Electronic Supplementary Material (ESI) for Nanoscale.

Supporting Information

Exploring the conformational dynamics of the SARS-CoV-2 SL4 hairpin by combining optical tweezers and base analogues

Vinoth Sundar Rajan[a,b,†], Anna Wypijewska del Nogal[b,†], Sune Levin[a], L. Marcus Wilhelmsson[b,*] and Fredrik Westerlund[a,*]

[†] These authors contributed equally to this work.

* Fredrik Westerlund and L. Marcus Wilhelmsson.
Email: fredrik.westerlund@chalmers.se and marcus.wilhelmsson@chalmers.se

This PDF file includes:

Supporting Information Text

1. Synthesis of SL4 constructs

Lyophilized oligonucleotides were dissolved in RNase-free water to a final concentration of 100 μM. Next, a digoxigenin-labelled tail of ~50 nt was added to the 3' end of the Oligo3. The 3' modification was performed at 37 °C using a Terminal Transferase kit (Merck, Ref. #: 3333566001) and at a digoxigenin-11-dUTP (DIG-dUTP; Roche, Ref. #: 11558706910) to dATP (Sigma-Aldrich, Ref. #: 11140965001) ratio of 1:10. The reaction was stopped after 15 min by adding 0.5 M EDTA, pH 8.0, yielding 5.0 mM Oligo3 with the 3' DIG-dUTP tail. The obtained product was purified at room temperature using QIAquick Nucleotide Removal Kit (Qiagen, Ref. #: 28304) with 30 μL of RNase-free water for elution. The purified tailed Oligo3 was subsequently diluted in RNase-free water to a concentration of 2.0 mM. To assemble the construct, three samples were annealed in parallel in a PCR machine (Fig. 2D): Oligo1 with Oligo4, Oligo2 alone, and the 3' DIG-dUTP tailed Oligo3 with Oligo4, in the splint buffer (33 mM Tris-HCl, 167 mM NaCl and 1.0 mM EDTA; pH 7.4). The annealing comprised the following steps: heating at 90 °C for 1.0 min, cooling from 90-80 °C in 10 sec, heating at 80 °C for 10 sec and cooling from 80-40 °C at 0.5 °C/10 sec. Next, the three samples were mixed at an equimolar ratio and cooled from 40-10 °C at 0.5 °C/20 sec. Finally, the two nicks in the construct (between Oligo1 and Oligo2, and between Oligo2 and the 3' DIG-dUTP tailed Oligo3; Fig. 2D) were ligated in the ligation buffer (50 mM Tris-HCl, 10 mM $MgCl_2$, 10 mM DTT, 1.5 mM ATP, pH 7.5), using T4 DNA Ligase (NEB, Ref. #: M0202S). Ligation was performed at 16 °C for 16 h and it was terminated by enzyme deactivation at 65 °C for 10 min. The construct was purified at room temperature using QIAquick Nucleotide Removal Kit (Qiagen, Ref. #: 28304) with 30 μL of RNase-free water for elution and annealed in the splint buffer as described above. All the sample preparation steps were carried out in RNase-free PCR tubes. The final construct at 0.10 mM concentration was aliquoted and stored at -80 °C until its use in the experiments.

The concentrations of the SL4 constructs were determined using the Beer-Lambert law and the following molar absorptivities at 260 nm: 1,405,100 M^{-1} cm^{-1} for WT, 1,404,500 M^{-1} cm^{-1} for UC, 1,408,300 M^{-1} cm^{-1} for US, 1,411,600 M^{-1} cm^{-1} for UD, and 1,411,600 M^{-1} cm^{-1} for LD. They were calculated as the sum of molar absorptivities of the 5' DNA (519,200 M^{-1} cm^{-1}), RNA (366,700 M^{-1} cm^{-1} for WT, 366,100 M^{-1} cm^{-1} for UC, 369,900 M^{-1} cm^{-1} for US, 373,100 M^{-1} cm^{-1} for UD, and 373,100 M^{-1} cm^{-1} for LD), and 3' DNA (519,200 M^{-1} cm^{-1}) fragments. For the DNA fragments, the molar absorptivity was approximated by a linear combination of molar absorptivities of individual nucleotides at 260 nm (ε_A = 15,300 M^{-1} cm^{-1}, ε_C = 7,400 M^{-1} cm^{-1}, ε_G = 11,800 M^{-1} cm^{-1}, ε_T = 9,300 M^{-1} cm^{-1})[1] and multiplied by 0.9 to account for base-stacking interactions. For the RNA fragment, the molar absorptivity was calculated using OligoAnalyzer (Integrated DNA Technologies. http://eu.idtdna.com/calc/analyzer (accessed on September 23 and December 30, 2021)) with the tC^O BA replaced by its unmodified counterpart (C), and next by correcting for the difference in molar absorptivity at 260 nm between C and tC^O (ε_{tCO} = 11,000 M^{-1} cm^{-1}).[2] Since the RNA sequence yields a double-helix hairpin, the calculated value was multiplied by 0.9 to account for the base-stacking interactions within the hairpin.[3]

2. Classification and analysis of the FDCs

The signature of each FDC acquired in the unfolding process (*i.e.*, pull curve) fell into one of the following categories:

- **Native**. In a native pull curve, unfolding transitions (rips) occurred between a fully folded molecule (hairpin, F) and its unfolded state (single strand, U), either directly (reflecting one-step unfolding) or through a partly folded state (intermediate, I; reflecting two-step unfolding). Moreover, both the one-step transitions and either step of the two-step transitions was reversible, which could be seen as a characteristic unfolding-folding pattern (multiple rips and zips between two states within a few piconewtons) called 'hopping'.[3-5]
- **UC's native**. For UC SL4, only two-step transitions were recorded and at significantly higher forces compared to the WT and tC^O-containing constructs.
- **Low**. A pull curve categorized as 'low' has the same features as the native curves, except that the first rip occurred at a significantly lower force (<8 pN). Also here, hopping between conformational states was observed.
- **Rescue**. In some pull curves, we observed a zip (folding) followed by a rip (unfolding). This type of pull curve has been reported in the literature and has been termed rescue.[6, 7] The rescue transition happens due to a hairpin starting from a misfolded conformation instead of the native one during unfolding (pull curve). The application of force allows the misfolded RNA/DNA to unfold partially or completely thereby forming the native hairpin in the zip transition first, with further force resulting in unfolding (zip), as is also observed in the native curves. The distribution in the number of rescued

nucleotides and the corresponding rescue force distribution is presented in ESI Fig. S1†. The distribution shows that the rescue can happen from a range of partially folded structures with a majority of them involving less nucleotides than the full hairpin, suggesting that they are rescued from a partially folded structure. It has to be noted that the misfolded state is attained after visiting different conformational during folding (Fig. 3A and ESI Fig. S15A-B†).

- **Locked**. In pull curves termed 'locked', we observed only two states, and a possible explanation for this misfolded conformation without rescue is presented in detail in ESI Section 5† and ESI Fig. S12-14†.

