

# **CSpritz - Accurate detection of protein disorder**

### **Version 1.2**

Title: PF3D7\_1227300.fasta emailaddress: mubasher.mohammed@su.se pid: 1976228859

**Status: finished** 

NB: The data will be kept for two weeks, after which it will be deleted without further notice.

### Available files: (help)

Your input parameters: HTML
Fasta sequences (amino, secondary structure, disorder, confidence): TXT
Disorder plot: PDF
Graph of PDB homologue(s) found:: PDF
Disorder Prediction (with disorder probability): TXT
Protein statistics: TXT

### Disordered residues and stats: (help)

MODIENTIVYERNINTYKENVERNYESPECHNILMY FYNPHEDFILLILITYMILEKYESYNKTHILLEEYTW   CCCCHHHHHHHHHHHHHHHCCCCCHHHHHHHHHHHHH	1	10	20	30	40	50	60
180   90   100   110   120   130							
10							
PRYVILSYTPLYVELDINNLYLFOMLLHKFNEKRRRENNMKKNVIDNNIYKKYRILMRINKKHGHVEAN HHHHHHHHCCHHHHHHCCCCHHHHHHHHHCCCCCCCCC	DDDDDDI	DDDD00000	000 <b>D</b> 0000000	0000000000	0000000000	000000DDDI	DDDDDDDD0000000
HARHHRHHCCCHHHRHHHCCCCHHHRHHHHCCCCCHHHRHHHHHH							
141							
141 150 160 170 180 190 200  LECVHKKNYEYMNINKLNKHMNICPNDNKCDSIYSLKKQTSEDIQKNEKQNDKLILPSFQGSLKKNIYM  EECCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC							
LECVEHENVEYNMINKLNKEHNICFNDNKCDSIYSLKKQTSEDIQKNEKQNDKLILPSFQGSLKKNIYM EECCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	000000	0000000D	0000000000	OOODDDDDDDI	וססססססססססס	וסססססססססס	
EECCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	141						
211 220 230 240 250 260 270  NNTYNRTICETNLFYKNNNNDHLKLKKNRSFITLCEDNLKKRKDIIKNMIKEMDNLYAHVYYWKYLTEIN. CCCCCEEBECCCCCCCCCCHHHCHCCCCCCCCCCHHHHHHHH							
211 220 230 240 250 260 270  NNTNRRICETNLFYKNNNNDBLKLKKNRSPIYLCEDNLKRRKDIIKNMIKEMDNLYAHVYYWKYLTELN CCCCCEEEECCCCCCCCCCCCCCCCCCCCCCCCCCC							
NNTINFICETNLFYKNNNNDHLKLKKNRSPIYLCEDNLKKRRDIIKAMIKEMDNLYAHVYWKYLTELN CCCCCEEEECCCCCCCCCCCCCCCCCCCCCCCCCCC	וטטטטטט	טטטטטטטטטט	וטטטטטטטטטט	וטטטטטטטטטט	וטטטטטטטטטט	וטטטטטטטטטטט	טטטטטטטטטטטטטטטטטט
CCCCCEEEECCCCCCCCCCCCCCCCCCHHCCCCCCCCCHHHHHH	211						
290   300   310   320   330   340							
KCNNNNNNDKNNNLEVNYAINKGQDIYLNIYKNIYNLYKYMNKKHKQNLEKMNMKKEQLSITHNNEMEN HHCCCCCCCCCCCCCCHHHHHCCCCHHHHHHHHHHHH							
