CHAPTER TWENTY-ONE

Designing Covalently Linked Heterodimeric Four-Helix Bundles

M. Chino*, L. Leone*, O. Maglio*,†, A. Lombardi*,1
*University of Napoli Federico II, Napoli, Italy
†Institute of Biostructures and Bioimages-IBB, CNR, Napoli, Italy
1Corresponding author: e-mail address: alombard@unina.it

Contents

Abstract

De novo design has proven a powerful methodology for understanding protein folding and function, and for mimicking or even bettering the properties of natural proteins. Extensive progress has been made in the design of helical bundles, simple structural motifs that can be nowadays designed with a high degree of precision. Among helical bundles, the four-helix bundle is widespread in nature, and is involved in numerous and fundamental processes. Representative examples are the carboxylate bridged diiron proteins, which perform a variety of different functions, ranging from reversible dioxygen binding to catalysis of dioxygen-dependent reactions, including epoxidation,

Methods in Enzymology, Volume 580
ISSN 0076-6879
http://dx.doi.org/10.1016/bs.mie.2016.05.036

desaturation, monohydroxylation, and radical formation. The "Due Ferri" (two-irons; DF) family of proteins is the result of a de novo design approach, aimed to reproduce in minimal four-helix bundle models the properties of the more complex natural diiron proteins, and to address how the amino acid sequence modulates their functions. The results so far obtained point out that asymmetric metal environments are essential to reprogram functions, and to achieve the specificity and selectivity of the natural enzymes. Here, we describe a design method that allows constructing asymmetric four-helix bundles through the covalent heterodimerization of two different α-helical harpins. In particular, starting from the homodimeric DF3 structure, we developed a protocol for covalently linking the two α_2 monomers by using the Cu(I) catalyzed azide–alkyne cycloaddition. The protocol was then generalized, in order to include the construction of several linkers, in different protein positions. Our method is fast, low cost, and in principle can be applied to any couple of peptides/proteins we desire to link.

1. INTRODUCTION

1.1 The Four-Helix Bundle: A Widespread Structural Motif

The four-helix bundle is a ubiquitous structural motif in nature, as it is found among a wide range of functionally diverse proteins and metalloproteins. For example, four-helix bundles are involved in the RNA-binding process (Banner, Kokkinidis, & Tsernoglou, 1987) and they are found in several proteins, such as growth hormones (De Vos, Ultsch, & Kossiakoff, 1992) and cytokines (Rozwarski et al., 1994). Numerous complex metalloproteins yet contain a simple four-helix bundle at the heart of the protein, where the metal cofactor (such as a heme, a dinuclear iron or copper site) necessary to accomplish functions is housed. A heme site is found in the electron transfer cytochrome *c′* (Weber et al., 1980) and cytochrome *b*562 (Mathews, Bethge, & Czerwinski, 1979). Diiron sites in the class of carboxylate-bridged diiron proteins are involved in dioxygen binding and activation (Lee & Lippard, 2003; Maglio, Nastri, & Lombardi, 2012). Hemerythrin and myohemerythrin (Stenkamp, 1994) reversibly bind and transport oxygen, whereas ferritins and bacterioferritins are devoted to ferroxidase activity and iron storage within the core of a polymeric four-helix bundle structure (Frolow, Kalb, & Yariv, 1994; Harrison & Arosio, 1996; Wahlgren et al., 2012). Diiron proteins also catalyze a diverse set of dioxygen-dependent reactions, including desaturation (acyl carrier Δ^9 desaturase), hydroxylation (catalytic component of bacterial monooxygenases), and radical formation (R2 subunit of ribonucleotide reductase) (Jordan & Reichard, 1998; Lindqvist, Huang, Schneider, & Shanklin, 1996; Sazinsky & Lippard,

2015; Sirajuddin & Rosenzweig, 2015). Dinuclear copper sites, housed into the interior of four-helix bundles, also play important roles in dioxygen binding and activation. Among them, hemocyanins reversibly bind dioxygen, catechol oxidase and tyrosinase further activate dioxygen for substrate hydroxylation or oxidation (Yoon, Fujii, & Solomon, 2009). Recently, a four-helix bundle protein, able to accumulate copper for particulate methane monooxygenase, was isolated from the methanotroph *Methylosinus trichosporium* OB3b (Vita et al., 2015), further expanding the repertoire of fundamental processes played by this protein scaffold. Due to its central role in Nature, numerous attempts have been made to construct artificial four-helix bundles by de novo design, not only to allow a better interpretation of the chemistry supported by the natural systems but also to develop novel proteins with programmed functions (Chino et al., 2015; Samish, MacDermaid, Perez-Aguilar, & Saven, 2011; Slope & Peacock, 2016; Yu et al., 2014).

1.2 Designing Functional Four-Helix Bundle Proteins

The four-helix bundle can be viewed as an α-helical coiled coil, which is, more generally, a super-secondary structure made up of α-helices packed together in a parallel or antiparallel orientation. Coiled coils amino acid sequences are usually described in terms of seven residues (heptad) repeats, since seven residues are present per two turns of the α-helix (Kohn & Hodges, 1998). This scaffold is very robust and thermodynamically stable, since it is able to tolerate multiple residue substitutions without disrupting the global three-dimensional fold. As a consequence, the four-helix bundle scaffold is of great interest in the field of protein design, as it represents a useful template for structure-to-function relationship analysis and for developing novel artificial metalloenzymes (Chino et al., 2015; Peacock, 2016). In principle, active site environment (first and second coordination sphere) can be modified to induce metal-binding selectivity and to finely tune the chemistry of the cofactor to achieve specific functions. This task often involves introducing asymmetry around the metal environment, thus representing a difficult challenge in the de novo design of α-helical coiled coils.

One possible strategy for developing an asymmetric four-helix bundle involves the noncovalent heterodimerization of four single α-helices or two helix–loop–helix (α_2) domains (Fig. 1A and B). This approach requires establishing a large energy gap to stabilize the desired heteromeric form respect to both homooligomeric folds, and any undesired heteromeric

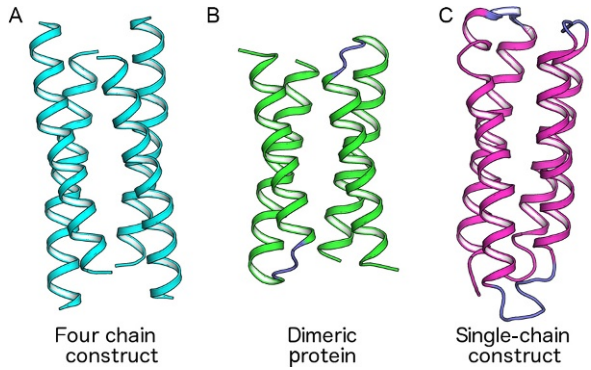

Fig. 1 Possible strategies for developing antiparallel four-helix bundles. (A) Tetramerization of four single α-helices. (B) Dimerization of two helix-loop-helix motifs. (C) Single-chain construct.

topology. Thus, the design methodology should include specific elements of both positive and negative design, to prevent alternate topologies from occurring (Grigoryan, Reinke, & Keating, 2009; Havranek & Harbury, 2003; Hill, Raleigh, Lombardi, & DeGrado, 2000). Even though the "rules" that guide oligomerization are now well established, all the interactions responsible for the pairing specificity are strictly dependent on slight variations of pH, ionic strength, and other physicochemical conditions of the environment (Fairman et al., 1996; Fry, Lehmann, Saven, DeGrado, & Therien, 2010; Marsh & DeGrado, 2002; Zhang et al., 2015). The noncovalently assembled complexes are generally not suitable for structural characterization, since it is difficult to completely avoid the presence of alternatively assembled species. On the other hand, heteromeric systems consisting of disconnected helices, which can be separately synthesized, purified, and combinatorially assembled, are well suited for the production of an array of any desired helical bundles from a significantly smaller number of peptides (Calhoun et al., 2005). A variety of de novo designed heteromeric two-stranded coiled coils (Litowski & Hodges, 2002; Thomas, Boyle, Burton, & Woolfson, 2013), three-helix (Chakraborty, Iranzo, Zuiderweg, & Pecoraro, 2012; Dieckmann et al., 1997), and four-helix bundles (Kaplan & DeGrado, 2004; Summa, Rosenblatt, Hong, Lear, & DeGrado, 2002) have been successfully developed and reported to date.

An alternative strategy to mimic the asymmetry of natural proteins in the context of designed coiled coils uses a single polypeptide chain (Fig. 1C), in which helices are connected by loops (Calhoun et al., 2003; Chakraborty et al., 2011; Smith & Hecht, 2011). Such proteins have generally

unambiguous three-dimensional structures, thus greatly facilitating structural analysis. Nevertheless, the design of large proteins requires methods that are computationally intensive. In particular, the choice of interhelical loops is crucial, since it greatly affects both the stability and flexibility of the bundle. Further, the complexity of a single-chain construct limits its applicability for catalytic screening purposes, aimed at evaluating how systematic changes in the sequence affect structure, substrate-binding, and catalytic properties.

