

University of Groningen

## Stability, conformational plasticity, oligomerization behaviour and equilibrium unfolding intermediates of the Ebola virus matrix protein VP40

Buzon, Pedro; Ruiz-Sanz, Javier; Martinez, Jose C.; Luque, Irene

*Published in:*  
Journal of Biomolecular Structure & Dynamics

*DOI:*  
[10.1080/07391102.2019.1671226](https://doi.org/10.1080/07391102.2019.1671226)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2020

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*  
Buzon, P., Ruiz-Sanz, J., Martinez, J. C., & Luque, I. (2020). Stability, conformational plasticity, oligomerization behaviour and equilibrium unfolding intermediates of the Ebola virus matrix protein VP40. *Journal of Biomolecular Structure & Dynamics*, 38(14), 4289-4303.  
<https://doi.org/10.1080/07391102.2019.1671226>

### Copyright

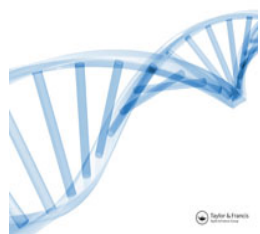
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/tbsd20>

# Stability, conformational plasticity, oligomerization behaviour and equilibrium unfolding intermediates of the Ebola virus matrix protein VP40

Pedro Buzon , Javier Ruiz-Sanz , Jose C. Martinez & Irene Luque

To cite this article: Pedro Buzon , Javier Ruiz-Sanz , Jose C. Martinez & Irene Luque (2020) Stability, conformational plasticity, oligomerization behaviour and equilibrium unfolding intermediates of the Ebola virus matrix protein VP40, Journal of Biomolecular Structure and Dynamics, 38:14, 4289-4303, DOI: [10.1080/07391102.2019.1671226](https://doi.org/10.1080/07391102.2019.1671226)

To link to this article: <https://doi.org/10.1080/07391102.2019.1671226>



View supplementary material [↗](#)



Accepted author version posted online: 30 Sep 2019.  
Published online: 09 Oct 2019.



Submit your article to this journal [↗](#)



Article views: 99



View related articles [↗](#)



View Crossmark data [↗](#)

## Stability, conformational plasticity, oligomerization behaviour and equilibrium unfolding intermediates of the Ebola virus matrix protein VP40

Pedro Buzon<sup>a</sup> , Javier Ruiz-Sanz, Jose C. Martinez  and Irene Luque

Faculty of Sciences, Department of Physical Chemistry and Institute of Biotechnology, University of Granada, Granada, Spain

Communicated by Ramaswamy H. Sarma

**Abbreviations:** CD: circular dichroism; DLS: dynamic light scattering; DSC: differential scanning calorimetry; DTT: dithiothreitol; EDTA: ethylenediaminetetraacetic acid; IPTG: isopropyl- $\beta$ -D-1-thiogalactopyranoside; LB: Luria Bertani; NMR: nuclear magnetic resonance; SH3: Src-homology 3 domain; PDZ: post-synaptic density disc-large zonula occludens domain

### ARTICLE HISTORY

Received 3 July 2019

Accepted 16 September 2019

### Introduction



Ebola virus is included in the Filoviridae family, together with Marburg virus, and classified as a category A pathogen by the NIH. Four different strains of Ebola virus have been identified, with Zaire species showing the highest level of fatality in humans. Ebola virus forms lipid-enveloped particles with filamentous morphology, where each virion includes one molecule of single-stranded, negative-sense RNA that encodes seven genes. Using a common strategy found in many other viruses, Ebola virus overcomes the limitations associated with a short genome by producing efficient multi-functional proteins.

VP40 is a multi-domain protein which is able to perform very diverse tasks during the viral life-cycle of Ebola virus. VP40 is a peripheral membrane protein that plays a central role in Ebola virus pathogenicity, being involved in budding, capsid assembly, cellular trafficking, and regulation of transcription; in fact, it is the most abundant protein found in virions (Bornholdt et al., 2013). The multifunctional character of VP40 has been widely associated with a remarkable conformational plasticity, which allows this protein to oligomerize in various forms, each with its respective three-dimensional arrangement of monomers. Thus, a number of crystallographic studies have shown a variety of oligomerization states that contribute to the viral life-cycle at different stages (Bornholdt et al., 2013; Radzimanowski, Effantin, & Weissenhorn, 2014; Saranya, Shankar, & Vijayakumar, 2019; Stahelin, 2014). According to the structural arrangement of VP40 monomers within such oligomers, an N-terminal domain organized by residues 1–195 and a C-terminal domain composed of residues 196–326 can be clearly distinguished. Molecular dynamics simulations have shown that

various salt-bridge interactions modulate the inter-domain coupling (Gc et al., 2016). These studies also describe that the disengagement of the C-terminal domain from the N-terminal one is a crucial step in the conformational transition of dimers to hexamers, both essential for cellular trafficking, matrix assembly, and virus budding (Bornholdt et al., 2013; Saranya et al., 2019; Stahelin, 2014). Besides, octameric-ring oligomers are involved in the regulation of the viral transcription (Bornholdt et al., 2013). In addition, two cationic patches and one hydrophobic loop within the C-terminal domain have been described as adaptable membrane interaction regions (Bornholdt et al., 2013; Radzimanowski et al., 2014).


Despite recent efforts, the energetic basis and mechanisms of the conformational transformations that VP40 undergoes during Ebola life-cycle still remain poorly understood. Thus, the profuse X-ray structural studies have to be complemented with biophysical experiments in solution to understand the dynamics and energetics of protein regions, that is, the cooperativity of the whole structural arrangement. Furthermore, the analysis of the oligomerization equilibria would be of invaluable help for the design of drug candidates to block the Ebola life-cycle at different stages, assuming that the oligomerization state is the main discriminator among VP40 functionalities (Stahelin, 2014).

In this work, we have analyzed such properties through the thermodynamic characterization of VP40 unfolding equilibrium under different buffered conditions by means of calorimetric and spectroscopic techniques. Multi-functional proteins are usually arranged in different structural domains. The three domains of monomeric streptokinase represent the simplest scenario, where each domain acts as an independent and cooperative structural unit (Azuaga, Dobson,

**CONTACT** Jose C. Martinez  [jcmh@ugr.es](mailto:jcmh@ugr.es)  Faculty of Sciences, Department of Physical Chemistry and Institute of Biotechnology, University of Granada, Avda. Fuentenueva, s/n, 18071 Granada, Spain.

<sup>a</sup>Present address: Molecular Biophysics, Zernike Instituut, Rijksuniversiteit Groningen, 9747 AG-Groningen, The Netherlands.

This article has been republished with minor changes. These changes do not impact the academic content of the article.

 Supplemental data for this article is available online at <https://doi.org/10.1080/07391102.2019.1671226>.

© 2019 Informa UK Limited, trading as Taylor & Francis Group

Mateo, & Conejero-Lara, 2002). Unfortunately, multi-domain proteins rarely present such a simplified scheme of unfolding equilibria, since the macrostates can be also involved in oligomerization equilibria, thus displaying a rather complex behaviour upon unfolding. The full thermodynamic characterization of the process typically becomes rather challenging in these cases. Nonetheless, the conformational and energetic information derived from these intricate equilibria is essential to understand, not only the folding plasticity and stability, but also the functionalities of these multi-domain proteins (Privalov, 1982). Despite the complex behaviour observed, we have been able to obtain a complete thermodynamic analysis of the thermal unfolding and stability of VP40.

## Results

### Characterization of VP40 conformational equilibrium at neutral pH by differential scanning calorimetry (DSC)

DSC experiments were carried out in 50 mM sodium phosphate 150 mM sodium chloride pH 7.5. Under these solvent conditions VP40 is a dimer (Bornholdt et al., 2013). We confirmed the dimeric state by size-exclusion chromatography (Figure S1).

The thermal unfolding profiles of VP40 obtained by DSC show two well-defined transitions (Figure 1). VP40 unfolding was fully irreversible when heated at 80 °C, however, it is highly reversible (at least 70%) when heating was stopped at 60 °C, before reaching the maximum of the second transition (Figure S2). These experiments thus reveal that the thermal denaturation of VP40 protein is partly reversible. In addition, we carried out DSC experiments at different scan rates and we did not observe thermal effects deriving from kinetic effects others than a slight shifting of no more than 4–5 degrees along the temperature axis when the heating rate was varied from 0.5 to 2 K·min<sup>-1</sup> (Figure S3). Shifting was more pronounced in the second transition. These observations drive to the conclusion that such kinetic effects do not have a relevant contribution to unfolding thermodynamics, since they do not distort significantly the shape of DSC endotherms, which will account at the end for the thermodynamic parameters derived from model analysis. In fact, it has been described elsewhere (Ibarra-Molero, Naganathan, Sanchez-Ruiz, & Munoz, 2016) that if the temperature-induced unfolding proceeds much faster than aggregation, thermodynamics can be applied to the data analysis even if reversibility is partial. Therefore, when relevant changes in shape and/or area are not taking place with scan rate, the thermodynamic information derived from these traces will not be too much affected by kinetic processes.

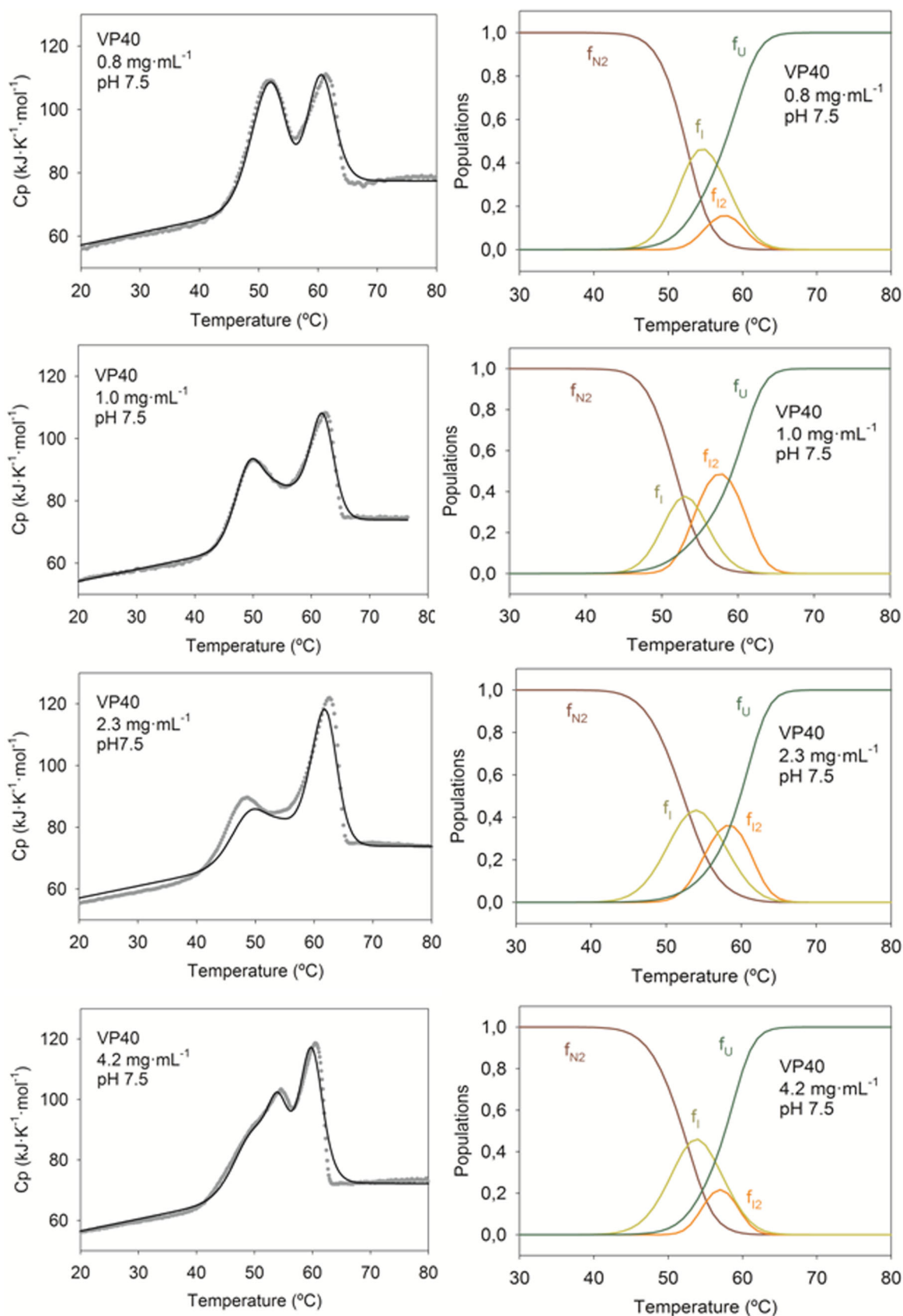
An inspection of Figure 1 shows that protein concentration affects the two unfolding transitions observed experimentally: the first transition moves towards lower temperatures when protein concentration increases, whereas the second roughly seems to be moving in the opposite direction although the changes are not very large. These features indicate that more than two macrostates are populated upon unfolding and that some of the intermediate states

