

ID CAF-; PATTERN DE NONE PA C-[TS]-C-[FLIV]-X-G-X(4,6)-C CC
 Randomized probability: min = 5.551e-09 max = 5.560e-09 . CC Patte
 rns In proteins query (PattInProt v5.4up, November 15th, 2011) CC C
 ontact Dr. Christophe Blanchet, christophe.blanchet@ibcp.fr, http://g
 bio-pbil.ibcp.fr/cblanchet CC Scan criterion: Similarity 100% CC
 Sequence databank: /db/uniprot/fasta/current/swissprot.fas CC Numbe
 r of sequences in databank: 561356 (Tue, 29 Sep 2020 10:07:29 AM UTC)
 . //

First step : before [\[go to EXTRACT\]](#).

SELECT SITES

of pattern 1 with 100% similarity.

```

-----
----- NPSA gnllunipspIP40410 [Bacteria] Iron-uptake system per
mease protein FeuB [Firmicutes][Bacillus subtilis] (length=334 residu
es). *****> PATTERN 1  Site : 282- 294, Identity vgsdysrli
p_CSCILGGIFLTLCDlasrfinyp -----
----- NPSA gnllunipspIP47104 [Eukar
yota] Regulator of V-ATPase in vacuolar membrane protein 1 [Fungi][Sa
ccharomyces cerevisiae] (length=1357 residues). *****> PATTERN 1 
Site : 602- 614, Identity aceintgimn_CTCIRGSSTGKLC_ivnstgkvm
-----
----- NPSA gnllunipspIA2VEC9 [Eukaryota] SCO-spondin [Metazoa
][Homo sapiens] (length=5150 residues). *****> PATTERN 1  Site :
3165- 3175, Identity gstvvrpcen_CSCVSGLIANC_sswpcaegep -----
-- NPSA gnllunipspIP00742 [Eukaryota] Coagulation factor X [Metazoa]
[Homo sapiens] (length=488 residues). *****> PATTERN 1  Site :
110- 121, Identity kckdglgeyt_CTCLEGFEGKNC_elftrklcsl -----
- NPSA gnllunipspIP02751 [Eukaryota] Fibronectin [Metazoa][Homo sapi
ens] (length=2477 residues). *****> PATTERN 1  Site : 2410- 242
1, Identity qkeylgaics_CTCFGQRGWRC_dncrrpggep -----
----- NPSA gn
llunipspIP04275 [Eukaryota] von Willebrand factor [Metazoa][Homo sap
iens] (length=2813 residues). *****> PATTERN 1  Site : 2281- 22
91, Identity wvdpqhpcqi_CTCLSGRKVNC_ttqpcptaka -----
----- NPSA gn
llunipspIP13611 [Eukaryota] Versican core protein [Metazoa][Homo sap
iens] (length=3396 residues). *****> PATTERN 1  Site : 3113- 31
24, Identity tcyptetsyv_CTCVPGYSGDQC_eldfdechsn -----
----- NPSA g
nllunipspIQ08431 [Eukaryota] Lactadherin [Metazoa][Homo sapiens] (le

```

ngth=387 residues). *****> PATTERN 1 Site : 55- 66, Identity
 vrgdvfpsy_tCTCLKGYAGNHC_etkcvpeplg1 -----
 ----- NPSA [gnllunipsplQ2I0M4](#) [Eukaryota] Leucine-rich repeat-containing protein 26 [Metazo
 a][Homo sapiens] (length=334 residues). *****> PATTERN 1 Site :
 47- 57, Identity slgapdcpev_CTCVPGGLASC_salslpavpp -----

 -- NPSA [gnllunipsplQ4LDE5](#) [Eukaryota] Sushi, von Willebrand factor t
 ype A, EGF and pentraxin domain-containing protein 1 [Metazoa][Homo s
 apiens] (length=3571 residues). *****> PATTERN 1 Site : 1293-
 1304, Identity icvdgvagyr_CTCVKG FVGLHC_etevnecqsn -----
 ----- NPSA
[gnllunipsplQ5IJ48](#) [Eukaryota] Protein crumbs homolog 2 [Metazoa][Ho
 mo sapiens] (length=1285 residues). *****> PATTERN 1 Site : 11
 59- 1170, Identity pceggspaan_CSCLEGLAGQRC_qvptlpcean -----

