Avidin and streptavidin ligands based on the glycoluril bicyclic system†

Pedro Hidalgo-Fernández, Eva Ayet, Ivan Canal and Joan-Antoni Farrera*

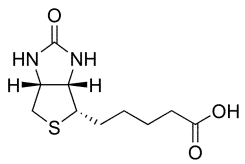
Received 7th April 2006, Accepted 26th June 2006
First published as an Advance Article on the web 14th July 2006
DOI: 10.1039/b605081f

Glycoluril derivatives with a carboxylic acid side chain have been synthesized and shown to bind to both avidin and streptavidin. Introduction of a valerate side chain in glycoluril led to an increased binding to both proteins only when the valerate group was bound to a N atom and with the proper stereochemistry [(+)-enantiomer]. On the other hand, introduction of the valerate side chain either on the bridgehead carbon or on the N atom with the opposite stereochemistry [(−)-enantiomer] led to a decrease in binding constant compared with unsubstituted glycoluril. Direct spectrophotometric competitive titration of each protein with a racemic ligand allowed measurement of the enantioselectivity of the ligand–protein complexation, together with the binding constant of the two enantiomers. In the case of the N-substituted glycoluril, the extension of the side chain by one methylene group, from valerate to caproate, led to an increase in the binding constant to both proteins. Docking studies using AutoDock 3.05 have been performed in order to predict the binding modes of these ligands to streptavidin. The effect of the stereochemistry and the position of the side chain on the binding constant to streptavidin is discussed in view of the predicted binding modes.

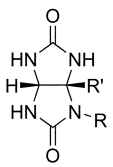
Introduction

Avidin (Av) and streptavidin (Sav) are tetrameric proteins, with one binding site per subunit, widely used in biochemical applications.[1] The basis of this prevalent use is the high association constant of the complexes that biotin (**1**) forms with both Av ($K_a \approx 10^{15}$ M^{-1}) and Sav ($K_a \approx 2.5 \times 10^{13}$ M^{-1}), which makes the binding process considered to be almost irreversible.[2] This strong binding between biotin (**1**) and Av or Sav has promoted the use of these systems in the construction of different supramolecular structures.[3–10]

Binding studies of biotin analogues to Av and Sav have suggested a major contribution from the ureido group, with a lower contribution from the valerate side chain and the thiolane ring.[11–15] X-Ray diffraction studies of both Av-**1** and Sav-**1** complexes showed a similar network of hydrogen-bonding interactions between the ureido group of biotin (**1**) and five protein residues, whereas the carboxylate group of biotin formed hydrogen-bonding interactions with five protein residues in the case of Av and with only two protein residues in the case of Sav. This different hydrogen-bonding network of the carboxylate, together with differences in the hydrophobic interactions to aromatic residues, is consistent with the stronger binding of biotin (**1**) to Av compared with Sav.[16]



1



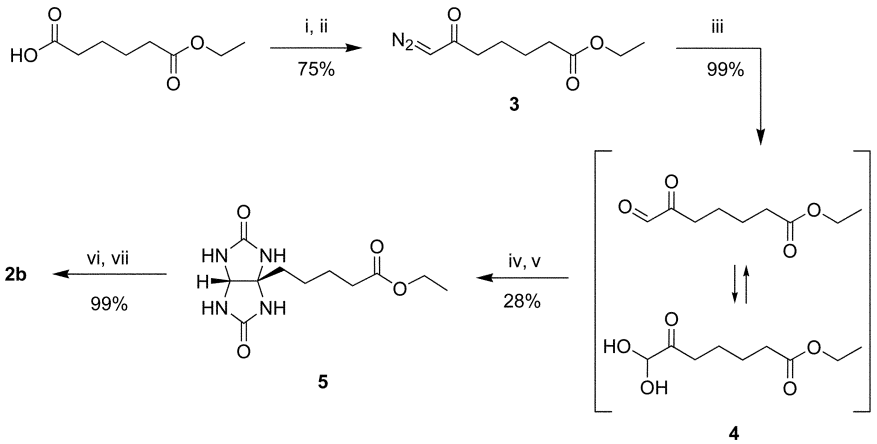
2a, R = H, R' = H
b, R = H R' = -$(CH_2)_4$-COOH
c, R = -$(CH_2)_4$-COOH R' = H
d, R = -$(CH_2)_5$-COOH R' = H

Av and Sav can also bind a large number of molecules with lower affinity than for biotin. One of these compounds is glycoluril (**2a**),[17] which was shown to bind Sav with an association constant of 4×10^5 M^{-1}.[18] X-Ray diffraction studies of the complex Sav–glycoluril showed hydrogen-bonding interactions between one of the ureido groups of glycoluril (**2a**) and the same five protein residues involved in the binding of biotin (**1**). The oxygen atom of the second ureido group of glycoluril also contributes to the binding with a hydrogen-bonding interaction to the same protein residue that binds to the sulfur atom of biotin. Considering the similar hydrogen-bonding pattern in both Sav–biotin and Sav–glycoluril complexes, it was concluded that the differences in binding constant between both complexes predominantly reflected the missing interactions due to the absence of the valerate side chain in glycoluril (**2a**).

The objective of the present work was to quantify the contribution to the ΔG° of the binding of a carboxylic acid side chain that is covalently linked to either a bridgehead carbon atom or to a nitrogen atom of glycoluril. For that reason we synthesized 5-(3,7-dioxo-2,4,6,8-tetraazabicyclo[3.3.0]oct-1-yl)pentanoic acid (**2b**), 5-(3,7-dioxo-2,4,6,8-tetraazabicyclo[3.3.0]oct-2-yl)pentanoic acid (**2c**),[19] and 6-(3,7-dioxo-2,4,6,8-tetraazabicyclo[3.3.0]oct-2-yl)hexanoic acid (**2d**).

Departament de Quimica Organica, Facultat de Química, Universitat de Barcelona, Martí i Franqués 1, 08028-Barcelona, Spain. E-mail: jfarrera@ub.edu

† Electronic supplementary information (ESI) available: binding models for competitive titrations with ligands, and binding isotherms for competitive titrations of (S)Av–HABA complexes with glycoluril-type ligands. See DOI: 10.1039/b605081f.



Scheme 1 *Reagents and conditions*: (i) $SOCl_2$; (ii) CH_2N_2, hexane–Et_2O (1 : 1), rt,1 h; (iii) dimethyldioxirane, acetone, 0 °C; (iv) H_2NCONH_2, H_2O–EtOH (2 : 1), pH = 3, 4 °C, 3 d; (v) pH = 1, 4 °C, 5 d; (vi) KOH–H_2O, rt, 24 h; (vii) pH = 1–2.

