|, ref|YP_147064.1|, ref|YP_001486753.1|, ref|NP_389494.1|, and ref|YP_091420.1|X (SEQ ID Nos: 1788-1792 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 100 depicts a sequence alignment between SEQ ID NO: 1803 (RAAC02854) and ref|YP_001125186.1|, ref|ZP_01775043.1|, ref|YP_175772.1|, ref|NP_243331.1|, and ref|YP_740910.1|X (SEQ ID Nos: 1805-1809 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 101 depicts a sequence alignment between SEQ ID NO: 1820 (RAAC03166) and ref|YP_001211938.1|, dbj|BAD22831.1|, ref|ZP_01666445.1|, ref|ZP_01665334.1|, and gb|EAY56013.1|X (SEQ ID Nos: 1822-1826 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

- 23 -

FIG. 102 depicts a sequence alignment between SEQ ID NO: 1854 (RAAC02961) and ref|ZP_02330756.1|, ref|ZP_02327778.1|, ref|ZP_02330395.1|, ref|ZP_02327484.1|, and ref|ZP_02326400.1|X (SEQ ID Nos: 1856-1860 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 103 depicts a sequence alignment between SEQ ID NO: 1871 (RAAC02202) and ref|YP_146129.1|, sp|O87703|DNLJ_BACST, gb|ABN05294.1|, ref|YP_001124385.1|, and ref|ZP_01169975.1|X (SEQ ID Nos: 1873-1877 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 104 depicts a sequence alignment between SEQ ID NO: 1888 (RAAC03682) and ref|ZP_01695687.1|, ref|ZP_01695971.1|, ref|ZP_01695982.1|, ref|ZP_01695655.1|, and ref|YP_430569.1|X (SEQ ID Nos: 1890-1894 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 105 depicts a sequence alignment between SEQ ID NO: 1922 (RAAC03770) and ref|ZP_01695687.1|, ref|ZP_01695971.1|, ref|ZP_01695982.1|, ref|ZP_01695655.1|, and ref|YP_430569.1|X (SEQ ID Nos: 1924-1928 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 106 depicts a sequence alignment between SEQ ID NO: 1939 (RAAC02738) and ref|NP_842969.1|, ref|YP_430640.1|, ref|ZP_02596019.1|, ref|YP_001666203.1|, and ref|YP_001512589.1|X (SEQ ID Nos: 1941-1945 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 107 depicts a sequence alignment between SEQ ID NO: 1956 (RAAC02514) and ref|YP_001114460.1|, ref|YP_001111555.1|, ref|YP_001112147.1|, ref|YP_001111684.1|, and ref|YP_001113963.1|X (SEQ ID Nos: 1958-1962 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 108 depicts a sequence alignment between SEQ ID NO: 1973 (RAAC02515) and ref|YP_001111555.1|, ref|YP_001111684.1|, ref|YP_001114460.1|,

- 24 -

ref|YP_001113963.1|, and dbj|BAD18231.1|X (SEQ ID Nos: 1975-1979 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 109 depicts a sequence alignment between SEQ ID NO: 1990 (RAAC02530)
5 and ref|YP_001212947.1|, ref|YP_431166.1|, ref|YP_754943.1|, ref|YP_754865.1|, and
ref|YP_001318017.1|X (SEQ ID Nos: 1992-1996 respectively). Amino acids conserved
among all sequences are indicated by a “*” and generally conserved amino acids are
indicated by a “:”.

FIG. 110 depicts a sequence alignment between SEQ ID NO: 2007 (RAAC02533)
10 and ref|YP_001039349.1|, ref|YP_001036724.1|, ref|YP_001039064.1|, ref|YP_076118.1|,
and ref|YP_076073.1|X (SEQ ID Nos: 2009-2013 respectively). Amino acids conserved
among all sequences are indicated by a “*” and generally conserved amino acids are
indicated by a “:”.

FIG. 111 depicts a sequence alignment between SEQ ID NO: 2024 (RAAC02534)
15 and ref|YP_076118.1|, ref|YP_074958.1|, ref|YP_001039349.1|, ref|YP_001036724.1|, and
ref|YP_001039064.1|X (SEQ ID Nos: 2026-2030 respectively). Amino acids conserved
among all sequences are indicated by a “*” and generally conserved amino acids are
indicated by a “:”.

FIG. 112 depicts a sequence alignment between SEQ ID NO: 2041 (RAAC02562)
20 and ref|YP_148733.1|, ref|YP_146224.1|, ref|ZP_02619781.1|, ref|ZP_02171259.1|, and
ref|NP_977687.1|X (SEQ ID Nos: 2043-2047 respectively). Amino acids conserved
among all sequences are indicated by a “*” and generally conserved amino acids are
indicated by a “:”.

FIG. 113 depicts a sequence alignment between SEQ ID NO: 2058 (RAAC03229)
25 and ref|YP_519534.1|, ref|ZP_01370818.1|, ref|ZP_01372264.1|, ref|YP_516922.1|, and
ref|YP_520815.1|X (SEQ ID Nos: 2060-2064 respectively). Amino acids conserved
among all sequences are indicated by a “*” and generally conserved amino acids are
indicated by a “:”.

FIG. 114 depicts a sequence alignment between SEQ ID NO: 2075
30 FIG. 124 depicts a sequence alignment between SEQ ID NO: 2092 (RAAC00160)
and ref|YP_001125159.1|, ref|ZP_02849289.1|, ref|ZP_02329219.1|, ref|YP_147039.1|, and
ref|NP_816723.1|X (SEQ ID Nos: 2094-2098 respectively). Amino acids conserved

- 25 -

among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 115 depicts a sequence alignment between SEQ ID NO: 2109 (RAAC03182) and ref|YP_146741.1|, ref|YP_148969.1|, ref|YP_001126171.1|, ref|ZP_02130848.1|, and
5 ref|YP_146154.1|X (SEQ ID Nos: 2111-2115 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 116 depicts a sequence alignment between SEQ ID NO: 2126 (RAAC03163) and ref|YP_148733.1|, ref|YP_146224.1|, ref|ZP_02171259.1|, ref|ZP_02619781.1|, and
10 ref|NP_977687.1|X (SEQ ID Nos: 2128-2132 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 117 depicts a sequence alignment between SEQ ID NO: 2143 (RAAC01387) and ref|YP_148733.1|, ref|YP_146224.1|, ref|ZP_02171259.1|, ref|ZP_02619781.1|, and
15 ref|YP_001396471.1|X (SEQ ID Nos: 2145-2149 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 118 depicts a sequence alignment between SEQ ID NO: 2160 (RAAC03385) and gb|AAL87775.1|AF403183_1286, ref|YP_517288.1|, ref|YP_517659.1|,
20 ref|ZP_01368657.1|, and ref|YP_001212990.1|X (SEQ ID Nos: 2162-2166 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 119 depicts a sequence alignment between SEQ ID NO: 2177 (RAAC03398) and ref|YP_001126171.1|, ref|YP_148969.1|, ref|YP_146154.1|, ref|YP_146741.1|, and
25 ref|ZP_02172080.1|X (SEQ ID Nos: 2179-2183 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 120 depicts a sequence alignment between SEQ ID NO: 2194 (RAAC03177) and ref|YP_148969.1|, ref|YP_146154.1|, ref|YP_001126171.1|, ref|YP_146741.1|, and
30 ref|ZP_02172080.1|X (SEQ ID Nos: 2196-2200 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

- 26 -

FIG. 121 depicts a sequence alignment between SEQ ID NO: 2211 (RAAC03588) and ref|YP_001036724.1|, ref|YP_001039064.1|, ref|YP_001039349.1|, ref|YP_074105.1|, and ref|YP_076073.1|X (SEQ ID Nos: 2213-2217 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 122 depicts a sequence alignment between SEQ ID NO: 2228 (RAAC03818) and ref|YP_074959.1|, ref|NP_634718.1|, gb|AAR99616.1|, sp|Q45618|TRA6_BACST, and ref|NP_616807.1|X (SEQ ID Nos: 2230-2234 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 123 depicts a sequence alignment between SEQ ID NO: 2245 (RAAC03819) and ref|YP_594046.1|, ref|YP_074959.1|, ref|YP_359963.1|, ref|YP_361300.1|, and ref|NP_622784.1|X (SEQ ID Nos: 2247-2251 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 124 depicts a sequence alignment between SEQ ID NO: 2262 (RAAC03823) and ref|ZP_01695687.1|, ref|ZP_01695971.1|, ref|ZP_01695982.1|, ref|ZP_01695655.1|, and ref|YP_430569.1|X (SEQ ID Nos: 2264-2268 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 125 depicts a sequence alignment between SEQ ID NO: 2279 (RAAC01171) and ref|ZP_02330348.1|, ref|ZP_02328298.1|, ref|ZP_02326599.1|, ref|YP_001664428.1|, and ref|YP_001664274.1| (SEQ ID Nos: 2281-2285 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 126 depicts a sequence alignment between SEQ ID NO: 2296 (RAAC03825) and ref|YP_519534.1|, ref|ZP_01370818.1|, ref|ZP_01372264.1|, ref|YP_516922.1|, and ref|YP_520815.1| (SEQ ID Nos: 2298-2302 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 127 depicts a sequence alignment between SEQ ID NO: 2313 (RAAC03826) and ref|ZP_02851608.1|, ref|ZP_02326599.1|, ref|YP_001307815.1|, ref|YP_001212943.1|,

- 27 -

and ref|YP_001113174.1| (SEQ ID Nos: 2315-2319 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 128 depicts a sequence alignment between SEQ ID NO: 2330 (RAAC02717) and ref|ZP_01002160.1|, ref|ZP_01592598.1|, ref|YP_001229345.1|, ref|YP_771767.1|, and ref|YP_245447.1| (SEQ ID Nos: 2332-2336 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 129 depicts a sequence alignment between SEQ ID NO: 2347 (RAAC01155) and ref|YP_145872.1|, ref|YP_001124150.1|, ref|ZP_02850412.1|, ref|NP_240902.1|, and ref|NP_466224.1| (SEQ ID Nos: 2349-2353 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 130 depicts a sequence alignment between SEQ ID NO: 2364 (RAAC03145) and dbj|BAA94830.1|, ref|ZP_01666433.1|, ref|YP_001213263.1|, ref|YP_387030.1|, and ref|YP_387237.1| (SEQ ID Nos: 2366-2370 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 131 depicts a sequence alignment between SEQ ID NO: 2381 (RAAC03325) and ref|YP_146741.1|, ref|YP_148969.1|, ref|YP_001126171.1|, ref|ZP_02130848.1|, and ref|YP_146154.1| (SEQ ID Nos: 2383-2387 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 132 depicts a sequence alignment between SEQ ID NO: 2398 (RAAC03376) and ref|ZP_02851608.1|, ref|ZP_02326599.1|, ref|YP_001664274.1|, ref|YP_001307815.1|, and ref|YP_001319448.1| (SEQ ID Nos: 2400-2404 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 133 depicts a sequence alignment between SEQ ID NO: 2415 (RAAC02657) and ref|YP_076090.1|, ref|YP_001516732.1|, ref|YP_001185431.1|, ref|YP_318565.1|, and ref|ZP_01061333.1| (SEQ ID Nos: 2417-2421 respectively). Amino acids conserved

- 28 -

among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 134 depicts a sequence alignment between SEQ ID NO: 2432 (RAAC01373) and ref|YP_001680037.1|, ref|YP_001505049.1|, ref|YP_482514.1|, ref|YP_481774.1|, and
5 ref|NP_215436.1| (SEQ ID Nos: 2434-2438 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 135 depicts a sequence alignment between SEQ ID NO: 2449 (RAAC00337) and ref|YP_001680296.1|, ref|YP_342400.1|, ref|NP_756021.1|, ref|YP_405147.1|, and
10 ref|NP_709160.1| (SEQ ID Nos: 2451-2455 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 136 depicts a sequence alignment between SEQ ID NO: 2466 (RAAC00506) and emb|CAD18993.1|, ref|ZP_00995572.1|, ref|YP_805441.1|, ref|YP_001115956.1|, and
15 ref|ZP_02134324.1| (SEQ ID Nos: 2468-2472 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 137 depicts a sequence alignment between SEQ ID NO: 2483 (RAAC00022) and ref|ZP_01695451.1|, ref|NP_391924.1|, ref|YP_001488901.1|, ref|NP_244897.1|, and
20 ref|YP_001423303.1| (SEQ ID Nos: 2485-2489 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 138 depicts a sequence alignment between SEQ ID NO: 2500 (RAAC00027) and ref|YP_001213441.1|, ref|NP_244917.1|, ref|YP_001377189.1|, ref|YP_149334.1|, and
25 ref|YP_077145.1| (SEQ ID Nos: 2502-2506 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 139 depicts a sequence alignment between SEQ ID NO: 2517 (RAAC01051) and ref|YP_074151.1|, ref|YP_001255315.1|, ref|YP_001392092.1|, gb|ACA43749.1|, and
30 ref|ZP_02619122.1|X (SEQ ID Nos: 2519-2523 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

- 29 -

FIG. 140 depicts a sequence alignment between SEQ ID NO: 2534 (RAAC01050) and ref|YP_878438.1|, ref|ZP_02621211.1|, ref|YP_001560315.1|, ref|YP_001392091.1|, and ref|YP_001255314.1|X (SEQ ID Nos: 2536-2540 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 141 depicts a sequence alignment between SEQ ID NO: 2551 (RAAC01009) and ref|YP_148023.1|, ref|YP_001126202.1|, ref|YP_001421643.1|, ref|ZP_01697284.1|, and gb|ABN10253.1|X (SEQ ID Nos: 2553-2557 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 142 depicts a sequence alignment between SEQ ID NO: 2568 (RAAC00998) and ref|ZP_02175216.1|, ref|YP_464174.1|, ref|YP_753805.1|, ref|ZP_02321813.1|, and ref|ZP_01575281.1|X (SEQ ID Nos: 2570-2574 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 143 depicts a sequence alignment between SEQ ID NO: 2585 (RAAC02359) and ref|NP_832076.1|, ref|YP_001645033.1|, ref|NP_844759.1|, ref|YP_001375058.1|, and ref|YP_535778.1|X (SEQ ID Nos: 2587 FIG. 153-2591 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 144 depicts a sequence alignment between SEQ ID NO: 2602 (RAAC00997) and ref|ZP_02854041.1|, ref|YP_075626.1|, ref|YP_753804.1|, ref|ZP_01667008.1|, and ref|YP_946103.1|X (SEQ ID Nos: 2604-2608 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 145 depicts a sequence alignment between SEQ ID NO: 2619 (RAAC02419) and ref|ZP_02330146.1|, ref|YP_001210709.1|, ref|ZP_02184702.1|, gb|AAD26564.1|AF124258_1, and ref|ZP_01828805.1|X (SEQ ID Nos: 2621-2625 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 145 depicts a sequence alignment between SEQ ID NO: 2636 (RAAC02417) and ref|NP_469419.1|, ref|ZP_02309926.1|, ref|ZP_01926077.1|, ref|ZP_01941236.1|, and

- 30 -

ref|YP_001111866.1|X (SEQ ID Nos: 2638-2642 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 147 depicts a sequence alignment between SEQ ID NO: 2653 (RAAC03180)
5 and ref|ZP_02442523.1|, ref|YP_001664041.1|, ref|YP_001210714.1|, ref|ZP_02589119.1|, and ref|NP_242309.1|X (SEQ ID Nos: 2655-2659 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 148 depicts a sequence alignment between SEQ ID NO: 2670 (RAAC03525)
10 and ref|ZP_02442523.1|, ref|YP_001210714.1|, ref|YP_001180871.1|, ref|YP_001662345.1|, and ref|ZP_02543721.1|X (SEQ ID Nos: 2672-2676 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 149 depicts a sequence alignment between SEQ ID NO: 2687 (RAAC03224)
15 and ref|YP_006607.1|, ref|YP_001038857.1|, ref|YP_146372.1|, ref|YP_001662865.1|, and ref|ZP_02781438.1|X (SEQ ID Nos: 2689-2693 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 150 depicts a sequence alignment between SEQ ID NO: 2704 (RAAC02915)
20 and ref|YP_001211829.1|, ref|ZP_02091210.1|, ref|YP_518738.1|, ref|ZP_02378091.1|, and ref|ZP_02454559.1|X (SEQ ID Nos: 2706-2710 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 151 depicts a sequence alignment between SEQ ID NO: 2721 (RAAC02943)
25 and ref|ZP_02850845.1|, ref|NP_243249.1|, emb|CAD56684.1|, ref|YP_079090.1|, and ref|YP_001421272.1|X (SEQ ID Nos: 2723-2727 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 152 depicts a sequence alignment between SEQ ID NO: 2738 (RAAC02234)
30 and ref|NP_242725.1|, ref|YP_175539.1|, ref|ZP_02170975.1|, ref|ZP_01696769.1|, and emb|CAJ73252.1|X (SEQ ID Nos: 2740-2744 respectively). Amino acids conserved

- 31 -

among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 153 depicts a sequence alignment between SEQ ID NO: 2755 (RAAC01662) and emb|CAK51299.1|, ref|YP_832554.1|, emb|CAI78402.1|, ref|YP_001362100.1|, and
5 ref|YP_712120.1|X (SEQ ID Nos: 2757-2761 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 154 depicts a sequence alignment between SEQ ID NO: 2772 (RAAC02171) and ref|ZP_01188667.1|, ref|ZP_02620185.1|, ref|YP_001512273.1|, ref|YP_877450.1|, and
10 ref|YP_001275096.1|X (SEQ ID Nos: 2774-2778 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 155 depicts a sequence alignment between SEQ ID NO: 2789 (RAAC01696) and sp|P80579|THIO_ALIAC, pdb|1NW2|A, pdb|1NSW|A, pdb|1RQM|A, and
15 ref|YP_703612.1|X (SEQ ID Nos: 2791-2795 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 156 depicts a sequence alignment between SEQ ID NO: 2806 (RAAC01724) and ref|NP_244119.1|, ref|ZP_02850022.1|, ref|YP_815236.1|, ref|NP_965472.1|, and
20 ref|YP_001422284.1|X (SEQ ID Nos: 2808-2812 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 157 depicts a sequence alignment between SEQ ID NO: 2823 (RAAC01817) and ref|ZP_02326197.1|, ref|YP_001212789.1|, ref|ZP_01666637.1|, ref|YP_001211675.1|, and
25 ref|ZP_01126596.1|X (SEQ ID Nos: 2825-2829 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 158 depicts a sequence alignment between SEQ ID NO: 2840 (RAAC01840) and ref|YP_001514193.1|, ref|YP_517020.1|, ref|YP_001317996.1|, ref|YP_001090064.1|, and
30 ref|ZP_01995293.1|X (SEQ ID Nos: 2842-2846 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

- 32 -

FIG. 159 depicts a sequence alignment between SEQ ID NO: 2857 (RAAC01875) and ref|YP_644758.1|, ref|YP_604970.1|, ref|YP_076566.1|, ref|NP_295807.1|, and ref|YP_357266.1|X (SEQ ID Nos: 2859-2863 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 160 depicts a sequence alignment between SEQ ID NO: 2874 (RAAC02539) and ref|ZP_01034116.1|, gb|ABY83631.1|, ref|ZP_02297879.1|, ref|YP_208280.1|, and ref|YP_001516905.1|X (SEQ ID Nos: 2876-2880 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 161 depicts a sequence alignment between SEQ ID NO: 2891 (RAAC02543) and ref|ZP_01042597.1|, gb|ABO14793.1|, ref|YP_001443312.1|, ref|ZP_01815366.1|, and ref|YP_205672.1|X (SEQ ID Nos: 2893-2897 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 162 depicts a sequence alignment between SEQ ID NO: 2908 (RAAC02564) and ref|ZP_01631840.1|, gb|AAZ73681.1|, ref|YP_001633718.1|, ref|YP_001275109.1|, and ref|ZP_01514022.1|X (SEQ ID Nos: 2910-2914 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 163 depicts a sequence alignment between SEQ ID NO: 2925 (RAAC02566) and ref|YP_853610.1|, ref|YP_512277.1|, ref|YP_001038857.1|, ref|ZP_01959153.1|, and ref|YP_001662865.1|X (SEQ ID Nos: 2927-2931 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 164 depicts a sequence alignment between SEQ ID NO: 2942 (RAAC02589) and ref|ZP_00591928.1|, ref|YP_001003150.1|, ref|NP_046614.1|, ref|YP_375842.1|, and ref|YP_001131112.1|X (SEQ ID Nos: 2944-2948 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 165 depicts a sequence alignment between SEQ ID NO: 2959 (RAAC02045) and ref|ZP_02848186.1|, ref|YP_080797.1|, ref|YP_001127101.1|, ref|YP_148915.1|, and

- 33 -

ref|YP_001488316.1|X (SEQ ID Nos: 2961-2965 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 166 depicts a sequence alignment between SEQ ID NO: 2976 (RAAC02635) and ref|YP_866618.1|, ref|YP_342704.1|, ref|NP_384606.1|, ref|YP_115396.1|, and ref|YP_911114.1|X (SEQ ID Nos: 2978-2982 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 167 depicts a sequence alignment between SEQ ID NO: 2993 (RAAC00088) and ref|NP_244431.1|, ref|ZP_02848186.1|, ref|ZP_02171648.1|, ref|ZP_02210735.1|, and ref|YP_176521.1|X (SEQ ID Nos: 2995-2999 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 168 depicts a sequence alignment between SEQ ID NO: 3010 (RAAC00111) and ref|YP_146448.1|, ref|NP_977168.1|, ref|ZP_02261191.1|, ref|ZP_02848045.1|, and ref|NP_981890.1|X (SEQ ID Nos: 3012-3016 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 169 depicts a sequence alignment between SEQ ID NO: 3027 (RAAC03161) and ref|ZP_01898092.1|, ref|ZP_01221581.1|, ref|YP_128524.1|, ref|ZP_01236658.1|, and ref|ZP_01161642.1|X (SEQ ID Nos: 3029-3033 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 170 depicts a sequence alignment between SEQ ID NO: 3044 (RAAC03110) and ref|YP_001567845.1|, gb|ACA00199.1|, ref|ZP_01731959.1|, ref|NP_490383.1|, and ref|YP_319960.1|X (SEQ ID Nos: 3046-3050 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 171 depicts a sequence alignment between SEQ ID NO: 3061 (RAAC03810) and emb|CAP00374.1|, ref|YP_001038857.1|, ref|NP_852746.1|, ref|ZP_01959153.1|, and ref|YP_853610.1|X (SEQ ID Nos: 3063-3067 respectively). Amino acids conserved

- 34 -

among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 172 depicts a sequence alignment between SEQ ID NO: 3078 (RAAC03316) and emb|CAP00374.1|, ref|YP_001038857.1|, ref|ZP_02327844.1|, ref|NP_852746.1|, and
5 ref|ZP_01959153.1|X (SEQ ID Nos: 3080-3084 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 173 depicts a sequence alignment between SEQ ID NO: 3095 (RAAC03018) and ref|YP_148854.1|, ref|YP_001127043.1|, ref|YP_080555.1|, ref|NP_693306.1|, and
10 ref|ZP_02170973.1|X (SEQ ID Nos: 3097-3101 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 174 depicts a sequence alignment between SEQ ID NO: 3112 (RAAC02718) and ref|YP_001178926.1|, ref|YP_001180877.1|, ref|YP_001219837.1|, and
15 ref|YP_001211485.1|, and ref|YP_001211518.1|X (SEQ ID Nos: 3114-3118 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 175 depicts a sequence alignment between SEQ ID NO: 3129 (RAAC01115) and ref|YP_001035109.1|, ref|ZP_01819917.1|, ref|YP_001450442.1|, ref|ZP_01818216.1|, and
20 ref|ZP_01830009.1|X (SEQ ID Nos: 3131-3135 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 176 depicts a sequence alignment between SEQ ID NO: 3146 (RAAC01119) and ref|NP_622644.1|, ref|YP_001664955.1|, ref|YP_076287.1|, ref|ZP_02616274.1|, and
25 ref|YP_001255422.1|X (SEQ ID Nos: 3148-3152 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

FIG. 177 depicts a sequence alignment between SEQ ID NO: 3163 (RAAC00203) and ref|YP_001433837.1|, ref|YP_001276310.1|, ref|ZP_01514627.1|, and
30 ref|YP_001636906.1|, and ref|YP_001102862.1|X (SEQ ID Nos: 3165-3169 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “:”.

- 35 -

FIG. 178 depicts a sequence alignment between SEQ ID NO: 3180 (RAAC01413) and ref[ZP_02849297.1], ref[NP_815033.1], ref[YP_796463.1], ref[NP_862606.1], and ref[ZP_00604769.1]X (SEQ ID Nos: 3182-3186 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 179 depicts a sequence alignment between SEQ ID NO: 3197 (RAAC01435) and ref[YP_359065.1], ref[YP_145897.2], ref[NP_387932.1], ref[YP_001124176.1], and sp[P42816]KPRS_BACCLX (SEQ ID Nos: 3199-3203 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 180 depicts a sequence alignment between SEQ ID NO: 3214 (RAAC01442) and ref[ZP_02170919.1], ref[YP_535778.1], ref[ZP_01862118.1], ref[NP_692713.1], and ref[YP_359077.1]X (SEQ ID Nos: 3216-3220 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 181 depicts a sequence alignment between SEQ ID NO: 3231 (RAAC00895) and ref[NP_846740.1], ref[ZP_00238564.1], ref[ZP_02583512.1], ref[NP_834002.1], and ref[ZP_02604064.1]X (SEQ ID Nos: 3233-3237 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 182 depicts a sequence alignment between SEQ ID NO: 3248 (RAAC03475) and ref[ZP_02442523.1], ref[YP_001210714.1], ref[YP_001180871.1], ref[YP_001662345.1], and ref[YP_001114120.1]X (SEQ ID Nos: 3250-3254 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

FIG. 183 depicts a sequence alignment between SEQ ID NO: 3265 (RAAC03560) and ref[YP_001114120.1], ref[NP_242309.1], ref[YP_146346.1], ref[YP_146347.1], and ref[YP_001662345.1]X (SEQ ID Nos: 3267-3271 respectively). Amino acids conserved among all sequences are indicated by a “*” and generally conserved amino acids are indicated by a “.”.

- 36 -

MODE(S) FOR CARRYING OUT THE INVENTION

Embodiments of the invention include genes and associated proteins related to genetic recombination from the thermoacidophile *Alicyclobacillus acidocaldarius*. Coding sequences for genes related to recombination were determined from sequence information generated from sequencing the genome of *Alicyclobacillus acidocaldarius*. These genes and proteins may represent targets and/or elements of transformation systems or vectors for genetic engineering for introducing nucleotide sequences of interest into *Alicyclobacillus acidocaldarius*, Gram positive thermophiles, or other organisms. Non-limiting examples of nucleotide sequences found within the genome of *Alicyclobacillus acidocaldarius*, and amino acids coded thereby, associated with recombination are listed in the sequence listing. Examples of these nucleotide sequences and the proteins they encode can be found in Table 1. Proteins related to recombination may be, without limitation, one of any of the following: ATP-dependent DNA helicase recG, ATP-dependent DNA ligase, ATP-dependent endopeptidase clp proteolytic subunit, Chromosome partitioning protein, Crossover junction endodeoxyribonuclease ruvC, Deoxyuridine 5'-triphosphate nucleotidohydrolase, DNA adenine methylase, DNA helicase, DNA integration/recombination/inversion, DNA polymerase I, DNA polymerase III beta chain, DNA polymerase IV, DNA primase, DNA repair protein radC, DNA replication and repair protein recF, DNA replication protein dnaD, DNA topoisomerase I, DNA/RNA helicase (DEAD/DEAH box family), DNA-binding protein HU, Fe-S oxidoreductase, Glycerophosphoryl diester phosphodiesterase, HNH endonuclease family protein, Holliday junction DNA helicase ruvB, Integrase/recombinase (XerC/CodV family), Ligase/carboxylase family protein, LtrC-like protein, Macrolide-efflux protein, NAD-dependent DNA ligase, Nicotinate phosphoribosyltransferase, nodulin-26, Phage antirepressor protein, Phage protein, Phosphinothricin N-acetyltransferase, Phosphohydrolase, RecA protein, Recombination protein recR, Replicative DNA helicase, Ribonucleoside-diphosphate reductase alpha chain, Ribonucleoside-diphosphate reductase beta chain, Ribose-phosphate pyrophosphokinase, Serine/threonine protein phosphatase, Single-strand DNA binding protein, Single-stranded DNA-binding protein, Site-specific recombinase, Site-specific resolvase / integrase, Thioredoxin, Thymidine kinase, Transcriptional regulator / Lex A repressor, Transcriptional regulator, Cro/CI family,

- 37 -

Transposase, TRSE protein, Two-component response regulator, Type II restriction-modification system methylation subunit, and others.

Embodiments of the invention relate in part to the gene sequences and/or protein sequences comprising genes and/or proteins of *Alicyclobacillus acidocaldarius*. Genes and proteins included are those which play a role in genetic manipulation. Intracellular enzyme activities may be thermophilic and/or acidophilic in nature and general examples of similar genes are described in the literature. Classes of genes, sequences, enzymes and factors include, but are not limited to, those listed in Table 1. Figures 1-183 provide sequence alignments between polypeptide sequences of the present invention and closely related proteins. Areas of high homology between the polypeptide sequences of the present invention and closely related proteins are indicative of functionality for polypeptide sequences of the present invention for the same purposes as the sequences to which they are aligned.

TABLE 1
***Alicyclobacillus acidocaldarius* coding sequences related to recombination**

Reference	Protein Sequence	Coding Sequence	Function
RAAC02304	SEQ ID NO: 137	SEQ ID NO: 138	Two-component response regulator
RAAC02309	SEQ ID NO: 222	SEQ ID NO: 223	TRSE PROTEIN
RAAC02315	SEQ ID NO: 324	SEQ ID NO: 325	Chromosome partitioning protein parA
RAAC02319	SEQ ID NO: 409	SEQ ID NO: 410	Transposase
RAAC02321	SEQ ID NO: 443	SEQ ID NO: 444	Ligase/carboxylase family protein
RAAC02324	SEQ ID NO: 494	SEQ ID NO: 495	Fe-S OXIDOREDUCTASE
RAAC02326	SEQ ID NO: 511	SEQ ID NO: 512	nodulin-26
RAAC02332	SEQ ID NO: 562	SEQ ID NO: 563	Phosphinothricin N-acetyltransferase
RAAC02292	SEQ ID NO: 647	SEQ ID NO: 648	Transcriptional regulator / Lex A repressor
RAAC02340	SEQ ID NO: 715	SEQ ID NO: 716	LtrC-like protein
RAAC02341	SEQ ID NO: 732	SEQ ID NO: 733	DNA repair protein radC
RAAC02342	SEQ ID NO: 749	SEQ ID NO: 750	DNA polymerase III, beta chain
RAAC02348	SEQ ID NO: 834	SEQ ID NO: 835	Single-stranded DNA-binding protein
RAAC02358	SEQ ID NO: 1004	SEQ ID NO: 1005	DNA/RNA helicase (DEAD/DEAH box family)
RAAC02361	SEQ ID NO: 1038	SEQ ID NO: 1039	DNA helicase II
RAAC02363	SEQ ID NO: 1072	SEQ ID NO: 1073	DNA polymerase IV
RAAC02318	SEQ ID NO: 1327	SEQ ID NO: 1328	Transposase
RAAC02319	SEQ ID NO: 1344	SEQ ID NO: 1345	Transposase

Reference	Protein Sequence	Coding Sequence	Function
RAAC02333	SEQ ID NO: 1361	SEQ ID NO: 1362	DNA integration/recombination/inversion protein
RAAC03173	SEQ ID NO: 1429	SEQ ID NO: 1430	Transposase
RAAC00757	SEQ ID NO: 1463	SEQ ID NO: 1464	Holliday junction DNA helicase ruvB
RAAC00756	SEQ ID NO: 1480	SEQ ID NO: 1481	Holliday junction DNA helicase ruvB
RAAC00755	SEQ ID NO: 1497	SEQ ID NO: 1498	Crossover junction endodeoxyribonuclease ruvC
RAAC01468	SEQ ID NO: 1514	SEQ ID NO: 1515	Site-specific recombinase
RAAC03178	SEQ ID NO: 1531	SEQ ID NO: 1532	Transposase
RAAC01937	SEQ ID NO: 1548	SEQ ID NO: 1549	DNA polymerase I
RAAC01372	SEQ ID NO: 1565	SEQ ID NO: 1566	Transposase
RAAC02377	SEQ ID NO: 1599	SEQ ID NO: 1600	Transposase
RAAC00037	SEQ ID NO: 1650	SEQ ID NO: 1651	ATP-dependent DNA helicase recG
RAAC00054	SEQ ID NO: 1667	SEQ ID NO: 1668	DNA replication and repair protein recF
RAAC03102	SEQ ID NO: 1684	SEQ ID NO: 1685	Transposase
RAAC03103	SEQ ID NO: 1701	SEQ ID NO: 1702	Transposase
RAAC03341	SEQ ID NO: 1718	SEQ ID NO: 1719	Transposase
RAAC02852	SEQ ID NO: 1786	SEQ ID NO: 1787	DNA topoisomerase I (EC 1280.1374.1276.1277)
RAAC02854	SEQ ID NO: 1803	SEQ ID NO: 1804	Integrase/recombinase (XerC/CodV family)
RAAC03166	SEQ ID NO: 1820	SEQ ID NO: 1821	Transposase
RAAC02961	SEQ ID NO: 1854	SEQ ID NO: 1855	DNA integration/recombination/inversion protein
RAAC02202	SEQ ID NO: 1871	SEQ ID NO: 1872	NAD-dependent DNA ligase
RAAC03682	SEQ ID NO: 1888	SEQ ID NO: 1889	Transposase
RAAC02738	SEQ ID NO: 1939	SEQ ID NO: 1940	DNA integration/recombination/inversion protein
RAAC02514	SEQ ID NO: 1956	SEQ ID NO: 1957	Transposase
RAAC02515	SEQ ID NO: 1973	SEQ ID NO: 1974	Transposase
RAAC02530	SEQ ID NO: 1990	SEQ ID NO: 1991	Transposase
RAAC02533	SEQ ID NO: 2007	SEQ ID NO: 2008	Transposase
RAAC02534	SEQ ID NO: 2024	SEQ ID NO: 2025	Transposase
RAAC02562	SEQ ID NO: 2041	SEQ ID NO: 2042	Transposase
RAAC03229	SEQ ID NO: 2058	SEQ ID NO: 2059	Transposase
RAAC00160	SEQ ID NO: 2092	SEQ ID NO: 2093	ATP-dependent DNA helicase recG
RAAC03182	SEQ ID NO: 2109	SEQ ID NO: 2110	Transposase
RAAC03163	SEQ ID NO: 2126	SEQ ID NO: 2127	Transposase
RAAC01387	SEQ ID NO: 2143	SEQ ID NO: 2144	Transposase
RAAC01171	SEQ ID NO: 2279	SEQ ID NO: 2280	Transposase
RAAC03825	SEQ ID NO: 2296	SEQ ID NO: 2297	Transposase
RAAC03826	SEQ ID NO: 2313	SEQ ID NO: 2314	Transposase

Reference	Protein Sequence	Coding Sequence	Function
RAAC02717	SEQ ID NO: 2330	SEQ ID NO: 2331	Transposase
RAAC01155	SEQ ID NO: 2347	SEQ ID NO: 2348	Recombination protein recR
RAAC03145	SEQ ID NO: 2364	SEQ ID NO: 2365	Transposase
RAAC03325	SEQ ID NO: 2381	SEQ ID NO: 2382	Transposase
RAAC03376	SEQ ID NO: 2398	SEQ ID NO: 2399	Transposase
RAAC02657	SEQ ID NO: 2415	SEQ ID NO: 2416	Type II restriction-modification system methylation subunit
RAAC01373	SEQ ID NO: 2432	SEQ ID NO: 2433	Site-specific resolvase / integrase
RAAC00337	SEQ ID NO: 2449	SEQ ID NO: 2450	DNA adenine methylase
RAAC00506	SEQ ID NO: 2466	SEQ ID NO: 2467	Serine/threonine protein phosphatase
RAAC00022	SEQ ID NO: 2483	SEQ ID NO: 2484	Replicative DNA helicase
RAAC00027	SEQ ID NO: 2500	SEQ ID NO: 2501	Single-strand DNA binding protein
RAAC01051	SEQ ID NO: 2517	SEQ ID NO: 2518	Ribonucleoside-diphosphate reductase alpha chain
RAAC01050	SEQ ID NO: 2534	SEQ ID NO: 2535	Ribonucleoside-diphosphate reductase beta chain
RAAC01009	SEQ ID NO: 2551	SEQ ID NO: 2552	DNA replication protein dnaD
RAAC00998	SEQ ID NO: 2568	SEQ ID NO: 2569	ATP-dependent DNA ligase
RAAC02359	SEQ ID NO: 2585	SEQ ID NO: 2586	DNA-binding protein HU
RAAC00997	SEQ ID NO: 2602	SEQ ID NO: 2603	ATP-dependent DNA ligase
RAAC02419	SEQ ID NO: 2619	SEQ ID NO: 2620	Site-specific recombinase
RAAC02417	SEQ ID NO: 2636	SEQ ID NO: 2637	Transcriptional regulator, Cro/CI family
RAAC03180	SEQ ID NO: 2653	SEQ ID NO: 2654	Transcriptional regulator, Cro/CI family
RAAC03224	SEQ ID NO: 2687	SEQ ID NO: 2688	Phage antirepressor protein
RAAC02915	SEQ ID NO: 2704	SEQ ID NO: 2705	Deoxyuridine 5'-triphosphate nucleotidohydrolase
RAAC02943	SEQ ID NO: 2721	SEQ ID NO: 2722	RecA protein
RAAC02234	SEQ ID NO: 2738	SEQ ID NO: 2739	Serine/threonine protein phosphatase
RAAC01662	SEQ ID NO: 2755	SEQ ID NO: 2756	Macrolide-efflux protein
RAAC02171	SEQ ID NO: 2772	SEQ ID NO: 2773	Phosphohydrolase
RAAC01696	SEQ ID NO: 2789	SEQ ID NO: 2790	Thioredoxin
RAAC01724	SEQ ID NO: 2806	SEQ ID NO: 2807	Thioredoxin
RAAC01817	SEQ ID NO: 2823	SEQ ID NO: 2824	Site-specific recombinase
RAAC01840	SEQ ID NO: 2840	SEQ ID NO: 2841	Nicotinate phosphoribosyltransferase
RAAC01875	SEQ ID NO: 2857	SEQ ID NO: 2858	Glycerophosphoryl diester phosphodiesterase
RAAC02539	SEQ ID NO: 2874	SEQ ID NO: 2875	Type II restriction-modification system methylation subunit
RAAC02543	SEQ ID NO: 2891	SEQ ID NO: 2892	DNA adenine methylase
RAAC02564	SEQ ID NO: 2908	SEQ ID NO: 2909	DNA adenine methylase
RAAC02566	SEQ ID NO: 2925	SEQ ID NO: 2926	Phage antirepressor protein
RAAC02589	SEQ ID NO: 2942	SEQ ID NO: 2943	DNA-binding protein HU

- 40 -

Reference	Protein Sequence	Coding Sequence	Function
RAAC02045	SEQ ID NO: 2959	SEQ ID NO: 2960	ATP-dependent endopeptidase clp proteolytic subunit clpP
RAAC02635	SEQ ID NO: 2976	SEQ ID NO: 2977	Type II restriction-modification system methylation subunit
RAAC00088	SEQ ID NO: 2993	SEQ ID NO: 2994	ATP-dependent endopeptidase clp proteolytic subunit clpP
RAAC00111	SEQ ID NO: 3010	SEQ ID NO: 3011	Site-specific recombinase
RAAC03161	SEQ ID NO: 3027	SEQ ID NO: 3028	DNA adenine methylase
RAAC03110	SEQ ID NO: 3044	SEQ ID NO: 3045	HNH endonuclease family protein
RAAC03316	SEQ ID NO: 3078	SEQ ID NO: 3079	Phage antirepressor protein
RAAC03018	SEQ ID NO: 3095	SEQ ID NO: 3096	Thioredoxin
RAAC02718	SEQ ID NO: 3112	SEQ ID NO: 3113	Phage protein
RAAC01115	SEQ ID NO: 3129	SEQ ID NO: 3130	Thymidine kinase
RAAC01119	SEQ ID NO: 3146	SEQ ID NO: 3147	Glycerophosphoryl diester phosphodiesterase
RAAC00203	SEQ ID NO: 3163	SEQ ID NO: 3164	Ribonucleoside-diphosphate reductase beta chain
RAAC01413	SEQ ID NO: 3180	SEQ ID NO: 3181	Serine/threonine protein phosphatase
RAAC01435	SEQ ID NO: 3197	SEQ ID NO: 3198	Ribose-phosphate pyrophosphokinase
RAAC01442	SEQ ID NO: 3214	SEQ ID NO: 3215	DNA-binding protein HU
RAAC00895	SEQ ID NO: 3231	SEQ ID NO: 3232	DNA primase
RAAC03475	SEQ ID NO: 3248	SEQ ID NO: 3249	Transcriptional regulator, Cro/CI family

The present invention relates to nucleotides sequences comprising isolated and/or purified nucleotide sequences of the genome of *Alicyclobacillus acidocaldarius* selected from the sequences of SEQ ID Nos. 2, 19, 36, 53, 70, 87, 104, 121, 138, 155, 172, 189, 206, 223, 240, 257, 274, 291, 308, 325, 359, 376, 410, 427, 444, 512, 529, 546, 597, 614, 648, 665, 682, 699, 716, 733, 750, 767, 784, 818, 835, 852, 886, 903, 920, 937, 954, 971, 988, 1005, 1039, 1056, 1073, 1090, 1107, 1124, 1141, 1158, 1175, 1192, 1209, 1226, 1243, 1260, 1277, 1294, 1311, 1328, 1345, 1362, 1379, 1396, 1413, 1430, 1447, 1464, 1481, 1498, 1515, 1532, 1549, 1566, 1583, 1600, 1634, 1651, 1685, 1702, 1719, 1787, 1804, 1821, 1872, 1889, 1923, 1940, 1957, 1974, 1991, 2008, 2025, 2042, 2059, 2093, 2110, 2127, 2161, 2178, 2195, 2212, 2229, 2246, 2263, 2280, 2314, 2331, 2348, 2365, 2382, 2399, 2416, 2433, 2450, 2467, 2484, 2501, 2518, 2535, 2552, 2569, 2586, 2603, 2620, 2637, 2654, 2671, 2688, 2705, 2722, 2739, 2773, 2790, 2824, 2841, 2858, 2926, 2943, 2960, 3011, 3028, 3062, 3079, 3096, 3113, 3130, 3147, 3164, 3181, 3198, 3232, 3249, 3266, and 3283 or one of their fragments.

- 41 -

The present invention likewise relates to isolated and/or purified nucleotide sequences, characterized in that they comprise at least one of: a) a nucleotide sequence of at least one of the sequences of SEQ ID Nos. 2, 19, 36, 53, 70, 87, 104, 121, 138, 155, 172, 189, 206, 223, 240, 257, 274, 291, 308, 325, 359, 376, 410, 427, 444, 512, 529, 546, 597, 614, 648, 665, 682, 699, 716, 733, 750, 767, 784, 818, 835, 852, 886, 903, 920, 937, 954, 971, 988, 1005, 1039, 1056, 1073, 1090, 1107, 1124, 1141, 1158, 1175, 1192, 1209, 1226, 1243, 1260, 1277, 1294, 1311, 1328, 1345, 1362, 1379, 1396, 1413, 1430, 1447, 1464, 1481, 1498, 1515, 1532, 1549, 1566, 1583, 1600, 1634, 1651, 1685, 1702, 1719, 1787, 1804, 1821, 1872, 1889, 1923, 1940, 1957, 1974, 1991, 2008, 2025, 2042, 2059, 2093, 2110, 2127, 2161, 2178, 2195, 2212, 2229, 2246, 2263, 2280, 2314, 2331, 2348, 2365, 2382, 2399, 2416, 2433, 2450, 2467, 2484, 2501, 2518, 2535, 2552, 2569, 2586, 2603, 2620, 2637, 2654, 2671, 2688, 2705, 2722, 2739, 2773, 2790, 2824, 2841, 2858, 2926, 2943, 2960, 3011, 3028, 3062, 3079, 3096, 3113, 3130, 3147, 3164, 3181, 3198, 3232, 3249, 3266, and 3283 or one of their fragments; b) a nucleotide sequence homologous to a nucleotide sequence such as defined in a); c) a nucleotide sequence complementary to a nucleotide sequence such as defined in a) or b), and a nucleotide sequence of their corresponding RNA; d) a nucleotide sequence capable of hybridizing under stringent conditions with a sequence such as defined in a), b) or c); e) a nucleotide sequence comprising a sequence such as defined in a), b), c) or d); and f) a nucleotide sequence modified by a nucleotide sequence such as defined in a), b), c), d) or e).

Nucleotide, polynucleotide, or nucleic acid sequence will be understood according to the present invention as meaning both a double-stranded or single-stranded DNA in the monomeric and dimeric (so-called in tandem) forms and the transcription products of said DNAs.

Aspects of the invention relate to nucleotide sequences which it has been possible to isolate, purify or partially purify, starting from separation methods such as, for example, ion-exchange chromatography, by exclusion based on molecular size, or by affinity, or alternatively fractionation techniques based on solubility in different solvents, or starting from methods of genetic engineering such as amplification, cloning, and subcloning, it being possible for the sequences of the invention to be carried by vectors.

Isolated and/or purified nucleotide sequence fragment according to the invention will be understood as designating any nucleotide fragment of the genome of

- 42 -

Alicyclobacillus acidocaldarius, and may include, by way of non-limiting examples, length of at least 8, 12, 20, 25, 50, 75, 100, 200, 300, 400, 500, 1000, or more, consecutive nucleotides of the sequence from which it originates.

Specific fragment of an isolated and/or purified nucleotide sequence according to the invention will be understood as designating any nucleotide fragment of the genome of *Alicyclobacillus acidocaldarius*, having, after alignment and comparison with the corresponding fragments of genomic sequences of *Alicyclobacillus acidocaldarius*, at least one nucleotide or base of different nature.

Homologous isolated and/or purified nucleotide sequence in the sense of the present invention is understood as meaning isolated and/or purified a nucleotide sequence having at least a percentage identity with the bases of a nucleotide sequence according to the invention of at least about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.6%, or 99.7%, this percentage being purely statistical and it being possible to distribute the differences between the two nucleotide sequences at random and over the whole of their length.

Specific homologous nucleotide sequence in the sense of the present invention is understood as meaning a homologous nucleotide sequence having at least one nucleotide sequence of a specific fragment, such as defined above. Said "specific" homologous sequences can comprise, for example, the sequences corresponding to the genomic sequence or to the sequences of its fragments representative of variants of the genome of *Alicyclobacillus acidocaldarius*. These specific homologous sequences can thus correspond to variations linked to mutations within strains of *Alicyclobacillus acidocaldarius*, and especially correspond to truncations, substitutions, deletions and/or additions of at least one nucleotide. Said homologous sequences can likewise correspond to variations linked to the degeneracy of the genetic code.

The term "degree or percentage of sequence homology" refers to "degree or percentage of sequence identity between two sequences after optimal alignment" as defined in the present application.

Two amino-acids or nucleotide sequences are said to be "identical" if the sequence of amino-acids or nucleotide residues, in the two sequences is the same when aligned for maximum correspondence as described below. Sequence comparisons between two (or more) peptides or polynucleotides are typically performed by comparing sequences of two

- 43 -

optimally aligned sequences over a segment or “comparison window” to identify and compare local regions of sequence similarity. Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman, *J Mol Biol*, 147, 195-197, by the homology alignment algorithm of Needleman and Wunsch, *J. Mol. Biol.* 48: 443 (1970), by the search for similarity method of Pearson and Lipman, *Proc. Natl. Acad. Sci. (U.S.A.)* 85: 2444 (1988), by computerized implementation of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, Wis.), or by visual inspection.

10 “Percentage of sequence identity” (or degree of identity) is determined by comparing two optimally aligned sequences over a comparison window, where the portion of the peptide or polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The
15 percentage is calculated by determining the number of positions at which the identical amino-acid residue or nucleic acid base occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity.

20 The definition of sequence identity given above is the definition that would be used by one of skill in the art. The definition by itself does not need the help of any algorithm, said algorithms being helpful only to achieve the optimal alignments of sequences, rather than the calculation of sequence identity.

25 From the definition given above, it follows that there is a well defined and only one value for the sequence identity between two compared sequences which value corresponds to the value obtained for the best or optimal alignment.

30 In the BLAST N or BLAST P “BLAST 2 sequence,” software which is available in the web site worldwideweb.blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&PROG_DEF=blastn&BLAST_PROG_DEF=megaBlast&SHOW_DEFAULTS=on&BLAST_SPEC=blast2seq&LINK_LOC=align2seq, and habitually used by the inventors and in general by the skilled person for comparing and determining the identity between

- 44 -

two sequences, gap cost which depends on the sequence length to be compared is directly selected by the software (i.e. 11.2 for substitution matrix BLOSUM-62 for length>85).

Complementary nucleotide sequence of a sequence of the invention is understood as meaning any DNA whose nucleotides are complementary to those of the sequence of the invention, and whose orientation is reversed (antisense sequence).

Hybridization under conditions of stringency with a nucleotide sequence according to the invention is understood as meaning hybridization under conditions of temperature and ionic strength chosen in such a way that they allow the maintenance of the hybridization between two fragments of complementary DNA.

By way of illustration, conditions of great stringency of the hybridization step with the aim of defining the nucleotide fragments described above are advantageously the following.

The hybridization is carried out at a preferential temperature of 65°C in the presence of SSC buffer, 1 x SSC corresponding to 0.15 M NaCl and 0.05 M Na citrate. The washing steps, for example, can be the following: 2 x SSC, at ambient temperature followed by two washes with 2 x SSC, 0.5% SDS at 65°C.; 2 x 0.5 x SSC, 0.5% SDS; at 65°C for 10 minutes each.

The conditions of intermediate stringency, using, for example, a temperature of 42°C in the presence of a 2 x SSC buffer, or of less stringency, for example a temperature of 37°C in the presence of a 2 x SSC buffer, respectively require a globally less significant complementarity for the hybridization between the two sequences.

The stringent hybridization conditions described above for a polynucleotide with a size of approximately 350 bases will be adapted by the person skilled in the art for oligonucleotides of greater or smaller size, according to the teaching of Sambrook et al., 1989.

Among the isolated and/or purified nucleotide sequences according to the invention, are those which can be used as a primer or probe in methods allowing the homologous sequences according to the invention to be obtained, these methods, such as the polymerase chain reaction (PCR), nucleic acid cloning, and sequencing, being well known to the person skilled in the art.

The terms "recombination," "genetic recombination," and "genetic engineering" are used interchangeable herein and refer to the process by which a first stretch of one or

- 45 -

more nucleic acids are removed from or added to a second stretch of nucleic acids. Such addition or removal may occur *in vivo* or *ex vivo*. The terms also refer to the introduction of coding sequences or genes into cells or organisms for the purposes of expression or regulation of other nucleotide sequences or polypeptides. The terms further refer to the alteration of nucleic acid structure or topology. By way of non-limiting example, altering the “twist,” supercoiling, helicity, separation and/or annealing of a nucleotide sequence.

Among the isolated and/or purified nucleotide sequences according to the invention, those are again preferred which can be used as a primer or probe in methods allowing the presence of SEQ ID Nos. 2, 19, 36, 53, 70, 87, 104, 121, 138, 155, 172, 189, 206, 223, 240, 257, 274, 291, 308, 325, 359, 376, 410, 427, 444, 512, 529, 546, 597, 614, 648, 665, 682, 699, 716, 733, 750, 767, 784, 818, 835, 852, 886, 903, 920, 937, 954, 971, 988, 1005, 1039, 1056, 1073, 1090, 1107, 1124, 1141, 1158, 1175, 1192, 1209, 1226, 1243, 1260, 1277, 1294, 1311, 1328, 1345, 1362, 1379, 1396, 1413, 1430, 1447, 1464, 1481, 1498, 1515, 1532, 1549, 1566, 1583, 1600, 1634, 1651, 1685, 1702, 1719, 1787, 1804, 1821, 1872, 1889, 1923, 1940, 1957, 1974, 1991, 2008, 2025, 2042, 2059, 2093, 2110, 2127, 2161, 2178, 2195, 2212, 2229, 2246, 2263, 2280, 2314, 2331, 2348, 2365, 2382, 2399, 2416, 2433, 2450, 2467, 2484, 2501, 2518, 2535, 2552, 2569, 2586, 2603, 2620, 2637, 2654, 2671, 2688, 2705, 2722, 2739, 2773, 2790, 2824, 2841, 2858, 2926, 2943, 2960, 3011, 3028, 3062, 3079, 3096, 3113, 3130, 3147, 3164, 3181, 3198, 3232, 3249, 3266, and 3283, and one of their fragments, or one of their variants such as defined below to be diagnosed.

The nucleotide sequence fragments according to the invention can be obtained, for example, by specific amplification, such as PCR, or after digestion with appropriate restriction enzymes of nucleotide sequences according to the invention, these methods in particular being described in the work of Sambrook et al., 1989. Such representative fragments can likewise be obtained by chemical synthesis according to methods well known to persons of ordinary skill in the art.

Modified nucleotide sequence will be understood as meaning any nucleotide sequence obtained by mutagenesis according to techniques well known to the person skilled in the art, and containing modifications with respect to the normal sequences according to the invention, for example mutations in the regulatory and/or promoter

- 46 -

sequences of polypeptide expression, especially leading to a modification of the rate of expression of said polypeptide or to a modulation of the replicative cycle.

Modified nucleotide sequence will likewise be understood as meaning any nucleotide sequence coding for a modified polypeptide such as defined below.

5 The present invention relates to nucleotide sequence comprising isolated and/or purified nucleotide sequences of *Alicyclobacillus acidocaldarius*, characterized in that they are selected from the sequences SEQ ID Nos. 2, 19, 36, 53, 70, 87, 104, 121, 138, 155, 172, 189, 206, 223, 240, 257, 274, 291, 308, 325, 359, 376, 410, 427, 444, 512, 529, 546, 597,614, 648, 665, 682, 699, 716, 733, 750, 767, 784, 818, 835, 852, 886, 903, 920, 937,
10 954, 971, 988, 1005, 1039, 1056, 1073, 1090, 1107, 1124, 1141, 1158, 1175, 1192, 1209, 1226, 1243, 1260,1277, 1294, 1311, 1328, 1345, 1362, 1379, 1396, 1413, 1430, 1447, 1464, 1481, 1498, 1515, 1532, 1549, 1566, 1583, 1600, 1634, 1651, 1685, 1702, 1719, 1787, 1804, 1821, 1872,1889, 1923, 1940, 1957, 1974, 1991, 2008, 2025, 2042, 2059, 2093, 2110, 2127, 2161, 2178, 2195, 2212, 2229, 2246, 2263, 2280, 2314, 2331, 2348,
15 2365, 2382, 2399, 2416, 2433, 2450, 2467, 2484, 2501, 2518, 2535, 2552, 2569, 2586, 2603, 2620, 2637, 2654, 2671, 2688, 2705, 2722, 2739, 2773, 2790, 2824, 2841, 2858, 2926, 2943, 2960, 3011,3028, 3062, 3079, 3096, 3113, 3130, 3147, 3164, 3181, 3198, 3232, 3249, 3266, and 3283 or one of their fragments.

Embodiments of the invention likewise relate to isolated and/or purified nucleotide
20 sequences characterized in that they comprise a nucleotide sequence selected from: a) at least one of a nucleotide sequence of SEQ ID Nos. 2, 19, 36, 53, 70, 87, 104, 121, 138, 155, 172, 189, 206, 223, 240, 257, 274, 291, 308, 325, 359, 376, 410, 427, 444, 512, 529, 546, 597,614, 648, 665, 682, 699, 716, 733, 750, 767, 784, 818, 835, 852, 886, 903, 920, 937, 954, 971, 988, 1005, 1039, 1056, 1073, 1090, 1107, 1124, 1141, 1158, 1175, 1192,
25 1209, 1226, 1243, 1260,1277, 1294, 1311, 1328, 1345, 1362, 1379, 1396, 1413, 1430, 1447, 1464, 1481, 1498, 1515, 1532, 1549, 1566, 1583, 1600, 1634, 1651, 1685, 1702, 1719, 1787, 1804, 1821, 1872,1889, 1923, 1940, 1957, 1974, 1991, 2008, 2025, 2042, 2059, 2093, 2110, 2127, 2161, 2178, 2195, 2212, 2229, 2246, 2263, 2280, 2314, 2331, 2348, 2365, 2382, 2399, 2416, 2433, 2450, 2467, 2484, 2501, 2518, 2535, 2552, 2569,
30 2586, 2603, 2620, 2637, 2654, 2671, 2688, 2705, 2722, 2739, 2773, 2790, 2824, 2841, 2858, 2926, 2943, 2960, 3011,3028, 3062, 3079, 3096, 3113, 3130, 3147, 3164, 3181, 3198, 3232, 3249, 3266, and 3283 or one of their fragments or one of their fragments; b) a

- 47 -

nucleotide sequence of a specific fragment of a sequence such as defined in a); c) a homologous nucleotide sequence having at least 80% identity with a sequence such as defined in a) or b); d) a complementary nucleotide sequence or sequence of RNA corresponding to a sequence such as defined in a), b) or c); and e) a nucleotide sequence
5 modified by a sequence such as defined in a), b), c) or d).

Among the isolated and/or purified nucleotide sequences according to the invention are the nucleotide sequences of SEQ ID Nos. 13-17, 30-34, 47-51, 64-68, 81-85, 98-102, 115-119, 132-136, 149-153, 166-170, 183-187, 200-204, 217-221, 234-238, 251-255, 268-272, 285-289, 302-306, 319-323, 336-340, 353-357, 370-374, 387-391, 404-408,
10 421-425, 438-442, 455-459, 472-476, 489-493, 506-510, 523-527, 540-544, 557-561, 574-578, 591-595, 608-612, 625-629, 642-646, 659-663, 676-680, 693-697, 710-714, 727-731, 744-748, 761-765, 778-782, 795-799, 812-816, 829-833, 846-850, 863-867, 880-884, 897-901, 914-918, 931-935, 948-952, 965-969, 982-986, 999-1003, 1016-1020, 1033-1037, 1050-1054, 1067-1071, 1084-1088, 1101-1105, 1118-1122, 1135-1139,
15 1152-1156, 1169-1173, 1186-1190, 1203-1207, 1220-1224, 1237-1241, 1254-1258, 1271-1275, 1288-1292, 1305-1309, 1322-1326, 1339-1343, 1356-1360, 1373-1377, 1390-1394, 1407-1411, 1424-1428, 1441-1445, 1458-1462, 1475-1479, 1492-1496, 1509-1513, 1526-1530, 1543-1547, 1560-1564, 1577-1581, 1594-1598, 1611-1615, 1628-1632, 1645-1649, 1662-1666, 1679-1683, 1696-1700, 1713-1717, 1730-1734,
20 1747-1751, 1764-1768, 1781-1785, 1798-1802, 1815-1819, 1832-1836, 1849-1853, 1866-1870, 1883-1887, 1900-1904, 1917-1921, 1934-1938, 1951-1955, 1968-1972, 1985-1989, 2002-2006, 2019-2023, 2036-2040, 2053-2057, 2070-2074, 2087-2091, 2104-2108, 2121-2125, 2138-2142, 2155-2159, 2172-2176, 2189-2193, 2206-2210, 2223-2227, 2240-2244, 2257-2261, 2274-2278, 2291-2295, 2308-2312, 2325-2329,
25 2342-2346, 2359-2363, 2376-2380, 2393-2397, 2410-2414, 2427-2431, 2444-2448, 2461-2465, 2478-2482, 2495-2499, 2512-2516, 2529-2533, 2546-2550, 2563-2567, 2580-2584, 2597-2601, 2614-2618, 2631-2635, 2648-2652, 2665-2669, 2682-2686, 2699-2703, 2716-2720, 2733-2737, 2750-2754, 2767-2771, 2784-2788, 2801-2805, 2818-2822, 2835-2839, 2852-2856, 2869-2873, 2886-2890, 2903-2907, 2920-2924,
30 2937-2941, 2954-2958, 2971-2975, 2988-2992, 3005-3009, 3022-3026, 3039-3043, 3056-3060, 3073-3077, 3090-3094, 3107-3111, 3124-3128, 3141-3145, 3158-3162, 3175-3179, 3192-3196, 3209-3213, 3226-3230, 3243-3247, 3260-3264, 3277-3281, and

- 48 -

3294-3298; or fragments thereof and any isolated and/or purified nucleotide sequences which have a homology of at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.6%, or 99.7% identity with the at least one of the sequences of SEQ ID Nos. 2, 19, 36, 53, 70, 87, 104,
5 121, 138, 155, 172, 189, 206, 223, 240, 257, 274, 291, 308, 325, 359, 376, 410, 427, 444, 512, 529, 546, 597, 614, 648, 665, 682, 699, 716, 733, 750, 767, 784, 818, 835, 852, 886, 903, 920, 937, 954, 971, 988, 1005, 1039, 1056, 1073, 1090, 1107, 1124, 1141, 1158, 1175, 1192, 1209, 1226, 1243, 1260, 1277, 1294, 1311, 1328, 1345, 1362, 1379, 1396, 1413, 1430, 1447, 1464, 1481, 1498, 1515, 1532, 1549, 1566, 1583, 1600, 1634, 1651,
10 1685, 1702, 1719, 1787, 1804, 1821, 1872, 1889, 1923, 1940, 1957, 1974, 1991, 2008, 2025, 2042, 2059, 2093, 2110, 2127, 2161, 2178, 2195, 2212, 2229, 2246, 2263, 2280, 2314, 2331, 2348, 2365, 2382, 2399, 2416, 2433, 2450, 2467, 2484, 2501, 2518, 2535, 2552, 2569, 2586, 2603, 2620, 2637, 2654, 2671, 2688, 2705, 2722, 2739, 2773, 2790, 2824, 2841, 2858, 2926, 2943, 2960, 3011, 3028, 3062, 3079, 3096, 3113, 3130, 3147,
15 3164, 3181, 3198, 3232, 3249, 3266, and 3283; or fragments thereof. Such homologous sequences can comprise, for example, the sequences corresponding to the genomic sequences *Alicyclobacillus acidocaldarius*. In the same manner, these specific homologous sequences can correspond to variations linked to mutations within strains of *Alicyclobacillus acidocaldarius* and especially correspond to truncations, substitutions,
20 deletions and/or additions of at least one nucleotide. As will be apparent to one of ordinary skill in the art, such homologues are easily created and identified using conventional techniques and publicly available computer programs such as BLAST. Accordingly, each homologue referenced above should be considered as set forth herein and fully described.

Embodiments of the invention comprise the isolated and/or purified polypeptides
25 coded for by a nucleotide sequence according to the invention, or fragments thereof, whose sequence is represented by a fragment. Amino acid sequences corresponding to the isolated and/or purified polypeptides which can be coded for according to one of the three possible reading frames of at least one of the sequences SEQ ID Nos. 2, 19, 36, 53, 70, 87, 104, 121, 138, 155, 172, 189, 206, 223, 240, 257, 274, 291, 308, 325, 359, 376, 410, 427, 444,
30 512, 529, 546, 597, 614, 648, 665, 682, 699, 716, 733, 750, 767, 784, 818, 835, 852, 886, 903, 920, 937, 954, 971, 988, 1005, 1039, 1056, 1073, 1090, 1107, 1124, 1141, 1158, 1175, 1192, 1209, 1226, 1243, 1260, 1277, 1294, 1311, 1328, 1345, 1362, 1379, 1396,

- 49 -

1413, 1430, 1447, 1464, 1481, 1498, 1515, 1532, 1549, 1566, 1583, 1600, 1634, 1651, 1685, 1702, 1719, 1787, 1804, 1821, 1872, 1889, 1923, 1940, 1957, 1974, 1991, 2008, 2025, 2042, 2059, 2093, 2110, 2127, 2161, 2178, 2195, 2212, 2229, 2246, 2263, 2280, 2314, 2331, 2348, 2365, 2382, 2399, 2416, 2433, 2450, 2467, 2484, 2501, 2518, 2535, 5 2552, 2569, 2586, 2603, 2620, 2637, 2654, 2671, 2688, 2705, 2722, 2739, 2773, 2790, 2824, 2841, 2858, 2926, 2943, 2960, 3011, 3028, 3062, 3079, 3096, 3113, 3130, 3147, 3164, 3181, 3198, 3232, 3249, 3266, and 3283.

Embodiments of the invention likewise relate to the isolated and/or purified polypeptides, characterized in that they comprise a polypeptide selected from at least one
10 of the amino acid sequences of SEQ ID Nos. 1, 18, 35, 52, 69, 86, 103, 120, 137, 154, 171, 188, 205, 222, 239, 256, 273, 290, 307, 324, 358, 375, 409, 426, 443, 511, 528, 545, 596, 613, 647, 664, 681, 698, 715, 732, 749, 766, 783, 817, 834, 851, 885, 902, 819, 936, 953, 970, 987, 1004, 1038, 1055, 1072, 1089, 1106, 1123, 1140, 1157, 1174, 1191, 1208, 1225, 1242, 1259, 1276, 1293, 1310, 1327, 1344, 1361, 1378, 1395, 1412, 1429, 1446, 1463,
15 1480, 1497, 1514, 1531, 1548, 1565, 1582, 1599, 1633, 1650, 1684, 1701, 1718, 1786, 1803, 1820, 1871, 1888, 1922, 1939, 1956, 1973, 1990, 2007, 2024, 2041, 2058, 2092, 2109, 2126, 2160, 2177, 2094, 2211, 2228, 2245, 2262, 2279, 2313, 2330, 2347, 2364, 2381, and 2398, 2415, 2432, 2449, 2466, 2483, 2500, 2517, 2534, 2551, 2568, 2585, 2602, 2619, 2636, 2653, 2670, 2687, 2704, 2721, 2738, 2772, 2789, 2823, 2840, 2857, 2925,
20 2942, 2959, 3010, 3027, 3061, 3078, 3095, 3112, 3129, 3146, 3163, 3180, 3197, 3231, 3248, 3265, and 3282; or one of their fragments.

Among the isolated and/or purified polypeptides, according to embodiments of the invention, are the isolated and/or purified polypeptides of amino acid sequence SEQ ID
25 Nos. 8-12, 25-29, 42-46, 59-63, 76-80, 93-97, 110-114, 127-131, 144-148, 161-165, 178-182, 195-199, 212-216, 229-233, 246-250, 263-267, 280-284, 297-301, 314-318, 331-335, 348-352, 365-369, 382-386, 399-403, 416-420, 433-437, 450-454, 467-471, 484-488, 501-505, 518-522, 535-539, 552-556, 569-573, 586-590, 603-607, 620-624, 637-641, 654-658, 671-675, 688-692, 705-709, 722-726, 739-743, 756-760, 773-777, 790-794, 807-811, 824-828, 841-845, 858-862, 875-879, 892-896, 909-913, 926-930,
30 943-947, 960-964, 977-981, 994-998, 1011-1015, 1028-1032, 1045-1049, 1062-1066, 1079-1083, 1096-1100, 1113-1117, 1130-1134, 1147-1151, 1164-1168, 1181-1185, 1198-1202, 1215-1219, 1232-1236, 1249-1253, 1266-1270, 1283-1287, 1300-1304,

- 50 -

1317-1321, 1334-1338, 1351-1355, 1368-1372, 1385-1389, 1402-1406, 1419-1423,
1436-1440, 1453-1457, 1470-1474, 1487-1491, 1504-1508, 1521-1525, 1538-1542,
1555-1559, 1572-1576, 1589-1593, 1606-1610, 1623-1627, 1640-1644, 1657-1661,
1674-1678, 1691-1695, 1708-1712, 1725-1729, 1742-1746, 1759-1763, 1776-1780,
5 1793-1797, 1810-1814, 1827-1831, 1844-1848, 1861-1865, 1878-1882, 1895-1899,
1912-1916, 1929-1933, 1946-1950, 1963-1967, 1980-1984, 1997-2001, 2014-2018,
2031-2035, 2048-2052, 2065-2069, 2082-2086, 2099-2103, 2116-2120, 2133-2137,
2150-2154, 2167-2171, 2184-2188, 2201-2205, 2218-2222, 2235-2239, 2252-2256,
2269-2273, 2286-2290, 2303-2307, 2320-2324, 2337-2341, 2354-2358, 2371-2375,
10 2388-2392, 2405-2409, 2422-2426, 2439-2443, 2456-2460, 2473-2477, 2490-2494,
2507-2511, 2524-2528, 2541-2545, 2558-2562, 2575-2579, 2592-2596, 2609-2613,
2626-2630, 2643-2647, 2660-2664, 2677-2681, 2694-2698, 2711-2715, 2728-2732,
2745-2749, 2762-2766, 2779-2783, 2796-2800, 2813-2817, 2830-2834, 2847-2851,
2864-2868, 2881-2885, 2898-2902, 2915-2919, 2932-2936, 2949-2953, 2966-2970,
15 2983-2987, 3000-3004, 3017-3021, 3034-3038, 3051-3055, 3068-3072, 3085-3089,
3102-3106, 3119-3123, 3136-3140, 3153-3157, 3170-3174, 3187-3191, 3204-3208,
3221-3225, 3238-3242, 3255-3259, 3272-3276, and 3289-3293; or fragments thereof or
any other isolated and/or purified polypeptides which have a homology of at least 80%,
81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%,
20 96%, 97%, 98%, 99%, 99.5%, 99.6%, or 99.7% identity with at least one of the sequences
of SEQ ID Nos. 1, 18, 35, 52, 69, 86, 103, 120, 137, 154, 171, 188, 205, 222, 239, 256,
273, 290, 307, 324, 358, 375, 409, 426, 443, 511, 528, 545, 596, 613, 647, 664, 681, 698,
715, 732, 749, 766, 783, 817, 834, 851, 885, 902, 819, 936, 953, 970, 987, 1004, 1038,
1055, 1072, 1089, 1106, 1123, 1140, 1157, 1174, 1191, 1208, 1225, 1242, 1259, 1276,
25 1293, 1310, 1327, 1344, 1361, 1378, 1395, 1412, 1429, 1446, 1463, 1480, 1497, 1514,
1531, 1548, 1565, 1582, 1599, 1633, 1650, 1684, 1701, 1718, 1786, 1803, 1820, 1871,
1888, 1922, 1939, 1956, 1973, 1990, 2007, 2024, 2041, 2058, 2092, 2109, 2126, 2160,
2177, 2094, 2211, 2228, 2245, 2262, 2279, 2313, 2330, 2347, 2364, 2381, 2398, 2415,
2432, 2449, 2466, 2483, 2500, 2517, 2534, 2551, 2568, 2585, 2602, 2619, 2636, 2653,
30 2670, 2687, 2704, 2721, 2738, 2772, 2789, 2823, 2840, 2857, 2925, 2942, 2959, 3010,
3027, 3061, 3078, 3095, 3112, 3129, 3146, 3163, 3180, 3197, 3231, 3248, 3265, and 3282;
or fragments thereof. As will be apparent to one of ordinary skill in the art, such

- 51 -

homologues are easily created and identified using conventional techniques and publicly available computer programs such as BLAST. Accordingly, each homologue referenced above should be considered as set forth herein and fully described.

Embodiments of the invention also relate to the polypeptides, characterized in that
5 they comprise a polypeptide selected from: a) a specific fragment of at least 5 amino acids of a polypeptide of an amino acid sequence according to the invention; b) a polypeptide homologous to a polypeptide such as defined in a); c) a specific biologically active fragment of a polypeptide such as defined in a) or b); and d) a polypeptide modified by a polypeptide such as defined in a), b) or c).

10 In the present description, the terms polypeptide, peptide and protein are interchangeable.

In embodiments of the invention, the isolated and/or purified polypeptides according to the invention may be glycosylated, pegylated, and/or otherwise post-translationally modified. In further embodiments, glycosylation, pegylation, and/or
15 other post-translational modifications may occur in vivo or in vitro and/or may be performed using chemical techniques. In additional embodiments, any glycosylation, pegylation and/or other post-translational modifications may be N-linked or O-linked.

In embodiments of the invention any one of the isolated and/or purified polypeptides according to the invention may be enzymatically or functionally active at
20 temperatures at or above about 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, and/or 95 degrees Celsius and/or may be enzymatically or functionally active at a pH at, below, and/or above 8, 7, 6, 5, 4, 3, 2, 1, and/or 0. In further embodiments of the invention, glycosylation, pegylation, and/or other posttranslational modification may be required for the isolated and/or purified polypeptides according to the invention to be enzymatically or
25 functionally active at pH at or below 8, 7, 6, 5, 4, 3, 2, 1, and/or 0 or at a temperatures at or above about 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, and/or 95 degrees Celsius.

Aspects of the invention relate to polypeptides that are isolated or obtained by purification from natural sources, or else obtained by genetic recombination, or alternatively by chemical synthesis and that they may thus contain unnatural amino acids,
30 as will be described below.

- 52 -

A “polypeptide fragment” according to the embodiments of the invention is understood as designating a polypeptide containing at least 5 consecutive amino acids, preferably 10 consecutive amino acids or 15 consecutive amino acids.

In the present invention, a specific polypeptide fragment is understood as designating the consecutive polypeptide fragment coded for by a specific fragment nucleotide sequence according to the invention.

“Homologous polypeptide” will be understood as designating the polypeptides having, with respect to the natural polypeptide, certain modifications such as, in particular, a deletion, addition, or substitution of at least one amino acid, a truncation, a prolongation, a chimeric fusion, and/or a mutation. Among the homologous polypeptides, those are preferred whose amino acid sequence has at least 80% or 90%, homology with the sequences of amino acids of polypeptides according to the invention.

“Specific homologous polypeptide” will be understood as designating the homologous polypeptides such as defined above and having a specific fragment of polypeptide according to the invention. In the case of a substitution, one or more consecutive or nonconsecutive amino acids are replaced by “equivalent” amino acids. The expression “equivalent” amino acid is directed here at designating any amino acid capable of being substituted by one of the amino acids of the base structure without, however, essentially modifying the biological activities of the corresponding peptides and such that they will be defined by the following. As will be apparent to one of ordinary skill in the art, such substitutions are easily created and identified using standard molecular biology techniques and publicly available computer programs such as BLAST. Accordingly, each substitution referenced above should be considered as set forth herein and fully described. Examples of such substitutions in the amino acid sequences SEQ ID Nos. 1, 18, 35, 52, 69, 86, 103, 120, 137, 154, 171, 188, 205, 222, 239, 256, 273, 290, 307, 324, 358, 375, 409, 426, 443, 511, 528, 545, 596, 613, 647, 664, 681, 698, 715, 732, 749, 766, 783, 817, 834, 851, 885, 902, 819, 936, 953, 970, 987, 1004, 1038, 1055, 1072, 1089, 1106, 1123, 1140, 1157, 1174, 1191, 1208, 1225, 1242, 1259, 1276, 1293, 1310, 1327, 1344, 1361, 1378, 1395, 1412, 1429, 1446, 1463, 1480, 1497, 1514, 1531, 1548, 1565, 1582, 1599, 1633, 1650, 1684, 1701, 1718, 1786, 1803, 1820, 1871, 1888, 1922, 1939, 1956, 1973, 1990, 2007, 2024, 2041, 2058, 2092, 2109, 2126, 2160, 2177, 2094, 2211, 2228, 2245, 2262, 2279, 2313, 2330, 2347, 2364, 2381, 2398, 2415, 2432, 2449, 2466, 2483, 2500, 2517,

- 53 -

2534, 2551, 2568, 2585, 2602, 2619, 2636, 2653, 2670, 2687, 2704, 2721, 2738, 2772, 2789, 2823, 2840, 2857, 2925, 2942, 2959, 3010, 3027, 3061, 3078, 3095, 3112, 3129, 3146, 3163, 3180, 3197, 3231, 3248, 3265, and 3282 may include those isolated and/or purified polypeptides of amino acid sequence SEQ ID Nos. 8-12, 25-29, 42-46, 59-63, 5 76-80, 93-97, 110-114, 127-131, 144-148, 161-165, 178-182, 195-199, 212-216, 229-233, 246-250, 263-267, 280-284, 297-301, 314-318, 331-335, 348-352, 365-369, 382-386, 399-403, 416-420, 433-437, 450-454, 467-471, 484-488, 501-505, 518-522, 535-539, 552-556, 569-573, 586-590, 603-607, 620-624, 637-641, 654-658, 671-675, 688-692, 705-709, 722-726, 739-743, 756-760, 773-777, 790-794, 807-811, 824-828, 841-845, 10 858-862, 875-879, 892-896, 909-913, 926-930, 943-947, 960-964, 977-981, 994-998, 1011-1015, 1028-1032, 1045-1049, 1062-1066, 1079-1083, 1096-1100, 1113-1117, 1130-1134, 1147-1151, 1164-1168, 1181-1185, 1198-1202, 1215-1219, 1232-1236, 1249-1253, 1266-1270, 1283-1287, 1300-1304, 1317-1321, 1334-1338, 1351-1355, 1368-1372, 1385-1389, 1402-1406, 1419-1423, 1436-1440, 1453-1457, 1470-1474, 15 1487-1491, 1504-1508, 1521-1525, 1538-1542, 1555-1559, 1572-1576, 1589-1593, 1606-1610, 1623-1627, 1640-1644, 1657-1661, 1674-1678, 1691-1695, 1708-1712, 1725-1729, 1742-1746, 1759-1763, 1776-1780, 1793-1797, 1810-1814, 1827-1831, 1844-1848, 1861-1865, 1878-1882, 1895-1899, 1912-1916, 1929-1933, 1946-1950, 1963-1967, 1980-1984, 1997-2001, 2014-2018, 2031-2035, 2048-2052, 2065-2069, 20 2082-2086, 2099-2103, 2116-2120, 2133-2137, 2150-2154, 2167-2171, 2184-2188, 2201-2205, 2218-2222, 2235-2239, 2252-2256, 2269-2273, 2286-2290, 2303-2307, 2320-2324, 2337-2341, 2354-2358, 2371-2375, 2388-2392, 2405-2409, 2422-2426, 2439-2443, 2456-2460, 2473-2477, 2490-2494, 2507-2511, 2524-2528, 2541-2545, 2558-2562, 2575-2579, 2592-2596, 2609-2613, 2626-2630, 2643-2647, 2660-2664, 25 2677-2681, 2694-2698, 2711-2715, 2728-2732, 2745-2749, 2762-2766, 2779-2783, 2796-2800, 2813-2817, 2830-2834, 2847-2851, 2864-2868, 2881-2885, 2898-2902, 2915-2919, 2932-2936, 2949-2953, 2966-2970, 2983-2987, 3000-3004, 3017-3021, 3034-3038, 3051-3055, 3068-3072, 3085-3089, 3102-3106, 3119-3123, 3136-3140, 3153-3157, 3170-3174, 3187-3191, 3204-3208, 3221-3225, 3238-3242, 3255-3259, 30 3272-3276, and 3289-3293. These equivalent amino acids may be determined either by depending on their structural homology with the amino acids which they substitute, or on

- 54 -

results of comparative tests of biological activity between the different polypeptides, which are capable of being carried out.

By way of nonlimiting example, the possibilities of substitutions capable of being carried out without resulting in an extensive modification of the biological activity of the corresponding modified polypeptides will be mentioned, the replacement, for example, of leucine by valine or isoleucine, of aspartic acid by glutamic acid, of glutamine by asparagine, of arginine by lysine etc., the reverse substitutions naturally being envisageable under the same conditions.

In a further embodiment, substitutions are limited to substitutions in amino acids not conserved among other proteins which have similar identified enzymatic activity. For example, one of ordinary skill in the art may align proteins of the same function in similar organisms and determine which amino acids are generally conserved among proteins of that function. One example of a program that may be used to generate such alignments is worldwideweb.charite.de/bioinf/strap/ in conjunction with the databases provided by the NCBI. Examples of such polypeptides may include, but are not limited to, those found in amino acid sequence SEQ ID Nos. 8-12, 25-29, 42-46, 59-63, 76-80, 93-97, 110-114, 127-131, 144-148, 161-165, 178-182, 195-199, 212-216, 229-233, 246-250, 263-267, 280-284, 297-301, 314-318, 331-335, 348-352, 365-369, 382-386, 399-403, 416-420, 433-437, 450-454, 467-471, 484-488, 501-505, 518-522, 535-539, 552-556, 569-573, 586-590, 603-607, 620-624, 637-641, 654-658, 671-675, 688-692, 705-709, 722-726, 739-743, 756-760, 773-777, 790-794, 807-811, 824-828, 841-845, 858-862, 875-879, 892-896, 909-913, 926-930, 943-947, 960-964, 977-981, 994-998, 1011-1015, 1028-1032, 1045-1049, 1062-1066, 1079-1083, 1096-1100, 1113-1117, 1130-1134, 1147-1151, 1164-1168, 1181-1185, 1198-1202, 1215-1219, 1232-1236, 1249-1253, 1266-1270, 1283-1287, 1300-1304, 1317-1321, 1334-1338, 1351-1355, 1368-1372, 1385-1389, 1402-1406, 1419-1423, 1436-1440, 1453-1457, 1470-1474, 1487-1491, 1504-1508, 1521-1525, 1538-1542, 1555-1559, 1572-1576, 1589-1593, 1606-1610, 1623-1627, 1640-1644, 1657-1661, 1674-1678, 1691-1695, 1708-1712, 1725-1729, 1742-1746, 1759-1763, 1776-1780, 1793-1797, 1810-1814, 1827-1831, 1844-1848, 1861-1865, 1878-1882, 1895-1899, 1912-1916, 1929-1933, 1946-1950, 1963-1967, 1980-1984, 1997-2001, 2014-2018, 2031-2035, 2048-2052, 2065-2069, 2082-2086, 2099-2103, 2116-2120, 2133-2137, 2150-2154, 2167-2171, 2184-2188, 2201-2205, 2218-2222,

- 55 -

2235-2239, 2252-2256, 2269-2273, 2286-2290, 2303-2307, 2320-2324, 2337-2341,
2354-2358, 2371-2375, 2388-2392, 2405-2409, 2422-2426, 2439-2443, 2456-2460,
2473-2477, 2490-2494, 2507-2511, 2524-2528, 2541-2545, 2558-2562, 2575-2579,
2592-2596, 2609-2613, 2626-2630, 2643-2647, 2660-2664, 2677-2681, 2694-2698,
5 2711-2715, 2728-2732, 2745-2749, 2762-2766, 2779-2783, 2796-2800, 2813-2817,
2830-2834, 2847-2851, 2864-2868, 2881-2885, 2898-2902, 2915-2919, 2932-2936,
2949-2953, 2966-2970, 2983-2987, 3000-3004, 3017-3021, 3034-3038, 3051-3055,
3068-3072, 3085-3089, 3102-3106, 3119-3123, 3136-3140, 3153-3157, 3170-3174,
3187-3191, 3204-3208, 3221-3225, 3238-3242, 3255-3259, 3272-3276, and 3289-3293.

10 Thus, according to one embodiment of the invention, substitutions or mutation may
be made at positions that are generally conserved among proteins of that function. In a
further embodiment, nucleic acid sequences may be mutated or substituted such that the
amino acid they code for is unchanged (degenerate substitutions and/mutations) and/or
mutated or substituted such that any resulting amino acid substitutions or mutation are
15 made at positions that are generally conserved among proteins of that function. Examples
of such nucleic acid sequences may include, but are not limited to, those found in are the
nucleotide sequences of SEQ ID Nos. 13-17, 30-34, 47-51, 64-68, 81-85, 98-102, 115-119,
132-136, 149-153, 166-170, 183-187, 200-204, 217-221, 234-238, 251-255, 268-272,
285-289, 302-306, 319-323, 336-340, 353-357, 370-374, 387-391, 404-408, 421-425,
20 438-442, 455-459, 472-476, 489-493, 506-510, 523-527, 540-544, 557-561, 574-578,
591-595, 608-612, 625-629, 642-646, 659-663, 676-680, 693-697, 710-714, 727-731,
744-748, 761-765, 778-782, 795-799, 812-816, 829-833, 846-850, 863-867, 880-884,
897-901, 914-918, 931-935, 948-952, 965-969, 982-986, 999-1003, 1016-1020,
1033-1037, 1050-1054, 1067-1071, 1084-1088, 1101-1105, 1118-1122, 1135-1139,
25 1152-1156, 1169-1173, 1186-1190, 1203-1207, 1220-1224, 1237-1241, 1254-1258,
1271-1275, 1288-1292, 1305-1309, 1322-1326, 1339-1343, 1356-1360, 1373-1377,
1390-1394, 1407-1411, 1424-1428, 1441-1445, 1458-1462, 1475-1479, 1492-1496,
1509-1513, 1526-1530, 1543-1547, 1560-1564, 1577-1581, 1594-1598, 1611-1615,
1628-1632, 1645-1649, 1662-1666, 1679-1683, 1696-1700, 1713-1717, 1730-1734,
30 1747-1751, 1764-1768, 1781-1785, 1798-1802, 1815-1819, 1832-1836, 1849-1853,
1866-1870, 1883-1887, 1900-1904, 1917-1921, 1934-1938, 1951-1955, 1968-1972,
1985-1989, 2002-2006, 2019-2023, 2036-2040, 2053-2057, 2070-2074, 2087-2091,

- 56 -

2104-2108, 2121-2125, 2138-2142, 2155-2159, 2172-2176, 2189-2193, 2206-2210,
 2223-2227, 2240-2244, 2257-2261, 2274-2278, 2291-2295, 2308-2312, 2325-2329,
 2342-2346, 2359-2363, 2376-2380, 2393-2397, 2410-2414, 2427-2431, 2444-2448,
 2461-2465, 2478-2482, 2495-2499, 2512-2516, 2529-2533, 2546-2550, 2563-2567,
 5 2580-2584, 2597-2601, 2614-2618, 2631-2635, 2648-2652, 2665-2669, 2682-2686,
 2699-2703, 2716-2720, 2733-2737, 2750-2754, 2767-2771, 2784-2788, 2801-2805,
 2818-2822, 2835-2839, 2852-2856, 2869-2873, 2886-2890, 2903-2907, 2920-2924,
 2937-2941, 2954-2958, 2971-2975, 2988-2992, 3005-3009, 3022-3026, 3039-3043,
 3056-3060, 3073-3077, 3090-3094, 3107-3111, 3124-3128, 3141-3145, 3158-3162,
 10 3175-3179, 3192-3196, 3209-3213, 3226-3230, 3243-3247, 3260-3264, 3277-3281, and
 3294-3298; or fragments thereof.

The specific homologous polypeptides likewise correspond to polypeptides coded
 for by the specific homologous nucleotide sequences such as defined above and thus
 comprise in the present definition the polypeptides which are mutated or correspond to
 15 variants which can exist in *Alicyclobacillus acidocaldarius*, and which especially
 correspond to truncations, substitutions, deletions, and/or additions of at least one amino
 acid residue.

“Specific biologically active fragment of a polypeptide” according to an
 embodiment of the invention will be understood in particular as designating a specific
 20 polypeptide fragment, such as defined above, having at least one of the characteristics of
 polypeptides according to the invention. In certain embodiments the peptide is capable of
 behaving as at least one of the types of proteins outlined in Table 1.

The polypeptide fragments according to embodiments of the invention can
 correspond to isolated or purified fragments naturally present in *Alicyclobacillus*
 25 *acidocaldarius* or correspond to fragments which can be obtained by cleavage of said
 polypeptide by a proteolytic enzyme, such as trypsin or chymotrypsin or collagenase, or by
 a chemical reagent, such as cyanogen bromide (CNBr). Such polypeptide fragments can
 likewise just as easily be prepared by chemical synthesis, from hosts transformed by an
 expression vector according to the invention containing a nucleic acid allowing the
 30 expression of said fragments, placed under the control of appropriate regulation and/or
 expression elements.

- 57 -

“Modified polypeptide” of a polypeptide according to an embodiment of the invention is understood as designating a polypeptide obtained by genetic recombination or by chemical synthesis as will be described below, having at least one modification with respect to the normal sequence. These modifications may or may not be able to bear on amino acids at the origin of specificity, and/or of activity, or at the origin of the structural conformation, localization, and of the capacity of membrane insertion of the polypeptide according to the invention. It will thus be possible to create polypeptides of equivalent, increased, or decreased activity, and of equivalent, narrower, or wider specificity. Among the modified polypeptides, it is necessary to mention the polypeptides in which up to 5 or more amino acids can be modified, truncated at the N- or C-terminal end, or even deleted or added.

The methods allowing said modulations on eukaryotic or prokaryotic cells to be demonstrated are well known to the person of ordinary skill in the art. It is likewise well understood that it will be possible to use the nucleotide sequences coding for said modified polypeptides for said modulations, for example through vectors according to the invention and described below.

The preceding modified polypeptides can be obtained by using combinatorial chemistry, in which it is possible to systematically vary parts of the polypeptide before testing them on models, cell cultures or microorganisms for example, to select the compounds which are most active or have the properties sought.

Chemical synthesis likewise has the advantage of being able to use nonnatural amino acids, or nonpeptide bonds.

Thus, in order to improve the duration of life of the polypeptides according to the invention, it may be of interest to use nonnatural amino acids, for example in D form, or else amino acid analogs, especially sulfur-containing forms, for example.

Finally, it will be possible to integrate the structure of the polypeptides according to the invention, its specific or modified homologous forms, into chemical structures of polypeptide type or others. Thus, it may be of interest to provide at the N- and C-terminal ends molecules not recognized by proteases.

The nucleotide sequences coding for a polypeptide according to the invention are likewise part of the invention.

- 58 -

The invention likewise relates to nucleotide sequences utilizable as a primer or probe, characterized in that said sequences are selected from the nucleotide sequences according to the invention.

It is well understood that the present invention, in various embodiments, likewise relates to specific polypeptides of *Alicyclobacillus acidocaldarius*, coded for by nucleotide sequences, capable of being obtained by purification from natural polypeptides, by genetic recombination or by chemical synthesis by procedures well known to the person skilled in the art and such as described in particular below. In the same manner, the labeled or unlabeled mono- or polyclonal antibodies directed against said specific polypeptides coded for by said nucleotide sequences are also encompassed by the invention.

Embodiments of the invention additionally relate to the use of a nucleotide sequence according to the invention as a primer or probe for the detection and/or the amplification of nucleic acid sequences.

The nucleotide sequences according to embodiments of the invention can thus be used to amplify nucleotide sequences, especially by the PCR technique (polymerase chain reaction) (Erlich, 1989; Innis et al., 1990; Rolfs et al., 1991; and White et al., 1997).

These oligodeoxyribonucleotide or oligoribonucleotide primers advantageously have a length of at least 8 nucleotides, preferably of at least 12 nucleotides, and even more preferentially at least 20 nucleotides.

Other amplification techniques of the target nucleic acid can be advantageously employed as alternatives to PCR.

The nucleotide sequences of the invention, in particular the primers according to the invention, can likewise be employed in other procedures of amplification of a target nucleic acid, such as: the TAS technique (Transcription-based Amplification System), described by Kwoh et al. in 1989; the 3SR technique (Self-Sustained Sequence Replication), described by Guatelli et al. in 1990; the NASBA technique (Nucleic Acid Sequence Based Amplification), described by Kievitis et al. in 1991; the SDA technique (Strand Displacement Amplification) (Walker et al., 1992); the TMA technique (Transcription Mediated Amplification).

The polynucleotides of the invention can also be employed in techniques of amplification or of modification of the nucleic acid serving as a probe, such as: the LCR technique (Ligase Chain Reaction), described by Landegren et al. in 1988 and improved by

- 59 -

Barany et al. in 1991, which employs a thermostable ligase; the RCR technique (Repair Chain Reaction), described by Segev in 1992; the CPR technique (Cycling Probe Reaction), described by Duck et al. in 1990; the amplification technique with Q-beta replicase, described by Miele et al. in 1983 and especially improved by Chu et al. in 1986, 5 Lizardi et al. in 1988, then by Burg et al. as well as by Stone et al. in 1996.

In the case where the target polynucleotide to be detected is possibly an RNA, for example an mRNA, it will be possible to use, prior to the employment of an amplification reaction with the aid of at least one primer according to the invention or to the employment of a detection procedure with the aid of at least one probe of the invention, an enzyme of 10 reverse transcriptase type in order to obtain a cDNA from the RNA contained in the biological sample. The cDNA obtained will thus serve as a target for the primer(s) or the probe(s) employed in the amplification or detection procedure according to the invention.

The detection probe will be chosen in such a manner that it hybridizes with the target sequence or the amplicon generated from the target sequence. By way of sequence, 15 such a probe will advantageously have a sequence of at least 12 nucleotides, in particular of at least 20 nucleotides, and preferably of at least 100 nucleotides.

Embodiments of the invention also comprise the nucleotide sequences utilizable as a probe or primer according to the invention, characterized in that they are labeled with a radioactive compound or with a nonradioactive compound.

20 The unlabeled nucleotide sequences can be used directly as probes or primers, although the sequences are generally labeled with a radioactive isotope (^{32}P , ^{35}S , ^3H , ^{125}I) or with a nonradioactive molecule (biotin, acetylaminofluorene, digoxigenin, 5-bromodeoxyuridine, fluorescein) to obtain probes which are utilizable for numerous applications.

25 Examples of nonradioactive labeling of nucleotide sequences are described, for example, in French Patent No. 78.10975 or by Urdea et al. or by Sanchez-Pescador et al. in 1988.

In the latter case, it will also be possible to use one of the labeling methods described in patents FR-2 422 956 and FR-2 518 755.

30 The hybridization technique can be carried out in various manners (Matthews et al., 1988). The most general method consists in immobilizing the nucleic acid extract of cells on a support (such as nitrocellulose, nylon, polystyrene) and in incubating, under

- 60 -

well-defined conditions, the immobilized target nucleic acid with the probe. After hybridization, the excess of probe is eliminated and the hybrid molecules formed are detected by the appropriate method (measurement of the radioactivity, of the fluorescence or of the enzymatic activity linked to the probe).

5 The invention, in various embodiments, likewise comprises the nucleotide sequences according to the invention, characterized in that they are immobilized on a support, covalently or noncovalently.

 According to another advantageous mode of employing nucleotide sequences according to the invention, the latter can be used immobilized on a support and can thus
10 serve to capture, by specific hybridization, the target nucleic acid obtained from the biological sample to be tested. If necessary, the solid support is separated from the sample and the hybridization complex formed between said capture probe and the target nucleic acid is then detected with the aid of a second probe, a so-called detection probe, labeled with an easily detectable element.

15 Another aspect of the present invention is a vector for the cloning and/or expression of a sequence, characterized in that it contains a nucleotide sequence according to the invention.

 The vectors according to the invention, characterized in that they contain the elements allowing the integration, expression and/or the secretion of said nucleotide
20 sequences in a determined host cell, are likewise part of the invention.

 The vector may then contain a promoter, signals of initiation and termination of translation, as well as appropriate regions of regulation of transcription. It may be able to be maintained stably in the host cell and can optionally have particular signals specifying the secretion of the translated protein. These different elements may be chosen as a
25 function of the host cell used. To this end, the nucleotide sequences according to the invention may be inserted into autonomous replication vectors within the chosen host, or integrated vectors of the chosen host.

 Such vectors will be prepared according to the methods currently used by the person skilled in the art, and it will be possible to introduce the resulting vectors into an
30 appropriate host by standard methods, such as, for example, lipofection, electroporation, conjugation, and thermal shock.

- 61 -

The vectors according to the invention are, for example, vectors of plasmid or viral origin. One example of a vector for the expression of polypeptides of the invention is baculovirus.

5 These vectors are useful for transforming host cells in order to clone or to express the nucleotide sequences of the invention.

The invention likewise comprises the host cells transformed by a vector according to the invention.

10 These cells can be obtained by the introduction into host cells of a nucleotide sequence inserted into a vector such as defined above, then the culturing of said cells under conditions allowing the replication and/or expression of the transfected nucleotide sequence.

15 The host cell can be selected from prokaryotic or eukaryotic systems, such as, for example, bacterial cells (Olins and Lee, 1993), but likewise yeast cells (Buckholz, 1993), as well as plants cells, such as *Arabidopsis sp.*, and animal cells, in particular the cultures of mammalian cells (Edwards and Aruffo, 1993), for example, Chinese hamster ovary (CHO) cells, but likewise the cells of insects in which it is possible to use procedures employing baculoviruses, for example sf9 insect cells (Luckow, 1993).

Embodiments of the invention likewise relate to organisms comprising one of such transformed cells according to the invention.

20 The obtainment of transgenic organisms according to the invention expressing one or more of the coding sequences of *Alicyclobacillus acidocaldarius* or part of the coding sequences may be carried out in, for example, rats, mice, or rabbits according to methods well known to the person skilled in the art, such as by viral or nonviral transfections. It will be possible to obtain the transgenic organisms expressing one or more of such coding sequences by transfection of multiple copies of such coding sequences under the control of a strong promoter of ubiquitous nature, or selective for one type of tissue. It will likewise be possible to obtain the transgenic organisms by homologous recombination in embryonic cell strains, transfer of these cell strains to embryos, selection of the affected chimeras at the level of the reproductive lines, and growth of said chimeras.

30 The transformed cells as well as the transgenic organisms according to the invention are utilizable in procedures for preparation of recombinant polypeptides.

- 62 -

It is today possible to produce recombinant polypeptides in relatively large quantity by genetic engineering using the cells transformed by expression vectors according to the invention or using transgenic organisms according to the invention.

5 The procedures for preparation of a polypeptide of the invention in recombinant form, characterized in that they employ a vector and/or a cell transformed by a vector according to the invention and/or a transgenic organism comprising one of said transformed cells according to the invention are themselves comprised in the present invention.

10 As used herein, "transformation" and "transformed" relate to the introduction of nucleic acids into a cell, whether prokaryotic or eukaryotic. Further, "transformation" and "transformed," as used herein, need not relate to growth control or growth deregulation.

Among the procedures for preparation of a polypeptide of the invention in recombinant form, the preparation procedures employing a vector, and/or a cell transformed by the vector and/or a transgenic organism comprising one of the transformed
15 cells, containing a nucleotide sequence according to the invention coding for a polypeptide of *Alicyclobacillus acidocaldarius*.

A variant according to the invention may consist of producing a recombinant polypeptide fused to a "carrier" protein (chimeric protein). The advantage of this system is that it may allow stabilization of and/or a decrease in the proteolysis of the recombinant
20 product, an increase in the solubility in the course of renaturation in vitro and/or a simplification of the purification when the fusion partner has an affinity for a specific ligand.

More particularly, the invention relates to a procedure for preparation of a polypeptide of the invention comprising the following steps: a) culture of transformed cells
25 under conditions allowing the expression of a recombinant polypeptide of nucleotide sequence according to the invention; b) if need be, recovery of the recombinant polypeptide.

When the procedure for preparation of a polypeptide of the invention employs a transgenic organism according to the invention, the recombinant polypeptide is then
30 extracted from said organism.

The invention also relates to a polypeptide which is capable of being obtained by a procedure of the invention such as described previously.

- 63 -

The invention also comprises a procedure for preparation of a synthetic polypeptide, characterized in that it uses a sequence of amino acids of polypeptides according to the invention.

5 The invention likewise relates to a synthetic polypeptide obtained by a procedure according to the invention.

The polypeptides according to the invention can likewise be prepared by techniques which are conventional in the field of the synthesis of peptides. This synthesis can be carried out in homogeneous solution or in solid phase.

10 For example, recourse can be made to the technique of synthesis in homogeneous solution described by Houben-Weyl in 1974.

This method of synthesis consists in successively condensing, two by two, the successive amino acids in the order required, or in condensing amino acids and fragments formed previously and already containing several amino acids in the appropriate order, or alternatively several fragments previously prepared in this way, it being understood that it
15 will be necessary to protect beforehand all the reactive functions carried by these amino acids or fragments, with the exception of amine functions of one and carboxyls of the other or vice-versa, which must normally be involved in the formation of peptide bonds, especially after activation of the carboxyl function, according to the methods well known in the synthesis of peptides.

20 Recourse may also be made to the technique described by Merrifield.

To make a peptide chain according to the Merrifield procedure, recourse is made to a very porous polymeric resin, on which is immobilized the first C-terminal amino acid of the chain. This amino acid is immobilized on a resin through its carboxyl group and its amine function is protected. The amino acids which are going to form the peptide chain are
25 thus immobilized, one after the other, on the amino group, which is deprotected beforehand each time, of the portion of the peptide chain already formed, and which is attached to the resin. When the whole of the desired peptide chain has been formed, the protective groups of the different amino acids forming the peptide chain are eliminated and the peptide is detached from the resin with the aid of an acid.

30 The invention additionally relates to hybrid polypeptides having at least one polypeptide according to the invention, and a sequence of a polypeptide capable of inducing an immune response in man or animals.

- 64 -

Advantageously, the antigenic determinant is such that it is capable of inducing a humoral and/or cellular response.

It will be possible for such a determinant to comprise a polypeptide according to the invention in glycosylated, pegylated, and/or otherwise post-translationally modified form used with a view to obtaining immunogenic compositions capable of inducing the synthesis of antibodies directed against multiple epitopes.

These hybrid molecules can be formed, in part, of a polypeptide carrier molecule or of fragments thereof according to the invention, associated with a possibly immunogenic part, in particular an epitope of the diphtheria toxin, the tetanus toxin, a surface antigen of the hepatitis B virus (patent FR 79 21811), the VP1 antigen of the poliomyelitis virus or any other viral or bacterial toxin or antigen.

The procedures for synthesis of hybrid molecules encompass the methods used in genetic engineering for constructing hybrid nucleotide sequences coding for the polypeptide sequences sought. It will be possible, for example, to refer advantageously to the technique for obtainment of genes coding for fusion proteins described by Minton in 1984.

The hybrid nucleotide sequences coding for a hybrid polypeptide as well as the hybrid polypeptides according to the invention characterized in that they are recombinant polypeptides obtained by the expression of the hybrid nucleotide sequences are likewise part of the invention.

The invention likewise comprises the vectors characterized in that they contain one of the hybrid nucleotide sequences. The host cells transformed by the vectors, the transgenic organisms comprising one of said transformed cells as well as the procedures for preparation of recombinant polypeptides using the vectors, the transformed cells and/or said transgenic organisms are, of course, likewise part of the invention.

The polypeptides according to the invention, the antibodies according to the invention described below and the nucleotide sequences according to the invention can advantageously be employed in procedures for the detection and/or identification of *Alicyclobacillus acidocaldarius*, in a sample capable of containing them. These procedures, according to the specificity of the polypeptides, the antibodies and the nucleotide sequences according to the invention which will be used, will in particular be able to detect and/or to identify *Alicyclobacillus acidocaldarius*.

- 65 -

The polypeptides according to the invention can advantageously be employed in a procedure for the detection and/or the identification of *Alicyclobacillus acidocaldarius* in a sample capable of containing them, characterized in that it comprises the following steps:
a) contacting of this sample with a polypeptide or one of its fragments according to the
5 invention (under conditions allowing an immunological reaction between said polypeptide and the antibodies possibly present in the biological sample); b) demonstration of the antigen-antibody complexes possibly formed.

Any conventional procedure can be employed for carrying out such a detection of the antigen-antibody complexes possibly formed.

10 By way of nonlimiting example, one method brings into play immunoenzymatic processes according to the ELISA technique, by immunofluorescence, or radioimmunological processes (RIA) or their equivalent.

Thus, the invention likewise relates to the polypeptides according to the invention, labeled with the aid of an adequate label, such as, of the enzymatic, fluorescent or
15 radioactive type.

Such methods comprise, for example, the following acts: deposition of determined quantities of a polypeptide composition according to the invention in the wells of a microtiter plate, introduction into the wells of increasing dilutions of serum, or of a biological sample other than that defined previously, having to be analyzed, incubation of
20 the microtiter plate, introduction into the wells of the microtiter plate of labeled antibodies directed against pig immunoglobulins, the labeling of these antibodies having been carried out with the aid of an enzyme selected from those which are capable of hydrolyzing a substrate by modifying the absorption of the radiation of the latter, at least at a determined wavelength, for example at 550 nm, detection, by comparison with a control test, of the
25 quantity of hydrolyzed substrate.

The polypeptides according to the invention allow monoclonal or polyclonal antibodies to be prepared which are characterized in that they specifically recognize the polypeptides according to the invention. It will advantageously be possible to prepare the monoclonal antibodies from hybridomas according to the technique described by Kohler
30 and Milstein in 1975. It will be possible to prepare the polyclonal antibodies, for example, by immunization of an animal, in particular a mouse, with a polypeptide or a DNA, according to the invention, associated with an adjuvant of the immune response, and then

- 66 -

purification of the specific antibodies contained in the serum of the immunized animals on an affinity column on which the polypeptide which has served as an antigen has previously been immobilized. The polyclonal antibodies according to the invention can also be prepared by purification, on an affinity column on which a polypeptide according to the invention has previously been immobilized, of the antibodies contained in the serum of an animal immunologically challenged by *Alicyclobacillus acidocaldarius*, or a polypeptide or fragment according to the invention.

The invention likewise relates to mono- or polyclonal antibodies or their fragments, or chimeric antibodies, characterized in that they are capable of specifically recognizing a polypeptide according to the invention.

It will likewise be possible for the antibodies of the invention to be labeled in the same manner as described previously for the nucleic probes of the invention, such as a labeling of enzymatic, fluorescent or radioactive type.

The invention is additionally directed at a procedure for the detection and/or identification of *Alicyclobacillus acidocaldarius* in a sample, characterized in that it comprises the following steps: a) contacting of the sample with a mono- or polyclonal antibody according to the invention (under conditions allowing an immunological reaction between said antibodies and the polypeptides of *Alicyclobacillus acidocaldarius* possibly present in the biological sample); b) demonstration of the antigen-antibody complex possibly formed.

The present invention likewise relates to a procedure for the detection and/or the identification of *Alicyclobacillus acidocaldarius* in a sample, characterized in that it employs a nucleotide sequence according to the invention.

More particularly, the invention relates to a procedure for the detection and/or the identification of *Alicyclobacillus acidocaldarius* in a sample, characterized in that it contains the following steps: a) if need be, isolation of the DNA from the sample to be analyzed; b) specific amplification of the DNA of the sample with the aid of at least one primer, or a pair of primers, according to the invention; c) demonstration of the amplification products.

These can be detected, for example, by the technique of molecular hybridization utilizing a nucleic probe according to the invention. This probe will advantageously be labeled with a nonradioactive (cold probe) or radioactive isotope.

- 67 -

For the purposes of the present invention, "DNA of the biological sample" or "DNA contained in the biological sample" will be understood as meaning either the DNA present in the biological sample considered, or possibly the cDNA obtained after the action of an enzyme of reverse transcriptase type on the RNA present in said biological sample.

5 A further embodiment of the invention comprises a method, characterized in that it comprises the following acts: a) contacting of a nucleotide probe according to the invention with a biological sample, the DNA contained in the biological sample having, if need be, previously been made accessible to hybridization under conditions allowing the hybridization of the probe with the DNA of the sample; b) demonstration of the hybrid
10 formed between the nucleotide probe and the DNA of the biological sample.

The present invention also relates to a procedure according to the invention, characterized in that it comprises the following acts: a) contacting of a nucleotide probe immobilized on a support according to the invention with a biological sample, the DNA of the sample having, if need be, previously been made accessible to hybridization, under
15 conditions allowing the hybridization of the probe with the DNA of the sample; b) contacting of the hybrid formed between the nucleotide probe immobilized on a support and the DNA contained in the biological sample, if need be after elimination of the DNA of the biological sample which has not hybridized with the probe, with a nucleotide probe labeled according to the invention; c) demonstration of the novel hybrid formed in act b).

20 According to an advantageous embodiment of the procedure for detection and/or identification defined previously, this is characterized in that, prior to act a), the DNA of the biological sample is first amplified with the aid of at least one primer according to the invention.

Embodiments of methods include methods of altering recombination inside or
25 outside of a cell, the methods comprising: providing a recombinant, purified, and/or isolated nucleotide sequence comprising a nucleotide sequence selected from the group consisting of a nucleotide sequences having at least 90% sequence identity to at least one of the sequences of SEQ ID Nos. 2, 19, 36, 53, 70, 87, 104, 121, 138, 155, 172, 189, 206, 223, 240, 257, 274, 291, 308, 325, 359, 376, 410, 427, 444, 512, 529, 546, 597, 614, 648,
30 665, 682, 699, 716, 733, 750, 767, 784, 818, 835, 852, 886, 903, 920, 937, 954, 971, 988, 1005, 1039, 1056, 1073, 1090, 1107, 1124, 1141, 1158, 1175, 1192, 1209, 1226, 1243, 1260, 1277, 1294, 1311, 1328, 1345, 1362, 1379, 1396, 1413, 1430, 1447, 1464, 1481,

- 68 -

1498, 1515, 1532, 1549, 1566, 1583, 1600, 1634, 1651, 1685, 1702, 1719, 1787, 1804,
1821, 1872, 1889, 1923, 1940, 1957, 1974, 1991, 2008, 2025, 2042, 2059, 2093, 2110,
2127, 2161, 2178, 2195, 2212, 2229, 2246, 2263, 2280, 2314, 2331, 2348, 2365, 2382,
2399, 2416, 2433, 2450, 2467, 2484, 2501, 2518, 2535, 2552, 2569, 2586, 2603, 2620,
5 2637, 2654, 2671, 2688, 2705, 2722, 2739, 2773, 2790, 2824, 2841, 2858, 2926, 2943,
2960, 3011, 3028, 3062, 3079, 3096, 3113, 3130, 3147, 3164, 3181, 3198, 3232, 3249,
3266, and 3283; and/or a recombinant, purified, and/or isolated polypeptide selected from
the group consisting of a polypeptide having at least 90% sequence identity to at least one
of the sequences of SEQ ID Nos. 1, 18, 35, 52, 69, 86, 103, 120, 137, 154, 171, 188, 205,
10 222, 239, 256, 273, 290, 307, 324, 358, 375, 409, 426, 443, 511, 528, 545, 596, 613, 647,
664, 681, 698, 715, 732, 749, 766, 783, 817, 834, 851, 885, 902, 819, 936, 953, 970, 987,
1004, 1038, 1055, 1072, 1089, 1106, 1123, 1140, 1157, 1174, 1191, 1208, 1225, 1242,
1259, 1276, 1293, 1310, 1327, 1344, 1361, 1378, 1395, 1412, 1429, 1446, 1463, 1480,
1497, 1514, 1531, 1548, 1565, 1582, 1599, 1633, 1650, 1684, 1701, 1718, 1786, 1803,
15 1820, 1871, 1888, 1922, 1939, 1956, 1973, 1990, 2007, 2024, 2041, 2058, 2092, 2109,
2126, 2160, 2177, 2094, 2211, 2228, 2245, 2262, 2279, 2313, 2330, 2347, 2364, 2381,
2398, 2415, 2432, 2449, 2466, 2483, 2500, 2517, 2534, 2551, 2568, 2585, 2602, 2619,
2636, 2653, 2670, 2687, 2704, 2721, 2738, 2772, 2789, 2823, 2840, 2857, 2925, 2942,
2959, 3010, 3027, 3061, 3078, 3095, 3112, 3129, 3146, 3163, 3180, 3197, 3231, 3248,
20 3265, and 3282 to a nucleotide sequence with which recombination event is desired.

Further embodiments of methods include placing a cell producing or encoding a
recombinant, purified, and/or isolated nucleotide sequence comprising a nucleotide
sequence selected from the group consisting of a nucleotide sequences having at least 90%
sequence identity to at least one of the sequences of SEQ ID Nos. 2, 19, 36, 53, 70, 87,
25 104, 121, 138, 155, 172, 189, 206, 223, 240, 257, 274, 291, 308, 325, 359, 376, 410, 427,
444, 512, 529, 546, 597, 614, 648, 665, 682, 699, 716, 733, 750, 767, 784, 818, 835, 852,
886, 903, 920, 937, 954, 971, 988, 1005, 1039, 1056, 1073, 1090, 1107, 1124, 1141, 1158,
1175, 1192, 1209, 1226, 1243, 1260, 1277, 1294, 1311, 1328, 1345, 1362, 1379, 1396,
1413, 1430, 1447, 1464, 1481, 1498, 1515, 1532, 1549, 1566, 1583, 1600, 1634, 1651,
30 1685, 1702, 1719, 1787, 1804, 1821, 1872, 1889, 1923, 1940, 1957, 1974, 1991, 2008,
2025, 2042, 2059, 2093, 2110, 2127, 2161, 2178, 2195, 2212, 2229, 2246, 2263, 2280,
2314, 2331, 2348, 2365, 2382, 2399, 2416, 2433, 2450, 2467, 2484, 2501, 2518, 2535,

- 69 -

2552, 2569, 2586, 2603, 2620, 2637, 2654, 2671, 2688, 2705, 2722, 2739, 2773, 2790,
2824, 2841, 2858, 2926, 2943, 2960, 3011, 3028, 3062, 3079, 3096, 3113, 3130, 3147,
3164, 3181, 3198, 3232, 3249, 3266, and 3283; and/or a recombinant, purified, and/or
isolated polypeptide selected from the group consisting of a polypeptide having at least
5 90% sequence identity to at least one of the sequences of SEQ ID Nos. 1, 18, 35, 52, 69,
86, 103, 120, 137, 154, 171, 188, 205, 222, 239, 256, 273, 290, 307, 324, 358, 375, 409,
426, 443, 511, 528, 545, 596, 613, 647, 664, 681, 698, 715, 732, 749, 766, 783, 817, 834,
851, 885, 902, 819, 936, 953, 970, 987, 1004, 1038, 1055, 1072, 1089, 1106, 1123, 1140,
1157, 1174, 1191, 1208, 1225, 1242, 1259, 1276, 1293, 1310, 1327, 1344, 1361, 1378,
10 1395, 1412, 1429, 1446, 1463, 1480, 1497, 1514, 1531, 1548, 1565, 1582, 1599, 1633,
1650, 1684, 1701, 1718, 1786, 1803, 1820, 1871, 1888, 1922, 1939, 1956, 1973, 1990,
2007, 2024, 2041, 2058, 2092, 2109, 2126, 2160, 2177, 2094, 2211, 2228, 2245, 2262,
2279, 2313, 2330, 2347, 2364, 2381, 2398, 2415, 2432, 2449, 2466, 2483, 2500, 2517,
2534, 2551, 2568, 2585, 2602, 2619, 2636, 2653, 2670, 2687, 2704, 2721, 2738, 2772,
15 2789, 2823, 2840, 2857, 2925, 2942, 2959, 3010, 3027, 3061, 3078, 3095, 3112, 3129,
3146, 3163, 3180, 3197, 3231, 3248, 3265, and 3282 in a environment comprising
temperatures at or above about 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, and/or
95 degrees Celsius and/or a pH at, below, and/or above 8, 7, 6, 5, 4, 3, 2, 1, and/or 0

The present invention provides cells that have been genetically manipulated to have
20 an altered capacity to produce expressed proteins. In particular, the present invention
relates to Gram-positive microorganisms, such as *Bacillus* species having enhanced
expression of a protein of interest, wherein one or more chromosomal genes have been
inactivated, and/or wherein one or more chromosomal genes have been deleted from the
Bacillus chromosome. In some further embodiments, one or more indigenous
25 chromosomal regions have been deleted from a corresponding wild-type *Bacillus* host
chromosome. In further embodiments, the *Bacillus* is an *Alicyclobacillus sp.* or
Alicyclobacillus acidocaldarius.

Additional embodiments, include methods of modulating recombination at
temperatures at or above about 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, and/or
30 95 degrees Celsius and/or at a pH at, below, and/or above 8, 7, 6, 5, 4, 3, 2, 1, and/or 0 via
providng a recombinant, purified, and/or isolated nucleotide sequence comprising a
nucleotide sequence selected from the group consisting of a nucleotide sequences having at

- 70 -

least 90% sequence identity to at least one of the sequences of SEQ ID Nos. 2, 19, 36, 53, 70, 87, 104, 121, 138, 155, 172, 189, 206, 223, 240, 257, 274, 291, 308, 325, 359, 376, 410, 427, 444, 512, 529, 546, 597, 614, 648, 665, 682, 699, 716, 733, 750, 767, 784, 818, 835, 852, 886, 903, 920, 937, 954, 971, 988, 1005, 1039, 1056, 1073, 1090, 1107, 1124, 1141, 1158, 1175, 1192, 1209, 1226, 1243, 1260, 1277, 1294, 1311, 1328, 1345, 1362, 1379, 1396, 1413, 1430, 1447, 1464, 1481, 1498, 1515, 1532, 1549, 1566, 1583, 1600, 1634, 1651, 1685, 1702, 1719, 1787, 1804, 1821, 1872, 1889, 1923, 1940, 1957, 1974, 1991, 2008, 2025, 2042, 2059, 2093, 2110, 2127, 2161, 2178, 2195, 2212, 2229, 2246, 2263, 2280, 2314, 2331, 2348, 2365, 2382, 2399, 2416, 2433, 2450, 2467, 2484, 2501, 2518, 2535, 2552, 2569, 2586, 2603, 2620, 2637, 2654, 2671, 2688, 2705, 2722, 2739, 2773, 2790, 2824, 2841, 2858, 2926, 2943, 2960, 3011, 3028, 3062, 3079, 3096, 3113, 3130, 3147, 3164, 3181, 3198, 3232, 3249, 3266, and 3283; and/or a recombinant, purified, and/or isolated polypeptide selected from the group consisting of a polypeptide having at least 90% sequence identity to at least one of the sequences of SEQ ID Nos. 1, 18, 35, 52, 69, 86, 103, 120, 137, 154, 171, 188, 205, 222, 239, 256, 273, 290, 307, 324, 358, 375, 409, 426, 443, 511, 528, 545, 596, 613, 647, 664, 681, 698, 715, 732, 749, 766, 783, 817, 834, 851, 885, 902, 819, 936, 953, 970, 987, 1004, 1038, 1055, 1072, 1089, 1106, 1123, 1140, 1157, 1174, 1191, 1208, 1225, 1242, 1259, 1276, 1293, 1310, 1327, 1344, 1361, 1378, 1395, 1412, 1429, 1446, 1463, 1480, 1497, 1514, 1531, 1548, 1565, 1582, 1599, 1633, 1650, 1684, 1701, 1718, 1786, 1803, 1820, 1871, 1888, 1922, 1939, 1956, 1973, 1990, 2007, 2024, 2041, 2058, 2092, 2109, 2126, 2160, 2177, 2094, 2211, 2228, 2245, 2262, 2279, 2313, 2330, 2347, 2364, 2381, 2398, 2415, 2432, 2449, 2466, 2483, 2500, 2517, 2534, 2551, 2568, 2585, 2602, 2619, 2636, 2653, 2670, 2687, 2704, 2721, 2738, 2772, 2789, 2823, 2840, 2857, 2925, 2942, 2959, 3010, 3027, 3061, 3078, 3095, 3112, 3129, 3146, 3163, 3180, 3197, 3231, 3248, 3265, and 3282 to a nucleotide sequence in with which a recombination event is desired.

In embodiments of the invention any one of the isolated and/or purified polypeptides according to the invention may be enzymatically or functionally active at temperatures at or above about 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, and/or 95 degrees Celsius and/or may be enzymatically or functionally active at a pH at, below, and/or above 8, 7, 6, 5, 4, 3, 2, 1, and/or 0. In further embodiments of the invention, glycosylation, pegylation, and/or other posttranslational modification may be required for

- 71 -

the isolated and/or purified polypeptides according to the invention to be enzymatically or functionally active at pH at or below 8, 7, 6, 5, 4, 3, 2, 1, and/or 0 or at a temperatures at or above about 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, and/or 95 degrees Celsius.

The invention is described in additional detail in the following illustrative
5 examples. Although the examples may represent only selected embodiments of the invention, it should be understood that the following examples are illustrative and not limiting.

EXAMPLES

10 **Example 1: Recombination Using Nucleotide and Amino Acid Sequences from *Alicyclobacillus acidocaldarius***

Provided in SEQ ID Nos. 2, 19, 36, 53, 70, 87, 104, 121, 138, 155, 172, 189, 206,
223, 240, 257, 274, 291, 308, 325, 359, 376, 410, 427, 444, 512, 529, 546, 597, 614, 648,
665, 682, 699, 716, 733, 750, 767, 784, 818, 835, 852, 886, 903, 920, 937, 954, 971, 988,
15 1005, 1039, 1056, 1073, 1090, 1107, 1124, 1141, 1158, 1175, 1192, 1209, 1226, 1243,
1260, 1277, 1294, 1311, 1328, 1345, 1362, 1379, 1396, 1413, 1430, 1447, 1464, 1481,
1498, 1515, 1532, 1549, 1566, 1583, 1600, 1634, 1651, 1685, 1702, 1719, 1787, 1804,
1821, 1872, 1889, 1923, 1940, 1957, 1974, 1991, 2008, 2025, 2042, 2059, 2093, 2110,
2127, 2161, 2178, 2195, 2212, 2229, 2246, 2263, 2280, 2314, 2331, 2348, 2365, 2382,
20 2399, 2416, 2433, 2450, 2467, 2484, 2501, 2518, 2535, 2552, 2569, 2586, 2603, 2620,
2637, 2654, 2671, 2688, 2705, 2722, 2739, 2773, 2790, 2824, 2841, 2858, 2926, 2943,
2960, 3011, 3028, 3062, 3079, 3096, 3113, 3130, 3147, 3164, 3181, 3198, 3232, 3249,
3266, and 3283 are a nucleotide sequences isolated from *Alicyclobacillus acidocaldarius*
and coding for the polypeptides of SEQ ID Nos. 1, 18, 35, 52, 69, 86, 103, 120, 137, 154,
25 171, 188, 205, 222, 239, 256, 273, 290, 307, 324, 358, 375, 409, 426, 443, 511, 528, 545,
596, 613, 647, 664, 681, 698, 715, 732, 749, 766, 783, 817, 834, 851, 885, 902, 819, 936,
953, 970, 987, 1004, 1038, 1055, 1072, 1089, 1106, 1123, 1140, 1157, 1174, 1191, 1208,
1225, 1242, 1259, 1276, 1293, 1310, 1327, 1344, 1361, 1378, 1395, 1412, 1429, 1446,
1463, 1480, 1497, 1514, 1531, 1548, 1565, 1582, 1599, 1633, 1650, 1684, 1701, 1718,
30 1786, 1803, 1820, 1871, 1888, 1922, 1939, 1956, 1973, 1990, 2007, 2024, 2041, 2058,
2092, 2109, 2126, 2160, 2177, 2094, 2211, 2228, 2245, 2262, 2279, 2313, 2330, 2347,
2364, 2381, 2398, 2415, 2432, 2449, 2466, 2483, 2500, 2517, 2534, 2551, 2568, 2585,

- 72 -

2602, 2619, 2636, 2653, 2670, 2687, 2704, 2721, 2738, 2772, 2789, 2823, 2840, 2857, 2925, 2942, 2959, 3010, 3027, 3061, 3078, 3095, 3112, 3129, 3146, 3163, 3180, 3197, 3231, 3248, 3265, and 3282, and respectively. The nucleotide sequences of SEQ ID Nos. 2, 19, 36, 53, 70, 87, 104, 121, 138, 155, 172, 189, 206, 223, 240, 257, 274, 291, 308, 325, 359, 376, 410, 427, 444, 512, 529, 546, 597, 614, 648, 665, 682, 699, 716, 733, 750, 767, 784, 818, 835, 852, 886, 903, 920, 937, 954, 971, 988, 1005, 1039, 1056, 1073, 1090, 1107, 1124, 1141, 1158, 1175, 1192, 1209, 1226, 1243, 1260, 1277, 1294, 1311, 1328, 1345, 1362, 1379, 1396, 1413, 1430, 1447, 1464, 1481, 1498, 1515, 1532, 1549, 1566, 1583, 1600, 1634, 1651, 1685, 1702, 1719, 1787, 1804, 1821, 1872, 1889, 1923, 1940, 1957, 1974, 1991, 2008, 2025, 2042, 2059, 2093, 2110, 2127, 2161, 2178, 2195, 2212, 2229, 2246, 2263, 2280, 2314, 2331, 2348, 2365, 2382, 2399, 2416, 2433, 2450, 2467, 2484, 2501, 2518, 2535, 2552, 2569, 2586, 2603, 2620, 2637, 2654, 2671, 2688, 2705, 2722, 2739, 2773, 2790, 2824, 2841, 2858, 2926, 2943, 2960, 3011, 3028, 3062, 3079, 3096, 3113, 3130, 3147, 3164, 3181, 3198, 3232, 3249, 3266, and 3283 are placed into expression vectors using techniques standard in the art. The vectors are then provided to cells such as bacteria cells or eukaryotic cells such as Sf9 cells or CHO cells. In conjunction with the normal machinery in present in the cells, the vectors comprising SEQ ID Nos. 2, 19, 36, 53, 70, 87, 104, 121, 138, 155, 172, 189, 206, 223, 240, 257, 274, 291, 308, 325, 359, 376, 410, 427, 444, 512, 529, 546, 597, 614, 648, 665, 682, 699, 716, 733, 750, 767, 784, 818, 835, 852, 886, 903, 920, 937, 954, 971, 988, 1005, 1039, 1056, 1073, 1090, 1107, 1124, 1141, 1158, 1175, 1192, 1209, 1226, 1243, 1260, 1277, 1294, 1311, 1328, 1345, 1362, 1379, 1396, 1413, 1430, 1447, 1464, 1481, 1498, 1515, 1532, 1549, 1566, 1583, 1600, 1634, 1651, 1685, 1702, 1719, 1787, 1804, 1821, 1872, 1889, 1923, 1940, 1957, 1974, 1991, 2008, 2025, 2042, 2059, 2093, 2110, 2127, 2161, 2178, 2195, 2212, 2229, 2246, 2263, 2280, 2314, 2331, 2348, 2365, 2382, 2399, 2416, 2433, 2450, 2467, 2484, 2501, 2518, 2535, 2552, 2569, 2586, 2603, 2620, 2637, 2654, 2671, 2688, 2705, 2722, 2739, 2773, 2790, 2824, 2841, 2858, 2926, 2943, 2960, 3011, 3028, 3062, 3079, 3096, 3113, 3130, 3147, 3164, 3181, 3198, 3232, 3249, 3266, and 3283. The polypeptides of SEQ ID Nos. 1, 18, 35, 52, 69, 86, 103, 120, 137, 154, 171, 188, 205, 222, 239, 256, 273, 290, 307, 324, 358, 375, 409, 426, 443, 511, 528, 545, 596, 613, 647, 664, 681, 698, 715, 732, 749, 766, 783, 817, 834, 851, 885, 902, 819, 936, 953, 970, 987, 1004, 1038, 1055, 1072, 1089, 1106, 1123, 1140, 1157, 1174, 1191, 1208, 1225, 1242, 1259,

- 73 -

1276, 1293, 1310, 1327, 1344, 1361, 1378, 1395, 1412, 1429, 1446, 1463, 1480, 1497, 1514, 1531, 1548, 1565, 1582, 1599, 1633, 1650, 1684, 1701, 1718, 1786, 1803, 1820, 1871, 1888, 1922, 1939, 1956, 1973, 1990, 2007, 2024, 2041, 2058, 2092, 2109, 2126, 2160, 2177, 2094, 2211, 2228, 2245, 2262, 2279, 2313, 2330, 2347, 2364, 2381, 2398, 5 2415, 2432, 2449, 2466, 2483, 2500, 2517, 2534, 2551, 2568, 2585, 2602, 2619, 2636, 2653, 2670, 2687, 2704, 2721, 2738, 2772, 2789, 2823, 2840, 2857, 2925, 2942, 2959, 3010, 3027, 3061, 3078, 3095, 3112, 3129, 3146, 3163, 3180, 3197, 3231, 3248, 3265, and 3282 are then isolated and/or purified. The isolated and/or purified polypeptides of SEQ ID Nos. 1, 18, 35, 52, 69, 86, 103, 120, 137, 154, 171, 188, 205, 222, 239, 256, 273, 290, 10 307, 324, 358, 375, 409, 426, 443, 511, 528, 545, 596, 613, 647, 664, 681, 698, 715, 732, 749, 766, 783, 817, 834, 851, 885, 902, 819, 936, 953, 970, 987, 1004, 1038, 1055, 1072, 1089, 1106, 1123, 1140, 1157, 1174, 1191, 1208, 1225, 1242, 1259, 1276, 1293, 1310, 1327, 1344, 1361, 1378, 1395, 1412, 1429, 1446, 1463, 1480, 1497, 1514, 1531, 1548, 1565, 1582, 1599, 1633, 1650, 1684, 1701, 1718, 1786, 1803, 1820, 1871, 1888, 1922, 15 1939, 1956, 1973, 1990, 2007, 2024, 2041, 2058, 2092, 2109, 2126, 2160, 2177, 2094, 2211, 2228, 2245, 2262, 2279, 2313, 2330, 2347, 2364, 2381, 2398, 2415, 2432, 2449, 2466, 2483, 2500, 2517, 2534, 2551, 2568, 2585, 2602, 2619, 2636, 2653, 2670, 2687, 2704, 2721, 2738, 2772, 2789, 2823, 2840, 2857, 2925, 2942, 2959, 3010, 3027, 3061, 3078, 3095, 3112, 3129, 3146, 3163, 3180, 3197, 3231, 3248, 3265, and 3282 are then 20 each demonstrated to have one or more of the activities provided in Table 1.

