The Journal of Antibiotics (2016) 69, 660–685

www.nature.com/ja

REVIEW ARTICLE

# Distribution of PASTA domains in penicillin-binding proteins and serine/threonine kinases of *Actinobacteria*

Hiroshi Ogawara[1,2]

**PASTA domains (penicillin-binding protein and serine/threonine kinase-associated domains) have been identified in penicillin-binding proteins and serine/threonine kinases of Gram-positive Firmicutes and *Actinobacteria*. They are believed to bind β-lactam antibiotics, and be involved in peptidoglycan metabolism, although their biological function is not definitively clarified. *Actinobacteria*, especially *Streptomyces* species, are distinct in that they undergo complex cellular differentiation and produce various antibiotics including β-lactams. This review focuses on the distribution of PASTA domains in penicillin-binding proteins and serine/threonine kinases in *Actinobacteria*. In *Actinobacteria*, PASTA domains are detectable exclusively in class A but not in class B penicillin-binding proteins, in sharp contrast to the cases in other bacteria. In penicillin-binding proteins, PASTA domains distribute independently from taxonomy with some distribution bias. Particularly interesting thing is that no *Streptomyces* species have penicillin-binding protein with PASTA domains. Protein kinases in *Actinobacteria* possess 0 to 5 PASTA domains in their molecules. Protein kinases in *Streptomyces* can be classified into three groups: no PASTA domain, 1 PASTA domain and 4 PASTA domain-containing groups. The 4 PASTA domain-containing groups can be further divided into two subgroups. The serine/threonine kinases in different groups may perform different functions. The pocket region in one of these subgroup is more dense and extended, thus it may be involved in binding of ligands like β-lactams more efficiently.**
*The Journal of Antibiotics* (2016) **69,** 660–685; doi:10.1038/ja.2015.138; published online 13 January 2016

## INTRODUCTION

PASTA domains (penicillin-binding protein and serine/threonine kinase-associated domains) were first identified in the C terminus of *Streptococcus pneumoniae* penicillin-binding protein PBP2x.[1] In the crystal structure, van der Waal's interactions between the β-lactam ring of cefuroxime, a second-generation cephalosporin, and one of two PASTA domains in PBP2x were observed.[2] In view of the structural similarity of the β-lactam ring of cefuroxime to the D-alanyl-D-alanine residues of the stem pentapeptide of peptidoglycan precursors, Yeats *et al.*[3] proposed that PASTA domains bind uncross-linked peptidoglycan. Subsequently, these domains were found in a variety of high MW penicillin-binding proteins (PBPs) as well as in serine/threonine kinases (STPKs) mainly from Gram-positive Firmicutes and *Actinobacteria*. The domains consist of 60–70 amino-acid residues and occur singly or as a few successive copies. Although PASTA domains show low amino-acid sequence similarity, they share strong structural conservation. Each domain has a globular fold consisting of three β-strands and one α-helix.

As for the interaction between PASTA domains in PBPs and β-lactam compounds, several studies have been published. The PASTA domains in PBP2x of *S. pneumoniae* bind the β-lactam antibiotic cefuroxime and a fluorescent penicillin Bocillin FL.[2,4] Furthermore, localization of PBP2x together with FtsZ and FtsW to cell division sites depends on its PASTA domains, but not on its transpeptidase activity.[5,6] As PASTA domains bind uncross-linked peptidoglycan,[7] it is suggested that localization of PBP2x is dependent on the localization of its substrate. In addition, alanine707 within the PASTA domain is important for stabilization.[8] On the other hand, the PASTA domain of *Mycobacterium tuberculosis* PBP PonA2 does not bind muropeptide nor does it bind the two β-lactam antibiotics cefuroxime and cefotaxime, or polymeric peptidoglycan.[9] In *Bacillus subtilis*, among 16 different PBPs only 2 of them, PBP2b and SpoVD, contain PASTA domains: two PASTA domains in PBP2b and one PASTA domain in SpoVD. However, the PASTA domain in SpoVD is not essential for cortex biosynthesis and not important for targeting SpoVD to the forespore outer membrane during sporulation.[10] Therefore, functionality of PASTA domains in PBPs remains controversial.

The functions of the PASTA domains in STPKs are more clear.[11,12] *S. pneumoniae* possesses only a single STPK StkP with four PASTA domains in its C-terminal region, and its corresponding phosphatase PhpP forms a functional pair with StkP.[13,14] The PASTA domains of StkP in *S. pneumoniae* were shown to bind synthetic and native peptidoglycan and β-lactam antibiotics.[7] In response to the binding, the PASTA domains are involved in the activation of StkP and substrate recognition.[15] Activated StkP phosphorylates cell division proteins DivIVA and FtsZ. FtsZ is a prokaryotic tubulin homolog.[15–18] Depending on the extracellular PASTA domains, StkP is recruited to cell division sites, interacts with FtsZ and is involved in the regulation of cell division and bacterial growth.[18–20] Therefore, cell wall

[1]HO Bio Institute, Tokyo, Japan and [2]Meiji Pharmaceutical University, Tokyo, Japan
Correspondence: Professor H Ogawara, HO Bio Institute, 33-9, Yushima-2, Bunkyo-ku, Tokyo 113-0034, Japan.
E-mail: hogawara@sc5.so-net.ne.jp
Received 22 August 2015; revised 15 November 2015; accepted 25 November 2015; published online 13 January 2016

biosynthesis and uncross-linked peptidoglycan are the signals for StkP to localize to cell division sites and stimulate its autophosphorylation activity.[16]

PknB (Rv0014c) is one of 11 STPKs but the only PASTA-containing STPK in *M. tuberculosis*.[21–23] The PknB of *M. tuberculosis* has four PASTA domains.[24] The extracellular PASTA domains bind muropeptides depending on the presence of specific amino acids at the second and third positions of the stem peptide.[25] In addition, the extracellular PASTA domain is required for proper localization of PknB to the mid-cell. The PASTA domains also have a role in stimulating growth by binding exogenous peptidoglycan fragments and are suggested to be a sensor for a signaling molecule that promotes initial growth.[24] PknB phosphorylates the DivIVA ortholog Wag31 in *M. tuberculosis* and *Corynebacterium glutamicum*,[26,27] and this phosphorylation may trigger remodeling of bacterial morphology.[28] Peptidoglycan fragments can induce germination of dormant *B. subtilis* spores through interactions with the PASTA domain-containing STPK, PkrC in *B. subtilis* and the serine/threonine kinase PrkC from *Staphylococcus aureus* can induce germination of dormant spores in *B. subtilis* through interaction with peptidoglycan fragments.[11,29]

*Actinobacteria*, especially *Streptomyces* species, are unique in that they are filamentous, soil-dwelling, Gram-positive bacteria and are characterized by their ability to undergo complex cellular differentiation similar to filamentous fungi.[30] Furthermore, *Streptomyces* species produce a wide variety of secondary metabolites including β-lactam antibiotics and enzymes.[31] Previously, roles of PBPs in *Actinobacteria* were reviewed[32] and STPKs in *Streptomyces coelicolor* A3(2) were discussed from an evolutionary point of view.[33] PASTA domains, as the name implies, are domains that suggest some type of relationship between PBPs and STPKs. Accordingly, the distribution of PASTA domain in PBPs and STPKs in *Actinobacteria* is of interest from the point of views of their functional role in peptidoglycan metabolism and their evolutionary origins. This review paper deals with this topic.

## DISTRIBUTION OF PASTA DOMAIN IN PBPS

Table 1 summarizes the genome size (Mb), G+C%, the number of proteins, the number of putative class A and class B PBPs, PBPs with PASTA domain, the number of putative STPKs and STPKs with PASTA domain(s) in 95 species of *Actinobacteria*. PBP and STPK genes were selected at two steps. From the database (Microbial Genomes in NCBI, http://www.ncbi.nlm.nih.gov/genomes/MICROBES), putative PBP and STPK genes were screened with keywords 'PBP', 'penicillin-binding protein', 'STPK' and 'protein kinase' at the first step. For each species, the screened sequences were analyzed at the second step by aligning and constructing phylogenetic trees by using ClustalW as implemented by MEGA together with the representative sequences of PBPs and STPKs.

It is intriguing that PBPs with PASTA domains are detected only in class A PBPs in *Actinobacteria*. No class B PBPs in *Actinobacteria* contain PASTA domains. This is in sharp contrast to the cases in *B. subtilis*, *Clostridium perfringens* and *S. pneumonia*. All of the PBPs with PASTA domains in these bacteria, that is, PBP2b (GenBank accession number: BSU15160, the same thereafter) and SpoVD (BSU15170) in *B. subtilis*, SpoVD (BAB80270), SpoVD (BAB81569) and SpoVD (BAB81586) in *C. perfringens* and PBP2x in *S. pneumoniae* belong to the class B PBPs. Class A PBPs have both transglycosylase and transpeptidase domains, whereas class B PBPs have only transpeptidase domain. Therefore, it is interesting to know the interaction between transglycosylase and PASTA domains in *Actinobacteria*.

In *Actinobacteria*, most PBPs with PASTA domains possess only one PASTA domain, with the exceptions of *Beutenbergia cavernae* DSM 12333 BCAV_4182, *Clavibacter michiganensis* subsp. *michiganensis* NCPPB 382 CMM_0919, *Isoptericola variabilis* 225 ISOVA_3000, *Leifsonia xyli* subsp. *xyli* str. CTCB07 LXX_03600, *Tropheryma whipplei* str. *Twist* TWT_0705 and *T. whipplei* TW08/27 TW_0722, where each has two PASTA domains. In contrast, BSU15160 in *B. subtilis* subsp. *subtilis* str. 168, BAB80270, BAB81569 and BAB81586 in *C. perfringens* str. 13 and PBP2x in *S. pneumoniae* R6 each possess two PASTA domains, whereas only BSU15170 in *B. subtilis* subsp. *subtilis* str. 168 has one PASTA domain. These two features, that is, that PBPs with a PASTA domain belong absolutely to class A PBPs, and that there is only one PASTA domain in PBPs, are quite characteristic properties in PBPs in *Actinobacteria*. In addition, none of the *Streptomyces* species analyzed here has PBPs with PASTA domains (Figure 1 and Table 1) as discussed later.

Genome sizes and G+C contents are not apparently related to the presence of PBPs with PASTA domain(s). For example, some bacterial species with small genome sizes such as *Bifidobacterium adolescentis* (2.09 Mb), *Mobiluncus curtisii* (2.15 Mb) and *Olsenella uli* (2.05 Mb) have PBPs without PASTA domains, but *Propionibacterium acnes* (2.49 Mb) and *T. whipplei* (0.93 Mb) each have a PBP with a PASTA domain. On the other hand, bacterial species with large genome sizes such as *Amycolatopsis mediterranei* (10.24 Mb), *Catenulispora acidiphila* (10.47 Mb) and *Streptosporangium roseum* (10.37 Mb) have no PBP with PASTA domains, whereas *Actinosynnema mirum* (8.25 Mb) and *Actinoplanes missouriensis* (8.77 Mb) each have a PBP with a PASTA domain. Similarly, some bacterial species with high G+C contents such as *Clavibacter michiganensis* (72.5%), *I. variabilis* (73.9%) and *Nocardiopsis dassonvillei* (72.7%), and some with low G+C contents such as *T. whipplei* (46.3%) have a PBP with PASTA domains, whereas others with high G+C contents such as *Geodermatophilus obscurus* (74.0%), *Micrococcus luteus* (73.0%) and *Streptomyces griseus* (72.2%) or with low G+C contents such as *Atopobium parvulum* (45.7%) and *Gardnerella vaginalis* (41.2%) have PBPs without PASTA domains. Moreover, the presence of PBPs with PASTA domains seems to be distributed independently from taxonomic classification with some distribution bias (Figure 1). All species belonging to *Propionibacteriales*, *Corynebacteriales* and *Micromonosporales* possess PBPs with a PASTA domain. However, PBPs in *Micrococcales* species are divided into two groups: PBPs with a PASTA domain and those without a PASTA domain. Similarly, in *Pseudonocardiales* (*Saccharopolyspora erythraea* NRRL 2338 and *A. mirum* DSM 43827) and *Coriobacteriales* (*Eggerthella lenta* DSM 2243), some PBPs have PASTA domains but others do not. Of particular interest is that none of the *Streptomyces* species analyzed here has PBPs with PASTA domains (Figure 1 and Table 1). However, antibiotic production, which is a notable feature of many *Streptomyces* spp., is presumably not related to this phenomenon, because *Saccharopolyspora erythraea*, which produces erythromycin, possesses a PBP with a PASTA domain. The reason for the lack of PASTA domains in the PBPs of *Streptomyces* spp. remains to be clarified. Furthermore, no PBP with PASTA domain was detectable in orders *Bifidobacteriales* and *Pseudonocardiales*.

The PASTA domains were searched for in the PBPs of all of the *Actinobacteria* analyzed in this paper by Blast analysis of NCBI. The secondary structures of these PASTA domains were analyzed by using PSIPRED software (http://bioinf.cs.ucl.ac.uk/psipred/). The 51 PASTA domains detected by these analyses each comprise a small globular fold consisting of three β-sheets and one α-helix (Figure 2). The lengths of amino-acid residues in the PASTA domains are similar to

**Table 1 The numbers and types of putative PBP and protein kinase genes**

| *Bacteria* | *Prefix* | *Genome size (Mb)* | *G+C %* | *No. of protein* | *Class A PBP* | *Class B PBP* | *PBP with PASTA domain*[a] | *No. of protein kinase* | *Protein kinase with PASTA domain*[a] |
|---|---|---|---|---|---|---|---|---|---|
| *Acidimicrobium ferrooxidans* DSM 10331 | AFER | 2.16 | 68.3 | 1964 | | 0089, 0769, 1250 | None | 4 | 0087(4), 0088(2), 1553(4) |
| *Acidothermus cellulolyticus* 11B | ACEL | 2.44 | 66.9 | 2157 | 2004, 2135 | 0020, 0751, 1004 | 2004(1) | 4 | 0019(3), 0986(4) |
| *Actinomyces urogenitalis* DORA_12 | Q605 | 2.6 | 68.4 | 3034 | AUC00015G0001, AUC00927G0001, AUC01040G0002 | AUC00266G0006 AUC00451G0003 | None | 5 | Q605_AUC00266G0004(4), Q605_AUC00956G0003(4), |
| *Actinoplanes missouriensis* 431 | AMIS | 8.77 | 70.8 | 8124 | 10850, 39800, 60090, 78300, 80960 | 00450, 15180, 71610 | 60090(1), 78300(1) | 26 | 470(3), 14740(4) |
| *Actinosynnema mirum* DSM 43827 | AMIR | 8.25 | 73.7 | 6916 | 0235, 5121, 7034 | 0023, 5772, 5886 | 0235(1) | 29 | 0021(4), 1396(4) |
| *Amycolatopsis mediterranei* U32 | AMED | 10.24 | 71.3 | 9228 | 3027, 8853, 9287, 9288 | 2045, 2965 | None | 25 | 0051(4), 2318(4) |
| *Amycolicicoccus subflavus* DQS3-9A1 | AS9A | 4.86 | 62.2 | 4705 | 0253, 2577 | 0033, 4146, 4510 | 0253(1) | 13 | 0031(4), 1301(4) |
| *Arcanobacterium haemolyticum* DSM20595 | ARCH_RS | 1.99 | 53.1 | 1732 | 08675 | 00530, 03395 | None | 3 | 00525(3), 03830(4) |
| *Arthrobacter aurescens* TC1 | AAUR | 5.23 | 62.4 | 4588 | 3369, 3416 | 0030, 1704, 3184, 4181 | None | 5 | 0028(3), 1694(4) |
| *Atopobium parvulum* DSM 20469 | Apar | 1.54 | 45.7 | 1353 | 1010 | 0480, 0673, 1344 | None | 3 | 1345(4) |
| *Beutenbergia cavernae* DSM 12333 | BCAV | 4.67 | 73.1 | 4197 | 0604, 4182 | 0028, 0389, 2416 | 0604(1), 4182 (2) | 9 | 0026(3), 1908(4) |
| *Bifidobacterium adolescentis* ATCC 15703 | BAD | 2.09 | 59.2 | 1632 | 0157, 1336 | 0040, 1107 | None | 7 | 0038(4), 0143(1), 1014(3) |
| *Brachybacterium faecium* DSM 4810 | Bfae | 3.61 | 72.0 | 3068 | 26280, 31830 | 06090, 10750, 26970 | 26280(1) | 9 | 14290(4), 26990(4) |
| *Brevibacterium casei* S18 | C272 | 3.66 | 68.1 | 3216 | 05329, 06024, 15065 | 05334 | None | 7 | 12727(1), 12732(3), 13024(3) |
| *Catenulispora acidiphila* DSM 44928 | CACI | 10.47 | 69.8 | 8913 | 1390 | 0037, 1307, 1448, 1656, 5826, 6659, 7282 | None | 52 | 6001(3) |
| *Cellulomonas flavigena* DSM 20109 | Cfla | 4.12 | 74.3 | 3678 | 3099, 3701 | 0027, 1590, 3460 | 3099(1), 3701 (1) | 9 | 0025(3), 2064(4) |
| *Clavibacter michiganensis* subsp. *michiganensis* NCPPB 382 | CMM | 3.4 | 72.5 | 3078 | 0915, 0919 | 0017, 1865 | 0919(2) | 8 | 1873(4) |
| *Conexibacter woesei* DSM 14684 | Cwoe | 6.36 | 72.7 | 5914 | 3542 | 0016, 1104, 1801, 2661, 3775, 5329 | None | 6 | 0017(4) |
| *Corynebacterium diphtheriae* NCTC 13129 | DIP | 2.49 | 53.5 | 2272 | 0298, 2294 | 0055, 1497, 1604, | 0298(1) | 4 | 0053(4), 1615(5) |
| *Corynebacterium glutamicum* ATCC 13032 | NCGL | 3.31 | 53.8 | 2959 | 0274, 2884 | 0042, 1933, 2084 | 0274(1) | 5 | 0040(4), 2095(5) |
| *Cryptobacterium curtum* DSM 15641 | Ccur | 1.62 | 50.9 | 1357 | 02290 | 06100, 09550 | None | 3 | 03300(4) |
| *Eggerthella lenta* DSM 2243 | ELEN_RS | 3.63 | 64.2 | 3049 | 02953, 11730 | 06285, 10550, 12605, 15235 | 02953(1) | 7 | 15240(3) |
| *Frankia alni* ACN14a | FRAAL | 7.50 | 72.8 | 6700 | 1281, 6546, 6857 | 1919, 2190, 5852, 6753 | 6546(1) | 50 | 3851(3), 4020(3), 5081(4), 6755(3) |
| *Frankia* sp. CcI3 | FRANCCI3 | 5.43 | 70.1 | 4499 | 0754, 4277, 4526 | 1214, 1409, 3641, 4434 | 4277(1) | 29 | 3077(4), 3623(2), 4436(3) |
| *Gardnerella vaginalis* HMP9231 | HMPREF9231 | 1.73 | 41.2 | 1299 | 1162, 1288 | 0015, 1089 | None | 4 | 0017(4), 0948(4) |