Results and discussion

Compound **2b** was obtained from 5-ethoxycarbonylpentanoic acid by reacting the corresponding acyl chloride with diazomethane to give the diazoketone **3** in 75% yield (Scheme 1). The diazoketone **3** was subsequently oxidized with dimethyldioxirane in acetone to yield glyoxal **4**, together with the corresponding hydrate.

Attempts to condense compound **4** to urea in standard conditions for the synthesis of glycolurils (toluene or benzene, with a catalytic amount of TFA and azeotropic removal of water) didn't afford glycoluril **5** due to competing side reactions involving glyoxal **4** (Table 1). These side reactions of glyoxal **4** could be minimized by reacting urea and glyoxal **4** at low temperature. However, in spite of water being formed in the condensation of urea and glyoxals, this type of reaction can proceed in aqueous or hydroalcoholic media with moderate to good yields.[20–22] Depending on the acidity of the aqueous medium the condensation of urea and glyoxal affords either glycoluril, when pH = 1–2, or the intermediate 4,5-dihydroxyimidazolidin-2-one, when pH = 3–4. The best result, 28% yield, was obtained when urea and glyoxal **4** were allowed to react in EtOH–H_2O (1 : 2) at 4 °C in two stages: first at pH = 3 in order to form the corresponding 4,5-dihydroxyimidazolidin-2-one, and then at pH = 1 in order to form glycoluril **5** (Table 1). Finally, saponification of the ester group of **5** afforded ligand **2b** in almost quantitative yield.

Compounds **2c**[19] and **2d** were obtained from ω-aminoacids which, by reaction with potassium cyanate in aqueous medium, afforded the corresponding ω-ureidocarboxylic acids. Condensation of these ω-ureidocarboxylic acids to 4,5-dihydroxyimidazolidin-2-one[21] afforded the ligands **2c** and **2d** in 45 and 50% yield respectively. Unlike compound **2b**, which is achiral, compounds **2c** and **2d** are chiral. Since the two enantiomers of a racemic ligand are expected to have different binding constants to a given protein, a small amount of racemic **2c** was resolved into its enantiomers by RP-HPLC using a chiral stationary phase.[23]

In order to prove that these ligands bind to Av and Sav we used the dye 2-(4′-hydroxyphenylazo)benzoic acid (HABA), which has been shown to bind to both Av and Sav at the same binding site as biotin. This dye, in its unbound form, has an absorption maximum at 348 nm, which shifts to 500 nm when bound to Av or Sav.[24] Addition of any of these ligands to an aqueous solution of HABA and either Av or Sav resulted in a decrease in the absorption at 500 nm and the corresponding increase in the absorption at 348 nm, proving that all of them bind to the same binding sites as HABA and biotin. Since the binding constants of HABA to Av and Sav were already known ($K_a = 1.7 \times 10^5$ and 7.3×10^3 M^{-1} respectively),[1,13] the binding constants of the new ligands to Av and Sav should also be measurable in spectrophotometric competition experiments with HABA. For every ligand, the plot of absorbance at 500 nm (A_{500}) *vs.* the ratio [ligand]/[subunits] was fitted to a 1 : 1 binding model assuming that each subunit of the protein behaves as an independent binding site, without cooperative effects.[25,26] The reasonably good fits obtained (shown in Fig. 1–3 and in the ESI†) between the experimental and calculated curves support our assumption of independent behaviour between the binding sites.

In order to compare binding constants of different compounds it is preferable that all of them are measured with the same technique and under the same experimental conditions. For that reason we measured the binding constant of glycoluril (**2a**) to Sav by the spectrophotometric competitive titration method, yielding a value which was *circa* one order of magnitude lower than the value

Table 1 Optimization of the synthesis of **5**

Entry	Solvent	Catalyst	Temperature/°C	Time	Yield (%)
1	Benzene	TFA	Reflux	16 h	*ca.* 1
2	H_2O–EtOH (2 : 1)	HCl (pH = 1)	50	16 h	*ca.* 1
3	H_2O–EtOH (2 : 1)	HCl (pH = 1)	25	24 h	5
4	H_2O–EtOH (2 : 1)	(i) HCl (pH = 3) (ii) HCl (pH = 1)	4 4	3 d 5 d	28

Table 2 Binding constants of glycoluril and its derivates to avidin and streptavidin[a]

	Avidin			Streptavidin		
Ligand	K_a/M^{-1}	ΔG/kJ mol^{-1}	S	K_a/M^{-1}	ΔG/kJ mol^{-1}	S
2a	$2.2 (\pm 0.6) \times 10^6$	-36.2 ± 0.7		$4.7 (\pm 0.5) \times 10^4$	-26.5 ± 0.3	
2b	$8.5 (\pm 1.0) \times 10^3$	-22.4 ± 0.3		76 ± 10	-10.7 ± 0.3	
(+)-**2c**	$2.6 (\pm 1.0) \times 10^7$	-42.3 ± 1.0	217	$1.8 (\pm 0.8) \times 10^6$	-35.7 ± 1.2	164
	$3.5 (\pm 2.0) \times 10^7$ [b]	-43.0 ± 1.6 [b]		$2.8 (\pm 1.0) \times 10^6$ [b]	-36.8 ± 0.9 [b]	
(−)-**2c**	$1.2 (\pm 0.3) \times 10^5$	-29.0 ± 0.6		$1.1 (\pm 0.2) \times 10^4$	-23.1 ± 0.5	
	$2.6 (\pm 2.0) \times 10^5$ [b]	-30.9 ± 2.5 [b]		$7.0 (\pm 2.0) \times 10^3$ [b]	-21.9 ± 0.7 [b]	
(±)-**2d**	$4.1 (\pm 2.0) \times 10^7$ [b]	-43.4 ± 1.3 [b]	103	$4.5 (\pm 2.0) \times 10^6$ [b]	-38.0 ± 1.7 [b]	265
	$4.0 (\pm 2.0) \times 10^5$ [b]	-32.0 ± 1.4 [b]		$1.7 (\pm 1.0) \times 10^4$ [b]	-24.1 ± 1.2 [b]	

[a] Determined by spectrophotometric competition with HABA, in 0.1 M phosphate buffer at pH = 7.0. [b] Obtained from the competitive tritation of the protein–HABA complex with the racemic ligand.

obtained by surface plasmon resonance using other experimental conditions[18] (Table 2). On the other hand, owing to the lack of a carboxylic acid side chain in glycoluril (**2a**), it was expected that a similar binding constant to both Av and Sav would be found. However, the binding constant to Av resulted in being almost two orders of magnitude higher than to Sav, as also happens in the complexes of biotin and other ligands.[1,14,24,27] In the case of glycoluril this difference in the binding constants to both proteins can only be ascribed to differences in the hydrophobic and van der Waals interactions of the bicyclic system, and to differences in the hydrogen-bonding interactions involving the ureido groups. Analogously, in the case of biotin, apart from the well established difference of interactions of the valerate side chain towards Av and Sav,[16] the interactions involving the bicyclic system of biotin could also make an important contribution to the difference of its binding constant to both proteins.

