Journal of Molecular Graphics and Modelling 28 (2010) 755–765

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

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

Molecular basis of the structural stability of a Top7-based scaffold at extreme pH and temperature conditions

Thereza A. Soares [a,b,*], Curt B. Boschek [a], David Apiyo [c], Cheryl Baird [a], T.P. Straatsma [a,**]

[a] Pacific Northwest National Laboratory, 902 Battelle Blvd., P.O. Box 999, MSIN K7-90, Richland, WA 99352, United States
[b] Department of Fundamental Chemistry, Federal University of Pernambuco, Av. Prof. Luiz Freire S/N, Cidade Universitária 50740-540, Recife, PE, Brazil
[c] Beckman Coulter Inc., 1000 Lake Hazeltine Dr. M/S R110B, Chaska, MN 55318, United States

ARTICLE INFO

Article history:
Received 15 October 2009
Received in revised form 29 December 2009
Accepted 31 January 2010
Available online 6 February 2010

Keywords:
Protein engineering
CD4 binding protein
Experimental interpretation
Chemical and thermal denaturation
Affinity reagents
Thermostable proteins
Temperature and pH-dependent conformations

ABSTRACT

The development of stable biomolecular scaffolds that can tolerate environmental extremes has considerable potential for industrial and defense-related applications. However, most natural proteins are not sufficiently stable to withstand non-physiological conditions. We have recently engineered the *de novo* designed Top7 protein to specifically recognize the glycoprotein CD4 by insertion of an eight-residue loop. The engineered variant exhibited remarkable stability under chemical and thermal denaturation conditions. In the present study, far-UV CD spectroscopy and explicit-solvent MD simulations are used to investigate the structural stability of Top7 and the engineered variant under extreme conditions of temperature and pH. Circular dichroism measurements suggest that the engineered variant $Top7_{CB1}$, like Top7, retains its structure at high temperatures. Changes in CD spectra suggest that there are minor structural rearrangements between neutral and acidic environments for both proteins but that these do not make the proteins less stable at high temperatures. The anti-parallel β-sheet is well conserved within the timescale simulated whereas there is a decrease of helical content when low pH and high-temperature conditions are combined. Concerted alanine mutations along the α-helices of the engineered Top7 variant did not revert this trend when at pH 2 and 400 K. The structural resilience of the anti-parallel β-sheet suggests that the protein scaffold can accommodate varying sequences. The robustness of the Top7 scaffold under extreme conditions of pH and temperature and its amenability to production in inexpensive bacterial expression systems reveal great potential for novel biotechnological applications.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

Antibodies have been the gold standard of binding proteins with desired specificities for more than one century [59,64]. Their remarkable success as binding proteins derives from the fact that they can be promptly generated against a broad range of target molecules with exceptional specificities, often in the low nanomolar to picomolar range [9,52]. However, antibodies are large molecules (four polypeptide chains) of complicated composition (glycosylated heavy chains), whose stability relies on the presence of disulfide bonds. As result, they tend to have relatively low expression levels, limited shelf life, and are not amenable to production in inexpensive expression systems such as bacteria. Alternatively, advances in combinatorial engineering methods [52,65,74] have led to the development of the so-called engineered protein scaffolds [73], *i.e.* small, non-immunoglobulin proteins that serves as a structural framework where new binding sites with tailored functionalities can be integrated [9,26,35,58].

Optimal protein scaffolds should exhibit several features [8,9,57]: thermodynamic and chemical stability, single polypeptide chain of small size, high bacterial expression for affordable production, functionality in the absence of disulphide bonds, availability of surface exposed loop regions tolerant to substitutions and of sufficient surface area to allow binding to target molecules with both high affinity and specificity. Several scaffold proteins such as affibodies [25,31,36], ankyrins [43], adnectins [44], avimers [72] and anticalins [58] have been successfully generated for medical applications, either in therapeutics or for *in vivo* diagnostics [26,57]. Indeed, adnectins based on the FN3 domain [30] and avimers based on multimers of human A domains targeted to IL-6 [71] are currently undergoing clinical trials [69]. Nevertheless, most natural proteins are not sufficiently stable to withstand non-physiological conditions such as extremes of pH

* Corresponding author at: Department of Fundamental Chemistry, Federal University of Pernambuco, Av. Prof. Luiz Freire S/N, Cidade Universitária 50740-540, Recife, PE, Brazil. Tel.: +55 81 3031 8041.
** Corresponding author.
E-mail addresses: thereza.soares@ufpe.br (T.A. Soares), tps@pnl.gov (T.P. Straatsma).

1093-3263/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2010.01.013

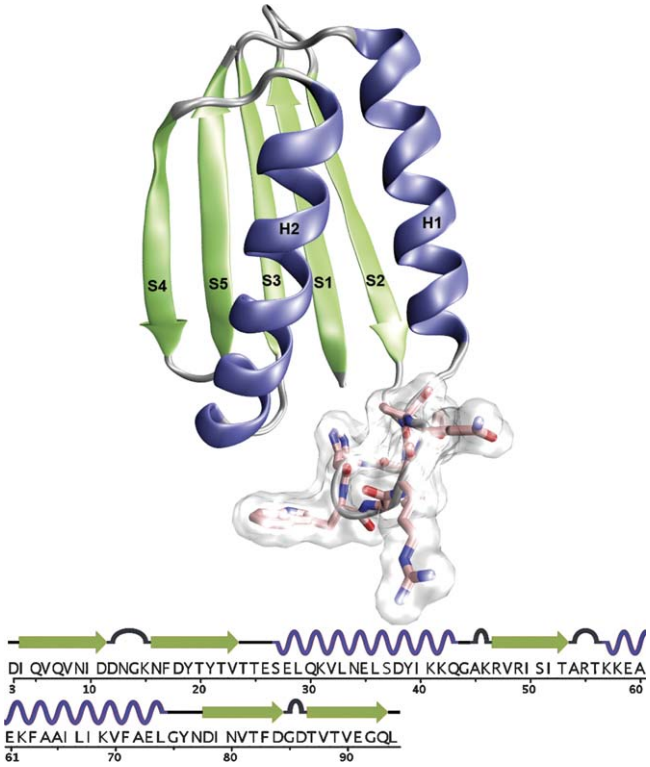

Fig. 1. Cartoon representation of the X-ray structure of the native Top7 (1QYS). The secondary structure elements are colored as function of the experimental B-factor values. Residues T25 and E26 indicate the site of insertion of the eight-residue loop in the engineered Top7 structures. Secondary structure definition according to Kabsch and Sander [40]. The illustration was prepared with the VMD software version 1.8 [38].

and temperature. This limitation has precluded the development of protein-based affinity reagents that can tolerate environmental extremes and extend the utility of these agents to applications outside of the highly controlled laboratory setting [53].

An alternative approach is the design of non-natural protein scaffolds stable under adverse environments [11]. A promising candidate for such scaffolds is the protein Top7, a *de novo* designed protein with a novel sequence and fold (Fig. 1) [48]. Top7 is a small α/β protein of 93 residues which exhibits remarkable stability under chemical (6 M GuHCL) and thermal denaturation (368.15 K) conditions [48,68]. It also exhibits significant mechanical stability and contains features shown by natural elastomeric proteins [48,68]. The chemical and thermal properties of Top7 has been linked to significantly less cooperative folding transitions than is observed for naturally occurring proteins of similar size [86]. It has been further shown that the three-dimensional structures of non-homologous amino acid sequences designed to adopt the Top7 fold exhibit high thermodynamic stability [19,77]. We have previously engineered a highly stable variant of Top7 capable to recognize the glycoprotein CD4 [11]. The novel functionality has been implemented by the insertion of an 8-residue sequence derived from the anti-CD4 13B8.2 paratope [2]. Chemical and thermal unfolding measurements showed that the engineered protein maintains stability at neutral pH, with no significant change in the free energy of unfolding when compared to native Top7. Of most relevance, ELISA experiments revealed that the engineered Top7 binds to CD4 whereas the native protein does not [11]. These findings demonstrated that Top7 displays considerable potential as an affinity reagent scaffold where binding sites for targets of interest can be grafted.

The identification of the structural principles underlying the stability of protein-based scaffolds is a crucial step for the rational design of novel functionalities. However, the elucidation of protein structure and dynamics through purely experimental approaches can be hindered by several limitations [6]. (i) The physical–chemical environment common for protein structural determination is fairly restricted, *i.e.* often crystalline phase at low temperature and neutral pH. (ii) Protein aggregation and insolubility can often curtail structural studies in solution by NMR spectroscopy [3,23,27,54]. Indeed, $Top7_{CB1}$ at low pH does aggregate at high concentrations making NMR measurements impracticable. (iii) The interpretation of such experiments is further complicated due to space/time ensemble averaging effects [81,82,84]. Molecular dynamics (MD) simulations are widely used to obtain information on the time-evolution of conformations at the atomic level and under varied environmental conditions [4,45,75,80]. Due to the unique property of probing space and time scales simultaneously, MD simulations offer a description of the structural changes, chemical pathways, and the thermodynamic character of cellular and physiological reactions that cannot be fully probed at the atomic level by experimentation alone [1,81,84]. It can also be used to probe the atomic details of conformational changes at different pH by representing the most predominant protonation states of ionizable groups at a given pH [10,32,39,45,50,61,75,76,79].

