

**Dataset S2: Representation of the sequences identified belonging to membrane proteins over the 50% threshold (Table 3), and their frequency in the “pan-surfome” of the 16 *Streptococcus pneumoniae* clinical isolates analyzed.**

Code:

| Sequence identified in # isolates | Colour |
|-----------------------------------|--------|
| 1                                 | X...X  |
| 2                                 | X...X  |
| 3                                 | X...X  |
| 4                                 | X...X  |
| 5                                 | X...X  |
| 6                                 | X...X  |
| 7                                 | X...X  |
| 8                                 | X...X  |
| 9                                 | X...X  |
| 10                                | X...X  |
| 11                                | X...X  |

**GI15902165**

**spr0121**

**surface protein pspA precursor [*Streptococcus pneumoniae* R6]**

```

MNKKKMILTS LASVAILGAG FVTSQPTFVR AEEAPVASQS KAEEKDYDAAV
KKSEAAKKHY EEAKKKAKEA QKKYDEDQKK TEDKAEAVKK ADEELQKANL
AVQKAYVEYR EAKDKASAEK KIEEAKRKQK EANKKFNEEQ AKVVPEAKEL
AATKQKAEKA KKDAEVAKEK YDKAVQEVEV EKNKILEQDA ENEKKIDVLQ
NKVADLEKGI ALYQNKVAEL NKEIARLQSD LKDAEENNVE DYIKEGLEQA
IADKKAEELAT TQQNIDKTQK DELDAELELE KVLATLDPEG KTQDELDKEA
AEDANIEALQ NKVADLENKV AELDKEVTRL QSDLKDAEEN NVEDYVKEGL
EKALTDKKVE LNNTQKALDT AQAIDTALN ELGPDGDEEE TPAPAPKPEQ
PAEQPKPAPA PQPAPAPKPE KTDDQQAEED YARRSEEEYN RLPQQQPPKA
EKPAPAPKPE QPVPAPKPGW KQENGMWYFY NTDGSMATGW LQNNGSWYYL
NSNGAMATGW LQYNGSWYYL NANGAMATGW AKVNGSWYYL NANGAMATGW
LQYNDWSWYYL NANGAMATGW LQYNGSWYYL NANGAMATGW AKVNGSWYYL
NANGAMATGW LQYNGSWYYL NANGAMATGW AKVNGSWYYL NANGSMATGW
VKDGDTSWYYL EASGAMKASQ WFKVSDKWYY VNGLGALAVN TTVDGYKVNA
NGEWV

```

**GI15902373**

**spr0329**

**penicillin-binding protein 1A [*Streptococcus pneumoniae* R6]**

```

MNKPTILRLI KYLSISFLSL VIAAIVLGGG VFFYYVSKAP SLSESCLVAT
TSSKIYDNKN QLIADLGSER RVNAQANDIP TDLVKAIVSI EDHRFFDHRG
IDTIRILGAF LRNLQSNLQ GGSALTQQLI KLTYFSTSTS DQTISRKAQE
AWLAIQLEQK ATKQEILTYI INKVYMSNGN YGMQTAAQNY YGKDLNLSL

```

PQLALLAGMP QAPNQYDPYS HPEAAQDRRN LVLSEMKNQG YISAEQYEKA  
 VNTPTIDGLQ SLKSASNYPA YMDNYLKEVI NQVEEETGYN LLTTGMDVYT  
 NVDQEAQKHL WDIYNTDEYV AYPDDELQVA STIVDVSNGK VIAQLGARHQ  
 SSVNSFGINQ AVETNRDWGS TMKPITDYAP ALEYGVYEST ATIVHDEPYN  
 YPGTNTPVYN WDRGYFGNIT LQYALQQSRN VPAVETLNKV GLNRAKTFLN  
 GLGIDYPSIH YSNAISSNTT ESDKKYGASS EKMAAAAYAAF ANGGTYKPM  
 YIHKVVFSDG SEKEFSNVGT RAMKETTAYM MTDMMKTVLS YGTGRNAYLA  
 WLPQAGKTGT SNYTDEEIEH HIKTSQFVAP DELFAGYTRK YSMAVWTGYS  
 NRLTPLVGNG LTVAAKVYRS MMTYLSEGSN PEDWNIPEGL YRNGEFVFKN  
 GARSTWSSPA PQQPPSTESS SSSSDSSTSQ SSSTTPSTNN STTTNPNNT  
 QQSNTTPDQQ NQNPPQAQP

