Biophysical Chemistry 230 (2017) 36–44

Contents lists available at ScienceDirect

Biophysical Chemistry

journal homepage: www.elsevier.com/locate/biophyschem

Ultracompact states of native proteins

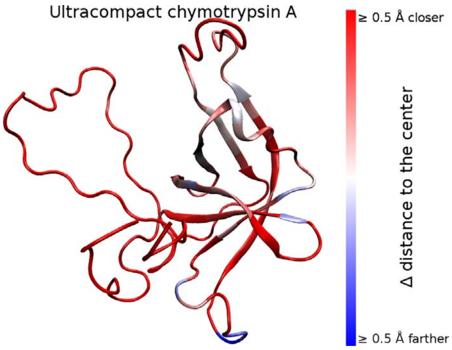
Leandro Grille Coronel[a], Juan P. Acierno[b], Mario R. Ermácora[a,b,*]

[a] Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, Roque Saenz Pena 352, B1876BXD Bernal, Pcia. de Buenos Aires, Argentina
[b] Instituto Multidisciplinario de Biología Celular, Conicet, Calle 526 y Camino Gral. Belgrano, B1906APO La Plata, Buenos Aires, Argentina

HIGHLIGHTS

- Inspection of circa 20000 X-ray structures confirmed cryogenic protein contraction.
- Ultracompact native states found by statistical analysis of radius of gyration
- Ultracompact states have shorter van der Waals contacts and hydrogen bonds.
- Ultracompact states have more van der Waals interactions.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Protein conformation
Cryocooling
X-ray crystallography
Radius of gyration
Protein glass transition
Protein energy landscape
Protein compaction
Protein intramolecular interactions

ABSTRACT

A statistical analysis of circa 20,000 X-ray structures evidenced the effects of temperature of data collection on protein intramolecular distances and degree of compaction. Identical chains with data collected at cryogenic ultralow temperatures (≤160 K) showed a radius of gyration (R_g) significantly smaller than at moderate temperatures (≥240 K). Furthermore, the analysis revealed the existence of structures with a R_g significantly smaller than expected for cryogenic temperatures. In these ultracompact cases, the unusually small R_g could not be specifically attributed to any experimental parameter or crystal features. Ultracompaction involves most atoms and results in their displacement toward the center of the molecule. Ultracompact structures on average have significantly shorter van der Waals and hydrogen bonds than expected for ultralow temperature structures. In addition, the number of van der Waals contacts was larger in ultracompact than in ultralow temperature structures. The structure of these ultracompact states was analyzed in detail and the implication and possible causes of the phenomenon are discussed.

1. Introduction

All the actual knowledge on the function of proteins is firmly grounded in the atomic description of the native state, which is generally defined in terms of atomic coordinates from X-ray diffraction data. Moreover, since the seminal contributions of Linus Pauling [1], protein folding theory and modeling rely on the energy of non covalent interactions estimates from inter atomic geometry and distances obtained from diffraction data.

A frequently overlooked aspect is the temperature dependence of diffraction data. Most of the structures in the RCSB Protein Data Bank (PDB; http://www.rcsb.org/; [2]) have been solved with data collection at cryogenic temperatures (≤200 K), after flash-freezing protein crystals. Although there is little doubt that this experimental condition permits a faithful representation of the native state at higher temperatures, the detailed effects of cryocooling on non covalent inter

[*] Corresponding author at: Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, Roque Saenz Pena 352, B1876BXD Bernal, Buenos Aires, Argentina
E-mail address: ermacora@unq.edu.ar (M.R. Ermácora).

http://dx.doi.org/10.1016/j.bpc.2017.08.004
Received 2 August 2017; Accepted 16 August 2017
Available online 18 August 2017

atomic distances need to be further investigated.

Several early studies assessed the effects of cryogenic temperatures on the atomic mobility in protein crystals. These studies confirmed that the atomic mean-square displacements are greatly reduced at temperatures below 200 K [3,4]. In addition, Hartman et al. reported in 1982 that the overall structure of metmyoglobin at 80 K was very similar to that at 300 K, but the volume of molecule was smaller [5]. According to these authors, the decrease in volume was accompanied by a shortening of non covalent intra molecular distances.

A few years latter, the volume contraction of metmyoglobin at cryogenic temperatures was further examined by Frauenfelder et al. [6]. It was found that the protein atoms were displaced toward the center of the molecule by 0.16 Å on average. Most importantly, most atoms were similarly involved in the displacement. The thermal compaction of metmyoglobin atoms was captured by the radius of gyration (R_g), which showed a contraction of 0.21 Å between 290 and 80 K. The authors also concluded that the compaction was not the result of an obliteration of the larger cavities within the molecule, but resulted from a general closer packing of all atoms.

In the early nineties, Earnest et al. [7] compared 120 and 295 K structures of rat trypsin, finding a reduction in the unit cell dimensions accompanied by significant decreases of 1.2% in molecular surface area and 0.2% in R_g at the lower temperature. Similarly, Tilton et al. [8] analyzed the structures of ribonuclease A at nine different temperatures ranging from 98 to 320 K, showing that the protein expands slightly (0.4% per 100 K) with increasing temperature and that this expansion was linear. Most inter atomic distances were involved in the change and this was evidenced by a significant change in R_g linearly related to temperature.

More recently, a survey of 15 crystal structures [9] showed that, on average and compared with room temperature, these proteins contracted by 1–2% at ∼100 K. The unit cell also contracted on average 4–5% in volume. Accordingly, the average Rg of the compacted proteins was 0.53% smaller than at room temperature. Cryocooling also increased the number of intramolecular van der Waals contacts. A subsequent analysis of 30 crystal structure cryo-room-temperature pairs essentially confirmed the above trend in structure parameters and, in addition, showed that the cryogenic structures have superior packing compared to the isomorphous high-resolution room-temperature structures [10].

The structural cryocooling effects raised the interest on the dynamic of protein. Several independent experimental techniques provided insights into a peculiar transition centered at 200–220 K. Above 200 K internal protein motions could not be modeled only as harmonic motions of individual atoms. Instead, collective motions of groups of atoms superimposed to simple vibrations had to be included in the models. This extra, high-order mobility above 200 K was invoked to explain the strong temperature dependence of the mean-square atomic displacements.

The characteristics of the broad transition between a temperature regime dominated by atom centered harmonic motions and another that included in addition correlated motions of groups of atoms were reminiscent of a liquid–glass transition, and it was termed the 'glass transition' in protein dynamics [11–17]. It has also been proposed that an additional protein transition takes place at about 110 K, correlated to a cryogenic phase transition of water from a high-density amorphous to a low-density amorphous state [18].

Binding and functional experiments across the glass transition temperature provided ground to the concept that conformational flexibility and adjustment are necessary for protein function [15]. Recent confluent methodological advances in X-ray crystallography, NMR and computer simulations are beginning to reveal the structural details of protein conformational dynamics at high resolution [19]. These advances make possible dynamic structural biology studies at atomic resolution, across many orders of magnitude of timescales, and at temperatures in the 100–300 K range, linking conformational variation to function.

The application of the above methodological advances enabled a recent study of the conformational dynamics of cyclophilin A from 100 K to room temperature [20]. The authors report that many alternative conformations in cyclophilin A are populated only at 240 K and above, and others remain populated or become populated at 180 K and below. These results suggest a conformational heterogeneity between 180 and 240 K, involving thermal deactivation and solvent-driven attenuation of protein motions in the crystal.

Although most of the crystallographic evidence for the existence of the glass transition was from the differential linear behavior of mean-square atomic displacements – which combines static and dynamic information – such transition can be captured by a purely static variable. Indeed, a biphasic behavior of R_g as a function of temperature was reported by Teeter et al.[14]: cambrin R_g remains constant from 100 to 160 K and increases linearly from 160 to 293 K.