In the present study, a combined experimental-computational approach was used to examine the effect of low pH, high temperature and concerted mutations on the stability and structural dynamics of the Top7-based scaffold [11]. Far-UV CD spectroscopy was used to evaluate the chemo and thermostability of the Top7 scaffold. Explicit-solvent MD simulations were performed to probe the atomic details of conformational changes of the engineered Top7 scaffold at extreme pH and temperature conditions. Our results support the occurrence of a well-defined fold for the engineered variant of Top7 in different conditions of pH and temperature. It further suggests that the conformational ensembles corresponding to the native and engineered variant of Top7 are similar under equivalent environmental conditions.

2. Materials and methods

2.1. Circular dichroism (CD) spectroscopy

The Top7 K39E K40E K55E (referred to here as Top7) and $Top7_{CB1}$ proteins were expressed and purified as previously described [11]. Proteins were buffer exchanged (HiTrap Desalting column, GE Healthcare) into phosphate buffered saline (PBS, 100 mM NaH_2PO_4 and150 mM NaCl) at pH 7.0 or pH 2.0. All measurements were made at protein concentrations around 10 μM. Protein concentrations were determined using the theoretical extinction coefficients (ε_{280} = 10,810 M^{-1} cm^{-1} for Top7 and 16,960 M^{-1} cm^{-1} for $Top7_{CB1}$). Spectra were acquired with an Aviv Model 410 circular dichroism spectrometer (Aviv Biomedical, Inc., Lakewood, NJ, USA) using a 0.1 cm cuvette, averaging time of 1 s and a bandwidth of 1 nm. Spectra were acquired in triplicate, averaged, buffer spectrum subtracted, and smoothed. Thermal scans were conducted with temperature increases at 1 °C/min with a 5 min incubation time at each set point before a 1 s measurement at 222 nm.

2.2. Molecular dynamics simulations

Nine atomistic simulations in explicit solvent were performed for three sequences of Top7 at temperatures of 298.15 and 400.15 K and at pH 2 and 7 (Table 1). These sequences corresponded to the wild-type protein (Top7), the loop insert mutant ($Top7_{CB1}$) and the loop insert mutant with six additional mutations ($Top7_{CB1a}$), namely E43A, D46A, E67A, E69A, EA81A.

Table 1
Simulated systems.

System	Sequence[a]	Average temperature [K]	pH	Simulation length [ns]	Counter-ions	Total number of solute atoms	Total number of atoms
Run1	$Top7_{wt}$	298.12	7	28	3 Na^+	1480	51,430
Run2	$Top7_{CB1}$	298.10	7	26	2 Na^+	1615	42,187
Run3	$Top7_{wt}$	399.67	7	28	3 Na^+	1480	51,430
Run4	$Top7_{CB1}$	399.98	7	26	2 Na^+	1615	42,187
Run5	$Top7_{wt}$	298.14	2	23	11 Cl^-	1502	51,386
Run6	$Top7_{CB1}$	298.15	2	27	13 Cl^-	1641	42,117
Run7	$Top7_{wt}$	399.95	2	28	11 Cl^-	1502	51,386
Run8	$Top7_{CB1}$	399.94	2	31	13 Cl^-	1641	42,117
Run9	$Top7_{CB1a}$	399.98	2	28	13 Cl^-	1614	42,084

[a] $Top7_{wt}$: wild type; $Top7_{CB1}$: insertion of residues FGHWVRQS between residues 25 and 26 of wild-type sequence; $Top7_{CB1a}$: mutation of residues Q43A, N46A, Q67A, Q69A, QA81A from $Top7_{CB1}$ sequence.

Initial coordinates were taken from the crystallographic structure of PDB code 1QYS determined at 2.5 Å of resolution [48]. The mutant structures were generated by insertion of an eight-residue loop between residues Thr25 and Glu26 and/or deletion of the side chains of glutamic/aspartic acid to create alanine residues. The loop sequence FGHWVRQS was derived from the CDR1-H region of the CD4 binding pocket on the anti-CD4 13B 8.2 mAb [2]. The insert was modeled into the wild-type crystallographic structure by elongation the region between Thr25 and Glu26 and was subjected to initial relaxation via MD simulations. Except for the loop region, backbone atoms were fixed while the temperature was raised from 0 to 400 K in 20 ps increments of 50 K and maintained at 400 K for a 2-ns simulation period. The adequacy of the modeled structure in representing the ensemble of possible conformations for the loop insert was validated through comparisons against loop conformers predicted by the colony energy approach [88]. This approach has been shown to successfully predict conformations of 5-, 6-, 7-, and 8-residue long loops in a dataset of 135 non-redundant proteins to average RMSD values of 0.085, 0.092, 0.123 and 0.145 nm, respectively [88]. In this work, the number of initial random conformations was set to 10,000 and the 10 best-scored structures were selected. The root-mean-square deviations between the backbone atoms of the modeled structure and the *ab initio* predicted conformers were below 0.025 nm. The protonation states of amino acids were determined based on their corresponding pK_a with the program propKa [51]. The predicted pK_a values are presented as Supplementary material.

The structures were solvated in a cubic box with edge dimensions of 8.0 nm. Water molecules within 0.28 nm of any atom in the solute were removed. Periodic boundary conditions and the SPC/E water model [83] were used to describe the solvent molecular interactions. The AMBER95 force field has been used to describe the solute molecular interactions [12]. MD simulations were carried out for the NPT ensemble with a time step of 1 fs during the equilibration and 2 fs during the production runs. The temperature of solute and solvent were separately coupled to a Berendsen thermostat [7] with a relaxation time of 0.1 ps. The pressure was maintained at 1.025×10^5 Pa by means of isotropic coordinate scaling with a relaxation time $\tau = 0.4$ ps. A time step of 2 fs was used to integrate the equations of motion based on the leapfrog algorithm [33]. The bond lengths between hydrogen and heavy atoms were constrained by using the SHAKE algorithm [66] with a tolerance of 10^{-3} nm. A short-range cutoff of 1.0 nm was used for all non-bonded interactions, and long-range electrostatic interactions were treated by the smooth PME method [20]. The equilibration procedure consisted of thermalization of the solvent, with the solute atoms fixed, for 20 ps at 298.15 K, followed by minimization of all solute atoms, keeping the solvent coordinates fixed, and then simulation of the complete system by raising the temperature either from 0 to 298.15 K or from 0 to 400 K in 20 ps increments of 50 K each of MD simulation. Data production was carried out for 20–30 ns and configurations of the trajectory were recorded every 0.4 ps. All simulations were performed with the NWChem program [78] on a Hewlett Packard high-performance supercomputer at the Pacific Northwest National Laboratory.

3. Results and discussion

The variant sequence of $Top7_{CB1}$ contains an insertion of an eight-residue loop between residues Thr25 and Glu26 (Fig. 1) [11]. The loop sequence FGHWVRQS was derived from the CDR1-H region of the CD4 binding pocket on the anti-CD4 13B 8.2 mAb [2]. The variant sequence of $Top7_{CB1a}$ is derived from the $Top7_{CB1}$ sequence plus mutations of residues Q43A, N46A, Q67A, Q69A, QA81A. These concerted mutations along the α-helices have been investigated in an attempt to increase the stability of these secondary structures under low pH conditions through the substitution of low pK_a residues by the helix stabilizer alanine amino acid.

Far-UV CD spectra of proteins are sensitive to changes in protein secondary structure and negative contributions at around 222 nm correspond largely to α-helical and, less significantly, β-sheet content. The CD spectrum of $Top7_{CB1}$ is similar to that of Top7 under identical conditions (Fig. 2) suggesting similar secondary structure content. The roughly 15% smaller amplitude of the mean residue ellipticity at 222 nm of $Top7_{CB1}$ compared to Top7 at pH 7.0 and 298 K is largely explained by the addition of the non-α/β structured $_{CB1}$ sequence. Both Top7 and $Top7_{CB1}$ lose amplitude at 222 nm when temperature increases from 298.15 to 368.15 K (25 to 95 °C) at both neutral and acidic pH but spectral shape is largely retained. Both Top7 and $Top7_{CB1}$ undergo a minor decrease in amplitude at 222 nm with the change from neutral to acidic pH but there is a pronounced change in spectral shape suggesting that some structural rearrangements occurs. Thermal unfolding generally occurs in a cooperative fashion with clearly defined folded and unfolded states separated by a steep unfolding transition. The loss in amplitude of CD at 222 nm of both Top7 and $Top7_{CB1}$ at neutral or acidic pH has a roughly linear correlation with increasing temperature from 298.15 to 368.15 K (Fig. 3) suggesting that the transition to an unfolded state would require higher temperatures. $Top7_{CB1}$ in the presence of a strong chaotropic agent (guanidine HCl) is denatured by 368.15 K and we have included this data (Fig. 3) for comparison.

