A robust strategy for the preparation of libraries of metallopeptides. A new paradigm for the discovery of targeted molecular imaging and therapy agents†

Andrea F. Armstrong,[a] Nick Oakley,[a] Shannon Parker,[a] Patrick W. Causey,[a] Jennifer Lemon,[a] Alfredo Capretta,[a] Craig Zimmerman,[b] John Joyal,[b] Francis Appoh,[a] Jon Zubieta,[c] John W. Babich,[b] Gurmit Singh[a] and John F. Valliant*[a]

Received (in Cambridge, UK) 25th June 2008, Accepted 15th August 2008
First published as an Advance Article on the web 24th September 2008
DOI: 10.1039/b810706h

A robust method for synthesizing structurally diverse metallopeptide libraries using a Re(I) complex of a non-natural amino acid was developed as a way to accelerate the discovery of novel molecular imaging probes.

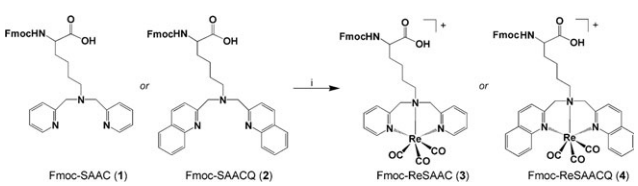
As technologies for identifying, quantifying and characterizing proteins from tissue samples generate complex proteome maps for diseases such as cancer, complementary methods are needed to examine the significance of the identified changes in receptor and enzyme levels in living subjects.[1] This realization has led to renewed interest in the radioimaging techniques single photon emission computed tomography (SPECT) and positron emission tomography (PET) which have the sensitivity necessary to monitor and quantify changes in the expression levels of specific proteins *in vivo*.[2,3] In order to use nuclear imaging methods for *in vivo* proteomics research and ultimately as clinical diagnostic tools, appropriate radiolabeled probes must first be developed.

Peptides are attractive vectors because sequences that bind virtually any protein target can be readily identified using phage display technologies.[4] In principle, a targeted molecular radioimaging agent can be prepared by attaching a chelate that can bind a radiometal to a lead peptide;[5,6] the chelate-conjugates can be often be prepared using standard automated peptide synthesis methods.[7] Unfortunately, developing new probes in this manner requires extensive and time-consuming optimization, largely because the core discovery methods for peptide-chelate bioconjugates tend to be linear in nature (*i.e.* one molecule at a time) and do not take into account the impact of the chelate–metal complex on the binding of the peptide to its target.

There is a need for a methodology that can be used to create libraries of structurally diverse metallopeptides which can be screened using modern high throughput methods to identify lead molecular imaging probes. To be successful, this methodology must produce non-radioactive metallopeptides that are structurally identical to the radiolabelled analogues, rather than simply peptides with a free chelate unit.

To move towards this type of discovery platform, we developed the single amino acid chelate ligands (Fmoc-SAAC **1** and Fmoc-SAACQ **2**, Q = quinoline, Scheme 1) which form stable complexes with technetium-99m, the most widely used radionuclide in diagnostic medicine, in the form of $[^{99m}Tc(CO)_3]^+$.[8,9] The non-radioactive Re(I) complexes of these ligands (Fmoc-ReSAAC **3** and Fmoc-ReSAACQ **4**), which serve as surrogates for their radioactive ^{99m}Tc analogues, are sufficiently stable that they can be incorporated into discrete peptides using standard automated synthesis and resin cleavage methods.[10,11]

The subsequent objective was to develop a means of producing libraries of structurally diverse metallated peptides containing ReSAAC(Q): these can then be screened in an analogous manner to that used to select traditional peptide leads. To construct a suitably diverse library, multi-gram quantities of **3** and **4** were required which are substantially greater than the small scale preparations reported previously.[8,9] The synthesis of ligands **1** and **2**[12] can readily be increased in scale such that a 30 g quantity of the free ligand is obtained from a one pot reaction using commercially available materials. Formation of the Re(I) chelate complexes (Scheme 1) was accomplished previously by heating a solution of the ligand with $[Et_4N]_2[Re(CO)_3Br_3]$ in boiling MeOH.[8,9] Alternatively, we discovered that combining a solution of **1** with $[Re(CO)_3(OH_2)_3]Br$[13] and heating in a microwave reactor for 11 min produces multi-gram quantities of the metal complex **3** in high purity and yield (89%). Although this material can be used without further purification,



Scheme 1 (i) $[Re(CO)_3]^+$, CH_3OH, 65 °C, 5 h; or microwave 110 °C, 11 min, CH_3CN–H_2O.