Recently, we showed that the compactness of the native state may be enhanced by protein engineering and established a new lower limit to the compactness of the Class-A β-lactamase fold [21]. In this work, we reported a 1.7-Å resolution X-ray structure of *Bacillus licheniformis* exo-small penicillinase mutant in which phenylalanine replaces wild-type tryptophan residues. The structure revealed no qualitative conformational changes compared with thirteen previously reported structures of the same protein, but it had a significantly smaller R_g. The importance of this finding is twofold. First, it suggests that temperature may not be the only factor involved in unusual protein compaction. Second, it shows that the subject can be further investigated by statistical analysis of the PDB structures.

In this work, we undertook a statistical survey of the PDB looking for unusually compact forms of protein structures. We will show that protein thermal contraction at cryogenic temperature is a generalized phenomenon. Furthermore, our analysis will show the existence of structures with a degree of compaction well beyond that attributable to normal thermal effects. The impact of ultracompaction on non covalent inter atomic distances will be also established.

2. Materials and methods

2.1. The analyzed set of protein chains

An initial list of protein chain IDs in the PDB was downloaded from PISCES (pdbaanr; 2015; [22]). The list includes 65,195 chain classes. Each chain class includes several experimental realizations of the same sequence in one or more PDB entries. The total number of chains in the initial set was 249,185, from 95,503 PDB entries. For instance, the hemoglobin W37A chain class includes four chains: 1A01 B, 1A01 D, 1A0 W B, and 1A0 W D from two different PDB entries (1A01 and 1A0A). All four hemoglobin chains have identical sequence and their structures can be considered experimental replicates of the same chain in different contexts.

The initial set was cleaned as follows: (*i*) only chains structures with a resolution better than 3.0 Å were retained, (*ii*) chains with missing atom coordinates, geometrical inconsistencies or other experimental anomalies were discarded, and (*iii*) chain classes including < 20 members were eliminated.

The final working set consisted of 631 chain classes, each with 20 or more replicates of structures from the same sequence chain. Summing all the classes, the sample space contained 19,393 chains from 7114 PDB entries.

2.2. Calculation of R_g

R_g describes the shape and size of a molecule by computing the dispersion of the individual atoms about either the mass or the geometrical center. In this work, R_g was calculated about the geometrical center, considering only main chain heavy atoms and disregarding mass differences. The geometrical center vector ($\mathbf{r}_c$) is defined as

$$\sum_i^N (\mathbf{r}_i - \mathbf{r}_c)^2/N = 0,$$

where $\mathbf{r}_i$ is the vector pointing to atom i, and N is the number of considered atoms.

With the defined $\mathbf{r}_c$, the scalar R_g is calculated as

$$R_g = \sqrt{\sum_i^N (\mathbf{r}_i - \mathbf{r}_c)^2}$$

2.3. Normalized R_g values

R_g values were normalized to compare chains from different chain classes as follows.

$$Z_i = (R_{g,i} - \overline{R_{g,j}})/S_{Rg,j}$$

where $R_{g,i}$ is the R_g of chain i from chain class j, $\overline{R_{g,j}}$ is the average R_g in chain class j, and $S_{Rg,j}$ is the standard deviation of R_g in chain class j.

2.4. Statistical analysis and molecular visualization

Statistical analysis were performed using *R* [23]. When appropriate, mean differences were tested with the Welch Two Sample *t*-test. Homoscedasticity was assessed by the Levene test. The non parametric Mann Whitney *U* test was used to compare the differences in R_g between populations of protein structures solved at different temperatures. Molecular visualization and calculations of inter atomic distances were performed using VMD [24]. VMD settings for hydrogen bonds measurements were donor–acceptor cutoff distance of 4 Å and 40 degrees maximum departure from 180 for the donor-hydrogen-acceptor angle. VMD setting for measuring van der Waals contacts between carbon atoms was a cutoff distance of 4 Å and only C atoms from different residues were considered.

3. Results

3.1. Thermally induced contraction of protein structure

To assess the general effect of temperature of data collection on the compactness of protein chains, we compared structures from different chain classes (*i.e.*, chains of different sequence). To that end, a *Z* value was calculated. *Z* is a normalized R_g that represents, in standard deviation units, the difference between the R_g of a particular structure and the average R_g of the structures of the chain class to which it belongs (see Section 2.3). To ensure that the differences in R_g were not the result of refinement artifacts, lengths of backbone covalent bonds were calculated for each structure, and the consistency of the values with standard bond lengths was verified [25]. The results grouped by temperature range are shown in Table 1.

The analysis showed that structures with diffraction data collected at ultralow temperatures (below 160 K; typically at 100 K), hereafter ‘ultralow temperature structures', have R_g values significantly smaller than protein structures solved with data obtained at moderate temperature (240 to 310 K), hereafter ‘moderate temperature structures' (Fig. 1). The difference between the means of ultralow and moderate temperatures was 0.44 $\overline{Z}$ units. This difference was significant with the Wilcoxon Mann Whitney nonparametric test ($P < 2.2 \times 10^{-16}$), and by ANOVA ($F = 205.6$; $P < 2 \times 10^{-16}$) and Tukey post hoc HSD test ($P < 0.05$).

Table 1
Temperature (K) of data collection and average *Z* values. Data from structures whose temperature of data collection was not reported are in the ‘not available’ column (n.a.). *N* is the number of chains in each temperature range. $\overline{Z}$ are averages of R_g normalized as described in Section 2.3. (*) The $\overline{Z}$ difference between ultralow (≤160 K) and moderate temperature (≥240 K) is statistically significant (see Section 3.1).

	≤160	161–220	≥240	n.a.	Total
N	14,509	274	2318	2292	19,393
$\overline{Z}$	−0.08	−0.01	0.37*	0.12	0.00

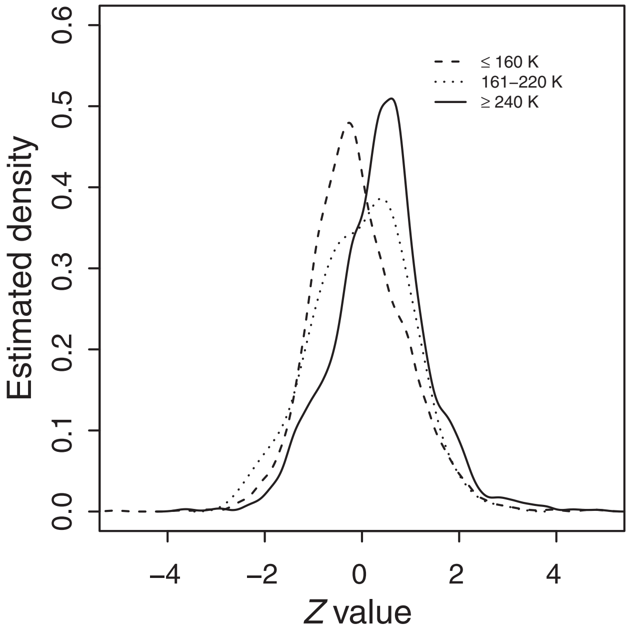


Fig. 1. The estimated density of *Z* values for different temperature ranges. *Z* measures in standard deviation units the distance of the R_g of a protein chain to the mean R_g of the chain class to which it belongs (see Section 2.3).

The structures solved at intermediate temperatures (161–220 K) evidenced a bimodal distribution, with maxima coincident with those of ultralow and moderate temperatures, respectively (Fig. 1).