The isolated and/or purified polypeptides of SEQ ID Nos. 1, 18, 35, 52, 69, 86, 103, 120, 137, 154, 171, 188, 205, 222, 239, 256, 273, 290, 307, 324, 358, 375, 409, 426, 443, 511, 528, 545, 596, 613, 647, 664, 681, 698, 715, 732, 749, 766, 783, 817, 834, 851, 885, 902, 819, 936, 953, 970, 987, 1004, 1038, 1055, 1072, 1089, 1106, 1123, 1140, 1157, 25 1174, 1191, 1208, 1225, 1242, 1259, 1276, 1293, 1310, 1327, 1344, 1361, 1378, 1395, 1412, 1429, 1446, 1463, 1480, 1497, 1514, 1531, 1548, 1565, 1582, 1599, 1633, 1650, 1684, 1701, 1718, 1786, 1803, 1820, 1871, 1888, 1922, 1939, 1956, 1973, 1990, 2007, 2024, 2041, 2058, 2092, 2109, 2126, 2160, 2177, 2094, 2211, 2228, 2245, 2262, 2279, 2313, 2330, 2347, 2364, 2381, 2398, 2415, 2432, 2449, 2466, 2483, 2500, 2517, 2534, 30 2551, 2568, 2585, 2602, 2619, 2636, 2653, 2670, 2687, 2704, 2721, 2738, 2772, 2789, 2823, 2840, 2857, 2925, 2942, 2959, 3010, 3027, 3061, 3078, 3095, 3112, 3129, 3146, 3163, 3180, 3197, 3231, 3248, 3265, and 3282 are demonstrated to have activity as at least

- 74 -

one of a ATP-dependent DNA helicase recG, ATP-dependent DNA ligase, ATP-dependent endopeptidase clp proteolytic subunit, Chromosome partitioning protein, Crossover junction endodeoxyribonuclease ruvC, Deoxyuridine 5'-triphosphate nucleotidohydrolase, DNA adenine methylase, DNA helicase, DNA

5 integration/recombination/inversion, DNA polymerase I, DNA polymerase III beta chain, DNA polymerase IV, DNA primase, DNA repair protein radC, DNA replication and repair protein recF, DNA replication protein dnaD, DNA topoisomerase I, DNA/RNA helicase (DEAD/DEAH box family), DNA-binding protein HU, Fe-S oxidoreductase, Glycerophosphoryl diester phosphodiesterase, HNH endonuclease family protein,

10 Holliday junction DNA helicase ruvB, Integrase/recombinase (XerC/CodV family), Ligase/carboxylase family protein, LtrC-like protein, Macrolide-efflux protein, NAD-dependent DNA ligase, Nicotinate phosphoribosyltransferase, nodulin-26, Phage antirepressor protein, Phage protein, Phosphinothricin N-acetyltransferase, Phosphohydrolase, RecA protein, Recombination protein recR, Replicative DNA helicase,

15 Ribonucleoside-diphosphate reductase alpha chain, Ribonucleoside-diphosphate reductase beta chain, Ribose-phosphate pyrophosphokinase, Serine/threonine protein phosphatase, Single-strand DNA binding protein, Single-stranded DNA-binding protein, Site-specific recombinase, Site-specific resolvase / integrase, Thioredoxin, Thymidine kinase, Transcriptional regulator / Lex A repressor, Transcriptional regulator, Cro/CI family,

20 Transposase, TRSE protein, Two-component response regulator, Type II restriction-modification system methylation subunit.

While this invention has been described in certain embodiments, the present invention can be further modified within the spirit and scope of this disclosure. This application is therefore intended to cover any variations, uses, or adaptations of the

25 invention using its general principles. Further, this application is intended to cover such departures from the present disclosure as come within known or customary practice in the art to which this invention pertains and which fall within the limits of the appended claims and their legal equivalents.

BIBLIOGRAPHIC REFERENCES

- Barany, F., 1991, *PNAS. USA*, 88: 189-193.
- Buckholz, R. G., 1993, Yeast systems for the expression of heterologous gene products.
Curr. Op. Biotechnology 4: 538-542.
- 5 Burg, J. L. et al., 1996, *Mol. and Cell. Probes*, 10: 257-271.
- Chu, B. C. F. et al., 1986, *NAR*, 14: 5591-5603.
- Duck, P. et al., 1990, *Biotechniques*, 9: 142-147.
- Edwards, C. P., and Aruffo, A., 1993, Current applications of COS cell based transient expression systems. *Curr. Op. Biotechnology* 4: 558-563.
- 10 Guateli, J. C. et al., 1990, *PNAS. USA*, 87: 1874-1878.
- Houben-Weyl, 1974, in *Methode der Organischen Chemie*, E. Wunsch Ed., Volume 15-I and 15-II, Thieme, Stuttgart.
- Innis, M. A. et al., 1990, in *PCR Protocols. A guide to Methods and Applications*, San Diego, Academic Press.
- 15 Kievitis, T. et al., 1991, *J. Virol. Methods*, 35: 273-286.
- Kohler, G. et al., 1975, *Nature*, 256(5517): 495-497.
- Kwoh, D. Y. et al., 1989, *PNAS. USA*, 86: 1173-1177.
- Luckow, V. A., 1993, *Baculovirus* systems for the expression of human gene products.
Curr. Op. Biotechnology 4: 564-572.
- 20 Matthews, J. A. et al., 1988, *Anal. Biochem.*, 169: 1-25.
- Merrifield, R. D., 1966, *J. Am. Chem. Soc.*, 88(21): 5051-5052.
- Miele, E. A. et al., 1983, *J. Mol. Biol.*, 171: 281-295.
- Olins, P. O., and Lee, S. C., 1993, Recent advances in heterologous gene expression in *E. coli*. *Curr. Op. Biotechnology* 4: 520-525.
- 25 Rolfs, A. et al., 1991, In *PCR Topics. Usage of Polymerase Chain reaction in Genetic and Infectious Disease*. Berlin: Springer-Verlag.
- Sambrook, J. et al., 1989, In *Molecular cloning: A Laboratory Manual*. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.
- Sanchez-Pescador, R., 1988, *J. Clin. Microbiol.*, 26(10): 1934-1938.
- 30 Segev D., 1992, in "Non-radioactive Labeling and Detection of Biomolecules," Kessler C. Springer Verlag, Berlin, New-York: 197-205.
- Urdea, M. S., 1988, *Nucleic Acids Research*, 16: 4937-4957.

- 76 -

Walker, G. T. et al., 1992, *NAR* 20: 1691-1696.

Walker, G. T. et al., 1992, *PNAS. USA*, 89: 392-396.

White, B. A. et al., 1997, *Methods in Molecular Biology*, 67, Humana Press, Totowa, NJ.

- 77 -

CLAIMS

What is claimed is:

- 5 1. An isolated or purified nucleic acid sequence comprising a nucleic acid sequence encoding a polypeptide selected from the group consisting of a polypeptides having at least 90% sequence identity to SEQ ID Nos. 1, 18, 35, 52, 69, 86, 103, 120, 137, 154, 171, 188, 205, 222, 239, 256, 273, 290, 307, 324, 358, 375, 409, 426, 443, 511, 528, 545, 596, 613, 647, 664, 681, 698, 715, 732, 749, 766, 783, 817, 834, 851, 885, 902, 819,
10 936, 953, 970, 987, 1004, 1038, 1055, 1072, 1089, 1106, 1123, 1140, 1157, 1174, 1191, 1208, 1225, 1242, 1259, 1276, 1293, 1310, 1327, 1344, 1361, 1378, 1395, 1412, 1429, 1446, 1463, 1480, 1497, 1514, 1531, 1548, 1565, 1582, 1599, 1633, 1650, 1684, 1701, 1718, 1786, 1803, 1820, 1871, 1888, 1922, 1939, 1956, 1973, 1990, 2007, 2024, 2041, 2058, 2092, 2109, 2126, 2160, 2177, 2094, 2211, 2228, 2245, 2262, 2279, 2313, 2330,
15 2347, 2364, 2381, 2398, 2415, 2432, 2449, 2466, 2483, 2500, 2517, 2534, 2551, 2568, 2585, 2602, 2619, 2636, 2653, 2670, 2687, 2704, 2721, 2738, 2772, 2789, 2823, 2840, 2857, 2925, 2942, 2959, 3010, 3027, 3061, 3078, 3095, 3112, 3129, 3146, 3163, 3180, 3197, 3231, 3248, 3265, and 3282.
- 20 2. The isolated or purified nucleic acid sequence of claim 1, wherein the polypeptide has activity at or below about pH 8.
3. The isolated or purified nucleic acid sequence of claim 1, wherein the polypeptide has activity at a temperature at or above about 50 degrees Celsius.
- 25 4. The isolated or purified nucleic acid sequence of claim 1, wherein the nucleic acid sequence is present in a vector.

- 78 -

5. An isolated or purified polypeptide comprising a polypeptide selected from the group consisting of polypeptide having at least 90% sequence identity to SEQ ID Nos. 1, 18, 35, 52, 69, 86, 103, 120, 137, 154, 171, 188, 205, 222, 239, 256, 273, 290, 307, 324, 358, 375, 409, 426, 443, 511, 528, 545, 596, 613, 647, 664, 681, 698, 715, 732, 749, 766, 783, 817, 834, 851, 885, 902, 819, 936, 953, 970, 987, 1004, 1038, 1055, 1072, 1089, 1106, 1123, 1140, 1157, 1174, 1191, 1208, 1225, 1242, 1259, 1276, 1293, 1310, 1327, 1344, 1361, 1378, 1395, 1412, 1429, 1446, 1463, 1480, 1497, 1514, 1531, 1548, 1565, 1582, 1599, 1633, 1650, 1684, 1701, 1718, 1786, 1803, 1820, 1871, 1888, 1922, 1939, 1956, 1973, 1990, 2007, 2024, 2041, 2058, 2092, 2109, 2126, 2160, 2177, 2094, 2211, 2228, 2245, 2262, 2279, 2313, 2330, 2347, 2364, 2381, 2398, 2415, 2432, 2449, 2466, 2483, 2500, 2517, 2534, 2551, 2568, 2585, 2602, 2619, 2636, 2653, 2670, 2687, 2704, 2721, 2738, 2772, 2789, 2823, 2840, 2857, 2925, 2942, 2959, 3010, 3027, 3061, 3078, 3095, 3112, 3129, 3146, 3163, 3180, 3197, 3231, 3248, 3265, and 3282.

6. The isolated or purified polypeptide of claim 5, wherein the polypeptide has activity at or below about pH 8.

7. The isolated or purified polypeptide of claim 5, wherein the polypeptide has activity at a temperature at or above about 50 degrees Celsius.

8. The isolated or purified polypeptide of claim 5, wherein the polypeptide is glycosylated, pegylated, or otherwise post-translationally modified.

9. The isolated or purified polypeptide of claim 5, wherein the polypeptide has an activity selected from the group consisting of ATP-dependent DNA helicase recG, ATP-dependent DNA ligase, ATP-dependent endopeptidase clp proteolytic subunit, Chromosome partitioning protein, Crossover junction endodeoxyribonuclease ruvC, Deoxyuridine 5'-triphosphate nucleotidohydrolase, DNA adenine methylase, DNA helicase, DNA integration/recombination/inversion, DNA polymerase I, DNA polymerase III beta chain, DNA polymerase IV, DNA primase, DNA repair protein radC, DNA replication and repair protein recF, DNA replication protein dnaD, DNA topoisomerase I, DNA/RNA helicase (DEAD/DEAH box family), DNA-binding protein HU, Fe-S

- 79 -

oxidoreductase, Glycerophosphoryl diester phosphodiesterase, HNH endonuclease family protein, Holliday junction DNA helicase *ruvB*, Integrase/recombinase (*XerC/CodV* family), Ligase/carboxylase family protein, LtrC-like protein, Macrolide-efflux protein, NAD-dependent DNA ligase, Nicotinate phosphoribosyltransferase, nodulin-26, Phage
 5 antirepressor protein, Phage protein, Phosphinothricin N-acetyltransferase, Phosphohydrolase, RecA protein, Recombination protein *recR*, Replicative DNA helicase, Ribonucleoside-diphosphate reductase alpha chain, Ribonucleoside-diphosphate reductase beta chain, Ribose-phosphate pyrophosphokinase, Serine/threonine protein phosphatase, Single-strand DNA binding protein, Single-stranded DNA-binding protein, Site-specific
 10 recombinase, Site-specific resolvase / integrase, Thioredoxin, Thymidine kinase, Transcriptional regulator / Lex A repressor, Transcriptional regulator, Cro/CI family, Transposase, TRSE protein, Two-component response regulator, Type II restriction-modification system methylation subunit activity.

15 10. A method of modulating or altering recombination, the method comprising: providing a recombinant, purified, and/or isolated nucleotide sequence comprising a nucleotide sequence selected from the group consisting of a nucleotide sequences having at least 90% sequence identity to at least one of the sequences of SEQ ID Nos. 2, 19, 36, 53, 70, 87, 104, 121, 138, 155, 172, 189, 206, 223, 240, 257, 274, 291, 308, 325, 359, 376,
 20 410, 427, 444, 512, 529, 546, 597, 614, 648, 665, 682, 699, 716, 733, 750, 767, 784, 818, 835, 852, 886, 903, 920, 937, 954, 971, 988, 1005, 1039, 1056, 1073, 1090, 1107, 1124, 1141, 1158, 1175, 1192, 1209, 1226, 1243, 1260, 1277, 1294, 1311, 1328, 1345, 1362, 1379, 1396, 1413, 1430, 1447, 1464, 1481, 1498, 1515, 1532, 1549, 1566, 1583, 1600, 1634, 1651, 1685, 1702, 1719, 1787, 1804, 1821, 1872, 1889, 1923, 1940, 1957, 1974,
 25 1991, 2008, 2025, 2042, 2059, 2093, 2110, 2127, 2161, 2178, 2195, 2212, 2229, 2246, 2263, 2280, 2314, 2331, 2348, 2365, 2382, 2399, 2416, 2433, 2450, 2467, 2484, 2501, 2518, 2535, 2552, 2569, 2586, 2603, 2620, 2637, 2654, 2671, 2688, 2705, 2722, 2739, 2773, 2790, 2824, 2841, 2858, 2926, 2943, 2960, 3011, 3028, 3062, 3079, 3096, 3113, 3130, 3147, 3164, 3181, 3198, 3232, 3249, 3266, and 3283 to a sequence in which a
 30 genetic recombination event is desired.

- 80 -

11. The method according to claim 10, wherein the modulation or alteration of recombination occurs at or below about pH 8.

12. The method according to claim 10, wherein the modulation or alteration of recombination occurs at a temperature at or above about 50 degrees Celsius.

13. The method according to claim 10, wherein the recombinant, purified, or isolated polypeptide is glycosylated, pegylated, or otherwise post-translationally modified.

14. The method according to claim 10, wherein the recombinant, purified, or isolated polypeptide has an activity selected from the group consisting of ATP-dependent DNA helicase recG, ATP-dependent DNA ligase, ATP-dependent endopeptidase clp proteolytic subunit, Chromosome partitioning protein, Crossover junction endodeoxyribonuclease ruvC, Deoxyuridine 5'-triphosphate nucleotidohydrolase, DNA adenine methylase, DNA helicase, DNA integration/recombination/inversion, DNA polymerase I, DNA polymerase III beta chain, DNA polymerase IV, DNA primase, DNA repair protein radC, DNA replication and repair protein recF, DNA replication protein dnaD, DNA topoisomerase I, DNA/RNA helicase (DEAD/DEAH box family), DNA-binding protein HU, Fe-S oxidoreductase, Glycerophosphoryl diester phosphodiesterase, HNH endonuclease family protein, Holliday junction DNA helicase ruvB, Integrase/recombinase (XerC/CodV family), Ligase/carboxylase family protein, LtrC-like protein, Macrolide-efflux protein, NAD-dependent DNA ligase, Nicotinate phosphoribosyltransferase, nodulin-26, Phage antirepressor protein, Phage protein, Phosphinothricin N-acetyltransferase, Phosphohydrolase, RecA protein, Recombination protein recR, Replicative DNA helicase, Ribonucleoside-diphosphate reductase alpha chain, Ribonucleoside-diphosphate reductase beta chain, Ribose-phosphate pyrophosphokinase, Serine/threonine protein phosphatase, Single-strand DNA binding protein, Single-stranded DNA-binding protein, Site-specific recombinase, Site-specific resolvase / integrase, Thioredoxin, Thymidine kinase, Transcriptional regulator / Lex A repressor, Transcriptional regulator, Cro/CI family, Transposase, TRSE protein, Two-component response regulator, Type II restriction-modification system methylation subunit activity.

FIG. 1

```
gb|AAV21825.1|-----SVSTTVSSTSS----PST
ref|XP_001743680.1|-----TTSTTSTTSTTSTNSTNSTT
ref|XP_001525241.1|-----
RAAC03697MLGCVKVGLGTGKEAIQRVLNWSCTHFIQRERFMMRKTIAAIVTLGTVSSMLAGTVYAAT
ref|YP_308280.1|-----TGSSFIEGNVYADDGIDLG
ref|ZP_01916690.1|-----SGPLQSGEQGQNTVTA

gb|AAV21825.1|TLTSTSTTNQTTNTTTASSTPANTTLTSTSTPTNQTTTTNSSG---TQTSNTSSTQPPTS
ref|XP_001743680.1|STTSTTSTTSTTSTTSTTSTTSTTSTTSTTSTTSTSTSTGNTSS---SAISNTTANSTDTS
ref|XP_001525241.1|--TSTTSTTSTTNTNVPANSTASTTTNST--ANSTTSTANNNSN---QLAQDKAAKYE---
RAAC03697NTTNATSTGNTANTTASATSNATSNTTGSAPTNPVLTQKSAVYQLIEVFNTIPSWPTGL
ref|YP_308280.1|WSTPITGDAVTPADINRPGNISGTVTTGAAPQLPNVLTDAEIA---AIVTDIQNETAFAA
ref|ZP_01916690.1|PNSSESGALTNSAPQNRSTSTSTSSNSPGSAPANPGIATQKSAPSEPVRPSRPTPPAPSGS
: : .. . . : : .: *

gb|AAV21825.1|QSLDTST-----
ref|XP_001743680.1|TSMSPSRT-NGSDTPSSPAPPA-----
ref|XP_001525241.1|-----
RAAC03697QSIDPATQNAQWDWKYKPIPPSQWVGLACVFPGDPESKYAADLGIQMTDPNGPVTAGQLA
ref|YP_308280.1|FSPGAI TRSSDWN IKYYPVPAS-----
ref|ZP_01916690.1|SGPQPAAPVSQ-----

gb|AAV21825.1|-----
ref|XP_001743680.1|-----
ref|XP_001525241.1|-----
RAAC03697QWIVDWEVKAR
ref|YP_308280.1|-----
ref|ZP_01916690.1|-----
```

2/228

FIG. 2

```
gb|EAZ41188.1|
RAAC02297
ref|YP_429214.1|
gb|AAR38445.1|
ref|ZP_01774730.1|
ref|ZP_01171092.1|
-----ARVVHREHAMRAQAG-----LDRVEQQQAAMGTGQFTGH
-MVPFDPSCHARFVHHRNHVLQAETGHLKGGDNGMGYVILALGLVPSVLLVMGAGQGEYA
-----LLRPAVERCIASWSG-----GALR-AIEQLGND----I
-----NGLLVGSYYTLLALGLS-----IIFSLGGVVNLAHGAFYAVGAYL
-----IARVAAVDRD-----QLLLVDET-----
KRNSFLSLYDMLLSEIEMMQQTRINAYSLIRTLKPYFKELDGTITRLLANWTNDKGPEY
.

gb|EAZ41188.1|
RAAC02297
ref|YP_429214.1|
gb|AAR38445.1|
ref|ZP_01774730.1|
ref|ZP_01171092.1|
GIERRRHRAAGVAFAHRLQEHRFDEPLVT-----
GIRTRARIAAGWYGAKMRVRKRLEDEQLVSLLRKSGTLKAYQYHYLRIGLTLVFLLMGV
NLKEADVLI-----VLMQIVEGGVAK-----
TVLFTDILGFG-----LAFLITPAVVG-----
-HAFRGRVAGG-----YLHRHLSQ-----
ALDIFAKEIGTNEAKSLANVLKKFDENKRD-----

gb|EAZ41188.1|
RAAC02297
ref|YP_429214.1|
gb|AAR38445.1|
ref|ZP_01774730.1|
ref|ZP_01171092.1|
-----CGILEHLAQAGL-----
VGLLHGRLPMLFPLVWVFGLEYRPPFMYYGFLALQKQAALERDKALYLLYRLLLQEA
-----LAGVMEEAL-----
-----LIGIAIEALFLR-----
-----ELPCVGDWVCLEK-----
-----TIILSLKGMEDMFIN-----
:
:

gb|EAZ41188.1|
RAAC02297
ref|YP_429214.1|
gb|AAR38445.1|
ref|ZP_01774730.1|
ref|ZP_01171092.1|
-----VVRRHGDH-----
AFRGRPLGVYDMLRRQLHRVPVLSPLERCLHDWVDDPAVALRRFGDEVGTSQAKALAHM
-----RLDEL-----
-----RLYNKDPILTLLFTFGLA-----
-----APGDDVGVVALLERRT-----
-----SQIENYRRKRKLYVDLAKL-----
.

gb|EAZ41188.1|
RAAC02297
ref|YP_429214.1|
gb|AAR38445.1|
ref|ZP_01774730.1|
ref|ZP_01171092.1|
LIEIEEAGVAVALDVLQTNLERFRADRIAAAFRAHLNTRSILATALTMLGLGATSFDFLMVV
-----
MSGEQALRLIFGASGIPFSIPEFLSGQLH-----
LLRRKSAGDAIEYQMIASNLD-----
PIKATHGLIILNFVVVIFMVSYLMKDS-----

gb|EAZ41188.1|
RAAC02297
ref|YP_429214.1|
gb|AAR38445.1|
ref|ZP_01774730.1|
ref|ZP_01171092.1|
-----
IQVYAGALMRASAGG
-----
-----
-----
-----
```


3/228

FIG. 3

```

ref|ZP_01900573.1|
RAAC02298
ref|ZP_01301851.1|
ref|ZP_01171091.1|
ref|ZP_02598168.1|
ref|ZP_01575699.1|

-----
MGGDIVANALNTLAQIVLYGVACAGLLWLFEQVWRDGWAWAKRTYYGIGHRQLLRVSEGF
-----

ref|ZP_01900573.1|
RAAC02298
ref|ZP_01301851.1|
ref|ZP_01171091.1|
ref|ZP_02598168.1|
ref|ZP_01575699.1|

-----
RGRQTTQLPWWMGWHDHLETLQLATLKRPHRDAVSRFIVGSVTA AVIVGVLT SVVTHNPF
-----
-----VSTFIVLSLLIGFLTFSLILVKYQDAF
-----HLELLLSSISKNSKKINVFNFLEASMLVFVITVSVLYFLIHDLV
-----NHIYKVLTA SLKEISDLGTYSFMIGSISLFIFS-FLTALKLFNFI

ref|ZP_01900573.1|
RAAC02298
ref|ZP_01301851.1|
ref|ZP_01171091.1|
ref|ZP_02598168.1|
ref|ZP_01575699.1|

-----
YALVMFALT VVAPYGVQLIRRYHMSIRNSYDIGTLLSVIVPEYRQHGSMLHALQ-STVE
-----
LGLIIGTVVSIIPYIILHVQLRNTRNAVGNQLTEIVELIIHAYGSSSSDMYQALK-VTQS
FALGIGITFGILPYVAIRYRLTMLRLKTQYAFLEIYHVLFQNYQSTSKDIYYTML-NAVK
ISLGIATIIAASPYVLLRLKLSVQIEGSYDANVLVPSITNEYKQHYFNMINAIENCAVR

ref|ZP_01900573.1|
RAAC02298
ref|ZP_01301851.1|
ref|ZP_01171091.1|
ref|ZP_02598168.1|
ref|ZP_01575699.1|

-----
-----IETISYRLSDHTTADDASR--YRSEDNVGEAWQKEPMTLRRLTLLMDELI I
HLPPSPIRRAVARLTDRLTDYTT PDDARRALHRFVKELGTSWAAQLASDIEHALLDGVNV
-----LGSRIAKKTLQAMEHH--LVYGAALAGAF-----LLMLAWAWFYF
HIAEKELRSILVRLISDLQTARTEEEMRLSIDLFI FT CGNSWSMRLGNIILKSYLHQENV
ETKNKELKHIYMKLLSSLQKDRGNLEFERAVNVFSYSINSTFAKRFAKLLKKAHIDRGDI
EDIGSYSRKNLFRLSLALKAYYSEEDLDKAIERFVFAYNTEWAMLLGLNIMAIHKGIVV
:      :      :      :      .

ref|ZP_01900573.1|
RAAC02298
ref|ZP_01301851.1|
ref|ZP_01171091.1|
ref|ZP_02598168.1|
ref|ZP_01575699.1|

DNAAIELLEKECSEEIDAAVTRYLA-----
E-ASLALLHKEFQEIEDARKSQNLARIDTLLISCVPFLMWPVMMV---LFYLFVSRNIFQ
TRTGRGLGSNQFLRGARFGTARQLRR-----LLW-----
MSALIQLQN-QMVNNQKMLEQEKASSYDAFVDATLT VILFPISLIG--AKFMTKPQSWMA
TMSLMDLNA-DIKRRKQDIQTDKTRKLETVILGYSPMVLFP-LMIF--LAYRISGVVDFW
SSGLEDILK-KLKDSSEQVEVSKRYNTDAFAIKFLLIPLYIGGILFSISTFDFTLRKYFE
:      .

ref|ZP_01900573.1|
RAAC02298
ref|ZP_01301851.1|
ref|ZP_01171091.1|
ref|ZP_02598168.1|
ref|ZP_01575699.1|

-----
YQFANPTGFTWFLLTLLATLGSFIIIGITFYKPKQDI
-----
LQFGEKATLMLFILT TIMVVISLLVGLIIRKPKNDL
YVFQQKTP IILFTISLVMSIFSVLTAIVMSKPRADI
YQFLNPIGLRTGIISFMGIIVSFISLRMIRKPKYDI

```

FIG. 4

```

ref|ZP_01171090.1|-----NSEWLERQ
ref|ZP_02598167.1|-----EKAKSVKQTHEFSEVLEIVKKS IENIYEDRGISSDEKIKRQ
RAAC02299MSVHELLPSEPFRADRYVLETQ RERAKDYGEVPFDRVIEAFQ QYLGEVGGKDDVDSQYL
ref|ZP_01575700.1|-----L
ref|ZP_02849387.1|-----SARLRLDQSAVTDPLDRGEAAGRGADFGRLAE EIRS YLAAPRGLTEEERQYS
ref|NP_627754.1|-----

ref|ZP_01171090.1|---HQAIIGDKVAIEYFLAEIEKVLRTK-NITSKDFFSFFNSLSEAI FHEIWGLSVLAKW
ref|ZP_02598167.1|EIEHNAILGDHEAEKILTKEIEKCLREQ-NLLDVKYPDFFD SLAQALFHEIYGFGAFYKW
RAAC02299HRKYRALIGDEAAKQYFIHRIHDFLRERPEFQ NTRYPRYPDLPEAIFQHALGFGPMSVW
ref|ZP_01575700.1|ERHRNAILGKPTENVYLKDKIREYLKAK- RLENERYPYTKNLIDAI FHENWGIAGIAEW
ref|ZP_02849387.1|ETLNRAVLGFASEREQVLAIIADRLIRLRIHQ LDGYKHPYATLAEALFAEVIGLNVLELV
ref|NP_627754.1|-----EQLRSELVSGSGPLEPL
: : . * :

ref|ZP_01171090.1|ERYPE-SEACCIRGTQLWIDIDG--QFVKQEEEFESLT VVERIKRAFVIRRPDSVINRES
ref|ZP_02598167.1|KKYPE-SVSASIIGKEIWFKING--KFVKQEEELRDEEH IYEIFRALEVGHKGLKINHEN
RAAC02299FANP--TESATVNGTQILFGVKGSNTKVLQPFADSIDQVKRLVRTLDRDPANQVNQTN
ref|ZP_01575700.1|MDMPE-SSSAKIIGDRIYFFING--KQVLKEQRISK-KRFEQLRQLM LDETNRANEN-
ref|ZP_02849387.1|LARKDGLEEIQVVGSIYEVRDG--QTMLSSYRFDHERDVERIQ QNLVLYNND-RINPRK
ref|NP_627754.1|LADPSVTDVLVSAPDRVWVDRGG--GLELAPVSFPDAAAVRRLA QRLAT-VAGRRLDDAR
: : * : : : : :

ref|ZP_01171090.1|PELEIEREDGSRITMIQ--PPRSRENYIMIRRFIVNKYS LHDQASR--GTIPQEDIPIFQ
ref|ZP_02598167.1|PRAEIEMKDGTRVNIIR--PPANLFPVIVFRRFIIKNFS FEEQARR--KTISSEDVELME
RAAC02299HWTQVDMINGTRVTIFA--PPLSETYVLVFRQYTFHRYT FEHEAEM--RTIPADSVEWWK
ref|ZP_01575700.1|YSELYMYSGERVTVYTGKVIDGQSVMVFRKYVVKVLT FEEQARR--GTIPVELVPALE
ref|ZP_02849387.1|RWAEVMLRDGSRVTMTG--FGFTSKPTLTIRFFT VRSFSLEALCSAPYHTLSLAMRNMLL
ref|NP_627754.1|PWVDARLPDGTRLHAVLP-PVAVDCTCLALRVVRPRAFTI RELAAA--GTVPPGGDRVLR
: . * * : : * . : . * :

ref|ZP_01171090.1|ALARTMANMIVAGRVRS AKSTFMTTLIGERDDSFVGAVLEK-HFEVALSKHFPNRL--FF
ref|ZP_02598167.1|IISQLYLNMI IAGHVESGKSTMLKTIFASRSPEKIAIC IET-SPESFLKKDFPDRL--VY
RAAC02299LLSRLMLTMVTTGIRRS GKTTLFKVIFGARDPNLEVVTVERGT FEAHLKRDFPERAGRII
ref|ZP_01575700.1|ALVNCGVKVAFIGPVRSGKSTMLLTWQLYEDPELEGVLIQT -DPEIRIHEVMPKAP--IM
ref|ZP_02849387.1|DVLEARFNLV IIGPTNSGKTHLMKALIAELPDEERIVTIEG-RFEMMLGRDFPMKN--TV
ref|NP_627754.1|ALLRARLSFLVSGGTSGSKTTL SALLGLVGPDERIVLAED--SAELRPDHPHVVRLET
: . . * * : : . : : *

ref|ZP_01171090.1|EIQAKEGD---LHKAI PRLRMEHDFVVVGEIRS-LEIEAYLQSTERGERGSLSTFHLT
ref|ZP_02598167.1|DMYTINGN---IEDVIYSALRTDHDY IIFQEV RG-IEADGAMKGAERGTTGMMMTYHIT
RAAC02299ALKSPIDE---MASLFPAFLRSDAHYMMVPEIRS-SEVDLLILSRE RGN-GCLASYHSP
ref|ZP_01575700.1|PLIAGGKE---LFELSS EILKSDADYLVVQEV RDGYTAYIAVEAANKGTNRLKITAHLS
ref|ZP_02849387.1|EYEADEDDPQHRAEQAFKLALRQSPQRIVHAEIRD-MDANIYVRAC TRGHSGSMTTVHAN
ref|NP_627754.1|RPANQEGAGLVLTLEDLVRQALRMRPDR LVVGEVRG-PEVVHLLAALNTGHEGGCGTVHAN
: : * : : * : : * :

ref|ZP_01171090.1|DVEQVVEQLARLTLEDFPTRRFEVEVERIARNID- IITMDTER-DRSKKRVVGVT EVIW
ref|ZP_02598167.1|DPSRTPEQLAQHIVDAYANR KLENIERRVAKNLD-LGIIMKNDE-PKNEKRLMSIYEICY
RAAC02299YVTNIPRELADLALENNPSRDYRATYIRTAQSLD-VAITMWED--PTGRKI VTVGYAYEF
ref|ZP_01575700.1|NPEDFCYDIANKIQGVFGG-NIDYQMV RVANSFN-FLFEMVQLPGNRSQKRLKSIYEIRY
ref|ZP_02849387.1|TLEDVPEAITDMCMLDGRGMNPERLT KRIA EYVTQVGIEMRYLG---GRRVIARIGELSW
ref|NP_627754.1|AAADVPARLEALGTT--AGLNRAALHSQ LAAALS-VVLHLVRDR--AGRRRIAEVHVLER
: : * : : : : :

ref|ZP_01171090.1|DNQQRHYTQDLIRYSKLKDKYYYSSNI-----
ref|ZP_02598167.1|DYKGDKAWINYLMKYNELDERWEYNSEVSEALIKKMK-----
RAAC02299DHETESYTVTTWMKYHRATDTWTFHAEIPPSMRARLEDTYPDVLAAFEAFQRLASAYPF
ref|ZP_01575700.1|DTEANVISYHKICEYNKDTDSWCFSYSVGNKVTELGEFENPKALEVYKNTL KFLSEKYP-
ref|ZP_02849387.1|AN--NEVNVRDWARFDETSEQWIYPEKPSPKARTRLSE-----
ref|NP_627754.1|D-----

ref|ZP_01171090.1|-----
ref|ZP_02598167.1|-----
RAAC02299QGEAKRRVKIGG
ref|ZP_01575700.1|-----
ref|ZP_02849387.1|-----
ref|NP_627754.1|-----

```

SUBSTITUTE SHEET (RULE 26)

6/228

FIG. 6

```

emb|CAE47778.1|-----VVVAK--EDIAEMQTIYDTMVETK
ref|NP_967133.1|-----VVVAK--EDIAEMQTIYDTMVETK
emb|CAE47790.1|-----VVVAK--EDIAEMQTVYDTMVETK
RAAC02301MKSWMRYTLATTLFVVGVGSGSIAYNQFVSPMLTSEWVYVAR--TQLPADTPIQASDVERI
ref|ZP_02756760.1|IKNKANILLQFSLFILFAGG--VFLETSQVQKPVAVYQYS--RNIPENTVIQKGDYIKT
ref|ZP_01171088.1|MKPWLKITLGVLLSVFVIAFIIVWDTVIKDKIDSVEVAIVRPGVVIEKNQVISKDLLMME
                                     *      :      :

emb|CAE47778.1|ELPADFIQPDAITIPDEIIIGNVAAPPIRKGMVVKNLLTPGPD TGISLQVAPSKRAVTI
ref|NP_967133.1|ELPADFIQPDAITIPDEIIIGNVAAPPIRKGMVVKNLLTPGPD TGISLQVAPSKRAVTI
emb|CAE47790.1|ELPADFIQPDAITVPDEIIIGNVAAPPIRKGMIVKNLLTPGPD TGISLQVAPSKRAVTI
RAAC02301RVPKSEVSPDAITNAEGLMGTYTAQVVDQNVLTATA---LDTEPDP---FTITPGTEDVPI
ref|ZP_02756760.1|FKPKDILTKSMITNEKDINGKILTNNVYRSEYAIAKNNLEDPKLDFAKIDLSNLRKVS I
ref|ZP_01171088.1|QRNRATLVEGTVDMDVVGYEAKQKLYGNSILSERDVEFIFFTP----DPEKGEAIRPI
               : . : . : * : .. . . *

emb|CAE47778.1|PVDEVRGVAKLIRPGDRIDIYA AVDSGKGVN-----
ref|NP_967133.1|PVDEVRGVAKLIRPGDRIDIYA AVDSGKGVN-----
emb|CAE47790.1|PVDEVRGVAKLIRPGDRIDIYA AVDSGKGVN-----
RAAC02301PSTWIASVSETLRQGDYVDLIPIAE PQAGSNNVGMLTTTSEASEFKHLLVLSVHTDNNAE
ref|ZP_02756760.1|PVEMKDAVGGNLKKGDRVDLTFVKQGD SKNNDS--DSFTYAKTFMQDVLVYNVVDG G K
ref|ZP_01171088.1|PASWIYASPSTIRKDHID-----
* . : : * : *

emb|CAE47778.1|-----
ref|NP_967133.1|-----
emb|CAE47790.1|-----
RAAC02301VTSQQTGAPQSVGARGNGSGVPASVDVKMTSAEAQALASLIQQKYQLLIVGVSDDNRTGG
ref|ZP_02756760.1|KYVDQTEGTQSLANEKGEVVESGSLSI VTVAVTAQQAE EI-----
ref|ZP_01171088.1|-----

emb|CAE47778.1|----
ref|NP_967133.1|----
emb|CAE47790.1|----
RAAC02301GAK
ref|ZP_02756760.1|----
ref|ZP_01171088.1|----

```

7/228

FIG. 7

```
emb|CAJ49597.1|-----
ref|YP_158155.1|-----
RAAC02302      MLYLAFFPADLRERYAESTWLVDKQVLETSFAPGDILLSSWVPHRAEQLQVIHRARQE
ref|YP_061819.1|-----
ref|NP_779769.1|-----
ref|ZP_01171087.1|-----EILRMIEQWRFQYNDSLRIVYICERERTDPLLGSIVARN-----VLDIFYTRSI

emb|CAJ49597.1|-----
ref|YP_158155.1|-----
RAAC02302      GARVIFVGSKEDETDWKRLCALGVYDFAFFGDEVVLSVLDLIEHPRTPLDVRAYVDE
ref|YP_061819.1|-----
ref|NP_779769.1|-----
ref|ZP_01171087.1|-----PTKLLINQILEPPK---FSNVAKIGVTDIDLAN-----LEYLEEQAEGSGIEDE

emb|CAJ49597.1|-----
ref|YP_158155.1|-----
RAAC02302      TVWHLAREDPVVVEVETPKADVDEPTQDEAPFRPRWRPRWGKPSPPAPERLQVVQPRLV
ref|YP_061819.1|-----
ref|NP_779769.1|-----
ref|ZP_01171087.1|-----LTEEHANQPAKPIDIPTSKSKKKVELPSLKLKPEFHIHVHKPSKIRTEKIAKAIDRKIV

emb|CAJ49597.1|-----
ref|YP_158155.1|-----
RAAC02302      VVLGLWPRAGVTTITYLLAQLFAKQLPTRSVACIEHPRPWRMWDYFQLDARMPAENYRH
ref|YP_061819.1|-----
ref|NP_779769.1|-----
ref|ZP_01171087.1|-----VVISPFERSGSTFVS---HQLAYQIAERKIGITYFENPFKRPYTYDRFGGHLVLPNFKS
                                :
emb|CAJ49597.1|-----
ref|YP_158155.1|-----
RAAC02302      ALAQGRADVVLADRVDTPMGPHYWVSGSYSGQ-----PKAVIRS--DTPLRQW
ref|YP_061819.1|-----
ref|NP_779769.1|-----
ref|ZP_01171087.1|-----LYTSQPNANIDIDRKWTVEGVQIQALNPILEM-----PYEEKDISIEKFLRLF
                                :
emb|CAJ49597.1|-----
ref|YP_158155.1|-----
RAAC02302      SEVAGKRVCMQSN-----NRAEALARRYGATVQT-----
ref|YP_061819.1|-----
ref|NP_779769.1|-----
ref|ZP_01171087.1|-----LSAGDTPILIVDIG-----ADRQRPIYDELLSIASHVLVVMDCD-----

emb|CAJ49597.1|-----
ref|YP_158155.1|-----
RAAC02302      VPYLAPEMMQALWAGQFPDVATLTSELHELTERVVTPLLPARGV
ref|YP_061819.1|-----
ref|NP_779769.1|-----
ref|ZP_01171087.1|-----
```

8/228

FIG. 8

```

ref|YP_001371728.1| -----
ref|ZP_00630666.1| -----
ref|YP_674884.1| -----
ref|YP_001235767.1| -----
RAAC02303 MKRIGLFRIPTEPKDVFPWLVKRLVLSLGLGGVLGYLALLVVALSLLGFLVGDYQANGRG
ref|NP_046584.1| -----

ref|YP_001371728.1| -----
ref|ZP_00630666.1| -----
ref|YP_674884.1| -----
ref|YP_001235767.1| -----
RAAC02303 TVENGNIVLSNATQARETALKNTYQQVADEWQQGLSASQIAQVEQQQVDLPAAVLMGIGK
ref|NP_046584.1| -----

ref|YP_001371728.1| -----
ref|ZP_00630666.1| -----
ref|YP_674884.1| -----
ref|YP_001235767.1| -----
RAAC02303 MINNLNPPNAQEYYDYLA PVYTWHTYIDVTITYQTVCSKNGCHVVTHEVDTPVTMLQTAN
ref|NP_046584.1| -----

ref|YP_001371728.1| -----
ref|ZP_00630666.1| -----
ref|YP_674884.1| -----
ref|YP_001235767.1| -----
RAAC02303 TWDGTLVDITYKWVTSMSGTTQNGVYTKKIELASSKRTYDWSRVWNLFahiPTAQHTHIKE
ref|NP_046584.1| -----

ref|YP_001371728.1| -----IAPLPGLAQIAVMGRAVA-VHP
ref|ZP_00630666.1| -----
ref|YP_674884.1| -----
ref|YP_001235767.1| -----
RAAC02303 SQEDEDILAGFIGAIDYTI SDPEVQKMVTSVLFPGDVVTDLP HVSTALPSGNAIQNI LH
ref|NP_046584.1| -----GIVTSTTSSGGTFSSTGGSYS GK

ref|YP_001371728.1| HGAHIEEASQRFGIPAAWIIAVMQVESAGNMHA-----ISSAGAMGLMQVMPSTWAE LR
ref|ZP_00630666.1| YAAQVTEASQRFGIPTTWIMAVMRTE SAGDLRA-----VSSAGAMGLMQVMPDTWAG LR
ref|YP_674884.1| YAANIHQASRRFGIPAAWIRAVLRTE SAGDVHA-----ISSEGAMGLMQVMPDTWAE LR
ref|YP_001235767.1| ----IEQASARFDVPAKWIRAVMHVESGGHEYMNG-HLTVSSAGAMGLMQLEPET YQEMA
RAAC02303 YKAEIEAAARMFNI PAVLIAGVMYQESGGNQLDPTGHVLTSSAGAMGLMQVEPSTAAG LT
ref|NP_046584.1| YSSYINSAASKYNVDPALIAAVIQQESGFNAKAR-----SGVGAMGLMQLPATAKSLG
      :  *:  ::  .  *  .*:  **.  .      *.  *****:  *  *  :

ref|YP_001371728.1| IR-HALGR----DPFEPRDNILAGTAYLREMWD RYGN-VAAMLAAYNAGPGRYDEYRATA
ref|ZP_00630666.1| IR-HGLGH----DPFEPRDNILAGAAYLREMWD RYGN-VAAMLAAYNAGPARYDGYRLAD
ref|YP_674884.1| VR-YRLGL----DPFDPRDNIIAGTAYLREMFD RYGN-VGAMLAAYNAGPDRYDEYLSKG
ref|YP_001235767.1| AR-YGLGP----DPFNPLDNIMAGTAYIHQMYE IYG--SPGFLAAYNAGPGRLDYLDYR
RAAC02303 INGVFVGSNWYADLSNPTTNILLGAELLSELYHE FNENVDETL SAYNAGPGAEEEA LSQG
ref|NP_046584.1| VN-----NAYDPYQNVMGGTKYLAQOLEKF GGNVEKALAAYNAGPG-----
      .      :  :*  *:  *:  :  :  .  :.      *:  *****

ref|YP_001371728.1| RPLP--AETR AYVAALT-----
ref|ZP_00630666.1| RPLP--AETR AYVAAL-----
ref|YP_674884.1| RALP--AETR AYVAALT-----
ref|YP_001235767.1| QPLP--AQTRHYVAMIAPQ-----
RAAC02303 YTVAQNSQTLEYVAAITQQWIPALEPYFGSL
ref|NP_046584.1| -----

```

9/228

FIG. 9

```

ref|YP_793245.1|
ref|NP_253469.1|
ref|YP_386759.1|
ref|ZP_01591801.1|
ref|YP_868126.1|
RAAC02304
-----
ULATRPSEUDMNASAERUGINSAPAGBEAZHYPTHETICALPRTEINPACGPSEUDMNAS
-----

ref|YP_793245.1|
ref|NP_253469.1|
ref|YP_386759.1|
ref|ZP_01591801.1|
ref|YP_868126.1|
RAAC02304
-----
AERUGINSACGBEAZHYPTHETICALPRTEINPAGPSEUDMNASAERUGINSASCRESIG
-----

ref|YP_793245.1|
ref|NP_253469.1|
ref|YP_386759.1|
ref|ZP_01591801.1|
ref|YP_868126.1|
RAAC02304
-----
-----AVIEALATLGDLRDNPRSRHLPRIERY
NIFICANCEE- IDENTITIESPSITIVESGAPSAVIEALATLGDLRDNPRSRHLPRIERY
-----SLASLAETRDNETGAHIMRTQAY
-----AMAKLAEFRDEDTGGHLEVRKEY
-----IIQKLGLAAEYRDNETGLHITRMSHY
-----MIIVHSLTSLFESIAFVAEYGEPAADRHVKRVATY
: . : . : * : * : *

ref|YP_793245.1|
ref|NP_253469.1|
ref|YP_386759.1|
ref|ZP_01591801.1|
ref|YP_868126.1|
RAAC02304
VRLLAEHLAAQRAFADELTPAVDILLSKSALLHDIGKVAVPDRVLLNPGQLDAADTALLQ
VRLLAEHLAAQRAFADELTPAVDILLSKSALLHDIGKVAVPDRVLLNPGQLDAADTALLQ
VRVLARQLRTHVRFRRSVLTPEGIEQLCKSAPLHDIGKVGVPDAILRQGGPLTPQEMDQIK
CRLLAEDLNHRHSPYSDLTAEFIDCIQHAAPLHDIGKVAIPDHILQKPEKLTPEEFDRMK
GRILAQYVCDSSQAWC-----EMYFTALPMHDIGKIGIPDSILSKPGKLTDDERRQMQ
ARFVGEHVLRWSPAA-----CERLALAALVHDVGKVVIPREILRKPGPLMPAERQYVQ
* . . . : : : * : * : * : :

ref|YP_793245.1|
ref|NP_253469.1|
ref|YP_386759.1|
ref|ZP_01591801.1|
ref|YP_868126.1|
RAAC02304
GHTRAGR DALASAERRLGQP-SGFLRFARQIAYSHHERWDGRGFPEGLAGERIPLAARIV
GHTRAGR DALASAERRLGQP-SGFLRFARQIAYSHHERWDGRGFPEGLAGERIPLAARIV
LHTVYGRDALRHASMR LGG--SSFLQMAEELVHTHHECWDGTGYPQQLTGEQIPVAGRIM
THTVIGADNLQLVYNNYPG--NLFVGMGIEIALYHHEQWDGSGYPDGLVGKNIPLPARIM
QHASFGAQILGDDD-----DPLLSLAKEIALYHHERWDGTGYPHRLAGEQIPLSARIA
CHTVYGRDMLVELARRYSGLDGGLFELAQQVALYHHERWDGDGYPEGLCREEIPLAARVV
* : * : * . : . : . *** *** * : * : . : . : . :

ref|YP_793245.1|
ref|NP_253469.1|
ref|YP_386759.1|
ref|ZP_01591801.1|
ref|YP_868126.1|
RAAC02304
ALADRYDELTSRHAYRPPLAHAEAVLLIQAGADSEFDPRLVEAFVAVADAFAEVAQ----
ALADRYDELTSRHAYRPPLAHAEAVLLIQAGAGSEFDPRLVEAFVAVADAFAEVAR----
ALADVYDALISRRCYKKVPVPHHQAREIILKGRGTRFDPAVVDAFLRAEHDFLAIAAAA----
ALADVYDALRSRDCYRKAMSHEQARTIILEGDGRHFDPEVVMFAFLRVETIFAKTAE----
AIADVFDALTSRDPYKVAWSTQKAFDYIEENAGTQFDPKLTRAFFIECKAQVLEVQQ----
HLVDVIDARLSRPYKPGQSWHQVKQALVEGMFLDFDGMVLTGLLEVESAFLELVQAQTY
: . * * * : * : . : . : . ** . : . : .

ref|YP_793245.1|
ref|NP_253469.1|
ref|YP_386759.1|
ref|ZP_01591801.1|
ref|YP_868126.1|
RAAC02304
-----
RQLVTI

```

10/228

FIG. 10

```

ref|YP_645800.1|      ----GHGMMGGWQGP GMMGSGPGAWAFLWMLVP-VLFWTGLLVLLVWAVLRLVPPSGG
ref|YP_950098.1|      ----GTPWMMNDY----GMGGMG-LAWIFWLLLIAGVILLAVVLVKAFTRNAGGNAPTDG
RAAC02305             MSVFGVSRMMGGYG--FGMMG-----LYWFLS-VLMLLAAIVLVIWAILRMKP----
ref|ZP_01697403.1|    -----MYGGY----GMMG-----VFSLIG-MIAQLVIFVLVIFLIVAGIKK---
ref|NP_111721.1|      IGIYGVG-YRGPY---FGMMGLP---YYGMYFFMP-IMAAISILVIILFIYLIAGAF---
dbj|BAB60367.1|       IGIYGVG-YRGPY---FGMMGLP---YYGMYFFMP-IMAAISILVIILFIYLIAGAF---
                        . :      ** *                ::  ::      :*  .

ref|YP_645800.1|      NGG-RE-APEEKAEIILRRRFARGEIDAE EYEERRRLLEEHHQ-
ref|YP_950098.1|      QGPPRPGAGPARAREILEERYARGEISTE EYQERRRTLEEDSR-
RAAC02305             FGSLAQHEPYDRAVETLRERYARGEISRE EYLERLEDIQDHRS
ref|ZP_01697403.1|    VSEFRPHQREDKSLDILRERYAKGEITE EEFKKMKKDLMD----
ref|NP_111721.1|      HDGSIMSDENSRATEILKERYAKGEISE E EYRKKLDELK-----
dbj|BAB60367.1|       HDGSIMSDENSRATEILKERYAKGEISE E EYRKKLDELK-----
                        .          :: : *..*:*:* ** :      *

```


11/228

FIG. 11

```
ref|ZP_01500882.1|
RAAC02306
gb|EAU91762.1|
ref|ZP_01076306.1|
ref|YP_173223.1|
ref|ZP_01035289.1|

-----
MRGIWHFIHHWYLKIPYSPAEHLFRHVWHMIPFWMPGFIALSIVPAG---YIIMILLSV
-----HFVHHR---SPREDAVPLIFAHGWPGNFTERPGEFSIARIARVY---HKLMVTLG-
-----IWSMVPFLYVIMILVSIIQPS---FFLFVLINV
-----FWG--PQTRAENVGLTLFWAG---WWPLILLA-
--HLAHHREEYLTDPYDDPEANYLDPTVWARLPLWVRAVLRINNTLAGRMLIGPMVAQAI

ref|ZP_01500882.1|
RAAC02306
gb|EAU91762.1|
ref|ZP_01076306.1|
ref|YP_173223.1|
ref|ZP_01035289.1|

-----AARFGTTFEGIFRQAVT
LFSGFEWFKYEKALEVIKERGRKYSYGLWLAYAIWAPLEWPRAATFGWTVVGWLTIRHF
-----YKHYVAQG-----GDWGSFILRS-----VAIQYPEAVVGLHIN---
MFG-WMGITWYMRITLYKEKAREYVQAARAQGAGLWR-----IILKHILPNTMVMIVTLA
-----FPFVGRLW-----CAYCPFMIYG-----ELVQWVSLKLWPRSL
FLRADWQAQKQGNRAVLHAWLWHIPAVALVLIWLSFASMPVWAYLLAVWLGLAVLKIRTF

ref|ZP_01500882.1|
RAAC02306
gb|EAU91762.1|
ref|ZP_01076306.1|
ref|YP_173223.1|
ref|ZP_01035289.1|

YKGRNRDTAWFSIVDSEWPA-LRAGFAE-WLD-----
WKKRRRRRIAWMSTLDKADLA-LMPGWAE-WLERQQRIEAVLDRTLPWRRAKAKKASSK
FPAAPIPSPKPNITLFWLA-IGWLTPE-EKKRINRIKW-----
PFGVVANISALTALDYLGLG-LMPPTPS-WGDLQKGKSNMD-----
PWPRAAAERWGGWFLFGLFA-LILLWEELWH--LEDVAWLS-----
LEHRAHERASGRTVIEDRGPLAWIFLNNNLHVHHMHPEVP---WYRLPALYAARR
```

. . . .

12/228

FIG. 12

```

ref|YP_503850.1|      -KISKKNYDLVFKEAFSIFDNRSFLAFLGID-LPPIISFLVTEIPEVETTDMMDLNFRLE
ref|YP_517477.1|      --ISYHNNDILMKVLAEQFKNKTLDFVGIK-TAKIKDLIPSVHPAVEANETRNDIIFLLE
ref|YP_001568284.1|  -----IISVKPIDIPVINVSNQNPDFVFELE
ref|NP_783815.1|      -----DTKIIAPANTELKTIDIKNTFTDYTFYTE
ref|YP_149134.1|      --ISHHAKDILFKSLSALYQNQALDVYGLHGLPRIKALLPNEFPSVRADERRADTVFLLLE
RAAC02289             MRIARSGNDIVAKHLTNALPGEVLSVIGIH-DAHVVRLPTLPTVEVRQFTDIMLELA
                        :               :               *       :

ref|YP_503850.1|      DGSILHLEEMNLSKRD-LIRFAHYDLRLFRYYD-----APVHTVVLTPADSSSGTKVL
ref|YP_517477.1|      DDTLLHLEFQTTAGEQD-LKRFLYYDARLVRRQE-----RKVHTTVIYSGRIEQARERL
ref|YP_001568284.1|  DNSLLHLEFQTTWKKAD-LLRFAQYDIALYQKER-----RRINTVVMYSGKYESAESL
ref|NP_783815.1|      NDDYLHFEFQTTNKEED-INRFLFYDASLFYKYG-----KKVNTLVVYSSDIKSKTKV
ref|YP_149134.1|      DDSILLLEYESNERFLDNHLKYLDYACRILHTYYQQEKIRIPRIVVIYTSDVTTARERL
RAAC02289             DGRLLHLEFQTT-REPN-LYRFGTYDWAMAERYK-----RPIRTVILYTRDVTEAPSEL
                        :.  * : * : .      :   : : *      :      :. : : .      :   :

ref|YP_503850.1|      DIGSLQYNVLQIVLSDRDGDALLSQMRAALEKGEPVNE---LELIFLPLMKSKLTKIELL
ref|YP_517477.1|      ECGSILYQVENIYMKHYNQDQEYNRLKHKIDNHQLLSETDTLKLIFLPLMKSEQKEEELA
ref|YP_001568284.1|  DMGSNKYKVQQIFMIKYDGIKRYEEIKEKIEKEEELTDKDLMDLVFLPLMRNEKSEEEVT
ref|NP_783815.1|      DAGSLKYEIKAFYMSSLNGDEEYNNLTKIDKGEDLTKEEILSLTFIPLMDSKEDKSTRT
ref|YP_149134.1|      DAGDVFLSSKAVLLGEFNGDAIFHAIEEKVHNGEPLTPEETMKLILVPLMHTRFDRQTM
RAAC02289             DAGSMRYAVENVYLGHMDGDGALETVKRHLAAHEWTEEDRVRLAFAFHMRFRERTREEAF
                        : * .      . :   : *      :. :   :      :   :   .      .

ref|YP_503850.1|      RRTIDLEKELPEKDLRNKVRELTILADKIVDQKILDELWEELRMFKVVKYAAEEKGMEKG
ref|YP_517477.1|      IQAAELAKAAPDEKTKLFAIAALIVITDKIMSESNNRKLLEVLKMTQIEQWIREEGRQE-
ref|YP_001568284.1|  KDVFEIAIKIPDEDKKEAVIGSLLGFSDNYVRDEYINELKEVIRMTKIGTSLFEEGVE--
ref|NP_783815.1|      IKSIELA EKMEENNTKLQ CITLLYAFLEKFGDAKSKKKFKEVFSMTTEIGRMIVEESIEKG
ref|YP_149134.1|      EKTIELAKAIGDEPKQLHIIAGVLTATDKFIDRSYAEKVKEWIKMNKVFRLLVEELEQEK
RAAC02289             GEIVEVVQRPDPVHEQNYLAALILGFSGRVMADEQKEQLRRVLEMTDLLRELEREFEEKG
                        : :      :   :      .      .      :. . : * :.      :. :

ref|YP_503850.1|      -----LEKGLEKGIGKKGMEKGKKQERETVAKNMLSLGIEDELI IKATGLDQSIIDKL
ref|YP_517477.1|      -----GELKGRRDEKRETAQTMLNLGMSPELIAKATKLPLEEILEM
ref|YP_001568284.1|  -----EGEKELTIKILNKRFGRRLTEEIKDRIREADKKTIDYI
ref|NP_783815.1|      RAEGIKKGIEEGIKGRTEGKTEGKSEILIKQLIKKF-KKVPEEYIQIKITLSIDTIDII
ref|YP_149134.1|      -----EEMLKVMQEKQAVQRAIQEKEQAVQRVIQEKEQAVQRA
RAAC02289             -----IQKGLQOGLQOQELQKAREIAHRLLRKGASVQEVVEITGLSSKDVEEI
                        .               :       .       :

ref|YP_503850.1|      KKSL-----
ref|YP_517477.1|      -----
ref|YP_001568284.1|  GDNL-----
ref|NP_783815.1|      ALEIFDMEDIK
ref|YP_149134.1|      IQ-----
RAAC02289             RQNLH-----

```

13/228

FIG. 13

```
ref|XP_001191064.1|
RAAC02307
ref|YP_303751.1|
ref|ZP_01697132.1|
ref|ZP_01222568.1|
ref|XP_001317319.1|

-----
MLGYEVIHMHGPGVFFMHLRHMLPWWSPAIIAWYGTLMWAREYPPQRYIEDIPASDGSVEV
-----
-----EFGQFDEWK-----DLTQLDWMLLD
-----VVDMHGNEVMLPMIQVLDMWYPNNLSFLGKLTHWGHQVVKFLTEDPREANTEGGV
-----YYN-----HYEEYIEDHKSGFE

ref|XP_001191064.1|
RAAC02307
ref|YP_303751.1|
ref|ZP_01697132.1|
ref|ZP_01222568.1|
ref|XP_001317319.1|

-----GN
FSRWFGMFFWALVMFYAMWFALPYPGVEILVDLCCIPVMLIVFGMLAIGWVVHLIDFGS
-SFWIGMVFF-----LLMF-----VAY
FETHRTFFQFT-----KALNAF-----YQN
FPAIFGTVMVLLMS-----IIVTPLGVMAAIY-----LHE
VNFRWQFGSWG-----KFSSY-----PPL

ref|XP_001191064.1|
RAAC02307
ref|YP_303751.1|
ref|ZP_01697132.1|
ref|ZP_01222568.1|
ref|XP_001317319.1|

YQ-RDSSVTQMOTCLEWESLERRRE---KSRLVMLYK-----IYKGKVGLKQ
YQ-ADEHLEKGLTCAEWEQFKRRREFRDFQLRVNAMYEAIEPERRKQRFIVFEGKRGKKR
YA-SGYGLSEIMIYPSIFSFG-----LVAFGMLG-----MGPVTIAV
TR-SLWRLDHDPEGFEWIDPNNRRQ-----SIIIFMR-----RGKRK
YA-GKNGFTKLIR-VAVINLAGVPS-----IVYGVFG-----LGFFV
FALVGAGITILIDMCVLSDLR-----ILFTIAN-----FAYIL

ref|XP_001191064.1|
RAAC02307
ref|YP_303751.1|
ref|ZP_01697132.1|
ref|ZP_01222568.1|
ref|XP_001317319.1|

ED--
GDRA
DSYG
GD--
YMV-
----
```

14/228

FIG. 14

```

ref|YP_001662357.1|      -----LKLKKNKEEA-----VKEE-----PELFKEK-VPTVVEVIS-SDGYLIE
ref|YP_001666189.1|      -----LKLKKNKEEA-----VKEE-----PELFKEK-VPTVVEVIS-PDGYLIE
ref|NP_623664.1|         -----KKKKQKTEK-----QKENNF--HDPFNH-VSDIKQIIA-SDGIYVD
ref|ZP_02755290.1|         -----KKKKEKSTAKIIDF-----NKDDILNETNLLDKGIPSLKELIS-PSSFEVD
RAAC02309                MTLFARKKKKKEKKRWSWLRHLRRQKEKRVLAQDPSVIGEHAPTLHDLMTLADGIEMR
ref|YP_001181426.1|      -----EFAPDIYELIL-PDTFHEE
                           . . . : : : . .

ref|YP_001662357.1|      KDYIRVG-KKFYRLYLVT-DFPQAVYVGWLEDIKLLG-NVDVSVHLTPVDARDAVNQLTR
ref|YP_001666189.1|      KDYIRVG-KKFYRLYLVT-DFPQAVYVGWLEDIKLLG-NVDVSVHLTPVDARDAVNQLTR
ref|NP_623664.1|         RNYMQIN-DKYVRIYYFT-DYPASVNIGWLEDLKLFS-NANISIHVPETNANAVKRLTT
ref|ZP_02755290.1|         ADYIKIG-KKYVRSFIMQ-GFPAQVYVGWLDISIFNYEGETDTTIIHDPTDDKIALEEYTR
RAAC02309                ASEIYVSPSGYTKTYVVT-GLPSTVHFGYFQRFFRVGADVHISLHVEPADSAVAMAKRTK
ref|YP_001181426.1|      ADYIYLGAENYCRVYVLDIDYPSEMYVGFFDSLDDKG-EIAISVFIEPRDTGQAIRDITR
                           . : . . . : : . . * : . * : : : : : : : : : * * : *

ref|YP_001662357.1|      KITQYQSQYMLDDQRGNIYELGILRKAVEDLERLRESIQMNRDKLYYTVAVSVSADSVE
ref|YP_001666189.1|      KITQYQSQYMLDDQRGNIYELGILRKAVEDLERLRESIQMNRDKLYYTVAVSVSADSVE
ref|NP_623664.1|         KITQLQSQMLMDEEEGNIAELGILQRASESLEALRENIQMNRDKLFYVTIVIAVYGDLSLA
ref|ZP_02755290.1|         KITQFEAQLMSEQEKGNIRNITRLRDKIQELYMERSKLERNTDKLFQVQIACNLYSDSL
RAAC02309                LMTKIEAEILAEQKAGTNKQIAFLQOEYQLLEKEREELRLGRERIFYATIILAVSSPNRQ
ref|YP_001181426.1|      RINEMKSNAMLQGE----PDYKLLAQIEWYDALRASLQRGIEKILMAQVFIVVYSKSLE
                           . : . : : : : : * : * : . . . : : . : . .

ref|YP_001662357.1|      ELERSSSKILEEVLAKSIRARRMFLRQSEAFKSFVPILEN--HCEDLSRNFNSGAAASLF
ref|YP_001666189.1|      ELERSSSKILEEVLAKSIRARRMFLRQSEAFKSFVPILEN--HCEDLSRNFNSGAAASLF
ref|NP_623664.1|         ELEHNSKTLEEVLASKSIKARSAIFRQEQGFKSLVPTAEN--YLADSSRNFNVGAASISLF
ref|ZP_02755290.1|         ELNKETQKIDNKLGRKMYMMPTYLRQDDTYKTVSPYGKS--YVEDMFRNFNSGALTSCF
RAAC02309                EFEAACERIEREGFEG-FLLEAFKEHDLGFRSVAPIGEN--ALRHP-IEMTASALANSF
ref|YP_001181426.1|      QLRENCIRITKFAESLGLNMRLCSFEQAKGYISCLPLGARNLKYIEKFRIMTTSATACLF
                           . : . : : : : : * : * : . . . : : . : . .

ref|YP_001662357.1|      PFASPEFSHKKGIPLGINLFTGSPVIFDNFIGPPFMTNYNLGFATSGAGKSFLIKLLSM
ref|YP_001666189.1|      PFASPEFSHKKGIPLGINLFTGSPVIFDNFIGPPFMTNYNLGFATSGAGKSFLIKLLSM
ref|NP_623664.1|         PFASPEYSHRQGPVGINLFTGSPVIFDPFIGPPVLPNANIAVFAQAGAGKSFMMKLLAL
ref|ZP_02755290.1|         PFYNSDIRHEKGVFCGINLSTCSPILID-FYDRSKLKNSNITVIGQPGTGKTFVFSLLTM
RAAC02309                PFTNARFSHEYGVPIGVWSSGHLNRYDAWHPK--LVNANGVIIKAGAGKSFLVKGLVA
ref|YP_001181426.1|      PVGNTDITYEEGFYIGYNAITNSPVFYNQFDRS--LPNPHMAIFGTTGAGKSTTMKTMLG
                           * . . . : . * . * : : : : : * : : . . * : : :

ref|YP_001662357.1|      RGALLGIKTVFVDPDGEYSRMVQKMGGSVVKISTETENIINPFDVEEEEDENE----IRY
ref|YP_001666189.1|      RGALLGIKTVFVDPDGEYSRMVQKMGGSVVKISTETENIINPFDVEEEEDENE----IRY
ref|NP_623664.1|         RSALLGVRTVFIDPEGEYKPVVDVGVVHVKLEPNCKHIIINPFDIEVDVDEG----IEF
ref|ZP_02755290.1|         RSALRGIRTVIDPESEYGRLEALGGSRIYLATDSKECINPFDIE-EEDELD----ENF
RAAC02309                RSAAMGIRHVIVDYEGETYPMVQALGGVVIRLDEHSPYKFNPFEEEEEEKQADGTRRF
ref|YP_001181426.1|      RMAAYGYNICVLDPEREYEKIINKLGKGIYIRIRPGEKLGINPFDVEIEEEEGK-----RF
                           * * * . . : * : * : : : * : : : : * : : : : :

ref|YP_001662357.1|      VDIVQKIAEIKALVG--MIVEGVMKDKISAEELAAVEEA VRDEYE-ARGINKSPESLYEK
ref|YP_001666189.1|      VDIVQKIAEIKALVG--MIVEGVMKDKISAEELAAVEEA VRDEYE-ARGINKSPESLYEK
ref|NP_623664.1|         IDLLQKANEIKGLVS--MIEKISKTPLTAYELSLIEETILEEYR-ARGITRNPESLYEQ
ref|ZP_02755290.1|         LPTGIKLV DVKGKIADLLNLIAMAGGLSGAQTSTVAQLLQELYE-SFGITKNPKSLYVT
RAAC02309                VDVEKISKDMERLIVS-MAHLHAGHDLLDSYTRNAINDMQLQELYERDFGFTSDPESLYER
ref|YP_001181426.1|      IDLNSKVADIRFILN--IISEYVYGRRLDGVQMAIERIVRKLYY-DKGITSDPESLYTS
                           : * : . . : : : : : : : * * : . . * : : * :

ref|YP_001662357.1|      N---IADSEEFYFK--KKKEMPTLSSFVERLKKNPRGER-----IAAVMKPFLKGGTILG
ref|YP_001666189.1|      N---IADSEEFYFK--KKKEMPTLSSFVERLKKNPRGER-----IAAVMKPFLKGGTILG
ref|NP_623664.1|         S---TQEIDGKYIIGTRQKQMPPTLSSFTKRLSEKPGGEK-----LAQILKPYLKGGTIMG
ref|ZP_02755290.1|         EP-YFDKNTKEFIIHEGRKKTMPISDFYNSLSYAEKNSDIEVMDLVRLKMFKKGGVYD
RAAC02309                HETWQRDAKGDRLVRRVKRPQPRFSDFYVCKLEERAASDS--RLEELAMRLRRFREGGTEG
ref|YP_001181426.1|      AR---TDEKGR LIVGKIKELPTLSELREELFKESETQA-----LAKAMEV FVGEGLS
                           . : : * : * : * : . : : : : : : : * .

```

```
ref|YP_001662357.1| IFDGQT--NVS-LEDSYLVNFDVSEIRDR-FLRTYAMYVITNWIWEKFKVDK-QKKRV
ref|YP_001666189.1| IFDGQT--NVS-LEDSYLVNFDVSEIRDR-FLRTYAMYVITNWIWEKFKVDK-QKKRV
ref|NP_623664.1| LFDGQS--NVD-LKDALLISFDVSGIMDE-FLKTYALYVVTNWWWEKFKVKNK-QKKFV
ref|ZP_02755290.1| LFDCYTSSNISNFKNSPIVTFDVSKLEEN-ILRPIGMYVALSWTWEKFGKKDPY-INKRI
RAAC02309 MFDCYS---NVELQDVPVVSFDLSHLPEKSMARLLGMQVVLEWIMEKFIKKNVH-LKKRV
ref|YP_001181426.1| LFDCQT---TIDADDKIIAFDLKEFERDNFLKFFASVNILSWIWNKFSNAKLGQKKVV
: ** : : : * : : : . * : * : . : * :

ref|YP_001662357.1| VVDEAWMFMKY-----KDTAYFLENLSRRARKRNTSLTVASQSFIEFSNSQEGKAVLTNS
ref|YP_001666189.1| VVDEAWMFMKY-----KDTAYFLENLSRRARKRNTSLTVASQSFIEFSNSQEGKAVLTNS
ref|NP_623664.1| IVDEAWMFMKY-----EHTANFLENLARRARKRNTSLTIASQSFMEFANSQQGRAVLTA
ref|ZP_02755290.1| VCDEAWMLVSKNMAGSEYSAKFLETASRRIRKRMGGLLVASQKFTEFVESTQGGAVLTSA
RAAC02309 VIDEAQKMLEH-----LYHAIFMEDVFRRIRKRSASAVAASQDFRKFAESEQGRAILQNS
ref|YP_001181426.1| IFDEAWMFTRY-----QASAEYLEFISRKGRKYKISLMIASQTLLEFLQNDTGRAIINMC
: *** : * : * * : * : . *** : * : . * : : .

ref|YP_001662357.1| AAVFLMKQAATDIDSVQEVFHLSSQGERDFTSCGVGEALFLVG-----
ref|YP_001666189.1| AAVFLMKQAATDIDSVQEVFHLSSQGERDFTSCGVGEALFLVG-----
ref|NP_623664.1| ATVFLLRQSPTDIDAVQEVFHLSSKGEREFLSSGIGEA-----
ref|ZP_02755290.1| ETKIFLGQDITDIDSIQNMFKLSEGEKLFLLRADKGEMLVRIQGESTVVVKVLAFFDEKKL
RAAC02309 DTKILLRQDKLDKEAVIENFGLEEHEFEELIAFRDQARWWVGGEVFNQLIPFADEFEL
ref|YP_001181426.1| ASKIILMQESDMAAQIAEFFRLSERTVEIITSAAQGGIINPREKIFIQVTPFDFFEWY
: : : * : : * * : : * :

ref|YP_001662357.1| -----
ref|YP_001666189.1| -----
ref|NP_623664.1| -----
ref|ZP_02755290.1| IEKR-----
RAAC02309 FTTRFVQSDAELAMQRRWLA
ref|YP_001181426.1| VTT-----
```

16/228

FIG. 15

```
ref|ZP_00874806.1|-----KAQKQEIKPSTVNTLAY
gb|ABV55445.1|-----KTQKQEISPSTVNTLAY
ref|NP_735797.1|-----LNLLMKPKTSSKQKKEKTKREVLP TTLNTLAY
gb|AAG38042.1|AF295925_7-----KEEVL PSTANTLSY
RAAC02310MGS AVKFIVLGAALAVMLVILILRDQRKQQGQGRSRKRKVT PSLQEF LPI
ref|XP_001247966.1|-----

ref|ZP_00874806.1|QGLFQNGLMQVSPSYFSQTYLLGDVNYQTVGLDDKG AIVEKYSD LINS LD
gb|ABV55445.1|QGLFQNGLMQVSPSYFSQTYLLGDVNYQTVGLDDKG AIVEKYSD LINS LD
ref|NP_735797.1|QGLFPNGLMQVSPDYFSQSYLLGDVNYQTVGLEEKEAIVEKYSN LINS LD
gb|AAG38042.1|AF295925_7QALYQNGLMQVKEDYFSQSYLLGDVNYQTVGLEDKG AII EKYSD LIKS LD
RAAC02310EGFHESGAVIVN-GRFRRLIRVGD LNL YAMSMD EIVAVRERFKAMLMRLD
ref|XP_001247966.1|-----QVRRAVVTAFMKHAQRQLEEGVSAGEAIKQLLT---
      . : . :: ::: : * . ::

ref|ZP_00874806.1|DKTNFQLTIFNQKVNLEKFRKSILYPLQE-----DGFDAYRDELNRMM
gb|ABV55445.1|DKTNFQLTIFNQKVNLEKFRKSILYPLQE-----DGFDYRDELNRMM
ref|NP_735797.1|DKTNFQLTIFNKKVNLEQFRKSVLYPLQE-----DGFD SYREELNRMM
gb|AAG38042.1|AF295925_7DQTNFQLTIFNKRNLNLEKFRHSVLYEEKE-----DGYDSYRKELNRMM
RAAC02310NP--FQISVQARRANYTDFVAYAESTIDESVKVYDNP AFTAYAEDLKTYL
ref|XP_001247966.1|-----KYKEMLARIV
                        * . * :

ref|ZP_00874806.1|DANLEAGENNFS AVKFLSFGKSD--QTPKLAFRSLSQIG EYFKSGFSEID
gb|ABV55445.1|DANLEAGENNFS AVKFLSFGKSD--QTPKLAFRSLSQIG EYFKSGFSEID
ref|NP_735797.1|DSNLEAGENNFS AVKLISFGKSD--QAPKLAYRSLSQIG EYFKSGLSEIE
gb|AAG38042.1|AF295925_7NQNLDSGENNFS AVKLISFGRKD--SNPKQAYRSLSQIG EYFKSGFSEID
RAAC02310REEAMKPRTDREN LIVIGVLPKVGG EDEKLQLERLAREQS FVESGSLSAMG
ref|XP_001247966.1|FDRETDEKPDQVDLLLL-----FQQDLVERSKGGTILLFVAKELYDLE
      . . : : : . . : : :

ref|ZP_00874806.1|VSLGLLGGEERVNV LADMLRGENHLPFSYKD-----
gb|ABV55445.1|VSLGLLGGEERVNV LADMLRGENHLPFSYKD-----
ref|NP_735797.1|ASFTLLTGEE RVNNLADMLRGENHLPFTYKDLV-----
gb|AAG38042.1|AF295925_7ARFESLAGEERVNLLADMLRGEHHL PFSYRDLTR-----
RAAC02310LPYEVLEPVQVVEAVQNFWNRERAVSQRYRDAVRRRTHAPRVDGLDVEVS
ref|XP_001247966.1|IVED-----EAFEQWWDNERSTATEALKKVRQQT-QPFIDWMMTEES
      : . : * . . .

ref|ZP_00874806.1|-----
gb|ABV55445.1|-----
ref|NP_735797.1|-----
gb|AAG38042.1|AF295925_7-----
RAAC02310DLVRTQEEEGKEA
ref|XP_001247966.1|DEEESDEESGEE-
```

17/228

FIG. 16

```
ref|YP_001205829.1| -----LPIVEGVTTVFGVRTPRNEYFTAVICWVA--IGLLWFLVNRTRRGKAVMAAS
ref|YP_001240062.1| -----LPIVEGVTTVFGVRTPRNEYFTAVICWVA--IGLLWFLVNRTRRGKSVMAAS
RAAC02311      -MLEVHSPIPMTEGETLVFGIRTKHIAHFSIGVCLTAPMAALAWFLLPLLHEPKLLVLFL
ref|YP_359336.1| -VLPFYVSIPVLESRLRLFFGGIGALAGFLFPVISGGL--ISAP-VSVPLWHTPLLPVLFL
ref|YP_001451893.1| -----ILIGVLGYALS FAMGVVDTAPIAEAHWFALPTFYTPR----FE
ref|YP_466026.1|  APILFSGGRPLAADLSDAVGRFG--LAPLVAAAGAAA--LRARWRAAPE-RRPEVLLLAT
                                     . *      :

ref|YP_001205829.1| MNPRGVTLGIELSS-----IYLTWV-----
ref|YP_001240062.1| MNPRGVTLGIELSS-----IYLTWV-----
RAAC02311      WGGAGVVFAKVPVAN-----RPLAEWLWLSFRYARRPKLILYDREFRIRVHRKRAAERWG
ref|YP_359336.1|  INSFGVAFAAIPALF-----APGSEILAFIFNLA-----
ref|YP_001451893.1| WFA---ILTILPAAL-----VVIAE-----
ref|YP_466026.1|  WGGLAFALALARNRFDVYAAAPLAVWCWLGLRHVQASRL-----
                                     :      :

ref|YP_001205829.1| --
ref|YP_001240062.1| --
RAAC02311      AR
ref|YP_359336.1|  --
ref|YP_001451893.1| --
ref|YP_466026.1|  --
```

18/228

FIG. 17

```

ref|ZP_02850831.1|      -----LSGFLADVSY--LSFVLNG-----
RAAC02312              MSMKRWITQVGKSLYAGSVVLASTAMSAEASTSGITNPLSSSSSS--VSSVLNQAL----
ref|NP_279998.1|        -----ATIGGIATTILGFSSQT--MAAILGGVFGLO-
ref|XP_001637270.1|     -----VGVLGTALN---
ref|YP_001506532.1|     -----QLGSSLLLNTIVLVLTATAGAASTPLAVARRALARPSLIALVLALALRPVG
ref|ZP_02248080.1|     -----GKLETIGSAMVAQNAAAAAPTPTTGVIPAAADEISV--LQAPLFTAYG---
                                   :      *

ref|ZP_02850831.1|     ----VTSGVFAL---FFLFY-LQR--VGTGSGDHSaelMRKKQAAGGASLWLLVTDVRLY
RAAC02312              --GWLDDAIFAVAGALFLFH-LYKAIIGLMAGSHHAQ---RREEAKSHLVVVAISGVMLG
ref|NP_279998.1|        -SAWLTYGIFAAIGAVFLVD-VWL---GLRSG-----IRNAARATMLLIGVAMAVL-
ref|XP_001637270.1|     ----LTDWIFAVSGGMFLFVSLAETMPELFASA--QREESRGKVLLAQNAGLLAG
ref|YP_001506532.1|     FPAWLDEALASIGATLTPLA-LFS--VGLQLRLPAVR-RWRRELALGLVVKLAVAPVVVL
ref|ZP_02248080.1|     ---TLYQQVSAEAAAVYDLF-VKT--LGVSAAGTYA----ATEAANSSAAALPLSGI---
                                   :      :      :      :      :      :      :      :

ref|ZP_02850831.1|     LFMALGGMLINLGNS-
RAAC02312              GGAILAGALYNLGKSF
ref|NP_279998.1|        ---VVVG-----
ref|XP_001637270.1|     -----FGTMLTLAAF-
ref|YP_001506532.1|     GCYALAGAL-----
ref|ZP_02248080.1|     -----

```


19/228

FIG. 18

```
ref|YP_245669.1|
RAAC02313
ref|ZP_02369868.1|
ref|YP_438666.1|
ref|NP_695275.1|
emb|CAB06069.2|

-----
MRQSSNAPDSVRGESWPKVGAICSFPSCTLVPSRPVQARGGEGMAGSTASNVTQTITQLG
--RGALAAARIGAAFDLPTLGRTLHVVCAPAIALYG-ATGDARWLLASIAIAIAVERVG
--RGALAAARIGAAFDLPTLGRTLHVVCAPTIALYG-ATGDARWLLASIAIAIAVERVG
-----VSVLG
-----

ref|YP_245669.1|
RAAC02313
ref|ZP_02369868.1|
ref|YP_438666.1|
ref|NP_695275.1|
emb|CAB06069.2|

-----AVLVLFGGFGTAIDVDGADQIIILGWGAVFSSILGIIGSVVVKSKSG--
FAP---VNPWALLNRVTSGFGHLIEVFAWVYFLSEIGVLAGALVWLVGSISSHARIK--
IAP---LGALAQGAAIIAGFLSLSCALNAWPAFVAGCAALAAAVALSRFGARLRSLGNF
IAP---LGALAQGAAIIAGFLSLSCALNAWPAFVAGCAALAAAVALSRFGARLRSLGNF
FVPHMFANPWGSLMAVLVALFLSVPAAFIYRKIRTRKGAAIGILVGAVLAIIVVALVSN--
SLP----STWR-IPSTCSEHSEGSDCSTNSNVTMSLLKLLDCFFRPFLPREGLKKIPS--

ref|YP_245669.1|
RAAC02313
ref|ZP_02369868.1|
ref|YP_438666.1|
ref|NP_695275.1|
emb|CAB06069.2|

--IGGAFMLIAAIIIGFLCI-----FAIYILPGALLAI-----
--RTGAHITLYAIVGFL-----AAVILPGVILAIIDSTFQR
VFIPSLYLTCEAAAAHLSAPRLVPYLCAAMLPPIALSCIATLR-
VFIPSLYLTCEAAAAHLSAPRLVPYLCAAMLPPIALSCIATLR-
----LIVTPIYAHMTYQAV-----AAMILP-ILLP-----
-----EKLSHIPELI-----EPWFIFSLVWSVGATGDH
:: .
```

20/228

FIG. 19

```
ref|ZP_01058859.1| -----IGEDILPENVD FEVIDGFTKVTDKDAAIYTRRL
RAAC02314      MERVEYVDRVKHVWSEYRTNDEELAFALGIDEEAESVDVSTKDGITNVT-----VFTTEA
ref|YP_754274.1| -----LSPNRWQHSLQVAET-ALQMAG---QYELELEQ
ref|XP_503244.1| -----NLTANVEIDGKQYNTFTTEPPKALAGER--AKVKFPIKD
ref|NP_218466.1| -----DLIVRLAKLKHVNPQVLQVMS----DSLHKKF
ref|NP_126488.1| -----ARKILEVLIDSDVNVYINDVVASEVIFIYLLKLTGKSYLTLKK
                                     .       :       .

ref|ZP_01058859.1| AKEEGFFLGNSAGAAIKGLLQLKQHFK-----
RAAC02314      VYHFGTFRADYVGHASRALVQLLQHFRVNLPIDFIVAHQTFHVYLKGGKVVAGEKEYPIA
ref|YP_754274.1| VYLTALLHDYAKGLSGQELLRLAEENNL---IEDEVDPDPLLH-----
ref|XP_503244.1| MTEFLHGGEENVMTIERIMTELERDPVLNVSGDYDMPKEQLRETAVARIAALS GHWKKDT
ref|NP_218466.1| AALHLSQRRDLGHA--VLAAILKKMER-----ATEHSILHALAEKN-----
ref|NP_126488.1| NPVIVRSVDKTSVYELLGMFKFLETNEFVFSIAKRLIDK-YGLLPNDALILATAIFY---
                                     :   :   .

ref|ZP_01058859.1| -----
RAAC02314      ERSEGYELVEPVEWMIASSVLDVVLRLAE EYEVTP EQIVESAVGTFYALLSVAEENDVEP
ref|YP_754274.1| -----
ref|XP_503244.1| EKEALL-----
ref|NP_218466.1| -----
ref|NP_126488.1| -----

ref|ZP_01058859.1| -----
RAAC02314      ETVISMLTELMHRHAEHTNEHELSTLNR
ref|YP_754274.1| -----
ref|XP_503244.1| -----
ref|NP_218466.1| -----
ref|NP_126488.1| -----
```

FIG. 20

```

ref|ZP_02854145.1|    ---MGTTISFGIQKGGVGKTTTTTAITSYILSKE-HKTLAVDFDSQGNLTRFLTQQNIYNF
RAAC02315             MERVGCTISVGLQKGGVGKSTTTALTSYILAEQGHRVLAVDFDSQGNLTQLLTQSPYDF
ref|YP_145847.1|      -----TITMGIQKGGCGKSTTTGVLAYLLSRDGYRVLAVDMDSQGNLTTELLSRKPSNEF
ref|YP_799230.1|      ---MGKIVSISNQKGGVGKTTTTINLAANLASIGKKVLIIDMDPQGNSSGSLGIEINTLV
ref|NP_714527.1|      --QMKGIVSISNQKGGVGKTTTTINLAANLASIEKKVLIIDMDPQGNSSGSLGLEIHKTN
ref|YP_536482.1|      ---MGTVIAIANQKGGVGKTTTTSVNLGACLARAGQKVLLIDTDAQGNATSGIGVRKHNI
                        :... **** *:***: . *: :.* :* *.*** : .

ref|ZP_02854145.1|    TEKTVLEAVKAKDPRPYIYKISDNLHILPAEDFLATFSRFLYQ-----EYQG-NKA
RAAC02315             VHRTSLEACKERDPRPYIHAI SDNLHLLPAEDFLSQFDKWIYT-----EVHVSQQM
ref|YP_145847.1|      TEKSVLEAMQERDPEPYIVKVNDRDLDPANNFLATFPRWIYTGETYLGKYIRYKG-KPT
ref|YP_799230.1|      KTSYELLGESSSTNECIQRTNVSNLHIIPSNINLSGAEADLLVED-----QRE
ref|NP_714527.1|      KTSYELLGGEASVNECIQRTNVSNLHIIPSNINLSGAEADLLAED-----QRE
ref|YP_536482.1|      NDVYDVIVSELPIREAIMPTYIDNLDVVPATIQLAGAEIELTAQM-----ARE
                        : : . ..*::*: *: :

ref|ZP_02854145.1|    LLLKETLDVVREQYNYITIDLPPHLDGQTINGLSASDYAVVLLQSEPFALDALDRYLEVL
RAAC02315             VILKNTLDVVKSDYDYILIDLPPNLGGLTLNGVCASDYCVVVCQSEPFAYDALDRYMEII
ref|YP_145847.1|      LILDDTLDKIRHRYDFIVIDTPPSLSEQTINALCASQVIMMFECSNWCYSAPVNFMESV
ref|YP_799230.1|      YRLKNAVSELRSEYDYILIDCPPSLGILTINALCAADSVMITLQTEYFALEGLTQLMKII
ref|NP_714527.1|      YRLKNAISDLRTEYDYILIDCPPSLGILTINALCAADSVMITLQTEYFALEGLTQLMKII
ref|YP_536482.1|      KKLYDAVQDVKEEYDFILIDCPPSLGILLTINAFTASDSILIPVQSEYYALEGLSQLMNTI
                        * :... :: *::* ** ** * . * *.. *:: :: : . . . . . : :

ref|ZP_02854145.1|    MGVOQKANSNMKLIGILSTMLDSRAAIDSEIINKARKDYEDVVFETIIRRRNRIKEFSLF
RAAC02315             QAAQQRVNPNLRIAGILISLLDARTAGNYITERIREEYQDFVFDTVIRRKSRRIIEFSVE
ref|YP_145847.1|      EGARVHGRHNTRLLGILRTMNDVRRNDKAFNEMIEEDYPNEVFKTIITRKAPIGRSLY
ref|YP_799230.1|      SLVQNQLNPSLELEGVLLTMFDKRTNLANQVAEDVKS YFKDKVYTTIIPRNVKLSE----
ref|NP_714527.1|      SLVQNQLNPSLELEGVLLTMFDKRTNLANQVAEDVKS YFKDKVYTTIIPRNVKLSE----
ref|YP_536482.1|      QLVQKHFNPDQLIEGVLMTMLDARTNLGNQVVEEVKKFFKEKVYKTVIPRNVRLSEAPSH
                        .: : . . . : *:* :: * * . . : .. : : *:*:* * . : .

ref|ZP_02854145.1|    GIEERT-KVDREALKYYKNFTKELISRVQK
RAAC02315             GIKIQT-KADREAIAMYESFVEELKARVSR
ref|YP_145847.1|      GFEENN-ELN-QALEQYENFYKEMMERV--
ref|YP_799230.1|      -----
ref|NP_714527.1|      -----
ref|YP_536482.1|      GMSIIDYDPRSRGAEEYEALAKEV-----

```

22/228

FIG. 21

```
ref|XP_624126.2|-----LSQPHTTIP-----A
RAAC02316-----MSAADKKKRFESLMPSTRTPPHKSEVIDTLMGQTYDADD
ref|ZP_01893908.1|-----PFMEEDTTP-----
ref|XP_001444409.1|-----LGISKTATE-----
ref|ZP_02180762.1|-----TYDQEGAVY-----
ref|ZP_02077766.1|LGSTQKEQKYEMEDYTESVQKQAEADELAKSLNDAEEAAKNSYADKIGDIGTVEKYLQK
.

ref|XP_624126.2|LHPPNVPLVSS-----SHTTSPAVASLGVTTPAAPSNGATTTTSTP-PSIPMAAYP-----Q
RAAC02316IGTPKTRMESLHKEVSHTSEVVYPMSNTTTTDDNGTTDNSFD-IAARIRAFEE--QRR
ref|ZP_01893908.1|-DTSSTSDSS-----DDDDDDDDSEEETESGPDPEETRLRFELLKEKLDAA-----A
ref|XP_001444409.1|---EQIKEAYA-----QKAQKLYPNVTTTSVAIND-VQAQQGFQDVAEFAVLS---QIQ
ref|ZP_02180762.1|-----THVAKD-----HQKKAMKFIQDQLFTTPTWMLDEDIFN----KIESAG-----S
ref|ZP_02077766.1|LREMSGETGYVDN---VNLAKMYVEEINSLLPETVQLTADGRVEWLKNTDAIYQEIAALK
. . .

ref|XP_624126.2|TQPSYPPLYTPYTAL-----
RAAC02316SRPSYDELYTKHTFLIRKDLMERLERASARHSRGFKTRFINELLEAGLTQLEIED
ref|ZP_01893908.1|SLAKHGRGHKKTQEALN-EL-----
ref|XP_001444409.1|SRNAYDLLNKEQPELLYGEEEMERYKQSFQRNDDG-----
ref|ZP_02180762.1|IERVRGMQTRTLNNMLDFGRMQRMLEN-----
ref|ZP_02077766.1|EEARVKAYQELYTQAVRDEIKARTDLTMTTEELNKKVQRLNELKADGISEDEIEE
```

FIG. 22

emb CAJ50746.1	-----AADLAQDTFVRVLRHRRELPGV
gb AAM28266.1	-----AADLAQDTFVRVLRHRRELPGV
RAAC02290	MTSGMRRFTNSSRTAEVETGMEGHYTYALMGPTSEHTVACDLAPDTFVEVLRGAMEAIGV
ref XP_816394.1	--RGAN-LSNEERTNETIVR-----QYETALENIIGTLQDRLTILRE
ref XP_001585185.1	-YIEMSAFGRESHPRVQALR-----HDRDRSPDPNATDPEVIEHLKT
gb EAU92316.1	---THSAFGPTVSFEE-----AQRIERETA EHLLKSRKLSLIV
emb CAJ50746.1	RE----PRAYLVTIAGRLLLNHYRRRS---LERAYEEALAVMPLESAPSPEQRLL-----
gb AAM28266.1	RE----PRAYLVTIAGRLLLNHYRRRS---LERAYEEALAVMPLESAPSPEQRLL-----
RAAC02290	RLDELVLRIYVQEGAGGWQLVHGRVKD---AREAQQTCLLPMPPE-PPSGDPEVLDDATY
ref XP_816394.1	EYAS-FLKVSEETANEKEVLNEHEQDSQVRVQTFRTAALTQRVFWSTREKEVPEVISE
ref XP_001585185.1	QLKDADDLMDLIVARYRVLENENGEPTERVEVAERLEECMDLVDGPLGQKEAFDN---
gb EAU92316.1	DLDTIVHATVDPPTVGEWIAEGEAWEAR--QERRNKVKTTTPDSDSDSDSDSDDEDEDE
emb CAJ50746.1	-----
gb AAM28266.1	-----
RAAC02290	QDIGYAALNLYRVLWQSRQRIKASSLTKVVRVQSLRALDERMRLVEQAVEGTEWWDLFR
ref XP_816394.1	LDN-----AAAAAEKSRAAVKVTLALLDSRLQSFSEQKE-----WM
ref XP_001585185.1	-----LKALLSESREALKVKMNEWNEKDKERDWAVES-----LE
gb EAU92316.1	CNPNWEALKDVKKFTLGPESFNAPSVKGRSGKGRMVEQEGCMYYIKPRPG-----W
emb CAJ50746.1	-----
gb AAM28266.1	-----
RAAC02290	LRYQQGKKLVACAMELHMSSRTLNRQISDMAGHVGRMLSTVLKERELAEIMQAAKRVPM
ref XP_816394.1	ETHQQSLEDLLVSLR-QLNKRAFERQLSLMNERVARPSST-----
ref XP_001585185.1	KLLKESEEREALAGLLHATQEDMRREGKRAAKVAEIEEK-----
gb EAU92316.1	KEFLENAAKKYEMHVTMTGTRAYAQEVCAAIDPDGKLFGRLLSRDESGSLTQKSLQRLF
emb CAJ50746.1	-----
gb AAM28266.1	-----
RAAC02290	PGKATARVIRLEQRELWS
ref XP_816394.1	-----
ref XP_001585185.1	-----
gb EAU92316.1	PCDTSMVVIIDDRADVW-

24/228

FIG. 23

```
ref|YP_001210712.1|
ref|ZP_01593342.1|
RAAC02317
ref|ZP_02727046.1|
ref|ZP_02758954.1|
ref|ZP_00235902.1|

-----
MWFKYTQKQTTQPCAGFSSLRTMGRFPSRRKVFFRPPLQGGTNMAYTPAKVRKGWSENW
-----KEW
-----W
-----F

ref|YP_001210712.1|
ref|ZP_01593342.1|
RAAC02317
ref|ZP_02727046.1|
ref|ZP_02758954.1|
ref|ZP_00235902.1|

----DAIFERHDLTTLQKLVIYIFCRRANNNG-QSTPSYDDIAKDCGCHRSSAIEAVNTL
---PNAISRSTELSSSTAKLVFGKLCQYAGQNG-QAYPSYKTLAGDVGVERRQAIRAVKEL
FRCPNSIYDIDYISGYAKAVYIFLCRCADGEG-QSFPGYNIAKAIGFGRTRTIEAIKEL
FWLENDLVDREDLGIYEKMIYIVLARYSDNES-CCFPSYKTIALKCGCSEKQAKSVVKIL
FWLENDLVDREDLGIYEKMIYIVLARYSDNES-CCFPSYKTIALKCGCSEKQAKSVVKIL
FMIDNEVIDNGELDVYAFKTYAVIVRYANKKTKSAFPSINTLAKRVGCGKKKVICIKIL
: : : : : : : * . . : * * . . : : *

ref|YP_001210712.1|
ref|ZP_01593342.1|
RAAC02317
ref|ZP_02727046.1|
ref|ZP_02758954.1|
ref|ZP_00235902.1|

GKMG-LVVKHRRKRHNGSDTSNMYVVFPP-----
VDYG-LIKPVGRKKGDGGFTSNIYAFLWHQTFSDGLTDPG----DKNDTRGGVTNVTFP
EQAGLLVKEHRRNEETGEYYSNLYTVIHPDDVAFRGGMSPHGIPMSRDDTPLSPDDIPMS
ENKGLIKKENRIKSNSNEKESNIYFVL-----
ENKGLIKKENRIKSNSNEKESNIYFVL-----
VEKGYVSKTLR-KDNKGDHLSNLYHLLPTSNISSKKQG-----
. * : : . **:* .:

ref|YP_001210712.1|
ref|ZP_01593342.1|
RAAC02317
ref|ZP_02727046.1|
ref|ZP_02758954.1|
ref|ZP_00235902.1|

QCHKRHHLS-----DMSP---KENHTRESNSETTTTEEIRLLLSGTPLSKISEKEVRI
PHDRRVSPGGREVRSIDQDPPTKIYIHTRSQAAATLATSNEHLTSNTALEEVCETVNEA
-----

ref|YP_001210712.1|
ref|ZP_01593342.1|
RAAC02317
ref|ZP_02727046.1|
ref|ZP_02758954.1|
ref|ZP_00235902.1|

LIKR-----HGNERVMAADIAAEKWRERKEIKNPGGYLQTFVNLVV-----
LVQLGIRAQKKTLHLWLQIATPEEIEAASLATREGVKSPAGYIGTVLRNGLVRVEKSTA
-----

ref|YP_001210712.1|
ref|ZP_01593342.1|
RAAC02317
ref|ZP_02727046.1|
ref|ZP_02758954.1|
ref|ZP_00235902.1|

-----
-----
KHKDPYNEFYKLFP
-----
-----
```

25/228

FIG. 24

```

ref|ZP_02171383.1|      -----VYLARGSTDLRKSIDGLAAIVQEGFELDPFSSSLFVFCNRYRDKIKILY
RAAC02319              MLAFDWTSDHRVYLACGATDMRKSIDGLAALVQASFQLDPFSPCLFVFCNQRDKLKIHL
ref|YP_431168.1|      -----RVYLACGATDLRKSIDGLAVLVKEGFELDPFSSCLFVFCNRRNRDKLKIHL
ref|YP_001212944.1|    -----RVYLAIGATDLRKSIDGLAVLVKEGFELDPFSSCLFVFCNRKCDKLKIHL
ref|YP_754944.1|      -----SNRQVYLACGSTDLRKSIDGLAVLVKEAFELNPFSPCLFVFCNRQRNKLKIQL
ref|YP_754864.1|      -----SNRLVYLACGSTDLRKSIDGLAVLVKESFHLDPFSPCLFVFCNRKRDKLKIQL
                        **** *:***:*****.:*:.*:***.***** *:***

ref|ZP_02171383.1|      WDHNGFWLYYRRLEKGRFPWPTSGSDEPMIITERQLRWLLDGLPLDQKAHR-----
RAAC02319              WSHNGFWLYYRRLEGRFDWPETGDAKTMVITRRELNWLLDGLPLEQPRAHRAVYVRSAL
ref|YP_431168.1|      WEHNGFWLYYRRLEKGRFVWPQDTTSSTITITRRELRWLLDGLPLKQPQAHPEVKARTIL
ref|YP_001212944.1|    WDHNGFWLHYRRLEKGRFHWPADAGSPTLVISRRELRWLLDGLSLKQPKAHPEVKARTIL
ref|YP_754944.1|      WDHNGFWLYYRRLEGRGKFEWPA-ADSQVVSISYREFRWLLDGLSLKQKQAHKAVKERTII
ref|YP_754864.1|      WEHNGFWLHYRRLEGRGKFDWPT-AHTDVVSISYREFRWLLDGLSLKQNQAHKAVKQR---
                        *.*****:*****:*:* **      : *: *: :.*****.*.* **

```

FIG. 25

ref ZP_01287154.1	-----RSKRRFWAAHVAAWEKSGLTQTAYCREHG-LSR
ref ZP_01287577.1	-----RSKRRFWAAHVAAWEKSGLTQTAYCREHG-LSR
ref YP_754945.1	-----MSSNERKAWWEERLAEHEASGQRTAWCEENS-ITP
ref YP_754863.1	-----MVNNDLRSLWEQRLADYETSGKSIATWCREHS-IRN
ref YP_431169.1	-----MTKAELQELWASRIAIEYKMSGQSVKEWCATHEGVSP
RAAC02320	MAFILPSVENVVIKPREGSCMREHMSHQRRELWRERIAAFYDSGQSASQFCAEHG-LKP
	. : * ::* ** :* : :
ref ZP_01287154.1	HAFGWWRKFRDQPASEQPLLQVPTVARPTAAT-AAADFSGLRLLLPRG-LQLEINQGF
ref ZP_01287577.1	HAFGWWRKFRDQPASEQPLLQVPTVTRPTAAT-AAADFSGLRLLLPRG-LQLEINQGF
ref YP_754945.1	RQFYWRRKLR-----TEHVEKEQPVKWLKYESRQLGIAGDAIAVHVGQATVEIRKGF
ref YP_754863.1	NQFYWRRKLR-----MDQVENNQPVKWLPLEVE--QANLAPGSIGVHVGQATVEIKPGF
ref YP_431169.1	RQLWYWLRYKN----QNGVLSAQSTRWLPVEIS-EQTSNVSNLLVRIGPAIIEVSPSF
RAAC02320	HQFWYWLRLRN----ETAPGTQATSFVSVVTTSSPSDASRSPLTLRVGSVEIDVHPGY
	. : :* :: : : . . . : : * ::: . :
ref ZP_01287154.1	DAATLGRVI-----
ref ZP_01287577.1	DAATLGRVI-----
ref YP_754945.1	DRELFCEIIQVLQTI-
ref YP_754863.1	DPHLLRQIVKVLQTI-
ref YP_431169.1	DPVLLSQVVKVLVALC
RAAC02320	DASTLAELIRLVMHVC
	* : .::

27/228

FIG. 26

```

ref|NP_982177.1|      TEINBACILLUSCEREUSSCRESIGNIFICANCEE--IDENTITIESPSITIVESGAPSQY
ref|ZP_02595431.1|  -----QY
RAAC02321             -----MNEKPPQA
ref|YP_001108426.1|  -----
emb|CAN89659.1|      -----
ref|NP_927486.1|      -----KKPSGS

ref|NP_982177.1|      VLVISPRDNE--VDLVKELGYGVILIRKNISFDEMFSVEVPVEIDLNDDEIVVGKCKELS
ref|ZP_02595431.1|  VLVISPRDNE--VDLVKELGYGVILIRKNISFDEMFSVEVPVEIDLNDDEIVVGKCKELS
RAAC02321             VVIGTREHEGEVELVRALGYRVLLNTKINIEDALVADVPVELDLNDETLVISKVIDLT
ref|YP_001108426.1|  -----ETNDVAALAEHVARIQ
emb|CAN89659.1|      -----IRCDTNSDAALRAALQDRF
ref|NP_927486.1|      IELKKAKELGAYIIFIGSR----KYYNKVSDNDLLYIDEFFEADTNDELVINMVEYIN
                        : * . :

ref|NP_982177.1|      EKYNVGVYTLNEYRIPLAAKVGEILEINSFLSYETAITCRNKKLARKKLNNNSNVSSVKF
ref|ZP_02595431.1|  EKYNVGVYTLNEYRIPLAAKVGEILEINSFLSYETAIICRNKKLARKKLNDNSNVSSVKF
RAAC02321             NHYYIKAVFTLNEYRVPLSARIAETLGLQRTISYEAQNCRNKKLTRRTLMRNGIQAMKF
ref|YP_001108426.1|  PVLGFDGVLTSDDYILPAVAHIAARLGLPGAP-PEAVERACSKDLTRRALREAGVPGPAF
emb|CAN89659.1|      RREEIAGVTTTSDFYVPAARLARWLGLPGNP-PEAVTACRDKSALRALLRRAGVHQPQY
ref|NP_927486.1|      AKKKIQGVITFMEFYVPLVAKVAETLGLKGIT-YENALRARNKHLMRESFRQKNVPIPKY
                        . * * : : * * : . * : : :

ref|NP_982177.1|      ILIRKLDNLNEKLEGFSFPVIVKPSNDSGSKNVYLCKDYNEVKQAVDVISNSKLNLVGQT
ref|ZP_02595431.1|  VLIRELDNLNEKLEDFSPVIVKPSNDSGSKNVYLCKDYNEVKQAVDVISHSKLNLVGQT
RAAC02321             ATVKTPAEALDVIADMDLPVVVKPANDAGSHLVYRCDTLQEVWEAVEAIAQTYPYNWVGQN
ref|YP_001108426.1|  ALAESAGLEEAALGYPLVLKPVLDLCAGMFVRKVTDDVELRDAYAAALRQFPVNARGQR
emb|CAN89659.1|      AVVREPGEVAAAVARTGLPCVVKPADDSGSVNVLLCTDEAQAQAERILAVTTNVRGMP
ref|NP_927486.1|      ALISNIDSAKIEAAKIGYPNIIKPINMAGSRGVLNRNDNVIERNFREVEITP-PFGVK
                        . * : : * : . * : : *

ref|NP_982177.1|      LDPEAIVEEYLDGPEYSIESYITIDGKTTIVGVTEKVVTFPFLSVEVGHNFPAVLEEDIET
ref|ZP_02595431.1|  LDPEAIVEEYLDGPEYSIESYITIDGKTTIVGVTEKVVTFPFLSVEVGHNFPAVLEEDIET
RAAC02321             RDEILIEEYLVGKEYSVEACTIQGETHILAITEKETTTNVSVEIGHTVPAILEEQVS
ref|YP_001108426.1|  RNPLVLEELLTGPEVSVETVTSRGRTSVIGVTDKSIAGAPWFVETGHMFPAALDLDER
emb|CAN89659.1|      TARTVLVEEYLDGPEYSVEMFSQDGEAVCVGITAKSVTGDPHFVEHRHLFPAPLPAATAE
ref|NP_927486.1|      KSSLFLIEEYLEGQEYSIESISFNGVNVNIVTITKKYVSSNGYFVELGHTLPANIPLOQKI
                        : : * * * * * : * . : : * * : * * * . * :

ref|NP_982177.1|      SIHETVINALDVIGVDFGVTHTEVKVTSDBGPKVIEVNARPGGDRITDLVEYVTGIDLRLRI
ref|ZP_02595431.1|  SIHETVINALDVIGVDFGVTHTEVKVTSDBGPKVIEVNARPGGDRITDLVEYVTGIDLRLRI
RAAC02321             SIHDLVKKALRALGVDCVTHTEIKLDGNSLRIVEVNARPGDKIPLLVRAVTGYDLREL
ref|YP_001108426.1|  AVVETAVAAVEALGLDNVVGHTLKLTPDGPRIVEVNPRPAGNQITELVRRVTGIDLAAA
emb|CAN89659.1|      RITETVMAALDAAGIRLGATHTVKLTGSGPAVIEINPRPAGGMIPELVRLASGVDLLEQ
ref|NP_927486.1|      EIEKVVTQALFALGIYNGGGHTEVKVTSDBGVKVIEVGARLGGDHIPELVEMATGIDMWKA
                        : . . * : . * : * : * : . . : : * . * . * . : * :

ref|NP_982177.1|      ALRINLGLPIKNSSCNKELVPSSSIRFLIADKEGYISFNEN----FRAESIKEIHWYVVK
ref|ZP_02595431.1|  ALRINLGLPIKNSSCNKELVSSSSIRFLIADKEGYISFNEN----FRAESIKEIHWYVVK
RAAC02321             ALHIALGGQFENAPRHEVLAPVAAIRFLIADKHGVVSYDSRGV--LEFEEMKYIEFFTKN
ref|YP_001108426.1|  YAQVALG-EEPALERAETGAGSAAISFLPPKEGVIGEVVGVDLDGADGVVDWKFKSS-
emb|CAN89659.1|      QLKAATG-QAPDLKPGHG--AHAGIQFLLADADGVLDVADGVARARAVDGVAVTVTVAP
ref|NP_927486.1|      VIQVSLN---ISPDLKKKFSKYAAISYITAPS-GIVKKVNYIKN-----GFIHFDVNV
                        : . . . : * : . * :

ref|NP_982177.1|      GERVNKTTSNFDRIGYYIVDGNKKESSKKIADSLNECFELAISEL
ref|ZP_02595431.1|  GERVNKTTSNFDRIGYYIVDGNKKESSKKIADSLNGCFELAISEL
RAAC02321             GYVVKTTSNYNRLGYFVVFGKNRDHVTNVRKILSKLNLRI
ref|YP_001108426.1|  GHRSGPATSNNNYLGHVMVTAPDPGRARAKAEALVGGDLVRYAE-
emb|CAN89659.1|      GAVVRRARSAGDRVGHVIACRPGPEQVTAALDEARDLLRLTVGE-
ref|NP_927486.1|      GEHIESLKNSSQRLEYAIACGVTAEQAEFESHHLKENIIIEIA--
                        * . : : : : . : : .

```

28/228

FIG. 27

```

ref|YP_504284.1|      -----SLYVAVFSTMLGIGIVIPLLPRFAETLGASGFGIGMIFSSFALSRAFAMPFFGRY
ref|ZP_02131576.1|    -KIFITLFFSIFTSVTGVGIVVPLLPVYANNLGASGFMVGMIFGSFSLSRILFLPYFGKR
ref|YP_001046337.1|    -RIYNVLFISVFATMLGLGIVSPLLPIYAENLGATGIWLGIIFSAFALSRSVFMPVIGRI
ref|ZP_02132246.1|    --ILVLLSVSVGVSMIGLGIWPLIPVYAVELGAGGFMVGLIIASFNLSKAAFGPFMGRF
ref|ZP_01288161.1|    -RLLLPLLFSVFIALLGIGIIVPVMPIFATSLGASGLALGFIIASFSTRGLCQPGVGML
RAAC02322             MKIVYTLSFAVFLLVGTGTGIVAPLIGPYAHSLGAGGFWIGLLFSGFYIVRLLVGTPTVGRL
                        * .: : * *: *: : * *** *: :*: : * : : . *

ref|YP_504284.1|      SDICGRRQFIIIGLLLYAIFSLLYVPAGSVLELSGIRFLQGIASAMVFPIAMAYIGDIAP
ref|ZP_02131576.1|    SDMIGRKPFITIGLFAYSLVSIAMFMSHGVTQLVVIRFIQGIASAMILPVCQAYIGEAP
ref|YP_001046337.1|    SDRRGRKWIILIGMFAYAVLSLAYIIVDSVYSLTAVRFAHGLASAMVVPIAMAYVADLSE
ref|ZP_02132246.1|    SDSLGRKKFITVGLIAYTCMSVMYVLGSAETLIAVRIFHGMASVMVPIAMALAADIAP
ref|ZP_01288161.1|    SDRWGRKGFLLGGLLIFGVVGLLIQAVSVENLVIRALHGVGSAMIVPVAMAYVSDIAP
RAAC02322             ADHQGPKRVLIYSLILYPLISIAYSWAHSVGVLFARLLHGLASAMMLEPMAMTYVGAITP
                        : * * : .: .: : .: . . . * * :*: :*: : . : :

ref|YP_504284.1|      PGMEGRYLGTFTSSFLGMGLGPFLGGIVTDIAGMDTAFICMGALTGVALLTCIFFLP--
ref|ZP_02131576.1|    QGREGFFMGLFNISMYSLSLGPILGGVVNDHFGLQTAFACMGVLSLLAFFMAQHFLPPV
ref|YP_001046337.1|    KGREGSHMGNFISMFLGMGMGPLLGGFLNDAFGLDSVFYVMAGLSAFATILVGISLP-E
ref|ZP_02132246.1|    KHELGLYMGTLNMAVMLGLGAGPVLGGALRDHFGMDSAFIAMGLLALITCILVVIVIPSD
ref|ZP_01288161.1|    VGEEGRYMGLLNTAIFAGIGSGPLLGGIFTDLWGMPSAFYAMSTLSFLALGLILLQMPAM
RAAC02322             PGQEGRYMGIYNTFLFIAGGIGPLLGGLLASELSPNAAFSLSLLAIGSLLLVIPLPR--
                        * .:* . : . . **.*** . . . :.* .: *: :

ref|YP_504284.1|      ----GYQGSRKEQSSILHLLIHPGLRIPLLYQMMNAFANGTFMVFLPVIAAHVGNLSAGE
ref|ZP_02131576.1|    KEE-PRQERGREPVTYRILAKNAGIMGAVLLRTTYTCIGAIWGFLPVYADAHFNLSSSA
ref|YP_001046337.1|    AKRGTFTVQEGNPVPMRKILTLPMRGVMVFAFISALGRGMMVFIPVFG-PLIAISPT
ref|ZP_02132246.1|    AETRSFKENQAR-SSVGQIVRHRIALGIIFMRFLAASQGQAVYTFPLLG-LKMDMSSSQ
ref|ZP_01288161.1|    PAASRTGNADGLFTVMGKMLANRSTTGLLLARMATMLIMVPTMAFLPLMTQAFQASGMQ
RAAC02322             VSSASKEKNGMNHWMGDLIISPGLALAVLNVENAVLDVVFQVSFFPVFA-QNRGLSLLA
                        : . *: *

ref|YP_504284.1|      TGLVISVSVLSTALLQRVCGRLADRFDKYLLIATGCLTVTVALALVPGFEGLWSYLFAL
ref|ZP_02131576.1|    IGFLIMLGVLTSGLVLTTPMGWLADRVDRKLLLCVIGSVMSISVAVYSFTFADGFWHLFASNM
ref|YP_001046337.1|    VGIVLSANTFLMALLQVPMGRITDTGNKVVLIVAGSAITALAIGAIFFSGSFGPLLATS
ref|ZP_02132246.1|    VGILLSANIFLIAFLQRFSGRLADRVNPKYLVIFGTFLAGFAVLGMPVMVGKFFLVVLVNI
ref|ZP_01288161.1|    IGTVIAVRTLINASLQGYCGRLADRYDKLRLRLRYGCLVISGVMCLIPLAGHYVWVLLLF
RAAC02322             IGFLIAINSIVIGASQVPCGWLVDRTNKYYLVLVSGIVTSVLLTMFPLCRKLWVITMLMI
                        * : : : . * : * : * . : .

ref|YP_504284.1|      LMGIGGGISVPAMYALVTIAG-RDVGQGAAMGTINMVMSMGMIISPVCGLFMDQTGISS
ref|ZP_02131576.1|    AFGIGGGIMAPALTAICVIEGHKERAMGSVMALLTIGHSLGMLAGSVATGVLMDFSLSA
ref|YP_001046337.1|    LVGIGGAIQQPAIMALTVDAG-RTIGMTSMGAYNTVFVGIGMIIAPLMGGVFM DYIGIDA
ref|ZP_02132246.1|    VMGAANGLALPGGLVITAQLG-RTMGMASLMSITDAAWSLGMIVSPILSGIILDVYGLPY
ref|ZP_01288161.1|    VLGLGEAIIWPTLGALATEEG-RVYGQGTMMGVFNLAMSAGVLGGALGAGFATDLLGLTW
RAAC02322             LIGFSSAITIAASSALSTMLG-RTGGMGHVMGFLNSATSFGMIVGPIVSGIILDKLVNYF
                        . * . : . . : * : . . * . . *: : .: * . * :

ref|YP_504284.1|      VFYLSAVIVLVVTPVF-----
ref|ZP_02131576.1|    AFTASAVLAILGTVVF-----
ref|YP_001046337.1|    VFYVGGAISSLGTGIFAVMMQRNARAADRKDIP----
ref|ZP_02132246.1|    VFLMGSAIIL-----
ref|ZP_01288161.1|    SFPVIGLLVLLLTFLAIGMIAADRLA-----
RAAC02322             TFFYNSIVWLVTMLFAMLWMVHNRRISKQDVPISY
                        * . : :

```

29/228

FIG. 28

```

ref|ZP_01287831.1| -----
ref|YP_076198.1| -----
ref|NP_634267.1| -----
ref|YP_446560.1| -----
ref|ZP_01106621.1| -----SSCCGATTPSNKVYNIMDDYTETDGYVEDADLGLGCGLPTEFAKI
RAAC02323 MRKNVNANCCSAEPEPLTPAKRVIKDFYDTVAVGSGPYASTDRFSPAAPTEKILEYVKR

ref|ZP_01287831.1| -----VLDLGGGGGDFCFLAARQVGETGRVIGVDMTPEMISQARANATKSGD---RNVEF
ref|YP_076198.1| -----VLDLGGGAGFDCFLAARQVGESGRVIGVDMTPEMLARARENARKGGF---ANVEF
ref|NP_634267.1| -----VLDLGGGGGDFCFLAAQKIGSSGKVIGVDMTLEMVEKAQANARKYGY---SNVEF
ref|YP_446560.1| -----RVLDLGGGAGMDAFVARRTVGPDPGHVHGVDFAEEMVAKARANADTLDY---DNVTF
ref|ZP_01106621.1| KKGD TVIDLGGGAGNDCFVARHETGSEGVLGIDFTPIMIEKARINAEKLG Y---NNVEF
RAAC02323 TNCRRILDVGC GGMTLLRMAQEHVSGVQFIGVDFSEKMIERARTSSLSLHDDL RKKIGF
      ::*:.* * : : ..*::: *: :*: .. : : *

ref|ZP_01287831.1| RLGEIENLPVADGAVDVIIISNCVINLSPEKRRVFAEAYRILKAGGRLAISDVVATAELPE
ref|YP_076198.1| RLGEIEHLPVADESVDVIIISNCVINLSPEKEQVFREAFRVL RPPGGRIAVADMVSLAPLPP
ref|NP_634267.1| RHGDIESLPVKDSSVDVIIISNCVINLAPDKEKVFREAFRVLKPEGRMYISDMVLLDELPE
ref|YP_446560.1| EVGDIEALPVESGAFDVILSNCVLNLVDPDKEAFAQMHRALRPPGGRFSVSDV VHAGALPD
ref|ZP_01106621.1| REGDIDAMPVSDEADVIVSNCVLNLVFNKNKVIGEMFRVLKPGGHFSVSDIVLVGNLPE
RAAC02323 FVANAESLPYMEGQDFVFSECVLNLI PEREKAIAEVMRVLAPGGMFVYTDFVAFSPISN
      . : : * . *...*:*: * *... : : * * . * : :*. * :.

ref|ZP_01287831.1| AVRKDMALYTGCIAGASLVSDIEQMLTEVGF-TEIRVSTKDESKSF-----
ref|YP_076198.1| EVREDLALYAGCVAGVATV GELRTMLTEAGF-VDI-----
ref|NP_634267.1| ELKNDSELLAGCIAGAVLKEEYLG LKKAGFSVEILNEDLDISKRQYRDLPVESLKLK--
ref|YP_446560.1| GLREAAELYVGC VAGAMERDMYLDRLREAGF-----
ref|ZP_01106621.1| ALKADAEMYAGCVAGAIQKEDY LKIIEDKGFMNLK-----
RAAC02323 SIRDNLNLVSGCRAGSKT LSENIRLLEETGFVKIECIDFTSDKNKRYADIMNESEQIRRE
      :: : ** ** : . **

ref|ZP_01287831.1| -----
ref|YP_076198.1| -----
ref|NP_634267.1| -----
ref|YP_446560.1| -----
ref|ZP_01106621.1| -----
RAAC02323 FEEFRNLYPDAAAF LDERVGYYLILGCKPI

```

FIG. 29

```

ref|ZP_02595423.1|      MNILLVSNFEGGFQPNNTIATAATPLVKAGFDVEILDITYVEG-LVEEKFKDKQLVAISVPL
ref|NP_982173.1|      MNILLVSNFEGGFQPNNTIATAATPLVKAGFNVEILDITYVEG-LVEEKFKDKQLVAISVPL
ref|ZP_02367476.1|      -----
RAAC02324              MKILLVSVFEGGFQPTIATAAGALSEIGVEIDALDVYIED-LRFNKVSDANLVAISVPL
ref|YP_001616264.1|    MNVLLVSTYESGFSQSLGTSVAAAHLLEADVVRNALDLSVTPTPDIDDLCRHDLVGFHFLPM
ref|YP_827637.1|      MRVLLVSTYEMGRQPFGLASPAAWLRAEGHEVTQADVSCPT-MPKDAVEAAGLIAFFLPM

ref|ZP_02595423.1|      FDAVTAGIEVAKKVAEINPDAHITFFGQHATINANRLAGRYSDSCISGEWEHPLTLLAKH
ref|NP_982173.1|      FDAVTAGIEVAKKVAEINPDAHITFFGQHATINANRLAGRYSDSCISGEWEHPLTLLAKH
ref|ZP_02367476.1|      -----
RAAC02324              FDSIQPAIALSKEIRKINPRTHVTFFGQHATIHAQKLASTYGDSCIRGDWEYPLINLARY
ref|YP_001616264.1|    FHSVPAAVRVAGRLRQQEKAPKIFFYGLYADLFREKLLGRHGDYVLGTDWEDQIVPLVK-
ref|YP_827637.1|      HTATKFLFLRLVDRVRVNPRAHLCAVGLYAPLNERLLRGAGVGTVLGGEFESGLRDLARR

ref|ZP_02595423.1|      LSGETQESLPGVLSAEQAIKGESVHPYMARNHLDIPSRHLLPALHKYPQKQINRLLGSDQ
ref|NP_982173.1|      LSGETQESLPGVLSAEQAIKGESVHPYMARNHLDIPSRHLLPALHKYPQKQINRLLGSDQ
ref|ZP_02367476.1|      -----
RAAC02324              LMGNNS-LVGVLLADDAKKGKSTPVYIGRNDLRVPARSILPPLDKYPQPHVERIMSGKQ
ref|YP_001616264.1|    -----RSPDPSVVQLQKRGFARQNRYPARQVLPPLGNYAKLVED---GAHM
ref|YP_827637.1|      LS-----SGDPAPAAPPVTISIERQQFLVPDRKGLPPLGAYQLVVG---NGTR

ref|ZP_02595423.1|      VVGSTEIARGCHHKCLYCSVFAAYDGKVILVPEEIVLQDVRNLVEGGMTHLTFTIDADFFN
ref|NP_982173.1|      VVGSTEIARGCHHKCLYCSVFAAYDGKVILVPEEIVLQDVRNLVEGGMTHLTFTIDADFFN
ref|ZP_02367476.1|      -----MYAAYDGKVIIMVTDIVVEDVRNLVKQMEHLTFTDAEFFN
RAAC02324              VVGATEIARGCHHKCLYCSVFAAYNGKVVLIPEDIVLEDAKNLVEQGMTHLSFIDADFFN
ref|YP_001616264.1|    LAGCVEATRGCAHCHTCPIPPVYGKVTIIPPEVVLADIDNLVLMGARHVSFVDPDFLN
ref|YP_827637.1|      RVGYTEASRGCKHLCRHCPVVPVYNGRFRIVQADIVLEDIRRQVAAGAAHITFGDPDFN
                        : ..*.*. : : :*: * . * * *::* *::*

ref|ZP_02595423.1|      AKYHGIKILRKLHEEFPETYDFTTRVDHILENKKTLAEMKELGVKFITSALEFPSEEV
ref|NP_982173.1|      AKYHGIKILRKLHEEFPETYDFTTRVDHILENKKTLAEMKELGVKFITSALEFPSEEV
ref|ZP_02367476.1|      AKNHGVRIMRRLHEEFPHTYDFTTRVDHILEHEDAIREMSGLGLRFTSALEFPQKVL
RAAC02324              AKWHGIKIIRKLHEKYPFLTYDFTTRVDHILENRDTFREMVNGLVKFVTSALEFPSEELL
ref|YP_001616264.1|    VPRHGLSIMKAVNDKYPFLTYDFVAKVSHFRRHEQYVRELAKLGLKFVLTAMEFNDNEVL
ref|YP_827637.1|      GPGHAVPIVEALHREWPWLSYDVTIKVEHLLKHRDLLPVLKETNCAVFTSAIESLDDEV
                        *: *. : : :* *::.. :*: :... : . *: :*: * :*:

ref|ZP_02595423.1|      DAVAKDTSVADIEQGIAYLREIDIKLNPTFIMFNPWTTTFEDLTFRSFVEDNELGNIIDP
ref|NP_982173.1|      DAVAKDTSVADIEQGIAYLREIDIKLNPTFIMFNPWTTTFEDLTFRSFVEDNELGNIIDP
ref|ZP_02367476.1|      DIVAKEISVDDIEMAIRRLKAIGVKLNPTFIMYNPWVSKEDILSFKAIFIERNDLEDVVD
RAAC02324              DEVEKQITVNDIREAIKFLREIGMKLNPTFIMFNPWTRLEDLIEFRAFIEDNNLQDLIDP
ref|YP_001616264.1|    DILKKKHIDIDLDWSIGLFHELGVHLKPTFVMVNPWAEVGDIMDLLEFVETRGLIDAVDP
ref|YP_827637.1|      VKLAKGHTRAGFLEALVLMRAVDLPLSPTFIPFHPWTTLESYREFLRTLAENGLASQITP
                        : * . : : : :*.***: **: . : : . * . : *

ref|ZP_02595423.1|      IQYETRLYLYKGSPLLHKQSIQDLE--LTEYEFHYDWKHPDSKIDELYFEMLTPPEEGIF
ref|NP_982173.1|      IQYETRLYLYKGSPLLHKQSIQDLE--LTEYEFHYDWKHPDSKIDELYFEMLTPPEEGIF
ref|ZP_02367476.1|      IQYETRLHLYKGSPLLNRASAGLK--LTEREFHFDWSDPDAVDEMYANVTPPEPGVF
RAAC02324              VQYQTRLHLYKGSPLLKSVKNLE--LTEREFHYEWKHPDPRVDELYSQLHKPVENGVM
ref|YP_001616264.1|    IQYKIRLLLFNNSPLMDSVGLYASL--GEESDYTTEWRHRNPAAVEELHREICRWVDEAV-
ref|YP_827637.1|      IQLAIRLLIPEGSLLMELPEVRALVGLFDARALSYPWRNPDPGLDRICATIQATIKRGEK
                        :* ** : :.* *: * : : : : : . .

ref|ZP_02595423.1|      KRCCLKC
ref|NP_982173.1|      KRCCLKC
ref|ZP_02367476.1|      KRCCLKC
RAAC02324              KRCCLKC
ref|YP_001616264.1|    -----
ref|YP_827637.1|      HR-----

```

SUBSTITUTE SHEET (RULE 26)

32/228

FIG. 31

```

ref|ZP_01090358.1|
ref|ZP_01856486.1|
ref|ZP_02736297.1|
ref|ZP_01311632.1|
ref|NP_982171.1|
RAAC02327
-----IR
-----IK
-----MTVR
-----MSCR
EINBACILLUSCEREUSSCRESIGNIFICANCEE- IDENTITIESPSITIVESGAPSTAR
-----MSVK
:

ref|ZP_01090358.1|
ref|ZP_01856486.1|
ref|ZP_02736297.1|
ref|ZP_01311632.1|
ref|NP_982171.1|
RAAC02327
FVMIGGFLGAGKTTTIGRLAQHYRDQGLNVGIVTNDQATDLVDTQMLRSQGFRVGEVAGA
FIMVGGFLGAGKTTTIGRLAKYYSQGLNVGVVTNDQAADLVDTNALRSQGLHVGGEVAGA
FVMVGGFLGAGKTTTIGRLARHYQQQKRVGVVTNDQAHDLVDTNTLRAQGLSVVEVPGA
LILVGGFLGAGKTTTLLAETAHKLSLQGLKVGLITNDQATNLVDTRMLVRSGAGVAEVSQS
LVLLGGFLGAGKTTTMIKSALKLEKEGYRVAIVTNDQKELIDTELARANGLESKVETGG
LVLLGGFLGAGKTTTLIRAAHMFQSAGQNVAVITNDQGTTELIDTELSRLNELNTEEVTTGG
:::***** : . * * .*:****. :*:** . **.*.

ref|ZP_01090358.1|
ref|ZP_01856486.1|
ref|ZP_02736297.1|
ref|ZP_01311632.1|
ref|NP_982171.1|
RAAC02327
CFCCNFNLTGTVEKLAAEDRPDIVIAEPVGSCTDLVATVVQPLVRMFDAQFDVAPYGV
CFCCHFNALMDTIEELGSEQKPDVILAEVGSCTDLVATVIQPIKRLFDADFSILPYTVL
CFCRFDDLVGRVMSLEAGERPDVILAEVGSCTDLVATVIQPLKDLYSGRFEVAPYAVL
CFCCNFQGLLDAMDQKKTTFDADIVLAEVGSCTDLSATIIQPLKDKLQSKLLISPLTVL
CFCCQFDDLYNNLTLMKEKQPDVILAEVGSCTDLAATVIQPLKQYADKFKTAPLTIV
CFCRFNELYDKLISLKSQFRPDVILAEVGSCTDLAATVIRPLKQYYGKEFDVAPLTIV
****.*: * . : * .*:****.***** *:***: : * ::

ref|ZP_01090358.1|
ref|ZP_01856486.1|
ref|ZP_02736297.1|
ref|ZP_01311632.1|
ref|NP_982171.1|
RAAC02327
LKPSHGLRILQGEDSGG--FSPKAAIYFKKQLEEADFVIINRIDELDAEKVDLLASLITE
MKPSHGLKILKNDKSGG--FSPKAAIYILKKQLEEADLILINRIDELSAEEVDEITALVNE
FKPSHGLRILRNQAAGG--FSPKAAIYFKKQLEEADAIVINRVDDELPAQIDELSLLVTO
ADPIKLGAILAG-GTAG--LHDDAAIYIRKQLEEADLILISKVDLLTLPVDDLLIRVRK
VEPARLLHELNLNDENTKPFSSQSVSYIFEKQLAEADIALNKVDRYSEEEIAKLKAYLQQ
VDPGRILSEIGPKEKIN--FSRSVSYIFDKQLSEADIALNKIDKYEERCINDLEDYLQK
.* : : : ..** *** ** : :*: : : :

ref|ZP_01090358.1|
ref|ZP_01856486.1|
ref|ZP_02736297.1|
ref|ZP_01311632.1|
ref|NP_982171.1|
RAAC02327
AFPGTPIILRTSAKTGAGFDALVELIDQGEFGKKILDIDYDVYAEGEAELGWLNSSLKAT
QFPGTPIVLRTSALTGEGFEPLLEFLEQDQDGFSGKILDIDYDIYAEGEAELGWLNSSVHIS
QFPGTPIVLRASATTGAGFDALTELLDQTAGFRKILDIDYDIYAEGEAELGWLNATARLT
AFPSAIVQPLSAERNVGLDYWLDHVLHKTAGNTLLDIDYDRYAEGEAVLGWLNARFELR
RYPQTIIQTFSAERGDNLALHTHTWLTDTLGGDKVLDIDYELYAEGEAQLAWMNLGDIK
RYPTAVIHTFSAHRGDNMERLLDLWRTLNISGEKALDIDYDKYAQGEAKLAWNLSDILS
:* : : ** . : : . * .*****:***:**.*:*

ref|ZP_01090358.1|
ref|ZP_01856486.1|
ref|ZP_02736297.1|
ref|ZP_01311632.1|
ref|NP_982171.1|
RAAC02327
ASEAF--DLDAFLMAIMNAMQKTLANSHAETAHLKTIGLWEG-FYGVANLVSSDTAPILS
AENSF--SLDQLLLDVISQLQTSFKEKSVETAHLKTIGLWEG-FFGVANLVSNDSERPRLS
SDEPF--DLDAALLTSVLGELAVVCRSLGGEV AHLKLGMDASAFAGNVISSDTAPKLS
TTAG---TWRTFAEAIMTRLHHQFAERNLPIGHVKLIVESSES-QVLFANLTGTEIEPKIR
AEQD--INPREWAEKLLDNLNKHFI REKMAIAHLKVHVGFEG-GYVKASMVQTDAPTFT
VDEANHFDAKLWMHMFLLDKLHEHFLREKMAIAHLKVYVGS DN-GFIKASIVETGDTPTYI
. : : .*: : : . *

ref|ZP_01090358.1|
ref|ZP_01856486.1|
ref|ZP_02736297.1|
ref|ZP_01311632.1|
ref|NP_982171.1|
RAAC02327
LPSNCQAKEADV VVNARVGISPEELRQVDDAIDAAAKKFGVRIERQQTQYFRPGRPVPT
LSSDCTVTEADLIVNARVACDPEALTAQVKQVLQNCQAQALNAKLEFRQTQSFRPGRPVPT
LPSGLRPREADLIVNARVAIDPAVLEAEVKRVVALES AKVGVIAEFRTSQSLRPGRPVPT
GEMTSTP-HATMTINARVETSPPELRQIIWQALQSSDGINVYQEOWN--CLKPGRPEPT
VENVKEGKEFRIVLNIRIEATPAIILNVVADSIGAINKELGSNWSETYNECFSPPLPKPV
VKEPVNGQEFRVVINIRIETEP ELLSLVVADALAIVNKEMKVWKQTYHECFSPPLPNPT
. : :* *: * * : : . : * * *.

ref|ZP_01090358.1|
ref|ZP_01856486.1|
ref|ZP_02736297.1|
ref|ZP_01311632.1|
ref|NP_982171.1|
RAAC02327
HR-----
HR-----
HR-----
YR-----
HRL-----
HRLRVAAK
:*

```

33/228

FIG. 32

```
ref|XP_955124.1|-----ELCDPVLEAN-GLKVAAKSTFGDM
ref|XP_763458.1|-----ELCDPVLEAN-GLKVAAKSTFGDM
ref|XP_666904.1|-----NTNLEEN-ELSDKTINLIQNL
RAAC02328      MIRGEEILLRKIVVEVLPKYSS-ENCCSVDGEETTIDCCNPLVEAN-ELKDKIKNVLGDL
ref|NP_982170.1|-----LIEIFPSVVTNQTD-CCDSDDNNEEGCCEESNQFFEESLELKEQLKEHFDKL
ref|XP_845342.1|-----LAEALERYSDSEQLCSCTLRVFSVFPSSPYVDMN-AIASLFKNTSESA
                  . . : . : . . .

ref|XP_955124.1| IIVELYKAAYVLGLWDVAAYNCLKT-CKRRGY-----
ref|XP_763458.1| IIVELYKAAYVLGLWDVAAYNCLKT-CKRRGY-----
ref|XP_666904.1| VIIHFSD--LIINQLPLNSVNHTESLCKER-----
RAAC02328      VVIHVFD--YTLSVDWSLAVNRLRQLFKERGFRLSEMDNILEVVTPVIVVNGNLVSFVC
ref|NP_982170.1| VNVHLYN--YELSMDRVLAQKKLSQLIQERGFNGNISE-DSVLRVYVTPAVVNGTLVSFAT
ref|XP_845342.1| VAIEAIH--FLCSVA--MSVKTLKQLIHTTGC--VQRVLEVMRKYAGNASIQEYACSLLS
                  : : . . : : :

ref|XP_955124.1|-----
ref|XP_763458.1|-----
ref|XP_666904.1|-----
RAAC02328      PFDFEKIVSRIMTMSYE
ref|NP_982170.1| KPQMESVL-----
ref|XP_845342.1| YLSFDSETITSFI----
```

FIG. 33

```

ref|YP_079109.1|      --INIRKAISADLPFILTIYNQGIEDRIATLEQDLKEMSDIEIWFQEHQGRYSVLVAE-S
RAAC02332             MMVTVRVATHRDLPSILAIYNQGIEDRVATLEQDLKMDYITNWFNEHTNRYPVFVAE-Y
ref|ZP_01858609.1|    --IIIRKAEIRDIDRILTIYNQGIEDRVATLETEPKDRTYMAEWLKAHQGRYTGIVAE-L
gb|AAV70501.1|        ----IRKATEQDAQEIMTIYNQGIEDRIATLETEIKTDKYVMEWLFQREKRYSVIVME-E
ref|YP_001319533.1|    MAYTVRIAELQDLKSITEIYNQGIEDRIATLETKIKSEAEMIPWLQORSEKHKVITIENE
ref|ZP_00539168.1|    ----IRPVKEDDLKDILAIYNQGIEDRIATLETDTKELSFMEDWYRERTPRYAGFVAY-E
                        : * . * * ***:*****:***** . * : * : : : .

ref|YP_079109.1|      KGEIVGWASLNPYSHRCAYQGVDLSVYVDRACRGKIGIGLLQLALEKTAKENSFYKIVL
RAAC02332             QNNIIGWADLHPYSHRCAYGGVAELSVYVHRGWRAKGVGQALLSALEAFARKHDFHKLVL
ref|ZP_01858609.1|    EGEVIGWASLNPYSPRKAYAGVADISVYIERNRSGKIGRKLSSLEVKARENGFHKLVL
gb|AAV70501.1|        NSHIVGWASINPYSHRCAYRGVGELSIIYIKREYRGKLGQKLLLALEKTGQKNEFYKFVL
ref|YP_001319533.1|    RDEVLGWASLNPFSRCCYDGVADFSIYIKRQMRGMGLGKLLKALIEVAREQDIHKLVL
ref|ZP_00539168.1|    GTTILGFISLDPYNPRPVYQSVGELSVYITRTHRGQGIGRQLLHVIEEHAIQTGFHKLIL
                        ::*: :. :*. * * .*: :*: * * .*: * * : . : :*: :*

ref|YP_079109.1|      FTFFPNELGQNLNKMGRYQVGIFKNQGILDGRFIDVMAMEKLL-----
RAAC02332             ATFFPFNSAGLALYRKMGFREVGVFMNHGRLDGKWVDVLWMEKLLLELKVE
ref|ZP_01858609.1|    FTFFPFNKLQQLYQKCGYREVGTFFKNQGILDGIYVDVMAMEKLL-----
gb|AAV70501.1|        FTFSFNNLGQGLYRKMGYREVGVFEKQGIMDGEHVDVMIMEKLLSVE--
ref|YP_001319533.1|    STFKSNEAGQRLYESFGFREVGTYKNQGILDGKFVDVTIMEKLL-----
ref|ZP_00539168.1|    FTFFPFNKIGQKLYIRSGFRIVGTFFKEQGMLNGYYVDVLAMEKLLPKTID
                        ** * . * * * *: * * * : :*: :*: :* * *****

```


35/228

FIG. 34

```

ref|YP_001358015.1|      ---ILIDTNVVRFLTRDNEKYFLKSVAIFQDIEDGKMEAMLMDFIVAEIVSVL-HRIYK
ref|ZP_02178354.1|      ----LIDANVILRYLLRDNKELYAKAEVFNDMMEGKTKILILESVIAEVVYVL-QRIYN
ref|YP_001356736.1|      ----LVDTNVLVRLFTQDNEEMFREAFSFFEKVANFEIQAVISEGVLLSWFVL-QKVYE
ref|YP_001485227.1|      ----VVDTNVIVRFLTNDDEEQSKKAFEIFKRAVNQEITLFLTPIVIECTWVLGMKRYG
RAAC02334                MKSVLLDTNVILRFLDDHPHSPAATNLFMKAEQGEIKLVIDPMIVAECYVLEGKVYQ
ref|ZP_02013298.1|      MKTWLLDANGILRFLRNDDAVQSPQSKLLERANAGEVRLISVLTVAEVFYAL-RASYK
                        :::* ::* : *.      : .:      : .:      : *  .*  *

ref|YP_001358015.1|      HSKKEIATTLKKLLLYDHLTYTENKLITFEALEIYAEKNIDFADAVLYAKQRLEGFEIISF
ref|ZP_02178354.1|      VSRKEISEILRKLIELRGVKVHNKGQMLNALEIFSEKKLDFVDCILCAHR--EENEVITF
ref|YP_001356736.1|      MKKDEIIKRLLTIVTLRNVILEDKLAFIEALHILKERNIDFIDAMLCVKSNIKGYKVFSF
ref|YP_001485227.1|      YSHEEIADNLTITITSPGVKTLES DVVLKALHEYKTNRIDFADAYLA AVSTKNGVPCLTW
RAAC02334                FSKELIASHLIPVLTHEGVQCEALMTVLDGLDIYVEHRLDFADAYLIA LARNHRFEVATF
ref|ZP_02013298.1|      MTRPAAAQVLMRFVLTGVAEVEREDVLLDALQRVVSANVDLGDAVLAVEAVRHGEGVASF
                        .:      *  .:      :...*.      .::* *. * .      .      ::

ref|YP_001358015.1|      DKDIEK-----
ref|ZP_02178354.1|      DKGLRR-----
ref|YP_001356736.1|      DNDIQR-----
ref|YP_001485227.1|      NSK-----
RAAC02334                DNKMARIENVVFQDL
ref|ZP_02013298.1|      DKDFSRTDV-----
                        :.

