Amino Acid Based Squaramides for Anion Recognition

Robert B. P. Elmes and Katrina A. Jolliffe

School of Chemistry, The University of Sydney, 2006, NSW, Australia.
Fax: +61 2 9351 3329; Tel: +61 2 9351 2297; E-mail: kate.jolliffe@sydney.edu.au

Abstract

Eight receptors **1** – **8** comprising an L-Lysine scaffold modified at the N- and C- termini with aliphatic alkyl chains and *N*,*N*'-alkyl amides, respectively and bearing squaramide moieties on the amino acid side chain were synthesised by a combination of solid and solution phase chemistries and shown to complex various anions in 0.5% H_2O in DMSO-d_6 solution. All of the receptors were found to bind SO_4^{2-}, Cl^-, AcO^- and BzO^- *via* hydrogen bond or acid base interactions with the squaramide protons, however **1** was found to bind to SO_4^{2-} *via* hydrogen bonds formed between the anion and both the squaramide and amide NH moieties. Moreover, modification of both the N- and C- termini of the amino acids with different alkyl substituents had a negligible effect on their anion binding properties while simultaneously conferring lipophilicities in a range that is optimal for molecules to behave as 'drug-like' systems as defined by Lipinski's rule of five. The results of this study demonstrate the versatility of such amino acid receptors as building blocks in the field of anion recognition.

Keywords

Anion binding; squaramide; amino acid

Introduction

An area of intense focus within supramolecular chemistry in recent years has been the development of artificial receptors for anion binding, sensing, extraction and transport.[1-5] In particular, significant effort has been directed towards the development of small molecules with the ability to selectively bind anions and facilitate their transport through lipid bilayers where the most effective anion transporters often show potent anticancer activity.[1,6]

In living systems, nature takes advantage of large peptides/proteins to successfully recognise and transport anions, using H-bonding interactions from various amino acid side chains (e.g. the OH groups of serine, threonine and tyrosine, the NH group in the indole moiety of tryptophan, and the guanidinium group of arginine) as well as from the amide backbone protons. However, synthetic peptide based anion receptors have received surprisingly little attention as membrane transport candidates. We have developed a number of highly efficient and selective anion receptors that can be constructed from a small number of amino acid building blocks to accommodate a variety of target anionic guests.[7-10] The ability of these receptors to discriminate between anions has inspired us to examine their use as anion transporters and we have recently reported the design and synthesis of peptide based cryptands which we demonstrated to act as efficient SO_4^{2-} transporters.[10] One of the key features limiting the use of such receptors as effective transporters of anions across membranes is the need for molecules with both suitable binding affinity and lipophilicity. Indeed, recent reports by Gale and Quesada[11,12] have highlighted the importance of receptor lipophilicity in transmembrane anion transport. We envisaged that the lipophilicity of our peptide based receptors could be readily tuned by the incorporation of various alkyl functionalities at the N- and C- termini. We chose to explore this using a series of amino acid receptors with squaramide–functionalised Lysine side chains, in which a variety of substituents were attached to the N- and C- termini to modulate receptor lipophilicity.

The squaramide functional group has recently been exploited in supramolecular chemistry for the design of anion receptors,[13-18] self-complementary molecular recognition motifs,[19] and most recently by Gale *et. al.* where the squaramide motif was found to be a potent functionality for transmembrane anion transport of both chloride and bicarbonate.[20] Notably, previous work indicates that the introduction of electron-withdrawing trifluoromethyl groups onto the squaramide aromatic substituents can be used to modify both lipophilicity and anion binding affinity.[10,13] Herein we

report the synthesis of eight squaramide-containing receptors **1** – **8** (Figure 1) and a study of their anion complexation properties using ^{1}H NMR spectroscopic methods.

[insert Figure 1 here]

Results and Discussion:

Receptor Design and Synthesis:

We chose L-Lysine (Lys) as a scaffold for the receptors as the side-chain of this amino acid can readily be functionalised with the squaramide anion binding motif and an amino acid of this type has previously been observed to bind to a range of anions.[10] The N- and C- termini of the side-chain functionalised Lys scaffold provide suitable handles for the introduction of aliphatic alkyl chains and *N,N'*-dialkyl amides at the N- and C- termini, respectively, thereby allowing systematic modification of the receptor lipophilicity without changing the anion binding motif. Calculated log P (*c*logP) and total polar surface area (TPSA) values were calculated for receptors **1** - **8** using Spartan '10 (Ghose–Crippen model) after structure minimisation using AM1 semi-empirical methods and the results indicate that receptors **1** – **8** (Table 1) cover a range of lipophilicities with *c*logP values ranging from 1.45 - 5.76. Notably, many of the receptors lie within the boundaries defined by Lipinski's rule of five, which dictates that a log P value ≤ 5 is optimal for molecules to behave as 'drug-like' systems; an important consideration in the context of the development of such anion receptor systems for biological applications.[21]

The synthesis of **1** - **8** was carried out using an Fmoc solid phase peptide synthesis (Fmoc-SPPS) strategy on Wang resin (Scheme 1) with orthogonal allyloxycarbonyl (Alloc) protection of the side chain amino groups. Resin loading was achieved by treatment of Fmoc-Lys(Alloc)-OH with diisopropylcarbodiimide (DIC) in the presence of DMAP followed by reaction with preswollen Wang resin in DMF. Fmoc deprotection (20% piperidine/DMF) preceded coupling with the

appropriate alkyl carboxylic acid [(*O*-benzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate (HBTU)/*i*Pr$_2$NEt] to install the desired N-terminal alkyl chain. The Alloc group was then removed by treatment with Pd(PPh$_3$)$_4$ in the presence of acetic acid and morpholine[22] before functionalisation of the side chain amine was achieved by reaction with either 3-ethoxy-4-(phenylamino)cyclobut-3-ene-1,2-dione (**1**), 3-ethoxy-4-(4-(trifluoromethyl)phenylamino)cyclobut-3-ene-1,2-dione (**2**) or 3-(3,5-bis(trifluoromethyl)phenylamino)-4-ethoxycyclobut-3-ene-1,2-dione (**3** – **8**) to install the variously functionalised squaramide moieties. Once the desired scaffold had been assembled, cleavage from the solid support was effected by treatment with a solution of TFA/TIS/H$_2$O (95/2.5/2.5; v/v) to yield the carboxylic acid precursors. The *N*,*N*'-dialkyl amide receptors (**1** – **8**) were successfully synthesised by coupling with the appropriate *N,N'*-dialkyl amine in the presence of 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) and *i*Pr$_2$NEt before being purified by flash column chromatography to yield **1** - **8** in isolated yields of 44 - 65%.