Structural properties derived from the MD simulations (Table 1) were monitored and are presented as time-series (Figs. 4–9) and averages (Table 2). Atom-positional root-mean-square-deviation (RMSD) and fluctuation (RMSF) of backbone atoms were calculated for the MD simulations with respect to the X-ray structure 1QYS (Figs. 4 and 5). RMSD and RMSF values are comparable for native and engineered Top7 under the corresponding pH and temperature conditions (Table 2). The highest average RMS deviation is observed for the $Top7_{CB1a}$ sequence under combined conditions of pH 2 and 400.15 K (Fig. 4). On average, RMSD values of the high-temperature simulations of $Top7_{CB1}$ and $Top7_{CB1a}$ diverge from the

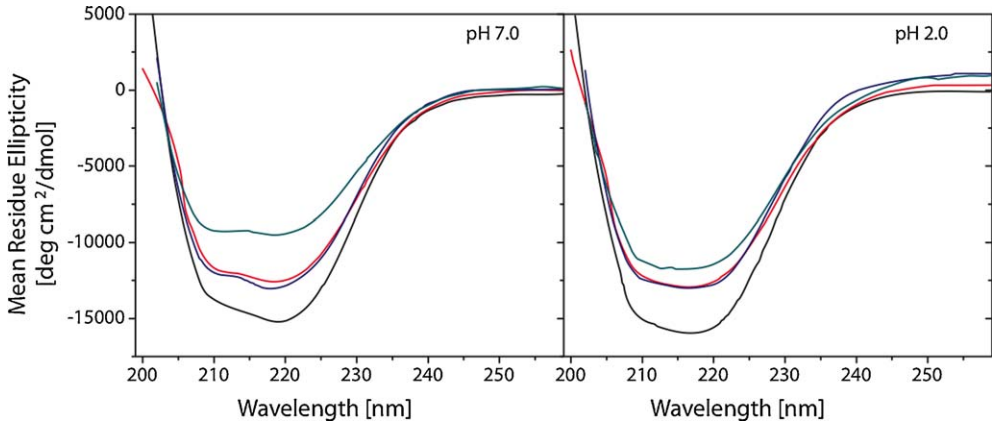

Fig. 2. CD spectra of Top7 and Top7$_{CB1}$ under extreme conditions. CD spectra for Top7 (black and red) and Top7$_{CB1}$ (blue and green) were acquired at 298.15 K (black and blue) and 368.15 K (red and green) at pH 7.0 (left panel) and pH 2.0 (right panel). Measurements were made in PBS at protein concentrations around 10 μM in a 0.1 cm cuvette. Lines represent the smoothed average of three spectra from which a buffer background has been subtracted.

native Top7 simulations at 400 K by 0.03–0.05 nm (pH2 and pH7) and 0.08 nm (pH2), respectively (Table 2). RMSF profiles show good agreement among the different MD simulations with the lowest RMSF values assigned to secondary structures and the highest ones to loop regions. The loop insert is the region of highest mobility in the engineered mutants, except for the Top7$_{CB1}$ simulation at 298.15 K and pH2. In this system the loop residues form four persistent hydrogen bonds that stabilize a particular loop conformation and lead to low atomic fluctuations (Supplementary material). Although some of these hydrogen bonds are also present in the other Top7$_{CB1}$ simulations, they exhibit shorter lifetimes, particularly at high temperature.

There is also a systematic increase in the radius of gyration (R_g) for all systems relative to initial X-ray structure (Table 2) with Top7 in solution adopting a less compact conformation than in the crystalline environment. The increase of the R_g of proteins in liquid phase compared to the corresponding crystalline phase has been widely reported in simulations of biomolecules [5,6,21,24,87,89].

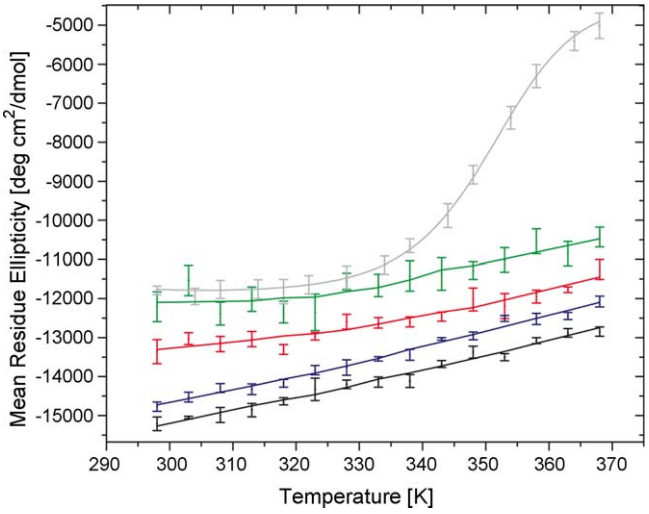

Fig. 3. Thermal denaturation of Top7 and Top7$_{CB1}$. CD at 222 nm was measured for Top7 (black and blue) and Top7$_{CB1}$ (red and green) in PBS at pH 7.0 (black and red) and pH 2.0 (blue and green). Thermal scans were conducted with temperature increases of 1 K/min with a 5 min incubation time at each set point and an averaging time of 1 s. Measurements were made in PBS at protein concentrations around 10 μM in a 0.1 cm cuvette. Error bars represent the deviation over the averaging time centered on each corresponding measurement of one representative measurement set. Lines are smoothed to aid in visualization. Top7$_{CB1}$ at pH 8 in the presence of 4 M guanidine (grey) is shown for comparison [11].

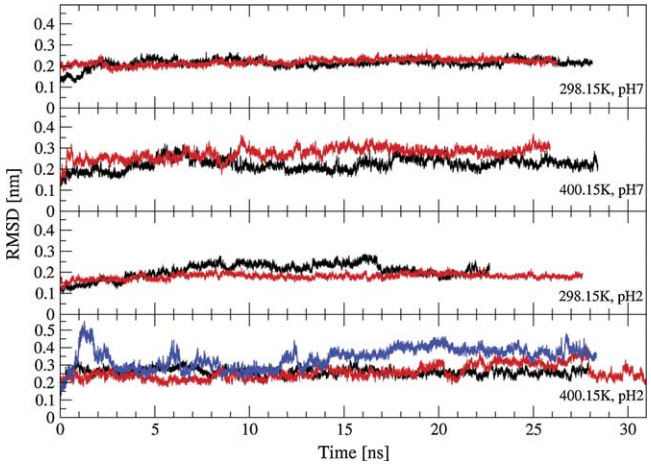

Fig. 4. Time-series of backbone root-mean-square-deviation from the X-ray structure of the native Top7 (1QYS) for Top7$_{wt}$ (black), Top7$_{CB1}$ (red) and Top7$_{CB1a}$ (blue). Rotational and translational fitting of pairs of structures was applied using Cα atoms. Residues 26–33 of Top7$_{CB1}$ and Top7$_{CB1a}$ were excluded from the RMSD calculations and structures fitting.

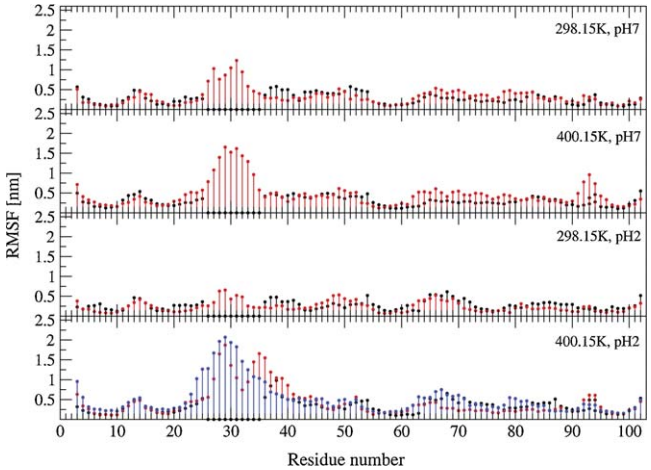

Fig. 5. Root-mean-square atom-positional fluctuations (RMSF) of Cα atoms of the MD simulations as function of residue sequence number. Values were calculated for the final 20 ns of simulation. Top7$_{wt}$ (black), Top7$_{CB1}$ (red) and Top7$_{CB1a}$ (blue). The sequence number of Top7$_{wt}$ (93 aa) was adjusted to match the numbering of the engineered sequences (100 aa).

It can be attributed to crystal-packing forces that restrict protein motions in the crystalline environment. Recently, this effect has been systematically characterized through MD simulations of long time scale [5,6]. A systematic characterization the effect of the loop-insert increases the R_g of native Top7 by 0.05–0.07 nm depending on the pH of the simulation (Table 2). The time-evolution of the secondary structure patterns is also equivalent among native and engineered Top7 ensembles (Fig. 6). The anti-parallel β-sheet is very stable under different pH and temperature conditions. The average β-sheet content in most MD-derived ensembles is around 34% in agreement with the β-sheet content in the crystallographic structure of native Top7 determined at pH 6.6 and 298.15 K [48]. The largest decrease of β-sheet content is seen for the $Top7_{CB1a}$ simulation at pH2 and 400.15 K where there is a loss of approximately four residues over the total five β-strands. The structural dynamics of the α-helices is more heterogeneous among the different simulations, being affected by protein sequence, temperature and pH conditions. An overall decrease of α-helix content occurs for structures in solution compared to the crystalline environment. This trend is more prominent for the native Top7 simulations (Table 2) whereas the engineered mutants exhibit slightly higher helix content if simulations under the same conditions are compared. At high temperature, there is a significant loss of helical content in the region corresponding to helix H1 for the native and engineered variants of Top7 whereas the loss of helical content along helix H2 is smaller for the variant sequences (Fig. 6). It is interesting to note that the N-terminal region of helix H1 of the native Top7, in particular residues 24–31, exhibit more elevated B-factors than the C-terminal half [48]. Recently, several non-homologous amino acid sequences created with a tetra-peptide fragment-based design algorithm led to protein variants that exhibited cooperative folding and high thermodynamic stability [19]. The NMR structure of the most stable of these variants (M7) revealed a well-ordered artificial protein that adopts the same fold as Top7 [77]. As observed for the X-ray structure of Top7, the N-terminal half of the M7 structure (*i.e.*, residues Asp3-Gly42 comprising the secondary structures S1-S2-H1) displays a significantly lower structural quality (average backbone RMSD = 1.26 ± 0.21 Å) compared to the C-terminal half (*i.e.*, residues Arg44-Arg91 comprising S3, S4, S5 and H2) with an average backbone RMSD of 0.70 ± 0.12 Å. Moreover, H/D exchange data also demonstrated that the majority of M7 residues in the C-terminal half of the protein show slowed amide proton exchange, whereas in the N-terminal half no slowed exchange was observed at all in the helix H1 and only for two residues in the β-sheet [77]. Altogether, these experimental data indicate that the C-terminal half of Top7 is more rigid, as evidenced by a stronger hydrogen bond network and in agreement with our MD simulations.