```

36/228

FIG. 35

```
dbj|BAF91394.1|-----KNMIPKISDFGMARIFARDETEANTKKVV
dbj|BAF91409.1|-----KNMIPKISDFGMARIFARDETEANTKKVV
gb|AAL17690.1|---ARGLVYLHQDSRFRIIHRDMKAGNILLDKNMIPKISDFGLARI IARDQTEASTDNPI
ref|NP_001058416.1|-----
dbj|BAD45624.1|-----
RAAC02335MHTSRGVGYMNPDEPVFEILALVWEA-----KRRLPKVSDYRRIARKRQTESPDHVIQRY

dbj|BAF91394.1|GTYG----YMSPEYTMHGIFSEKTDVFSFGVIVLEIVSGKKNK-----
dbj|BAF91409.1|GTYG----YMSPEYAMHGIFSEKSDVFSFGVIVLEIVSGKKNR-----
gb|AAL17690.1|GTYG----YMSPEYAMYGILSEKTDVFSFGVIVLEIVTGKKNR-----
ref|NP_001058416.1|-----
dbj|BAD45624.1|-----
RAAC02335GTWSRVLEDFQAYFTLHGYP TIDPDLFGALFSRAKVVRDKRGRWYLRVSDDTTQIEIL

dbj|BAF91394.1|-----GLYNLSFENNLLSYVWSQWK---EGRAL EIVDPVIVDSLSSLP
dbj|BAF91409.1|-----GLYNLSFENNLLSYVWSQWK---EGRAL EIVDPV IIDSLSLSP
gb|AAL17690.1|-----GFYQSTPEDNLVCYAWTHWA---QGRAL EIVDPVIVDSL S---
ref|NP_001058416.1|-----TVIMRCTQVALLCAQE-----DPADRPTMTDVTAMLDSESIML
dbj|BAD45624.1|-----TVIMRCTQVALLCAQE-----DPADRPTMTDVTAMLDSESIML
RAAC02335RSFMRRSAATLVNRGKPSLYLR CYDVLICAYEHACKTEPVFRPTVDFVRGYIDTHSHFR
.      :  *:.      *      *:  *

dbj|BAF91394.1|STFQPQEV LKCIQIGLLCVQEHA-----
dbj|BAF91409.1|STFQPQEV LKCIQIGLLCVQEHA-----
gb|AAL17690.1|STFQPKEV LKCIQIGLLCIQERAVHRPTMSSV-----
ref|NP_001058416.1|SDPKEPTELTHGGAPVLTGHRHILVTQAR-----
dbj|BAD45624.1|SDPKEPTELTHGGAPVLTGHRHILVTQAR-----
RAAC02335KDPSGRDRLTLTGPLVPECHDFLVALGARNTRVSRVKDSYRMNVHAGSLRRIREALYPFG
.      *.      :      :

dbj|BAF91394.1|-----
dbj|BAF91409.1|-----
gb|AAL17690.1|-----
ref|NP_001058416.1|-----
dbj|BAD45624.1|-----
RAAC02335CVCNPEIRARIFTV
```

37/228

FIG. 36

```

ref|YP_517489.1|
ref|ZP_01370335.1|
ref|YP_517477.1|
gb|ACA46983.1|
ref|YP_001568284.1|
RAAC02336
-----
-----DILMKVLAEQFKNKTLDVFGIKTAKIKDLIPSVHPAVEANETRNDIIFLLEDDTLL
---KDVLFKTLSEMFKDKALIMYGLDYPKIVEMLPNEFPEVKADERRADSIFLLEDGSIL
-----DFVFELEDNSLL
MPVQDTLMKALTRSLSGGALDVLGVHGVLDLVEPFATELP---ANTLRMDRVWRTVDGNLF

ref|YP_517489.1|
ref|ZP_01370335.1|
ref|YP_517477.1|
gb|ACA46983.1|
ref|YP_001568284.1|
RAAC02336
-----YDARLASRQE-----RQIRTFVVYSGHIEQAKERLDCGSI
-----YDARLASRQE-----RQIRTFVVYSGHIEQAKERLDCGSI
HLEFQTT-AGEQDLKRFLYDARLVRQE-----RKVHTIVIYSGRIEQARERLECGSI
LLEYESNNRITENMYKYIDYVLRISRKYEEENKSIKKINVAVIYASNIERAEDHFNIGSV
HLEFQTTWKKAD-LLRFAQYDIALYQKER-----RRINTVVMYSGKYESAESELDMGSN
HLEFQTK--RESTLHRFLEYDARLANEHR-----TRIRTVVLYHASVASAPDELDIGTA
* : . :...*: * :...*:

ref|YP_517489.1|
ref|ZP_01370335.1|
ref|YP_517477.1|
gb|ACA46983.1|
ref|YP_001568284.1|
RAAC02336
LYQVENIYMKDYNGDQEYNRLKAKIESGQLNETDTLKLIFLPLMKSRQOEEDLAIQAAE
LYQVENIYMKDYNGDQEYNRLKAKIESGQLNETDTLKLIFLPLMKSRQOEEDLAIQAAE
LYQVENIYMKHYNGDQEYNRLKHKIDNHQLLSETDTLKLIFLPLMKSEQKEEELAIQAAE
GIDVKSVMKNYDGTIYKEVEQKIKNGINLKDIDLMLNLIILLPLMKSIKDKHELKDTIE
KYKVQQIFMIKYDGIKRYEEIKEKIEKEEELTDKDLMDLVFLPLMRNEKSEEEVTKDVFE
LYRVENVFLRHLGDQALNEVEHHLRAG-QWEPGDRMRLGLALNMR-LENQAHAFERVRE
*:...:.* :...: * : * : * :...:.*

ref|YP_517489.1|
ref|ZP_01370335.1|
ref|YP_517477.1|
gb|ACA46983.1|
ref|YP_001568284.1|
RAAC02336
LAKAT-DEKTKVFAIAALIVITDKIISESNKRKLEVLK-MTQIEQWIREEGREEGRQEG
LAKAT-DEKTKVFAIAALIVITDKIMSESNKRKLEVLK-MTQIEQWIREEGREEGRQEG
LAKAAPDEKTKLFAIAALIVITDKIMSESNKRKLEVLK-MTQIEQWIREEG----RQEG
LAKEVKDEKNQYFIIAGVLTSTDKFIDEKYANTVRSWLR-MTKVEK-IFEKE----KEEA
LAIKIPDEDKKEAVIGSLLGFSDNYVRDEYINELKEVIR-MTKIGTSLFEEG----VEEG
LIPQVPDETERDLVVSAILVLGDQGLTEEQRALLRKELRNVSKLAEELYEEG----REEG
* ** : :...: * : : . : . : : : : * : :*.

ref|YP_517489.1|
ref|ZP_01370335.1|
ref|YP_517477.1|
gb|ACA46983.1|
ref|YP_001568284.1|
RAAC02336
ELKGRQEEKRETARTMLSMGMSPEVIAKATKLSQEEILRIEKEMKN
ELKGRQEEKRETARTMLSMGMSPEVIAKATKLSQEEILRIEKEMKN
ELKGRRDEKRETAQTMLNLGMSPELIAKATKLPLEEILEMAK----
MNQAEKSKAIEIAKNLMDI-LSVEMIAKKTGLSIEEVEKL-----
-----
RMVGRMEERMEIALNMLRKGMSVEDISDTTGLTKQIEELARKQSH

```

38/228

FIG. 37

gb|AAB91591.1|
ref|NP_391247.1|
ref|YP_001422657.1|
ref|YP_093160.1|
ref|NP_391246.1|
RAAC02292

MSPFGQQLRELRRARKLTVNQLAVYSGISSATISKIENGKRGTTPKPATIKKLA AVLKVPY
MSPFGQQLRELRRARKLTVNQLAVYSGISSATISKIENGKRGTTPKPATIKKLA AVLKVPY
MNTFGKQLRELRRARKLTVNQLAVYSGVSSATISRIENGHRGIPKPATIRKLADTLKIPY
MTNFGHHLRQLRERKKLTVNQLAMYSGVSSAGISRIENGKRGVKPKPATIRKLADALKVPY
MESFGEQLRALREERKKLTVNQLATYSGVSAAGISRIENGKRGVKPKPATIKKLAEALKIPY
MSQFGQYLRKLRKERNLTINQLALYSGVSSALISRIENGQRGRPKPDTLKKLASALKVPY
* *. ** *. ::::**** *:*: * *:****: * * * *:**** .**.*

gb|AAB91591.1|
ref|NP_391247.1|
ref|YP_001422657.1|
ref|YP_093160.1|
ref|NP_391246.1|
RAAC02292

ENLMAAAGHIRAFPEEIREASES-----
ENLMAAAGHIQAFPEEIREASE-----
EELMARAGHIKAFQEEIRETSSES-----
EELMASAGYISAS--TVQEARS-----
EGLMYKAGYIEE-----
EDLLHAGVLNEQISRTSESRLKPVDPSWYKRQVPIPVLGSI RAGTPVEMLALNSSEFV
* *: ** :

gb|AAB91591.1|
ref|NP_391247.1|
ref|YP_001422657.1|
ref|YP_093160.1|
ref|NP_391246.1|
RAAC02292

LVDSDLLGNHEGFALEVVGDSMIGDYIFPGDLVIVKYTSNFS PQDICVVAINGEEATLKR

gb|AAB91591.1|
ref|NP_391247.1|
ref|YP_001422657.1|
ref|YP_093160.1|
ref|NP_391246.1|
RAAC02292

VKCQGDICILTPSNPSMEPMVYNSVDVHVIGVVVEVRRRLRNK

FIG. 38

```

ref|YP_195796.1|-----
ref|NP_817052.1|-----
ref|NP_664934.1|-----
ref|ZP_02626811.1|MYEEYLKHFDLHVSDDIKNFVEETLKDKEYLFVRNEYEPESISNLKVKVQYGYCSHCRKE
gb|ACA41259.1|-----AHFPAEVPEHVQFEADEKALLLSRYIFVSSGKN-----RTGYCTHCRNT
RAAC02337-----MSHDFFAHFPTEPSAALIEYARDVAFLRSRYLFVWRDGK-----VQMAYCTHCRET

ref|YP_195796.1|-----
ref|NP_817052.1|-----
ref|NP_664934.1|-----
ref|ZP_02626811.1|FKTG---EYLKHNEVIICPNCGAEIMVKNIGRGKKCLMNDFCFYWFDKSVLDPEIVTCKG
gb|ACA41259.1|FPIE---NPIKQNGKGECPACFSKCRYKKAWLGRKALIDTACILHCQKSVLNPSVTVVEW
RAAC02337YPLGSDKDTRYHNEQATCRQCGSTCVVKHRGRGRRLIDTAYIVWYEKSMVDAQTVVATW

ref|YP_195796.1|-----
ref|NP_817052.1|-----
ref|NP_664934.1|-----
ref|ZP_02626811.1|YYFSKKYDEDYKNPKYEYNLSLYVFNTKDNTSKMFKNWYTVESWEERSSIFDFNINWL
gb|ACA41259.1|LYASRDYRESFENVSTEYAAVARYVFDYDRKVCQKIE-GGSGCRWWTESGFSTPSFQLNR
RAAC02337YYAWRDYSGDFHDVETQFQPRARYVFPGR-GGSMMRLDWSGEWQPRNVHPLPVGMMGW

ref|YP_195796.1|-----
ref|NP_817052.1|-----
ref|NP_664934.1|-----
ref|ZP_02626811.1|GNKLCYCSFESLDEAVKDTIFQYSNYKRRCGYKS-IVRYLDAFNKYQWIEQLEKMGFDRL
gb|ACA41259.1|SFMSRNIMHGTFFEDGISGTPFQYSGWQR-YDYED-VLKYLPLVAKYPSVEILTAKGLDSF
RAAC02337RTVDVWCSHESIQTAVAGTPLQYSGWETYCDQYNTLLTFFDLASRYPCVEYLTKLGSRL

ref|YP_195796.1|-----
ref|NP_817052.1|-----
ref|NP_664934.1|-----
ref|ZP_02626811.1|VETLENRWTDDYSINRYGKDIFFKKLKLNRGEVKQLTNEQKRELDSEFVLRVYQLNRDNGFN
gb|ACA41259.1|VKAKIYGYKTYSAVNWSKSKLHHIFKMSKQDLQMLREKNFVSSVYVYVGFNFNAWVQQW
RAAC02337VMAKLDGLRTFGAIHWRGRTMEQVLRMPRADVHAFRKLADVIEP-LSLSRYQTWRRLGWK

ref|YP_195796.1|-----APMINEIEKYVHYSQIKQLPKEVNLT-----KFQKWFIRKGE
ref|NP_817052.1|-----APMINEIEKYVHYSQIKQLPKEVNLT-----KFQKWFIRKGE
ref|NP_664934.1|-----PTCCCUSPYGENESMGASSCRESIGNIFICANCEE-IDENTIT-----IESPSITIVES
ref|ZP_02626811.1|PSFVEAYEVYEMCRYSDSSGLNNLIKIVGTNKVIKYEKQYNSF-----NESNSYYNKAS
gb|ACA41259.1|RKEKSKMNIDELQKNMQIFSVDKDLTFKFKFIRRFSTSLHRLFNANKQFEKDEKHFIRRHQ
RAAC02337VSPEEAHLLRELLIPHWHREIQVRASLASEVEIAKYLLKQLRKG-----EYRSISY

ref|YP_195796.1|RFDYYMDYLMLEELNTPLNNDVLYPENLQVAHDNAMNTLNLLKSEIEEKQYQERKNQI
ref|NP_817052.1|RFDYYMDYLMLEELNTPLNNDVLYPENLQVAHDNAMNTLNLLKSEIEEKQYQERKNQI
ref|NP_664934.1|GAPSYIDYLSMLDELGIEIDTDNLIMPDKDLVKAHDNVVKLVNQKKSEIEKRKFKNRLKSL
ref|ZP_02626811.1|VISKWNDYLNCKEKLNWSIRDKSILLPKNLVKAHNHTTSLINAEKNKADQKIKRRLPGL
gb|ACA41259.1|VLITWSDYLNDCQKLIHID-EAIIFFKNLRKAHEETIKRVKHYEDELMRKKAKERYEKI
RAAC02337ALIEWIDYLRFCSELGIPLNTTRNVFPSNLREMHDEMMRRVRIKRDERLNAQIQARLAEI
: *** .:* : : *:* *: . . . : : * :

ref|YP_195796.1|KALEAEIDDLLFLTPHSLQE-IIQEGSILRHCVGSQHYIERHTQGKTTIVFIRKEKPD
ref|NP_817052.1|KALEAEIDDLLFLTPHSLQE-IIQEGSILRHCVGSQHYIERHTQGKTTIVFIRKEKPD
ref|NP_664934.1|SKYEQTIGDYCFRAPVDSGE-LIREGKTLSHCVGSARYTQAHASGKTTIIFIRRKSDSDK
ref|ZP_02626811.1|KKKYFFKDKDFFIRPAESSEDLNKGGLNHCN-APHYMKPYANKETDILMIRRIDNPTV
gb|ACA41259.1|KHFEFELGQLKIVVPYTFKE-IIDEGNKLSHCV--GGYAEHADGQTTILFVRNIKEPDE
RAAC02337EPFGYT-DEHFMIRPARSVQELFDEGQALHHCVG--SYAAKYASGETNLFLMRRVSAPDT
. . . : * : : : * * * * : . . : * : : : . . .

ref|YP_195796.1|PYFTLEYRNQQVIQIQGKCNR-----
ref|NP_817052.1|PYFTLEYRNQQVIQIQGKCNR-----
ref|NP_664934.1|PFYTMKYKAGHIVQVRGKHNSATEEVQKVVDQWLAIVNKNYKHA-----
ref|ZP_02626811.1|PLVTMEIKNGQVKQAYGKNNDTIPKKDVEKFIEKFKTEILEKINSSKKNK--
gb|ACA41259.1|SFYTVFVKNEKVLQVRGKIKNKPATEDVQTFIDEFKKQKLTKKARK-----
RAAC02337PYCTVEMRDSALIQAARGERNLLREDEQTFVDQFVRHVARMHKRAQRKAS
. *:* : : * * :

```

FIG. 39

```

RAAC02338
gb|ACA41261.1|
ref|ZP_02626812.1|
ref|XP_001701427.1|
sp|Q2M3V2|ANR43_HUMAN
ref|NP_664935.1|
PRTEINSTREPTCCCUSPYGENESSI-GBAATUNKNWNPHAGEPRTEIN

RAAC02338
gb|ACA41261.1|
ref|ZP_02626812.1|
ref|XP_001701427.1|
sp|Q2M3V2|ANR43_HUMAN
ref|NP_664935.1|
STREPTCCCUSPYGENESMGASGBABFPHAGEPRTEINSTREPTCCCUSP

RAAC02338
gb|ACA41261.1|
ref|ZP_02626812.1|
ref|XP_001701427.1|
sp|Q2M3V2|ANR43_HUMAN
ref|NP_664935.1|
-----MYEKAVEKITL-EMSANEHHPYVQVIG-----
-----MVEQAMAKLSA-ELRVERVNPYVAAVG-----
-----AVEKIKE-EMS-KESNPYIAYIG-----
-----DMAARYLSTQE-ALLVATMSIHNGYGI-----
-----VLGFLQEHGGKVRN-SELLSRFKPLLDAGDP-----
YGENESMGASSCRESIGNIFICANCEIDENTITIESPSITIVESGAPSEK

RAAC02338
gb|ACA41261.1|
ref|ZP_02626812.1|
ref|XP_001701427.1|
sp|Q2M3V2|ANR43_HUMAN
ref|NP_664935.1|
--RFLLVHLEQHPEHAKN-----LLEEKTIILKSLDTM
--NFLMDFLEDNPSSAEK-----FLAEGKTIIGSMKEV
--NYVLENIEVNKPAAEK-----Iavgsktisesfkkv
--RLTWTQRPPRPQPPNR-----LAPGAGTIGSSGASP
--RGRAARRDRFKQFVNN-----VAVVKELDGVKFVVL
ALSKMLDEMNNKKSAAEDAIHNWLSQDDEELFKGILDDKKS IKDAMKYC
:

RAAC02338
gb|ACA41261.1|
ref|ZP_02626812.1|
ref|XP_001701427.1|
sp|Q2M3V2|ANR43_HUMAN
ref|NP_664935.1|
RRYAETQRVGNVAVISDADGFGIVLQYFDCWDGKPFEPPEPQPPARVAA
R-----
QGEARKIAQNGIAMLAEREVISIVTKYFE-----
QASPRAS--GTVPSAGTTAAASDAD-----
RKKPRPPEPEPAPFGPPGAAAQPSKPTSTVLPRSAS--APGAPPLVRV--
INQAQKQKTGNCAMVDDKTVFGWVRVYFT--GKTKKIEPVQATVTVSQE

RAAC02338
gb|ACA41261.1|
ref|ZP_02626812.1|
ref|XP_001701427.1|
sp|Q2M3V2|ANR43_HUMAN
ref|NP_664935.1|
PTRVTSAKSAQASKKSTAPALTQLSLFDDAEGGEAL
-----
-----
-----
-----
NTKVIPKK----NKKSKGVVEGQLDLF-----

```

FIG. 40

```

gb|AAU83457.1|      --IKGKVVLIIPFPFDDL SAEKVRPAVCLTEPIGPHHHVILAFITSQIPQKLETDLVLDA
emb|CAJ70907.1|      -----KVVLVNFPFDDLSTSKVRPAVCLTNPIGSHSHVILAFISSRIPSDLLETDLIIDT
RAAC02339            MKMKHKIVLIIPFPFDDLQSHKVRPALCLTDPISAHDHVVVAFISSQVPTHLLPTDIVLRA
ref|YP_183482.1|      MNLKGGKFLVLPFPFTNLKATKLRLPALVLYE---GKEDVVVAFVSSRIDTFDPGTDVRIET
ref|ZP_00514953.1|    ---KGDIVLVFPFPFTDLSTTKLRPAVVLWVDL-SGIDVTVCFISSQNVNHVTPEEFVIET
ref|YP_322920.1|      ---KGDIVLVFPFPFTDLTGTKLRPGLILWVDN-SGNDITLCFISSQDVTNLTSGEFVLDP
                      ..***: *** :*  *:***.: *      .: .:***:      :. : .

gb|AAU83457.1|      GQPDFALVGLNVSSTLRRLHRLMTVTTSLILRELGELPFDMKTEVDKRLRLKLFV-----
emb|CAJ70907.1|      SHEGFSGTGLKVASTLRRLHRLMTVTTSLCQRELGELSPKFLSEVNNKLKLF-----
RAAC02339            GTDAFAPTGLKVSSVIRLHRMVSMTTRIIRYELGHIADETRDLVESKLIQLFNLQHRLR
ref|YP_183482.1|      SHPQFRGIGLKVPSIIKLTKIATLHKGLLIGVLGELPEDLVQEVNSKL-----
ref|ZP_00514953.1|    TNSEFSKTGLKVTSKVRVSRMVTIERNLITRRLGKLDISLLNKLNDCLKRVFQL-----
ref|YP_322920.1|      SDAEFAGTGLKLISKVRVTRLVTLERRLITRRIGKLGVNQIQQLNLLMIQAFQL-----
                      *   **.: *  .: : : : :   :   :*.: .   : : :

```

```

ref|ZP_02432977.1|      --EKLKEITDRLEQGIAELFDSERYREYLKVMKSFHNYSFRNTVLIAMQKPDASLVAGFS
ref|ZP_02085861.1|      --EKLKEITDRLEQGIAELFDSERYREYLKVMKSFHNYSFRNTVLIAMQKPDASLVAGFS
ref|ZP_02207628.1|      --EKLKEITDRLEQGIAELFDSERYKEYLKVMSKFHNYSFRNTVLIAMQKPDASLVAGFS
ref|ZP_02429891.1|      --EKLKEITDRLEQGITELFDSERYKEYLRVMSKFHNYSFNNTLLIAMQKPDASLIAGFN
ref|ZP_01731985.1|      ---KTALAFQOOLEQGIAELLESQDWQRYLKVQSEFHNYSFNNTLLILSQFPEASRVAGYQ
RAAC02340                MNEKVKAAAMERLEHGLETLTLLSTEWRKFLQFQAQAFHHYSFSNTLLIMCQKPDATYVAGYN
                        *      :.***.*:  *:.:  :.:.*:.  :  **:* **.* ** * *.*: :*.:.

ref|ZP_02432977.1|      AWKNNFERNVMKGQKGIKI IAPSPYKIKQEMQKIDPHTQKPVIGKDGKPVTEEEKVTIPA
ref|ZP_02085861.1|      AWKNNFERNVMKGQKGIKI IAPSPYKIKQEMQKIDPHTQKPVIGKDGKPVTEEEKVTIPA
ref|ZP_02207628.1|      AWKNNFERNVMKGQKGIKI IAPSPYKIKQEMQKIDPHTQKPIIGKDGKPVTEEEKVTIPA
ref|ZP_02429891.1|      AWKNNFGRNVMMKGQKGIKI LAPSPFKIKKEMEKIDPQTQKPVIGKDGKPVTEEEKITIPA
ref|ZP_01731985.1|      HWQ--ELGRQVKKGSKSIKILAP-----LKCK-----IEKENDNGELEAKTGIFG
RAAC02340                TWR--ELGRYVKKGEGHIEIFAP-----LLKKSKNKADIPRESEETSQEENKRIIYG
                        *:  :  * * ** .:.*:***      *      .:      * :  * .

ref|ZP_02432977.1|      YKVVSVDVDSQTEGKELPDIAVDE--LTGDVDRYKDFFAALEKTSVPVPIAFENIEGGSHG-
ref|ZP_02085861.1|      YKVVSVDVDSQTEGKELPDIAVDE--LTGDVDRYKDFFAALEKTSVPVPIAFENIEGGSHG-
ref|ZP_02207628.1|      YKVVSVDVDSQTEGKELPDIAIDE--LPGDVDRYKDFFAALEKTSVPVPIAFENIEGGSHG-
ref|ZP_02429891.1|      FKMVSVDVDSQTEGKEIPNIAVDS--LTGDVERYKDVFAALEKTSVPVPVGFEKIEGGAHG-
ref|ZP_01731985.1|      FRTVNIFDISQTQGEDLPERTSP--LTGDDDDGLIDRLMAFSLNNNVPVFFKGLLGNANGC
RAAC02340                YRIVYVDVDSQTDGKPLPTVESPQIISGSD--LYEKLQVCP---YPVSEVVSLSGARG-
                        :. : * :***:*:  : * :      :.***:  .      * :      * :. *

ref|ZP_02432977.1|      --YYHLEDKRIAINEG--MSELQTLKTAIHEIAHAKLHDIDLNAPKDEQQPHVDRRTREV
ref|ZP_02085861.1|      --YYHLEDKRIAINEG--MSELQTLKTAIHEIAHAKLHDIDLNAPKDEQQPHVDRRTREV
ref|ZP_02207628.1|      --YYHLEDKRIAINEG--MSELQTLKTAIHEIAHAKLHDIDLNAPKDEQQPHIDRRTREV
ref|ZP_02429891.1|      --YYHLEDKRIALDEG--MSELQTLKTLIHEIAHAKLHDIDLNAPLEDLENRDRRTREV
ref|ZP_01731985.1|      CRYDALSGHPIEIVDPLLPKQHQAKTLCHEIGHSLLSHR-----TQYNNHIPRSKAEL
RAAC02340                -EFD--LETNQIKIVQT--LPEAHKAKTLIHEWAHGLLHTN-----GPLN--PAHKPLIEL
                        :  * . :  * :      :. :  **  ** .*  **      :      :  * :

ref|ZP_02432977.1|      EAESVAYTVQCQHYGLDTSDDYSFGYVAGWSSGRELS--ELKSSLETIRSAAAEIINSIDENL
ref|ZP_02085861.1|      EAESVAYTVQCQHYGLDTSDDYSFGYVAGWSSGRELS--ELKSSLETIRSAAAEIINSIDENL
ref|ZP_02207628.1|      EAESVAYTVQCQHYGLDTSDDYSFGYVAGWSSGRELS--ELKSSLETIRSAAAEIINSIDENL
ref|ZP_02429891.1|      QAESIAYTVQCQHYGLDTSDDYSFGYVAGWSAGRELA--ELKSSLETIRSTAAEIINSIDEHI
ref|ZP_01731985.1|      EAESVAFIVLNYFGIDSRDYSFPYVAGWQQGEDALENLRQSGMRIQKAANKVI-----
RAAC02340                EAESTAFVVS HALGLD TTDYSFAYIAGWS--GKEAVNALKACGTRIQQAA NSILLALEDQA
                        :*** * :  * :  *.*: ***** * :***. * :      * :  .      * :. * :. :

ref|ZP_02432977.1|      AELQKA-
ref|ZP_02085861.1|      AELQKA-
ref|ZP_02207628.1|      AELQKA-
ref|ZP_02429891.1|      AELQKA-
ref|ZP_01731985.1|      -----
RAAC02340                QDFKTAM

```


SUBSTITUTE SHEET (RULE 26)

FIG. 43

```

ref|ZP_01273840.1|      -----LNNVLRAISSKTTIPILTGLKMVVN--EDNIVLTGSNSDITIESVIN
ref|YP_001270615.1|      -----LNNVLRAISSKTTIPILTGLKMVVN--EDNIVLTGSNSDITIESVIN
ref|NP_783868.1|        MKFTINRSFAFIKELNNVQRAISSKTTIPILTGLKLDVN--TDAITLTGSDADISIEETIP
ref|YP_805310.1|        MKFTITRSTFLKTLNDVSRRAISTKTTIPILTGLKIVLN--DSGLILTGSDADISIESRIN
ref|YP_803555.1|        MKFTIARNTFIKLNNAVQRAISSKTTIPILTGLKIEAE--EDRLILTGSNTDISIETTIS
RAAC02342                MKATFRHGDLEKILRQLLRVPVNSTNKAVLKHLVLEAKREEDRVDFYASSEDMSLRRLC
                        * . : * . . . . * . : * . : : . : : . * . * : : . :

ref|ZP_01273840.1|      ANDADNDLTIEDTGAIVLPARFFSDIVKKLPDKKVTIEVTSGFQADITSGSAKFQINGQD
ref|YP_001270615.1|      ANDADNDLTIEETGAIVLPARFFSDIVKKLPDKKVTIEVTSGFQADITSGSAKFQINGQD
ref|NP_783868.1|        ASDDNNTLVVEDAGSIVLPARFFSEIVKKLPEDTMTVNVVDGFGQTQITSGAASFTINGLD
ref|YP_805310.1|        ATDENNDLQIGSTGEIVLPARFFSEIVKRLPESTMTELVKDNFQTVITSGASEFTINGLD
ref|YP_803555.1|        TNDEEAHLNIESTGAVVLTARFFGDIVKKLPDNTMLLEVKGDFQTTITSGASEFKINGLD
RAAC02342                VEATDSSVEIARSGSCILLPAKELYEVIKRRANGPITVETTTD--RTVITFGKTKFELVGLQ
                        . : : : : * : * . * : : : : * . : : * * * : * : * :

ref|ZP_01273840.1|      AENFPHLPEIETNKS-VTLPNDILKEVIRQTVIAVSKQESRPILAGVHMTLKDGLTAVA
ref|YP_001270615.1|      AENFPHLPEIETNKS-VTLPNDILKEVIRQTVIAVSKQESRPILAGVHMTLKDGLTAVA
ref|NP_783868.1|        PENYPHLPEIDTTNT-ITLAGDVLKELIGQTVIAVSNQESRPILTGPHFILANGEFLAVA
ref|YP_805310.1|        ANNYPRLPEITADAA-LSVSADVLRLQINQTVIAVSNQESRPILTGPHILTITGDQLVAVA
ref|YP_803555.1|        ANNFPHLPEVDSQTT-ITLKADVSEMINETVIAVSNQESRPILTGPHIFTIEGNKLSAVA
RAAC02342                PRLFPQFYGDADDEVTTATVLAHLYRLIRRTSYAAGKSQTRPILTGVLTLADGHLSSAVA
                        .. : : : : : : : : : : * * * . . . . . : : : : : * * *

ref|ZP_01273840.1|      TDSHRLAQRKVVLENIDNGIDFDVIIIPGKSMEELSGMI--SDVHEDVQM QVTENQVLFIF
ref|YP_001270615.1|      TDSHRLAQRKVVLENIDNGIDFDVIIIPGKSMEELSGMI--SDVHEDVQM QVTENQVLFIF
ref|NP_783868.1|        TDSHRLSQRRIKLP EANN-ANYDVIIIPGKSLTELSRMI--GDNNPDVQMRLSENQVLFVL
ref|YP_805310.1|        TDSHRLAQRSLTLPTASA-SDYDIIIPGKSLTELSRML--SDDVEKIEIRIAENQVLFVF
ref|YP_803555.1|        TDSHRLSQRVINLPEPIK-MPEDIIIPGKSLLTELSRML--DDEIDTITLQISENQALFTF
RAAC02342                TDALRLAQYTVSCEDVNG-EDRQLVIPA VLLD TLAALPASDDDEVTF TLTGTSSTVSW
                        ** : ** : * : : : : * : : : : . . .

ref|ZP_01273840.1|      GNTHFYSSL--LEGNYPETSQ LIPQTADTTVELEAGTFLSSIERASLLSHESRNDVVKLS
ref|YP_001270615.1|      GNTHFYSSL--LEGNYPETSQ LIPQTADTTVELEAGTFLSSIERASLLSHESRNDVVKLS
ref|NP_783868.1|        GNTSFYSSL--LEGNYPDTSR LIPKESNTTVEISAPALSAAIERASLLSHESRNNVVRFS
ref|YP_805310.1|        GQTAFYSSL--LEGNYPDTSR LIPKESNTTVEISAPALSAAIERASLLSHESRNNVVRFS
ref|YP_803555.1|        DETLFYSSL--LEGLYPDTSR LIPKESSTEMEF EAPELLASIERASLLSHAGRN NVVKLT
RAAC02342                GDDAFHMA LRGLEGTYPD TARLIPERTAH RVVVERQALLTACERVAILSEADHQ---RAE
                        . : * : * * * * : : : : : : : : : : : : : :

ref|ZP_01273840.1|      LKPSENLVRI SGDSPDIGTV EEEVVTSALD GNDLEISFNP NYMKDALRSFG-QATIKISF
ref|YP_001270615.1|      LKPSENLVRI SGDSPDIGTV EEEVVTSALD GNDLEISFNP NYMKDALRSFG-QATIKISF
ref|NP_783868.1|        VNPTDKTITIFGNSPDVGEVTEQLQPTDL SGDELEISFNP DYMKDALRSFG-QAMIKISF
ref|YP_805310.1|        LNIADQKATIYGNSPDVGNVEEVLSFNKLSGEDLEISFNP DYMKDALQGFQ-QTAIEVDF
ref|YP_803555.1|        INAANKQAIISGDSPEVGNVEEVVTKDIVGEDLEISFNP DYMKDALRSFG-HTSIKMAF
RAAC02342                FRFTMSGLTVSATSTQYGHAEETLETIQGT KDDVELLCNVHYWIAALKALD GIAQVEIGL
                        .. : . : * . : * . : : : * : * * : : : : : : :

ref|ZP_01273840.1|      TSPLRPFTLVPTEDQ-----
ref|YP_001270615.1|      TSPLRPFTLVPTEDQ-----
ref|NP_783868.1|        TMALRPFTLVPT E EGFENFIQLI-----
ref|YP_805310.1|        TAPLRPFTLVPTEDK-----
ref|YP_803555.1|        TSALRPFTLVPTEDQ-----
RAAC02342                NGPLQPCLLRPVGE EGVGLIATVARSSAPTETKRQAS
                        . . * : * * . :

```

45/228

FIG. 44

```
ref|YP_429218.1|-----VQRALEMAVRAAAMQVTPDSQAAGHPR
ref|YP_001112194.1|-----AAKSAVMQVTKESQADGRPR
ref|YP_001112320.1|-----ALALAARSAAGQVVPVSNASNNPR
RAAC02293MKWIIMALSLVVLQFWTHEQVLNLEMETVVYHRIVNAMVLASQDAVEDVIPSSSTANGEPI
ref|ZP_01171099.1|-----LNFTFDSTATRGLRDSLELATHDAGLQIYNEELVNGNIV
ref|ZP_02598174.1|-----LTLNMNIATYAKSSKYLKEDLEVAVHDASLELKKDELANGKLV
                                * : * :: . . .

ref|YP_429218.1|INIVAANIAFRRELASNGLDANTLAPLKGSAMKTRP-----
ref|YP_001112194.1|INTTNAQAVFQQQLAKNLGLDETTLNPLSGS-----
ref|YP_001112320.1|VHADNAHSTFKKILAKNLGLDEVTLNPLPGS-----
RAAC02293FDQTEAAQTFRMTLANNLGLDPNTLQPLSNSTFHVAPKILDEEFYDWSNT-TFPYHYVNS
ref|ZP_01171099.1|FDQSEARRVLEESLKKNLLLDAS-LQPKPDSFFQDTVQIKFVDYLDNNTPEFPTNYINE
ref|ZP_02598174.1|FDPIKAKQTFKESLKDNTGFQEG-TDYK-----ILEFQVLDQSNS-KFPVKYKAN
.. * ... * .* ::

ref|YP_429218.1|-----
ref|YP_001112194.1|-----
ref|YP_001112320.1|-----
RAAC02293AYGINETLDAPSMVVVVQFTMPSYAANVPPFTITVPMVQSYAGS
ref|ZP_01171099.1|QYDIVDTVDGPAIVVVLETTGPRY-----
ref|ZP_02598174.1|TLKFQDTFQNP TLVAIIETTTKKY-----
```

FIG. 45

```

ref|ZP_02327783.1|-----IKFKTLFLHNFKSHRDLE
ref|ZP_02846176.1|-----IRLIQEQLINYAGIKDRT
ref|ZP_00235040.1|SCRESIGNIFICANCEE-IDENTITIESPSITIVESGAPSMKIVFKQLTLENFKNHKNLV
ref|NP_623604.1|-----MIIKSITLKNFKSHKNTI
RAAC02346-----MKILSIQLENFRSFTEAS
ref|ZP_02082031.1|-----MKIKGVTFDFRNHKAPQ
: : : : .

ref|ZP_02327783.1|INFG-ELTKITGENTKGKSSILEAIPWLFYSVDMLG-----SKSDPTPINYEYDHTLVKL
ref|ZP_02846176.1|ITFG-NVTNLSGKNGEGKSSIGGAPIWILFGKDLYGNDYTKDKYSPRPSNYKYDRVYASI
ref|ZP_00235040.1|VDYE-QVTQISGKNGFGKTSIGEAVTWLLYGTDLLG-----TKIEPQPLG-TEEEVHVS
ref|NP_623604.1|INFNDKNTVIYGDNGTGKTSIGEAIWCLTGANLFG----TENVTNKLVITIGKNEMSVTL
RAAC02346FQFH-DITVISGHNGAGKSTLAEAVVWCLFGTDIAG----RQKQDEKLMRLGKRMVTV
ref|ZP_02082031.1|SYTFGDISYITGHNGTGKTTMAHGICYALYGVSYG-----EQKIERLMNEKATGTQVQL
. : : *. * **:: . : : . * : :

ref|ZP_02327783.1|HFA-VDGKDILLGRGIE-KGKATYYINEVPAKAKKEYEELVKSLFDKDLFLSLYNPSYYFT
ref|ZP_02846176.1|LLS-IDGTEYKFAREIDEKKNNNFVNDIPKSATDFSAAVVALITQDEFMSLYFPAYFFG
ref|ZP_00235040.1|LIN-ADGKDLLLTKKQK--KTAKYAINVPRKATEFADMIDSLFEKNLFYSLSPGYFFS
ref|NP_623604.1|VIE-KDGKEYEITRSK--KNEIETINGVKSTQIDLYTQFVQDKDIFFTVFNPLYFTT
RAAC02346TWL-IHGKSVVISRTRASRQGSTLLVNGKRAQPGQIEGWFGTVQE---FLSVFVPGYFSS
ref|ZP_02082031.1|DFTDQNGTTHTLIRNRS-GDKTALLLDGYTVRQGDIDR---IFCDKETFLSMFNPTYLTE
. *. : : : : * : : * *

ref|ZP_02327783.1|LKWNEQRELLRLRYVSAPANKEVFQAQLPKQQAELG-----
ref|ZP_02846176.1|LKWTEQRELLMKGATAPLTKAVLKEIEKLYADKLE-----
ref|ZP_00235040.1|QHWQTQREQLLSYVTEPGEKEVLEEMNEIDRTLLS-----
ref|NP_623604.1|LAPRDAKAILYKVLPEVSNEEVFEKLSPEVAETLKNNGFNNTFIEKQREILKDWDGL
RAAC02346LEPKEAKTVLSRCVPDIPKEDVLARMTSVHASMLA-----
ref|ZP_02082031.1|RLGEKGRPLILKYLQPVSANAVLEQMSETYREYLDGIDLNTFPPEAKLKEFRG-----
: : : * : : *

ref|ZP_02327783.1|-----ELVKKHSLADLEKIHNRDNKNKKDKAYIAAQSRKTTLI-----
ref|ZP_02846176.1|-----PLLKKNLSLDLESKHKQDKTRLEKAHTEAGGATKKLR-----
ref|ZP_00235040.1|-----TELNKHLLDDLEAVNRETFKNSDKQYERASERVTLK-----
ref|NP_623604.1|LKSEGNIILRQIINAKIPEMKSFDETQLKELEKKLMEFQKNQNIIEKELAILK-----
RAAC02346-----RDQFVMGLDSIEFAMQKVRDEIKECEAERLRLEG-----
ref|ZP_02082031.1|-----AVRQAEQEAYLQGNIDSFEEASRTAEQKLSELYADKAAIEKKRKVL
. : : . : :

ref|ZP_02327783.1|--EQIQLLPQSPPNLKN--IQDDYDRFLAEIKVIDSMLPKADENNGK--IIKLESQIEH
ref|ZP_02846176.1|--EMLERMPAVEGELTE--LETQAEALRVEIAKEDAIVAAAWQTNTS--YRELESALFY
ref|ZP_00235040.1|--EQLSNASEVNMDIKE--ITEQKDALIAERTAIELKEDKNVQLRND--YADAEQKINA
ref|NP_623604.1|--EQYKNPPIEKPLQKDTSLQKTREELLQEQYKQIQEQIKNLKPQYIT--CNKCGNKIDV
RAAC02346--QCQAYQAVLRRGEPQ----PYVPSVTDEERARYEAAKRELMELEAS--QGNRKERLRD
ref|ZP_02082031.1|SDKQFEGIEVEDLSIQRNMLLEKLSVVPKGENPKVTQLQEKIEELRHKPYVSKYTQAMAE
: : : * : :

ref|ZP_02327783.1|LKEQRDQIKEQIRELHN--EPIDD--RCRACGQTLQGEAKAAAEADQRRVDQLIEEFNN
ref|ZP_02846176.1|TQHEVDSSKAAPALKD--EAIND--TCRTCAQPLQEDAIAKAVIADKEKRIAAYRSKHSN
ref|ZP_00235040.1|LKERILRKREEALNVRE--QKIEE--NCEYCGQTLQGDSIEHAIQYRKEHYNRLVTAGKI
ref|NP_623604.1|TAKEKQLLLAKLQEIKE--KGTKT--AAELKAVIEENEKAEKEFKVKVNYRKALEIKIK
RAAC02346LYARRDSLGRAFRALRDSLPQADT--HCHTCGQPLPEDQAARILQETIAQKRKASLAKMKE
ref|ZP_02082031.1|NAAEVKNLSERYKALVNRVKGLRPGAQCPACLMRITEQNLPEVRNRLAELKSLAEQGKE
. : : . : :

ref|ZP_02327783.1|VIT-----KRKELEAHLTQIEHI
ref|ZP_02846176.1|LQS-----IRDEAKAALAAAKWV
ref|ZP_00235040.1|MVE-----ELEASKARLAKLENP
ref|NP_623604.1|RLE-----NALNTNANKEQIQQI
RAAC02346LYD-----EGNQVQAEIAKLEAM
ref|ZP_02082031.1|RVAQGKELAGMDYQSKTVFEQFKADDLKKLYEELKGLKTEAAGKTGDGELRAALDKVEEL
. * :

```

47/228

```

ref|ZP_02327783.1|-----DVSKQREEVRELEVRLDQLDEVISNHKSREQLQG--QIDHAETDEAAILTS
ref|ZP_02846176.1|-----DVTERQANVRALEDQRDLITEKIRAHKDRARFEV--ELQQSEETEATTLIS
ref|ZP_00235040.1|-----EKNFDRVKYKEIDEKILELSGYIQSVGQTEKLHK--QIADAELEQQRIKQ
ref|NP_623604.1|-----MQQIERLRFEERE-VISHNESVKALLRQKEEAKK--KLKEVEEDIKLAQQ
RAAC02346PDYDAPHPELVEFVQTMEARLKDEHYREVAYAAQLRAYEQAKG--HFTQAQEDLQATAEH
ref|ZP_02082031.1|QK-----YGNLDEDEYTELNCLSAELAGIEAQIQAVQDMCDERKLEDAYAQQEVCKGQ
. . . . .

ref|ZP_02327783.1|LNESIFIIDAIKAFAAKEAEMMADKVQALFTTSLRLFKTNKTDGEIKPDFEIMDGKPY
ref|ZP_02846176.1|LRESTLIIDAIKAYKAKEAELQAAEQSNFTRLVRLFKYVKSNDAYEPDFSIQMDGKDY
ref|ZP_00235040.1|RNKSQSIVEAIKRFKAKRSDLMVGKVNGLFENITIKLYEVLKN-GTEKPTFEVEWQQKPY
ref|NP_623604.1|INEVKTLIEYAKAFNAKKLELEATEINKYLNKVSLQLWKIVQSTGEIKDDFKILYDGKEF
RAAC02346LEGLKQRLQALQEFREFEYLRAQHEQLNGLFRHVAIHLMDVNKETGEVREAFRIEWKGRPY
ref|ZP_02082031.1|ILKYKNVISALGEFICKRTEIAVSSLQ--MPNVKIKLFDVVRSTGEVINVFKFTYKGRDY
:. : : : : * . . . :

ref|ZP_02327783.1|RKLSLSESI RAGLELRDVLSSQSGIIAPCMVDNAESITQFKQP--NGQLIVSRV-----
ref|ZP_02846176.1|AFLSTGEKIAAGLELAEVLHKQSGIIAPVFIDN-----
ref|ZP_00235040.1|SKLSTAEEKI IAGIEFANALS LKAETI IPLFADNAESVIELPKP--TGQLITATVKKTKFT
ref|NP_623604.1|NILSYSERIKAGLEIANLIMGLTKIKFP I FIDNAESI-----
RAAC02346RLLSYSEKIRCDIEIGRALAQAKGEAMPVYVDNAESVQRIMDETFSGQVIAAYVADGPLT
ref|ZP_02082031.1|STLSLSEKTLAGIEITAMIRKITGIDCPICVDNTESIAAFNSVSMPSQTLLLRFVKGQPL
** .* .:.* : * **

ref|ZP_02327783.1|-----
ref|ZP_02846176.1|-----
ref|ZP_00235040.1|VKGVSE----
ref|NP_623604.1|-----
RAAC02346VSKMPEAQGA
ref|ZP_02082031.1|TVQSRNRA--

```

FIG. 46

```
ref|YP_113896.1|-----PWLNVNVAKG-DLKLVGVEA
RAAC02347MRPSTLYYNGVQMNPFVRGGPSMEDKKVLKTVIRNGVTFDNYPVYVSEAYG-DSYLMKHEE
ref|XP_001383704.2|-----NKSVPFSFIIRKPINMNGYDYSVFTNYPTIDLQHGEH
ref|XP_001664270.1|-----SDSSG-DDHLMKPEE
ref|XP_761114.1|-----KEIADRGAKADKEWTELFKAYG-EKYPNEHAE
ref|XP_001015776.2|-----VQRNGDLESKKISKVVLKSG-ITFLDIQIK

VTSEQAACRTAEWEKAADRAPAGLIG---PTQLDLPATAPAEERLMS-----
LAEQIASCIPQAWRKAA-RFDCNLIA---EFQDDMDETSEDEQRLLSKFERWMSRHN
LPPAIKESQIKSWESAERISTNVIFGNSSESEDDIDELDEDEHTEASRSSKVNSRKN
LDLRSNFFDKIESTKRQSGEPAPQFD--CNAGMRLSDSSDDDDQAEFIPSSIQK--
IARRIAGKLPEGWEKS-----
VQRDLKLGYTENFRLEF-REGVEMFE-----NDIPFLQNNCYLLASMGEEYDSTS-
:
```

FIG. 47

```

ref|YP_503850.1|      -KISKKNYDLVFKEAFSIFDNRSALFLGID-LPPIISFLVTEIPEVETTDMMDLNFRLE
ref|YP_517477.1|      --ISYHNNDILMKVLAEQFKNKTLDFVGIK-TAKIKDLIPSVHPAVEANETRNDIIFLLE
ref|YP_001568284.1|  -----IISVKPIDIPVINVSNQNPDFVFELE
ref|NP_783815.1|      -----DTKIIAPANTELKTIDIKNTFTDYTFYTE
ref|YP_149134.1|      --ISHHAKDILFKSLSALYQNQALDVYGLHGLPRIKALLPNEFPSVRADERRADTVFLLE
RAAC02289             MRIARSGNDIVAKHLTNALPGEVLSVIGIH-DAHVVRLPTLPTVEVRQFTDIMLELA
                        :               :               *       :

ref|YP_503850.1|      DGSILHLEEMNLSKRD-LIRFAHYDLRLFRYYD-----APVHTVVLTPADSSSGTKVL
ref|YP_517477.1|      DDTLLHLEFQTTAGEQD-LKRFLYYDARLVRRQE-----RKVHTTVIYSGRIEQARERL
ref|YP_001568284.1|  DNSLLHLEFQTTWKKAD-LLRFAQYDIALYQKER-----RRINTVVMYSGKYESAESL
ref|NP_783815.1|      NDDYLHFEFQTTNKEED-INRFLFYDASLFYKYG-----KKVNTLVVYSSDIKSKTKV
ref|YP_149134.1|      DDSILLLEYESNERFLDNHLKYLDYACRILHTYYQQEKIRIPRIVVIYTSDVTTARERL
RAAC02289             DGRLLHLEFQTT-REPN-LYRFGTYDWAMAERYK-----RPIRTVILYTRDVTEAPSEL
                        :.  * : * : .      :  ::  *      :      :.  ::.  .      :

ref|YP_503850.1|      DIGSLQYNVLQIVLSDRDGDALLSQMRAALEKGEPVNE---LELIFLPLMKSKLTKIELL
ref|YP_517477.1|      ECGSILYQVENIYMKHYNQDQEYNRLKHKIDNHQLLSETDTLKLIFLPLMKSEQKEEELA
ref|YP_001568284.1|  DMGSNKYKVQQIFMIKYDGIKRYEEIKEKIEKEEELTDKDLMDLVFLPLMRNEKSEEEVT
ref|NP_783815.1|      DAGSLKYEIKAFYMSSLNGDEEYNNLTKIDKGEDLTKEEILSLTFIPLMDSKEDKSTRT
ref|YP_149134.1|      DAGDVFLSSKAVLLGEFNGDAIFHAIEEKVHNGEPLTPEETMKLILVPLMHTRFDRQTM
RAAC02289             DAGSMRYAVENVYLGHMDGDGALETVKRHLAAHEWTEEDRVRLAFAFHMRFRERTREEAF
                        : * .      . :      : *      :.  :      :      :.  :      .

ref|YP_503850.1|      RRTIDLEKELPEKDLRNKVRELTILADKIVDQKILDELWEELRMFKVVKYAEKGMKEG
ref|YP_517477.1|      IQAAELAKAAPDEKTKLFAIAALIVITDKIMSESNNRKLLEVLKMTQIEQWIREEGRQE-
ref|YP_001568284.1|  KDVFEIAIKIPDEDKKEAVIGSLLGFSDNYVRDEYINELKEVIRMTKIGTSLFEEGVE--
ref|NP_783815.1|      IKSIELA EKMEENNTKLQCITLLYAFLEKFGDAKSKKKFKEVFSMTTEIGRMIVEESIEKG
ref|YP_149134.1|      EKTIELAKAIGDEPKQLHIIAGVLTATDKFIDRSYAEKVKEWIKMNKVFRLVVEELEQEK
RAAC02289             GEIVEVVQRPDPVHEQNYLAALILGFSGRVMADEQKEQLRRVLEMTDLLRELEREFEEKG
                        ::      :      :      .      .      :.  . : * :.      :.  :

ref|YP_503850.1|      -----LEKGLEKGIKKGMKEGKKQERETVAKNMLSLGIEDELI IKATGLDQSIIDKL
ref|YP_517477.1|      -----GELKGRRDEKRETAQTMLNLGMSPELIAKATKLPLEEILEM
ref|YP_001568284.1|  -----EGEKELTIKILNKRFRRLTEEIKDRIREADKKTIDYI
ref|NP_783815.1|      RAEGIKKGIEEGIKGRTEGKTEGKSEILIKQLIKKF-KKVPEEYIQIKITLSIDTIDII
ref|YP_149134.1|      -----EEMLKVMQEKQAVQRAIQEKEQAVQRVIQEKEQAVQRA
RAAC02289             -----IQKGLQOGLQOQELQKAREIAHRLRLKRGASVQEVVEITGLSSKDVEEI
                        .               :      .      .      :

ref|YP_503850.1|      KKSL-----
ref|YP_517477.1|      -----
ref|YP_001568284.1|  GDNL-----
ref|NP_783815.1|      ALEIFDMEDIK
ref|YP_149134.1|      IQ-----
RAAC02289             RQNLH-----

```

FIG. 48

ref NP_607899.1	-----
ref NP_269831.1	SINGLE-STRANDDNA-BINDINGPRTEINSTREPTCCCUSPYGENESMGASREFYPSIN
gb AAF98351.1	-----
ref ZP_01171110.1	-----
ref XP_414088.2	-----
RAAC02348	-----
ref NP_607899.1	-----
ref NP_269831.1	GLE-STRANDDNA-BINDINGPRTEINSTREPTCCCUSPYGENESMGASREFYPSINGLE
gb AAF98351.1	-----
ref ZP_01171110.1	-----
ref XP_414088.2	-----
RAAC02348	-----
ref NP_607899.1	-----
ref NP_269831.1	-STRANDDNA-BINDINGPRTEINSTREPTCCCUSPYGENESMGASREFYPSINGLE-ST
gb AAF98351.1	-----
ref ZP_01171110.1	-----
ref XP_414088.2	-----
RAAC02348	-----
ref NP_607899.1	-----
ref NP_269831.1	RANDDNA-BINDINGPRTEINSTREPTCCCUSPYGENESMGASSPPSSBSTRPSINGLE-
gb AAF98351.1	-----
ref ZP_01171110.1	-----
ref XP_414088.2	-----
RAAC02348	-----
ref NP_607899.1	-----
ref NP_269831.1	STRANDEDDNA-BINDINGPRTEINSSBHELIX-DESTABILIZINGPRTEINSPSSBS
gb AAF98351.1	-----
ref ZP_01171110.1	-----
ref XP_414088.2	-----
RAAC02348	-----
ref NP_607899.1	-----
ref NP_269831.1	TRPSINGLE-STRANDEDDNA-BINDINGPRTEINSSBHELIX-DESTABILIZINGPRT
gb AAF98351.1	-----
ref ZP_01171110.1	-----
ref XP_414088.2	-----
RAAC02348	-----
ref NP_607899.1	-----
ref NP_269831.1	EINSPQXASSBSTRPSINGLE-STRANDEDDNA-BINDINGPRTEINSSBHELIX-DEST
gb AAF98351.1	-----
ref ZP_01171110.1	-----
ref XP_414088.2	-----
RAAC02348	-----
ref NP_607899.1	-----
ref NP_269831.1	ABILIZINGPRTEINGBAAKPUTATIVESINGLESTRANDBINDINGPRTEIN-PHAGEA
gb AAF98351.1	-----
ref ZP_01171110.1	-----
ref XP_414088.2	-----
RAAC02348	-----

ref NP_607899.1	-----
ref NP_269831.1	SSCIATEDSTREPTCCCUSPYGENESMGASGBAAMPUTATIVESINGLESTRANDDNABI
gb AAF98351.1	-----
ref ZP_01171110.1	-----
ref XP_414088.2	-----
RAAC02348	-----
ref NP_607899.1	-----
ref NP_269831.1	NDINGPRTEINSTREPTCCCUSPYGENESMGASDBBACPUTATIVESINGLESTRANDBI
gb AAF98351.1	-----
ref ZP_01171110.1	-----
ref XP_414088.2	-----
RAAC02348	-----
ref NP_607899.1	-----
ref NP_269831.1	NDINGPRTEINPHAGEASSCIATEDSTREPTCCCUSPYGENESSI-GBAATPHAGESIN
gb AAF98351.1	-----
ref ZP_01171110.1	-----
ref XP_414088.2	-----
RAAC02348	-----
ref NP_607899.1	-----
ref NP_269831.1	GLE-STRANDDNABINDINGPRTEINSTREPTCCCUSPYGENESMGASGBAAXPHAGESI
gb AAF98351.1	-----
ref ZP_01171110.1	-----
ref XP_414088.2	-----
RAAC02348	-----
ref NP_607899.1	-----
ref NP_269831.1	NGLE-STRANDDNABINDINGPRTEINSTREPTCCCUSPYGENESMGASGBAAZPHAGES
gb AAF98351.1	-----
ref ZP_01171110.1	-----
ref XP_414088.2	-----
RAAC02348	-----
ref NP_607899.1	-----
ref NP_269831.1	INGLE-STRANDDNABINDINGPRTEINSTREPTCCCUSPYGENESMGASGBABFSINGL
gb AAF98351.1	-----
ref ZP_01171110.1	-----
ref XP_414088.2	-----
RAAC02348	-----
ref NP_607899.1	-----
ref NP_269831.1	ESTRANDBINDINGPRTEINSTREPTCCCUSPYGENESMGASGBABFSINGLESTRANDB
gb AAF98351.1	-----
ref ZP_01171110.1	-----
ref XP_414088.2	-----
RAAC02348	-----
ref NP_607899.1	-----
ref NP_269831.1	INDINGPRTEINSTREPTCCCUSPYGENESMGASGBABFSINGLESTRANDBINDINGPR
gb AAF98351.1	-----
ref ZP_01171110.1	-----
ref XP_414088.2	-----
RAAC02348	-----
ref NP_607899.1	-----
ref NP_269831.1	TEINSTREPTCCCUSPYGENESMGASGBABFSINGLESTRANDBINDINGPRTEINSTRE
gb AAF98351.1	-----
ref ZP_01171110.1	-----
ref XP_414088.2	-----
RAAC02348	-----

52/228

```
ref|NP_607899.1|-----GAPSVTEV
ref|NP_269831.1|PTCCCUSPYGENESMGASSCRESIGNIFICANCEIDENTITIESPSITIVESGAPSVTEV
gb|AAF98351.1|-----VTEV
ref|ZP_01171110.1|-----VTEV
ref|XP_414088.2|-----
RAAC02348-----MAEV
```

```
ref|NP_607899.1|VADNFQMLESRATREGGSTGSFNGGFNNNTSSSNSYSAPAQQTPNFGRDDSPFGNSNPMD
ref|NP_269831.1|VADNFQMLESRATREGGSTGSFNGGFNNNTSSSNSYSAPAQQTPNFGRDDSPFGNSNPMD
gb|AAF98351.1|VADNFQMLESRATREGGSTGSFNGGFNNNTSSSNSYSAPAQQTPNFGRDDSPFGNSNPMD
ref|ZP_01171110.1|IARYVQFLDSKKNNQSQEQRPE--NNTPSGRNSNNNPYAD-PYFEGEP-----
ref|XP_414088.2|-----KYAEEGEGKQSEEEAKASA-----KIKKNSAPLPAEQ-P-----
RAAC02348VAEGVRFLDRGEGNQSASASAAP-----PVKNNTSPAPSKQ-PDFGPYDDPFADDLPPG
          :  :      . : . .          *:  *  :  *
```

```
ref|NP_607899.1|ISDDDLPF
ref|NP_269831.1|ISDDDLPF
gb|AAF98351.1|ISDDDLPF
ref|ZP_01171110.1|ISDDDLPF
ref|XP_414088.2|-----
RAAC02348FEED-LPF
```

[illegible]

PGAEKILVLLGLTS-EYEIIEKIENYEKGIFA-YTVRCILS--KNGKKVTEGLGSCNSKE
PGAEKILMLMGLTS-EYNIIEKIEDYDKGIFA-YTIKCILR--KNGQKITEGVGSCNSKE
PGGEKICMMFGLNP-EYEFLERTEDYKDGFFA-YNIKCTLY--RNGNPVSQGVGNCNSME
PGAEKLTDFGFSK-QVEILNRNENWENGLFN-YEVKVS LINKRNQLIEAEGIGCCNSKE
SGAEKLCDIFGFSK-KIEILNRVEDWEKALFH-YEIKTILINKKTGLIEAEGIGSCNNRE
PGAELKCDVYGLSAGEAHIEFTRDDTKTPIYISYRVSPLISRIDGKIMVGVSANSWE
* ** *

DKYRWRWVSEKDLPPNVDKDMLSKTNEYG---QKLYRIENDEIFTQANTILKIAKKRAQ
DKYRWRWVKEDDLPMGIDKDAVSKVDNYG---HTKYKVENDDICSQANTILKMAKKRAQ
KKYRY--INVDAPVDGIDPSTVEKVTTRYG---TVKYKIPNPHIADLVNTILKMAKKRAF
KRYRS-----QDPYN-----TVNTVLKMAKKRAL
RKYKN-----QDGYS-----IVNNILKMAKKRAF
KKYRWRWVTERELPDLPKETLKQATFDGKYGSYIKYRIPNDEIDDLNTLLKMAKKRAL
*: *:****

```

IDAVLTVAALSEIFTQDVEDMQEF---LQNEQLETMKAEEAVNVKVTFGKHKGKTLGEIY
VDATLTVASLSEIFTQDIEDMAQF---QERE-
IDAVLQVASLSDVFTQDLEEMQE-----
IDAVLSATRSSGLFTQDVEDLDISESWVRSPQPQATKTVRGVKPD-SLSVANITQPQLRK
IDAVLSATRSSGLFTQDMED-----
IDAVLSATRSSLALTQDIEDLED---LTIRQPESVDRRQRTQPDRSTGSPDISQAQKNL
:*. * .: :*****:

```

SQAPDYVQWLAQNARNNDVLRKAANMVMNGKGNESQQEAQRSEDL-----

 VQALAEAIGMTPHQLQALIQDMFRVNHTTRL SKNQASALIQHLLS-----

 ILRKAQERGMDTEALNAFVLQATGKGELDSLTKOEASKLIDLLTSGKPRER

54/228

FIG. 50

emb CAJ73677.1	-----MTNYNFGEILLKFPYSDGKGDAKRPVVVLAQTDL-EDIVTAKVTSM
emb CAJ73386.1	-----GEIILVRFPHDTLQDISKRPALILYDSGD-QDILIARITTO
ref YP_001655174.1	-----PQKTKVYMTI IKGGEFWVAKILFTDGTSSKKRPVLILWVDG--DDVVVTAVTSA
RAAC03270	MPFTMKPSDTYWNLMNVRRGQIYLMKVQFSDLSGEKIRPVVIGTDRVDDDDVTVVFTSS
emb CAJ74660.1	-----MQYPFSGVSVQKVRPAVVVNAPHVSHDLLVVPLTSK
ref ZP_00516046.1	-----MANYWKNDVILVRYPFSDLSNSKIRPAIIVSSPHTSQDVLITPLTSK
	: .:. **.:. .*: . :*:
emb CAJ73677.1	-EQRGEYDIAIDAWKDVGLLYPSVVRIDKLATLSKQKRVKRFGLVGSYKPEIISRIK--
emb CAJ73386.1	-EYTTGTDYKIVEWKSCGLLVESFIRISKQATIEKKYVIKPLGTLAEELNAVKSIIKN-
ref YP_001655174.1	-MPRTQTDILLKDWANSGLRVPSTVRLSRLDCLEKNLLIAKLGRISNEDAVYLLRAWDDY
RAAC03270	-PPRLRYDVQITEWSVAGLLKPSTVRASKFLT VHKGRFLKPLGTLTSDLQAVMTAVRSY
emb CAJ74660.1	TTSLLKGEFVLTDWTLAGLNVVTAVKRG-IYTVHQSLI IKRLGKLSHADSQKIEVALKEW
ref ZP_00516046.1	TQLLLEGEFILSDWKKAGLNVETAVKRG-IYTIYQGLIILKKVGKLVADTTKLEQSLRQW
	: : * ** : :. : : . * :
emb CAJ73677.1	--
emb CAJ73386.1	--
ref YP_001655174.1	I-
RAAC03270	LL
emb CAJ74660.1	L-
ref ZP_00516046.1	L-

55/228

FIG. 51

sp P26545 VE2_HP5B	-STTRSRSTSVGKTRALTSRSRS-RGRSPSTCRRGG---GRSPRRRSRSPSTYSSCTTQR
ref NP_041368.1	-STTRSRSTSLTKTRALTSRSRS-RGRSPPTCRRGG---GRSPRRRSRSPSTSSSCTTQR
ref XP_001371550.1	-SRTRSPARRGGRSRSRTPARRGGRSRSRTPARRGGRSRSRTPARRGRS---RSRTPAR
sp P36786 VE2_HP19	RSRSKSKSRSRSRSLSSNRRS-RSKSR---RKAS-----TTRGRGRGSPTATSDQSSR
RAAC03271	MSRTRSTATNAGKKAARSFFTGH-CAESRSRQKPN---CTRRGHGSATYTPQNGSPPP
ref XP_001131003.1	-TREARARPPPPESAARARVTHK-AADPRPGVSRAT----TAREQRGETHGAPYAPAPP
	: : : . . : : : . . : : .
sp P26545 VE2_HP5B	SQRARAESPTTRGARGSRG--SRGGSRGGRG-----RRRGRSSSSSS-----
ref NP_041368.1	SQRARAESSTTRGARGSRG--SRGGSRGGRG-----RRRGRSSSSSS-----
ref XP_001371550.1	RGRTRSRTPARRGRSRTRSPVRRGRSHSRTP-----QRRGRSGSSSER---
sp P36786 VE2_HP19	SPSATSTTSLRSRGSSRVGRSRG-RSRVG-----RSRGRGKRSR-----
RAAC03271	NSPSLASRRASRGACQKAWTHWHGDVSAKQCSIGFGKNTRGMSASSASRPSPG
ref XP_001131003.1	-PPPRQPKRSERTR-----HRGA-----
	. : * : *

56/228

FIG. 52

```
ref|YP_941474.2| -----
gb|AAZ42391.1| -----
prf||2123261AD -----
ref|NP_498368.2| -----
RAAC02294 MCSRLRDSRLRPRKALCRAADGARGRSSRMSAII FALMFFFILSWNYQPYMDVINGARL
ref|ZP_01171098.1| -----FTPFFVYLDLHR

ref|YP_941474.2| -----VSEAAFRQYSKTKDPETLVQYIATMLSSSSNHVING-ITMR--SGSP-I
gb|AAZ42391.1| -----VSEAAFRQYSKTKDPETLVQYIATMLSSSSNHVIING-ITMR--SGSP-I
prf||2123261AD -----SHMKKVIVFILLY-P--FLLLWNVIISEKYLDFIFGG-FVISYIKRVPWA
ref|NP_498368.2| -----SHMKKVIVFILLY-P--FLLLWNVIISEKYLDFIFGG-FVISYIKRVPWA
RAAC02294 EYLQAVAETAISEAKIKGYFSATDLSNIQQAVATHLGYVPVSEVQVQG-TTLPTRGNP-I
ref|ZP_01171098.1| AVVALVLQQGLKEAAIEGYFSQEIVEGMKDTLEQDYKFERNLIEITPDSSPQTRGEY-L
. : .

ref|YP_941474.2| DLEEVPLAVTF-YVMA-----
gb|AAZ42391.1| DLEEVPLAVTF-YVMA-----
prf||2123261AD SLSKFQIISYVFTFSVTLVTN-----
ref|NP_498368.2| SLSKFQIISYVFTFSVTLVTN-----
RAAC02294 DIQISIPTHINL-FSMAPASNQATLTAYESADSEALNPS
ref|ZP_01171098.1| EVEISVPRGPMFILNIFNQGP-----
... : .
```

57/228

FIG. 53

```

ref|YP_430185.1| -----
emb|CAO81523.1| -----
ref|YP_001233893.1| -----
ref|YP_001603689.1| -----
ref|YP_719187.1| -----
RAAC02353 MNEVNTMENLRAYVNTFAEGSRSTDPIQLSVDELKMAARGWKRAKTVIHVHLLGMQQKL

ref|YP_430185.1| -----PASLVERGYVEDLNAYLGQTLRLRVIELDRSKNKV
emb|CAO81523.1| -----
ref|YP_001233893.1| -----DTFDLYIERYEDKDGS
ref|YP_001603689.1| -----EDRDGSI
ref|YP_719187.1| -----DAVEDGFGET
RAAC02353 VDGTTRDILVGDIGPAKVMLPVDPEYSALEDGEDPMSLTDRWICGIVEDFDLQDEGDSTI

ref|YP_430185.1| VLSRKAILEEEYEKQR-QATWNSLEVQVRKGIVRRLTNFGAFVDLGGVDGLLHVSEISW
emb|CAO81523.1| -----LEEMREQERKETALAKLVGDIVKGKVLRTTFFGAFIDLGGIEGLMHVSEISW
ref|YP_001233893.1| VLSR----EKARREEAWTALERAYEQQARVNGVIHGRVKGKGGFTVDLGGATAFLPGSQVDI
ref|YP_001603689.1| VLSR----EKARREEAWTNLEKAFEGNQRVNGTIYGRVKGKGGFTVDLGGAMAFLPGSQVDI
ref|YP_719187.1| KLSR----EKAVRQESWIDLEKAYEEQATVIGLINGKVKGKGGFTVELNGVRAFLPGSLVDI
RAAC02353 LLNR----KKGLERLRELNAERVS KPGNRAIGVIVGIRRGAYVLNVGGYTALMPKAWYDW
      ::  ..      :      * :      .  ::.*  ::  :  .

ref|YP_430185.1| GRVEHPRDALSEGQEIIEVKILGIDREEGKVS LGRKQLLPNP----WDTAAERYPVG TIVE
emb|CAO81523.1| QHIVRPQDELKKQEIIEVKILDIKGE--KIALSRKVLLDP----FEVAMKELHEGDIIN
ref|YP_001233893.1| RPVRDVTPLMGVQQ--PFQILKMDRARGNIVVSRRAVLEETRAEQRSELIQGLKEGQILD
ref|YP_001603689.1| RPVRDVTPLMGVPQ--PFQILKMDRARGNIVVSRRAVLE-----
ref|YP_719187.1| RPTREADHLLGKEQ--EFKVIKLDQKRNNVVVSRRAVIESENSQEREELANLAEGSEVK
RAAC02353 DDSKRDQGTIGEEFPVQIQPSKVEDR---IVVSRCHLMENP----NVPSSLRFRDRTILR
      :      ..      :      : ..*  ::

ref|YP_430185.1| GKILRLAP-----FGAFVEVEPGIEGLVHISQLA-DRHVDKPEDVVSIGDIIPVKVL
emb|CAO81523.1| CRVLR LHN-----FGAF AELKPGVEGLIPISEMSRNRNISHPRDIVKEGDWVQVQIL
ref|YP_001233893.1| GVVKNITD-----YGAFVDLG-GVDGLLHVTDIA-WRRINHPAEALTI-----
ref|YP_001603689.1| -----
ref|YP_719187.1| GTVKNLTD-----YGAFVDLG-GVDGLLHITDMA-WKRVKHPGEIVSVGDEITVKVL
RAAC02353 ATVAFIRNGLIRA EVYPGFLVSVDPVILRQIPKPGDRITVRILGQNKNGYGMMVDHQPO

ref|YP_430185.1| G----
emb|CAO81523.1| ----
ref|YP_001233893.1| ----
ref|YP_001603689.1| ----
ref|YP_719187.1| KFDKD
RAAC02353 SVG--

```

58/228

FIG. 54

```

ref|YP_001251565.1| -----PVYIPQDVRVGHTFVVGTTRVGKTRLASILINQDIRNGD--
ref|YP_126374.1| -----PVYIPQDVRVGHTFVVGTTRVGKTRLASILINQDIRNGD--
ref|YP_123382.1| -----PVYIPQDVRVGHTFVVGTTRVGKTRLASILINQDIRNGD--
ref|YP_001251158.1| -----VYIPQEVVRVGHTFVVGTTRVGKTRLASILINQDIRNGD--
RAAC02354 MVTHMLTDIKLERHAVPYAPVWIPGRERTGNVLVEGGPGTGKTYCLKTMLHQDIQAMVDG
ref|YP_001438903.1| -----IILEDVTGAGKTEAALVLAHRLMAAGQAR
                        :: . . . ***      : :: :

ref|YP_001251565.1| ----AVIVVDPKGDQDLVRDMMACKVSGRTEDFKIVH--LGFPEQSAQYNPLKNFDQIS
ref|YP_126374.1| ----AVIVVDPKGDQDLVRDMMACKVSGRTEDFKIVH--LGFPEQSAQYNPLKNFDQIS
ref|YP_123382.1| ----AVIVVDPKGDQDLVRDMKAACKVSGRTEDFKIVH--LGFPEQSAQYNPLKNFDQIS
ref|YP_001251158.1| ----AVIVVDPKGDLDLVRDMYSACKASGRLHDFRVVH--LGFPELSAHYNPLKNYDQVS
RAAC02354 QQDCRMIVISPEGSMDIEDCTSRQVDWIKIYSRALH--VSSPYEKAFFSALRNLLMVS
ref|YP_001438903.1| ----GLYIGLP--TMATANAMYARMSQAWLRRLYREGSHPSLVLAHSARKLSAGFNASIWV
                        : : * . : . . * : . . . *

ref|YP_001251565.1| -EVATRITDAISAE-----
ref|YP_126374.1| -EVATRITDAISAE-----
ref|YP_123382.1| -EVATRITDAISAE-----
ref|YP_001251158.1| -EVATRVTDAIQAE-----
RAAC02354 PDDSQSTCDRVLAEYGLLKPFLYCSRQNTKIVSLYTGYPEIDTFIGAAFVHYLACNMSKI
ref|YP_001438903.1| HELLPNDSGDEAAAYEGCAAWFAQSPKKALLAETGVGTLDQAMMAVMAFKH---QNLHLL
                        : . *

ref|YP_001251565.1| -----
ref|YP_126374.1| -----
ref|YP_123382.1| -----
ref|YP_001251158.1| -----
RAAC02354 TMKPTILYVDELHRYAAYAPHAVAKLFECGAEHGISLIAAVQSTEQLNTAESPCLSSELVN
ref|YP_001438903.1| GLNDKVLIADIEIHSYDAYMSHVVEKLVETRARYGN---ATILLSATLSQAQRDRLIAAFY

ref|YP_001251565.1| -----
ref|YP_126374.1| -----
ref|YP_123382.1| -----
ref|YP_001251158.1| -----
RAAC02354 RNTRFHVKMRTSEDRTLR LAPNQALWITSTTRHVIEVSPQFRN
ref|YP_001438903.1| KG-----LNTTRESPRLGPDDYPWIT-----

```


59/228

FIG. 55

```

ref|ZP_02602342.1| -----
ref|ZP_02597242.1| -----
ref|ZP_02524501.1| -----
gb|ACA42232.1| -----
ref|ZP_02758276.1| -----
RAAC02355 MYALAYWRAKRKEHRRMPEVPALVRVSIHSTWFMVVIGAIGAFALYRSVFRALGRMAQL

ref|ZP_02602342.1| -----VEIGPHIEHKEMVRIKGDRTLN-----CIIIGPIGSGK-----
ref|ZP_02597242.1| -----VEIGPHIEHKEMVRIKGDRTLN-----CIIIGPIGSGK-----
ref|ZP_02524501.1| -----VDIGPHIKHKEMIRIKGDRTLN-----GIIIGPIGSGK-----
gb|ACA42232.1| -----LILFVLPALISGLMIFIQVRDFMVH-----KDMLSESFMTW-----
ref|ZP_02758276.1| -----KIESINGEAGETLDTFVFIPLKDRFLH-----MLILGPTGCGK-----
RAAC02355 GTLAVIAMAGFGVVKLHGLIAAWQNRLHPSTPKAATTATNSHAIINPLTSGTGPKPLS
: : : :

ref|ZP_02602342.1| -----TSSLIIPMINQDLHWMARFI
ref|ZP_02597242.1| -----TSSLIIPMINQDLHWMARFI
ref|ZP_02524501.1| -----TSSLIIPMINQDLHWMVRFI
gb|ACA42232.1| -----KAPIIRRFTHKKLLGTADI
ref|ZP_02758276.1| -----TSQTIIPMINQDLQNK
RAAC02355 ALNATNPLVHENAWVIALALVAGVLGLLAFFVYRGLSPTAKAQFLQRMQATDFDRSTHRR
: : : :

ref|ZP_02602342.1| N-----KFINVFKKKDYHTEEVKGTFLN-----
ref|ZP_02597242.1| N-----KFINVFKKKDYHTEEVKGTFLN-----
ref|ZP_02524501.1| N-----KFENTYKKNNYDTEEVKGTFLN-----
gb|ACA42232.1| IGYQFKTMIP-----IVLKEAQRFLHEAVIGATGSGKTSTALLRIAQDLIN
ref|ZP_02758276.1| -----
RAAC02355 DDDPNKFSLVGVEVGIRKDNRPRIEGKDRFINTLVLGSTGTGKTSRIMLKAVYQDLRS

ref|ZP_02602342.1| -----GVTVIEPSNDLCQKVFKLVQAHKIPESAVYYIDPTNPDTKNINILRGPVDKV
ref|ZP_02597242.1| -----GVTVIEPSNDLCQKVFKLVQAHKIPESAVYYIDPTNPDTKNINILRGPVDKV
ref|ZP_02524501.1| -----GITVIEPSNDLCQKVFKLVQAHKIPESIIYYIDPTNPDTKNINILRGPVDKV
gb|ACA42232.1| IATGRKMGMLVFLEPKGDGVDDVLKMCCKLKIPDEKIKVIDATKAFTIKENPFIFIGPSAPA
ref|ZP_02758276.1| -----IGVTVLEPKSDLAEKVYQMAKIHNR-----EVIYFNPLADCPYFNPLFGHESEV
RAAC02355 MANG-TPMDVIAMDPDGGFAQAAVNMANQLGVET-IIMDLRGTMPTSVSFPFGGEIADI
: : * : : : : : : : *

ref|ZP_02602342.1| AEVFAMVIQGLSESNNAFFEQAQRNHLKQHIYLLKLHNPQKD-----
ref|ZP_02597242.1| AEVFAMVIQGLSESNNAFFEQAQRNHLKQHIYLLKLHNPQKD-----
ref|ZP_02524501.1| AEVFAMVIQGLSESNNAFFEQAQRNHLKQHIYLLKLHNPQKD-----
gb|ACA42232.1| AATFEGTINALSGDQDDFYKGQONEAASTLTKLAKIAFGEKTNIFHIQRMFSDPRYLANI
ref|ZP_02758276.1| IENMATTFNMLNPDSQFFKDMSDGLIRKSVKLLKRLYGDDA-----
RAAC02355 IDNVRAALQEKMGKQDGGFFQNAQDDLVRTVIQVQVPLWPEAD-----
: : : * : : : :

ref|ZP_02602342.1| -----VTFDDLIEMYD-----DVEHVHRMHKLLKVQVEKLYDFVQSGAAS
ref|ZP_02597242.1| -----VTFDDLIEMYD-----DVERVHRMHKLLKVQVEKLYDFVQSGAAS
ref|ZP_02524501.1| -----VTFDDLIDMYD-----DVEHVHRMHKLLKVQVEKLYDFVQSGVAS
gb|ACA42232.1| VESIREQITTREKQFREQKDLIMLLDSLGSVEHIEQTEHELMFKLKNLSNSSAINENQ
ref|ZP_02758276.1| -----TLIDLND-----LVWNSNGIGKKIVNEFSKLPVKNPQLQ
RAAC02355 -----FLQFADLVLT-----DPLHFRAICSMVQDCAAQEAGTAKKKRKS
: : : :

ref|ZP_02602342.1| R-----DQKNEYKIKGIHEWFDNTIREKLDFOGEPAVYKSGKYRGQPMHYDR-----
ref|ZP_02597242.1| R-----DQKNEYKIKGIDWFDNTIREKLDFOGEPAVYKSGKYRGQPMHYDR-----
ref|ZP_02524501.1| R-----DQKNEYKIKGIDWFDNTIREKTGFQGEPAVYKSGKYRGKPMHYDR-----
gb|ACA42232.1| KKLNEIQELKKRNLLRAQQIEIDNTEQIIFYFENEVLLYKVDQRTQQPILYPKNHIYAN
ref|ZP_02758276.1| Q-----KENEEIAIWFLTDYYSGMTGDRKGT-----
RAAC02355 E-----MDEAMSMWEHERPEVEARFHLTPHEKSMVLSAARSFLMDTATEQK-----
: : : :

```


61/228

FIG. 56

```
ref|NP_347717.1|-----AVGLAVGVVLGALM
ref|YP_429187.1|-----IGRGIVG---
ref|YP_423535.1|-----MMVVSHVVVLGVSA
RAAC02356MMGRSHMAIGAVGAVAATPLVLHERWESLRDLLTHPWASMPHIVVVQAAFVAATTVVGALV
ref|YP_645289.1|-----
ref|ZP_01860459.1|-----IGAGGISGLI

ref|NP_347717.1|PDIDTSRSKVSYNLRKIV-----VYAALVIVGAYILMANFYPEVEKTNL
ref|YP_429187.1|-----TAILGAWHQTR-----
ref|YP_423535.1|WTLAQRAG-----LAVPLGPEGAAAAALG-----
RAAC02356PDLDQQDAKLTYTIEIVFGLPVLALAIVVMVLMHWATSLTAWGIALLLMFI FGAAHNVTR
ref|YP_645289.1|-----
ref|ZP_01860459.1|PDLDVDGKLSNTFTSSYK-----FLMTLVQFIGVAVIVYSWWAGTDQEMW

ref|NP_347717.1|FLKNMGVYKLDYIPYNLKLSNAGLLIIVGCIIFSFKTKHRSFSHSILGLILFTVGVKLL
ref|YP_429187.1|-----LPLLLPAG-----VAFVALAFLPHRGITHSALGLVLAWLAIKTL
ref|YP_423535.1|-----ALLPDIDHPSSWLGRRLWPISKPISMVLGHRGLTHSLLAVMGGLAVLMLV
RAAC02356MLGLGALATILLDLAYHHRMPMEAAVLLAAMVATMPAKHRTFTTHSLLGLAVFGAGCYLS
ref|YP_645289.1|-----LRAT-----SFVLVRTVGHRTLTHSLLGLALFCAPVWLL
ref|ZP_01860459.1|KG-----IAAGLGIMLVSSFIKKRHLLTVAGGGALAGGMFLAETWLQLLGIYIVIASL
                                     :::: * :

ref|NP_347717.1|LG-----NIFIY---FAVGFI SHMVADTFTNSGIEVFYPIKKKISLKL VHTGSMLDHFTG
ref|YP_429187.1|G-----WPQVP---FLIGYGIHLAEDLLTPSGIPLLYPWQERMRIPLAQTGGIIDRTL G
ref|YP_423535.1|EPGRGLVRLAEP---LALGYLSHLAADALTPAGVPLLWPWKQRFVGLCSTGGVMEWL V-
RAAC02356EPALSHLHLGVAAYGLILGYVLHMAADFIAG-GVPLLWPWGKRQGVHLVKSFSFSAVDYLIG
ref|YP_645289.1|LGG--YPAFALA---LAAGYASHLLADALNTRGVPLLWPLGK-----
ref|ZP_01860459.1|VSHRSYTHSLIG--LAFFAVIAYYLQASIKTEGILLACTGG--YVSHLIADMKWLPFNKR
                                     .   :   :   * :   :   .

ref|NP_347717.1|GLAFIVFLAIIL-
ref|YP_429187.1|LAALALFL-----
ref|YP_423535.1|-----
RAAC02356GIGIFTFVGLALV
ref|YP_645289.1|-----
ref|ZP_01860459.1|GIKLFLPVSRKE-
```

62/228

FIG. 57

```
ref|NP_217638.1|      PRTEINTBHGMYCBACTERIUMTUBERCULSISSTRHAARLEMGBABQHYPHETICALP
ref|ZP_00876805.1|    -----
ref|NP_856790.1|      -----
gb|EAY58379.1|        -----
ref|ZP_01505670.1|    -----
RAAC02357             --MVCDMAIRYDGWSIAAQYRRCGKSACRVCREGPGHGPYWGSKTVDGRRLTKYFGKVP

ref|NP_217638.1|      RTEINMRAMYCBACTERIUMTUBERCULSISHRAGBABRHYPHETICALPRTEINTBFG
ref|ZP_00876805.1|    -----
ref|NP_856790.1|      -----
gb|EAY58379.1|        -----
ref|ZP_01505670.1|    -----
RAAC02357             PVEQEIAQDEPSVLKELSRLEENENLRAQVAQLQAEALALRTPPALSDPPHPLEETKEA

ref|NP_217638.1|      MYCBACTERIUMTUBERCULSISFSCRESIGNIFICANCEIDENTITIESPSITIVESGA
ref|ZP_00876805.1|    -----
ref|NP_856790.1|      -----
gb|EAY58379.1|        -----
ref|ZP_01505670.1|    -----
RAAC02357             SQDELTIHEQPRRRASVRSSDELTKADVEAIWCEETEGLEEDPDILAEVHDALVRNLFPV

ref|NP_217638.1|      PSSISENYRRCGKPN--CVCAQEGHPGHGPRYL-----WTRTVAGRGTKGRQ
ref|ZP_00876805.1|    --SISENYRRCGKPN--CVCAQEGHPGHGPRYL-----WTRTVAGRGTKGRQ
ref|NP_856790.1|      --SISENYRRCGKPN--CVCAQEGHPGHGPRYL-----WTRTVAGRGTKGRQ
gb|EAY58379.1|        --SISENYRRCGKPN--CVCAQEGHPGHGPRYL-----WTRTVAGRGTKGRQ
ref|ZP_01505670.1|    --SLIERYKRCGKPG--CKCADG--PGHGPKYY-----LSVSFPGRRPQMDY
RAAC02357             PTQDTTRLKTGTNPRKPYVCPFRATKRGGRLTFGSADKLVRTAIPWLIQSVKHQREWREL
                        .   .   :   :*   *   .   *   .   :   .   :

ref|NP_217638.1|      LSVEEVDKVRaelanyhrfa-----QVSEQIVAVNEAI
ref|ZP_00876805.1|    LSVEEVDKVRaelanyhrfa-----QVSEQIVAVNEAI
ref|NP_856790.1|      LSVEEVDKVRaelanyhrfa-----QVSEQIVAVNEAI
gb|EAY58379.1|        LSVEEVDKVRaelanyhrfa-----QVSEQIVAVNEAI
ref|ZP_01505670.1|    VPQADYTDVAEHLanyhrvr-----EIIIEICEINREL
RAAC02357             SRRRQQDRIaellstyrrmdtLELVHLVERMERAILMGDRtGWTvrLEELDEHVEVIRQV
                        :   :   *:.*:*.   :   :   :   :

ref|NP_217638.1|      CEARPPNP
ref|ZP_00876805.1|    CEARPPNP
ref|NP_856790.1|      CEARPPNP
gb|EAY58379.1|        CEARPPNP
ref|ZP_01505670.1|    LRRR----
RAAC02357             LQERNT--
                        .   *
```

```
ref|YP_655149.1|-----ELRPYQREAVEAIESHWSQGVTTRVGVVLPTGTGKSTVIGRTAVNGYQN  
ref|NP_943831.1|-----MRPYQREAVEAIESHWSQGVTTRVGVVLPTGTGKSTVIGRTAVNGYQN  
ref|YP_001456771.1|-----ELRAYQTEAVNAVLGEWVDGK-RTCVLVLTGTGKSTVIKLAIEIAYKA  
RAAC02358ALNDVEGISVAIQLRPYQQAAVDAFFQALAEGRKRQLIVLPTGAGKTIVFGSVARRFHEE  
ref|YP_024839.1|-----LRPYQEKALEGIEKAEREGVRRLVVLPTGTGKTUVFSHGIKNRADR  
ref|XP_748956.1|-----SPKLVLRDYQEECIQSVLKYLDEGHKRLGISLATGAGKTIVFTTELIGRIPSR  
:* ** .:. : * * : *.***: :  
  
ref|YP_655149.1|R---EPVLMVAHRGELIDQMAGTIFEVDPSIPRSHVGIVRAEMDDHSAPIVVATLQTLAT  
ref|NP_943831.1|R---EPVLMVAHRGELIDQMAGTIFEVDPSIPRSHVGIVRAEMDDHSAPIVVATLQTLAT  
ref|YP_001456771.1|G---QRVILLAHRRRELLDQMAQSIRMVAPQIPTDDIGFVQAERDQPERPIVCASFQTLMS  
RAAC02358VSRERPILVIAHRTELLDQAEQKIHFVWP--EAFIGRIQGARNEQLGDVLLASTQTIVA  
ref|YP_024839.1|G---RSLVLVHRDELVRQTIKIGMVAP---ELTTGVVKADENGLDADVVASVQTAQV  
ref|XP_748956.1|NEIGDKSLIIVHRKELVEQAAQHCRRAYP--DRTVEIEMGHSHASGAGDIVVASVQTLTR  
:::.* ** : * . * : :: *: **  
  
ref|YP_655149.1|AHRREAVGFRR-----RILWDEVHHAGAEGFHHTTFTELGGYTDL----FAGFTATMRR  
ref|NP_943831.1|AHRREAVGFRR-----RILWDEVHHAGAEGFHHTTFTELGGYTDL----FAGFTATMRR  
ref|YP_001456771.1|ASRLQAVGERT-----VVLVDEVHHSAASTYAEILSAPN-FDGAF----KAGFTATLQR  
RAAC02358GRRIPQPG-----LIITYDECHHSRAEGALGVLERLGVFESDGPP--LLGVTATPSR  
ref|YP_024839.1|DRRLAQVLVEAAKRSPFGTVVWDEAHHPAPSWTKVLTGLGSFNYPGYPL--TVGFTATPER  
ref|XP_748956.1|GNRLAKFDPKR---FKLLLVDEAHHIVASSYREV LKHFGADETSADS PVLVGVVSATFSR  
* : ** * * : . *: ** *  
  
ref|YP_655149.1|DDKGKSPVGLGDVIEKV-----VYEKDILWAIDS GYLVRPRGLTVRIKNLNALDDVRTVA  
ref|NP_943831.1|DDKGKSPVGLGDVIEKV-----VYEKDILWAIDS GYLVRPRGLTVRIKNLNALDDVRTVA  
ref|YP_001456771.1|ADGG-----LADYWD SI-----AFERDLRWALDEGFLVP PQGKTVVIPGLDTTN-IKLRN  
RAAC02358SDRT-----ELGDI FEHL-----TYERTILDIMMDGYLS D VRGVKVEV PGLNLGA-IRTTA  
ref|YP_024839.1|DKKTLGVWERLAAAFMSIREAIYGNGKRKG DGHEGGYLP I LP AVV VET EMDLTR-VRK TG  
ref|XP_748956.1|SDGLK-----LGAAIDHI-----VYHKDYIDMINDNWLAN-AVFTTVRSEANLSKVKKDSF  
. : : :.* : : :  
  
ref|YP_655149.1|GDFQQSDLAEVM--EAATEYVVDAIKLHAADRR-PIIFAASVDAAHHIADALTAADFPAV  
ref|NP_943831.1|GDFQQSDLAEVM--EAATEYVVDAIKLHAADRR-PIIFAASVDAAHHIADALTAADFPAV  
ref|YP_001456771.1|GDYAAGDLS EVM--MSSVDSTVEA IH THAPDDR-MLIFGAGVEHC QALS DTLSATGIHTA  
RAAC02358GDYN SKDSL SY VM NIESALDAVVD AVVTHAPGRK-CLVFAVDVKHAHALAE RFQ KRGIACA  
ref|YP_024839.1|GDFSEGDLGREMEESGAIAQIADAYLINAHDRK-GVAFTPTVATAHALAAALCARGIRTE  
ref|XP_748956.1|GDFAIGPLSKAVNTENVNNITVRAWLANAQDRKSTLVFCVDVAHTKALTETFRNYGIDAR  
** : * . : . * : * .* : : * * : : : :.  
  
ref|YP_655149.1|AVTGSM S Y A ERQP V YEAYRN GTAKAL VT VQ VL TE GA DF M CD CV VL AR PT R SR NL Y SQ MI  
ref|NP_943831.1|AVTGSM S Y A ERQP V YEAYRN GTAKAL VT VQ VL TE GA DF M CD CV VL AR PT R SR NL Y SQ MI  
ref|YP_001456771.1|LVVGST S SEER TEL FEE FTAG RV Q AL VT VQ VL TE GT DL PAC DC V VL AR PT RS AV LF T Q MV  
RAAC02358AVD GAM K AE ER AAIL QA FA EN RL CV LV NC Q IL TE GY D Q PD VC V VI AR PT RS Q AL Y V Q MV  
ref|YP_024839.1|ALDGT TH KDERRAILRLR L KTGETQ V VTNC GV LT EG FD EPS I SC V V AR P TF KH GL Y V Q MI  
ref|XP_748956.1|YITAKTPKDVRMEQLRAF RN GEY PV LL NC GL FTE GT DI PN ID CV LL AR PT R SR N LI IQ MI  
: . * . . . : . : * * * . * : : : :  
  
ref|YP_655149.1|GRALRLYDGKQDALVLDLAGSS-----RSMKLVNLTQLV-----  
ref|NP_943831.1|GRALRLYDGKQDALVLDLAGSS-----RSMKLVNLTQLV-----  
ref|YP_001456771.1|GRALRLHQDKNDALVLDLAGST-----RDVAMVTLSSLV-----  
RAAC02358GRALRLHPDKTDALVLDLTGAS-----DDKSLQTFARLMRTQRKTA  
ref|YP_024839.1|GRGTRLYPGKKDLMLIDLVAAS-----RRHEFVG YVD LG-LDL DEG  
ref|XP_748956.1|GRGLRLYPGDKEDCHI DM VAT L NT G VL ST P TL FGL HP DE I L Q NA KA KL DR M PLE K TT TG  
** . * : * * : : * : : :
```

WO 2011/139994

PCT/US2011/034852

64/228

ref|YP_655149.1|
ref|NP_943831.1|
ref|YP_001456771.1|
RAAC02358
ref|YP_024839.1|
ref|XP_748956.1|

THALVGAEEGEDAVPMEDGESVGEWLTRVAQKRELAEQVAQAINLFANRSRYRWVRVKDN
KKPKEGEPERQACPTCEEPCEVTEHRCALCHRYLPVAVTAEGGSRHENCQANGSGRVN--
TGTALAGEETEPPPTPDDLVDNLTFTKYDTIYDLIADLKSEKHIRSLSPHVWVRVGDHR

ref|YP_655149.1|
ref|NP_943831.1|
ref|YP_001456771.1|
RAAC02358
ref|YP_024839.1|
ref|XP_748956.1|

FAIAYGHDGWAYLYRDGDEFWVLELKNEKFMPLHDRSIPLEYAQGVVEGFLSLFESSLI
-----VFGESRLRWLPVGPAPWVLGAGKEIVVMVPEGVDTWKLAAYE
YVLSDASGWLTIDKEEDTSRAYRPRLNPDPTS IANTPDEPYIFTVRHVAKFKNSDDSIMH

ref|YP_655149.1|
ref|NP_943831.1|
ref|YP_001456771.1|
RAAC02358
ref|YP_024839.1|
ref|XP_748956.1|

TKEADWRNAPMTERQKYVLQKYRIRYDDTWTRGMAADALGQRFQAAKRVRVLQKNFDAQ--
NGRVEVLHEEIPSDWAMGIGEDRAKAFQKLVERQARWLNEPVSI SQKGRLVREGLPEK--
TRPRIATAPDFETALRAADTFATKEFEERYVSVRQTWRQLPATEAQVRFLNKAKVRHGS

ref|YP_655149.1|
ref|NP_943831.1|
ref|YP_001456771.1|
RAAC02358
ref|YP_024839.1|
ref|XP_748956.1|

KWRDVLAQPHGQAWLEHRLMALRARAQSRARVAQTS
HLPRVKTKGEAADLLTRISGKRAVRKL-----
IQRKHLTRGQAADLITKLKFGGKQRFEARRAERL--

FIG. 59

```

ref|YP_001086797.1|
ref|ZP_01805266.1|
ref|ZP_01966753.1|
ref|ZP_01188985.1|
ref|YP_001681547.1|
RAAC02361
-----
MDTQCIEVRCGSARHHSYEDLFTPEVQSLFEECYALGVEVRCLCNDRPVPMLRRVRI

ref|YP_001086797.1|
ref|ZP_01805266.1|
ref|ZP_01966753.1|
ref|ZP_01188985.1|
ref|YP_001681547.1|
RAAC02361
-----
RPPTYTVVTNPLGIHHFNCPRFRQSRHPSASHKTRDAQAVETTPNHIEEWLPHRRIFVPT

ref|YP_001086797.1|
ref|ZP_01805266.1|
ref|ZP_01966753.1|
ref|ZP_01188985.1|
ref|YP_001681547.1|
RAAC02361
-----
LDTLNPAQREAVEKTEGPVLILAGAGSGKTKVLTTRIAYLIEDKGVOAPNILAITFTNKA
LDTLNPAQREAVEKTEGPVLILAGAGSGKTKVLTTRIAYLIEDKGVOAPNILAITFTNKA
-DTLNKPQKEAVFHTEGPLLILAGAGSGKTRVLTHRIAYLIEEKGVNPNILAITFTNKA
LTGLNPEQKKAHEHFEGLLILAGAGSGKTRVLTHRIAYLIENYGVNPLQILAVTFTNKA
LHSLNPVQREAVLHQEGPLLLLAGAGSGKTRVLTHRIHGLIEQCRVSPFHILAITFTNKA
LDQLNDAQREAATHKNGPCMVVAAAGSGKTAMLIARIQHILIN-QGVRPGDILACTFTRKA
** *::*. : : ** :::*.***** : * ** :***: * . .*** **.***

ref|YP_001086797.1|
ref|ZP_01805266.1|
ref|ZP_01966753.1|
ref|ZP_01188985.1|
ref|YP_001681547.1|
RAAC02361
-----
ANEMRERVEQNIGPETKDMWISTFHSCCVRILRKDINKIGYNRSFVIYDSADQVTLVKDC
ANEMRERVEQNIGPETKDMWISTFHSCCVRILRKDINKIGYNRSFVIYDSADQVTLVKDC
AEEMRQRVDSLVGIGAESIWVSTFHSVCVRILRRYIDRLGYDNRFTIYDQKTLKKEV
AGEMKERVNDLLGGMAGDLWVSTFHSICARILRKEIGKIGYDNNFVIFDQKQLISRI
AAEMRERLGRIGPRAHDVWVSTFHSCTMRILRKDGEKLGDRSFVIYDYDDQRLKEC
AQEMTDRLLLAAVGERGKAVTIGTISVAYRMVTPELG--EDWRVLSDPWTMIERVLEEPS
* ** :*: : * : :. : ** . *:: : : : * .

ref|YP_001086797.1|
ref|ZP_01805266.1|
ref|ZP_01966753.1|
ref|ZP_01188985.1|
ref|YP_001681547.1|
RAAC02361
-----
LKELNLSKDVFEKPAVISASISGAKDKLYTPKQFKDINMADNRMVKIADIYALYQDRLKRN
LKELNLSKDVFEKPAVISASISGAKDKLYTPKQFKDINMADNRMVKIADIYALYQDRLKRN
CRKTDIDTKRFRKERMILLSVISSAKNEMILPEEFELNAGGDFVQLKIAKVYKEYEAQMRAN
LKELNLDPKKTRPRAILSEISRAKNELIDPRSYANNVG-DYFQDITARIYPLYQERLKES
LKELNIDEKRFKPQAVGAAISSAKNRLVGPVAFERQAY-DHFAQISAKVYHIYQKKLKAH
GKNPHGVGVPVVKLGEAISAIKAKADALGP-----HQVSDPLTKVYAAETLKAER
:: . . : * : ** * . : : * :

ref|YP_001086797.1|
ref|ZP_01805266.1|
ref|ZP_01966753.1|
ref|ZP_01188985.1|
ref|YP_001681547.1|
RAAC02361
-----
SALDFDDLILKTVELFKANDEVLAYRYSRFRYIMVDEYQDTSKAQYELIKLLAREHQNIC
SALDFDDLILKTVELFKANDEVLAYRYSRFRYIMVDEYQDTSKAQYELIKLLAREHQNIC
NALDFDDLILKTVELFKANDEVLAYRYSRFRYIMVDEYQDTSKAQYELIKLLAREHQNIC
NALDFDDLILKTVELFKANDEVLAYRYSRFRYIMVDEYQDTSKAQYELIKLLAREHQNIC
NALDFDDLIMKTIEVFVDNPMVLEYQERFKYILVDEYQDVNFAQYKLVQLLANKYRNL
NAMDFDDLILVNGVCLFREFPVLDNYQDRFRYIHVDEYQDTNHAQYVVLKLLADKYRNL
KTLDFFEDMILHAIRLFTDEAFKRWQRWRVYVMVDEFQDTNTAQWLFLLELVKAHNNLF
.:**:*::: : : . : : ** : ** : ** : ** : ** : ** : ** : ** : ** :

ref|YP_001086797.1|
ref|ZP_01805266.1|
ref|ZP_01966753.1|
ref|ZP_01188985.1|
ref|YP_001681547.1|
RAAC02361
-----
VVGDDQSIYGWRGADIRNILEFEKDYDNVHVVKLEQNYRSTQVILDAANKVISNNIERK
VVGDDQSIYGWRGADIRNILEFEKDYDNVHVVKLEQNYRSTQVILDAANKVISNNIERK
VVGDDQSIYKFRGANIRNILDFEKEFSDAKVIKLEQNYRSVGNVLEVANSVIRNNKGRK
VVGDPDQGIYGRGADIRNILDFEKEFSDAKVIKLEQNYRSVGNVLEVANSVIRNNKGRK
VVGDDQSIYGWRGADIQNILDFERYPEAVVLKLEQNYRSTGKILEAANAVVGNNGRGRK
VVGDDWQSIYIFRGRAPDLMEFLRRFPDAKRVTLTNYRSHDLIVDVGRRIIRLNDGHQ
**** *.** :*** : * . : . : ** ***** :..... : * :

ref|YP_001086797.1|
ref|ZP_01805266.1|
ref|ZP_01966753.1|
ref|ZP_01188985.1|
ref|YP_001681547.1|
RAAC02361
-----
RKK---LWSEKKEGELIKIQLTGSEIEADFIADSIAQIARKENRPYKDFAVLYRANAQA
RKK---LWSEKKEGELIKIQLTGSEIEADFIADSIAQIARKENRPYKDFAVLYRANAQA
EKT---LWTDNEKGEKIRLRQFDTADEAQFIAEDIKDETAQG-ANYSDHAVLYRTNAQS
EKR---LWTKRGKGEDLKLYVAFDDKDEASYVCRKIKELKREKNYKFSDFAVLYRTNSQS
SKK---LWTQNPSPGQPIVAYQGETEHDEARYIVRTIKRLSESENRPYRDFAILYRTNAQS
LPKRVAHRSMPEGAIAQIVTVRSDLLEEARFVAQELQRLRKEHGVSWSDCAVLYRTNIQS
. . * : ** : : : . : * ** : ** : ** :

```

ref|YP_001086797.1|
 ref|ZP_01805266.1|
 ref|ZP_01966753.1|
 ref|ZP_01188985.1|
 ref|YP_001681547.1|
 RAAC02361

RPVEDALNRSQIPYNIYGGTKFYERKEIKDLLAYLRVIQNPQDDISIKRIINVPRRGIGL
 RPVEDALNRSQIPYNIYGGTKFYERKEIKDLLAYLRVIQNPQDDISIKRIINVPRRGIGL
 RLLEEKFVAMNIPYKIVGGINFYSRREIKDVL SYLKTIDNGKDDLAVRRIINVPKRGIGL
 RSVEEMMVKYAIPYQIVGGFRFYDRMEIKDILAYLRVIYNPSDEVSLRIINRPKRGIGQ
 RVLEEHFMYAGIPYRIFGGLRFYERKEIKDIVAYLRFISNPADAVSFRVNVNPKRGIGD
 RLFEALADADIPYHVVGDKHFYESPDVKIILDYLRTTQDTS DPTVWGHLLNRPKRYIPI
 * .*: : ***.: *. **. :*: : **: : * :*: *:* *

ref|YP_001086797.1|
 ref|ZP_01805266.1|
 ref|ZP_01966753.1|
 ref|ZP_01188985.1|
 ref|YP_001681547.1|
 RAAC02361

RTIEKIEDRANLKQESIYSVLIDIETNSDISTKARASISGFVDIIGTLRTIKEVYPVSKL
 RTIEKIEDRANLKQESIYSVLIDIETNSDISTKARASISGFVDIIGTLRTIKEVYPVSKL
 TTINRIQESAAAARGIGFYDALSAPDLIPGIGRSASK-LDSFAALIEYFKGRSEESGVTDL
 GTISKLSRYARERGISLYKAGTEAESNPYLTA SFKKRVKAFFDLLEELREKSETLSIDL
 ATVQKLEHADSEGWTVG EALARVGEVPGLSR-AVKALS AFGQMIEELRREAPSLLV TQI
 DVVHEVQHGG-----WEAVVAHPKCRAFVTTIDTLRRIE--PSKAI
 .: .: . * : : :

ref|YP_001086797.1|
 ref|ZP_01805266.1|
 ref|ZP_01966753.1|
 ref|ZP_01188985.1|
 ref|YP_001681547.1|
 RAAC02361

IEKVLDTTGYMDELVEIRNKNEKDLTGKGEEAQDRIDN LREFISIALEFESSNDDTYENK
 IEKVLDTTGYMDELVEIRNKNEKDLTGKGEEAQDRIDN LREFISIALEFESSNDDTYENK
 LTEVIEKTGYTESLEADDP-----EELEARVQNI DELVSKAAVYEESCSDRGERP
 THQVVTRTGYQRELNEEGT-----QQARNRLENIQELFS---VIEEFMKG-NENK
 VEAILNRTGYVRELEAEKT-----EEAKGRIENIKEFLS---VTKEFDRT-ADDK
 QWLVDNHPGLVRQODEDEP-----IKWVDSL IAS-----ASRYK
 : .* . : : : :

ref|YP_001086797.1|
 ref|ZP_01805266.1|
 ref|ZP_01966753.1|
 ref|ZP_01188985.1|
 ref|YP_001681547.1|
 RAAC02361

DLETFLTSIALTSESND-EDNDRVSLMTIHTSKGLEFPV VFLTGMEEGLFPISRAIKSM
 DLETFLTSIALTSESND-EDNDRVSLMTIHTSKGLEFPV VFLTGMEEGLFPISRAIKSM
 TLSGFLEEVALVADIDSVAEDRDYVILMTLHSAKGLEFPV VLAGMEDGLFPSYMSISGD
 TLGAFLEEVSLISDVNMDNQNVTIMTLHSAKGLEFPV VFIIGMEEGLFPHANSMM--
 TLEEFLAGVSLVSDTDNNEDED AVVLMTHSAKGLEFPV VFVAGMEEGVFPHSRVQF--
 TVASFLRFVDWIE-KSKEPKDEAVQIMTIHKAKGLEWTTVFVAGLAEGLLPHKKALK--
 : ** : : . : * ***:*.****: . **: * :*: *

ref|YP_001086797.1|
 ref|ZP_01805266.1|
 ref|ZP_01966753.1|
 ref|ZP_01188985.1|
 ref|YP_001681547.1|
 RAAC02361

SDSQIEEERRLCYVGITRAKEELM-----
 SDSQIEEERRLCYVGITRAKEELM-----
 DPEELEERRLCYVGVTRAEEKLTLCAR-----
 DHEELEERRLCYVGITRADELYLTRAR-----
 EETQVEEERRLCYVAITRERERLYLARA-----
 -GEELEERRLCYVAATRARDNLYLMAAKWYGDKEREVSRYVNAVKNP
 :. ** *****. ****.:* :

FIG. 60

```

ref|ZP_02406784.1|      ----RRHARRRGGGDDRRRIAVVEQRLHARQIVVVHVDRAHLMRIRAEELAAEPRRE-VLLQ
ref|YP_335221.1|      ----RRHARRRGGGDDRRRIAVVEQRLHARQIVVVHVDRAHLMRIRAEELAAEPRRE-VLLQ
ref|ZP_02485527.1|      ----RRHARRRGGGDDRRRIAVVEQRLHARQIVVVHVDRAHLMRIRAEELAAEPRRE-VLLQ
ref|ZP_02509777.1|      ----RRHARRRGGGDDRRRIAVVEQRLHARQIVVVHVDRAHLMRIRAEELAAEPRRE-VLLQ
RAAC02362               MFPVPYILLIRTYVRIRGVRMNRPTALLRQLLILELIQAHIVRERARVKAQMRAEGLHIV
ref|YP_157691.1|      -----VQILDVRLKRVDL---PLEVSESVYRRMEAERKRVANELRSEGGATA
                        : : .      : : .      ::  *  . :  *  *   :

ref|ZP_02406784.1|      EREREHVQQPDVRVVRAHERDHVQPAFAQQQLQ--AERA-----
ref|YP_335221.1|      EREREHVQQPDVRVVRAHERDHVQPAFAQQQLQ--AERA-----
ref|ZP_02485527.1|      EREREHVQQPDVRVVRAHERDHVQPAFAQQQLQ--AERA-----
ref|ZP_02509777.1|      EREREHVQQPDVRVVRAHERDHVQPAFAQQQLQ--AERA-----
RAAC02362               ERQDGDMDIRVEFRVGDHYDEAVFMRKMLEAEAAANRAKRTGMISR
ref|YP_157691.1|      EKIRADADRQREVIIAEAYRDAQAKGAGDAKAT-----GI---
*:      .      : : .      :      :      :      :      :