A third exploited strategy to obtain heteromeric four-helix bundles involves the covalent binding onto a predefined molecular scaffold. Mutter and colleagues introduced the concept of template assembled synthetic proteins (TASP) (Mutter, 2013; Mutter & Tuchscherer, 1997), which have been successfully adopted as scaffold for recognition and coupling of exogenous ligands (Monien, Drepper, Sommerhalter, Lubitz, & Haehnel, 2007; Rau, DeJonge, & Haehnel, 2000; Rau & Haehnel, 1998). Following the pioneering works of Mutter and coworkers, which adopted a properly designed cyclic decapeptide as template for assembling a variety of tertiary structures (Mutter et al., 1988), Haehnel and coworkers developed modular organized proteins (MOPS), for selectively binding metal cofactors, such as heme and copper ion (Monien et al., 2007; Rau, DeJonge, & Haehnel, 1998; Rau et al., 2000; Rau & Haehnel, 1998; Schnepf, Haehnel, Wieghardt, & Hildebrandt, 2004). A suitable chemoselective synthetic strategy was developed in order to control the identity and directionality of the helical segments, and obtain the desired heteromers.

A further approach for the design of heteromers retraces the way chosen by Nature, by building side chain/side chain covalent ligation through disulfide bond formation. Several computational methods have been developed so far, for predicting which pairs of residues, once mutated to cysteines, are suitable to form a disulfide bond. The used algorithms are derived from the analysis of side chains packing preferences of cysteine pairs involved in disulfide bonds, as found in crystal structures (Burton, Oas, Fterke, & Hunt, 2000; Craig & Dombkowski, 2013; Hazes & Dijkstra, 1988; Pabo & Suchanek, 1986; Sowdhamini et al., 1989). Despite inspired by Nature, the applicability of this strategy is limited, due to the stringent geometrical requirements for disulfide bond formation.

Recently, we implemented a novel design method to obtain an asymmetric four-helix bundle through the covalent heterodimerization of two different α-helical hairpins (Chino et al., 2016). This strategy aims at realizing an easy-to-screen system in a robust covalent framework, thus merging

the advantages of using self-assembled monomers and single-chain constructs. We selected an efficient and orthogonal chemistry, to properly bind two different monomers in native conditions. In 2001 Kolb, Finn, and Sharpless published their seminal paper on the application of powerful and selective reactions to join small units through heteroatom links, and coined the term "Click Chemistry" (Kolb, Finn, & Sharpless, 2001). Several research groups explored the different applications of the Click Chemistry in numerous fields, such as drug discovery and synthesis (Galibert et al., 2010; Góngora-Benítez, Cristau, Giraud, Tulla-Puche, & Albericio, 2012; Valverde, Vomstein, Fischer, Mascarin, & Mindt, 2015) and polymer bioconjugation (Marine, Song, Liang, Watson, & Rudick, 2015; Rachel & Pelletier, 2016; Shu, Tan, DeGrado, & Xu, 2008). In particular, the use of Cu(I)-catalyzed azide–alkyne cycloaddition (CuAAC) has been largely employed as amide bond surrogate in the generation of α-helical and β-turn pseudopeptides (Beierle et al., 2009; Horne, Yadav, Stout, & Ghadiri, 2004), in TASP based molecular assemblies (Avrutina et al., 2009), as well as in strategies of peptide stapling, a macrocyclization process, where an intramolecular linkage is introduced to constrain the peptide in the desired α-helical conformation (Jacobsen et al., 2011; Lau, Wu, de Andrade, Galloway, & Spring, 2015; Scrima et al., 2010). Finally, it is also notable the work of Kolmar and coworkers (Empting et al., 2011), who constructed a triazole bridge, in replacement of a disulfide bond, employing the Ru(II)-catalyzed azide–alkyne cycloaddition (RuAAC).

The wide range of Click Chemistry applications prompted us to test this reaction in the selective intermolecular chemical ligation of two functionalized α_2 peptides, to generate heterodimeric proteins. One of the major advantages of the Click Chemistry-based methodology relies on its orthogonality to the chemistry involved in solid-phase peptide synthesis. Once chosen the binding positions, the peptides to be ligated can be easily functionalized during the synthetic step, by introducing noncanonical amino acids bearing the azide and alkyne moieties in their sequences (Fig. 2). CuAAC provides a simple to perform coupling process, leading to a thermally and hydrolytically stable triazole connection between the

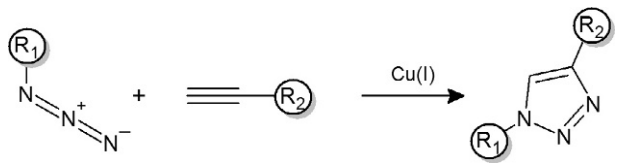

Fig. 2 Copper catalyzed 1,3-dipolar cycloaddition.

peptides. Moreover, the triazole ring of the linker could be introduced as a ligand in the metal-binding site.

Finally, given the large number of commercially available azide and alkyne building blocks, the designer can finely control the length and flexibility of the linker by choosing different pairs of functionalized amino acids.

In this chapter, we describe the developed computational protocol, first applied to the DF3 structure, as a specific case study. As logical extension of the method, a more general protocol is also reported, aimed at including the construction of several linkers, in different protein positions, and at fulfilling as many as possible designer needs.

2. SELECTION OF THE BEST DOCKING HOTSPOT GIVEN A PREDEFINED ANCHOR BOLT

In this section, we define a method for the rational design of a covalent attachment between the two subunits of the de novo designed DF3 protein. DF3 is made up of two identical 48-residue helix–loop–helix (α_2) motifs, able to specifically self-assemble into an antiparallell four-helix bundle, in the presence of metal ions (Faiella et al., 2009). The diiron form of the DF3 protein is able to perform phenol oxidase activity, rivaling natural counterparts in terms of catalytic efficiency. In order to move from oxidase to monooxygenase activity, a careful observation of natural monooxygenases structures, such as methane and toluene monooxygenases, points out that symmetry is broken in proximity of the active site (Friedle, Reisner, & Lippard, 2010). Starting from this observation, we developed a new asymmetric family of DF compounds, named DF-Click (Chino et al., 2016). In order to accomplish this task, we adopted a protocol for the design and synthesis of a covalent linkage between the two α_2 subunits based on Click Chemistry. This allowed us to generate a hybrid variant between a self-assembled heteromer and a single-chain protein, preserving the pros of dimeric derivatives, ie, simplified synthesis and structural data interpretation, ability to quickly generate several compounds for catalytic screening, and to finely tune the active site properties.

Here, we will first trace the steps leading to DF-Click, rationalizing, and elucidating the design process. In the next section, we will generalize our approach both in terms of different linkers and binding positions.

In the first member of DF-Click family, we used propargyl glycine (Pra) and 6-azidohexanoic acid (6aha) as the alkyne and the azide moiety, respectively, for the Click reaction. 6aha is often used to functionalize amino

groups (Witte et al., 2013); in DF-Click family, it has been used as N-capping reagent of one subunit, replacing the N-terminal acetyl group. Once located the azide moiety, the next step required searching for the best position to mutate to Pra residue, in order to obtain efficient triazole formation, as well as the lowest perturbation of the global protein fold. This can be obtained when: (a) the interresidue distance is comparable with the distance spanned by the linker; (b) the conformation adopted by the linker is energetically favorable, otherwise the bundle could be strained to allow the linker to reach a more stable conformation. One possible way to address these requirements is to compare the geometrical parameters of the linker with those calculated for each possible binding position on the protein. A key step is to define a suitable description of the binding geometry, as well as to consider all the most favorable conformations of the linker. To accomplish these tasks, we define three binding parameters, which are calculated both for the protein and for the linker. Then, we compare these parameters to find the best candidates for synthesis. Two pivot bonds are chosen to describe the binding geometry, the C1–C2 bond of 6aha and the Cα–Cβ bond of Pra, which are compared to the C–C_{methyl} bond of the terminal acetyl group and the Cα–Cβ bond of each residue from the other subunit, respectively. The geometrical parameters, illustrated in Fig. 3, are: (1) d, the C–C pivotal distance; (2) θ, the angle described by the first pivot bond and the first atom of the second pivot; (3) θ', the angle described by the second pivot bond and the first atom of the first pivot.

Upon generation of the "clicked" model, we evaluate the designed linker conformations in terms of RMSD from the energetically favorable starting conformations. A detailed step-by-step procedure will clarify this general approach.