[a] *Departments of Chemistry and Medical Physics and Applied Radiation Sciences, McMaster University, Hamilton, ON, Canada. E-mail: valliant@mcmaster.ca; Fax: 905-522-7776; Tel: 905-525-9140 ext. 22840*
[b] *Molecular Insight Pharmaceuticals Inc., Cambridge, MA 02142, USA*
[c] *Department of Chemistry, Syracuse University, Syracuse, NY 13244-4100, USA*

† Electronic supplementary information (ESI) available: Details of experimental procedures. See DOI: 10.1039/b810706h

Table 1 Peptide libraries[17]

Peptide sequence	Amino acids used
H_2N-X_1-X_2-X_3-X_4-X_5-Leu-Trp-Ser-ReSAAC	All L-amino acids except Cys
H_2N-X_1-X_2-X_3-X_4-X_5-X_6-X_7-ReSAACQ	All L-amino acids except Cys
(H_2N-X_1-X_2-X_3-X_4-X_5)$_2$-Lys-ReSAAC	All L-amino acids except Cys
ReSAAC-x-Cha-F-s-x-Y-L-W-S	All amino acids except Cys
ReSAAC-X-Cha-F-s-x-Y-L-W-S	All amino acids except Cys
x-ReSAAC-x-Cha-F-s-x-Y-L-W-S	All amino acids except Cys

Peptides are written with the *N*-termini on the left and are linked to a Tentagel S-amine resin *via* the terminal acid residue; X denotes a position of variation; capital letters are used for L-amino acids.

the product purity can be increased to >99% using the Biotage automated reversed-phase chromatography system.

With access to large quantities of the protected metallo-amino acid analogues **3** and **4** in hand, their use in preparing libraries of resin-bound peptides was explored. The one-bead one-compound (OBOC) approach, in which each resin bead within the library contains multiple copies of a unique peptide, was employed because it is efficient, robust and amenable to using non-natural amino acids such as the Fmoc-ReSAAC(Q) units.[14] Peptide libraries were prepared (Table 1) using a shaker platform fitted with multiple polyethylene reaction vessels containing fine frits and the appropriate valves needed to facilitate washing and filtering the resins (see ESI† for details).[15] To confirm the effectiveness of the apparatus, a series of discrete model peptides containing both the ReSAAC and ReSAACQ ligands were prepared, monitoring each coupling and deprotection reaction to ensure they went to completion. Particular attention was given to steps involving the Fmoc-ReSAAC(Q) ligands. No difference was found in the times required for coupling or deprotection of the organometallic complexes **3** and **4** compared to natural amino acids, indicating that special reaction conditions were not required for these units. The identities of the model peptides were confirmed by on-bead sequencing[16] which further confirms the integrity of the synthetic methodology.

The synthesis of a peptide library consisting of the sequence H_2N-X_1-X_2-X_3-X_4-X_5-Leu-Trp-Ser-ReSAAC, where X represents the various positions for randomization, was then undertaken using a TentaGel S NH_2 resin. The Leu-Trp-Ser sequence was retained throughout as it is common to peptide ligands for urokinase plasminogen activator receptor (uPAR),[18] which is overexpressed in aggressive cancers and therefore an attractive target for developing probes to monitor metastases.[19] The library was constructed using a ''mix-and-split'' strategy in which the distribution of the beads was accomplished using isotonic solutions of DMF–CH_2Cl_2 (2 : 1 v/v). Variation in this case included 19 different natural amino acids (cysteine is excluded) resulting in 2 476 099 unique metallopeptides. Sufficient resin was used to ensure adequate statistical coverage of all possible combinations of amino acids; the entire synthesis was complete in less than 36 h.

Of equal importance is the ability to generate more focused libraries because it provides an opportunity to optimize the structure of probes based on peptides that have known affinity for a particular target. One of the benefits of this optimization strategy is that following screening, investigators will have a pool of potential leads that vary in composition and hence charge and lipophilicity. This information will be invaluable during the subsequent phase of probe development in that it can be used to select compounds that are the most likely to have the desired pharmacokinetic properties and route of clearance for the intended application which are key factors in finding translatable molecular imaging probes. To demonstrate the ability of our approach to achieve this objective, three focused ReSAAC-containing libraries (Table 1) were prepared based on the parent sequence X-x-Cha-F-s-x-Y-L-W-S (Cha = cyclohexyl-alanine), several variants of which are reported to bind to uPAR.[18]

From the various libraries prepared, several samples of resin were subjected to on-bead sequencing; all were found to contain the expected number of amino acids with no deletions detected. In addition, infrared spectra of various resin samples displayed strong CO stretching absorptions due to the $[Re(CO)_3]^+$ unit, affirming the presence of ReSAAC(Q) in the metallopeptides. The presence of the metal complexes could also be confirmed visually, as beads containing the Re ligands were much more intensely coloured than the pale yellow resin supporting natural peptides.