**Table 1 (Continued)**

| Bacteria | Prefix | Genome size (Mb) | G+C % | No. of protein | Class A PBP | Class B PBP | PBP with PASTA domain[a] | No. of protein kinase | Protein kinase with PASTA domain[a] |
|---|---|---|---|---|---|---|---|---|---|
| *Gardnerella vaginalis* ATCC 14019 | HMPREF0421 | 1.67 | 41.4 | 1365 | 20263, 20394 | 20139, 20438 | None | 5 | 20142(4), 20603(4) |
| *Geodermatophilus obscurus* DSM 43160 | GOBS_RS | 5.32 | 74.0 | 4811 | 24855 | 16015, 19365 | None | 6 | 00320(1) |
| *Gordonia polyisoprenivorans* VH2 | GPOL | 5.67 | 67.0 | 5110 | 07230, 49200 | 00240, 20210, 27800 | 07230(1) | 23 | 00220(4) |
| *Isoptericola variabilis* 225 | ISOVA | 3.31 | 73.9 | 2881 | 2701, 3000 | 0021, 1284, 2892 | 3000(2) | 6 | 0019(4), 1812(4) |
| *Janibacter* sp. HTCC2649 | JNB | 4.23 | 68.4 | 4100 | 13593 | 00615, 05649, 12079 | None | 9 | 13208(3) |
| *Jonesia denitrifican* DSM 20603 | JDEN_RS | 2.75 | 58.4 | 2481 | 02160, 12675 | 05355 | None | 5 | 00860(4), 07220(4) |
| *Kineococcus radiotolerans* SRS30216 | KRAD | 4.96 | 74.2 | 4681 | 0429, 4341 | 0073, 0475, 3205 | 0429(1) | 6 | 0072(3), 3232(4) |
| *Kitasatospora setae* KM-6054 | KSE | 8.78 | 74.2 | 7566 | 27750, 36430, 38960, 59840 | 26130, 39410, 46160, 46190, | None | 52 | 21520(3), 39400(3) |
| *Kocuria rhizophila* DC2201 | KRH | 2.7 | 71.2 | 2357 | 04490 | 14880, 20650 | None | 4 | 14970(3), 20660(1), 20670(3) |
| *Kribbella flavida* DSM 17836 | Kfla | 7.58 | 70.6 | 6943 | 0444, 2401, 7017 | 0061, 2302, 2880, 6892 | 0444(1) | 20 | 0060(3), 2845(4) |
| *Kytococcus sedentarius* DSM 20547 | KSED | 2.79 | 71.6 | 2554 | | 00200, 09230, 16620 | None | 3 | 00180(1) |
| *Leifsonia xyli* subsp. *xyli* str. CTCB07 | LXX | 2.58 | 67.7 | 2028 | 02090, 03600, 05450, 23190 | 00230, 15320 | 03600 (2),05450(1) | 6 | 00210(3), 00220(1), 15400(4) |
| *Micrococcus luteus* NCTC 2665 | Mlut | 2.5 | 73.0 | 2236 | 18460 | 00770, 13660 | None | 6 | 00750(3), 00760(2), 13750(3) |
| *Micromonospora aurantiaca* ATCC 27029 | MICAU | 7.03 | 72.8 | 6222 | 3350, 4230, 4961, 5144, 5927, 6271 | 0098, 4478, 5070 | 5927(1) | 17 | 0100(3), 4509(4) |
| *Mobiluncus curtisii* ATCC 43063 | HMPREF0573 | 2.15 | 55.4 | 1909 | 10708, 10751 | 10644, 11576 | None | 3 | 10646(2), 11797(4) |
| *Mycobacterium avium* 104 | MAV | 5.48 | 69.0 | 5120 | 0071, 0446 | 0020, 2330, 3723 | 0446(1) | 11 | 0017(4) |
| *Mycobacterium bovis* BCG str. Tokyo 172 | JTY | 4.37 | 65.6 | 3944 | 0051, 3742 | 0016, 2174, 2881 | 3742(1) | 11 | 0014(4) |
| *Mycobacterium leprae* TN | ML | 3.27 | 57.8 | 1605 | 2308, 2688 | 0018, 0908, 1577 | 2308(1) | 4 | 0016(4) |
| *Mycobacterium tuberculosis* H37Rv | RV | 4.41 | 65.6 | 3906 | 0050, 3682 | 0016c, 2163c, 2864c | 3682(1) | 11 | 0014c(4) |
| *Nakamurella multipartita* DSM 44233 | NAMU | 6.06 | 70.9 | 5240 | 0707 | 0079, 2190, 3930 | 0707(1) | 17 | 0077(4), 2392(5), 3222(5), 4030(1) |
| *Nocardia farcinica* IFM 10152 | NFA | 6.29 | 70.7 | 5934 | 03390, 55490, 55570 | 820, 17600, 18430, 41160 | 03390(1) | 21 | 800(4), 17330(4) |
| *Nocardioides* sp. JS614 | NOCA | 5.29 | 71.5 | 4909 | 0326, 4676 | 0024, 3069, 3462, 4600 | 0326(1) | 12 | 0022(3), 3086(4), 3705(1) |
| *Nocardiopsis dassonvillei* subsp. *dassonvillei* DSM 43111 | Ndas | 6.54 | 72.7 | 5497 | 0250, 0388, 4845, 5307 | 0890, 2552, 3385, 3720, 5248 | 5307(1) | 35 | 0036(3), 31139(4), 5246(3) |
| *Olsenella uli* DSM 7084 | OLSU_RS | 2.05 | 64.7 | 1727 | 02875 | 04295, 05540, 08820 | None | 4 | 08825(4) |
| *Propionibacterium acnes* 266 | PAZ | 2.49 | 60.0 | 2345 | 01380, 22310 | 01980, 08010 | 22310(1) | 4 | c01970(3) |
| *Propionibacterium acnes* KPA171202 | PPA | 2.56 | 60.0 | 2297 | 0126, 2149 | 0185, 0752 | 2149(1) | 4 | 0184(3) |
| *Renibacterium salmoninarum* ATCC 33209 | RSAL33209 | 3.16 | 56.3 | 3507 | 2795 | 2241, 2500, 2891 | None | 5 | 1800(2), 2512(4), 2892(2), 2893(3) |
| *Rhodococcus erythropolis* PR4 | RER | 6.90 | 62.3 | 6437 | 04630, 58380, 58990 | 00300, 10660, 25560, 35580 | 04630(1) | 8 | 00280(4), 35760(4) |
| *Rothia dentocariosa* ATCC 17931 | HMREF0733 | 2.51 | 53.7 | 2217 | 10478 | 10948, 11665 | None | 4 | 10946(3), 10947(2) |
| *Rubrobacter xylanophilus* DSM 9941 | Rxyl | 3.23 | 70.5 | 3140 | 1310, 2308 | 0022, 1138, 1498 | None | 8 | 0021(4) |
| *Saccharomonospora azurea* NA-128 | SACAZ | 4.76 | 70.0 | 4376 | 02388, 02721 | 01210, 02813 | None | 21 | 01512(4), 02811(4) |
| *Saccharomonospora viridis* DSM 43017 | SVIR | 4.31 | 67.3 | 3828 | 36250, 39340 | 24950, 33330 | None | 16 | 00360(4), 27990(4) |

**Table 1 (Continued)**

| *Bacteria* | *Prefix* | *Genome size (Mb)* | *G+C %* | *No. of protein* | *Class A PBP* | *Class B PBP* | *PBP with PASTA domain*[a] | *No. of protein kinase* | *Protein kinase with PASTA domain*[a] |
|---|---|---|---|---|---|---|---|---|---|
| *Saccharopolyspora erythraea* NRRL 2338 | SACE | 8.21 | 71.1 | 7197 | 0314, 0385, 6352, 7356 | 0046, 5864, 5990 | 0314(1) | 28 | 0044(4), 1710(4) |
| *Salinispora arenicola* CNS-205 | SARE | 5.79 | 69.5 | 4917 | 3240, 3923, 4021, 4796, 5078 | 0051, 3444 | 4796(1) | 10 | 0053(3), 3480(4) |
| *Salinispora tropica* CNB-440 | STROP | 5.18 | 69.5 | 4536 | 3015, 3548, 3639, 4354, 4560 | 0046, 3218 | 4354(1) | 9 | 0048(3), 3253(4) |
| *Sanguibacter keddieii* DSM 10542 | Sked | 4.25 | 71.9 | 3710 | 04860, 37820 | 00200, 22860, 36830 | 37820(1) | 6 | 00180(4), 15820(4) |
| *Segniliparus rotundus* DSM 44985 | SROT_RS | 3.16 | 66.8 | 2999 | 05555 | 01875, 11430 | 05555(1) | 4 | 01870(4) |
| *Slackia heliotrinireducens* DSM 20476 | SHEL | 3.17 | 60.2 | 2766 | 08350 | 08800, 14210 | None | 13 | 27070(4) |
| *Stackebrandtia nassauensis* DSM 44728 | Snas | 6.84 | 68.1 | 6379 | 1056, 5713 | 2665, 6473 | None | 43 | 1251(1), 1807(1), 2077(1), 2079(1), 2132(1), 2149(1), 2236(1), 3561(1), 5130(1), 5305(1), 6471(3) |
| *Streptomyces albulus* NK660 | DC74 | 9.37 | 72.3 | 8086 | 3326, 4231, 4477, 5204, 7647 | 3128, 3135, 3598, 4185, 5459 | None | 40 | 2595(4), 4186(4) |
| *Streptomyces albus* J1074 | XNR | 6.84 | 73.3 | 5832 | 1770, 2736, 2983, 4127 | 1496, 2096, 2097, 3038, 4337, 4789 | None | 25 | 3037(4), 3064(1), 4768(4) |
| *Streptomyces avermitilis* MA-4680 | SAV | 9.12 | 70.7 | 7676 | 3225, 4294, 4423, 4583, 5179, 7219 | 2952, 3603, 3604, 4339, 5458, 6116, 6387 | None | 33 | 4338(4), 4371(1), 6092(4) |
| *Streptomyces bingchenggensis BCW-1* | SBI | 11.94 | 70.8 | 10 022 | 03076, 04174, 05361, 05810, 06697, 09068 | 02283, 04376, 05407, 06233, 07119, 07873 | None | 67 | 05406(4), 07851(4) |
| *Streptomyces cattleya N*RRL8057 | SCAT | 8.09 | 73.0 | 7470 | 1929, 2889, 3140, 3906 | 0768, 1207, 1730, 1901, 3088, 4153, 5676 | None | 18 | 1232(4), 3053(1), 3089(4) |
| *Streptomyces clavuligerus ATCC27064* | SCLAV | 8.56 | 72.4 | 7287 | 2006, 2887, 3942 | 1087, 1301, 1774, 2276, 2947, 4154, 4179, 4180, 4198 | None | 33 | 1326(4), 2946(4), 2991(1) |
| *Streptomyces coelicolor* A3(2) | SCO | 9.05 | 72.0 | 8152 | 2897, 3580, 3901, 5039 | 1875, 2090, 2608, 3156, 3157, 3771, 3847, 4013, 5301 | None | 37 | 2110(4), 3821(1), 3848(4) |
| *Streptomyces collinus Tu365* | B446 | 8.38 | 72.6 | 7113 | 15140, 19060, 23580 | 09755, 10960, 13820, 16355, 19320, 24955 | None | 39 | 11080(4), 19315(4), 19450(1) |
| *Streptomyces davawensis* JCM4913 | BN159 | 9.56 | 70.6 | 8616 | 3357, 4150, 4546, 5391 | 3075, 4478, 5121, 5122, 5684, 6352, 6632 | None | 35 | 4396(1), 4479(4), 6328(4) |
| *Streptomyces ghanaensis* ATCC 14672 | SSFG | 8.51 | 72.2 | 7891 | 02387, 02608, 03635, 04479 | 02394, 03587, 04216, 04217, 04765, 05266 | None | 35 | 03552(1), 03588(4) |
| *Streptomyces griseoflavus* Tu4000 | SSRG | 8.05 | 71.7 | 6337 | 02182, 02879, 03203, 03961 | 01957, 03076, 03158, 03705, 03706, 04177, 04634, 04850 | None | 35 | 03115(1), 03159(4), 04610(4) |
| *Streptomyces griseus* subsp. *griseus* NBRC 13350 | SGR | 8.55 | 72.2 | 7136 | 2494, 3341, 3679, 4647 | 2203, 3726, 4232, 4340, 4934, 5621 | None | 30 | 3725(4), 5391(4) |
| *Streptomyces hygroscopicus* subsp. *jinggangensis* 5008 | SHJG | 10.38 | 71.9 | 9107 | 3853, 4373, 5171, 5432, 6136 | 3336, 4100, 4627, 4628, 5219, 6411 | None | 43 | 3594(4), 5218(4), 5252(1) |
| *Streptomyces lividans* TK24 | SLIV | 8.35 | 72.2 | 7396 | 13190, 18760, 20390, 23205 | 11865, 18345, 19035, 19450, 21910, 21915, 24635, 28340 | None | 35 | 19030(4), 19175(1), 27155(4) |
| *Streptomyces rimosus* subsp. *rimosus* ATCC10970 | SRIM | 9.5 | 71.9 | 8411 | 00295, 08328, 13873, 22689 | 00065, 04191, 06646, 15770, 26297, 31850, 31885 | None | 40 | 00070(4), 07563(1), 24996(4) |