3.2. Ultracompact states of native proteins

The results presented in the precedent section established that R_g calculation allows detection of very subtle compaction effects. To search for compaction effects larger than those caused by the thermally-induced dynamic glass transition described above, we examined the most deviating cases in the left tail of the ultralow temperature relative frequency curves in Fig. 1. We found a number of protein structures with $Z < -3$ that could be ultracompact states. However, the *Z* values in Fig. 1 and Table 1 were calculated for each class with a single standard deviation and a single mean for all temperatures, and this could have led to an underestimation of *Z* values. A more accurate calculation of *Z* values was performed using the means and standard deviations corresponding to each range of temperatures. One example of such recalculation is shown in Fig. 2 for the chain class trypsin. In this example, the density plot for the ultralow temperatures (≤160 K) exhibits structures with *Z* values well beyond the left tail of the curve.

Using the recalculated *Z* values for all classes, we identified 23 cases with $Z < -4$ (Table 2), corresponding to eight different protein folds.

The ultracompact structures listed in Table 2 have R_g values significantly smaller than the ultralow temperature structures of their respective classes. Thus, the thermal contraction at cryogenic temperature described in Section 3.1 cannot suffice to explain the ultracompaction. Trivial explanations for these observations were discarded by the quality control applied to all the structures analyzed. Namely, structures with geometrical inconsistencies or refinement artifacts were not included in this study. Moreover, since only backbone atoms of sequentially identical chains were used to calculate R_g values, side chain rotameric differences cannot explain the deviant results.

A preliminary inspection of the structures in Table 2 showed that

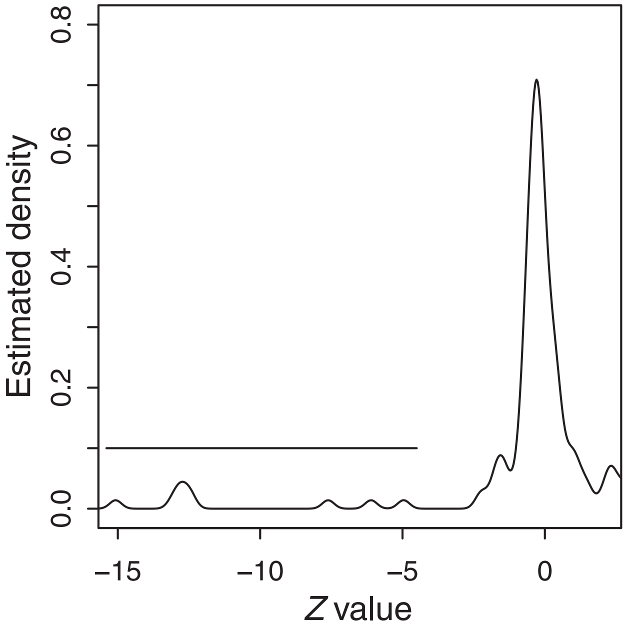


Fig. 2. The estimated density of Z values for trypsin class at ≤160 K. The horizontal line encompasses the ultracompact structures defined by Z values ≤ -4.

Table 2
Ultracompact structures. Z values were calculated differently than in Table 1: whereas in Table 1Z values were calculated with a single mean and standard deviation for all temperatures, here the mean and standard deviation of ultralow temperature was used (see Section 3.2). Although formally B2MG human 99 and B2MG human 100 pertain to different classes, the only difference between them is that in the latter the structure includes an initial methionine residue. Trypsin 1QL7 pertains to moderate temperature structures, however, it was included in the table because it exhibits a very high degree of compaction.

Class	Chain	T (K)	Z
B2MG human 99	3VFN B	100	−4.3
B2MG human 99	1T1Z B	100	−4.2
B2MG human 100	3UTT B	100	−5.5
B2MG human 100	3UTT G	100	−4.3
B2MG mouse	1K8D B	100	−5.0
Cyclin A2	2WMA B	100	−8.2
Cyclin A2	2WMA D	100	−5.6
Trypsin	1QL7 A	287	−9.4
Trypsin	1AQ7 A	93	−5.0
Trypsin	2A7H A	100	−6.1
Trypsin	2G81 E	100	−7.6
Trypsin	3GY2 A	100	−12.8
Trypsin	3GY3 A	100	−12.5
Trypsin	3GY5 A	100	−13.1
Trypsin	3GY6 A	100	−12.4
Trypsin	3GY8 A	100	−12.8
Trypsin	3RU4 T	100	−15.1
Chymotrypsin A	3RU4 D	100	−9.6
Lysozyme	3IJV A	100	−10.7
Fab heavy	2DWD A	100	−4.9
Fab light	2DWD B	100	−5.5
CF VII heavy	2FLR H	130	−6.4
HLA II DRα	2Q6 W A	98	−14.8

the conformational change involved in ultracompaction is peculiar, for it results from the movement of most backbone atoms toward the center of the molecule modifying neither the local nor the overall fold. In other words, the conformational change is not the result of rigid body reorganization of sub domains or hinge-like displacements.

Since in ultracompact structures the movement toward the center involved most of the backbone atoms, its global impact on non covalent interactions was examined. The mean hydrogen bond length was shorter in ultralow temperature structures (≤160 K) than in moderate (≥240 K) temperature structures (3.041 *vs* 3.051 Å, respectively; $P < 3.2 \times 10^{-18}$). The same comparison between ultralow temperature and ultracompact structures also yielded a significant difference (3.041 *vs* 2.971 Å, respectively; $P < 3.4 \times 10^{-8}$).

The mean van der Waals C–C bond length was not significantly different in ultralow compared with moderate temperature structures. However, the comparison between ultralow temperature and ultracompact structures yielded a significant difference (3.746 *vs* 3.737 Å, respectively; $P = 0.0002$).

To compare the number of noncovalent interactions at different data collection temperatures and degrees of compaction, data were normalized as Z-values, as in the case of R_g. The number of hydrogen bonds was not significantly different in moderate temperature, ultralow temperature, or ultracompact structures. However, the van der Waals C–C contacts were significantly more at ultralow temperature compared with moderate temperature ($P < 4.4 \times 10^{-105}$) and in ultracompact compared with in ultralow temperature structures ($P < 1.3 \times 10^{-12}$).

The conformational changes involved in ultracompaction will be described with more detail in the next sections for each of the ultracompact structures listed in Table 2.

3.3. Beta-2-microglobulin (B2MG)

B2MG is the light chain of the class I major histocompatibility complex (MHC-I). It is also found in serum as a stand alone domain. Associated to the heavy (alpha) chain of the MHC-I, it functions in the presentation of peptides to the T-cell receptor of CD8-bearing T lymphocytes and killer inhibitory receptors on natural killer cells. Circa five hundred structures in the PDB include one or more B2MG chains with over thirty different quaternary architectures (InterPro IPR015707; [26]). In this study, 582 B2MG chains in three different classes were analyzed. The two classes of human B2MG (UniProt [27] P61769,) listed in Table 2 only differ in the conservation of the initiating methionine residue, and although they were considered separately in the statistical analysis, for all practical purposes can be considered as a single class. The third class corresponds to mouse B2MG.

Human B2MG exhibits four ultracompact structures (PDB ID: 3VFN, 1T1Z, and two chains in 3UTT). 3VFN corresponds to the heterodimer formed with the alpha chain of HLA class I MHC B35 R151A mutant complexed with an Epstein Bar virus peptide [28]. The generalized movement of backbone atoms toward the center of the molecule for the B2MG chain in 3VFN can be appreciated in Fig. 3, compared with the average distance to the geometric center of the ultralow temperature structures of human B2MG class.

Another way to visualize the compaction effect along the main chain is with a ribbon representation and a color scale based on the distance-to-the-center difference between two structures. In the case of PDB ID: 3VFN ($Z = -4.3$) the difference with PDB ID: 3VFP ($Z = -0.3$) is shown in Panel A of Fig. 4.