[insert Scheme 1 here]

Anion Binding:

In order to assess the anion binding properties of this family of receptors a number of ^{1}H NMR spectroscopic titration experiments in 0.5% H$_2$O in DMSO-d_6 were conducted. Initially qualitative measurements with receptors **1** - **3** were undertaken using ^{1}H NMR screening experiments in which 10 eq. of a range of anions [Cl$^-$, F$^-$, SO$_4^{2-}$, HSO$_4^-$, H$_2$PO$_4^-$, AcO$^-$, BzO$^-$, NO$_3^-$, *p*-toluenesulfonate (OTs$^-$) as their tetrabutylammonium salts] were added to the receptors in solution. These preliminary results indicated significant changes of the spectra of **1** - **3** in the presence of Cl$^-$, F$^-$, SO$_4^{2-}$, H$_2$PO$_4^-$, AcO$^-$, BzO$^-$ culminating in large downfield shifts or disappearance of both of the squaramide NH protons (NH$_A$ and NH$_B$) and, in some cases, the amide NH proton (NH$_C$).

Conversely, only minor changes were observed in the presence of HSO_4^-, NO_3^- and TsO^- suggesting little interaction of these anions with **1** – **3**. For the trifluoromethylaniline functionalised squaramide derivatives **2** and **3**, there was also a large degree of peak broadening observed for the NH_A and NH_B signals in the presence of F^-, SO_4^{2-}, $H_2PO_4^-$, AcO^- and BzO^- indicating that these receptors showed possible acid–base interactions with these anions (F^-, AcO^- and BzO^-)[23] or alternatively that more complex binding events (e.g proton-proton transfer between bound and unbound $H_2PO_4^-$) or slow exchange processes (SO_4^{2-}) are occuring.[10,24] In contrast, the addition of Cl^-, SO_4^{2-}, AcO^- and BzO^- to receptor **1** resulted in downfield shifts of NH_A and NH_B as well as a large down field shift for the backbone amide NH_C in the case of SO_4^{2-}.

In order to investigate the anion binding affinities more closely we conducted quantitative binding studies with **1** - **3** in the presence of Cl^-, SO_4^{2-}, AcO^- and BzO^-. Representative spectra for titration of **1** with AcO^- and SO_4^{2-} are shown in Figure 2, illustrating the significant downfield shifts of the squaramide proton signals and the varying response of backbone amide protons. Unfortunately, titration of **2** with SO_4^{2-} and AcO^- and **3** with SO_4^{2-}, AcO^- and BzO^- led to peak broadening, preventing an association constant from being determined in these cases. In all other cases the observed changes to NH_A and NH_B were fitted to a 1:1 binding model using Hyperquad©[25] to give apparent stability constants which are summarised in Table 1.

Table 1 Summary of the anion binding affinities (K_a), *c* log P and TPSA [Å^2] values for receptors **1**–**8**.[a, b]

Receptor	K_a (M^{-1})				*c* log P	PTSA (Å^2)
	Cl^-	AcO^-	BzO^-	SO_4^{2-}		
1	180	3600	3150	> 10^4	3.19	88.62
2	209	-[c]	3715	-[c]	4.11	89.32
3	383	-[c]	-[c]	-[c]	5.03	90.62
4	380	-[c]	-[c]	-[c]	4.20	88.74
5	417	-[c]	-[c]	-[c]	3.36	89.47
6	457	-[c]	-[c]	-[c]	1.45	91.13
7	389	-[c]	-[c]	-[c]	5.71	90.60
8	394	-[c]	-[c]	-[c]	5.76	90.97

[a] Determined at 300 K. Data was fitted to a 1 : 1 binding model. K_a values are an average obtained from monitoring both squaramide NH signals. Errors < 15%. [b] Anions added as their tetrabutylammonium salts. [c] Addition of SO_4^{2-}, AcO^- or BzO^- resulted in peak broadening and prevented an association constant from being determined.

A general trend was observed for Cl^- binding where the receptor bearing the most electron withdrawing substituents, **3**, was found to bind Cl^- most strongly followed by **2** and finally **1** (**3** +

Cl^- K_a = 383 M^{-1}, **2** + Cl^- K_a = 209 M^{-1}, **1** + Cl^- K_a = 180 M^{-1}). For titrations with AcO-, BzO- and SO_4^{2-}, a comparison of relative anion affinities could only be made for receptor **1** (receptors **2** and **3** exhibited behaviour suggesting deprotonation on addition of these more basic anions). Receptor **1** was found to bind SO_4^{2-} with an affinity too high to measure accurately by NMR and was also shown to exhibit higher affinity for SO_4^{2-} than for AcO^- or BzO^- (**1** + SO_4^{2-} $K_a > 10^4$ M^{-1} while **1** + AcO^- K_a = 3600 M^{-1} and **1** + BzO^- K_a = 3150 M^{-1}). Notably, the addition of SO_4^{2-} resulted in a significant downfield shift ($\Delta\delta$ = 0.55) (Figure 2) of the amide proton of **1**, indicating the possible formation of a hydrogen bonding interaction between this proton and SO_4^{2-}. In contrast, the addition of Cl^-, AcO^- and BzO^- to **1** resulted in a very minor shift of this proton (< 0.2 ppm), mirroring our results observed with dipeptide based anion receptors[10] and suggesting that the higher affinity observed for binding of SO_4^{2-} over that for other anions may be a result of the formation of an additional hydrogen bond to this anion.

[insert Figure 2 here]

We next wished to evaluate the effect of modification of the N- and C- termini of our amino acid based scaffold. Additional Cl^- titrations were carried out using **4 - 8**, to probe the effect of increased receptor lipophilicity on binding behavior while keeping the binding motif constant. For the series **3 - 6** in which the N-terminal substituent was varied, a modest increase in affinity for Cl^- was observed on decreasing the length of the chain from decanoyl (**3**: K_a = 387) to acetyl (**6**: K_a = 457) (Table 1), while receptors **3**, **7** and **8** in which the C-terminal substituent was varied had essentially identical Cl^- affinities (within experimental error) indicating that the C-terminal amide substituent had a neglible impact on binding affinity. These results indicate that the lipophilicity of these amino acid receptors can be tuned with negligible imact on binding affinity and opens the possibility of introducing other functionalities onto the amino acid based anion receptors or introducing amino

acid based receptors to larger receptor scaffolds or solid supports without having a detrimental effect on their anion recognition ability.

