Supplementary data

Multiple sequence alignment (block 1, residues 1–50; secondary structure elements α1, α2, α3):

```
gi|119691
                                     α1           α2                        α3
                                   eeeeee      eeeeeeee               eeeeeeeee
             1          10        20          30          40          50
gi|119691     ......MAD.......VD.ELE.....DPMEEMTSYTFARFLRSPETEAFVRNLDRPPQMPAMRFVYLYC
gi|82503243   ......MAD.......VD.ELE.....DPMEEMTSYTFARFLRSPETEAFVRNLDRPPQMPAMRFVYLYC
gi|123845633  ......MAD.......VD.ELE.....DPMEEMTSYTFARFLRSPETEAFVRNLDRPPQMPAMRYVYLYC
gi|139424516  ......MAD.......VD.ELE.....DPMEEMTSYTFARFLRSPETEAFVRNLDRPPQMPAMRYVYLYC
gi|228537     ......MAD.......VD.ELE.....DPMEEMTSYTFARFLRSPETEAFVRNLDRPPQMPAMRYVYLYC
gi|51518056   ......MAD.......VD.ELE.....DPMEEMTSYTFARFLRSPEIEDFVRDLDRPPQMPAMRYVYLYC
gi|24943119   ......MGS.......QHVEME.....DPMINMTTFTFARFLRSSDAVRFASEIREQPRMPTMRYLYLYC
gi|20453824   M.......DILDNKTLLD.EASNLEIGEQLKKISEYTFSNFLRSVQVQEAIANKTIVPEMPTMRFVYLYY
gi|27452874   M.......DILDNKTFLD.EVSNLEIGEQLKKISEYTFSNFLRSAPVQEAIASKTIVPEMPAMRFVYLYY
gi|30348540   M.......DLFSEESPIN.EIGNMDMSDQQTQLCSSSFSHFLKHPKVQHFISTYSELVKMPTIRYVYFYY
gi|9625993    M.......DLFSEESPIN.EIGNMDMTDQQTQLCSSSFSHFLKHPKVQHFISTYSELVKMPTIRYVYFYY
gi|27452835   M.......DILDNRTLLE.EAHNLEIGEQLKKMAEYTFSNFIRSKPVQEAIAKKCFLPEMPTMRFVYIYY
gi|10140958   M.......DFLHKATLLE.QTLNMEIGEQVKKLATYTFSNFIRCPEVQQAIASGEIRNDMPHLRFVYIFY
gi|9631229    M.......DLFSEESLIN.EIGNLDMCDQQTRLCASSFSNFLKHNKVQHFISTYSELVKMPTIRYVYFYY
gi|9628040    M.......DFFSRVPLYD.ELSGLSQPETKCKLEEFTFSNFLKARRVREFLR.AHAVAKLPAMRYVYVYY
gi|9629578    MEGSII.LDFFEKDDLLE.VLGNMARENQLESLATLNFSAFIRSPQVQELLATSKTSVRISDMRLTYYYF
gi|83642873   M.......DFLHKATLLE.QTITLDIEDQVKKLSTYTFSNFIRCPEVQQAIGDGTIQRDMPHFRFVYIFY
gi|78127870   M.......DFLHKATLLE.QTITLDIEDQVKKLSTYTFSNFIRCPEVQQAIGDGTIQRDMPHFRFVYIFY
gi|262285082  MERSIT.LDFFAKDDLLE.LLGNMAREDQLESLATLNFSAFIRSPQVQDLLATSETHVRIADMRLTYYYF
gi|134038663  MERSIT.LDFFAKDDLLE.LLGNMAREDQLESLATLNFSAFIRSPQVQDLLATSETHVRIADMRLTYYYF
gi|13095614   MTMA.SPVDFFDSQPLLD.EMDTIDLDAQSRKITEFTFSSFLGHSRIQQFMSTCNVIPRMPAMRYMYFYY
gi|2246484    MEATPTPADLFSEDYLVD.TLDGLTVDDQQAVLASLSFSKFLKHAKVRDWCAQAKLQPSMPALRMAYNYF
gi|1718290    MEATPTPADLFSEDYLVD.TLDGLTVDDQQAVLASLSFSKFLKHAKVRDWCAQAKIQPSMPALRMAYNYF
gi|139472831  MEATPTPADLFSEDYLVD.TLDGLTVDDQQAVLASLSFSKFLKHAKVRDWCAQAKIQPSMPALRMAYNYF
gi|46519394   M.......DFFSDEPMVQ.EMALLDIDEQQRHLSKMSLANFLKHERVRAFFNDNKKKISMPAIRFVYNFY
gi|18653844   M.......DFFSDEPMVQ.EMALLDIDEQQRLLSKMSLANFLKHERVRAFFSDNKKEISMPAIRFVYNFY
gi|7330027    .........MVQ.EMALLDIDEQQRLLSKMSLANFLKHERVRAFFSDNKKEISMPAIRFVYNFY
gi|321496589  MTPQ...IDFFSKKSLVD.CLDGMEHEEQIDCISGLRFSAFLQTPEVQSLLSLSSTRVRVAPLRMTYYYF
