ARTICLE

OBC
www.rsc.org/obc

# Using combinatorial methods to arrive at a quantitative structure–stability relationship for a new class of one-armed cationic peptide receptors targeting the C-terminus of the amyloid β-peptide †

**Carsten Schmuck * and Martin Heil**
*Institut für Organische Chemie, Universität Würzburg, Am Hubland, 97074 Würzburg, Germany. E-mail: schmuck@chemie.uni-wuerzburg.de; Fax: + 49 931 8884625*

***Received 21st November 2002, Accepted 20th December 2002***
***First published as an Advance Article on the web 27th January 2003***

A new class of one-armed tripeptide based cationic guanidiniocarbonyl pyrrole receptors is shown to strongly bind the tetrapeptide L-Val-L-Val-L-Ile-L-Ala, representing the C-terminus of the amyloid β-peptide even under polar conditions. A medium sized combinatorial library of 125 receptors was synthesized on a solid support and their binding properties determined on bead using a quantitative fluorescence assay. The binding constants are in the order of $10^3$–$10^4$ $M^{-1}$ (in the presence of a formate counter ion in methanol) for the most efficient ones but differ by more than a factor of 100 among the 125 library members. Based on the binding data of 12 receptors a structure–stability relationship was established for peptide binding by this new receptor class. Complex formation is controlled by a fine balanced interplay of hydrophobic and electrostatic interactions with none of these two interactions alone being strong enough to ensure complexation under these polar conditions.

## Introduction

A detailed quantitative understanding of peptide–peptide interactions not only increases our general knowledge of protein structure and stability but also may help better understand the origin of some of today's most devastating diseases. The self-aggregation of small peptides is responsible both in animals and humans for a variety of neuro degenerative diseases such as Scrapie, BSE, Creutzfeldt-Jakob or Alzheimer's disease (AD).[1] For example, the senile plaques found within the brain of AD patients consist of an insoluble aggregate of a 39–42 residue peptide, called amyloid β-peptide (Aβ).[2] Though the exact correlation between plaque formation and the onset of the disease is not quite clear yet, prevention of Aβ self-aggregation might be a strategy to delay the onset of neuronal degeneration.[3] In the last few years, various small molecules such as rifampicin, Congo red, benzofurans or oligopetides, representing fragments of Aβ itself, have been shown to inhibit Aβ aggregation both *in vitro* and in cell assays.[4] Unfortunately, very little to nothing is known about their mechanism of action, their binding specificities or the molecular basis of their interaction with Aβ.[5] It is not even certain in some cases that a specific complexation between Aβ and the inhibitor actually occurs.[4c] This lack of understanding and of quantitative experimental binding data is a major obstacle in the design of more specific amyloid inhibitors for a potential future therapeutic use. Therefore, a supramolecular structure–stability relationship for molecular interactions with the amyloid peptide is highly needed to identify the structural and thermodynamic parameters that control its self-association and hence also possible interactions with aggregation inhibitors. To arrive at such a quantitative understanding of peptide binding, relative binding constants of a whole series of structurally closely related receptors are needed to correlate binding affinity with structure. In this context, we wish to report here how a medium sized combinatorial library[6] of a new class of one-armed cationic peptide receptors was effectively used to study their binding to a model tetrapeptide representing the C-terminus of Aβ.

† Electronic supplementary information (ESI) available: quantitative on bead binding assay. See http://www.rsc.org/suppdata/ob/b2/b211425a/

## Results and discussion

The general design of our receptor library was based on the observation that the C-terminal sequence of Aβ (-Val[39]-Val[40]-Ile[41]-Ala[42]) is one of two domains being mainly responsible for its self-aggregation.[7] It is thought to promote the formation of aggregated β-sheets stabilized through a combination of H-bonds and hydrophobic interactions.[8] In our receptors (Fig. 1) a tripeptide unit was chosen to provide the necessary