Interestingly, PDB ID: 3VFN and 3VFP are two of a series of ten structures reported by the same group and with identical experimental crystallization and data collection parameters, refinement procedures and crystal properties (identical space group and unit cell geometry and dimensions) [28]. These structures only differ in a single amino acid substitution in the associated alpha chain of the complex or in the ligand peptide – as they were designed to characterize the interaction between the alpha chain and the bound antigen. Furthermore, the mutations and the binding site as a whole are very far away from the B2MG subunit. Thus, the fact that only one of the ten variants resulted in ultracompaction suggests that the phenomenon depends of very subtle differences in the experimental conditions in the preparation of the diffracting crystal.

The second ultracompact structure of human B2MG (PDB ID: 1T1Z) in Table 2 provides another interesting example of the subtleness of the ultracompaction phenomenon. It belongs to a subset of eight human B2MG structures reported by the same group and corresponding to a

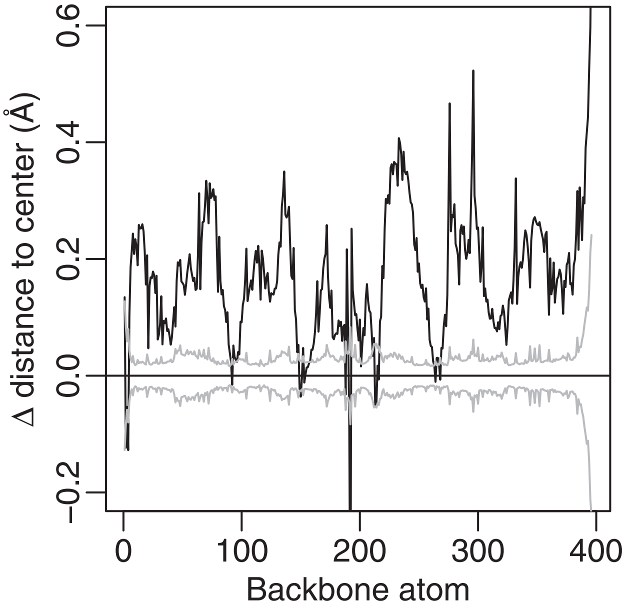


Fig. 3. Distance difference plot. The distance to the center of backbone heavy atoms was calculated for PDB ID: 3VFN (human B2MG; Table 2), and the difference in angstroms with the average distance of corresponding atoms of ultralow temperature structures for human B2MG class was plotted as a heavy line. Positive values indicate the closeness to the center. For comparison, the same procedure was applied individually to all ultralow temperature B2MG structures and the 2 ± *SEM* range was plotted in gray.

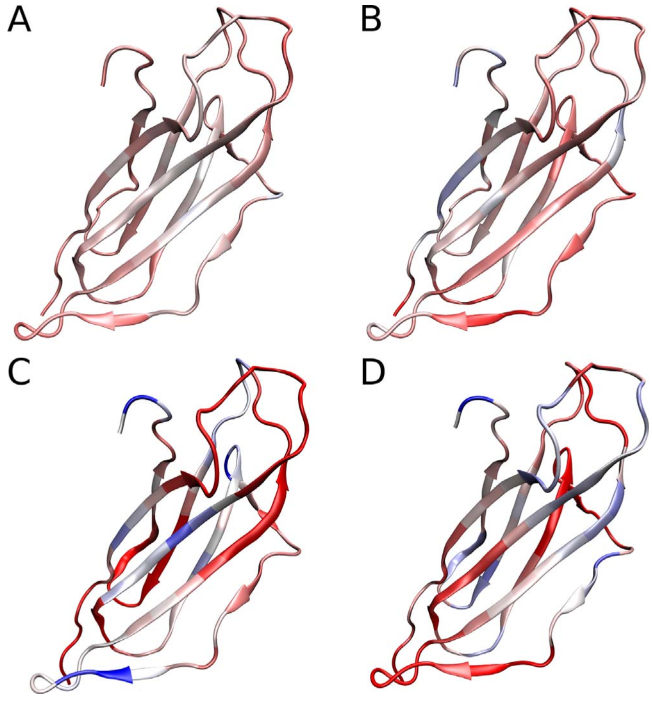


Fig. 4. Ribbon representation of ultracompact human B2MG (Table 2). The difference in angstroms in the distance to the center compared with ultralow temperature references is shown in a blue-white-read scale spanning −0.5–0.5 Å. Most of the ultracompact structure backbone atoms are located closer to the center (red) compared with the references, only a few short segments moved outward (blue) or are unchanged (white). *Panel A*, PDB ID: 3VFN *vs* 3VFP; *Panel B*, PDB ID: 1TZT *vs* 2FZ3; *Panels C, D*, PDB ID: 3UTT (chains B and G, respectively) *vs* 3VFP. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

HIV Gag-derived peptide variants complexed to HLA-A2 class I MHC [29]. All experimental parameters for the X-ray study of these peptide variants were identical, and the only relevant difference in the results was that the wild type peptide complex crystal was orthorhombic, whereas all the seven mutated peptide variants were monoclinic. Yet, the only ultracompact structure in this subset was PDB ID: 1T1Z, which corresponds to a peptide variant and a monoclinic crystal.

The plot of atom-by-atom distance-to-center difference for PDB ID: 1T1Z (not shown) revealed a generalized compaction along the sequence, similar to that shown in Fig. 3. The color-scale representation of the distance-to-the-center difference between PDB ID: 1T1Z ($Z = -4.2$) and PDB ID: 2FZ3 ($Z = 0.0$) is shown in Fig. 4, Panel B. The latter structure was chosen as a reference because is at the center of the ultralow temperature Z distribution for the corresponding B2MG class.

The third and fourth ultracompact structures of human B2MG in Table 2 are in a single unit cell from PDB ID: UTT. This entry corresponds to a complex of the T-cell receptor expressed by the beta-cell-specific cytotoxic CD8$^+$ T-cell clone 1E6, a peptide from the signal peptide of human insulin, and HLA-A2 class I MHC [30]. Its unit cell contains a dimer of the complex and, accordingly, two chains of 2BMG. Both chains exhibit most of its backbone atoms displaced to the center in the atom-by-atom distance to center plot (not shown). The distance color scale representations of the secondary structure elements are shown in Fig. 4, Panels C and D.

Comparing the color scale representations for the four structures in Fig. 4, it becomes clear that, despite some overall similarity, the pattern of compaction along the chain is not the same in all four structures. That is, the conformational change that leads to ultracompaction is not a transition between two well-defined states. On the contrary, the results suggest heterogeneous compaction modes, as if each particular structure were sampling a conformational ensemble of ultracompact structures.

Another observation worth of note is that among the stand-alone 2BMG there was none ultracompact (data not shown). However, this could simply be the consequence of the reduced number of uncomplexed structures compared with the large number of those in complex with HLA-A2 class I MHC.

The structure of ultracompact mouse B2MG PDB ID: 1K8D ($Z = -5.0$; Table 2; [31]) was analyzed using PDB ID: 3P4N ($Z = 0.0$) as the reference. Both, the distance-to-center difference plot and the color coded ribbon structure (not shown) were in agreement with the above described behavior of the human variants.

3.4. Cyclin

Cyclin controls both the G1/S and the G2/M phases of the cell cycle by forming complexes with the cyclin-dependent protein kinases CDK1 or CDK2. There are more than 90 PDB entries related to human cyclin chain (UniProt P20248). However, due to the rigorous quality control and selection criteria implemented in this work, the cyclin class examined for ultracompact states included only 21 chains. All of them were complexes with kinase CDK2 with or without additional ligands.