KCNNNNNNDKNNNLEVNYAINKGQDIYLNIYKNIYNLYKYMNKKHKQNLEKMNMKKEQLSITHNNEMEN HHCCCCCCCCCCCCCCHHHHHCCCCHHHHHHHHHHHH		200	200	21.0	200	222	240
HHCCCCCCCCCCCHHHHHHCCCCHHHHHHHHHHHHHHH							
351         360         370         380         390         400         410           351         360         370         380         390         400         410           1NEKKIHKKNEEYHKNVNDMVFKHIAEKNALCRHIEHEMNIIQQINMKEYKENIFYIFSIISKNKQNLS         HHCCCCCCCCCHHHHHHHHHHHHHHHHHHHHHHHHHH							
INEKKIHKKNEEYHKNVNDMVFKHIAEKNALCRHIEHEMNIIQQINMKEYKENIFYIFSIISKNKQNLS HHCCCCCCCCCCHCCCHHHHHHHHHHHHHHHHHHHHH							
INEKKIHKKNEEYHKNVNDMVFKHIAEKNALCRHIEHEMNIIQQINMKEYKENIFYIFSIISKNKQNLS HHCCCCCCCCCCHCCCHHHHHHHHHHHHHHHHHHHHH	251	260	270	200	200	400	410
HHCCCCCCCCCHCCCHHRHHHHHHHHHHHHHHHHHHHH							
### ### ##############################							
421 430 440 450 460 470 480  PPLPVEEERDIQKEVRPYHNEEKKKKHNNNNNYNNICNNNNNVSNINNSGCNNFNNMCENMNDLSK CCCCCHHHHHHHHCCCCCCCCCCCCCCCCCCCCCCCC							
PPLPVEEERDIQKEVRPYHNEEKKKKHNNNNNYNNICNNNNNNNNINNSGCNNFNNMCENMNDLSK CCCCCHHHHHHHHHCCCCCCCCCCCCCCCCCCCCCCC	יסססססס	טטטטטטטטטט	DOOOOOOD	000000000000000000000000000000000000000	000000 <mark>0</mark> 000	DDD00000000	300000000000000000000000000000000000000
CCCCCHHHHHHHHCCCCCCCCCCCCCCCCCCCCCCCCC	421						
######################################							
491 500 510 520 530 540 550  KCKDIKENYIYNDKNNITMEDKIKTYYHIEECYYHYKYIIKCFHFKYLYELIKSKIMIYNYLHFHIKKS CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC							
KCKDIKENYIYNDKNNITMEDKIKTYYHIEECYYHYKYIIKCFHFKYLYELIKSKIMIYNYLHFHIKKS CCCCCCCCCCCCCCCCCCCCCCCCHHHHEEEEHHHHHHHH	וטטטטטט	טטטטטטטטטט	וטטטטטטטטטט	וטטטטטטטטטט	וטטטטטטטטטט	וטטטטטטטטטט	וטטטטטטטטטטטטטטטט
CCCCCCCCCCCCCCCCCCCHHHHHEEEEHHHHHHHHHH	491						
DDDDDDDDDDDDDDDDDDDDDDDDDDDDDOOOOOOOOO							
561 570 580 590 600 610 620  YNIMNVSIIFSLGEIYHIFEYDKKKNIQFQDNFFTTLRKMQIYNYNKYYQYMMNNEKGPDKYFNSFQIY  HHCCCEEEEEECHHHHHHHHHHCCCCCHHHHHHHHHH							
YNIMNVSIIFSLGEIYHIFEYDKKKNIQFQDNFFTTLRKMQIYNYNKYYQYNMNNEKGPDKYFNSFQIY HHCCCEEEEEECHHHHHHHHHHCCCCCHHHHHHHHHHH	וססססססו	וסססססססססס	וטסטטטטטטטטט	DD000000000	0000000000	0000000000	000000000000000000000000000000000000000