Conclusions:

In conclusion, we have developed an efficient strategy for the synthesis of amino acid based squaramide anion receptors **1** – **8** using a combination of solid and solution phase chemistries. We have demonstrated that receptors **1** – **8** can efficiently bind anions in aqueous DMSO solution with selective SO_4^{2-} binding observed in the case of **1**, which was attributed to hydrogen bonds being formed between the anion and both the squaramide and amide NH moieties. The ability to attenuate the molecular properties of these receptors without affecting their anion recognition properties has also been demonstrated and we envisage that this work will facilitate the potential use of peptide based anion receptors in biological or environmental applications. We are continuing to explore these and other aspects of peptide based anion complexation and the results of these studies will be reported in due course.

Experimental Section:

General Remarks:

Optical rotations were obtained using a Perkin Elmer model 341 polarimeter at 589 nm and 20 °C, using the indicated spectroscopic grade solvent. ^{1}H NMR (500 MHz) and ^{13}C{^{1}H} NMR (125 MHz) were determined on a Bruker Avance DPX 500 spectrometer. Chemical shifts for ^{1}H NMR are reported in parts per million (ppm) downfield shift from TMS using the residual solvent peak of dimethyl sulfoxide-*d*$_6$ (δ_H 2.50 ppm) as internal references. The data are reported as chemical shift (δ), multiplicity (br = broad, s = singlet, d = doublet, t = triplet, dd = doublet of doublets, dt = doublet of triplets, m = multiplet), coupling constant (*J* Hz), relative integral and assignment where

possible. Chemical shifts for $^{13}C\{^{1}H\}$ NMR are reported in ppm, relative to the central line of a septet at δ = 39.51 ppm for dimethyl sulfoxide-d_6.

Infrared absorption spectra were recorded on a Bruker Alpha-E FT-IR spectrometer using attenuated total reflection (ATR) of a thin film. FT-IR are reported in wavenumbers (cm^{-1}). HRMS (ESI) were obtained using a Bruker BioApex Fourier Transform Ion Cyclotron Resonance mass spectrometer (FTICR) with an analytical ESI source, operating at 4.7 T and reported as *m/z* (relative intensity). Commercial materials were used as received unless otherwise noted. Amino acids, coupling reagents and resins were obtained from Novabiochem or GL Biochem. Anhydrous CH_2Cl_2 was obtained by distillation over CaH_2 prior to use. Anhydrous DMF was purified by passage through neutral alumina using an Innovative Technology, Inc., PureSolv™ solvent purification system. HPLC grade DMF used for solid phase peptide synthesis was obtained from LabScan or Merck. Analytical TLC was performed using precoated silica gel plates (Merck Kieselgel 60 F254). Squarate monoesters were synthesised as previously described.[15]

Synthesis:

1. Loading of amino acid onto Wang resin.

Under an atmosphere of nitrogen, Wang (100-200 mesh; resin capacity 1.1 mmol g^{-1}) was swollen in a sinter-fitted syringe in anhydrous DMF for 1 h. Simultaneously, a solution of Fmoc-Lys(Alloc)-OH (4 equiv. relative to resin capacity) was dissolved in anhydrous CH_2Cl_2 before DIC (2 equiv. relative to resin capacity) was added dropwise at 0°C and the resulting solution was stirred for 1 hr. The CH_2Cl_2 was removed under reduced pressure before the residue was dissolved in the minimum of anhydrous DMF. The resin was drained and treated with the amino acid solution and DMAP (0.1 equiv. relative to resin capacity) in anhydrous DMF and the resulting suspension was gently agitated at rt for 24 h under an atmosphere of nitrogen. The resin was drained and washed sequentially with DMF (5 × 10 mL), CH_2Cl_2 (5 × 10 mL) and DMF (5 × 10 mL).

2. N-terminal Fmoc deprotection.

The resin-bound peptide was agitated in a solution of 10% piperidine in DMF (2 × 5 mL × 15 min) and then drained and washed sequentially with DMF (5 × 5 mL), CH_2Cl_2 (5 × 5 mL) and DMF (5 × 5 mL). The resulting resin-bound amine was used immediately in the next coupling step.

3. Estimation of resin loading.

The drained Fmoc deprotection solution was diluted with a solution of 10% piperidine in DMF so that the maximum concentration of the fulvenepiperidine adduct was in the range of 2.5–7.5 × 10^{-5} M. A sample of this solution was transferred to two matched 1 cm quartz glass cuvettes and the UV-Vis absorbance at 301 nm was measured, using the solution of 10% piperidine in DMF as a reference. The absorbance values were used to calculate the loading, using ε = 7800 $M^{-1\ cm-1}$.

4. Manual SPPS peptide coupling.

Under an atmosphere of nitrogen, a solution of carboxylic acid (5 equiv. relative to loading), HBTU (5 equiv. relative to peptide) and *i*Pr_2NEt (10 equiv. relative to peptide) in anhydrous DMF was added to the resin and the resulting suspension was agitated at rt for 24 h. The resin was then washed sequentially with DMF (5 × 10 mL), CH_2Cl_2 (5 × 10 mL) and DMF (5 × 10 mL).

5. Allyloxycarbonyl deprotection.

The resin was swollen at rt for 15 min in $CHCl_3$/morpholine/acetic acid (90:5:5; 10 mL). $Pd(PPh_3)_4$ (0.15 equiv. relative to loading) was added and the suspension was agitated at rt for 3 hrs. The resin was drained and washed sequentially with a solution of DMF/diethyldithiocarbamic acid-3-water/triethylamine (25 mL; 225 mg; 250 μL) (5 × 5 mL), 0.5% triethylamine in DMF (5 × 10 mL), DMF (5 × 10 mL), CH_2Cl_2 (5 × 10 mL) and DMF (5 × 10 mL).