Hydrophobic, hydrophilic and total solvent accessible surface areas (SASA) were also calculated for the MD-derived ensemble of structures and average values are presented in Fig. 8. The loop-insert residues led to a maximum increase of the wild-type total SASA in approximately 11% and 7% for simulations at 298.15 and 400.15 K, respectively. In the absence of the loop-insert residues, the total SASA for the native and engineered variant of Top7 in simulations at 400.15 K are nearly indistinguishable. The total SASA for the engineered variant $Top7_{CB1}$ in simulations at 298.15 K are also analogous to their counterpart at higher temperatures while being on average 4.0% larger than the total SASA for the wild-type simulations under corresponding conditions. Given the similarity of the ensemble-averaged structures for native and engineered variants of Top7 in simulations at 298.15 K (Figs. 6 and 7), this difference in total SASA, if significant at all, is expected to be confined to the side chain of residues in the loop-insert region. The SASA analysis, consistently with the previous molecular properties analysis, indicates that the loop insert is located entirely on the surface of the protein where it adopts a solvent exposed conformation.

An accurate description of the native state structure of proteins can be obtained from the matrix of its native contacts. Contact maps contain the essential geometrical definition of the topology of the native structure, providing a quantitative means to compare different conformations, structures and interaction patterns in proteins [34,85]. Contact maps of the distance between α-carbon atoms were calculated for the conformational ensemble of structures corresponding to the last 5 ns of simulation for the nine systems (Fig. 9). The contact maps of native and engineered Top7 simulations exhibit similar patterns of Cα-atom distances at different pH or temperature. In the context of comparison between contact maps of native and engineered Top7, differences are confined to the region of loop insert (residues 26–33 in $Top7_{CB1}$). In the native Top7 the region corresponding to the S2-loop-H1 motif (residues 20–40) is spatially close to the S3-loop-H2 region (residues 52–68). These same contacts are also present in the simulations of the engineered variants of Top7, though the sequence number differs from the native Top7 sequence due to the loop insert (Fig. 9). In the Top7 variants, the S2-loop-H1 region corresponds to residues 20–25 and 34–48, and the S3-loop-H2 one to residues 60–76. Significant differences between the contact maps of the native and engineered Top7 ensembles are limited to atomic distances between the loop-insert and the S3-loop-H2 region; these differences are all the most noticeable for high temperature ensembles. High-temperature MD simulations can overcome limited sampling by accelerating transitions over high-energy barriers through the addition of kinetic energy to the system. These simulations will increase the thermal motions of the system, which often result in an increase of atomic displacement in flexible regions of protein structure.

The loop-insert is fairly flexible and exhibits larger atomic displacement compared to the much shorter S2-loop-H1 region in the native structure (Fig. 5). The increase of the contact distances between Cα-atoms in the loop-insert and in the S3-loop-H2 region at high temperature is indicative of a less structured, more "open" conformation with respect to the protein surface. It further supports the assertion that the loop insert is located entirely on the surface of the protein. As for the scaffold structure, the analysis of the contact maps show that the overall fold, packing contacts and tertiary structure are equivalent for native and engineered Top7 ensembles under equivalent environmental conditions. Likewise, the concerted mutations in the two helices (E43A, D46A, E67A, E69A, EA81A) do not lead to significant changes in tertiary structure with respect to $Top7_{wt}$ and $Top7_{CB1}$ under similar conditions of pH and temperature and for the time-scale simulated.

The design process of novel and highly stable proteins typically disfavors buried polar interactions and favors peripheral hydrophobic residues [47]. Top7 is an unambiguous example of such principle. The MD simulations of the native and engineered variants of Top7 show that, despite a decrease of helical content at high temperature, the anti-parallel β-sheet remains folded and forms extensive hydrophobic contacts with the two helices and interconnecting loops (Fig. 9). Because of the non-specific nature of hydrophobic interactions, non-native packing arrangements of the β-sheet with the partially unfolded helices are relatively stable as observed in the MD-derived ensembles. Indeed, the folding of Top7 is significantly less cooperative than the folding of similarly sized natural proteins with formation of non-native stable conformations at equilibrium and of multiple fragments that are stable in isolation [67,86]. In support to our argument, the formation of these non-native stable states has been attributed to the highly hydrophobic core of Top7 [67,86]. Furthermore, the existence of multiple non-native structure states during the folding of Top7

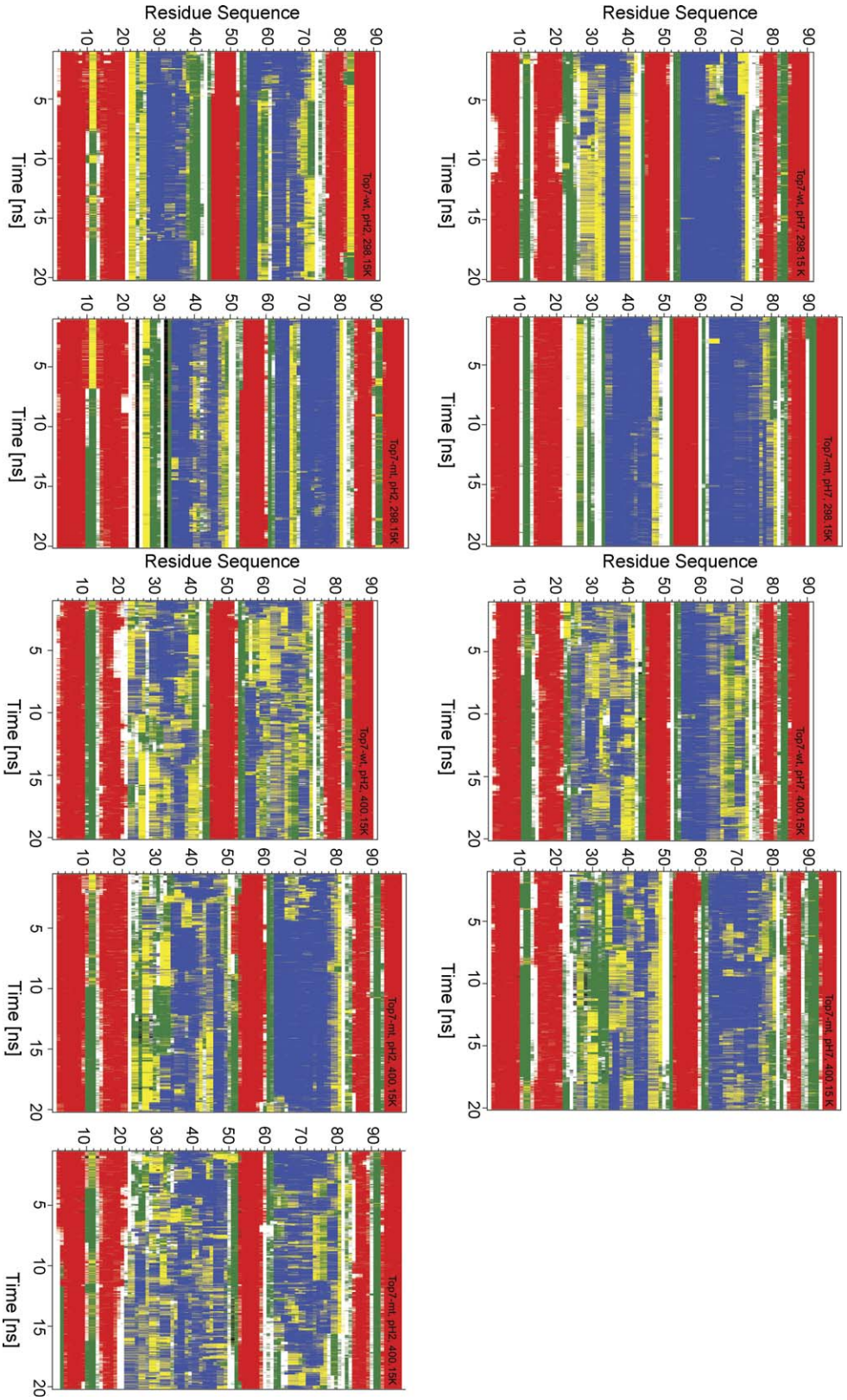

Fig. 6. Time-evolution of the secondary structure elements for MD simulations of $Top7_{wt}$, $Top7_{CB1}$ and $Top7_{CB1a}$ under different conditions of pH and temperature and for the last 20 ns of simualtions. Color patterns are blue for helices, red for β-sheet, green for bend, yellow for turn and white for coil and black for β-bridge. Secondary structure definition according to Kabsch and Sander [40]. Analyses were carried out over intervals of 5 ps.