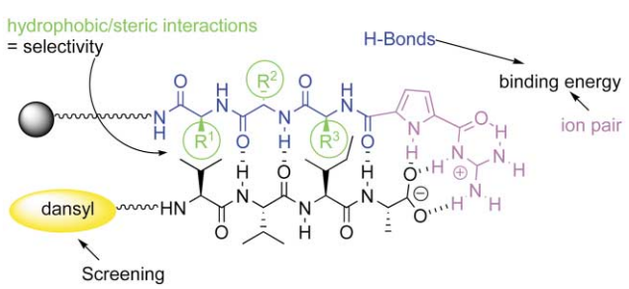


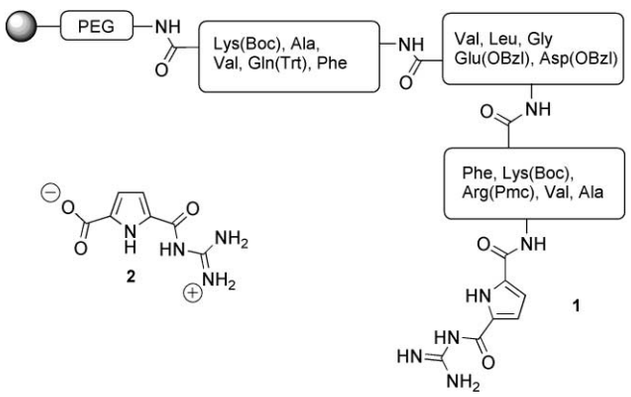

**Fig. 1** A tripeptide-based library of cationic guanidiniocarbonyl pyrrole receptors **1** of the general structure Amino-TentaGel-AA[1]-AA[2]-AA[3]-Gua (AA = amino acid, Gua = guanidiniocarbonyl pyrrole cation) designed for the binding of *L*-Val-*L*-Val-*L*-Ile-*L*-Ala, a tetrapeptide representing the C-terminus of Aβ.

binding sites for the formation of a hydrogen bonded antiparallel β-sheet with the backbone of this tetrapeptide substrate. To ensure strong complexation in polar solvents even for such a short β-sheet, a carboxylate binding site in the form of a cationic guanidiniocarbonyl pyrrole group was introduced. As we could show, these are among the most efficient binding motifs for carboxylates even in aqueous solutions known so far.[9] A combinatorial variation[10] of the three amino acids in the receptor side chain can then be used to identify receptors in which additional hydrophobic and steric interactions between these side chains and the substrate further enhance the binding within the β-sheet and also render the recognition event selective for this specific tetrapeptide.

The peptidic nature of the receptor also allowed the use of well elaborated solid phase peptide chemistry for the synthesis. In a first step, 125 different linear tripeptides were synthesized on Amino-TentaGel™ as the solid support[11] according to a

DOI: 10.1039/b211425a

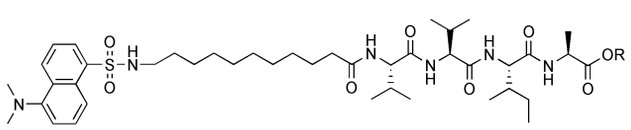
standard Fmoc-protocol using the split-mix-approach[12] in combination with the IRORI™-radio frequency tagging technology.[13] In each of the three coupling steps five different amino acids were used, with ten amino acids in total, selectively chosen to provide a range of structurally varying hydrophobic or steric interactions. In the second step each of the various 125 different tripeptides was coupled with the zwitterion **2** using PyBOP in DMF to yield the desired guanidiniocarbonyl pyrrole receptors **1** (Scheme 1).[14]




**Scheme 1**

The binding properties of these 125 different receptors were then studied in an on bead fluorescence binding assay.[15] For this purpose a fluorescence label in the form of a dansyl group was attached *via* a C11 alkyl spacer to the N-terminus of the tetrapeptide Val-Val-Ile-Ala.[16]

After mixing aliquots of the solid phase bound receptors (formate salts), the entire library was incubated for 18 hours with a 13 μmolar solution of the labeled substrate **3** ($NMe_4^+$ salt) in methanol (Scheme 2). Methanol was chosen as the




**3**: R = H
**4**: R = Me

**Scheme 2**

solvent to provide a polar environment in which both electrostatic and hydrophobic interactions can contribute to the binding and in which both, the receptors and the labeled substrate, are reasonable soluble. After the supernatant solution was washed off the library was screened under UV light. Those receptors that are capable of binding the substrate under these experimental conditions show the characteristic fluorescence activity of the dansyl label whereas those which do not bind the tetrapeptide remain dark. Indeed, the binding assay showed significant and highly selective complexation of the labeled tetrapeptide **3** by approximately 7% of the 125 different receptors (Fig. 2).