6. Preparation of substituted N,N'-squaramides

The resin was swollen at rt for 30 min in DMF before the addition of the appropriate squarate monoester (1.8 equiv. relative to loading) and triethylamine (3 equiv. relative to loading). The

suspension was agitated at 80°C for 24 hr, drained and washed sequentially with DMF (5 × 10 mL), CH_2Cl_2 (5 × 10 mL) and DMF (5 × 10 mL).

7. Cleavage of intermediates from the resin.

The resin was treated with a mixture of TFA/TIS/H_2O (95/2.5/2.5; v/v) 5 mL) the resulting suspension was agitated at rt for 15 min. The solutions were drained and the resin was washed with CH_2Cl_2 (3 × 5 mL) before all washings were combined and the solvent was removed under reduced pressure to afford the crude product which was used without further purification.

8. Preparation of N,N'-alkyl amides

To a stirred solution of the acid (1 equiv.) and the amine (5 equiv.) in anhydrous CH_2Cl_2 (0.15 M) was added EDC (2.5 equiv.) and *i*Pr_2NEt (5 equiv.) under an atmosphere of nitrogen. The resulting mixture was stirred at rt for 24 h before being quenched with sat. aq. NH_4Cl solution (10 mL). The organic phase was washed with sat. aq. NH_4Cl solution (2 × 10 mL), 0.2 M HCl (3 × 10 mL) and brine (10 mL). The organic layer was then dried over $MgSO_4$ and the solvent was removed under reduced pressure.

_N_-(1-(Dimethylamino)-6-(3,4-dioxo-2-(phenylamino)cyclobut-1-enylamino)-1-oxohexan-2-yl)decanamide (1)

Fmoc-Lys(Alloc)-OH was loaded onto Wang resin (500 mg, resin capacity 1.1 mmol g^{-1}) according to general procedure 1. The N-terminal Fmoc protecting group was then removed (general procedure 2) and the appropriate carboxylic acid was coupled to the resin-bound amine according to the manual SPPS procedure 4. The side chain Alloc group was subsequently removed according to procedure 5 before installation of the appropriate squaramide moiety according to procedure 6. The resulting intermediate was cleaved from the resin according to general procedure 7 to afford the carboxylic acid derivative which was used in the final step without further purification. The *N*,*N*'-alkyl amide functionality was finally introduced according to procedure 8 before subjection of the

crude material to flash chromatography eluting with 5% MeOH in EtOAc to yield **1** as a pale beige solid after lyophilisation (0.91 g, 59%). $[\alpha]^{20}_{D}$ = + 7.1 (c 0.17, MeOH); ^{1}H NMR (500.13 MHz, DMSO-*d*$_6$): 0.84 (t, *J* = 6.9 Hz, 3H), 1.20 – 1.61 (m, 20H), 2.07 (m, 2H), 2.80 (s, 3H), 3.01 (s, 3H), 3.58 (br s, 2H), 4.70 (q, *J* = 8.3 Hz, 1H), 7.02 (t, *J* = 7.1 Hz, 1H), 7.33 (app t, *J* = 8.1 Hz, 2H), 7.43 (d, *J* = 8.1 Hz, 2H), 7.62 (br s, NH), 7.97 (d, *J* = 8.6 Hz, NH), 9.61 (br s, NH); ^{13}C NMR (125.76 MHz, DMSO-*d*$_6$): 14.4, 22.6, 22.6, 25.8, 29.1, 29.1, 29.2, 29.4, 30.8, 31.6, 31.8, 35.4, 35.6, 37.0, 44.0, 48.3, 118.4, 123.0, 129.8, 139.6, 164.0, 169.7, 171.9, 172.4, 180.6, 184.4; HRMS (ESI) calcd. for $C_{28}H_{42}N_4O_4Na$ $[M + Na]^+$ 521.3098, found 521.3100.; IR (thin film) ν_{max} 3286, 2926, 2854, 1794, 1621, 1589, 1544, 1502, 1457 cm^{-1}.

***N*-(1-(Dimethylamino)-6-(3,4-dioxo-2-(4-(trifluoromethyl)phenylamino)cyclobut-1-enylamino)-1-oxohexan-2-yl)decanamide (2)**

Procedure as for **1** above gave the desired compound as a pale beige solid after lyophilisation (0.137 g, 65%). $[\alpha]^{20}_{D}$ = + 8.1 (c 0.12, MeOH); ^{1}H NMR (500.13 MHz, DMSO-*d*$_6$): 0.83 (t, *J* = 7.1 Hz, 3H), 1.18 – 1.62 (m, 20H), 2.07 (m, 2H), 2.80 (s, 3H), 3.00 (s, 3H), 3.59 (br s, 2H), 4.70 (q, *J* = 8.6 Hz, 1H), 7.60 (d, *J* = 7.8 Hz, 2H), 7.68 (d, *J* = 7.8 Hz, 2H), 7.73 (br s, NH), 7.97 (d, *J* = 7.8 Hz, NH), 9.93 (br s, NH); ^{13}C NMR (125.76 MHz, DMSO-*d*$_6$): 14.4, 22.5, 25.8, 29.1, 29.1, 29.4, 30.7, 31.2, 31.6, 31.7, 35.5, 35.6, 37.0, 39.5, 39.6, 39.8, 40.0, 40.1, 40.2, 40.3, 40.4, 40.5, 40.6, 44.1, 48.3, 118.3, 122.6, 123.5 (q, *J* = 33.6 Hz), 126.0, 127.1, 143.1, 163.3, 170.2, 171.8, 172.4, 185.2 ; HRMS (ESI) calcd. for $C_{29}H_{41}N_4O_4F_3Na$ $[M + Na]^+$ 589.2972, found 589.2970.; IR (thin film) ν_{max} 3302, 2926, 2855, 1797, 1631, 1574, 1460, 1337, 1163, 1116, 1071 cm^{-1}.