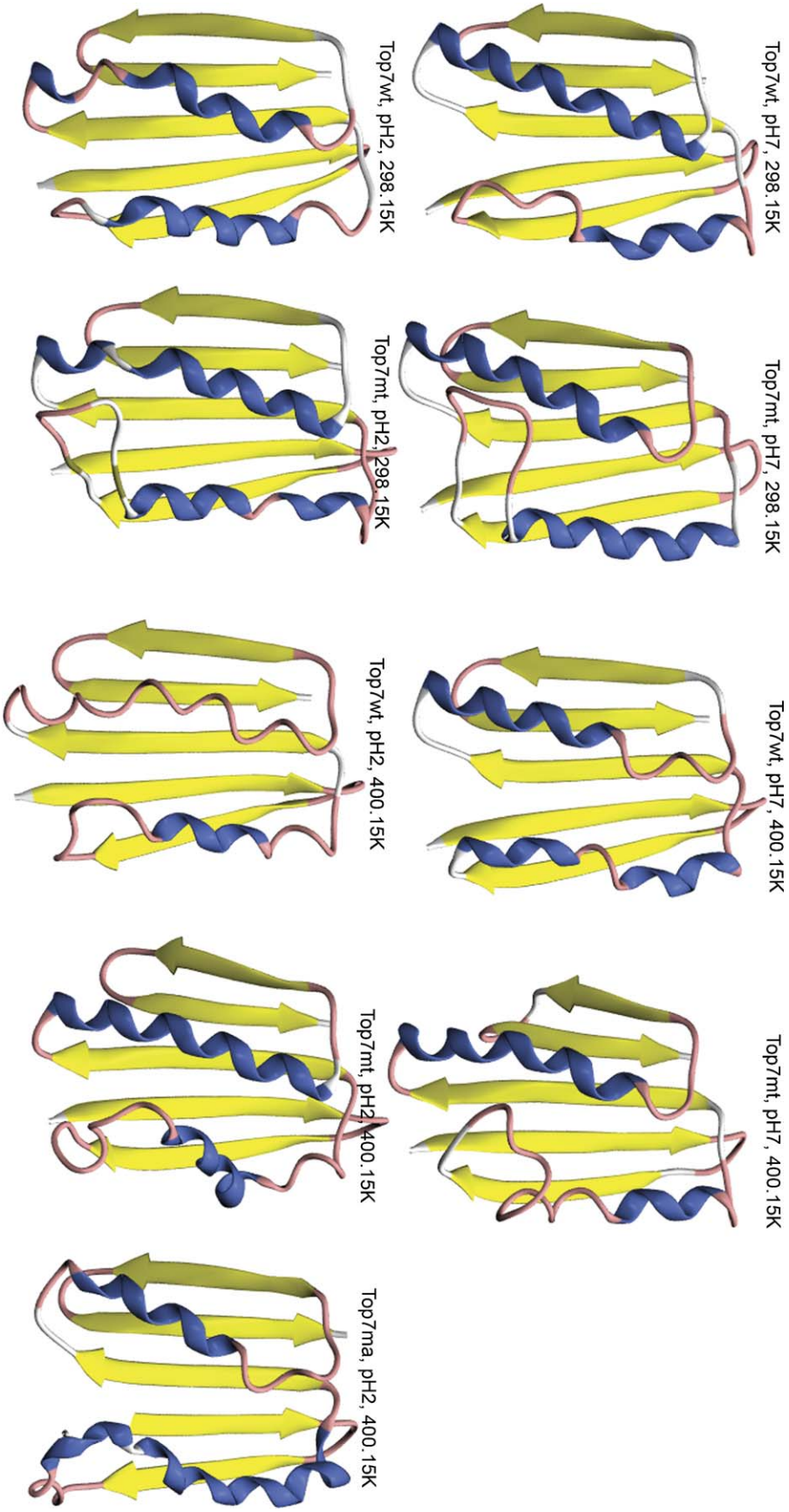

Fig. 7. MD-derived average conformations of $Top7_{wt}$, $Top7_{CB1}$ and $Top7_{CB1a}$. Structures were averaged over the last 20 ns of simulation and intervals of 5 ps. The illustration was prepared with the VMD software version 1.8 [38].

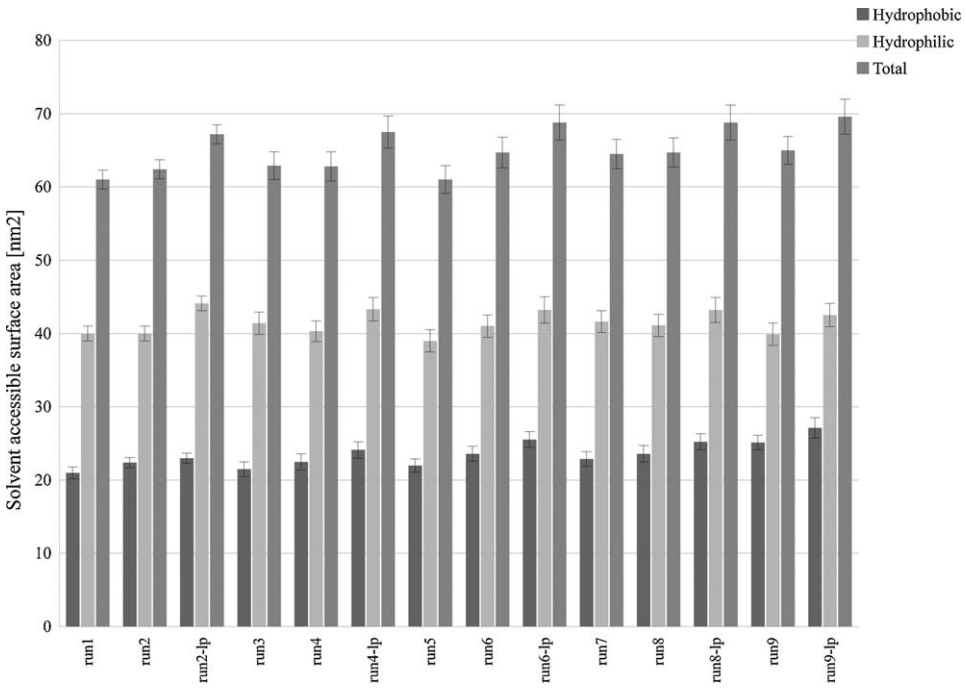

Fig. 8. Solvent accessible surface area for $Top7_{wt}$, $Top7_{CB1}$ and $Top7_{CB1a}$. Averages were carried out over intervals of 2 ps and along the last 20 ns of simulation. Two sets of values are reported for $Top7_{CB1}$ and $Top7_{CB1a}$, accordingly to the exclusion (run2-lp, run4-lp, run6-lp, run8-lp, run9-lp) or inclusion (run2, run4, run6, run8, run9) of the loop residues in the calculations. See Table 1 for simulated systems.

could also explain its tolerance to single or multiple mutations. Watters et al. [89] have examined the folding kinetics of 11 point mutants of Top7, mostly involving side-chain truncations like the ones performed for the engineered $Top7_{CB1a}$. Strikingly, these mutations did not significantly destabilize the non-native states, but led to even more complex folding kinetics presumably because more non-native states became populated. Since the folding kinetics of the Top7 variant sequences has not been characterized, caution is needed before any extrapolation can be made from the complex folding behavior of the native Top7. However, the similarity of the native and engineered Top7 conformational ensembles under equivalent conditions suggests that their folding behavior may not differ appreciably.

It is illustrative to compare the structural features of the heat-resistant $Top7_{CB1a}$ with the cold-active α-amylase from the Antarctic bacterium *Pseudoalteromonas haloplanktis*, one of the best

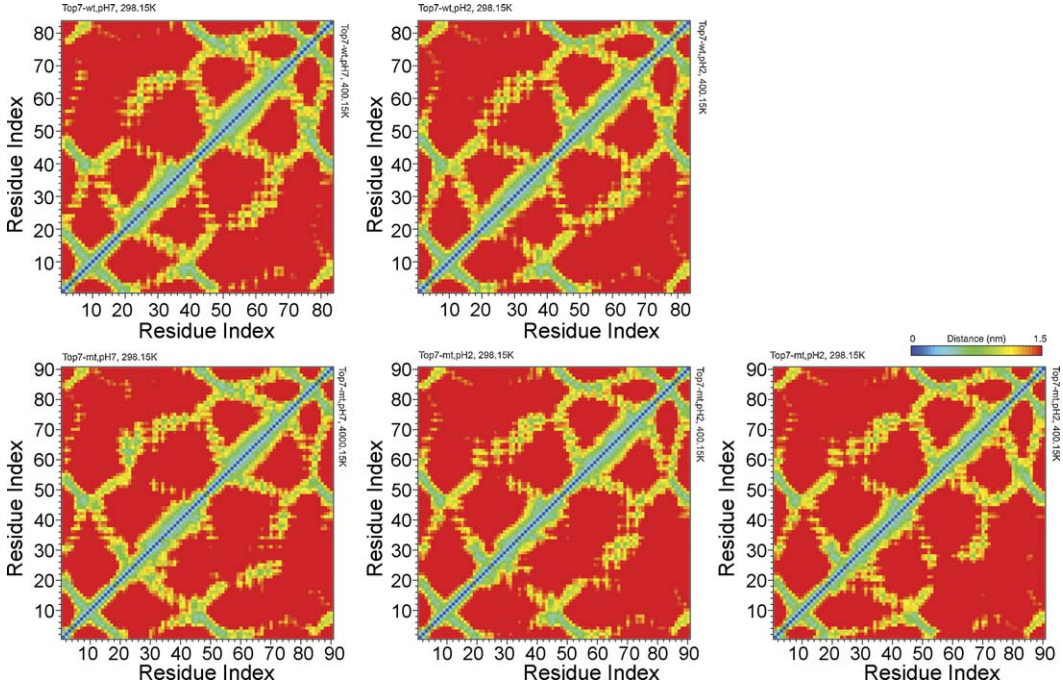

Fig. 9. Maps of average residue-residue distances for Cα-atom pairs within 4.5 nm of distance in MD-generated ensembles of $Top7_{wt}$, $Top7_{CB1}$ and $Top7_{CB1a}$. For each system the top moiety of the contact map corresponds to the simulation at 298.15 K and the bottom to the simulation of the same system at 400.15 K. For the case of $Top7_{CB1a}$ where there is only one simulation at 400.15 K, comparisons are made with the simulation of $Top7_{CB1}$ at 298.15 K. Distances between residues A_i and A_j are the smallest distance between any pair (i,j) of atoms where $i \in A_i$, $j \in A_j$. Averages were carried out over the final 20 ns of simulation.