 NPSA [gnllunipsplQ5T1H1](#) [Eukaryota] Protein eyes shut homolog [Metazo
 a][Homo sapiens] (length=3165 residues). *****> PATTERN 1 Site
 : 394- 405, Identity tckkcekdy_CSCISGFTEKNC_ekaidhckll -----

 ---- NPSA [gnllunipsplQ5VY43](#) [Eukaryota] Platelet endothelial aggrega
 tion receptor 1 [Metazoa][Homo sapiens] (length=1037 residues). ****
 *> PATTERN 1 Site : 380- 391, Identity slschpmnge_CSCLPGWA
 GLHC_nescpqdthg -----
 ----- NPSA [gnllunipsplQ701N2](#) [Eukaryota] Kerat
 in-associated protein 5-5 [Metazoa][Homo sapiens] (length=237 residue
 s). *****> PATTERN 1 Site : 202- 212, Identity yccqsscckp
 _CSCFSGCGSSC_cqsscypcc Site : 202- 213, Identity yccqssc
 kp_CSCFSGCGSSC_cqsscypccc -----
 ----- NPSA [gnllunipsplQ8TER0](#) [Eukar
 yota] Sushi, nidogen and EGF-like domain-containing protein 1 [Metazo
 a][Homo sapiens] (length=1413 residues). *****> PATTERN 1 Site
 : 297- 308, Identity dcvtgnpsy_CSCLSGFTGRRRC_hldvnecsq -----

 ---- NPSA [gnllunipsplQ99466](#) [Eukaryota] Neurogenic locus notch homol
 og protein 4 [Metazoa][Homo sapiens] (length=2003 residues). *****>
 PATTERN 1 Site : 103- 114, Identity spspltpsfl_CTCLPGFTGER
 C_qakledpcpp -----
 ----- NPSA [gnllunipsplQ9UM47](#) [Eukaryota] Neurogen
 ic locus notch homolog protein 3 [Metazoa][Homo sapiens] (length=2321
 residues). *****> PATTERN 1 Site : 873- 884, Identity sc
 qdgvgsfs_CSCLPGFAGPRC_ardvdeclsn -----
 ----- NPSA [gnllunipsplP07207](#)
 [Eukaryota] Neurogenic locus Notch protein [Metazoa][Drosophila melan
 ogaster] (length=2703 residues). *****> PATTERN 1 Site : 1046-
 1057, Identity scidgingyn_CSCLAGYSGANC_qyklnkcdsn -----
 ----- NPS

[A gnlunipsplP10040](#) [Eukaryota] Protein crumbs [Metazoa][Drosophila melanogaster] (length=2146 residues). *****> PATTERN 1 Site : 1900- 1911, Identity fnastgnnft_CTCVPGFEGPLC_dipfceitpc -----