**GI15902378**

**spr0334**

**hypothetical protein spr0334 [Streptococcus pneumoniae R6]**

MSKRRNRHK KEAQEPQFDF DEAKEELTVGQ AIRKNEEVEA GVLPEDESILD  
 KYVKQHRDEI EADKFATRQY KKEEFVETQS LDDLIQEMRE AVEKSEASSE  
 EVPSSDILL PLPLDDEEQG LDPLLLDDEN PTEMTEEVVEE EQNLSRLDQE  
 DSEKSKKGF ILTVLALVSV IICVSAYYVY RQVARSTKEI ETSQSTTANQ  
 SDVDDFNTLY DAFYTNSNKT ALKNSQFDKL SQLKTLLDKL EGSREHTLAK  
 SKYDSLATQI KAIQDVNAQF EKPAIVDGVL DTNAKAKSDA KFTDIKTGNT  
 ELDKVLDKAI SLGKSQQTST SSSSSQTS SSSSQASSNT TSEPKPSSSN  
 ETRSSRSEVN MGLSSAGVAV QRSASRVAYN QSAIDDSNNS AWFADGVLE  
 QILATRSRSG YITGDQYILE RVNIVNGNGY YNLYKPDGTY LFTLNCKTGY  
 FVGNAGHAD DLDY

**GI15902951**

**spr0907**

**pneumococcal histidine triad protein D precursor [Streptococcus pneumoniae R6]**

MKINKKYLAG SVAVLALSVC SYELGRHQAG QVKKESNRVS YIDGDQAGQK  
 AENLTPDEVK KREGINAEQI VIKITDQGYV TSHGDHYHY NGKVPYDAII  
 SEELMKDPN YQLKSDIVN EIKGGYVIKV DGKYYVYLKD AAHADNIRTK  
 EEIKRQKQER SHNHNSRADN AVAAARAQGR YTTDDGYIFN ASDIEDTGD  
 AYIVPHGDHY HYIPKSDLSA SELAAAQAYW NGKQGSRPSS SSSHNANPAQ  
 PRLSENHNLVTPTYHQNG ENISSLLREL YAKPLSERHV ESDGLIFDPA  
 QITSRTANGV AVPHGDHYF IPYSQLSPL EKLARIIPLR YRSNHWVPDS  
 RPEQSPQST PEPSPSPQPA PNPQPAPSNP IDEKLVKEAV RKVGDGYVFE  
 ENGVPRYIPA KDLAETAAG IDSKLAKQES LSHKLGAKKT DLPSSDREFY  
 NKAYDLLARI HQDLLDNKGR QVDFEALDNL LERLKDVSSE KVKLVDDILA  
 FLAPIRHPER LGKPNQAITY TDDEIQVAKL AGKYTTEDGY IFDPRDITSD  
 EGDAYVTPHM THSHWIKKDS LSEAERAAAQ AYAKEKGLTP PSTDHQDSGN  
 TEAKGAEAIY NRVKAACKVP LDRMPYNLQY TVEVKNGLI IPHYDHYHNI  
 KFEWFDEGLY EAPKGYSLED LLATVKYYVE HPNERPHSDN GFGNASDHVQ  
 RNKNGQADTN QTEKPNEEKP QTEKPEEDKE HDEVSEPTHP ESDEKENHVG  
 LNPSADNLYK PSTDTEETEE EAEDTTDEAE IPQVEHSVIN AKIAEAEALL  
 EKVTDSIRQ NAVETLTGLK SLLLLGTKDN NTISAEVDSL LALLKESQPT  
 PIQ

GI15903413

spr1370

hypothetical protein spr1370 [Streptococcus pneumoniae R6]

MSEKSREEEK LSFKEQILRD LEKVKGYDEV LKEDAVVRT PANEPSAEEL  
MADSLSTVEE IMRKAPTVPPT HPSQGVPAASP ADEIQRETPTG VPSHPSQDVP  
SSPAEESGSR PGPGPVRPKK LEREYNETPT RVAVSYTAE KKAEQAGPET  
PTPATETVDI IRDTSRRSRR EGAKPAKPKK EKKSHVKAFV ISFLVFLALL  
SAGGYFGYQY VLDSLLPIDA NSKKYVTVGI PEGSNVQEIG TTLEKAGLVK  
HGLIFSFIYAK YKNYTDLKAG YYNLQKSMST EDLLKELQKG GTDEPQEPVL  
ATLTIPEGYT LDQIAQTVGQ LQGDFKESLT AEAFLAKVQD ETFISQAVAK  
YPTLLES LPV KDSGARYRLE GYLFPATYSI KESTTIESLI DEMLAAMDKN  
LSLYYSTIKS KNLTVNELLT IASLVEKEGA KTEDRKLIAG VFYNR LNRDM  
PLQSNIAILY AQGKLGQNIS LAEDVAIDTN IDSPYVYKN VGLMPGPVDS  
PSLDAIESSI NQTKSDNLYF VADVTEGK VY YANNQEDHDR NVAEHVNSKL  
N

