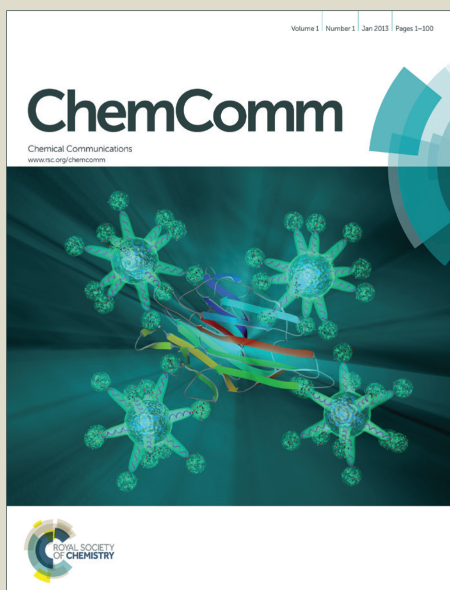
ChemComm

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

www.rsc.org/chemcomm

Journal Name

RSCPublishing

COMMUNICATION

Cite this: DOI: 10.1039/x0xx00000x

Protein Engineering with Artificial Chemical Nucleases

Received 00th January 2012,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

Ruth Larragy,[a] Jenny Fitzgerald[a], Andreea Prisecaru,[b] Vickie McKee,[b] Paul Leonard,[a,c]* and Andrew Kellett[b]*

Herein we report the application of oxidative artificial chemical nucleases as novel agents for protein engineering. The complex ion $[Cu(Phen)_2(H_2O)]^{2+}$ (CuPhen; Phen = 1,10-phenanthroline) was applied under Fenton-type conditions against a recombinant antibody fragment specific for prostate-specific antigen (PSA) and compared against traditional DNA shuffling using DNase I for the generation of recombinant mutagenesis libraries. We show that digestion and re-annealment of single chain variable fragment (scFv) coding DNA is possible using CuPhen. Results indicate recombinant library generation in this manner may generate novel clones—not accessible through the use of DNase I—with CuPhen producing highly PSA-specific binding antibodies identified by surface plasmon resonance.

Protein engineering is a well-established field for improving the performance of target proteins in areas such as antibody affinity or enzyme catalytic activity.[1-5] However, some of these processes, for example DNA shuffling—involving enzymes such as DNase I—generate a limited diversity of mutants due to the biased nature of their nuclease activity.[6,7] Small molecule artificial metallo-nucleases (AMNs), such as the chemical nuclease $Cu(Phen)_2$ cation (Phen = 1,10-phenanthroline)[8] have already found biochemical application within molecular biology (*e.g.* DNA footprinting)[9] and may provide unique features for unbiased genetic shuffling.[10] At the centre of AMN-induced activity by **CuPhen** lies the Haber-Weiss process (Figure 1B): copper ions generate Fenton-type products (the hydroxyl radical and the hydroxide anion), from superoxide radicals and hydrogen peroxide, under aerobic conditions in the presence of a reductant such as ascorbate.[11-13] This radical chemistry, combined with the nucleic acid binding affinity of **CuPhen**, provide the reactive platform for DNA digestion through nucleoside H-atom abstraction.[14,15] In the context of DNA shuffling, **CuPhen** may also provide additional variation due to mutagenic base oxidation reactions (*e.g.* 8-oxo-deoxyguanosine; 8-oxo-dG),[16] however, it has not yet been established if digestion products from **CuPhen** are suitable for fragment annealing through primerless PCR (polymerase chain reaction)[17] (Figure 1C). Thus, the aim of this study was to determine whether **CuPhen** could induce DNA scission in a totally unbiased fashion, allowing true randomisation of target gene sequences that could potentially generate new, and diverse, prostate specific antigen (PSA)—the major biomarker used in detection of prostate cancer—recombinant antibody mutants that are not accessible under current established techniques.