 - [NPSA gnlunipsplP10079](#) [Eukaryota] Fibropellin-1 [Metazoa][Strongylocentrotus purpuratus] (length=1064 residues). *****> PATTERN 1
 Site : 276- 287, Identity icvdgvmfe_CTCLAGFTGVRV_evnidecasa
 Site : 542- 553, Identity tctdqngyv_CTCVPGYTGVLV_etdineca
 Site : 618- 629, Identity acmdvngfv_CTCLPGWEGTNC_eintdecass -----
 ----- [NPSA gnlunipsplP49013](#) [Eukaryota] Fibropellin-3 [Metazoa][Strongylocentrotus purpuratus] (length=570 residues). *****
 > PATTERN 1 Site : 276- 287, Identity icvdginqfa_CTCLPGYTGILC_etdinecass -----
 ----- [NPSA gnlunipsplP84031](#) [Eukaryota] U7-cte nitoxin-Pr1a [Metazoa][Phoneutria reidi] (length=58 residues). ****
 *> PATTERN 1 Site : 38- 48, Identity wfggkwhplg_CSCVYGDKYIC_ekkkkecpnv -----
 ----- [NPSA gnlunipsplQ5Y4U3](#) [Eukaryota] U9-aga toxin-Ao1a [Metazoa][Agelena orientalis] (length=110 residues). ****
 *> PATTERN 1 Site : 70- 80, Identity cscplfgalg_CSCIIGDAMVC_vrkkkecrts -----
 ----- [NPSA gnlunipsplQ6TMJ6](#) [Eukaryota] SrfA-induced gene K protein [Amoebozoa][Dictyostelium discoideum] (length=426 residues). *****> PATTERN 1 Site : 388- 399, Identity sspfiicngt_CTCLPGFSGNDC_tlcngngevi -----
 ----- [NPSA gnlunipsplO19045](#) [Eukaryota] Coagulation factor X [Metazoa][Oryctolagus cuniculus] (length=490 residues). *****> PATTERN 1 Site : 110- 121, Identity tckdglgmyt_CSCVEGYEQDC_epvtrklcsl -----
 ----- [NPSA gnlunipsplP02801](#) [Eukaryota] Metallothionein-1B [Metazoa][Equus caballus] (length=61 residues). *****> PATTERN 1 Site : 5- 15, Identity <mdpn_CSCVAGESCTC_agsckckqcr -----
 ----- [NPSA gnlunipsplP07589](#) [Eukaryota] Fibronectin [Metazoa][Bos taurus] (length=2478 residues). *****> PATTERN 1 Site : 2411- 2422, Identity qkeylgaics_CTCFQGQRGWRC_dncrrpgaep -----
 ----- [NPSA gnlunipsplP68303](#) [Eukaryota] Metallothionein-2 [Metazoa][Macaca fascicularis] (length=61 residues). *****> PATTERN 1 Site : 5- 15, Identity <mdpn_CSCVAGDSCTC_agsckckck -----
 ----- [NPSA gnlunipsplP68304](#) [Eukaryota] Metallothionein-2 [Metazoa][Chlorocebus aethiops] (length=61 residues). *****> PATTERN 1 Site : 5- 15, Identity <mdpn_CSCVAGDSCTC_agsckckck -----

----- [NPSA gnllunipsplP81282](#) [Eukaryota] Versican core protein [Metazoa][Bos taurus] (length=3381 residues). ****
 *> PATTERN 1 Site : 3098- 3109, Identity tcyptetsyv_CTCVPGYS
 GDRC_eldfdechn -----

----- [NPSA gnllunipsplP98167](#) [Eukaryota] SCO-spondin [Metazoa][Bos taurus] (length=5146 residues). *****> PATTERN
 1 Site : 3144- 3154, Identity gsavlrcpen_CSCVSLITNC_tswpckegqp -----

----- [NPSA gnllunipsplQ28295](#) [Eukaryota] von Willebrand factor [Metazoa][Canis lupus familiaris] (length=2813 residues). *****
 > PATTERN 1 Site : 255- 265, Identity pfvalcertl_CTCVQMECP
 C_avlleyarac Site : 2281- 2291, Identity wvpahqpcqi_CTCLSGR
 KVNC_tlqpcptara -----

----- [NPSA gnllunipsplQ28749](#) [Eukaryota] Fibronectin [Metazoa][Oryctolagus cuniculus] (length=100 residues). *****
 > PATTERN 1 Site : 33- 44, Identity qkeylgaics_CTCFGGQRG
 WRC_dncrrpgvep -----

----- [NPSA gnllunipsplQ28833](#) [Eukaryota] von Willebrand factor [Metazoa][Sus scrofa] (length=2482 residues). *****>
 PATTERN 1 Site : 1950- 1960, Identity wvpdhqpcqi_CTCLSGRRV
 C_tlqpcptara -----

----- [NPSA gnllunipsplQ28858](#) [Eukaryota] Versican core protein [Metazoa][Macaca nemestrina] (length=862 residues). **
 **> PATTERN 1 Site : 742- 753, Identity tcyptetsyv_CTCVPG
 YSGDQC_eldfdechn -----