**Table 1 (Continued)**

| *Bacteria* | *Prefix* | *Genome size (Mb)* | *G+C %* | *No. of protein* | *Class A PBP* | *Class B PBP* | *PBP with PASTA domain*[a] | *No. of protein kinase* | *Protein kinase with PASTA domain*[a] |
|---|---|---|---|---|---|---|---|---|---|
| *Streptomyces scabiei* 87.22 | SCAB | 10.15 | 71.5 | 8746 | 33601, 41401, 56801, 64431 | 10101, 29591, 45551, 53611, 53621, 60051, 70631 | None | 43 | 0931(1), 45201(1), 67711(4) |
| *Streptomyces* sp. PAMC26508 | F750 | 7.63 | 71.1 | 7073 | 2719, 3337, 3596, 4580 | 1743, 2434, 2998, 3546, 4834, 6320 | None | 27 | 1975(4), 3512(1), 3547(4) |
| *Streptomyces* sp. SirexAA-E | SACTE | 7.41 | 71.7 | 6357 | 2371, 3027, 3329, 4291 | 1307, 1519, 2029, 2618, 2701, 3283, 4532 | None | 32 | 1542(4), 3284(3) |
| *Streptomyces sviceus* ATCC 29083 | SSEG | 9.31 | 69.8 | 8205 | 01073, 07525, 03439, 04164 | 00010, 00011, 00733, 01896, 09019, 09517, | None | 40 | 02705(4), 06024(4) |
| *Streptomyces venezuelae* ATCC 10712 | SVEN | 8.23 | 72.4 | 7448 | 2646, 3350, 3677, 4705 | 1522, 1745, 2386, 2985, 3631, 4995 | None | 37 | 1769(4), 3592(1), 3632(4) |
| *Streptomyces violaceusniger* Tu4113 | STRVI | 11.14 | 71.0 | 8985 | 1350, 2314, 3845, 8252, 9005 | 0275, 1135, 3190, 7171, 7897, 7904 | None | 38 | 0274(4), 7194(4) |
| *Streptomyces viridochromogenes* DSM 40736 | SSQG | 8.65 | 71.1 | 7714 | 02328, 02941, 03901, 04279, 05113 | 01781, 02628, 03242, 03243, 03958, 05348 | None | 37 | 02054(4), 03956(4), 03996(1) |
| *Streptomyces zinciresistens* K42 | SZN | 8.22 | 72.5 | 7579 | 06389, 16730, 18682, 28493 | 02952, 10458, 13352, 17932, 18819, 22026 | None | 39 | 08009(4), 17937(4), 21616(1) |
| *Streptosporangium roseum* DSM 43021 | SROS | 10.37 | 70.9 | 8975 | 2902, 3010, 8177, 9363 | 0113, 1441, 1456, 2864, 3583, 4062, 7279, 7683, | None | 71 | 0111(3), 1975(3), 2780(4), 8314(1) |
| *Thermobifida fusca* YX | TFU | 3.64 | 67.5 | 3087 | 0570, 3097 | 1416, 2475 | None | 23 | 1041(4), 3066(3) |
| *Thermobispora bispora* DSM 43833 | TBIS | 4.19 | 72.4 | 3546 | 0195, 1426, 3106, 3566 | 0053, 0796, 1401, 1685, 1727, 2465 | None | 29 | 0055(3), 1337(4), 2915(1) |
| *Thermomonospora curvata* DSM 43183 | TCUR | 5.64 | 71.6 | 4890 | 1026, 1268, 4921, 4955 | 0065, 1542, 2932, 4002 | 1026(1), 4921(1) | 49 | 0063(3), 3036(4) |
| *Tropheryma whipplei str.* Twist | TWT | 0.93 | 46.3 | 808 | 0705 | 0222, 0776 | 0705(2) | 4 | 216(3), 778(2) |
| *Tropheryma whipplei* TW08/27 | TW | 0.93 | 46.3 | 783 | 0722 | 0548, 0787 | 0722(2) | 4 | 555(4), 789(2) |
| *Tsukamurella paurometabola* DSM 20162 | TPAU | 4.48 | 68.4 | 4242 | 3939, 4192 | 0029, 0349, 1690, 2652, 3973 | 3939(1) | 12 | 0027(4), 2669(5) |
| *Xylanimonas cellulosilytica* DSM 15894 | XCEL_RS | 3.83 | 72.5 | 3413 | 07030, 15260, 16845 | 06420, 16455 | 16845(1) | 6 | 00100(3), 10040(4) |

Abbreviation: PASTA domain, penicillin-binding protein and serine/threonine kinase-associated domains.
[a]The number within parenthesis is the number of PASTA domains.

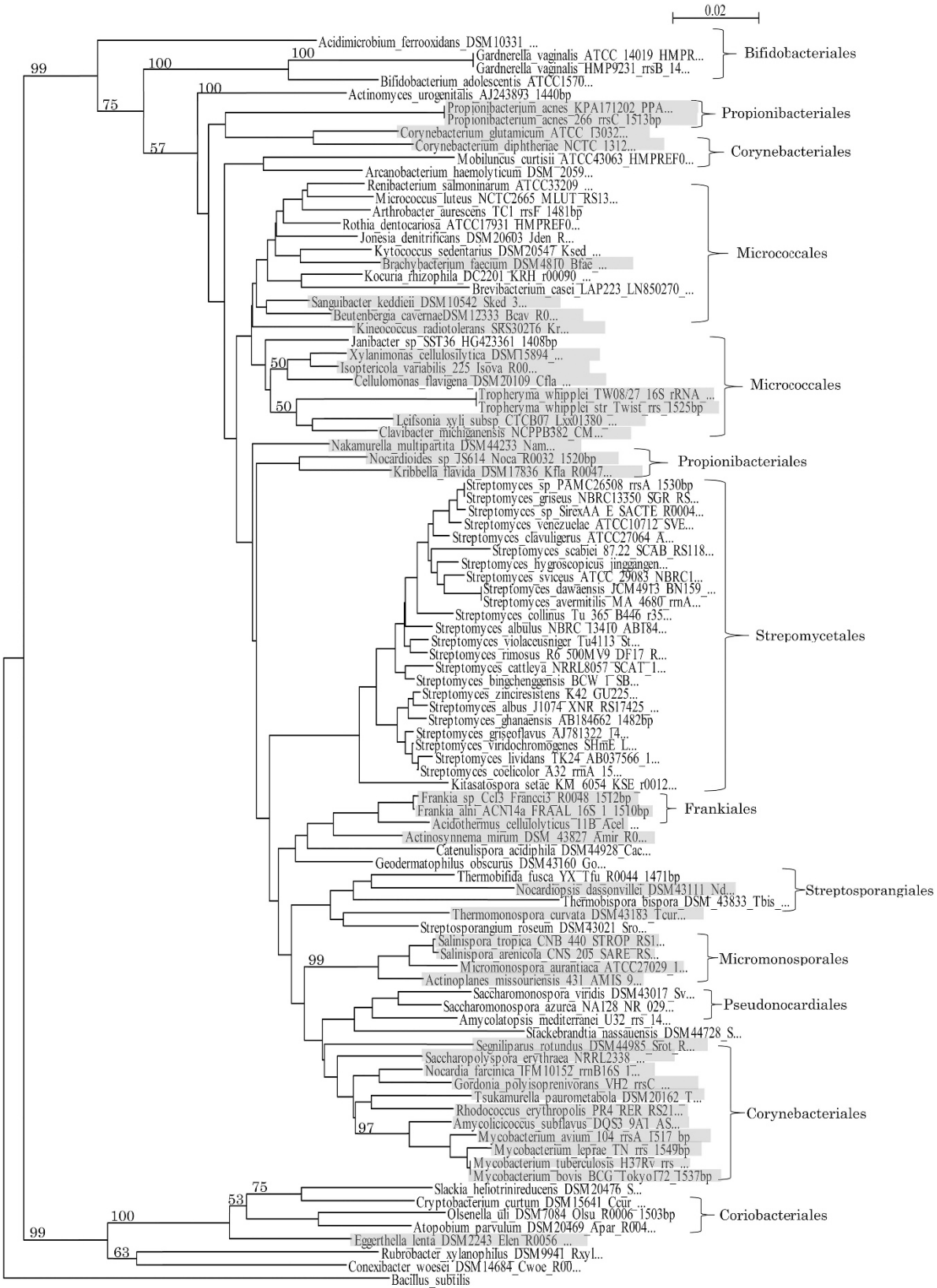


**Figure 1** Phylogenetic tree of 95 species of *Actinobacteria* on the basis of nucleotide sequences of their 16S ribosomal RNA. Bacterial species having penicillin-binding proteins (PBPs) with PASTA domain (penicillin-binding protein and serine/threonine kinase-associated domains) are marked with red. The tree was constructed by using ClustalX 2,[54] and *B. subtilis* ribosomal RNA as the outgroup. Bootstrap probabilities are indicated at the nodes. The names of actinobacterial species, strain numbers, accession numbers and number of base pairs of 16S ribosomal RNAs are listed in Supplementary Table S2. A full color version of this figure is available at the *Journal of Antibiotics* journal online.

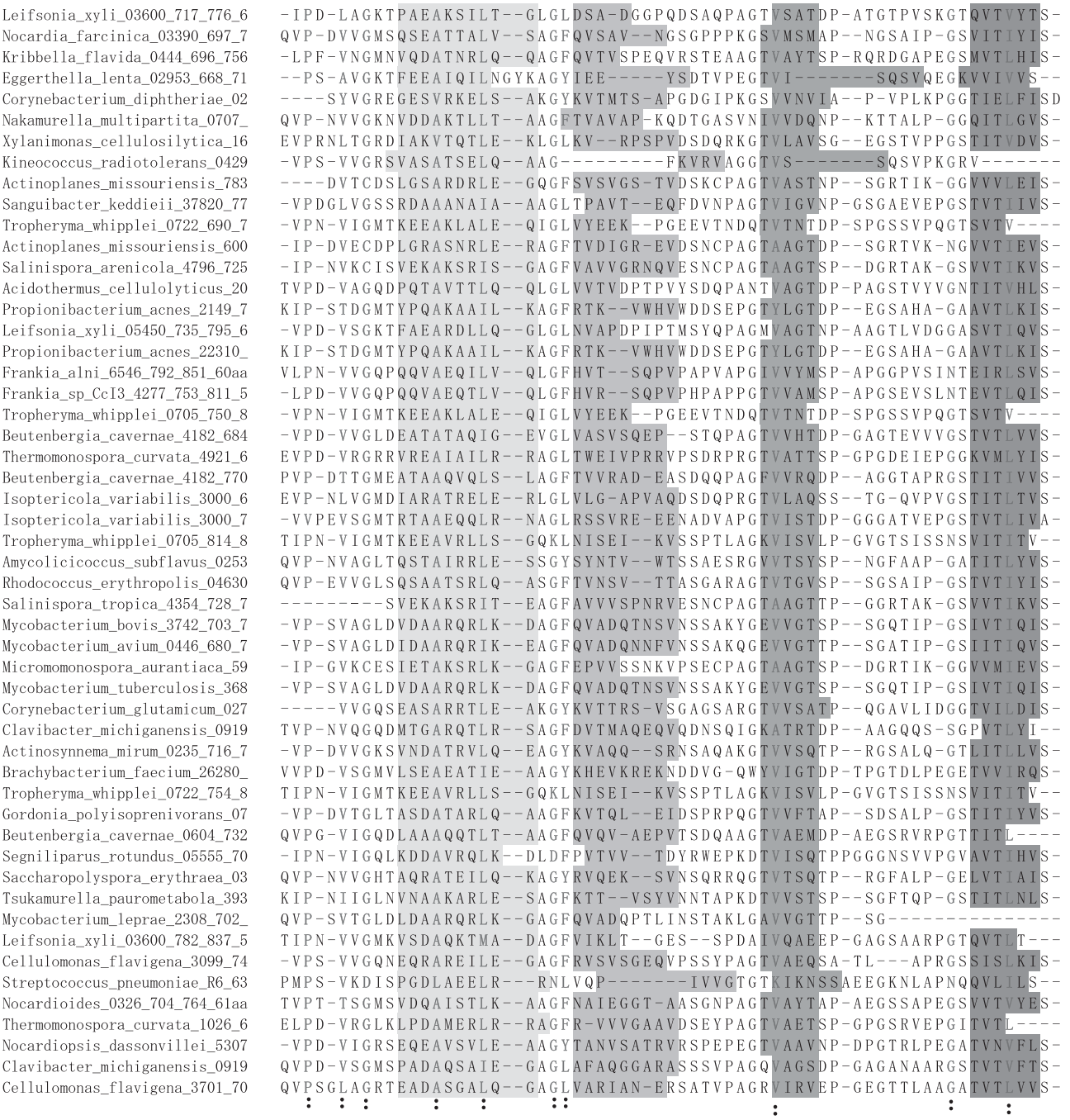


**Figure 2** Amino-acid alignment of 51 PASTA domains (penicillin-binding protein and serine/threonine kinase-associated domains) in penicillin-binding proteins (PBPs) of *Actinobacteria*. The amino-acid sequences are aligned by using ClustalX 2.[54] The conserved amino acids are marked with red and colon. α-Helix are marked with yellow, and three β-sheets are with cyan, green and magenta, respectively. A full color version of this figure is available at the *Journal of Antibiotics* journal online.

other PASTA domains, that is, from 40 (*Kineococcus radiotolerans* KRAD_0429 719–758, the numbers at the end indicate the number of amino-acid sequences of the PASTA domain, the same hereafter) to 62 (*Acidothermus cellulolyticus* ACEL_2004 703–764; *Cellulomonas flavigena* Cfla_3701 701–762; *C. michiganensis* CMM_0919 710–771 and *Thermomonospora curvata* TCUR_4921 666–727) (Supplementary Table S1). In addition, conserved amino-acid residues, shown in red and marked with colons in Figure 2, are detectable in most PASTA domains. Moreover, helical domains are also preserved (Figure 2). A phylogenetic tree was constructed on the basis of their amino-acid sequences (Figure 3). Many PASTA domains from the same order form clusters in the tree: PASTA domains within the *Micrococcales* group into four clusters with *T. whipplei* (four domains) forming one cluster, *C. michiganensis* and *C. flavigena* in a second cluster, *Xylanimonas cellulosilytica*, *I. variabilis* and *B. cavernae* (one domain each) in a third cluster and *B. cavernae* (two domains) and *C. michiganensis* (one domain) in a fourth cluster; *Micromonosporales* are represented in a single cluster comprising *Salinispora tropica*,

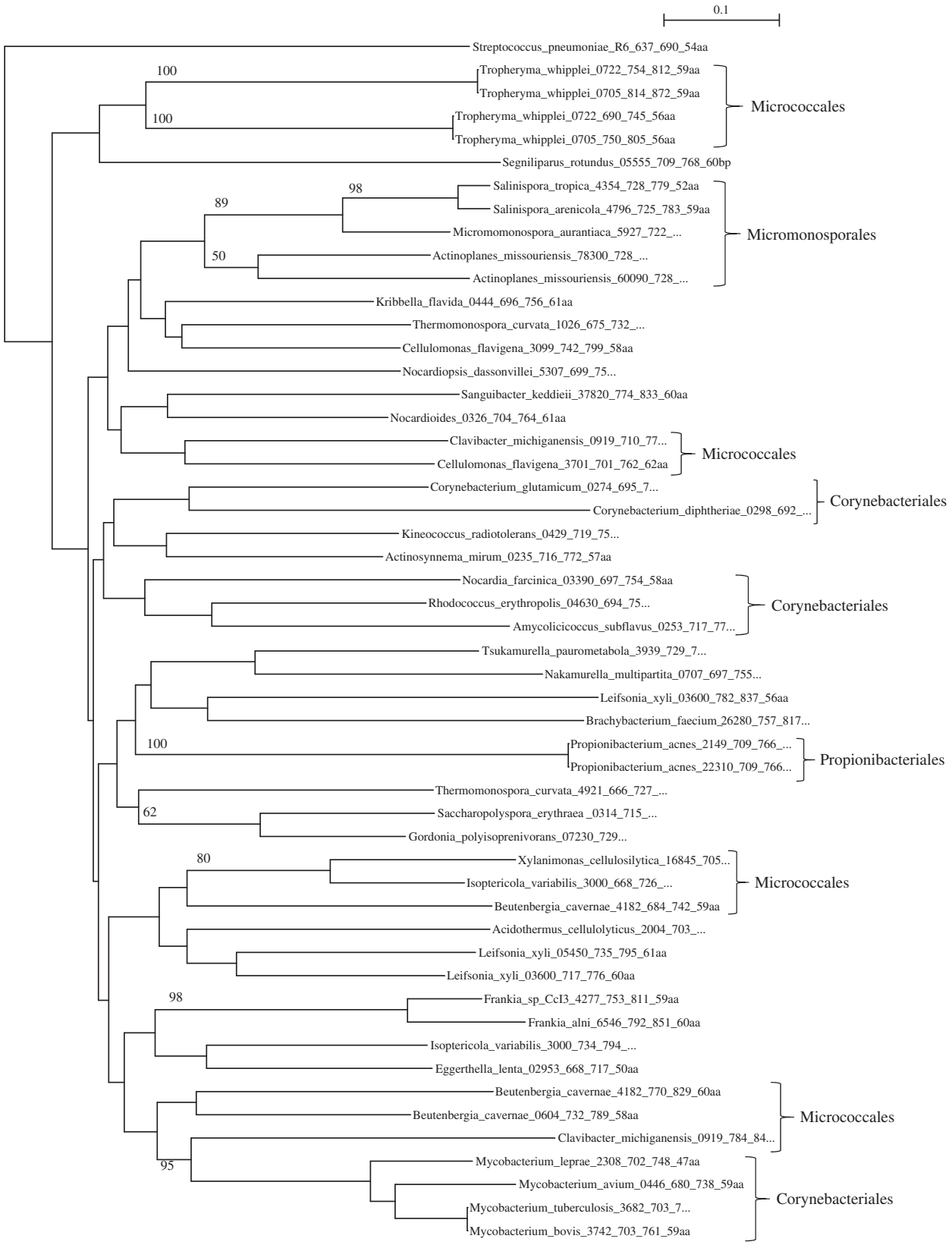


**Figure 3** Phylogenetic tree of 51 PASTA domains (penicillin-binding protein and serine/threonine kinase-associated domains) in penicillin-binding proteins (PBPs) of *Actinobacteria* on the basis of their amino-acid sequences. Bacterial species in Figure 1 are marked in circled numbers. Actinobacterial orders are indicated. The tree was constructed by using MEGA4,[55] and *S. pneumonia* R6 637–690 as the outgroup. Bootstrap probabilities are indicated at the nodes.