```

Multiple sequence alignment (block 2, residues 60–90; secondary structure elements α4, α5; Motif I[3], Ia):

```
gi|119691
                  α4                          α5
             eeeeee    eeeeeeee                         eeeeeeeee
                 60        70        80                 90
gi|119691     LCKQIQEFSGETGFCDFVSSLVQENDSKD.................GPSLKSIYWGLQ
gi|82503243   LCKQIQEFSGETGFCDFVSSLVQENDSKD.................GPSLKSIYWGLQ
gi|123845633  LCKQIQEFSGETGFCDFVSSLVQENDSQD.................GPSLKSIYWGLQ
gi|139424516  LCEQIQEFSGETGFCDFVSSLVQENDSQD.................GPSLKSIYWGLQ
gi|228537     LCKQIQEFSGETGFCDFVSSLVQENDSQD.................GPSLKSIYWGLQ
gi|51518056   LCKQIQEFSGETGFCDFVSSLIQENDSQE.................SPSLISIYRGLQ
gi|24943119   LCKEIQEFSGDTAFCDFVTALIPENIRET.................PPSLRAVYRALA
gi|20453824   LFKKIQSFIGDEQILRVYDIEFKKNNIET.................PTLKDVVSLCN
gi|27452874   LFRKIQSFIGDEQILRVYDTEFKKNNIEI.................PTLKDVVSMCN
gi|30348540   LFKKIGGFIGNEKIGTYFSKNVCNNIAAK.................GV.PKLADVYKACE
gi|9625993    LFKKIGGFIGNEKIGTYFSKNVCNNIAAK.................GV.PKLADVYKACE
gi|27452835   LFKKIYSFIGDEQVLHKYETEFKKNSIKN.................PTLKDVVSVCS
gi|10140958   LFKKIGDFIGDTEVFGIYEKVMGVERANS.................AKLVDVYKACT
gi|9631229    LFRKIGGFIGNDKISNYFSENVYCDFESN.................SI.PKLADVYKACE
gi|9628040    LFSRIGDYIGDDTVLAIFGEFMGVSDAGKARGEGDGESDVHGEGEEEGDGGGEAEPGVGPDIPEVYRVCE
gi|9629578    LFLRLNEYIGNTAIMGVFKDMMHLTDNSE.................VSAVYAACR
gi|83642873   LFRKIHTFIGDSEVLDVYTQAMGSECAST.................AKLPDVVRACV
gi|78127870   LFRKIHTFIGDSEVLDVYTQAMGSECAST.................AKLPDVVRACV
gi|262285082  LFLRLNEYIGNTAIMGVFKDMMHLSDNSA.................VAAVYAACR
gi|134038663  LFLRLNEYIGNTAIMGVFKDMMHLSDNSA.................VAAVYAACR
gi|13095614   LFKKIGEFIGNNDIVKFYEDKVFDKYNPP.................G...SIYEVYMACH
gi|2246484    LFSKVGEFIGSEDVCNFFVDRVFGGVRLL.................DVASVYAACS
gi|1718290    LFSKVGEFIGSEDVCNFFVDRVFGGVRLL.................DVASVYAACS
gi|139472831  LFSKVGEFIGSEDVCNFFVDRVFGGVRLL.................DVASVYAACS
gi|46519394   LFAKVGDFIGNTEVYDFYVSCVFRGRRLT.................LLSDVYDACL
gi|18653844   LFAKVGDFIGNTDVYDFYVTCVFRGRRLT.................RLSEVYDACL
gi|7330027    LFAKVGDFIGNTDVYDFYVTCVFRGRRLT.................RLSEVYDACL
gi|321496589  LFRRIGEFMGDDSVCATFEATINISSDSG.................LNDVYSACA

                        ┌──────────── Motif I[3] ────────────┐        ┌── Ia ──┐
```

Multiple sequence alignment (block 3, residues 100–160; secondary structure elements α6, α7, β1, α8):

```
gi|119691
                   α6              α7        β1       α8
             eeeeeeeeeeee      eeeeee    eeee→   eeeeeeee..e           eee