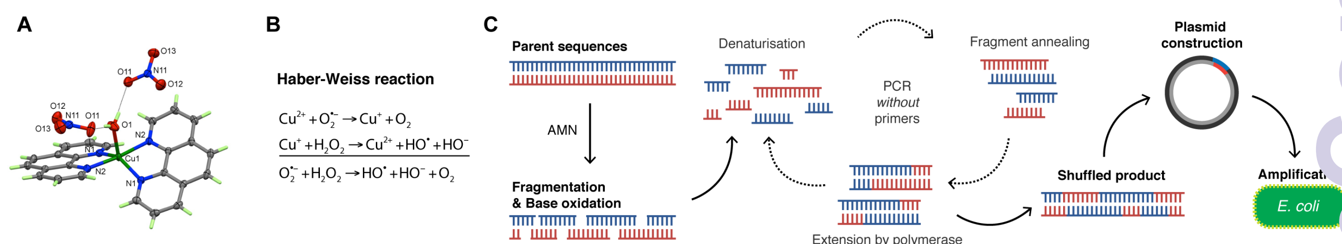


Figure 1. A. Crystal structure of $[Cu(Phen)_2(H_2O)](NO_3)_2$ **(CuPhen)**, colour scheme: copper, green; oxygen, red; nitrogen, blue, **B**. Production of hydroxyl radical ($HO^{\bullet}$) through Haber-Weiss-type reaction, and **C**. Experimental protocol design for DNA shuffling random mutagenesis of two scFv genes with DNase I and **CuPhen**.

This journal is © The Royal Society of Chemistry 2012

ChemComm Accepted Manuscript

$[Cu(Phen)_2(H_2O)](NO_3)_2$ (**CuPhen**, Figure 1A) was generated by ethanolic reflux of copper(II) nitrate with 1,10-phenanthroline[18] and crystallised upon slow evaporation (supplementary S1). The copper ion five-coordinate, with distorted square-pyramidal geometry; the cation lies on a 2-fold axis passing through the copper and O1. The coordinated water molecule is H-bonded to two (equivalent) nitrate anions and within the crystal lattice packing is governed by π-π-stacking of the Phen rings (supplementary S2). The mechanistic and stability aspects of the **CuPhen** complex have been reported extensively by this group.[15,16,18] Prior to DNA shuffling experiments, the selection and generation of a suitable α-PSA single chain variable fragment (scFv)—a bifunctional fusion protein having both antigen-binding capacity and biomarker activity applied for one-step immunodetection of biological agents[19]—was evaluated using the enzyme-linked immunosorbent assay (ELISA) against immobilised PSA using a pComb vector[20] (a small piece of DNA specifically designed for the construction of antibody libraries) scFv library stock previously generated at this institute (supplemental S4). From the selection of positive fluorescent signals against PSA (Table S3), plasmid DNA of a selected clone (from well F8, Figure S3) was prepared, sequenced by Eurofins Genomics GmbH (Figure S1) and translated into protein sequence using ExPASy protein translation tool (Figure 2A pComb_F8). scFv F8 DNA required for subsequent DNA shuffling experiments was prepared using PCR with CSC-F and CSC-B primers (Figure 2).