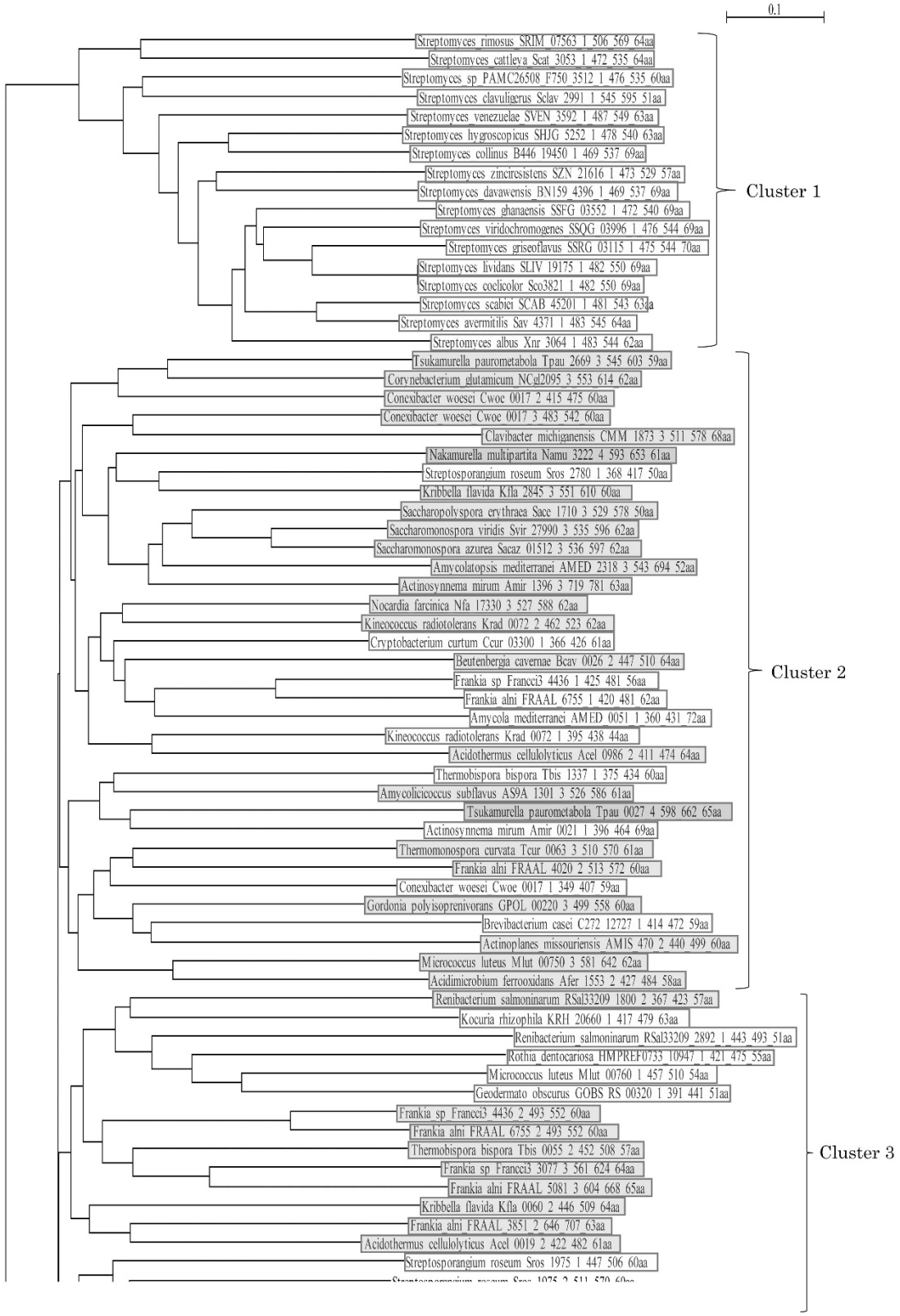


**Figure 4** Phylogenetic tree of 677 PASTA domains (penicillin-binding protein and serine/threonine kinase-associated domains) in serine/threonine kinases (STPKs) on their amino-acid sequences. These sequences are tentatively classified into 21 clusters. Position 1 PASTA domain (PASTA 1) were marked with yellow, position 2 PASTA (PASTA 2) with blue, position 3 PASTA domains (PASTA 3) with green, position 4 PASTA domains (PASTA 4) with red and PASTA 5 with white, respectively. The numbering of the PASTA domains are from N- to C-terminal regions in STPKs. The tree was constructed by using ClustalX 2.[54] A full color version of this figure is available at the *Journal of Antibiotics* journal online.