***N*-(6-(2-(3,5-Bis(trifluoromethyl)phenylamino)-3,4-dioxocyclobut-1-enylamino)-1-(dimethylamino)-1-oxohexan-2-yl)decanamide (3)**

Procedure as for **1** above gave the desired compound as a pale beige solid after lyophilisation (0.20 g, 63%). $[\alpha]^{20}_{D}$ = + 9.0 (c 0.13, MeOH); ^{1}H NMR (500.13 MHz, DMSO-*d*$_6$): 0.83 (t, *J* = 6.8 Hz, 3H), 1.12 – 1.65 (m, 20H), 2.06 (m, 2H), 2.80 (s, 3H), 3.00 (s, 3H), 3.59 (br s, 2H), 4.68 (m, 1H), 7.63 (s, 1H), 7.96 (d, *J* = 8.5 Hz, NH), 8.01 (br s, NH), 8.07 (s, 2H), 10.54 (br s, NH); ^{13}C NMR (125.76 MHz, DMSO-*d*$_6$): 14.4, 14.4, 22.5, 22.5, 25.8, 29.0, 29.1, 29.2, 29.3, 30.6, 31.5, 31.7, 35.4, 35.6, 37.0, 44.1, 48.3, 115.0, 118.3, 123.6 (q, *J* = 274.7 Hz), 131.8 (q, *J* = 30.5 Hz), 141.8, 162.8, 170.3, 171.8, 172.4, 180.7, 185.2 ; HRMS (ESI) calcd. for $C_{30}H_{40}N_4O_4F_6Na$ $[M + Na]^+$ 657.2846, found 657.2843.; IR (thin film) ν_{max} 3396, 2928, 2857, 1631, 1610, 1568, 1455, 1381, 1278, 1180, 1134 cm^{-1}.

***N*-(6-(2-(3,5-Bis(trifluoromethyl)phenylamino)-3,4-dioxocyclobut-1-enylamino)-1-(dimethylamino)-1-oxohexan-2-yl)octanamide (4)**

Procedure as for **1** above gave the desired compound as a pale beige solid after lyophilisation (0.104 g, 65%). $[\alpha]^{20}_{D}$ = + 5.9 (c 0.15, MeOH); ^{1}H NMR (500.13 MHz, DMSO-*d*$_6$): 0.80 (t, *J* = 6.7 Hz, 3H), 1.15 – 1.58 (m, 16H), 2.04 (m, 2H), 2.79 (s, 3H), 2.98 (s, 3H), 3.58 (br s, 2H), 4.67 (q, *J* = 8.5 Hz, 1H), 7.63 (s, 1H), 7.68 (br s, NH), 7.95 (d, *J* = 8.4 Hz, NH), 8.01 (s, 2H), 10.14 (br s, NH); ^{13}C NMR (125.76 MHz, DMSO-*d*$_6$): 14.2, 14.3, 22.2, 22.5, 22.5, 25.4, 25.7, 28.8, 29.0, 30.7, 31.2, 31.3, 31.6, 35.4, 35.6, 37.0, 44.2, 48.2, 115.1, 118.4, 123.7 (q, *J* = 277.3 Hz), 131.8 (q, *J* = 31.8 Hz), 141.6, 162.7, 170.3, 171.8, 172.3, 180.8, 185.3; HRMS (ESI) calcd. for $C_{28}H_{36}N_4O_4F_6Na$ $[M + Na]^+$ 629.2533, found 629.2535.; IR (thin film) ν_{max} 3280, 2928, 2861, 1796, 1631, 1580, 1461, 1434, 1380.

6-(2-(3,5-Bis(trifluoromethyl)phenylamino)-3,4-dioxocyclobut-1-enylamino)-2-hexanamido-*N*,*N*-dimethylhexanamide (5)

Procedure as for **1** above gave the desired compound as a pale beige solid after lyophilisation (0.99 g, 54%). $[\alpha]^{20}_D$ = + 3.5 (c 0.15, MeOH); ^{1}H NMR (500.13 MHz, DMSO-*d*$_6$): 0.80 (t, *J* = 7.0 Hz, 3H), 1.15 – 1.61 (m, 12H), 2.06 (m, 2H), 2.81 (s, 3H), 3.00 (s, 3H), 3.60 (br s, 2H), 4.68 (q, *J* = 8.5 Hz, 1H), 7.65 (s, 1H), 7.75 (br s, NH), 8.02 (s, 2H), 8.07 (d, *J* = 8.3 Hz, NH),10.20 (br s, NH); ^{13}C NMR (125.76 MHz, DMSO-*d*$_6$): 13.8, 21.7, 22.0, 24.9, 30.2, 30.8, 31.1, 34.9, 35.1, 36.5, 43.7, 47.8, 114.6, 118.0, 123.1 (q, *J* = 270.9 Hz), 131.4 (q, *J* = 35.3 Hz), 141.1, 162.3, 169.8, 171.3, 171.8, 180.3, 184.8; HRMS (ESI) calcd. for $C_{26}H_{32}N_4O_4F_6Na$ $[M + Na]^+$ 601.2220, found 601.2220.; IR (thin film) ν_{max} 3292, 2958, 2927, 1794, 1609, 1568, 1460, 1381, 1278, 1180, 1134.

2-Acetamido-6-(2-(3,5-bis(trifluoromethyl)phenylamino)-3,4-dioxocyclobut-1-enylamino)-*N*,*N*-dimethylhexanamide (6)

Procedure as for **1** above gave the desired compound as a pale beige solid after lyophilisation (0.123 g, 55%). $[\alpha]^{20}_D$ = + 3.4 (c 0.12, MeOH); ^{1}H NMR (500.13 MHz, DMSO-*d*$_6$): 1.31 – 1.64 (m, 6H), 1.81 (s, 3H), 2.81 (s, 3H), 3.01 (s, 3H), 3.60 (br s, 2H), 4.68 (m, 1H), 7.65 (s, 1H), 7.71 (br s, NH), 7.97 (d, *J* = 8.5 Hz, NH), 8.02 (s, 2H), 10.17 (br s, NH); ^{13}C NMR (125.76 MHz, DMSO-*d*$_6$): 22.1, 22.3, 30.3, 31.2, 35.1, 36.5, 43.7, 48.0, 114.6, 118.0, 123.1 (q, *J* = 269.4 Hz), 131.3 (q, *J* = 36.7 Hz), 141.2, 162.3, 168.9, 169.8, 171.3, 180.4, 184.8; HRMS (ESI) calcd. for $C_{22}H_{24}N_4O_4F_6Na$ $[M + Na]^+$ 545.1594, found 545.1595.; IR (thin film) ν_{max} 3292, 2943, 1795, 1611, 1563, 1458, 1380, 1279, 1180, 1129.