In order to assure that the observed fluorescence activity of these beads is really due to a selective complexation of **3** by the receptor a series of control experiments was performed (Fig. 3): A) The labeled tetrapeptide does not bind to the unmodified Amino-TentaGel™. Hence, the observed fluorescence activity is not due to an unspecific interaction with the solid support itself. B) The dansylated spacer alone does not bind to the receptor library, showing that the binding takes place indeed between the peptide part of the substrate and the receptor. C) The percentage of receptors that bind the substrate is concentration dependant: at high concentrations (> 260 μM) nearly all of the library members bind the substrate, which shows that the

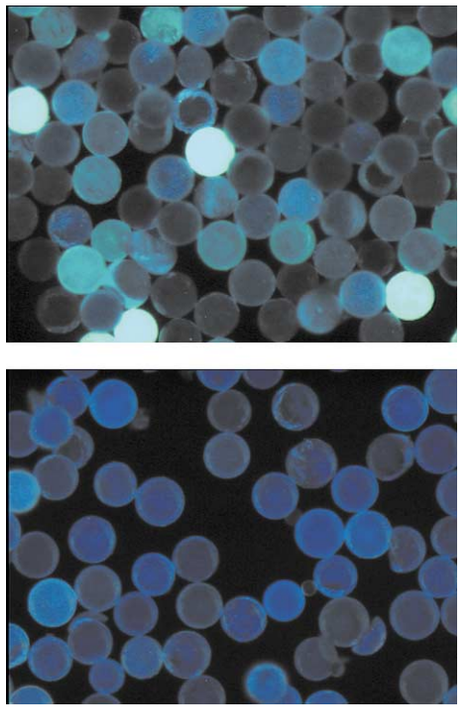


**Fig. 2** On bead binding assay in methanol. The strong fluorescence activity indicates selective binding of the charged substrate **3** to some of the cationic receptors within the library (top). The neutral methyl ester **4** is only weakly bound under similar conditions (bottom).

observed binding specificity is not due to a selective quenching of the dansyl fluorescence rather than a selective binding.

Interestingly, the non-charged methyl ester of the substrate **4** requires much higher concentrations (> 26 μmol) to show only a weak and rather unselective binding to the receptor library (Fig. 2), suggesting that side chain interactions alone are not strong enough in methanol to form a stable complex. However, at low substrate concentrations the negatively charged carboxylate substrate **3** is selectively bound only by some and not all of the receptors, although the ion pairing with the guanidiniocarbonyl pyrrole unit is the same for all the 125 receptors. Therefore, the ionic interaction alone is not sufficient either to provide strong complexation. Hence, the binding of the tetrapeptide **3** by this new receptor class **1** requires both electrostatic as well as hydrophobic interactions at least under the conditions of our assay (μmolar concentrations in methanol). The binding is driven by a combination of various weak noncovalent interactions in a similar way that peptide binding is considered to take place in natural systems.

We then selected 12 library members, that showed both strong (6 receptors) and weak binding (6 receptors). The strong binding ones were selected from an assay performed under dilute conditions, under which only the best receptors are capable of binding, whereas the weak binding ones are those that even at high substrate concentrations do not bind the substrate.[17] Their structures were decoded by a read out of the IRORI™-radio tags (Table 1). As the use of the IRORI™-microreactor technique provides enough material (*ca.* 30 mg resin ≈ 6 μmol of each receptor), a direct quantitative determination of their binding constants on the solid support in a second experiment could be performed (for further information see the ESI †).[18]

As these data show (Table 1), the binding is very efficient with relative association constants of $10^4$ $M^{-1}$ compared to formate for the best receptors, L-Val-L-Val-L-Val-Gua and L-Phe-L-Val-L-Val-Gua, respectively (entry 1 + 2). In this assay the binding is measured in the presence of a formate counter ion, which also binds to the guanidiniocarbonyl pyrrole cation though to a lower extent than other carboxylates.[10] Therefore, the absolute