The glycoluril derivatives with a valerate side chain bound to a bridgehead carbon atom (**2b**) or to a nitrogen atom [(+)-**2c** and (−)-**2c**] also have a higher binding affinity to Av than to Sav (Fig. 1 and Table 2). However, the difference in the ΔG of binding ($\Delta\Delta G$) for a given ligand to Av and Sav is about two times smaller in the N-substituted glycolurils (+)-**2c** and (−)-**2c** than in the C-substituted glycoluril **2b** (6.6, 5.9 and 11.7 kJ mol^{-1} respectively).

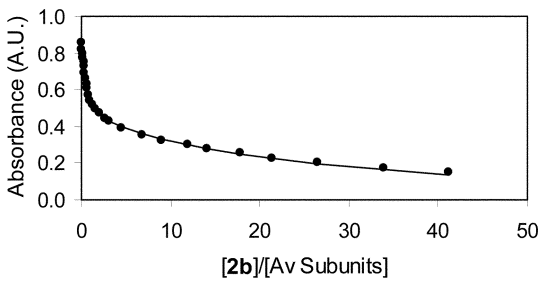


Fig. 1 Absorption change at 500 nm in the titration of Av 9.2 μM (tetramer) and HABA 34.7 μM with ligand **2b** in phosphate buffer 0.1 M, pH = 7.0. The solid line represents the fit of the data to the (1 : 1) binding model.

For the chiral compounds **2c** the (+)-enantiomer was found to bind to both Av and Sav about 13 kJ mol^{-1} stronger than the (−)-enantiomer (Fig. 2 and Table 2), which is an indication of the better fit of the (+)-enantiomer inside the binding pocket of both proteins. The position and stereochemistry of the bicyclic system where the valerate side chain is bound has an important

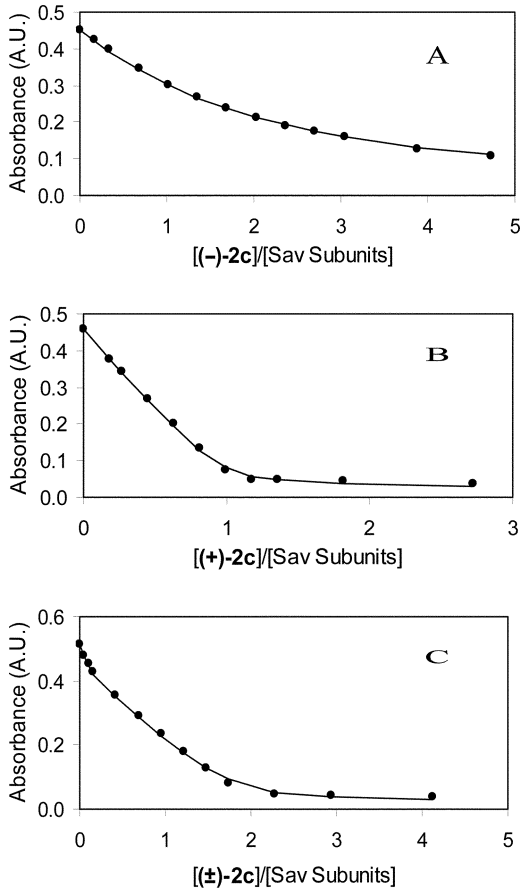


Fig. 2 Absorption change at 500 nm in the titrations of the complex Sav–HABA with ligands **2c** in phosphate buffer 0.1 M at pH = 7.0. **A**: Sav 11.8 μM (tetramer) and HABA 59.7 μM with ligand (−)-**2c**. **B**: Sav 11.9 μM (tetramer) and HABA 59.7 μM with ligand (+)-**2c**. **C**: Sav 12.1 μM (tetramer) and HABA 67.6 μM with ligand (±)-**2c**. The solid line represents the fit of the data to the (1 : 1) binding model, and to the model considering different binding constants for each enantiomer in titration **C**.

influence on the free energy of binding of the resulting ligand to Av and Sav. The binding is weakest when the valerate side chain is bound to a glycoluril bridgehead carbon (**2b**). On the other side, the strongest binding is obtained when the valerate side chain is bound to a glycoluril nitrogen atom with the proper configuration in the stereogenic centers [(+)-**2c**]. The decrease in the free energy of binding on going from **2b** to (+)-**2c** is of 19.9 kJ mol^{-1} for

Av and of 25.0 kJ mol^{-1} for Sav. Smaller decreases in the free energy of binding are obtained on going from **2b** to (−)-**2c**, being 6.6 kJ mol^{-1} for Av and of 12.4 kJ mol^{-1} for Sav. It is clear from these results that introduction of a valerate side chain in glycoluril will lead to an increase in the binding constant to Av and Sav only when the side chain is bound to a nitrogen atom and with the proper stereochemistry [(+)-enantiomer]. This increase in binding afinitty could be due to new hydrogen-bonding interactions of the carboxylate group together with van der Waals interactions of the valerate side chain.[25,28] In other cases, either with the valerate side chain bound to the bridgehead carbon or to the nitrogen atom with the wrong stereochemistry [(−)-enantiomer], there will be a decrease in binding constant compared to unsubstituted glycoluril (**2a**).

Although the side chain in **2b** and **2c** – valerate – is the same as in biotin (**1**), the constitution and stereochemistry at the linkage point between the bicyclic system and the side chain is different for these compounds. Owing to these differences, the valerate side chain might not be the most appropriate one for substituted glycolurils in order to make hydrogen-bonding interactions to specific residues of Av and Sav. In the case of biotin it has been found that shortening or extending the carboxylic acid side chain by a single methylene group leads to a markedly different network of hydrogen-bonding interactions with Av and Sav.[16] For that reason we were interested in finding out whether an extension of the carboxylic acid side chain of N-substituted glycolurils would lead to an increase in the binding constant to Av or to Sav. Ligand **2d**, which has five methylene groups in the side chain, was obtained as a racemate. We were interested to know, without having to ressolve the racemic **2d**, whether any one of its enantiomers would bind to Av or to Sav stronger than the enantiomers of **2c**, and also the enantioselectivity of the binding to each protein.

