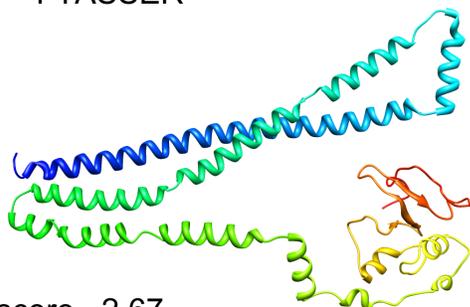


Supplementary figures

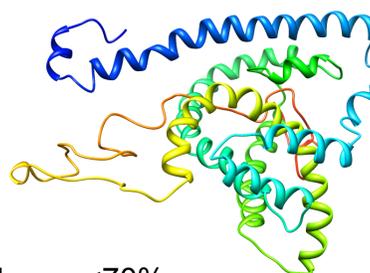
Figure S1. Ribbon diagrams of the gp44 structure predictions. Predictions from I-TASSER (A), Phyre2 (B), IntFOLD5 (C) and RaptorX (D), rainbow colored from the N-terminus (blue) to the C-terminus (red). The reliability measure given by the program for each prediction is listed.

A I-TASSER



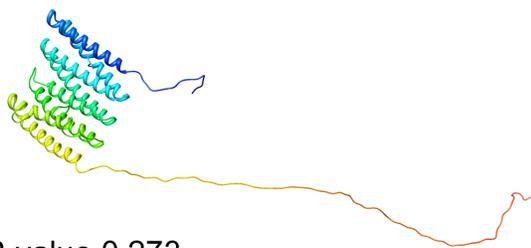
C-score -2.67

B Phyre2



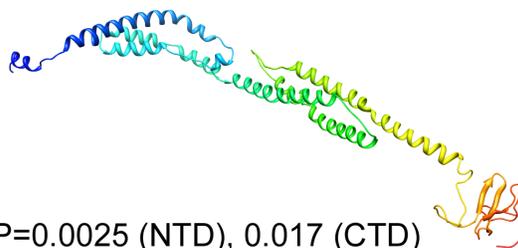
Confidence <70%

C IntFOLD



P-value 0.273

D RaptorX



P=0.0025 (NTD), 0.017 (CTD)

Figure S2. Mass spectrometry analysis. (A-C) LC-MS analysis of the high molecular weight band corresponding to full-length gp44 (A), the intermediate molecular weight band corresponding to gp44-NTD (B) and the low molecular weight band corresponding to gp44-CTD (C), extracted from SDS-PAGE gel of the SEC-purified material (Fig. 2A). The peptides identified in each band are highlighted in yellow over the gp44 sequence. (Methionines and cysteines are shown in green.) (D) Histogram of the number of measured spectra corresponding to the gp44-NTD (blue) and gp44-CTD (orange) in the three bands.

A gp44 high molecular weight (HMW) band

83 exclusive unique peptides, 114 exclusive unique spectra, 391 total spectra, 279/337 amino acids (83% coverage)

```

M P N K N T Q E Y W E E R G R K A I E N E L K R D K T K A E E I E R I L N M M I K R I E K E I N A F I V K Y G D F A G V
T L Q E A Q K I I D E F D V K A F Q E E A K R L V E N K D F S D R A N E E L K K Y N T K M Y V S R E Q M L K I Q I E F L
I A Y A T A Q T E L S M R E Y F E S T A Y R V F S D Q A G I L G E G V Q V A K E V I D T I V D T Q F H G V V W S E R L W
T N T E A M K Q E V E E I I A N V V I R G R H P N E Y V K D M R K H L N K F E G T A R Q K T A A I K S L L Y T E S A R V
H A Q S S I D S M K E I S P E G Y Y M Y I A K I D N R T T K V C K G L N G E I F K V K D A K I G V N F Y P M H I N C R S
D E A L L P K S M W P K K P S K K R K T K Y F G G K V K S G D H H H H H H