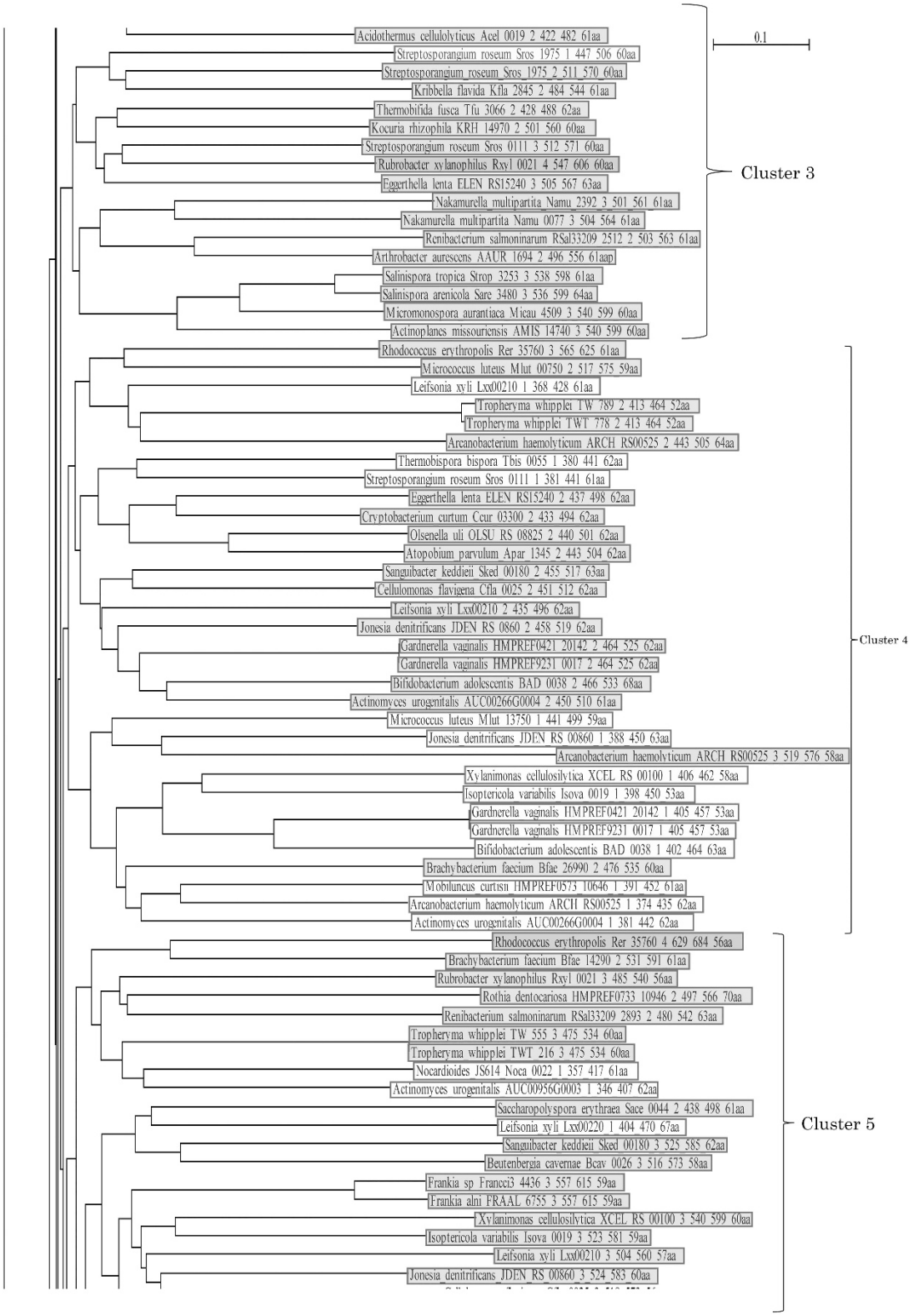


**Figure 4** Continued

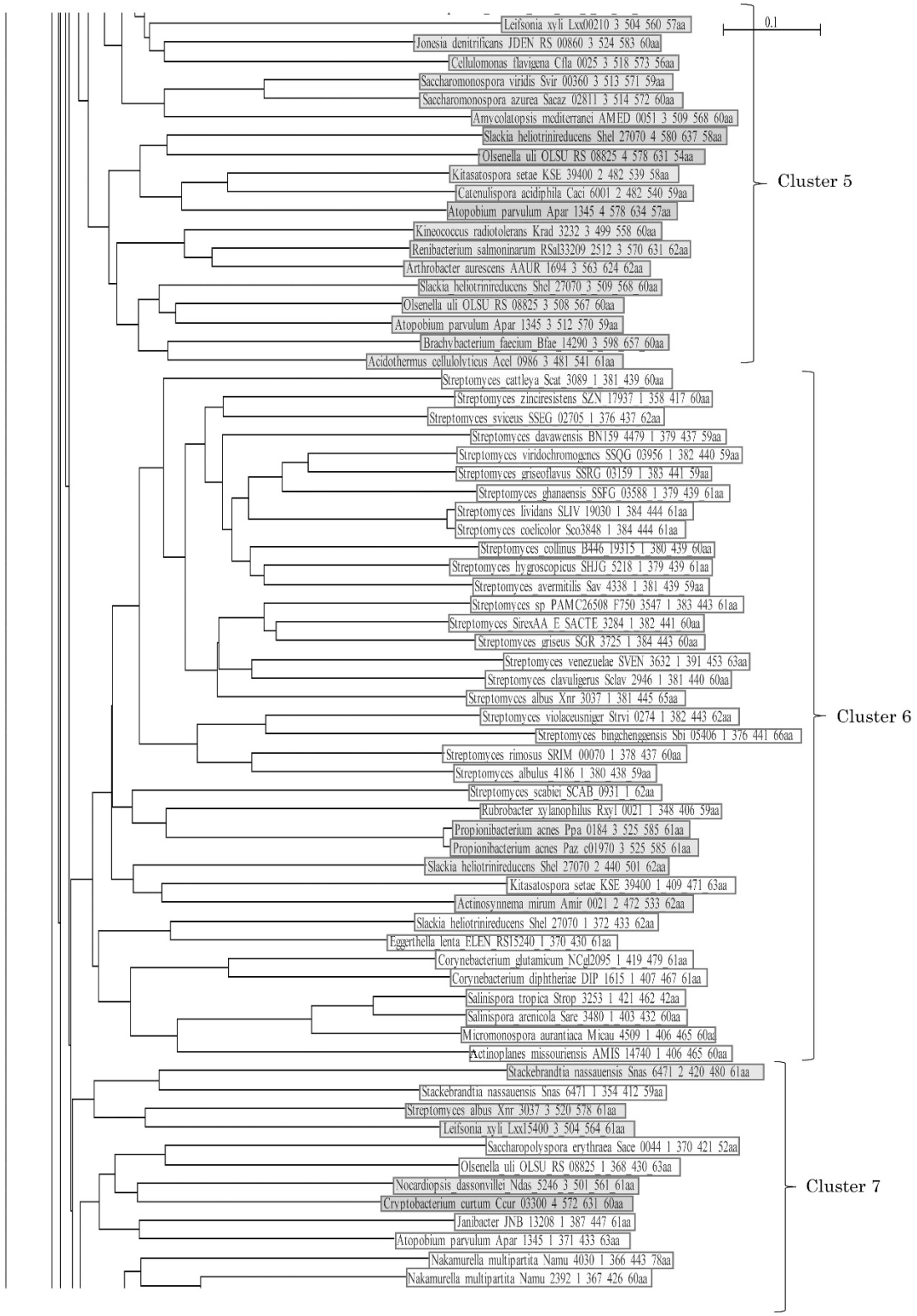


**Figure 4** Continued

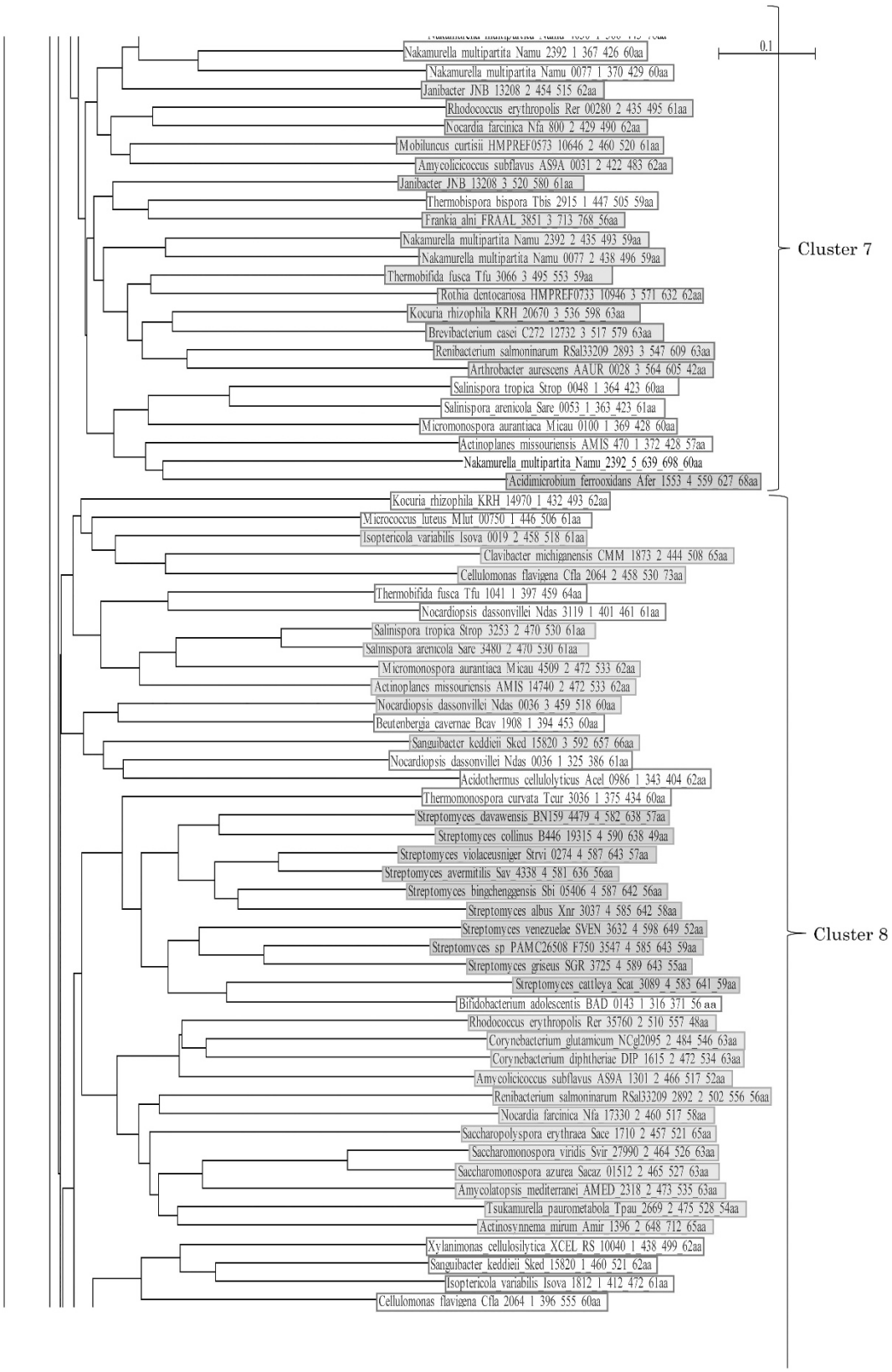


**Figure 4** Continued

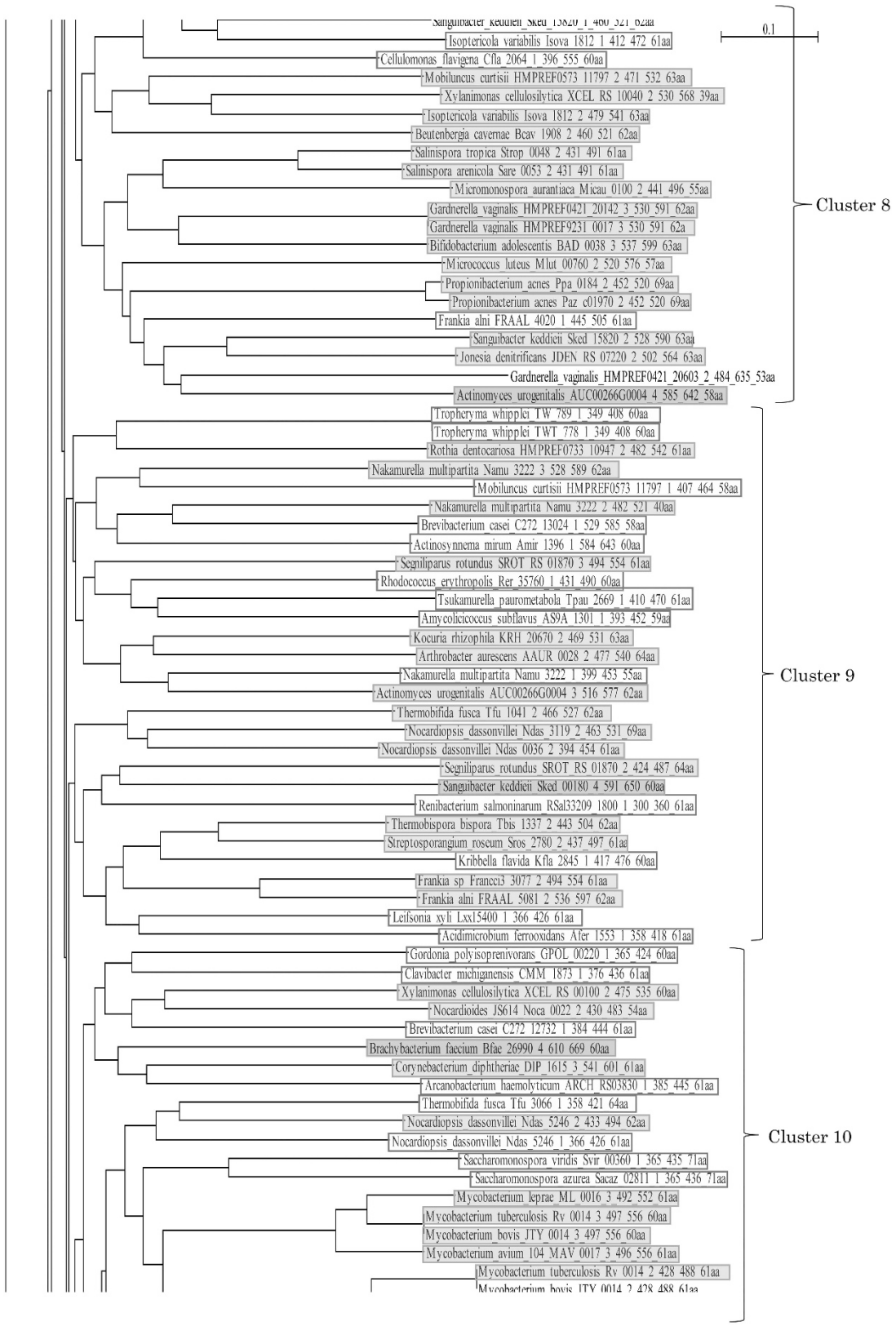


**Figure 4** Continued

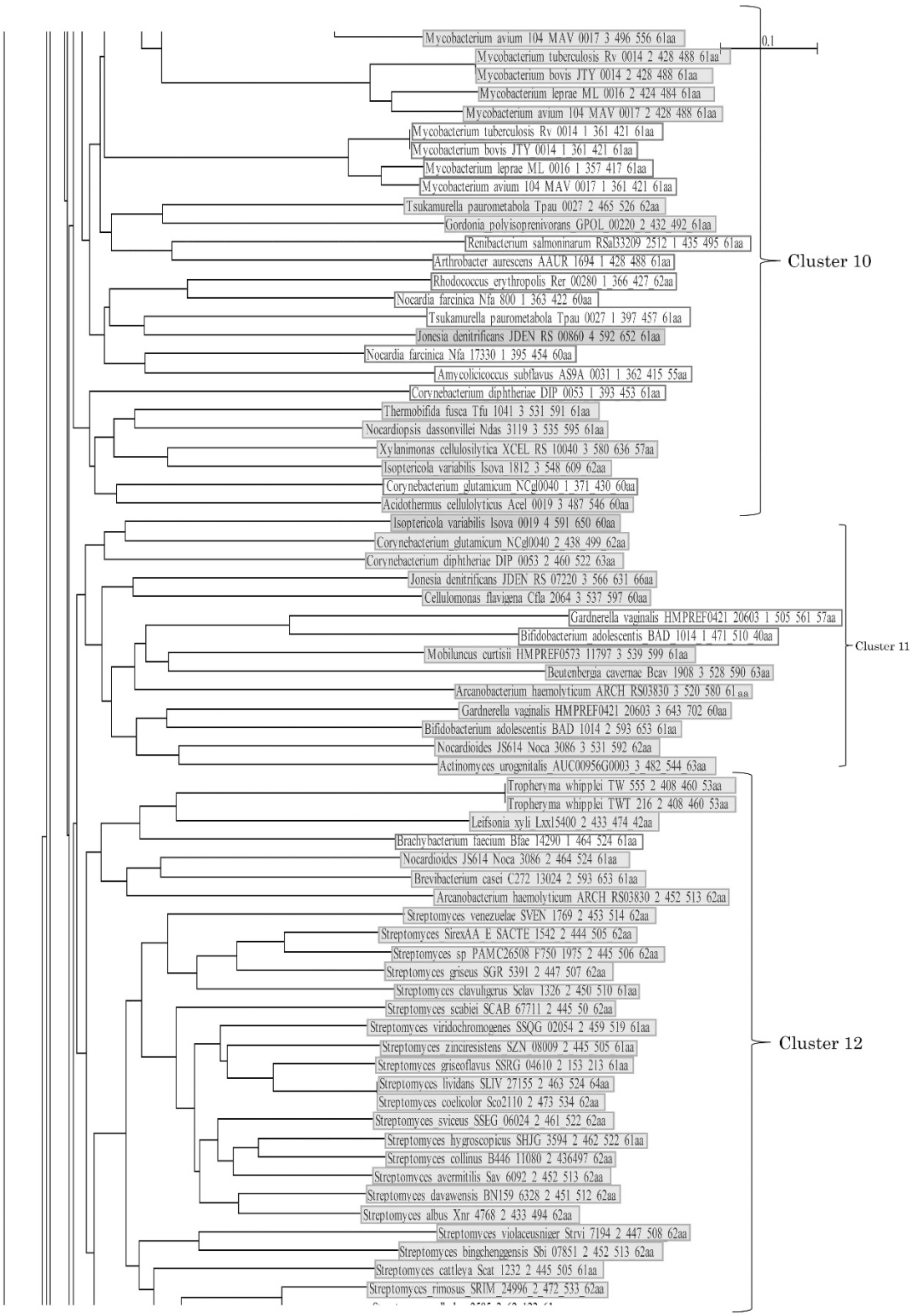


**Figure 4** Continued

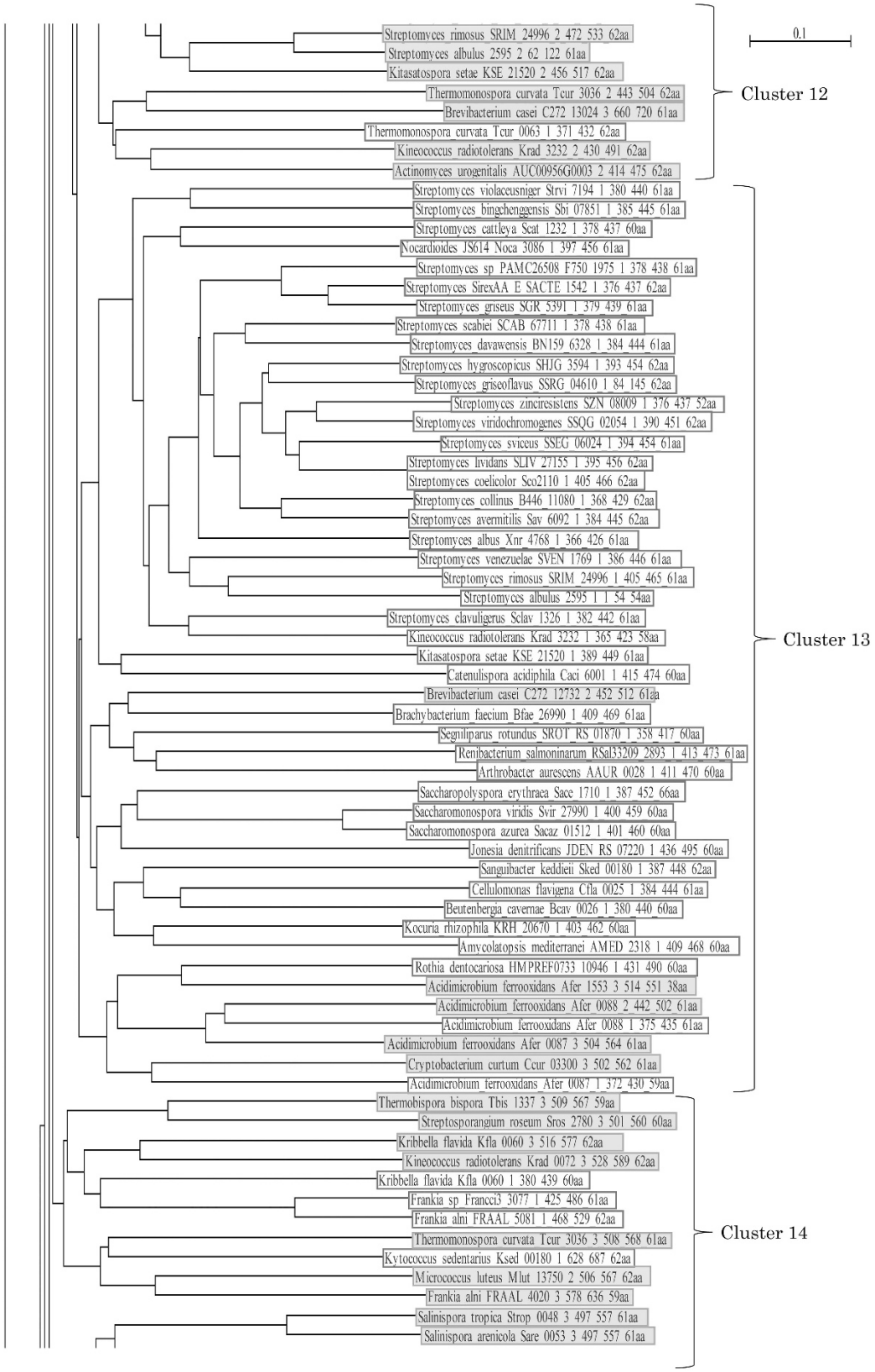


**Figure 4** Continued

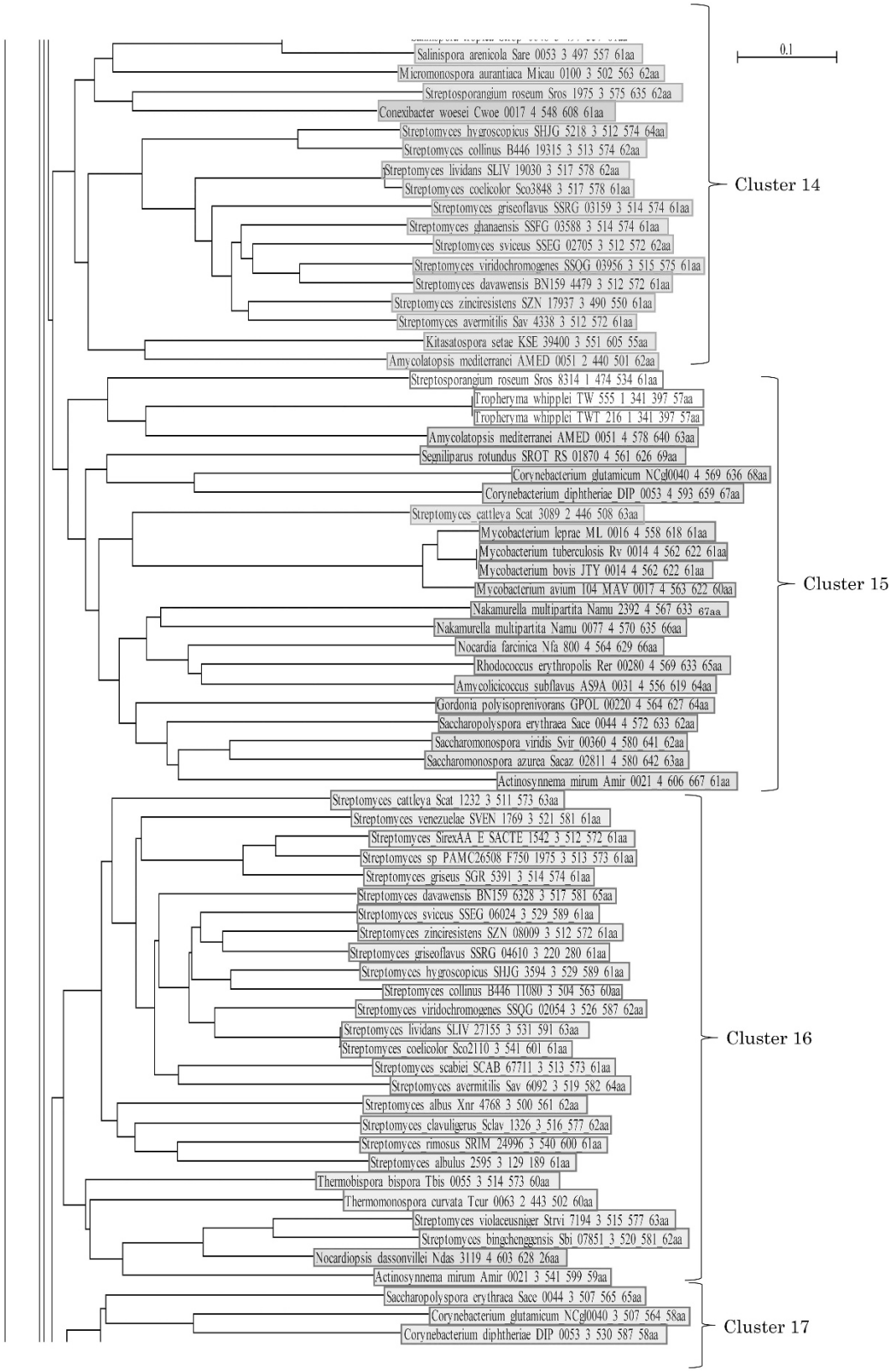


**Figure 4** Continued

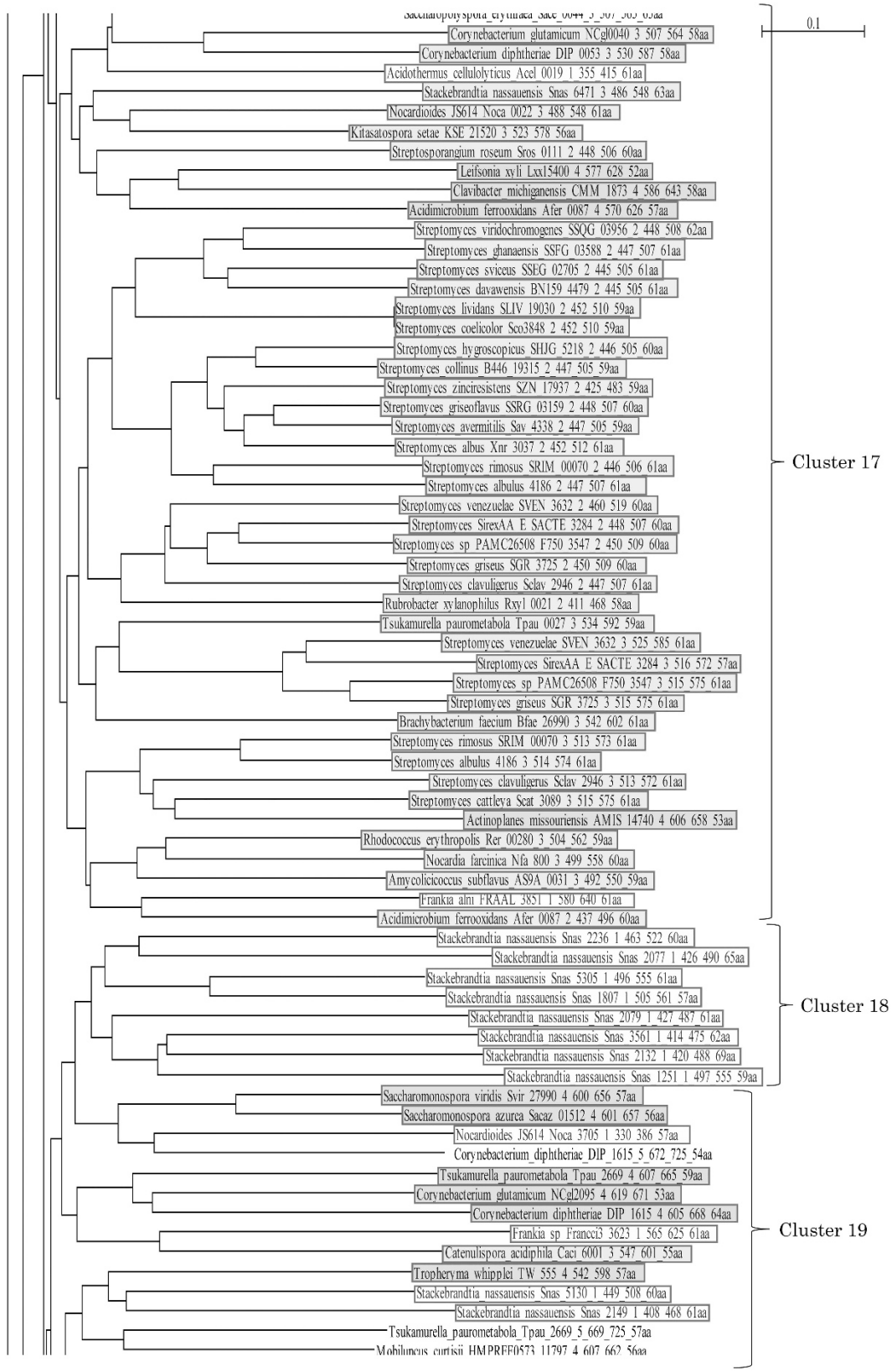


**Figure 4** Continued

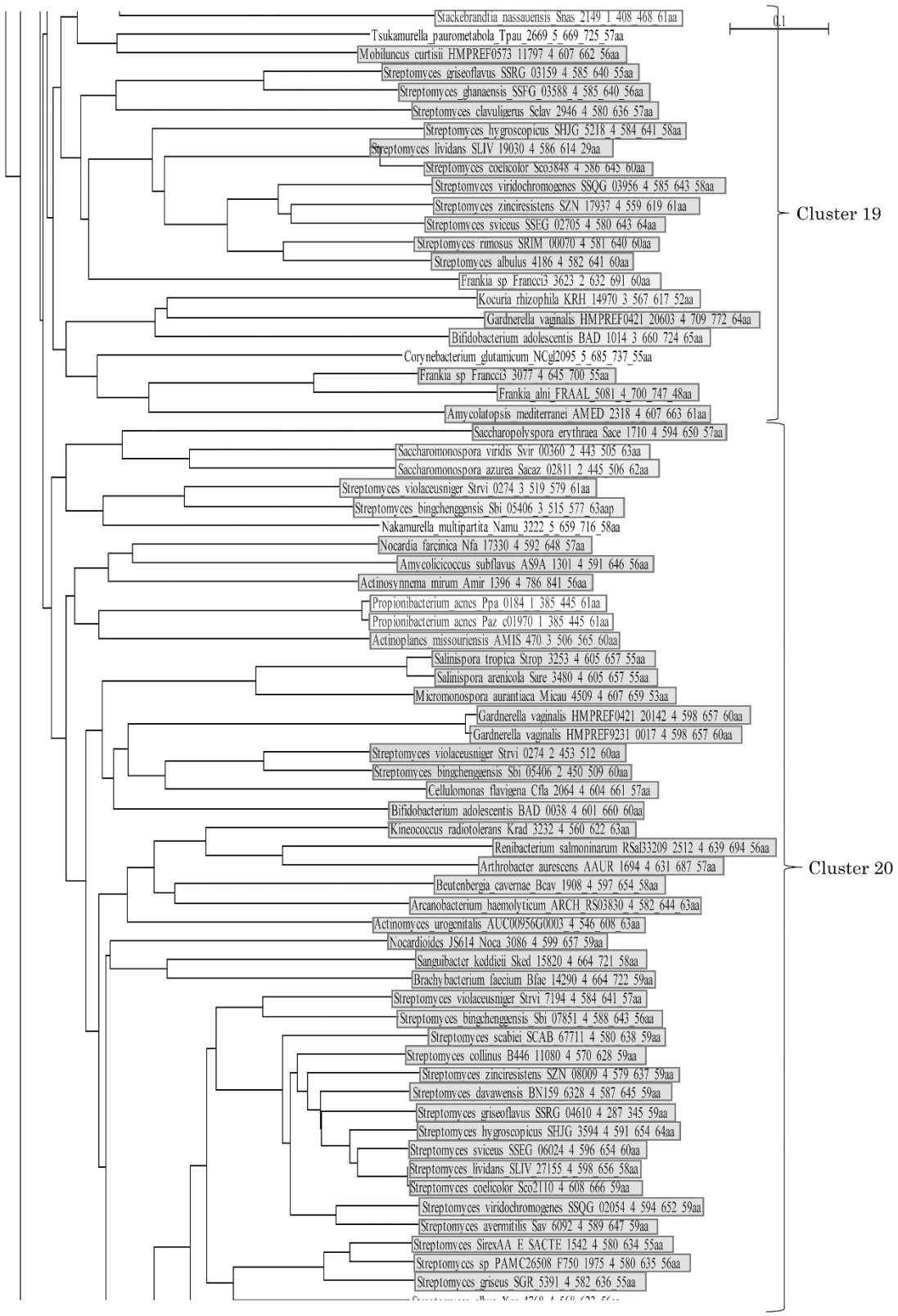


**Figure 4** Continued

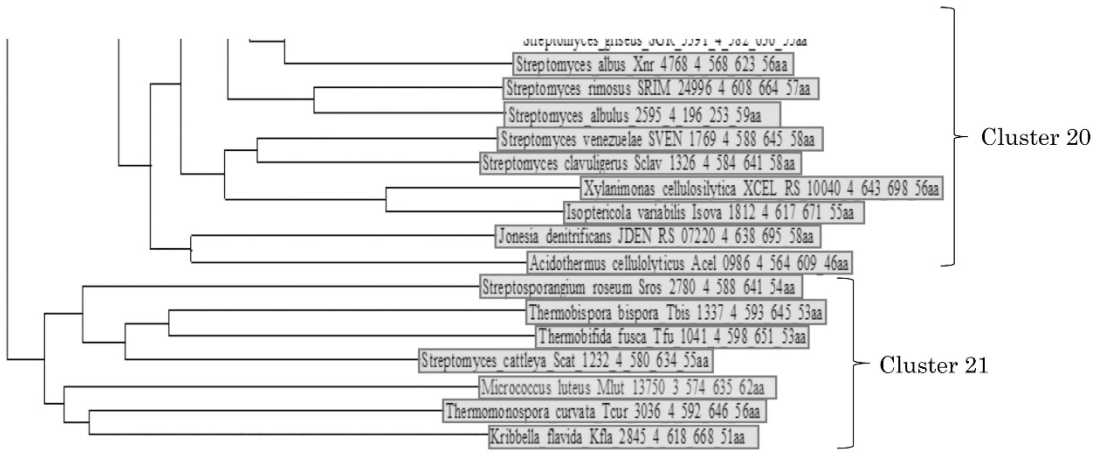


**Figure 4** Continued

*Salinispora arenicola*, *Micromonospora aurantiaca* and *A. missouriensis* (two domains). *Corynebacteriales* group into three clusters with *C. glutamicum* and *Corynebacterium diphtheria* domains in one subgroup, *Nocardia farcinica*, *Rhodococcus erythropolis* and *Amycolicicoccus subflavus* domains in a second subgroup, and *Mycobacterium leprae*, *Mycobacterium avium*, *Mycobacterium tuberculosis* and *Mycobacterium bovis* domains as the third subgroup. Finally, *P. acnes* (two domains) forms another subgroup (the order *Propionibacteriales*). Interestingly, there are numerous examples: two domains each in *C. flavigena*, *C. michiganensis*, *I. variabilis* and *T. curvata* , and three domains in *L. xyli* of domains that are found in a single species, but are located within different clusters. In addition, PBPs in *Micrococcales* and *Corynebacteriales* form several branches in the tree, respectively. Taken together with the fact that conserved amino-acid residues appear in most PASTA domains (Figure 2), these results suggest that some domains evolve independently from taxonomic classification, whereas others evolve vertically along with the species. Similar topological trees were obtained by using maximum parsimony and UPGMA (Unweighted Pair Group *Method* with Arithmetic Mean) methods.

On the other hand, a phylogenetic tree constructed by using the amino-acid sequences of PBPs excluding PASTA domain shows clusters following taxonomic classification (Supplementary Figure S1). This is also the cases with trees constructed using transglycosylase (Supplementary Figure S2) as well as transpeptidase (Supplementary Figure S3) domains, confirming that at least some PASTA domains evolve independently of taxonomy and the amino-acid residues in both transglycosylase and transpeptidase domains are more conserved than those in PASTA domains.

## DISTRIBUTION OF PASTA DOMAIN IN STPKS

In comparison with the PBPs, all bacterial species analyzed in this paper contain 1 to 11 protein kinases,[34] and each of these can have 0 to 5 PASTA domains (Table 1). That is, STPKs can be divided into two groups: PASTA-containing STPKs and those without PASTA domains. In general, however, most STPKs do not have PASTA domains. The STPKs in *S. coelicolor* SCO4423 (AfsK), SCO6681 (RamC), SCO3821 (PksC), SCO2110 (PkaF) and SCO3848 (PknB) have 0, 0, 1, 4 and 4 PASTA domains, respectively. Figure 4 shows a phylogenetic tree constructed by using amino-acid sequences of the 677 PASTA domains of STPKs in *Actinobacteria*. These PASTA domains were searched for in the STPKs of all of the *Actinobacteria* analyzed in this paper by Blast analysis. The secondary structures of these PASTA domains