There are many examples in the literature on the simultaneous determination of two binding constants in a single titration of an enantiomerically pure host with a racemic guest, or *vice versa* an enantiomerically pure guest with a racemic host, when the two diastereomeric complexes show separate signals. This is the case in ^{1}H NMR spectroscopy of some host–guest complexes[29] and affinity capillary electrophoresis of some diastereomeric ion-pairs.[30,31] ^{1}H NMR titration of a racemic host with a racemic guest leading to an averaged chemical shift can also give the two binding constants of the diastereomeric complexes, but only when the ratio between both binding constants can be obtained from a separate experiment.[32] However, it is very common to have a racemic ligand that binds to an enantiomerically pure chiral receptor (*e.g.* a protein) leading to a single averaged response for the two diastereomeric complexes. In this case it is also possible to simultaneously obtain the binding constants of both enantiomeric ligands by means of a single titration of the enantiomerically pure receptor with the racemic ligand.[33] Obviously this methodology does not tell us to which enantiomer of the ligand each binding constant corresponds. But it has the great advantage of giving us the binding constant of both enantiomeric ligands, as well as the enantioselectivity, in a single titration of the receptor with the racemic ligand, without having to separate the two enantiomers. The application of this methodology should allow the fast screening of racemic ligands. Only those racemic ligands showing the appropriate binding constants (or activities) and enantioselectivity would proceed to the next, time- and resources-consuming, stage of racemic resolution. Finally, titration with the enantiomerically pure optimal ligands would allow confirmation of the previously found binding constants as well as to their assignment to the corresponding enantiomer of the ligand.

We have applied this methodology to the competitive spectrophotometric determination of binding constants, although its general character makes it also applicable to other methods, either direct or competitive. In order to check this methodology we first used racemic **2c**, because it had already been resolved and the binding constants of both enantiomers to Av and Sav had been independently measured. Competitive titration of the complex Sav–HABA with racemic **2c** afforded an experimental curve (Fig. 2C) that was fitted to a 1 : 1 binding model. But this time the binding model included two different binding constants – one for each enantiomer of the ligand. Non-linear fitting of the calculated absorbance to the experimental absorbance at 500 nm led to the binding constants of 2.8×10^6 and 7.0×10^3 M^{-1} for the two enantiomers, which agree reasonably well with the values obtained in the titrations with the pure enantiomers (Table 2). In order to further check this method with another protein, the complex Av–HABA was titrated with racemic **2c**. Non-linear fitting of the experimental curve (see ESI†) to the same 1 : 1 binding model led to the binding constants of 3.5×10^7 and 2.6×10^5 M^{-1} for the two enantiomers (Table 2). The reasonably good agreement between the binding constants obtained by titration with each enantiomerically pure ligand and by a single titration with the corresponding racemate, with two different proteins, supports the validity of this methodology.

When this methodology was applied to the competitive titration of Av–HABA with racemic **2d**, a K_{major} of 4.1×10^7 M^{-1} and a K_{minor} of 4.0×10^5 M^{-1} were obtained. In the most likely case where the K_{major} corresponded to the **2d**-enantiomer with the same stereochemistry as (+)-**2c**, the extension of the side chain by one methylene group would lead to an increase in the binding constant of both enantiomers to Av. The increase in stability ($\Delta\Delta G$) of the Av complexes would be higher for the weak-binding enantiomer (3.0 kJ mol^{-1}) than for the strong-binding enantiomer (1.1 kJ mol^{-1}).

In an analogous competitive titration of Sav–HABA with racemic **2d** (Fig. 3), a K_{major} of 4.5×10^6 M^{-1} and a K_{minor} of 1.7×10^4 M^{-1} were obtained. Assuming that the K_{major} corresponded to the **2d**-enantiomer with the same stereochemistry than (+)-**2c**, the

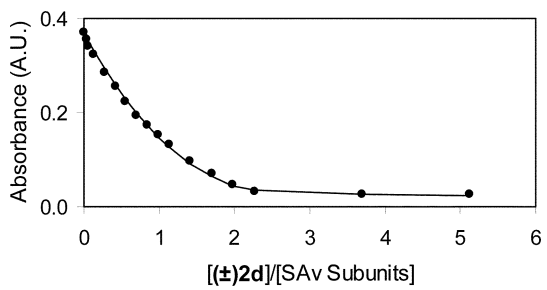


Fig. 3 Absorption change at 500 nm in the titration of Sav 10.6 μM (tetramer) and HABA 52.7 μM with ligand (±)-**2d** in phosphate buffer 0.1 M, pH = 7.0. The solid line represents the fit of the data to the (1 : 1) binding model considering different binding constants for each enantiomer.

extension of the side chain by one methylene group would also lead to an increase in the binding constant of both enantiomers to Sav. But now the increase in stability ($\Delta\Delta G$) of the Sav complexes would be higher for the strong-binding enantiomer (2.3 kJ mol^{-1}) than for the weak-binding enantiomer (1.0 kJ mol^{-1}).

Apart from this increase in the binding constant to both Av and Sav, the extension of the side chain by one methylene group leads also to changes in the enantioselectivity S (defined as the ratio of the binding constants of both enantiomers to a given receptor). While in the case of Av, the extension of the side chain by one methylene group leads to decrease in the enantioselectivity, in the case of Sav it has the opposite effect (Table 2). This different behaviour is consistent with the markedly different effect of the extension of the side chain of biotin on the network of hydrogen-bonding interactions to Av, compared to Sav.[16]

We were interested in predicting the binding mode of ligands **2a**–**d** to Sav, in order to explain the major effects of the stereochemistry and the position of the side chain on the binding constant. Docking programs have been used to predict the binding site and bound conformation of flexible ligands to macromolecules of known structure, which are maintained rigid. Among these programs, AutoDock[34–36] (Lamarckian genetic algorithm) has been successfully used in reproducing the experimental binding conformation of a wide range of ligand–protein complexes.[37] In this sense, AutoDock has been shown to reproduce the experimentally observed binding conformation of biotin (**1**) to Sav with a rmsd of 0.66 Å.[34] Since the binding of biotin (**1**) to Sav leads to an ordering of a surface protein loop,[12,38] we assumed that the protein conformation in the Sav-**2** complexes would be more similar to that in the biotin complex (Sav-**1**) than to that in the uncomplexed Sav. Consequently, we proceeded to dock ligands **2a**–**d** to the complexed Sav conformation.