GI15903916

spr1875

hypothetical protein spr1875 [Streptococcus pneumoniae R6]

MKKRMLLAST VALSFAPVLA TQAEVLWTA RSVEQIQNDL TKTDNKTSYT  
VQYGD TLSTI AEALGVDVTV LANLNKITNM DLIFPETVLT TTVNEAEEVT  
EVEIQTPQAD SSEEVTTATA DLTTNQVTVD DQTVQVADLS QPIAEAPKEV  
ASSSEVTKTV IASEEVAPST GTSVPEEQTA ETSSAVAEAA PQETTPAEKQ  
ETQTSPPQAAS AVEATTTTSE AKEVASSNGA TAAVSTYQPE ETKIISTTYE  
APAA PDYAGL AVAKSENA GL QPQTAAFKEE IANLFGITSF SGYRPGDSD  
HGKGLAIDFM VPERSELGDK IAEYAIQNMA SRGISYIIWK QRFYAPFDSK  
YGPANTWNPM PDRGSVTENH YDHVHVS MNG

GI15904086

spr2045

serine protease [Streptococcus pneumoniae R6]

MEANMKHLKT FYKKWFQLLV VIVISFFSGA LGSFSITQLT QKSSVNNSNN  
NSTITQTAYK NENSTTQAVN KVKDAVSVI TYSANRQNSV FGNDTDTDS  
QRISSESGSV IYKNDKEAY IVTNNHVI NG ASKVDIRLSD GTKVPGEIVG  
ADTFSDIAVV KISSEKVTTV AEFGDSSKLT VGETAIAIGS PLGSEYANTV  
TQGIVSSLNR NVSLKSEDGQ AISTKAIQTD TAINPGNSGG PLINIQQQVI  
GITSSKIATN GGTSVEGLGF AIPANDAINI IEQLEKNGKV TRPALGIQMV  
NLSNVSTSDI RRLNIPSNVT SGVIVRSVQS NMPANGHLEK YD VITKVDDK  
EIASSTDLQS ALYNHSIGDT IKITYYRNGK EETTSIKLNK SSGDLES