Table 2
Average properties derived from MD simulations[a].

System[b]	Sequence	RMSD [nm]	R_g [nm]	Residues per secondary structure [%]		Hydrogen bonds[c] (pairs within 0.35 nm)	SASA [nm^2]		
				α-Helix	β-Strand		Hydrophobic	Hydrophilic	Total
X-ray[d]	1QYS	0.00	1.26	31	34	56	21.0 ± 0.8	40.0 ± 1.0	61.0 ± 1.3
Run1	$Top7_{wt}$	0.22 ± 0.02	1.30 ± 0.01	26.4 ± 3.2	34.8 ± 2.0	50 ± 3.7	23.0 ± 0.7	44.1 ± 1.0	67.2 ± 1.3
Run2	$Top7_{CB1}$	0.22 ± 0.01	1.35 ± 0.01	28.4 ± 2.2	33.8 ± 0.5	50 ± 2.9	21.5 ± 1.0	41.4 ± 1.5	62.9 ± 1.9
Run3	$Top7_{wt}$	0.22 ± 0.03	1.34 ± 0.01	23.7 ± 4.5	34.5 ± 1.7	44 ± 3.6	24.1 ± 1.1	43.3 ± 1.6	67.5 ± 2.2
Run4	$Top7_{CB1}$	0.26 ± 0.03	1.36 ± 0.01	24.7 ± ± 4.2	30.1 ± 1.4	45 ± 3.8	22.0 ± 0.9	39.0 ± 1.5	61.0 ± 1.9
Run5	$Top7_{wt}$	0.21 ± 0.04	1.30 ± 0.01	23.8 ± ± 2.9	37.6 ± 1.8	46 ± 3.4	25.5 ± 1.1	43.2 ± 1.8	68.8 ± 2.4
Run6	$Top7_{CB1}$	0.18 ± 0.01	1.37 ± 0.01	25.7 ± 2.5	34.5 ± 1.6	53 ± 3.1	22.9 ± 1.0	41.6 ± 1.5	64.5 ± 2.0
Run7	$Top7_{wt}$	0.26 ± 0.02	1.32 ± 0.01	17.8 ± 4.8	34.3 ± 2.3	42 ± 3.7	25.2 ± 1.1	43.2 ± 1.7	68.8 ± 2.4
Run8	$Top7_{CB1}$	0.26 ± 0.04	1.36 ± 0.03	27.9 ± 3.7	33.4 ± 2.3	50 ± 4.3	27.1 ± 1.4	42.5 ± 1.6	69.6 ± 2.4
Run9	$Top7_{CB1a}$	0.34 ± 0.06	1.38 ± 0.03	26.5 ± 5.2	29.6 ± 2.9	47 ± 4.6	21.0 ± 0.8	40.0 ± 1.0	61.0 ± 1.3

[a] Reported values are averages over the final 20 ns of the respective MD simulation. RMSD values for all backbone atoms except residues 26–33 of the mutants $Top7_{CB1}$ and $Top7_{CB1a}$; R_g and SASA values for all heavy-atoms; hydrogen-bond values calculated for main chain atoms (backbone, Cβ and hydrogen atoms).
[b] See Table 1 for description of simulated systems.
[c] Donor–acceptor distance cutoff of 0.35 nm and acceptor-hydrogen-donor angle cutoff of 30°.
[d] X-ray structure solved at 2.5 Å resolution [24].

characterized psychrophilic enzymes [13–18,22,28,29,37,55,70,90]. Although the psychrophilic α-amylase displays high sequence and structural similarity to its mesophilic porcine counterpart, the former exhibit longer loops that extend further away from the core of the barrel where the active site is located. Recently, it has been shown that the highly solvent-accessible loops of the psychrophilic α-amylase exhibit significantly greater flexibility than the mesophilic counterpart. Like $Top7_{CB1a}$, the comparatively larger flexibility is restricted to the loop regions whereas the structural core is more rigid for the psychrophilic and mesophilic α-amylase [62]. The structural rigidity of Top7 appears to be a precondition to tolerate very high temperatures without unfolding. However, specific binding requires some flexibility to recognize and interact with the biding partner. It appears that the heat-resistant $Top7_{CB1a}$ and psychrophilic α-amylase have adopted a similar strategy to conciliate thermal stability and functionality.

4. Conclusions

The exposure of immobilized biomolecules to changes in conditions such as pH and temperature can lead to significant loss of activity and is a source of potential problems in the design of biosensor. High temperatures can lead to denaturation of the molecule, and low pH conditions can result in protein conformations that may not be favorable for target molecule binding. We have previously shown that the robust properties of Top7, coupled with the ease in production, make it a suitable scaffold for use as an affinity reagent for applications outside of the highly controlled laboratory setting described [11]. Here, these studies have been expanded to include the associated effects of low pH, high temperature and multiple mutations on the stability of the Top7-based scaffold though far-UV CD spectroscopy in combination with a series of nine atomistic MD simulations in explicit solvent. The designed Top7 exhibits a well-defined fold as indicated by CD measurements at different pH (2 and 7) and temperature (298.15–368.15 K) conditions and has a similar structural content to that in native Top7. Several structural properties relevant for the molecular characterization of the conformational ensemble have been compared for three different systems: the native Top7, the eight-residue loop insert mutant $Top7_{CB1}$ and the loop insert mutant plus six additional mutations, $Top7_{CB1a}$. It has been found that the insertion of an eight-residue loop into the structure of Top7 does not adversely affect the global fold or structural stability of the native structure, even under extreme conditions of pH and temperature. It has also been proposed that the molecular basis of the high chemo- and thermostability, and the mutational tolerance of the Top7 scaffold is the very stable, hydrophobic β-sheet core. In fact, the highly hydrophobic core of Top7 has been previously linked to the high degeneracy of stable non-native states observed during the folding kinetics of Top7 and single-point variants [67,86].

Protein-based drugs represent an increasing percentage of new pharmaceuticals being brought to clinical trials each year. A protein drug must have a sufficiently long lifetime in the body in order to be effective and a protein's resistance to proteolysis has been found to generally correlate with its thermodynamic stability [41,42,46,49,56,60,63]. The stability at acidic pH and small size of Top7 suggest the potential use of theraputics based on the Top7 scaffold in non-traditional routs of administration for protein drugs, namely oral and inhalational.

Acknowledgments

This research was supported by the D.O.E. Office of Advanced Scientific Computing Research. The authors acknowledge the William R. Wiley Environmental Molecular Sciences Laboratory for the computational resources required for this work. Pacific Northwest National Laboratory is operated for the U.S. Department of Energy by Battelle.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2010.01.013.

References

[1] S.A. Adcock, J.A. McCammon, Molecular dynamics: survey of methods for simulating the activity of proteins, Chem. Rev. 106 (2006) 1589–1615.
[2] C. Báes, L. Briant-Longuet, M. Cerutti, F. Heitz, S. Troadec, M. Pugniáere, F. Roquet, F. Molina, F. Casset, D. Bresson, et al., Mapping the paratope of anti-cd4 recombinant fab 13b8.2 by combining parallel peptide synthesis and site-directed mutagenesis, J. Biol. Chem. 278 (2003) 14265–14273.
[3] F. Baneyx, M. Mujacic, Recombinant protein folding and misfolding in *Escherichia coli*, Nat. Biotechnol. 22 (2004) 1399–1408.
[4] R. Baron, D. Bakowies, W.F. van Gunsteren, Carbopeptoid folding: effects of stereochemistry, chain length, and solvent, Angew. Chem. Int. Ed. 43 (2004) 4055–4059.
[5] R. Baron, J.A. McCammon, Dynamics, hydration, and motional averaging of a loop-gated artificial protein cavity: the w191g mutant of cytochrome c peroxidase in water as revealed by molecular dynamics simulations, Biochemistry 46 (2007) 10629–10642.
[6] R. Baron, S.E. Wong, C.A. de Oliveira, J.A. McCammon, E9-im9 colicin dnase-immunity protein biomolecular association in water: a multiple-copy and accelerated molecular dynamics simulation study, J. Phys. Chem. B 112 (2008) 16802–16814.