***N*-(6-(2-(3,5-Bis(trifluoromethyl)phenylamino)-3,4-dioxocyclobut-1-enylamino)-1-(diethylamino)-1-oxohexan-2-yl)decanamide (7)**

Procedure as for **1** above gave the desired compound as a pale beige solid after lyophilisation (0.160 g, 44%). $[\alpha]^{20}_D$ = + 3.5 (c 0.17, MeOH); ^{1}H NMR (500.13 MHz, DMSO-*d*$_6$): 0.83 (t, *J* = 7.0 Hz, 3H), 0.98 (t, *J* = 7.0 Hz, 3H), 1.11 – 1.55 (m, 25H), 2.07 (m, 2H), 3.14 (m, 2H), 3.58 (d, *J* = 6.2 Hz,

2H), 4.62 (q, *J* = 8.7 Hz, 1H), 7.61 (s, 1H), 7.98 (d, *J* = 8.3 Hz, NH), 8.10 (s, 2H), 8.23 (br s, NH), 10.82 (br s, NH); ^{13}C NMR (125.76 MHz, DMSO-d_6): 12.8, 13.9, 14.5, 22.0, 25.3, 28.6, 28.9, 30.2, 31.2, 31.7, 34.9, 41.2, 43.6, 47.8, 114.3, 117.8, 123.2 (q, *J* = 272.8 Hz), 131.3 (q, *J* = 33.0 Hz), 141.6, 162.5, 170.0, 170.7, 171.9, 180.3, 184.6; HRMS (ESI) calcd. for $C_{32}H_{44}N_4O_4F_6Na$ $[M + Na]^+$ 685.3159, found 685.3158.; IR (thin film) ν_{max} 3291, 2928, 2857, 1793, 1609, 1563, 1451, 1331, 1180, 1134 cm^{-1}.

N-(6-(2-(3,5-Bis(trifluoromethyl)phenylamino)-3,4-dioxocyclobut-1-enylamino)-1-oxo-1-(piperidin-1-yl)hexan-2-yl)decanamide (8)

Procedure as for **1** above gave the desired compound as a pale beige solid after lyophilisation (0.185 g, 50%). $[\alpha]^{20}_D$ = - 3.5 (c 0.15, MeOH); ^{1}H NMR (500.13 MHz, DMSO-d_6): 0.83 (t, *J* = 6.9 Hz, 3H), 1.17 – 1.57 (m, 26H), 2.07 (m, 2H), 3.39 – 3.58 (m, 6H), 4.69 (q, *J* = 8.5 Hz, 1H), 7.62 (s, 1H), 7.98 (d, *J* = 8.7 Hz, NH), 8.10 (s, 2H), 8.19 (br s, NH), 10.74 (br s, NH); ^{13}C NMR (125.76 MHz, DMSO-d_6): 13.9, 22.0, 24.0, 25.3, 25.3, 26.0, 26.1, 28.6, 28.6, 28.7, 28.7, 28.9, 30.1, 31.2, 31.4, 35.0, 42.4, 43.6, 45.8, 47.6, 114.4, 117.7, 123.1 (q, *J* = 274.3 Hz), 131.4 (q, *J* = 31.7 Hz), 141.5, 162.4, 169.4, 169.9, 171.7, 180.2, 184.6; HRMS (ESI) calcd. for $C_{33}H_{44}N_4O_4F_6Na$ $[M + Na]^+$ 697.3159, found 697.3160.; IR (thin film) ν_{max} 3290, 2928, 2856, 1657, 1461, 1380, 1277, 1180, 1133 cm^{-1}.

NMR Binding Studies

NMR titrations were performed by additions of aliquots of the putative anionic guest as the tetrabutylammonium (TBA) salt (0.15 – 0.2 M), in a solution of the receptor (2.5×10^{-3}M) in 0.5% H_2O in DMSO-d_6 to a 2.5×10^{-3} M solution of the receptor in 0.5% H_2O in DMSO-d_6. Typically, up to 9-12 equivalents of the anion were added to the solution. Both salt and receptor were dried under high vacuum prior to use. ^{1}H NMR spectra were recorded on a Bruker Avance III 500

spectrometer at a frequency of 500.13 MHz and calibrated to the residual protio solvent peak in DMSO-d_6 (δ = 2.50 ppm). Stack plots were made using MestReNova Version 6.0. Where possible and when the change in chemical shift was larger than 0.02 ppm, non-linear curve fitting of the experimentally obtained titration isotherms (equivalents of anion *vs*. chemical shift of the squaramide NH protons or amide NH protons) using the commercially available software program HypNMR® (Hyperquad® package) enabled the calculation of association constants (K_a/M^{-1}) using a 1:1 model.

clogP Calculations

Calculated log P (*c* log P) and total polar surface area (TPSA) values were calculated using Spartan '10 (Ghose–Crippen model) for Windows (Wavefunction, Inc. Irvine, CA). after minimisation using AM1 semi-empirical methods.

Acknowledgements

The Australian Research Council (DP140100227) and the University of Sydney are acknowledged for financial support.

References:

1.Gale, P. A.; Perez-Tomas, R.; Quesada, R. *Acc. Chem. Res.* **2013**, *46*, 2801-2813.

2. Duke, R. M.; Veale, E. B.; Pfeffer, F. M.; Kruger, P. E.; Gunnlaugsson, T. *Chem. Soc. Rev.* **2010**, *39*, 3936-3953.

3. Fowler, C. J.; Haverlock, T. J.; Moyer, B. A.; Shriver, J. A.; Gross, D. E.; Marquez, M.; Sessler, J. L.; Hossain, M. A.; Bowman-James, K. *J. Am. Chem. Soc.* **2008**, *130*, 14386-14387.

4. Gale, P. A.; Caltagirone, C. *Chem. Soc. Rev.* **2014**.

5. Caballero, A.; Zapata, F.; Beer, P. D. *Coord. Chem. Rev.* **2013**, *257*, 2434-2455.

6. Gale, P. A. *Acc. Chem. Res.* **2011**, *44*, 216-226.

7. Dungan, V. J.; Ngo, H. T.; Young, P. G.; Jolliffe, K. A. *Chem. Commun.* **2013**, *49*, 264-266.

8. Young, P. G.; Jolliffe, K. A. *Org. Biomol. Chem.* **2012**, *10*, 2664-2672.

9. Young, P. G.; Clegg, J. K.; Jolliffe, K. A. *Supramol. Chem.* **2011**, *24*, 77-87.

10. Elmes, R. B. P.; K. Y. Yuen, K.; Jolliffe, K. A. *Chem. Eur. J.* **2014**, *20*, 7373-7380.

11. Saggiomo, V.; Otto, S.; Marques, I.; Felix, V.; Torroba, T.; Quesada, R. *Chem. Commun.* **2012**, *48*, 5274-5276.

12. Busschaert, N.; Bradberry, S. J.; Wenzel, M.; Haynes, C. J. E.; Hiscock, J. R.; Kirby, I. L.; Karagiannidis, L. E.; Moore, S. J.; Wells, N. J.; Herniman, J.; Langley, G. J.; Horton, P. N.; Light, M. E.; Marques, I.; Costa, P. J.; Felix, V.; Frey, J. G.; Gale, P. A. *Chem. Sci.* **2013**, *4*, 3036-3045.