                100       110       120       130       140       150       160
gi|119691     EATDEQRTVLCSYVESMTRGQSENLMWDILRNGIISSSKLLSTIK..NGPTKVFEPAPISTNHYFGGPVA
gi|82503243   EATDEQRTVLCSYVESMTRGQSENLMWDILRNGIISSSKLLSTIK..NGPTKVFEPAPISTNHYFGGPVA
gi|123845633  EATDEQRTVLCSYVESMTRGQSENLMWDILRNGIISSSKLLSTIK..NGPTKVFEPAPISTNHYFGGPVA
gi|139424516  EATDEQRTVLCSYVESMTRGQSENLMWDILRNGIISSSKLLSTIK..NGPTKVFEPAPISTNHYFGGPVA
gi|228537     AATDEQRTVLCSYVESMTRGQSENLMWDILRNGIISSSKLLSTIK..NGPTKVFEPAPISTNHYFGGPLA
gi|51518056   EATDEQRTVLCSYVESMTRGQSENLMWDILRNGIISSSKLLSTIK..NGPTKVFEPAPISTNHYFGGPVA
gi|24943119   RTTEREKTQLCSYIESMTRGQSENLMWDILRNGIISSSKLLSTLK..NGPMKLFEPNPISNNHYVGGPIA
gi|20453824   KMDITTKTDICLMIENMTRGQSENILWEILRDGVISSSKLLKVIKQQATDNKIFDPLPIQKNHYVASPIA
gi|27452874   KMDITTKTDICLMIENMTRGQSENILWEILRDGVISSSKLLKVIKQQAADNKIFDPLPIQKNHYVASPIA
gi|30348540   KMNLRQQSEICLLIEEVTRGQYLNSLWDALRDGTISSSKFYWATKKQNSTKKIFEPWPIKNDYYVAGPLA
gi|9625993    KMNLRQQSEICLLIEEVTRGQYLNSLWDALRDGTISSSKFYWATKKQNSTKKIFEPWPIKNDYYVAGPLA
gi|27452835   KMDLTTRTDICMMIENMTRGQSDNLLWEILRDGVISSSKLLKQQASDSKIFDPLPIQKNHYVASPVA
gi|10140958   EMRPHIQAQICEHIERLTRGQGENVLWEVLRDGVISSSKLLKFVKQQTPDSKIFNPIPIQKNHYVASPVA
gi|9631229    KMNLTQQSKICLLIEEVTRGQYLNSIWDALRDGTISSSKFYWATKKQNSTKKIFEPWPIKNDYYVAGPLA
gi|9628040    CLPVRTKCRISLAIEAITRGQYENNLWEIFRDGIISSSKFYHAVRQQNSSKKLFQPWPIVNNYYPLSPLA
gi|9629578    DVTPDIKYAVCQRIEALTRGQQDNELWDILRDGMISSSKFQWAVKQHSINKKIFNPQPIKVNHYVAGPLA
gi|83642873   AMKPPLKAKICEHIERLTRGQSDNVLWEVLRDGVISSSKLLKFVKQQTPDSKIFHPIPIQKDHYVASPVA
gi|78127870   AMKPPLKAKICEHIERLTRGQSDNVLWEVLRDGVISSSKLLKFVKQQTPDSKIFHPIPIQKDHYVASPVA
gi|262285082  EVTPDIKYAVCQRIEALTRGQQDNELWDILRDGMISSSKFQWAVKQQSINKKLFNPQPIKVNHYFAGPLA
gi|134038663  DVTPDIKYAVCQRIEALTRGQQDNELWDILRDGMISSSKFQWAVKQQSINKKLFNPQPIKVNHYFAGPLA
gi|13095614   HMDLYKQYAICLLLESITREQHLSTLRNGIISSSKMHWVIKQRKTSKKIFEPWPIKNNHFVAGPLA
gi|2246484    QMNAHQRHHICCLVERATSSQSLNPVWDALRDGIISSSKFHWAVKQQNTSKKIFSPWPITNNHFVAGPLA
gi|1718290    QMNAHQRHHICCLVERATSSQSLNPVWDALRDGIISSSKFHWAVKQQNTSKKIFSPWPITNNHFVAGPLA
gi|139472831  QMNAHQRHHICCLVERATSSQSLNPVWDALRDGIISSSKFHWAVKQQNTSKKIFSPWPITNNHFVAGPLA
gi|46519394   NMHPHDRHHVCALIEQVTRGQNINPLWDALRDGIISSSKFHWAIKQQNSSKKIFNPWPIVNNHFIAGPLA
gi|18653844   NMHPHDRHHVCALIEQVTRGQNINPLWDALRDGIISSSKFHWAIKQQNSSKKIFNPWPIVNNHFVAGPLA
gi|7330027    NMHPHDRHHVCALIEQVTRGQNINPLWDALRDGIISSSKFHWAIKQQNSSKKIFNPWPIVNNHFVAGPLA
gi|321496589  NTIPEILYGICLKIEAMTRGQQANGLWDILRDGVSSSRFHWAVKKSMGSKKLFVQWPIQSNHYVAGPLA
```

```
                 Motif Ia²                                Motif II³                      III³
                      α9            β2      β3      β4      β5        β6
gi|119691      ΩΩΩΩΩΩΩΩΩΩΩΩΩΩΩ      ──→     ──→     ──→     ──→      ───→