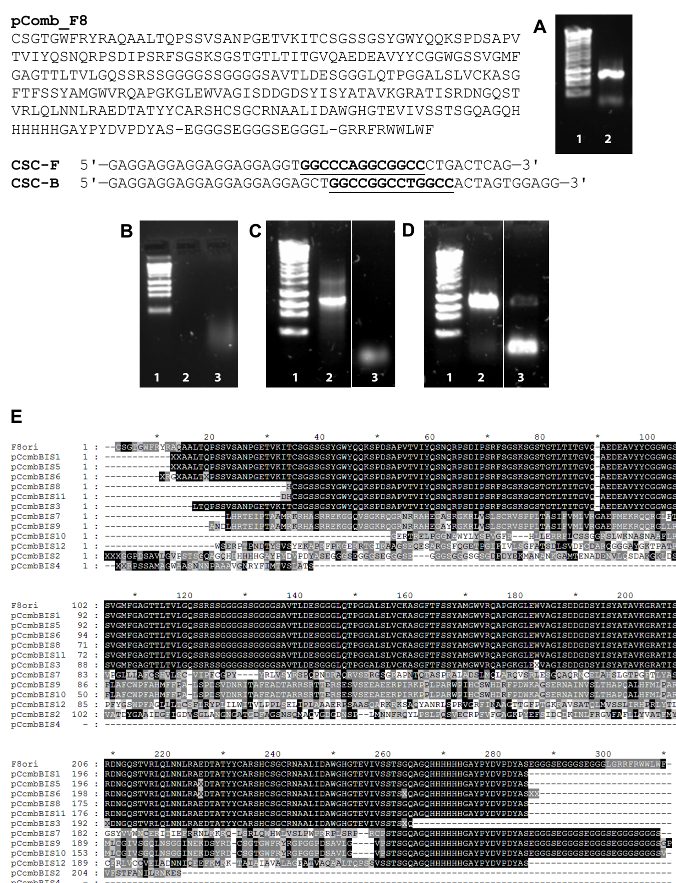


Figure 2. scFv clone F8 (pComb_F8) DNA translated to protein sequence along with forward and reverse primer DNA sequences (CSC-F and CSC-B). **A.** Agarose gel electrophoresis of scFv clone F8 DNA (lane 2), **B.** Digestion of F8 DNA with DNase I (lane 2) and with 1 uM **CuPhen** in the presence of 1 uM of Na-*L*-ascorbate (lane 3), **C.** Amplification of self-annealing scFv DNA digested with DNase I (lane 2) and **CuPhen** (lane 3), **D.** Further amplification of scFv DNA digested with DNaseI (lane 2) and CuPhen (lane 3), Note **A.** – **D.** Lane 1: 1 kb ladder. **E.** Protein alignment of sequencing results for 12 **CuPhen** digested library sample clones.

Next, F8 DNA was digested using DNase I (NEB); the reaction mixture was heated to 37 °C before the addition of DNase I and the reaction was carried out over 20 min. Concomitantly, F8 DNA was digested by 1 μM of the AMN **CuPhen** with 1 mM of Na-L-ascorbate at pH 7.2 in the presence of 25 mM NaCl. Samples were incubated at 37 °C for 45 min and quenched with EDTA and neocuproine (supplementary S3). Digested DNA from each process were ethanol precipitated, re-suspended in H_2O and analysed by 2% gel electrophoresis (Figure 2B). There were no parent (i.e. 750 bp) fragments visible in either sample. Digested DNA was next self-annealed using primerless PCR (Figure 2C), carried out for 45 cycles using the method outlined in Figure 1C. The resulting PCR products were then diluted 1:25 in fresh PCR reaction mixture with scFv sequences being amplified using CSC-F and CSC-B primers. Initially, 15 cycles were used to amplify from the re-annealed DNase I and **CuPhen** digests but due to low levels of 750 bp fragment recovery from the **CuPhen**-digested DNA, this was increased to 30 cycles to improve overall DNA yield (Figure 2D).

*Sfi*I restriction digests were then carried out to prepare pComb3XSS, DNase I, and **CuPhen** scFv DNA samples for ligation (required during plasmid construction in Figure 1C, supplementary S4). DNA and vector samples were ligated using T4 DNA ligase—an enzyme that catalyses the joining of DNA sequences *via* phosphodiester bond formation—and transformed into freshly prepared electrocompetent *E. coli* XL-1 blue cells. The resulting libraries were of sufficient number and quality to proceed with analysis. Plasmid DNA of representative clones from each library were prepared and sequenced to check for desired mutations compared to the original F8 clone. Sequencing of the DNase I library revealed 7 of the 9 sequenced samples were unique and different from the original—albeit with very low levels of mutation within each translated protein sequence (Figure S4, Table S4). There were 5 unique sequences among the 12 **CuPhen** fragments (Table S5). Further, the **CuPhen** library appeared to have a high level of mutation (Figure 2E) in unique clones, however, almost half matched the original F8 sequence.