**Table 1** Relative association constants (in $M^{-1}$) determined in an on bead binding assay for the complexation of **3** by receptors **1** (amino-TentaGel-$AA^1$-$AA^2$-$AA^3$-Gua, formate salt) in methanol [error limits were estimated to be ± 25%]

| Entry | $AA^1$ | $AA^2$ | $AA^3$ | $K_{ass}$ |
|---|---|---|---|---|
| 1 | Phe | Val | Val | 9800 |
| 2 | Val | Val | Val | 9300 |
| 3 | Phe | Val | Lys(Boc) | 4600 |
| 4 | Val | Val | Lys(Boc) | 4500 |
| 5 | Val | Val | Phe | 3700 |
| 6 | Lys(Boc) | Leu | Val | 1800 |
| 7 | Lys(Boc) | Val | Phe | 370 |
| 8 | Val | Glu(OBzl) | Lys(Boc) | 340 |
| 9 | Phe | Glu(OBzl) | Val | 250 |
| 10 | Lys(Boc) | Glu(OBzl) | Lys(Boc) | 130 |
| 11 | Phe | Glu(OBzl) | Phe | 120 |
| 12 | Val | Glu(OBzl) | Phe | 80 |

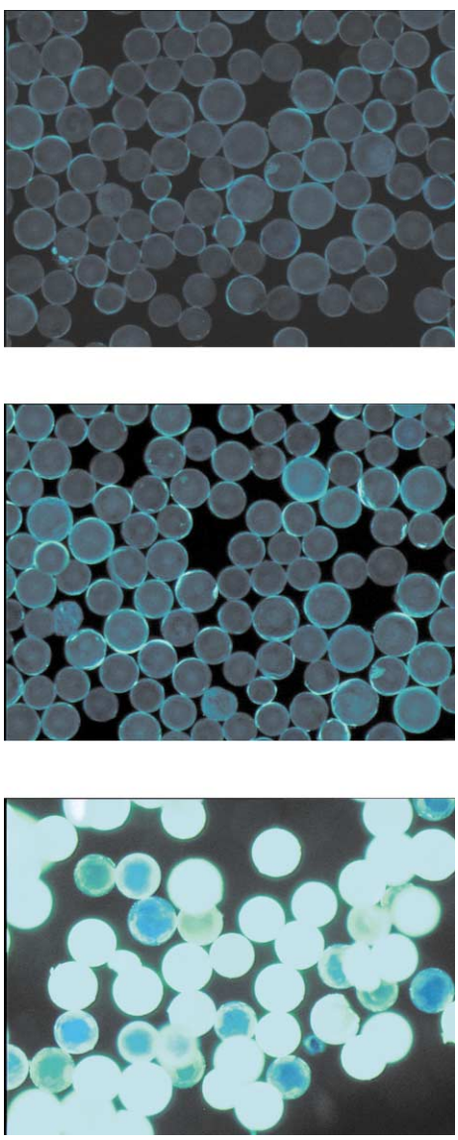


**Fig. 3** Control experiments. Binding of the labeled substrate to Amino-TentaGel™ (top); of the dansyl label to the receptor library (middle) and of the labeled substrate to the receptor library at high concentrations (below).

binding constants between the tetrapeptide **3** and the receptors **1** are actually even stronger than suggested by these relative numbers. Furthermore, small changes in the receptor structure have pronounced effects on the binding properties: Even among these 12 receptors, the association constants differ by a factor of more than 100! Hence, this newly developed class of one-armed cationic peptide receptors not only efficiently binds peptides under highly competitive conditions (polar solvent, formate counter ion) but also allows tuning of the binding affinity over two orders of magnitude by changing its structure; a necessary prerequisite to achieve selective binding of different substrates by this receptor class in the future.

By comparing the relative binding data for these 12 receptors it is now possible to identify structural features within this receptor class that are important for the binding of this peptide. For example, exchanging valine for phenylalanine does not alter the binding affinity significantly: neither at the first position ($AA^1$) where the binding constants stays the same (compare entry 1 + 2; 3 + 4 or 11 + 12), nor at position three ($AA^3$) where the complex stability is only reduced by a factor of two (compare entry 2 + 5 or 9 + 11). In contrast to this, exchanging valine in the first position for N-Boc protected lysine reduces the binding affinity dramatically up to a factor of 10 (compare entry 5 + 7). This is in good agreement with model studies which suggested that hydrophobic interactions with Val(39), which is opposite to this position of the receptor in the proposed β-sheet, are especially important for the Aβ self-aggregation.[4a,19] The side chain of lysine is probably too small and flexible to provide enough hydrophobic shielding of the isopropyl group of Val(39) in the complex. However, upon the same exchange next to the guanidinium cation (position three: $AA^3$), the binding constant just drops by a factor of two (compare entry 1 + 3 or 2 + 4). Probably because of the nearby ion pair, hydrophobic interactions to Val(41) are not equally important for complex stability here. Interestingly, a glutamate benzyl ester at position two ($AA^2$) seems to be incompatible with complex formation at least with this specific substrate. The binding constants are only in the order of $10^2$ $M^{-1}$, a value similar to the one expected for simple ion pairing between the carboxylate and the guanidinium cation. A reason for the pronounced effect of the glutamate at this position is not quite clear yet.

