ChemComm

Selective recognition of carbonic anhydrase using transition metal complexes†

Bidhan C. Roy, Md. Abul Fazal, Shuguang Sun and Sanku Mallik*

Department of Chemistry, North Dakota State University, Fargo, ND 58105, USA.
E-mail: sanku_mallik@ndsu.nodak.edu

Received (in Bloomington, IN, USA) 15th November 1999, Accepted 17th February 2000
Published on the Web 17th March 2000

A metal complex was designed to bind to the protein carbonic anhydrase strongly and selectively in aqueous medium.

Strong and selective binding to a pre-determined protein by synthetic receptors is of paramount importance for drug design,[1] protein purification[2] and protein sensing.[3] The recognition of active sites of enzymes leading to very effective inhibitors has been developed to a high level of sophistication. [4] Recognition of protein surface patterns by polypeptides has been used as a model system for studying protein–protein interactions[5] or to design enzyme inhibitors.[6] These studies employed multiple hydrogen bonding or ionic interactions as the basis of recognition. Herein, we report our results on selective recognition of a protein, carbonic anhydrase (CA, bovine erythrocyte) based on its *surface histidine pattern*, employing Cu^{2+}–histidine interactions in water (25 mM HEPES buffer, pH 7.0, 25 °C). To our knowledge, this is the first report of protein surface recognition using the surface histidine pattern of the protein. Metal–histidine interactions have been used for protein purification using immobilized metal affinity chromatography (IMAC).[2] However, in IMAC, the proteins are distinguished based on their surface histidine contents.

Metal–ligand interactions have several advantages in recognition compared to hydrogen bonding, ion-pair or other weak interactions. Metal–ligand interactions are stronger especially in aqueous media.[7] A variety of transition metal ions are available and are used by nature in biological systems for molecular recognition and catalysis.[8] This allows fine tuning of each of the interactions and opens the possibility of catalysis, following recognition. The spectroscopic properties of the metal ions can be utilized to monitor the binding process and to get structural information about the resultant complex.[9]

CA has five histidines (3, 10, 15, 17, 64) exposed on the surface (determined using the modeling software insight-II and discover, version 98.0, Molecular Simulations Inc., Burlington, MA).[10] The distance amongst the histidines 3, 10, 17 (or 15) are *ca*. 16 Å. This pattern was mapped into the pattern of Cu^{2+} ions by the tris–Cu^{2+} complex **2**. Another shorter tris–Cu^{2+} complex **3** (distance amongst the Cu^{2+} ions: *ca*. 12 Å) was synthesized for comparative studies. A mono-Cu^{2+} complex **1** served as the control for the recognition studies. Structures of the metal complexes **1**, **2** and **3** are shown in Fig. 1. These complexes were modeled using the software Spartan (version 5.0.3, Wavefunction Inc., Irvine) employing the Merck molecular mechanics force field. After energy minization, the resultant structures were subjected to systematic conformational searches to identify the lowest energy structures.

Two other proteins with different surface histidine patterns were used as controls for these studies, chicken egg albumin (CEA) and chicken egg lysozyme. CEA has six histidines on the surface (22, 23, 329, 332, 363, 371); lysozyme (chicken egg) has only one histidine (15) exposed to surface.[10] All of these three proteins are known to interact with transition metal ions through the surface exposed histidines.[11] CEA has a higher number of surface exposed histidines than the target protein (CA) but their distributions are different. If the recognition is based on the *contents* of surface exposed histidine residues (*e.g*. IMAC), then CEA is expected to bind more strongly compared to CA. Since the studies reported here rely on the histidine pattern, CEA was chosen as a control protein.

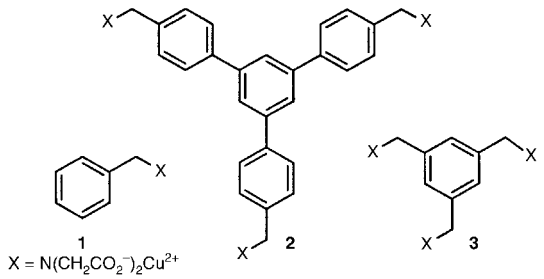

Fig. 1 Structures of the tris–Cu^{2+} complexes **2**, **3** and the control **1** used for recognition of carbonic anhydrase.

