SUPPLEMENTARY DATA

Molecular dynamics analysis of conformational change of paramyxovirus F protein during the initial steps of membrane fusion.

Fernando Martín-García[1,2], Jesús Ignacio Mendieta-Moreno[1,2] , Jesús Mendieta[1,2] and Paulino Gómez-Puertas[1].

[1]*Centro de Biología Molecular "Severo Ochoa" (CSIC/UAM), C/ Nicolás Cabrera, 1, Cantoblanco, 28049 Madrid, Spain.*
[2] *Biomol-Informatics SL, Parque Científico de Madrid, C/ Faraday, 7, Cantoblanco, 28049 Madrid, Spain.*

Contents:

1. Supplementary Materials and Methods

1.1. Analysis of MD trajectories. To ensure that sufficient sampling has been done, a series of analysis of the MD trajectories were performed: temporal profiles of root mean square deviation (RMSD), secondary structure and radius of gyration, as well as root mean square fluctuation (RMSF) profiles of residues during MD (Figures S2 to S4). In addition, non-weighted covariance matrix of Cα atoms of the structure were calculated to obtain cosine content (c_i) [1] of the three first principal components (Table S1). This value ranges between 0 (no cosine) and 1 (perfect cosine). Values close to 1 are representative of random motion, and therefore, of insufficient sampling.

2. Supplementary Figures

Figure S1. Refolding of domain III of paramyxovirus F protein.

Paramyxovirus F protein has been crystallized in pre-fusion and post-fusion states for two different members of the *Paramyxoviridae* family: simian virus 5 (PDB ID: 2B9B) and parainfluenza 3 virus (PDB ID: 1ZTM) respectively. In order to compare how the structural changes exhibited by domain III of F protein could be extrapolated to other members of the family, a concise multiple alignment of some representative sequences is included. **A**. Multiple sequence alignment of domain III of F protein from representative members of the *Paramyxoviridae* family (*HRSV1* and *HRSVL*: human respiratory syncytial virus groups B (strain 18537) and A (strain Long) respectively; *SV5*: simian virus 5 (strain W3); *NDVA*: Newcastle disease virus (strain chicken/Australia-Victoria/32); *9PARA*: human parainfluenza virus 1; *SENDA*: Sendai virus (strain Hamamatsu); *PI3H4*: human parainfluenza 3 virus (strain Wash/47885/57)). Elements of the secondary structure of crystallized pre-fusion (PDB ID: 2B9B) and post-fusion (PDB ID: 1ZTM) F protein are included (arrows: beta-

sheets; rectangles: alpha-helices). Position of fusion peptide (F-pep) is indicated. Purple box surrounds the HRA region that undergoes the principal structural change. **B**. Domain III in the pre-fusion state of the crystallized structure from simian virus 5 (PDB ID: 2B9B). **C**. Domain III in the post-fusion state of the crystallized structure from human parainfluenza 3 virus (PDB code: 1ZTM). Proteolytic cleavage site (B) and disulfide bond (B, C) are indicated. Figure plots were generated using PyMOL Molecular Graphics System, Schrödinger, LLC.

Figure S2. Analysis of monomeric domain during steered-MD simulation.

Temporal profiles of Cα root mean square deviation (RMSD, **A**) and radius of gyration (**B**); and root mean square fluctuation (RMSF, **C**) profile of all residues in the domain III of the monomeric structure. Data correspond to the MD simulation shown in Figure 1 and Figure 2B in the manuscript.

Figure S3. Analysis of trimeric domain during MD simulation of coiled-coil formation.

Temporal profiles of Cα root mean square deviation (RMSD, **A**) and radius of gyration (**B**); and root mean square fluctuation (RMSF, **C**) profile of all residues in the domain III of the trimeric structure. Data correspond to the MD simulation shown in Figure 3 in the manuscript. Simulation has been extended up to 20 ns to get a broader view of the stability of the obtained structure after conformational change of the trimer.

Figure S4. Analysis of complete protein during 30 ns of unrestricted MD simulation.

Temporal profiles of Cα root mean square deviation (RMSD, **A**), radius of gyration (**B**) and secondary structure content (**C**) of the protein in the pre-fusion state (starting structure: PDB ID: 2B9B; resolution: 2,85 Angstroms). Root mean square fluctuation

(RMSF, **D**) profile of residues in the domain III of the protein in the pre-fusion state. Data correspond to 30 ns of unrestricted MD simulation.