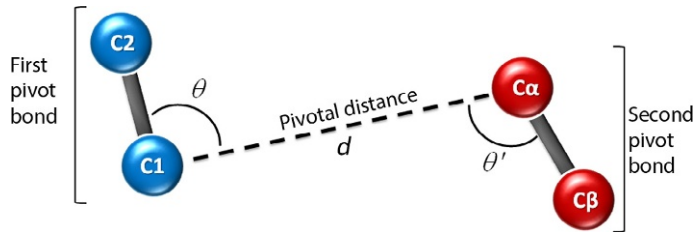

Fig. 3 Geometrical parameters calculated for the linker and for any pair of binding positions on the protein. For the linker, the first and the second pivot are the C1–C2 bond of 6aha and Cα–Cβ bond of Pra, respectively. For the protein, the first pivot is the C–C_{methyl} bond of N-terminal acetyl group, and the second pivot is the Cα–Cβ bond of each selected residue from the other subunit.

2.1 Structure Preparation of the Target Protein

1. Retrieve the coordinates of the protein structure that you wish to adopt as a template. You can choose either an X-ray or a NMR structure, since no limitation is imposed by the method. In the protocol described in the following, we used the NMR model of the di-Zn-DF3, which can be downloaded from PDB (PDB ID: 2kik). We limited the search to the first model of the NMR bundle for the sake of brevity; in principle, the protocol can be adopted for each model of the bundle. It is worth to say that performing the protocol for multiple NMR models may result in a more exhaustive search, since the analysis on alternative conformations could result in a higher number of hotspots.
2. Remove water molecules, if present.
3. Add hydrogen atoms to the structure using the preferred software. In our case, we used Accelrys Discovery Studio 3.0 (DS3) (Accelrys Software Inc., 2012).
4. Save the structure as pdb format, and open it in PyMOL (DeLano, 2002).

2.2 Structure Preparation of the Linker

1. Manually generate the linker coordinates. Specifically, combine 6aha and Pra in the triazole form of the linker. We used DS3 to perform this task.
2. Perform a fast minimization cycle to clean the linker geometry. We adopted a DREIDING force field (Mayo, Olafson, & Goddard, 1990) to perform the minimization. DS3 uses a predefined function, activated by clicking *Structure*|*Clean Geometry* in the menu bar.
3. Save the linker coordinates in sdf format.
4. Submit this file to the Frog2 web server (http://bioserv.rpbs.univ-paris-diderot.fr/services/Frog2/; Miteva, Guyon, & Tufféry, 2010) to generate the 3D conformation ensemble, which will be analyzed in Sections 2.3 and 2.4. In our example, 50 conformations have been generated in pdb format, by imposing a minimization cycle for each of them.
5. Open the structural ensemble in PyMOL. Fig. 4 shows the Frog2 output we used for this protocol.

2.3 Performing the Geometrical Parameter Calculations

This protocol relies on the evaluation of three geometrical parameters, which have been calculated for: (i) each acetyl/residue pair of the protein and (ii) each linker conformation. Two PyMOL scripts have been produced

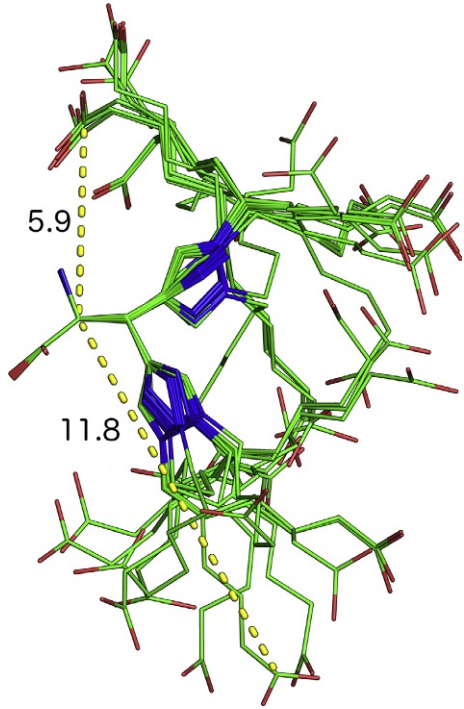

Fig. 4 Bundle of the 50 conformers generated by Frog2 (Miteva et al., 2010) for the Pra/6aha linker, showing the minimum and the maximum distance spanned by this linker. Conformers are fitted toward Pra backbone, for clarity.

to perform the calculation. They can be copied and saved as .py script files and run under the PyMOL environment with the command "*run/path/to/script.py*."

A first script (script 1 reported below) has been used for the protein; it generates two files: "distca.txt" containing the pivotal distances d for each residue, and "angcbcaca.txt" containing θ, θ' angles.

```
#script 1 starts here
from pymol import cmd
#get the model coordinates for the 4 atoms. caA and cbA are the acyl
carbon and the methyl carbon of acetyl residue of the A subunit. caB
and cbB are the alpha and the beta carbon coming from B subunit.
caA = cmd.get_model("(n. C and c. A and resn ace)")
caB = cmd.get_model("(n. CA and c. B)")
cbA = cmd.get_model("(n. CH3 and c. A and resn ace)")
cbB = cmd.get_model("(n. CB+HA2 and c. B)")
```

```
#this is to generate a file with caA-caB distances
outFile = open("./distca.txt", 'w')
for atA in caA.atom:
      for atB in caB.atom:
            outFile.write( "%s %s %s %s %s\n" %(atA.resn, str(atA.
resi), atB.resn, str(atB.resi), str(cmd.get_distance("(id %s)"
% atA.id, "(id %s)" % atB.id))))
outFile.close()
#this is to generate a file for theta and theta prime.
outFile = open("./angcbcaca.txt", 'w')

outFile.write("CHAIN A (THETA)\n")

for acaA in caA.atom:
       for acbA in cbA.atom:
             if acaA.resi == acbA.resi:
                   for acaB in caB.atom:
                         outFile.write( "%s %s %s %s %s\n" %(acaA.
resn, str(acaA.resi), acaB.resn, str(acaB.resi), str(cmd.get_angle
("(id %s)" % acbA.id, "(id %s)" % acaA.id, "(id %s)" % acaB.id))))

outFile.write("CHAIN B (THETA')\n")

for acaA in caA.atom:
      for acaB in caB.atom:
            for acbB in cbB.atom:
                  if acaB.resi == acbB.resi:
                        outFile.write( "%s %s %s %s %s\n" %(acaA.
resn, str(acaA.resi), acaB.resn, str(acaB.resi), str(cmd.get_angle
("(id %s)" % acbB.id, "(id %s)" % acaB.id, "(id %s)" % acaA.id))))
outFile.close()
```

A second script (script 2 reported below) has been used for the linker.

```
#script 2 starts here
from pymol import cmd

#Before starting the script, generate four selections in pymol GUI, as
follows: sele1 (C1 of 6aha), sele2 (C2 of 6aha), sele3 (CA of Pra), sele4
(CB of Pra).
```

```
outFile = open("./distca_linker.txt", 'w')
for i in range(1,51):
     outFile.write("%s %s\n" %(str(i),
cmd.get_distance("sele1","sele3",int(i))))
outFile.close()

outFile = open("./theta_linker.txt", 'w')
for i in range(1,51):
     outFile.write("%s %s %s\n" %(str(i),
cmd.get_angle("sele2","sele1","sele3",int(i)),
cmd.get_angle("sele4","sele3","sele1",int(i))))
outFile.close()
```

Also this script generates two files "distca_linker.txt" and "theta_linker.txt," containing the three selected geometrical parameters (d, θ, and θ', illustrated in Fig. 3). Table 1 reports the statistics of the d parameter calculated for the linker ensemble. It is evident that the linker is able to span a wide range of distances from 5.9 to 11.8 Å, with a mean value of 9.0 Å. The agreement between residue pairs and linker conformations will be evaluated in Section 2.4.

2.4 Data Analysis

Each subunit of DF3 is composed of 48 residues, and for each of them the three geometrical parameters (d, θ, and θ') have been calculated. Each triad of parameters has to be compared with the 50 triads calculated for the linker, for a total of 2400 deviations to be analyzed. Before generating this huge amount of data, it is appropriate to filter only for those residues that fall in the range of distances spanned by the linker (5.9–11.8 Å). Sorting the DF3 residues of one subunit according to the d-value (ie, the distance between each residue of one subunit and the N-terminal acetyl of the other subunit) results in identifying only three residues, within the maximum distance spanned by the Pra/6aha linker: Ala20 (11.5 Å), Tyr23 (10.9 Å), and Thr24 (12.0 Å). For these three residues, deviations for d, and θ, θ' parameters have been calculated as follows:

Table 1 Descriptive Statistics for the Conformers of the Pra/6aha Linker

Total number of conformers	Minimum Distance (Å)	Maximum Distance(Å)	(Mean Distance ± Standard Deviation) (Å)
50	5.9	11.8	9.0 ± 1.5

$$\mathrm{dev}(d) = d_{\mathrm{linker}} - d_{\mathrm{protein}}$$

$$\mathrm{rmsd}(\theta) = \sqrt{\left(\theta_{\mathrm{linker}} - \theta_{\mathrm{protein}}\right)^2 + \left(\theta'_{\mathrm{linker}} - \theta'_{\mathrm{protein}}\right)^2}$$

Table 2 shows the five best conformers in terms of dev(d) for the three evaluated binding positions. Tyr23 gives the best fitness with the linker, in terms of both distance and angles, resulting, for the conformer 28 (numbered as in the Frog2 output), in a dev(d) equal to zero and in a rmsd(θ) of only ~6°. In a refined protocol, it would be desirable to define a threshold for the calculated deviations to discard all the unproductive pairs. However, before taking a general rule, best models for each of the three residue positions have to be evaluated.