The signature of relax FDCs, illustrating the folding process, was well-resolved at high forces. At low forces (close to 0 pN), the resolution of the optical tweezers instrument did not allow for the distinction of the folding events (ESI Fig. S15†). Therefore, as described previously in the literature,[8, 9] the relax FDCs were classified between different folding pathways based on the outcome obtained in the consecutive pull FDC. The features of native, low, and locked relax curves show similar steps as in the native, low, and locked pull FDCs, respectively, except that in reverse direction. A misfolding curve showed no zipping to the native state, which was achieved, however, in the consecutive pull curve, as described before.[6, 7]

The unfolding (F_U) and folding (F_F) forces were determined as force jumps between folded (F), intermediate (I), and unfolded (U) conformations. The unfolding (E_U) and folding (E_F) extensions of the molecule for each of the transitions, at a given force, were obtained using the formula:[10]

$$E(F) = \frac{\Delta f}{k_{eff}} + x_d(F) \quad (1)$$

where Δf is the measure of a force jump between two states in an FDC, k_{eff} is the effective stiffness of the hairpin in the folded state (equals the slope of an FDC), and x_d is the double helix diameter. The x_d was modeled as a single bond of a length, d = 2.0 nm and the mathematical equation for this model is given by:

$$x_d(F) = d\,[coth(\frac{Fd}{k_BT}) - \frac{k_BT}{Fd}] \quad (2)$$

where k_B is the Boltzmann constant and T is the absolute temperature (298 K). The force and extension frequency distributions were presented in force (bin size of 0.5 pN) and extension (bin size of 0.5 nm) histograms, respectively. The errors reported for each bin in the histograms were obtained by calculating the error of the relative frequency of forces/extensions in each bin over different molecules.

3. Identification of all and first transitions in FDCs

The relax and pull FDCs of the SARS-CoV-2 SL4 constructs were analysed with custom-written MATLAB programs to extract (un)folding forces and extensions. As the first step in the analysis, all datapoints in an FDC were classified between folded, intermediate, and unfolded states, and all transitions between these states were identified. To achieve this, a linear fit was initially made to datapoints in the folded-state (6-9 pN; blue line) and unfolded-state (18-25 pN; red line) regions of the FDCs, respectively. An intermediate-state region was then hallmarked by a line in the middle between the two linear fits, and each datapoint was assigned to either the folded, intermediate, or unfolded state, depending on which line it was the closest to. Hereafter, the blue line was fitted again (ESI Fig. S16†), this time taking into account all folded-state datapoints – from both the folded-state region and the transition region (*i.e.,* the region where conformational transitions occur; 9-18 pN). By analogy, the red-line fitting (ESI Fig. S16†) was repeated for all unfolded-state datapoints – from the unfolded-state and transition regions. Next, datapoints were assigned between the different states once more. Finally, a linear fit was made for intermediate state datapoints (ESI Fig. S16†, black line) to better visualize the change in the stiffness of the intermediate with distance (black line has lower slope than blue and red lines).

The detection of a short-lived intermediate state is limited by the acquisition rate (in this study, 1 kHz) of the instrument. We here defined a single-step transition as a transition where the distance between the folded and unfolded states is < 1 nm (ESI Fig. S2†). Two-step transition, on the other hand, is characterized by a distance between the first (F-I in unfolding/U-I in folding) and the second step (I-U in unfolding/I-F in folding) of ≥ 1 nm. The lifetime of the intermediate state for the WT SL4 ranged from less than 5 ms to 244 ms (unfolding) and 326 ms (folding), respectively. The distribution of lifetimes of the intermediate state using first transitions for the WT SL4 showed that ~90% of the intermediates existed for less than 5 ms (ESI Fig. S3†). This is much higher than what would be expected from an exponential distribution of all lifetimes and suggests that

there are two different native processes identified as single-steps and two-steps. Whereas UC SL4 show an exponential distribution with ~10% of intermediates in the range ≤ 5 ms suggesting the existence only one native pathway confirming our experiments (ESI Fig. S4†). The intermediate's lifetime distribution of tC^O-modified SL4 constructs also showed similar inference as WT SL4 (ESI Fig. S5-S7†). Short (less than 2 frames, *i.e.,* 2 datapoints) intermediate states, as well as short folded or unfolded states transitioning directly back and forth from an intermediate state, were considered as noise. The unfolding and folding forces for all transitions between the three states were used to calculate the extensions as presented in the Materials and Methods section. In addition, the first transition in each FDC, either single-step or two-step, was identified and displayed separately. The non-native FDCs, *i.e.*, misfolding, rescue, low, and locked, were manually separated from native FDCs prior to the analysis.

4. Force-dependent state probability analysis

Similarly, as in the algorithm for the identification of transitions in FDCs, to obtain the probability of each conformational state, a single FDC was first divided into different force regions (the same as in the section above) and each datapoint was then assigned to either folded, intermediate, or unfolded state and refitted again. After data points' assignment, the occurrence of each state for a bin size of 0.5 pN was calculated for each molecule. Finally, after combining all measured molecules for the specific construct, the occurrence of different conformational states was plotted for both pull and relax FDCs (ESI Fig. S17†).

5. Hypothetical model of the locked conformation observed for UD SL4

The WT SARS-CoV-2 SL4 structure shown in Fig. 2A and ESI Fig. S12A† is the most thermodynamically favoured structure with ΔG = -24.3 kcal/mol (UNAFold). It is identical with the structure determined experimentally.[11, 12] The second most stable structure predicted in UNAFold with ΔG = -18.8 kcal/mol is shown in ESI Fig. S12B†. It has two hairpins: a smaller hairpin (12 nt) and a larger hairpin (23 nt) separated by six unpaired nucleotides.

The difference in free energy between the two conformations ($\Delta\Delta G$) is 5.5 kcal/mol, making the folding of WT SL4 to the second conformation highly unfavourable. Possibly, this second structure could be the locked state in the (un)folding of UD SL4, which was observed in a large fraction of FDCs (21%; Fig. 3A and 3B). The unfolding transitions observed in locked FDCs were around 13.0 ± 0.6 pN (ESI Fig. S13A†) and might correspond to the unfolding of the larger hairpin (23 nt). The transitions corresponding to the unfolding of the smaller hairpin (12 nt) probably occurred at very low forces, which made them difficult to detect. We estimated that the number of nucleotides released in the unfolding of the locked state was 26.0 ± 1.6 nt (ESI Fig. S13B†), which agrees with the size of the larger hairpin (23 nt) in ESI Fig. S12B†. This might suggest that the two tC^Os incorporated into the upper stem of UD SL4 significantly decrease the $\Delta\Delta G$ between the two conformations compared to WT SL4.

Supporting Figures

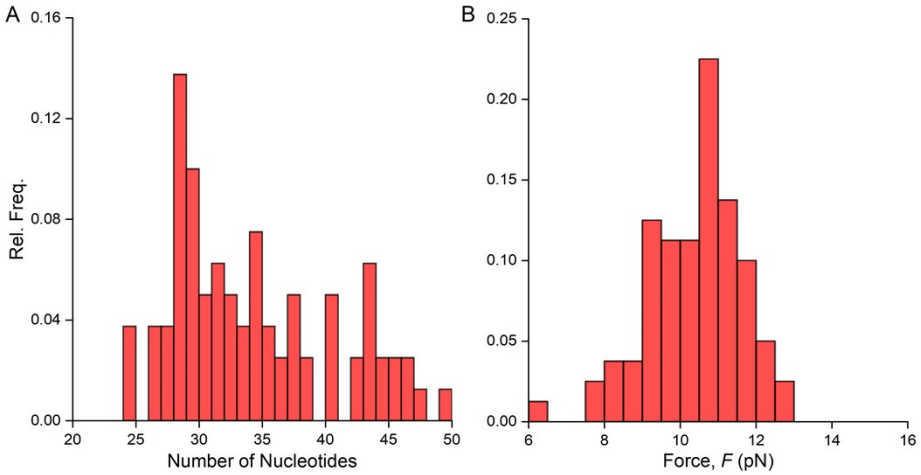


Fig. S1 Histograms of the number of nucleotides rescued (A) and rescue force (B) estimated from the WT SARS-CoV-2 SL4 misfolding-rescue FDCs. Rel. Freq. – relative frequency.