GI15900016

SP\_0071

**immunoglobulin A1 protease [Streptococcus pneumoniae TIGR4]**

MSRKSIGEKR HSFMRKLSV GLVSVTVSSF FLMSQGIQSV SADNMESPIH  
YKYMTEGKLT DEEKSLLEVA LPQLAEESDD TYYLVSRSQQ FLPNTGFNPT  
VGTFLFTAGL SLLVLLVSKR ENGKKRLVHF LLLTSMGVQL LPASAFGLTS  
QILSAYNSQL SIGVGEHLPE PLKIEGYQYI GYIKTKKQDN TELSRTVDGK  
YSAQRDSQPN STKTSDEVVHS ADLEWNQGGG KVSLOGEASG DDGLSEKSSI  
AADNLSSNDS FASQVEQNPD HKGESVVRPT VPEQGNPVA TTVQSAEEEV  
LATTNDRPEY KLPLETKGTQ EPGHEGEAAV REDLPVYTKP LETKGTQGGP  
HEGEAAVREE EPAYTEPLAT KGTQEPGHEG KATVREETLE YTEPVATKGT  
QEPEHEGEAA VEEELPALEV TTRNRTEIQN IPYTTEEIQD PTLKNNRRI  
ERQQQAGTRT IQYEDYIVNG NVVETKEVSR TEVAPVNEVV KVGTLVKVKP  
TVEITNLTKV ENKKSITVSY NLIDTTSAYV SAKTQVFHGD KLVKEVDIEN  
PAKEQVISGL DYYTPYTVKT HLTYNLGENN EENTETSTQD FQLEYKKIEI  
KDIDSVELYG KENDRYRRL SLSEAPTDTA KYFVKVKS DR FKEMYPVKS  
ITENTDGTYK VTVAVDQLVE EGTGDKYKDDY TFTVAKSKAE QPGVYTSFKQ  
LVTAMQSNLS GVYTLASDMT ADEVSLGDKQ TSYLTGAFTG SLIGSDGTKS  
YAIYDLKKPL FDTLNGATVR DLDIKTVSAD SKENVAALAK AANSANINNV  
AVEGKISGAK SVAGLVSAT NTVIENSSFT GKLIANHQDS NKNDTGGIVG  
NITGNSSRVN KVRVDALIST NARNNNQTAG GIVGRLENGA LISNSVATGE  
IRNGQGYSRV GGIVGSTWQN GRVNNVSVN DVG DGYVITG DQYAAADVKN  
ASTSVDNRKA DRFATKLSKD QIDAKVADYG ITVTLDDTGQ DLKRNLRVD  
YTRLNKAEEA RRVAYSNIK LMPFYNKDLV VHYGNKVATT DKLYTTELLD  
VVPMKDDEVV TDINNKNSI NKVMLHFKDN TVEYLDVTFK ENFINSQVIE  
YNVTGKEYIF TPEAFVSDYT AITNNVLS DL QNVTLNSEAT KKVLGAANDA  
ALDNLYLDRQ FEEVKANIAE HLRKVLAMDK SINTTGDGVV EYVSEKIKNN  
KEAFMLGLTY MNRWYDINYG KMNTKDLSTY KFDENGNET STLDTIVALG  
NSGLDNLRAS NTVGLYANKL ASVKGEDSVF DFVEAYRKLFLPNKTNNEWF  
KENTKAYIVE MKSDIAEVRE KQESPTADRK YSLGVYDRIS APSWGHKSML  
LPLTLPEES VYISSNMSTL AFGSYERYRD SVDGVILSGD ALRTYVRNRV  
DIAAKRHRDH YDIWYNLLDS ASKEKLFERSV IVYDGFNVKD ETGRTYWARL  
TDKNIGSIKE FFGPVGKWYE YNSSAGAYAN GSLTHFVLDR LLDAYGTSVY  
THEMVHNSDS AIYFEGNGRR EGLGAELYAL GLLQSVDSVN SHILALNTLY  
KAEKDDLNRH HTYNPVERFD SDEALQSYMH GSYDVMYTLD AMEAKAILAQ  
NNDVKKKWF R KIENYYVRDT RHNKDTHAGN KVRPLTDEEV ANLTSLSLI  
DNDIINRRSY DDSREYKRNG YYTISMFSVP YAALSNSKGA PGDIMFRKIA  
YELLAEKGYH KGFLPYVSNQ YGAEAFASGS KTFSSWHGRD VALVTDDLVF  
KKVFNGEYSS WADFKKAMFK QRIDKQDNLK PITIQYELGN PNSTKEVTIT  
TAAQMQLIN EAAAKDITNI DRATSHTPAS WVHLLKQKIY NAYLRTTDDF  
RNSIYK

GI15902056

spr0012

**cell division protein FtsH [Streptococcus pneumoniae R6]**

MKKQNNGLIK NPFLWLLFIF FLVTGFQYFY SGNNSSGGSQQ INYTELVQEI  
TDGNVKELTY QPNGSVIEVS GYKPNPKTSK EGTGIQFFTP SVTKVEKFTS  
TILPADTTVS ELQKLATDHK AEVTVKHES SGIWINLLVS IVPFGILFFF  
LFSMMGNMGG GNGRNPMSFG RSKAKAANKE DIKVRFS DVA GAEEEEQELV

EVVEFLKDPK RFTKLGARIP AGVLLEGPPG TGKTLAKAV AGEAGVPFFS  
 ISGSDVEMF VVGASRVRS LFEDAKKAAP AIIFIDEIDA VGRQGVGLG  
 GGNDEREQTL NQLLIEMDGF ENEGIIVIA ATNRSDVLDP ALLRPGRFDR  
 KVLVGRPDVK GREAILKVHA KNKPLAEDVD LKLVAQQTG FVGADLENVL  
 NEAALVAAR NKSIIASDI DEAEDRV IAG PSKKDKTVSQ KERELVAYHE  
 AGHTIVGLVL SNARVVHKVT IVPGRAGGY MIALPKEDQM LLSKEDMKEQ  
 LAGLMGGRVA EEIIFNVQTT GASNDFEQAT QMARAMVTEY GMSEKLGVPVQ  
 YEGNHAMLGA QSPQKSISEQ TAYEIDEEVR SLLNEARNKA AEIIQSNRET  
 HKLIAEALLK YETLDSTQIKALYETG KMPE AVEEESHALS YDEVKSKMND  
 EK