2.5 Generation of the Best Candidates for Click Reaction

DS3 has been used to generate the three desired structural models. The procedure described in the following has been performed with the implemented superimposition routine, even though other docking software could be adopted.

Table 2 Deviations of d and θ Parameters for the Five Best Conformers of the Pra/6aha Linker Respect to the Three Selected Binding Positions

Binding Position	Conformer n°	dev(d) (Å)	rmsd(θ) (°)
A20	10	0.2	35.8
A20	20	0.3	48.9
A20	28	–0.6	67.5
A20	27	–0.6	40.9
A20	46	–0.6	42.6
Y23	28	0	5.6
Y23	27	0	25.8
Y23	46	0	23.8
Y23	25	–0.1	6.3
Y23	37	–0.1	8.7
T24	20	–0.2	56.3
T24	10	–0.3	43.8
T24	28	–1.1	71.3
T24	27	–1.1	48.9
T24	46	–1.1	50.2

1. Load the protein structure and the best linker conformer in the same Molecule Window. We loaded conformer 10 for Ala20 hotspot, conformer 28 for Tyr23, and conformer 10 for Thr24 (see Table 2).
2. Add tethers between corresponding atoms, giving the command *Structure|Superimpose|Add Tether* (Fig. 5A). We tethered the main chain atoms of Pra and the carboxyl atoms of 6aha.
3. Set rotatable bonds on the linker, giving the command *Structure|Superimpose|Add Rotatable Bonds* (Fig. 5B). We set all bonds between the two pivotal bonds as free.
4. Perform superimposition with flexible torsions, giving the command *Structure|Superimpose|Molecular Overlay*. Fig. 5C displays the settings adopted to perform the superimposition.
5. Repeat the step 4 until the superimposed linker converges to a final invariable conformation.

2.6 Evaluation of the Best Candidates for Synthesis

In DS3, the molecular overlay with flexible torsions relies on a non-deterministic algorithm; thus, the final results may depend on the starting conformation of the linker adopted. For this reason, careful analysis of the resulting models should be made. In particular, the models could present some forbidden dihedrals, or the final superimposition could be not satisfactory as the tethered atoms are too far from the desired positions.

We found out that the best and more reproducible results are obtained when the final linker conformation is close to the starting one. Thus, the final linker models can be classified according to the RMSD respect to the starting conformer coordinates. This ranking has the double advantage of checking the superimposition task, as well as the goodness of the resulting conformation, since the starting conformers can be considered as minima in the energy landscape of the linker. Table 3 reports the deviations with respect to the starting conformations for the three models as obtained in Section 2.5. Comparing the RMSD values among the three models, it results that the linker bound in position 23 gives the lowest RMSD. A careful inspection of the three models (Fig. 6), further confirms the goodness of the resulting linker structures: the best model in terms of RMSD does not present any violation in the dihedrals.

We have already successfully synthesized a covalent heterodimeric DF analog, named DF-Click1, using the best linker described earlier (Chino et al., 2016), corroborating the design process. Given the success of this

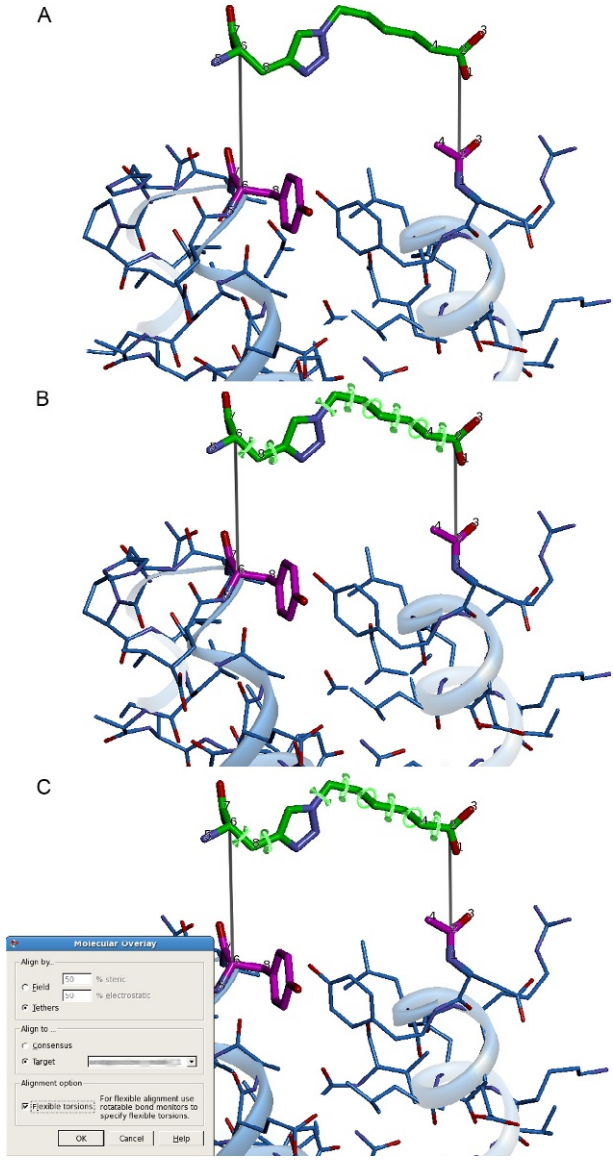

Fig. 5 Discovery Studio steps for flexible alignment between DF3 (Y23/N-terminal acetyl pair) and Pra/6aha linker with rotatable bonds. (A) Tether assignments for the defined fixed atom set. (B) Definition of the rotatable bonds of the linker molecule. (C) Molecular overlay settings.

Table 3 RMSD Values Calculated for the Conformations of the Pra/6aha Linker Before and After Flexible Alignment

Binding Position	Starting Conformer	RMSD (Å)
A20	10	1.08
Y23	28	0.91
T24	20	1.42

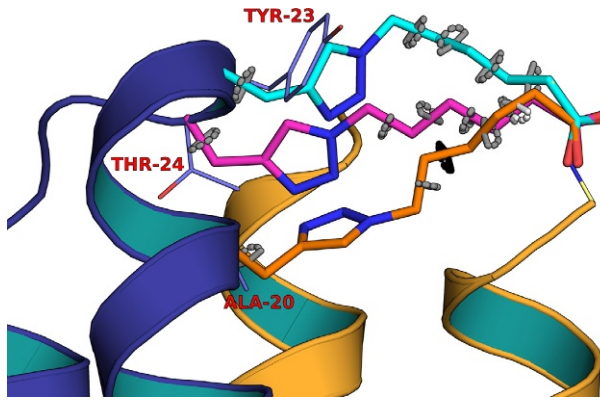

Fig. 6 Stick representation of the three possible docking hotspot positions of Pra/6aha on DF3: A20, Y23, and T24. Eclipsed torsions are highlighted in *black*, gauche torsions in *white*, *trans* torsions in *gray*.

approach, we have generalized the method, giving the designer the possibility to select both docking hotspots, and a wider set of linker options.

3. BROADENING THE HOTSPOT AND LINKER SELECTIONS

Keeping fixed one docking position (in our case the N-terminus of one subunit) limits the search output to only those positions that are in proximity of the chosen hotspot. This option greatly simplifies the design process; however, it may result reductive, since it may not be generally applicable to any protein scaffold of interest. Furthermore, one may be interested in stapling specific positions at a predefined distance, narrowing down the choice of suitable linkers. To meet these requirements, a generalized method has been defined (Fig. 7), which allows designing the link between any given

residue pair and to adopt any linker combination. Three linker combinations have been considered to demonstrate the goodness of this approach. Covalently linked models will be then generated and evaluated, and all the steps for their design will be discussed.

