This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. This Accepted Manuscript will be replaced by the edited, formatted and paginated article as soon as this is available.

You can find more information about Accepted Manuscripts in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal’s standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.
Graphical Abstract

The NGAL binding siderophore Catechol was directly purified from human urine with columns and identified by LC-HR-ESI-MS, NMR spectroscopy methods.
Purification and Structural Characterization of “Simple Catechol”, the NGAL-Siderocalin Siderophore in Human Urine

Guan-Hu Bao1*, Jonathan Barasch2, Jie Xu1, Wei Wang1, Feng-Lin Hu1, and Shi-Xian Deng2

1Anhui Agricultural University, 130 West Changjiang Road, Hefei, Anhui Province, 230036, China.
2College of Physicians and Surgeons of Columbia University, 630 West 168th St, New York, USA.

* Corresponding Author
Professor Guan-Hu Bao
State Key Laboratory of Tea Plant Biology and Utilization, Anhui Agricultural University
Laboratory Homepage: http://tealab.ahau.edu.cn/
Office Phone: 0551-65786401
Office Fax: 0551-65786765
baoguanhu@ahau.edu.cn
Abstract

The identification of ligands that bind the protein Neutrophil Gelatinase-Associated Lipocalin (NGAL, Siderocalin, Lipocalin-2) have helped to elucidate its function. NGAL-Siderocalin binds and sequesters the iron loaded bacterial siderophore enterochelin (Ent), defining the protein as an innate immune effector. Simple metabolic catechols can also form tight complexes with NGAL-Siderocalin and ferric iron, suggesting that the protein may act as an iron scavenger even in the absence of Ent. While different catechols have been detected in human urine, they have not been directly purified from a biofluid and demonstrated to ligate iron with NGAL-Siderocalin. This paper describes a “natural products” approach to identify small molecules that mediate iron binding to NGAL-Siderocalin. A 10K filtrate of human urine was subjected to multiple steps of column chromatography and reverse-phase HPLC, guided by NGAL-Siderocalin-iron binding assays and LC-MS detection. The co-factor forming a ternary structure with iron and NGAL-Siderocalin was identified as authentic simple catechol (dihydroxybenze) by ESI-HR-Mass, UV, and NMR spectrometric analysis. Comparison of the binding strengths of different catechols demonstrated that the vicinal-dihydroxyl groups were the key functional groups and that steric compatibilities of the catechol ring have the strongest effect on binding. Although catechol was a known NGAL-Siderocalin co-factor, our purification directly confirmed its presence in the urine as well as its capacity to serve as an iron trap with NGAL-Siderocalin.

Keywords: NGAL, Siderocalin, Lipocalin-2, Catechol, Purification, Binding, Ferric Iron, Enterochelin
Introduction

Neutrophil gelatinase-associated lipocalin (NGAL, now known as Siderocalin) was first purified by two independent groups [1,2] from neutrophil granules. Like other members of the lipocalin family, NGAL-Siderocalin contains a calyx composed of eight anti-parallel β-sheets. However, the relationship of its structure to its proposed function in tissue damage, cell differentiation, or inflammation has been largely unknown, and major advances have resulted from the identification of its ligands which bind to the calyx. For example, the epochal article [3] demonstrating the interaction between NGAL-Siderocalin and bacterial siderophores Enterochelin (Ent) and 2,3-dihydroxybenzoic acid (DHBA) resulted in a surge of research on the bacteriostatic effects of NGAL-Siderocalin in Gram-Negative infections [3-6].

In addition to blocking bacterial growth by sequestering iron loaded Ent or its component DHBA, NGAL-Siderocalin may also have a function in the absence of bacteria, since it is highly expressed in embryonic yolk sac, some epithelial tubules, and bone marrow and in the adult kidney, heart, and retina damaged by ischemia-reperfusion injury and by chemical toxins [7,8]. Additionally, NGAL-Siderocalin was also reported to regulate diverse processes such as cell apoptosis [9], cell growth [7] and the rescue of ischemic injury [8, 10-12], but the mechanisms remain unknown. All functions of the lipocalin proteins relate to their association with low molecular weight ligands, suggesting that NGAL-Siderocalin has additional binding partners other than Ent and DHBA. If the additional cofactors associate with iron, then NGAL-Siderocalin may chelate iron even in the nominal absence of a bacterial infection [7,8,13,14]. Siderophore-like ligands are necessary for this premise since NGAL-Siderocalin does not directly bind iron. Additionally, because NGAL-Siderocalin is rapidly secreted from cells, we must assume that iron will be chelated extracellularly in its oxidized form (ferric iron, Fe$^{3+}$). In sum, attempts to find iron binding siderophores have been the most important tasks in NGAL-Siderocalin research since the structure of the bacteria ligand, Ent, was elucidated by Strong, Raymond and their colleagues [3,15].