In the first place, biotin (**1**) and glycoluril (**2a**) were docked to Sav in order to test the reliability of the docking formalism and of the assumed protein conformation. The predicted binding mode of biotin (**1**) and glycoluril (**2a**) reproduced their crystallographic complex coordinates[18,38] with rmsd values of 0.37 and 1.29 Å respectively. When **2b** was docked to Sav, the best scored binding mode had the bicyclic system rotated approximately 107° about the bridge bond, compared with the bicyclic systems of bound biotin (**1**) and bound glycoluril (**2a**). In this binding mode (Fig. 4A) only one of the urea groups in **2b** was almost superimposable to that in biotin (**1**) or in glycoluril (**2a**), with the possibility of forming similar hydrogen-bonding interactions. Besides, the valerate side chain, although following a similar direction to that in biotin, in **2a** might not be long enough to form hydrogen-bonding interactions with Asn 49 and Ser 88. On the other hand, the most populated cluster of docked **2b**, which did not contain the lowest energy conformation, had binding modes in which the bicyclic system of **2b** was slightly tilted in relation to that of biotin (Fig. 4B). The lower stability of this binding mode might come from less effective hydrogen bonding and van der Waals interactions.

When (1*R*,5*S*)-**2c** and (1*R*,5*S*)-**2d**, the ligands with a stereochemistry more similar to that of biotin (**1**), were docked to Sav, the predicted binding mode showed for both ligands an orientation within the binding site similar to that of crystallographic biotin (**1**) (Fig. 4C). For these ligands the best scored conformation corresponded to the most populated cluster.

In the case of (1*S*,5*R*)-**2c**, maintaining the valerate side chain in a position similar to that of bound biotin would imply that its bicyclic system could not be superimposable to that of its enantiomer, or to that of bound biotin. Instead, the docking results predicted that the bicyclic system of (1*S*,5*R*)-**2c** would be almost superimposable to that of its enantiomer (Fig. 4D), while the valerate side chain would lie in a different position to that of its enantiomer, or to that of biotin. In this binding mode, apart from potentially less favourable van der Waals interactions, one hydrogen-bonding interaction might be lost because of the valerate side chain substitution on the N atom close to Ser 45.

The results obtained by docking studies could explain the major differences in binding affinities, although it is difficult to extract conclusions about the predicted absolute binding free energies and hydrogen-bonding distances because the structure of the protein was maintained rigid. More detailed studies based on molecular dynamic methods are in progress.

Conclusion

In this paper we describe the synthesis of glycoluril derivatives with a carboxylic acid side chain, showing that they bind to both avidin and streptavidin at the same binding site as biotin. Glycoluril derivatives with a valerate side chain bound either to a bridgehead carbon atom or to a nitrogen atom have a higher binding affinity to Av than to Sav. Moreover, the position and stereochemistry of the bicyclic system where the valerate side chain is bound has an important influence on the binding affinity of the resulting ligand. Introduction of a valerate side chain in glycoluril led to an increased binding to both proteins only when the valerate group was bound to a N atom and with the proper stereochemistry [(+)-enantiomer]. On the other hand, introduction of the valerate side

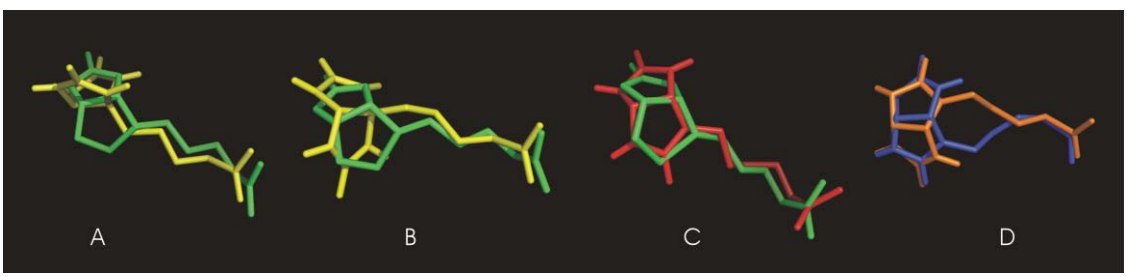


Fig. 4 Superimposition of binding modes of ligands **2b**–**d** with Sav as found by AutoDock. A: lowest energy binding mode of **2b** (yellow) and crystallographic biotin (green); B: most populated cluster of **2b** (yellow) and crystallographic biotin (green); C: lowest energy binding mode of (1*R*,5*S*)-**2d** (red) and crystallographic biotin (green); D: lowest energy binding modes of (1*R*,5*S*)-**2c** (blue) and (1*S*,5*R*)-**2c** (orange).

chain either on the bridgehead carbon or on the N atom with the opposite stereochemistry [(−)-enantiomer] led to a decrease in binding constant compared to the unsubstituted glycoluril. Direct spectrophotometric competitive titration of each protein with a racemic ligand allowed measurement of the enantioselectivity of the ligand–protein complexation, together with the binding constant of the two enantiomers. In the case of the N-substituted glycoluril, the extension of the side chain by one methylene group, from valerate to caproate, led to an increase in the binding constant to both proteins. The effect of the stereochemistry and the position of the side chain on the binding constant to streptavidin could be explained in view of the binding modes predicted from docking studies using AutoDock 3.05.

Experimental

General

^{1}H NMR spectra were recorded on a Varian Gemini (200 MHz) or Varian Mercury spectrometer (400 MHz) in the indicated solvent. Chemical shifts were reported in parts per million (ppm) relative to tetramethylsilane (0.0 ppm), $CDCl_3$ (7.26 ppm), or DMSO-d_6 (2.50 ppm) as an internal standard. ^{13}C NMR spectra were recorded on a Varian Gemini (50 MHz) or Varian Mercury spectrometer (100 MHz). Chemical shifts were given in parts1 per million (ppm) relative to DMSO-d_6 (39.5 ppm) as an internal standard. Infrared spectra were recorded on a Perkin-Elmer 681 instrument. UV-Vis spectra were recorded on a Varian CARY 500 Scan spectrometer. Mass spectra were recorded on a Waters Micromass ZQ spectrometer for ESI spectra. Analytical thin layer chromatography was performed on Merck Silica Gel 60 F254 plates and visualized with UV light and developed by exposure to Cl_2 (gas) prior to soaking in a solution of 4,4′-methylenebis(*N*,*N*-dimethylaniline). Reverse phase HPLC was performed on a C_{18} Kromasil silica-based column. Melting points were determined in a variable temperature optical microscope and are uncorrected. Optical rotation was measured on a Perkin Elmer 241 MC polarimeter.