undergo oligomerization. It is well established that for oligomerization equilibria an increase in protein concentration leads to a gain in population of the oligomeric species (Filimonov & Rogov, 1996; Murciano-Calles, Cobos, Mateo, Camara-Artigas, & Martinez, 2010). The simplest scheme accounting for these observations must involve an intermediate macrostate  $I$  with a defined oligomeric state  $n$ , that is,  $1/2 N_2 \rightleftharpoons 1/n I_n \rightleftharpoons U$ . Nevertheless, the application of this model could not reproduce the shapes of VP40 experimental traces. We decided to include an additional intermediate macrostate on the scheme,  $1/2 N_2 \rightleftharpoons 1/n I_n \rightleftharpoons I \rightleftharpoons U$ , where  $I$  represents a monomeric intermediate of VP40.

The stoichiometry,  $n$ , of the association-dissociation equilibria can be accurately obtained from the evaluation of protein concentration effects on DSC traces through model analysis of the whole set of traces (Filimonov & Rogov, 1996; Murciano-Calles et al., 2010). We performed a global fitting of our experiments to the  $1/2 N_2 \rightleftharpoons 1/n I_n \rightleftharpoons I \rightleftharpoons U$  model considering a common heat capacity function for the four macrostates, and leaving the values of mid-point temperatures, enthalpies and  $n$  (characterizing the  $I_n$  state) as adjustable parameters. After the fitting session we obtained a value of  $n = 1.8 \pm 0.4$ . Thus, we established the model described by the Scheme 1 in the Materials and Methods section, being the one used to analyze the experimental DSC traces.

Although in a first approach we carried out a global fitting strategy to establish the stoichiometry and the definitive thermodynamic model, we decided to analyze each DSC experiment individually using Scheme 1 in order to assess if the model was able to describe the intricate shapes of each endotherm. In addition, since the model parameters cannot be affected by protein concentration when the rest of experimental conditions are kept (Filimonov & Rogov, 1996), the differences observed among experiments analyzed individually would reveal the true uncertainties associated to these fitting results. The best fit of each DSC experiment is seen in the left panels of Figure 1, where we observe that the model can acceptably describe the complex unfolding of VP40. The major deviations appear at the end of the traces where the irreversible/kinetic phenomena mainly occur.

According to the equations described in the Materials and Methods section we have 14 floating parameters characterizing the thermodynamic model of Scheme 1: 8 describing the heat-capacity linear functions of the four macrostates; besides 3 mid-point temperatures and 3 enthalpy changes describing the three equilibria selected from Scheme 1. The thermodynamic magnitudes with pure fitting errors are displayed in Table 1, where the good  $R$ - and  $R$ -square values show that convergence has been achieved in all examples despite the number of floating parameters. Only in the DSC experiment at 1.0 mg·mL<sup>-1</sup> the fitting errors of the independent terms ( $a$ ,  $c$ ,  $e$ ,  $g$ ) included in the four heat capacity functions were too high. Thus, we decided to fix them to the average value calculated from the other traces. As a consequence, the fitting errors of enthalpy values are noticeably higher than in the other examples (Table 1), despite the quality of the fitting is comparable in all DSC traces of Figure 1.

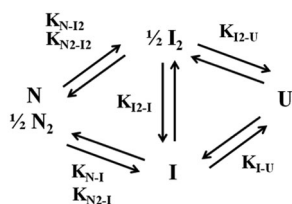


**Figure 1.** Left panels: thermal unfolding profiles of VP40 monitored by DSC as a function of protein concentration. The best individual fittings to the model of Scheme I are shown as black solid lines through the experimental data (gray symbols). Right panels: distribution of populations of the different VP40 states for the different protein concentrations assayed. Brown lines show the respective populations of the native state ( $f_{N2}$ ), light-green lines the corresponding to the monomeric intermediate ( $f_i$ ), orange lines the ones of the dimeric intermediate ( $f_{i2}$ ), and dark-green lines the respective of the unfolded state ( $f_U$ ). Experimental conditions were 50 mM sodium phosphate 150 mM sodium chloride pH 7.5.



In any case, the different parameters compare well in the interval of protein concentrations assayed, although their variability is clearly higher than the reflected by fitting errors. The major discrepancies appear between the enthalpies of the three equilibria (Table 1). Nevertheless, the sum of the three enthalpic contributions for every experiment compare better, which indicates that the thermodynamic parameters reporting for the global unfolding ( $N \rightleftharpoons U$ ), including the stability of the protein, can be estimated more precisely than those of partial equilibria.

Let us remark that all individual fittings were carried out without any parameter restrictions; in spite of which, we have obtained a solid convergence within the results (Figure 1 and Table 1). Therefore, the model is capable of describing the protein-concentration dependency of the DSC traces. The experimental deviations from ideal protein-concentration behaviour would arise from the irreversible phenomena. As an alternative explanation, such deviations might be due to the population of other intermediates upon thermal unfolding, but we are convinced that the robustness of our fitting results does not justify the inclusion of additional *I*-macrostates in the unfolding Scheme 1, since they are not expected to populate significantly. In fact, the population of the two already considered intermediate macrostates is not more than 40% at any temperature (Figure 1). In addition, we appreciate that the monomeric intermediate populates more than the dimeric one (with the exception of the trace at 1 mg·mL<sup>-1</sup>, although in this case the analysis gave rise to some troubles already described which can account for these



**Scheme 1.** The four-state equilibrium model for the analysis of VP40 DSC traces.

**Table 1.** Thermodynamic parameters of the thermal unfolding of the dimeric VP40 domain, obtained in 50 mM sodium phosphate 150 mM sodium chloride pH 7.5 at different protein concentrations. Parameters were obtained from the analysis of DSC experiments using Scheme 1 and the sequence of equilibria described by Equation (13).<sup>a</sup>

VP40 concentration (mg·mL <sup>-1</sup> )	0.8	1.0 <sup>b</sup>	2.3	4.2
<i>a</i> (kJ/K <sup>1</sup> ·mol <sup>1</sup> )	-58.09 ± 0.05		-57.22 ± 0.03	-59.42 ± 0.09
<i>b</i> (kJ/K <sup>2</sup> ·mol <sup>1</sup> )	0.3933 ± 0.0003	0.3843 ± 0.0004	0.3899 ± 0.0003	0.3956 ± 0.0003
<i>c</i> (kJ/K <sup>1</sup> ·mol <sup>1</sup> )	103.23 ± 0.10		109.30 ± 0.06	95.44 ± 0.10
<i>d</i> (kJ/K <sup>2</sup> ·mol <sup>1</sup> )	-0.2177 ± 0.0005	-0.216 ± 0.011	-0.2086 ± 0.0001	-0.2417 ± 0.0003
<i>e</i> (kJ/K <sup>1</sup> ·mol <sup>1</sup> )	96.16 ± 0.12		98.27 ± 0.06	98.68 ± 0.15
<i>f</i> (kJ/K <sup>2</sup> ·mol <sup>1</sup> )	-0.2100 ± 0.0002	-0.301 ± 0.018	-0.2101 ± 0.0001	-0.1973 ± 0.0004
<i>g</i> (kJ/K <sup>1</sup> ·mol <sup>1</sup> )	79.45 ± 0.11		75.88 ± 0.08	74.23 ± 0.08
<i>h</i> (kJ/K <sup>2</sup> ·mol <sup>1</sup> )	(-5.8091 ± 0.0000)·10 <sup>-3</sup>	(-6.4157 ± 0.0006)·10 <sup>-3</sup>	(-5.9020 ± 0.0000)·10 <sup>-3</sup>	(-5.8679 ± 0.0000)·10 <sup>-3</sup>
<i>T</i> <sub>N2-12</sub> (K)	328.74 ± 0.08	327.0 ± 0.4	328.28 ± 0.08	328.27 ± 0.05
<i>T</i> <sub>12-1</sub> (K)	328.05 ± 0.20	324.5 ± 1.5	328.09 ± 0.12	329.99 ± 0.09
<i>T</i> <sub>1-U</sub> (K)	329.72 ± 0.06	329 ± 1	330.76 ± 0.06	329.46 ± 0.03
$\Delta H_{N2-12}(T_{N2-12})$ (kJ·mol <sup>1</sup> )	381.5 ± 0.5	377 ± 17	293.1 ± 0.3	374.6 ± 0.3
$\Delta H_{12-1}(T_{12-1})$ (kJ·mol <sup>1</sup> )	-201.4 ± 0.5	-175 ± 75	-209.8 ± 0.2	-220.9 ± 0.3
$\Delta H_{1-U}(T_{1-U})$ (kJ·mol <sup>1</sup> )	342.6 ± 0.7	465 ± 15	456.7 ± 0.4	422.3 ± 0.5
$\Delta G_{N-U}(298)$ (kJ·mol <sup>1</sup> )	41.1	38.5	43.1	49.0

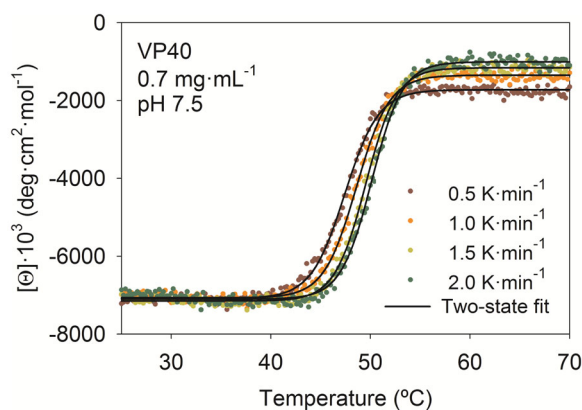
The error intervals are pure fitting standard errors. The values of thermodynamic magnitudes for the  $I_2 \rightleftharpoons I$  equilibrium were estimated for a  $P_{ref} = 100 \mu\text{M}$ .  $R^-$  and  $R^+$  parameters were: 0.994 and 0.989 for 0.84 mg·mL<sup>-1</sup> protein concentration; 0.997 and 0.993 for 1.0 mg·mL<sup>-1</sup>; 0.986 and 0.972 for 2.3 mg·mL<sup>-1</sup>; and 0.992 and 0.983 for 4.2 mg·mL<sup>-1</sup>.