were analyzed by using PSIPRED software (http://bioinf.cs.ucl.ac.uk/psipred/). Similar to the STPKs with PASTA domains of other bacteria, all of the STPKs with PASTA domains in *Actinobacteria* detected in this work comprise the serine/threonine protein kinase domain in the N-terminal region, and the PASTA domains locate in the far most C-terminal region. The number of amino-acid residues in PASTA domains and the composition of the secondary structures in PASTA domains are also similar (e.g. Figure 5). When these 677 sequences are tentatively categorized into 21 clusters, PASTA domains from the same relative positions within the STPKs in *Actinobacteria* have a tendency to be located within the same clusters. For example, many PASTA domains from position 1 (PASTA 1) fall into clusters 1, 6, 9, 10 and 13, with clusters 1 and 13 being especially prominent; those from position 2 (PASTA 2) fall into clusters 3, 4, 8, 9, 12 and 17, with 4, 9, 12 and 13 being especially prominent; those from position 3 (PASTA 3) fall into clusters 2, 3, 5, 7, 10, 11, 14, 16 and 17, with 5, 11, 14 and 16 especially prominent; and those from position 4 (PASTA 4) fall into clusters 8, 15, 19, 20 and 21, with 19, 20 and 21 being prominent. This is in accordance with the suggestion of Jones and Dyson[5] that individual PASTA domains evolve position-dependently. Cluster 18 consists of only PASTA domains from *Stackebrandtia*, but these STPKs contain only one PASTA domain. STPKs, *P. acnes* Ppa_0181, *P. acnes*_Paz c01970 and *A. missouriensis* AMIS_470 from cluster 20 and *M. luteus* Mlut_13750 in cluster 21 have only three PASTA domains. Therefore, their amino-acid sequences are more similar to those of position 4 (PASTA 4) instead of position 1 (PASTA 1) or 3 (PASTA 3), although they group into cluster 20 or 21, which consists of many position 4 (PASTA 4) sequences. Similarly, position 2 PASTA (PASTA 2) and position 1 PASTA (PASTA 1) domains of *M. luteus* Mlut_13750 belong to clusters 14 and 4, respectively. Similarly, STPKs, *Stackebrandtia nassauensis* Snas_5130, *S. nassauensis* Snas_2149 and *Nocardioides* JS614 Noca_3705 group with cluster 19, although they have position 1-type PASTA (PASTA 1) domains. These STPKs possess only one PASTA domain. Although some position 2 and position 3 PASTA domains of *Streptomyces violaceusniger* (Strvi_7194 and Strvi_7194, respectively) and of *Streptomyces bingchenggensis* (Sbi_07851 and Sbi_07851, respectively) cluster together with other position 2 or position 3 PASTA domains of *Streptomyces* species (cluster 12 and cluster 16, respectively), other position 2 and position 3 PASTA domains of *S. violaceusniger* (Strvi_0274 and Strvi_0274, respectively) and of *S. bingchenggensis* (Sbi_05406 and Sbi_05406,

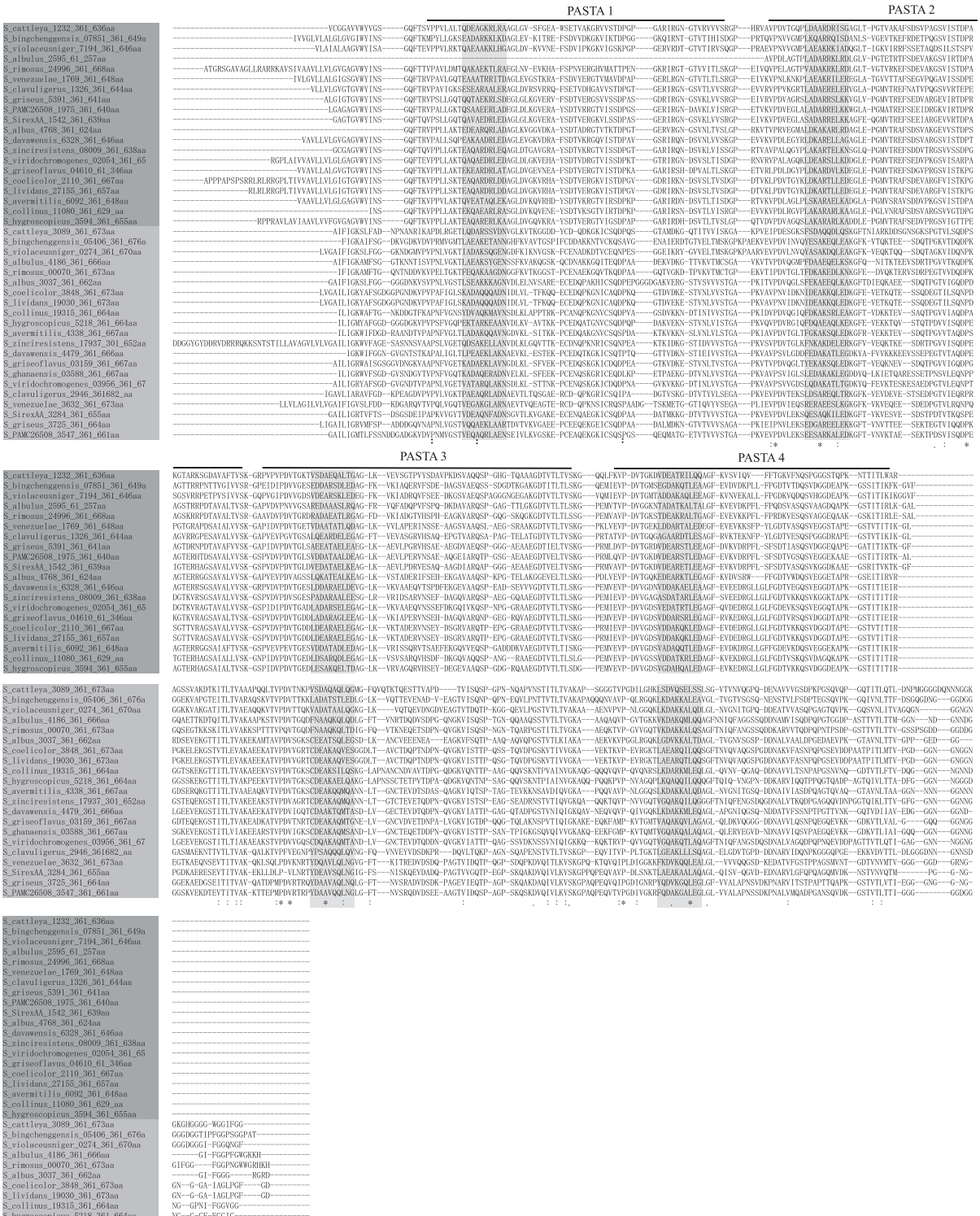


**Figure 5** Amino-acid alignment of 41 PASTA domains (penicillin-binding protein and serine/threonine kinase-associated domains) in serine/threonine kinases (STPKs) of *Streptomyces*. The amino acid are aligned by using MUSCLE.[56] Approximately four PASTA regions are indicated. Four α-helix are marked with yellow, and the two classes in Table 2 are shown in green and blue, respectively. The conserved amino-acid residues are indicated with dot, colon or star. A full color version of this figure is available at the *Journal of Antibiotics* journal online.