A search for iron co-factors, “mammalian siderophores” has languished soon after the hypothesis was first posited [16]. A mammalian siderophore with molecular weight ~1500 Da was isolated from horse liver in 1980; the siderophore bound iron and stimulated bacterial growth by promoting iron uptake [17]. More recently, several groups launched related exploration to identify NGAL-Siderocalin co-factors with varying degrees of success. They identified different types of catechols such as norepinephrine [18], 2,5-DHBA [19] and simple catechol [14] and they proposed that in combination with NGAL-Siderocalin, each may bind iron either intracellularly (2,5-DHBA) or extracellularly (norepinephrine, simple catechol), providing an iron buffer and potentially an inhibitor of iron donation for bacterial growth.

While it is the case that the simple catechols can form tight ternary complexes with
NGAL-Siderocalin and ferric iron even in the circulation in vivo [14], catechol has not been directly purified from human biofluids and shown to interact with NGAL-Siderocalin and iron, albeit that it has been detected in the urine by quantitative analyses [14]. In addition, while the other two ligands can interact with iron, the affinity of these interactions may be limited. For example, 2,5-DHBA interacts with iron only in the salicylate mode (between the 2-hydroxyl and carboxylic acid groups) which has a lower affinity for iron than vincinal diols. In fact, 2,5-DHBA was extensively tested by physiological analyses and has not been confirmed to adequately bind or transport iron [20-23]. Further, while each of these catechols must associate into tetrameric complexes (catechol$_3$:iron) both 2,5-DHBA and norepinephrine are sterically inhibited within the NGAL-Siderocalin calyx. The 5-hydroxyl group clashes with the wall of the NGAL-Siderocalin [21,22]. Similarly, the 3,4 positioning of norepinephrine’s diol, may result in steric hindrance for its side chain.

In sum, ongoing studies have identified endogenous iron chelators, some of which may contribute to our understanding of the role of NGAL-Siderocalin in non-bacteriostatic mammalian processes. In the current study, simple catechol was isolated from human urine through extensive column chromatography and its structure was established by high-performance liquid chromatography (HPLC)-high resolution electrospray ionization (HR-ESI) mass spectrometry (MS) and NMR spectroscopy methods. The coordination with iron and NGAL-Siderocalin was detected by filtration methods. These analyses demonstrate conclusively that endogenous simple catechol is present in human urine, and that it binds NGAL-Siderocalin, and acts as a siderophore trapping iron with NGAL-Siderocalin.

**Results**

Repeated Liquid Chromatography Identifies a NGAL-Siderocalin Binding Fraction.

Preliminary fractionation of the urine was conducted by Mori et al. [8] who demonstrated molecules less than 10 KDa associate iron with NGAL-Siderocalin. These co-factors occupied the same binding sites that dock the bacterial siderophore, Ent, because mutations of amino acids K125 and K134 reduced iron binding. Further purification was conducted to identify these molecules.

200 L human urine was collected and frozen at -16 °C for three days. Since the task is to identify small molecular siderophores from human urine, all urine was pooled together. The urine was then quickly thawed and filtered through a series of filter papers, Whatman, 0.45 μm, and 0.22 μm Millipore filters, and finally a 10 KDa filter under nitrogen (4 °C). The filtrate was subjected to repeated MCI (a kind of gel) column chromatography, which was eluted stepwise to give five fractions: F1-F5.
Figure 1 The general purification procedures of NGAL binding siderophores: from raw urine to pure catechol by different column chromatography.