Materials

Avidin (affinity purified, Sigma), streptavidin (affinity purified, Sigma), 5-ethoxycarbonylpentanoic acid (Aldrich) and glycoluril (Acros) were used as received. 4,5-Dihydroxyimidazolidin-2-one,[21] 6-ureidohexanoic acid,[39] dimethyldioxirane[40] and 5-(3,7-dioxo-2,4,6,8-tetraazabicyclo[3.3.0]oct-2-yl)pentanoic acid (**2c**)[19] were prepared acording to literature procedures.

(+)- and (−)-5-(3,7-Dioxo-2,4,6,8-tetraazabicyclo-[3.3.0]oct-2-yl)pentanoic acid (2c)

Racemic resolution of 5-(3,7-dioxo-2,4,6,8-tetraazabicyclo-[3.3.0]oct-2-yl)pentanoic acid at milligram scale was achieved by HPLC using as chiral stationary phase a Chirobiotic® T (ASTEC Inc, USA) column.[23] (+)-**2c**: $[\alpha]_D^{22} = +50.5$ ($c = 0.12$, MeOH); (−)-**2c**: $[\alpha]_D^{22} = -48.5$ ($c = 0.09$, MeOH).

6-(3,7-Dioxo-2,4,6,8-tetraazabicyclo[3.3.0]oct-2-yl)hexanoic acid (2d)

6-Ureidohexanoic acid (1.74 g, 10 mmol) was dissolved in 60 ml of water at 90 °C. The pH of the solution was adjusted to 1 with conc. HCl and 4,5-dihydroxyimidazolidin-2-one (903 mg, 10.5 mmol) was added. The resulting mixture was heated to reflux for 1 h and was then cooled to 4 °C and the precipitate collected after 24 h and recrystallized from water. Yield: 50%; mp: 179–183 °C. ^{1}H NMR (200 MHz, DMSO-d_6): 7.40 (s, 1H, N*H*); 7.23 (s, 2H, N*H*); 5.20 (m, 2H, –C*H*–C*H*–); 3.19–3.05 (m, 1H, N–C*H*H); 2.97–2.83 (m, 1H, N–CH*H*); 2.17 (t, J = 7 Hz, 2H, CH_2–COOH); 1.44 (m, 4H, CH_2); 1.18 (m, 2H, CH_2). IR (KBr): ν_{max}/cm^{-1}: 3290, 1740, 1708, 1682, 1509. ESI-MS: 279.10 ([M + Na]$^+$). Elemental analysis (%) calc. for $C_9H_{14}N_4O_4 \cdot 0.6H_2O$: C 44.97, H 6.49, N 20.98, O 27.56; found: 44.98, H 6.51, N 20.69, O 27.82.

Ethyl-7-diazo-6-oxoheptanoate (3)

Ethyl 5-(chloroformyl)pentanoate (1.203 g, 6.25 mmol) was dissolved in 30 ml of anhydrous hexane under a nitrogen atmosphere and added dropwise to a solution of diazomethane in Et_2O (80 ml, 15.1 mmol). The reaction mixture was stirred for 1 h at room temperature and then the excess diazomethane and the solvent were removed under reduced pressure to yield 1.199 g of a yellow oil. The crude diazoketone was purified by column chromatography (gradient from hexane–AcOEt 3 : 1 to 1 : 1) to yield 0.912 g of a yellow oil (75%). IR (film) ν_{max}/cm^{-1}: 3094, 2981, 2940, 2871, 2104, 1731, 1641. ^{1}H NMR (200 MHz, $CDCl_3$): 5.26 (s broad, 1H, –C*H*=N_2), 4.13 (q, 2H, –CH_2–O–CO–), 2.32 (m, 4H, CH_2–COOEt and CH_2–CO–CH=N_2), 1.66 (m, 4H, CH_2–CH_2). 1.25 (t, 3H, CH_3–CH_2–O).

Ethyl-5-(3,7-dioxo-2,4,6,8-tetraazabicyclo[3.3.0]oct-1-yl)pentanoate (5)

Ethyl 7-diazo-6-oxoheptanoate (42 mg, 0.21 mmol) was dissolved in 5 ml of acetone, and cooled with an ice bath. Dimethyldioxirane solution (1.5 equiv.) was added and evolution of nitrogen was observed. The reaction mixture was allowed to warm to room temperature and react for 30 min. The solvent was removed under reduced pressure and the desired glyoxal **4** was obtained as an oil, pure enough for the next reaction. The glyoxal **4** was dissolved in a H_2O–EtOH mixture (2 : 1) and cooled on an ice bath. Urea (61.4 mg, 1.02 mmol) was added and the pH of the solution was adjusted to 3 with 0.05 M HCl. The reaction mixture was allowed to react for 3 d at 4 °C. Then the pH of the solution was adjusted to 1 with 0.1 M HCl, and the reaction was kept at 4 °C for a further 5 d. The solvents were removed under reduced pressure and the solid obtained was loaded (as a solid dispersion in silica) into a chromatography column and eluted with an AcOEt–MeOH gradient. The product was then further purified by RP-HPLC (40% MeOH in H_2O isocratic) to yield 16 mg of a white solid (28%); mp: 255–258 °C. ESI-MS: 293.1 ([M + Na]$^+$). IR (KBr): ν_{max}/cm^{-1}: 3225, 2941, 1730, 1681, 1507. ^{1}H NMR (200 MHz, DMSO-d_6): 7.16 (s, 2H, N*H*); 7.06 (s, 2H, N*H*); 4.85 (s, 1H, NH–C*H*–NH); 4.03 (q, J = 7.2 Hz, 2 H, –CH_2–O–CO), 2.26 (t, J = 7 Hz, 2H, –CH_2–COOEt); 1.62 (m, 2H, –CH_2–CH_2–COOEt); 1.48 (t, J = 7 Hz, 2H, –C–CH_2–CH_2); 1.26 (m, 2H, C–CH_2–CH_2–CH_2); 1.16 (t, J = 7.2 Hz, 3H, O–CH_2–CH_3).