Two of the 21 cyclin chains analyzed are ultracompact forms (Table 2; PDB ID: 2WMA, chains B and D with $Z = -8.2$ and -5.6, respectively). Both were compared with PDB ID: 1QMZ chain B ($Z = 0.4$). This structure was chosen as reference because it has a Z value close to the center of the distribution, and its experimental parameters and results were very similar to those of the ultracompact structure.

The atom-by-atom distance to the center plot (not shown) and the color scale representation in Fig. 5 indicate the movement of most atoms in both cyclin chains in PDB ID: 2WMA toward the domain centers. However, there are a number of differences in the red-shifted color patterns between chains B and D. This again suggests that these structures are sampling an ensemble of similar conformations, rather than representing a change between a single, normally condensed state and a single ultracompact state. Interestingly, the two cyclin chains are ultracompact structure. Which suggests that ultracompaction affects the asymmetric unit as a whole.

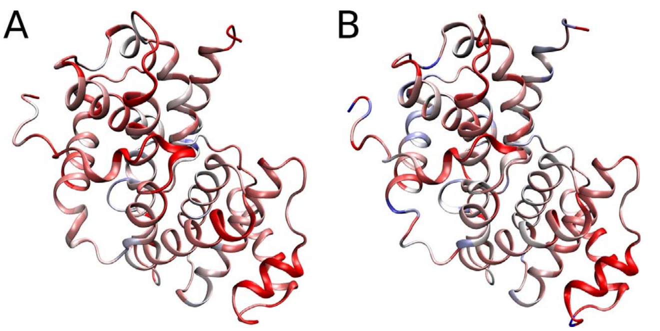


Fig. 5. A ribbon representation of ultracompact human cyclins. The distance to the center is color coded as in Fig. 4. *Panel A*, PDB ID: 2WMA chain B *vs* 1QMF; *Panel B*, PDB ID: 2WMA chain D *vs* 1QMF. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.5. Trypsin

The class bovine trypsin (UniProt P00760) included 261 chains, and among them ten ultracompact structures were found (Table 2). The first case of ultracompact structure in the class corresponded to PDB ID: 1QL7 ([32], $Z = -9.4$). It is the only case in Table 2 with diffraction data collection at moderate temperature. This structure was compared with 1GI6 ($Z = 0.4$), also with data collection at moderate temperature. 1GI6 was used as the reference because it had experimental details for data collection, crystallographic parameters, and resolution very similar to those of 1QL7.

1QL7 plot of the atom-by-atom difference distance to the center (not shown) corroborated that a large number of backbone atoms move toward the center as shown in the color-scale Fig. 6, Panel A. However, small regions exhibit blue color, indicating that they move away from the center. As context information, it can be added that trypsin ligands in 1QL7 are calcium, sulphate, and the protease inhibitor [4-(6-chloro-naphthalene)-piperazin-1-yl]-(3,4,5,6-tetrahydro-2H-[1,4′]bipyridinyl-4-yl)-methanone; whereas in the reference structure 1GI6, ligands are calcium, sulphate, and the protease inhibitor 2-(2-hydroxy-phenyl)-1H-indole-5-carboxamidine.

The color-coded distance difference between trypsins chains in PDB ID: 1AQ7 ([33], $Z = -5.0$) and 1P2 K ($Z = 0.0$) is shown in Fig. 6, Panel B. 1AQ7 structure is a complex between trypsin and the modified peptide inhibitor aeuroginosin 98-B; whereas 1P2 K describes a complex between trypsin and aprotinin, calcium, and sulphate. Both structures were solved at ultralow temperature, and the normally-compacted reference for this – and for some of the other trypsin examples below – was chosen because its $Z = 0$ value is at the center of the distribution for the compared structures. The plot of the atom-by-atom delta distance to the center (not shown) corroborated that a large number of backbone atoms in 1AQ7 are closer to the center than in the reference. Nevertheless, as shown in the color-scale figure, the few atoms that move away from the center tend to cluster in the N-terminal domain of the molecule.

The color-coded difference comparison of PDB ID: 2A7H and 2G81 ($Z = -6.1$ and $Z = -7.6$, respectively) with 1P2 K ($Z = 0.0$) is shown in Fig. 6, Panels C and D respectively. The coloring pattern is similar to that of 1QLZ and 1AQ7 (Fig. 6, Panels A and B), indicating that they are similarly compacted. 2AH7 structure ligands are calcium, and chloride; whereas 2G81 describes a complex between trypsin and Bowman-Birk inhibitor plus calcium, sulphate and acetate.

The color-coded differences for PDB ID: 3GY2, 3GY3, 3GY5, 3GY6, and 3GY8 contrasted with the reference 3GY4 are shown in Fig. 6, Panels E–I. The compaction pattern for these structures is very similar, with almost all atoms greatly displaced toward the center. Seven structures ‘3GY2–3GY8’ were solved by the same group and with similar experimental conditions and results [34]. Five are ultracompact structures, whereas 3GY7 and 3GY4 are normally compacted. No relation between ultracompaction and crystal properties was observed. Space groups for these ultracompact structures were $P3_121$ or $P2_12_12_1$; and the space group of the two non ultracompact structures was $P2_12_12_1$. All of them have an inhibitor, calcium, and sulphate as ligands. In 3GY2, 3GY5, 3GY6 and 3GY8 the inhibitor is 1,3-tris-(4′ amidinophenil)triazine; in 3GY4 and 3GY7 the inhibitors are *p*-amino benzamidine and benzamidine, respectively; and in 3GY3 the inhibitor is 1,5-bis(4-aminophenoxy)pentane.

