Journal of Structural Biology 185 (2014) 215–222

Contents lists available at SciVerse ScienceDirect

Journal of Structural Biology

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

Structure-based non-canonical amino acid design to covalently crosslink an antibody–antigen complex

Jianqing Xu [a,1], Drew Tack [b,1], Randall A. Hughes [b,c,1], Andrew D. Ellington [b,c,*], Jeffrey J. Gray [a,*]

[a] Department of Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, MD 21218, USA
[b] Department of Chemistry and Biochemistry, Institute for Cellular and Molecular Biology, The University of Texas at Austin, Austin, TX 78712, USA
[c] Applied Research Laboratories, The University of Texas at Austin, Austin, TX 78758, USA

ARTICLE INFO

Article history:
Available online 13 May 2013

Keywords:
Computer-aided design
Structure-based design
Rosetta
Antibody
Non-canonical amino acid
Crosslinking
Binding affinity
L-DOPA

ABSTRACT

Engineering antibodies to utilize non-canonical amino acids (NCAA) should greatly expand the utility of an already important biological reagent. In particular, introducing crosslinking reagents into antibody complementarity determining regions (CDRs) should provide a means to covalently crosslink residues at the antibody–antigen interface. Unfortunately, finding the optimum position for crosslinking two proteins is often a matter of iterative guessing, even when the interface is known in atomic detail. Computer-aided antibody design can potentially greatly restrict the number of variants that must be explored in order to identify successful crosslinking sites. We have therefore used Rosetta to guide the introduction of an oxidizable crosslinking NCAA, L-3,4-dihydroxyphenylalanine (L-DOPA), into the CDRs of the anti-protective antigen scFv antibody M18, and have measured crosslinking to its cognate antigen, domain 4 of the anthrax protective antigen. Computed crosslinking distance, solvent accessibility, and interface energetics were three factors considered that could impact the efficiency of L-DOPA-mediated crosslinking. In the end, 10 variants were synthesized, and crosslinking efficiencies were generally 10% or higher, with the best variant crosslinking to 52% of the available antigen. The results suggest that computational analysis can be used in a pipeline for engineering crosslinking antibodies. The rules learned from L-DOPA crosslinking of antibodies may also be generalizable to the formation of other crosslinked interfaces and complexes.

1. Introduction

Antibodies are key components of the immune system with broad diversity to recognize a variety of antigens. Antibody-based therapeutic, diagnostic, and industrial applications frequently require antibodies having high stability and strong binding affinity. With the development of computational techniques and a number of successful experiences in protein modeling and design (Lippow and Tidor, 2007; Mandell and Kortemme, 2009), computational antibody design has begun to play an important role in predicting improvements to antibody function. Computational design of antibodies has been used to enhance binding affinity (Barderas et al., 2008; Clark et al., 2006; Lippow et al., 2007), to improve stability by improvement of thermal/aggregation resistance (Chennamsetty et al., 2009; Miklos et al., 2012), and to alter binding specificity (Farady et al., 2009), and others (Caravella et al., 2010; Kuroda et al., 2012; Midelfort et al., 2004; Pantazes and Maranas, 2010).

To date, though, most computational design methods have focused on manipulating the twenty natural proteogenic amino acids to modify molecular forces such as electrostatics (Lippow et al., 2007), hydrophobic interactions (Chennamsetty et al., 2009), hydrogen bonds (Clark et al., 2006), and salt bridges (Miklos et al., 2012). However, recent advances in engineering the translation system have now allowed for the site-specific insertion of non-canonical amino acids (NCAAs) with a variety of functionalities into proteins with good efficiency (Wang et al., 2006; Xie and Schultz, 2006). Such NCAAs can be used to improve the stability and pharmacokinetics of therapeutic proteins (Cho et al., 2011), to augment binding (Liu et al., 2009), and to provide a myriad of chemical handles to study protein structure and function (Jones et al., 2010; Tsao et al., 2006; Zhang et al., 2002).

* Corresponding authors at: Department of Chemistry and Biochemistry, Institute for Cellular and Molecular Biology, The University of Texas at Austin, Austin, TX 78712, USA (A.D. Ellington), Department of Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, MD 21218, USA (J.J. Gray).
E-mail addresses: ellingtonlab@gmail.com (A.D. Ellington), jgray@jhu.edu (J.J. Gray).
[1] These authors contributed equally.

1047-8477/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jsb.2013.05.003

The generation of protein–protein crosslinks by inserting NCAAs into proteins could prove useful for a variety of applications. To this end, a number of crosslinking-capable NCAAs have been incorporated into proteins in a site-specific manner utilizing an array of functionalized amino acids. These crosslinking functionalities include photo-crosslinkable aryl-azides (Chin et al., 2002b), benzophenones (Chin et al., 2002a) and diazirines (Ai et al., 2011) as well as the oxidizable crosslinker, L-DOPA (Alfonta et al., 2003). While any of the crosslinkers might benefit from a quantitative placement methodology, we chose L-DOPA because the periodate induced oxidation allowed for more control over the crosslinking conditions relative to photo-inducible crosslinkers that have been found to spuriously crosslink during sample handling (Chin et al., 2002b). In addition, the nucleophile-driven cross-linking mechanism of L-DOPA has been extensively characterized with a variety of proteinaceous nucleophiles (Liu et al., 2006).