```

B gp44 intermediate molecular weight (IMW) band

72 exclusive unique peptides, 95 exclusive unique spectra, 347 total spectra, 279/337 amino acids (83% coverage)

```

M P N K N T Q E Y W E E R G R K A I E N E L K R D K T K A E E I E R I L N M M I K R I E K E I N A F I V K Y G D F A G V
T L Q E A Q K I I D E F D V K A F Q E E A K R L V E N K D F S D R A N E E L K K Y N T K M Y V S R E Q M L K I Q I E F L
I A Y A T A Q T E L S M R E Y F E S T A Y R V F S D Q A G I L G E G V Q V A K E V I D T I V D T Q F H G V V W S E R L W
T N T E A M K Q E V E E I I A N V V I R G R H P N E Y V K D M R K H L N K F E G T A R Q K T A A I K S L L Y T E S A R V
H A Q S S I D S M K E I S P E G Y Y M Y I A K I D N R T T K V C K G L N G E I F K V K D A K I G V N F Y P M H I N C R S
D E A L L P K S M W P K K P S K K R K T K Y F G G K V K S G D H H H H H H

```

C gp44 low molecular weight (LMW) band

29 exclusive unique peptides, 39 exclusive unique spectra, 88 total spectra, 143/337 amino acids (42% coverage)

```

M P N K N T Q E Y W E E R G R K A I E N E L K R D K T K A E E I E R I L N M M I K R I E K E I N A F I V K Y G D F A G V
T L Q E A Q K I I D E F D V K A F Q E E A K R L V E N K D F S D R A N E E L K K Y N T K M Y V S R E Q M L K I Q I E F L
I A Y A T A Q T E L S M R E Y F E S T A Y R V F S D Q A G I L G E G V Q V A K E V I D T I V D T Q F H G V V W S E R L W
T N T E A M K Q E V E E I I A N V V I R G R H P N E Y V K D M R K H L N K F E G T A R Q K T A A I K S L L Y T E S A R V
H A Q S S I D S M K E I S P E G Y Y M Y I A K I D N R T T K V C K G L N G E I F K V K D A K I G V N F Y P M H I N C R S
D E A L L P K S M W P K K P S K K R K T K Y F G G K V K S G D H H H H H H

```

D Spectral counts

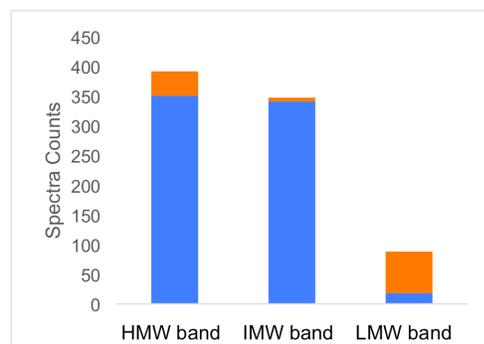


Figure S3. Effect of proteinase K treatment. Agarose gel electrophoresis of 25 nM *cl-cro* DNA incubated with varying concentrations of gp44 followed by T5 exonuclease treatment with (left) and without (right) proteinase K (PrK) treatment. The -fold molar excess of protein to DNA is indicated above the gel. The lane labeled, DNA has no gp44 and no exonuclease; lane 0 has T5 exonuclease, but no gp44. The arrows indicate the position of the *cl-cro* DNA band (CC) and the DNA-gp44 complexes that did not enter the gel (*).