3. Supplementary Table

Table S1. Cosine content (c_i) of the first three principal components for monomeric domain and trimeric domain (coiled-coil formation) in stable phase of MD simulations.

	PC-1	PC-2	PC-3
monomeric domain	0.0013	0.1877	0.0419
trimeric domain	0.1019	0.0108	0.0114

Cosine content was calculated according to Hess, 2000 [1]. The results, far from a perfect cosine ($c_i = 1$) indicate that the trajectories have reached an overall sufficient sampling.

4. References.

[1] Hess, B. (2002). Convergence of sampling in protein simulations. Phys Rev E Stat Nonlin Soft Matter Phys 65, 031910

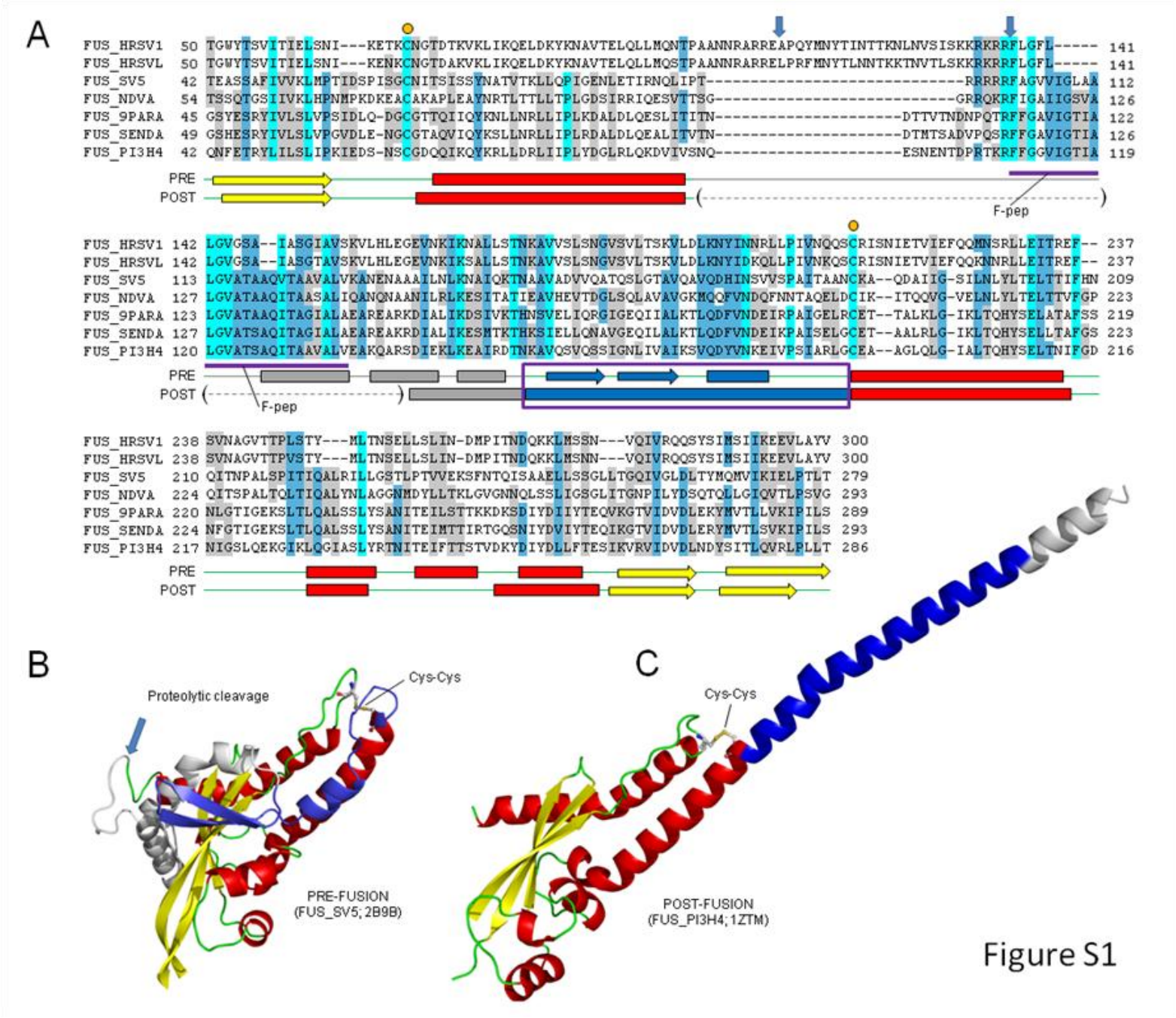
A
FUS_HRSV1 50 TGWYTSVITIELSNI---KETKCNGTDTKVKLIKQELDKYKNAVTELQLLMQNTPAANNRARREAPQYMNYTINTTKNLNVSISKKRKRRFLGFL----- 141
FUS_HRSVL 50 TGWYTSVITIELSNI---KENKCNGTDAKVKLIKQELDKYKNAVTELQLLMQSTPAANNRARRELPRFMNYTLNNTKKTNVTLSKKRKRRFLGFL----- 141
FUS_SV5 42 TEASSAFIVVKLMPTIDSPISGCNITSISSYNATVTKLLQPIGENLETIRNQLIPT----------RRRRRFAGVVIGLAA 112
FUS_NDVA 54 TSSQTGSIIVKLHPNMPKDKEACAKAPLEAYNRTLTTLLTPLGDSIRRIQESVTTSG----------GRRQKRFIGAIIGSVA 126
FUS_9PARA 45 GSYESRYIVLSLVPSIDLQ-DGCGTTQIIQYKNLLNRLLIPLKDALDLQESLITITN----------DTTVTNDNPQTRFFGAVIGTIA 122
FUS_SENDA 49 GSHESRYIVLSLVPGVDLE-NGCGTAQVIQYKSLLNRLLIPLRDALDLQEALITVTN----------DTMTSADVPQSRFFGAVIGTIA 126
FUS_PI3H4 42 QNFETRYLILSLIPKIEDS-NSCGDQQIKQYKRLLDRLIIPLYDGLRLQKDVIVSNQ----------ESNENTDPRTKRFFGGVIGTIA 119
PRE
POST
F-pep
FUS_HRSV1 142 LGVGSA--IASGIAVSKVLHLEGEVNKIKNALLSTNKAVVSLSNGVSVLTSKVLDLKNYINNRLLPIVNQQSCRISNIETVIEFQQMNSRLLEITREF-- 237