13. Elmes, R. B. P.; Turner, P.; Jolliffe, K. A. *Org. Lett.***2013**, *15*, 5638-5641.

14. Jin, C.; Zhang, M.; Wu, L.; Guan, Y.; Pan, Y.; Jiang, J.; Lin, C.; Wang, L. *Chem. Commun.* **2013**, *49*, 2025-2027.

15. Rostami, A.; Colin, A.; Li, X. Y.; Chudzinski, M. G.; Lough, A. J.; Taylor, M. S. *J. Org. Chem.* **2010**, *75*, 3983-3992.

16. Prohens, R.; Martorell, G.; Ballester, P.; Costa, A. *Chem. Commun.* **2001**, 1456-1457.

17. Neus Piña, M.; Carmen Rotger, M.; Costa, A.; Ballester, P.; Deyà, P. M. *Tetrahedron Lett.* **2004**, *45*, 3749-3752.

18. Soberats, B.; Martínez, L.; Sanna, E.; Sampedro, A.; Rotger, C.; Costa, A. *Chem. Eur. J.* **2012**, *18*, 7533-7542.

19. P. Davis, A.; M. Draper, S.; Dunne, G.; Ashton, P. *Chem. Commun.* **1999**.

20. Busschaert, N.; Kirby, I. L.; Young, S.; Coles, S. J.; Horton, P. N.; Light, M. E.; Gale, P. A. *Angew. Chem. Int. Ed.* **2012**, *51*, 4426-4430.

21. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Delivery Rev.* **1997**, *23*, 3-25.

22. Kates, S. A.; Daniels, S. B.; Albericio, F. *Anal. Biochem.* **1993**, *212*, 303-310.

23. Amendola, V.; Fabbrizzi, L.; Mosca, L. *Chem. Soc. Rev.* **2010**, *39*, 3889-3915.

24. Gale, P. A.; Hiscock, J. R.; Jie, C. Z.; Hursthouse, M. B.; Light, M. E. *Chem. Sci.* **2010**, *1*, 215-220.

25. Gans, P.; Sabatini, A.; Vacca, A. *Talanta* **1996**, *43*, 1739-1753.

Figure and Scheme captions

Figure 1: Structures of squaramide based amino acid receptors **1** – **8** highlighting the positions of NH_A, NH_B and NH_C.

Scheme 1: General synthetic approach to anion receptors **1** - **8**; (i) 20% v/v piperidine in DMF , (ii) carboxylic acid, HBTU, *i*Pr_2NEt, DMF, (iii) $Pd(PPh_3)_4$, AcOH, morpholine, $CHCl_3$, (iv) ethyl squarate monoamide, Et_3N, EtOH, (v) TFA/H_2O/triisopropylsilane (95:2.5:2.5 v/v), (vi) amine, EDC, *i*Pr_2NEt, CH_2Cl_2.

Figure 2: Stack plots of ^{1}H NMR spectra of **1** (2.5×10^{-3} M) upon addition of (a) $(TBA)_2SO_4$ and (b) TBAOAc (0 – 12 eq.) in 0.5% H_2O in DMSO-d_6 at 25 °C. (c) Comparison isotherms of **1** (NH_A) in the presence of increasing concentrations of SO_4^{2-} (♦), Cl^- (■), BzO^- (▲) and AcO^-(×) and (d) Comparison isotherms of **1** (NH_C) in the presence of increasing concentrations of SO_4^{2-} (♦), Cl^- (■), BzO^- (▲) and AcO^-(×)

Graphical abstract

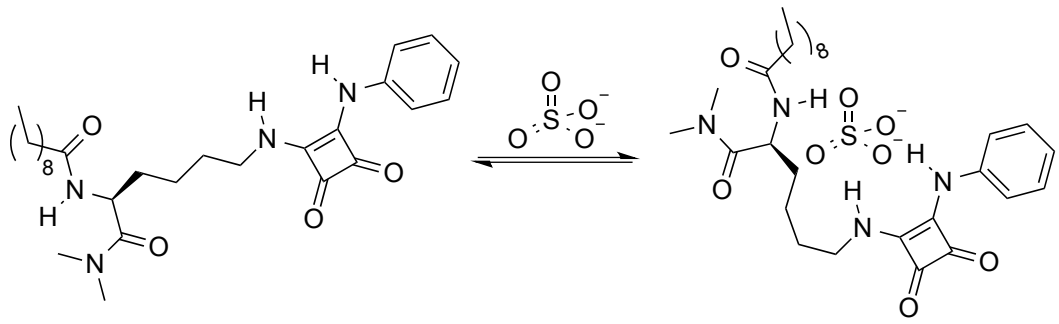
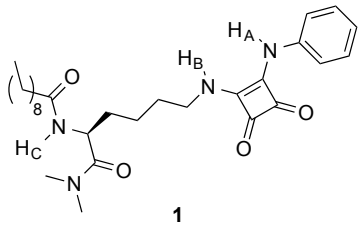
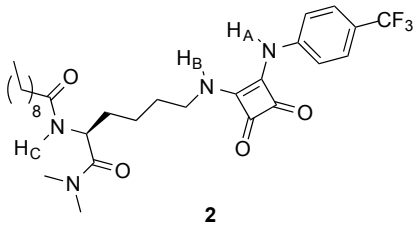