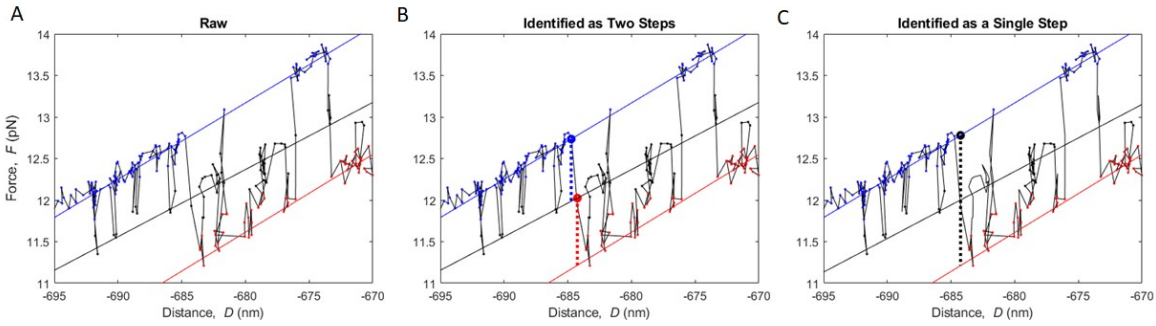


Fig. S2 Zoom in of an exemplary native pull FDC of the WT SARS-CoV-2 SL4 illustrating the criteria used in the identification of single vs two steps. (A) Raw datapoints before the identification of a transition type. (B) Misidentification of two steps as the first transition (F-I) and the second one (I-U) occurred without a significant change in distance/extension (< 1 nm). (C) Identification of a single step after excluding the false-positive intermediate. The same criteria were used in the identification of the number of steps in relax FDCs.

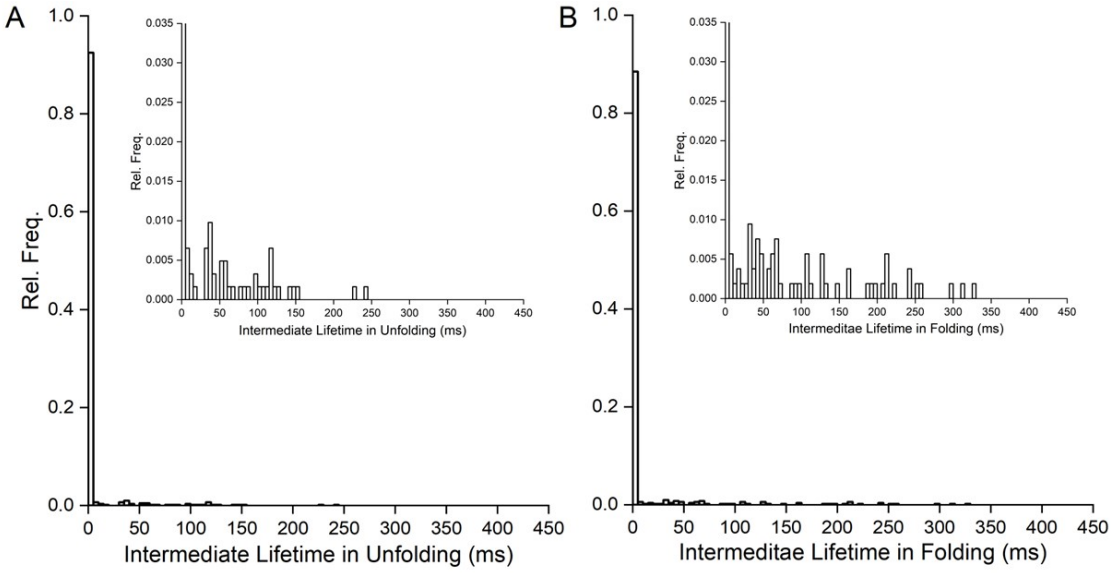


Fig. S3 Histogram of the intermediate lifetime in the native unfolding (A) and folding (B) WT SARS-CoV-2 SL4 with the zoom in the inset. Rel. Freq. – relative frequency.

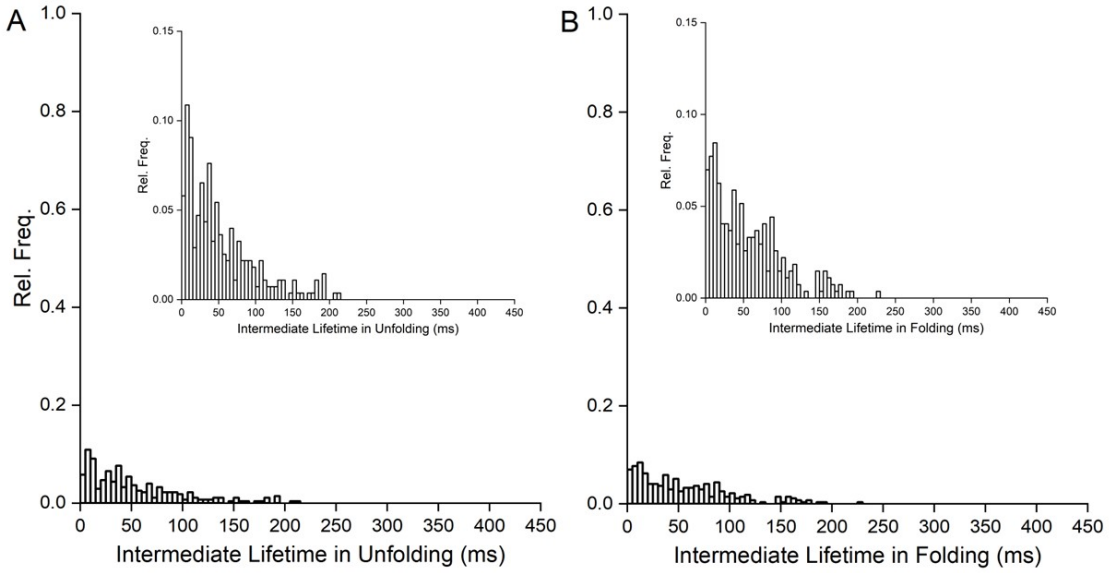


Fig. S4 Histograms of the intermediate lifetime in the native unfolding (A) and folding (B) of the UC SARS-CoV-2 SL4 with the zoom in the inset. Rel. Freq. – relative frequency.

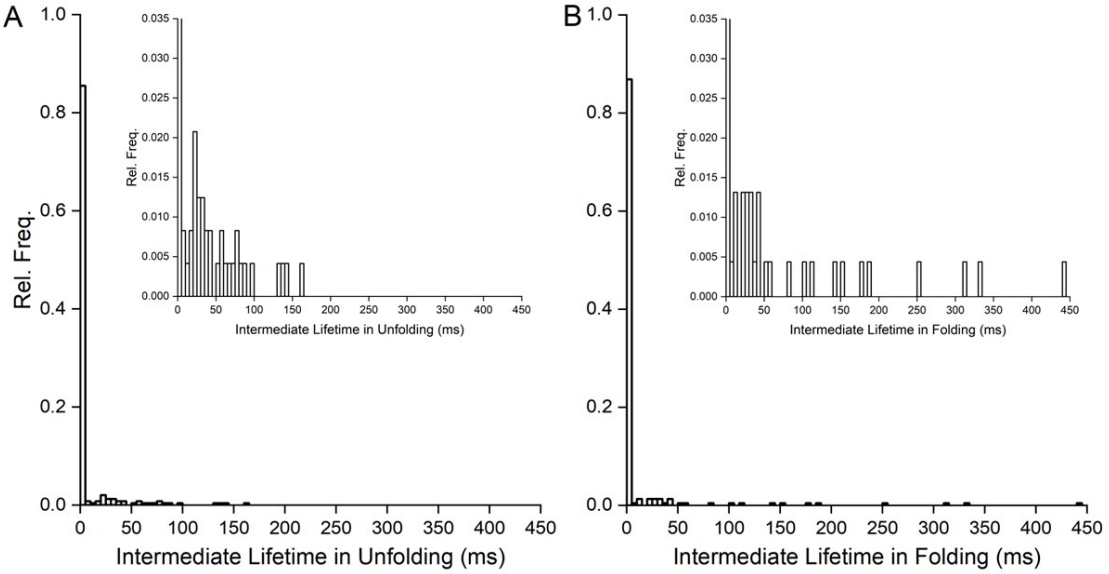


Fig. S5 Histograms of the intermediate lifetime in the native unfolding (A) and folding (B) of the US SARS-CoV-2 SL4 with the zoom in the inset. Rel. Freq. – relative frequency.