FUS_HRSVL 142 LGVGSA--IASGTAVSKVLHLEGEVNKIKSALLSTNKAVVSLSNGVSVLTSKVLDLKNYIDKQLLPIVNKQSCRISNIETVIEFQQKNNRLLEITREF-- 237
FUS_SV5 113 LGVATAAQVTAAVALVKANENAAAILNLKNAIQKTNAAVADVVQATQSLGTAVQAVQDHINSVVSPAITAANCKA--QDAIIG-SILNLYLTELTTIFHN 209
FUS_NDVA 127 LGVATAAQITAASALIQANQNAANILRLKESITATIEAVHEVTDGLSQLAVAVGKMQQFVNDQFNNTAQELDCIK--ITQQVG-VELNLYLTELTTVFGP 223
FUS_9PARA 123 LGVATAAQITAGIALAEAREAKKDIALIKDSIVKTHNSVELIQRGIGEQIIALKTLQDFVNDEIRPAIGELRCET--TALKLG-IKLTQHYSELATAFSS 219
FUS_SENDA 127 LGVATSAQITAGIALAEAREAKRDIALIKESMTKTHKSIELLQNAVGEQILALKTLQDFVNDEIKPAISELGCET--AALRLG-IKLTQHYSELLTAFGS 223
FUS_PI3H4 120 LGVATSAQITAAVALVEAKQARSDIEKLKEAIRDTNKAVQSVQSSIGNLIVAIKSVQDYVNKEIVPSIARLGCEA--AGLQLG-IALTQHYSELTNIFGD 216
PRE
POST
F-pep
FUS_HRSV1 238 SVNAGVTTPLSTY---MLTNSELLSLIN-DMPITNDQKKLMSSN---VQIVRQQSYSIMSIIKEEVLAYV 300
FUS_HRSVL 238 SVNAGVTTPVSTY---MLTNSELLSLIN-DMPITNDQKKLMSNN---VQIVRQQSYSIMSIIKEEVLAYV 300
FUS_SV5 210 QITNPALSPITIQALRILLGSTLPTVVEKSFNTQISAAELLSSGLLTGQIVGLDLTYMQMVIKIELPTLT 279
FUS_NDVA 224 QITSPALTQLTIQALYNLAGGNMDYLLTKLGVGNNQLSSLIGSGLITGNPILYDSQTQLLGIQVTLPSVG 293
FUS_9PARA 220 NLGTIGEKSLTLQALSSLYSANITEILSTTKKDKSDIYDIIYTEQVKGTVIDVDLEKYMVTLLVKIPILS 289
FUS_SENDA 224 NFGTIGEKSLTLQALSSLYSANITEIMTTIRTGQSNIYDVIYTEQIKGTVIDVDLERYMVTLSVKIPILS 293
FUS_PI3H4 217 NIGSLQEKGIKLQGIASLYRTNITEIFTTSTVDKYDIYDLLFTESIKVRVIDVDLNDYSITLQVRLPLLT 286
PRE
POST
B
Proteolytic cleavage
Cys-Cys
PRE-FUSION
(FUS_SV5; 2B9B)
C
Cys-Cys
POST-FUSION
(FUS_PI3H4; 1ZTM)