```

FIG. 61

```

ref|ZP_02329650.1|      -KRVIF-LADCQSFYASVEKAHHPEYRNR-----PLVVAGDPARRSGIVLAACP
ref|YP_001421775.1|      -----LVDMQSFYASVEKAEAPHLKSR-----PVIVSGDPERRSGVLAACP
ref|YP_177318.1|         -----LVDMESFYASIEHAANPQYDGR-----PLVVSGDVNRRSGVILAAACP
ref|NP_243607.1|         MDKVIF-MVDMESFFASVERANHPELSGR-----PLLVS
GDPERRSGVILAAACP
ref|YP_001681084.1|      -----LADMNSFYASVEMAHNP
TLRGQ-----PVLVCGDPERRHGII
LAASRE
RAAC02363               MTKLIYGLVDMQSFYASCEVASREEYAARRKEFDDSTDP
PLVVSGDPARRSGIILAAATPT
                        :.* :*. ** * * : : :*. ** ** * :*. **
ref|ZP_02329650.1|      AKKYGITTAE
RLGEAINKCPDLVVVRPRMQEYIKVSLQITEILQSYTDLVEPLSVDEQHL
ref|YP_001421775.1|      AKRYGVKNAERLWEAQAKCPDAVIVRPRMQRYIDVSVMITELFERYTDLVEPYSIDEQFL
ref|YP_177318.1|      AKAKGVRNAERLFEAQQKCPDLVVVKPHMQRYVDISLQISKILGTFTDLVEPYSIDEQFM
ref|NP_243607.1|      AKARGVTNGERLWEAQQKCP
EAVVVRPHMQQYVTVSVQITEILERFTDLIEPFSIDEQFM
ref|YP_001681084.1|      AKRRGVKTAMTVGEARAICPD
AVCVRPRMSLYLEVSWQIQQIARTLSPLVEPYSVDELEFI
RAAC02363               AKRYGVENAMRLGEALRLYPR
LIVVRPHMAFYLVHSVRIQMLMQQCFPFQEQFSVDEGFI
** * : . : * * * : * : * : * : * : * : * : * : * :
ref|ZP_02329650.1|      DVTASIKLLGSPQEI
AKSIQSRVWNETGVYTRIGISENKVLAKMACDNFAKKNGDGIFYL
ref|YP_001421775.1|      DVTGSRRLFGDPFTIAKSIQQA
IMREFGIYARVGIGPNKALSKMACDHFAKKNASGIHRL
ref|YP_177318.1|      DVTGSQKLF
GPPYEIAEKVKQAIMDRFGVKARVGIGENKVLAKLACDNFAKK
SSEGIYWL
ref|NP_243607.1|      DVTHSQRLFGAPREIAQKVQQA
IWHETGVRARIGMGESKVLAKMACDNFAKKMPSGVFHL
ref|YP_001681084.1|      DVRGAEHIWGD
AVEAAGIFRQRVWEAVKVPCSVGVGPNKFLSKMACDVEAKK
SPSGVALW
RAAC02363               AFPYPSNLFDPDIAAARNLQARI
WDQFRIRARIGLAPNKWLAKMAN-KAAKKTGGIVWW
. . : . * : : : * : . * : * : * : * :
ref|ZP_02329650.1|      PKTEMEQKLWPLPVNKMYHIGSRMTRHLKRMGIHTIGDLACASVLR
LQKRWGINGEVIWR
ref|YP_001421775.1|      DMSNIRQDLWPLPVGKLF
GIGKRMEHHLRRMGISTIGGLAGHPAELLKKRWGINGELLQR
ref|YP_177318.1|      RKDSLDDLWCLPIEKLFGVGRKMSVHLRNMGI
RTIGQLAQTGDARIKKRFGVHGQVLWM
ref|NP_243607.1|      TKERMERLLWPLPIECLYGVGRQMTKYFRNQGIR
TIGQLANTSLERIKGKWGVNGHVLWL
ref|YP_001681084.1|      RVDDVEAKLHPLPIGKMF
MVGSRMERHFRNMGLLTIGDLAHYPVDYLMRRFGLRGAVYHN
RAAC02363               REEDIPSVLHPLPVEEMWGLKRRAEVLR
RKFKCETIGDVARLPVGLKAEFGVWGEVIHR
: * ** : : : : : * : * : * : * :
ref|ZP_02329650.1|      IANGLDDSPVTPDT-YIGQKGVGHQMTLPRDYMTINELLVPLLELSELVCQRCRAKGYMG
TARGIDPSFVTVNT-HSRQKAIGHNMTLPRDYSRFEDIKVVLELSEEVARRARFKQYIG
ref|YP_001421775.1|      SANGEDYSPVTRAA-HARRKGYGNMTLPRDYVKKEDICV
VLELCEEVCARLRQDGWMG
ref|YP_177318.1|      TAHGIDPSFVTPHS-HDKQKGIGHG
MTLPHDYVKAEDIHVVLLELCEEVCARRAHLMG
ref|NP_243607.1|      LAWGRDGSFVRPDS-LEQTKSVGH
SVTLPRDYHHSDDIELVLELTDEVCRARKLGKAG
ref|YP_001681084.1|      WANGIDVSDINPDSYHAPHKGF
SHRTTLPRDFYERSEIAVVILELLDEVCHRLRQAHQAG
RAAC02363               * * * * : : * . : * : * : * : * : * : * :
ref|ZP_02329650.1|      QVVSVCQGTDFDYPTGFHRQMKLEDP
TNLSDEVNRAVLLFKRHWNGLPVRKISVSLTG
ref|YP_001421775.1|      HTVSVSIRGADFEFFSDFHRQRKLVSP
TNFGMDIFKTAVKLFKEHWNGEPVRSAGVLSQ
ref|YP_177318.1|      STVLSVNGADFEV
EKRGFHRQYTI PFETNITMEVYEAACALLERFDGYP
IRRLSIGVSN
ref|NP_243607.1|      RTVAIGVSGANMETPTGFHRQMKLTNHTNITMEVYEGAATL
FERFDGKPIRRLHVNLSN
ref|YP_001681084.1|      RTVSVGLR--AYDLTRGFYRQTTLPS
PSNLSGPVFEKAQQLFRRHWDGRPVRTVTVDLSG
RAAC02363               RRVGLGLTYEGLTG--GFYRARTLPRATNDPAELYP
VLLALLDEHWDGSGVRAVSVAVDM
* . . . * : * : * : * : * : * :
ref|ZP_02329650.1|      LVRDDTFQLVLFED---RMKKLAL
EKALDGIKDRFGNASIMRAVSLTAAGQAKDRSMKIG
ref|YP_001421775.1|      LEPCDYVQLSLFDA---QEKKISL
GKVLDDIHERYGPASLLHAASLTEAGQAFHRAEKIG
ref|YP_177318.1|      LQSDQNWQLSLFDDNASRDL
SLTIGYVMDGIRQYKGLAIQRASSLQKASQLRDRSQKIG
ref|NP_243607.1|      LTSDEAWQLSFFGN---RDR
AHQLGYTMDTIKEKFGDTAIRRAVSFLSASQAEERAKKIG
ref|YP_001681084.1|      LEEDGTLQQELFRD---MDRQDRV
SRTMDRIRDAMGTTAIVRASLLPAGQADRRAVKIG
RAAC02363               LQFRETQLQLSLFEN---V
PARTRLYETVDEIRARFGETSIMRAVSLTRAGQLRERSLRIG
* * : * : : * : * : * : * : * : * :
ref|ZP_02329650.1|      GHY-
ref|YP_001421775.1|      GHY-
ref|YP_177318.1|      GHY-
ref|NP_243607.1|      GHY-
ref|YP_001681084.1|      GHY-
RAAC02363               GHYA
***

```

FIG. 62

```

ref|ZP_01860132.1|      -RDRGRIKWTS--MMLPEHVKVLRDWAKEDSYEQRKEIDEQHLEELSEITAEAMEYGRIV
ref|ZP_01171904.1|      IRDRGRIKWTS--MMLPEHVKLLRDWAKEDTYEKPKELEQOLEQMNETLAEAMEFGQAV
ref|YP_079689.1|        LRDRGSIKWVS--MMLPEHVELLREYHESFQKIKKPILDEQKYEEFNEIICEAMAENRFL
ref|ZP_02329649.1|      -----LWESSRMMLPEHREQLEQRRRELKHKHAKPLLDEQRLEELSTIILNYALATKHKV
ref|YP_001487332.1|      -----MWESSRMMLPEHREQLLAQKRKKKEYTPPPLSTDQLEEMNFLITQSITEDQAI
RAAC02364               MNIRDGNI FEAMRLVLP EHRALMAQIQ RERMKRKR PMLTEERLEEMQYVLSEAI REGRIV
                        : :  ::****  :    ..      :  :: *::  ::  : :

ref|ZP_01860132.1|      TITHYVGRRHELLIGRI-----
ref|ZP_01171904.1|      AITHFRTHRHELVI GNIHYWDEIGQKLHVIDHFG EVHRIPLNTVADV-----
ref|YP_079689.1|        QFAYYRQGEVKTLAGRIHYADALKREL RIVSRAD EICILKIEDIIEIEYD--
ref|ZP_02329649.1|      RFTVYDVYEDQHIAGVLIKYDPLTRSLGVI SEANKAMHIMLENIIDVRLE--
ref|YP_001487332.1|      CVTYAAAGRKEQFWGWVKT IHYETQRIKIVND-EDVLNLSLQQ-----
RAAC02364               RVTMFTPERDVVLVGRVS--ARGREL R-VRTAAGVHIVDVRDVVG VEMERS
                        .:      .    . * :

```

FIG. 63

ref YP_161675.1	-----
RAAC02366	MRSSKTMHKEGDLLVVDWTWTGREERFSYDEASAKRSEWLSQYENRERYRFYCLCRGHERI
gb ABH06559.1	-----
gb AAB95339.1	-----
ref YP_001202661.1	-----
ref ZP_01614696.1	-----
ref YP_161675.1	-----
RAAC02366	RLHLVKREHWHLASNPQARLHAPHCGFYRNEVVRIVERRRKKQFGIETVETVNEKGERVA
gb ABH06559.1	-----
gb AAB95339.1	-----
ref YP_001202661.1	-----
ref ZP_01614696.1	-----
ref YP_161675.1	-----
RAAC02366	SLIFHVDDFFSADQEQTEEHEPPTPNMSSAAETERTILPEKRIIVRGERKERGKLTFSGF
gb ABH06559.1	-----
gb AAB95339.1	-----
ref YP_001202661.1	-----
ref ZP_01614696.1	-----
ref YP_161675.1	-----
RAAC02366	LREWYRVGLQWYEHHQQRQAKNVSELLYGMWRVMLDGTITFHDGKDPALLFIPNRYVEP
gb ABH06559.1	-----
gb AAB95339.1	-----
ref YP_001202661.1	-----
ref ZP_01614696.1	-----
ref YP_161675.1	-----
RAAC02366	WHGERQKIVVGQLGDRSRSGSGEFWRVDGIYAKGTSLRANWNVLVSEVHLCPNPHIGALV
gb ABH06559.1	-----
gb AAB95339.1	-----
ref YP_001202661.1	-----
ref ZP_01614696.1	-----
ref YP_161675.1	-----
RAAC02366	ALRVKHETDQLMSVCPAKQALIVERDGCWVDSLVEYQFHEILMNGLQRRPDVKVEKPLE
gb ABH06559.1	-----EQTEKSHPRSP-----NVLSVALSQRTTAPEEELNP
gb AAB95339.1	-----
ref YP_001202661.1	DNLSEYHSILMTSEDP-----DPDLEGEAAKEAREKDEADD
ref ZP_01614696.1	---YRNEVMWLYQHKN-----KHRYRVKNLQDALEQG-YTVAA
ref YP_161675.1	-----PDFIL-TDVAPEVVIEVLGMSGNADYDARIAEKRAHY--LAS-G-IPLLEW
RAAC02366	GLVEYDGRRPDYVLRMEGKPLIIEIWGMSGKTDYDESKEKRQAFYRQLEKRGELRFLEW
gb ABH06559.1	KILALQNAQRKRKIEHDGS--LFQAVGIGTLLQQPDDRATTS-----LSWKR
gb AAB95339.1	-----EEQTKGATFES-----VAREWHAAN-KKWTEEHSRRV-----LKSLED
ref YP_001202661.1	RRLSYLGTQLDLLLNQKKP--LQKALWGISDEILALDTKEERQA-----
ref ZP_01614696.1	HVGSYYGEQFEHLKNIYPK---QIIFTGAAKTRFALLAKQRV-----
ref YP_161675.1	D-----
RAAC02366	DSRNPRERA AVLRLNIGLWMMK
gb ABH06559.1	KGCKSSEQNGMEQKTIIFIPS
gb AAB95339.1	N-----LF---
ref YP_001202661.1	-----
ref ZP_01614696.1	-----

FIG. 64

```

ref|ZP_02093159.1| -----
ref|ZP_02091713.1| -----
ref|ZP_02423704.1| -----A
ref|ZP_02026447.1| -----
ref|YP_001127515.1| -----PEAIEE
RAAC02367 MYDGRQTVVREATALQQPQWTRDIVMVNPNELVPHPRRDEAGERYDTPEAMDNLPTNLVS

ref|ZP_02093159.1| ---SVEQYGVLSPLIARPRP--EGGYEIIISGHRRQHAAQLAGLDTLFPVIVRQMDDDAAVL
ref|ZP_02091713.1| ---SVEQYGVLSPLIARPRP--EGGYEIIISGHRRQHAAQLAGLDTLFPVIVRQMDDDAAVL
ref|ZP_02423704.1| LAESIKMHGVVSPPIVRPLENTADEYEIIISGHRRVMASRKAGITEVPALVVS LDRDAAAI
ref|ZP_02026447.1| LTESIRERGVLLPILVRKTN--DEEYEIIISGHRRTHAARLAGLEKVPVVI IRELSNDDATI
ref|YP_001127515.1| LKQSILQHGILOPLIVRRSL---KGFEIVVGERRYRAAKEANLPSVPVVVRELTD EQMME
RAAC02367 LQANILQNGIREPLLVQRST----NILMTGHFRAVAIQAGWTEVPVQYLDVTDEEAYA
      .:  *:  *:::  ::  *.  *  .:  *.  :*.  .:  :

ref|ZP_02093159.1| LMVDSNLQR-----ENILPSERAFAYKMKLEALKNQGARSDLTSVQVAP-----KL
ref|ZP_02091713.1| LMVDSNLQR-----ENILPSERAFAYKMKLEAIERTVGRPKNVG-QVVPD---YFGKR
ref|ZP_02423704.1| VLVDNLHR-----EHILPSEKAFAYKMKAEALAHKGYRTDLTSVQVAP-----KL
ref|ZP_02026447.1| VMVDSNIQR-----EEILPSEKAYAFQMKLEAIHHKGIKG-----AE
ref|YP_001127515.1| FALLENLQR-----EDLNPIEEAMAYKMLMDKLH-----L
RAAC02367 IMLADNWERNTGIMEDYMAVARNMFYFALRLKLLNEIDDSVDISPSRKQEGTVPSFSSE
      . :  .*  .*      : :  . .      : :  . :

ref|ZP_02093159.1| STEKIGEEVGMSKDNVKRYIRLTNLVPELLDMVDEKKIAFNPAVELSYLDEAQQRDFLE-
ref|ZP_02091713.1| STEIVAEGTGESYKQVQRFIRLTNLIP ELLDMVDEKKIAFNPAVELSYLDESQQRDFLE-
ref|ZP_02423704.1| ATEQIAEDAGTSKDTIKRYIRLTNLIP EILQYVDDGRIAFTPAVELSYLNEQEYDILL--
ref|ZP_02026447.1| SREVVG EANGLSGRQVSRYIKLTNLLPELLEMVDKKKIAIKLAVEIAELSESEQQEILDY
ref|YP_001127515.1| TQEEIASRVGKSRPHIANHLRLSLPPEVQKLLIDGTL SMHGGRALLGLKKKGKMK SIVE
RAAC02367 IVHLVSKKFFERSRAVVRKHFALLKLIPELQHWISEKKIGFEGGAMLAGMSEQAQRDFMRD
      . :..      *      : :..  *  .*  **:  . :  . :..  . :  :..  : . :

ref|ZP_02093159.1| -----AMNDTQNA
ref|ZP_02091713.1| -----AMQDTQNA
ref|ZP_02423704.1| -----
ref|ZP_02026447.1| FNLGYKVS-----LEQVKAIKNKEKS
ref|YP_001127515.1| R-----TVREGLNV
RAAC02367 YADWPKISDSDIKAFRNTWESIVERQTNGLGAAEKPI SIPEPSSETLVAEWEEMRSDSV

ref|ZP_02093159.1| PSLSQA-----
ref|ZP_02091713.1| PSLSQA-----
ref|ZP_02423704.1| -----
ref|ZP_02026447.1| TDIIERPEEKTKESTKVTISRKKLKQYFP-----ENYTKAEMENI IYQLLEKWKSDG
ref|YP_001127515.1| RQLEKLVQQMNENVSRETSKRKPPEKSVF-----IRESESLRE-----
RAAC02367 PHITKYASDAYELSVVADGGRHQPMYSAEKRSISIVDVP EHKTL LARNRLALLVKKQRA

ref|ZP_02093159.1| -----
ref|ZP_02091713.1| -----
ref|ZP_02423704.1| -----
ref|ZP_02026447.1| YDI-----
ref|YP_001127515.1| -----
RAAC02367 MERMTNELSQQA EALMEEIKEGGLOEMMEDIEQLEFHTKGMLAL IENLKLHDREAVQRLS

ref|ZP_02093159.1| -----
ref|ZP_02091713.1| -----
ref|ZP_02423704.1| -----
ref|ZP_02026447.1| -----
ref|YP_001127515.1| -----
RAAC02367 DIFEFEFEDHA

```

FIG. 65

```
ref|ZP_00960984.1| -PQICKGAPMAPRKL-YAGAKLRETRSRLGLTQKDFAAARLGVSLPYLNQMENNRRPVSTT
ref|ZP_01903846.1| -----AKLRETRQRLGLTQKDFATKLGVS LPYLNQMENNRRPVSTT
ref|ZP_01880414.1| -----AKLRETRQRLGLTQKDFAGKLGVS LPYLNQMENNRRPVSTT
ref|ZP_01035782.1| -----AKLRETRQRLGLTQKDFAGKLGVS LPYLNQMENNRRPLSTT
RAAC02370 MPRLSKDALVPPEKRDTFIAWLINKRDELGFTQKDMADFLGLSLSYNAIENRKRNL SAK
ref|ZP_01742943.1| -----RRIDLGLRQGD LALKVGISPAYLN LIEHNKRRI G GK
* ** : * * : * : * : * : * : . .

ref|ZP_00960984.1| VVLALAQE-----
ref|ZP_01903846.1| VVLALAQE-----
ref|ZP_01880414.1| VVLALAQE-----
ref|ZP_01035782.1| VVLALAQE-----
RAAC02370 LVIQLASKLPQGPRFAVQLLGPEVVGKYMNL IQLPPEKEEEYEAE DALLDAIQ QREKSAT
ref|ZP_01742943.1| LLVSLAATLAVEPIALTQGGAAAVVAGLKNAASNA PQIDVEEANVDQFL-----
::: **

ref|ZP_00960984.1| -----
ref|ZP_01903846.1| -----
ref|ZP_01880414.1| -----
ref|ZP_01035782.1| -----
RAAC02370 DYLGLLRSVLHHSNAVMAVSVPNLCWFGLTPEDVAIITFPVTDLKALIPQQLVYVVDRTA
ref|ZP_01742943.1| -----

ref|ZP_00960984.1| -----
ref|ZP_01903846.1| -----
ref|ZP_01880414.1| -----
ref|ZP_01035782.1| -----
RAAC02370 QAGEFAFVRALHGKELTSVESTLMMMLDLPRDTTQLFEWGIDPRRALMSWNADDPNLMVC
ref|ZP_01742943.1| -----

ref|ZP_00960984.1| -----
ref|ZP_01903846.1| -----
ref|ZP_01880414.1| -----
ref|ZP_01035782.1| -----
RAAC02370 EVVTICKASPSRIVPLTSPGYEEWSLLERKTILOQGGIVYNRLRQEMKETIQKFRRIMATQ
ref|ZP_01742943.1| -----
```

FIG. 66

```

ref|ZP_01551668.1|
RAAC02371
ref|ZP_02756730.1|
ref|YP_521772.1|
ref|ZP_02833143.1|
ref|YP_887014.1|

-----
MRMGKADEWEARSEQRADEVSRALVRMGVMTHGQLTRLLGINPKSERDLIRRWDRFEEQ
-----

ref|ZP_01551668.1|
RAAC02371
ref|ZP_02756730.1|
ref|YP_521772.1|
ref|ZP_02833143.1|
ref|YP_887014.1|

-----
-----IKQNNFSEAIKLGESLIANSQGLDHIRINKK-----
HNVAKWVASSRYSVPGNRVRLVRLTDAALTEWAEREGIRARNRVKPRMLAQTINMGEV
-----
-----YAIIGKRIRNYRKAGYSQEALAKKAGLFHAYLG-----
-----MNIGMAREQDKLAMQSSLPTEAQRVLVREHPCKSVELLR-----
---MADLQAPEWVGAECAARRTLSPRKLSTMKAPVIFANEVATG-----
-KNPRLADDIVSAVAELGPATAGQIEAHLEAEPRGRKGPWWDRSDT-----

ref|ZP_01551668.1|
RAAC02371
ref|ZP_02756730.1|
ref|YP_521772.1|
ref|ZP_02833143.1|
ref|YP_887014.1|

-----
-----LIECSMMLSQIDKANLYLKLSLNDPDYKFYQALIYQKLSRNEKAMK---I
LIALRRDGALYDEWDMMLPQDGDGLHAWLVKQDDPDYRMGLFLLPTRLTQEEQGKKGLI
-----
-----QIERGESKASLRSIFKIANALEMPLEILFENIQNEKDPE---
-----LMGVSDQDQLDIVRWHHEVNESHGLAHNVMARLLRLADGFV---
-----LFGHLVGAIAGGAVYRKSTFLLDSLKGQILPEWLTIEEHPHL---
-----KWVAEALFAAGIILTATRVGFARHYDLAENVLPPEVLAREVGDE---
:      ::      :

ref|ZP_01551668.1|
RAAC02371
ref|ZP_02756730.1|
ref|YP_521772.1|
ref|ZP_02833143.1|
ref|YP_887014.1|

YQELIR-----SGYRNA
YHGVIRRVIQNTKIRDITLFLVPRKYITVALRLLSYIEQVGGGLYVLPFEAFKKNPGWYLD
-----

ref|ZP_01551668.1|
RAAC02371
ref|ZP_02756730.1|
ref|YP_521772.1|
ref|ZP_02833143.1|
ref|YP_887014.1|

SIYQN-----LGYEFASLG-----QHKLAQNYYTEAINI-----
SIYRGESQQRASLIDVLRPVRKLNMPLOYQFAALVQLPDGRYRMIDTYTSGDVKRVQNW
-----
-----TLSSEA-----YELIDSLTSKEQKAIKLL
-----AKMAARKTRLAMSPLGAAK
-----RKGLASSPFDSEGVRTER
-----EAVRELALRAATAL-----

ref|ZP_01551668.1|
RAAC02371
ref|ZP_02756730.1|
ref|YP_521772.1|
ref|ZP_02833143.1|
ref|YP_887014.1|

RFVLGYQVPNTGQVAGADVIVFDETMREGLDAVLRMKSKKRFDVSMVEVHAWPEGMVT
KEII EYR-----
SVFLG-----
RDIVKDGVL-----

ref|ZP_01551668.1|
RAAC02371
ref|ZP_02756730.1|
ref|YP_521772.1|
ref|ZP_02833143.1|
ref|YP_887014.1|

-----
PGRDELRFDAFDEVDDSHIDWDRAWLNEPW
-----

```

FIG. 67

```
ref|YP_324842.1| -----
ref|NP_486002.1| -----
RAAC02372      MAKRMHPVNKLNLEMLKNSNDPLDMLRQWGILMSTVYETLTQLTMNEDDE
ref|ZP_01733540.1| -----
ref|XP_660834.1| -----
sp|Q7ZXB1|MCM7B_XENLA -----GL

ref|YP_324842.1| -----
ref|NP_486002.1| -----
RAAC02372      LVTEYLSGI-EELDDKIQEYLYEEAAARIAHFDERTSKVVFDQWPTGLCL
ref|ZP_01733540.1| -----NPVHHYKMNVLGLQQNDYADFIVVEDLVNFKVSKTF
ref|XP_660834.1| -----SRNDQELAKYGAVALLRLAMLVGALVDAEERRQELQ---
sp|Q7ZXB1|MCM7B_XENLA LSETYLESHRLVKMNKTEDDELGTEELSEEELRQITEEDFYEKLAASIAP

ref|YP_324842.1| -----HGSKDQQHPLYNR---DRPF-I-----
ref|NP_486002.1| -----HGSKDQQHPLYNR---DRPF-I-----
RAAC02372      VPSNFGFEIRNLPVWMLGHGAKRKHHPPLYHR---DRPF-IRRRQLWYWL
ref|ZP_01733540.1| -----ELVAENGISNVKHPFDR---PNNFNITTKTISDFE
ref|XP_660834.1| -----GNGRSVSYSIAWMS---EEQQ---QRLQEIL
sp|Q7ZXB1|MCM7B_XENLA -----EYGHEDVKKALLLLLVGGVDHSPRGMKIRGNINVC
                        :

ref|YP_324842.1| -----
ref|NP_486002.1| -----
RAAC02372      VHELKWNLYLNSVASWDDSLLPKNPPRYAHMNIRFMSRNDYEIRDLNRYV
ref|ZP_01733540.1| -----IES-TASKIRVIE-----
ref|XP_660834.1| -----GQRHPETYLSVRY-----
sp|Q7ZXB1|MCM7B_XENLA LMGDPGVAKSQLLSYIDR-----

ref|YP_324842.1| -----
ref|NP_486002.1| -----
RAAC02372      SLPMIVNALVANGLIQSDRPGYFTYSVEWVQKHEQDEVQEPTTTLRVRYL
ref|ZP_01733540.1| -----
ref|XP_660834.1| -----
sp|Q7ZXB1|MCM7B_XENLA -----

ref|YP_324842.1| -----
ref|NP_486002.1| -----
RAAC02372      EREPNVSTL
ref|ZP_01733540.1| -----
ref|XP_660834.1| -----
sp|Q7ZXB1|MCM7B_XENLA -----
```


FIG. 68

```

ref|YP_001236354.1|      -----VVRTGAAMTIRFRRITDG---DLPAIRNWLH-QPHIRQW---WGDPEQY
RAAC02296                MGEFMCHPPVLRIGARREIRGRRIREGRHPGLSDLREFLLTEPHIRHR----QADHDDHH
ref|XP_001563017.1|      -----HLSSHVAER-NDQHER---VDDWRGVPPKLPGSSYR----YEERHPQQ
ref|ZP_01776409.1|      ----LCQPP--RLHDRRRHRGRRLREQG--RPGQIRGVLDLHQDGRQQPLEAERDQPEVT
ref|ZP_01565636.1|      -----RREQHDQRQHGR---RADFHPVLLADEHHGQR---AADEQDHR
ref|XP_001615133.1|      -----KSSLKGAKKKKGRGSYTDDLAEHETTWEKKRR---RSAEGDNHD
                           :
ref|YP_001236354.1|      GLIDGDRDEP-----
RAAC02296                EQPDDDLPHPPR---LLSCST-IRFESEHFTPRVRHIFIHLSRA
ref|XP_001563017.1|      RHMAHHHHQPA-----
ref|ZP_01776409.1|      APPDIGLKPPLKGTGSARCRSPFSELAARLRPKKWRLFVELH---
ref|ZP_01565636.1|      PREPREQPHPQR-----QVRAARFE-----
ref|XP_001615133.1|      EQEDDEEYEPKR-----KVPRKMDDDYEHT-----
                           *

```

FIG. 69

```

ref|YP_001376930.1|
RAAC02373
gb|AAW81277.1|
ref|YP_001312077.1|
ref|ZP_02852259.1|
ref|YP_001642790.1|

-----KLKTYKVEG
MLTIGLDVGNISIGLCVRDGTTLVQDTTPSVYGRVDPTRQVLSVPGKSAPRKVDVFTFGG
-----

---VGIDLGYG YVKFIDGKEP---KMFPSVVG YGNSQKYKSALQLDLNPLDDLQIKIGD
-----VYTYEG

TEYVWGDDIIKVN--TLNTYAQQNRYKTNQYKTLISKIALAEMAATNVKS--YDEILVI
EHFVLGYKNVHAMHSTPIGAYDREQRYASRQFETLAKLALLDAATRTGRTG--VIEVVVA
-----RYNSEAFKLLANFALGLLASDFKIANNQVLEVVT
-----VRVI
EHFFIGDLAIRQSEVASRSLGKDRSQDKNARVLMLTALSLLSSWDKQ-----GFNLV
IRYVIGEAGQ-----ISSARNDDRYSSAGYRTETILAISSQLVKDGS-----EIVVG
. .

TGVPSSEIG---TKAVDEIKEVYQGAHDLVNGKKVSINVDVIVLAQFVGTVMRSRYLDE
CGTPSEDFT---TRTVEIMQRWFSEPV TGAKNGEQVVMIKRLEVIPQPFVFLDAYLDQ
AGLPTGDYAD--QERLRSLLKVLGQHQVTIDQIVTVRVRKVYILPQPIGTLYNELDD
TGLPAQFFA---EQKNSLIKALENRRVFMKLNGENRSFTITKVIVFPQSAG-----
TGLPTNFYAAFAEEWESTLNGEFKTKMKIGGKTQERSFQIEEVTTLPPFGTLYDQVLNS
TGLPSEDYKN--GDNHEKVKRNLVGEHTVQIDGKTKTFSILRVYTPMQPIGSVNVRIYD
* *: . : . : * .

DGFVADDTYED--MTVGIIDIGTGTDLDVISMLRREKE-STSVPKGMHDVYEPIVAKIK
DGLVVDDEELEK--QDVLVIDSGSGTLDLSEIHRLELTR--QTSIPAGLNDVYQLILEEIR
QGFIKNKALLD--EKVGIVDVGGGTILIDTILNFELSGKNRQQFNTGVNDLYEAIASRIE
-LFLYDKSLVE--KDTLVVDIGGGTLDIAYMSNGQFKEG--RTYPLGVNPTYDVLLQELT
VGKVVDRDLTD--SKIGIVDIGFKTTDLAVSDGMEFINPLSFSTTGLSNVNLVNEKLR
NLKVRKDMESERTARKLVIDIGFGTTDVCEAEGLRIVRY--DGVQVGMLEANRIKDELS
. . : * * : . * : .

K-----ETS---ATINDYKLEKVFEEGAYQASKRMDPID-FNDE--KTASIKE
R-----EEPKVYATAYDLEAQLRAQDGAQEFWFEYGALR-MNITNLRERAMRQ
G-----DVS----LYQLEKELRHGNQQHQWSYRFSKNRQDDITELVGKESDR
K-----YGVN-----YSNRMKAEQIIADKAI FVEGKE-IDVSKDIDNVLSL
H-----EFKID-----REEHQLDDCINSQKIMVAGKS-EDISSWVREALQT
KRGARGIVSLLHMDTLLRNAKREYVKDEFTDKEILSKVIEIGGKE-YEIKDLMEQALEY
. : .

VYDFIVNGVNNAWKTFDRFDEVLSVDGGANTF-HELLEEWIGKVTKLEES--QTANVEGF
VWDRMQQGIQYAYPDRSSFGFRVILAGGSGEAF-RNYFLAWMPSIRIAPEP--QLAVARGL
FTRRLVANVTSTLKNLDSIDTLFFTGGSANLINQKILKTTFTNAAIVKDT--EVANVNGF
RAGEIINAIAKQAFPEQSKYSRFRVFIGGGALLLKNY-----
VSDKISVEIESKWDYRD-FDTLLLTGGGGEML-YPYLKDKFPNLVLVEDP--QTANVRGY
TARIVMQRVNDNLGYVLKDYDVVLTGGGSLAL-HKYIKPYLTGVNTKAEQGAQTANVKGY
: : . . . . * .

YRYGKFEVGEE-
YKYALAQGAEES
YKYGLSQQVQE-
-----
QKLA-----
TKYAMIQDAK--

```

FIG. 70

```
ref|XP_001467069.1| LVSTWPFQSRRLFLESRIQLQTADMQAEDMISGDGGS-----QHIEDDSVSVTPSTS-----
ref|YP_001208199.1| -----QGAGGL-----SGMQSSSPSSSSSTP-----
ref|YP_946581.1| -----RRLAERLTARLNEIYAERQLMAVGGSEWAAPGAAPQSNPGIPASTA-----
ref|YP_001376929.1| -----MANKTYLLSYDDVLDKDIKEWLESL-----PRNRKAEMVRAI-----
RAAC02374 -----MTIRRYVLSYDDEYDREIAEVIEQT-----PKRRRAERVRLILLGIQAEA
gb|EAU86007.1| -----QTVIVKRRRAKPEPTGYDGNPTNAAASP-----AFSTTTQLTASPPSL-----

ref|XP_001467069.1| -----APS-PAPRSAAATGAPPLAPR-----
ref|YP_001208199.1| -----SSTGLEPRSGGAGGGQPLMPNVRIPT
ref|YP_946581.1| -----TPAHMMAQAGSSTGADPVLAR-----
ref|YP_001376929.1| -----
RAAC02374 EMRGEGGMKPVQTGSPASDDSSIPSPQTPQSGSSTGKRPVIPRITPKS
gb|EAU86007.1| -----FPNSPPTGTSSGIRPLIPHIVPPS
```

FIG. 71

```
ref|XP_975359.1|
RAAC02375
ref|ZP_02840410.1|
ref|ZP_01467536.1|
ref|YP_001016790.1|
ref|XP_001315633.1|

-----
MCKMWTKRWTDFLAVSILTATVTGCGTIPHAVSQTSSTAI PAPSVHAVVDRLPGHLQAKL
-----
-----IPAKAGEAHLERLLRHVHSR-
-----QGKLNRIDVASKRVTPIPFHV
-----PLRLRCWRISMRTMVQRAAAE
-----LQAEFINFENAEN

ref|XP_975359.1|
RAAC02375
ref|ZP_02840410.1|
ref|ZP_01467536.1|
ref|YP_001016790.1|
ref|XP_001315633.1|

-----RRKSKEENK
IGFDVRKKRVSHLPIAVKGTNVLYISPMEQYAIQQFQQIWPRLKAYPVVWTVGVTQAQAE
AGLEAAPSRIED-----QLDHV-----ARTVKGRNL
KGTRTLFARVHS-----PQAQAVVPERFPVKMLRWVQVSPKGDR
LGIEVIPSRRTY-----ALLDWLAERE-----RDVYPLEEG
IGFSLKQS-----SKLVIS-----EQRSNKNYQ

ref|XP_975359.1|
RAAC02375
ref|ZP_02840410.1|
ref|ZP_01467536.1|
ref|YP_001016790.1|
ref|XP_001315633.1|

LDREKSDYAPESVPIVGEE-----NGVHYEDGHFWMEVPG-LLPESEED-----
MVWRKEGYPGDPLPSQQTEYVSQTIPTPDAYHRDGKAWVEVPG-ILPASQLEDWVTFFR
LFVVADEIPASPALQLLR-----RLRAQHEILWLTIR-----DAE-----
VVYQALGHLYTRELPDGLP-----RRLTKQTEHFEMHPS-FSRDGKSIV-----
YMAGPLAPPPTPIPTFPVP-----LPEAVRGDAWSWASLPLGLLREAQ-----
KFWQRL-WSAMTMPLALWEG-----ATSKKIEMIYQRD-----
:
```

FIG. 72

```
ref|ZP_01091610.1| -----
RAAC03273      MDMAWEPSQDPYTLLQDFSCFYGTDVKGPNSVLTQKDMMAIRNNIVEVSRGWRQLGPNKV
ref|YP_944003.1| -----
ref|YP_502758.1| -----
ref|YP_462360.1| -----
gb|EAU81483.1| -----

ref|ZP_01091610.1| -----
RAAC03273      ELLEPLMDPSAVSGAFV-YLAKHHEDWSLSDWPKVYLAFIRTVDSQTLTFEPNGNVIYTY
ref|YP_944003.1| -----
ref|YP_502758.1| -----PSEVSGNFLSALNKLSDDYSKIITSIESGFSTNLDSIIQLLERKGQIILYG
ref|YP_462360.1| -----
gb|EAU81483.1| -----

ref|ZP_01091610.1| -----RL
RAAC03273      KRGTGLTPNVTGVFEKGVDTVYSSYSYSDQIGVEAYKEDDPSLNSKIFYEPDYQVFSPKT
ref|YP_944003.1| -----
ref|YP_502758.1| -----PPGTGKT-----YSVQSLIK
ref|YP_462360.1| -----HSYR
gb|EAU81483.1| -----

ref|ZP_01091610.1| HRSNYTYTYT---ASGGTLSFSQWACVALTFDG-----YQYRFYKNGFLISTH
RAAC03273      HLSNYHYSYPGYVPPSNGQIAFELWGTVCDPYTARVGGAI FIEGMYEFTYY-NGHLVNAH
ref|YP_944003.1| -----SYSYN---DEVGTLDY-ILASASLK-----DKIVDSMYWNIN
ref|YP_502758.1| KTASTPYIST---TKDYGNIQFFWLITYKDKYDDIS-----TINSKSIYTYQWKS
ref|YP_462360.1| AGAGLTYNLT--EISDLG-LDYAYSKSDYDEEDK-----VDSTKDTVSMAYR
gb|EAU81483.1| ----YPGYIP----PSSAQIESALRGTIYDALIRR-----IMWEGVRGKWGVPST
. . .

ref|ZP_01091610.1| GIGGSFVNNPAASVWIGD---
RAAC03273      QIGQESENYPSEVTRYGD---
ref|YP_944003.1| SSESSLFEYSTKYTGYP---
ref|YP_502758.1| YNYSKYFYMINEGDYFAIYLW
ref|YP_462360.1| HFFRERRDIVTVTPFYSRY--
gb|EAU81483.1| HQGEDGEERYSEIDTWG---
```

FIG. 73

```

ref|YP_074959.1|
ref|ZP_01846154.1|
ref|YP_594046.1|
RAAC02967
sp|Q45618|TRA6_BACST
ref|YP_828009.1|
-----
NTEGRASECATALYTICREGINMETHYLBACTERIUMSP-GBACAINTEGRASECATALY
-----

ref|YP_074959.1|
ref|ZP_01846154.1|
ref|YP_594046.1|
RAAC02967
sp|Q45618|TRA6_BACST
ref|YP_828009.1|
-----
TICREGINMETHYLBACTERIUMSP-GBACAHYPHETICALPRTEINMMETHYLBACTE
-----

ref|YP_074959.1|
ref|ZP_01846154.1|
ref|YP_594046.1|
RAAC02967
sp|Q45618|TRA6_BACST
ref|YP_828009.1|
-----MAP-----YRPMAKAAATVRF
RIUMSP-SCRESIGNIFICANCEE-IDENTITIESPSITIVESGAPSLRPSAPLPIERRF
---VRPFRR-----TQVSAARVTTRF
---MKPLR-----PVVSAKATERF
---MKPFR-----ETAKKKYTVRY
---PWR-----EEGRERAFVRF
*                                     *:

ref|YP_074959.1|
ref|ZP_01846154.1|
ref|YP_594046.1|
RAAC02967
sp|Q45618|TRA6_BACST
ref|YP_828009.1|
ETPPGKQAQVDWADFGYIEVD--GRRLKLYCFIMVLAYS RAMYLEFVTATDMKTFMRCHI
ETPPGEQAQVDLAREFEVVFADPEGVTRIVWL FAMVLGHSRYLWARFVHQLQTVLRCHI
ETAPGEQAQVDFGRYSYLNLE--GQTRSIWAFVMVLGWSRALYVEFIRKADTASFIRCHL
ESDPGEQAQIDLGAFLYYDSH--GQRRTIWAFAMVLAYS RMLYVEFIKAADQLHILQALR
ETLPGEQMQVDWKEVGEVVIE--GKKVKLSLFVATLGYSRMKYAVFTTSQDQEHLMECLI
ETGPGEQSQMDWGHEFGN--WD--GR--RLYGFALTLCYSRMRYIEFTQSQDIHLLACMV
*: **: * *: * . * : * . * ** : * * * . : .

ref|YP_074959.1|
ref|ZP_01846154.1|
ref|YP_594046.1|
RAAC02967
sp|Q45618|TRA6_BACST
ref|YP_828009.1|
NAFKFFGGVPHEILYDNVKT VVKDRDDNRPFVNERFLDFSSYYGFRPRLCLPYHAWTKG
AAFQALGGAPREILYDRMKTAVIGEDPDGLVIYNRSLDLARHYGFLPRACRPYRAKTKG
NAFAYFGGMTQSILYDNTKQVVLERDETGPVWNPQFLDFSLRLGFSIRLCRPYRPRTKG
NDLEFFGGVPRVMSDNCSP LVVANDGQGHVDWQPAYLDFAKFYGFVPKACRPRRSRTKG
QSFKYFGGVPKKVLFDMNKT VTDGREQ-GVVKWNQRFSEFASYYGFIPKVC RPYPYRAQTKG
HAFRYFGGVTVESVLTDRMKT VLDQGTG-GELHFNQKFLQFAAYYGFVPRVCRPYRPETKG
: : ** . : * * . . . . : : : : ** : * * : . ***

ref|YP_074959.1|
ref|ZP_01846154.1|
ref|YP_594046.1|
RAAC02967
sp|Q45618|TRA6_BACST
ref|YP_828009.1|
KVERPVRYIRQNFW---
KVERPFRYLREDFW---
KVESGVGYVEKNFW---
KVERPIRYIRDSFWPVA
KVERAIQYIMDHFY---
KIESSVRFVKQNFWP--
*: * . : : . * :

```

FIG. 74

```

ref|YP_148969.1| -----YLQLAHNEWDPKAKYAKAKVIYSFGREDEVDRAVLER
ref|YP_001126171.1| -----YLQLAHNEWDPKAKYAKAKVIYSFGREDEVDRAVLER
ref|YP_146154.1| -----
ref|YP_146741.1| -----MYIRRVTRKNKDGTTVAYLQLAHNEWDPKAKYAKAKVIYSFGREDEVDRAVLER
RAAC03589 MPYHVGMYIRVIRRNKNGSVTGYVQLAHNFRDPNTGQPKAKVLYTFGREDEIDLEALRR
ref|ZP_02130848.1| -----MYIRTISRKNKDGSKVEYVQLAHNYRDPKSKQARAEVLYSFGRKDQLDMEAIRR

ref|YP_148969.1| LAKSISRFLSPEQAWIEITLTGEVSDDFQFQSSKRLGGAWLLDQLWRQLGLGEILHSLFA
ref|YP_001126171.1| LAKSISRFLSPEQAWIEITLTGEVSDDFQFQSSKRLGGAWLLDQLWRQLGLGEILHSLFA
ref|YP_146154.1| -----VC
ref|YP_146741.1| LAKSISRFLSPEQAVEVEKLTGEASDDFQFQSKHLGGVWLLDQLWRQLGLGEILHSLFT
RAAC03589 LAQSIHRFVG--DEFTAGRGQSERIQTTLLDS-RPMGGAYLLDELWRQLGLDEVLRERLA
ref|ZP_02130848.1| LAKSVRFIAKTGDVETQCKLQFPGEDVRFVESRPMGGVFVLKKIWDRLRISECLDKALA

ref|YP_148969.1| SRHHQIPLERLIFAMVANRALHPSSKLAMEEWVEKDVH IPHLPQVASHQLYRAMDELLAV
ref|YP_001126171.1| SRHHQIPLERLIFAMVANRALHPSSKLAMEEWVEKDVH IPHLPQGASHQLYRAMDELLAV
ref|YP_146154.1| SRHHQIPLERLIFAMVANRALHPSSKLAMEEWVEKDVH IPHLPQVASHQLYRAMDELLAV
ref|YP_146741.1| SRHHQISLERLIFAMVANRALHPSSKLAMEEWVEKDVY IPHLPQVASHQLYRAMDELLAV
RAAC03589 DRKFKAVERVIFAMVANRALAPSSKLAMEEWVDREVALPGMTELDVWQAYRAMDFLHDV
ref|ZP_02130848.1| DRQYTAPIGDAVFAMVANRALAPDSKLAVEDWAAKDVHLELDQPLKVQHLYRAMDFLEN
.*.: : :***** *.*****:*. :*: : :***** *

ref|YP_148969.1| QPELERQVFHAVADLLNLEVDLIYFDTTSSYFEVDPSETPEGESLRKQGFSGDKRPDLVQ
ref|YP_001126171.1| QPELERQVFHAVADLLNLEVDLIYFDTTSSYFEVDPSETPEGESLRKQGFSGDKRPDLVQ
ref|YP_146154.1| QPELERQVFHAVADLLNLEVDLIYFDTTSSYFEVDPSETPEGESLRKQGFSGDKRPDLVQ
ref|YP_146741.1| QPELERQVFHAVADLLNLEVDLIYFDTTSSYFEVDPSETPEGESLRKQGFSGDKRPDLVQ
RAAC03589 AEDLQYEVFRRVSDLLNLDVLLFFDFTTSTYFET---EDESDDSLRRKGYSKDHPRDLFQ
ref|ZP_02130848.1| QEAIQKEVFWSTANLLNLEVDLVFFDFTTSTYFER---DEEDEEGLKRYGHSKDKRKDLFQ
: : ** : :*****:*****:*** : . :*.: *.***** * ** *

ref|YP_148969.1| IVIGLAVTREGVPIRAWVWPGNTMDMTVIKQVKQDLIGWKLGRVISVMDRGFSSEENLRI
ref|YP_001126171.1| IVIGLAVTREGVPIRAWVWPGNTMDMTVIKQVKQDLIGWKLGRVISVMDRGFSSEENLRI
ref|YP_146154.1| IVIGLAVTREGVPIRAWVWPGNTMDMTVIKQVKQDLIGWKLGRVISVMDRGFSSEENLRI
ref|YP_146741.1| IVIGLAVTREGVPIRAWVWPGNTMDMTVIKQVKQDLIGWKLGRVISVMDRGFSSEENLRI
RAAC03589 VVIGLAVTRDGIPIVRCWTWPGNTADMSVVEEVKQDLIGWRLGRVITVVDGRGVSESNLRI
ref|ZP_02130848.1| VVGLAVTKEGLPIRSWVFPNTPDVNTVEQIQKEMNDWKLGRVWVWAMDRGMTSEENRAI
*:*****:*.*****:*.***** *:.....: :*.*****: :*****: **.* **

ref|YP_148969.1| LQQAGGHYIVGEKMRSGKAAVEEALNRRGRYQQVRENLHIKEIIVGDGEARQRYVLVYNP
ref|YP_001126171.1| LQQAGGHYIVGEKMRSGKAAVEEALNRRGRYQQVRENLHIKEIIVGDGEARQRYVLVYNP
ref|YP_146154.1| LQQAGGHYIVGEKMRSGKAAVEEALNRRGRYQQVRENLHIKEIIVGDGEARQRYVLVYNP
ref|YP_146741.1| LQQAGGHYIVGENMRSGKAAVEEALSRRGRYHEVDENLHIKEIIVGDGEARQRYVLVYNP
RAAC03589 LQRAGGHCIAGEKMTSGKPAVEAALARPGRFRELRPNLKVEVVGDGEARVRYVLAFNP
ref|ZP_02130848.1| LQRGGGNYILGEKLR-GSNMSKAVLSPGRFTTVRDNLIEIKVETAGDGACRRRYVIVRNP
**:.**:* ***: * . :.* ***: : **:.*** .* ***: . **

ref|YP_148969.1| SEAEQRKEREKLLSILKEELEGRLQLPNEAHHKATCRLRSHPSYGYLRQLKDGTLRID
ref|YP_001126171.1| SEAEQRKEREKLLSILKEELEGRLQLPNEAHHKATCRLRSHPSYGYLRQLKDGTLRID
ref|YP_146154.1| SEAEQRKEREKLLSILKEELEGRLQLPNEAHHKATCRLRSHPSYGYLRQLKDGTLRID
ref|YP_146741.1| GEAEQRKERETLLESILKEELEGRLQLPNEAHHKATCRLRSHPSYGYLRQLKDGTLRID
RAAC03589 EEAKRDEARREAMLRELRLMELERLKEIQGEAHTKAHCRLASHPTFKRYLQDRWGNLRID
ref|ZP_02130848.1| KQVKRDQATRERLIRAEQIEAIGDLTGKKHTKAACALLSHRSMGKYVRELKSGKLGKIN
:.*.: ** :. .*: :* :* ** * ***: :*.: :*.*.:

ref|YP_148969.1| KQAVREAKEYDGKYLIRTSDDTSLAEDVALGYQLVDIEQAFRTLK-----
ref|YP_001126171.1| KQAVREAKEYDGKYLIRTSDDTSLAEDVALGYQLVDIEQAFRTLK-----
ref|YP_146154.1| KQAVREAKEYDGKYLIRTSDDTSLAEDVALGYQLVDIEQAFRTLK-----
ref|YP_146741.1| KQAVRDAKEYDGNYLIRTSDDTSLAEDVAIGYQLVDIEQAFRTLK-----
RAAC03589 PEAVRQAAHLDGKYLIRTSDDTSLTEDVALGYQLLMVESALRTLKKFWTSVP
ref|ZP_02130848.1| KAKITEEEKLDGKYLSCSDDTSLPEEIALGYQLLEVERAFRTLK-----
: : :**.*: *****:*****: ** *.*

```

FIG. 75

```
ref|ZP_01665148.1| -----
ref|YP_001111903.1| -----
ref|YP_753434.1| -----
ref|YP_753226.1| EFYPHYPTHETICALPRTEINSWLSYNTRPHMNASWLFEISUBSPWLFEISTRGETTING
ref|YP_752864.1| -----
RAAC03695 -----

ref|ZP_01665148.1| -----
ref|YP_001111903.1| -----
ref|YP_753434.1| -----
ref|YP_753226.1| ENGBABIHYPTHETICALPRTEINSWLSYNTRPHMNASWLFEISUBSPWLFEISTRGETT
ref|YP_752864.1| -----
RAAC03695 -----

ref|ZP_01665148.1| -----
ref|YP_001111903.1| -----
ref|YP_753434.1| -----
ref|YP_753226.1| INGENGBABIHYPTHETICALPRTEINSWLSYNTRPHMNASWLFEISUBSPWLFEISTRG
ref|YP_752864.1| -----
RAAC03695 -----

ref|ZP_01665148.1| -----
ref|YP_001111903.1| -----
ref|YP_753434.1| -----
ref|YP_753226.1| ETTINGENGBABIHYPTHETICALPRTEINSWLSYNTRPHMNASWLFEISUBSPWLFEIS
ref|YP_752864.1| -----
RAAC03695 -----

ref|ZP_01665148.1| -----
ref|YP_001111903.1| -----
ref|YP_753434.1| -----
ref|YP_753226.1| TRGETTINGENGBABIHYPTHETICALPRTEINSWLSYNTRPHMNASWLFEISUBSPWLFE
ref|YP_752864.1| -----
RAAC03695 -----

ref|ZP_01665148.1| -----EAXTRANSPSASEISFAMILYPRTEINTHERMSINU-----
ref|YP_001111903.1| -----
ref|YP_753434.1| -----
ref|YP_753226.1| EISTRGETTINGENGBABIHYPTHETICALPRTEINSWLSYNTRPHMNASWLFEISUBSP
ref|YP_752864.1| -----
RAAC03695 -----

ref|ZP_01665148.1| -----SCARBXYDIVRANSNRGBEAXTRANSPSASEISFAMILYPRTEINTHERMSINU
ref|YP_001111903.1| -----
ref|YP_753434.1| -----IHYPTHETICALPRTEINSWLSYNTRPHMNASWLFEISU
ref|YP_753226.1| WLFEISTRGETTINGENGBABIHYPTHETICALPRTEINSWLSYNTRPHMNASWLFEISU
ref|YP_752864.1| -----
RAAC03695 -----MLPRRFFMDN-----

ref|ZP_01665148.1| -----SCARBXYDIVRANSNRSCRESIGNIFICANCEE-IDENTITIESPSITIVESGAPS
ref|YP_001111903.1| -----
ref|YP_753434.1| BSPWLFEISTRGETTINGENSCRESIGNIFICANCEE-IDENTITIESPSITIVESGAP-
ref|YP_753226.1| BSPWLFEISTRGETTINGENSCRESIGNIFICANCEE-IDENTITIESPSITIVESGAP-
ref|YP_752864.1| -----
RAAC03695 -----TVDFLEPFAESRVLVRQG----FFEHKGLHPNLRERKWYPANHFKGAKSN
```


ref|ZP_01665148.1|
 ref|YP_001111903.1|
 ref|YP_753434.1|
 ref|YP_753226.1|
 ref|YP_752864.1|
 RAAC03695

MYIRQKPLFSFDITLMQYQPKTRLAMVFESIDLHPLLKTLPIKSIR--GPKGYSSAALIKA
 MYLLQPNLFSFEELLKFEPETKLQKVLSDLSPALN-VVKRAVL--GPKGHCVGNMIRA
 -SIRQNCIFSFEDALKIQPKSRLEKIINTLDLKPVLCKLDKPGEIRVGPKPYPAYAMLNA
 -SIRQNCIFSFEDALKIQPKSRLEKIINTLDLKPVLCKLDKPGEIRVGPKPYPAYAMLNA
 --IRQGYVFSFEDAINLQPRSRLEIVLATLDFDDVITVLDKENKQHRGPTGYPFESKLNA
 VYVRQAWLFSFDEWMEMSPCERRELFSTLDLSPYAAKLRSSTPQ--GAKPISREAILRA
 : * :***: :: . * : . : **: : *.. :.*

ref|ZP_01665148.1|
 ref|YP_001111903.1|
 ref|YP_753434.1|
 ref|YP_753226.1|
 ref|YP_752864.1|
 RAAC03695

FLAMRLCSIPTVTLLVERLKTDLVFRYECGFSLEQVPVSLATFSRFFQKIAETDS-----
 LVAKQLEQIPTVAALVKRLSNDIRFRFQCGFSLSKPIPSESTFSRLIQKLAATDETNKTE
 LIAMRLENMGFTTQLVERLTYDPLHRYVCGFEPFGTAPSKSCFSRFYSKLAQSGC-----
 LIAMRLENMGFTTQLVERLTYDPLHRYVCGFEPFGTAPSKSCFSRFYSKLAQSGC-----
 LIAMRVYNMATFTTELVERLTHDPVLRVNCGFDVFGKVPSIATFSRFYERLTQSEV-----
 FLAAPLEGISTFTQLHRRLESDLRFYQCGFSLHESIPSVSTLSRVFQAIVDKGV-----
 ::* : :*. : * .** * :*: ***. ** : :*. . :. .

ref|ZP_01665148.1|
 ref|YP_001111903.1|
 ref|YP_753434.1|
 ref|YP_753226.1|
 ref|YP_752864.1|
 RAAC03695

--LQVLFSAVDTAIQDKVISGEVVAIDASAIDSYEKVPVKKELNSNGDSATWGAKLDT
 SVMKQIFDNLVKSADMGIDSNCAIDSSKIDAYEKSPPKDLQGD-KTANWGAKRDT
 --LETLFTSLVKQAEEMGLLDLSSVAIDATKVEAYEKSVPKNI IQDGNVADWGIKSDTN
 --LETLFTSLVKQAEEMGLLDLSSVAIDATKVEAYEKSVPKNI IQDGNVADWGIKSDTN
 --LRELFFKKQVTTAESMGLIDTSSIAIDASKVDANEKSVPRKNI KDDGQSANWGSKLDTN
 --AATLFAELVRQCRDEGLIEGEHVAIDSTAIHAYERKHPRSGVQPS-NRANWGAKFDAF
 : * * . . :. . :***: :. : * : *.. : . . * * * * *:

ref|ZP_01665148.1|
 ref|YP_001111903.1|
 ref|YP_753434.1|
 ref|YP_753226.1|
 ref|YP_752864.1|
 RAAC03695

GNQHVWFGYKHLHVAIDTKSELPIAVKVTANRNDVTQAIPLMDEIK-----HQPKYXCM
 GNQISWFGYKAHVAIDCHSELPIAIMVTANRNDVTQAIPLMDEIK-----HQPKYXCM
 GNPIKWFGYKHLHIGTDVKSGLPIAMKVTANRNDVTQAIPLMDEIK-----HQPKYXCM
 GNPIKWFGYKHLHIGTDVKSGLPIAMKVTANRNDVTQAIPLMDEIK-----HQPKYXCM
 GNQITWFGYKHLHIGTDVKSGLPIAMKVTANRNDVTQAIPLMDEIK-----HQPKYXCM
 GNKLAWFGYKHLHIGTDVKSGLPIAMKVTANRNDVTQAIPLMDEIK-----HQPKYXCM
 ** ***** *..* : * * * : :***. * * : : : : : *

ref|ZP_01665148.1|
 ref|YP_001111903.1|
 ref|YP_753434.1|
 ref|YP_753226.1|
 ref|YP_752864.1|
 RAAC03695

DMGYDAKDQVYQAA-ERQAQAIIPNRRKEKLPPEGMDENRTPCTSMGYPMVYWGECEREK
 DMGYDSKDIYSVVMNDFNAQAIIPNRRKEKLPPEGMDENRTPCTSMGYPMVYWGECEREK
 DAGYDHREIYSVIRDKYHAQAIIPNRRKEKLPPEGMDENRTPCTSMGYPMVYWGECEREK
 DAGYDHREIYSVIRDKYHAQAIIPNRRKEKLPPEGMDENRTPCTSMGYPMVYWGECEREK
 DAGYDQKSIYELIRDKYHAQAIIPNRRKEKLPPEGMDENRTPCTSMGYPMVYWGECEREK
 DAGYDQTKNYEAAAR-ALGAQAIIPNRRKEKLPPEGMDENRTPCTSMGYPMVYWGECEREK
 * * * * . * . *****. * * * * * : * * * : *..**

ref|ZP_01665148.1|
 ref|YP_001111903.1|
 ref|YP_753434.1|
 ref|YP_753226.1|
 ref|YP_752864.1|
 RAAC03695

GILKFRCPHVCGKVNCPNGSAWCSPSNYGLVIKKKVEDDPRSFCTPHRGTRWEKLYAER
 GTNKYRCFPHVMGKCDPCFYSWCSPPSYGLVVKTKVKDDPRMNCIPARGTKNWQSLYNKR
 GVNKFRCPHIMGKCDPCFYSWCSPPSYGLVVKTKVKDDPRMNCIPARGTKNWQSLYNKR
 GVNKFRCPHIMGKCDPCFYSWCSPPSYGLVVKTKVKDDPRMNCIPARGTKNWQSLYNKR
 GVNKFRCPHIMGKCDPCFYSWCSPPSYGLVVKTKVKDDPRMNCIPARGTKNWQSLYNKR
 GVNKFRCPHIMGKCDPCFYSWCSPPSYGLVVKTKVKDDPRMNCIPARGTKNWQSLYNKR
 AWLKFRCPHATGQVDCFLGMAACSASYGMVVKKHIDEDVRRYANPHRGSRWTKMLYDER
 . * :*** * : : * * : * * * * : * * : * * * : * * *

ref|ZP_01665148.1|
 ref|YP_001111903.1|
 ref|YP_753434.1|
 ref|YP_753226.1|
 ref|YP_752864.1|
 RAAC03695

TSVERAFSRLKEQLGANT-VRVQGIKKVTAHLMLCCIALLAGTIAVNR-----
 TSVERCFGRKQHLGANS-IRTRGLEKVTLHITLSCIALLAGTIAVAKTKRIEQAA
 TYSERCFGRKQHLGANS-IRTRGLEKVTLHITLSCIALLAGTIAVAKTKRIEQAA
 TYSERCFGRKQHLGANS-IRTRGLEKVTLHITLSCIALLAGTIAVAKTKRIEQAA
 TSVERCFGRKQHLGANS-IRTRGLEKVTLHITLSCIALLAGTIAVAKTKRIEQAA
 TAVERCFGRKQHLGANS-IRTRGLEKVTLHITLSCIALLAGTIAVAKTKRIEQAA
 * * * . * : * : : : * * : * : * : * : * : * : *

FIG. 76

```

ref|YP_754943.1|-----LNAKVWYEEQFRLSRQKQFGASSEK-----
ref|YP_754865.1|ANCEE-IDENTITIESPSITIVESGAPSLNAKVWYEEQYRLSRQKQFGTSSEK-----
ref|YP_431166.1|-----MRQOVAYLEEQIHLLRHRLFGASSEKRRKTQA-----
RAAC02318-----MQKKKFGTSSEK-----T
ref|ZP_02171171.1|-----IKQQNQWLMEQFRLKHKQYGASSEQ-----
ref|YP_519650.1|

ref|YP_754943.1|---TTPEQINLFNEAEDITDPKLEEPS-----LETVTYQRKKKQAGQREDKDKD
ref|YP_754865.1|---TTPEQINLFNEAEDIVDPEIKEPD-----IETITYERKKKQPGQKADKDKD
ref|YP_431166.1|-----
RAAC02318ESDSVQLSLFNEAEVEADAQSSEETGEADTKAPSEDVETETITYERRKPRAVRERDAWLY
ref|ZP_02171171.1|DERFEQGSFLFNESEKEQDAEEEEPT-----VEAITYERKKKRRKARKD---LT
ref|YP_519650.1|---HSAEQMDLFNEAEATAELSAPEPP-----LTEVKTTHYRKRTRLTTDK---LP

ref|YP_754943.1|L--PVEIIEYRLEKHEQICPCCQGELHEMSIQVRHEIKIIPAQAISVKHVQYIYACRRCE
ref|YP_754865.1|L--PVEVIEYRLLLEHEQVCPCCQGSLEHMSIQIRQEIQVIPAQVKVQHVQYIYSCRQCE
ref|YP_431166.1|-----MSTEVRQELKIIPAQVKVVKHIRYVYACRHCE
RAAC02318QGEADEVVVEYRLSDDERVCPCAGELHEMSREITRRVKIIPAQMKKVEYVRYVYACRHCE
ref|ZP_02171171.1|ENLYTETVTYTLFVEDQVCSNCELHIMKTQVKDELEIIPAQVKKRYETIYSCRHCE
ref|YP_519650.1|EDLPVEVMEYELPEPERRCPECSGELHTMGRDIWEELKIIPAKAVIVRHQHVYACRCCE
                                *  ::  .:::****:  .:  :*:** **

ref|YP_754943.1|KENITTFPIIKAEMPKPILPGSLASPSILAYIMDQKYTNSLPLYRQEQQFSRLGIELSRQT
ref|YP_754865.1|KENITTFPIIKAQMPNPILPGSLASPSILAYIMDQKYTNSMPLYRQEQQFSRLGIELSRQT
ref|YP_431166.1|REELTTPVVTAPMPAPVLPSPVSPSLLAYVMHQQYGEGLPLYRQEQQFKSLGLELSRQT
RAAC02318AQDVETFPVVRAPTQKPVQAKSLATPEAVAYVMTKKFVDGMPLYRQEQQFARHGYPVLSRQT
ref|ZP_02171171.1|RTGTRNPIVKAPSPERFPFPGSLASPSIVSYMINQKFVQGVPLYRQEQQEFKRMVPIISRQT
ref|YP_519650.1|ATSEHVPMPLKAQRPSFVIKGGFACAETIAHLAVQKFMMSPLYRQEQQEWWKHNGILLSRQT
                                *:  *  *  .  .  .  ::::  *:  .  *****:  .  :****

ref|YP_754943.1|MANWLLAAADPWLKIYDRLHEQLLEKDIILHADETTLQVLKEPGRRAESKSYMWLRYT-G
ref|YP_754865.1|MANWVLNVADPWLKIYDRLHVELLDRIILHADETTLQVLKEPGRSAETKSYMWLRYT-G
ref|YP_431166.1|LANWVLHGANTWLTHIYDRLHEYLLKRDILHADETTLQVLREPGRGAATKSLWLRYT-G
RAAC02318LANWVVAHAETWLEPLIYAKLRQVLLAQRYLHADETTLQVLHEAGRAAQTSYMWVYRS-G
ref|ZP_02171171.1|MSNWIIIEASEQMLEPIWDLMIRILTSLDVLHADETTVQVLKEDGKEAAKSYMWLRYRS-G
ref|YP_519650.1|MSNWLIKASQDWLEPLYEAMKLRLCEHDVLHGDETTLQVLKEPGKTAQSKSYMWLRYRTSG
::*:::  ::  *  ::  :  *  **.**:***:*  *  *  :*:***:  *

ref|YP_754943.1|RDGPPIVLYDYQTTRASKHPDSFLSGFKGYLQTDGYSYGSLT-SVTLACWAHARRKFT
ref|YP_754865.1|RDGPPIVLYEYQTTRASKHPDRFLSGFKGYLQTDGYSAYGKLT-GITLVGCWAHARRKFT
ref|YP_431166.1|RDGPSIVLYDYQTTRASKHPCRFLAGFKGYLHVDGYAGYNELP-DVTLVGCWAHARRKFD
RAAC02318TDEPPVLYDYQETRNAEHPQRFLAGFQGYLHVDGYAGYEGLP-DVTLACWAHARRKFD
ref|ZP_02171171.1|SHDVPIVLYDYQPGRASKYPRRFLEGFTGYLHVDGYGGYHALKPKVELVGCWAHARRKFF
ref|YP_519650.1|EAKHPLVIYEQPDRKHNPQTFLKEFSGYLHTDGYEAYHKLPGNIIIVGCAAHLLRRKFF
.:*:***  *  :*:  **  *  ***:***  .  *  *  :  :.*  **  ****

ref|YP_754943.1|EALKALPAEQKDKPVAASI-----
ref|YP_754865.1|EALKALPAAQKDKPVAA-----
ref|YP_431166.1|EALKALPEDKRNAAVAA-----
RAAC02318EALKAVPPKERKGTAAEAIARWKPTCASLTTSHTPRSPRTFQTSKE
ref|ZP_02171171.1|DAVQTLPPDDRSTTSAK-----
ref|YP_519650.1|DALKTLPKDRQADSNAAGVAY-----
.:*:***  .  **

```

FIG. 77

```

ref|ZP_02171383.1|      -----VYLARGSTDLRKSIDGLAAIVQEGFELDPFSSSLFVFCNRYRDKIKILY
RAAC02319              MLAFDWTSDHRVYLACGATDMRKSIDGLAALVQASFQLDPFSPCLFVFCNRQDKLKIILH
ref|YP_431168.1|      -----RVYLACGATDLRKSIDGLAVLVKEGFELDPFSSCLFVFCNRRDKLKIILH
ref|YP_001212944.1|    -----RVYLAGATDLRKSIDGLAVLVKEGFELDPFSSCLFVFCNRKCDKLKIILH
ref|YP_754944.1|      -----SNRQVYLACGSTDLRKSIDGLAVLVKEAFELNPFSPCLFVFCNRQNKLIILQ
ref|YP_754864.1|      -----SNRLVYLACGSTDLRKSIDGLAVLVKESFHLDPFSPCLFVFCNRKRDKLKIILQ
                        *****:***:*****.:*: .*:***.:***** :*:***

ref|ZP_02171383.1|      WDHNGFWLYYRRLEKGRFPWPTSGSDEPMIITERQLRWLLDGLPLDQKGAHR-----
RAAC02319              WSHNGFWLYYRRLEGRFDWPETGDAKTMVITRRELNWLLDGLPLEQPRAHRAVYVRSAL
ref|YP_431168.1|      WEHNGFWLYYRRLEKGRFVWPQDTTSSTITITRRELRWLLDGLPLKQQAHPKARTIL
ref|YP_001212944.1|    WDHNGFWLHYRRLEKGRFHWPADAGSPTLVISRRELRWLLDGLSLKQPKAHPKARTIL
ref|YP_754944.1|      WDHNGFWLYYRRLEGRGFEPWA-ADSQVVSISYREFRWLLDGLSLKQQAHKAVKERTII
ref|YP_754864.1|      WEHNGFWLHYRRLEGRKFDWPT-AHTDVVSISYREFRWLLDGLSLKQQAHKAVKQR---
*.*****:*****:*** **      : *: *: :.*****.*.* **

```

```

ref|ZP_02734990.1|          -----DIGSEADRYVEKSRAPNTRRAQSRSDWKFSSWCACAKYARSP-----
ref|YP_713924.1|          -----ELAARVGDYARASRSASTWRAYDSDLRQFRAWCAARRPAAPSA
RAAC02333                   MQEIMDTNVSSTIHVGLTELSASANKYVLNSKARNTIRAYQSDWRSFCTWCDERHLSS--
ref|YP_008142.1|          -----DLVQTAKEYATFARSFNTNKSYSRDWDDFVFWCQEKNLRP-----
ref|ZP_01265219.1|        -----ITDIKALQEETLLNLQSSKANNTVRAYKSDFNDFGLFCAKNGFKS--
ref|YP_266430.1|          -----ITDINALKKETLLNLQSSKSINTARAYKSDFTDFSLFCVKNNGFKS--
                               :      .      ::: .* :: **  .*  :*  .
ref|ZP_02734990.1|          LPAAPDPTVAYYLAQRSQE-----LKTSTLQRRLMSISDAHRTAGFDSPTKSAQ
ref|YP_713924.1|          LPATAATVAGYLATLADAG-----YKPSTIRRRALAIsvAHQLAQHPNPAAPAE
RAAC02333                   LPAEPKTVALLYLSMDADRG-----YRTSTIGRHMISIGLAHRTKGFSPSTSDET
ref|YP_008142.1|          LPALPQTIVVYLISRADNAWINQKGLQKPLKISSLSRRLTSISQAHKLAKQPFDKNCPE
ref|ZP_01265219.1|        LPSEPKIVSLYLTHLSTKD-----VRMSTLKRRLVSIGVIHKLKGHYLDTKHPS
ref|YP_266430.1|          LPSEPKIVSLYLTYLSTKE-----VKMSTLKRRLVSIGVIHKLKGHYLDTKHPA
                               **: . : ** : : *: : *: :*: *:
ref|ZP_02734990.1|          VKLVWAGIRRDKGVAQNKKPPTLTKHIREMVEHLPQG-----LLGVRDRALLLLGFAGA
ref|YP_713924.1|          VGAVWDGIRRTGRVPRTRKTAALDLDLTVRVVAGLRDDD-----LADIRDRALLLVGFAGC
RAAC02333                   VRAVWRGIRNTLGVAPQKGSPILVEDLRRMLQHVPND-----LMGLRDRALLLIGFAGA
ref|YP_008142.1|          IQEVWKGIKNKLGSQAQIRKDPILLDDLKRMIESINNDNSKANSLSGMRDKALLLLGFVGA
ref|ZP_01265219.1|        IVENIMGIKRRKGSIQRGKKPLLINNLKKIINVIDQOKKEE--IKKLDRDSIILIGFSGG
ref|YP_266430.1|          IIVENIMGIKRRKGSIQGKKPLLINNLKQIIDVIDQOKKEE--IKKIRDRSIILIGFSGG
                               :      ***. * * * . : : : : : : : :*: :*: :*: *
ref|ZP_02734990.1|          MRRSELVIGIDATDVALTDEGLVVTIRKSKTDQVQEGRTLGIPIYGEHEGTCPVRAVLAWVD
ref|YP_713924.1|          LRRSELVGLDVLADLVETADGLILTVRRSKTDQEGEGALVGIAYGSYRPTCPVRAWRAWAQ
RAAC02333                   FRRSEIVALNVEDVEFVREGLVITLRRSKTDQEGEGRKVGIPIYGSFIETCPVRALQAWLN
ref|YP_008142.1|          FRRSELVSLTIDDIKFVREGLQITLRSKTDQEGEKRIIPIYGSNILTQPVRTLNDWLD
ref|ZP_01265219.1|        FRRNEIVSLDYEDLDVQEGKLINLRRSKTDQFGEFVKGLPYFQNPQCPVISIHKWIE
ref|YP_266430.1|          FRRNEIVSLDFDDLDFVSEGLKINIKRSKTDQFGEFTKALPYFDSSQYCPVSLRNWLE
                               :*: :*: : : * : . :*: :*: :*: :* : . :* : * :
ref|ZP_02734990.1|          QAAILEGPLFRSVNKHGHVMGTRLSDRTVAEVVKRSLVAAGHTARGYAAHSLRAGLITQA
ref|YP_713924.1|          AADLRQGFAPRAVSRHGHVGATRLYPGTSVARVVRQRVAAAGLDPADFAGHSLSRSGFATAA
RAAC02333                   ATGIDSGPLFRRVTKGHDVDRDARLSDKTVARIVKRYVRLIGLDQRHFAGHSLRAGLATSA
ref|YP_008142.1|          CSKISEGLLFRPINRHGQIMDKALTSKSVALIIKRKNHLE-NQKNSFSGHSLRAGFATTA
ref|ZP_01265219.1|        ISKIKSGALFRRFIKSGKSIDNRLTDQTVLIVKEYLKLKLAGIDCKNYSGHSLRSGFATSA
ref|YP_266430.1|          ISRIKSGSVFRRFIKSGSKLSENRLTDQTVALLIKEYLSMLGIDTKNYSGHSLRSGFATSA
                               : : .* ** . : . : * :*: :*: :*: :*: :*: :*: :*
ref|ZP_02734990.1|          AIAGASDRDIQDQSGHKSLLVMRRYIRDGSLF-RQNVAAKVGL
ref|YP_713924.1|          ARAGVADRSIMRQGRWRSSASLDGYVRAGRLFDRDNPSGRVGL
RAAC02333                   ALAGVSESDIMAQTGHRSPMMVRRYIRDSNLF-RSNAAARIGL
ref|YP_008142.1|          AIFGVPEHLIMKQTHKSFDITIRRYI-----
ref|ZP_01265219.1|        AESGAEEERSIMAMTGHKSTEMVRRYIKEANLF-KNNALKNI--
ref|YP_266430.1|          AESGVEERNIMAMTGHKSTEMVRRYIKDANLF-RNNALNKI--
                               * * * : * * * : * : * :

```

FIG. 79

```

ref|YP_001036724.1| TRANSPTSASEMUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPTSASE
ref|YP_001039064.1| -----
ref|YP_001039349.1| -----
RAAC03703 -----
ref|YP_076118.1| -----
ref|YP_074958.1| -----

```

```

ref|YP_001036724.1| MUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPTSASEMUTATRTYPE
ref|YP_001039064.1| -----
ref|YP_001039349.1| -----
RAAC03703 -----
ref|YP_076118.1| -----
ref|YP_074958.1| -----

```

```

ref|YP_001036724.1| CLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPTSASEMUTATRTYPECLSTRIDIUM
ref|YP_001039064.1| -----
ref|YP_001039349.1| -----
RAAC03703 -----
ref|YP_076118.1| -----
ref|YP_074958.1| -----

```

```

ref|YP_001036724.1| THERMCELLUMATCCGBABNTRANSPTSASEMUTATRTYPECLSTRIDIUMTHERMCELLU
ref|YP_001039064.1| -----
ref|YP_001039349.1| -----
RAAC03703 -----
ref|YP_076118.1| -----
ref|YP_074958.1| -----

```

```

ref|YP_001036724.1| MATCCGBABNTRANSPTSASEMUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABN
ref|YP_001039064.1| -----
ref|YP_001039349.1| -----
RAAC03703 -----
ref|YP_076118.1| -----
ref|YP_074958.1| -----

```

```

ref|YP_001036724.1| TRANSPTSASEMUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPTSASE
ref|YP_001039064.1| -----
ref|YP_001039349.1| -----
RAAC03703 -----
ref|YP_076118.1| -----
ref|YP_074958.1| -----

```

```

ref|YP_001036724.1| MUTATRTYPECLSTRIDIUMTHERMCELLUMATCCSCRESIGNIFICANCEE-IDENTIT
ref|YP_001039064.1| -----
ref|YP_001039349.1| -----
RAAC03703 -----
ref|YP_076118.1| -----
ref|YP_074958.1| -----

```

```

ref|YP_001036724.1| IESPSITIVESGAPSMATNNRMALLEQLSKYVVEKDKDFLKEALTLLINALMDAEVTSII
ref|YP_001039064.1| -----MATNNRMALLEQLSKYVVEKDKDFLKEALTLLINALMDAEVTSII
ref|YP_001039349.1| -----MATNNRMALLEQLSKYVVEKDKDFLKEALTLLINALMDAEVTSII
RAAC03703 -----MTNTKSRRHATMASLNSFAVLEWIRKMQDQIDFLRELMQLVTQFLIDAEAAEKI
ref|YP_076118.1| -----SNTSQRGNPQVTNFRIALEELLRK-SGVDDVDFLREGVRVLAQGLMELEVSQQI
ref|YP_074958.1| -----LEELLRK-TGVDDDFLREGLRVLAQGLMELEVSQRI
: * : * : : * : : : * : * :

```

ref YP_001036724.1	GAEKYERNNNRNNYRNGYRLREWDTRVGTQLQSLIPKLRHGSYFPSLLEPRKMSEKALLNV
ref YP_001039064.1	GAEKYERNNNRNNYRNGYRLREWDTRVGTQLQSLIPKLRHGSYFPSLLEPRKMSEKALLNV
ref YP_001039349.1	GAEKYERNNNRNNYRNGYRLREWDTRVGTQLQSLIPKLRHGSYFPSLLEPRKMSEKALLNV
RAAC03703	GAERYERTESRVTQRNGYRSRAWDTRLGTVDLKI PKLRQGSFFPSILEPRRRAEQALASV
ref YP_076118.1	GAERYERSERSNYRNGYRERQWDTRVGTIDLQIPKLRKGSYMPSPLEPRRRAEKALVAV
ref YP_074958.1	GADRYERSAERSTYRNGYRERQWDTRVGTIDLQIPKLRKGSYMPSPLEPRRRAEKALVAV
	:. * . ***** * *****:***:*. *****:***:*** *****: :*:** *
ref YP_001036724.1	VQEAYVHGVSTRKVDELVEALGMKGIDKSEVSRIKQLDEFVVEEFKNRRLEGEYPYLWLD
ref YP_001039064.1	VQEAYVHGVSTRKVDELVEALGMKGIDKSEVSRIKQLDEFVVEEFKNRRLEGEYPYLWLD
ref YP_001039349.1	VQEAYVHGVSTRKVDELVEALGMKGIDKSEVSRIKQLDEFVVEEFKNRRLEGEYPYLWLD
RAAC03703	IQEAYVKGVSTRKVDDLVRALGLDGIKSEVSRLCQLIDEEVRQFKERPLEREYPYVWLD
ref YP_076118.1	VQEAYIQGVSTRKVDELVQALGMTGVSKSQVSRLCAELDEVVEAFNRPLEGRYPYVWLD
ref YP_074958.1	VLEAYVNGVSTRKVDDLQVQALGMTGVSKSQVSRLCAELDEVVQAFNRPLESRYPYVWLD
	: ***:*****:***.***: *.***:***:.. ** *. *: * ** .***:***
ref YP_001036724.1	ATFPKVREGGRVCSMALVIAVGVNQGEREILGFDVGMSSEDGAFWEEFLRRLVARGLKGV
ref YP_001039064.1	ATFPKVREGGRVCSMALVIAVGVNQGEREILGFDVGMSSEDGAFWEEFLRRLVARGLKGV
ref YP_001039349.1	ATFPKVREGGRVCSMALVIAVGVNQGEREILGFDVGMSSEDGAFWEEFLRRLVARGLKGV
RAAC03703	ATFPKVREGGRVQSMALVIAIGVTDGTGEREVLGFDVGTSEDGAFWSDFLRSLKARGLRGV
ref YP_076118.1	AKYEKVRENGRVSSMALVIAMGVREDGEREILGLDVGPSSEDGAFWTAFLRQLVARGLKGV
ref YP_074958.1	AKYVKVRENGRVCTMALVAVGVREDGDREVLGLDVGPSSEDGAFWTAFLRQLVTRGLKGV
	*.: ****.*** :****:***: *.***:***:*** ***** *** * :***:***
ref YP_001036724.1	RLVISDAHEGLKAAIKKILTGSAWQRCRVHFMNRNLSQVPKHYQGMVSSIIRTIFAQNDQ
ref YP_001039064.1	RLVISDAHEGLKAAIKKILTGSAWQRCRVHFMNRNLSQVPKHYQGMVSSIIRTIFAQNDQ
ref YP_001039349.1	RLVISDAHEGLKAAIKKILTGSAWQRCRVHFMNRNLSQVPKHYQGMVSSIIRTIFAQNDQ
RAAC03703	RLVVSDAHAGLRQAISEVLGTATWQRCVKVHTIRNLSQVPKREQSMVASIIRTIFTQPTQ
ref YP_076118.1	LLVISDNHVGLEAIRTVFSGASWQRCRVHFMNRNLLGYVPKNLQSMVASAVRTIFAQPDQ
ref YP_074958.1	LLAISDSHVGLQEAIRTVLSGASWQSRVHFMNRNLLGYVPKHWQSMVAAVRTIFAQPDQ
	*.:** * **: * * :*:***:*** :***.***. ****. *.***: :****:* *
ref YP_001036724.1	ESAREQLRHVVDELKNRFPKAMKILEEAEIEILAYMAFFREHWAQIHSTNPLERLNREIR
ref YP_001039064.1	ESAREQLRHVVDELKNRFPKAMKILEEAEIEILAYMAFFREHWAQIHSTNPLERLNREIR
ref YP_001039349.1	ESAREQLRHVVDELKRRFPKAMKILEEAEIEILAYMAFFREHWAQIHSTNPLERLNREIR
RAAC03703	EAAREQLRRVVAELRGFRPKAMDILEAEEDVLAFMALPIEHWRQICSTNPLERLNREMR
ref YP_076118.1	QAAKSQLAVVVENLRKQFPRAAQLLEDAEEDILAYMAFFTEHWRRLHSTNPLERLNREIG
ref YP_074958.1	QAARRQLAVVADNLRQFPRAAQLLEAEEDILAYMAFFTEHWRQLHSTNPLERLNREIG
	:*: ** *. *: :***:*. **:***:***:*** * ** : :*****:
ref YP_001036724.1	RRTDVVCIFFPNREAVIRLVGAMLMEQNDEWKVG-RRYFSLESMSKITSINEFTLTPVALL
ref YP_001039064.1	RRTDVVCIFFPNRKAVIRLVGAMLMEQNDEWKVG-RRYFSLESMSKITSINEFTLTPVALL
ref YP_001039349.1	RRTDVVCIFFPNREAVIRLVGAMLMEQNDEWKVG-RRYFSLESMSKITSINEFTLTPVALL
RAAC03703	RRMDVVGIFPNRASVVRLAGAILQEQQHEEWLVS-RRYFSLESMAKLPNRPF-LAAEAML
ref YP_076118.1	RRTEIVGIFPNREALIRLAGAVMIEQQEWMTPAPRRYFSQASMAKLYAHDPSLGRPELL-
ref YP_074958.1	RRTDVVGIFPNREAFIRLAGAVLIEQQDEWTAAPRRYFSQASMAKL-----
	** :*: ***** :*:***:***: ***:*** .. ***** ***:*
ref YP_001036724.1	HK
ref YP_001039064.1	HK
ref YP_001039349.1	HK
RAAC03703	QK
ref YP_076118.1	--
ref YP_074958.1	--

FIG. 80

```

ref|YP_001039349.1|
ref|YP_001036724.1|
ref|YP_001039064.1|
RAAC03568
ref|YP_076118.1|
ref|YP_074958.1|

```

TRANSPSASEMUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPSASE

```

ref|YP_001039349.1|
ref|YP_001036724.1|
ref|YP_001039064.1|
RAAC03568
ref|YP_076118.1|
ref|YP_074958.1|

```

MUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPSASEMUTATRTYPE

```

ref|YP_001039349.1|
ref|YP_001036724.1|
ref|YP_001039064.1|
RAAC03568
ref|YP_076118.1|
ref|YP_074958.1|

```

CLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPSASEMUTATRTYPECLSTRIDIUM

```

ref|YP_001039349.1|
ref|YP_001036724.1|
ref|YP_001039064.1|
RAAC03568
ref|YP_076118.1|
ref|YP_074958.1|

```

THERMCELLUMATCCGBABNTRANSPSASEMUTATRTYPECLSTRIDIUMTHERMCELLU

```

ref|YP_001039349.1|
ref|YP_001036724.1|
ref|YP_001039064.1|
RAAC03568
ref|YP_076118.1|
ref|YP_074958.1|

```

MATCCGBABNTRANSPSASEMUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABN

```

ref|YP_001039349.1|
ref|YP_001036724.1|
ref|YP_001039064.1|
RAAC03568
ref|YP_076118.1|
ref|YP_074958.1|

```

TRANSPSASEMUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPSASE

```

ref|YP_001039349.1|
ref|YP_001036724.1|
ref|YP_001039064.1|
RAAC03568
ref|YP_076118.1|
ref|YP_074958.1|

```

MUTATRTYPECLSTRIDIUMTHERMCELLUMATCCSCRESIGNIFICANCEE-IDENTIT

```

ref|YP_001039349.1|
ref|YP_001036724.1|
ref|YP_001039064.1|
RAAC03568
ref|YP_076118.1|
ref|YP_074958.1|

```

-----MATNNRMALLEQLSKYVVEKDKDFLKEALTLLINALMDAEVTSII
IESPSITIVESGAPSMATNNRMALLEQLSKYVVEKDKDFLKEALTLLINALMDAEVTSII
-----MATNNRMALLEQLSKYVVEKDKDFLKEALTLLINALMDAEVTSII
----MTNTKSRRHATMASLNSFAVLEWIRKMQDQIDFLRELMQLVTQFLIDAEAAEKI
-----SNTSQRGNPQVTNFRIALEELLRK-SGVDDVDFLREGVRVLAQGLMELEVSQOI
-----LEELLRK-TGVDDDFLREGLRVLAQGLMELEVSQRI
: * : * : * : * : * : * : * : * : *

GAEKYERNNNRNNYRNGYRLREWDTRVGTLQLSI PKLRHGSYFPSLLEPRKMSEKALLNV
GAEKYERNNNRNNYRNGYRLREWDTRVGTLQLSI PKLRHGSYFPSLLEPRKMSEKALLNV
GAEKYERNNNRNNYRNGYRLREWDTRVGTLQLSI PKLRHGSYFPSLLEPRKMSEKALLNV
GAERYERTESRVTQRNGYRSRAWDTRLGTVDLKIPKLRQGSFFPSI LEPRRAEQALASV
GAERYERSERSNYRNGYRERQWDTRVGTIDLQIPKLRKGSYMPSWLEPRRAEKALVAV
GADRYERSAERSSTYRNGYRERQWDTRVGTIDLQIPKLRKGSYMPSWLEPRRAEKALVAV
* : * * * * * : * * * * * * : * * * * * : * * * * * : * * * * * : * * * * *

```
ref|YP_001039349.1|
ref|YP_001036724.1|
ref|YP_001039064.1|
RAAC03568
ref|YP_076118.1|
ref|YP_074958.1|
```

```
VQEAYVHGVSTRKVDELVEALGMKGIDKSEVSRI SKQLDEFVEEFKNRRLEGEYPYLWLD
VQEAYVHGVSTRKVDELVEALGMKGIDKSEVSRI SKQLDEFVEEFKNRRLEGEYPYLWLD
VQEAYVHGVSTRKVDELVEALGMKGIDKSEVSRI SKQLDEFVEEFKNRRLEGEYPYLWLD
IQEAYVKGVSTRKVDDLVRALGLDGISKSEVSRLCQLIDEEVRQFKERPLEREYPYVWLD
VQEAYIQGVSTRKVDELVLQALGMTGVSKSQVSRLCAELDEVVEAFNRNPLEGRYPYVWLD
VLEAYVNGVSTRKVDLVLQALGMTGVSKSQVSRLCAELDEVVQAFNRNPLESRYPYVWLD
: ***:*****:***: * :***:***: : ** * : * : * : * : * : *
```

```
ref|YP_001039349.1|
ref|YP_001036724.1|
ref|YP_001039064.1|
RAAC03568
ref|YP_076118.1|
ref|YP_074958.1|
```

ATFPKVVREGGRVCSMALVIAVGVNQQGEREILGFDVGMSEDGAFWEEFLRRLVARGLKG
ATFPKVVREGGRVCSMALVIAVGVNQQGEREILGFDVGMSEDGAFWEEFLRRLVARGLKG
ATFPKVVREGGRVCSMALVIAVGVNQQGEREILGFDVGMSEDGAFWEEFLRRLVARGLKG
ATFPKVVREGGRVQSMALVIAIGVTDTGEREVLGFDVGTSEDGAFWSDFLRSLKARGLRGV
AKYEKVRENGRVSSMALVIAMGVREDGEREILGLDVGPSEDGAFWTAFLRQLVARGLKG
AKYVKVRENVRCTMALVVAVGVREDGDREVLGLDVGPSEDGAFWTAFLRQLVTIRGLKG
* - * - * - * - * - * - * - * - * - * - * - * - * - * - * - * - * - *

```
ref|YP_001039349.1|
ref|YP_001036724.1|
ref|YP_001039064.1|
RAAC03568
ref|YP_076118.1|
ref|YP_074958.1|
```

```
RLVLSDAHEGLKAAIKKILTGSAWQRCRVHFMRNVLSQVPKHYYQGMVSSIIRTI FAQN DQ
RLVLSDAHEGLKAAIKKILTGSAWQRCRVHFMRNVLSQVPKHYYQGMVSSIIRTI FAQN DQ
RLVLSDAHEGLKAAIKKILTGSAWQRCRVHFMRNVLSQVPKHYYQGMVSSIIRTI FAQN DQ
RLVVSDAHAGLRQAISEVL TGATWQRCKVHTIRNVLSQVPKREQSMVASIIRTI FTHPTQ
LLVISDNHVGLREAIRTVFSGASWQRCRVHFMRNLLGYVPKNLQSMVSAAVRTI FAQP DQ
LLAISDSHVGLEAIRT VLSGASWQRSRVHFMRNLLGYVPKKHWQSMVA AAVRTI FAQP DQ
* . * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
```

```
ref|YP_001039349.1|
ref|YP_001036724.1|
ref|YP_001039064.1|
RAAC03568
ref|YP_076118.1|
ref|YP_074958.1|
```

ESAREQL-
ESAREQL-
ESAREQL-
EAGREQLP
QAAKSQL-
QAARRQL-
.: : **

FIG. 81

```

ref|YP_516922.1| MFARIKTAYNRDGSPPRYLQIVESRREEGKVRQKVLNLRVEDLQNGKLDDLIRSLAKF
ref|YP_519084.1| MFARIKTAYNRDGSPPRYLQIVESRREEGKVRQKVLNLRVEDLQNGKLDDLIRSLAKF
ref|ZP_01370818.1| MFARIKTAYNRDGSPPRYLQIVESRREEGKVRQKVLNLRVEDLQNGKLDDLIRSLAKF
ref|ZP_01372264.1| MFARIKTAYNRDGSPPRYLQIVESRREEGKVRQKVLNLRVEDLQNGKLDDLIRSLAKF
ref|YP_519534.1| MFARIKTAYNRDGSPPRYLQIVESRREEGKVRQKVLNLRVEDLQNGKLDDLIRSLAKF
RAAC03707 MFAQIVSTKKPDGKTYKYLHIVESYREGRTVKRRVASLGNISQYSEREIEQIIRTLES
***:* : : **.. :***:*** ** .*::: :..**.: : : :***:* :

ref|YP_516922.1| SDKLAVVDAEDLFADWSKEYGPSMVFNRLWDNLGLHKILDGLFSERELSDIVREAIFCM
ref|YP_519084.1| SDKLAVVDAEDLFADWSKEYGPSMVFNRLWDNLGLHKILDGLFSERELSDIVREAIFCM
ref|ZP_01370818.1| SDTLAVVDAEDLFADWSKEYGPSMVFNRLWENLGLHTILDGLFNERDLSIDVQEAIFCM
ref|ZP_01372264.1| SDTLAVVDAEDLFADWSKEYGPSMVFNRLWENLGLHKIFAGLFNERDLSIDVQEAIFCM
ref|YP_519534.1| SDTLAVVDAEDLFADWSKEYGPSMVFNRLWENLGLHKILSGLFNERELSDIVQEAIFCM
RAAC03707 LQHR-TTGSLEDFAEQVVLHFGVPYVQFLWNQLGLTEAIRDALRAREVTFDVARYVQAM
: :..: **: *: :.* . *.. **:*** : : : *::*** : :.*

ref|YP_516922.1| VLNRLTEPTSKLGVSDWKDSVYRPDF--ETLKLHHFYKAIDFLDENKDTIEEQLFHHHTN
ref|YP_519084.1| VLNRLTEPTSKLGVSDWKDSVYRPDF--ETLKLHHFYKAIDFLDENKDTIEEQLFHHHTN
ref|ZP_01370818.1| VLNRLTEPTSKLGVSDWKDSVYRPEF--ESLKLHHFYKAIDFLDENKDTIEEQLFHHHTN
ref|ZP_01372264.1| VLNRLTEPTSKLGVSDWKDSVYRPEF--ESLKLHHFYKAIDFLDENKDTIEEQLFHHHTN
ref|YP_519534.1| VLNRLTEPTSKLGVSDWKDSVYRPEF--ESLKLHHFYKAIDFLDENKESIEEQLFHHQTN
RAAC03707 VIHRLVDPSSKRLRFHTLDDLYLPDWGGEFWQLQHFYRALDYLDIKPQLERVLARLTD
*:**.:***: : . *.:* *: : *.**:***:***: : * :*. *: : *:

ref|YP_516922.1| LFTQQQLDLVFFDTTSTYVEGDAGAFDLLEYGHSKDHRPDRLOVMIGILMSRDGIPIAHHV
ref|YP_519084.1| LFTQQQLDLVFFDTTSTYVEGDAGAFDLLEYGHSKDHRPDRLOVMIGILMSRDGIPIAHHV
ref|ZP_01370818.1| LFTQQQLDLVFFDTTSTYVEGDAGAFDLLEYGHSKDHRPDRLOVMIGILMSRDGIPIAHHV
ref|ZP_01372264.1| LFTQQQLDLVFFDTTSTYVQGDAGAFDLLEYGHSKDHRPDRLOVMIGILMSRDGIPIAHHV
ref|YP_519534.1| LFTQQQLDLVFFDTTSTYVEGDAGAFDLLEYGHSKDHRPDRLOVMIGILMSRDGIPIAHHV
RAAC03707 LLNFRLSIVLYDLTSTHLGHA--CPLGEHGYSRTHRPDLQVELGLLVTPGIPITHEV
*.: *.**.:* **.:*. * *.:*: :*** ** *:*: : :***:*.

ref|YP_516922.1| FPGNTPDADFIEAVSDLKKRFNIQRVIVVGDRGMMGKRTLELLEELQLQYILGVRMRN-
ref|YP_519084.1| FPGNTPDADFIEAVSDLKKRFNIQRVIVVGDRGMMGKRTLELLEELQLQYILGVRMR-
ref|ZP_01370818.1| FPGNTSDADFIEAVSDLKKRFTIQRVIVVGDRGMMGKRTLELLEELQLHYILGVRMRN-
ref|ZP_01372264.1| FPGNTPDADFIEAVSDLKKRFTIQRVIVVGDRGMMGKRTLELLEELQLHYILGVRMRN-
ref|YP_519534.1| FPGNTPDADFIEAVRDLKKRFTIQRVIVVGDRGMMGKRTLELLEELQLHYILGVRMRN-
RAAC03707 FAGNVSDKQTVPDILKRLKEDFAVEQCVMVFGDRGMVTEKNMALMAEAGFPYIVGFHKRGR
*.*.*.*.: : : **: * : : :.*****: :*: * : :*:*: *

ref|YP_516922.1| --IKAGPDLANSPEPYAFIKDNLKVKEVVHQ-----EKRYIVCLNEEEAKRDQLVR
ref|YP_519084.1| --IKAGPDLANSPEPYAFIKDNLKVKEVVHQ-----EKRYIVCLNEEEAKRDQLVR
ref|ZP_01370818.1| --VKAGPELATSPEPYVFTKDNLKVKEVLHQ-----GKRYIVCLNEEEAKRDQWVR
ref|ZP_01372264.1| --IKAGPELAASAQYPFVKDNLKVKEVLHQ-----EKRYIVCLNEEEAKRDQLVR
ref|YP_519534.1| --IKAGPELAASPEPYTFLKDNLKVKEVLHQ-----EKRYIVCLNEEEAKRDQLVR
RAAC03707 IVSDALLEQFADVNAHYELKDNRLRYLEVPAAVDDVEKAEGVRYILCYNPEKARQDAAFR

ref|YP_516922.1| E-----
ref|YP_519084.1| E-----
ref|ZP_01370818.1| E-----
ref|ZP_01372264.1| E-----
ref|YP_519534.1| E-----
RAAC03707 ESALEEAETGLKALAESLAKPKRGRKPTDKGVMLKVADLLTRKGVEAFFQVDYKDGILTY

ref|YP_516922.1| -----
ref|YP_519084.1| -----
ref|ZP_01370818.1| -----
ref|ZP_01372264.1| -----
ref|YP_519534.1| -----
RAAC03707 RRDEDAIT

```

FIG. 82

```

ref|ZP_02171383.1|      -----VYLARGSTDLRKSIDGLAAIVQEGFELDPFSSSLFVFCNRYRDKIKILY
RAAC03173              MLAFDWTSDHRVYLVCGATDMRKSIDGLAALVQASFQLDPFSPCLFVFCNQRDKLKIHL
ref|YP_431168.1|      -----RVYLACGATDLRKSIDGLAVLVKEGFELDPFSSCLFVFCNRRNRDKLKIHL
ref|YP_001212944.1|    -----RVYLAIGATDLRKSIDGLAVLVKEGFELDPFSSCLFVFCNRKCDKLKIHL
ref|YP_754944.1|      -----SNRQVYLACGSTDLRKSIDGLAVLVKEAFELNPFSPCLFVFCNRQRNKLKIILQ
ref|YP_754864.1|      -----SNRLVYLACGSTDLRKSIDGLAVLVKESFHLDPFSPCLFVFCNRKRDKLKIILQ
                        ***. *:***:*****.:*: .*:***.***** *:***

ref|ZP_02171383.1|      WDHNGFWLYYRRLEKGRFPWPTSGSDEPMIITERQLRWLLDGLPLDQKGAHR-----
RAAC03173              WSHNGFWLYYRRLEGRFDWPETGDAKTMVITRRELNWLLDGLPLEQPKAHRAVPVRSAL
ref|YP_431168.1|      WEHNGFWLYYRRLEKGRFVWPQDTSSTITITRRELRWLLDGLPLKQPKAHPEVKARTIL
ref|YP_001212944.1|    WDHNGFWLHYRRLEKGRFHWPADAGSPTLVISRRELRWLLDGLSLKQPKAHPEVKARTIL
ref|YP_754944.1|      WDHNGFWLYYRRLEGRGKFEWPA-ADSQVVSISYREFRWLLDGLSLKQKQAHKAVKERTII
ref|YP_754864.1|      WEHNGFWLHYRRLEGRGKFDWPT-AHTDVVSISYREFRWLLDGLSLKQKQAHKAVKQR---
                        *.*****:*****:*:* **      : *: *:***:*****.*.* **

```

FIG. 83

```

ref|NP_634718.1|-----MLKTEEWLLIRDLYSQGFSISEIA
ref|NP_616807.1|-----MLKTEEWLSIRDLYSQGFSISEIS
gb|AAR99616.1|-----IKEMYERGMSISDIA
sp|Q45618|TRA6_BACST-----IKEMYERGMSISDIA
ref|YP_074959.1|-----MLRSGETLEIRQMYAGGLSISEIA
RAAC02966MRI PQDHRPQFLKIIGMNSPPPCRIVAPVIRSWEVPMREDERMEIRQLYEAGVSISELA
                                     *:::*  *.***:::

```

```

ref|NP_634718.1|KQTGFDRKTVRKYLRLKTLPEPQKRSG-----RKSCLDPFKPYIQEKLKEGPYT-----
ref|NP_616807.1|RRTGYARETVRKYLKKKTAPEPQKRPP-----KPSCLDPFKPYIQEKLKEGPYT-----
gb|AAR99616.1|RELGIDRKTVRKYIHSPNPPSKSKRKQ-----RKSCLDPFKPYLQKRM-----
sp|Q45618|TRA6_BACSTRELGIDRKTVRKYIHSPNPPSKSKRKQ-----RKSCLDPFKPYLQKRM-----
ref|YP_074959.1|RRTGRDRKTIRKWLRTNTMPKPAKRK-----RSSMLDQHEAFTLEQMOKGVTSASK---
RAAC02966RRFGYDRKTIRSALNSSLEEKQGERASRGERKKGSKLEPYKDYVKQRMQLGVSTLNEFCE
:. *  *::*. :. . : * : * *: : : : :

```

```

ref|NP_634718.1|-----
ref|NP_616807.1|-----
gb|AAR99616.1|-----
sp|Q45618|TRA6_BACST-----
ref|YP_074959.1|-----
RAAC02966RFGSRAIPAVSPSSVSS

```

FIG. 84

```

ref|YP_148444.1|      MEERLVSGEVLGEETALEPSLRPQYLHEYIGQDKIKENLKVFI EAAKLREETLDHVLLYG
ref|NP_980798.1|      MDERLLSGESAYEDADLEYSLRPQTLRQYIGQDKAKHNLEVFIEAAKMREETLDHVLLYG
sp|Q9KDI8|RUVB_BACHD MEERMVSAEAQTEEA AVEQGIRPHSFEQYIGQEKVKQNLKVFI EAAKMREEALDHVLLYG
ref|YP_430545.1|      --ERLVAGNLHNEDQELELSLRPRCLAEYIGQEHVKETLGIFIQ AARERGEALDHVLLYG
ref|YP_001211577.1|   MKDRLISAVARPEDADVDTSLRPRLLAEYIGQEKVKETISVFIQ AARGRGEPLDHVLLFG
RAAC00757             MDERLISA EWMREDAQLDT-IRPRFLDDYIGQRAAVENLRIFIQ AAKERGEPLDHVLLYG
                        *: : : :      *: : :      : : : :      : : : :      : : : :      : : : :      : : : :      : : : :
ref|YP_148444.1|      PPGLGKTTLAVIIANEMGVKL RATS G PALERPGDLAALLTSLEPGDVL FIDEI HRLPRAV
ref|NP_980798.1|      PPGLGKTTLANIIANEMGVNVRTTS G PAIERPGDLAAVLTS LQPGDVL FIDEI HRLHRSI
sp|Q9KDI8|RUVB_BACHD PPGLGKTTTLSTIIANELGVQMR TTS G PAIERPGDLAAILTALEPGDVL FIDEI HRLNRMV
ref|YP_430545.1|      PPGLGKTTLAGIIANELGVQLRV TTS G PALERAGDLAAILTNLQPRDVL FIDEI HRLPRQV
ref|YP_001211577.1|   PPGLGKTTLANIIANEMGV SIRTTS G PAVERPGLD AAILTSLSQGDIL FIDEI HRLSRTV
RAAC00757             PPGLGKTS LAMIIANELGVQIRVT S G PAIERAGDLAAILTNLQPGDVL FIDEI HRLSPSV
                        *****: : *****: : *****: : *****: : *****: : *****: : *****: :
ref|YP_148444.1|      EEVLYPAMEDYCLDITIGKGP DARTLRDLPPFTLVGAT TRAGALSAPLRDRFGVISRLE
ref|NP_980798.1|      EEVLYPAMEDFCLDIVIGKGPSARSVR L DLPFTLVGAT TRAGALSAPLRDRFGVLSRLE
sp|Q9KDI8|RUVB_BACHD EEVLYPAMEDYCIDIVIGKGPTARSVR L DLPFTLVGAT TRAGMLSSPLRDRFGVMARLE
ref|YP_430545.1|      EEIYPAMEDFVLDI ILGKGPGARSIR L DLPFTLVGAT TRAGLLSSPLRDRFGVINSRLE
ref|YP_001211577.1|   EEVLYPAMEDYALDIVIGKGP GARSIRLELP RFTLVGAT TRAGLLTSPLRDRFGVISRLE
RAAC00757             EEVLYPAMEDFAIDIVIGKGPSARSVR L DLPFTLVGAT TRAGLLSHPLRDRFGVMLHLD
                        **: *****: : ** : **** *: : : : : ** *: *****: : *: *****: : *:
ref|YP_148444.1|      YYHVDQLAQIIERAAA I LQIGIEREAALELARRARGTPRIANRLLRRVRDFAQVRGEGGI
ref|NP_980798.1|      YYTVDQLSEIVERTA E VFEVEIDSLAALEIARRARGTPRIANRLLRRVRDFAQVRGNGTV
sp|Q9KDI8|RUVB_BACHD YYNVEELTTIIERTATIFDTELERDASIEIARRS RGT PRIANRLLRRVRDFAQVSGDMRI
ref|YP_430545.1|      FYQVAEELEEII RRAATILQVAIEPEGAREIARRARGTPRVANRLLKRVRDYAEIRAGGVI
ref|YP_001211577.1|   YYRPEDLV LIVNRSARILGIEITAE GAFEIARRS RGT PRIANRLLKRVRDYAQVRANGVI
RAAC00757             YYPVRDLAEIVKRNARILQLSITEDGCAE IARRARGTPRIANRLLKRVRDIAQVAGWPEI
                        : * : * *: * * : : : : : : : : : : : : : : : : : : : : : : : :
ref|YP_148444.1|      TLPLAVEALERLQVDR LGLDQIDHKLLSAMI EK FAGGFPVGLET LA AVIGEEAQTIEEVYE
ref|NP_980798.1|      TMEITQMALELLQVDK LGLDHDHDKLLLGII EKFRGGFPVGLETVSATIG EESHTIEDVYE
sp|Q9KDI8|RUVB_BACHD SSSRAIESLERLQVDR LGLDHDHDKLIK GIMTKFN GGPVGLETISATIG EETDTIEEVYE
ref|YP_430545.1|      TREVAREALELLQVDAAGLDSSDR LLLTLIRKFNGGFPVGLET LAAAI SEEPDTIEDVYE
ref|YP_001211577.1|   TCEVAVEALKFLEVDPLGLDFADR LLLTIIQKFGGGFPVGLET IATAVNEEPETVEDVYE
RAAC00757             DAARAAEALAQ LHV DPLGLDATDKRILEAAMD KFGGGFPVG LDTLAAAVGE EPSTLEDVYE
                        : : * * * * * : : : : : : : : : : : : : : : : : : : : : :
ref|YP_148444.1|      PYLMQIGLLQRTPRGRVVT PAAYTHLGMEVP--
ref|NP_980798.1|      PYLLQIGFLQRTPRGRIVT PLAYEHFGMEVP--
sp|Q9KDI8|RUVB_BACHD PYLLQIGFLQRTPRGRVVT PLAYEHFNMEVPN--
ref|YP_430545.1|      PFLQMGYLQRTPRGRVATPGAYAHLG-----
ref|YP_001211577.1|   PYLIQLGMLARTPRGRVTTPLAFRHLG-----
RAAC00757             PYLLQIGFLKRTPRGRVVMPSAYRHLGRAVPSG
                        *: *: *: * * *****: . * *: * : .

```

FIG. 85

```

ref|YP_234498.1| -----MIGRLRGSIAEKQPPHLVLDVNGVGYEVEVPMTTYRLPHVGETVTLHTHLVV
ref|YP_275913.1| -----MIGRLRGSIAEKQPPHLVLDVNGVGYELEVPMTTYRLPHVGETVTLHTHLVV
ref|NP_793742.1| -----MIGRLRGAIAEKQPPHLVLDVNGVGYELEVPMTTYRLPHVGETVTLHTHLVV
ref|YP_001186768.1| -----VIGRLRGTLAEKQPPHLVLDVNGVGYELEVPMTTYRLPAVGETLTLHTHQVV
ref|ZP_00991066.1| -----MIGRLRGTLIEKQPPELLIEVSGVGYEVQMPMSCFYELPNVGEEAIIYTHFVV
RAAC00756 MGTGAFSVIAFLRGRVAFLLGPGYVDLDVRDVGYRVHVCDRDQAAL-NLGDTAFLYTHHHV
          :*. *** : * : :*.***.:. * :*: :.* *

ref|YP_234498.1| REDAHLLYGFYEKRERELFRELIRLNGVGPKLALALMSGLEVDELVRCVQAQDTSALTRI
ref|YP_275913.1| REDAHLLYGFYEKRERELFRELIRLNGVGPKLALALMSGLEVDELVRCVQAQDTSALTRI
ref|NP_793742.1| REDAHLLYGFYEKRERELFRELIRLNGVGPKLALALMSGLEVDELVRCVQAQDTSALTRI
ref|YP_001186768.1| REDAHLLYGFFYEKRERELFRELIRLNGVGPKLALALMSGLEVDELVRCVQAQDTAALTKV
ref|ZP_00991066.1| REDAQLLYGFNTVKERALFREVIVKANGVGPKLGLGILSGMTASQFVQSVEREDISTLVKL
RAAC00756 REDGWALYGFETIEERALFERLVAVSGIGPKLALQVIGAAGVEIVAAILAEDAESLSRL
          ***. **** .** **.:. :*.***.* :.. .:.* :. :* :* :.

ref|YP_234498.1| PGVGKKTAERLLVELKDRFKAWELPGTFTLVSNNGPNQAEP-VASAESDAVSALISLGKY
ref|YP_275913.1| PGVGKKTAERLLVELKDRFKAWESLPGTFTLVSNNGPNQAEP-VASAESDAVSALISLGKY
ref|NP_793742.1| PGVGKKTAERLLVELKDRFKAWESLPGTFTLVSNNGPNHADP-VASAESDAVSALISLGKY
ref|YP_001186768.1| PGVGKKTAERLLVELKDRFKAWETMPAIATLVVE-PRLGAT-VSSAENDAVSALISLGKY
ref|ZP_00991066.1| PGVGKKTAERLVEMKDRKLGWAGDLFTPATDAAPMDSMPTVHDAEEEAVSALLALGKY
RAAC00756 PGIGRKLASRLVVELREKLDDLAPALRAPAPASSPQ-----GSAEEDAVSALVALGYR
          **.*: * **.*:***:~::~. . * .*:*****:***:

ref|YP_234498.1| PQEASKAVSAIKEKDLSSADLIRRAL-----
ref|YP_275913.1| PQEASKAVSAIKEKDLSSADLIRRAL-----
ref|NP_793742.1| PQEASKAVSAIKEKDLSSADLIRRAL-----
ref|YP_001186768.1| PQEASRAVAAVKEDGLSSEDLIRRAL-----
ref|ZP_00991066.1| PTQASKVVVAQVAKGMTSEQLIREAL-----
RAAC00756 PREAEEAVSAVGRGRQSVEDTIKAALTYLYARDARTEPLSQP
          * :*...*: : . : :* : **

```

FIG. 86

```

ref|YP_644098.1|      -----IILGIDPGTATMGWGVVRQ-EGSRLRYVQHGAITTP
RAAC00755             -----MESLRILGVDPGLARLGFGLIERGPGDSLHHVAHGCIETG
ref|YP_074988.1|      -----MRILGIDPGTARMGYGIVEDAGAGRERAVEYGCLETP
ref|ZP_02848139.1|    -----MRVLGIDPGIAIAGFGFIDK-DGHKLKPVQYGCIQTA
ref|YP_518701.1|      EE-IDENTITIESPSITIVESGAPSILGIDPGTAIMGYGLIEK-KGNRLFPVDYACWRTP
ref|YP_001180347.1|    -----MRVLGIDPGIALTGYGIIESKNGSEFKVIDYGRIETS
                        :*:*** * *:*. : . : : *

ref|YP_644098.1|      SGWGMFRRDLRLFAGVTELRGYRPSAVAVEELFFNTNVTTAITVGQARGVALLAAYRAG
RAAC00755             ADTPLPERLRHIFQQLTELCTHRPAVMAVEELFFSRNTTTAFTVGQARGVALLAGAEAG
ref|YP_074988.1|      PDMRPELRRLQALYRGLADLMARHRPDALAVEELFFGRNVTTAIHVGQARGVLLAAADSG
ref|ZP_02848139.1|    AHTPQEERLVQIYDSAVALMDKYKPDTVAVEKLFFNRNVTTAFVVGQARGVILAAAQRG
ref|YP_518701.1|      AHTPMPERLLMLYHEIEAYIKEKQPHHVAVEELFFNRNTTTAISVGQARGVLLAAAQCG
ref|YP_001180347.1|    SSLKKSMLLLHLYTELCSIIISLYQPDVVAIEELFFSKNSKTAITIGEARGVILTCIQNN
                        . ** : : * :*:***. * .*: :*:*** :*: .

ref|YP_644098.1|      VEVFEYTPLQVKQAITSYGRADKRQVQEMVRALLGLRSIPRP-----
RAAC00755             LAVMEYTPMQVKQAVTGYGRADKRQVQEMVRLLLRLTSVPKPDDAADALAVAIHAHAGR
ref|YP_074988.1|      VPVREFTPMQVKMAVTGYGRANKAQVQRMVQALLGLPEIPKP-----
ref|ZP_02848139.1|    LPVAEYTPLQVKQAVVGYGKAEKRQVQEMVKMFLKLSAIPKP-----
ref|YP_518701.1|      LPVYETPLQVKQAVAGYGRADKQIQQMVRALLGLQEIPKP-----
ref|YP_001180347.1|    LSIYETPLQVKQSITGYGRADKTQIQKMVKSLLGLSEIPKP-----
                        : : **:*** : : .*:***: * *:*.*** : * * :*:

ref|YP_644098.1|      -----
RAAC00755             IGELEARLAAGKRTGWVWERGRFR
ref|YP_074988.1|      -----
ref|ZP_02848139.1|    -----
ref|YP_518701.1|      -----
ref|YP_001180347.1|    -----

```

FIG. 87

```

ref|YP_146341.1|-----VFIYLRKSRKDIEEEKKAAESGASYDTLQHRDNLAVARKEGHNIL
ref|YP_001373830.1|-----VFIYLRKSRKDLEEEKKAMEHGQHYDTLERHRTQLELAHKEHNII
ref|ZP_02261478.1|-----VCIYLRKSRKDVEEERRAIEEGSSYNALERHHRKRLFAIAKAENHNII
ref|ZP_02257063.1|-----VCIYLRKSRKDVEEERRAIEEGSSYNALERHHRKRLFAIAKAENHNII
ref|YP_001513188.1|-----MEKVSIIYLRKSRADIEAESRGE-----TLAKHRSTLLKVAKEKKLSIV
RAAC01468MTPVSTFTGLEHVAIYLRKSRADLEAEARGE-----TLTKHRRALLELAKQYHSID
          * ***** * : * * : .          : * : * * : : * : *

ref|YP_146341.1|GIFEEIVSGESIAERSEIQKLLRELETGVADAVLMDIDRLGRGMDLQDQGILDRAFRYSG
ref|YP_001373830.1|DIFEEVVSGEYISERPMQKLLREVETGIADAVLMDLDRGRGMDVQDGTIYRVFRYSE
ref|ZP_02261478.1|DIFEEVASGESIQERPQMQLLRKLEGNEIDGVLVIDLDRGRGMDLQDQGILDRAFRYSS
ref|ZP_02257063.1|DIFEEVASGESIQERPQMQLLRKLEGNEIDGVLVIDLDRGRGMDLQDQGILDRAFRYSS
ref|YP_001513188.1|KIYEEIVSGESLMHRPEMLELLKNVSESKKYDAVLMDVDRLGRGNMQEQGLILETFKKSC
RAAC01468HIYEEIVSGELIVDRPEVQRLHAVREGKYSAVLMDIDRLGRGNQIDQGI IQAFKQSG
          * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *

ref|YP_146341.1|TKIITPTEVYDPESETWELVFGVKSIVSREELKVITKRLQGGRRDSAAKGRSISKPPPYG
ref|YP_001373830.1|TFIITPTEVINPNDENQELTFSIKSLIAREELKTIKRMQRGRKASAKEGKSISRVPAYG
ref|ZP_02261478.1|TKIITPTDVPDDESWELVFGIKSLISRQELKSITKRLQNGRIDSVEGKHHIGKKPPYG
ref|ZP_02257063.1|TKIITPTDVPDDESWELVFGIKSLISRQELKSITKRLQNGRIDSVEGKHHIGKKPPYG
ref|YP_001513188.1|TKIITPRKIYNLDEEFDEEYSEFEAFMARKELKMINRRLQNGRVSIEDGNYISPNNPYG
RAAC01468TLIITPRKVYNLEDELDEEFSEFEQFMARRELKIITRRMQRGRKLSAKEGKSITPYVPFG
          * * * * * : : : * * : : : * * * * : * * * * : * * * : * : * : *

ref|YP_146341.1|YLRDEQ---LKLYDPDTPSVWVVKIFEMVRDGHGRQAIAAELDRGVKPPDEKRAFWSPS
ref|YP_001373830.1|YLRDNN---LKLYDPDEKSWVPIKIFELMANGIGRQAIAQELDRGLIAPPEGE--YWNPS
ref|ZP_02261478.1|YLKDEN---LRLYDPDEKAWIVKKIFELMCDGKGRQMIAAELDRGLIDPPVTTRGAWDSS
ref|ZP_02257063.1|YLKDEN---LRLYDPDEKAWIVKKIFELMCDGKGRQMIAAELDRGLIDPPVTTRGAWDSS
ref|YP_001513188.1|YEIDEGKDYRTLKPHPEQAEEVKMIFEWYVSGLGSGKIANKLNDLGYKSYTGIP--WRSS
RAAC01468YKRDEN---LKLYDPDTPAPIVRQIFEWSAEGLGIIKIAKLLNEMGVAPRSCG--WQRT
          * * : * * * * : : * * * * * : * * * * : * * : * : * :

ref|YP_146341.1|TISAIKNEVYLGHIIWGKVYIKQNGRYKRK---KMPRERWYVKENAHAPLVSKELWEA
ref|YP_001373830.1|TISSIIKNEVYLGHIIWGKIRYIKQNGKYIRK---KVSKEWRQHDNAHPPLVSEELFQK
ref|ZP_02261478.1|TITSIIKNEVYTGIVWGKFKHKKRNGKYTRH---KNPQEKWIMYENAHAPIISKELFDA
ref|ZP_02257063.1|TITSIIKNEVYTGIVWGKFKHKKRNGKYTRH---KNPQEKWIMYENAHAPIISKELFDA
ref|YP_001513188.1|SVLNMLKNLVYTGKVVWGRKDIKKSTEVGKVKDTVTRPKEEWIIADGKHPALVSEELFEK
RAAC01468TVQHILKNEVYLGRIVWDKKRDVKTTEG-KYR-SIKRPREEWIVHENAHEPIISQELWDR
          : : : * * * * : : : * : : : : * * : : * : * : * : * :

ref|YP_146341.1|ANKAYRSRWRPSTVESKPLANPLAGLLKCEVCGYTMWYQPRKDRPHPLVRCNPCKKGQVQ
ref|YP_001373830.1|ANTAHSKRWRPPTIKTKKLSNPLAGLLCELCHGSMLYQPRKDRPNPQVRCVQPSCKGVQ
ref|ZP_02261478.1|ANEAHSSRHKPAVITSKELTNPLAGILKCKLCGYTMLIQTRKDRPHNYLRCNNPACKGKQ
ref|ZP_02257063.1|ANEAHSSRHKPAVITSKKLTNPLAGILKCKLCGYTMLIQTRKDRPHNYLRCNNPACKGKQ
ref|YP_001513188.1|AQEILNKRYHVPYQLENGITNPLAGILRCNCGASMVLRPYPDKD-HQVMCYNN-CGNKS
RAAC01468VVEIRKVNDR-TKDNLDLKNPFAGILRCKQCGRVMKRPFRPNRNGDTLQCYTTGCPTR
          . : : : * * * * : : : * : : : * : : : * : :

ref|YP_146341.1|KGALLPLVEEKILQSLAEFVDQFEVQE---DRKEQRSV-----IPLKQKAIEKKEKELRE
ref|YP_001373830.1|KGASLALVEQRILDLGLKQIIIESFEIQENMVQKQKRNKNN-----IHLQKQALEKKEQOMIN
ref|ZP_02261478.1|KQSVFNLVEEKLLYSLOQIVDEYQAQKVEEVEIDDSKL-----ISFKEKAIISKEKELKE
ref|ZP_02257063.1|KQSVFNLVEEKLLYSLOQIVDEYQAQKVEEVEIDDSKL-----ISFKEKAIISKEKELKE
ref|YP_001513188.1|---SKLKYVEKEVLAGLAEWCLQYRAQWDIDNKSRRKSKSLTSSIPVLEKAVENLQKELVE
RAAC01468--VLLSRLEDRLVLSIEEFVQSYAAQSSVKKRADNRQKKLEA----LHRQKKSISQKLSK
          : : * : * : : : : * : : : : : : : : : : : :

ref|YP_146341.1|LHKQKDALHDLLERGVYTIETFLERQQTIVSRIKKTQEEIDQLREEIAKEQLKEKNISEY
ref|YP_001373830.1|LQKQKNNLHDLLEKGVYDVETFLERQKSIAVRLKTTQKAIIEELKHEIKKILEKEKHIHEF
ref|ZP_02261478.1|LQAQKGNLHDLLEQGIYTVEIFLERQKNLVERITSIENDIEVLQKEIETEIQIHEHNKTEF
ref|ZP_02257063.1|LQAQKGNLHDLLEQGIYTVEIFLERQKNLVERITSIENDIEVLQKEIETEIQIHEHNKTEF
ref|YP_001513188.1|LENQKNSLHDFLERGIYDVDTYLSRQNLATRIDATKVSLSKAKLVLGQEMQREKAQIDI
RAAC01468LETQKNRFLDFLEQGIYDVQTFIERSKLVEQIDQAKEELKICEQAIEREMLQQQHEEEL
          * * * * * : * : * : * : : : : : : : : : : : : : : :

ref|YP_146341.1|VPTVKKVLDAYRLTDDVEKKNRLLKSVELEKATYLRKPEWTKKQDFTIQLYPRI
ref|YP_001373830.1|VPRIKNVLEAYATNDIEKKNRLLKSVELEKATYLRKKIWQRKDEFLIELYTRI
ref|ZP_02261478.1|IPALKTVIESYHKTTNIELKNQLLKTILSTVTTYRHPDWKTN-EFEIQVYFKI
ref|ZP_02257063.1|IPALKTVIESYHKTTNIELKNQLLKTILSTVTTYRHPDWKTN-EFEIQVYFKI
ref|YP_001513188.1|IPRIEKVLDVYPKINDPAHKNELLKSVDYAAYSKDYSK-RNDDFSRLRF---
RAAC01468IPAISEAIAMYRTASNAEIKNMALKSVIKEIHYYRPRSWSKSKEFEIDYFRI
          : * : : : * : : * * * : : : : : : : : : : : :

```

FIG. 88

| | |
|--------------------|---|
| ref YP_753434.1 | ----- |
| ref YP_753226.1 | EFYPHYPTHEICALPRTEINSWLSYNTRPHMNASWLFEISUBSPWLFEISTRGETTING |
| ref YP_752864.1 | ----- |
| ref ZP_01665148.1 | ----- |
| ref YP_001111903.1 | ----- |
| RAAC03178 | ----- |
| | |
| ref YP_753434.1 | ----- |
| ref YP_753226.1 | ENGBABIHYPTHEICALPRTEINSWLSYNTRPHMNASWLFEISUBSPWLFEISTRGETT |
| ref YP_752864.1 | ----- |
| ref ZP_01665148.1 | ----- |
| ref YP_001111903.1 | ----- |
| RAAC03178 | ----- |
| | |
| ref YP_753434.1 | ----- |
| ref YP_753226.1 | INGENGBABIHYPTHEICALPRTEINSWLSYNTRPHMNASWLFEISUBSPWLFEISTRG |
| ref YP_752864.1 | ----- |
| ref ZP_01665148.1 | ----- |
| ref YP_001111903.1 | ----- |
| RAAC03178 | ----- |
| | |
| ref YP_753434.1 | ----- |
| ref YP_753226.1 | ETTINGENGBABIHYPTHEICALPRTEINSWLSYNTRPHMNASWLFEISUBSPWLFEIS |
| ref YP_752864.1 | ----- |
| ref ZP_01665148.1 | ----- |
| ref YP_001111903.1 | ----- |
| RAAC03178 | ----- |
| | |
| ref YP_753434.1 | ----- |
| ref YP_753226.1 | TRGETTINGENGBABIHYPTHEICALPRTEINSWLSYNTRPHMNASWLFEISUBSPWLFE |
| ref YP_752864.1 | ----- |
| ref ZP_01665148.1 | ----- |
| ref YP_001111903.1 | ----- |
| RAAC03178 | ----- |
| | |
| ref YP_753434.1 | ----- |
| ref YP_753226.1 | EISTRGETTINGENGBABIHYPTHEICALPRTEINSWLSYNTRPHMNASWLFEISUBSP |
| ref YP_752864.1 | ----- |
| ref ZP_01665148.1 | EAXTRANSPSASEISFAMILYP-----RTEINTHERMSINU----- |
| ref YP_001111903.1 | ----- |
| RAAC03178 | ----- |
| | |
| ref YP_753434.1 | -----IHYPTHEICALPRTEINSWLSYNTRPHMNASWLFEISU |
| ref YP_753226.1 | WLFEISTRGETTINGENGBABIHYPTHEICALPRTEINSWLSYNTRPHMNASWLFEISU |
| ref YP_752864.1 | ----- |
| ref ZP_01665148.1 | -----SCARBXYDIVRANSNRGBEAXTRANSPSASEISFAMILYPRTEINTHERMSINU |
| ref YP_001111903.1 | ----- |
| RAAC03178 | -----MLLSNGTTYMQERVHWGVHHIAPVNCMHLHNIKFFRSESTRFPKN- |
| | |
| ref YP_753434.1 | ----- |
| ref YP_753226.1 | BSPWLFEISTRGETTINGENSCRESIGNIFICANCEE-IDENTITIESPSITIVESGAP- |
| ref YP_752864.1 | BSPWLFEISTRGETTINGENSCRESIGNIFICANCEE-IDENTITIESPSITIVESGAP- |
| ref ZP_01665148.1 | -----SCARBXYDIVRANSNRSCRESIGNIFICANCEE-IDENTITIESPSITIVESGAPS |
| ref YP_001111903.1 | ----- |
| RAAC03178 | -----FVRYLNEPFAESRVLVLRQG-----FFEHKEGLHPNLRERKWYPANHFKGAKSN |

ref|YP_753434.1| -SIRQNCIFSFEEDALKIQPKSRLEKIINTLDLKPVLCKLDKPGEIRVGPKPYPAYAMLNA
 ref|YP_753226.1| -SIRQNCIFSFEEDALKIQPKSRLEKIINTLDLKPVLCKLDKPGEIRVGPKPYPAYAMLNA
 ref|YP_752864.1| --IRQGYVFSFEEDAINLQPRSRLEIVLATLDFDDVITVLDKENKQHRGPTGYPFESKLNA
 ref|ZP_01665148.1| MYIRQKPLFSFDTLMQYQPKTRLAMVFESIDLHPLKTLPIKS--IRGPKGYSSAALIK
 ref|YP_001111903.1| MYLLQPNLFSFEELLKFEPETKLQKVLVLDLSPALNVVKRAV---LGPKGHCVGNMIRA
 RAAC03178 VYVRQAWLFSFEDEMEMSPCERRELFSTLDLSPYAAKLRSTP--QGAKPISREAILRA
 : * :***: :: .* : .: :*: : *.. :.*

ref|YP_753434.1| LIAMRLENMGFTTQLVERLTYDPLHRYVCGFEPFGTAPSKSCFSRFYSKLAQS-----
 ref|YP_753226.1| LIAMRLENMGFTTQLVERLTYDPLHRYVCGFEPFGTAPSKSCFSRFYSKLAQS-----
 ref|YP_752864.1| LIAMRVYNMATFTQLVERLTHDPSVLRVYNGCFDVFGEKVPSTATFSRFRYERLTQS-----
 ref|ZP_01665148.1| FLAMRLCSIPTVTLLVERLKTDLVFRYECGFSLEQPVPSLATFSRFFQKIAET-----
 ref|YP_001111903.1| LVAKQLEQIPTVAALVKRLSNDIRFRFQCGFSLSKPIPSESTFSRLIQKLAATDETNTKE
 RAAC03178 FLAAPLEGISTFTQLHRRLESCLRFRYQCGFSLHESIPSVSTLSRVFQAIVDK-----
 :.* : :*. : *.* * :*: ***. ** : :*. . :.