(Figure 1). Using the Mori et al. 2005 assay, 100 µg -25 µg of each fraction was combined with recombinant human NGAL-Siderocalin (22 KDa) and $^{55}$Fe, and the 10
K retentate of $^{55}$Fe was compared to assays which lacked aliquots of urine (“apo-NGAL-Siderocalin”) as a negative control or contained enterochelin as a positive control for iron binding (“Ent”; Figure 2). Figure 2 shows that Fractions F2-F4 had binding activity even when assayed at low concentration (25 μg; the incubation volume was 100 μL), particularly fraction F2.

![Figure 2](image-url)

**Figure 2** Formation of NGAL-Siderocalin:Siderophore:$^{55}$Fe$^{3+}$ complexes by (A) five fractions of the first MCI gel Liquid Chromatography (LC) column compared with the negative control, apoNGAL-Siderocalin and with the positive control, NGAL-Siderocalin:Ent; (B) the Sephadex LH-20 LC fractions; (C) the first HPLC fractions; (D) the Silical Gel LC fractions.

Further fractionation of MCI F2 using Sephadex LH-20 column chromatography produced fractions SF-1 to SF-6, and the undissolved fraction SF-7. Their NGAL-Siderocalin binding activity was tested at 50 μg (Figure 2B) which showed that SF3 through SF7 had greater specific activity than the starting fraction from MCI F2. The selected active SF1-6 fractions were then subjected to reverse phase chromatography, producing five fractions HPLC 1-5, the first three demonstrated NGAL-Siderocalin binding activity at 25 μg (Figure 2C). Finally, the first three active HPLC fractions were subjected to a silical gel column chromatography, providing 15 fractions SSF1 to SSF14, and a precipitate. Only 10 μg of Fractions 1-3 were necessary to demonstrate intensive NGAL-Siderocalin binding activities equivalent to Ent (Figure 2D).

**Further Purification Identifies Catechol**
**Figure 3** LC-HR-ESI-MS\(^{-}\) (negative mode) of the SSF1 fraction and the standard catechol positive control: (A) Total Ion Chromatography of SSF1, (B) Total Ion Chromatography of standard catechol, (C): ESI (negative mode)-HR-MS of the peak at RT=12.211 min. (D) Automatic calculation of the molecular formula by G3335AA MassHunter Qualitative Analysis Software B.01.03 based on high resolution (HR)-MS.

LC-HR-ESI-MS (Negative mode) demonstrated that the mass at retention time (RT) 12.211 min. (the active fraction) was 109.0298 [M-H]\(^{-}\} (Calc 109.02951), which suggested that it had the formula C\(_6\)H\(_6\)O\(_2\) and four degrees of unsaturation (Figure 3).

There are three simple phenols o-dihydroxybenzene (1,2-dihydroxybenzene, Catechol), m-dihydroxybenzene (1,3-dihydroxybenzene, Resorcin), and p-dihydroxybenzene (1,4-dihydroxybenzene, Hydroquinone) with the same formula. However, as can be seen in Figure 4, the peak at Retention Time (RT) = 12.211 min coincided with the catechol (o-dihydroxybenzene) standard and not with the m- or p-dihydroxybenzene standards. Our identification was further validated by demonstrating catechol in many different active fractions of human urine by HPLC-photodiode array detector (PAD)-high resolution (HR)-ESI MS. These urines derived from normal patients and patients affected by acute or chronic diseases of the kidney.
Figure 4 Separated chromatograms listing LC-DAD (280 nm) of SSF1 (S, Green), o-(Catechol, dark blue), m- (Resorcin, light blue), and p- (Hydroquinone, orange) dihydroxybenzenes.

Confirmation of Catechol with UV and NMR Spectroscopy

We purified catechol from the SSF1 fraction by a second HPLC step which produced a highly purified peak (Figure 5A) corresponding to authentic catechol (Figure 5B). The UV spectrum of the purified target demonstrates a peak at 274 nm, which is the characteristic peak of catechol (Figure 5C). The $^1$H NMR (DMSO-$d_6$) showed four aromatic hydrogen ($\delta$ ppm): 6.569 (2H, H-3,6), 6.690 (2H, H-4, 5), 8.811 (2H, OH-1,2), which further confirmed that the structure is catechol (Figure 5D).