Complexes **2** and **3** were synthesized from the corresponding bromides and diethyl iminodicetate, using K_2CO_3 as the base in acetonitrile solvent. The ester groups were hydrolyzed by LiOH in methanol–water at 25 °C.[12] The complexes were found to be stable in aqueous buffer solution (25 mM HEPES buffer, pH 7.0), at room temperature, in air for more than a month. It is reported in the literature that the IDA–Cu^{2+} complex binds to proteins (pH 7.0) primarily through the surface exposed histidine residues of proteins. Primary amino groups and carboxylate residues on the protein surface only play weak, secondary roles.[13]

Binding studies were conducted in water (25 mM HEPES buffer, pH 7.0; [protein] = 100 μM, [**1**] = 0–1.2 mM; [**2**] or [**3**] = 0–500 μM, 25.0 °C) and were followed by isothermal titration microcalorimetry[14] (ITC-4200, Calorimetry Scientific Corporation, Provo, UT). The software provided by the instrument manufacturer (Bindworks 3.0) was used for data and error analysis. The raw data were corrected for the heat of dilution of the appropriate Cu^{2+}-complex before the curve fitting process (data for titrations are available as electronic supplementary information†).

Results for the binding experiments (stoichiometry, binding constant and enthalpy) are shown in Table 1. The control **1** showed low affinity ($K < 4000\ M^{-1}$) for the proteins tested. With the complex **2**, CA was found to bind tightly ($K \approx 300\,000\ M^{-1}$). CEA showed weak binding with complex **2** ($K < 1000\ M^{-1}$) under the experimental conditions. Lysozyme was found to precipitate when [**2**] > 25 μM. With the shorter complex **3**, the affinity for CA was found to decrease ($K \approx 75\,000\ M^{-1}$) compared to that of **2**. Both lysozyme and CEA bound weakly to complex **3** (no precipitation). Complex **2** bound CA more strongly compared to the control **1** (90 : 1) or the shorter tris–Cu^{2+} complex **3** (4 : 1). This slight selectivity of CA for **2** (over **3**) may be due to greater strain in binding for **3** ($\Delta H = -15.1$ kcal mol^{-1} for **2**; -12.2 kcal mol^{-1} for **3**) to the protein. Complex **2** also showed very good selectivity for CA when compared to CEA (300 : 1).

In order to demonstrate selective binding of **2** to CA, a mixture of the protein and the control **1** ([CA] = 100 μM; [**1**] = 100 μM) was titrated with **2** ([**2**] = 0–500 μM). No change in the binding constant of **2** with CA was observed. When a

† Electronic supplementary information (ESI) available: scheme for the synthesis of **2**, **3** and ITC titration data (raw and processed) for the results reported in Table 1. See http://www.rsc.org/suppdata/cc/a9/a909001k/

DOI: 10.1039/a909001k

Table 1 Binding parameters for the receptors **2**, **3** and control **1** with carbonic anhydrase, chicken egg albumin and lysozyme

Protein	Stoichiometry (n)	K/M^{-1}	$-\Delta H$/kcal mol^{-1}
Lysozyme	**1**: 1.2 ± 0.3	(1.6 ± 0.2) × 10^3	7.0 ± 0.5
	2: nd	Precipitation	—
	3: —	< 10^3	—
Carbonic anhydrase	**1**: 0.90 ± 0.03	(3.5 ± 0.7) × 10^3	39.2 ± 2.0
	2: 0.94 ± 0.006	(299 ± 30) × 10^3	15.1 ± 0.3
	2: 0.81 ± 0.02 (pH 6.0)	(199 ± 35) × 10^3 (pH 6.0)	10.8 ± 0.5 (pH 6.0)
	2: 0.82 ± 0.01 (pH 8.0)	(97 ± 16) × 10^3 (pH 8.0)	10.2 ± 0.3 (pH 8.0)
	3: 1.3 ± 0.01	(75 ± 6) × 10^3	12.7 ± 0.2
	3: 1.3 ± 0.04 (pH 6.0)	(27.5 ± 4) × 10^3 (pH 6.0)	12.2 ± 0.7 (pH 6.0)
	3: 1.3 ± 0.02 (pH 8.0)	(45 ± 5) × 10^3 (pH 8.0)	20.6 ± 0.6 (pH 8.0)
Albumin	**1**: —	< 10^3	—
	2: —	< 10^3	—
	3: 1.6 ± 0.07	(16 ± 2.5) × 10^3	20.9 ± 1.5