**GI15902130**

**spr0086**

**hypothetical protein spr0086 [Streptococcus pneumoniae R6]**

MKQEFESND FVKTTSKNKP EEQAQEVADK AEETIADLDT PIEKNTOLEE  
 EVSQAEVELE SQQEEKIETP EDSEAR TKIE EKKASNSTEE EPDLSKETEK  
 VTIAEESQEA LPQQKATTKE PLLISKLES PYIPDQAPKS RDKWKEQVLD  
 FWSWLVEAIK SPTSKLETSI THSYTAFLLL ILFSASSFFF SIYHIKHAYY  
 GHIASINSRF PEQLAPLTLF SIVSILVATT LFFFSFLLGS FVVRRFIHQE  
 KDWTLDKVLQ QYSQLLAIPF FLTAIASFFA FFDSLRF TAL LCVISIGIIL  
 LASLHIITRP SQASETDSFY QLFLSVLVNG VIILLFFVAE VALIGDYLRI  
 LAFL

**GI15902625**

**spr0581**

**Zinc metalloprotease [Streptococcus pneumoniae R6]**

MFKKDRFSIR KIKGVVGSVF LGSLLMAPSV VDAATYHYVN KEIISQEAKD  
 LIQTGKPDNRN EVVYGLVYQK DQLPQTGTEA SVLTAFGLLT VGSLLLIYKR  
 KKIASVFLVG TMGLVVLPSA GAVDPVATLA LASREGVVEM EGYRYVGYLS  
 GDILKTLGLD TVLEETSAKP GEVTVVEVET POSTTNQEQA RTENQVVETE  
 EAPKEEAPKT EESPKEPKS EVKPTD DTLP KVEEGKEDSA EPAPVEEVGG  
 EVESKPEEKV AVKPESQPSD KPAAESKVEQ AGEPVAPRKD EQAPVEPENQ  
 PEAPPEEKAV EETPKQEST PDKAEETVE PKEETKTAKG TQEEGKEGQA  
 PVQEVNPEYK VTTGTVEKST ESELDFTEV VPDDTKYVDE EVVERQ GSKG  
 VQVTKTYET VEVVETDKVL STTTEVKTPV VPKVVKKGTK PVETREEVIP  
 FATKEQEDDT LKRGTRQVAQ EGVNGKKQIT ETYKTIRGEK TNEAPTVEET  
 VLQAPQDEII KKGTKGLEKP TLQWANTEKD VLKKSATASY TLTKPAGVEI  
 KSIKLALKDK DGQLVKEVTV AENNLNATLD KLYYYQGYTL STTMVYDRGE  
 GEE TEKLEDK QIQDLKKVE IKNIKETSLM NVDAEGNETD KSLLEKPTD  
 VSQLYLRVTT HDNKVTRLAV SSVVEVVVDG KTLYKVVAKA PDLVQRRADD  
 TLSEEVVHYF EKQLPKVNNV YYNFNELVKD MQANPMGEFK LGADLNAVNV  
 KPAGKAYVMA KFRGTLSSVE NHQYTIHNL RPLFNEAEGA TLKNFN LGNV  
 NINMPWADKV APIGNMFKKS TLENIKVVGVS VTGNNDVTGA VNKLDEANMR  
 NVAFIGKINS LGDKGWSG LVSESWSRNT DSVYFDGDIV GNNKFGGLV  
 AKVNHGSNQW DVKQKGRLTN SVVKGTMTLK NHGQSGGLVH ENYDWGWVEN  
 NISMVKVNNG EIMYGSISID GDPYFGFDYF KNNYYVKDVA TGESTYKRK

QIQSISQAEA DAKIANMGIT ANTFAIQDPV VNKLNRIDR DSEYKAIQDY  
 QETRNLAYRN LEKLQPFYK EWIVNQGNKL TDESNLVKKT VLSVTGMKSG  
 QFVTDLSSVD KIMIHADGT KEEFGVSAIS DSRVKQVKEY NVDDLGVVYT  
 PNMVDKNRDS LITKVKEKLS SVALDSAEVK SITNNPASLY LEESFAEVRE  
 TLDKLVKSLI ENEDHQLNSD EVAEKALLKK VEDNKAKIIL ALTYLNRYYG  
 IDYDGLNFKH LMMFKPDFYG KTPSILDFLI RIGSAEKNLK GDRSLEAYRE  
 VIGGTIGKGE LNGLLGYNMR LFTKYTDLND WFIHAAKNVY VSEPETTTED  
 FKDKRHRIYD GLNNDVHGRM ILPLLNLKKA HIFVISTYNT IAFSSFEKYG  
 KNTEEERNAY KAEIDRVAKA QORYLDFWSR LALPKVRNQL LKSQNSVPTP  
 VWDNQVYVGL GGANRMGYGD GGRVVTVPVRE LFGPTDRWHQ INWNMGAMAK  
 IYERPWKDDQ VYFVMTNMME PFGISAFTHE TTHVNDRMAY YGGDWHREGT  
 DLEAFAQGM LQTPDKSTTNG EYGALGINMA YERKNDGEQL YNYDPEKLDL  
 REKIDSVMKN YNESMMLLDY LEASAVIRQN LSDNSKWFKK MDKEWRTNAD  
 RNRLIGEPHQ WDKLRDLTEE EKKLPIIDSID KLVENNFVTL HGMPKNGRYR  
 TEGFDSSYQP VNMMAGVFGG NTSKSTVCSI SFKHNAFRMW GYYGYENGFI  
 PYVSNKLGKA ANKENKLLG DDFIIKK VSK NQFQNLLEWK KHWHYHEVYDK  
 AQKGFVEIEV DGVKISTYAO LQSLFEEAVS KDLAGMDDKN IKNHYQYTEN  
 LKWKIYKQLL KNTDGFSSDL FTAPQA