L-DOPA has previously been used to successfully crosslink the monomeric domains of a dimeric sortase A for structural studies (Umeda et al., 2009), to enhance the affinity of low-affinity peptide probes for a kinase SH3 bioassay (Umeda et al., 2010), and to site-specifically label proteins with polysaccharides (Ayyadurai et al., 2011). While previously reported uses of L-DOPA as a site-specific crosslinker have yielded examples of effective crosslinking (as shown by SDS–PAGE or Western blot analyses), the actual efficiencies of crosslinking have never been reported (Burdine et al., 2004; Umeda et al., 2009, 2010). These previous reports indicated that it was possible to place L-DOPA by intuition, but did not provide more quantitative assessments of what parameters impacted crosslinking efficiency (Umeda et al., 2009, 2010).

In this paper, we explore how the Rosetta suite of computational protein design tools might be used to predict the site-specific, functional incorporation of L-DOPA into an antibody, allowing it to crosslink to its cognate antigen. A better understanding of where and how to insert crosslinking moieties into an antibody combining site could lead to the development of tools for validating antibody–antigen structural models (Pimenova et al., 2008) and to reagents capable of binding analytes with extremely high affinities and specificities Kim and Yoon (2010).

As a proof-of principle demonstration, we chose a complex with a known structure, the anti-anthrax antibody M18 bound to anthrax protective antigen (PA) (Leysath et al., 2009). PA is a component of the tripartite toxin secreted by *Bacillus anthracis* which binds to cellular receptors, and assists host cellular targeting and transport of the lethal factor (LF) and edema factor (EF) into cytoplasm (Moayeri and Leppla, 2004; Young and Collier, 2007). M18 is a neutralizing antibody (Leysath et al., 2009) derived by directed evolution from monoclonal antibody 14B7 (Harvey et al., 2004; Little et al., 1988), which binds to the fourth domain of PA (PAD4), and effectively blocks PA binding to cellular receptors such as CMG2 to mitigate anthrax toxicity.

2. Methodology

2.1. Computational methods

2.1.1. Creation of models of L-DOPA antibody mutants in complex with antigen

Models for various mutants of the antibody–antigen complex were created using Rosetta (Leaver-Fay et al., 2011) with L-DOPA placed in various positions within the antibody paratope. Coordinates for the wild-type M18-PAD4 complex were downloaded from the Protein Data Bank (Berman et al., 2000) (PDB ID 3ETB). To remove crystal packing effects and obtain a Rosetta-minimized reference structure, fixed-backbone side-chain packing and minimization (1000 decoys) on the entire protein complex was performed using Rosetta's score12. The lowest-scoring structure was used for the calculations for the predictive introduction of L-DOPA into the CDRs of M18. The Rosetta 3.4 (revision 51671, available at www.rosettacommons.org) command line used to run the "fix_bb" protocol was:

```
fixbb.linuxgccrelease -s Crystal.pdb -nstruct 1000
 -use_input_sc
 -minimize_sidechains
 -run:multiple_processes_writing_to_one_directory
 -packing:repack_only -ex1 -ex2aro
```

For each interface Lys on the antigen, neighboring antibody residues within 10 Å (C_β–C_β distance) were selected as potential mutation sites. Each antibody residue within the cutoff distance was substituted to L-DOPA individually, followed by fixed-backbone side-chain packing (20 decoys) of the nearby residues (<10 Å) to accommodate the local changes. For these and any further calculations where L-DOPA is present, Rosetta uses the molecular mechanics-based scoring function (mm_std) and associated NCAA rotamer library (Renfrew et al., 2012).

To carry out these calculations, the position of the L-DOPA mutation and the positions of the neighboring residues were specified in a "resfile", and the same "fixbb" protocol read the "resfile" and substituted the target residue to L-DOPA, followed by side chain repacking including all the neighboring residues. A Rosetta 3.4 (revision 51671) command line example is:

```
fixbb.linuxgccrelease -s Best_Prepacked.pdb
 -nstruct 20 -use_input_sc
 -resfile 315_LYS_J_679—139_SER_H_31.resfile
 -score:weights mm_std
 -minimize_sidechains -ex1 —ex2
```

2.1.2. Model relaxation with crosslink constraint

Some measures are performed on a structure where the L-DOPA is artificially constrained to be proximal to the target lysine residue. For these calculations, we employed an empirically-determined linear constraint potential,

$$E_{\text{constr}} = -100 + 400 \times |d_{XL} - 3.5|$$

where d_{XL} is the distance in Ångstroms between C_γ atom on the L-DOPA ring (the atom bound to the C_β atom) and the Lys N_ϵ atom, and E_{constr} is the constraint energy in Rosetta Energy Units (REU). The constraint weights were chosen to bring the L-DOPA and Lys into proximity, in order to evaluate interface compatibility. This constraint energy was not included in the final calculated interface score. All the neighboring residues within 10 Å (C_β–C_β distance) of the L-DOPA/Lys pair were repacked to accommodate the constraint. The constrained conformation was generated using a custom PyRosetta script with PyRosetta 2.012, revision 51671 (PyRosetta available at www.rosettacommons.org, script available upon request).

2.1.3. Crosslinking distance

After selecting the L-DOPA position and the target Lys, all the distances of potential crosslinking atom pairs (lysine N_ϵ atom and L-DOPA atoms C_2, C_5, and C_6) were evaluated, and the one with the minimum value represented the crosslinking distance (d_{XL}).