mixture of **2** and CA ([CA] = 100 μM; [**2**] = 100 μM) was titrated with control **1** ([**1**] = 0–1 mM), no binding was observed. Similar results were obtained with the complex **3**. When a mixture of the three tested proteins (100 μM each) was titrated with complex **2**, the binding constant remained unchanged (280 000 ± 30 000). If carbonic anhydrase was not included in the mixture (*i.e*. CEA and lysozyme, 100 μM each), very weak affinity (< 1000) was detected. The binding selectivity reported here (for complex **2**) is distinctly different compared to that observed with a random distribution of copper(II) ions (on chelating Sepharose fast flow with iminodiacetate–Cu^{2+}, pH 7.0, 22 °C, either in equilibrium binding experiments or in chromatograhy).[15]

To demonstrate the role of Cu^{2+}–histidine interactions in the recognition process, we have studied the binding between CA and complex **2** by EPR spectroscopy in the solution phase (9.4 GHz, [CA] = 1.2 mM; [**2**] = 600 μM, 25.0 °C). Upon addition of the protein, the $g_{\|}$ value of Cu^{2+} ions of **2** decreased from 2.290 to 2.261. These values match well with the reported $g_{\|}$ values of the iminodiacetate–Cu^{2+} complex, free and bound to myoglobin through the histidine residues (2.288 and 2.264, respectively).[16] Also ITC titrations failed to detect any binding between the metal-free ligands (for **1**, **2** and **3**) and the proteins.

The binding of CA with the metal complexes **2** and **3** were followed by UV–VIS spectrometry. Upon addition of CA to a solution **2** (or **3**, [**2** or **3**] = 0.5 mM; [CA] = 0.8 mM, 25 mM HEPES buffer, pH 7.0, 25 °C, 10 × 100 μL additions), the absorption maxima progressively shifted from 727 to 660 nm. This indicated the coordination of one imidazole group (of histidine residues) per copper(II) ion of the complexes.[17] Analyses[18] of the resultant titration curves corroborated the ITC results. Circular dichroism studies (80 μM protein, 0–500 μM of **2** or **3**, 25 mM phosphate buffer, pH 7.0, 23.0 °C) indicated that the proteins were not unfolding in presence of the control **1** or receptors **2** or **3**.

Thus these studies demonstrate that the tris–Cu^{2+}complex **2** binds carbonic anhydrase strongly and with good selectivity compared to two other proteins with different histidine patterns on the surface. It should be noted that the reported method of protein recognition is applicable to proteins of known structures possessing histidine residues on the surface. With rapidly increasing number of solved protein structures, the method has wide applicability.

This work was supported by NSF-CAREER award (CHE-9896083) and an NIGMS - AREA grant (1R15 59594-01, NIH). The microcalorimeter was purchased through an NSF-EPSCoR award to North Dakota. We thank Professor Kenton Rodgers for help with the EPR experiments.

Notes and references

1 H.-J. Bohm and G. Klebe, *Angew. Chem., Int. Ed. Engl.*, 1996, **35**, 2588; R. E. Baine and S. L. Bender, *Chem. Rev.*, 1997, **97**, 1359.

2 S. Vunnum, V. Natarajan and S. Cramer, *J. Chromatogr. A*, 1998, **818**, 31; K. M. Muller, K. M. Arndt, K. Bauer and A. Pluckthun, *Anal. Biochem.*, 1998, **259**, 54; J. Mahiou, J. P. Abastado, L. Cabanie and F. Godeau, *Biochem. J.*, 1998, **330**, 1051.

3 F. Loscher, T. Ruclestuhl and S. Seeger, *Adv. Mater.*, 1998, **10**, 1005; S. Okada, S. Peng, W. Spevak and D. Charych, *Acc. Chem, Res.*, 1998, **31**, 229; I. Vikham and W. M. Albers, *Langmuir*, 1998, **14**, 3865; D. Elbaum, S. K. Nair, M. W. Patchan, R. B. Thompson and D. W. Christainson, *J. Am. Chem. Soc.*, 1996, **118**, 8381.