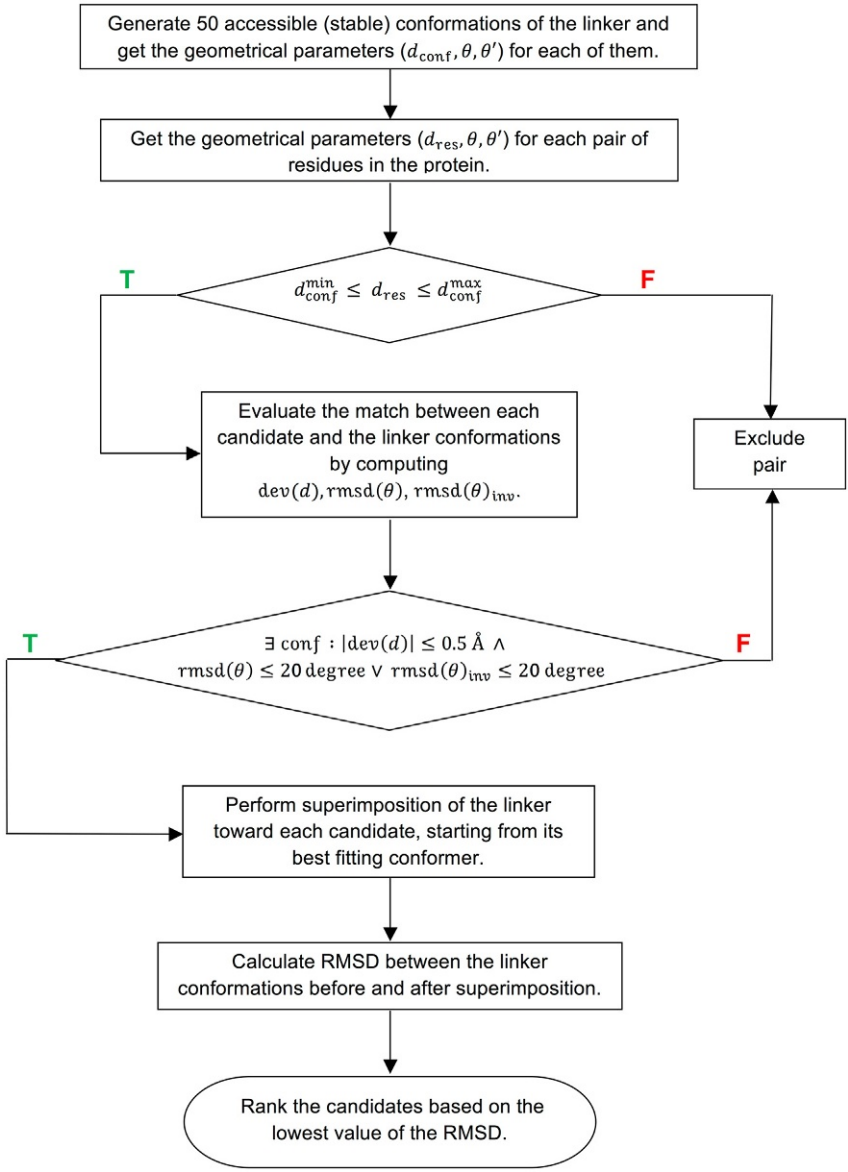

Fig. 7 General method flowchart.

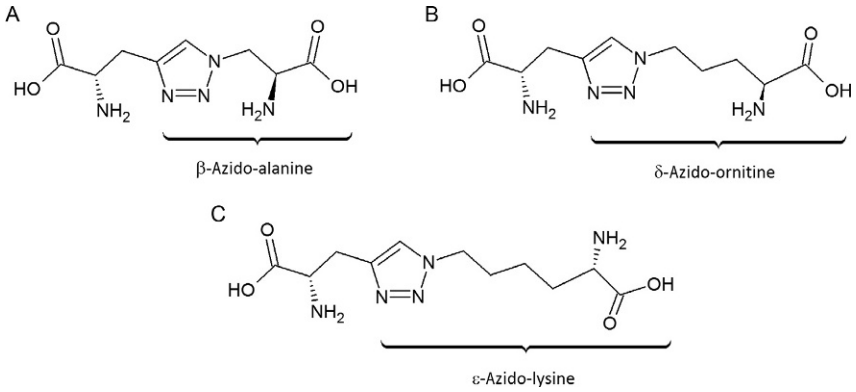

Fig. 8 Chemical structures of the selected linkers. Each of them is composed by a Pra residue and an azido-amino acid of different side chain length. (A) β-Azido-alanine, (B) δ-Azido-ornitine, (C) ε-Azido-lysine.

3.1 Structure Preparation of Any Target Protein

For any protein we wish to use as template, the structure preparation can be performed following the steps described in Section 2.1

3.2 Structure Preparation of the Linker Library

Three linker combinations have been chosen: Pra/β-azido-alanine (azido-Ala), Pra/δ-azido-ornitine (azido-Orn), and Pra/ε-azido-lysine (azido-Lys); their structures are shown in Fig. 8. These linkers have been chosen as they offer a wide range of distances between the binding residues. Steps to generate the set of conformations for each of them are reported in Section 2.2.

3.3 Performing the Geometrical Parameter Calculations for Each Residue Pair

A PyMOL script (script 3) has been developed to calculate the geometrical parameters, when the two pivotal bonds are both Cα–Cβ bonds, each from a different subunit.

```
#script 3 starts here
from pymol import cmd
#get the model coordinates for the 4 atoms. caA and cbA are the alpha
carbon and the beta carbon in the A subunit. caB and cbB are the alpha
and the beta carbon in the B subunit.
caA = cmd.get_model("(n. CA and c. A)")
caB = cmd.get_model("(n. CA and c. B)")
cbA = cmd.get_model("(n. CB+HA2 and c. A)")
```

```
cbB = cmd.get_model("(n. CB+HA2 and c. B)")
#this is to generate a file with caA-caB distances
outFile = open("./distca.txt", 'w')
for atA in caA.atom:
      foratB in caB.atom:
            outFile.write( "%s %s %s %s %s\n" %(atA.resn, str(atA.
resi), atB.resn, str(atB.resi), str(cmd.get_distance("(id %s)"
% atA.id, "(id %s)" % atB.id))))
outFile.close()
#this is to generate a file for theta and theta prime.
outFile = open("./angcbcaca.txt", 'w')

outFile.write("CHAIN A (THETA)\n")

for acaA in caA.atom:
      for acbA in cbA.atom:
            if acaA.resi == acbA.resi:
                  for acaB in caB.atom:
                        outFile.write( "%s %s %s %s %s\n" %(acaA.
resn, str(acaA.resi), acaB.resn, str(acaB.resi), str(cmd.get_angle
("(id %s)" % acbA.id, "(id %s)" % acaA.id, "(id %s)" % acaB.id))))
outFile.write("CHAIN B (THETA')\n")

for acaA in caA.atom:
      for acaB in caB.atom:
            for acbB in cbB.atom:
                  if acaB.resi == acbB.resi:
                        outFile.write( "%s %s %s %s %s\n" %(acaA.resn,
str(acaA.resi), acaB.resn, str(acaB.resi), str(cmd.get_angle("(id
%s)" % acbB.id, "(id %s)" % acaB.id, "(id %s)" % acaA.id))))
outFile.close()
```

The above reported script generates two files with the three geometrical parameters (d, θ, and θ') for each residueA–residueB pair. Given the great amount of data, we suggest ordering the script result in any spreadsheet software.

For each linker ensemble, the same script reported in Section 2.3 can be adopted to calculate the geometrical parameters for the linker. Before running the script, it is needed to create in the PyMOL GUI four selections with the four pivotal atoms.

3.4 Data Analysis

Table 4 summarizes the linker distances of the three ensembles, adopted in this procedure. Only pairs of residues whose Cα–Cα distance falls in the range spanned by the linkers have been taken into account, and for each of them the scores defined in Section 2.4 have been calculated. In this case, also a third score has been calculated, which considers the switch between the alkyne and the azide moieties:

$$\mathrm{rmsd}(\theta)_{\mathrm{inv}} = \sqrt{\left(\theta_{\mathrm{linker}} - \theta'_{\mathrm{protein}}\right)^2 + \left(\theta'_{\mathrm{linker}} - \theta_{\mathrm{protein}}\right)^2}$$

Only residue pairs showing at least one linker conformer meeting the required deviations have been considered as candidates for the design step. In particular, only those conformers whose dev(d) value was equal or lower than 0.5 Å and at least one of rmsd(θ) or rmsd(θ)$_{\mathrm{inv}}$ was lower than 20 degree have been selected. These thresholds have been adopted taking into account the results obtained from the previous designs. Further, since the design protocol does not consider backbone flexibility (Butterfoss & Kuhlman, 2006), it is not suggested to narrow down these thresholds. The numbers of candidate pairs for each linker, resulting from this analysis, are reported in Table 5. Thresholds filtered out 67% of candidates on average, resulting in less than 20 candidates for the Pra/azido-Lys linker. Tables 6–8 display the scores of the best matching conformation for each candidate pair of

Table 4 Descriptive Statistics for the 50 Conformers of Each of the Selected Linkers

Linker	Minimum Distance (Å)	Maximum Distance (Å)	Mean Distance ± Standard deviavtion (Å)
Pra/azido-Ala	5.9	7.0	6.6 ± 0.4
Pra/azido-Orn	6.8	9.4	8.5 ± 0.8
Pra/azido-Lys	5.4	10.7	8.2 ± 1.2

