Design, expression and characterization of a highly stable tetratricopeptide-based protein scaffold for phage display application

Edyta Petters¹, Daniel Krowarsch² and Jacek Otlewski³

¹Department of Protein Engineering and ²Department of Protein Biotechnology, Faculty of Biotechnology, University of Wrocław, Wrocław, Poland

INTRODUCTION

Engineered antibodies and their fragments for years have been the most successful and commonly used tools for biomolecular binding applications (Skerra, 2003). However, despite their undoubted merits, they remain rather difficult to obtain due to complicated molecular composition, posttranslational modifications and large size. Alternative protein scaffolds are a viable alternative to natural or recombinant antibodies (Skerra, 2007; Löfblom et al., 2011). Unlike to immunoglobulins, alternative scaffolds can be easily obtained in large quantities in inexpensive bacterial expression systems. However, one of the major challenges in design of such polypeptides is to engineer monomeric and highly stable scaffold. Increased stability of designed molecule is required to accommodate multiple substitutions, which are introduced to provide molecular interaction with defined target. Moreover, high stability ensures resistance to protease action and allows prolonged storage of such protein (Markert et al., 2001). Using phage, mRNA or yeast display polypeptide libraries based on designed stable scaffold can be subsequently used to select sequences, that recognize many molecular targets with high affinity and specificity (Binz et al., 2005).

Alternative scaffolds have been developed based on single structural domains or repeated polypeptide modules. The most successful examples of molecular binders based on single domain are: affibody (Löfblom et al., 2010), adnectin (Lipovsek et al., 2011), or anticalins (Skerra, 2008). The second large group of alternative scaffolds includes repetitive sequence motifs. These polypeptides are built of linearly arranged, small (typically 20-50 amino acids in length) structural motifs. Repetitive protein scaffolds have been already shown to successfully interact with proteins or peptides. Designed ankyrin repeat protein, binds HER2 with picomolar affinity (Zahnd et al., 2007), designed CTPR390+ protein containing three tetratricopeptide repeats interacts with the C-terminal peptide of Hsp90 with a $K_D$ of 1 μM (Cortajarena et al., 2008), and designed armadillo repeat protein binds neurotensin with a $K_D$ of 7 μM (Varadamsetty et al., 2012).

The tetratricopeptide repeat is a small, 34-amino acid, structural motif present in organisms from bacteria to humans (Sikorski et al., 1990). Proteins harboring repetitive TPR motifs mediate protein-protein interactions in a broad range of biochemical processes (D’Andrea & Regan, 2003). Large diversity in molecular targets shows an enormous potential of the TPR motif to mediate specific and strong interactions with different ligands. The number of repeats in TPR proteins varies from 3 to 16. Three consecutive repeats are most commonly found, and represent the smallest possible number of TPR motifs capable of ligand binding. Almost all solved TPR structures possess an additional capping helix at C-terminus, which seems to be essential for the solubility and stability of the domain (D’Andrea & Regan, 2003).

The TPR motifs are composed of helix-turn-helix elements, where $\alpha$-helices (termed A and B) are packed in anti-parallel fashion. Adjacent TPR motifs are parallel and adopt right-handed super-helical conformation, forming an amphipathic groove. The internal surface is formed mainly by the side-chains from helix A, whereas the opposite site of the protein contains amino acids from both helices A and B (D’Andrea & Regan, 2003).

Key words: tetratricopeptide repeat; alternative scaffold; consensus design; protein stability

Received: 27 May, 2013; revised: 12 October, 2013; accepted: 12 November, 2013; available on-line: 17 December, 2013

Abbreviations: TPR, tetratricopeptide repeat; dTPR, designed TPR protein containing three repeated TPR motifs; CD, circular dichroism; DSC, differential scanning calorimetry; GdmCl, guanidine hydrogenchloride; TEV, tobacco etch virus
Small size, repetitive nature, defined ligand binding area and highly degenerate sequence make TPR motif an interesting candidate for an alternative protein scaffold. In this paper we present the design, expression and biophysical characterization of a novel, extremely stable designed TPR (dTPR), which could serve as a molecular platform for combinatorial library selection.

MATERIAL AND METHODS

Design of the TPR consensus sequence and structure modeling. We collected 413 TPR sequences from SMART database (Schultz et al., 1998) and aligned them with ClustalW (Larkin et al., 2007). Amino acid frequency at all positions were calculated. The protein model was built using crystal structure of CTPR3 as a template (PDB ID: 1NA0) with Swiss-Model sever (Arnold et al., 2006).