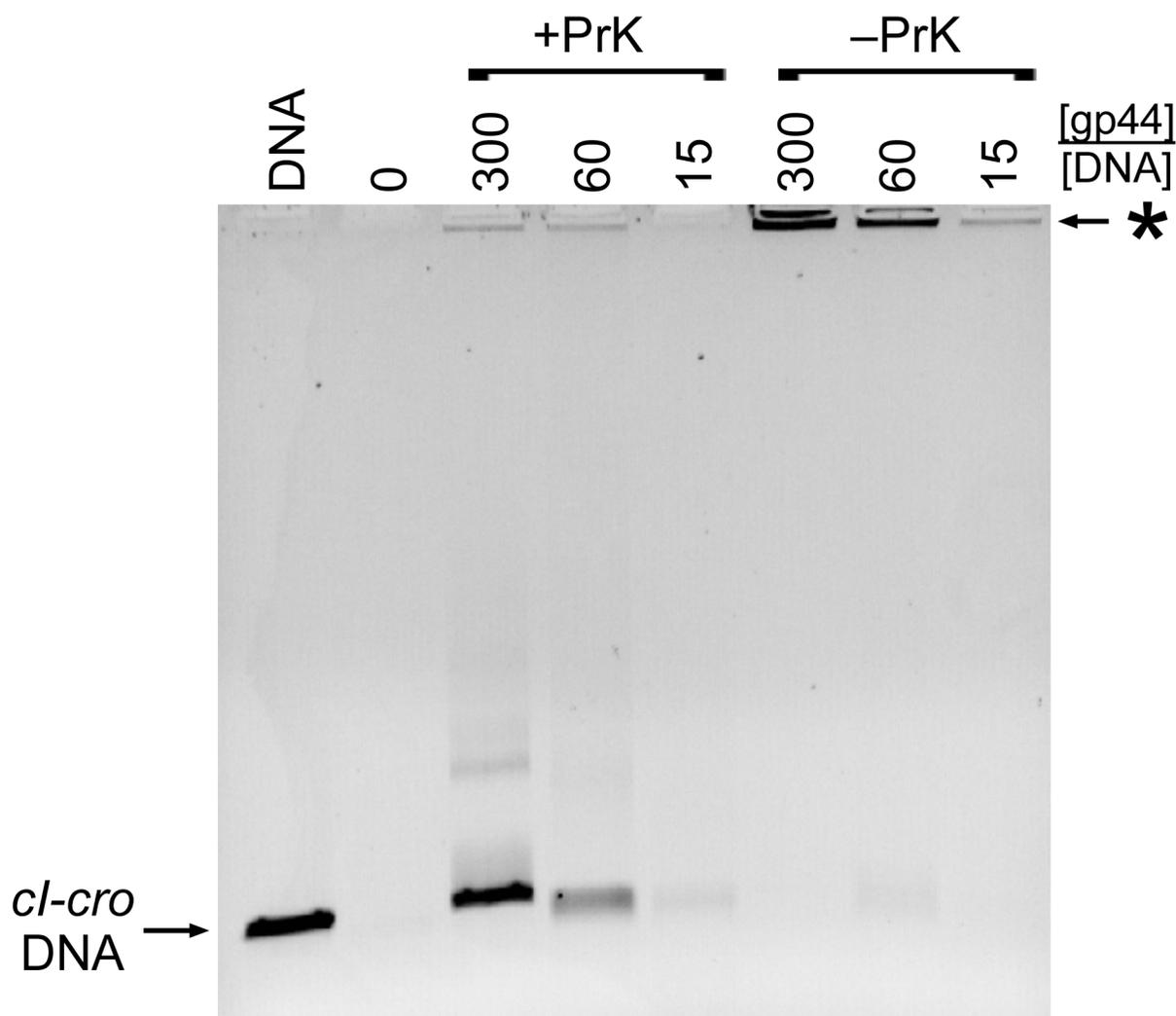


Figure S4. Protection of non-specific DNA. Agarose gel electrophoresis of 25 nM tail fiber (TF) DNA incubated with varying concentrations of gp44 or CI, after incubation with T5 exonuclease, followed by proteinase K treatment, showing that gp44, but not CI, protects TF DNA against T5 degradation. The -fold molar excess of protein to DNA is indicated above the gel. The lane labeled DNA has no gp44 or CI and no T5 exonuclease; lane 0 has T5 exonuclease, but no gp44 or CI.