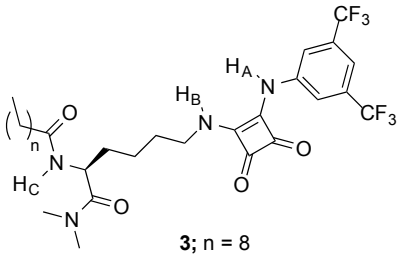
Figure 1



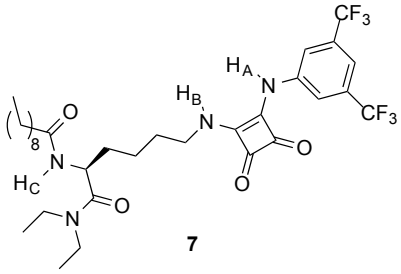
1



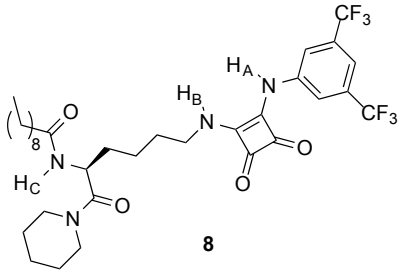
2



3; n = 8
4; n = 6
5; n = 4
6; n = 0



7



8

Scheme 1

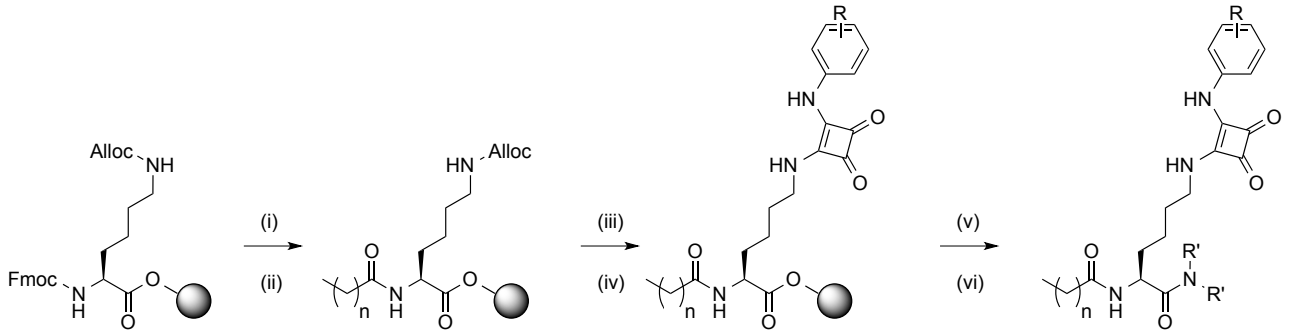
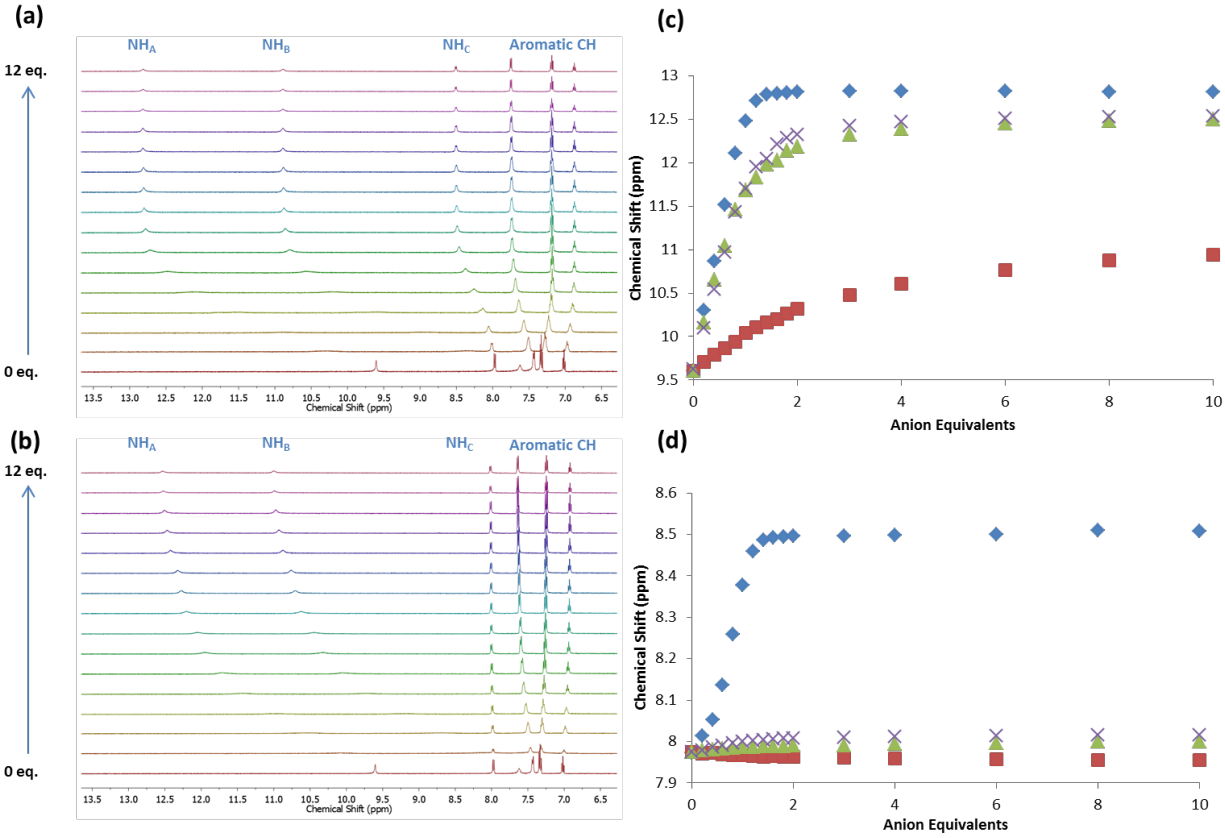


Figure 2


(a)
12 eq.
0 eq.
NHA
NHB
NHC
Aromatic CH
13.5 13.0 12.5 12.0 11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5
Chemical Shift (ppm)
(b)
12 eq.
0 eq.
NHA
NHB
NHC
Aromatic CH
13.5 13.0 12.5 12.0 11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5
Chemical Shift (ppm)
(c)
Chemical Shift (ppm)
13
12.5
12
11.5
11
10.5
10
9.5
0 2 4 6 8 10
Anion Equivalents
(d)
Chemical Shift (ppm)
8.6
8.5
8.4
8.3
8.2
8.1
8
7.9
0 2 4 6 8 10
Anion Equivalents