4 A. C. Bishop, C. Kung, K. Shah, L. Witucki, K. M. Shokat and Y. Liu, *J. Am. Chem, Soc.*, 1999, **121**, 627; G. F. Short, M. Lodder, A. L. Laikhter, T. Arslan and S. M. Hecht, *J. Am. Chem. Soc.*, 1999, **121**, 478; M. Beuck, *Angew. Chem., Int. Ed.*, 1999, **38**, 631; R. Zutshi, M. Brickner and J. Chmielewski, *Curr. Opin. Chem. Biol.*, 1998, **2**, 62.

5 S. Hirota, M. Endo, K. Hayamizu, T. Tsukazaki, T. Takake, T. Kohzuma and O. Yamaguchi, *J. Am. Chem. Soc.*, 1999, **121**, 849.

6 H. S. Park, Q. Lin and A. D. Hamilton, *J. Am. Chem. Soc.*, 1999, **121**, 8; Y. Hamuro, C. Calama, H. S. Park and A. D. Hamilton, *Angew. Chem., Int. Ed. Engl.*, 1997, **36**, 2680.

7 P. A. Frey and W. W. Cleland, *Bioorg. Chem.*, 1998, **26**, 175.

8 S. J. Lippard and J. M. Berg, *Principles of Bioorganic Chemistry*, University Science Books, Mill Valley, CA, 1994, pp. 349–376; H. Dugas, *BioOrganic Chemistry: A Chemical Approach to Enzyme Action*, Springer, New York, NY, 1996, pp. 388–460.

9 M. J. Kendrick, M. T. May, M. J. Plishka and K. D. Robinson, *Metals in Biological Systems*, Ellis Horwood, New York, NY, 1992, pp. 17–56.

10 The pdb files were obtained from the pdb server. The URL on the internet is: http://www.rcsb.org/pdb/. The files used for modeling are: 1can.pdb (for carbonic anhydrase), 1ova.pdb (for albumin) and 1azf.pdb (for lysozyme).

11 P. J. Sadler and J. H. Viles, *Inorg. Chem.*, 1996, **35**, 4490; A. Singh, M. Markowitz, L.-I. Tsao and J. Deschapms, *ACS Symp. Ser.*, 1993, **556**, 252.

12 The synthetic scheme for the synthesis of **2** and **3** is available as electronic supplementary material.† The tribromides were prepared following reported procedures. Y. Kim and R. Beekerbauser, *Macromolecules*, 1994, **24**, 1968; W. P. Cochrane, P. L. Pauson and T. S. Stevens, *J. Chem. Soc. C*, 1968, 630. All new compounds gave satisfactory ^{1}H and ^{13}C NMR spectra. The Cu^{2+} complexes were characterized by elemental analysis. For **1**: Anal. Calc. for $C_{11}H_{11}CuNO_4$: C, 46.39; H, 3.89; N, 4.92. Found: C, 46.06; H, 3.58; N, 4.62. For **2**: Anal. Calc. for $C_{39}H_{33}Cu_3N_3O_{12}\cdot 3H_2O$: C, 49.60; H, 3.74; N, 4.45. Found: C, 49.68; H, 3.90; N, 4.52. For **3**: Anal. Calc. for $C_{21}H_{21}Cu_3N_3O_{12}\cdot 3H_2O$: C, 33.54; H, 3.62; N, 5.59. Found: C, 33.53; H, 3.59; N, 5.54%.

13 W. Jiang, B. Graham, L. Spiccia and M. T. W. Hearn, *Anal. Biochem.*, 1998, **255**, 47.

14 M. Doyle, *Curr. Opin. Biotech.*, 1997, **8**, 31; I. Wadso, *Chem. Soc. Rev.*, 1997, 79.

15 T. W. Hutchens, T.-P. Yip and J. Porath, *Anal. Biochem.*, 1988, **170**, 168.

16 D. R. Schnek, D. W. Pack, D. Y. Sasaki and F. H. Arnold, *Langmuir*, 1994, **10**, 2382.

17 P. K. Dhal and F. H. Arnold, *Macromolecules*, 1992, **25**, 7051.

18 K. A. Connors, *Binding Constants*, Wiley, New York, 1987.

Communication a909001k