nature structural & molecular biology

Supplementary information

https://doi.org/10.1038/s41594-024-01355-x

Resolving chaperone-assisted protein folding on the ribosome at the peptide level

In the format provided by the authors and unedited

Name	DHFR length (aa)	Amino acid sequence
1-37^{RNC}	1-37	MSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFICTTGKLPVPWPTLVTTLTYGVL CFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILG HKLEYNFNSHNVYITADKQKNGIKAYFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSVL SKDPNEKRDHMVLLEDVTAAGITHGMDELYK**GSGS**LEVLFQGP**GSGSGSGSGSGSGSGSGSGSGSGSGS** **ENLYFQGSG**MISLIAALAVDRVIGMENAMPWNLPADLAWFKRNTLNWWWPRIRGPPGS
1-64^{RNC}	1-64	MSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFICTTGKLPVPWPTLVTTLTYGVL CFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILG HKLEYNFNSHNVYITADKQKNGIKAYFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSVL SKDPNEKRDHMVLLEDVTAAGITHGMDELYK**GSGS**LEVLFQGP**GSGS**MISLIAALAVDRVIGMENAMPW NLPADLAWFKRNTLNKPVIMGRHTWESIGRPLPGRKNIILSSWWWPRIRGPPGS
1-106^{RNC}	1-106	MSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFICTTGKLPVPWPTLVTTLTYGVL CFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILG HKLEYNFNSHNVYITADKQKNGIKAYFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSVL SKDPNEKRDHMVLLEDVTAAGITHGMDELYK**GSGS**LEVLFQGP**GSGS**MISLIAALAVDRVIGMENAMPW NLPADLAWFKRNTLNKPVIMGRHTWESIGRPLPGRKNIILSSQPGTDDRVTWVKSVDEAIAACGDVPEI MVIGGGRVYEQFLPKWWWPRIRGPPGS
1-126^{RNC}	1-126	MSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFICTTGKLPVPWPTLVTTLTYGVL CFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILG HKLEYNFNSHNVYITADKQKNGIKAYFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSVL SKDPNEKRDHMVLLEDVTAAGITHGMDELYK**GSGS**LEVLFQGP**GSGS**MISLIAALAVDRVIGMENAMPW NLPADLAWFKRNTLNKPVIMGRHTWESIGRPLPGRKNIILSSQPGTDDRVTWVKSVDEAIAACGDVPEI MVIGGGRVYEQFLPKAQKLYLTHIDAEVEGDTHFPWWWPRIRGPPGS
1-126^{RNC} mutant (V40A, V75H, I91L)	1-126	MSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFICTTGKLPVPWPTLVTTLTYGVL CFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILG HKLEYNFNSHNVYITADKQKNGIKAYFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSVL SKDPNEKRDHMVLLEDVTAAGITHGMDELYK**GSGS**LEVLFQGP**GSGS**MISLIAALAVDRVIGMENAMPW NLPADLAWFKRNTLNKPAIMGRHTWESIGRPLPGRKNIILSSQPGTDDRVTWHKSVDEAIAACGDVPEL MVIGGGRVYEQFLPKAQKLYLTHIDAEVEGDTHFPWWWPRIRGPPGS
FL+28^{RNC}	1-159	MSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFICTTGKLPVPWPTLVTTLTYGVL CFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILG HKLEYNFNSHNVYITADKQKNGIKAYFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSVL SKDPNEKRDHMVLLEDVTAAGITHGMDELYK**GSGS**LEVLFQGP**GSGS**MISLIAALAVDRVIGMENAMPW NLPADLAWFKRNTLNKPVIMGRHTWESIGRPLPGRKNIILSSQPGTDDRVTWVKSVDEAIAACGDVPEI MVIGGGRVYEQFLPKAQKLYLTHIDAEVEGDTHFPDYEPDDWESVFSEFHDADAQNSHSYCFEILERR**G** **SGGDLGSGGYLGSGGLDGSGGVEGSGGFL**WWWPRIRGPPGS
FL+38^{RNC}	1-159	MSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFICTTGKLPVPWPTLVTTLTYGVL CFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILG HKLEYNFNSHNVYITADKQKNGIKAYFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSVL SKDPNEKRDHMVLLEDVTAAGITHGMDELYK**GSGS**LEVLFQGP**GSGS**MISLIAALAVDRVIGMENAMPW NLPADLAWFKRNTLNKPVIMGRHTWESIGRPLPGRKNIILSSQPGTDDRVTWVKSVDEAIAACGDVPEI MVIGGGRVYEQFLPKAQKLYLTHIDAEVEGDTHFPDYEPDDWESVFSEFHDADAQNSHSYCFEILERRG SGGDLGSGGYLGSGGLDGSGGVEGSGGFLWWWPRIRGPPGS
FL+58^{RNC}	1-159	MSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFICTTGKLPVPWPTLVTTLTYGVL CFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILG HKLEYNFNSHNVYITADKQKNGIKAYFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSVL SKDPNEKRDHMVLLEDVTAAGITHGMDELYK**GSGS**LEVLFQGP**GSGS**MISLIAALAVDRVIGMENAMPW NLPADLAWFKRNTLNKPVIMGRHTWESIGRPLPGRKNIILSSQPGTDDRVTWVKSVDEAIAACGDVPEI MVIGGGRVYEQFLPKAQKLYLTHIDAEVEGDTHFPDYEPDDWESVFSEFHDADAQNSHSYCFEILERR**G** **SGGDLGSGGYLGSGGLDGSGGVEGSGGFLGSGGLDGSGGLYGSGGVEGS**WWWPRIRGPPGS
FL+58(GS)^{RNC}	1-159	MSKGEELFTGVVPILVELDGDVNGHKFSVRGEGEGDATNGKLTLKFICTTGKLPVPWPTLVTTLTYGVL CFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGTYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILG HKLEYNFNSHNVYITADKQKNGIKAYFKIRHNVEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSVL SKDPNEKRDHMVLLEDVTAAGITHGMDELYK**GSGS**LEVLFQGP**GSGS**MISLIAALAVDRVIGMENAMPW NLPADLAWFKRNTLNKPVIMGRHTWESIGRPLPGRKNIILSSQPGTDDRVTWVKSVDEAIAACGDVPEI MVIGGGRVYEQFLPKAQKLYLTHIDAEVEGDTHFPDYEPDDWESVFSEFHDADAQNSHSYCFEILERR**G** **SGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGS**WWWPRIRGPPGS

Supplementary Table 1. DHFR RNC sequences.

The N-terminal muGFP (green), functioning as an affinity tag, precedes a 3C cleavage sequence (blue) that is flanked on either side by $(GS)_2$ linkers (bold). The subsequent DHFR sequence fragment is followed by a C-terminal stall-inducing sequence derived from *M. succiniciproducens* SecM[41]. Structural analysis of ribosomes stalled on wild-type *E. coli* SecM indicates that various stalling scenarios are possible[65]. When describing our RNCs we assume that the ribosome stalls at the first available opportunity, with the penultimate $tRNA^{Pro}$ in the A-site and the first 8 residues from the stall-inducing sequence occupying the exit tunnel. In the full-length RNCs an additional disordered linker precedes the stall sequence (bold). In 1-37^{RNC}, the GS linker following the 3C cleavage sequence is extended to improve capture by the GFP affinity resin. Sites of point mutations are underlined.