----- [NPSA gnllunipsplQ5R6R1](#) [Eukaryota] Protein eyes shut homolog [Metazoa][Pongo abelii] (length=3164 residues)
 . *****> PATTERN 1 Site : 394- 405, Identity tckkwekdyh_C
 SCISGFTEKNC_ekaidhcrll -----

----- [NPSA gnllunipsplQ6AWT8](#) [Eukaryota] Guanine nucleotide-binding protein subunit gamma 3 [Viridiplantae][
 Arabidopsis thaliana] (length=251 residues). *****> PATTERN 1 S
 ite : 172- 182, Identity cccprlscps_CSCFRGCWCSC_pdmsscpcsc ----

----- [NPSA gnllunipsplQ6NPN4](#) [Eukaryota] LysM domain-containing GP I-anchored protein 3 [Viridiplantae][Arabidopsis thaliana] (length=42
 3 residues). *****> PATTERN 1 Site : 251- 262, Identity y
 altaghcvaq_CSCVLGSRSMYC_epasisvsccs -----

----- [NPSA gnllunipsplA2AVA0](#)
 [Eukaryota] Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1 [Metazoa][Mus musculus] (length=3567 resid
 ues). *****> PATTERN 1 Site : 1292- 1303, Identity tctdglas
 yr_CTCVKGYMGVHC_etdvnecqss -----

----- [NPSA gnllunipsplD3ZHH1](#) [Eukaryota] Delta-like protein 4 [Metazoa][Rattus norvegicus] (length=686 r
 esidues). *****> PATTERN 1 Site : 311- 322, Identity csns
 gprgyt_CTCLPGYTGEHC_elelskcasn -----

----- NPSA [gnllunipsplP04937](#) [Eukaryota] Fibronectin [Metazoa][Rattus norvegicus] (length=2477 residues). *****> PATTERN 1 Site : 2409- 2420, Identity qkeylgai cs_CTCFGGQRGWRC_dncrrpgae -----

----- NPSA [gnllunipsplP0C6B8](#) [Eukaryota] Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1 [Metazoa][Rattus norvegicus] (length=3564 residues). *****> PATTERN 1 Site : 1288- 1299, Identity tctdglasyr _CTCVSGYVGVHC_etdvnecqss -----

----- NPSA [gnllunipsplP11276](#) [Eukaryota] Fibronectin [Metazoa][Mus musculus] (length=2477 residues). *****> PATTERN 1 Site : 2409- 2420, Identity qkeylgai cs_CTCFGGQRGWRC_dncrrpgae -----

----- NPSA [gnllunipsplQ5ZQU0](#) [Eukaryota] Sushi, nidogen and EGF-like domain-containing protein 1 [Metazoa][Rattus norvegicus] (length=1403 residues). *****> PATTERN 1 Site : 297- 308, Identity dcvtgnpsyt_CSCLAGFTGRRRC_hldvnecash -----

----- NPSA [gnllunipsplQ61982](#) [Eukaryota] Neurogenic locus notch homolog protein 3 [Metazoa][Mus musculus] (length=2318 residues). *****> PATTERN 1 Site : 874- 885, Identity scqdgvgfsfs_CSCLDGFAGPRC_arvdclss -----

----- NPSA [gnllunipsplQ700K0](#) [Eukaryota] SCO-spondin [Metazoa][Rattus norvegicus] (length=5141 residues). *****> PATTERN 1 Site : 3149- 3159, Identity gsavilpcen_CSCVSLITNC_sswpceegap -----

----- NPSA [gnllunipsplQ70E20](#) [Eukaryota] Sushi, nidogen and EGF-like domain-containing protein 1 [Metazoa][Mus musculus] (length=1403 residues). *****> PATTERN 1 Site : 297- 308, Identity dcvtgnpsyt_CSCLAGFTGRRRC_hldvnecash -----

----- NPSA [gnllunipsplQ8CG65](#) [Eukaryota] SCO-spondin [Metazoa][Mus musculus] (length=4998 residues). *****> PATTERN 1 Site : 3000- 3010, Identity gsavhlp cen_CSCISGLITNC_sswpceegap -----