<sup>b</sup>The *a*, *c*, *e*, *g* values in this example were fixed to the average value of the respective values shown in Tables 1 and 2.

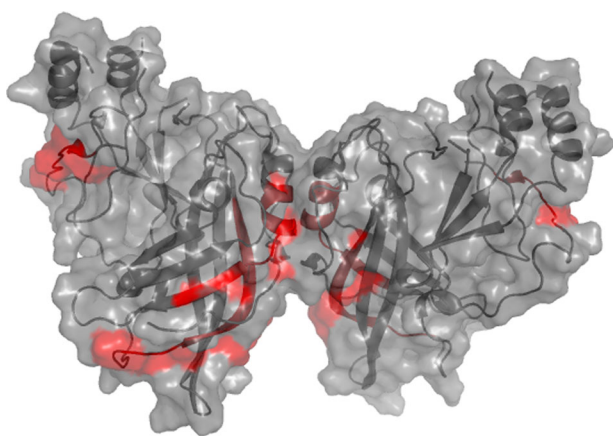
differences). Besides, the population analysis of VP40 reveals the proper sequence of events upon unfolding. Thus, we can clearly observe in the right panels of Figure 1 that the monomeric intermediate populates at lower temperatures than the dimeric one, being  $1/2 N_2 \rightleftharpoons I \rightleftharpoons 1/2 I_2 \rightleftharpoons U$  the most appropriate sequence of equilibria with increasing temperature.

### Structural characterization and oligomerization analysis of VP40 conformational equilibrium

Despite the complex energetic profile of VP40 shown by DSC thermograms, when analyzing the thermal unfolding in 50 mM sodium phosphate 150 mM sodium chloride pH 7.5 by far-UV circular dichroism (CD) the experiments simplify dramatically (Figure 2). The transitions in Figure 2 reflect scan-rate effects similar to those observed by DSC. However, their S-shapes are characteristic of a simple two-state unfolding evolution, where the mid-point temperature,  $T_m$ , is roughly approaching to the maximum temperature of the first transition of the DSC thermograms. This coincidence



**Figure 2.** Thermal denaturation profiles followed by circular dichroism (CD) signal at 222 nm of VP40 in 50 mM sodium phosphate 150 mM sodium chloride pH 7.5 (symbols). Protein concentration was 0.7 mg·mL<sup>-1</sup>. The lines through experimental data correspond to the best fitting of the data to the two-state equilibrium model.



**Figure 3.** The Tango analysis of beta-aggregation-prone regions in dimeric VP40. The surface representation of the VP40 structure (PDB code: 4LDI) is represented in gray, as well as the secondary structure elements, while the residues with some propensity to aggregate are shown in red. The image was done using PyMol (Delano Scientific).

indicates that the whole structural change happens upon the  $N_2 \rightleftharpoons 2I$  endothermic heat evolution, upon which more than 80% of secondary-structure content is lost (Figures 2 and S4). These features also point towards a structural similarity between the  $I$  and  $I_2$  macrostates, since the association-dissociation component described by the  $2I \rightleftharpoons I_2$  equilibrium and the evolution of the intermediates to the unfolded state,  $I_2 \rightleftharpoons 2U$ , were not revealed by this structural technique, indicating that both equilibria are not accompanied of noticeable changes in secondary structure.

Considering the low secondary structure content of the intermediates, the Tango algorithm was applied in order to identify the protein regions driving oligomerization of these intermediates. Actually, it was capable of predicting the main association region of an unfolding intermediate in the PDZ3 domain of PSD95, being confirmed by NMR (Murciano-Calles et al., 2010). Tango calculations with VP40 sequence showed that beta-aggregation propensities are negligible over the whole protein with the exception of a few short sequences, including residues 112–123 (TAAIMLASYTIT) with roughly 50% propensity, residues 171–183 (YFTFDLTALKLIT) with 30%, and residues 303–307 (LTMVI) with 20%. An alternative analysis of the amyloidogenic regions using Waltz web server confirmed and extended one of the regions, comprising residues 157–162 and 170–175 (FLQEFV and QYFTFD), and revealed another one comprising residues 236–241 and 245–253 (KIQAIM and QDFKIVPID). The latter region is located close to the region of 20% propensity for beta-aggregation, and it was also predicted by Tango but with very low beta-aggregation propensity (<5%). The tridimensional arrangement of these residues in the X-ray structure of dimeric VP40 (PDB code: 4LDI; Figure 3) shows that the majority of residues are not on the surface of the protein. Hence, the increase in solvent exposure of these aggregation-prone regions upon the  $N_2 \rightleftharpoons 2I$  transition may develop in the protein some tendency to self-associate.

To obtain additional experimental evidence of the association/dissociation phenomena observed in the DSC analysis, we performed DLS experiments under the same experimental conditions, heating the VP40 samples at 20  $\mu$ M from 25 to

70 °C at a heating rate of 1 K $\cdot$ min $^{-1}$ . At 25 °C we observed a single species with  $4.5 \pm 1.0$  nm hydrodynamic radius, determined as a dimer according to the molecular weight calculation. Dimers disappeared from the solution at roughly 40 °C, giving place to other species of higher sizes (Figure S5). These results qualitatively agree with the DSC population analysis (Figure 1), although we did not find evidence of any dimeric species others than the native state. The S-shape of the unfolding curve drawn from the thermal evolution of the hydrodynamic radius (Figure S5) reflects that the size increases progressive and cooperatively until the unfolded state, which displays the largest aggregation propensity. The  $T_m$  value is very similar to that derived from the CD-thermal unfolding curve (Figure 2), revealing a certain coupling between the main structural unfolding event,  $N_2 \rightleftharpoons 2I$ , and the arrangement of the resulting intermediates as oligomers,  $2I \rightleftharpoons I_2$ . Actually, according to the population analysis performed from DSC results (Figure 1), this conversion takes place within few degrees.

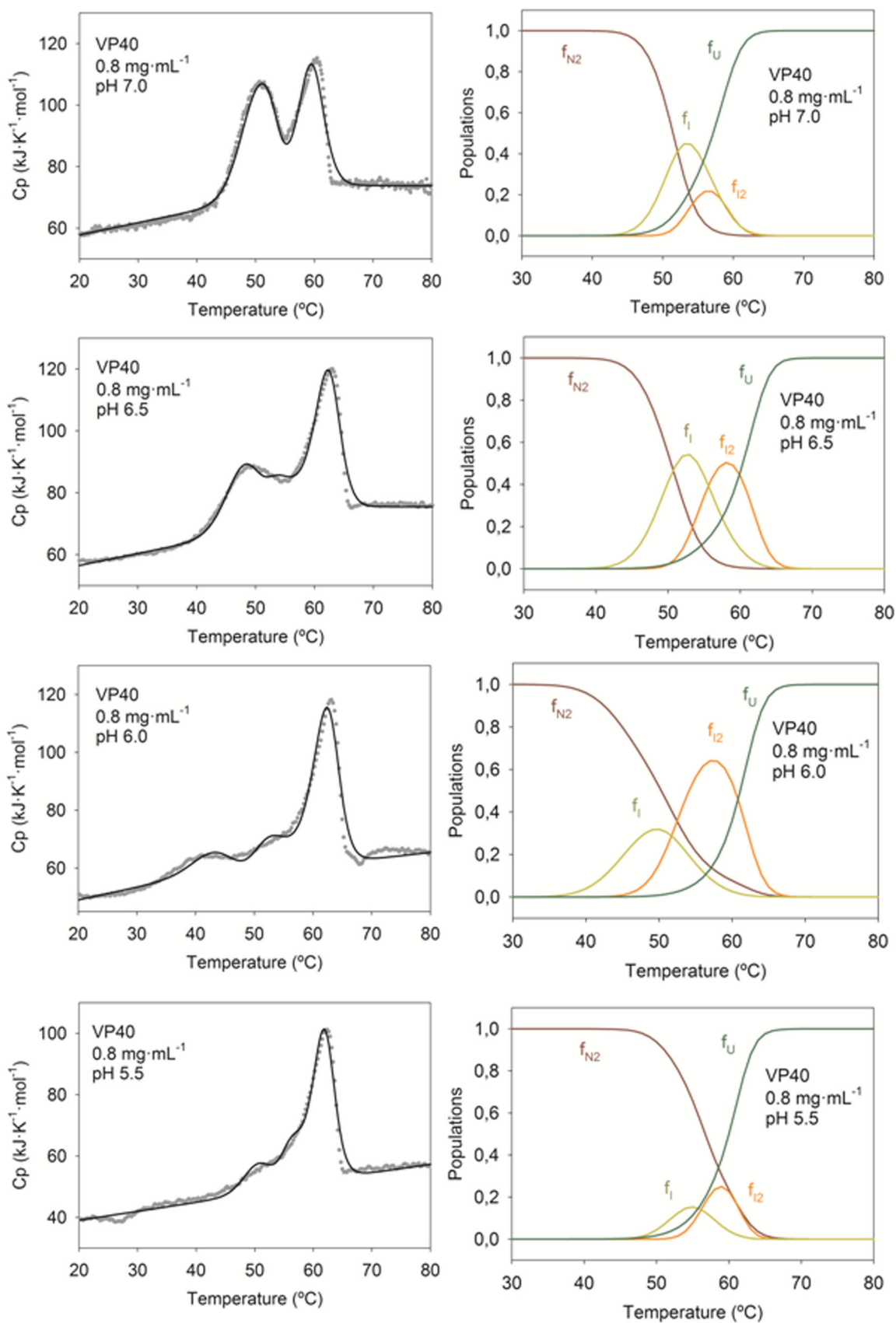
In summary, even though the thermodynamic analysis should be performed using the simplest model that can explain DSC behaviour, we cannot discard the possibility of having an even more complex unfolding, where oligomerization events result from a mixture of dimers and bigger associations. We could not find any evidence of fibrils formation upon incubation, which may well indicate that such aggregates are mostly amorphous arrangements.

#### pH dependence of VP40 conformational equilibrium

When we extended our DSC unfolding experiments to buffered solutions at acidic pH values we observed a similar behaviour than at neutral pH (Figure 4). Nevertheless, at the lower pH conditions the biphasic behaviour almost disappeared showing a main endotherm preceded of small heat effects. In any case, all traces can be described by the model of Scheme 1, being the quality of the fittings surprisingly good, as revealed by Figure 4 and by their  $R$ -square parameters (Table 2). Only the DSC experiments at pH 6.0 and 5.5 showed high fitting errors of the independent terms ( $a$ ,  $c$ ,  $e$ ,  $g$ ) corresponding to the four heat capacity functions, so we fixed them to an average value as we previously did.

The population analysis keeps the tendency to favour the  $I_2$  state instead of  $I$  when pH decreases (Figure 4), maintaining the same order of appearance with temperature than at pH 7.5. Thus, we can conclude that the sequence of equilibria remains unaltered with pH, which strongly suggests that electrostatics do not play a relevant role in the association/dissociation of  $I$ -macrostates.