## GI15903086

spr1042

### immunoglobulin A1 protease [Streptococcus pneumoniae R6]

MEKYFGEKQE RFSFRKLSVG LVSATISSLF FMSVLASSSV DAQETAGVHY  
 KYVADSELSS EEKQLVYDI PTYVENDDET YYLVYK LNSQ NQLAELPNTG  
 SKNERQALVA GASLAALGIL IFAVSKKKVK NKTVLHLVLV AGMNGVLVS  
 VHALENHLLL NYNTDYELTS GEKLPLPKEI SGYTYIGYIK EGKTTSDFEV  
 SNQEKSAAATP TKQKQVDYV TPNFVDHPST VQAIQEQTPV SSTKPTEVQV  
 VEKPFSTELI NPRKEEKQSS DSQEQLAEHK NLETKKEEKI SPKEKIGVNT  
 LNPQDEVLSG QLNKPELLYR EETIETK IDF QEEIQENPDL AEGTVR VQKE  
 GKLGKKEVEIV RIFSVNKEEV SREIVSTSTT APSPRIVEKG TKKTQVIKEQ  
 PETGVEHKDV QSGAIVEPAI QPELPEAVVS DKGEPEVQPT LPEAVVTDKG  
 ETEVQPESPD TVVSDKGEPE QVAPLPEYKG NIEQVKPETP VEKTKEQGPE  
 KTEEVPVKPT EETPVNPNEG TTEGTSIQEA ENPVQPAEES TTNSEKVPD  
 TSENTGGEVS SNPSDSTTSV GESNKPEHND SKNENSEKTV EEVVPNNEG  
 TVEGTSNQET EKPVPQAEET QTN SGKIANE NTGEVSNKPS DSKPPVEESN  
 QPEKNGTATK PENSGNTTSE NGQTEPEKKL ELRNVSDIEL YSQTNGTYRQ  
 HVSLDGIPEN TDYFVKVKS SAFKDVYIPV ASITEEKRNG QSVYKITAKA  
 EKLOQELENK YVDNFTFYLD KKAKEENTNF TSFSNLVKAI NQNPSTGYHL  
 AASLNANEVE LGPDRSRYK DTFTGRLIGE KDGNKYAIYN LKKPLFENLS  
 GATVEKLSLK NVAISGKNDI GSLANEATNG TKIKQVHVDG VLAGERGVGG  
 LLAKADQSSI AESSFKGRIV NTYETTDAYN IGGLVGHLTG KNASIAKSKA  
 TVTISSNTNR SDQTVGGLAG LVDQDAHIQN SYAEGDINN V KHFGKVAGVA  
 GYLWDRSAGE EKHAGELTNV LSDVNVNNGN AITGYHYTGM KVANTFSSKA  
 NRVFNVTLEK DEVVSKESFE ERGTMLDASQ IVSKKAEINP LTLPTVEPLS  
 TSGKKDSDFS KIAHYQANRA LVYKNIEKLL PFYKSTIVK YGNLVKENS L  
 LYQKELLSAV MMKDDQVITD IVSNKQTANK LLLHYNDHSS EKFDLKYQTD  
 FANLAEYNLG NTGLLYTPNQ FLYDRDSIVK EVLPELQKLD YQSDAIRKTL  
 GISPEVKLLE LYLEDQFSKT QONLGDLSLK LLSADAGLAS DNSVTRGYLV  
 DKIKNNKEAL LLGLTYLERW YNFNYGQVNV KDLVMYHPDF FGKGNTPSLD  
 TLIELGKSGF NLLAKNNVD TYGISLASQH GATDLFSTLE HYRKVFLPNT  
 SNNDWFKSET KAYIVEEKST IEEVKTQQL AGTKYSIGVY DRITSATWKY  
 RNMVLP LLTL PERSV FVIST MSSLGFGAYD RYRSSDHKAG KALNDFVEEN  
 ARETAKRQD HYDYWYRILD EQSREKLYRT ILLYDAYKFG DDTTSGKATA  
 EAKFDSSNPA MKNFFGFPVGN KVVHNQHGAY ATGDGVYYMS YRMLDKDGA I