Table 5 Number of Residue Pairs Whose Geometrical Parameters Match with the Low-Energy Conformations of Each Linker

Linker	Matching Pairs (Evaluated Pairs)
Pra/azido-Ala	2 (4)
Pra/azido-Orn	10 (42)
Pra/azido-Lys	19 (81)

Table 6 Scores of the Best Fitting Conformation of the Pra/azido-Ala Linker for Each of the Selected Candidate Pairs of Binding Residues

Binding Positions	Best Conformer	dev(d) (Å)	rmsd(θ) (°)	rmsd(θ)$_{inv}$ (°)
V28–W42	50	0.1	12.9	23.9
I32–H39	6	0.3	1.9	15.7

Table 7 Scores of the Best Fitting Conformation of the Pra/azido-Orn Linker for Each of the Selected Candidate Pairs of Binding Residues

Binding Positions	Best Conformer	dev(d) (Å)	rmsd(θ) (°)	rmsd(θ)$_{inv}$ (°)
D35–H39	1	0.2	16.6	18.0
K31–W42	44	0.4	15.4	17.3
Y2–Q16	41	–0.2	18.6	14.8
I32–K38	46	–0.2	11.5	12.8
L6–A20	8	–0.2	16.4	14.6
N26–I46	10	0.2	10.8	18.2
K31–H39	46	–0.3	17.1	16.0
T24–I46	50	0.4	18.5	39.7
E36–E36	46	–0.5	16.4	16.8
Y2–A20	50	0.3	44.2	19.6

Table 8 Scores of the Best Fitting Conformation of the Pra/azido-Lys Linker for Each of the Selected Candidate Pairs of Binding Residues

Binding Positions	Best Conformer	dev(d) (Å)	rmsd(θ) (°)	rmsd(θ)$_{inv}$ (°)
D35–H39	40	–0.1	11.5	19.6
K31–W42	40	0.1	21.4	13.9
V28–T45	3	0.0	13.7	55.9
T24–L43	23	–0.1	19.5	67.4
E5–Q16	12	–0.3	38.1	17.1
K31–D35	33	0.0	28.6	13.7
Y2–Q16	38	0.1	7.2	11.1
I32–L43	34	–0.3	16.4	43.4

Continued

Table 8 Scores of the Best Fitting Conformation of the Pra/azido-Lys Linker for Each of the Selected Candidate Pairs of Binding Residues—cont'd

Binding Positions	Best Conformer	dev(d) (Å)	rmsd(θ) (°)	rmsd(θ)$_{inv}$ (°)
K31-K38	12	–0.4	17.8	7.2
L3-T24	33	0.4	15.0	5.4
Y2-A20	34	–0.3	8.7	35.4
A20-L43	12	–0.3	13.6	26.2
E10-G13	22	0.1	4.3	10.1
E10-I32	30	0.3	42.3	7.5
L6-Q16	34	–0.2	40.0	13.1
T24-I46	31	–0.2	30.9	5.2
Y2-I19	30	0.3	49.1	6.4
I32-W42	34	0.2	35.0	13.8
K31-H39	22	–0.4	21.3	7.7

residues. The reported conformers have been chosen to perform superimposition in the design step.

3.5 Generation of the Best Candidates for the Identified Residue Pairs

The desired models can be generated following the steps described in Section 2.5, with the only exception that in this case tethers for the superimposition can be imposed in the main chain for both ends of the linker.

3.6 Evaluation of the Best Models Amenable for the Selected Linkers

As discussed in Section 2.6, RMSD between the starting and the modeled linker coordinates can be adopted to rank the designed models. For the sake of brevity, Table 9 summarizes the results for the two best models for each designed linker. The average RMSD value of 1.1 Å is in line with the values obtained from the previous designs (see Table 3). The three best designs, one for each linker, are shown in Fig. 9. As expected by the RMSDs, the Pra/azido-Lys linker gives the best designed structure in terms of dihedrals, even though all of them may be considered as good candidates for synthesis.

Table 9 RMSD Values Calculated for the Conformations Before and After Flexible Alignment, for Each of the Three Linkers

Linker	Binding Positions	Starting Conformer	RMSD (Å)
Pra/azido-Ala	H39-I32′	6	1.22
Pra/azido-Ala	V28-W42′	3	1.35
Pra/azido-Orn	Y2-A20′	50	1.07
Pra/azido-Orn	L6-A20′	1	1.25
Pra/azido-Lys	V28-T45′	3	0.65
Pra/azido-Lys	T24-L43′	23	1.26

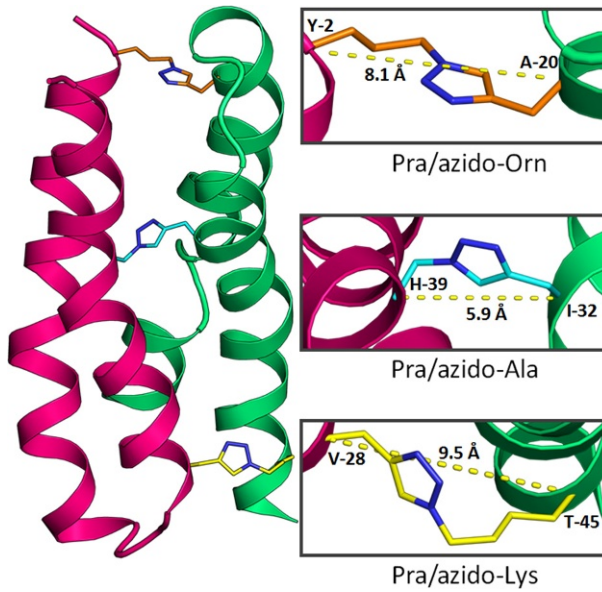

Fig. 9 Representation of the three best linker models (in terms of RMSD), along the bundle structure. *Inlets* show the details for each designed linker.

4. CONCLUDING REMARKS

In this chapter, we have described the steps leading to the generation of heterodimeric DF proteins through the rational design of covalent linkage between the two subunits. Asymmetrization of the active site in DF proteins has been already achieved in the tetrameric construct by self-assembly

(Kaplan & DeGrado, 2004) and in the single-chain construct (Reig et al., 2012), resulting in the modulation of the catalytic properties. The presented design methodology fills the gap between these two extremes, as it allows designing asymmetric models in the framework of the dimeric constructs, by means of a very simple and reliable approach. Oligomerization of two or more smaller subunits is frequently preferred in Nature to achieve complex protein structures, as supermolecular assembly is relatively simple and economical (Boersma & Roelfes, 2015). The protocol here described keeps the advantages given by the oligomerization of small subunits by fusing them in rationally designed positions that do not alter the global folding of the four-helix bundle. These advantages have been proven particularly remarkable in the study of the first DF-Click analog, as it showed complete reduction of dioxygen, coupled to the oxidation of a phenolic substrate leading to only one specific product (Chino et al., 2016).

The described linking moieties are based on the widely adopted CuAAC, which holds the advantage to be orthogonal to peptide chemistry; nonetheless this method can be efficiently applied to any class of linker, widening the applicability of the methodology. We kept this method as simple as possible, largely using simple scripts, with easily accessible softwares. This is particularly favorable as this protocol does not rely specifically on the four-helix bundle scaffold, and it can be freely applied to covalently link any protein/protein interface through a structure-based approach.

It is worth to say that there is still room for further improvements to make the methodology more accessible and reliable. In the actual implementation, this methodology does not consider explicitly backbone flexibility. One possible way to circumvent this limitation is to apply the method to a previously generated ensemble of protein structures, or to evaluate the protein flexibility of the finally designed dimers through molecular dynamics simulations. In the next future, we aim to integrate all the design tasks in an open-source environment, with the creation of a freely accessible web server.

ACKNOWLEDGMENTS

We wish to thank Flavia Nastri for critically reading the manuscript, and Fabrizia Sibillo for helping with editing.

The STRAIN PROJECT (POR Campania FSE 2007/2013 CUP B25B0900000000), which has provided a postdoctoral fellowship to M.C., and the Cost Action CM1003 (Biological Oxidation Reactions—Mechanisms and Design of New Catalysts) are kindly acknowledged.

REFERENCES

Accelrys Software Inc. (2012). *Discovery Studio, Release 3.0*. San Diego: Accelrys Software Inc.

Avrutina, O., Empting, M., Fabritz, S., Daneschdar, M., Frauendorf, H., Diederichsen, U., et al. (2009). Application of copper(i) catalyzed azide–alkyne [3+2] cycloaddition to the synthesis of template-assembled multivalent peptide conjugates. *Organic & Biomolecular Chemistry*, *7*(20), 4177–4185.

Banner, D. W., Kokkinidis, M., & Tsernoglou, D. (1987). Structure of the ColE1 rop protein at 1.7 A resolution. *Journal of Molecular Biology*, *196*(3), 657–675.