The heat capacity functions for the different macrostates remain the same (Tables 1 and 2), with the exception of the trace at pH 5.5. The disappearance of the first transition with pH depletion is reflected in the decrease of  $\Delta H_{N_2-I_2}(T_{N_2-I_2})$  values that actually vanish at pH 5.5. On the other hand, the second transition slightly increases, revealed by the growing values of  $\Delta H_{I-U}(T_{I-U})$ , whereas the  $T_{I-U}$  values are not affected by pH. The fade-away of the  $\Delta H_{N_2-I_2}(T_{N_2-I_2})$  enthalpy indicates that the native state has vanished. In fact, the hydrodynamic



**Figure 4.** Left panels: thermal unfolding profiles of VP40 monitored by DSC as a function of pH. Protein concentration was kept at  $0.8\text{ mg}\cdot\text{mL}^{-1}$ . Experimental conditions were 50 mM sodium phosphate 150 mM sodium chloride at pH values 6.5 and 7.0; 50 mM MES 150 mM sodium chloride pH 6.0; and 50 mM sodium acetate 150 mM sodium chloride pH 5.5. The best fitting to the model of Scheme 1 is shown as black solid lines through the experimental data (gray symbols). Right panels: distribution of populations of the different VP40 states for the different pH values assayed. Brown lines show the respective populations of the native state ( $f_{N2}$ ), light-green lines the corresponding to the monomeric intermediate ( $f_i$ ), orange lines the ones of the dimeric intermediate ( $f_{i2}$ ), and dark-green lines the respective of the unfolded state ( $f_U$ ).



**Table 2.** Thermodynamic parameters of the thermal unfolding of the dimeric VP40 domain, obtained at protein concentration of 0.85 mg·mL<sup>-1</sup>. Experimental conditions were 50 mM buffer 150 mM sodium chloride at different pH values. Parameters were obtained from the analysis of DSC experiments using Scheme 1 and the sequence of equilibriums described by Equation (13).<sup>a</sup>

Buffer pH	Sodium phosphate pH 7.0	Sodium phosphate pH 6.5	MES pH 6.0 <sup>b</sup>	Sodium acetate pH 5.5 <sup>b</sup>
<i>a</i> (kJ/K <sup>1</sup> ·mol <sup>1</sup> )	-58.514 ± 0.08	-62.4 ± 0.4		
<i>b</i> (kJ/K <sup>2</sup> ·mol <sup>1</sup> )	0.3969 ± 0.0005	0.4091 ± 0.0013	0.40 ± 0.10	0.29 ± 0.03
<i>c</i> (kJ/K <sup>1</sup> ·mol <sup>1</sup> )	102.04 ± 0.15	110.4 ± 0.8		
<i>d</i> (kJ/K <sup>2</sup> ·mol <sup>1</sup> )	-0.2222 ± 0.0004	-0.219 ± 0.002	-0.2 ± 1.0	-0.3 ± 5.1
<i>e</i> (kJ/K <sup>1</sup> ·mol <sup>1</sup> )	94.54 ± 0.09	94.6 ± 0.6		
<i>f</i> (kJ/K <sup>2</sup> ·mol <sup>1</sup> )	-0.2146 ± 0.0002	-0.2129 ± 0.0016	-0.2 ± 3.5	-0.3 ± 2.2
<i>g</i> (kJ/K <sup>1</sup> ·mol <sup>1</sup> )	75.74 ± 0.09	74.8 ± 0.2		
<i>h</i> (kJ/K <sup>2</sup> ·mol <sup>1</sup> )	-5.6469·10 <sup>-3</sup> ± 0.0000	-7.5807·10 <sup>-3</sup> ± 0.0001	7.0·10 <sup>-3</sup> ± 0.05	0.28 ± 0.04
<i>T</i> <sub>N2-12</sub> (K)	327.40 ± 0.09	325.58 ± 0.09	324.2 ± 1.5	333.9 ± 0.9
<i>T</i> <sub>12-1</sub> (K)	326.1 ± 0.2	324.04 ± 0.11	320.5 ± 2.2	308 ± 59
<i>T</i> <sub>I-U</sub> (K)	328.89 ± 0.07	329.11 ± 0.06	328.3 ± 1.2	328.9 ± 1.0
$\Delta H_{N2-12}(T_{N2-12})$ (kJ·mol <sup>1</sup> )	374.9 ± 0.8	371.1 ± 2.9	227 ± 73	-0.9 ± 12.1
$\Delta H_{12-1}(T_{12-1})$ (kJ·mol <sup>1</sup> )	-197.4 ± 0.2	-236.9 ± 2.2	-201 ± 69	182 ± 215
$\Delta H_{I-U}(T_{I-U})$ (kJ·mol <sup>1</sup> )	385.5 ± 0.5	541.2 ± 3.0	575 ± 125	374 ± 29
$\Delta G_{N-U}(298)$ (kJ·mol <sup>1</sup> )	48.0	51.6	52.0	48.9

<sup>a</sup>The error intervals are pure fitting standard errors. The values of thermodynamic magnitudes for the I<sub>2</sub> ⇌ I equilibrium were estimated for a P<sub>ref</sub> = 100 μM. R<sup>1</sup> and R<sup>2</sup> parameters were: 0.991 and 0.981 for pH 7.0; 0.988 and 0.976 for pH 6.5; 0.976 and 0.952 for pH 6.0; and 0.991 and 0.982 for pH 5.5.

<sup>b</sup>The a, c, e, g values in these examples were fixed to the average value of the respective values shown in Tables 1 and 2.

radius of VP40 at pH 5.5 (and below) increases (Table S1), indicating that the native state may not be fully folded. In addition, although these changes might be mostly due to tertiary structure, the far-UV CD spectrums at the different pH values show small but visible differences between pH 5.5 (and below) and the others (Figure S6).

It is quite probable that at 5.5 and lower pH values the salt-bridges responsible of the integrity of VP40, through the assembling of the N- and C-terminal domains (Gc et al., 2016) (see Introduction), would not be as stable as they are at neutral pH due to the intervention of protonation equilibriums of the charged residues organizing such salt-bridges. Thus, the X-ray structure of dimeric VP40 (PDB code: 4LDI) shows that the rest of acidic residues are widespread through the protein surface and exposed to the solvent (Figure S7).

### Characterization of monomeric L117R-VP40 conformational equilibrium at neutral pH

Full-length VP40 has been reported to be a dimer in solution at neutral and acidic pH. The dimeric interface of VP40 is located in a hydrophobic region of the N-terminal domain involving the residues A55, H61, F108, T112, A113, M116 and L117 (Figure S7). The introduction of a positively charged residue at the dimer interface through the L117R mutant disrupts the VP40 quaternary structure producing a monomeric variant (Bornholdt et al., 2013).

In order to investigate the unfolding behaviour of monomeric VP40, we set up the experimental conditions for L117R-VP40 assays to 50 mM sodium phosphate 150 mM sodium chloride pH 7.5. Under these conditions, a hydrodynamic radius of 3.4 ± 0.3 nm was estimated at 25 °C by DLS for L117R-VP40, confirming its monomeric character. Such monomeric state was also confirmed by size-exclusion chromatography (Figure S1).

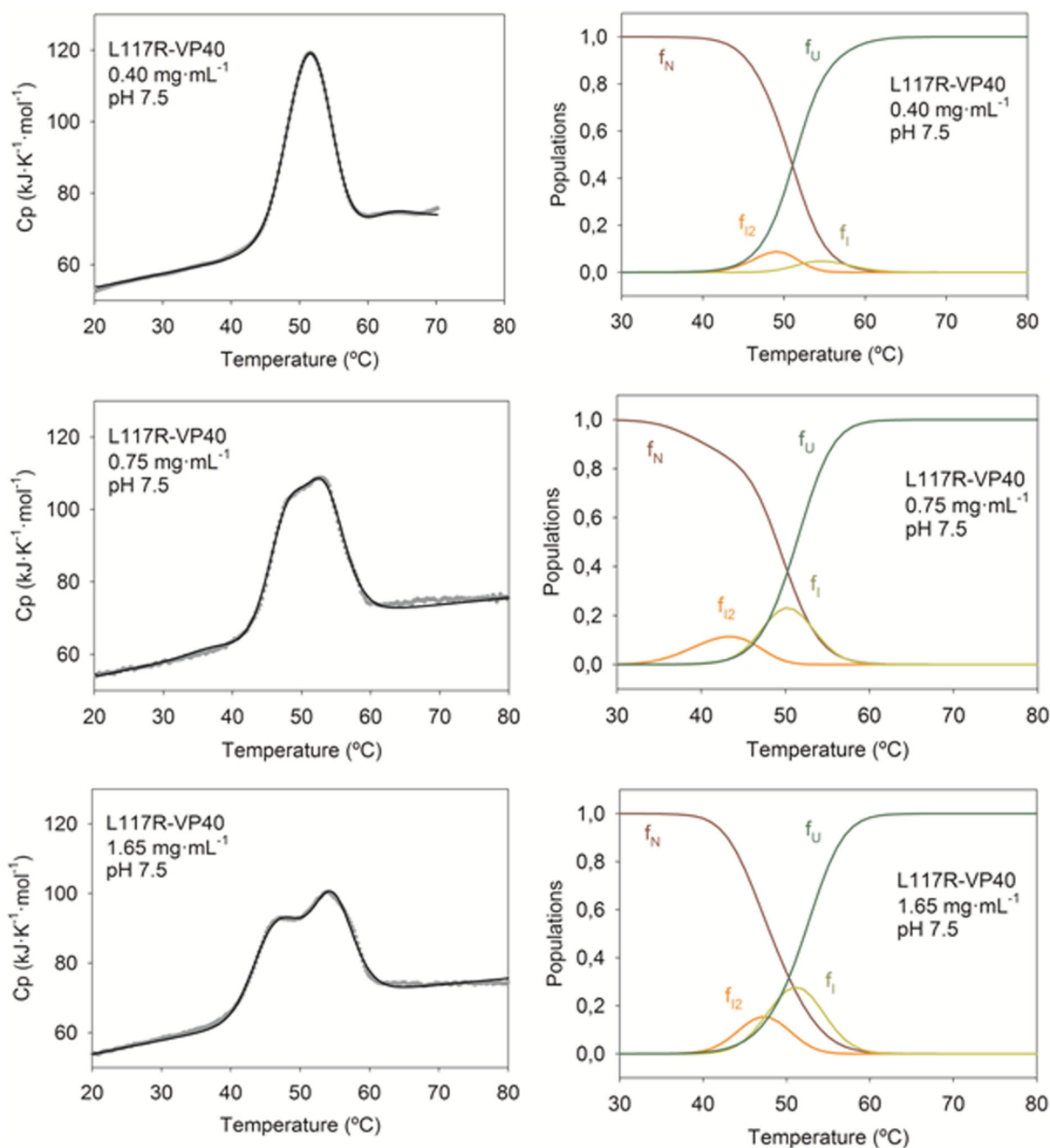
DSC experiments were carried out at three different protein concentrations. In spite of its monomeric character,

L117R-VP40 thermal unfolding shows a concentration dependency (Figure 5): at the lowest protein concentrations L117R-VP40 presents a single transition that separates into two well defined peaks when concentration increases. Thus, the biphasic character essentially remains, but appears at higher protein concentrations. The second transition presents a mid-point temperature very similar to that of the wild type VP40 and is only barely affected by protein concentration. Conversely, the first transition moves towards lower temperatures in the L117R-VP40 mutant.

The same arrangement of equilibriums describing dimeric VP40 unfolding (Scheme 1) is able to fully describe the thermal unfolding of monomeric L117R-VP40. In order to determine the stoichiometry of the process, we similarly performed a global fitting considering common heat capacity functions for the four macrostates in all traces of Figure 5, keeping the values of mid-point temperatures, enthalpies and *n* (characterizing the I<sub>n</sub> state) free to float. The results for L117R-VP40 confirmed also a value of *n* = 2. The quality of the global fitting to the model shown by Scheme 1 (see Materials and Methods section for details) is seen in Figure 5, observing that this scheme is capable of describing the protein-concentration dependency of DSC traces. The thermodynamic parameters with fitting standard errors are displayed in Table 3, where we appreciate that convergence has been achieved satisfactorily, having reasonable error intervals.