**Table 2 Distribution of PASTA domains of STPKs of *Streptomyces* species in the phylogenetic tree**

| *STPK* | *Position 1* | *Position 2* | *Position 3* | *Position 4* |
|---|---|---|---|---|
| *Class 1* | | | | |
| *Streptomyces albulus* 4186 | 6[a] | 17 | 17 | 19 |
| *Streptomyces albus* 3037 | 6 | 17 | 7 | 8 |
| *Streptomyces avermitilis* 4338 | 6 | 17 | 14 | 8 |
| *Streptomyces bingchenggensis* 05406 | 6 | 20 | 20 | 8 |
| *Streptomyces cattleya* 3089 | 6 | 15 | 17 | 8 |
| *Streptomyces clavuligerus* 2946 | 6 | 17 | 17 | 19 |
| *Streptomyces coelicolor* 3848 | 6 | 17 | 14 | 19 |
| *Streptomyces collinus* 19315 | 6 | 17 | 14 | 8 |
| *Streptomyces davawensis* 4479 | 6 | 17 | 14 | 8 |
| *Streptomyces ghanaensis* 03588 | 6 | 17 | 14 | 19 |
| *Streptomyces griseoflavus* 03159 | 6 | 17 | 14 | 19 |
| *Streptomyces griseus* 3725 | 6 | 17 | 17 | 8 |
| *Streptomyces hygroscopicus* 5218 | 6 | 17 | 14 | 19 |
| *Streptomyces lividans* 19030 | 6 | 17 | 14 | 19 |
| *Streptomyces rimosus* 00070 | 6 | 17 | 17 | 19 |
| *Streptomyces* PAMC26508 3547 | 6 | 17 | 17 | 8 |
| *Streptomyces* SirexAA 3284 | 6 | 17 | 17 | None[b] |
| *Streptomyces venezuelae* 3632 | 6 | 17 | 17 | 8 |
| *Streptomyces violaceusniger* 0274 | 6 | 20 | 20 | 8 |
| *Streptomyces viridochromogenes* 03956 | 6 | 17 | 14 | 19 |
| *Streptomyces zinciresistens* 17937 | 6 | 17 | 14 | 19 |
| *Class 2* | | | | |
| *Streptomyces albulus* 2595 | 13 | 12 | 16 | 20 |
| *Streptomyces albus* 4768 | 13 | 12 | 16 | 20 |
| *Streptomyces avermitilis* 6092 | 13 | 12 | 16 | 20 |
| *Streptomyces bingchenggensis* 07851 | 13 | 12 | 16 | 20 |
| *Streptomyces cattleya* 1232 | 13 | 12 | 16 | 21 |
| *Streptomyces clavuligerus* 1326 | 13 | 12 | 16 | 20 |
| *Streptomyces coelicolor* 2110 | 13 | 12 | 16 | 20 |
| *Streptomyces collinus* 11080 | 13 | 12 | 16 | 20 |
| *Streptomyces davawensis* 6328 | 13 | 12 | 16 | 20 |
| *Streptomyces griseoflavus* 04610 | 13 | 12 | 16 | 20 |
| *Streptomyces griseus* 5391 | 13 | 12 | 16 | 20 |
| *Streptomyces hygroscopicus* 3594 | 13 | 12 | 16 | 20 |
| *Streptomyces lividans* 27155 | 13 | 12 | 16 | 20 |
| *Streptomyces rimosus* 24996 | 13 | 12 | 16 | 20 |
| *Streptomyces* PAMC26508 1975 | 13 | 12 | 16 | 20 |
| *Streptomyces* SirexAA 1542 | 13 | 12 | 16 | 20 |
| *Streptomyces venezuelae* 1769 | 13 | 12 | 16 | 20 |
| *Streptomyces violaceusniger* 7194 | 13 | 12 | 16 | 20 |
| *Streptomyces viridochromogenes* 02054 | 13 | 12 | 16 | 20 |
| *Streptomyces zinciresistens* 08009 | 13 | 12 | 16 | 20 |

Abbreviations: PASTA domains, penicillin-binding protein and serine/threonine kinase-associated domains; STPK, serine/threonine kinases.
[a]Number of cluster in Figure 4.
[b]Position 4 is missing.

respectively) behave independently from other position 2 or position 3 PASTA domains of *Streptomyces* species and belong to cluster 20.

*C. diphtheria* DIP_1615, *C. glutamicum* NCgl_2095, *Nakamurella multipartite* Namu_2392, *N. multipartite* Namu_3222 and *Tsukamurella paurometabola* Tpau_2669 each have five PASTA domains. The position 5 PASTA domains of these species typically fall into cluster 19 or cluster 20, which are otherwise dominated by position 4 PASTA domains, indicating that most position 5 PASTA domains are similar to those of position 4. Only *N. multipartite* Namu_2392 forms an exception by falling into cluster 7.

The position 1 PASTA domains (PASTA 1) of *Streptomyces* cluster into clusters 1, 6 and 13 in Figure 4; position 2 domains fall into clusters 12 and 17; position 3 domains into clusters 14, 16 and 17; and position 4 domains into clusters 8, 19 and 20. Interestingly, all STPKs having position 1-type PASTA domains that cluster into cluster 1, in Figure 4, retain only one PASTA domain. The other STPKs of *Streptomyces* can be divided into two classes (Table 2): PASTA 1 of one class group with cluster 6 (see Figure 4), while those of the second class cluster into cluster 13 in Figure 4. For those STPKs in which PASTA 1 cluster to cluster 6, their PASTA 2 clusters to cluster 17 or 20, their PASTA 3 cluster to clusters 7, 14, 17 or 20 and their PASTA 4 cluster to cluster 8 or 19 (Table 2). On the other hand, when PASTA 1 clusters to cluster 13, the PASTA 2 clusters exclusively to cluster 12, PASTA 3 absolutely to cluster 16 and PASTA 4 to cluster 20 or 21 (see


**Figure 6** Comparison of the secondary and quaternary structures of serine/threonine kinases (STPKs) of *S. coelicolor* SCO3848 and *S. coelicolor* SCO2110 by using Phyre2,[37] and *M. tuberculosis* of PknB (PDB code number, 2KUI) as a template. (**a**) Comparison of the secondary structure of *S. coelicolor* SCO3848 and the template. Blue arrows indicate β-sheets, and green coils show α-helix. Light brown and brown indicate inserted and deleted residues. (**b**) Comparison of the secondary structure of *S. coelicolor* SCO2110 and the template. Symbols are the same as in **a**. (**c**) ProQ2 quality assessment of *S. coelicolor* SCO3848. ProQ2 is a model quality assessment algorithm that uses support vector machines to predict local as well as global protein models.[57] N and C terminals are indicated. (**d**) ProQ2 quality assessment of *S. coelicolor* SCO2110. (**e**) Pocket detection of *S. coelicolor* SCO3848 by using fpocket2 program.[39] Only pocket-detected regions are shown. (**f**) Pocket detection of *S. coelicolor* SCO3848 by using fpocket2 program.[39] A full color version of this figure is available at the *Journal of Antibiotics* journal online.

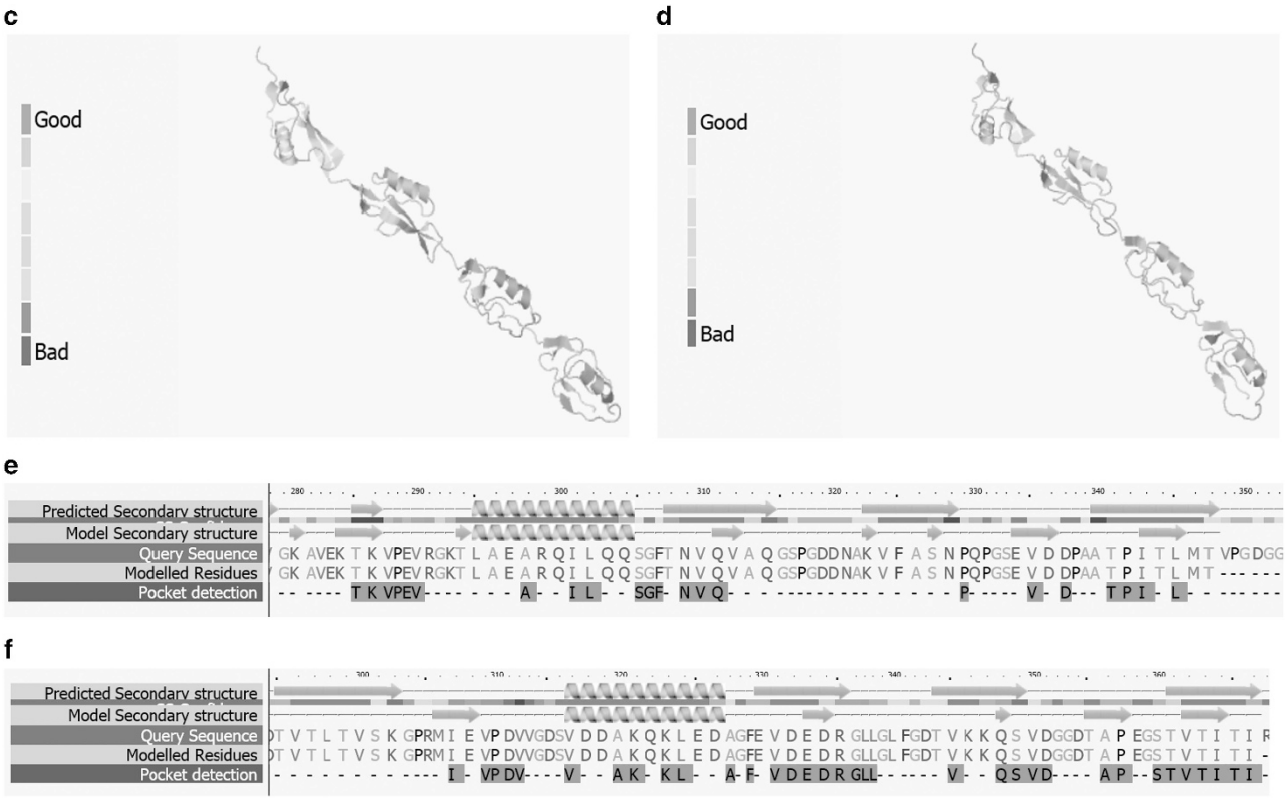


**Figure 6** Continued

Figure 4). A *S. coelicolor* STPK (SCO3848, PknB) in class 1 and that (SCO2110, PkaF) in class 2 (Table 2) are reported to be involved in the regulation of central carbon metabolism, carbon flux and of biosynthesis of the antibiotic actinorhodin and in the regulation of morphogenesis and actinorhodin production, respectively,[35,36] indicating that STPKs in class 1 and class 2 in Table 2 show different functions and those in the same group may carry out similar biological functions. These STPKs have four PASTA domains. No report has been published on the biological function of another STPK (SCO3821, PksC) in *S. coelicolor* carrying one PASTA domain.

A phylogenetic tree constructed by using the amino-acid sequences of these PASTA domains in *Streptomyces* shows that these sequences are clearly divided into two classes (Supplementary Figure S4), which correspond to the classes in Table 2. Furthermore, the amino-acid sequence alignment of the PASTA domains of these STPKs also shows two classes (Figure 5), indicating that some parts of the amino-acid sequences evolved from at least two origins. However considering that certain amino-acid residues are conserved throughout the sequences (Figure 5), one original PASTA domain is presumed to have been quadruplicated to form the four prototype PASTA domains, and then some parts were further modified in two independent directions, indicating that the PASTA domains in STPKs of *Streptomyces* species evolved position-dependently. However, when a phylogenetic tree was constructed with PASTA domains in STPKs of five *Streptomyces* species, some close interrelationships were observed between different positions (Supplementary Figure S5), that is, PASTA domains in different positions were located close together in the phylogenetic tree. For example, PASTA domains in position 1 (cluster 6 in Figure 4) and position 3 (cluster 14) of class 1 and, those in position 3 (cluster 16 in Figure 2) and position 4 (cluster 20 in Figure 4) and those in position 2 (cluster 17 in Figure 4) and position 3 (cluster 17 in Figure 4) of class 2 (Table 2). Similar phenomena are also observed in Figure 4. These results indicate that the PASTA domains in STPKs of *Streptomyces* species evolved position-independently as well and support the idea that one original PASTA domain is presumed to have been quadruplicated to form the four prototype PASTA domains.

Another interesting observation is that almost the entire amino-acid sequences of two STPKs, *Streptomyces albulus* DC74_2595 and *Streptomyces griseoflavus* SSRG_04610, are occupied by four PASTA domains. The secondary structures of these PASTA domains are similar to other PASTA domains, that is, that they are composed of one α-helix and three β-sheets. However, in the case of *S. albulus* DC74_2595 (Supplementary Figure S6A), the gene sequence for this kinase is preceded on the chromosome by another protein kinase encoding gene and the genes overlap by 3 bp. (Supplementary Figures S6A–C), suggesting that these protein kinases interact with each other. On the other hand, in the case of *S. griseoflavus* SSRG_04610 (Supplementary Figure S7A), the gene encoding this protein kinase has an atypical start codon (GAG) (Supplementary Figures S7B and C) and the preceding putative open reading frame ends (TAG) just before the putative start codon (GAG) of SSRG_04610 (Supplementary Figure S7C). This preceding putative open reading frame (Supplementary Figure S7D) apparently also encodes a protein kinase from its Blast analysis (Supplementary Figure S7E). These results suggest that *S. albulus* DC74_2595 and *S. griseoflavus* SSRG_04610 are not protein kinases but complex PASTA domains, and have their roles by acting in concert with the preceding protein kinases, respectively.

## SECONDARY AND QUARTERNARY STRUCTURE PREDICTION

When the secondary and quarternary structures of *S. coelicolor* STPKs SCO3848 and SCO2110 were analyzed using Phyre2,[37] as

representatives of these two classes, the NMR structure of the PASTA domain of *M. tuberculosis* of PknB (PDB code number, 2KUI) was selected as a best fit template in both cases. In addition, the secondary structures of both sequences fit very closely to that of template in both cases (Figures 6a and b), and the ProQ2 quality assessments[38] were also very similar (Figures 6c and d). Significant difference were only observed in Pocket detection,[39] where the pocket region in SCO2110 was more compact and extended than that in SCO3848 (Figure 6f), suggesting that the SCO2110 pocket may bind β-lactam antibiotics more efficiently. These pocket regions are within the PASTA 4 domains.