Three rounds of phage display—an *in vitro* process that facilitates antibody selection whereby a gene encoding the protein of interest is inserted into a phage coat protein gene, causing the phage to "display" the protein on its outer surface that results in direct connection between genotype and phenotype[20]—were carried out on the DNase I library and two rounds on the **CuPhen** libraries (Round 1, 10 μg/ml PSA; Round 2, 5 μg/ml PSA; Round 3, 2 μg/ml PSA). 190 colonies from the final round of output plates from each library were selected and screened against 50 μl of PSA immobilised at 1 μg/ml using the 'on plate' growth method.[21] There were 63 positive readings from the DNase I output plates including the original F8 clone (Table S6). These 63 clones were re-grown for antibody titration with 34 clones selected for sequencing. There were 57 positive clones on the **CuPhen** plates (Table S7) with the majority sent for sequencing. A comparison of diversity within sequences of DNase I and **CuPhen** phage display output clones is shown in Table S8. Clones from the DNase I library had greater levels of diversity among sequenced clones when compared to the **CuPhen** library. This is likely to lower levels of scFv DNA re-annealing/recombining within the AMN digested library induced by higher levels of digestive DNA damage as compared with DNase I. Indeed, we anticipated this result given the known DNA oxidative damage effects of this agent.[16] Further, the requirement for higher PCR cycle numbers for scFv fragment recovery from AMN digested DNA (30

This journal is © The Royal Society of Chemistry 2012
ChemComm Accepted Manuscript

cycles) compared to the DNase I digested DNA (15 cycles) indicates there is a higher chance of re-amplification of identical DNA sequences, which may contribute to lower overall sequence diversity. The process of phage display for library panning purposes would also eliminate any clones with low specificity/affinity for PSA, perhaps due to high levels of sequence shuffling, while enriching strong binders, many of which could have identical sequences. Although DNase I digested clones have higher rates of mutation within this sequence, these mutations may have occurred in areas of the protein not involved in antigen binding and may not necessarily contribute to improvements in antigen binding affinity.

The unique HA (hemagglutinin) tagged—an extensively used epitope tag for expression vectors—scFv clones from the **CuPhen** and DNase I libraries were expressed overnight in 20 ml of media supplemented with carbenicillin (50 µg/ml) incubated at 37 °C with shaking. The cells were spun out, re-suspended in 2 ml of PBS, and then lysed using freeze thaw (x3). Cell debris was removed by centrifugation and the scFv containing lysate was stored for Biacore analysis[22]; this technology uses the phenomenon of surface plasmon resonance to monitor changes in the refractive index of a surface correlating in this instance to antibody-antigen interaction and is currently the most widely used technique for high-resolution antibody characterisation.[23] Anti-HA antibody was immobilised onto the surface of a CM5 chip using standard amine coupling. scFv containing lysate was injected over the surface of the coated chip and captured by the anti-HA antibody. PSA was passed over the chip surface, captured by the immobilised scFv and left to dissociate in buffer (HEPES buffered saline). Any scFv-PSA complex that remained bound to the surface immobilised anti-HA antibody was removed using a 20 mM NaOH wash with runs being carried out in duplicate. The % of bound PSA remaining following dissociation was calculated and compared to the original F8 clone (Table 1). Five of the top ten best performing clones, including the top four, were from the **CuPhen**-digested library. There did not appear to be any mutational 'hot spots' within the sequences contributing to enhanced PSA binding activity indicating that specific amino acid mutations perhaps give rise to critical modifications to protein folding. It is also possible that due to the unbiased nature of AMN DNA digestion, output clones from this library contain DNA sequence mutations not accessible through digestion with DNase I, which ultimately contribute to tighter binding of PSA.

Table 1. Characterisation of scFv clones bound to PSA by CM5 on-chip analysis. Calculation and ranking of % scFv left bound to PSA following dissociation in buffer.