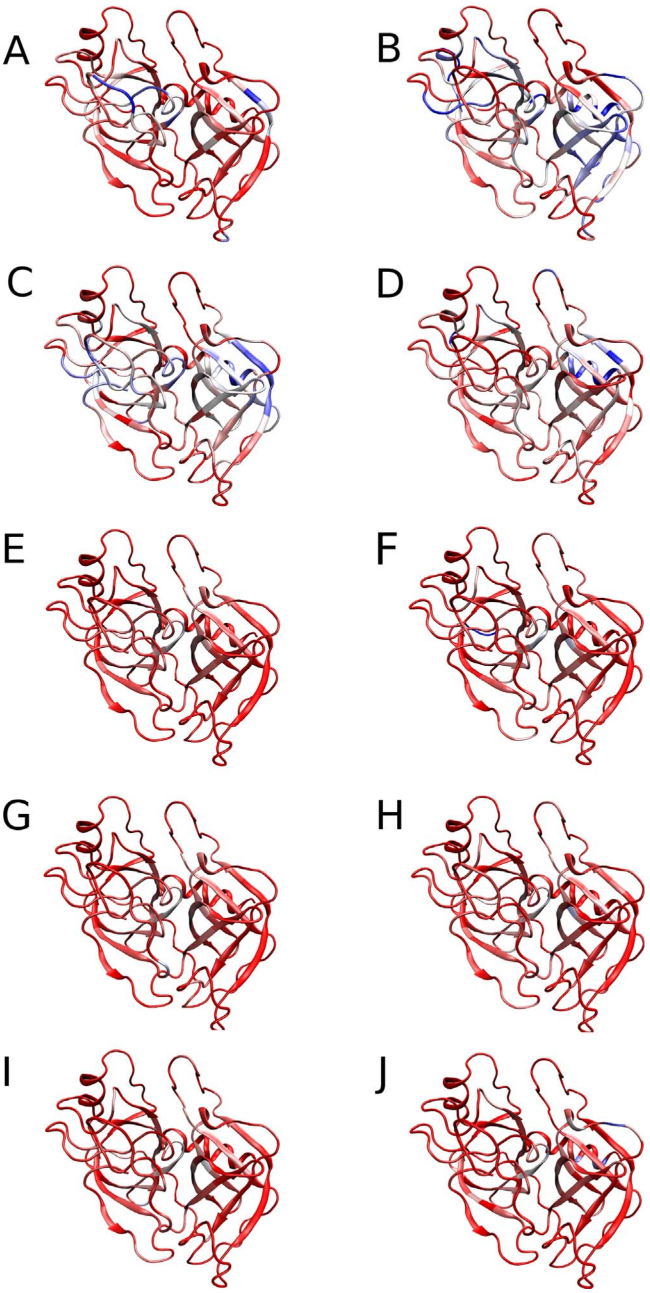


Fig. 6. A ribbon representation of ultracompact human trypsin chains. The difference with the reference in distance to the center is color coded as in Fig. 4. *Panel A*, PDB ID: 1QL7 *vs* 1GI6; *Panel B*, PDB ID: 1AQ7 *vs* 1P2K; *Panel C*, PDB ID: 2A7H *vs* 1P2K; *Panel D*, PDB ID: 2G81 *vs* 1P2K; *Panel E*, PDB ID: 3GY2 *vs* 3GY4; *Panel F*, PDB ID: 3GY3 *vs* 3GY4; *Panel G*, PDB ID: 3GY5 *vs* 3GY4; *Panel H*, PDB ID: 3GY6 *vs* 3GY4; *Panel I*, PDB ID: 3GY8 *vs* 3GY4; *Panel J*, PDB ID: 3RU4 *vs* 1P2K. The rationale for choosing the reference structures is explained in the text. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Finally, the difference in compaction between PDB ID: 3RU4 ($Z = -15.1$) and 1P2 K ($Z = -0.6$) is shown in Fig. 6, Panel J. 3RU4 describes a complex of Bowman-Birk inhibitor with both, bovine

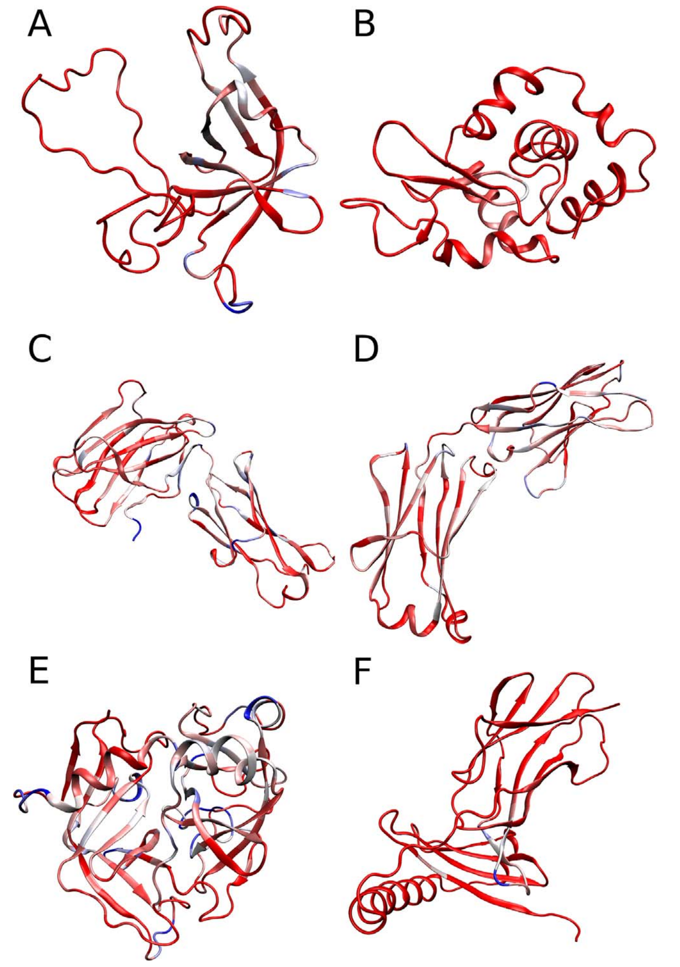


Fig. 7. A ribbon representation of different ultracompact structures. The difference with the reference in distance to the center is color coded as in Fig. 4. *Panel A*, chymotrypsin A, PDB ID: 3RU4 *vs* 1YPH; *Panel B*, lysozyme, PDB ID: 3IJV *vs* 3A8Z; *Panel C*, antibody Fab heavy chain, PDB ID: 2DWD *vs* 2HG5; *Panel D*, antibody Fab, light chain, PDB ID: 2DWD *vs* 1ZWI; *Panel E*, coagulation factor, PDB ID: 2FLR *vs* 2C4F; *Panel F*, HLA II DRα, PDB ID: 2QW6 chain A *vs* 4E41 chain A. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

trypsin and chymotrypsin A. Similarly to the above, the coloring pattern for this structure shows that almost all the atoms are closer to the center than in the reference. As in the ultracompact structure of cyclin described above, chymotrypsin A in this complex with trypsin is also an ultracompact structure (Table 2). Which again suggests that ultracompaction is an overall effect on the asymmetric unit as a whole, and not a specific feature of each chain.

3.6. Chymotrypsin A, lysozyme, Fab heavy chain, Fab light chain, coagulation factor VII heavy chain, and HLA II DRα

The curated class chymotrypsin A (UniProt P00766, chain B, residues 16–146), contains 24 chains. Of these, eight correspond to data collection at ultralow temperature. Among the latter, one ultracompact structure was found: PDB ID: 3RU4 (Table 2; $Z = -9.6$). 1YPH ($Z = -0.3$) was chosen as the reference for the class. The ultracompact structure is a triple complex with trypsin and Bowman-Birk inhibitor, for which the trypsin chain was found to be also ultracompact (see above). The difference in compaction between these two structures is shown in Fig. 7, Panel A. With the exception of very few residues, all the chain is either unchanged or displaced toward the center.

Chicken egg lysozyme (UniProt P00698) class comprises 221 chains with data collection at ultralow temperature. Among these, a single ultracompact structure was found, PDB ID: 3IJV, with a Z value of -10.7. The structure chosen as a reference was PDB ID: 3A8Z ($Z = 0.0$). The ultracompact structure and the reference have nearly identical experimental parameters and crystal data. As can be seen in Fig. 7, Panel B, all backbone atoms in 3IJV are strongly shifted toward the center. The distance plot corroborated this finding (not shown).

The mouse antibody Fab light and heavy chain classes included in this study (Table 2) are in 17 PDB entries of triple complexes with the membrane protein potassium channel KcsA [35]. All of them have nearly identical experimental parameters and crystal data. Only one of these entries, PDB ID: 2DWD, was found to contain ultracompact chains. Fab heavy and light chains in 2DWD ($Z = -4.9$ and -5.5, respectively) were compared with the references 2HG5 ($Z = 0.3$) and 1ZWI ($Z = 0.7$), respectively (Fig. 7 Panels C and D). As in all the above cases, most atoms in the 2DWD chains are closer to the center than in their respective references. Interestingly, the third chain in the complex, the potassium channel KcsA, was also ultracompact (not shown), reinforcing the notion that ultracompaction affects the unit cell as a whole.

Coagulation factor VII (CF VII; UniProt P08709) initiates the extrinsic pathway of blood coagulation. It is a serine protease that circulates in the blood as a zymogen. After processing, it becomes converted in heavy and light chains. CF VII heavy chain class collected for this study includes 37 chains of which 32 were solved with data collection at ultralow temperature. Among the latter, one ultracompact structure was found: PDB ID: 2FLR ([36], $Z = -6.4$; Table 2). The reference for this structure was 2C4F ($Z = 0.1$) and the color coded compaction is shown in Fig. 7, Panel E. Although it is not as predominant as in the previous examples, the red color of most of the backbone reveals the overall movement of the atoms toward the center. A similar conclusion can be drawn examining the distance-to-the center plot (not shown). In the entry 2FLR, a second chain in the unit cell, the CF VII light chain, is also ultracompact. However, this finding must be interpreted with caution because the light chain exhibits conformational variation in the analyzed set of structures (not shown).