5-(3,7-Dioxo-2,4,6,8-tetraazabicyclo[3.3.0]oct-1-yl)pentanoic acid (2b)

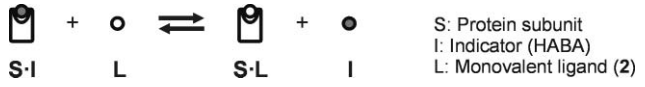
Ethyl-5-(3,7-dioxo-2,4,6,8-tetraazabicyclo[3.3.0]oct-1-yl)pentanoate (24 mg, 0.088 mmol) was dissolved in 8 ml of an aqueous solution of KOH (0.012 M, 0.096 mmol). The reaction mixture was stirred overnight at room temperature. Then, 0.3 ml of a KOH solution (0.098 M, 0.029 mmol) were added and the reaction was stirred for a further 6 h. The reaction mixture was then acidified to pH 2 with 0.1 M HCl and evaporated *in vacuo*. The solid residue was redissolved in the minimum quantity of water (about 8 ml) and loaded onto a column packed with MCI Gel CHP20P resin, washed with 50 ml of water and then the product eluted with water–MeOH (2 : 1, v/v), yielding 22 mg of a white solid (quantitative); mp: 243–245 °C. ESI-MS: 265.0 [M + Na^+]. IR (KBr): ν_{max}/cm^{-1}: 3220, 2948, 1745, 1683, 1558, 1508. ^1H NMR (400 MHz, DMSO-d_6): 7.16 (s, 2H, N*H*); 7.06 (s, 2H, N*H*); 4.85 (s, 1H, NH–C*H*–NH); 2.18 (t, J = 7 Hz, 2H, CH_2–COOH); 1.62 (m, 2H, CH_2–CH_2–COOH); 1.48 (t, J = 7 Hz, 2H, –C–CH_2–CH_2); 1.26 (m, 2H, –C–CH_2–CH_2–CH_2). ^{13}C (100 MHz, DMSO-d_6): 175.8, 161.9, 77.1, 69.3, 39.4, 35.1, 25.9, 23.6. Elemental analysis (%) calc. for $C_9H_{14}N_4O_4{\cdot}0.5H_2O$: C 43.03, H 6.02, N 22.30, O 28.66; found: 43.54, H 5.64, N 22.05, O 28.77.

General procedure for the spectrophotometric competitive titrations

Spectrophotometric competitive titrations were performed on a 1.5 mL cell of 1 cm pathlength. Aliquots of a 0.1–3.0 mM solution of the ligand (either single or racemic) in phosphate buffer (0.1 M, pH = 7.0) were added to a 5–15 μM solution of the protein and 30–70 μM of HABA in phosphate buffer (0.1 M, pH = 7.0). UV-Vis absorption spectra were recorded from 700 to 200 nm, 5 min after each addition, in order to monitor the changes in absorbance at 500 and 348 nm (due to the displacement of the protein-bound dye by the ligand).

(a) Binding model for a competitive titration with an enantiomerically pure ligand (or a single ligand). This binding model assumes that the four identical subunits of each protein (Av or Sav) behave independently, without cooperative effects. This assumption allows us to consider the competition of the indicator (I) and the ligand (L) for the tetrameric protein as equivalent to a competition of I and L for the monomeric protein subunits (S).

Since Av and Sav have only one binding site per subunit, the equations of the 1 : 1 binding model[25,26] can be used (see Scheme 2 and ESI†).

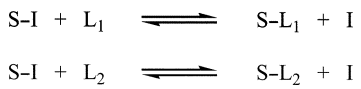


Scheme 2

Least-squares fitting of the calculated absorbance (A_{calc}) [eqn. (1)] to the experimental absorbance (A_{exp}) led to the optimal value for the binding constant between S and L. The reasonably good fit between the experimental and calculated curves supports our assumption of independent behaviour between the binding sites.

$$A_{calc} = \varepsilon_{I(500)}[I] + \varepsilon_{S\cdot I(500)}[S\cdot I] \quad (1)$$

(b) Binding model for a competitive titration with a racemic ligand (or a mixture of two ligands). This binding model also assumes that the four identical subunits of each protein (Av or Sav) behave independently, so that the equations of the 1 : 1 binding model[25,26] can also be used. But now there are two ligands (L_1 and L_2) that compete with the indicator (I) for the protein subunits (S) (Scheme 3). It was assumed that there was no interaction between the two ligands L_1 and L_2, and that they would bind the same protein with different binding constants (K_{L_1S} and K_{L_2S} respectively, see ESI†).



Scheme 3

In order to simultaneously measure the binding constant of both enantiomeric ligands to a given protein by means of a single titration with the racemic ligand, the general procedure for spectrophotometric competitive titrations was also followed.

Least-squares fitting of the calculated absorbance (A_{calc}) [eqn. (1)] to the experimental absorbance (A_{exp}) afforded in a single experiment the binding constant of each enantiomeric ligand. This method can also be applied to non-racemic mixtures of two ligands, as long as their molar ratio is known. The reasonably good fit between the experimental and calculated curves supports our assumption of independent behaviour between the binding sites.

Docking studies

AutoDock 3.05[34–36] was used in the docking studies. The X-ray structure of the dimeric streptavidin–biotin complex was retrieved from the Protein Data Bank[41] (entry code: 1SWE[38]), water molecules were removed and the resulting pdb file was converted to the tetrameric form by symmetry operations using the Swiss-PdbViewer.[42] Polar hydrogens were added and Kollman charges were assigned to the macromolecule using the AutoDockTools program.[43] One of the four biotin molecules was removed from its binding site and affinity grids of 80 × 80 × 80 points, separated 0.375 Å and centered in the empty binding site, were constructed using the Autogrid program. Charges and active torsions were added to the ligands using the AutoDockTools program and non-polar hydrogens were merged with the carbons. The Lamarckian genetic algorithm (LGA) was used in order to perform the conformational search. For each ligand, 50 independent docking runs of 100 individuals and 10 000 generations were performed using the default values for the AutoDock parameters. Results were clustered and ranked according native AutoDock scoring function.

Acknowledgements

This research was supported by Generalitat de Catalunya (2005SGR000653 and 2002BEAI400180) and by Ministerio de Ciencia y Tecnologia (BQU2002-02424) of Spain. Both P. H.-F. and E. A. thank Generalitat de Catalunya for predoctoral fellowships, and I. C. thanks Universitat de Barcelona for a predoctoral fellowship. We thank Professor Roeland J. M. Nolte,

Dr Maria Lluïsa Pérez, Dr Jaume Farràs and Dr Oriol Pineda for helpful discussions.

