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.

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

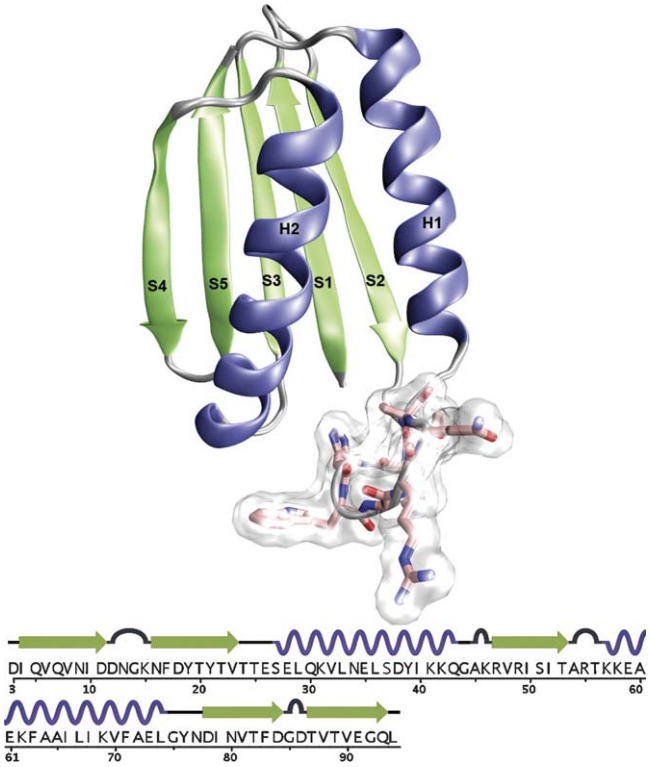

Fig. 1. Cartoon representation of the X-ray structure of the native Top7 (1QYS). The secondary structure elements are colored as function of the experimental B-factor values. Residues T25 and E26 indicate the site of insertion of the eight-residue loop in the engineered Top7 structures. Secondary structure 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.