HLA-DR – class II major histocompatibility complex (MHC-II) cell surface receptor – is a heterodimer comprising alpha and beta chains in complex with the corresponding peptide antigen. It participates in the antigen presentation to the T-cell receptor. HLA II DRα (UniProt P01903) class in this study comprises 12 chains in eight PDB entries. Among these, only one ultracompact state was found (PDB ID: 2Q6 W chain A ([37], $Z = -14.8$, Table 2)). As the reference state, we chose chain A in PDB ID: 4E41 with a Z value of -0.2. The overall compaction of HLA II DR alpha chain in 2Q6 W compared with the reference is illustrated in Fig. 7, Panel F. It is worthy of note that in the unit cell of both the ultracompact and its reference there are two HLA II DR alpha chains. Chain D in the ultracompact structure was rejected by the quality filter used to assemble the class because of four missing atoms in the backbone. Nevertheless, a very reliable estimation of the degree of compaction could still be obtained, and this evidenced that both HLA II DRα chains in 2Q6 W are ultracompact.

Unfortunately, the number of chains included in HLA II DRα class is rather small, and the experimental conditions and results are dissimilar for them. For instance, the ultracompact structure is the only one for which crystals were grown at pH 4.4. Moreover the beta chain in the asymmetric unit is not the same for all the crystals. This introduces some uncertainty in regards of the origin of the observed compaction for this structure.

4. Discussion

In this work, we report the statistical analysis of the degree of compaction of 19,393 protein chains. We found that R_g is significantly smaller at cryogenic than at moderate temperatures. The average behavior of Rg in the large set of analyzed structures confirmed the generality of the previously reported 'dynamic glass transition' at cryogenic temperatures [11,13-15]. The mean feature of this transition is the suppression of large correlated motions of atoms and the consequent predominance of simpler harmonic vibrations.

In addition, our survey identified a number of cases with a compaction much higher than the expected from the 'dynamic glass transition'. The existence of ultracompact cases suggests that proteins can contract beyond the limits normally observed at cryogenic temperatures.

We examined the ultracompact cases searching for experimental factors that could have caused the effect. Trivial explanations – like systematic errors in the resolution of the structures or primary structure errors – were discarded considering the quality control applied in the selection of the chains included in the survey.

Instrumental settings for X-ray data collection, crystal and unit cell parameters, crystal growth procedure, crystallization methods, refinement procedures, and resolution also failed to provide a convincing explanation for the observed effect. For the statistical analysis, the structures were grouped in classes of identical chains (*i.e.*, no mutations within a class were allowed), and only backbone heavy atoms were used to calculate *Rg*. Thus, chain chemical heterogeneity cannot explain the results. Remarkably, several examples were found of series of structures solved by the same laboratory with identical or nearly identical experimental conditions giving rise to ultracompaction only in one member of the series.

A class in the analyzed data set comprises chemically identical protein chains, monomeric or associated to other chains in a variety of domain architectures. In addition, these chains formed complexes with different ligands, ions, and additives. Thus, the context of each chain structure within a class was not the same. However, we found ultracompact structures within the same class with different context. Thus the context differences neither explain the observed ultracompaction.

Due to the sparsity of the data, we could not statistically test if ultracompaction is a property of all the different protein chains in a unit cell. However, in the few cases for which we were able to make the comparison, different chains in the same unit cell were found to be ultracompact. Thus, it is likely that ultracompaction, when present, affects the crystal as a whole.

The experimental factor that give rise to ultracompaction is elusive and could not be identified by the statistical survey. Since the crystal conditioning that allow X-ray data collection is very elaborated and prone to hysteresis, small variations in the procedure might cause the cryogenic effect to reach its maximum only in a small fraction of the solved structures [38]. In any case, the results provided by the statistical survey of the PDB make possible the design better controlled experiments aimed to identify the factors that leads to ultracompaction.

Despite the origin of ultracompaction, its realization rises a number of questions. One is whether ultracompaction is an energy stabilization correlate. Noncovalent bond lengths and number are widely used as main criteria to asses the energy of a protein conformation. In this regard, we found that ultracompact structures on average have significantly shorter van der Waals and hydrogen bond interactions. In addition, the number of van der Waals contacts was larger in ultracompact than in ultralow temperature structures. Thus, it is very likely that the ultracompact state is further low in enthalpy scale than ultralow temperature states. A compensating decrease in the entropy for the ultracompact state is also conceivable due to the restrictions to the movements in overpacked structures. The following question is then why cryocooling only exceptionally captures the ultracompaction effect. Certainly the answer to this question will require case-by-case molecular dynamic analyses and experiments aimed to reproduce reliably the ultracompaction effect.

Another related question is whether the different conformations sampled in the survey – moderate temperature, ultralow temperature, and ultracompact – are local minima at the bottom of the rugged energy landscape of protein folding. Again, this question cannot be solved solely from the results of the survey of X-ray structures. However, conformational states sampled in kinetic or non equilibrium experiments can be considered connected by a an hypothetical reversible process and endowed a priory of thermodynamic significance. Thus, the ultracompact structures identified in this work constitute valuable starting points for molecular dynamic simulations aimed to characterize the energetic of the local minima at the bottom of the folding landscape.

Are the ultracompact forms observed in this work populated to any extent at physiological temperatures or cryocooling just generates an entirely new free energy landscape? In this regard, we found only one case of ultracompact structure at moderate temperature (see Table 2). The significance of this single case remains to be established. However, it illustrates the need of controlled experiments aimed to repeatably reproduce the ultracompaction effect and identifying the experimental factors responsible for it. These experiments would expand the scope of current efforts to characterize the conformational heterogeneity of the native state [39].

Acknowledgments

This work was supported by grants from CONICET (PIP-GI 11220130100277CO) and the Universidad Nacional de Quilmes (PP 53/1002). Leandro Grille Coronel is the recipient of a Universidad Nacional de Quilmes fellowship.

References

[1] L. Pauling, R.B. Corey, H.R. Branson, The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain, Proc. Natl. Acad. Sci. U.S.A. 37 (4) (1951) 205–211.

[2] H. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. Bhat, H. Weissig, I. Shindyalov, P. Bourne, The protein data bank, Nucleic Acids Res. 28 (2000) 235–242.

[3] H. Frauenfelder, G.A. Petsko, D. Tsernoglou, Temperature-dependent X-ray diffraction as a probe of protein structural dynamics, Nature 280 (5723) (1979) 558–563.

[4] J. Walter, W. Steigemann, T. Singh, H. Bartunik, W. Bode, R. Huber, On the disordered activation domain in trypsinogen: chemical labelling and low-temperature crystallography, Acta Crystallogr., Sect. B 38 (5) (1982) 1462–1472.

[5] H. Hartmann, F. Parak, W. Steigemann, G. Petsko, D.R. Ponzi, H. Frauenfelder, Conformational substates in a protein: structure and dynamics of metmyoglobin at 80 K, Proc. Natl. Acad. Sci. U.S.A. 79 (16) (1982) 4967–4971.

[6] H. Frauenfelder, H. Hartmann, M. Karplus, I. Kuntz Jr, J. Kuriyan, F. Parak, G.A. Petsko, D. Ringe, R.F. Tilton Jr, Thermal expansion of a protein, Biochemistry 26 (1) (1987) 254–261.

[7] T. Earnest, E. Fauman, C.S. Craik, R. Stroud, 1.59 Å structure of trypsin at 120 K: comparison of low temperature and room temperature structures, Proteins 10 (3) (1991) 171–187.