ref|YP_753434.1| GCLETFTSLVKQAEEMGLLDLSSVAIDATKVEAYEKSVPKNI IQDGNVADWGIKSDTN
 ref|YP_753226.1| GCLETFTSLVKQAEEMGLLDLSSVAIDATKVEAYEKSVPKNI IQDGNVADWGIKSDTN
 ref|YP_752864.1| EVLRELFFKQVTTAESMGLIDTSSIAIDASKVDANEKSVPKNIKDDGQSANWGSKLDTN
 ref|ZP_01665148.1| DSLQVLFSAVDTAIQDKVISGEVVAIDASAIDSYEKVPKKEKLSNGDSATWGAKLDTH
 ref|YP_001111903.1| SVMKQIFDNLVKSADMDGLIDSNCAIDSSKIDAYEKSVPKDLQGD--KTANWGAQRDTH
 RAAC03178 GVAATLFAELVRQCRDEGLIEGEHVAIDSTAIHAYERKHPRSGVQPS--NRANWGAQFADF
 :* * . . :.. : :***: :. : * :*. : . . * * * * *:

ref|YP_753434.1| GNPIKWFGYKLIHIGTDVKSGPLIAMKVTTPANYSDSSVALELVEKCCANTQS--KIVYFLM
 ref|YP_753226.1| GNPIKWFGYKLIHIGTDVRSGLPIAMKVTTPANCSDSSVALELVEKCCANTQS--KIVYFLM
 ref|YP_752864.1| GNQITWFGYKLIHIGTDVKSELPAVALSITPASTNDGIMAESILEECSNNLNS--KPQYYLM
 ref|ZP_01665148.1| GNQHVWFGYKLIHIGTDVKSELPIAVKVTTPANRNDVTQAIPMLDEIKH-----QPKYYCM
 ref|YP_001111903.1| GNQISWFGYKAHVAVDCHSELPIAIMVTTPANTHDAKMAIPLIELVNKALEDSKKPKYYTM
 RAAC03178 GNKLAWFGYKVHLAVDAKSELPMALTVTTPANVYDGEMAIPLMEELHHHDWR---IRFVLM
 ** ***** *..* :* ***: :***. * * :.: : *

ref|YP_753434.1| DAGYDHREIYSVIRDKYHAQAI IALNKRGAQPPPEGFDWDGTPICSARYRMVYWGGSY--Q
 ref|YP_753226.1| DAGYDHREIYSVIRDKYHAQAI IALNKRGAQPPPEGFDWDGTPICSARYRMVYWGGSY--Q
 ref|YP_752864.1| DAGYDQKSIYELIRDKYKAQAI IPLNHRGAKEPPEGLDWDATPICSAGYRMAYWGGS--N
 ref|ZP_01665148.1| DMGYDAKDVYQAAYER--QAQAI IPLNRRKEKLPEGMENRTPCISMGYPMVYWGGEREK
 ref|YP_001111903.1| DMGYDSKDIYSVMNDFNAQAI IPINSRGSKDHPEGDFDGTPTICSMGQRMVFGSDAKA
 RAAC03178 DAGYDQTKNYEAARAL--GAQAI IPMNRRNEKEPPEGMDFDGTPTCTMGYRMTYWGVS--N
 * *** . * . *****.* * * *** * : * * : *.:**

ref|YP_753434.1| GVNKFRCPHIMGKDCPFGSAWCSDSNYGMVVKTKVKDDPRLFSSPHRGSANWQKQYNLR
 ref|YP_753226.1| GVNKFRCPHIMGKDCPFGSAWCSDSNYGMVVKTKVKDDPRLFSSPHRGSANWQKQYNLR
 ref|YP_752864.1| GVNKFRCPHVMGKDCPFGSSWCSDSNYGMVVKTRARQDSRLFIVPHRGTSNWKLLYNKR
 ref|ZP_01665148.1| GILKFRCPHVCVKVNCENGSAWCSPSNYGLVIKKKVEDDPRSFCTPHRGTREREKLYAER
 ref|YP_001111903.1| GTNKYRCPHVMGKDCPFGSAWCSPSYGLVVKTKVKDDPRMNCIPARGTKNWQSLYNKR
 RAAC03178 AWLKFRCPHATGQVDCPLGMAACSASNYGMVVKHLDVRRYANPHRGSRTWKMLYDER
 . * :**** * : :*** : * * * : * * : * * : * * : * * : *

ref|YP_753434.1| TYSERCFSRFKENLGLDGLNVRKITKVETHAYLCAITMIAAVIAINQ-----
 ref|YP_753226.1| TYSERCFSRFKENLGLDGLNVRKITKVETHAYLCAITMIAAVIAINQ-----
 ref|YP_752864.1| TSVERCFGRLEHLETLGLNVRGIKKVKTAYLSVITMIASVIAINKDKSSTDIA
 ref|ZP_01665148.1| TSVERAFSRLKEQLGANT--VRVQGIKKVTAHMLCCIALLAGTIAVNR-----
 ref|YP_001111903.1| TSVERCFGRLEKQHLGANS--IRTRGLEKVTLHITLSCIALLAGSIAVAKTKRIEQAA
 RAAC03178 TAVERCFARLKEWLTLDG--VHVRGIEKVTVHAYIHAIVLLASALAMHRTNRIEQVA
 * **.*.**: * : :.. : ** * : *.:*. :*: :

FIG. 89

```

ref|YP_001126744.1|
gb|AAB52611.1|
sp|P52026|DP01_BACST
gb|ABM97416.1|
dbj|BAF33373.1|
RAAC01937
-----MNFEVNIINSLKCVNSACVPLGVPRFSC
MNFHPRHGPVRTFWHLSAPPRRTYSLHGLRSMNFVNIINSLKCVNSACVPLGVPRFSC

ref|YP_001126744.1|
gb|AAB52611.1|
sp|P52026|DP01_BACST
gb|ABM97416.1|
dbj|BAF33373.1|
RAAC01937
-----KLVLIDGNSVAYRAFFALPLLHNDKG
-----KLVLIDGNSVAYRAFFALPLLHNDKG
-----NKLVLIDGNSVAYRAFFALPLLHNDKG
-----NKLVLIDGNSVAYRAFFALPLLHNDKG
RHKTTTRADKGNVIEWSGTTVRQTARSLGWPMPSKLVLIDGNSILYRAFFALPPLTARDG
RHKTTTRADKGNVIEWSGTTVRQTARSLGWPMPSKLVLIDGNSILYRAFFALPPLTARDG
*****:***** * . *

ref|YP_001126744.1|
gb|AAB52611.1|
sp|P52026|DP01_BACST
gb|ABM97416.1|
dbj|BAF33373.1|
RAAC01937
IHTNAVYGFTMMLNKILAEEQPTHLVAFDAGKTTFRHETFDQYKGGROQTPELSEQFP
IHTNAVYGFTMMLNKILAEEQPTHLVAFDAGKTTFRHETFDQYKGGROQTPELSEQFP
IHTNAVYGFTMMLNKILAEEQPTHLVAFDAGKTTFRHETFDQYKGGROQTPELSEQFP
IHTNAVYGFTMMLNKILAEEQPTHLVAFDAGKTTFRHETFDQYKGGROQTPELSEQFP
TPTNAVYGFTTMI LRLMSDEKPTH LAVAFD KSKTTFRHADFAAYKGTROETPDELVQQFP
TPTNAVYGFTTMI LRLMSDEKPTH LAVAFD KSKTTFRHADFAAYKGTROETPDELVQQFP
***** *: :::::*****: ***** * *** **:* ** :***

ref|YP_001126744.1|
gb|AAB52611.1|
sp|P52026|DP01_BACST
gb|ABM97416.1|
dbj|BAF33373.1|
RAAC01937
LLRELLKAYRIPAYELDHYEADDIIGTLAARAEQEGFEVKIISGDRDLTQLASRHVTVDI
LLRELLKAYRIPAYELDHYEADDIIGTLAARAEQEGFEVKIISGDRDLTQLASRHVTVDI
LLRELLKAYRIPAYELDHYEADDIIGTLAARAEQEGFEVKIISGDRDLTQLASRHVTVDI
LLRELLRAYRIPAYELNYEADDIIGTLAARAEQEGFEVKIISGDRDLTQLASRHVTVDI
LARRTLEALSIPMVEIDQYEADDVIGTLAKRAAEAGFDVRVVS GDKDLLQLVDDRIHVLL
LARRTLEALSIPMVEIDQYEADDVIGTLAKRAAEAGFDVRVVS GDKDLLQLVDDRIHVLL
* * . * . * * * *: :::*****:*****: * * . * * *: :::*****: * * . * . : : * :

ref|YP_001126744.1|
gb|AAB52611.1|
sp|P52026|DP01_BACST
gb|ABM97416.1|
dbj|BAF33373.1|
RAAC01937
TKKGITDIEPYTPETVREKY-GLTPEQIVDLKGLMGDKSDNIPGVPGIGECTAVKLLKQF
TKKGITDIEPYTPETVREKY-GLTPEQIVDLKGLMGDKSDNIPGVPGIGECTAVKLLKQF
TKKGITDIEPYTPETVREKY-GLTPEQIVDLKGLMGDKSDNIPGVPGIGECTAVKLLKQF
TKKGITDIEPYTPETVREKY-GLTPEQIVDLKGLMGDKSDNIPGVPGIGECTAVKLLKQF
TRKGITEMEHFDEQAVARRYPGLKPAQVIDLKGLMGDPSDNIPGVPGVGEKTALKLLASF
TRKGITEMEHFDEQAVARRYPGLKPAQVIDLKGLMGDPSDNIPGVPGVGEKTALKLLASF
* .*****: * : : * . * * . * *: :::*****: *****:*****: * * . *

ref|YP_001126744.1|
gb|AAB52611.1|
sp|P52026|DP01_BACST
gb|ABM97416.1|
dbj|BAF33373.1|
RAAC01937
GTVENVLASIDEVKGEKLEKLNLRQHRDLALLSKQLASICRDAPVELSLDDIVYEGQDREK
GTVENVLASIDEVKGEKLEKLNLRQHRDLALLSKQLASICRDAPVELSLDDIVYEGQDREK
GTVENVLASIDEIKGEKLEKLNLRQHRDLALLSKQLAAICRDAPVELSLDDIVYEGQDREK
GTVENVLASIDEIKGEKLEKLNLRQHRDLALLSKQLAAICRDAPVELSLDDIVYEGQDREK
GTVEGVYDHLDEVQGQKLRLERLEQHREDAFLSKRLATACDAPIEVDLETTRYEGPDPAR
GTVEGVYDHLDEVQGQKLRLERLEQHREDAFLSKRLATACDAPIEVDLETTRYEGPDPAR
****. * *: :::*****: * . * . * *: :::*****: *****:*****: * * . *

ref|YP_001126744.1|
gb|AAB52611.1|
sp|P52026|DP01_BACST
gb|ABM97416.1|
dbj|BAF33373.1|
RAAC01937
VIALFKELGFQSFLEKMAAPAAEGEKP-----LEEMEFAIVDVI-----TEE
VIALFKELGFQSFLEKMAAPAAEGEKP-----LEEMEFAIVDVI-----TEE
VVALFQELGFQSFLEKMAVQTDGEKP-----LAGMDFAIADSV-----TDE
VVALFQELGFQSFLEKMESSPSEEEKP-----LAKMAFTLADRV-----TEE
AIAWFRELDFRSLVDKISEEMSHDSTPTPSPAAASGASSEWSSFAYGLIEDAGAWQEAI
AIAWFRELDFRSLVDKISEEMSHDSTPTPSPAAASGASSEWSSFAYGLIEDAGAWQEAI
. : * * . * . * . * . * : : : : : : : : : : : : : : : :

ref|YP_001126744.1|
gb|AAB52611.1|
sp|P52026|DP01_BACST
gb|ABM97416.1|
dbj|BAF33373.1|
RAAC01937
MLADKAALVVEVMEENYHDAPIVIGIALVNEHGRFFMRPETALADSQFLAWLADETKKKSM
MLADKAALVVEVMEENYHDAPIVIGIALVNEHGRFFMRPETALADSQFLAWLADETKKKSM
MLADKAALVVEVVDNYHDAPIVIGIALVNEHGRFFMRPETALADPKFLAWLGDETKKKSM
MLADKAALVVEVMEENYHDAPIVIGIAVNEHGRFFMRPETALADPQFVAVWLGDDETKKKSM
SFSEPVGMMDLADPDYHRAEIRGMAVATPKRAYYVRFGERLELSDVRPWLVS-D-RPKVA
SFSEPVGMMDLADPDYHRAEIRGMAVATPKRAYYVRFGERLELSDVRPWLVS-D-RPKVA
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

ref|YP_001126744.1|
 gb|AAB52611.1|
 sp|P52026|DP01_BACST
 gb|ABM97416.1|
 dbj|BAF33373.1|
 RAAC01937

FDKRAVVALKWKGIELRGVA--FDLLLAAYLLNPAQDAGDIAAVAKMKQYEAVRSDEAV
 FDKRAVVALKWKGIELRGVA--FDLLLAAYLLNPAQDAGDIAAVAKMKQYEAVRSDEAV
 FDSKRAAVALKWKGIELRGV--FDLLLAAYLLDPAQAAGDVAAVAKMHQYEAVRSDEAV
 FDSKRAAVALKWKGIELCGVS--FDLLLAAYLLDPAQGVDDVAAAAMKQYEAVRSDEAV
 FDLKSMAFALDAHGIIGLTSECGWQDVKLAAYLLNPQDG-----EVELSDVFA
 FDLKSMAFALDAHGIIGLTSECGWQDVKLAAYLLNPQDG-----EVELSDVFA
 ** * : ** * . * : ***** : * : ** .

ref|YP_001126744.1|
 gb|AAB52611.1|
 sp|P52026|DP01_BACST
 gb|ABM97416.1|
 dbj|BAF33373.1|
 RAAC01937

YKGKVKRSLPDEQTLAEHLVRKAAAIWALEQPFMDLRNNEQDQLLTKEQLAAILAEM
 YKGKVKRSLPDEQTLAEHLVRKAAAIWALEQPFMDLRNNEQDQLLTKEQLAAILAEM
 YKGKAKRTVPDEPTLAHLVRKAAAIWALEEPLMDLRRNEQDRLLTEQLAGILANM
 YKGKAKRAVPDEPVLAEHLVRKAAAIWALERPFLDELRRNEQDRLLVELEQLPLSSILAEM
 RERGQELPAWEEGEREKWLAYTASQLPPLFESLAYTIRMQEMERLYQEVFLAFVLAKM
 RERGQELPAWEEGEREKWLAYTASQLPPLFESLAYTIRMQEMERLYQEVFLAFVLAKM
 : * : . : * : * . * : : * . . : : * : * : * : * : * : * : *

ref|YP_001126744.1|
 gb|AAB52611.1|
 sp|P52026|DP01_BACST
 gb|ABM97416.1|
 dbj|BAF33373.1|
 RAAC01937

EFTGVNVDTKRLEQMGSELAELRAIEQRIYELAGQEFNINSPKQLGVILFEKLQPLVLK
 EFTGVNVDTKRLEQMGSELAELRAIEQRIYELAGQEFNINSPKQLGVILFEKLQPLVLK
 EFTGVKVDTKRLEQMGAELEQLQAVERRIYELAGQEFNINSPKQLGTVLFDKQLPLVLK
 EFAGVKVDTKRLEQMGSELAELRTVEQRIYELAGQEFNINSPKQLGVILFEKLQPLILK
 EITGFYVNREKLVAFGQELTERIKRITQEIYDLAGTSFNLNINSPKQLGEILFDKGLPLALK
 EITGFYVNREKLVAFGQELTERIKRITQEIYDLAGTSFNLNINSPKQLGEILFDKGLPLALK
 * : * . * : : * : * * : * : : : : * : * : * : * : * : * : *

ref|YP_001126744.1|
 gb|AAB52611.1|
 sp|P52026|DP01_BACST
 gb|ABM97416.1|
 dbj|BAF33373.1|
 RAAC01937

KTKTGYSTSADVLEKLAPHHEIVENILHYRQLGKLQSTYIEGLLKVVRPDTGKVHTMFNQ
 KTKTGYSTSADVLEKLAPHHEIVENILHYRQLGKLQSTYIEGLLKVVRPDTGKVHTMFNQ
 KTKTGYSTSADVLEKLAPHHEIVEHILHYRQLGKLQSTYIEGLLKVVHPVTGKVHTMFNQ
 KTKTGYSTSADVLEKLAPHHEIVENILHYRQLGKLQSTYIEGLLKVVRPDTKKVHTIFNQ
 KTKTGYSTSADVLEKLAPHHEIVQKILDYRLLAKLQSTYIEGLLKVIRKETGRVHTRFHQ
 KTKTGYSTSADVLEKLAPHHEIVQKILDYRLLAKLQSTYIEGLLKVIRKETGRVHTRFHQ
 ***** : * : * : * : * : * : * : * : * : * : * : * : *

ref|YP_001126744.1|
 gb|AAB52611.1|
 sp|P52026|DP01_BACST
 gb|ABM97416.1|
 dbj|BAF33373.1|
 RAAC01937

ALTQTGRLSSAEPNLQNIPIRLEEGRKIRQAFVPSEPDWLI FAADYSQIELRVLAHIADD
 ALTQTGRLSSAEPNLQNIPIRLEEGRKIRQAFVPSEPDWLI FAADYSQIELRVLAHIADD
 ALTQTGRLSSVEPNLQNIPIRLEEGRKIRQAFVPSEPDWLI FAADYSQIELRVLAHIAED
 ALTQTGRLSSSTEPNLQNIPIRLEEGRKIRQAFVPSEPDWLI FAADYSQIELRVLAHIAED
 TLTATGRLSSSTEPNLQNIPIRLEEGRRLRQVFEPYKDWVIFAADYSQIELRILAHISGD
 TLTATGRLSSSTEPNLQNIPIRLEEGRRLRQVFEPYKDWVIFAADYSQIELRILAHISGD
 : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *

ref|YP_001126744.1|
 gb|AAB52611.1|
 sp|P52026|DP01_BACST
 gb|ABM97416.1|
 dbj|BAF33373.1|
 RAAC01937

DNLIEAFQRDLDIHTKTAMDIFHVSEEEVTANMRRQAKAVNFGIVYGISDYGLAQNINIT
 DNLIEAFQRDLDIHTKTAMDIFHVSEEEVTANMRRQAKAVNFGIVYGISDYGLAQNINIT
 DNLIEAFRRGLDIHTKTAMDIFHVSEEDVTANMRRQAKAVNFGIVYGISDYGLAQNINIT
 DNLMEAFRRDLDIHTKTAMDIFQVSEEDVTANMRRQAKAVNFGIVYGISDYGLAQNINIS
 EALIDAFRRDMDIHTRTAADVFEPPEQVTSLMRRQAKAVNFGIVYGISDFGLAQNINIP
 EALIDAFRRDMDIHTRTAADVFEPPEQVTSLMRRQAKAVNFGIVYGISDFGLAQNINIP
 : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *

ref|YP_001126744.1|
 gb|AAB52611.1|
 sp|P52026|DP01_BACST
 gb|ABM97416.1|
 dbj|BAF33373.1|
 RAAC01937

RKEAAEFIERFYFASFPGVKQYMENIVQEAQKQGYVTLLHRRRYLPDITSRNFNVRSAE
 RKEAAEFIERFYFASFPGVKQYMENIVQEAQKQGYVTLLHRRRYLPDITSRNFNVRSAE
 RKEAAEFIERFYFASFPGVKQYMDNIVQEAQKQGYVTLLHRRRYLPDITSRNFNVRSAE
 RKEAAEFIERFYFASFPGVKRYMENIVQEAQKQGYVTLLHRRRYLPDITSRNFNVRSAE
 QKEAKRFIESYFEKFPVKRYMDEIVKQARERGVYVTLMNRRRYLPDIHSRNYQLRSFAE
 QKEAKRFIESYFEKFPVKRYMDEIVKQARERGVYVTLMNRRRYLPDIHSRNYQLRSFAE
 : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *

ref|YP_001126744.1|
 gb|AAB52611.1|
 sp|P52026|DP01_BACST
 gb|ABM97416.1|
 dbj|BAF33373.1|
 RAAC01937

RTAMNTPIQGSAAIDIKKAMIDLARLKEEQQLARLLLQVHDELILEAPKEEIERLCELV
 RTAMNTPIQGSAAIDIKKAMIDLARLKEEQQLARLLLQVHDELILEAPKEEIERLCELV
 RTAMNTPIQGSAAIDIKKAMIDLSVRLREERLQARLLLQVHDELILEAPKEEIERLCLRV
 RMAMNTPIQGSAAIDIKKAMIDLARLKEERLQARLLLQVHDELILEAPKEEIERLCLRV
 RTAMNTPIQGSAAIDIKLAMVRIDRAMRDAQMDARMLLQVHDELIFECPKDELAALVVLV
 RTAMNTPIQGSAAIDIKLAMVRIDRAMRDAQMDARMLLQVHDELIFECPKDELAALVVLV
 * : * : * : * : * : * : * : * : * : * : * : * : * : * : *

ref|YP_001126744.1|
 gb|AAB52611.1|
 sp|P52026|DP01_BACST
 gb|ABM97416.1|
 dbj|BAF33373.1|
 RAAC01937

PEVMEQAVTLRVPLKVDYHYGPTWYDAK
 PEVMEQAVTLRVPLKVDYHYGPTWYDAK
 PEVMEQAVTLRVPLKVDYHYGPTWYDAK
 PEVMEQAVTLRVPLKVDYHYGPTWYDAK
 RDNMENAMTLSVPLKVDYHYGPTWYDAK
 RDNMENAMTLSVPLKVDYHYGPTWYDAK
 : * : * : * : * : * : * : * : * : * : * : * : *

FIG. 90

```

ref|ZP_00651175.1|      -----AHRIALDPNNVQATHLSRVAGVARFAYNWALAEWRHQYEACTLDSALPKPSQHSL
gb|ACA11657.1|          -----AHRIALDPNNVQATHLSRVAGVARFAYNWALAEWRHQYEACTLDSALPKPSQHSL
ref|YP_001681573.1|    MKINRAYRYELKPNVAQRILLAQHAGTARFAYNWGLARRITLYEAEKKS-----TNAIAQ
RAAC01372               MKIHRAIRYELAPNRMQRVMLAKHAGVARFAYNWGLARRIALHEEAGQS-----TNAIEQ
ref|YP_473713.1|      -----NNRQKTLLAAQHAGVARHAYNWGLAICKQAVESKQKL-----PTAIDL
ref|YP_475384.1|      -----NNRQKTLLAAQHAGVARHAYNWGLAICKQAVESKQKL-----PTAIDL
                        * *      : : **.*.*.*.*.*.* *      *      ..

ref|ZP_00651175.1|    RRQLNAIKREQFPWMGEVTKNAPQMAIIQLGQAFQNFAG-----RAKYPKFRKKGAHDR
gb|ACA11657.1|        RRQLNAIKREQFPWMGEVTKNAPQMAIIQLGQAFQNFAG-----RAKYPKFRKKGAHDR
ref|YP_001681573.1|    HRELNLQKQTEFPWMYEVSKCAPQEALRDLDKAFKNFFRGLKLGQKVGFPFRKKKGQDDS
RAAC01372               HRELNLRLKKTDFPWMYEVSKCAPQEALRDLDRAFQHFRRGLNEGRKVGFPFRKKKGRRDS
ref|YP_473713.1|      HKRLVAEVKKENPWYQVSKCAPQALRNLEQAFKRWRS-----LGKFPRFRKKGVRDS
ref|YP_475384.1|      HKRLVAEVKKENPWYQVSKCAPQALRNLEQAFKRWRS-----LGKFPRFRKKGVRDS
                        : : *      : : **      : : * * * * * : : * : * : : *      : : * : * : * * *

ref|ZP_00651175.1|    FTLTNDQFDLDASRIRIPRLGWVRMRETLRFAGRIMSATVSRVAARWFVSITVDVPDPSSH
gb|ACA11657.1|        FTLTNDQFDLDASRIRIPRLGWVRMRETLRFAGRIMSATVSRVAARWFVSITVDVPDPSSH
ref|YP_001681573.1|    FRLT-GAIKVNKGAVQLPRLGVIRLKEEPWIRGRILSATVSRVADRWVSLACEVEIPDP
RAAC01372               FRLT-GSIRVLDNAVQLPRLGRIRLKEKPYVEGRILSATVSKREADRWVSLATETETIPDP
ref|YP_473713.1|      FYLE-GSIRISGDRIKVPILGWLRCAELLPTA-TPKNVVISLRAGHWYVSFKYEAPAP-Q
ref|YP_475384.1|      FYLE-GSIRISGDRIKVPILGWLRCAELLPTA-TPKNVVISLRAGHWYVSFKYEAPAP-Q
                        * *      . : : . : : * * * : * *      . . . . . * : * : * : : . *

ref|ZP_00651175.1|    LPQAEHQGVVGVLDGLVLAALATLSTGETICGPRPHRALLGRVRRLSRSVSRKVKDSANRHK
gb|ACA11657.1|        LPQAEHQGVVGVLDGLVLAALATLSTGETICGPRPHRALLGRVRRLSRSVSRKVKDSANRHK
ref|YP_001681573.1|    APVLGP--IVGIDVGLNHFAVLSDGTKVEAPKPLDKFLKRLRLRSKKHSRKQKGSSTNRKK
RAAC01372               VPPSGE--PVGVDLGVSWFLTLSDGTKIKAPKPLRRYLRLRLERLSKRHSRKKPGSRNRKK
ref|YP_473713.1|      VEKTGE--VVGVDLGINRLATCSDGEVFENPKPYKAKKRLARLQRRLSRKQKGSANRKK
ref|YP_475384.1|      VEKTGE--VVGVDLGINRLATCSDGEVFENPKPYKAKKRLARLQRRLSRKQKGSANRKK
                        ** : * : : : . * *      . * : *      * : * : : * * * . * * : *

ref|ZP_00651175.1|    AKATLAHLHARIAAIRLDALHKLTSDLTRRFTTIGIENLNKGMVKNRHLARSIADMGGF
gb|ACA11657.1|        AKATLAHLHARIAAIRLDALHKLTSDLTRRFTTIGIENLNKGMVKNRHLARSIADMGGF
ref|YP_001681573.1|    SALALARLHRRIRNIRQDFLHKLTTLNLAKTKEIVMEDLNVRGMMRNGLARHIADVGWG
RAAC01372               SALALARLHRRIRNIRQDFLHKVTTELAKTKEIVMEDLNVRGMMRNGLARHIADVGWG
ref|YP_473713.1|      AVAQLAKAHKRVADIRQDNLHKLTITYLAKKYRVVVIEDLQVKNLLKNHKLALAGALSDCGFY
ref|YP_475384.1|      AVAQLAKAHKRVAGIRQDNLHKLTITYLAKKYRVVVIEDLQVKNLLKNHKLALAGALSDCGFY
                        :      ** : * : : * * * * * : * : :      : : * : * : * : : * * * : * : * :

ref|ZP_00651175.1|    EFRRQLEYKAAMRGQVVVADRFFASSKMCSTCGHTLKEPLSVRQWACLGCGRHHRDV
gb|ACA11657.1|        EFRRQLEYKAAMRGQVVVADRFFASSKMCSTCGHTLKEPLSVRQWACLGCGRHHRDV
ref|YP_001681573.1|    EFRRQLGYKTAWYGSKLTLVHRFYPSKTCACGYVREELPLSVREWDPCSGVHHRDH
RAAC01372               EFRRMLTYKCAWYGSELIAPRFYASSKTCACGYVISELPLSVREWTCFACSTRHHRDI
ref|YP_473713.1|      EFRRQLEYKARLYGCQVVVADRFYPSQLCSRCGHR-QKMPQLQERMFCPCCGLELDRDL
ref|YP_475384.1|      EFRRQLEYKARLYGCQVVVADRFYPSQLCSRCGHR-QKMPQLQERMFCPCCGLELDRDL
                        **** * * *      * : : . ** : * : * * * : : * : * : * * : * * . * *

ref|ZP_00651175.1|    NAAVNLKNMA--VSSTVS--ACGEEGTG-----
gb|ACA11657.1|        NAAVNLKNMA--VSSTVS--ACGEEGTG-----
ref|YP_001681573.1|    NAAKNL-----
RAAC01372               NAAKNLLRIG--TASSAGSDACGDPGGGALGC
ref|YP_473713.1|      NAALNLLRWYRTTSSSGSDACGDPGG-----
ref|YP_475384.1|      NAALNLLRWYRTTSSSGSDACGDPGG-----
                        *** **

```

FIG. 91

```

ref|ZP_01695687.1|-----VGPSQLLSKLCDEIELEKTINDLVKWDVSRCHISPGTRIKALVLN
ref|ZP_01695982.1|-----VGPSQLLSKLCDEIELEKTINDLVKWDVSRCHISPGTRIKALVLN
ref|ZP_01695971.1|-----VGPSQLLSKLCDEIELEKTINNVLKWDVSRCHISPGTRIKALVLN
ref|ZP_01695655.1|-----PQSCKLCDEIELEKTINNVLKWDVSRCHISPGTRIKALVLN
ref|YP_430569.1|-----MRFFRAGPAALISRLCDVLKIAEIIDAVVDWDPAQCHLSPGNRVKALIIN
RAAC00062MREAFQLFGPVRSYVMGPAPVLARLIDELKWVEIIDFVFRP--DSKLSVGLRTKALLVN
      .: * * : : * : . *      .: * * * * : : *

ref|ZP_01695687.1|ILCSGKPLYKVHEFYQNLDSEMLF--DTSVSPDQLNDDALGRALDYLYEA-EAWKVYSTL
ref|ZP_01695982.1|ILCSGKPLYKVHEFYQNLDSEMLF--DTSVSPDQLNDDALGRALDYLYEA-EAWKVYSTL
ref|ZP_01695971.1|ILCSGKPLYKVHEFYQNLDSEMLF--DTSISPDQLNDDALGRALDYLYKA-EAWKVYSTL
ref|ZP_01695655.1|ILCSGKPLYKVHEFYQNLDSEMLF--DTSISPDQLNDDALGRALDYLYKA-EAWKVYSTL
ref|YP_430569.1|LLVDREALYHVERFYENQDLEVLFGAEQQVRPEDFNDDALGRALDKLFTSGQLKKLFSSI
RAAC00062IGTNREALYRVVEFYAQRDVEVLLG--SGVSADDLHDDALARALDALYDA-GLEALYARI
      : . : . * : * : : * * : : : : * * * * * : : : : :

ref|ZP_01695687.1|ALKTLKKLNLP-----IGILHNDTTSISVYGEYKQPKEDGLQITYGSKAHRPDLKQIVL
ref|ZP_01695982.1|ALKALKKKLNLP-----IGILHNDTTSISVYGEYKQPKEDGLQITYGSKAHRPDLKQIVL
ref|ZP_01695971.1|ALKALKKKLNLP-----IGVLHNDTTSISVYGEYKQPKEDGLQITYGSKAHRPDLKQIVL
ref|ZP_01695655.1|ALKALKKKLNLP-----IGVLHNDTTSISVYGEYKQPKEDGLQITYGSKAHRPDLKQIVL
ref|YP_430569.1|ALTAATHNVS-----IAGIHVDTTSISVQAYDG--EGDLDTITFGFSKDHPRDLKQFLI
RAAC00062ALHTLRRLRVLSDSNELIPIHADTTSLSMTGEYLD--QTAFRIDRGFSKDHPRDLKQIVF
      * * : . : : : * * * * : * * : : * * * * * * * : :

ref|ZP_01695687.1|GMGVTPERIPILAKVENGNTSDKSWNVEFIQKMRKILSHEDWKNNLIYQADSALITTENLA
ref|ZP_01695982.1|GMGVTPERIPILAKVENGNTSDKSWNVEFIQKMRKILSHEDWKNNLIYQADSALITTENLA
ref|ZP_01695971.1|GMGVTPERIPILAKVENGNTSDKSWNVEFIQKMRKILSHEDWKNNLIYQADSALITTENLA
ref|ZP_01695655.1|GMGVTPERIPILAKVENGNTSDKSWNVEFIQKMRKILSHEDWKNNLIYQADSALITTENLA
ref|YP_430569.1|GLTVNRDGLPILAQSLDGNSSDKSWYPQVIEELVQTFKPEKLKEVIFVADCALVTKDNL
RAAC00062GL-CTVHGLGLCANVNPGLDDHTWNFENIQQLLSQLDEETRKRSVYVADAALVTKDNL
      * : . . : : * : * * . : * : : : . . * * . : * * * * : : *

ref|ZP_01695687.1|EIQQQN-----LSFISRLPDTFGLSTELKKEAWLLNN-WERVGSLSNKKDAAIYQIQAFE
ref|ZP_01695982.1|EIQQQN-----LSFISRLPDTFGLSTELKKEAWLLNN-WERVGSLSNKKDAAIYQIQAFE
ref|ZP_01695971.1|EIQQQN-----LSFISRLPDTFGLSTELKKEAWLLNN-WERVGSLSNKKDAAIYQIQAFE
ref|ZP_01695655.1|EIQQQN-----LSFISRLPDTFGLSTELKKEAWLLNN-WERVGSLSNKKDAAIYQIQAFE
ref|YP_430569.1|LLVQEEGNKPALQFISLLPENFGLNKEIKAEAFRTGT-WQEI GKLSPKKDAACYKSQSFV
RAAC00062LLAED-----FHFISRLPGTYKLSLSEDLKRAAWEKENSWEKVGRLAEAEADSAHYRIQAFR
      : : : : : * * * * : * . : * * : . * : : * * : : * * * : *

ref|ZP_01695687.1|RQIQNLFPYRFLVVHNSNNLDQRKEKTLNRAIEKEEIQKKEIEKLSKVIFHCKEDALEAIQ
ref|ZP_01695982.1|RQIQNLFPYRFLVVHNSNNLDQRKEKTLNRAIEKEEIQKKEIEKLSKVIFHCKEDALEAIQ
ref|ZP_01695971.1|RQIQNLFPYRFLVVHNSNNLDQRKEKTLNRAIEKEEIKLKEIEKLSKVIFHCKEDALEAIQ
ref|ZP_01695655.1|RQIQNLFPYRFLVVHNSNNLDQRKEKTLNRAIEKEEIQKKEIEKLSKVIFHCKEDALEAIQ
ref|YP_430569.1|REIDGRDYRLIVVHSTTLDKREKNSLLKKWAKQREVLEKAAKDLRRPFACKADARKAIE
RAAC00062RTLYGRTYRFVVVRSSSIDTRKERKLEVLKREKAALEKAAKAMSQNVYSCEQDAQMAMQ
      * : . * * : * * : * * * * . . : . * * : : * : : * * * : :

ref|ZP_01695687.1|SFKKKQKASYFCYKLVQMSDQPIKRRKGRPKKDEQTK-REQIYQIQLESFEKDHFIE
ref|ZP_01695982.1|SFKKKQKASYFCYKLVQMSDQPIKRRKGRPKKDEQTK-REQIYQIQLESFEKDHFIE
ref|ZP_01695971.1|SFKKKQKASYFCYKLVQMSDQPIKRRKGRPKKDEQTK-REQIYQIQLESFEKDHFIE
ref|ZP_01695655.1|SFKKKQKASYFCYKLVQMSDQPIKRRKGRPKKDEQTK-REQIYQIQLESFEKDHFIE
ref|YP_430569.1|LFLREYR-----
RAAC00062TFMHEHRATLHPISARICAEQVQAKRARRGRPRKDDPPFPVHTQYRVEVAILPPSEERVQ
      * : : :

ref|ZP_01695687.1|NKKKMLSTFVLITNKLDEETLSNQEVLRVYKGQSAETRFRLIKDSQMIDAIYLTTPERV
ref|ZP_01695982.1|NKKKMLSTFVLITNKLDEETLSNQEVLRVYKGQSAETRFRFRLIKDSQMIDAIYLTTPERV
ref|ZP_01695971.1|NKKKMLSTFVLITNKLDEETLSNQEVLRVYKGQSAVETRFRFRLIKDSQMIDAIYLTTPERV
ref|ZP_01695655.1|NKKKMLSTFVLITNKLDEETLSNQEVLRVYKGQSAETRFRFRLIKDSQMIDAIYLTTPERV
ref|YP_430569.1|-----
RAAC00062QWREKEATFVLITDIRDDQRVSDQILRLYKEQHEVEARFRYLKSPYHVGPYIYLHKPTRV

```

```
ref|ZP_01695687.1| EALGIVYVMALLIYGILEYRVRKELKEKNLSLILKGKRKLYQPTGQALLEQLEDITVILI
ref|ZP_01695982.1| EALGIVYVMALLIYGILEYRVRKELKEKNLSLILKGKRKLSQPTGQALLEQLEDITVILI
ref|ZP_01695971.1| EALGIVYVMALLIYGILEYRVRKELKEKNLSLILKGKRKLYQPTGQALLEQLEDITVILI
ref|ZP_01695655.1| EALGIVYVMALLIYGILEYRVRKELKEKNLSLILKGKRKLYQPTGQALLEQLEDITVILI
ref|YP_430569.1| -----
RAAC00062 KAFGFVMLLSLLLYSVLEYLIREKMKRETEPLMLPGNRKSFRTGLAILEMLDGVTTVHM
```

```
ref|ZP_01695687.1| NQNQQKIRLLPDNIDSQAKKIIELCGYDLSIY-----
ref|ZP_01695982.1| NQNQQKIRLLPDNIDSQAKKIIELCGYDLSIY-----
ref|ZP_01695971.1| NQNRQKLRLLPDNIDSQAKKIIELCGYDLSIY-----
ref|ZP_01695655.1| NQNRQKLRLLPDNIDSQAKKIIELCGYDLSIY-----
ref|YP_430569.1| -----
RAAC00062 QVGDTWQRVPATPHNPQIMRVLKLLNMDLSIYTEAQKTA
```

FIG. 92

```
ref|ZP_01665148.1| -----
RAAC02377 -----
ref|YP_754667.1| -----
ref|YP_753226.1| EFYPHYPTHETICALPRTEINSWLSYNTRPHMNASWLFEISUBSPWLFEISTRGETTING
ref|YP_752864.1| -----
ref|YP_001111903.1| -----

ref|ZP_01665148.1| -----
RAAC02377 -----
ref|YP_754667.1| -----
ref|YP_753226.1| ENGBABIHYPTHETICALPRTEINSWLSYNTRPHMNASWLFEISUBSPWLFEISTRGETT
ref|YP_752864.1| -----
ref|YP_001111903.1| -----

ref|ZP_01665148.1| -----
RAAC02377 -----
ref|YP_754667.1| -----
ref|YP_753226.1| INGENGABIHYPTHETICALPRTEINSWLSYNTRPHMNASWLFEISUBSPWLFEISTRG
ref|YP_752864.1| -----
ref|YP_001111903.1| -----

ref|ZP_01665148.1| -----
RAAC02377 -----
ref|YP_754667.1| -----
ref|YP_753226.1| ETTINGENGABIHYPTHETICALPRTEINSWLSYNTRPHMNASWLFEISUBSPWLFEIS
ref|YP_752864.1| -----
ref|YP_001111903.1| -----

ref|ZP_01665148.1| -----
RAAC02377 -----
ref|YP_754667.1| -----
ref|YP_753226.1| TRGETTINGENGABIHYPTHETICALPRTEINSWLSYNTRPHMNASWLFEISUBSPWLFE
ref|YP_752864.1| -----
ref|YP_001111903.1| -----

ref|ZP_01665148.1| -----EAXTRANSPSASEISFAMILYPRTEINTHERMSINUSCA
RAAC02377 -----
ref|YP_754667.1| -----
ref|YP_753226.1| EISTRGETTINGENGABIHYPTHETICALPRTEINSWLSYNTRPHMNASWLFEISUBSP
ref|YP_752864.1| -----
ref|YP_001111903.1| -----

ref|ZP_01665148.1| RBXYDIVRANSNRGBEAX-----TRANSPSASEISFAMILYPRTEINTHERMSINU
RAAC02377 -----
ref|YP_754667.1| -----
ref|YP_753226.1| WLFEISTRGETTINGENGABIHYPTHETICALPRTEINSWLSYNTRPHMNASWLFEISU
ref|YP_752864.1| -----
ref|YP_001111903.1| -----

ref|ZP_01665148.1| ----SCARBXYDIVRANSNRSCRESIGNIFICANCEE- IDENTITIESPSITIVESG-AP
RAAC02377 -----M
ref|YP_754667.1| -----
ref|YP_753226.1| BSPWLFEISTRGETTINGENSCRESIGNIFICANCEE- IDENTITIESPSITIVESGAPS
ref|YP_752864.1| -----
ref|YP_001111903.1| -----
```

```

ref|ZP_01665148.1| SMDMGYDAKD VYQAAY-ERQAQAI I PLNRRKEKL PPEGMDENRTPTCSMGYP MVYWG CER
RAAC02377 LMDAGYDQTKNYEAR-ALGAQAI I PMNRRNEKEPPEGMDFDGT PRCTMGYRMTYWG AEG
ref|YP_754667.1| LMDAGYDHREIYSVIRDKYHAQAI I ALNKRGA KQPPEGFDWDGT PICSARYRMVYWG S--
ref|YP_753226.1| LMDAGYDHREIYSVIRDKYHAQAI I ALNKRGA KQPPEGFDWDGT PICSARYRMVYWG S--
ref|YP_752864.1| LMDAGYDQKSIYELIRKDYKAQAI I PLNHRGA KEPEGLDWDAT PICSAGYRMAYWGG--
ref|YP_001111903.1| -MDMGYDSKDIYSVVMNDFNAQAI I PINSRGS KDHPGCD FDTPTCSMGQRMVFWGSDA
** *** . * . *****.:* * * *** * : ** *: *.:**

ref|ZP_01665148.1| EKGILKFRCPHVCGKVNCPNGSAWCS P S NYGLVIKKKVEDDPRS FCTPHRG TREWEKLYA
RAAC02377 D--WLKFRCPHATGQVDCPLGMAACSASNYGMVVKRLDEDVRRYANPHRGSR TWKMLYD
ref|YP_754667.1| YQGVNKFRCPHIMGKDCPFGSAWCSDSNYGMVVKTKVKDDPRLFSSPHRG SANWQKQYN
ref|YP_753226.1| YQGVNKFRCPHIMGKDCPFGSAWCSDSNYGMVVKTKVKDDPRLFSSPHRG SANWQKQYN
ref|YP_752864.1| SNGVNKFRCPHVMGKDCPFGSSWCSDSNYGMVVKTRARQDSRLFIVPHRGTSNWKLLYN
ref|YP_001111903.1| KAGTNKYRCPHVMGKDCPYGSAWCS P S SYGLVVKTKVKDDPRMNCI PARGTKNWQSLYN
*:**** *: : ** * : ** *.**:*:*.: : * * * ** : * : *

ref|ZP_01665148.1| ERTSVERAFSRLKEQLGANT-VRVQGIKKVTAHMLLCCIALLAGTIAVNR-----
RAAC02377 ERTAVERCFARLKEWLTLDG-VHVRGIEKVTAHAYINASVLLASALAMHRTNR IEQVA
ref|YP_754667.1| LRTYSERCFSRFKENLGLEDGLNVRKITKVETHAYLCAITMIAAVIAINQ-----
ref|YP_753226.1| LRTYSERCFSRFKENLGLEDGLNVRKITKVETHAYLCAITMIAAVIAINQ-----
ref|YP_752864.1| KRTSVERCFGRLEHLGLETGLNVRGIKKVKTHAYLSVITMIASVIAINKDKSSTDIA
ref|YP_001111903.1| KRTSVERCFGRLEHLGANS-IRTRGLEKVTLHITLSCIALLAGSIAVAKTKRIEQAA
** **.*.*:*: * : :...: ** * : :...:.*:*: :

```


FIG. 93

```

ref|YP_754865.1| ANCEE-IDENTITIESPSITIVESGAPSEEQYRLSRQKQFGTSSEKTTPEQIN-----
ref|YP_754943.1| -----EEQFRLSRQKQFGASSEKTTPEQIN-----
ref|YP_431166.1| -----MEEQIHLLRHRLFGTSSEKRRKPQTPDPSIQLP
RAAC03117 -----MOKKKFGTSSEKTERFEQGS-----
ref|ZP_02171171.1| -----EEQFRLSQAQKYGSSEKTPDQMS-----
ref|YP_001318017.1|

ref|YP_754865.1| LFNEAEDIVDPEIKEPDIE-----TITYERKKKQPGQKADKLKDL--PVEVI
ref|YP_754943.1| LFNEAEDITDPKLEEPSLE-----TVTYQRKKKQAGQREDKLKDL--PVEII
ref|YP_431166.1| -----
RAAC03117 LFNEAEVEADAQASEETVEGNAEATSEGVETETITYERRKPRAAREDAWLYQGEADEVV
ref|ZP_02171171.1| LFNESEKEQDAAEEPTVE-----AITYERKKKKRKARKD---LTENLYTETV
ref|YP_001318017.1| IFNEAEKFSVKLDEEPEAE-----EVLTKRRTGKSKSKKK---YEDLPIEEV

ref|YP_754865.1| EYRLLEHEQVCPCCQGSLSHEMSTQIRQEIKVIPAQVKVQHVQYIYSCRQCEKENITT--
ref|YP_754943.1| EYRLEKHEQICPCCQGELHEMSIQVRHEIKIIPAQAISVKHVQYIYACRRCEKENITT--
ref|YP_431166.1| -----MSTEVRELKIIIPAQVKVVKHIRYVYACRHCEREELTT--
RAAC03117 EYRLSDDERVCSKCAGELHEMSREITRRVKIIPAQMKEVYVRYVYTCRHCEAQDVET--
ref|ZP_02171171.1| TYTLFVEDQVCSNGELHIMKTQVKDELEIIPAEVKVKRYETIYSCRHCERTGTRN--
ref|YP_001318017.1| HYTLSDEERQCPKCDHTLHEMKTEVRKELKIIIPAQVKVHHIKQVYACRGDAIDSDNGG
* . : . : : * : . : . : * : * :

ref|YP_754865.1| PIIKAQMPNPILPGSLASPSILAYIMDQKYTNMPLYRQEQQLSRLGIELSRQTMANWVL
ref|YP_754943.1| PIIKAEMPKPILPGSLASPSILAYIMDQKYTNMPLYRQEQQFSRLGIELSRQTMANWLL
ref|YP_431166.1| PVVTAPMPAPVLPSPVSPSLLAYVMHQKYGEGLPLYRQEQQFKSLGLELSRQTLANWVL
RAAC03117 PVVRAPMPKPVQAKSLATPEAVAYVMAKKFVDGMPLYRQEQQFARHGYPILSRQTLANWVV
ref|ZP_02171171.1| PIVKAPSPERPFPGLSASPSIVSYMINQKFVQGVPLYRQEQEFKRMVPIISRQTMNSNWI
ref|YP_001318017.1| TIITAPMPKPVLPSPVSPVSLAFIMENKYNQALPLYRQEASFVNYGIDLSRQNMASWII
. : * * . * . : . : : : * : . : : : * : * : . : : * : * :

ref|YP_754865.1| NVADPWLKIIYDRLHVELLDRLDILHADETTLQVLKEPGRSAETKSYMWLRYTGRD-GPPI
ref|YP_754943.1| AAADPWLKIIYDRLHEQLLEKDILHADETTLQVLKEPGRRAESKSYMWLRYTGRD-GPPI
ref|YP_431166.1| HGANTWLTHIYDRLHEYLLKRDILHADETTLQVLREPGREAAATKSLFWLYRTGRD-GPSI
RAAC03117 HAAETWLEPLYAKLRQVLLAQRYLHADETTLQVLHEAGRAAQTSYMWVYRSSMN-GPPL
ref|ZP_02171171.1| EASEQMLEPIWDLIRILTSLDVLHADETTVQVLKEDGKEAAAKSYMWLRYSGSH-DVPI
ref|YP_001318017.1| QGAEKWLSPLYDRMHTLHKQSPVIAHDESPLKVLDEKDK--SQSYMWLRYATAETSEYPI
. : * . : . : * . : : : : * : . : : : * : * : . :

ref|YP_754865.1| VLYEYQTTTRASKHPDRFLSGFKGYLQTDGYSAYGKLT-GITLVGCWAHARRKFTEALKAL
ref|YP_754943.1| VLYDYQTTTRASKHPDSFLSGFKGYLQTDGYSYGYSLT-SVTLGACWAHARRKFTEALKAL
ref|YP_431166.1| VLYDYQTTTRASKHPCRFLAGFKGYLHVDGYAGYNELP-DVTLVGCWAHARRKFDEALKAL
RAAC03117 VLYDYQETRSAEHPRRFLAGFQGYLHVDGYAGYEGLP-DVTLVGCWAHARRKFDEALKAV
ref|ZP_02171171.1| VIYDYQPGRASKYPRRFLGFTGYLHVDGYGGYHALPKPKVELVGCWAHARRKFFDAVQTL
ref|YP_001318017.1| YLYEYQPSRAKKHPQFLEGFTGFLQTDGYAGYNGVE-NVVQVGC LAHARRKYTDAIKAL
: * : * : : * * * * : * : : * : * : : * : :

ref|YP_754865.1| PAAQKDKPVAASVGLCYNRLFATERQLKD--VSDKERYDKRLEKSKPLLDEFYIWLKKQ
ref|YP_754943.1| PAEQKDKPVAASIGLYCNKLFATERQLKD--MSCQERYEKRLTRPLLNEFYAWLKRQ
ref|YP_431166.1| PEDKRNAAVAAREGLEFCNRLFATIERDLKD--KTPEERYQLRQVRSKPVLD A FLAWLKTQ
RAAC03117 PSKERKKGKTAEEGLSYCNALYAVEKKLKN--ASAEERQVRMAKSKPILDAFLAWLEKQ
ref|ZP_02171171.1| PDDRSTTSAAKGLNAIDELYRIEREIQNEYKTPPEEFYEARKERIEPPLEAFSAWVESM
ref|YP_001318017.1| PEGSDVSLTKANEGLSLLRKIYRLEKSFKE--MEPEVRYEARIEQTQPVLDAYKTWLEVE
* * * * : : : * : * : * :

ref|YP_754865.1| KQQTLPKSTFGQAITYCLNQWDCLNSFLLDGRLEIDNNRAERSIKPFVIGRKNWLFNTNP
ref|YP_754943.1| RQQTLPKSMFGQAITYCLNQWDALNTFLLDGRLEIDNNRAERSIKPFVIGRKNWLFNTNP
ref|YP_431166.1| KSRVLPKSSFGQAINYCLGQWDKLTAFQLDGRLELDNNRSERSIKPFVIGRKNWLFNTNP
RAAC03117 EQQVLPKSALGRAVSYVLKQWPKLIRYVENGYLEIDNNRCERSLKPFFVIGRKNWLFNTNP
ref|ZP_02171171.1| KPILSKSLLGKAVIYASNQMEHLRTFLKDGRLAIDNNRAERSIKPFVIGRKNWIFNTNP
ref|YP_001318017.1| EKRTLPKSKLQGAISYSLKQWDKLA AFMKDGRIADNNLAERGKIPFVLGRKNYLFKSP
. : * : * : * : * : : : * : * : * : * : * : * : * : * :

```

| | |
|--------------------|--|
| ref YP_754865.1 | RGARGSAIIYSVIETAKENNLKPYNYMFYLFQLPNVDITGDQAAIDRLLPWSDTLPE--- |
| ref YP_754943.1 | RGARGSAIIYSIIETAKENNLKPYDYLVYLFQLPNVDITSDQTAVDRIMPWSDTLPEGCR |
| ref YP_431166.1 | RGAKASAITYSIIETAKDNGLNPFQYLIYLFERLPNLDLKDKDALDQLLPWSASLP---- |
| RAAC03117 | RGARASAVTYSIVETAKENGLNPTAYLTYLFERMPNIDLKDEAAFEALLPWSEGLPEGIR |
| ref ZP_02171171.1 | RGAKSSSIIYSMIETAKENQLKPQAYLNYLFENLP---SSKQSEMEQFLPWSDSLPRMIF |
| ref YP_001318017.1 | KGATASALCYSIETAKANKLIPFQYLTYLFEQLPNLDIEDPEALDAMPLPWAESLPNEVR |
| | :** .*: : **: :***** * * * *: *****.:* . .: :*: ** |
| ref YP_754865.1 | --- |
| ref YP_754943.1 | I-- |
| ref YP_431166.1 | --- |
| RAAC03117 | VRK |
| ref ZP_02171171.1 | V-- |
| ref YP_001318017.1 | --- |

FIG. 94

```

ref|NP_780819.1|      ----IQIFQNYSVFIVLGLIIATIILFFIVVVQAKAINRLEKRYRKFMRGVDNKNLEELI
ref|YP_699899.1|      ----LNILTQYSTYITIGLIVLVLIQFILLIVALRSLSKVENKFRKIMRGVNNKNLEELI
ref|YP_001514390.1|  MAFITNILEEYRDILFIASLGLNLTIIFLIINMGISSNLKEKYRKLVRGTDGKQIESIL
ref|YP_077153.1|      -----LLQODPLLVAFTALGFCIVAIVIMLIVLVRQSILLRRYSLLRGNTNASLEDLL
RAAC00037             MAMGLSILQPYALDIALFSGILAIICLVIASVALSRARLKRKFNRLKEVTSAADLERVF
ref|YP_001663996.1|  ----LDIISQATLIILFLSVLSIIEIIFILIINGKFLRLNRTYNKI IKTLEKGDVFDIF
                        ::      : :      :: :: :      : . . . .      .: ::

ref|NP_780819.1|      NTYLDKVDKASEECQYAKELYKSLEDRLNICVQKIAIIRYRAFEDVGSDSLFSFVALLDYK
ref|YP_699899.1|      NSYLDKVEEVKDDSEETLETNKMLKAQIEKCTQKVS VIRYKAFEDVGSDSLFSFVALLDGE
ref|YP_001514390.1|  FEHLDKIEDVHQRLNQFEGKLEIFNNRLSFCVQVRVGIIRYNAFDDTGSDLSYSIALLDEN
ref|YP_077153.1|      IQQQQATADLRAAQESIRRRSLDESASQKYLQRIGIVRYNAFPDVGADLSFSCALLDGE
RAAC00037             EETKDAVRKLEMKLREAEHLRIVEEALQSKVSTPAILRYNAFAEVLGSDLSYSVALIDGK
ref|YP_001663996.1|  SRILTENEEIKNKLDKLRMDLNSLDKETKTAIKKVGIVRYNAFSDVGSDSLFSFVALLDSN
                        .          ..      .      .      .:*.** :*:*:* *:* * :

ref|NP_780819.1|      DSGIIITGIYGRNESTTYAKPIDKGISRYELSEEE----NHVLKEAM-----
ref|YP_699899.1|      NNGVILTGIYGRDYSTTYAKPIDKGISRYDLSEEE----LHVLNAAM-----
ref|YP_001514390.1|  NDGIITGIYGRIETVSYAKPVKNGVSNYSLSVEE----LQALERA-----
ref|YP_077153.1|      DNGVVVTSLYGRSECRTYAKPIRGSSSSYALTDEE----KQALRLA-----
RAAC00037             GDGVVITSIYGREDSVTYGKPVQGGDSPYMLTEERAVIEEALRGAPRRRTTAQIS
ref|YP_001663996.1|  DNGIVLSGIYGRNETATFAKPIERGQSKYPLSAEE----VQAIERAKRKA-----
                        ..*:::***      ::** : * * * * : *      .... *

```

FIG. 95

```
ref|NP_387885.1|
ref|YP_001419683.1|
ref|YP_077286.1|
ref|ZP_02326643.1|
ref|YP_803557.1|
RAAC00054
```

[illegible]

```
ref|NP_387885.1|
ref|YP_001419683.1|
ref|YP_077286.1|
ref|ZP_02326643.1|
ref|YP_803557.1|
RAAC00054
```

```

IRWDKDYAKIEGRVMKQNGAIPMQLVLSKKGKKGVNHIEQQKLSQYVGALNTIMFAPED
IRWDEDYAKIEGRVMKRNGDIPMQLVLSKKGKKGVNHIEQQKLSQYVGALNTIMFAPED
IRWDEDYAKIEGRVICKNGSVPIQLVLSKKGKKGVNHIEQQKLSQYVGAVNTIMFAPED
IQWEGESALLQGDVEKKYGSYSLDLAISSKGGKAKINGLEQKKLSQFICALNVVLFAPED
ISWKTKAARVSGSVQKEHTVTPLEINLSSKGKNAKVNHLEQSRLSQYVGQLNVILFAPED
IRWEQDRARILLEASTRYGDRRLTLELGPERRAFANGVQVGRMTEFVGQVQVVLAPED
*      *      *      *      *      *      *      *      *      *

```

```
ref|NP_387885.1|
ref|YP_001419683.1|
ref|YP_077286.1|
ref|ZP_02326643.1|
ref|YP_803557.1|
RAAC00054
```

LNLVKGSPQVRRRFLDMEIGQVSPVYLHDLISLYQKILSQRNHFLKQLQ-TRKQTDRTMLD
LNLVKGSPQVRRRFLDMEIGQVSAVYLYDLISLYQKILSQRNHFLKQLQ-SRKQTDRTMLD
LNLVKGSPQVRRRFLDMEIGQVSPVYLHDLISLYQKILSQRNHFLKQLQ-TRKQTDQTMMLD
LEIIKGNPGIRRRFLDMEIGQVYPGYLYDLISQYQKVLQQRNMMKKAFAPSAAHAAML
LSIVKGSPAVRRKFIDMEFGQMSSKYLNSAQYRSVLKQRNQYIKQLQFNPKG-DQVYLD
LDLVKGSPRVRRRFLDTELQMEPLYLHHLISLYNRALLQRNRWLKTAP---LSPDDDVLA
* : * * * * : * * * * : * * : * : * * * * : * * * * : *

```
ref|NP_387885.1|
ref|YP_001419683.1|
ref|YP_077286.1|
ref|ZP_02326643.1|
ref|YP_803557.1|
RAAC00054
```

VLTDLQVEVAAKVVVKRLQFTAQLEKWAQPIHAGISRGLEELTLKYHTALD-VSDPLDLS
VLTDLQIEAAKVVAKRLQFTAQLEKWAQPIHSGISRGLEELTLKYHTALD-VSDPKDLS
VLTEQLTEFAAKVVMKRLQFVDQLEKWAQPIHSGISRGLEELTLKYHTSLH-VSDSPDLS
IWNEQLAQFGVKIMKKRQNFIIKKLQNWAEQIHDGITNGGEELTI RYQPSFA-VQDFEDET
VLSDQLAAHGAEEIFQRIQFLKKLEKWSQEVHKEISQGEKLSFQYVSPIS-SDQADTTE
TFDRQIAFHGAHVHRLRLFLARLAYAARIYSDIASGREEFALAYRSSVSGVEEGMSVE
* : * * * * * : * * * * * : * * * * * : *

```
ref|NP_387885.1|
ref|YP_001419683.1|
ref|YP_077286.1|
ref|ZP_02326643.1|
ref|YP_803557.1|
RAAC00054
```

KIGDSYQEAFSKLREKEIERGVTLSGPHRDDVLFFVYNGRDVQTYGSQGQORTTALSCLKA
KIGNSYQESFSKLKEKEIERGVTLFGPHRDDVLFFVYNGRDVQTYGSQGQORTTALSCLKA
KMINSYQETFSKLKDKEIERGVSLSGPHRDDVLFFVYNGRDVQTYGSQGQORTTALSCLKA
VLMEQFMIKLSQIKDQEIRRGVSLAGPHRDDLLFYINDKEVQTYGSQGQORTTALSCLKA
KIYAALQALFQKQREKELQQGKTLVGPHLDDVRFMVNDKNVSTFGSQGQORTTALSCLKA
EMADTVQRALEKNRAQDLRFGTTSAGPHRDDILLFLDGREVHTAASQGQORTTALSCLKA
: : : : : * : : : : : * : : : : : * : : : : : *

```
ref|NP_387885.1|
ref|YP_001419683.1|
ref|YP_077286.1|
ref|ZP_02326643.1|
ref|YP_803557.1|
RAAC00054
```

[illegible]

```
ref|NP_387885.1|
ref|YP_001419683.1|
ref|YP_077286.1|
ref|ZP_02326643.1|
ref|YP_803557.1|
RAAC00054
```

GMFRVQNGALVK---
GMFRVENGTLVK---
EIFRVENGTLSD---
SVFRV-----
HVFNIDHGVLMQSK-
RLFRVCSGIIAEERV
.*.*

FIG. 96

```

gb|AAR99616.1|-----IKEMYERGMSISDIA
sp|Q45618|TRA6_BACST-----IKEMYERGMSISDIA
ref|YP_074959.1|-----MLRSGETLEIRQMYAGGLSISEIA
ref|YP_075129.1|-----MYAAGLSISEIA
RAAC03102MRIPODHRPTFLKIIIGMNSPPPCRIVAPVIRSWEVFVMREDERMEIRQLYEAGVSISELA
ref|YP_359963.1|-----LHELQARGKSIRAIA
: * * * :*

gb|AAR99616.1|RELGIDRKTVRKYIHSPNPPSKSKR-----KQRKSKLDPFKPYLQKRMLLEDGVFNSEKLF
sp|Q45618|TRA6_BACSTRELGIDRKTVRKYIHSPNPPSKSKR-----KQRKSKLDPFKPYLQKRMLLEDGVFNSEKLF
ref|YP_074959.1|RRTGRDRKTIRKWLRTNTMPKPAK-----RKRSSMLDQHEAFTLEQMOK-GVTSASKML
ref|YP_075129.1|RRTGRDRKTIRKWLRTNAMPKPAK-----RKRSSMLDQHKAFILEQMOK-GVTNASKML
RAAC03102RRFGYDRKTIRNALNSSVEEKQGERASRGERKRGSKLEPYKDYVKQRMQL-GVLNAERIL
ref|YP_359963.1|RETGHSRNTVRKYLRAGIPEKPR-----PKRGSKLDPYKDTIQELMNL-GIFNCEVIY
* . * .*:*:*: ::: . :* * *: .: : * *: ... :

gb|AAR99616.1|FEIRQQGYTGGKTILKDYMKPFRETAKKKYTVRYETLPGEQMQVD-----
sp|Q45618|TRA6_BACSTFEIRQQGYTGGKTILKDYMKPFRETAKKKYTVRYETLPGEQMQVD-----
ref|YP_074959.1|YLLQQRGFGKGKIRIVRAFMAPYRPMAKAAATVRFETPPGKQAQVDWADFGYIEVDGRR--
ref|YP_075129.1|YLLQQRGFGKGKIRIVRAFMAPYRPMAKAAATVRFETPPGKQAQVDWADFGYIEVDGRR--
RAAC03102REIREQGYTGGITVLRREFMKPLRPVVSAKATERYESDPGEQAQIDLGAFPPYDSDHGQRRRT
ref|YP_359963.1|ERIKEEGYTGGRTILRDYVRQFRPPKQVPAVCRYETKPGQQAQVDWGEYTYID-----
:::.*.* ::: :: * . .*:*: **:**:*

gb|AAR99616.1|-
sp|Q45618|TRA6_BACST-
ref|YP_074959.1|-
ref|YP_075129.1|-
RAAC03102I
ref|YP_359963.1|-

```

FIG. 97

```
ref|YP_074959.1|      MVLAYSRA MYLEFVTATDMKTFMRCHINAFKFFGGVPHEILYDNVKT VVKDRD-----
ref|YP_594046.1|      MVLGWSRALYVEFIRKADTASFIRCHLNAFAYFGGMTQSILYDNTKQVVLERDETG----
ref|ZP_02516401.1|    -VLGYSRMLYIEFTDNMRYDTLETCHRNAFRFFGGVPREVLYDNMKT VVLQRD-----
ref|ZP_02335796.1|    -VLGYSRMLYIEFTDNMRYDTLETCHRNAFRFFGGVPREVLYDNMKT VVLQRD-----
ref|ZP_02563000.1|    -VLGYSRMLYIEFTDNMRYDTLETCHRNAFRFFGGVPREVLYDNMKT VVLQRD-----
RAAC03103             MVLAYS RMLYVEFIKAADQLHILQALRNALEFFGGVPRVMLS DNC SPLVVANDGQGHVDC
                        **.:** :*:**      : . **: :***.: : * ** . : * . *

ref|YP_074959.1|      -----
ref|YP_594046.1|      -----
ref|ZP_02516401.1|    -----
ref|ZP_02335796.1|    -----
ref|ZP_02563000.1|    -----
RAAC03103             NRLISILPSSTDSCPRHVGLAGAAPRAR
```

FIG. 98

```
ref|YP_001039349.1|
ref|YP_001036724.1|
ref|YP_076075.1|
ref|YP_074958.1|
ref|YP_076118.1|
RAAC03341
-----
TRANSPSASEMUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPSASE
-----

ref|YP_001039349.1|
ref|YP_001036724.1|
ref|YP_076075.1|
ref|YP_074958.1|
ref|YP_076118.1|
RAAC03341
-----
MUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPSASEMUTATRTYPE
-----

ref|YP_001039349.1|
ref|YP_001036724.1|
ref|YP_076075.1|
ref|YP_074958.1|
ref|YP_076118.1|
RAAC03341
-----
CLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPSASEMUTATRTYPECLSTRIDIUM
-----

ref|YP_001039349.1|
ref|YP_001036724.1|
ref|YP_076075.1|
ref|YP_074958.1|
ref|YP_076118.1|
RAAC03341
-----
THERMCELLUMATCCGBABNTRANSPSASEMUTATRTYPECLSTRIDIUMTHERMCELLU
-----

ref|YP_001039349.1|
ref|YP_001036724.1|
ref|YP_076075.1|
ref|YP_074958.1|
ref|YP_076118.1|
RAAC03341
-----
MATCCGBABNTRANSPSASEMUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABN
-----

ref|YP_001039349.1|
ref|YP_001036724.1|
ref|YP_076075.1|
ref|YP_074958.1|
ref|YP_076118.1|
RAAC03341
-----
TRANSPSASEMUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPSASE
-----

ref|YP_001039349.1|
ref|YP_001036724.1|
ref|YP_076075.1|
ref|YP_074958.1|
ref|YP_076118.1|
RAAC03341
-----
MUTATRTYPECLSTRIDIUMTHERMCELLUMATCCSCRESIGNIFICANCEE-IDENTIT
-----

ref|YP_001039349.1|
ref|YP_001036724.1|
ref|YP_076075.1|
ref|YP_074958.1|
ref|YP_076118.1|
RAAC03341
-----
-----MATNNRMALLEQLSKYVVEKDKDFLKEALTLLINALMDAEVTSII
IESPSITIVESGAPSMATNNRMALLEQLSKYVVEKDKDFLKEALTLLINALMDAEVTSII
-----SRRRIPOVTTFRIALEEELLRK TGVD-DFDFLREGLRVLAQGLMELEVSQRI
-----LEEELLRK TGVD-DFDFLREGLRVLAQGLMELEVSQRI
-----SNTSQRGNPQVTNFRIALEEELLRKSGVD-DVDFLREGVRVLAQGLMELEVSQQI
-----MTNTKSRRHATMASLNSFAVLEWIRKMQDQIDFLRELMQLVLTQFLIDAEAAEKI
: * : * : ****: : : *:: *:: *
:

ref|YP_001039349.1|
ref|YP_001036724.1|
ref|YP_076075.1|
ref|YP_074958.1|
ref|YP_076118.1|
RAAC03341
GAEKYERNNNRNNYRNGYRLREWDTRVGTQLSLIPKLRHGSYFPSLLEPRK
GAEKYERNNNRNNYRNGYRLREWDTRVGTQLSLIPKLRHGSYFPSLLEPRK
GADRYERSAERSTYRNGYRERQWDTRVGTIDLQIPKLRKGSYMP SFLEPRR
GADRYERSAERSTYRNGYRERQWDTRVGTIDLQIPKLRKGSYMP SFLEPRR
GAERYERSSERSNYRNGYRERQWDTRVGTIDLQIPKLRKGSYMP SWLEPRR
GAERYERTESRVTORNGYRSRAWDTRLGTVDLKI PKLRQGSFFPSI LEPRR
**::***. * . ***** * ****:***:*.*****:***:***:***:

```

```
ref|YP_001125185.1|MSDYLIVIVESPTKAKTIERYLGKKYTVKASMGHVRDLPKSQMG  
ref|YP_147064.1|-----MSDYLIVIVESPTKAKTIERYLGKKYTVKASMGHVRDLPKSQMG  
ref|NP_389494.1|-----MSDYLIVIVESPAKAKTIERYLGKKYVKV KASMGHVRDLPKSQMG  
ref|YP_091420.1|-----MSEYLVIVESPAKAKTIERYLGKKYVKV KASMGHVRDLPKSQMG  
ref|YP_001486753.1|MADYLVIVESPAKAKTIERYLGKKYVKV KASMGHIRDLPKSQLG  
RAAC02852MNQGGQLPRLRRMEDRGVS DYLVIVESPAKAKTI GKYLGSKYT VKASMGHVRDLPKSQLG  
::*****:***** :***.*.*****:*****:*  
  
ref|YP_001125185.1|VDINDGYTPKYITIRGKGQVIKELKTAACKAKKVFLAADPDREGEIAIWHLANMLDLDIH  
ref|YP_147064.1|VDIDHGYPEKYITIRGKGQVIKELKTAACKAKKVFLAADPDREGEIAIWHLAHMLDLDIH  
ref|NP_389494.1|VDIEQNFEFKYITIRGKGPVLKELKTAACKAKKVYLAADPDREGEIAIWLHLSDLIDLN  
ref|YP_091420.1|VDIEQNFEFKYITIRGKGPVLKELKTAACKAKKVYLAADPDREGEIAIWHLAHSLDLIDLN  
ref|YP_001486753.1|VDTEHDFEPYITIRGKGPVLKELKTAACKAKKVYLAADPDREGEIAIWHLAHSLDLIDLN  
RAAC02852VDVEHGFEPKYITIRGKGDV IKALREASKAKKVYLAADPDREGEIAIWHLQHLLDLNPD  
** :.: * :***** * : * : * :*****:*****:***** : ***:  
  
ref|YP_001125185.1|SDCRVVFHEITKDAIKQS FQQPRSINMNLVDAQQARRVLDRLVGYNISPLLWKKVKKGLS  
ref|YP_147064.1|SDCRVVFHEITKDAIQQS FQQPRAINMNLVDAQQARRVLDRLVGYNISPLLWKKVKKGLS  
ref|NP_389494.1|SDCRVVFNEITKDAIKESFKHPRMINMDLVDAQQARRILDRLVGYKISPILWKKVKKGLS  
ref|YP_091420.1|SDCRVVFNEITKDAIKESFKHPRMINMDLVDAQQARRILDRLVGYKISPILWKKVKKGLS  
ref|YP_001486753.1|SDCRVVFNEITKDAIKDSFKHPRMINMDLVDAQQARRILDRLVGYKISPILWKKVKKGLS  
RAAC02852DDCRVVFHEITKDAVKQAQFQHPRKINLDLVNAQQTRRILDRLVGYKLS PLLWRVKVKKGLS  
.*****:*****:.:*:** *:**:*****:*****:***:***:*****  
  
ref|YP_001125185.1|AGRVQSVALRLIIDREKEIRQFQPEEYWTIQAKFQKGKEAFAASFYGVDGEKLELKTEAD  
ref|YP_147064.1|AGRVQSVALRLIIDREIREIFQPEEYWTIQATFQKEGETFAASFYSIDGQKRDLKTEAD  
ref|NP_389494.1|AGRVQSVALRLIIDREKEINDFKPEEYWTIDGTFLKGQETFEASFFGKNGKKLP LNSEAD  
ref|YP_091420.1|AGRVQSVALRLIIDREKEINDFIPEEYWTIEGSFLKGKETFEASFFGINGEKHQ LKSEDD  
ref|YP_001486753.1|AGRVQSVALRLIIDRENEINEFKPEEYWTIDGTFLKGKESFEASFFGVNGKKHPLKTED  
RAAC02852AGRVQSVALRLI VDRENI RTFQPEEYWT VDAHGOVHGKKLVARFYGYGEEKT PLPNEAA  
*****:*****:***.*. * *****:.. : * *. . : * * .:  
  
ref|YP_001125185.1|VKEVLARLNGRAFTVTKVT KRERRRSVPVPFFTSSLQQEAARKLNFR TKTMMAVQQLYE  
ref|YP_147064.1|VKAVLDRLNGTAFVAVTKVT KRERKRSVPVPFFTSSLQQEAARKLNFR TKTMMAIQQLYE  
ref|NP_389494.1|VKEILS QLKNQYTVEKVTKKERKRN PALFFTSTLQQEAARKLNFR AKTTMMIAQQLYE  
ref|YP_091420.1|VKQILSR IKG NKFEVKKVTKKERKRN PALFFTSTLQQEAARKLNFR AKTTMMIAQQLYE  
ref|YP_001486753.1|VKEILSKLKGSKFSVEKVTKKERKRN PAVFFTSTLQQEAARKLNFR AKTTMMIAQQLYE  
RAAC02852VRELLSVEGHS LVVRVKSERKRN PAFFTSTSSLQQEAARKLNFRAKTYMTMQIAQQLYE  
*: * :.: * * .* * **:*.*. *****:*****:** : *** :*****  
  
ref|YP_001125185.1|GIDLGEGTVGLITYMR TDSTR IAETAQQEEAAAY IEATFGAQYVHQEK RKEKKSTNAQDA  
ref|YP_147064.1|GIDLGSEGT VGLITYMR TDSTRVAETAQQEEAAAY IEATFGAMYVNQEKR EKKSTNAQDA  
ref|NP_389494.1|GIDLGREGTVGLITYMR TDSTRISNTAVDEAAAF IDQTYGKEFLGGKRKP AKKNENAQDA  
ref|YP_091420.1|GIDLGKEGT VGLITYMR TDSTRISNTAQEEAAS FIDQTYGKEFLGSKRP AKKNENAQDA  
ref|YP_001486753.1|GIDLGKEGT VGLITYMR TDSTRISNTAE EVSAF IDQTYGKNFLNTTKRTVKKNENAQDA  
RAAC02852GLDVPEGGT VGLITYMR TDSTR IAQSAQEEARAY IRQSFGEDYVPDR PRQYAKNEDAQDA  
*.: * :*****:*****:.:* :*. :.* : * : : * . :****  
  
ref|YP_001125185.1|HEAIRPTS AF RDPDKVKPYLTRDQFR LYKL IWERFIASQMAAAVLD TMSVELENNGLVFR  
ref|YP_147064.1|HEAIRPTS AF RDPDKVKPYLTRDQFR LYKL IWERFIASQMAAAVLD TMSVELENNGVFR  
ref|NP_389494.1|HEAIRPTS VL RKPS ELKAVLGRD QRL YKL IWERFVASQMAPAVLD TMSVDLTNNGLTFR  
ref|YP_091420.1|HEAIRPTS VL RKPS DLKAVLGRD QRL YKL IWERFVASQMAPATLD TMSVDLENGLTFR  
ref|YP_001486753.1|HEAIRPTS TLRK PADVKHVLSRD QRL YKL IWERFVASQMAPAVLD TMSVDLDNNGLTFR  
RAAC02852HEAIRPTS VMRH PDRLKD HLSRD QRYLY KL IWERFVASQMESAVLD TT SVDLEANGAWFR  
*****:.*.* : * * *** *****:***** :*.*** **: * ** **  
  
ref|YP_001125185.1|ASGSKVKFP GF MKVYIEGTD DQT EEQD-RL LPDLEE GEAVESEMIE PKQH FTQPPPRYTE  
ref|YP_147064.1|ASGSKVKFP GF MKVYIEGTD DQT EEQD-R ILPDLEEE ETVESETIESKQH FTQPPPRYTE  
ref|NP_389494.1|ANGSKVKFSG FMKVYVEG KDDQMEEKD-RMLPDLQEGD TVLSKD IEPEQH FTQPPPRYTE  
ref|YP_091420.1|ANGSKVKFSG FMKVYVEG KDDQMEEKN-KMLPDLAEGD TVLSKD IEPEQH FTQPPPRYTE  
ref|YP_001486753.1|ANGSKVKFP GF MKVYVEG KDDQLEEKD-KMLPDLKEGD TVLSKD IEPEQH FTQPPPRYTE  
RAAC02852ATGSVVRF PGFMALYTEGRDDD ADEEGKLLPPLQEGD VVQVKTWKPEQH FTQPPPRYTE  
* * * * * * * * * * * * * * * * * * * * * * * * * * * * *
```


| | |
|--------------------|---|
| ref YP_001125185.1 | ARLVKLTLEELGIGRPSTYAPTLDTIQKRNYYVLENKRFVPTTELGEIVVELILEFFPEIID |
| ref YP_147064.1 | ARLVKLTLEELGIGRPSTYAPTLDTIQKRNYYVLENKRFVPTTELGEIVVELMLEFFPEIID |
| ref NP_389494.1 | ARLVKLTLEERGIGRPSTYAPTLDTIQRRGYVALDNKRFVPTTELGGQIVLDLIMEFFPEIIN |
| ref YP_091420.1 | ARLVKLTLEELGIGRPSTYAPTLDTIQKRGYVALDNKRFIPTTELGEIVLDLIMEFFPEIIN |
| ref YP_001486753.1 | ARLVKLTLEELGIGRPSTYAPTLDTIQKRGYVALDNKRFIPTTELGEIVLNLIIEFFPEIIN |
| RAAC02852 | STLVKAMEELGIGRPSTYAPTIDILLKRGYVTLQKRFVPTTELGEIVVNILKEHFPQLID
: ***.:** *****:* : :*.**.*:***:*****:***::: *.**.:*: |
| | |
| ref YP_001125185.1 | VEFTAKMEKELDEIEEGKVEWIKVVDEFYREFEKLKVAEKEMREVEIKDEPAGIDCDVC |
| ref YP_147064.1 | VEFTAKMEKELDEIEEGKVEWIKVVDEFYREFEKLKVAEKEMRAVEIKDEPAGIDCEVC |
| ref NP_389494.1 | VEFTAKMERDLHDHVEEGNTEWVKIIDNFYTDFEKRVKKAEMKEVEIEPEYAGEDCELC |
| ref YP_091420.1 | VEFTAKMEKELDDVEDGNIQWVQIIDSFYKDFEKRVEKAEAEQMVEIEPEYAGVDCEAC |
| ref YP_001486753.1 | VEFTAKMEKELDSVEEGTIEWVRIIDSFYQDFAKRVEKAEAEQMVEIEPEYAGVDCEEC |
| RAAC02852 | VSFTADMESRLDKVEEGNANWIELLDQFYHDFEKDLKKAESALGHVELKDEVSDVRCEKC
*.***.** ** :*:.* :*:.:*.** :* * : : ** : **.: * :. * : * |
| | |
| ref YP_001125185.1 | GSPMVYKMGRFGKFIACSNFPECRHTKPIVKEIGVKCPKCHEGNIVERNSKRKRIFYGCD |
| ref YP_147064.1 | GSPMVYKMGRFGKFIACSNFPECRHTKPIVKEIGVKCPKCREGNIVERSTKRKRIFYGCD |
| ref NP_389494.1 | SSPMVYKMGRYGFKLACSNFPDCRNTKPIVKQIGVKCPSCGEGNIVERKSKKKRIFYGCD |
| ref YP_091420.1 | GHPMVYKMGRYGFMACSNFPDCRNTKPIVKEIGVKCPSCKTGNIVERKSKKKRIFYGCD |
| ref YP_001486753.1 | GHPMVYKMGRYGFMACSNFPDCRNTKPIVKDIGVKCPTCHEGNIVERKSKKKRIFYGCD |
| RAAC02852 | GRLMVYKTGRYGFKLACPGFPECRNTKPIVKEIPVSCPKCGKP-LVERKKGSRKRVFYGCS
. **** **.:***:*.**.:**.:***:***:*.** *.*.* :***. * :.:****. |
| | |
| ref YP_001125185.1 | RFPECDFVSWDKPLARPCPKCGLLVEKKLKKGVQVQCT----- |
| ref YP_147064.1 | RFPDCDFVSWDKPLARPCPKCAGLLVEKVLKKGVQVQCT----- |
| ref NP_389494.1 | RYPDCEFVSWDKPIERKCPKCGKMLVEKKLKKGIQVQC----- |
| ref YP_091420.1 | RYPECEFVSWDKPLERKCPKCEDMLVEKKLKKGVQVQCVN----- |
| ref YP_001486753.1 | RFPECEFVSWDKPIERKCPKCNMLVEKKLKKGMQVQCVN----- |
| RAAC02852 | GYPECDYVLWQRPTGQTCPVCGHPMIEKGGKGKTVVVCSNEKAHPMVATEAQAK
:*.***:**:.* : ** * : : ** * * * |

```
ref|YP_001125186.1|-----ALQLFLEYLQIEKNYSQYITVICRYRDIEQFLQFMNEEGIDEL-NEVAYSADSVRLYL  
ref|NP_243331.1|-----SPHTIVNYELDLRHRFRDFMEQQSIPSF-AAVSYA FVRHYL  
ref|YP_175772.1|-----TIDHYRKDVQQFAAFMSAAGIAEI-KDVKHQDVRLF  
ref|ZP_01775043.1|-----SPHTLAAYRS DLEQFLSFLAREREAPRAEEVDHLAIRRYL  
ref|YP_740910.1|-----SPHTLQAYQRDLTRFHWHWCAENGLADR-DAVSAHDIRRF  
RAAC02854MTQSVGD AVKVFLDDAALRFSPTVRSYGQDLEAFRWQLDDRGVYDL-DALSTRDVRMHA  
*: *: *: *: : : : *  
  
ref|YP_001125186.1|TKLYGQQQLASRS VARKISSLR SFYKFLREGWTAENPFALAALPKKEQKIPNF LYREEE  
ref|NP_243331.1|TVLYEQEYARSTVSRKLSTLR SFYQFLVREKWVMENPFLLAHTPGKVKKLP SFLYEEME  
ref|YP_175772.1|SELVDKKYARKSIARKLSALRSFGKF LMEEGHISENPF LHTHPKQKTRLPTFLYEEME  
ref|ZP_01775043.1|AQLH-KGCAKSSIGR KLSAIRALFRYLMREGKLEKNPAELVSTPPKKEKR LPFHLNIDQVS  
ref|YP_740910.1|AARHRQGLAPGSVQRTLSSLRS LF RYLVREGRLTGNPAEGVAAPRRPRRLPGVLSPDEAA  
RAAC02854SDLLAKGA AKSSVARRLSLCIRTFLRFC AERG WVRQVMAKNVRLPKDRRLPRYLHEEEVA  
: : * ::*: *: *: : .. : *: **: * : :  
  
ref|YP_001125186.1|ALFRVNDGNTAVGQRNAALLELLYATGARVSECCHIRLSDLDFAASTVL IHGKGNKQRYV  
ref|NP_243331.1|QLLDALNGDSPLQLRNRAFETIYASGLRVSECCGLKLQDVDSLIGTFVFVGKGRKERYV  
ref|YP_175772.1|HWLLALPANKPLEKRDKAIELLYATGM RVSECSMLALDQWD RVSETVRVFGKGRKERYV  
ref|ZP_01775043.1|ALVTAPAGSSGLPLRDRAVLETLYSCGIRVSELTGMNVGDMDLAAGLARVMGKG GKERLV  
ref|YP_740910.1|RLLEGSPEDDPLALRDRALYELIYSSGLRLAEAVGLDLGR LDLTEGLVEVVGKGAKTRRV  
RAAC02854ALIDHVGGDDFVALRDRALLEFLYATGVRVSECVHLDIGDLDSL AGFARVLGKGGRERYV  
. . : *: *: *: *: *: *: : : * . : *** : * *  
  
ref|YP_001125186.1|PFGRPAREALERYIGGGRREL VGKLPADHR--YLFVNARGNPLTPRGVRYILDRIVETA  
ref|NP_243331.1|PIGSFACDAIQEYIENGREKLLKKS KSVDP L PGDLFLN YRGGPLTERGV RKILHQALDQAA  
ref|YP_175772.1|PVGKMVASAVESYIHEARP KLLR--ANDPTS HLFNLNRGGALS DR SIRKIVEKRLDEAA  
ref|ZP_01775043.1|PVGSCARSALAA YLAER-----ADPGPAEPLILN-ARGGRLTRRSVARIVDAHMLLIA  
ref|YP_740910.1|PVGGKAREALQAWLAVRP-----ALAGADEPA VFVSQRGGRLSARSVQARLARLATLSG  
RAAC02854MVGRRAVDALRRYLPLR-----DRMARCSAVFINRRGGRLTD RSVRVLERRIQEV  
. * * *: : : : *: *. *: *: : :  
  
ref|YP_001125186.1|LTQNISP HVLRHTFATHLLNEGADLR SVQELLGHASLSSTQVYTHVTKDRL-----  
ref|NP_243331.1|LSTRVSPHS LRHSFATHLLNNGADLR VVQDLLGHENLSTTQVYTHVTKDRL-----  
ref|YP_175772.1|MOKKISPHAIRHSFATHLLNAGADLR AVQELLGHQS LKTTQVYTHVSKERL-----  
ref|ZP_01775043.1|AMRKVSPHTLRHTFATHLLEGGADLR AIQELLGHASLSSTQKYTHVSIDKL-----  
ref|YP_740910.1|VGRFPVHPMLRH SFASHLLESSGDLRAVQELLGHADIATTQVYTHLDFQHL-----  
RAAC02854GLRSIHVHGLRH SFATHMLNNGADLR SVQELLGHASLSSTQIYTHTSREQIARAYYAAHP  
: * :*:*: *: *: .*** *:*** : :** ** : :*  
  
ref|YP_001125186.1|-----  
ref|NP_243331.1|-----  
ref|YP_175772.1|-----  
ref|ZP_01775043.1|-----  
ref|YP_740910.1|-----  
RAAC02854RARRGT OGRESSEDGV
```

DINACTIVATEDDERIVATIVESPELTMACULUMTHERMPRPINICUMSIDBBAFTTRANS

PSASEANDINACTIVATEDDERIVATIVESPELTMACULUMTHERMPRPINICUMSIDBB

AFTRANSPSASEANDINACTIVATEDDERIVATIVESPELTMACULUMTHERMPRPINIC

-----MAQYQINIDSKILHQQL

UMSISCRESIGNIFICANCEE-IDENTITIESPSITIVESGAPSMQYQITVNQELLHQQL
-----MAQYQITLDSQLLHQQL
-----MVECETEKKPLHQEG-YRMAQYQITVDDEILEGL
-----MAELNLTLPDILLPNL

```

FLSNAKDAAMASFLEEVLNQILQAQATEQLRAERYERTDERQGYRNGVYPHRLTTRVGTI
-----
FLSNNKDSGIAALLESVLNQILQAQATEQLQAGPYERTDERKGYRNGTYPHQLTTRVGTI
FLSNAHDGAVAKLLESVLNQVLQAQATEQLGAEPYERTEERQGYRNGTYPHQLTTRVGTI
FQ---RDGGVAKLVESVVNQILQAQVQEQQLKAAPYERTKERQGYRNGTVTRTLTTRVGR
LTE---GGDGLKLVESVLNOVLEAOMTEHLGADRHERTEERAGYRNGVRERTLTRVGTV

```

```
TLKVPFRFDGKFSTEMFARYQRSEQALVLTLMEMVVNGVSTRKVAQITEELCGSEFSKST
-----MFARYQRSEQALVLTLMEMVVNGVSTRKVAQITEELCGSEFSKST
TLRVPRFRNGQFSTEMFARYQRSEQALVLTLMEMVINGVSTRKISQITEELCGAEFSKST
TLRVPIRNGKFSTELFARYQRSEQALVLALMEMVVNGVSTRKVAQITEELCGTEFSKST
VLRVPRVRNGQFSTELFARYQRSEQAFIALMEMVVNGVSTRDVAKITEQLCGTTFSKST
ILRVPTRDGSFSTDLFKRYQRSEQALVLSMMEMVVQGVSTRKANVITEELCGTRFSKST
.* *****.....*****..*.*.*.*.*.*.*.*.*.*
```

[illegible][illegible]

| | |
|--------------------|--|
| ref ZP_01666445.1 | HFIRNILDATPKSLQPEILARVRAILAAPDRETAVMLLNETLAAYETKAPKAMAILEAGF |
| ref ZP_01665334.1 | HFIRNILDATPKSLQPEILARVRAILAAPDRETAVMLLNETLAAYETKAPKAMAILEAGF |
| ref YP_001211938.1 | HFMRNIMDSAPKSVKEELYPRRLAILDAPDIGSARLLLNQTLAEFEKKAPRAMRVLEMGE |
| dbj BAD22831.1 | HFIRNILDATPKELQDEVHWSVRAILDAPDLDTARLLLNQVLETYETKAPKAMAILEAGF |
| RAAC03166 | HFMRNLLDATPKALQEEVYQQVRAVLDAPDLKTARLLKDAFVEAYAEKAPKAVQVLEDGF |
| gb EAY56013.1 | HFLRNILGHAPASQRGPLAQALSRLFRSETMEEARMVRNEILRTFEKKAPKAMECLEEGF |
| | **:*:*:. : * : : : : * : : : : : : ***:*: ** ** |
| ref ZP_01666445.1 | DDAIAVLALPEKYRKRLRTTNGVERLNEEVRRRERVIRIFPNRASALRLIGALLMEIDDK |
| ref ZP_01665334.1 | DDAIAVLALPEKYRKRLRTTNGVERLNEEVRRRERVIRIFPNRASALRLIGALLMEIDDK |
| ref YP_001211938.1 | DDATAVLVLPPEKYRLRLRTTNGVERLIEEVRRRERVIRIFPNRESVRLIGALLMEIDDK |
| dbj BAD22831.1 | EDATAVLLLPEKYRKRLRTTNALERLNEEIIRRERVIRIFPNRESAMRLIGALLMEID-- |
| RAAC03166 | DDVTAVLVLPERYRRRLRTTNGVERLNEEIIRRERVIRIFPNRESAIRLLGALLMEIDEE |
| gb EAY56013.1 | DETLNILTFFPKKYRVRLRSTNSQERLNEEIIRRERVIRIFPNRESAIRLLIGALLSEFHEQ |
| | ::. : * :*:** ***:*. *** **:*****. *:***:*** *:. |
| ref ZP_01666445.1 | WASGKKYLDMTDYWDWRERQ----- |
| ref ZP_01665334.1 | WASGKKYLDMTDYWDWRERQ----- |
| ref YP_001211938.1 | WAAGKKYLDMAEYLQWQKEQKHDR----- |
| dbj BAD22831.1 | ----- |
| RAAC03166 | WTTGRKYLMNDEYEAWKKAQEASRGSTQACAATA |
| gb EAY56013.1 | WSTGKKYLDMTEYHEWKKQESFKTSSTLA----- |

FIG. 102

```

ref|ZP_02327778.1|MDELKIHKHEWKKE--KLFVGDKWRGGDKQYVFHAGFGKPLYHTYPTQWWGEFIKRHNLK
ref|ZP_02327484.1|MDELKVHEYEWKKE--KLSVGDKWRGGDRQYVFHAGFGKPFHHTYPTTEWWNGFTKRHNLK
ref|ZP_02330756.1|MVELKEYEREWKKE--KLSVGDKWIGGDRRYVFHAGLGKPYFYSPSEWWSKFIKRHDLK
ref|ZP_02330395.1|MQELKVHEREWKKE--KLFMGDKWLGGDRQYVFHAGYGKPFHTTPTKWWRSFISRNKLN
ref|ZP_02326400.1|-DELSKFEIWNNE--KETAGQKWEERKHSFIFHNGLGKPFYRTVPSQRWLQFIRANNLP
RAAC02961MNELSKYREIWIKEWHLQQSGKWQGGEKQFLFHNGFGEKYYPSTPSLHWRRLFDDKHGLP
      ** . . * : * : . ** . : : * * * : . : * : * * : *

ref|ZP_02327778.1|RVRFHDLRHSSATLLIEAGASMKAIOERLGHSKHQTTADIIYAHITKKVSRETAEKFDKFA
ref|ZP_02327484.1|RVRFHDLRHSSATLLIEAGASMKAQQRLGHSKHQTTADIIYAHVTKKVSRDTAEKFNKFA
ref|ZP_02330756.1|RVRFHDLRHSSATLLIEAGASMKAQQRLGHSKHQTTADVYAHVTKKVSRDTAEKFDKFA
ref|ZP_02330395.1|YIRLHDLRHSSATLLIEAGAPMKAIQKRLGHSKHQTTADIIYAHVTKKVSRDTAEKFDKFA
ref|ZP_02326400.1|HIRLHDLRHTTVATLLLEEGVRLKVIQERHGHANYQTTADIIYSHVTKRLTEDAVDKFEKFG
RAAC02961RIRLHDLRHTTATILREDGADLKSIOERLRHTRLSTADLYTHETEAVSRETADRLEKLN
      *:*****: **:* * *. :* :*: * *. . ***:*** * : : : : : : : :

ref|ZP_02327778.1|P-----
ref|ZP_02327484.1|PNNIRPQSV
ref|ZP_02330756.1|PDSIRPQSV
ref|ZP_02330395.1|PNNIRPQSV
ref|ZP_02326400.1|P-----
RAAC02961PFRSRSQSI
      *
```

FIG. 103

```

ref|YP_146129.1|
sp|O87703|DNLJ_BACST
gb|ABN05294.1|
ref|YP_001124385.1|
ref|ZP_01169975.1|
RAAC02202
-----
MGARRDRRQVGRGRIAGTRRALRRPHRREAAVCQVRADSQGGSLSGFKHPMQCVIRSVMR

ref|YP_146129.1|
sp|O87703|DNLJ_BACST
gb|ABN05294.1|
ref|YP_001124385.1|
ref|ZP_01169975.1|
RAAC02202
-----
--QQAERRAAELRELLHRYGYEYYVLD RPSVPDAEYDRLMQELMAIEEQYPE
--QQAERRAAELRELLHRYGYEYYVLD RPSVPDAEYDRLMQELMAIEEQYPE
--QQAERRAAELRELLHRYGYEYYVLD RPSVPDAEYDRLMQELMAIEEQYPE
--QQAERRAAELRELLHRYGYEYYVLD RPSVPDAEYDRLMQELMAIEEQYPE
-----MDLQNAEARVKELHNLNLYGYEYYVLDKPSVPDSEYDRLKELMELENEHPQ
VVDAKSALSLEEARAKVLRQIEYHNRKYLEDNPEISDAEWDALMRDLIELERKYPE
      :*. * . *:: . . :*: .*. :*: * :*: :*. :*:

ref|YP_146129.1|
sp|O87703|DNLJ_BACST
gb|ABN05294.1|
ref|YP_001124385.1|
ref|ZP_01169975.1|
RAAC02202
-----
LKTSDSPTQRIGGPPEAFRKVTHVPMMSLANAFDEGDLRDFDRRVQEVGE-AAYVCE
LKTSDSPTQRIGGPPEAFRKVAHRVPMMSLANAFEGGDLRDFDRRVQEVGE-AAYVCE
LKTSDSPTQRIGGPPEAFRKVTHVPMMSLANAFDEGDLRDFDRRVQEVGE-AAYVCE
LKTSDSPTQRIGGPPEAFRKVTHVPMMSLANAFNEGDLRDFDRRVQEVGE-AAYVCE
LKTADSPTQVRVGGEILDMFEKVEHQTPMLSLGNFNEEDLRSFDRRVQTAGENLSYVCE
LVDPASPTQVRVGAPALEGFAKVVHEVPMLSLANAYSTEDLLDWRDRRVQAVGDDVRYVCE
* . *****:*. * : * * * . *****:*. ** :***** .*: ****

ref|YP_146129.1|
sp|O87703|DNLJ_BACST
gb|ABN05294.1|
ref|YP_001124385.1|
ref|ZP_01169975.1|
RAAC02202
-----
LKIDGLAVSVRYEDGYFVQGATRGDGTGEDITENLRTIRSLPLRLKEPVSLEARGEAFM
LKIDGLAVSVRYEDGYFVQGATRGDGTGEDITENLKTIRSLPLRLKEPVSLEARGEAFM
LKIDGLAVSVRYEDGYFVQGATRGDGTGEDITENLRTIRSLPLRLKEPVSLEARGEAFM
LKIDGLAVSVRYEDGYFVQGATRGDGTGEDITENLKTIRSLPLRLNEPVSLEARGEAFM
LKIDGLAVALKYEDGLFIQGATRGDGTIGEDITVNLRTIRSIPLRLSEPVSIEVRGEAFM
LKVDGLAVALRYQDGRVLVGLATRGDGSVGEDITANIRTIRNVPLELSEPVSLEVRGEAYM
** :*****::*:** : *****: ***** *:*****:*. *.*****:*.*****:

ref|YP_146129.1|
sp|O87703|DNLJ_BACST
gb|ABN05294.1|
ref|YP_001124385.1|
ref|ZP_01169975.1|
RAAC02202
-----
PKASFRLRLNEERKARGEELFANPRNAAAGSLRQLDPKVAASRQLDLFVYGLANAEEELGIE
PKASFRLRLNEERKARGEELFANPRNAAAGSLRQLDPKVAASRQLDLFVYGLADAEEALGIA
PKASFRLRLNEERKARGEELFANPRNAAAGSLRQLDPKVAASRQLDLFVYGLANAEEELGIE
PKASFRLRLNEERKARGEELFANPRNAAAGSLRQLDPKVAASRQLDLFVYGLANAEEELGIE
PGRSFEALNKGKEERGEPEFANPRNAAAGSLRQLDPRIAASRNLDIFLYGISNTGDTGVE
PKREFMRLNELREQQGEPLFANPRNAAAGSLRQLDPVAASRRLGVIVYQLVRAEAHGCE
* . * **:: : ** *****:***** :*****:*. :*: : : *

ref|YP_146129.1|
sp|O87703|DNLJ_BACST
gb|ABN05294.1|
ref|YP_001124385.1|
ref|ZP_01169975.1|
RAAC02202
-----
SHSEALDYLQALGFKVNPERRRCANIDEVIAFVNEWHEKRPQLPYEIDGIVIKVDSFAQQ
SHSEALDYLQALGFKVNPERRRCANIDEVIAFVSEWHDKRPQLPYEIDGIVIKVDSFAQQ
SHSEALDYLQALGFKVNPERRRCANIDEVIAFVNEWHEKRPQLPYEIDGIVIKVDSFAQQ
SHSAALSYLQSLGFKVNPERRRCATIDEVIAFVNEWHEKRPQLPYEIDGIVIKVDSFAQQ
SHSAGLDYLETLGFKTNKERRKCASIEEVIEYVSSWTEKRPNLDYDIDGIVIKVDSLALQ
THSQALDYVARLGLPAHRERHVCANIEDVIAIIEAWADKRHELPHYATDGMVVKVDSLALQ
:** .*: ** : ** : ** :*:** :. * ** :* * ** :*****: *

ref|YP_146129.1|
sp|O87703|DNLJ_BACST
gb|ABN05294.1|
ref|YP_001124385.1|
ref|ZP_01169975.1|
RAAC02202
-----
RELGATAKSPRWAIAYKFPAAEEVTTLIGIEVNVGRTGVVTPTAILEPVRVAGTTVQRAT
RALGATAKSPRWAIAYKFPAAEEVTTLIGIEVNVGRTGVVTPTAILEPVRVAGTTVQRAT
RELGATAKSPRWAIAYKFPAAEEVTTLIGIEVNVGRTGVVTPTAILEPVRVAGTTVQRAT
RQLGATAKSPRWAIAYKFPAAEEVTTLIGIEVNVGRTGAVTPTAILEPVRVAGTTVQRAT
EELGATAKSPRWAIAYKFPAAEEVTTLRDIELSVGRTGVVTPTAILEPVRVAGTTVQRAS
ARLGATAKSPRWAIAYKYAAEQAETTLRAIELNVGRTGVVTPTAVFDPVQLAGTTVSRAS
*****:*****:*. ** ** :*****:*****:*****:*****:

ref|YP_146129.1|
sp|O87703|DNLJ_BACST
gb|ABN05294.1|
ref|YP_001124385.1|
ref|ZP_01169975.1|
RAAC02202
-----
LHNEDFIREKDIRIGDAVI IKKAGDIIPEVVGVVVDRRDGDETPFAMPTHCPECESELVR
LHNEDFIREKDIRIGDAVI IKKAGDIIPEVVGVVVDRRDGDETPFAMPTHCPECESELVR
LHNEDFIREKDIRIGDAVI IKKAGDIIPEVVGVVVDRRDGDETPFAMPTHCPECESELVR
LHNEDFIREKDIRIGDAVI IKKAGDIIPEVVGVVVDRRDGDETPFAMPTHCPECESELVR
LHNEDLIREKDIKIDKVVVKKAGDIIPEVNVVLAEQRTGEEREFFIMPTHCPECGSELVR
LHNEDLVREKDIRVGDVIVVQKAGDIIPEVIRSLPERTEPLFRMPETCPQCGSRLVR
*****:*****:** :*:*****: : : * * * * * :* * *

```

ref|YP_146129.1|
sp|O87703|DNLJ_BACST
gb|ABN05294.1|
ref|YP_001124385.1|
ref|ZP_01169975.1|
RAAC02202

LEGEVALRCLNPNCPAQLRERLIHFASRAAMNIEGLGEKVVTQLFNAGLVRDVADLYRLT
LEGEVALRCLNPNCPAQLRERLIHFASRAAMNIEGLGEKVVTQLFNAGLVRDVADLYCLT
LDGEVALRCLNPKCPAQLRERLIHFASRAAMNIEGLGEKVVTQLFNAGLVHDVADLYRLT
LDGEVALRCLNPKCPAQLRERLIHFASRAAMNIEGLGEKVVTQLFNAGLVHDVADLYQLT
LEGEVALRCINPKCPAQIREGLIHVFSRNAMNIDGLGEKVVSQLFKELIKDVADLYKLT
LDGEVAWRCINPDCPALLREGLIHFCSDAMNIEGLGEQWITVLLDRGLVRTHADLYRLR
*:**** *:*** ** :** ***** ** *****:****: :: *: *: : **** *

ref|YP_146129.1|
sp|O87703|DNLJ_BACST
gb|ABN05294.1|
ref|YP_001124385.1|
ref|ZP_01169975.1|
RAAC02202

KEQLIGLERMGEKSATNLLAAIEASKQNSLERLLFGLGIRYVGAKAAQLLAEHFETMERL
KEQLVGLERMGEKSAANLLAAIEASKQNSLERLLFGLGIRYVGAKAAQLLAEHFETMERL
KEQLIGLERMGEKSATNLLAAIEASKQNSLERLLFGLGIRYVGAKAAQLLAEHFETMERL
KEQLVGLERMGEKSAANLLAAIEASKQNSLERLLFGLGIRYVGAKAAQLLAEHFETMERL
HEALIGMERMGEKSVNNLVQAIEASKQNSLEKLLFGLGIRHVGAKAAKTLAQHFSSMDNL
KADLVQLDRMGDKLADKLLHNIQESKRNSLERLLFGLGIRHVGEKAAKTLAEHFVTIDAL
: *: :****: * . :*: *: **:*****:*****:*** **: **:*** :: *

ref|YP_146129.1|
sp|O87703|DNLJ_BACST
gb|ABN05294.1|
ref|YP_001124385.1|
ref|ZP_01169975.1|
RAAC02202

ERATKEELMAVPEIGEKMAITAFAAQPEATELLQELRAYGVNMAYKGPKRSAEAPADS
ERATKEELMAVPEIGEKMAITAFAAQPEATELLQELRAYGVNMAYKGPKRSAEAPADS
EKATKEELMAVPEIGEKMAISIIAFFSQPEAAELLHELRLYGVNMAYKGPKRSAEAPADS
EAATKDELMVPEIGEKMAISITTFYSQPEAVELLNELRTYGVNMAYKGRKRTAETPASS
MAASEDDLTAINIEGKMGAGIIVAFQEAEKELIRELKESGVNMEYKGPKPVAEEDSDS
MSASEEDLMAVPDIPGKVAQSIRQYFDTPRVRQLIQELKDLGVNMTYLGPKVSDGP---
*:**** *: **:*** :* :* .. :*:****: **** * * : :

ref|YP_146129.1|
sp|O87703|DNLJ_BACST
gb|ABN05294.1|
ref|YP_001124385.1|
ref|ZP_01169975.1|
RAAC02202

AFAGKTVVLTGKLASMSRNEAKEEIERLGGRVTGSVSRSTDVLIAGEDAGSKLEKAQQLG
AFAGKTVVLTGKLASMSRNEAKEEIERLGGRVTGSVSRSTDVLIAGEDAGSKLEKAQQLG
AFAGKTVVLTGKLASMSRNEAKEEIERLGGRVTGSVSRSTDVLIAGEDAGSKLEKAQQLG
VLGKTVVLTGKLASMSRNEAKEEIERLGGRVTGSVSRSTDVLIAGEDAGSKLDKAQQLG
FFAGKTIVLTGKLSILSRNEAKEIEALGGNVSGSVSKKTDLLIAGEDAGSKLAKAQDLG
-LAGKTVVLTGVLQAADRKQATAWIEQMGGKVASSVSAKTDVLIAGDKAGSKLAKAQEIL
:****:**** * .*:*. ** :****:*** .***:***:***** **::

ref|YP_146129.1|
sp|O87703|DNLJ_BACST
gb|ABN05294.1|
ref|YP_001124385.1|
ref|ZP_01169975.1|
RAAC02202

-----IEIWDESRLFQEIN-----
-----IEIWDESRLFQEIN-----
-----IEIWDESRLFQEIN-----
-----IEIWDETRFLQ-----
-----IDIWNEEQVLEE-----
RNHPDAKLEIWDEAAFLRLVDEAGLR
::***: * :..

FIG. 104

```

ref|ZP_01695687.1| -----VGPSQLLSKLCDEIELEKTINDLVKWDVSRCHISPGTRIKALVLNLCSGK
ref|ZP_01695982.1| -----VGPSQLLSKLCDEIELEKTINDLVKWDVSRCHISPGTRIKALVLNLCSGK
ref|ZP_01695971.1| -----VGPSQLLSKLCDEIELEKTINNVLKWDVSRCHISPGTRIKALVLNLCSGK
ref|ZP_01695655.1| -----PQSCKLCDEIELEKTINNVLKWDVSRCHISPGTRIKALVLNLCSGK
ref|YP_430569.1| ----MRFFRAGPAALISRLCDVLKIAEIIDAVVDWDPAQCHLSPGNRVKALIINLLVDRE
RAAC03682 MFGPVRSYVMGPAPVLARLIDELKWVEIIDFVPRP--DSKLSVGLRTRKALLVNIGTNR
          .:* *: : *: .*      .::* * * *::*: : :

ref|ZP_01695687.1| PLYKVHEFYQNLDSEMLF--DTSVSPDQLNDDALGRALDYLYEA-EAWKVYSTLALKTLK
ref|ZP_01695982.1| PLYKVHEFYQNLDSEMLF--DTSVSPDQLNDDALGRALDYLYEA-EAWKVYSTLALKTLK
ref|ZP_01695971.1| PLYKVHEFYQNLDSEMLF--DTSISPDQLNDDALGRALDYLYKA-EAWKVYSTLALKTLK
ref|ZP_01695655.1| PLYKVHEFYQNLDSEMLF--DTSISPDQLNDDALGRALDYLYKA-EAWKVYSTLALKTLK
ref|YP_430569.1| ALYHVERFYENQDLEVLFGAEQQVRPEDFNDDALGRALDKLFTSGQLKKLFSSIALTAA
RAAC03682 ALYRVEEFYAQRDVEVLG--SGVSADDLHDDALARALDALYDA-GLEALYARIALHTLR
          .*:.*.*: * *: : : : :*:.*.* * : : : :*: :

ref|ZP_01695687.1| KLNLP-----IGILHNDTTSISVYGEYKQPKEDGLQITYGYSKAHRPDLKQIVLGMGVTP
ref|ZP_01695982.1| KLNLP-----IGILHNDTTSISVYGEYKQPKEDGLQITYGYSKAHRPDLKQIVLGMGVTP
ref|ZP_01695971.1| KLNLP-----IGVLHNDTTSISVYGEYKQPKEDGLQITYGYSKAHRPDLKQIVLGMGVTP
ref|ZP_01695655.1| KLNLP-----IGVLHNDTTSISVYGEYKQPKEDGLQITYGYSKAHRPDLKQIVLGMGVTP
ref|YP_430569.1| THNVS-----IAGIHVDTTSISVQAYDG--EGDLITFGFSKDHPRDLKQIFLIGLTVNR
RAAC03682 RLRVLSDSNELIPIHADTTSLSMTGEYLD--QTAFRIDRGFSKDHPRDLKQIVFGL-CTV
          . : : : * *::*: * * : : * *::*: * *::*: * *::*: * *::*: *

ref|ZP_01695687.1| ERIPILAKVENGNTSDKSWNVEFIQMRKILSHEDWKNLIYQADSALITTENLAEIQQQN
ref|ZP_01695982.1| ERIPILAKVENGNTSDKSWNVEFIQMRKILSHEDWKNLIYQADSALITTENLAEIQQQN
ref|ZP_01695971.1| ERIPILAKVENGNTSDKSWNVEFIQMRKILSHEDWKNLIYQADSALITTENLAEIQQQN
ref|ZP_01695655.1| ERIPILAKVENGNTSDKSWNVEFIQMRKILSHEDWKNLIYQADSALITTENLAEIQQQN
ref|YP_430569.1| DGLPILAQSLDGNSSDKSWYPQVIEELVQTFKPEKLEKIVFVADCALVTKDNLALLVQEE
RAAC03682 HGLGLCANVNPGLNDHTWNFENIQQLLSQLEETRKRVSYYADAALVTKDNLELLAEED
          . : : * : ** .*: : *::: . . * * . : **.*.*.*: : :

ref|ZP_01695687.1| -----LSFISRLPDTFGLSTELKKEAWLLNN-WERVGSLSNKKDAAIYQIQAFERQIQNL
ref|ZP_01695982.1| -----LSFISRLPDTFGLSTELKKEAWLLNN-WERVGSLSNKKDAAIYQIQAFERQIQNL
ref|ZP_01695971.1| -----LSFISRLPDTFGLSTELKKEAWLLNN-WERVGSLSNKKDAAIYQIQAFERQIQNL
ref|ZP_01695655.1| -----LSFISRLPDTFGLSTELKKEAWLLNN-WERVGSLSNKKDAAIYQIQAFERQIQNL
ref|YP_430569.1| GNKPALQFISLLPENFGLNKEIKAEAFRTGT-WQEIGKLSPPKDAACYKSQSFVREIDGR
RAAC03682 -----FHFISRLPGTYKLSDELKRAAWEKENSWEVGRLAEAEDSAHYRIQAFRTLYGR
          : *** * : * . : * * : . * . : * * : * : * * : * :

ref|ZP_01695687.1| PYRFLVVHSSNNLDQRKEKTLNRAIEKEEIQLKKEIEKLSKVIHFCKEDALEAIQSFKKKQ
ref|ZP_01695982.1| PYRFLVVHSSNNLDQRKEKTLNRAIEKEEIQLKKEIEKLSKVIHFCKEDALEAIQSFKKKQ
ref|ZP_01695971.1| PYRFLVVHSSNNLDQRKEKTLNRAIEKEEIKLKKEIEKLSKVIHFCKEDALEAIQSFKKKQ
ref|ZP_01695655.1| PYRFLVVHSSNNLDQRKEKTLNRAIEKEEIQLKKEIEKLSKVIHFCKEDALEAIQSFKKKQ
ref|YP_430569.1| DYRLIVVHSTTLDRKRENSLLKKWAKQREVLEKAAKDLRRPFACKADARKAIELFLREY
RAAC03682 TYRFFVVRSSSLDTRKERKLKEVLKREKAALEKAAKAMSQNVYSCEQDAQMAMQTFMHEH
          *:.*.*.*.* * . : . * : * : : * : * : * : * :

ref|ZP_01695687.1| KASYFCYKLVQMSDQPIKRRKGRPRPKDEQTK-REQIYQIQLESFEKDHDFIENKKML
ref|ZP_01695982.1| KASYFCYKLVQMSDQPIKRRKGRPRPKDEQTK-REQIYQIQLESFEKDHDFIENKKML
ref|ZP_01695971.1| KASYFCYKLVQMSDQPIKRRKGRPRPKDEQTK-REQIYQIQLESFEKDHDFIENKKML
ref|ZP_01695655.1| KASYFCYKLVQMSDQPIKRRKGRPRPKDEQTK-REQIYQIQLESFEKDHDFIENKKML
ref|YP_430569.1| R-----
RAAC03682 RATLHPISARICAEQVQAKRARRGRPRKDDPPPPVHTQYRVEVAILPPSEERVQOWREKE
          :

ref|ZP_01695687.1| STFVLITNKLDEETLSNQEVLRVYKQGSAAETFRFLIKDSQMIDAIYKTPERVEALGIV
ref|ZP_01695982.1| STFVLITNKLDEETLSNQEVLRVYKQGSAAETFRFLIKDSQMIDAIYKTPERVEALGIV
ref|ZP_01695971.1| STFVLITNKLDEETLSNQEVLRVYKQGSAAVETFRFLIKDSQMIDAIYKTPERVEALGIV
ref|ZP_01695655.1| STFVLITNKLDEETLSNQEVLRVYKQGSAAETFRFLIKDSQMIDAIYKTPERVEALGIV
ref|YP_430569.1| -----
RAAC03682 ATFVLITDIRDDQRVSDQILRLYKEQHEVEARFRYLKSPYHVGPIYLHKPTRVKAFGFV

```


ref|ZP_01695687.1|
ref|ZP_01695982.1|
ref|ZP_01695971.1|
ref|ZP_01695655.1|
ref|YP_430569.1|
RAAC03682

YVMALLIYGILEYRVKELKEKNLSLILKGKRKLYQPTGQALLEQLEDITVILINQNQQK
YVMALLIYGILEYRVKELKEKNLSLILKGKRKLSQPTGQALLEQLEDITVILINQNQQK
YVMALLIYGILEYRVKELKEKNLSLILKGKRKLYQPTGQALLEQLEDITVILINQNRQK
YVMALLIYGILEYRVKELKEKNLSLILKGKRKLYQPTGQALLEQLEDITVILINQNRQK

MLLSLLLYSVLEYLIREKMKRETEPLMLPGNRKSFRTGLAILEMLDGVTTVHRQVGDTW

ref|ZP_01695687.1|
ref|ZP_01695982.1|
ref|ZP_01695971.1|
ref|ZP_01695655.1|
ref|YP_430569.1|
RAAC03682

IRLLPDNIDSQAKKIIELCGYDLSIY-----
IRLLPDNIDSQAKKIIELCGYDLSIY-----
LRLLPDNIDSQAKKIIELCGYDLSIY-----
LRLLPDNIDSQAKKIIELCGYDLSIY-----

QRVPATPHNPQIMRVLKLLNMDLSIYTEAQKTA

FIG. 105

```

ref|ZP_01695687.1| -----VGPSQLLSKLCDEIELEKTINDLVKWDVSRCHISPGTRIKALVLNLCSGK
ref|ZP_01695982.1| -----VGPSQLLSKLCDEIELEKTINDLVKWDVSRCHISPGTRIKALVLNLCSGK
ref|ZP_01695971.1| -----VGPSQLLSKLCDEIELEKTINNVLKWDVSRCHISPGTRIKALVLNLCSGK
ref|ZP_01695655.1| -----PQSCKLCDEIELEKTINNVLKWDVSRCHISPGTRIKALVLNLCSGK
ref|YP_430569.1| ----MRFFRAGPAALISRLCDVLKIAEIIDAVVDWDPAQCHLSPGNRVKALIINLLVDRE
RAAC03770 MFGPVRSYVMGPAPVLARLIDELKWVEIIDFVPRP--DSKLSVGLRTRKALLVNIGTNR
          .:* * : : * : . *      . : * * * * : : : : . :

ref|ZP_01695687.1| PLYKVHEFYQNLDSEMLF--DTSVSPDQLNDDALGRALDYLYEA-EAWKVYSTLALKTLK
ref|ZP_01695982.1| PLYKVHEFYQNLDSEMLF--DTSVSPDQLNDDALGRALDYLYEA-EAWKVYSTLALKTLK
ref|ZP_01695971.1| PLYKVHEFYQNLDSEMLF--DTSISPDQLNDDALGRALDYLYKA-EAWKVYSTLALKTLK
ref|ZP_01695655.1| PLYKVHEFYQNLDSEMLF--DTSISPDQLNDDALGRALDYLYKA-EAWKVYSTLALKTLK
ref|YP_430569.1| ALYHVERFYENQDLEVLFGAEQQVRPEDFNDDALGRALDKLFTSGQLKKLFSSIALTAA
RAAC03770 ALYRVEEFYAQRDVEVLLG--SGVSADDLHDDALARALDALYDA-GLEALYARIALHTLR
          .*:.*.*** : * * : : : : : : : : : : : : : : : : : : : : : :

ref|ZP_01695687.1| KLNLP-----IGILHNDTTSISVYGEYKQPKEDGLQITYGYSKAHRPDLKQIVLGMGVTP
ref|ZP_01695982.1| KLNLP-----IGILHNDTTSISVYGEYKQPKEDGLQITYGYSKAHRPDLKQIVLGMGVTP
ref|ZP_01695971.1| KLNLP-----IGVLHNDTTSISVYGEYKQPKEDGLQITYGYSKAHRPDLKQIVLGMGVTP
ref|ZP_01695655.1| KLNLP-----IGVLHNDTTSISVYGEYKQPKEDGLQITYGYSKAHRPDLKQIVLGMGVTP
ref|YP_430569.1| THNVS-----IAGIHVDTTSISVQGAIDG--EGDLITFGFSKDRPDLKQIFLIGLTVNR
RAAC03770 RLRVLSDSNELIPIHADTTSLSMTGEYLD--QTAFRIDRGFSKDRPDLKQIVFGL-CTV
          . :      : : * * : : : : : : : : : : : : : : : : : : : : : :

ref|ZP_01695687.1| ERIPILAKVENGNTSDKSWNVEFIQMRKILSHEDWKNLIYQADSALITTENLAEIQQQN
ref|ZP_01695982.1| ERIPILAKVENGNTSDKSWNVEFIQMRKILSHEDWKNLIYQADSALITTENLAEIQQQN
ref|ZP_01695971.1| ERIPILAKVENGNTSDKSWNVEFIQMRKILSHEDWKNLIYQADSALITTENLAEIQQQN
ref|ZP_01695655.1| ERIPILAKVENGNTSDKSWNVEFIQMRKILSHEDWKNLIYQADSALITTENLAEIQQQN
ref|YP_430569.1| DGLPILAQSLDGNSSDKSWYPQVIEELVQTFKPEKLEKIVFVADCALVTKNLALLVQEE
RAAC03770 HGLGLCANVNPGLNDHTWNFENIQQLLSQLEETRKRVSYYVADAALVTKNLLELLAEED
          . : : * : ** . : * : : : : . . * * . : : **.*.*. : : : : : : : :

ref|ZP_01695687.1| -----LSFISRLPDTFGLSTELKKEAWLLNN-WERVGSLSNKKDAAIYQIQAFERQIQNL
ref|ZP_01695982.1| -----LSFISRLPDTFGLSTELKKEAWLLNN-WERVGSLSNKKDAAIYQIQAFERQIQNL
ref|ZP_01695971.1| -----LSFISRLPDTFGLSTELKKEAWLLNN-WERVGSLSNKKDAAIYQIQAFERQIQNL
ref|ZP_01695655.1| -----LSFISRLPDTFGLSTELKKEAWLLNN-WERVGSLSNKKDAAIYQIQAFERQIQNL
ref|YP_430569.1| GNKPALQFISLLPENFGLNKEIKAEAFRTGT-WQEIIGKLSPPKDAACYKSQSFVREIDGR
RAAC03770 -----FHFISRLPGTYKLSDELKRAAWEKENSWEVGRLAEAEDSAHYRIQAFRTLYGR
          : *** * : : * . : : * : . * . : * * : : * * * : : * : .

ref|ZP_01695687.1| PYRFLVVHSSNNLDQRKEKTLNRAIEKEEIQLKKEIEKLSKVIHFCKEDALEAIQSFKKKQ
ref|ZP_01695982.1| PYRFLVVHSSNNLDQRKEKTLNRAIEKEEIQLKKEIEKLSKVIHFCKEDALEAIQSFKKKQ
ref|ZP_01695971.1| PYRFLVVHSSNNLDQRKEKTLNRAIEKEEIKLKKEIEKLSKVIHFCKEDALEAIQSFKKKQ
ref|ZP_01695655.1| PYRFLVVHSSNNLDQRKEKTLNRAIEKEEIQLKKEIEKLSKVIHFCKEDALEAIQSFKKKQ
ref|YP_430569.1| DYRLIVVHSTTLDKRKENSLLKKWAKQREVLEKAAKDLSSRPFACKADARKAIELFLREY
RAAC03770 TYRFVVRSSSLDTRKERKLKEVLKREKAALEKAAKAMSQNVYSCEQDAQMAMQTFMHEH
          **: : **.*.*** **.*. . : . * : : : : : : : : : : : : : : : : : :

ref|ZP_01695687.1| KASYFCYKLVQMSDQPIKRRKGRPRPKDEQTK-REQIYQIQLESFEKDHDFIENKKML
ref|ZP_01695982.1| KASYFCYKLVQMSDQPIKRRKGRPRPKDEQTK-REQIYQIQLESFEKDHDFIENKKML
ref|ZP_01695971.1| KASYFCYKLVQMSDQPIKRRKGRPRPKDEQTK-REQIYQIQLESFEKDHDFIENKKML
ref|ZP_01695655.1| KASYFCYKLVQMSDQPIKRRKGRPRPKDEQTK-REQIYQIQLESFEKDHDFIENKKML
ref|YP_430569.1| R-----
RAAC03770 RATLHPISARICAEQVQAKRARRGRPRKDDPPPPVHTQYRVEVAILPPSEERVQQWREKE
          :

ref|ZP_01695687.1| STFVLITNKLDEETLSNQEVLRVYKQGSAAETFRFLIKDSQMIDAIYKTPERVEALGIV
ref|ZP_01695982.1| STFVLITNKLDEETLSNQEVLRVYKQGSAAETFRFLIKDSQMIDAIYKTPERVEALGIV
ref|ZP_01695971.1| STFVLITNKLDEETLSNQEVLRVYKQGSAAVETFRFLIKDSQMIDAIYKTPERVEALGIV
ref|ZP_01695655.1| STFVLITNKLDEETLSNQEVLRVYKQGSAAETFRFLIKDSQMIDAIYKTPERVEALGIV
ref|YP_430569.1| -----
RAAC03770 ATFVLITDIRDDQRVSDQILRLYKEQHEVEARFRYLKSPYHVGPIYLHKPTRVKAFGV

```

ref|ZP_01695687.1|
ref|ZP_01695982.1|
ref|ZP_01695971.1|
ref|ZP_01695655.1|
ref|YP_430569.1|
RAAC03770

YVMALLIYGILEYRVKELKEKNLSLILKGKRKLYQPTGQALLEQLEDITVILINQNQQK
YVMALLIYGILEYRVKELKEKNLSLILKGKRKLSQPTGQALLEQLEDITVILINQNQQK
YVMALLIYGILEYRVKELKEKNLSLILKGKRKLYQPTGQALLEQLEDITVILINQNRQK
YVMALLIYGILEYRVKELKEKNLSLILKGKRKLYQPTGQALLEQLEDITVILINQNRQK

MLLSLLLYSVLEYLIREKMKRETEPLMLPGNRKSFRTGLAILEMLDGVTTVHMQVGDTW

ref|ZP_01695687.1|
ref|ZP_01695982.1|
ref|ZP_01695971.1|
ref|ZP_01695655.1|
ref|YP_430569.1|
RAAC03770

IRLLPDNIDSQAKKIIELCGYDLSIY-----
IRLLPDNIDSQAKKIIELCGYDLSIY-----
LRLLPDNIDSQAKKIIELCGYDLSIY-----
LRLLPDNIDSQAKKIIELCGYDLSIY-----

QRVPATPHNPQIMRVLKLLNMDLSIYTEAQKTA

FIG. 106

ref|YP_001666203.1|
ref|YP_001512589.1|
ref|NP_842969.1|
ref|ZP_02596019.1|
ref|YP_430640.1|
RAAC02738

EINTEGRASEFAMILYBACILLUSANTHRACISSTRAREFZPPRPHAGELAMBDAASIT

ref|YP_001666203.1|
ref|YP_001512589.1|
ref|NP_842969.1|
ref|ZP_02596019.1|
ref|YP_430640.1|
RAAC02738

E-SPECIFICRECMBINASEPHAGEINTEGRASEFAMILYBACILLUSANTHRACISSTR

ref|YP_001666203.1|
ref|YP_001512589.1|
ref|NP_842969.1|
ref|ZP_02596019.1|
ref|YP_430640.1|
RAAC02738

AREFZPPRPHAGELAMBDAASITE-SPECIFICRECMBINASEPHAGEINTEGRASEFA

ref|YP_001666203.1|
ref|YP_001512589.1|
ref|NP_842969.1|
ref|ZP_02596019.1|
ref|YP_430640.1|
RAAC02738

MILYBACILLUSANTHRACISSTRAREFZPPRPHAGELAMBDAASITE-SPECIFICRE

ref|YP_001666203.1|
ref|YP_001512589.1|
ref|NP_842969.1|
ref|ZP_02596019.1|
ref|YP_430640.1|
RAAC02738

CMBINASEPHAGEINTEGRASEFAMILYPRTEINBACILLUSANTHRACISTSIANKVSK

ref|YP_001666203.1|
ref|YP_001512589.1|
ref|NP_842969.1|
ref|ZP_02596019.1|
ref|YP_430640.1|
RAAC02738

II-IGBAAPPRPHAGELAMBDAASITE-SPECIFICRECMBINASEPHAGEINTEGRAS

ref|YP_001666203.1|
ref|YP_001512589.1|
ref|NP_842969.1|
ref|ZP_02596019.1|
ref|YP_430640.1|
RAAC02738

EFAMILYBACILLUSANTHRACISSTRAMESGBAATPRPHAGELAMBDAASITE-SPEC

ref|YP_001666203.1|
ref|YP_001512589.1|
ref|NP_842969.1|
ref|ZP_02596019.1|
ref|YP_430640.1|
RAAC02738

IFICRECMBINASEPHAGEINTEGRASEFAMILYBACILLUSANTHRACISSTRAMESAN

ref|YP_001666203.1|
ref|YP_001512589.1|
ref|NP_842969.1| CESTRGBAATPRPHAGELAMBDA SITE-SPECIFIC RECOMBINASE PHAGE INTEGRASE
ref|ZP_02596019.1|
ref|YP_430640.1|
RAAC02738

ref|YP_001666203.1|
ref|YP_001512589.1|
ref|NP_842969.1| SEFAMILYBACILLUS ANTHRACIS STRS TERNEGBEDRPRPHAGELAMBDA SITE-S
ref|ZP_02596019.1|
ref|YP_430640.1|
RAAC02738

ref|YP_001666203.1|
ref|YP_001512589.1|
ref|NP_842969.1| PECIFIC RECOMBINASE PHAGE INTEGRASE FAMIL
ref|ZP_02596019.1|
ref|YP_430640.1|
RAAC02738

ref|YP_001666203.1|
ref|YP_001512589.1|
ref|NP_842969.1| EDRPRPHAGELAMBDA SITE-SPECIFIC RECOMBINASE PHAGE INTEGRASE FAMIL
ref|ZP_02596019.1|
ref|YP_430640.1|
RAAC02738

ref|YP_001666203.1|
ref|YP_001512589.1|
ref|NP_842969.1| YBACILLUS ANTHRACIS STRAGBEDRPRPHAGELAMBDA SITE-SPECIFIC RECOMB
ref|ZP_02596019.1|
ref|YP_430640.1|
RAAC02738

ref|YP_001666203.1|
ref|YP_001512589.1|
ref|NP_842969.1| INASE PHAGE INTEGRASE FAMIL
ref|ZP_02596019.1|
ref|YP_430640.1|
RAAC02738

ref|YP_001666203.1|
ref|YP_001512589.1|
ref|NP_842969.1|
ref|ZP_02596019.1|
ref|YP_430640.1|
RAAC02738

-----YSIVVDIGRDEN-GKRKQKWISGFKTKKEAEK
-----KKNHKWYIVVDS-IDEN-GKRKRKWISGFNTQKEAEK
CEE-IDENTITIESPSITIVESGAPSEKWSFTIDIGKDPITGKRKQKTASGFKTKKEAER
-----TVKKDDTKTKTWYFVLTHGKKED-GKPRQFKRGFKTKQEAELK
-----WTAQITVGIDHE-GKQRRLTLYG-KSKKEVLE
-----MTTGRVFSYTTKSNEKKYGYTITITPADPQT-GRYPQVTRKGFKSERAARKS
: : *: : * : : .

ref|YP_001666203.1|
ref|YP_001512589.1|
ref|NP_842969.1|
ref|ZP_02596019.1|
ref|YP_430640.1|
RAAC02738

ALAEIIAKIEKNELELTN-LTLGQFLNKWYETHCELKLA PKTLKSYKELIELYFKPYLGD
KLPEILQQLNTNTFVDTNNITLRDYMQLWLEDYAKPNVAPT TYDKYTYAAN-KVSEILGN
ACNELIHQFNTGSLVDDKNFTLSEYLQEWLENTAKQVRRETTFTNYKRAINSRIIPVLGS
ALHELEHSLMTGTIYAPT KMLYGEYLLNQWLEDKQTKVKQT LQTYRWLVEKHII PAIGN
KLTQALYQQONGGFIEPSKITVEQWLNRLWLT DYAKPHLRQSTWESYETVLRHLV IPTLGS
AMMEHLRSLEDTPVHVAKSFTLHEWASRWFADKS-SSLRPWTRRTYHHVLQAYILPELKG
: . . : : : * . * . . : *

ref|YP_001666203.1|
ref|YP_001512589.1|
ref|NP_842969.1|
ref|ZP_02596019.1|
ref|YP_430640.1|
RAAC02738

IELTKLKPVMIQNYNKLK-ELGLSDTTINYHHRVLKSALKKAVVWQLISKNPCDYVEPP
MKLQQIKPIHIQQLVNKLN-NSDVT PSTVLSYYRVLNTAINQAVKWQFIQYNPCVAVTPP
HKLKDLKPLHGQRFVKS LI-DEGLSPAYIEYIFIVLKGSL EDAVRWELLFKNP FQHV EIP
MELTQLTPMAIQKLYNQLTKEKALSDENIQKVHTLINDSLKKAERWGIISKNPASLVDRP
IPLKKLQPADIQRLYASKL-ESGLSPTRVRYIHVVLHEAMSQARESGLLLQNPT EAAKPP
MPISKIEARHIQALYRKLS--NRLKPVTVHRIHVLKTCLLAAVKAGYLDKSPFLNVEPP
: . : * . : : : : * : : * . *

ref|YP_001666203.1|
 ref|YP_001512589.1|
 ref|NP_842969.1|
 ref|ZP_02596019.1|
 ref|YP_430640.1|
 RAAC02738

KKNKNEITVWSINDVKKAKE--IFKDTPIYLHFMALALYTGMRRIGEICGLKWEDIDFNNKT
 RNTKNKMMILDQDEIQILLD--KSKDHVLYPVIVLALLCGLRRGEILGLQWENIDFFSGV
 RPRKVVNSTWSIEETKKFLNRTKFENVIIYHLFLLALNTGMRRGEILGLKWKNFDLNEGK
 KAEKKEIKVWDVKEVQTFLK-HAQSHSRYYIAFLALTTGMRRGEILGLRWRDVFETGC
 RHPKKKVQPLNPEQVKRFLE--TAKQDPLYPAFLALGTGLRRGEILGLRWQDLDLQKGI
 EHKTTPPKPVLVNDFAFLRLAWLREHRTSYMAAFLAIHTGMRMGEIAGLQWRDIDLDTGV
 * . ** : * : * * * * : * : . *

ref|YP_001666203.1|
 ref|YP_001512589.1|
 ref|NP_842969.1|
 ref|ZP_02596019.1|
 ref|YP_430640.1|
 RAAC02738

CTVKRQYQQVGGKEIHK-EPKSETSIIRVIPLHSDVIEVLKEEKKKQLQNRMLLGEKYNKK
 IHLENNLMANNESILK-ETKTSTGRRAVDISSNVVEVLKKVKKQKMSYKLLYGSSYHDS
 ISVTETLIYDENGFRFT-EPKTHGSKRLISIDQNLCKEFKSYKAKQNEFKLLFGQSYEDN
 IRITQTLSS-DGKEILP-YTKTKSGSRTVDLPEETIIQLKKHRKLISEKLEAGSEVYNN
 LQVRQSLIRTRGLKFE-EPKTEKSRRIPLPPSVVAALKRHKAWVNQNKILILGPDYEDH
 IQLETRTRYRPKGGQDFLGPCKTFGSRRIIVVTREVVDLRRWK--QSQQEIERESWTPES
 : . : . * : . * : : . : :

ref|YP_001666203.1|
 ref|YP_001512589.1|
 ref|NP_842969.1|
 ref|ZP_02596019.1|
 ref|YP_430640.1|
 RAAC02738

YEGYISVWEDGRMKTPEYVSKKFSKILKAYPELPQIRFHDRLRHSCASFLVQAGVPMKVVS
 --NFVCTWEDGRPFREDYIPKAFKILVS-ANLPKIRFHDRLRTHASILLKSGIHPKVQVQ
 --DLVFAKETGQPILPRTMTTTFNQFIKK-ADVPQIRFHDRLRTHATILLKLGINPKIVS
 -WDLVVCTELGTPTNKSNIIRSFNSI IKK-AKIPKIRFHDRLRTHATILLQLGVNPKIVS
 --DLVFPVENGRPRDPKGFAEYFNRLLDK-AGLPHIRLHDLRTHATILLLEGVHPKVQVQ
 ---FVVRLPNSAPPSPASFNNAIQNARKE-LGLPPVSFHLRTHATWLLSGVDLKIIVS
 : . . . : * : : * : * : * : * : *

ref|YP_001666203.1|
 ref|YP_001512589.1|
 ref|NP_842969.1|
 ref|ZP_02596019.1|
 ref|YP_430640.1|
 RAAC02738

EILGHSQIGITMDLYSHVLLDSKKEAIKKLEEYLQ-----
 ERLGHSSISITLDTYSHLVPSLQSAAEKMATM-----
 ERLGHSSIKTTLDTYSHVTIDMQESAVLKLSEALKS-----
 ERLGHADVRLTDLTYSHLLPSMQKDTAIFGEML-----
 ERLGHSTVSITLDIYSHILPGLQEKAAERIDGLLQPK-----
 ERLGHSSITITADIYAHVTDALQREAIEKLQRMMSRRTNNSGSDDEEDL
 * * * : * * * : * : : : :

FIG. 108