	Average % retained	Times better than F8	Rank
CuPhen F8	71.69 %	2.19	1
CuPhen E5	66.03 %	2.02	2
CuPhen E7	63.27 %	1.93	3
CuPhen E12	62.10 %	1.90	4
DNaseI 2D12	57.27 %	1.75	5
DNaseI 2C7	53.35 %	1.63	6
DNaseI 2F1	52.63 %	1.61	7
DNaseI 2D2	52.18 %	1.59	8
CuPhen F3	50.73%	1.55	9
DNaseI 2B6	50.63%	1.55	10

In the context of control and optimisation of DNA digestion by **CuPhen**, this group have previously reported the application of 'on-chip' microfluidics using the Agilent Bioanalyzer 2100.[14] Here, a linear segment of dsDNA can be exposed, kinetically, to a copper nuclease and the digestion process monitored at specific time points by quenching AMN activity with neocuproine prior to separating digestion fragments by capillary electrophoresis. This method not only facilitates the study of chemical nuclease activity at specific intervals of time, but also could easily be applied in the optimisation of digestion processes based on fragment size distribution, **CuPhen** concentration, pH, exogenous ascorbate concentration *etc*. To further illustrate the versatility of this technique, **CuPhen** (5 µM) with Na-*L*-ascorbate (100 µM) was kinetically exposed to dsDNA fragments of increasing CG content generated *via* PCR (36% 120 bp; 50% 140 bp; 63% 160 bp) (supplementary S-5). DNA digestion was monitored simultaneously at selected time points between 0-60 min with results shown in Figure S5 A and B. Using peak height analysis, it was possible to identify the smooth nuclease efficiency of **CuPhen** as a function of base composition; here, all three sequences are degraded over 1 h with the highest percentage GC fragment (63%) being degraded more efficiently than those of higher AT content by 30 min (Figure S5 B). Furthermore, since **CuPhen** is known to oxidatively damage duplex DNA from the minor groove, we suggest that additional control on the digestion process can be made through the introduction of recognition elements that prime or inhibit this binding site. Recently, we have established that the introduction of the major groove binder, methy green, primes or enhances oxidative dsDNA damage by **CuPhen** while the minor groove binder, netropsin, inhibits its activity.[16] In that study we also monitored the degree of oxidative lesion formation by **CuPhen** through the use of an 8-oxo-2′-deoxyguanosine (8-oxo-dG) ELISA protocol. From the perspective of protein engineering, therefore, it is possible to monitor the number of oxidative lesions—or mutations being introduced into single chain variable fragment (scFv) coding DNA.

Conclusions

Although more in-depth analysis will be required to evaluate the commercial application of these results in comparison to other well-established techniques such as staggered extension process (StEP) recombination, data here indicates that artificial chemical nucleases may become a valuable tool in the generation of recombinant mutagenesis libraries through the unbiased introduction of mutations within a specific gene sequence. This method could be applied, not only in the context of recombinant antibodies, but to proteins in any area of research for directed protein evolution leading to improved products and processes: particularly within the field of biomedical diagnostics.

Notes and references

[a] Biomedical Diagnostics Institute, Dublin City University, Dublin 9, Ireland.

[b] School of Chemical Sciences and National Institute for Cellular Biotechnology, Dublin City University, Dublin 9, Ireland. Tel.: +353-1-7005461, Email: andrew.kellett@dcu.ie

[c] Current address: Vaccinogen Ireland R&D Company, Hamilton R&D Building, Dublin 9, Ireland.

Electronic Supplementary Information (ESI) available: Experimental procedures and biological evaluation studies. CCDC 1403896. For ESI and crystallographic data see DOI: 10.1039/c000000x/

The PSA antibody used here was kindly donated by Prof. Richard O'Kennedy of Dublin City University. This work was funded by Science Foundation Ireland (13/TIDA/B2682) and the Irish Research Council (GOIPG/2013/937). AK acknowledges funding from the Marie Skłodowska-Curie Innovative Training Network (ITN) ClickGene (H2020-MSCA-ITN-2014-642023).

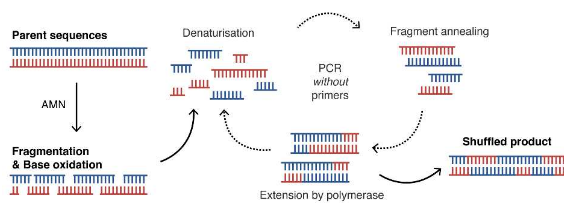
1 J. E. Thompson, T. J. Vaughan, A. J. Williams, J. Wilton, K. S. Johnson, L. Bacon, J. A. Green, R. Field, S. Ruddock, M. Martins, A.