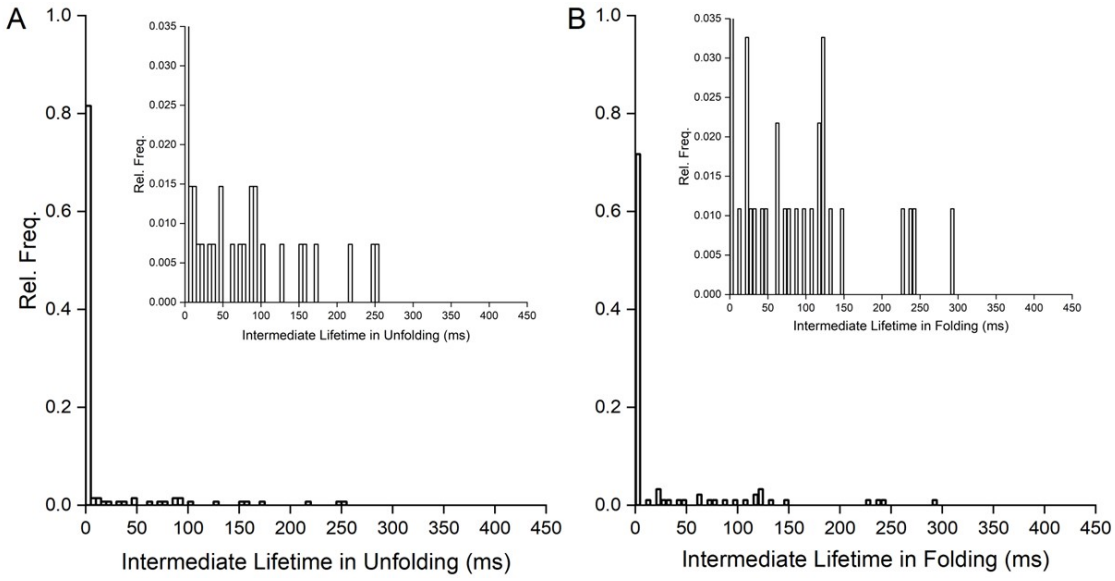


Fig. S6 Histograms of the intermediate lifetime in the native unfolding (A) and folding (B) of the UD SARS-CoV-2 SL4 with the zoom in the inset. Rel. Freq. – relative frequency.

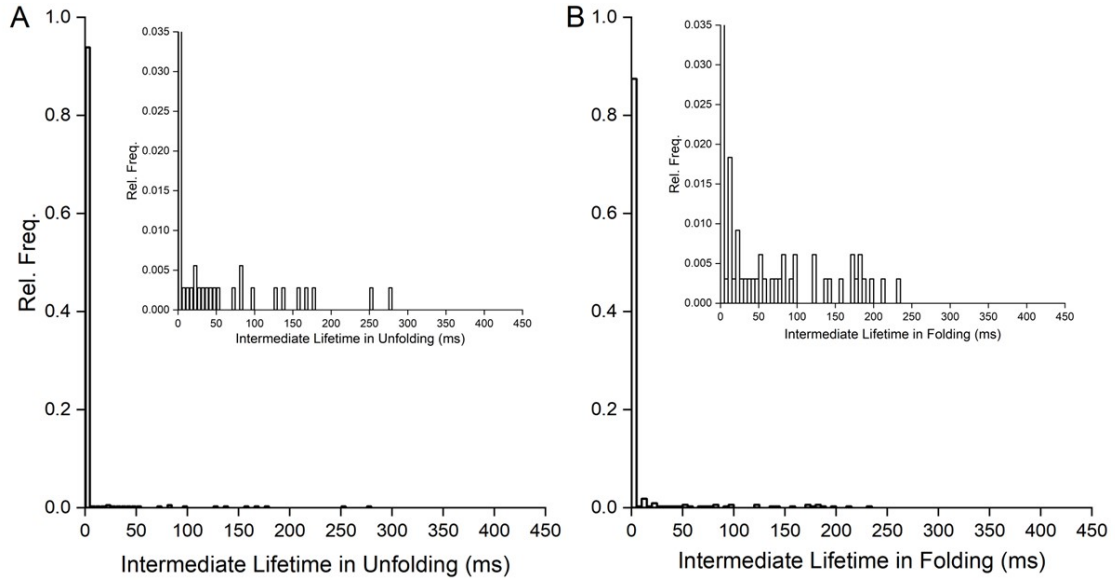
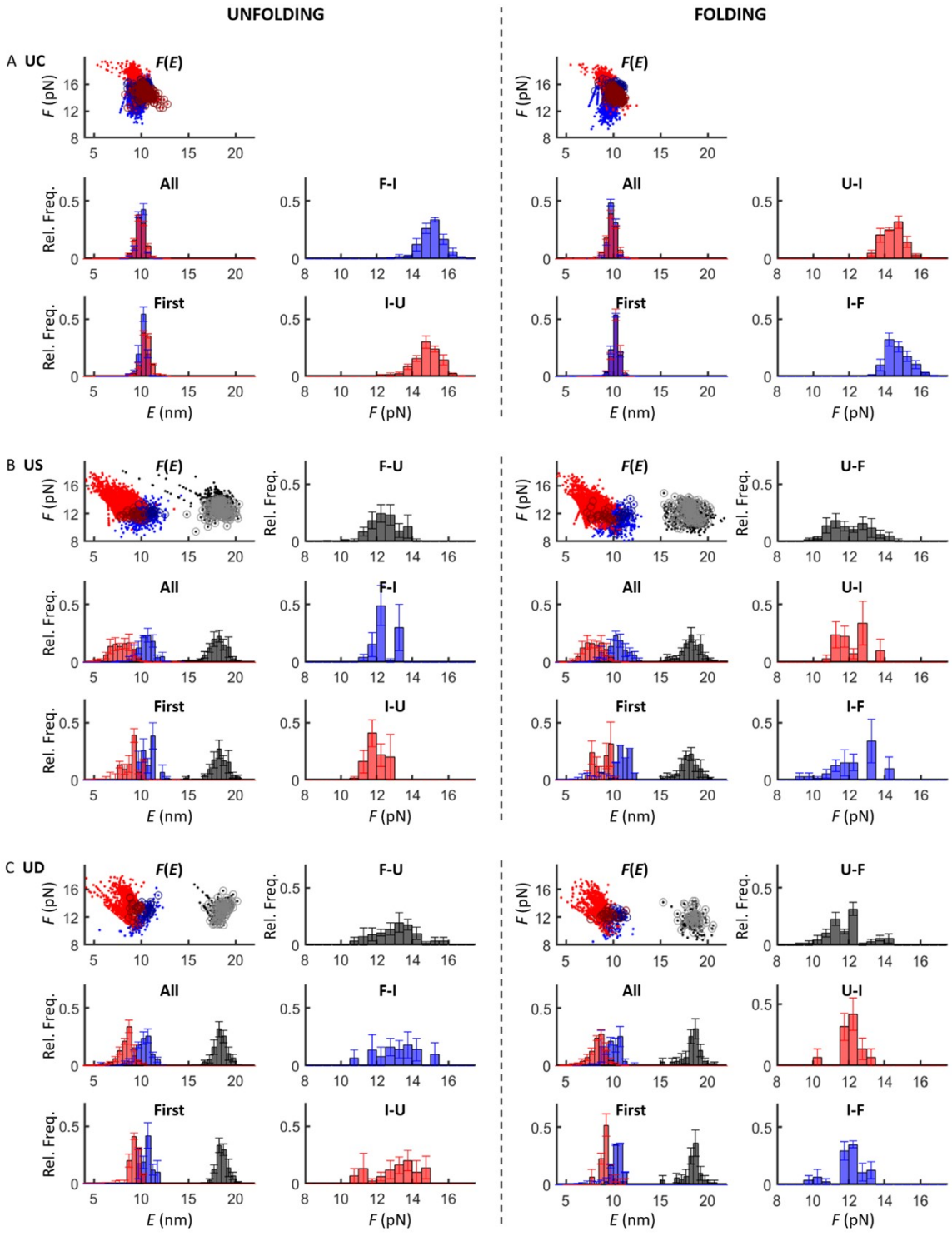