Figure S1

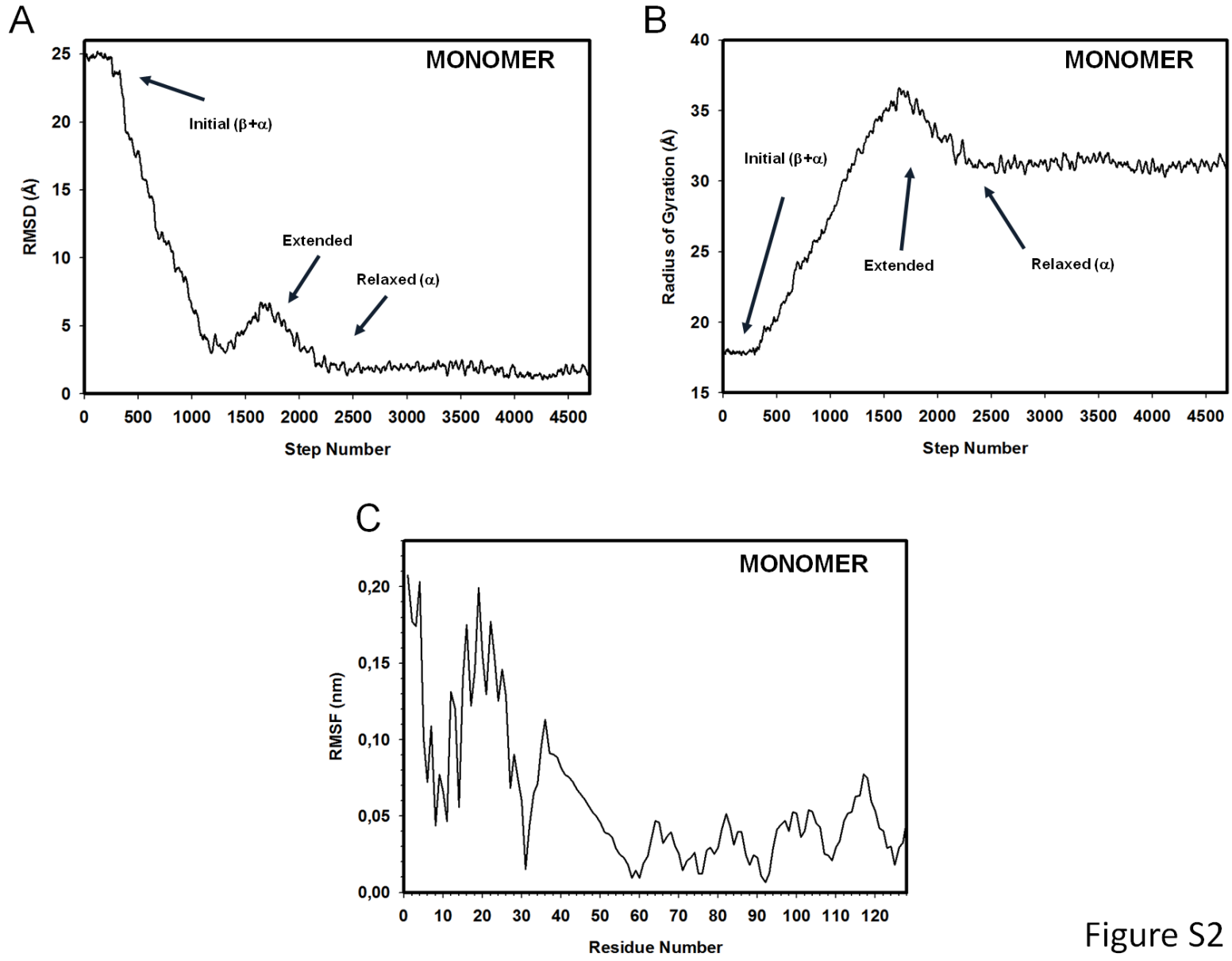
A
MONOMER
Initial (β+α)
Extended
Relaxed (α)
RMSD (Å)
Step Number
B
MONOMER
Initial (β+α)
Extended
Relaxed (α)
Radius of Gyration (Å)
Step Number
C
MONOMER
RMSF (nm)
Residue Number