As molecular modelling studies suggest (Macromodel 8.0,[20] Amber*, GB/SA water solvation), the complex formed between the substrate and *e.g.* the receptor L-Val-L-Val-L-Val-Gua indeed adopts a hydrogen bonded beta-sheet with additional hydrophobic and ionic interactions as depicted in Fig. 4.

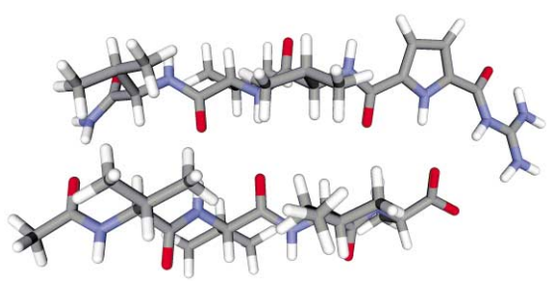


**Fig. 4** Proposed structure for the complex formed between the receptor (top) and the tetrapeptide substrate (below). A hydrogen bonded β-sheet is held together by additional hydrophobic interactions between the side chains and the ion pair between the carboxylate and the guanidinium cation (on the right side).

In conclusion, we have shown here that even within a moderate sized combinatorial library of 125 different members receptors can be identified, that bind with high affinity to a tetrapeptide representing the C-terminal end of the Alzheimer amyloid peptide Aβ. A quantitative structure–stability relationship using the relative binding constants of 12 structurally related receptors revealed a strong effect of side chain interactions at position $AA^1$ opposite to the first valine within the β-sheet complex on its stability. Hence, the recognition of the C-terminal Aβ fragment seems to be controlled by a fine balanced interplay between electrostatic and hydrophobic interactions. Such information helps to increase our general

knowledge of the molecular basis of peptide–molecule interactions and can be useful for the design of specific amyloid inhibitors in the future.

## Acknowledgements

Financial support of this work by the DFG (SCHM 1501/2–1) and the Fonds der Chemischen Industrie is gratefully acknowledged. We thank Professor Dr. A. Berkessel (Universität Köln) for technical support with the combinatorial assays.

## References

1 (*a*) P. Kurosinski, M. Guggisberg and J. Götz, *Trends Mol. Med.*, 2002, **8**, 3–5; (*b*) L. Dumery, F. Bourdel, Y. Sousssan, A. Fialkowsky, S. Viale, P. Nicolas and M. Reboud-Ravaux, *Pathol. Biol.*, 2001, **49**, 72–85; (*c*) R. Knight, *Proteomics*, 2001, **1**, 763–766; (*d*) F. Edenhofer, S. Weiss, E.-L. Winnacker and M. Famulok, *Angew. Chem.*, 1997, **109**, 1748–1769.

2 (*a*) E. Zerovnik, *Eur. J. Biochem.*, 2002, **269**, 3362–3371; (*b*) J. Hardy and D. J. Selkoe, *Science*, 2002, **297**, 353–356; (*c*) B. Austen and M. Manca, *Chem. Brit.*, 2000, 28–31; (*d*) L. Gopinath, *Chem. Brit.*, 1998, 38–40; (*e*) P. T. Lansbury, Jr., *Acc. Chem. Res.*, 1996, **29**, 317–321; (*f*) B. A. Yankner, *Neuron*, 1996, **16**, 921–932.

3 P. T. Lansbury, Jr., *Curr. Opin. Chem. Biol.*, 1997, **1**, 260–267.

4 (*a*) K. Watanabe, K. Nakamura, S. Akikusa, T. Okada, M. Kodaka, T. Konakahara and H. Okuno, *Biochem. Biophys. Res. Commun.*, 2002, **290**, 121–124; (*b*) C. Hetenyi, Z. Szabo, E. Klement, Z. Datki, T. Körtvelyesi, M. Zarandi and B. Penke, *Biochem. Biophys. Res. Commun.*, 2002, **292**, 931–936; (*c*) T. L. Lowe, A. Strzelec, L. L. Kiessling and R. M. Murphy, *Biochemistry*, 2001, **40**, 7882–7889; (*d*) C. Soto, *J. Mol. Med.*, 1999, **5**, 343–350.