Fig. S7 Histograms of the intermediate lifetime in the native unfolding (A) and folding (B) of the LD SARS-CoV-2 SL4 with the zoom in the inset. Rel. Freq. – relative frequency.


UNFOLDING
FOLDING
A UC
B US
C UD
F(E)
All
First
F-I
I-U
F-U
U-I
I-F
U-F
F (pN)
E (nm)
Rel. Freq.

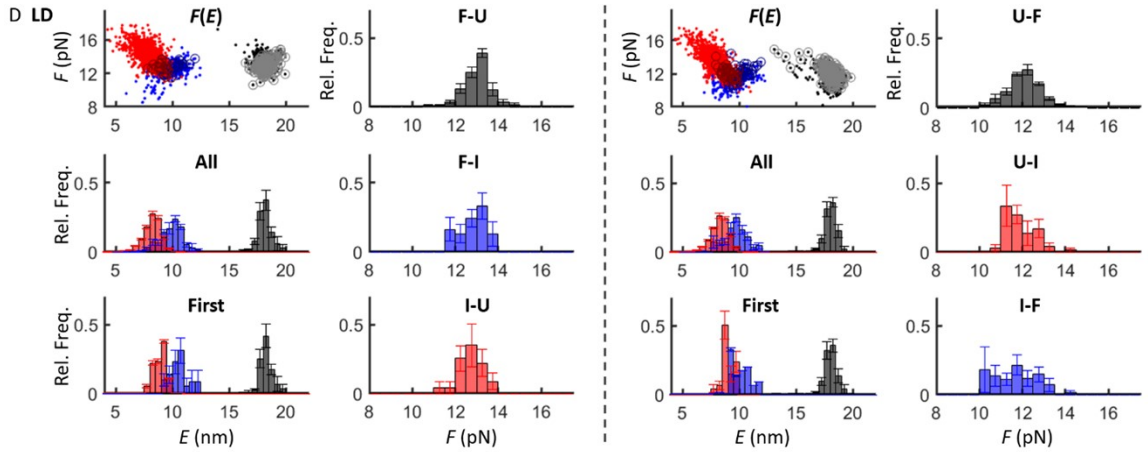


Fig. S8 Force-extension (*F*(*E*)) plots with the corresponding force (*F*) and extension (*E*) histograms in native (un)folding of the different SARS-CoV-2 SL4 constructs. (A) UC, (B) US, (C) UD, and (D) LD. The transitions occur between F – folded, I – intermediate, and U – unfolded states, and are shown in grey – F-U/U-F transitions, blue – F-I/I-F transitions, and red – I-U/U-I transitions. In *F*(*E*) plots, dots – all transitions and circles – first transitions. Rel. Freq. – relative frequency. The corresponding data for WT SL4 can be seen in the main text, Fig. 4.

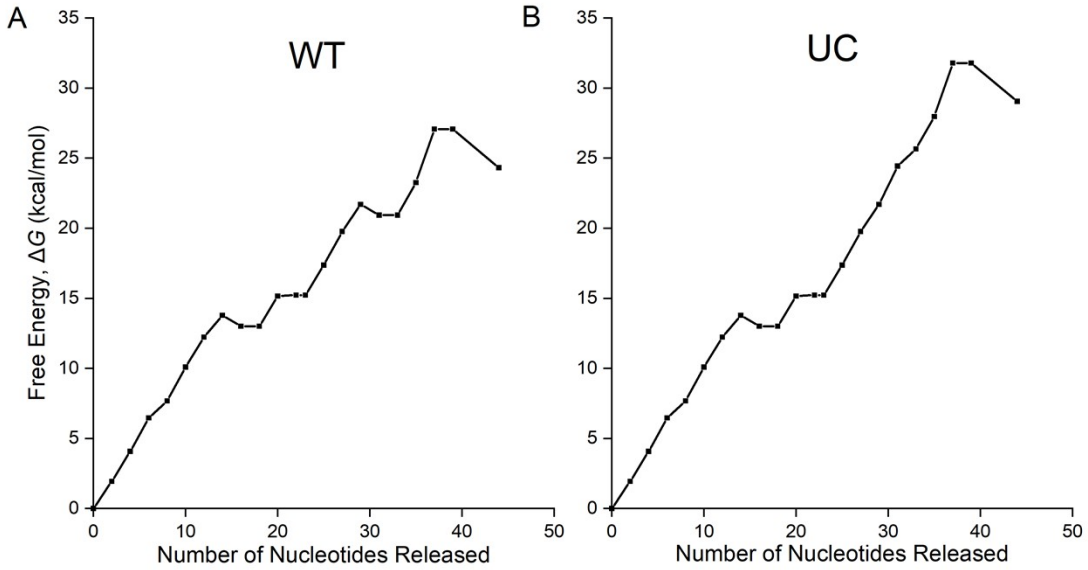


Fig. S9 The free energy landscape (FEL) at zero force for the WT (A) and UC (B) SARS-CoV-2 SL4 obtained in UNAFold by summing base pairs and stacking energies. Both WT and UC show an intermediate state embracing half of the hairpin (around 22 nt out of 44), which agrees with our optical tweezers' measurements (Table 1 in the main text).