As we have already seen for dimeric VP40, when analyzing the thermal unfolding of L117R-VP40 under the same experimental conditions using far-UV CD, the experiments display a simple two-state unfolding evolution (Figure 6), pointing towards a structural similarity between the I and the I<sub>2</sub> macrostates. Thus, the far-UV CD transition does not reflect any additional structural changes, related with the association/dissociation of the intermediates and their evolution to the unfolded state.

The population analysis shown in Figure 5 describes the proper sequence of events upon thermal unfolding of L117R-VP40. Thus, the three equilibriums arrange with temperature



**Figure 5.** Left panels: thermal unfolding profiles of L117R-VP40 monitored by DSC as a function of protein concentration. Experimental conditions were 50 mM sodium phosphate 150 mM sodium chloride pH 7.5. The best fitting to the model of Scheme 1 is shown as black solid lines through the experimental data (gray symbols). Right panels: distribution of populations of the different VP40 states for the different protein concentrations assayed. Brown lines show the respective populations of the native state ( $f_N$ ), light-green lines the corresponding to the monomeric intermediate ( $f_I$ ), orange lines the ones of the dimeric intermediate ( $f_{I_2}$ ), and dark-green lines the respective of the unfolded state ( $f_U$ ).

as  $N \rightleftharpoons 1/2 I_2 \rightleftharpoons I \rightleftharpoons U$ , where conformational transitions are accompanied by association/dissociation thermal effects, thus revealing the complexity of the process. The only relevant difference we appreciate when the populations for dimeric VP40 and L117R-VP40 are compared (right panels of Figures 1, 4 and 5) is related with the alternated thermal stabilities displayed by the intermediates. Thus, for the wild type the monomeric species is the most populated at low temperatures, whereas for the L117R mutant the most populated species at low temperatures is the dimer. Interestingly, the region with highest beta-aggregation propensity predicted by Tango (112-TAAIMLASYTIT-123) partially overlaps the

hydrophobic interface of the dimeric arrangement, which would somehow explain the different thermal stabilities found for  $I$  and  $I_2$  species, since the mutation of L117 by the charged Arg may interfere with oligomerization, thus destabilizing the  $I_2$  state.

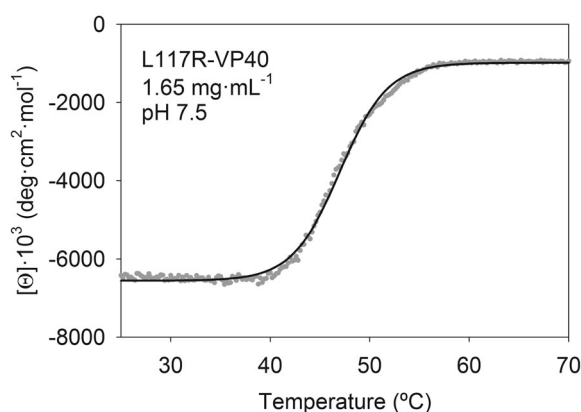
## Discussion

Protein size, as well as multi-domain arrangement, can be determinant in the folding/unfolding behaviour of globular proteins. From a conformational point of view, we can

**Table 3.** Thermodynamic parameters of the thermal unfolding of the monomeric L117R-VP40 domain, obtained in 50 mM sodium phosphate 150 mM sodium chloride pH 7.5 at different protein concentrations. Parameters were obtained from the analysis of DSC experiments using Scheme 1 and the sequence of equilibriums described by Equation (1).<sup>a</sup>

L117R-VP40 concentration (mg·mL <sup>-1</sup> )	1.65	0.75	0.40
<i>a</i> (kJ/K <sup>1</sup> ·mol <sup>-1</sup> )	-62.6 ± 0.2	Idem	Idem
<i>b</i> (kJ/K <sup>2</sup> ·mol <sup>-1</sup> )	0.3972 ± 0.0007	Idem	Idem
<i>c</i> (kJ/K <sup>1</sup> ·mol <sup>-1</sup> )	-19.9 ± 0.3	Idem	Idem
<i>d</i> (kJ/K <sup>2</sup> ·mol <sup>-1</sup> )	0.1216 ± 0.0003	Idem	Idem
<i>e</i> (kJ/K <sup>1</sup> ·mol <sup>-1</sup> )	43.96 ± 0.15	Idem	Idem
<i>f</i> (kJ/K <sup>2</sup> ·mol <sup>-1</sup> )	(-1.6389 ± 0.1) · 10 <sup>-6</sup>	Idem	Idem
<i>g</i> (kJ/K <sup>1</sup> ·mol <sup>-1</sup> )	9.96 ± 0.04	Idem	Idem
<i>h</i> (kJ/K <sup>2</sup> ·mol <sup>-1</sup> )	0.1859 ± 0.0002	Idem	Idem
<i>T</i> <sub>N-1</sub> (K)	323.74 ± 0.05	326.06 ± 0.06	330.84 ± 0.06
<i>T</i> <sub>1-12</sub> (K)	321.37 ± 0.07	322.55 ± 0.11	331.55 ± 0.09
<i>T</i> <sub>12-U</sub> (K)	321.89 ± 0.04	322.13 ± 0.07	324.63 ± 0.03
Δ <i>H</i> <sub>N-1</sub> ( <i>T</i> <sub>N-1</sub> ) (kJ·mol <sup>-1</sup> )	210.9 ± 0.6	99 ± 2	227.5 ± 0.1
Δ <i>H</i> <sub>1-12</sub> ( <i>T</i> <sub>1-12</sub> ) (kJ·mol <sup>-1</sup> )	-287.7 ± 1.1	-360 ± 6	-285.0 ± 0.2
Δ <i>H</i> <sub>12-U</sub> ( <i>T</i> <sub>12-U</sub> ) (kJ·mol <sup>-1</sup> )	404.7 ± 0.8	548 ± 2	368.1 ± 0.2
Δ <i>G</i> <sub>N-U</sub> (298) (kJ·mol <sup>-1</sup> )	27.6	30.2	28.9

The error intervals are pure fitting standard errors. The values of thermodynamic magnitudes for the I ⇌ I<sub>2</sub> and I<sub>2</sub> ⇌ U equilibriums were estimated for a P<sub>ref</sub> = 100 μM. R- and R<sup>2</sup>- parameters were 0.997 and 0.98 respectively.



**Figure 6.** Thermal denaturation profile followed by circular dichroism signal at 222 nm of L117R-VP40 in 50 mM sodium phosphate 150 mM sodium chloride pH 7.5 (gray symbols). Protein concentration was 1.65 mg·mL<sup>-1</sup>. The line through experimental data corresponds to the best fitting of the data to the two-state equilibrium model.

assume that small globular proteins are usually integrated by a single cooperative unit, which is well justified by natural evolution to avoid the population of misfolded species that will drive to a multitude of organic and functional complications in living organisms. Thus, a simple two-state model characterizes the equilibriums of many small proteins, like the exhaustively studied hen egg-white lysozyme and bovine pancreatic ribonuclease A (~14 kDa) (Privalov, 1979), and small protein domains such as the Src Homology 3 (SH3; ~7 kDa) (Viguera, Martinez, Filimonov, Mateo, & Serrano, 1994).

Nonetheless, most of the known proteins do not follow such ideal two-state behaviour. In some cases, equilibrium intermediates have been found to be populated to some extent in small proteins and domains under certain experimental conditions. When size does not increase by many kDa, such intermediates become compulsory in the unfolding pathway, even though the equilibrium conditions usually weaken due to irreversible protein aggregation upon thermal denaturation.

When facing more complex equilibriums, the thorough thermodynamic analysis has demonstrated to be very useful,

not only to understand the energetic aspects of the unfolding equilibriums, but also to derive information on the structural organization of the different cooperative subunits that may unfold more or less independently. It is well known that the partition function of the system can be obtained by a double integrative calculation of experimental DSC traces. The partition function includes all the information on thermodynamic states of the protein in the temperature range studied. Traditionally, such information has been obtained from DSC experiments through a deconvolution analysis, where no assumptions have to be made on the number of states of the system and on their equilibriums. In practice, the deconvolution analysis is able to discern how many two-state transitions involve the DSC experiment, assuming that everyone will correspond to a cooperative assembling of the protein.

Pepsinogen is perhaps the best example of such deconvolution studies (Privalov, 1982). It is a globular protein of 40 kDa like VP40; however, it displays a single cooperative transition that cannot be depicted by a two-state model. The deconvolution analysis showed that the transition can be described by two two-state equilibriums characterized by different temperatures and enthalpies, which were thus ascribed to two different structural domains. These results were supported by parallel studies with pepsin, differing from pepsinogen in a fragment of 44-residues, where both transitions were clearly distinguished in the DSC traces, indicating that both cooperative domains unfold more or less independently (Privalov, 1982). Similar conclusions were obtained from papain studies, where the structural information available confirmed the two structural domains revealed by DSC (Privalov, 1982). Even more surprising is the DSC study with intact plasminogen, revealing seven transitions (Privalov, 1982), where the parallel study of several protein fragments was an invaluable help to identify such cooperative units into the structure.

Nevertheless, the deconvolution analysis does not consider the possibility of oligomerization of macrostates, describing all equilibriums to occur among monomeric assemblies. Therefore, the analyses mentioned above have been performed under the assumption that protein

concentration does not alter the DSC unfolding traces of the proteins under study. Today we know that proteins can develop misfolding routes where intermediates and unfolded states play a predominant role, mainly because their poor structural arrangement is prone to self-aggregation. Let us remark that a similar thermal unfolding behaviour to the described here for VP40 has been previously seen in other examples, as the already mentioned PDZ domains (Murciano-Calles et al., 2010), CheY (Filimonov et al., 1993), or Cro repressor (Fabian et al., 1999; Filimonov & Rogov, 1996), where the unfolding intermediates and their association/dissociation processes were also studied in depth by a variety of biophysical approaches, being able even to organize into amyloid fibrils and annular arrangements (Murciano-Calles et al., 2010). What is surprising in all these examples is that the two unfolding transitions observed by DSC do not correspond with the denaturation of protein structural domains. According to spectroscopic experiments, the first transition, related to the equilibrium between the native and one of the intermediate states, involves the most relevant structural changes, whereas the second transition mainly consist of heat effects derived from association/dissociation phenomena of the intermediates. The protein concentration dependence of VP40 DSC traces reflects such oligomerization phenomena, being demonstrated by DLS as well. Consequently, our unfolding experiments with VP40 are not an exception of this rule, showing also that circular dichroism is unable to distinguish between the  $I$ ,  $I_2$  and  $U$  states, displaying all of them a poor structural content. Nevertheless, the fact that two structural domains are clearly observed in the crystallographic structures of VP40, could lead to the wrong assumption that the unfolding of each structural domain may well correspond with each one of both DSC transitions. Since it has been described that the inter-domain coupling is achieved by various salt-bridge interactions (Gc et al., 2016), we can conclude that these interactions are strong enough to build a single cooperative unit in VP40, being both structural domains undistinguishable by the unfolding experiments carried out in this work, even those at acidic pH values.