Figure 5 Purification of catechol from SSF1 by HPLC: (A) HPLC spectrum; (B)
HPLC analysis of purified catechol eluting at RT = 20.7min, overlaying the first peak found in A.; (C) UV spectrum of purified catechol demonstrating its characteristic peak at 274 nm; (D) $^1$H NMR of purified catechol (DMSO- $d_6$ as solvent: solvent $\delta$ ppm: 2.477).

**Discussion**

It was reported that NGAL-Siderocalin weakly binds catechol ($K_d = 0.2 \pm 0.06 \mu M$) but that the addition of ferric iron markedly enhanced catechol binding ($K_d = 2.1 \pm 0.5 \text{ nM}, 0.4 \pm 0.2 \text{ nM}$), generating NGAL-Siderocalin:Catechol:Fe$^{3+}$ complexes [14]. While the co-occurrence of iron and catechol has been detected by quantitative analyses, particularly in urine [24-29], direct purification of free catechol from a biofluid is necessary to demonstrate its authenticity and chemical structure. It is particularly important to determine whether catechol is oxidized or whether catechol is derivatized by bulky side chains. These ongoing studies are necessary to identify a co-factor that ligates iron in the NGAL-Siderocalin calyx in the nominal absence of Enterochelin.

We isolated authentic, non-oxidized and non-derivatized catechol from human urine. This was surprising because catechol may oxidize to semiquinone or quinone but the presence of authentic catechol suggests that the reduced state may be maintained in the urine by a variety of reductants [30,31]. In addition, catechol has been reported to be derivatized by sulfation upon transport across the proximal tubule into the urine [32], and most quantitative assessments have demonstrated that 90 - 95% of urinary catechol is sulfated (non-sulfated component: 1-10 μM). However, most importantly, we isolated the non-derivatized form of catechol. Therefore, Urinary catechol (free one is around 1-10 μM), NGAL (around 0.1 μM at disease condition), were found to be highly abundant to bind iron [14].

The specific chemistry of catechol derivatization is critical to iron and NGAL-Siderocalin binding efficiency. We previously tested catechol derivatives to determine the importance of its structural components (Figure 6). Based on binding assays, we found that the ortho-dihydroxy group is the key binding group: methylation or sulfonation of one or both of the hydroxyl groups reduced its binding activity (compare dihydroxybenzene (compound 1) to methoxy- (compound 4), dimethoxy- (compound 5), cyclic sulfated - (compound 6), or monosulfate - benzene (compounds 7). Separation of the diol by a sulfated hydroxyl group (pyrogallol-2-sulfate, compound 8) also reduced binding compared with pyrogallol (compound 3). We also found that steric compatibilities dominate whether catechol can serve as a co-factor. For example, bulky side groups placed anywhere on the ring were inhibitory to binding such as the addition of methoxy- or carboxylic acid- ortho to the diol (compare 2,3-dihydroxybenzoic acid (compound 2) or 3-methoxycatechol (compound 9) with 2,3-dihydroxybenzene, compounds 1 with 2, 9). Addition of side groups in the para-position were even more highly inhibitory compared with bulky
Figure 6. The binding strength of catechol and its derivatives (% retention of $^{55}\text{Fe}^{3+}$), was tested with NGAL-Siderocalin, $^{55}\text{Fe}^{3+}$ and 10 μM compounds (see Materials and Methods). Compounds 1: catechol, 2: 2,3-DHBA, 3: pyrogallol, 4: guaiacol, 5: 1,2-dimethoxybenzene, 6: catechol cyclic sulfate, 7: catechol sulfate sodium, 8: pyrogallol-2-Sulfate, 9: 3-methoxycatechol, 10: pyrogallol-4-Sulfate, 11: gallic acid, 12: 3-methylcatechol, 13: 4-methylcatechol, 14: 3,4-dihydroxybenzoic Acid (A) Basic siderophores (B) Protected hydroxyl group(s) dramatically decrease binding strength.
(C) Bulky steric groups decrease the binding strength (D) The para-position substitution has a greater effect than the ortho-position substitute (E) The steric compatibilities may have greater importance than electrostatic interactions. [14]