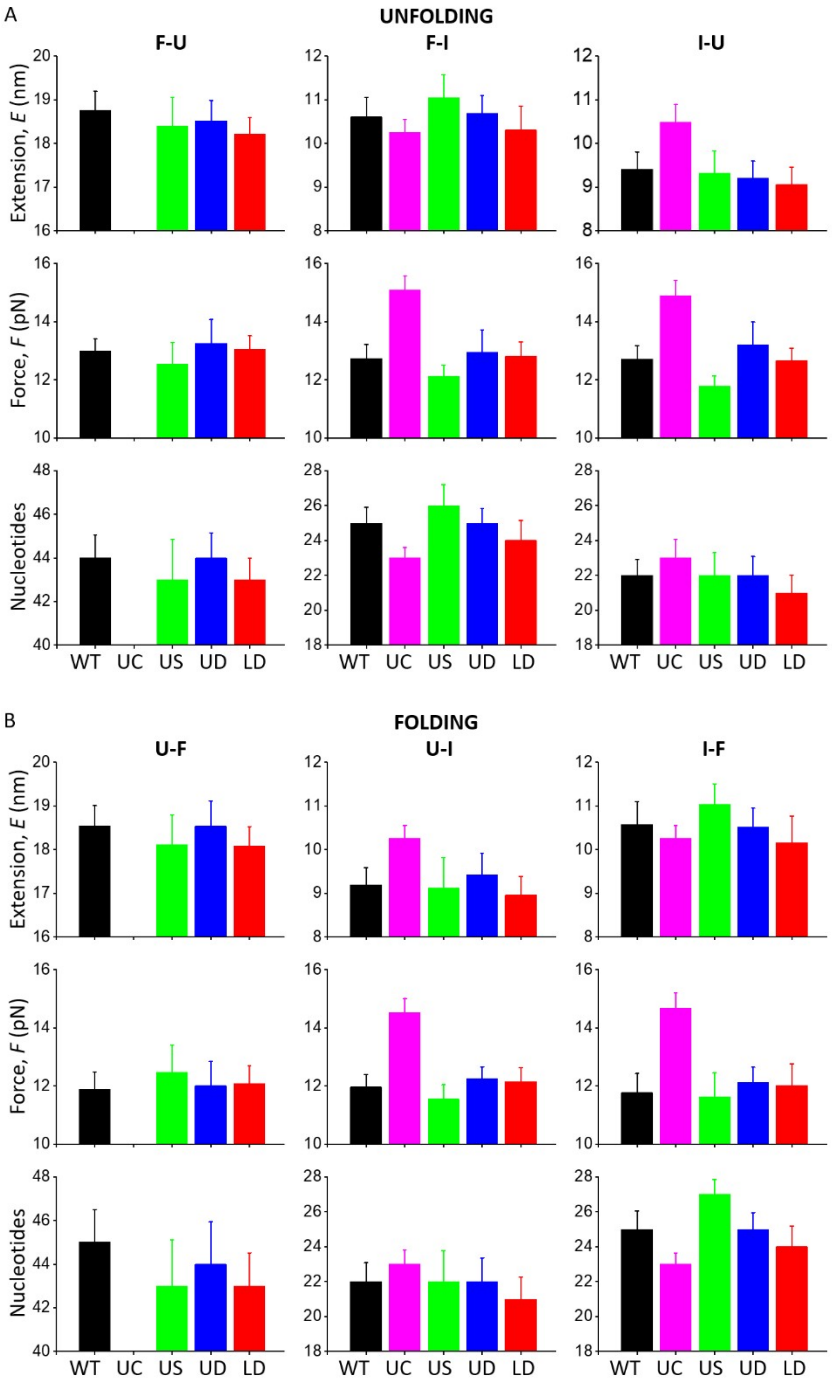


Fig. S10 Median extension (top), force (middle) and nucleotide number (bottom) for the different SARS-CoV-2 SL4s: WT– black, UC – pink, US – green, UD – blue, and LD – red. (A) Unfolding and (B) Folding. The median values can be seen in Table 1 in the main text.

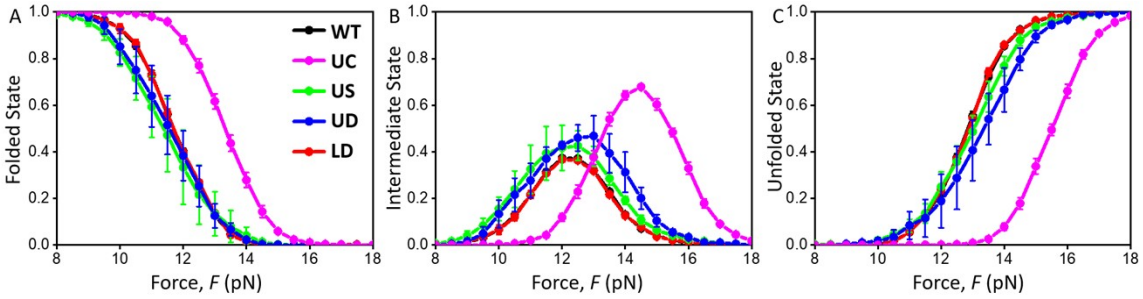


Fig. S11 Force-dependent probability of folded (A), intermediate (B), and unfolded (C) states during folding for each SARS-CoV-2 SL4. Note the overlap of the WT (in black) and LD (in red) results in (A-C). Colour legend given in (A) also applies to (B) and (C). The corresponding data for unfolding can be seen in Fig. 6 in the main text.

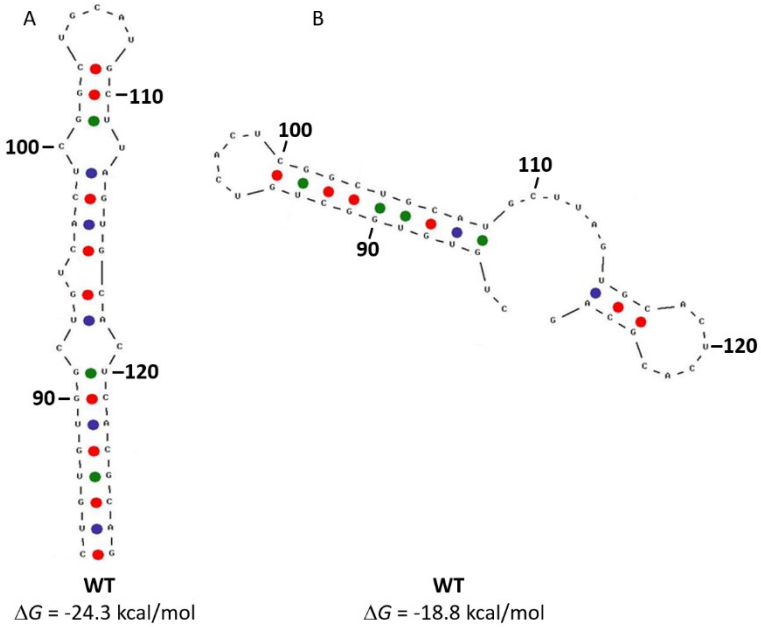


Fig. S12 The first (A) and the second (B) lowest free energy secondary structures of the WT SARS-CoV-2 SL4 predicted using UNAFold.

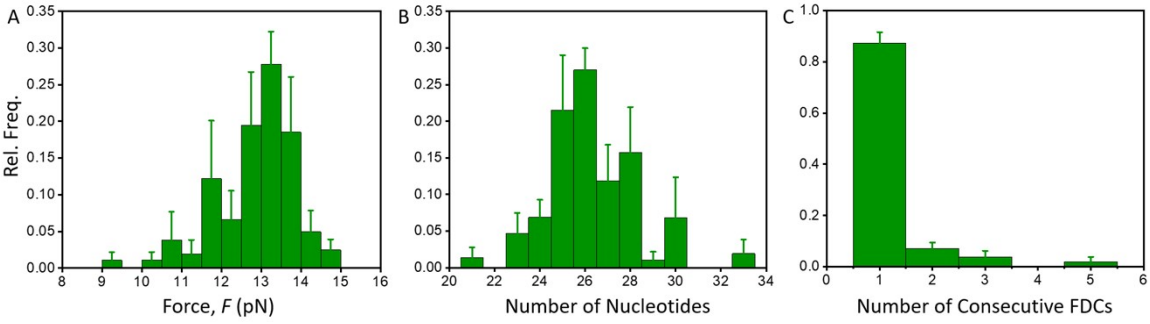


Fig. S13 (A) Unfolding force and (B) number of unfolded nucleotides in the locked-locked pathway of UD SL4. (C) Occurrence of consecutive locked-locked FDCs of UD SL4. Rel. Freq. – relative frequency.