TYTHEMTHDS DQDIYLGGYG RRNGLGPEFF AKGLLQAPDQ PSDATITINS  
 ILKHSKSDST EGSRLQVLDP TERFQNAADL QNYVHNMF DL IYMMEYLEGQ  
 SIVNKL SVYQ KMAALRKIEN KYVKDPADGN EVYATNVVKE LTEAEARNLN  
 SFESLIDHNI LSAREYQSGD YERNGYT IK LFAPISALS SEKGT PGDLM  
 GRIAYELLA AKGFKDMVP YISNQYEEDA KQGGQTINLY GKERGLVTDE  
 LV LKKVFDGK YKTWAEFKTA MYQERVDQFG NLKQVTFKDP TKPWPSYGTK  
 TINNVDELQA LMDQAVLKDA EGPRWSNYDP EIDSAVHKLK RAIFKAYLDQ  
 TNDFRSSIFE NKK

**GI221231868**

**SPN23F\_10590**

**Zinc metalloprotease ZmpD [Streptococcus pneumoniae ATCC 70069]**

MSL LKKDKF S IRKIKGIVGS VFLGSL LFAP SVVGASTYHY LDYSSLTQTE  
 RDQLKQGRPD ESKEYALDY EKDALPNTGS SQSIMTALGL LAIGSLIVII  
 TKDNRNKKIA TFLIVGATGL VTLSTASALN LNANIHESGR DGVLQISGYR  
 YVGYLELDDK TVSSVSPAST VSPVEQPKVV TEKGEPEVHE KPDYTPQIGA  
 NLVEPEVHEK LAYTEPVGTT GVDENGNLIE PPVNDIPEYT EPVGT TGVD E  
 NGNLIEPPVS DIPEYTEPIS TVSEVASERE ELPSLHTDIR TETIPKTTIE  
 ESDPSKFIGD DSVRQVGEDG ERQIVTSYEE LHGK KISDPV ETVTILKEMK  
 PKILVKGTKE KPKEKTAPVL TLDRTNTNVL NRSATLSYHL VNTDGVTINK  
 ITATIKD GNE IVKTVDLTSE QLDKQVEDLK FYKDYKIETT MTYDRGKGEE  
 TATLEEKPLR LDLKKVEIKN IASTNLVKVN DDGTETPSDF MTEKPSDEDV  
 KKMYLKITSR DNKVTRLAVD KIEEVTEEGK KLYKITAEAQ DLIQHTDPTK  
 VRNKYVHYIE KPVPKVDVY YNFKELVDAM NADKNGTFKI GADLNATNVP  
 TPNKQYVPGT FKGHLSSVDG KQYTIHNIAR PLFDRVENG S VKNINLGNVD  
 INMPWADGIA PVANMVKNAT VEDVKVTGNV VANNNIAGIV NKIDSGGQLT  
 NVAFIGNLTG VGDKGQY MAG IAGEIWRGNL AKAYVEADIV ANRARIGGLV  
 AKTDNGNDSM GIGKYGSIRK SVTKGTIKTK VLFETGGFIN SNLPFGKLED  
 NISMMRVENG EEFFGSSDL D YDGGYFTNGW LERNFVVKGV SSGKHSYKRS  
 RDKIKEISQD EANKRIANFG LTADKYEINE PVVNRNLNRLT RREDEYKSTQ  
 DYKSERDLAY RNIEKLQPFY NKEWIVNQGN KLAEDSNLAK KEVLSVTGMK  
 DGQFVTDLSD IDKIMVHYAD GTEKEMDVTK NTDSKVQQVR EYSVSGLGDV  
 VYTPNMVVKN RDKLIADVKS QLSSVELISQ EVRDLMSRRD KPAENTDERK  
 NGYIKDLYLE ESFAEVKQNL DKLVKSLVEN EDHQLNGDEA AIKSLKKVE  
 TNKAKIMMAL TYLNRYDDIK YGDISIKNIM MFKPDFYGKT PSVIDRLINI  
 GSSEK NLKGD RTQDAYREII AGNTGKSNLR NFLEYNMRLF TEDKDINDWF  
 IHSAKNVYVS EPKTTNTELK DKRHRVFDGL DNGVHGRMIL PLLTLKDAHM  
 FLISTYNTMA YSSFEKYGKH TEEARNEFKT KIDEVAHAQQ TYLDFWSRLA  
 LPNVRDRLLK SQNMVPTPVW DNQTYNGSPV GRRGFDSKGN PIAPIRELYG  
 PTWRHHDRDW RMGAMASIFP NPNNDKVL F MVTDMISPF G ISAFTHETTH  
 VNDRMLYFGG HKHRQGT DVE AYAQGMLQTP DSSTNGEYG ALGINMAYHR  
 PNDGNQWYNP DPDKLKRDD IDRYMRNYNE AMMLLDHVEA DAVLPKIKGD  
 NSKWFKKIDK EMRSK IQYND LLGPNQWDSI RDLKDEEKVM TLSSVNDLVD  
 NNFMTKHG NP GNGRYPEDF TPNSAYVNVN MMAGIYGGNT SQGAPGSLSF  
 KHNAFRMWGY YGYENGFISY VSNKYKAEAD KNNHGLLSDK LIINKVSKGN  
 FNTLEEWKRH WYGEVLAKAK KGFEAIDIDG VHISNYDEL R PLFDKAVEED  
 LKKPDDFSHT VALKSKVFKA LLKNTDGF FN KLFKEDI