```
ref|YP_001111555.1| TRANSPSASEISFAMILYPRTEINDESULFTMACULUMREDUCENSMI-GBABTRANSPS
ref|YP_001114460.1| -----
ref|YP_001111684.1| -----
ref|YP_001113963.1| -----
dbj|BAD18231.1| -----
RAAC02515 -----

ref|YP_001111555.1| ASEISFAMILYPRTEINDESULFTMACULUMREDUCENSMI-GBABTRANSPSASEISFA
ref|YP_001114460.1| -----
ref|YP_001111684.1| -----
ref|YP_001113963.1| -----
dbj|BAD18231.1| -----
RAAC02515 -----

ref|YP_001111555.1| MILYPRTEINDESULFTMACULUMREDUCENSMI-GBABTRANSPSASEISFAMILYPRT
ref|YP_001114460.1| -----
ref|YP_001111684.1| -----
ref|YP_001113963.1| -----
dbj|BAD18231.1| -----
RAAC02515 -----

ref|YP_001111555.1| EINDESULFTMACULUMREDUCENSMI-SCRESIGNIFICANCEE-IDENTITIESPSIT
ref|YP_001114460.1| -----
ref|YP_001111684.1| -----
ref|YP_001113963.1| -----
dbj|BAD18231.1| -----
RAAC02515 -----

ref|YP_001111555.1| IVESGAPSFEIEQGDEELTTHSGLALIGALLTNTKIKTRLNNTMSAEQKK-PHISNGSVG
ref|YP_001114460.1| -----FEIEQGDEELTTHSGLALIGALLTNTKIKTRLNNTMSAEQKK-PHISNGSVG
ref|YP_001111684.1| -----FEIEQGDEELTTHSGLALIGALLTNTKIKTRLNNTMSAEQKK-PHISNGSVG
ref|YP_001113963.1| -----FEIEQGDEELTTHSGLALIGALLTNTKIKTRLNNTMSAEQKK-PHISNGSVG
dbj|BAD18231.1| -----VKFILENSDDTLTTHSGLGLIGLLLSKTKFKHFRFSSLKVPEIKSSPKILNGDVT
RAAC02515 -----MRFIIIEESDEVIVTHSGMTLVGVLLDKTRIGERLNHTRLPGMGK-PNISNRDVA
* :*:.*. :.*****: *: * ** :*: : *:. . . *: * * . *

ref|YP_001111555.1| IAYIGLLCQGKSDFDHIEPFREDNFFAISLNKETPSSPTLRQLDMVAKDKQWNNILLE
ref|YP_001114460.1| IAYIGLLCQGKSDFDHIEPFREDNFFAISLNKETPSSPTLRQLDMVAKDKQW-----
ref|YP_001111684.1| IAYIGLLCQGKSDFDHIEPFREDNFFAISLNKETPSSPTLRQLDMVAKDKQWNNILLE
ref|YP_001113963.1| IAYIGLLCQGKSDFDHIEPFREDNFFAISLNKETPSSPTLRQLDMVAKDKQW-----
dbj|BAD18231.1| TSYVGLLCQGKNDFDHIEQFRDDPFFYRALDIKLPSSPTLRQLDQIAKVTGW-----
RAAC02515 YSYIGLLCQGKTDQFDHIEAFRDDEFTIALQVDNVPSSPTLRQLDMVAGKSGWESILRE
*:*****.***** **: * ** :*: : .***** :* . *

ref|YP_001111555.1| ESAGLIKKTNAPLTPVYLGO-----
ref|YP_001114460.1| -----
ref|YP_001111684.1| ESAGLIKKTNAPLTPVYLGLENRT--
ref|YP_001113963.1| -----
dbj|BAD18231.1| -----
RAAC02515 ESARLLRALDVTLHPIELGEPART
```


FIG. 109

```

ref|YP_001212947.1| -----IKYCLSQWDKLEAFLODGRLELDNNRSERSIK
ref|YP_431166.1| -----INYCLGQWDKLTAFLODGRLELDNNRSERSIK
ref|YP_754943.1| -----ITYCLNQWDALNTFLLDGRLEIDNNRAERSIK
ref|YP_754865.1| ANCEE-IDENTITIESPSITIVESGAPSITYCLNQWDCINSFLLDGRLEIDNNRAERSIK
RAAC02530 -----MSYCLKQWPKLVRYMEDGHLEIDNNRCERSLK
ref|YP_001318017.1| -----ISYSLKQWDKLAAFMKDGRIADNNLAERGIK
                        :.*.* ** * : : **:: :*** .**.:*

ref|YP_001212947.1| PFVIGRKNWLFANTPRGARASAIIVYSIVETAKENGLNPFHYLSYLFKLPNLDTKDENAL
ref|YP_431166.1| PFVIGRKNWLFANTPRGAKASAITYSIIETAKDNGLNPFQYLIYLFERLPNLDLKDGDAL
ref|YP_754943.1| PFVIGRKNFLFSNTPRGARGSAIIYSIIETAKENNLKPYDYLVLFEQLPNVDTSDQTAV
ref|YP_754865.1| PFVIGRKNWLFNTNTPRGARGSAIIYSVIETAKENNLKPYNYMFYLFQQLPNVDTGDQAAI
RAAC02530 PFVIGRKNWLFANTPRGARASAIAYSIVETAKENGLNPFAYLEYLFKLPNMDTDDKTAM
ref|YP_001318017.1| PFVLGRKNYLFKSPKGATASALCYSIETAKANKLIPFQYLYLFQQLPNLDIEDPEAL
***:****:***::*:** .**: **:**** * * *: *: ****:****:* * *:

ref|YP_001212947.1| DKLLPWSDSLPSVCRVVK
ref|YP_431166.1| DQLLPWSASLP-----
ref|YP_754943.1| DRIMPWSDTLPEGCR---
ref|YP_754865.1| DRLLPWSDTLP-----
RAAC02530 AALLPWSETLPAHIRRRK
ref|YP_001318017.1| DAMLPWAESLPNEVRHK-
                        :*: :**

```

FIG. 110

```
ref|YP_001036724.1| TRANSPSASEMUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPSASE
ref|YP_001039064.1| -----
ref|YP_001039349.1| -----
RAAC02533 -----
ref|YP_076118.1| -----
ref|YP_076073.1| -----
```

```
ref|YP_001036724.1| MUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPSASEMUTATRTYPE
ref|YP_001039064.1| -----
ref|YP_001039349.1| -----
RAAC02533 -----
ref|YP_076118.1| -----
ref|YP_076073.1| -----
```

```
ref|YP_001036724.1| CLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPSASEMUTATRTYPECLSTRIDIUM
ref|YP_001039064.1| -----
ref|YP_001039349.1| -----
RAAC02533 -----
ref|YP_076118.1| -----
ref|YP_076073.1| -----
```

```
ref|YP_001036724.1| THERMCELLUMATCCGBABNTRANSPSASEMUTATRTYPECLSTRIDIUMTHERMCELLU
ref|YP_001039064.1| -----
ref|YP_001039349.1| -----
RAAC02533 -----
ref|YP_076118.1| -----
ref|YP_076073.1| -----
```

```
ref|YP_001036724.1| MATCCGBABNTRANSPSASEMUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABN
ref|YP_001039064.1| -----
ref|YP_001039349.1| -----
RAAC02533 -----
ref|YP_076118.1| -----
ref|YP_076073.1| -----
```

```
ref|YP_001036724.1| TRANSPSASEMUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPSASE
ref|YP_001039064.1| -----
ref|YP_001039349.1| -----
RAAC02533 -----
ref|YP_076118.1| -----
ref|YP_076073.1| -----
```

```
ref|YP_001036724.1| MUTATRTYPECLSTRIDIUMTHERMCELLUMATCCSCRESIGNIFICANCEE-IDENTIT
ref|YP_001039064.1| -----
ref|YP_001039349.1| -----
RAAC02533 -----
ref|YP_076118.1| -----
ref|YP_076073.1| -----
```

```
ref|YP_001036724.1| IESPSITIVESGAPSLDEFVVEEFKNRRLEGEYPYLWLDATFPKVREGGRVCSMALVIAVG
ref|YP_001039064.1| -----LDEFVVEEFKNRRLEGEYPYLWLDATFPKVREGGRVCSMALVIAVG
ref|YP_001039349.1| -----LDEFVVEEFKNRRLEGEYPYLWLDATFPKVREGGRVCSMALVIAVG
RAAC02533 -----MDETQQQFKERPLEREYPYVWLDATFPKVREGGRVQSMALVIAIG
ref|YP_076118.1| -----LDEVVVEAFRNRPLEGRYPYVWLDKYEKVRENGRVSSMALVIAMG
ref|YP_076073.1| -----LDEVVQAFRNRPLEGRYPYVWLDKYEKVRENGRVSSMALVVAVG
: ** *: *: * ** .***:****.: ****.*** *****: **
```

VNQQGEREILGFDVGMSEDGAFWEEFLRRLVARGLGKGVRLVISAHEGLKAAIKKILTGS
VNQQGEREILGFDVGMSEDGAFWEEFLRRLVARGLGKGVRLVISAHEGLKAAIKKILTGS
VNQQGEREILGFDVGMSEDGAFWEEFLRRLVARGLGKGVRLVISAHEGLKAAIKKILTGS
VSDTGEREVLGFDVGTSEDGAFWTDFLRDLKTRGLRGVRLVVSADHAGLRQAI AEVLTTGA
VREDGEREILGLDVGPSDEGAFWTAFLRQLVARGLGKGVLLVSDNHVGLREAI RTVFSGA
VREDGDREILGLDVGPSDEGAFWTAFLRQLVARGLGKGVLLAISDSHVGLQEAI RTVLSGA
* : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *

```
ref|YP_001036724.1|
ref|YP_001039064.1|
ref|YP_001039349.1|
RAAC02533
ref|YP_076118.1|
ref|YP_076073.1|
```

```
AWQRCRVHFMRNVLSQVPKHYQG MVSSI IRTIFA QNDQESAREQLRHVVDELKNRFPKAM
AWQRCRVHFMRNVLSQVPKHYQG MVSSI IRTIFA QNDQESAREQLRHVVDELKNRFPKAM
AWQRCRVHFMRNVLSQVPKHYQG MVSSI IRTIFA QNDQESAREQLRHVVDELKKRFPKAM
TWQRCKVHAIRNVLSQVPKKEQPMMASILRTIFTQPTQDAAREQLRRVVGEFRRRYPKAM
SWQRCRVHFMNRLLGYVPKNLQSMVA AVRTI FAQPDQQA AKS QLAVVVENLRKF PRAA
SWQRCRVHFMRNLLSYVPKHWSMVAA VRTI FAQPDQQAARRQLAVVDNLRPQF PRAA
****:* :*:.* **** * *: : ***** *: **:** * . : : : ** *
```

```
ref|YP_001036724.1|
ref|YP_001039064.1|
ref|YP_001039349.1|
RAAC02533
ref|YP_076118.1|
ref|YP_076073.1|
```

```
KILEEAAAAIILAYMAFFPREHWAQIHSTNPLERLNREIRRRTDVVCIFPNREAVIRLVGAM
KILEEAAAAIILAYMAFFPREHWAQIHSTNPLERLNREIRRRTDVVCIFPNRKAVIRLVGAM
KILEEAAAAIILAYMAFFPREHWAQIHSTNPLERLNREIRRRTDVVCIFPNREAVIRLVGAM
AIIADAEEDVLAFMALPFEHWRQICSTNPLERLNREMRRRMNVVGI FPNRDSVVRLVGAI
QLLEDAEEDI LAYMAFFT EHWRR LHSTNPLERLNREIGR RTEIVGI FPNREALIRLAGAV
QLLEAAEDDILAYMAFFT EHWRRQHSTNPLERLNREIGR RTDVVGI FPNREAVIRLAGAV
*: *: *: *: *: *: *: *: *: *: *: *: *: *: *: *: *: *: *: *: *: *: *: *
```

```
ref|YP_001036724.1|
ref|YP_001039064.1|
ref|YP_001039349.1|
RAAC02533
ref|YP_076118.1|
ref|YP_076073.1|
```

```
LMEQNDEWKVG--RRYFSLESMSKITSINEFTLTPVALLHK
LMEQNDEWKVG--RRYFSLESMSKITSINEFTLTPVALLHK
LMEQNDEWKVG--RRYFSLESMSKITSINEFTLTPVALLHK
LQEQHEEWLVLS--RRYFSLESMAKLKPQ-QPMLSAQALLQK
MIEQQEEDWMTAPRRYFSQASMAKLYAH-DPSLGRPELL--
LIEQQDEWTAAPRRYFSQASMAKLYTS-NQSVSGTDFL--
*:***:**:*****:***:
*:***:**:*****:***:
```

FIG. 111

```
ref|YP_076118.1| -----
ref|YP_074958.1| -----
ref|YP_001036724.1| TRANSPTSASEMUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPTSASE
ref|YP_001039064.1| -----
ref|YP_001039349.1| -----
RAAC02534 -----
```

```
ref|YP_076118.1| -----
ref|YP_074958.1| -----
ref|YP_001036724.1| MUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPTSASEMUTATRTYPE
ref|YP_001039064.1| -----
ref|YP_001039349.1| -----
RAAC02534 -----
```

```
ref|YP_076118.1| -----
ref|YP_074958.1| -----
ref|YP_001036724.1| CLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPTSASEMUTATRTYPECLSTRIDIUM
ref|YP_001039064.1| -----
ref|YP_001039349.1| -----
RAAC02534 -----
```

```
ref|YP_076118.1| -----
ref|YP_074958.1| -----
ref|YP_001036724.1| THERMCELLUMATCCGBABNTRANSPTSASEMUTATRTYPECLSTRIDIUMTHERMCELLU
ref|YP_001039064.1| -----
ref|YP_001039349.1| -----
RAAC02534 -----
```

```
ref|YP_076118.1| -----
ref|YP_074958.1| -----
ref|YP_001036724.1| MATCCGBABNTRANSPTSASEMUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABN
ref|YP_001039064.1| -----
ref|YP_001039349.1| -----
RAAC02534 -----
```

```
ref|YP_076118.1| -----
ref|YP_074958.1| -----
ref|YP_001036724.1| TRANSPTSASEMUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPTSASE
ref|YP_001039064.1| -----
ref|YP_001039349.1| -----
RAAC02534 -----
```

```
ref|YP_076118.1| -----
ref|YP_074958.1| -----
ref|YP_001036724.1| MUTATRTYPECLSTRIDIUMTHERMCELLUMATCCSCRESIGNIFICANCEE-IDENTIT
ref|YP_001039064.1| -----
ref|YP_001039349.1| -----
RAAC02534 -----
```

```
ref|YP_076118.1| -----LEELLRKSGVDDVDFLREGVRVLAQGGLMELEVSQQI
ref|YP_074958.1| -----LEELLRKTVGDDDFLREGVRVLAQGGLMELEVSQRI
ref|YP_001036724.1| IESPSITIVESGAPSMATNNRMALLEQLSKYVVEKDKDFLKEALTLLINALMDAEVTSII
ref|YP_001039064.1| -----MATNNRMALLEQLSKYVVEKDKDFLKEALTLLINALMDAEVTSII
ref|YP_001039349.1| -----MATNNRMALLEQLSKYVVEKDKDFLKEALTLLINALMDAEVTSII
RAAC02534 -----MALLNSFAVLEWIRKMQDEHDHIDVLRLEMLQVAQFLIDAEAVEKI
** : : .. *.* : : : * : * . *
```

ref|YP_076118.1|
 ref|YP_074958.1|
 ref|YP_001036724.1|
 ref|YP_001039064.1|
 ref|YP_001039349.1|
 RAAC02534

GAERYERSSERSNYRNGYRERQWDTRVGTIDLQIPKLRKGSYMPSWLEPRRRAEKALVAV
 GADRYERSAERSTYRNGYRERQWDTRVGTIDLQIPKLRKGSYMPSFLEPRRRAEKALVAV
 GAEKYERNNNRNNYRNGYRLREWDTRVGTQLSIPKLRHGSIYFPSLLEPRKMSEKALLNV
 GAEKYERNNNRNNYRNGYRLREWDTRVGTQLSIPKLRHGSIYFPSLLEPRKMSEKALLNV
 GAEKYERNNNRNNYRNGYRLREWDTRVGTQLSIPKLRHGSIYFPSLLEPRKMSEKALLNV
 GAERYERTDSRVTRQNGSRSTWDTLGTVDLKIPLRQGSFFPSLLEPRRRAEQALAAV
 :*. * . *** * * *****:***:*.*****:***:*** *****: :*:** *

ref|YP_076118.1|
 ref|YP_074958.1|
 ref|YP_001036724.1|
 ref|YP_001039064.1|
 ref|YP_001039349.1|
 RAAC02534

VQEAYIQGVSTRKVDELVQALGMTGVSKSQV
 VLEAYVNGVSTRKVDDLQALGMTGVSKSQV
 VQEAYVHGVSTRKVDELVEALGMKGIDKSEV
 VQEAYVHGVSTRKVDELVEALGMKGIDKSEV
 VQEAYVHGVSTRKVDELVEALGMKGIDKSEV
 IQEAYVKGVSTRKVDDLVRALGLEGISKSEV
 : ***:*****:***.***: *:***:*

```
ref|YP_148733.1|      -MDMTLTAKIKIYPTAEQAEVLKATLSAYRQACNAVSVVIFDT-KVL
ref|YP_146224.1|      -----MDMTLTAKIKIYPTAEQAEVLKATLSAYRQACNAVSVVIFDT-KVL
ref|ZP_02171259.1|    -----MKLTLTAKTKIVPTVEQEDTLRKTAAHYRDACNAVSEVVFDE-NTL
ref|NP_977687.1|      -----TITAKIQIHVSDNQAESLKITTNAYRKACNWLSKHIFET-KNL
RAAC02562             MSAIPPPSEGEVNGMELTRTVRVRLDPTPEQASALALTI EANRQALDYVSRIFAER-RIR
ref|ZP_02619781.1|    -----MEIVRTLKIKLNINQEEKNMNSTLEAFLKALNYASQVAFNNGEIT
                        . * : :: . :: . : * * .* : * *: .

ref|YP_148733.1|      AQAKLHDMTYRLLRSNYALRSQMAQSVIKTVIARYRSLKSNGHEWTLVRFKKPEYDLVWN
ref|YP_146224.1|      AQAKLHDMTYRLLRSNYALRSQMAQSVIKTVIARYRSLKSNGHEWTLVRFKKPEYDLVWN
ref|ZP_02171259.1|    VQAKLHKRTYRELRSFTFGLKLSQMAQSALKTVIATYKTNQSNGHERSQVFKKPFYDLVFN
ref|NP_977687.1|      NQVNLNNLYYSDLRNQFRLKSQMAQSVMKTVIARYKSAKSNGHEWSLIDFKLAEYDLVWN
RAAC02562             NAFDLQRIVYRTIRTRFGLRAQMACTVCRTVAAVYKSMKSNGNANALAKFKHAKPVFQWN
ref|ZP_02619781.1|    NKPKLQKLVDLRLTKFNLKSQMAVNTCTTVCSSYVTQHSNKVFNSLAVYKSPKAIYSYG
                        .*: * :*. : :*:*** .. ** : * : :** : : * :. : .

ref|YP_148733.1|      RDYSIVQG--LFSVNTLEGRIKVPFEPKG-MEPYFDGSWTFGTAKLVYKHKNFFLHIPMT
ref|YP_146224.1|      RDYSIVQG--LFSVNTLEGRIKVSFEPKG-MEPYFDGSWTFGTAKLVYKHKNFFLHIPMT
ref|ZP_02171259.1|    RDYSLTKG--LFSVNTLEGRIKVPFYTEG-LEPYFDGTWTFGTAKLVHKYKWFLLHIPVS
ref|NP_977687.1|      RDYSLTKN--QFSVNTLEGRRLKNYERKA-MKKYFNSTWKFGTAKLVHKYKWFLLHIPMT
RAAC02562             KDYVLRSG--LAYITTL DGRLLRI PFRVEPPYQRYLQDGTWTFGAAELVQKRHWGLHVSVS
ref|ZP_02619781.1|    RDYSFLNDGQTISINTIAKRIKVFHKVNNYFKKYLTKEWSFGSLEIVERENSYYAHITVS
                        :** : .. :*: **: : : : : * : *.*: **: : : :*:...

ref|YP_148733.1|      KTIPTVDEHNIRQVVGVDVGVNFLAAAYDSQGKTTFFNGRKIKHMRACYKRMKRTLQQKG
ref|YP_146224.1|      KTIPTVDEHNIRQVVGVDVGVNFLAVAYDSQGKTIFFNGRKIKHMRACYKRMKRTLQQKG
ref|ZP_02171259.1|    KDVQEANLDNINQVVGIDMGVNFTATTYDSNGQTRFFNGKQIKHKRACYKQMRKELQQKQ
ref|NP_977687.1|      KEYQTLDFADVNNIVGVDLGINFLATTYDSQSKTTFFYNGNIVKHKRGKFKATRKLQLQTRQ
RAAC02562             KTVDPDP-VGDFDAVIGVDQGMRLVTAS-CGNEVMFIRGGKVKQTRLRYVRLRASLQRKG
ref|ZP_02619781.1|    KEVQEKPLNEFKNIIGIDLGNFIATFYDSKGKIKFFKGRYLKDMRAKYTRLRKLQLRKG
                        * .. :*:** * * .. :.. * .* :*. * :. * ** :

ref|YP_148733.1|      TASARRKLKTIGQRENRMWTDVNHAVTKALV---RQYGERTLFVLEDLTGIREKTERVRI
ref|YP_146224.1|      TASARRKLKTIGQRENRMWTDVNHAVTKALV---RQYGERTLFVLEDLTGIREKTERVRI
ref|ZP_02171259.1|    TPSARRKLKKIGQRENRMWQNNINCYISKALV---EQYKDTLFVMEDLTGVRNATEKVRV
ref|NP_977687.1|      TPSSRRKKIQIGSRENRYVDVNHQITKALV---EAYPRGTMFVLEDLTGVSRAKTEKVRV
RAAC02562             TRSAKRRLKAIGRRSRFMTDVNHQIAKAVVRFAKAQGQRPLIVLEDLTGS-NLSVRFRM
ref|ZP_02619781.1|    TYNAHKKIIVTINQREQRTMTYINHKKISKEIV--EYAKQNNAIAMENLTGI-NLSCVKVK
                        * ..::: * . **.* : :*: :*: * . ..:::*** . : :.

ref|YP_148733.1|      HDRYETVSWAFYQFRQLLEYKARLHGSKVMVAVPHYTSLTCPKCGHTEKANRNKRTHTF
ref|YP_146224.1|      HDRYETVSWAFYQFRQMLEYKARLHGSKVIVAVPHYTSLTCPKCGHTEKANRKRTHTF
ref|ZP_02171259.1|    KDRYQTVSWSFNDLRQKIA YKAQKTGALGIAVDPGYTSQTCPKCGHTEKANRNKKTHTFC
ref|NP_977687.1|      KNRYVSVSWAFYQFRQMLEYKAEELNGQKVIVVDPKYTSQTCPKCGNIEKANRNKKLHTFK
RAAC02562             KDQYWRMSWAFRQLADFIYKAEVGIAMVIDPTGTSETCPKCGHCEAANRSRKRHEFR
ref|ZP_02619781.1|    DNRYRYKVSFAFNQLQQFIEYKAVQAGLKVIYINPKYTSQTCPICGHIHKDNRNKKLHIFK
                        .:.* :***: : : : *** * : : * ** *** **: . **: : * *

ref|YP_148733.1|      CRTCGYTSNDDRIGAMNLQRKGIE-----
ref|YP_146224.1|      CRTCGYTSNDDRIGAMNLQRKGIE-----
ref|ZP_02171259.1|    CQTCQYTSNDDRIGAMNLQRKGI-----
ref|NP_977687.1|      CKNCQYQSNDDRISAMNLRKGIK-----
RAAC02562             CKKCGYRCNDDLAASRVIAQKGLECLRQSQSA
ref|ZP_02619781.1|    CRACGCTLNDDLIGAKNIQHKGYD-----
                        * : * *** : : : : * : : * **
```

FIG. 113

```

ref|YP_516922.1|      ---MFARIKTAYNRDGSPPRYLQLVESRREEGKVRQKVLGNLGRVEDLQNGKLDDDLIRSL
ref|YP_520815.1|      -----
ref|YP_519534.1|      ---MFARIKTAYNRDGSPPRYLQLVESRREEGKVRQKVLGNLGRVEDLQNGKLDDDLIRSL
ref|ZP_01372264.1|    ---MFARIKTAYNRDGSPPRYLQLVESRREEGKVRQKVLGNLGRVEDLQNGKLDDDLIRSL
ref|ZP_01370818.1|    ---MFARIKTAYNRDGSPPRYLQLVESRREEGKVRQKVLGNLGRVEDLQNGKLDDDLIRSL
RAAC03229             MILLFAQIVSTKKPDGKTYKYLHIVESYREGRTVKKRRVASLGNISQYSEREIEQIIRTL

ref|YP_516922.1|      AKFSDKLAVVDAEDLFADWSKEYGPSMVFNRLWDNLGLHKILDGLFSEERELSIDVREAI
ref|YP_520815.1|      -----LWDNLGLHKILDGLFSEERELSIDVREAI
ref|YP_519534.1|      AKFSDTLAVVDAEDLFADWSKEFGPSMVFRRLWENLGLHKILSGLFNERELSIDVQEI
ref|ZP_01372264.1|    AKFSDTLAVVDAEDLFADWSKEFGPSMVFRRLWENLGLHKIFAGLFNERDLSIDVQEI
ref|ZP_01370818.1|    AKFSDTLAVVDAEDLFADWSKEFGPSMVFRRLWENLGLHTILDGLFNERDLSIDVQEI
RAAC03229             ESLLQHRTTG-SLEDFAEQQVLHFGVPYVQFLWNQLGLTEAIRDALRAREVTFDVARVY
                        *:*** : . : *::** . :

ref|YP_516922.1|      FCMVLNRLTEPTSKLGVSDWKDSVYRPDF--ETLKLHHFYKAIDFLDENKDTIEEQLFH
ref|YP_520815.1|      FWMVLNRLTEPTSKLGVSDWKDSVYRPDF--ETLKLHHFYKAIDFLDENKDTIEEQLFH
ref|YP_519534.1|      FCMVLNRLTEPTSKLGVSDWKDSVYRPEF--ESLKLHHFYKAIDFLDENKESIEEQLFH
ref|ZP_01372264.1|    FCMVLNRLTEPTSKLGVSDWKDSVYRPEF--ESLKLHHFYRAIDFLDENKDTLEEQLFFH
ref|ZP_01370818.1|    FCMVLNRLTEPTSKLGVSDWKDSVYRPEF--ESLKLHHFYKAIDFLDENKDTLEEQLFFH
RAAC03229             QAMVIHRLVDPSSKLRLEFHTLDDLYLPDWGGEPWQLQHFYRALDYLDIKPQLERVLYAR
                        *:***:*** : . *.***: *. :*:***:***: * :* :*. *: :

ref|YP_516922.1|      HTNLFTQQQLDLVFFDFTTSTYVEGDAGAFDLLEYGHSKDHDPDRQLQVMIGILMSRDGIPIA
ref|YP_520815.1|      HTNLFTQQQLDLVFFDFTTSTYVEGDAGAFDLLEYGHSKDHDPDRQLQVMIGILMSRDGIPIA
ref|YP_519534.1|      QTNLFTQQQLDLVFFDFTTSTYVEGDAGAFDLLEYGHSKDHDPDRQLQVMIGILMSRDGIPIA
ref|ZP_01372264.1|    HTNLFTQQQLDLVFFDFTTSTYVQGDAGAFDLLEYGHSKDHDPDRQLQVMIGILMSRDGIPIA
ref|ZP_01370818.1|    HTNLFTQQQLDLVFFDFTTSTYVEGDAGAFDLLEYGHSKDHDPDRQLQVMIGILMSRDGIPIA
RAAC03229             LTDLLNFRLSLVLYDLTSTHLHG--HACPLGGHGYSRTHRPDLQVELGLLVTPEGIPIT
                        *:*. :*.***:*.***:*. * * :*:*. **** * *:***: :****:

ref|YP_516922.1|      HHVFPGNTPDPTDAFIEAVSDLKKRFNIQRVIVVGDGRGMMGKRTLELLEELQLQYILGVRM
ref|YP_520815.1|      HHVFPGNTPDPTDAFIEAVSDLKKRFNIQRVIVVGDGRGMMGKRTLELLEELQLQYILGVRM
ref|YP_519534.1|      HHVFPGNTPDPTDAFIEAVRDLKKRFTIQRVIVVGDGRGMMGKRTLELLEELQLHYILGVRM
ref|ZP_01372264.1|    HHVFPGNTPDPTDAFIEAVSDLKKRFTIQRVIVVGDGRGMMGKRTLELLEELQLHYILGVRM
ref|ZP_01370818.1|    HHVFPGNTPDPTDAFIEAVSDLKKRFTIQRVIVVGDGRGMMGKRTLELLEELQLHYILGVRM
RAAC03229             HEVFAGNVSDKQTPVDILKRLKEDFAVEQCVFVGDGRGMVTEKNMALMAEAGFPYIVGFHK
                        *.**.*.***: : : ***: * : : :*****: : : : * : * : ***:

ref|YP_516922.1|      RN---IKAGPDLANSPEPYAFIKDNLKVKEVVHQ-----EKRYIVCLNEEEAKRDQ
ref|YP_520815.1|      RN---IKAGPDLANSPEPYAFIKDNLKVKEVVHQ-----EKRYIVCLNEEEAKRDQ
ref|YP_519534.1|      RN---IKAGPELAASPEPYTFLKDNLKVKEVLHQ-----EKRYIVCLNEEEAKRDQ
ref|ZP_01372264.1|    RN---IKAGPELAASAQPYPFVKDNLKVKEVLHQ-----EKRYIVCLNEEEAKRDQ
ref|ZP_01370818.1|    RN---VKAGPELATSPEPYVFTKDNLKVKEVLHQ-----GKRYIVCLNEEEAKRDQ
RAAC03229             RGRIVSDALLEQFADVNAYHELKDNRLYLEVPAASVDDVEKAEGVRYILCYNPEKARQDA
                        *. * : . :.* ****: ** ****:* * ***:

ref|YP_516922.1|      LVREQIEIKLRS-----KLEHG-----SIKDLIGNSEYKKYINVTAEAAAT
ref|YP_520815.1|      LVREQIEIKLRS-----KLEHG-----SIKDLIGNSEYKKYINVTAEAAAT
ref|YP_519534.1|      LVREQIEMKLRS-----KLEHG-----SIKDLIGHSEYKKYINVTAEAAAT
ref|ZP_01372264.1|    LVREHIEMKLRS-----KLEHG-----GIKDLIGHSEYKKYINVTAEAAAT
ref|ZP_01370818.1|    WVREQIEVKLRS-----KLEHG-----SIKDLIGHSEYKKYLNVSAAEAAAT
RAAC03229             AFRESALEEAAETGLKALAESLAKPKRGRKPTDKGVMLKVADLLTRKGVEAFFQVDYKDG
                        .** : . : * : * : ***: . : : :* : .

ref|YP_516922.1|      IN----TDKCLKQAADFVGLYILQTNTELPTEEVATAYRDLWQIERAFRNKSTLDRPVY
ref|YP_520815.1|      IN----TDKCLKQAADFVGLYILQTNTELPTEEVATAYRDLWQIERAFRNKSTLDRPVY
ref|YP_519534.1|      MN----TDKCLKQAADFVGLYILQTNTELPTEEVATAYRDLWQIERAFRNKSTLDRPVY
ref|ZP_01372264.1|    IN----TDKCLKQAADFVGLYILQTNTELPTEEVATAYRDLWQIERAFRNKSTLDRPVY
ref|ZP_01370818.1|    IN----TDKCLKQAADFVGLYILQTNTELPTEEVATAYRDLWQIERAFRNKSTLDRPVY
RAAC03229             LTYRDEDAITKEALRDGKFLIRTNLPAADVQSYKTLMGIERAFHQIKNFLDVGPYIY
                        :. * : : * : ** : : :***:***: :*. :* : *****: :*. ** :*

```



```

-----VTAVKGI GEETAALADLGITTVGELLNYPYRYDDYEQKDLAAVRHEEKVTV
-----VTAVKGI GEETAALADIGIATVGELLAYAPYRYDDYEQKDLAAVRHEEKVTV
MTLEQYSVRQIKGVSAPKEELHAFGVHTVADLLDYFFRYEDYRIRELA EIKDGEKVTV
MDLNQIPVTRVSGVGPQKAEDL AALGIHTAAQLIAYFFRYEDYRLRDLSEVKDGERITV
-----VSVLPGVGPKRAENLQELGIATIEDLLTYYPFYDDIQEKDLSEIQDEKVTL
MSLRSLSVRALPGVGPQKERAELGIRTVDDLLHTYFPRYDERAEKFPSEWRDGRVTA
      *  :  *      *  :  *  :  *  :  *  :  *  :  *  :  *

```

EGKVHSAPLLTTYGKKKSRLSFRLLTG-RYLITVVCFNRPYLGKGIITLNETVTTVIGKWDR
EGKVYSAPLLTTYGKKKSRLSFRMLAD-RYLITVVCFNRPYLKEKIALNETVTTVIGKWDR
QGTIRNGILQRYGKNKSRLTCKIEVD-HMLVTAVWFNHRFLQGQLTPGREIMLTGKWEQ
QGNLIGAPVLQRYGR-KSRLSCKVVID-HFFVTAVWFNRPFCLKDKLASGEIIRLTGKWDA
KGLVVSEAVVSRYGYKKSRLTFRMMQE-HAVINVSFFNQFFLKDKVVLSEEIAVYGKWDA
RAVVEG-PVQVRWRGSKSIMTARVRVDGQHFPVCLVESQHYLRSLSDGRFIVVTGKWNE
.....**:::..*::::***:

[illegible]

PALRRAYRLIDKQEA VRALHFFPRS REELHQARRRLVYEEFLLYQLKM QALRKVMRDERRG
PALRRAYRLVDKQE ALRALHFFPRTREELHQARRRLI YEEFLLYQLKM QALRRLMRDERRG
HELV ERHGLMARDA VQR IHDPEEVKEGLEARRRMVYEEL FLFQLKLQAYS LTRKRGDG
AD FVKRYQLLPRKQ AVALIHQPS GLEDGKQARRRMVYEELFFFQLKM HAFKAITRKRADG
NDLLEKYRLMPR KEAMWAMHFSPN PESHQAKRRVVFE EFLFLQLKM OGLKKQEAKENG
YALVRKYRLWTH RDA LFGMHRP SLEDVRQARRRVFE EFLFLQIQ LQWLRAK-REEPAG

: : * : : * : : * : : * : : * : : *

I I H S F P E E Q L A S F L S G L P F S L T N A Q R R V I R E I L D D M R A P Q M N R L L Q G D V G S G K T V V A A V
V V H S F S E E R L S S F L S G L P F V L T N A Q R R V I G E I L A D M R S P Q M N R L L Q G D V G S G K T V V A A V
I V H Q V N G E T I R N F A A T L P F E L T D G Q K K V V N E I M S D M R Q P A A M N R L L Q G D V G S G K T V V A A I
L A Q Q V D L P K I R S F V K S L P F E L T P S Q K Q V V G E I L H D M Q P Y T M N R L L K G D V G A G K T V V A A T
L A I Q Y D V D R L K T F T Q G L P F E L T G A Q K K V T N E I C R D L R S P K H M Q R L L Q G D V G S G K T V V A A I
R A Q F V P S D A L T A F E A L L P G P M T N A Q R R A C E D I L R D Q R F V P M T R L I Q G D V G S G K T W V A L F

: : * * * * * * * * *

ALYAAVLSGFQGALMVPTEILAEQHVRSLAELFADTGVTVELLTSSVKGKRRKELLAKLE
ALYAAALSGFQGALMVPTEILAEQHARSLAELFADTDVTIALLTSSVKGKQRKAVLAELE
ALYCTIKAGHQGALMVPTEILAEQHLRSLQKLFADTGIEVALLTGSLTEKKRRDVLAGLQ
ALYATVTAGCQGALMVPTEILAEQHKKLSLRMFKPYGIETALLTGSSTDKKREILAGVQ
ALYATMTAGFQGALMVPTEILAQQHMESLQQLFDPLEVRTALLTGSTKTERRLILEELA
ACFAVHLARGQSALMAPTEILAEQHARLAHELLGSAGVRVELLTGSVTGRDRVLAGLA
* : : : * : : : * : : : * : : : * : : : * : : : * : : : * : : :

[illegible]

```
RTLAITAFGDM DVSVLDEMPAGRKKVETYWVKHNQFARVLDFIEKELRRGHQAYVICPLI
RTLAITAFGDM DVSVLDEMPAGRKKVETYWVKHHQFSRVLDFIEKELRRGHQAYVICPLI
RTLAITAFGDM DVSTIKERPHGRPKIKTYWVKHDDMMERV LGFIRREVGEGRQAYVICPLI
RTLAITAFGD LDVSTIREMPKGKRPIKTYWVHHDDL ERLV LFGFIQKEAAGRQAYVICPLI
RTLAITAYGMD VSIIDEMPAGRIPIETRWIRPPQLD TVLEWMEKELARGHQAYIIICPLI
RTLALAIYGMD VSILNELPKGRKPVTIAVP SKDD ET VLRLIRRELARGHQAYIVAPAI
***** : * : * : * : * : * : * : * : * : *
```

ref|YP_001125159.1|
ref|YP_147039.1|
ref|ZP_02849289.1|
ref|ZP_02329219.1|
ref|NP_816723.1|
RAAC00160

EESEKLDVQNAIDVHSQLVYYRGKYEVGLMHGRLSADEKEAVMRAFSENRIHVLVSTTV
EESDKLDVQNAIDVHSQLVHYYRGKYEIGLMHGRLSADEKERVMAFSENRIHVLVSTTV
EESDKLDVQNAIDLVYQMQQAFP-DLKVGLLHGRLSASEKDEVMRGFGANETQLLVATTV
EESDKLDVQNAIDVHVQLQQHFP-DLNIGLLHGRLTAAEKDDTMRSPKDGSIQVLVSTTV
EESEALDVKNATEIFEHMQSFYSPRYQVGLLHGKMKNQEKDDIMQEFKDNQLQLLVSTTV
EASERDDVASVTELYERVREHLA-GFRVELLHGRLMPADSADKERMMAFRDGDHALVATTV
* *: ** .. :: :. :****: **: * : **:**

ref|YP_001125159.1|
ref|YP_147039.1|
ref|ZP_02849289.1|
ref|ZP_02329219.1|
ref|NP_816723.1|
RAAC00160

VEVGVNVPNATVMVIYDAERFGLAQLHQLRGRVGRGDAQSYCILIADPKSEIGKERMHIM
VEVGVNVPNATVMVIYDAERFGLAQLHQLRGRVGRGDAQSYCILIADPKSEVGKERMIM
VEVGVDVPNATLIIIMDAERFGLSGLHQLRGRVGRGAHQSFVVLVADPKSENGRERMKVM
IEVGVDVPNATLMMVYDADRFGLSQLHQLRGRVGRGEHQSFVLIADPKTEVGKERMQAM
IEVGVNVPNATVMLIMDADRFGLSQLHQLRGRVGRGSSASCYILVANPKNEMGVERMKIM
IEVGIDVPNATVMAIYGAERFGLAQLHQLRGRVGRGPHPSYCLLIHDASSEARARIETM
:***:*****: : .:*****:***** ***: :...* . *: *

ref|YP_001125159.1|
ref|YP_147039.1|
ref|ZP_02849289.1|
ref|ZP_02329219.1|
ref|NP_816723.1|
RAAC00160

TETTDGFFVLAEKDLELRGPGDFFGKQSGLPFQFGDPVHDYRILEVARRDAAKLVSSAA
TETADGFFVLAEKDLELRGPGDFFGKQSGLPFQFGDPVHDYRILEVARRDAAKLVSSAA
TETNDGFEVSRDLELRGPGDFFGKQSGLPFKIADMAAEYEMLELARDAAELTGRDD
TETTDGFEIARDLELRGPGDFFGKQSGLPFRIADLMCDFEIMEIARDDAADLVAKPE
TETTDGFFVLSERDLELRGPGEVFGARQSGVPQFAVGDIIVTDFNILEVARQEASALWKVKE
LQTDGFEIAERDLELRGPGELFGLRQSGLPFALGDLARDYRIMEVAREEALALLRRDD
:* **: :.:*****:.* :***:* * . * :.:*:** :* *

ref|YP_001125159.1|
ref|YP_147039.1|
ref|ZP_02849289.1|
ref|ZP_02329219.1|
ref|NP_816723.1|
RAAC00160

FWRDEAYAGLRAELE-----
FWRDEAYAGLRAELE-----
FWTNPLFERMR-----
FWTAAEYVPLRQFLQK-----
WWQYPAYQGLANRVKPDQDEAAQFFD
FWYAPWAEGLRNALKEAMDKVSYRD
:* :

FIG. 115

```

ref|YP_148969.1| -----YLQLAHNEWDPKAKYAKAKVIYSFGREDEVDRAVLERLAKSIS
ref|YP_001126171.1| -----YLQLAHNEWDPKAKYAKAKVIYSFGREDEVDRAVLERLAKSIS
ref|YP_146154.1| -----
ref|YP_146741.1| MYIRRVTRKNKDGTTVAYLQLAHNEWDPKAKYAKAKVIYSFGREDEVDRAVLERLAKSIS
RAAC03182 MYIRVIRRNKNGSVTGYVQLAHNFRDPNTGQPKAKVLYTFGREDEIDLEALRRLAQSIH
ref|ZP_02130848.1| MYIRTISRKNKDGSKVEYVQLAHNYRDPKSKQARAEVLYSFGRKDQLDMEAIRRLAKSVE

ref|YP_148969.1| RFLSPEQAWIEITLTGEVSDDFQFQSSKRLGGAWLLDQLWRQLGLGEILHSLFASRHHQI
ref|YP_001126171.1| RFLSPEQAWIEITLTGEVSDDFQFQSSKRLGGAWLLDQLWRQLGLGEILHSLFASRHHQI
ref|YP_146154.1| -----VCSRHHQI
ref|YP_146741.1| RFLSPEQAWIEKLTGEASDDFQFQSCCKHLGGVWLLDQLWRQLGLGEILHSLFSTRHHQI
RAAC03182 RFVGDEFTAGRGQSEGIQT---TLSDSRPMGGAYLLDELWRQLGLDEVLRERLADRKFKA
ref|ZP_02130848.1| RFLAKTGDVETQCKLQFPGEDVRFVESRPMGGVFLKKIWDRLRISECIDKALADRYTA
. . * . .

ref|YP_148969.1| PLERLIFAMVANRALHPSSKLAMEEWVEKDVHIPHLPOVASHQLYRAMDELLAVQPELER
ref|YP_001126171.1| PLERLIFAMVANRALHPSSKLAMEEWVEKDVHIPHLPOGASHQLYRAMDELLAVQPELER
ref|YP_146154.1| PLERLIFAMVANRALHPSSKLAMEEWVEKDVHIPHLPOVASHQLYRAMDELLAVQPELER
ref|YP_146741.1| SLERLIFAMVANRALHPSSKLAMEEWVEKDVYIIPHLPOVASHQLYRAMDELLAVQPELER
RAAC03182 AVERVIFAMVANRALAPSSKLAMEEWDREVALPGMTELDQVQAYRAMDFLHDVAEDLQY
ref|ZP_02130848.1| PIGDAVFAMVANRALAPDSKLAVEDWAAKDVHLELDQPLKVQHLRYRAMDFLENQEIQK
. : : ***** * . ***** : * : : : ***** * : :

ref|YP_148969.1| QVFHAVADLLNLEVDLIYFDTTSSYFEVDPSETPEGESLRKQGFSGDKRDLVQIVIGLA
ref|YP_001126171.1| QVFHAVADLLNLEVDLIYFDTTSSYFEVDPSETPEGESLRKQGFSGDKRDLVQIVIGLA
ref|YP_146154.1| QVFHAVADLLNLEVDLIYFDTTSSYFEVDPSETPEGESLRKQGFSGDKRDLVQIVIGLA
ref|YP_146741.1| QVFHAVADLLNLEVDLIYFDTTSSYFEVDPSETPEGESLRKQGFSGDKRDLVQIVIGLA
RAAC03182 EVFRRVSDLLNLDVLLFFDFTTSTYFET---EDESDDSLRRKGYSKDRPDLQVQVIGLA
ref|ZP_02130848.1| EVFWSTANLLNLEVDLVFFDFTTSTYFER---DEEDEGLKRYGHSKDKRKDLQVQVVGLA
:* * . : ***** : ***** : * . : . : * . : * * * : * : * : *

ref|YP_148969.1| VTREGVPIRAWVWPNTMDMTVIKQVKQDLIGWKLGRVISVMDRGFSSEENLRILQQAGG
ref|YP_001126171.1| VTREGVPIRAWVWPNTMDMTVIKQVKQDLIGWKLGRVISVMDRGFSSEENLRILQQAGG
ref|YP_146154.1| VTREGVPIRAWVWPNTMDMTVIKQVKQDLIGWKLGRVISVMDRGFSSEENLRILQQAGG
ref|YP_146741.1| VTREGIPIRAWVWPNTMDMTVIKQVKQDLIGWKLGRVISVMDRGFSSEENLRILQQAGG
RAAC03182 VTRDGIPIRCWTWPNTADMSVVEVKQDLIGWRLGRVITVVDGRGFVSESNLRLQIRAGG
ref|ZP_02130848.1| VTKEGLPIRSWVFPNTPDVNTVEQIQKEMNDWKLGRVWVWAMDRGMTSEENRILQIRGGG
* : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *

ref|YP_148969.1| HYIVGEKMRSGKAAVEEALNRRGRYQQVRENLIKEIIVGDGEARQRYVLVYNPSEAEQ
ref|YP_001126171.1| HYIVGEKMRSGKAAVEEALNRRGRYQQVRENLIKEIIVGDGEARQRYVLVYNPSEAEQ
ref|YP_146154.1| HYIVGEKMRSGKAAVEEALNRRGRYQQVRENLIKEIIVGDGEARQRYVLVYNPSEAEQ
ref|YP_146741.1| HYIVGENMRSGKAAVEEALSRRGRYHEVDENLIKEIIVGDGEARQRYVLVYNPGEAEQ
RAAC03182 HYIAGEKMTSGKPAVEAALARPGRFRELRLPNLKVKEVVVDGEARVRYVLAFFNPPEAKRD
ref|ZP_02130848.1| NYILGEKLR-GSNMSKAVLGSPGRFTTVRDNLEIKEVTAGDGACRRRYVIVRNPKQVKRD
:* * * : * . : * * * : * : * : * : * : * : * : * : * : * : * : * : *

ref|YP_148969.1| RKEREKLLLESKEELEGLRQLPNEAHHKATCLRSHPSYGKYLRLQDKGTLRIDKQAVRE
ref|YP_001126171.1| RKEREKLLLESKEELEGLRQLPNEAHHKATCLRSHPSYGKYLRLQDKGTLRIDKQAVRE
ref|YP_146154.1| RKEREKLLLESKEELEGLRQLPNEAHHKATCLRSHPSYGKYLRLQDKGTLRIDKQAVRE
ref|YP_146741.1| RKERETLLESKEELEGLRQLPNEAHHKATCLRSHPSYGKYLRLQDKGTLRIDKQAVRD
RAAC03182 EARREAMLRELRLMELRLKELQGEAHTKAHCLASHPTFKRYLQDRWGNLRIDPEAVRQ
ref|ZP_02130848.1| QATRERLIRRAEQEIEAIGDLTGKKHTKAACALLSHRSMGKYVRELKSGKLKINKAKITE
. * * : . * : * : * : * : * : * : * : * : * : * : * : * : * : * :

ref|YP_148969.1| AEKYDGKYLIRTSDDTLAEDVALGYKQLVDIEQAFRTLKSTLELRPMYHRLEDRIRAHV
ref|YP_001126171.1| AEKYDGKYLIRTSDDTLAEDVALGYKQLVDIEQAFRTLKSTLELRPMYHRLEDRIRAHV
ref|YP_146154.1| AEKYDGKYLIRTSDDTLAEDVALGYKQLVDIEQAFRTLKSTLELRPMYHRLEDRIRAHV
ref|YP_146741.1| AEKYDGNLYIRTSDDTLAEDVAIGYKQLVDIEQAFRTLKSTLELRPMYHRLEDRIRAHV
RAAC03182 AAHLDGKYLIRTSDDTLSTEDVALGYKQLLMVESAFRTLKTTLDIRPVYHRKDERIRSHV
ref|ZP_02130848.1| EEKLDGKYLSCSDDTLSPPEIALGYKQLLEVERAFRTLKSTLDLRPVYHRKDERIRSHV
: * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *

```

```
ref|YP_148969.1|      LLSWLALLLVRIVEIRTHESWPKVR-----DECE-----
ref|YP_001126171.1|   LLSWLALLLVRIVEIRTHESWPKVR-----DECE-----
ref|YP_146154.1|      LLSWLALLLVRIVEIRTHESWPKVR-----DECE-----
ref|YP_146741.1|      LLSWLALLLVRIVEIRTHESWPKVR-----DECE-----
RAAC03182             LLCWLALLLVRIAIEVQIGRSWPDIRSHVQGNAPSDAFRRDRCAAHRNDRGAAGDPASLED
ref|ZP_02130848.1|    TLCWLALLLVRLIELETGMTWNQVR-----
                      *.*****: *:. :* .:*
```

```
ref|YP_148969.1|      -----
ref|YP_001126171.1|   -----
ref|YP_146154.1|      -----
ref|YP_146741.1|      -----
RAAC03182             QGTAEDPEGRTSNTRAVDTTTRVIHRSHESTSDMGLWDFV
ref|ZP_02130848.1|    -----
```

FIG. 116

```

ref|YP_148733.1| -----MDMTLTAKIKIYPTAEQAEVLKATLSAYRQACNAVSVVIFDTKVLA-QAKLHDM
ref|YP_146224.1| -----MDMTLTAKIKIYPTAEQAEVLKATLSAYRQACNAVSVVIFDTKVLA-QAKLHDM
ref|ZP_02171259.1| -----MKLTTLTAKTKIVPTVEQEDTLRKTAAHAYRDACNAVSEVVFDENTLV-QAKLHKR
ref|NP_977687.1| -----TITAKIQIHVSDNQAESLKITTNAYRKACNWLKSHIFETKNLN-QVNLNNL
ref|ZP_02619781.1| -----MEIVRTLKIKLNINQEEKNMINSTLEAFLKALNYASQVAFNNGEITNPKPKLQKL
RAAC03163 MREVNGMELTRTIRVRLEPTPEQASALARTVEANRQALDYVSRIAFAFERKIRN-RVALHRA
      . * : : . : . : * * . * : * * : : * :

ref|YP_148733.1| TYRLRSNYALRSQMAQSVIKTVIARYRSLKSNNGHEWTLVRFKKPEYDLVWNRDYSIVQG
ref|YP_146224.1| TYRLRSNYALRSQMAQSVIKTVIARYRSLKSNNGHEWTLVRFKKPEYDLVWNRDYSIVQG
ref|ZP_02171259.1| TYRELSTFGLKSQMAQSALKTVIATYKTNQSNNGHERSQVQFKKQFDLVFNDRDYSITKG
ref|NP_977687.1| YYSDLRNQFRLKSQMAQSVMTVIARYKSAKSNNGHEWSLIDFKLAEYDLVWNRDYSITKN
ref|ZP_02619781.1| VYDDLRTKFNLSQMAVNTCTTVCSYVTQHSNKVFNLSLAVYKSPKAIYSYGRDYSFLND
RAAC03163 AYRTIRTRFGLRAQMACTVCRVAVGVYKSMKSNNGEDTPAEFKRAKPVFQWNKDYVLRNG
      * : * . : * : * * . . . * . * : : * * : : * : : : * : : :

ref|YP_148733.1| --LFSVNTLEGRIKVPFEPK-GMEPYFDGSWTFGTAKLVYKHNKFFLHIPMTKTIPTVDE
ref|YP_146224.1| --LFSVNTLEGRIKVSFEPK-GMEPYFDGSWTFGTAKLVYKHNKFFLHIPMTKTIPTVDE
ref|ZP_02171259.1| --LFSVNTLEGRIKVPFYTE-GLEPYFDGTWTFGTAKLVHKKYKWFLLHIPVSKDVQEANL
ref|NP_977687.1| --QFSVNTLEGRLKLNRYERK-AMKKYFNSTWKFGTAKLVHKKYKWFLLHIPMTKEYQTLDF
ref|ZP_02619781.1| GQTISINTIAKRIKVFHFKVNNYFKKYLTKESFSGSLEIVERENSYYAHITVSKEVQEKPL
RAAC03163 --VAHITTIDGRLRIPFRVEPPYRRLYQDGTWTFGAELVRKRHWYLVHVSVSKSVPD-HA
      : . * : * : : : : . * : * . * : : * : : : * : : : *

ref|YP_148733.1| HNIRQVVGVDVGVNFLAAAYDSQGKTTFFNGRKIKHMRKAKYKMRKTLQQKGTASARRKL
ref|YP_146224.1| HNIRQVVGVDVGVNFLAVAYDSQGKTTFFNGRKIKHMRKAKYKMRKTLQQKGTASARRKL
ref|ZP_02171259.1| DNINQVVGIDMGVNFTATTYDSNGQTRFFNGKQIKHMRKAKYKMRKELQQKQTPSARRKL
ref|NP_977687.1| ADVNNIVGVLDGINFLATTYDSQSKTTFYNGNIVKHKRGKFKATRKQLQTRQTPSSRRKKI
ref|ZP_02619781.1| NEFKNIIGIDLGQNFIAIFYDSKGIKFFKGRYLDKMDRAKYTRLRKLQKQKGTYNHKKI
RAAC03163 DDFDAVIGVDQGMRFIVTAS-CGDQVMFIRGGRVKQTRLRYVRLRASLQKQKGTSAKRRKL
      : . : : * * . * . . . : * . * : * . * : : * * : * : : : :

ref|YP_148733.1| KTIGQRENRMWTDVNHAVTKALVRQYGER---TLFVLEDLTGIREKTERVRIHDRIYETVS
ref|YP_146224.1| KTIGQRENRMWTDVNHAVTKALVRQYGER---TLFVLEDLTGIREKTERVRIHDRIYETVS
ref|ZP_02171259.1| KKIGQRENRMWQNIYNCISKALVEQYQKD---TLFVLEDLTGVRNATEKVRVKDRYQTVS
ref|NP_977687.1| KQIGSRENRYVTDVNHQITKALVEAYPRG---TMFVLEDLTGVRNATEKVRVKDRYQTVS
ref|ZP_02619781.1| VTINQREQRTMTYINHKSKEIVEYAKQNN--AIIAMENLTGIN-LSCKVKKNRYRVS
RAAC03163 KAIGRESRFMTDVNHQIAKAVVRFAQAQGORPLIVLEDLTGSN-LSVRFLKDRYWRMS
      * . * . * : : * : : * : * . : : : * : * : : * : : : *

ref|YP_148733.1| WAFYQFRQLEKARLHGSKVMVAPHYTSLTCPKCGHTEKANRNKRTHTFCCRTCGYTS
ref|YP_146224.1| WAFYQFRQMLEYKARLHGSKVIVVAPHYTSLTCPKCGHTEKANRNKRTHTFCCRTCGYTS
ref|ZP_02171259.1| WSFNDLRQKIAYKAQKTGALGIAVDPGYTSQTCPKCGHTEKANRNKRTHTFCCQTCQYTS
ref|NP_977687.1| WAFYQFRQMLEYKAEKNGQKVIIVDPKYTSQTCPKCGHTEKANRNKRTHTFCKKNCQYQS
ref|ZP_02619781.1| WAFNQLQQFIEYKAVQAGLKVIIYNPKYTSQTCPICGHIHKDNRNKKLHIFKCRACGCTL
RAAC03163 WAFRQLTDFIRYKAEAGIAVMFIDPTGTSETCPKCGHCETANRSRKRHEFRCKKCGYRC
      * : * : : : * * . : : * * * * * : . . * : : : * * : *

ref|YP_148733.1| NDDRIGAMNLQRKGIE-----
ref|YP_146224.1| NDDRIGAMNLQRKGIE-----
ref|ZP_02171259.1| NDDRIGAMNLQRKGI-----
ref|NP_977687.1| NDDRISAMNLHRKGIK-----
ref|ZP_02619781.1| NDDLIGAKNIQHKGYD-----
RAAC03163 NDDLAAASRVIAQKGLECLRQSQA
      *** . : : : *

```

FIG. 117

```

ref|YP_148733.1|      --YSIVQGLFSVNTLEGRIKVPFEPKG-MEPYFDGSWTFGTAKLVYKHNFLLHIPMTKT
ref|YP_146224.1|      --YSIVQGLFSVNTLEGRIKVSFEPKG-MEPYFDGSWTFGTAKLVYKHNFLLHIPMTKT
ref|ZP_02171259.1|    --YSLTKGLFSVNTLEGRIKVPFYTEG-LEPYFDGTWTFGTAKLVHKYGKWFLLHIPVSKD
ref|YP_001396471.1|    --YSLNKNIFSINSIQRLKIPYQSKG-MEKYFDGSYSFGTAKLVYKFNKYFLHIPVTKD
ref|ZP_02619781.1|    -----INTIAKRIKVHFKVNNYFKKYLTKEWSFGSLEIVERENSYYAHITVSKE
RAAC01387              MVYVLRSGLAHITTFDGRRLRIPYHVEPPYLRYLQDGWTFGAELVQKRHRWYLVHSVSKT
                        :.: : *.: : : : * : :.: : : * : : : *.: : *
ref|YP_148733.1|      IPTVDEHNIRQVVGVDVGVNFLAAAYDSQGKTTFFNGRRIKHMRAKYKMRKTLQQKGTA
ref|YP_146224.1|      IPTVDEHNIRQVVGVDVGVNFLAVAYDSQGKTIFFFNGRRIKHMRAKYKMRKTLQQKGTA
ref|ZP_02171259.1|    VQEANLDNINQVVGIDMGVNFTATTYDSNGQTRFFNGKQIKHKRAKYKMRKELQQKQTP
ref|YP_001396471.1|    YPQTTPFEINKIVGIDLGINFLLATTYDSFGKATFYQGRHIKAKRGHYKILRKQLQECGSK
ref|ZP_02619781.1|    VQEKPLNEFKNIIGIDLGQNFIFATFYDSKGKIKFFKGRYLKDMRAKYTRLRKQLQRKGTY
RAAC01387              APD-PVGDFDAVIGIDQGMRLITASCGN-QVMFIRGGKVKQTRLRYVRLRASLQRKGTR
                        : : :.: * * * . * . : * . * : * * : * * . :
ref|YP_148733.1|      SARRKLKTIGQRENRMWTDVNHAVTKALVRQYG---ERTLFVLEDLTGIREKTERVRIHD
ref|YP_146224.1|      SARRKLKTIGQRENRMWTDVNHAVTKALVRQYG---ERTLFVLEDLTGIREKTERVRIHD
ref|ZP_02171259.1|    SARRKLKKTIGQRENRMWQNYNCISKALVEQYG---KDTLFVMEDELTVGRNATEKVRVKD
ref|YP_001396471.1|    SAKRRIKSIGSRENRVSDINHQTSLVDKYG---TNTLFVLEDLTNVRTATEKVTINN
ref|ZP_02619781.1|    NAHKKIVTINQREQRTMTYINHKISKEIVEYAKQ--NNAI IAMENLTGIN-LSCKVKKDN
RAAC01387              SAKRRLKAIGRRESRFMADVNHQIAKAVIRFAQTQGRPLIVLEDLTGSN-LSVRFRLLK
                        .*:.: : * . *. * : :.: : * : : :.: : *:.*. . : : . :
ref|YP_148733.1|      RYETVSWAFYQFRQLLEYKARLHGSKVMVAPHYTSLTCPKCGHTEKANRNKRTHTFCCR
ref|YP_146224.1|      RYETVSWAFYQFRQMLEYKARLHGSKVIVVAPHYTSLTCPKCGHTEKANRKKRTHTFCCR
ref|ZP_02171259.1|    RYQTVSWSFNDLRQKIAYKAQKTGALGIAVDPGYTSQTCPKCGHTEKANRNKKTHTFCCQ
ref|YP_001396471.1|    RYVSVSWAFYQFRQLLEYKAQMNGSIVMAVNPKYTSQTCPKCGHIEKANRDKNKHLIFKCK
ref|ZP_02619781.1|    RYYRVSWAFNQLQFFIEYKAVQAGLKVIIYNPKYTSQTCPICGHIHKNRNRKHLIFKCK
RAAC01387              QYWRMSWSFRQLADFIYKAEAGIAVMLVDPAGTSETCPKCGHCEPANRSRKRHEFRCK
                        : * :.: : * : : : * * : : * * * * * . *.: : * * * :
ref|YP_148733.1|      TCGYTSNDDRIGAMNLQRKGIE-----
ref|YP_146224.1|      TCGYTSNDDRIGAMNLQRKGIE-----
ref|ZP_02171259.1|    TCQYTSNDDRIGAMNLQRKGI-----
ref|YP_001396471.1|    NCSYQSNDDRIGAINLWRKGIEYIEQS---
ref|ZP_02619781.1|    ACGCTLNDDLIGAKNIQHKGYD-----
RAAC01387              RCGYRCNDDLAASRVIAQKGLECLRQSQSA
                        * * * * . : : :.: : *

```

FIG. 118

| | |
|---------------------------|--|
| gb AAL87775.1 AF403183_11 | ----- |
| ref ZP_01368657.1 | REFZPTRANSPSASEISDESULFITBACTERIUMHAFNIENSEDCB-GBE |
| ref YP_517288.1 | ----- |
| ref YP_517659.1 | ----- |
| ref YP_001212990.1 | ----- |
| RAAC03385 | ----- |
| gb AAL87775.1 AF403183_11 | ----- |
| ref ZP_01368657.1 | ATTRANSPSASEISDESULFITBACTERIUMHAFNIENSEDCB-GBEATT |
| ref YP_517288.1 | ----- |
| ref YP_517659.1 | ----- |
| ref YP_001212990.1 | ----- |
| RAAC03385 | ----- |
| gb AAL87775.1 AF403183_11 | ----- |
| ref ZP_01368657.1 | RANSPSASEISDESULFITBACTERIUMHAFNIENSEDCB-GBEATTRAN |
| ref YP_517288.1 | ----- |
| ref YP_517659.1 | ----- |
| ref YP_001212990.1 | ----- |
| RAAC03385 | ----- |
| gb AAL87775.1 AF403183_11 | ----- |
| ref ZP_01368657.1 | SPSASEISDESULFITBACTERIUMHAFNIENSEDCB-GBEATTRANSPS |
| ref YP_517288.1 | ----- |
| ref YP_517659.1 | ----- |
| ref YP_001212990.1 | ----- |
| RAAC03385 | ----- |
| gb AAL87775.1 AF403183_11 | ----- |
| ref ZP_01368657.1 | ASEISDESULFITBACTERIUMHAFNIENSEDCB-GBEATTRANSPSASE |
| ref YP_517288.1 | ----- |
| ref YP_517659.1 | ----- |
| ref YP_001212990.1 | ----- |
| RAAC03385 | ----- |
| gb AAL87775.1 AF403183_11 | ----- |
| ref ZP_01368657.1 | ISDESULFITBACTERIUMHAFNIENSEDCB-GBEATTRANSPSASEISD |
| ref YP_517288.1 | ----- |
| ref YP_517659.1 | ----- |
| ref YP_001212990.1 | ----- |
| RAAC03385 | ----- |
| gb AAL87775.1 AF403183_11 | ----- |
| ref ZP_01368657.1 | ESULFITBACTERIUMHAFNIENSEDCB-GBEATTRANSPSASEISDESU |
| ref YP_517288.1 | ----- |
| ref YP_517659.1 | ----- |
| ref YP_001212990.1 | ----- |
| RAAC03385 | ----- |
| gb AAL87775.1 AF403183_11 | ----- |
| ref ZP_01368657.1 | LFITBACTERIUMHAFNIENSEDCB-GBEATTRANSPSASEISDESULFI |
| ref YP_517288.1 | ----- |
| ref YP_517659.1 | ----- |
| ref YP_001212990.1 | ----- |
| RAAC03385 | ----- |

| | |
|---------------------------|--|
| gb AAL87775.1 AF403183_11 | ----- |
| ref ZP_01368657.1 | TBACTERIUMHAFNIENSEDCB-SCRESIGNIFICANCEE-IDENTITIE |
| ref YP_517288.1 | ----- |
| ref YP_517659.1 | ----- |
| ref YP_001212990.1 | ----- |
| RAAC03385 | ----- |
| gb AAL87775.1 AF403183_11 | -----MLQHNSLPEQHQNQLSLIFSSSLKLSQLLRAAGIRKSY |
| ref ZP_01368657.1 | SPSITIVESGAPSMMLQHNSLPEQHQNQLSLIFSSSLKLSQLLRAAGIRKSY |
| ref YP_517288.1 | -----MLQHNSLPEQHQNQLSLIFSSSLKLSQLLRAAGIRKSY |
| ref YP_517659.1 | -----MLQHNSLPEQHQNQLSLIFSSSLKLSQLLRAAGIRKSY |
| ref YP_001212990.1 | -----QFYSRVDNFLSRHEIGKILRRSNFQKEK |
| RAAC03385 | -----MVRDSQHPDQLQSMVTAFFMEYHIGKLLRQSNITKQA |
| | * . : : : . : : * : : * |
| gb AAL87775.1 AF403183_11 | GVSSFVVFQIIFQLVFQGRNLFRLLEGSRAESLPGKDVVYRFLNDSRYNW |
| ref ZP_01368657.1 | GVSSFVVFQIIFQLVFQGRNLFRLLEGSRAESLPGKDVVYRFLNDSRYNW |
| ref YP_517288.1 | GVSSFVVFQIIFQLVFQGRNLFRLLEGSRAESLPGKDVVYRFLNDSRYNW |
| ref YP_517659.1 | GVSSFVVFQIIFQLVFQGRNLFRLLEGSRAESLPGKDVVYRFLNDSRYNW |
| ref YP_001212990.1 | GFSCLELFKFIFLLVFKGKNLYRTLQSEAEPRPEKDTIYRFLNSFRYNW |
| RAAC03385 | GIPVLEVFRLLFALVFHQRSIKRVLEQLAMN-QFGKDTVYRFLNSPRHNW |
| | * . : : * : * * : : * * * : * : * : * : * : * |
| gb AAL87775.1 AF403183_11 | RRFYQLLSLKMVGRFEK-LTSAQRIRVFIVDDSVMERERSKKVELLARVF |
| ref ZP_01368657.1 | RRFYQLLSLKMVGRFEK-LTSAQRIRVFIVDDSVMERERSKKVELLARVF |
| ref YP_517288.1 | RRFYQLLSLKMVGRFEK-LTSAQRIRVFIVDDSVMERERSKKVELLARVF |
| ref YP_517659.1 | RRFYQLLSLKMVGRFEK-LTSAQRIRVFIVDDSVMERERSKKVELLARVF |
| ref YP_001212990.1 | RKFLLLSSSVINETIEPLTSRNWKNVLVDDSLYSRNRKAVELLARVK |
| RAAC03385 | RRFLLLLSSAVVRRTAR-LTSEDRAADVIVDDSLFSRSRSKKVELLAKVY |
| | * : * : * : . . * : * : * : * : * : * : * : * |
| gb AAL87775.1 AF403183_11 | DHV----- |
| ref ZP_01368657.1 | DHV----- |
| ref YP_517288.1 | DHV----- |
| ref YP_517659.1 | DHV----- |
| ref YP_001212990.1 | DHV----- |
| RAAC03385 | DHV----- |
| | DHVLRSNLLTEQPASIPSGDYFFDGMVYSPAGVNRPYAHPDKGYGISPY |
| | *** |
| gb AAL87775.1 AF403183_11 | ----- |
| ref ZP_01368657.1 | ----- |
| ref YP_517288.1 | ----- |
| ref YP_517659.1 | ----- |
| ref YP_001212990.1 | ----- |
| RAAC03385 | VTWHRVGNQDLHCQIDYDSQSILEAPAWAVGSGVGALVGMVTDGNLTG |
| gb AAL87775.1 AF403183_11 | ----- |
| ref ZP_01368657.1 | ----- |
| ref YP_517288.1 | ----- |
| ref YP_517659.1 | ----- |
| ref YP_001212990.1 | ----- |
| RAAC03385 | SVIGAIIGGALGAYYGAQVERLADENGCIWFTIDNPPIVNLGTWWAPEW |
| gb AAL87775.1 AF403183_11 | ----- |
| ref ZP_01368657.1 | ----- |
| ref YP_517288.1 | ----- |
| ref YP_517659.1 | ----- |
| ref YP_001212990.1 | ----- |
| RAAC03385 | YMDVRYIELGPWATTNFYYYLGGVY |

FIG. 119

```

ref|YP_001126171.1| MDELLAVQPELERQVFHAVADLLNLEVDLIYFDTTSSYFEVDPSETPEGESLRKQGFSGD
ref|YP_148969.1| MDELLAVQPELERQVFHAVADLLNLEVDLIYFDTTSSYFEVDPSETPEGESLRKQGFSGD
ref|YP_146154.1| MDELLAVQPELERQVFHAVADLLNLEVDLIYFDTTSSYFEVDPSETPEGESLRKQGFSGD
ref|YP_146741.1| MDELLAVQPELERQVFHAVADLLNLEVDLIYFDTTSSYFEVDPSETPEGESLRKQGFSGD
ref|ZP_02172080.1| MDELLSVQKQLEDVDFHDVSNLLNLEVDLIYFDTTSSYFEVASDETPEEDDFRLQGYSGD
RAAC03398 MDLFLHDVAEDLQYEVFRRVSDLLNLDVDDLFFDTTSTYFETE-DES--DDSLRRKGYSGD
* * * *: *: .** *: :****:****:****:****. *: :.:* :*:**

ref|YP_001126171.1| KRPDIVQIVIGLAVTREGVPIRAWVWPGNTMDMTVIKQVKQDLIGWKLGRVISVMDRGFS
ref|YP_148969.1| KRPDIVQIVIGLAVTREGVPIRAWVWPGNTMDMTVIKQVKQDLIGWKLGRVISVMDRGFS
ref|YP_146154.1| KRPDIVQIVIGLAVTREGVPIRAWVWPGNTMDMTVIKQVKQDLIGWKLGRVISVMDRGFS
ref|YP_146741.1| KRPDIVQIVIGLAVTREGVPIRAWVWPGNTMDMTVIKQVKQDLIGWKLGRVISVMDRGFS
ref|ZP_02172080.1| KRSDIVQVIGLAVTREGVPIKVTWTPGNTMDMNVVEEVKDKLMGWRLGRVVKVMDRGFS
RAAC03398 HRPDLFQVIGLAVTRDGIPVRCWTWPGNTADMSVVEEVKQDLIGWRLGRVITTVDRGFV
*: ** *:*****:*.*: :.***** ** *: :****:****:****:****. *:*****

ref|YP_001126171.1| SEENLRILQQAGGHYIVGEKMRSGKAAVEEALNRRGRYQQVRENLHIKEIIVGDGEARQR
ref|YP_148969.1| SEENLRILQQAGGHYIVGEKMRSGKAAVEEALNRRGRYQQVRENLHIKEIIVGDGEARQR
ref|YP_146154.1| SEENLRILQQAGGHYIVGEKMRSGKAAVEEALNRRGRYQQVRENLHIKEIIVGDGEARQR
ref|YP_146741.1| SEENLRILQQAGGHYIVGENMRSGKAAVEEALSRRGRYHEVDENLHIKEIIVGDGEARQR
ref|ZP_02172080.1| SEKNLRILQRGAGHYIIGERMRSKKDVEEALSKRGRFHKVRENLHVKEIVGDGEARKR
RAAC03398 SESNLRILQRAGGHCIAGEKMTSGKPAVEAALARPGRFRELRPNLKVKEVVVDGEARVR
**.******:..** * **.* ** ** ** : **::: **::** :***** *

ref|YP_001126171.1| YVLVYNPSEAERQKEREKLLSKEELEGLRQLPNEAHHKATCRLRSHPSYGYLRQLK
ref|YP_148969.1| YVLVYNPSEAERQKEREKLLSKEELEGLRQLPNEAHHKATCRLRSHPSYGYLRQLK
ref|YP_146154.1| YVLVYNPSEAERQKEREKLLSKEELEGLRQLPNEAHHKATCRLRSHPSYGYLRQLK
ref|YP_146741.1| YVLVYNPGEAERQKEREKLLSKEELEGLRQLPNEAHHKATCRLRSHPSYGYLRQLK
ref|ZP_02172080.1| YVIAFNPDIAIRDREKRQEIVRSVEEQLHHLKQLPNEAHHKQACALRAHKVYGYIRQLK
RAAC03398 YVLAFNPEEAKRDEARREAMLRELRLMELERLKLQGEAHTKAHCRSLASHPTFKRYLKQDR
**.:** ** *: .*: :.:.. :*. :*: * .*** * * * : :*: * :

ref|YP_001126171.1| DGTLRIDKQAVREAKEYDGKYLIRTSDDTLAEDVALGYKQLVDIEQAFRTLKSTLELRP
ref|YP_148969.1| DGTLRIDKQAVREAKEYDGKYLIRTSDDTLAEDVALGYKQLVDIEQAFRTLKSTLELRP
ref|YP_146154.1| DGTLRIDKQAVREAKEYDGKYLIRTSDDTLAEDVALGYKQLVDIEQAFRTLKSTLELRP
ref|YP_146741.1| DGTLRIDKQAVRDAEKYDGNLIRTSDDTLAEDVAIGYKQLVDIEQAFRTLKSTLELRP
ref|ZP_02172080.1| DGTCLKLNKQAVRDAEKYDGKYLIRTSDDTLAEDVALGYKQLQVEDAFRTMKTTLVLRP
RAAC03398 WGNLRIDPEAVRQAAHLDGKYLIRTSDDTLSTEDVALGYKQLLMVESAFRTLKTTLDIRP
*.*::: :***:* : **.****** *****.*****: *.*****:*** :**

ref|YP_001126171.1| MYHRLEDRIRAHVLLSWLALLLVRIVEIRTHESWPKVRDECERLMLGHFSSKNGDLYQRT
ref|YP_148969.1| MYHRLEDRIRAHVLLSWLALLLVRIVEIRTHESWPKVRDECERLMLGHFSSKNGDLYQRT
ref|YP_146154.1| MYHRLEDRIRAHVLLSWLALLLVRIVEIRTHESWPKVRDECERLMLGHFSSKNGDLYQRT
ref|YP_146741.1| MYHRLEDRIRAHVLLSWLALLLVRIVEIRTHESWPKVRDECERLMLGHFSSKNGDLYQRT
ref|ZP_02172080.1| MYHRLEDRIRAHVIISWLALLLVRIEELDTNESWNTVRNIQRLQAGHFTTSDGDLYRTT
RAAC03398 MYHRKDERIRSHVLLCWLALLLVRIAEVQTRSWPDIRSHMQAMHQVTKSTPDGIVVQRT
**** :****:***:..*****: * : * .** * : : : : : * : : *

ref|YP_001126171.1| ELTAKQALFLAALGLEPPPKILGIHPRT---
ref|YP_148969.1| ELTAKQALFLAALGLEPPPKILGIHPRT---
ref|YP_146154.1| ELTAKQALFLAALGLEPPPKILGIHPRT---
ref|YP_146741.1| ELTAKQALFLAALGLEPPPKILGIHPRT---
ref|ZP_02172080.1| TPTAKQKEIFRNTGTEIPPEILAIRPKS---
RAAC03398 ETTEVQWEILRALRITEPPRILRMEPRTRL
* * : : **.* ** :.*:

```

SUBSTITUTE SHEET (RULE 26)

FIG. 121

```
ref|YP_074105.1|
ref|YP_076073.1|
ref|YP_001036724.1|
ref|YP_001039064.1|
ref|YP_001039349.1|
RAAC03588
-----
-----
TRANSPSASEMUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPSASE
-----
-----
```

```
ref|YP_074105.1|
ref|YP_076073.1|
ref|YP_001036724.1|
ref|YP_001039064.1|
ref|YP_001039349.1|
RAAC03588
-----
-----
MUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPSASEMUTATRTYPE
-----
-----
```

```
ref|YP_074105.1|
ref|YP_076073.1|
ref|YP_001036724.1|
ref|YP_001039064.1|
ref|YP_001039349.1|
RAAC03588
-----
-----
CLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPSASEMUTATRTYPECLSTRIDIUM
-----
-----
```

```
ref|YP_074105.1|
ref|YP_076073.1|
ref|YP_001036724.1|
ref|YP_001039064.1|
ref|YP_001039349.1|
RAAC03588
-----
-----
THERMCELLUMATCCGBABNTRANSPSASEMUTATRTYPECLSTRIDIUMTHERMCELLU
-----
-----
```

```
ref|YP_074105.1|
ref|YP_076073.1|
ref|YP_001036724.1|
ref|YP_001039064.1|
ref|YP_001039349.1|
RAAC03588
-----
-----
MATCCGBABNTRANSPSASEMUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABN
-----
-----
```

```
ref|YP_074105.1|
ref|YP_076073.1|
ref|YP_001036724.1|
ref|YP_001039064.1|
ref|YP_001039349.1|
RAAC03588
-----
-----
TRANSPSASEMUTATRTYPECLSTRIDIUMTHERMCELLUMATCCGBABNTRANSPSASE
-----
-----
```

```
ref|YP_074105.1|
ref|YP_076073.1|
ref|YP_001036724.1|
ref|YP_001039064.1|
ref|YP_001039349.1|
RAAC03588
-----
-----
MUTATRTYPECLSTRIDIUMTHERMCELLUMATCCSCRESIGNIFICANCEE-IDENTIT
-----
-----
```

```
ref|YP_074105.1|
ref|YP_076073.1|
ref|YP_001036724.1|
ref|YP_001039064.1|
ref|YP_001039349.1|
RAAC03588
-----
-----
-----VISDAHEGLRQAI AEVLQGA AWQRCRVHFM RNLLGYVPKPAQSMV
-----ISDSHVGLQEAIRTVLSGASWQRCRVHFM RNLLSYVPKHWQSMV
IESPSITIVESGAPSVISDAHEGLKAAIKKILTGS AWQRCRVHFM RNVL SQVPKHYQGMV
-----VISDAHEGLKAAIKKILTGS AWQRCRVHFM RNVL SQVPKHYQGMV
-----VISDAHEGLKAAIKKILTGS AWQRCRVHFM RNVL SQVPKHYQGMV
-----MVSDAHAGLRQAI SEVLTGATWQRCVKVHTIR NVLSQVPKREQSMV
:***: * *: ** : * *: *****: ** : * *: * . *** * . **
```


FIG. 122

```
ref|NP_634718.1|-----MLKTEEWLLIRDLYSQGFSISEIA
ref|NP_616807.1|-----MLKTEEWLSIRDLYSQGFSISEIS
ref|YP_074959.1|-----MLRSGETLEIRQMYAGGLSISEIA
gb|AAR99616.1|-----IKEMYERGMSISDIA
sp|Q45618|TRA6_BACST-----IKEMYERGMSISDIA
RAAC03818MRI PQDHRPQFLKIIGMNSPPPCRIVAPVIRSWEVPMREDERMEIRQLYEAGVSISELA
                                         *:::*  *.***:::

ref|NP_634718.1|KQTGFDRKTVRKYLRLKTLPEPQ--KRSGRKSKLDPFKP----
ref|NP_616807.1|RRTGYARETVRKYL-----
ref|YP_074959.1|RRTGRDRKTIRKWLRTNTMPKPA--KRKR-----
gb|AAR99616.1|RELGIDRKTVRKYIHSPNPPSKS--KRKQ-----
sp|Q45618|TRA6_BACSTRELGIDRKTVRKYIHSPNPPSKS--KRKQ-----
RAAC03818RRFGYDRKTIRNALNSSLERSRAQERRVGSEKVPSPTRIT
:. *  *::*: : :
```

FIG. 123

```

ref|YP_594046.1|      LKGRIEQGVLSAVVLFREVQERGYQGQYTVVKDFVRPFRRTQVSAARVTTRFETAPGEQA
RAAC03819             MKQRMQLGVLNAERILREIREQGYTGGITVLRREFMKPLR--PVVSAKATERYESDPGEQA
ref|YP_074959.1|      --EQMQKGVTSASKMLYLLQQRGFKGKIRIVRAFMAPYR--PMAKAAATVRFETPPGKQA
ref|YP_359963.1|      IQELMNLGIFNCEVIYERIKEEGYTGGRTILRDYVRQFR--PPKQVPAVCRYETKPGQQA
ref|YP_361300.1|      IQELMNLGIFNCEVIYERIKEEGYTGGRTILRDYVRQFR--PPKQVPAVCRYETKPGQQA
ref|NP_622784.1|      ----INMGIFNCEVIYERIKEEGYTGGKTILRDYVRQFR--PSKHIQAVCRYETKSGEQA
                        :: *: .. : :...*: * ::: : * .. *:: :*:**

ref|YP_594046.1|      QVDFGRYSYLNLE-GQTRSIWAFVMVLGWSRALYVEFIRKADTASFIRCHLNAFAYFGGM
RAAC03819             QIDLGAFFPYDSH-GQRRTIWAFAMVLAYSRLMYVEFIKAADQLHLILQALRNALEFFGGV
ref|YP_074959.1|      QVDWADFGYIEVD-GRRLKLYCFIMVLAYSRAMYLEFVTATDMKTFMRCHINAFKFFGGV
ref|YP_359963.1|      QVDWGEYTYIDEETGEIRKLYVFVMVLGYSRAIYVEFTNRCDVRTFIRCLIHGFEYFGGV
ref|YP_361300.1|      QVDWGEYTYIDEETGEIRKLYVFVMVLGYSRAIYVEFTNRCDVHTFIRCLIHGFEYFGGV
ref|NP_622784.1|      QVDWGEYNYIDQETGEVRKLYLFVMVLGYSRAMYVEFTNRCDVHTFNRCILIHGFEYFGGV
*: * . : * : . * . ::: * ***.:** :*:** * : .. :.. :***:

ref|YP_594046.1|      TQSILYDNTKQVVLERDETG-----
RAAC03819             PRVMSLDNCSPLVVANDGQGHVDCNRLISILPSSTDSCPRHVLAGAAPRAR
ref|YP_074959.1|      PHEILYDNVKTIVVKDRD-----
ref|YP_359963.1|      TDIVLTDRMKTIVIL-----
ref|YP_361300.1|      TDIVLTDRMKTIVIL-----
ref|NP_622784.1|      TDVVLTD RMKTIVIIIGTDANKKPIWNAVFEDLAATLGFVPR-----
. : * * . . : :

```

FIG. 124

```

ref|ZP_01695687.1|      LVLNIIICSGKPLYKVHEFYQNLDSSEMLF--DTSVSPDQLNDDALGRALDYLYEA-EAWKV
ref|ZP_01695982.1|      LVLNIIICSGKPLYKVHEFYQNLDSSEMLF--DTSVSPDQLNDDALGRALDYLYEA-EAWKV
ref|ZP_01695971.1|      LVLNIIICSGKPLYKVHEFYQNLDSSEMLF--DTSISPDQLNDDALGRALDYLYKA-EAWKV
ref|ZP_01695655.1|      LVLNIIICSGKPLYKVHEFYQNLDSSEMLF--DTSISPDQLNDDALGRALDYLYKA-EAWKV
ref|YP_430569.1|        MLVNIGTNREALYRVEEFYAQRDVEVLLG--SGVSADDLHDDALARALDALYDA-GLEAL
RAAC03823                :::* : . :.*:*.*.* : * *:* : . :.:*****.*.* * : :

ref|ZP_01695687.1|      YSTLALKTLKKNLNP-----IGILHNDTTSISVYGEYKQPKEDGLQITYGYSKAHRPDLK
ref|ZP_01695982.1|      YSTLALKALKKNLNP-----IGILHNDTTSISVYGEYKQPKEDGLQITYGYSKAHRPDLK
ref|ZP_01695971.1|      YSTLALKALKKNLNP-----IGVLHNDTTSISVYGEYKQPKEDGLQITYGYSKAHRPDLK
ref|ZP_01695655.1|      YSTLALKALKKNLNP-----IGVLHNDTTSISVYGEYKQPKEDGLQITYGYSKAHRPDLK
ref|YP_430569.1|        FSSIALTAATHNVS-----IAGIHVDTTSISVQAYDG--EGDLDTTFGFSKDHHRPDLK
RAAC03823                YARIALHTLRLRLVSDSNELIPIHADTTSLSMTGEYLD--QTAFRIDRGFSKDHHRPDLK
                        : : * : . : : * *****: * * : : * *:* ** *****

ref|ZP_01695687.1|      QIVLGMGVTTERIPILAKVENGNTSDKSWNVEFIQMKRKILSHEDWKNLIYQADSALITT
ref|ZP_01695982.1|      QIVLGMGVTTERIPILAKVENGNTSDKSWNVEFIQMKRKILSHEDWKNLIYQADSALITT
ref|ZP_01695971.1|      QIVLGMGVTTERIPILAKVENGNTSDKSWNVEFIQMKRKILSHEDWKNLIYQADSALITT
ref|ZP_01695655.1|      QIVLGMGVTTERIPILAKVENGNTSDKSWNVEFIQMKRKILSHEDWKNLIYQADSALITT
ref|YP_430569.1|        QFLIGLTVNRDGLPILAQSLDGNSSDKSWYPQVIEELVQTFKPEKLEKIVFVADCALVTK
RAAC03823                QIVFGL-CTVHGLGLCANVNPGLDDHTWNFENIQQLLSQLDEETRKRSVYVADAALVTK
                        *::*: . . : : * : ** .*: : *:: . . * * . : **.*.*.

ref|ZP_01695687.1|      ENLAEIQQQN-----LSFISRLPDTFGLSTELKKEAWLLNN-WERVGSLSNKKDAIYQI
ref|ZP_01695982.1|      ENLAEIQQQN-----LSFISRLPDTFGLSTELKKEAWLLNN-WERVGSLSNKKDAIYQI
ref|ZP_01695971.1|      ENLAEIQQQN-----LSFISRLPDTFGLSTELKKEAWLLNN-WERVGSLSNKKDAIYQI
ref|ZP_01695655.1|      ENLAEIQQQN-----LSFISRLPDTFGLSTELKKEAWLLNN-WERVGSLSNKKDAIYQI
ref|YP_430569.1|        DNLAALLVQEEGNKPALQFISLLPENFGLNKEIKAEAFRTGT-WQEIIGLSPKKDAACYKS
RAAC03823                DNLELLAEED-----FHFISRLPGTYKLSEDLKRAAWEKENSWEKVGRLAEAEADSAHYRI
                        **: : : : : *** ** . : * .:* * : . *::* * : :*: * :

ref|ZP_01695687.1|      QAFERQIQNLPIRFLVVSNNLDQRKEKTLNRAIEKEEIQLKKEIEKLSKVIHFCKEDAL
ref|ZP_01695982.1|      QAFERQIQNLPIRFLVVSNNLDQRKEKTLNRAIEKEEIQLKKEIEKLSKVIHFCKEDAL
ref|ZP_01695971.1|      QAFERQIQNLPIRFLVVSNNLDQRKEKTLNRAIEKEEIKLKKEIEKLSKVIHFCKEDAL
ref|ZP_01695655.1|      QAFERQIQNLPIRFLVVSNNLDQRKEKTLNRAIEKEEIQLKKEIEKLSKVIHFCKEDAL
ref|YP_430569.1|        QSFVREIDGRDYRLIVVHSTTLDRKENSLLKKWAKQREVLEKAAKDLRRPFACKADAR
RAAC03823                QAFRTLYGRTYRFVVVRSSSLDRKERKLKEVLKREKAALEKAAKAMSQNVYSCEQDAQ
                        *: * * : . *::*:.*.*.* **.* . . : . ** : : : : * : **

ref|ZP_01695687.1|      EAIQSFKKKQKASYFCYKLVQMSDQPIKRRKGRPKKDEQTK-REIYQIQLESFEKDH
ref|ZP_01695982.1|      EAIQSFKKKQKASYFCYKLVQMSDQPIKRRKGRPKKDEQTK-REIYQIQLESFEKDH
ref|ZP_01695971.1|      EAIQSFKKKQKASYFCYKLVQMSDQPIKRRKGRPKKDEQTK-REIYQIQLESFEKDH
ref|ZP_01695655.1|      EAIQSFKKKQKASYFCYKLVQMSDQPIKRRKGRPKKDEQTK-REIYQIQLESFEKDH
ref|YP_430569.1|        KAIELFLREYR-----
RAAC03823                MAMQTFMHEHRATLHPISARICAEQVQAKRARRGRPRKDDPPPPVHTQYRVEVAILPPSE
                        *: * : : :

ref|ZP_01695687.1|      DFIEKKKKMLSTFVLITNKLDEETLSNQEVLRVYKGQSAATRFRLIKDSQMIDAIYLT
ref|ZP_01695982.1|      DFIEKKKKMLSTFVLITNKLDEETLSNQEVLRVYKGQSAATRFRLIKDSQMIDAIYLT
ref|ZP_01695971.1|      DFIEKKKKMLSTFVLITNKLDEETLSNQEVLRVYKGQSAVETRFRLIKDSQMIDAIYLT
ref|ZP_01695655.1|      DFIEKKKKMLSTFVLITNKLDEETLSNQEVLRVYKGQSAATRFRLIKDSQMIDAIYLT
ref|YP_430569.1|        -----
RAAC03823                ERVQQWREKEATFVLITDIRDDQVRVSEQILRLYKEQHEVEARFRYLSKSPYHVGPYIYLHK

ref|ZP_01695687.1|      PERVEALGIVYVMALLIYGILEYRVRKELKEKNLSLILKGKRKLYQPTGQALLEQLEDIT
ref|ZP_01695982.1|      PERVEALGIVYVMALLIYGILEYRVRKELKEKNLSLILKGKRKLSQPTGQALLEQLEDIT
ref|ZP_01695971.1|      PERVEALGIVYVMALLIYGILEYRVRKELKEKNLSLILKGKRKLYQPTGQALLEQLEDIT
ref|ZP_01695655.1|      PERVEALGIVYVMALLIYGILEYRVRKELKEKNLSLILKGKRKLYQPTGQALLEQLEDIT
ref|YP_430569.1|        -----
RAAC03823                PTRVKAFGFVMLLSLLLYSVLEYLIREKMKRETEPLMLPGNRKSFRTGLAILEMLDGV

ref|ZP_01695687.1|      VILINQNOQKIRLLPDNIDSQAKKIELCGYDLSIY-----
ref|ZP_01695982.1|      VILINQNOQKIRLLPDNIDSQAKKIELCGYDLSIY-----
ref|ZP_01695971.1|      VILINQNRQKRLRLPDNIDSQAKKIELCGYDLSIY-----
ref|ZP_01695655.1|      VILINQNRQKRLRLPDNIDSQAKKIELCGYDLSIY-----
ref|YP_430569.1|        -----
RAAC03823                TVHMQVGDWTQVRPATPHNPQIMRVLKLNNMDLSIYAESQKTA

```

```

ref|ZP_02330348.1|      MGLWTKQQQLREFIKENNLVTAQDAQNALKELFAETIQEMLEAELDTHLGYEKHEVKAKKT
ref|ZP_02328298.1|      MGLWTKQQQLREFIKENNLVTAQDAQNALKELFAETIQEMLEAELDTHLGYEKHEVKAKKT
ref|ZP_02326599.1|      MGLWTKQQQLREFIKENNLVTAQDAQNALKELFAETIQEMLEAELDTHLGYEKHEVKAKKT
RAAC01171                MELLSSKEQIRQLIRDGKLKDIHDVQSMKDLFASTIQEMLEAELNTHLGYAKYDAKHKDT
ref|YP_001664428.1|      MSLLTKEQLRNFISENNIQSIPDLYTSLKNLFKDTIQEMLEAELSTELGYERYDKKDKDT
ref|YP_001664274.1|      MSLLTKEQLRNFISENNIQSIPDLYTSLKNLFKDTIQEMLEAELSTELGYERYDKKDKDT
                          * * :*:*:*:*: * : : * . *:** .*****.*.*** : : * *.
                          *

ref|ZP_02330348.1|      PNSRNGR-SHKTVVSEYGEQQIAVPRDRMGEGFEPLVVKKHQSNVTGIEDQIVALYAK---
ref|ZP_02328298.1|      PNSRNGR-SHKTVVSEYGEQQIAVPRDRMGEGFEPLVVKKHQSNVTGIEDQIVALYAK---
ref|ZP_02326599.1|      PNSRNGR-SHKTVVSEYGEQQIAVPRDRMGEGFEPLVVKKHQSNVTGIEDQIVALYAK---
RAAC01171                DNARNHGGAKRVTQSELGDIDIALPRDRKGEGFEPLIVQKRQKRMPISIEEQVIALYLKDFS
ref|YP_001664428.1|      QNSRNGY-TQKTVKTKQFGEMEIDIPDRRQGEFEPKIVPKYKRDISGIEEKVIALYAR---
ref|YP_001664274.1|      QNSRNGY-TQKTVKTKQFGEMEIDIPDRRQGEFEPKIVPKYKRDISGIEEKVIALYAR---
                          *:*** :*:** : : * : * :**** ***** : * * : :..**:::*** :

ref|ZP_02330348.1|      -----
ref|ZP_02328298.1|      -----
ref|ZP_02326599.1|      -----
RAAC01171                KTHPKVVRTALQRGFFAV
ref|YP_001664428.1|      -----
ref|YP_001664274.1|      -----

```


FIG. 126

```

ref|YP_516922.1|      ---MFARIKTAYNRDGSPPRYLQLVESRREEGKVRQKVLGNLGRVEDLQNGKLDDDLIRSL
ref|YP_520815.1|      -----
ref|YP_519534.1|      ---MFARIKTAYNRDGSPPRYLQLVESRREEGKVRQKVLGNLGRVEDLQNGKLDDDLIRSL
ref|ZP_01372264.1|    ---MFARIKTAYNRDGSPPRYLQLVESRREEGKVRQKVLGNLGRVEDLQNGKLDDDLIRSL
ref|ZP_01370818.1|    ---MFARIKTAYNRDGSPPRYLQLVESRREEGKVRQKVLGNLGRVEDLQNGKLDDDLIRSL
RAAC03825             MILLFAQIVSTKKPDGKTYKYLHIVESYREGRTVKRRVASLGNISQYSEREIEQIIRTL

ref|YP_516922.1|      AKFSDKLAVVDAEDLFADWSKEYGPSMVFNRLWDNLGLHKILDGLFSEERELSIDVREAI
ref|YP_520815.1|      -----LWDNLGLHKILDGLFSEERELSIDVREAI
ref|YP_519534.1|      AKFSDTLAVVDAEDLFADWSKEFGPSMVFRRLWENLGLHKILSGLFNERELSIDVQEI
ref|ZP_01372264.1|    AKFSDTLAVVDAEDLFADWSKEFGPSMVFRRLWENLGLHKIFAGLFNERDLSIDVQEI
ref|ZP_01370818.1|    AKFSDTLAVVDAEDLFADWSKEFGPSMVFRRLWENLGLHTILDGLFNERDLSIDVQEI
RAAC03825             ESLLQHRTTG-SLEDFAEQQVLHFGVPYVQFLWNQLGLTEAIRDALRAREVTDFVARYV
                        *:*** : . : *::** . :

ref|YP_516922.1|      FCMVLNRLTEPTSKLGVSDWKDSVYRPDF--ETLKLHHFYKAIDFLDENKDTIEEQLFH
ref|YP_520815.1|      FWMVLNRLTEPTSKLGVSDWKDSVYRPDF--ETLKLHHFYKAIDFLDENKDTIEEQLFH
ref|YP_519534.1|      FCMVLNRLTEPTSKLGVSDWKDSVYRPEF--ESLKLHHFYKAIDFLDENKESIEEQLFH
ref|ZP_01372264.1|    FCMVLNRLTEPTSKLGVSDWKDSVYRPEF--ESLKLHHFYKAIDFLDENKDTLEEQLFFH
ref|ZP_01370818.1|    FCMVLNRLTEPTSKLGVSDWKDSVYRPEF--ESLKLHHFYKAIDFLDENKDTLEEQLFFH
RAAC03825             QAMVIHRLVDPSSKRLRFLHTLDDLPLDWGGEPWQLQHFYRALDYLDIKPQLERVLAR
                        *:***:*** : . *.***: *. :*:***:***: * :*. *: :

ref|YP_516922.1|      HTNLTQQQLDLVFFDTTSTYVEGDAGAFDLLEYGHSKDHDPDRQLQVMIGILMSRDGIPIA
ref|YP_520815.1|      HTNLTQQQLDLVFFDTTSTYVEGDAGAFDLLEYGHSKDHDPDRQLQVMIGILMSRDGIPIA
ref|YP_519534.1|      QTNLTQQQLDLVFFDTTSTYVEGDAGAFDLLEYGHSKDHDPDRQLQVMIGILMSRDGIPIA
ref|ZP_01372264.1|    HTNLTQQQLDLVFFDTTSTYVQGDAGAFDLLEYGHSKDHDPDRQLQVMIGILMSRDGIPIA
ref|ZP_01370818.1|    HTNLTQQQLDLVFFDTTSTYVEGDAGAFDLLEYGHSKDHDPDRQLQVMIGILMSRDGIPIA
RAAC03825             LTDLLNFRLSLVLYDLTSTHLHG--HACPLGEHGYSRTHRPDLQVELGLLVTPEGIPIT
                        *:*. :*.***:*.***:*. * * *:***: ***** *:***: :****:

ref|YP_516922.1|      HHVFPGNTPDPTDAFIEAVSDLKKRFNIQRVIVVGDGRGMMGKRTLELLEELQLQYILGVRM
ref|YP_520815.1|      HHVFPGNTPDPTDAFIEAVSDLKKRFNIQRVIVVGDGRGMMGKRTLELLEELQLQYILGVRM
ref|YP_519534.1|      HHVFPGNTPDPTDAFIEAVRDLKKRFTIQRVIVVGDGRGMMGKRTLELLEELQLHYILGVRM
ref|ZP_01372264.1|    HHVFPGNTPDPTDAFIEAVSDLKKRFTIQRVIVVGDGRGMMGKRTLELLEELQLHYILGVRM
ref|ZP_01370818.1|    HHVFPGNTPDPTDAFIEAVSDLKKRFTIQRVIVVGDGRGMMGKRTLELLEELQLHYILGVRM
RAAC03825             HEVFAGNVSDKQTPVDILKRLKEDFAVEQCVFVGDGRGMVTEKNMALMAEAGFPYIVGFHK
                        *.**.*.***: : : ***: * : : :*****: : : : * : * : ***:

ref|YP_516922.1|      RN---IKAGPDLANSPEPYAFIKDNLKVKEVVHQ-----EKRYIVCLNEEEAKRDQ
ref|YP_520815.1|      RN---IKAGPDLANSPEPYAFIKDNLKVKEVVHQ-----EKRYIVCLNEEEAKRDQ
ref|YP_519534.1|      RN---IKAGPELAASPEPYTFLKDNLKVKEVLHQ-----EKRYIVCLNEEEAKRDQ
ref|ZP_01372264.1|    RN---IKAGPELAASAQPYPFVKDNLKVKEVLHQ-----EKRYIVCLNEEEAKRDQ
ref|ZP_01370818.1|    RN---VKAGPELATSPEPYVFTKDNLKVKEVLHQ-----GKRYIVCLNEEEAKRDQ
RAAC03825             RGRIVSDALLEQFADVNAYHELKDNRLYLEVPAASVDDVEKAEGVRYILCYNPEKARQDA
                        *. . * : . :* ****: ** *****: : : * : * : ***:

ref|YP_516922.1|      LVREQIEIKLRS-----KLEHG-----SIKDLIGNSEYKKYINVTAEAAAT
ref|YP_520815.1|      LVREQIEIKLRS-----KLEHG-----SIKDLIGNSEYKKYINVTAEAAAT
ref|YP_519534.1|      LVREQIEMKLRS-----KLEHG-----SIKDLIGHSEYKKYINVTAEAAAT
ref|ZP_01372264.1|    LVREHIEMKLRS-----KLEHG-----GIKDLIGHSEYKKYINVTAEAAAT
ref|ZP_01370818.1|    WVREQIEVKLRS-----KLEHG-----SIKDLIGHSEYKKYLNVSAAEAAAT
RAAC03825             AFRESALEEAAETGLKALAESLAKPKRGRKPTDKGVMLKVADLLTRKGVEAFFQVDYKDG
                        .** : . : * : * : ***: . : : :* : .

ref|YP_516922.1|      IN----TDKCLKQAAVFDGLYILQTNTELPTEEVATAYRDLWQIERAFRNKSTLDRPVY
ref|YP_520815.1|      IN----TDKCLKQAAVFDGLYILQTNTELPTEEVATAYRDLWQIERAFRNKSTLDRPVY
ref|YP_519534.1|      MN----TDKCLKQAAVFDGLYILQTNTELPTEEVATAYRDLWQIERAFRNKSTLDRPVY
ref|ZP_01372264.1|    IN----TDKCLKQAAVFDGLYILQTNTELPTEEVATAYRDLWQIERAFRNKSTLDRPVY
ref|ZP_01370818.1|    IN----TDKCLKQAAVFDGLYILQTNTELPTEEVATAYRDLWQIERAFRNKSTLDRPVY
RAAC03825             LTYRDEDAITKEALRDGKFLIRTNLPAADVAQSYKTLMGIERAFHQIKNFLDVGPIY
                        :. * : : * : ** : : :***:***: :** :* : *****: :. ** :* *

```


[illegible]

FIG. 128

```
ref|ZP_01592598.1|-----RYDRKATIVTSQLPIKAWYDAMQDPTLADAILDRL-----
RAAC02717      MNKLGGAGRNLPAAPRYALNSTCVASQLPIEAWYDTFADPTVADAVLDRLAKCSQALVEK
ref|ZP_01002160.1|-----RYERKSTVITSQLPIEKWYDIADPTLADAILDRL-----
ref|YP_771767.1|-----RYEVGSTLITSQLPIDAWHDVIGEPTFADAILDRL-----
ref|YP_001229345.1|-----RHGITSTIIASQIPTEKWHDAIGDPTIADAVLDRLVHNAHMITMK
ref|YP_245447.1|-----SLLITSQYPQEKWYELFADPTIADAILDRIIHKSHTLQLK
                  :  : ** * . *:: : : **.**:***:

ref|ZP_01592598.1|-----
RAAC02717      AITKELIEGRRN
ref|ZP_01002160.1|-----
ref|YP_771767.1|-----
ref|YP_001229345.1|-----
ref|YP_245447.1|      GESMRKVRARRS
```

FIG. 129

ref|YP_145872.1|
ref|YP_001124150.1|
ref|NP_240902.1|
ref|NP_466224.1|
ref|ZP_02850412.1|
RAAC01155

INATINPRTEINRECRLISTERIAMNCYTGENESREFZPRECMBINATINPRTEINRECR

ref|YP_145872.1|
ref|YP_001124150.1|
ref|NP_240902.1|
ref|NP_466224.1|
ref|ZP_02850412.1|
RAAC01155

LISTERIAMNCYTGENESREFZPRECMBINATINPRTEINRECRLISTERIAMNCYTGEN

ref|YP_145872.1|
ref|YP_001124150.1|
ref|NP_240902.1|
ref|NP_466224.1|
ref|ZP_02850412.1|
RAAC01155

ESSREFZPRECMBINATINPRTEINRECRLISTERIAMNCYTGENESHBPREFZPRECMB

ref|YP_145872.1|
ref|YP_001124150.1|
ref|NP_240902.1|
ref|NP_466224.1|
ref|ZP_02850412.1|
RAAC01155

INATINPRTEINRECRLISTERIAMNCYTGENESFSL-REFZPRECMBINATINPRTEIN

ref|YP_145872.1|
ref|YP_001124150.1|
ref|NP_240902.1|
ref|NP_466224.1|
ref|ZP_02850412.1|
RAAC01155

RECRLISTERIAMNCYTGENESLREFZPRECMBINATINPRTEINRECRLISTERIAMNC

ref|YP_145872.1|
ref|YP_001124150.1|
ref|NP_240902.1|
ref|NP_466224.1|
ref|ZP_02850412.1|
RAAC01155

YTGENSEFSRLR-REFZPRECMBINATINPRTEINRECRLISTERIAMNCYTGENESFSL-

ref|YP_145872.1|
ref|YP_001124150.1|
ref|NP_240902.1|
ref|NP_466224.1|
ref|ZP_02850412.1|
RAAC01155

REFZPRECMBINATINPRTEINRECRLISTERIAMNCYTGENESFSL-REFZPRECMBIN

ref|YP_145872.1|
ref|YP_001124150.1|
ref|NP_240902.1|
ref|NP_466224.1|
ref|ZP_02850412.1|
RAAC01155

ATINPRTEINRECRLISTERIAMNCYTGENESFSL-SPQYXRECRLISMRECMBINATIN

ref|YP_145872.1|-----
ref|YP_001124150.1|-----
ref|NP_240902.1|-----
ref|NP_466224.1|PRTEINRECRSPQWRECRLLISMFRECMBINATINPRTEINRECREMBCADRECRLISTER
ref|ZP_02850412.1|-----
RAAC01155-----

ref|YP_145872.1|-----
ref|YP_001124150.1|-----
ref|NP_240902.1|-----
ref|NP_466224.1|IAMNCYTGENESGBAATRECMBINATINPRTEINRECRLLISTERIAMNCYTGENESSTRB
ref|ZP_02850412.1|-----
RAAC01155-----

ref|YP_145872.1|-----
ref|YP_001124150.1|-----
ref|NP_240902.1|-----
ref|NP_466224.1|FGBEALRECMBINATINPRTEINRECRLLISTERIAMNCYTGENESSTRAFGBEBARECMB
ref|ZP_02850412.1|-----
RAAC01155-----

ref|YP_145872.1|-----
ref|YP_001124150.1|-----
ref|NP_240902.1|-----
ref|NP_466224.1|INATINPRTEINRECRLLISTERIAMNCYTGENESGBEBARECMBINATINPRTEINRECR
ref|ZP_02850412.1|-----
RAAC01155-----

ref|YP_145872.1|-----
ref|YP_001124150.1|-----
ref|NP_240902.1|-----
ref|NP_466224.1|LISTERIAAMNCYTGENESSGBEBARECMBINATINPRTEINRECRLLISTERIAMNCYTGE
ref|ZP_02850412.1|-----
RAAC01155-----

ref|YP_145872.1|-----
ref|YP_001124150.1|-----
ref|NP_240902.1|-----
ref|NP_466224.1|NESGBEBARECMBINATINPRTEINRECRLLISTERIAMNCYTGENESFSLN-GBEBAREC
ref|ZP_02850412.1|-----
RAAC01155-----

ref|YP_145872.1|-----
ref|YP_001124150.1|-----
ref|NP_240902.1|-----
ref|NP_466224.1|MBINATINPRTEINRECRLLISTERIAMNCYTGENESFSLN-GBEBARECMBINATINPR
ref|ZP_02850412.1|-----
RAAC01155-----

ref|YP_145872.1|-----
ref|YP_001124150.1|-----
ref|NP_240902.1|-----
ref|NP_466224.1|EINRECRLLISTERIAMNCYTGENESFGBEBARECMBINATINPRTEINRECRLLISTERIA
ref|ZP_02850412.1|-----
RAAC01155-----

ref|YP_145872.1|-----YPEPLSK
ref|YP_001124150.1|-----YPEPISK
ref|NP_240902.1|-----YPEPIAK
ref|NP_466224.1|MNCYTGENESHBPBSCRESIGNIFICANCEE-IDENTITIESPSITIVESGAPSYPEPITK
ref|ZP_02850412.1|-----YPEPIAK
RAAC01155-----MYGYPEPVAR

****:::

ref|YP_145872.1|
ref|YP_001124150.1|
ref|NP_240902.1|
ref|NP_466224.1|
ref|ZP_02850412.1|
RAAC01155

LIDSEFMKLPGIGPKTAARLAFHVLAMKEDTVLEFAKALVDVKRHHYCTICGHITDTPC
LIDSEFMKLPGIGPKTAARLAFHVLAMKEDTVLEFAKALVDVKRHHYCTICGHITDTPC
LIEGFMRLPGIGPKTASRLAFFVLEMKEDDVLDFAKALVNVKRKLTYCSVCHNITDTPC
LIDSEFMKLPGIGPKSAARLAFYVLDKEDDVLDFAKALVDKRNLSFCSVCGHITDKDPC
LIDAFTRLPGVGPKTAARLAFHVLKEDDVIDFAKALVSVKRNLTYSVCCNITDTPC
AIEHFMKLPGVGPKTAARLAFHVLSEADVKAFKALIDLKTGLTECAVCCNITEASPC
*: * :***:***:*.***.* * * * : * : * : * : * : *

ref|YP_145872.1|
ref|YP_001124150.1|
ref|NP_240902.1|
ref|NP_466224.1|
ref|ZP_02850412.1|
RAAC01155

YICKDERRDRTTICVVQDPKDVIAERMKEYNGLYHVLHGAISPMEGIGPEDIKIAELLT
YICKDERRDRTMICVVQDPKDVIAEMKKEYNGLYHVLHGAISPMEGIGPEDIKIAELLA
RICEDSKRDESVICVVQDAKDVIAMEKKEYHGKYHVLHGAISPMDGIGPEDIKIPELIK
YICADTSRDRSVICVVQESKDVIAEMKMRDFHGLYHVLHGTISPMDGIGPEDINIPDLLK
RICQDKTRDNVICVVQESKDLVAMERTKEFQGFYHVLQGAISPIEGIGPDQIKIAELLR
AICRDPRRDRVICVVQEPDVIAMERTREYHGLYHVLHGAISPMEGVGPQDIRIRELVT
* * * * . * * * : : : * : : : * : : : * : : : * : : * : :

ref|YP_145872.1|
ref|YP_001124150.1|
ref|NP_240902.1|
ref|NP_466224.1|
ref|ZP_02850412.1|
RAAC01155

RLQDETQEVILATDPNIEGEATAMYISRLKPTGIKVTRIAHGLPVGGDLEYADEVTLS
RLQDETIQEVILATDPNIEGEATAMYLRLKPTGIKITRIAHGLPVGGDLEYADEVTLS
RLQDDTIQEVIVATNPNTIEGEATAMYISRLVKPTGIKVTRIAHGLPVGGDLEYADEVTLS
RLQDDTIEEVILATNPNVEGEATAMYISRLKPSGIKVTRIAHGLPVGGDLEYADEVTLS
RLSDERVQELILATNPNIEGEATAMYISRLVKPFGIRVTRIAHGLPVGGDLEYADEVTLS
RVGENEIEEVILATNPNVEGEATAMYISRLKPFQVKLTRIAHGLPVGGDLEYADEVTIA
*: : : : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *

ref|YP_145872.1|
ref|YP_001124150.1|
ref|NP_240902.1|
ref|NP_466224.1|
ref|ZP_02850412.1|
RAAC01155

KALEGRREL
KALEGRREL
KAMEGRREL
KAMEGR--
KALEGRREL
KALEGRRAL
.**

FIG. 130

| | |
|--------------------|--|
| dbj BAA94830.1 | ----- |
| RAAC03145 | ----- |
| ref ZP_01666433.1 | ----- |
| ref YP_001213263.1 | ----- |
| ref YP_387030.1 | PSASEPRTEINDESULFVIBRIDESULFURICANSGREFYPPUTATIVETRANSPSASEP |
| ref YP_387237.1 | ----- |
| | |
| dbj BAA94830.1 | ----- |
| RAAC03145 | ----- |
| ref ZP_01666433.1 | ----- |
| ref YP_001213263.1 | ----- |
| ref YP_387030.1 | RTEINDESULFVIBRIDESULFURICANSGREFYPPUTATIVETRANSPSASEPRTEIND |
| ref YP_387237.1 | ----- |
| | |
| dbj BAA94830.1 | ----- |
| RAAC03145 | ----- |
| ref ZP_01666433.1 | ----- |
| ref YP_001213263.1 | ----- |
| ref YP_387030.1 | ESULFVIBRIDESULFURICANSGBABBPUTATIVETRANSPSASEPRTEINDESULFV |
| ref YP_387237.1 | ----- |
| | |
| dbj BAA94830.1 | ----- |
| RAAC03145 | ----- |
| ref ZP_01666433.1 | ----- |
| ref YP_001213263.1 | ----- |
| ref YP_387030.1 | IBRIDESULFURICANSGBABBPUTATIVETRANSPSASEPRTEINDESULFVIBRIDE |
| ref YP_387237.1 | ----- |
| | |
| dbj BAA94830.1 | ----- |
| RAAC03145 | ----- |
| ref ZP_01666433.1 | ----- |
| ref YP_001213263.1 | ----- |
| ref YP_387030.1 | SULFURICANSGBABBPUTATIVETRANSPSASEPRTEINDESULFVIBRIDESULFUR |
| ref YP_387237.1 | ----- |
| | |
| dbj BAA94830.1 | ----- |
| RAAC03145 | ----- |
| ref ZP_01666433.1 | ----- |
| ref YP_001213263.1 | ----- |
| ref YP_387030.1 | ICANSGBABBPUTATIVETRANSPSASEPRTEINDESULFVIBRIDESULFURICANS |
| ref YP_387237.1 | ----- |
| | |
| dbj BAA94830.1 | ----- |
| RAAC03145 | ----- |
| ref ZP_01666433.1 | ----- |
| ref YP_001213263.1 | ----- |
| ref YP_387030.1 | GBABBPUTATIVETRANSPSASEPRTEINDESULFVIBRIDESULFURICANSGBABBP |
| ref YP_387237.1 | ----- |
| | |
| dbj BAA94830.1 | ----- |
| RAAC03145 | ----- |
| ref ZP_01666433.1 | ----- |
| ref YP_001213263.1 | ----- |
| ref YP_387030.1 | UTATIVETRANSPSASEPRTEINDESULFVIBRIDESULFURICANSGBABBPUTATIV |
| ref YP_387237.1 | ----- |

dbj|BAA94830.1|
RAAC03145
ref|ZP_01666433.1|
ref|YP_001213263.1|
ref|YP_387030.1|
ref|YP_387237.1|

PGVPRWEAKS-VVEVRVGM DGQVWLWD--QGRWPCVETQ--ATQTPAPTTAKKEAAPA
QHRETLAPRS-TVEVRVTTNGELWIVAG-QGRLYRLEPCS--KSPAVKRREDATPTPSR
-GLPAYFPDT-TVLVHVHPDGKLRASYPKQNVVYDLTRVS--EPVRTKSQPKKAGPPP
GKTMPIAPKA-KITVLKSPQFGLKVQYG--SSIYDIEILE--QLP-PKDIAPGQPRQPR
RGRTRLLRRGKAVTVVEKLDDEL MALADGEAFVLVPVVTGKSSEEPASEEKS LARGA QER
RGRTRLLRRGKAVTVVEKLDDEL MALADGEAFVLVPVVTGKSSEEPASEEKS LARGA QER
: * : :

dbj|BAA94830.1|
RAAC03145
ref|ZP_01666433.1|
ref|YP_001213263.1|
ref|YP_387030.1|
ref|YP_387237.1|

SPRKPAANHPWRK-----
QPYKPPADHPWRRMTLGRPKPKLPTPSAE AISP
TPRRPAPDHPWRK-----
KPVKPAENHPWRKTKTTTFP-----
EQTSPAAESSW-----
EQTSPAAESSW-----
* . : . *

FIG. 131