Having established a means of producing libraries of simple linear peptides, the technique was expanded to facilitate the generation of collections of dimeric ReSAAC-containing peptides. Bioconjugates containing more than one sequence can be used to enhance affinity for a particular target, facilitate cell penetration or alter pharmacokinetics.[20] The dimeric peptide library (H_2N-X_1-X_2-X_3-X_4-X_5)$_2$-Lys-ReSAAC-linked to a Tentagel resin was constructed off of a lysine anchor, with one chain appended to each of the *N*α- and *N*ε- amines. Randomization along the two chains was done simultaneously and as a result, each of the dimers produced contained identical sequences on both arms. For this library, 19 natural amino acids were again used, resulting in 19^5 = 2 476 099 unique sequences, which the quantity of resin employed (2 g) was capable of supporting.

One attractive feature of the ReSAACQ moiety **4** as compared to ReSAAC unit **3** is that it is luminescent and can be visualized by fluorescence microscopy. Thus, identifying a ''hit'' from a library of peptides containing **4** will generate both a lead SPECT imaging agent (^{99m}TcSAACQ analogue) and an isostructural optical probe (ReSAACQ analogue) having identical binding properties for the target of interest. In this instance, a more highly randomized library was produced using seven positions for variation (Table 1). Although the quantity of resin employed was not sufficient to ensure complete statistical coverage of the expected 19^7 (893 871 739)

unique sequences, it is likely not necessary to encompass all possible structural space in order to identify a lead sequence, from which a more focused library can be developed.

With several peptide libraries prepared, selecting a means of deconvoluting the vast number of peptides and identifying lead sequences became key.[18,21–23] It is important to note that finding molecules with high affinity for the protein of interest is at the heart of identifying a molecular imaging agent that can achieve high target-to-non-target ratios.[24,25] Libraries can be incubated with isolated proteins or with whole cells, and, because the targeting vectors are based on peptides, identifying the chemical structure of a "hit" is much less problematic than it is when screening combinatorial libraries derived from small molecules.

Although manual screening of bead libraries is an established technique, a model automated screening study was conducted using COPAS sorting instrument.[26] Here, a mixture of beads containing a peptide sequence (e-Cha-F-s-y-Y-L-W-S) known to bind *u*PAR[18] and a series of control beads were mixed and co-incubated with recombinant human soluble *u*PAR. The beads were washed, incubated with biotinylated anti-*u*PAR antibody, and washed again prior to incubation with FITC conjugated streptavidin (see ESI† for details). The COPAS system was able to identify the beads bearing the *u*PAR binding sequence (verified by amino acid sequencing[27]) and the controls in a quantitative manner based on fluorescence associated with target binding. The autofluorescence of the beads and that of the peptides were measured and did not interfere with the assay method. This screening approach is highly efficient and can be applied to a wide range of protein targets in addition to *u*PAR, providing a means of discovering probes for the ever increasing number of proteins of interest identified by disease-specific proteomic studies.

In summary, a new platform for preparing and identifying novel peptide-based Re/Tc molecular imaging probes was developed. This approach is unique in that by incorporating the metal into the peptide backbone during the construction of the library, screening data takes into account the presence of the chelate complex. With the advent of a method for producing large quantities of the rhenium complexes **3** and **4**, custom metallopeptide libraries can be prepared in a timely manner. In tandem with the ability of the COPAS screening system to sort millions of beads per day, the reported approach stands as a platform for discovering new imaging probes that is far superior in efficiency than traditional linear discovery methods. The associated radiochemistry needed to produce the technetium(I) core[28] $[^{99m}Tc(CO)_3]^+$ and the radiolabeled forms of lead agents[29] have been developed previously, thereby making it practical to convert hits from the library into imaging probes for preclinical and ultimately clinical testing. It will also be possible to use the reported chemistry to create radiotherapeutic analogues by using the β-emitting isotopes $^{186/188}Re$ in place of ^{99m}Tc.

This study was conducted with the support of the Ontario Institute for Cancer Research (OICR) through funding provided by the province of Ontario as well as the Canadian Institutes of Health Research (CIHR) provided by the government of Canada.