[8] R.F. Tilton Jr, J.C. Dewan, G.A. Petsko, Effects of temperature on protein structure and dynamics: X-ray crystallographic studies of the protein ribonuclease-A at nine different temperatures from 98 to 320 K, Biochemistry 31 (9) (1992) 2469–2481.

[9] D.H. Juers, B.W. Matthews, Reversible lattice repacking illustrates the temperature dependence of macromolecular interactions, J. Mol. Biol. 311 (4) (2001) 851–862.

[10] J.S. Fraser, H. van den Bedem, A.J. Samelson, P.T. Lang, J.M. Holton, N. Echols, T. Alber, Accessing protein conformational ensembles using room-temperature X-ray crystallography, Proc. Natl. Acad. Sci. U.S.A. 108 (39) (2011) 16247–16252.

[11] I. Iben, D. Braunstein, W. Doster, H. Frauenfelder, M. Hong, J. Johnson, S. Luck, P. Ormos, A. Schulte, P. Steinbach, et al., Glassy behavior of a protein, Phys. Rev. Lett. 62 (16) (1989) 1916–1919.

[12] W. Doster, S. Cusack, W. Petry, Dynamical transition of myoglobin revealed by inelastic neutron scattering, Nature 337 (6209) (1989) 754–756.

[13] Y. Miyazaki, T. Matsuo, H. Suga, Low-temperature heat capacity and glassy behavior of lysozyme crystal, J. Phys. Chem. B 104 (33) (2000) 8044–8052.

[14] M. Teeter, A. Yamano, B. Stec, U. Mohanty, On the nature of a glassy state of matter in a hydrated protein: relation to protein function, Proc. Natl. Acad. Sci. U.S.A. 98 (20) (2001) 11242–11247.

[15] D. Ringe, G.A. Petsko, The 'glass transition' in protein dynamics: what it is, why it occurs, and how to exploit it, Biophys. Chem. 105 (2) (2003) 667–680.

[16] W. Doster, The protein-solvent glass transition, Biochim. Biophys. Acta 1804 (1) (2010) 3–14.

[17] J.R. Lewandowski, M.E. Halse, M. Blackledge, L. Emsley, Direct observation of hierarchical protein dynamics, Science 348 (6234) (2015) 578–581.

[18] C.U. Kim, M.W. Tate, S.M. Gruner, Protein dynamical transition at 110 K, Proc. Natl. Acad. Sci. U.S.A. 108 (52) (2011) 20897–20901.

[19] H. Van Den Bedem, J.S. Fraser, Integrative, dynamic structural biology at atomic resolution–it's about time, Nat. Methods 12 (4) (2015) 307–318.

[20] D.A. Keedy, L.R. Kenner, M. Warkentin, R.A. Woldeyes, J.B. Hopkins, M.C. Thompson, A.S. Brewster, A.H. Van Benschoten, E.L. Baxter,

M. Uervirojnangkoorn, et al., Mapping the conformational landscape of a dynamic enzyme by multitemperature and XFEL crystallography, eLife 4 (2015) e07574.

[21] V.A. Risso, J.P. Acierno, S. Capaldi, H.L. Monaco, M.R. Ermácora, X-ray evidence of a native state with increased compactness populated by tryptophan-less *B. licheniformis β*-lactamase, Protein Sci. 21 (7) (2012) 964–976.

[22] G. Wang, R.L. Dunbrack Jr, PISCES: recent improvements to a PDB sequence culling server, Nucleic Acids Res. 33 (2005) W94–8.

[23] R language and Environment 2012, ISBN 3-900051-07-0, http://www.R-project.org/.

[24] W. Humprey, A. Dalke, K. Schulten, VMD-visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38.

[25] R.A. Engh, R. Huber, Accurate bond and angle parameters for X-ray protein structure refinement, Acta Crystallogr. Sect A 47 (4) (1991) 392–400.

[26] R.D. Finn, T.K. Attwood, P.C. Babbitt, A. Bateman, P. Bork, A.J. Bridge, H.-Y. Chang, Z. Dosztányi, S. El-Gebali, M. Fraser, et al., InterPro in 2017-beyond protein family and domain annotations, Nucleic Acids Res. 45 (2016) D190–D199.

[27] UniProt Consortium, UniProt: the universal protein knowledgebase, Nucleic Acids Res. 45 (2017) D158–D169.

[28] Y.C. Liu, Z. Chen, S.R. Burrows, A.W. Purcell, J. McCluskey, J. Rossjohn, S. Gras, The energetic basis underpinning T-cell receptor recognition of a super-bulged peptide bound to a major histocompatibility complex class I molecule, J. Biol. Chem. 287 (15) (2012) 12267–12276.

[29] E. Martinez-Hackert, N. Anikeeva, S.A. Kalams, B.D. Walker, W.A. Hendrickson, Y. Sykulev, Structural basis for degenerate recognition of natural HIV peptide variants by cytotoxic lymphocytes, J. Biol. Chem. 281 (29) (2006) 20205–20212.

[30] A.M. Bulek, D.K. Cole, A. Skowera, G. Dolton, S. Gras, F. Madura, A. Fuller, J.J. Miles, E. Gostick, D.A. Price, et al., Structural basis for the killing of human beta cells by CD8+ T cells in type 1 diabetes, Nat. Immunol. 13 (3) (2012) 283–289.

[31] X.-l. He, P. Tabaczewski, J. Ho, I. Stroynowski, K.C. Garcia, Promiscuous antigen presentation by the nonclassical MHC Ib Qa-2 is enabled by a shallow, hydrophobic groove and self-stabilized peptide conformation, Structure 9 (12) (2001) 1213–1224.

[32] M.T. Stubbs, S. Reyda, F. Dullweber, M. Möller, G. Klebe, D. Dorsch, W.W. Mederski, H. Wurziger, pH-Dependent binding modes observed in trypsin crystals: lessons for structure-based drug design, ChemBioChem. 3 (2–3) (2002) 246–249.

[33] B. Sandler, M. Murakami, J. Clardy, Atomic structure of the trypsin-aeruginosin 98-B complex, J. Am. Chem. Soc. 120 (3) (1998) 595–596.

[34] C.S. Perilo, M.T. Pereira, M.M. Santoro, R.A.P. Nagem, Structural binding evidence of the trypanocidal drugs Berenil® and Pentacarinate® active principles to a serine protease model, Int. J. Biol. Macromol. 46 (5) (2010) 502–511.

[35] S. Yohannan, Y. Hu, Y. Zhou, Crystallographic study of the tetrabutylammonium block to the KcsA K+ channel, J. Mol. Biol. 366 (3) (2007) 806–814.

[36] J.R. Riggs, H. Hu, A. Kolesnikov, E.M. Leahy, K.E. Wesson, W.D. Shrader, D. Vijaykumar, T.A. Wahl, Z. Tong, P.A. Sprengeler, et al., Novel 5-azaindole factor VIIa inhibitors, Bioorg. Med. Chem. Lett. 16 (12) (2006) 3197–3200.

[37] C.S. Parry, J. Gorski, L.J. Stern, Crystallographic structure of the human leukocyte antigen DRA, DRB3* 0101: models of a directional alloimmune response and autoimmunity, J. Mol. Biol. 371 (2) (2007) 435–446.

[38] B. Halle, Biomolecular cryocrystallography: structural changes during flash-cooling, Proc. Natl. Acad. Sci. U.S.A. 101 (14) (2004) 4793–4798.

[39] D.A. Keedy, H. Van Den Bedem, D.A. Sivak, G.A. Petsko, D. Ringe, M.A. Wilson, J.S. Fraser, Crystal cryocooling distorts conformational heterogeneity in a model Michaelis complex of DHFR, Structure 22 (6) (2014) 899–910.