Beierle, J. M., Horne, W. S., van Maarseveen, J. H., Waser, B., Reubi, J. C., & Ghadiri, M. R. (2009). Conformationally homogeneous heterocyclic pseudotetrapeptides as three-dimensional scaffolds for rational drug design: Receptor-selective somatostatin analogues. *Angewandte Chemie International Edition*, *48*(26), 4725–4729.

Boersma, A. J., & Roelfes, G. (2015). Protein engineering: The power of four. *Nature Chemistry*, *7*(4), 277–279.

Burton, R. E., Oas, T. G., Fterke, C. A., & Hunt, J. A. (2000). Novel disulfide engineering in human carbonic anhydrase II using the PAIRWISE side-chain geometry database. *Protein Science*, *9*(4), 776–785.

Butterfoss, G. L., & Kuhlman, B. (2006). Computer-based design of novel protein structures. *Annual Review of Biophysics and Biomolecular Structure*, *35*(1), 49–65.

Calhoun, J. R., Kono, H., Lahr, S., Wang, W., DeGrado, W. F., & Saven, J. G. (2003). Computational design and characterization of a monomeric helical dinuclear metalloprotein. *Journal of Molecular Biology*, *334*(5), 1101–1115.

Calhoun, J. R., Nastri, F., Maglio, O., Pavone, V., Lombardi, A., & DeGrado, W. F. (2005). Artificial diiron proteins: From structure to function. *Peptide Science*, *80*(2–3), 264–278.

Chakraborty, S., Iranzo, O., Zuiderweg, E. R. P., & Pecoraro, V. L. (2012). Experimental and theoretical evaluation of multisite cadmium(ii) exchange in designed three-stranded coiled-coil peptides. *Journal of the American Chemical Society*, *134*(14), 6191–6203.

Chakraborty, S., Yudenfreund Kravitz, J., Thulstrup, P. W., Hemmingsen, L., DeGrado, W. F., & Pecoraro, V. L. (2011). Design of a three-helix bundle capable of binding heavy metals in a triscysteine environment. *Angewandte Chemie International Edition*, *50*(9), 2049–2053.

Chino, M., Leone, L., Maglio, O., Pavone, V., Nastri, F., & Lombardi, A. (2016). *Enhancing the selectivity of* de novo *due ferri proteins through heterodimer formation*. Manuscript in preparation.

Chino, M., Maglio, O., Nastri, F., Pavone, V., DeGrado, W. F., & Lombardi, A. (2015). Artificial diiron enzymes with a de novo designed four-helix bundle structure. *European Journal of Inorganic Chemistry*, *2015*(21), 3371–3390.

Craig, D. B., & Dombkowski, A. A. (2013). Disulfide by Design 2.0: A web-based tool for disulfide engineering in proteins. *BMC Bioinformatics*, *14*, 346.

De Vos, A. M., Ultsch, M., & Kossiakoff, A. A. (1992). Human growth hormone and extracellular domain of its receptor: Crystal structure of the complex. *Science*, *255*(5042), 306–312.

DeLano, W. L. (2002). *The PyMOL molecular graphics system*. San Carlos, CA, USA: De Lano Scientific.

Dieckmann, G. R., McRorie, D. K., Tierney, D. L., Utschig, L. M., Singer, C. P., O'Halloran, T. V., et al. (1997). De novo design of mercury-binding two- and three-helical bundles. *Journal of the American Chemical Society*, *119*(26), 6195–6196.

Empting, M., Avrutina, O., Meusinger, R., Fabritz, S., Reinwarth, M., Biesalski, M., et al. (2011). "Triazole bridge": Disulfide-bond replacement by ruthenium-catalyzed formation of 1,5-disubstituted 1,2,3-triazoles. *Angewandte Chemie International Edition*, *50*(22), 5207–5211.

Faiella, M., Andreozzi, C., de Rosales, R. T. M., Pavone, V., Maglio, O., Nastri, F., et al. (2009). An artificial di-iron oxo-protein with phenol oxidase activity. *Nature Chemical Biology*, *5*(12), 882–884.

Fairman, R., Chao, H.-G., Lavoie, T. B., Villafranca, J. J., Matsueda, G. R., & Novotny, J. (1996). Design of heterotetrameric coiled coils: Evidence for increased stabilization by Glu−Lys+ ion pair interactions. *Biochemistry*, *35*(9), 2824–2829.

Friedle, S., Reisner, E., & Lippard, S. J. (2010). Current challenges of modeling diiron enzyme active sites for dioxygen activation by biomimetic synthetic complexes. *Chemical Society Reviews*, *39*(8), 2768–2779.

Frolow, F., Kalb, A. J., & Yariv, J. (1994). Structure of a unique twofold symmetric haem-binding site. *Nature Structural Biology*, *1*(7), 453–460.

Fry, H. C., Lehmann, A., Saven, J. G., DeGrado, W. F., & Therien, M. J. (2010). Computational design and elaboration of a de novo heterotetrameric α-helical protein that selectively binds an emissive abiological (porphinato)zinc chromophore. *Journal of the American Chemical Society*, *132*(11), 3997–4005.

Galibert, M., Sancey, L., Renaudet, O., Coll, J.-L., Dumy, P., & Boturyn, D. (2010). Application of click–click chemistry to the synthesis of new multivalent RGD conjugates. *Organic & Biomolecular Chemistry*, *8*(22), 5133.

Góngora-Benítez, M., Cristau, M., Giraud, M., Tulla-Puche, J., & Albericio, F. (2012). A universal strategy for preparing protected C-terminal peptides on the solid phase through an intramolecular click chemistry-based handle. *Chemical Communications*, *48*(17), 2313.

Grigoryan, G., Reinke, A. W., & Keating, A. E. (2009). Design of protein-interaction specificity gives selective bZIP-binding peptides. *Nature*, *458*(7240), 859–864.

Harrison, P. M., & Arosio, P. (1996). The ferritins: Molecular properties, iron storage function and cellular regulation. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, *1275*(3), 161–203.

Havranek, J. J., & Harbury, P. B. (2003). Automated design of specificity in molecular recognition. *Nature Structural & Molecular Biology*, *10*(1), 45–52.

Hazes, B., & Dijkstra, B. W. (1988). Model building of disulfide bonds in proteins with known three-dimensional structure. *Protein Engineering*, *2*(2), 119–125.

Hill, R. B., Raleigh, D. P., Lombardi, A., & DeGrado, W. F. (2000). *De novo* design of helical bundles as models for understanding protein folding and function. *Account of Chemical Research*, *33*(11), 745–754.

Horne, W. S., Yadav, M. K., Stout, C. D., & Ghadiri, M. R. (2004). Heterocyclic peptide backbone modifications in an α-helical coiled coil. *Journal of the American Chemical Society*, *126*(47), 15366–15367.

Jacobsen, Ø., Maekawa, H., Ge, N.-H., Görbitz, C. H., Rongved, P., Ottersen, O. P., et al. (2011). Stapling of a 310-helix with click chemistry. *The Journal of Organic Chemistry*, *76*(5), 1228–1238.

Jordan, A., & Reichard, P. (1998). Ribonucleotide reductases. *Annual Review of Biochemistry*, *67*(1), 71–98.

Kaplan, J., & DeGrado, W. F. (2004). De novo design of catalytic proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(32), 11566–11570.

Kohn, W. D., & Hodges, R. S. (1998). De novo design of α-helical coiled coils and bundles: Models for the development of protein-design principles. *Trends in Biotechnology*, *16*(9), 379–389.

Kolb, H. C., Finn, M. G., & Sharpless, K. B. (2001). Click chemistry: Diverse chemical function from a few good reactions. *Angewandte Chemie International Edition*, *40*(11), 2004–2021.

Lau, Y. H., Wu, Y., de Andrade, P., Galloway, W. R. J. D., & Spring, D. R. (2015). A two-component "double-click" approach to peptide stapling. *Nature Protocols*, *10*(4), 585–594.

Lee, D., & Lippard, S. J. (2003). 8.13—Nonheme di-iron enzymes. In J. A. McCleverty & T. J. Meyer (Eds.), *Comprehensive coordination chemistry II* (pp. 309–342). Oxford: Pergamon.

Lindqvist, Y., Huang, W., Schneider, G., & Shanklin, J. (1996). Crystal structure of delta9 stearoyl-acyl carrier protein desaturase from castor seed and its relationship to other di-iron proteins. *The EMBO Journal*, *15*(16), 4081–4092.

Litowski, J. R., & Hodges, R. S. (2002). Designing heterodimeric two-stranded α-helical coiled-coils. Effects of hydrophobicity and α-helical propensity on protein folding, stability, and specificity. *Journal of Biological Chemistry*, *277*(40), 37272–37279.