----- NPSA [gnllunipsplQ8VIK5](#) [Eukaryota] Platelet endothelial aggregation receptor 1 [Metazoa][Mus musculus] (length=1034 residues). *****> PATTERN 1 Site : 636- 647, Identity hsschpsdgt_CSCLAGWTGPDC_seacppghwg -----

----- NPSA [gnllunipsplQ9JI71](#) [Eukaryota] Delta-like protein 4 [Metazoa][Mus musculus] (length=686 residues). *****> PATTERN 1 Site : 311- 322, Identity csngpkgyt_CTCLPGYTGEHC_elglskcasn -----

----- NPSA [gnllunipsplQ9R0M0](#) [Eukaryota] Cadherin EGF LAG seven-pass G-type receptor 2 [Metazoa][Mus musculus] (length=2919 residues). *****> PATTERN 1 Site : 1817- 1828, Identity ycsndwdsys_CSCVLGYGDNC_tnvcdlnpce -----

----- [NPSA gnllunipsplQ9R172](#) [Eukaryota] Neurogenic locus notch homolog protein 3 [Metazoa][Rattus norvegicus] (length=2319 residues). *****> PATTERN 1 Site : 875- 886, Identity scqdg vgsfs_CSCLSGFAGPRC_ardvdeclss -----

----- [NPSA gnllunipsplA2RUV0](#) [Eukaryota] Neurogenic locus notch homolog protein 1 [Metazoa][Xenopus tropicalis] (length=2522 residues). *****> PATTERN 1 Site : 242- 253, Identity crqtdtsye_CTCLPGFSGQNC_eeidncpsn -----

----- [NPSA gnllunipsplA6MFK7](#) [Eukaryota] Venom prothrombin activator vestarin-D1 [Metazoa][Demansia vestigiata] (length=473 residues). *****> PATTERN 1 Site : 110- 121, Identity tckdgigsyt_CTCLAGYEGKNC_ehdllkscr -----

----- [NPSA gnllunipsplA6MFK8](#) [Eukaryota] Venom prothrombin activator vestarin-D2 [Metazoa][Demansia vestigiata] (length=471 residues). *****> PATTERN 1 Site : 110- 121, Identity tckdgigsyt_CTCLPGYEGKNC_ehvvvksr -----

----- [NPSA gnllunipsplB8JJ71](#) [Eukaryota] Protein eyes shut homolog [Metazoa][Danio rerio] (length=2904 residues). *****> PATTERN 1 Site : 484- 495, Identity sqttapphym_CTCLPGYTGPHYC_eaevnecds Site : 1284- 1295, Identity vcqdlvngfq_CSCVPGYFGPHC_nldvncdss -----

----- [NPSA gnllunipsplP21783](#) [Eukaryota] Neurogenic locus notch homolog protein 1 [Metazoa][Xenopus laevis] (length=2524 residues). *****> PATTERN 1 Site : 242- 253, Identity crqtdtsyd_CTCLPGFSGQNC_eeidncpsn -----

----- [NPSA gnllunipsplP25155](#) [Eukaryota] Coagulation factor X [Metazoa][Gallus gallus] (length=475 residues). *****> PATTERN 1 Site : 110- 121, Identity qckdglgsyt_CSCLDGYQGKNC_efvipkycki -----

----- [NPSA gnllunipsplP46530](#) [Eukaryota] Neurogenic locus notch homolog protein 1 [Metazoa][Danio rerio] (length=2437 residues). *****> PATTERN 1 Site : 242- 253, Identity cvqtsdttht_CSCLPGFTGQTC_ehvnvdctqh -----

----- [NPSA gnllunipsplQ1L658](#) [Eukaryota] Coagulation factor X isoform 2 [Metazoa][Pseudonaja textilis] (length=463 residues). *****> PATTERN 1 Site : 110- 121, Identity ickdgigsyt_CTCLSGYEGKNC_ervlykscr -----