[7] H.J.C. Berendsen, J.P.M. Postma, W.F. Vangunsteren, A. Dinola, J.R. Haak, Molecular-dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.
[8] H.K. Binz, P. Amstutz, A. Plueckthun, Engineering novel binding proteins from non-immunoglobulin domains, Nat. Biotechnol. 23 (2005) 1257–1268.
[9] H.K. Binz, A. Plueckthun, Engineered proteins as specific binding reagents, Curr. Opin. Biotechnol. 16 (2005) 459–469.
[10] E. Bismuto, E. Di Maggio, S. Pleus, M. Sikor, C. Rocker, G.U. Nienhaus, D.C. Lamb, Molecular dynamics simulation of the acidic compact state of apomyoglobin from yellowfin tuna, Proteins: Struct. Funct. Bioinformatics 74 (2009) 273–290.
[11] C.B. Boschek, D.O. Apiyo, T.A. Soares, H.E. Engelmann, N.B. Pefaur, T.P. Straatsma, C.L. Baird, Engineering an ultra-stable affinity reagent based on top7, Protein Eng. Des. Sel. 22 (2009) 325–332.
[12] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz, D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules, J. Am. Chem. Soc. 117 (1995) 5179–5197.
[13] S. D'Amico, P. Claverie, T. Collins, D. Georlette, E. Gratia, A. Hoyoux, M.A. Meuwis, G. Feller, C. Gerday, Molecular basis of cold adaptation, Philos. Trans. R. Soc. Lond. B Biol. Sci. 357 (2002) 917–925.
[14] S. D'Amico, T. Collins, J.C. Marx, G. Feller, C. Gerday, Psychrophilic microorganisms: challenges for life, EMBO Rep. 7 (2006) 385–389.
[15] S. D'Amico, C. Gerday, G. Feller, Dual effects of an extra disulfide bond on the activity and stability of a cold-adapted alpha-amylase, J. Biol. Chem. 277 (2002) 46110–46115.
[16] S. D'Amico, C. Gerday, G. Feller, Structural determinants of cold adaptation and stability in a large protein, J. Biol. Chem. 276 (2001) 25791–25796.
[17] S. D'Amico, C. Gerday, G. Feller, Temperature adaptation of proteins: engineering mesophilic-like activity and stability in a cold-adapted alpha-amylase, J. Mol. Biol. 332 (2003) 981–988.
[18] S. D'Amico, J.C. Marx, C. Gerday, G. Feller, Activity–stability relationships in extremophilic enzymes, J. Biol. Chem. 278 (2003) 7891–7896.
[19] R. Dalluege, J. Oschmann, O. Birkenmeier, C. Luecke, H. Lilie, R. Rudolph, C. Lange, A tetrapeptide fragment-based design method results in highly stable artificial proteins, Proteins: Struct. Funct. Bioinformatics 68 (2007) 839–849.
[20] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh ewald method, J. Chem. Phys. 103 (1995) 8577–8593.
[21] H. Fan, A.E. Mark, J. Zhu, B. Honig, Comparative study of generalized born models: protein dynamics, Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 6760–6764.
[22] G. Feller, S. D'Amico, A.M. Benotmane, F. Joly, J. Van Beeumen, C. Gerday, Characterization of the c-terminal propeptide involved in bacterial wall spanning of alpha-amylase from the psychrophile alteromonas haloplanctis, J. Biol. Chem. 273 (1998) 12109–12115.
[23] A.C. Fisher, W. Kim, M.P. DeLisa, Genetic selection for protein solubility enabled by the folding quality control feature of the twin-arginine translocation pathway, Protein Sci. 15 (2006) 449–458.
[24] T. Fox, P.A. Kollman, The application of different solvation and electrostatic models in molecular dynamics simulations of ubiquitin: how well is the X-ray structure ''maintained''? Proteins: Struct. Funct. Bioinformatics 25 (1996) 315–334.
[25] M. Friedman, E. Nordberg, I. Hoiden-Guthenberg, H. Brisimar, G.P. Adams, F.Y. Nilsson, J. Carlss, S. Stahl, Phage display selection of affibody molecules with specific binding to the extracellular domain of the epidermal growth factor receptor, Protein Eng. Des. Sel. 20 (2007) 189–199.
[26] M. Gebauer, A. Skerra, Engineered protein scaffolds as next-generation antibody therapeutics, Curr. Opin. Chem. Biol. 13 (2009) 245–255.
[27] G. Georgiou, P. Valax, Expression of correctly folded proteins in *Escherichia coli*, Curr. Opin. Biotechnol. 7 (1996) 190–197.
[28] D. Georlette, V. Blaise, T. Collins, S. D'Amico, E. Gratia, A. Hoyoux, J.C. Marx, G. Sonan, G. Feller, C. Gerday, Some like it cold: biocatalysis at low temperatures, FEMS Microbiol. Rev. 28 (2004) 25–42.
[29] C. Gerday, M. Aittaleb, M. Bentahir, J.P. Chessa, P. Claverie, T. Collins, S. D'Amico, J. Dumont, G. Garsoux, D. Georlette, et al., Cold-adapted enzymes: from fundamentals to biotechnology, Trends Biotechnol. 18 (2000) 103–107.
[30] E.V. Getmanova, Y. Chen, L. Bloom, J. Gokemeijer, S. Shamah, V. Warikoo, J. Wang, V. Ling, L. Sun, Antagonists to human and mouse vascular endothelial growth factor receptor 2 generated by directed protein evolution in vitro, Chem. Biol. 13 (2006) 549–556.
[31] C. Gronwall, E. Snelders, A.J. Palm, F. Eriksson, N. Herne, S. Stahl, Generation of affibody (r) ligands binding interieukin-2 receptor alpha/cd25, Biotechnol. Appl. Biochem. 50 (2008) 97–112.
[32] W. Gu, T.T. Wang, J. Zhu, Y.Y. Shi, H.Y. Liu, Molecular dynamics simulation of the unfolding of the human prion protein domain under low ph and high temperature conditions, Biophys. Chem. 104 (2003) 79–94.
[33] R.W. Hockney, The potential calculation and some applications, in: B. Alder, S. Fernbach, M. Rotenberg (Eds.), Methods in Computational Physics, Academic Press, New York/London, 1970, pp. 135–211.
[34] L. Holm, C. Sander, Mapping the protein universe, Science 273 (1996) 595–602.
[35] R.J. Hosse, A. Rothe, B.E. Power, A new generation of protein display scaffolds for molecular recognition, Protein Sci. 15 (2006) 14–27.
[36] W. Hoyer, C. Gronwall, A. Jonsson, S. Stahl, T. Hard, Stabilization of a beta-hairpin in monomeric alzheimer's amyloid-beta peptide inhibits amyloid formation, Proc. Natl. Acad. Sci. U.S.A. 105 (2008) 5099–5104.
[37] A. Hoyoux, V. Blaise, T. Collins, S. D'Amico, E. Gratia, A.L. Huston, J.C. Marx, G. Sonan, Y. Zeng, G. Feller, et al., Extreme catalysts from low-temperature environments, J. Biosci. Bioeng. 98 (2004) 317–330.
[38] W. Humphrey, A. Dalke, K. Schulten, Vmd: visual molecular dynamics, J Mol. Graphics 14 (1996) 33.
[39] H. Jang, N. Michaud-Agrawal, J.M. Johnston, T.B. Woolf, How to lose a kink and gain a helix: Ph independent conformational changes of the fusion domains from influenza hemagglutinin in heterogeneous lipid bilayers, Proteins: Struct. Funct. Bioinformatics 72 (2008) 299–312.
[40] W. Kabsch, C. Sander, Dictionary of protein secondary structure–pattern-recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577–2637.
[41] K.H. Khoo, A. Andreeva, A.R. Fersht, Adaptive evolution of p53 thermodynamic stability, J. Mol. Biol. 393 (2009) 161–175.
[42] C. Klein, G. Georges, K.P. Kunkele, R. Huber, R.A. Engh, S. Hansen, High thermostability and lack of cooperative DNA binding distinguish the p63 core domain from the homologous tumor suppressor p53, J. Biol. Chem. 276 (2001) 37390–37401.
[43] A. Kohl, H.K. Binz, P. Forrer, M.T. Stumpp, A. Pluckthun, M.G. Grutter, Designed to be stable: crystal structure of a consensus ankyrin repeat protein, Proc. Natl. Acad. Sci. U.S.A. 100 (2003) 1700–1705.
[44] A. Koide, S. Koide, Monobodies:, Antibody mimics based on the scaffold of the fibronectin type iii domain, Methods Mol. Biol. 352 (2007) 95–109.
[45] D.B. Kony, P.H. Hunenberger, W.F. van Gunsteren, Molecular dynamics simulations of the native and partially folded states of ubiquitin: influence of methanol cosolvent, ph, and temperature on the protein structure and dynamics, Protein Sci. 16 (2007) 1101–1118.
[46] J.M. Kowalski, R.N. Parekh, J. Mao, K.D. Wittrup, Protein folding stability can determine the efficiency of escape from endoplasmic reticulum quality control, J. Biol. Chem. 273 (1998) 19453–19458.
[47] B. Kuhlman, D. Baker, Exploring folding free energy landscapes using computational protein design, Curr. Opin. Struct. Biol. 14 (2004) 89–95.
[48] B. Kuhlman, G. Dantas, G.C. Ireton, G. Varani, B.L. Stoddard, D. Baker, Design of a novel globular protein fold with atomic-level accuracy, Science 302 (2003) 1364–1368.
[49] W.S. Kwon, N.A. Da Silva, J.T. Kellis, Relationship between thermal stability, degradation rate and expression yield of barnase variants in the periplasm of *Escherichia coli*, Protein Eng. Des. Sel. 9 (1996) 1197–1202.
[50] B. Li, V. Daggett, The molecular basis of the temperature- and ph-induced conformational transitions in elastin-based peptides, Biopolymers 68 (2003) 121–129.
[51] H. Li, A.D. Robertson, J.H. Jensen, Very fast empirical prediction and interpretation of protein pKa values, Proteins: Struct. Funct. Bioinf. (2005) 704–721.
[52] N. Lonberg, Fully human antibodies from transgenic mouse and phage display platforms, Curr. Opin. Immunol. 20 (2008) 450–459.
[53] S. Luetz, L. Giver, J. Lalonde, Engineered enzymes for chemical production, Biotechnol. Bioeng. 101 (2008) 647–653.
[54] S.C. Makrides, Strategies for achieving high-level expression of genes in *Escherichia coli*, Microbiol. Rev. 60 (1996) 512–538.
[55] J.C. Marx, T. Collins, S. D'Amico, G. Feller, C. Gerday, Cold-adapted enzymes from marine antarctic microorganisms, Mar. Biotechnol. (NY) 9 (2007) 293–304.
[56] S. Mayer, S. Rudiger, H.C. Ang, A.C. Joerger, A.R. Fersht, Correlation of levels of folded recombinant p53 in *Escherichia coli* with thermodynamic stability in vitro, J. Mol. Biol. 372 (2007) 268–276.
[57] S.D. Nuttall, R.B. Walsh, Display scaffolds: protein engineering for novel therapeutics, Curr. Opin. Pharmacol. 8 (2008) 609–615.
[58] P.A. Nygren, A. Skerra, Binding proteins from alternative scaffolds, J. Immunol. Methods 290 (2004) 3–28.
[59] J.P. Overington, B. Al-Lazikani, A.L. Hopkins, Opinion—how many drug targets are there? Nat. Rev. Drug Discov. 5 (2006) 993–996.
[60] A.A. Pakula, R.T. Sauer, Amino acid substitutions that increase the thermal stability of the lambda cro protein, Proteins: Struct. Funct. Bioinf. 5 (1989) 202–210.
[61] S. Park, J.G. Saven, Simulation of ph-dependent edge strand rearrangement in human beta-2 microglobulin, Protein Sci. 15 (2006) 200–207.
[62] M. Pasi, L. Riccardi, P. Fantucci, L. De Gioia, E. Papaleo, Dynamic properties of a psychrophilic alpha-amylase in comparison with a mesophilic homologue, J. Phys. Chem. B 113 (2009) 13585–13595.
[63] S. Patel, T.T.T. Bui, A.F. Drake, F. Fraternali, P.V. Nikolova, The p73 DNA binding domain displays enhanced stability relative to its homologue, the tumor suppressor p53, and exhibits cooperative DNA binding, Biochemistry 47 (2008) 3235–3244.
[64] J.M. Reichert, V.E. Valge-Archer, Outlook—development trends for monoclonal antibody cancer therapeutics, Nat. Rev. Drug Discov. 6 (2007) 349–356.
[65] A. Rothe, R.J. Hosse, B.E. Power, In vitro display technologies reveal novel biopharmaceutics, FASEB J. 20 (2006) 1599–1610.
[66] J.P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Numerical-integration of cartesian equations of motion of a system with constraints—molecular-dynamics of n-alkanes, J. Comp. Phys. 23 (1977) 327–341.
[67] M. Scalley-Kim, D. Baker, Characterization of the folding energy landscapes of computer generated proteins suggests high folding free energy barriers and cooperativity may be consequences of natural selection, J. Mol. Biol. 338 (2004) 573–583.
[68] D. Sharma, O. Perisic, Q. Peng, Y. Cao, C. Lam, H. Lu, H. Li, Single-molecule force spectroscopy reveals a mechanically stable protein fold and the rational tuning of its mechanical stability, Proc. Natl. Acad. Sci. U.S.A. 104 (2007) 9278–9283.
[69] C. Sheridan, Pharma consolidates its grip on post-antibody landscape, Nat. Biotechnol. 25 (2007) 365–366.