Figure S2

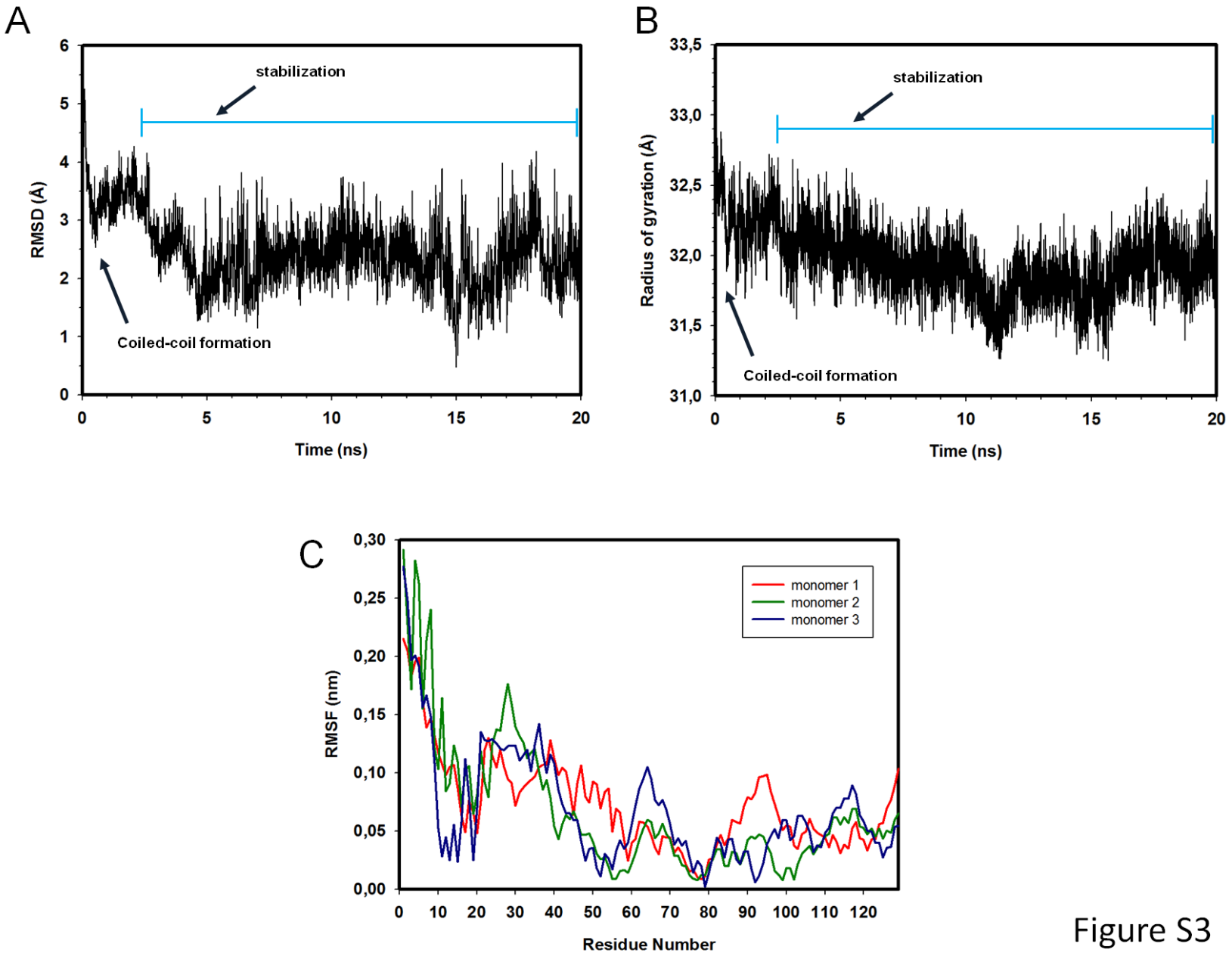
A
RMSD (Å)
Time (ns)
stabilization
Coiled-coil formation
B
Radius of gyration (Å)
Time (ns)
stabilization
Coiled-coil formation
C
RMSF (nm)
Residue Number
monomer 1
monomer 2
monomer 3