HHCCCEEEEECHHHHHHHHHCCCCCHHHHHHHHHHHHH	561	570	580	590	600	610	620
OOOOOOOOOOOOOOOODDDDDDDDDDDDDDDDDDDDDD	YNIMNV	SIIFSLGEI	YHIFEYDKKKI	NIQFQDNFFT'	TLRKMQIYNY	икуудуимии	EKGPDKYFNSFQIYE
631 640 650 660 670 680 690  EKKKKVTNVLEIQKDEDILKKDVNTNEWANPTGLKKKILHLLKGGTQINKDNEHMNKFVPIYNNSLLVN  HCCCCCCCEEECCCCCHHHCCCCCCCCCCCCHHHHHHHH							
EKKKKVTNVLEIQKDEDILKKDVNTNEWANPTGLKKKILHLLKGGTQINKDNEHMNKFVPIYNNSLLVN HCCCCCCEEEECCCCHHHCCCCCCCCCCCCCHHHHHHHH	000000	000000000	OOODDDDDDDI	DDDD000000	000000DD000	OODDDDDDDDI	ומממממממממממממ
HCCCCCCCEEECCCCCHHCCCCCCCCCCCCCCHHHHHHHH	631						
DDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD							
701 710 720 730 740 750 760  ILSGYNAHDIKEQKGDDSLKCEESFYFDENGYFKYKKKRKGAKEFFNHNYSSTYNYINNNTLYFMNNK HCCCCCCCHHHHCCCCCCCCCCCCCCCCCCCCCCCHHHHCCCC							
ILSGYNAHDIKEQKGDDSLKCEESFYFDENGYFKYKKRKRGAKEFFNHNYSSTYNYINNNTLYFMNNK HCCCCCCCHHHHCCCCCCCCCCCCCCCCCCCCCCCC	DDDDDDI	וססססססססס	ומממממממממ	ומממממממממ	ומממממממממ	ומממממממממ	ומממממממממממממ
HCCCCCCCHHHHCCCCCCCCCCCCCCCCCCCCCCCCCC	701	710	720	730	740	750	760
DDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD							
771 780 790 800 810 820 830 DENIYNNNNNNNDIKINKLTIKDKKLNCLILFINEDISNYTKENVYNNIFNISLTKMDFIFPTIKEQ CCHHHECCCCCCCCCCCCCCCCCCCCCCCCCCCCCHHHHHH							
DENIYNNNNNNNNDIKINKLTIKDKKLNCLILFINEDISNYTKENVYNNIFNISLTKMDFIFPTIKEQ CCHHHECCCCCCCCCCCCCCCCCCCCCCCCCCCHHHHHHH	DDDDDDI	וססססססססס	וסססססססססס	ומממממממממ	וסססססססססס	וסססססססססס	000000000DDD
CCHHHECCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	771	780	790	800	810	820	830
DDDDDDDDDDDDDDDDDDDDDDDDDDDOOOOOOOOODDDOOOO							
841 850 860 870 880 890 900 HVVKEMYKGKKEKLFSDDLYKKSIKNICNDISEKNESLSYGTIIITRHKNLKIAMQGKNNFKENHQTGD HHHHHHHCCCCCEEEECCCCCCCCCCCCCCCCCCCCCC		CCCCCCCC					
HVVKEMYKGKKEKLFSDDLYKKSIKNICNDISEKNESLSYGTIIITRHKNLKIAMQGKNNFKENHQTGD HHHHHHHCCCCCEEECCCCCCCCCCCCCCCCCCCCCCC					0000DDD0000	0000000000	000000000000000000000000000000000000000
HVVKEMYKGKKEKLFSDDLYKKSIKNICNDISEKNESLSYGTIIITRHKNLKIAMQGKNNFKENHQTGD HHHHHHHCCCCCEEECCCCCCCCCCCCCCCCCCCCCCC		מסססססססס	וטטטטטטטטטטט				
HHHHHHHCCCCCEEECCCCCCCCCCCCCCCCCCCCCCCC							
DDOOODDDDDDDDDDDDDDDDDDDDDDDDDDDDDOOODDDOODDDD	DDDDDDI 841	850	860	870	880	890	900