                     170           180     190     200     210       220
```

```
gi|119691      FGLRCEDTVKDIVCKLIC.GDASANRQFGFMISPTDGIFGVSLDLCVNV.ESQGDFILFTDRSCIYEIKC
gi|82503243    FGLRCEDTVKDIVCKLIC.GDASANRQFGFMISPTDGIFGVSLDLCVNV.ESQGDFILFTDRSCIYEIKC
gi|123845633   FGLRCEDTVKDIVCKLIC.GDASANRQFGFMISPTDGIFGVSLDLCVNV.ESQGDFILFTDRSCIYEIKC
gi|139424516   FGLRCEDTVKDIVCKLIC.GDASANRQFGFMISPTDGIFGVSLDLCVNV.ESQGDFILFTDRSCIYEIKC
gi|228537      FGLRCEDTVKDIVCKLIC.GDASANRQFGFMISPTDGIFGVSLDLCVNV.ESQGDFILFTDRSCIYEIKC
gi|51518056    FGLRCEDTVKDIVCKLIC.GDAAANRQFGFMISPTDGIFGVSLDLCVNV.ESRGDFIVFTDRSCIYEIKC
gi|24943119    FGVRNENTVKTLLTQLIAKESWSNTTNFGFMLSPIDGIFGVSLDMCLKASVDVHNKVMFSSLTEIYEIKC
gi|20453824    FGVRNENTVKTLLTQLIAKESWSNTTNFGFMLSPIDGIFGVSLDMCLKASVDVHNKVMFSSLTEIYEIKC
gi|27452874    FGVRNENTVKLLLTQLIAKESWSNTTNFGFMLSPIDGIFGVSLDMCLKASVDVHNKVVFSSLTEIYEIKC
gi|30348540    FGLRCEEVIKTVLNELICTPKQASCFDCGFMQSPLDGIFGVSLDYCTNVETNKDNLLVFHPDTEVYEIKS
gi|9625993     FGLRCEEVIKTVLNELICTPKQASCFDCGFMQSPLDGIFGVSLDYCTNVETNKDNLLVFHPDTEVYEIKS
gi|27452835    FGVRNENTVKTLLTQLIAKESWSNVTNFGFMLSPLDGIFGVSLDMCLKSSVDAENKVVFSSLTEIYEIKC
gi|10140958    FGVRNETVVKKLISELVVEEGIGCVTEFGFMLSPNDGIFGVSLDMCTNASMSDHNTVEFTSTTTIYEIKC
gi|9631229     FGLRCEEVIKTVLNELVCTSKQVGCFDCGFMQSPLDGIFGVSLDYCTNVETNKDNLLVFHPDTEVYEIKC
gi|9628040     FGLRCEDAVKTLLAEFVC.GRKEVMCDVGFLQSPKDGIFGVSLDMCANVSVGRDNLLEFRADAEIYEIKC
gi|9629578     FGIRCEQTVKKILSELIH.PNLPRFHDCGFLPSAIDGIFGVSLDTAFNVFTDSSGLVHFEPDSIVYEIKS
gi|83642873    FGVRNESTVKTLVSELVLEENVKCITDFGFMLSPNDGIFGVSLDMCANACLTSDNIVQFEATTHIYEIKC
gi|78127870    FGVRNESTVKTLVSELVLEENVRCITDFGFMLSPNDGIFGVSLDMCANACLTSDNIVQFEATTQIYEIKC
gi|262285082   FGIRCEQTVKKILSELIH.PSLPRFHDCGFLPSAIDGIFGVSLDTAFNVFTDSCGLVHFEPDSIVYEIKS
gi|134038663   FGIRCEQTVKKILSELIH.PSLPRFHDCGFLPSAIDGIFGVSLDTAFNVFTDSCGLVHFEPDSIVYEIKS
gi|13095614    FGLRCEGIVKSILINIIY.PNTPNCIDYGFMSPLDGIFGVSLDFCTNISHDENGMLIFEPDCCVYEIKC
gi|2246484     FGLRCEEVVKTLLATLLH.PDETNCLDYGFMQSPQNGIFGVSLDFAANVKTDTEGRLQFDPNCKVYEIKC
gi|1718290     FGLRCEEVVKTLLATLLH.PDETNCLDYGFMQSPQNGIFGVSLDFAANVKTDTEGRLQFDPNCKVYEIKC
gi|139472831   FGLRCEEVVKTLLATLLH.PDEANCLDYGFMQSPQNGIFGVSLDFAANVKTDTEGRLQFDPNCKVYEIKC
gi|46519394    FGLRCEDVVKKILATLLH.PGEAHCENYGFMQSPHNGVFGVSLDFGINVKSNPKDGLEFHPDCKIYEIKC
gi|18653844    FGLRCEKVVKKILATLLH.PGEAHCENYGFMQSPLNGVFGVSLDFGINVRSDPKDGLEFHPDCKIYEIKC
gi|7330027     FGLRCEEVVKKILATLLH.PGEAHCENYGFMQSPLNGVFGVSLDFGINVRSDPKDGLEFHPDCKIYEIKC
gi|321496589   FGLRCEETVKTLLSVLIH.PNVPCYNNCGFLPSACDGIFGVSLDSAFNAYIDADGSVNFEPDTIVYEIKS
```

```
                    Motif III³                                 Motif IV³
                                   α10          α11          β7          β8