The unfolding of dimeric wild type VP40, as well as of monomeric L117R-VP40, can be summarized by the equilibrium Scheme 1, despite the different oligomeric nature of their native states. Thus, the oligomerization of the native state does not seem to alter the unfolding pathway of VP40. The population analysis shows that monomeric intermediates appear at higher temperatures than dimers in L117R-VP40, whereas the opposite is true for wild type VP40 (right panels of Figures 1, 4 and 5). We cannot rely with certainty on these small differences in thermal stability of both intermediates, despite the persistency in all the panels of Figures 1, 4 and 5, since the CD analysis reveals that they are both very similar structurally. In any case, Tango analysis recognizes the dimerization interface to partly overlap with the major aggregation region of VP40 (residues 112–123; TAAIMLASYTIT) with roughly 50% propensity. The mutation of L117 by the charged Arg may interfere with oligomerization, thus destabilizing the  $I_2$  state. Therefore, these observations point

towards the relevance of this region in the dimerization arrangements of both the native and intermediate states. We can also presume that other region(s) will participate in the arrangement of the  $I_2$  state, since the L117R mutation was enough to disrupt the dimeric native state, but unable to prevent the formation of the  $I_2$  state upon thermal unfolding.

The conformational stability of dimeric VP40 at neutral pH can be estimated from the average of the  $\Delta G_{N-U}(298)$  values collected in Table 1 to be around 40–50 kJ·mol<sup>-1</sup>. This value is noticeably higher than the 25–30 kJ·mol<sup>-1</sup> exhibited by the monomeric L117R-VP40 under the same solvent conditions (Table 3), indicating that the quaternary structure contributes significantly to the whole stability of the protein, which is not surprising from a thermodynamic point of view in a protein prone to oligomerize in different ways. The comparison of far-UV circular dichroism spectrums of both proteins supports such argument, at least in the sense of a gain in secondary structure of the dimer (Figure S8). From a thermodynamic point of view, the stabilization of the quaternary structure can be mostly of enthalpic nature, since the unfolding enthalpies (estimated as the sum of enthalpic contributions collected in Tables 1 and 3 respectively) are clearly higher for the dimer than for the monomer. Accordingly, molecular dynamics studies have found that five salt-bridges are responsible of the interactions between monomers, being them surrounded of a hydrophobic environment afforded by the VP40 dimer interface (Gc et al., 2016).

The stability value is comparable to the one obtained for the 414 residues-long streptokinase (Azuga et al., 2002), where the three structural domains were isolated and studied by DSC, as well as for the above mentioned pepsinogen (Privalov, 1982), of similar size as VP40. Nonetheless, this conformational stability is not far away from the obtained for PDZ domains (Murciano-Calles et al., 2010), Chey (Filimonov et al., 1993) or Cro-repressor (Fabian et al., 1999; Filimonov & Rogov, 1996), even though VP40 triplicates their size. Protein stability is not expected to increase with protein size indefinitely, mainly when unfolding allows the organization of equilibrium intermediates in detriment of folding cooperativity. Actually, the integrative DSC studies of Privalov and Pfeil (Privalov, 1982) with medium- and large-size proteins indicate that the upper limit of  $\Delta G_{N-U}(298)$  values does not exceed 60 kJ·mol<sup>-1</sup>. Still, the total Gibbs energy will not account for the true stability of the native state when constituted by several cooperative subunits, being defined by the stability of the least stable one. In the case of VP40 the DSC and CD experiments do not show such partitioning of Gibbs energies among structural domains, vanishing most of the structural content in a single and cooperative equilibrium, and being the second transition an energetic consequence of the association/dissociation equilibria of intermediates and unfolded states. Therefore, we believe that these additional energetic contributions would also be included in our stability estimations.

Additional analyses of the DSC traces point towards the reasons of this relatively low stability of VP40. Thus, theoretical heat capacity functions of the native and unfolded states



can be obtained from molecular weight and amino acid composition, respectively (Cobos et al., 2009). These functions are well reproduced by small globular proteins but it is not the case of VP40 (Figure S9), which shows experimental values of  $C_{pU}(T)$  that are lower than the predicted values. This can be easily explained by the irreversible phenomena that occur upon unfolding. However, the values of experimental  $C_{pN}(T)$  and  $C_{pN2}(T)$  functions are higher than predicted and indicate a high conformational flexibility of the native state, since they reflect dynamic fluctuations within microstates, which would be higher than usual in VP40, thus being the native state not as well defined as in small globular proteins (Cobos et al., 2009). Since the slopes of these pre-transition heat capacities parallel the predicted one, we can conclude that, despite the high flexibility, such dynamic fluctuations would have a non-cooperative nature, being thus established between iso-enthalpic conformations, and suggesting a well-defined energetic barrier for the equilibrium between the native and intermediate macrostates in VP40 (Cobos et al., 2009). Supporting these arguments, we have obtained very low values for specific enthalpies, which have been described to converge to  $55 \text{ J}\cdot\text{g}^{-1}$  at  $110^\circ\text{C}$ . Our calculations with wild type VP40 and L117R-VP40 provide values not higher than  $10 \text{ J}\cdot\text{g}^{-1}$ , although the high error associated to this long extrapolation of the unfolding enthalpy can question such value. In any case, the total unfolding enthalpy derived from Tables 1 to 3 points in the same direction, being surprising that the unfolding enthalpy of VP40 is comparable to those of typical small globular proteins as lysozyme or ribonuclease A (Privalov, 1979). A similar picture has been described for Cro repressor (Filimonov & Rogov, 1996), where the structural analysis has revealed an unusually high exposure of hydrophobic groups to the solvent in the native state.

The explanation of the marginal stability displayed by all proteins, independently of their sizes, is obviously related to the functional character of protein structures. Thus, proteins need a well defined arrangement, stable against environmental fluctuations, but adaptable in time. As reported by Privalov (Privalov, 1982), the reserve of stability must exceed the energy of thermal motion, RT, so no less than  $12 \text{ kJ}\cdot\text{mol}^{-1}$  at  $25^\circ\text{C}$ . The cooperation of 30 residues can be enough to achieve such energy. The upper limit of a cooperative unit has been established by these studies in roughly 300 residues, which is approximately the size of VP40 (326 amino acids). The auto-assembling nature of proteins may well preclude bigger cooperative units, giving rise to multi-domain arrangements with longer sequences.

## Conclusions

In summary, the thermodynamic analysis reveals that VP40 behaves as a single-domain protein that unfolds by a complex scheme where association/dissociation equilibria of the different macrostates occur. Thus, the N- and C-terminal domains seen by the crystallographic studies do not appear to be independent cooperative units, since when the salt-bridges responsible of their mutual interaction are titrated

the native state of the protein vanishes. Let us remark that the structural studies describe that the disengagement of the C-terminal domain from the N-terminal one is a crucial step in the conformational transition of dimers to other functional oligomeric arrangements (Bornholdt et al., 2013; Stahelin, 2014). Within this context, the electrostatics of the VP40-domains interface might be modulated by the physiological conditions and the membrane micro-environments, as well as by the spatial- and time-dependent interaction with the respective partners, which at last will induce the conformational changes driving to each specific functionality. Thus, the interaction of the C-terminal domain with the membrane has been described as responsible of hexamerization of VP40 from the dimeric form. Molecular dynamics have also shown the relevant role of D45-K326 salt-bridge in guiding domain association (Gc et al., 2016).

## Materials and methods

### Protein samples

The full-length wild-type VP40, matrix protein of Zaire Ebola virus (strain Mayinga), and the L117R-VP40 mutant, were expressed using pETM-11 vector (EMBL Core Facilities) in *Rosetta2(DE3)* *E. coli* cells (Novagen). The poly-histidine tagged constructs were expressed in bacterial cultures, LB-Kanamycin medium, at  $37^\circ\text{C}$  up to an optical density at 600 nm of approximately 0.3, when the temperature was lowered to  $25^\circ\text{C}$ . Then, the cultures were induced with 0.5 mM IPTG at an optical density at 600 nm of 0.4–0.5, and allowed to grow at  $25^\circ\text{C}$  overnight. Bacterial pellets were resuspended in lysis buffer, 50 mM  $\text{NaH}_2\text{PO}_4$  300 mM sodium chloride 20 mM imidazole pH 8.0, including a Protease Inhibitor Cocktail tablet (Roche); and lysed by sonication. The clarified lysate was bound in a 5 mL prepacked Ni-Sepharose column (His Trap HP from GE Healthcare). The column was washed twice with washing buffer 1 (50 mM  $\text{NaH}_2\text{PO}_4$  300 mM sodium chloride 20 mM imidazole pH 8.0), as well as with washing buffer 2 (50 mM  $\text{NaH}_2\text{PO}_4$  300 mM sodium chloride 50 mM imidazole pH 8.0), and then the protein of interest was eluted with elution buffer (50 mM  $\text{NaH}_2\text{PO}_4$  300 mM sodium chloride 250 mM imidazole pH 8.0). The aliquots containing wild type VP40 and L117R-VP40 proteins were buffer exchanged using desalting columns PD-10 (GE Healthcare) into 50 mM sodium phosphate 0.5 mM EDTA 1 mM DTT pH 8.0. Under these conditions, the poly-histidine tag was removed from the protein sequence by proteolysis with Tobacco Etch Virus protease. The product was dialyzed into 10 mM HEPES 100 mM sodium chloride pH 7.5, and loaded into a strong cation exchange column (HiTrap SP HP from GE Healthcare). Proteins were eluted by performing a buffer gradient up to 15% of 10 mM HEPES 1 M sodium chloride pH 7.5 buffer. Finally, size exclusion chromatography in 10 mM HEPES 300 mM sodium chloride pH 7.5, using a Superdex 200 16/60 column (GE Healthcare), was performed. Fractions corresponding to wild type VP40 and L117R-VP40 were kept at  $4^\circ\text{C}$ , and used within few days.

A theoretical extinction coefficient of  $19940 \text{ cm}^{-1}\cdot\text{M}^{-1}$  was calculated for both, VP40 and L117R-VP40 in ProtParam-



ExPASy server. Molecular masses of 35368 Da (VP40) and 35411 Da (L117R-VP40) were obtained by MALDI-TOF experiments carried out at the Scientific Instrumentation Services (CIC) of the University of Granada. Experimental samples were always prepared by extensive dialysis against the respective buffer at 4 °C.

### General description of the experimental procedures

Differential scanning calorimetry (DSC) experiments were carried out indistinctly in an automated VP-DSC instrument from Microcal INC (Northampton, MA), as described previously (Viguera et al., 1994). DSC traces were generated at 2 K·min<sup>-1</sup>, unless specified otherwise. Dynamic light scattering (DLS) and circular dichroism (CD) experimental details can also be obtained from previous references, as well as the two-state analysis conducted in the CD-thermal denaturation experiments (Cobos et al., 2009). The molecular weight of oligomers was calculated from their respective hydrodynamic radii by using the “globular protein model” included in the software of the DLS instrument (DynaPro, DYNAMICS V6, Wyatt Technology Corporation, Santa Barbara, CA). Analytical size exclusion chromatography combined with DLS measurements were performed in a Superdex 200 10/30 HR (GE-Healthcare, Fairfield, CT) coupled to a Zetasizer μV (Malvern Panalytica). The column was calibrated by means of different markers of known molecular weight: blue dextran (to evaluate the void volume; 2000 kDa), cytochrome C (12.4 kDa), bovine serum albumin (66 kDa), myoglobin (17 kDa), aprotinin (6.5 kDa) and pepsin (35 kDa).

### Evaluation of thermodynamic magnitudes by a four-state model analysis of DSC traces

The established four-state model including monomeric and dimeric intermediates can be mathematically defined as it is in Scheme 1:

Any alternative order of equilibriums connecting the four macrostates are completely equivalent from a thermodynamic point of view, giving rise to absolutely equal fitting results, but referring to different equilibrium aspects.