Total amino acids:	1162
Total % disorder:	65.74
Total no. of disordered regions > 30 amino acids:	7
Total no. of disordered regions > 50 amino acids:	4
Number of disordered segments:	28
Length distribution of segments (N to C	10 1 11 1 163 20 1 37 10 1 2 102 11 2 149 30 3 2 32 3 74 25 6 1 7 2
terminal order):	47 11

13/01/2022, 09:07 CSpritz PID: 1976228859

NNDNNNNKNNKNDNNNNNNNKDT KNKEENNT VVNNHNT.DNT.NVT DEDNEFEST.T.NMDEOVMETSHEST N 981 990 1000 1010 1020 1030 1040 EKMFPKKKKKIYIDIIFHVLIPKKINRKSFKIILFSLFKILDLCKLYNIYCINCSFPYVHDIVYKNKRP 1070 1080 1090 1100 1051 1060 IIYSFIYSFIMTIMKYIKKGOPEDHNNTVRIFHFIFPPLOFYEKLTTEOMENKNDGKNOFKENINKNYDN нинининининининини 1140 KENNINYDYIHNVDKTKISNITYFIDKLIHDLKEKYMLIESI 

### Disordered segment motifs: (help)

59 ELM motif(s) in total found in disordered residues

## SS Motif for disorder segment 1:

(start,end): (1,10)

seq : MVDIKNIVVR S.S. : СССИНИНИНИ alpha: 024789999 beta : 000000000 coil : 9752100000

### SS Motif for disorder segment 2:

(start,end): (52,62)

52 61 seq : KKYESYNKTIH S.S.: HHHHHCCCCCC alpha: 98776431122 beta : 0000000011 coil : 01123568765

#### SS Motif for disorder segment 3:

(start,end): (99,261)

118 128 138 : HKFNEKRKRENNMKKNVIDNNIYKKYRILNRINKNGHVEANKLHCVHHKNVEYMNINKLNKHMNICFNDN 5432466777765553333347887777765421000112111111111234444555544443332111 beta: 00000000000000000000000011111100003542244432000122221100000001121100 3466533212233445555552111110122467885335544456777643233333444455445788 coil : 188 198 178 208 KCDSIYSLKKOTSEDIOKNEKONDKLILPSFOGSLKKNIYMKNNTNRRICETNLFYKNNNNDHLKLKKNR 1123565543214567665544334442222112222222111000000011223211112344554321 beta: 00000001111000000000000010000000012332100135554322222100000100000 7776333345665322234445554446666777666544567888643345554467887654334678 coil : 248 SPIYLCEDNLKKRKDIIKNMIKE : СССССИНИНИНИНИНИНИНИНИ s.s. alpha: 01112577788888999999999 coil: 88665422211111000000000

### SS Motif for disorder segment 4:

(start,end): (282,301)

#### SS Motif for disorder segment 5:

(start,end): (330,366)

Graph of Probability(helix, strand, coil)
PDF
Amino SS Probability(helix, strand, coil)
TXT

Linear motif(s):

Graph of Probability(helix, strand, coil)
PDF
Amino SS Probability(helix, strand, coil)

Linear motif(s):

Graph of Probability(helix, strand, coil)
PDF
Amino SS Probability(helix, strand, coil)
TXT

#### Linear motif(s):

residues (156,165)

residues (151,154) LIG\_SH2\_GRB2 =

LIG MAPK 1=

KLNKHMNICF residues (170,173)  $LIG_PDZ_3 = CDSI$ residues (178,184)  $LIG_FHA_2 =$ KOTSEDI residues (181,184) LIG PDZ 3 = SEDIresidues (186,196) LIG\_MAPK\_1 = KNEKQNDKLIL residues (191,194) **LIG\_PDZ\_3 = NDKL** residues (193,212) LIG\_IQ = KLILPSFOGSLKKNIYMKNN residues (224,227) LIG\_SH2\_GRB2 = residues (229,232) LIG PDZ 3 = NDHLresidues (233,243) LIG\_MAPK\_1 = KLKKNRSPIYL residues (236,241) LIG\_WW\_4 = **KNRSPI** residues (242,245) LIG\_SH2\_STAT5 =

Graph of Probability(helix, strand, coil)
PDF
Amino SS Probability(helix, strand, coil)
TXT

 $LIG_PDZ_3 = EDNL$ 

 $LIG_PDZ_3 = KDII$ 

residues (245,248)

residues (252,255)

#### Linear motif(s):

Graph of Probability(helix, strand, coil)
PDF
Amino SS Probability(helix, strand, coil)