GI307128262

SP670\_2141

TMP repeat family [Streptococcus pneumoniae 670-6B]

MATLDELKVM IDAEIAPFRK KMK **EVENQVK** **GTSDQV** KNAT AKVREQSSSI  
GSAFGKLAKF AGFAILGKKL **LDVGMYSTQT** **ALEVSASMNQ** IKRQMGESSQ  
SFLKVVNDNA NAMNMGVGEA TNYGAVYSNL FSGFIKDTNK LSAYTAKMLQ  
TSAVVAEGSG RTITDVMERI R **SGLLGNTA** **IEDLGINVNV** **AMIESTEAFK**  
**K**FANGQSWQQ LDYQTQQQIR LMAILEQATA **KYGD**TLNSV **NGR**ISLFSKL  
MKDAALNLGN SMLPIINAIM PVLNSFAMVL KNVTAKLAEF IALMFNKKAT  
VK **DGVGGAVG** **DMGNAMK**DAA GGAGDLADAV DDAGDSAGGL ADNLGDSAKN  
AKKAAKELLG LLGFDEINIL QKPKDDDAGG SGGGGKGGK **KGGGGP**FKD  
**ILPEVELTDM** **DNKFKSIFDG** **LGDK**LKGLFD LFKKGFDAAF RPEGIKRIKT  
ALDQIAKTMG EIATDPR **VVN** **AFNR**MAEKIA YALGQVTGSI TTIGLGIGVF  
LAESIANGLG RQKERIIRAL VALFDNVGNL SEAVGNIAQD FSSAFYDVIT  
STGAVRIGSA IVSTLLSLTS TIVEVGSKLA GSLFKGFVKV VVTSAPKISS  
VFQSLLDTVV PVFESIERSV NKFGDGLSRV YDEHVVPAIN SIANAFNGLI  
DIIQILWENS WQPF AEFLSG VFGVSIEGIS DLLGGLLAT LGLLADAIKL  
VADGFTVFSW WCKENKEPIL ALITTWQTIN FLSWAEQAGG LAGAFSLLGS  
KVSLIVGGIK NLGLAIKALT FDKLVSFGET IYLNTLYAKD FVVNSGKTIA  
QLGKTALELG KSALAWTAHA AKMGLATAAE FAHSVAAGVA TAATWAFNAA  
LAVLTSPITW IIAAIAALIA IGVLLYQNWV TVVEFAKTAW QGLCDFISGI  
CRAIGEFFSG LWTKLQEIFE PIGQWFGKEF QQAWDAIVNI FSGIGEWFSG  
VFQGAWDAIV NIFTPIGSWF **GQRWADVTSA** **LANIGAWFTD** **IFQK**AWTGLT  
NIFSKLGLWF GERWADVTSV LANVSSWFGN MFTSAYNAVK NAFSSIGGFF  
SGVWSTVQSI FVNAGQKVG S AVGGAFKSAV NAVLGTIENV VNGFIGMING  
VLGVVRNLPV LGWVGSVSTV SLPRLARGGI VDSPTIAMIG EAGKEAVVPL  
ENTGFIQTLG RVVSSAVVNA MAGISPQGGF SSDGDIVIQI AGHEFGRVAI  
QEINKEHERA GQTLKKI