```

ref|YP_148969.1| -----YLQLAHNEWDPKAKYAKAKVIYSFGREDEVDRAVLER
ref|YP_001126171.1| -----YLQLAHNEWDPKAKYAKAKVIYSFGREDEVDRAVLER
ref|YP_146154.1| -----
ref|YP_146741.1| -----MYIRRVTRKNKDGTTVAYLQLAHNEWDPKAKYAKAKVIYSFGREDEVDRAVLER
RAAC03325 MSYHVDMYIRVIRRNKNGSVTGYVQLAHNYRDPQTGQPKAKVLYTFGREDEMDLEALRR
ref|ZP_02130848.1| -----MYIRTISRKNKDGSKVEYVQLAHNYRDPKSKQARAEVLYSFGRKDQLDMEAIRR

ref|YP_148969.1| LAKSISRFLSPEQAWIEITLTGEVSDDFQFQSSKRLGGAWLLDQLWRQLGLGEILHSLFA
ref|YP_001126171.1| LAKSISRFLSPEQAWIEITLTGEVSDDFQFQSSKRLGGAWLLDQLWRQLGLGEILHSLFA
ref|YP_146154.1| -----
ref|YP_146741.1| LAKSISRFLSPEQAWIEVEKLTGEASDDFQFQSSCKHLGGVWLLDQLWRQLGLGEILHSLFT
RAAC03325 LAQSIHRFVG--DEFTSGRGQSEAIQTTLDS-RPMGGAYLLDELWRQLELDEVLRERLV
ref|ZP_02130848.1| LAKSVERFLAKTGDVETQCKLQFPGEDVRFVESRPMGGVFLKKIWDRLRISECLDKALA

ref|YP_148969.1| SRHHQIPLERLIFAMVANRALHPSSKLAMEEWVEKDVH IPHLPQVASHQLYRAMDELLAV
ref|YP_001126171.1| SRHHQIPLERLIFAMVANRALHPSSKLAMEEWVEKDVH IPHLPQGASHQLYRAMDELLAV
ref|YP_146154.1| -----LERLIFAMVANRALHPSSKLAMEEWVEKDVH IPHLPQVASHQLYRAMDELLAV
ref|YP_146741.1| SRHHQISLERLIFAMVANRALHPSSKLAMEEWVEKDVY IPHLPQVASHQLYRAMDELLAV
RAAC03325 DRKFKAVERVIFAMVANRALAPSSKLAMEEWVEREVALPGMTELD VWQAYRAMDVLHDV
ref|ZP_02130848.1| DRQYTAPIGDVAFAMVANRALAPDSKLAVEDWAAKDVHLELDQPLKVQHLYRAMDFLEN
: : ***** *:*****:*. :*: : : ***** *

ref|YP_148969.1| QPELERQVFHAVADLLNLEVDLIYFDTTSSYFEVDPSETPEGESLRKQGFSGDKRDPDLVQ
ref|YP_001126171.1| QPELERQVFHAVADLLNLEVDLIYFDTTSSYFEVDPSETPEGESLRKQGFSGDKRDPDLVQ
ref|YP_146154.1| QPELERQVFHAVADLLNLEVDLIYFDTTSSYFEVDPSETPEGESLRKQGFSGDKRDPDLVQ
ref|YP_146741.1| QPELERQVFHAVADLLNLEVDLIYFDTTSSYFEVDPSETPEGESLRKQGFSGDKRDPDLVQ
RAAC03325 TEELQYEVFRRVSNLLNLDVLLFFDFTTSTYFETE--DEPE-DGLRRKGYSKDHDPDLFQ
ref|ZP_02130848.1| QEAIQKEVFWSTANLLNLEVDLVFFDFTTSTYFERD---EEDDEGLKRYGHSKDKRKDLFQ
: : ** :*:*****:*****:*** : : :*: ***** * **

ref|YP_148969.1| IVIGLAVTREGVPIRAWVWPGNTMDMTVIKQVKQDLIGWKLGRVISVMDRGFSSEENLRI
ref|YP_001126171.1| IVIGLAVTREGVPIRAWVWPGNTMDMTVIKQVKQDLIGWKLGRVISVMDRGFSSEENLRI
ref|YP_146154.1| IVIGLAVTREGVPIRAWVWPGNTMDMTVIKQVKQDLIGWKLGRVISVMDRGFSSEENLRI
ref|YP_146741.1| IVIGLAVTREGIPIRAWVWPGNTMDMTVIKQVKQDLIGWKLGRVISVMDRGFSSEENLRI
RAAC03325 VVIGLAVTRDGI PVRCWTWPGNTADMSVVEEVKQGLIGWRLGRVITVVDGRGVSESNLRI
ref|ZP_02130848.1| VVGLAVTKEGLPIRSWVFPNTPDVNTVEQIQKEMNDWKLGRVWVWAMDRGMTSEENRAI
*:*****:*.*****:*.***** *:.....: : :*.*****: :****: ** *

ref|YP_148969.1| LQQAGGHYIVGEKMRSGKAAVEEALNRRGRYQQVRENLHIKEIIVGDGEARQRYVLVYNP
ref|YP_001126171.1| LQQAGGHYIVGEKMRSGKAAVEEALNRRGRYQQVRENLHIKEIIVGDGEARQRYVLVYNP
ref|YP_146154.1| LQQAGGHYIVGEKMRSGKAAVEEALNRRGRYQQVRENLHIKEIIVGDGEARQRYVLVYNP
ref|YP_146741.1| LQQAGGHYIVGENMRSGKAAVEEALSRRGRYHEVDENLHIKEIIVGDGEARQRYVLVYNP
RAAC03325 LQRAGGHYIAGEKMNSGKSAVEVALARPGRFRELRPNLKVKEVVGDGEARVRYVLAFNP
ref|ZP_02130848.1| LQRGGGNYILGEKLR-GSNMSKAVLSPGRFTTVRDNL EIKVTAGDGACRRRYVIVRNP
**:.**:* **:. * : * ** : : **:.** : *** * **:. **

ref|YP_148969.1| SEAEQRKEREKLLSKEELEGRLQLPNEAHHKATCRLRSHPSYGYLRQLKDGTLRID
ref|YP_001126171.1| SEAEQRKEREKLLSKEELEGRLQLPNEAHHKATCRLRSHPSYGYLRQLKDGTLRID
ref|YP_146154.1| SEAEQRKEREKLLSKEELEGRLQLPNEAHHKATCRLRSHPSYGYLRQLKDGTLRID
ref|YP_146741.1| GEAEQRKERETLLESKEELEGRLQLPNEAHHKATCRLRSHPSYGYLRQLKDGTLRID
RAAC03325 EEAKRDEARREAMLRELRLMELERLKEIQGEAHTKAHCRLASHPTFKRYLQDRWGNLRID
ref|ZP_02130848.1| KQVKRDQATRERLIRAEQIEAIGDLTGKKHTKAACALLSHRSMGKYVRELKSGKLKIN
: :*: . ** :. : **: : * : * ** * ** : :*: : **:*:

ref|YP_148969.1| KQAVREAKEYDGKYLIRTSDDTLAEDVALGYQLVDIEQAFRTLKSTLELRPMYHRLED
ref|YP_001126171.1| KQAVREAKEYDGKYLIRTSDDTLAEDVALGYQLVDIEQAFRTLKSTLELRPMYHRLED
ref|YP_146154.1| KQAVREAKEYDGKYLIRTSDDTLAEDVALGYQLVDIEQAFRTLKSTLELRPMYHRLED
ref|YP_146741.1| KQAVRDAKEYDGNYLIRTSDDTLAEDVAIGYQLVDIEQAFRTLKSTLELRPMYHRLED
RAAC03325 PEAVRQAAHLDGKYLIRTSDDTLSTEDVALGYQLLMVESAFRTLKTTLDIRPMYHRKDE
ref|ZP_02130848.1| KAKITEEEKLDGKYLLSCSDDTLSP EIALGYQLLEVERAFRTLKSTLDLRPVYHRKDE
: : : **:*: *****:*****: * *****:*****:*** :

```

ref|YP_148969.1|
ref|YP_001126171.1|
ref|YP_146154.1|
ref|YP_146741.1|
RAAC03325
ref|ZP_02130848.1|

RIRAHVLLSWLALLLVRIVEIRTHESWPKVRDECERIMLGHFSSKNGDLYQRTeltaKQA
RIRAHVLLSWLALLLVRIVEIRTHESWPKVRDECERIMLGHFSSKNGDLYQRTeltaKQA
RIRAHVLLSWLALLLVRIVEIRTHESWPKVRDECERIMLGHFSSKNGDLYQRTeltaKQA
RIRAHVLLSWLALLLVRIVEIRTHESWPKVRDECERIMLGHFSSKNGDLYQRTeltaKQA
RIRSHVLLCWLALLLVRIA EVQTGR TWS DIRSHMQAMHQVTKSTPDGMVVQRTETT EVQR
RIRSHVTL CWLALLLVRLIELETGMTWNQVRRILERLHMGEFFLNNSRILQRTeltaQDQN
: ** *.**: *:.* :* .:* : : :. : **** * *

ref|YP_148969.1|
ref|YP_001126171.1|
ref|YP_146154.1|
ref|YP_146741.1|
RAAC03325
ref|ZP_02130848.1|

LFLAALGLEPPPKILGIHPRT---
LFLAALGLEPPPKILGIHPRT---
LFLAALGLEPPPKILGIHPRT---
QLFAALGLEPPPKILGIHPR----
EILRALRIKEPPRIKVEPRTRGL
KLLKKLKIKPPPLIKKID-----
:: * :: ** * :.

FIG. 132

```

ref|ZP_02851608.1|      MNLIDKKKKVREMMKEGKLDVNDIQDLLKEQFGELIEEMLEGELDHELGYSKYDYREKET
ref|YP_001307815.1|      MSLSRDELVKLILSNSDIKTTEDIQNTLKD LFGGVLQOMLESEMESHLYAKHDYENKNT
ref|ZP_02326599.1|      MGLWTKQQLREFIKENNLVTAQDAQNALKELFAETIQEMLEAELDTHLGYEKHEVKAKKT
RAAC03376                MELLKSKEQIRQLIRDGKLDIHDVQSMKDLFASTIQEMLEAELDTHLGYAKYDAKHKDT
ref|YP_001664274.1|      MSLLTKEQLRNFISENNIQSIPDLYTSLKNLFKDTIQEMLEAELSTELGYERYDKKDKDT
ref|YP_001319448.1|      MSTVSKKVLREMITGGDLKTAGDLQSYLKELFKDTLQEMLEAEIESDLGYEKGRKNKNT
*      .: :: :: .:      *      **: *      :*:***.*: .*** : : .*.

ref|ZP_02851608.1|      TNSRNGK-REKQLKSNYGNLEIEVPRDREGFEFEPQVVKKNQRDVSSIDDQVLSMYAKGMT
ref|YP_001307815.1|      SNSRNGK-STKTMKSNLGLFDLDVPRDREGSFEPQVVKKHQTDVSHLES AVIGMYAKGMT
ref|ZP_02326599.1|      PNSRNGR-SHKT VVSEYGEQQIAVPRDRMGFEFEPVVKKHQSNVTGIEDQIVALYAKGVS
RAAC03376                DNARNGHGAKRTVQSELGDIDIALPRDRKGFEFEPVVKKHQSNVTGIEDQIVALYAKGVS
ref|YP_001664274.1|      QNSRNGY-TQKT VTKTQFGEMEIDI PRDRQGEFEPKIVPKYKRDISGIEEKVIALYARGMS
ref|YP_001319448.1|      QNRRNGY-SEKTVKSKFGEMEIEVPRDRNGFEFEPVVPKNKRDISGIEEKVISLYARGMS
*      ***      : : :: *      : : :*** *.* ** * : : .: .: .:***.*:

ref|ZP_02851608.1|      VRDIQTHLQELYGVDASPTLISGITDKVIPLIKQWNRPLSRIYAHVMDAVHYKVRQDG
ref|YP_001307815.1|      TRDIATQINDIYGMDASPTLISNITDKVIPMLKEWQSRPLESIYPIIFMDAIHFVKRKDN
ref|ZP_02326599.1|      TREIQDHLQNLGYIEVSPTLISNVTNKIVPLIKQWNRPLQSVYAVVFLDAIHFKVKQDG
RAAC03376                TRDIQDHLQQIYGVDMSPTLVSNLTDRLLPRIQEWQNRPLHPLYTVVFLDAIHFKVREEG
ref|YP_001664274.1|      TRDIHDQIKDLYGIELSAEMVSKITERIVPEIKQWNRPLEKIYTFIFMDAIHYKVRTDG
ref|YP_001319448.1|      TRDIHDQIQDIYGIELSAEMVSKITDKVIPQVKEWQNRALAIYPFVMDAIHYKVREDG
.*.*      :*:***.*: * . : * :*:***.*:***.*.*      :*. :*:***.*: :.

ref|ZP_02851608.1|      RIVNKAAYMAIGIDLDGMKDVLGWIWIGENESSKYWLKIINELKNRGVEDILIVSIDGLKG
ref|YP_001307815.1|      AIVSKAAYAVIGVNLGKDKVLGIWIGASESPKYWLLVLNKLNRGINDILIAACVDGLNG
ref|ZP_02326599.1|      AIVSKAAYMIGIDLDGNKDVLMWIGENESSKFVLSVLNDLKNRGVQDILITCVDNLNG
RAAC03376                RLVSKAAYMIGIDIEGQKDVLGWIWIGSESSKFVLSVLNDLKNRGVQDVLVFDNLNG
ref|YP_001664274.1|      HIINRAAYVVLGVTIEGIKDVLGWIWIGENESSRFLVGLNKLNRGVEDVIVFSDGLTG
ref|YP_001319448.1|      QIKSKAAYVVLGIAMDGMKDILGIWIGSESSKFVLSVLNDLKNRGVNDVLIFSVDGLAG
: .:*** .: : : * :*:***.*:***.*.*:***.*.*      :*:***.*: :. *.*

ref|ZP_02851608.1|      FEDAIHAVYPQTEIQSCI IHQIRNSTRYISYKDRKEFCSDLKNVYRAPTEEVALIELDKL
ref|YP_001307815.1|      FKEAIKAVFPNTEIQRCI IHQIRNSSKYSYKDLKAFNTDLKS VYKAPTEDEVALAELDNL
ref|ZP_02326599.1|      FSEAITASYPKTEIQKCI IHQIRNSTRYVS YKDLKKVTADLKPIYKASTEEAAVLELDRF
RAAC03376                FSEAITAACFPQSDVQKCIHVQIRNLSRYVS YKDFKAVAAALKPIYQAPTEEAALMELDQF
ref|YP_001664274.1|      IKEAIQAAFPKSEIQRCI IHQLRNCFKYVS YKHLKEFSKDFKAVYQSANEIARDEFEKL
ref|YP_001319448.1|      MKEAIQASFPKSEIQRCV IHQLRYSFKYVNYKDRKEFAKDFKEVYTAVNEKAGHEKLMEL
.:*** * :*:***.*:***.*.*:***.*.*      :*:***.*: :. *.*

ref|ZP_02851608.1|      EEKWGDKEYEISIRSWRNNDKLSVMFKYPEEVRKLIYTNNSMESYNRQLRKVTKS KSIFF
ref|YP_001307815.1|      EEKWGGKYLIAIKSWRNNDKLSVMFKYPEEVRKLIYTNNSMESYNRQLRKVTKS KSIFF
ref|ZP_02326599.1|      EEVWGSKYPLIIRSWRNNDKLSVMFKYPEEVRKLIYTNNSMESYNRQLRKVTKS KSIFF
RAAC03376                ERGWGARYPLCVKSWRDNWTELATFYRYPVEMRRIMYTTNIIIEGYHRQLRKATKGKSMFP
ref|YP_001664274.1|      KNKWQNLYPYAIKSWENNWDVLS PFYKFPPEVRKIMYTTNIIIEGFHRQLRKVTKS KSIFF
ref|YP_001319448.1|      ENKWGKKYPYAIKSWDANWDVLS PFYKFPPEVRKIMYTTNIIIEGLHRQFRKVTKS KSIFF
.: * * :*:***.*:***.*.*:***.*.*      :*:***.*: :. *.*

ref|ZP_02851608.1|      TDESLLKMLYLATMDITKKWMTMRKNWAQILGQLSIYFEGRI-----
ref|YP_001307815.1|      NDESLLKILYLATIDITKKWTQGIKQWQAQILAQLSIFFEGR-----
ref|ZP_02326599.1|      SDEALLK-----
RAAC03376                NDEALLKMLYLATMELTRKWTMRVANWG TILGQLAIYFGDRVTPYIP
ref|YP_001664274.1|      SDEALEKMLYLVTMNVLKKWTVRYKNWDIVLNQLIIMYPGRLEKYL-
ref|YP_001319448.1|      SDQALEKMLFLASQNMKKWTLSHRNWDIVLNQLMIFFEGRLT-----
.*:*** *

```

FIG. 133

```

ref|YP_001516732.1| -----KTRLNGICPYFTMFPLDFPYSILEE
ref|YP_318565.1| -----
ref|YP_076090.1| -----PHLRNALCPYFTMFPLDFPTRYLAQ
RAAC02657 MRKRRIAMHVDPNRTDRGAAASQMEIWSKRRSSDPRLRINAISPYFTMFPLAFPLRVLAG
ref|YP_001185431.1| -----
ref|ZP_01061333.1| -----QVTPWADKSDSDPLHSLCSYLGAFFPSLAKYFIKY

ref|YP_001516732.1| HSGRGEWVLDPFCEGRGTTIYASRLLGMPISIGIDSSPVATAISEAKLVNIKPGHIVSTAIK
ref|YP_318565.1| -----
ref|YP_076090.1| -ARPGEWVLDPFCEGRGTTNYAARLHGLPTVGIDSSPVAVAIARAKLVQVTPDELIAEAEER
RAAC02657 -AERG-WVLDPFCEGRGTTNFAARLAGFPTVGVDINPIAVAIQAQAKLASTTVHGVVERCQQ
ref|YP_001185431.1| -----IIDPFCEGRGTSMFARAKLGLKAWGIDSSPVATAIARAKLASCSKEDIIDLARD
ref|ZP_01061333.1| FTDENDLVFDPFSGRGTTILESRIILNRKISIGSDLNPIALALS KAKSHKLKKKDIIDRIDE

ref|YP_001516732.1| ILKNAEESDIPTG---EFWELAFHKNVLSLCKFRE-SFLRNCRSDSRKALRAIILGA
ref|YP_318565.1| -----
ref|YP_076090.1| ILAG-PEPADVPEG---HFWDLAYHPTLLDICRLRE-ALLRDCEAPVRIALRGLILGI
RAAC02657 ILES-EAPGDVPDG---EFWSWCFHQKTLEDVCRLRQ-ALS-DVQTEEDIVLRALVLGV
ref|YP_001185431.1| LLQN--AATDMPES---EFFSKLYTSQTLKDVICALREGLLSLAHETDASVMLRALVLGA
ref|ZP_01061333.1| LESDYDYALYLPEAQSESEDIHLIFHQATIAQLCYLKD--ILLTSKSDIDEFLIGAILGI

ref|YP_001516732.1| LHGPRPKSKQS--YFSNQSQRTYAPKPNYAVNYWKRKGLLPEEVDVIEIIREKADRYFGL
ref|YP_318565.1| -----
ref|YP_076090.1| LHGPRTRRAPS--YLSNQMPRTYATKPDPAVRFWVRGMRPEEVDVLDVLARRARYSFAA
RAAC02657 LHGPRNKGLPS--YLSNQMPRTYATKPDAAVRYWKSRIHPVYVDVLDVVRRAEHVLSH
ref|YP_001185431.1| LHGPLNKSOLDTAIYFSNQMPRTFASKPDYSVRYWDQRSIVPPAISVLDVLKRLDR-IPT
ref|ZP_01061333.1| MHGGERKDGTSG-YLSISMPNTFMSPEYVRRFVQTKELNRVKNRVFDILKEKVERVFSK

ref|YP_001516732.1| EKSQGMG---KIIVGDSRNNK---YFQKIQSQVNWVITSPPPYGMSSYIPDQWLRSWFLG
ref|YP_318565.1| -----KFDWIIITSPPPYGLRTRYLPDQWIRNWFGL
ref|YP_076090.1| LPPFVPG---CILQADSREPG---AIPVVDGGYRWVITSPPYLGMRSYQPDQWLRWFLG
RAAC02657 VPDKVPD---FVRLADSRTLA---PEDFAMR-FRYVITSPPYFGMKTYVSDHWRHWFGL
ref|YP_001185431.1| LDEQFTGSFHQIIIEGDSQLEI---TRAKVADDFSIVITSPPPYGMKTYVQDQWIRNWFGL
ref|ZP_01061333.1| HKSPEKES--YIFECDAKEISKSEKLKKYQGVLDLITSPPYLGIVNYAKQNWIRSWFLD
::***** *: . * ::*: * ***.

ref|YP_001516732.1| DSSQVNYSSPEQLNHLSPDHFSSALKKVVWENSSDICAQNATLVIRFGSINTRNVD-ALEI
ref|YP_318565.1| GPSVVDYDTSEGQLSHNGRDRFIADLRQVWANVGARCRPGATLVIRFGSIGDRLVEDPAQL
ref|YP_076090.1| GEPTVRYIHEGQVSHLQG-EYQRDLATVWRNVAARCRPGTRLIIRFGCLPTLPCN-PKQV
RAAC02657 GPPYVDYASDPQLGRPSLLAFIEGLREVWTRVSQVCEPGALLVVRFGAIPSYEVD-PIEV
ref|YP_001185431.1| GASTVDYTSQPQLEHGGVTTFAESLGKVVWRNMADTRADSLRMFIRFGIIPSAKID-AKKI
ref|ZP_01061333.1| SDP-----
. .

ref|YP_001516732.1| IKKSLC--QTRWKIKNIKSAGFASQGRRQAIHINSKVSNPREEYDIWANL---
ref|YP_318565.1| VSASLE--ETGWVTTEILHADNAAARGRRQADTFHRKRSAPYDEVVWA-----
ref|YP_076090.1| LEETLARSQVPRIVEVTGAGIPRHRRAQVQFEKTRP-AHEEIDLVAVLGE-
RAAC02657 LRASLS--GTPWKERRLCDAGSASNGRRQADQFQFVRSRATLEVLDYASLEEA
ref|YP_001185431.1| MQMSLEASGKEWKIISIRSAETAAGKRQATQMKSSS-AAIEYDFH-----
ref|ZP_01061333.1| -----

```

FIG. 134

| | |
|--------------------|--|
| ref YP_001505049.1 | ----- |
| ref YP_481774.1 | ----- |
| ref YP_482514.1 | ----- |
| ref NP_215436.1 | MYCBACTERIUMTUBERCULSISFREFZPHYPTHETICALRESLVASEMYCBACTERIUM |
| ref YP_001680037.1 | ----- |
| RAAC01373 | ----- |
| | |
| ref YP_001505049.1 | ----- |
| ref YP_481774.1 | ----- |
| ref YP_482514.1 | ----- |
| ref NP_215436.1 | TUBERCULSISSTRHAARLEMREFZPHYPTHETICALRESLVASEMYCBACTERIUMTUB |
| ref YP_001680037.1 | ----- |
| RAAC01373 | ----- |
| | |
| ref YP_001505049.1 | ----- |
| ref YP_481774.1 | ----- |
| ref YP_482514.1 | ----- |
| ref NP_215436.1 | ERCULSISHRAEMBCABPSSIBLERESLVASEMYCBACTERIUMTUBERCULSISHRVGB |
| ref YP_001680037.1 | ----- |
| RAAC01373 | ----- |
| | |
| ref YP_001505049.1 | ----- |
| ref YP_481774.1 | ----- |
| ref YP_482514.1 | ----- |
| ref NP_215436.1 | AAKISRESLVASEMYCBACTERIUMTUBERCULSISDCCEMBCADPSSIBLERESLVASE |
| ref YP_001680037.1 | ----- |
| RAAC01373 | ----- |
| | |
| ref YP_001505049.1 | ----- |
| ref YP_481774.1 | ----- |
| ref YP_482514.1 | ----- |
| ref NP_215436.1 | MYCBACTERIUMBVISAFEMBCALPSSIBLERESLVASEMYCBACTERIUMBVISBCGST |
| ref YP_001680037.1 | ----- |
| RAAC01373 | ----- |
| | |
| ref YP_001505049.1 | ----- |
| ref YP_481774.1 | ----- |
| ref YP_482514.1 | ----- |
| ref NP_215436.1 | RPASTEURPGBEAYHYPTHETICALPRTEINTBCGMYCBACTERIUMTUBERCULSISCG |
| ref YP_001680037.1 | ----- |
| RAAC01373 | ----- |
| | |
| ref YP_001505049.1 | ----- |
| ref YP_481774.1 | ----- |
| ref YP_482514.1 | ----- |
| ref NP_215436.1 | BEBAHYPTHETICALRESLVASEMYCBACTERIUMTUBERCULSISSTRHAARLEMGBAB |
| ref YP_001680037.1 | ----- |
| RAAC01373 | ----- |
| | |
| ref YP_001505049.1 | ----- |
| ref YP_481774.1 | ----- |
| ref YP_482514.1 | ----- |
| ref NP_215436.1 | QISRESLVASEMYCBACTERIUMTUBERCULSISHRAGBABRHYPHETICALRESLVAS |
| ref YP_001680037.1 | ----- |
| RAAC01373 | ----- |

MYCROBACTERIUM TUBERCULOSIS FSCRESIGNIFICANCEE-IDENTITIESPSITIVE

-----MDNDVHL

```
ref|YP_001505049.1|
ref|YP_481774.1|
ref|YP_482514.1|
ref|NP_215436.1|
ref|YP_001680037.1|
RAAC01373
```

```
-----MNLKEWAESQGVAYVTAQRWFHAGKLPVPARKVG-GLILVG-EPDRPATSGTVVV
-----MNLKEWAESQGVAYVTARRWYAAGKLPVPARKVG-GLILVG-EPDQPTGDGLTAV
-YDVRVNLKEWAAANGVGYTTARRWYRDGLLPVPARKVG-GLVLVD-ESTVPAGRPFVAVV
SGAPSMNLDWAESVGVNRHTAYRWRFREGTLFVPAERVG-RLILVKTAASASAAAAGVVL
-----MKLSEWAKKGITYRTAWQFWFKAGKLPVPAVMQPTGTILIQ--EGG-KHEGKVAL
VYFISMKLSDWARKNGISYKTARWVKEGRMPVFPEQTSPSGTILVH--EPEPSTTNNAVAL

:::*:*:*:*:*:*:*:*:
```

```
ref|YP_001505049.1|
ref|YP_481774.1|
ref|YP_482514.1|
ref|NP_215436.1|
ref|YP_001680037.1|
RAAC01373
```

[illegible]

```
ref|YP_001505049.1|
ref|YP_481774.1|
ref|YP_482514.1|
ref|NP_215436.1|
ref|YP_001680037.1|
RAAC01373
```

```

IVVEHRDRFARFGAEYVEAALAANGRLLLVDPAEVDDDLVRDITEILTSMCARLYGRR
IVVEYRDRFARFGAEYVEAALSAQGRLLLVDPGEVDDDLVGDVTEILTSLCARLYGRR
IVVEHRDRFARFGAEYVEAALSAQGRLLLVDPGEVDDDLVGDVTEILTSLCARLYGRR
IVVEHRDLARFVGEHLEAALSAQGRRIVVADPGETDDLVCDMIEVLTGMCARLYGRR
IVVEHRDRLMRFGFEYVESALVAQGRRIVVVDQSELKDDLVQDMIEVLTSFCARLYGRR
IVVEHRDRLMRFGFEYVEAALAAQGRRIIVVEPGEVKDDLVDQDMIEVLTSFCARLYGRR
****.***.***.***.***.***.***.***.***.***.***.***.***.***.***.***

```

```
ref|YP_001505049.1|
ref|YP_481774.1|
ref|YP_482514.1|
ref|NP_215436.1|
ref|YP_001680037.1|
RAAC01373
```

```
AANRRARRAV-----
AVNRRATRAV-----
AANRAARAV-----
ARNRAMRAVTEAKREPGAG--
AANKAKKAMEAMQCED-----
ARHRAKRALEVLEREDSSGVSV
*  ::*  ::
```


FIG. 135

```

ref|NP_756021.1|      DNAADENINEMETHYLASEESCHERICHIACLICFTGBABEDNAADENINEMETHYLASE
ref|NP_709160.1|      -----
ref|YP_405147.1|      -----
ref|YP_001680296.1|   -----
RAAC00337             -----
ref|YP_342400.1|      -----

```

```

ref|NP_756021.1|      ESCHERICHIACLIUTIGBABGDNAADENINEMETHYLASEESCHERICHIACLIGBABD
ref|NP_709160.1|      -----
ref|YP_405147.1|      -----
ref|YP_001680296.1|   -----
RAAC00337             -----
ref|YP_342400.1|      -----

```

```

ref|NP_756021.1|      NAADENINEMETHYLASEESCHERICHIACLIAPECSRESIGNIFICANCEE-IDENTI
ref|NP_709160.1|      -----
ref|YP_405147.1|      -----
ref|YP_001680296.1|   -----
RAAC00337             -----
ref|YP_342400.1|      -----

```

```

ref|NP_756021.1|      TIESPSITIVESGAPSRAFLKWAGGKYPLDDIKRHLPKGECLVEPFVAGSVFLNTDFS
ref|NP_709160.1|      -----RAFLKWAGGKYPLDDIKRHLPKGECLVEPFVAGSVFLNTDFS
ref|YP_405147.1|      -----RAFLKWAGGKYPLDDIKRHLPKGECLVEPFVAGSVFLNTDFS
ref|YP_001680296.1|   -----RPFLKWAGNKYRIIHRIRAVLPAGNRLIEPFVGSAAVFLNTSFK
RAAC00337             ---MSRLVRPRGEAAGMKPFLKWAGGKYRLLPYIQRALPPGRRRIEFPVGSAGVFLNTSYD
ref|YP_342400.1|      -----RPFLKWAGNKYRLLTRIISVLPFGKRLIEPFAGSAAFLFLNAEYE
                        :.*****.*. :. *  ** *. *:***.*.:***:..

```

```

ref|NP_756021.1|      RYILADINSDILSYLNIVKMRTDEYVQAARELFVPETNCAEVYYQFREEFNKSQDPFRRA
ref|NP_709160.1|      RYILADINSDILSYLNIVKMRTDEYVQAARELFVPETNCAEVYYQFREEFNKSQDPFRRA
ref|YP_405147.1|      RYILADINSDILSYLNIVKMRTDEYVQAARELFVPETNCAEVYYQFREEFNKSQDPFRRA
ref|YP_001680296.1|   ENLLADCNGDLIRLYQTVKTGGEDFIAYCRSFFTEENNRPERYYALREIFNTTDDGDLKS
RAAC00337             AYVLS DINADVIALYRVLSLHGEAFIEACRRLFVPENNCPEVYYALRDEFNASRDLVRRRA
ref|YP_342400.1|      RYWVNDINPDIALYQILQKEGKEFIHYAGCLFTHNNTPKAYYRLRARFNTTADVAEKA
                        : * * *: * *. :. . :. :*. .*. :. ** : * ** : * :.

```

```

ref|NP_756021.1|      VLFLYLNRYGYNGLCRYNLRGEFNVFPFGRYKKPYFPEAELYHFAEKAQNAFFYCESYADS
ref|NP_709160.1|      VLFLYLNRYGYNGLCRYNLRGEFNVFPFGRYKKPYFPEAELYHFAEKAQNAFFYCESYADS
ref|YP_405147.1|      VLFLYLNRYGYNGLCRYNLRGEFNVFPFGRYKKPYFPEAELYHFAEKAQNAFFYCESYADS
ref|YP_001680296.1|   ALFIYLNHRHSYNGLCRYNASGKFNAPAGRYKKPYFPERELRFFMAKARDAVFLCQDFTQT
RAAC00337             ALFVYLNHRHGYNGLCRYNAEGRFNVFPFGRYKRPYFPEREMRYFCEKARRAEFLCVDREV
ref|YP_342400.1|      ALFVYLNHRHGYNGLCRYSGSGIFNVFPFGRYQRPYFPSKEMMAFHIKARRARFTCLDFRKY
                        .**.:***.:*****. * **.* ***.:***. *: * **: * * * .: .

```

```

ref|NP_756021.1|      MARADDASVVYCDPPYAPLSATANFTAYHTNSFTLEQQAHAEIAEGLVDRHIPVLISNH
ref|NP_709160.1|      MERADDASVVYCDPPYAPLSATANFTAYHTNSFTLEQQAHAEIAEGLVERHIPVLISNH
ref|YP_405147.1|      MARADDASVVYCDPPYAPLSATANFTAYHTNSFTLEQQAHAEIAEGLVERHIPVLISNH
ref|YP_001680296.1|   MGEARPGDVVYCDPPYVPLSRTSNFTGYSAGGFGEAEQWKLAELARQLADRGVSVLISNH
RAAC00337             MRMAEPGDVVYCDPPYVPLSKTANFRQYAAQGFGEAEQDLARIAEELADRGIPVLISNH
ref|YP_342400.1|      LARIRRGITIVYADPPYTPLSQTACFTHYSGNGFGPEEQKALTQSAQRLAKRGIPVLISNH
                        : . :*.*****.*. * * * . * : * *:. * .*. :.*****

```

```

ref|NP_756021.1|      DTMLTREWYQRAKLHVVKVRRSISNNGGTRKKVDELLALYKP---
ref|NP_709160.1|      DTMLTREWYQRAKLHVVKVRRSISNNGGTRKKVDELLALYKP---
ref|YP_405147.1|      DTMLTRKWYQRAKLHVVKVRRSISNNGGTRKKVDELLALYKP---
ref|YP_001680296.1|   DTDFTQKAYDAARIERFPVQRFVSCDGANRGKAGELLALF-----
RAAC00337             LTPFTERAYRRACRQLQVRRRAISCGARRGYAPEVLALFAPRSR
ref|YP_342400.1|      DTPSVRQAYRNAQLTGFSVTRLISCKGTQRTPA NEVLALF-----
                        * ... * *. . * * :...* * . *.*:

```

FIG. 136

```

emb|CAD18993.1|
ref|ZP_00995572.1|
RAAC00506
ref|YP_001115956.1|
ref|ZP_02134324.1|
ref|YP_805441.1|

MTNRTAVLSDIHGVLPALAEVLAEPVRA-ADRVVLTGDIACGPQ-PAEVLDLLTALGD-
---RVAVLSDVHGVLPVLEAVLAEPDVAS-ADLIVVTGDHTAGPM-PVETLDALTALDG-
MPARVAALYDIHGNLAALDAVLQEVERAG-ADALVFGGDLAFGPQ-PRQVLERVMSLGG-
---KIAALSDIHGNLAALDAVLDDVRRRG-ADVIVNLGDIVSGALHPAETADRLIALD--
---RIALISDIHGNLDALETVLKDMKSAGGIDQIICLGDVANLGPDPSCSVVARLRDVN--
---KIAALYDIHGNYPALKEVLKQVKKLS-PDLVVLGGDLIAGPM-PLETSLSLKHVSTT
: * : *: * . * . * : : * * * . : :

emb|CAD18993.1|
ref|ZP_00995572.1|
RAAC00506
ref|YP_001115956.1|
ref|ZP_02134324.1|
ref|YP_805441.1|

-RVTWVAGNADRELVE---FRRGVRETIPDPGPWAARQLRPDHLELLASLPLSVRLPVA
-RVVLVRGNADREAVA---IRSG--DDSPHPEALWAAHALREDQIAVLSALPHVPTIDIA
-PVSFVRGNTDREVAERHGVEQGVRGWIAEVN-AWCHDQLTPDQREFLLSQKAHVTLHVE
--LPTIRGNHERQLLER-----TRDAMGLSDRWADTLHERHRAWLAALPERATLG--
--CPCVLGNHDAFLFDRDLHQYTEVDYIIIEAVDWCRDQLTQDDIDWMGDFHRVLEVLMD
FKTVGIMGNNDQDIVDIYAKKRVGLSRKATEQLTWIANQLSYKQVSFLRNLLPSISIG--
: * : * * . : :

emb|CAD18993.1|
ref|ZP_00995572.1|
RAAC00506
ref|YP_001115956.1|
ref|ZP_02134324.1|
ref|YP_805441.1|

GLGTVLFCCHATPRDDE---EVVVVDSRPDRWREVDGLG--PDVDAVICGHTHMP-FVR
GFGEVLFCGHSRDDD---EVVLVDSPTAWQNALSIVP--DTVQTIVCGHTHMP-FVR
GLGDVLFFVHGSPRSDE---EAIRRDTPAEILPMVAHV--ESI--IVCGHTHIQ-FDR
--DDVLLVHGTPASDLDFLETVTTRDGCRAATPDEIAQRAEDEPASLILCGHTHVPRVAR
KAGSLIAFHGSPLSHM---DNILSRTPDAEVDRLIQGLN---GEILACGHTHIQ-MLR
---NYFFCHAVKNDNT---TVFSPQONKAYIETLFKGVQ---ESYIICGHTHIQ-FEL
: * . . . : * * * :

emb|CAD18993.1|
ref|ZP_00995572.1|
RAAC00506
ref|YP_001115956.1|
ref|ZP_02134324.1|
ref|YP_805441.1|

LAHGRLVVNPGSVGMPYG-RSGAHWALLGPG-VDLRRTFYDTDAAIARLTRDCGYPAIAE
LVDRTVINPGSIGMPYG-RAGGAWALLDRGQVSLRHTIDLDVACARIVAESGYPDRAA
MVAGKRNVNAGSVGLPSA-ARGACWALIGPE-IELRETLYDFERAAAEIR-KSGAPKAGE
LADGRTIVNPGSVGLPA-----
QHRGRLIINPGSVGLP-----
SLPNKKIINAGSIGMEFSNQFGAQWLWDDNRIEYKRTIFN-QQAAIQLISQTEYPFKNE
: : : * . * : *

emb|CAD18993.1|
ref|ZP_00995572.1|
RAAC00506
ref|YP_001115956.1|
ref|ZP_02134324.1|
ref|YP_805441.1|

WAD-----
WADE-----
FADDILNPPVAGP
-----
-----
FIANNLRSTIS--

```

FIG. 137

```

ref|YP_001488901.1|-----RLPPQNI EAEQAVLGAVFLEPSALT LASEVLIPEDFYRM
ref|YP_001423303.1|-----RLPPQNI EAEQAVLGAVFLOPSALT LASEVLIPDDFYRM
ref|NP_391924.1|-----RLPPQNI EAEQAVLGAI FLQPSALT LASEVLIPDDFYRM
ref|NP_244897.1|-----RTPPQNI EAEQAVLGAI FLEDHALVTASERLAPEDFYRA
ref|ZP_01695451.1|-----RIPPQNI EAEQAVLGAI FLEPSAFITASEILTAEDFYRN
RAAC00022MRNMAAEETLWRDREA EEAALRMPHNLEAEQAVLGAMLISPD AVEAMEILEADDFYRS
          * *: *: *****: :. * . * * * .: ****

ref|YP_001488901.1|SHQKIYNAMLVLGDRGEPVDLVTVTSELANT-DLLEEVGGISYLTDIANSVPTAANIEYY
ref|YP_001423303.1|SHQKIYNAMLVLGDRGEPVDLVTVTSELANT-DLLEEVGGVSYLTDIANSVPTAANIEYY
ref|NP_391924.1|SHQKIYNAMLVLGDRGEPVDLVTVTSELANT-DLLEEVGGISYLTDIANSVPTAANIEYY
ref|NP_244897.1|AHQRIYQVMLDLAEKGEPVDLVTVTAELQDR-KALDDVGGVAYLGLDANAVPTAANVEYY
ref|ZP_01695451.1|AHQKIFQVFAKLNDAGKAVDLVTVAEELSAT-RQLEDAGGLSYLSELAASVPTAANIAYY
RAAC00022AHQAIYRAIREVYEAGDPVDIITVASRLRTYGDVLDVAVGGPEYLAADLAAMPTALHVVHY
          **: *: :. : : * . *: *: : * * : * : * : : * : : *

ref|YP_001488901.1|AKIVEEKSILRRLIRTATTIAQDGYTREDEVEDLLSDAEKTIMEVAQRKNSGAFQNIKDV
ref|YP_001423303.1|AKIVEEKSILRRLIRTATSIAQDGYTREDEVEDLLSDAEKTIMEVAQRKNSGAFQNIKDV
ref|NP_391924.1|AKIVEEKSILRRLIRTATTIAQDGYTREDEVEDLLSEAEKTIMEVAQRKNTSAFQNIKDV
ref|NP_244897.1|SKIIVEEKSLLRRLIRVATNIAQEGYASEEEVDAILDDAEKTI LDVAQRKNSSAFISIKDV
ref|ZP_01695451.1|AGIVA EKALLRRLIRTATHIAQEGYTREDDVDELLDEAERSIMEVAQRKNAGDFKHIDV
RAAC00022AEIVREKALLRRIISAGTRIAEAA YEPDASAMEVLADAERLVLELSQHQRTRDFTHISDV
          : * * *: *: *: * . * *: : * : : * : : * : : * : : * * *

ref|YP_001488901.1|LVQTYDNIEQLHNRKGDITGIPTGFSELD RMTAGFQRNDLIIVAARPSVGKTAFALNIAQ
ref|YP_001423303.1|LVQTYDNIEQLHNRKGDITGIPTGFIELDRMTAGFQRNDLIIVAARPSVGKTAFALNIAQ
ref|NP_391924.1|LVQTYDNIEQLYNRKGDITGIPTGFTELD RMTAGFQRNDLIIVAARPSVGKTAFALNIAQ
ref|NP_244897.1|LVETYDKIEMLQNKGEITGIPTGFKDLDRMTAGFQRNDLIIVAARPSVGKTAFALNISQ
ref|ZP_01695451.1|LVSTYDNIEMLHNRKGDVTGIPTGFYELDRMTAGFQRNDLIIVAARPSVGKTAFALNIAQ
RAAC00022LQTTFERIEQLYESDGNITGVPTGYSDDLDRMTSGFQKSDLIIVAARPSVGKTAFALNIAQ
          * *: :. * * : * : *: *: *: : *****: *****: *****: *****: *

ref|YP_001488901.1|NVATKTDESVAIF SLEMGEQLVMRMLCAEGNINAQNLR TGNLTEEDWGKLT MAMGSLN
ref|YP_001423303.1|NVATKTDESVAIF SLEMGAELVMRMLCAEGNINAQNLR TGNLTEEDWGKLT MAMGSLN
ref|NP_391924.1|NVATKTDESVAIF SLEMGAELVMRMLCAEGNINAQNLR TGNLTEEDWGKLT MAMGSLN
ref|NP_244897.1|NVATKTDENVAIF SLEMGAQLVQRMLCAEGNIDAQRMRTGALTEEDWQKLT MAMGSLAR
ref|ZP_01695451.1|NVGTKTEENVAIF SLEMGAELVMRMLCAEGNIDAQRLRTGALTDEEDWRKLT MAMGSLN
RAAC00022NVAVRAGLPVAIF SLEMSKDLVQRMLCAEAFIDGHKLRNGTLDDEEDWPKLSMGVTTLN
          ** . : : * * * * * . * * * * * . *: :. : * * : * * * * : : * : .

ref|YP_001488901.1|SGIFIDDTPGIRVSEIRSKCRRLKQENGLGMILIDY LQLIQSGRSSDNRQQEVSEISRA
ref|YP_001423303.1|SGIFIDDTPGIRVSEIRSKCRRLKQENGLGMILIDY LQLIQSGRSSDNRQQEVSEISRA
ref|NP_391924.1|SGIYIDDTPGIRVSEIRAKCRRLKQESGLGMILIDY LQLIQSGRSSDNRQQEVSEISRE
ref|NP_244897.1|AGIYIDDTPGIKVNEIRAKCRRLKQEKGLGMILIDY LQLIQGNRSGENRQQEVSEISRS
ref|ZP_01695451.1|SGIYIDDTPGVRVTEIRSKCRRLKQEHGLGMVVIDY LQLIQGSARSREN RQQEVSEISRS
RAAC00022SPIYIDDTPGITVPEMRSKLRLKLEHGLGFVVIDY LQLIHGRMAGENRQQEISDISRS
          : *: *****: * *: * * * * * * * *: : *****: * : : *****: * : *

ref|YP_001488901.1|LKSLARELEV FVIALS QLSRGVEQRQDKRPMSDIRESGS IEQDADIVAFLYRDDYDKE
ref|YP_001423303.1|LKALARELEV FVIALS QLSRGVEQRQDKRPMSDIRESGS IEQDADIVAFLYRDDYDKE
ref|NP_391924.1|LKSIARELQVPVIALS QLSRGVEQRQDKRPMSDIRESGS IEQDADIVAFLYRDDYDKE
ref|NP_244897.1|LKGLARELEV FVIALS QLSRGVESRQDKRPMSDIRESGS IEQDADIVAFLYRDDYDKE
ref|ZP_01695451.1|LKALARELEV FVIALS QLSRGVEQRQDKRPMSDIRESGS IEQDADIVAFLYREDY YEQD
RAAC00022LKQLARELEV PIALAQLSR SVEQRQDKRPMSDIRESGS IEQDADVVAFLYRDDY YNPD
          ** : * * *: *: *: * * . * * . * * * * * : * * * * * : * * * * * : :

ref|YP_001488901.1|SENKNIIEIII AKQRNGPVGTVSLAFVKEYNKFVNLER-----
ref|YP_001423303.1|SENKNIIEIII AKQRNGPVGTVSLAFVKEYNKFVNLER-----
ref|NP_391924.1|TENKNIIEIII AKQRNGPVGTVSLAFVKEYNKFVNLER-----
ref|NP_244897.1|TENQNIIEIII AKQRNGPVGTVELAFIKEYNKFVNLD RHDE-
ref|ZP_01695451.1|TENKNIIEIII AKQRNGPVGTVQLAFIKEYNKFVNLER-----
RAAC00022TENPNVVEVII AKQRNGPTGKIELVFLKNFNKFVNLERAHAE P
          : * * *: : *: *****: * :. * *: : *****: *

```

FIG. 138

```
ref|YP_001213441.1| -----MLNKVILIGRLTQDPELRYTPGGVAVARFTLAVNRARLNKQGERETDFIDVVV
RAAC00027 MNAEGTSMNLRVILIGRLTADPELRYTNNGTAVASFTLAVDRMRSGPNGERQTDFFINVVV
ref|YP_077145.1| -----MLNSVVLIGRLTKDPELRYTPSGKAVATLRLAVDRGTVNQQGERETDFIDIVV
ref|NP_244917.1| -----MLNRVVLVGRLLTRDPELRYTPNGVAVANFTLAVNRPFSSNQQGEREADFINCVV
ref|YP_149334.1| -----MINRVILVGRLLTRDPELRYTPSGVAVATFTLAVNRPFSSNQQGERETDFIQCVV
ref|YP_001377189.1| -----MMNRVILVGRLLTKDPELRYTPNGVAVATFTLAVNRAFTNQQGEREADFINCVI
*:* *:*.***** *:***** .* *** : ***:* . :****:***: *:

ref|YP_001213441.1| WQKQAEYCANYIRKGRLLVAVEGRLLQVRSYDDSQGIRRKAAEVVAETVRFLDR-----
RAAC00027 WQKQAEIYAQYLQKGRLLAAVDGRLQIRSYDNRDGQVRVVAEVVAETVRFLDRGPDQAQGS
ref|YP_077145.1| WEKQAEVANYLQKGRLLVAVQGRLLQIRQYTTQDGGQKREKAEVVATTVRFLDSARD-----
ref|NP_244917.1| WRKQAEVANYLLKKGSLAGVDGRLQTRSYDNNEGR-----
ref|YP_149334.1| WRRQAEVANYFLKKGSLAGVDGRLQTRSYENQEGRRVYVTEVVADSVQFLEPKGT-SEQR
ref|YP_001377189.1| WRKQAEVANYLLKKGSLAGVDGRLQTRNYEGQDGRRVYVTEVLAESVQFLEPRNSGGEQR
*.:*** *:.:** *:.:***:*.*. *.* :*

ref|YP_001213441.1| -----
RAAC00027 GYSAAG-----AQTRQQRPTPSSAPPFEDDPFADDSQLIDISEDLLPF
ref|YP_077145.1| -----
ref|NP_244917.1| -----
ref|YP_149334.1| GATAGGYYGDPFFPGQDQNHQYPNEKGFGRIIDDPFANDGQPIDISDDDLPF
ref|YP_001377189.1| GSFNQQPSGAGFGNQGSNPFQSGNSGFTK-NDDPFSNVGQPIDISDDDLPF
```

FIG. 139

```

ref|YP_001255315.1|      -----GAPSIKKRDGQYEPLQVEKTKK
gb|ACA43749.1|           -----IKKRDGQYEPLQVEKTKK
ref|YP_001392092.1|      -----IKKRDGQYEPLQVEKTKK
ref|ZP_02619122.1|      -----IKKRDGQYESLQVEKTKK
ref|YP_074151.1|         -----PERASEFIAELR-----
RAAC01051                MQMVHPSDTSTGLMERTGQGELRSGRAHLEEWRDRVKTFTTVEKRDGRIEPLSLDKIFA
                           : *

ref|YP_001255315.1|      MVKLACEGIEGCDPLELELDSRIQFRDGMTTKEIQRTLIQTAEKVIQNSKDNNGNNIKK
gb|ACA43749.1|           MVKLACEGIEGCDPLELELDSRIQFRDGMTTKEIQRTLIQTAEKVIQNSKDNNGNNIKK
ref|YP_001392092.1|      MVKLACEGIEGCDPLELELDSRIQFRDGMTTKEIQRTLIQTAEKVIQNSKDNNGNNIKK
ref|ZP_02619122.1|      MVKLACEGIEGCDPLELELDSRIQFRDGMTTKEIQRTLIQTAEKVIQNSKDNNGNNIKK
ref|YP_074151.1|         -----PLLHATTDRSQOMDLAARLAAEKTSVEEPD-----
RAAC01051                FLKRGCEAAEGCDPERLAADVWPQLRDMRTSEIARVAIQAAVEKTSVAEPG-----
                           ::      .:      : * *.      .

ref|YP_001255315.1|      TNANWQYVAARLLCFDLYKEAKISRHYNSFGYGDYYQLVKKLVKIK-----LYGEYLIQ
gb|ACA43749.1|           TNANWQYVAARLLCFDLYKEAKISRHYNSFGYGDYYQLVKKLVKIK-----LYGEYLIQ
ref|YP_001392092.1|      TNANWQYVAARLLCFDLYKEAKISRHYNSFGYGDYYQLVKKLVKIK-----LYGEYLIQ
ref|ZP_02619122.1|      TNANWQYVAARLLCFDLYKEAKISRHYNSFGYGDYYQLVKKLVKMK-----LYGEYLIQ
ref|YP_074151.1|         ----WQYVAARLYLQKLYGEAARNRGYHQPGYGDFYELVRQLHAMGDEAGNRVYGAYMIE
RAAC01051                ----WQYVAARLLLFDAYKQASRNRGRHLGYGDFYQLIVELERIG-----RYGTYIRQ
                           ** :*****      * : *      * *.      ***** : * : :      : * * : :

ref|YP_001255315.1|      NYSDEEIKELAKYIVPERDELFNIEGLKLLNDRYLIKGNNGEILELPQERFMTIAMHLAI
gb|ACA43749.1|           NYSDEEIKELAKYIVPERDELFNIEGLKLLNDRYLIKGNNGEILELPQERFMTIAMHLAI
ref|YP_001392092.1|      NYSDEEIKELAKYIVPERDELFNIEGLKLLNDRYLIKGNNGEILELPQERFMTIAMHLAI
ref|ZP_02619122.1|      NYSDEEIKELAKYIVPERDELFNIEGLKLLNDRYLIKGNNGEILELPQERFMTIAMHLAI
ref|YP_074151.1|         AYSEEEIRELGRYIRPERDELFTYVGLLHLADRYLIKGFHGEILELPQERYMHIAMHLAS
RAAC01051                HYSDEEIAELAAAIKPERDYLFNIVGLKQLLDRYAIRNLDNGIMELPQELFMGVSMHLAM
                           ** : ** **      ** ***** ** * ** * * : . . ***** : * : :*****

ref|YP_001255315.1|      PEGDKKVFYAKKFFYDLLSELKVTVATPTLGNAAGTFFYQLSSCFISVVGDNLWSIYDVNQK
gb|ACA43749.1|           PEGDKKVFYAKKFFYDILSELKVTVATPTLGNAAGTFFYQLSSCFISVVGDNLWSIYDVNQK
ref|YP_001392092.1|      PEGDKKVFYAKKFFYDLLSELKVTVATPTLGNAAGTFFYQLSSCFISVVGDNLWSIYDVNQK
ref|ZP_02619122.1|      PEGDKKVFYAKKFFYDLLSELKVTVATPTLGNAAGTFFYQLSSCFISVVGDNLWSIYDVNQK
ref|YP_074151.1|         VESDR-VTWAKRFYDVLRSQEMTVATPTFRNAATPLPQLSSCFIDTVDDSLQSIYDTNQS
RAAC01051                KEKDK-VAWAKRFYDMLSTLQATVATPTLSNARKPFHQLSSCFIDMPEDDLVSIIYATDEA
                           * * : * : ** :***** : * : ***** : * . * ***** : * . * ** * : :

ref|YP_001255315.1|      FAQVSKHGGALGIYTGKIRALNSEIRGHKNASGGVVPWIRLYNDTAVAVDQLGKRKGGAA
gb|ACA43749.1|           FAQVSKHGGALGIYTGKIRALNSEIRGHKNASGGVVPWIRLYNDTAVAVDQLGKRKGGAA
ref|YP_001392092.1|      FAQVSKHGGALGIYTGKIRALNSEIRGHKNASGGVVPWIRLYNDTAVAVDQLGKRKGGAA
ref|ZP_02619122.1|      FAQVSKHGGALGIYTGKIRALNSEIRGHKNASGGVVPWIRLYNDTAVAVDQLGKRKGGAA
ref|YP_074151.1|         FAQVSKHGGGMGIYIGKLRARGSAIRGRKGAAAGVWPVRNYNDTAVAVDQLGARKGAVS
RAAC01051                FARVSKFGGGMGIYVGKIRARGSAIRNHPGASGGVVPWIRNFNNTAVSCNQLGMRAGAAA
                           ** :*** . * . :*** ** : * * * : . : . ** : * * : :***** : *** * * . :

ref|YP_001255315.1|      ITLDIWHKDI FDFLDLKTNNGDDRRKAHDI FPSVSI PDLFMKRLEKRESWSLFDPIYIVEK
gb|ACA43749.1|           ITLDIWHKDI FDFLDLKTNNGDDRRKAHDI FPSVSI PDLFMKRLEKRESWSLFDPIYIVEK
ref|YP_001392092.1|      ITLDIWHKDI FDFLDLKTNNGDDRRKAHDI FPSVSI PDLFMKRLEKRESWSLFDPIYIVEK
ref|ZP_02619122.1|      ITLDIWHKDI FDFLDLKTNNGDDRRKAHDI FPSVSI PDLFMKRLEKRESWSLFDPIYIVEK
ref|YP_074151.1|         IWLDVWHKEIFEFLALKLNNGDDRMRAHDI FPGVCVPDAFMRAVEQDADWHLFCPHEVRT
RAAC01051                VYLDVWHKDI ILDFLQLRTNNGDERMKAHDI FPGVCIPDLFMRVVEERGMYLFCPYEVRK
                           : ** :*** : * : * * : ***** : ***** . : * * : * :      * * * * : * .

ref|YP_001255315.1|      IMGYKLEDYFDDEDRKEFTNKYLECERNNTNIPRDTVPTLDIMKKLMKSAVETGTPFIFFR
gb|ACA43749.1|           IMGYKLEDYFDDEDRREFTNKYLECERNNTNIPRDTVPTLDIMKKLMKSAVETGTPFIFFR
ref|YP_001392092.1|      IMGYKLEDYFDDEDRREFTNKYLECERNNTNIPRDTVPTLDIMKKLMKSAVETGTPFIFFR
ref|ZP_02619122.1|      IMGYKLEDYFDDEDCREFTNKYLECERNNTNIPRDTVPTLDIMKKLMKSAVETGTPFIFFR
ref|YP_074151.1|         VMGFSLEDAGWEEWERRYR-LCVEAAHADKLDATVVRARQVAKAMLKSQYETGGPFLLFFR
RAAC01051                VMGFSLEDWSGEEEFERRY-ACVENP---NLPRVEIPAIEIMKRIMQSAFETGTPFIFFR
                           ** : * * : * * : : * : * : * : * : * : * : * : * : * : * : * :

```

DTVNKANPNKHKGMIIYSSNLCHEIAQNMSESRLLEEEIIDNGYSEIVQRVKAGDMVTCN
DTVNKANPNKHKGMIIYSSNLCHEIAQNMSESRLLEEEIIDNGYSEIVQRVKAGDMVTCN
DTVNKANPNKHKGMIIYSSNLCHEIAQNMSESRLLEEEIIDNGYSEIVQRVKAGDMVTCN
DTVNKANPNKHKGMIIYSSNLCHEIAQNMSESLLEEEIIDENGYSEIVQRVKAGDMVTCN
DTVNRLNPKNHAGMVYCSNLCHEIAQNMSPTRLIEKTDGEG---VITYRWQPGDFVVCN
DTANRLNPKNHAGMVYCSNLCHEIAQNMSPSRRIETEEGG---IITMKTEAGDFVVCN
.*:*****.****** ** ** *::*:*:::*:*:*****

```
ref|YP_001255315.1|
gb|ACA43749.1|
ref|YP_001392092.1|
ref|ZP_02619122.1|
ref|YP_074151.1|
RAAC01051
```

LNSINLSKVKKEE-FNECIPFQIRMLDNVISLNKLPVKEAKVTS DKYRAIGLGTSGYHHF
LNSINLSKVKKEE-FNECIPFQIRMLDNVISLNKLPVKEAKVTS DKYRAIGLGTSGYHHF
LNSINLSKVKKEE-FNECIPFQIRMLDNVISLNKLPVKEAKVTS DKYRAIGLGTSGYHHF
LNSINLSKVKKEE-FSECIPLOIRMLDNVISLNKLPVKEAKVTS DKYRAIGLGTSGYHHF
LASINLSKVHTEEKIAEVVPLAIRMLDNVIDLNHPVPOAKITNNKYRAIGLVGHGYHQM
LSSLNLGRCDRLDSIREIVRRQIRAMDNDVIDLNHYVPVQAAVTNNKYRAVGLGVSGYHQY
* *:***::: : : * : ** : ***** ** : * :. :. *****.***.***:

```
ref|YP_001255315.1|
gb|ACA43749.1|
ref|YP_001392092.1|
ref|ZP_02619122.1|
ref|YP_074151.1|
RAAC01051
```

LANNKIRWESDEHIKVADEIYEEIAYIAIKSSMELAKEKGSYPAFKDESEWETGKYFERRG
LANNKIRWESDEHIKVADEIYEEIAYIAIKSSMELAKEKGSYPAFKDESEWETGKYFERRG
LANNKIRWESDEHIKVADEIYEEIAYIAIKSSMELAKEKGSYPAFKDESEWETGKYFERRG
LANNKISWESDEHIKVADEIYEEIAYIAIESSMELAKEKGSYPAFKGSWETGKYFERRG
LAELGIHWESEDHLRKADEVFELLNYYAVKTSIELAAEKGCYPLCKGSDWETGEYFALRG
LAEKGIPEWSEAHVKHADELFEWINFPAIEASMELAREKGPYPLFEGSDWQGTGRYFDLRG
* *

```
ref|YP_001255315.1|
gb|ACA43749.1|
ref|YP_001392092.1|
ref|ZP_02619122.1|
ref|YP_074151.1|
RAAC01051
```

YNS---ERWKKLQSNIKKYGMRNGYITAIAPTGSTSNIANTTAGIDPVFKKFFMEEKKG
YNS---ERWKKLQSNIKKYGMRNGYITAIAPTGSTSNIANTTAGIDPVFKKFFMEEKKG
YNS---ERWKKLQSNIKKYGMRNGYITAIAPTGSTSNIANTTAGIDPVFKKFFMEEKKG
YNS---ERWKKLQSNIKKYGMRNGYITAIAPTGSTSNIANTTAGIDPVFKKFFMEEKKG
YES---DRWQALRRQAAEVGLRNAYLLAIAPTSSTSLLCGTTASVDPIYDRVYNEGKKD
YRTRPGGPDWDGLRREVAEHGVRNAYLVAIAPTSSTSLIAGSTAGIDPVYARFFLEEKKN
* : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *

```
ref|YP_001255315.1|
gb|ACA43749.1|
ref|YP_001392092.1|
ref|ZP_02619122.1|
ref|YP_074151.1|
RAAC01051
```

SFTPKTAPDLNEENFWYYKEAHTIDQQWSIKACAVRQKHIDQAQSFNLYITPEIKAKEIL
SFTPKTAPDLNEENFWYYKEAHTIDQQWSIKACAVRQKHIDQAQSFNLYITPEIKAKEIL
SFTPKTAPDLNEENFWYYKEAHTIDQQWSIKACAVRQKHIDQAQSFNLYITPEIKAKEIL
SFTPKTAPDLNEENFWYYKEAHTIDQQWSIKACAVRQKHIDQAQSFNLYITPEIKAKEIL
QVIPLAAPGLSPKTYLYYKPAHQIDQTWSIRAAGVRQRHIDQSQSFNLYIRPDIKGRDFL
GVVPEQTAPNLNDKTFWYYKEAHTIDQRWSEIACAARQHVDSQSFSNLYITPGISARQFL
* . * * * . : . * * * * * * * * * * * * . * . * * * * * * * * * * * * . * . : . *

```
ref|YP_001255315.1|
gb|ACA43749.1|
ref|YP_001392092.1|
ref|ZP_02619122.1|
ref|YP_074151.1|
RAAC01051
```

```
NMYIESWKQGVKTIYYVRNKSLEMD--ECTSCSS
NMYMESWKQGVKTIYYVRNKSLEMD--ECTSCSS
NMYMESWKQGVKTIYYVRNKSLEMD--ECTSCSS
NMYMESWRQGVKTIYYVRNRSLEMD--ECTSCSS
NLYMQAWKNGLKTVYYVRSRSLLEVTEAECEACQA
DLYLLAWKRGRLKTVYYVRSKSVEV--EDCVACS
```

MBTULINUMASTRATCCGBABSTRIBNUCLESIDE-DIPHSPHATEREDUCTASEBETASU

BUNITCLSTRIDIUMBTULINUMASTRHALLSCRESIGNIFICANCEE-IDENTITIESP

```
-----IFNENGNR--GTESMINGNTTNLREWNRIKYNWANELYRTMLNNFWIPE
-----IFNENGNR--GTESIINGSTTNLREWNRIKYNWANELYRNMLNNFWIPE
-----IFNEKGQR--GTESMINGNTTNLREWNRIKYSWASDFYRTMLNNFWIPE
SITIVESGAPSIFNEKGQR--GTESMINGNTTNLREWNRIKYSWASDFYRTMLNNFWIPE
-----MEKKKIFNEYGDR--GTDCLIKGNTTNLREWNRIKYDWARLIYRTMLNNFWIPE
----MQLKRTLFLNAEGDRDWGKRRIIGGNTTNMIELNNKIKEYAWKLYRAMNNFWIPE
      : * : * * : * : * : * : * : * : * : *
```

```
EISLNEDVKQFPYLTGGERNAFDKIIISFLNFLDSVQSENLPNISRYITAPEVSSLLNVQT
EISLNEDVKQFSYLTGGERNAFDKIIISFLNFLDSIQSENLPNISRYITAPEVSSILNIQA
EISLNEDIKQFPYLTGGERNAFDKIIISFLNFLDSVQSENLPNISRYITAAEVSSLLNIQT
EISLNEDIKQFPYLTGGERNAFDKIIISFLNFLDSVQSENLPNISRYITAAEVSSLLNIQT
EISLQEDVKQFPYLTGGERNAFDKIIAFLNFLDSIQSENLPNISRYITAPEVSSLLNVQA
EIPLADARQYDLLSDDERRSYNKVSFLIFMDSLTLHNLPNINEYITAPEVNLCLTVHA
* * * * * * * * * * * * * * * * * * * * * * * * * * * *
```

```
FQEEIHAQSYSYILDVTNPITRDKIYDQWREDKNLLERNKFIAGIYQSFNENPTTENFL
FQEEIHAQSYSYILDVTNPITRDKIYDQWREDKNLLERNKFIAGIYQRFNENPTTENFL
FQEEIHAQSYSYILDVTNPITRDKIYDQWREDEHLLERNKFIAGIYEKFNKEPEIHNFL
FQEEIHAQSYSYILDVTNPITRDKIYDQWREDEHLLERNKFIAGIYEKFNKEPEIHNFL
FQEEIHAQSYSYILDVTNPVTRDRIYDIWREDKYLLERNQFIANIYQRFSDPEPSDENFV
FQEAHVHSQSYGYVLDSVVNPATRDQIYNENWRNDDHLLRNQFITDLFEEFIENPTDRNLM
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
```

```
RTIMANYILEGIYFYSGFSFFYTTLARQGKMTATSTIFKYINRDEITHLILFQNI IKELKN
RTIMANYILEGIYFYSGFSFFYTTLARQGKMTATSTIFKYINRDEITHLILFQNI IKELKN
RAIMANYILEGIYFYSGFSFFYTTLARQGKMTATSTIFKYINRDEVTHLVLFQNI IKELKN
RAIMANYILEGIYFYSGFSFFYTTLARQGKMTATSTIFKYINRDEVTHLVLFQNI IKELKN
RVVFANYILEGIYFYSGFSFFYTTLARQGKMTATSTIFKYINRDEITHLVLFQNIIMKELRV
KITVANYILEGVFYAGFGFFFLGRQGMKLGTVSEIKFIQDELTHLAI FRNLYHEIIR
:  * * * * * : * * * * * : * * * * * : * * * * * : * * * * * : *
```

[illegible]

GLEVLYPEITEHPMEWIENFSKLNNTKTDFFEAKVTNYTKAAAFDFFDL
GLEVLYPEITEHPMEWIENFSKLNNTKTDFFEAKVTNYTKAAAFDFFDL
GLKPLYPEINKHMPMEWIDSFSKLNSTKTDFFEAKVTNYTKAAAFDFFDL
GLKPLYPEINKHMPMEWIDSFSKLNSTKTDFFEAKVTNYTKAAAFDFFDL
GLEILYPEITKHMPMEWIESFSNINSTKTDFFEAKVTNYTKAAVFDFFDL
NLPILYPEIKEHPMPWMSFAAINSHKTDFFEQRVTNYTKAGNLNWDDL

* * * * *

FIG. 141

```

ref|YP_148023.1|-----AEWLAQGSIAVPKLLLGHY
ref|YP_001126202.1|-----AEWLAQGSVVVPKLLLDHY
ref|ZP_01697284.1|-----LTVPQLLKKY
ref|YP_001421643.1|-----SIPNLLLTHY
gb|ABN10253.1|-----IPNLLLTHY
RAAC01009MDLWARSRPRSDSVSPHALPNGALSARGGERMEPSGRQGGNSDYLSAPFVAVPCDLLRRF
: * * * :

ref|YP_148023.1|KQLGLGEGELVLLLHMQSFLE-EGVLFPTPAELAERMTVSAAECEMEMVRRLQKGMIAIE
ref|YP_001126202.1|KQLGLSEGELVLLLHMQSFLE-EGIVFPTPAELAEMTVSAAECEMEMVRRLQKGMVIE
ref|ZP_01697284.1|KQLGLNETELVLLQVYSFLK-EGNAFPTPEDLAGRLTIPESMCVSIILRLIQHQFLSIE
ref|YP_001421643.1|RQLGLNETELILLIKIMHLE-KGSYFPTPFELQSGMSISAEECTSCLRMFIQKGLFIE
gb|ABN10253.1|KQLGLNETELILLIKIMHLE-KGSYFPTPNQLQEGMSISVEECTNRLRMFIQKGLFIE
RAAC01009AQLGLHPHELVLVLLQILASQTEGTTELSPELGERCGMSSKEVMACTVERLVTEGFLAIG
**** **::**:: : : * : * : * : : : : : : : *

ref|YP_148023.1|EHTDEQGIRNEKYTLEPLWEKLVHHLTYTQAAQQGELGRQEEEEESLYTVFEQEFGRPLSPF
ref|YP_001126202.1|EHTDERGVRGEKYTLEPLWEQLVHHLTYAQTVDGQIGQQKEEESLYTIFEQEFGRPLSPF
ref|ZP_01697284.1|EGEGEGGILYKYSIMPLWAKLADEFIYEKKQDELEKSVNEETDLYTTFEQEFGRPLSPL
ref|YP_001421643.1|ECEDHNGIKFEKYSIQPLWAKLYDYMQHSQNETQERTSEREQKSLYTIFFFFEGRPLSPL
gb|ABN10253.1|ECEDQNGIKFEKYSIQPLWGKLYEYIQLAQTTQERKAEGEQKSLYTIFFFFEGRPLSPL
RAAC01009ERYDDQGAHVITYFDLQPLWDLKGRRRMVQPPP-----AEKDLVSLFEFEFGRPLSSL
* .. * : : *** : * : . * : *****.:

ref|YP_148023.1|ECETLAMWIDQDGHEPAIIKAALREAVLSGKLNFRYIDRILFEWKNGIRT-----
ref|YP_001126202.1|ECETLSMWIDQDGHEPAIIKAALREAVLSGKLNFRYIDRILFEWKNGIRT-----
ref|ZP_01697284.1|ECETLAMWIDQDGQSPDLIKAALREAVISGKLNFRYIDRILFEWKNGIRT-----
ref|YP_001421643.1|ECETLAIWQDQDQHDAILIKHALKEAVLSGKLSFRYIDRILFEWKNGFKT-----
gb|ABN10253.1|ECETLAIWQDQDQHDAILIKHALKEAVLSGKLSFRYIDRILFEWKNGLKT-----
RAAC01009ECDQLRAWLGEHGYPEWLVVEALKESVLANKYSFRYIDRVLYNWQKNNVRSRQDLEQYRA
** : * * .. : : **:*:*:*.* .*****:*:*:*:*.:

ref|YP_148023.1|-----
ref|YP_001126202.1|-----
ref|ZP_01697284.1|-----
ref|YP_001421643.1|-----
gb|ABN10253.1|-----
RAAC01009QYRERQAQWRGEPQAGQSKPRSRPASPRTAAREPVRDERYASFYELFPD

```


FIG. 142

```

ref|ZP_02175216.1| -----DKVFYPEAGFTKGE
ref|YP_464174.1| -----DKVFYPEAGFTKGE
ref|ZP_02321813.1| -----DKVFYPEAGFTKGE
ref|YP_753805.1| -----AMAGREIK-----LTNLDRLVLPEDGYCKRD
ref|ZP_01575281.1| -----DKLEWPEAGITKLE
RAAC00998 MLEMGSKGRKNGIQVSSLIATGSSAHARREVKRMKRHESPTTLTHGDKIYFPAAGLCKRD
                                         *: : * * * :

ref|ZP_02175216.1| VVDYYRRVAPVLLPHLRDRPLTLKRYPEGVDGPHFYEKRCPRHRPDWFRTEAIWSEGNQD
ref|YP_464174.1| VVDYYRRVAPVLLPHLRDRPLTLKRYPEGVEGPHFYEKRCPRHRPDWFRTEAIWSEGNQE
ref|ZP_02321813.1| VVDYYRRVAPVLLPHLRDRPLTLKRYPEGVDGPHFYEKRCPRHRPDWFRTEAIWSEGNQD
ref|YP_753805.1| LVEYYTAIFPYMLPHLSERPLVFTRYPRGIGESFYQKNAPEGLPQWIKTFTTWAGS-DGD
ref|ZP_01575281.1| YVKTMTKLAPFLIKYSKHRMLTSIRYPHGINDKSFFQKEKPQGTPEWVETVEFN-----
RAAC00998 YMTYLAQIGEHLVRHLRHRPVTLLVRCPDGVAGRRFYQRHLPHPAPPSLPRREVQG-----
      :      :      : : . * : . * * * :      * : : . * * * .

ref|ZP_02175216.1| YIDYCVVDDLSSLVWLASIADLELHPSLSLVDDV--ERPTALVFDLDPGPPADLLACCEV
ref|YP_464174.1| YIDYCVVDDLSSLVWLASIADLELHPSLSLVDDV--ERPTALVFDLDPGPPADLLACCEV
ref|ZP_02321813.1| YIDYCVVDDLSSLVWLASIADLELHPSLSRVDDV--ERPTALVFDLDPGPPAEALLACCEV
ref|YP_753805.1| SKNYVLVQQTVDLMWLANLACIEIHPWLSQINSI--EYPDFIVFDLDPSEQSTFEQVISV
ref|ZP_01575281.1| QKNYINLNSAATLVWLCTQAALHLTSFNVHEKP--NHPSSLVFDLDPDDDLHFEDVAEL
RAAC00998 ERPLIAIPDVETLLYYGNLGAIEFHAGLHLVQGPQAGCPTALTFDLDPSPDRDFERVREL
      : . * : : . . : * : . :      * : : * * * . :      :

ref|ZP_02175216.1| ALLLRLLAALGLEAFPKSGSGKGMQLYVPLSG-ATYADTKPFAHAVARLLERRHPQLVV
ref|YP_464174.1| ALLLRLLAALGLEAFPKSGSGKGMQLYVPLSG-ATYADTKPFAHAVARLLERRHPQLVV
ref|ZP_02321813.1| ALLLRLLAALGLEAFPKSGSGKGMQLYVPLSG-ATYADTKPFAHAVARVLERRHPQLVV
ref|YP_753805.1| ARLLELMDLSLSRVYPKTSAGAKGLHLYLPPIAEGFTYSQIRRVQAQAMAEMVCQVIPDIAT
ref|ZP_01575281.1| AGRIHETLEALGIMDFIKTSGATGLQIFVPVAAKFDYDTARSLNEFFAQYFAEKLSTVT
RAAC00998 ALCLREVLRLGLDGVAKTTGASGLQVLVPLEEPLPYAVTRPVVNFVAAYCASRWPDLAT
* : : . * : :      * : : * : : : : * : : . . * . . .

ref|ZP_02175216.1| ERMAKALRGKVLVDWSQNDPHKTTVCVYSLRARPRPTVSTPLRWAEVEKAVRTRDAGGL
ref|YP_464174.1| ERMAKALRGKVLVDWSQNDPHKTTVCVYSLRARPRPTVSTPLRWAEVEKAARTRDAGGL
ref|ZP_02321813.1| ERMAKALRGKVLVDWSQNDPHKTTVCVYSLRARPRPTVSTPLRWAEVEKAVRTRDAGGL
ref|YP_753805.1| TERALKHRGPRVYLDYLQNLGKTVCAAYSVRPHKGAPCSTPIEWQELES-IRPDQ----
ref|ZP_01575281.1| IERMKKKREGKIYFDWQMQMTGKSMITAYSARAVKSAAVSAPIEWSELND-VRPEM----
RAAC00998 TERRVRHRGRRVYVDAPQHGPTRTLIAAYSVRAVEKALVSAPITWAELEYGVSPDA----
      . * : : * * : : . * * . . * : : * : : .

ref|ZP_02175216.1| VFEAAAVLRRVERVGDLFAPVLTLRQLPA-----
ref|YP_464174.1| VFEAAAVLRRVERVGDLFAPVLTLRQLPA-----
ref|ZP_02321813.1| VFEAGAVLRRVERVGDLFAPVLTLRQLPA-----
ref|YP_753805.1| -FTIKTLPERLQQVGD-----
ref|ZP_01575281.1| -FTLKNIINRLEQKGDIFE-----
RAAC00998 -FDLHVVPDRLERVGDWLELAARHPASRVRAIHDALPPPWRQTRVRSV
      * : * : : **

```

FIG. 143

ref|NP_832076.1|
ref|YP_001645033.1|
ref|NP_844759.1| ELREFZPCGBACTERIALNUCLEIDDNA-BINDINGPRTEINBACILLUSANTHRACISS
ref|YP_001375058.1|
ref|YP_535778.1|
RAAC02359

ref|NP_832076.1|
ref|YP_001645033.1|
ref|NP_844759.1| TRAREFYPDNA-BINDINGPRTEINHUBACILLUSTHURINGIENSISSTRALHAKAMRE
ref|YP_001375058.1|
ref|YP_535778.1|
RAAC02359

ref|NP_832076.1|
ref|YP_001645033.1|
ref|NP_844759.1| FZPDNA-BINDINGPRTEINHUBACILLUSANTHRACISSSTRAREFZPDNA-BINDINGP
ref|YP_001375058.1|
ref|YP_535778.1|
RAAC02359

ref|NP_832076.1|
ref|YP_001645033.1|
ref|NP_844759.1| RTEINHUBACILLUSCEREUSAHREFZPDNA-BINDINGPRTEINHUBACILLUSCEREU
ref|YP_001375058.1|
ref|YP_535778.1|
RAAC02359

ref|NP_832076.1|
ref|YP_001645033.1|
ref|NP_844759.1| SAHREFZPDNA-BINDINGPRTEINHUBACILLUSANTHRACISSSTRAREFZPDNA-BIN
ref|YP_001375058.1|
ref|YP_535778.1|
RAAC02359

ref|NP_832076.1|
ref|YP_001645033.1|
ref|NP_844759.1| DINGPRTEINHUBACILLUSANTHRACISSSTRAREFZPDNA-BINDINGPRTEINHUBAC
ref|YP_001375058.1|
ref|YP_535778.1|
RAAC02359

ref|NP_832076.1|
ref|YP_001645033.1|
ref|NP_844759.1| ILLUSCEREUSWREFZPDNA-BINDINGPRTEINHUBACILLUSCEREUSNVH-REFZPD
ref|YP_001375058.1|
ref|YP_535778.1|
RAAC02359

ref|NP_832076.1|
ref|YP_001645033.1|
ref|NP_844759.1| NA-BINDINGPRTEINHUBACILLUSCEREUSHREFZPDNA-BINDINGPRTEINHUBAC
ref|YP_001375058.1|
ref|YP_535778.1|
RAAC02359

ref|NP_832076.1|
ref|YP_001645033.1|
ref|NP_844759.1|
ref|YP_001375058.1|
ref|YP_535778.1|
RAAC02359

ILLUSCEREUSBBREFZPDNA-BINDINGPRTEINHUBACILLUSANTHRACISTSIAK

ref|NP_832076.1|
ref|YP_001645033.1|
ref|NP_844759.1|
ref|YP_001375058.1|
ref|YP_535778.1|
RAAC02359

VSKII-IGBAAPDNA-BINDINGPRTEINHUBACILLUSANTHRACISSTRAMESGBAAS

ref|NP_832076.1|
ref|YP_001645033.1|
ref|NP_844759.1|
ref|YP_001375058.1|
ref|YP_535778.1|
RAAC02359

DNA-BINDINGPRTEINHUBACILLUSCEREUSATCCGBAATDNA-BINDINGPRTEINH

ref|NP_832076.1|
ref|YP_001645033.1|
ref|NP_844759.1|
ref|YP_001375058.1|
ref|YP_535778.1|
RAAC02359

UBACILLUSANTHRACISSTRAMESANCESTRGBEALDNA-BINDINGPRTEINHUBACI

ref|NP_832076.1|
ref|YP_001645033.1|
ref|NP_844759.1|
ref|YP_001375058.1|
ref|YP_535778.1|
RAAC02359

LLUSCEREUSGGBAATDNA-BINDINGPRTEINHUBACILLUSANTHRACISSTRSTERN

ref|NP_832076.1|
ref|YP_001645033.1|
ref|NP_844759.1|
ref|YP_001375058.1|
ref|YP_535778.1|
RAAC02359

EGBAATDNA-BINDINGPRTEINHUBACILLUSTHURINGIENSISSESVARKNKUKIAN

ref|NP_832076.1|
ref|YP_001645033.1|
ref|NP_844759.1|
ref|YP_001375058.1|
ref|YP_535778.1|
RAAC02359

STR-GBAAUDNA-BINDINGPRTEINHUBACILLUSCEREUSELGBABKHUFAMILYDNA

ref|NP_832076.1|
ref|YP_001645033.1|
ref|NP_844759.1|
ref|YP_001375058.1|
ref|YP_535778.1|
RAAC02359

-BINDINGPRTEINBACILLUSTHURINGIENSISSTRALHAKAMGBEDRDNA-BINDIN

ref|NP_832076.1|
ref|YP_001645033.1|
ref|NP_844759.1|
ref|YP_001375058.1|
ref|YP_535778.1|
RAAC02359

GPRTEINHUBACILLUSANTHRACISSTRAGBEDRDNA-BINDINGPRTEINHUBACILL

-----SISATCC-----
 USANTHRACISSTRAGBEDRDNA-BINDINGPRTEINHUBACILLUSANTHRACISSTRA


```

SCRESIGNIFICANCEE-IDENTITIESPSITIVESGAPSMNKTELKVNVAQSADISQKD
-----MNKTELKVNVAQSADISQKD
SCRESIGNIFICANCEE-IDENTITIESPSITIVESGAPSMNKTELKVNVAQSADISQKD
-----MNKTELKVNVAQSADISQKD
-----NKAALIERVAEKTGLTKKD
-----MNKRDLLRKTAETGLSQKD
          **  *  . . . *  . . . :  :  **

```

```

ASAAVQSVFDTIANALQSGDKVQLIGFGTFEVRERSARTGRNPQTGEEIQIAAGKVPAFK
ASAAVQSVFDTIANALQSGDKVQLIGFGTFEVRERSARTGRNPQTGEEIQIAAGKVPAFK
ASAAVQSVFDTIATALQSGDKVQLIGFGTFEVRERSARTGRNPQTGEEIQIAAGKVPAFK
ASVAVQSVFDTITNALQNGDKVQLIGFGTFEVRERAARTGRNPQTGEEIQIAAGKVPAFK
ATVAVDAVFETIQDALVDGEKVQLIGFGNFEVRERAARKGRNPQTGEEIEIPASKVPAFK
CEAVINTLFDITRIKTVESGEKVQIIGFGTFELRERAARTARNPRTGEAVEVPARRVPAFK

```

AGKELKEAVK-
AGKELKEAVK-
AGKELKEAVK-
AGKELKEAVK-
PGKSLKDAVK-
PGAELKQAVQV
* ** **

FIG. 144

```

ref|ZP_01667008.1| -----EIKWDGIRAVIYLASDNLRILSRNLKDVTRQYPEL
ref|YP_946103.1| -MPNYSFPLATSGTTADLRGDDWLYELKWDGIRAIITGTESRIRLMSRNGNDLTAAYPEL
ref|YP_075626.1| ---PFPPMLLER-AEAPFADDDWVYQVKWDGVRNLTVEGGRVRHWSRRLRERTALFPEF
RAAC00997 -MPFFRPFEPISSTEVPVG-DAWIAQVKWDGVRVAEVSQDGVQIWNHRHRLRTDRYPEI
ref|ZP_02854041.1| MEFFQIQAMEPIPRQSVPTGDDWVYQIKWDGIRILTFFDGNISIKLRTKKGFKRTTEYPEL
ref|YP_753804.1| -----RSMDPINCEQAFDSDDFLYQVKWDGVRMLVAVAGEQVSLLNKRGNLRSRQYPEL
                                :*:***:* : : : : :*:

ref|ZP_01667008.1| MPLTQALAGHSVILDGEIVAFDINGRPS-----FARLQNRMGSLASDRT
ref|YP_946103.1| TDRACWPDGD-FVADGEIVALGKGSRPD-----FGRLQLRMNLVKAAD
ref|YP_075626.1| DGLAAALPGRRVLDGEIIVLQ-DGRPS-----FSAVLERDLAGSTAG
RAAC00997 ASALRPFRGC--AFDGEVVALT-QGRP-----FYRVLRRDRAVASHE
ref|ZP_02854041.1| NELTKEFGNTTWLLDGEIIVANQTDKSDHPKFITTNESCEEVGANFFHVLKRDRTKNPT-
ref|YP_753804.1| QNLPLNIRASTAVLDGEIIVLREGKPS-----FPAVMQRDNCRNPMK
                                ***:.. . * : *

ref|ZP_01667008.1| IARKTTEIPATYIIFDLLYIDGQSTLSLPYIERRRLLEEIDLN---GPNWQTPAYKTG-Q
ref|YP_946103.1| IERARATVPVQMLFDLLYDDGTDLGSLPFHERRERLSGFAERWRVGCPLHLSAVLDH-D
ref|YP_075626.1| AARRR---PATLMLFDLLEWDGRELVDVPLAERLALLEAVVPP---DEAWQVVSSFPQSD
RAAC00997 IEQLVREIPVWYAVFDVVRMGGEWVGWPLAKRLDWMLEHLSG-----VEHVLTSSEGEVD
ref|ZP_02854041.1| -QKLLKTFPVRYKVFDDLMDNQVLVNEPFIKRRELLNKHFNN-----NKFLQIIRSYS
ref|YP_753804.1| IQHLSKTLAISYMFVDDLNLNGRDLRSASLVDRLSQLAELFDN-----QGCLYLVSFSQ
                                : . :*:.. . . . * : :

ref|ZP_01667008.1| GRELLASRRLGLEGIVAKKLDISIYLPKRPPTWLKIKNVRROEFVIGGWLPQGTRTGM
ref|YP_946103.1| VEDLMTSAAELGLEGVMAKKADSRVYIGRRSRSWIKLEQSQEVVVGWRPGAGARAGT
ref|YP_075626.1| GPALYAAAVARGLEGVAKRRASRYVPGARSRDWLKVRRSEMLAVVGYTN----PTGR
RAAC00997 TEALFAATKQLGLEGIVCKRVDSTYAPGGKDGWVVKVHERNAIAAVGGVVY----RDGA
ref|ZP_02854041.1| GLSLFNITSQQNHEGIVAKKSNSIYHPGKKHSEWFKIKHLKYLQAYLGGVIVK----ENQ
ref|YP_753804.1| GTILFDSVQRAGLEGIVAKKNSFYRPGKQHDDWFKIKCRCSQACLVGGYTLR----GKQ
                                * . **::*: * * * : *.** : *

ref|ZP_01667008.1| IGALLLGYDRTPEAKAGQPQRLLFAGAVGTG-FTHHTLKKLQHLK--PLEQAEPP-
ref|YP_946103.1| FGALLLGIPD-----GDKLHYVGRVGSF-KDWQLRDIEMKLE--PLAIVESP-
ref|YP_075626.1| PGGLLLGAYR-----DGRRLYIGRVGSF-LSGADLAAIRHHLPSGFCPLAEIPR
RAAC00997 PNALMLGLFD-----DKARLFHVGNAGAGRLPRRAWRLVERVLS---HETSLQP
ref|ZP_02854041.1| IKSLLLGQKS-----DLADYKLNIGRASTG-LTQKELNMLKDFAFKNRVSNPFGN
ref|YP_753804.1| VNALLLG-----VMRDGSLSFAGKASNG-LDAEQQLQILSETLPRLEVKDSPF--
                                .*:** . * . * : : .

ref|ZP_01667008.1| FAESP-----PLKGAVFVKPALVGEVEFTEWTPGGTLRHP-----
ref|YP_946103.1| FADIPR---EDAAGATWVSPELVAEVTFGEWTPGRLRHPV-----
ref|YP_075626.1| LRERFG---GDPGPVWVPEPLLTVRVSFTEWTEEGRLRDPVVV-----
RAAC00997 FARAP---KAVRTCRFAHPVAVKVRFAEWTPKGTMRHPVLLAELDIPATPAACSFAQA
ref|ZP_02854041.1| LSKVDLELADSKESVVFSLPQLKLQVKYTNWTPGGNLRHPVI-----
ref|YP_753804.1| -----LEPTPAGSHYITPQLAVEVEYLEWTDLSLHLRFVI-----
                                : * : . * : ** : * *

ref|ZP_01667008.1| --
ref|YP_946103.1| --
ref|YP_075626.1| --
RAAC00997 DP
ref|ZP_02854041.1| --
ref|YP_753804.1| --

```

```

---MRVAIYIRVSTEEQAQYGYSIDAQKERLIAYCTSQGWTDYKIYIDDG
MLLMTVALYCRVSTDEQAEHGFSIDNQKERLIAFCKSQGWDQYRLYIDDG
----TVACYIRVSTDEQAEQGISIPAQKSRLLAFCRSQGWNLYGYFYIDDG
-----ALYIRVSTMEQAKEGYSIPAQTDKLKAFAKADMAVAKVYTDPG
-----VAIYVRVSTQDQAEEGYSIEEQIDKLTKYCDIKDWYVYIYKDPG
-----SKDKLSSYCDIKDWNVYKVYTDGG
          :  : *   :  :      . * * *

```

YTGTKMMRPALNRLRIHIEDDKIDLVLVVVYKLDRLSRKQLDVLVYLLEEVFE
 YTGTNMDRPAQRLIKHIKEKKIDTVIVYKLDRLSRKQRDVLVYLLEDVFD
 YSGKDLDRPAMRRLIGDTGAKKFDTVLVLKLDRLSRRQKQDVLFLLEDVFE
 FSGAKMERPALQEMISDIQNKKIDVVLVYKLDRLSRSQKNTLYLIEDVFL
 FTGSTIDRPGMKKLNVDSKQHKFDTVLVYKLDRLSRSQKDTLYLIEDVFA
 FSGSNTDRPALESLIKDAKKRKFDTVLVYKLDRLSRQKDTLHLIEDVFI
 : * * * : : * . * * * * * * * * * : * . : * * : *

KHNVGFSKATEPFETTTTFFGKAMIGILAVFAQLERDMIVERTTIGRRQRV
KNNVVFKSATEPFDSTPLGKAMLGILAVFAQLERDTIIERLSTGQKLRV
PGGVGFSKVTESEFDTTTPFGKAALGMMAVFAQLERETIVERVRMAKKESA
KNNVDIFISMQESFDTSTPPGRATIGMLSVAFAQLERDTERMHMGRTERA
KYNVDFVLSLENFDTSTAFGKAMIGILSVFAQLEREQITERMQMGKVGRA
KNIEFLSLQENFDTSTFPFKAMIGLLSVFAQLEREQIKERMQLGKIGRA

. * * * * . : *

[illegible]

IAEWAAER----TKARVIDHSVIRDILSRPVYTGKLSNAGNVVDGKHEAI
LSDWIESQ----TSARYWDHARIDMLTRETYYTGVMPYGTLRSTEIAPP
IAEELEHTGVPGPTEKWNKAFVRKILTNFVYAGLIRHRENLYPGRHDP
ITKRMVAK----YPDKVKTLTIVKYALTNPPLYIGKISWDGKVVYDGHHSPI
LRDKLNDEG-HLGKDI PWSYRAIRS ILDNPVYAGYMNFKGEVYKGNHDAI
LRDDLNEK---YPKHVPWSYRAVRT ILDNPVYAGFNQYKGEIYPGNHEPI
: : : * * * *

[illegible]

IVHVKRITRKYGKEYLYELYACKNQHVR-----KKDRNNNCSLGYIRREK
FIHVIRKK---GPYVYYLYGCRSQHQRPKKVPRAPKQACTVGYNREE
VKNVWQNHPTNPKKVTRYVVCYSQDRAGG---HMVRNPTCRCGYKHGGA
--EVFRYVSG-GKKYRYNYMCRSVKKMLP---SLVKDWNCKQPSLRQEV
-LEIMLGNIRKDGTRLKKYKCTNRIVKSR--VTFYNDNKKCTSGFYHMND
-LKIMLGVKRKDGSRLLKKECHQRHPRTLRGVTTYNDNKKCDSGFYKDK
:
*
:

```
VEKFVIEQIKSYT--TDDVLIKQITNEKNKFKEN--DESALNNLESQ LKK
LESWVIEQVKRIARLDAEEFEHQIANIETNSDE--TEDTLRLDRKNLS D
VEEQVLRELFCYS--YDRRLRRALAEALAAADNRAFLRELAQARKDLAA
VEKKVIDSLKSLDFKKIERELKQVENKT-----KSKITTIINNQISK
LEDYVINEVSKLQ--KNP---ETIFERTEIKN----ENHTEIYEHRIVV
LEAYVLKEISKLQ--DDADYLDKIFSGDNAET----IDR-ESYKKQIEE
*:  *:  :
```

ref|ZP_02330146.1|
RAAC02419
ref|YP_001210709.1|
gb|AAD26564.1|AF124258_1
ref|ZP_02184702.1|
ref|ZP_01828805.1|

VLTGLENLYDAIESGEIKASSISDRIRKLEEQRDVLENHIDEIKDNT PQN
IEHKLNRWYEAFFEEGTIDATQLRDRIKHLEEEERRRLQTNIEELQDTLVKD
TEKKLERWYDAFEKGALEGDQLTERVKGLHQKAYLRGQIAEMESRLKEG
KHNEKQKILDLYQYGTDFVTMLNERMKKIDNEINALTANIANLEGTKSES
IDKKIKRLSDLYMNDLINLEDMQFKAKALKDEKDHLNKLNHSEHKLENV
LSKKLSRLNDLYIDDRITLEELQSKSAEFISMRGTLETELENDPALRKNK
.. : . : : : . * . : :

ref|ZP_02330146.1|
RAAC02419
ref|YP_001210709.1|
gb|AAD26564.1|AF124258_1
ref|ZP_02184702.1|
ref|ZP_01828805.1|

---ELNIQEFNIFIKEIGEAWDYLTEDQKALIRKAFRSVTLYKD-KEIK
---TQPQRVFN-AAKLIEEAWDYMTFDEKKNVLRAAISYIEIPRKGQPPV
KERQTSVEEMMGILQDFPRIWENATAEERREIVVNTVKAVKV-----
-----LINKLETLKTFNWETETTENKILIIKEFVERIEL-----
-----KKDKAILFLKDIENNINDETYENKKKIINVLIKKVDVKANEIKVI
-----RKADMRKLLN--AEKVFSMDYENQKVLVRRLLINKVKVTAEDIVIN
::: :: . : :

ref|ZP_02330146.1|
RAAC02419
ref|YP_001210709.1|
gb|AAD26564.1|AF124258_1
ref|ZP_02184702.1|
ref|ZP_01828805.1|

VEWN-
IHWNV

WN---
WKI--

FIG. 146

```

ref|ZP_02309926.1|      ----FSKRLSELRKKKGFSQYKLADDELGFSRGQVANYEQGTREPDYQTLKIAEFFNVST
ref|ZP_01941236.1|      ----FSKRLSELRKKKGFSQYKLADDELGFSRGQVANYEQGTREPDYQTLKIAEFFNVST
ref|ZP_01926077.1|      ----FSKRLSELRKKKGFSQYKLADDELGFSRGQVANYEQGTREPDYQTLKIAEFFNVST
ref|NP_469419.1|         ----FSKRLSELRKKKGFSQYKLADDELGFSRGQVANYEQGTREPDYQTLKIAEFFNVST
ref|YP_001111866.1|     -----MTQEQLAQQLGFTRGQVSNYEQGSREPDFETLKKIADFFKVTT
RAAC02417               MMTTFGERLAQLRRSKGLSQYALAEQLKMTRGQIANYEQGTREPDIE TLKKLADFFDVSI
                        ::*  **::*  :*****:*****:*** **::**.*:

ref|ZP_02309926.1|      DYL-LGR---DDNNLADTIAAHID--SNATEEDIKEILAYIEEKR-KEHANEKEINITE
ref|ZP_01941236.1|      DYL-LGR---DDNNLADTIAAHID--SNATEEDIKEILAYIEEKR-KEHANEREINITE
ref|ZP_01926077.1|      DYL-LGR---DDNNLAETIAAHID--SNASEEDIKEILAYIEEKR-KEHVNEEEINITE
ref|NP_469419.1|         DYL-LGR---DDNNLADTIAAHID--SNASEEDMKEILAYIEEKR-KEHANEEDIDITD
ref|YP_001111866.1|      DYM-LGR---TDD-----PTPVDKLIELSALAGQQK-FDPMK-----E
RAAC02417               DFLVLGKPNVSDFNGLTNEVKRTLEALSQMSVEKQQEVADFAEYLR SKEEQPVVEYDVRE
                        *: : ** :      :      .  .  * :      :      :

ref|ZP_02309926.1|      IASKEDDEEINKFVDE--NEDFKVVAARVM-----
ref|ZP_01941236.1|      IASKEDEAVDKFVEE--NEDFKAVAARVM-----
ref|ZP_01926077.1|      IASKEDDAVDKFVEE--NEDFKAVAARVM-----
ref|NP_469419.1|         IAAKKDADVAKFVEE--NPDFKAVAARVM-----
ref|YP_001111866.1|      LPPEAQRSLEDFID-----YLMRK-----
RAAC02417               IAA NMEKALYAHGDEDLIQHFEIMRRVIRRYDERASQDQQNR
                        :...: : : : :      *

```


FIG. 147

```
ref|ZP_02442523.1|      ---FQKRLKELRNAKGTSQIAIAAALGITDRGYRKYEAGDSEPTLSVIALADYFDVSLD
ref|YP_001210714.1|    ---FAKRLSFLITKNKLSKQAVANAINVSRPAVSQFANGENLPSVEKLIADFFDVSLD
RAAC03180               MSSFPERLSELLSATNSTKRALARAIGISERMIQYYITGAKSPTLDVLVAMADYFNVGLD
ref|YP_001664041.1|    -----KRIKELRKKKGITQKELASYLGISDRAVGYYESGQRTPPPDILQKIADFFNVSTD
ref|ZP_02589119.1|    -----RIKSLRKKENLTQQLAEKIGVSQRMIGYYESEERFPPHDVLTKLADCFSVSAD
ref|NP_242309.1|      ---FPERLRYLRKKHGLTMKELGKKINVAESTISGYENGNRKPMDTLVKMAEYFNSSTD
                        *:  * .      :   :.  :.:      :       *   . :  :*: *. . *

ref|ZP_02442523.1|      YLCG-----
ref|YP_001210714.1|    YLVG-----
RAAC03180               YLAGRSDDPTPPPRSPSSGWDG
ref|YP_001664041.1|    YLLG-----
ref|ZP_02589119.1|    YLLG-----
ref|NP_242309.1|      YLLG-----
                        ** *
```

FIG. 148

```
ref|ZP_02442523.1|      --IFQKRLKELRNAKGTSQIAIAAALGITDRGYRKYEAGDSEPTLSVIALADYFDVSLD
ref|YP_001662345.1|      ---FSQRLRQLREEKGLLQKDVAKILGITPSAYGYEQGKREPSMEVLKKLSDFFNVSID
ref|YP_001210714.1|      --IFAKRLSFLITKNKLSKQAVANAINVSRPAVSQFANGENLPSVEKLIADFFDVSLD
RAAC03525               MTLFSQRLTELLDRTGTTRRSFAQALGVSERMVQYYITGKKDPTVETLIAIADFFDVSLD
ref|YP_001180871.1|      MNLFRRFLKELREEKNISRDLAEILGVSTQTIANYENGHREPNFDTLLKIADYFGVTVD
ref|ZP_02543721.1|      --MIGEKIKELRKNNKITQEQLGNAIGVSKMAISYFEKGKKSPGRESLEKIADYFNVTVD
                        :  ::  *    .    :  ..  ::::      :  *.  *    .  :  ::*:*.*:  *

ref|ZP_02442523.1|      YLCGRSDDPAR-
ref|YP_001662345.1|      YLLGRTD-----
ref|YP_001210714.1|      YLVGRSDDPRR-
RAAC03525               YLVGRSENPERK
ref|YP_001180871.1|      YLIGRSE-----
ref|ZP_02543721.1|      YLLGRSEDPE--
                        **  **::
```

FIG. 149

```

ref|YP_001038857.1|-----YKGKQVRTFIIDGEPWWVAKDVC
ref|YP_146372.1|-----NQLQKVFIYSGSQVRTIIKDDEVWVFAKDVC
ref|YP_001662865.1|-----YEGNTVRTVMKDGPNPWWVLKDVC
RAAC03224MCFDPDRRGDLQLQDLHRRDTMQPVDMDGNQTMMEWMFEGHRIRVVMINDEPWWVAKDVC
ref|YP_006607.1|-----ENHPVRVVLVNGEPWVFAKDVC
ref|ZP_02781438.1|-----FDSVNVVRVVLNGDPWVFAKDVC
.. :*.. :.: *:* **:*

ref|YP_001038857.1|DILELGDTHKAMERLDEDERNTIPVTDLSLGRQLQETVYVNEAGLYNLILGSRKQE-----A
ref|YP_146372.1|EILDIADARKAVQRLDEDESLIPVTDLSLGRKQETFIVNEPGLYTLILGSRKSE-----A
ref|YP_001662865.1|SVLDIGNSRDVMARLDSDEKG-VDIIDTPGGKQEVSIINESGLYSVILVSRKPE-----A
RAAC03224EALQIANSRDAVSRLDEDEKNTVAITDGNRGPNNTTIINEAGLYQLTFTSRVDT-----A
ref|YP_006607.1|DALKLVSRLKALSSLDDEKNTVTLSDGNRGPNMSIISSEGLYTLILRCRDAVKQGTTA
ref|ZP_02781438.1|VALEISNSRDALKALDADEKKTVALSYGIRGNPNHSLISESGFYKLIARSRKAVTPGTFA
*.: :.: :. ** **: : : : :.:*.*:* : . * *

ref|YP_001038857.1|KEFKRWITHEVIPQIRKTGIY-----
ref|YP_146372.1|KQFKRWVTHEVIPTIRKTGGY-----
ref|YP_001662865.1|KKFKRWVTHEVLPISIRRHGLYATDELLANPDFLIQ--ALQELKAERAKNAELTTTISIQE
RAAC03224KRFRRWLAHEVLPISIRKTGEYKTPGRRQECDIAAKQVAVMEMKARTEQAKLLVDAVHRLE
ref|YP_006607.1|WRFRKWVTNEVLPAIRKSGEY-----
ref|ZP_02781438.1|HRFSNWVFRNVIPGIRKTGAYGIP-WGALQDFSRRKEQYQLSASQKGRELQACKRKKREL
.* .*: .:.* **: * *

ref|YP_001038857.1|-----
ref|YP_146372.1|-----
ref|YP_001662865.1|QQIAEM-----
RAAC03224HRLSDLLIERILIASTNLMAGYDAIGIPTDGPSFTEPGRTRFPDLKPYWKPF
ref|YP_006607.1|-----
ref|ZP_02781438.1|EEEEKRL-----

```

FIG. 150

```

ref|ZP_02378091.1| -----DYLPAYATTGSAGLDLRACLDAPLTLKPGETALVPT
ref|ZP_02454559.1| -----DYLPKYATTGSAGLDLRACLDAPVTLKPGDTALVPT
RAAC02915 MCVVVMGGWEMEVKQIVSPLLTAADDLPQYATQGSAGMDLRACLEAPRVVRPGEIVPVPT
ref|YP_001211829.1| -----AISLPYATDGSAGLDLPACLDEPLAVPPGARVKVPT
ref|ZP_02091210.1| -----VPAYATPGAAAADLCAVLDAPLTVAPMQRVLVPT
ref|YP_518738.1| -----PSITIVESGAPSAPKLPQYATPGAAGVDLQASLDQELTIEPGQIVKIPT
                        : * * * * * : * * * * : . : * . : * *

ref|ZP_02378091.1| GLAIHLADPGYAALILPRSGLGHKHGIVLGNLVGLIDSDYQGELMISTWNRGQTEFVLNP
ref|ZP_02454559.1| GLAIHLADPGYTALILPRSGLGHKHGIVLGNLVGLIDSDYQGELMISTWNRGQTEFALNP
RAAC02915 GLAIQLPRRDAVALVYARSGLAAKHGIALANGVGVIDSDYTGEIVVPLHHFGSKEFVLQP
ref|YP_001211829.1| GIAIEIPHRNIAGLVFPRSGLASKHGISLANAVGVIDSDYKGEIVIAVFNQSDQEYLIKPF
ref|ZP_02091210.1| GLAIELPGAHSVALVYARSGLSIKHGLCMANGVGVDSDYRGELKVPVNLGAEAYTIQPF
ref|YP_518738.1| GLAIELPHAGVGAFVFARSGLASKYGLALANGVGVIDSDYRGEILVAVINQSEPFVVKD
*: ** : . . : . : * * * * . * : * : . : * * : * * * * * : . . : . : : :

ref|ZP_02378091.1| FERLAQLVIVPVVQAQFNIVDDFAASDRGAGGFGSTGR-
ref|ZP_02454559.1| FERLAQLVIVPVVQARFNLVDDFAQSERGAGGFGSTGR-
RAAC02915 KERIAQLVIAPIYVARLVAVDHLGPTARGAGGFGSTGRV
ref|YP_001211829.1| GERVAQLVFPVFTATLDVVENLNCSSRGEGGFGSTGRI
ref|ZP_02091210.1| GERVAQLCIAPVYTAAAFVPAEELGDTQRGVGGFGSTGK-
ref|YP_518738.1| GDRIAQMVFPLPVFIGEFYLADQLDETGRGCGGFGSTG--
*: ** : : * : . : . : : : : : : : * * * * *

```

```

-----GAPSMDSRQAALDMALKQIEKQFGKGSIMKLGEQTETRIST
-----MSDRQAALDMALKQIEKQFGKGSIMKLGEKTDTRIST
-----MNRDQAALDMALKQIEKQFGKGSIMKLGEQTEKRIST
-----MSDRKAALDMALRQIEKQFGKGSIMKLGEQAEQRVST
-----MSDRRAALEMALRQIEKQFGKGSIMKLGESTHMQVET
MFGSRRRRGSRAYALEEIEGWCVMGDKKAALQALRKIEKEFGKGSIMRLGEAAAMTVET
          * *:****: **:****:*****:*** : :*

```

```
VPSGSLALDAALGVGGYPRGRIEEEVGPESSGKTTVALHAIAEVQQQGGQAAFIDAEHAL
VPSGSLALDTALGIGGYPRGRIEEEVGPESSGKTTVALHAIAEVQEKGGQAAFIDAEHAL
IPSGSLALDIALGVGGYPRGRVIEVGPESSGKTTVALHAIAEVQQQGGQAAFIDAEHAL
ISSGALALDIALGVGGYPRGRVIEVGPESSGKTTVALHAIAEVQRNGGQAAFIDAEHAL
VPSGSIALDIALGIGGLPRGRIECYGPESSGKTTVALHAIAEVQRIGGQAAFIDAEHAL
VPTGSIALDIALGVGLPRGRIVEIYGPESSGKTTVALHVVAEVQKLGQAAFIDAEHAL
.:*.:*****.* *****.* *****.*.:*****.*.:*****.*
```

[illegible]

GD SHVGLQARLMSQALRKLSGAINKSKTIAIFINQIREKVGVMFGNPETTPGGRALKFYS
GD SHVGLQARLMSQALRKLSGAINKSKTIAIFINQIREKVGVMFGNPETTPGGRALKFYS
GD SHVGLQARLMSQALRKLSGAINKSKTIAIFINQIREKVGVMFGNPETTPGGRALKFYS
GD SHVGLQARLMSQALRKLSGAINKSKTIAIFINQIREKVGVMFGNPETTPGGRALKFYS
GD SHVGLQARLMSQALRKLGGAISKSKTIAIFINQLREKVGVMFGNPETTPGGRALKFYS
GD SHVGLQARLMSQALRKLAGAISKSKTIAIFINQIREKVGVMFGNPETTPGGRALKFYA

[illegible]

ELDIVQKSGAWYSYQEERLGQGRENAKOFLKENKDILLMIQEIQIREHY-----
ELDIVQKSGSWSYEEERLGQGRENAKOFLKENKDIMLMIEQEQIREYY-----
ELDIVQKSGSWSYSNDERLGQGRENAKOFLKENTDIRQEIAGQVREH-----
ELDIVQKSGAWYSFNDERLGQGRENAKOFLKENPEAAEIESRIREHYGLN-----
EMDIVQKSGAWFSYNGDR LGQGRENAKOFLKDHPEVA AAVIEKQIRE-----
EIDVVQKSGAWYSFG EERLGQGRENAKOYLKEHP IEAQIEA AVREYFHVNPAKL VAAA
.:*****.*:* : *****.*.*.* : * : *

AEDADDDADDAFDEFDAF

FIG. 152

```

ref|NP_242725.1|-----MKAAFISDIHGNAQALEAVLTDIESKQVDQIIVLGDICFR
ref|ZP_01696769.1|-----MKLAFISDIHGNAHALEAVLEDIEQKRADKIFVLGDLICFR
ref|YP_175539.1|-----MRFAFLSDIHGNATALEAVLNDLQTKQIDQTYIILGDLICFR
RAAC02234MIARMETRTMDVTTMGRGVSMRLAFFSDVHGNEALALDAVIADLRQVGC DG VYVLGDLAFLR
ref|ZP_02170975.1|-----MRLALLSDIHGNEAALRAVLEDLSRKNASHVAVLGDISYR
emb|CAJ73252.1|-----MKILIISDIHGNEAALKAVLEET----ADMIFCLGDIVNY
* * : : * : * * * * * * : : . * * :

ref|NP_242725.1|GIEPKRSLELVRS LQVPVVKGNAD EWVVRGIRKGEV PDSVLEVMRIEREWTY GKLDES DI
ref|ZP_01696769.1|GPEPQRAYEMVMALNTEVIKGNAD EWVYRGVGMNEVPEKAYEMMNRRERDWT VSRMEQNAV
ref|YP_175539.1|GPEPKRVLELVQASGARVIKGNAD EWLVRGFKDGEVPVERLEM FNKERDWT LGRLSQEDL
RAAC02234GYAPKACVEKVAEVADKVIRGNAD EWVVRGVRPGEVPDERRAGMDEEAA FARGLLAREEL
ref|ZP_02170975.1|GPKPKEC L D L I R E L H G K V I K G N A D E W L I R G I R E G E L P Q Q A F A I M Q R E Q A W S Y G K M T D E G L
emb|CAJ73252.1|GPYPKECIEKIRKLT DKIVRGNHDNAI GKNMECG-CSEKYKALSDQ GKIFTKTILDAGEK
* * : : : : * * : : . . . : :

ref|NP_242725.1|DFLKQLPTVHTFSLSDTWNVLCFHATPTS LFDIVTPTAADHV VKEKLMAQNQANLYLYGH
ref|ZP_01696769.1|ESLHQLPEEVKYEYGG- IKIHGFHATPYSLFENVPPDSENSKLKEKLMQED-ADIYLYAH
ref|YP_175539.1|DYLKNLPTELVIDDGQDLVIHAFHASPSGLFTAIDGD-ETEKIENELMVRDEADIYIYGH
RAAC02234EYLANLP LLLQ-EESPFGRWLA FHATPLDPFPVVAADAPDDDDIESRI VAGQDARLYLYGH
ref|ZP_02170975.1|HYLNQLPTELEIPLTNRIQLYATHAFPD DLFKVIPEHAENSAFDAFFEHNPRAMYYAYGH
emb|CAJ73252.1|EFLANLP LTLNTEAGGVTFLLSHGSPGDIYKYLRPEVSDSEMEDELKGVR-ANIVLLGH
. * : * * : . . : : . . : * . *

ref|NP_242725.1|IHLPYVRFDIGKIVANLGSVGLPFDGVCQASYLLVDGAE EQFSVTIQRVSYDIHSVCEQI
ref|ZP_01696769.1|IHLPFIRTFDGKTFVNLGSVGLPFDGI AKASYAMVEIGDHDYQVSNVRVSYDVQKTIGLF
ref|YP_175539.1|THHPYVRSLHGKNVINTG SVGMPPFDGHPLASYI ILDIEDGSHSVQLNRVPYNREHVVEIY
RAAC02234IHVPYVRDIRGRTVVNLGSVGMPPFDGVPQASYVILHVEDVDVFRVEHRRVPYDVEAACRRY
ref|ZP_02170975.1|IHIPHRNITGNTLLNTG SIGLPFDGNPDASYVMLERTNDSISTSFHRVAYDIEKAVFDL
emb|CAJ73252.1|THLPVVRKVGEITVVNPGSVGQPRDGIPLASYAIWKDG----ALEIKRVPYDIDATARGL
* * : * . . * * : * * * * * * : . * * : . .

ref|NP_242725.1|EASDYPNPTF--LQNV LKHGKP-
ref|ZP_01696769.1|EASDYPNKDK--LIDILTRAK--
ref|YP_175539.1|KQSGYPNIEA--MSAVI-----
RAAC02234DEIGYPAAEM--MKRVLRTARPV
ref|ZP_02170975.1|KDTDYPEDAAPLLES IYRTGK--
emb|CAJ73252.1|QHTTIPADHVARLEEILRKGR--
. * : :

```

FIG. 153

```

emb|CAK51299.1|      -----RVRFLDTRPLRSSQSFRDLWIGTSVSQVGGQIANMAVLAQVWDLTGSPVGT
emb|CAI78402.1|      -----RVRFLDTRPLRSSQSFRDLWIGTSVSQVGGQIANMAVLAQVWDLTGSPVGT
ref|YP_001362100.1|  -----VDTRPLRSSPAFRRLWIGTTAAAFGGQVAVVAVLYQVWNLTGSTVWT
RAAC01662            MPMTLQAVFRSAMLDTPLRTSPAFLRLWIGSGISTLGGTMTFFAVMLQVYRATHSTLDV
ref|YP_832554.1|      -----LADITPLRESPDFRRLWLGSVSNLGSQTLTVAVSLEVYRLTQDSLYV
ref|YP_712120.1|      -----LADLSPLREHPAYRRLWVGETISALGSQITATAVLLQVFAVTRSSFQV
                        *   ***       :   ***:   :   .*.   :   **   :*:   *   . . .

emb|CAK51299.1|      GAIGLATGLPMVLFGLLGGTLADTFDRAVVRATTAGQLLAAAGLCAQALADNRN--VLL
emb|CAI78402.1|      GAIGLATGLPMVLFGLLGGTLADTFDRAVVRATTAGQLLAAAGLCAQALADNRN--VLL
ref|YP_001362100.1|  GAIGIATAVPTIVGGIVGGTLADTLDRRLVLVTTTVSVLAALSALQAVAGAAT--VLL
RAAC01662            GLTGLASVLPGLFVLLGGSVGDVRDRKLALLATSGQMLVSALLAVQALIAWSA--LWL
ref|YP_832554.1|      GLLSIFALVPLVLGGLLGGSIADAHDRRKVALLATMMLWLTTAGLAAQAWLGLGN--VWL
ref|YP_712120.1|      GLVSLVALVPLVAGGLFGGAIVDAVDRRLAMITSGGLAVVSAFFLLTLGAVDDVAWP
                        *   .:   :   :*   :   .***:   *   ***   .:   :   :   :   :   :

emb|CAK51299.1|      LLALVAMGTSSGALGAPARRTFPVRLLPGDQVAAGLALTNVSFQAAMLAGPAMAGLIIAH
emb|CAI78402.1|      LLALVAMGTSSGALGAPARRTFPVRLLPGDQVAAGLALTNVSFQAAMLAGPAMAGLIIAH
ref|YP_001362100.1|  VLVLVAAQTSAAAVGAAARRTFISRLLPDDLVRAGVALNHIGFQIAMLTGPAVAGVVLAA
RAAC01662            IYALLCAQSVLGAINAPARRTFVPRLLPKEQVRAGQTLNTLAIRFGEVAGPSLAGVIAF
ref|YP_832554.1|      LYVLVAVQSGAQAINQPARSAIIPALVRKELLPAANALSMITFGLGMTAGPLLAGVLVAV
ref|YP_712120.1|      LMCLVAVQSALLAADQFARRAMTPNLIPLEHLPAASALAQIGGTAAGVFGLIAGVSVAA
                        :   *.   :   :   *   .***:   :   :   :   :   :   :   :   :   :   :

emb|CAK51299.1|      WDLAAYATQAVAMVLSMLTVLRLPAMRPEGADAAGGRRRPERG-----GWRIVL
emb|CAI78402.1|      WDLAAYATQAVAMVLSMLTVLRLPAMRPEGADAAGGRRRPERG-----GWRIVL
ref|YP_001362100.1|  GGLRAAYLTDAVGLTVSLYGVLRPLPAIRPQ-SLADTGTARAARGGLPAVRATWEGWRYLL
RAAC01662            AGLGWCYACDAATFLAALYAIYRLPPMLPEANAKQRSMLDGVME-----GLRFLG
ref|YP_832554.1|      VGFGWYITLDVASFAFAFWALLRLPPMPGKTTTHRAGLRSVVEG-----FRFLG
ref|YP_712120.1|      GGFSIAYAIDLVTFAAPLYGLARLPAMPPRGGGRGAGVASAEG-----LRFLR
                        .:   *   :   .:   :   :   :   :   :   :   :   :   :   :   :

emb|CAK51299.1|      RRPTLWGSMATDLSATLLAMPVALFPLVNEIRFGGNPQTLGLFLSAVAVGGITAGLLSGT
emb|CAI78402.1|      RRPTLWGSMATDLSATLLAMPVALFPLVNEIRFGGNPQTLGLFLSAVAVGGITAGLLSGT
ref|YP_001362100.1|  RRPVLRGCLATDLAATVLAMPIALFPAINAQRFDEEATLGLFLSAIAVGGLLAGFTSGT
RAAC01662            RSRALLGAMLADLSVTVLGVPTALLPALVSERFGRPETLGLMMGATGIGGLVILALSGP
ref|YP_832554.1|      TRPNLRMTFVIDLVAMIFAQPRALMPAIGAVMIGGGEATVGVLLASTAVGAFLAGLFSGP
ref|YP_712120.1|      GQPVLMTFVVDIIAMVFGMPRALFPALAEGRFGGGSATAGAMYSAVAVGSLLAAGLSGP
                        *   :   *   :   .:   :   *   ***:   :   :   :   :   :   :   :   :

emb|CAK51299.1|      VTRWHRGGLVQMSAAGVWGLALACFGLAGPLW-----LALGCL-AVAGAADTV
emb|CAI78402.1|      VTRWHRGGLVQMSAAGVWGLALACFGLAGPLW-----LALGCL-AVAGAADTV
ref|YP_001362100.1|  LTRARRTGRVMLIAAATWGTSLTTFALVDSLPL-----ATLACL-AVAGAADTV
RAAC01662            VRHIRSEGKAILVACAAWAIAAGGLGLAPFVW-----LCVALL-GLMGAADSI
ref|YP_832554.1|      LGGIRRQGSAAVVSVMGWGASIAFGGLVVVLAGRSGADGVTTWLLPALCCALAGIADSV
ref|YP_712120.1|      LGGVRRQGLAVVVSIVLWGGAIAGFGLAHDIA-----LGLALL-ALAGAADMV
                        :   :   *   .:   :   :   *   :   :   .*.   :   :   :   :   :   :   :

emb|CAK51299.1|      SVVTRSALVQLETPDAYRGRVSSVEHVIGVAGPELGNFRGGLLASATSASFSLVFGGLSA
emb|CAI78402.1|      SVVTRSALVQLETPDAYRGRVSSVEHVIGVAGPELGNFRGGLLASATSASFSLVFGGLSA
ref|YP_001362100.1|  AVISRGITIVQLSTPDAFLGRVSAAGETVGVAGPGLGNARAGAVAGVTSTAVSSLTGGLAC
RAAC01662            LVVMRSTLIQQTDPDLGRRISSVDYLVGTGGPQLGNLRAGLVGSLPPGASIAVGGGLST
ref|YP_832554.1|      SAVFRTTILQAATPDHLRGRQLQGVFVVVAGGPRIGDMLAGGGTKLLSEGWVLLLGGLL-
ref|YP_712120.1|      SAILRNAILNVATPDEMGRQLQGVFLVVVTGGPRLGDLEAGSVAAVTSPAFSVVSGLLAG
                        .:   *   :   :   :   ***   ***:   . . .   :   .***:   :   :   :   :   :

emb|CAK51299.1|      ILAIAAV-----
emb|CAI78402.1|      ILAIAAV-----
ref|YP_001362100.1|  VLAVAVIAATTPALRR-----
RAAC01662            LAAILCVGWFPPELRRQAEEGVGASD
ref|YP_832554.1|      --CIAVAWTV-----
ref|YP_712120.1|      IGCLLLATAAVPALIRYDARAAV----
                        .:

```

FIG. 154

```

ref|ZP_02620185.1| -----MALYAISDLHLSLN-----CDKPMDFVGDHWMNHDNRIKENW
ref|YP_877450.1| -----MALYAISDLHLSLN-----CDKPMDFVGDHWMNHDNRIKENW
ref|YP_001512273.1| -----MAIFAIGDLHLSGY-----SNKPMDFGEHWTEHDKKIMESW
ref|ZP_01188667.1| -----IYAIGDLHLSFENKVIPGDWEKVSQYKPMFLSGDKWVEHYRKIYKNW
RAAC02171 MISWNGQGTTVAIYAIADLHLDTS-----QSKPMDFGHEWRDHAEKIAHHW
ref|YP_001275096.1| -----VWSISDLHLSFA-----RPKPMDFGSRWKDHPERIAAAW
                        :::*.****.                ****.:* . * : * . : *  *

ref|ZP_02620185.1| INKITNEDTVLIAGDISWSMKMEDGMADLEWIKLPGRKII SKGNHDYWWGS--ISKLNS
ref|YP_877450.1| INKITNEDTVLIAGDISWSMKMEDGMADLEWIKLPGRKII SKGNHDYWWGS--ISKLNS
ref|YP_001512273.1| QKNVKDEDAVLIPGDISWAMTLEDAKIDLNWIADLPQKYLIRGNHDYWWGS--LTKLNS
ref|ZP_01188667.1| KAEVTCKDLVLVPGDISWAMKLEEAVIDLEFIGSMGKKIFIRGNHDYWWGS--ISKVRS
RAAC02171 QDQIHDDDIVLIPGDISWAMKLEEAAPDLVWIGRLPGRKVLIRGNHDFWWGG--IQVRVK
ref|YP_001275096.1| RARVKPDDVLLAGDTSWAMKLQDALVDLQWIAALPGRKII SRGNHDYWWSSERTNRVRR
                        .: . * **:.** **:.*.:. . ** : * :*: * : :****:***. . :.

ref|ZP_02620185.1| LY-DDMLFIQNNFFIYEDYAICGTRGWTPP-ADKYS-EHDDKIYKREQIRLKISLDAAK-
ref|YP_877450.1| LY-DDIRFIQNNFFVYEDYAICGTRGWNPP-TDKYS-EHDDKIYKREQIRLRISLDAAK-
ref|YP_001512273.1| LF-DSMHFIQNNFFTYNQYAICGTRGWNCNPNHYKFT-EHDGKIFTREVNRLLSLKAAC-
ref|ZP_01188667.1| ILPDGCFALQNDCLEFEGVSI TGTRGWICPNEDNFT-EHDEKVIYKREVNRLKLSLESIK-
RAAC02171 ALPPRMALQNDCLVLDNVCFAGTRGWTLPHHPSYNAEQDEPILKREILRLELSLKAAC-
ref|YP_001275096.1| SLPPGIDILEASAI DIGEAVVCATRGWNTPETPGFQESTDRPYERELMRDLTALAAAQH
                        :: . : . . **** * : . * ** ** : * :

ref|ZP_02620185.1| -KAGYEKIIIVMIHYPPVNDKFKQKNELTEIFNEYNVEKVIYGHHLGP-SLKTI FEGKHEGV
ref|YP_877450.1| -KAGFEKIIIVMVHYPPVNDKFEKTELTEIFNEYNVEKVIYGHHLGP-SLKNI FEGEHEGV
ref|YP_001512273.1| -EKGYEDIIVMLHYPPPTNDKLEPSLFTEMLEKYKVQVYVYGHHLGETSYDAGLKGEYNGV
ref|ZP_01188667.1| -NTD-KKKIVMMHYMPVNENHEHNDFIKVMIDYNVDICIYGHHLGEDSHKIRLPEEKWGI
RAAC02171 -KEG-KPILCLMHYPPVDSNHPHSPFHELLAAYGVRACVYGHHLGP-AHRFAFNGEIDSV
ref|YP_001275096.1| LASGKRPIIVMIHFPPFAGRRP-TEFARRIAAKAAACVYGHHLRPEDWAVATQGLVDGV
                        . . : :*: * . . : . : : ***** . :

ref|ZP_02620185.1| EYIMTSCDYIDFDPIKI-----
ref|YP_877450.1| EYIMTSCDYIDFDPIKI-----
ref|YP_001512273.1| YYNLVSCDYAGF-----
ref|ZP_01188667.1| RFYLVSSDFLNFKPLRI-----
RAAC02171 RYQLVSCDYLQFIPWKIPEEWLQP
ref|YP_001275096.1| YYQLTSCDYLGFGP-----
                        : :*.** : *

```


FIG. 156

```

ref|ZP_02850022.1|      MNKLTTD--AAFREAVQGDGITVVVFKTTWCADCHYIDPFMPDVEQQYNGKFSFHEIDRD
RAAC01724              MEEIRTS--ERYREAVQ-SGRVVVEFYATWCPDCRRIEPYLGWEEKYREQFTMVRVNRD
ref|NP_244119.1|        MEQIKTL--EQFQQVKN-QENVVFLFSADWCPDCRVIEPFLPELEQTY-DEYQFYVNRD
ref|YP_001422284.1|     MKKIEST--EELQKAVQ-DDWAVFMFSADWCPDCRFVEPFLPELEADF-PEFTYYYVDRD
ref|YP_815236.1|        MEEIKELTPEKLKEITA-NGKVVLFLATWCPDCRFLDPFLPQIEKDN-SDAKFYKIDRD
ref|NP_965472.1|        MEEIKELTPEKLKEITK-NGKVVLFLSATWCPDCRFLDPFLPQIEKDN-PDAKFYKIDRD
                        *:::      ::      .  .*. * : **.*: :::: : *      .      ::**

ref|ZP_02850022.1|      ELPDLCSELNILGIPSFIAFRNGKELIRYVNKLRTREEIEQFLDRA--
RAAC01724              EVPDLAEELQILGIPTFLVYDQGREVKRLFSRDAKSKEQVEQFLDQAYA
ref|NP_244119.1|        DFIELCQELDIFGIPSFIFYNGEERSRFVSKDRKTKEEIERFLTEA--
ref|YP_001422284.1|     QFIDTCAEWEIYGIPSFVVYQDGKEIGRLVNKDRKTKEEIEQFLTDSL
ref|YP_815236.1|        GSIDVAKELNIFGIPSFVVYQDGKEIGRLVNKDRKTKEEVENFLN----
ref|NP_965472.1|        GSIDVAKELNIFGIPSFVVYQDGKEIGRLVNKDRKTKEEVENFLN----
                        : . * : * **.*: : *.* * ..: *:::*.**

```

FIG. 157

```

ref|YP_001212789.1|      -MRAAIYLRVSTEDQAKHGYSLPDQRTACREKAAALGAG--EVMEFADEGISGELLDRPG
ref|YP_001211675.1|      -MRAAVYVRVSTEDQARHGYSLQEQKEACRCRAVDLGAK--TVLEFADEGVSGATLDRPG
ref|ZP_01666637.1|      ----AIYARVSTDEQARTGYSLQDQVNQCRRKLLSLGLS--NIKEYIDDGYSGEFLDRPA
ref|ZP_02326197.1|      -MHAAIYVRVSTGLQAVEGTSLETQLDYCLRKAFELGISQNDIHIYREEGASGEDLDRPA
RAAC01817                MMWTAIYTRVSTEHQAGHGLDVQREACVQYALSGLGVAPHDIRLYEEAGGSGEDMDRPE
ref|ZP_01126596.1|      MKIAAIYARVSSDQQKGAN-TIASQTAALIEFAREQGFTVPDEWIIIEDEGFSASLLRPG
                        *: * ***: * . : * * : * * : **

ref|YP_001212789.1|      LSALREAVKAGQLDLVVCDFDPDLRLARKLAHQLIITDEIEKAGVRLEFVNF-EWQNTPDGR
ref|YP_001211675.1|      LQGLRELIRSGQIDLVVVRDPDLRSRKLSHQLILTEEIEKAGVRLEFLDF-DWKDTPDGR
ref|ZP_01666637.1|      LSRLRDDLRAGLIKTVMVYDPDLRSRNLTNQLIIADEIEKYGAKLEFTG-SYDASPEGR
ref|ZP_02326197.1|      LNRLRQDVASGTFSLILTHPDRLTRDLTDKLFICRELESRNIRLVFVDT-EYKNTPEGQ
RAAC01817                LLRLDDVRRGLVDRVVVKHPDLRSRVADKAIVVRELSACGVKLHFVDVFNWDESDEAV
ref|ZP_01126596.1|      LERLRDLAAEGQIQAVLIHSPDLRSKYAYQVLLTEEFARHGVEAIFIKA-PHSATPEDQ
                        * * : * .. : : ****.* : : : * : . . * : . : :

ref|YP_001212789.1|      LFYALRGAI AEYEKEKIRERTVRGKLQKARLGGLPVQADSYGYIID-----NGTVKP
ref|YP_001211675.1|      LFYAIRGAI AEFKEKIRERMARGKTQKAKQGGMPIGFYNYGYVYEPE-----TGKVRL
ref|ZP_01666637.1|      LFFSIRGAI AEFKEKIRERSRLRGKRAKVLSGKPLFGRDPYGYICDRD-----TGQYVI
ref|ZP_02326197.1|      LFFNLMSVIAQYELSLIKKRTVRGRLKAVEKENKIMPRTAPYGYDII-----GSKLVI
RAAC01817                LLFHVISSIAEYELRQIRRTLAGKLKAVR-GGKVMPSGVDPYGYRYE-----DGRYVV
ref|ZP_01126596.1|      LMLQFQGMIAEYERAQILERSRRGKRHRAKSGEISVLGG-APYGYRYIRKMPETPARYEI
                        *: . . **:* * . * *: . . . * .