[70] K.S. Siddiqui, G. Feller, S. D'Amico, C. Gerday, L. Giaquinto, R. Cavicchioli, The active site is the least stable structure in the unfolding pathway of a multidomain cold-adapted alpha-amylase, J. Bacteriol. 187 (2005) 6197–6205.
[71] J. Silverman, Q. Lu, A. Bakker, W. To, A. Duguay, B.M. Alba, R. Smith, A. Rivas, P. Li, H. Le, et al., Multivalent avimer proteins evolved by exon shuffling of a family of human receptor domains, Nat. Biotechnol. 23 (2005) 1556–1561.
[72] J. Silverman, Q. Lu, A. Bakker, W. To, A. Duguay, B.M. Alba, R. Smith, A. Rivas, P. Li, H. Le, et al., Multivalent avimer proteins evolved by exon shuffling of a family of human receptor domains (vol 23, pg 1556, 2005), Nat. Biotechnol. 24 (2006) 220–1220.
[73] A. Skerra, Engineered protein scaffolds for molecular recognition, J. Mol. Recogn. 13 (2000) 167–187.
[74] A. Skerra, Imitating the humoral immune response, Curr. Opin. Chem. Biol. 7 (2003) 683–693.
[75] T. Soares, M. Christen, K.F. Hu, W.F. van Gunsteren, Alpha- and beta-polypeptides show a different stability of helical secondary structure, Tetrahedron 60 (2004) 7775–7780.
[76] J. Sorensen, D. Hamelberg, B. Schiott, J.A. McCammon, Comparative md analysis of the stability of transthyretin providing insight into the fibrillation mechanism, Biopolymers 86 (2007) 73–82.
[77] C. Stordeur, R. Dalluege, O. Birkenmeier, H. Wienk, R. Rudolph, C. Lange, C. Luecke, The nmr solution structure of the artificial protein m7 matches the computationally designed model, Proteins: Struct. Funct. Bioinformatics 72 (2008) 1104–1107.
[78] T.P. Straatsma, E. Aprà, T.L. Windus, E.J. Bylaska, W. de Jong, S. Hirata, M. Valiev, M. Hackler, L. Pollack, R. Harrison, et al., Nwchem, a Computational Chemistry Package for Parallel Computers, Version 4.6, Pacific Northwest National Laboratory, Richland, WA, 2004.
[79] J.Z. Tan, K.H.G. Verschueren, K. Anand, J.H. Shen, M.J. Yang, Y.C. Xu, Z.H. Rao, J. Bigalke, B. Heisen, J.R. Mesters, et al., Ph-dependent conformational flexibility of the sars-cov main proteinase (m-pro) dimer: molecular dynamics simulations and multiple X-ray structure analyses, J. Mol. Biol. 354 (2005) 25–40.
[80] D. Trzesniak, R.D. Lins, W.F. Van Gunsteren, Protein under pressure: molecular dynamics simulation of the arc repressor, Proteins: Struct. Funct. Bioinformatics 65 (2006) 136–144.
[81] W. van Gunsteren, D. Bakowies, R. Burgi, I. Chandrasekhar, M. Christen, X. Daura, P. Gee, A. Glattli, T. Hansson, C. Oostenbrink, et al., Molecular dynamics simulation of biomolecular systems, Chimia 55 (2001) 856–860.
[82] W.F. van Gunsteren, D. Bakowies, R. Baron, I. Chandrasekhar, M. Christen, X. Daura, P. Gee, D.P. Geerke, A. Glattli, P.H. Hunenberger, et al., Biomolecular modeling: goals, problems, perspectives, Angew. Chem. Int. Ed. 45 (2006) 4064–4092.
[83] W.F. van Gunsteren, H.J. Berendsen, Thermodynamic cycle integration by computer simulation as a tool for obtaining free energy differences in molecular chemistry, J. Comput. Aided Mol. Des. 1 (1987) 171–176.
[84] W.F. van Gunsteren, J. Dolenc, A.E. Mark, Molecular simulation as an aid to experimentalists, Curr. Opin. Struct. Biol. 18 (2008) 149–153.
[85] M. Vendruscolo, E. Kussell, E. Domany, Recovery of protein structure from contact maps, Fold. Des. 2 (1997) 295–306.
[86] A.L. Watters, P. Deka, C. Corrent, D. Callender, G. Varani, T. Sosnick, D. Baker, The highly cooperative folding of small naturally occurring proteins is likely the result of natural selection, Cell (2007) 128.
[87] P.L. Wintrode, D.Q. Zhang, N. Vaidehi, F.H. Arnold, W.A. Goddard, Protein dynamics in a family of laboratory evolved thermophilic enzymes, J. Mol. Biol. 327 (2003) 745–757.
[88] Z. Xiang, C.S. Soto, B. Honig, Evaluating conformational free energies: the colony energy and its application to the problem of loop prediction, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 7432–7437.
[89] H.B. Yu, M. Amann, T. Hansson, J. Kohler, G. Wich, W.F. van Gunsteren, Effect of methylation on the stability and solvation free energy of amylose and cellulose fragments: a molecular dynamics study, Carbohydr. Res. 339 (2004) 1697–1709.
[90] L. Zecchinon, P. Claverie, T. Collins, S. D'Amico, D. Delille, G. Feller, D. Georlette, E. Gratia, A. Hoyoux, M.A. Meuwis, et al., Did psychrophilic enzymes really win the challenge? Extremophiles 5 (2001) 313–321.