#### Monomer version



We can define the equilibrium constants that characterize the processes of Scheme 1 as

$$K_{N-I_2} = \frac{[I_2]^{1/2}}{[N]} \quad K_{I_2-I} = \frac{[I]}{[I_2]^{1/2}} \quad K_{I-U} = \frac{[U]}{[I]} \quad (2)$$

$$K_{N-I} = \frac{[I]}{[N]} \quad K_{I-I_2} = \frac{[I_2]^{1/2}}{[I]} \quad K_{I_2-U} = \frac{[U]}{[I_2]^{1/2}}$$

in which it is obvious that  $K_{N-U} = K_{N-I_2} \cdot K_{I_2-I} \cdot K_{I-U} = K_{N-I} \cdot K_{I-I_2} \cdot K_{I_2-U}$ , being  $K_{I_2-I} = 1/K_{I-I_2}$ , and therefore  $\Delta H_{N-U} = \Delta H_{N-I_2} + \Delta H_{I_2-I} + \Delta H_{I-U} = \Delta H_{N-I} + \Delta H_{I-I_2} + \Delta H_{I_2-U}$  and  $\Delta H_{I_2-I} = -\Delta H_{I-I_2}$ . This means that the thermodynamic parameters of any

arrangement of equilibriums in the model can easily be obtained from these simple mathematical relationships.

We used the order established by Equation (1) for curve fitting analysis to simplify the mathematical formulation. The populations or fractions of the different species can be defined as

$$f_N = \frac{[N]}{P} \quad f_{I_2} = \frac{2[I_2]}{P} \quad f_I = \frac{[I]}{P} \quad f_U = \frac{[U]}{P} \quad (3)$$

where  $P$  is the total protein concentration expressed as monomer in the experiment and  $f_N + f_{I_2} + f_I + f_U = 1$  at any temperature. By replacing the concentration values into the equilibrium constants of Equation (2) we obtain

$$K_{N-I} = \frac{f_I}{f_N} \quad K_{I-I_2} = \frac{f_{I_2}^{1/2}}{(2 \cdot P)^{1/2} \cdot f_I} \quad K_{I_2-U} = \frac{(2 \cdot P)^{1/2} \cdot f_U}{f_{I_2}^{1/2}} \quad (4)$$

which we can combine to solve the fraction of protein in the monomeric intermediate state,  $f_I$ . We thus have

$$[2 \cdot P \cdot (K_{N-I} K_{I_2}^2)] \cdot f_I^2 + [1 + K_{N-I} + K_{N-I} \cdot K_{I-I_2} \cdot K_{I_2-U}] \cdot f_I - K_{N-I} = 0 \quad (5)$$

To solve Equation (5) we use the well known expression

$$f_I = \frac{-B \pm \sqrt{B^2 - 4 \cdot A \cdot C}}{2 \cdot A} \quad (6)$$

where  $A = [2 \cdot P \cdot (K_{N-I} K_{I_2}^2)]$ ;  $B = [1 + K_{N-I} + K_{N-I} \cdot K_{I-I_2} \cdot K_{I_2-U}]$ ; and  $C = -K_{N-I}$ . Considering that the equilibrium constants are always higher/equal than/to zero, we opted by the positive solution of Equation (6) to avoid negative and unrealistic values for  $f_I$ , assumed to range between 0 and 1.

This leaves the temperature dependencies of thermodynamic functions to be estimated. The heat-capacity functions of different equilibrium states have been defined by linear functions of temperature

$$C_{pN} = a + (b \cdot T) \quad C_{pI} = c + (d \cdot T) \quad C_{pI_2} = e + (f \cdot T)$$

$$C_{pU} = g + (h \cdot T) \quad (7)$$

The enthalpy functions for the three equilibriums in Equation (1) were obtained by integrating Kirchoff's equations as

$$\Delta H_{N-I} = \Delta H_{N-I}(T_{N-I}) + [(c - a) \cdot (T - T_{N-I})]$$

$$+ \left[ \frac{(d - b) \cdot (T^2 - T_{N-I}^2)}{2} \right]$$

$$\Delta H_{I-I_2} = \Delta H_{I-I_2}(T_{I-I_2}) + [(e - c) \cdot (T - T_{I-I_2})]$$

$$+ \left[ \frac{(f - d) \cdot (T^2 - T_{I-I_2}^2)}{2} \right]$$

$$\Delta H_{I_2-U} = \Delta H_{I_2-U}(T_{I_2-U}) + [(g - e) \cdot (T - T_{I_2-U})]$$

$$+ \left[ \frac{(h - f) \cdot (T^2 - T_{I_2-U}^2)}{2} \right] \quad (8)$$

where  $T_{N-I}$ ,  $T_{I-I_2}$  and  $T_{I_2-U}$  are the midpoint temperatures of the three equilibriums of Equation (1) respectively (corresponding to population values of 0.5 for the respective macrostates), whereas  $\Delta H_{N-I}(T_{N-I})$ ,  $\Delta H_{I-I_2}(T_{I-I_2})$  and  $\Delta H_{I_2-U}(T_{I_2-U})$  are

the corresponding enthalpy values. In a similar way, the entropy functions can be calculated as

$$\begin{aligned}\Delta S_{N-I} &= \Delta S_{N-I}(T_{N-I}) + \left[ (c - a) \cdot \ln\left(\frac{T}{T_{N-I}}\right) \right] \\ &+ [(d - b) \cdot (T - T_{N-I})] \\ \Delta S_{I-I_2} &= \Delta S_{I-I_2}(T_{I-I_2}) + \left[ (e - c) \cdot \ln\left(\frac{T}{T_{I-I_2}}\right) \right] \\ &+ [(f - d) \cdot (T - T_{I-I_2})] \\ \Delta S_{I_2-U} &= \Delta S_{I_2-U}(T_{I_2-U}) + \left[ (g - e) \cdot \ln\left(\frac{T}{T_{I_2-U}}\right) \right] \\ &+ [(h - f) \cdot (T - T_{I_2-U})]\end{aligned}\quad (9)$$

Where

$$\begin{aligned}\Delta S_{N-I}(T_{N-I}) &= \frac{\Delta H_{N-I}(T_{N-I})}{T_{N-I}} \\ \Delta S_{I-I_2}(T_{I-I_2}) &= \frac{\Delta H_{I-I_2}(T_{I-I_2})}{T_{I-I_2}} + R \cdot \ln(1/\sqrt{P_{ref}}) \\ \Delta S_{I_2-U}(T_{I_2-U}) &= \frac{\Delta H_{I_2-U}(T_{I_2-U})}{T_{I_2-U}} + R \cdot \ln(\sqrt{P_{ref}})\end{aligned}\quad (10)$$

The second term of the association-dissociation entropy function is responsible for the protein concentration dependencies shown in Figure 5, giving rise to a unique set of thermodynamic parameters compatible with all thermal transitions at different protein concentrations. Thus the parameters shown in Table 3 for this equilibrium correspond to a hypothetical experiment made at  $P_{ref}$  concentration. This value has been established arbitrarily as 100  $\mu$ M. Finally, the Gibbs energies can be easily calculated from the well known thermodynamic relations  $\Delta G = \Delta H - T \cdot \Delta S$  and  $\Delta G = -R \cdot T \cdot \ln K$ .

The temperature function of the enthalpy describing the thermal unfolding of L117R-VP40 according to Equation (1) can be expressed as

$$\begin{aligned}H &= H_N + [\Delta H_{N-I} \cdot f_I] + [(\Delta H_{N-I} + \Delta H_{I-I_2}) \cdot f_{I_2}] \\ &+ [(\Delta H_{N-I} + \Delta H_{I-I_2} + \Delta H_{I_2-U}) \cdot f_U]\end{aligned}\quad (11)$$

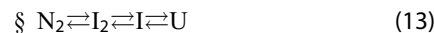
and the corresponding heat-capacity function, which will be fitted to the experimental traces, can easily be obtained by numerical methods, since

$$C_p = \left( \frac{\partial H}{\partial T} \right)_p \quad (12)$$

For the analysis of the L117R-VP40 traces shown in Figure 5 and obtained in 50 mM sodium phosphate 150 mM sodium chloride pH 7.5 at different protein concentrations we assume that a single set of thermodynamic magnitudes should be compatible with all the experimental traces, as it is considered by the model. Therefore, any differences in such thermodynamic magnitudes are consequence of the changes in equilibrium, solely induced by variations in protein concentration. Thus we undertook an overall analysis to obtain the heat-capacity  $C_{pN}(T)$ ,  $C_{pI}(T)$ ,  $C_{pI_2}(T)$  and  $C_{pU}(T)$  linear functions (Equation (7)), as well as the  $T_{N-I}$ ,  $T_{I-I_2}$  and  $T_{I_2-U}$ , and  $\Delta H_{N-I}(T_{N-I})$ ,  $\Delta H_{I-I_2}(T_{I-I_2})$  and  $\Delta H_{I_2-U}(T_{I_2-U})$  parameters related to the three equilibria shown in Equation (1). We carried out

a global fitting analysis, taking as common the heat-capacity linear functions, whereas the rest of these 14 parameters were fitted individually for each experiment in an attempt to assess the realistic error associated to these values, compared in Table 3, which includes the fitting errors. Calculations were made using Sigma Plot 2000 (Systat Software Inc.).

### Dimer version



We can define the equilibrium constants that characterize the processes of Scheme I as

$$\begin{aligned}K_{N_2-I_2} &= \frac{[I_2]^{1/2}}{[N_2]^{1/2}} \quad K_{I_2-I} = \frac{[I]}{[I_2]^{1/2}} \quad K_{I-U} = \frac{[U]}{[I]} \\ K_{N_2-I} &= \frac{[I]}{[N_2]^{1/2}} \quad K_{I-I_2} = \frac{[I_2]^{1/2}}{[I]} \quad K_{I_2-U} = \frac{[U]}{[I_2]^{1/2}}\end{aligned}\quad (14)$$

in which it is obvious that  $K_{N_2-U} = K_{N_2-I_2} \cdot K_{I_2-I} \cdot K_{I-U} = K_{N_2-I} \cdot K_{I-I_2} \cdot K_{I_2-U}$ , being  $K_{I_2-I} = 1/K_{I-I_2}$ , and therefore  $\Delta H_{N_2-U} = \Delta H_{N_2-I_2} + \Delta H_{I_2-I} + \Delta H_{I-U} = \Delta H_{N_2-I} + \Delta H_{I-I_2} + \Delta H_{I_2-U}$  and  $\Delta H_{I_2-I} = -\Delta H_{I-I_2}$ . This means that the thermodynamic parameters of any arrangement of equilibria in the model can easily be obtained from these simple mathematical relationships.