gi|119691      ●                ΩΩΩΩΩΩΩΩΩ    ΩΩΩΩΩΩΩΩ         ─→          ─→
                     230         240          250      260  270         280        290
```

```
gi|119691      RFKYLFSKSEFDPIYPSYTALYKRPCKRSFIRFINSIARPTVEYVPDGRLPSEGDYLLTQDEAWNLK.DV
gi|82503243    RFKYLFSKSEFDPIYPSYTALYKRPCKRSFIRFINSIARPTVEYVPDGRLPSEGDYLLTQDEAWNLK.DV
gi|123845633   RFKYLFSKSEFDPIYPSYTALYKRPCKRSFIRFINSIARPTVEYVPDGRLPSEGDYLLTQDEAWNLK.DV
gi|139424516   RFKYLFSKSEFDPIYPSYTALYKRPCKRSFIRFINSIARPTVEYVPDGRLPSEGDYLLTQDEAWNLK.DV
gi|228537      RFKYLFSKSEFDPIYPSYTALYKRPCKRSFIRFINSIARPTVEYVPDGRLPSEGDYLLTQDEAWNLK.DV
gi|51518056    RFKYLFSKSEFDPIYPSYTALYERPCKRSFIRFINSIARPTVEYVPDGRLPSEGDYLLTQDEAWNLK.EV
gi|24943119    RFKYLFSKSEFDPIYPSYTALYQNPSKKSFIRFINSIARPTIEYVADGRLPSESDYLLTRDESWNLK.EV
gi|20453824    RYKYLFSKSEFDPIYTKYDALYNNPCKATLVDFISSIPKPAVEHVPRGRVPTQNDYLLSFDKMWNFNPRQ
gi|27452874    RYKYLFSKSEFDPIYTKYDALYNSPCKATLVDFISSIPKPAVEHVPRGRVPTQNDYLLSFDKIWNFNPRQ
gi|30348540    RFKYLFDKSECDTLYKKYKELYSNPCVKTLIKFIFSVSRPAIEFVPSGRLPSESDYLLAYDEEWNLR.PT
gi|9625993     RFKYLFDKSECDTLYKKYKELYSNPCVKTLIKFIFSVSRPAIEFVPSGRLPSESDYLLAYDEEWNLR.PT
gi|27452835    RYKYLFSKSEFDHIYKKYDQLYNNPCKSTLVDFISSIPKPAVEHVPRGKVPSQNDYLLSFDKLWNFNPRQ
gi|10140958    RYKYLFSKCDYDPIYQAYQKLYNSPGRQELDFIQSIQKPTVEYVSRGRLPTQNDYLLSFDRSWDFGPPK
gi|9631229     RFKYLFAKTDCDMLYKKYKDLYTNPCLKTLIKFIFSVSRPAIEFVPPGKLPSETDYLLAYDEEWNLR.PT
gi|9628040     RFKYNYSKIECDPLYQKYVSLYNSPSKTTLIRFLSGINRPAVEYVPPGKLPTKNDFLLTSDRDWDLS.PK
gi|9629578     RYKYQFSKSEFDGLAKKYLDLYKNPCEATFIKFINCISRPAVEYVPHGKLPSESDYLITHSMNWKTS.NK
gi|83642873    RYKYLFAKCDYDPIYNYYLALYNAPGRQPLIDFIQAIQKPAVEYVPRGRLPTQNDYLLSFDKCWDFGPPR
gi|78127870    RYKYLFAKCDYDPIYNYYLALYNAPGRQPLIDFIQAIQKPAVEYVPRGRLPTQNDYLLSFDKCWDFGPPR
gi|262285082   RYKYQFSKSEFDGLSKKYVDLYKNPSEATFIKFINCISRPAVEYVPSGKLPSESDYLITHSTNWKTN.NK
gi|134038663   RYKYQFSKSEFDGLSKKYVDLYKNPSEATFIKFINCISRPAVEYVPSGKLPSESDYLITHSTNWKTN.NK
gi|13095614    RFKYMFSKSECDPLYGKYVSLYQNPNKKNLINFILSVSRPAVEFVAPGGIPSERDFLLTHGLEWRWE.PP
gi|2246484     RFKYTFAKMECDPIYAAYQRLYEAPGKLALKDFFYSISKPAVEYVGLGKLPSESDYLVAYDQEWEAC.PR
gi|1718290     RFKYTFAKMECDPIYAAYQRLYEAPGKLALKDFFYSISKPAVEYVGLGKLPSESDYLVAYDQEWEAC.PR
gi|139472831   RFKYTFAKMECDPIYAAYQRLYEAPGKLALKDFFYSISKPAVEYVGLGKLPSESDYLVAYDQEWEAC.PR
gi|46519394    RFKYTFSKMECDPIYAAYAKLYQKPSMQTLKGFLYSISKPAIEFVGEDKLPSEADYLVAYDKEWEVC.PR
gi|18653844    RFKYTFSKMECDPIYAAYAKLYQKPSMQTLKGFLYSISKPAIEFVGERLPSESDYLVAYDKEWEVC.PR
gi|7330027     RFKYTFSKMECDPIYAAYAKLYQKPSMQTLKGFLYSISKPAIEFVGERLPSESDYLVAYDKEWEVC.PR
gi|321496589   RYKYLFSKSECHVLSKHYTALYDNPSPANFFRFMMAIQKPGVDFLPPGKHPTERDYLLTSNEKCHMG.PK
```

```
                           Motif V³                                       Motif VI³
                                α12          β9        β10   β11     β12          α13