5 (*a*) Y. Takahashi, A. Ueno and H. Mihara, *ChemBioChem*, 2002, **3**, 637–642; (*b*) K. Watanabe, T. Segawa, K. Nakamura, M. Kodaka, T. Konakahara and H. Okuno, *J. Pept. Res.*, 2001, **58**, 342–346; (*c*) L. O. Tjernberg, D. J. E. Callaway, A. Tjernberg, S. Hahne, C. Lilliehöök, L. Terenius, J. Thyberg and C. Nordstedt, *J. Biol. Chem.*, 1999, **18**, 12619–12625.

6 For related work on other peptide substrates see *e.g.*: (*a*) K. Jensen, T. M. Braxmeier, M. Demarcus, J. G. Frey and J. D. Kilburn, *Chem. Eur. J.*, 2002, **8**, 1300–1309; (*b*) R. Xuo, G. Greiveldinger, L. E. Marenus, A. Cooper and J. A. Ellman, *J. Am. Chem. Soc.*, 1999, **121**, 4898–4899; (*c*) For a review see: M. W. Peczuh and A. D. Hamilton, *Chem. Rev.*, 2000, **100**, 2479–2493.

7 J. T. Jarret, E. P. Berger and P. T. Lansbury, Jr., *Biochemistry*, 1993, **32**, 4693–4697.

8 M. O. Chaney, S. D. Webster, Y.-M. Kuo and A. E. Roher, *Protein Engineering*, 1998, **11**, 761–767.

9 (*a*) C. Schmuck, *Chem. Commun.*, 1999, 843–844; (*b*) C. Schmuck, *Eur. J. Org. Chem.*, 1999, 2397–2403; (*c*) C. Schmuck, *Chem. Eur. J.*, 2000, **6**, 709–718.

10 (*a*) L. A. Thompson and J. A. Ellman, *Chem. Rev.*, 1996, **96**, 555–600; (*b*) F. Balkenhohl, C. von dem Busche-Hünnefeld, A. Lansky and C. Zechel, *Angew. Chem.*, 1996, **108**, 2436–2487; (*c*) G. Lowe, *Chem. Soc. Rev.*, 1995, 309–317.

11 TentaGel was used as the solid support to allow a screening of the receptor library not only in methanol but later on for also other studies in aqueous polar solvents.

12 K. S. Lam, M. Lebl and V. Krchnak, *Chem. Rev.*, 1997, **97**, 411–448.

13 A. W. Czarnik, *Curr. Opin. Chem. Biol.*, 1997, **1**, 60–66.

14 Due to the low solubility of zwitter ion **2** in DMF a modified coupling protocol was necessary, using an excess of 2 equivalents of **2**, PyBOP as the coupling reagent in 5% NMM in DMF at room temperature and reaction times of 20 h. To ensure quantitative yields the coupling was performed twice.

15 W. C. Still, *Acc. Chem. Res.*, 1996, **29**, 155–163.

16 The synthesis of the dansylated substrate was performed on solid support. Using a Fmoc-protocol $H_2N$-Val-Val-Ile-Ala-OH was synthesized on Wang resin beginning at the C-terminus. The free amino end was then reacted with the dansylated C11 spacer to give, after cleavage (TFA, $CH_2Cl_2$ 50:50), the desired fluorescence labeled substrate **3** in analytically pure form. The corresponding methyl ester **4**, which was needed as a control, was obtained *via* esterification of the free acid using methanol and thionyl chloride.

17 The well binding receptors were selected using a 13 μmolar solution of the labeled substrate, whereas the weak binding receptors were selected from an assay using a 260 μmolar solution.

18 S. S. Yoon and W. C. Still, *Tetrahedron*, 1995, **51**, 567–578.

19 P. T. Lansbury, Jr., P. R. Costa, J. M. Griffiths, E. J. Simon, M. Auger, K. J. Halverson, D. A. Kocisko, Z. S. Hendsch, T. T. Ashbury, R. G. S. Spencer, B. Tidor and R. G. Griffin, *Nat. Struct. Biol.*, 1995, **2**, 990–998.

20 F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caufiled, G. Chang, T. Hendrickson and W. C. Still, *J. Comput. Chem.*, 1990, **11**, 440–467.