... groups in the ortho-position. Examples include (i) inhibition of binding by the addition of either a carboxyl group or a sulfonate group to trihydroxybenzene (gallic acid and pyrogallol-4-sulfate compared to pyrogallol, compounds 3 vs 10, 11) (ii) the weaker binding activity of 4-methylcatechol compared to 3-methylcatechol (compounds 12, 13), and (iii) the positioning of the side groups in 3,4-dihydroxybenzoic acid which inhibits activity to a greater extent than 2,3-dihydroxybenzoic acid (compounds 2, 14). Lastly, it is the case that electron donating groups which should enhance binding such as methyl, methoxy, and even hydroxyl groups actually reduce binding activity, reflecting the dominance of steric compatibilities compared with electrostatic effects (compare compounds 1 vs 9). In sum, the binding of catechol is nearly restricted to its parent compound with significant restriction on modifications to the diol group as well as at most positions around the ring.

Our data confirm that simple catechol could be stably obtained from human urine. Catechol was nearly unique among small molecules: it can act as an extracellular co-factor for iron chelation within the scaffolding provided by NGAL-Siderocalin protein. 2,3-dihydroxybenzoic acid and pyrogallol may play a similar role but are less effective NGAL-Siderocalin binding partners. Finally, it should be noted that the co-occurrence of NGAL-Siderocalin, catechol and iron is found in human urine after the kidney is injured.

Materials and Methods

Chemicals

HPLC-grade acetonitrile was purchased from TEDIA (USA). Formic acid was of HPLC grade and was purchased from Agilent (Santa Clara, CA). Pure distilled water was purchased from Watersons (China). 0.45 μm, 0.22 μm stericup filter and 10 K filter film were bought from Millipore Corporation Company. Standard catechol, resorcin, hydroquinone were purchased from Shanghai ZhongTai Chemical Co. Ltd. Sephadex LH-20 and Whatman Filter paper were from GE Healthcare Life Sciences, MCI (CHP20/P120) gel was purchased from Mitsubishi Chemical Co., Silical gel was from Qingdao Marine Chemical Co. Ltd. All other reagents used in this study were of analytical grade. 1H NMR was recorded in dimethyl-d6 sulfoxide (DMSO-d6) with Agilent DD2 600 Hz NMR.

Urine collected, frozen and filtration

200 L Urine was collected in the early morning and frozen at -16 °C for three days. Since the task is to identify small molecular siderophores from human urine, all urine
was pooled together. The frozen urine was melted and filtered through Whatman filter paper quickly, then filtered through 0.45 μm, 0.22 μm stericup filters and then 10 KDa filters at 4 °C. De-identified urine samples were collected according to the Institutional Review Board Protocols IRB-AAAE9967 and IRB-AAAC4097. IRB approval was granted by Columbia University Medical Center IRB to collect the de-identified urine used for this analysis. The approval period began 02/27/2007 under protocol AAMAE9967 and continues under protocol AAAC4097 through 12/14/2015. The requirement to obtain written documentation of informed consent has been waived by the IRB in accordance with 45 C.F.R. § 46.117(c).

**MCI column chromatography**

The filtrate was loaded repeatedly onto a MCI column (800 × 70 mm ID) (8 L filtrate per column) and eluted stepwise CH₃OH/dH₂O [0:1 (total 200 L), 1:2 (160 L), 2:1 (160 L) and 1: 0 (5.0 L)] and finally 100% Acetone (4.0 L) to give five fractions F1 to F5.

**Sephadex LH-20 chromatography**

Sephadex LH-20 column chromatography (500 g Sephadex LH20, 800 × 70 mm ID) of the active fractions F2 was serially eluted with 100% methanol as eluting solvent, to give SF1 (400mL MeOH, 6.6 g), SF2 (second 400 mL MeOH, 7.4 g), SF3 (further 250mL MeOH, 6.2 g), SF4 (further 500 mL MeOH, 2.5545 g), SF5 (further 500 mL MeOH, 0.2371 g), SF6 (further 2 L MeOH, 0.4774 g), and the undissolved fraction SF7 (0.5622 g).