gi|119691      ●             ΩΩΩΩΩΩΩΩΩΩ      ──→        ─→    ─→      ──→      ΩΩΩΩΩΩΩΩΩΩΩ
                     300          310        320       330   340     350         360
```

```
gi|119691      RKRKLGPGHDLVADSLAANRGVESMLYVMTDPSENAGRIGIKDRVPVNIFINPRHNYFYQVLLQYKIVGD
gi|82503243    RKRKLGPGHDLVADSLAANRGVESMLYVMTDPSENAGRIGIKDRVPVNIFINPRHNYFYQVLLQYKIVGD
gi|123845633   RKRKLGPGHDLVADSLAANRGVESMLYVMTDPSENAGRIGIKDRVPVNIFINPRHNYFYQVLLQYKIVGD
gi|139424516   RKRKLGPGHDLVADSLAANRGVESMLYVMTDPSENAGRIGIKDRVPVNIFINPRHNYFYQVLLQYKIVGD
gi|228537      RKRKLGPGHDLVADSLAANRGVESMLYVMTDPSENAGRIGIKDRVPVNIFINPRHNYFYQVLLQYKIVGD
gi|51518056    RKRKLGPGHDQVAAALAANRGVESMLYVMTDPSENAGRIGIKDRVPVNIFINPRHNYFYQVLLQYKIVGD
gi|24943119    RRRKLGEGHDTVVTCLDYNRGAESMLYVMNDPSENSGIIGIKDRVPVDVFINPRHNYFYQVLLQYKIVGD
gi|20453824    KRRKMTNIHKLTEQCMKYNCYTESKVIILTDPALTSGNICIKDTFFVDLYINPRHAYYYQILLQYKIVTN
gi|27452874    KRRKMTNIHKLTEQCMKYNCYTESKVIILTDPALTSGNICIKDTFFVDLYINPRHAYYYQILLQYKIVTN
gi|30348540    KKRKLNASHEMIKKCIEYNSYAGSQIYILSDPAENGQITVKTKFKAGIFMNPRETYFYQVALQHRVVQS
gi|9625993     KKRKLNASHEMIKKCIEYNSYAGSQIYILSDPAENNGQITVKSKFKAGIFMNPRHTYFYQVALQHRVVQS
gi|27452835    KRRKMTNTHKVTEQCMKYNCYTESKVIILTDPALTSGNICIKDTFFVDLYINPRHAYYYQILLQYKIVTN
gi|10140958    RKRKLTSGHKITEQCMKYNCYTESKVIILTDPALTSGKIEVKGSFFVDIYINPRHAYYHQCMLQYKIVTN
gi|9631229     KKRKLTADHDMIKKCMEYNSYSCSWVYILSDPAENNGIISIKSKFKAGIFMNPRHTYFYQVALQHRVVQS
gi|9628040     RKRNLTPAHKSLYDCIRANEHASSQVLILSDPSETEGKIDIKARFDVDIFINPEHSYFYQILLQYKVVWD
gi|9629578     RKRKITDSHICLKKCLLHNMYQQSTVYILSDPSETEGKITIKASFPIDVFVNPAHNYFYQALQNMVVQD
gi|83642873    RKRRLVSGHKVTEQCMKYNCYTESKVLVLTDPALTSGKIEVKNSFFVDIYINPRHAYYHQCMLQYKIVTN
gi|78127870    RKRRLVSGHKVTEQCMKYNCYTESKVLVLTDPALTSGKIEVKNSFFVDIYINPRHAYYHQCMLQYKIVTN
gi|262285082   RKRNITDSHICLKKCLLHNMYQQSTVYILSDPSETSGKITVKASFPIDVFVNPNENYFYQVALQNMVVQD
gi|134038663   RKRNITDSHICLKKCLLHNMYQQSTVYILSDPSETSGKITVKASFPIDVFVNPNENYFYQVALQNMVVQD
gi|13095614    RKRKTVKSTNWIIECIKYNSCVESDVYVLTDPSITNGNITIKSHFKADLFVNVRESYFFQVLLQSSIVEE
gi|2246484     KKRKLTPLHNLIRECILHNSTTESDVYVLTDPQDTRGGQISIKARFKANLFVNVRHSYFYQVLLQSSIVEE
gi|1718290     KKRKLTPLHNLIRECILHNSTTESDVYVLTDPQDTRGGQISIKARFKANLFVNVRHSYFYQVLLQSSIVEE
gi|139472831   KKRKLTPLHNLIRECILHNSTTESDVYVLTDPQDTRGGQISIKARFKANLFVNVRHSYFYQVLLQSSIVEE
gi|46519394    KKRRLTAVHHLVKKCMIHNSTAPSDVYILSDPQETGGQINIKAHLSANLFINVREPYYYQVLLQSLVVQE
gi|18653844    KKRRLTAVHHLVKKCMIHNSTAPSDVYILSDPQETGGQINIKAHLSANLFINVREPYYYQVLLQSLVVQE
gi|7330027     KKRRLTAVHHLVKKCMIHNSTAPSDVYILSDPQETGGQINIKAHLSANLFINVREPYYYQVLLQSLVVQE
gi|321496589   KRKNLTSQHHLLQQCLIHNKDQKSILHVLTDPSITQGGIISIKQTVPVDIYINPAHPYFFQVCLQHMVVQD
```

```
                        VI                                                      Motif VII 3