13/01/2022, 09:07

CSpritz PID: 1976228859 TXT 330 339 349 359 : EKMNMKKEQLSITHNNEMENIINEKKIHKKNEEYHKN sea : нинининноссссссинининносссссссинссс Linear motif(s): alpha: 9998767654322234678887543221111455322 beta: 00000000001000000000000011110000001 residues (335,341) LIG\_MAPK\_1 = coil: 0001222345666665321112455667777544665 KKEQLSI  $LIG_PDZ_3 = KEQL$ residues (336,339) residues (347,350)  $LIG_PDZ_3 = MENI$ SS Motif for disorder segment 6: Graph of Probability(helix, strand, coil) (start,end): (375,384) PDE Amino SS Probability(helix, strand, coil) seq : IAEKNALCRH s.s. : нинининин alpha: 9998888888 Linear motif(s): beta : 0000000000 coil: 0001111111 Graph of Probability(helix, strand, coil) Amino SS Probability(helix, strand, coil) TXT SS Motif for disorder segment 7: Linear motif(s): (start,end): (416,517) residues (418.424) LIG SH3 3 = **LSLPPLP** 455 ONLSLPPLPVEEERDIOKEVRPYHNEEKKKKHNNNNNYNNICNNNNNNVSNINNSGCNNFNNMCENMN residues (421,424) LIG WW 2 = PPLPresidues (424,427) LIG\_TRAF2\_1 = alpha: 10000000456777654321001112122221111234444432211111122211111112223444432 residues (438,441) LIG\_SH2\_GRB2 = 8778999994332223445677877777766678887544322356776544456788877666544566 505 495 515 : DLSKHKCKDIKENYIYNDKNNITMEDKIKTYY residues (454,457) LIG\_SH2\_GRB2 = VNNI alpha: 33333222333333221111111566533333 residues (495,499) beta: 00000010000112332100022000123443 LIG\_APCC\_KENbox\_2 = IKENY coil: 66556666555544345777766333332223 residues (499,502) LIG\_SH2\_STAT5 = YIYN residues (506,512) LIG\_FHA\_2 = NITMEDK residues (510,513) LIG\_PDZ\_3 = EDKI Graph of Probability(helix, strand, coil) SS Motif for disorder segment 8: (start,end): (579,589) Amino SS Probability(helix, strand, coil) TXT 579 588 : FEYDKKKNIOF : нинсссссини Linear motif(s): s.s. alpha: 65433332456 beta: 21100000111 residues (583,589) LIG MAPK 1 = coil: 12366666432 KKKNIOF Graph of Probability(helix, strand, coil) Amino SS Probability(helix, strand, coil) TXT SS Motif for disorder segment 9: Linear motif(s): (start,end): (608,756) residues (622,625) LIG\_SH2\_STAT5 = 627 YYOYNMNNEKGPDKYFNSFOIYEEKKKKVTNVLEIOKDEDILKKDVNTNEWANPTGLKKKILHLLKGGTO residues (632,642) LIG\_MAPK\_1 = KKKKVTNVLEI alpha: 4443322110001222234445554211111333322245654432222332012346677777521011 residues (635,641) LIG\_FHA\_1 = coil: 3223467788987655653221234676555222356643344556776657877643211111368875 KVTNVLE residues (645,648) LIG\_PDZ\_3 = DEDI 707 717 : INKDNEHMNKFVPIYNNSLLVNNILSGYNAHDIKEOKGDDSLKCEESFYFDENGYFKYKKKRKGAKEFFN residues (665,671) LIG\_MAPK\_1 = KKKILHL 111123433211122224566655431114455554211112222333222211222322222345542 residues (667,671) LIG\_CYCLIN\_1 = KILHL coil: 5677654566565446642212334577654433456888765444444567775434455666532236 residues (692,695) LIG SH2 GRB2 = : HNYSSTYNY s.s. : cccccccc residues (716,719) LIG PDZ 3 = DDSLalpha: 111122233 LIG\_RB = LKCEE residues (719,723) beta : 111111222 residues (723,727) coil: 776665543 LIG\_BRCT\_BRCA1\_1 = ESFYF residues (726,729) LIG\_SH2\_STAT5 =

SS Motif for disorder segment 10:

(start,end): (767,796)

old.protein.bio.unipd.it/cspritz/work/pid\_1976228859/batch/PF3D7\_1227300.fasta\_cspritz.html

YFKY

LIG SH2 STAT5 =

residues (732,735)

Graph of Probability(helix, strand, coil)

Amino SS Probability(helix, strand, coil)