Cloning and molecular biology. dTPR sequence (containing three repeated TPR motifs with N-capping sequence and solvating helix at the C-terminus) was optimized for E. coli expression, synthesized and cloned into pPCR-Script vector by Sloaning BioTechnology (Germany). At the N-terminal sequence of dTPR we added a cleavage sequence for TEV (tobacco etch virus) protease: NLYFQGSS. Recombination sites (attB) for the Gateway cloning system were added on both sides of the gene. dTPR was cloned to pDONR201 vector in Gateway BP reaction, and subsequently into the destination vector pDST15 in Gateway LR reaction.

Expression and purification of dTPR. E. coli BL21 (DE3) RIL strain was transformed with the pDST15 vector containing dTPR. Colonies were selected on LB-agar plates containing 34 µg/ml chloramphenicol and 100 µg/ml ampicillin. Transformants were grown in 200 ml of LB medium (100 µg/ml of ampicillin and 34 µg/ml chloramphenicol) overnight at 37°C, 200 rpm, then 25 ml of the overnight culture was transferred to 1 l of LB medium and grown as above to an OD600 of 0.8. Protein expression was induced with isopropyl-β-D-thiogalactopyranoside (IPTG) added to 0.6 mM and the bacteria were further cultivated overnight at 27°C. Cells were harvested by centrifugation (4°C, 3300 × g, 8 min), pellet from 4 l of culture was resuspended in 80 ml of ice-cold lysis buffer (150 mM NaCl, 5 mM EDTA, 50 mM Tris–HCl, pH 8.0), sonicated and centrifuged at 14 500 × g for 1 hour at 4°C. The supernatant was applied on Glutathione Sepharose column for 1.5 h at 4°C, and the resin was washed with 2 l of washing buffer (250 mM NaCl, 2.5 mM EDTA, 50 mM Tris–HCl, pH 8.0). Bound proteins were eluted with 60 ml of 100 mM NaCl, 25 mM reduced glutathione, 25 mM Tris, pH 8.0 and analyzed by 12% SDS-PAGE gel. Fractions with the highest protein concentration were pooled and GST-tag was cleaved with TEV protease (4°C, 48 h). dTPR was separated from GST-tag on Glutathione Sepharose. Concentrated dTPR was applied on Superdex 75 prep-grade column (GE Healthcare) equilibrated with 100 mM NaCl, 50 mM phosphate, pH 6.3 and eluted at 1 ml/min at 4°C. The homogeneity of the purified protein was analyzed by Superdex 75 analytical column (GE Healthcare). Protein identity was assessed with AB 4800+ MALDI TOF/TOF (Applied Biosystems) mass spectrometer.

Circular dichroism (CD) measurements. CD spectra were acquired on a Jasco J-715 spectropolarimeter in a 1 mm cuvette with the dTPR concentration of 1.4×10^{-6} M at 21°C that allowed us to measure good quality CD spectra within the range of 195 to 265 nm. The protein was dialyzed to the following buffers: 5 mM glycine at pH between 2.0 and 3.0, 5 mM citrate, pH 4.0, or 10 mM sodium phosphate at pH between 5.0 and 12.0. Spectra were averaged from three separate scans with a slit width set to 2 nm and a response time of 1 s. Estimation of a secondary structure content was performed with K2D2 software (Perz-Iratxeta & Andrade-Navarro, 2008).

CD-monitored thermal denaturation was carried out in 5 mM glycine, pH 2.0 and 3.0 or in 10 mM sodium phosphate, pH 10.0 or 11.0 with the protein concentration of 3.6×10^{-6} M. Thermal scans were performed in a 1 cm cuvette following ellipticity at 222 nm using a response time of 16 s. An automatic Peltier accessory PFD 350S allowed continuous monitoring of the thermal transition at a constant rate of 1°C/min. The data were analyzed using PeakFit software (Jandel Scientific Software) assuming a two-state reversible equilibrium transition as described previously (Zakrzewska et al., 2005).

Chemical denaturation of dTPR with guanidine chloride (GdmCl) was monitored with a Jasco J-715 spectropolarimeter at 21°C. Protein samples were incubated with a various concentrations of GdmCl in 10 mM Tris pH 7.5 or 10 mM glycine 2.0 for 24 h at 21°C. To ensure an adequate signal-to-noise ratio we used protein concentration of 3.6×10^{-6} M. The transitions were monitored by the changes of the CD signal at 222 nm with a 2 nm bandwidth. The free energy change of denaturation in the absence of denaturant (ΔG_{H2O}) was determined by fitting CD intensity changes as a function of GdmCl concentration, as described previously (Wezner-Ptasinska et al., 2011).