ChemComm Accepted Manuscript

R. Pope, P. R. Tempest and R. H. Jackson, *J. Immunol. Methods*, 1999, **227**, 17–29.
2 R. Schier, J. Bye, G. Apell, A. McCall, G. P. Adams, M. Malmqvist, L. M. Weiner and J. D. Marks, *J. Mol. Biol.*, 1996, **255**, 28–43.
3 R. Schier, A. McCall, G. P. Adams, K. W. Marshall, H. Merritt, M. Yim, R. S. Crawford, L. M. Weiner, C. Marks and J. D. Marks, *J. Mol. Biol.*, 1996, **263**, 551–567.
4 M. Ohlin, H. Owman, M. Mach and C. A. Borrebaeck, *Mol. Immunol.*, 1996, **33**, 47–56.
5 A. Wörn and A. Plückthun, *Biochemistry*, 1999, **38**, 8739–8750.
6 D. H. Sutton, G. L. Conn, T. Brown and A. N. Lane, *Biochem. J.*, 1997, **321 (Pt 2)**, 481–486.
7 W. Mäueler, G. Bassili, C. Epplen, H. G. Keyl and J. T. Epplen, *Chromosome Res.*, 1999, **7**, 163–166.
8 T. B. Thederahn, M. D. Kuwabara, T. A. Larsen and D. Sigman, *J. Am. Chem. Soc.*, 1989, **111**, 4941–4946.
9 a) J. Gallagher, C. B. Chen, C. Q. Pan, D. M. Perrin, Y. M. Cho and D. S. Sigman, *Bioconjug. Chem.*, 1996, **7**, 413–420.
10 S. Basak and V. Nagaraja, *Nucleic Acids Res.*, 2001, **29**, 105.
11 J. A. Imlay, *Annu. Rev. Microbiol.*, 2003, **57**, 395–418.
12 C. Von Sonntag, *Free-Radical-Induced DNA Damage and Its Repair*, Springer Science & Business Media, Berlin, Heidelberg, 2006.
13 B. Halliwell and J. Gutteridge, *Biochem. J.*, 1984, **219**, 1–14.
14 Z. Molphy, A. Prisecaru, C. Slator, N. Barron, M. McCann, J. Colleran, D. Chandran, N. Gathergood and A. Kellett, *Inorg. Chem.*, 2014, **53**, 5392–5404.
15 A. Prisecaru, V. McKee, O. Howe, G. Rochford, M. McCann, J. Colleran, M. Pour, N. Barron, N. Gathergood and A. Kellett, *J. Med. Chem.*, 2013, **56**, 8599–8615.
16 Z. Molphy, C. Slator, C. Chatgilialoglu and A. Kellett, *Front. Chem.*, 2015, **3**, 28.
17 P. S. Daugherty, G. Chen, B. L. Iverson and G. Georgiou, *Proc. Natl. Acad. Sci. U.S.A.*, 2000, **97**, 2029–2034.
18 A. Prisecaru, M. Devereux, N. Barron, M. McCann, J. Colleran, A. Casey, V. McKee and A. Kellett, *Chem. Comm.*, 2012, **48**, 6906–6908.
19 N. E. Weisser and J. C. Hall, *Biotechnol. Adv.*, 2009, **27**, 502–520.
20 C. F. Barbas, D. R. Burton, J. K. Scott and G. J. Silverman, *Phage Display: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, 2001.
21 W. L. Guo, P. Leonard and R. O'Kennedy, *J. Immunol. Methods*, 2010, **359**, 61–64.
22 P. Leonard, S. Hearty and R. O'Kennedy, in *Protein Chromatography*, Humana Press, Totowa, NJ, 2010, vol. 681, pp. 403–418.
23 S. Hearty, P. Leonard and R. O'Kennedy, *Methods Mol. Biol.*, 2012, **907**, 411–442.

ChemComm Accepted Manuscript

This journal is © The Royal Society of Chemistry 2012



The process of protein engineering using artificial chemical nucleases is reported using the Cu(II)-*bis*-1,10-phenanthroline complex.

ChemComm Accepted Manuscript