----- [NPSA gnllunipsplQ1L659](#) [Eukaryota] Coagulation factor X isoform 1 [Metazoa][Pseudonaja textilis] (length=483 residues). *****> PATTERN 1 Site : 110- 121, Identity tckdgigsyt_CTCLSGYEGKNC_eyvlykscr -----

----- [NPSA gnllunipsplQ2PC93](#) [Eukaryota] SCO-spondin [Metazoa][Gallus gallus] (length=5255 residues). *****> PATTERN 1 Site : 1983- 1993, Identity

y gqrmklncl_CTCCLDGQPRRC_rhnpacsvsc Site : 3239- 3249, Identity gdvvpqgchn_CSCIAGRVTNC_sqedcgdvdg -----

----- NPSA gnlunipsplQ4QXT9 [Eukaryota] Coagulation factor X [Metazoa][Tropidechis carinatus] (length=483 residues). *****> PATTERN 1 Site : 110- 121, Identity tckdgigsyt_CTCCLAGYEGKNC_qvlyqscr -----

----- NPSA gnlunipsplQ56VR3 [Eukaryota] Venom prothrombin activator pseutarin-C catalytic subunit [Metazoa][Pseudonaja textilis] (length=467 residues). *****> PATTERN 1 Site : 110- 121, Identity ickdgigsyt_CTCCLSGYEGKNC_ervlykscr -----

----- NPSA gnlunipsplQ58L95 [Eukaryota] Venom prothrombin activator omicarin-C catalytic subunit [Metazoa][Oxyuranus microlepidotus] (length=467 residues). *****> PATTERN 1 Site : 110- 121, Identity tckdgigsyt_CTCCLFGYEGKNC_ervlykscr -----

----- NPSA gnlunipsplQ58L96 [Eukaryota] Venom prothrombin activator oscutarin-C catalytic subunit [Metazoa][Oxyuranus scutellatus] (length=467 residues). *****> PATTERN 1 Site : 110- 121, Identity tckdgigsyt_CTCCLSGYEGKNC_ervlykscr -----

----- NPSA gnlunipsplQ90953 [Eukaryota] Versican core protein [Metazoa][Gallus gallus] (length=3562 residues). *****> PATTERN 1 Site : 3278- 3289, Identity tcyprgsfyi_CTCCLPGFNGEQC_eldidecqs -----

----- NPSA gnlunipsplA3EXG6 [Viruses] Spike glycoprotein [Riboviria][Bat coronavirus HKU9] (length=1274 residues). *****> PATTERN 1 Site : 1240- 1252, Identity vimlmcmtnc_CSCFKGMCDRRC_cgsydsydd -----

NPSA gnlunipsplP59594 [Viruses] Spike glycoprotein [Riboviria][Human SARS coronavirus] (length=1255 residues). *****> PATTERN 1 Site : 1223- 1235, Identity tillcmtsc_CSCLKGACSCGSC_ckfdeddsep -----

----- NPSA gnlunipsplQ0Q475 [Viruses] Spike glycoprotein [Riboviria][Bat coronavirus 279/2005] (length=1241 residues). *****> PATTERN 1 Site : 1209- 1221, Identity tillcmtsc_CSCLKGACSCGSC_ckfdeddsep -----

----- NPSA gnlunipsplQ3I5J5 [Viruses] Spike glycoprotein [Riboviria][Bat coronavirus Rp3/2004] (length=1241 residues). *****> PATTERN 1 Site : 1209- 1221, Identity tillcmtsc_CSCLKGACSCGSC_ckfdeddsep -----

----- NPSA gnlunipsplQ3LZX1 [Viruses] Spike glycoprotein [Riboviria][Bat coronavirus HKU3] (length=1242 residues). *****> PATTERN 1 Site : 1210- 1222, Identity tillcmtsc_CSCLKGACSCGSC_ckfdeddsep -----

----- Total number of sequences in PA

TTINPROT run 1 : 72 Total number of patterns found in PATTINPROT run
1 : 78 Total number of selected patterns for pattern 1 : 78

Figure 1S. The original PattInProt analysis using the pattern C-[TS]-C-[FLIV]-X-G-X(4,6)-C.