gi|119691                                 β13              β14    β15          β16
                       ΩΩΩ            →  β13  →         →β14→  →β15→       →  β16  →     ΩΩΩΩΩΩ
                       370      380      390      400      410      420      430

gi|119691     YVRHSGGGKP.GRDCSPRVNIVTAFFRKRSPLDPATCTLGSD.LLLDASVEIPVAVLVTPVVLPDSVIRK
gi|82503243   YVRHSGGGKP.GRDCSPRVNIVTAFFRKRSPLDPATCTLGSD.LLLDASVEIPVAVLVTPVVLPDSVIRK
gi|123845633  YVRHSGGGKP.GRDCSPRVNIVTAFFRKRSPLDPATCTLGSD.LLLDASVEIPVAVLVTPVVLPDSVIRK
gi|139424516  YVRHSGGGKP.GRDCSPRVNIVTAFFRKRSPLDPATCTLGSD.LLLDASVEIPVAVLVTPVVLPDSVIRK
gi|228537     YIRHSGGGKP.GRNCSPRVNIVTAFFRKRSPLDPATCTLGSD.LLLDASVEIPVAVLVTPVCLPDSIVRK
gi|51518056   YVRHSGGGAA.DRDCQPRVNIVTAFFRKRSPLDPPTCTLGSD.LLLDASVEIPVAVLVTPVVLPDSVIRE
gi|24943119   YIRW.GSQSK.KTNPKPKVNIVTAFFRKRSSSDPALCTIGPQ.VPLDTSVEIPTPIYIPPPSVVKE
gi|20453824   YIQFSQN..SGSKLGKPETYIVTAFFRKRDSSDFKRTYIKTESNVLEPSLEIPVLLIITPVFIPHDPLIA
gi|27452874   YIQFSQN..SGSKLGKPETYIVTAFFRKRDSSDFKRTYIKTESNVLDPSLEIPVLLIITPVFIPHDPLIA
gi|30348540   YIGLSES..P.RSLGTQKNFIVSSFFRKRHFSDPPVCYIGKK.Q.LEKTVEIPVFIIITPVYIPRSALLE
gi|9625993    YIGLSES..P.KSLGTQKNFIVSSFFRKRHFSDPPVCYVGKK.Q.LEKTVEIPVFIIITPVYIPRSALLE
gi|27452835   YIQFSHN..SGSKLGSPSTYIVTAFFRKRDSSDFKRTYINQEKNLLDPSLEIPVLLIITPVFIPHDPLIA
gi|10140958   YVQLTKG..DSCKHTHPGVFLVSAFFRKRNSADFPKTYIKTERSFVDASVEIPVLLIITPVFVPHGPLVD
gi|9631229    YIGLSES..P.RSLGTQKNFIVSSFFRKRHFSDPQLCYIGKT.Q.LEKTVEIPVFIIVTPVYIPRSALLE
gi|9628040    YIQYHSS....PGLGSLKTFIVSGFFRKRNHSDPLECSIGGR.GTLDSACEIPVLLILTPVYIPHCVVTE
gi|9629578    YIEFGQG..VYKRLGHQKNFIASGFFRKRHFSDPAVCSIGNC.GKLDTTDEIPVALIITPVRIPSTVLHE
gi|83642873   YVQLSRD..HSCRHTRPSVFLVSAFFRKRNSADFPKTYIKTERGFLDCSVEIPVLLIITPVFIPHEPLVE
gi|78127870   YVQLSRD..HSCRHTRPSVFLVSAFFRKRNSADFPKTYIKTERGFLDCSVEIPVLLIITPVFIPHEPLVE
gi|262285082  YIEFGQG..VYKRLGHQKNFIASGFFRKRHFSDPPVCTIGEC.VKLDTTDEIPVALIITPVRIPSTVLNE
gi|134038663  YIEFGQG..VYKRLGHQKNFIASGFFRKRHFSDPPVCTIGEC.VKLDTTDEIPVALIITPVRIPSTVLNE
gi|13095614   YIQFSPS..T.KTLGSQKNFIVSAFFRKRNFKDPLTCTLGDTREVLKETVEIPVMIIITQVRIPKFILKE
gi|2246484    YIGLDSG..I.PRLGSPKYYIATGFFRKRGYQDPVNCTIGGD.A.LDPHVEIPTLLIVTPVYFPRGAKHR
gi|1718290    YIGLDSG..I.PRLGSPKYYIATGFFRKRGYQDPVNCTIGGD.A.LDPHVEIPTLLIVTPVYFPRGAKHR
gi|139472831  YIGLDSG..I.PRLGSPKYYIATGFFRKRGYQDPVNCTIGGD.A.LDPHVEIPTLLIVTPVYFPRGAKHR
gi|46519394   YISLSKG..T.KNLGTQKNFIATGFFRKRQFQDPSNCTIGEF.APLDPHVEIPTLLIVTPVYFPSVAKHQ
gi|18653844   YISLSKG..T.KNLGTQKNFIATGFFRKRQFQDPSCCTIGEF.APLDPHVEIPTLLIVTPVYFPSVAKHQ
gi|7330027    YISLSKG..T.KNLGTQKNFIATGFFRKRQFQDPSCCTIGEF.APLDPHVEIPTLLIVTPVYFPSVAKHQ
gi|321496589  YIAFKES..Y.TRLGHQKNFIVSAFFRQRHFSDPKECYIGDL.AKIEAGREIPVALLITPVRIPSTVLAE


                        α14