Notes and references

1 S. R. Cherry, *Phys. Med. Biol.*, 2004, **49**, R13–R48.
2 A. Saleem, N. Charnley and P. Price, *Eur. J. Cancer*, 2006, **42**, 1720–1727.
3 P. J. Cassidy and G. K. Radda, *J. R. Soc. Interface*, 2005, **2**, 133–144.
4 G. Smith and V. Petrenko, *Chem. Rev.*, 1997, **97**, 391–410.
5 S. Liu and D. Edwards, *Chem. Rev.*, 1999, **99**, 2235–2268.
6 S. M. Okarvi, *Cancer Treat. Rev.*, 2008, **34**, 13–26.
7 W. E. P. Greenland, K. Howland, J. Hardy, I. Fogelman and P. J. Blower, *J. Med. Chem.*, 2003, **46**, 1751–1757.
8 K. Stephenson, J. Zubieta, S. R. Banerjee, M. Levadala, L. Taggart, L. Ryan, N. McFarlane, D. R. Boreham, K. P. Maresca, J. W. Babich and J. F. Valliant, *Bioconjugate Chem.*, 2004, **15**, 128–136.
9 K. Stephenson, S. R. Banerjee, T. Besanger, O. Sogbein, M. Levadala, N. McFarlane, J. A. Lemon, D. R. Boreham, K. P. Maresca, J. Brennan, J. W. Babich, J. Zubieta and J. F. Valliant, *J. Am. Chem. Soc.*, 2004, **126**, 8598–8599.
10 P. Schaffer, J. A. Gleave, J. A. Lemon, L. C. Reid, L. K. K. Pacey, T. H. Farncombe, D. R. Boreham, J. Zubieta, J. W. Babich, L. C. Doering and J. F. Valliant, *Nucl. Med. Biol.*, 2008, **35**, 159–169.
11 S. R. Banerjee, K. P. Maresca, L. Francesconi, J. F. Valliant, J. W. Babich and J. Zubieta, *Nucl. Med. Biol.*, 2005, **32**, 1–20.
12 M. K. Levadala, S. R. Banerjee, K. P. Maresca, J. W. Babich and J. Zubieta, *Synthesis*, 2004, **11**, 1759–1766.
13 N. Lazarova, S. James, J. W. Babich and J. Zubieta, *Inorg. Chem. Commun.*, 2004, **7**, 1023–1026.
14 K. Lam, M. Lebl and V. Krchnak, *Chem. Rev.*, 1997, **97**, 411–448.
15 G. Liu and K. Lam, in *Combinatorial Chemistry*, ed. H. Fenniri, Oxford University Press, New York, 2000, pp. 33–49.
16 On-bead sequencing (Edman degradation) analyses provided by the NRC Biotechnology Research Institute, Montréal, QC, Canada.
17 Throughout this article, capital letters denote L-amino acids, lowercase letters indicate D-amino acids.
18 M. Ploug, S. Ostergaard, H. Gardsvoll, K. Kovalski, C. Holst-Hansen, A. Holm, L. Ossowski and K. Dano, *Biochemistry*, 2001, **40**, 12157–12168.
19 S. Knör, S. Sato, T. Huber, A. Morgenstern, F. Bruchertseifer, M. Schmitt, H. Kessler, R. Senekowitsch-Schmidtke, V. Magdolen and C. Seidl, *Eur. J. Nucl. Med. Mol. Imaging*, 2008, **35**, 53–64.
20 L. Schaffer, R. Brissette, J. Spetzler, R. Pillutla, S. Ostergaard, M. Lennick, J. Brandt, P. Fletcher, G. Danielsen, K. Hsiao, A. Andersen, O. Dedova, U. Ribel, T. Hoeg-Jensen, P. Hansen, A. Blume, J. Markussen and N. Goldstein, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 4435.
21 D. H. Lau, L. Guo, R. Liu, A. Song, C. Shao and K. S. Lam, *Biotechnol. Lett.*, 2002, **24**, 497–500.
22 O. H. Aina, J. Marik, R. Liu, D. H. Lau and K. S. Lam, *Mol. Cancer Ther.*, 2005, **4**, 806–813.
23 S. Aggarwal, S. Janssen, R. M. Wadkins, J. L. Harden and S. R. Denmeade, *Biomaterials*, 2005, **26**, 6077–6086.
24 W. C. Eckelman, *Nucl. Med. Biol.*, 1994, **21**, 759–769.
25 W. C. Eckelman, M. R. Kilbourn and C. A. Mathis, *Nucl. Med. Biol.*, 2006, **33**, 449–451.
26 COPAS: Complex Object Parametric Analyzer and Sorter: http://www.unionbio.com/.
27 Sequencing of selected beads was performed on an Applied Biosystems Procise Sequencer at the Hospital for Sick Kids Advanced Protein Technology Centre, Toronto, ON, Canada.
28 R. Alberto, K. Ortner, N. Wheatley, R. Schibli and A. P. .Schubiger, *J. Am. Chem. Soc.*, 2001, **123**, 3135–3136.
29 K. A. Stephenson, S. R. Banerjee, N. McFarlane, D. R. Boreham, K. P. Maresca, J. W. Babich, J. Zubieta and J. F. Valliant, *Can. J. Chem.*, 2005, **83**, 2060–2066.