## CONCLUSION

PASTA domains are detected in PBPs as well as STPKs. They are proposed to bind β-lactam antibiotics and peptidoglycan fragments, and also to be involved in cell wall metabolism. However, the functions of PASTA domains in PBPs and those in STPKs are thought to be different. For example, the PASTA domain of *S. pneumoniae* PBP2x (2 PASTA domains) binds to β-lactam antibiotics as well as peptidoglycan fragments,[2] whereas that of *M. tuberculosis* PBP PonA2 (1 PASTA domain) does not bind β-lactam antibiotics, muropeptides, nor polymeric peptidoglycan. No report has been published on the binding property of other PBPs. However, it is interesting to know whether the PASTA domains in PBPs with two PASTA domains such as *B. cavernae* DSM 12333 BCAV_4182, *I. variabilis* 225 ISOVA_3000 bind β-lactam antibiotics and muropeptides. On the other hand, PASTA domains of STPKs such as *M. tuberculosis* PknB with four PASTA domains are proposed to bind β-lactams and peptidoglycan fragments.[9,40] While the PASTA domains of *S. pneumoniae* PBP2x form a compact conformation, those of STPKs display an extended conformation.[24] The conformational differences between PASTA domains in PBPs and those in STPKs and the number of PASTA domains may reflect their functional differences. Considering that *Actinobacteria*, especially *Streptomyces* species, show complex morphogenesis and produce various antibiotics including β-lactam antibiotics, and that STPKs with PASTA domains are involved in the signal transduction leading to morphogenesis, cell wall metabolism[17,40–44] and microbial resistance,[45–50] it is of interest to know the distribution of PASTA domains in PBPs and STPKs in *Actinobacteria*. The results showed that PASTA domains in PBPs distribute independently of taxonomy with some distribution bias. Intriguingly though, no *Streptomyces* species possess PBPs with PASTA domains. In contrast, STPKs in *Streptomyces* do contain PASTA domains, and on this basis, can be divided into three groups: one PASTA-containing STPKs, four PASTA-containing STPKs and those without PASTA domain. Four PASTA-containing STPKs can be further resolved into two classes.

Only a few reports have been published on their biological function. SCO4423 (AfsK) without PASTA domain is a global regulator of secondary metabolism in *S. coelicolor*.[51] SCO6681 (RamC), another STPK without PASTA, is reported to be involved in aerial mycelium formation and sporulation but not in secondary metabolism.[52,53] SCO2110 (PkaF), a 4 PASTA-containing STPK, is related to morphogenesis and actinorhodin biosynthesis but not to undecylprodigiosin production,[35] and SCO3848 (PknB), another 4 PASTA-containing STPK, is reported to deregulate central carbon metabolism, with carbon flux diverted to biosynthesis of actinorhodin.[36] PknB (Rv0014c) of *M. tuberculosis*, a 4 PASTA domain-containing STPK, phosphorylates Wag31 and triggers remodeling of bacterial morphology.[28] No report has been published on the function of SCO3821 (PksC), a 1 PASTA domain-containing STPK. Therefore, available data are too limited to conclude definitely the functional roles of PASTA domains in PBPs as well as in STPKs at the present time. The pocket region in SCO2110 (PkaF) was more compact and extended than that in SCO3848 (PknB) and might bind β-lactam antibiotics more efficiently. Further research is needed on PASTA domain to clarify the biological functions of PBPs and STPKs in *Actinobacteria*, especially in the field of cell wall metabolism and antibiotic resistance.

## CONFLICT OF INTEREST

The author declares no conflict of interest.

## ACKNOWLEDGEMENTS

Thanks are due to Susan Jensen of University of Alberta for critical reading of the manuscript and helpful comments.

1 Pares, S., Mouz, N., Pétillot, Y., Hakenbeck, R. & Dideberg, O. X-ray structure of Streptococcus pneumoniae PBP2x, a primary penicillin target enzyme. *Nat. Struct. Biol.* **3**, 284–289 (1996).
2 Gordon, E., Mouz, N., Duée, E. & Dideberg, O. The crystal structure of the penicillin-binding protein 2x from *Streptococcus pneumoniae* and its acyl-enzyme form: implication in drug resistance. *J. Mol. Biol.* **299**, 477–485 (2000).
3 Yeats, C., Finn, R. D. & Bateman, A. The PASTA domain: a β-lactam-binding domain. *Trends Biochem. Sci.* **27**, 438–440 (2002).
4 Maurer, P., Todorova, K., Sauerbier, J. & Hakenbeck, R. Mutations in Streptococcus pneumoniae penicillin-binding protein 2x: importance of the C-terminal penicillin-binding protein and serine/threonine kinase-associated domains for beta-lactam binding. *Microb. Drug Resist.* **18**, 314–321 (2012).
5 Jones, G. & Dyson, P. Evolution of transmembrane protein kinases implicated in coordinating remodeling of Gram-positive peptidoglycan: inside versus outside. *J. Bacteriol.* **188**, 7470–7476 (2006).
6 Peters, K. *et al. Streptococcus pneumoniae* PBP2x mid-cell localization requires the C-terminal PASTA domains and is essential for cell shape maintenance. *Mol. Microbiol.* **92**, 733–755 (2014).
7 Maestro, B. *et al.* Recognition of peptidoglycan and β-lactam antibiotics by the extracellular domain of the Ser/Thr protein kinase StkP from Streptococcus pneumoniae. *FEBS Lett.* **585**, 357–363 (2011).
8 Schweizer, I., Peters, K., Stahlmann, C., Hakenbeck, R. & Denapaite, D. Penicillin-binding protein 2x of *Streptococcus pneumoniae*: the mutation Ala707Asp within the C-terminal PASTA2 domain leads to destabilization. *Microb. Drug Resist.* **20**, 250–257 (2014).
9 Calvanese, L. *et al.* Structural and binding properties of the PASTA domain of PonA2, a key penicillin binding protein from *Mycobacterium tuberculosis*. *Biopolymers* **101**, 712–719 (2014).
10 Bukowska-Faniband, E. & Hederstedt, L. The PASTA domain of penicillin-binding protein SpoVD is dispensable for endospore cortex peptidoglycan assembly in *Bacillus subtilis*. *Microbiology* **161**, 330–340 (2015).
11 Ruggiero, A., De Simone, P., Smaldone, G., Squeglia, F. & Berisio, R. Bacterial cell division regulation by Ser/Thr kinases: a structural perspective. *Curr. Protein Pept. Sci.* **13**, 756–766 (2012).
12 Ruggiero, A. *et al.* X-ray structural studies of the entire extracellular region of the serine/threonine kinase PrkC from *Staphylococcus aureus*. *Biochem. J.* **435**, 33–41 (2011).
13 Nováková, L. *et al.* Characterization of a eukaryotic type serine/threonine protein kinase and protein phosphatase of *Streptococcus pneumoniae* and identification of kinase substrates. *FEBS J.* **272**, 1243–1254 (2005).
14 Osaki, M. *et al.* The StkP/PhpP signaling couple in *Streptococcus pneumoniae*: cellular organization and physiological characterization. *J. Bacteriol.* **191**, 4943–4950 (2009).
15 Nováková, L. *et al.* Identification of multiple substrates of the StkP Ser/Thr protein kinase in *Streptococcus pneumoniae*. *J. Bacteriol.* **192**, 3629–3638 (2010).
16 Beilharz, K. *et al.* Control of cell division in *Streptococcus pneumoniae* by the conserved Ser/Thr protein kinase StkP. *Proc. Natl Acad. Sci. USA* **109**, E905–E913 (2012).
17 Fleurie, A. *et al.* Mutational dissection of the S/T-kinase StkP reveals crucial roles in cell division of *Streptococcus pneumoniae*. *Mol. Microbiol.* **83**, 746–758 (2012).
18 Giefing, C., Jelencsics, K. E., Gelbmann, D., Senn, B. M. & Nagy, E. The pneumococcal eukaryotic-type serine/threonine protein kinase StkP co-localizes with the cell division apparatus and interacts with FtsZ *in vitro*. *Microbiology* **156**, 1697–1707 (2010).
19 Fleurie, A. *et al.* Interplay of the serine/threonine-kinase StkP and the paralogs DivIVA and GpsB in pneumococcal cell elongation and division. *PLoS Genet.* **10**, e1004275 (2014).
20 Morlot, C. *et al.* Interaction of penicillin-binding protein 2x and Ser/Thr protein kinase StkP, two key players in *Streptococcus pneumonia* R6 morphogenesis. *Mol. Microbiol.* **90**, 88–102 (2013).
21 Av-Gay, Y. & Everett, M. The eukaryotic-like Ser/Thr protein kinases of *Mycobacterium tuberculosis*. *Trends Microbiol.* **8**, 238–244 (2000).

22 Baer, C. E., Iavarone, A. T., Alber, T. & Sassetti, C. M. Biochemical and spatial coincidence in the provisional Ser/Thr protein kinase interaction network of *Mycobacterium tuberculosis*. *J. Biol. Chem.* **289**, 20422–20433 (2014).
23 Cole, S. T. *et al.* Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537–544 (1998).
24 Barthe, P., Mukamolova, G. V., Roumestand, C. & Cohen-Gonsaud, M. The structure of PknB extracellular PASTA domain from *Mycobacterium tuberculosis* suggests a ligand-dependent kinase activation. *Structure* **18**, 606–615 (2010).
25 Mir, M. *et al.* The extracytoplasmic domain of the *Mycobacterium tuberculosis* Ser/Thr kinase PknB binds specific muropeptides and is required for PknB localization. *PLoS Pathog.* **7**, e1002182 (2011).
26 Kang, C. M. *et al.* The *Mycobacterium tuberculosis* serine/threonine kinases PknA and PknB: substrate identification and regulation of cell shape. *Genes Dev.* **19**, 1692–1704 (2005).
27 Schultz, C. *et al.* Genetic and biochemical analysis of the serine/threonine protein kinases PknA, PknB, PknG and PknL of *Corynebacterium glutamicum*: evidence for non-essentiality and for phosphorylation of OdhI and FtsZ by multiple kinases. *Mol. Microbiol.* **74**, 724–741 (2009).
28 Molle, V. & Kremer, L. Division and cell envelope regulation by Ser/Thr phosphorylation: *Mycobacterium* shows the way. *Mol. Microbiol.* **75**, 1064–1077 (2012).
29 Shah, I. M., Laaberki, M. H., Popham, D. L. & Dworkin, J. A eukaryotic-like Ser/Thr kinase signals bacteria to exit dormancy in response to peptidoglycan fragments. *Cell* **135**, 486–496 (2008).
30 Chater, K. F., Biró, S., Lee, K. J., Palmer, T. & Schrempf, H. The complex extracellular biology of *Streptomyces*. *FEMS Microbiol. Rev.* **34**, 171–198 (2012).
31 Liras, P. & Martin, J. F. Gene clusters for β-lactam antibiotics and control of their expression: why have clusters evolved, and from where did they originate? *Int. Microbiol.* **9**, 9–19 (2006).
32 Ogawara, H. Penicillin-binding proteins in Actinobacteria. *J. Antibiot.* **68**, 223–245 (2015).
33 Ogawara, H., Aoyagi, N., Watanabe, M. & Urabe, H. Sequences and evolutionary analyses of eukaryotic-type protein kinases from *Streptomyces coelicolor* A3(2). *Microbiology* **145**, 3343–3352 (1999).
34 Pereira, S. F., Goss, L. & Dworkin, J. Eukaryote-like serine/threonine kinases and phosphatases in bacteria. *Microbiol. Mol. Biol. Rev.* **75**, 192–212 (2011).
35 Oh, E. A. *et al.* Characterization of the autophosphorylating kinase, PkaF, in *Streptomyces coelicolor* A3(2) M130. *Arch. Microbiol.* **193**, 845–856 (2011).
36 Jones, G, Del Sol, R., Dudley, E. & Dyson, P. Forkhead-associated proteins genetically linked to the serine/threonine kinase PknB regulate carbon flux towards antibiotic biosynthesis in *Streptomyces coelicolor*. *Microb. Biotechnol.* **4**, 263–274 (2011).
37 Kelley, L. A. & Sternberg, M. J. Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc.* **4**, 363–371 (2009).
38 Ray, A., Lindahl, E. & Wallner, B. Improved model quality assessment using ProQ2. *BMC Bioinform.* **13**, 224 (2012).
39 Le Guilloux, V., Schmidtke, P. & Tuffery, P. Fpocket: an open source platform for ligand pocket detection. *BMC Bioinform.* **10**, 168 (2009).
40 Chawla, Y. *et al.* Protein kinase B (PknB) of *Mycobacterium tuberculosis* is essential for growth of the pathogen *in vitro* as well as for survival within the host. *J. Biol. Chem.* **289**, 13858–13875 (2014).
41 Fiuza, M. *et al.* From the characterization of the four serine/threonine protein kinases (PknA/B/G/L) of *Corynebacterium glutamicum* toward the role of PknA and PknB in cell division. *J. Biol. Chem.* **283**, 18099–18112 (2008).
42 Kang, C. M., Nyayapathy, S., Lee, J. Y., Suh, J. W. & Husson, R. N. Wag31, a homologue of the cell division protein DivIVA, regulates growth, morphology and polar cell wall synthesis in mycobacteria. *Microbiology* **154**, 725–735 (2008).
43 Parikh, A., Verma, S. K., Khan, S., Prakash, B. & Nandicoori, V. K. PknB-mediated phosphorylation of a novel substrate, *N*-acetylglucosamine-1-phosphate uridyltransferase, modulates its acetyltransferase activity. *J. Mol. Biol.* **386**, 451–464 (2009).
44 Squeglia, F. *et al.* Chemical basis of peptidoglycan discrimination by PrkC, a key kinase involved in bacterial resuscitation from dormancy. *J. Am. Chem. Soc.* **133**, 20676–20679 (2011).
45 Beltramini, A. M., Mukhopadhyay, C. D. & Pancholi, V. Modulation of cell wall structure and antimicrobial susceptibility by a *Staphylococcus aureus* eukaryote-like serine/threonine kinase and phosphatase. *Infect. Immun.* **77**, 1406–1416 (2009).
46 Dias, R., Félix, D., Caniça, M. & Trombe, M. C. The highly conserved serine threonine kinase StkP of *Streptococcus pneumoniae* contributes to penicillin susceptibility independently from genes encoding penicillin-binding proteins. *BMC Microbiol.* **9**, 121 (2009).
47 Hall, C. L., Tschannen, M., Worthey, E. A. & Kristich, C. J. IreB, a Ser/Thr kinase substrate, influences antimicrobial resistance in *Enterococcus faecalis*. *Antimicrob. Agents Chemother.* **57**, 6179–6186 (2013).
48 Kristich, C. J., Wells, C. L. & Dunny, G. M. A eukaryotic-type Ser/Thr kinase in *Enterococcus faecalis* mediates antimicrobial resistance and intestinal persistence. *Proc. Natl Acad. Sci. USA* **104**, 3508–3513 (2007).
49 Tamber, S., Schwartzman, J. & Cheung, A. L. Role of PknB kinase in antibiotic resistance and virulence in community-acquired methicillin-resistant *Staphylococcus aureus* strain USA300. *Infect. Immun.* **78**, 3637–3646 (2010).
50 Pensinger, D. A. *et al.* Selective pharmacologic inhibition of a PASTA kinase increases *Listeria monocytogenes* susceptibility to β-lactam antibiotics. *Antimicrob. Agents Chemother.* **58**, 4486–4494 (2014).
51 Umeyama, T., Lee, P. C. & Horinouchi, S. Protein serine/threonine kinases in signal transduction for secondary metabolism and morphogenesis in *Streptomyces*. *Appl. Microbiol. Biotechnol.* **59**, 419–425 (2002).
52 Kodani, S. *et al.* The SapB morphogen is a lantibiotic-like peptide derived from the product of the developmental gene ramS in *Streptomyces coelicolor*. *Proc. Natl Acad. Sci. USA* **101**, 11448–11453 (2004).
53 O'Connor, T. J., Kanellis, P. & Nodwell, J. R. The ramC gene is required for morphogenesis in *Streptomyces coelicolor* and expressed in a cell type-specific manner under the direct control of RamR. *Mol. Microbiol.* **45**, 45–57 (2002).
54 Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
55 Tamura, K., Dudley, J., Nei, M. & Kumar, S. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**, 1596–15119 (2007).
56 Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
57 Capra, J. A. & Singh, M. Predicting functionally important residues from sequence conservation. *Bioinformatics* **23**, 1875–1882 (2007).

Supplementary Information accompanies the paper on The Journal of Antibiotics website (http://www.nature.com/ja)