Maglio, O., Nastri, F., & Lombardi, A. (2012). Structural and functional aspects of metal binding sites in natural and designed metalloproteins. In A. Ciferri & A. Perico (Eds.), *Ionic interactions in natural and synthetic macromolecules* (pp. 361–450). Hoboken, NJ: John Wiley & Sons, Inc.

Marine, J. E., Song, S., Liang, X., Watson, M. D., & Rudick, J. G. (2015). Bundle-forming α-helical peptide–dendron hybrid. *Chemical Communications*, *51*(76), 14314–14317.

Marsh, E. N. G., & DeGrado, W. F. (2002). Noncovalent self-assembly of a heterotetrameric diiron protein. *Proceedings of the National Academy of Sciences*, *99*(8), 5150–5154.

Mathews, F. S., Bethge, P. H., & Czerwinski, E. W. (1979). The structure of cytochrome b562 from Escherichia coli at 2.5 A resolution. *Journal of Biological Chemistry*, *254*(5), 1699–1706.

Mayo, S. L., Olafson, B. D., & Goddard, W. A. (1990). DREIDING: A generic force field for molecular simulations. *The Journal of Physical Chemistry*, *94*(26), 8897–8909.

Miteva, M. A., Guyon, F., & Tufféry, P. (2010). Frog2: Efficient 3D conformation ensemble generator for small compounds. *Nucleic Acids Research*, *38*(Suppl. 2), W622–W627.

Monien, B. H., Drepper, F., Sommerhalter, M., Lubitz, W., & Haehnel, W. (2007). Detection of heme oxygenase activity in a library of four-helix bundle proteins: Towards the de novo synthesis of functional heme proteins. *Journal of Molecular Biology*, *371*(3), 739–753.

Mutter, M. (2013). Four decades, four places and four concepts. *CHIMIA International Journal for Chemistry*, *67*(12), 868–873.

Mutter, M., Altmann, E., Altmann, K.-H., Hersperger, R., Koziej, P., Nebel, K., et al. (1988). The construction of new proteins. Part III. Artificial folding units by assembly of amphiphilic secondary structures on a template. *Helvetica Chimica Acta*, *71*(4), 835–847.

Mutter, M., & Tuchscherer, G. (1997). Non-native architectures in protein design and mimicry. *Cellular and Molecular Life Sciences CMLS*, *53*(11–12), 851–863.

Pabo, C. O., & Suchanek, E. G. (1986). Computer-aided model-building strategies for protein design. *Biochemistry*, *25*(20), 5987–5991.

Peacock, A. F. (2016). Recent advances in designed coiled coils and helical bundles with inorganic prosthetic groups—From structural to functional applications. *Current Opinion in Chemical Biology*, *31*, 160–165.

Rachel, N. M., & Pelletier, J. N. (2016). One-pot peptide and protein conjugation: A combination of enzymatic transamidation and click chemistry. *Chemical Communications*, *52*(12), 2541–2544.

Rau, H. K., DeJonge, N., & Haehnel, W. (1998). Modular synthesis of de novo-designed metalloproteins for light-induced electron transfer. *Proceedings of the National Academy of Sciences*, *95*(20), 11526–11531.

Rau, H. K., DeJonge, N., & Haehnel, W. (2000). Combinatorial synthesis of four-helix bundle hemoproteins for tuning of cofactor properties. *Angewandte Chemie International Edition*, *39*(1), 250–253.

Rau, H. K., & Haehnel, W. (1998). Design, synthesis, and properties of a novel cytochrome b model. *Journal of the American Chemical Society*, *120*(3), 468–476.

Reig, A. J., Pires, M. M., Snyder, R. A., Wu, Y., Jo, H., Kulp, D. W., et al. (2012). Altering the O_2-dependent reactivity of de novo due ferri proteins. *Nature Chemistry*, *4*(11), 900–906.

Rozwarski, D. A., Gronenborn, A. M., Clore, G. M., Bazan, J. F., Bohm, A., Wlodawer, A., et al. (1994). Structural comparisons among the short-chain helical cytokines. *Structure*, *2*(3), 159–173.

Samish, I., MacDermaid, C. M., Perez-Aguilar, J. M., & Saven, J. G. (2011). Theoretical and computational protein design. *Annual Review of Physical Chemistry*, *62*(1), 129–149.

Sazinsky, M. H., & Lippard, S. J. (2015). Methane monooxygenase: Functionalizing methane at iron and copper. In P. M. H. Kroneck & M. E. S. Torres (Eds.), *Sustaining life on planet earth: Metalloenzymes mastering dioxygen and other chewy gases* (pp. 205–256): Switzerland: Springer International Publishing.

Schnepf, R., Haehnel, W., Wieghardt, K., & Hildebrandt, P. (2004). Spectroscopic identification of different types of copper centers generated in synthetic four-helix bundle proteins. *Journal of the American Chemical Society*, *126*(44), 14389–14399.

Scrima, M., Le Chevalier-Isaad, A., Rovero, P., Papini, A. M., Chorev, M., & D'Ursi, A. M. (2010). CuI-catalyzed azide–alkyne intramolecular i-to-(i+4) side-chain-to-side-chain cyclization promotes the formation of helix-like secondary structures. *European Journal of Organic Chemistry*, *2010*(3), 446–457.

Shu, J. Y., Tan, C., DeGrado, W. F., & Xu, T. (2008). New design of helix bundle peptide−polymer conjugates. *Biomacromolecules*, *9*(8), 2111–2117.

Sirajuddin, S., & Rosenzweig, A. C. (2015). Enzymatic oxidation of methane. *Biochemistry*, *54*(14), 2283–2294.

Slope, L. N., & Peacock, A. F. A. (2016). De novo design of xeno-metallo coiled coils. *Chemistry—An Asian Journal*, *11*(5), 660–666.

Smith, B. A., & Hecht, M. H. (2011). Novel proteins: From fold to function. *Current Opinion in Chemical Biology*, *15*(3), 421–426.

Sowdhamini, R., Srinivasan, N., Shoichet, B., Santi, D. V., Ramakrishnan, C., & Balaram, P. (1989). Stereochemical modeling of disulfide bridges. Criteria for introduction into proteins by site-directed mutagenesis. *Protein Engineering*, *3*(2), 95–103.

Stenkamp, R. E. (1994). Dioxygen and hemerythrin. *Chemical Reviews*, *94*(3), 715–726.

Summa, C. M., Rosenblatt, M. M., Hong, J.-K., Lear, J. D., & DeGrado, W. F. (2002). Computational de novo design, and characterization of an A2B2 diiron protein. *Journal of Molecular Biology*, *321*(5), 923–938.

Thomas, F., Boyle, A. L., Burton, A. J., & Woolfson, D. N. (2013). A set of de novo designed parallel heterodimeric coiled coils with quantified dissociation constants in the micromolar to sub-nanomolar regime. *Journal of the American Chemical Society*, *135*(13), 5161–5166.

Valverde, I. E., Vomstein, S., Fischer, C. A., Mascarin, A., & Mindt, T. L. (2015). Probing the backbone function of tumor targeting peptides by an amide-to-triazole substitution strategy. *Journal of Medicinal Chemistry*, *58*(18), 7475–7484.

Vita, N., Platsaki, S., Baslé, A., Allen, S. J., Paterson, N. G., Crombie, A. T., et al. (2015). A four-helix bundle stores copper for methane oxidation. *Nature*, *525*(7567), 140–143.

Wahlgren, W. Y., Omran, H., von Stetten, D., Royant, A., van der Post, S., & Katona, G. (2012). Structural characterization of bacterioferritin from blastochloris viridis. *PloS One*, *7*(10), e46992.

Weber, P. C., Bartsch, R. G., Cusanovich, M. A., Hamlin, R. C., Howard, A., Jordan, S. R., et al. (1980). Structure of cytochrome c′: A dimeric, high-spin haem protein. *Nature*, *286*(5770), 302–304.

Witte, M. D., Theile, C. S., Wu, T., Guimaraes, C. P., Blom, A. E. M., & Ploegh, H. L. (2013). Production of unnaturally linked chimeric proteins using a combination of sortase-catalyzed transpeptidation and click chemistry. *Nature Protocols*, *8*(9), 1808–1819.

Yoon, J., Fujii, S., & Solomon, E. I. (2009). Geometric and electronic structure differences between the type 3 copper sites of the multicopper oxidases and hemocyanin/tyrosinase. *Proceedings of the National Academy of Sciences*, *106*(16), 6585–6590.

Yu, F., Cangelosi, V. M., Zastrow, M. L., Tegoni, M., Plegaria, J. S., Tebo, A. G., et al. (2014). Protein design: Toward functional metalloenzymes. *Chemical Reviews*, *114*(7), 3495–3578.

Zhang, Y., Bartz, R., Grigoryan, G., Bryant, M., Aaronson, J., Beck, S., et al. (2015). Computational design and experimental characterization of peptides intended for pH-dependent membrane insertion and pore formation. *ACS Chemical Biology*, *10*(4), 1082–1093.