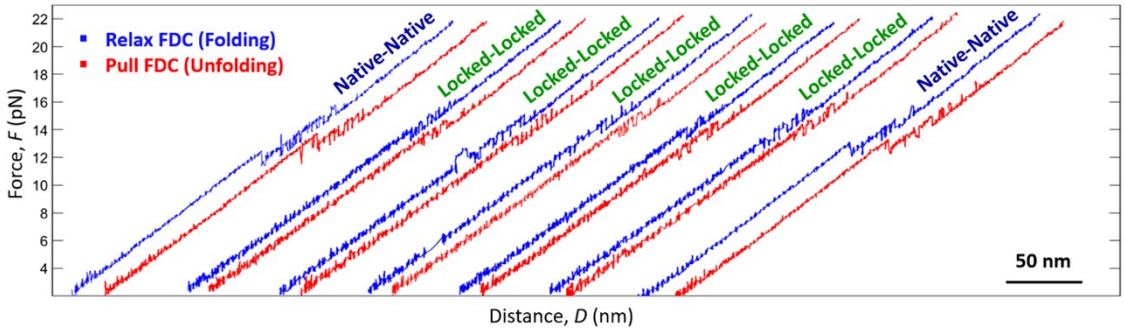


Fig. S14 Example of consecutive relax-pull FDCs for UD SL4 with the locked-locked pattern repeating.

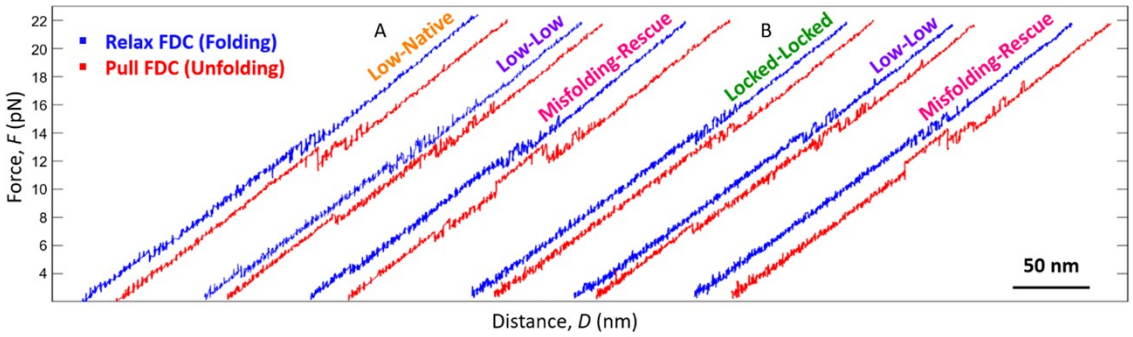


Fig. S15 Classification of folding pathways (blue FDCs) based on the outcome in unfolding (red FDCs). (A) Zip at a low force in relax FDCs led to different folded conformations (native or not) distinguished only in the consecutive pull FDC. (B) Flat/gradually changing bottom of relax FDCs corresponded to different conformations distinguished only in the consecutive pull FDC.

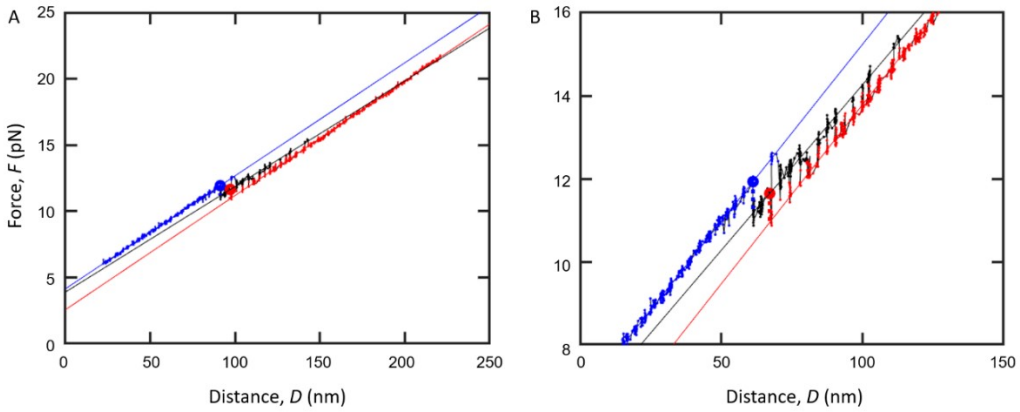


Fig. S16 (A) An exemplary pull FDC of the WT SARS-CoV-2 SL4, illustrating the classification of datapoints between the folded, transition and unfolded regions, linear fits (blue and red lines), and the change in the intermediate stiffness (the slope of the black line). The site of the first transition (two-step) is marked by circles. (B) Zoom in on the transition region.

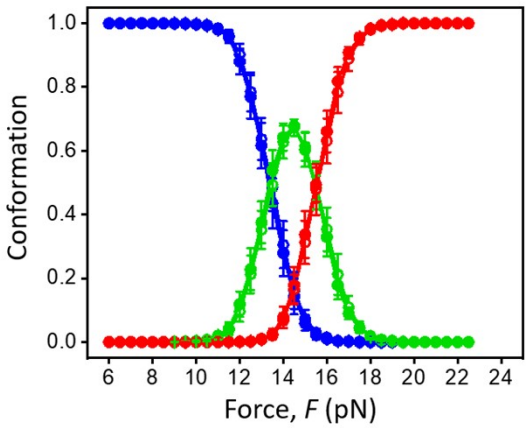


Fig. S17 An example of force-dependent state probability analysis in unfolding (unfilled circles) and folding (filled circles) for the UC SARS-CoV-2 SL4. Unfolded state – red, intermediate state – green, and folded state – blue.

Supporting Tables

Table S1 Oligonucleotide sequences used in the optical tweezers experiments.[a]

Name	Nucleic Acid	Length (nt)	5'-3' Sequence
Oligo1	DNA	35	Biotin-AGTTAGTGGTGGAAGCACAGTGCCAGCGCAGTTAG
Oligo2 WT	DNA-RNA-DNA Hybrid	74	Phosphate-CGGTGGAAACACAGT-**CUGUGUGGCUGUCACUCGGCUGCAUGCUUAGUGCACUCACGCAG**-AGTTAGTGGTGGAAG
Oligo2 UC	DNA-RNA-DNA Hybrid	74	Phosphate-CGGTGGAAACACAGT-**CUGUGUGGCUGUCACUCGGCUGCAUGCU<u>G</u>AGUGCACUCACGCAG**-AGTTAGTGGTGGAAG
Oligo2 US	DNA-RNA-DNA Hybrid	74	Phosphate-CGGTGGAAACACAGT-**CUGUGUGGCUGUCACUXGGCUGCAUGCUUAGUGCACUCACGCAG**-AGTTAGTGGTGGAAG
Oligo2 UD	DNA-RNA-DNA Hybrid	74	Phosphate-CGGTGGAAACACAGT-**CUGUGUGGCUGUCAXUCGGCUGCAUGXUUAGUGCACUCACGCAG**-AGTTAGTGGTGGAAG
Oligo2 LD	DNA-RNA-DNA Hybrid	74	Phosphate-CGGTGGAAACACAGT-**CUGUGUGGCUGUCACUCGGCUGCAUGCUUAGUGCACUXACGXAG**-AGTTAGTGGTGGAAG
Oligo3	DNA	35	Phosphate-CACAGTGCCAGCGCAGTTAGCGGTGGAAACACAGT
Oligo4	DNA	50	ACTGTGTTTCCACCGCTAACTGCGCTGGCACTGTGCTTCCACCACTAACT

[a] The RNA fragment of Oligo2 corresponding to the WT or modified SARS-CoV-2 SL4 sequence is highlighted in bold. The U112G mutation in Oligo2 UC is underlined. The tC^O substitutions of C in US, UD and LD are indicated by X.

Table S2 Statistical analysis of the fraction of the native-native (un)folding pathway in SL4 constructs, corresponding to Fig. 3B.[a]

SARS-CoV-2	SARS-CoV-2 SL4			
SL4	WT	US	UD	LD
WT				
US	*			
UD	*	*		
LD	NS	*	*	

[a] Performed in MATLAB using Wilcoxon rank sum test [13] for each molecule of a minimum of 50 FDCs, * $p < 0.05$ and not significant (NS) $p > 0.05$.