Differential scanning calorimetry (DSC) measurements. DSC experiments were performed on a Nano DSC II (Calorimetry Sciences Corp.) in the temperature range from 25 to 125°C at 7.16×10^{-3} M protein concentration (total cell volume 323 µl). The scan rate was 1.0°C/min under pressure excess of 2.5 atm. Combination of appropriate protein concentration and scan rate allowed us to determine a reversible, two state dTPR denaturation. Before DSC run protein was extensively dialyzed against 10 mM Gly-Gly buffer, pH 7.5. Baseline was determined by running the calorimeter with both cells filled with the dialysis buffer. Reversibility of denaturation was checked through repeated heating of the protein sample. Denaturation parameters were calculated from thermogram analysis using CpCalc software (Calorimetry Sciences Corp.).

RESULTS AND DISCUSSION

Design of TPR scaffold

In order to produce a folded and stable protein based on a TPR motif, we applied a method that combines consensus sequence design (Steipe et al., 1994) with protein charge neutralization (Cortajarena et al., 2004). We have successfully applied this approach to design stable scaffold based on leucine-reach repeat (Wezner-Ptasinska et al., 2011) as well as to increase protein stability (Zakrzewska et al., 2005). Proteins based on idealized TPR motif were previously designed by Regan group (Main et al., 2003) and the approach was based on statistical analysis of TPR sequences in the Pfam database. For each position in the TPR motif ratio of the percentage of occurrence of an amino acid at a given position to its percentage occurrence within the TPR sequences parameter called global propensity was calculated. Res-
Residues with the highest global propensity were placed at given position of the designed TPR motif.

In our study, consensus sequence of the TPR motif was based on sequences collected from the SMART database. Family of 413 non-redundant amino acid sequences of TPR motif were aligned with ClustalW, and amino acid frequencies at a particular position were calculated. The most frequent amino acids were used to define the 34 amino acids sequence of TPR (Fig. 1A).

The strongest preference (>40% occurrence of a particular amino acid at a given position) was observed at 10 positions: 7, 8, 11 in helix A, 20, 24, 27, 28, 30 in helix B, 15 and 32 in turns between helices (Fig. 1A). Compared to the sequence of idealized TPR motif (Main et al., 2003) we observed four amino acid sequence differences (Fig 1A, B). At positions 4 and 12 and 14 the most frequently occurring amino acid according to our analysis was leucine instead of, respectively: tryptophan, tyrosine and methionine. At position 25 we identified glutamate instead of previously reported glutamine (Main et al., 2003). Since we found almost the same percentage occurrence of leucine (26.39%) and tyrosine (25.91%) at position 4, we introduced tyrosine, expecting that hydroxyl group of the side chain could form additional hydrogen bond with surrounding residues stabilizing the protein conformation.

Since three arrayed TPRs are mostly observed in proteins and probably represent the minimal number of TPR domains required for ligand binding (D’Andrea & Regan, 2003), we constructed dTPR protein containing three repeated motifs. To increase the stability and solubility of the designed protein, we added N-capping sequence (GNS-) and solvating helix (-AEAKQNLG-NAKQKQG) at the C-terminus of arrayed TPRs, as previously reported by Main et al. (2003) (Fig. 1D, E).

As the final step in the design of dTPR, we modified the charge of the protein. The net charge of natural TPRs tends to be near zero. It was also shown, that charge neutralization of designed TPR polypeptide increased the protein stability (Cortajarena et al., 2004).
Expression and purification of dTPR

dTPR was overexpressed as a GST-tag fusion in E. coli BL21 (DE3) RIL strain. The highest protein expression yield was observed after 16h incubation at 37°C upon induction with 0.6 mM IPTG. Protein purification was performed based on affinity and size exclusion chromatography (Fig. 2A, B). The yield was about 10 mg of purified protein per 1 l of culture. The homogeneity of purified protein per 1 l of culture. The homogeneity of the expressed protein was confirmed by size-exclusion chromatography (Fig. 2A, B). The yield was about 10 mg of protein per 1 l of culture.

Within the sequence of designed TPR motif, we noticed eight negatively charged amino acids (5 × Glu and 3 × Asp) and two positively charged lysines (Fig 1A). The theoretical net charge at neutral pH of the construct composed of three TPR repeats, N-capping sequence and solvating helix is equal to −16. To neutralize the net charge of designed TPR we substituted selected glutamates or aspartates with lysines. We located these electrostatic substitutions outside the ligand binding surface: at positions 16 and 18 in all three TPR motifs we introduced lysine instead of aspartic acid and at position 2 in TPR motifs 1 and 3 we placed lysine instead of glutamic acid. The total net charge of final dTPR construct was zero.

The dTPR protein sequence differs at 20 positions as compared to the sequence designed by Main et al. (2003) or at 28 positions as compared to the charged modified TPR presented by Cortajarena et al. (2004).