                       ΩΩΩΩΩΩΩΩΩΩΩΩΩΩ
                       440      450      460      470

gi|119691     TLSTAAGSWKAYADNTFDT.APWVPSGLFADDESTP
gi|82503243   TLSTAAGSWKAYADNTFDT.APWVPSGLFADDESTP
gi|123845633  TLSTAAGSWKAYADNTFDT.APWVPSGLFADDESTP
gi|139424516  TLSTAAGSWKAYADNTFDT.APWVPSGLFADDESTP
gi|228537     ELNTAAGSWKAYADDTFDT.APWVPSGLFADDESTP
gi|51518056   TLNAAAGSWKAYADNTFDT.APWVPSGLFAGDESTP
gi|24943119   TLATSAESGRAYAENTFES.APWVPSSLFAGAEYTP
gi|20453824   TLKKAIGFWQKSVQEEFPF.PPWACSSLCALGDITP
gi|27452874   TLQKAIGFWQKSVQEEFPF.SPWACSSLCALGDITP
gi|30348540   TISQAVNFWEESAKEAFTE.YPWAPCALFANGDLTP
gi|9625993    TISQAVNFWEESAKEAFTE.YPWAPCALFANGDLTP
gi|27452835   TLEKAIEFWQKSVQEEFPF.SPWASSSLCALGDITP
gi|10140958   TLEQAIKFWQVAVKEEFNH.WPWAPTSLSAVGDVTP
gi|9631229    SISKAVNFWEESAKETFTE.YPWAPCSLFANGDLTP
gi|9628040    SLKKASQYWNQSAEEEFSH.PPWVSSSLFADDGMTP
gi|9629578    YLKKAIDFWNQCAEANFDH.IPWGQLSDVARKPITP
gi|83642873   TLEQAIGFWQTAVEEEFCH.TPWAPTSLSAVGDITP
gi|78127870   TLEQAIGFWQTAVEEEFCH.TPWAPTSLSAVGDITP
gi|262285082  YIRKAIDFWNKCAEENFEH.SPWGRLSDVARKAITP
gi|134038663  YIRKAIDFWNKCAEENFEH.NPWGRLSDVARKAITP
gi|13095614   NMRKATTYWADCSEKTFTHSPWVTGLHLAVGKSMTP
gi|2246484    LLHQAANFWSRSAKDTFPY.IKWDFSYLSANVPHSP
gi|1718290    LLHQAANFWSRSAKDTFPY.IKWDFSYLSANVPHSP
gi|139472831  LLHQAANFWSRSAKDTFPY.IKWDFSYLSANVPHSP
gi|46519394   LVKQATEFWAASASEAFPE.LPWDLSSLCANAPPTP
gi|18653844   LVKQATEFWAASAREAFPE.LPWDLSSLCANAPPTP
gi|7330027    LVKQATEFWAASAREAFPE.LPWDLSSLCANAPPTP
gi|321496589  YLQSATDFWNKCSKETFKC.APWVQYSDYAGKTLTP
```
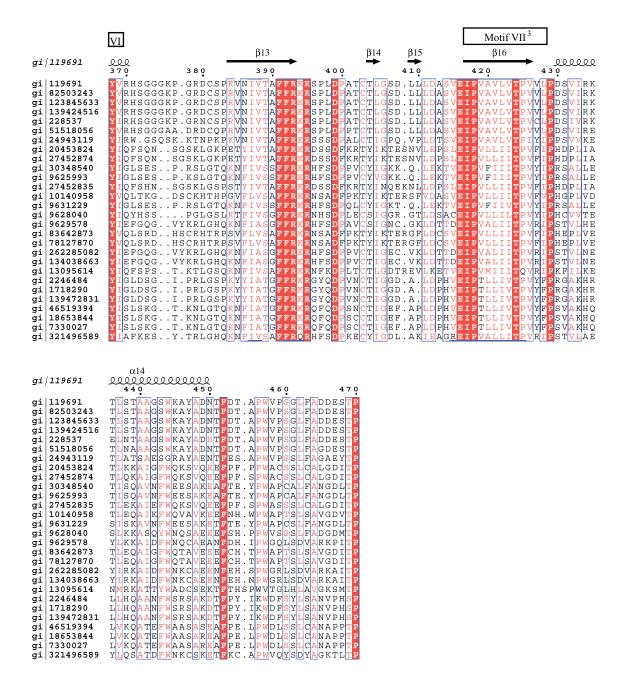
Figure S1. Alignment of gamma-herpesvirus exonuclease sequences

ESPript (1) was used to prepare this figure. Strictly conserved residues are printed on a red background, residues where some conservative substitutions occur are printed in red. The secondary structure elements are derived from the crystallographic structure of EBV BGLF5 (pdb entry 2W45 (2)). The conserved motifs are indicated above the alignment (adapted from (2,3)). The mutants under study are colored as follows: P158 mutants in dark blue, D203S mutant in purple, K231M mutant in yellow, and G305C mutant in red.

Supplemental references

1. Gouet, P., E. Courcelle, D. I. Stuart, and F. Metoz. 1999. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics. 15:305-308.

2. Buisson, M., T. Geoui, D. Flot, N. Tarbouriech, M. E. Ressing, E. J. Wiertz, and W. P. Burmeister. 2009. A bridge crosses the active-site canyon of the Epstein-Barr virus nuclease with DNase and RNase activities. J. Mol. Biol. 391:717-728.

3. Goldstein, J. N. and S. K. Weller. 1998. The exonuclease activity of HSV-1 UL12 is required for in vivo function. Virology 244:442-457.