ref|YP_001212789.1|      HPVESEVVKMIFRWFTTEDIGVNGVAARLSEQGIPTRKGRP-AWQRCVVKTILTNPMYVG
ref|YP_001211675.1|      HETEAQVVEEIFKWFVQEDIGINGVAKRLNEAEVPSRK GK--RWHKQVVRQVLVNPVYKG
ref|ZP_01666637.1|      NEEEAQIVRLIFKLYTENRYSVAKLHAQLKAMGVVNRSGK--PFSLSVLDHILANEMYAG
ref|ZP_02326197.1|      NEEEARFVRYIYEWYVHQWRTIRQIGEKLVELGAIPKRKESRSWSASSIQRIILTSEIYIG
RAAC01817                VPEEAEIVRLIYQWYIGDMSLRRAIERLDAMGVPTKTRQSARWHHSTVARILDNPLYRG
ref|ZP_01126596.1|      DAAEAAVVRLVFEKYTVDDLISGAIARLLREMGPPTRRRVTR-RWERSVWGMRLRNPAYKG
                        *: *. :. : : : : * : : : : * . * *

ref|YP_001212789.1|      TFYNNRRDCRGTCYNK-----HLPPEKRVKVKEKPGEEWIAVPVPAIIDRETWEKAQE
ref|YP_001211675.1|      TWQYKD-----ICIPVPAIIDEAVWLKAQE
ref|ZP_01666637.1|      TKWYFQKYQKTVG-----QKKRKVLKRNVNEWVSIDVPAIIDKETFQKAAE
ref|ZP_02326197.1|      RYYNNRRKTGKVKGQK-----TPSGSNRKLLEWRKEEDWIRVEVPAIIDTGIYEQAMQ
RAAC01817                TWYNNRRRTRKRQGVRG-----ARRGRGRQVVAVRDPGEWIAVSVPVVDPLARSVAL
ref|ZP_01126596.1|      TACFNKTQVGPRQKVTKPFRLSGRSVHGEKTSASHERPREEWIEVPVPAI VSEETFALAAE
                        : : : **.:. :

ref|YP_001212789.1|      KIKAAARLWSGWSREEYLLSGLISCADCGNTMHGAVKTKSGG-----
ref|YP_001211675.1|      KIRGARRLWAGQRKHDYLLSGIVTCGECGQTMGTG-VYSKWWN-----
ref|ZP_01666637.1|      VRRQNKVTAKRNTKFEYLLSGIIKCPKCGYAMRG---TRFPK-----
ref|ZP_02326197.1|      Q-RQRNRKKS GHVKESYLLRGLIRCCEGCRSWQATSYS-----
RAAC01817                R-KQGS-PRGGSTSAHTFLSGKLVCACHCQRAWRHEACRTAGGIVRKFRRPPAERGADPCA
ref|ZP_01126596.1|      RLADNKRFAFPRRTIEPSLVQGLVSCRKCGYALYRASTRTSAR-----
                        : : * : * . * :

ref|YP_001212789.1|      -----
ref|YP_001211675.1|      -----
ref|ZP_01666637.1|      -----
ref|ZP_02326197.1|      -----
RAAC01817                -----GRADP-----
ref|ZP_01126596.1|      YRCARVPADEIERAVVRELVRMRALGIGRELAAGGRASPGDPLMEDVARLRREMRAAVE
                        -----

ref|YP_001212789.1|      -----SRERGYTCVKTAAGAS-----NQGC
ref|YP_001211675.1|      -----KKDRRYTCFKGYQGAR-----HRGC
ref|ZP_01666637.1|      -----RNNKDYAYYVCSAYVN-----AYEC
ref|ZP_02326197.1|      -----MSGAKKYLCCYRCPNKTPTATFG-----TTRLI
RAAC01817                RRDRARELYLRGHLDRTAALRAMSEEARRLREGRAEACRLCHARQRLNAWEDLWTALET
ref|ZP_01126596.1|      -----KIHYRCLGSDGWRHLG-----GSV

ref|YP_001212789.1|      RPIKRVPAEIVEKAVW-----
ref|YP_001211675.1|      LPSKYVLAGYVESAVWEQVKDWLQDPGALASEAYASSPRVED-----
ref|ZP_01666637.1|      DNRRCVPSQELDEAVWQEI VNMFFKKS-----
ref|ZP_02326197.1|      CTAPSLKAEWLDQNVWKSVL-----
RAAC01817                HDGVELRRLWLEAAVSRVEMDASGGEVVLTFLFARIGEQTGARVDDTLAT
ref|ZP_01126596.1|      CDSRPIRQDLLDHIVWQEVMRLIADPGL-----
                        : : : *

```

```

ref|YP_001514193.1|-----
ref|YP_001317996.1|-----
ref|YP_001090064.1|-----
ref|ZP_01995293.1|-----
ref|YP_517020.1|-----
RAAC01840      MTHGGGVLRFARLSRVSSSSRGRSHDQALSDRVGGGDQIVAGGRGDRRRSRLRGGHPPH

ref|YP_001514193.1|-----
ref|YP_001317996.1|-----
ref|YP_001090064.1|-----
ref|ZP_01995293.1|-----
ref|YP_517020.1|-----
RAAC01840      TLRRVSSSSRSGRRGRPARAGSGGRSPGELPRLHGSLDCRSAGRHHELCPASAAAQRLRHR

ref|YP_001514193.1|-----
ref|YP_001317996.1|-----
ref|YP_001090064.1|-----
ref|ZP_01995293.1|-----
ref|YP_517020.1|-----
RAAC01840      VRERRRGELRRRPGSLSGRVVLRAPRRRVCRGRSRRPRVARRAVRGLAGQAHARVRGVR

ref|YP_001514193.1|-----
ref|YP_001317996.1|-----
ref|YP_001090064.1|-----
ref|ZP_01995293.1|-----
ref|YP_517020.1|-----
RAAC01840      ICILEQCD-BEMBCAPUTATIVENICTINATEPHSPHTRIBSYLT-----
IAARGHGHGAPRPAPRSKARDAKSYGRDRRGEEPANARRGRAAPGLRGRRRAHRLLGIRS

ref|YP_001514193.1|-----
ref|YP_001317996.1|-----
ref|YP_001090064.1|-----
ref|ZP_01995293.1|-----
ref|YP_517020.1|-----
RAAC01840      QRLGHRRGGAFGPGSRRRHPVDARRTLFAPRAGRVRGIRWRDAPGARGASCSRMSKAISR

ref|YP_001514193.1|-----
ref|YP_001317996.1|-----
ref|YP_001090064.1|-----
ref|ZP_01995293.1|-----
ref|YP_517020.1|-----
RAAC01840      KNLTLTLDLYQLTMMNGYLKNGADENVVIFD
RNLTLTLDLYQLTMMNGYFENNSHEDTVIFD
EE-IDENTITIESPSITIVES---GAPSRNLTLLTDLYQLTMLNGYFEKNIHEDIVFD
RNLTLTLDLYQLTMMQGYEKGQONEN-VIFD
VALLTDLYQLTMMQGYQNGYENKEAVFD
AKDGDGEMDIKERARARVEEIGALPVYRDRHLTLLTDLYQLTMMYGHFRAGRHETRNVFD
*****:***: *: . . .: .: **

ref|YP_001514193.1|-----
ref|YP_001317996.1|-----
ref|YP_001090064.1|-----
ref|ZP_01995293.1|-----
ref|YP_517020.1|-----
RAAC01840      LFFRTNPCNNSYTMIAGIEQVIDYIENLKFDEESLVYLKGLNLFDDDEFIDYLRNFKFTGT
LFFRKNPCNNSFTIIAGIEQVIDYIENLGFTEEDIQYLRSLNLFGEAFLDMLKQFTFTGT
MFFRKNAACDGGYTIVCGIDQVVEYIDNLHFSDEDELEYLKNLNLFSDFLKFKEFKFTGD
VFFRQNPCNNGYSVCAGLDQVIDYIKNLHFTYDDVDYLRGLGIFKEDFLHYLSGFHFSGD
LYFRKIPSGGGYVIAAGLEQVVEYIENLRFSSDMAYLRGLNIFDEGFLNLLKDFRFGHD
LFYRKNPCNGYVIAAGLEQVVWYIYNLAFSEDDLAYLRSLGMFSEDFLSYLRHFRFRGD
:::* .....: .:****: ** ** * ..: **:.*: * : * : *

ref|YP_001514193.1|-----
ref|YP_001317996.1|-----
ref|YP_001090064.1|-----
ref|ZP_01995293.1|-----
ref|YP_517020.1|-----
RAAC01840      IYGVDEGTVMFPYEPILRVKAPVIAQALIETLLNIVNFQSLIATKAARICSAAD-----
IYGVQEGSIMFPHEPILRVKASVLQAALIETALLNMINFQSLIATKASRIVEAAK-----
IYAVEEGTIMFPNEPLITVKAPLYQAALIETALLTIVNFQSLIATKASRVCFAAQ-----
IYAIPEGTVVFPKPEPLLKVAPIMEAQVETAILNIINHQSIIATKTSTRIVFAAN-----
IDAVPEGTVVFPYEPILRVKGRILEAQALIETALLNIINFETLIATKASRVVAAAG-----
VYAVPEGTVVFPNEPILRVEGPVIAEVQLIESAVLAFINHQSLIATKARRIVEAARTNVRH
.: : ****: ** * : : ****: ** * : : ****: ** * : : ****: **

```

[illegible]

```
ref|YP_001514193.1|
ref|YP_001317996.1|
ref|YP_001090064.1|
ref|ZP_01995293.1|
ref|YP_517020.1|
RAAC01840
```

```

SELEAFRAYARSYPNSTLLVDVTYNVLHSGVPNAITVFNELKEQGYPEK--GIRIDSGDI
SELEAFRAYAKAYDPKCLLLVDTYDTLKSGIPNALTVFNELREKGYEPK--GIRIDSGDI
NELEAFQAYADVYPDKCLLLVDTYDVLNSGVPNAIKVFNISEKGYKPM--GIRLDSGDL
DEYTAFTKYAEMYPDNCTLLVDTYDTLKSGVPNAIRVFQEFKDAGKPLIKYGIRLDSGDL
DELEAFRAYARTFPDQCLLLVDVTYNVLHSGVPNAIKVGLELEAEGHRLF--GIRIDSGDI
SEYEAFAQYAAAFDNTVLLVDVTYDLRSGVPNAIRVGLEMKAAGQSLK--AIRIDSGDI
.*  **::*  :.*  *****.*.***::*:  ::  *  .***:****.

```

```
ref|YP_001514193.1|
ref|YP_001317996.1|
ref|YP_001090064.1|
ref|ZP_01995293.1|
ref|YP_517020.1|
RAAC01840
```

[illegible]

```
ref|YP_001514193.1|
ref|YP_001317996.1|
ref|YP_001090064.1|
ref|ZP_01995293.1|
ref|YP_517020.1|
RAAC01840
```

GGVYKLSAIE-VDGQIIPKIKISENPEKITNPGYKKVVRIYDKENKKAQADLIMLDSEEI
GGVYKLSAVE-KSGTLIPKMKLSDNPGKITNPGYKKVIRIYEAENNRAQADLILLEEEI
GGVYKLAASC-EKGVLEPKIKISENPEKINNPGYKKVIRIYN-EDNKAEADLIMLHDEVI
GGVYKLAAIQNEEGEFVPKIKISENTEKITNPGNKTIYRIYEKASGKIKADLICFADEVI
GGVYKLSAEG-QKGVFEPRLKVSENISKITNPGIKKIVRFYD-RRGKAMADLIALEDEHF
CGVYKLAQQ-FGDRMEPRIKISENPNKITNPGKKKVLRLYV--NGSATADLIALDEEY
* * * * *

```
ref|YP_001514193.1|
ref|YP_001317996.1|
ref|YP_001090064.1|
ref|ZP_01995293.1|
ref|YP_517020.1|
RAAC01840
```

NTNEPLTIFHPIYTWKKKTFRN--YETKDLLVPIYEQGLVYERKTIKEIRKHAQSELD
DTTKPLTIFHPLYTWKKKTFSN--YKVRMLMPIYEDGKLVPYPRKTVNEIRAYVRQELST
DESKPLEIFHPTYYWTKVFTN--YKVKELLKPLYIKGRCKYNKKAVLEIKNHVQYELST
DPKQDLLLLFDPMDTWKKTLAGGTYNVREILLPIFKNGECLYKSPTLKEIAAYCREEKDT
E--EPLTIFDPIETWKRKTLTD--FKTRELLRPVFRGGRRVYELPYLKD IQTYAQHECET
DPSEPLFHFPIHTYKRKVQN--YEMEELLRPVFGGELVYELPTQVEIQARVETQLSA
:
: * * * * * : * * * * * : * * * * * : * * * * * : *

```
ref|YP_001514193.1|
ref|YP_001317996.1|
ref|YP_001090064.1|
ref|ZP_01995293.1|
ref|YP_517020.1|
RAAC01840
```

LWPEYRRLNRPQLYKVDLSKKLWDLKNQMMDQFR-
LWPEYKRLNRPQLYKVDLSQKLWSLKHNMIQSLKK
IWEQYKRLSKPHIYKVDLSRNLWYLKTQMI-----
LWDETKRLFYPHRVYVDLSQKLYAVKQSLLDQ---
LWDEVKRLVNPHRYIVDLSPKLFELRQSLLLE---
FSSEVRHLNPHIEHYHVDLSKPLWDLKQRLLHEWR

: : * *: **** *: : :

FIG. 159

```

ref|YP_604970.1|-----AHQGGEERWPSNTMLAYRNAALGVDMLDLDLHA
ref|NP_295807.1|-----AHQGGEALWPSNTLTAYRGAVALGVDLLEMDLHA
RAAC01875MARAVEEFVEGRMQMDGTRRPRVRIFAHRGASREAPENTLPAFQIALRQGADVIETDVHW
ref|YP_076566.1|-----PRPVVLAHQGASGHAPSNTMEAFRLALEQADILELDVHM
ref|YP_644758.1|-----FAHRGASARAPENTLEAFRLGVEAGAGGLEMDLHM
ref|YP_357266.1|-----PRPRLFGHRGNSGDFPENTLPAFSDALACGTPYLELDVRA
                          .*: * . * .*: *: . * . :*: *:

ref|YP_604970.1|TRDGALVLSHDETLDRLTDTQGRIADMTLAQVLAADAGYAFTPDGGKTFPFRGQGVVVAQ
ref|NP_295807.1|TRDGALVLSHDETLDRLTDTQGRIADLTLEQVLRADAGYTLTPAGGSGFPFRGQGVVVAQ
RAAC01875TKDRELIVICHDPRIIDAVSDGTGAIADMTYEELLRYDFGYRFTRDGGRTYPYRGCGIRMAT
ref|YP_076566.1|TRDGVVVVSHDETIDRMSDGTGLIKEMTLAELRRYDFGYRFTPDGGLTYPYRGKGVTIPT
ref|YP_644758.1|TRDGEIVVIHDHTVDRTTDGSGAVREMSLRELRLDAGYRFPDGGGRSHPYRGRGLRVPT
ref|YP_357266.1|TKDDEVVVIHDESLLRTCGIDRPVAGLTFAELQNYDAGATFTPDQGRSYPHRGCGIRVPL
*: * *: * *: . : : : * * : * .*: * *: .

ref|YP_604970.1|LAEVLAAFFNMPLTIEIKQTSPLAAPFCKALRDAGVTDHVIVASFSDKALNEFRAACPE
ref|NP_295807.1|LKDVLREFPNTPLIIELKQASPSIAAPFCAELRRAGATGRVIAASFSDAALNEFRRLCPE
RAAC01875LREALRAFFGAHFNVDLKPKCPHVG-LFLRMLEEEDSLERVTLASFHHRTLVEARARCPR
ref|YP_076566.1|LEEVLQAFPEVPVNIIEIKQADPPMEAQLWELIQRYGAEDRVIVASFHGTVAKRWRDLAGD
ref|YP_644758.1|LREVEAEFPRAAVNMEIKEFPVPGIERRVLEVIDGAGARGRTLVAADFHDGIVRRFREASGG
ref|YP_357266.1|LADVLQAFPRALFNIEIKQETPAMETATLEVIIRADMTEQVLLAAENDAVMARLRPLCQT
* :. ** . :*: * : : . :. *: . * .

ref|YP_604970.1|V--ITSMTEKELRPLVLLSKVGLA--HLAPLPGRAAQVPVRSNG---IEVVTPAFVVRAM
ref|NP_295807.1|V--MTSMTERELRPLVLLSKVGLS--ALAPTPGQVAQVPVRSNG---IEIVTPAFVVRAL
RAAC01875L--KTSASPFVEVARLAAMARAGWERGGRRMPFQAVQVPLRQYG---VRVVTPRFIEFA
ref|YP_076566.1|R--VATSAPVEHMYLVAAHYLSHLD--RLYAPAHDAFQVPVAQKAGPLTVRFDTERFLRMA
ref|YP_644758.1|E--VATSASRLLEVAFYALVRLGLW--RFARPGYAALQIPVRRRG---IELITPSLVAAA
ref|YP_357266.1|AGIPTSFYSYGELVSFFTWLQAGGQ--SPYHPPAQAALQIPETYEG---QTLVTPQSIAAA
      : : * : . . *: . * :

ref|YP_604970.1|HARGVAVHVWTINDPAEMRRLIAMGVDGIITDRPDL-----
ref|NP_295807.1|HARGVAVQVWTIDDEAEMRRLVQMGVDGVITNRPDL-----
RAAC01875HRLGLEVDVWTVDDPNVIRGLIRQGVGDGIVTNSPQILQKLLLELEQ
ref|YP_076566.1|ERVNVAVHYWTINDEDEMRRLYQLGAHGIITDYP-----
ref|YP_644758.1|HALGVRVDAWTINEPEEMRRLGLGLVDVVMTRDPE-----
ref|YP_357266.1|HTLGIEIHVWTVNQAQDMERLLRMGVDGVMSDRPAV-----
. . :. **: :. * *.. :*: *

```

FIG. 160

```

ref|ZP_01034116.1|-----MRSVELFVGAG
ref|ZP_02297879.1|-----MRAIELFAGAG
RAAC02539MRHYAFRRMLTMWTRWMAIETVLFRKQLVGFEGNRHDLKNSYSGERSVVRMRSVELFVGAG
gb|ABY83631.1|-----MKSVELFAGAG
ref|YP_208280.1|-----MKSLEIFSGAG
ref|YP_001516905.1|-----KLSCLELFTGAG
: .:*:* ***

ref|ZP_01034116.1|GLGIGVSQAGFRPAAVMDWRWACDTLRENKERGLDPIAHWPI-HEGDIRQDFDGTVDGT
ref|ZP_02297879.1|GLGMGVSRAGFTPAVVEWDRWCCDTIRENREKGIASLAGWPMPIEGDVRGVNFRGFEGK
RAAC02539GLAMGISNAGFRHVGLYEWDRYACDTIRFNKERNVGPVRDWPI-YQLDVRSDFTQYRG-
gb|ABY83631.1|GLAMGCEIAGFEHLAVVEWDKWACDTVRENKKS GFPLSDWDL-FEGDVREFDWSKIPKG
ref|YP_208280.1|GLAKGLELAGFQHASFIELNKDACNSLRNFP-----KLVIYQGDVADFDLSSQEG-
ref|YP_001516905.1|GLAKGLEKAGVQHTAFVEWNKDACITLANNYS-----QLVHNVDIRTFKFSQFGH-
**.*.***..: :.*: :*. : :*: ..

ref|ZP_01034116.1|VDLVTGGPPCQPFMSGGRHRAFLDGRDMFPQAIRAVRELRPRAFIFENVKGLTRSSFANY
ref|ZP_02297879.1|LDLVTGGPPCQPFSLGGKHQAADRRDMWPEAVRAVRETRPRAFIFENVKGLTRESFATY
RAAC02539IELLAGPPCQPFSLGGKHRGREDHRNMFPEMIRAVREIQPKVLIENVKGLLRRESFAKY
gb|ABY83631.1|IDLLAGPPCQPFSLGGKHANSDSRDMFPATAEAIHQIRPKAFIVENVKGLTRATFATY
ref|YP_208280.1|IEVIAGPPCQPFSLGGKHLAHEDRDMFPHAVRYVEYYRPAKAFIFENVKGLLRKSFADY
ref|YP_001516905.1|VDIVSGPPCQPFMSGKKHGNMDKRMFPYACKAISVCTPKAFVENVKGLLRKSFSSY
: : : :*****:***.*.*:***. : :* : : :*****.**: *

ref|ZP_01034116.1|LEYIRLQLTYPDLVAKRDEEWLAHLARLEDHHT--KGTEKGLRYRVVMRVLNSANYGVPO
ref|ZP_02297879.1|VSHIVLQMTYPDIVAKSEEGWEGHLQRLERHHTS--RRRVTGLEYRVVYRVLNAANHGVPO
RAAC02539FEYILLQICYPEIVPENEDWTDHLRLEQQHT--KGRYHGLSYRVVFRLLNAADYGVPO
gb|ABY83631.1|FQYIQLQLEFPEVAPRNGEDWSEHLKRLQEEKTSKNGRGLTYNVLATLVNAADYGIPO
ref|YP_208280.1|FEYILLRLTYPNLQILQNEWDWGHILRLKEIEFN---LYKGIKYKVSQQLNAADYGVPO
ref|YP_001516905.1|FEYILLRLRYPDLSLQASERWEDHLARLEKIHTS--GKYDGIKYNVVRFLVDAANYGVPO
..:* *: :*: :*. *.**: :*. * : : : :*: :*

ref|ZP_01034116.1|RRERVFLVGFRADTGIEWHFPPKPTHSRDALLWSQWRDEVYWDLHRVARK--NRP-EGGAA
ref|ZP_02297879.1|RRERVVVFGRSDDLKWAFFAETHSLDALLWDQVHGD-YWDRHRVRA--DRN-IGDRY
RAAC02539KRERVFIVGFRSDDLNIWSFNPPTHSFDALLYDQWVTGEYWERHRVAKK--HRPDI PKL
gb|ABY83631.1|KRERVFIVGFRSDDLNIWSFNPPTHSFDALLYDQWVTGEYWERHRVAKK--HRPDI PKL
ref|YP_208280.1|KRERVFIVGFRSDDLNIWSFNPPTHSFDALLYDQWVTGEYWERHRVAKK--HRPDI PKL
ref|YP_001516905.1|KRERVFIVGFRSDDLNIWSFNPPTHSFDALLYDQWVTGEYWERHRVAKK--HRPDI PKL
:****.:***.*.*.*** :**.*.* : : : :** :

ref|ZP_01034116.1|KAR--SLKIADRPDL-EPWLTVRDAISDLPDPEHAPGTARGFHDHRFQPGARSYAGHTGS
ref|ZP_02297879.1|RQR--AARLGQPDPT-LPWRTTRDAIADLPDPEREPKQSMTYLNHRFPQARSYPGHTGS
RAAC02539KKRIESLRMGLRLIQAPWVTVRDAISGLPDPESE--QDCGVPNHRFPNGARVYPGHTGS
gb|ABY83631.1|VSR--IAKLRKNPPTDLPWKTVRDAIADLPDPETK--AAKMIRNHGFQAGARTYPGHTGS
ref|YP_208280.1|IAEKLQKKYGIFEPEKKPWQTVRDTLSDI PHP-LGN---HKITGHEYRDGARIYPGHTGS
ref|YP_001516905.1|RAKQLMQQPSLFSPLPEWKTVRDQIGKLPAPDIQ---SFDNEHVLREGARSYPGHTGS
. : : **.*.* : .:*.* * .***.******

ref|ZP_01034116.1|PLDEPAKTLKAGVHGVPGGENMLRRPDGVSRYFTIRESARLQTFPDDMVHGSWTETMRQ
ref|ZP_02297879.1|PLDEPAKTLKAGVHGVPGGENMLLRPDGVSRYFTVRESARLQTFPDDFRLHGSWSEAMRQ
RAAC02539PLDEPAKTLKAGDHGVPGGENMLVKPDGTVRYFTVRESARLQTFPDEYVFSGSWTESMRQ
gb|ABY83631.1|PLDLPAKTLKAGAHGVPGGENMLVRDDGVSRYFTVRESARIQTFPDSYRFHGSWTETMRQ
ref|YP_208280.1|GIDEPSKTIKAGGHGVPGGENMIRYDDGTVRYFTSYEAKLLQTFPEEFVISGAWGEAMRQ
ref|YP_001516905.1|YIDMPSKALKAGDHGVPGGENMIRYQDGRIRYTTFEAKRIQTFPDNYRISGSWTEAMRQ
:* *: : :*** *****: ** :***.**: :*****. : :*.**:***

ref|ZP_01034116.1|LGNAVPCQLARIVASGVRHKL-----
ref|ZP_02297879.1|LGNAVVELAHVASSVNAHLQFKAS-----
RAAC02539LGNAVVPVRLAHFITKDITRLEWSVADAGKAV
gb|ABY83631.1|LGNAVVPVHLAYCVASSVAEKL-----
ref|YP_208280.1|IGNAVPVKLSEILGKHL-----
ref|YP_001516905.1|IGNAVPVELGYRIANSI-----
:*****.*. : . :

```

FIG. 161

gb|ABO14793.1|
 ref|YP_001443312.1|
 ref|ZP_01815366.1|
 ref|YP_205672.1|
 ref|ZP_01042597.1|
 RAAC02543

---MKKQRAFLKWAGGKYGLVEDIQRLPPARKLVEPFVGAGSVFLNTDYDHYLLADINP
 ---MKKQRAFLKWAGGKYGLVEDIQRLPPARKLVEPFVGAGSVFLNTDYDHYLLADINP
 ---MKKQRAFLKWAGGKYGLVEDIQRLPPARKLVEPFVGAGSVFLNTDYEQYLLADINP
 ---MKKHRAFLKWAGGKYSLVEEIQRLPDARKLIEPFVGAGSVFLNTDYDHYLLADINP
 ---MKRNRAFLKWAGGKYGLVEPIKACLPEGRKLIEPFVGAGSVFLNTDYDRYLLNDINP
 MQSMPIRRSYLQWPGGKSHIVSALRAFLPPGRRLIEPFVGASSVFLNTDYPEYLLGDANP
 * .*:*:*.*** :*. :. ** .*:*:*****.***** .*** * **

gb|ABO14793.1|
 ref|YP_001443312.1|
 ref|ZP_01815366.1|
 ref|YP_205672.1|
 ref|ZP_01042597.1|
 RAAC02543

DLINLYNLLKERPEEYISEAKRWFVAENNRKEAYLNIRAEFNKTDDVMYRSLAFLYMNRF
 DLINLYNLLKERPEEYISEAKRWFVAENNRKEAYLDIRAEFNKTDDVMYRSLAFLYMNRF
 DLINLYNLLKTDPEYIAEAKRWFCEPNNRKEVFLDIRAQFNDTDNVMYRSLAFLYMNRF
 DLINLYNHLKDEPERFIEDARALFQPEYNKKEVYLALRVEFNQCTDTRYRSLFLYMNRF
 DLITLYQFVKRRPKTFIADARRLFTARNNQSDAYYALRKSFNESNDPYRSLFLYLNRF
 DLILVHRTLQAYGEAFIGACRELFPENNAPERYYALRQEFNATQNLWRRAILFVYLNRF
 *** :. :. : : * . : * . * : : : * .** : * : * : * .**

gb|ABO14793.1|
 ref|YP_001443312.1|
 ref|ZP_01815366.1|
 ref|YP_205672.1|
 ref|ZP_01042597.1|
 RAAC02543

GFNGLCRYNKKGGFNVFFGSYKKPYFPEAELEFFAEKAKKATFVCEGYPETFSRARKGSV
 GFNGLCRYNKKGGFNVFFGSYKKPYFPEAELEFFAEKAKKATFVCEGYPETFSRARKGSV
 GFNGLCRYNKKGGFNVFFGSYKKPYFPEAELEFFAEKAKKATFVCEGYHETFSRARKGCV
 GFNGLCRYNKKGGFNVFFGSYKKPYFPEAELEFFAEKAKKATFVCEGYMDTFKRARKGAV
 GYNGLCRYNSRGIFNVFFGDYRKPYFPEAEIEVFAEKAKRAKFCVCSFEQVFRARKGDV
 GIHGLMRYNRRGAFNTFFGYRRKIYFPEAEMRLFAEHAKRASVHADFRDLMSQARPGDV
 * : ** ** : * ** .*** : * ** ** : : .***.***.*** : : : : ** * *

gb|ABO14793.1|
 ref|YP_001443312.1|
 ref|ZP_01815366.1|
 ref|YP_205672.1|
 ref|ZP_01042597.1|
 RAAC02543

VYCDPPYAPLSNTANFTSYAGNGFTLDDQAALADIAEKAATERGIPVLISNHDTTLTRRL
 VYCDPPYAPLSNTANFTSYAGNGFTLDDQAALADIAEKAATERGIPVLISNHDTTLTRRL
 VYCDPPYAPLSNTANFTSYAGNGFTLDDQAALADIAEKAATERGIPVLISNHDTTLTRRL
 VYCDPPYAPLSNTANFTSYAGNGFTLDDQAALADIAEKAATERGIPVLISNHDTTLTRRL
 IYCDPPYAPLVQASNTSYATGGFSLDDQSELARHAIKRAKRAHIVLISNHDTPLTRAL
 VYCDPPYVPLSDTANFTSYATGGFSLDDQSELARHAIKRAKRAHIVLISNHDTPLTRAL
 :*****.*** : :***.*** .** : : : ** * . * . * . * . * . * . *

gb|ABO14793.1|
 ref|YP_001443312.1|
 ref|ZP_01815366.1|
 ref|YP_205672.1|
 ref|ZP_01042597.1|
 RAAC02543

YHGAELNVVKVRTIS-RNGSGRNKVDELLALFRAP
 YHGAELNVVKVRTIS-RNGSGRNKVDELLALFRAP
 YHGAELNVVKVRTIS-RNGAGRNKVDELLALF---
 YHGATLNTIKVRTIS-RNGAGRNKVDELMALFHAP
 YERAKISSLKVARNIS-QKGDARKPVAELLALF---
 YRGAQLVSVRAPRNIGVQHRQAARTAEELAIIFWPP
 * . * : : . * . * . : . . . * : * : *

FIG. 162

```

ref|YP_001275109.1| -----RRERKRVIAFGWYGGKYSHLDWLLPLLPPFCHHY
ref|ZP_01514022.1| -----KRSKKRIAFGWYGGKYSHLDWLLPLLPSCHHY
ref|YP_001633718.1| -----QRRTRRLIAFGWYGGKYSHLDWLLPLLPPCHHY
ref|ZP_01631840.1| -----KKIAFGWYGGKYSHLDWLLPLLPTITHY
gb|AAZ73681.1| -----KNKLIAFGWYGGKYSHLDWLLPLLPKTTHY
RAAC02564 MGVMLLSFIEGLPPFTALMIDHVSTARQKPRRKRIAFGWYGGKYSHLSWLLPLLPRAHY
: :*****.***** **

ref|YP_001275109.1| CEPFGGSAAVLNFNREPSPVETYNDIDGEVNVFFRVLRDDPERLVRRAIGLTPFSREEFAIA
ref|ZP_01514022.1| CEPFGGSAAVLNFNREPSPVETYNDIDGDVVHFFRVLRDEPERLVRRAIGLTPFSREEFAIA
ref|YP_001633718.1| CEPFGGSAAVLLNRDPAPVETYNDIDGEVTNFFRVLRDQPDQLTRAIALTPFSREEFALA
ref|ZP_01631840.1| CEPFGGSAAVLLNREPAPVETYNDIDGQVVNFFRVLRDQDELIAIGLTPFSREEFRIA
gb|AAZ73681.1| CEPFGGSASILINREPSPVETYNDIDGELVNFFRVLRDEKNEILIRAIATPFSRSEFELA
RAAC02564 CEPFGGSAAVLLNRDPSPVETYNDLDGEVNVFFRVLREQDALIEAIALTPFSREEFVRA
*****:*.**:*.*****:*.:.*****: : * .*.*****.* *

ref|YP_001275109.1| CEIDPA-LDPVERARRFYVRARQVRTGLAQSATLGRWANCKNTSRSGMAGAVSRWLGAIVE
ref|ZP_01514022.1| CEIDPS-LDPVECARRFYVRARQVRTGLAQSATLGRWANCKNTSRAGMAGAVSRWLGAIE
ref|YP_001633718.1| CEIDPS-LEPIERARRFYVRARQVRTGLAQSATLGRWANCKQTSRSGMAGAISRWFGAIE
ref|ZP_01631840.1| IAKDETGLSDLEKARRFFVRARQVRTGLAQTSIGRWAHCKLTSRAGMAGAVSRWLGSVE
gb|AAZ73681.1| ISKDTTNLSNLERARRFFIRARQVRTGLAQTSAGRWAHCLLTSRAGMAGAVSRWLGSID
RAAC02564 IVEPTGGLTELERARRFYVRARQARMALAQATATPGRWAACRNTSQRGMAGNVSRWLGSAA
* :* *****:*****.* *****: ***** * **: ***** :*****:

ref|YP_001275109.1| DLPEIALRLLRVQIENRPAIEVIRLYDSADTLFYCDPPYIHETRGDDNAYAYEMNYEEHC
ref|ZP_01514022.1| DLPDIALRLLRVQIENRPSIEVIRLYDSPDTLFYCDPPYVHDTWGDNNAYAYEMNNEEHR
ref|YP_001633718.1| DLPDIAIRLLRVQIENRPAIEVIRLYDSPDTLFYCDPPYVHETRGDDNAYAYEMSNEEHQ
ref|ZP_01631840.1| HLPFIVQRLLRVQIENDTAIDIIQRYDSSETLFYCDPPYPHDSRGDSKAYAYEMTDEEHR
gb|AAZ73681.1| DLSKIVQRLQRVQIENSSSFDIIRRYDSEETLFYCDPPYPHNTRGDSKAYAYEMTDEEHR
RAAC02564 NLAEIAARLVRVQIEHRPAMDVIRAYDSPDTLFYCDPPYIHATRTDPRAYRYEMSDEEHR
.*..*. ** *****: .:..*: ** :***** * : * .** *** :**

ref|YP_001275109.1| ELAMSLNSVRGLVAISGYECDLMNDLYPSSRWYKSVAPSKTIHSTKDKRTEILWTNY
ref|ZP_01514022.1| QLALVLNSVQGFVAISGYDCDLMDLYPSPNWKNIAPTCTIHSK-----
ref|YP_001633718.1| QLADVLNSVQGFVAISGYDCDLMDLYSPKWKIKHVGPTRTIHSKDKQRTTEILWTNY
ref|ZP_01631840.1| QLAGVLRSVKGKVALSGYDCTLMQELYG--DWNCIKAPQKQCHSIKELRQEVLVVNY
gb|AAZ73681.1| KLAEIVHNIKGVAVSGYECTLMNELYG--DWKQISAPTKSCHSVKKPRTERLWVNY
RAAC02564 ELARVLHNVGRGKVAISGYHCDLMDELYR--DWRCIDAPPKQCHASKRWRQALWVNY
:** :..*: *******.* *****: * .* : * : *

```

FIG. 163

```
ref|YP_853610.1|-----FESQADIRVIVINGEPWFIASDVCRAIGIANHRDAVR
ref|YP_512277.1|-----FESQADIRVIVINGEPWFIASDVCRAIGIANHRDAVR
ref|ZP_01959153.1|-----NSTPIRVQVINNEPWFVAKDVCIDLGISKYRDAIA
ref|YP_001038857.1|-----GKQVRTFIIDGEPWWVAKDVCDILELGDTHKAME
ref|YP_001662865.1|-----GNTVRTVMKDGPNPWWVLKDVCSVLDIGNSRDVMA
RAAC02566      MPFCPADIALRRGRSLMQQTWAVFANGAQVRVFWVDGEPWFDAVGVCAMGLRNIEKAIR
                  :*.  :.:***:  .**  : : . . . . :

ref|YP_853610.1|      KLDDDEKG-VASTDTPGGEQESIIISESGLYTLILRCRAVTPGTIPYRFRKWVTGEVLP
ref|YP_512277.1|      KLDDDEKG-VASTDTPGGEQESIIISESGLYTLILRCRAVTPGTIPYRFRKWVTGEVLP
ref|ZP_01959153.1|      RLDDDEGC-PIEVDTLGGMQKMAAVNESGLYTLILQSR---KPEAKP--FRKWVTSEVLP
ref|YP_001038857.1|      RLDEDERNTIPVTDLSLGRQLQETYVVNEAGLYNLILGSR---KQEAKE--FKRWITHEVIP
ref|YP_001662865.1|      RLDSDEKG-VDIIDTPGGKQEVSIINESGLYSVILVSR---KPEAKK--FKRWVTHEVLP
RAAC02566      RLDDDEKG-LVTVDNAGGREEILVVRESGMLRLALAGR---EPHARA--FQRWVREVLV
                  :**. **      *. *  ::  : *: *:  : * *      :      *: *: . ** : *

ref|YP_853610.1|      -QIRRTGSYIKNSLPQEERIKMVA-----
ref|YP_512277.1|      -QIRRTGSYIKNSLPQEERIKMVA-----
ref|ZP_01959153.1|      -----
ref|YP_001038857.1|      -QIRKTGIYALEPKQLLAVAIIEAQKIIEEQDR-----
ref|YP_001662865.1|      -SIRRHGLYATD--ELL-----
RAAC02566      LAIRGQRPNALEEQRLLEKVSLLVEILAAYQDR LSDGGIEHLAKTIANLLT
```

FIG. 164

```

ref|YP_375842.1|-----MSKAEELVEKIAAQAKLTKVDAERAVNAFINVVTSSLKGGD
ref|YP_001131112.1|-----MSKAEELVEKIAAQANLTKVDAEKSVNAFINVVTSSLKAGD
ref|ZP_00591928.1|-----LMSKAEELVEKIASQAGLTKADAERAVNSFVSVVTDLSLKAGE
ref|YP_001003150.1|-----MNKADLAEEKVAAETGVSKRVATDAVSAVFTGIEESLASGE
ref|NP_046614.1|-----GAPSMNKTELIKVAEKQGVSKKEGAPSVKVFDTISEALKSGE
RAAC02589      MCCLTGLERMRDLLEKERELVNKGELVAEVQARVGLPKSQVLQVLNTFCEVTTERLQAGE
                  :.* :* :: . :.*      :. . . * .*:

ref|YP_375842.1|      DVTLVGFGTFTTGDRAARQGRNPQTGKAITIPAKKVVKFKPGKALKDEV-
ref|YP_001131112.1|    DVTLVGFGTFTTGDRAERQGRNPQTGKTITIPAKTVVKFKPGKALKDEV-
ref|ZP_00591928.1|    DVTLVGFGTFSVGERAERQGRNPQTGETITIAARKAVKFKPGKALKKEEVD
ref|YP_001003150.1|    DVSIPGFGKFAVVARPERQGRNPQTGELIDIPAGMNVRFKPGAPLKRSVD
ref|NP_046614.1|      KVSIPGFGTFEVRERAARKGRNPQTGEEIDIPATKAPAFKPAKALKDAV-
RAAC02589      EVSLPPLGKFQYVMRSARRQRNPQTGEMIDVPEKATVRFPSGALKGRVN
.*:: :*. * . *: *****: * :.      *:* .** *

```

FIG. 165

```

ref|YP_001127101.1|-----YLIPTVIEQTNRGERAYDIYSRLLKD
ref|YP_148915.1|-----YLIPTVIEQTNRGERAYDIYSRLLKD
ref|YP_080797.1|SIGNIFICANCEE- IDENTITIESPSITIVESGAPSIPTVIEQTNRGERAYDIYSRLLKD
ref|YP_001488316.1|-----IPTVIEQTNRGERAYDIYSRLLKD
ref|ZP_02848186.1|-----VPIVVEQTNRGERSYDIYSRLLKD
RAAC02045-----MSSTWPGGEGGCCMYYPYVIEQTSRGERTYDIYSRLLKD
                                     : * *:***.****:*****

ref|YP_001127101.1|RIIFLGSPIDDQVANSIVSQLLFLAEDPEKDISLYINSPGGSITAGLAIYDTMQFIKPD
ref|YP_148915.1|RIVFLGSPIDDQVANSIVSQLLFLAEDDPKDISLYINSPGGSITAGLAIYDTMQFIKPD
ref|YP_080797.1|RIIMLGSAIDDNVANSIVSQLLFLEAEDPEKDISIYINSPGGSITAGMAIYDTMQFIKPD
ref|YP_001488316.1|RIIMLGSAIDDNVANSIVSQLLFLEAEDPEKDISIYINSPGGSITAGMAIYDTMQFIKPK
ref|ZP_02848186.1|RIIFLGSAIDDDVANSVIAQLLFLAADDPEKDIHLYINSPGGSVTAGMAIYDTMQYIKPD
RAAC02045RIVILGTPIDDQVANSIVAQLLFLAADDPEKDIQMYINSPGGSVTAGMAIYDTMQHIRPA
***:***.***:****:***** *:***:*** :*****:***:*****.*:*

ref|YP_001127101.1|VSTICIGMAASMGAFLLAAGAKGKRFALPNSEVMIHQPLGGAQQGQATEIEIAAKRILFLR
ref|YP_148915.1|VSTICIGMAASMGAFLLAAGAKGKRFALPNSEIMIHQPLGGAQQGQATEIEIAAKRILFLR
ref|YP_080797.1|VSTICTGMAASMGAFLLAAGEKGKRYALPNSEVMIHQPLGGAQQGQATEIEIAAKRILSLR
ref|YP_001488316.1|VSTICIGMAASMGAFLLAAGEKGKRYALPNSEVMIHQPLGGAQQGQATEIEIAAKRILSLR
ref|ZP_02848186.1|VSTICMGMAASMGSLLLTAGAKGKRFALPNAEVMIHQPLGGVVRGQASDIKIHADWILKTK
RAAC02045VSTLCIGMAASMGAFLLAAGEKGKRYALPNAEIMIHQPLGGVEGQASDIKIHAEWILKTK
***.* *****:***:*** *****:*****:*****.***:***.* ** :

ref|YP_001127101.1|DKLNRILAENTGQPIEVIERDTRDRNFMTAQKAQEYGIIDRVL-----
ref|YP_148915.1|DKLNRILSENTGQPIEVIERDTRDRNFMTAQKAMEYGIIDRVL-----
ref|YP_080797.1|DKLKNILAERTGQPLEVIERDTRDRNFKTAEEAKEYGLIDKVL-----
ref|YP_001488316.1|DKLNQVLAERTGQPIEVIERDTRDRNFKTAEEALQYGLIDKVL-----
ref|ZP_02848186.1|QKLNEIYVERTGQPYEKIDRDTDRDNFMSAEDALNYGLIDKVI-----
RAAC02045DKLNRLLSERTGQPLEVIERDTRDRHFMTAEEAKAYGLIDEVLHPLHFPVR
:***.: *.***** * *:*****:* :*:.* **:***.*:

```

FIG. 166

```

ref|YP_866618.1| -----IKAPKKLIEVALPLDDINAAAAREKSIRHG-HPSTLHLWWARRPLAAARAVLFA
ref|YP_342704.1| -----EIKTPKKLIEVALPLDDINTAAAREKSIRHG-HPSTLHLWWARRPLAAARAVLFA
ref|NP_384606.1| -----VRTPKKLIEVALPLDAINEAAAAREKSIRHG-HPSTLHLWWARRPLAAARAVIFA
ref|YP_911114.1| -----IKSPRKLIEVALPLDAINAAACAYEKMPGIGAHPRGIHLWWARRPLAAARAVLFA
ref|YP_115396.1| -----VKAPKKLIEVALPLDAINEASAREKSIRHG-HPSTLHLWWARRPLAAARAVIFA
RAAC02635 MYVVPKVKAPKKLIEVTLPLDDINAEAAAREKSIRHG-HPSTLHLWWARRPLAAARAVLFA
               :::::*****:**** ** . * ** * ** :*****:***

ref|YP_866618.1| QMVNDP----GGERG--YYAGTKKAQADAEREELFKI IRELVLWENTNNEEVLNKARAAI
ref|YP_342704.1| QMVNDP----GYQQGEGFKYGVNKKAEIKREKLFQI IRDLVKWENTNNEEVLNRAREAI
ref|NP_384606.1| QMVNDPSWKWELEHPGEIPPNNIKASWAASRNRLFAI IKDLVEWENTTNEEVLEKARAEI
ref|YP_911114.1| QLVNDP----GYQQGCGFKYGVNKKAEAI ERKRLFKI IEELVLWENTTNEEVLERARVEI
ref|YP_115396.1| QMVNDP----GYQQGGGFRYGVNKEKAQLERERLFI IEELVQWENTNNEAVLSRARAEI
RAAC02635 QLVNDP----GGERG--YKPGMTREQAQKERERLFGIMRRLVKWENSNDDEEVLREAREEI
               *:****      ::      . : . . :*:*** *: . ** ***:.* ** .** *

ref|YP_866618.1| RKSWRETCELNK----GKSGFDPDKLPFAHDPFAGGGAI PLEAQRGLMESHASDLNPVAV
ref|YP_342704.1| WESWRETCHLNRNHPQAAELFNPDKLPFAHDPFAGGGAI PLEAQRGLGESYASDLNPVAV
ref|NP_384606.1| RKSWRETCDLNKDHPQAAELFNPERLPFAHDPFAGGGAI PLEAQRGLGLASYASDLNPVAV
ref|YP_911114.1| RRSWREVCELNKEHPQAAELFNPEKMPAFHDPFAGGGAI PLEAQRGLGLEAYASDLNPVAV
ref|YP_115396.1| WKSWRETCELNKNHPCAAELFNPDKLPFAHDPFAGGGAI PLEAQRGLGESYASDLNPVAV
RAAC02635 WKSWRETCEMNK----GLPGFDPDKLPFAHDPFAGGGI PLEAQRGLGLEAYASDLNPVAV
               .****.*.:*:      . :*::::*****:*****: :*****

ref|YP_866618.1| LINKAMIEIPPKFAGRKFPVGPIPEGEKQG-----RMESDWPGATGLAEDVRRYGHWMRE
ref|YP_342704.1| MINKAMIEIPPKFAGQRPVGPPLPQGEKQG-----KLMDWDSGARGLAEDVRRYGHWMRE
ref|NP_384606.1| LINKATIEIPPKFAGRPVPVNPPEARTSRD-----AWSKQWFRAQGLAEDVRRYGRWMMRM
ref|YP_911114.1| LINKAMIEIPPKFAGRPVGPPEIENKQKSL---ELPKIWPGATGLAEDVRRYGSWMRD
ref|YP_115396.1| TINKAMIEIPPRFAGRAPVGPVPPSPDGRGVGGEGFLAQDWAGAKGLAEDVRRYGAWMRS
RAAC02635 LINKALIEIPPKFSGRPFPVNPPEARSKVG-----LLEQEWKGASGLADIRYYGEWMRD
               **** *****:*.*: **.*      . * * ***:.* ** ***

ref|YP_866618.1| EAFKRIGHLYPQVEITAEMAKERPDLKGIVGQKLTVIAWLWARTVRSNPAPFASHIAPLV
ref|YP_342704.1| EAFERIGNLYPRIKITQEMVAERPDLKPYQGQELTVIAWLWARTVKSNPAPFASHADIPLA
ref|NP_384606.1| VAQKRIGHLYPPVEITSDMAKERPDLPLVGQQLTVIAWLWARTVKSNPAPFRHVDVPLA
ref|YP_911114.1| EAQKRIGHLYPPVEVTEMARERSDLKPLVGKLTVIAWLWTRTVKSNPAPFTHVNVPLV
ref|YP_115396.1| EAEKRIGHLYPQVEVTRELAQGRQDLQPLVGQKLTVIAWLWARTVKSNPAPFASHVEVPLA
RAAC02635 EAFKRIGHLYPQVELPPEYG-----GGKATVIAWIWARTVKSNPAPYSHVDVPLV
               * :***:*** :.: . : * : *****:*****: * :**

ref|YP_866618.1| SSFLLSTKKGKEAYIEPVVD-----ANSYFVSVKKGTPS-----KDSAR
ref|YP_342704.1| SSFLLSTKKGKESYVNPLVE-----GHNYQFEVCMGVPP-----AEARN
ref|NP_384606.1| SSFVLSTKTGAAYVDVID-----KDTYHFAVRSGRAP-----SQARE
ref|YP_911114.1| SSYILANKDGKEVYVKPVIE-----GDKYYFLIKNGTPT-----TEAKD
ref|YP_115396.1| STFVLSSKAGKEAYVQPMISPLPLGEGLVGRAGSEGYRFTVQVAGTPGFDKADYARAKS
RAAC02635 RSFWLSTKKGKEVWVEPVIHE-----DGSGYHFEVRTGKP-----KIE
               :: *.:* * * :.:*:      * * : .

ref|YP_866618.1| GTSAGK-RGGFRCIFSDAPIDYNYIRDEGSAGRIGTRLMAIVAEGVRGRIYLSATPELEI
ref|YP_342704.1| GTKLGR-GANFTCLLSDTPIDPKYIYAQAQSGNLGQRLMAVVAEGKSGRIYLTPTAEMEQ
ref|NP_384606.1| GTKFSR--GNFRCLLSQAPIDGDYIKAEAKAGRMGERLMAIVAEGRNRIYLS PSSDQEN
ref|YP_911114.1| GTKATVRGANFRCIVSGAVIGSDYIKTEGKAGRIGQRLMAIVAEGTRGRVYLPPTLDAET
ref|YP_115396.1| GTKLAR-GANFECLLSNTPIEPNHIYTEANAGRMGARLMAIVAEGARGVYLPPLPEHEA
RAAC02635 GTVNR--GGTCIMSGSPIFFDYIRAEGKAERMGQRLMTVVVEGPRGRLYIAPTDDMEG
               **      . *: * : * :.* :.: .:* *****:*** *:.* : *

ref|YP_866618.1| IANSAK-PEWSPDVKLHGKCR-VNVSNYGLDVYSDLFTRPQLVALTTFSNLVQEARVKAV
ref|YP_342704.1| AASAAS-PDWKPDALMPENPRWFSPPMYGMKSYGDLFTPRPQLVALNTFSDLVQEACYKAI
ref|NP_384606.1| IASAAH-PDWKPDVEFFQALGFRVGNYGMTKWSDLFTARQLVALTTFTELIAEVKQIV
ref|YP_911114.1| VAKDAA-PTWRPSGDVPERLTGGTCVPYGLKEWGDIFTPRPQLVALTILSDLLSEVRELVR
ref|YP_115396.1| IARQAQ-PEWKPEVAMPDNPRWFSPPYGLKNYGDLFTPRPQLVALTTFSDLVDAIERCR
RAAC02635 MALSLKEPDGLPQTQLPKRALGFRVQEGYGMTMHKHLFTNRQLVALKTFSDLVVEAREKVY
               * * * . . ** : .:** ***** :.:* :.

```

ref|YP_866618.1| NDAKITGMADDGMGIDEGGGFAGAYGDAVAVYLGFIIVDKVSESLSTICTWSSSPKNE---
ref|YP_342704.1| ADAKAAGMTDDGIGIDDDGGRGATAYGDALAVYLTFAINKLADRGSTICTWSSSRSS----
ref|NP_384606.1| ADAIAAGMIDDQTGLDKGGEGASAYAEAVSVYLAIALSRLTIDCNALCRWEVTKT-----
ref|YP_911114.1| EDALASGLSDSEKLGDDGGSGVIAYAEAVSVYLAFAINRCADFCNSVTRWVPGNQ-----
ref|YP_115396.1| RDAAGLEPDDGVPLDAGGTGATAYAQAVGVYLAIAISRFSDRNNISICTWDSGPTGTAKS
RAAC02635 EDALSSGMTDDGIGLEDGGIGARAYADAVAVYLGVLVLDKADYWSSICSWNSPKE-----
* * : * : * : * * * . * : * : * : * : * : *

ref|YP_866618.1| -----LIVSTFRRQAI PMTWDFGEANPFANSSGSLEKIVPAVSKVIKTSLCGSVDGN
ref|YP_342704.1| -----TRNTFGRQAI PMTWDFAEPNPLSDSTGNFMGGIGWANDVLSRMIP-SSGGI
ref|NP_384606.1| -----QVRNLFQRQAI PMLWDFAEENNVEFGGAAGDYIISLGNMVKALE-KLPARDPGV
ref|YP_911114.1| -----KVMNLFQKQAI PMTWDPYEAAILADTVGGFAPASKYVADCIG-KLSPAAGVF
ref|YP_115396.1| TGGSARTASLRNLFARQAI PMAWDFGEANPFSDSGGGFSSAFEWIEPAVR-SLRGGCAGY
RAAC02635 -----LVRNTFSRQAI PMVWDFAEANPFSDSTGNWGMVMDWVWKAVERNAPIKTTLGE
* : * : * : * : * : * : * : *

ref|YP_866618.1| AIQFDARTVNL-S-DRVSTDPPIYDNIGYADLSDDFFVWSRRALKSIFPSLYSTLAVPKA
ref|YP_342704.1| AVQQDAATQNI SAEKVISTDPPIYDNIGYADLSDDFFVWMRRSLKSFYPSLEFATMAVPKA
ref|NP_384606.1| SRQQAQTQSISAGKAI STDPPIYDNIGYADLSDDFFVWLREPLRSIYPLGFATVVTPKA
ref|YP_911114.1| ASQADAQTQSISMAKII STDPPIYDNIGYADLSDDFFVWLRLSLRFLPLGFSTITVPKV
ref|YP_115396.1| GDGADAQTQTL SRDKVSTDPPIYDNIGYADLSDDFFVWLRLSLKPIFPGLYATLAVPKA
RAAC02635 AHLADAQTQSISVGKIVSTDPPIYDNIGYADLSDDFFVWLRLSLNPLPLGFATLMTPE
* * * : * : * : * : * : * : * : *

ref|YP_866618.1| EELVATPYRHGSKEEAEAFFMNGMICAINNFANQAHPSFPVTIYYAFKQSETK-ETGTTSS
ref|YP_342704.1| EELVAIPYRHGTKEKAETFFLDGMTQAIHNMADKGHAPFPVSIYYAFKQSETK-EGATSN
ref|NP_384606.1| EELVATPARHGGSEAEELFFLGGMTAAMQRLAELAHPTPTIYYAFKQSETESDGTGSS
ref|YP_911114.1| EELVAAPYRHGTQKKAETFFLTGMTETAIHNLAEQAHPEGPVTIYYAFKQSETSGQDGTSS
ref|YP_115396.1| EELVATPYRHGSKEEAEAFFLDGMRRALKNLAEQAHAPFPVTIYYAFKQSETTDAAGTSS
RAAC02635 DELVAAPERHQSKGDANSFFIHGMSKAMSTLANSTHPFPVTIYYAFKQSDTE-NEGTSS
* * * : * : * : * : * : * : * : *

ref|YP_866618.1| TGWETFLEAVIQAGFGITGTWPMRTERGARSIGIGANALASSIILVCRKRDNSAESISRR
ref|YP_342704.1| TGWETFLEAVIRAGFSIDGTWPMRTEMNSRMIGSGTNALASSVVLVCKKREIEAESISRR
ref|NP_384606.1| TGWETFLEDAVIRSGALTGTWPMRTELGNRMRGQDSNALASSIVMVCRRPATAETVSRR
ref|YP_911114.1| PGWVTFLSAVLSAGFAIVGTWPLRSEQEFMIGMGANALASSIVLVCRRSADAPSVSRR
ref|YP_115396.1| TGWETFLEQAVLDAGFALTGTWPMRTELGNRMIGAGTNALASSIVLVCRRATDAPTASRR
RAAC02635 EGWVAFLEALLQAGFAVGTWPLRTEMNSRMRLDSNALASSIVVLCRRRPAETAETISRR
* * : * : * : * : * : * : * : * : *

ref|YP_866618.1| QFQRELREILPEALETMIGG---KEGASP-----VAPVDLAQASIGPGMAVYSKYAAV
ref|YP_342704.1| DFQRELREQMPDALEAMIGG---ETGTTTP-----IAPVDLAQAAIGPGMAIFSKYEAV
ref|NP_384606.1| AFLRELNQVLPEALDEMTRG---SGDDRS-----VAPVDLSQAIIGPGMAVFSKYAAV
ref|YP_911114.1| EFIRELNGVLPEALDEMTRG---SGEERS-----VAPVDLSQAIIGPGMAVFSKYSAV
ref|YP_115396.1| EFLRELNATLPEAIADMGADPSPQPLSPRERGGRVAPVDLSQAIIGPGMAIFSKYAAV
RAAC02635 AFLRELNTVIPDALSDMMHG---GPDRSP-----IAAVDLQQAIGPGMAVFSKYAAV
* * * : * : * : * : * : * : * : *

ref|YP_866618.1| LNQDGNPMSVHDALILINREITDFTLTPDSGSFDNDTLFCSTWFDQYGWKAGPFGGEADTSL
ref|YP_342704.1| LNQDGSRMSVHDALILINRAITEYLSPESGSFDADTQFCSSWFDQYGWSTGPFGEANVLA
ref|NP_384606.1| LEADGTPMTVQAALRLINR---FLAEDD---FDHDSQFCLHWFEQYGWKAGPFGGEADTSL
ref|YP_911114.1| LEADGTPMSVRTALQLINR---FLAEDD---FDPDTQFCLHWFEQYGWENENLFGEADVLA
ref|YP_115396.1| LEADGTPMTVKTALALINR---FLAEDD---FDHDTQFCLHWFEQYGWASGKYGEADVLA
RAAC02635 LEADGSPMTVKTALQLINRVVDAYLHASEAEVDADTLFCINWFDQYGWSEADFGRAVDLA
* : * : * : * : * : * : * : *

ref|YP_866618.1| RAKGTSVDGVQEAGVVQSGGGKVRLLFKWDEYDPPDWDPKKNRTPVWEALHHLIRALNKDG
ref|YP_342704.1| QAKGTTVDGVNTAGVVS GGGKVRLLKWAIEYADWDPIKDNRTPIWEACHQMIRSLNNQG
ref|NP_384606.1| RAKGTSVDGVKQSGVIYASGGIVRLLKWAIEYPSDWDVPDGRPLVWEALHHLIRIFKSEG
ref|YP_911114.1| RSKSTSVDAMKEAGVLQSGSGKVRLLKWAIEYPSDWDPRDTRKMPVWEALHQLIRALKQGG
ref|YP_115396.1| RAKGTAVDALVAAGVAESA KGSVRLLKWEYPADWSPESDTRTPWEALHQLIRALNQAG
RAAC02635 RAKGTSVDAVRAAGVLTADHGKVRLLRWQDYPSDWSPGKDYDTPVWEALHHLIRALQDQD
* : * : * : * : * : * : * : *

ref|YP_866618.1| ESVS GGLLARMPERAEAIRQLAYHLYTL CERKKWADDARAYNELITSWHGIAAASHEVGH
ref|YP_342704.1| ESAAGELLAKMPEKGEPIRQLAYHLYTL CERKKWAEDARAYNELIGSWHAIVTASHEVGH
ref|NP_384606.1| ESGAGNVLA AAVA KAEAPRQLAYRLYTL CERAGWAEDARAYNDIITSWSAIESAAAA--
ref|YP_911114.1| ESASGALLAALGGKAEAVRQLAYRLYTL CERLGQAEDARAYNELITSWTGIESVA-----
ref|YP_115396.1| ETEAGRLLARMPARAEP IRLAYRLYTL CERKGWAEDARAYNELVTAWSGIEQAANEAGV
RAAC02635 EQAAGALLAGMYSVSESIRSLAYRLYTL CERKGWAEDARAYNELIASWDAIEVAQKAGL
* : * : * : * : * : * : * : *

| | |
|-----------------|------------|
| ref YP_866618.1 | LGTQCTLDL- |
| ref YP_342704.1 | SGSQAEGLD |
| ref NP_384606.1 | ----- |
| ref YP_911114.1 | ----- |
| ref YP_115396.1 | VGAQMQLL- |
| RAAC02635 | SGSQLTLDVE |

FIG. 167

```

ref|NP_244431.1|      MNLIPTVIEQTNRGERAYDIYSRLLKDRIIMLGTAIDDNVANSIVAQLLEFLQAEDPKDI
ref|YP_176521.1|      MNLIPTVIEQTNRGERAYDIYSRLLKDRIIMLGSGVDDNVANSIVAQMLFLQAEDPEKDI
ref|ZP_02171648.1|    MNLVPTVIEQTNRGERAYDIYSRLLKDRIIMLGTPIDDNVANSIVAQLLEFLAADDPKDI
RAAC00088             MSLIPYVIEQTSRGERSYDIYSRLLKDRIIFLGTPIDDDVANAVVAQLLEFLAADDPKDI
ref|ZP_02848186.1|    MNLVPIVVEQTNRGERSYDIYSRLLKDRIIFLGSAIDDDVANSVIAQLLEFLAADDPEKDI
ref|ZP_02210735.1|    MALVPTVVEQTNRGERAYDIYSRLLKDRIIFLSDEVNDTTASLVVAQLLEFLAEDPKDI
* * * * *:***.***:*****:*.  : * .*. :*:*** *:***:***

ref|NP_244431.1|      SLYINSPGGSITAGMAIYDTMQYIKPNVSTICIGMAASMGAFLLAAGAKGKRFALPNSEV
ref|YP_176521.1|      SLYINSPGGSITAGMAIYDTMQFIKPDVSTICIGMAASMGAFLLTAGAKGKRMALPNSEV
ref|ZP_02171648.1|    SLYINSPGGSITAGMAIYDTMQFIKPKVQTCIGMAASMGAFLLTAGEPGKRYALPNSEV
RAAC00088             QMYINSPGGSVSAGLAIYDTMQHIKPDVSTCMGMAASMAAVLLAAGAKGKRYALPNSEV
ref|ZP_02848186.1|    HLYINSPGGSVTAGMAIYDTMQYIKPDVSTICMGMAASMGSLLLTAGAKGKRFALPNAEV
ref|ZP_02210735.1|    HLYINSPGGSVTAGMAIYDTMQYIKPDVSTICIGMAASMGAFLLNAGAKGKRFALPNSEI
:*****:*.***:*****.***.*.***:*****.:** ** *** *****:

ref|NP_244431.1|      MIHQPLGGTRGQASDIEIHTRRILEMRETLNRILAERTGQPLEQIAKDTDRDNFMTAEKA
ref|YP_176521.1|      MIHQPLGGMQGQAADMEIHARRIQMREKLNQIMAERSGQPYERIARDTDRDNFMTAEQA
ref|ZP_02171648.1|    MIHQPLGGAQGQASDIEIHAKRIVEMKEKLNQILAERTGQDIEQIRKDTDRDNFLSAAQA
RAAC00088             MIHQPLGGARGQASDIEIHARHILKTRERLNRLAERTGQPLERVAQDTRDNFMSAEEA
ref|ZP_02848186.1|    MIHQPLGGVRGQASDIKIHADWILKTKQKLNEIYVERTGQPYEKIDRDTDRDNFMSAEDA
ref|ZP_02210735.1|    MIHQPLGGAKGQATDIEIHAKWILKIKERLNKILSERTGQPIEKIQEDTERDNFMSAQEA
***** :***.*:***: *:: : *.* **.* *:: .**:*::* . *

ref|NP_244431.1|      REYGLIDKVI--
ref|YP_176521.1|      KEYGLIDKVI--
ref|ZP_02171648.1|    KEYGLIDEVM--
RAAC00088             KAYGLIDEVIYR
ref|ZP_02848186.1|    LNYGLIDKVI--
ref|ZP_02210735.1|    KEYGLIDEVI--
*****:*.

```



```

ref|YP_146448.1|      --ERSIALYCRVSTDEQAREGISLDEQOERLKAYCRAMGWSEEPFLFIDDGYSAKNLD RP
ref|ZP_02848045.1|    ---RFVALYCRVSTDEQAREGVSLDEQVERLKAYCRAMGWSEQVQIYIDDGYSAKNLD RP
ref|NP_977168.1|      -SSQFIALYCRVSTDDQAKEGVSL EEQOERLEAYCRAMGWGEDVVVYVEDGYS AKSTDRP
ref|ZP_02261191.1|    -SSQFIALYCRVSTDDQAKEGVSL EEQOERLEAYCRAMGWGEDVVVYVEDGYS AKSTDRP
RAAC00111             MTERHLALYVRVSTEEQAASGHS LRQEQLVEYARSHGFPR-YELYCDDGYS GKSLNRP
ref|NP_981890.1|      -SKRKVAIYIRVSTEEQANEGFSLGAQEDY LKQYANTKG YEV-YDIYIDDGYS GKDYNRP
                        : :*: * ****: :** . * ** * : * * . : : * : : : ****. . : **

ref|YP_146448.1|      ELNRL LKEVKKGTISKILVT KLDRLSRRLD LLLKLIDMFQEHNV SFISISES-FDTNTPS
ref|ZP_02848045.1|    ELKRL LHAVKSGAVSRIMVT KLD RMSRRLD LLLNIIDL FQDHEV SFV SISES-FDTNTPS
ref|NP_977168.1|      HLKRLIEDV KAGEIKKIIVTKLD RMSRKL D LLLTLIDL FQEH DVSFV SISES-FDTNTPS
ref|ZP_02261191.1|    HLKRLIEDV KAGEIKKIIVTKLD RMSRKL D LLLTLIDL FQEH DVSFV SISES-FDTNTPS
RAAC00111             AMERLRADIRAGCVLG VVTTTRIDRL TRSVAD FARLVD E MNLHG VFYRSTRQN-FEISTAM
ref|NP_981890.1|      EIQR LFKDLYQE KFTGILVKSVDRI SRVSDVT KLIDDV LAPRKCYVLVSDNNLDSSTTG
                        : : ** : : . : : . : ** : : * . : : * . : : . : : . : *

ref|YP_146448.1|      GRLTLQVLGAVAEFERERIRERV FENMLHAARQ GKWLTQSPYGYRLQNKELVIYEPEAQI
ref|ZP_02848045.1|    GRLTLQVLGAVAEFERERIRERV FENMLHAATTG KWL TQSPYGYDLVDKSLVNIVESQI
ref|NP_977168.1|      GRLTLQVLGAVAEFERERIRERV FENMYHAASKGKWL TQHPFGYRLNKELVTYENEAKI
ref|ZP_02261191.1|    GRLTLQVLGAVAEFERERIRERV FENMYHAASKGKWL TQHPFGYRLNKELVTYENEAKI
RAAC00111             GRLVAQMLS VIAEFEREMIAERVHENLLAMAQRGELATKPPFGYRLVNGRLTPDPREARW
ref|NP_981890.1|      GTAFIQLLATFAEYERSMIVQRVKAGMSKRAELGYWNGGHVGLGYDVKNKELKVNKEEA EV
                        * : : * . . : ** : : * * : : * : : * : :

ref|YP_146448.1|      VRNVYDWYLNKGLGYAIAK K LNEEGIPSRQKK EWSIRA IKMLTNPVYKGT LFWNRMD S
ref|ZP_02848045.1|    VKRVYD MYLNEGLGFYSIAKRLNEEGVPSRHRKEWSIRA IKMLTNPAYKGT LVVNRKDS
ref|NP_977168.1|      VRDIFIWYTQEGHGFY AIAK K LNL LGIPSRQNKQWSIRS VKLLSNPVYKGT VVWNRKDR
ref|ZP_02261191.1|    VRDIFIWYTQEGHGFY AIAK K LNL LGIPSRQNKQWSIRS VKLLSNPVYKGT VVWNRKDR
RAAC00111             VREG-ARLLLTGASPRDVAMEFNALGVRTKTGR LWTDR TVRTLFSNPALAGVS VVWNRKT
ref|NP_981890.1|      VRKIFHLRT-EGKGYKSIALILNQ QGFKTKRGR LFSIGTIKTILENPLYIGYCRWGRHKN
                        * : : * . : * : * * . : : : : : : : : ** * * .

ref|YP_146448.1|      SKSKRVKKDDADWV VLEDACPAI IDQATWEQVQKR--VNK-QQMAPRAQTS PHLLGGLLK
ref|ZP_02848045.1|    SKTQRKDKDEAEWVSVEECLPVI IDQPMWDKVQQR--INQPSNIAPRAKTS PHLLGGM LR
ref|NP_977168.1|      SKKKEQFKKEDEWVIQENAAPAIVTDELWNEAQKR--MQT-STLAPRAQTS PHLLGGLLK
ref|ZP_02261191.1|    SKKKEQFKKEDEWVIQENAAPAIVTDELWNEAQKR--MQT-STLAPRAQTS PHLLGGLLK
RAAC00111             QGSRRTEREPT EWCVEGAHEA ILSREEFEALRGM--LERRRALS PRSQGSKRPLAGIAR
ref|NP_981890.1|      WNVERRKGT TDEYK FVGKHEPIVSEDLWKRVQGVNYAHKKS I SKNRNFNGEFVLSGILR
                        . . : : : : * : : : . * . * . : :

ref|YP_146448.1|      CGRCGSGMSIGWSGSAKNRYR----VYRCSANKNKGTCTSKQYRAD DVESWFKHG LLLKLA
ref|ZP_02848045.1|    CGNCGSGMSIGFSGSPNKRYR----VYRCSANKNKGTCTSKQYRA-----
ref|NP_977168.1|      CGKCGSGMSIGWSGSTKNRYR----VYRCSANKNKGTCTSKQYKA-----
ref|ZP_02261191.1|    CGKCGSGMSIGWSGSTKNRYR----VYRCSANKNKGTCTSKQYKA-----
RAAC00111             CGFCGSAMYAGWQVSKRGHERQKRPIYRCGRYVTGGGCTPNQVDASALEMLVVEELLRRV
ref|NP_981890.1|      CPSCGAGM-
                        * ** : *

ref|YP_146448.1|      DSINSQMIPHLIQKAAENEESTIEQKIRTAKNRYKRKVEAYTAGLIELEDLNEEKKRMEQ
ref|ZP_02848045.1|    -----
ref|NP_977168.1|      -----
ref|ZP_02261191.1|    -----
RAAC00111             RPELEFPVVSAYRAHAASNELRTVQRKR RVALARLSRLAEAYASGVVSEEA FSREQVRLLR
ref|NP_981890.1|      -----

ref|YP_146448.1|      II-----
ref|ZP_02848045.1|    -----
ref|NP_977168.1|      -----
ref|ZP_02261191.1|    -----
RAAC00111             VLEACREAEAAQAAALSTTDGWD AWVREL CARLAGD PHLER ALFTTVVREV RVFRPKRSRE
ref|NP_981890.1|      -----

ref|YP_146448.1|      -----
ref|ZP_02848045.1|    -----
ref|NP_977168.1|      -----
ref|ZP_02261191.1|    -----
RAAC00111             VDVDLVLRLE
ref|NP_981890.1|      -----

```

FIG. 169

```

ref|ZP_01236658.1| -----RAFLKWAGGKYSLVEDIKRHLPPARKLVEPFVVGAGSVFLNTDYDHYLLADI
ref|ZP_01161642.1| -----RAFLKWAGGKYSLVEDIKRHLPPARKLVEPFVVGAGSVFLNTDYDHYLLADI
ref|ZP_01221581.1| -----RAFLKWAGGKYSLVEEIQRHLPPARKLVEPFVVGAGSVFLNTDFEQYLLADI
ref|YP_128524.1| -----RAFLKWAGGKYSLVEEIQRHLPPARKLVEPFVVGAGSVFLNTDFEQYLLADI
ref|ZP_01898092.1| -----MQKTRAFLKWAGGKYSLVDHLREKLPAGKRLVEPFVVGAGSVFLNTDYDEYLLNDI
RAAC03161 MEVQFMTVLRSYLQWPGGKSHIVTALRAFLPPGRRLIEPFVVGAGSVFLNANYPEYLLGDA
               *::*:*.*** :* :: **..*:*****:.. :*** *

ref|ZP_01236658.1| NPDLINLYNILKHEPERYIEDARKLFTPEYNNKEAYLKIREEFNASDDPYLRSLYFLYMN
ref|ZP_01161642.1| NPDLINLYNILKHEPERYVEDARKLFTPEYNNKEAYLKIREEFNASDDPYLRSLYFLYMN
ref|ZP_01221581.1| NPDLINLYNILKDDPEKYVEDVRKLFTPEFNQKEEYLKIREEFNKTDQDPYVRSLYFLYMN
ref|YP_128524.1| NPDLINLYNILKDDPEKYIEDVRKLFTPEFNQKEEYLKIREEFNKTDQDPYVRSLYFLYMN
ref|ZP_01898092.1| NPDLINMYKILQHKEQFIADAQRFTTPEFNDKERYYKIREKFNKTSNPYQRSIMFLYMN
RAAC03161 NGDLILTHQMLQAHGEAFIEACRDLFVFNNTPDRIYELRDEFNSTADPWRKAILFVYLN
               * *** :::*: . * :: : :*.** * : * :::*:** : :*: :::*:**

ref|ZP_01236658.1| RHGFNGLCRYNKKGGFNVFPFGSYKKPYFPEAEMYYFSEKAKRATFVCEGYLQTFSTRARKG
ref|ZP_01161642.1| RHGFNGLCRYNKKGGFNVFPFGSYKKPYFPEAEMYYFSEKAKRATFVCEGYLQTFSTRARKG
ref|ZP_01221581.1| RHGFNGLCRYNKKGGFNVFPFGSYKKPYFPEKEMYFFAEKAKRATFVCEGYQQTFSRARKG
ref|YP_128524.1| RHGFNGLCRYNKKGGFNVFPFGSYKKPYFPEKEMYFFAEKAKRATFVCEGYQQTFSRARKG
ref|ZP_01898092.1| RHGFNGLCRYNKSGGFNVFPFGSYKKPYFPLKELRFFAEKSKKATFICESYSVDVYKRLGSD
RAAC03161 RHGFHGLCRYTRNGDFNVSFGRYKTVYFPEEEMRYFAEKAKQATFVCADFRDLLERVEPG
               ****:*****.:*.***.** *. *** *: :*:**:*:**:* :. : .* .

ref|ZP_01236658.1| CVVYCDPPYAPLSSTANFTSYSGNGFSLDDQATLADVAEKAAAERDIAVLISNHD TALTR
ref|ZP_01161642.1| CVVYCDPPYAPLSSTANFTSYSGNGFSLDDQATLADVAEKAAAERDIAVLISNHD TALTR
ref|ZP_01221581.1| CVVYCDPPYAPLSSTANFTSYAGNGFSLDDQADLADVAEKAATEREIPVLISNHD TKLTR
ref|YP_128524.1| CVVYCDPPYAPLSSTANFTSYTGNGFSLDDQADLADVAEKAATEREIPVLISNHD TKLTR
ref|ZP_01898092.1| DVVYCDPPYAPLSRTASFTSYASGNGFSLDDQALLAKVSRETSQERNIPILISNHD IPLTR
RAAC03161 DVVYCDPPYVPLSRTSNFTEYAPTRFSWRDHTQLAGYAQALT-ERGVTVVISNHRRAVE
               *****.*** *:.***.*. : ** *:: ** :. : ** ::::***** ..

ref|ZP_01236658.1| RLYTGADLSVVKAKRTIS-RNGSGRKNKVDLALF---
ref|ZP_01161642.1| RLYTGADLSVVKVVRTIS-RNGSGRKNKVDLALF---
ref|ZP_01221581.1| RLYHGADLSVVKVVRTIS-RNGSGRKNKVDLALFKS-
ref|YP_128524.1| RLYHGADLSVVKVVRTIS-RNGSGRKNKVDLALFKS-
ref|ZP_01898092.1| ELYHGSTFEVVQVKRTIS-RNAGKRKNKVDLALYR--
RAAC03161 SLYRGAEIHVVEAPRNIVNRYRQATHVIEELIAVFRAL
               ** *: : **:. *. * : :::*:**::

```

FIG. 170

```

ref|YP_001567845.1|
RAAC03110
gb|ACA00199.1|
ref|NP_490383.1|
ref|YP_319960.1|
ref|ZP_01731959.1|

-----MTSLSQGRRAIGLRYPGHDTLECASQLQALSRAVKQAYGGEASAARMTSRDNIGEAHVTA
-----

ref|YP_001567845.1|
RAAC03110
gb|ACA00199.1|
ref|NP_490383.1|
ref|YP_319960.1|
ref|ZP_01731959.1|

-----RVFVLDKNKQPLMPCHPARARELLKKGKAAVFRYPFTIILKDR-EG
ERRDEEVTLMRQNRVLVLDTHRRPLMPCHPARARQLLKAGRASVFRYPFTIILHDR-NG
-----RVLVLDKNRQPLMPCHPARARQLLRGKAAVFRFPFTIILKDR-ES
-----KVFEIDNEKRSLNPIHPAQARQLLKRGKAAVFRFPFTIIVVKESYTD
-----KVFEINKEKRPLNPIHPAQARQLLRSGKAAVFRFPFTIIVKESNVN
-----VFVLSKDKRPLSPCHQARARKLLKQGKAAVRYKYPFTIILKEEVNP
*:*:.....* * * *:*:***: *:*:***: .*****:

ref|YP_001567845.1|
RAAC03110
gb|ACA00199.1|
ref|NP_490383.1|
ref|YP_319960.1|
ref|ZP_01731959.1|

GDTQPIQVKIDPGSKITGVTLVGDFKN-GKKVIWGAEIHHRGQSIKKALDTRRGVRRSRR
GDVQPMQLKLAPGSKTGVVALVADFQR-GKTVVWAAEIQHRGDQIRQALLRRRLRRGRR
GVTQPTVLKLDPGSNTTGLALVADFARRGKVLVWAAELRHRGAAIRKALADRRARRRFR
ATVSPRLRLKFDPGAKYTGIALVNDVNG---EVVFAELKHRRGFAIRDALTSRRQLRRSRR
TNVSPRLRLKLDPGAKTTGIAIINDAND---EVVFAELKHRRGFAIRDAITSRRQLRRSRR
NKPKECRKIKLDPGAKVTGIALIQ--GD---RVIFGAELEHRRGFRKIDSLESRRQLRRGRR
. * :*:***: *****: :*:***:*** *::: ** ** **

ref|YP_001567845.1|
RAAC03110
gb|ACA00199.1|
ref|NP_490383.1|
ref|YP_319960.1|
ref|ZP_01731959.1|

NRKIRYRIARFDNRKRSKGLPPLISRVENILTWIKRIRRFSPITGISLELVRFDTOQL
YRKTRYRKPRFDNRRLLEGWLPPLMSRVHNIETWVARLRRWAPITHLSMQLVSFDTOQL
-FKLRYRPARFNNRTRPSGWLAPSLQHRVDTTMTWVQRIMRWTPVTSLSQELVRFDTOAM
SRKTRYRQPRFFNRTRPQGWLAPSLQSRVENIKTWVNRRLKRVAPITAISSQELVKFDTOQL
NRKTRYRQPRFLNRTPKEGWLAPSLQSRVENIKTWVSKLRKRVAPITAISSQELVKFDMQLI
GRKTRYRKPRFLNRTRPKGWLPPSLQSRVENINTWVNRRLKRLCPVDSISVELVRFDTOQL
* *** .** ** : .***.*** **.. **: :: : * : * :** ** * :

ref|YP_001567845.1|
RAAC03110
gb|ACA00199.1|
ref|NP_490383.1|
ref|YP_319960.1|
ref|ZP_01731959.1|

QDPEINGIEYQRTLYGYEIKYLLKWKGRKCVYCGKENVPLEIEHIVPKSKGGS DRVSN
QYPEISGEEYQQTLYGYEVREYLLKWKGRKCAVCGAENVPLEIDHVI PRSRGGS DRVSN
QNPEISGVEYQQTLYGYEVREYLLKWKVRKCAVCGAENVPLTIDHHPKSKGGS DRVSN
RHPDIQGKEYQQTLAGFETRQYFLEKWNRECAVCGIKDVPLQIEHHPKSKGGSNSITN
RRPDIQGKEYQQTLAGFETRQYFLEKWNRECAVCGIKDVPLQIEHHPKSKGGSNSITN
ENPEISGVEYQQTLYGYEVREYLLKYNRTCYCGAKEVPFEIEHIVPKSKGGS DRVSN
. *:*.* ***:*** ** * :*:***: * * ** . :***: *:*: *:*:***: :*:

ref|YP_001567845.1|
RAAC03110
gb|ACA00199.1|
ref|NP_490383.1|
ref|YP_319960.1|
ref|ZP_01731959.1|

LTLACHECNQKKGNQSIIEFLTNNPERLKQIKSESKRPLKDTAALNATRWYIFNQLRGNS
LVLACHRCNQAKANRPEVEFLAHDQERLHRIQAQLKLSLRDAAAMNATRWAVFRRLKDT-
LTLACFCPCNQKGNRDVAEFLDRDPKRLARIEAQRRAPLQDAAAVNSIRWALWRNLVAT-
LTLSCCKCNLKKGTQDIKNFLKKDKSRLEKILALTKKSLADAAAVNTTRFALLEVLKST-
LTLSCCKCNLKKGNQDIRDYLEKDKPRLEKILALTKKPLADAAAVNATRFALLYALKST-
LTLSCCKCNQKKGNRSVEEFLKGKPEVLKRVKANAKKPLRDAAAVNAARWAVFNRLKAS-
*:*.* ** *::: :*: . * :: : : . * *:*:***: *:*: * .

ref|YP_001567845.1|
RAAC03110
gb|ACA00199.1|
ref|NP_490383.1|
ref|YP_319960.1|
ref|ZP_01731959.1|

LTAGKEELPIEVGTGGRTKYNRETQNYPKKHWDAAACVGESGQNVQIEPDMQVLEIKAMG
-----GLPVEAGSGGRTKYNRFSQGYPAKWIDAAACVGESGQCVRLDTQMQLTIVAKG
-----GLDVEVGTGGRTKYNRSQQYKPEHWIDAAACVGESGANVRLDPFHKPLAIQAMG
-----GLPVEIGSGGLTKFNRTQQGLNKS HWLDACCVGASTP-ILKIKGIRLLITANG
-----GLPIETGSGGLTKFNRTQQALNKS HWLDACCVGASTP-ILKIKGIRLLITANG
-----GLEVEAGSGGLTKFNRTKHELPHKTHWLDAAACVGKSTSPSKLRLVQKPLKIKATG
* :* *:* ** ** * : * :*:***.*** * : * * * *

ref|YP_001567845.1|
RAAC03110
gb|ACA00199.1|
ref|NP_490383.1|
ref|YP_319960.1|
ref|ZP_01731959.1|

HGMRRMCFVDKYGFPPKYPKERTYMGYKTGDIVLAVIPKGKNMGIHIGRIAIRHRPSFL
HGKRQRCITDKYGFPRSHAPSSRSYMGFRTGDLVQAHVPRGKYAGTHVGRIVIRHRPSFQ
HGERQARLNRNRYGFPRGYKLARSYLGFRGTGDVVFCAKPK---IEVFAGRVAIRFRPSFV
HGTRQSCRTDKYGFPIRHCSRTEKFHFQGTGDIVKAVVTGKKGKIGTYIGRIATRATGSFN
HGTRQSCRTDKYGLPVRHCSRTEKFHFQGTGDIVKAVVTGKKGKIGTYIGRIATRATGSFN
HGTRQMCRTNKYGFPTRYVPHKKDVKGFTGDIVKAVVTGKKGKIGTYIGRVAVRSTGSFN
** * : . :*:***: * : : * :*:***: * . : . : . ***. * **

ref|YP_001567845.1|
RAAC03110
gb|ACA00199.1|
ref|NP_490383.1|
ref|YP_319960.1|
ref|ZP_01731959.1|

LNGVG---DVHPKYLTLQKNDGYGY---
LNGF---NVHPKYLKVLQRGDGYGYAME
VTGAD-RKVESHPKYLSMIQKGDGYAYA--
ISTVQGLIQGISHKYCKTIHKKDGYAYAI-
ISTAKGLIQGISHKYCKTIHKKDGYAYGI-
ISTAKGVIQGINHKYCRQIHKKDGYSYA--
.. ** :*:***.*

```

FIG. 171

```
ref|NP_852746.1|-----IRTLAINNEPWFIAKDVCD AIGIDNNRKALLALDEDEK-GVTLSNTL
ref|YP_853610.1|-----FESQADIRVIVINGEPWFIASDVCRAIGIANHRDAVRKLDDEK-GVASTDTP
emb|CAP00374.1|-----FEQNSQIRIIMINSNPWFVAKDICDALGLSNHRDAISKLDKDEK-GVALTDTL
ref|ZP_01959153.1|-----NSTPIRVQVINNEPWFVAKDVCDVLGISKYRDAIARLDDDEG-CPIEVDTL
ref|YP_001038857.1|-----GKQVRTFIIDGEPWWVAKDVCDILELGDTHKAMERLDEDERNTIPVTDLSL
RAAC03810MQQTWAVFANGTQVRVFWVDGEPLFDAMGVCEAVGLRNVEKALRRLDDDEQ-GSVILEGL
          :*   :::* : * .:* : : . .*: **.**   :

ref|NP_852746.1|GGKQEMNIISESGLYTLILRCDAVKKGSVPHRFRKWVTAEVLPAIRKTGKYEAKTTVDD
ref|YP_853610.1|GGEQESIIIESGLYTLILRCDAVTPGTIPYRFRKWVTGEVLPQIRRTGSYIKNSLPQE
emb|CAP00374.1|GGQQELSVINESGMYALVMRSRDAMKEGTPQHKFRKWVTSEVLPSIRKTGKYEAP-----
ref|ZP_01959153.1|GGMQKMAAVNESGLYTLILQSR---KPEAKP--FRKWVTSEVLPSIRKKGYG-----
ref|YP_001038857.1|GRLQETYVVNEAGLYNLILGSR---KQEAKE--FKRWITHEVIPQIRKTGIYALEPKQL-
RAAC03810DGREEIHVVRESGMLRLSLVGK-----DEHARALQRWATREVLPSVIRTGRYGE PDIEQE
.   ::   :*:* : * : :   ::* * **:* : :.* *

ref|NP_852746.1|RTGLRNA-----
ref|YP_853610.1|E-----
emb|CAP00374.1|-----
ref|ZP_01959153.1|-----
ref|YP_001038857.1|-----
RAAC03810QLQLQKATLLFNTLAMFRDRLFDETQNLANAMADVLRG
```

FIG. 172

emb|CAP00374.1|
 ref|NP_852746.1|
 ref|ZP_01959153.1|
 ref|YP_001038857.1|
 ref|ZP_02327844.1|
 RAAC03316

-----FEQNSQIRIIMINSNPWFVAKDICDALGLSNHRDAISKL
 -----IRTLAINNEPWFIAKDVCDAGIDNNRKALLAL
 -----NSTPIRVQVINNEPWFVAKDVCVLGISKYRDAIARL
 -----GKQVRTFIIDGEPWWVAKDVCDILELGDTHKAMERL
 -----GKETNHMNQLQVFNFTGKDVRVVVKDGHPPWWVAKDVSELLGFRMASDFTRTL
 MSCTTHSIGRRRSPMQQTWAVFANGTQVRVFWVDGEPLFDAMGVCEAVGLRNVEKALRRL
 : * : . . * : * . . . : : : . *

emb|CAP00374.1|
 ref|NP_852746.1|
 ref|ZP_01959153.1|
 ref|YP_001038857.1|
 ref|ZP_02327844.1|
 RAAC03316

DKDEKGVALTDLGGQQLSVINESGMYALVMRSRDAMKEGTPQHKFRKWTSEVLPSI
 DEDEKGVTLNLTGGQQLSVINESGMYTLILRCDAVKKGSVPHRFRKWTAEVLPAI
 DDDEGCPIEVDLTGGMQMAAVNESGLYTLILQSRKPEAKP-----FRKWTSEVLPSI
 DEDERTIPVTDLSLRLQETVYVNEAGLYNLILGSRKQEAKE-----FKRWITHEVIPQI
 DDDEKDTQIVRTPGGNQEVTTINESGLYSAILKSRKPEAKQ-----FKRWVTHEVLPAI
 DDDEQGSVILEGLDGREEIHVVRESGMLRLSLVGKDEHARA-----LQRWATREVLPSV
 * . ** . : : * : * : : . . : : * * * : *

emb|CAP00374.1|
 ref|NP_852746.1|
 ref|ZP_01959153.1|
 ref|YP_001038857.1|
 ref|ZP_02327844.1|
 RAAC03316

RKTGKYEAP-----
 RKTGKYEAKTTVDDRTGLRNA-----
 RKKGYG-----
 RKTGIYALE-----PKQL-----
 RKTGMYATDELLDDPELLIQAVTKL-----
 IRTGRYGEPIEQELQLQKATLLFNTLAMFRDRLFDETVQNLANAMADVLRG
 : . * *

FIG. 173

| | |
|--------------------|---|
| ref YP_148854.1 | -----DRNDVMRVVEVEPLLALYLYTPLCGTCQLARRM |
| ref YP_001127043.1 | -----PIERTEIQQTITEQPLMALYLYTPLCGTCQLAKRM |
| ref YP_080555.1 | -----LKRLEHETCL-LYVYTPFCGTCQLAGRM |
| ref NP_693306.1 | -----VI--YIYTPFCGTCVSARAM |
| RAAC03018 | MVRANVQGPRQVHTEGRRIPMFTPFPGDRPEALGTLPEALV--FVHTPLCGTCQLARRM |
| ref ZP_02170973.1 | -----KGVL--FLHTPLCGTCCKMAGQF |
| | : : : * : * : * |
| ref YP_148854.1 | LTVVEQLF-PALPFYETDINYIPEQAVAWKIESVPCLLLFDRDGTVAGKWFHSPVPLYE |
| ref YP_001127043.1 | LTVVEELF-PALPFYETDVNYIPEQAVVWKIESVPCLLLFVDGTMVGKWFHSPVPLYE |
| ref YP_080555.1 | LEVVDMM-KDIPFYKNNLNYSFSAKQOEIESVPCFIFYKKGRVVKKEYAFHSVSYLYD |
| ref NP_693306.1 | LENIERLHRQDI-FYEMNAAVHPEYMDNQIESVPCLLFVNGEIVERIYTFYSTANIYD |
| RAAC03018 | LEVVDAAHTNRLPLYDLNLAPEAMQSWKIESVPALLYVKNGEIAHVQYRFGDVVDLAD |
| ref ZP_02170973.1 | LQIAEQIP-DIPEIFSMDLNMAPHIAKDWKIESVPCLVVIDDGNVTDRLMYAFKSIPDI-- |
| | * : : : * : * : * : * |
| ref YP_148854.1 | VIQACL-- |
| ref YP_001127043.1 | EIQARLPR |
| ref YP_080555.1 | TIKQNL-- |
| ref NP_693306.1 | YLIQY--- |
| RAAC03018 | AIRQFLER |
| ref ZP_02170973.1 | ----- |

FIG. 174

```
ref|YP_001178926.1|      -----FPAVFEPGENGGYCVTFPDLPGCITEGETLEEALYMAKEALELHLYGLEEDNEEI
ref|YP_001211485.1|      -----FPAIFEVGEIKGYCVTFPDLPGCITEGNTLEEALQMAAREALELHLYGMEEDEDEI
ref|YP_001180877.1|      -----FPAIFEPAEKEGYVVTFPDLPGCITEGKNLEEALYMAKEALELHIFGLEEDGEEI
ref|YP_001219837.1|      -----YPALFEAYEDGGYTVSFPDLPGCITEGDSLTEALYMAKEALELFLWNMEDDNEEI
ref|YP_001211518.1|      -----YPAIFDYAED-GISVEFPDLPGCYTCGDNDDEALYMAKEALALHLYGLEEDGFPI
RAAC02718                MATRGYPAlFDPYEDGSFVVLFPDFPGCVTQGCNLDDAVAMAEALTLQLYGIERDGEDL
                           :*: *:  *  .  *  *:*:*: *  *  .  *: *: *:*: *  *: *: *  *  :
                           :

ref|YP_001178926.1|      PEPTPPEKLDIP-----
ref|YP_001211485.1|      PSPTPPEERVDVPEKGF----VAIVEARMPL---
ref|YP_001180877.1|      PEPTPPEKLN-----
ref|YP_001219837.1|      PEPTPPEKI-----
ref|YP_001211518.1|      PEPSPASRIKVEPNQV-----
RAAC02718                PEPSDPSHVTVPPEVQGWVMLIHPRPELFHP
                           *. *:  .  *: :
```

FIG. 175

```

ref|YP_001035109.1|-----MAQLYYKYGTMTNSGKTIEILKVA
ref|YP_001450442.1|-----MAQLYYKYGTMTNSGKTIEILKVA
ref|ZP_01818216.1|-----MAQLYYRYGTMTNSGKTIEILKVA
ref|ZP_01830009.1|-----MAQLYYRYGTMTNSGKTIEILKVA
ref|ZP_01819917.1|-----MAQLYYRYGTMTNSGKTIEILKVA
RAAC01115MERLARDDGHSARRLVRARGRVTCITRSREPERGIDVAKLYFRFGQMNASKSIQLLTVA
:*****:* **:.*****:*.**

ref|YP_001035109.1|HNYEEQGKGSVVIMTSAVDTRDGVGVSSRIGMKRNAIAIEDQTDIFGYIQQLPEKPYCVL
ref|YP_001450442.1|HNYEEQGKGVVIMTSAVDTRDGVGVSSRIGMKRQAMAIEDDDILGYIKNLPEKPYCIL
ref|ZP_01818216.1|YNYEEQGKGVVIMTSALDTRDGVGVSSRIGMKRPALAEETTDIFGYIRDLPEKPCCVL
ref|ZP_01830009.1|YNYEEQGKGVVIMTSALDTRDGVGVSSRIGMKRPALAEETTDIFGYIRDLPEKPYCVL
ref|ZP_01819917.1|YNYEEQGKGVVIMTSALDTRDGVGVSSRIGMKRPALAEETTDIFGYIRGLPEKPYCVL
RAAC01115HSYEEQGKKVALFTPAIDDRYGRGVIASRVGIAKRAVVVEEDTDLFACVR--AQMPDCVL
:***** *.:*.*.* * * * :*****: * :.:*: **:. :. : * *.*

ref|YP_001035109.1|IDEAQFLKRHHVYDLARVVDELDPVMAFGLKNDFRNELFEGSKHLLLLADKIEEIKTIC
ref|YP_001450442.1|IDEAQFLKRHHVYDLARVVDELDPVMAFGLKNDFRNELFEGSKHLLLLADKIEEIKTIC
ref|ZP_01818216.1|VDEAQFLKRHHVYDLARVVDELDPVMAFGLKNDFRNELFEGSKYLLLLADKIDEIKTIC
ref|ZP_01830009.1|VDEAQFLKRHHVYDLARVVDELDPVMAFGLKNDFRNELFEGSKYLLLLADKIDEIKTIC
ref|ZP_01819917.1|VDEAQFLKRHHVYDLARVVDELDPVMAFGLKNDFRNELFEGSKYLLLLADKIDEIKTIC
RAAC01115VDEVQFLRAYHVRQLARVADDLDIPVIMYGLLKDYRNRLFEGSEAAILWADRIEEIKTIC
:*.***: **.*****.*:*.***: **.***.*****: * **.*:*****

ref|YP_001035109.1|QY--CSRKATMVLRTDNGKPVYDGEQIKIGGHETYISVCRKHYFNPDIK-----
ref|YP_001450442.1|QY--CSRKATMVLRTDHGKPVYDGEQIQIGGNETYIPVCRKHYFKPDI-----
ref|ZP_01818216.1|QY--CKKKATMVLRTQDGLPVYDGEQIQIGGNETYISVCRKHYFAPEI-----
ref|ZP_01830009.1|QY--CKKKATMVLRTQDGLPVYDGEQIQIGGNETYISVCRKHYFAPEI-----
ref|ZP_01819917.1|QY--CKKKATMVLRTQDGLPVYDGEQIQIGGNETYISVCRKHYFAPEI-----
RAAC01115AHPGCDRKATMILKVKGDRPVYEGEQIEVGGNDLYKSVCRKHYFHPDVEGLMRATGP
: *.:*****:*.***.*****:*****: * .***** *:

```


FIG. 176

```

ref|NP_622644.1|
ref|YP_001664955.1|
ref|ZP_02616274.1|
ref|YP_001255422.1|
ref|YP_076287.1|
RAAC01119

```

TEINCLSTRIDIUMBTULINUMASTRATCCGBABSGLYCERPHSPHRYLDIESTERPHSP

```

ref|NP_622644.1|
ref|YP_001664955.1|
ref|ZP_02616274.1|
ref|YP_001255422.1|
ref|YP_076287.1|
RAAC01119

```

HDIESTERASEFAMILYPRTEINCLSTRIDIUMBTULINUMASTRHALLSCRESIGNIFI

```

ref|NP_622644.1|
ref|YP_001664955.1|
ref|ZP_02616274.1|
ref|YP_001255422.1|
ref|YP_076287.1|
RAAC01119

```

TLVIAHRGDSKNVPENTIAAFKRAMELGADG
LVIAHRGDSRNAPENTLASFKSALEMGADG
TKILAHRGASAYAPENTMAAFKKAIEMNADG
CANCEE-IDENTITIESPSITIVESGAPSTKILAHRGASAYAPENTMAAFKKAIEMNADG
IGHRGAAGTHPENTMVSFRRAFELGVDG
MIRELLMRDDTLFIAHRGASAECPENTLPAFVRAAELGADM
: . * * * : * * * * : * * * . . *

```

ref|NP_622644.1|
ref|YP_001664955.1|
ref|ZP_02616274.1|
ref|YP_001255422.1|
ref|YP_076287.1|
RAAC01119

```

IELDVQLTKDGHLLVVIHDETVDRTTNGEGFVKDFTLEEIKKLDAGIKFGEKFAGERIPTL
IELDVQLTKDQGLVVIHDERVDRTTDGIGFVKDFTLKEIKRLDAGIKFDKKFAGERIPSL
IELDVHLSKDGIVYIIHDERVDRTTDGKGEVKDFSLYELKKLDAGSWFSDEYKDEKIPTL
IELDVHLSKDGIVYIIHDERVDRTTDGKGEVKDFSLDELKKLDAGSWFSYEEKDEKIPTL
IEFDVHRTADGHLVVIHDPVDRTTDGSGGLVMAMTLEQIRALDAGSWKDFRFAGERVPTL
IELDVRLTGDDGVVVLHDPVDRRTNGSGLAADMRLADLRKLDAGSWFDARFRGTTLIPTL
* . * * : * * * : * * * * * * * * : * * * : * * * : * * *

```

ref|NP_622644.1|
ref|YP_001664955.1|
ref|ZP_02616274.1|
ref|YP_001255422.1|
ref|YP_076287.1|
RAAC01119

```

YEVFELIGDKDFLVNIEIKSGIVLYP-GIEEKLIKAIKEYNFEERVIISSFNHYSRLDVK
YEVFELIGHKNFIVNIEIKSGIVLYP-GIEEKLIKAIEDYDFEDRVVISSFNHYSRLDVK
EELLKLIKNTELYLNIEIKAGYRVYP-NIEEKVIAMIEKYKMLDRVVISSFDHYSRLRVK
EELLKLIKNTELYLNIEIKAGYRVYP-NIEEKVIAMIEKYKMLDRVVISSFDHYSRLRVK
QELVRATP-PGVRLYLELKAGSVHYP-GIEQELVDLLKAEGVLDRTQVSSFDHRLVKLK
DEVFATVP--HMCLNIELKTSPVAHTRQLIRRVLGAIYRHNARDRVLISSFDHAALAEVR
* . . : * * * . : . : . : : * . : * * * * * * : *

```

ref|NP_622644.1|
ref|YP_001664955.1|
ref|ZP_02616274.1|
ref|YP_001255422.1|
ref|YP_076287.1|
RAAC01119

```

KMAPHLKIGLLYQCGLVEPWHMALRMEAYSLHPFYFNIIPELVEGCKKNGVKLFPTVDR
RMAPELKIGLLYQCGLVEPWHMAIRMKAYSLHPFYFNIIPELVEGCKKNNIKLFPPTVDR
EINSNIKTGMLYEALYEPWEYARSIKVEALHPNYITLTKEFIDKAYINSFEVNSYTIND
EINSNIKTGMLYEALYEPWEYARSIKVEALHPNYITLTKEFIDKAYINSFEVNPYTIND
EICPELPLGMLTSSNLVDPVGMAKAIACEAIHPAWPWTPEYVEAAHAAGLVNCWTVND
RFDRDIAIGVLFTRGLLEPIQLAERLGAQSLHPDIEHLDPLFVAEAKSHHLLAVYAWTAKS
. : . : * * * * * : * * * : : . : : * .

```

ref|NP_622644.1|
ref|YP_001664955.1|
ref|ZP_02616274.1|
ref|YP_001255422.1|
ref|YP_076287.1|
RAAC01119

```

KEDMERMIKAGVDGIITDD-----
KEDMEMMIKAGVDGIITDD-----
ETNMKSLIKNKVTSII-----
ETNMESLIENKITSII-----
PLALALMQNAGVDGIMTD-----
ELAVNRALACQVSGIILDDLRLRQIATKAVPSPTPPSA
: : * :

FIG. 177

```

ref|YP_001433837.1|-----VFVTTS-RGLRRDSVPMRLF EKAKRLGVWNPSDIDFSQD
ref|YP_001276310.1|-----TTS-RGLRRDTVPMRLF EKAKRLGVWNPSDIDFSQD
ref|ZP_01514627.1|-----TTS-RGLRRNSPPMRLF EKAKQFGIWNPSLLDFSQD
ref|YP_001636906.1|-----TTS-RGLNRNSPPMRLF EKAKRFGIWNPSLIDLSRD
RAAC00203MVGVEFDVLEATPMALTEMPRVFQTTSEKRLNHSILPMRLYHKAKKLTWDPRDIDFTKD
ref|YP_001102862.1|-----LNWDSLPLRLFGKGN-AKFWDPADIDFTRD
* . . *::: *:: *:: *:::

ref|YP_001433837.1|MRDWQTLTDDERDLVLRLTSLFQAGEEAVTLDLLPLVATIAAEGRIEEELFLTSFLWEEA
ref|YP_001276310.1|VRDWQTLTADERDLVLRLTSLFQAGEEAVTLDLLPLVATIAAEGRIEEELFLTSFLWEEA
ref|ZP_01514627.1|IDDNRLSDEERDLLRLTALFQAGEEAVTLDLLPLIKAIADHGRLEEELFLTTFLFEEA
ref|YP_001636906.1|IRDWEQLRDEERDLLRLTALFQAGEEAVTLDLLPLIRTIAAREGRLEEELYLTTFLFEEA
RAAC00203KEDWQGMNEAERRVILRLCSLFVAGEEAVTLDLLPLIMAVAREGRLEEELYLTTFLFEEA
ref|YP_001102862.1|AEDWQGLTEEEERRSVAMLC SQFIAGEEAVTQDLQPFMAAAEGRFGDEMYLTQFCFEEA
** : ** : * : * ***** ** *:: ::: *::: *::: *::: *:::

ref|YP_001433837.1|KHTDFFNRFLTEVAGVTGDL SHYHTAN--YRALIHEALPAAMERLRDDPSPEAQAIASLT
ref|YP_001276310.1|KHTDFFNRFLTEVTGVTGDL SHYHTSN--YRALIHEALPAAMGRLRDDPSPEAQAIASLT
ref|ZP_01514627.1|KHTDFFARFISEVARVDPDL SRYHTPS--YRALIYDALPRAMQRLHDPSPFNLAESLT
ref|YP_001636906.1|KHTDFFARFISEVARVDPDL SRYHTAS--YRALIYEALPAALHRLDHDPSPLNLAESLT
RAAC00203KHTEVFRRLDEVAGETSDLSVFHKDH--YRKIFYEYLPQAMGRLVHDPSPPEAQAEASVT
ref|YP_001102862.1|KHTQVFRWLMDAVG-LTGDLHSHVAENPGYRAIFYEELPRSLNALHDDPSPANQVRASVT
***:.* :: * ** . ** :::: ** :: * .**** .***:

ref|YP_001433837.1|YNMIVEGVLAETGYHAYFTMLERNLMPGTRKGIGLLKQDESRHIAYGVYLLSRLINAEF
ref|YP_001276310.1|YNMIVEGVLAETGYHAYFTMLERRDLMPGTRKGIGLLKQDESRHIAYGVYLLSRLINAEF
ref|ZP_01514627.1|YNMIVEGVLAETGYHGYFTILDTHNLMPGVREGIRLLKQDESRHIAYGIYLLSRLIAADR
ref|YP_001636906.1|YNMIVEGVLAETGYHAYFTILDSQDIMPGTREGIRLLKQDESRHIAYGVYLLSRLIATDP
RAAC00203YNMIVEGVLAETGYAFYITALQENKLPGLIQAIRLLQRDESRHIGYTYLLSRLISENA
ref|YP_001102862.1|YNHVVEGTALTGYFAWQKICRSRGILPGMQEVVRRIGDERRHMAWGTFTCRRHVADE
** :***.** ***.: . .. :** : : : ** **.:* : * : :

ref|YP_001433837.1|RLWDVVVERMNELVVYAMGVIDD AFACYEV-----VPFGLSHDDFMQFAMG
ref|YP_001276310.1|RLWDVIVERMNELVVYAMGVIDD AFACYEV-----VPFGLSHDDFMHFAMG
ref|ZP_01514627.1|QIWDHIVERMNELVIHALGVIDE IFTSYDT-----MPFGLQIETFTSTFALN
ref|YP_001636906.1|QVWDHIVERMNELVLYAMGVIDE IFCYDP-----MPFNLQIETFTSTFALN
RAAC00203HIWDVVNRINLLLPHAIGVVQDLN AELDTEEAQKLAEELGITDVPFGLNPDEFVAYAQK
ref|YP_001102862.1|SNWDVVQE QMQHLLPLAVTQIQWRPEDAPEE-----TPFRLDIDELAAAYASD
** : .::: * : * : : ** * . : : :

ref|YP_001433837.1|QFERRLDRIERARNQTVE-----
ref|YP_001276310.1|QFQRRLDRIERARDQTLESI-----
ref|ZP_01514627.1|QFQRRNLNRLELACQQSIAEI-----
ref|YP_001636906.1|QFQRRNLNRLELASRQSMNEI-----
RAAC00203QFATRLGVLRARGKSVDEIYYYTEETVGV EQ
ref|YP_001102862.1|RAGRRLGAISAARGVPVEQI-----
: ** . : . * : :

```

FIG. 178

```

ref|ZP_02849297.1|      -----ERTLVISDIHGICYDAFNRLLERVNYEPVKDQLILLGDFVSKGPDSKLVVEQSAS
RAAC01413              MGSTPPERTIFISDIHGHLRPFARLLERLGYRPDRDQLVLVGDYISGGPDSLGLVRLVRN
ref|NP_815033.1|        -----ISDIHGQADLFDALLT--DYDPVEHQLVLIIGDLNDRGSDSKACFLKGKE
ref|ZP_00604769.1|      -----IGDVHGKHELFEKLIH--YYDPKIHQLVLLGDLNDRGPRTKECFLLGKQ
ref|YP_796463.1|        -----VSDIHGEYKKFKEILK--YWDSNRQQQLILLGDLCDRGLQSYECFYLAKE
ref|NP_862606.1|        -----VGDIHGGEYKKFKEILK--YWDSNRQQQLILLGDLCDRGLQSYECFYLAKE
                        .*:.*      *   :   :   .   .*:.*:*      . *   :   .

ref|ZP_02849297.1|      LVREHGAIHAVQGNHDERFVDIVRNKTKEAREKFFRHGGRQTVYSYLKGSMDTGSEEDLLN
RAAC01413              LCSEG-AVALRGNHEEAUVNWMR-----RGMKPLGPVRDSLY
ref|NP_815033.1|        LVEQHGAIVYLRGNHEEYFLQFLQNPEDWFAGYVRNGGKETIESLLHPGATAEYSPTMAL
ref|ZP_00604769.1|      LVEETGAIYLRGNHEEYFLQFLQSPEDWYAPYMRNGGKETTESLIHPGASEEYSPTISM
ref|YP_796463.1|        LCDNYGAILIKGNHEDLFLKFLNKTEDFKENYIKNGGLKTLESF---GYSENNTFKDIVL
ref|NP_862606.1|        LCDNYGAILIKGNHEDLFLKFLNKTEDFKENYIKNGGLKTLESF---GYSENNTFKDIVL
                        *   :   *: :*:*: :   . . .      *       .   :

ref|ZP_02849297.1|      RLRETVLERYPYHADFLQPLPYYYEDKHFIYVHAGLNPYPNWKQPKRDFLYIKDSFHK
RAAC01413              R----AIAADASLAAFLSTLPYAWEGERWVAVHAGIDPKPHWRQTAKRDLLTIRERFYA
ref|NP_815033.1|        MIR----SRYPELIDFLTKRPLYFEWQHLYLFVHAGVDLTMEDWRQTAPKDFLWIREPFHQ
ref|ZP_00604769.1|      MIQ----SRYSELIKFLVQRPLYEYEWENYLFVHAGVDLTKKDWRKTSLRDFIWIREFPHE
ref|YP_796463.1|        DIK----KNNDKLIEFLTYLPNFYEWNDYIFVHAGVNLKINNWKDTSIRDFMWIREDFHF
ref|NP_862606.1|        DIK----KNNDKLIEFLTYLPNFYEWNDYIFVHAGVNLKINNWKDTSIRDFMWIREDFHF
                        **      *   :*   : :   *:*: :   .*: . . :*:*: *: * :

ref|ZP_02849297.1|      SDTVVNKVVVFGHTKTVDLH---GRADVWYGKGKIGIDGGCASGLQLNALEIIGEEVRSF
RAAC01413              RPHRVGKLVVFGHTPCVVLH---GTHDVWYGSDKVGIDGGARHGGQVNALVDAGGALTST
ref|NP_815033.1|        GKNNTGKTIVFGHTITPMLHGDMQTTDLWQSDGKIGIDGGAIFGGSVHGV-----
ref|ZP_00604769.1|      GKNNTGKTIVFGHTITPMLHGDMQTTDLWIQDHKIGMDGGAIVFGGSLHGV-----
ref|YP_796463.1|        TPNRLNKTIVFGHTETKILN-KNNKYDIWIHDNKIGIDGGAIVGGYLYGVI-----
ref|NP_862606.1|        TPNRLNKTIVFGHTETKILN-KNNKYDIWIHDNKIGIDGGAIVGGYLYGVI-----
                        .* :*****      *:      *: *   . *:*:***. *   :   :

ref|ZP_02849297.1|      SVP-
RAAC01413              AEPV
ref|NP_815033.1|        ----
ref|ZP_00604769.1|      ----
ref|YP_796463.1|        ----
ref|NP_862606.1|        ----

```

FIG. 179

| | |
|----------------------|--|
| ref YP_145897.2 | ----- |
| sp P42816 KPRS_BACCL | ----- |
| ref YP_001124176.1 | ----- |
| ref NP_387932.1 | SISFALLSTERICINHIBITINANDACTIVATINPDBDKUACHAINACRYSTALSTRUCT |
| RAAC01435 | ----- |
| ref YP_359065.1 | ----- |
| | |
| ref YP_145897.2 | ----- |
| sp P42816 KPRS_BACCL | ----- |
| ref YP_001124176.1 | ----- |
| ref NP_387932.1 | URESFBACILLUSSUBTILISPHSPHRIBSYLPYRPHSPHATESYNTHETASEMLECULA |
| RAAC01435 | ----- |
| ref YP_359065.1 | ----- |
| | |
| ref YP_145897.2 | ----- |
| sp P42816 KPRS_BACCL | ----- |
| ref YP_001124176.1 | ----- |
| ref NP_387932.1 | RBASISFALLSTERICINHIBITINANDACTIVATINPDBDKUBCHAINBCRYSTALSTR |
| RAAC01435 | ----- |
| ref YP_359065.1 | ----- |
| | |
| ref YP_145897.2 | ----- |
| sp P42816 KPRS_BACCL | ----- |
| ref YP_001124176.1 | ----- |
| ref NP_387932.1 | UCTURESFBACILLUSSUBTILISPHSPHRIBSYLPYRPHSPHATESYNTHETASEMLEC |
| RAAC01435 | ----- |
| ref YP_359065.1 | ----- |
| | |
| ref YP_145897.2 | ----- |
| sp P42816 KPRS_BACCL | ----- |
| ref YP_001124176.1 | ----- |
| ref NP_387932.1 | ULARBASISFALLSTERICINHIBITINANDACTIVATINPDBIBSACHAINAPHSPHRI |
| RAAC01435 | ----- |
| ref YP_359065.1 | ----- |
| | |
| ref YP_145897.2 | ----- |
| sp P42816 KPRS_BACCL | ----- |
| ref YP_001124176.1 | ----- |
| ref NP_387932.1 | BSYLDIPHSPHATESYNTHETASEINCMPLXWITHCADMIUMINSPDBIBSBCHAINBP |
| RAAC01435 | ----- |
| ref YP_359065.1 | ----- |
| | |
| ref YP_145897.2 | ----- |
| sp P42816 KPRS_BACCL | ----- |
| ref YP_001124176.1 | ----- |
| ref NP_387932.1 | HSPHRIBSYLDIPHSPHATESYNTHETASEINCMPLXWITHCADMIUMINSEMBCAAUN |
| RAAC01435 | ----- |
| ref YP_359065.1 | ----- |
| | |
| ref YP_145897.2 | ----- |
| sp P42816 KPRS_BACCL | ----- |
| ref YP_001124176.1 | ----- |
| ref NP_387932.1 | NAMEDPRTEINPRDUCTBACILLUSSUBTILISDBBAAPHSPHRIBSYLPYRPHSPHATE |
| RAAC01435 | ----- |
| ref YP_359065.1 | ----- |

SYNTHETASEBACILLUSSUBTILISEMBCABPHSPHTRIBSYLPYRPHSPHATESYNTHETASE

```
ref|YP_145897.2|
sp|P42816|KPRS_BACCL
ref|YP_001124176.1|
ref|NP_387932.1|
RAAC01435
ref|YP_359065.1|
```

TASEBACILLUSSUBTILISSUBSPSUBTILISSTRSCRESIGNIFICANCEE-IDENTI

```
ref|YP_145897.2|
sp|P42816|KPRS_BACCL
ref|YP_001124176.1|
ref|NP_387932.1|
RAAC01435
ref|YP_359065.1|
```

```

-----LKLFLALNSNMKLAKETAEVMGIELGKCSVSRFSDGEIQINI
-----LKLFLALNSNMKLAKETAEVMGIELGKCSVSRFSDGEIQINI
-----LKLFLALNSNMKLAKQIAETMGIELGKCSVSRFSDGEIQINI
TIESPSITIVESGAPSDKNLKIFSLNSNPELAKEIADIVGVQLGKCSVTRFSDGEVQINI
-----MAIDADLKLVTGNANPALAQEIADYIGVPLADCQVGRFSDGEVRIKL
-----DLKIFSGNANPDLAREIASFLGVEVGDAVRFSRSDGEIQVKI
*:      *:      *:      *:      *:      *:      *:      *:      *:

```

```
ref|YP_145897.2|
sp|P42816|KPRS_BACCL
ref|YP_001124176.1|
ref|NP_387932.1|
RAAC01435
ref|YP_359065.1|
```

```

EESIRGDDVFVIQSTSVFVNEHLMELLIMIDALKRASARTINIVMPYYGYARQDRKARSF
EESIRGDDVFVIQSTSVFVNEHLMELLIMIDALKRASARTINIVMPYYGYARQDRKARSF
EESIRGDDVFVIQSTSVFVNEHLMELLIMIDALKRASARTINIVMPYYGYARQDRKARSF
EESIRGDCDYIIQSTSDPVNEHIMELLIMVDALKRASAKTINIVIPYYGYARQDRKARSF
GESVRGANVFIIQPTSQPVNEHLMELLILMDAVKRASARTINIVIPYYGYARQDRKARAR
NESVRGADVFIIQPTCTPVNENLMELLILIDAIRRASARITAVLPYFGYARQDRKTRAR
**** :*: * :*: *: *: *: *: *: *: *: *: *: *: *: *: *: *: *:

```

ref|YP_145897.2|
sp|P42816|KPRS_BACCL
ref|YP_001124176.1|
ref|NP_387932.1|
RAAC01435
ref|YP_359065.1|

```
EPITAKLVANLLETAGASRVITLDLHAPQIQGFFDIPIDHLMGVPIILADYFKSKQLEDIV
NPITAKLVANLLETAGASRVITLDLHAPQIQGFFDIPIDHLMGVPIILADYFKSKQLEDIV
EPITAKLVANLLETAGASRVITLDLHAPQIQGFFDIPIDHLMGVPIILADYFKSKQLDDVV
EPITAKLFANLLETAGATRVIALDLHAPQIQGFFDIPIDHLMGVPIILGEYFEGKNLEDIV
DPITAKLVANLLQTAGASRVISMDLHAGQIQGFFDIPVDHLIGMPIILADYFLDKQIENPV
DPITAKLVSNLITVSGARRVIAMD L H A G Q I Q G F F D I P V D H L V G V P I L A K Y F N E K G L E N K V
***** : . * . : . * . ***** : . * . ***** : . * . : . *
```

```
ref|YP_145897.2|
sp|P42816|KPRS_BACCL
ref|YP_001124176.1|
ref|NP_387932.1|
RAAC01435
ref|YP_359065.1|
```

VVSPDHGGVTRARKLADRLKAPIAIIDKRRPKPNVAEVMNIVGQVAGKTAILIDDIIDTA
VVSPDHGGVTRARKLADRLKAPIAIIDKRRPKPNVAEVMNIVGQVAGKTAILIDDIIDTA
IVSPDHGGVTRARKLADRLKAPLAIIDKRRPRPNAEVMNIIIGQVKGTAILIDDIIDTA
IVSPDHGGVTRARKLADRLKAPIAIIDKRRPRPNVAEVMNIVGNIIEGKTAILIDDIIDTA
VVSPDMGGVTRARAFERLGAPLAIIDKRRPDANVAQVMNIIIGDIEGKTAILIDDMIDTA
VVSPDLGGVTRARDLAERIGAPIAIIDKRRPEPNVAEIMNIIIEGKTVIMIDDIIDTA
***** *

ref|YP_145897.2|
sp|P42816|KPRS_BACCL
ref|YP_001124176.1|
ref|NP_387932.1|
RAAC01435
ref|YP_359065.1|

GTITLAANALAENGAKEVYACCTHPVLSPAIERIQNSRIKELVVTNSIALPEEKKIDKI
 GTITLAANALAESGAKEVYACCTHPVLSPAIERIQSSKIKELVVTNSIALPEEKKIDKI
 GTITLGANALVENGAKEVYACCTHPVLSPAIERIQNSKIKELVVTNSIALPEEKKIDKI
 GTITLAANALVENGAKEVYACCTHPVLSPAVERINNSTIKELVVTNSIKLPEEKKIERF
 GTITAGAAELLKRGARGVYACCTHPVLSPGQVERLQNSAIEVVVTNTIALPEHKRIDKI
 GTITQGAQALMERGAKEVYVCTHPILSPGAVDRLANAPIKEVVVTNTIPLPEKKLPKI
 *** * * * * * ** *

```
ref|YP_145897.2|
sp|P42816|KPRS_BACCL
ref|YP_001124176.1|
ref|NP_387932.1|
RAAC01435
ref|YP_359065.1|
```

VELSVAPLIAEAI SRVYEMKSVSVLFD
VKLSVRPLIAEAITRVYEMKSVSVLFD
VELSVAPLIAEAITRVYEMQSVSVLFD
KQLSVGPLIAEAI IRVHEQQSVSYLF-
KVLSVAELIAEAI IRVHTQSVSQLFD
KVLVAPLMGEAIRRIHEDLSVSELF-
*** * . *** * . *** **

FIG. 180

```

ref|ZP_02170919.1|      MNKTELINAVAESADLSKKDATSAVDVFEVITDSLKKGDKVQLIGFGNFEVRERAARKG
ref|ZP_01862118.1|      MNKTELINAVAEAAELSKKDATKAVDAVFESIQDALANGDKVQLIGFGNFEVRERAARKG
ref|NP_692713.1|        MNKTDLVNAVAEKSELSKKDATKAVDAVFESVMDSLKNGEKVQLIGFGNFEVRERSARKG
ref|YP_535778.1|        -NKAALI ERVAEKTGLTKKDATVAVDVFFETIQDALVDGEKVQLIGFGNFEVRERAARKG
ref|YP_359077.1|        MNKAELVSVIAEKAEMTKKDAEKALNAVLA AIEEALKKGEKVQLVGFGTFEVRERAARKG
RAAC01442               MNKMELINRVAEKTNLKKKDAESAVNAVFEIIEEALANGEKVQIIIGFGTFETR SRAARSG
                        **  *:. : ** : :.***** *: : ** : : : * .*:*****:***.***.**:**.*

ref|ZP_02170919.1|      RNPQTGEEIEIPASNVP AFKPGKALKDAVK
ref|ZP_01862118.1|      RNPQTGEEIEISASKVP AFKPGKALKDAVK
ref|NP_692713.1|        RNPQTGEEIEIPASKVP AFKPGKALKD---
ref|YP_535778.1|        RNPQTGEEIEIPASKVP AFKPGKSLKDAVK
ref|YP_359077.1|        RNPQTGQEIEIPASKVP VF KPGKLLKE---
RAAC01442               RNPQTGEVIEIPASTVPAFKPGNKLKEVTR
                        *****: ***.**.**.*****: **:

```

```
ref|ZP_02583512.1|      LKSLSQEFSYSMETLLNQLHQYRKEQKVQQKQIKQVAKPSQVVQTKPKLTGFERAEREII
ref|NP_834002.1|       LKSLSQEFSYSMETLLNQLHQYRKEQKVQQKQIKQVAKPSQVVQTKPKLTGFERAEREII
ref|ZP_00238564.1|     LKSLSQEFSYSMETLLNQLHQYRKEQKVQQKQVQVSKPSQIVQTKPKLTGFERAEREII
ref|ZP_02604064.1|     LKSLSQEFSYSMETLLNQLHQYRKEQKVQQKQVQVSKPSQIVQTKPKLTGFERAEREII
ref|NP_846740.1|       LKSLSQEFSYSMETLLNQLHQYRKEQKVQQKQVQVSKPSQIVQTKPKLTGFERAEREII
RAAC00895              LRNLSQEFNLSVETLKEEMRGFAKPLRRRSPPREVETP--WVNSPRAAKGKDQVSLRII
*: *****:          *:          *:          *:          *:          *:          *
```

ref|ZP_02583512.1|
ref|NP_834002.1|
ref|ZP_00238564.1|
ref|ZP_02604064.1|
ref|NP_846740.1|
RAAC00895

YHMLQSAEVAVR-MESHIEDFHTEEHKGILYELYAYYEKGNPSVGTFLSWLSDEKLKNI
YHMLQSAEVAVR-MESHIEDFHTEEHKGILYELYAYYEKGNPSVGTFLSWLSDEKLKNI
YHMLQSPEVAVR-MESHIEDFHTEEHKGILYELYAYYEKGNPSVGTFLSWLSDEKLKNI
YHMLQSPEVAVR-MESHIEDFHTEEHKGILYELYAYYEKGNPSVGTFLSWLSDEKLKNI
YHMLQSPEVAVR-MESHIEDFHTEEHKGILYELYAYYEKGNPSVGTFLSWLSDEKLKNI
QAALFSQEAAEYLMEKGVTELAHPLHTALLSHVYSWRLEQPGQPPSAFIDRLEDEELVRL
* * *. * ** . : :: * . : * . : * : : : . . : * : . * . * : * . :

ref|ZP_02583512.1|
ref|NP_834002.1|
ref|ZP_00238564.1|
ref|ZP_02604064.1|
ref|NP_846740.1|
RAAC00895

ITDISTDEF--INPEYTEEVLOQGHLETIRRHOEKLEKMEIIFKVKQMEKTDPVEAAKYYV
ITDISTDEF--INPEYTEEVLOQGHLETIRRHOEKLEKMEIIFKVKQMEKTDPVEAAKYYV
ITDISTDEF--INPEYTEEVLOQGHLETIRRHOEKLEKMEIIFKIKQMEKTDPVEAAKYYV
ITDISTDEF--INPEYTEEVLOQGHLETIRRHOEKLEKMEIIFKIKQMEKTDPVEAAKYYV
ITDISTDEF--INPEYTEEVLOSHLETIRRHOEKLEKMEIIFKIKQMEKTDPVEAAKYYV
ASSLLFDEPPEITVELVEDYLR-ALELFRLEQELKEKVRANNEAEAAGDQEKREIKLQV
: . : ** * . * . * : * : ** : * . * * * : . : . : . * *

ref|ZP_02583512.1|
ref|NP_834002.1|
ref|ZP_00238564.1|
ref|ZP_02604064.1|
ref|NP_846740.1|
RAAC00895

AYLQNQKA-----
AYLQNQKA-----
AYLQNQKA-----
AYLQNQKA-----
AYLQNQKA-----
EWIQSQMATMKQPRALHAE
: . : * . * *

FIG. 182

```
ref|YP_001180871.1|      ----MNLFRFRLKELREEKNISRSDLAELGVSTQTIANYENGHREPFDTLKADIYFG
ref|YP_001114120.1|      ---VMSTIGERLRQLRKEKNINREDVANLLGVTVRSVTNYESGQRNLDPDQLIALADYYD
ref|YP_001662345.1|      --EKLNPFSQRLRQLREEKGLLQKDVAKILGITPSAYGYEQQGREPSMEVLKKLSDDFFN
ref|YP_001210714.1|      -----IFAKRLSFLITKNKLSKQAVANAINVSRPAVSQFANGENLPSVEKLIADFFD
RAAC03475                MEEVMTLFSQRLTELLDRTGTTRRSFAQALGVSERMVQYYITGKKDPTVETLIAIADFFD
ref|ZP_02442523.1|      -----IFQKRLKELRNAKGTSQIAIAAALGITDRGYRKYEAGDSEPTLSVIALADYFD
                        :  **  *   .   :  .*  :.:  :   :  *.   .   :  :*:..

ref|YP_001180871.1|      VTVDYLIGRSE-----
ref|YP_001114120.1|      VSMDYLTGRSDNPK--
ref|YP_001662345.1|      VSIDYLLGRTD-----
ref|YP_001210714.1|      VSLDYLVGRSDDPRR-
RAAC03475                VSLDYLVGRSENPERK
ref|ZP_02442523.1|      VSLDYLCGRSDDPAR-
                        *:*** ***:
```

FIG. 183

```
ref|YP_001114120.1|      --GERLRQLRKEKNINREDVANLLGVTVRSVTNYESGQRNLDPDQLIALADYYDVSM DYL
RAAC03560                MNGERLRALRKSLGLKRDEVAQAIGVTPRIITDYETETKRPTLDTAVKLADFFNVSLDYL
ref|YP_001662345.1|      ---QRLRQLREEKGLLQKDVAKILGITPSAYGYEQGKREPSMEVLKKLSDFFNVSIDYL
ref|NP_242309.1|         ---ERLRYLRKKHGLTMKELGKKINVAESTISGYENGNRKPDMDTLVKMAEYFNSSTDYL
ref|YP_146346.1|         --GDRLRKLRQEKKLTQEELGKKINVTKVSISGYENGNRTPDTETLQKLADFFNVTTDYL
ref|YP_146347.1|         ----RLKMLRLQKKLTHQDMADFLGITRQGYSKYENGQSQPDIDTINKLAEFFNVTTDYL
                        **: ** . : .:.. :.:  **      : :.:. : ***

ref|YP_001114120.1|      TGRSDNPKPPQTTTIDDE---
RAAC03560                VGRDDPTPPKRSSPSNEPGS
ref|YP_001662345.1|      LGRTDIRSP-----
ref|NP_242309.1|         LGRTEEP-APYQKQPPEKQ--
ref|YP_146346.1|         LGRTDHPNPPDQDDIPEE---
ref|YP_146347.1|         LGRTDDPTPPEQDDIPEE---
                        **: : .
```