Biophysical properties

Conformation of dTPR protein (containing three repeated TPR motifs with N-capping sequence and solvating helix at the C-terminus) was analyzed by circular dichroism (CD) measurements. Far-UV spectra of the designed protein showed high helical content (Fig. 3A) with typical minima at 208 and 220 nm and maximum at 190 nm. CD spectra analysis with CD2D software revealed 84% of α-helix. We also collected CD spectra of dTPR at various pHs. Even at pHs 2.0 and 11.0 the shape of CD spectrum is very similar to that at pH 7.0 (Fig. 3A) showing that the protein is extremely pH resistant.

The conformational stability of dTPR was analyzed by thermal and chemical denaturation. Using CD spectroscopy, we observed highly cooperative and reversible unfolding transitions at pHs 2.0, 3.0, 10.0 and 11.0 (Fig. 3B, Table 1). The denaturation transition at pH 7.5 could not be monitored by CD as the transition exceeded 100°C. The lowest stability, with a midpoint of 60.4°C was observed at pH 11.0 (Table 1).

We also determined the denaturation parameters of the dTPR using DSC (Fig. 3C). The thermodynamic values calculated from thermogram showed denaturation temperature 104.5°C, calorimetric enthalpy (ΔH_m) equal to 122.1 kcal/mol and van’t Hoff enthalpy (ΔH_vH) 120.8 kcal/mol. Very close values of calorimetric and van’t Hoff enthalpy indicate a two-state mechanism of denaturation of the dTPR (Privalov, 1979). Additionally, heat capacity change (ΔC_vH) was calculated directly from thermogram to be 0.85 kcal/molK. The denaturation parameters allowed us to calculate the free energy change of denaturation ΔG_m (T) from the Gibbs-Helmholtz equation: ΔG_m (T) = ΔH_vH / T - ΔC_vH(T_m-T) + T ln (T/T_m) (Becktel & Schellman, 1987). The calculated value of ΔG_m at 21°C equals to 18.1 kcal/mol that is extremely high, compared to many naturally occurring proteins (Makhatadze & Privalov, 1995).

In the next step of thermodynamic stability analysis of TPR, we performed chemical denaturation using GdmCl (Fig. 3D). We found that the protein is very resistant to denaturant, and the free energy change of denaturation in the absence of denaturant (ΔG^0_reduced) equals to 15 kcal/
Design and expression of a highly stable TPR-based protein scaffold

mol at pH 7.5, concentration of GdmCl at midpoint was 4.05 and m value equals to 3.7 kcal/mol M. Taking into account different methods of denaturation and long extrapolation of experimental data, we conclude that both DSC and chemical denaturation provide coherent values of free energy change of denaturation at pH 7.5. Compared to pH 7.5, the $\Delta G^\text{H2O}$ at pH 2.0 was lower only by 1.5 kcal/mol, concentration of GdmCl at midpoint was 4.73 M and m value $= 2.85$ kcal/mol M.

Compared with other tetratricopeptide repeat proteins the dTPR is significantly more stable. The TPR domains of natural proteins e.g: TPR domain of PP5, TPR1 of Hop, TPR2A of Hop or 3-TPR of Vpu binding protein, show $T_{\text{den}}$ about 50°C (Cortajarena & Regan, 2006). The CTPR3 protein (containing three TPR repeats) obtained in Regan laboratory has $T_{\text{den}}$ lower by 21.5°C at pH 6.3, as compared to dTPR designed by us at pH 7.5, and free energy change of unfolding lower by 5 kcal/mol (Main et al., 2003). The modified sequence of CTPR3 designed to bind Hsp90 that was “charge neutralized” (Cortajarena et al., 2004) was still at least 10°C or 3 kcal/mol less stable than our dTPR. Comparison to other designed repeated proteins the dTPR shows favorable thermodynamic properties. Highly stable variant of armadillo repeat protein shows $T_{\text{den}} = 85.5°C$ (Alfarano et al., 2012), proteins containing three ankyrin repeats selected from combinatorial library (variant E3_5) or designed by full consensus approach (NISC) denature above 90°C or even above 100°C respectively (Binz et al., 2003; Wetzel et al., 2008). Finally, a stable scaffold based on six leucine-rich repeats (dVLR), that we designed previously shows $T_{\text{den}} = 73.9°C$ at pH 6.0 (Wezner-Ptasinska et al., 2011).

Favorable biophysical properties of dTPR protein may provide substantial advantage in terms of generation of specific ligand-binding molecules. Stable scaffold could accommodate a number of mutations providing interaction with defined ligand, and still preserve its structural stability.

Acknowledgments

The work was supported by the Wroclaw Research Center EIT + under the project “Biotechnologies and advanced medical technologies – BioMed” (POIG 01.01.02-02-003/08/00) financed by the European Regional Development Fund (Operational Program Innovative Economy, 1.1.2)

REFERENCES