Figure S3

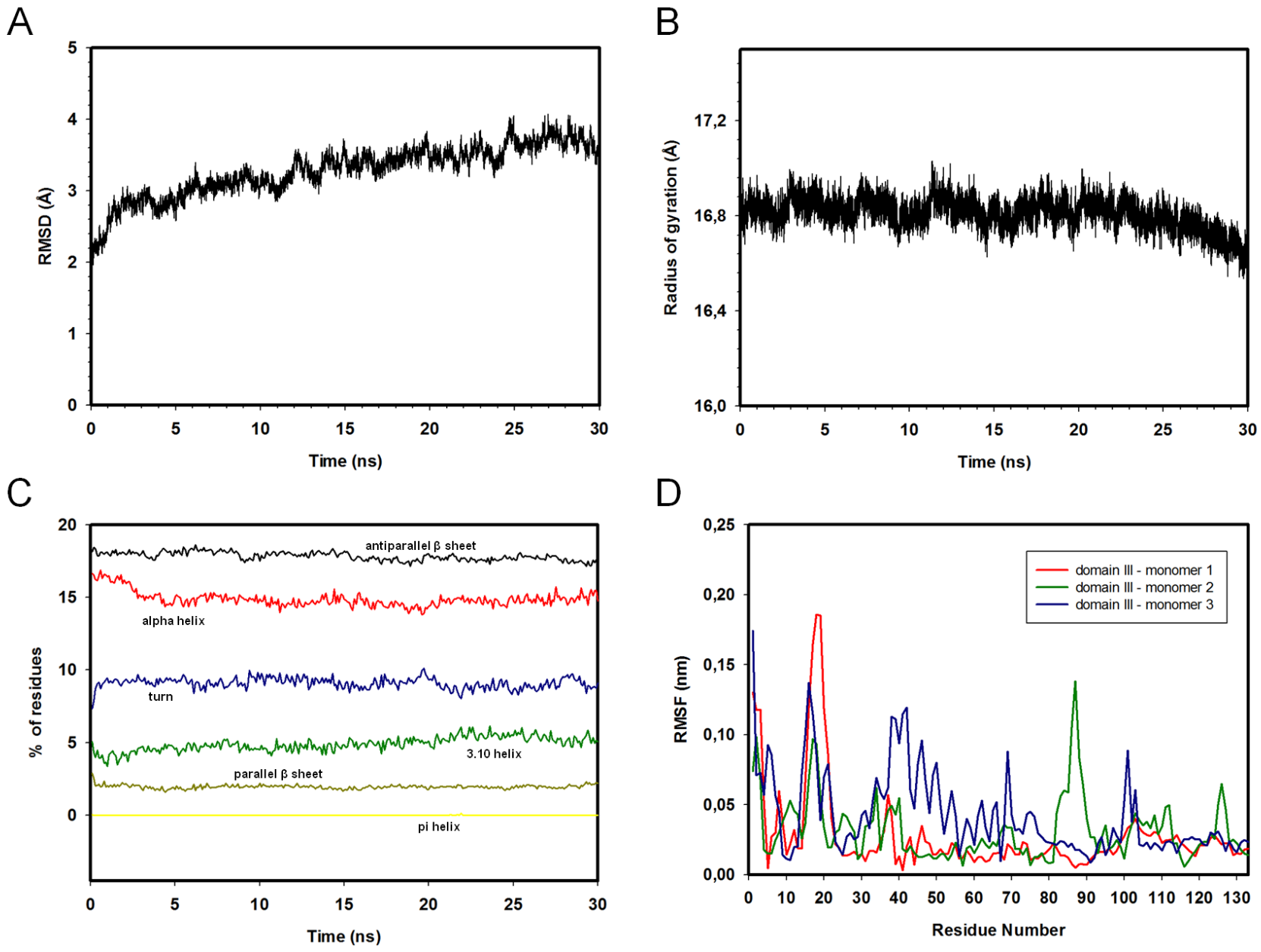
A
RMSD (Å)
Time (ns)
B
Radius of gyration (Å)
Time (ns)
C
% of residues
antiparallel β sheet
alpha helix
turn
3.10 helix
parallel β sheet
pi helix
Time (ns)
D
RMSF (nm)
domain III - monomer 1
domain III - monomer 2
domain III - monomer 3
Residue Number

Figure S4