We used the scheme in Equation (13) for curve fitting analysis to simplify the mathematical formulation. The populations or fractions of the different species can be defined as

$$f_{N_2} = \frac{2[N_2]}{P} \quad f_{I_2} = \frac{2[I_2]}{P} \quad f_I = \frac{[I]}{P} \quad f_U = \frac{[U]}{P} \quad (15)$$

where  $P$  is the total protein concentration expressed as monomer in the experiment and  $f_{N_2} + f_{I_2} + f_I + f_U = 1$  at any temperature. By replacing the concentration values into the equilibrium constants of Equation (14) we obtain

$$K_{N_2-I_2} = \frac{f_{I_2}^{1/2}}{f_{N_2}^{1/2}} \quad K_{I_2-I} = \frac{(2 \cdot P)^{1/2} \cdot f_I}{f_{I_2}^{1/2}} \quad K_{I-U} = \frac{f_U}{f_I} \quad (16)$$

which we can combine to solve the fraction of protein in the monomeric intermediate state,  $f_I$ . We thus have

$$\begin{aligned}[2 \cdot P \cdot (1 + K_{N_2-I_2}^2)] \cdot f_I^2 + [K_{N_2-I_2}^2 \cdot K_{I_2-I}^2 \cdot (1 + K_{I-U})] \\ \cdot f_I - K_{N_2-I_2}^2 \cdot K_{I_2-I}^2 = 0\end{aligned}\quad (17)$$

To solve Equation (17) we use the well known expression

$$f_I = \frac{-B \pm \sqrt{B^2 - 4 \cdot A \cdot C}}{2 \cdot A} \quad (18)$$

where  $A = [2 \cdot P \cdot (1 + K_{N_2-I_2}^2)]$ ;  $B = [K_{N_2-I_2}^2 \cdot K_{I_2-I}^2 \cdot (1 + K_{I-U})]$ ; and  $C = -K_{N_2-I_2}^2 \cdot K_{I_2-I}^2$ . Considering that the equilibrium constants are always higher/equal than/to zero, we opted by the positive solution of Equation (18) to avoid negative and unrealistic values for  $f_I$ , assumed to range between 0 and 1.

This leaves the temperature dependencies of thermodynamic functions to be estimated. The heat-capacity functions of different equilibrium states have been defined by

linear functions of temperature

$$C_{pN2} = a + (b \cdot T) \quad C_{pI2} = c + (d \cdot T) \quad C_{pI} = e + (f \cdot T)$$

$$C_{pU} = g + (h \cdot T) \quad (19)$$

The enthalpy functions for the three equilibria in Equation (13) were obtained by integrating Kirchoff's equations as

$$\Delta H_{N2-I2} = \Delta H_{N2-I2}(T_{N2-I2}) + [(c - a) \cdot (T - T_{N2-I2})]$$

$$+ \left[ \frac{(d - b) \cdot (T^2 - T_{N2-I2}^2)}{2} \right]$$

$$\Delta H_{I2-I} = \Delta H_{I2-I}(T_{I2-I}) + [(e - c) \cdot (T - T_{I2-I})]$$

$$+ \left[ \frac{(f - d) \cdot (T^2 - T_{I2-I}^2)}{2} \right]$$

$$\Delta H_{I-U} = \Delta H_{I-U}(T_{I-U}) + [(g - e) \cdot (T - T_{I-U})]$$

$$+ \left[ \frac{(h - f) \cdot (T^2 - T_{I-U}^2)}{2} \right] \quad (20)$$

where  $T_{N2-I2}$ ,  $T_{I2-I}$  and  $T_{I-U}$  are the midpoint temperatures of the three equilibria of Equation (13) respectively (corresponding to population values of 0.5 for the respective macrostates), whereas  $\Delta H_{N2-I2}(T_{N2-I2})$ ,  $\Delta H_{I2-I}(T_{I2-I})$  and  $\Delta H_{I-U}(T_{I-U})$  are the corresponding enthalpy values. In a similar way, the entropy functions can be calculated as

$$\Delta S_{N2-I2} = \Delta S_{N2-I2}(T_{N2-I2}) + \left[ (c - a) \cdot \ln\left(\frac{T}{T_{N2-I2}}\right) \right]$$

$$+ [(d - b) \cdot (T - T_{N2-I2})]$$

$$\Delta S_{I2-I} = \Delta S_{I2-I}(T_{I2-I}) + \left[ (e - c) \cdot \ln\left(\frac{T}{T_{I2-I}}\right) \right]$$

$$+ [(f - d) \cdot (T - T_{I2-I})]$$

$$\Delta S_{I-U} = \Delta S_{I-U}(T_{I-U}) + \left[ (g - e) \cdot \ln\left(\frac{T}{T_{I-U}}\right) \right]$$

$$+ [(h - f) \cdot (T - T_{I-U})] \quad (21)$$

where

$$\Delta S_{N2-I2}(T_{N2-I2}) = \frac{\Delta H_{N2-I2}(T_{N2-I2})}{T_{N2-I2}}$$

$$\Delta S_{I2-I}(T_{I2-I}) = \frac{\Delta H_{I2-I}(T_{I2-I})}{T_{I2-I}} + R \cdot \ln(\sqrt{P_{ref}}) \quad (22)$$

$$\Delta S_{I-U}(T_{I-U}) = \frac{\Delta H_{I-U}(T_{I-U})}{T_{I-U}}$$

The second term of the association-dissociation entropy function is responsible for the protein concentration dependencies shown in Figure 1, giving rise to a unique set of thermodynamic parameters compatible with all thermal transitions at different protein concentrations. Thus the parameters shown in Table 1 for this equilibrium correspond to a hypothetical experiment made at  $P_{ref}$  concentration. This value has been established arbitrarily as 100  $\mu$ M. Finally, the Gibbs energies can be easily calculated from the well known thermodynamic relations  $\Delta G = \Delta H - T \cdot \Delta S$  and  $\Delta G = -R \cdot T \cdot \ln K$ .

The temperature function of the enthalpy describing the thermal unfolding of WT-VP40 according to Equation (13) can be expressed as

$$H = H_{N2} + [\Delta H_{N2-I2} \cdot f_{I2}] + [(\Delta H_{N2-I2} + \Delta H_{I2-I}) \cdot f_I]$$

$$+ [(\Delta H_{N2-I2} + \Delta H_{I2-I} + \Delta H_{I-U}) \cdot f_U] \quad (23)$$

and the corresponding heat-capacity function, which will be fitted to the experimental traces, can easily be obtained by numerical methods, since

$$C_p = \left( \frac{\partial H}{\partial T} \right)_p \quad (24)$$

For the analysis of the VP40 traces shown in Figure 1 obtained at different protein concentrations we assume that a single set of thermodynamic magnitudes should be compatible with all the experimental traces, as it is considered by the model. Therefore, any differences in such thermodynamic magnitudes are consequence of the changes in equilibrium, solely induced by variations in protein concentration. Thus we undertook an overall analysis to obtain the heat-capacity  $C_{pN2}(T)$ ,  $C_{pI2}(T)$ ,  $C_{pI}(T)$  and  $C_{pU}(T)$  linear functions (Equation (19)) for each experiment independently, as well as the  $T_{N2-I2}$ ,  $T_{I2-I}$  and  $T_{I-U}$ , and  $\Delta H_{N2-I2}(T_{N2-I2})$ ,  $\Delta H_{I2-I}(T_{I2-I})$  and  $\Delta H_{I-U}(T_{I-U})$  parameters related to the three equilibria shown in Equation (1). All these 14 parameters were fitted individually for each experiment in an attempt to assess the realistic error associated to these values, compared in Table 1, which includes the fitting errors. Calculations were made using Sigma Plot 2000 (Systat Software Inc.).

## Disclosure statement

There are no conflicts of interest to declare

## Funding

This work was supported by the Spanish Ministry of Economy under grants [BIO2012-39922-CO2 and BIO2016-78746-C2-1-R], as well as by the Fondo Europeo de Desarrollo Regional (FEDER). Mass spectrometry measurements were performed at the Center of Scientific Instrumentation of the University of Granada.

## ORCID

Pedro Buzon  <http://orcid.org/0000-0002-1282-9006>

Jose C. Martinez  <http://orcid.org/0000-0003-2657-2456>

## References

- Azuaga, A. I., Dobson, C. M., Mateo, P. L., & Conejero-Lara, F. (2002). Unfolding and aggregation during the thermal denaturation of streptokinase. *European Journal of Biochemistry*, 269(16), 4121–4133. doi:10.1046/j.1432-1033.2002.03107.x
- Bornholdt, Z. A., Noda, T., Abelson, D. M., Halfmann, P., Wood, M. R., Kawaoka, Y., & Saphire, E. O. (2013). Structural rearrangement of Ebola virus VP40 begets multiple functions in the virus life cycle. *Cell*, 154(4), 763–774. doi:10.1016/j.cell.2013.07.015
- Cobos, E. S., Iglesias-Bexiga, M., Ruiz-Sanz, J., Mateo, P. L., Luque, I., & Martinez, J. C. (2009). Thermodynamic characterization of the folding equilibrium of the human Nedd4-WW4 domain: At the frontiers of cooperative folding. *Biochemistry*, 48(36), 8712–8720. doi:10.1021/bi9007758

- Fabian, H., Falber, K., Gast, K., Reinstadler, D., Rogov, V. V., Naumann, D., ... Filimonov, V. V. (1999). Secondary structure and oligomerization behavior of equilibrium unfolding intermediates of the lambda CRO repressor. *Biochemistry*, 38(17), 5633–5642. doi:10.1021/bi982120d
- Filimonov, V. V., Prieto, J., Martinez, J. C., Bruix, M., Mateo, P. L., & Serrano, L. (1993). Thermodynamic analysis of the chemotactic protein from *Escherichia coli*, CheY. *Biochemistry*, 32(47), 12906–12921. doi:10.1021/bi00210a045
- Filimonov, V. V., & Rogov, V. V. (1996). Reversible association of the equilibrium unfolding intermediate of lambda CRO repressor. *Journal of Molecular Biology*, 255(5), 767–777. doi:10.1006/jmbi.1996.0062
- Gc, J. B., Johnson, K. A., Husby, M. L., Frick, C. T., Gerstman, B. S., Stahelin, R. V., & Chapagain, P. P. (2016). Interdomain salt-bridges in the Ebola virus protein VP40 and their role in domain association and plasma membrane localization. *Protein Science*, 25(9), 1648–1658. doi:10.1002/pro.2969
- Ibarra-Molero, B., Naganathan, A. N., Sanchez-Ruiz, J. M., & Munoz, V. (2016). Modern analysis of protein folding by differential scanning calorimetry. *Methods Enzymology*, 567, 281–318.
- Murciano-Calles, J., Cobos, E. S., Mateo, P. L., Camara-Artigas, A., & Martinez, J. C. (2010). An oligomeric equilibrium intermediate as the precursory nucleus of globular and fibrillar supramacromolecular assemblies in a PDZ domain. *Biophysical Journal*, 99(1), 263–272. doi:10.1016/j.bpj.2010.04.003
- Privalov, P. L. (1979). Stability of proteins: Small globular proteins. *Advances in Protein Chemistry*, 33, 167–241.
- Privalov, P. L. (1982). Stability of proteins. Proteins which do not present a single cooperative system. *Advances in Protein Chemistry*, 35, 1–104.
- Radzimanowski, J., Effantin, G., & Weissenhorn, W. (2014). Conformational plasticity of the Ebola virus matrix protein. *Protein Science*, 23(11), 1519–1527. doi:10.1002/pro.2541
- Saranya, V., Shankar, R., & Vijayakumar, S. (2019). Structural exploration of viral matrix protein 40 interaction with the transition metal ions (Ag(+) and Cu(2+)). *Journal of Biomolecular Structure and Dynamics*, 37(11), 2875–2896. doi:10.1080/07391102.2018.1498803
- Stahelin, R. V. (2014). Could the Ebola virus matrix protein VP40 be a drug target? *Expert Opinion on Therapeutic Targets*, 18(2), 115–120. doi:10.1517/14728222.2014.863877
- Viguera, A. R., Martinez, J. C., Filimonov, V. V., Mateo, P. L., & Serrano, L. (1994). Thermodynamic and kinetic analysis of the SH3 domain of spectrin shows a two-state folding transition. *Biochemistry*, 33(8), 2142–2150. doi:10.1021/bi00174a022