References

1 N. M. Green, *Methods Enzymol.*, 1990, **184**, 51–67.
2 N. M. Green, *Adv. Protein Chem.*, 1975, **29**, 85–133.
3 M. Ahlers, W. Mueller, A. Reichert, H. Ringsdorf and J. Venzmer, *Angew. Chem.*, 1990, **102**, 1310–1327; M. Ahlers, W. Mueller, A. Reichert, H. Ringsdorf and J. Venzmer, *Angew. Chem., Int. Ed. Engl.*, 1990, **29**(11), 1269–1285.
4 L. Haeussling, B. Michel, H. Ringsdorf and H. Rohrer, *Angew. Chem.*, 1991, **103**, 568–572; L. Haeussling, B. Michel, H. Ringsdorf and H. Rohrer, *Angew. Chem., Int. Ed. Engl.*, 1991, **30**(5), 569–572.
5 H. Ebato, J. N. Herron, W. Mueller, Y. Okahata, H. Ringsdorf and P. Suci, *Angew. Chem.*, 1992, **104**, 1064–1066; H. Ebato, J. N. Herron, W. Mueller, Y. Okahata, H. Ringsdorf and P. Suci, *Angew. Chem., Int. Ed. Engl.*, 1992, **31**(8), 1087–1090.
6 W. Mueller, H. Ringsdorf, E. Rump, G. Wildburg, X. Zhang, L. Angermaier, W. Knoll, M. Liley and J. Spinke, *Science (Washington, D. C.)*, 1993, **262**, 1706–1708.
7 K. Fujita, S. Kimura, Y. Imanishi, E. Rump, J. van Esch and H. Ringsdorf, *J. Am. Chem. Soc.*, 1994, **116**, 5479–5480.
8 H. Ringsdorf and J. Simon, *Nature*, 1994, **371**, 284.
9 P. Ringler, W. Muller, H. Ringsdorf and A. Brisson, *Chem.–Eur. J.*, 1997, **3**, 620–625.
10 J. M. Hannink, J. J. L. M. Cornelissen, J. A. Farrera, P. Foubert, F. C. De Schryver, N. A. J. M. Sommerdijk and R. J. M. Nolte, *Angew. Chem., Int. Ed.*, 2001, **40**, 4732–4734.
11 O. Livnah, E. A. Bayer, M. Wilchek and J. L. Sussman, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 5076–5080.
12 P. C. Weber, D. H. Ohlendorf, J. J. Wendoloski and F. R. Salemme, *Science (Washington, D. C.)*, 1989, **243**, 85–88.
13 P. C. Weber, J. J. Wendoloski, M. W. Pantoliano and F. R. Salemme, *J. Am. Chem. Soc.*, 1992, **114**, 3197–3200.
14 P. C. Weber, M. W. Pantoliano, D. M. Simons and F. R. Salemme, *J. Am. Chem. Soc.*, 1994, **116**, 2717–2724.
15 L. Pugliese, A. Coda, M. Malcovati and M. Bolognesi, *J. Mol. Biol.*, 1993, **231**, 698–710.
16 Y. Pazy, T. Kulik, E. A. Bayer, M. Wilchek and O. Livnah, *J. Biol. Chem.*, 2002, **277**, 30892–30900.
17 A. E. Rowan, J. A. A. W. Elemans and R. J. M. Nolte, *Acc. Chem. Res.*, 1999, **32**, 995–1006.
18 B. A. Katz, B. Liu and R. Cass, *J. Am. Chem. Soc.*, 1996, **118**, 7914–7920.
19 A. N. Kravchenko, E. Y. Maksareva, P. A. Belyakov, A. S. Sigachev, K. Y. Chegaev, K. A. Lyssenko, O.V. Lebedev and N. N. Makhova, *Russ. Chem. Bull.*, 2003, **52**, 192–197.
20 A. Wu, J. C. Fettinger and L. Isaacs, *Tetrahedron*, 2002, **58**, 9769–9777.
21 E. Grillon, R. Gallo, M. Pierrot, J. Boileau and E. Wimmer, *Tetrahedron Lett.*, 1988, **29**, 1015–1016.
22 S. Gautam, R. Ketcham and J. Nematollahi, *Synth. Commun.*, 1979, **9**, 863–870.
23 K. Steinbach and S. Berger, *Chromatographia*, 1997, **46**, 92–94.
24 N. M. Green, *Biochem. J.*, 1965, **94**, C23.
25 H.-J. Schneider and A. K. Yatsimirsky, *Principles and Methods in Supramolecular Chemistry*, John Wiley & Sons, Chichester, 2000.
26 K. A. Connors, *Binding constants. The measurement of molecular complex stability*, John Wiley & Sons, New York, 1987.
27 J.-A. Farrera, P. Hidalgo-Fernandez, J. M. Hannink, J. Huskens, A. E. Rowan, N. A. J. M. Sommerdijk and R. J. M. Nolte, *Org. Biomol. Chem.*, 2005, **3**, 2393–2395.
28 D. H. Williams and M. S. Westwell, *Chem. Soc. Rev.*, 1998, **27**, 57–64.
29 G. Tarkanyi, *J. Chromatogr., A*, 2002, **961**, 257–276.
30 P. Bartak, P. Bednar, L. Kubacek, M. Lammerhofer, W. Lindner and Z. Stransky, *Anal. Chim. Acta*, 2004, **506**, 105–113.
31 J. P. Liu, H. Coffey, D. J. Detlefsen, Y. Li and M. S. Lee, *J. Chromatogr., A*, 1997, **763**, 261–269.
32 A. Dalla Cort, J. I. M. Murua, C. Pasquini, M. Pons and L. Schiaffino, *Chem.–Eur. J.*, 2004, **10**, 3301–3307.
33 E. Buergisser, R. J. Lefkowitz and A. DeLean, *Mol. Pharmacol.*, 1981, **19**, 509–512.
34 G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew and A. J. Olson, *J. Comput. Chem.*, 1998, **19**, 1639–1662.
35 G. M. Morris, D. S. Goodsell, R. Huey and A. J. Olson, *J. Comput. Aided Mol. Des.*, 1996, **10**, 293–304.
36 D. S. Goodsell and A. J. Olson, *Proteins: Struct., Funct., Genet.*, 1990, **8**, 195–202.
37 B. D. Bursulaya, M. Totrov, R. Abagyan and C. L. Brooks, 3rd, *J. Comput. Aided Mol. Des.*, 2003, **17**, 755–763.
38 S. Freitag, I. Le Trong, L. Klumb, P. S. Stayton and R. E. Stenkamp, *Protein Sci.*, 1997, **6**, 1157–1166.
39 G. R. Stark, *Biochemistry*, 1965, **4**, 1030–1036.
40 W. Adam, J. Bialas and L. Hadjiarapoglou, *Chem. Ber.*, 1991, **124**, 2377.
41 H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235–242.
42 N. Guex and M. C. Peitsch, *Electrophoresis*, 1997, **18**, 2714–2723.
43 M. F. Sanner, *J. Mol. Graphics Modell.*, 1999, **17**, 57–61.