**First HPLC purification method**

SF fractions SF3-SF7 were pass through an HPLC column (Phenomenex selectosil 5μm, C-18, 250 × 10 mm i. d.) to give five fractions HPLC-1 (15% aqueous methanol), HPLC-2 (30% aqueous methanol), HPLC-3 (50% aqueous methanol), HPLC-4 (75% aqueous methanol), HPLC-5 (100% methanol).

**Silical gel column chromatography**

Silical gel column chromatography of the fractions HPLC1-HPLC3 produced 11 fractions SSF1-14 and the precipitate (fast washed in 30 min with different ratios of hexane:acetone varying from 4:1 to 1:2; and finally with 100% methanol).

**Second HPLC purification methods for purification of catechol from SSF1**

Purification of catechol was carried out using HPLC (SHIMADZU LC-20A) with pump (LC-20AD), diode array detector (SPD-M20A), connector (CBM-20A), column oven (CTO-10AS), auto sampler (SIL-20A), and a SP ODS-A C18 column (250 mm × 4.6 mm ID, 5 μm, H&E Co., Ltd, P. R. China). Column temperature was set at 30 °C. The isolation utilized mobile phase A (Distilled Water) and mobile phase B (Methanol). The gradient of mobile phase B was as follows: 0-3 min, 5%; 3-20 min, from 5% to 50%; 20-30 min. from 50% to 70%; 25-26 min, from 55% to 100%; 26-40 min, from 100% to 100%; then kept at 5% for 10 min.
LC-MS data acquisition

HPLC-photodiode array detector (PAD) - electrospray ionization (ESI) - high resolution mass spectrometry (HRMS) analysis was carried out on an Agilent 6210 time-of-flight LC/MS system with a binary high-pressure mixing pump, auto sampler, column oven, PAD, time-of-flight (TOF) MS with an ESI source and an Agilent workstation. Analytical separations were carried out on an Agilent Poroshell 120 column (C-18, 2.7 μm, 100 × 3.0 mm i. d.). Liquid samples were concentrated with a Senko R 206 rotary evaporator and lyophilized with a Laboconco freeze dryer. LC parameters were as follows: injection volume, 5 μL; column temperature, 40 °C; flow rate, 0.3 mL/min; and the eluates were monitored with a PAD at full length scan from 200 nm to 600 nm. The mobile phase A = 0.1% formic acid in water, B = 0.1% formic acid in acetonitrile, and gradient elution was carried out: 10 - 30% B for 0 - 5 min; 30-70% B for 5 - 40 min; 70 - 100% B for 40 - 41 min; 100% B for 41 - 51 min. The column was reconditioned with initial gradient for 12 min. The mass spectrometer parameter settings used for the measurement were as follows: ionization mode, positive and negative; gas temperature, 350 °C; drying gas, 12 L/min; nebulizer pressure, 45 psi; capillary voltage, 4000 V in positive mode and 3500 V in negative mode; fragmentor voltage, 215 V in positive mode and 170 V in negative; skimmer voltage, 60 V; OCT 1 RF, 250 V. Data acquisition was performed in the m/z range 50 - 1100 Da.

Binding Assays

NGAL-Siderocalin (10 μM), \(^{55}\text{Fe}^{3+}\) (1 μM + cold FeCl\(_3\), 9 μM) and urine fractions (10-100 μg/100μL), catechols (10 μM) or the positive control Ent (10 μM) were incubated in 150 mM NaCl, 20 mM Tris (pH 7.4), room temperature (RT) for 60 min. The buffers (pH7.4) utilized Tris Base titrated with HCl and NaOH. The mixture was then washed three times with the Tris buffer on an Ultrapel-10K Amicon Ultra (Millipore, Ireland Ltd) and retained \(^{55}\text{Fe}\) was measured with a scintillation counter.

Acknowledgement

Financial assistance was received with appreciation from National Natural Science Foundation of China 81170654/H0507, Undergraduate Innovation Project from Anhui Province and Anhui Agricultural University, Anhui Major Demonstration Project for Leading Talent Team on Tea Chemistry and Health, and Program for Changjiang Scholars and Innovative Research Team in University IRT1101. Financial support also derived from NIH 5R01DK073462 (JB).
References