Table S3 Statistical analysis of the fraction of the low-native (un)folding pathway in SL4 constructs, corresponding to Fig. 3B.[a]

SARS-CoV-2	SARS-CoV-2 SL4			
SL4	WT	US	UD	LD
WT				
US	NS			
UD	NS	*		
LD	NS	NS	*	

[a] Performed in MATLAB using Wilcoxon rank sum test [13] for each molecule of a minimum of 50 FDCs, * $p < 0.05$ and not significant (NS) $p > 0.05$.

Table S4 Statistical analysis of the fraction of the low-low (un)folding pathway in SL4 constructs, corresponding to Fig. 3B.[a]

SARS-CoV-2	SARS-CoV-2 SL4			
SL4	WT	US	UD	LD
WT				
US				
UD	NS			
LD	NS		NS	

[a] Performed in MATLAB using Wilcoxon rank sum test [13] for each molecule of a minimum of 50 FDCs, * $p < 0.05$ and not significant (NS) $p > 0.05$. Note: Low-low pathway was not observed in US.

Table S5 Statistical analysis of the fraction of the misfolding-rescue (un)folding pathway in SL4 constructs, corresponding to Fig. 3B.[a]

SARS-CoV-2	SARS-CoV-2 SL4			
SL4	WT	US	UD	LD
WT				
US	*			
UD	*	*		
LD	NS	*	NS	

[a] Performed in MATLAB using Wilcoxon rank sum test [13] for each molecule of a minimum of 50 FDCs, * $p < 0.05$ and not significant (NS) $p > 0.05$.

Table S6 Number of transitions (*T*) between F and I *vs* I and U states in native (un)folding normalized by the number of FDCs.[a]

SARS-CoV-2 SL4	No of Molecules	No of FDCs	Unfolding			Folding		
			T_{F-I}	T_{I-U}	T_{I-U}/T_{F-I}	T_{U-I}	T_{I-F}	T_{U-I}/T_{I-F}
WT	9	1093	1.1	7.7	7.0	9.6	1.0	9.6
UC	5	548	12.3	9.2	0.7	9.5	12.6	0.7
US	5	481	3.1	14.2	4.6	13.4	3.0	4.5
UD	5	228	1.5	9.3	6.2	10.5	1.4	7.5
LD	6	686	1.2	7.2	6.0	8.9	1.2	7.4

[a] Measured in 50 mM Tris-HCl, 140 mM NaCl, 5.0 mM KCl, 1.0 mM EDTA and 5.0 mM $MgCl_2$, pH 7.4, at room temperature.

Table S7 Statistical analysis of the fraction of single- vs two-step native unfolding in SL4 constructs, corresponding to Fig. 5A.[a]

SARS-CoV-2 SL4	SARS-CoV-2 SL4			
	WT	US	UD	LD
WT				
US	NS			
UD	*	NS		
LD	NS	NS	*	

[a] Performed in MATLAB using Wilcoxon rank sum test [13] for each molecule of a minimum of 50 FDCs, * $p < 0.05$ and not significant (NS) $p > 0.05$.

Table S8 Statistical analysis of the fraction of single- vs two-step native folding in SL4 constructs, corresponding to Fig. 5B.[a]

SARS-CoV-2 SL4	SARS-CoV-2 SL4			
	WT	US	UD	LD
WT				
US	NS			
UD	*	NS		
LD	NS	NS	*	

[a] Performed in MATLAB using Wilcoxon rank sum test [13] for each molecule of a minimum of 50 FDCs, * $p < 0.05$ and not significant (NS) $p > 0.05$.

SI References

1. A. Wypijewska del Nogal, A. F. Fuchtbauer, M. Bood, J. R. Nilsson, M. S. Wranne, S. Sarangamath, P. Pfeiffer, V. S. Rajan, A. H. El-Sagheer, A. Dahlen, T. Brown, M. Grotli and L. M. Wilhelmsson, *Nucleic Acids Res*, 2020, **48**, 7640-7652.
2. M. Bood, A. Wypijewska del Nogal, J. R. Nilsson, F. Edfeldt, A. Dahlen, M. Lemurell, L. M. Wilhelmsson and M. Grotli, *Sci Rep*, 2021, **11**, 9396.
3. A. Wypijewska del Nogal, V. Sundar Rajan, F. Westerlund and L. M. Wilhelmsson, *Int J Mol Sci*, 2021, **22**.
4. J. Liphardt, B. Onoa, S. B. Smith, I. Tinoco, Jr. and C. Bustamante, *Science*, 2001, **292**, 733-737.
5. P. T. Li, C. Bustamante and I. Tinoco, Jr., *Proc Natl Acad Sci U S A*, 2006, **103**, 15847-15852.
6. P. T. Li, C. Bustamante and I. Tinoco, Jr., *Proc Natl Acad Sci U S A*, 2007, **104**, 7039-7044.
7. L. Pekarek, M. M. Zimmer, A. S. Gribling-Burrer, S. Buck, R. Smyth and N. Caliskan, *Nucleic Acids Res*, 2023, **51**, 728-743.
8. Y. J. Wu, C. H. Wu, A. Y. Yeh and J. D. Wen, *Nucleic Acids Res*, 2014, **42**, 4505-4515.
9. J. Li and H. Li, *Chem Sci*, 2022, **13**, 7498-7508.
10. A. Alemany and F. Ritort, *Biopolymers*, 2014, **101**, 1193-1199.
11. C. Iserman, C. A. Roden, M. A. Boerneke, R. S. G. Sealfon, G. A. McLaughlin, I. Jungreis, E. J. Fritch, Y. J. Hou, J. Ekena, C. A. Weidmann, C. L. Theesfeld, M. Kellis, O. G. Troyanskaya, R. S. Baric, T. P. Sheahan, K. M. Weeks and A. S. Gladfelter, *Mol Cell*, 2020, **80**, 1078-1091 e1076.
12. A. Wacker, J. E. Weigand, S. R. Akabayov, N. Altincekic, J. K. Bains, E. Banijamali, O. Binas, J. Castillo-Martinez, E. Cetiner, B. Ceylan, L. Y. Chiu, J. Davila-Calderon, K. Dhamotharan, E. Duchardt-Ferner, J. Ferner, L. Frydman, B. Furtig, J. Gallego, J. T. Grun, C. Hacker, C. Haddad, M. Hahnke, M. Hengesbach, F. Hiller, K. F. Hohmann, D. Hymon, V. de Jesus, H. Jonker, H. Keller, B. Knezic, T. Landgraf, F. Lohr, L. Luo, K. R. Mertinkus, C. Muhs, M. Novakovic, A. Oxenfarth, M. Palomino-Schatzlein, K. Petzold, S. A. Peter, D. J. Pyper, N. S. Qureshi, M. Riad, C. Richter, K. Saxena, T. Schamber, T. Scherf, J. Schlagnitweit, A. Schlundt, R. Schnieders, H. Schwalbe, A. Simba-Lahuasi, S. Sreeramulu, E. Stirnal, A. Sudakov, J. N. Tants, B. S. Tolbert, J. Vogele, L. Weiss, J. Wirmer-Bartoschek, M. A. Wirtz Martin, J. Wohnert and H. Zetzsche, *Nucleic Acids Res*, 2020, **48**, 12415-12435.
13. C. F. Hsu, K. C. Chang, Y. L. Chen, P. S. Hsieh, A. I. Lee, J. Y. Tu, Y. T. Chen and J. D. Wen, *Nucleic Acids Res*, 2021, **49**, 6941-6957.