13/01/2022, 09:07 CSpritz PID: 1976228859

TXT 767 776 786 seq : NNKMDENIYNNNNNNNNNNNNLKINKLTIKDK : CCCCCCHHHECCCCCCCCCEECCCCCCC Linear motif(s): alpha: 322333433332221111111111111111 beta: 100001233321111111244433343210 residues (771.774) LIG PDZ 3 = DENIcoil: 576655322245666677643344544677 residues (775,778) LIG\_SH2\_GRB2 = residues (790,796) LIG FHA 2 = KLTIKDK SS Motif for disorder segment 11: Graph of Probability(helix, strand, coil) (start,end): (810,812) PDF Amino SS Probability(helix, strand, coil) TXT seq : SNY s.s. : ccc alpha: 333 **Linear motif(s):** beta : 001 Graph of Probability(helix, strand, coil) PDF SS Motif for disorder segment 12: Amino SS Probability(helix, strand, coil) (start.end): (847.878) TXT **Linear motif(s):** seq : YKGKKEKLFSDDLYKKSIKNICNDISEKNESL residues (851,854) LIG\_PDZ\_3 = KEKL alpha: 520000000112222222221011211110 beta: 00001355421001111222221100000122 residues (853,858) LIG\_14-3-3\_3 = KLFSDD coil: 36887533467766666554456877678766  $LIG_PDZ_3 = SDDL$ residues (856.859) residues (875,878) LIG\_PDZ\_3 = NESL SS Motif for disorder segment 13: Graph of Probability(helix, strand, coil) (start.end): (882.884) Amino SS Probability(helix, strand, coil) TXT seq : TII S.S. : CEE alpha: 000 Linear motif(s): beta : 488 coil : 510 Graph of Probability(helix, strand, coil) SS Motif for disorder segment 14: Amino SS Probability(helix, strand, coil) (start.end): (889.962) 229 292 908 918 928 938 948 Linear motif(s): : KNLKIAMOGKNNFKENHOTGDNNNDNNNNKNNKNDNNNNNNKDIKNKEENNIYYNNHNLDNLNYIPFD residues (889,893) LIG\_CYCLIN\_1 = KNLKI coil: 6642234678886555667888876666666666666665544444334543456666666556766 residues (901.905) LIG\_APCC\_KENbox\_2 = FKENH seq : NFFF S.S. : CCCC residues (943,946) LIG\_SH2\_GRB2 = alpha: 2224 beta : 0111 residues (949,952) **LIG\_PDZ\_3 = LDNL** coil: 6654 residues (954,957) LIG\_SH2\_STAT5 = SS Motif for disorder segment 15: Graph of Probability(helix, strand, coil) (start,end): (968,992) Amino SS Probability(helix, strand, coil) : PEOVMRTSHFSINEKMFPKKKKKKI : HHHHHHCCCCECCCCCCCCCC S.S. Linear motif(s): alpha: 6787642122223332111112110 beta: 0000000023332100000011135 residues (968,971) LIG\_PDZ\_3 = PEQV coil: 3111246754334456877766653 SS Motif for disorder segment 16: Graph of Probability(helix, strand, coil) (start,end): (1004,1009) PDF Amino SS Probability(helix, strand, coil) : KKINRK · cccccc alpha: 333244 **Linear motif(s):** beta : 000000 coil: 656655 SS Motif for disorder segment 17: Graph of Probability(helix, strand, coil) (start,end): (1070,1076) Amino SS Probability(helix, strand, coil) TXT seq : GQPEDHN s.s. : cccccc Linear motif(s): alpha: 0001110 beta : 0000000

coil: 8987778

Graph of Probability(helix, strand, coil)
PDF
Amino SS Probability(helix, strand, coil)
TXT

#### **Linear motif(s):**

Graph of Probability(helix, strand, coil)

Amino SS Probability(helix, strand, coil)

**Linear motif(s):** 

SS Motif for disorder segment 18: (start,end): (1093,1139)

SS Motif for disorder segment 19:

(start,end): (1151,1161)

1151 1160
seq : DLKEKYMLIES
S.S. : HHHHHHHHHCC
alpha: 87665444421
beta : 00000123331
coil : 12333432236

(c) Ian Walsh, Alberto J. Martin and Silvio Tosatto for Biocomputing UP, 11/2016